Tue. Jun 4, 2019

Room D

Organized Session | Organized Session | [OS] OS-10

[1D2-OS-10a] 不動産と AI(1)

清田 陽司（（株）LIFULL）、山崎 俊彦（東京大学）、諏訪 博彦（奈良先端科学技術大学院大学）、清水 千弘（日本大学）、橋本 武彦（GA technologies）

1:20 PM - 3:00 PM Room D (301B Medium meeting room)

[1D2-OS-10a-01] Assessment of Thermal Insulation and Sound Performance of Apartments Based on IoT Sensing

〇Toshihiko Yamasaki[1], Yuki Obuchi[1], Yuan Lin[1], Ryoma Kitagaki[2], Satoshi Toriumi[3], Mikihisa Hayashi[3], Ai Sakai[4], Nobuhito Haga[4], Shimpei Nomura[4], Yoichi Ikemoto[4] （1. The University of Tokyo, 2. Hokkaido University, 3. Future Standard Co., Ltd., 4. Recruit Sumai Company Ltd.）

1:20 PM - 1:40 PM

[1D2-OS-10a-02] Data Collection for the Comfort Level of Rental Property by IoT Sensing

〇Hirohiko Suwa[1], Atsushi Otsubo[1], Yugo Nakamura[1], Masahito Noguchi[2] （1. Nara Institute of Science and technology, 2. LIFULL Co., Ltd.）

1:40 PM - 2:00 PM

[1D2-OS-10a-03] Business utilization of real estate image classification system using deep learning

〇Tomoya Tsukahara[1], Kodai Sudo[1] （1. BrainPad Inc.）

2:00 PM - 2:20 PM

[1D2-OS-10a-04] Proposal of Railway Route Search System for Finding Living Place

〇Keisuke Kikuchi[1], Kenichiro Kobayashi[2], Takehiko Hashimoto[2], Yasufumi Takama[1] （1. Tokyo Metropolitan University, 2. GA Technologies）

2:20 PM - 2:40 PM

[1D2-OS-10a-05] Evaluation of Rent Prediction Models using Floor Plan Images

〇Ryosuke Hattori[1], Kazushi Okamoto[1], Atsushi Shibata[2] （1. The University of Electro-Communications, 2. Advanced Institute of Industrial Technology）

2:40 PM - 3:00 PM

Organized Session | Organized Session | [OS] OS-10

[1D3-OS-10b] 不動産と AI(2)

清田 陽司（（株）LIFULL）、山崎 俊彦（東京大学）、諏訪 博彦（奈良先端科学技術大学院大学）、清水 千弘（日本大学）、橋本 武彦（GA technologies）

3:20 PM - 5:00 PM Room D (301B Medium meeting room)

[1D3-OS-10b-01] (Invited talk) Recent studies on real estate tech in China

〇Xiangyu Guo[1] （1. Fudan University）

3:20 PM - 4:00 PM

[1D3-OS-10b-02] Two-Step estimation model using machine learning at detached house price

〇Yusuke Takahashi[1] （1. FUJITSU CLOUD TECHNOLOGIES LIMITED）

4:00 PM - 4:20 PM

[1D3-OS-10b-03] Statistical modeling of the transition time of an occupation of rental rooms by using the housing information website data

〇Hayafumi Watanabe[1], Yu Ichifuji[2], Masahito Suzuki[3], Satoshi Yamashita[4] （1. Kanazawa university, 2. Nagasaki university, 3. UD Asset Valuation Co., Ltd., 4. The institute of statistical mathematics）

4:20 PM - 4:40 PM

[1D3-OS-10b-04] Efforts on UX innovation in real estate field by utilizing AI

Yoji Kiyota[1], 〇Satoshi Shiibashi[1], Takeshi Ninomiya[1], Takao Yokoyama[1], Takuro Hanawa[1], Takeshi Eto[1], Akiko Yokoyama[1], Kei Kikuchi[1], Musashi Kobayashi[1], Akane Kameda[1], Kazuki Takigawa[1], Yusuke Saito[1], Kazushi Katayama[1] （1. LIFULL Co., Ltd.）

4:40 PM - 5:00 PM

Room E

Organized Session | Organized Session | [OS] OS-3

[1E2-OS-3a] AI における離散構造処理と制約充足(1)

波多野 大督（理化学研究所）、蓑田 玲緒奈（（株）ベイシスコンサルティング）

1:20 PM - 3:00 PM Room E (301A Medium meeting room)

[1E2-OS-3a-01] (Invited talk) Sentential Decision Diagrams and related topics

〇Masaaki Nishino[1] （1. NTT Communication

©The Japanese Society for Artificial Intelligence

Science Laboratories）
1:20 PM - 2:00 PM

[1E2-OS-3a-02] MCSes Enumeration with the Glucose SAT Solver
〇Miyuki Koshimura[1], Ken Satoh[2] （1. Kyushu University, 2. National Institute of Informatics）
2:00 PM - 2:20 PM

[1E2-OS-3a-03] A SAT-based CSP Solver sCOP and its Results on 2018 XCSP3 Competition
〇Takehide Soh[1], Daniel Le Berre[3,4], Mutsunori Banbara[2], Naoyuki Tamura[1] （1. Kobe University, 2. Nagoya University, 3. CRIL-CNRS, UMR 8188, 4. Université d'Artois）
2:20 PM - 2:40 PM

[1E2-OS-3a-04] An Analysis of Entry and Exit Data in Office by Decision Tree Learning Using Clustering Factor Matrix from Non-negative Multiple Matrix Factorization
〇Seidai Kojima[1], Hayato Ishigure[1], Miwa Sakata[1], Atsuko Mutoh[1], Koichi Moriyama[1], Nobuhiro Inuzuka[1] （1. Nagoya Institute of Technology）
2:40 PM - 3:00 PM

Organized Session | Organized Session | [OS] OS-3

[1E3-OS-3b] AI における離散構造処理と制約充足(2)

波多野 大督（理化学研究所）、蓑田 玲緒奈（（株）ベイシスコンサルティング）
3:20 PM - 4:00 PM Room E (301A Medium meeting room)

[1E3-OS-3b-01] A Branch-and-bound Algorithm for Determining Discrete Changes for Hybrid Constraint Solver HyLaGI
〇Masashi SATO[1], Kazunori UEDA[1] （1. Waseda University）
3:20 PM - 3:40 PM

[1E3-OS-3b-02] Dynamic Reduction of Guarded Constraints for the Hybrid Systems Modeling Language HydLa
〇Takafumi Horiuchi[1], Kazunori Ueda[1] （1. Waseda University）
3:40 PM - 4:00 PM

Room F

Organized Session | Organized Session | [OS] OS-17

[1F3-OS-17a] 農業と AI(1)

小林 一樹（信州大学）、竹崎 あかね（農研機構革新工学センター）
3:20 PM - 5:00 PM Room F (302B Medium meeting room)

[1F3-OS-17a-01] (Invited talk) Issues of field high throughput plant phenotyping based on images
〇Seishi Ninomiya[1] （1. University of Tokyo）
3:20 PM - 3:40 PM

[1F3-OS-17a-02] Bird Damage Prevention System Utilizing Deep Learning based on Birds' Behavior
〇Kazuki Kobayashi[1], Fumiya Shimobayashi[1], Kazunori Terada[2], Takefumi Yoshikawa[3], Hiroyuki Sato[4], Hiroyuki Tsuchiya[4], Kanokwan Atchariyachanvanich[5] （1. Shinshu University, 2. Gifu University, 3. Toyama Prefectural University, 4. Marimo Electronics Co., Ltd., 5. Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang, Thailand）
3:40 PM - 4:00 PM

[1F3-OS-17a-03] Training Data Augmentation for Hidden Fruit Image Segmentation by using Deep Learning
〇Ryoma Takai[1], Kazuki Kobayashi[1] （1. Shinsyu University）
4:00 PM - 4:20 PM

[1F3-OS-17a-04] *Q*-learning with Neural Network for Automatic Irrigation Control of Crops
〇Shuto Namba[1], Junpei Tsuji[1], Masato Noto[1] （1. Kanagawa University）
4:20 PM - 4:40 PM

[1F3-OS-17a-05] Crop Yield Estimation for Hydroponic Tomatoes using regional CNNs
〇Yuri Isoyama[1], Fumiyo Emura[1], Hirohisa Satoh[1], Takashi Shinozaki[2,3] （1. Kyowa Co., Ltd., 2. Japanese Society for Artificial IntelligenceCiNet, National Insitute of Information and Communications Technology, 3. Graduate School of Information Science and Technology, Osaka University）
4:40 PM - 5:00 PM

Organized Session | Organized Session | [OS] OS-17

[1F4-OS-17b] 農業と AI(2)

小林 一樹（信州大学）、竹崎 あかね（農研機構革新工学センター）
5:20 PM - 7:00 PM Room F (302B Medium meeting room)

©The Japanese Society for Artificial Intelligence

[1F4-OS-17b-01] A Comparative Analysis of Labor Force Survey Data among Vegetable Cultivation Systems based on Agriculture Activity Ontology
〇Akane Takezaki[1], Kaoru Maeyama[3], Joo Sungmin[2], Hideaki Takeda[2], Tomokazu Yoshida[1] （1. National Agriculture and Food Research Organization , 2. National Institute of Informatics , 3. Iwate Agricultural Research Center）
5:20 PM - 5:40 PM

[1F4-OS-17b-02] A Method for Judging the Semantic Similarity between Crops from Farm Management Articles based on Agricultural Knowledge Graph
〇Sungmin Joo[1], Hideaki Takeda[1], Akane Takezaki[2], Tomokazu Yoshida[2] （1. National Institute of Informatics, 2. The National Agriculture and Food Research Organization）
5:40 PM - 6:00 PM

[1F4-OS-17b-03] Analysis of Konbu-Dashi by multi-band optical method from a viewpoint of miscellaneous taste
〇Takaharu Kameoka[1], Takumi Taguchi[1], Eriko Nishikawa[1], Ryoei Ito[1], Atsushi Hashimoto[1], Noriyuki Yugawa[2], Nobuaki Obiki[2] （1. Mie University, 2. Tsuji Culinary Institute）
6:00 PM - 6:20 PM

[1F4-OS-17b-04] The individual identification of cattle using LP-residual signal extracted from cattle sound
〇Yurie Iribe[1], Mako Soga[1], Tomoki Kojima [2], Tatsuaki Masuda[2] （1. Aichi Prefectural University, 2. Aichi Agricultural Research Center）
6:20 PM - 6:40 PM

[1F4-OS-17b-05] Suppression of false alarm using crowdsourcing in calving detection system
〇Yusuke Okimoto[1], Susumu Saito[1,2], Teppei Nakano[1,2], Makoto Akabane[1,2], Tetsunori Kobayashi[1], Tetsuji Ogawa[1] （1. Waseda University, 2. Intelligent Framework Lab）
6:40 PM - 7:00 PM

Room G

Organized Session | Organized Session | [OS] OS-13

[1G3-OS-13a] “ナッジ”エージェント：人をウェルビーイングへと導くエージェントの構築へ向けて(1)

小野 哲雄（北海道大学）、今井 倫太（慶應義塾大学）、植田 一博（東京大学）、山田 誠二（国立情報学研究所
3:20 PM - 5:00 PM Room G (302A Medium meeting room)

[1G3-OS-13a-01] “ Nudge” Agent
〇Tetsuo Ono[1] （1. Hokkaido University）
3:20 PM - 3:40 PM

[1G3-OS-13a-02] (Invited talk) Cognitive science of choice and manipulation
〇Ayumi Yamada[1] （1. The University of Shiga Prefecture）
3:40 PM - 4:20 PM

[1G3-OS-13a-03] How do we harness others’ opinions?
〇Itsuki Fujisaki[1], Hidehito Honda[2], Kazuhiro Ueda[1] （1. University of Tokyo, 2. Department of Psychology）
4:20 PM - 4:40 PM

[1G3-OS-13a-04] Nudge with a forward posture
〇Masaru Shirasuna[1], Hidehito Honda[2], Kazuhiro Ueda[1] （1. the University of Tokyo, 2. Yasuda Women’s University）
4:40 PM - 5:00 PM

Organized Session | Organized Session | [OS] OS-13

[1G4-OS-13b] “ナッジ”エージェント：人をウェルビーイングへと導くエージェントの構築へ向けて(2)

小野 哲雄（北海道大学）、今井 倫太（慶應義塾大学）、植田 一博（東京大学）、山田 誠二（国立情報学研究所
5:20 PM - 6:40 PM Room G (302A Medium meeting room)

[1G4-OS-13b-01] Framework for Human-Agent Team using Implicit Information Showing via Behavior
〇Ryo Nakahashi[1], Seiji Yamada[2,1] （1. The Graduate University for Advanced Studies(Sokendai), 2. National Institute of Informatics）
5:20 PM - 5:40 PM

[1G4-OS-13b-02] To determine the display timing of driving assistant using Nudge
〇Kouichi Enami[1], Michita Imai[1], Kohei Okuoka[1], Shohei Akita[1] （1. Keio

©The Japanese Society for Artificial Intelligence

University）

5:40 PM - 6:00 PM

[1G4-OS-13b-03] Adaptive Trust Calibration for Human-AI Collaboration

〇Kazuo Okamura[1], Seiji Yamada[2,1] （1. SOKENDAI (The Graduate University for Advanced Studies), 2. National Institute of Informatics）

6:00 PM - 6:20 PM

[1G05-07-4] Discussion / Conclusion

6:20 PM - 6:40 PM

Room P

Organized Session | Organized Session | [OS] OS-21

[1P3-OS-21] プロセス中心のシステムデザインとラーニングアナリティクス

緒方 広明（京都大学）、近藤 伸彦（首都大学東京）、瀬田 和久（大阪府立大学）、平嶋 宗（広島大学）、松居 辰則（早稲田大学）、堀口 知也（神戸大学）

3:20 PM - 5:00 PM Room P (Front-left room of 1F Exhibition hall)

[1P3-OS-21-01] Data-driven Infrastructure for Evidence-based Education and Learning

〇Hiroaki Ogata[1], Majumdar Rwitajit [1], Gökhan Akçpinar[1], Brendan Flanagan[1] （1. Kyoto University）

3:20 PM - 3:40 PM

[1P3-OS-21-02] An Analysis of Learning Processes using Scrapbox and Learning Outcomes

〇Nobuhiko Kondo[1], Toshiharu Hatanaka[2], Takeshi Matsuda[1] （1. Tokyo Metropolitan University, 2. Osaka University）

3:40 PM - 4:00 PM

[1P3-OS-21-03] Diagnosing and Promoting Self-Directed Investigative Learning on the Web

〇Yoshiki Sato[1], Akihiro Kashihara[1], Shinobu Hasegawa[2], Koichi Ota[3], Ryo Takaoka[4] （1. The University of Electro-Communications, 2. Japan Advanced Institute of Science and Technology, 3. Japan Institute of Lifelong Learning, 4. Yamaguchi University）

4:00 PM - 4:20 PM

[1P3-OS-21-04] A Lecture Substitution Robot for Diagnosing and Reconstructing Presentation Behavior

〇Akihiro Kashihara[1], Tatsuya Ishino[1], Mitsuhiro Goto[2] （1. The University of Electro-Communications, 2. NTT Service Evolution Laboratories）

4:20 PM - 4:40 PM

[1P06-09-5] Discussion / Conclusion

4:40 PM - 5:00 PM

©The Japanese Society for Artificial Intelligence

Wed. Jun 5, 2019

Room D

Organized Session | Organized Session | [OS] OS-1

[2D4-OS-1a] 計算社会科学(1)

鳥海 不二夫（東京大学）、笹原 和俊（名古屋大学）、吉田 光男（豊橋技術科学大学）、瀧川 裕貴（東北大学）、榊 剛史（ホットリンク）、高野 雅典（サイバーエージェント）

3:20 PM - 5:00 PM Room D (301B Medium meeting room)

[2D4-OS-1a-01] Proposal of the Method to Construct Landscape Map for Massive Information Propagation on Social Media

〇Takeshi Sakaki[1,2], Fujio Toriumi[2], Masanao Ochi[2] （1. Hottolink, Inc., 2. Univ. of Tokyo）

3:20 PM - 3:40 PM

[2D4-OS-1a-02] The emergence and evolution of communication systems in dilemma games with messaging

〇Naoki Inoue[1], Junya Morita[1] （1. Shizuoka University）

3:40 PM - 4:00 PM

[2D4-OS-1a-03] Analysis of viewers’ comments on Internet TV news

〇Tomosato Nishi[1], Yuki Ogawa[1], Hiromitsu Hattori[1], Fumiaki Taka[2], Masanori Takano[3], Soichiro Morishita[3] （1. Ritsumeikan University, 2. Kanagawa University, 3. CyberAgent, Inc.）

4:00 PM - 4:20 PM

[2D4-OS-1a-04] (Invited talk) Can business cards network approach unresolved problems in social science?

〇Naoki Maejima[1], Tomonori Manabe[1] （1. Sansan, Inc.）

4:20 PM - 5:00 PM

Organized Session | Organized Session | [OS] OS-1

[2D5-OS-1b] 計算社会科学(2)

鳥海 不二夫（東京大学）、笹原 和俊（名古屋大学）、吉田 光男（豊橋技術科学大学）、瀧川 裕貴（東北大学）、榊 剛史（ホットリンク）、高野 雅典（サイバーエージェント）

5:20 PM - 7:00 PM Room D (301B Medium meeting room)

[2D5-OS-1b-01] Analysis of personal attribute relationships based on online research and social data

〇Baofa Du[1], Kazutoshi Sasahara[1,2] （1. Nagoya University, 2. JST PRESTO）

5:20 PM - 5:40 PM

[2D5-OS-1b-02] Retrieving and analyzing interpersonal relationship from large-scale mobility data.

〇Kimitaka Asatani[1], Fujio Toriumi[1], Masanao Ochi[1], Junichiro Mori[1], Ichiro Sakata[1] （1. University of Tokyo）

5:40 PM - 6:00 PM

[2D5-OS-1b-03] Proposal of a method to estimate insurance claims caused by heavy rain

〇Takeshi Okazaki[1], Oriol Gaspa Rebull[1] （1. Aon Benfield Japan Ltd.）

6:00 PM - 6:20 PM

[2D5-OS-1b-04] Detection of indirect dominance by China on the global shareholding network

〇Takayuki Mizuno[1,2], Mitsuhiro Odaka[1,2] （1. National Institute of Informatics, 2. School of Advanced Sciences, SOKENDAI）

6:20 PM - 6:40 PM

[2D5-OS-1b-05] Expressing Modern/Oldfashioned Racism depending on News Contents of Internet Television

〇Masanori Takano[1,4], Fumiaki Taka[2], Soichiro Morishita[1], Tomosato Nishi[3], Yuki Ogawa[3] （1. CyberAgent, Inc., 2. Kanagawa University, 3. College of Information Science and Engineering, Ritsumeikan University, 4. The University of Tokyo）

6:40 PM - 7:00 PM

Room E

Organized Session | Organized Session | [OS] OS-9

[2E4-OS-9] 顔文字の科学

奥村 紀之（大手前大学）、Michal Ptaszynski（北見工業大学）、Rzepka Rafal（北海道大学）

3:20 PM - 5:00 PM Room E (301A Medium meeting room)

[2E4-OS-9-01] Analysis of Net Slang Expressing laugh at Comments on Nico Nico Douga

〇Shiori Aoyagi[1], Yasuo Kawai[1] （1. Graduate School of Information and Communications, Bunkyo University）

3:20 PM - 3:40 PM

[2E4-OS-9-02] A Multi-class Classification based on the Original Form of Kaomoji using Neural Network

©The Japanese Society for Artificial Intelligence

○Noriyuki Okumura[1], Rei Okumura[2] （1. Otemae University, 2. National Institute of Technology, Akashi College）
3:40 PM - 4:00 PM

[2E4-OS-9-03] Study on Potential of Automatic Emoticon Generation Based on Emoticon Parts
○Tomoaki Sakai[1], Michal Edmund Ptaszynski[1], Fumito Masui[1] （1. kitami institute of technology）
4:00 PM - 4:20 PM

[2E4-OS-9-04] Convolutional Neural Network for Chinese Sentiment Analysis Considering Chinese Slang Lexicon and Emoticons
○Da Li[1], Rafal Rzepka[1], Michal Ptaszynski[2], Kenji Araki[1] （1. Hokkaido University, 2. KITAMI Institute of Technology）
4:20 PM - 4:40 PM

[2E10-13-5] Panel discussion
4:40 PM - 5:00 PM

Room F

Organized Session | Organized Session | [OS] OS-5

[2F3-OS-5a] 複雑化社会における意思決定・合意形成のための AI技術(1)

福島 俊一（科学技術振興機構）、福田 直樹（静岡大学）、伊藤 孝行（名古屋工業大学）
1:20 PM - 3:00 PM Room F (302B Medium meeting room)

[2F3-OS-5a-01] Innovation of Large-scale Consensus Support System based on Agent Technologies:A Large-scale Social Experiment using the Automated Facilitation Agent
○Takayuki Ito[1], Daichi Shibata[1], Shota Suzuki[1], Naoko Yamaguchi[1], Tomohiro Nishida[1], Kentaro Hiraishi[1], Kai Yoshino[1] （1. Nagoya Institute of Technology）
1:20 PM - 1:40 PM

[2F3-OS-5a-02] (Invited talk) Towards Realization of Collective Intelligence based on Communication-field Mechanism Design
○Tadahiro Taniguchi[1] （1. Ritsumeikan University）
1:40 PM - 2:00 PM

[2F3-OS-5a-03] (Invited talk) Market design: Designing social mechanisms/institutions based on game theory
○Makoto Yokoo[1] （1. Graduate School and Faculty of Information Science and Electrical Engineering, Kyushu University）
2:00 PM - 2:20 PM

[2F3-OS-5a-04] A Preliminary Design and Potential Issues of LOD-based Inter-Domain Querying for a Large-Scale Consensus Support System
○Naoki Fukuta[1] （1. Shizuoka University）
2:20 PM - 2:40 PM

[2F3-OS-5a-05] Opinion dynamics theory considering interpersonal relationship of trust and distrust and media effects
○Akira Ishii Ishii[1], Yasuko Kawahata[2] （1. Tottori University, 2. Gunma University）
2:40 PM - 3:00 PM

Organized Session | Organized Session | [OS] OS-5

[2F4-OS-5b] 複雑化社会における意思決定・合意形成のための AI技術(2)

福島 俊一（科学技術振興機構）、福田 直樹（静岡大学）、伊藤 孝行（名古屋工業大学）
3:20 PM - 5:00 PM Room F (302B Medium meeting room)

[2F4-OS-5b-01] An annotation method for a discussion model based on the IBIS structure
○Naoko Yamaguchi[1], Tomohiro Nishida[1], Daichi Shibata[1], Shota Suzuki[1], Kai Yoshino[1], Kentaro Hiraishi[1], Takayuki Ito[1] （1. Nagoya Institute of Technology）
3:20 PM - 3:40 PM

[2F4-OS-5b-02] Verification of discussions using consensus building support system organized by city development
○Tomohiro Nishida[1], Takanori Ito[1], Kai Yoshino[1], Daichi Shibata[1], Naoko Yamaguchi[1], Shota Suzuki[1], Kentaro Hiraishi[1], Takayuki Ito[1] （1. Nagoya Institute of Technology）
3:40 PM - 4:00 PM

[2F4-OS-5b-03] Analysis of Discussions with Dealing Rights to Speak Mobile Application
○Hirofumi Masui[1], Takahiro Oshima[1], Tadahiro Taniguchi[1] （1. Ritsumeikan University）
4:00 PM - 4:20 PM

[2F4-OS-5b-04] Filtering of Impertinent Remarks using distributed expression
○Kentaro Hiraishi[1], Daichi Shibata[1], Tomohiro

©The Japanese Society for Artificial Intelligence

Nishida[1], Naoko Yamaguchi[1], Shota Suzuki[1], Kai Yoshino[1], Ahmed Moustafa[1], Takayuki Ito[1]
（1. Nagoya Institute of Technology）
4:20 PM - 4:40 PM

[2F4-OS-5b-05] Issues of social acceptability of autonomous system from the view point of robots
〇Shinichi Fukuzumi Fukuzumi[1], Osamu Sakura[2,1], Hiroshi Nakagawa[1] （1. RIKEN, 2. University of Tokyo）
4:40 PM - 5:00 PM

Room G

Organized Session | Organized Session | [OS] OS-2

[2G3-OS-2a] データ流通社会における技術基盤と異分野連携(1)

早矢仕 晃章（東京大学）、大澤 幸生（東京大学）
1:20 PM - 2:40 PM Room G (302A Medium meeting room)

[2G3-OS-2a-01] Data Jackets Evolving and Connecting via History of IMDJ
〇Yukio Ohsawa[1], Teruaki Hayashi[1] （1. The University of Tokyo）
1:20 PM - 1:40 PM

[2G3-OS-2a-02] Data visualization method according to thinking process on the data distribution platform with blockchain
Kosuke Miyahara[1], Yusuke Ejiri[1], Eiji Ikeda[1], 〇Hiromichi Sasaki[1] （1. Fujitsu Limited）
1:40 PM - 2:00 PM

[2G3-OS-2a-03] Efficient Privacy-Preserving Prediction for Three-Layer Feedforward Neural Networks Using Ring-LWE-based Homomorphic Encryption
〇Takehiro Tezuka[1], Lihua Wang[1,2], Takuya Hayashi[2], Sangwook Kim[1], Tomoya Tamei[1], Toshiaki Omori[1], Seiichi Ozawa[1] （1. Kobe University, 2. National Institute of Information and Communications Technology）
2:00 PM - 2:20 PM

[2G3-OS-2a-04] Network Growth Process on Cross-disciplinary Data in the Data Platform
〇Teruaki Hayashi[1], Yukio Ohsawa[1] （1. The University of Tokyo）
2:20 PM - 2:40 PM

Organized Session | Organized Session | [OS] OS-2

[2G4-OS-2b] データ流通社会における技術基盤と異分野連携(2)

早矢仕 晃章（東京大学）、大澤 幸生（東京大学）
3:20 PM - 4:40 PM Room G (302A Medium meeting room)

[2G4-OS-2b-01] Proposing the Enactment of Data Registration
〇Makoto Koike[1,2] （1. Makoto Koike Microwave Research Institute, 2. Micro Patent Office）
3:20 PM - 3:40 PM

[2G4-OS-2b-02] Design of Data Flow Process among Multiple Organizations for PR Measurement
〇Kunihiko Ueshima[1], Taito Tosaka[2], Kohei Taniguchi[3], Teruaki Hayashi[4], Yukio Ohsawa[4]
（1. Japan Data Exchange, Inc., 2. OZMA, Inc., 3. Kikkoman Corporation, 4. School of Engineering, The University of Tokyo）
3:40 PM - 4:00 PM

[2G4-OS-2b-03] (Invited talk) Expectation for data trading
〇Aki Ando[1] （1. ABeam Consulting Ltd.）
4:00 PM - 4:20 PM

[2G05-07-4] Panel discussion
4:20 PM - 4:40 PM

Room I

Organized Session | Organized Session | [OS] OS-15

[2I3-OS-15a] 人と AIが織りなす新たなエコシステム(1)

中川 裕志（理化学研究所）、山川 宏（ドワンゴ人工知能研究所／全脳アーキテクチャ・イニシアティブ）、浅田 稔（大阪大学）、井上 智洋（駒沢大学）、江間 有沙（東京大学）、金井 良太（アラヤ）、高橋 恒一（理化学研究所）、萩田 紀博（ATR）、堀川 優紀子（ATR）、松原 繁夫（京都大学）
1:20 PM - 3:00 PM Room I (306+307 Small meeting rooms)

[2I3-OS-15a-01] Future Prototyping Methodology for Knowledge Co-creation
〇MIWA NISHINAKA[1], Hideaki Takeda[2], Kunio Shirahada[3], Yusuke Kishita[4], Hisashi Masuda[5]
（1. The Graduate University for Advanced Studies, 2. National Institute of Informatics, 3. Japan Advanced Institute of Science and Technology, 4. The University of Tokyo, 5. Kyoyto University）
1:20 PM - 1:40 PM

[2I3-OS-15a-02] Collective responsibility under AI

©The Japanese Society for Artificial Intelligence

network
〇shigeo KAWASHIMA[1] （1. Aoyama Gakuin Women's Junior College, RIKEN）
1:40 PM - 2:00 PM

[2I3-OS-15a-03] Analysis of the AI and Society Meeting in Japan
〇Hitomi Sano Sano[1] （1. AI and Society lab in Keio university）
2:00 PM - 2:20 PM

[2I3-OS-15a-04] Gaming Space as an Online Society, and A Consideration about Its AI Implementation for Rapid Transformation Tolerance
〇Shinya Umeno[1], Masaya Nakamoto[1], Tetsuya Nohama[1] （1. Mynet.ai Inc.）
2:20 PM - 2:40 PM

[2I01-04-5] Discussion / Conclusion
2:40 PM - 3:00 PM

Organized Session | Organized Session | [OS] OS-15

[2I4-OS-15b] 人と AIが織りなす新たなエコシステム (2)

中川 裕志（理化学研究所）、山川 宏（ドワンゴ人工知能研究所／全脳アーキテクチャ・イニシアティブ）、浅田 稔（大阪大学）、井上 智洋（駒沢大学）、江間 有沙（東京大学）、金井 良太（アラヤ）、高橋 恒一（理化学研究所）、萩田 紀博（ATR）、堀川 優紀子（ATR）、松原 繁夫（京都大学）
3:20 PM - 5:00 PM Room I (306+307 Small meeting rooms)

[2I4-OS-15b-01] A Study on Formation of Social Consensus on Personal Data Utilization in AI Society
〇Etsuko Tane[1] （1. The University of Tokyo）
3:20 PM - 3:40 PM

[2I4-OS-15b-02] AI as agency in agency law
〇Ryota Akasaka[1] （1. National Institute of Advanced Industrial Science And Technology）
3:40 PM - 4:00 PM

[2I4-OS-15b-03] (Invited talk) Legal Issues toward the New Ecosystem Interwoven between Human and AIs
〇Satoshi Narihara[1,2] （1. Kyushu University, 2. The RIKEN Center for Advanced Intelligence Project）
4:00 PM - 4:40 PM

[2I05-07-4] Discussion / Conclusion
4:40 PM - 5:00 PM

Room J

Organized Session | Organized Session | [OS] OS-19

[2J4-OS-19a] 「プロジェクション科学」の発展と応用 (1)

岡田 浩之（玉川大学）、鈴木 宏昭（青山学院大学）
3:20 PM - 5:00 PM Room J (201B Medium meeting room)

[2J11-13-1] Aim of the session
3:20 PM - 3:40 PM

[2J4-OS-19a-01] Towards the mechanism for projection
〇Hiroki Yokoyama[1], Hiroyuki Okada[1], Hiroaki Suzuki[2] （1. Tamagawa University, 2. Aoyama Gakuin University）
3:40 PM - 4:00 PM

[2J4-OS-19a-02] Physics Projection
〇Naoto Iwahashi[1], Hideaki Negoro[1], Soichi Kawano[2] （1. Okayama Prefectural University, 2. Luke System）
4:00 PM - 4:20 PM

[2J4-OS-19a-03] (Invited talk) JIZAI body
〇Masahiko Inami[1] （1. The University of Tokyo）
4:20 PM - 5:00 PM

Organized Session | Organized Session | [OS] OS-19

[2J5-OS-19b] 「プロジェクション科学」の発展と応用 (2)

岡田 浩之（玉川大学）、鈴木 宏昭（青山学院大学）
5:20 PM - 7:00 PM Room J (201B Medium meeting room)

[2J5-OS-19b-01] Prospects of Studies on Elastic Limb(s) Illusion based on Muscletension-Elastication Correalation
〇Kenri Kodaka[1] （1. Graduate School of Design and Architecture, Nagoya-city University）
5:20 PM - 5:40 PM

[2J5-OS-19b-02] The effect of delayed visual feedback on the full body illusion
〇Itsuki Ohtsuka[1], Sotaro Shimada[1] （1. Meiji University）
5:40 PM - 6:00 PM

[2J5-OS-19b-03] Body Schema Calibration under Virtual Reality
〇Noriki Mochizuki[1], Ryo Suzuki[1], Eimei Oyama[2], Sousuke Nakamura[1] （1. Hosei University, 2. National Institute Advanced

©The Japanese Society for Artificial Intelligence

Industrial Science and Technology）
6:00 PM - 6:20 PM

[2J5-OS-19b-04] The existence of a competitor increases enthusiasm when playing video games:
〇Nobuyuki Kawai[1], Ryuzaburo Nakata[1], Kenta Kubo[1] （1. Nagoya University）
6:20 PM - 6:40 PM

Room K

Organized Session | Organized Session | [OS] OS-16

[2K4-OS-16a] 臨床の知(1)

諏訪 正樹（慶應義塾大学）、藤井 晴行（東京工業大学）、加藤 文俊（慶應義塾大学）
3:20 PM - 4:40 PM Room K (201A Medium meeting room)

[2K4-OS-16a-01] Consideration on Onigokko as a healthcare content for children in artificial intelligence society
〇Yuushi Hiramine[1,2] （1. Onigokko Association, 2. Onigokko Research Institute）
3:20 PM - 3:40 PM

[2K4-OS-16a-02] Types of Knowledge Emerging from Photos and Language Expression of Experience
〇Haruyuki Fujii[1], Sadayo Hirata[2], Ken-icih Shinozaki[3] （1. Tokyo Institute of Technology, 2. Shibaura Institute of Technology, 3. Nihon University）
3:40 PM - 4:00 PM

[2K4-OS-16a-03] Process of Comprehension of Constructing Dry Stone Walls in Izena in Ryukyu
〇Kenichi Shinozaki[1], Sadayo Hiram[2], Haruyuki Fujii[3] （1. Nihon University, 2. Shibaura Institute of Technology, 3. Tokyo Institute of Technology）
4:00 PM - 4:20 PM

[2K4-OS-16a-04] Toward the Aesthetic Appreciation of the Haptics and Tastes
〇Hiroki Fukushima[1] （1. Kyushu Women's University）
4:20 PM - 4:40 PM

Organized Session | Organized Session | [OS] OS-16

[2K5-OS-16b] 臨床の知(2)

諏訪 正樹（慶應義塾大学）、藤井 晴行（東京工業大学）、加藤 文俊（慶應義塾大学）
5:20 PM - 6:40 PM Room K (201A Medium meeting room)

[2K5-OS-16b-01] Construction of Design Community by designer and users of place for expression
〇maiko kobayakawa[2,3], takeshi sunaga[1], sunao maruyama[1], tomoki hirano[2], kosuke yamada[4], kakuichi nishimura[3], kentaro watanabe[3], sachiko fujimitsu[5] （1. Tokyo University of the arts, 2. Tokyo University of the Arts Graduate School, 3. National Institute of Advanced Industrial Science and Technology, 4. Keio University, 5. Saga University Hospital ）
5:20 PM - 5:40 PM

[2K5-OS-16b-02] On learning through metaphorical imaginations
〇Fumitoshi Kato[1] （1. Keio University）
5:40 PM - 6:00 PM

[2K5-OS-16b-03] Segmental response of the human body to changes in the interrelationship with the environment
〇KAZUYA TANAKA[1], Tomoaki Atomi[2], Wataru Takano[3], Soichiro Fujiki[4], Katsuya Hasegawa[5], Miho Shimizu[6], Yoriko Atomi[6] （1. Teikyo University of science, 2. Kyorin University, 3. Osaka University, 4. Dokkyo Medical University, 5. Japan Aerospace Exploration Agency, 6. Tokyo University of Agriculture and Technology）
6:00 PM - 6:20 PM

[2K5-OS-16b-04] Savage Experiment
〇Takahito Horiuchi[1], Masaki Suwa[2] （1. Graduate School of Media and Governance, Keio University, 2. Faculty of Environment and Information Studies, Keio University）
6:20 PM - 6:40 PM

©The Japanese Society for Artificial Intelligence

Thu. Jun 6, 2019

Room D

Organized Session | Organized Session | [OS] OS-4

[3D3-OS-4a] 自律・創発・汎用 AIアーキテクチャ(1)

栗原 聡（慶應義塾大学）、川村 秀憲（北海道大学）、津田 一郎（中部大学）、大倉 和博（広島大学）

1:50 PM - 3:10 PM Room D (301B Medium meeting room)

[3D3-OS-4a-01] Story Association Mediated by Individual and General Concepts

〇Taisuke Akimoto[1] （1. Kyushu Institute of Technology）

1:50 PM - 2:10 PM

[3D3-OS-4a-02] How "intelligence" is called as "intelligence"?

〇Masayuki Yoshinobu[1] （1. Freelance）

2:10 PM - 2:30 PM

[3D3-OS-4a-03] Double Articulation Analyzer with Prosody for Unsupervised Word Discovery

〇Yasuaki Okuda[1], Ryo Ozaki[1], Tadahiro Taniguchi[1] （1. Ritsumeikan University）

2:30 PM - 2:50 PM

[3D3-OS-4a-04] Generating Collective Behavior of a Robotic Swarm in a Two-landmark Navigation Task with Deep Neuroevolution

〇Daichi Morimoto[1], Motoaki Hiraga[1], Kazuhiro Ohkura[1], Yoshiyuki Matsumura[2] （1. Hiroshima University, 2. Shinsyu University）

2:50 PM - 3:10 PM

Organized Session | Organized Session | [OS] OS-4

[3D4-OS-4b] 自律・創発・汎用 AIアーキテクチャ(2)

栗原 聡（慶應義塾大学）、川村 秀憲（北海道大学）、津田 一郎（中部大学）、大倉 和博（広島大学）

3:50 PM - 5:10 PM Room D (301B Medium meeting room)

[3D4-OS-4b-01] Introducing a Call Stack into the RGoal Hierarchical Reinforcement Learning Architecture

〇Yuuji Ichisugi[1], Naoto Takahashi[1], Hidemoto Nakada[1], Takashi Sano[2] （1. National Institute of Advanced Industrial Science and Technology (AIST), 2. Department of Computer and Information Science, Faculty of Science and Technology, Seikei University）

3:50 PM - 4:10 PM

[3D4-OS-4b-02] Avoiding catastrophic forgetting in echo state networks by minimizing the connection cost

〇Yuji Kawai[1], Yuho Ozasa[1], Jihoon Park[1], Minoru Asada[1] （1. Osaka University）

4:10 PM - 4:30 PM

[3D05-06-3] Special program

4:30 PM - 5:10 PM

Room E

Organized Session | Organized Session | [OS] OS-12

[3E3-OS-12a] 画像と AI（ MIRU2019プレビュー）(1)

長原 一（大阪大学）、川崎 洋（九州大学）、岡部 孝弘（九州工業大学）

1:50 PM - 3:10 PM Room E (301A Medium meeting room)

[3E3-OS-12a-01] Convolutional Neural Network for Image Recognition and Visual Explanation

〇Takayoshi Yamashita[1] （1. Chubu University）

1:50 PM - 2:30 PM

[3E3-OS-12a-02] Adaptive selection of auxiliary tasks in UNREAL

〇Hidenori Itaya[1], Tsubasa Hirakawa[1], Yamashita Takayoshi[1], Fujiyoshi Hironobu[1] （1. Chubu University）

2:30 PM - 2:50 PM

[3E3-OS-12a-03] ShakeDrop Regularization for ResNet Family

〇Yoshihiro Yamada[1], Masakazu Iwamura[1], Koichi Kise[1] （1. Osaka Prefecture University）

2:50 PM - 3:10 PM

Organized Session | Organized Session | [OS] OS-12

[3E4-OS-12b] 画像と AI（ MIRU2019プレビュー）(2)

長原 一（大阪大学）、川崎 洋（九州大学）、岡部 孝弘（九州工業大学）

3:50 PM - 5:10 PM Room E (301A Medium meeting room)

[3E4-OS-12b-01] Neural 3D Mesh Renderer

〇Hiroharu Kato[1], Yoshitaka Ushiku[1], Tatsuya Harada[1] （1. The University of Tokyo）

3:50 PM - 4:10 PM

[3E4-OS-12b-02] Between-class Learning for Image Classification

〇Yuji Tokozume[1], Yoshitaka Ushiku[1],

©The Japanese Society for Artificial Intelligence

Tatsuya Harada[1,2] （1. The University of Tokyo, 2. RIKEN）
4:10 PM - 4:30 PM

[3E4-OS-12b-03] A Generative Framework for Creative Data Based on the Generative Adversarial Networks
〇Riku Fujimoto[1], Takato Horii[1], Tatsuya Aoki[1], Takayuki Nagai[1,2] （1. The University of Electro-Commnications, 2. Osaka University）
4:30 PM - 4:50 PM

[3E4-OS-12b-04] Modeling of prejudice eyes by designers
〇Ryuichi Ishikawa[1], Kou Izumi[1], Hidekazu Hayashi[1], Hiroyuki Fukuda[2] （1. DentsuDigital,Inc, 2. Dentsu,Inc）
4:50 PM - 5:10 PM

Room F

Organized Session | Organized Session | [OS] OS-14

[3F3-OS-14a] 人狼知能と不完全情報ゲーム(1)

稲葉 通将（広島市立大学）、片上 大輔（東京工芸大学）、狩野 芳伸（静岡大学）、大槻 恭士（山形大学）
1:50 PM - 3:30 PM Room F (302B Medium meeting room)

[3F3-OS-14a-01] Analyzing Gestures in Real-World Werewolf Game
〇Shutarou Takayama[1], Hirotaka Osawa[1] （1. University of Tsukuba）
1:50 PM - 2:10 PM

[3F3-OS-14a-02] “Goodness” analysis of a Werewolf Game’s player based on a biological signal
〇Hirotaka Yamamoto[1], Nagisa Munekata[1] （1. Kyoto Sangyo University）
2:10 PM - 2:30 PM

[3F3-OS-14a-03] Simulation of Strategic Evolution in 5-player Werewolf
〇Atsushi Takeda[1], Fujio Toriumi[1] （1. The University of Tokyo）
2:30 PM - 2:50 PM

[3F3-OS-14a-04] Simulation Analysis based on Behavioral Experiment of Cooperative Pattern Task
〇Norifumi Watanabe[1], Kota Itoda[2] （1. Musashino University, 2. Keio University）
2:50 PM - 3:10 PM

[3F3-OS-14a-05] Analyis of Strategies in Werewolf Game by 3 Players Considering Voting Behaviour of Other Players
〇Hinako Tamai[1] （1. Nara Women's University）
3:10 PM - 3:30 PM

Organized Session | Organized Session | [OS] OS-14

[3F4-OS-14b] 人狼知能と不完全情報ゲーム(2)

稲葉 通将（広島市立大学）、片上 大輔（東京工芸大学）、狩野 芳伸（静岡大学）、大槻 恭士（山形大学）
3:50 PM - 5:30 PM Room F (302B Medium meeting room)

[3F4-OS-14b-01] Development and evaluation of the game agent to change confidence of estimating in Cooperative game Hanabi
〇Eisuke Sato[1], Hirotaka Osawa[1] （1. University of Tsukuba）
3:50 PM - 4:10 PM

[3F4-OS-14b-02] Role estimation in Werewolf games using a selective desensitization neural network
Masahiro Saito[2], Seiryu Mishina[1], 〇Ken Yamane[1] （1. Teikyo University, 2. Fujisoft Incorporated）
4:10 PM - 4:30 PM

[3F4-OS-14b-03] Artificial Intelligence for Deducing Roles of Players in the Werewolf Game using Information about Conversations among Players
〇Munemichi Fukuda[1], Hajime Anada[1] （1. Tokyo City University）
4:30 PM - 4:50 PM

[3F4-OS-14b-04] Extraction of Interpretable Rules for Role and Team Estimation in AI Wolf
Yuki Omura[1], Wataru Sakamoto[1], 〇Tomonobu Ozaki[1] （1. Nihon University）
4:50 PM - 5:10 PM

[3F06-09-5] Discussion / Conclusion
5:10 PM - 5:30 PM

Room G

Organized Session | Organized Session | [OS] OS-18

[3G3-OS-18a] 感情と AI(1)

日永田 智絵（電気通信大学）、堀井 隆斗（電気通信大学）、長井 隆行（大阪大学）
1:50 PM - 3:30 PM Room G (302A Medium meeting room)

[3G3-OS-18a-01] (Invited talk) Predictive coding of interoception and affective emergence

©The Japanese Society for Artificial Intelligence

○Hideki Ohira[1] （1. Nagoya University）
1:50 PM - 2:30 PM

[3G3-OS-18a-02] Action Selection based on Somatic Marker Hypothsis
○Chie Hieida[1], Takato Horii[1], Takayuki Nagai[1,2] （1. The University of Electro-Communications, 2. Osaka University）
2:30 PM - 2:50 PM

[3G3-OS-18a-03] Active perception based on free-energy minimization on restricted Boltzmann machines
○Takato Horii[1], Takayuki Nagai[1,2] （1. The University of Electro-Communications, 2. Osaka University）
2:50 PM - 3:10 PM

[3G3-OS-18a-04] Extraction of Neuroscientific Findings by Visualization of Deep Neural Network
○Kazuki Sakuma[1], Junya Morita[1], Taiki Nomura[2], Takatsugu Hirayama[3], Yu Enokibori[3], Kenji Mase[3] （1. Faculty of Informatics, Shizuoka University, 2. Graduate School of Information Science, Nagoya University, 3. Graduate School of Informatics, Nagoya University）
3:10 PM - 3:30 PM

Organized Session | Organized Session | [OS] OS-18

[3G4-OS-18b] 感情と AI(2)

日永田 智絵（電気通信大学）、堀井 隆斗（電気通信大学）、長井 隆行（大阪大学）
3:50 PM - 5:30 PM Room G (302A Medium meeting room)

[3G4-OS-18b-01] Automatic Detection of Insincere Utterances with LSTM
○Kazuya Mio[1], Aya Ishino[2], Kazuya Mera[1], Toshiyuki Takezawa[1] （1. Graduate School of Information Sciences, Hiroshima City University, 2. Faculty of Media Business, Hiroshima University of Economics）
3:50 PM - 4:10 PM

[3G4-OS-18b-02] A Study on Concept Acquisition Method for User Preference in Dialogue for Empathy
○Takahisa Uchida[1,2], Takashi Minato[2], Yutaka Nakamura[1], Yuichiro Yoshikawa[1], Hiroshi Ishiguro[1,2] （1. Osaka University, 2. Advanced Telecommunications Research Institute International）
4:10 PM - 4:30 PM

[3G4-OS-18b-03] A personalized model to estimate emotion of individual based on observed facial expression and generation of robot behavior
○Kazumi Kumagai[1], Ikuo Mizuuchi[1] （1. Tokyo University of Agriculture and Technology）
4:30 PM - 4:50 PM

[3G4-OS-18b-04] Human inference system derived from emotion as a value calculation system hypothesis
○Masahiro Miyata[1], Takashi Mory[2] （1. Graduate School of Engineering, Tamagawa University, 2. College of Engineering, Tamagawa University）
4:50 PM - 5:10 PM

[3G4-OS-18b-05] Agents which generate and estimate emotions based on logical model
○Mai Tsukamoto[1] （1. Nara Women's University）
5:10 PM - 5:30 PM

Room L

Organized Session | Organized Session | [OS] OS-22

[3L3-OS-22a] 創作者と人工知能が創る創作の未来(1)

上野 未貴（豊橋技術科学大学）、森 直樹（大阪府立大学）、はたなか たいち（(株)クリエイターズインパック）
1:50 PM - 3:30 PM Room L (203+204 Small meeting rooms)

[3L3-OS-22a-01] (Invited talk) The Future of Animation Industry Drawn by Creators and Artificial Intelligence
○Taichi Hatanaka[1] （1. Creators in Pack Inc.）
1:50 PM - 2:10 PM

[3L3-OS-22a-02] A Proposal of a Mathematical Story Generation Method Based on Hero’s Journey
○noguchi katsuhiro[1] （1. comic artist）
2:10 PM - 2:30 PM

[3L3-OS-22a-03] Narrative Structure Analysis Punchlines of SF Genre within the Flash Fiction of Shinichi Hoshi
○Shuuhei Toyosawa[1], Hajime Murai[1] （1.

©The Japanese Society for Artificial Intelligence

Future University Hakodate）
2:30 PM - 2:50 PM

[3L3-OS-22a-04] Story Creation System based on Sentence Similarity for Supporting Contents Creation
〇Haruka Takahashi[1], Miki Ueno[1], Hitoshi Isahara[1] （1. Toyohashi University of Technology）
2:50 PM - 3:10 PM

[3L3-OS-22a-05] Classification method for four-scene comics based on Creative Viewpoint
〇Naoki Mori[1] （1. Osaka Prefecture University）
3:10 PM - 3:30 PM

Organized Session | Organized Session | [OS] OS-22

[3L4-OS-22b] 創作者と人工知能が創る創作の未来(2)

上野 未貴（豊橋技術科学大学）、森 直樹（大阪府立大学）
3:50 PM - 5:30 PM Room L (203+204 Small meeting rooms)

[3L4-OS-22b-01] Autonomous Collection and Feedback of Dialogue Profile for Dialogue Contents Circulation
Masaki Mori[1], 〇Akinobu Lee[1] （1. Nagoya Institute of Technology, Japan）
3:50 PM - 4:10 PM

[3L4-OS-22b-02] A Discussion about Jazz Improvisation System based on Studies with First-Person's View
〇Daichi Ando[1] （1. Tokyo Metropolitan University）
4:10 PM - 4:30 PM

[3L4-OS-22b-03] Of application of artificial intelligence to fashion design
〇Yoko Fujishima[1,3], Osamu Sakura[2,3] （1. Graduate School of Interdisciplinary Information Studies, The University of Tokyo, 2. Interfaculty Initiative in Information Studies, The University of Tokyo, 3. RIKEN-AIP）
4:30 PM - 4:50 PM

[3L4-OS-22b-04] Tone Pasting Using cGANs with Tone Feature Loss
〇Koki Tsubota[1], Kiyoharu Aizawa[1] （1. The University of Tokyo）
4:50 PM - 5:10 PM

[3L4-OS-22b-05] Creative Future will be Created by Contents Creators and Artifical Intelligence
〇Miki Ueno[1] （1. Toyohashi University of Technology）
5:10 PM - 5:30 PM

Room P

Organized Session | Organized Session | [OS] OS-20

[3P3-OS-20] 脳波から音声言語情報を抽出・識別する

新田 恒雄（早稲田大学/豊橋技術科学大学）、桂田 浩一（東京理科大学）、入部 百合絵（愛知県立大学）、田口 亮（名古屋工業大学）
1:50 PM - 3:30 PM Room P (Front-left room of 1F Exhibition hall)

[3P3-OS-20-01] Describing Brain Activity Evoked by Speech Stimuli
〇Rino Urushihara[1], Ichiro Kobayashi[1], Hiroto Yamaguchi[2,3], Tomoya Nakai[2,3], Shinji Nishimoto[2,3] （1. Ochanomizu University, 2. National Institute of Information and Communications Technology, 3. Osaka University）
1:50 PM - 2:10 PM

[3P3-OS-20-02] Development of Syllable Labelling Tool for Electroencephalogram Data
〇Mingchuan Fu[1], Ryo Taguchi[1], Kentaro Fukai[2], Kouichi Katsurada[2], Tsuneo Nitta[3,4] （1. Nagoya Institute of Technology, 2. Tokyo University of Science, 3. Waseda University, 4. Toyohashi University of Technology）
2:10 PM - 2:30 PM

[3P3-OS-20-03] Word Recognition from speech-imagery EEG
〇Satoka Hirata[1], Yurie Iribe[1], Kentaro Fukai[2], Kouichi Katsurada[2], Tsuneo Nitta[3,4] （1. Aichi Prefectural Univ., 2. Tokyo Univ. of Science, 3. Waseda Univ., 4. Toyohashi Univ. of Tech.）
2:30 PM - 2:50 PM

[3P3-OS-20-04] Syllable recognition in speech-imagery EEG
〇Kentaro Fukai[1], Hidefumi Ohmura[1], Kouichi Katsurada[1], Tsuneo Nitta[2,3] （1. Tokyo University of Science, 2. Waseda University, 3. Toyohashi University of Technology）
2:50 PM - 3:10 PM

[3P3-OS-20-05] BCI Studies on Extraction of Spoken Language representation from Speech Imagery EEG

©The Japanese Society for Artificial Intelligence

○Tsuneo Nitta[1], KENTAROU FUKAI[2], KOUICHI KATSURADA[2], YURIE IRIBE[3], RYOU TAGUCHI[4], SHUNJI SUGIMOTO[5], JUNSEI HORIKAWA[5] （1. Waseda University, 2. Tokyo Science University, 3. Aichi Prefectural University, 4. Nagoya Institute of Technology, 5. Toyohashi University of Technology）

3:10 PM - 3:30 PM

©The Japanese Society for Artificial Intelligence

Fri. Jun 7, 2019

Room D

Organized Session | Organized Session | [OS] OS-6

[4D2-OS-6] 人工知能と倫理

神崎 宣次（南山大学）、久木田 水生（名古屋大学）、服部 宏充（立命館大学）

12:00 PM - 1:20 PM Room D (301B Medium meeting room)

[4D2-OS-6-01] Trust in Medical AI

〇Arisa Ema[1], Katsue Nagakura[2], Takanori Fujita[3] （1. The University of Tokyo, 2. M3, Inc., 3. Keio University）

12:00 PM - 12:20 PM

[4D2-OS-6-02] A Study on Constructing the Ethical Design Sciences

〇Kaira Sekiguchi[1], Koichi Hori[1,2] （1. Graduate School of Engineering, The University of Tokyo, 2. Center for Advanced Intelligence Project, RIKEN）

12:20 PM - 12:40 PM

[4D2-OS-6-03] Practice of a support for creating self-regulations on profiling

〇Fumiko Kudo[1] （1. Institute of Economics Chukyo University）

12:40 PM - 1:00 PM

[4D2-OS-6-04] AI in Sci-fi Movies from the Viewpoint of Human Centered Design

〇Shigeyoshi IIZUKA[1] （1. Kanagawa University）

1:00 PM - 1:20 PM

Room E

Organized Session | Organized Session | [OS] OS-7

[4E2-OS-7a] AIの法学の応用(1)

佐藤 健（国立情報学研究所）、新田 克己（国立情報学研究所）

12:00 PM - 1:40 PM Room E (301A Medium meeting room)

[4E2-OS-7a-01] (Invited talk) Offender profiling for criminal investigation and application of machine learning

〇Zaitsu Wataru[1] （1. Forensic Science Laboratory, Toyama Prefectural Police Headquarters）

12:00 PM - 12:40 PM

[4E2-OS-7a-02] Extracting Important Sentences with Random Forest for Statute Summarization

〇Yasuhiro Ogawa[1,2], Michiaki Satou[2], Takahiro Komamizu[1,2], Katsuhiko Toyama[1,2] （1. Information Technology Center, Nagoya University, 2. Graduate School of Informatics, Nagoya University）

12:40 PM - 1:00 PM

[4E2-OS-7a-03] Legislation Supporting Systems by AI

〇Tokuyasu Kakuta[1] （1. Chuo University）

1:00 PM - 1:20 PM

[4E2-OS-7a-04] People's Expectation and Attitude toward AI-assisted Courts

〇Shozo Shozo OTA[1] （1. The University of Tokyo）

1:20 PM - 1:40 PM

Organized Session | Organized Session | [OS] OS-7

[4E3-OS-7b] AIの法学の応用(2)

佐藤 健（国立情報学研究所）、新田 克己（国立情報学研究所）

2:00 PM - 3:40 PM Room E (301A Medium meeting room)

[4E3-OS-7b-01] Question Analysis and Rule Design for Generating PROLEG Rules towards Automatic Legal Bar Exam Solver

〇Ryuji Hayashi[1], Yoshinobu Kano[1] （1. Shizuoka University）

2:00 PM - 2:20 PM

[4E3-OS-7b-02] Problem classification of legal bar exam and structured dictionary design for human relationships in legal documents

〇Naoki Kiyota[1], Yoshinobu Kano[1] （1. Shizuoka University）

2:20 PM - 2:40 PM

[4E3-OS-7b-03] Reasoning on a bipolar argumentation framework for a law

〇Tatsuki Kawasaki[1], Sosuke Moriguchi[1], Kazuko Takahashi[1] （1. Kwansei Gakuin University）

2:40 PM - 3:00 PM

[4E3-OS-7b-04] Towards expanding the scope of trials via machine learning

〇Tatsuro Kawamoto Kawamoto[1], Masami Yasumoto[1], Yoichi Motomura[1] （1. National Institute of Advanced Industrial Science and Technology）

3:00 PM - 3:20 PM

[4E3-OS-7b-05] Implementation of Criminal Procedure using PROLEG Technology

〇Ken Satoh[1], Yoshiaki Nishigai[2] （1. National

©The Japanese Society for Artificial Intelligence

Institute of Informatics, 2. Nihon University)
3:20 PM - 3:40 PM

Room F

Organized Session | Organized Session | [OS] OS-11

[4F2-OS-11a] 社会的信号処理と AI(1)

岡田 将吾（北陸先端科学技術大学院大学）、石井 亮（NTT）
12:00 PM - 1:40 PM Room F (302B Medium meeting room)

[4F2-OS-11a-01] The Interaction between Dancer and DJ in Performance Scene of Breakdance
〇Daichi Shimizu[1], Takeshi Okada[1] (1. The University of Tokyo)
12:00 PM - 12:20 PM

[4F2-OS-11a-02] Estimating Verbal・ Nonverbal Skills in Business Presentation
〇Yagi Yutaro[1], Okada Shogo[1], Shiobara Shota[2], Sugimura Sota[2] (1. Japan Advanced Institute of Science and Technology, 2. SoftBank Corp.)
12:20 PM - 12:40 PM

[4F2-OS-11a-03] Measuring affective sharing between two people from simultaneously recorded EEG signals
〇Taiki Kinoshita[1], Hiroki Tanaka[1], Koichiro Yoshino[1], Satoshi Nakamura[1] (1. Nara Institute of Science and Technology)
12:40 PM - 1:00 PM

[4F2-OS-11a-04] Estimating performance of group interaction based on product dimension
〇Go Miura[1], Shogo Okada[1] (1. Japan Advanced Institute of Science and Technology)
1:00 PM - 1:20 PM

[4F2-OS-11a-05] Use of Dialogue History for User's Interest Level Estimation Based on Multimodal Information
〇Haruto Nishimoto[1], Ryu Takeda[1], Kazunori Komatani[1] (1. The Institute of Scientific and Industrial Research (ISIR), Osaka University)
1:20 PM - 1:40 PM

Organized Session | Organized Session | [OS] OS-11

[4F3-OS-11b] 社会的信号処理と AI(2)

岡田 将吾（北陸先端科学技術大学院大学）、石井 亮（NTT）
2:00 PM - 3:00 PM Room F (302B Medium meeting room)

[4F3-OS-11b-01] Considering Idling Motion Generation using RNNs for Virtual Conversational Agents
〇Hung-Hsuan Huang[1,2], Masato Fukuda[1], Toyoaki Nishida[1,2] (1. RIKEN Center for Advanced Intelligence, 2. Graduate School of Informatics, Kyoto University)
2:00 PM - 2:20 PM

[4F3-OS-11b-02] Identifying Discourse Boundaries in Group Discussions using Multimodal Features
Ken Tomiyama[1], 〇Fumio Nihei[1], Yutaka Takase[2], Yukiko Nakano[2] (1. Graduate School of Science and Technology, Seikei University, 2. Faculty of Science and Technology, Seikei University)
2:20 PM - 2:40 PM

[4F3-OS-11b-03] A corpus-based typology of reactions to apologies in Japanese}
〇Masashi Inoue[1], Kaori Suzuki[1] (1. Tohoku Institute of Technology)
2:40 PM - 3:00 PM

Room G

Organized Session | Organized Session | [OS] OS-8

[4G2-OS-8a] マイニングと知識創発(1)

砂山 渡（滋賀県立大学）、加藤 恒昭（東京大学）、西原 陽子（立命館大学）、森 辰則（横浜国立大学）、高間 康史（首都大学東京）
12:00 PM - 1:40 PM Room G (302A Medium meeting room)

[4G2-OS-8a-01] Proposal of Summarization Support System Based on Mixed-Initiative Interaction
〇Hiroto Kato[1], Yasufumi Takama[1] (1. Tokyo Metropolitan University)
12:00 PM - 12:20 PM

[4G2-OS-8a-02] *differential* PLSA
〇Koji Nomori[1] (1. Analytics Design Lab Inc.)
12:20 PM - 12:40 PM

[4G2-OS-8a-03] Basic Analysis of Descriptions on Medicine-care in Nursing Records
〇Seiya Uesaka[1], Yoshinori Hijikata[1], Muneou Suzuki[2] (1. Kwansei Gakuin University , 2. University of Miyazaki Hospital)
12:40 PM - 1:00 PM

[4G2-OS-8a-04] Extraction of Classification Patterns from

©The Japanese Society for Artificial Intelligence

Deep Learning Networks

〇MASAYUKI ANDO[1], Yoshinobu Kawahara[3,4], Wataru Sunayama[2], Yuji Hatanaka[2] （1. Graduate School of Engineering, The University of Shiga Prefecture, 2. School of Engineering, The University of Shiga Prefecture, 3. Institute of Mathematics for Industry, Kyushu University, 4. RIKEN Center for Advanced Intelligence Project）

1:00 PM - 1:20 PM

[4G2-OS-8a-05] Attachment of “ Cuteness” to Dialogue Sentences Using Deep Learning

〇Yusuke Goto[1], Wataru Sunayama[2], Yuji Hatanaka[2], Kazunori Ogohara[2] （1. Graduate School of Engineering, The University of Shiga Prefecture, 2. School of Engineering,The University of Shiga Prefecture）

1:20 PM - 1:40 PM

Organized Session | Organized Session | [OS] OS-8

[4G3-OS-8b] マイニングと知識創発(2)

砂山 渡（滋賀県立大学）、加藤 恒昭（東京大学）、西原 陽子（立命館大学）、森 辰則（横浜国立大学）、高間 康史（首都大学東京）

2:00 PM - 2:40 PM Room G (302A Medium meeting room)

[4G3-OS-8b-01] AI Player Creation Reflecting Personal and Characters in PvP Games

〇Toshinobu Sakata[1], Wataru Sunayama[1], Yuji Hatanaka[1], Kazunori Ogohara[1] （1. The university of shiga prefecture）

2:00 PM - 2:20 PM

[4G3-OS-8b-02] Predictability of Number of Visitors from GPS Location Big Data

Hirotoshi Yanagi[1,3], 〇Yoshitoshi Nakahara[2], Takahiro Hoshino[2,3] （1. Keio University, Graduate school of Economics, 2. Keio University, Faculty of Economics, 3. Riken AIP）

2:20 PM - 2:40 PM

©The Japanese Society for Artificial Intelligence

Organized Session | Organized Session | [OS] OS-10

[1D2-OS-10a] 不動産と AI(1)

清田 陽司（（株）LIFULL）、山崎 俊彦（東京大学）、諏訪 博彦（奈良先端科学技術大学院大学）、清水 千弘（日本大学）、橋本 武彦（GA technologies）

Tue. Jun 4, 2019 1:20 PM - 3:00 PM Room D (301B Medium meeting room)

[1D2-OS-10a-01] Assessment of Thermal Insulation and Sound Performance of Apartments Based on IoT Sensing

〇Toshihiko Yamasaki[1], Yuki Obuchi[1], Yuan Lin[1], Ryoma Kitagaki[2], Satoshi Toriumi[3], Mikihisa Hayashi[3], Ai Sakai[4], Nobuhito Haga[4], Shimpei Nomura[4], Yoichi Ikemoto[4] （1. The University of Tokyo, 2. Hokkaido University, 3. Future Standard Co., Ltd., 4. Recruit Sumai Company Ltd.）

1:20 PM - 1:40 PM

[1D2-OS-10a-02] Data Collection for the Comfort Level of Rental Property by IoT Sensing

〇Hirohiko Suwa[1], Atsushi Otsubo[1], Yugo Nakamura[1], Masahito Noguchi[2] （1. Nara Institute of Science and technology, 2. LIFULL Co., Ltd.）

1:40 PM - 2:00 PM

[1D2-OS-10a-03] Business utilization of real estate image classification system using deep learning

〇Tomoya Tsukahara[1], Kodai Sudo[1] （1. BrainPad Inc.）

2:00 PM - 2:20 PM

[1D2-OS-10a-04] Proposal of Railway Route Search System for Finding Living Place

〇Keisuke Kikuchi[1], Kenichiro Kobayashi[2], Takehiko Hashimoto[2], Yasufumi Takama[1] （1. Tokyo Metropolitan University, 2. GA Technologies）

2:20 PM - 2:40 PM

[1D2-OS-10a-05] Evaluation of Rent Prediction Models using Floor Plan Images

〇Ryosuke Hattori[1], Kazushi Okamoto[1], Atsushi Shibata[2] （1. The University of Electro-Communications, 2. Advanced Institute of Industrial Technology）

2:40 PM - 3:00 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

IoTセンシングによる不動産物件の断熱・防音性能評価

Assessment of Thermal Insulation and Sound Performance of Apartments Based on IoT Sensing

山崎 俊彦, 大渕 友暉, 林 遠 *1
Toshihiko Yamasaki, Yuki Obuchi, Yuan Lin

北垣 亮馬 *2
Ryoma Kitagaki

鳥海 哲史, 林 幹久 *3
Satoshi Toriumi, Mikihisa Hayashi

酒井 藍, 芳賀 宣仁, 野村 眞平, 池本 洋一 *4
Ai Sakai, Nobuhito Haga, Shimpei Nomura, Yoichi Ikemoto

*1東京大学
The University of Tokyo

*2北海道大学
Hokkaido University

*3株式会社フューチャースタンダード
Future Standard Co., Ltd.

*4株式会社リクルート住まいカンパニー
Recruit Sumai Company Ltd.

We have been developing IoT sensors for the evaluation of the comfort levels of real estate building properties. In this paper, we additionally evaluated thermal insulation performance and sound performance. We are now conducting a large-scale experiments using 60+ properties to assess the system' usability and reliability.

1. はじめに

不動産物件を選ぶ際の判断基準として、物件の場所、価格、間取り、住みやすさなど様々なものがある。場所、価格、間取りなどに関しては消費者が定量的・客観的に評価が可能である。しかし、物件の住みやすさについては物件情報から類推する、物件を扱う担当者に質問する、直接物件を見に行くなど手間を掛けないと情報が得られないばかりか、定性的な評価が一般的であった。我々は、不動産物件の住み心地・快適度に関係する様々な要因を計測可能な IoT センサを独自設計し定量的に評価する試みを行ってきた [大渕 16, 大渕 17a, Obuchi 18]。また、開発したシステムは保育や老健施設でのセンシングにも展開している [大渕 17b]。

我々の IoT センサシステムの実用化に向けてセンサの種類や計測方法を工夫し、新たに断熱・防音性能の定量化を行った。さらに、空き物件や帰省中の不在物件を利用した大規模実験を行った。本稿ではその概要について紹介する。

2. センサの構成

独自開発した IoT センサ (図 1) は温湿度、照度、加速度、紫外線、CO2、PM2.5 を計測できるセンサを搭載しており、Raspberry Pi 3 Model B で I2C インターフェースにより制御している。また、補助的に OMRON の 2JCIE-BL01 も利用し、Bluetooth により Raspberry Pi 側に取得される。計測したデータはフューチャースタンダード社の SCORER*1 というプラットフォームを利用して一括管理する。現地に据え付け、もしくはモバイル型の WiFi があればリアルタイムでデータはアップロードされ、そうでない場合はインターネットがある環境で電源 ON になったときに自動的にアップロードされる。そのため、センサから手間を掛けてデータを回収する必要がない利点がある。

連絡先: 山崎俊彦，東京大学大学院情報理工学系研究科電子情報学専攻，113-8656 東京都文京区本郷 7-3-1，yamasaki@hal.t.u-tokyo.ac.jp

*1 https://www.scorer.jp/

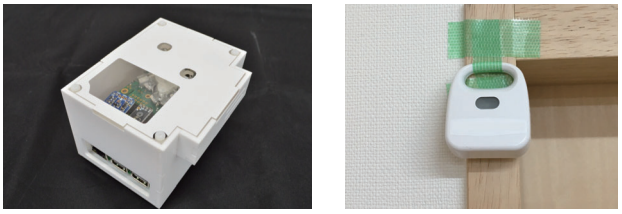


図 1: 独自開発した IoT センサと補助的に用いるオムロンの環境センサ

3. 断熱性能・防音性能の推定

3.1 断熱性能

不動産物件の省エネ性能は建築物の Q 値 (熱損失係数) や UA 値 (外皮平均熱貫流率) によって評価可能である。しかし、不動産物件の所有者や取扱業者のもつ物件の情報整備状態 (図面、設備情報、施工状態) の不十分さや、評価に必要な測定条件、測定業者の手続きが複雑である点から現実的ではない。そこで本研究では、センサを設置することで建築物の簡易な省エネ性能を評価し、不動産取引において省エネ性能の目安になる評価値を提案する。

時刻 t における室内・室外の温度をそれぞれ $T_i(t), T_o(t)$ とする。部屋のみかけの熱伝導率を λ、みかけの外皮厚を d、部屋のみかけの比熱を C すると以下の 2 つの式が成り立つ。

$$-\frac{dQ(t)}{dt} = \frac{\lambda}{d}(T_i(t) - T_o(t)) \quad (1)$$

$$\frac{dQ(t)}{dt} = C\frac{dT_i(t)}{dt} \quad (2)$$

式を整理して

$$D = \frac{\lambda}{dC} = \frac{-\frac{T_i(t)}{dt}}{T_i(t) - T_o(t)} \quad (3)$$

が部屋の熱変動のしやすさであり、一般的にはみかけの熱拡散率に近い指標となる。

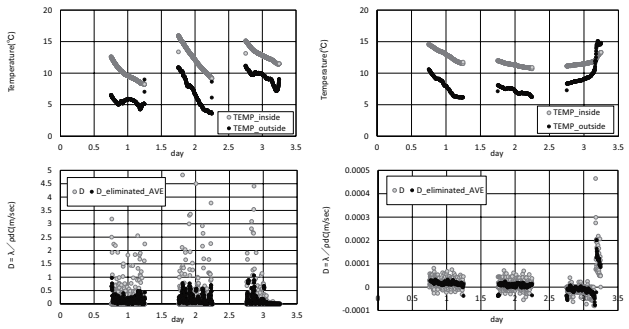
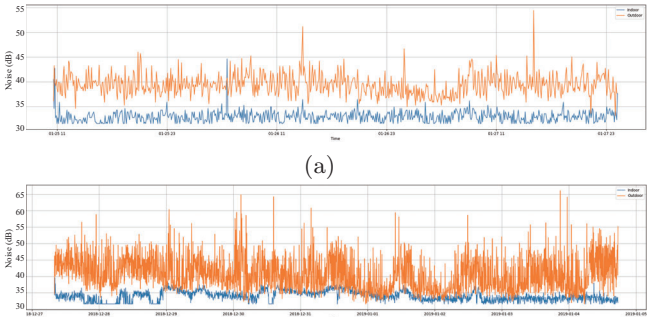


図 2: 断熱性能の簡易指標 D の計算結果。



(a)

(b)

図 3: 騒音の計測値。

室内外の床から一定距離離した場所に防水・通気加工を施したセンサを置き、深夜帯の日照のない時間帯を利用してこの D 値の平均値・標準偏差を算出して指標とする。

3.2 防音性能

建物の防音性能は外部からくるノイズレベル及び音の種類に依存するため、一般化は難しい。そのため、室内外で観測される騒音レベルを逐次計測しそれを可視化することで防音性能の評価とする。

なお、上下左右の部屋からの騒音については発生頻度の問題や騒音伝搬のメカニズムが外からの騒音と異なるなどの理由により評価が難しいので対象外とした。現在、評価方法を検討中である。

4. 測定実験

4.1 実験規模

省エネ基準は最近では 1992 年、1999 年、2013 年に改正されており、その区間内に建築された物件であれば、地域区分、構造種別が同一ならば省エネ性能はおおむね類似しているとみなすことができる。そこで、地域を関東一都三県に限定し、1992-1999 年築、1999-2013 年築、2013-2018 年築の 3 つの期間について、木造、鉄筋、RC の 3 種類の構造全てをカバーするように測定を行った。測定は空き物件、利用者が帰省などで一時的に不在としている物件を中心に行った。現在、まだ測定を継続中の物件もあるため最終的な物件数は確定していないが 60-70 件の物件について測定できる予定である。なお、空き物件と家具あり物件では計測値に差が出ることも予想され、今後詳しく検証していく。

4.2 測定結果

断熱性能の簡易指標 D の計算結果を図 2 に示す。左カラムが軽量鉄骨物件 (1996)、右カラムが木造物件 (2015) である。上段が計測した屋内外の温度の時系列変化を示し、下段がそこから計算した D 値を示している。これを見ると、D 値は木造物件 (2015) のほうが格段に小さく、その分断熱性能が優れていることがわかる。冷暖房費はこの断熱性能と部屋の広さ、用いる冷暖房器具がわかるとおおよその参考値が出せるため、簡易な省エネ性能の指標として利用できる。

騒音の計測結果を図 3 に示す。横軸は時間である。図 3(a) は比較的外が静かで 40dB 程度で安定しているため、室内も 32dB 付近で安定している。一方、図 3(b) では外で 45-60dB の騒音が発生し、それにつられて室内の騒音レベルも少し高めになっていることがわかる。また、夜間にも騒音が発生している様子が見て取れる。屋外の騒音レベルは、建物の防音性能だけでなく、窓を開放したときの騒音レベルを推定するのにも利用できる。

5. まとめ

本稿では、我々が従来より開発している住心地評価・可視化用 IoT センサを用いて、断熱・防音性能解析のための簡易的な定量化指標と実測例を示した。断熱・防音性能ともに物件により大きな性能差があることが可視化された。断熱・防音性能はこれまで消費者には不明であり、物件同士の客観的な比較も難しかった。このような情報は消費者にとって有益な参考情報になる可能性がある。また、不動産物件所有者・管理者にとっても、例えばリノベーションによる性能向上などを客観的に評価するための参考情報となりうると考えている。現在実証のための大規模実験を実施中で、近い将来の実用化を目指している。

参考文献

[大渕 16] 大渕友暉, 山崎俊彦, 相澤清晴, 鳥海哲史, 林幹久, “不動産物件の快適度評価のための IoT センサ実装と評価,” ITE 冬期大会, 11C-4, 2016.

[大渕 17a] 大渕友暉, 山崎俊彦, 相澤清晴, 鳥海哲史, 林幹久, “IoT センサを用いたマンション物件計測と快適度評価,” JSAI, 1H2-OS-15a-4, 2017.

[大渕 17b] 大渕友暉, 山崎俊彦, 鳥海哲史, 林幹久, 野澤祥子, 高橋翠, 遠藤利彦, 秋田喜代美, “保育施設における IoT カメラを用いた環境・行動解析,” IEICE-MVE, 信学技報, vol. 117, no. 217, MVE2017-15, pp.7-11, 2017.

[Obuchi 18] Yuki Obuchi, Toshihiko Yamasaki, Kiyoharu Aizwa, Satoshi Toriumi, Mikihara Hayashi, “Measurement and Evaluation of Comfort Levels of Apartments Using IoT Sensors,” ICCE2018, pp. 864-869, 2018.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

IoT センシングによる賃貸物件快適度推定のためのデータ収集

Data Collection for the Comfort Level of Rental Property by IoT Sensing

諏訪博彦[*1], 大坪敦[*1], 中村優吾[*1], 野口真史[*2]
Hirohiko Suwa, Atsushi Otsubo, Yugo Naskamura, Masahito Noguchi

[*1] 奈良先端科学技術大学院大学
Nara Institute of Science and Technology

[*2] 株式会社 LIFULL
LIFULL Co., Ltd.

Abstract: When searching for a rental property, information such as rent, breadth, time to station and age of a building is used. These indicators are quantitative data and can be compared. However, some conditions such as quiet and sunny are often described in comments in the remarks column are difficult to compare because there is ambiguity in the state recognition of each property. Therefore, indices that can quantitatively evaluate noise and sunrise are required. For indexing, it is necessary to collect data, but because the target is vacant, a data collection device cannot use power from the outlet. Therefore, it is necessary to construct an IoT device capable of sensing environmental information without supplying power from the outlet. In this paper, we develop an environmental information sensing device under the constraint that there is no electricity in the vacant rental property and a no Internet environment. As a result of the demonstration experiment based on the cooperation of real estate agents, we showed the possibility that the proposed device can collect data.

1. はじめに

賃貸物件を探索する際に，物件探索者（借主）は，場所，賃料，広さ，築年数などを検索条件として探索する．一方で，移住したあとに問題になるものとして，騒音や日当たりなどがある．借主は，物件探索時に内見するとはいえ，複数回内見する借主はまれであり，昼夜，平日/休日など，条件をかえて内見することは困難である．そのため，騒音や日当たりを認識するためには，「閑静な住宅街」「日当たり良好」などの物件に対する定性的なコメントや，主要採光面，階数などから推定せざるを得ない．しかしこの方法では，各物件の状態認識にあいまい性が残り，明確な比較は困難である．そのため，騒音や日当たりについても定量的に評価できる指標が求められている．

しかし，これらの情報すべてが物件探索ポータルサイトに明示されえいるわけではない．騒音や日当たりについては，「閑静な住宅街」「繁華街そば」，「日当たり良好」などの定性的なキャッチコピーから想像したり，主要採光面の情報から想像したりすることしかできない．そのため，これらの情報については，実際に現地へ赴き内見して獲得することが必要になるが，時間帯や季節によって変化する条件であり，正確に把握することは困難またはコストを必要とする．

この問題に対し，我々は IoT デバイスを用いた指標化を試みた[諏訪 18a]．具体的なデバイスとして，オムロンの環境センサと，Raspberry Pi を組み合わせたデバイスを開発した．構築した IoT デバイスを用いて，条件の異なる 4 件の賃貸物件についてセンシングを実施した結果，(1)静穏性，(2)防音性，(3)採光性，(4)断熱性について，想定通りに判定できることを明らかにした．

しかしながら，いくつかの課題も明らかとなった．1 つ目の課題は，デバイスのエナジーハーベスト化である．センシング対象物件は空き家であるため，基本的にはコンセントから電力を供給することはできない．そのためモバイルバッテリーを使用したが 3 日間しか稼働しなかった．週末効果などを把握するためにも，最低 1 週間の連続稼働が必要である．

2 つ目の課題は，データのアップロードの効率化である．1 つ目の課題と同様に，センシング対象物件は空き家であるためインターネット環境は存在しない．プロトタイプデバイスでは，Raspberry Pi にデータを蓄積し，手作業でデータの抽出を行った．しかしながら，実運用を考えた場合，この方法は非効率である．そのためより効率的なアップロード方法の検討が必要である．

3 つの目の課題は，操作の容易性である．構築したデバイスは，Raspberry Pi により稼働するため，その制御に一定のデバイス操作スキルが必要であった．しかし実社会で利用可能とするには，その作業を誰にでも，特に不動産屋の職員が実施可能である必要がある．

本研究では，実社会において大規模にかつ効率的に新たな賃貸物件指標を収集・活用することを目指している．そのために，本稿では，先行研究をベースとし，データ集時に必要なデバイスの要件の再抽出，プロトタイプデバイスの改良を行う．さらに，ある不動産屋の協力のもと実施したデータ収集実験について述べ，大規模化・効率化の課題について議論する．

2. 関連研究

直接的に物件探索に影響を及ぼす要因について検討した調査として，2017 年賃貸契約者に見る部屋探しの実態調査[Sumo 17]がある．この調査によると，物件探索者が部屋探しの際に重視した条件は，「家賃」が 74.7%と他の項目より突出して高く，以下「最寄り駅からの時間」（58.6%），「通勤・通学時間」（57.8%），「路線・駅やエリア」（54.7%），「間取り」（53.1%）になっている．また，大手不動産ポータルサイトの関係者によれば，移住したあとに問題になるものとして，騒音や日当たりなど指摘されている．

また，物件探索に影響を与える要因を間接的に示唆している研究として，不動産価格推定に関する研究がある．価格は，物件の良し悪しを示す代替指標と考えられ，価格に影響を及ぼす要因は，借主が物件探索時に考慮している要因と考える．

Wu[Wu 04]らは，台湾での住宅選定に影響があるといわれる風水に着目し，不動産価格推定に取り組んでいる．機械学習の特徴量には，一般的な建物固有の属性に加え，風水におけるタブーを変数として設けている．その結果，風水のタブーを考慮した方がより良好な推定結果となることを明らかにしている．

連絡先：氏名，所属，住所，電話番号，Fax 番号，電子メールアドレスなど

Chiarazzo [Chiarazzo 14] らは，ターラント市（イタリア）において，交通システムと地域毎の環境の質が不動産価格に深く関係していると考え，人工ニューラルネットワーク（ANN）を用いて検証を行っている．特徴量には，立地条件や建物の構造に加え，交通に関する属性として駅やバス停までの距離などが，環境汚染に関する属性として，SO2，NOX，NO，NO2，CO，PM10 の値と最大値がそれぞれ含まれている．この研究でも各属性の感度分析を行った結果，SO2 の最大値が全 42 ある属性の内の 8 番目に重要であることを明らかにしている．

これらの調査や研究結果は，騒音や日当たり，風水や周辺環境などが家賃推定に影響があることを示しており，借主は定量的には示されていないこれらの要因も考慮して物件探索を実施していると考えられる．しかしながら，現在の不動産ポータルサイトにおいて，これらの要因を定量的に比較することは困難である．

この問題に対して，山崎らのグループは，物件に関する指標の計測と快適度評価を試みている[大渕 17，Obuchi18]．山崎らのグループは，これまで限定的な時間帯に短時間のみしか計測できていなかった物件の様々な価値を IoT 技術にて定量化することを目的に，日照，温度・湿度，騒音，振動などを取得する IoT センサを開発している．実際にあるマンション物件で計測を行った結果，同じ物件でも階層や窓の有無により快適度が異なることを示している．

しかしながら，山崎らの研究においては電源をコンセントから確保した状態でデータ収集実験を行っている．これは，実社会で計測しようとした場合，制約になる．実運用時は，空き物件の電気は使用できず，デバイス自体に電源供給機能を持たせるか，エナジーハーベストなデバイスの構築が求められる．また，山崎らは快適度を評価するための指標について検討しているが，我々は物件探索時に比較可能な指標を構築しようとする点で視点が異なる．本稿の貢献は，実運用に向けてすべての物件において指標化するための課題について，実データ収集に基づいて議論する点である．

3. 評価指標およびデータ収集システム

本章では，収集対象とする評価指標およびデータを確認するとともに，データ収集時の要件について整理する．その上で，データ収集システムについて説明する．

3.1 評価指標および要件

先行研究[諏訪 18a]に基づいて，評価指標は静穏性，防音性，採光性，断熱性の 4 つに設定した．各指標において収集すべきデータの種類は以下の通りである．

- 静穏性：部屋の内外の音を測る
- 防音性：部屋の内外の音を測る
- 採光性：部屋の内外の明るさを測る
- 断熱性：部屋内外の温湿度を測る

また，これらの指標は時間帯による変化についても計測する必要があり，その日の天気や週末効果の影響を考慮すると 1 週間程度の計測期間が必要である．そのため，データ収集デバイスに必要な要件を以下の通りとした．

1. 電源供給環境を必要としない．
2. ネット接続環境を必要としない．
3. 手軽に持ち運び可能である．
4. 誰でも利用可能である．
5. 音量，明るさ，温度，湿度が計測可能である
6. 数日間の連続計測が可能である

3.2 システム概要

図 1 は，データ収集システムの概要である[諏訪 18b]．センシング部分は，オムロン 環境センサ 2JCIE-BL01 を用いている．このセンサは，温度・湿度・照度（明るさ）・気圧・音・UV，加速度が収集できる．またリチウム電池(CR2032)1 個で稼働し，1 分間隔でデータを取得しても，1 ヶ月程度稼働することが確認されている．

データをアップロードするゲートウェイには，Smartphone を用いる．センサとゲートウェイとの通信には BLE を用いる．センシングデータは，オムロン環境センサ内のフラッシュメモリに蓄積される．その後，Smartphone によりそうさすることで BLE を用いてデータを抽出し，Smartphone の回線を利用してクラウドにアップロードする．Smartphone の操作については，専用のアプリケーションを開発し，設置・回収の操作を誰でも可能にしている．

3.3 センサの設置およびデータ収集プロセス

図 2 は，センサ設置時のアプリの遷移を示している．センサ設置時は，まず部屋の内外に環境センサを設置する．次にアプリを起動し，センサ設置ボタンをタッチする．アプリは，設置され

図 1：システム概要

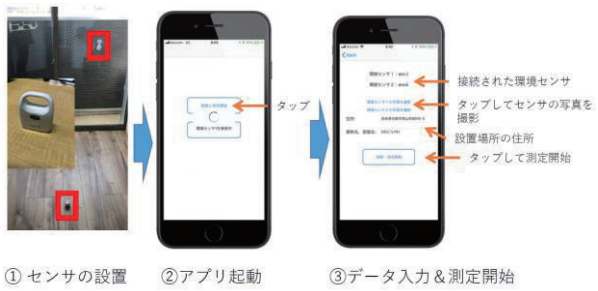


図 2 センサ設置プロセス

図 3：センサ回収プロセス

たセンサを認識し，センサ番号を表示する．この時，アプリは Smartphone の時間に基づいて環境センサの時間を合わせる．また，設置者はアプリのカメラ機能を利用して，設置状況を写真に撮影する．最後に測定開始ボタンをタップすることで測定が開始される．

また，図 3 は，センサ回収時のアプリの遷移を示している．センサ回収時は，アプリを起動し，センサ回収ボタンをタップする．これにより，アプリは環境センサと BLE 通信を開始し，環境センサのフラッシュメモリに蓄積されたデータを抽出する．抽出されたデータは，スマートフォン回線を利用して，自動的にサーバにアップロードされる．最後にセンサを回収し，測定が終了される．

4. データ収集実験

本章では，データ収集実験およびその結果について述べる．

4.1 実験概要

我々のシステムを用いて実際に評価用のデータが収集できることを確認するために，ある不動産屋に協力を依頼してデータ収集実験を行った．実験概要は以下の通りである．

- 日時:2018 年 11 月 26 日（月）～12 月 17 日（月）
- 場所:関東近県（東京都，神奈川県）
- 対象:実際の空き物件
- 物件数:47 件
- 1 件当たりデータ収集期間:1 週間程度
- デバイス設置者:不動産屋職員

4.2 収集キットおよび設置インストラクション

図 4 は，データ収集キットである．センサ 2 個，センサ操作用スマホ 1 台（充電器付き），室外設置用（窓設置）フック，設置距離確認用メジャー，実施マニュアル，データ収集案内標識がい 1 セットとなっている．実施マニュアルは，図 2，図 3 の手順を示したものである．

設置インストラクションは，不動産会社の代表者のみに説明し，実際に設置行く職員には，代表者から説明していただいた．そのため，設置者とデバイス開発者は直接面談や説明は一切していない．これにより，実際に使用する場面を想定した設置インストラクション環境を構築している．

4.3 データ収集結果

実験の結果，47 件の物件にセンサを設置することができた．収集データを確認したところ，47 件中 8 件のデータが終了時にアップロードができていなことが確認された．

図 4：データ収集キット

収集されたデータの総数は，1,240,474 データである．また，収集されたデータの日数は述べ 280 日分である．

我々の過去の研究を含め，空き物件における日当たりや騒音の程度を指標化する実験は，指標化の可能性を探るものであり，実社会における実現可能性について議論しているものは見当たらなかった．本実験において，実在する不動産屋がデータ収集を実現できたことは，空き物件における快適度などを現実社会において指標化するための可能性を示唆したと考える．

5. まとめ

本稿では，空き物件には電気がないこと，インターネットがないことを制約とした環境情報センシング装置を開発した．開発したデバイスにより，実社会においてデータ収集可能かどうかを明らかにするために，実際の不動産業者に協力を依頼し，データ収集実験を行った．不動産業者の協力による実証実験の結果，延べ 280 日分のデータが収集できた．ことから，提案システムが実社会においてデータを収集できる可能性を示されたと考える．

参考文献

[Chiarazzo 14] Vincenza Chiarazzo, Leonardo Caggiani, M. M. M. O.: A Neural Network based model for real estate price estimation considering environmental quality of property location, Transportation Research Procedia 3, pp. 810–817 (2014).

[大渕 17] 大渕友暉，山崎俊彦，相澤清晴，鳥海哲史， 林幹久：IoT センサを用いたマンション物件計測と快適度評価，第 31 回人工知能学会全国大会（JSAI2017) (2017).

[Obuchi 18] Obuchi, Y., Yamasaki, T., Aizawa, K., Toriumi, S. and Hayashi, M.: Measurement and evaluation of comfort levels of apartments using IoT sensors, 2018 IEEE International Conference on Consumer Electronics (ICCE), pp. 1–6 (online), DOI: 10.1109/ICCE.2018.8326169 (2018).

[Sumo 17] SUMO 編集部:きっかけは？重視する条件は？857 人に聞いた引越し･住み替えの実態調査 2017 (https://suumo.jp/article/oyakudachi/oyaku/chintai/frdata/hikkoshi-sumikae2017/ 2018/5/14 参照).

[諏訪 18a] 諏訪博彦，中村優吾，野口真史：IoT センシングによる新 たな賃貸物件探索指標の検討，マルチメディア、分散、協 調とモバイル (DICOMO2018) シンポジウム (2018).

[諏訪 18b] 諏訪博彦，大坪敦，中村優吾，野口真史，新たな賃貸物件探索指標のための IoT センシングデバイスの検討，グループウェアとネットワークサービスワークショップ 2018 (2018).

[Wu 09] Chih-Hung Wu, Chi-Hua Li, I.-C. F. C.-C. H. W.-T. L. C.- H. W.: HYBRID GENETIC-BASED SUPPORT VECTOR REGRESSION WITH FENG SHUI THEORY FOR APPRAISING REAL ESTATE PRICE, 2009 First Asian Conference on Intelligent Information and Database Systems, pp. 295–300 (2009).

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

深層学習を用いた不動産画像の分類システムのビジネス適用

Business utilization of real estate image classification system using deep learning

塚原　朋也[*1]
Tomoya Tsukahara

須藤 広大[*2]
Kodai Sudo

[*1] 株式会社ブレインパッド
BrainPad Inc.

[*2] 株式会社ブレインパッド
BrainPad Inc.

In the real estate industry, they need to reduce the cost of operations by using deep learning. This paper mentions real estate image classification system we built for Daito Trust Co., Ltd.

1.　はじめに

近年、深層学習を活用し、ビジネスに適用するというニーズが多くなってきている。画像認識や画像分類では、Google 社によるサービスなど、機械学習の専門知識がなくてもカスタマイズされた機械学習モデルの構築が可能となっている。このコモディティ化の流れにより、深層学習の活用を益々後押しする要因になっていると考えられる。

2.　物件写真自動掲載システムについて

不動産業界において、深層学習の活用により業務におけるコスト削減を図りたいというニーズがあり、大東建託株式会社向けに不動産画像の分類システム（物件写真自動掲載システムと呼ぶ）を構築し、ビジネス適用に至った。

2.1　背景・目的

大東建託株式会社では、営業スタッフが不動産画像を目視でリビングやキッチン、玄関、洗面所などの分類種別を設定し、1 枚ずつ手作業で基幹システムへ登録していた。この登録作業は物件 1 件あたり 5～10 分程を要し、その登録総数は年間 30 万件近くにのぼるため、その膨大な作業時間の効率化が課題となっていた。

上記の単純作業に要する工数を削減することで、営業スタッフの工数を付加価値の高い業務へ振り分けるため、不動産画像の分類および基幹システムへの登録を自動化するシステムを構築する。

2.2　機能内容

物件写真自動掲載システムでは、次に挙げる機能を備えている。

(1)　不動産画像のアップロード

- 物件ごとのフォルダに関連する複数の不動産画像を纏めて、アップロードする。
- 基幹システムとの API 連携により基幹システムに存在する物件かどうかをチェックする。このチェックは大東建託株式会社が管理している物件どうかを確認するために行うものである。

(2)　不動産画像の分類

- 深層学習活用による不動産画像の分類として、部屋に関連する 21 分類器による分類を行う。クラスの内訳は、外観、その他共有部分、その他部屋・スペース、トイレ、洗面所、収納、バルコニー、庭、セキュリティ、眺望、キッチン、その他設備、エントランス、ロビー、リビング、居室、風呂、その他（画像）、玄関、周辺環境、駐車場である。
- 上記以外は、ルールベースによる分類を行う。

(3)　不動産画像の登録

基幹システムへ不動産画像を登録する。その後、基幹システムを経由して不動産物件サイトに掲載される。

2.3　期待効果

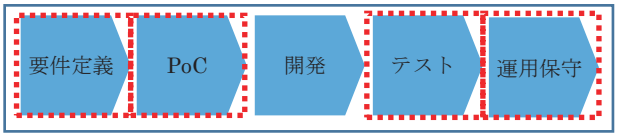
営業スタッフの作業時間が 1 物件あたり約 70%短縮され、1 ヶ月あたりに換算すると約 3,000 時間の作業時間削減に寄与することが見込まれる。

3.　ビジネス適用までの進め方

第 1 段階として、不動産画像のアップロード、分類、登録といった一連の業務フローを満たす WEB アプリケーションを実装し、限定的な範囲でリリースする。限定的な範囲とは、一部の営業店を指す。このリリースにより、全国営業店へのビジネス適用に向けて必要になる要求事項がより一層明確なものとなる。

次の段階として、全国営業店へのビジネス適用に向けた追加機能をリリースする。UI 改善や新たなルールベースによる画像分類の追加が主な対応となる。

上記の 1 つのリリースにつき、次に挙げる 5 つの工程を踏む。



ここでは、要件定義、PoC、テスト、運用保守の部分にフォーカスする。

3.1　要件定義

顧客の要求事項からシステムに実装するべき機能要件や性能要件を定義する。特に、対象の分類種別を精査し、その精度目標について顧客の合意を得ることが重要である。要件定義工程の取り組みは、最も特筆すべき点であるため詳細を後述したい。

3.2　PoC

受領した不動産画像の重複を削除し、モデル学習用のデータセットを構築する。その後、対象の分類種別を定義しモデル構築および検証を実施する。なお、不動産画像の分類種別のラベルは、受領データをそのまま流用する。

連絡先：塚原朋也，株式会社ブレインパッド，03-6721-7842，
tomoya.tsukahara@brainpad.co.jp

対応スピードとコストを両立することは重要な要素であるが、Google 社のサービスである Cloud AutoML が有効である。同程度の精度を発揮することを確認済みである。

3.3 テスト

システム観点に加えて分析モデル観点のテストが必要になる。後者では、要件定義工程にて顧客と合意の得られた対象の分類種別とその精度目標を達成していることを確認する。

検証方法は、予測したクラスと実際の正解クラスとの混同行列を作成して、正解率を評価する。物件写真自動掲載システムを介した場合は、分類結果と不動産画像を目視で確認する。

3.4 運用保守

全国営業店へのビジネス適用を実現するには、システムの安定運用が必須条件となる。季節性が高くデータの増減が著しいものであるためスケールアウトの仕組みやシステムリソース監視、プロセス監視、アラート通知といった非機能要件を満たす。

分類精度の監視や今後必要になる思われる再学習の頻度を見据えて、分類結果とそれらに必要な履歴管理の仕組みやサービス提供後のランニングコストを調整し、顧客の合意を得ることも必要になる。

その他に、モックアップ、α版、β版、正式版といった形で段階的にシステムを成長させながら顧客要望を具体化することが有効であり、リリース自動化の仕組みを準備することで効率的にリリースを実施することができる。

4. 要件定義工程での取り組み

要件定義工程の取り組みについて、特筆すべき点は次の通りである。

4.1 深層学習活用による画像分類に関する要件定義

対象の分類種別を精査し、その精度目標を設定する。不動産画像の分類結果の具体例や精度の良し悪しが発生する原因を顧客に説明し、対応スコープについて顧客の合意を得る。

例えば、精度が良いものの代表例として、トイレ、洗面所、セキュリティ、キッチン、風呂、周辺環境などがあり、人が容易に判断出来そうな分類種別は、90%以上の高精度となる。逆に精度が悪いものの代表例として、その他画像、ロビーなどがあり、人が容易に判断出来ず、分類種別のラベル付け基準がぶれやすい分類種別は、50%を下回る低精度となる。その他に、受領データに誤ったラベル付けがされた不動産画像が含まれているため、ミスリードを引き起こしている可能性があることを顧客に説明する。

4.2 ルールベースによる画像分類に関する要件定義

深層学習活用による画像分類で対応困難なものはルールベールによる画像分類を行う。

例えば、リビング、居室、その他部屋スペースの3つの分類種別に関しては、リビング付き物件の間取り（e.g. 1LDK、2LDK）の不動産画像をリビングに分類し、リビングなし物件の間取り（e.g. 1DK、2DK）の不動産画像を居室に分類するための工夫が必要になる。基幹システムとの API 連携により間取り情報（e.g. 1LDK、2LDK、1DK、2DK）を取得し、その値を判断基準にして居室をリビングへ、またはリビングを居室へ上書きする。その他部屋スペースは、強制的に居室に上書きする。

間取りに関しては 100%の分類精度が要求される。物件写真自動掲載システムでは、間取りは必ずファイルの拡張子が gif となるため、これを判断基準とする。他に、ファイル名の prefix（e.g. madori-*.jpeg）で判断する方法も考えられる。

建物外観と部屋外観に関しては、建物と部屋それぞれに外観がある。部屋に関連する21分類器を活用して、外観の分類スコアが最も高い不動産画像を建物外観、次に分類スコアの高い不動産画像を部屋外観とする。

4.3 機能仕様に関する要件定義

不動産画像の分類は、全分類種別に対して精度 100%を保つことは現実的でないため、営業スタッフが手作業で分類種別の変更を行い、リカバリー可能な UI を実装する。

不動産画像の基幹システムへの登録では、不動産画像の登録順番（e.g. 1 枚目、2 枚目、3 枚目）、登録種別の優先順位（e.g. 1 枚目はリビングと居室のいずれか、2 枚目はキッチン、3 枚目は風呂）といった業務仕様を満たす。パズルゲームのように不動産画像のドラッグ&ドロップによる登録順番入れ替えといったモダンな UI を導入し、システムの利便性向上を図りつつ作業時間の効率化に繋げる。

5. 考察

実運用では、深層学習活用による画像分類とルールベースによる画像分類の両者をバランス良く組み合わせて、最大限の分類精度を発揮することが重要になる。また、業務効率化においては、モダンな UI を導入することは依然として欠かせないものである。

5.1 今後の課題

これまで述べてきたもののうち、今後、対応するべき課題を次に挙げる。

(1) 不動産画像の分類種別のラベル付け基準のぶれ

- 分析結果は不動産画像の分類種別のラベル付けの精度の依存度が大きく、50%を下回る低精度となる場合がある。
- 今後は、明らかにミスリードを起こしそうなデータを修正し、誤った分類種別が付与されたデータを学習データから除外するように対応したい。

(2) 人手を介する分析モデルのテスト

- 分析モデルのテストにて目視で確認する部分が残る。
- 今後は、数理統計的な判断基準を取り入れることによる分析モデルのテストの自動化を検討したい。

(3) 分析モデルの運用保守

- システムの長期運用に際して、モデル学習時点と物件写真自動掲載システムに実際にアップロードされる時点の不動産画像に乖離が発生する可能性がある。
- 分析モデルの精度低下を検知する仕組みや再学習を実施した場合のデータセットと分析モデルを管理できるように対応したい。

6. 謝辞

物件写真自動掲載システムは、大東建託株式会社の協力による。

参考文献

[BrainPad.inc 2018] BrainPad Inc.: ブレインパッド、大東建託へ AI を活用した賃貸物件の画像分類システムを構築・提供, http://www.brainpad.co.jp/news/2018/06/11/7695.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

居住候補地発見支援のための鉄道経路探索システムの提案

Proposal of Railway Route Search System for Finding Living Place

菊池　圭祐 Keisuke Kikuchi*1　小林 賢一郎 Kenichiro Kobayashi*2　橋本 武彦 Takehiko Hashimoto*2　高間　康史 Yasufumi Takama*1

*1首都大学東京大学院システムデザイン研究科 Graduate School of System Design, Tokyo Metropolitan University　*2株式会社 GA technologies GA Technologies

This paper proposes a route search system for supporting users to find their living place based on railway information. When we want to find a living place, the access to the place of work or school is usually considered. However, usual railway search systems are designed to search a route from departure station to destination at a specified time. Given the destination station and conditions such as commuting time and the number of connection, the proposed method finds multiple stations satisfying the conditions. This paper describes the search algorithm and explains the feedback about a prototype system from the salespersons in a real-estate company.

1.　はじめに

本稿では, ユーザが居住地を探す作業の支援を目的として, 目的地志向の路線探索手法を提案する. 一般に, 居住地を選択する際には, 最寄り駅から勤務地や学校などの目的地までのアクセスの良さを考慮することが多い. しかし, 現在, 広く普及している路線検索システムは, 出発地・目的地間の経路を求めることを目的としているため, 居住候補地を探す際のような, 指定した目的地に到達可能な候補地を複数探索する用途には適していないという問題点がある.

この問題点に対して, 提案手法では, ユーザの目的地から, 所要時間および乗換回数などの条件を満たす全ての駅および使用路線を検索可能とする. また, 提案手法を用いた検索システムのプロトタイプを実装し, 不動産会社の営業担当者とのヒアリングにより, 路線検索システムの有効性を検証する.

2.　関連研究

一般的な路線探索は, 基本的に出発駅から目的駅への最短経路問題として捉えることができる. そのため, 探索はダイクストラ法 [Dijkstra 59] などのコストを考慮した探索アルゴリズムに基づいて行われる. ダイクストラ法とは, グラフネットワークにおいて, 単一始点最短経路問題を解くための最良優先探索のアルゴリズムである. ダイクストラ法を用いた路線探索に関する研究として, 運賃設定の異なる複数の鉄道会社を含む鉄道ネットワーク上での運賃計算アルゴリズムの報告 [池井 08] などがある. 現状, 一般に普及している乗換案内サービスで用いられている具体的な探索アルゴリズムやノウハウは, 公開されていない部分が多いとされているが, 研究では, 所要時間, 乗換回数, 料金などの複数の実用的な制約条件から, 優良経路を高速検索する手法などが提案されている [半田 05].

3.　提案手法

3.1　提案システム

路線探索は, 東京都都内の 654 駅, 80 路線のデータを対象として行う. 提案する路線探索システムでは, ユーザの目的地の最寄り駅, 移動時間上限および路線の乗換回数上限の設定を入力として, 条件を満たして目的地に到着可能な全ての駅を居住候補地として求め, 経路情報とともにユーザに提示する. また, 提案手法は, 乗換案内サービスのように特定日時での移動経路を調べることを目的としたものではないため, 出発時刻・到着時刻の時間指定は行わなず, 朝の通勤時における一般的な所要時間に基づき探索する. 具体的には, 朝の通勤通学時間を想定して, 平日 AM8:00 から AM9:00 の間に運行している路線ダイヤの一般的な移動時間を設定する.

3.2　路線ネットワーク

駅間の接続情報および路線情報をもとに, 図 1 のような路線ネットワークを構築する. 路線ネットワークでは, ノードを駅, エッジを移動可能な路線, エッジコストを駅間移動の所要時間とする. また, 隣接駅ノード間で複数路線が運行する場合には, 基本的に最短の移動時間となる路線とコストを設定する. また, 同一路線区間のエッジには, 路線の駅順に沿った向きの情報を付与する. これは, 以下の探索アルゴリズムにおいて, 不正乗車となる経路を判別することを目的として設定している. また, 同一路線のみで, 乗換なしで移動できる駅間には, その区間の所要時間の合計値をコストとして, 直接エッジを接続する. 図 1 に, 路線ネットワークの例を示す. 図 1 では, A 駅から C 駅までが同一路線で移動できる駅となっている. この場合, 隣接する駅間のエッジに加えて, A 駅から C 駅間を所要時間の合計値である 7 分をコストとして, エッジを接続する. これによって, 路線ネットワーク上での移動に関して, エッジ数と乗換回数が一致する. このネットワーク構造より, 以降に説明する探索アルゴリズムにおいて, 所要時間と乗換回数を識別することが可能になる.

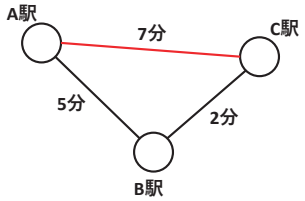


図 1: 路線ネットワーク例

3.3　探索アルゴリズム

3.2 節の路線ネットワークを利用した探索アルゴリズムを説明する. 探索は始点となる駅ノードを指定し, 所要時間の上限エッジコストおよび乗換回数の上限回数を設定して, 条件を満たす

連絡先: 高間康史, 首都大学東京大学院システムデザイン研究科, 〒191-0065 東京都日野市旭が丘 6-6, ytakama@tmu.ac.jp

到達可能な駅ノードへの最短経路探索となる．最短経路探索はダイクストラ法を用いて行う．なお，ダイクストラ法による探索は Python のグラフ分析パッケージ networkX のライブラリを用いて，始点ノードから各ノードまでの最短経路および最短距離を求める．以下に，アルゴリズムの擬似コードを示す．PahtsA, PathsB, ResultPaths は経路を要素とするリストであり，重複した要素は持たない．PathsA, PathsB は事前に計算しておく．また，アルゴリズム上での枝切りおよび再探索の処理内容の例を図 2 に示す．図 2 では，A 駅から D 駅間に B, C 駅を経由する各駅停車が運行しているほか，B, C 駅に停車しない快速があるような状況を想定している．この時，B 駅から A 駅に戻って快速を利用する経路は，A-B 間と A-D 間の向きが逆となるため候補外となり，再探索となる．A-D 間の経路を枝切りし，再探索された B-D 間の経路コストは 12 分から 14 分となる．

Algorithm 1 探索アルゴリズム

```
G ← グラフデータ読み込み
PathsA ← 始点から各ノードまでの所要時間最短経路
PathsB ← 始点から各ノードまでのエッジ数最短経路
ResultPaths = []
while PathsA do
    if エッジ数 ≤ 乗換回数上限 then
        ResultPaths に経路を追加
    end if
end while
while PathsB do
    if 経路の総コスト ≤ 所要時間上限 then
        ResultPaths に経路を追加
    end if
end while
while ResultPaths do
    if 経路に逆向き移動のエッジが含まれる then
        対象エッジをグラフ G から枝切り
        再探索 ⇒ 再探索結果の総コスト・総エッジ数を取得
        if (再探索の総コスト > 所要時間上限) ∨
            (再探索の総エッジ数 > 乗換回数上限) then
                経路を ResultPaths から削除
        end if
    end if
end while
```

図 2: 枝切りと再検索の例 (B 駅→ D 駅の経路)

4. 実装結果

3 節で述べた路線探索手法を用いて，実装した検索システムのインタフェース画面を図 1，検索を行った例を表 2 に示す．検索例では，東京駅から乗換回数 2 回以下で 40 分以内の条件を入力とした．また，検索結果は一部を抜粋して表記している．

不動産企業の営業担当者とのヒアリングでは，路線探索の提案システムを実際に使用してもらい，実際の営業現場での観点から利用可能性を評価した．

営業担当者からは，現状の現場で使用している検索手法よりも提案システムの方が効率的に条件を満たす駅を探すことが可能だと評価された．また，営業担当者だけでなく，顧客側の利用を視野に入れた場合，居住候補地として選択可能な駅および地理的な範囲を，地図上で可視化できた方が直感的な理解につながるという意見を頂いた．また，提案システムでは，目的駅の設定は 1 駅となっているが，複数駅を入力として，条件を満たす駅を AND 検索できた方が，一人暮らしの個人ユーザだけではなく，夫婦などのユーザでも効率的な検索を行うことができる，という意見も頂いた．

図 3: 提案システムのインタフェース

表 1: 提案手法での検索例

到達駅	所要時間	乗換回数	使用路線
有楽町駅	2 分	0 回	JR 山手線
神田駅	2 分	0 回	JR 中央線
⋮	⋮	⋮	⋮
末広町駅	10 分	1 回	JR 中央線, 東京メトロ銀座線
⋮	⋮	⋮	⋮
泉岳寺駅	20 分	1 回	JR 京浜東北線, 都営浅草線
⋮	⋮	⋮	⋮
沼袋駅	40 分	2 回	東京メトロ丸ノ内線, 東京メトロ東西線, 西武新宿線

5. おわりに

本稿では，ユーザが居住地を探す作業の支援を目的として，目的地志向の路線探索手法，統計データに基づく類似地域検索手法を提案した．路線探索手法について検証を行った結果，不動産企業の営業現場にて，顧客の設定条件に基づき居住候補地となるを探す際に有効であることを確認した．居住候補地の選択には，その街の印象や家賃相場，住みやすさなどの情報も重要であるため，今後は地域の印象や特性も考慮した検索機能を検討する予定である．

参考文献

[Dijkstra 59] E.W. Dijkstra, “A note on two problems in connection with graphs,” Numerische Mathematik, Vol. 1, No. 1, pp. 269-271, 1959.

[池井 08] 池上敦子，森田隼史，山口拓真，菊地丞，中山利宏，大倉元宏，鉄道運賃計算のための最安運賃経路探索:複数の鉄道会社を含む場合，日本オペレーション・リサーチ学会和文論文誌，Vol. 51, pp. 1-24, 2008.

[半田 05] 半田恵一，田中俊明，乗換え案内サービスにおける経路探索手法，電子情報通信学会論文誌，Vol. J88-D-I , No. 10, pp. 1525-1533, 2005.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

間取り図を用いた賃料予測モデルに関する一検討

Evaluation of Rent Prediction Models using Floor Plan Images

*1服部 凌典 Ryosuke Hattori　*2岡本 一志 Kazushi Okamoto　*3柴田 淳司 Atsushi Shibata

*1電気通信大学 情報理工学部
Faculty of Informatics and Engineering, The University of Electro-Communications

*2電気通信大学 大学院情報理工学研究科
Graduate School of Informatics and Engineering, The University of Electro-Communications

*3産業技術大学院大学 産業技術研究科
Graduate School of Industrial Technology, Advanced Institute of Industrial Technology

This study constructs rent prediction models with/without floor plan images in order to validate whether such images contribute the prediction accuracy. In addition, applications of PCA (principal component analysis) and convolutional neural network are considered as a feature extractor from floor plan images. The prediction accuracy is measured using properties of 90,000 rental housings in Tokyo. In the experimental results, the root mean squared error values of the prediction model with floor plan images and PCA tend to be higher than without floor plan images. This suggests that the use of floor plan images contributes to accuracy of rent prediction.

1. はじめに

賃貸物件において，属性が全て同じ物件は数少なく，類似した属性を持つ物件でも築年数や階数，間取りなどのどこかに違いがある．特に，同一の間取りの規格（1K や 2LDK など）であったとしても，各部屋の配置や面積などは物件毎に異なる．また，日本では賃貸物件を探す際に間取り図を見る習慣があり [清田 2017]，多くの物件では間取り図が用意されている．

物件属性は賃料に影響を与えると考えられており [河合 2008, 石島 2010, 福井 2018]，賃料決定の際に利用可能な手法としてヘドニック・アプローチがある．間取り図の記載内容も物件属性のひとつと見なすこともできるが，これまでのヘドニック・アプローチによる賃料予測の事例では考慮されてきていない．

本研究では，間取り図を用いた賃料の線形回帰モデル構築し，間取り図を用いない線形回帰モデルと予測精度を比較することで，間取り図の活用が賃料予測に与える影響を明らかにする．前者の線形回帰モデルでは，間取り図から抽出した特徴量を利用するが，特徴抽出器として，主成分分析（PCA: principal component analysis）と畳み込みニューラルネットワーク（CNN: convolutional neural network）の適用を検討する．構築したモデルの予測精度は，LIFULL HOME'S データセットに含まれる東京都の賃貸物件 9 万件を用い，モデルの出力と真の賃料との平均二乗誤差を求めることで検証する．

2. 構築する賃料予測モデル

ヘドニック・アプローチとは，ある商品の価格をその商品の属性の価値に関する集合とみなし，商品価格を線形回帰モデルで予測する手法である [河合 2008]．本研究では，入力ベクトル $\boldsymbol{x} \in \mathbb{R}^d$ から予測賃料 $\hat{y}$ を

$$\hat{y} = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x} + \beta \tag{1}$$

で求めるモデルを扱う．ここで，$\boldsymbol{\alpha} \in \mathbb{R}^d$ を偏回帰係数，β を定数項とし，データから予測誤差を最小化するよう推定する．

本研究では，間取り図を用いる場合と用いない場合の予測モデルを構築する．物件属性（築年数や階数，専有面積など）を $\boldsymbol{u} \in \mathbb{R}^m$，間取り画像を $\boldsymbol{v} \in \mathbb{R}^n$，関数 f_θ をパラメータ θ を持つ画像特徴抽出器としたとき，間取り図を用いる場合は $\boldsymbol{x} = [\boldsymbol{u}, f_\theta(\boldsymbol{v})]$ を，用いない場合は $\boldsymbol{x} = \boldsymbol{u}$ として扱うことになる．　また，特徴抽出器 f_θ には PCA と CNN を適用する．

2.1 PCA による間取り図からの特徴抽出

特徴抽出器 f_θ に PCA を用いる場合，学習データに含まれる間取り図の集合から変換行列 $\theta \in \mathbb{R}^{k\times n}$ を推定し，$f_\theta(\boldsymbol{v}) = \theta\boldsymbol{v}$ とすることで間取り図の特徴ベクトルを得ることができる．変換行列 θ の推定には教師信号（賃料）は必要ないため，$\boldsymbol{x} = [\boldsymbol{u}, f_\theta(\boldsymbol{v})]$ の算出後に，式 (1) の $\boldsymbol{\alpha}, \beta$ を推定することになる．また，$f_\theta(\boldsymbol{v})$ の結果を逆変換することで，次元圧縮後の間取り図を近似的に復元でき，次元圧縮の傾向を画像として確認することができる．

2.2 CNN による間取り図からの特徴抽出

CNN による間取り図用の特徴抽出器 f_θ では VGG16 を利用する．VGG16 は，ImageNet データセットを用いて学習した 13 層の畳み込み層と 3 層の全結合層を持つニューラルネットワークである．VGG16 の出力は 1,000 クラスの確率分布であり，層数が増加することにより計算コストも高くなってしまうため，本研究では，VGG16 の後ろ 5 層を削除し新たに k 次元の全結合層を追加する．なお，VGG16 の結合係数は固定とし，追加した k 次元の全結合層への結合行列 θ のみを学習することとする．

この特徴抽出器を用いる場合，結合行列 θ の推定には教師信号が必要となるため，本研究では，間取り図 $\boldsymbol{v}$ を VGG16 に入力し得られた特徴ベクトル $f_\theta(\boldsymbol{v})$ と物件属性ベクトル $\boldsymbol{u}$ の連結層と賃料の出力層（1 次元）を持つニューラルネットワークを構築する．そのため，PCA を用いる場合と異なり，賃料を用いてパラメータ $\boldsymbol{\alpha}, \beta, \theta$ を同時に推定するモデルとなる．

連絡先: 岡本 一志，電気通信大学 大学院情報理工学研究科 情報学専攻，東京都調布市調布ヶ丘 1-5-1，Tel.: 042-443-5280，E-mail: kazushi@uec.ac.jp

3. 賃料予測モデルの精度評価実験

3.1 使用データセット

本研究では，国立情報学研究所の情報学研究データリポジトリで公開されている LIFULL HOME'S データセットを使用し，構築した賃料予測モデルの予測精度を検証する．このデータセットは，2015 年 9 月時点で不動産住宅情報サイト LIFULL HOME'S に掲載されていた全国 533 万件の物件データで構成されており，賃料，面積，立地，築年数，間取り，建物構造，諸設備などの物件属性や，間取り図（120 × 120 ピクセルの）や外観写真，内観写真などの 8,300 万の画像ファイルが含まれている．このうち，東京都の 9 万件の物件について，賃料と共益費を足したものを目的変数とし，市区郡町村・徒歩距離・建物構造・建物面積/専有面積・建物階数（地上）・建物階数（地下）・築年月・新築・未入居フラグ・部屋階数・契約期間（年）・駐車場料金・引渡/入居時期の 12 変数を物件属性として利用する．なお，賃料が欠損している物件は除外する．

本研究では，9 万件の物件を K / R / DK/ LD / LDK の 5 種類の間取り規格で層別し，間取り規格毎に賃料予測モデルを構築する．ただし，間取りが LD の物件は 35 件のみであるため，間取りが K / R / DK / LDK の物件を使用する．

3.2 実験条件

間取り図を予測モデルに用いる場合，特徴抽出器のハイパーパラメータとして，次元数 k を設定する必要がある．このため，パラメータ調整用に検証用データを，構築した予測モデルの評価用にテストデータをそれぞれ用意する．具体的には，各間取り規格において，物件を 4:1 の割合で検証用データとテストデータに分割し，パラメータ調整は検証用データに 10 分割交差検証法を適用することで行う．テストデータに対する予測では，検証用データ全てを学習に利用する．なお，本研究においては，CNN の全結合層の次元数 k の調整は計算時間の関係から行えておらず，$k = 64$ に固定している．PCA については，$\{64, 128, 256, 512, 1024, 2048\}$ の次元数の中から適切な次元数を探索している．また，モデルの予測精度を評価指標には平均二乗誤差（RMSE: root mean squared error）を用いる．

4. 予測精度の比較

PCA を用いた特徴抽出器において，検証用データにおける最良の平均 RMSE をとるのは，間取り規格が K の物件で 10548.18（$k = 512$），R の物件で 8789.96（$k = 2048$），DK の物件で 11069.98（$k = 1024$），LDK の物件で 19668.87（$k = 2048$）であることを確認している．そのため，PCA の次元数 k にはこれらの値を利用する．

表 1 に間取り規格が LDK の物件における RMSE を示す．表 1 のテストデータの欄では，間取り図に PCA を適用した予測モデルの RMSE が最も高く，間取り図に CNN を適用した予測モデルの RMSE が最も低くなっている．同様の傾向は，間取り規格が R / DK の物件についても確認している．さらに，全ての間取り規格の物件において，PCA を適用する場合が CNN を適用する場合に比べて RMSE が高くなっている．このことは，賃料予測に間取り図を使うことよりも，今回のニューラルネットワークの構造に課題があるためと考える．また，間取り規格が LDK の物件について，PCA を適用した間取り図の復元例を図 1 に示す．図 1 より，PCA による間取り図の次元圧縮自体は適切に行えていることが確認できる．

PCA を用いた予測モデルと間取り図を用いない予測モデルのテストデータにおける RMSE の差は，間取り規格が K の物件で-86，R の物件で 490，DK の物件で 199，LDK の物件で 495 となっている．間取り図の利用が賃料の予測精度の向上に寄与していると考えるが，その効果の分析はまだ詳細には行えておらず，今後明らかにしていく予定である．

表 1: 間取り規格が LDK の物件に対する RMSE 値（検証用データについては交差検証法適用時の平均と 95% 信頼区間）

	検証用データ	テストデータ
w/ 間取り図（PCA, k=2,048）	19668.87 [19265.94, 20071.80]	19646.72
w/ 間取り図（CNN, k=64）	23590.04 [23230.41, 23949.65]	22878.80
w/o 間取り図	20227.00 [19927.19, 20526.87]	20141.84

(a) 入力画像　(b) 復元画像

図 1: PCA（$k = 2048$）により次元圧縮した間取り図の復元例

5. おわりに

本研究では，間取り図を用いた賃料の線形回帰モデルを構築し，間取り図を用いない線形回帰モデルとの予測精度の比較を行っている．東京都の賃貸物件 9 万件を利用した賃料予測実験より，間取り図に PCA を適用した線形回帰モデルが間取り図を利用しない線形回帰モデルよりも予測精度（RMSE）が高い傾向にあり，間取り図の利用が賃料の予測精度の向上に寄与していると考える．加えて，間取り図からの特徴抽出には PCA が VGG16 ベースのニューラルネットワークよりも RMSE が高くなることも確認している．

参考文献

[清田 2017] 清田 陽司，山崎 俊彦，諏訪 博彦，清水 千弘：不動産と AI，人工知能，Vol.32，No.4，pp.529–535（2017）

[河合 2008] 河合 伸治：ヘドニック・アプローチによる賃貸住宅価格の価格決定要因の推定 –西武池袋線の賃貸住宅を事例として–，ソシオ サイエンス，Vol.14，pp.49–63（2008）

[石島 2010] 石島 博，谷山 智彦：個別性と歪みを考慮した住宅価格分析とパーソナルファイナンスへの応用，ファイナンシャル・プランニング研究，Vol.10, pp.4–17（2010）

[福井 2018] 福井 光，阪井 一仁，南村 忠敬，三尾 順一，木下 明弘，高田 司郎：レインズのニューラルネットワークを用いた不動産価格査定について，人工知能学会全国大会論文集，Vol.JSAI2018, No.4A2-03，pp.1–4（2018）

謝辞

本研究では，国立情報学研究所の IDR データセット提供サービスにより株式会社 LIFULL から提供を受けた「LIFULL HOME'S データセット」を利用した．

Organized Session | Organized Session | [OS] OS-10

[1D3-OS-10b] 不動産と AI(2)

清田 陽司（（株）LIFULL）、山崎 俊彦（東京大学）、諏訪 博彦（奈良先端科学技術大学院大学）、清水 千弘（日本大学）、橋本 武彦（GA technologies）

Tue. Jun 4, 2019 3:20 PM - 5:00 PM Room D (301B Medium meeting room)

[1D3-OS-10b-01] (Invited talk) Recent studies on real estate tech in China

〇Xiangyu Guo[1] （1. Fudan University）

3:20 PM - 4:00 PM

[1D3-OS-10b-02] Two-Step estimation model using machine learning at detached house price

〇Yusuke Takahashi[1] （1. FUJITSU CLOUD TECHNOLOGIES LIMITED）

4:00 PM - 4:20 PM

[1D3-OS-10b-03] Statistical modeling of the transition time of an occupation of rental rooms by using the housing information website data

〇Hayafumi Watanabe[1], Yu Ichifuji[2], Masahito Suzuki[3], Satoshi Yamashita[4] （1. Kanazawa university, 2. Nagasaki university, 3. UD Asset Valuation Co., Ltd., 4. The institute of statistical mathematics）

4:20 PM - 4:40 PM

[1D3-OS-10b-04] Efforts on UX innovation in real estate field by utilizing AI

Yoji Kiyota[1], 〇Satoshi Shiibashi[1], Takeshi Ninomiya[1], Takao Yokoyama[1], Takuro Hanawa[1], Takeshi Eto[1], Akiko Yokoyama[1], Kei Kikuchi[1], Musashi Kobayashi[1], Akane Kameda[1], Kazuki Takigawa[1], Yusuke Saito[1], Kazushi Katayama[1] （1. LIFULL Co., Ltd.）

4:40 PM - 5:00 PM

©The Japanese Society for Artificial Intelligence

3:20 PM - 4:00 PM (Tue. Jun 4, 2019 3:20 PM - 5:00 PM Room D)

[1D3-OS-10b-01] (Invited talk) Recent studies on real estate tech in China

〇Xiangyu Guo[1] （1. Fudan University）

Keywords: real estate tech, real estate finance, urban economics

With efforts to improve productivity through introduction of technologies in the real estate industry globally, progress in sharing of information and data by technologies is thriving in the Chinese market as well. This presentation shows the latest situations of the Real Estate Tech in China centering on real estate property price estimation.

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

戸建住宅価格における機械学習を用いた2段階推計モデル

Two-Step estimation model using machine learning at detached house price

高橋 佑典 *1
Yusuke Takahashi

*1富士通クラウドテクノロジーズ株式会社
FUJITSU CLOUD TECHNOLOGIES LIMITED

This paper reports the result of trying to estimate the detached house price with a two step model. By handling the advantages of linear model and nonlinear model in combination, we can expect explanation possibility and estimation accuracy for model when estimating house price. Experiments compare the methods by machine learning and the effects of explanatory variables to be input to the nonlinear model.

1. はじめに

住宅価格の分析を行った従来の研究では，多くが線形モデルとして推定している．その背景として，シンプルなモデルであればモデルの推計価格に対して明確な説明が可能であるという点が大きい．しかし，専有面積や築年数などでは，住宅価格に与える影響が区間によって異なり，非線形になると考えられ，従来の研究でもその存在が示されている．また，線形モデルとして推定する場合，住宅価格を決定する上で主要な要因である築年数と建築年代といった多重共線性を持つ説明変数を同時に投入してしまうと，モデルの推計値が不安定になってしまうことも挙げられる．

そこで，先行研究のうち [Shimizu 14] では，非線形モデルとして推計するため，ノンパラメトリックなモデルである連続量 dummy モデル（DmM）と，AIC を評価指標とした Switching RegressionModel（SWR），一般加法モデル（GAM）が用いられ，非線形性を考慮した推計が行われている．非線形モデルを扱う課題として，モデルが複雑になり再現性が低下することが考えられる．

本研究では，線形モデルと非線形モデルの課題に対処するため，モデルを 2 段階に分けて戸建住宅における取引価格の推計を行う．1 段階目のモデルでは線形モデルとして推計し，モデルに対する説明性を担保する．2 段階目のモデルでは 1 段階目の誤差項を目的変数として，非線形なモデルによって推計を行う．本研究の最終的な目的は，線形モデルと非線形モデルを組み合わせた二段階のモデルによって推計することで，住宅価格の推計時に，説明性と推計精度を向上させることである．

実験では，最初に線形モデルと機械学習による非線形モデルを作成し，誤差率の分布を比較することで，非線形性をうまく表現できる手法の選択を行う．次に，研究背景で述べた非線形性の考慮を機械学習を用いて行い，先行研究で扱われている一般加法モデル（GAM）の結果と比較する．評価指標としては機械学習で用いられている指標である平均平方二乗誤差（RMSE），平均平方二乗誤差率（RMSPE）を用いる．

連絡先: 高橋 佑典，富士通クラウドテクノロジーズ株式会社，東京都中央区銀座 7 丁目 16 番 12 号　G-7 ビルディング，03-6281-5710，yus-takahashi@fujitsu.com

2. 関連研究

2.1 非線形性

[Shimizu 14] では東京 23 区の中古マンションを対象に，線形モデルをベースに置き，物件の平米単価と各説明変数間の関係を非線形モデルとして推計している．具体的には，ノンパラメトリックなモデルである連続量 dummy モデル（DmM）と，AIC を評価指標とした Switching Regression Model（SWR），一般加法モデル（GAM）を用いて中古マンションの取引価格の主要要因が持つ非線形性を明らかにしている．

[小野 02] ではリクルート社の「週刊住宅情報」に掲載された東京 23 区の中古マンションデータを扱い，中古マンション価格に影響を与える要因のうち，「建築後年数」と「最寄り駅までの徒歩時間」に非線形性が存在することを明らかにしている．

2.2 機械学習

[福井 18] では中古マンションの取引データをもつ，レインズの成約データに対してニューラルネットワークを用いて，不動産査定価格を分類問題として解いている．

[大和 18] では不動産情報サイト SUUMO におけるデータを扱い，家賃の推計を行う際に変数の非線形性を表現できる手法として，線形回帰モデルの代わりに決定木をベースとしたランダムフォレストを用いている．

本研究では，関連研究で明らかにされている，住宅価格に影響を与える要因に非線形性が存在する点と，それらの非線形性を機械学習によって表現し，推計価格の精度向上が図れるかを調査する．また，先行研究では主に中古マンションの取引価格データが扱われていることに対し，本研究では戸建住宅の中古取引価格データを扱う．

3. 実験

3.1 データセット

本研究においては，国土交通省が提供している土地総合情報システムのうち，不動産取引価格情報ダウンロードサービス（[土地総合情報システム](http://www.land.mlit.go.jp/webland/download.html)）から取得したデータを用いた．取得したデータは，取引時期が 2005 年第 3 四半期から 2018 年第 3 四半期である東京 23 区内の取引価格情報である．本研究では，戸建住宅の中古取引価格データを扱いたいため，データの種類が「宅地（土地

と建物)」，用途が「住宅」，地域が「住宅地」となっているレコードのみを抽出した．この不動産取引情報には，不動産取引において一般的な取引時期，延床面積，土地面積，最寄り駅までの徒歩分数といった情報が取引価格とともに含まれる．表 1 に本研究で扱うデータの要約統計量を示す．各データはそれぞれ TS:最寄り駅までの徒歩分数，P:成約価格，L:土地面積，S:延床面積，RW:前面道路幅員，BLR:建ぺい率，FRA:容積率，CY:建築年，A:建築後年数，WOOD:木造ダミーである．また，要約統計量で示したデータ以外にも用途地域や市区町村コードといったデータも含まれている．

表 1: データの要約統計量

	count	mean	std	min	25%	50%	75%	max
TS	14424	10.65654	5.411991	1	7	10	14	29
P	14424	48846850	17609200	21000000	35000000	46000000	60000000	99000000
L	14424	93.0633	31.69272	50	70	90	110	200
S	14424	97.72463	24.75509	50	85	95	105	200
RW	14424	4.693795	1.288334	2.5	4	4	5.5	9
BLR	14424	56.56129	7.888145	30	50	60	60	80
FRA	14424	174.0294	76.58877	60	100	150	200	800
CY	14424	1996.819	13.50506	1955	1988	2000	2007	2017
TDY	14424	2011.655	3.736598	2005	2008	2012	2015	2018
A	14424	14.83555	13.07827	1	2	12	23	50
WOOD	14424	0.8840821	0.3201376	0	1	1	1	1

3.2 二段階推計モデル

本研究では，モデルを二段階に分けて推計している．理由として，1 段階目に線形モデルを推計して説明性を担保しつつ，2 段階目のモデルでは多重共線性を持ってしまう要因や非線形性を持つと思われる要因を投入し，1 段階目の誤差を予測し補正することで，モデル全体としての予測精度の向上を図る．

step1.

$$log(y) = \alpha + \beta_1 a_1 + \beta_2 a_2 + \beta_3 a_3 + \cdots + \mu_1$$

step2.

$$\mu_1 = \gamma + \delta_1 b_1 + \delta_2 b_2 + \delta_3 b_3 + \cdots + \mu_2$$

線形モデルを二段階に分けて推計する場合，上記のような 2 つの線形モデルが作られる．本研究では，1 段階目では線形モデルを推計し取引価格を予測する．2 段階目では 1 段階目の予測誤差を目的変数として，非線形なモデルを推計し，取引価格に対する最終的な予測を行う．

本研究では，戸建住宅の取引データに対して一般加法モデル（GAM）を用いて住宅価格に影響を与える主要な要因の非線形性を確認した．一般加法モデル（GAM）を適用した結果を図 1 に示す．

一般加法モデル（GAM）を適用した結果から，極端な傾きの変化は見られなかったものの，S:延床面積，L:土地面積，A:建築後年数，といった主要な要因に対して，非線形性を確認することができた．

3.3 手法の比較

本研究では，線形モデルである OLS をベースラインとし，比較対象として，機械学習の手法である決定木をもとにしたランダムフォレストと xgboost を選択した．また，ニューラルネットワークによる手法として多層パーセプトロンを選択した．以下で各手法についての概要を説明する．

OLS は定数項（切片）と説明変数の係数によって値を予測する線形モデルであり，最小二乗法によって係数と切片を決定する．ランダムフォレストは，決定木とバギングを組み合わせた手法であり，決定木を大量に生成し、各決定木の結果を集計

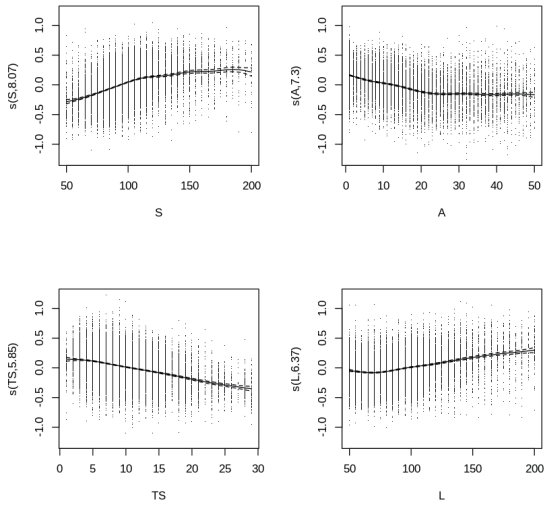


図 1: 各変数と取引価格の関係:GAM

して予測を行う．各決定木は独立して異なる特性を持つように学習する．xgboost は，決定木とブースティングを組み合わせた手法であり，決定木を逐次的に増やしていき、生成済みの決定木の誤差を補正するように、新たな決定木を生成し学習を進めていく．多層パーセプトロンは，入力、中間、出力の 3 層からなるニューラルネットワークの手法。バックプロパゲーションを用いた学習を行う．

機械学習においては，学習に使ったデータセットだけに過度に適合したパラメータが学習されてしまい，テストデータに対して性能が出ない過学習と呼ばれる状態に陥ることがある．過学習を抑えるために，決定木による手法では特徴量の数や生成する木の深さを設定することが考えられる．また NN による手法では，学習の過程において大きな重みを持つことに対してペナルティを課す Weight decay と呼ばれる正則化の手法が存在する．

本研究における機械学習手法では，グリッドサーチによるパラメータチューニングを行い，過学習を抑制するためにパラメータを定めてからモデルの構築を行った．

3.4 評価

まず，手法間の誤差率分布を比較した結果を図 2 に示す．データは学習データとテストデータを 8:2 でランダムに分割し，各手法において同じ説明変数を投入したモデルで一つ一つの物件に対して取引価格の予測を行い，予測価格/実際の取引価格で定義される誤差率を算出した．その実験を 200 回繰り返し行い，誤差率の平均値の分布を作成した．

結果から，OLS や決定木をベースにした手法は予測結果が大きくはずれてしまう可能性が少ないことに対して，ニューラルネットワークをベースにした多層パーセプトロンでは誤差率の平均値の分布に外れ値が見られた．このことから多層パーセプトロンでは推計結果に大外れが生じてしまうことが確認できた．

3.5 2 段階目投入変数

第 3.4 節の結果から，2 段階目の手法として大外れが少なく，過学習も抑制できていると考えられた xgboost を選択した．2 段階目の機械学習モデルに投入する変数として，他の説明変数への影響が明確でない STATION:最寄り駅ダミーや

CENSUS:大字・通称レベルの地域ダミーを投入してその効果を確認した．また，S:延床面積，L:土地面積，A:建築後年数，TS:最寄り駅までの徒歩分数を説明変数として単独投入し，その効果を確認した．結果を箱ひげ図として図 3 と図 4. に示す．また，各評価指標の値としては表 4. と表 3 に示す．

これらの結果から評価指標を見ると，機械学習における非線形性の考慮が，僅かだが精度向上に寄与していることが確認できた．

4. おわりに

本研究では，戸建住宅データにおける取引価格に影響を与える主要要因に対して，機械学習での非線形性を考慮した 2 段階推計を行った．実験結果より，手法間の比較では，ニューラルネットワークをベースにした手法よりも，決定木をベースにした手法の方が，誤差率のばらつきが少ないことがわかった．また，2 段階推計においては，2 段階目のモデルに個別に説明変数を投入し，その補正結果を調べた．結果は一般加法モデル（GAM）で確認した非線形性を考慮し，精度向上が確認できた．また，合わせて戸建住宅価格に影響を与える最寄り駅ダミーと大字・通称ダミーを二段階目の説明変数として投入し，精度向上を確認できた．今後は，今回得られた手法ごとの誤差率分布をもとに，非線形なモデルを推計する手法ごとの特性を明らかにすることや，モデルに投入する説明変数の数を増やしていき，その効果を検証することが考えられる．

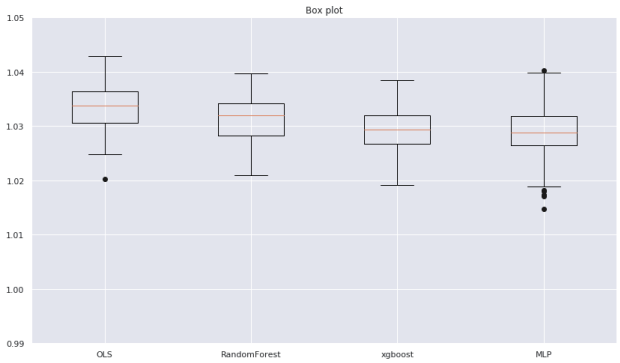


図 2: 誤差率の平均値の分布

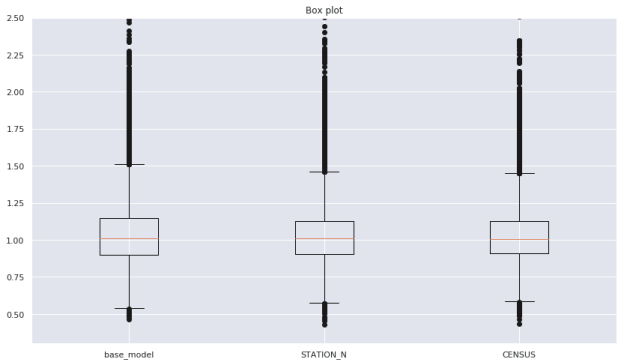


図 3: ダミー変数の効果

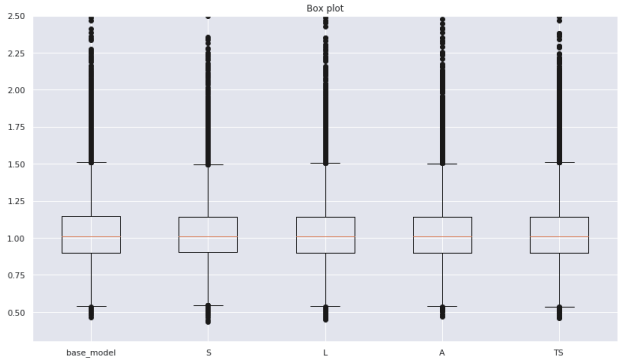


図 4: 説明変数ごとの効果

表 2: 2 段階目投入変数別の効果（ダミー変数）

	RMSE	RMSPE
BaseModel	10038752	5.1895
STATION	9566984	4.7900
CENSUS	9364350	4.6412

表 3: 2 段階目投入変数別の効果（連続量）

	RMSE	RMSPE
BaseModel	10038752	5.1895
S	9942438	5.1025
L	9976350	5.1309
A	9902803	5.1253
TS	10033856	5.1733

参考文献

[Shimizu 14] Shimizu, C., Karato, K. and Nishimura, K.: Nonlinearity of housing price structure assessment of three approaches to nonlinearity in the previously owned condominium market of Tokyo, Int. J. Housing Markets and Analysis, Vol. 7, No. 4, pp. 459-488（2014）.

[小野 02] 小野宏哉，高辻秀興，清水千弘 :「品質を考慮した中古マンション価格モデルの推定」，麗澤経済研究，第 10 巻第 2 号，pp.81-102，2002.

[福井 18] 福井光，阪井一仁，南村忠敬，三尾順一，木下明弘，田司郎 :「レインズのニューラルネットワークを用いた不動産価格査定について」，The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018.

[大和 18] 大和大祐，野村眞平 :「SUUMO でのビッグデータ活用事例」, 日本不動産学会誌／第 31 巻第 1 号, pp.78-83, 2017.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Web 不動産データを用いた空物件が入居されるまでの期間に関するデータ特性を考慮した統計モデリング

Statistical modeling of the transition time of an occupation of rental rooms by using the housing information website data

渡邊隼史 *1 Hayafumi Watanabe　一藤裕 *2 Yu Ichifuji　鈴木雅人 *3 Masahito Suzuki　山下智志 *4 Satoshi Yamashita

*1金沢大学 Kanazawa University　*2長崎大学 Nagasaki University　*3UD アセットバリュエーション UD Asset Valuation Co., Ltd.

*4統計数理研究所 The institute of statistical mathematics

The apartment loan is a loan for rentals such as for condos, apartments. This loan is a very large loan which is the account for a percentage of more than 10 percents of the whole banks' loan. However, a risk model of the apartment loan with the appropriate accuracy has not been provided in Japan mainly due to the lack of data Thus, in order to develop the risk model, we analyzed the duration in which a vacant room become occupied by a tenant by using the housing information website data, as a first step. As a result, it was found that (i) This duration can be explained by the geometric distribution, and (ii) The mixture geometric regression model considering nonlinear effects can describe the data properties. In addition, coefficients of this model roughly consistent with the empirical common senses.

1. はじめに

アパートローンは，賃貸物件向けの融資であり，銀行が貸し出す全与信額の 10 %を超える巨大な融資であるにもかかわらず，これまで十分な精度のリスク計量化モデルが考案されてこなかった．アパートローンのリスクは,（1）賃貸物件の経営自体による資金不足リスク（2）賃貸経営以外の事業によるリスクの 2 種類の要因から構成される.（2）については，一般的な信用リスクモデルやデータベースが整っており，それを用いてある程度の精度をもってリスク計量が可能である．一方（1）については，賃貸不動産の空室データベースが整っておらず，十分な精度のモデルが提供されていなかった．そこで現在（1）について空占室データベース構築と評価モデルの開発を目指しプロジェクトとして研究を行っている．なお，本研究プロジェクトの最終的な目的の一つは構築したアパートローンリスク計量モデルを CRD 協会（銀行にリスクデータベースとリスクモデルを提供する組織）を通して，民間金融機関に提供することにある．

前学会より，著者らは，アパートローンリスクの評価法の開発の第一歩として,「物件の埋まりやすさ」や「埋まりにくさ」をアパートのもつ特性から不動産情報サイトのデータ (以下 Web データ) から評価するモデルの構築について報告している．前学会では,（1）Web データは，棟のレベルでは，銀行の融資物件からの層化抽出データである鑑定士による調査データ（後述）と統計的には性質が類似しており，アパートローン用の空室率解析に利用可能性があること.（2）簡単なロジティック回帰モデルである程度は空物件が埋まる状況が説明可能なことを報告した．ただし，特に（2）について，簡単な線形ロジスティック回帰モデルでは現象記述にはやはり限界があり，モデルの説明力のうち 9 割以上が「築年数」が占めるという問題点が残った（この結果は，不動産の経験や常識とあまり整合性が高いとはいえない）.

そこで，今学会では，上記の問題を克服するためのモデルの精密化について報告する．まず，築年数の説明力が大きい理由の考察，次に，埋まる期間の確率的構造の解析を行った．さらに，それらに基づき，ロジスティック回帰モデルよりデータ構造に適合した空室が埋まる期間を記述する統計モデルを提案し，その性質を調べた．結果，データ解析では，築年数は線形の主成分解析においてデータの特性を最も説明する第一主成分であること，データ上の「空物件が埋まる期間」は幾何分布と対応する特性をもつこと；モデル化では，提案モデルは，単純なロジスティック回帰に比べて，説明力が築年数以外にも分散した常識と大きくは矛盾しない回帰係数をもつこと等を示した．また，それらの考察により物件が埋まる期間の変動がどの程度がランダムに決定し，どの程度が物件自身の特性によって決まるかの割合を見積もることが出来る可能性を示唆した．

連絡先: 渡邊隼史, hayafumi.watanabe@gmail.com

2. 関連研究

本稿で扱う研究は，データを用いた不動産評価の研究の一種といえる．データを用いた不動産評価，特に，不動産価格評価の研究は 1970 年代に主な手法が確立し，現在も大規模データや機械学習等の技術を用いて精度の改善，網羅性やリアルタイム性の性能上の向上の努力が続けられている [清田 17, 清水 17].

一方，価格ではなく，データを用いた個別物件の空占遷移予測について国内における学術論文はデータ入手の困難もあり多くは見つけられなかった．その例としては，籠，高辻，小野らがモンテカルロシミレーションによる空室率を考慮したアパートリスク計量の理論的な研究を行っている [籠 00]．また，実データを用いた例では小林の研究がある [小林 16]．小林の研究では仲介管理会社 3 件の約 6 7 3 件のデータについて空室期間等を実データから推定しそれに基づく将来収益予測モデルの構築を行っている．

本研究は実データ解析のため小林の研究に近い．小林の研究として比べた新たな貢献としては，より様々な物件要因を考慮していること，よりスケーラブルな Web データを利用すること，また不動産鑑定士の現地調査のサーベイデータ（業者による偏りが少ないデータ）と比較することで業者やサービス等

のバイアスの存在の有無を確認していること等があげられる．

3. データの取得

本研究では２つのデータを用いた．一つはメインの Web 不動産サイトデータであり，もう一つは補助的に用いる不動産鑑定士による現地調査データである．不動産鑑定士による調査データを利用する主な目的は，以下の２点である．一点目は，Web では得にくい情報（既に埋まっている物件の情報，管理の悪さなどネガティブな情報等）の効果の影響調査のため，二点目は，２つのデータを比較することで Web サイトデータのサンプリングバイアスについての情報を得るためである．

Web データは，大手不動産のサイトのある県のある路線地域について，2014 年 11 月 28 日から 2015 年夏まで，各月の 8 日， 18 日， 28 日の情報を取得した．一方，鑑定士によるサーベイデータは対応する地域についてある物件名簿より駅別に層化抽出を行った物件に対して同観測期間に３か月ごとに空室状況や物件の状況等を現地調査して得られたデータである．

本研究の Web データ解析では「空室が占室がなる」を「Web ページから物件が消える」で大まかに近似できるという仮定のもとで研究を行っている．

4. データ解析とその結果

4.1 主成分分析による築年数の説明力の高さの考察

なぜ築年数の説明力が高くなるかの原因を考察するため Web データに関する主成分分析を行った．結果，意味の解釈可能な主成分が６成分とれた．この６成分により，データ変動の意味が解釈できた６成分は，表記を,「第〇主成分 (寄与率，累積寄与率): 主成分のネーミング（寄与率が大きい要因)」としたとき，第１主成分 (6.98%，6.98%): 築年数・築年月系， 第２主成分 (5.89%，12.8%) : 水道光熱系（上下水道，空調等），第３主成分 (4.99%，17.7%): 建物構造・高さ（建築構造，高さ等），第４主成分 (3.42%，21.2%) : 各部屋の広さ（面積，賃料等），第５主成分 (3.09%， 24.3%) 入居条件（楽器，ペット，事務所可否等），第６主成分 (2.24%，26.6%) : 周辺環境 (スーパー，コンビニ，駅距離，最寄り駅等) であった．以上より，築年数が線形の範囲では，各物件の特徴をもっとも代表する変数ということが確認でき，これが空物件が入居するときに築年数が説明力が高い一つの要因と考えられる．ただし，ほかの変数の説明力も小さくないためこれだけが要因でないことも同時に考察される．

4.2 同じ物件の個別の部屋の埋まるまでの期間の解析

モデル構築するため非説明変数にあたる「埋まるまでの期間」の確率的特性を調べた．具体的には，物件内のすべての部屋が観測できる鑑定士データを用いた分散解析を行った．分散分析では，各棟ごとに埋まるまでの期間の平均 λ とその分散 σ^2（同じ物件内の埋まるまでの期間の平均からのばらつき）の関係性を調べた． その結果，大まかには，２つの量は $\sigma^2 = \lambda^2 + \lambda$ という関係性を持つことがわかった（加分散が存在する）．この結果は，非常に粗くいうと，全く同じ特性をもつ部屋でもあっても，埋まる期間のばらつきは，その物件の埋まる期間程度はあるということになる（例えば，平均１００日埋まる特性もつ部屋であっても，ばらつきが１００日なので，偶然性によって０日で埋まることもあれば 200 日程度で埋まる可能性もある [ただし，厳密言うと，埋まる期間の分布の形状は対称形をでないので完全にはこの解釈どおりにはならない]）．なお，このばらつき方は同じ棟の各部屋はランダムに埋まるという幾何過程で説明できるため，これをモデル化に利用する．

4.3 データ特性にあったモデル化と結果

上記のデータ解析等をもとに，より現実のデータ構造にあった埋まる期間のモデル（入居期間モデル）を提案した．モデルは前学会と比べて以下の特性をもつ．

- 非説明変数は，埋まる期間（前学会モデルの非説明変数は， ３か月で空物件が占か埋まらないかの二値変数）
- モデルは，混合幾何回帰モデル（前学会は，ロジスティック回帰モデル）
- 連続的な説明変数を非線形の効果を記述できるように順序カテゴリカル変数化した（前学会は，線形効果のみ）．また,「隣接するカテゴリの回帰系数の値は近い」という事前知識を事前分布という形で導入した（Bayesian lasso）．

なお，１点目はデータの情報量をより活用するため，２点目は上記のデータ解析の結果得られた構造にモデルが適合するように，３点目は，データ構造をよりモデルが柔軟に記述できるように導入した．

結果，Web データからモデルの回帰係数を推定したところ，例えば，築年数以外にも，駅やコンビニからの距離が近いほど埋まる期間の回帰係数が小さい [埋まりやすい]，建物が高いほど回帰係数が大きい [埋まりにくい]，クローゼットがあるほうが回帰係数が小さい [埋まりやすい] など，常識との相違が大きくない有意に説明力をもつ回帰係数が得られた（c.f., 単純なロジスティック回帰モデルではほとんどの説明力が築年数になる）．また，推定されたモデルパラメータの結果により，非常に大まかな目安として，変動要因別の大きさは (i)Web で可観測の物件特性効果 [± 50 日程度]，(ii)Web では非観測の物件特性効果 [± 10 日程度], (iii) 物件要因以外のランダム変動 [± 150 日程度] と見積もれる可能性があることを示唆した．

5. まとめ

Web 不動産賃貸募集データを用いた「空物件が埋まるまでの期間」のデータ特性に合わせた統計モデル（混合幾何回帰モデル）を提示した．結果（１）モデルの回帰係数は大まかには常識とは外れない（２）モデルにより,「Web で観測できる物件特性効果」,「Web では非観測の物件特性効果」,「物件要因以外のランダム効果」の大きさをそれぞれわけて見積もれる可能性を示した．ただし，まだ，賃料に関係する量（要因調整後も高賃料ほど埋まりやすい）などモデル解釈にはいくつかの課題が残るため，今後とも，その原因の解明が必要である．

参考文献

[清田 17] 清田陽司・山崎俊彦・諏訪博彦・清水千弘:不動産と AI，人工知能，32 巻，4 号，pp. 529–535 (2017)

[清水 17] 清水千弘: 不動産ビッグデータでみる不動産価格の決まり方，日本不動産学会誌，32 巻，1 号，pp. 45–51 (2017)

[籠 00] 籠義樹・高辻秀興・小川祐哉，空室率と賃料の変動過程を考慮した不動産投資モデルの期待収益率と VaR に関する研究，麗澤経済研究，8 巻，2 号，pp. 99–113 (2000)

[小林 16] 小林 秀二: 入退去ミクロ分析による不動産評価における将来予測の捉え方：生存時間解析データによる家賃キャッシュ・フローの無条件確率遷移, ファイナンシャル・プランニング研究 16 号, pp. 18-27(2016)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

AI活用による不動産分野のUX革新の取り組み

Efforts on UX innovation in real estate field by utilizing AI

清田 陽司 *1 Yoji Kiyota　椎橋 怜史 *1 Satoshi Shiibashi　二宮 健 *1 Takeshi Ninomiya　横山 貴央 *1 Takao Yokoyama　塙 拓朗 *1 Takuro Hanawa　衛藤 剛史 *1 Takeshi Eto
横山 明子 *1 Akiko Yokoyama　菊地 慧 *1 Kei Kikuchi　小林 武蔵 *1 Musashi Kobayashi　亀田 朱音 *1 Akane Kameda　瀧川 和樹 *1 Kazuki Takigawa　齋藤 裕介 *1 Yusuke Saito
花多山 和志 *1 Kazushi Katayama

*1株式会社 LIFULL
LIFULL Co., Ltd

This presentation describes attempts to innovate user experience (UX) in real estate field by artificial intelligence technologies. (1) We applied deep learning to property photographs to improve UX, both by implementing automated annotation of photographs, and by creating a new UX "search properties simply by holding a smartphone camera over the street corner". (2) For the UX problem of real estate sales transaction that it is difficult to properly pricing, we developed a service that provides reference prices of apartments in Japan on the map, using a reference price calculation algorithm by machine learning.

1. はじめに

スマートフォンやクラウドサービスなどのテクノロジーの急速な普及は，金融・流通・求職・医療・運輸など，あらゆる業界に変革をもたらしているが，その背景には，テクノロジーによって創出されるユーザ体験 (User eXperience，以下 UX と略す) が存在する．テクノロジーがビジネス上の効果を発揮するためには，業界のプレーヤーの大部分がそのテクノロジーを採用する必要があるが，導入コストや業界の慣習，規制など，普及を阻む大きな壁が存在することも多い．普及にあたっての壁を越えるための大きな原動力の一つが，優れた UX である．たとえば，Uber などの配車サービスは,「スマートフォンアプリを起動するだけで配車できる」という UX を提供することにより，爆発的に普及した．送金や小口融資，電子商取引などの分野でも，UX の優劣が市場での勝敗を左右するという時代に入っている．

不動産情報サイトの世界でも，UX の差別化による競争が非常に激しくなっている．人工知能学会全国大会オーガナイズドセッション「不動産と AI」でも，サイト上での物件推薦 [杉浦 17] やチャットによる接客 [大浜 17] など，UX 向上にフォーカスした発表が行われている．しかし，現実の不動産情報サイトは，紙媒体時代の不動産広告文化の名残で，依然として「都道府県」「路線・駅」「市区町村」を最初に選ばせる形であり，UX テクノロジーの最近の進歩 (深層学習による画像認識，AR 機能など) はまだ十分に取り込めていない．

不動産業界全体を見渡せば，物件探しは依然として「紙」「FAX」「電話」の存在を前提とした煩雑な UX になっている．一例を挙げれば，不動産情報サイトで見つけた物件への問い合わせを行った後は，まず店舗に足を運んで，プリントアウトされた物件情報チラシによる説明を受け，気に入った物件まで車で案内してもらうのが通例である．ようやく物件が決まると，ふたたび店頭で重要事項説明を受け，たくさんの書類に記入・押印することを求められる．対面でやりとりを続けることによるお互いの安心感はあるものの，スマートフォンやタブレット，VR (仮想現実) などのテクノロジーの導入によって UX を改善できる余地は大いにある．

不動産業界全体の UX 向上は，業務の生産性向上とセットで考えていく必要があるだろう．日米両国の不動産流通市場のコスト構造を分析した清水らの研究 [清水 04] は，日本における不動産流通コストを高めている要因として，登録免許税などの税金のほか，物件調査にかかるコストの大きさを指摘している．物件価値の算定など，取引全般にわたる不透明性を軽減し，業界全体で共有可能なデータベースの整備を進めていくことが，業務の効率化，ひいては UX の向上には欠かせない．

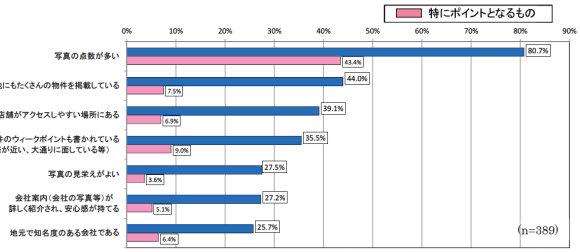
本稿では，不動産・住宅情報サイト LIFULL HOME'S における AI を活用した UX 向上の取り組み事例として，不動産物件画像への深層学習の適用による 2 つのアプリケーションと，不動産価格推定アルゴリズムを用いた参考価格算出システムを紹介する．その上で，不動産業界全体の UX を向上していくために何が必要かについて論じる．

2. 不動産物件画像への深層学習の適用

不動産情報サイトの UX において，大きなウェイトを占めるのが物件画像である．不動産情報サイトのユーザを対象としたアンケート [不動 16] では，不動産会社を選ぶ際のポイント (複数回答可) を尋ねた結果として,「写真の点数が多い」が 1 位の 80.7%,「写真の見栄えが良い」も 5 位の 27.5%となっており，物件画像が非常に重視されている傾向がうかがえる (図 1)．

LIFULL では，画像情報を含む不動産情報を活用するための知見を学術コミュニティに広く求めるとともに，AI 技術の応用研究の発展に貢献することを主な目的として，2015 年より国立情報学研究所 (NII) の協力を得て「LIFULL HOME'S データセット」を提供開始した [清田 17]．その後，データセットに含まれる約 8300 万点の不動産物件画像への研究コミュニ

連絡先: 清田 陽司，株式会社 LIFULL AI 戦略室，〒 102-0083 東京都千代田区麹町 1-4-4，Tel: 03-6774-1629，E-mail: KiyotaYoji@lifull.com



RSC によるアンケート調査結果 [不動 16] の 7 ページより引用

図 1: 不動産情報サイト利用者が不動産会社を選ぶ際のポイント (複数回答可，物件を契約した人を対象)

ティからの関心の高さや深層学習を利用した研究の成果に刺激を受け，不動産物件画像への深層学習適用を実サービスに展開する取り組みに全社的に注力している．

物件画像に関する UX の向上は，不動産事業者向け (BtoB) と一般消費者向け (BtoC) の両輪で進めていく必要がある．不動産情報サイトに掲載されている物件画像の品質を向上していくためには，不動産事業者からの協力が不可欠である．また，一般消費者の不動産物件探しの習慣を変えていく上でも，最新の UX テクノロジーの取り込みを積極的に進めていく必要がある．

上記の考え方にもとづき，LIFULL HOME'S では中古売買物件を扱う会員不動産会社向けのデータ入稿画面に深層学習を実装し，入稿作業の負担軽減と情報品質向上に活用するとともに，一般消費者向けの物件探しスマートフォンアプリにて，「かざして検索」と称する機能を深層学習の活用により実現し，新たな物件探しの UX を創出するという試みを行っている．以下に，それぞれの取り組みの概要を説明する．

2.1 BtoB での取り組み：不動産物件画像への自動タグ付け

LIFULL HOME'S では，会員である不動産会社が購入者 (入居者) の募集広告として出稿している物件情報を掲載し，物件を探しているユーザ向けに提供している．ユーザにとっては質の高い画像が多数掲載されていることが満足できる住まい探し体験に重要である一方で，不動産会社にとっては多数の画像データの入稿作業，とくに画像種別のタグ付け作業が非常に大きな負担となっており，正確なタグ付けがなされた質の高い画像情報を増やす上で障害になっていた．画像種別タグは，より価値の高い物件情報を優先的にユーザに提示することを目的に，画像種別の多様性が物件リストのランキングにも用いられており，正確なタグ付けは UX の観点でもきわめて重要である．

そこで，LIFULL HOME'S に蓄積されている膨大な画像データ資産を活用し，深層学習により画像種別の高精度な推定モデルを構築し，中古売買物件を扱う不動産会社向けの画像データ入稿画面における画像種別の自動推定・入力機能として導入した (図 2)．導入にあたっては，画像種別の付与率および付与精度，不動産会社の担当者による活用率などの具体的な指標を定めて継続的な検証を行った．導入の具体的な成果としては，不動産会社の入稿作業の負担が大幅に軽減 (1 物件あたり 導入前 40〜50 秒→導入後 2 秒) するとともに，LIFULL HOME'S に掲載された画像のうち，画像種別のタグ付けがなされた画像の割合が 2 %増加するなどの効果が得られた．

図 2: 不動産物件画像への自動タグ付け機能

2.2 BtoC での取り組み：スマートフォンアプリへの「かざして検索」機能の実装

現在，住まい探しにスマートフォンを利用するユーザの割合は 50 代以下の世代では 8 割を超えている [不動 18]．しかし，不動産情報サービスのスマートフォン (アプリを含む) の UX は，紙媒体の不動産広告文化のなごりで依然として「都道府県」「路線・駅」「市区町村」を最初に選ばせる形であり，UX テクノロジーの最近の進歩 (深層学習による画像認識，AR 機能など) を十分に取り込めていなかった．また，現状の不動産情報サービスでは物件の周辺環境に関して，満足できるレベルの情報をユーザに提供できていなかった．

そこで，街角で気になる物件の建物にスマートフォンのカメラをかざすだけで，その物件の入居募集情報を手軽に確認できる「かざして検索」を，LIFULL HOME'S のスマートフォンアプリ (iOS および Android) に AR 機能として実装した (図 3)．技術的には，カメラの画像のどの位置に建物があるのか，その建物の種類は何か (マンション，アパートなど) を深層学習，物体認識などの画像処理技術を用いて判定するとともに，GPS による位置情報，地磁気センサによる方角情報などを組み合わせて用いることにより実現した．また，シームレスな操作感を実現するため，画像処理はスマートフォン上のみで実行するとともに，建物にかざしている間の AR 表現や建物情報が見つかった場合のフィードバックなど，細部にわたるデザイン，UX の向上にも注力した．詳細な技術的内容については SlideShare の発表資料を公開している [塙 18] ので，参照されたい．

「かざして検索」機能は，2018 年 6 月のリリース以来，「街のその場の雰囲気を感じながら，カメラをかざすだけで部屋探しができる」という新たな体験価値が注目され，多くのメディアに取り上げられた [CNET Japan 18]．スマートフォンアプリのインストール数も大幅な伸びを記録し，多くの住まい探しユーザに新たな価値を届けることができている．また，スマートフォン (Android) のプラットフォーム事業者である Google 社からも，人工知能技術や AR 機能を活用した優れた事例として紹介されており，日本発のイノベーティブな事例として大きな注目を集めている．

図 3: 「かざして検索」機能

3. 不動産参考価格算出システム「LIFULL HOME'S プライスマップ」の開発

不動産業界では依然として人力主体で行われている業務も多い．なかでも，不動産物件の売買取引において，売り手と買い手の双方が納得できる適正な価格づけを行うプロセスでは，以下に挙げるような困難があり，AI 活用が進んでいなかった．

1. まったく同一の物件は存在しない
   たとえば同一のマンション内の同じ間取りの 2 つの部屋を比較してみても，日照・景観・騒音レベルなどが異なる．

2. 多数の属性が価格に影響する
   物件価格に影響を与える代表的な要因としては，面積，建物構造，築年数，交通アクセスなどが挙げられるが，ほかにも様々な要因が影響する．たとえば，ファミリー向け物件であれば学区や治安などが重視されるし，物件のメンテナンスの良否も物件価格に大きな影響を与える．さらに，地区内における属性の分布の変化 (たとえば新築分譲マンションの建設ラッシュなど) も無視できない．これらの多数の要因を考慮した価格推定モデルの構築は容易ではない．

3. リアルタイムな取引価格情報が得られない
   不動産価格の査定においては，近隣の取引事例や公示価格などのデータが参考とされるが，これらのデータに相場の変化が反映されるまでには数ヶ月以上のタイムラグが存在する．実際の取引価格と公示価格などの不動産鑑定価格を時系列に調べた清水らの研究 [Shimizu 14] によれば，1990 年代のバブル経済期には不動産鑑定価格は実際の市場価格の半分から 6 割程度であり，バブル崩壊後には 2 割程度高い価格がつけられていた．

しかしながら，複雑な物件属性の分布をうまく表現できるアプローチなどの実証研究が進められるとともに，前述のように大量の不動産物件情報が Web 広告としてリアルタイムに入手可能になったことから，近年は AI による不動産価格推定サービスの開発が世界中で進められつつある．代表事例としては，米国の大手不動産情報サイト運営会社の Zillow 社が開発している Zestimate*1 が挙げられる．Zillow 社は，不動産流通市場を効率化することを目的として，全米の 1 億件以上の物件に対して Zestimate を参考価格として提示するサービスを提供している．

*1 https://www.zillow.com/zestimate/

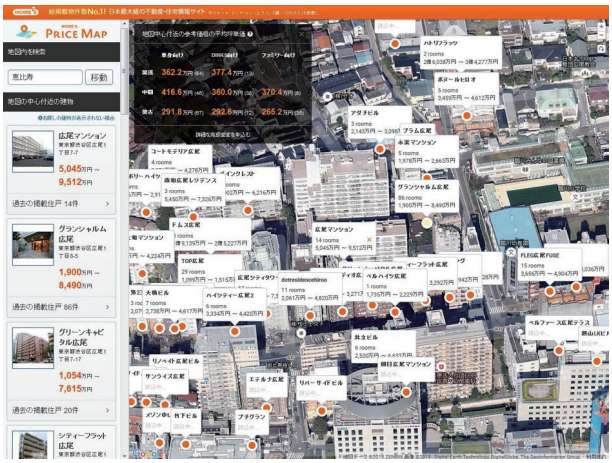


図 4: 「LIFULL HOME'S プライスマップ」の画面

そこで，日本全国の不動産物件価格の推計とその情報のオープン化を AI 活用により進め，不動産流通市場における UX を改善し，ひいては空き家問題などの社会的問題の解消に資することを目指して，サービスの開発に着手した．このサービスは「LIFULL HOME'S プライスマップ」(以下,「プライスマップ」と略す) として，2015 年よりインターネット上で公開している (図 4)*2．

プライスマップにおける不動産参考価格算出に用いられているデータは，LIFULL HOME'S に掲載された物件データである．これまでに掲載した物件情報はデータベースに蓄積しており，過去 10 数年分のデータが存在する．このうち，AI (機械学習) による参考価格算出に使用している掲載物件データ数は 2018 年 11 月現在売買物件で 300 万件超，賃貸物件においては 6,000 万件を超え，今後も増えていく見込みである．

プライスマップの開発にあたっては，不動産物件価格推定の推定手法を長年研究している清水千弘氏の助言を受け，一般的な不動産鑑定フローにならって LIFULL HOME'S の掲載物件データから「立地要素」「時間要素」「物件属性要素」の 3 要素を考慮してリアルタイムに参考価格を計算する手法を実装している．また，推定手法としては，線形回帰モデルを基本としている．推定手法の基本的な考え方については横山らによる既発表 [横山 17] を参照されたい．

4. 不動産業界全体の UX を向上していくために必要なこと

不動産業界における UX は，インターネットおよび Web の発達に牽引されて向上してきた．大量の不動産データが Web 上に集積することが，AI による不動産物件価格推定サービスの開発や，深層学習や VR などのテクノロジーを活用したイノベーションを促してきている．また，高度成長期には合理的だった業界の仕組みが，急速な少子高齢化の進展などによって矛盾をきたす中，異業種やスタートアップなど，業界外からのチャレンジャーの参入が，変革を促している．

一方で，Web に集積されている不動産データは，不動産ストック全体からみればごく一部のデータにすぎない．Web に掲載されている物件価格は売主の希望価格であり，実際の取引価格とはずれがある．しかし，物件価格推定サービスの多く

*2 https://www.homes.co.jp/price-map/

は，データの取得を不動産情報サイトからのスクレイピングなどに依存していることが指摘されており，算出された不動産物件価格の信頼性やデータ取得の安定性に難がある．

不動産価格推定のアルゴリズムは，研究コミュニティを通じたデータベース共有を軸に発達してきた．多大なコストをかけてデータベースを整備している不動産業界のプレーヤーが外部にデータベースを提供してきた背景には，データベース整備の営みへのリスペクトがある．スクレイピングに依存したデータ取得と商業利用が野放図に広がることは，データベース整備へのリスペクトを失わせるだけでなく，信頼性に欠ける不動産物件価格の指標が広まってしまい，最終的には不動産業界全体への信頼を失わせることにもなりかねない．

また，将来にわたって満足できる不動産物件選びのためには，ローンや保険などの金融サービス，医療・介護・教育などの社会的サービス，地域コミュニティの在り方なども考慮に入れる必要があるため，金融業界，医療業界，介護業界，地方自治体などとの密接な連携が必要となるが，現時点ではそのような連携は希薄である．現在起きている過剰な不動産投資の問題も，業界の壁をまたいだ連携の仕組みがあれば，未然に防げた可能性もある．

今後，不動産業界全体で優れた UX が広く提供されるようになるためには，業界全体をあげて網羅性の高い不動産物件データベースの整備や，治安・防災・教育などの住生活に関わるオープンデータの充実を図るとともに，さまざまなステークホルダーとの連携を深めていくことが大切になるだろう．

参考文献

[CNET Japan 18] CNET Japan, : LIFULL、スマホをかざして見つける新しい部屋探し–空き室情報を AR で表示 , https://japan.cnet.com/article/35119662/ (2018)

[Shimizu 14] Shimizu, C., Karato, K., and Nishimura, K.: Nonlinearity of housing price structure Assessment of three approaches to nonlinearity in the previously owned condominium market of Tokyo, *International Journal of Housing Markets and Analysis*, Vol. 7, No. 4, pp. 459–488 (2014)

[横山 17] 横山 貴央, 清水 千弘 : 都市間比較を目的とした住宅価格指数の整備に関する研究, 第 31 回人工知能学会全国大会 (JSAI 2017) 予稿集, pp. 1H2–OS–15a–2 (2017)

[杉浦 17] 杉浦 大樹, 野村 眞平 : SUUMO での不動産データ活用の取り組みと未来, 第 31 回人工知能学会全国大会 (JSAI 2017) 予稿集, pp. 1H3–OS–15b–4in1 (2017)

[清水 04] 清水 千弘, 西村 清彦, 浅見 泰司 : 不動産流通システムのコスト構造–不動産取引コストの把握, 住宅土地経済, No. 51, pp. 28–37 (2004)

[清田 17] 清田 陽司, 石田 陽太 : 学術コミュニティへのデータセット提供を通じた不動産領域におけるオープンイノベーション推進, 人工知能, Vol. 32, No. 4, pp. 584–589 (2017)

[大浜 17] 大浜 毅美 : 不動産仲介マーケティングのためのユーザ行動予測, 第 31 回人工知能学会全国大会 (JSAI 2017) 予稿集, pp. 1H3–OS–15b–3 (2017)

[塙 18] 塙 拓朗 : LIFULL HOME'S「かざして検索」リリースの裏側, https://www.slideshare.net/takurohanawa/lifull-homes-112519081 (2018)

[不動 16] 不動産情報サイト事業者連絡協議会 :「不動産情報サイト利用者意識アンケート」2016 年度調査結果, https://www.rsc-web.jp/pre/img/161027.pdf (2016)

[不動 18] 不動産情報サイト事業者連絡協議会 :「不動産情報サイト利用者意識アンケート」2018 年度調査結果, https://www.rsc-web.jp/pre/img/181025.pdf (2018)

Organized Session | Organized Session | [OS] OS-3

[1E2-OS-3a] AI における離散構造処理と制約充足(1)

波多野 大督（理化学研究所）、蓑田 玲緒奈（（株）ベイシスコンサルティング）

Tue. Jun 4, 2019 1:20 PM - 3:00 PM Room E (301A Medium meeting room)

[1E2-OS-3a-01] (Invited talk) Sentential Decision Diagrams and related topics

〇Masaaki Nishino[1] （1. NTT Communication Science Laboratories）

1:20 PM - 2:00 PM

[1E2-OS-3a-02] MCSes Enumeration with the Glucose SAT Solver

〇Miyuki Koshimura[1], Ken Satoh[2] （1. Kyushu University, 2. National Institute of Informatics）

2:00 PM - 2:20 PM

[1E2-OS-3a-03] A SAT-based CSP Solver sCOP and its Results on 2018 XCSP3 Competition

〇Takehide Soh[1], Daniel Le Berre[3,4], Mutsunori Banbara[2], Naoyuki Tamura[1] （1. Kobe University, 2. Nagoya University, 3. CRIL-CNRS, UMR 8188, 4. Université d'Artois）

2:20 PM - 2:40 PM

[1E2-OS-3a-04] An Analysis of Entry and Exit Data in Office by Decision Tree Learning Using Clustering Factor Matrix from Non-negative Multiple Matrix Factorization

〇Seidai Kojima[1], Hayato Ishigure[1], Miwa Sakata[1], Atsuko Mutoh[1], Koichi Moriyama[1], Nobuhiro Inuzuka[1] （1. Nagoya Institute of Technology）

2:40 PM - 3:00 PM

©The Japanese Society for Artificial Intelligence

1:20 PM - 2:00 PM (Tue. Jun 4, 2019 1:20 PM - 3:00 PM Room E)

[1E2-OS-3a-01] (Invited talk) Sentential Decision Diagrams and related topics

〇Masaaki Nishino[1] （1. NTT Communication Science Laboratories）

Keywords: Sentential Decision Diagrams, Discrete Structure, Binary Decision Diagrams

Sentential Decision Diagrams （ SDD）は近年提案された決定グラフの一種であり，二分決定グラフ (BDD）と同等の演算性能を保持しつつ， BDDよりも簡潔な表現が可能という特徴をもつ．本講演では SDDの紹介と，講演者らの最近の研究成果である Zero-suppressed SDD （ ZSDD）や， SDDの効率的な構築方法などを紹介する．

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

SAT ソルバー Glucose を用いた MCS 列挙

MCSes Enumeration with the Glucose SAT Solver

越村 三幸 *1　佐藤 健 *2
Miyuki Koshimura　Ken Satoh

*1九州大学 大学院システム情報科学研究院
Faculty of Information and Electrical Engineering, Kyushu University

*2国立情報学研究所
National Institute of Informatics

Enumerating all Maximal Satisfiable Subsets (MSSes) or all Minimal Correction Subsets (MCSes) of an unsatisfiable CNF Boolean formula is a cornerstone task in various AI domains. This paper considers MCSes enumeration with a SAT solver. We aim to develop a procedure which outperforms several MCSes enumerators proposed so far. The paper presents a basic enumeration procedure and compares it with a state-of-the-art enumerator `Enum-ELS-RMR-Cache`. The experimental results show that the proposed procedure is more efficient than `Enum-ELS-RMR-Cache` to solve Partial-MaxSAT instances but it is inefficient than `Enum-ELS-RMR-Cache` to solve plain MaxSAT instances.

1. はじめに

解を一つも持たない制約集合の MSS（Maximal Satisfiable Subset）あるいは MCS（Minimal Correction Subset）を求めることは，人工知能の様々な分野で重要とされている [Grégoire 18]．本稿では，SAT ソルバーを利用した MCS の列挙を論ずる．同様の研究は多くあるが，我々は，効率性の点でそれらを上回る列挙手続きの提案と実装を目的としている．本稿ではその基本手続きを示し，その実装の性能を現時点で最高性能を示していると思われる `Enum-ELS-RMR-Cache`[Grégoire 18] と比較する．

2. 準備

本稿では，問題は節集合で与えられるものとする．節（clause）はリテラル（literal）の選言（論理和 $\vee$）でリテラルはブール変数あるいはその否定（$\neg$）である．ブール変数の集合を Var とする．変数の値割当て μ は，Var から $\{0,1\}$ の写像である．節集合 Σ に現れる変数の値割当て μ で Σ の全ての節を 1 にするものがある時，Σ は充足可能（satisfiable），そのような μ がない時，充足不能（unsatisfiable）であるという．

本稿では二種類の節，ハード節とソフト節を扱う．ハード節は必ず成り立って欲しい制約，ソフト節は出来るだけ成り立って欲しい制約を表す．ハード節の集合 Σ_1 とソフト節の集合 Σ_2 からなる節集合 $\Sigma\ (= \Sigma_1 \cup \Sigma_2)$ に対して，MSS と MCS は次のように定義される．

定義 1 (MSS) Σ の *MSS*（*maximal satisfiable subset*）Φ は，Σ_1 を含む Σ の部分集合（$\Sigma_1 \subseteq \Phi \subseteq \Sigma$）で，充足可能，かつ，$\forall \alpha \in \Sigma \setminus \Phi, \Phi \cup \{\alpha\}$ は充足不能，を満たすものである．

MSS は充足可能性を保ちつつ，Σ_1 を Σ_2 のソフト節でぎりぎりまで拡大した節集合である．

定義 2 (MCS) Σ の *MCS*（*minimal correction subset*）Ψ は，Σ の部分集合（$\Psi \subseteq \Sigma$）で，その補集合，つまり $\Sigma \setminus \Psi$ が Σ の *MSS* であるものである．

連絡先: 越村三幸，九州大学大学院システム情報科学研究院，〒 819-0395 福岡市西区元岡 744，092-802-3599，koshi@inf.kyushu-u.ac.jp

MSS は Σ_1 を含むので，MCS は Σ_2 の部分集合である．MCS は Σ_2 の部分集合で，それを取り除くと Σ が充足可能になる，ぎりぎりまで小さい節集合である．

SAT ソルバーを利用した MCS 列挙プログラムではしばしば，ソフト節に対し節選択変数（clause selector）と呼ばれる変数が次のように導入される．それぞれのソフト節 $\alpha \in \Sigma_2$ に対し，新変数 s_α を導入し，その否定を元のソフト節に加えた $\alpha \vee \neg s_\alpha$ を作る．これにより，新しい節集合 $\Sigma_2^S = \{\alpha \vee \neg s_\alpha \mid \alpha \in \Sigma_2\}$ が得られる．そして，SAT ソルバーを，Σ そのものでなく，$\Sigma_1 \cup \Sigma_2^S$ の充足可能性判定に利用する．

$s_\alpha = 1$ の仮定の下では，$\alpha \vee \neg s_\alpha$ と元の節 α は充足可能性が一致するので，$s_\alpha = 1$ とすることを $\alpha \vee \neg s_\alpha$ を活性化（activate）する，という．逆に $s_\alpha = 0$ とすることを非活性化（deactivate）する，という．MSS を求めることは，$\Sigma_1 \cup \Sigma_2^S$ の充足可能性を保ちつつ，活性化された選択変数をぎりぎりまで増やす，ことに対応する．言い換えると，MCS を求めることは，充足可能性を保ちつつ，非活性化された選択変数をぎりぎりまで減らす，ことに対応する．このことを利用して，次節のアルゴリズムは MCS を列挙する．

3. MCS 列挙アルゴリズム

Algorithm 1 に提案アルゴリズムの概観を示す．6 行目の $SAT(\Sigma_1 \cup \Sigma_2^S, A)$ が SAT オラクルを示し，A に含まれるリテラルの値が全て 1 と仮定した時の $\Sigma_1 \cup \Sigma_2^S$ の充足可能性を判定する．充足可能であれば st が $TRUE$ となり，見つけた変数の値割当て μ が得られる．充足不能であれば，st が $FALSE$ となる．

本アルゴリズムは，SAT オラクルが見つけた変数の値割当てを出発点として MCS を探していく．アルゴリズム中の A と B は選択変数の集合であり，A に対応するソフト節の集合 $\{\alpha \mid \mu(s_\alpha) = 1\}$（以降 A_{MSS}）が MSS の候補であり，B に対応するソフト節の集合 $\{\alpha \mid \mu(s_\alpha) = 0\}$（以降 B_{MCS}）が MCS の候補である．このアルゴリズムでは，SAT オラクルが充足可能を返す限り，6〜10 行目が繰り返され，B は段々と小さくなり B_{MCS} は MCS に近づいていく．逆に A は段々と大きくなり A_{MSS} は MSS に近づいていく．10 行目の $\bigvee_{s_\alpha \in B} s_\alpha$ は，B_{MCS} より大きな集合を以降の MCS の探索から除外す

Algorithm 1 Enum-MCS (Σ の全ての MCS を列挙する)

入力： $\Sigma (= \Sigma_1 \cup \Sigma_2)$（充足不可能な節集合，$\Sigma_1$ はハード節の集合，Σ_2 はソフト節の集合）

出力： Σ の全ての MCS

1: $\Sigma_2^S \leftarrow \{\alpha \vee \neg s_\alpha \mid \alpha \in \Sigma_2\}$; // s_α は α の選択変数
2: $S \leftarrow \{s_\alpha \mid \alpha \in \Sigma_2\}$; // 選択変数の集合
3: $A \leftarrow \emptyset$; // MSS 候補に対応する選択変数の集合
4: $B \leftarrow \emptyset$; // MCS 候補に対応する選択変数の集合
5: **while true do**
6: $(st, \mu) = SAT(\Sigma_1 \cup \Sigma_2^S, A)$ // SAT オラクル
7: **if** $st = TRUE$ **then**
8: $A \leftarrow \{s_\alpha \mid \mu(s_\alpha) = 1\}$
9: $B \leftarrow \{s_\alpha \mid \mu(s_\alpha) = 0\}$
10: $\Sigma_1 \Leftarrow \Sigma_1 \cup (\bigvee_{s_\alpha \in B} s_\alpha)$ // 阻止節
11: **else if** $A = \emptyset$ **then**
12: **return**
13: **else**
14: $output(B)$; // B_{MCS} が MCS
15: $A \leftarrow \emptyset$;
16: **end if**
17: **end while**

表 1: 変数と節の個数の平均

	変数の数	節の数	
		ハード節	ソフト節
MS	156,844	0	496,827
PMS	16,032	100,134	10,759

るための節である．

SAT オラクルの答えが充足不能なら，その時の B_{MCS} より小さな MCS 候補はない，ことになり，この B_{MCS} が MCS であることわかる（14 行目）．15 行目の $A \leftarrow \emptyset$ は，MSS 候補がない状態から再び探索を始めることを表す．SAT オラクルの答えが充足不能で A が空なら，もうこれ以上 MCS はないことが分かり，手続きを終了する（11，12 行目）．

なお，本アルゴリズムは，各選択変数に否定をつけたリテラル集合 $S^- = \{\neg s_\alpha \mid \alpha \in \Sigma_2\}$ に関する極小モデルを列挙している，と見なすことができ，[Koshimura 09] の極小モデル生成と手続き的には同じである．

4. 計算機実験

前節のアルゴリズムを SAT ソルバー Glucose 3.0[Audemard 09, Eén 03] を用いて実装した．評価には，[Grégoire 18] と同じ 1090 個のインスタンスを用いた．その内訳は，ソフト節のみからなるインスタンス（以降 MS）が 493 個，ソフト節とハード節からなるインスタンス（以降 PMS）が 597 個である．

実験には，メモリ 32Gb を有する Intel Xeon E3-1246v6(3.70GHz) プロセッサ上の Ubuntu 18.04 を用いた．1 インスタンス当たりの制限時間と制限メモリは，[Grégoire 18] と同じく，それぞれ，30 分，8Gb とした．

表 1 は，MS，PMS，それぞれのインスタンス当たりの変数の数と節の数の平均を示している．MS の方が，変数の数，節の数のいずれも平均的には多いことが分かる．表 2 に，上記の制限時間及び制限メモリで列挙した MCS の総和を MS，PMS 毎に示した．比較のため [Grégoire 18] で提案された MCS 列挙プログラム `Enum-ELS-RMR-Cache`*1 の列挙数も示した．これは，現段階で最も高速に MCS を列挙するプログラムの一つである．

表 2: MCS 総列挙数（単位：千個）

	`Enum-ELS-RMR-Cache`	提案手法
MS	349,155	60,173
PMS	494,308	637,800

表から，MS に対しては `Enum-ELS-RMR-Cache` が，PMS に対しては提案手法が優位であるのが分かる．現実的な問題はハード節とソフト節の双方を用いて表現されることが多いと思われるので，この結果から，提案手法の方が現実的な問題に適用するには優位である，と言ってもいいだろう．

提案手法の性能が MS に対して芳しくない原因の一つとして，MS のソフト節が多い，ことが考えられる．提案方式では，1 つのソフト節に対して 1 つの選択変数が導入される．現実装では，選択変数の連言を Glucose の assumption に設定し，MCS を求める．ソフト節が増えると連言，つまり assumption も多くなるが，これが多くなると SAT ソルバーの性能が落ちていくことが知られている．MS に対する性能を上げるには，これに何らかの対応が必要だと思われる．先駆的な研究 [Audemard 13] を足がかりとしたい．

5. おわりに

SAT オラクルを利用して MCS を列挙する手続きを提案し，SAT ソルバー Glucose を用いて実装した．`Enum-ELS-RMR-Cache` との比較実験では，提案手法は PMS に対しては優位であったが，MS に対しては劣っていた．本提案手法は，`Enum-ELS-RMR-Cache` の手続きと比べると簡潔である．それにも関わらず，PMS に対して優位性を示せたことは本手続きの将来性の高さを示していると思われる．今後は，実験結果を詳細に分析し，手続きの改良手法を考えていきたい．

謝辞：本研究は JSPS 科研費 JP16K00304，JP17K00307 の助成を受けたものです．

参考文献

[Audemard 09] G. Audemard, L. Simon: Predicting Learnt Clauses Quality in Modern SAT Solver, IJCAI-09, pp.399-404, 2009.

[Audemard 13] G. Audemard, J.-M. Lagniez, L. Simon: Improving Glucose for Incremental SAT Solving with Assumption: Application to MUS Extraction, SAT 2013, pp.309-317, 2013.

[Eén 03] N. Eén, N. Sörensson: An Extensible SAT-solver, SAT-03, pp.502-518, 2003.

[Grégoire 18] É. Grégoire, Y. Izza, J.-M. Lagniez: Boosting MCSes Enumeration, IJCAI-18, pp.1309-1315, 2018.

[Koshimura 09] M. Koshimura, H. Nabeshima, H. Fujita, R. Hasegawa: Minimal Model Generation with Respect to an Atom Set, FTP 2009, pp.49-59, 2009.

*1 http://www.cril.fr/enumcs から入手した．なお，実験で用いた 1090 個のインスタンスも同サイトから入手した．

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

A SAT-based CSP Solver sCOP and its Results on 2018 XCSP3 Competition

Takehide Soh*1 Daniel Le Berre*2 Mutsunori Banbara*3 Naoyuki Tamura*1

*1Kobe University *2CRIL-CNRS, UMR 8188, Université d'Artois *3Nagoya University

Constraint Satisfaction Problem (CSP) is the combinatorial problem of finding a variable assignment which satisfies all given constraints over finite domains. CSP has a wide range of applications in the research domains of Artificial Intelligence and Operations Research. XCSP3 is one of major constraint languages that can describe CSPs. More than 23,000 instances over 105 series are available in the XCSP3 database. In 2018, the international XCSP3 competition was held and 18 solvers participated.

This paper describes the under development CSP solver sCOP and its results on the 2018 XCSP3 competition. sCOP is a SAT-based solver which encodes CSPs into SAT problems and finds a solution using SAT solvers. Currently, sCOP equips the order and log encodings, and uses off-the-shelves backend SAT solvers. We registered sCOP to two competition tracks—CSP-Standard-Sequential and CSP-Standard-Parallel—of the 2018 XCSP3 competition and won both tracks.

1. Input of sCOP—XCSP3 Language

This section explains the XCSP3 language [Boussemart 17], an input of the CSP solver sCOP. XCSP3 is an XML based constraint language that can describe CSPs. Let's start with an easy example of XCSP3 and then explain its constraints.

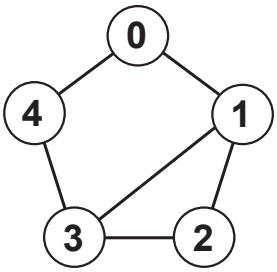
1.1 Example

The graph coloring problem (GCP) is a problem whose goal is to assign a color to each node in a given graph such that there is no two same-colored nodes connected by an edge. Figure 1 shows an instance of GCP.

This GCP can be represented by a CSP. We introduce five variables n_0, n_1, n_2, n_3, n_4. Each variable has the same domain $\{0, 1, 2\}$ which represents different colors like red, green, and blue. We also introduce the conjunction of the six inequalities representing the constraints of the GCP instance.

$$n_0 \neq n_1 \qquad n_0 \neq n_4 \qquad n_1 \neq n_2$$
$$n_1 \neq n_3 \qquad n_2 \neq n_3 \qquad n_3 \neq n_4$$

This CSP can be represented by the XCSP3 instance in Fig. 1. Line 1 describes the format name "XCSP3" and the type "CSP" of the given problem. Line 2 to 4 describe the five integer variable by using array. The domain of those array variables is set to the range from 0 to 2, and their identifier is set to "n". Line 5 to 14 describe the conjunction of the six inequalities.




```
<instance format="XCSP3" type="CSP">
  <variables>
    <array id="n" size="[5]"> 0..2 </array>
  </variables>
  <constraints>
    <intension>
      and(ne(n[0],n[1]),
          ne(n[0],n[4]),
          ne(n[1],n[2]),
          ne(n[1],n[3]),
          ne(n[2],n[3]),
          ne(n[3],n[4]))
    </intension>
  </constraints>
</instance>
```

Figure 1: (top) GCP instance (bottom) its XCSP3 representation

1.2 Constraints in XCSP3

In the previous example, constraints are represented by using *intentional* constraints (ne). Except intensional constraints, we can use *extensional* and *global* constraints in XCSP3. This section explains each kind of constraints.

Intentional constraints are constraints using the arithmetic, logical, and comparison operators. As arithmetic operators, `add` (addition), `sub` (subtraction), `mul` (multiplication), `div` (integer division), `abs` (absolute value), etc. can be used. As logical operators, `not` ($\neg$), `and` ($\wedge$), `or` ($\vee$), `xor` ($\oplus$), `iff` ($\Leftrightarrow$), `imp` ($\Rightarrow$), etc. can be used. As comparison operators, `le` ($\leq$), `lt` ($<$), `ge` ($\geq$), `gt` ($>$), `ne` ($\neq$), `eq` ($=$), etc. can be used. For instance, a constraint $(2x + 3 < y) \vee z \geq 1$ can be represented in the following intensional constraints of XCSP3.

```
<intension>
  or(lt(add(mul(2,x),3),y),ge(z,1))
</intension>
```

Contact: Takehide Soh, Information Information Science and Technology Center, Kobe University, 1-1 Rokkodai, Nada, Kobe, 657-8501, soh@lion.kobe-u.ac.jp

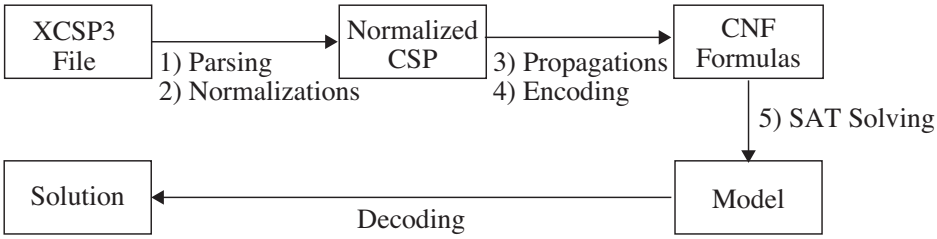


Figure 2: Framework of sCOP

Extensional constraints represent feasible (or infeasible) value combinations of variables extensionally. It is also called *table* constraints. As the name suggests, constraints are often represented by tables. For instance, one of the inequalities in the previous GCP example, $n_0 \neq n_1$ is represented by the following tables.

(Support Table for $n_0 \neq n_1$)

n_0	n_1
0	1
0	2
1	0
1	2
2	0
2	1

(Conflict Table for $n_0 \neq n_1$)

n_0	n_1
0	0
1	1
2	2

There are two types of extensional constraints: *support* and *conflict*. The former represents feasible value combinations of variables and the latter represents infeasible ones. Obviously, they complement each other and one of them is enough to represent constraints. The conflict table above can be represented in the following extensional constraints of XCSP3.

```
<extension>
  <list> n[0] n[1] </list>
  <conflicts> (0,0)(1,1)(2,2) </conflicts>
</extension>
```

Global constraints represent comparatively complex but useful constraints to model applications. In XCSP3, 19 global constraints are available. For instance, *allDifferent* is such a global constraint ensuring mutually different values must be assigned to given variables. An implicit allDifferent constraints between nodes 1, 2, and 3 in the previous GCP instance can be represented in the following constraints of XCSP3.

```
<allDifferent>
  n[1] n[2] n[3]
</allDifferent>
```

Other than above, XCSP3 provides more advanced constructs for symbolic/set/real/graph/stochastic/qualitative variables, soft constraints, mono and multi objective optimization, etc. In XCSP3 competitions, a set of basic constructs named XCSP3-Core is used. Interested reader is referred to the full specification [*1].

2. SAT-based CSP Solver sCOP

sCOP [Soh 18] is a SAT-based CSP Solver written in Scala. Given a XCSP3 instance file, sCOP finds a solution using a SAT solver. Figure 2 shows the framework of sCOP.

Following Sugar [Tamura 08] and Diet-Sugar [Soh 17], sCOP encodes CSPs (XCSP3 instances) into SAT problem in conjunctive normal form (CNF) using (i) the order encoding [Tamura 09, Tamura 13] and (ii) the log encoding for Pseudo-Boolean (PB) constraints [Soh 17]. Then, sCOP invokes a SAT solver that would find a model if any. The obtained model is decoded into a solution of the XCSP3 instance. sCOP is publicly available in its web page [*2].

1) Parsing. Parsing of XCSP3 formatted file is done by using an XCSP3 official tool XCSP3-Java-Tools [*3]. Currently, sCOP accepts all constraints in the XCSP3-core language[*1].

2) Normalization. Before pre-processing, all constraints are translated into intensional constraints. This normalization is processed as follows. Global constraints are translated into intensional constraints by a straightforward way but we use extra pigeon hole constraints for allDifferent constraints as in Sugar [Tamura 08]. Extensional constraints are translated into intensional constraints by using a variant of multi-valued decision diagrams. This is a difference to ones in Sugar. All intentional constraints are normalized to be in the form of CNF using Tseitin transformation. Literals of this CNF-CSP are linear comparisons $\sum_i a_i x_i \geq k$ where a_i's are integer coefficients, x_i's are integer variables and k is an integer constant.

3) Pre-processing: propagation. Constraint propagations are executed to the normalized CSP (clausal CSP, i.e., in the form of CNF over linear comparisons $\sum_i a_i x_i \geq k$) to remove redundant values, variables, and linear comparisons. Currently, it is done by using an AC3 like algorithm.

4) Encoding into SAT. In sCOP, the order encoding [Tamura 09, Tamura 13] and the log encoding are used. The order encoding uses propositional variables $p_{x \geq d}$'s meaning $x \geq d$ for each domain value d of each integer variable x. To encode linear comparisons, Algorithm 1 of the literature [Tamura 13] is used in sCOP. The log encoding uses a binary representation of integer variables. There are several ways to encode linear comparisons by using those propositional variables. In sCOP, we replace all integer variables with its binary representation—it gives us a set of PB constraints. We then encode those PB constraints into CNF formulas by using the BDD encoding [Eén 06]. sCOP basically uses the order encoding but uses the log encoding in case that the huge number of clauses is expected to be encoded. For this expectation, the idea of domain product criteria [Soh 17] is used.

5) SAT Solving. By using DIMACS CNF files, sCOP's backend

*1 http://www.xcsp.org/specifications

*2 https://tsoh.org/sCOP/

*3 https://github.com/xcsp3team/XCSP3-Java-Tools

SAT solver is switchable. In the 2018 XCSP3 competition, sCOP uses MapleCOMSPS and glucose-syrup. A SAT solver MapleCOMSPS *4 is used for sequential CSP solving. It is the winning solver on the main track of the SAT competition 2016. It also shows a good performance for solving CSP instances encoded by sCOP. A SAT solver glucose-syrup *5 is used for parallel CSP solving. It is the winning solver on the parallel track of the SAT competition 2017.

Example. After saving the XCSP3 instances of Figure 1 as a plain text file `gcp.xml`, the execution of sCOP returns the following results using a default SAT solver.

```
$ java -jar scop.jar gcp.xml
<... encoding information ...>
s SATISFIABLE
v <instantiation>
v     <list>n[0] n[1] n[2] n[3] n[4]</list>
v     <values>2 0 2 1 0</values>
v </instantiation>
```

3. Results on 2018 XCSP3 Competition

International competitions of CSP solvers using the XCSP language have been held since 2005. The first series were held with earlier versions of XCSP languages and called International CSP Solver Competition (CSC). CSCs were held in 2005, 2006, 2008 and 2009. After some break period, it is re-started in 2017 with XCSP3. In those competitions, solvers are submitted from research institutions over Europe and North America.

3.1 Overview of 2018 XCSP3 Competition

In 2018 XCSP3 Competition, there are the following 8 tracks.

	Solver	Seq/Par	Timeout	#Ins.	#Sol.
CSP	Standard	Sequential	40 min.	236	14
		Parallel	40 min.	236	4
COP		Sequential	40 min.	346	8
		Parallel	40 min.	346	1
		Sequential	4 min.	346	7
		Parallel	4 min.	346	1
CSP	Mini	Sequential	40 min.	176	11
COP		Sequential	40 min.	188	7

CSP and *COP* are categories for Constraint Satisfaction Problem and Constraint Optimization Problem. There are two solver categories: *Mini* is a category for "mini solvers" which is comparatively small and simple solvers. Also, those mini solvers must be open source software. *Standard* is a category for other solvers. *Sequential* and *Parallel* are categories for sequential solvers and parallel solvers that can use eight CPU cores. For COP category, there is a "Fast" solver track whose goal is to evaluate solvers computing a good quality solution within a comparatively short time—4 minutes. In the other tracks, 40 minutes for each instance are given to solvers. About benchmark, 236 instances are selected for CSP categories and 346 instances are selected for COP categories. In case for "mini solvers", benchmark instances are limited to contain only intensional constraints or some basic global constraints.

*4 https://sites.google.com/a/gsd.uwaterloo.ca/maplesat/

*5 http://www.labri.fr/perso/lsimon/glucose/

Table 1: Ranking of CSP-Standard-Sequential Track

Rank	Solver	#Solved	S/U	%VBS
—	VBS (Virtual Best Solver)	163	103/60	100
1	scop-order+maple	146	92/54	90
2	scop-both+maple	140	87/53	86
3	PicatSAT	138	85/53	85
4	Mistral-2.0	116	80/36	71
5	Choco-solver-4.0.7b seq	115	77/38	71
6	Concrete-3.9.2	92	64/28	56
7	OscaR - Conf. Ord. Res.	90	62/28	55
8	Concrete-3.9.2-SuperNG	84	55/29	52
9	Sat4j-CSP	83	40/43	51
10	OscaR - Conf. Order.	81	51/30	50
11	cosoco-1.12	79	53/26	48
12	BTD	76	31/45	47
13	BTD_12	76	32/44	47
14	macht	66	33/33	40

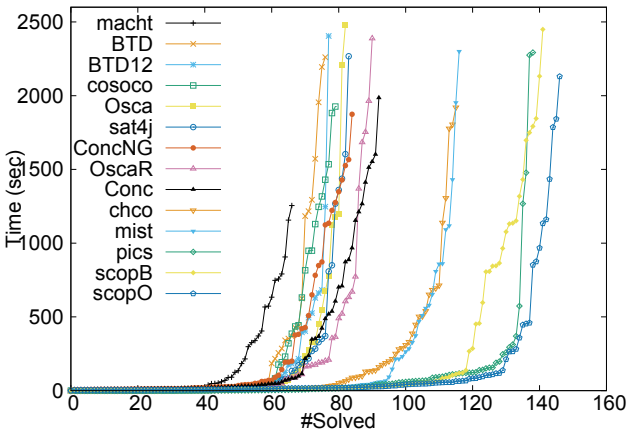


Figure 3: Cactus Plot of CSP-Standard-Sequential Track

Among those eight tracks, two tracks were canceled because there was only one solver registered.

sCOP registered to CSP-Standard-Sequential and CSP-Standard-Parallel tracks. The former is popular and the most competitive track in a sense of the number of participated solvers.

3.2 Results of sCOP

Table 1 shows the solver ranking of the CSP-Standard-Sequential track with regards to the number of solved instances. The first row shows the abstract VBS (virtual best solver) which is the collection of the best solver for each instance. The third column shows the number of instances solved and the fourth column shows the number of satisfiable and unsatisfiable instances solved. The fifth column shows the percentages of the number of solved instances w.r.t. VBS. Among all 14 solvers, sCOP using the order encoding solved the most number of instances. sCOP using the both of the order and log encodings takes the second place.

Figure 3 shows the cactus plot whose x-axis is the number of instances solved and y-axis is the CPU time in seconds. The meaning of the cactus plot is that "each of x instances were solved within y CPU seconds". Note that "which x instances" are different for each solver. For each plot, being righter means solvers solve more instances, being lower means solvers solve instances faster.

Table 2: Number of Instances Solved: Figures are organized by instance series.

Series Name	#Ins.	Choco-solver-4.0.7b	Mistral-2.0	PicatSAT	scop-both+maple	scop-order+maple	Ctrs.
Bibd	12	4	**8**	**8**	**8**	**8**	pure int
CarSequencing	17	11	6	**17**	12	**17**	sup.
ColouredQueens	12	3	**4**	**4**	3	3	pure int
Crossword	13	**5**	3	2	3	3	sup.
Dubois	12	7	7	**12**	**12**	**12**	sup.
Eternity	15	**7**	**7**	6	6	6	sup.
Frb	16	3	3	3	**5**	**5**	conf./sup.
GracefulGraph	11	**6**	**6**	**6**	**6**	**6**	pure int
Haystacks	10	4	2	**10**	**10**	**10**	conf./sup.
Langford	11	7	**9**	**9**	**9**	**9**	pure int
MagicHexagon	11	4	**5**	3	3	3	pure int
MisteryShopper	10	**10**	**10**	**10**	**10**	**10**	sup.
PseudoBoolean-dec	13	4	6	4	7	**8**	pure int
Quasigroups	16	6	6	7	**8**	**8**	pure int
Rlfap-dec-scens11	12	11	10	**12**	**12**	**12**	pure int
SocialGolfers	12	6	6	**8**	**8**	**8**	pure int
SportsScheduling	10	3	**4**	**4**	**4**	**4**	sup.
StripPacking	12	7	8	**11**	**11**	**11**	sup.
Subisomorphism	11	**7**	6	2	3	3	conf./sup.
Total	236	115	116	138	140	**146**	

Table 2 shows the number of instances solved. We picked up the top five solvers and all figures are organized by 19 instance series. The second column "#Ins." denotes the number of instances included in each series. The last column "Ctr." denotes types of constraints: "pure int" means instances consist of only intensional and global constraints, "sup." means instances containing extensional constraints of supports, "conf." means instances containing extensional constraints of conflicts. According to the results, sCOP is particularly better than others in Pseudo-Boolean series and Quasigroups. Oppositely, it is worse than others in Subisomorphism.

In CSP-Standard-Parallel track, four solvers are registered. sCOP using the order encoding is also better than other solvers.

All information contains, i) how instances are selected, ii) descriptions of instance series and solvers, are available in the competition proceedings [Lecoutre 18].

4. Conclusion

This paper describes the under development SAT-based CSP solver sCOP and its results during the 2018 XCSP3 Competition. The sCOP solver is written in Scala and currently uses the order and log encodings to solve XCSP3 instances. The results of 2018 XCSP3 Competition showed that sCOP is superior to the other state-of-the-art XCSP3 solvers in terms of the number of solved instances within the given time limit. Future work is as follows. Adapting COP is important. To improve performance, implementing more SAT encodings and their hybridization are necessary. To enhance usability, supporting other constraint languages such as MiniZinc or Sugar's language are also important future work.

References

[Boussemart 17] Boussemart, F., Lecoutre, C., Audemard, G., and Piette, C.: XCSP3 An Integrated Format for Benchmarking Combinatorial Constrained Problems: XCSP3 Specifications—Version 3.0.5, `http://xcsp.org/format3.pdf` (2017)

[Eén 06] Eén, N. and Sörensson, N.: Translating Pseudo-Boolean Constraints into SAT, *Journal on Satisfiability, Boolean Modeling and Computation*, Vol. 2, No. 1-4, pp. 1–26 (2006)

[Lecoutre 18] Lecoutre, C. and Roussel, O.: XCSP3 Competition 2018 Proceedings, `https://www.cril.univ-artois.fr/~lecoutre/papers/XCSP3_2018_Proceedings.pdf` (2018)

[Soh 17] Soh, T., Banbara, M., and Tamura, N.: Proposal and Evaluation of Hybrid Encoding of CSP to SAT Integrating Order and Log Encodings, *International Journal on Artificial Intelligence Tools*, Vol. 26, No. 1, pp. 1–29 (2017)

[Soh 18] Soh, T., Berre, D. L., Banbara, M., and Tamura, N.: sCOP: SAT-based Constraint Programming System, in *XCSP3 Competition 2018 Proceedings*, pp. 93–94 (2018)

[Tamura 08] Tamura, N. and Banbara, M.: Sugar: a CSP to SAT Translator Based on Order Encoding, in *Proceedings of the 2nd International CSP Solver Competition*, pp. 65–69 (2008)

[Tamura 09] Tamura, N., Taga, A., Kitagawa, S., and Banbara, M.: Compiling Finite Linear CSP into SAT, *Constraints*, Vol. 14, No. 2, pp. 254–272 (2009)

[Tamura 13] Tamura, N., Banbara, M., and Soh, T.: PBSugar: Compiling Pseudo-Boolean Constraints to SAT with Order Encoding, in *Proceedings of the 25th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2013)*, pp. 1020–1027 (2013)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

非負値多重行列因子分解の因子行列を用いたクラスタリングと決定木学習によるオフィスの入退室データの分析

An Analysis of Entry and Exit Data in Office by Decision Tree Learning Using Clustering Factor Matrix from Non-negative Multiple Matrix Factorization

小島世大 Seidai Kojima　石榑隼人 Hayato Ishigure　坂田美和 Miwa Sakata　武藤敦子 Atsuko Mutoh　森山甲一 Koichi Moriyama　犬塚信博 Nobuhiro Inuzuka

名古屋工業大学 大学院 工学研究科 情報工学専攻
Department of Computer Science, Graduate School of Engineering, Nagoya Institute of Technology

Recently, IC card systems are popular and their log data are used for analyzing human behaviors. In this paper, we extract user behavior patterns using Non-negative Multiple Matrix Factorization (NMMF) and propose an analysis method to analyze patterns and attribute information by decision tree learning using clustering factor matrix. We examine our proposed method using actual entry and exit data and confirm the effect.

1. はじめに

近年 IC カードの普及に伴い、IC カード利用履歴を用いた人の行動分析の研究が増えている [1][2]。さらに、電子錠と IC カードにより、人の入退を制御する入退室管理システムの導入が増えている。入退室管理システムは、電子錠に備えられたカードリーダにかざされた IC カードの情報を読み取り通行履歴として記録する。現在では、企業は単純な入退室制御だけではなく、社員や部門毎の入退室傾向の分析や社員がどの部屋にいるのかを把握したいなどの要望がある [3]。また、近年、センサノード及びネットワークの技術が急速に発展している。環境に設置されたセンサノードと人が身に着けるウェアラブルセンサーの併用は、企業におけるオフィス内のワーカーの行動のモニタリングへ利用されている。センサデータを用いたオフィスワーカーの行動分析 [4] もされているが、これには特別な装置を必要とする。これまでに我々はクラスタリングによる入退室データからの移動時間パターンの分析をしてきた [5][6]。しかし、移動時間パターンの中で解釈が困難なものが存在した。そこで本研究では、非負値多重行列因子分解（Non-negative Multiple Matrix Factorization,NMMF)[7] を用いてユーザの行動パターンを抽出し、行動パターンと属性情報の関係を分析する手法を提案する。ＮＭＭＦを用いることで移動時間の特徴量だけではなく、部屋間の組に関する特徴量の行列を加えた社員の行動パターンの分析が可能となる。最後に提案手法を用いて多くの組織で普及が進んでいる入退室管理システムから得られる入退室データの分析を行う。

2. 関連研究

IC カードの利用者数の増加により、IC カード利用履歴データが人の行動分析手段として注目を集めている。 鈴木ら [2] は交通 IC カードの利用履歴を用いて人が駅の改札口を出て、次の改札口に入るまでの間にその人の滞在目的があるとし、生活パターンを定量的に表す生活行動属性を提案した。 カード利用者の生活圏 (駅) によらず似た生活パターンを持つ人を抽出しマーケティング等へ活用し得ると示した。また、嶋本ら [1] は英国・ロンドンで導入されている Oyster Card の 4 週間分の利用履歴データを用いて公共交通の変動を把握した。料金支払い形態に応じて利用者の利用属性を分類することで，1 人あたりの利用回数の変動の 4 割以上を説明できることを示した。

幸島ら [8] は消費者行動パターン抽出の為に、従来の NMMF にユーザや商品などの属性情報に関する入力行列を加えた、属性情報を考慮した非負値多重行列因子分解法（Non-negative Micro Macro Mixed Matrix Factorization,NM4F）の提案をした。提案手法と従来の NMMF の定量評価や因子行列をクラスタに分けて提案手法の定性評価を行っていた。

古川ら [9] はインターネットの YouTube にアップロードされた国内外のチェリスト 46 名に対して、著者が目視により様々な属性についてのデータを獲得し、クラスタリングと決定木分析などのデータ分析の手法を用いてチェリストの分類を試みた。

本研究では NMMF により抽出されたパターンと属性情報の関係を分析できる方法を提案する。本提案はＮＭＭＦにより得られた因子行列をクラスタに分けて属性の分析をする点で [8] と似ているが、クラスタリング結果を決定木学習に用いる点で異なる。また、クラスタリングと決定木分析を用いる点で [9] と似ているが、クラスタリングの特徴量ベクトルに因子行列を用いる点で異なる。

3. 提案手法

NMMF の因子分解結果を用いたクラスタリングと決定木学習によるユーザの頻出パターンと属性情報の関係を分析する手法を提案する。提案手法の流れを図 1 に示す。ＮＭＭＦとは複数の入力行列を同時に分解し、頻出パターンを同時に抽出する方法である。本提案の想定する入力行列は購買履歴データや入退室データなどから得られるユーザの購買パターンや行動パターンなどを表現する複数の特徴行列である。購買データであればユーザ×商品の購買回数の行列とユーザ×場所の訪問回数のデータである。まず NMMF により複数の入力行列から頻出パターンを抽出する。ユーザの属性情報との関係を分析する為に、因子分解により得られた因子行列の各行を特徴量ベクトルとし、クラスタリング手法を用いてユーザのグループ分けを行う。クラスタリングにより得られたグループ分け結果を元に、説明変数を属性情報とし、目的変数をクラスタ（グループ）に属するユーザを 1、その他のユーザを 0 とし、各クラスタ（グループ）に対して決定木学習を用いて分類木を作成する。各クラスタ中心をクラスタ（グループ）に含まれるユーザの特徴とし、因子行列と分類木を照らし合わせることで、頻出パターンと属性情報の関係を分析する。

連絡先: 小島世大, 名古屋工業大学 大学院, 〒 466-8555 愛知県名古屋市昭和区御器所町, s.kojima.571@nitech.jp

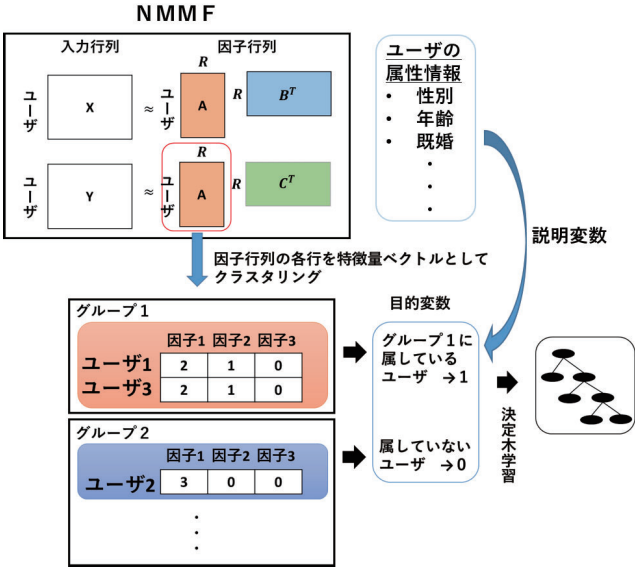


図 1: 提案手法の流れ

4. 提案手法を用いた入退室データの分析

4.1 入退室データ

本研究では (打刻日、打刻時刻、場所、操作、社員 ID) の 5 つの属性で構成された入退室データを扱う。操作には、部屋を退室もしくは入室したかが記録されている。移動時間に着目するために、退室と入室がセットになっているものを結合し、(移動開始日、移動開始時刻、移動時間、移動元、移動先、社員 ID) の 6 つの属性で構成されたデータ構造（以後移動データと呼ぶ）に変換する。移動時間は（入室時刻 − 退室時刻）で計算する。分析には協力企業の 2016 年 6 月の 332 名の移動データの内、移動時間が 90 分以内の移動のみを用いた。

4.2 提案手法に用いる入力行列の算出

社員ごとの移動時間に関する特徴量を表す行列（社員×移動時間）と場所に着目した部屋間の組に関する特徴量を表す行列（社員×部屋間の組）の二つの特徴行列を算出をする。

4.2.1 移動時間特徴行列 X

移動時間に関するの特徴量を表す行列 X（社員×時間区分）を算出する。社員 i の移動時間特徴行列の要素 X_{ij} は時間区分 j（時間帯及び移動時間の区分）の移動回数の割合とする [5]。分析に用いる時間区分を 9 種類 (時間帯 (午前・昼休憩・午後の 3 次元) ×移動時間 (0-5 分・5-20 分・20-90 分の 3 次元)) で表した（表 1 左）。

4.2.2 部屋間の組特徴行列 Y

部屋間の組に関する移動の特徴行列 Y（社員×部屋間の組）を算出する。社員 i の部屋間の組特徴行列の要素 Y_{ik} は部屋間の組（ex, A 室から B 室 →(A,B)）の移動回数の割合とする。例えば、ある社員が A 室から B 室の移動を 10 回しており、ある期間の総移動回数が 10 回の場合、部屋間の組特徴行列の要素の値は 1.0 となる。入力行列 Y の部屋間の組は 6 種類（a:移動前と移動後同じ部屋、b:移動前と移動後同じビル、同じ階で違う部屋、c:移動前と移動後同じビルで違う階、d:第 1 ビルと第 2 ビルの間の移動、e:第 1 ビルと第 3 ビルの間の移動、f:第 2 ビルと第 3 ビルの間の移動）で表した（表 1 右）。

表 1: 時間区分と部屋間の組

時間区分			部屋間の組	
記号	時間帯	移動時間	記号	
11	午前	0–5 分	a	移動前と移動後同じ部屋
12	午前	5–20 分	b	移動前と移動後同じビル、
13	午前	20–90 分		同じ階で違う部屋
21	昼休憩	0–5 分	c	移動前と移動後同じビルで
22	昼休憩	5–20 分		違う階
23	昼休憩	20–90 分	d	第 1 ビルと第 2 ビルの間の移動
31	午後	0–5 分	e	第 1 ビルと第 3 ビルの間の移動
32	午後	5–20 分	f	第 2 ビルと第 3 ビルの間の移動
33	午後	20–90 分		

4.3 NMMF によるパターン抽出

分析に用いる NMMF は [7] を参考にした。本節では入退室データを NMMF で因子分解した行列分解形、分析に用いる NMMF、因子行列の初期値の設定、因子数の決定方法と最後に抽出した因子行列について説明をする。

4.3.1 行列分解形

4.2 節で算出した入力行列 X、Y に NMMF を用いると図 2 のように因子分解できる。因子行列 A、B の積が入力行列 X

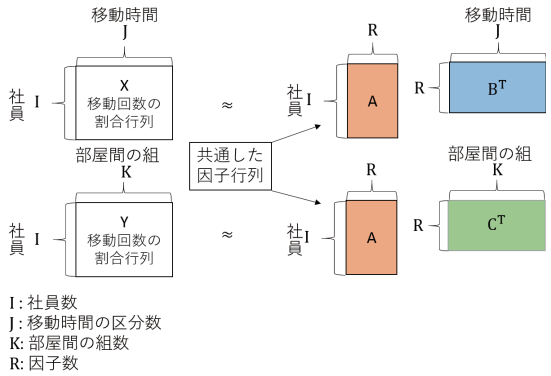


図 2: 行列分解形

の分解結果であり、因子行列 A と C の積が入力行列 Y の分解結果である。因子行列 A は各社員がどの因子にどの程度基づくのかを示す。因子行列 B は各因子の代表的な割合の多い移動時間、因子行列 C が各因子の代表的な割合の多い部屋間の組である。

4.3.2 定式化

図 2 の近似 ≈ の尺度には、ユークリッド距離を利用した。

$$d_{EU}(x_{ij}||\hat{x}_{ij}) = \frac{1}{2}(x_{ij} - \hat{x}_{ij})^2. \quad (1)$$

d は行列の要素同士の距離であり、行列同士の距離 $\mathcal{D}$ は式 (2) のように表せる。

$$\mathcal{D}_{EU}(X||\hat{X}) = \sum_{i,j=1}^{I,J} d_{EU}(x_{ij}||\hat{x}_{ij}). \quad (2)$$

ここで、入力行列 X と因子行列の積 $\hat{X}$ の距離を $\mathcal{D}(X||\hat{X})$（入力行列 Y も同様 $\mathcal{D}(Y||\hat{Y})$）とし、NMMF は式 (3) に示す最適化問題を解くことで因子行列を出力する。因子行列を更新していく上で入力行列と因子行列の積の誤差（距離）を損失関数の値とする。

$$\arg\min_{A,B,C}\{\mathcal{D}_{EU}(X||\hat{X}) + \mathcal{D}_{EU}(Y||\hat{Y})\} \quad s.t. A, B, C \geq 0. \quad (3)$$

この最適化問題を解くアルゴリズムは複数存在するが、ここでは実装上の簡易さから利用されることの多い式 (4)-(6) の乗法更新則に基づくアルゴリズムを利用した。

$$a_{ir} \leftarrow a_{ir} \frac{\sum_{j=1}^{J} x_{ij} b_{jr} + \sum_{k=1}^{K} y_{ik} c_{kr}}{\sum_{j=1}^{J} \hat{x}_{ij} b_{jr} + \sum_{k=1}^{K} \hat{y}_{ik} c_{kr}}. \tag{4}$$

$$b_{jr} \leftarrow b_{jr} \frac{\sum_{i=1}^{I} x_{ij} a_{ir}}{\sum_{i=1}^{I} \hat{x}_{ij} a_{ir}}. \tag{5}$$

$$c_{kr} \leftarrow c_{kr} \frac{\sum_{i=1}^{I} y_{ik} a_{ir}}{\sum_{i=1}^{I} \hat{y}_{ik} a_{ir}}. \tag{6}$$

4.3.3 因子行列の初期値の設定

入力行列 X,Y は要素が割合で表してあることからの各行の合計 1 になることが分かる。そこで、本論文では更新をスムーズに進めるために、因子行列の初期値を式 (7)-(9) に示すように設定した。

$$A \sim Uniform\left\{0, \frac{1}{\sqrt{J*R}} + \frac{1}{\sqrt{K*R}}\right\}. \tag{7}$$

$$B \sim Uniform\left\{0, \frac{2}{\sqrt{J*R}}\right\}. \tag{8}$$

$$C \sim Uniform\left\{0, \frac{2}{\sqrt{K*R}}\right\}. \tag{9}$$

この初期値は入力行列の各行の合計が 1 になることを考慮し、因子行列同士を掛け合わした行列の各行の合計が 1 に近くなるように設定している。これによって、入力行列に近い初期値を与えることが可能となる。

4.3.4 因子数の決定方法

因子行列は因子数を増やすほど損失関数の値は減少していくため、損失関数の値の減少幅が明らかに小さくなった場合、そこを最適な因子数とする [10]。入力行列 X, Y を因子分解した結果、因子数は損失関数の値の減少幅が明らかに下がった因子数 9 とする。因子分解により得られた 3 つの因子行列 A,B,C のヒートマップを図 3 に示す。色の濃いマス程値が大きく、色の薄いマス程値が小さくなるように表示している。

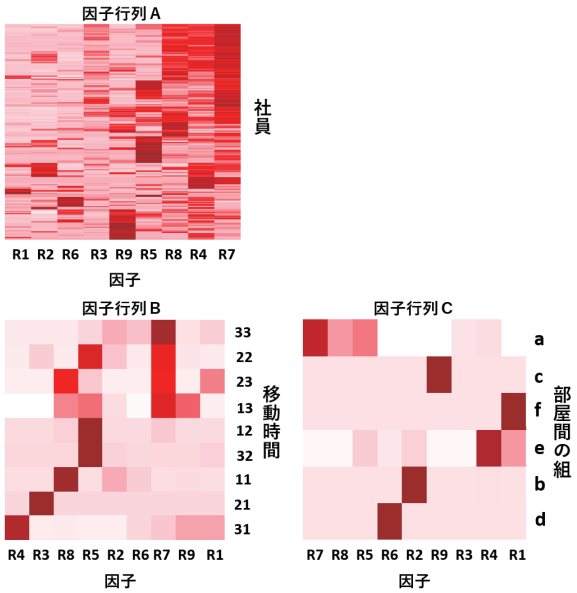


図 3: NMMF による因子分解の結果

表 2: クラスタ（グループ）内に含まれる社員数

クラスタ番号	1	2	3	4	5	6	7	8	9	10	11
社員数	26	14	8	13	11	5	46	29	21	51	108

4.4 因子行列によるクラスタリング

社員がどの因子にどの程度基づくのかを示す因子行列 A(図 3）の各行を社員の特徴量ベクトルとし、クラスタリングを用いた社員のグループ分けを行う。クラスタリングには k-means 法を用いて、クラスタ数は elbow 法を用いて決定した。結果クラスタ数 11 で社員のクラスタリングを行った。クラスタ中心のヒートマップを図 4 に、クラスタに含まれる社員数を表 2 に示す。

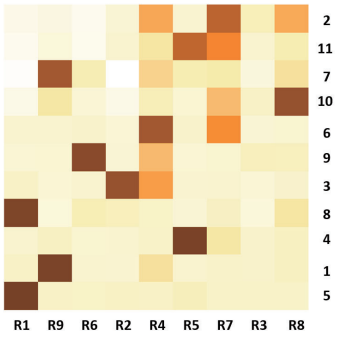


図 4: 因子行列 A をクラスタリングした結果のクラスタ中心のヒートマップ、行がクラスタ番号、列が因子

4.5 決定木学習を用いた分類木の作成

次に、クラスリング結果を用いて決定木学習を行う。決定木学習に用いる説明変数を社員属性（性別、部署、職種、採用種別、年代、社歴）とし、最大の階層を 3 とした。全クラスタの分類木の内、クラスタ人数の最も多いクラスタ 11 の結果を図 5 に示す。

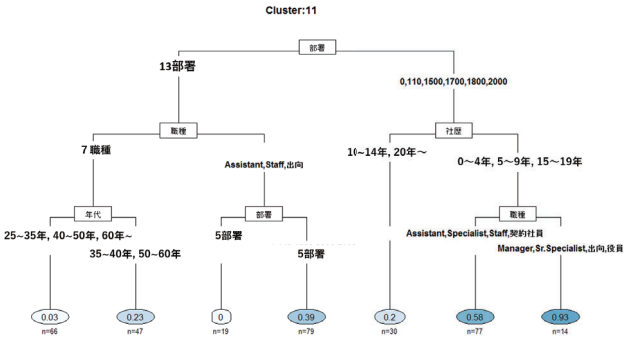


図 5: クラスタ 11 の分類木

4.6 考察

各クラスタがどの因子に当てはまるのかを図 4 及び因子行列 B と C（図 3）から特徴を見て、決定木学習により得られた分類ルールの内、カバー率の高いものから社員属性とパターンの関係を考察していく。特に社員属性の影響によりパターンが分かれていると考えられる結果のクラスタ 3、4、11 について考察する。クラスタ 3 とクラスタ 11 は社員属性の影響によりパターンが異なると考えられる結果である。またクラスタ 4 とクラスタ 11 は最も当てはまる因子は同じであるが、クラスタ 11 はさらに少し当てはまる因子を持っているため、クラスタ 4 との比較を行う。

- クラスタ 3 の 8 名の社員は主に因子 2 に当てはまり、因子 4 も少し当てはまっていた。因子 2 は、移動時間の割合が均一であることから移動時間がばらばらな因子であると読み取れる。また、移動には同じ階の違う部屋へ移動することが多かった。分類木はある一つの部署で職種が Manager、Sr. Specialist、Staff である 15 名が 27 %の確率でクラスタ 3 に分類されていた。以上から、この部署の役職の高い社員 (Manager、Sr. Specialist) と Staff は基本的には同じ階の会議室とオフィス（オフィス以外は会議室が多い）を往復することが多く、移動の途中でトイレ休憩や小休憩などを挟みながら仕事をしていると推測できる。

- クラスタ 4 の 13 名の社員は主に因子 5 に当てはまっていた。同様にクラスタ 11 も因子 5 に当てはまるが、クラスタ 4 とは違い因子 7 も少し当てはまっていた。因子 5 は午前は比較的に短い移動をしており、昼休憩と午後は 5–20 分の移動が多く、同じ部屋へ戻ることが多かった。分類木は、ある 3 つの部署で職種が Assistant、Manager、Specialist の年代が 25–30 歳、35–40 歳,、50–55 歳の 10 名が 40 %の確率でクラスタ 4 に分類されていた。以上から、この三つの部署の 25–30 歳、35–40 歳、50–55 歳の役職の高い社員（Manager、Specialist）はクラスタ 3 の役職の高い社員とは異なり、基本は自分のオフィスで仕事をしており、午前はトイレ休憩など移動に時間をかけていない。しかし、午後になるに連れ、小休憩をとっている（少し長めの移動（5-20 分）をしていることから）と推測できる。

- クラスタ 11 の 108 名の社員もクラスタ 4 と同様に因子 5 に当てはまり、因子 7 も少し当てはまっていた。因子 5 はクラスタ 4 にて示した通りの特徴を持つ。因子 7 は午前と午後に 20-90 分の移動をする割合が多く、昼休憩では比較的に移動に時間をかけており、同じ部屋に戻ってくる移動が多かった。分類木は、ある 6 つの部署で社歴が–9 年もしくは 15–19 年の職種が Manager、Sr. Specialist、出向、役員の 14 名が 93 %の確率でクラスタ 11 に分類されていた。以上から、これらの部署の役職の高い社員（Manager、Sr. Specialist、役員）もクラスタ 4 の社員と同様な仕事のスタイルであると考えられるが、因子 7 の特徴も持っていることから、たまに社外で会議があったり、昼休憩にランチをとり行くことがあると推測できる。

以上より、役職が同じ場合でも部署で行動パターンが異なることや、クラスタ中心を各クラスタの特徴ととらえることでクラスタ 4 とクラスタ 11 の比較のように、一番当てはまる因子は同じ場合でも、少し異なる特徴を持っていることを発見できることがわかった。その他にも、ビル間の移動に時間をかけない社員など様々な社員属性とパターンの関係がみられた。

5. まとめと今後の方針

本研究では NMMF により抽出されたパターンと属性情報の関係を分析する手法を提案した。提案手法を用いてオフィスの入退室データを用いて社員の移動時間と部屋間の組に関する行動パターンと社員属性の分析を行ったところ、入退室データの分析により、役職が同じ場合でも部署で行動パターンが異なるなど、提案手法を用いて属性情報とＮＭＭＦにより得られたパターンの関係を分析することの有効性を確認した。しかし、クラスタ数の決定方法に elbow 法を用いたが、分析結果の中で同じ特徴を持ったクラスタがあった為、最適なクラスタ数とは言い切れないと考える。

今後の方針としては、クラスタ数の決定方法の見直しや決定木学習の最大の階層の指定方法などの課題が挙げられる。

謝辞

本研究を進めるにあたり、入退室データを提供して頂いた協力企業に感謝の意を表する。本研究は JSPS 科研費 JP18K18160 の助成を受けたものです。

参考文献

[1] 嶋本寛, 北脇徹, 宇野伸宏, 中村俊之, “IC カード利用履歴データを用いた公共交通需要変動分析”, 土木学会論文集 D3(土木計画学), Vol.70, NO.5(土木計画学研究・論文集第 31 巻) , pp.605-610, 2014.

[2] 鈴木敬, 相薗敏子, “交通 IC カード利用履歴を用いた生活行動属性指標の提案”, 信学技報, IEICE, Technical Report, LOIS2011-84, 2012.

[3] 佐藤雅之, 及川和彦, 永嶋規充,“入退室管理システムにおける通行履歴の応用”, 情報処理学会第 77 回全国大会, 2G-05.

[4] 岡田将吾, 神谷祐樹, 佐藤祐作, 藤田義弘, 山田敬嗣, 新田克己,“センサ環境を利用したオフィスワーカーの行動パターン分析”, 第 27 回人工知能全国大会, 1C4-1in, 2013.

[5] Seidai Kojima, Hayato Ishigure, Miwa Sakata, Atsuko Mutoh, Koichi Moriyama and Nobuhiro Inuzuka,“An Analysis Method of Traveling-Time Patterns Between Rooms from Entry and Exit Data of Office Workers”, 2018 IEEE 7th Global Conference on Consumer Electronics(GCCE2018), pp.267-270.

[6] 小島世大, 石榑隼人, 坂田美和, 武藤敦子, 森山甲一, 犬塚信博,“オフィスワーカーの入退室データを用いた移動時間パターンの分析”, 第 32 回人工知能全国大会, 302-OS-1b-01, 2018.

[7] 幸島匡宏, 松林達史, 澤田宏 “複合データ分析技術と NTF［1］——複合データ分析技術とその発展——”, 電子情報通信学会誌 Vol. 99, No. 6, 2016.

[8] 幸島匡宏, 松林達志, 澤田宏 “属性情報を考慮した消費者行動パターン抽出のための非負値多重行列因子分解法”, 人工知能学会論文誌 30 巻 6 号 SPI-G, 2015.

[9] 古川康一, 升田俊樹, 西山武繁 “チェロ演奏動画の目視によるデータ獲得と演奏スタイルの分類”, 第 30 回人工知能学会, 1M4-OS-14a-3, 2016.

[10] 安川武彦 “非負値行列因子分解を用いたテキストデータ解析”, 計算機統計学, 第 28 巻・第 1 号: 2015, pp.41-55.

Organized Session | Organized Session | [OS] OS-3

[1E3-OS-3b] AI における離散構造処理と制約充足(2)

波多野 大督（理化学研究所）、蓑田 玲緒奈（（株）ベイシスコンサルティング）

Tue. Jun 4, 2019 3:20 PM - 4:00 PM Room E (301A Medium meeting room)

[1E3-OS-3b-01] A Branch-and-bound Algorithm for Determining Discrete Changes for Hybrid Constraint Solver HyLaGI

〇Masashi SATO[1], Kazunori UEDA[1] （1. Waseda University）

3:20 PM - 3:40 PM

[1E3-OS-3b-02] Dynamic Reduction of Guarded Constraints for the Hybrid Systems Modeling Language HydLa

〇Takafumi Horiuchi[1], Kazunori Ueda[1] （1. Waseda University）

3:40 PM - 4:00 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

ハイブリッド制約処理系 HyLaGI における 分枝限定法を用いた離散変化時刻導出手法

A Branch-and-bound Algorithm for Determining Discrete Changes for Hybrid Constraint Solver HyLaGI

佐藤柾史　上田和紀
Masashi SATO　Kazunori UEDA

*1 早稲田大学大学院基幹理工学研究科情報理工・情報通信専攻
Graduate School of Fundamental Science and Engineering, Waseda University

*2 早稲田大学理工学術院情報理工学科
Faculty of Science and Engineering, Waseda University

Hybrid systems are dynamical systems involving continuous and discrete changes. Various models such as cyber-physical systems and control systems can be discribed as hybrid systems. We are developing HydLa, a modeling language of hybrid system and its symbolic simulator HyLaGI. HyLaGI performs exhaustive search to find time points of discrete changes, but its computation becomes a bottleneck for some programs having a large number of guarded constraints. In this research, we propose a efficient search method using a branch-and-bound algorithm and implement a prototype. The result of simple experiment shows that our approach reduces most of the search cost in the motivating example.

1. はじめに

ハイブリッドシステム [4] は時間進行に伴って連続変化と離散変化を繰り返す系であり, 物理系やサイバーフィジカルシステム等に広く適用できる概念である.

ハイブリッドシステムのシミュレーションとそれによる検証のための処理系として, HyLaGI[2] がある. HyLaGI は, HydLa[1] で記述されたプログラムを入力として, 不確定性をはらんだハイブリッドシステムの記号実行を実現する. HydLa では, ハイブリッドシステムを制約 (時相論理式や等式/不等式, 微分方程式) を用いてモデリングする. HyLaGI は記述された制約を HydLa の意味論に従って解くが, 記号を残した誤差の無いシミュレーションには様々な困難がある.

本研究は, HyLaGI における離散変化の要因が多数ある問題のシミュレーションの高速化を目的とする. 現在の HyLaGI では, 離散変化発生の原因をしらみつぶしに探索するため, その数に比例して実行時間が増大することが分かっている. 制約で記述された多数の離散変化条件を包含する緩和問題を作成し, 分枝限定法を用いた効率的な解探索をする手法の提案とプロトタイプ実装を行い, 目的とする例題での実行時間の抑制を実現した.

2. HyLaGI における離散変化時刻導出手続き

HydLa における離散変化は, 論理包含で記述された制約の前件 (ガード条件と呼ぶ) が成立するようになる, または成立しなくなることで発生する. これは, 時間の関数である変数の式とそれを参照 *1 するガード条件の元で時刻を最小化する連続最適化問題に帰着できる. HyLaGI では, 記述された各ガード条件についての最小化問題の作成と Mathematica の組み込み関数である Minimize を用いた求解を行うことで, 離散変化時刻を導出している. この手続きは軽いものではなく, ガード条件が多数ある場合はしらみつぶしに探索するため, 実行における明確なボトルネックとなることがある.

連絡先: 佐藤柾史 早稲田大学基幹理工学研究科情報理工・情報通信専攻, 〒 169-8555 新宿区大久保 3-4-1 63 号館 5 階 02 号，03-5286-3340，masashi(at)ueda.info.waseda.ac.jp

*1 並行制約プログラミングでは ask と呼ばれる

```
 1 // #define N 6
 2
 3 INIT_X(v) <=> (x = 0 & x'=v).
 4 X_MOVE <=> [](x'' = 0).
 5 INIT_Y(h) <=> (y = h & y' = 0).
 6 FALL(g) <=> [](y'' = -g).
 7 WALL(w) <=> [](x- = w & 0<=y-<=2*N => x' = -x'-).
 8
 9 BOUNCE_ON_STEP_HOR(cornerx) <=>
10   []( (y- = N - cornerx-) & (cornerx- <= x- < cornerx- + 1)
11     => (y' = -9/10 * y'-) ).
12 BOUNCE_ON_STEP_VER(cornerx) <=>
13   []( (x- = cornerx-) & (N - cornerx- < y- <= N - cornerx- + 1)
14     => (x' = -x'-) ).
15 BOUNCE_HOR := { BOUNCE_ON_STEP_HOR(i) | i in {0..N} }.
16 BOUNCE_VER := { BOUNCE_ON_STEP_VER(i) | i in {0..N} }.
17
18 INIT_X(1), INIT_Y(N + 3),
19 (FALL(9.8) << BOUNCE_HOR),
20 (X_MOVE << BOUNCE_VER << WALL(0)).
```

図 1: 階段を跳ねる質点の HydLa プログラム

2.1 最適化の対象とする例題

図 1 は, 階段を跳ねる質点をモデリングした HydLa プログラムである. HydLa プログラムの詳細は [1] を参照してほしい. 10, 13 行目が段の横, 縦を記述した制約である. 1 行目のマクロにより, 階段の段数を指定する.

図 2 は, このプログラムの階段の段数 N を変えながら HyLaGI で実行した際の階段の段数と実行時間の関係である. 質点が 10 回跳ねるまでにかかる時間を測定している. まず, 離散変化時刻の導出が全実行時間の殆どを占めていることが分かる. また, 離散変化時刻の導出にかかる時間は階段の段数に比例して増加することが分かる.

3. 緩和問題と分枝限定法

分枝限定法 [3] とは, ナップサック問題や巡回セールスマン問題といった離散最適化において, 効率的な枝刈りを行うための探索アルゴリズムである. 本章では, 分枝限定法の詳細を述

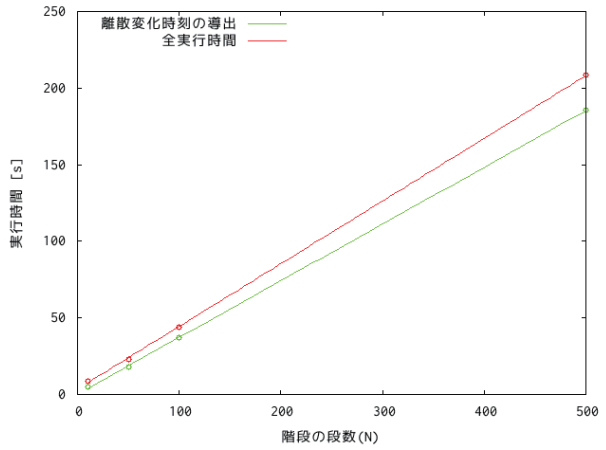


図 2: 階段を跳ねる質点のプログラムにおける段数と実行時間の関係

べる.

3.1 緩和問題

緩和問題とは, 最適化問題において目的関数を変えずに制約条件を緩めた問題のことである. 緩めた制約条件は, 元の問題の制約条件を包含する. 緩和問題について定理 1 が成立する. なお, 以降の記述において, 制約をそれを満たす集合で表す場合があることに注意してほしい.

定理 1. 制約条件 $\boldsymbol{x} \in X_p$ 下で $f(\boldsymbol{x})$ を最小化する問題 P と, 制約条件 $\boldsymbol{x} \in X_{rp}$ 下で $f(\boldsymbol{x})$ を最小化する問題 RP において, $X_p \subset X_{rp}$ が成立するならば, P の最適値 z_p* と RP の最適値 $z_{rp}*$ について, $z_p* \geq z_{rp}*$.

証明. $f(x_p*) = z_p*$ を満たす P の最適解 x_p* について, $x_p* \in X_p \subset X_{rp}$ より, x_p* は RP の実行可能解.
$\therefore z_p* \geq z_{rp}$ □

3.2 分枝限定法

分枝限定法では離散最適化問題 (最小化問題について説明) を, 部分問題への分割操作 (分枝) と部分問題の下界を計算して枝刈りを行う操作 (限定) を繰り返すことで最適値の探索を行う.

分枝操作では, $f(x)$ を $x \in X$ の元で最小化する問題において, X をいくつかの部分集合 $X_1, X_2, \ldots, X_n$ に分割して各々の元で $f(x)$ を最小化する問題に変換する. そして, 各部分問題の最適値を比較して, その中で最小の値, すなわち $min(f(x_1* \in X_1), f(x_2* \in X_2), \ldots, f(x_n* \in X_n))$ が元の問題の最適値となる.

分枝操作を再帰的に適用して元の問題の最適解を探索するのだが, 限定操作では, 各部分問題の最適値の下界 *2 を求める. 導出した下界となる値は, 以下のような探索の枝刈りに用いる.

1. 探索における暫定的な最適値 $z*$ が求まっている時, ある部分問題の最適値の下界 z_{min} が導出された. この時, $z* \leq z_{min}$ を満たすならば, その部分問題の最適値は暫定解を更新し得ないため探索を終了する.

2. ある部分問題 (制約は X) の最適値の下界 $f(x_{min}) = z_{min}$ が導出された. この時, $x_{min} \in X$ を満たすならば, その部分問題の解は x_{min} であるため, 解を保存して探索を終了する.

*2 必要ならば上界も求める

部分問題の最適値の下界は, 例えば部分問題を緩和問題に変換してそれを解くことで導出できる. 緩和問題の定理 1 より, 緩和問題の最適値は元の問題の下界となる. 探索の効率化という観点でいうと, 緩和問題の性質として, 元の問題と比べて十分に解きやすいことや下界が元の問題の解に近くなりやすい (あるいは一致しやすい) ことが要求される.

4. HyLaGI の離散変化時刻導出手続きにおける分枝限定法の適用

HydLa プログラムに出現する変数にかかる時間変化の等式制約 $S \Leftrightarrow x_1 = f_1(t) \wedge x_2 = f_2(t) \wedge \cdots \wedge x_n = f_n(t)$ とそれらを参照するガード条件の集合 $G = (g_1, g_2, \ldots, g_m)$ を用いて, HyLaGI の離散変化時刻の導出問題は以下の最小化問題として記述できる.

$$\text{Minimize } t \text{ such that } S \wedge (g_1 \vee g_2 \vee \cdots \vee g_m)$$

HyLaGI では論理和を分解して全てのガード条件についての最小化問題の求解と各々の最適値の比較を行っているが, この問題について分枝限定法を適用して離散変化時刻 (言い換えると, 原因となるガード条件) を効率よく探索することを考える.

4.1 分枝限定法を用いた離散変化時刻導出アルゴリズム

アルゴリズム 1 は, 分枝限定法を用いて離散変化時刻を導出する非決定アルゴリズムである. この関数は, ガード条件の集合と変数の時間変化の等式, 記号パラメータを受け取り, 記号パラメータに対応する離散変化時刻を返す. ここで, 記号パラメータは変数の持つ不確実性を保存しており, パラメータの取る値によって離散変化の原因が異なるかもしれないことを考慮する必要がある.

4.2 節で具体例を用いた説明を行うが, まずは一般的なアルゴリズムについて説明する. 15 行目の *CalculateRelaxedGuards* 関数は, ガード条件の集合を入力してそれらの論理和を論理包含する制約を返す. すなわち, アルゴリズム上の記号 $(g_1, g_2, \ldots, g_n) = G_{curt}$ と出力 $(h_1, h_2, \ldots, h_m) = H$ について, $\bigvee_{i=1}^{n} g_i \Rightarrow \bigvee_{j=1}^{m} h_j$ が成立する. この操作は分枝限定法の分枝, および限定操作のための緩和問題の作成に対応する. 多数のガード条件を幾つかのグループに振り分けて (分枝), 各々のグループに所属するガード条件を包含する新しい制約の作成を行う. 探索対象の多数のガード条件をひとまとめにすることが目的のため, n より m がずっと小さいことが望ましい. 16 行目のループにおいて, 緩和した制約下で t を最小化する問題 (部分問題) を解く. 17 行目の *FindMinTime* は作成した各々の緩和問題の求解を行う関数で, 部分問題の下界となる時刻を求める手続き (限定操作) に対応する. パラメータによる非決定性があるため, パラメタの取る値に対応した複数の解を返す場合がある. 19-20 行目の *CheckGuardSat* 関数は, 17 行目で導出した緩和問題の解が元の問題の実行可能解であるかの確認を行う. 具体的には, 部分問題に所属するガード条件を満たすかをチェックし, 満たされていたらその部分問題の最適解が求まることになる. この手続きも非決定性を伴うため, パラメタの取る値に対応してガードが満たされるか否かを返す. 7 行目, 10 行目は, 分枝限定法における枝刈りである. 7 行目は, 下界が暫定値を上回った部分問題の探索を打ち切る. 10 行目は, 19-20 行目で最適解が判明した場合に, パラメタに対応する最適値が見つかったことになり *3, 探索を打ち切る.

*3 最適値の下界でソートされた優先キューを用いた探索のため, 最初に求まった解がそのパラメタに対応する最も良い解である

アルゴリズム 1 分枝限定法を用いた離散変化時刻導出問題の非決定アルゴリズム

input:
　G : ガード条件
　S : 制約条件
　P : 記号定数の条件
output:
　最小離散変化時刻
　記号定数の条件

1: **function** FindMinTimeUsingBranchAndBound(G, S, P)
2: 　$PQ := PriorityQueue()$
3: 　$Sol := \{\}$
4: 　$PQ.push(\langle 0, G, P, false\rangle)$
5: 　**repeat**
6: 　　$\langle T_{curt}, G_{curt}, P_{curt}, tf_{curt}\rangle := PQ.pop()$
7: 　　**if** $P_{curt} = false$ **then**
8: 　　$continue$
9: 　　**end if**
10: 　　**if** $tf_{curt} = true$ **then**
11: 　　$Sol := Sol \cup \{\langle T_{curt}, P_{curt}\rangle\}$
12: 　　$PQ := \bigcup_{\langle T,g,p,tf\rangle \in PQ}\{\langle T, g, \neg P_{curt} \wedge p, tf\rangle\}$
13: 　　$continue$
14: 　　**end if**
15: 　　$H := CalculateRelaxedGuards(G_{curt})$
16: 　　**for** $h \in H$ **do**
17: 　　$MinResult := FindMinTime(Subst(h, S), P_{curt})$
18: 　　**for** $\langle T_{min}, p\rangle \in MinResult$ **do**
19: 　　$CGSResult :=$
20: 　　$CheckGuardSat(Subst(S, t = T_{min}), G_{curt}, p)$
21: 　　**for** $\langle tf, p\rangle \in CGSResult$ **do**
22: 　　$PQ.push($
23: 　　$\langle T_{min}, \{g | g \in G_{curt} \wedge g \Rightarrow h\}, p, tf\rangle)$
24: 　　**end for**
25: 　　**end for**
26: 　　**end for**
27: 　**until** $PQ = \{\}$
28: 　**return** $GetElement(Sol)$
29: **end function**

4.2 アルゴリズムの具体化

アルゴリズム 1 は, あらゆるガード条件の入力に対応した一般的なアルゴリズムである. 一般的に分枝限定法が有効に働くかどうかは分枝と限定の方法に依り, このアルゴリズムでは 15 行目の *CalculateRelaxedGuards* の実装に強く依存する. ここでは, ガード条件をどのようにグルーピングし, どのように緩和するかについて一切言及しておらず, 実装にあたっては具体化が必要である. HydLa の制約として記述できる (不) 等式のクラスに制限はなく [1], あらゆるガード条件の入力に対して有効に働く *CalculateRelaxedGuards* 関数を実装することは現実的ではない. そこで, 本研究では 2 次元の幾何問題を記述した HydLa プログラムを動機とし, それに対して有効に働くアルゴリズムの具体化を試みる. 以下に実装の方針を列挙する. アルゴリズムの詳細は省略する.

- 入力されるガード条件は, ある変数 x, y のどちらかまたは両方を参照した, 一次方程式/一次不等式に制限する.
- 分枝操作では, 変数の時間変化の式 $x = f(t)$ における $x = f(0)$ を超平面として x-y 座標平面を分割し, それぞれに包含されるガード条件を同じグループとする. ガード条件が境界面をまたぐ場合は, どちらかに所属させるか独立させる (後述).
- 制約の緩和においては, 同じグループに所属するガード条件を凸包で包含することで実現する.

以上の方針に従って, 図 1 のプログラムにおける離散変化時刻の導出をアルゴリズム 1 で実行した際の探索の様子を図 3 に示す. なお, プログラムの質点の初期位置は階段の左端の壁だが, 図解を分かりやすくするために, 数フェーズ経過して質点が階段の上にある状況から次に跳ねる段を探索する場合を考える.

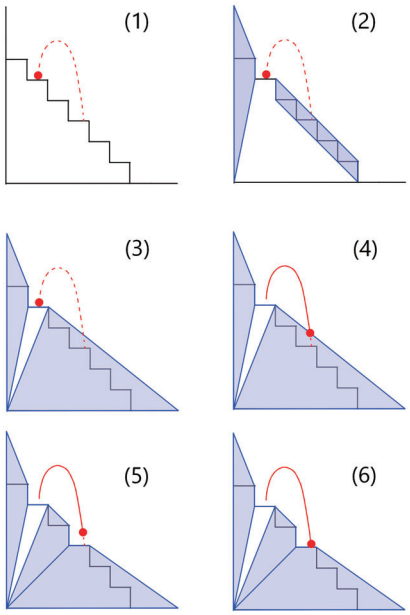


図 3: 階段を跳ねる質点のプログラムに提案アルゴリズムを適用した解探索の流れ

図 3 の (2)-(3) は, アルゴリズムの 15 行目の関数に対応する. まず, 質点の初期位置の x 座標の左右で空間を 2 分し, それらの領域に包含されるガード条件を同一のグループとして 2 分割する. 次に, それぞれのグループに属するガード条件を内包する凸包 (凸多角形) を作成する (2). ここで, 左右の空間にまたがるガード条件について考える. 図では, 床と質点が乗っている段を表す制約は, 左右の空間にまたがる. これらはどちらかの凸包にマージを試みるが, この際に凸包が質点の初期位置を包含しない場合に限りマージする. 凸包が包含する領域でガード条件を緩和するが, 質点の初期位置を包含してしまうと緩和問題の制約が最初から満たされることになり, 導出される時刻の下界が真の最適値と極端に離れてしまい役に立たなくなってしまうので, これを防ぐために行う. なお, マージの結果初期位置が凸包の境界になる場合も, 質点の移動する方向によっては同じ問題が発生するため, 失敗とする. また, マージに失敗した場合は独立した凸包にする. この場合, 床は右の凸包にマージできるが, 質点が乗る辺は独立させる. その結果, (3) のようになる. 左右の青い領域が示す凸包とマージに失敗した質点が乗るガード条件の, 3 つに対応する制約が関数の戻り値として返される.

	元の実装 [s]	提案手法 [s]
1 回目の離散変化時刻の探索 (N=100)	1.4072	0.15838
10 回目の離散変化時刻の探索 (N=100)	7.6225	0.808
1 回目の離散変化時刻の探索 (N=500)	6.5186	0.56365
10 回目の離散変化時刻の探索 (N=500)	38.597	1.2795

表 1: 提案手法を図 1 のプログラムに適用した実験結果

(4) は, アルゴリズムの 17-20 行目に対応する. それぞれの凸包が示す領域と変数の時間変化の式の元で時刻を最小化する問題を解く (17 行目). 右の凸包の元での最小化問題のみに最適値がある. 凸包は包含するガード条件を緩和した制約のため, この最適値は凸包が包含するガード条件の元で解いた最小化問題の最適値の下界となる. また, 導出された最適解が凸包が包含しているガード条件を満たすかを確認する (19-20 行目). (4) より, ガード条件を包含した際の近似誤差が原因で最適値が求まっているため, この最適値は元の問題の最適値と一致しない. そのため, 探索を続行する.

(5), (6) は, 今までの操作を右の凸包が包含するガード条件に対して再帰的に行っている. (6) では, 求めた最適解がガード条件に含まれることが図から分かり, 導出された下界が元の問題の最適値と一致する. 従って, 最適値を更新して探索を終了する.

5. 実装と例題を用いた実験

アルゴリズム 1 と, その 15 行目を具体化した節 4.2 のアルゴリズムを Mathematica でプロトタイプ実装をした. そして, 図 1 のプログラムにおいて離散変化時刻を特定するまでにかかる時間を元の HyLaGI の実装と比較した. 表 1 はその結果である.

表の各行はそれぞれ, 階段の段数が 100・500 段で 1・10 回目の離散変化時刻の導出にかかる時間を比較したものである. 階段の段数を比較したのは提案手法の実行にかかる時間と階段の段数の関係を確認するためである. また, ハイブリッドシステムの記号シミュレーションでは時間経過に伴って微分方程式の初期値が複雑化していく傾向があり, 跳ねるボールのように変数の時間変化の式が 2 次である場合もそれに当てはまる. 式の複雑化は離散変化時刻の導出にかかるコストにも強く影響するため, 離散変化の回数に対しても比較を行っている.

どの場合でも, 一桁の高速化が実現された. 階段の段数が増えたり, シミュレーションのフェーズが進行したりしたほうがコストが増加するのは同じであるが, その関係は元の実装ほど単純でない. そこで, アルゴリズム 1 の主要な関数に費やした累計時間の内訳を図 4 に示す. まず, 凸包作成の初期化処理 *4 では階段を表す制約のソートを行うため, 時間計算量は O(NlogN) である. また, 凸包は逐次添加法で作成するため, O(N) である. 図 4 のグラフは概ねそれに倣っている. 次にグラフから読み取れる重要な特性は, *FindMinTime* にかかる時間が階段の段数に依らない点である. 元の実装では, 最適化問題を作成して解く *FindMinTime* 関数が離散変化時刻導出における実行時間の内訳の全てであり, そのコストは段数に比例していたが, 分枝限定法が良く機能すれば *FindMinTime* の回数は階段の段数に対して定数オーダーで収めることができる. アルゴリズムが良く機能するか否かは, 不要な部分問題をいかに枝刈りできるかや, 最適解をいかに早く見つけて部分問題の探索を打ち切ることができるかに依る. ガード条件が一次でありそれらを凸包で包含する手法では, 凸包の辺とガード条件の境界が重なり, そこが最適解となった場合に探索を打ち切ることができる. 本実験では図 3 のように早々と最適解を求めることができるため, *FindMinTime* のコストは元の問題と比べて十分に小さく, そのために緩和問題の作成と求解のコストを加味しても高速化が実現できた.

*4 シミュレーションで一度行えば良い

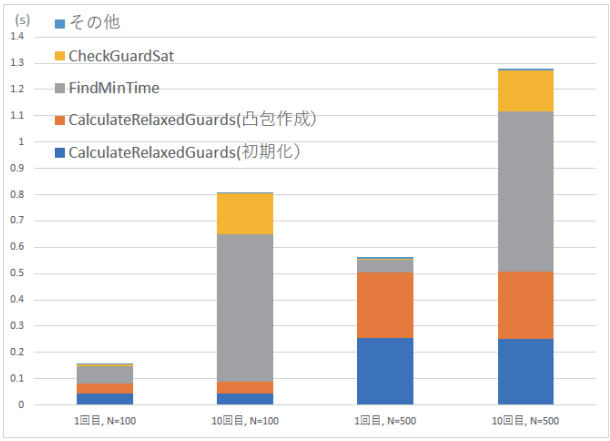


図 4: 実験結果の内訳

6. まとめと今後の課題

HyLaGI の離散変化時刻の導出手続きに分枝限定法を適用することで, 動機となる例題において一桁の高速化を実現した.

今後の課題としては, まず実験例題を増やして対象とする多くの例題で有効に機能するかどうかを確認する必要がある. モデルが変われば探索の様子も大きく変わるため, 一つの例題で有効性を主張することはできない. 加えて, 4.2 節の分枝の方針は単純であり, 部分問題の大きさが偏ってしまう場合がある. 探索木が偏り, 探索が葉まで到達する最悪のケースでは, ガード条件を全探索することになるだけでなく余計な緩和問題を解くことになるため, 元の実装より悪化することは明白である. より広いモデルに適用するためには, 部分問題の大きさをバランスさせる工夫が必要である.

参考文献

[1] 上田和紀, 石井大輔, 細部博史 : 制約概念に基づくハイブリッドシステムモデリング言語 HydLa, SSV2008(第 5 回システム検証の科学技術シンポジウム), 2008.

[2] Matsumoto, Shota et al. HyLaGI: Symbolic Implementation of a Hybrid Constraint Language HydLa. Electr. Notes Theor. Comput. Sci. 317 (2015): 109-115.

[3] Land, Ailsa H., and Alison G. Doig. An automatic method of solving discrete programming problems. Econometrica: Journal of the Econometric Society (1960): 497-520.

[4] Lunze, J : Handbook of Hybrid Systems Control : Theory, Tools, Applications, Cambridge University Press, 2009.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Dynamic Reduction of Guarded Constraints for the Hybrid Systems Modeling Language HydLa

Takafumi Horiuchi Kazunori Ueda

Department of Computer Science and Engineering, Waseda University

HydLa is a language for modeling hybrid systems—dynamical systems that intermix discrete and continuous behavior. Its adoption of a constraint-based framework benefits the language in various ways, such as allowing a concise representation of systems and performing error-free high precision simulations. In spite of all the advantages, the computations among sets of constraints become a bottleneck in simulation time when handling some large-scale models. The purpose of this research lies in providing a method of improving the computational efficiency and the scalability of the language and its simulator. This is achieved by considering the monotonic aspects in HydLa models to dynamically reduce the size of guarded constraints. Results show that this approach is effective for models that contain multiple objects represented by guard conditions. As for the model evaluated in an experiment in the research, the overall computational time has reduced to approximately half the original length.

1. Introduction

In recent years, there has been a rapid growth of interest in hybrid systems [Lunze] — systems that intermix discrete and continuous behavior. HydLa [Ueda et al.], which is the subject language of this research, is a modeling language for hybrid systems. A notable feature of HydLa is that it allows a concise representation of hybrid systems by directly handling high-level descriptions of mathematical and logical formulas as the source program. This feature is realized by the adoption of a constraint-based framework. In HydLa, constraints are the basic components that specify the behavior of models. These constraints are structured in the form of a constraint hierarchy [Borning et al.], among which the consistency is retained when processing the model.

The guaranteed-accuracy implementation of HydLa is called HyLaGI [Matsumoto], which takes a HydLa model as the input, simulates the model, and outputs the solution trajectory. HyLaGI has strong computational features such as performing error-free symbolic calculations and handling uncertain parameters. However, it is known that the computations on constraint sets become a bottleneck when the number of guarded constraints (i.e. constraints enabled when guard conditions are entailed) is large.

In this paper, we observe the relation between the size of guarded constraints and simulation time and propose a method for improving the efficiency of simulations by considering the monotonic behavior in HydLa models to dynamically reduce the number of guarded constraints.

2. Guarded Constraints in Simulations

In the current simulation algorithm of HyLaGI, the simulation time increases in relation to the size of guarded constraints that appears in the model. This becomes a severe bottleneck when processing large-scale models that contain a large number of objects represented by guarded constraints. The correlation between these two elements can be observed in a series of simulations of the model shown in Figure 1. This HydLa model represents a ball bouncing on a flat surface, where the surface is split into N pieces, each of which has the length of 1 unit. Constraints INIT_X, INIT_Y, and FALL define the initial and default behavior of the ball and BOUNCE describes the behavior of the ball when it collides with the surface at height 0. In the constraint declaration at line 9, FALL is assigned a weaker priority than BOUNCE, therefore FALL is temporarily unadopted when the two constraints conflict at the timing of the bounce. The pieces of the surface are generated in constraint BOUNCES with a list notation. For simplicity, the coefficient of restitution, initial horizontal velocity, and initial height of the ball are all set to 1. The model is simulated until time N, where the ball travels from the left end to the right end of the split surface.

```
1  // #define N 100
2
3  INIT_X(v)   <=> x=0 & [](x'=v).
4  INIT_Y(h)   <=> y=h & y'=0.
5  FALL        <=> [](y''= -9.8).
6  BOUNCE(l,r) <=> []((y- = 0) & (l <= x- < r)
7                  => y'= -1.0*y'-).
8  BOUNCES := {BOUNCE(i,i+1) | i in {0..N-1}}.
9  INIT_X(1), INIT_Y(1), (FALL << BOUNCES).
10
11 // #hylagi -t 100
```

Figure 1: Model of Bouncing Ball on Split Surface

As an experiment, the model is simulated multiple times, each time with different values of N ranging from 10 to 200. This setting will change the number of guarded constraints in the model, with the increase of one for each increment on the value of N. Other conditions are consistent throughout

Contact: Takafumi Horiuchi, Department of Computer Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo, 169-8555, Japan, Bldg.63, 5F-02, 03-5286-3340, horiuchi(at)ueda.info.waseda.ac.jp

the experiment. The result[*1] of the experiment is shown in Figure 4 (circle plots). These data indicate that the increase in the value of `N` leads to a longer simulation time. The regression line among the plotted data (original) is represented by the following equation:

$$SimulationTime = 1.1463\,N^2 + 1.5554\,N + 0.4949 \quad (1)$$

Thus, it can be inferred that, for this model, the simulation time increases in proportion to the square of the number of guarded constraints. In general, this behavior can be a bottleneck in the simulation of large-scale models that contain a large number of guarded constraints.

3. Location of the Bottleneck

To identify the location of the bottleneck in the current simulation algorithm, the distribution of time consumption among different computational units was measured. In a simulation of the split surface model where `N` was set to 100, the entire simulation time was 1.23×10^8 seconds. The most time consuming procedure was *FindMinTime*, a function that calculates the minimum satisfiable time of the next discrete change, which consumed 1.17×10^8 seconds (i.e. 96% of the entire simulation time). Taking this into account, the computation of *FindMinTime* is the location of the aforesaid bottleneck problem.

4. Reduction of Guarded Constraints

The computation of *FindMinTime* is performed by checking the satisfiability of conditions in antecedents of all the guarded constraints that appears in the model. This can be a drawback in terms of efficiency of simulations. For example, when considering a model of a ball bouncing down a staircase (Figure 2), the ball will never interact with the steps which it has already passed. In this case, it is unnecessary to check the satisfiability of guards that represent the steps behind the current position of the ball. The present simulation algorithm does not consider these situations, resulting in redundant computations when iterating through all of the guards during the computations of *FindMinTime*.

This inefficiency in the current simulation algorithm leads to the proposal of this research—dynamically reducing the number of guarded constraints. The aim of this proposal is to avoid redundant calculations when evaluating the entailment of guards and improving the computational efficiency of simulations.

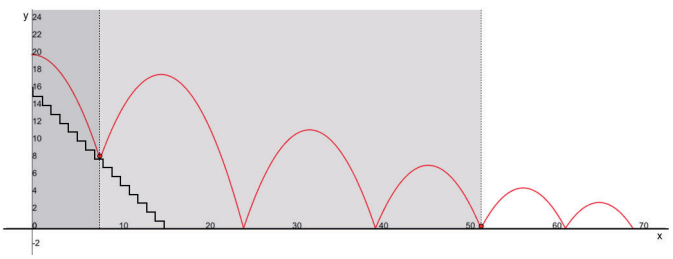


Figure 2: Concept of Omitting Verbose Guard Evaluations

The removal of guarded constraints is *safe only when it is certain that the guard condition will never be satisfied in the future.* Inadequate removal of guards will lead to wrong simulation results, which must be avoided. The decision of whether the guards are removable or not can be made by taking into account the concept of monotonicity, described by the following propositions:

$$\forall t \in (\alpha, \beta)\ \frac{d}{dt}x(t) \geq 0 \quad (2)$$

$$\forall t \in (\alpha, \beta)\ \frac{d}{dt}x(t) \leq 0 \quad (3)$$

Proposition (2) ensures the monotonic increase of variable x in the time interval (α, β) and proposition (3) ensures the monotonic decrease. The proposed method considers monotonicity of variables from two approaches, one dealing with uniform monotonicity and another with alternating monotonicity.

4.1 Approach to Uniform Monotonicity

The first approach considers the monotonicity of variables that are uniform throughout the simulation. This situation corresponds to the case where α and β are set to 0 and $MaxT$ (i.e. maximum simulation time), respectively, in propositions (2) and (3). For instance, this behavior is exhibited in the model of a ball bouncing on split surface, where variable `x` keeps increasing from the beginning to the end of simulation. The conceptual diagram of this approach is illustrated in Figure 3, representing the case where a variable is monotonically increasing throughout the simulation.

In order to apply this method for the dynamic reduction of guarded constraints, the truth/falsity of the uniform monotonicity must be evaluated, which can be achieved by using model checking techniques. After evaluating the uniformly monotonic behavior of variables, that information is used to determine the removable guarded constraints.

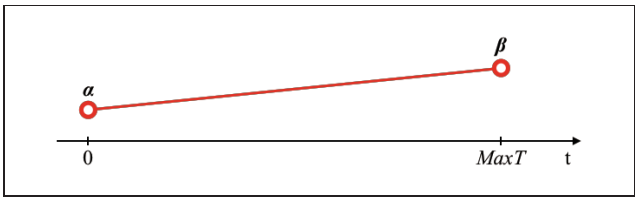


Figure 3: Concept of Approach to Uniform Monotonicity

4.2 Approach to Alternating Monotonicity

The types of model that the above-mentioned uniform approach can make use of is limited to cases where monotonic aspects are preserved from the beginning to the end of simulations. Variables that alternate its monotonic behavior during simulations would not be covered in this approach. To enhance the uniform approach and to enable the handling of alternating monotonicity of variables, the method can be enhanced by the integration with assertion

*1 Execution environment: CentOS 7.4.1708, AMD Ryzen Threadripper 1950X 16-Core Processor, 64GB Memory, clang 6.0.0 compiler, Mathematica 11.3.0.

techniques. The basic concept of this method is equivalent to the uniform approach, where, in the beginning of the procedure, the uniformity of monotonicity is assumed to hold throughout the simulation. The notable feature of this method is that the uniformity is constantly asserted by statements `ASSERT(x'>=0)` and `ASSERT(x'<=0)`. When violations on these conditions are detected, the simulation is restarted from that time point with the assumptions on monotonicity inverted.

Illustrated in Figure 5 is the concept of the enhanced method. α and β in the diagram corresponds to those in propositions (2) and (3). In the beginning of the procedure, α is set to 0 and β is set to $MaxT$. The properties are *not* determined to be held at this point. This is different from the case in the method for uniform monotonicity, where either of the propositions (2) or (3) were determined to be held from the computations made prior to the simulation. On the occurrence of assertion violations, that time point is set as the new α and the procedure continues. This process is repeated until the simulation time reaches $MaxT$. The applicable guarded constraints are removed and reset dynamically during the simulation by using these information of monotonicity.

5. Experimental Results

The approach to uniform monotonicity was implemented and its effectiveness was evaluated with the model of a ball bouncing on split surface (Figure 1). The comparison between the simulation time of the original algorithm of HyLaGI and the algorithm with the proposed method is shown in Figure 4. The regression line for the result of the proposed method is represented by the following equation:

$$SimulationTime = 0.6083\,N^2 + 0.8763\,N + 1.691 \qquad (4)$$

By comparing equations (1) and (4), we can infer that the simulation time with the proposed method is half the length of the original. Since the method aims to improve simulation efficiency by reducing the size of guarded constraints, this proposal is expected to be effective for models that contain large number of guarded constraints.

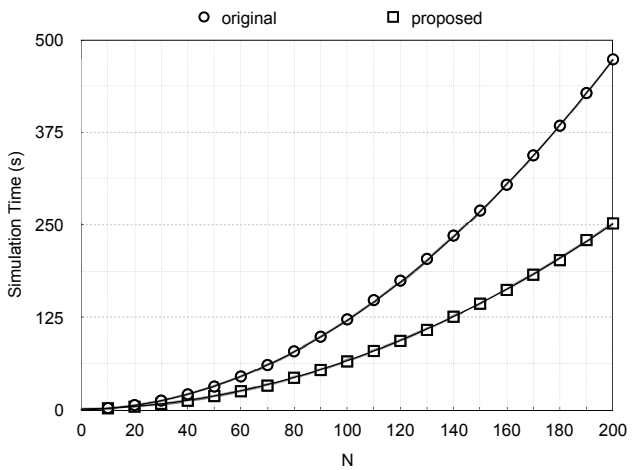


Figure 4: Simulation Time of Split Surface Model

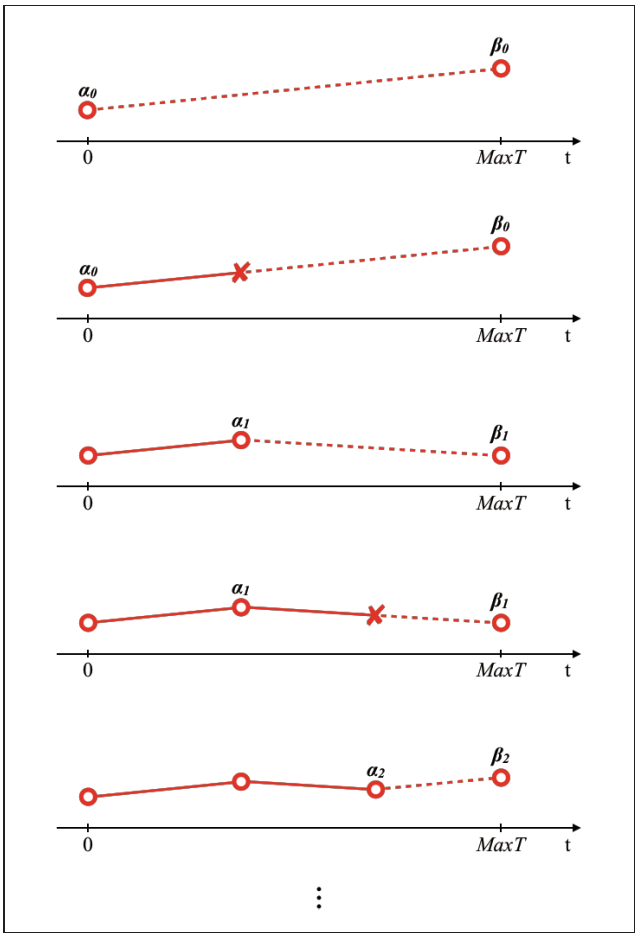


Figure 5: Concept of Approach to Uniform Monotonicity

6. Conclusion and Future Work

In this research, we proposed an optimization technique for the simulation algorithm of HyLaGI. This was achieved through dynamically reducing the size of guarded constraints that appears in the model, based on monotonicity.

Future work of this research includes evaluating the alternating monotonicity approach. Another work to put effort into is exploiting invariants other than monotonicity, for this would enlarge the scalability of this research.

References

[Lunze] J. Lunze, "Handbook of Hybrid Systems Control: Theory, Tools, Applications", Cambridge University Press, 2009.

[Ueda et al.] Ueda, K., Matsumoto, S., Takeguchi, A., Hosobe, H. and Ishii, D.: "HydLa: A High-Level Language for Hybrid Systems", In Proc. Second Workshop on Logics for System Analysis, 2012, pp.3-17.

[Borning et al.] Borning, A., Freeman-Benson, B. and Wilson, M., "Constraint Hierarchies", Lisp and Symbolic Computation, Vol.5, No.3, 1992, pp.223-270.

[Matsumoto] Shota Matsumoto: "Validated Simulation of Parametric Hybrid Systems Based on Constraints", Graduate School of Fundamental Science and Engineering, Waseda University, Doctoral Thesis, 2017.

Organized Session | Organized Session | [OS] OS-17

[1F3-OS-17a] 農業と AI(1)

小林 一樹（信州大学）、竹崎 あかね（農研機構革新工学センター）

Tue. Jun 4, 2019 3:20 PM - 5:00 PM Room F (302B Medium meeting room)

[1F3-OS-17a-01] (Invited talk) Issues of field high throughput plant phenotyping based on images

〇Seishi Ninomiya[1] （1. University of Tokyo）

3:20 PM - 3:40 PM

[1F3-OS-17a-02] Bird Damage Prevention System Utilizing Deep Learning based on Birds' Behavior

〇Kazuki Kobayashi[1], Fumiya Shimobayashi[1], Kazunori Terada[2], Takefumi Yoshikawa[3], Hiroyuki Sato[4], Hiroyuki Tsuchiya[4], Kanokwan Atchariyachanvanich[5] （1. Shinshu University, 2. Gifu University, 3. Toyama Prefectural University, 4. Marimo Electronics Co., Ltd., 5. Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang, Thailand）

3:40 PM - 4:00 PM

[1F3-OS-17a-03] Training Data Augmentation for Hidden Fruit Image Segmentation by using Deep Learning

〇Ryoma Takai[1], Kazuki Kobayashi[1] （1. Shinsyu University）

4:00 PM - 4:20 PM

[1F3-OS-17a-04] *Q*-learning with Neural Network for Automatic Irrigation Control of Crops

〇Shuto Namba[1], Junpei Tsuji[1], Masato Noto[1] （1. Kanagawa University）

4:20 PM - 4:40 PM

[1F3-OS-17a-05] Crop Yield Estimation for Hydroponic Tomatoes using regional CNNs

〇Yuri Isoyama[1], Fumiyo Emura[1], Hirohisa Satoh[1], Takashi Shinozaki[2,3] （1. Kyowa Co., Ltd., 2. Japanese Society for Artificial IntelligenceCiNet, National Insitute of Information and Communications Technology, 3. Graduate School of Information Science and Technology, Osaka University）

4:40 PM - 5:00 PM

©The Japanese Society for Artificial Intelligence

3:20 PM - 3:40 PM (Tue. Jun 4, 2019 3:20 PM - 5:00 PM Room F)

[1F3-OS-17a-01] (Invited talk) Issues of field high throughput plant phenotyping based on images

〇Seishi Ninomiya[1] （1. University of Tokyo）

Keywords: Agricultural information, high throughput plant phenotyping

野外圃場で植物の形質を高速に評価する植物フェノタイピング技術開発が急速に進んでいる．とくに，画像によるフェノタイピングは AIやドローン技術の発展にも支えられ，人による目視判断を代替するばかりか，精度や安定性においても人を凌駕するようになってきている．しかし，同時に多くの課題も浮き彫りになっている．本講演では，それらの課題と今後の展望について議論する．

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

深層学習を用いた鳥行動に基づく追い払いシステムの開発

Bird Damage Prevention System Utilizing Deep Learning based on Birds' Behavior

小林 一樹 *1 Kazuki Kobayashi　下林 史弥 *1 Fumiya Shimobayashi　寺田 和憲 *2 Kazunori Terada　吉河 武文 *3 Takefumi Yoshikawa　佐藤 寛之 *4 Hiroyuki Sato
土屋 博之 *4 Hiroyuki Tsuchiya　Kanokwan Atchariyachanvanich*5

*1信州大学 Shinshu University　*2岐阜大学 Gifu University　*3富山県立大学 Toyama Prefectural University　*4マリモ電子工業株式会社 Marimo Electronics Co., Ltd.
*5Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang, Thailand

This paper proposes a novel system to prevent bird damage to crops by driving birds away from agricultural fields. Bird damage is a major issue in agriculture, and many tools have previously been introduced to protect crops from wild birds. However, hazing tools such as alarms and distress calls are quite simple, for example, playing sound periodically, and birds eventually habituate to such stimuli. Thus, such hazing tools have only short-term usefulness. The proposed system adaptively produces stimuli according to bird behavior in real time by utilizing deep learning. The proposed system can sustain its effectiveness for a much longer term.

1. はじめに

近年，有害鳥獣による農作物への被害が深刻化しており，様々な対策が講じられている．たとえば，電気柵は，農作物の防護に有効な手段であるが，農業従事者の減少や高齢化を背景に設置後の管理が課題となっている [長門 11]．また，シカやイノシシ等の被害防止策として，罠とセンサユニットとを組み合わせた研究が行われており，捕獲した動物数を深度画像を用いて自動カウントし，それを管理者に通知するシステムが開発されている．

一方で，鳥類による農作物への被害を防止するシステムについては事例が少なく，現状では，爆音機やディストレスコールと呼ばれる忌避音を再生する装置による追い払いが行われている [Berge 07]．これらの手法は，鳥が刺激に慣れてしまい被害が減少しにくい問題がある [Summers 85]．

そこで本研究では，鳥の行動に基づいた追い払いにより，刺激への慣れを防止し，従来の手法と比較して持続的な効果を持つ追い払いシステムを提案する．

2. 鳥行動に基づく追い払いシステム

提案手法では，画像から鳥を検出し，農園に侵入した個体のみをターゲットとして追い払い刺激を与えるアプローチをとる．提案するシステムの構成を図 1 に示す．以下では，主たる構成要素である高精細モニタリングモジュール，深層学習を用いた鳥追跡モジュール，鳥追い払い機器制御モジュールの役割と動作について説明を行う．

2.1 高精細モニタリングモジュール

高精細モニタリングでは，独自設計した高速無線通信機器を使用し，鳥追跡モジュールへの低遅延低欠損な動画像送信を実現する．監視用カメラには，広範囲の鳥を検出するための HD 画質の全天球カメラを使用した．

連絡先: 小林一樹，信州大学 学術研究院 工学系，〒 380–0928 長野県長野市若里 4–17–1，kby@shinshu-u.ac.jp

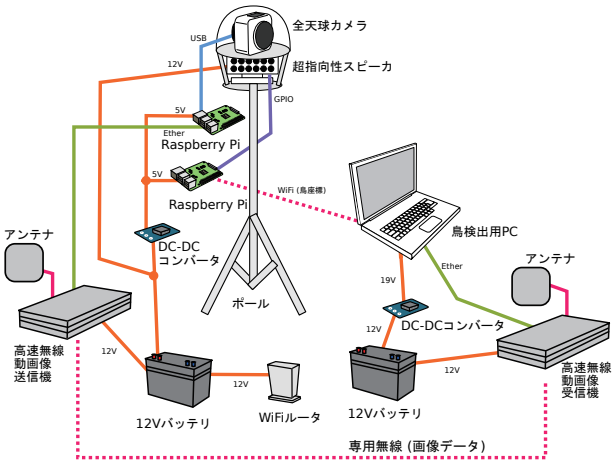


図 1: 追い払いシステムの構成

2.2 深層学習を用いた鳥追跡モジュール

鳥追跡モジュールでは，高精細モニタリングデバイスから送信された画像を用い鳥追跡を行う．鳥の検出と追跡には畳み込みニューラルネットワーク (CNN) とパーティクルフィルタを組み合わせ，パーティクルのクラスタリングにより複数個体同時追跡を実現した．図 2 複数個体の鳥の追跡結果を示す．鳥が検出された際には，その座標が鳥追い払い機器制御モジュールに送信される．

2.3 鳥追い払い機器制御モジュール

鳥追い払い機器制御モジュールでは，鳥追跡モジュールから送信された鳥座標を用い，鳥追い払い装置を制御する．カメラ座標系における鳥座標を，実世界座標系に変換し，カメラに対して鳥がどの方位にいるかを特定した上で，追い払い機器に合わせてた制御を行う．

ここでは，鳥追い払い装置として，従来の爆音機や無指向性

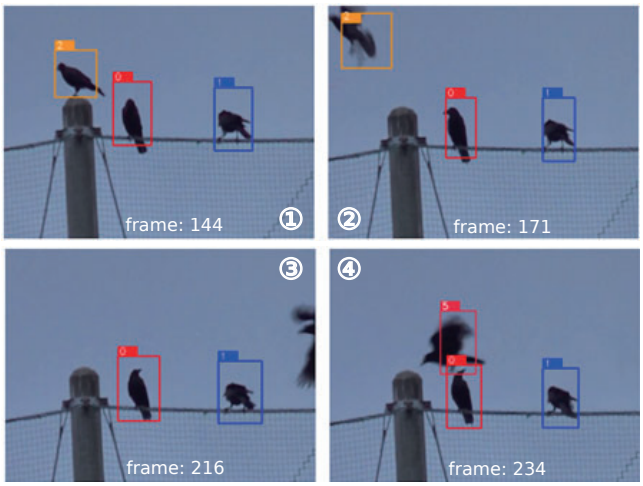


図 2: 複数個体の鳥を追跡する様子

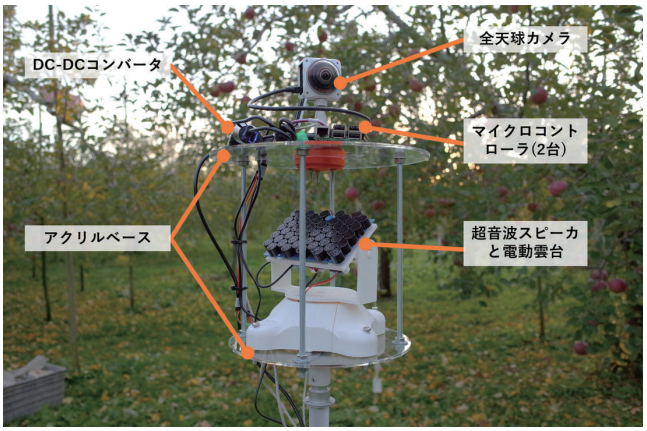


図 3: 開発した鳥追い払いシステム

スピーカと比較して，周辺への影響の少ない超指向性スピーカーを採用した．鳥が圃場の外にいる場合には，追い払い装置による刺激が与えられないため，慣れを防止した持続的な追い払いが可能だと考えられる．図 3 に開発した鳥追い払いシステムを示す．スピーカは 2 自由度の電動雲台に固定され，鳥検出座標に基づいてスピーカが回転される．

3. 検証実験

電動雲台の追尾性能を評価するため，鳥に見立てた AR マーカをカメラで認識し超指向性スピーカを制御する実験を行った．

3.1 実験方法

AR マーカの設置地点において，騒音計による音圧測定を行い，追尾対象を超指向性スピーカの指向角の範囲に収めることができるかを検証した．カメラ–AR マーカ間の距離は 2 メートルから 17.5 メートルの範囲であり，この間の 7 地点において測定を行った．各地点では追尾時の音圧に加え，環境音と，参照値としてスピーカ音圧が最大となるように手動調整した音圧も計測した．

3.2 実験結果

図 4 に音圧測定を行った結果を示す．各地点では 4 回ずつ測定を行い，開始時はスピーカ面をマーカに対して 90° に向けた初期位置からマーカ追尾動作を実行し，スピーカ音を測定した．

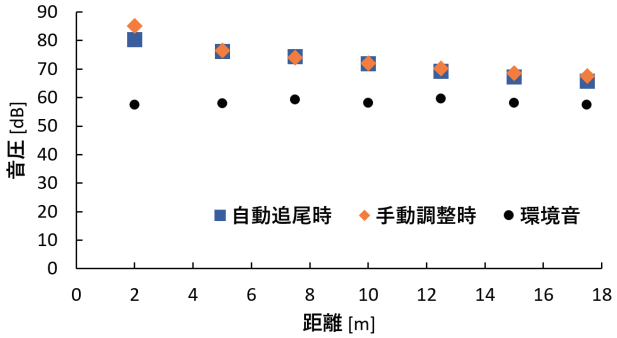


図 4: 音圧測定結果

4. 考察

超指向性スピーカの追尾性能を音圧測定を通して評価したところ，すべての計測地点において参照値と同様の音圧が確認された．実際に圃場で運用するためには，複数個体に対応した電動雲台制御が必要となる．たとえば，すべての個体を網羅するような回転パターンや，鳥の動きに追従した連続的な動きを実装する必要がある．

現状では，鳥検出モジュールと鳥追い払いモジュールの連携が部分的にしか実装されておらず，今後リアルタイムな追い払い動作の実現に向けた実装を進める必要がある．

5. まとめ

本研究では，農園における鳥害防止のための追い払いシステムを提案した．農園での鳥の追い払いでは，鳥が刺激に慣れてしまい，効果が継続する時間が短い問題ある．提案手法は，この問題の解決のために，深層学習によって鳥の位置検出を行い，鳥の行動に合わせて追い払い刺激を与えるアプローチをとった．電動雲台と超指向性スピーカを組み合わせたシステムを開発し，刺激提示範囲を限定できることを実験により確認した．今後，より効果的な追い払い刺激の提示方法として，ドローンを用いた手法を実装する予定である．

謝辞

本研究の一部は総務省 戦略的情報通信研究開発推進事業 (SCOPE) により実施されました．

参考文献

[Summers 85] Summers, R. W.: The effect of Scarers on the presence of Starlings (Sturnus vulgaris) in Cherry Orchards. Crop Protection, Vol.4, No.4, pp.520-528 (1985).

[Berge 07] Berge, A., Delwiche, M., Gorenzel, W. P. and Salmon T.: Bird Control in Vineyards Using Alarm and Distress Calls. American Journal of Enology and Viticulture, Vol. 58, pp.135-143 (2007).

[長門 11] 長門雄治，吉仲怜：鳥獣被害対策における電気柵管理の実態と方向性，農業経営研究，Vol.49，No.2，pp.105–110 (2011).

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

農園画像における深層学習を用いた隠れ果実領域の抽出

Training Data Augmentation for Hidden Fruit Image Segmentation by using Deep Learning

高井　亮磨[*1]
Ryoma Takai

小林　一樹[*2]
Kazuki Kobayashi

[*1] 信州大学　工学部
Faculty of engineering, Shinshu University

[*2] 信州大学　学術研究院
Academic Assembly, Shinshu University

This paper proposes a method to detect apple fruits areas from field monitoring images. The proposed method can detect the fruits areas even if they are hidden by leaves or other fruits. We use a deep neural network with a way to create a large amounts of artificial field images as training data. Since the developed system uses the fruit and leaf image components and can easily retrieve their exact areas in the creation process, the created data include annotations of hidden fruits. The experimental results showed that the proposed method showed promising results in terms of accuracy compared to that of the COCO pre-trained model.

1.　はじめに

近年，IoT や ICT を用いて農作物の生育状態を抽出し，生産性の向上を目指す研究が行われている．たとえば，深層学習を用いた果樹の果実検出 [Bargoti 2017] に関する研究では，深層学習を利用した物体検出手法である Faster R-CNN[Ren 2015]を用いて，リンゴやマンゴー，アーモンドの農園画像から果実を検出している．これらの研究では果実を囲む矩形領域を検出している．もし果実領域を検出できれば，正確なサイズと形状とを詳細に把握でき，農園管理に有用である．しかし，多くの果樹は果実を覆い隠すように葉を茂らせているため，先行研究では葉で隠れた果実サイズの検出までは実現されていない．

そこで本研究では，定点カメラで観測した農園モニタリング画像から，果実の一部が隠れていても領域推定が可能な手法を提案する．提案手法では，深層学習を用いた隠れた果実領域抽出のための効率的な訓練データ拡張システムを開発する．

2.　隠れ果実領域の深層学習

深層学習での物体検出には大量の訓練データが必要であり，COCO [Lin 2014] のような既存の大規模データセットを用いて学習を行うことが多い．しかし，これらのデータセットには，隠れた果実領域を検出するためのデータが存在しない．大量の訓練データを人手で生成する作業には多大な労力を要するとともに，隠れた領域に関しては，人間が正確に領域を特定できない問題がある．そこで，本研究では実際の農園モニタリング画像をベースとして，人工的に大量の訓練データを生成するアプローチをとる．

2.1　果実領域の検出

農園観測画像から果実の領域を抽出するために，畳み込みニューラルネットワークによる物体検出手法である Mask R-CNN [He 2017] を用いる．Mask R-CNN は，画像の中の任意の対象について，その形状に合わせて領域を抽出する手法である．

本研究においては隠れている果実領域を推定するために，果実が隠れている観測画像に対して，隠れていない状態の果実領域を学習させる．

連絡先：　高井亮磨，信州大学　工学部，長野県長野市若里 4-17-1，E-mail 15t5047d@shinshu-u.ac.jp

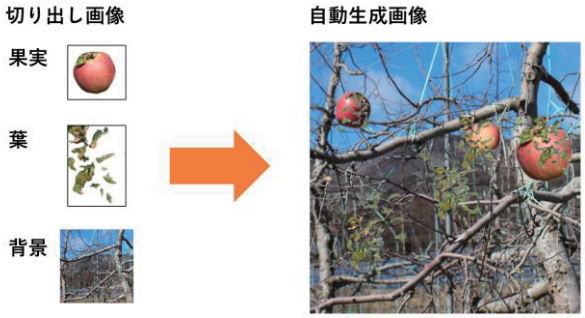


図 1: 合成画像の生成方法

2.2　訓練データ自動生成

生成する訓練データは，合成農園画像と，果実領域を示すバイナリマスクの 2 つである．図1に合成農園画像の生成方法を示す．まず，実際の農園モニタリング画像から手動で果実画像，葉画像，背景画像の切り出しを行う．このとき，切り出された画像は PNG 形式の RGBA 画像とし，果実画像と葉画像の背景は透明化されている．次に，切り抜いた画像の輝度や角度，拡大率をランダムに変更したものをランダムな位置に配置することで，合成農園画像を生成する．

バイナリマスクとは，白黒の 2 値で表現された対象物の領域を表すデータである．Mask R-CNN では，対象ごとにバイナリマスクを用意する必要がある．

果実ごとのバイナリマスクは，図 2 に示すように，果実画像を背景画像上に配置する際に生成する，具体的には，果実画像の不透明領域と透明領域とで区別してバイナリ画像を生成し，全画素値を 0 で初期化した合成農園画像と同じサイズの画像上に配置する．バイナリ果実画像を配置する座標は，果実画像を配置した座標と同一とする．

バイナリマスクは果実領域を表しているが，生成した合成農園画像では，果実画像を配置した後に葉画像や，ほかの果実画像が配置されるため，果実部分には一部隠れた領域が生じる．このような手順によって合成農園画像とバイナリマスクを大量に生成することで，隠れた果実の領域であっても元の領域を抽出可能な訓練データが得られる．

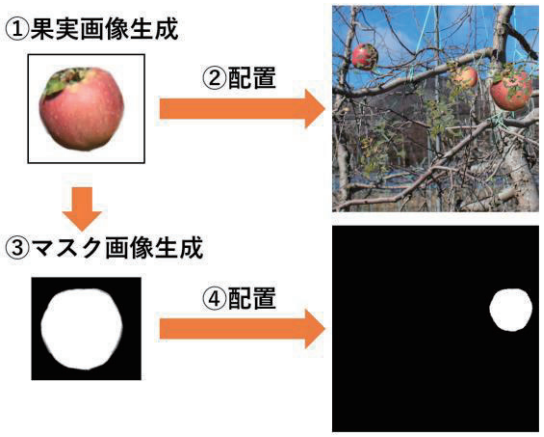


図 2: 果実ごとのバイナリマスクの生成

表 1: 果実検出精度の比較

	COCO	提案手法
AP	0.215	0.367
$AP^{IoU=.50}$	0.432	0.769
$AP^{IoU=.75}$	0.195	0.315
$AR^{max=100}$	0.384	0.492

3. 実験

提案手法による隠れ果実領域の学習の有効性を検証するために，比較実験を行った．比較対象は，COCO データセットを用いた学習済みモデル[1]とした．この学習済みモデルは，80 種類のクラスに対応している．提案手法では，この COCO 学習済みモデルに対して，上記の手法で生成した隠れ果実領域を含む 3000 枚の訓練データを追加学習した．

学習モデルの検証用データとして，実際のリンゴ農園におけるモニタリング画像を 9 枚用い，そこから 90 枚に分割して切り出した画像を用いた．この 90 枚の画像に対して，リンゴの果実部分を手作業でアノテーション付けした．このとき，一部分が隠れている果実においては，作業者の判断で元の形を推定して領域を決定した．

果実の検出精度を確認するために COCO の精度検証手法である Average Precision (AP) と Average Recall (AR) [Lin 2014] を採用した．表 1 に各指標に対する果実の検出精度を示す．

4. 考察

4.1 隠れ果実領域の学習の有効性について

表 1 の結果より，各評価項目で提案手法のスコアが COCO 条件よりも上回っており，より正しく果実領域を抽出できていることがわかる．また，図 3 に示すように実際の検出領域を比較したところ，提案手法では果実の上に重なっている葉の影響を受けずに隠れた果実領域をより高精度に特定できていることがわかる．このように，提案手法によって隠れた果実領域の学習が可能であり，未学習の農園画像に対しても有効であることが示唆された．

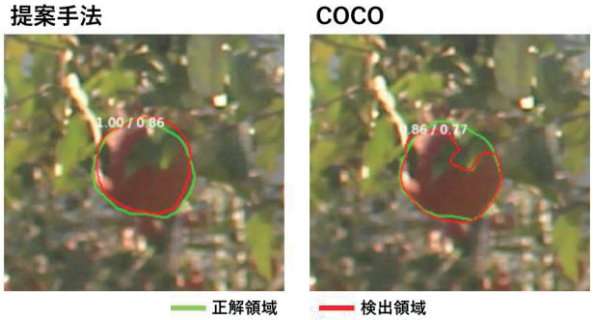


図 3: 隠れ果実領域の検出例(緑 : 真値，赤 : 推定値)

4.2 検証用のデータについて

隠れた果実領域は真値の取得が難しく，現状は作業者 1 名による手作業で対象領域を抽出し真値として扱っている．そのため，隠れた果実領域についてのデータの客観性が高いとは言い難く，より厳密な検証のためには，複数人での同じ対象の領域抽出を実施するのが望ましい．

5. まとめ

本研究では，果実の生長の把握のために，深層学習を用いて農園モニタリング画像から隠れた領域を含めた果実領域を抽出するための訓練データの生成方法を提案した．提案手法の有効性を検証するため，COCO 学習済みモデルに対して，生成した訓練データを学習させた際の果実検出精度の変化を評価した．今後，葉によって隠れている果実領域の抽出精度の向上や，検出した果実領域を用いた自動訓練システムなどの開発にも取り組む予定である．また，果実領域の正確な面積が求まれば，果実の生育情報抽出に有用であるため，正しい果実領域と予測した果実領域との面積比の平均値を求めることで，果実領域の面積抽出精度を評価する予定である．

参考文献

[Bargoti 2017] Bargoti, S., Underwood, J.: Deep fruit detection in orchards. In 2017 IEEE International Conference on Robotics and Automation (ICRA), pp. 3626-3633 (2017)

[Ren 2015] Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks, Neural Information Processing Systems, Vol.1, pp.91-99, (2015)

[Lin 2014] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L.: Microsoft COCO: common objects in context, In Proc. of the European Conference on Computer Vision, pp.740-755, (2014)

[He 2017] He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask R-CNN, In Proc. of the International Conference on Computer Vision, pp.1059-1067, (2017)

[1]https://github.com/matterport/Mask_RCNN/releases/download/v2.0/mask_rcnn_coco.h5

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

農作物の自動潅水制御に向けたニューラルネットワークを用いたQ学習

Q-learning with Neural Network for Automatic Irrigation Control of Crops

難波 脩人 Shuto Namba　辻 順平 Junpei Tsuji　能登 正人 Masato Noto

神奈川大学大学院工学研究科電気電子情報工学専攻
Graduate School of Electrical, Electronics and Information Engineering, Kanagawa University

There are many studies to reproduce cultivation methods based on experience and intuition of skilled farmers by supervised learning. Farmer's knowledge has an impact on the collecting data to conduct supervised learning, and it is difficult to judge whether such knowledge is a true optimal solution in agricultural work. So, we focused on reinforcement learning (RL) which learns based on reward given from the environment as a learning method independent the knowledge like experience and intuition. There are few cases of study using RL, and also the effectiveness has not been enough clarified. In this paper, we tried to cultivate Komatsuna based on simple Q-learning as a preparation to apply RL to the plant cultivation. As a result of Q-learning for a Komatsuna, it was possible to confirm the state of giving water according to the height of plant.

1. はじめに

日本の農業が抱える問題の一つに農業従業者の高齢化による離農率の上昇があげられる．それに伴い，熟練農家の長い年月をかけて培った技術が継承されることなく失われていくことが懸念されており，熟練農家の知見を残すためにも技術継承は農業分野における課題である．こうした問題を解決するため，センサや IoT 技術を農業に生かしたスマート農業が注目されている．これらの分野では機械学習を用いて従来経験と勘によって行われてきた農作業の方法を再現する取り組みが行われている．特に機械学習の中でも，センサやカメラから収集した温湿度や水分量，草姿の変化などのデータを用いて教師あり学習によって熟練農家の知見を再現しようとする研究が盛んに行われている．例えば，センサを用いた研究では茎の太さの変化を測定することで植物に付加されているストレスを推定する研究 [Gallardo 06]，露地栽培における土壌水分量や水耕栽培における養液の濃度を測定することで自動潅水制御を行う研究 [Abidin 14] などがある．一方でカメラを用いた研究では植物の画像から病害診断を行う研究 [Mohanty 16]，葉のしおれ具合を推定することからストレス推定を行う研究 [Kaneda 17] などがある．農業のような複雑な要因が絡む作業において，農家の行う栽培方法が最適かどうか判断することは難しい．一方で，センサやカメラを用いてデータを収集する際に植物のどこに注目し，どのデータが重要であるか判断するために熟練農家の知見が必要不可欠である．さらに，実際の栽培環境から教師あり学習に必要なデータを採取するには限界があることが知られている．そのため，熟練農家の知見は教師データの質に大きく影響し，採取したデータに基づいて学習した結果が最適かどうかの判断は困難である．

そのような背景のもと，熟練農家の知見である経験と勘に依存せずに環境から与えられる報酬に基づいて学習を実行する学習手法として強化学習があげられる．実際に植物栽培に強化学習を適用した研究としてシミュレーションによる収穫量の最適化 [Sun 17] や栽培を繰り返すことによる養液供給の最適化 [Wakahara 10] などを目的としたものがある．実際の植物栽培に強化学習を適用した研究に焦点を当てた場合，栽培期間の長さや状態の時間変化など植物のもつ特徴によってロボティクスで用いられるような手法をそのまま適用することが困難であることが知られている．Wakahara らの研究は植物の栽培に強化学習を適用した際に生じるこれらの問題を植物の状態遷移や報酬の与え方から定義することで解消している．しかしながら，その事例は少なく，有効性は十分明らかにされていない．

本研究では，植物の栽培に強化学習を適用するための準備として Q 学習に基づく小松菜の栽培を実践する．植物の状態は時間変化することから同じ状態は存在せず，全ての状態に対して Q テーブルを作成することが困難である．そのため，ニューラルネットワーク（NN）を用いた Q 関数の作成手法を採用する．植物栽培における NN を用いた Q 学習の有効性や課題を検討する．

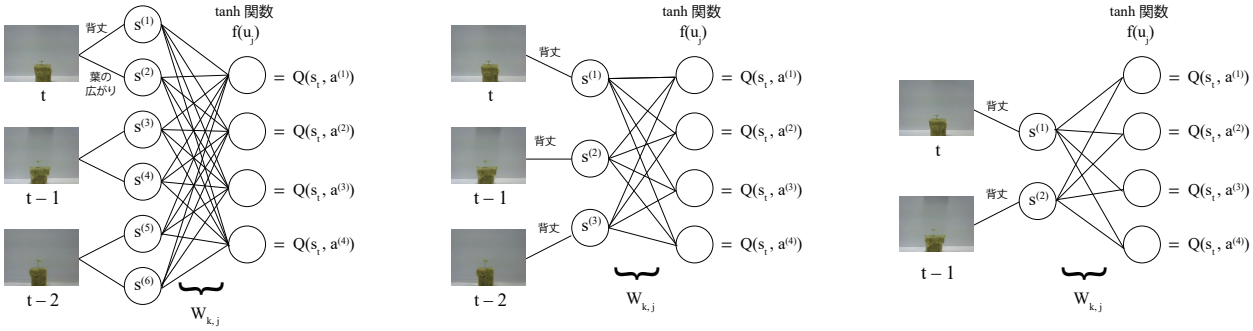
2. 植物栽培の最適化に向けた強化学習

2.1 ニューラルネットワークを用いた Q 学習

本研究では植物がよりよく成長するための潅水行動の最適化に向けて強化学習の手法の中でも最もよく利用される Q 学習を用いる．Q 学習は状態と行動の各組に対して Q 関数を探索と活用を繰り返すことで更新し，エージェントが最適な行動を選択するよう学習させる手法である．

一方，植物栽培の最適化を強化学習によって行うにあたり重要な点として状態の定義があげられる．植物の状態は時間とともに変化することから同じ状態は 1 試行の中に 1 度しか現れない．つまり，Q 学習ではすべての状態に対して Q テーブルを構築することが現実的ではない．さらに，植物の状態は連続量であることから状態を Q 関数で表すことが困難である．このような場合において重み W_{kj} を用いて Q 関数の作成に関数近似を用いる手法が知られている．本研究における Q 関数の導出概要を図 1 に示す．関数近似には出力層に tanh 関数を用いた NN を使用している．複雑な NN で学習できるほど十分なデータが集められるとは限らないため，NN の構造は単層構造とした．Q 関数は tanh 関数を用いて近似するため，

連絡先: 難波脩人，神奈川大学大学院工学研究科電気電子情報工学専攻　能登研究室，〒 221-8686 神奈川県横浜市神奈川区六角橋 3-27-1，電話：045-481-5661



(a) 6 入力　(b) 3 入力　(c) 2 入力

図 1: 関数近似を用いた Q 関数の作成概要

$f(u_j) = \tanh(u_j)$ としたとき，以下の関係式が成り立つ．

$$Q(s, a^{(j)}) = f(u_j) \tag{1}$$

$$u_j = \sum_{k=1}^{n} W_{kj} s^{(k)} + b_j \tag{2}$$

ここで時刻 t の状態 S_t の入力は 3 パターン用意し，それぞれ $n = 6$ 入力，3 入力，2 入力としている．植物の状態を定義する際に時刻 t における背丈や葉の広がりのみから状態を定義した場合，現在の状態に至るまでの遷移が反映されないことが考えられる．例えば，時刻 $t-1$ で健康な状態と比較して枯れているとみられる植物がある行動によって時刻 t で回復した場合と時刻 $t-1$ で健康な植物が同じ行動で健康なままの場合を考える．時刻 t の植物の状態のみを反映した s_t と時刻 t と $t-1$ の状態を反映した s_t とではより詳細な植物の状態を表すためには，後者の状態のほうが望ましいと考えられる．また，植物の状態は背丈だけに現れるのではなく，子葉の成長具合にも反映される．そのため，6 入力の場合は時刻 t，$t-1$，$t-2$ の各時刻における植物の背丈と葉の広がりを表した 2 次元ベクトルを 3 つ並べた 6 次元ベクトルとして $s_t = (s^{(1)}, s^{(2)}, s^{(3)}, s^{(4)}, s^{(5)}, s^{(6)})$ で表される．3 入力の場合は時刻 t，$t-1$，$t-2$ の各時刻における植物の背丈のみの 1 次元ベクトルを 3 つ並べた 3 次元ベクトルとし，2 入力の場合は時刻 t，$t-1$ における植物の背丈を 2 つ並べた 2 次元ベクトルとしている．学習時における重み W_{kj} の更新式は以下の式 (3) のようになる．また，式 (3) の δ_t の導出は式 (4) に示す．

$$W_{kj} \leftarrow W_{kj} + \alpha \delta_t \frac{\partial f(u_j)}{\partial W_{kj}} \tag{3}$$

$$\delta_t = r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \tag{4}$$

α は学習率，γ は割引率とし，それぞれ $\alpha = 0.5$，$\gamma = 0.9$ と設定している．報酬の定義はどの入力パターンに対しても同様に時刻 t と $t-1$ における植物の背丈を比較し，以下のように定義する．

$$r_t = \begin{cases} +0.2 & \text{（時刻 } t \text{ で成長している）} \\ 0 & \text{（変化なし）} \\ -0.2 & \text{（時刻 } t \text{ でしおれている）} \end{cases}$$

行動（action）は $a^{(0)}$ = “0 ml”，$a^{(1)}$ = “10 ml”，$a^{(2)}$ = “15 ml”，$a^{(3)}$ = “30 ml” の潅水を 4 パターン用意し，状態 s_t で action を起こした場合に次の状態 s_{t+1} に遷移するのは水を与えてからの変化に時間がかかることから 2 日後と間隔を開けるよう設定している．農作物の栽培期間は 1 か月から数か月かかるものがほとんどであり，種を植えてから収穫までを 1 試行とした場合の学習効率は悪い．

本研究における 1 試行は学習効率を上げるために栽培期間を 2 週間として学習している．強化学習の適用先は小松菜とし，状態は種を植えてから一律 30 ml の潅水を行った 4 日後を s_1 と定義する．

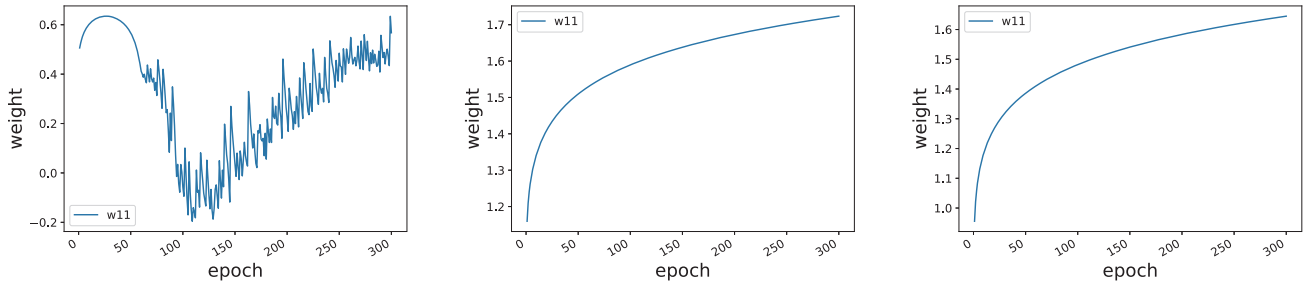
2.2 experience replay を用いた学習

1 試行やステップ数は学習の効率を上げるために 2 週間の 7 ステップと定義している．一方，シミュレーションに強化学習を適用した場合などと比較すると十分なデータ数が集まらず，植物に対する学習が最適解に収束することは困難であると考えられる．仮にステップ数を長くすることを考えると植物は成長段階によって要求する水分量が変化することから行動によって与える水分量を変化させる必要がある．強化学習を植物に適用していくためには植物の成長段階に合わせた学習手法が必要であると考えられる．本研究では植物が十分なデータを確保しつつ学習が収束していくことを目的とし，experience replay (ER)[Lin 92] を採用する．ER はエージェントが経験したサンプル $\{(s_t, a_t, r_t, s_{t+1})\}_{t=1}^{T}$ をメモリに保存する．このとき，T はサンプル数の合計を表している．一定量のサンプルが保存されたところでメモリからサンプルをランダムに抽出し，学習の入力として再利用することで学習の効率化を図っている．つまり，メモリにサンプルを保存するために人の手によって小松菜の栽培を行うことでサンプルを一定量収集し，メモリに保存した栽培データを用いて重み W_{kj} を更新する．

3. 実験および結果

ER を植物栽培に対して適用する際にメモリにどの程度のデータ数を保存し，かつデータをどの程度学習させることで植物が最適な方策を得るかを評価するため，3 つの栽培期間に分けてデータ採取を行った．小松菜は室内栽培用ハウスキットの中でロックウールと呼ばれる培養地に植えている．

- 2018 年 12 月 1 日から 2018 年 12 月 13 日（1 回目）
  発芽しない個体や枯れる個体が現れることもあり，小松菜 48 株を栽培した結果，データ採取ができた個体は 28 株であった．このときのサンプル数の合計は $T = 168$ とした．



(a) 6 入力時の重みの更新の様子　(b) 3 入力時の重みの更新の様子　(c) 2 入力時の重みの更新の様子

図 2: ER による各入力に対する重み W_{kj} の更新の様子（$k = 1, j = 1$ の場合）

- 2018 年 12 月 14 日から 2018 年 12 月 26 日（2 回目）
  小松菜 16 株栽培した結果，15 株からデータを採取し，サンプル数の合計は 1 回目の結果と合わせて $T = 256$ とした．
- 2019 年 1 月 14 日から 2019 年 1 月 26 日（3 回目）
  小松菜 48 株栽培した結果，40 株からデータを採取し，サンプル数の合計は 1 回目，2 回目と合わせて $T = 496$ とした．

3.1　サンプル数 $T = 168$ に対する学習結果

実験では 3 つの栽培期間で採取したデータの各サンプル数に対して 6 入力，3 入力，2 入力の入力を試すことで植物に強化学習を適用するために必要なデータ数及び，状態の定義など ER を用いた場合にどの手法が最適であるかを明らかにする．まず，サンプル数 $T = 168$ に対する ER の学習成果を評価するため訓練データに使用したデータを入力し，予測された action と実際の action を比較した．その結果，スペースの都合上結果は除くが，どの入力の場合に対しても偏った行動しか示せず，学習が成功した様子は確認できなかった．

3.2　サンプル数 $T = 256$ に対する学習結果

次に $T = 256$ のサンプルを訓練データとして ER を行い，各入力に対する学習結果を評価するため，同様に action の比較を行った．このとき，3 入力と 2 入力の学習結果からは状態における action の選択にばらつきが生じていることが確認できた．学習段階の様子を確認するため，各入力に対する重み W_{kj} の更新の様子を図 2 に示す．各グラフはそれぞれの重みに対して $k = 1, j = 1$ とした場合の値を表しており，6 入力では重みが収束していない様子が確認できる．一方で 3 入力と 2 入力の場合においてはどちらも収束している様子が明らかである．

3.3　サンプル数 $T = 496$ に対する学習結果

最後に $T = 496$ のサンプルを訓練データとして ER を行った結果に対して同様の評価を行った．6 入力に対しては $T = 168, 256$ のときと同様にデータ数を増やしたところで改善は見られず，学習回数を増やしたとしても重みが収束する様子は確認できなかった．3 入力と 2 入力に対しては $T = 256$ と同様に重みの更新は収束している様子が確認できた．しかしながら，ER を行うにあたり，メモリに保存するデータは少ないほうが学習の効率は良いことが考えられるため，本研究では $T = 256$ の場合に着目する．

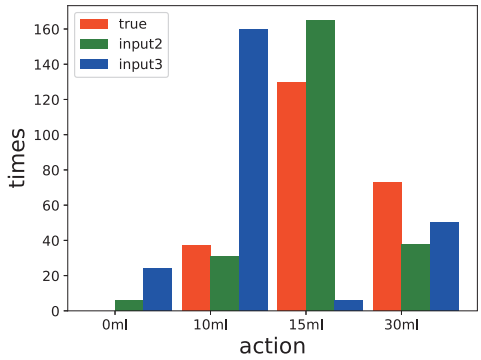


図 3: action の出現頻度

3.4　サンプル数 $T = 256$ かつ 2 入力に対する学習結果

3 入力と 2 入力における状態の定義としてどちらが優れているかを評価するために，訓練データで使用したデータを NN に入力した際に出力される action のばらつきをまとめてグラフに示す（図 3）．グラフの true は人の手によって小松菜を栽培した際に与えていた action を示している．input2 や input3 は強化学習を用いて学習した NN に訓練データを入力した際に得られた action を表しており，後ろの数字はそれぞれ状態の入力数を表している．横軸は action の種類，縦軸は各 action の出現頻度を表している．グラフからわかる通り，input3 の action は true の action と比較すると $a^{(1)}$ = “10 ml” の頻度が多く，$a^{(2)}$ = “15 ml” の頻度が少ないことが明らかである．また，input2 の action と true の action を比較すると $a^{(1)}$ や $a^{(2)}$ の頻度はほとんど一緒であることが明らかであった．このことから 3 入力と 2 入力の場合では学習は同様に収束しているにもかかわらず，2 入力の方が人の手による栽培と近い行動を選択していることが明らかである．

2 入力に対する強化学習の評価としてどのような方策に基づいて action を決定しているかを明らかにするために状態の入力に対する行動の選択分布をグラフに示す．図 4 は強化学習による行動選択，図 5 は人の手による行動選択を表しており，横軸は入力時の現在の状態を表し，縦軸は次の状態を表している．強化学習による学習結果から得られた action は greedy 戦略に基づいて選択した結果を表しており，入力に対して単純に最適な action を選択した場合の結果のみを表している．

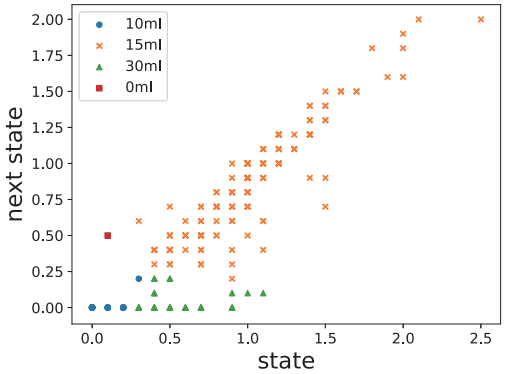


図 4: action の選択分布（ER による結果）

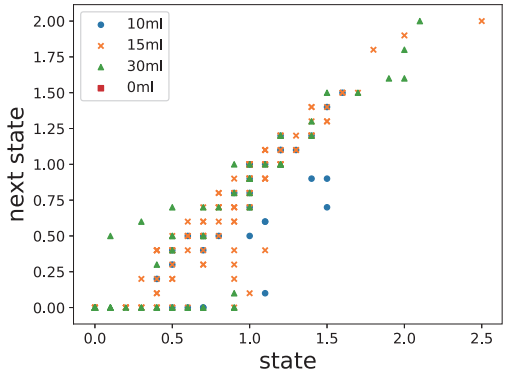


図 5: action の選択分布（人の手による結果）

4. 考察

図 4，5 から植物の状態を 2 入力とした場合の学習結果と人の手による栽培結果を比較すると，学習結果は人間と近い行動を選択している傾向がみられる．しかしながら，人の手による栽培では現在の状態と次の状態の組み合わせが同じ値をとっている場合においても異なる行動を選択している場合がある．原因として，人の手による栽培では植物の背丈や葉の広がり以外にもロックウールの乾き具合などを参考に与える水の量を決定しているが強化学習による Q 学習ではロックウールの乾き具合など，入力以外の外的要因を考慮していないことがあげられる．Q 学習を植物栽培に適用した結果，図 4 のような分布がみられたことから植物は背丈の高さを基準に与える水の量を決めていることが明らかになった．この結果は一般的な Q 学習では困難であった植物の栽培に強化学習を適用する際の手法として NN を用いた手法が有効であることを示唆している．強化学習は環境に応じてエージェント自身が最適だと思う行動を選択することでその環境における最適化を求める手法であるため，今回のように人の手によって栽培することで得た行動を真値として比較することは必ずしも正当な評価方法とは言いきれない．しかしながら，図 3 に示したようにあまりにも人の手による栽培結果とかけ離れている場合は学習結果に妥当性はないと考えられる．人の手による栽培から得られたデータは現状の暫定解に過ぎず，強化学習によって学習した結果をもとに実際に小松菜の栽培を繰り返すことでエージェントが選択する行動の良し悪しを評価することが必要であると考える．

5. おわりに

本研究では農作物の自動潅水制御に向けて実際に小松菜の栽培に強化学習を適用することからその有効性や課題の検討を目的とした．植物の特徴である状態の時間変化や種を植えてから収穫までの栽培期間の長さを考慮するため，関数近似による Q 学習を採用して実験を行っている．さらに，栽培データの少なさをカバーしつつ，学習が収束していくことを狙いとして関数近似には NN の単層構造を採用し，ER による学習を行った．NN の入力には状態を 6 次元ベクトル，3 次元ベクトル，2 次元ベクトルと 3 パターン用意し，状態の定義として適切なものを明らかにした．実験を行った結果，状態の入力を 2 次元ベクトルとし，ER に用いるためのメモリの保存数 $T = 256$ とした場合の学習結果が人の手による栽培方法と最も近い行動を選択する様子が確認できた．よって，植物の栽培に NN を用いた Q 学習を適用するためには状態の入力は 2 入力，メモリの保存数は $T = 256$ 程度とした場合，望ましいと考えられる action を選択することが明らかになった．しかしながら，人の手による栽培と比較したとき，強化学習による学習結果からは人間が判断材料としている複雑な要因を考慮している様子は見られず，改善の余地がある．

参考文献

[Abidin 14] Abidin, M. S. B. Z., Shibusawa, S., Ohaba, M., Li, Q., and Bin Khalid, M.: Capillary Flow Responses in A Soil–plant System for Modified Subsurface Precision Irrigation, *Precision Agriculture*, Vol. 15, No. 1, pp. 17–30 (2014)

[Gallardo 06] Gallardo, M., Thompson, R., Valdez, L., and Fernández, M.: Response of Stem Diameter Variations to Water Stress in Greenhouse-grown Vegetable Crops, *The Journal of Horticultural Science and Biotechnology*, Vol. 81, No. 3, pp. 483–495 (2006)

[Kaneda 17] Kaneda, Y., Shibata, S., and Mineno, H.: Multi-modal Sliding Window-based Support Vector Regression for Predicting Plant Water Stress, *Knowledge-Based Systems*, Vol. 134, No. 15, pp. 135–148 (2017)

[Lin 92] Lin, L.-J.: Self-improving Reactive Agents Based on Reinforcement Learning, Planning and Teaching, *Machine Learning*, Vol. 8, No. 3-4, pp. 293–321 (1992)

[Mohanty 16] Mohanty, S. P., Hughes, D. P., and Salathé, M.: Using Deep Learning for Image-based Plant Disease Detection, *Frontiers in Plant Science*, Vol. 7, pp. 1–10 (2016)

[Sun 17] Sun, L., Yang, Y., Hu, J., Porter, D., Marek, T., and Hillyer, C.: Reinforcement Learning Control for Water-Efficient Agricultural Irrigation, in *Proc. of ISPA/IUCC 2017*, pp. 1334–1341 (2017)

[Wakahara 10] Wakahara, T. and Mikami, S.: Adaptive Nutrient Water Supply Control of Plant Factory System by Reinforcement Learning, in *Proc. of SCIS & ISIS 2010*, pp. 1020–1025 (2010)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

周年栽培トマトの果実検出システムの検証

Crop Yield Estimation for Hydroponic Tomatoes using regional CNNs

磯山 侑里[*1] 江村 文代[*1] 佐藤 裕久[*1] 篠崎 隆志[*2*3]
Yuri Isoyama Fumiyo Emura Hirohisa Satoh Takashi Shinozaki

[*1] 協和株式会社
Kyowa Co. Ltd.,
[*2] 国立研究開発法人情報通信研究機構脳情報通信融合研究センター
Japanese Society for Artificial IntelligenceCiNet, National Institute of Information and Communications Technology
[*3] 大阪大学大学院情報科学研究科
Graduate School of Information Science and Technology, Osaka University

Tomatoes are important plants in greenhouse cultivation from the rise in their agricultural productivity by spread of environment management system, and its crop yield estimation is crucial for efficient cultivation and controlled distribution. We develop a crop yield estimation system for hydroponic tomatoes using regional convolutional neural networks and verify the detection score and estimation accuracy.

1. はじめに

農業は近年，計画生産，計画出荷の需要が高まっている．特に施設栽培のトマトでは収穫量の増減による機会損失や収穫物の廃棄が課題になっており，中長期的な収穫量の予測は喫緊の課題となっている．一週間先の収穫量を予測するためには果房についている果実数，およびその状態を把握することが好ましいが，現状では栽培管理者が目視で判断しており，定量的な評価が行われていないのが実状である．

近年の人工知能技術は，特に画像処理分野での進歩が著しく，畳み込みニューラルネットワーク(Convolutional Neural Network，以下 CNN) による画像分類を発端に，画像検出に応用した Regional CNN(以下 RCNN)と呼ばれる技術が急速に発展してきている．RCNN を農業に応用した先行研究としては，リンゴの果実検出を行った報告 [小林 18] がなされているが，より高い空間密度で結実するトマトに関しては十分な研究が行われていない．そこで本研究では栽培中のトマトの画像から翌週の収穫量を予測することを目的として，様々な分野で有効性が確認されつつある RCNN である Faster RCNN (以下 F-RCNN) [Ren 15] および Single Shot Detector (以下 SSD) [Liu 16] を用いたトマトの果実検出システムを構築し，その検出精度と収穫量予測における精度を検証した．

2. 手法

2.1 データセット

2018 年 10 月から 2019 年 2 月の間に周年栽培のミニトマトの収穫時に果実が含まれる果房をデジタルカメラで撮影した．撮影した画像（4608×3456 ピクセル）から，トマトが結実している高さで 1600×1200 ピクセルの画像を 3 枚切り出し，800×600 ピクセルに縮小した．切り出された画像に映っている果実を，専用に作成されたタグ付けソフトウェアを用いて，収穫作業者が赤（収穫できる状態），オレンジ（成熟途中），緑（未成熟）の 3 種類にタグ付けした（図 1）．画像の総数は 107 枚であり，73 枚を学習用，25 枚を検証用，9 枚を評価用に分割して用いた．

連絡先：磯山侑里，〒569-1136 大阪府高槻市郡家新町 85-1

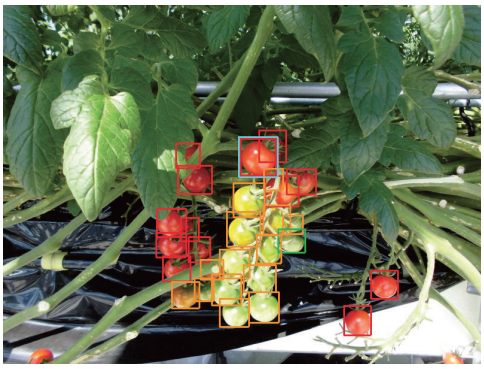


図 1. タグ付けされた撮影画像

2.2 RCNN

RCNN の実装には Preferred Networks 社製の Chainer [Tokui 15] およびに ChainerCV [Niitani 17] を用いた．学習およびに評価の各過程は NVIDIA 社製 GeForce 1070Ti を搭載したワークステーションによって行った．

F-RCNN では 800×600 ピクセルの画像を入力し，ミニバッチサイズを 1 として，3000 試行の学習を行った．SSD では 800×600 ピクセルの画像から中央の 512×512 ピクセルを切り出した後，ミニバッチサイズを 8 として，1500 試行の学習を行った．

学習は PASCAL VOC で学習済みのモデルからの転移学習によって行った．重みの更新には通常の Stochastic Gradient Descent (以下 SGD)を用い，学習係数は初期値を 0.01 とし，全試行の 1/3 ごとに学習係数を 1/10 にしたものを用いた．画像検出の精度は検証用データに対する mean Average Precision (以下 mAP)によって確認し，収穫量予測には評価用データに対して 10 分割交差検証を行い，評価用の 9 枚の画像毎，およびに画像全体を通しての検出数を確認した．

3. 実験及び結果

図 2 に学習過程における損失，およびに検証用画像に対する mAP の一例を示す．10 試行の学習に対する mAP は F-RCNN が 52.9±2.7%，SSD が 63.3±1.5%であった． 1 試行当たりの学習に要した時間は FRCNN が 1190±6 秒，SSD が 1216±8 秒であった．

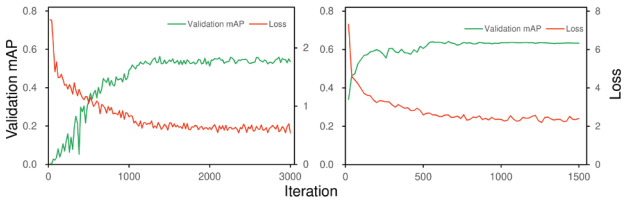


図 2.　学習過程の損失，およびに精度の変化

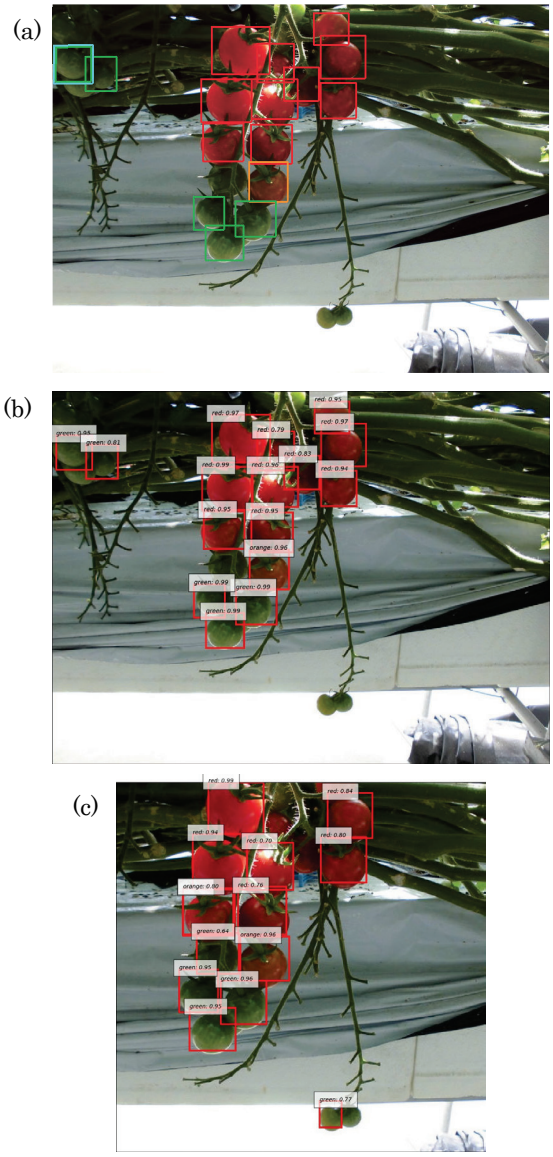


図 3.　学習済みモデルによる画像検出の結果．(a) 正解データ，(b) F-RCNN，(c)SSD

学習が完了したそれぞれのモデルを用いて，評価用画像のトマト果実検出を行った結果を図 3 に示す．図 3(a)が正解データ，図 3(b)が F-RCNN での検出結果，図 3(c)が SSD での検出結果となっている．図の様に赤，オレンジ，緑の果実を十分な精度で検出できることが確認された．

果実検出の結果から，それぞれの手法における赤，オレンジ，緑の果実の収穫量予測を，画像毎およびに評価画像全体を通しての検出数によって行った．表 1 に画像毎の収穫量予測の一例として，代表的な試行における予測結果の画像間平均の結果を示す．SSD では赤とオレンジの判断間違いが F-RCNN より高い傾向にあり，特にオレンジの果実を赤と判断することが多いため，赤の果実の予測値が 100%を越える結果となった．一方で F-RCNN では正解率の平均値自体は高い値になっているものの，葉に隠れた場合の検出率が低くなる傾向があった．表 2 に全画像を通しての収穫量予測の試行間平均の結果を示す．1 枚ごとの評価に比べて，特に検出数の標準偏差が低減され，安定した精度での予測が可能であることが確認された．

表 1.　画像毎の収穫量予測の 1 例(%) (n=9)

	赤	オレンジ	緑
F-RCNN	94.7±24.2	94.9±36.3	88.9±33.3
SSD	109.2±41.5	78.3±55.5	79.5±40.0

表 2.　画像全体を通しての収穫量予測(%) (n=10)

	赤	オレンジ	緑
F-RCNN	89.5±10.8	97.5±7.7	92.3±14.1
SSD	81.6±13.8	85.5±9.9	80.0±21.1

4.　考察

本研究では独自に構成したトマト画像のデータセットを用いて，F-RCNN と SSD による周年栽培トマトの果実検出システムを構築，検出精度と収穫量予測の精度についての検証を行った．その結果，SSD に比べて F-RCNN のほうがより安定した収穫量の予測が可能であった．さらに 1 枚毎の画像を用いた検出システムより複数の画像から検出した果実の総数で評価する方が検出数の精度が高かった．また，その標準偏差の小ささから農場全体の収穫量予測としては，予測値に一定の係数を乗じることによって一定の精度の予測が可能であることが示唆された．

RCNN は現在も発展中の技術であり，本研究で用いたデータセットや評価手法をより新しい RCNN に適用することによって更なる精度の向上が期待される．より高い精度の熟度の判定を可能とするためにカラーチャートを用いた色補正の導入などと併せて，さらなる改良を行ってゆく予定である．

5.　参考文献

[Girshick 2014] Girshick, R., Donahue, J., Darrell, T. and Malik, J.:Rich feature hierarchies for accurate object detection and semantic segmentation, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2014)

[小林　18]　小林一樹，中村俊輝: 深層学習を用いた果実生育情報抽出 Web アプリケーション，*第 32 回人工知能学会* (2018)

[Ren 15] Ren, S., He, K., Girshick, R. and Sun., J.: Faster R-CNN: Towards real-time object detection with region proposal networks, *Neural Information Processing Systems (NIPS)* (2015)

[Liu 16] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed S., Fu, C. and Berg, A.C.: SSD: Single shot multibox detector, *European Conference on Computer Vsion (ECCV)* (2016)

[Tokui 15] Tokui, S., Oono, K., Hido, S. and Clayton, J.: Chainer: a Next Generation Open Source Framework for Deep Learning, *Proc. Of Workshop on Machine Learning Systems (LearningSys)* (2012)

[Niitani 17] Niitani, Y., Ogawa, T., Saito, S. Saito, M.: ChainerCV: a Library for Deep Learning in Computer Vision, *ACM Multimedia* (2017)

Organized Session | Organized Session | [OS] OS-17

[1F4-OS-17b] 農業と AI(2)

小林 一樹（信州大学）、竹崎 あかね（農研機構革新工学センター）

Tue. Jun 4, 2019 5:20 PM - 7:00 PM Room F (302B Medium meeting room)

[1F4-OS-17b-01] A Comparative Analysis of Labor Force Survey Data among Vegetable Cultivation Systems based on Agriculture Activity Ontology

〇Akane Takezaki[1], Kaoru Maeyama[3], Joo Sungmin[2], Hideaki Takeda[2], Tomokazu Yoshida[1] （1. National Agriculture and Food Research Organization , 2. National Institute of Informatics , 3. Iwate Agricultural Research Center）

5:20 PM - 5:40 PM

[1F4-OS-17b-02] A Method for Judging the Semantic Similarity between Crops from Farm Management Articles based on Agricultural Knowledge Graph

〇Sungmin Joo[1], Hideaki Takeda[1], Akane Takezaki[2], Tomokazu Yoshida[2] （1. National Institute of Informatics, 2. The National Agriculture and Food Research Organization）

5:40 PM - 6:00 PM

[1F4-OS-17b-03] Analysis of Konbu-Dashi by multi-band optical method from a viewpoint of miscellaneous taste

〇Takaharu Kameoka[1], Takumi Taguchi[1], Eriko Nishikawa[1], Ryoei Ito[1], Atsushi Hashimoto[1], Noriyuki Yugawa[2], Nobuaki Obiki[2] （1. Mie University, 2. Tsuji Culinary Institute）

6:00 PM - 6:20 PM

[1F4-OS-17b-04] The individual identification of cattle using LP-residual signal extracted from cattle sound

〇Yurie Iribe[1], Mako Soga[1], Tomoki Kojima [2], Tatsuaki Masuda[2] （1. Aichi Prefectural University, 2. Aichi Agricultural Research Center）

6:20 PM - 6:40 PM

[1F4-OS-17b-05] Suppression of false alarm using crowdsourcing in calving detection system

〇Yusuke Okimoto[1], Susumu Saito[1,2], Teppei Nakano[1,2], Makoto Akabane[1,2], Tetsunori Kobayashi[1], Tetsuji Ogawa[1] （1. Waseda University, 2. Intelligent Framework Lab）

6:40 PM - 7:00 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

経営指標を利用した農作業基本オントロジーに基づく野菜栽培の作業分析

A Comparative Analysis of Labor Force Survey Data among Vegetable Cultivation Systems based on Agriculture Activity Ontology

竹崎あかね[*1] 前山薫[*2] 朱成敏[*3] 武田英明[*3] 吉田智一[*1]
Akane TAKEZAKI Kaoru MAEYAMA Sungmin JOO Hideaki TAKEDA Tomokazu YOSHIDA

[*1] 農研機構 National Agriculture and Food Research Organization
[*2] 岩手県農業研究センター Iwate Agricultural Research Center
[*3] 国立情報学研究所 National Institute of Informatics

Abstract: A comparative analysis of labor requirement among several cropping systems is important when crop rotation systems are designed to reduce production cost. We proposed a comparative analysis of labor requirement among vegetable cultivation systems based on Agriculture Activity Ontology (AAO). AAO was used to establish correspondences among terms to describe work in labor force survey as a core vocabulary. Working time aggregated on each AAO - defined activity enabled to compare the labor requirement among vegetable cultivation systems.

1. はじめに

農業経営では，所得最大化のために，所有する土地，施設，農業機械，労働力や，圃場環境などの資源を考慮して，作物，品種，栽培体系等を選択し作付計画を策定することが重要となる．近年では，土地・施設・機械稼働率の向上や，労働平準化による生産コスト低減を実現するため，栽培期間の異なる作物や品種を複数組み合わせた輪作体系が増えており，計画策定にあたって考慮すべき条件は増える傾向にある．

都道府県が発行する農業経営指標は，作物，品種，栽培技術，経営規模等を想定し，作物生産に要した費用，得られる収益等を示したもので，作付計画策定の際には参考データとして活用できる．農業生産者は，圃場環境が類似する都道府県だけでなく，希望する営農条件を想定した都道府県の農業経営指標も要望すると推測する．例えば，新技術導入の際には，圃場環境が異なっていても新技術を想定した都道府県の農業経営指標を参考にするであろう．農業生産者にとっては，都道府県，作物，品種，栽培体系間で農業経営指標を比較検討し作付計画を策定することが望ましいが，その前提となる農業経営指標のオープンデータ化やデータ連携は進んでいない．

我々は，これまでデータ連携に必要な共通語彙として農作業基本オントロジー（Agriculture Activity Ontology，AAO）を構築公開してきた [朱 2016]．また連携したいデータの名称を AAO に対応付けることで，比較する際のデータ名の前処理が軽減することを示した[朱 2017]．都道府県の農業経営指標については，6 都道府県の水稲経営指標を対象に作業データ名が異なっていても，AAO を共通語彙に利用することでデータ連携が可能であることを確認した[竹崎 2017]．本研究では，作付計画の際に複数作物の経営指標を比較検討する場面を想定し，AAO を共通語彙とした作業時間の比較分析を提案する．

2. AAO に基づく作業時間の野菜栽培体系間比較

2.1 AAO の概要

AAO は農作業名を，目的，行為，対象，副対象，場所，手段，機資材，対象作物，時期，作業条件の 10 属性と，その属性値を用いて定義する．例えば，“刈取り”は「農作業∩∀目的．収穫∩∀行為．刈り取る」から「農作業であり，かつその目的は必ず収穫であり，行為は刈り取りであるものである」と定義される．AAO は，下位の農作業概念を上位概念から継承，細分化，追加した属性とその属性値で定義することで，明確な基準に従った階層構造をもつ．なお，最上位から第 5 階層程度までの上位概念は，農作業分類のために設定した抽象的概念であり，作業目的の値を細分化することで階層化している．AAO はイネ，ムギ，ダイズ，野菜，果樹等，主要作物で行われる農作業名（概念）を 475 収録する（2018 年 5 月時点）．

2.2 農業経営指標の概要

提供を受けた岩手県農業技術体系データ 2015（以下農業経営指標とする）のうち，野菜栽培体系の作業時間データを対象とした．岩手県の農業経営指標は，作物，品種，栽培技術，経営規模等を想定した栽培体系を設定し，作業ごとの時期や時間（以下作業時間），使用資材，使用機材等を整理したものである．作業は栽培暦に従って分類されており，最上位の「作業項目1」，その下位分類の「作業項目2」，さらに下位分類の「技術の内容」の 3 階層である．作業時間の集計単位である「技術の内容」のデータ名を AAO の農作業名に対応づけした（図1）．栽培暦に従い分類される「技術の内容」のデータ名は，作業目的を分類基準とする AAO と対応がとれないことがあった．そこで「技術の内容」に複数目的の作業が含まれた場合は，単一目的の作業に分けて AAO に対応付けし，その作業時間は，「技術の内容」全体の作業時間を，分類した作業数で除して算出した．

2.3 野菜栽培体系間比較

2.2 節の対応付けに基づき作業時間を集計する簡易な Excel マクロを作成し検証した．図1では AAO の第 3 階層の作物生産を目的とした作業を，第 4 階層の作物生育制御，作物生産環境制御，収穫調整，作物生産支援を目的とした作業に分類し作業時間を集計した結果を示す（図2，第 4 階層）．作物生育制御を目的とした作業（“作物生育制御作業”）はイチゴ（施設），ナス（露地），生食トマト（施設），キュウリ（施設），キュウリ（露地）で多く，果菜類であってもピーマン（施設，露地），加工トマト（露地）で少なかった．AAO の下位階層を確認し，“作物生育制御作業”が多い要因を解析した．第 5 階層では栄養成長期の作業（“栄養成長制御作業”）の時間が多く，“栄養成長制御作業”の下位階層（第 6 階層）では草姿調整を目的とした作業（“草姿

連絡先：竹崎あかね，農研機構革新工学センター，〒305-0856
茨城県つくば市観音台 1-31-1，E-mail: akane@affrc.go.jp

調整作業”）で時間が多かった．次に，労働平準化の重要な指標となる旬別作業時間を比較した（図3）．草姿調整作業が多い 5 つの栽培体系で比較したところ，イチゴでは，ピーク作業時間は少ないものの作業期間が長いことが特徴的であった．以上から，農業経営指標のデータ名が多様であっても AAO に対応付けて作業時間を集計することで作業目的や旬別に野菜栽培体系間の比較が可能になること，AAO の下位階層で集計すれば具体的作業での比較も可能になることを確認した．

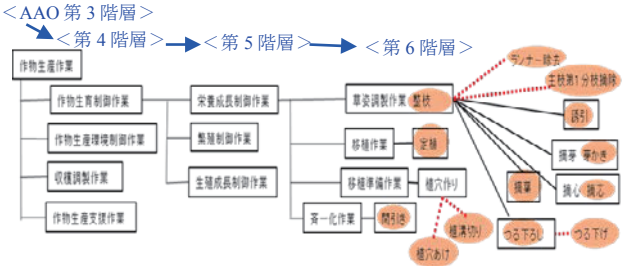


図1．農業経営指標データ名の AAO への対応付け例
□AAO 農作業名，同じ概念の用語は空白区切りで表示，●農業経営指標データ名
■AAO 農作業名と農業経営指標データ名が一致

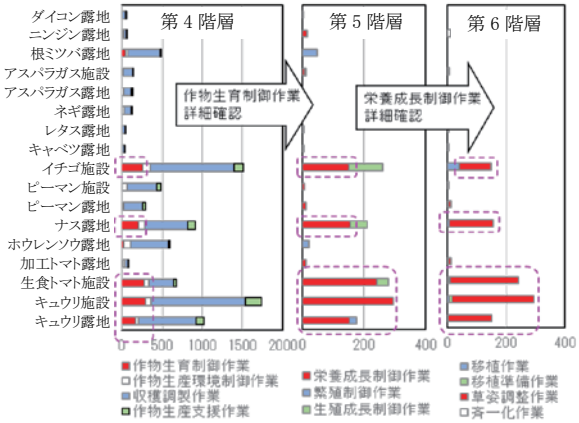


図 2．AAO に基づく作業時間の野菜栽培体系間比較
図1の対応付け結果に基づき，AAO 第 4 階層，第 5 階層（“作物生育制御作業”の下位階層），第 6 階層（“栄養成長制御作業”の下位階層）で作業時間を集計し比較．
横軸は 10a 当たりの作業時間を示す．

3. 農林統計に基づく作業時間の野菜品目間比較

農林水産省では，野菜，果樹，花き等の品目ごとの経営実態を把握するため，2017 年まで農業生産者の経営収支を全国規模で調査し，品目別経営統計調査として公表していた[農林水産省 2017]．品目別経営統計調査にある作業時間のデータ名（以下統計データ名とする）は，2.2 節の農業経営指標と比較すると作業時間の集計単位は大きいが，栽培暦に従った類似の分類であった．そこで，品目別経営統計調査を農業現場の実態を反映した全国的基準と位置づけ野菜品目間比較を試みた．まず統計データ名と AAO 農作業名を対応づけした後（図4），統計データ名に対応する農業経営指標のデータ名を，AAO との対応関係（図1）に基づき抽出した．統計データ名単位で作業時間を集計することで，品目別経営統計調査を基準とした野菜品目間比較が可能となった（データ略）．以上から AAO に対応づけた基準での作業時間の比較も可能であることを確認した．

4. 考察と今後の課題

本研究では 1 県のデータを対象に AAO に基づいた作業時間の野菜栽培体系間比較を行った．今後は作業時間の都道府県間比較，農業生産者と AAO に対応付けた指標との比較を行い，AAO に基づく作業時間の解析事例を蓄積する予定である．今回作成した Excel マクロで作業時間を集計する際には，あらかじめデータ名を手動で AAO に対応付けする必要があった．このような作業は，今後事例が蓄積し対応付けルールの学習が進むことで，自動化する可能性がある．

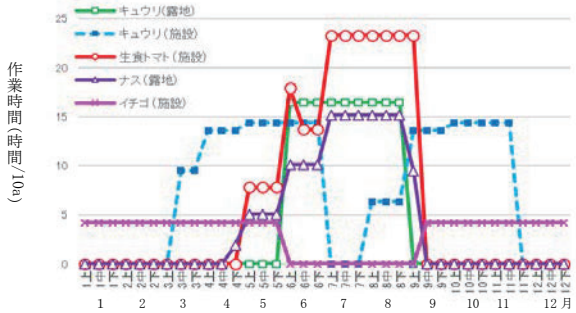


図 3．AAO に基づく旬別作業時間の野菜栽培体系間比較
図1の対応付け結果に基づき，“草姿調整作業”の作業時間を旬別に集計し比較．

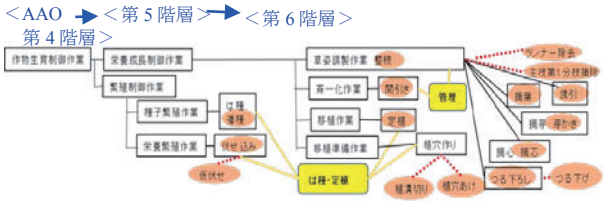


図 4．統計データ名の AAO への対応付け，および農業経営指標データとの関係例
□AAO 農作業名，同じ概念の用語は空白区切りで表示，●農業経営指標データ名
■AAO 農作業名と農業経営指標データ名が一致，■統計データ名

5. まとめ

作付計画の際に複数作物の経営指標を比較検討する場面を想定し，AAO を共通語彙に利用した作業時間分析を提案した．農業経営指標の作業時間は，データ名を変換しないと野菜栽培体系間で比較ができなかった．AAO に対応付けて作業時間を集計することで，データ名を変更せずに，作業目的や旬別による野菜栽培体系間の比較が可能になること，AAO の下位階層で集計すれば具体的作業の比較も可能になること，AAO に対応づけた他の基準での比較も可能であることを確認し，作業時間の分析が簡便化すると結論づけた．

謝辞

本報告は内閣府～農水省予算により生研センターが管理運営する「SIP（戦略的イノベーション創造プログラム）次世代農林水産業創造技術」での研究成果に基づく．

参考文献

[農林水産省 2017] 農林水産省:農業経営統計調査，平成 19 年度品目別経営統計，2017. <http://www.maff.go.jp/j/tokei/kouhyou/noukei/hinmoku/> 2019 年 2 月 8 日参照.

[朱 2016] 朱成敏，小出誠二，武田英明，法隆大輔，竹崎あかね，吉田智一：記述論理に基づく農作業オントロジーの 設計と応用，第 38 回人工知能学会セマンティックウェブ とオントロジー研究会（SIGSWO），06, 2016.

[朱 2017] 朱成敏，武田英明，法隆大輔，竹崎あかね，吉田智一：標準語彙に基づく農業データの連携と統計への活用，第 31 回人工知能学会全国大会，2E3-OS-36a-4, 2017.

[竹崎 2017] 竹崎あかね，法隆大輔，朱成敏，武田英明,吉田智一：農作業基本オントロジーを基盤とする水稲技術経営指標データの連携，第 31 回人工知能学会全国大会，2E3-OS-36a-2, 2017.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

農業ナレッジグラフを用いた営農記事からの農作物の関係の発見

A Method for Judging the Semantic Similarity between Crops from Farm Management Articles based on Agricultural Knowledge Graph

朱 成敏 *1 Sungmin JOO　武田 英明 *1*2 Hideaki TAKEDA　竹崎 あかね *3 Akane TAKEZAKI　吉田 智一 *3 Tomokazu YOSHIDA

*1国立情報学研究所 National Institute of Informatics　*2総合研究大学院大学 SOKENDAI University(The Graduate University for Advanced Studies)

*3農業・食品産業技術総合研究機構 National Agriculture and Food Research Organization

Terminology in the agricultural field are difficult to process data automatically since they are often used in various expressions. It is especially difficult for crops whose names vary depending on the agriculture activity, edible parts or cultivation method in which various names are used. This study suggests semantic analysis using agricultural knowledge graph to solve these problems. This study also confirmed that the use of knowledge graph's semantic structure not only enabled the extraction of agriculture activity and the crop name clearly, but also enabled agriculture-specific analysis.

1. はじめに

農業分野の用語には同じ意味にも拘らず，様々な表記が存在している．例えば，農作業の一つである「荒代」は「かじり」や「荒代かき」のように習慣によって異なる名称が使われる場合や農業 ICT システムでは「あらしろ」や「荒しろ」など異なる表記で処理される場合もある．農作物名は地域や利用部位，栽培方法などの基準によって名称が異なる場合や，品種名や食品名が農作物の名称として使われる場合もある．ICT システムのデータや関連文書にはこのような場合，異なる用語として処理される可能性がある．こういった農業分野の語彙が持つ表記の多様性は農業情報の利活用において妨げとなる．

一方，様々な表記が収録されていて，かつその意味関係まで記述されている知識体系があればそれを参照することによって上記の問題は解決できる．例えば,「荒代」の場合は同義語として「荒代かき」と「かじり」を持ち,「あらしろ」や「荒しろ」などで表記される基準があれば，ICT システムはこれらの名称を一括して「荒代」として処理することができる．また，農作物の場合も「レッドクイーン」や「シャインマスカット」のような品種名と一般的な名称である「ブドウ」との関連性が定義された基準があれば，システムは「レッドクイーン」と「シャインマスカット」を「ブドウ」の一種として扱うことができる．このように名称に対する基準があれば農業データやコンテンツに対する正確な情報の抽出や分析など高度な利用が可能となる．

そこで，本研究では農業語彙が持つ表記の多様性に対応し，農業コンテンツの分析に農業ナレッジグラフを用いることを提案する．農業ナレッジグラフは農業分野を対象に構築された知識体系であり，農業における概念と概念間の関係性について定義をしている．それぞれの概念は表記を持っており，表記は見出し語である代表表記以外にも同義語も収録されている．また， 関連情報体系と連携されており，様々な情報を活用することもできる．本研究では農業関連の文書を対象とし，農業ナレッジグラフを用いて農作業と農作物の名称を抽出する．そして，抽出された農作業と農作物の共起を用いて農作業と農作物の関連性を発見する．発見された関連性から農作物同士の関連性を推測する．

連絡先: 朱成敏，国立情報学研究所，〒 101-8430 東京都千代田区一ッ橋 2-1-2, joo@nii.ac.jp

2. 農業ナレッジグラフ

筆者らは農業 ICT システムのデータ連携のために農業分野のナレッジグラフを構築してきた [朱 18]．本章では農作業を体系化した農作業基本オントロジー (AAO，Agriculture Activity Ontology) と農作物の語彙を整理し，Linked Data 化した農作物語彙体系 (CVO，Crop Vocabulary) について概観し，それぞれの意味構造について述べる．

2.1 農作業基本オントロジー

農業 ICT システム間の相互運用性を確保するために農作業の標準語彙として筆者らは農作業基本オントロジーを開発し，推進してきた [朱 16]．農作業基本オントロジーは農作業概念を定義するために目的，行為，対象，副対象，場所，手段，機資材，作物，時期，作業条件の 10 項目を属性を用い，それぞれの属性が持つ値の包含関係から農作業概念を体系化した．また，記述論理による設計を行い，矛盾や重複のない論理性も確保した．最新版である Ver2.01 には 475 語の農作業名称が収録されており，それぞれの農作業名称は固有の名前空間 (URI) を持つ．共通農業基盤 *1 では農作業基本オントロジーを用いた語彙変換 API，用語集などのサービスを提供しており，Turtle/RDF と CSV 形式の関連データも公開している．農作業基本オントロジーでは農作業に対し代表表記，同義語などの別名，英名の３つの表記を与えている．また，これらの表記は共通農業基盤にて提供している語彙変換 API を用いて容易に変換することができる．図 1 は「荒代」の定義と表記を表した例である．

2.2 農作物語彙体系

筆者らは農作物名称の標準語彙として農作物語彙体系を構築した [竹崎 17]．農作物語彙体系は植物学的分類に基づいて様々な農作物名を分類し，それぞれの農作物名は同義語，英名，学名を基本情報として収録している．また，既存語彙である「農薬登録における適用作物名」の作物名,「農産物等の食品分類表」の食品名,「日本食品標準成分表」の食品名，情報

*1 CAVOC, http://cavoc.org

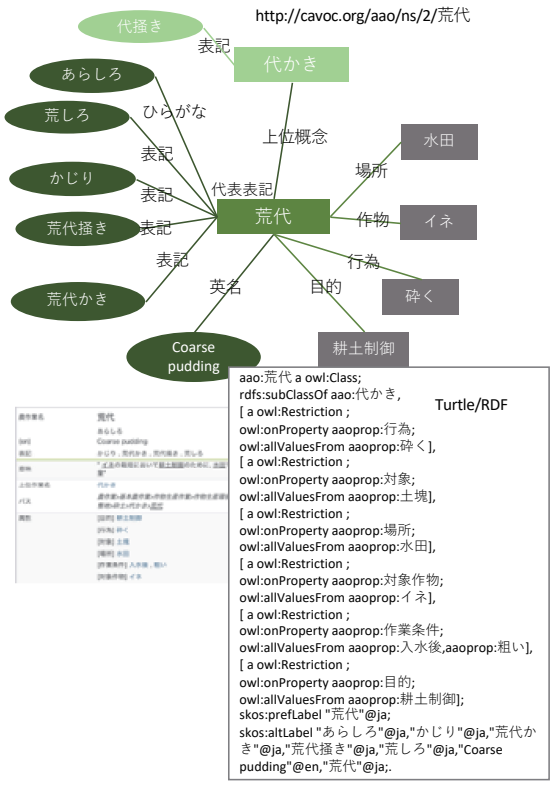


図 1: 「荒代」における情報の意味関係.

体系である NCBI の Taxonomy ID，Wikipedia の項目名との対応関係を調査し，関連情報として収録した．最新版である Ver1.52 では 1,249 語が収録されており，共通農業基盤にて公開している．それぞれの農作物名は固有の名前空間を持ち，収録情報を閲覧することが可能である．これらの情報は CSV と Turtle/RDF 形式でも公開されており，機械可読性も確保している．農作物語彙体系では農作物名に対し，英名や科名，学名のような植物学的情報，別名，一般的な名称にあたる総称を上位概念として持つ．また，既存の語彙体系に収録されている名称も収録されており，様々な語彙の意味関係を把握することができる．図 2 は「シャインマスカット」に関連する意味関係を表す例である．

3. 農作物間の関係性発見

本章では実際の新聞記事のデータから発見された農作業名称と農作物名称の関係性について考察し，作物の生産過程の類似性から農作物同士の関係性を判定する手法について述べる．

3.1 営農記事の分析

本研究では農業コンテンツとして営農に関連する記事を用いる．営農関連記事は研究目的での利用許可を得た 2014 年 4 月から 2017 年 3 月までの日本農業新聞 *2 の営農面の記事 3,479 件である．オープンソース形態素解析エンジンである Mecab[MeCab 13] を用いて本文の形態素分析を行い，品詞が名詞の場合に農作業基本オントロジーと農作物語彙体系の名称を抽出した．抽出対象となる名称は代表表記と同義語を含む全ての関連語彙である．実行結果，99 件の農作業名，292 件の農作物名がそれぞれ 1,371 件，1,265 件の記事から発見された．

*2 日本農業新聞，https://www.agrinews.co.jp/

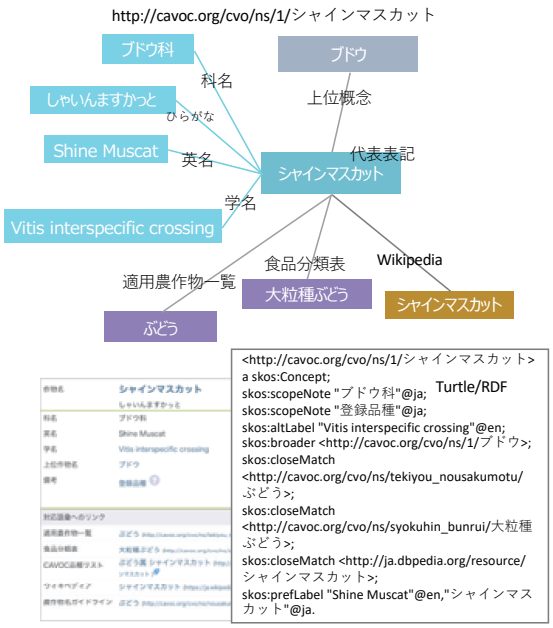


図 2: 「シャインマスカット」における情報の意味関係.

3.2 表記の調整

そして，出現頻度と共起頻度を確認するために発見された農作業と農作物の名称に対して表記の整理を行った．同じ作業にも関わらず異なる表記で書かれており，別の作業として処理される場合に備えて，代表表記に変換する処理を農作業基本オントロジーの語彙変換 API を用いて行った．例えば，同じ農作業である「播種」と「種まき」がそれぞれ異なる名称として含まれていた記事が一部あり，農作業基本オントロジーの代表表記である「は種」にまとめた．農作物の場合は品種や用途，利用部位によって異なる名称になる場合が多く，代表表記に統一するために総称にあたる上位概念の名称を用いてまとめた．この処理によって「シャインマスカット」や「レッドクイーン」など品種名は総称である「ブドウ」として処理された．

3.3 関連農作業の抽出

次は農作物名と農作業名の共起を確認した．農作物単位で共起した農作業名を抽出し，関連農作業のリストを作成した．最も多くの農作業と共起した農作物は「イネ」であり，52 件であった．一方，共起した農作業が 5 件以下の農作物も 79 件があり，農作物 1 件にあたり農作業の平均共起数は 12.49 件であった．

3.4 農作物間の類似性判別

本研究では関連農作業名のリストを用いて農作物同士の類似性を判別するために Jaccard 係数を用いる．共起した農作業の集合 $Activities_a$ を持つ農作物 $crop_a$ と集合 $Activities_b$ と共起した農作物 $crop_b$ の Jaccard 係数は次のように計算できる．

$$Jaccard(crop_a, crop_b) = \frac{|Activities_a \cap Activities_b|}{|Activities_a \cup Activities_b|} \quad (1)$$

本実験では共起した農作業名称が 5 件以上の農作物を対象にし，Jaccard 係数を求めた．表 1 はコムギ，タマネギ，トマトと関連度がある上位 5 件の農作物の順位である．比較のた

め農作物間の共起数による順位も表示した．実験の結果，コムギと最も類似性を持つ農作物はオオムギであり，タマネギはネギ，トマトの場合はナスに判明された．表 2 はコムギ，タマネギ，トマトと最も関連があると判明した農作物の関連農作業のリストである．

表 1: 農作物間の類似判別.

(a) コムギ

#	Jaccard 係数		共起数	
1	オオムギ	0.6563	イネ	70 回
2	ダイズ	0.6364	ダイズ	47 回
3	トウモロコシ	0.5938	ジャガイモ	17 回
4	イネ	0.5789	テンサイ	15 回
5	キャベツ	0.5428	ソバ	14 回

(b) タマネギ

#	Jaccard 係数		共起数	
1	ネギ	0.7274	イネ	11 回
2	ホウレンソウ	0.6957	トマト	11 回
3	ジャガイモ	0.6400	ネギ	9 回
4	レタス	0.6087	ニンジン	9 回
5	キュウリ	0.5770	ダイズ	6 回

(c) トマト

#	Jaccard 係数		共起数	
1	ナス	0.7083	イチゴ	41 回
2	ピーマン	0.6207	キュウリ	40 回
3	イチゴ	0.6061	イネ	32 回
4	ウンシュウミカン	0.5833	ナス	29 回
5	CO	0.5517	チャ	19 回

コムギは 31 件の農作業と，オオムギの 22 件の農作業と共起しており，22 件全部コムギの農作業と一致した．一方，イネと共起した 52 件の農作業の中でコムギと一致する農作業は 25 件だった．一致した農作業名はオオムギより 2 件多かったが，27 件の農作業が一致していないことが判明したので，農作業の類似性はオオムギとコムギの方が強い類似性を持つと考えられる．採種，追肥，融雪の 3 つの農作業はコムギのみ共起していたことがわかった．

タマネギは 21 件の農作業名との共起が確認された．ネギの 17 件の共起農作業の中で排水作業を除く 16 件が一致した．一方，イネは 20 件の共起農作業の一致が確認されたが，関連性がない農作業が 32 件発見された．今回の実験ではコムギとタマネギの場合，最も共起した農作物はイネであることがわかった．イネは農作業名称が発見された 2,505 件の記事の中で 674 件の記事に出現しており，共起頻度が農作物の中で最も多かったと考えられる．

トマトは 28 件の農作業と共起しており，選別作業，意見交換，土寄せの 4 件以外はナスとイチゴの農作業と一致した．ナスは 22 件農作業の中で 20 件が，イチゴは 23 件の中で 19 件がトマトの農作業と一致している．実際，イチゴは Jaccard 係数による順位でも 3 番目である．

3.5 農作物間の類似性による関連記事の提案

前節で求めた農作物間の類似性を用い，類似性の高い農作物に関する記事を関連記事として提案した．Jaccard 係数の上位 3 件までの農作物を選択し，該当する農作物の出現頻度が高い記事を掲載日が近い順で提示した．図 3 は「タマネギ」に関する記事の例である．1 の (b) の結果より「ネギ」と「ホウレンソウ」,「ジャガイモ」に関する記事が「タマネギ」に関する記事と共に関連記事として表示されている．

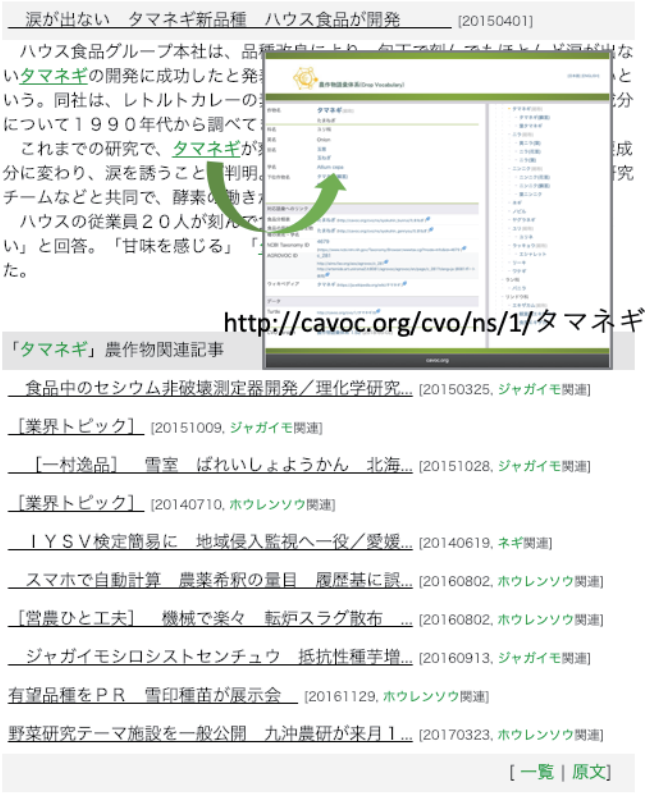


図 3: 農作物間の類似性による関連記事リスト.

4. 考察と今後の課題

今回，農作物名と共起した農作業の集合を Jaccard 係数を用いて類似性を確認した．この場合，どちらかの農作業の数が少なく，さらに出荷作業や収穫作業のような栽培において基本となる農作業名の場合は強い類似性を見せた．そのため比較的に関連農作業の数が少ない農作物が上位になることが多かった．一方，麦踏みのように特定農作物のための農作業やハウス栽培の農作業など農作物の栽培特徴が反映されたことも確認できた．

営農記事は栽培や経営など幅広い情報が書かれているため農作業と農作物の関連性が一般的な記事より多く含まれていると考えられる．この場合は共起頻度も関連性を判別する重要な因子となる．今回は Jaccard 係数を用いて類似性を判別したが，共起頻度を重み付けとして取り込む場合，改善の可能性がある．今後の課題として取り組む予定である．

今回は農作業基本オントロジーの語彙変換を用いて同義語を代表表記に整理し，また，農作物語彙体系の階層構造を用いて利用部位や栽培方法などによる別名と品種名を総称に変換して共起を確認した．この調整作業により抽出する単語の数や集計による手間が軽減されたと考えられる．

農作業名称の中では収穫や運搬のように農業以外にも使われる単語が多く含まれている．こういう単語の共起は本研究の提案手法の性能に影響を与える可能性が高い．形態素解析方法の改善や文脈情報を用いて作業名を特定する必要がある．今回は農作業に注目して農作物の類似性判別したが，今後は農機具や資材など様々な農業関連語彙を用いて類似性の判別を行いたい．また，そのために農業ナレッジグラフの拡張する予定して

表 2: 農作物名と共起した農作業名のリスト.

農作物名	共起した農作業名
コムギ	収穫作業, 評価作業, 受粉, 施肥, 雑草抑制作業, 出荷作業, は種, 移植作業, 定植, 育苗, 出荷調製作業, 排水作業, 耕耘, 選別作業, 稲刈り, 採種, 生物制御作業, 基肥, せん定, 追肥, 出芽, 中耕, 運搬作業, 土寄せ, 刈取り, 田植え, 麦踏み, 融雪, 誘引, 意見交換, 砕土
オオムギ	移植作業, 育苗, 出荷作業, 収穫作業, 出荷調製作業, 耕耘, 施肥, 排水作業, 雑草抑制作業, 選別作業, 評価作業, は種, 出芽, 生物制御作業, 基肥, 麦踏み, 中耕, 運搬作業, 定植, 刈取り, 稲刈り, 田植え
イネ	追肥, 生物制御作業, 収穫作業, 意見交換, 評価作業, せん定, 受粉, 摘果, 施肥, 雑草抑制作業, 出荷作業, は種, 出芽, 中耕, 土寄せ, 定植, 育苗, 耕耘, 鎮圧, 排水作業, 給水作業, 出荷調製作業, 移植作業, プラスチックマルチング, 基肥, かん水, 物理的雑草抑制作業, 砕土, 運搬作業, 選別作業, モニタリング作業, 稲刈り, 間伐, 誘引, 包装, 代かき, 摘粒, 清掃作業, 草姿調整作業, 田植え, 催芽, 保水作業, 刈取り, 暖房, ハロー作業, 風制御作業, 遮光, 脱穀, 精米, 中干し, 冷房, 手取除草
タマネギ	収穫作業, は種, 中耕, 土寄せ, 施肥, 基肥, 追肥, 生物制御作業, 移植作業, 定植, 育苗, 出荷作業, 選別作業, プラスチックマルチング, 運搬作業, 評価作業, 給水作業, かん水, 出荷調製作業, 遮光
ネギ	は種, 定植, 排水作業, 出荷調製作業, 移植作業, 育苗, 出荷作業, 遮光, 収穫作業, 生物制御作業, 選別作業, 土寄せ, 給水作業, かん水, 評価作業, 施肥
イネ	追肥, 生物制御作業, 収穫作業, 意見交換, 評価作業, せん定, 受粉, 摘果, 施肥, 雑草抑制作業, 出荷作業, は種, 出芽, 中耕, 土寄せ, 定植, 育苗, 耕耘, 鎮圧, 排水作業, 給水作業, 出荷調製作業, 移植作業, プラスチックマルチング, 基肥, かん水, 物理的雑草抑制作業, 砕土, 運搬作業, 選別作業, モニタリング作業, 稲刈り, 間伐, 誘引, 包装, 代かき, 摘粒, 清掃作業, 草姿調整作業, 田植え, 催芽, 保水作業, 刈取り, 暖房, ハロー作業, 風制御作業, 遮光, 脱穀, 精米, 中干し, 冷房, 手取除草
トマト	接ぎ木, 移植作業, 定植, 生物制御作業, 収穫作業, 施肥, かん水, 出荷作業, は種, 排水作業, 育苗, 評価作業, 暖房, 出荷調製作業, 換気, 選別作業, 遮光, 包装, 意見交換, モニタリング作業, プラスチックマルチング, 送風, 誘引, 受粉, 保温, 土寄せ, 草姿調整作業, 冷房
ナス	接ぎ木, 生物制御作業, 収穫作業, 出荷作業, 定植, せん定, 施肥, 保温, プラスチックマルチング, 育苗, 誘引, 受粉, 出荷調製作業, 暖房, 換気, かん水, 雑草抑制作業, 評価作業, 移植作業, 意見交換, 草姿調整作業, モニタリング作業
イチゴ	収穫作業, 定植, 生物制御作業, 育苗, 出荷作業, 出荷調製作業, プラスチックマルチング, 暖房, 評価作業, 換気, 予冷, かん水, 草姿調整作業, せん定, 施肥, 送風, 受粉, 冷房, 包装, は種, 遮光, 追肥, 摘果

いる.

5. おわりに

農作物の栽培には特定の農作業，または複数農作業の組み合わせが用いられており，農作物の特徴として用いることができる．そこで，本研究では農作業に注目して農作物間の類似性の判別を提案した．営農記事から農作業名称と農作物名称を抽出し，共起を基準にその関連性を定義した．農作業と農作物の抽出は農業ナレッジグラフを用いて行い，同義語や意味関係を考慮した名称の抽出ができた．今後は共起頻度による重み付けや農作業と農作物以外の農業関連語彙を用い，本研究の提案手法を改善していきたい.

謝辞

本研究で利用した記事データは株式会社日本農業新聞により提供されたものである.

本研究（の一部）は，総合科学技術・イノベーション会議のSIP（戦 略的イノベーション創造プログラム）「次世代農林水産業創造技術」（管理法人：農研機構生物系特定産業技術研究支援センター）の支援を受けて行った.

参考文献

[朱 18] 朱成敏, 武田英明, 竹崎あかね, 吉田智一： 農業データ連携のためのナレッジグラフに基づく標準語彙の運用, 2018 年度人工知能学会全国大会 (第 32 回), No. 2G2-OS-10b-01, 2018.

[朱 16] 朱 成敏, 小出 誠二, 武田 英明, 法隆 大輔, 竹崎 あかね, 吉田 智一：記述論理に基づく農作業オントロジーの設計と応用, 第 38 回人工知能学会セマンティックウェブとオントロジー研究会（SIGSWO）, 06, 2016.

[竹崎 17] 竹崎あかね, 朱成敏, 武田英明, 吉田智一： 農業 IT システム間のデータ連携を促進する農作物語彙体系の構築, 電子情報通信学会技術研究報告，知的環境とセンサネットワーク研究会, vol. 117, no. 310, ASN2017-J27, pp. 98-99, (2017).

[MeCab 13] MeCab: Yet Another Part-of-Speech and Morphological Analyzer, <http://taku910.github.io/mecab/> 2019 年 2 月 1 日参照.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

雑味の視点からみた昆布だしの光分析手法による解析

Analysis of Konbu-Dashi by multi-band optical method from a viewpoint of miscellaneous taste

亀岡孝治[1] 田口拓実[1] 西川恵梨子[1] 伊藤良栄[1] 橋本篤[1] 湯川徳之[2] 大引伸昭[2]
Takaharu Kameoka Takumi Taguchi Eriko Nishikawa Ryoei Ito Atsushi Hashimoto Noriyuki Yugawa Nobuaki Obiki

1. 三重大学大学院生物資源学研究科
Graduate School of Bioresources, Mie University
2. エコール辻大阪(辻調グループ)
Tsuji Culinary Institute

When evaluating Dashi in Japanese cuisine, the chef uses the word umami and miscellaneous taste. Umami has been used as an important factor. On the other hand, there is no definite definition of miscellaneous taste, and quality evaluation viewed from miscellaneous taste has hardly been performed. In this research, therefore, the objective was to clarify the characteristics of the Konbu-Dashi which the chef is conscious of miscellaneous taste. We established a quality evaluation method of Konbu-Dashi by using Quantitative Descriptive Analysis (QDA) used to express the overall taste perceived by human beings, and made a comprehensive evaluation. As a result of the multi-spectroscopic analysis, the characteristic of Konbu-Dashi that felt miscellaneous taste had "little umami, relatively much minerals and saccharides" was confirmed. In addition, in the QDA, evaluation terms were created by using 32 kinds of Konbu-dashi to capture the characteristics of both delicious and not delicious Konbu-dashi. From the results using the QDA, 1) focusing on fragrance and flavor as an approach to harshness, 2) the necessity to consider the influence of minerals such as potassium, etc. were derived. From now on, it is necessary to develop machine learning and analysis using AI.

1. はじめに

日本料理の出汁を評価する際，料理人はうま味と雑味のバランスを重視する．基本五味として発見されたうま味は数多くの研究が行われている(たとえば，二宮 2010)が，雑味の視点から昆布出汁を評価する研究はほとんど行われていない．素材を生かすことを重視する日本料理において，うま味と雑味のバランスを考えた調理法を探ることは有用であると考えられるが，そのためには雑味の特徴を明らかにする必要がある．

そこで本研究では，昆布出汁における雑味の特徴を明らかにすることを目的とした．具体的には，高速液体クロマトグラフィー(HPLC)・中赤外分光分析・蛍光X線分光分析を用いたアミノ酸・糖・ミネラル成分の分析と，ヒトが感じる総合的な味覚を表現するための記述型官能評価法(QDA 法)(Stone 1974)の確立およびその手法による出汁の品質評価を行った．

2. 調理条件と呈味成分，雑味評価の関係

調理条件を変えた 5 種類の昆布出汁に対して，雑味評価と機器分析を行い，雑味を感じる昆布出汁の呈味成分の特徴把握を試みた．出汁材料には真昆布(北海道南茅部産) 20g，南アルプスの天然水(サントリー社製) 1000g を用い，図 1 に示す調理方法で出汁を作成した．各出汁サンプルに対して，有機物の分析に中赤外分光分析，ミネラル情報取得のために蛍光 X 線分光分析，うま味成分(アミノ酸)分析(グルタミン酸:Glu，アスパラギン酸:Asp)のために HPLC による分析を行った． 出汁の品質決定において重要な雑味の有無を，辻調理師専門学校の料理人が判定した．すべての昆布出汁で顕著なカリウムの存在が認められた．は図 2 に一例として Rh 基準で正規化した蛍光 X 線分光スペクトルのカリウムのピーク(3.4keV)を示した．

雑味評価の結果，刻んだ昆布による昆布出汁⑤⑥に雑味が認められた．昆布出汁①②⑦⑧には雑味が感じられなかったが，有機成分バランス， 元素バランス， アミノ酸量の単独データでは，昆布出汁⑤⑥との違いを把握することが難しかった．そこで，中赤外分光スペクトルの糖や脂質などのピーク波数の吸光度，

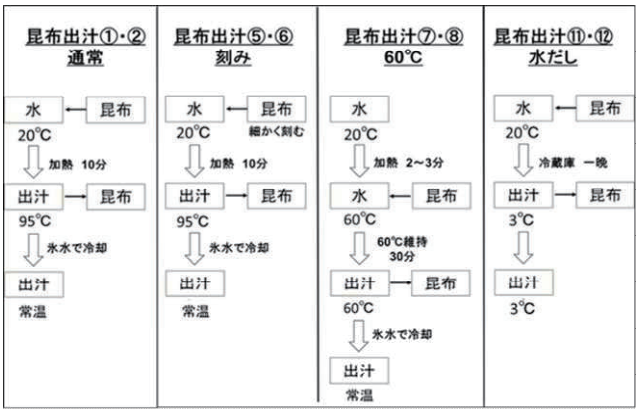


図 1 昆布出汁の調理方法

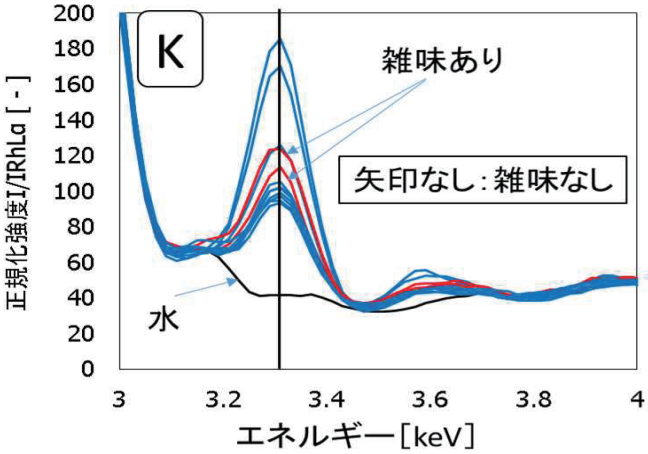


図2 各出汁の蛍光 X 線分光スペクトル(3.4keV)

蛍光 X 線スペクトルのカリウム(K)，塩素(Cl)に対応するエネルギーでの正規化強度，及び Asp･Glu の濃度をデータセットとして用い，正規化の後に主成分分析を行った(図3)．また，図には因子負荷ベクトルを示した．雑味を感じた昆布出汁⑤⑥には「旨味が少なく，ミネラルや糖が相対的に多い」という特徴が確認された(図3)．この結果，雑味はうま味成分が少ない時に生じる可能性があることが示唆されると共に，ミネラルや糖も雑味に関与している可能性が示された．

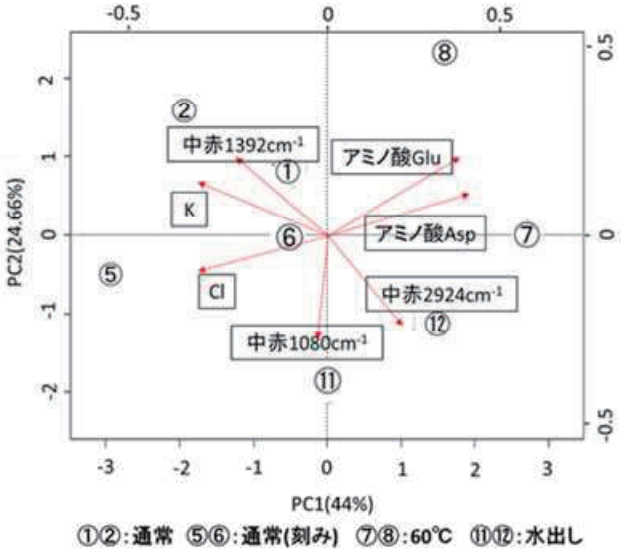


図 3 昆布出汁の主成分分析結果

3. 昆布出汁の官能評価手法の構築

昆布出汁の詳細な官能特性を表現するため，QDA 法を用いた評価手法の構築を試みた(表 1)．本研究では，おいしくない昆布出汁の特徴も捉えるために，出汁に向いていない昆布やタブーとされている調理条件を含めた 32 種類の昆布出汁を用いて評価用語の作成を行った．材料には，株式会社神宗から取り寄せた真昆布，利尻昆布，日高昆布，羅臼昆布，厚葉昆布(釧路産， 根室産)，長昆布(釧路産，根室産)の 8 種類と南アルプスの天然水(サントリー社製)を用いた．それぞれの昆布の形状(そのまま，刻み)と調理方法(60℃1 時間，水出し)を組み合わせ，計 32 種類のだしを作製した．パネリストは辻調理師専門学校の教員と株式会社神宗の職員がおこなった．特性表現用語を作成し， 香り 9 個，風味 17 個，味 6 個，食感 3 個の計 25 個の評価用語が得られた(表 2)．言葉だしと全体討議を経て，パネリストの 2/3 が共通認識を持つ用語を対象に，実際のサンプルを用いた強度評価の練習を行った．その後，パネリストが共通した認識を有した特性表現用語を採用した．構築した評価用語を用いて， 真昆布 60℃1 時間抽出の出汁，刻んだ真昆布 60℃1 時間抽出の出汁， 長昆布の水出汁の比較を行った．評価値を主成分分析した結果を図4に示す．図中の p はパネリスト番号を示す． 主に第 1 主成分にて，真昆布と長昆布の差が明確であった． 第 1 主成分は，フルーティーな風味と甘味が負に寄与し，その他の項目が正に寄与しており，スコアが低いほど上品な真昆布の特徴が表れていると考 えられる．また，真昆布出汁における刻みの有無の影響は，全体の傾向では確認できないが，各パネリストのスコアに着目すると，共通の傾向(刻んだ出汁の方がスコアが高い)が認められた．以上の結果から，構築した QDA 法にて品種の識別および刻みの影響を捉えることができ，QDA 法による昆布出汁評価の有効性が示唆された．

表1 QDA 法の構築手順

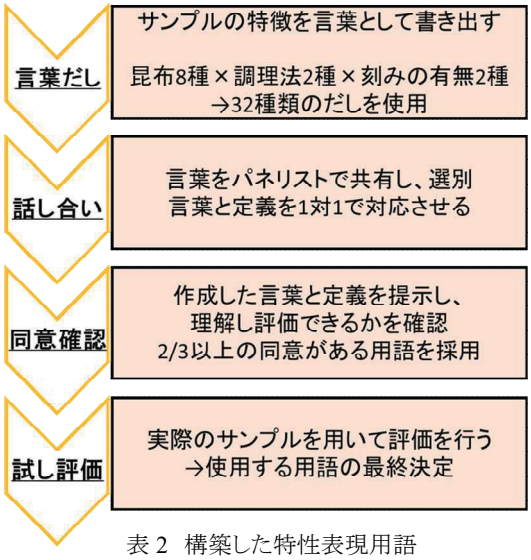


表 2 構築した特性表現用語

属性	特性表現用語	定義もしくはリファレンス
香り	潮の 香り	海岸に行ったときに感じる、臭くない海の 香り
	磯の 香り	生臭い岩場の 香り、漁港の 香り、腐敗した 海藻
	木の 香り	加工した 木材、新築の 家、材木、丸太を積んだ 匂い
	落葉のような 香り	湿った 落ち 葉、木材
	青草の 匂い	畳、イグサ
	硫黄の 匂い	温泉、ゆで 過ぎた 卵黄
	生臭い	魚のドリップ
	ムレタ感じ	湿気がある 所、じめじめした 感じ、鼻につく
	金属臭	鉄が 錆びたような
風味	潮の 風味	海岸に行ったときに感じる、臭くない海の 香り
	塩素の 風味	プール
	フルーティーな 風味	柑橘系
	大豆、味噌の 風味	みそ
	香ばしい 風味	醤油のような、日本酒を焦がしたような、甘味を含む
	ニンジンのような 風味	生ニンジン、野菜の断面、野菜の 発酵手前
	粘土のような 風味	油粘土
味	甘味	淡い甘味、自然な甘み、板ガムの 最初の 甘味
	塩味	塩、塩化ナトリウム
	苦味	塩化カリウム、バナナの 白いスジ、ミネラル 感、硬水
	うま 味	グルタミン 酸、アスパラギン 酸
	化学調味料のような 味	人工的な 調味料の 味
	後味	口に 残る 感覚の 総称
食感	ぬめり	オクラのぬめり
	濃厚さ	味の 濃さ
	ミネラル 感	軟水と硬水の 違いのような、口当たりが 重い、舌に 残る 感覚

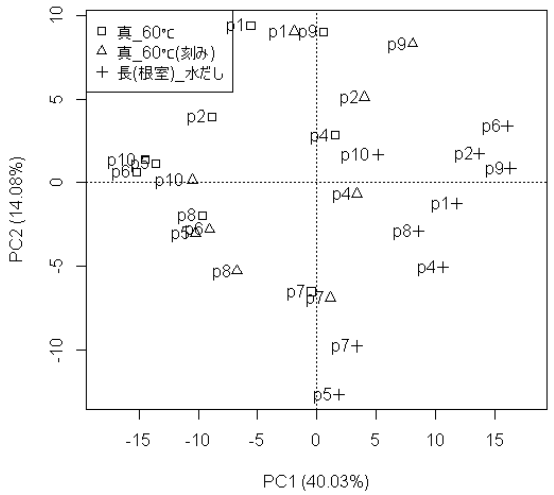


図 4 評価値の主成分分析結果

4. 雑味の検証

機器分析の結果，作業仮説として得られた雑味の特徴，すなわち「雑味はうま味成分が少ない時，他の成分が相対的に際立つことで感じる相対的な味覚」の仮説検証実験を行った．

予備実験を行い，うま味が少なく雑味を感じることが確認された日高昆布の出汁に対して，うま味成分の Glu と Asp を添加し，雑味の減少や消失を検討した．うま味成分の添加には，グルタミン酸ナトリウム水和物（日理化学株式会社，食品添加物規格），アスパラギン酸ナトリウム水和物（日理化学株式会社，食品添加物規格）を用いた．評価の際は，うま味成分が多く含まれ，雑味の少ない真昆布出汁を基準として使用した．真昆布出汁，日高昆布出汁，うま味成分を添加した日高昆布の雑味評価および QDA 法による官能評価を行うことで，雑味の仮説検証や特徴把握を試みた．

12 名のパネリストによる雑味評価では，真昆布に雑味を感じた人は 3 名と少なく，日高昆布に雑味を感じた人が 7 名と多いことが示された．さらに，日高昆布に Glu，Asp を添加しても雑味を感じた人数は 6 名でほとんど減少しなかった．よって，うま味成分の添加によって成分のバランスを変化させても雑味に影響はなく，雑味が相対味覚という仮説は立証できなかった．

QDA 法による各出汁の評価値を主成分分析した結果を図 5 に示す．第 1 主成分（PC1）の寄与率が 49．93%，第 2 主成分（PC2）が 11．84%で，両軸で累積寄与率 61．77%の説明が可能となった．PC1 では，すべての特性が正に寄与した．その中でも，「甘味」「うま味」の寄与は小さい傾向であった．

PC2 では，香り・風味が正に寄与し，味・食感が負に寄与した．ここから， PC1 は「甘味やうま味を除く総合的な味覚の強度」，PC2 は「香りと味のバランス」と解釈した．図4より，PC1 に関して，真昆布は負の方向に，日高昆布は正の方向に配置し，日高昆布は真昆布よりも「味の強度」が全体的に強いことが示唆された．しかし， 日高昆布のみに着目すると，Glu や Asp を添加しても PC1 のスコアに変化はなく，「味の強度」の変化は示されなかった．つまり，うま味成分を添加することによる総合味覚の強度に変化は生じないことが示された．PC2 に関しては，日高昆布が負の方向に大きくばらついた．この結果，日高昆布は香りに対して味の強度が強く，バランスの崩れた出汁であることが示唆された．

日高昆布のみに着目すると，PC1 と同様に，Glu や Asp を添加しても変化することはなく，「過度に強い味や食感」が弱まることは示されなかった．つまり，うま味成分を添加することによる総合味覚のバランスは補正されないことが示された．また，この実験系における成分分析の結果と各官能評価項目の評価値からは，雑味を感じる出汁にはミネラルが多く含まれること，塩味や苦味が強い傾向が確認された．

本研究では，雑味の定義につながる明確な結果は得られなかったが，雑味へのアプローチとして，QDA 法から裏付けされるように香り・風味に着目することや，昆布出汁の成分に関しては蛍光 X 線分光分析で計測されたカリウムを主とするミネラル類の影響，中赤外分光分析で得られた多糖類の影響などを考慮する必要性が明らかとなった．

QDA 法からも裏打ちされたが，昆布出汁の評価はフレーバー（味覚と香りの同時知覚）で行われるため，香りの評価は極めて重要であると考えられる．赤外分光分析では，同じ計測手法で液体成分と香り成分が同時に計測できるため，中赤外分光分析による気相の計測が有用であるが，出汁の香りが極めて繊細で有るため高い計測感度が必要となる．

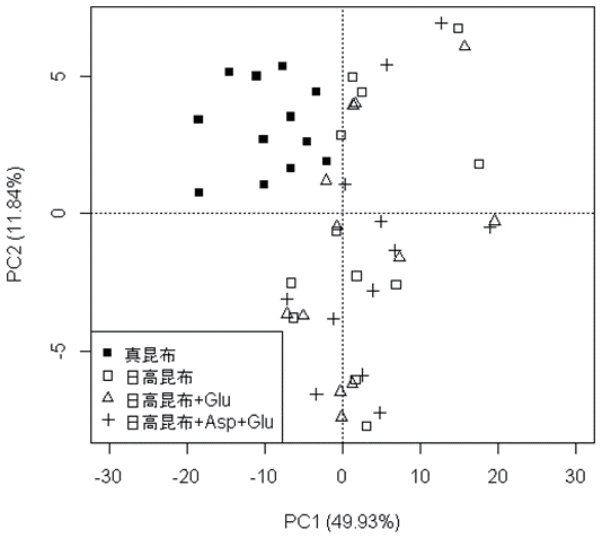


図 5 QDA 法による各出汁の評価値の主成分分析結果

5. おわりに

試し評価の際，各パネリストの評価値にばらつきがあったが，全パネリストの傾向は一致しており，成分分析の結果とも対応した結果が得られた．そのため，評価用語の定義に対する認識は一致しており，成分の違いの傾向も捉えることが可能であるが，パネリスト間で評価尺度の認識が一致していないと考えられる．今後，より詳細な官能評価を行っていく上で，精度よく評価をしていくためには，試し評価とキャリブレーションを繰り返し，パネリスト間の評価尺度を擦り合わせる必要がある．

また，カリウムが多量に含まれることで雑味が生じる傾向が認められたが，ミネラルは昆布だしらしい味に必要な要素である．そこで，昆布だしらしい味から雑味に変化するときのミネラル量およびその他成分の関係性を評価することで，より雑味の原因解明につながると考えられる．

本研究では，雑味を探ることが目的だったため，真昆布，日高昆布，長昆布の評価を行った．今後，各種昆布だしの特徴を表現するには，利尻昆布や羅臼昆布を含めて評価を行い，昆布全体での各昆布だしの相対的な位置を評価する必要がある．

また，データ数が充分得られた段階で，機械学習や AI を用いた評価を行う必要があるため，これらの解析に向けたデータ収集とデータクレンジングが必要である．

参考文献

［二宮，2010］二宮くみ子：日本及び西洋料理における'だし'に関する研究, 広島大学博士論文 , 2010.

［Stone 1974］Stone, H., Sidel, J. L., Oliver, S., Woolsey, A., Singleton, R.C.: Sensory evaluation by quantitative descriptive analysis, Food Technol, 28（11）:24, 1974.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

鳴き声の線形予測残差を用いた牛の個体識別

The individual identification of cattle using LP-residual signal extracted from cattle sound

入部 百合絵[*1] Yurie Iribe　曽我 真子[*1] Mako Soga　兒嶋 朋貴[*2] Tomoki Kojima　増田 達明[*2] Tatsuaki Masuda

[*1] 愛知県立大学 Aichi Prefectural University　[*2] 愛知県農業総合試験場 Aichi Agricultural Research Center

In order to reduce the time and effort of the individual identification, a method for easily identifying individuals of animals using ICT is required. Our research have focused on animal call as one of the features useful for individual identification. Although there is research on individual identification using linear prediction coefficients extracted from animal's voice in previous research, the correct rate is not sufficient. Therefore, it is necessary to perform individual identification using various acoustic features useful for individual identification as well as the linear prediction coefficients. In this research, we extract various kinds of acoustic features suitable for individual identification of cattle, and identify individuals using these acoustic features of cattle sound.

1. はじめに

近年，牛などの家畜を飼養している畜産農家は，高齢化や若者の就業数の減少などの理由により年々減少する傾向にある．しかし，飼養される家畜の数は一定であるため，農家一戸当たりの家畜の数は年々増加する傾向にある．また，家畜の個体識別や状態把握を行う際，飼養者は家畜を捕獲して識別や計測を行うための機器を取り付ける必要があるため，飼養者と家畜の両者に負担がかかる．このことから畜産農業の分野においても ICT による支援が重要と考えられる．特に個体識別は，家畜個々の健康状態や発情時期を把握するために必要となるため，本研究では ICT を用いて個体識別を簡便かつ正確に行うための手法を提案する．特に本研究では牛が自然に発する情報の一つである鳴き声を用いた個体識別法を提案する．牛の鳴き声は人間と同様に発声方法や声質に違いがある可能性が高いため，個体識別の判定材料として有用であると考えられる．

牛の鳴き声を用いた個体識別には石井らによる線形予測係数を用いた研究がある[石井 00]．個体識別や状態識別を目的とし，鳴き声の周波数特性の線形予測係数を用いて，マハラノビスの汎距離を判別基準として判別分析による識別を行っている．識別率は個体によってばらつきが大きく，平均は 59.8%という結果であった．また，牛の鳴き声から得られた音響情報を用いて分析を行っているが，使用した音響情報は線形予測係数のみであるため，他にも有用な音響情報を抽出する必要がある．また，池田らは個体識別を目的とし，2 頭の牛に対して線形予測誤差，基本周波数，およびスペクトル勾配を用いた識別を行っている[池田 01]．これらの特徴量を用いることで 2 頭の牛の識別が可能であることを示しているが，複数頭の牛に対して個体識別を行う必要がある．

図 1 実験風景

本研究では，牛の個体識別に有用な音響的特徴を明らかにするとともに，複数の牛に対し個体識別を行う．

表 1 収集した牛の鳴き声

牛番号	発声回数(回)	発声時間(s)
牛 1	136	246
牛 2	45	53
牛 3	46	70
牛 4	24	38
牛 5	2	3
牛 6	6	10

連絡先：入部百合絵，愛知県立大学 情報科学部，愛知県長久手市茨ヶ廻間 1522-3，iribe@ist.aichi-pu.ac.jp

2. 鳴き声からの音響的特徴量抽出と分析

2.1 収集した牛の鳴き声

本研究では，愛知県農業総合試験場のホルスタイン種乾乳牛 6 頭から 1 日約 4～6 時間程度の収録を 5 日間に亘って実施した．収録実験の様子を図 1 に示す．収録機器はビデオカメラ 2 台とレコーダー1 台(TASCOM PCM-RECORDER) である．

収集した牛の鳴き声データは，牛が鳴いていない部分や背景雑音によって鳴き声がかき消されている部分を消去し，計 259 回の鳴き声を得た．前処理として，鳴き声データに対してスペクトル引き算法による雑音除去を施した．ラベリングは収録と同時にビデオカメラで撮影した映像をもとに，目視により鳴いている牛を特定し，牛毎に鳴き声を分類した．特に，牛は鳴く際に首を高く上げて遠くに向かって口を開ける，腹部を大きく膨らませるため，これらの状態を参考に鳴き声データに対し牛番号を付与した．本研究に用いる牛の鳴き声データの内訳を表 1 に示す．

2.2 抽出する音響的特徴量

人間と同様に牛の鳴き声にも声帯情報と声道情報が含まれ

ていると考えられる．そこで，声帯特性に関わる特徴として基本周波数 F0，線形予測残差（以降，残差波形）の MFCC(Mel-Frequency Cepstrum Coefficients)および ΔMFCC を牛の鳴き声から抽出する．また，声道の長さや発声時の調音運動にも個体性が現れると考えられるため，声道特性に関わる情報として，音声波形の MFCC と ΔMFCC を抽出した．加えて，話者認識に用いられている RMS energy（パワー）の変化幅や残差波形の RMS energy につても確認した．一方，人と同じように牛の鳴き声も音韻毎に異なる特性を持っている可能性があるため，牛の鳴き声を/m/と/o/の区間に分け，音韻毎に MFCC と ΔMFCC を求めた．

2.3 分析結果と考察

牛の鳴き声に含まれる音響的特徴量が個体毎に異なっているかを確認するため，多重比較検定 (Tukey-Kramer 法) を用いて分析を行った．有意差の認められた特徴量の結果を表 2 に，/m/と/o/に対する分析結果を表 3 に示す．表中の低次数とは MFCC の 1 次～6 次，高次数とは MFCC の 7 次～12 次を示している． また，有意差の確認されたその他の特徴量に関しては紙幅の関係で割愛する．

予備実験より RMS energy の最大値や最小値には有意差は認められなかったが，RMS energy の高低差は有意な差が確認された．RMS energy の高低差は個体の鳴き方によって変化が現れやすいため，個体性が現れたと考えられる．また，残差波形の MFCC と ΔMFCC に有意差が生じたことから，牛毎に声帯の特性が異なる可能性が示唆され，残差波形が個体識別に有用であることが分かった．音声波形の MFCC にも有意な差が認められたため，牛の声道の長さや形状，発声動作に個体性が現れる可能性も示された．

表 3 より，/m/と/o/に対する分析では MFCC の高次数に有意差が認められている．MFCC の高次数は人間の感情や状態推定に用いられる音響情報であり，ここに有意差が生じていることから，今回の収録では牛の状態の違いが MFCC の高次数に現れた可能性がある．

3. 鳴き声の音響的特徴量を用いた個体識別

3.1 識別方法

本研究では 2.3 の分析結果で明らかとなった有意差の認められた 93 次元の特徴量を用いて個体識別を行った．用いた特徴量は単位やスケーリングが異なるため，特徴量に対し標準化処理を行った．事前に実施した複数の機械学習による予備実験を受けて，識別には SVM（Support Vector Machine）の線形カーネルを採用した．評価方法は 10-fold cross validation である．また，比較のため先行研究 と同様に線形予測係数のみを用いた識別を行った．結果を表 4 に示す．また，牛毎の識別結果を確認するため表 5 に Confusion Matrix を示す．

3.2 識別結果

提案手法では約 96.8%という高精度な結果を得ることができた．また，先行研究の結果は約 67.3%であった．この結果より，本研究の提案手法は従来の研究よりも約 30%も識別率を向上させることができた． 線形予測係数は MFCC と同じように声道情報の一つであるが，本研究では MFCC などの声道情報だけでなく，残差波形などの音源に関わる声帯情報も用いている．このことから，牛の個体識別には声道情報だけでなく声帯情報も用いることが有効であることが分かった．特に声帯特性を示す残差波形の MFCC やパワーが牛の個体識別に有用であることを確認した．

表 2 有意差の認められた特徴量

	p 値	
音響的特徴量	低次数	高次数
音声波形の MFCC		5.74E-3**
音声波形の ΔMFCC	6.22E-3**	6.70E-4**
残差波形の MFCC		1.42E-2*
残差波形の ΔMFCC	2.60E-3**	1.03E-2*
RMS energy(高低差)	2.38E-12**	

*p<0.05， **p<0.01

表 3 鳴き声/m/と/o/に対する分析結果

音響的特徴量	/m/	/o/
MFCC(高次数)	2.63E-2*	8.35E-3**
Δ MFCC(高次数)	1.26E-3**	3.30E-7**

*p<0.05， **p<0.01

表 4 個体識別の結果

	Correct rate
提案手法	96.8%
先行研究	67.3%

表 5 Confusion Matrix

		予測値					
		牛 1	牛 2	牛 3	牛 4	牛 5	牛 6
真値	牛 1	135	0	1	0	0	0
	牛 2	1	43	1	0	0	0
	牛 3	2	1	43	0	0	0
	牛 4	3	0	0	21	0	0
	牛 5	0	0	0	0	1	1
	牛 6	2	0	0	0	0	3

4. おわりに

本研究では牛の鳴き声から個体間で有意差の認められた音響的特徴量を用いて個体識別を行った．識別の結果，提案手法は 96.8%という高い識別率を得ることができた．また，先行研究の線形予測係数のみを用いた場合よりも，識別率が約 30%向上したことを確認した．これにより，鳴き声の音響情報には声道情報だけでなく声帯情報も用いることが個体識別に有用であることが明らかとなった．

実際の飼育環境を考えると多数頭での個体識別が必要であるため，長期間に亘り鳴き声データを収集するとともに，多頭数での高精度な牛の個体識別が今後の課題である．

参考文献

[石井 00] 石井洋平，池田善郎：鳴き声による肉用牛の個体識別および状態識別，農業機械学会誌，第 62 巻 1 号，p.50-58，2000．

[池田 01] 池田善郎，石井洋平：コンピュータによる牛音声の理解，農業機械学会誌，第 63 巻，1 号，p.4-9，2001．

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

肉牛の分娩検知システムにおけるクラウドソーシングを用いた誤通報の抑制

Suppression of false alarm using crowdsourcing in beef cattle delivery detection system

沖本祐典 *1 Yusuke OKIMOTO　斎藤奨 *1 Susumu SAITO　中野鐵兵 *1*2 Teppei NAKANO　赤羽誠 *1*2 Makoto AKABANE　小林哲則 *1 Tetsunori KOBAYASHI
小川哲司 *1 Tetsuji OGAWA

*1早稲田大学 Waseda University　*2知能フレームワーク研究所 Intelligent Framework Lab

In order for cattle farmers to detect calving sign beforehand and assist it to reduce risks of fatal accidents, recent work proposed a pattern recognition approach based on video information from cameras. To reduce false alarms given by the pattern recognizer, crowdsourcing can be used for double-checking the result of the automatic event detection. However, calving sign detection from videos is not a common task for crowd workers, where most of them are not experts of cattle farming; it is therefore not clear about how microtasks can be designed for the workers and thus their answers contribute to better prediction accuracy. In the present study, a calving detection system of beef cattle is designed aiming for real-world deployment. Exposure of the amnion and allanto from the buttocks of cattle is considered as a sign of calving and identified by the crowdworkers in microtasks designed. As a result of simulation evaluation of detecting two birth prognostic events, precision obtained when using only the pattern recognizer were 0.049 and 0.22, whereas in the case of using crowdsourcing it improved to 0.91 and 0.83, respectively. This result demonstrated that verification of the pattern recognition result by crowdworkers successfully reduced detection errors.

1. はじめに

本研究では，パターン認識器とクラウドソーシングによるラベリングの仕組みを同時に用いた肉牛の分娩検知システムを設計・実装し，検知結果をリアルタイムでクラウドワーカーに検証させることで，誤通報の抑制が可能であるか検証を行った．

肉牛の繁殖農家にとって，分娩間近の牛の見回りは大きな負担になっている．実際，分娩時の子牛の死亡事故は非常に多いため [Mee 13]，農家は見回りの中で分娩の予兆や開始に関わる牛の変化を見つけ，分娩に立ち合い，状況に応じて介助を行う．そこで，現在農家が目視で確認している予兆や変化を，パターン認識技術を用いてカメラ映像から自動で検出し，通知するシステムを構築できれば，農家の負担を大きく軽減できる可能性がある [Saint-Dizier 15, 沖本 18]．

映像による肉牛の分娩検知システムにおいては分娩の見逃しは致命的であり，最小限にすることが求められる．その一方で，見逃しと誤検出はトレードオフの関係にあり，見逃しを抑えようとすると誤検出が多くなる．また，パターン認識器の性能が十分でない場合，見逃しと誤検出のトレードオフの関係はより強くなる．つまり，十分なデータが集まっていないシステム運用初期においては見逃しの件数を小さくするため検出の閾値を低めに設定せざるを得ないが，結果として農家に対する分娩の誤通報が多くなる可能性がある．このとき，頻繁な誤通報は農家にとって大きなストレスになると考えられる．

見逃しを最小限にしながら誤通報を抑制する手段として，クラウドソーシングをパターン認識器の出力の検証に用いる方法が提案されている [Saito 16]．見逃し件数を小さくするように設計されたパターン認識器がイベントを検知したとき，クラウドソーシングの作業者 (クラウドワーカー) に人手による二重チェックを依頼する．二重チェックの結果が負であると判断された場合はパターン認識結果は誤りとみなされ，通報は行われなくなる．このとき，クラウドソーシングによる誤検出の抑制を実現するためには，クラウドワーカーの回答が信頼できることが前提となる．しかし，肉牛およびその分娩の映像は，多くのクラウドワーカーには馴染みがないため，分娩を判断する根拠として畜産の専門家ではないクラウドワーカーでも判断可能なものを選び，分娩検知システムを設計する必要がある．

本研究では，パターン認識器とクラウドソーシングによるラベリングの仕組みを同時に用いた肉牛の分娩検知システムを設計・実装し，検知結果をリアルタイムでクラウドワーカーに検証させることで，誤通報の抑制が可能であるか検証を行った．分娩開始の目印は，尿膜・羊膜と呼ばれる組織の露出とする．このとき，牛舎に設置されたカメラから得られる画像を入力として，ニューラルネットワークに基づく分娩検知器により尿膜・羊膜を検出し，検出結果の検証をクラウドワーカーに依頼する．シミュレーション実験として，運用初期段階を想定したシステムの検知結果を得た後に，検証による正確なラベルを得た上で通知する場合と検知結果をそのまま通知する場合における最終的な分娩開始検知性能を画像単位の適合率と再現率により評価した．

本稿の構成は次の通りである．まず，2. で，映像による分娩検知システムの概要を述べる．次に，3. で，システムの各モジュールについて説明する．4. では，構築したシステムを用いて行った実験について述べ，クラウドソーシングによって誤検出を抑制できるか検証する．最後に，5. で本稿のまとめを述べる．

2. 分娩検知システムの概要

本章では，本研究で構築する，肉牛の分娩検知システムの概要 について述べる．

連絡先: 沖本 祐典, 早稲田大学基幹理工学研究科，東京都新宿区早稲田町 27， okimoto@pcl.cs.waseda.ac.jp

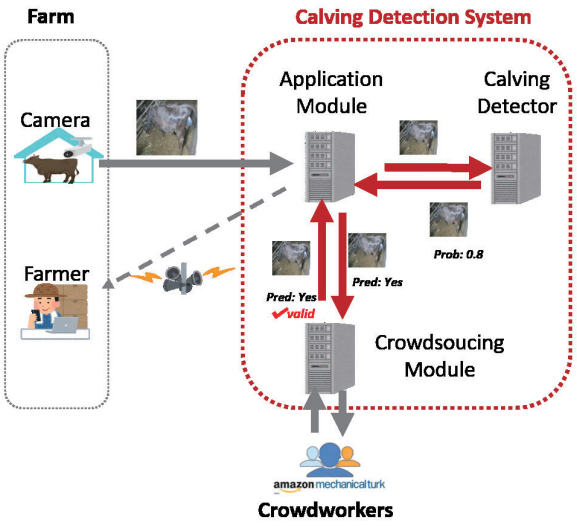


図 1: 映像による分娩検知システムの構成．アプリケーションモジュール（Application Module），分娩検知器（Calving Detector），クラウドソーシングモジュール（Crowdsourcing Module）の 3 つの部分からなる．

このシステムは，肉牛の尿膜・羊膜の露出（図 2）をカメラの映像から検出することで，分娩検知を行う．システムは，分娩検知器，クラウドソーシングモジュール，アプリケーションモジュールの 3 つのモジュールからなる（図 1）．分娩検知器は，画像から肉牛の領域を検出し，その肉牛で尿膜・羊膜の露出が発生している確率を出力する．アプリケーションモジュールは，牧場，分娩検知器，クラウドソーシングモジュールをつなぐ役割を持つ．また，アプリケーションモジュールは，尿膜・羊膜の露出の確率についての閾値 p_{th} をもち，分娩検知器が出力した確率がその閾値 p_{th} 以上だった場合，システムとして画像で尿膜・羊膜の露出が発生していると判断する．クラウドソーシングモジュールは，システムが尿膜・羊膜の露出が発生していると判断した画像について，その検出結果が正しいかクラウドワーカーに検証を依頼する．

分娩の見逃しを小さくするため，アプリケーションモジュールが持つ，尿膜・羊膜の露出の確率の閾値 p_{th} は低く設定する．この結果，分娩検知器による尿膜・羊膜の露出と誤検出は多くなる．ここで，尿膜・羊膜の検出結果をクラウドソーシングモジュールを通てクラウドワーカーにより検証してもらうことで，誤検出を訂正する．この，分娩検知器とクラウドソーシングを併用することにより，分娩の見逃しを低く抑えながら，同時に誤検出も抑制する．

3. 分娩検知システムの設計

本章では，肉牛の分娩検知システムの各モジュールについて説明する．

3.1 分娩検知器

本節では，本分娩検知システムで分娩の根拠として用いる肉牛の尿膜・羊膜の露出について説明したのち，尿膜・羊膜を検出する分娩検知器の構成について述べる．

3.1.1 分娩時における尿膜・羊膜の露出

肉牛の尿膜・羊膜の露出の画像例を図 2 に示す．羊膜は分娩 3 時間から 7 時間前に露出する．また，尿膜は分娩 2 時間

(a) 平常状態　(b) 尿膜の露出　(c) 羊膜の露出

図 2: 尿膜・羊膜の画像例．尿膜・羊膜ともに水風船状の外見を持ち，牛の産道から露出する．尿膜は黒っぽい色を，羊膜は白っぽい色をしている．

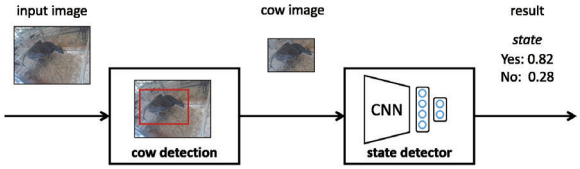


図 3: ニューラルネットワークによる分娩検知器の構成．まず牛の領域を検出し，次に尿膜・羊膜が露出しているか識別する．

半から 5 時間前に露出する [Saint-Dizier 15]．両者とも，水風船状の外見をしており，尿膜は黒っぽく，羊膜は白っぽい色をしている．

3.1.2 ニューラルネットワークによる分娩検知器

尿膜・羊膜検出による分娩検知器は，画像から牛を検出する部分と，検出した牛が尿膜・羊膜を露出しているか識別する部分からなる（図 3）．牛領域の検出には，ニューラルネットワークを用いた高精度な物体検出アルゴリズム YOLOv2 [Redmon 17] を用いる．尿膜・羊膜の露出の識別には，畳み込みニューラルネットワーク (Convolutional Neural Networks, CNN) を用いる．

3.2 クラウドソーシングモジュール

クラウドソーシングモジュールは，尿膜・羊膜を露出していると判断された牛の画像領域を受け取り，その正しさをクラウドワーカーによって検証する．

クラウドソーシングのプラットフォームには Amazon Mechanical Turk を用いる．タスク画面には，尿膜・羊膜が露出している判定された画像が複数枚一度に提示される．クラウドワーカーには，画面の画像から，尿膜・羊膜の露出している画像を選択するよう求められる．

クラウドワーカーの回答の信頼性を担保するため，一画面に提示される画像には，正解ラベルが既に付与された画像も含まれている．正解ラベルを持つ画像をクラウドワーカーが間違えた場合，その画面で得られた回答は，信頼できないとしてすべて無効とする．画像ごとに有効な回答が一定数集まったら，多数決をとって検証結果とする．

3.3 アプリケーションモジュール

アプリケーションモジュールは，各モジュール間をつなげる役割を持つ．具体的には，以下の処理を行う．

- 牧場のカメラからのインターネットを介した画像の受け取る．
- 分娩検知器へ牧場のカメラ画像を渡す．
- 分娩検知器から，牛の画像領域と，領域ごとの尿膜・羊膜の露出の確率を受け取る．

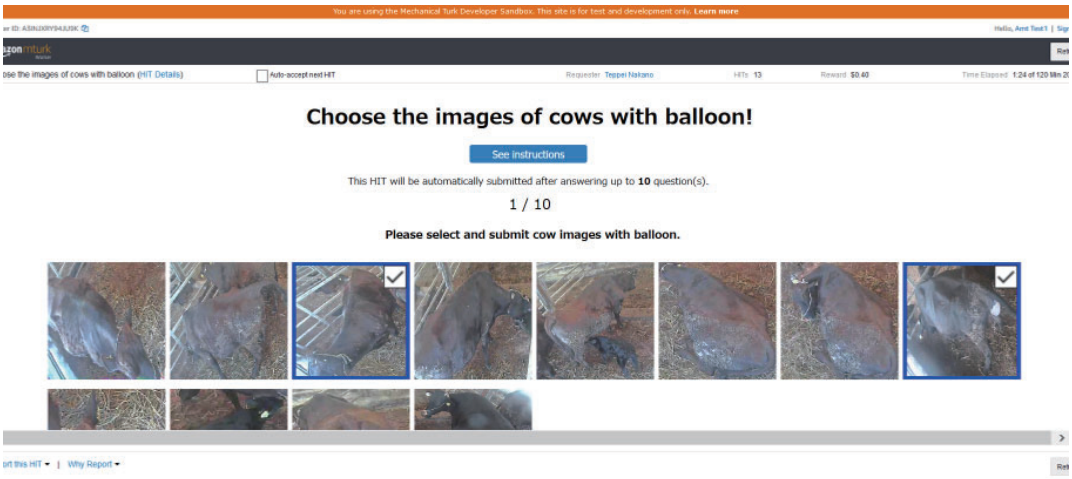


図 4: クラウドワーカーによる尿膜・羊膜の露出の検証タスク画面．複数提示された画像から，尿膜・羊膜が露出している画像を選択する．あらかじめ正解ラベルが付与された画像も提示され，それらの画像の回答をチェックし，ワーカーの信頼性を確認する．

- クラウドソーシングモジュールに，尿膜・羊膜が露出していると判断した牛画像領域を渡す
- クラウドソーシングモジュールから，尿膜・羊膜が露出していると判断した牛画像領域の検証結果を受け取る．

を行う．アプリケーションモジュールは，分娩検知器から受け取った牛領域の尿膜・羊膜が露出している確率が，あらかじめ定めてある閾値 p_{th} 以上であった場合，クラウドソーシングモジュールに検証を依頼する．先述した通り，分娩の見逃しを小さくするため，閾値 p_{th} は低く設定しておく．

4. 実験

誤検出の抑制を，通報をリアルタイムにクラウドワーカーに検証させることで実現できるか調べるため，システムを実験あらかじめラベリングしたデータセットを入力として，分娩検知システムを稼働させ，クラウドソーシングを用いた場合と用いない場合について，尿膜・羊膜の露出の，画像単位での適合率と再現率を測定した．

検知結果をリアルタイムでクラウドワーカーに検証させることで，誤通報の抑制が可能であるか検証を行うため，あらかじめラベリングしたデータセットを入力として，シミュレーション実験として，分娩開始の目印には，尿膜・羊膜の露出を用い，

運用初期段階を想定したシステムの検知結果を得た後に，検証による正確なラベルを得た上で通知する場合と検知結果をそのまま通知する場合における最終的な分娩開始検知性能を画像単位の適合率と再現率により評価した．

4.1 使用データ

実験には，鹿児島県の肉牛の畜産農家で収録した黒毛和種の映像を用いる．映像の収録は 2016 年 11 月から開始し，現在も継続して行われている．カメラで取得した映像は，1 秒ごとに 1 枚の画像として，分娩検知システムが稼働するクラウドサーバー上の転送され，保管される（図 5）．

4.1.1 分娩検出器の訓練データセット

分娩検知器の訓練データには，2017 年 3–8 月に収録した画像を用いた．分娩 5 時間前から分娩時まで画像から 6 秒毎に 1 枚ずつ選びそれらの画像から牛領域を切り出し，尿膜・羊膜の露出についてアノテーションした．アノテーションには Amazon Mechanical Turk を用いた．各画像ごとに 2 人のワーカーの回答を集め，両者の回答が一致したものをラベルとして用いる．尿膜・羊膜が露出している正例画像の枚数と，露出していない負例画像の枚数は，それぞれ 3806 枚である．正例と負例の数が一致するように，ランダムアンダーサンプリングを行った．

図 5: 収録された画像例．

表 1: システム検証用データの枚数．

	正例	負例
分娩 1	125	16829
分娩 2	941	26446

4.1.2 システム検証用データ

クラウドソーシングによる適合率向上実験の評価データには，2017 年 9–12 月に収録した画像を用いた．計 2 回の分娩シーンが含まれており，分娩 6 時間前から分娩時までのすべての画像を用いた．枚数を表 1 に示す．

4.2 実験設定

本節では，実験における，分娩検知システムの各モジュールにおける設定を述べる．

4.2.1 分娩検知器

牛を検出する YOLOv2 の重みには，MS COCO [Lin 14] で学習した既存のモデルを用いた．尿膜・羊膜を検出する CNN は，畳み込み層は ResNet50 [He 16] と同様，それに続く全結合層は 1024unit の 1 層の構造のものを用いた．出力層は，尿膜・羊膜が露出しているかどうかの二値識別を行う．畳込み層は ImageNet[Deng 09] で学習済みの重みを初期値とし，先述した訓練データを用いてファインチューニングした．

アプリケーションモジュールが持つ，検知器の出力確率の閾値 p_{th} には 0.5 を用いる．

4.2.2 クラウドソーシングモジュール

クラウドワーカーには，一画面に 10 枚の画像を提示した．そのうち 2 枚は，品質保証用の正例 1 枚，負例 1 枚のアノテー

表 2: システムを稼働させた際の画像単位の適合率と再現率

	クラウドソーシング	適合率	再現率
分娩 1	なし	0.046	1.0
	あり	**0.91**	1.0
分娩 2	なし	0.22	0.90
	あり	**0.83**	0.70

ション済み画像である．1 画像につき 3 人分のワーカーの有効回答を集め，多数決を行った．

4.3 実験結果

分娩検知システムを，2 回の分娩データセットで稼働させたときの，画像単位の適合率と再現率を表 2 に示す．

適合率は，クラウドソーシングによる検証を行わなかった場合，それぞれ 0.049，0.22 であった．一方で，クラウドソーシングによる検証を行った場合，それぞれ 0.91, 0.83 であり，クラウドソーシングによる検証によって，それぞれ 0.86, 0.61 改善した．また，再現率は，クラウドソーシングによる検証を行わなかった場合はそれぞれ 1.0，0.90 で，クラウドソーシングによる検証を行った場合もそれぞれ 1.0, 0.70 だった．

以上より，映像による分娩検知システムにおいて，尿膜・羊膜の露出を分娩の情報を用い，クラウドソーシングによる分娩検知器の出力を検証することで，見逃しを少なく（再現率を高く）保ちつつ，誤検出を抑制（適合率を高く）できることが確認できた．

クラウドソーシングを用いない場合，適合率がかなり小さい．再現率が高くなるよう閾値を設定したことが，理由の一つである．加えて，牛の動きが緩慢であり，同じような画像が連続していることも理由として考えられる．一つ画像が正例として間違えて検出した場合，連続している同じような画像をすべて正例として間違え，まとまって誤検出が発生していると考えられる．これは，クラウドソーシングで得られた検証結果を用いて分娩検知器を逐次更新を行うことで，改善できる可能性がある．

また，分娩 2 において，クラウドソーシングの検証により，再現率が 0.20 悪化している．クラウドワーカーが，尿膜・羊膜が露出している画像を，間違えて露出していないと判断しているが原因である．これは，画面設計の変更や説明を詳しくすることにより，タスクの改善によって，改善できると考えられる．

5. 結論

本研究では，パターン認識器とクラウドソーシングによるラベリングの仕組みを同時に用いた肉牛の分娩検知システムを設計・実装し，検知結果をリアルタイムでクラウドワーカーに検証させることで，誤通報の抑制が可能であるか検証を行った．分娩開始の目印には，尿膜・羊膜と呼ばれる組織の露出を用いた．このとき，牛舎に設置されたカメラから得られる画像を入力として，ニューラルネットワークに基づく分娩検知器により尿膜・羊膜を検出し，検出結果の検証をクラウドワーカーに依頼する．シミュレーション実験として，運用初期段階を想定したシステムの検知結果を得た後に，検証による正確なラベルを得た上で通知する場合と検知結果をそのまま通知する場合における最終的な分娩開始検知性能を画像単位の適合率と再現率により評価した．実験の結果，パターン認識器のみの場合の適合率が 0.049 と 0.22 であったのに対し，クラウドソーシングを用いた場合は 0.91 と 0.83 に向上した．この結果からパターン認識結果をクラウドワーカーに検証させることで，適切に誤検出を抑制できることを示した．

今後は，クラウドソーシングで得られた検証結果を用い，分娩検知器を逐次更新することによる，分娩検知器の適合率の改善と，クラウドワーカーに対するタスク設計の改善による，最終的な分娩開始検知における再現率の改善について検討を行う．

6. 謝辞

本研究を行うにあたり，牛のモニタリング動画を提供してくださった東村牧場様，そして有益な議論を頂いたファーマーズサポート株式会社様，鹿児島頭脳センター様に感謝申し上げます．

参考文献

[Deng 09] Deng, J., Dong, W., Socher, R., Li, L., Li, K., and Fei-Fei, L.: ImageNet: A large-scale hierarchical image database, in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255 (2009)

[He 16] He, K., Zhang, X., Ren, S., and Sun, J.: Deep Residual Learning for Image Recognition, in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778 (2016)

[Lin 14] Lin, T., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C.: Microsoft COCO: Common Objects in Context, in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 740–755 (2014)

[Mee 13] Mee, J.: Why Do So Many Calves Die on Modern Dairy Farms and What Can We Do about Calf Welfare in the Future?, *Animals*, pp. 1036–1057 (2013)

[Redmon 17] Redmon, J. and Frhadi, A.: YOLO9000: Better, Faster, Stronger, in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7263–7271 (2017)

[Saint-Dizier 15] Saint-Dizier, M. and Chastant-Maillart, S.: Methods and on-farm devices to predict calving time in cattle, *The Veterinary Journal*, Vol. 205, pp. 349–356 (2015)

[Saito 16] Saito, S., Kobayashi, T., and Nakano, T.: in *the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion(CSCW)*, pp. 393–396 (2016)

[沖本 18] 沖本 祐典, 菅原 一真, 斎藤 奨, 中野 鐵兵, 赤羽 誠, 小林 哲則, 小川 哲司 : 牛の分娩予兆として映像から観測可能な状態の検知, 2018 年度人工知能学会全国大会（第 32 回）(2018)

Organized Session | Organized Session | [OS] OS-13

[1G3-OS-13a] “ナッジ”エージェント：人をウェルビーイングへと導くエージェントの構築へ向けて(1)

小野 哲雄（北海道大学）、今井 倫太（慶應義塾大学）、植田 一博（東京大学）、山田 誠二（国立情報学研究所

Tue. Jun 4, 2019 3:20 PM - 5:00 PM Room G (302A Medium meeting room)

[1G3-OS-13a-01] “ Nudge” Agent

〇Tetsuo Ono[1] （1. Hokkaido University）

3:20 PM - 3:40 PM

[1G3-OS-13a-02] (Invited talk) Cognitive science of choice and manipulation

〇Ayumi Yamada[1] （1. The University of Shiga Prefecture）

3:40 PM - 4:20 PM

[1G3-OS-13a-03] How do we harness others’ opinions?

〇Itsuki Fujisaki[1], Hidehito Honda[2], Kazuhiro Ueda[1] （1. University of Tokyo, 2. Department of Psychology）

4:20 PM - 4:40 PM

[1G3-OS-13a-04] Nudge with a forward posture

〇Masaru Shirasuna[1], Hidehito Honda[2], Kazuhiro Ueda[1] （1. the University of Tokyo, 2. Yasuda Women’s University）

4:40 PM - 5:00 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

“ナッジ ”エージェント: 人をウェルビーイングへと導くエージェントの提案

“Nudge” Agent: Proposal of an Agent that Guides People to Well-being

小野 哲雄[*1]
Tetsuo Ono

[*1] 北海道大学 大学院情報科学研究院
Faculty of Information Science and Technology, Hokkaido University

“Nudge” advocated in behavioral economics is known as a strategy that guides people to health and well-being with “a little modest way.” Thaler is known as the Nobel Prize in Economics winner in 2017 with this achievement. Now that redesign (reconstruction) of social system is required due to the rapid development of information technology such as AI and IoT, in this research we regard “nudge” as “mental navigation” through implementing as a “Nudge Agent” by using the basic technology of Human-Agent Interaction (HAI) which we have been carrying out. Specifically, by using HAI-based technology, we construct a “Nudge Agent” that can manipulate the cognitive environment (selection architecture) that directs human decision-making and conduct field experiments at actual home and commercial facilities. Then, we will examine whether this agent can “guide” people’s decision to health and happiness to some extent.

1. はじめに

行動経済学において提唱されている“ナッジ” (nudge) は「ほとんど気づかないくらいにささやかな方法」で人々を健康や幸福へと導く戦略 [Thaler 08] として知られ，Thaler は当該研究の業績により 2017 年にノーベル経済学賞を受賞した．

AI や IoT などの情報技術の急速な発展により，社会システムのリデザイン (再構築) が求められている現在，本研究では“ナッジ”を「心のナビゲーション」として捉え，それを我々がこれまで約 20 年間にわたり研究を進めてきた Human-Agent Interaction (HAI) の基盤技術を用いることで“ナッジ”エージェントとして実装することにより，人々を「良い行動」へと導くことが可能な，新しい社会的インタラクションの枠組みを提案する．

具体的には，HAI 基盤技術を用いることにより，人の意思決定を方向づける認知的環境 (選択アーキテクチャ) を操作可能な“ナッジ”エージェントを構築し，実際の家庭や商業施設などでのフィールド実験をとおして，このエージェントが人々の意思決定を健康や幸福へと“それとなく”導くことが可能かどうかを検証する．

本稿では，2 節において“ナッジ ”エージェントの構想に至った背景について述べる．この研究を進めるうえで，基礎的な側面として，3 節において認知的なメカニズムの解明とそのモデル化について述べる．さらに，応用的な側面として，4 節において“ナッジ ”エージェントの実装方法について述べる．以上の考察を踏まえ，5 節では本研究の今後の展望について述べる．

2. 研究の背景

行動経済学における“ナッジ”は，複数の選択肢がある場合，どういう意思決定をすれば多くの人にとっての効用 (満足) が高まるかを，特に「人の心のメカニズム」に焦点を当てることにより，選択の余地を残しながらも“それとなく”支援しようというものである [Thaler 08]．たとえば，カフェテリアでサラダなどの健康によい品を利用者が取りやすい最初の方に配置することで，無意識のうちに健康によい食べ物を多く取らせたり，書類の選択肢の初期値にあらかじめ同意のチェックを入れることで，年金への加入者を大幅に増やすことに成功している [Hagman 15]．

“ナッジ”エージェントの構想に至った背景は以下の 3 点にまとめることができる．第 1 は大量のデータの必要性に関する問題である．現在，AI に関する研究は隆盛を極め，社会からの期待も大きい．しかし，現時点での AI 技術は大量のデータを必要とし，データがなければ学習できず，データのないまったく新しい事象へは対処もできない．データが活用できる事象へは当然，現在の手法を積極的に活用すべきであるが，データを得るためのセンシング手法もいまだ確立されていない人の意思決定過程への適用は難しい．

第 2 はシステム化，一般化の必要性である．行動経済学では，人間の意志決定の過程に注目し，人間の行動を観察することで実証的に社会現象や経済行動を捉えようとしている．たとえば，“ナッジ”を効果的に利用した事例を数多く収集し，研究報告ではそれらを紹介している．しかし，そのような人間の行動に一定の法則性を見出し，それを一般化し，システムとして捉えようという研究事例は少ない．このため，“ナッジ”の一時的な効果を検証するのみとなったり，それに対する「慣れ」や「飽き」が効果を減じているとの報告もある．

第 3 に，AI 技術などにより評価関数が定められないような，さまざまな価値観が混在する状況下での意思決定や合意形成の問題を支援する方法がまだ見つかっていない [JST CRDS 17]．“ナッジ”エージェントに関する議論を進める中で，この問題も研究の射程におさめたいと考えている．

3. 認知的なメカニズムの解明とそのモデル化

従来の“ナッジ”に関連した研究では，実験室内において質問紙やディスプレイ上の選択肢の内容や構成を変えることにより条件を設定し，その条件間における実験参加者の回答や反応の違いを調べるものが多かった．つまり，選択者の自由を残しながら，選択肢の内容や構成など意思決定を方向づける認知的環境 (選択アーキテクチャ) [Thaler 08] を変えることにより，それに対して選択者がどのような反応をするかを調べる研究が主であった．

言うまでもなく，実際の社会的なインタラクションでは，うなづきなどの応答の時間的なタイミングや視線などの空間的参照を用いた社会的シグナルが意思決定に大きな影響を与えている

連絡先：小野哲雄，北海道大学 大学院情報科学研究院，札幌市北区北 14 条西 9 丁目，tono@ist.hokudai.ac.jp

[植田 16]. つまり，従来の行動経済学における“ナッジ”に関する研究では，選択肢や物の配置などの「静的な」選択アーキテクチャのみを考慮し，社会的シグナルなどの「時間的・空間的」に変化する「動的な」選択アーキテクチャを扱ってこなかった．本研究では後に 4 節で述べるように，最新の AI 技術を用いた動作計測および解析装置により人の動作を計測・解析することにより，社会的シグナルを含んだ「時間的・空間的」に変化する「動的な」選択アーキテクチャにおける意思決定過程の分析をとおして，“ナッジ”の認知的メカニズムの解明を目指す．さらに，この研究成果とこれまでの研究で明らかになったバイアスの機能に基づき，計算機上に実装可能な“ナッジ”のモデル化を行う．

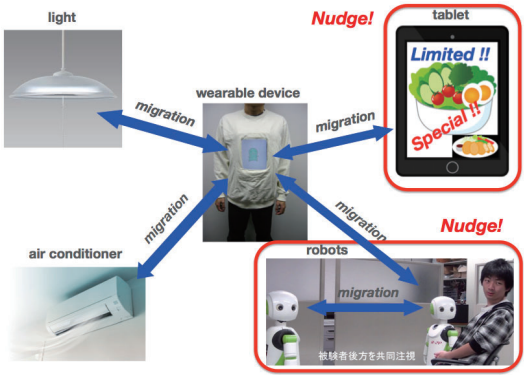


図 1 エージェントの「移動」による“ナッジ ”の実装

4. “ナッジ”エージェントの実装

本研究では，人々をウェルビーイングへと導く意思決定を促す“ナッジ”エージェントの実装方法を提案する．“ナッジ”エージェントとは，人とのインタラクションにおいて，情報的，時間・空間的な要因(「動的な」選択アーキテクチャ) を変化させることで認知的な自動システムに働きかけ，人の意思決定を“それとなく”導くシステムである．

“ナッジ”エージェントの実装では，知覚的コントラストとリズム構造に注目し，これらを HAI 基盤技術を用いて実装する．ここで知覚的コントラストとは精神物理学の領域で確立された原理であり，提示される順番(時間的な差異) や空間的配置が我々の認識に影響を与えるというものである．たとえば，高額な商品 A を見たあとにある商品 B を見ると，商品 A を見なかった場合よりも商品 B が安く感じる現象であり，あらゆる種類の知覚に適用可能である．またリズム構造とは，2 つ以上の系が引き込み合う(entrainment) ことにより創発する構造であり，この構造を用いてその後の人の身体動作や注意を制御することができる[Ono 16].

具体的な事例を用いて，“ナッジ”エージェントの実装に関する説明を行う(図 1). ここでは HAI 基盤技術のうち，「移動」(agent migration) 機構を基礎とした ITACO システム[小川 06] の技術を用いる．ITACO システムでは，ユーザの選好(判断基準や趣味など) を学習済みのエージェントが環境内にあるさまざまなメディアに「移動」することにより，ユーザに対して，文脈に応じた適切な支援を行うことができる．この「移動」の利点は，環境内にある「メディアの効果的な利用」とユーザに対する「一貫性のある(consistent) 支援」である．

いま，病気治療のためカロリー摂取制限を受けているユーザを想定する．エージェントはユーザの選好を学習済みであり，環境内にあるさまざまなメディアや IoT 機器に「移動」しながら，“ナッジ”を提供することができる．たとえば，図 1 に示すように，エージェントはユーザのウェアラブル PC からレストランのメニューを表示するタブレットに「移動」し，低カロリーメニューの配置やデザインを変更することにより(知覚的コントラスト)，ユーザの意思決定に働きかえ，無意識のうちに低カロリーメニューの選択を促す．さらに別の状況では，エージェントがロボットに「移動」し，共同注意の機構により関係(リズム構造) を構築した後，視線誘導の機能を用いて，ユーザの視線(注意) を高カロリーの飲み物の広告から低カロリーの飲み物へ誘導し，自然にそれを選択するように促す．これらの“ナッジ”の実現方法は HAI に関する我々の先行研究の知見に基づくものである．

本研究で提案する“ナッジ”エージェントのポイントは，ユーザは予め意思決定の方向性は決めているということである．先ほどの事例では，病気治療のためカロリーの摂取制限には事前に同意している点である．つまり，本研究で実現する“ナッジ”は無意識的に人をある意思決定へと誘導するわけではなく，意思決定の方向性は同意のうえ，その後の意思決定を後押ししていくようなエージェントおよび環境知能を実現しているのである．

本研究の成果は，“ナッジ”エージェントとしての社会的応用に留まらず，既存の研究分野である，対話研究(説得過程の解明) や自然言語研究(意味理解のバイアス効果) に貢献すると考える．特に，実際の家庭や商業施設などでのフィールドにおいて，「動的な」選択アーキテクチャに注目し，これらの点を明らかにしようという試みは新しい取り組みである．

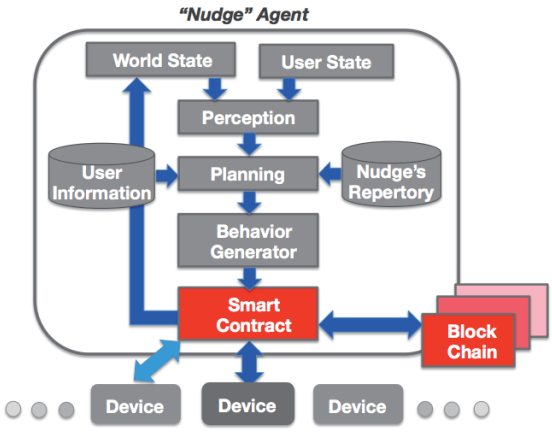


図 2 “ナッジ ”エージェントのシステムの概要

図 2 に“ナッジ”エージェントシステムの概要を示す．ここでは，図 1 と同様に，病気治療のためカロリー摂取制限を受けているユーザを想定して説明を行う．

<u>① Nudge’s Repertory モジュールに基づく文脈に応じた“ナッジ”の提供</u>

図 1 に示したように，カロリー摂取制限を受けているユーザには，メニューの配置やデザインの変更および視線 (注意) の誘導によりユーザの意思決定にはたらきかけ，無意識のうちに低カロリーメニューの選択を促す．これらの“ナッジ”の実現は，User Information と Nudge’s Repertory を用いて Planning モジュールで動作の計画がなされ，Behavior Generator で実際のデバイスに応じたエージェントの振る舞いが決定される．

具体的には，Nudge’s Repertory には活性伝搬モデルを実装し，現在の文脈で適用可能な“ナッジ”の提供を実現する．さらに，ユーザの情報(User Information)も取り込まれ，Planning モジュールにおいて最適なエージェントの振る舞いが決定される．その振る舞いが Behavior Generator により実現される．この事例では，カロリーを制限しなくてはいけないため，エージェント

はタブレットのメニューに「移動」し，コントラストの原理を用いてそのデザインを変更することにより低カロリーメニューの選択を促す．

②どのようなデバイスにもエージェントの移動が可能

この“ナッジ”エージェントはスマートコントラクトとブロックチェーンの技術を用いることにより，どのようなデバイスへも「移動」可能となるように設計する．これは，我々の提案してきた HAI 基盤技術を発展させることにより実装可能となる．具体的には，従来は事前にエージェントの「移動」先を指定する必要があったが，上記の技術により，プロトコルの異なる機器同士が自律的にネットワークを構築することが可能となったため本システムが実現可能となった．つまり，本エージェントシステムのポイントは，エージェントがメディアへの「移動」をとおして，そのメディアの表現力を有効に活用し，かつ一貫性を持って，“ナッジ”をユーザに的確に提供できることである．この実装された“ナッジ”が人の意思決定にどのような影響を与えるかについてフィールド実験をとおして明らかにしていく．

5. 今後の展望

本稿では，HAI 基盤技術を用いることにより，人の意思決定を方向づける認知的環境 (選択アーキテクチャ) を操作可能な“ナッジ”エージェントを提案してきた．この考え方に至った経緯を振り返りつつ，本研究の今後の展望について考えてみたい．

ここでは，「環世界」，「コミュニケーションの“膜”」，「集団の意思決定」の 3 つの観点から議論を行う．

5.1 「環世界」

ユクスキュルの提唱する「環世界」 (Umwelt) とは，すべての動物はそれぞれに種特有の知覚世界をもって生きており，その主体として行動しているという考え方である [ユクスキュル 05] ．つまり，動物は彼が「トーン」 (ton) と呼ぶフィルタを用いて環境を「意味」付けながら行動していると考えることもできる．言うまでもなく，視覚や聴覚を持たないが嗅覚や触覚のすぐれているマダニと，我々人間が見ている世界はまったく異なる．さらに，同じ人間であっても，そのときの状況や文脈，生理的な状態，本人の趣味や嗜好によっても見える世界は異なるであろう．

このような「環世界」を仮定すると，当然それは意思決定を方向づける認知的環境 (選択アーキテクチャ) にも大きな影響を与えることになる．各種の生物の知覚フィルタが創る「環世界」と意思決定の関係，および「環世界」間の相互作用を考えることは大変興味深い研究テーマである．

5.2 コミュニケーションの“膜”

「環世界」のような主体的な空間およびコミュニケーションの「場」のような複数人によって創発される相互作用空間は科学的な興味とともに，工学的な応用としても重要である．

[小野 14] は，HAI 研究の立場から，コミュニケーションと社会的空間の利用に焦点をあて，拡張社会空間システムを提案している．さらに，このシステムが創り出す“膜”状の相互作用空間は，人とのインタラクションにおけるロボットの文脈適応的な振る舞いを実現し，また人と人のコミュニケーションにおける創造的な情報支援を可能にするなど，人とロボットの共創インタフェースとしての役割を果たす可能性があるとしている．

具体的には，図 3 では，ロボットは 2 人の対話の活性度を計算することにより社会的相互作用空間を認識し，その空間を回避する行動をとることができる．また，図 4 左では，相互作用を行なっている人のみが参加できる社会的空間が形成される．さらに，その社会的空間内でのみ情報を共有することができ，その空間外からはその情報にアクセスすることができない（図 4 右）．このような空間はあたかもコミュニケーションの“膜”のように機能する．

[津田 12] は，コミュニケーションの神経機構の数理的研究の立場から，「コミュニケーションの膜」としての「SF 膜 (SF-brane)」を提案している．この SF 膜によって，脳は孤立(solipsism) から解放されコミュニケーションによって初めて意味を生成できるようになるとし，このとき“膜”は内と外のインタフェースの役割を果たし，共創の場として機能すると述べている．

図 3 拡張社会空間を用いたロボットの文脈適応的な振る舞い

図 4 拡張社会空間の実装(左)と“膜”内のユーザのみへの情報提供(右)

5.3 集団の意思決定

前節で述べたコミュニケーションの“膜”は，膜外との「相互浸透」的な情報交換により，集団の意思決定にも活用できるのではないかと考えている．

社会的価値観やさまざまな人間の価値観が混在している状況下での合意形成は非常に難しい問題であり[JST CRDS 17]，現在，情報科学の技術を用いた有効な支援策は見つかっていない．本来，論理的な解決が難しい問題である以上，トップダウンな解決方法の提案はなじまないのではないかと考えている．

このような論理的な合意形成が難しい状況においては，コミュニケーションの“膜”の相互浸透のアナロジーによるボトムアップなアプローチは問題解決の 1 つの方法を示すのではないかと考えている．

コミュニケーションの“膜”の相互浸透のメカニズムは，[Spencer-Brown 72] による考えが示唆を与えるのではないかと考えている．当然，彼の提案する体系が古典命題論理やブール代数を再構成しているに過ぎないという批判はあるが，数学的な道具立てを整えることにより，彼の「思想」を記述することが可能なのではないかと考える．また，彼の「思想」自体が“膜”間の相互浸透のメカニズムを考えるとき，示唆を与えるものであると考える．

6. おわりに

本稿では，行動経済学において提唱されている“ナッジ”を HAI 基盤技術を用いることにより“ナッジ”エージェントとして実装し，人々を健康や幸福へと“それとなく”導くことが可能な，新しい社会的インタラクションの枠組みを提案した．さらに本提案が，AI や IoT などの情報技術の急速な発展により，社会システ

ムのリデザイン (再構築) が求められている現在においてどのような有効性があるかについて議論した.

今後は, 本提案システムのプロトタイプを用いて, 実際の家庭や商業施設などでのフィールド実験をとおして, 提案したエージェントシステムが人々の意思決定を健康や幸福へと“それとなく”導くことが可能かどうかについて検証していく予定である.

参考文献

[Hagman 15] Hagman, W. et al. Public Views on Policies Involving Nudges, *Review of Philosophy and Psychology*, Volume 6, Issue 3, pp 439–453 (2015).

[JST CRDS 17] 科学技術振興機構 研究開発戦略センター, 複雑社会における意思決定・合意形成を支える情報科学技術 https://www.jst.go.jp/crds/pdf/2017/SP/CRDS-FY2017-SP-03.pdf

[小川 06] 小川浩平, 小野哲雄(2006). ITACO: メディア間を移動可能なエージェントによる遍在知の実現, ヒューマンインタフェース学会論文誌, Vol. 8, No. 3, pp. 373-380.

[小野 13] 小野哲雄, 今吉晃.「空気を読むロボット」:コミュニケーション空間を利用した人を動かす HAI デザイン, 人工知能学会誌, Vol. 28, No. 2, pp. 284-289 (2013).

[小野 14] 小野 哲雄, 今吉 晃, 兼古 哲也. 信学技報 114(85), pp. 11-16, 2014-06-16.

[Ono 16] Ono, T., Ichijo, T., Munekata, N. (2016). Emergence of Joint Attention between Two Robots and Human using Communication Activity Caused by Synchronous Behaviors, IEEE RO-MAN 2016, pp. 1187-1190.

[Spencer-Brown 72] Spencer-Brown, G. (1972). *Laws of Form*, Julian Press.

[Thaler 08] Thaler, R. H. & Sustein, C. R. *Nudge: Improving Decisions About Health, Wealth, and Happiness,* Yale University Press (2008).

[津田 12] 津田一郎. 脳を創造する共創の場, 計測と制御, 51 巻, 11 号, pp. 1056-1058 (2012) .

[植田 16] 植田一博, 小野哲雄, 今井倫太, 長井隆行, 竹内勇剛, 鮫島和行, 大本義正. 意思疎通のモデル論的理解と人工物設計への応用, 人工知能学会誌, Vol. 31, No. 1, pp. 3-10 (2016).

[ユクスキュル 05] ユクスキュル, J. v., クリサート, G. 生物から見た世界, 岩波書店 (2005).

3:40 PM - 4:20 PM (Tue. Jun 4, 2019 3:20 PM - 5:00 PM Room G)

[1G3-OS-13a-02] (Invited talk) Cognitive science of choice and manipulation

The rocky road of nudge to well-being

〇Ayumi Yamada[1] （1. The University of Shiga Prefecture）

Keywords: nudge, behavioral economics, well-being, choice behavior, manipulation

Advances in behavioral economics and psychology have revealed the mechanism of human decision-making, and driven increasingly widespread practices of nudging people toward better choices. This lecture will span a range of theories, issues, and approaches that represent the frontier of nudge research. The lecture offers a general introduction of the idea of nudging, and touches on potential pitfalls as well as benefits that might result from nudging. It also provides a discussion of how the idea of nudging could be developed as tools for improved social interacting.

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

人は他者の意見をどう活用しているか：商品の性質に関する検討

How do we harness others' opinions?: An investigation on product types

藤崎　樹[*1]
Itsuki Fujisaki

本田　秀仁[*2]
Hidehito Honda

植田　一博[*1]
Kazuhiro Ueda

[*1] 東京大学大学院総合文化研究科
Graduate School of Arts and Sciences, The University of Tokyo #1

[*2] 安田女子大学心理学部
Department of Psychology, Yasuda Women's University #2

With the rising of information technology, we can conveniently harness many others' opinions. However, it remains unclear how we utilize them. Online review sites, such as Amazon.com, display a lot of products. Especially, it is well-known that products could have two types of consumption mode: hedonic or utilitarian. In this paper, we examined how these product types affected people's behavior on others' opinions. Concerning others' opinions, we focused on a rating distribution. In our experiment, participants conducted a binary product choice task. The task composed of single product name and two rating distributions which had similar average of ratings but different variance of ratings (high / low). The results showed that product types affected people's choices on rating distributions. For hedonic products, participants tended to prefer high-variance rating distribution more often than for utilitarian products. We discussed how the results emerged and illustrated how our findings could 'nudge' people.

1.　イントロダクション

近年，情報技術の台頭により，私たちは他人の意見を大量に，かつ手軽に入手できるようになった．しかし，それでは人が，その他人の意見をどのように活用しているかについては明らかになっていない点が多い．すでに先行研究[Fujisaki et al., 2018]では，Amazon に代表されるレイティングサイトを題材に，商品を誰のために購入するか（自分／他人）によって，意見の活用の仕方が変化することを明らかにしている．しかしその実験では，商品がただ一つ（コーヒーメーカー）に限定されており，多種多様な商品が持つ性質に関しては考慮されていなかった．

この点，消費者心理学の分野では，商品の性質によって感情ベースの消費（例：コメディー映画，以下，感情価と表記），実用ベースの消費（例：ドキュメンタリー映画，以下，実用性と表記）という，2種類の消費形態のいずれかが促されうることが知られている[e.g., Hirschman and Holbrook, 1982]．特に先行研究[Sen and Lerman, 2007]では，商品に対する否定的なレビューについて，感情価ではレビュワーにその原因が帰属され，実用性では商品自体に原因が求められることを明らかにしている．しかし[Sen and Lerman, 2007]では，レビューはただ1つに限定されていた．

そこで本研究では，多数の評価について，商品の性質がその捉え方に与える影響を検討する．具体的には，Amazon や食べログ等でも呈示されている，評価分布（[e.g., Fujisaki et al., 2018]）を対象とした．評価分布には，たとえ評価の平均値が近くとも，評価が安定した分散が小さい（以下，低分散と表記）ものもあれば，評価が分かれる分散が大きい（以下，高分散）ものもありうる（図1）．そこで，この高分散が持つ少数の低評価に関して，[Sen and Lerman, 2007]知見を基に以下の仮説を立てた．

仮説：感情価では，高分散の低評価の原因がレビュワーに帰属される．その結果，その低評価が無視される格好となり，高分散の選択率が高くなる．それに対し実用性では，その原因が商品自体に求められる．その結果，その低評価を無視できず，高分散の選択率が低くなる．

連絡先：藤崎樹，東京大学大学院総合文化研究科，東京都目黒区駒場 3-8-1，E-mail: bpmx3ngj@gmail.com

行動実験では，商品名を呈示した上で，分散が異なる2つの評価分布のうち一方を選択させる商品選択課題を実施し，この仮説を検証した．

2.　行動実験

2.1　参加者

43 名の大学生・大学院生が実験に参加した（年齢：$M = 19.79$, $Sd = 1.88$; 性別：女性 = 19 名，男性 = 24 名）．

2.2　刺激

まず，評価分布を先行研究[Fujisaki et al., 2018]に準拠して作成した．各評価分布は，Amazon のものと同様に評価値が星で表され，評価の数はバーの長さで示される．評価値は10段階に設定した（1が最低，10 が最高）そして，一方が高分散，もう一方が低分散からなる評価分布のペアを 20 個作成した．評価の平均値は，Amazon の値を参考に，いずれも 10 段階中7程度に設定した．なお，この評価の平均値は，高分散のほうが低分散よりも常に 0.5 高い．高分散の形状としては，「多数の高評価，少数の低評価，ほとんどない中程度の評価」から構成される，実

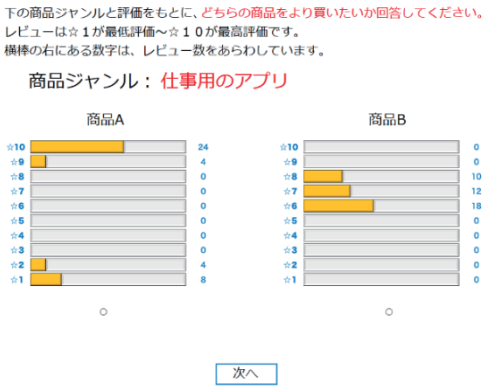


図 1. 実験画面． 商品ジャンルの性質は実用性．商品 A は高分散，商品 B は低分散．レビュー数は 40 のものを例示している．

際のレイティングサイトでもよく見られる J-shaped[e.g., Hu et al., 2009]と呼ばれるものを採用した．また，評価分布の総レビュー数としては4, 20, 40, 100 の4段階を設定し，ペア内の総レビュー数は同じとした．

次に，呈示する商品名を先行研究[Lu et al., 2016]に準拠して作成した．9つのカテゴリー（例：映画）について，感情価を1つ（例：コメディー），実用性を1つ（例：ドキュメンタリー），合計 18 個の商品名を設けた．

2.3 手続き

全参加者共通で商品選択課題を行わせた（図 1）．PC 上に商品名と評価分布のペアを呈示し，買いたい方を選択させた．課題は全 90 試行からなり，18 個の商品名が5回ずつ呈示された．なお，各試行で呈示される評価分布のペア，商品名の呈示順，呈示位置の左右は参加者ごとにランダマイズした．

また，課題前に「2つの商品は似たような性質を持ち，同じくらいの値段である」と想定するよう教示した．

3. 結果

結果を図 2 に示した．図 2 は，商品の各性質（感情価／実用性）における高分散の選択率を，参加者ごとにプロットしたものである．分析の結果，仮説通り，感情価で実用性よりも高分散の選択率が有意に高いことが明らかになった（t = 3.80, df = 42, p < .001）．また，効果量としては中程度のものが確認された（*Cohen's d* = 0.58[Cohen, 1988]）．なお，高分散の選択率の平均値は，感情価・実用性ともにチャンスレベルを上回っていた（感情価 = 0.60; 実用性 = 0.51）．これは，高分散の評価の平均値を，低分散に対し高く設定したためであると考えられる．

また，レビュー数の影響は，先行研究[Fujisaki et al., 2018]と同様に観察されなかった．具体的には，レビュー数を固定効果，参加者をランダム効果とした GLMM（統計ソフト R の lme4 を使用）による尤度比検定を行ったところ，感情価・実用性のいずれにおいても選択の予測率は改善されなかった（感情価: $\chi 2(df = 3) = 3.44$, $p = 0.33$; 実用性: $\chi 2(df = 3) = 3.12$, $p = 0.37$）．

4. 総合討論

情報技術の台頭を通じて，私たちは他人の意見を大量に，かつ手軽に入手できるようになった．本研究では，レイティングサイトにおける評価分布を題材に，商品の性質がそうした他人の意見の活用のしかたに与える影響を検討した．その結果，分散の高い評価分布について，感情ベースの消費を促す商品で，実用ベースの消費を促す商品よりも選ばれやすくなることが明らかになった．これは，高分散の評価分布が持つ少数の低評価に関して，以下の認知的なメカニズムが働いたためであると考えられる．まず，感情ベースの消費を促す商品では，低評価の原因がレビュワーに帰属された．その結果，その低評価が無視される格好となり，高分散の選択率が高くなった．それに対し実用性では，その原因が商品自体に求められた．結果，その低評価を無視できず，高分散の選択率が相対的に低くなった．

先行研究[Analytis et al., 2018]が議論しているように，人が他人の意見をどのように活用するか検討することで，その意見をどう提示すべきかについて示唆を与えることができる．本研究の知見の直接的な適用としては，レイティングサイトにおいて，感情（実用）ベースの消費を促す商品では，高（低）分散の評価分布を持つものを優先的に推薦すべきである，といったことが挙げられる．このように，人が他人の意見をどう捉えているのかを，認知心理学の観点から検討することを通じて，人々をよりよい消費活動へと「ナッジ」できる可能性がある．

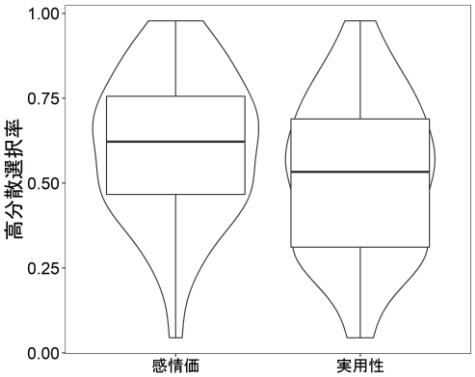


図 2. 実験結果

参考文献

[Analytis et al., 2018] Analytis, P. P., Barkoczi, D., & Herzog, S. M.: Social learning strategies for matters of taste. Nature Human Behaviour, 2, 415-424, 2018.

[Cohen, 1988] Cohen, J. Statistical Power Analysis for the Behavioral Sciences, 2nd Edn. Mahwah, NJ: Lawrence Erlbaum Associate, 1988.

[Fujisaki et al., 2018] 藤崎樹・本田秀仁・植田一博:「少数派」の意見が気になる時：商品選択の対象が評価分布への好みに与える影響．日本社会心理学会第 59 回大会発表論文集, 89, 2018.

[Hirschman and Holbrook, 1982] Hirschman, E. C., & Holbrook, M. B.: Hedonic Consumption: Emerging Concepts, Methods, and Propositions. Journal of Marketing, 46, 92–101, 1982.

[Hu et al., 2009] Hu, N., Zhang, J., & Pavlou, P. A.: Overcoming the J-shaped distribution of product reviews. Communications of the ACM, 52, 144-147, 2009.

[Lu et al., 2016] Lu, J., Liu, Z., & Fang, Z.: Hedonic products for you, utilitarian products for me. Judgment & Decision Making, 11, 2016.

[Sen and Lerman, 2007] Sen, S., & Lerman, D.: Why are you telling me this? An examination into negative consumer reviews on the web. Journal of interactive marketing, 21, 76-94, 2007.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

前傾体勢に伴うナッジ ~意思決定環境の操作から態度の変容を促す~

Nudge with a forward posture
~manipulation of a decision-making environment will enhance willingness to change one's attitude~

白砂　大[*1]
Masaru Shirasuna

本田　秀仁[*2]
Hidehito Honda

植田　一博[*1]
Kazuhiro Ueda

[*1] 東京大学
The University of Tokyo

[*2] 安田女子大学
Yasuda Women's University

Based on previous nudging and embodied cognition approaches, we proposed a new type of nudge: manipulating people's body posture by changing a decision-making environment could lead decisions in a certain direction. In this paper, our purposes were to investigate whether a forward posture really affect decision making, and if yes, what types of decisions could be likely to be influenced. Through two behavioral experiments, we found that participants' forward posture could enhance their decisions, especially willingness to change their own attitude. Our findings will provide an opportunity to consider the design of the real-world environment and artifacts for leading people to their predictable decisions.

1.　はじめに:「体勢」に着目したナッジ

近年，人間の行動や選択を特定の方向に導く手法として，「ナッジ(nudge)」の研究が注目されている[Thaler 08]．一般的なナッジの手法としては，「選択の構造(choice architecture)」を操作することで，意思決定を変化させるものとされてきた．

本研究では，意思決定環境を操作することによる新たなナッジのアプローチを検証した．具体的には，環境によって人間の「体勢」を変えることで，意思決定の変化を促すアプローチである．認知科学の分野では，心的状態と身体状態との関連について長年，議論されており，特定の身体状態が特定の心的状態を作り出しうるという「身体性認知(embodied cognition)」の研究も多く存在する[Ackerman 10, Bargh 12, Strack 88]．さらに近年，実際の対面での提案場面(具体的には，旅行販売員が顧客に旅行プランを提案する場面)において，「前傾体勢」の表出と，提案されたプランに対する興味との，強い相関も示唆されている[Honda 16]．一般には，「興味を示した(心的状態)から，前傾になった(身体状態)」という関係が想定されるだろう．しかし身体性認知研究の知見を踏まえると，「前傾(身体状態)が，興味(心的状態)を促した」という逆の関係も想定される．すなわち，前傾という体勢を作り出すことで，特定の望ましい意思決定を促せるかもしれない．

体勢と意思決定とを関連づけた研究として，[白砂 18]では，座面が前方に傾いた椅子(前傾椅子．詳細は後述)を用いた行動実験から，座面における圧力分布の重心が前方に置かれていた人の方が，後方に置かれていた人よりも，ポジティブな意思決定を行う可能性を示していた．しかしこの研究では，次の 2 点において問題があったと考えられる．第 1 の問題は，[白砂 18]自身も指摘していたように，椅子だけでは前傾体勢を十分に作り出せなかった可能性である．実際に，前傾椅子に座った実験参加者が全員，座面の重心位置を前方に置いていたわけではなかった．そのため，椅子だけでは，体勢を十分に操作できていなかったと考えられる．また第 2 の問題は，「課題実施中」における座面の圧力データを用いていた点である．なぜなら，参加者が課題に熱心に取り組んでいたために前傾となった可能性があり，実験環境(椅子)が前傾体勢を作り出していたかどうかについては分からないためである．環境が体勢を操作できていたかどうかを検証するためには，「課題開始時」における重心を考慮するべきだと考えられる．

そこで本研究では，参加者が確実に前傾体勢となるような実験環境を設定し，「課題開始時」における座面の圧力分布を測定することで(いずれも詳細は後述)，上記 2 つの問題点を改善した．そのうえで，「環境から形成された前傾体勢が，特定の意思決定を促すか」，またその場合に「どのような意思決定が促されるのか」の 2 点を検証することを目的に，行動実験を実施した．

連絡先：白砂大，東京大学大学院総合文化研究科，〒153-8902　東京都目黒区駒場 3-8-1，m.shirasuna1392@gmail.com

2.　実験 1

環境により作られた体勢が意思決定に与える影響について，[白砂 18]の実験環境を修正する形で検証した．具体的には，実験参加者がより確実に前傾体勢となるような実験環境を設定したうえで，体勢を，「課題開始時」における座面の重心座標から定義した．そのうえで，この実験 1 に新たに参加した参加者の回答と，[白砂 18]の「通常体勢群」に割り当てられた参加者の回答とを比較した．

2.1　方法

(1)　実験参加者

大学生 16 名が実験に参加した．うち 8 名は女性であり，*M*age = 19.6, *SD*age = 1.75 であった．本研究の実験目的に気づいた参加者は皆無であった．

図 1: 座面の傾きが 10 度の前傾椅子(左)および座面の傾きが 0 度の通常椅子(右)．いずれもコクヨ株式会社製．

(https://www.kokuyo-furniture.co.jp/manabi/product/campus_up/)

表 1: 実験 1 で用いた課題の概要と, その仮説

番号	課題テーマ	概要	仮説(前傾群の結果)
1	臓器提供意思を問う課題	「死後に臓器を提供する意思」を, 「0 絶対に提供したくない—100 非常に提供したい」で回答する	より提供する意思が高い
2	募金金額を問う課題	絶滅危惧動物であるハワイアンモンクアザラシの説明文を読み, その保護のための寄付金額を「0 円—1 万円」で回答する	より寄付金額が高い
3	議論の賛否を問う課題	ある議論の記事(題材は「滋賀県の名称変更」)を読み, その賛否について「0 非常に反対—50 どちらでもない—100 非常に賛成」で回答する	より「賛成」の態度を示す
4	最後通牒課題	「5000 円を自分と相手にどう分配するか」という状況で, 以下 10 通りの分配額を受け入れるか否か, それぞれ回答する: (自分, 相手) = (2500, 2500), (2750, 2250), …, (4000, 1000 円)	より低い金額でも受け入れる

課題 1~3 では, 参加者は visual analog scale (VAS)を用いて回答した.

(2) 実験環境および装置

参加者の体勢をより確実に前傾にするため, 以下のような実験環境を設定した. 参加者は, 座面が前方に 10 度傾いた椅子(図 1 左. 以下「前傾椅子[1]」と呼称)に座った. 椅子の前脚を机の脚から 30cm 離したところに, さらに, iPad を机の端から 15cm 離したところに, それぞれ設置. 参加者には, 椅子および iPad を実験中に動かさないようにすることを教示した. また, 課題中に参加者が座面のどこに重心をかけているかを確かめるため, 椅子の座面には薄型のシートセンサーを敷いた. 実験中は, 10 秒ごとに, 座面の圧力分布がシートセンサーに記録された.

(3) 実験課題および手続き

[白砂 18]と同様の意思決定課題 4 題を用いた. 具体的な概要および仮説を表 1 に示す. 参加者は, iPad 端末にて, 課題に回答した. 全ての参加者は 15 分以内に回答を終了した.

2.2 結果および考察

初めに, 実験環境から参加者の前傾体勢を作り出せていたかどうかを確認した. 具体的には, シートセンサーに記録された圧力分布に基づき, 課題開始時における座面の重心座標を参加者ごとに算出し, その座標が座面の前方に置かれていたかを確認した. シートセンサーの座標の範囲は, 前後方向が「1(前)から 30(後)まで」, 左右方向が「1(左)から 24(右)まで」であった. [白砂 18]の参加者について, 課題開始時の前後方向における重心座標の中央値を算出したところ, 13.95 であった. そこで本研究では, この値を, 座面の「前方」および「後方」を分ける基準座標とした. 実験1の参加者 16 名のうち, シートセンサーのデータに不備があった 2 名を除く 14 名について, 座面の重心座標を算出したところ, すべての参加者が, 基準座標 13.95 よりも前方に重心をかけていた(M = 9.71). このことから, 今回の実験環境により, 参加者の前傾体勢を作り出せていたと考えられる.

続いて, 体勢の違いによる意思決定課題の回答を比較した[2]. 実験 1 の参加者 16 名から得られたデータを「前傾群」と定義し, 比較対象として, [白砂 18]にて「通常体勢群」すなわち重心を座面の後方(前後方向の重心座標が, 基準座標である 13.95 より大きい)に置いていた参加者 16 名のデータを用いた. 本研究の実験 1 では, この 16 名を「通常群」と定義した. 結果として, 課題 1(臓器提供意思を問う課題)において, 前傾群と通常群との間に有意傾向が見られ, 前傾群の方が比較的, 提供意思が高かった(図 2. 課題 1: $W = 83.5$, $p = .096$, $d = 0.29$; 課題 2: $W = 121$, $p = .801$, $d = 0.04$; 課題 3: $W = 147$, $p = .485$, $d = 0.12$; 課題 4: $W = 94.5$, $p = .753$, $d = 0.06$). 前傾群と通常群との間で回答に差が見出されたことから, 前傾という体勢が, 意思決定に少なからず影響を与えることが示唆される.

では, 前傾体勢は, 具体的にどのような意思決定に影響を与えるのだろうか. 有意傾向が見られた臓器提供意思を問う課題は, 他の 3 課題と比較して, 他人のために何かを行おうとする, すなわち「自身の態度を変えようとする」意思を反映していると考えられた. そこで次の実験 2 では, 新たな意思決定課題を作成・実施し, この点を検証した.

3. 実験 2

実験 1 の結果を踏まえて, 実験 2 では, 前傾体勢が「自身の態度を変えようとする意思」を促すかどうかについて検証した. 具体的には, 「他人の推定をどの程度受け入れるか」を検証する advice-taking 課題[Yaniv 00]を題材とし, 前傾体勢および通常体勢との間で回答を比較した.

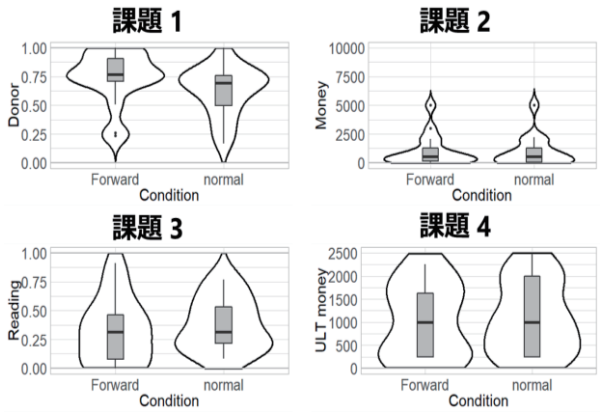


図 2: 実験 1 の結果. 左が前傾群(実験 1 の参加者), 右が通常群([白砂 18]における通常体勢群)の回答値の分布を示す.

[1] 本研究では, この「前傾椅子(図 1 左)」に対し, 座面の傾きが 0 度の通常の椅子(実験 2 で使用)を「通常椅子(図 1 右)」と表記する. いずれの椅子もコクヨ株式会社製であり, [白砂 18]で用いられたものと同一である.

[2] 本研究では, データに正規分布を仮定できなかったため, 2 群間の比較にはすべてノンパラメトリック検定の一種である「マン・ホイットニーの U 検定」を適用した.

3.1 方法

(1) 実験参加者および条件

大学生 10 名が実験に参加した．うち 4 名は女性であり，*M*age = 20.7，*SD*age = 2.16 であった．実験 1 と同様，本研究の実験目的に気づいた参加者は皆無であった．条件としては，実験 1 の実験環境(iPad を机の端から 15cm，前傾椅子を机の脚から 30cm，それぞれ離す)を「前傾条件」，[白砂 18]における「通常群」の実験環境(iPad を机の端から 10cm，通常椅子を机の脚から 10cm，それぞれ離す)を「通常条件」とした．各条件には，5 名ずつの参加者がランダムに割り当てられた．

(2) 実験課題，手続きおよび仮説

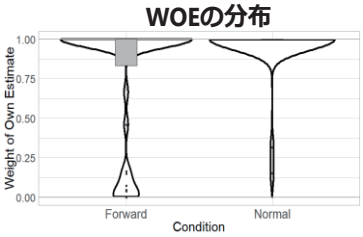
「他人の推定をどの程度受け入れるか」を測る課題として，[Gino 08]および[Yaniv 00]の advice-taking 課題を改題したものを用いた．課題は，「歴史上の出来事が，西暦何年に起こったか」を推定する一般知識課題であった．参加者は，呈示された出来事(例「アメリカにおける奴隷解放宣言の発表」，「樺太・千島交換条約の締結」)について，「その出来事が起こった西暦年」および「95%の確率で正答が含まれると思う範囲(下限と上限)」を回答するよう求められた．設問は，世界の歴史 5 問および日本の歴史 4 問の，計 9 問であった．

この実験課題は，2 つのフェーズ[3]から構成された．第 1 フェーズは実験冒頭に実施され，参加者は 9 つの各出来事の西暦年および範囲について，何も参照せず自身で推定を行った(このときの回答を「最初の推定値」と表記する)．第 2 フェーズは実験末尾に実施され，参加者は第 1 フェーズと同じ課題に回答した．ただしこのとき，参加者には「他人の推定値」が呈示されており，参加者はこの値を参照しながら再度，推定を行うことができた(このときの回答を「2 回目の推定値」と表記する)．なお，この「他人の推定値」は，実際には実験者により生成された乱数であった[4]．仮説として，前傾群の方が，通常群よりも他人の推定値をより受け入れる(通常群の方が，より最初の推定値に固執しやすい)ことが予想された．

3.2 結果および考察

初めに，実験環境から参加者の体勢を操作できていたかどうかについて，実験 1 と同様の手続きにより確認した．課題開始時において，前傾群，通常群における前後方向の重心座標の平均値はそれぞれ 8.88，15.9 であり，前傾群の方が有意に前方へ重心を置いていた ($W = 0$, $p < .01$, $d = 0.84$)．このことから，今回の実験環境も，体勢を操作するうえで十分に機能していたことが考えられた．

続いて，体勢による「他人の推定値」を受け入れる程度を検証した．測定指標として，[Yaniv 00]による Weight of Own Estimate (WOE)を採用した．WOE は「| 他人の推定値 − 2 回目の推定値 | ÷ | 他人の推定値 − 最初の推定値 |」で定義され，「最初の推定値と他人の推定値が異なる」かつ「2 回目の推定値が『他人の推定値』と『最初の推定値』の間に位置する」ときに適用可能となる[5]．WOE の値が大きいほど，「自身の最初の推定に固執している程度が高い」ことを意味し，最初の推定値から動かない(他人の推定値を受け入れない)場合は WOE = 1，完全に他人の推定値に合わせた場合は WOE = 0 となる．前傾群および通常群について，この WOE の平均値を比較したところ，前傾群の方が有意に低かった(図 3 上; $W = 709$, $p < .01$, $d = 0.30$)．このことから，前傾群の方が他人の推定値を受け入れやすいという仮説は，支持されたといえる．



WOEを予測する回帰モデル

独立変数	回帰係数	P値	変数選択(Stepwise)後の係数
切片	0.702	< .001	0.707
実験環境 (通常)	0.178	< .01	0.180
自身の正確さ	-0.003	< .001	-0.003
自身の情報性	0.001	.146	0.001
他人の正確さ	-0.012	.270	(選択されず)
他人の情報性	0.003	< .05	0.002

図 3: 実験 2 の結果．上のグラフでは，左が前傾群，右が通常群の回答値の分布を示す．

しかし先行研究[Yaniv 97]において，一般知識課題に対する推定の際には，「正確さ(accuracy)」と「情報性(informativeness)」も重要な要素であることが議論されている．今回の課題では，正確さとは，「出来事が起こった西暦年」に対する「推定値と正答との誤差(大きいほど，正確さに乏しい)」であり，情報性とは「正答が含まれると思う範囲の大きさ(大きいほど，情報性に乏しい)」である．そこで，[Yaniv 00]と同様の回帰モデルから，WOE を予測した．具体的には，従属変数を WOE とし，独立変数に「実験環境(前傾または通常)」のほか，「自身の正確さ(| 正答 − 最初の推定値 |)」，「自身の情報性(| 最初の下限値 − 最初の上限値 |)」，「他人の正確さ(| 正答 − 他人の推定値 |)」，および「他人の情報性(| 他人の下限値 − 他人の上限値 |)」を投入した．結果として，これら 4 つの独立変数が存在してもなお，「実験環境」の要因が有意に WOE を予測した．さらに，[Yaniv 00]と同様の stepwise による変数選択を実施した結果，「自身の正確さ」，「自身の情報性」，「他人の情報性」に加えて，「実験環境」も選択された．またその際の「通常条件」における実験環境の回帰係数は正の値(0.180)を示しており，通常群の方がより自身の最初の推定値に固執しやすいことが確認された(図 3 下)．

以上から，環境により作られた体勢が，自身の態度を変えようとする意思，特に他人の推定を受け入れようとする意思を促すことが示唆された．

4. 総合考察

本研究では，前傾体勢が意思決定に影響を与えるかどうか，また，与える場合に，どのような意思決定が促されるのかについて検証した．実験 1・2 を通して，環境から形成された前傾体勢が意思決定に影響を与えうること，特に「自身の態度を変えようとする意思」が促進されることが示唆された．先述の通り，実験目的に気づいた参加者は皆無であったことから，今回見られた効果は，当人が気づかない程度の環境操作であっても十分に生じるといえる．

[3] 第 1 フェーズと第 2 フェーズの間で，実験参加者は他の実験課題にも回答していた．これらの課題を含めて，すべての参加者は，30 分以内には回答を終了していた．

[4] 「平均 = 正答の西暦年，標準偏差 = 50」の正規分布に従う乱数を 1000 個発生させ，その平均値を呈示した．

[5] 今回の実験 2 において，WOE を適用できたケースは全体の 92.1%であった．以降は，これらのケースを分析対象とした．

これまでのナッジ研究では，選択の構造や情報呈示の仕方を操作することで，特定の意思決定を促すものであった[Johnson 03]．また近年の身体性認知研究では，体勢(特に，座面の圧力分布)から人間の内的状態を予測したり[D'Mello 07]，実世界においても体勢が特定の行為を引き起こすことを示唆したり[Yap 13]するものも見られる．しかし本研究のように，環境を操作して体勢を変えることで特定の意思決定を促すという，いわば「環境による体勢変化」を媒介としたナッジのアプローチは，我々の知る限り存在しない．本研究は，望ましい意思決定を促すための，環境や人工物のデザインを考える契機となることが期待されるだろう．

5. 謝辞

本研究は，東京大学とコクヨ株式会社との共同研究として実施された．また，東京大学植田一博研究室の今泉拓氏(修士課程 2 年)には，実験プログラムなどの面において多大なご協力を頂いた．ここに謝意を記す．

参考文献

[Ackerman 10] J. M. Ackerman, C. C. Nocera, and J. A. Bargh, "Incidental haptic sensations influence social judgments and decisions.," *Science*, vol. 328, no. 5986, pp. 1712–5, 2010.

[Bargh 12] J. A. Bargh and I. Shalev, "The substitutability of physical and social warmth in daily life.," *Emotion*, vol. 12, no. 1, pp. 154–162, 2012.

[D'Mello 07] S. D'Mello, P. Chipman, and A. Graesser, "Posture as a Predictor of Learner's Affective Engagement," in *Proceedings of the 29th Annual Conference of the Cognitive Science Society*, 2007, pp. 905–910.

[Gino 08] F. Gino, "Do we listen to advice just because we paid for it? The impact of advice cost on its use," *Organ. Behav. Hum. Decis. Process.*, vol. 107, no. 2, pp. 234–245, 2008.

[Honda 16] H. Honda, R. Hisamatsu, Y. Ohmoto, and K. Ueda, "Interaction in a Natural Environment: Estimation of Customer's Preference Based on Nonverbal Behaviors," in *Proceedings of the Fourth International Conference on Human Agent Interaction (HAI '16)*, 2016, pp. 93–96.

[Johnson 03] E. J. Johnson and D. Goldstein, "Do Defaults Save Lives?" *Science*, vol. 302, no. 5649, pp. 1338–1339, 2003.

[白砂 18] 白砂大, 本田秀仁, 植田一博, "前傾姿勢は意思決定に影響を与えるか?～前傾椅子を用いた実験的検討～" *日本認知科学会第 35 回大会*, 2018, pp. 204–206.

[Strack 88] F. Strack, L. L. Martin, and S. Stepper, "Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis.," *J. Pers. Soc. Psychol.*, vol. 54, no. 5, pp. 768–777, 1988.

[Thaler 08] R. H. Thaler and C. R. Sunstein, *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Penguin Books, 2008.

[Yaniv 97] I. Yaniv and D. P. Foster, "Precision and accuracy of judgmental estimation," *J. Behav. Decis. Mak.*, vol. 10, no. 1, pp. 21–32, 1997.

[Yaniv 00] I. Yaniv and E. Kleinberger, "Advice Taking in Decision Making: Egocentric Discounting and Reputation Formation," *Organ. Behav. Hum. Decis. Process.*, vol. 83, no. 2, pp. 260–281, 2000.

[Yap 13] A. J. Yap, A. S. Wazlawek, B. J. Lucas, A. J. C. Cuddy, and D. R. Carney, "The Ergonomics of Dishonesty: The Effect on Incidental Posture on Stealing, Cheating, and Traffic Violations," *Psychol. Sci.*, vol. 24, no. 11, pp. 2281–2289, 2013.

Organized Session | Organized Session | [OS] OS-13

[1G4-OS-13b] “ナッジ”エージェント：人をウェルビーイングへと導くエージェントの構築へ向けて(2)

小野 哲雄（北海道大学）、今井 倫太（慶應義塾大学）、植田 一博（東京大学）、山田 誠二（国立情報学研究所

Tue. Jun 4, 2019 5:20 PM - 6:40 PM Room G (302A Medium meeting room)

[1G4-OS-13b-01] Framework for Human-Agent Team using Implicit Information Showing via Behavior

〇Ryo Nakahashi[1], Seiji Yamada[2,1] （1. The Graduate University for Advanced Studies(Sokendai), 2. National Institute of Informatics）

5:20 PM - 5:40 PM

[1G4-OS-13b-02] To determine the display timing of driving assistant using Nudge

〇Kouichi Enami[1], Michita Imai[1], Kohei Okuoka[1], Shohei Akita[1] （1. Keio University）

5:40 PM - 6:00 PM

[1G4-OS-13b-03] Adaptive Trust Calibration for Human-AI Collaboration

〇Kazuo Okamura[1], Seiji Yamada[2,1] （1. SOKENDAI (The Graduate University for Advanced Studies), 2. National Institute of Informatics）

6:00 PM - 6:20 PM

[1G05-07-4] Discussion / Conclusion

6:20 PM - 6:40 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

行動からの暗黙的な情報相互伝達を利用した Human-Agent Team フレームワーク

Framework for Human-Agent Team using Implicit Information Showing via Behavior

中橋 亮 *1　山田 誠二 *2*1
Ryo Nakahashi　Seiji Yamada

*1総合研究大学院大学
The Graduate University for Advanced Studies (Sokendai)

*2国立情報学研究所
National Institute of Informatics

The problem that autonomous agents and humans cooperate to solve one task is one big theme in the field of Human Agent Interaction. We are interested in cooperative agents that improve human plans naturally not just assisting human plans. As a situation where such an agent is useful, there is situation where a person and an agent have only part of information to achieve tasks. We developed a framework of agents that cooperates on the premise that humans and agents implicitly communicate information through actions with each other in this situation. We formulated the target situation as a Human-Agent Team problem and developed an agent planning method for this problem. This method is composed of two parts, the model that human's predict other purpose and the planning algorithms under the models, which are realized by improving existing methods called CIRL and Bayesian Inverse Planning, respectively. We evaluate our method through participant experiments that humans achieve simple tasks with autonomous agents and confirmed that our method has good performance of collaborative work.

1. はじめに

人工知能・Human-Agent Interaction の分野において，一つの作業において人間と協力しながら一つの作業 (タスク) を行うような協調エージェントの開発は一つの大きなテーマである．スムーズな協調作業のためには人間から逐一行動の指示を待つのではなく，エージェントは現在の状況から自分の行うべき行動を自律的に決定する必要があつ．このような自律協調エージェントへのアプローチとして，人間の行動からその目的を予測し，その上で自分の行うべき行動を決定する自律エージェントなどが考えられるが，このアプローチは人間が行おうとしている手順を補助するという目的を達成するということにフォーカスしており，そもそもの手順が間違っていたり，非効率であったりした場合に効率的なタスクの達成を実現することができない．この状況を改善するためには，協調エージェントに人間の行動手順そのものを正しく修正する機能が必要になる．しかし，人間は例え全体のタスク実行のパフォーマンスが下がっても自らタスクの主導権を持つほうが満足度が高いという研究結果が存在する [Nikolaidis 17] ように，人間の手順をエージェントが強引に変更することで人間にとってエージェントの印象を下げるリスクがある．したがって，エージェントが明に手順変更の指示を行わず，人間に自律的に手順を変更するよう促すことが重要になる．このような暗黙的な手順変更のためには，エージェントは人間の手順や目的を推測する必要があり，また．人間はエージェントの行動からエージェントの意図を読み取り，手順を変更してもらう必要がある，このような協調の系として人間とエージェントがお互いの行動を観察し，それに基づいて意図や目的を推測・伝達しあい，それに基づいて協調動作を実現するフレームワークの開発に取り組んでいる．協調の媒介に行動の観察を利用している理由として，人間は他人の行動から無意識に意図を推定していまい [Baker 09]，他の媒介を利用したときに行動から無意識に推定された意図と認知的に競合が起きる可能性を回避するため，また行動観察における情報伝達スピードが速いこと，通信のための装置が何もいらないことなどがある．これらのメリットにより，例えば通信プロトコルの構築が難しい自動運転車と人間が操作している車との協調や，通信インフラを当てにできない大規模災害時の自立ロボットと人との協調など，有用なシチュエーションが多く存在すると考えている．

先に述べたような人間の行動手順が間違っている典型的な原因の一つは，人間が最善の行動計画に必要な情報すべてを把握しておらず，情報不足により最善な行動を計画できない場合である．我々はこのような状況における協調問題を Human-Agent Team(HAT) 問題として定式化した．HAT は人間とエージェントが協調して一つのタスクを解く問題の定式化の一つである，Cooperative Inverse Reinforcement Learning(CIRL) Game [Hadfield-Menell 16] の拡張として定式化される．HAT において，人間，エージェントはタスクの状態や状態遷移に関しては完全な情報を持つが，報酬に関しては報酬を定義づけるパラメータの一部しか知らないものとする．この報酬パラメータが人間やエージェントがもつ相手が知らない情報や意図に対応する．HAT の問題を解くためには相手の行動から相手がもつ報酬に関するパラメータを推測する必要がある．

我々は上記のような HAT 問題を解くためのフレームワークを検討している．このフレームワークは (1) 人間のエージェントに対する報酬パラメータ予測モデルと (2) 人間の予測モデルを組み込んだ，エージェントの協調行動のための行動ポリシー計算アルゴリズムという 2 つの要素から構成されている．(1) に関してはは人間がエージェントの行動をみてエージェントがどのような報酬パラメータを持っているのかを推測するモデルの構築することであり,「心の理論」の研究で良く用いられている Bayesian Inverse Planning[Baker 09] を基に，人間の限定合理性を考慮した新たな予測モデルを考案した [Nakahashi 18]．(2) に関しては人間が (1) のモデルをもってエージェントの報酬パラメータを予測するという前提の元，効率的に協調する行動ポリシーを計算するアルゴリズムのことであり，CIRL game を効率的に解くための手法 [Malik 18] を我々の HAT 向けに

連絡先: 中橋 亮，総合研究大学院大学 複合科学研究科 情報学専攻 山田研究室，東京都千代田区一ツ橋 2-1-2，ryon@nii.ac.jp

拡張することで実現した．

また，我々の手法の有用性を確認するため，災害救助をモチーフとした人間とエージェントの協調タスクを作成し，Web 上で参加者実験を行い，実際に人間とエージェントの協調タスクのパフォーマンスを測定した．その結果，我々の手法がその他の比較アルゴリズムに対して高いパフォーマンスを持つことが確認できた．

2. 問題設定

2.1 CIRL game

我々の HAT の定式化の説明の前に，基となっている CIRL game の定式化について解説する．CIRL game において，人間とエージェントは一つの報酬関数を共有しているが，人のみがその報酬の情報を完全に知っており，エージェントは一部しか知らない．つまり報酬関数は報酬パラメータによって定義づけられており，その報酬パラメータは人間のみが知っている．共有された報酬関数における収益を最大化するように人間とエージェントが行動を計画することによって，人間とエージェントの協調戦略，すなわちエージェントは人間の行動から報酬パラメータを推測し，また人間はエージェントに自らの報酬パラメータを伝えるような行動が獲得される．

今回，CIRL game を次のタプル表現で定義する．$< \mathcal{S}, \mathcal{A}^H, \mathcal{A}^A, T, \Theta^H, R >$．$\mathcal{S}$ はタスクの状態空間 $s \in \mathcal{S}$，$\mathcal{A}^H$，$\mathcal{R}^A$ はそれぞれ，人間，エージェントの行動空間 $a^H \in \mathcal{A}^H$，$a^A \in \mathcal{A}^A$，T はタスクの状態遷移関数 $T(s'|s, a^H, a^A)$，Θ^H は人間のみが知っている報酬パラメータの集合 $\theta^H \in \Theta^H$，R は報酬パラメータで条件付けられた報酬関数 $R(s, a^H, a^A; \theta^H)$ を表している．

2.2 HAT 問題の定式化

我々は CIRL game を拡張して HAT 問題の定式化を行った．変更点としてはエージェントのみが知っている報酬パラメータを導入したことである．これにより，HAT は次のタプルで定義される．$< \mathcal{S}, \mathcal{A}^H, \mathcal{A}^A, T, \Theta^H, \Theta^A, R >$. $\theta^A \in \Theta^A$ はエージェントのみが知っている報酬パラメータの集合であり，R は各種パラメータで条件付けられる報酬関数 $R(s, a^H, a^A; \theta^H, \theta^A)$ に変更される．残りの定義は CIRL と同等である．

HAT の目的は今までの人間とエージェントの行動から，最適なエージェントの行動を求めることである時刻 T において，今までのタスクの状態列を $\mathbf{s}_{1:T}$，人間，エージェントの行動列を $\mathbf{a}^H_{1:T}$，$\mathbf{a}^A_{1:T}$ とすると，行動ポリシー $p(a^A|s, \theta^R, \mathbf{s}_{1:T}, \mathbf{a}^H_{1:T}, \mathbf{a}^A_{1:T})$ を求めることである．

3. 手法

3.1 CIRL game の行動ポリシー計算アルゴリズム

まず，我々の行動ポリシー計算アルゴリズムの基となる，CIRL game における行動ポリシー計算アルゴリズムを解説する．CIRL はマルチエージェントの協調問題で利用される Decentlized POMDP の一種であると捉えられる．この問題は NEXP-hard のクラスに属し，解を求めるのに非常に多くの計算を要する．しかし，POMDP における Value Itaration の定式化をを一部修正することにより現実的な時間で CIRL の解を求める方法が提案されている [Malik 18]．CIRL game をエージェントが行動の主体をから見た POMDP 問題として捉えると，CIRL のタスクの状態，人間の報酬パラメータの空間の直積空間 $\mathcal{S} \times \Theta^H$ が POMDP の真の状態空間，人間の行動空間が POMDP の観測空間に対応付けられる．ただし，POMDP と異なり，人間の行動確率 $P(a^H|\{s, \theta^H\}, a^R)$ が一意に決定されない (すなわち CIRL は POMDP ではない) ため，純粋な POMDP の手法では解は得られない．ここで CIRL game の目的は収益の最大化であることに着目し，人間も自らの報酬パラメータに従って最も期待報酬が高くなるように行動すると仮定，すなわち，下記の確率に従って行動するとする．

$$P(a^H|\{s, \theta^H\}, a^A) = \mathbb{1}(a^H = \arg \max_{a'^H \in \mathcal{A}^H} Q(s, a'^H, a^A; \theta^H))$$

ここで $Q(s, a^A, a'^H; \theta^H)$ は人間の報酬パラメータが θ^H の場合において，タスク状態 s のときに人間・エージェントの行動 a'^H，a^A が行われた場合の行動価値関数である．

この修正を施すことにより，CIRL の行動計画を POMDP の従来の POMDP の解法を用いて求めることが出来る．

3.2 HAT のの行動ポリシー計算アルゴリズム

上記の CIRL の行動計画に更に修正を施すことで HAT の行動計画に対しても同様に POMDP の解法を用いることが出来る．CIRL との相違点はエージェントしか知らない報酬パラメータの存在である．CIRL ではこれが存在していなかったため，人間が行動価値関数を正確に計算できるという仮定が置けたが，HAT においてその仮定は利用できない．よって人間がエージェントの報酬パラメータを推測し，それに基づいて行動価値関数を計算するというモデルに置き換える．これによって行動価値関数の計算は下記の式で表現されるようになる．

$$Q(s, a'^H, a^A; \theta^H)) = \sum_{\theta^A \in \Theta^A} p(\theta^A|s, a^A) Q(s, a'^H, a^A; \theta^H, \theta^A))$$

ここで $p(\theta^A|s, a^H)$ は人間のエージェントに対する報酬パラメータの予測モデルであり，3.3 章で詳細を解説する．

3.3 報酬パラメータ予測モデル

人間のエージェントに対する報酬パラメータの予測モデルは，人がエージェントの行動から，エージェントの報酬パラメータを推測するモデルである．これは人が他人の行動からその人の目的を推定するということであり，認知科学では「心の理論」の文脈で盛んに研究されている．そのなかで，Bayesian Inverse Planning[Baker 09] に基づくモデルが良く用いられている．このモデルは，ある状態 s で行動 a をとったときにその目的 θ は下記のように計算されるとする．

$$\begin{aligned} p(\theta|s, a) &\propto p(a|s, \theta) \\ p(a|s, \theta) &= \frac{\exp(\beta Q(s, a; \theta))}{\sum_a \exp(\beta Q(s, a; \theta))} \end{aligned} \tag{1}$$

このアプローチは人は他人の行動モデルを「行動価値が高い行動ほと高い確率で行われる」，すなわち，他人の行動モデルに合理性を仮定して，その行動確率を尤度としたベイズ推定を行動の目的を推定している．

Bayesian Inverse Planning は様々な問題に使える汎用的な手法だが，一部の複雑な問題には適用が難しい．これは，Bayesian Inverse Planning が人間が全ての目的に対して全ての行動の尤度，つまりどれほど効果的かが計算できているという，合理性に関する情報が完全なことを前提においているが，人間はかならずしも完全に合理的に計算ができない限定合理性を持つからである．具体的には Bayesian Inverse Planning に置いてある行動 a を観測した場合，ある目的 θ に対し，今とった行動より有用な別の行動 a' が存在すれば，a を観測後の目

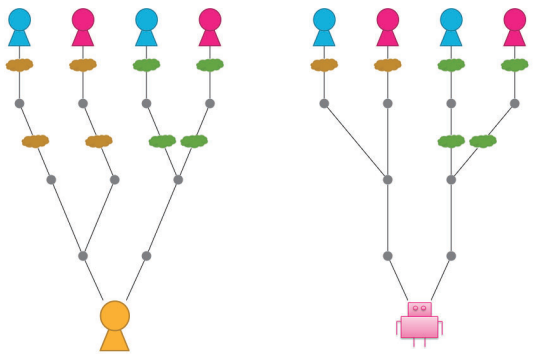


図 1: 参加者実験で用いた HAT タスク

的 θ の予測確率は低くなる．なぜならばもし人が目的 θ を達成しようとするのであれば，その行動 a' を実行すべきであるからである．このような考えは考慮すべき目的や行動の数が膨大になればなるほど，実際の人間には難しくなる．我々はこのようなズレを緩和する限定合理性を組み込んだモデルとして Predictability Bias というバイアスを想定した予測モデルを考案した [Nakahashi 18]．その予測モデルは下記である．

$$p(\theta|s,a) = \frac{\exp(\beta Q(s,a;\theta))}{\sum_a \exp(\beta Q(s,a;\theta))} \quad (2)$$

これは，端的には今他人が行った行動が目的に対して如何に有効であったかという軸で目的を推測するモデルである．すなわち，ある行動 a を観測した場合にある目的 θ に対してより有用な行動 a' があろうとも，純粋に行動 a が目的 θ に有効であれば，目的 θ の予測確率を上げる．他の行動を精査せず，ある行動に対し直感的に関連性の強い (予測しやすい) 目的に強い予測確率を割り当てる手法のため，Predictability Bias に基づくモデルと呼んでいる．

4. 実験

我々の手法の精度を確認するため，人間がエージェントと協調して一つの目的を達成する協調タスクを作成し，それを用いて参加者実験を行った．

4.1 作成したタスク

図 1 が作成したタスクである．このタスクは災害時の被災者救助のメタファーとなっており，参加者とエージェントが手分けして被災者をできるかぎり多く救出する (被災者の位置まで移動する) ことを目的とする．ただし，被災者は 2 種類のタイプ (青，赤) が存在しており，どちらを救出すべきかはタスクの開始時に参加者に指示されるが，エージェントはその情報を持たない．また，災害のため一部の道路は汚染されており，その道路を通過することで参加者・エージェントはペナルティを受ける．汚染に関しても 2 種類 (茶，緑) が存在しており，どちらのほうが多くのペナルティを受けるかどうかに関しての情報はエージェントのみがわかっており，参加者に伝えられることはない．このタスクの解決のためには参加者とエージェントが，救助すべき被災者と回避すべき汚染の別々の情報を推測しあう必要がある．また，タスクの実行においてエージェント・参加者の順で交互に 1 歩ずつ動く (灰色のドットに移動する・もしくは被災者にたどり着く) ものとし，上方向にしか動けないものとする (つまり，後戻りはできない)．

4.2 実験シナリオ

我々の手法の汎用性を示すため，3 つの異なる協調シナリオを考案し，それぞれその協調動作がタスクの最適計画となるタスクを作成した．考案した協調シナリオは下記の 3 つである．

1. エージェントは参加者がどの種類の被災者を助けようとしているかによって，汚染に関する情報を変更する必要がある．すなわち，エージェントは汚染の情報を伝える前に参加者の目的を推測する必要がある．
2. エージェントは参加者の被災者の目的を推測する前にいち早く汚染に関する情報を伝える必要がある．
3. エージェントは参加者の被災者の目的を推測する前にいち早く汚染に関する情報を伝える必要がある．ただし，自らの報酬パラメータを正しく伝えるのではなく，あえて間違った報酬パラメータを伝える

図 1 はシナリオ 3 の例である．エージェントは緑の汚染のほうが茶色の汚染よりひどいことを知っているものとする．エージェントは自らの情報を正しく伝えるため茶色の汚染を避けた右に動くべきだが，その結果参加者が同様に右に向かってしまうと，エージェントが青と赤の被災者を選択するときに参加者がどちらに向かっているかを判別することができない．よって参加者の被災者の情報を必要なタイミングで手に入れるために，あえて左に向かい参加者の行動を左に誘導することが最も良い協調行動になる．

4.3 実験アルゴリズム

実験のためのアルゴリズムは我々の手法で 2 つ．対抗手法で 2 つの計 4 つ準備し，各シナリオと各アルゴリズムの全ての組み合わせで実験を行った．利用したアルゴリズムは下記の 4 つである．

(A) 我々の手法 (人間の報酬パラメータ予測モデルに Predictability Bias モデルを用いたもの)

(B) 我々の手法 (人間の報酬パラメータ予測モデルに Predictability Bias モデルを用いず，Bayesian Inverse Planning モデルを用いたもの)

(C) Optimistic CIRL モデル (人間はエージェントの報酬パラメータの正解を知っていると仮定して CIRL を利用)

(D) Min-Max モデル (人間がもっとも不都合に動いたときの影響が最小になるように行動)

4.4 実験手法

Yahoo クラウドソーシングを用いて，実験を行った．実験は Web 上で行い，インストラクションを行った後，理解度確認テストを 2 種行い，それに正解した被験者のみ実実験に進めるようにした．この結果 71 名 (男性 36 名，女性 21 名，無回答 14 名) の被験者の結果を集めた．

4.5 実験結果

図 2 は各シナリオ，各アルゴリズムにおいて，参加者とエージェントが協調した場合の獲得したスコアの平均値である．Oracle は参加者とエージェントがそれぞれの報酬パラメータを知っていて，完全に協調が成功した場合の想定スコア，すなわち各シナリオでの理論的な最高スコアの値を示している．この結果が示すように我々の手法，とくに Predictability Bias を導入したモデルが他のアルゴリズムに比べ，高いスコアを獲得できていることが分かる．

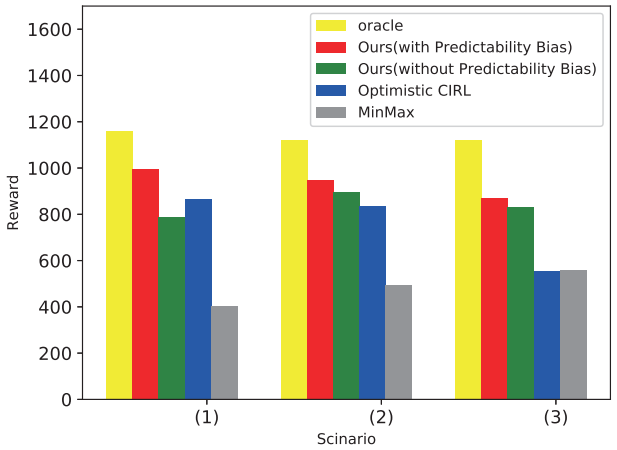


図 2: 各シナリオにおける各手法の性能評価

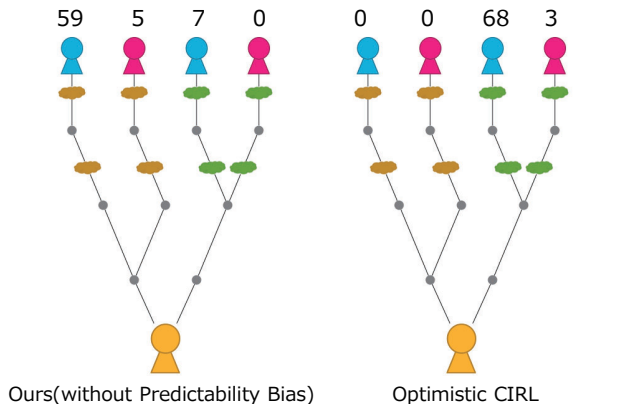


図 3: シナリオ 3 の例における参加者の最終目的地

具体例として，図 3 にシナリオ 3 の例における我々の手法と Optimistic CIRL における各被験者の最終目的値の違いを示した．この結果が示すように我々の手法は被験者がほとんど一番左に被験者に向かっているのに対し，Optimistic CIRL は左から 3 番目に向かっている被験者がほとんどである．これは我々の手法においてはエージェントが図 2 において 4.2 で説明したように左に向かい，結果参加者が一番左の被災者に誘導されたことを示している．一方，Optimistic CIRL に関しては参加者がエージェントの報酬パラメータを知っていることを想定しているため，エージェントはあえて汚染のひどい左に行くモチベーションがなく，右に向かって移動する．結果，被験者も右に誘導されている．

また，本稿では紙面の都合上省略するが，我々の手法が対抗手法に比べてスコアの分散が低い，すなわち安定した結果を返すことができること，行動計画で計算した期待値 (初期状態における状態価値関数の値) が実際のスコアに比べて近い，すなわちより精度高く参加者の行動を予測できていることの確認も行っている．さらに被験者に対して 7 段階でエージェントに関して協調していると感じるかどうかのアンケートも行い，Predictability Bias を導入した我々の手法が他の手法に対して被験者が優位に協調していると感じたという結果も得ている．

5. まとめ

本稿では，エージェントと人間がそれぞれ，タスクに関する情報を一部しか知らない状態で協調して一つの問題を解くような環境を想定し，Human-Agent Team(HAT) 問題として定式化した．さらに，HAT においてお互いの行動からお互いの情報を推測し，効率的に人間と協調するためのエージェントのフレームワークについて解説した．このフレームワークは (1) 人間のエージェントに対する報酬パラメータ予測モデルと (2) 人間の予測モデルを組み込んだ，エージェントの行動ポリシー計算アルゴリズムの 2 つからなり，それぞれ，Bayesian Inverse Planning，CIRL という 2 つの既存研究をそれぞれ，限定合理性の導入，人間の予測モデルの導入という改善を施すことで HAT に対して高い精度を出すように改善した．また，災害救助をモチーフとした協調タスクにおいて，実際の人間と比べて高いパフォーマンスが出ることを確認した．本研究はまだ初期段階であり，実用までに様々な課題が存在する．最も大きいものは現在の人間，エージェントそれぞれが持つ情報は離散情報のみであり，非常に限定されたものである．これらを連続情報化，多次元化と拡張していくことがより現実的なアプリケーションの応用を見据えたときに重要になる．また，より個人にカスタマイズされた協調動作の獲得のために，今までの人の行動から人の特有の報酬情報を逆強化学習などの手法を用いて事前学習しておき，その結果を協調プランニングに導入するという方向性も今後の研究の方向性としては非常に興味深い．

謝辞

本研究は JSPS 科研費 JP26118005 の助成を受けた。

参考文献

[Baker 09] Baker C. L., Saxe R., and Tenenbaum, J. B. : Action understanding as inverse planning, in *Cognition*, vol 1(4), pp. 329–349 (2009)

[Hadfield-Menell 16] Hadfield-Menell D., Dragan A. D., Abbeel P., and Russell S. : Cooperative Inverse Reinforcement Learning, in *Advances in neural information processing systems*, pp. 3909–3917 (2016)

[Malik 18] Malik D., Palaniappan M., Fisac J. F., Hadfield-Menell D., Russell S., and Dragan A. D. : An Efficient, Generalized Bellman Update For Cooperative Inverse Reinforcement Learning, in *Proceedings of the 35th International Conference on Machine Learning*, pp. 3394–3402 (2018)

[Nakahashi 18] Nakahashi R., and Yamada S.: Modeling Human Inference of Others' Intentions in Complex Situations with Plan Predictability Bias, in *Proceedings of the 40th Annual Conference of the Cognitive Science Society*, pp. 2147–2152 (2018)

[Nikolaidis 17] Nikolaidis S., Zhu Y. X., Hsu D., and Srinivasa S.: Human-Robot Mutual Adaptation in Shared Autonomy, in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 294–302 (2017)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

ナッジ効果を用いた運転支援システムの提示タイミングの決定

To determine the display timing of driving assistant using Nudge

榎波 晃一 *1 Kouichi Enami　今井 倫太 *1 Michita Imai　奥岡 耕平 *1 Kohei Okuoka　秋田 祥平 *1 Shohei Akita

*1慶應義塾大学 理工学部 情報工学科
Keio University, Faculty of Science and Technology, Department of Information and Computer Science

In this aging society, to improve convenience of electric wheelchairs is an issue. Especially how to adapt automatic operating systems, which has developed in recent years, to electric wheelchairs is important. In this paper, we propose Mizusaki system, a screen agent that informs drivers gain changes in an electric wheelchair with automatic gain adjustment system. Mizusaki system, based on the Nudge effect, aims to improve operability and safety not with directly teaching how to drive, but with presenting surrounding environment and internal information. Since Mizusaki system uses a screen effect that is easily recognized even in peripheral vision such as vection and color, it can attract the attention of the driver even while driving, and expect that it will not request gazing at unnecessary time. In designing Mizusaki system, in order to investigate the optimum presentation timing, we conducted an evaluation experiment of the system by changing the presentation timing. As a result of the experiment, we found that when presenting at an earlier timing, drivers give better impressions.

1. はじめに

電動車椅子の操作を支援する運転支援エージェントは, 電動車椅子の需要に伴い求められている. エージェントは運転者の運転と周囲環境に基づき, 運転者に様々な指示を与えることで操作性・安全性の向上を図ることができる. また, 衝突防止や障害物検知といった運転支援システムと組み合わせることで, 運転者とシステムの連携を高めるといったことも可能である.

運転支援エージェントの態度に関しては, 既に研究がなされている. Tanaka らは運転者の評価が低い運転指導員と高い指導員を比較し, 評価が低い指導員が運転者に運転方法を直接指示する発言を多用しているのに対し, 評価が高い指導員は注意喚起を多く行っていることを示した [田中 他 15]. エージェントにおいても, 運転者にそれとなく最適な行動を示唆する「ナッジ効果」[Sunstein 14] に基づいた設計を行うことで運転者の評価を高められるのではないかと思われる.

ナッジ効果の研究 [リチャード 他 09] では, 人の選択行動へ介入する様々なさりげないデザインが紹介されている. 保険の選択や投資の選択といった場面を例に, 人の認知バイアスの観点からナッジ効果の有効性を解説している.

一方で, 本稿が対象とする運転支援は契約の選択行動の場面と異なり時間的制約が強く影響する場面である. 運転操作では走行状態の予測の元で人が乗り物を操作しており, 実時間で予測と選択行動が行われている. 運転支援エージェントによる運転支援を考える際には, 人の予測の部分に介入しナッジを与えることが重要であると考える. 特に本稿では, 自律的に車椅子の操作ゲインを変更し個人適応する車椅子 [Furuya *et al.*] に運転支援エージェントを導入することを考えており, 車椅子の自律的なゲイン変化によって生じる恐れがある人の予測と車椅子の操作ゲインのズレの解消にナッジ効果を用いることができると考えている.

本稿では, 電動車椅子に画面内エージェントを搭載し, 地図情報から取得できた周辺状況と自動調節される速度情報を伝達する Mizusaki システムを提案する. Mizusaki システムでは, 自動速度調節という運転支援システムと組み合わせることを想定し, 速度調節がなされる度に速度が変化する旨と「なぜ速度を変化する必要があるのか」という理由をタブレット端末に表示されたエージェントが運転者に伝達する. Mizusaki システムは実験環境として汎用的な Android タブレットを使用しており, 特別な設備やメンテナンスを必要としない. また, 音声による通知ではなく画面の変化による通知を行うため, 騒音の出る環境や音が出せない環境においても問題なく扱うことが出来る. 動きや色彩の変化は周辺視野においても感知しやすいことから, エージェントは「ベクション」と「色彩」を利用して伝達を行い, 中心視野で見なくとも変化が読み取れるようになっている.

ナッジ効果を利用した運転支援エージェントの組み立てにおいて, 一つの重要な要素が「情報の通知タイミング」である. 運転支援エージェントからの通知タイミングを調整することで, 車椅子の操作ゲインに対する人の予測を修正するナッジ効果が得られると考えられる. 運転者はエージェントから状況の変化を事前に伝えられることで, 後に起きる変化に対してスムーズに対応することができる. 状況の変化の例としては, 車椅子の速度の変化, 道路状況の変化などが挙げられる. 通知タイミングが遅すぎる, あるいは早すぎることは運転者のスムーズな対応を阻害することになる. なぜなら, 遅すぎる通知は運転者にとって対応するための十分な時間を設けないことに繋がり, 早すぎる通知は通知から変化までの長時間運転者を緊張状態に晒すことになり不要な注意を惹いてしまうためである. 本稿では Mizusaki システムを作成するにあたり, 通知を遡る時間と運転者の印象について比較し, 最適な提示時間を調べる.

2. Mizusaki システム

Mizusaki システムは, 電動車椅子及び電動車椅子の前方に取り付けられたタブレットで動作する. また,Mizusaki システムは [Furuya *et al.*] に代表される操作ゲイン調整システムと連携して動作させることを前提としている. Mizusaki システ

連絡先: 榎波晃一, 慶應義塾大学 理工学部 情報工学科 今井研究室, 神奈川県横浜市港北区日吉 3-14-1 26-203, enami@ailab.ics.keio.ac.jp

ムの機能は, ゲイン調整システムによって変化するゲインを変化前に運転者に通知することである. 通知は「変化予定のゲインのレベル」と「ゲインが変化する理由」の 2 つから構成される. これにより, 運転者のゲイン調整システムへの適応コストを下げ, 信頼を向上することを目的とする. Mizusaki システムの画面構成例を図 1 に示す.

図 1: Mizusaki システムの画面構成例

画面を構成するオブジェクト, 画面効果を以下のように呼称する.

1. キャラクター
2. ベクション
3. 吹き出し
4. 背景色

ベクション及び背景色は「ゲインが変化する理由」の通知, キャラクター及び吹き出しは「変化予定のゲインのレベル」を通知するために使われる. それぞれの詳細を以下に記載する.

2.1 キャラクター

画面中央に存在する, 人形の 3D モデルを指す. モデルには Unity の Humanoid avatar が導入されており, アニメーションが適用できる他, 三次元座標を指定して注視させることが可能である. 表示されているモデルは平時は常に画面奥を向き, 画面奥に向かって走るアニメーションが適用されている. 後述する通知機能の発火時には運転者の方, すなわちキャラクターにとって後方を向くことで, 運転者に伝えたいことがあると知らせる.

2.2 ベクション

キャラクターの下方, 走っている床にあたる部分に表示される立方体のパーティクルの集合を指す. ベクションは射出速度を変えることでゲインの変化を表現する他, 背景色と連動して色を変える. ベクションの射出速度をゲインと比例させることで, 運転者が加速・減速を直感的に理解することを目指す.

2.3 吹き出し

キャラクターの上方に吹き出しを表示する. 常時表示されているわけではなく, 通知機能の発火時のみ表示する. 吹き出しには任意の画像が描画できるが, 本稿においては社会的コンテキストが浸透しており人間工学に基づいてデザインされているとされる日本国の交通標識を採用する.

2.4 背景色

画面上方の色を変化させることで, ゲインの変化を通知する. 色は青･緑･赤の 3 色を用いているが, これは [Sakurai *et al.* 02] によって得られた結果を参考に, 判別しやすい色を選択した.

3. 評価実験

本稿では Mizusaki システムの通知機能における最適な提示タイミングを得るために実験を行った. Mizusaki システムを提示タイミングを変えて実装した車椅子を参加者に乗り比べてもらい, 最適な提示タイミングを調べる.

3.1 実験環境

本稿で提案する Mizusaki システムは, ゲイン自動調整機能付き電動車椅子 (ゲイン調整車椅子) に搭載することを前提としている. そのため, 本章ではゲイン調整車椅子に関しても共に述べる.

ゲイン調整車椅子は, 周辺環境, または運転者の運転特性, あるいは双方をセンシングし, 運転者にとって安全で運転しやすいゲインに, 運転者の指示なしに調整する機能を付与した電動車椅子を指す. ゲイン調整車椅子においては, 操作は運転者の持つジョイスティックにて行われる. ゲインとは, ジョイスティックの入力値に対してどれくらいモーターを動かすかの係数である. ジョイスティックを同じだけ傾けたときの電動車椅子の速さは, ゲインに比例する. [Furuya *et al.*] ではレーザーレンジファインダーとユーザーのジョイスティック操作の傾向からゲインを調整している一方で, 本稿では実験への影響を最小限に留めるべく, あらかじめマッピングしたゲインマップへの参照を行う方式としている. 図 2 に本実験で使用したゲインマップを示す. 赤で示された範囲がゲイン 0.7, 青で示された範囲がゲイン 1.5, どちらでもない範囲がゲイン 1.0 である.

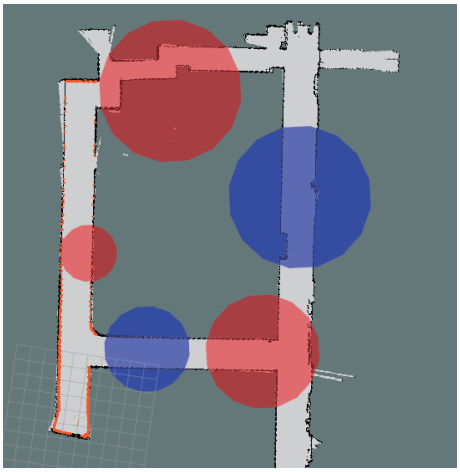


図 2: ゲインマップ

3.2 実験条件

本実験は 20〜24 歳の男性参加者 10 名, 女性参加者 2 名に,Mizusaki システムを搭載した電動車椅子で室内コースを

走ってもらうというものである．コース上にはT字路, ジグザグ道, 障害物といったゲインを上下させる要因が存在し, 計10回のゲインの変動と, それに伴うMizusakiシステムの通知が発生する．条件の順序は, 順序効果を考慮しラテン方格法に基づいて並べ替えている．4条件は,

- 提示時間 -3.0 秒
- 提示時間 -1.5 秒
- 提示時間 0.0 秒
- 提示時間 +1.5 秒

である．提示時間は「実際にゲインの変化が発生する時間との差」で表している．すなわち提示時間-3秒はゲインが変化する3秒前にMizusakiシステムでの通知を行うことを表す．実験参加者ごとに各条件の車椅子を操縦し, 計4回コースを走行してもらった．実験後表1に示すアンケートを取り, 参加者がシステムに対して抱いた印象を調べる．

3.3 結果

図3にアンケート結果のグラフを載せる．

「車椅子に意図を感じた」「車椅子に安心感を抱いた」を除くすべての質問において,-3.0秒が最もよい印象を与え,+1.5秒が最も悪い印象を与えている．「車椅子に安心感を抱いた」においても,+1.5秒が最も悪い印象を与えており,0.0秒,-3.0秒,-1.5秒が続いている．「ナビに情報を表示するタイミングは適切であった」という質問に対しては,+1.5秒と-3.0秒,+1.5秒と0.0秒の間には有意な差があった．「ナビは信頼できる」という質問に対しては,+1.5秒と-3.0秒の間で有意な差があった．

3.4 考察

全体を通して, 提示する時間がより早ければ早いほど, 運転者に良い印象を与えていることがわかった．特にゲインが変化した後に通知を行う+1.5秒や0.0秒に関しては,「すでに自分がゲインの変化を知ってしまっていたため, エージェントが意味をなしていない」とするコメントが複数寄せられた．ゲインの変化前に通知を行った方が良い印象を与えやすく, 実際にゲインが変化するより前に通知できるような, レンジが広く予測可能なセンシングを行うシステムが求められている．

今回の実験では,-3.0秒より前に通知を行う場合の実験は行わなかった．しかし, 今回の実験の結果,-3.0秒がおおむね最も優れている提示時間だと示されたため, 最適な提示時間を得るためには-3.0秒より前に関しても調べる必要がある．ただし, より前の提示時間を実現するためには, より長く精度の高い状態予測が必要である．

一部参加者からは「ベクションと背景色の提示はキャラクターと吹き出しの提示と趣旨が異なるため, 変えても良いのではないか」という意見が出された．今回の実験では, ベクション, 背景色, キャラクター, 吹き出しといった全ての画面効果を同時に提示した．しかし, ベクション・背景色と比較し, キャラクターと吹き出しは性質が大きく異なる．ベクションや背景色は, 既存研究により周辺視野においても変化が知覚しやすいことが分かっている [Sakurai *et al.* 02][福田 79]．対して, キャラクターの振る舞いや吹き出しといったものは周辺視野においても特別認識しやすいとは言えず, 詳細な情報を読み取るためにはある程度の注視を必要とする．ベクション・背景色が周辺視野において運転者の注意を惹く役割を果たし, キャラクター・吹き出しは情報を詳細に伝達する役割を果たしている．また, キャラクター・吹き出しが提示する情報である「ゲインを変化させる理由」というのは, ゲインが変化するということ, 及び変化するゲインの値と比べ即時で伝える必要性はあまりない．そのため, ベクション・背景色とキャラクター・吹き出しの提示タイミングは同期するのではなく, 独立して考慮する必要があると思われる．

4. おわりに

本稿では, ゲイン調整機能付き電動車椅子に搭載し, ゲインが変化する値と理由を運転者に知らせるMizusakiシステムを提案する．Mizusakiシステムの評価の一部として, システムの提示タイミングを変化させる実験を行った．結果として, 実際にゲインが変化する3秒前1.5秒後の範囲においては, 提示タイミングは早ければ早いほど運転者に良い印象を与える可能性が示唆された．Mizusakiは, 運転支援エージェントからの情報提示タイミングの形で, 車椅子の状態に対する人の予測を修正し, 人の運転行動を変化させる実時間のナッジ効果を与えることができた．

参考文献

[Furuya *et al.*] Seigo Furuya, and Michita Imai. 車椅子のための個人の操作特性と環境情報に基づいたコントロールゲイン調整システム．

[Sakurai *et al.* 02] Masato Sakurai, Takayuki Koseki, Hirofumi Hayashi, and Miyoshi Ayama. Color appearance in peripheral vision: Effects of test stimulus and surround luminance. *Journal of the Illuminating Engineering Institute of Japan*, pp. 9–18, 2002.

[Sunstein 14] Cass R. Sunstein. Nudging: A very short guide. *J. Consumer Pol'y*, p. 583, 2014.

[リチャード 他 09] リチャードセイラー, キャスサンスティーン．実践 行動経済学．日経BP社, 7 2009.

[田中 他 15] 田中貴紘, 米川隆, 吉原佑器, 竹内栄二郎, 山岸未沙子, 高橋一誠, 青木宏文, 二宮芳樹, 金森等．高齢ドライバ支援エージェントの提案－運転指導員による指導方法の分析－．日本知能情報ファジィ学会 ファジィ システム シンポジウム 講演論文集, Vol. 31, pp. 375–378, 2015.

[福田 79] 福田忠彦．運動知覚における中心視と周辺視の機能差．*The Institute of Image Information and Television Engineers*, pp. 479–484, 1979.

表 1: アンケートの項目一覧

Q1	車椅子の動きを快適に感じた
Q2	車椅子を便利に感じた
Q3	車椅子に安心感を抱いた
Q4	車椅子に意図を感じた
Q5	車椅子が自分に合わせていると感じた
Q6	この車椅子を今後も利用したい
Q7	ナビに情報を表示するタイミングは適切であった
Q8	ナビは信頼できる
Q9	ナビは親しみが持てる
Q10	ナビを今後も利用したい
Q11	ナビは役に立つ
Q12	ナビによって安全運転ができた
Q13	ナビは不快ではなかった
Q14	ナビは邪魔ではなかった
Q15	ナビの伝えたいことはすぐに理解できた

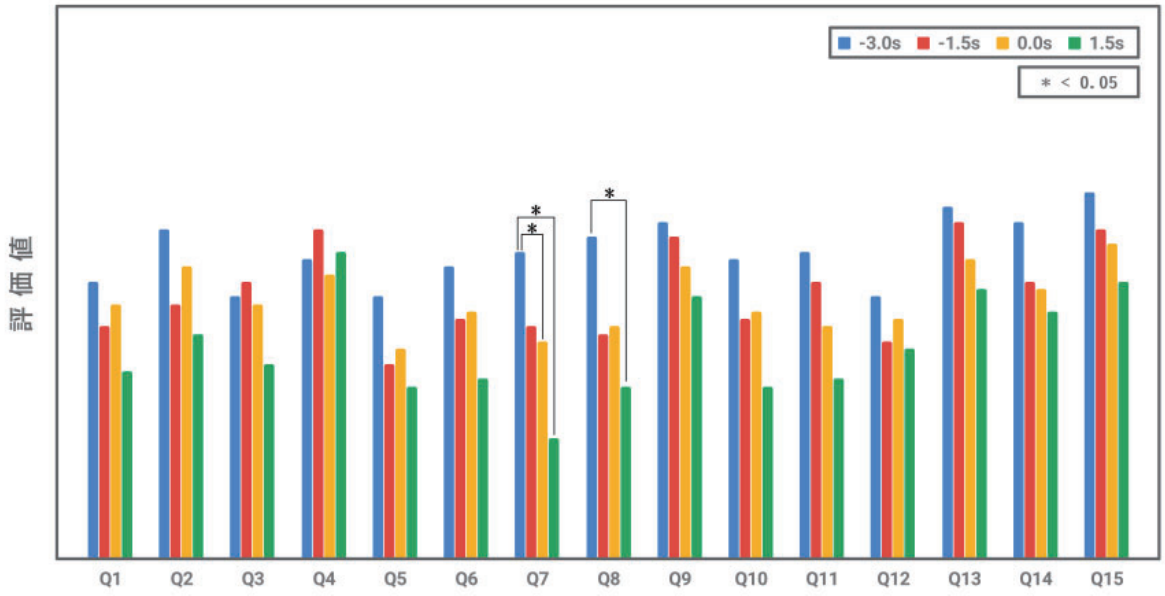


図 3: Mizusaki システム試乗後のアンケート結果

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

人間とエージェントとの協調作業における適応的な信頼較正手法の提案

Adaptive Trust Calibration for Human-AI Collaboration

岡村和勇 *1 山田誠二 *2*1
Kazuo Okamura Seiji Yamada

*1総合研究大学院大学
SOKENDAI (The Graduate University for Advanced Studies)

*2国立情報学研究所
National Institute of Informatics

Poor trust calibration in human-AI collaboration often degrades the total system performance in terms of safety and efficiency. Existing studies have primarily examined the importance of system transparency in maintaining proper trust calibration, with little emphasis on how to detect over-trust and under-trust nor how to recover from them. With the goal of addressing these research gaps, we propose a novel method of adaptive trust calibration, which consists of a framework for detecting the status of calibration and cognitive cues called "trust calibration cues". Our framework and four types of trust calibration cues were evaluated in an online experiment with a drone simulator. The result showed that presenting the simple cues at the time of over-trust could significantly promote the trust calibration.

1. 背景

自律的システムと人間の協調作業においては，システムに対して人間が持つ信頼が重要なパラメータとなることが知られている．人間が持つ信頼とシステムの実際の能力がバランスしていない状態，すなわち不信あるいは過信状態においては，効率の低下や危険な結果を招いてしまうことがある．これは例えばドライバーと自動運転システムの関係を考えれば容易に理解できる．

このバランスを取ることを信頼較正（trust calibration）と呼び，[Lee 04] は "the correspondence between a person's trust in an agent and the agent's capabilities"と定義している．適切な信頼較正がなされていれば，過信や不信状態に陥ることが無く，人間とエージェントの協調作業のパフォーマンスが安全かつ最大となりえる．先行研究 [Visser 14, Lyons 17] においては，適切な信頼較正の維持のためには，システムの状態を透過的に提示すること (system transparency) の重要性が示されている．例えば，自動運転システムの不確実さをメータパネル内に表示することで，ドライバーの信頼を適切に維持できることが示されている． [Helldin 13, Jung 15]

しかしながら，これらは過信や不信に陥らないようにするための方策についての研究であり，筆者の知る限り，不適切な信頼較正状態になっているかどうかの判定や，過信不信状態からの復帰についての研究はあまり行われていない．本研究では，この２つの観点から，適応的に信頼構成を促す新たな手法を提案し，ドローンシミュレータを用いたオンライン実験による評価を行った．

1.1 提案手法

本研究にて提案する，適応的な信頼較正手法は，不適切な信頼較正状態を判定するフレームワークと，人間に対して適正な信頼較正状態への復帰を促す認知的手がかり（信頼較正 CUE と呼ぶ）から構成される．

連絡先: 岡村和男, 総合研究大学院大学 情報学専攻,
東京都千代田区一ツ橋 2 丁目 1-2, ok@nii.ac.jp

1.2 信頼較正フレームワーク

人間と自律的エージェントの協調作業を想定する．ここで実行すべきタスクは，人間が手動で行ってもエージェントに実行させても良いが，その選択は人間が行うものとする．この協調作業において，まず以下のパラメータを確率として定義する．

- P_{auto}: エージェントによるタスク実行の成功確率（エージェントのタスク実行能力を表す）
- P_{trust}: 人間による P_{auto} の推定値（人間によるエージェントに対する信頼を表す）

P_{auto} はエージェントの内外状況（外部環境変化，故障など）により変化するものとする．もし信頼較正が適切に行われていれば，P_{trust} も P_{auto} の変化に従って変化し，同じ値になるが，$P_{trust} > P_{auto}$ のままであれば過信状態，$P_{trust} < P_{auto}$ のままであれば不信状態であると言える．これらは意味的には明瞭な定義であるが，実際には P_{trust} の計測が困難であり，このままでは状態の判定を行うことは難しい．そこで，本フレームワークでは，P_{self} を導入し，過信および不信を以下のように定義する．

- P_{self}: 人間によるタスク実行の成功確率（タスク実行に関する自信を表す）

過信 エージェントに対する信頼が，人間によるタスク実行の自信よりも大きい，かつ，人間によるタスク実行の自信のほうがエージェントのタスク実行能力よりも大きい時，過信状態にあると言う

$$(P_{trust} > P_{self}) \wedge (P_{self} > P_{auto}) \quad (1)$$

不信 エージェントに対する信頼が，人間によるタスク実行の自信よりも小さい，かつ，人間によるタスク実行の自信のほうがエージェントのタスク実行能力よりも小さい時，不信状態にあると言う

$$(P_{trust} < P_{self}) \wedge (P_{self} < P_{auto}) \quad (2)$$

式 (1) と (2) における左側の不等式は，人間の行為の観察により真偽を判定することができる．自動化エージェントにタス

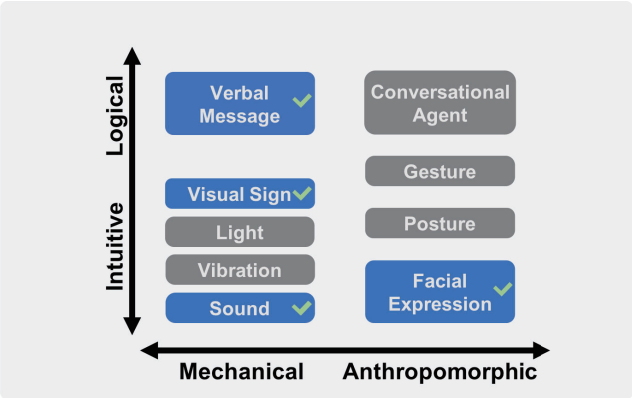


図 1: 信頼較正 CUE の分類

クを委任する行為が，人間の自信 (self-confidence) とエージェントに対する信頼との関係によって説明できることが [Vries 03, Gao 06, Yang 17] に示されている．人間がエージェントにタスクを委任した場合，$P_{trust} > P_{self}$ であるとみなすことができる．また，エージェントに委任しなかった場合は，$P_{trust} < P_{self}$ と推定される．このように式 (1) と (2) の左不等式の真偽は，人間の行為から推定可能である．従って，残りの右不等式の真偽を推定することが可能であれば，過信・不信を判定可能である．

人間の行為からの信頼推定に関しては， [Hergeth 16] が，視線の移動に注目した手法を提案している．この手法も適用範囲が広いと考えられるが，エージェントへの委任行為の観察に基づいた手法は実現の容易さにおいて優位性があると考えらえる．

1.3 信頼較正 CUE

過信や不信の状態とは，一種の認知的なバイアス状態であると考えられ，それゆえに抜け出すのは簡単ではない．そこで，過信や不信状態において，適正な信頼較正を促すためのきっかけとなる cue の提示を提案する．我々はこれを信頼較正 CUE と呼ぶ．

信頼に影響を与える情報提示の方法については多くの研究がなされている． [Visser 14] は，“trust cue’ を用いたシステム設計のガイドラインを提案している．また，[Komatsu 10] は，直感的な通知手法として，“artificial subtle expressions” (ASE) を提案している．[Cowell 05] は，会話エージェントのノンバーバル行為について論じ，顔の表情が信頼形成に対して有意であったことを示している．これら先行研究を参考に，信頼に影響を与える情報提示のカテゴリーを，機械的から擬人的，直感的から論理的の二軸にてまとめたものが図 1 である．本研究では，Visual，Sound，Verbal Message，Facial Expression の 4 つのカテゴリに着目し，信頼較正 CUE としての有効性を評価することとした．

1.4 適応的信頼較正

本研究が提案する適応的信頼較正は以下の手順にて実現される．

1. 人間の委任行為を観察する
2. 信頼較正フレームワークの式 (1) と (2) を評価する
3. 過信・不信が判定された場合，信頼較正 CUE を提示する
4. 最初から繰り返す

適切な信頼較正のためにシステムの透過性の維持の重要性を主張するこれまでの研究と異なり，本手法では，過信あるいは不信の判定時に信頼較正 CUE を提示することで適切な信頼較正を促すことを目的とする．

図 2: ドローンシミュレータ画面

図 3: 自動診断画面，選択画面，手動診断画面

2. 実験

提案手法を評価するため，Web ブラウザ上で動作するドローンシュミレータを開発し，オンライン実験を実施した．図 2 にその画面を示す．今回の実験では，過信状態を対象とした．

設定した協調タスクは，道路表面の異常（穴，くぼみなど）診断である．ドローンには，自動診断機能が搭載されており，決められたチェックポイントに到達すると，実験参加者に，図 3 中央に示したように自動診断（ドローンに任せる）か手動診断（自ら異常診断を行う）の選択が要求される．画面上の天候，視界，画面下部に常時表示されるドローンの自動診断機能の性能メーター（ P_{auto} を表す）が選択のための材料となる．自動を選ぶと，図 3 左に示されている自動診断結果が 3 秒間表示される．手動を選ぶと図 3 右に示されている画面が表示され，画面の写真を見て参加者が手動で診断する．道路上の異常は，濃い茶色の箇所として道路上に表示される．実験では手動での診断の正解率が 75%以上になるように問題の難易度を事前に調整した．

本実験では，4 種類 (Visual, Audio, Verbal, Anthro) の信頼較正 CUE の効果を評価した．Visual CUE は赤い逆三角形の警告表示，Audio CUE は [Komatsu 10] にて提案された（周波数が 400Hz から 250Hz へ減る）オーディオ ASE，Verbal

図 4: 実験で評価した 4 種類の信頼較正 CUE

CUE は「その選択は良くないかもしれません」という文字列を表示する吹き出し，Anthro CUE は正面に目を表示したドローンのアニメーションをそれぞれ使用した．

実験はマクロミル社の提供するクラウドソーシングサービスを使って実施され，194 名がオンラインで参加した (平均年齢 (標準偏差) = 44.35(14.10) 才, うち女性 98 名)．参加者は，1 種類の信頼較正 CUE がそれぞれ提示される 4 つの実験群と，信頼較正 CUE 提示無しの実験群の計 5 つのどれかにランダムに割り当てられた．

実験参加者は，まず最初に，実験目的と操作方法，および手動診断の正解率は平均 75%であること，自動診断の正解率は極めて優秀であるが天候に左右されること，を画面上で学んだ．次にビープ音の再生による音量調整と操作練習を行った後に，実験を開始した．実験においては，チェックポイント 6ヶ所を診断完了するまでは良好な天候とし，その間の P_{auto} を 100%とした．6ヶ所目の診断完了直後より，天候を悪化させた．具体的には雷雨の効果音再生と画面暗転を行い，P_{auto} を 50%に低下させた．実験を通じて P_{self} は P_{auto} に比べ，それほど大きくは低下しないという仮定をおいた．これにより $P_{self} - P_{auto}$ の符号は，好天候時と悪天候時で反転すると想定し実験システムにあらかじめ組み込んだ．

本実験では，信頼較正フレームワークの式 (1) の値が，最新を含む連続した 3ヶ所のチェックポイントでのウィンドウ中で 2 回以上真となった際に過信であると判断し，信頼較正 CUE が提示された．具体的には，自動選択のボタンが押された直後に，その行為にて過信判定された場合，ボタンのそばに 3 秒間 CUE が表示（Audio CUE の場合は再生）された．

実験コース上には，30ヶ所のチェックポイントが設けられたが，そのうち 22ヶ所をクリアするか制限時間 8.5 分を超えるかのどちらかで実験終了とした．好天候と悪天候の期間が少なくとも同等になるよう，実験参加者は少なくとも 15ヶ所以上のチェックポイントを診断完了することが期待された．

全てのキーボード入力およびマウスクリックがタイムスタンプと共に記録された．

3. 結果

参加者のうち，117 名が 15ヶ所以上のチェックポイントを診断完了した (平均年齢 (標準偏差) = 43.25(14.01) 才, 範囲 20 − 69 才, 女性 50 名)．その内訳は，NoCUE(CUE 提示無し) 実験群 28 名，Visual CUE 実験群 18 名，Audio CUE 実験群 22 名，Verbal CUE 実験群 30 名，Anthro CUE 群 19 名となった．

チェックポイントの 1ヶ所目から 6ヶ所目までを区間 1（好天候），7〜9ヶ所目までを区間 2（悪天候 CUE 提示なし）*1，10〜15ヶ所目までを区間 3（悪天候 CUE 提示あり）とし，各実験群の区間 1 と 3 における平均手動選択率を求めた．なお，手動診断の際の平均正解率は 88%となり，想定通り 75%を上回った．

測定されたデータに基づき，信頼較正 CUE を第 1 要因（CUE の種別により 5 水準），チェックポイント区間を第 2 要因（前述の区間 1 と区間 3 の 2 水準，対応あり），そして従属変数を平均手動選択率とする，2 要因混合計画の ANOVA による検定を行った．

第 1 要因（信頼較正 CUE の種類）に主効果 ($F(4, 112) = 3.59, p = 0.009, \eta_p^2 = 0.08$) が認められた．また，第 2 要因

*1 チェックポイントが 3 ヶ所のため，区間 2 の選択行為より前に信頼較正 CUE は提示されない

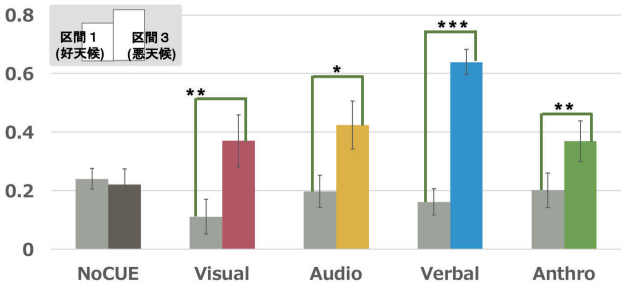


図 5: 信頼較正 CUE 毎の手動選択率（区間 1 と区間 3）

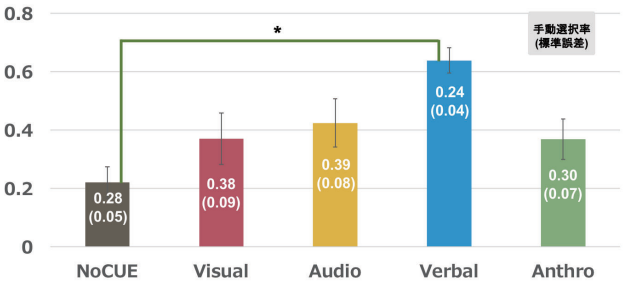


図 6: 信頼較正 CUE 毎の手動選択率（区間 3 での多重比較）

にも強い主効果 ($F(1, 112) = 53.00, p < 0.001, \eta_p^2 = 0.32$) が認められた．さらに，2 要因の間に交互作用が認められた ($F(4, 112) = 5.29, p = 0.0006, \eta_p^2 = 0.13$)．

4 種類全ての信頼較正 CUE において，それぞれ好天候であった区間 1 と悪天候の区間 3 の手動選択率の間に統計的に優位な差を認めた．Visual CUE で ($F(1, 17) = 9.20, p = 0.0075, \eta_p^2 = 0.31$)，Audio CUE で ($F(1, 21) = 5.54, p = 0.03, \eta_p^2 = 0.17$)，Verbal CUE で ($F(1, 29) = 67.79, p < 0.001, \eta_p^2 = 0.69$)，Anthro. CUE で ($F(1, 18) = 8.55, p = 0.009, \eta_p^2 = 0.28$) の結果を得た．一方で NoCUE 実験群では有意差は認められなかった (図 5)．下方検定として Holm 法で adjust した多重比較を区間 3 の信頼較正 CUE に対して行った結果 (図 6)，NoCUE と Verbal CUE との間に有意差 ($t(112) = 4.90, adjusted.p < 0.01, \alpha = 0.05$) が認められた．

4. 考察

信頼較正 CUE を提示しなかった NoCUE 実験群においては，好天候の区間と悪天候の区間にて手動選択率に有意差が認められなかった．今回の実験において，悪天候時の画面変化，雷雨効果音，そして自動診断性能メーターの指示値の低下は極めて明示的で認知しやすいものと考えられるが，NoCUE 実験群の参加者は好天時と同様にドローンへのタスク委任を継続した．本実験群の結果より，NoCUE 実験群の参加者は過信状態に入っており区間 3 までの間では不適切な信頼較正状態から自力で抜け出すことができなかったものと見做すことができる．

4 種類の信頼較正 CUE を提示した実験群においては，いずれの実験群においても区間 1 より区間 3 の手動選択率が有意に増加し，信頼較正 CUE の効果が認められた．すなわち信頼較正が促され，より適切な状態へ移行したと見做すことができる．

やや意外なことに，Verbal CUE が最も高い効果を示した．よりはっきりと目立つがシンプルな警告の意味に留まる他の CUE と比べると，Verbal CUE には，表示する警告文字列中に「選択」という単語が含まれているおり，このことが参加者

の次の委任行為（自動と手動の選択）に対して影響を与えた可能性が推測される．Audio CUE にも，より大きな効果が期待されていたが，Verbal CUE を超えることはなかった．ひとつには実験環境でのドローンの飛行音や悪天候時の雷雨の効果音など他の音源が干渉した可能性が考えられる．また，Anthro CUE も，実験前の期待に反して，それほど大きな効果は認められなかった．このことは単に画面上での認識度合いが大きいだけでは信頼較正 CUE として不十分であることを示している．

以上より，常時表示される手段でシステムの透明性を維持するだけでは，例えそこに必要と考えられる情報が提示されていても，過信状態から抜け出すことが困難である場合が存在することと，過信状態において信頼較正 CUE を提示することで行動を変化させることができることが示された．また，次の行為を連想する情報を含む信頼較正 CUE が，そうでない CUE より有効に働く場合があることを観測することができた．

5. まとめ

本研究では，信頼較正フレームワークと信頼較正 CUE を用いて適応的な信頼較正を促す手法を提案し，ドローンシミュレータを用いた，過信状態の検出と信頼較正の促進実験を実施し，提案手法の有効性を示すことが出来た．

今回の実験では，先行研究にても多く用いられている発見タスク（あるいは偵察タスクとも呼ばれる）での過信状態を最もシンプルなシナリオにて評価を行ったが，不信状態を含むより多様なシナリオでの検証が必要である．また，異なるタスク要因（負荷，結果のリスクなど），異なるユーザ要因（事前知識やシステムを信頼しやすい傾向の有無など）での検証も必須である．また，フレームワークにおける P_{self} の扱いの検討およびさらなる信頼較正 CUE の評価なども今後の課題である．

6. 謝辞

本研究は JSPS 科研費 JP26118005 の助成を受けた．

参考文献

[Cowell 05] Cowell, A. J. and Stanney, K. M.: Manipulation of non-verbal interaction style and demographic embodiment to increase anthropomorphic computer character credibility, *International Journal of Human-Computer Studies*, Vol. 62, No. 2, pp. 281–306 (2005)

[Gao 06] Gao, J. and Lee, J. D.: Extending the Decision Field Theory to Model Operators' Reliance on Automation in Supervisory Control Situations, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, Vol. 36, No. 5, pp. 943–959 (2006)

[Helldin 13] Helldin, T., Falkman, G., Riveiro, M., and Davidsson, S.: Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving, *Proceedings of the International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 210–217 (2013)

[Hergeth 16] Hergeth, S., Lorenz, L., Vilimek, R., and Krems, J. F.: Keep your scanners peeled: Gaze behavior as a measure of automation trust during highly automated driving, *Human Factors*, Vol. 58, No. 3, pp. 509–519 (2016)

[Jung 15] Jung, M. F., Sirkin, D., Gür, T. M., and Steinert, M.: Displayed uncertainty improves driving experience and behavior: The case of range anxiety in an electric car, *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems*, pp. 2201–2210 (2015)

[Komatsu 10] Komatsu, T., Yamada, S., Kobayashi, K., Funakoshi, K., and Nakano, M.: Artificial Subtle Expressions: Intuitive Notification Methodology of Artifacts, *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems*, pp. 1941–1944 (2010)

[Lee 04] Lee, J. D. and See, K. A.: Trust in automation: Designing for appropriate reliance, *Human Factors*, Vol. 46, No. 1, pp. 50–80 (2004)

[Lyons 17] Lyons, J. B., Sadler, G. G., Koltai, K., Battiste, H., Ho, N. T., Hoffmann, L. C., Smith, D., Johnson, W., and Shively, R.: *Shaping trust through transparent design: theoretical and experimental guidelines*, Vol. 499, Springer (2017)

[Visser 14] Visser, de E. J., Cohen, M., Freedy, A., and Parasuraman, R.: A design methodology for trust cue calibration in cognitive agents, *Proceedings of the International Conference on Virtual, Augmented and Mixed Reality*, pp. 251–262 (2014)

[Vries 03] Vries, de P., Midden, C., and Bouwhuis, D.: The effects of errors on system trust, self-confidence, and the allocation of control in route planning, *International Journal of Human-Computer Studies*, Vol. 58, No. 6, pp. 719–735 (2003)

[Yang 17] Yang, J. X., Unhelkar, V. V., Li, K., and Shah, J. A.: Evaluating Effects of User Experience and System Transparency on Trust in Automation, *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*, pp. 408–416 (2017)

6:20 PM - 6:40 PM (Tue. Jun 4, 2019 5:20 PM - 6:40 PM Room G)

[1G05-07-4] Discussion / Conclusion

Discussion / Conclusion

©The Japanese Society for Artificial Intelligence

Organized Session | Organized Session | [OS] OS-21

[1P3-OS-21] プロセス中心のシステムデザインとラーニングアナリティクス

緒方 広明（京都大学）、近藤 伸彦（首都大学東京）、瀬田 和久（大阪府立大学）、平嶋 宗（広島大学）、松居 辰則（早稲田大学）、堀口 知也（神戸大学）

Tue. Jun 4, 2019 3:20 PM - 5:00 PM Room P (Front-left room of 1F Exhibition hall)

[1P3-OS-21-01] Data-driven Infrastructure for Evidence-based Education and Learning

〇Hiroaki Ogata[1], Majumdar Rwitajit [1], Gökhan Akçpinar[1], Brendan Flanagan[1] （1. Kyoto University）

3:20 PM - 3:40 PM

[1P3-OS-21-02] An Analysis of Learning Processes using Scrapbox and Learning Outcomes

〇Nobuhiko Kondo[1], Toshiharu Hatanaka[2], Takeshi Matsuda[1] （1. Tokyo Metropolitan University, 2. Osaka University）

3:40 PM - 4:00 PM

[1P3-OS-21-03] Diagnosing and Promoting Self-Directed Investigative Learning on the Web

〇Yoshiki Sato[1], Akihiro Kashihara[1], Shinobu Hasegawa[2], Koichi Ota[3], Ryo Takaoka[4] （1. The University of Electro-Communications, 2. Japan Advanced Institute of Science and Technology, 3. Japan Institute of Lifelong Learning, 4. Yamaguchi University）

4:00 PM - 4:20 PM

[1P3-OS-21-04] A Lecture Substitution Robot for Diagnosing and Reconstructing Presentation Behavior

〇Akihiro Kashihara[1], Tatsuya Ishino[1], Mitsuhiro Goto[2] （1. The University of Electro-Communications, 2. NTT Service Evolution Laboratories）

4:20 PM - 4:40 PM

[1P06-09-5] Discussion / Conclusion

4:40 PM - 5:00 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Data-driven Infrastructure for Evidence-based Education and Learning

Hiroaki Ogata[*1] Rwitajit Majumdar[*1] Gökhan Akçpinar[*1,2] Brendan Flanagan[*1]

[*1] Kyoto University, Japan [*2] Hacettepe University, Turkey

Current eLearning infrastructures often include a Learning Management System (LMS), various ubiquitous and classroom learning tools, Learning Record Stores (LRS) and Learning Analytics Dashboards (LAD). Applying Learning Analytics (LA) methods to process data collected within such infrastructure can support various stakeholders. Learners can reflect on learning experiences, teachers can refine their instructional practices, and researchers can study the dynamics of the teaching-learning process with it. While LA platforms gather and analyze the data, there is a lack of specific design framework to capture the technology-enhanced teaching-learning practices. We proposed the Learning Evidence Analytics Framework (LEAF) and in this paper focus on the computational support for evidence extraction and analysis in a data-driven educational scenario.

1. Introduction

The concept of Evidence-Based Practices has its root in medicine and coined by doctors at McMaster University in Hamilton, Ontario in early 1990s (Kvernbekk T., 2017). According to Kvernbekk, EBP involves the use of the best available evidence to bring about desirable outcomes, or conversely, to prevent undesirable outcomes. Davies, P. (1999) reviews the concept of evidence-based practices in education. He proposes that evidence in the context of education needs to be established where its lacking and can be used in the following four ways:

1. Pose an answerable question about education;
2. Know where and how to find evidence systematically and comprehensively using the electronic (computer-based) and non-electronic (print) media;
3. Retrieve and read such evidence competently and undertake critical appraisal and analysis of that evidence according to agreed professional and scientific standards;
4. Organize and grade the power of this evidence and determine its relevance to their educational needs and environments.

While literature takes various theoretical perspective on Evidence-based education (Davies, P. 1999), Research-based education (Hargreaves, 1996), Literature-based education (Hammersley, 1997), Context-sensitive practice (Greenhalgh and Worrall, 1997) they mostly debate about rigorous studies to establish causalities similar to medical practices. What is missing is any research agenda of how technology can support the process and relevant discussions regarding issues in the current age of data-driven education. This position paper focuses on the notion of evidence-based education in the age of e-learning. Technology now supports logging of teaching-learning (TL) interactions and Learning Analytics has matured tremendously over the period to provide robust methods to analyze and predict learning behaviors and outcomes in different TL contexts. Hence there is relevance in rethinking about the question Davies (1999) asked regarding "What is evidence?" and how the four objectives can be supported by technology. This would push the boundaries of learning analytics and move towards an evidence-based education system that can assist the various stakeholders in the teaching-learning scenarios.

In the Learning Analytics community, SOLAR, the term evidence has recently come up in the context of a workshop in LAK 18 regarding evidence-based institutional LA policy (Tsai Y.S., Gašević D., Scheffel, M., 2018, sheilaproject.eu) and in LAK 17 by work presented by Ferguson & Clow (2017) where they introduce Learning Analytics Community Exchange (LACE) project's Evidence Hub. The Evidence Hub (http://evidence.laceproject.eu/) followed the evidence-based medicine paradigm to synthesize published LA literature and meta-analyze four propositions about learning analytics: whether they support learning, support teaching, are deployed widely, and are used ethically. But neither of the works look at technological affordances required to extract evidence of learning from logged data and make it available for the practitioners to adopt in their own context. We proposed a technological design framework, LEAF for evidence-based education and learning in this data driven age to find evidence of learning from the logged data of teaching-learning interactions. Figure 1 gives an overview (Ogata et.al. 2018; Majumdar R., 2018a).

2. Learning Evidence Analytics Framework

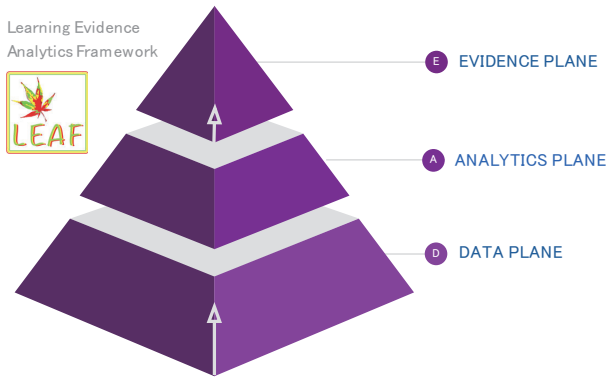


Figure 1. Extraction of evidence from teaching-learning data

Our Learning Evidence Analytics Framework (LEAF), is a technological design framework to support evidence-based education by integrating with existing learning analytics infrastructure. Figure 2 gives an overview of the components of

Contact: Hiroaki Ogata, Kyoto University, Kyoto, ogata.hiroaki.3e@kyoto-u.ac.jp

LEAF. We follow the DAPER (Data-Analysis-Planning-Execution monitoring-Reflection) model of data driven activities (Majumdar et.al 2018b) to guide the activity flow within the described framework.

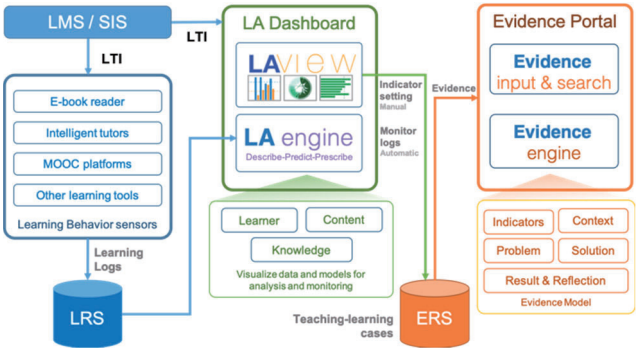


Figure 2. Components of Learning Evidence Analytics Framework (LEAF)

The data plane gathers data streams from various sources of teaching learning tools. For example, in our BookRoll system we can get learners reading activity data, their assignment score data from different learning management systems such as Moodle or Sakai, their health and physical activity data from GOAL system. Such a pool of learning behavior sensors, would also enable various connected services to create richer teaching-learning ecosystems. Data is stored in the Learning Record Store (LRS).

The power of such a data-driven ecosystem would be harnessed at the analytics plane by creating models and services for smartly supporting the stakeholders. For example, the data synchronized in our system can be used to create a context model for the learner. Based on that, analysis of the health and learning balance of the individual is possible. The objective of such measurement is focused to support learners to be aware of their status rather than to be directed by the technology. Similarly, in the case of the teachers, analytics can assist by suggesting intervention possibilities for automatically detected low engaged students. Learning Analytics Dashboards plays an important role there to present the analysis to the end users.

We seek to extract an evidence plane capturing effectiveness of the decisions taken by the users assisted by analytics models. Such a plane could be created by capturing meta-knowledge of results of applying models along with the broader perspective of having the context of the collected data. This evidence plane aims to keep the human agency in loop by capturing the user's reflection of the process. It captures the context, problems, indicators of the problem, the solution selected by the user, and collates its results. Based on the available infrastructure and information we aim to automatically extract some of these data. Else the user provides it through an evidence input portal.

3. Conclusion

Some of the research issues and challenges are enlisted for further investigation:

- How to extract evidences from data?
- How to design data format of evidences?
- How to evaluate evidences (rate them, evaluate similarities, meta-analysis, etc)?
- How to support search or context-aware recommendation of evidences?
- How to support teachers and students to apply evidences in their context?

While their exists endeavors to synthesize evidence from literature, LEAF aims to extend that and extract evidences from log data, considering contextual teaching-learning practices and harnessing the power of learning analytics methods and infrastructures. Our research agenda would give a fresh perspective on Davies' four use of evidence in education in this technology enhanced data-driven age.

Acknowledgement

This research was partly supported by JSPS Grant-in-Aid for Scientific Research (S) Grant Number 16H06304, Research Activity Start-up Grant Number 18H05746 and NEDO Special Innovation Program on AI and Big Data 18102059-0.

References

Davies, P. (1999). What is evidence‐based education? British journal of educational studies, 47(2), 108-121.

Ferguson, R., & Clow, D. (2017). Where is the evidence? : A call to action for learning analytics. In Proceedings of the 7th International Learning Analytics & Knowledge Conference (pp. 56-65). ACM.

Flanagan, B., & Ogata, H. (2017) Integration of Learning Analytics Research and Production Systems While Protecting Privacy. In Proceedings of the 25th ICCE 2017, New Zealand, Nov 2018

Tsai Y.S., Gašević D., Scheffel, M. (2018). Developing an evidence-based institutional learning analytics policy. In the 8th International Learning Analytics & Knowledge Conference. ACM.

Greenhalgh, T. and Worrall, J.G. (1997) From EBM to CSM: The evolution of context-sensitive medicine, Journal of Evaluation in Clinical Practice, 3, (2), 105–8.

Hargreaves, D.H. (1996) Teaching as a Research-Based Profession: Possibilities and Prospects. Cambridge Teacher Training Agency Annual Lecture.

Kvernbekk, T. (2017) Evidence-Based Educational Practice, in Oxford research encyclopedias,
DOI: 10.1093/acrefore/9780190264093.013.187

Majumdar R. (2018), Supporting Data-Driven Decision Making by Learners and Teachers, In Early Career Workshop Proc. of 26th ICCE, Manila, Philippines, Nov 2018.

Majumdar R., Yuan Yuan Yang, Huiyong Li, Akçapınar G., Flanagan B., Ogata H., GOAL: A System to Support Learner's Acquisition of Self Direction Skills, Proceedings of the 26th International Conference on Computers in Education (ICCE2018), pp. 406-415.

Ogata H., Majumdar R., Akçapınar G., Hasnine M.N., Flanagan B., Beyond Learning Analytics: Framework for Technology-Enhanced Evidence-Based Education and Learning, Proceedings of the 26th International Conference on Computers in Education (ICCE2018), pp. 486-489.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Scrapboxを用いたオンラインノートの学習記録と学習成果の分析

An Analysis of Learning Processes using Scrapbox and Learning Outcomes

近藤 伸彦 *1 Nobuhiko Kondo　畠中 利治 *2 Toshiharu Hatanaka　松田 岳士 *1 Takeshi Matsuda

*1首都大学東京 Tokyo Metropolitan University　*2大阪大学 Osaka University

In order to improve students' learning activities, it is important for teachers to understand how they learn both inside/outside classes and to support them in line with each student's actual learning process. In this research, we analyzed recorded data on an online note system, Scrapbox, to clarify students' learning processes and outcomes. The data were created by students as the performance of tasks on Scrapbox and collected as its log files. As a result, we could quantify and visualize some students' knowledge network building on online notes. In addition, it was implied that a deep learning with consciousness of the association among different knowledge had a possibility to enhance student ' s self-evaluation of his/her understanding.

1. はじめに

大学教育におけるアクティブラーニングの導入は，2008 年の「学士課程答申」，2012 年の「質的転換答申」あたりを契機として国内でも広く関心がもたれ，これまでに多くの理論的考察と実践が行われてきた [溝上 14]．これは，大学教育の質保証のひとつとして，知識獲得だけでなく技能や汎用的能力の獲得もまた大学での学習成果として求められるようになったことがその背景のひとつである．

質保証と密接に関連する機能として，学内外のデータの分析をもとに機関の意思決定を支援する IR（Institutional Research）がある．このうち，教育・学習に関する IR である「教学 IR」において近年とくに重要視されるものに「学習成果の可視化」がある．現時点では多くの場合，成績や就職状況等のデータ，学生調査による間接評価，あるいは標準テストによる直接評価などによって学習成果の可視化が行われるが，こうした時間粒度の荒いマクロなデータのみから，学生が実際に「いかに学んでいたか」という点について詳細に分析するのは難しい．しかしながら，学生の学びの具体的な改善のためには，そうした授業外も含めた学習のようすを把握し，個に応じた支援を行うことが重要であると考えられる．さらに，はじめに述べたように，質保証の観点から身につけるべきとされる論理的思考力のような汎用的能力をいかに身につけさせるかという点からも，いかに思考し，いかに学んでいるかという学習プロセスを適切に把握することは本来重要であるものと思われる．

一方で，ラーニングアナリティクスをはじめとする教育・学習データ分析がこの 10 年ほどの間に発展している．この分野では，LMS（Learning Management System）等の ICT システムのログやセンサーデータ等から学習状況を分析・可視化し，学習者の支援に役立てることが目指されている．教学 IR においても，ラーニングアナリティクス等の方法をうまく融合させることで，学習プロセスと学習成果を結びつけたきめ細やかな分析とそれに基づく学習支援が可能になると考えられる．

そこで本研究では，ICT を活用した授業を通して得られた比較的規模の大きなデータから学習プロセスをモデル化し，これとなんらかの学習成果との関連を総合的に分析することを考える．

連絡先: 近藤伸彦，首都大学東京，192-0397 東京都八王子市南大沢 1-1，kondo@tmu.ac.jp

本稿では，著者の近藤が担当する首都大学東京の全学共通科目におけるアクティブラーニング型授業を対象とする．本授業では，近年注目される情報整理ツールである Scrapbox[Nota 19] を活用して，学生相互に参照可能なオンラインノートを作成するという学習活動を核とした授業を行っている．このオンラインノート作成では，授業で扱うテーマに関連する情報を学生自ら調べ，Scrapbox にまとめるものとしている．Scrapbox の学習記録データを用いることで，Scrapbox の機能を用いたリンク構造により知識をネットワーク化するようすを分析することが可能であると考えられ，学習プロセスのひとつのモデル化ができることが期待される．

本稿では，Scrapbox およびこれを用いた本授業の概要を紹介したのち，Scrapbox によるオンラインノートの学習記録と学習成果をあわせて分析・可視化した結果を示す．

2. Scrapboxを用いたオンラインノート作成

2.1 Scrapbox

本研究が対象とする授業では，学習活動のための ICT ツールとして Nota 社の開発による Scrapbox[Nota 19] を用いている．Scrapbox は，端的に言えば「カード型のスムーズ Wiki」と表現できる情報整理・思考整理のためのクラウドツールである [倉下 18]．テキストベースの簡易な入力記法をもち，リンクとハッシュタグにより階層でなくフラットに情報を蓄積・構造化するため，容易にかつスケーラブルに知識をネットワーク化できる点が大きな特徴である．

Scrapbox は，図 1 のように「ページ」とよばれる単位で情報の入力を行い，このページ群を図 2 のように「プロジェクト」として管理する．図 1 に示されるように，文章中にリンクやハッシュタグを埋め込むと，ページ下部にそのリンク先やハッシュタグのページが表示され，さらに同一のリンクやハッシュタグをもつページ群も並べて表示される．つまり，文章中に適当なリンクやハッシュタグを埋め込むことで，関連するページとページをつないでいくことができ，階層構造をもたないフラットなネットワークとして情報整理がなされる．

Scrapbox の教育利用についてはまだ注目され始めたばかりであるが，たとえば [塩澤 18] のように，講義やゼミにおける活用事例は増加しているようであり，アクティブラーニングとしての高い効果も報告されつつある．

図 1: オンラインノートのページとリンクの例

図 2: Scrapbox によるオンラインノートの例

2.2 Scrapbox を用いたデータリテラシーの授業

本研究は，著者の近藤が担当する首都大学東京の全学共通科目「教養としてのデータサイエンス」を対象とする．本授業は，データリテラシーを現代的教養として身につけることをめざし，以下の 6 つのテーマを設定して，それぞれおよそ授業 2 回ずつを用いて講義・演習を行った．

- テーマ 1：グラフ等によるデータの可視化
- テーマ 2：データの分布と数値要約
- テーマ 3：データの相関
- テーマ 4：統計的検定
- テーマ 5：機械学習
- テーマ 6：進化計算

本授業はアクティブラーニング型授業を志向し，学生の能動的な学びを引き出すためのツールとして Scrapbox を活用した授業のデザインを試みており，以下のような特徴をもつ．

1. 調べ学習によるオンラインノートの作成
2. オンラインノートの相互参照による共同体としての知識構築
3. オンラインノートを活用した演習課題への取組み
4. 評価基準の明確化と学習プロセスの可視化によるゲーミフィケーション要素

本稿ではこのうち 1 に着目し，Scrapbox の学習記録データからオンラインノートの作成のようすを分析する．

2.3 本授業におけるオンラインノート作成の進め方

本授業では，先述の 6 つのテーマそれぞれについてまずヒントとしての概念マップ（テーマ 1 の例を図 3 に示す.）を配付し，このキーワードやキーワード間のつながりを参考に学生自ら書籍や Web で調べ学習をしつつ，学生ごとに用意した Scrapbox プロジェクトに適宜ページを作成して調べた内容をまとめるものとした．この学生ごとの Scrapbox プロジェクトをオンラインノートと呼んでいる．学習の導線として，概念マップにおけるキーワードの色や表示の濃淡で調べ学習の優先度を暗示している．

たとえば，図 1，図 2 のように,「平均値」「擬似相関」といったページを作成し，その内容を各ページに入力する．Scrapbox では，文字列を [ ] で囲むとリンクとなり，その文字列のページへのハイパーリンクが生成される．また#から始まる文字列はハッシュタグとなり，同じハッシュタグの記載されたページが関連ページとして下部に表示される．オンラインノート作成においては，作成したページごとに，関連するテーマ名のハッシュタグを「#テーマ 1」のように記載するものとした．このリンクとハッシュタグによってページ間が関係づけられ，全体がネットワーク構造をとるようになる．

リンクについては，Scrapbox の使用方法の導入を行った第 2 回授業において簡単に説明したほか，オンラインノートの相互閲覧という学習活動を各テーマの演習回において行うことで，他の学生のオンラインノート作成のようすから Scrapbox そのものの使い方を相互に学べるようにした.「このようにリンクを貼るとよい」といった指示は明示的には与えなかった．

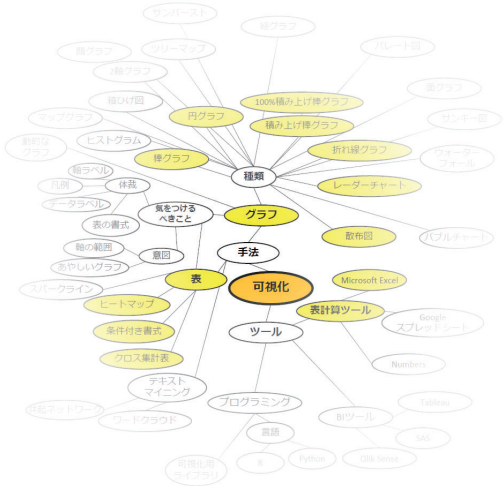


図 3: ヒントとしての概念マップの例（テーマ 1）

3. オンラインノートの学習記録に基づく学習のプロセスと成果の分析

本節では，前節にて述べた本授業について，2018 年度後期の授業を通して得られたデータの分析結果を報告する．本稿で用いたデータは，2018 年度後期の本授業履修者のうち，本研究におけるデータの使用について同意が得られた 46 名についての Scrapbox の学習記録データおよび授業後のアンケート回

答データである．学習記録データは，第 15 回授業から 1 週間後の時点のものを用いた．

3.1 オンラインノートのリンク構造からみる学習プロセス

ページ間のリンク構造をグラフとして可視化するツール [daiiz 19] を用いて，例として A〜C の 3 名のオンラインノートのリンク構造を可視化したものを図 4〜6 に示す．この図では，ひとつの円がひとつのページあるいはハッシュタグに対応し，短縮されたページタイトルが表示される．円の大きさは，対応するページから他のページへのリンクの数および他のページからリンクされた数の合計に比例する．なお，図中に示した「テーマ*」の文字は，本ツールで生成した図に著者が追加したものであり，各テーマに関連するページのクラスタのおおよその位置を示すものである．これらの図における各テーマ名の大きな円はハッシュタグのページである．

表 1 には，この A〜C の 3 名について，オンラインノートの作成状態を定量化したものと，授業後アンケートの回答による各テーマの理解度の自己評価をそれぞれ示す．

オンラインノート作成状態の指標は次の 4 つである．「文字数合計」は，すべてのページの入力文字数の合計である．ただし，URL が記載されている場合はその文字数を除いている．「リンク数＋被リンク数合計」は，他のページへリンクした数および他のページからリンクされた数をすべてのページについて合計したものであり，ネットワーク全体としてのページ間のつながりの豊かさを示すものである．「同一テーマへのリンク数」は同一のテーマのハッシュタグがついたページへリンクした数，「別テーマへのリンク数」は異なるテーマのハッシュタグがついたページへリンクした数を，それぞれすべてのページについて合計したものである．後者が大きいほど，異なるテーマであっても同一の概念が使われることが意識化できていることが示唆され，「一見異なる知と知」を結びつける学びを行っているようすをある程度反映すると思われる．以上はいずれもハッシュタグがついたページ，すなわち 6 つのテーマいずれかに関連する内容のページのみを対象としたものである．

「テーマ*自己評価」は，授業終了時のアンケートにおいて，各テーマの理解度を自己評価した値である．いずれも 5 段階で，「かなり理解している」は 5，「ある程度理解している」は 4，「どちらともいえない」は 3，「あまり理解していない」は 2，「ほとんど理解していない」は 1 にそれぞれ対応させて表記している．文字数，リンク数のようなオンラインノートの作成状態のようすと自己評価との関係については次節にて述べる．

学生 A・学生 B は比較的ページ間のリンクが豊かである例である．学生 A は同一テーマ間のリンクが多く，図 4 を見ても各テーマのクラスタの境界が比較的明瞭である．一方学生 B は別テーマ間のリンクが多く，図 5 でもクラスタがきれいに分かれずに別テーマをつなげてネットワークを構築しているようすが窺える．

逆に学生 C はページ間リンクを全く作成していない例である．また学生 C の図においてテーマ 1・2・3・4・6 からつながっているページは，ハッシュタグがついていない単なるハブページであるので，これを除くと各テーマのクラスタは完全に独立している．クラスタの独立は，あるページからリンクをたどって別テーマのページへ移動することができないことを意味しており，知識のネットワークが分断されているといえる．

このように，Scrapbox の学習記録データを定量化・可視化することで，知識獲得における概念のネットワーク化のようすを知ることができる．

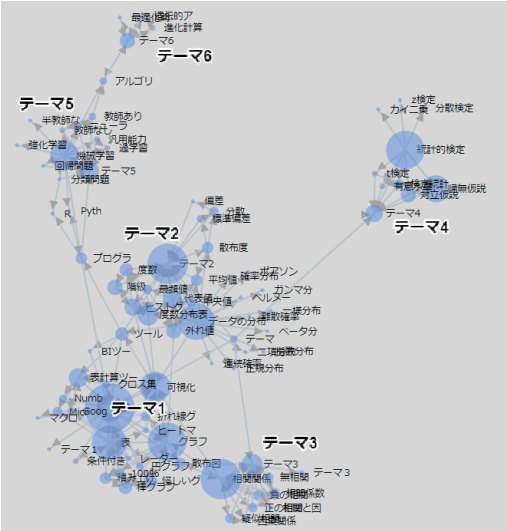


図 4: 学生 A のオンラインノートのリンク構造

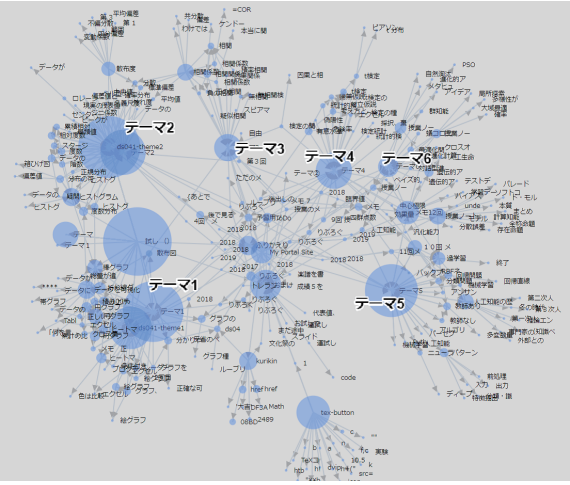


図 5: 学生 B のオンラインノートのリンク構造

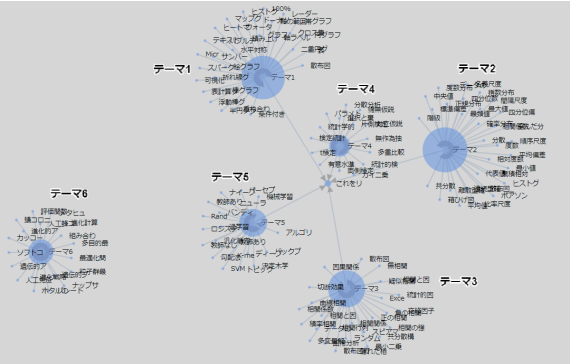


図 6: 学生 C のオンラインノートのリンク構造

表 1: オンラインノート作成状態の指標値と各テーマ理解度の自己評価（自己評価は 5 段階で 5 が最高）

学生	文字数合計	リンク数＋被リンク数合計	同一テーマへのリンク数	別テーマへのリンク数	テーマ 1 自己評価	テーマ 2 自己評価	テーマ 3 自己評価	テーマ 4 自己評価	テーマ 5 自己評価	テーマ 6 自己評価
A	17,260	316	146	18	5	5	5	5	5	5
B	35,657	255	105	36	5	4	4	2	5	4
C	18,698	0	0	0	4	4	4	4	4	4

表 2: オンラインノート作成状態指標値と自己評価値の相関係数（テーマ別）

	テーマ 1	テーマ 2	テーマ 3	テーマ 4	テーマ 5	テーマ 6	関連ページ計
文字数合計との相関係数	0.252	0.248	0.230	0.041	-0.012	0.063	0.132
リンク数＋被リンク数合計との相関係数	0.357	0.299	0.345	0.068	0.429	0.138	0.430

3.2 オンラインノートの作成スタイルと自己評価の関係

3.1 で示した学習プロセスのようすは，学生ごとのオンラインノートの作成スタイルともいえる．ここでは，このスタイルと授業終了時の自己評価との関係を分析する．3.1 で示した「文字数合計」および「リンク数＋被リンク数合計」をテーマ別に算出し，これらと自己評価値とのピアソンの積率相関係数を求めた．その結果を表 2 に示す.「関連ページ計」は，テーマ 1〜6 いずれかのハッシュタグがついたページについてのそれぞれの合計値と自己評価値との相関係数である．また,「関連ページ計」についての散布図を図 7 に示す.

学習内容の理解のしやすさはテーマごとに異なっていたと考えられ，とくにテーマ 4，6 は比較的難易度が高く自己評価が全体的に低めの値となっており，相関も比較的弱くなっているが，全体的に,「文字数」と自己評価の相関よりも,「リンク数＋被リンク数」と自己評価の相関のほうが強い傾向にある．本授業では，文字数を可視化して学生にフィードバックしこれを成績評価の一部としているため，文字数に対してはインセンティブが働いているが，リンクについては評価や可視化を行っていないため，リンクを設定する行為はあくまでも自発的なものである．結果として，単純に文字数（学習量）のみを意識してページ間（概念間）の関連性を考えず個別に調べ学習をするスタイルよりも，リンクを貼りながらページ間（概念間）の有機的なつながりを意識して調べ学習をするスタイルのほうが「理解できた感覚」が生まれやすかった可能性がある．

このように，学習プロセスのデータに基づく知識構築のスタイルと学習成果とを統合的にモデル化することができれば，これを学習成果につながるような学習プロセスへの適切な支援に応用できることが示唆される．

4. おわりに

本研究では，Scrapbox を用いてオンラインノートを作成する学習活動を核とした授業を対象に，オンラインノートの学習記録データに基づいて，学習プロセスと学習成果を関連付けた分析を行った．Scrapbox の機能とログデータを用いることで，オンラインノート上に知識のネットワークを構築する際のスタイルを定量化・可視化することができ，また異なる知の関連付けを意識した深い学びを行うほど理解の自己評価が高くなる傾向が見出された．

今後はさらに詳細な分析と学習プロセスのモデル化をめざし，学習成果に結びつく個に応じた学習支援のあり方を模索したいと考えている．

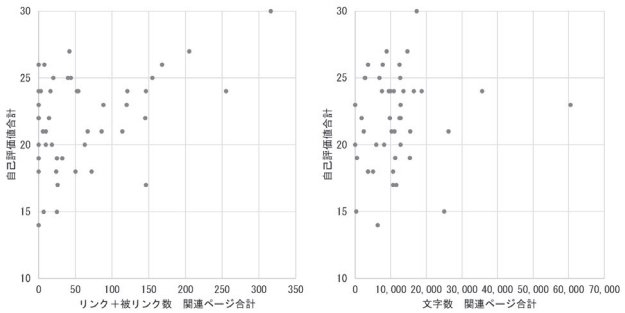


図 7: オンラインノート作成状態指標値と自己評価値の散布図

謝辞

首都大学東京 2018 年度後期開講科目「教養としてのデータサイエンス」履修者のみなさまに謝意を表します.

参考文献

[溝上 14] 溝上慎一: アクティブラーニングと教授学習パラダイムの転換, 東信堂（2014）.

[Nota 19] Nota Inc.: Scrapbox, https://scrapbox.io/product/ （参照日: 2019/2/8）.

[倉下 18] 倉下忠憲: Scrapbox 情報整理術, C&R 研究所 (2018).

[塩澤 18] 塩澤一洋: 教育における IT 利用に関する著作権法改正案と Scrapbox によるアクティブ・ラーニングの効用, 成蹊法学, 第 88 号, pp.149-188 (2018).

[daiiz 19] daiiz： Scrapbox Drinkup/ページの繋がりの可視化, https://scrapbox.io/scrapbox-drinkup/ページの繋がりの可視化_(daiiz)（参照日: 2019/2/8）.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

主体的な Web 調べ学習プロセスの診断に基づく支援手法

Diagnosing and Promoting Self-Directed Investigative Learning on the Web

佐藤 禎紀[*1] Yoshiki Sato　柏原 昭博[*1] Akihiro Kashihara　長谷川 忍[*2] Shinobu Hasegawa　太田 光一[*3] Koichi Ota　鷹岡 亮[*4] Ryo Takaoka

[*1] 電気通信大学大学院 The University of Electro-Communications

[*2] 北陸先端科学技術大学院大学 Japan Advanced Institute of Science and Technology

[*3] 日本生涯学習総合研究所 Japan Institute of Lifelong Learning

[*4] 山口大学 Yamaguchi University

In Web-based investigative learning, learners are expected to construct wider and deeper knowledge with navigating Web resources/pages. On the other hand, learners need to create a learning scenario with decomposing into related ones as sub-questions in order to elaborate the initial question. In our previous work, we have built a model of Web-based investigative learning and developed the system named iLSB, which scaffolds the investigation for learners. However, learners often investigate unrelated questions even if they use iLSB. This suggests the necessity of diagnosing learner-created scenario to present the results as feedback. On the other hand, it prevents learners from self-directed investigation. Toward this issue, we aim to diagnose learner-created scenario without prevention of self-directed investigation with LOD and to present the diagnosed results on the scenario. This paper also reports a case study. The results suggest that feedback of appropriateness of question decomposition can promote reflection and contribute to creating more appropriate scenario.

1. はじめに

近年, 21 世紀型スキルと呼ばれる情報活用能力が重要視されており[1], Web 調べ学習はその習得に適した学習である. Web 調べ学習では, 与えられた課題(初期課題)に対して, キーワード検索だけでなく, Web を探索して, 網羅的, 体系的な知識を構築する. 一方, 学習項目や順序(学習シナリオ)は与えられておらず, 学習者は知識構築と次に学ぶ課題の展開(課題展開)を並行して行うため, 認知的負荷が高い[2].

筆者らは, Web 調べ学習における学習プロセスをモデル化し, そのモデルに沿った学習を促すシステム iLSB(interactive Learning Scenario Builder)を開発した[3].

一方, 学習者は iLSB を用いても学習課題と関係ない課題を展開する場合があり, 学習シナリオの診断が必要である. 一般的には解となるシナリオ(解シナリオ)と学習者の学習シナリオを比較することで診断するが, Web 調べ学習のような主体性が重視される学習では, 解シナリオの定義が難しく, 仮に定義しても解シナリオに沿った学習を誘導し, 学習者の主体性を阻害する.

そこで, 本研究では Web 上の関連データをリンク付けした仕組みである LOD(Linked Open Data)を用いて, 学習者の主体性を阻害せずに課題展開を診断し, より妥当なシナリオ作成を促す手法を提案する. また, 本手法の評価実験を行った結果, 妥当なシナリオ作成を促す上で有効であることがわかった.

2. Web 調べ学習モデル

筆者らは, Web 調べ学習を以下の 3 フェイズからなるプロセスとしてとらえた Web 調べ学習モデルを提案した. 学習者は, この 3 フェイズを繰り返し, 最終的に学習シナリオを学習課題の木構造で作成することを想定している.

連絡先:佐藤 禎紀
電気通信大学 大学院 情報理工学研究科 情報学専攻
Email: yoshiki.sato@uec.ac.jp

(1) Web リソース探索フェイズ

学習課題を表すキーワード(課題キーワード)を用いて Web を探索し, 学習に用いる Web リソース(学習リソース)を収集する.

(2) Navigational Learning フェイズ

(1)で収集した学習リソースをページナビゲーションしながら, 学んだ項目を関連付けし, 知識構築する.

(3) 課題展開フェイズ

(2)で構築した知識から, 初期課題を理解する上でさらに学ぶべき項目を選択し, それを部分課題として展開する.

一方, 学習者は Web 調べ学習モデルに沿って学習を進めても, 必ずしも学習課題と関連する課題を展開するとは限らない. そこで本研究では, 学習の主体性を損なわずに学習者の課題展開を診断・促進する手法を提案する.

3. 課題展開の診断手法

3.1 LOD (Linked Open Data)

LOD とは, Web 上の関連データをリンク付けし, 公開している仕組みのことであり, リンク付けしたデータ群はネットワークとして表現できる. LOD の主な例として DBpedia Japanese や GeoNames などがあるが, 本研究では日本語版 Wikipedia を LOD として表現した DBpedia Japanese を用いる.

DBpedia Japanes のデータは RDF と呼ばれる, 主語, 述語, 目的語の形式で表現され, SPARQL と呼ばれるクエリ言語を用いて取得可能である. SPARQL を用いると, キーワード間の関係やキーワードに対する関連項目の取得が可能である.

3.2 診断の枠組み

診断手順を図1に示す. 診断機能は Firefox のアドオンとして開発した iLSB の機能として実装した.

学習者は iLSB を用いて Web 調べ学習を行い, 学習者が課題展開した際, iLSB の診断機能は DBpedia Japanese に課題キ

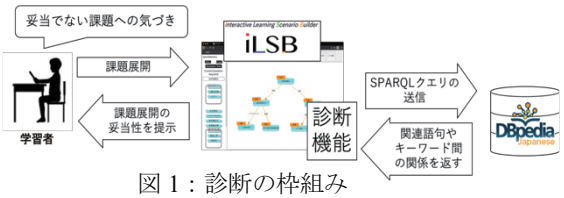


図 1：診断の枠組み

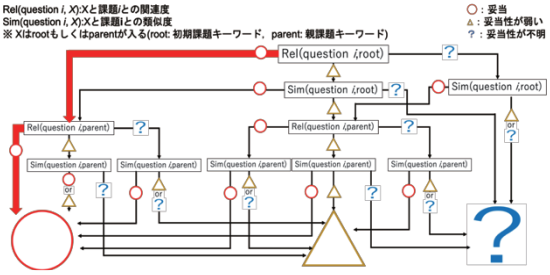


図 2：診断アルゴリズム

ーワード間の関係や関連項目を取得する SPARQL クエリを送信する．DBpedia Japanese から得られた結果から課題展開の妥当性を「妥当」，「妥当性が弱い」，「妥当性が不明」の3段階で診断し，学習者に提示する．提示された結果をもとに学習者は，課題展開を見直し，より妥当な学習シナリオ作成を行うことを想定している．

4. 課題展開妥当性の計算方法

課題展開の妥当性は，課題キーワード間の関係から求められる関連度と課題キーワードの関連語句比較から求められる類似度から決定する．これらを，展開した課題キーワードと初期課題キーワード間，展開した課題キーワードと親の課題(親課題)キーワード間において算出し，課題展開の妥当性を決定する．

4.1 課題キーワード間の関連度の算出

課題キーワード間の関連度は，DBpeida Japanese における 2 つの課題キーワード間の距離，経路数から求められ，最終的に予め設定した閾値をもとに「関連あり」，「関連度が弱い」，「関連度は不明」の 3 段階で求める．

4.2 課題キーワード間の類似度の算出

課題キーワード間の類似度は，DBpedia Japanese から得られた関連語句の比較によって求められる．具体的に，DBpedia Japanese から得られた関連語句を形態素解析し，各課題キーワードに対する関連単語集合を作成する．これらを Simpson 係数で集合の類似度を算出し，予め設定した閾値をもとに「類似」，「類似度が弱い」，「類似度が不明」の3段階で求める．

4.3 診断アルゴリズム

初期課題キーワードと展開した課題キーワード間と，親課題キーワードと展開した課題キーワード間の関連度，類似度の結果から課題展開の妥当性を，図2に示す診断アルゴリズムで「妥当」，「妥当性が弱い」，「妥当性が不明」の3段階で決定する．

5. 評価実験

提案手法が，学習者の課題展開の見直しを促し，学習者の妥当な学習シナリオ作成に有効か評価するため，理工系大学生大学院生の被験者 16 名に対して実験した．被験者に診断機能なし iLSB を用いて Web 調べ学習を行ってもらい，学習終了後，提案手法で計算された課題展開の妥当性を提示した．その結果をもとに診断機能あり iLSB を用いて学習シナリオの修正を

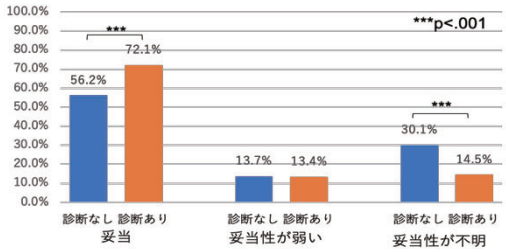


図 3：システムの提示結果

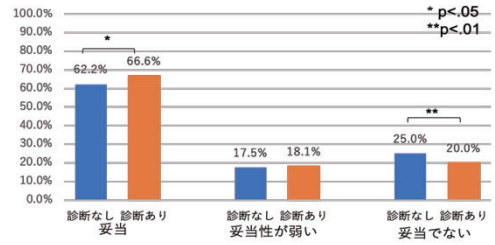


図 4：人手の評価結果

行ってもらった．診断機能なし iLSB で作成された学習シナリオ(診断なしシナリオ)と診断機能あり iLSB で作成された学習シナリオ(診断ありシナリオ)における，診断機能が提示した診断結果の割合を比較した．結果は図 3 のようになり，片側 t 検定を行なった結果，「妥当」と診断されたものは 0.1%水準($t(15)=-3.90$, $p<.001$)で有意に増加し，「妥当性が不明」と診断されたものも 0.1%水準($t(15)=4.52$, $p<.001$)で減少した．また，診断なしシナリオと診断ありシナリオを筆者ら3名が信頼できるリソースで課題展開の妥当性を「妥当」，「妥当性が弱い」，「妥当でない」の3段階で評価し，割合を比較した．結果は図4のようになり，片側 t 検定を行なった結果，「妥当」と評価された課題は 5%水準で有意($t(15)= -1.76$, $p<.05$)に増加し，だった．一方妥当でないと評価された課題は 1%水準で有意($t(15)= 2.79$, $p<.01$)に減少した．以上から提案手法は学習者の課題展開への見直しを促し，妥当なシナリオ作成に有効であることが示された．

6. 結論

Web 調べ学習のような主体的学習での診断が難しいという問題に対し，LOD を用いて課題展開を診断することで，主体性を損なわずに学習者の学習シナリオを診断する手法を提案した．評価実験の結果，提案手法は学習者の妥当な学習シナリオ作成に有効であることがわかった．

今後は，妥当でないと診断された課題に対してどこが妥当でないか提示する手法の提案などが挙げられる．

謝辞

本研究の一部は，JSPS 科研費基盤研究(B)(No.17H01992)の助成による．

参考文献

[1] Anainadou,K. and M. Claro: 21st Century Skills and Competeences for New Millennium Learners in OECD Countries: OECF Education Working Papers, No.41, OECD Publishing, 2009

[2] Zumbach, Joerg and Maryam: Cognitive load in hypermedia reading comprehension: Influence of text type and linearity, Computers in Human Behavior, 2008

[3] Kashihara, A. and Akiyama N.:Learning Scenario Creation for Promoting Investigative Learning on the Web, The Journal of Information and Systems in Education, Vol 15 Issue1, pp.62-72, 2017

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

プレゼンテーション動作を診断・再構成する講義代行ロボット

A Lecture Substitution Robot for Diagnosing and Reconstructing Presentation Behavior

石野 達也[*1] Tatsuya Ishino　後藤 充裕[*2] Mitsuhiro Goto　柏原 昭博[*1] Akihiro Kashihara

[*1] 電気通信大学 大学院情報理工学研究科 The University of Electro-Communications　[*2] NTT サービスエボリューション研究所 NTT Service Evolution Laboratories

In lecture with presentation slides such as e-Learning lecture, it is important for lecturers to control their non-verbal behavior involving gaze, gesture, and paralanguage. However, it is not so easy even for skilled lecturers to properly use non-verbal behavior in their lecture to facilitate learners' understanding. This paper proposes robot lecture, in which a robot substitutes for a human lecturer, and reconstructs his/her non-verbal behavior to enhance his/her lecture. Towards such reconstruction, we have designed a model of presentation behavior. This paper also reports a case study with the system, whose purpose was to ascertain the benefits of robot lecture. The results suggest that the robot lecture involving reconstruction promotes participants' understanding of the lecture slides. This also indicates the validity of the presentation behavior model. In addition, gaze, face direction, and pointing gesture for keeping and controlling learners' attention in the robot lectures could be more acceptable and understandable.

1. はじめに

近年，プレゼンテーション(以下，プレゼン)は，様々な場において盛んに行われており，発表する場や聴衆に応じて，伝達目的が異なる．本研究の対象である大学の講義や e-Learning 講義では，講師が学習者に対して，講義内容全体をわかりやすく伝達することを目的としている．通常，講義プレゼンでは，講義スライドと口頭説明によって学習者へ講義内容の伝達が行われる．その際，学習者の内容理解を促すためには，学習者の注意を講師が意図した場所へと集めることが重要であり，指差しや視線などの非言語動作の活用が求められる．

しかしながら，講義スライドを十分に準備した経験豊富な大学講師であっても，講義プレゼン実施時に，非言語動作を適切に用いることは容易ではない．例えば，内向的な講師は手元の PC を常に見ながら講義を行ったり，ジェスチャを使わずに講義を進める傾向がある．その結果，講義内容がうまく学習者に伝わらず，理解が不十分なまま講義が終わってしまう．

そこで，本研究では，e-Learning 講義において講師の代わりにロボットが適切な非言語動作を伴いながらプレゼンを実施する代講ロボットを開発した．本システムは，講師のプレゼンをもとに非言語動作の診断・再構成を行い，講義プレゼンをエンハンスすることが特徴である．

2. 講義プレゼンテーション

2.1 プレゼンテーションにおける非言語動作

一般に，講義プレゼンでは，イラスト，キーワードで表現された講義スライドの内容を指差しながら説明したり，スライドに陽に表現されていない事項について口頭で説明を行いながら講義を進めていく．Kamide らの研究では，ロボットプレゼンタによる視線の向きや指差しにより，聴衆の注意を意図した場所へ誘導できることが確認されている[Kamide 2014]．また，有馬は初任教師と熟練教師の授業中での視線の向きと思考について比較しており，熟練教師は意図的な視線行動を多く実施すると分析している [有馬 2014]．これらの研究から，聴衆の注意誘導にはプレゼン中の視線や指差しなどのジェスチャをやみくもに実施するのではなく，講師の意図を考慮して適切に実施することが重要であると言える．

連絡先：柏原昭博，電気通信大学，東京都調布市調布ヶ丘 1- 5-1，akihiro.kashihara@inf.uec.ac.jp

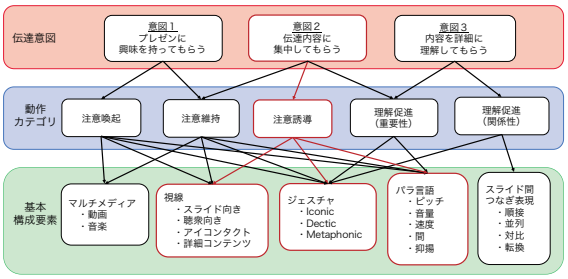


図 1 プレゼンテーション動作モデル

2.2 プレゼンテーション動作モデル

このような観点から，本研究では講師の伝達意図から適切な非言語動作を導き出すプレゼン動作モデルを提案している [Ishino 2018]．本モデルは，図1に示す通り，3層構造のモデルであり，講師の伝達意図と動作カテゴリ，基本構成要素の対応関係から，プレゼン時の非言語動作の組み合わせを決定する．本モデルを用いることで，例えば講師の意図が「集中してほしい」の場合，動作カテゴリとして「注意誘導」をする必要があり、動作カテゴリを表現するために，「視線やジェスチャ，パラ言語」の利用が必要ということを導きだすことが可能となる．

3. ロボットによる代講システム

図2に示す通り，本システムはロボットによる講義プレゼン代行を，3つのフェイズで実現する．フェイズ1では，講義スライドのスクリーンキャプチャやスライドの切り替えタイミング，口頭説明の音声の録音データを記録する．また，プレゼン中のジェスチャや視線などの講師の動作を Microsoft 社の Kinect によってモーションデータとして記録する．フェイズ2ではモーションデータからジェスチャを推定し，スライドデータからはスライド内に書かれているテキスト情報と赤字や太字などの修飾情報を取得してスライド内の重要箇所を推定する．また，オーディオデータからは，パラ言語とオーラルテキストデータを取得してオーラル

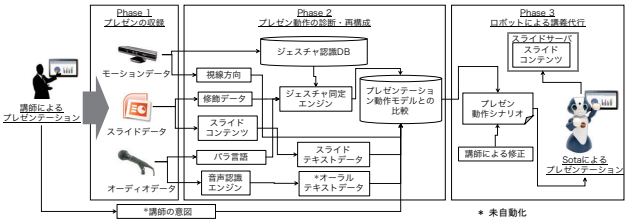


図 2 講義代行システムの全体構成図

での重要箇所を推定する．そして，視線，ジェスチャ，パラ言語のデータとモデルとを照合し，不適切な動作を診断して，適切な動作へ再構成する．フェイズ3では，前フェイズの再構成結果に基づき，NTT の R-env:連舞 [松元 2016]を用いて Vstone 社の Sota を動作させプレゼン代講を実施する．

4. ケーススタディ

4.1 ケーススタディ1：代講ロボットに対する印象

e-Learning 講義において，再構成されたロボットの動作が講師の動作よりも学習者の注意を制御し，理解を促す効果があるかを印象評価アンケートにより検証した．被験者は理工系大学生および大学院生31名であった．本ケーススタディでは，講師によるプレゼンを録画したビデオ(ビデオ条件)と講師のプレゼンに基づいて再構成された動作を行う代講ロボット(ロボット条件)の2つの条件を設定し，無作為に，19名の被験者を群1に，12名の被験者を群2に割り当てた．本ケーススタディの代講ロボットはジェスチャと視線を再構成対象としており，パラ言語は対象外であった．また，講義プレゼンは約5分間で，両条件のプレゼンでは同じコンテンツを使用した．各被験者はこれらの条件の下で講義を2回受講した後，自由記述ありの11問の2択アンケートに答えた．アンケートはわかりやすさ，集中度合い，視線，モチベーションの維持の4つの観点から作成し，ロボットと講師のどちらがよかったかを選択してもらった．

全被験者の回答結果を図3に示す．ロボット条件は11の質問のうち9つでビデオ条件よりも良くなる傾向があり，残りの2つの質問（Q4,Q6）では，ビデオ条件はロボット条件よりも良い傾向だった．片側 χ^2 検定の結果から，Q8～Q11 で有意差が認められた（$p<0.01$）．これらの結果から，講義ロボットによって再構成された学習者の注意を維持および制御するための視線および指差しが効果的であることが示唆された．一方で自由記述から，本ケーススタディで再構成していないパラ言語が講義内容の理解促進に非常に重要であることが明らかとなった．

4.2 ケーススタディ2：代講ロボットによる理解促進効果

ケーススタディ1の結果を踏まえて，代講ロボットによる理解促進効果を理解度テストによって検証した．被験者は大学生および大学院生22名であった．本ケーススタディでは，講師によるプレゼンを録画した動画（講師の動画），講師のプレゼンを再現する単純再現ロボット（再現ロボット），講師のプレゼンに基づいて再構成された動作を行う代講ロボット(再構成ロボット)の3つの条件を設定し，順序効果を考慮して3～4名の被験者を6群に分けて配置した．再構成ロボットと単純再現ロボットを比較することで，モデルの有効性を検証することができ，再構成ロボットと講師の動画を比較することで，見た目や次元が違う場合でも理解度が向上するかを確認することができる．本ケーススタディの再構成ロボットはジェスチャ，視線，パラ言語を再構成対象とした．また，講義プレゼンは約6分間で内容の異なるコンテンツを3つ用意した．被験者は各講義を受講した後，各講義で全6問の理解度テストに取り組んだ．

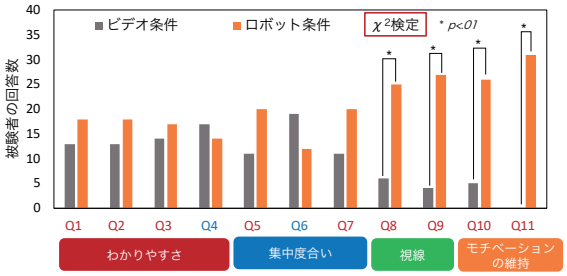


図 3 各質問項目の結果

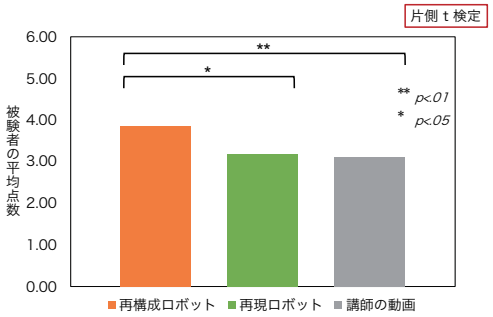


図 4 理解度テストの平均点数

プレゼンタごとの理解度テストの平均点数を図4に示す．理解度テストの結果から，再構成ロボットと再現ロボットで対応のある t 検定を行ったところ，5%水準で有意な差が認められた（片側検定，$t(21)= 1.925$ $p<.05$）．再構成ロボットの平均点数が再現ロボットよりも有意に高いため，プレゼンテーション動作モデルの有効性が確認できたと言える．また，再構成ロボットと講師の動画で対応のある t 検定を行ったところ，1%水準で有意な差が認められた（片側検定，$t(21)= 2.590$ $p<.01$）．この結果から，再構成された視線やジェスチャ，パラ言語が次元や見た目が違う場合でも学習者の講義内容の理解を促すことに貢献していることが示唆された．

5. まとめ

本稿では，講師のプレゼンを再構成する代講ロボットと再構成に必要となるプレゼン動作モデルを提案した．本システムを用いてケーススタディを実施し，本モデルの有効性が確認でき，学習者の理解促進に貢献した．今後は，インタラクティブな講義のために，学習者の状態に合わせて非言語動作を動的に変化させることを目指している．

謝辞

本研究の一部は，JSPS 科研費挑戦的研究（萌芽）(No.18K19836)の助成による．

参考文献

[Kamide 2014] Kamide, H., Kawabe, K., Shigemi, S., and Arai, T: Nonverbal behaviors toward an audience and a screen for a presentation by a humanoid robot, Artificial Intelligence Research, Vol.3, No.2, pp.57-66 (2014).

[有馬 2014] 有馬 道久：授業過程における教師の視線行動と反省的思考に関する研究：熟練教師と初任教師の比較を通して，広島大学大学院教育学研究科紀要，第一部，Vol.63，pp.9-17 (2014).

[Ishino 2018] Ishino, T., Goto, M., and Kashihara, A.: A Robot for Reconstructing Presentation Behavior in Lecture, 6th International Conference on Human-Agent Interaction, pp. 67-75 (2018).

[松元 2016] 松元 崇裕，松村 成宗，細淵 貴司，望月 崇由，吉川 博，山田 智広：「R-env:連舞」クラウド対応型インタラクション制御技術，人工知能学会全国大会論文集，JSAI2016 巻 (2016).

4:40 PM - 5:00 PM (Tue. Jun 4, 2019 3:20 PM - 5:00 PM Room P)

[1P06-09-5] Discussion / Conclusion

Discussion / Conclusion

©The Japanese Society for Artificial Intelligence

Organized Session | Organized Session | [OS] OS-1

[2D4-OS-1a] 計算社会科学(1)

鳥海 不二夫（東京大学）、笹原 和俊（名古屋大学）、吉田 光男（豊橋技術科学大学）、瀧川 裕貴（東北大学）、榊 剛史（ホットリンク）、高野 雅典（サイバーエージェント）

Wed. Jun 5, 2019 3:20 PM - 5:00 PM Room D (301B Medium meeting room)

[2D4-OS-1a-01] Proposal of the Method to Construct Landscape Map for Massive Information Propagation on Social Media

〇Takeshi Sakaki[1,2], Fujio Toriumi[2], Masanao Ochi[2] （1. Hottolink, Inc., 2. Univ. of Tokyo）

3:20 PM - 3:40 PM

[2D4-OS-1a-02] The emergence and evolution of communication systems in dilemma games with messaging

〇Naoki Inoue[1], Junya Morita[1] （1. Shizuoka University）

3:40 PM - 4:00 PM

[2D4-OS-1a-03] Analysis of viewers’ comments on Internet TV news

〇Tomosato Nishi[1], Yuki Ogawa[1], Hiromitsu Hattori[1], Fumiaki Taka[2], Masanori Takano[3], Soichiro Morishita[3] （1. Ritsumeikan University, 2. Kanagawa University, 3. CyberAgent, Inc.）

4:00 PM - 4:20 PM

[2D4-OS-1a-04] (Invited talk) Can business cards network approach unresolved problems in social science?

〇Naoki Maejima[1], Tomonori Manabe[1] （1. Sansan, Inc.）

4:20 PM - 5:00 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

ソーシャルメディア上の大規模情報拡散に関する俯瞰的可視化手法の提案

Proposal of the Method to Construct Landscape Map for Massive Information Propagation on Social Media

榊 剛史 *1,*2 Takeshi Sakaki　鳥海 不二夫 *2 Fujio Toriumi　大知 正直 *2 Masanao Ochi

*1株式会社ホットリンク Hottolink, Inc.　*2東京大学 The University of Tokyo

In recent years, a lot of people pay attention to negative aspects of information publishing by individuals, such as fake news, flames and echo-chamber phenomena. However, information propagation on social media has become a large-scale and complicated phenomenon, and it is difficult for human beings to intuitively understand the whole picture.

Therefore, we propose a method to visualize information propagation on social media in an intuitively understandable form in this paper. Specifically, in the proposed method a clustering method is applied recursively to a social network on Twitter and a hierarchical structure of user communities is constructed. We use the hierarchical structure to visualize information propagation. By using a hierarchical structure, we transform complex phenomena into simpler structures, thereby encouraging human intuitive understanding. We applied the proposed method to the actual case of large information propagation on Twitter and confirmed that reasonable visualization result can be obtained.

1. はじめに

ソーシャルメディアによる情報拡散が実社会に大きな影響を与えるようになった現在，その正の側面・負の側面の双方が顕在化している．特に近年は，炎上やデマ，フェイクニュース負の側面が注目されている．

それに伴い，ソーシャルメディア上での情報拡散を分析し，炎上やフェイクニュースを検知したり，そのメカニズムを明らかにしようとする研究が数多く行われている．それらのうちの代表的なアプローチは２つに分類できると考えられる．

１つ目はエージェントシミュレーションにより情報拡散の基礎的なメカニズムを明らかにしようとするアプローチである．２つ目は，実際のデータを元に情報拡散に関わるユーザの特徴を明らかにしたり，炎上やデマなどの現象を検知しようとするアプローチである．これらのアプローチは情報拡散現象に共通する抽象的な知見を得る意味では有用なアプローチである．しかし，個々の情報拡散の実際の事例について，どのように拡散したかを全体を俯瞰するためには適していない，近代科学において，ある現象を扱うための基礎的なアプローチは「データに基づいて客観的に観測すること」である．しかし，大規模・複雑化したソーシャルネットワーク上での情報拡散を適切に観測するアプローチは確立していない．

本研究では，Twiiter 上の情報拡散を人間が直観的に理解可能な形で可視化する手法を提案する．ソーシャルメディア上での情報拡散を観測可能にすることで，情報拡散を俯瞰的に理解し，分析や対策を容易に行えるようにすることが提案手法の目的である，

本研究では，図１のように Twitter 上の社会ネットワークを階層構造化することで，大規模かつ複雑なソーシャルメディア上の情報拡散をより単純に可視化することを目指す．具体的には，まず Twitter 上の情報拡散行動からユーザ同士の情報拡散ネットワークを構築する．構築したネットワークに，社会ネットワーク分析のクラスタリング手法である Louvain 法を再帰的に適用することで Twitter 上のコミュニティ構造を階層化する．各コミュニティに自然言語処理の特徴語算出手法である tf-idf によるラベリングを行う．さらに Circle − Packing という可視化手法を適用することで，直観的に理解可能なソーシャルメディア上の情報拡散の俯瞰図を作成する．本論文では，いくつの情報拡散事例を提案手法により可視化し，妥当な結果が得られていることを定性的に検証する．今後は提案手法による俯瞰図について定量的に妥当性を評価することを目指していきたい．

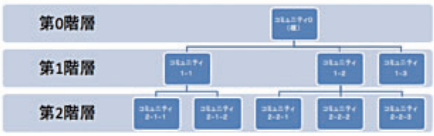


図 1: ソーシャルメディア上の社会ネットワークの階層構造化

2. 関連研究

SNS 上での情報拡散とは，実際には社会ネットワーク上での事象であるため，社会ネットワークの可視化手法が適用できると考えられる．社会ネットワークの代表的な可視化手法としては，クラスタリングを用いた可視化手法があげられる．柴田らの研究なので論文の引用ネットワークに Girvan-Newman 法を適用し，そこから得られるコミュニティ構造を用いて論文ネットワークを可視化している [Shibata 08].

しかし，Twitter 上での大規模な情報拡散においては下記のような理由から，既存手法を利用することが難しい．

1. 既存研究では，ノード数ががたかだか数万のネットワークを扱っていたため，全ノードを点として表示しても，人間が理解可能な画像サイズに収めることができる．一方，SNS 上での情報拡散に関わるユーザ（ノード）の数は数十万〜数百万規模であり，全ノードを点として表示することが難しい．

2. 論文ネットワークの可視化の場合，ジャンル毎に論文の

連絡先: 榊 剛史，株式会社ホットリンク, 東京都千代田区富士見１−３−１１ デュープレックスビズ５Ｆ

部分集合を抽出した上で可視化を行うため，可視化を行う度に結果が異なってしまう．そのため，論文の部分集合毎に比較することが難しい．一方，SNS 上での情報拡散の場合，各情報拡散の事例同士を比較することが難しい．そのため，比較可能な形で可視化することが求められる．

SNS 上の大規模拡散の可視化が，1. ノード数が大規模である，2. 個別の拡散事例を比較する必要がある，という特徴を持つことから，既存手法の適用が困難である，

本研究では，これらの問題点を解消するための新たな可視化手法を提案する．

3. 提案手法

本研究では，Twitter 上の大規模な情報拡散を直観的に理解可能な形で可視化する手法を提案する．

前説において．SNS 上の大規模拡散の可視化には，1. ノード数の大規模性，2. 個別の拡散事例を比較する必要性，という 2 つの問題点があることを述べた．これらの問題点を解決する手段について述べる．

1. については，階層構造を取り入れるで解決することを目指す．複雑系科学においては,「複雑なシステムは一般に階層をもった構造を持つ」とされている [ハーバート・A. 99]．本稿では，この発想を取り入れ，複雑な構造を階層構造により表現することで，直観的に理解可能な形で可視化することを目指す

2. については，Twitter 上の定常的な社会ネットワーク構造を用いることで解決することを目指す．Twitter 全体の定常的なユーザ同士のインタラクション情報から社会ネットワークを構築した後，Twitter 全体を俯瞰できるような俯瞰図の作成を行う．俯瞰図上で，様々な情報拡散の事例を可視化することで，個別の情報拡散の事例を比較することを目指す，ただし，Twitter 全体を対象とすることが困難であるため，本研究では日本語 Twitter ユーザを対象とする．

実際に構築した日本語 Twitter ユーザ全体の俯瞰図は図 2 である．以下では，図 2 を Twitter 俯瞰図と呼ぶこととする．

以下，本節では，Twitter 俯瞰図の見方および構築手順について説明する．

3.1 Twitter 俯瞰図の見方

Twitter 俯瞰図の見方について説明する．

Twitter 俯瞰図とは，図 1 のように階層構造化された Twitter ユーザネットワークのコミュニティ分布を可視化したものである．階層ごとに 1 枚の俯瞰図が作成される．図 2(a) は第 1 階層の俯瞰図であり，図 2(b)　は第 5 階層の俯瞰図である．用いるネットワークは個別の情報拡散事例のデータから構築されたものではなく，ある一定期間の全ユーザ（後述するが，本稿では日本語 Twitter ユーザ）の定常的な情報拡散データから構築されたものである．つまり，Twitter 俯瞰図は，その期間における定常的なユーザの情報拡散関係を表すデータと言える．

図 2(a) は，第 1 階層を表した俯瞰図である．図 2(a) 中の円が一つのコミュニティを表す．また円のサイズは，コミュニティに含まれるユーザ数により相対的に決定される．コミュニティに属するユーザ数が多いほどサイズが大きくなる．また丸ごとにラベルが付与されている．コミュニティ内のユーザプロフィール文から抽出した特徴語が用いられている（ただし，第 1 階層は得られた特徴語をの内容から，人手で再ラベリングを行った）．図 2 ではラベルの文字が読めないが，実際はブラウザで俯瞰図で表示し，マウスオーバーしたコミュニティのラベル名をポップアップで表示する仕様である．円の位置関係は適切に可視化できるように自動的にに決定されたものである．そのため，円同士の近さや隣接性とコミュニティ同士の関係性にはなにも相関性はない．つまり，Twitter 俯瞰図は，Twtiter ユーザの定常的な情報拡散に基づくコミュニティ分布のうち，規模と特徴ラベルを可視化したものである，1 つ下の層では，個別の円，つまり個別コミュニティの中のコミュニティ分布が再帰的に表現されている．

Twitter 俯瞰図に対し，個別の事例を用いて各円の塗色を決定する．各円の塗色は, 対応するコミュニティに属するユーザのうち，個別の拡散事例に関わったユーザの数で相対的に決定される，つまり，色が濃ければ濃いほど，関わったユーザが多いことがわかる．つまり，これが特定の円のみが濃い色で塗られていれば，ある情報拡散に関わったユーザ群はそのコミュニティに偏っていることが分かる．逆に，複数の円が薄い色で塗られていれば，そのユーザ群は複数のコミュニティに広く広がっていることとなる

Twitter 俯瞰図が階層化されているメリットは，全体的な俯瞰とより細部の俯瞰が両方実現できることである．全体を俯瞰したければ，第 1 階層の俯瞰図を見ればよい．より細部を見たければ第 2 階層，第 3 階層と下の層の俯瞰図を見ればよい．

次に具体的事例に基づいて Twitter 俯瞰図の見方を説明する．図 3 では，著者本人（tksakaki）のフォロワーリストに基づいて，Twitter 俯瞰図に色を塗っている．

まず，第 1 階層について図 3 と図 2(a) を比較すると,「ミドルエイジ」のクラスタに多くのフォロワーがおり，逆にそれ以外のクラスタには殆どフォロワーがいないことがわかった，つまり，自分は日本語 Twitter ユーザ上では、非常に偏った領域のユーザにフォローされていることがわかる．次に第 5 階について，図 3 と図 2(b) を比較すると，第 5 階層では，1 つのコミュニティの色が濃くなっている．これは「研究科の一覧,21 世紀 COE プログラム・プログラム一覧, 日本学術振興会賞」というラベルを持ったコミュニティである．つまり，自分は「ミドルエイジ」「のクラスタの中でも，研究者関連の語をプロフィールに持つユーザに多くフォローされていることが分かる．これより，自分は Twitter の中でも研究者コミュニティに属していると判断することができる．

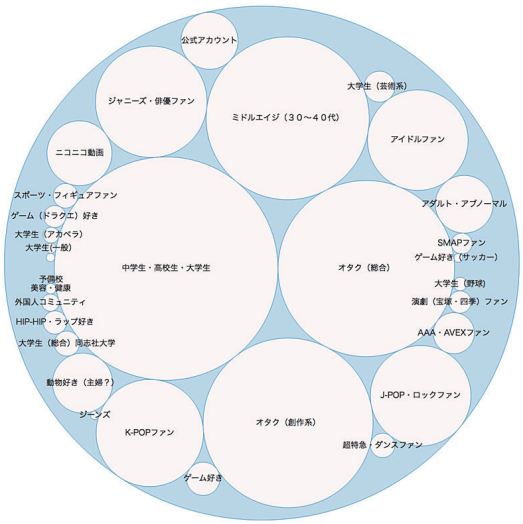
このように Twitter 俯瞰図を用いて個別の情報拡散事例に関わったユーザのコミュニティ分布を階層的に可視化することで，情報拡散の多様性と規模を直観的に理解できると考えられる．

3.2 俯瞰図構築の手順

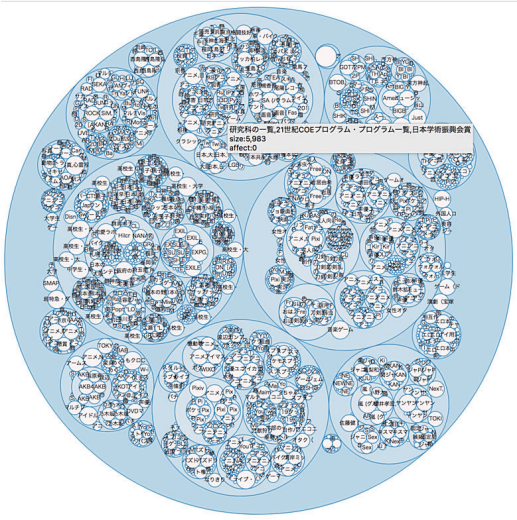
ここでは，Twitter 俯瞰図の構築手順について説明する．大まかな手順は以下の通りである．

1. 一定期間の Twitter ユーザ同士のインタラクション情報から，Twitter ユーザ間のネットワークを構築する
2. 構築したネットワークに対して，クラスタリング手法を適用し，ユーザをコミュニティに分割する
3. 得られたコミュニティについてラベルを付与する
4. コミュニティ規模が大きい場合，再帰的に 2，3 を適用する（再帰的なクラスタリングのことを以下、サブクラスタリングと呼ぶ）
5. 得られた階層的なコミュニティ構造について, 階層構造に適した可視化手法を適用する

以下，各ステップについて詳細に述べる．



(a) 深さ 5 の Twitter 俯瞰図



(b) 深さ 1 の Twitter 俯瞰図

図 2: Twitter 俯瞰図

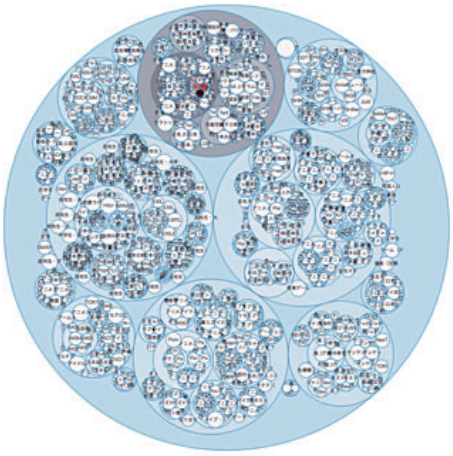


図 3: tksakaki のフォロワー群による俯瞰図

3.2.1 ネットワーク構築

Twitter ユーザの同士のインタラクションからネットワークを構築する，Twitter ユーザネットワークを構築するために手掛かりとして用いられるインタラクションはいくつかあるが，本稿では，リツイート情報を用いる，リツイートとは，1 クリックで，他のユーザの投稿をコピーし，元投稿のユーザを明示した上で，自身に向けてリンクを張っているユーザに対して再発信できる機能である．その手軽さ故に多くのユーザが使っている機能であり，またデータ収集も容易であることから，リツイート情報を用いることとした．

3.2.2 クラスタリング

本稿では，社会ネットワークのクラスタリングに用いられる Modularity-Optimization の手法を採用する，Modularity Q とはクラスタの結合度合いを表す指標であり，式 1 で表される，これが高ければ高いほどネットワーク全体のよくクラスタリングされていると考えられる．本稿では，既存研究で使われていた Louvain 法を用いる [Blondel 08]．式 1 において，$e_{i,j}$ は「コミュニティ$_{i,j}$ に属するノード同士が繋がるリンク数の全リンク数に占める割合」，a_i^2 は $e_{i,i}$ の期待値を表す．

$$Q = \sum_i \left(e_{i,i} - a_i^2\right) \tag{1}$$

クラスタリングにより得られるクラスタをコミュニティとして抽出する．

3.2.3 コミュニティへのラベル付与

本稿では，各コミュニティに関連するドキュメント集合をコミュニティ文書と見なし，各コミュニティ文書に特徴的な語を抽出し，それによりクラスタをラベリングする．

最初に各コミュニティに含まれる全ユーザのプロフィール文を結合して，コミュニティ文書とした，

次に同じ階層にいる全コミュニティのコミュニティ文書を全文書集合として，tf-idf 法を適用することで，各コミュニティ文書に特徴的なキーワードを抽出し，コミュニティのラベルとして用いた．tf-idf 法とは，自然言語処理でよく使われる特徴語抽出手法であり，全文書集合に対して，個別の文書に特徴的な語を抽出する機能である．

ただし，第 1 階層・第 2 階層のコミュニティはサイズが大きいものが多く，説明力が高い特徴語を付与することが難し買ったため，得られた他特徴語から人手で再度ラベルを付与した．

3.2.4 サブクラスタリング

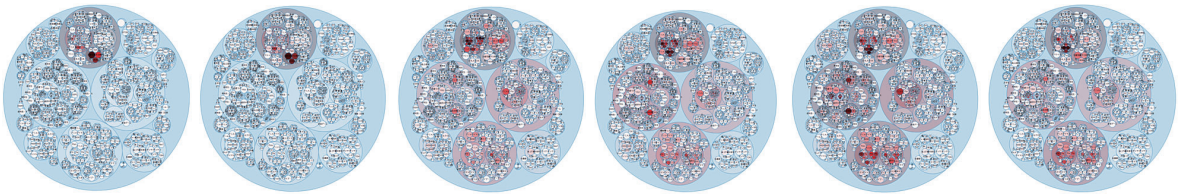
本稿では，サブクラスタリングについて説明する．2. において RT ネットワークからクラスタを抽出したが，可視化の観点からは，得られたコミュニティはサイズが大きく，これだけでは不適であった．

そこで，1 回目のクラスタリングで得られたコミュニティのうち，サイズが閾値 th_{clsize} を超えるものについては，再度クラスタリングを行う，そうして 2 回目のクラスタリングで得られたクラスタのうち，サイズが閾値 th_{clsize} を超えるものについては再度クラスタリング行う，という手順で，再帰的にクラスタリングを適用する．

これにより一定規模以下のコミュニティが大量に得られることとなる．本稿では最初のクラスタリング結果を第 1 階層，それ以降の深さ n のクラスタリング結果を第 n 階層と呼ぶこととする．ただし，深さは最大 th_{depth} になるようにし，それ以上はクラスタリングを適用しないこととした．本校では，$th_{depth} = 5$，$th_{clsize} = 1000$ とした．

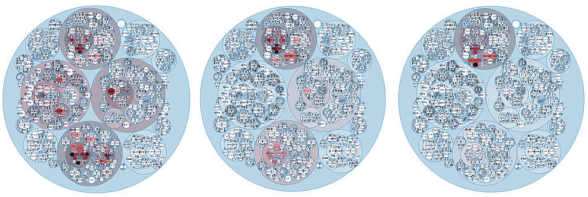
3.2.5 可視化

本稿では，階層構造の可視化手法の筆頭である Circle Packing を用いることとした．具体的には，階層構造を可視化するためのパッケージである．ここでは D3.js に準備されている



(a) 1 月 3 日　(b) 1 月 4 日　(c) 1 月 5 日　(d) 1 月 6 日　(e) 1 月 7 日　(f) 1 月 8 日以降

図 4: ZOZOTOWN 事例における俯瞰図の変化



(a) マクドナルド　(b) 御嶽山　(c) オリンピック

図 5: 大規模バーストにおける俯瞰図

Circle Packing のモジュールを用いて可視化を行った *1.

4. 事例紹介

今回は，我々の既存研究で用いたバースト現象のデータに加え，直近のバースト現象のデータを用いて，可視化を行う．

本稿では下記のようなデータを用いて，Twitter ユーザネットワークの構築・コミュニティの抽出およびふ Twitter 俯瞰図作成を行った．

投稿期間 2017 年 5 月 1 日〜7 月 31 日

利用データ 日本語 10%サンプリング

抽出投稿 公式 RT すべて

4.1 利用データ

我々の既存研究では，トピックエントロピーという値を用いて話題の多様性を評価している [鳥海 17]．本稿ではこれらのデータのうち，多様性が高いと評価された事例としてマクドナルドの異物混入事件（以下，マクドナルド事例）と御嶽山噴火前後のバースト（事例「以下，御嶽山事例），多様性が低いと評価された事例としてオリンピックエンブレム盗作騒動（以下，オリンピックエンブレム事例）の 3 つの事例を用いる．また直近のデータとして，ZOOTOWN 前澤社長の 1 億円プレゼントの事例（以下，ZOZOTOWN 事例）を用いる．特に ZOZOTOWN 事例では，キャンペーン前後の俯瞰図の時間変化を確認する．表 1 にデータセットの概要を示す．

表 1: データセットの概要

事例	件数	期間
マクドナルド	959,792	2014.12-2015.01
御嶽山	1,097,091	2014.09-2014.10
オリンピック	3,161,798	2015.07-2015.09
ZOZOTOWN	6,971,571	2019.01-2019.01

4.2 事例分析

図 5 に過去論文のデータを可視化した結果を示す．過去論文で多様性が高いと評価されたマクドナルド事例，御嶽山事例においては，確かに Twitter 俯瞰図上の幅広いユーザが反応していることがわかる．特に 1 層目に注目しても，「中学生・高校生・大学生」「ミドルエイジ」「「オタク（総合）「オタク（創作系）などである．一方，オリンピック事例においては，反応したユーザがミドルエイジに偏っている．第 2 階層以降に注目すると，ミドルエイジ（政治・社会・経済：左翼系）」など政治に強い興味を持つコミュニティに特に偏って拡散していることが分かる．

図 4 に ZOZOTOWN 事例の俯瞰図の時系列変化を示す，お年玉キャンペーンが行われる前の 1 月 3 日，4 日においては，「ミドルエイジ」のユーザが多く反応しているのがわかる，しかし，5 日にキャンペーンが始めると，一気に反応しているユーザが広がるのが分かる．5 日と 7 日を比較すると，キャンペーン開始時よりも終了時の方がさらに情報がより広く拡散していることが分かる．また，8 日以降も ZOZOTOWN が幅広く話題になっていることがわかる，このように ZOZOTOWN の事例では，キャンペーン当初から幅広いユーザが ZOZOTWON についてつぶやいており，またキャンペーン後にもその影響が残っていることが分かる．

5. おわりに

本研究では，SNS 上での大規模な情報拡散を直観的に理解するために，コミュニティ抽出と特徴語抽出の手法を組み合わせることで，Twitter 全体を俯瞰する俯瞰図を構築する手法を提案した．ただし，本論文では事例を通じた定性的な妥当性評価しか行っておらず，その評価も十分ではない．今後この手法について，定量的な評価を行い，手法の妥当性を検証していきたい，また既存研究で提案しているトピックエントロピやトピッククラスタリングの手法を組み合わせることで，より直観的にわかりやすい可視化を目指していきたい．

参考文献

[Blondel 08] Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E.: Fast unfolding of communities in large networks, *Journal of Statistical Mechanics: Theory and Experiment*, Vol. 2008, No. 10 (2008)

[Shibata 08] Shibata, N., Kajikawa, Y., Takeda, Y., and Matsushima, K.: Detecting emerging research fronts based on topological measures in citation networks of scientific publications, *Technovation*, Vol. 28, No. 11, pp. 758 – 775 (2008)

[ハーバート・A. 99] ハーバート・A. サイモン F システムの科学 〔第 3 版〕, パーソナルメディア (1999)

[鳥海 17] 鳥海不二夫, 榊剛史 F バースト現象におけるトピック分析, 情報処理学会論文誌, Vol. 58, No. 6, pp. 1287–1299 (2017)

*1 https://beta.observablehq.com/@mbostock/d3-circle-packing

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

メッセージ付きジレンマゲームにおけるコミュニケーションシステムの成立と変化

The emergence and evolution of communication systems in dilemma games with messaging

井上直紀 *1　森田純哉 *1
Naoki Inoue　Junya Morita

*1静岡大学情報学部
Faculty of Informatics, Shizuoka University

This research focuses on the emergence and evolution of human communication system, following Machiavellian intelligence hypothesis, which claims social intelligence resulted from the successful exploitation of rivals for personal benefits. Based on the framework of experimental semiotics, we design a dilemma game with messaging as a means to explore what kind of communication system is emerged and how it evolves in the laboratory experiment. In this game, graphical symbols are generated through interaction between pairs of players, and a dilemma of sharing and monopoly of rewards occurs. We consider that such a situation produces concealment of the betrayal intention utilizing the ambiguity of the symbols. The emergence and evolution of the novel communication system was examined with a small laboratory experiment. Our data suggested that this game causes ambiguity of communication protocol like polysemy.

1. 序論

人間は社会生活において相手の考えを推定しながら自己の利益を追求するようなコミュニケーションを行っている．コミュニケーションにおいて，他者がどのように考えているかをその人に成り代わって想定することは，高度な知能を必要とする．人間がこのような高い知能を得た理由として，マキャベリ的知性仮説が提案されている．この仮説は，人間などの類人猿の知能の進化の原因を，集団社会の中での他者推定のコミュニーションに求めるものである．そのようなコミュニケーションができる知能を持った集団の環境において，それぞれの個体の適応性が向上することで，現在の人間のような複雑なコミュニケーションを操る種が誕生したと考えられている [Byrne 95].

一方，近年は言語コミュニケーションの起源と進化過程に関する認知科学的研究が盛んに行われている [Scott-Phillips 10]. 例えば，Galantucci は通常のコミュニケーションが制限されたビデオゲーム環境下で協調的な課題を課すことで，二者間の相互作用によって，通常の言語のような図形と意味の対応づけのシステム（以下,「コミュニケーションプロトコル」という）が新規に成立することを確認した．[Galantucci 05, Galantucci 09]. Galantucci らの研究を受け，金野らは Galantucci らの実験課題をより制限することで分析を容易にする実験課題を設計した．2×2 マスのマップ上からペアがそれぞれ 1 回の移動によって，共通の場所で出会う協調ゲームを課題とした．初期状態において，ペアに参加するプレイヤーは互いの位置を知ることができず，移動に先立って，5 種類の図形を二つ組み合わせてメッセージとして交換するとで，共通の移動先を決定することを求められた．メッセージに使用される図形は，実験開始の当初は意味が付与されておらず，課題を繰り返すことで徐々に意味づけが行われ，コミュニケーションプロトコルとして成立することが確認された [金野 16]．以上のように認知科学的視点から言語コミュニケーションの成立と進化に関して数多くの議論が行われてきた．しかしながら，これらの先行研究は扱う課題の状況が単純という問題がある．特に従来の研究では，コミュニケーションを行う実験参加者の保持する目標が協調的な関係にある課題のみを扱ってきた．これに対し，より現実の社会に近いモデルを考えるのであれば，協調だけでなく裏切りが行えるようなジレンマ要素を含む課題の設定が必須と言えるだろう．

本研究ではマキャベリ的知性仮説に従いつつ，既存の協調コミュニケーションに関する研究を拡張した実験を行う．それにより，人間の知能の進化過程の示唆に有用な知見を得ることを目指す．それに対するより具体的な問いとして本研究では，通常の言語コミュニケーションが制限された環境下でジレンマ状況を与えることでどのような言語的なコミュニケーションプロトコルが成立し得るのか探索し，参加者同士の相互作用によってどのような変化が起きるのか検討する．

2. 方法

実験は 3 ペア，6 人を対象に実施した．実験参加者は 20〜24 歳 ($M_{age} = 21.7, SD_{age} = 1.25$) の男性であった．本実験はどのような条件下でコミュニケーションプロトコルが成立，進化するのかを明らかにすることを目的としている．よって実験環境に複数のパラメータを設けることでその条件を探索的に検討した．

本実験では金野らの研究を踏襲し，実験課題としてゲームを用いた．ゲームは複数ラウンドから構成され，各ラウンドは「部屋割り当て」,「メッセージ交換」,「移動」の 3 フェーズから構成される．本研究では金野らの研究を拡張し，マップを 3×3 に拡大し，メッセージとして使用できる図形の数を 10 種類に増やした．さらにジレンマ状況を構築するため，2 名のプレイヤーに加え,「報酬」を新たにマップに与えた．ゲーム中，各プレイヤーは隣の部屋に存在するプレイヤー，あるいは報酬を視認することが可能であった．両プレイヤーのメッセージ交換・部屋移動が終わると次のラウンドがはじまり，2 名のプレイヤーのいずれかが報酬の配置された部屋に移動すると各プレイヤーおよび報酬はマップ上に再割り当てされ，次のラウンドが始まる．このような状況の中で，報酬の独占（報酬を一人で取得する）と共有（報酬を二人が同時に取得する）のジレンマが生じる設計を目指す．

連絡先: 井上直紀，静岡大学 情報学部，〒 432-8011 静岡県浜松市中区城北 3-5-1，ia15003@s.inf.shizuoka.ac.jp

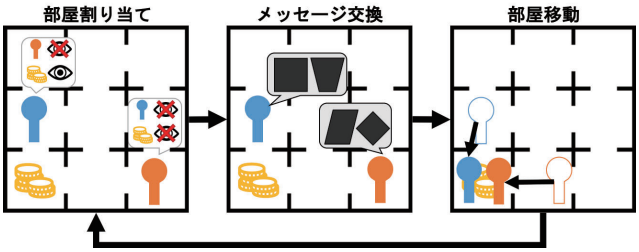


図 1: 実験課題

表 1: 実験の流れ

時間 (m)	フェーズ	
5	実験全体の教示	
10	協調ゲームの説明	協調フェーズ
30	協調ゲームの実施	協調フェーズ
10	協調ゲームのアンケート	協調フェーズ
5	ジレンマゲームの説明	ジレンマフェーズ
30	ジレンマゲームの実施	ジレンマフェーズ
10	ジレンマゲームのアンケート	ジレンマフェーズ
120	所要時間合計	

実験の流れを表 1 に示す．本実験は大きく「協調フェーズ」と「ジレンマフェーズ」の二つに分けられる．まず，協調フェーズで「共有でしか得点が得られない（二人が同時に報酬を取らなければならない）」という条件の中で，上記の課題を遂行する．協調フェーズは，プレイヤー間で互いの意思疎通をある程度可能な状態にするコミュニケーションシステムを構築するために設定する．次に，ジレンマフェーズで「共有することも独占することもできる（二人同時に報酬を取ることも一人で取ることもできる）」という条件で課題に取り組む．これにより裏切りの意図の隠蔽などの騙しのコミュニケーションの生起を図る．ジレンマフェーズでは独占によって得られる得点 (m:monopoly) と共有によって得られる得点 (s:sharing) の比率が実験参加者の行動に大きな影響を与えると考えられる．よって本研究では実験参加者に m が s の 2 倍であると教示した $m > s$ 条件と m と s が等しいと教示した $m = s$ 条件，m が s の 1/2 であると教示した $m < s$ 条件を設定した．なお，いずれの条件においても報酬を取らなかったラウンドは失敗（f:failure）とし，f は 0 ポイントとした．

協調フェーズではまずゲームのルールの説明後，課題における目標を「二人の得点の最大化を目指すこと」と教示したその後前述の条件下で課題を実施し，2 名のプレイヤーが 6 回報酬を共有することに成功した後，もしくはゲーム開始から 30 分経過した後にゲームを終了し，実験内容に関するアンケートを行なった．ジレンマフェーズではまず報酬の取得に関わるルールの変更について説明した後，課題における目標を「自分の得点の最大化を目指すこと」と教示した．ジレンマフェーズにおけるゲームは，ゲーム開始から 30 分経過後に終了した．実験の最後に，ジレンマフェーズでの振る舞いに関わるアンケートが実施された．

なお，それぞれのゲーム実施中は自分の考えをメモ用紙に記入するよう教示した．また，バイアスになることを避けるため，教示内ではゲームの基本ルールと得点が得られる条件のみ説明し，ゲームが協調課題であるかジレンマ課題であるかの説明は行っていない．

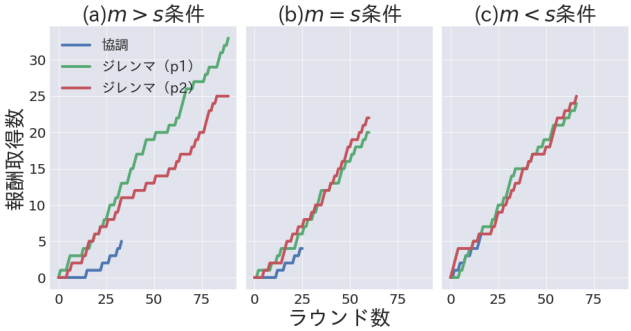


図 2: 報酬取得数の推移

3. 結果

3.1 行動分析

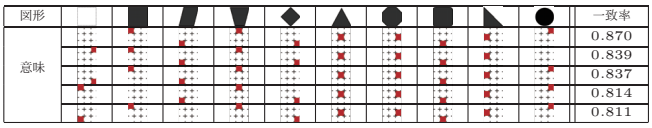
各条件において，ペアを組んだ二人のプレイヤー (p1,p2) が協調フェーズとジレンマフェーズで報酬を取得した累積回数の推移を図 2 に示す．図 2a から，$m > s$ 条件のジレンマフェーズにおいて他条件に比べプレイヤー間の報酬を取得した回数に大きな差が生じていることが確認される．これに関連して，ジレンマフェーズ後のアンケートにおいて，$m > s$ 条件の p1 は自分のことだけを考えたと回答し，p2 は相手が協力しなかったため次第に意思疎通の気力が失われたと回答した．さらに，$m = s$ 条件の p1 は「相手が 1 回嘘をついたのでそれから相手を信じなくなりました.」と回答し，$m < s$ 条件の p2 は実験後の内省報告において相手と協調フェーズで形成したコミュニケーションプロトコルを逆手にとって，でたらめなメッセージを送ることで撹乱しようとしていたという趣旨の回答をした．

3.2 メッセージ分析

メッセージ分析について，鮫島らは実験記号論を利用した協調ゲームにおける記号生成確率モデルを用いて実験参加者のメッセージデータの分析を行った [鮫島 16]．本研究では鮫島らの分析手法を参考にし，モデルと実際のデータの比較を行うことでメッセージデータを分析した．本実験では実験参加者にメッセージを送る際に自分の中でルールを決めて送信するよう教示した．そのため，まずプレイヤーが決めたルールを行動データから導くことができるのかを検討する．さらに，行動とメッセージの組み合わせが他のプロトコルに当てはまる可能性がないか，つまり形成された言語にどの程度多義性が見られるのかを検討する．

本来，新規なコミュニケーションにおいては，コミュニケーションプロトコルとして想定し得る図形の意味づけの組み合わせ全ての中から相手とのやり取りによって徐々に候補を絞っていくと考えられる．しかしながら，全ての可能性を考慮すると，候補となる組み合わせ数が極めて膨大になり，分析が困難である．そのためまず，アンケートから得られた情報を元に，本実験において成立したコミュニケーションプロトコルを分類した．アンケート結果によると $m > s$ 条件のプレイヤー双方ともに各マスに空白以外の 9 種類の図形を当てはめるように意味づけすることでコミュニケーションを取ろうとしていた（以下,「絶対座標タイプ」という）．また $m < s$ 条件の p1 は，自分から見た報酬の相対的な位置を上下左右の 4 方向をそれぞれ 4 種類の図形に当てはめ，自分が報酬を発見している（報酬に隣接した部屋にいる）ときとそうでないときにそれぞれのフラグに 2 種類の図形を使い，それを組み合わせることで意思疎通を図ろうとしていた．その他のプレイヤーは分析が困難な，

表 2: $m > s$ 条件のゲーム全体の平均一致率上位 5 位までの候補プロトコル



図形											一致率
意味											0.870
											0.839
											0.837
											0.814
											0.811

より複雑なプロトコルでコミュニケーションを取ろうとしていた．本項ではアンケートから得られたコミュニケーションプロトコルのタイプのうち絶対座標タイプについてのみ分析した．

前述の通り，絶対座標タイプは 9 種類で行き先を表すことができるので，送ることができる図形の二つの組み合わせのうち片方のみで行き先を表すことができる．本実験においても $m > s$ 条件の片方のみ実質的なメッセージとして使用していた（他方は協調ゲームの 89.7%のラウンドにおいて同一の図形（空白）を送り続けていたため使用していないと判断した）．よって，絶対座標タイプにおいてある片方のみ使用していたと仮定したとき，想定可能なコミュニケーションプロトコルの組み合わせは 10 種類の図形を 9 種類の行き先+未使用図形 1 種類=計 10 種類の意味から選ぶ順列組み合わせであり，${}_{10}\mathrm{P}_{10} = 10! = 3628800$（通り）となる．

本研究で行った分析では得られた実験データを分割し，ウィンドウごとにずらして比較するスライディングウィンドウ法を用いた．本研究では実際の行動データ（実際の行き先）と各条件，各プレイヤーの各プロトコルを用いるモデル（あるプロトコルを用いたときメッセージから想定される行き先）との一致率をウィンドウ幅 10 ストライド 1 で分析したものをヒートマップで可視化した（図 3）. 緑枠で囲まれた部分が協調ゲーム，赤枠で囲まれた部分がジレンマゲームの結果である．

この図において，y 軸は，本研究において正解プロトコルになりうる全てのコミュニケーションプロトコル（図形の意味づけ）約 360 万通りのゲーム全体での平均一致率上位から順に並べたものである．一方 x 軸は前述したウィンドウごとにまとめて分割された各ラウンドでとった行動（実際の行き先）である．この図は凡例から分かる通り，一致率が 1 に近づくほど白く明るく，0 に近づくほど黒く暗く着色されている．つまり図 3 において，各ヒートマップ内で明るい行があればその行が一致率の高いコミュニケーションプロトコルであり，明るい列があれば，多くのコミュニケーションプロトコルに当てはまるような行動をとっていたという意味となる．

なお，絶対座標タイプであった $m > s$ 条件において，この分析手法を用いて，もっともゲーム全体の平均一致率が高いものを参照すると，$m > s$ 条件のプレイヤーがアンケートにて回答したコミュニケーションプロトコル（以下,「正解プロトコル」という）と同一のものであった．以上のことからこの分析手法は一定程度信頼できると言える．正解プロトコルを含め，ゲーム全体の平均一致率が上位 5 位までのプロトコルを表 2 に示す．

また，これらのメッセージ分析に関連するアンケート結果として，コミュニケーションの内容について問う設問に対し,「意味が分かった図形と分からない図形が最後まで混在していた」（$m < s$ 条件 p2）,「自分と相手で同じ記号に対して別の意味で使用していた」（$m = s$ 条件 p1）といった回答があった．

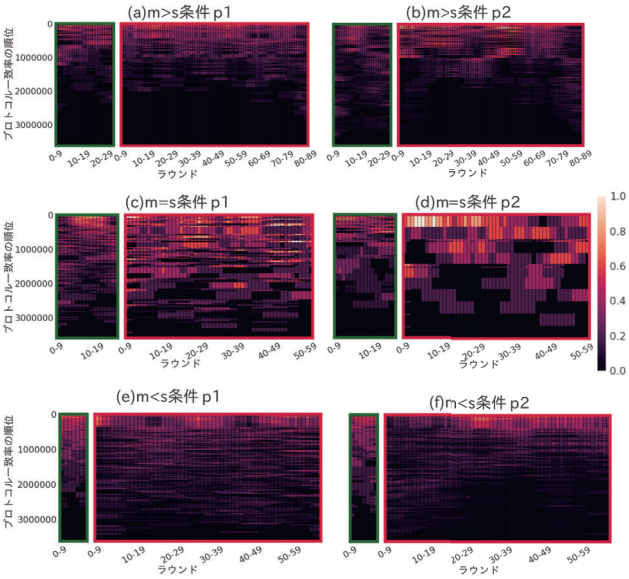


図 3: 各プレイヤーにおける各プロトコルと実際の行動の一致率

4. 考察

4.1 パラメータの調節による変化について

$m > s$ 条件の報酬取得数の推移（図 2a）から，このペアにおけるプレイヤー間の関係性が，搾取をする側（報酬を独占する側）とされる側に分かれていたと解釈できる．この解釈は，アンケート結果からも支持される．これにより，ゲーム内における行動およびコミュニケーションの内容が，実験で設定されるパラメータによって変化する可能性が示唆された．これは $m > s$ 条件のアンケート結果からも支持される．また他の主観評価からも，協調ゲームからジレンマゲームにかけて，ジレンマ要素の追加によって相手を裏切って報酬を騙し取るようなコミュニケーションが生起していたと考えられる．

4.2 言語の曖昧性について

言語の曖昧性について関連する結果について，$m > s$ 条件のメッセージと行動の一致率（図 3a,b）を協調フェーズとジレンマフェーズで比較すると，協調ゲームの序盤はヒートマップが全体的に暗い．これは候補となるプロトコルが分散していたことを意味する．しかしながら，協調ゲームの終盤からジレンマフェーズにかけては明暗が明確に別れ，候補となるプロトコルが収束していることがわかる．一方で最終盤においてもいくつか明るい部分は確認できるため，複数のプロトコルが候補として推定できる．このことは $m = s$ 条件と $m < s$ 条件からも確認できる．これらの条件のペアは本来絶対座標とは全く異なるタイプのプロトコルでやりとりしていた．しかしながら図 3c–f を確認すると絶対座標タイプが候補になり得る可能性も一定程度残されていることが示される．この結果は同じ図形に対して全く異なる意味づけを行っているにもかかわらず，異なる意味に解釈することもできるような状況が生起している可能性を示唆する．これはアンケートにおいても同様の趣旨の記述があったことからも支持される．

これらの結果から実験で生成されたコミュニケーションプロトコルがある一つの図形に対し，曖昧な意味づけや，多義性を持った意味づけがなされていた事例があったことを示している．

5. 結論

5.1 まとめ

本研究では 6 名 3 ペアを対象に参加する予備的な実験を行った．実験ではメッセージ付きジレンマゲームを用いることで，コミュニケーションプロトコルが成立し，変化する過程を観察した．実験で得られた行動データおよびメッセージデータの分析やアンケートの回答により，実験内で成立した言語的コミュニケーションプロトコルには多義語のような曖昧性が生じていたことが示された．また，協調フェーズにおいて二者間の合意のもと生成されたコミュニケーションプロトコルを，ジレンマフェーズにおいて意図的に異なる意味で用いることで，相手に対する裏切りの意図の隠蔽が行われていた可能性が示唆された．

5.2 今後の課題

本実験は予備的な実験として行われたものである．そのためゲーム課題の設計におけるパラメータ部分に幅をもたせて設定した．特に各ペアにおいて得点の条件を変えたため，比較や検討が困難になっている．また実験対象が 3 ペア 6 人と母数が少ないため，実験結果が実験課題や条件ではなく，実験参加者の特性に依存する可能性があり，データの信頼性の観点で疑問が残る．よってさらに実験を行うことでより多くのデータを収集する必要がある．その際は本研究で得られた知見を元により目的を絞った課題設計および実験環境を整備することが今後の課題となる．

参考文献

[Byrne 95] Byrne, R. W.: *The thinking ape: Evolutionary origins of intelligence*, Oxford University Press on Demand (1995)

[Galantucci 05] Galantucci, B.: An experimental study of the emergence of human communication systems, *Cognitive science*, Vol. 29, No. 5, pp. 737–767 (2005)

[Galantucci 09] Galantucci, B.: Experimental semiotics: A new approach for studying communication as a form of joint action, *Topics in Cognitive Science*, Vol. 1, No. 2, pp. 393–410 (2009)

[Scott-Phillips 10] Scott-Phillips, T. C. and Kirby, S.: Language evolution in the laboratory, *Trends in cognitive sciences*, Vol. 14, No. 9, pp. 411–417 (2010)

[金野 16] 金野武司, 森田純哉, 橋本敬：言語的コミュニケーションシステムの創発に関する実験的アプローチ, 計測と制御, Vol. 53, No. 9, pp. 801–807 (2016)

[鮫島 16] 鮫島和行, 金野武司, 李冠宏, 奥田次郎, 森田純哉, 橋本敬 他：コーディネーションゲームにおける記号生成確率モデルと 「意味」の推定, 研究報告数理モデル化と問題解決 (MPS), Vol. 2016, No. 5, pp. 1–6 (2016)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

インターネットテレビのニュース番組におけるコメント内容の分析

Analysis of viewers' comments on Internet TV news

西 朋里 *1 Tomosato Nishi　小川 祐樹 *1 Yuki Ogawa　服部 宏充 *1 Hiromitsu Hattori　高 史明 *2 Fumiaki Taka　高野 雅典 *3 Masanori Takano　森下 壮一郎 *3 Soichiro Morishita

*1立命館大学情報理工学部
College of Information Science and Engineering, Ritsumeikan University

*2神奈川大学
Kanagawa University

*3株式会社サイバーエージェント
CyberAgent, Inc.

In recent years, the global spread of the Internet has altered the influence of media on the formation of people's knowledge and concern. Therefore, clarifying the effect of new media is critical.

This study aims at clarifying the relevance between viewers' comments and news coverage contents on Internet TV. Specifically, we analyzed the words of viewers' comments using TF-IDF and multivariate analysis. The results indicated that critical and hateful remarks were being posted on high-entertainment soft news.

1. はじめに

近年，インターネットを中心とする多様な情報環境の浸透により，人々の知識や関心の形成に対する，既存のマスメディアによる影響が変化しつつある．以前は地上波テレビなどといったマスメディアが人々の共通した情報源となっており，個々の知識や関心の形成に大きな影響を持っていると考えられていた [1][2][3]．しかし，近年ではインターネットを通じた情報が人々の新たな情報源となる機会も多くなってきている．例えば，AbemaTV をはじめとするインターネットテレビもその一つである．

インターネットテレビは，時間や場所に依存せずに視聴できることや，自身の意見をコメントとして投稿できるなど地上波のテレビとは異なる特性を持っている．そういった特性を好み，インターネットテレビを主な情報源としている人が一定数いると考えられる．しかし，インターネットテレビに対する分析は少なく，視聴者に与える影響は明らかになっていない．そのため，インターネットテレビの影響分析を進めるにあたって，まず視聴者に対する効果を明らかにする必要がある．

メディア視聴者に対する効果を分析する手法として質問紙調査が挙げられる．質問紙調査は回答者の心理的態度を測定するための有効な方法ではあるが，自己報告による回答であるため行動量に関しては客観的なデータではないという点で課題がある．そこで，本論文では株式会社 AbemaTV より提供された行動ログデータを分析に利用し，行動ログデータに含まれる視聴者のコメントを発言と見て分析を行う．行動ログデータはユーザに対して正確な情報を持っているため，分析を通じてユーザの具体的な行動を把握できるというメリットがある．ニュース番組は放送内容から主に 2 種類に分けられ，事実の伝達を目的とするハードニュースと政治家のスキャンダルなどバラエティ要素を交えて政治情勢を伝える娯楽的要素の強いソフトニュースがある [4]．それぞれの放送内容から視聴者の政治知識に与える影響が異なる [5][6] ことから，ハード/ソフトニュース視聴による視聴者の発言の違いがあると考える．

よって，本論文ではインターネットテレビに投稿されるコメントを視聴者の発言として分析し，ハード/ソフトニュースとコメントの関連性を明らかにすることを目的とする．

連絡先: 西朋里，立命館大学情報理工学部，滋賀県草津市野路東 1-1-1, is0320kp@ed.ritsumei.ac.jp

2. 関連研究

ニュース視聴が人々の知識や関心に与える影響について様々な研究がなされており，その一つとして，地上波のテレビ番組において前述したハードニュース，ソフトニュースのそれぞれが視聴者に及ぼす効果についての研究がなされている．稲増ら [5] の研究では，番組内容を詳細に記したテキスト情報をもとに番組をハード／ソフトニュースに分類し，これらのニュースへの接触と政治知識や政治意識との関連性を分析している．この研究では，ハードニュースへの接触が多いほど選挙への関心と政治への関与が高くなる結果が報告されている．稲増ら [5] はこの結果に対して，政党の主張や政策争点を取り上げて政治を報道するような番組であれば，選挙への関心や政治への関与と結びつきうるという見解を述べている．また，政治知識の少ない群において，ソフトニュースの視聴が多いほど政治関心が高く，政治知識の多い群においては，ソフトニュースの視聴が多いほど，若干ではあるが政治関心が低いという結果が出ている．つまり，ニュース視聴の効果は個人の政治知識の多寡によって決定される部分があるという調査結果が出ている．政治意識に対するマスメディアの効果に個人のメディア・リテラシーと批判的思考態度の影響を調査した末吉ら [6] の研究も存在する．この研究では，日常的に視聴しているニュースと政治意識に関する質問紙調査を行い，ニュース視聴が視聴者の思考に与える影響を調査している．この研究によると，ハードニュース視聴は政治意識を高めるという前述した研究と同様の結果が報告されている．ソフトニュース視聴は政治知識の多寡によらず政治への関与に対して負の効果を持つことが示されている．

以上の関連研究から，視聴者の発言に着目するとハードニュースは政治知識に関するコメントが多く，ソフトニュースは政治的要素と関連性の薄いコメントが多くなると考えられる．

3. ハード／ソフトニュースにおけるコメントの分析

本論文では，インターネットテレビに投稿されるコメントをテキストデータとして用意し，コメントに含まれる単語からハード/ソフトニュースそれぞれの発言の違いについて分析する．以下にデータセットの構築と分析手法について述べる．

3.1 データセット

本論文では，株式会社 AbemaTV より提供を受けた AbemaTV の視聴ログデータを用いて分析を行う．AbemaTV とは，株式会社 AbemaTV が運営するインターネットテレビ局であり，独自のニュース番組や生放送番組など多彩な番組を 24 時間，無料で閲覧できる．配信される動画は，スマートフォン・タブレットに加え，PC や TV アプリケーション (e.g., Apple TV, Google Cast) などで視聴が可能である．提供を受けた視聴ログデータは，2018 年 6 月 1 日から 2018 年 6 月 30 日の期間に放送された番組を対象としたデータであり，放送された番組のデータと投稿されたコメントのデータを含んでいる．なお，提供されたデータはユーザ情報を匿名化しており，コメントした個人を特定できない仕様となっている．AbemaTV では放送チャンネルが設定されており，本論文ではニュース番組を放送している AbemaNews チャンネルを対象とする．

AbemaTV では，スロットと呼ばれる放送枠が採用されており，スロットに各番組を当てはめることで放送を行なっている．本論文では，スロットごとにコメントを集約したものを 1 文書とする文書データを作成する．スロットごとにコメントを集約する理由は，1 コメントを 1 文書として取り扱った場合，1 コメントが短いこととコメント数が多いことから単語の出現頻度が低くなり，分析で使用する TF-IDF の結果に影響が出るためである．それぞれの文書データに対して MeCab*1 を用いた形態素解析を行い，単語ごとの分かち書きに変換する．形態素解析をするにあたって，辞書に含まれない単語の除去，不要な英数字や特殊文字の除去，ストップワードの除去を行った．辞書には NEologd*2 を利用し，ストップワードには内容推測に適したキーワード抽出のための日本語ストップワード [7] と Slothlib のストップワード *3 を利用した．上記の手順を踏まえて作成したデータを文書データとして分析に用いる．

加えて，AbemaNews チャンネルで放送されている番組を [6] を参考に，出演者が意見することが多く番組内での企画が多い番組をソフトニュース，それ以外の番組をハードニュースに分類した．

表 1: Abema-News 内の番組のハード/ソフトニュース分類

ハードニュース	ソフトニュース
AbemaNews24 時	AbemaMorning
AbemaPrime	AbemaSports！
けやきヒル’sNews	W の悲喜劇
テレメンタリー	みのもんたのよるバズ
報道ステーション	グッド！モーニング
	スーパー J チャンネル
	ワイド！スクランブル
	千原ジュニアのキング・オブ・ディベート
	日曜スクープ
	羽鳥慎一モーニングショー

*1 http://taku910.github.io/mecab/
*2 https://github.com/neologd/mecab-ipadic-neologd
*3 OSDN.net

3.2 ロジスティック回帰によるハード／ソフトニュースに関連する単語の分析

ハード/ソフトニュースとコメントの関連性を分析するに当たって，ニュース番組に投稿されるコメントを構成する単語に着目し，単語とハード/ソフトニュースの関連性を分析する．分析手法の概要を図 1 に示す．上記で作成した文書データを TF-IDF を用いて文書ベクトルに変換した．この文書ベクトルは，各文書ごとに出現する単語の TF-IDF 値を示している．また，この文書ベクトルは出現頻度が 1%以上の単語のみで構成した．各単語を説明変数 x，ハード/ソフトニュースを目的変数 y として，L1 正則化ロジスティック回帰分析を用いて回帰係数を算出した．このとき，ハードニュースを $y = 1$，ソフトニュースを $y = 0$ と設定した．算出した回帰係数を指数変換することで各単語のオッズ比を算出した．算出したオッズ比を評価指標として利用し，オッズ比が 1 超過の単語はハードニュースと関連があり，1 未満の単語はソフトニュースと関連があるとする．

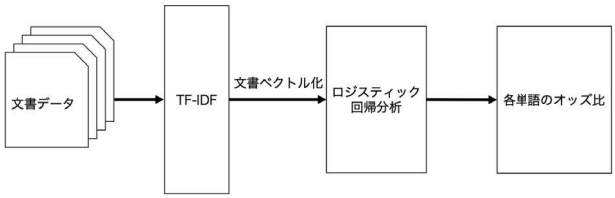


図 1: 分析手法の概要

4. 分析結果

文書データに対して，TF-IDF を用いて 784 単語の TF-IDF 値を算出し，その上位 20 単語を表 2 を示す.「日本」,「韓国」,「アメリカ」などの海外に関する単語が高い TF-IDF 値を示しており，これは外交に関する報道の影響と思われる.「お前」,「かわいい」,「こいつ」，などの日常会話で用いる単語が高い TF-IDF 値を示しており，これはコメント投稿が日常会話と近い感覚で行われているからと考えられる．時期的な影響として，2018 年 6 月には大阪北部地震，FIFA ワールドカップ，米朝首脳会談があったため,「地震」,「サッカー」,「トランプ」等の単語が抽出されている．

次に，文書ベクトルからロジスティック回帰分析によって求めた単語ごとのオッズ比を見る．オッズ比の上位 20 単語と下位 20 単語を表 3, 表 4 に示す．オッズ比の上位 20 単語では,「ご飯」,「優しい」,「復活」といった単語が強い関連性を示しているが，理由は放送内容から考察する必要がある．また,「小松」,「柴田」,「アナウンサー」などの人名と思われる単語が高い値を示している．これは，ハードニュースに属する番組のみに出演しているアナウンサーや芸能人による影響だと考えられる．加えて，前述したように地震などの単語が上位語に含まれているのは時事の影響と考えられる．そして,「総理」,「左翼」という単語から，わずかに政治的内容と関連があるように考えられる．オッズ比の下位 20 単語では,「殺害」,「貧乏」,「低位」といったネガティブな印象を持つ単語や「おかしい」,「ウヨ」,「アホ」といった批判的な意味を持つ単語が抽出されている．また,「韓国人」,「半島」などの単語が下位語に含まれており，これは差別的なコメントの影響と考えられる．よって，ソフトニュースと批判的な単語や差別的な単語が関連性があるという結果を得た．

表 2: TF-IDF 値の上位 20 単語

単語	TF-IDF 値
日本	0.048021
安倍	0.032962
お前	0.026305
かわいい	0.022149
日本人	0.019947
こいつ	0.019704
在日	0.019503
仕事	0.019459
韓国	0.018963
怖い	0.018854
可愛い	0.018433
問題	0.018233
トランプ	0.017683
今日	0.017119
サッカー	0.017017
好き	0.016934
地震	0.016881
子供	0.016148
すごい	0.015899
ええ	0.015734

表 3: オッズ比の上位 20 単語

単語	オッズ比
小松	2.298781
ご飯	2.213871
復活	2.184056
総理	2.044065
柴田	2.043512
最強	1.930302
証拠	1.917079
地震	1.870551
洗脳	1.847198
ピンク	1.822233
自称	1.820762
先生	1.797857
アナウンサー	1.781289
左翼	1.743556
英語	1.732251
うまい	1.729077
開発	1.696934
優しい	1.665322
興味	1.646466
オッサン	1.621961

表 4: オッズ比の下位 20 単語

単語	オッズ比
貧乏	0.349703
意味	0.359977
殺害	0.392264
ウヨ	0.392374
韓国人	0.392751
えな	0.406409
低位	0.412941
半島	0.443341
スマホ	0.445431
結果	0.483959
ども	0.491187
アメリカ	0.491433
かっこいい	0.520964
ブロック	0.527381
な!	0.534321
しん	0.534575
おかしい	0.536381
アホ	0.540805
彼氏	0.550286
揺れ	0.558001

5. 考察

ハードニュース視聴においては，外交問題や国政などの具体的な政治に関する単語が多くなる傾向は見られなかったが，表 3 よりアナウンサー等の番組出演者に関する単語がハードニュース視聴の発言傾向として見られた．逆に，ソフトニュースの中には特定の芸能人が持つ冠番組があったが，表 4 から見て取れるように特定の人物には含まれていない．これは，ハード/ソフトニュースで番組出演者の印象が異なり，ハードニュースでは番組出演者に対する視聴者の関心が強いと考えられる．

次に，ソフトニュースのコメントから批判的な単語が抽出された理由として，出演者と視聴者の影響が考えられる．ソフトニュースは，報道内容に出演者の考えや意図を加えて視聴者に情報を与えるため，視聴者が持つ意見とギャップが生まれる可能性が生じ，批判的な発言 (コメント) に結びついたと考えられる．また，放送内で特定の出演者による過激な他者批判や扇動的な発言などが行われた場合，その意見に同調する視聴者によって批判的なコメントが投稿されている可能性も考えられる．加えて，韓国・朝鮮人に関する報道に対して保守派の視聴者による批判的なコメントや差別的なコメントが投稿されている可能性も考えられる．視聴者層として，批判や煽りを好む視聴者がソフトニュースを積極的に視聴しコメントを行っていることが考えられる．その結果，批判や煽りのコメントが増えていくため，より視聴層を偏らせている可能性が考えられる．[6] に倣って質問紙調査を行う場合，こういった視聴者の心理的態度の測定を検討する必要がある．

6. おわりに

本論文では，インターネットテレビを通じたニュース視聴による視聴者の発言に着目し，AbemaTV のニュース番組に投稿されるコメント内容の分析を行った．その結果，ソフトニュースは批判的な単語や差別的な単語と関連があることがわかり，ハードニュースは放送内容とわずかに関連があることがわかった．

今後の展望として，本論文ではコメントデータのみを用いて分析を行ったが，先行研究が実施しているように視聴者に対する質問紙調査を行い，視聴者層や政治的態度を踏まえてニュース視聴の効果を明らかにする必要がある．また，番組の放送内容を踏まえたコメントの分析，ユーザごとの視聴傾向を考慮した分析を行う必要もある．さらに，コメントの影響を明らかにするために，視聴者が他者のコメントを見たことによる思考の変化について分析を進める必要がある．

参考文献

[1] Bennett, S. E., Flickinger, R. S., Baker, J. R., Rhine, S. L. and Bennett, L. L.: Citizens' knowledge of foreign affairs, The Harvard International Journal of Press/Politics, vol. 1, pp. 10-29, (1996).

[2] Mutz, D. C. and Martin, P. S.: Facilitating communication across lines of political difference: The role of mass media, American Political Science Review, vol. 95, No. 1, pp. 97–114, (2001).

[3] Baum, M. A.: Soft news goes to war: Public opinion and American foreign policy in the new media age. Princeton, NJ: Princeton University Press, (2003).

[4] 萩原滋, ニュース番組の内容と形式 : 娯楽化傾向の検証と番組の類型化, 萩原滋編著『変容するメディアとニュース報道』, 丸善, pp.67-114, (2001).

[5] 稲増一憲, 池田謙一, 多様化するテレビ報道と, 有権者の選挙への関心および政治への関与との関連: 選挙報道の内容分析と大規模社会調査の融合を通して, 社会心理学研究, 第 25 巻, 第 1 号, pp. 42-52, (2009).

[6] 末吉南美, 三浦麻子, 政治意識へのマスメディア効果に対するメディア・リテラシーと批判的思考態度の影響, 関西学院大学心理科学研究, 第 39 号, pp. 81-88, (2013).

[7] 國府久嗣, 山崎治子, 野坂政司, 内容推測に適したキーワード抽出のための日本語ストップワード, 日本感性工学会論文誌 原著論文, 第 12 巻, 第 4 号, pp.511-518, (2013).

4:20 PM - 5:00 PM (Wed. Jun 5, 2019 3:20 PM - 5:00 PM Room D)

[2D4-OS-1a-04] (Invited talk) Can business cards network approach unresolved problems in social science?

〇Naoki Maejima[1], Tomonori Manabe[1] （1. Sansan, Inc.）

Keywords: Social Network Analysis, Business Card, Job Change

名刺交換に基づくビジネスネットワークは、従来の社会科学では明らかにできなかった経済的・社会的行為や価値に関する問題への計算社会科学的アプローチを可能にする。本講演では、転職や企業成長、ブランド価値、業界による「人脈作り」の違いなど、様々な観点から、名刺交換ネットワークを理解することでどのような新たな知見を創出できるのかについて報告する。

©The Japanese Society for Artificial Intelligence

Organized Session | Organized Session | [OS] OS-1

[2D5-OS-1b] 計算社会科学(2)

鳥海 不二夫（東京大学）、笹原 和俊（名古屋大学）、吉田 光男（豊橋技術科学大学）、瀧川 裕貴（東北大学）、榊 剛史（ホットリンク）、高野 雅典（サイバーエージェント）

Wed. Jun 5, 2019 5:20 PM - 7:00 PM Room D (301B Medium meeting room)

[2D5-OS-1b-01] Analysis of personal attribute relationships based on online research and social data

〇Baofa Du[1], Kazutoshi Sasahara[1,2] （1. Nagoya University, 2. JST PRESTO）

5:20 PM - 5:40 PM

[2D5-OS-1b-02] Retrieving and analyzing interpersonal relationship from large-scale mobility data.

〇Kimitaka Asatani[1], Fujio Toriumi[1], Masanao Ochi[1], Junichiro Mori[1], Ichiro Sakata[1] （1. University of Tokyo）

5:40 PM - 6:00 PM

[2D5-OS-1b-03] Proposal of a method to estimate insurance claims caused by heavy rain

〇Takeshi Okazaki[1], Oriol Gaspa Rebull[1] （1. Aon Benfield Japan Ltd.）

6:00 PM - 6:20 PM

[2D5-OS-1b-04] Detection of indirect dominance by China on the global shareholding network

〇Takayuki Mizuno[1,2], Mitsuhiro Odaka[1,2] （1. National Institute of Informatics, 2. School of Advanced Sciences, SOKENDAI）

6:20 PM - 6:40 PM

[2D5-OS-1b-05] Expressing Modern/Oldfashioned Racism depending on News Contents of Internet Television

〇Masanori Takano[1,4], Fumiaki Taka[2], Soichiro Morishita[1], Tomosato Nishi[3], Yuki Ogawa[3] （1. CyberAgent, Inc., 2. Kanagawa University, 3. College of Information Science and Engineering, Ritsumeikan University, 4. The University of Tokyo）

6:40 PM - 7:00 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Analysis of personal attribute relationships based on online research and social data

Du Baofa[1] Kazutoshi Sasahara[1,2]

[1] Graduate School of Informatics Nagoya University [2] JST PRESTO

Although social data can be an important source of information for studying personal attributes, it remains unclear which kinds of attributes can be inferred from social data and meaningfully used in the social sciences. In response, we analyzed the relationships between a variety of attributes, including political alignment, morality, food identity, and brand preferences, based on social data (i.e., Twitter text) and online survey data. Our results revealed a relationship between political alignment (liberal and conservative) and morality, food identity, and brand preferences. In addition, liberal people were more likely to belong to the food left-wing category. As for brand preferences, the conservative people showed a higher preference for Japanese companies than liberal people.

1. Introduction

Although inferring personal attributes from social data is important in research in the social sciences as well as application in social media, complex relationships in person attributes remain unstudied. In our research, we analyzed social data—namely, the tweets of Twitter users—in relation to their personal attributes, as identified by an online survey. In particular, we investigated the relationships between each user's political alignment, morality, food identity, and brand preferences (Fig. 1).

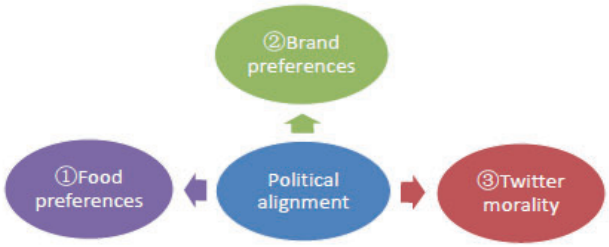


Figure 1. Research Diagram

The framework for our research draws from two important concepts. The first, Moral Foundations Theory developed by Jonathan Haidt [1], holds that every person innately has five moral foundations that allow him or her to intuitively judge moral situations. They are:

- Harm (i.e., disliking the pain of others and feeling responsible for protecting vulnerable people);
- Fairness (i.e., taking the right or just action based on accepted rules);
- Ingroup loyalty (i.e., being loyal to social ingroups, including family and nation);
- Authority (i.e., respecting and obeying tradition and legitimating authority);
- Purity (i.e., feeling antipathy for disgusting things and contamination).

We measured those five moral foundations in Japanese-language texts with reference to the Japanese Moral Foundations Dictionary (J-MFD) [2].

Contact: Du Baofa, Graduate School of Informatics, Nagoya University, du.baofa@h.mbox.nagoya-u.ac.jp

The second concept is food identity, which at base proposes two types of people. Whereas proponents of the so-called "food left-wing" typically advocate natural foods and health consciousness (e.g., vegetarians and vegans), proponents of the so-called "food right-wing" generally consume any available food products and enjoy eating fast food. Sasahara (2018) has shown that food identity can serve as an appropriate proxy for personal attributes and can offer insights into potential buying patterns [3].

Mobilizing those two concepts, we examined the relationship between the self-reported political alignment and food identity, brand preferences, and morality.

2. Method

2.1 Data

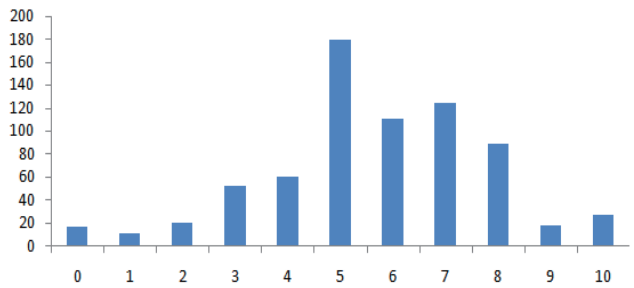


Figure 2. Distribution of users with different political alignment on an 11-point scale (0: very liberal, 10: very conservative)

First, we gathered the responses of a sample of 703 participants from online users to a questionnaire addressing political alignment, food preferences, and brand preferences in statements that participants rated on an 11-point scale (0 = *very liberal*, 10 = *very conservative*). Participants also indicated their political alignment on the same scale (Fig. 2). We placed ones whose scores ranged from 0 to 2 in the liberal group and ones whose scores ranged from 8 to 10 in the conservative group. As we were only focusing on people from these two groups, we excluded the ones whose scores ranged from 3 to 7.

Second, upon receiving permission from the participants, we also collected Twitter data representing each user by using Twitter's API. After gathering all posts (i.e., tweets, retweets, and replies) from 650 active Twitter accounts with more than 2,000 posts, we quantified the tweets in terms of moral foundations based on the J-MFD.

2.2 Data Analysis

We analyzed the relationships among political alignment, food identity, and brand preferences as indicated in the questionnaire data for each group of participants (i.e., the liberal group and the conservative group). Our calculations prioritized the value of political alignment in identifying links among other attributes.

For morality analysis, we used MeCab to separate Japanese-language texts into words and extracted the moral loadings from each participant's tweets with reference to the J-MFD. The moral loading was the percentage of words in the J-MFD also used in tweets. Based on the moral loading, we pinpointed aspects of morality in relation to political alignment.

3. Results

Figure 3 shows the relationship between political alignment and food preferences, the latter of which we divided into three types: fast food (i.e. food right-wing), healthy food (i.e., food left wing), and basic food (i.e., base ingredients). Our results reveal a relationship between political alignment and food preferences; liberal people prefer healthy foods more than conservative ones, whereas conservative people prefer fast food more than liberal ones. The groups did not indicate a significant difference in preferences for basic food.

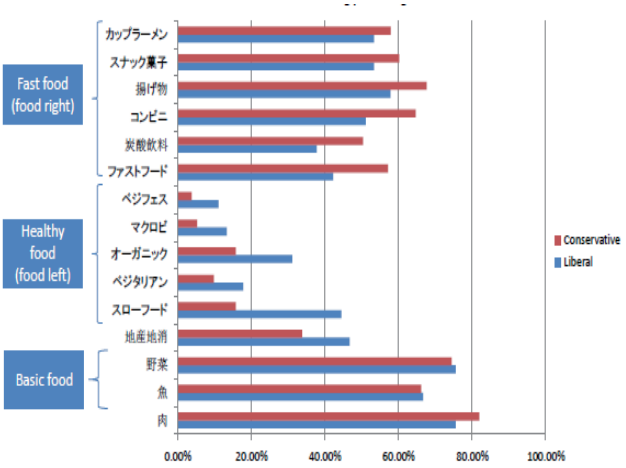


Figure 3. Food liking percentage

We also analyzed the difference in preferences for specific brands between the two groups. Companies from Japan such as Nissan and Toyota were more preferred in the conservative group, whereas the majority of participants who preferred popular brands such as IKEA and Facebook were in the liberal group (Fig. 4).

Figure 4. Brand liking percentage

To measure moral loadings, we separated tweet texts into words and computed the frequencies of words related to "virtue" and "vice" as defined in the J-MFD. Figure 5 shows that participants in the liberal group had higher virtue loadings than ones in the conservative group in most of the foundations except purity. Furthermore, the frequencies of virtue words related to ingroup and authority were greater than any of the other moral foundations. By contrast, vice-related words indicated harm more than any other moral foundation (Fig. 6), with fairness, with the minimum proportion, as a remote second. By political alignment, participants in the liberal group had higher vice-relate moral loadings for all of the five moral foundations, especially for harm and fairness.

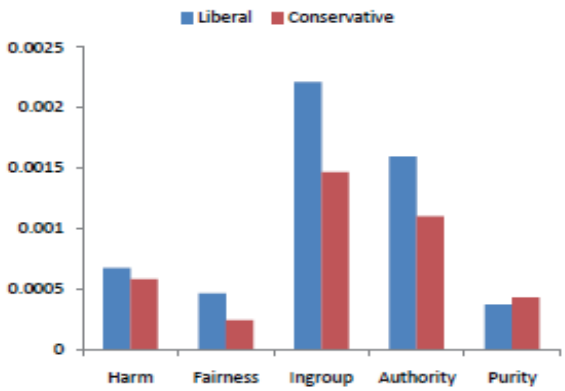


Figure 5.Virtue loadings by moral foundation

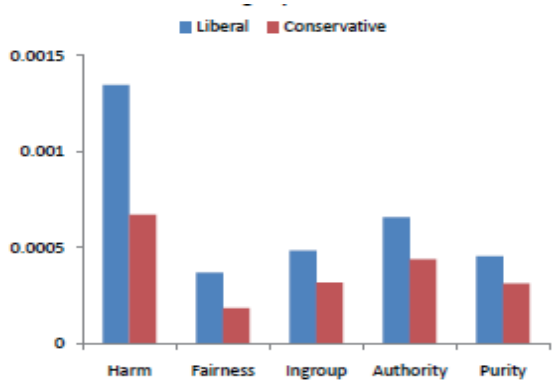


Figure 6. Vice loadings by moral foundations

Last, we visualized word frequency using word cloud (Fig. 7). For ingroup, “社会,” ”国,” ”国民,” and “家族” often appeared in the tweets of participants in the liberal group, while, for harm, “虐待,” “被害,” “战争,” and “平和” appeared frequently for them as well. In such tweets, participants in the liberal group often demonstrated their morality via ingroup loyalty (i.e., virtue), authority (i.e., virtue), and harm (i.e., vice).

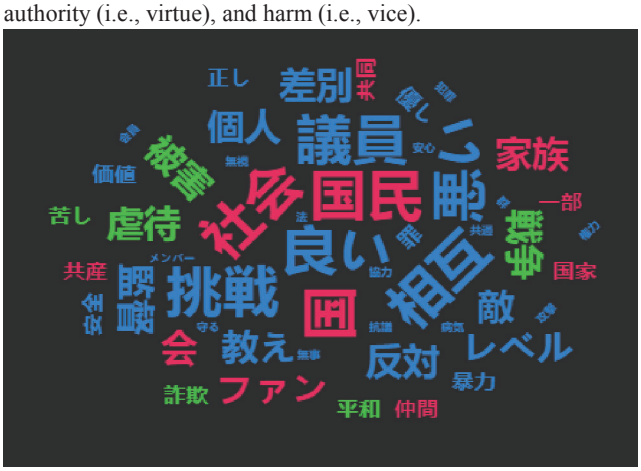


Figure 7. Morality word cloud (red: ingroup virtue, green: harm vice, blue: other foundations)

4. Discussion

In our examination of relationships among the personal attributes of political alignment, food identity, and brand preferences, preference for fast food was more popular than one for healthy food. Participants in the liberal group were more likely to belong to the food left-wing, whereas participants did not differ in terms of their preference for basic food. Meanwhile, regarding brand preferences, participants in the conservative group more than ones in the liberal group preferred Japanese companies (e.g., Toyota). Combining food preferences and brand preferences, participants in the conservative group who were more likely to prefer fast food were also likely to prefer fast-food brands.

In sum, liberal people and conservative people showed starkly different food and brand preferences. Accordingly, our results indicate a relationship among political alignment, food identity, and brand preference. As for personal attributes related to morality, liberal people express their feelings more actively than conservative ones, as well as demonstrated higher morality loadings. Words related to politics closely also show a high frequency of occurrence.

5. References

[1]. Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P.H., Moral foundations theory: The pragmatic validity of moral pluralism, In advances of experimental social psychology 47, pp.55−130, 2013.

[2]. Matsuo, A., Sasahara, K., Taguchi, Y., Karasawa, M., Development and Validation of the Japanese Moral Foundations Dictionary, arXiv:1804.00871, 2018

[3]. Sasahara, K., You are what you eat: A social media study of food identity, arXiv:1808.08428, 2018

Acknowledgement

This research was supported by JSPS/MEXT KAKENHI Grant Number JP17H06383 in #4903, JST PRESTO Grant Number JPMJPR16D6, and JST CREST Grant Number JPMJCR17A4.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

移動データを用いた人間関係の抽出とその性質の分析

Retrieving and analyzing interpersonal relationship from large-scale mobility data.

浅谷　公威 *1　　鳥海　不二夫 *2　　大知　正直
Kimitaka Asatani　　Fujio Toriumi　　Masanao Ochi

森　純一郎　　坂田　一郎
Junichiro Mori　　Ichiro Sakata

This paper is extended abstract of our recently published paper[Asatani 18]. Interpersonal relationship affects our daily activity. In this paper, we consider the link between them by analyzing large-scale mobility data. In the past studies of human mobility, owing to the difficulty of associating mobility data with interpersonal relationship data, previous studies on the link between interpersonal relationships and mobility are limited to the specific activities of particular users. In this paper, we propose a method for detecting interpersonal relationships from mobility data, while distinguishing these relationships from those of familiar strangers such as commuters. In the method, persons who take diverse variations within the same activities are recognized as a pair. From IC card data covering the daily mobility of six million people over three years, we detected millions of frequently co-located pairs. Under certain conditions, most of the detected pairs are confirmed as not being familiar strangers, but rather to have an interpersonal relationship. Next, we analyzed the detected pairs and found that the density of the relationships between groups was divided by gender and age and was found to be asymmetric by gender. For example, an elderly male person is not likely to take trips as a pair with a same-gender elderly person, and this result is data-based evidence for the isolation of retired men. In addition, group trips are confirmed to have an extraordinal character and sometimes converge spatiotemporally. These findings indicate that interpersonal relationship is a strong factor to determine their mobility and group observation is potentially useful for event detection.

1.　はじめに

本論文は [Asatani 18] で発表した内容の Extended abstract である．人の移動を理解し予測することは，感染症阻止，交通の最適化，イベント企画などの多くの応用につながる．これまでの研究の大部分は，人間の移動行動 [Gonzalez 08] や場所の特性 [Giannotti 11] に焦点をあてることで人間の移動を理解・予測してきた．いくつかの応用的な研究は，犯罪の可能性 [Bogomolov 14] と土地利用 [Grauwin 17] のパターンは人間の移動性に基づいて推定できること示している．しかし，さまざまな動機から生じる人間の移動は完全には理解・予測することはできない [Song 10]．その大きな要因の一つは，人間の移動はしばしば人間関係によって動機付けられることであると考えられる．

ある先行研究はデータ解析により社会的相互作用が人間の移動の予測に重要な情報であることを示している [De Domenico 13]．社会的つながりと人間の移動性との関係についても解析が進んでいる．[Fan 17]．しかし，これらの研究の対象期間や対象となる人間は非常に限られている．そこで，データを用いない理論的なアプローチによる人間関係と移動についての考察も行われきた [Miller 05]．このように人間の移動と人間関係に多くの注意が払われているにもかかわらず，一般の人間の日常におけるそれらに関する理解はあまり進んでいない．たとえば，どのような人々が一緒に移動しているか，人間関係が移動に与える影響は何か，グループでの移動の特性などの問いに対する考察は進んでいない．

より広い文脈で考えると，人間関係は人間社会の基本的要素であり，個人の精神的および肉体的健康と強く関連している [Newman 12]．これまで多くの研究者は，オンラインコミュニティや画像共有サービス [Cheung 15] における隠れた人間関係の推定や，人間関係とインターネット上での活動との関連を調査してきた [Pflieger 10]．ただし，これらのデータは各ユーザーの特定の活動に限定されており，日々の日常の行動がよく表されているデータを用いて人間関係を理解することが必要である．そのようなデータの一つとして，広範囲にわたって収集された人間の移動データがあげられる．例えば，大都市に住む人々にとって，鉄道システムからの移動データは彼/彼女らの日常生活の活動の多くを反映している [Shimosaka 15]．したがって，人間関係データを人間関係と結び付けることが出来れば，私たちが日々の人間の相互作用をよりよく理解できるだろう．しかしながら，データの制限とプライバシーの問題から，このようなデータを組合わせることには現実的に難しいと考えられる．

本研究では，人間の移動と人間関係を紐づけたデータがない状況で人間の移動と人間関係に関する上述の問を考察するために，人間の移動そのものから人間関係を推定する方法を開発した．データとして，関西圏 (大阪，神戸，奈良，京都を含む日本で 2 番目に大きい都市圏) から 600 万人を対象として 3 年間に収集された 22 億の移動ログを用いた．電車による移動は，IC カードを使用して電車で移動する人々の日々の活動に不可欠であるため，これらのデータはさまざまな人間の日常的/非日常な移動を含む活動の多くを反映している．本研究では，一定数以上の出発/目的地の組合わせにおいて同一トリップ (同じ目的地と同じ出発地をそれぞれ同時間) を行っている二人の人間の間に友人関係があるとみなした．

定義した手法を用いて移動データから頻繁に共存する数百万もの人間関係を抽出した．また，意図しない同一トリップを頻繁にする関係のない同乗者 (通勤客など) と人間関係がある二人の組合せの間を区別するために，タイム時間ウィンドウ幅の変化に関して検出された二人の組合せの違いを観察した．ある条件下では，時間ウィンドウ幅の増加が検出された二人の組合せに与える影響が小さく，検出された二人の組合せの大部分は一緒に移動する強い意志を持つ人間関係を持っていると推測された．抽出した関係性を使い，大都市における我々は大規模な人間関係グラフを観察することにより，性別による人間関係の違いを明らかにした．さらに，誰かと一緒に移動することによる人間の行動の変化 (例えば普段行かないところに行くなど) や，グループでの移動 (人間関係がある人との同時移動) が時空間的に収束することを明らかにした．

連絡先: 浅谷公威，東京大学工学系研究科，東京都文京区本郷 7-3-1，asatani@gmail.com

2. データ

分析に用いるデータは，2014 年 4 月から 2016 年 3 月までの間の関西圏 (大阪，神戸，奈良，京都を含む日本で 2 番目に大きな都市圏) における 600 万人の匿名ユーザーによる 22 億回の移動である．関西の人口の約 25%にあたる 500 万人が 3 年間で 100 回以上鉄道を使用している．また，各移動は乗車駅と降車駅における乗車/降車の時間が分単位で記録されている．また，全体の 33%のユーザーに関しては 10 年ごとの年齢ラベルと性別の情報が含まれている．また，データにはどのユーザー間に人間関係があるかを示す情報は含まれていない．

3. 手法

我々は人間関係を抽出するために異なる 2 つの方法を開発した．それぞれの手法はともに，頻繁に同時移動する二人の組合せを抽出している． 1 つ目は「同一トリップ検出」である．この方法では，少なくとも t_{id} の同一トリップをする二人の組合せが，頻繁に同時移動する二人の組合せとして検出される．時間ウィンドウ幅が $w = 1(min)$ の場合，各個人が同じ目的地と同じ出発地をそれぞれ分単位で同時刻に入場/退場していれば同一トリップとなる． $w \geq 2$ の場合，二人の個人の出発時刻と到着時刻の差が $w - 1(min)$ 以内の場合に同一トリップと見なされる．少なくとも t_{id} の同一トリップを共有する人 x , y は頻繁に同時移動する二人の組合せとみなされる． 1. この式の $IT^{w}(x, y)$ は，時間ウィンドウ幅 w でのユーザー x と y の間の同一の移動の有無 (1,0) を意味している．

$$P_{id}^{t_{id},w}(x,y) \quad = \begin{cases} 1 & n(IT_{(x,y),w}) \geq t_{id} \\ 0 & \text{else} \end{cases} \tag{1}$$

同一トリップ検出法では，同じ通勤ルートを共有するなどの見知らぬ同乗者が二人の組合せとして検出される可能性がある．仮に十分に大きい t_{id} の値をとったとしても，毎日同じルートの時間厳守の通勤者たちの間に人間関係があると判断してしまう．そのことを避けるため，十分に長期間にわたるデータがある場合に人間関係のある人々は多くの種類の移動を一緒にすると仮定し，少なくとも t_{ur} の異なる種類の目的地/出発地の組合わせで同一トリップを行う二人の組合せを検出する「ユニークルート検出法」を提案した．より正式には，少なくとも t_{ur} の目的地/出発駅のセットで同一トリップを行うユーザー x, y は，人間関係がある二人の組合せと見なす．(式 2 を参照)．この式で，$ITS_{(x,\ y),\ w}$ は，ユーザー x と y が時間ウィンドウ幅 w で同一トリップになる目的地と出発地の駅の組 (順不同) を意味する．

$$P_{ur}^{t_{ur},w}(x,y) = \begin{cases} 1 & n(\{ITS_{(x,y),w}\}) \geq t_{ur} \\ 0 & \text{else} \end{cases} \tag{2}$$

4. Results

4.1 人間関係の抽出

はじめに，頻繁に同時移動する二人の組合せの存在確率を考察する．図 1(a) は，時間ウィンドウ幅 w を変化させながら，(a) 同一トリップ手法で t_{id} を変えながら検出された頻繁に同時移動する二人の組合せ数 N_{id}，(b) ユニークルート手法で t_{ur} を変えながら検出された頻繁に同時移動する二人の組合せ数 N_{ur} を示している．図 (a) に着目すると $w = 1$ では，t_{id} を増やしたとき N_{id} にべき乗則に近く減少する．(a) の他の条件や (b) ではそれよりも若干急峻に減少する．一人一人がランダムに移動をすると仮定すると t_{id} や t_{ur} が大きくなったときの N_{ur} や N_{id} の減少は指数関数的であり，条件下の実測値のそれよりもはるかに緩く減少している．したがって，頻繁に同時移動する二人の組合せが何かしらの要因で数多く存在していると考えられる．

頻繁に同時移動する二人の組合せが人間関係がある人同士なのか通勤などによる見知らぬ同乗者であるかを考察するために，検出された二人の組合せの数に対する時間ウィンドウ幅 w の影響を考察した．見知らぬ同乗者の二人の組合せは列車に乗ったり降車したりする時間お互いに調整して合わせることはないため，w が大きくなるにつれて検出数は大幅に増加すると考えられる．一方，人間関係のある二人の組合せの検出数は w の変化に大きく影響を受けないと考えられる．彼/彼女らが同時に移動する場合，多くの場合はわずか数秒以内に改札を通過するためである．したがって，検出された二人の組合せから見知らぬ同乗者がすでに除外されている場合，時間ウィンドウ幅 w が大きくなっても検出される二人の組合せ数はそれほど増えないと考えられる．

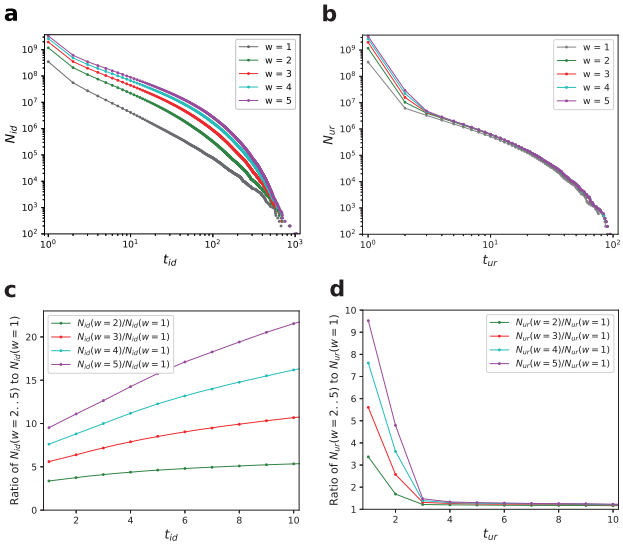


図 1: **(a，b) 同一トリップ検出法 (a) とユニークルート検出法 (b) で検出された，頻繁に同時移動する二人の組合せの数．**横軸は，同一トリップ検出法の各しきい値 t_{id}，t_{ur}，縦軸は頻繁に同時移動する二人の組合せの数を示す．各図では，時間ウィンドウ幅 w を 1..5 に変えた結果を異なる線でプロット．**(c，d) $w = 2..5$ における頻繁に同時移動する二人の組合せの数の $w = 1$ の場合に対する割合**

図 1(a，c) より，同一トリップ手法では時間ウィンドウ幅 w の増加にともなって検出された頻繁に同時移動する二人の組合せ N_{id} の数が大幅に増加することがわかる．このことにより，同一のトリップ方法によって検出された検出された二人の組合せには，見知らぬ同乗者が多く含まれていると考えられる．その一方で，ユニークルート法の場合，w の変化に伴う検出二人の組合せ数 N_{ur} の増加は，$t_{ur} \geq 3$ ではそれほど大きくはない (図 1) (b，d))．この場合，検出された二人の組合せには見知らぬ同乗者がほとんど含まれていないと考えられる．言い換えれば，これらの検出された対の間の移動の時間差は短い時間に収束し，各人は正確に同じ時間に移動するという強い動機を持っていると推測される．

また，$t_{ur} \geq 3$ においても $w = 2$ での二人の組合せ検出数が $w = 1$ よりも 2 割ほど大きい．しかし，$w = 3, 4, 5$ と w を増やしたときの増加は少ない．この理由は，友人関係があるユーザーが同時移動するときに改札を通る時間が記録単位である分をまたがることが一定の確率で存在するからだと考えられる．したがって,$w = 1$ ではなく $w = 2$ を適切なパラメータと考える．

以上より，$t_{ur} \geq 3$ のユニークルート法で頻繁に同時移動する二人の組合せのほとんどが，意図的に同一トリップをする人間関係を持っている二人の組合せだと推測できる．これらの理

由から，本論文ではパラメータ $t_{ur} = 3, w = 2$ を指定した一意のルート方法を使用する．

4.2 年齢性別による人間関係の密度

どのようなタイプの人々の間に人間関係が存在する確率が高いかを確認するために，グループ間の関係の相対密度 $\bar{D}(i,j)$ を図 2 にプロットした．$\bar{D}$ は全ユーザー間のリンク密度 D に対する，対象とするグループ間のグループ密度 $D(i,j)$ の割合である．

同年齢のグループの組合わせの間には，多くの人間関係が存在する． 10 代から 20 代の人々は同じ性別のグループを形成する傾向があるが，30 歳以上の人々は同じ年齢の男性/女性グループを形成する．この傾向は人間の生活を反映している：人は彼/彼女の友達と育ち，後でパートナーと家族を持っている．これらの結果は一般的に想像できる予想と一致している．ただし，性別および老齢学の観点から，人間関係密度の結果から考えられる重要な点は以下である．

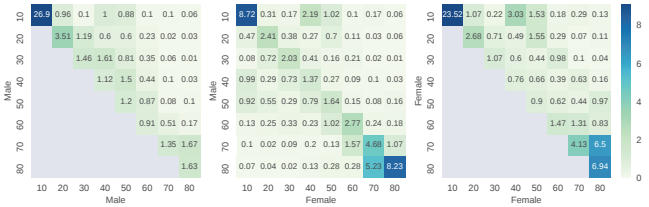


図 2: $\bar{D}$: **性別・年齢別のグループ間の関係の相対密度** 左から，男性/男性，男性/女性，女性/女性の年代別のグループの関係の相対密度 $\bar{D}$ をプロットした．

1. 40〜50 歳の年齢層の女性と 10 代の若者の人間関係の密度は，40〜50 歳の年齢層の男性と 10 代の若者のほぼ 2 倍である．これは，父親が母親ほど子供の世話をすることに関わっていないという事実を反映している．

2. 高齢者 (60 歳以上) では，男性/男性の人間関係は，女性/男性および女性/女性の人間関係と比較してまれである．この傾向は，年配の男性の方が同行する友達が少ないという事実を反映している

3. 年配の男性と若い女性の間の人間関係は，若い男性と年配の女性の間のそれらよりも多く観察される．これは男性と女性の関係の非対称性を反映している．

パラメータ t_{ur} の値を 2 から 10 に変更しても同様の傾向の結果が得られた．これらの結果は，日本の日常生活の活動における著しい性差を強調しており，日本におけるジェンダーによる役割の差を示した研究や報告 [Schwab 16] を裏付けるものである．また，高齢男性が同性の友人と移動をすることはあまりないという結果は高齢男性の孤立の証拠である．これは，老齢学の専門家によるアンケート調査に基づく観察と一致している [Beach 14]．このように，大量の移動データを調査することによって，社会学および老年学の分野で提案された仮説を大規模なデータから確認したといえる．

4.3 グループ移動の特性

大都市における個人の移動は日々の日課により条件付けられている． [Wang 15, Gonzalez 08, Song 10]．ここでは，人間関係が人間の移動における他の決定要因であることを実証するために，はじめにグループ移動の非日常な性質を考察する．グループ移動とは人間関係が推測できる二人の組合せと一緒に移動した移動である．両側グループ移動とはグループ移動のうち乗車駅も降車駅が同じ移動，片側グループ移動とは乗車駅か降車駅のどちらかのみが同じ移動をさす．移動の戻り時間 (すなわち，同じ駅への最初の到着時間とその次の到着時間との間の差) は，その移動の非日常性を性質を表す．移動の戻り時間が 1 日〜数日と短ければ，その移動は日常的な移動と推測される．また，グループ移動におけるある駅への戻り時間の計測には，対象者のその後の全移動を使用していおり，グループ移動だけに限ってはいない．また，ユーザーが同じ日に駅へ何度か移動をする場合には，それらを単一の移動と見なす．戻り時間の（累積）確率分布は図 3 にプロットした．

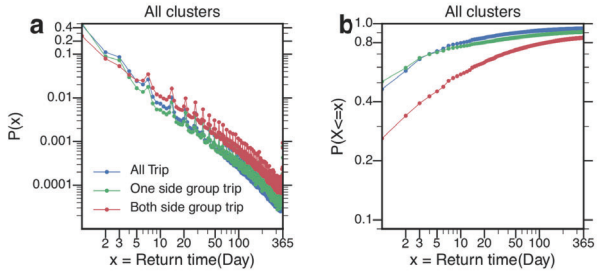


図 3: **(累積) 戻り時間の確率分布．** (a)：全移動の戻り時間の確率分布． (b)：全移動の累積確率分布． 各図において，両側グループ移動 (赤)，片側グループ移動 (緑)，全移動 (青) がプロットされている．

図 3 より，戻り時間の確率分布はべき乗則に従っていることが分かる．この結果は先行研究 [Song 10] の結果と一致している．べき乗則に比べ，7 の倍数の日付と 3 日後の戻り時間の確率は高い値をとり，これは人間の週次の行動や通勤・通学における土日休みのスケジュールを反映しているといえる．

図 3(a,b)，帰路時間の累積確率分布における全移動とグループ移動の違いを明確に示している．全体の移動の 43%(図の青線 3(b)) では翌日に同じ駅に戻るのに対し，両側グループ移動はわずか 26%(図の赤線) である．したがって，両側のグループ移動非日常のものであると考察される．この結果は人間関係が非日常的な移動をもたらすことを示している．

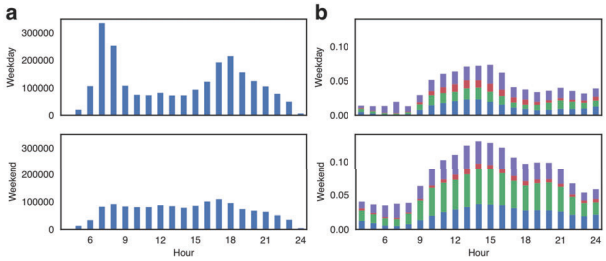


図 4: **(a) 時間別移動数**平日と週末 (休日を含む) で計算します **(b) グループ移動の時間比率**各図で，平日/週末の値を上下にプロット．

人間関係のある人々との同時移動が非日常的な移動である確率が高いことの要因を考察するために，いつ，どのような場所へ知り合いとで一緒に移動するのかを調べた．ユニークルート手法では，全移動の 9.7%がグループ移動として識別され，これらは特定の時間に発生する傾向がありた．図 4(b) に示すように，両側のグループ移動の割合は平日の日中と週末に増加した．これらの時間には一人での移動が多い通勤が少ないためだと考えられる．また，グループ移動の平均時間は 25 分程と，全移動の平均時間の 28 分より短い結果となった．これは，グループ/二人の組合せの意思決定プロセスの中で，人々が長旅のための高コストを避けていることを示唆している．

4.4 グループ移動の時間空間的な集中

以上の分析でグループ移動は非日常的な性質を持っていることが分かった．ここでは，グループ移動先が特定の時刻に特定の駅に時間空間的に集中していることを示す．大きな競技場の近くの駅 (図 5(a)) と他の駅 (図 5(b,c) のグループ移動数と全移動数の時系列を観察した．図 5(a) では，いくつかの明確なピークが見て取れるが，これらはスタジアムでイベントが開催された日とほぼ正確に一致している．一方で，1 日の総移動数のピークはイベントが発生した人との相関は比較的少ない．

また，都市の住宅街駅 (b,c) ではグループ移動数を観察することによって，2 月上旬 (b) と 12 月中旬 (c) にそれぞれイベントが発生したことがわかる．

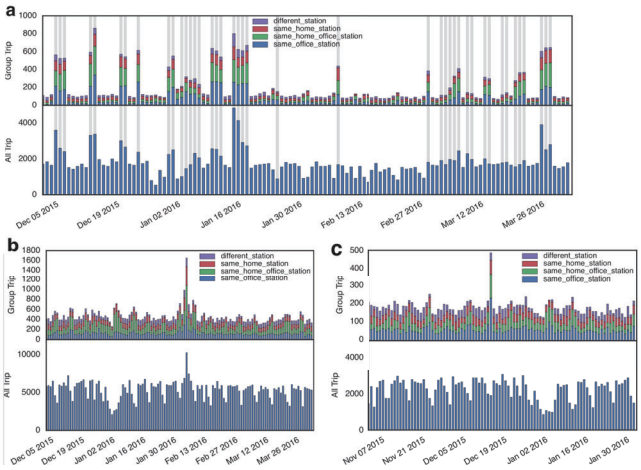


図 5: **1 日のグループ移動数と総移動数．** 1 日のグループ移動数 (上) と総移動数 (下) をプロット．(a) イベント会場の近くの駅．実際にイベントがひらかれた日を灰色でハイライトしている． (b，c) 住宅地にある駅．

これらの観察から，モビリティを介した人間の相互作用の一時的な収束が存在する．したがって，友達との移動はイベント検出の信頼できる指標である．つまり，ある駅における友人と一緒に降りる人の割合は，駅の近くでイベントが発生していることを示す一つの手法として有効である．

5. 考察

本稿では，大規模モビリティデータから人間関係を抽出する手法を提案した．提案手法において多様な種類の同一移動をとる人々が二人の組合せとして検出される．実際のデータを分析したところ，検出された二人の組合せには見知らぬ同乗者はほとんど含まれていないと推定された．このアプローチは人間の相互作用をデータから推測することに応用可能だと考えらる．鉄道移動データに加えて，記録される情報の量が増加すると，潜在的に人間の相互作用に関する情報，例えばＧＰＳログデータ，ウェブサーバの活動や購入ログが蓄積されている．いくつかの研究では，ソーシャルネットワークや画像共有サービスに基づいて，ユーザー間の関係を推定しようとしている [Cheung 15]．本研究で提案した，同じ活動グループの人々を誤って含まずに友人を検出する方法は，これらのデータが十分に長い期間にわたって記録されていれば，他のデータセットにも適用可能と考えられる．

以前のモビリティに関する研究では，人間-場所の関係性を考慮することにより人間の移動を理解・予測してきた [Wang 15]．視点を変えて考えると，多くの移動は友人や会社・学校の人と合うことを目的として行われるので，人間の移動は人間-人間の関係をベースに理解することも可能である．本稿では，人間関係がモビリティのもう一つの決定要因であることを確認した．我々の結果が人間の移動性のある基本的な側面を明らかにするためのエージェントベースのアプローチに貢献することを期待している．

さらに，移動から推測された人間関係の密度は性別や年齢によって異なり，性差による日常生活の違いがデータから読み取れた．例えば，年配の女性と比較して，年配の男性が互いに交流することはめったにない．年配の男性の孤立は私たちの社会の中で重要な課題である [Beach 14]．これに加えて，30〜50 歳の女性と若者の間に存在する関係の確率は，30〜50 歳の男性と若者のそれとは有意に異なる．中年の男性が中年の女性より若い人と移動する可能性が低いという結果は，子育てにおける性差を表している．この結果は，政策決定やイベント計画に適用可能な証拠となりうるだろう．

参考文献

[Asatani 18] Asatani, K., Toriumi, F., Mori, J., Ochi, M., and Sakata, I.: Detecting interpersonal relationships in large-scale railway trip data, *Journal of Computational Social Science*, Vol. 1, No. 2, pp. 313–326 (2018)

[Beach 14] Beach, B. and Bamford, S.: Isolation: the emerging crisis for older men, *London: Independent Age* (2014)

[Bogomolov 14] Bogomolov, A., Lepri, B., Staiano, J., Oliver, N., Pianesi, F., and Pentland, A.: Once upon a crime: towards crime prediction from demographics and mobile data, in *ACM ICMI*, pp. 427–434ACM (2014)

[Cheung 15] Cheung, M., She, J., and Jie, Z.: Connection discovery using big data of user-shared images in social media, *IEEE Transactions on Multimedia*, Vol. 17, No. 9 (2015)

[De Domenico 13] De Domenico, M., Lima, A., and Musolesi, M.: Interdependence and predictability of human mobility and social interactions, *Pervasive and Mobile Computing*, Vol. 9, No. 6, pp. 798–807 (2013)

[Fan 17] Fan, C., Liu, Y., Huang, J., Rong, Z., and Zhou, T.: Correlation between social proximity and mobility similarity, *Scientific Reports*, Vol. 7, (2017)

[Giannotti 11] Giannotti, F. and Nanni, M. e.: Unveiling the complexity of human mobility by querying and mining massive trajectory data, *VLDB*, Vol. 20, No. 5, pp. 695–719 (2011)

[Gonzalez 08] Gonzalez, M. C., Hidalgo, C. A., and Barabasi, A.-L.: Understanding individual human mobility patterns, *nature*, Vol. 453, No. 7196, p. 779 (2008)

[Grauwin 17] Grauwin, S. and Szell, M. e.: Identifying and modeling the structural discontinuities of human interactions, *Scientific Reports*, Vol. 7, (2017)

[Miller 05] Miller, H. J.: Necessary space—time conditions for human interaction, *Environment and Planning B: Planning and Design*, Vol. 32, No. 3, pp. 381–401 (2005)

[Newman 12] Newman, P. and Matan, A.: Human mobility and human health, *Current Opinion in Environmental Sustainability*, Vol. 4, No. 4, pp. 420–426 (2012)

[Pflieger 10] Pflieger, G., Rozenblat, C., Mok, D., Wellman, B., and Carrasco, J.: Does distance matter in the age of the Internet?, *Urban Studies*, Vol. 47, No. 13, pp. 2747–2783 (2010)

[Schwab 16] Schwab, K., Samans, R., and Zahidi, e.: The Global Gender Gap Report 2016, in *World Economic Forum* (2016)

[Shimosaka 15] Shimosaka, M., Maeda, K., Tsukiji, T., and Tsubouchi, K.: Forecasting urban dynamics with mobility logs by bilinear Poisson regression, in *Proceedings of the UbiComp 2015*, pp. 535–546ACM (2015)

[Song 10] Song, C., Qu, Z., Blumm, N., and Barabási, A.-L.: Limits of predictability in human mobility, *Science*, Vol. 327, No. 5968, pp. 1018–1021 (2010)

[Wang 15] Wang, Y. and Yuan, N. J. e.: Regularity and conformity: Location prediction using heterogeneous mobility data, in *SIGKDD*, pp. 1275–1284ACM (2015)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

豪雨による支払件数を予測する被害想定手法の提案

Proposal of a method to estimate insurance claims caused by heavy rain

岡崎 豪[*1]
Takeshi Okazaki

レブル オリオル[*1]
Oriol Gaspa Rebull

[*1] エーオンベンフィールドジャパン株式会社
Aon Benfield Japan Ltd.

In order to facilitate flood and landslide damage assessment, we propose an evaluation method that can predict the total number of damaged houses with high accuracy immediately after a disaster. To improve accuracy, we narrowed the exposure down by using low-lying and landslide high-risk areas and apply the precipitation of 15-year return period as a threshold to the original precipitation to take account of disaster preparedness by each local government. When applied to two heavy rain events occurred in 2017 and 2018 and validated the method, we confirmed that it can be well reproduced in any prefecture.

1. はじめに

1.1 目的

2018 年は平成 30 年 7 月豪雨に始まり，台風第 21 号を含めた 5 つの台風の上陸，大阪北部地震と北海道胆振東部地震が発生するなどの多くの災害に見舞われた年であった．これら一連の自然災害による損害保険業界の支払件数は約 138 万件となり，東日本大震災による地震保険の支払件数の約 1.6 倍の規模に達した[日経新聞 18]．損害保険会社や共済団体（以下，まとめて損保業界と呼ぶ）はこれら支払いのため，膨大な人員を投入した．被災地の社員だけでは対応が困難であることから，現地などに対策本部を設置し，他県の社員や損害保険鑑定人が被災地に応援要員として向かった．これら対応の意思決定は即座に行う必要があることから，発災直後に被害の規模を推定する被害想定の重要性が高まっている．被害の規模が大きくなるほど過去の経験からの推定が難しくなるが，意思決定のためにはそれを可能な限り正確に予測することが求められる．筆者らはこの被害想定のため，台風被害想定システムと呼ぶシステムを開発した[岡崎 17]．しかし，同システムは台風による強風の被害を対象としており，洪水や土砂災害などは対象としていない．そこで，本研究ではそれら水災の被害想定を可能とするシステムの開発に取り組んだ．

1.2 評価手法の課題

本研究の評価対象とする水災は，中小規模の外水氾濫，内水氾濫，及び土砂災害とする．一般に外水氾濫を評価する場合，降水量を入力データとした流出解析により河川の流量を評価し，破堤地点を選定したのち，2 次元不定流モデルによる氾濫解析により氾濫域を評価するプロセスとなる．内水氾濫を評価する場合は，流出解析と氾濫解析，そして下水道などの排水能力を考慮する管渠内解析から評価するプロセスとなる．国土交通省によるハザードマップの作成指針においても同様の手法が示されている[国土交通省 15]，[国土交通省 10]．しかし，前者の同手法には河道のデータ，後者には下水道などの排水路のデータが必要であり，それらデータが整備されていない河川や地域には適用できない．平成 29 年九州北部豪雨や平成 30 年 7 月豪雨においては，河道データが整備されていない 2 級河川の氾濫や内水氾濫が各地で多発し，累積で大きな被害となっている．データが整備されていない河川や地域は多く，現状では，それら被害を前述の手法から評価することは困難である．そこで，降水量から被害を直接的に評価する統計的なアプローチで取り組んだ．

図 1 に平成 30 年 7 月豪雨の罹災率の分布を示す．罹災率とは火災保険の契約件数を分母とし，水災による支払いが発生した件数を分子としたときの比率であり，それを市区町村別に表示したものである．広島県，岡山県，福岡県，愛媛県で大きな罹災率となっていることが確認できる．一方，図 2 に平成 30 年 7 月豪雨が発生した期間中の日降水量の最大値の分布を示す．図 1 と図 2 を比較すると，岡山県と広島県の罹災率は高いが，降水量は大きい値ではない．一方，高知県の罹災率は低いが，降水量は大きい値となっている．これら結果から，降水量と罹災率との相関は高くないことが分かる．

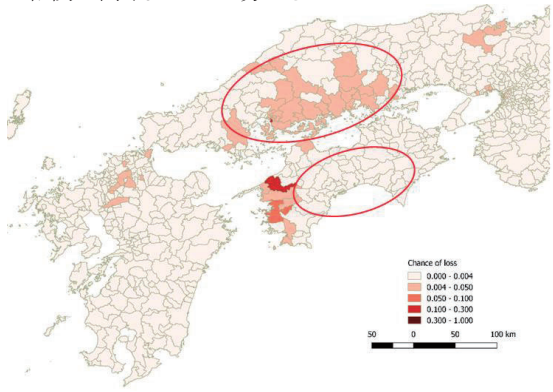


図 1　平成 30 年 7 月豪雨による罹災率の分布

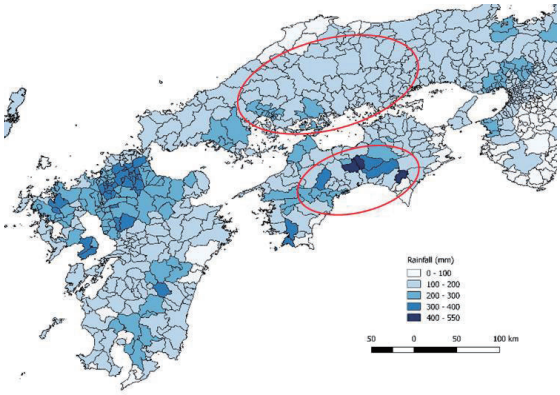


図 2　平成 30 年 7 月豪雨による日降水量の最大値の分布

連絡先：岡崎 豪，takeshi.okazaki@aon.com，エーオンベンフィールドジャパン株式会社，03-4589-4237，東京都千代田区永田町 2-10-3 キャピトルタワー11 階

これら図のように，従来は降水量から被害を直接的に評価することは難しいと考えられてきた．本研究ではこの課題に対して，2つの手法を適用することで解決を図った．一つは洪水や土砂災害が発生する危険性の高い地域の空間情報を用いて，母集団に修正を加える手法である．もう一つは降水量に閾値を適用することで，各自治体の災害対策の効果を反映させる手法である．以下，それぞれの手法を述べたのち，平成 29 年九州北部豪雨と平成 30 年 7 月豪雨に対して適用したときの検証結果を示す．

2. 被害想定の手法

罹災率と降水量との相関を高くするため，母集団と降水量のそれぞれに変更を加えた．以下，それぞれの手法を示す．

2.1 母集団の変更

図 1 はすべての契約データを分母として罹災率を評価したが，それら母集団には高台などの水災リスクが低い契約も含まれる．水災リスクが低い契約データを母集団から除くことで，罹災率と降水量との相関が高くなると考えられることから，土砂災害危険箇所と低位地帯と呼ばれるデータを用いて，母集団に変更を加えた．土砂災害危険箇所とは，都道府県が指定する土石流危険渓流，地すべり危険箇所，急傾斜地崩壊危険箇所及び雪崩危険箇所の範囲をベクターフォーマットでデータ化したものであり，低位地帯は周辺部よりも標高が低く，排水が困難である地帯をデータ化したものである．いずれも国土交通省が website で公表・配布している[国土交通省 10]．

図 3 の左図は日本のすべての建物を表示したものであり，この分布に土砂災害危険箇所を重ね合わせ，土砂災害危険箇所に含まれる建物のみを抽出した結果が図 3 の右上である．全国の建物約 4,915 万棟に対して，土砂災害危険箇所に含まれる建物は約 370 万棟であり，8%の割合に相当する．同様に低位地帯に含まれる建物のみを抽出した結果を図 3 の右下に示す．低位地帯に含まれる建物は約 590 万棟であり，12%の割合に相当する．土砂災害危険箇所の総面積は，低位地帯よりも広いが，低位地帯は人口密集地の沿岸部や河口部に多いため，建物数の割合は低位地帯の方が多い結果となる．母集団の変更においては，これら土砂災害危険箇所と低位地帯のいずれかの領域に含まれる契約を対象とした．

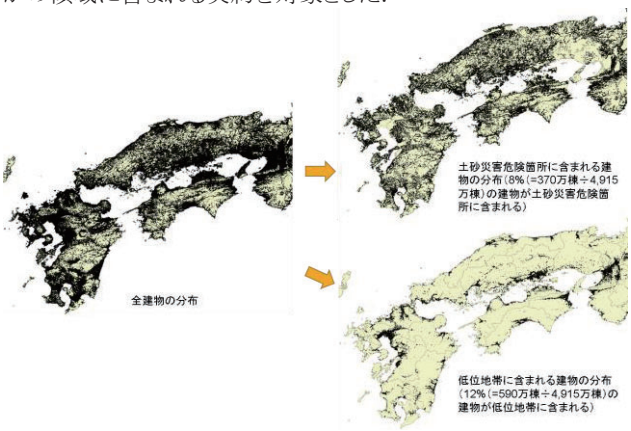


図 3 土砂災害危険箇所と低位地帯を用いた母集団の変更

2.2 降水量データの変更

高知県，宮崎県，徳島県などでは毎年のように大雨が降るが，必ずしも大きな被害とはならない．これは地方整備局や自治体が長年に渡ってその大雨に対応した河川整備や土砂災害対策を施してきたためである．そのため，たとえ大雨が降っても，その地域で発生しやすい規模であれば，相応の対策が施されていると考えられる．本研究では，この地域ごとの対策の度合いを降水量の閾値として反映させた．すなわち，降水量に閾値を適用し，降水量が閾値以上であれば被害が大きくなる手法とした．閾値の計算のため，まず気象庁から日本の全観測地点の 1976～2018 年（43 年間）の降水量を日別に取得した．次に，日降水量の各年の最大値を算出し，それらプロットに一般化極値分布（GEV）と対数正規分布を適用した．適用後，最小二乗法により地点毎にいずれかの分布を選定した．そして，それら分布から再現期間 15 年の日降水量を算出し，閾値とした．再現期間には 10, 15, 20, 25, 30, 35, 40, 45, 50 年の 9 パターンを適用した後，罹災率と最も相関が高かった 15 年を選定した．

母集団に変更を加えた後の罹災率の分布を図 4，降水量に閾値を適用した後の値の分布を図 5 に示す．これらと図 1, 2 を比較すると，広島県，岡山県，福岡県は罹災率が高くなっており，降水量も相対的に大きくなっていることが確認できる．また，高知県は罹災率が低く，降水量も小さくなっている．以上から罹災率と降水量との相関が向上したことが確認できる．

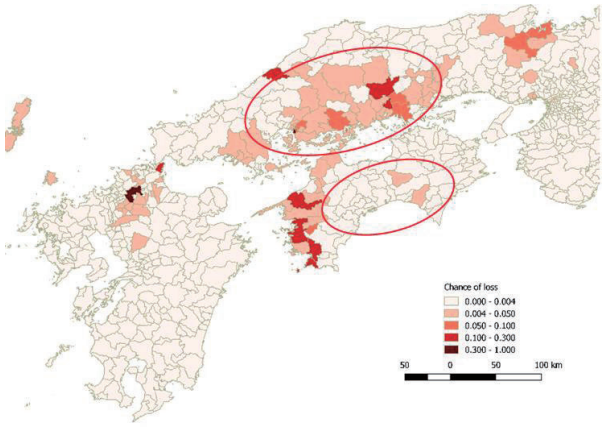


図 4 母集団を変更した後の罹災率の分布

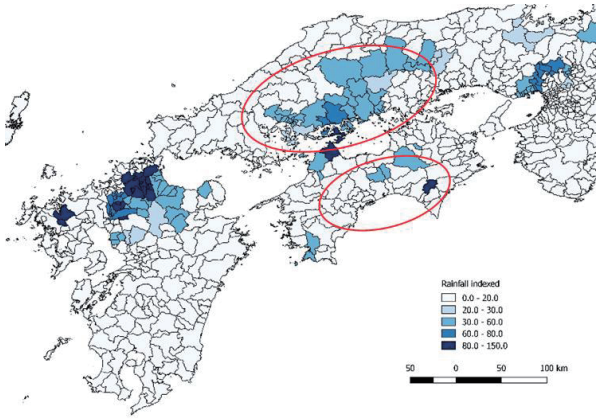


図 5 閾値を適用した後の降水量の分布

2.3 罹災率曲線の構築

母集団と降水量に変更を加えた後のデータを用いて，罹災率曲線を構築した．罹災率曲線は建物構造別，保険の目的別（建物，家財）に構築した．住宅物件の木造を例とした結果を図 6 に示す．横軸は閾値を適用した後の降水量，縦軸は母集団に変更を加えた後の罹災率である．プロットは市区町村毎の値である．近似曲線には，複数の曲線を用いたのち，最も再現性の高い結果が得られた決定木分析（Decision Tree）を適用した．

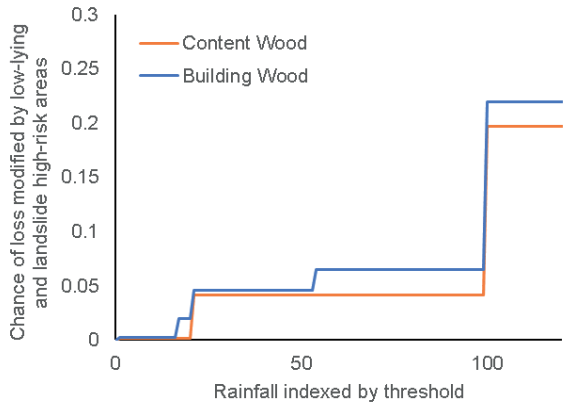


図 6　構築した罹災率曲線（木造建物を対象とした例）

3.　検証

本手法を平成 30 年 7 月豪雨と平成 29 年 7 月九州北部豪雨に適用し，その予測誤差を検証した．以下，比較結果と今後の課題を示す．

3.1　比較結果

平成 30 年 7 月豪雨と平成 29 年 7 月九州北部豪雨を対象として，実際の支払件数（以下，実績件数と呼ぶ）と本研究の手法により再現した件数（以下，予測件数と呼ぶ）を比較した．図 7 に平成 30 年 7 月豪雨の全国計の比較結果を示す．図 8 に平成 29 年 7 月九州北部豪雨の全国計の比較結果を示す．これら図は風災，水災，大規模な外水氾濫の内訳を示している．実績件数の内訳において，風災とは強風による支払件数を意味する．これら豪雨において，支払保険金は水災が大きな割合を占めたが，件数は風災が大きな割合を占めた．

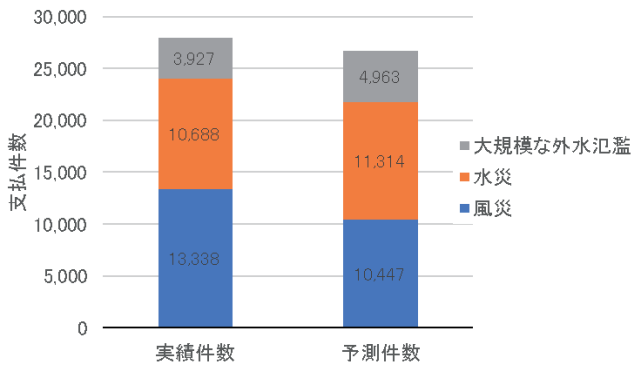


図 7　平成 30 年 7 月豪雨における実績と予測の比較

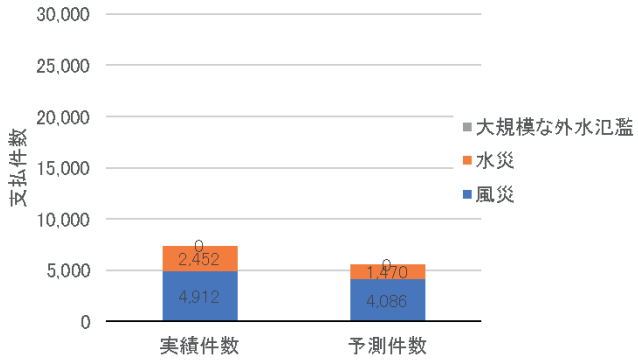


図 8　平成 29 年 7 月九州北部豪雨における実績と予測の比較

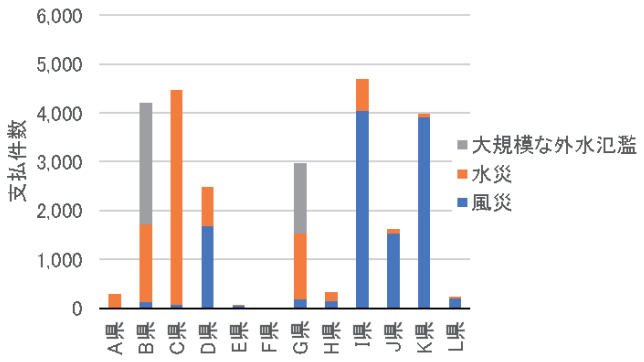


図 9　平成 30 年 7 月豪雨の都道府県別の実績件数

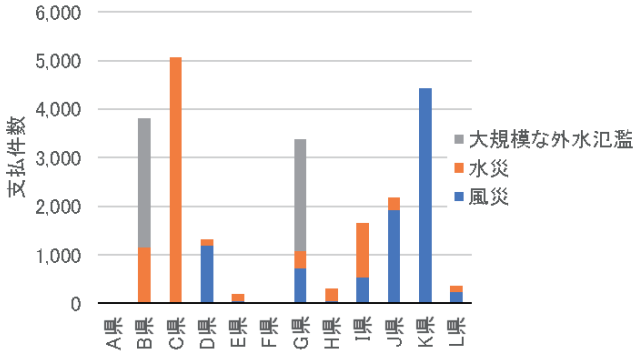


図 10　平成 30 年 7 月豪雨の都道府県別の予測件数

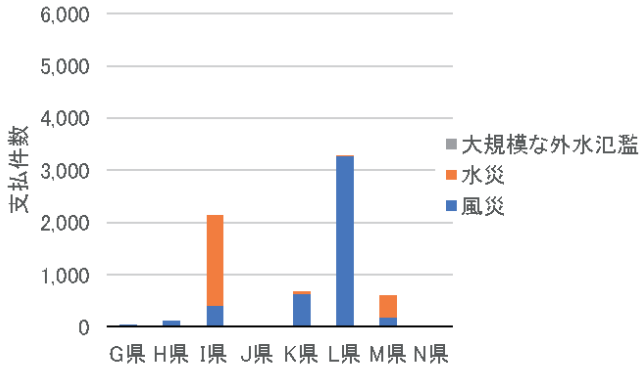


図 11　平成 29 年 7 月九州北部豪雨の都道府県別の実績件数

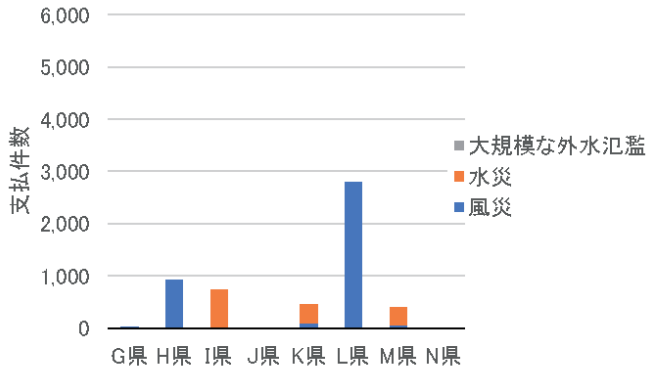


図 12　平成 29 年 7 月九州北部豪雨の都道府県別の予測件数

実績件数の水災とは中小規模の外水氾濫，内水氾濫，土砂災害による支払件数を意味する．ただし，一級水系の本川やそ

れに準じる支川が氾濫するような大規模な外水氾濫による件数は含めていない．平成 30 年 7 月豪雨では，岡山県倉敷市の小田川と愛媛県大洲市の肱川が氾濫しており，それら市で発生した件数は「大規模な外水氾濫」の区分とした．

予測件数の内訳において，風災は前述の台風被害想定システムを用いて評価した結果である．水災は本研究で提案した手法から評価した結果である．大規模な外水氾濫は，国土地理院による氾濫域図を用いて評価した結果である．図 7 と 8 の結果から，予測件数は実績件数をその内訳を含めて，よく再現できていることが確認できる．

次に都道府県別の内訳を示す．図 9 は平成 30 年 7 月豪雨の実績件数，図 10 は同豪雨の予測件数，図 11 は平成 29 年 7 月九州北部豪雨の実績件数，図 12 は同豪雨の予測件数である．いずれの県においても予測件数は実績件数に近い結果となっており，都道府県別の内訳でもよく再現できていることが確認できる．一部，風災の件数が過小評価となっている県があるが，それらは今後に罹災率曲線を見直す予定である．以上から，母集団に変更を加え，また降水量に閾値を適用することで，降水量から支払件数を高い精度で予測できることを確認した．

3.2 今後の課題

被害想定の精度を向上させるための大きな課題は，気象庁の観測所が少ない点である．現状の降水量を観測する AMeDAS の観測地点数は約 1,300 である．市区町村の区分が約 1,700 あることから，市区町村の区分よりも少ない設置数である．市区町村よりも細かい区分で被害想定を行うためには，局地的な降水量の情報が必要となる．しかし，気象庁の観測機器には 5 年に一度再検定が必要であり，その再検定を含めた維持費用が高いことから，設置数が急に増える見込みはない．この課題を解決する手法として，被災地での写真や映像などにディープラーニングによる画像判別を適用し，降水量や被害の度合いを推定する方法がある．2018 年に岡山県倉敷市で洪水が発生した際，国土地理院は一般人が投稿した画像を参考として浸水の水際の位置を特定している[朝日新聞 18]．この事例のように写真や動画の活用は有効であり，被害想定にも応用できると考える．

4. まとめ

豪雨が発生したときの損保業界の意思決定を支援することを目的として，発災直後に高い精度で被災状況を予測できる水災被害想定の手法を提案した．精度を向上させるため，母集団と降水量に変更を加えた．母集団の変更においては，土砂災害危険箇所と低位地帯のいずれかの領域に含まれる契約を対象とした．降水量の変更においては，再現期間 15 年の日降水量を閾値として適用し，地方整備局や各自治体による河川整備や土砂災害対策の度合いを反映させた．平成 30 年 7 月豪雨と平成 29 年 7 月九州北部豪雨を対象として，実際の支払件数と本研究の手法により予測件数を比較したところ，総支払件数，都道府県別の内訳においても，実績件数をよく再現できていることを確認した．

参考文献

[日経新聞 18] 日経新聞, 損害保険支払額, 今年度 1.3 兆円に 7 年ぶり高水準, 2018 年 12 月 21 日付, アクセス日時 2019 年 1 月 30 日
https://www.nikkei.com/article/DGKKZO39194410Q8A221C1EE9000/

[岡崎 17] 岡崎 豪, レブル オリオル, ディープラーニングによる建物特性の抽出と台風被害想定の手法, 人工知能学会全国大会論文集, 2017, JSAI2017 巻, 第 31 回全国大会(2017), セッション ID 3P1-NFC-00a-4, p. 3P1NFC00a4, 公開日 2018/07/30, https://www.jstage.jst.go.jp/article/pjsai/JSAI2017/0/JSAI2017_3P1NFC00a4/_article/-char/ja

[国土交通省 15] 国土交通省, 洪水浸水想定区域図作成マニュアル（第 4 版）, 2015,
www.mlit.go.jp/river/shishin_guideline/pdf/manual_kouzuishinsui_1507.pdf

[国土交通省 10] 国土交通省, 内水ハザードマップ作成の手引き（案）, 2010,
http://www.mlit.go.jp/report/press/city13_hh_000063.html

[国土交通省 10] 国土交通省, 国土数値情報, 2010,
http://nlftp.mlit.go.jp/ksj/index.html

[朝日新聞 18] 朝日新聞, 西日本豪雨の浸水域, 推定図に SNS 活用 国土地理院, アクセス日時 2019 年 1 月 30 日
https://www.asahi.com/articles/ASL8F3QMQL8FUJHB001.html

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

グローバル株所有ネットワークでの中国による間接支配の見える化

Detection of indirect dominance by China on the global shareholding network

水野　貴之*1　　小髙　充弘 1
Takayuki Mizuno　　Mitsuhiro Odaka

1 国立情報学研究所, 総合研究大学院大学複合科学研究科
National Institute of Informatics, School of Advanced Sciences, SOKENDAI (The Graduate University for Advanced Studies)

We investigate indirect dominance by China on the global shareholding network using a dataset that contains information on shareholding for about 300 million firms and 140 million people worldwide. First, we extend the percolation model on the network to detect the dominance relationships as a subgraph (sub-network) on the global shareholding network. This model decides whether to be able to dominate companies (to percolate nodes), not only by the direct ownership, but also indirect ownership through the dominated companies. Next, we observe world dominance by Chinese owners and American companies (investment companies and GAFA) by controlling the ownership ratio required for domination. When Chinese owners cooperate, world domination is greatly expanded. China dominates 47% of all worldwide employees by an indirect shareholding ratio over 1%. Chinese domination on Tax Haven areas, France, South Africa, Nigeria is prominent. France has many stocks of companies in Africa, the former colony. Domination has a hierarchical structure. That is, China dominates African companies through French companies. The mining industry is particularly under great domination. However, Chinese dominance of the world is strong, but dominance by American companies is even stronger.

1. はじめに

中国政府は一帯一路構想を立ち上げ，中国の国有企業や富豪らが，南太平洋，南シナ海，アジア全域，ロシア，アフリカ，欧州のインフラへの投資や政治家への資金提供を積極的におこない，大規模経済圏を構築しようとしている．近年，それらによる弊害が発生してきた．例えば，豪州では，資金提供を受けた政治家らが，南シナ海問題や自由貿易協定（FTA）などで，中国に望ましい方向に政策を誘導しているという事態が発生している [1]．また，融資が返せない途上国では，政治や資源に対する中国支配が強まっている [2]．このように，想定を超える中国マネーにより，各国の主権が弱まっている．

本稿では，グローバルで複雑な株所有ネットワークを介した中国による間接的な世界支配を検出するために，支配関係に拡張したパーコレーションモデルを提案する．このモデルに従い，中国と米国について，株主らの協力関係を反映した間接的な世界支配の強さについて示す．

中国は世界の全株式会社の従業員の 47%を，1%の間接的な株所有率により支配している．特に，タックスヘブン地域や，フランス，南アフリカ，ナイジェリアでの支配が目立つ．このように中国の支配は強まっているが，アメリカの支配は，その比ではない．米国の僅か4つの投資会社は，その 5 倍の所有比率で47%の従業員を支配している．本発表では，これらの結果に加えて，異なる期間のデータを用いて間接的支配の時間変化や，国会議員データを用いて政治との繋がりについて報告する．

2. グローバルな株所有ネットワークの作成

我々は，Bureau van Dijk 社が提供する全世界 3 億社と 1.4 億人が収録された 2016 年の企業情報データベースを用いる．株式の保有比率が記載された企業の株主リストと，企業や人の保有銘柄リストが含まれている．我々は，これらのリストを全ての企業に関して繋ぎ合わせることで，グローバルな株所有ネットワークを構築した．

連絡先：水野貴之，国立情報学研究所，東京都千代田区一ツ橋２－１－２．mizuno[at]nii.ac.jp

3. パーコレーションモデルによる間接支配の抽出

ネットワーク上のパーコレーションモデルを応用することで，中国が直接的・間接的に支配する企業群を，グローバル株所有ネットワークの部分グラフ（部分ネットワーク）として抽出する．

ネットワーク上のシンプルなパーコレーションモデルでは，各リンクに値が設定されている．そして，あるノードが，そのノードに接続されたリンクを経由して隣のノードに影響をあたえるかどうかは，そのリンクの値が設定した閾値以上であるのかだけに依存する．しかし，このシンプルなモデルでは支配関係を抽出できない．

図1は，ある3社間の株所有ネットワークである．B 社は A 社の株の 30%を，C 社の株の 80%を保有している．そして，C 社は A 社の株の 40%を保有している．支配に必要な閾値を 50%に設定して，B 社からのシンプルなパーコレーションモデルを実行すると，B 社は C 社のみを支配下に置く．しかしながら，現実は，B 社は支配下にある C 社の持つ A 社の 40%の株を利用することができ，B 社が持つ A 社の 30%の株とあわせて間接的に70%を保有することにより，A 社を支配できる．つまり，支配関係を抽出するためには，パーコレートの閾値を，B 社が直接所有する A 社の株を表すリンクに対して設定するのではなく， B 社が支配下企業を介して間接的に保有する A 社の株も合計して，そのリンクの値の合計に対して閾値を設定するパーコレーションモデルにする必要がある，この間接的な影響を加味したパーコレーションモデルは，反復計算を繰り返すことで，B 社が支配する全ての企業を抽出することができる．

図2は，支配に必要な閾値を 50%に設定したときの，間接的な影響を加味したパーコレーションモデルにより抽出された中国政府による企業支配を表す．中国政府は，複雑な株所有ネットワークを通じて多くの企業を支配下においていることが分かる．

4. 支配の条件と抽出結果

株式企業は，様々な所有率で株主に支配されている．例えば，株主は，所有率 2/3 以上で単独での取締役の解任や会社の合

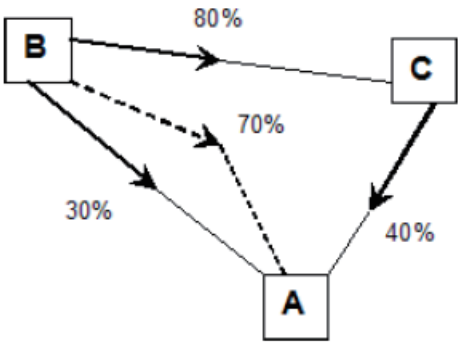


図1　ある3社間の株所有ネットワーク

図2　中国政府による直接的・間接的な企業支配

併解散を決めることができる．また，所有率 1/3 以上であれば，これらの決議について拒否権を発動できる．所有率 50%以上では，普通決議を単独で決めることができる．1%や 3%，10%でも会社を動かす様々な権利が得られる．このように，様々な種類の支配関係が存在する．

複数の株主が協力して会社を支配する場合もある．例えば，妻と夫が，それぞれ，会社 A の株を保有しており，それらの株を合わせることによって会社 A を支配する場合がある．中国政府が，中国人富豪に協力を要請することもある．

世界における企業サイズの分布は，ファットテール従っているから，多くの従業員は大企業で働いている [3]．我々は，支配における企業サイズの違いを考慮するために，正規従業員数で支配の規模を評価する．本データセットに含まれる株式企業で働く正規従業員の合計は約4億人である．我々は，図3で，中国政府，中国株主連合（全ての中国人株主と非株式組織），GAFA (Google, Apple, Facebook, Amazon とそれらのオーナー)，米国投資会社 (Vanguard, Berkshire Hathaway, Bank of America, Blackrock)について，それぞれ支配を調査した．

中国政府は，株式所有率 50%で，世界の全株式会社の従業員の 5%を支配している．中国株主が協力し合うことで，支配率は 8%まで上昇する．支配に必要な閾値を 1%以下に下げると

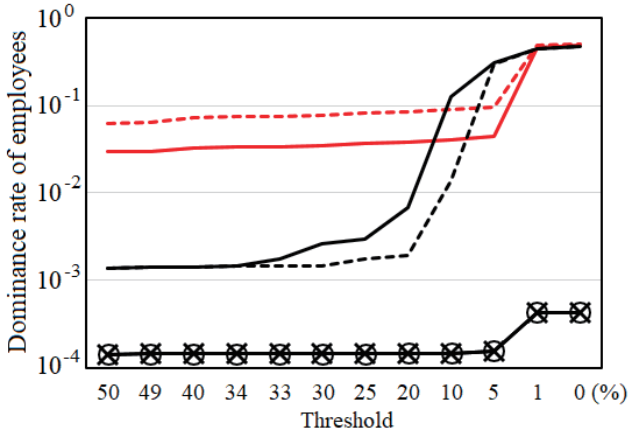


図3　中国政府などによる従業員数支配率．赤実線：中国政府，赤破線：中国株主連合，黒実線：米国投資会社（協力），黒破線：米国投資会社（非協力），○：GAFA（協力），×：GAFA（非協力）．

支配の規模を急激に広げ，中国政府と中国株主連合は，全従業員の 47%を支配する．米国による世界支配は中国の比ではない．アメリカの僅か4社の投資会社が協力するだけで，閾値 10%で中国を逆転する．そして，全世界の 40%の従業員を，株式所有率 5%で支配している．GAFA による株所有による企業支配は，それほど大きくはない．しかも，彼らは，支配企業を明確に棲み分けしており，協力をしたとしても，支配は拡大しない．GAFA は投資ではなくビジネス上の連携をおこなうために，企業支配をおっていると考えられる．

5.　まとめと課題

本稿では，グローバルで複雑な株所有ネットワークを介した中国による間接的な世界支配を検出するために，支配関係に拡張したパーコレーションモデルを提案した．中国は全世界の従業員の 47%を，1%の間接的な株所有率により支配している．しかし，米国企業による支配は中国の比ではない．

我々は，国別に中国による支配率を調査した．タックスヘブン地域や，フランス，南アフリカ，ナイジェリアでの支配が目立つ．フランスは旧植民地であるアフリカの企業の株を多く持っている．支配には階層構造があり，中国は，フランス企業を介してアフリカを支配している．アフリカでは，マイニング産業が大きな中国支配を受けている．

本講演では，これらの結果に合わせて，2009 年の株所有データで同じ解析をすることにより，中国の海外支配の時間変化についても報告する．

本研究は科学研究費補助金 16H05904 の支援を受けている．

参考文献

[1] C. Hamilton, A. Joske (2018) Silent Invasion: China's Influence in Australia. Hardie Grant.

[2] J. Hurley, S. Morris, G. Portelance (2018) Examining the Debt Implications of the Belt and Road Initiative from a Policy Perspective. CGD Policy Paper 121, 1-37.

[3] S. Fujimoto, A. Ishikawa, T. Mizuno, T. Watanabe (2011) A New Method for Measuring Tail Exponents of Firm Size Distributions. Economics 5, 2011-19.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

現代的/古典的レイシズムの表出におけるニュース番組の影響: インターネットテレビに投稿される差別的コメントの分析

Expressing Modern/Oldfashioned Racism depending on News Contents of Internet Television

高野雅典 *1,*2 Masanori Takano　高史明 *3 Fumiaki Taka　森下壮一郎 *1 Soichiro Morishita　西朋里 *4 Tomosato Nishi　小川祐樹 *4 Yuki Ogawa

*1株式会社サイバーエージェント CyberAgent, Inc.　*2東京大学 The University of Tokyo　*3神奈川大学 Kanagawa University

*4立命館大学情報理工学部 College of Information Science and Engineering, Ritsumeikan University

We analyzed comments to news programs on Internet television to explore relationships between news contents and modern/oldfashioned racism on expressing prejudice against Zainichi Koreans in Japan. We found that users tended to express both prejudices; nevertheless, it depended on news programs whether users expressed modern or oldfashioned racism.

1. はじめに

匿名のネット空間（e.g. Twitter）では現実の社会では受け入れられないような偏見が表出されることが少なくない [Munger 17, 高 15]．そのような表出されるレイシズムは古典的レイシズムと現代的レイシズムの 2 種類に分けられる [Kinder 81, McConahay 86]．古典的レイシズムとは，被差別民族は能力や倫理観が劣っているという信念に基づく偏見である．現代的レイシズムとは，差別はもはや解決しているにも関わらず被差別民族が恵まれていなかったり権利を求めるのはただの怠惰であり，本人の責任であるという信念に基づく偏見である．この 2 つは元々はアメリカにおける黒人に対する偏見を理解するために考案された枠組みであるが，日本における在日コリアンに対する偏見も同じ枠組みで理解できることが示されている [高 15]．

在日コリアンに対するオンラインでの差別発言はニュース記事をきっかけとすることが多い [高 15]．本研究では在日コリアンに対する差別発言において古典的レイシズムと現代的レイシズムの表出とニュースコンテンツの関係を明らかにすることが目的である．そのために我々はインターネットテレビ局 AbemaTV*1 におけるニュースチャンネルの各番組につけられた在日コリアンに関係するコメントを分析する．AbemaTV には約 20 の各種専門チャンネル（ニュース，スポーツなど）があり，番組表にしたがってリニア放送がパーソナルコンピュータやタブレット，スマートフォンを使って視聴できる Web サービスである．放送されるコンテンツやユーザ体験はテレビに近い．ユーザは放映中の番組についてリアルタイムにコメントをつけることができる．本研究はこのコメントを分析する．

2. データとモデル

以下の手順で現代的/古典的レイシズムを含む差別発言を抽出した．

1. 2018/1/17〜2018/8/31 に AbemaTV のニュースチャンネルに投稿されたコメントからコリアン関連のコメントを抽出（基準は表 1 を参照 *2）
2. [高 15] の分類コードを拡張したもの（表 2）を用いて各コメントに現代的/古典的レイシズムのラベルを付与（両方含まれていなければ付与しない）
3. 現代的/古典的レイシズムの両方のラベルが付与されたコメントが少なくとも 1 回以上出現する番組に対するコメントのみ抽出
4. 現代的/古典的レイシズムの両方のラベルが付与されたコメントを少なくとも 1 回以上投稿し，かつ，2 つ以上の番組を視聴したユーザのコメントのみ抽出

表 1: コリアン関連のコメントを抽出した条件．これらの単語をコメントが含んでいればコリアン関連のコメントとした．コリアン関連の名詞とインターネット空間でしばしば使われる差別的なスラングを用いた．* はワイルドカードである．ただしコリアン以外の在日外国人に関する単語や韓国人俳優名（ユチョン氏など）がコメントに含まれる場合は除去した．

コリアン関連コメントの抽出基準単語リスト
韓国人　朝鮮人 在日　チョン　キムチ*臭 火病 ファビョ ＜｀∀´＞ ニダー

抽出したコメントから「ユーザ IDi，視聴した番組 j，現代的レイシズムの有無 $y_m \in \{0,1\}$，古典的レイシズムの有無 $y_o \in \{0,1\}$，視聴時間の対数値 t_{ij}」というデータセットを作成した．このデータセットは 49,894 行で，599 人のユーザと 1078 番組で構成される．コリアン関係のコメントにおける現代的/古典的レイシズムを含むコメントの比率はそれぞれ 9.18% と 5.88%であった．y_* は番組 j を視聴したときの差別発言の有無である（1 が有，0 が無）．このときユーザ単位での現代的レイシズムと古典的レイシズムの表出確率の相関は弱かった（ピアソン積率相関係数が 0.140*** *3）．

連絡先: takano_masanori@cyberagent.co.jp

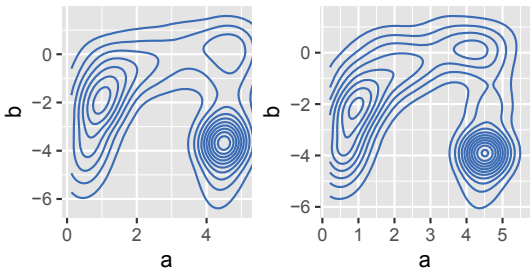
*1　https://abema.tv

*2　本研究の分析対象には差別的な語を用いたコメントも含まれるが，これらを伏せることは現状の認識の妨げになるため，研究の資料的価値に鑑みて抽出に用いた単語はそのまま掲載した．

*3　*** は t 検定で $p-\text{value} < 0.001$ であることを表す．

表 2: 現代的/古典的レイシズムの分類コード（[高 15] の拡張）．これらの単語をコメントが含んでいれば各分類のレイシズムのラベルを付与した．したがって現代的レイシズムと古典的レイシズムのラベルは排他的ではない．

分類	単語
現代的レイシズム	特権 生活保護 生保 受給 人権 通り名 通名 参政権 投票 選挙権 差別 年金
古典的レイシズム	犯罪 事件 逮捕 悪事 凶悪 強姦 レイプ 犯す 犯人 襲う 強盗 殺人 容疑 殺害 暴行 被害 整形 劣る 劣等 ヤクザ 火病 ファビョ



(a) 現代的レイシズム　　(b) 古典的レイシズム

図 1: a と b の密度分布

以下の統計モデルを用いて，現代的/古典的レイシズム（y_m, y_o）それぞれについてユーザ i の個人属性としての差別発言のしやすさ（θ_i）とニュース番組 j の差別発言投稿に関する性質（識別力 a_j と喚起力 b_j）を評価した．

$$y_* \sim B(p_*) \tag{1}$$

$$\mathrm{logit}(p_*) = a_j(\theta_i + b_j + \beta t_{ij}) \tag{2}$$

ここで $B(p)$ は従うベルヌーイ分布，t_{ij} はユーザ i の番組 j の視聴時間を調整するための共変量を表す．モデルは項目反応理論の 2 変数ロジスティクスモデル [Birnbaum 68] の拡張である．

パラメータの推定には Stan を用い，2000 回サンプリングを行って，その後半の 1000 回分を分析に用いた．事前分布は $a \sim \mathrm{N}(0,5)$（ただし $a \geq 0$），$b \sim \mathrm{N}(0,5)$，$\theta \sim \mathrm{N}(0,1)$，$\beta \sim \mathrm{N}(0,5)$ とした（$N(\mu,\sigma)$ は平均 μ，標準偏差 σ の正規分布）．以上のようにしてパラメータ推定をした結果，全ての RHat は 1.1 以下になり，十分に収束していると言える．

3. 分析結果

現代的レイシズムと古典的レイシズムの間の関連を番組とユーザの性質それぞれについて知るために a, b, θ の分布の平均値についてピアソン積率相関係数を使って評価した．その結果，a は 0.158^{***}，b は 0.202^{***}，θ は 0.771^{***} であった．すなわち，二種類のレイシズムはユーザ（θ）に関しては強い相関を示した．一方で，番組の性質（a と b）は相対的に相関は弱かった．

番組の性質は a と b によって大きく 3 つのクラスに分かれた．図 1 に現代的/古典的レイシズムにおける (a,b) の密度分布をそれぞれ示す．ここでは $a = 3$，$b = -1.5$ をしきい値として以下のようにクラス分けをした．

表 3: 現代的/古典的レイシズムのクラス間の共起頻度

現代的＼古典的	1	2	3	合計
1	124	42	156	322
2	47	54	105	206
3	114	92	344	550
合計	285	188	605	1078

- **1)** 識別力 a が大きく喚起力 b が小さいクラス．これは極端な態度（θ が大きい）の人だけがレイシズムを表出する番組クラスである．
- **2)** a が大きく b も高いクラス．これはあまりレイシズム傾向を持たないユーザでもレイシズムを表出しやすい番組クラスである．このクラスは現代的レイシズムのほうが多く見られた．
- **3)** a が小さく b は中程度のクラス．これはレイシズムはある程度喚起するが実際に表出するか否かのゆらぎは大きい番組クラスである．

この 3 つのクラス分けは現代的/古典的レイシズムで類似の傾向を示したが，両レイシズムにおけるクラス間の関連はあまり強いものではなかった．表 3 に各クラス間の共起頻度を示す．番組における現代的/古典的レイシズムの相関を評価するために Cohen の κ 係数を計算した所 0.14 と小さかった．

4. 議論

本研究では現代的/古典的レイシズムの表出要因を番組と人に分離するモデルを提案し，AbemaTV の番組につけられた差別的コメントに適用して分析した．その結果，先行研究 [McConahay 86, 高 15] でも観測されたユーザ単位での現代的レイシズムと古典的レイシズムの相関を確認した．一方で，そのユーザが表出するレイシズムとは番組の内容に依存していた．

レイシズムが表出された番組の性質にも 3 つのクラスが存在することが示された．1 つは極端な態度のユーザのみが態度を表出する番組（クラス 1），1 つはユーザの態度が弱くてもレイシズムの表出を喚起してしまう番組（クラス 2，3）である．これはレイシズムの表出を禁止する場合の対策を番組単位（前者）かユーザ単位（後者）のどちらの観点で実施すべきかが，番組内容によって異なることを示唆する．

各クラスの性質は現代的/古典的レイシズムで類似の性質を示した一方で，クラス間の相関は弱かった．これは番組内容がユーザの現代的/古典的レイシズムに関する偏見をどのように刺激するか？ が異なることを示唆する．レイシズム表出の対策のためには，具体的な番組内容と現代的/古典的レイシズムの関連を明らかにすることが必要である．

参考文献

[Birnbaum 68] Birnbaum, A.: Some Latent Trait Models and Their Use in Inferring an Examinee's Ability, in *Statistical Theories of Mental Test Scores*, pp. 397—-479, Addison-Wesley, Reading, MA, USA (1968)

[Kinder 81] Kinder, D. R. and Sears, D. O.: Prejudice and politics: Symbolic racism versus racial threats to the good life.,

Journal of Personality and Social Psychology, Vol. 40, No. 3, pp. 414–431 (1981)

[McConahay 86] McConahay, J. B.: Modern racism, ambivalence, and the Modern Racism Scale, in J. F. Dovidio, and S. L. Gaertner, eds., *Prejudice, discrimination, and racism*, pp. pp. 91—-125, Academic Press, San Diego, CA, US (1986)

[Munger 17] Munger, K.: Tweetment Effects on the Tweeted: Experimentally Reducing Racist Harassment, *Political Behavior*, Vol. 39, No. 3, pp. 629–649 (2017)

[高 15] 高 史明: レイシズムを解剖する: 在日コリアンへの偏見とインターネット, 勁草書房 (2015)

Organized Session | Organized Session | [OS] OS-9

[2E4-OS-9] 顔文字の科学

奥村 紀之（大手前大学）、Michal Ptaszynski（北見工業大学）、Rzepka Rafal（北海道大学）

Wed. Jun 5, 2019 3:20 PM - 5:00 PM Room E (301A Medium meeting room)

[2E4-OS-9-01] Analysis of Net Slang Expressing laugh at Comments on Nico Nico Douga

〇Shiori Aoyagi[1], Yasuo Kawai[1] （1. Graduate School of Information and Communications, Bunkyo University）

3:20 PM - 3:40 PM

[2E4-OS-9-02] A Multi-class Classification based on the Original Form of Kaomoji using Neural Network

〇Noriyuki Okumura[1], Rei Okumura[2] （1. Otemae University, 2. National Institute of Technology, Akashi College）

3:40 PM - 4:00 PM

[2E4-OS-9-03] Study on Potential of Automatic Emoticon Generation Based on Emoticon Parts

〇Tomoaki Sakai[1], Michal Edmund Ptaszynski[1], Fumito Masui[1] （1. kitami institute of technology）

4:00 PM - 4:20 PM

[2E4-OS-9-04] Convolutional Neural Network for Chinese Sentiment Analysis Considering Chinese Slang Lexicon and Emoticons

〇Da Li[1], Rafal Rzepka[1], Michal Ptaszynski[2], Kenji Araki[1] （1. Hokkaido University, 2. KITAMI Institute of Technology）

4:20 PM - 4:40 PM

[2E10-13-5] Panel discussion

4:40 PM - 5:00 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

ニコニコ動画のコメントにおける笑いを表現するネットスラングの分析

Analysis of Net Slang Expressing laugh at Comments on Nico Nico Douga

青柳志織[*1]
Shiori Aoyagi

川合康央[*2]
Yasuo Kawai

[*1] 文教大学 情報学研究科

Graduate School of Information and Communications, Bunkyo University

Abstract: Internet communication is developed day by day in the information society. Among these communications, the Internet's own slang called "net slang" is often used. In this research, we focused on net slang expressing laugh in net slang and aimed to clarify whether their meaning usage is the same. By comparing net slang, we found that the type of laugh and the subject of laugh tend to be different in each slang.

1. はじめに

現在，スマートフォンや PC の普及により，様々な世代の人々がインターネットを生活の一部として活用している．それに伴いインターネット上で書記言語を用いたテキストコミュニケーションが日々盛んに行われている．近年では，Skype や Discord を代表とする無料通話アプリケーションやボイスチャット等による音声コミュニケーションも可能となってきたが，現在でもテキストコミュニケーションの場は大いに活用されている．5ちゃんねる（旧：2ちゃんねる）や爆サイ.com といった大型電子掲示板，LINE などのテキストチャット，twitter や Facebook といった身の回りのニュースを発信し，コミュニケーションを交わすことができる SNS，コンテンツに関する評価や意見交換をすることが可能な YouTube やニコニコ動画が代表とされる動画共有サービスなど，テキストを用いたコミュニケーションを行う情報サービスは多く存在し，趣味の発信や娯楽，連絡手段など，生活の一部として，多くの人々に活用されている．

これらのテキストコミュニケーションの中では，“ネットスラング”と呼ばれるインターネット独自の俗語がしばしば用いられている．ネットスラングは様々なきっかけで生まれ，人々の間で流行し，インターネットならではの表現として定着し，また衰退していく．インターネットにおけるコミュニケーションが盛んな現代において，ネットスラングは新たな言葉の創造の場としての役割を担っていると考えられる．また，ネットスラングの変遷からインターネットコミュニケーションにおいて人々が何を求めているのかを観察したいと考えている．本稿では，ネットスラングの中でも笑いを表現するネットスラングについて注目する．

2. 笑いを表現するネットスラング

テキストコミュニケーションにおいて文に笑いの要素を付加するネットスラングを，笑いを表現するネットスラングとする．本稿では「笑」の漢字を含む「（笑）」を代表とし，その派生形である「w」，「草」を対象とした．対象のスラングを使ったものは以下のような例がみられる。「（笑）」，「w」については「現代日本語書き言葉均衡コーパス（BCCWJ）」で，「草」については「ニコニココメント等データ」から引用した

（1）今日はクリスマスという事で…アキバに行ってきます（笑）
（BCCWJ 2008）

（2）ワタシも笑っちゃいましたわwww
（BCCWJ 2005）

（3）本番のハードル上がりまくってて草生える
（NII 18）

以上のように「（笑）」，「w」は記号のように用いられ，「草」は普通名詞と変わらない用いられ方をしている。SNS やブログサイトを対象とした「w」に関する意味用法の研究を行った岩崎らによると，「w」には書き手が文章の内容が笑いを生じさせるものであることを示し，その笑いにも自嘲的な笑いや皮肉的な笑い，苦笑，微笑などさまざまな笑いを示しているとされている [Iwasaki 17] ．本研究では，「笑い」から派生したこれらのネットスラングの意味用法が本当に同じなのかを明らかにすることを目的とした．

3. 研究方法

3.1 調査対象

本研究では調査対象を「ドワンゴ」が提供している動画共有サービス「ニコニコ動画」のコメントを調査の対象にした。「ニコニコ動画」は 2006 年にサービスが開始され，ユーザが動画上にリアルタイムでコメントを書き込める独自のシステムが人気を集めている．インターネットの普及から，「5 ちゃんねる(旧2ちゃんねる)」などの大型電子掲示板サイトや，「あやしいわーるど」に代表されるアンダーグラウンドサイトでは，多くのネットスラングが生み出されてきた．「ニコニコ動画」にはこれらのネットスラングが多く流入され，現在のサブカルチャー界隈の文化を築いている．多くのネットスラングの使用例が抽出できると考えられる。コメントデータは国立情報学研究所のダウンロードサービスにより株式会社ドワンゴから提供を受けた「ニコニコ動画コメント等データ」を利用している．

3.2 データの収集方法

2016 年から 2018 年までの 6 月の月間総合ランキング上位 10 位までの動画を対象とし，期間は「ニコニコ動画」の月間コメントの期間に合わせ，毎年 5 月 1 日（AM6:00）から 6 月 1 日（AM5:59）までに設定し，その間に投稿されたコメントを抽出した．カテゴリはすべてのカテゴリが対象のカテゴリ合算に設定し

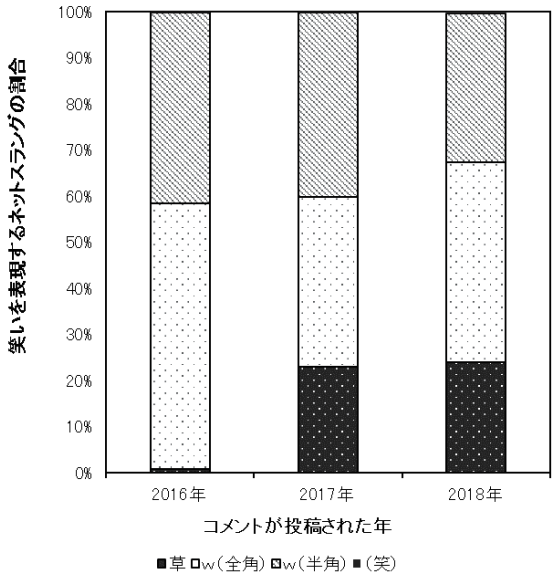
連絡先：青柳志織，文教大学情報学研究科，
b8g51001@shonan.bunkyo.ac.jp

た. 2018 年に関しては, 月間総合ランキング 2 位動画のコメントデータが大きすぎたため期間中のコメントを抽出することができなかった. そのため, 2 位を除いた 1 位から 11 位までの 10 個の動画を対象とした. さらに,「(笑)」,「w(全角)」,「w(半角)」,「草」の含まれるコメントのみを抽出したデータを各年用意した. ただし「w」には複数個並べて笑いの度合いを強くする使用法がある.「w」を並べる数に際限がないため, 分析作業を考慮し「w(全角)」,「w(半角)」に共通して4つ以上「w」が並ぶ場合,「w」の数の上限を3つまでに設定した. また,「草」や「w」を含む「草履」や「GW(ゴールデンウイーク)」などの笑いを表現するネットスラングとして機能していない単語が含まれたコメントに関しては文脈から確認し, 除外した.

表 1 データ収集の結果

	期間中抽出されたコメント		対象のスラングが含まれたコメント			
	期間中抽出されたコメントの合計	対象のスラングが含まれたコメント	草	w(全角)	w(半角)	(笑)
2016年	150,923	53,821	403	31,057	22,252	109
2017年	405,018	34,094	7,822	12,607	13,599	66
2018年	46,595	9,064	2,168	3,936	2,934	26

表 2 データの割合



データ収集によって以上のデータセットが作成された(表 1). 表 2 は対象の笑いを表現するネットスラングが含まれたコメントの, スラングごとの内訳を表したものである. 2018 年は抽出できたコメント数が非常に少なくなった. 2 位の動画のコメントが抽出できなかったことと, 9 位, 11 位の動画コメント数が非常に少なかったことが理由と考えられる.「w(全角)」,「w(半角)」を含むコメントはどの年でも多く抽出された.「草」を含むコメントの割合が多く見られるようになったのは 2017 年からである.「(笑)」に関してはどの年においても出現数はかなり少なかった. 笑いを表現するネットスラングの比較には収集できたデータが多く,「草」を含むコメントの割合が増えた 2017 年のデータを使用する. 本研究ではユーザーローカル テキストマイニングツールを利用して分析をおこなう. ユーザーローカルの「2つの文章を比較」機能を使用し, 特徴的な単語を抽出し, 各笑いを表現するネットスラングの特色を調べる.

4. 結果

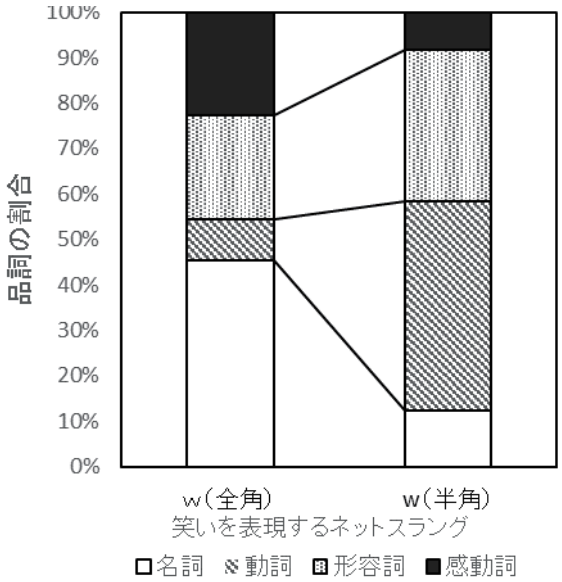
4.1 分析 1「w」同士の比較

「w(全角)」と「w(半角)」の比較をおこなった. コメントから抽出された単語の意味がどれだけポジティブ・ネガティブであるかの比較マップ(以下, ユーザーローカルでの名称に合わせ「ネガポジマップ」とする)においては大きな差は出なかった. 表3は「w(全角)」と「w(半角)」のデータに出現する単語を, それぞれどちらのデータに偏って出現しているか(以下,「単語分類」とする)で, グルーピングした表である. 個別の単語に傾向は見られないが,「w(全角)」には名詞, 感動詞が,「w(半角)」には動詞, 形容詞が多く抽出された(表 4). 岩崎らの研究では名詞につく「w」は単純な笑いと皮肉な笑いを示し, 感動詞につく「w」は単純な笑いを示す. また動詞につく場合, 自嘲や自虐的な笑いを示し, 形容詞においては他者への嘲笑が多いとされている. 本研究では「w」が付与された単語ではなく,「w」が含まれるコメントの中から抽出された単語のため条件は異なるが,「w(全角)」は単純な笑いを示すとき,「w(半角)」は嘲笑的な笑いを示すときに多く使われているといえるだろう.

表 3 「w (全角)」と「w (半角)」の単語分類

全角w	w(半角)
OrGA(名詞)	ウッホホホーホホwww(名詞)
動画(名詞)	草(名詞)
LUCIFER(名詞)	100万(名詞)
編集(名詞)	過ぎる(動詞)
ナターリア(名詞)	知る(動詞)
ちゃん(名詞)	聞こえる(動詞)
二次会(名詞)	ふく(動詞)
グラードン(名詞)	分かる(動詞)
うめ(名詞)	くう(動詞)
グレープ(名詞)	使う(動詞)
でる(動詞)	がる(動詞)
避ける(動詞)	持つ(動詞)
こわい(形容詞)	聞く(動詞)
こい(形容詞)	釣る(動詞)
でかい(形容詞)	早い(形容詞)
長い(形容詞)	くろい(形容詞)
ずるい(形容詞)	少ない(形容詞)
ありがとう(感動詞)	高い(形容詞)
あかん (感動詞)	酷い(形容詞)
お疲れ様 (感動詞)	かっこよい(形容詞)
すまん(感動詞)	わるい(形容詞)
いやー(感動詞)	嬉しい(形容詞)
	うわー(感動詞)
	なんてこった(感動詞)

表 4 「w (全角)」と「w (半角)」の単語分類結果の割合



4.2 分析 2 「草」と「w」の比較

分析2では笑いを表現するネットスラングの中でも 2017 年から多く出現するようになった「草」とどの年においても多く出現された「w」の比較をおこなった.「w」の中でも 2017 年により多くデータが抽出できた「w(半角)」を対象とする. ネガポジマップでは「草」寄りの単語はネガティブが 7 単語, ポジティブが 3 単語であり,「w(半角)」寄りの単語はネガティブが 21 単語, ポジティブが 6 単語であった(図 1). ネガティブ:ポジティブの比が「草」は 7:3,「w(半角)」では 7:2 で「w(半角)」の方がポジティブな単語の割合が低かった. また,「草」においては「草」自体や,「草生える」など「草」と共に使用される語が抽出されている.「草」自体がネガティブ寄りのものと認識されていると考えられる.「w(半角)」においては「w(半角)」自体がネガティブであるわけではなくネガティブな単語と共に現れるケースが多く見られた.

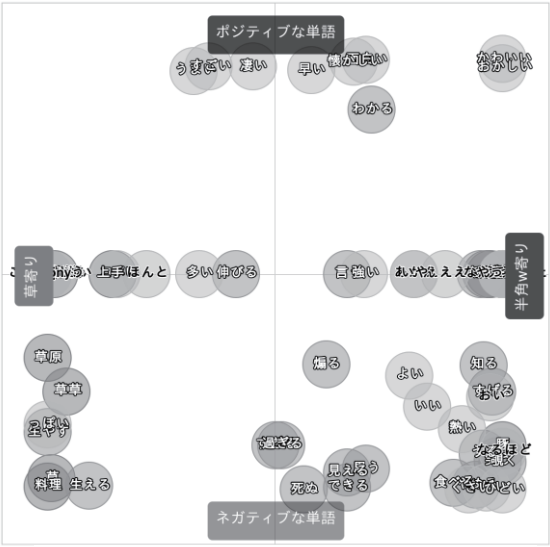


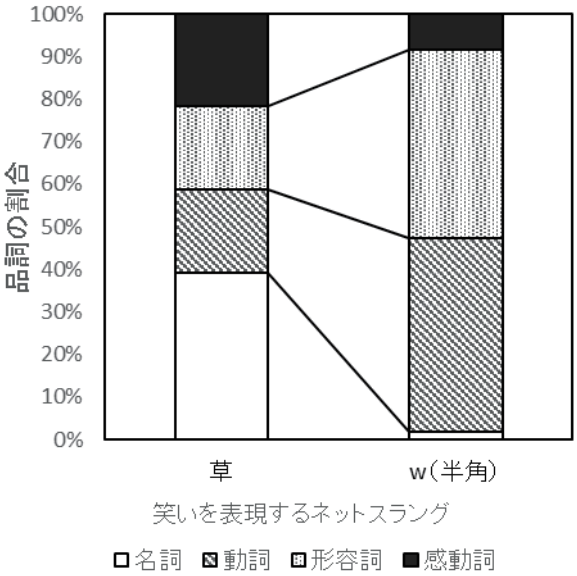
図 1 「草」と「w (半角)」のネガポジマップ

表 5「草」と「w (半角)」の単語分類

草		w(半角)	
草(名詞)	多い(形容詞)	w(名詞)	いい(形容詞)
コメ(名詞)	うまい(形容詞)		面白い(形容詞)
職人(名詞)	上手い(形容詞)	やめる(動詞)	熱い(形容詞)
普通(名詞)	思い(形容詞)	言う(動詞)	やばい(形容詞)
OrGA(名詞)	くろい(形容詞)	知る(動詞)	くさい(形容詞)
顔(名詞)	っぽい(形容詞)	すげる(動詞)	懐かしい(形容詞)
流石(名詞)	やすい(形容詞)	食べる(動詞)	よい(形容詞)
動画(名詞)	貴い(形容詞)	思う(動詞)	強い(形容詞)
声(名詞)	新しい(形容詞)	わかる(動詞)	良い(形容詞)
タグ(名詞)		できる(動詞)	怖い(形容詞)
若干(名詞)		見える(動詞)	わろい(形容詞)
草原(名詞)	ほんと(感動詞)	うる(動詞)	若い(形容詞)
ライ(名詞)	ありがとう(感動詞)	とる(動詞)	かっこいい(形容詞)
ルシファー(名詞)	お疲れ様(感動詞)	いく(動詞)	かわいい(形容詞)
NAVI(名詞)	わーい(感動詞)	とく(動詞)	長い(形容詞)
料理(名詞)	ごめんなさい(感動詞)	聞こえる(動詞)	おかしい(形容詞)
LUCIFER(名詞)	kq(感動詞)	笑う(動詞)	おもろい(形容詞)
まみれ(名詞)	お疲れさま(感動詞)	覗く(動詞)	ひどい(形容詞)
	すごーい(感動詞)	ひく(動詞)	あつい(形容詞)
	おめでとう(感動詞)	きつける(動詞)	優しい(形容詞)
生える(動詞)	anthony(感動詞)	いける(動詞)	おおい(形容詞)
増える(動詞)		あむ(動詞)	黒い(形容詞)
まくる(動詞)		泣ける(動詞)	大人しい(形容詞)
泣く(動詞)		踊れる(動詞)	すっぱい(形容詞)
生やす(動詞)		しとる(動詞)	なつかしい(形容詞)
散らかす(動詞)		しらむ(動詞)	こわい(形容詞)
湧く(動詞)		乗る(動詞)	おい(感動詞)
沸く(動詞)		諦める(動詞)	えええ(感動詞)
踊る(動詞)		乗り上げる(動詞)	すまん(感動詞)
			うううう(感動詞)
			なんてこった(感動詞)

単語分類においては,「草」では名詞, 感動詞が,「w(半角)」では形容詞, 動詞が多く出現した(表 5, 6).「草」と「w(全角)」を比較した場合においても「草」の含まれるコメントのほうが名詞,「w(全角)」では形容詞が多く出現する傾向がみられた.「草」は笑いの対象が名詞である際に用いられ,「w(半角)」は名詞そのものを笑いの対象にするよりは, 行為や状態が笑いの対象である場合に多く用いられるものと思われる.

表 6 「草」と「w (半角)」の単語分類結果の割合



4.3 分析 3 「(笑)」について

「(笑)」について収集できたデータ数が非常に少なく他の笑いを表現するネットスラングとの比較が難しかった.「(笑)」が含まれたコメントの抽出数は年々減少しており, 使用されなくなりつつあるネットスラングであるのかもしれない.「草」は普通名詞のように使用するため他の笑いを表現するネットスラングと使用方法が異なっているが,「(笑)」と「w」は文末に記号のようにつけるなど使用方法が似ている.「w」が含まれるコメントが多く抽出されることから「w」を使うのが主流と思われる.「(笑)」が含まれるコメントを見ていくと以下のような名詞に直に付与されているものがあった.

(4) ユーチューバー(笑)とかいうクソどもと比べ物にならんくらいおもろいわ

(NII 18)

(5) ちゃんと書いてあったろ？同じ番組のナレーター(笑)って

(NII 18)

これらは「(笑)」が付与された名詞を嘲笑する意味合いで使われていた. ミーハーな女性の別称に使われる「スイーツ(笑)」というネットスラングがある[Kindaichi 09]. 上記のコメントはこのスラングを真似た用法と思われる.「スイーツ(笑)」や,「ま～ん(笑)」など,「(笑)」には蔑称として使われる例があるため,「(笑)」は付与された名詞を侮蔑する役割として使用されていると考えらえられる.

5. まとめ

本研究では笑いを表現するネットスラング，「w（全角）」，「w（半角）」，「草」，「（笑）」の比較をおこない，意味用法が同じなのかを明らかにすることを目的とした．ニコニコ動画のランキング上位の動画のコメントを抽出し，笑いを表現するネットスラングの含まれるデータを用意し，ユーザーローカルテキストマイニングを使用して分析をおこなった．「w（全角）」は単純な笑いを，「w（半角）」は嘲笑的な笑いに使用される傾向があった．また多く抽出された「草」と「w」の比較においては，「草」は「草」自体がネガティブな単語とされ，「w（半角）」はネガティブな単語が含まれるコメントに出現する傾向が表れた． また単語分類の結果から笑いを表現するネットスラングによってコメント内に出現する品詞が異なることがわかった．笑いの対象が何かによって使用されるネットスラングが変わると考えられる．

本研究で対象とした笑いを表現するネットスラングは全般的にネガティブな意味や嘲笑，侮蔑的な笑いとして用いられている傾向があった．インターネットコミュニケーションの場においての笑いは嘲笑や批判的なものが多いと言えるのではないだろうか．

参考文献

[Iwasaki 17]岩崎 真梨子, 前田 梨沙, 川島 大樹：若者が着目するインターネット上の表現―ネットスラングと方言，八戸工業大学紀要，36 巻，pp. 41-56，2017.

[Kindaichi 09]金田一「乙」彦：オタク語辞典，美術出版，2009.

[BCCWJ]国立国語研究所：KOTONOHA「日本語書き言葉均衡コーパス」少納言，http://www.kotonoha.gr.jp/shonagon/

[NII 18] 国立情報学研究所: ニコニコ動画コメント等データ,2018.

[User Local] ユーザーローカル：テキストマイニングツール，https://textmining.userlocal.jp/

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

ニューラルネットワークによる顔文字の原形に基づくマルチクラス分類

A Multi-class Classification based on the Original Form of Kaomoji using Neural Network

奥村紀之 *1　　奥村嶺 *2
Noriyuki OKUMURA　　Rei OKUMURA

*1大手前大学 現代社会学部
Otemae University, Faculty of Modern Social Studies

*2明石工業高等専門学校 専攻科 機械・電子システム工学専攻
National Institute of Technology, Akashi College, Advanced Cource, Mechanical and Electronic System Engineering

In this paper, we propose a multi-class classification method for Kaomoji using feed forward neural network. Neural network has some units in each layer, but the suitable number of units is not clear. This research investigated the relation between the number of units and the accuracy of multi-class classification method.

1. はじめに

本稿では，顔文字の分類タスクの一つである，顔文字の原形推定に関する調査報告を行う．顔文字は，半角文字であるASCII 文字に加え，日本語のような全角文字も含めた様々な組み合わせにより表現され，現在確認されているだけでも既に10 万種類を超えている．また，Unicode の合成により，1 行に収まりきらない顔文字を合成することも可能となってきており，風間らの研究によって数百万種規模の顔文字が抽出可能だと言われている．一方で，顔文字を包括的に扱うための枠組みは確立されておらず，大規模な顔文字の辞書を保有しているに過ぎない状況である．

本稿では，多種多様な顔文字をまとめ上げるために顔文字の原形を定義し，任意の顔文字からその原型を推定するタスクに取り組んでいる．すでに顔文字の原形推定の手法についてはニューラルネットワークを用いた手法を提案しているが，中間層に用いるべきユニット数の妥当性が検証されていない．そこで，ある顔文字が単一の原形に集約されない（複数の原形を有する）可能性を踏まえ，マルチクラス分類における中間層のユニット数を調査している．

調査実験の結果，中間層 1 層の場合において 6,500 ユニットを準備することで，学習時間，正解率の観点から優れた結果が得られることが分かった．

2. 関連研究

Bedrick らは，PCFG による顔文字の検出を試みている．[Bedrick 12] 顔文字を抽出する対象としては Twitter に投稿された記事であるが，PCFG のルールを設定しているのみで，顔文字そのもの基本形などを定義しているわけではない．

風間らは，顔文字検出アルゴリズムを提案している．[風間 16]Bedrick らの研究と同様に，Twitter などの SNS に投稿された記事を解析し，顔文字様の文字列を抽出する手法について述べている．風間らの手法により，大量の顔文字様の文字列を抽出することは可能となるが，顔文字の構成をルール化しているのみで，本稿のような顔文字のまとめ上げには向かない．

Ptaszynski らは，顔文字を包括的に扱うための枠組みとして CAO システムを提案している．[Ptaszynski 10] CAO システムは，顔文字様の文字列を抽出することはもとより，抽出された顔文字が表現している感情を抽出することも目的としている．特に，顔文字の特徴として，目-口-目の並び（トリプレット）を顔文字の基本として抽出している．顔文字を表す基本の文字列を定義している点では本研究と類似しているが，我々のように対称性を考慮する [奥村 16] などの詳細な定義を行っておらず，やはりまとめ上げには向かない．

なお，著者のチームでは，ニューラルネットワークによる顔文字の分類手法の開発を進めている．[奥村 18] これまでの成果として，任意の顔文字から推定される唯一の原形を抽出することは 70%程度の正解率で可能となっているが，複数の原形を有する顔文字に対するマルチクラス分類は実装されていない．

本稿では，顔文字様の文字列を原形と称するクラスにラベリングし，複数のクラスに属する場合においても推定可能なシステムの構築を目指す．また，顔文字様の文字列から原形を推定するために必要な，ニューラルネットワークの中間層のユニット数について検証している．

3. マルチクラス分類可能な顔文字の原形抽出

従来の我々の手法では，3,110 種類存在している顔文字の原形のいずれかを出力することを目的としていた．しかし，例えば，(￣▽￣)ノ”o(*‥*)oゞ(￣▽￣) ヨチヨチのような顔文字の場合，顔を表すと考えられる文字列が 3 種類も含まれているため，いずれの原形を抽出して良いか判断できない．したがって，いずれか一つの原形を抽出するだけではまとめ上げとしても不十分であり，顔文字様の文字列に含まれている全ての顔とその原形を同定する必要がある．

本稿では，文字 Embedding を利用した固定長入力のフィードフォワードニューラルネットワークを利用し，マルチクラス分類を実装している．先ほどの例であれば，(￣▽￣)，(‥)，(￣▽￣) の 3 種類（厳密には (￣▽￣) が 2 回出現しているため 2 種類）の顔文字の原形を抽出することが出来れば正解とするようなモデルが必要となる．

連絡先: 奥村紀之，大手前大学現代社会学部，兵庫県伊丹市稲野町 2-2-2，TEL 072-770-6334，FAX 072-770-6916，noriyuki@otemae.ac.jp

3.1 顔文字のマルチクラス分類による原形推定

本稿で構成するニューラルネットワークは中間層 1 層の単純なモデルである．一方で，入力層に与えるべき情報は One-hot ベクトルでは不十分であることが分かっており，文字 Embedding を利用した入力の整形を行っている．本稿執筆時点で収集している顔文字データベースに登録されている顔文字の中で最長のものは 65 文字であるため，Embedding 層への入力は 65 ユニットとしている．Embedding 層へ入力された各文字は，顔文字の左側から順に 100 次元のベクトルへと変換され，順に結合される．65 文字に満たない顔文字については，不足分をゼロパディング（NULL 文字）で固定長の入力ベクトルを生成する．図 1 に本稿で採用したニューラルネットワークの構成を示す．

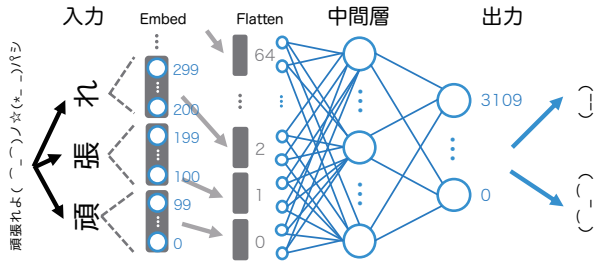


図 1: 顔文字の原形推定モデル

本稿では，図 1 における中間層のユニット数を変化させ，正解率（後述）による評価と学習時間の双方から，適切なユニット数を導出したい．なお，ユニット数の目安として，（入力ユニット数 ＋ 出力ユニット数）× 2/3 という目安があるという議論がなされている *1 が，解決すべき問題の複雑さ，学習データのサイズなど様々な要素に左右されるため，経験則の域を出ない．

3.2 評価方法

これまで，顔文字の原形推定における評価尺度は，出力された原形群の中に一つでも正しい原形が含まれていれば正解とカウントし，算出されていた．しかし，マルチクラス分類においては，出力された原形群に全ての原形が含まれているかを評価する必要がある．本稿では，評価尺度の難易度を 3 つに分類し，従来の評価方法（Easy），出力されるべき原形と出力された原形群に含まれる正解との比率による評価（Normal），顔文字様の文字列に含まれている原形全てを過不足なく推定することができるかという評価尺度（Hard）の 3 種類の評価方法によって傾向を調査する．

例えば，正解の原形が (-_-) (゜-゜) (._.) の 3 種類であり，推定された結果が (-_-) (._.) (_ _) だとすると，Easy では正解としてカウントされ，Normal では 2/3 という正解率が算出され，Hard では不正解としてカウントされることとなる．評価は以下の式 1，2，3 により行う．

$$Eval_{Easy} = \frac{\text{一つでも正解が出力された顔文字の数}}{N} \tag{1}$$

$$Eval_{Norm} = \frac{1}{N}\sum \frac{\text{出力された正解の数}}{\text{出力されるべき原形の数}} \tag{2}$$

$$Eval_{Hard} = \frac{\text{全ての正解が出力された顔文字の数}}{N} \tag{3}$$

*1 https://stackoverflow.com/questions/10565868/multi-layer-perceptron-mlp-architecture-criteria-for\-choosing-number-of-hidde

$Eval_{Easy}$ では，全評価対象に対する正解数の比率を算出し，$Eval_{Norm}$ では全評価対象における正解率の平均を求め，$Eval_{Hard}$ では $Eval_{Easy}$ と同様の比率を計算し，ユニット数による正解率の変化を調査する．

なお，評価では 28,296 組の顔文字様の文字列と原形のペアを使用し，10 分割交差検定により，平均正解率を算出している．以降の節では，10 分割交差検定の平均正解率の推移と学習時間を検討する．

4. 実験結果

本節では，3.2 節で述べた手法で図 1 を評価した結果について述べる．中間層のユニット数は 500 を最小とし，10,000 まで 500 刻みで検証した結果を示している．

4.1 損失の評価

図 2，図 3 に学習データおよび評価データにおける損失の推移を示す．縦軸は損失関数からの出力，横軸は Epoch 数を示す．

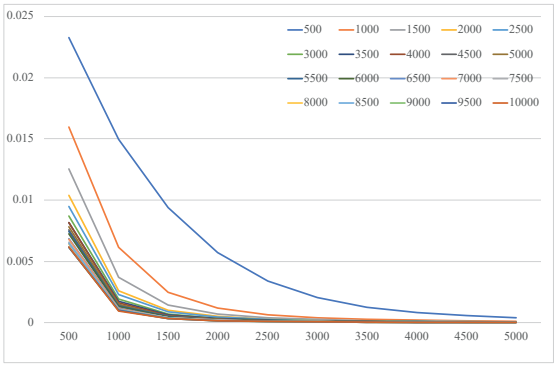


図 2: 学習データの損失

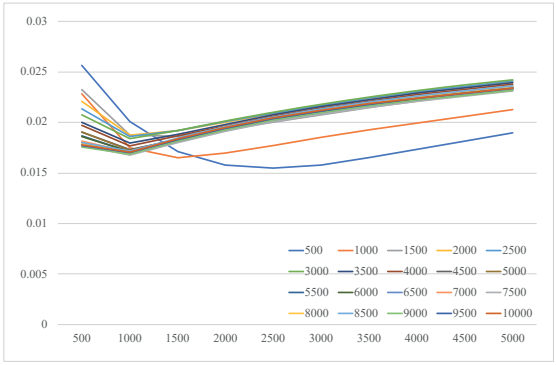


図 3: 評価データの損失

4.2 学習データにおける評価値の推移

図 4，図 5，図 6 に学習データにおける Easy，Normal，Hard の 3 手法による評価値の推移を示す．縦軸は正解率，横軸は Epoch 数を示す．

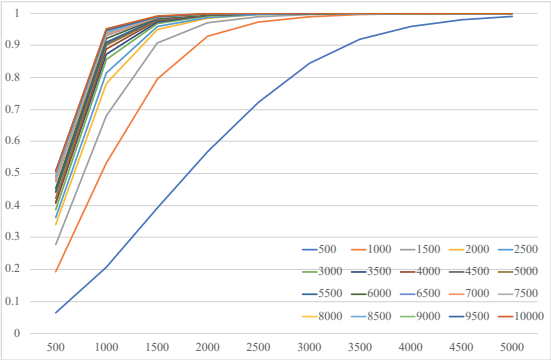


図 4: 学習データの Easy 評価

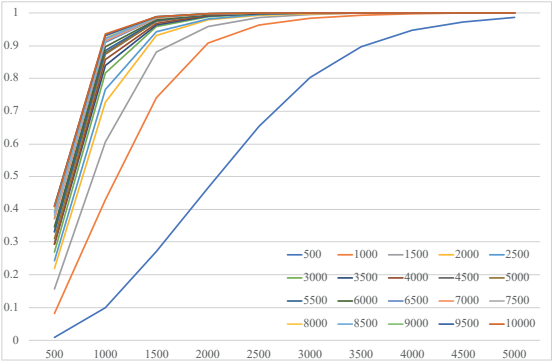


図 6: 学習データの Hard 評価

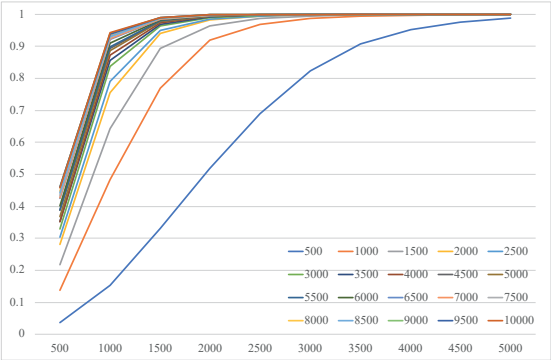


図 5: 学習データの Normal 評価

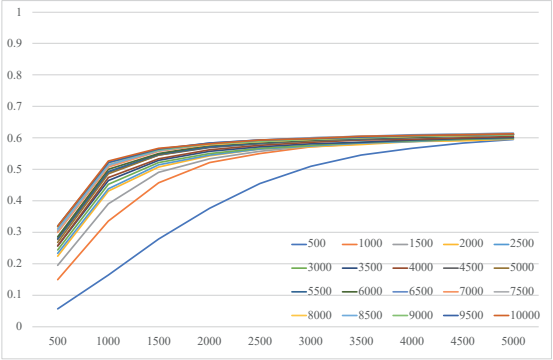


図 7: 評価データの Easy 評価

4.3 評価データにおける評価値の推移

図 7，図 8，図 9 に学習データにおける Easy，Normal，Hard の 3 手法による評価値の推移を示す．縦軸は正解率，横軸は Epoch を示す．

4.4 評価時間

図 10 に 10 分割交差検定に要した時間の推移を示す．縦軸は時間（秒）を示している．

5. 考察

4.1 節に示した図 2，図 3 から，学習は 1000Epoch 程度までで急速に進み，以降は評価データにおける損失が増加していくことから，本稿で採用しているニューラルネットワークのモデルでは 1,000Epoch 程度の学習までで十分であることが分かる．ユニット数による違いは，500 ユニット，1,000 ユニットのみ学習の進みが遅い傾向にあるが，それ以外のユニット数では同等の学習過程であることが分かる．

4.2 節に示した Easy，Normal，Hard のそれぞれの評価を見てみると，いずれの評価においてもユニット数が 500，1,000，1,500 の場合を除いて同等の傾向を示していることが分かる．1000Epoch 付近の評価値からも，少なくとも 2,000 ユニット以上の中間層を準備することによって，学習データに対しては 90%以上の正解率を得られており，学習が収束に向かっていることが読み取れる．

4.3 節に示した Easy，Normal，Hard それぞれの評価を確認すると，Easy の評価のみ 10%程度正解率が高くなるものの，Normal，Hard については大差のない評価が得られていることが分かる．このことから，マルチクラス分類を行う場合において，顔文字様の文字列の中に含まれていると思われる原形の総数以上に推定され出力されることは少なく，推定により出力される原形の個数は顔文字様の文字列に含まれている原形の総数（正解となる原形の個数）に等しいことが読み取れる．一方で，正解率は 50%程度と決して高くはなく，今後の性能向上が求められる．なお，1,000Epoch 付近の正解率を確認してみると，10,000 ユニットの場合が最も正解率が高く 42%ほどであるが，3.1 節で述べた経験則に基づいて算出したユニット数が 6500 (=(6500+3110) × 2/3=6406 ≒ 6500) の場合の正解率を確認してみると 41%程度とその差はほとんどないことが分かる．また，全体的な傾向としては，ユニット数が増えるごとに正解率はわずかながら上昇するが，経験則によるユニット数の場合と比較して正解率が低くなる場合もあるため，ユニット数を増やせば増やすほど良いというわけではないことも分かる．

最後に，4.4 節に示した学習に要する時間について検討する．学習に要する時間は図 10 に示したとおり，ユニット数が増えるごとに増大している．なお，ユニット数が 2,500 の場合に

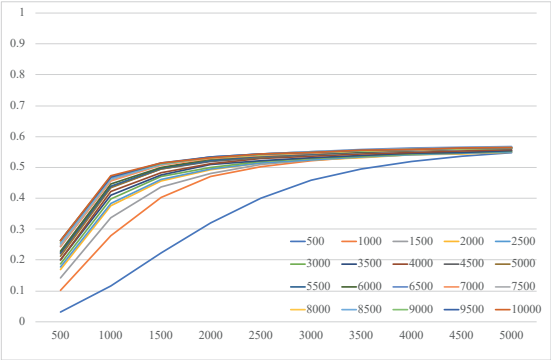


図 8: 評価データの Normal 評価

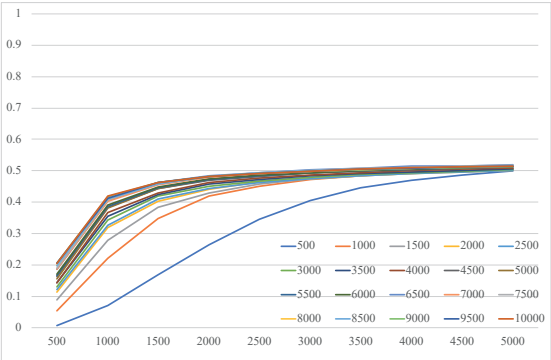


図 9: 評価データの Hard 評価

ついては，計算途中に何らかの割り込みが生じたため，ユニット数が 10,000 の場合をはるかに上回る時間が経過していたため，結果の図からは除外している．10 分割交差検定に要した時間のみを基準にすれば，ユニット数は少なければ少ないほど学習時間の観点からは効果的であるが，正解率は向上しづらいため，経験則に基づくユニット数付近のユニット数を採用するのが良いと考えられる．また，ユニット数が 9,000 以上になると，急激に所要時間が増えてしまう傾向があるため，本稿で採用しているモデルに置いては，8,000 以下のユニット数が妥当だと考えられる．

6. まとめ

本稿では，顔文字の原形推定のためのフィードフォワードニューラルネットワークにおける中間層のユニット数について調査した．経験則に基づくユニット数の最適値について実験的に確認し，およそ経験則から外れることのないモデルであることが分かった．中間層のユニット数を 6,500 とした場合，3,500Epoch 以上の学習でどの評価方法においても 50%を超えることが確認されているが，学習に時間がかかりすぎてしまうという問題が発生する．

今後の課題としては，入力層に与えるべき情報として文字 Embedding を実施しているが，顔文字のデータベースに出現する文字しか対応が出来ていないため，あらゆる文字へ対応

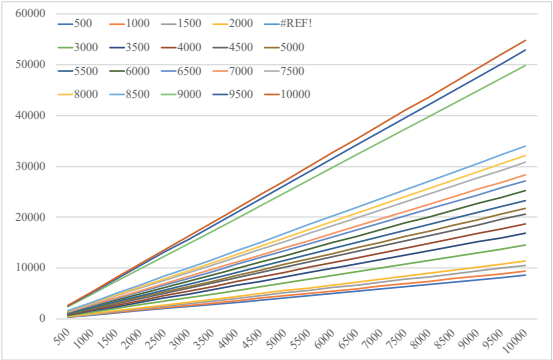


図 10: 10 分割交差検定に要した時間

可能な Embedding 手法を考える必要がある．また，入力が 6,400 次元の固定長であるため，リカレントニューラルネットワークのような可変長の入力に対応可能なモデルを適用して，任意の長さの顔文字に対応可能なシステムへと拡張する必要がある．同様に，本稿では中間層を 1 層としているが，多層化した場合の正解率についても調査を行い，顔文字の分析に使用するモデルとしての層数とユニット数の関係を明らかにしていきたい．

謝辞

本研究は JSPS 科研費 18K11455 の助成を受けている．

参考文献

[Bedrick 12] Bedrick, S., Beckley, R., Roark, B., and Sproat, R.: Robust kaomoji detection in Twitter, in *Proceedings of the Second Workshop on Language in Social Media*, pp. 56–64, Association for Computational Linguistics (2012)

[Ptaszynski 10] Ptaszynski, M., Araki, K., Dybala, P., Rzepka, R., and Maciejewski, J.: CAO: A Fully Automatic Emoticon Analysis System Based on Theory of Kinesics, *IEEE Transactions on Affective Computing*, Vol. 1, pp. 46–59 (2010)

[奥村 16] 奥村 紀之：顔文字の原形抽出, 言語処理学会年次大会, Vol. NLP2016, pp. P1–1 (2016)

[奥村 18] 奥村 嶺, 奥村 紀之：文字 Embedding を使用したニューラルネットワークによる顔文字の原形推定, 人工知能学会全国大会論文集, Vol. JSAI2018, pp. 2B405–2B405 (2018)

[風間 16] 風間 一洋, 水木 栄, 榊 剛史：Twitter における顔文字を用いた感情分析の検討, 人工知能学会全国大会論文集, Vol. JSAI2016, pp. 3H3OS17a4–3H3OS17a4 (2016)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

顔文字パーツに基づいた顔文字の自動生成の可能性に関する調査

Study on Potential of Automatic Emoticon Generation Based on Emoticon Parts

酒井 伴彰 *1　プタシンスキ ミハウ *1　桝井 文人 *1
Tomoaki Sakai　Michal Ptaszynski　Fumito Masui

北見工業大学 情報システム工学科
Department of Computer Science, Kitami Institute of Technology

In this paper, we present our preliminary study in research on generating emoticons completely automatically for a context specified by emotions expressed in a sentence. To verify whether it is even feasible to generate comprehensible emoticons fully automatically, we firstly generate a number of emotion-specific emoticons according to a specified algorithm and then perform a survey in which we ask users about the affective impressions of the artificially generated emoticons. The survey revealed similarities between what were the emotions perceived in those emoticons by human participants and their main emotion score assigned automatically according to the emoticon parts that were used in the generation of those emoticons. This suggests it should be possible to generate comprehensible emoticons fully automatically in the future.

1. はじめに

近年, メールを始め，Facebook，Twitter，LINE などのような Social Networking Services（SNS）を通じたオンラインコミュニケーションが広く普及している．しかし，長年使用されてきたインターネット上の文字ベースのみでのコミュニケーションは実際の会話とは異なり，表情や話し方等の非言語情報が与えられず, 誤解が生じることがある．非言語情報の不足の中で行われる情報伝達を支援するため, 顔文字や絵文字, スタンプ, アスキーアートなどが提案された．また，絵文字，高画質画像（スタンプ，ステッカーなど）や単純なアニメーション（GIF）などによって多彩にしていく一方，文字のみでできた顔文字が人気を維持し古くから使用されている [山本&木村 2016, プタシンスキら 2017]．顔文字とは記号や文字で構成された表情や姿勢を表現するものであり，テキストでのコミュニケーションで補足的な情報を文に付与することができる．

しかし，顔文字は数多くの記号や特殊文字から成立されており，日常コミュニケーションの中でユーザは自発的に新しい顔文字を作ったり, それを使い続けたりしているため，ユーザによって顔文字の理解度が異なる．意味の曖昧な顔文字が使われた場合，誤解を防ぐためにはその曖昧性を解消する必要があると論じられる [石井ら 2017]．一方，より広い範囲のユーザにとって分かりやすい顔文字を文脈に応じて自動的に生成することができたらインターネット上のコミュニケーションの円滑化へ大きく貢献できることが見込まれる．

そこで，本研究では文の感情的な文脈に合わせた顔文字パーツに基づいた顔文字の生成が可能であるかをアンケート調査を用いて調査する．また, 今回は, 日本語や東アジアの文字コードの使用圏などで使用される (^_^) のような正位置の形式の顔文字を対象とする．

この論文は以下の通り構成されている 2. 章では顔文字の生成に関する従来研究について説明する．3. 章では本研究で提案する顔文字の生成のためのアルゴリズムを紹介する．また，4. 章ではパーツから自動生成した顔文字では既存の顔文字と同様の分かりやすさが得られるかについてアンケートをもとに確証する．5. 章ではアンケートの結果を分析して考察を行う．最後に 6. 章では論文をまとめ今後の課題を検討する．

2. 従来研究

顔文字の自動生成に関する研究としては，以下のような研究が挙げられる．[中村 2002] は，顔文字をパーツ（口，目，など）に分割し，機械学習モデルを作成し，対話システムにおいてユーザ入力文に合った顔文字の自動生成を行った．しかし，顔文字と感情表現間の関連性の調査は小規模のサンプルで行ったため結果には不十分な点があった．

中村の研究以来，顔文字の自動生成と異なる顔文字の自動推薦の研究が盛んとなった．その中，文中の感情解析をベースにした顔文字の自動推薦 [Urabe ら 2015] や，幅広い情報の種類（感情，あいさつ，眠気など）を用いた顔文字推薦の研究 [江村&関 2012] が挙げられる．

上記の研究と違い，[Ptaszynski ら 2010] は顔文字の自動解析を行うため顔文字に対して [中村 2002] と似たアプローチをとり，大規模の顔文字データベースをパーツ分割にしオープンソースとして公開した *1．本研究では，この顔文字データベースを用いて調査を行うことにした．

3. 顔文字生成アルゴリズム

3.1 CAO システムの顔文字パーツのデータベース

[Ptaszynski ら 2010] は，顔文字をパーツ（目，口など）に分割し顔文字データベースを構築し，文中顔文字抽出，感情解析を行う CAO システムを提案した．データベース構築のためには，インターネット上にある複数の顔文字辞書を参照し，各辞書内に顔文字が付与されているラベル（たとえば，“ 挨拶 ”，“ 突っ込み ”，“ 励まし ”，“ 怒る ”，“ 泣く ”，“ 笑う ” など）の中から感情関連のラベルのみ使い,「喜・怒・哀・怖・恥・好・厭・昂・安・驚」という 10 種類の感情種類に分類した 1 万以上の顔文字を収集した．さらに，これらの 1 万の顔文字を「口」,「目」などのパーツに自動的に分割し，感情種類別顔文字パーツのデータベースを構築した．

連絡先:

*1　http://arakilab.media.eng.hokudai.ac.jp/ ptaszynski/repository/cao

3.2 顔文字パーツに基づいた顔文字生成アルゴリズム

顔文字データベースを元に顔文字生成アルゴリズムを考えた．[Urabe ら 2015] と同様に，[Ptaszynski ら 2017] の感情解析システム ML-Ask*2 を用いて，文中表出される感情的文脈を解析し，生成する顔文字の感情種類を確定する．さらに，顔文字データベースを参照し，文の感情的文脈に合った感情種類のデータベース内の頻出の顔文字パーツを組み合わせて以下のように新しく顔文字を生成する．

CAO システム [Ptaszynski ら 2010] のデータベースには,「喜」「安」「昂」「好」「驚」「怖」「嫌」「怒」「恥」「哀」の 10 種類の各感情のパーツごとに使用頻度の高いものから順に並べられている. 例として表 1，表 2 はデータベースの一部を表した. このデータベースから 10 種類の各感情で使用頻度の高い顔文字のパーツを組み合わせ，その使用頻度の値の合計が高いものの中から上位 10 個の顔文字を用いてランダムに選択する．

しかし, このように新しく生成した顔文字は「｀ д´」のように普段括弧 () の中身だけのため，最後に，生成した顔文字の両面に括弧を付与し「(｀ д´)」のような () がついたものを最終的に生成した顔文字とする．

4. 調査設計

3. 章で提案した顔文字の自動生成アルゴリズムの有効性を検証するためには以下のようにアンケート調査を設計した．

4.1 アンケート用設問

アンケートの設問について以下で述べる. 顔文字データベースから，1 つの感情種類につき 2 つ，計 20 個の顔文字を 1 セットとし，アンケートを 4 セット作成した．アンケートに参加する被験者の回答としては，各顔文字が表す感情種類について回答してもらう形式にした.

アンケートは選択形式で行う. しかし，回答選択肢として顔文字データベースにあるすべての感情種類を用いた場合，被験者への負荷となり回答の品質への影響が起こるため，4 つだけの選択肢を与えた．

具体的には，アンケートの回答には (1) 元となった感情種類（正解扱い）,(2) 類似した感情種類,(3) 反対の感情種類, そして (4) その他という 4 つの選択肢を与えた. そこで，類似した感情種類及び反対の感情種類は，[Ptaszynski ら 2017] のように [Russell 1980] の 2 次元感情モデルの上にマッピングした ML-Ask の感情種類を参照しその中から選択した．また，感情種類のラベルとしては，元のデータベースのラベルを使った場合，普段使われていない文字（例：昂など）が含まれるため，感情解析システムの感情表現データベースを参照し分かりやすい表現を回答とした．このように，例えば,「怒」の感情種類から顔文字を作った場合, 選択肢には (1)「怒り」と (2)「怖い」と (3)「安らぎ」と (4)「その他」にする. また，正解が分からないように，(1)〜(3) の回答を問題ごとにランダムの順番にした．そして，その中に被験者の考えに合う感情種類がなかった場合のためには,【その他】という選択肢を与え，自分の考えに合った顔文字の感情種類を記述してもらうことにした．

このように，4 セットのアンケート *3 ができ，調査を行った．

*2 https://github.com/ptaszynski/mlask
*3 https://goo.gl/forms/5NHHfu6lRsEZtZC13
https://goo.gl/forms/5JXkaAJfOcKHkWV82
https://goo.gl/forms/bygq0jFMpLQcEGzA3
https://goo.gl/forms/qTgQ4D88mWbLA5Os2

表 1: 顔文字データベースの一部 (怒り)

頻度	パーツ	頻度	パーツ
42	д	53	｀ ´
36	2	47	° °
34	_	46	- -
29	д	42	´ ｀
24	ω	40	- -

表 2: 顔文字データベースの一部 (喜び)

頻度	パーツ	頻度	パーツ
97	∀	81	^ ^
64	▽	77	^ ^
58	_	68	´ ｀
54	∇	59	° °
49	ω	53	- -

5. 結果と考察

5.1 被験者について

予備調査として各セットで回答者数を最低 20 人を目指していたところ第 1 セットは 42 人となり，回答者数が段階的に下がりセット 4 は丁度 20 人となった．今後回答者数を増やし各セットで 100 人からの回答を収集することを目指す．また，男女別割合は各セットで凡そ半分となったが，男性回答者の方が数人多かったため今後より均等の割合になるように調整して，統計的信頼性を高め男女別の顔文字の理解度について調べたい．各セットの回答者の統計情報を表 3 に表した．

表 3: セット毎の回答者の統計情報.

セット	回答者数	年齢（代/%） 10 代/20 代/30 代/40 代	男女別割合 男/女
1	42	4.8%/64.3%/11.9%/19.0%	57.1%/42.9%
2	32	6.3%/62.5%/15.6%/15.6%	59.4%/40.6%
3	26	3.8%/65.4%/19.2%/11.5%	57.7%/42.3%
4	20	5.0%/65.0%/15.0%/15.0%	55.0%/45.0%

5.2 アンケート結果と考察

顔文字データベース内のの出現頻度をもとにして生成した顔文字の感情の種類と実際に回答者が感じる感情の種類の差を調べた. 結果は以下のとおりとなった．

アルゴリズムの通りに生成した顔文字の感情種類が多くの場合，大半の回答者の回答と同様となった．例えば，(^∀^) の顔文字の感情種類は「喜」となっていたところ，92.9%の回答者は「嬉しい」と回答した．また，アルゴリズムと一致した回答を選ばなかった回答者でも，類似した感情種類を選んだことも多かった（例:「喜」ではなく「好」）．その現象は，[Ptaszynski ら 2010] の研究でも取り上げられ，人間によって感情の理解が多少異なっているが，[Russell 1980] の 2 次元感情モデルで顔文字の表す感情種類を抽象化させたら，ほとんどの場合，具体的な感情の種類が異なっていても次元は合っていることが確認されている．さらに，(^∀^) の顔文字には,「煽り」や「皮肉」という回答もあったため，今後，[江村&関 2012] の研究でも触れられていた「動作」の表現として顔文字や [村上ら 2011] の強調している顔文字における皮肉の役割に関する研究にも応用することが考えられる．上記と同様に，(゚д゚) の顔文字の大半の回答は「驚き」となったが，そのほかには，次元が合っ

ている「怒り」や「焦り」も見られた．また，構成記号が少しだけ異なっていても印象が異なる例もあった．例えば，(゜д゜) の場合は，口の部位は大きさのみが異なっている（ 凵 対 д ）が，第 2 印象は，前者の「焦り」から「怖い」に代わっている．その説明としてはこのようなことが考えられる．[Ptaszynski ら 2010] の研究では，顔文字は表情や姿勢など，身体言語（ボディランゲージ）と同様な扱いすることができると強調されている．また，実際の表情と同じように，表情筋のわずかな変更によって印象（読み取られる感情の種類）が変わることがあり [Birdwhistell 1952]，顔文字も，インターネット上のボディランゲージとして扱うことが可能であることの証明だと考えられる．

さらに，(｀凵´) のように回答が 100 %一致した例もあった．

アルゴリズムの通りに生成した顔文字の感情種類と回答者の回答が一致した顔文字の例を図 1 に表した．

また，上記の結果と違い，回答者の回答が分散した例もあった（図 2）．

しかし，その中，3 種類の回答パターンが見られた．まず，(´凵｀) のように，分散は見られたが，1 つか 2 つの感情種類が大部分の回答を覆い，さらに 2 次元での一致が高かった．次に，(￣ω￣) のように，結果が大きく分散した例もあった．しかし，そのような例の場合でも，2 次元の抽象化では，偏りの傾向が見られた（例 :「喜び」,「好き」,「安らぎ」というポジティブな印象）．最後に，(~2~) や (￣ 2 ￣) のような顔文字はの場合は，結果が分散し，それぞれの回答が異なっている印象を表した．その原因としては，表情と結びつきにくい記号が考えられる．[石井ら 2017] によると，顔文字には，表情と結びつきにくい記号（ ~ や数字の 2 など）が顔文字の構成に利用されたら，その顔文字全体の理解度が下がり，印象も想像しにくい．さらに，石井らは，その問題を解決するように顔文字の意味の曖昧性を定量化しスコアの低い顔文字の共通点を探し，原因となる記号を確定してる．今後，本研究のフォローアップとして顔文字の自動生成を行う時は，顔文字データベース内の記号の出現頻度のみならず，出力する前に，出力候補の顔文字の意味の曖昧性を算出し，より曖昧性のない顔文字を利用することが望ましい．

6. まとめ

本研究では，顔文字の自動生成に向けて，感情種類別で構成された顔文字のパーツを含むデータベースに基づいて新しい顔文字を生成するためのアルゴリズムを提案し，その正確性を検証した．検証するためには，人間回答者に各顔文字の感情的印象（感情の種類）についてアンケートにて尋ね，それがデータベース内でパーツに付与されている感情種類と一致しているかを調べた．

アンケートの結果としては，回答者が認識した顔文字の感情的印象とアルゴリズムに沿って生成した顔文字の感情の種類が大幅一致したことを確認した．なお，人間にとって理解可能な顔文字の完全自動生成は可能であること確認した. 今後，提案したアルゴリズムの自動化を行い，従来研究にて提案されてきた顔文字推薦と異なり，文の感情を考慮したパーツからの顔文字の自動生成を行うシステムの実装を行う予定である．

また，より信頼性の高いシステムを実装するためには，今後アンケートの回答者数を現在の 20〜40 人ほどから，各セットで 100 人まで増やしアンケート結果の統計的信頼性を上げる予定である．また，男女別の顔文字の理解度についても調べる予定である．そして，[石井ら 2017] の顔文字の意味の曖昧性推定手法を応用し，一般ユーザにとってより分かりやすい顔文字の提供を行いたい．

参考文献

[山本&木村 2016] 山本恭子, & 木村昌紀. 協調的なメール交換過程における顔文字の返報性規範の影響. 感情心理学研究, 23(Supplement), ps12 (2016).

[プタシンスキら 2017] プタシンスキミハウ, 奥村紀之, & ジェプカラファウ. 顔文字の現象および研究の概観: 記号の遊びが科学されるようになった道 (特集 顔文字の科学: Web 上の非言語表現・行動に関する新研究分野の誕生). 人工知能: 人工知能学会誌: Journal of the Japanese Society for Artificial Intelligence, 32(3), 333-341, (2017).

[石井ら 2017] 石井 直人, 桝井 文人, プタシンスキ ミハウ. 言語表現性に基づく顔文字の意味の曖昧性の自動推定. In Proceedings of The 31st Annual Conference of the Japanese Society for Artificial Intelligence (JSAI 2017), pp. 1-4, paper ID: 2O2-OS-22a-1.

[中村 2002] 中村純平. 感情モデルと DialogueAct を用いて自由対話システムの応答文に合った顔文字を出力する手法. 情報処理学会第 64 回全国大会, 2, 25-26, (2002).

[Urabe ら 2015] Urabe, Y., Rzepka, R., & Araki, K. Comparison of Emoticon Recommendation Methods to Improve Computer-Mediated Communication. In Recommendation and Search in Social Networks (pp. 23-39), (2015).

[江村&関 2012] 江村優花, & 関洋平. (2012). マイクロブログにおける感情・コミュニケーション・動作タイプの推定に基づく顔文字の推薦. 自然言語処理, 19(5), 401-418.

[Ptaszynski ら 2010] Ptaszynski, M., Maciejewski, J., Dybala, P., Rzepka, R., & Araki, K. CAO: A fully automatic emoticon analysis system based on theory of kinesics. IEEE Transactions on Affective Computing, (1), 46-59, (2010).

[Ptaszynski ら 2017] Ptaszynski, M., Dybala, P., Rzepka, R., Araki, K., & Masui, F. ML-Ask: Open Source Affect Analysis Software for Textual Input in Japanese. Journal of Open Research Software, 5(1), (2017).

[Russell 1980] Russell, J. A. A circumplex model of affect. Journal of personality and social psychology, 39(6), 1161, (1980).

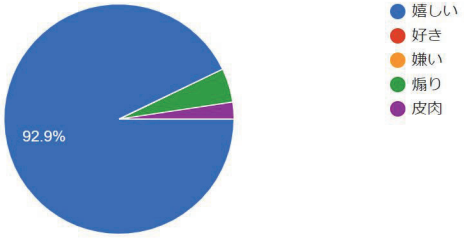
[村上ら 2011] 村上浩司, 山田薫, & 萩原正人. 顔文字情報と文の評価表現の関連性についての一考察. 言語処理学会第 17 回年次大会発表論文集, 1155-1158, (2011).

[Birdwhistell 1952] Birdwhistell, R. L. Introduction to kinesics. University of Louisville, (1952).

図 1: アルゴリズムの通りに生成した顔文字の感情種類と回答者の回答が一致した顔文字の例.

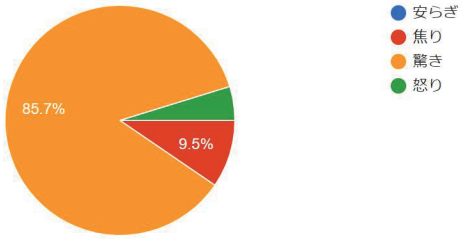
1. (^∀^)

42 responses



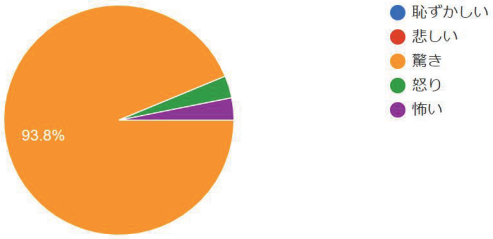
11. (゜Д゜)

42 responses



12. (゜д゜)

32 responses



19. (`Д´)

41 responses

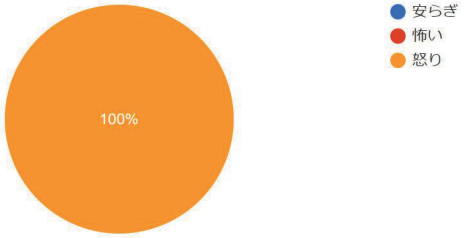
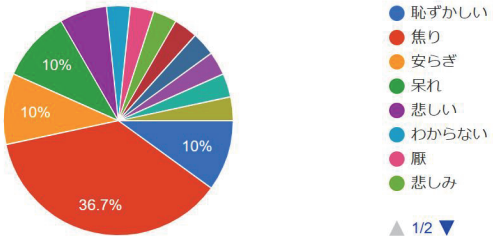


図 2: 回答が分散した顔文字の例.

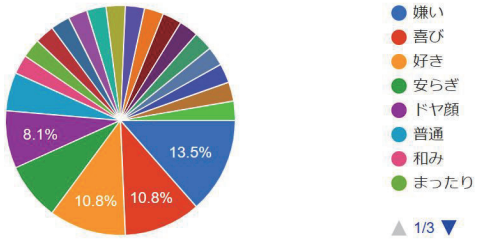
13. (´Д`)

30 responses



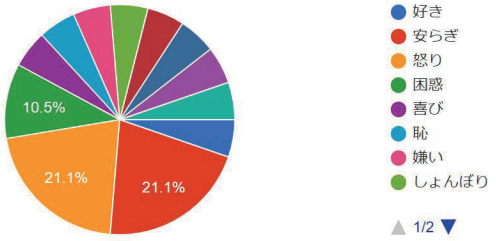
4. (￣ω￣)

37 responses



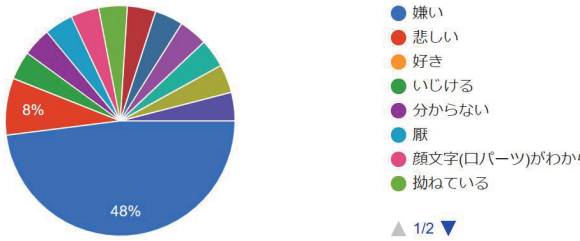
9. (~2~)

19 responses



(￣2￣)

25 responses



The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Convolutional Neural Network for Chinese Sentiment Analysis Considering Chinese Slang Lexicon and Emoticons

Da Li[*1] Rafal Rzepka[*1] Michal Ptaszynski[*2] Kenji Araki[*1]

[*1] Graduate School of Information Science and Technology, Hokkaido University
[*2] Department of Computer Science, Kitami Institute of Technology

Nowadays, social media have become the essential part of our lives. Internet slang is an informal language used in everyday online communication which quickly becomes adopted or discarded by new generations. Similarly, pictograms (emoticons/emojis) have been widely used in social media as a mean for graphical expression of emotions. People can convey delicate nuances through textual information when supported with emoticons, and the effectiveness of computer-mediated communication is also improved. Therefore, it is important to fully understand the influence of Internet slang and emoticons on social media. In this paper, we propose a convolutional neural network model considering Internet slang and emoticons for sentiment analysis of Weibo which is the most popular Chinese social media platform. Our experimental results show that the proposed method can significantly improve the performance for predicting sentiment polarity.

1. Introduction

Today, many people share their lives with their friends by posting status updates on Facebook, sharing their holiday photos on Instagram or tweeting their views via Twitter or Weibo - the biggest Chinese social media network that was launched in 2009. Social media data contain a vast amount of valuable sentiment information not only for the commercial use, but also for psychology, cognitive linguistics or political science [Li 18a].

Sentiment analysis of microblogs became an important area of research in the field of Natural Language Processing. Study of sentiment in microblogs in English language has undergone major developments in recent years [Peng 17]. Chinese sentiment analysis research, on the other hand, is still at early stage [Wang 13] especially in the domains of lexicons and emoticons.

Pictograms (emoticons/emojis) have been widely used in social media as a mean for graphical expression of emotions. For example, 😂 "face with tears of joy", an emoji that means that somebody is in an extremely good mood, was regarded as the 2015 word of the year by The Oxford Dictionary [Moschini 16]. In our opinion ignoring emoticons in sentiment research is unjustifiable, because they convey a significant emotional information and play an important role in expressing emotions and opinions in social media [Novak 15, Guibon 16].

Internet slang is ubiquitous on the Internet. The emergence of new social contexts like micro-blogs, discussion groups or social networks has enabled slang and non-standard expressions to abound on the Web. Despite this, slang has been traditionally viewed as a form of non-standard language, a form of language that is not the focus of linguistic analysis and has largely been neglected [Kulkarni 17].

Furthermore, we also noticed that when people use new words and pictograms, they tend to express a kind of humorous emotion which is difficult to be easily classified as positive or negative. It seems that some emoticons are used just for fun, self-mockery or jocosity which expresses an implicit humor which might be characteristic to Chinese culture. Figure 1 shows an example of a Weibo microblog posted with emoticons and Internet slang. In the third line of the post, 柠檬人 (*ning meng ren*) is a new word appeared in early 2019 on Chinese social media which means "lemon man". Accordingly, to meet this new popular phrase, 🍋 was added to the pictogram repoitoare by social media companies in January 2019. This lemon with a sad face also called "lemon man" which expresses the same emotion as slang *ning meng ren* – "sour grapes" or "jealous of someone's success". This entry seems to express a humorous nuance of a pessimistic attitude. Emoticons and slang seem to play an important role in expressing this kind of emotions. There is a high possibility that this phenomenon can cause a significant difficulty in sentiment recognition task.

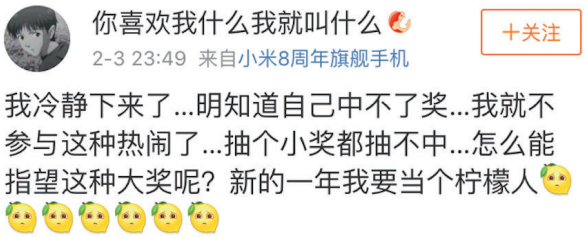


Figure 1: Example of Weibo post with Internet slang and emoticons. The entry says "I have calmed down. Knowing that I won't win the prize, I don't participate in this excitement. I even can't win a small prize. How can I expect a big one? I want to be a lemon man in the new year".

To address this phenomenon, in this paper we focus on the Internet slang and emoticons used on Weibo in order to establish if both slang and emoticons improve sentiment analysis by recognizing humorous entries which are difficult to polarize. To perform experiments, we built a Chinese Internet slang lexicon and a Chinese emoticon lexicon. Because the emoticons probably play a more important role in expressing emotion than textual features, we also analyzed the characteristics of this particular set of emoticons, report on their evaluation while dividing them into three categories: positive, negative and humorous. We also noticed that

Table 1: Examples of Emoticons Conveying Humor Typical for Chinese Culture.

Emoticon	Humorous {%}	Negative {%}	Positive {%}
🙂	41.7	25.0	33.3
🤧	91.7	8.3	0.0
🐮	83.3	0.0	16.7
🥰	75.0	8.3	16.7

Table 2: Examples of Chinese Emoticon Lexicon.

Emoticon	Textual Feature	Emotion/Implication
😏	[阴险]	"smirking"
🤭	[挖鼻]	"nosepick"
🤮	[污]	"filthy"
😍	[舔屏]	"screen lick"

among the resources of Chinese social media sentiment analysis, the labelled Weibo data containing emoticons are extremely rare which makes considering them in machine learning approaches difficult. To resolve this problem, we applied the emoticons polarity and utilized both lexicons with convolutional neural network (CNN) in a way which allows sentiment analysis on smaller annotated data sets. Our experimental results show that the proposed method can significantly improve the performance for predicting sentiment polarity on Weibo.

2. Related Research

In 2017, Felbo and others [Felbo 17] proposed a powerful system utilizing emoji in Twitter sentiment analysis model called DeepMoji. They trained 1,246 million tweets containing one of 64 common emoticons by Bi-directional Long Short-Term Memory (Bi-LSTM) model and applied it to interpret the meaning behind the online messages. DeepMoji is also one of the most advanced sarcasm-detecting model, sarcasm reverses the emotion of the literal text, therefore sarcasm-detecting capability can play a significant role in sentiment analysis, especially in case of social media. Although sarcasm and irony tend to convey negative emotions in general, we found that in Chinese social media (Weibo in our example), in addition to the expression of positive and negative emotions, people tend to express a kind of humorous emotion that escapes the traditional bi-polarity.

In our early research [Li 18b], we analyzed the usage of 67 emoticons with facial expression used on Weibo. We asked 12 Chinese native speakers to label these emoticons by applying one of three following categories: positive, negative and humorous. We have confirmed that 23 emoticons can be considered more as humorous than positive or negative. On this basis, we applied the emoticons polarity (see Table 1) in a Long Short-Term Memory recurrent neural network for sentiment analysis of undersized labelled data.

Chinese Internet slang is defined as an informal language used to express ideas on the Chinese Internet in response to events, to mass media and foreign cultures. It also expresses a natural human desire to simplify and update language. In [Li 19], we collected 448 frequent Internet slang expressions and created a slang lexicon (examples are shown in Table 2), then we converted the 109 Weibo emoticons into textual features creating Chinese emoticon lexicon (examples are shown in Table 3). To test the influence of slang and emoticons on sentiment analysis task, we also utilized both lexicons with several machine learning-based classifiers for detecting humorous expressions on Chinese social media.

3. Convolutional Neural Network Approach

Inspired by the above mentioned works on Internet slang and emoticons, in this paper, we utilized both lexicons and emoticon polarity with convolutional neural network (CNN) for sentiment classification of Weibo undersized labelled data.

In the first step, we added the Chinese slang lexicon and Chinese emoticon lexicon to segmentation tool for matching new words and emoticons. Then we used the updated tool to segment the sentences of a large data set. Secondly, we applied the segmentation results into the word embedding tool for training word vectors. Then, we applied the word embedding model which considered Internet slang and emoticons to train a CNN model with training data to learn an output representation. Next, we input testing data into deep learning model, and we use a softmax classifier to obtain the predicted results and output their probability. Since we assume that emoticons have relatively greater impact than textual features on emotional expression, we set a hyperparameter for each text and emoticon, and calculate the summation of both features' polarities with the hyperparameter. Finally, we can obtain the sentiment probability of a Weibo post which considers the effect of emoticons and Internet slang.

In order to verify the validity of our proposed method, we performed series of experiments described below.

3.1 Preprocessing

Initializing word vectors with those obtained from an unsupervised neural language model is a popular method to improve performance in the absence of a large supervised training set. For our experiment we collected a large dataset (7.6 million posts) from Weibo API from May 2015 to July 2017 to be used in calculating word embeddings. Firstly, we deleted the images, and videos treating them as noise. Secondly, we applied Chinese Internet slang lexicon and Chinese emoticon lexicon [Li 19] into the dictionary of Python Chinese word segmentation module Jieba*1. Next, we used Jieba to segment the sentences of the microblogs, and applied the segmentation results into the word2vec model [Mikolov 13] for training word vectors. The vectors have dimensionality of 300 and were trained using the continuous skip-gram model.

Next, we collected 4,000 Weibo posts containing ambiguous (🙂, 🐮, 😂, 🤷, 😏, 🤔, 🤧, 🙄) emoticons, ensuring each entry has only one emoticon of given type (cases with more emoticons of the same type were allowed). To use these posts as our training data, we asked three Chinese native speakers to annotate them into three categories: "positive", "negative", and "humorous". After one annotator labelled polarities of all posts, two other native speakers confirmed correctness of his annotations. Whenever there was a disagreement, all decided the final polarity through discussion.

*1 https://github.com/fxsjy/jieba

Table 3: Examples of our Chinese Internet Slang Lexicon.

Type	Examples (Origin)	English Translation
Numbers	233 (哈哈哈)	"laughter"
Chinese contractions	人艰不拆 (人生已经如此的艰难有些事情就不要拆穿)	"Life is so hard that some lies are better not exposed."
Slang derived from foreign language	欧尼酱 (お兄ちゃん)	"O-niichan" ("Brother" in Japanese)

3.2 Convolutional Neural Network

Convolutional neural networks (CNN) utilize layers with convolving filters that are applied to local features [LeCun 98]. Originally invented for computer vision, CNN models have subsequently been shown to be effective for NLP and have achieved excellent results in semantic parsing [Yih 14], search query retrieval [Shen 14], sentence modeling [Kalchbrenner 14], and other traditional NLP tasks.

The equations of the CNN are as follows: $x_i \in \mathbb{R}^k$ is the k-dimensional word vector corresponding to the i-th word in the sentence. A sentence of length n (padded where necessary) is represented as described in [Kim 14]:

$$x_{1:n} = x_1 \oplus x_2 \oplus \ldots \oplus x_n \tag{1}$$

where $\oplus$ is the concatenation operator. In general, let $x_{i:i+j}$ refer to the concatenation of words $x_i, x_{i+1}, ..., x_{i+j}$. A convolution operation involves a filter $w \in \mathbb{R}^k$, which is applied to a window of h words to produce a new feature. For example, a feature c_i is generated from a window of words $x_{i:i+h-1}$ by

$$c_i = f(wx_{i:i+h-1} + b) \tag{2}$$

Here $b \in \mathbb{R}$ is a bias term and f is a non-linear function such as the hyperbolic tangent. This filter is applied to each possible window of words in the sentence $\{x_{i:h}, x_{2:i+1}, \ldots, x_{n-h+1:n}\}$ to produce a feature map:

$$c = [c_1, c_2, \ldots, c_{n-h+1}] \tag{3}$$

with $c \in \mathbb{R}^{n-h+1}$. We then apply a max-overtime pooling operation [Collobert 11] over the feature map and take the maximum value $\hat{c} = max\{c\}$ as the feature corresponding to this particular filter. The idea is to capture the most important feature, one with the highest value, for each feature map. This pooling scheme naturally deals with variable sentence lengths. We experimented with the CNN architecture and a model with 10 epochs and the performance achieved the highest value when the dropout rate was 0.5, the filter size was 32 and strides number was 2. The validity of the model was examined by holdout method (90%/10%, training/validation). The activation functions, we used $RELU$ in general, and the network output activation function was softmax.

3.3 Emoticon polarity

In order to predict sentiment category of Weibo posts considering the influence of emoticons for Chinese social media sentiment analysis, we assign the probability of the deep learning model's softmax output $S(z_i)$ a hyperparameter λ_1. At the same time, we apply the labelled emoticons [Li 18b] as polarity P_e, and assign a hyperparameter λ_2. P becomes the final probability output of the classification:

$$P = \lambda_1 S(z_i) + \lambda_2 P_e \tag{4}$$

where the summation of λ_1 and λ_2 is equal to 1.

Table 4: Comparison F scores results of three CNN approaches.

	Humorous	Negative	Positive
CNN	70.48%	66.67%	56.00%
CNN+Lexicons	70.59%	70.45%	70.58%
CNN+Lexicons+Polarities	**76.84%**	**74.69%**	**73.56%**

3.4 Performance Test

Using a trained word2vec model, we passed word vectors of training data into the three deep learning models to train the model. We collected and annotated 180 Weibo entries with the eight emoticons mentioned above as a testing set, deleting images and videos. Then we used the proposed method to calculate probability of each category and confirmed the precision, recall and F1-score. Because we assumed that in emotion expression emoticons might play a greater role than text, in our experiment, we set the hyperparameters λ_1 and λ_2 to 0.4 and 0.6 respectively. We compared the results of sentiment classification by CNN only, CNN considering Internet slang and emoticons lexicons only, and our proposed methods: CNN model considering Internet slang and emoticons lexicons and emoticon polarities. Table 4 describes the results of F1-score with above methods.

The results proved that our proposed method is more effective than a) convolutional neural network only categorization and b) convolutional neural network approach considering just slang and emoticon lexicons. Limited to small annotated data, the precision of the sentiment classification was relatively low, but by considering Internet slang and emoticons, the F1-score of each classifier outperformed previous method. Our proposed approach has improved the performance showing that low-cost, small-scale data labeling is sufficient to outperform widely used state-of-the-art when emoticon and slang information is added to the learning process.

4. Discussion

In our proposed approach, we focused on emoticons and Internet slang in microblogs and investigated how adding these features separately and together influences the previously proposed method

for recognizing humorous posts which are problematic when it comes to semantic analysis.

Error analysis showed that some posts were wrongly predicted due to ambiguous usage of emoticon which brought clearly negative impact on the results. In Figure 2 we show an example of such misclassification into "positive" category annotated as "humorous" by annotators. 😏 was considered as more positive than humorous by our annotators (67%/0%/33%, positive/negative/humorous). It seems that this particular user wrote a joke just for fun, however, our proposed method was misguided by this "smirking" emoticon. Therefore, we plan to increase the number of evaluators for annotating Weibo emoticons in fine-grained humorous emotion to enhance the reliability of the polarity of emoticons.

Post: 吃饱了就有力气减肥了 😏😏
Pinyin: *Chi bao le jiu you li qi jian fei le* 😏😏
Segmentation: 吃饱了/就/有/力气/减肥/了/[阴险]/[阴险]
Translation: When your stomach is full, you get the strength to reduce weight 😏😏

Figure 2: Example of wrong classification into "positive" category.

5. Conclusions and Future Work

In this paper, we applied the emoticons polarity, Chinese Internet slang lexicon and Chinese emoticon lexicon with a convolutional neural network model for sentiment analysis of undersized labelled data. Our experimental results show that the proposed method can significantly improve the F1-score for predicting sentiment polarity on Weibo.

For improving the performance of our proposed method, in the near future we are going to increase the size of both slang lexicon and training dataset to improve further results. We also plan to test other deep learning approaches to compare the classification results.

6. Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 17K00295.

References

[Li 18a] Li, D., Rzepka, R., Araki, K.: Preliminary Analysis of Weibo Emojis for Sentiment Analysis of Chinese Social Media, Proceedings The 32th Annual Conference of the Japanese Society for Artificial Intelligence, 1J3-04 (2018)

[Peng 17] Peng, H., Cambria,E., Hussain, A.: A review of sentiment analysis research in Chinese language, Cognitive Computation, 2017, 9(4): 423-435 (2017)

[Wang 13] Wang, X., Zhang, C., Ji, Y.: A depression detection model based on sentiment analysis in micro-blog social network, Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, Berlin, Heidelberg, 2013: 201-213 (2013)

[Moschini 16] Moschini, I.: The 'Face with Tears of Joy' Emoji, A Socio-Semiotic and Multimodal Insight into a Japan-America Mash-Up, HERMES-Journal of Language and Communication in Business, 2016 (55): 11-25 (2016)

[Novak 15] Novak, P.K., Smailović, J., Sluban,B. et al.: Sentiment of emojis, PLOS One, 2015, 10(12): e0144296 (2015)

[Guibon 16] Guibon, G., Ochs, M., Bellot, P.: From Emojis to Sentiment Analysis, WACAI 2016 (2016)

[Kulkarni 17] Kulkarni, V., Wang, W.Y.: TFW, DamnGina, Juvie, and Hotsie-Totsie: On the Linguistic and Social Aspects of Internet Slang, arXiv preprint arXiv:1712.08291, 2017 (2017)

[Felbo 17] B. Felbo, A. Mislove, A. Søgaard, et al. "Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm." arXiv preprint arXiv:1708.00524 (2017)

[Li 18b] Li, D., Rzepka, R., Ptaszynski, M., Araki K.: Emoticon-Aware Recurrent Neural Network Model for Chinese Sentiment Analysis, in The Ninth IEEE International Conference on Awareness Science and Technology, iCAST 2018 (2018)

[Li 19] Li, D., Rzepka, R., Ptaszynski, M., Araki K.: A Novel Machine Learning-based Sentiment Analysis Method for Chinese Social Media Considering Chinese Slang Lexicon and Emoticons, The AAAI-19 Workshop on Affective Content Analysis, AffCon 2019 (2019)

[Mikolov 13] T. Mikolov, K. Chen, G. Corrado, et al. "Efficient estimation of word representations in vector space." arXiv preprint arXiv:1301.3781 (2013)

[LeCun 98] LeCun Y., Bottou L., Bengio Y., et al.: Gradient-based learning applied to document recognition, Proceedings of the IEEE, 1998, 86(11): 2278-2324 (1998)

[Yih 14] Yih W., He X., Meek C.: Semantic parsing for single-relation question answering, Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). 2014, 2: 643-648 (2014)

[Shen 14] Shen Y., He X., Gao J., et al.: Learning semantic representations using convolutional neural networks for web search, Proceedings of the 23rd International Conference on World Wide Web. ACM, 2014: 373-374 (2014)

[Kalchbrenner 14] Kalchbrenner N., Grefenstette E., Blunsom P.: A convolutional neural network for modelling sentences, arXiv preprint arXiv:1404.2188 (2014)

[Kim 14] Kim Y.: Convolutional neural networks for sentence classification, arXiv preprint arXiv:1408.5882 (2014)

[Collobert 11] Collobert R., Weston J., Bottou L., et al.: Natural language processing (almost) from scratch, Journal of Machine Learning Research, 2011, 12(Aug): 2493-2537 (2011)

4:40 PM - 5:00 PM (Wed. Jun 5, 2019 3:20 PM - 5:00 PM Room E)

[2E10-13-5] Panel discussion

Panel discussion

©The Japanese Society for Artificial Intelligence

Organized Session | Organized Session | [OS] OS-5

[2F3-OS-5a] 複雑化社会における意思決定・合意形成のための AI技術(1)

福島 俊一（科学技術振興機構）、福田 直樹（静岡大学）、伊藤 孝行（名古屋工業大学）

Wed. Jun 5, 2019 1:20 PM - 3:00 PM Room F (302B Medium meeting room)

[2F3-OS-5a-01] Innovation of Large-scale Consensus Support System based on Agent Technologies:A Large-scale Social Experiment using the Automated Facilitation Agent

〇Takayuki Ito[1], Daichi Shibata[1], Shota Suzuki[1], Naoko Yamaguchi[1], Tomohiro Nishida[1], Kentaro Hiraishi[1], Kai Yoshino[1] （1. Nagoya Institute of Technology）

1:20 PM - 1:40 PM

[2F3-OS-5a-02] (Invited talk) Towards Realization of Collective Intelligence based on Communication-field Mechanism Design

〇Tadahiro Taniguchi[1] （1. Ritsumeikan University）

1:40 PM - 2:00 PM

[2F3-OS-5a-03] (Invited talk) Market design: Designing social mechanisms/institutions based on game theory

〇Makoto Yokoo[1] （1. Graduate School and Faculty of Information Science and Electrical Engineering, Kyushu University）

2:00 PM - 2:20 PM

[2F3-OS-5a-04] A Preliminary Design and Potential Issues of LOD-based Inter-Domain Querying for a Large-Scale Consensus Support System

〇Naoki Fukuta[1] （1. Shizuoka University）

2:20 PM - 2:40 PM

[2F3-OS-5a-05] Opinion dynamics theory considering interpersonal relationship of trust and distrust and media effects

〇Akira Ishii Ishii[1], Yasuko Kawahata[2] （1. Tottori University, 2. Gunma University）

2:40 PM - 3:00 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

エージェント技術に基づく大規模合意形成支援システムの創成：自動ファシリテーションエージェントを用いた大規模社会実験

Innovation of Large-scale Consensus Support System based on Agent Technologies: A Large-scale Social Experiment using the Automated Facilitation Agent

伊藤孝行 *1 Takayuki Ito　柴田大地 *1 Daichi Shibata　鈴木祥太 *1 Shota Suzuki　山口直子 *1 Naoko Yamaguchi　西田智裕 *1 Tomohiro Nishida　平石健太郎 *1 Kentaro Hiraishi　芳野魁 *1 Kai Yoshino

*1名古屋工業大学
Nagoya Institute of Technology

Crowd-scale discussion platforms are receiving great attention as potential next-generation methods for democratic citizen platforms. Such platforms require support functions that can efficiently achieve a consensus, reasonably integrate ideas, and discourage flaming. We are developing several crowd-scale discussion platforms and conducting social experiments with private citizens. The aim of this paper is to present an implementation of the automated facilitation agent that facilitate online crowd discussion and demonstrate the result of our large scale social experiment with Nagoya local government. Our automated facilitation agent extracts discussion structures from text discussions, analyzes them, and posts facilitation messages. The experiment results demonstrate that the automated facilitation agent worked quite well.

1. はじめに

オンライン上のクラウド（Crowd）スケールの議論の支援が次世代の民主主義プラットフォームとして注目を集めている [Malone 18]．Twitter や Facebook などの SNS によって，インターネットで何万人，何百万人という人たちの意見を収集できるようになってきている．これらの意見をうまくまとめ，何百万人という人たちの合意を形成できる可能性がありえる．大規模な合意を形成できれば，これまでには不可能だった，大規模な人数による意思決定が可能になる．

そこで，筆者らは複数の実フィールドでオンラインの議論支援システム Collagree を開発し，幾つかの社会実験を行い有用性を確認している [伊藤 17, 伊藤 15, 伊美 15, Ito 14, Ito 15, Sengoku 16, Ito 18]．例えば，2013 年の名古屋市の市民からの意見集約や愛知デザインリーグなどの取り組みである．ここでは，炎上を防ぎながら議論を適切にリードするために人間のファシリテータがオンライン議論をファシリテートした．Collagree では，人間のファシリテータを支援するための様々な技術的工夫を行なった．

課題は，規模が非常に大きいことから，人手でファシリテーションを行うのが困難な点である．オンラインで２４時間休むことなく議論を管理することは非常に難しい．さらに，議論の内容が複雑になればなるほど，それを追い続けるのは多大な負担があることがわかっている．そこで本研究では，自動ファシリテーションエージェントを実装することで，大規模な人数の人たちの意見を効率的に収集し，合意を形成するシステムを実現する．

本稿では，第２章で大規模合意形成支援システム D-Agree について述べる．第３章で自動ファシリテーションエージェントの実装について説明し，第４章では 2018 年の名古屋市との共同社会実験の概要とその結果について述べる．最後に第５章で本論文をまとめる．

連絡先: 伊藤孝行，名古屋工業大学，名古屋市昭和区御器所町名古屋工業大学，052-735-7968, ito.takayuki@nitech.ac.jp

2. 大規模合意形成支援システム D-Agree

本研究で実装した大規模合意形成支援システム D-Agree は，ウェブ上の掲示板型の議論支援システムである．図 1 に概要を示す．システムでは，テーマ毎に掲示板が用意されており，参加者はテーマ毎にフリーテキストを入力し議論を行う．自動ファシリテーションエージェントは，テキストから議論の構造を抽出し，抽出した構造に基づいてファシリテーションを進める．議論の構造としては IBIS(Issue-Based Information System) という創造的かつ建設的な議論を進めるための構造を採用している．議論を活性化するための議論ポイント [伊美 15] や，議論において頻出する単語を表示するキーワードクラウド機能などが掲示板の機能として提供される．

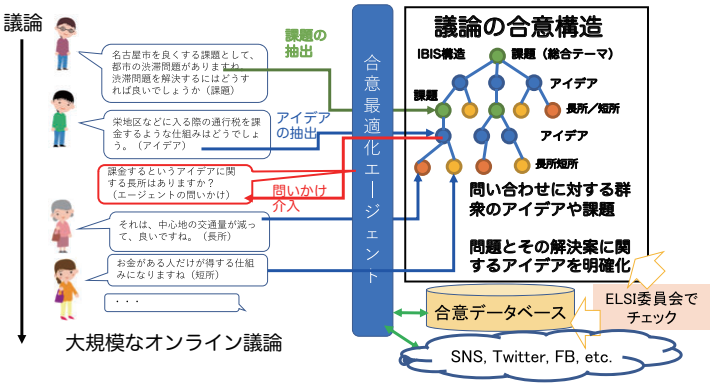


図 1: 大規模合意形成システムの概要

IBIS 議論構造は，Issue，Position(Idea), Argument(Pros, Cons) からなる議論構造である [Kunz 70, Conklin 89]．Issue（課題・論点）を表現するクェリが構造として含まれていることが特徴である．テーマに対する様々な論点をあげ，それに対する個々人の Position（視点）をあげさせ，それに基づいて Argument（討論）を進めるという基本的な仕組みになっている．オンライン議論の実験を数多く行なっているが，経験的に，議論に関してより高いリテラシのある参加者の議論は自然に

IBIS のような議論構造になり，そうでない場合はどちらかというと Twitter のツィートのような形になる．dd 自動ファシリテーションエージェントは，抽出した IBIS 構造に基づいて，ファシリテーション投稿を行う．例えば，参加者からアイデアが投稿されたと判定した時は，そのアイデアの長所（pros）や短所（cons）を尋ねるなどの投稿を行う．

3. 自動ファシリテーションエージェント

3.1 概要

自動ファシリテーションエージェントの基本機能は，IBIS の構造に基づいて，議論中の文をベースに構造化する．構造化した議論に基づいて，ファシリテーション投稿を生成し投稿する．また，炎上を引き起こす可能性のある不適切な発言のフィルタリングも行う．

3.2 アノテーション

上記の深層学習に基づく抽出機構のトレーニングデータとして，実際に議論を行なってデータを獲得した．議論データの詳細は次の通りである：合計議論回数 38 回，タグ付けされた文節の合計数 4,972，平均議論システム開放時間 7.81 時間 (15 分〜144 時間)，平均議論参加者数 13 名 (4〜115 名)，平均投稿数 101.8 ／回 (ファシリテータ発言含む) ．議論参加者はおもに大学生，ファシリテーションは研究員が行った．議題は，教育，災害対策，環境保全，観光推進等多岐にわたる．38 回の議論データのアノテーションは，2 名の研究員および 3 名の大学生・大学院生によって行われた．2 名の研究員が，ラベルの定義考察等を含むアノテーション作業の全体管理を行った．アノテーション作業行程の詳細については論文 [山口 19, Yamaguchi 18] を参照されたい．

3.3 議論構造抽出

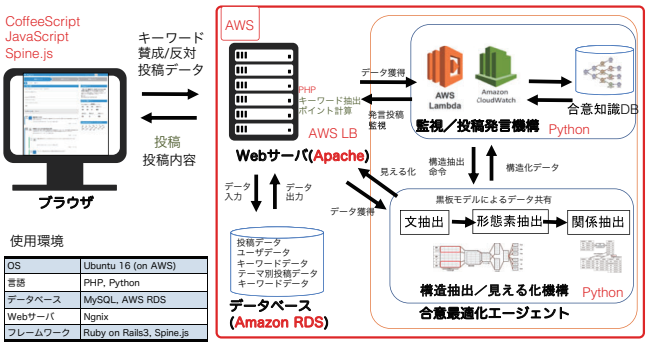
IBIS の基本構造の 4 コンポーネントとそのリンク関係について議論構造の抽出を行なった．手法としては深層学習の一つである BiLSTM を用い，コンポーネントのラベル判定を行ない，リンクについては，回帰としてリンク先を予測している．精度としては，コンポーネントのラベル判別で F 値で 80%以上，リンクの予測で 90%以上を得ている．詳細は論文 [Suzuki 19] を参照されたい．本成果は Argumentation Mining の先端技術 [Stab 17] に匹敵する精度であり，Argumentation Mining の技術としても重要である．

3.4 不適切発言フィルタリング

議論において，不適切な言葉による発言は，公序良俗に反するだけでなく，炎上の原因にもなる．そこで，本システムでは，不適切な言葉についてのフィルタリングを行い，発言を一旦非表示にする機能を実装している．非表示にされた発言は，管理者によって再表示することができる．フィルタリングアルゴリズムとしては，doc2vec を用いて文書をベクトル化し，コサイン類似度を用いたアンサンブルでの分類を行なっており，既存研究を上回る精度でのフィルタリングを実現している．本実験の学習データとして，Collagree や D-Agree 上の投稿データ，また大規模オンライン掲示板である 5 ちゃんねる掲 示板からデータを収集した．正常発言を 3,200 件，性的発言を 2,400 件，暴力的発言を 1,920 件，合計 7,520 件の投稿文書を用いて，学習データの生成を行なった．結果として，F 値は 0.962 を示している．詳細は論文??を参照されたい．

3.5 AWS での実装

図 2 にシステムアーキテクチャを示す．本システムは，AWS 上に実装されており世界中から利用することができる．エージェントは，構造抽出部と投稿生成部から構成される．構造抽出部は上で述べた構造抽出機構が利用されている．投稿生成部は 200 以上の投稿ルールにを用いて，抽出された議論構造と状況から投稿を生成する．AWS の CloudWatch など各機能を効率的に機能を利用し，スケーラブルにエージェントが動作する仕組みを実現している．詳細は論文 [柴田 19] を参照されたい．



使用環境

OS	Ubuntu 16 (on AWS)
言語	PHP, Python
データベース	MySQL, AWS RDS
Webサーバ	Ngnix
フレームワーク	Ruby on Rails3, Spine.js

図 2: システムアーキテクチャ

4. 名古屋市との大規模社会実験

4.1 実験概要

自動ファシリテーションエージェントを導入した合意形成支援システムを用いて名古屋市次期総合計画の中間案に対する市民からの意見集約についての社会実験を行なった．期間は 2018 年 11 月 1 日から 2018 年 12 月 7 日までである．

本意見集約では，名古屋市の将来像を考える上での都市像として 5 つのテーマについて意見を集めた．この 5 つのテーマについては，名古屋市が行政としてまとめたものであり，次期総合計画の根幹をなすテーマである．名古屋市の次期総合計画は，名古屋市の事業にも直結するものであり，重要なテーマが慎重に定義されている． 5 つのテーマは，図 3 の右側にある通り,「テーマ 1 ：人権が尊重され，誰もがいきいきと暮らし，活躍できるまち」,「テーマ 2 ：安心して子育てができ，子供や若者が豊かに育つまち」,「テーマ 3 ：人が支え合い，災害に強く安心・安全に暮らせるまち」，および「テーマ 4 ：快適な都市環境と自然が調和したまち」,「テーマ 5 ：魅力と活力にあふれ、世界から人や企業をひきつける、開かれたまち」である．

5 つのテーマについて，テーマ 1 とテーマ 2 は，人間のファシリテータがファシリテーションを行なった．テーマ 3 とテーマ 4 は，自動ファシリテーションエージェントがファシリテーションを行なった．テーマ 5 については，人間のファシリテータと自動ファシリテーションエージェントが共同でファシリテーションを行なった．

2018 年 11 月 1 日から最初の 1 ヶ月は市民にシステムを使って自由に議論していただいた．2018 年 12 月 1 日から 1 週間は，名古屋市との協議の上，投稿された議論を人間がまとめ，まとめた内容に対して参加者から支持表明をいただくこととした．議論した結果は IBIS の構造に基づいて整理されており，まとめをする上での支援が可能であった．

4.2 結果の分析

表 1 に各テーマの投稿の数を示す．それぞれ，合計，ファシリテータ（人間 FA と自動 FA），および参加者の投稿数を示す．ここで特に参加者の投稿数に注目すると，テーマ 3，テー

図 3: 名古屋市実験のチラシ

マ４，およびテーマ５の投稿数が，テーマ１とテーマ２の投稿数より多いことがわかる．すなわち，今回の実験では，自動ファシリテーションエージェントがファシリテーションをしたテーマの方が，参加者の投稿を増加させる傾向にあったことがわかる．

表 1: 実験結果：投稿の数

テーマ	投稿			
	合計	人間FA	自動FA	参加者
1: 人間FA	81	43	0	38
2: 人間FA	56	21	0	35
3: 自動FA	88	0	24	64
4: 自動FA	70	0	18	52
5: 自動&人間 FA	137	17	21	99
合計	432	81	63	288

図４に，アンケート結果の１つを示す.「時期総合計画に関する議論に満足できましたか？」という，本社会実験におけるオンラインの議論に満足できたか，について問う質問であった．20 の有効回答があり，図４に示すように，自動ファシリテーションエージェントによるファシリテーションが行われたテーマ３，テーマ４，およびテーマ５について，満足度は，人間のファシリテータによってファシリテーションが行われたテーマ１とテーマ２と同等の結果が得られている．

4.3 自動ファシリテーションの事例

図５に，本社会実験において，本システムで自動ファシリテーションが実際に行われている事例の一つを示す．ここでは，課題（課題１）を投稿しており，これを自動ファシリテーションエージェントが「課題」と認識した上で,「解決するために何が必要でしょうか」と問い合わせている．それに対して，参加者がアイデア（アイデア１）を投稿している．

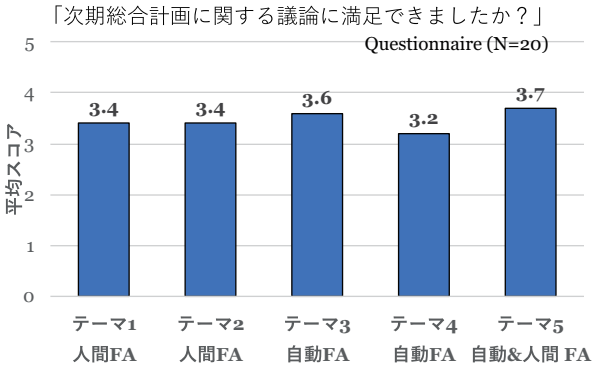


図 4: 実験結果：アンケート結果

5. まとめと今後の課題

本論文では，オンラインの大規模合意形成支援システムのための自動ファシリテーションエージェントの実装と，その実践的応用として，名古屋市との協力による名古屋市次期総合計画の中間案に対する市民からの意見集約実験について示した．自動ファシリテーションエージェントの実装は IBIS 構造に基づいて，実際の議論の意味的構造を抽出し，構造に基づきファシリテーションの投稿を行うことができる．今回は，慎重なアノテーション，深層学習モデルのチューニング，フィルタリング機能，AWS を用いたスケーラブルなシステム実装，によって，実際の市民を対象とした社会実験に，知能処理技術によるファシリテーションが可能であることを示すことができた．大規模合意形成支援は次世代の民主主義プラットフォームにもなりえる上，知能処理技術がどの程度人間社会の意思決定に関与すべきかという観点から ELSI についても議論を進めている [浜田 18]．今後さらに，自動ファシリテーションエージェントが議論に関与する際どの程度まででであれば社会は受容するのか，様々な条件があるが調査を進める．

謝辞

本論文で紹介した研究内容は，JST CREST「エージェント技術に基づく大規模合意形成支援システムの創成：代表伊藤孝行」（グラント番号 JPMJCR15E1）に支援を受けている研究の一部である．本論文の研究内容に寄与していただいているメンバー全員に感謝の意を表する．

参考文献

[Conklin 89] Conklin, J. and Begeman, M. L.: gIBIS: A tool for all reasons, *Journal of the American Society for Information Science*, Vol. 40, No. 3, pp. 200–213 (1989)

[Ito 14] Ito, T., Imi, Y., Ito, T., and Hideshima, E.: COLLAGREE: Facilitator-mediated Large-scale Consensus Support System, in *The 2nd Collective Intelligence Conference* (2014)

[Ito 15] Ito, T., Imi, Y., Sato, M., Ito, T., and Hideshima, E.: Incentive Mechanism for Managing Large-Scale Internet-Based Discussions on COLLAGREE, in *The 3nd Collective Intelligence Conference* (2015)

図 5: 自動ファシリテーションの成功事例の 1 つ

[Ito 18] Ito, T.: Towards Agent-based Large-scale Decision Support System: The Effect of Facilitator, in *The 51st Hawaii International Conference on System Sciences (HICSS2018)* (2018)

[Kunz 70] Kunz, W. and Rittel, H. W.: Issues as elements of information systems, Technical report (1970), CiteSeerX 10.1.1.134.1741

[Malone 18] Malone, T. W.: *Superminds: The Surprising Power of People and Computers Thinking Together*, Little, Brown and Company (2018)

[Sengoku 16] Sengoku, A., Ito, T., Takahashi, K., Shiramatsu, S., Ito, T., Hideshima, E., and Fujita, K.: Discussion Tree for Managing Large-Scale Internet-based Discussions, in *Collective Intelligence 2016* (2016)

[Stab 17] Stab, C. and Gurevych, I.: Parsing argumentation structures in persuasive essays, *Computational Linguistics*, Vol. 43, No. 3, pp. 619–659 (2017)

[Suzuki 19] Suzuki, S., Yamaguchi, N., Nishida, T., Moustafa, A., Shibata, D., Yoshino, K., Hiraishi, K., and Ito, T.: Extraction of Online Discussion Structures for Automated Facilitation Agent, 2019 年度 人工知能学会全国大会 (第 33 回) (2019)

[Yamaguchi 18] Yamaguchi, N., Ito, T., and Nishida, T.: A Method for Online Discussion Design and Discussion Data Analysis, in *The Thirteenth International Conference on Knowledge, Information and Creativity Support Systems (KICSS-2018)* (2018)

[伊藤 15] 伊藤 孝行, 奥村 命, 伊藤 孝紀, 秀島 栄三：多人数ワークショップのための意見集約支援システム Collagree の試作と評価実験:〜議論プロセスの弱い構造化による意見集約支援〜, 日本経営工学会論文誌, Vol. 66, No. 2, pp. 83–108 (2015)

[伊藤 17] 伊藤 孝行, 藤田 桂英, 松尾 徳朗, 福田 直樹：エージェント技術に基づく大規模合意形成支援システムの創成—自動ファシリテーションエージェントの実現に向けて—, 人工知能学会誌, Vol. 32, No. 5 (2017)

[伊美 15] 伊美 裕麻, 伊藤 孝行, 伊藤 孝紀, 秀島 栄三：オンラインファシリテーション支援機構に基づく大規模意見集約システム COLLAGREE—名古屋市次期総合計画のための市民議論に向けた社会実装, 情報処理学会論文誌, Vol. 56, No. 10, pp. 1996–2010 (2015)

[山口 19] 山口直子, 西田智弘, 柴田大地, 鈴木祥太, 芳野魁, 平石健太郎, 伊藤孝行：IBIS 構造に基づく議論モデルにおけるアノテーション手法の提案, 2019 年度 人工知能学会全国大会 (第 33 回) (2019)

[柴田 19] 柴田 大地, 山口 直子, 西田 智弘, 鈴木 祥太, 芳野 魁, 平石 健太郎, 伊藤 孝行：オンライン上における大規模議論支援のための IBIS 構造に基づく自動ファシリテータの実装, 2019 年度 人工知能学会全国大会 (第 33 回) (2019)

[浜田 18] 浜田 良樹, 松尾 徳朗, 伊藤 孝行：ELSI 委員会による合意形成システムと社会との融和, 人工知能学会論文誌, Vol. 33, No. 5, pp. 1–4 (2018)

1:40 PM - 2:00 PM (Wed. Jun 5, 2019 1:20 PM - 3:00 PM Room F)

[2F3-OS-5a-02] (Invited talk) Towards Realization of Collective Intelligence based on Communication-field Mechanism Design

〇Tadahiro Taniguchi[1] （1. Ritsumeikan University）
Keywords: Communication field, Mechanism design

本招待講演では多人数のコミュニケーションを間接的に制御するコミュニケーション場のメカニズムデザインの概念について紹介する。複数の人間が参画し話し合い、会議、相談等をしている状況は、自律的分散的なマルチエージェントシステムとして捉えられる。各主体はあくまで自らの効用に基づいて行動するために、その振る舞いを誘導するためには適切な誘引や制約の設計が重要である。ビブリオバトルや発話権取引などの事例に触れながら概説する。

2:00 PM - 2:20 PM (Wed. Jun 5, 2019 1:20 PM - 3:00 PM Room F)

[2F3-OS-5a-03] (Invited talk) Market design: Designing social mechanisms/institutions based on game theory

〇Makoto Yokoo[1] （1. Graduate School and Faculty of Information Science and Electrical Engineering, Kyushu University）
Keywords: game theory, mechanism design, market design, multiagent systems

"Market Design" is a research field that examines how to design a new market or improve an existing market such that a certain design goal is satisfied. This research field is influenced by micro economics, in particular, game theory. Here, the meaning of a "market" is very broad; it includes a spectrum auction, in which a government allocates the licenses to use specific spectrum bandwidth to companies, or a market without monetary transfer, such as a kidney exchange program, or a school choice program, in which children/parents can choose public schools they want to attend. Now, market design has become a interdisciplinary research topic that is relevant to computer science and information systems. In this lecture, I will describe an overview of market design research.

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

合意形成過程の蓄積・知識化のための LOD 技術に基づく分野横断型議論過程検索機構の設計と課題

福田 直樹 *1
Naoki FUKUTA

*1静岡大学 学術院情報学領域
Colledge of Informatics, Shizuoka University

In this paper, we present our preliminary idea and design issues of a mechanism with a set of software and tools for querying online discussions and arguments with linked open data accesses over different domains to make it as a great knowledge repository. The prototype design is based on Linked Data and Knowledge Graph structure and it includes mechanisms which aim to realize queries over different domains for the stored structured discussion data.

1. 議論過程の分野を超えた知識化の重要性

実世界とインターネット上などのオンライン上の議論とをあわせたハイブリッドな議論について，その実現についての試み [Ito 16] や，その実際の活用の事例に関する考察が行われてきている [Kawase 18]．すでに文献 [福田 18] でも述べてきたが，それら 1 つ 1 つの議論には，興味深く目を向けるべきものが多くあると考えられ，ある 1 つの大きなテーマ内であれば，ファシリテーターなどをうまく動員してそれらをうまく結びつけられる可能性も十分に考えられる．一方で，その議論をさらに深め確実なものとするには，時期やテーマを超え，議論の過程やその際の情勢，その後の経過などを含めて，より広い範囲での議論の参照が重要となってくる [福田 18] 点は，本研究におけるこれまでに一貫した視点である．特に，その議論の背景に深い歴史的背景やそれに起因する対立などがある場合には，その重要性大きいと考えられ [福田 16][福田 17b]，そのための議論過程の蓄積と知識化のための基盤の検討を進めてきた [福田 18]．

すでにこれまでの検討経過でも述べてきたとおり [福田 18]，文脈に基づいて社会的合意のための議論を深いレベルで行うためには，過去の社会的合意形成問題における種々の議論がそこに適切に継承・参照されることが望ましい [武田 13] と考えられる．一方で，それらの議論の重要な部分には，その対象となる分野や時代に固有の事象の扱いが，深く関わってくる場合があるのではないかと本研究では考えている．社会的合意を目指す当事者の個人的な利害関係や立場が原因となって，武田の指摘するような「囚人のジレンマ的な状況で相互の議論を噛み合う状態になるようにお互いが譲歩できない状況」[武田 13] を正確に理解するためには，そこでの固有な利害関係やその背景理解しようとせず，抽象的な議論のみで理解を行うアプローチには，限界があるのではないかと考える．そうした固有の事情を省略することなく，別の分野や時代に議論された場合ではそれらの固有な事情がどういうものにマッピングされるのかという，具体的な対象をその検討に含めた状態での，議論内容の相互参照とそこからの知識の活用が，重要であると考えられる．これらが，本研究の重要な動機付けとなっている．

連絡先: 福田直樹，静岡大学 学術院情報学領域，432-8011 浜松市中区城北 3-5-1，fukuta (at) inf.shizuoka.ac.jp

2. 議論過程の知識化に向けた機構設計の拡張と課題

本研究で拡張の対象となる，議論過程の知識化に向けた課題に対応するための基盤は，こらmでに設計を進めてきた社会的合意のための議論の過程に対する関係性の付与・閲覧支援機構 [福田 17a] であり，さらにそれを複数の異なる議論過程に拡張し，議論過程や関連するデータの閲覧支援に適用可能にするための機構の設計の概要について，文献 [福田 17b] および文献 [福田 18] で述べている．

議論過程の閲覧支援としては，すでに文献 [福田 17b] において，その初期の設計の概要を示している．本研究で主に扱う議論過程は基本的に [福田 17b] と同様の前提をおいている．その前提とは，議論が，その議論当初には実空間における対面での議論が組み合わされる場合もあるが，発言の投稿とそれに対する返信という構造を基本として，Web 上におけるオンライン議論として構成され，それらに「いいね」等の付加情報が付与される形で構成される [Ito 17]，ということである．また，その議論の過程では，ファシリテーターと呼ばれる議論の整理や進展を促す役割のユーザにより，通常のユーザとは異なる権限により議論の方向性などに対する調整が行われる場合があることも，前提とする．

長期的に議論過程を蓄積していった場合には，これらの前提が時間の経過や環境の変化によって異なるものに変容・多様化していくことが想定される．そうした変化に寛容な閲覧支援技術の実現が，重要な課題の 1 つとなる [福田 18]．そうであるがゆえに，それらの中の相違点を，その議論の前提や対象が異なるもの同士の間でも，何らかの説明可能な形で取り出して検討できることには，意義があると考える．

そのためのアイデアの 1 つとして，本研究では，ある議論から一部分を取り出すという操作の説明を抽象的なクエリ言語で表現することを考える．たとえば，その議論過程が，Linked Data（あるいは Knowledge Graph）の形式で表現できると仮定する場合には，そのための標準クエリ言語である SPARQL がそのクエリ表現言語の選択肢として考えられる．ここで，そのクエリを直接的に分野に横断的になるように抽象化するというのではなく，クエリ自身を必要に応じて別の分野・オントロジーに基づく対象へと変換できるようにすることで，そのクエリの持つドメイン固有の詳細性を維持しながらも，それらを相互に比較検討可能にすることが，できるのではないか．それが，本研究の狙いである．

図 1 は，ある議論過程における議論関係を，クエリ表現へ

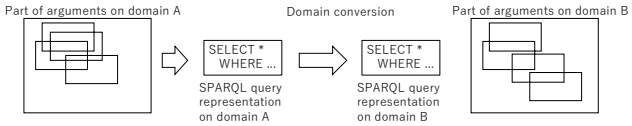


図 1: 議論過程のクエリ表現への変換を通じたドメイン変換

の変換を通じて別の議論ドメインに変換し利用する過程の概念図を示している．

この実現に当たっては，まず，クエリの表現自身を容易に行えるようにするために，これまでに共同で研究を進めてきている，クエリ記述支援における動的な概念マッピング生成に基づく支援技術（たとえば，[Adachi 17]，[Adachi 18] など）の活用が考えられる．これに加えて，すでに記述されたクエリを，別の分野・オントロジーで構成された対象に対するクエリに変換し，その特性をできるだけ維持し相互に関連づけた状態とするための手法も，開発を進めてきている [Adachi 18]．

議論内に見られる要素間の種々の関係を踏まえた閲覧支援は重要であり [福田 17b]，本アプローチでは，その一つの方法として，別々の分野・時期に行われた議論同士の関連性を効果的にかつ説明可能な形で取り出して再利用することを狙っているが，その実現に当たっては，すでに文献 [福田 18] でのべたような，議論過程の効果的な蓄積と再利用性向上のためのアノテーションの品質の管理と向上が重要となる．Carroll らによって提案された関係性そのものに ID を付与する手法 [Carroll 05] と併用することで，アノテーションに対するメタなアノテーションが可能となる [福田 17a] ことはその一助となるかもしれないが，アノテーションの複雑化は同時にそれを高い品質で付与することの難しさにもつながると考えられる．その関連づけに対するメタなアノテーションの基盤となるオントロジーそのものも，ニーズや社会の変化にあわせて深化し，多様化することが考えられる [福田 18]．そのようなオントロジーの変更や変化への許容性と高速な計算可能性の両立は，本研究でも継続的に重要な課題の 1 つとなる．

3. おわりに

議論などの理解の支援では，論証内のミクロな係り受け関係をより上位の抽象度の関係でとらえていくことと同時に，それぞれの信念や立場に基づいたときに，それらがどのように見えるのかを理解していくことが，重要な一助となると考えられることは，これまでも指摘してきた [福田 18]．このためには，本研究で試みるような，ことなる場面で行われた議論どうしを近い視点から再構成するようなアプローチと同時に，同一の議論に対しての異なる視点からの議論の見え方に関しての論理的な構造をふまえた分析・閲覧支援を同時に行えるようにしていきたい．効果的なアノテーションの実現と，それに基づく論理的矛盾なども踏まえた分析の支援を行えるようにすることは，今後の課題である．

謝辞

本研究の一部は，JST CREST JPMJCR15E1 の支援を受けたものである．

参考文献

[Adachi 17] Adachi, T. and Fukuta, N.: A Mapping-enhanced Linked Data Inspection and Querying Support System using Dynamic Ontology Matching, in *Proc. of 2nd International Workshop on Platforms and Applications for Social problem Solving and Collective Reasoning(PASSCR2017)*, pp. 1191–1194 (2017)

[Adachi 18] Adachi, T. and Fukuta, N.: A Query Anonymization Approach using Ontology Mappings, in *Proc. of Joint International Workshop on Practical Application of Ontology for Semantic Data Engineering(PAOS2018) and Platforms and Applications for Social Problem Solving and Collective Reasoning(PASSCR2018)*, pp. 21–32 (2018)

[Carroll 05] Carroll, J. J., Bizer, C., Hayes, P., and Stickler, P.: Named Graphs, Provenance and Trust, in *Proc. of the International World Wide Conference (WWW2005)*, pp. 613–622 (2005)

[Ito 16] Ito, T., Otsuka, T., Kawase, S., Sengoku, A., Shiramatsu, S., Matsuo, T., Oishi, T., Fujita, R., Fukuta, N., and Fujita, K.: Preliminary Results on A Large-scale Cyber-Physical Hybrid Discussion Support Experiment, in *Proc. of 11thInternational Conference on Knowledge, Information and Creativity Support Systems (KICSS 2016)* (2016)

[Ito 17] Ito, T., Otsuka, T., Kawase, S., Sengoku, A., Shiramatsu, S., Ito, T., Hideshima, E., Matsuo, T., Oishi, T., Fujita, R., Fukuta, N., and Fujita, K.: Experimental Results on Large-scale Cyber-Physical Hybrid Discussion Support, *International Journal of Crowd Science*, Vol. 1, (2017)

[Kawase 18] Kawase, S., Ito, T., Otsuka, T., Sengoku, A., Shiramatsu, S., Matsuo, T., Oishi, T., Fujita, R., Fukuta, N., and Fujita, K.: Cyber-Physical Hybrid Environment using a Largescale Discussion System Enhances Audiences' Participation and Satisfaction in the Panel Discussion, *IEICE Transaction on Information and Systems*, Vol. E101-D, No. 4, pp. 847–855 (2018)

[武田 13] 武田 徹：原発議論はなぜ不毛なのか, 中央公論新社 (2013)

[福田 16] 福田 直樹：合意形成支援システムの高度化のための LOD 連携機構の概念設計, 2016 年人工知能学会全国大会講演論文集, pp. 1N4–OS–19a–2 1–2 (2016)

[福田 17a] 福田 直樹：議論に基づく合意形成支援における議論構造を考慮した議論過程 LOD の閲覧支援機構の設計, セマンティック Web とオントロジー研究会, pp. SIG–SWO–041–07 1–2 (2017)

[福田 17b] 福田 直樹：合意形成過程の再利用に基づく LOD 技術による議論過程検索提示機構の設計, 2017 年人工知能学会全国大会講演論文集, pp. 1P3–4 1–2 (2017)

[福田 18] 福田 直樹：合意形成過程の蓄積・知識化のための LOD 技術に基づく議論過程クラウドの設計と課題, 2018 年人工知能学会全国大会講演論文集, pp. 1D3–OS–28b–05 (2018)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

信頼・不信とメディア効果を考慮したオピニオンダイナミクス理論

Opinion dynamics theory considering interpersonal relationship of trust and distrust and media effects

石井　晃 *1　　川畑泰子 *2
Akira Ishii　　Yasuko Kawahata

*1鳥取大学　　*2群馬大学
Tottori University　　Gunma University

Bounded Confidence Model is well-known for opinion dynamics theory that considers opinions as continuous values. However, this does not include confrontation of people. In this paper, we express a theory which expresses the relation of trust and distrust between people by the sign of interpersonal relationship coefficient and extends the value of opinion on one dimensional axis from positive to negative. In addition, the influence on the mass media was also taken into consideration, and efforts were made to consider. In this paper, we will show examples of 2, 3, 20 and 300 people in computational simulation. The distribution of opinions calculated from this calculation may be comparable to the distribution of opinions measured on social networks.)

1.　はじめに

オピニオンダイナミクスは社会（あるいは少数のグループ内）での意見の合意を目指すプロセスを解析する理論として古くから様々な面から研究されている。 [?, ?, ?, ?, ?, ?]. 意見の合意にゲーム理論を使う試みも行われた [?]。オピニオンダイナミクス理論の総合報告も参照されたい [?]。

定量的な解析に向いたオピニオンダイナミクスの理論は、大きく分けて 2 つの流れがある。1 つの手法は意見を 1 と 0、あるいは 1 と－1 の二値とした離散的な意見を扱う理論である。アメリカやフランスの大統領選挙、あるいは BREXIT に見られる国民投票などは投票でかならずどちらか一方に絞って投票するため、この二値の理論は応用しやすい。もう 1 つの手法は意見を一次元（あるいは多次元）の連続値で考える理論である。合意形成などはこちらのやりかたで考えることが多い。

離散的な二値の理論としては、代表的なものに Galam による磁性物理学理論の Ising 模型を応用した理論がある [?, ?, ?]。この理論では、人々の意見は 1 人 1 人の意見が I_i で表せるとし、意見交換による相互作用を $-J\sum_i\sum_j I_iI_j$ と表し、これが社会の conflict になるとする。conflict が低くなるように社会の動力学が動き、$J>0$ ならば、隣り合う人同士の意見が同じに揃う方が conflict は低くなる。

また、理論物理学の繰り込み群の考え方を利用した局所多数決モデルも賛成と反対の二値である [?, ?]。この理論で、Galam はイギリスの EU 離脱（BREXIT）の国民投票 [?] やトランプの大統領当選 [?] の解析を発表している。

一方、意見を連続値とした理論の代表的なものに Deffuant-Weisbuch Model や Heselmann-Krause Model として知られる Bounded Confidence Model が挙げられる [?, ?, ?] 。Deffuant-Weisbuch Model の拡張として意見の対立による 2 極化を扱った試みはあるが [?, ?, ?]、Hegselmann-Krause Model では合意形成のみを念頭に置いている。意見は 0 から 1 までの連続的な値を取るとするが、 1 は賛成で、 0 は無関心であり、反対意見を想定していないことに注意されたい。Deffuant-Weisbuch Model の拡張 [?, ?, ?] ではN人の各人相互の影響の受け方は一定であったり、お互いの意見の相違の度合いから決まるとしていて、各人が元々固有に信頼していたり、反発を抱いていたりという要素はあらわに入っていない。

本研究では、理論的にN人の相互の意見交換を設定しやすい Hegselmann-Krause のモデル [?] に、人々が反発する要素やマスメディアの影響などを含める形で拡張する。

連絡先: 石井晃, 鳥取大学, 鳥取市湖山町南 4-101, 0857315629, ishii@tottori-u.ac.jp

2.　理論

我々の数理モデルは、意見を連続値とした理論の代表的な bounded confidence model[?] から出発する。 N 人の人を想定し、 $1\le i\le N$ として時刻 t でのエージェント i の意見を $I_i(t)$ と書く。 Hegselmann-Krause [?] の理論では、エージェント i の意見は次のように定義される。

$$I_i(t+1)=\sum_{j=1}^{N}D_{ij}I_j(t) \tag{1}$$

ここで、$I_i(t)$ の値域は $0\le I_i(t)\le +\infty$ である。Hegselmann-Krause の理論では、ここで N 人の中の全ての組み合わせ i,j について、係数 $D_{ij}\ge 0$ である。 彼らの定義では $D_{ij}=0$ はエージェント i の意見はエージェント j の意見の影響は受けないことを意味する。なお、この Hegselmann-Krause の理論では意見を強く持つ人から無関心な人までの間での合意形成を念頭に置いていた。

現実の社会において人々の中に対立が無いことは存在しない。より一般的に社会の中の意見交換では、全員が合意形成する場合だけでなく、中には反発する人もいて、対立が存在するのが普通である。昨年、石井ー川畑によって対立を採り入れた理論が提出された [?]。 石井ー川畑の理論では社会の各自の意見 $I_i(t)$ は他の人の意見に影響され、またマスメディアの影響 $A(t)$ を受けるとして、次の形になる。

$$\frac{dI_i(t)}{dt}=c_iA(t)+\sum_{j=1}^{N}D_{ij}I_j(t) \tag{2}$$

ここで係数 D_{ij} はエージェント i から見たエージェント j に対する信頼度である。

石井ー川畑の理論 [?] では、意見は賛成と反対の二値でも Bounded Confidence Model での $0\le I_i(t)\le +\infty$ でもなく、

$I_i(t)$ の値域は $-\infty \leq I_i(t) \leq +\infty$ と拡張し、正の値と負の値がそれぞれある問題に関する賛成と反対に相当するとする。

また、人と人との意見交換の際にどのような影響を受けるかを定義する係数 D_{ij} は人と人との間の信頼と不信を表すとする。即ち $D_{ij} > 0$ であれば、エージェント i はエージェント j を信頼していて、$D_{ij} < 0$ であればエージェント i はエージェント j に不信感を抱いていると考える。また、係数 D_{ij} は非対称とし、エージェント i から見たエージェント j の信頼・不信とエージェント j から見たエージェント i の信頼・不信は異なる数値で、片方が正でもう片方は負というのはあり得るとする。実際の人間関係でそれは十分にあり得るであろう。

先に提出された石井ー川畑の理論 [?] では、実際には同調や反発が時間とともに、無限に進行する欠点がある。例えばエージェント A とエージェント B が反発する場合、どこまでも反発して両者の意見は正の無限大と負の無限大に発散してしまう。実際にはあまりにも意見が違う場合、お互いに無視するであろう。また、もう一つの欠点として、自分と同じ意見の人と共鳴する面も考慮する。エージェント A とエージェント B がまったく同じ意見の場合、石井ー川畑の理論 [?] では両者がどんどん正の意見の方向（あるいは負の意見の方向）に発散していく。実際には同じ意見であれば、それほど影響は受けないであろう。これは謝辞にも記したパリ政治学院の S Galam 教授からも指摘された。そこで、 式（**??**）の $D_{ij}I_j(t)$ を次のように変更して上に述べた２つの足りない効果を採り入れれた石井の修正が提案された [?]。

$$D_{ij}\Phi(I_i, I_j)(I_j(t) - I_i(t)) \tag{3}$$

$$\Phi(I_i, I_j) = \frac{1}{1 + exp(\beta(|I_i - I_j| - b))} \tag{4}$$

ここで $\Phi(I_i, I_j)$ は $|I_i - I_j| = b$ 付近でスムーズにカットオフが働く関数であり、統計物理学で言うフェルミ分布関数と同じ関数形である。さらに $I_j(t) - I_i(t)$ によって、 $I_i(t)$ は自分と同じ意見 $I_j(t)$ には影響されないという影響が入っている。

マスメディアの発信や社会のフレームワークを形成する発信者 (行政、公的機関など) の情報は社会生活をする私たちの中で伝播する。本論文の理論ではそうしたメディアの影響はヒット現象の数理モデル [?, ?] にならって採り入れた [?, ?]。 $A(t)$ を時刻 t におけるメディアからの情報とし、これが人ごとに受け取り方が異なる可能性を考慮して係数 c_i を付加する。つまり 係数 c_i は各エージェントごとにメディアの受け取り方が異なる効果を含めている。もし 係数 c_i が正ならばエージェント i はメディア情報の影響を受けて意見を変更し、逆に係数 c_i が負ならば、エージェント i はメディアからの情報に反発して意見を変更することを意味する。

最終的に、本論文で計算に用いるのは以下の方程式である。

$$\frac{dI_i(t)}{dt} = c_i A(t) + \sum_{j=1}^{N} D_{ij}\Phi(I_i(t), I_j(t))(I_j(t) - I_i(t)) \tag{5}$$

実際のソーシャルメディア上のデータが１日ごと、あるいは１時間ごとであれば、差分形式では以下のようになる。

$$\Delta I_i(t) = c_i A(t)\Delta t + \sum_{j=1}^{N} D_{ij}\Phi(I_i(t), I_j(t))(I_j(t) - I_i(t))\Delta t \tag{6}$$

以下ではこの式を用いて数値計算を行う。

3. 計算

この新しいオピニオンダイナミクス理論による計算例を示す。計算例は２人の場合、３人の場合、そして多人数の計算例として３００人の場合の計算例を示す

3.1 ２人の場合

まず、二人の場合で計算をしてみよう。二人をA，Bとしよう。二人の場合、式（**??**）は以下の連立方程式となる。

$$\begin{aligned}\Delta I_A(t) &= -\alpha I_A(t)\Delta t + c_A Adv(t)\Delta t \\ &+ D_{AB}\Phi(I_A(t), I_B(t))(I_B(t) - I_A(t))\Delta t\end{aligned} \tag{7}$$

$$\begin{aligned}\Delta I_B(t) &= -\alpha I_B(t)\Delta t + c_B Adv(t)\Delta t \\ &+ D_{BA}\Phi(I_B(t), I_A(t))(I_A(t) - I_B(t))\Delta t\end{aligned} \tag{8}$$

係数 $D_{AB} > 0$、$D_{BA} > 0$ と設定すると、両者の意見の中間に収束するように二人の意見が推移していくことが計算結果（Fig.**??**) からわかる。これはお互いに信頼している友人同士の意見交換と考えると自然に理解されたと推察される結果となった。合意に達したあとにわずかに正の方向に意見が動いているのは、マスメディアの影響 $Adv = 0.5$ による。

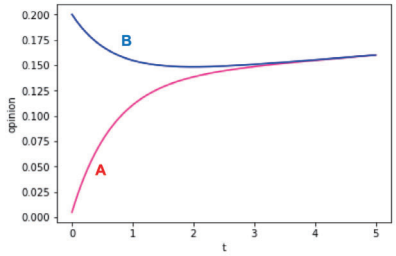


図 1: N=2 の場合の計算。Adv= 0.5, $D_{AB} = 1.0$, $D_{BA} = 0.5$. 初期値は $I_A(0) = 0.005$, $I_B(0) = 0.2$.

次に二人で、お互いに不信感を抱いて反発している場合、係数 $D_{AB} < 0$、$D_{BA} < 0$ と設定する。この場合は Fig.**??** に示すように、お互いに反発するけれど、ある程度意見が離れてしまったら、お互いを無視する形で影響を受けなくなっていることが再現された。

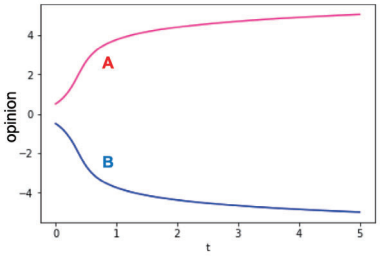


図 2: N=2 の場合の計算。 Adv= 0.5, $D_{AB} = -2.0$, $D_{BA} = -2.0$. 初期値は $I_A(0) = 0.005$, $I_B(0) = 0.2$

3.2　３人の場合

次に３人の場合を考えよう。この場合も式（**??**）で、これを３人で連立させて考える。ここでは３人は A,B,C の３エージェントで、A と B はお互いに不信感を抱いていて反発しているが、A も B もエージェント C には大きな信頼を持っているとする。つまり、C は反発対立している A と B の仲介者の役割になる。計算した結果が Fig.**??**である。A と B の C に対する信頼（係数 D_{AC} 、D_{BC} の値）が大きく正となると、対立していた A と B が C の意見に収束していく様子がわかる。これは仲介者が対立している両者から絶大な信頼を寄せられていると仲介が上手く行くという、日常生活でも経験するような事に対応していることが推察される。

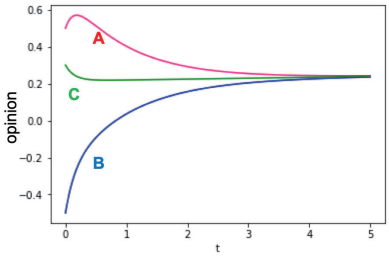


図 3:　３人の場合の計算。 Adv $= 0.005$, $b = 5.$, $\beta = 1.0$, $D_{AB} = -2.0$, $D_{AC} = 5.0$, $D_{BA} = -2.0$, $D_{BC} = 5.0$, $D_{CA} = 1.0$, $D_{CB} = 1.0$. 初期値は $I_A(0) = 0.5$, $I_B(0) = -0.5$, $I_C(0) = 0.2$.

前述のような事由を、実際に起きた事例に置き換えて解釈すると、例えば幕末の日本史で、A を薩摩藩、B を長州藩、C を坂本龍馬と考えると、薩長同盟を仲介した坂本龍馬に両藩から寄せられた信頼の大きさが理解できるケースである（実際の日本史研究では坂本龍馬の果たした役割の大きさには諸説あるが、ここでは単純な通説に従った）。また、現代史で言えば、人種差別政策撤廃後の南アフリカ共和国大統領のネルソン・マンデラ氏の役割が C に似ていると推察する。ネルソン・マンデラは人種差別政策撤廃後に黒人優位の政策を採らず、黒人も白人も平等にという宥和政策に徹して、南アフリカ共和国をうまくひとつにまとめた実績がある。

3.3　300 人の場合

次に、人数を増やして３００人とした場合の計算を示そう。300 人で計算した理由は、意見の分布が上手く描ける程度に多数の人間という程度の意味合いで、300 という数字自体に意味はない。計算にあたって、係数 D_{ij} は全て、－１～１の乱数で決めている。結果は Fig.**??**に一例を示す。左図が時間経過による 300 人の意見の軌跡で、青線は任意抽出の一人の軌跡をわかりやすくするために色を変えてある。右図は計算の最終の時点でこのシミュレーション計算から算出した意見の分布である。

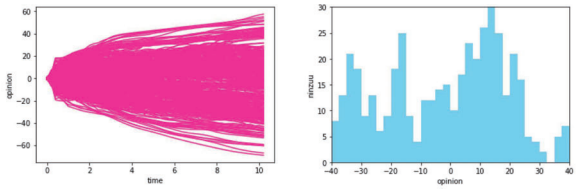


図 4: 300 人の場合の計算。係数 D_{ij} は全て－１～１の乱数で決めている。左図が時間経過による 300 人の意見の軌跡である。右図は計算の最終の時点での意見の分布。

この計算からわかるのは、式（**??**）で求められた分布はある時間から先は有限の範囲内に留まるということである。

3.4　カリスマ的存在がいた場合

次に、多くの人から強い信頼感を持たれているカリスマ的な人物が１人いる場合を考えよう。計算にあたって、係数 D_{ij} でカリスマ的な人物に対して残り 299 人からは高い信頼を寄せられているとする。Fig.**??**の例なら 5.0 の値である。他の全ての係数 D_{ij} は、－１～１の範囲の乱数で決めている。計算結果の一例を Fig.**??**に示す。左図が時間経過による 300 人の意見の軌跡で、青線はカリスマの人の軌跡である。カリスマの人も他の人の意見への対応は、同じく－１～１の範囲の乱数で決めている。右図は計算の最終の時点でこのシミュレーション計算から算出した意見の分布である。計算結果を見ると、カリスマ的人物の意見に近い意見の人の分布は大きいことがわかる。

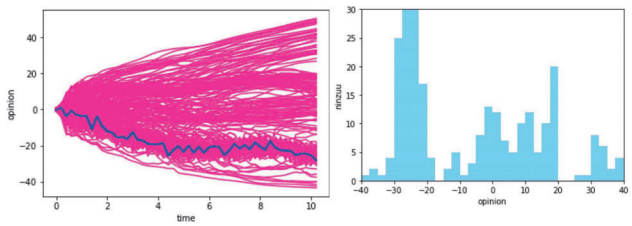


図 5: 300 人の場合の計算。 係数 D_{ij} は 299 人からカリスマの一人に対しての信頼度は 5.0 で、それ以外の係数は全て－１～１の乱数で決めている。左図が時間経過による 300 人の意見の軌跡で、青線はカリスマの人の軌跡。右図は計算の最終の時点での意見の分布。

4.　結論

Bounded Confidence Model を拡張する形で、社会の人々の間の信頼関係と不信関係を同時に考慮し、さらにメディア効果も含めた新しいオピニオンダイナミクス理論を提案した。3 人の計算では仲介者が反発する両者から強い信頼を寄せられていると仲介が成功することが計算から示された。300 人で計算すると、意見の分布はクラスターごとに分かれる傾向が見出された。また、全員から強い信頼を寄せられる人がいると、その人の意見に近い人が多くなる傾向があることが計算で示された。

謝辞

本研究には、パリ大学パリ政治学院の Serge Galam 教授と行われた議論が大きく貢献しているので、Galam 教授に深く

感謝します。また、本研究の計算面で鳥取大の堀尾優輝君に感謝します。

参考文献

[1] French J R P (1956) A formal theory of social power., Psychological Review 63. pp. 181-194.

[2] Harary F (1959) “ A criterion for unanimity in French ’ s theory of social power ”. In Cartwright D (Ed.), Studies in Social Power. Institute for Social Research, Ann Arbor.

[3] Abelson, R P (1964), “ Mathematical models of the dis-tribution of attitudes under controversy ”. In Frederiksen, N and Gulliksen H (Eds.), Contributions to Mathemat-ical Psychology, New York, NY: Holt, Rinehart, and Winston.

[4] De Groot M H (1974) Reaching a consensus. J. Amer.Statist. Assoc. 69. pp. 118 - 121.

[5] Lehrer K (1975) Social consensus and rational agnoiology. Synthese 31. pp. 141 - 160.

[6] Chatterjee S (1975) Reaching a consensus: Some limit theorems. Proc. Int. Statist. Inst. pp.159 -164.

[7] 谷本圭志、喜多秀行、三ッ国 篤志「合意形成の場における雰囲気の形成と その下での住民の発言行動に関するゲーム論的考察」 土木計画学研究 ・論文集 Vol.18no.1 2001 年 9 月, 89 ー 95

[8] Sîrbu A., Loreto V., Servedio V.D.P., Tria F. (2017) Opinion Dynamics: Models, Extensions and External Effects. In: Loreto V. et al. (eds) Participatory Sensing, Opinions and Collective Awareness. Understanding Complex Systems. Springer, Cham

[9] Galam, Physica A 238, 66 (1997).

[10] Sznajd-Weron and J. Sznajd, Int. J. Mod. Phys. C 11,1157 (2000)

[11] Sznajd-Weron, M. Tabiszewski, and A. M. Timpanaro, Europhys. Lett. 96, 48002 (2011).

[12] Galam S, ”Application of statistical physics to politics” Physica A: Statistical Mechanics and its Applications 274, 1999, Pages 132-139

[13] Galam S, ”Real space renormalization group and totalitarian paradox of majority rule voting” Physica A: Statistical Mechanics and its Applications 285, Issues 1-2, 15 September 2000, Pages 66-76

[14] Galam S, “ Are referendums a mechanism to turn our prejudices into rational choices? An unfortunate answer from sociophysics ” Chapter 19 of The Routledge Handbook to Referendums and Direct Democracy edited by Laurence Morel and Matt Qvortrup, (Taylor & Francis, London, 2017)

[15] Galam, S, ”The Trump phenomenon: An explanation from sociophysics”, Int. J.Mod. Phys.B31 (2017) 1742015

[16] Hegselmann R and U Krause, ”Opinion Dynamics and Bounded Confidence Models, Analysis, and Simulation” Journal of Artificial Society and Social Simulation 5 (2002)

[17] Guillaume Deffuant, David Neau, Frédéric. Amblard, and Gérard Weisbuch. Mixing Beliefs among Interacting Agents. Advances in Complex Systems, 3:87-98, 2000.

[18] Gérard Weisbuch, Guillaume Deffuant, Frédéric Amblard, and Jean-Pierre Nadal. Meet, Discuss and Segregate! Complexity, 7(3):55-63, 2002.

[19] Wander Jager and Frédéric Amblard, ”Uniformity, Bipolarization and Pluriformity Captured as Generic Stylized Behavior with an Agent-Based Simulation Model of Attitude Change”, Computational and Mathematical Organization Theory 10 (2004) 295-303

[20] Wander Jager and Frédéric Amblard, ”Multiple attitude dynamics in large populations” presented in the Agent 2005 Conference on Generative Social Processes, Models and Mechanisms, October13-15, 2005 at The University of Chicago.

[21] Evaguenii Kurmyshev, Héctor A. Juárez, Ricardo A. González-Silva, ”Dynamics of bounded confidence opinion in heterogeneous social networks: Concord against partila antagonism” Physica A 390 (2011) 2945-2955

[22] A Ishii and Y Kawahata, “ Opinion Dynamics Theory for Analysis of Consensus Formation and Division of Opinion on the Internet ”, Proceedings of The 22nd Asia Pacific Symposium on Intelligent and Evolutionary Systems (IES2018), 71-76; arXiv:1812.11845 [physics.soc-ph]

[23] A. Ishii, ”Opinion dynamics theory considering trust and suspicion in human relations” Proceeding of 19th International Conference on Group Decision and Negotiation in 2019 a Joint GDN-EWG/BOR meeting, submitted

[24] Ishii A, Arakaki H, Matsuda N, Umemura S, Urushidani T, Yamagata N and Yoshida N, “ The ‘hit’ phenomenon: a mathematical model of human dynamics interactions as a stochastic process ”, *New Journal of Physics* 14 (2012) 063018

[25] A Ishii and Y Kawahata, ”Sociophysics Analysis of the dynamics of peoples’ interests in society” Frontiers in Physics, 08 October 2018 — https://doi.org/10.3389/fphy.2018.00089

Organized Session | Organized Session | [OS] OS-5

[2F4-OS-5b] 複雑化社会における意思決定・合意形成のための AI技術(2)

福島 俊一（科学技術振興機構）、福田 直樹（静岡大学）、伊藤 孝行（名古屋工業大学）

Wed. Jun 5, 2019 3:20 PM - 5:00 PM Room F (302B Medium meeting room)

[2F4-OS-5b-01] An annotation method for a discussion model based on the IBIS structure

〇Naoko Yamaguchi[1], Tomohiro Nishida[1], Daichi Shibata[1], Shota Suzuki[1], Kai Yoshino[1], Kentaro Hiraishi[1], Takayuki Ito[1] （1. Nagoya Institute of Technology）

3:20 PM - 3:40 PM

[2F4-OS-5b-02] Verification of discussions using consensus building support system organized by city development

〇Tomohiro Nishida[1], Takanori Ito[1], Kai Yoshino[1], Daichi Shibata[1], Naoko Yamaguchi[1], Shota Suzuki[1], Kentaro Hiraishi[1], Takayuki Ito[1] （1. Nagoya Institute of Technology）

3:40 PM - 4:00 PM

[2F4-OS-5b-03] Analysis of Discussions with Dealing Rights to Speak Mobile Application

〇Hirofumi Masui[1], Takahiro Oshima[1], Tadahiro Taniguchi[1] （1. Ritsumeikan University）

4:00 PM - 4:20 PM

[2F4-OS-5b-04] Filtering of Impertinent Remarks using distributed expression

〇Kentaro Hiraishi[1], Daichi Shibata[1], Tomohiro Nishida[1], Naoko Yamaguchi[1], Shota Suzuki[1], Kai Yoshino[1], Ahmed Moustafa[1], Takayuki Ito[1] （1. Nagoya Institute of Technology）

4:20 PM - 4:40 PM

[2F4-OS-5b-05] Issues of social acceptability of autonomous system from the view point of robots

〇Shinichi Fukuzumi Fukuzumi[1], Osamu Sakura[2,1], Hiroshi Nakagawa[1] （1. RIKEN, 2. University of Tokyo）

4:40 PM - 5:00 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

IBIS 構造に基づく議論モデルにおけるアノテーション手法の提案

An annotation method for a discussion model based on the IBIS structure

山口　直子* 西田　智裕* 柴田　大地* 鈴木　祥太* 芳野　魁* 平石　健太郎* 伊藤　孝行*
Naoko Yamaguchi Tomohiro Nishida Daichi Shibata Shota Suzuki Kai Yoshino Kentaro Hiraishi Takayuki Ito

*名古屋工業大学
Nagoya Institute of Technology

We propose an annotation method for making training data to develop an automated facilitation agent on the online discussion system. The targets of our developing online discussion system are mainly tackling social wicked problems, therefore in order to progress discussions efficiently, we introduced the Issue-Based Information System's approach into our discussion structuring model. We conducted online discussions based on that model and to collect discussion data. In order to make training data, we annotated those collected data with our defined components. We could achieve to make a prototype automated facilitation agent with using our original annotated data.

1. はじめに

本稿では，オンライン議論システム上の自動ファシリテーションエージェント開発のための，学習用データ作成におけるアノテーションの提案手法について論じる．現在我々は，社会問題等を対象とした大規模議論を，より効率的，かつ個々の意見を尊重して行うオンラインの議論システムを開発中である [Ito 2014,2018]．本システムの重要な機能の一つに「自動ファシリテーション」機能がある．この機能は，基本的には対面式議論におけるファシリテータの役割と同様に，意見の促進，整理，参加者間の相互理解を支援し議論を円滑に進行する役割をコンピュータで実現することを想定している．但し，コンピュータが行うからこそ可能となる議論進行の技法を我々は模索している．

自動ファシリテーションエージェント開発の第一段階として，我々は 2018 年 10 月に，独自に開発したオンライン議論システム「D-Agree」及び「HAMAgree」[注 1] 上で，コンピュータが参加者の意見より特定の議論構造を構成する要素を判別し，その要素から次の構成要素の意見を促進させる発言を出すしくみの実装を果たした．この開発には深層学習を用い [Suzuki 2019]，その学習用データは，D-Agree 上で行った議論データを手動でアノテーションをし，作成した．

2.2 にて詳細を後述するが，現在我々は自動ファシリテーションの手法に関して，Rittle らによる Issue-Based Information System (IBIS) [Kuntz 1970][Rittle 1988]，および IBIS を応用した Dialogue Mapping の概念[Conklin 2005]を主軸に置いている．したがって，アノテーションでは，議論の各意見を IBIS の構成要素をラベルとし，ラベル付けを行った．

2. 提案手法

2.1 オンライン議論システム概要

アノテーションを行ったデータは，すべて D-Agree 上で行った議論内で投稿された意見を用いたものである．ここで D-Agree の特徴的機能を，図 1 を用いていくつか紹介する．図1は実際の議論画面である．まずひとつめの特徴として，スレッド形式議論について説明する．D-Agree では議論参加者が自身で議題に対する意見を自由に入力し投稿していく．投稿には，新規投稿と返信の二つの形式がある。新規投稿は議題に対して新たな小テーマを作成する際に行う．新規投稿を行うと，ひとつのスレッドが形成される．他の参加者（新規投稿者も可）は作成されたスレッドの小テーマを読み，関心を持った場合，その意見に対して自らの意見を，返信機能を用いて投稿をする．このようにしてスレッド内で小テーマに対する議論が展開されていく．スレッド形式を採用したことにより，参加者は明示的構造から解放され，より自由な意見を投稿することが可能である．

図 1：D-Agree 上での議論スレッド画面

二つ目の特徴が，ファシリテータによる議論進行である．ファシリテータも参加者と同様に，新規投稿と返信機能を用いて議論に介入する．議論参加者には予めファシリテータが存在し，そのコメントに注目するように伝達する．ファシリテータ役を人間かコンピュータが担うことは，設定によって変更が可能である．データ作成用議論では，ファシリテーションはすべて人間が行った．

連絡先：山口　直子，名古屋工業大学，〒466-8555 愛知県名古屋市昭和区御器所町，TEL/FAX: 052-735-7407，E-mail: yamaguchi.naoko@nitech.ac.jp

注1「D-Agree」「HAMAgree」は同機能を備えるオンライン議論システムである．D-Agree が原型システムの名称であり，HAMAgree は 2018 年 11 月 1 日～12 月 7 日の間行われた名古屋市次期総合計画中間案市民意見聴取社会実験の際に使用したシステムの名称である．

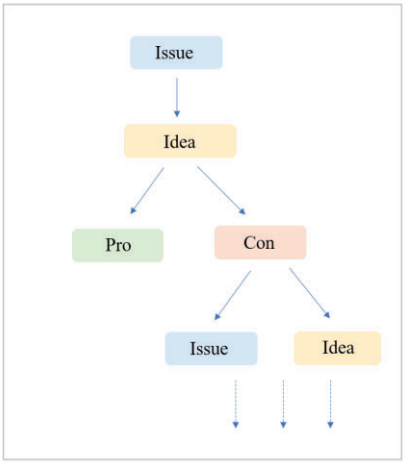


図 2:IBIS 構造に基づく議論モデルフロー

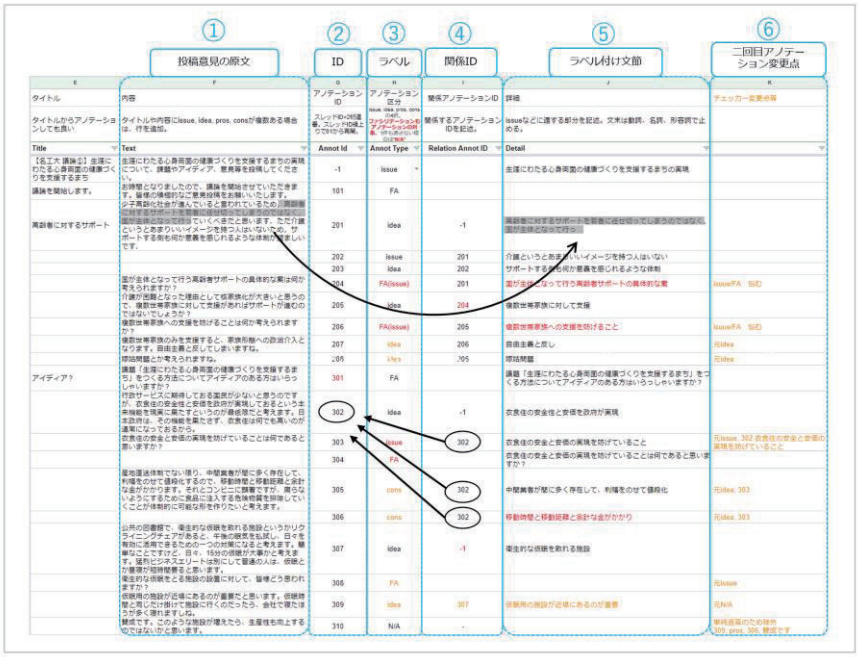


図 3:作成したアノテーションシート

2.2 議論モデル

我々は自動ファシリテーションエージェントのプロトタイプを作成するにあたり，モデルとする議論構造を定義する必要があった．開発初期段階のためシンプルで，しかし議論として意味のある構造を検討した結果，IBIS 手法の応用に至った[Yamaguchi 2018]．IBIS 手法は，その名前の通り，議論において課題に重点を置き，課題を解決するためのアイディアを引き出し，さらに出されたアイディアに対する意見を引き出していく形式による議論設計の概念である．IBIS 手法に基づく議論では，課題を中心に意見を展開していくため，現実的に創造的な議論を展開していくことが可能である．また，IBIS 手法はコンピュータ上での応用を想定して研究された手法であるため，各意見のノードとリンク関係が明確にされており，我々の自動ファシリテーションエージェント開発に適していた．但し Rittle らの IBIS 原案[Kuntz 1970][Rittle 1988]では，構成要素が Issue, Position, Argument の三つで，Position と Argument をシンプルなノードとして取り扱うことが出来なかった為，開発初期段階ではよりシンプルな Conklin による Dialogue Mapping の手法[Conklin 2005]も応用することにした．Dialogue Mapping における議論の構成要素は，Question, Idea, Pro, および Con の四つであり，各要素の定義は文字通りである．

Dialogue Mapping 手法の応用に際し，Question の定義を単純に疑問文と設定してしまうと，原案である IBIS の課題重点式議論から逸脱してしまうと考え，我々の議論モデルでは，構成要素を Issue, Idea, Pro, Con と設定した．ファシリテータは議論において，これらの要素を図 2 のフローで構成させていくことを意図しながら参加者の意見投稿を促進させるよう努めた．

2.3 アノテーション

(1) アノテーション用議論データ

2.2 にて説明した議論モデルをもとに，我々は D-Agree 上で日本語の議論を行い，アノテーション用議論データを収集した．オープンデータではなくオリジナルデータを作成した理由は，より効率的に理想的な議論モデルに則した議論データを収集するため，および，D-Agree 上でのファシリテータと参加者のインタラクションを含めた議論進行の様子を観察するためである．自動ファシリテーションエージェントのプロトタイプ開発に用いた議論データの詳細は次の通りである：合計議論回数 38 回，タグ付けされた文節の合計数 4,972，平均議論システム開放時間 7.81 時間(15 分～144 時間)，平均議論参加者数 13 名(4～115 名)，平均投稿数 101.8／回(ファシリテータ発言含む) 議論参加者はおもに大学生，ファシリテーションは研究員が行った．議題は，教育，災害対策，環境保全、観光推進等多岐にわたる．

(2) アノテーション方法

今回の 38 回の議論データのアノテーションは，2 名の研究員および 3 名の大学生・大学院生によって行われた．2 名の研究員が，ラベルの定義考察等を含むアノテーション作業の全体管理を行った．アノテーション作業行程は次の通りである：(1)議論データを D-Agree データベースより CSV 形式にダウンロードし，各投稿をスレッド毎に時系列に配置し，Google スプレッドシートへアップロードする．(2)一人目のアノテータが各投稿を読み，議論構成要素に該当する文節を抜き出し，ラベルを付与する．(3)二人目のアノテータが，一人目のアノテーションデータを全て確認し，自身の考えと異なる部分に修正を施す．修正を施した箇所はすべてのアノテータが確認出来るよう文字色を変え，修正を行った理由を記述する．アノテーションデータの均質化と全アノテータの要素判別力を確認するため，2 回のアノテーション作業の内，最低 1 回は研究員によって行われた．

アノテーション作業の全体のフローについては以上であるが，次にアノテーションシート上での作業について実際のアノテーションシート(図 3)を用いて説明する．

(1) ①「投稿意見の原文」欄に記されている文章を読み，議論モデル構成要素が含まれている文節の有無を確認する．
(2) ラベル要素を含む文節が存在する場合，該当する文節を抜き出し，⑤「ラベル付け文節」欄に入力する．同時に該当ラベルを③「ラベル」欄に入力する．尚，一投稿に議論構成要素が含まれない場合は，「N/A」ラベルを付与する．
(3) ラベル付けされた文節に対し，ID を付与する． ID は 100 番台から採番し，時系列に数字を増やしていく．また，アノテーションの効率化のため，スレッドが変わるごとに，100 番台の数字も増やしていく．ID は②「ID」欄に入力する．

(4) 2 番目以降のアノテーションより，各ラベル付けされた文節に対し，関係 ID も付与する．これは各ノードの親子関係も学習用データとして利用するためである．関係 ID は④「関係 ID」欄に入力する．

以上の(1)～(4)が基本的なアノテーション作業のフローである．二人目のアノテータはこれらの情報を確認し，修正を施した際は，⑥「2 回目アノテーション変更点」欄に修正理由等を記す．

(3) ラベルの定義

我々は 38 回分の議論データのアノテーション作業を通じ，最終的に以下の通り，IBIS 概念に基づく我々の議論モデルを構成する要素を定義した．

Issue: 課題を解消するための提案を引き出す質問
Idea: 課題を解消するための提案
Pro: 提案の長所や良い側面
Con: 提案の短所や負の側面

尚，ラベルの種類には上記の議論構成要素以外に以下の 2 つが存在する．

FA: 上記ラベルに該当しないファシリテーションの投稿
N/A: 該当せず

ファシリテーションの投稿が議論構成要素に該当する場合は，議論構成要素ラベルを優先的に付与する．

以上のラベル定義に基づき行ったアノテーションにおいて，アノテータ間一致度は，Fleiss’ κ =0.66 を記録した．

3. 学習用データの有効性

提案手法により作成したアノテーションデータは，自動ファシリテーション機能の一部である議論内発言の要素判別とファシリテーションコメント生成の開発に利用した．その結果，コンピュータによる各要素の判別正解率は Argumentation Mining 分野において十分な精度を達成し [Suzuki 2019]，名古屋市次期総合計画中間案に係る市民意見聴取において使用した HAMAgree システム上で自動ファシリテーションエージェントの実装が実現した．また，自動ファシリテーション機能の評価実験[注 2] を行った際，「この議論においてファシリテーションは必要であったと思うか」とのアンケート質問に対し，5 段階評価で，人間 4.22 に対し，コンピュータ（自動ファシリテータ） 3.63 の平均スコアを記録した．これらの結果より，今回開発された自動ファシリテーションは，ファシリテータとしての役割を果たしたものと考える．つまり，開発に利用した提案アノテーション手法を用いて作成した学習用データは有効性の高いデータであることが示された．

4. まとめ

本研究では，自動ファシリテーションエージェント開発のための学習用データ作成に係り，シンプルかつ現実的に創造的議論を展開していく IBIS 手法を応用した，オンライン議論システム上の議論モデル，およびそのモデルに基づくアノテーション手法を構築した．この提案手法により作成した学習データは自動ファシリテーションエージェントの開発に有効なものであることが示されたが，今後の機能性の向上，および国際的に利用できるシステム開発の為に我々は次の課題を掲げる．(1) ラベルの種類の拡張：より意義のあるファシリテーション発言の生成の為，発言の分類数を増やす必要がある．エージェントによる要素判定率とのバランスを考慮しながら，最大限の議論構成要素のラベルを拡張する． (2) 英語議論データのアノテーション：システムの国際的社会応用を考慮し，早急に英語議論データのアノテーションを行う必要がある．英語議論のアノテーションにあたり，日本語議論のアノテーションとの相違点を明らかにし，議論モデルおよびラベルの定義を明確に行っていく．

注 2 開発した自動ファシリテーションの有効性を確認するために評価実験を行った．被験者平均 21 名を各回 2 グループに分け，人間とコンピュータのファシリテーションによる議論を行った．被験者はすべての議論において，人間とコンピュータのどちらがファシリテーションを行うかの情報は与えられていない．議論は 1 時間×2 回，45 分間×4 回の全 6 回を行い，各議論の終了後にファシリテーションの定性的評価を測る質問を中心としたアンケート調査を行った．また，議論データより，議論における人間とコンピュータのファシリテーションの効果に関する定量的調査も行った．

謝辞

本研究内容は，JST CREST「エージェント技術に基づく大規模合意形成支援システムの創成：代表伊藤孝行」（グラント番号 JPMJCR15E1）に支援を受けている研究の一部である．

参考文献

[Ito 2014] Takayuki Ito, Yuma Imi, Takanori Ito, and Eizo Hideshima: COLLAGREE: A Facilitator-mediated Large-scale Consensus Support System, Collective Intelligence 2014, June 10-12, 2014. MIT Cambridge, USA.

[Ito 2018] Takayuki Ito: Towards Agent-based Large-scale Decision Support System: The Effect of Facilitator, The 51st Hawaii International Conference on System Sciences, Hilton Waikoloa Village, USA, January 3-6, 2018.

[Kuntz 1970] Kunz, Werner and Rittel, Horst W. J. Issues as elements of information systems (Working paper). Berkeley: Institute of Urban and Regional Development, University of California, Berkeley, July 1970.

[Rittle 1988] Rittel, Horst W. J. and Noble, Douglas E.: Issue-based information systems for design (Working paper), Berkeley: Institute of Urban and Regional Development, University of California, Berkeley. OCLC 20155825. 492. Retrieved 2017-05-26. Originally presented to the ACADIA '88 Conference, Association for Computer Aided Design in Architecture, University of Michigan, October 1988.

[Conklin 2005] Jeff Conklin: Dialogue Mapping Building Shared Understanding of Wicked Problems, John Wiley & Sons, Inc. New York, NY, USA ©2005 ISBN:0470017686

[Yamaguchi 2018] Naoko Yamaguchi, Takayuki Ito, Tomohiro Nishida: A Method for Online Discussion Design and Discussion Data Analysis, The Thirteenth International Conference on Knowledge, Information and Creativity Support Systems (KICSS-2018) November 15-17, 2018 Pattaya, Thailand

[Suzuki 2019] Shota Suzuki, Naoko Yamaguchi, Tomohiro Nishida, Ahmed Moustafa, Daichi Shibata, Kai Yoshino, Kentaro Hiraishi, Takayuki Ito: Extraction of Online Discussion Structures for Automated Facilitation Agent, 2019 年度 人工知能学会全国大会 (第 33 回), 2019 年 6 月 4 日(火)～6 月 7 日(金) 新潟県新潟市（朱鷺メッセ 新潟コンベンションセンター）（投稿中）

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

まちづくり団体の主催による合意形成支援システムを用いた議論の検証

Verification of discussions using consensus building support system organized by city development association

西田　智裕[*1] Tomohiro Nishida　伊藤　孝紀[*1] Takanori Ito　芳野　魁[*1] Kai Yoshino　柴田　大地[*1] Daichi Shibata

山口　直子[*1] Naoko Yamaguchi　鈴木　祥太[*1] Shota Suzuki　平石　健太郎[*1] Kentaro Hiraishi　伊藤　孝行[*1] Takayuki Ito

[*1] 名古屋工業大学
Nagoya Institute of Technology

We have developed the consensus building support system using web technology and have been conducting verification of social experiment by online discussion with the theme of city development. In the past social experiments, we have organized discussions. However, when disseminating the systems to real society in the future, other organizations/person will organize the online discussions. Therefore, we conduct social experiments that a city development association organize online discussions. We verify the effect and the problem about the social experiments. From this verification, we found that incentives have a major influence on participant behavior.

1. はじめに

筆者らは，2012 年から現在にかけて，web 技術を活用した合意形成支援システム COLLAGREE（以下，支援システム）を開発し，まちづくりをテーマとしたオンライン議論による社会実験の検証をおこなってきた[伊美 15, 伊藤 17, 西田 18]. 伊美らの名古屋市と実施したインターネットタウンミーティング[伊美 15]では，オンライン議論においてファシリテータが重要であることを明らかにした. 伊藤らの愛知県と実施した AICHI デザインリーグ[伊藤 17]では，オンライン議論において，参加者である自治体職員や学生から，高い満足を得られることがわかった. 筆者らの名古屋市青年会議所と実施したまちづくりワークショップ[西田 18]では，対面式とオンライン議論の組み合わせにより，参加者による提案とその満足度を向上させることを明らかにした.

これまでの社会実験では，自治体や非営利団体などと協力して実施してきた. そのため，オンライン議論におけるテーマ設定，ファシリテーション，対面式議論やプレゼンテーションとの連携などについて，筆者らが中心となって主催してきた. しかし，今後に支援システムを実社会へ展開する場合，筆者ら以外の団体や個人がオンライン議論を主催することになる.

そこで本研究では，まちづくり団体の 1 つである公益社団法人名古屋青年会議所（以下 JC 名古屋）の主催による支援システムを用いたオンライン議論の社会実験を実施する. この社会実験における支援システムの効果と課題について検証することを，本研究の目的とする.

2. 合意形成支援システムの概要

本研究で扱った支援システムにおける画面の一部を図 1 に示す. 画面上部に「議論フェイズ」「「投稿欄」参加者からの「投稿」がスレッド形式で時系列に表示される.「投稿」へは「返信投稿」することができる. また，画面右側には議論を支援する「議論ポイント機能」と「キーワード表示機能」が配置されている. さらに，投稿について賛同されている参加者を知るための「賛同ボタン機能」がある. 支援システムに搭載した機能の詳細を表 1 に示す.

図 1　支援システムにおける画面の一部

表 1　支援システムに搭載した機能

システム機能	内容
(1)画像の添付機能	投稿するときに，画像を添付することができる機能.
(2)賛同ボタン機能	投稿されたコメントに対して，賛同を示すことができる機能.
(3)議論ポイント機能	議論の貢献度に応じて参加者がポイントを獲得できる機能. 投稿数や他参加者からの返信コメント数，賛同数によってポイントが得られる. 獲得したポイントのランキングを表示することで，議論の活性化を目指す.
(4)キーワード提示機能	議論中に，頻繁に出現したキーワードを表示する機能.
(5)議論ツリー機能	議論の返信関係をもとに，議論の流れを自動で可視化する機能. 参加者は，議論の全体像や論点が把握でき，議論内容の把握を支援することができる. ファシリテータの利点として，議論構造を共有することで，参加者の認識を統一することができる点がある.
(6)リマインドメール機能	自分の投稿に対する返信があった場合にメールにて通知する機能.
(7)コアタイム機能	設定した時間帯において，議論ポイントを 2 倍獲得できる機能. 参加者が集って議論する時間帯を設定することにより，議論の活性化を目指す.
(8)投資機能	ファシリテータが選んだ投稿について，参加者らが 1000 ポイントを 100 ポイント刻みで投資する機能. 投資結果により投稿の重み付けを可視化し，合意案の決定を支援する.

連絡先：西田智裕，名古屋工業大学コレクティブインテリジェンス研究所，名古屋市昭和区御器所町，052-735-7269，nishida.tomohiro@nitech.ac.jp

図２　コアタイムの表示

図３　ナゴヤ未来会議のホームページ

議論を円滑に進行し，合意形成に向けて深い議論がされるよう調整する役割としてファシリテータが参加する．ファシリテータは，議論に対して中立的な立場を保ちながら議論を進行する．議論は「発散フェイズ」「収束フェイズ」「評価フェイズ」の順に，進めることができる．切り替えは，ファシリテータが議論に毎に判断しておこなう．「発散フェイズ」では，参加者が議論テーマ（以降，テーマ）について幅広く投稿し，様々な可能性について検討する．「収束フェイズ」では，「発散フェイズ」の投稿について検討を重ねて絞り込んでいく．「評価フェイズ」では，これまでの

図４　フォーラムの様子

投稿について，参加者らが 1000 ポイントを 100 ポイント刻みで投資することにより評価し，その結果について議論をおこない，合意案を決定する．

支援システムには，議論の活性化を目的とした「議論ポイント機能」がある．議論ポイントは，投稿，返信投稿，賛同などにより得られる．また「コアタイム機能」もある[Nishida 17]．コアタイムの時間帯は，図 2 に示す告知が画面左上に表示される．コアタイムの時間帯では，「議論ポイント機能」における獲得ポイントを倍増させる．コアタイムはファシリテータにより，任意に設定できる．

3. オンライン議論の概要

JC 名古屋の主催により「ナゴヤ未来会議」という名のホームページを設置して，オンライン議論を開催した（図 3）．このホームページでは，支援システムにおけるテーマの表示や，議論へアクセスする機能の他に，「ナゴヤ未来会議」の趣旨，JC 名古屋の説明などもある．議論期間は，2018 年 3 月 7 日から 7 月 29 日

表２　議論テーマ毎の集計結果

No.	議論テーマ	開始	終了	期間	投稿数		閲覧数		賛同数		閲覧人数		最終フェイズ	投資人数
					出現数	構成比(%)	出現数	構成比(%)	出現数	構成比(%)	出現数	構成比(%)		
1	ナゴヤの「好きなところ」について語ろう	3/7	3/20	14	42	1.5	669	5.4	150	0.9	74	5.7	評価	2
2	小学校の部活廃止ってどう思う	3/8	3/21	14	33	1.2	478	3.8	92	0.5	64	4.9	評価	1
3	子育てを楽しむコツについて語ろう！！	3/22	4/16	26	20	0.7	249	2.0	28	0.2	49	3.8	評価	1
4	久屋大通の再生案募集！	3/23	4/22	31	18	0.6	271	2.2	20	0.1	56	4.3	評価	3
5	憲法って何のためになるの？	3/23	4/22	31	12	0.4	139	1.1	21	0.1	44	3.4	発散	-
6	初めて名古屋にくる外国人に名古屋の魅力を説明することになりました。あなたならどう説明しますか	3/23	4/22	31	18	0.6	172	1.4	15	0.1	50	3.8	発散	-
7	障がい者が住みやすい名古屋って？	3/23	4/14	23	19	0.7	168	1.4	6	0.0	40	3.1	評価	0
8	性的マイノリティの方が住みやすい名古屋って？	3/23	4/14	23	9	0.3	173	1.4	3	0.0	43	3.3	評価	0
9	名古屋の良いところ、悪いところを語ろう！	3/25	4/25	32	25	0.9	236	1.9	17	0.1	65	5.0	発散	-
10	親しい友人に同性愛をカミングアウトされたら受け入れられますか？	4/14	4/26	13	11	0.4	137	1.1	18	0.1	48	3.7	発散	-
11	障がい者に対する必要な配慮と特別扱いの違いは何ですか？	4/14	4/26	13	8	0.3	108	0.9	7	0.0	34	2.6	発散	-
12	スマホを育児に使ってもいい？	4/16	4/30	15	12	0.4	190	1.5	29	0.2	54	4.2	発散	-
13	名古屋城にエレベーターは必要か	4/24	5/24	31	15	0.5	211	1.7	23	0.1	56	4.3	発散	-
14	政治に参画していますか？していないなら、その理由は？？	5/29	6/28	31	31	1.1	298	2.4	90	0.5	52	4.0	発散	-
15	日本は民主主義だと思いますか？あなたにとっての民主主義とは？	6/12	7/11	30	14	0.5	299	2.4	42	0.2	35	2.7	評価	2
16	久屋大通の再開発案、大募集！	6/19	7/18	30	22	0.8	303	2.4	77	0.4	57	4.4	発散	-
17	学校教育でどの様にLGBTを教えるべきでしょうか？	6/25	7/17	23	20	0.7	326	2.6	66	0.4	56	4.3	発散	-
18	男性の仕事と育児の両立、どうしてますか？？？	7/6	7/25	20	7	0.2	99	0.8	16	0.1	33	2.5	発散	-
19	日常生活で憲法を意識することはありますか？	7/6	7/13	8	6	0.2	122	1.0	12	0.1	24	1.8	評価	0
20	（メインホール優先シート争奪）あなたならナゴヤのどんな魅力を発信する？	7/12	7/19	8	1788	62.9	4932	39.7	13084	75.3	145	11.1	評価	27
21	（サブディスカッション優先シート争奪）海外に出る、来訪外国人とお話するなどして「名古屋ってこんないいところがあったんだ」とか「名古屋ってここがダメだ」と思ったことを教えて下さい	7/12	7/19	8	274	9.6	893	7.2	1031	5.9	65	5.0	収束	-
22	ナゴヤにカジノは必要か？	7/17	7/25	9	13	0.5	120	1.0	28	0.2	24	1.8	発散	-
23	（メインホール優先シート争奪、第２弾）知事と市長にどんなまちづくりの課題について議論して欲しいですか？	7/23	7/27	5	365	12.8	1423	11.5	2446	14.1	77	5.9	評価	25
24	外国人の参政権！？	7/24	7/27	4	1	0.0	38	0.3	0	0.0	10	0.8	発散	-
25	（メインディスカッション専用）みんなで創ろうナゴヤの未来！！	7/29	7/29	1	60	2.1	370	3.0	44	0.3	46	3.5	発散	-
合計				474	2843	100	12424	100	17365	100	1301	100		61

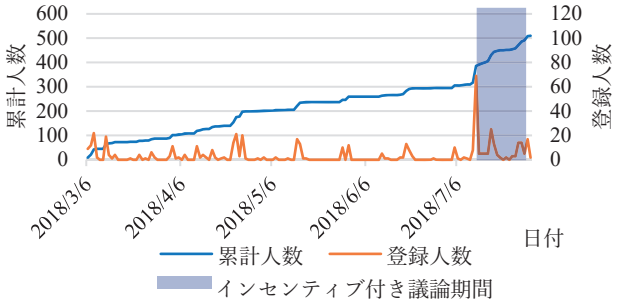


図 5　アカウント登録人数と累計人数

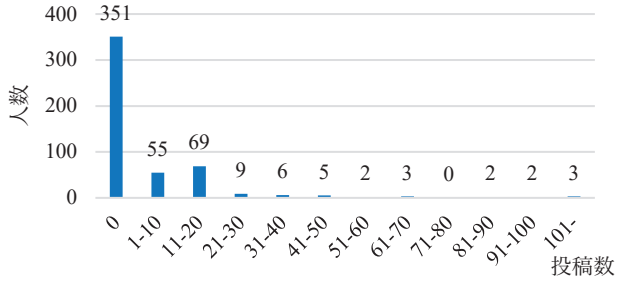


図 6　投稿数毎の登録者分布

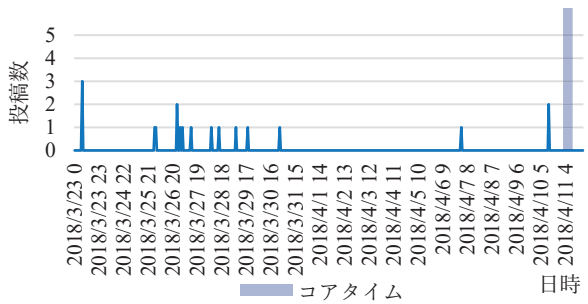


図 7　テーマ 4 における日毎の投稿数分布

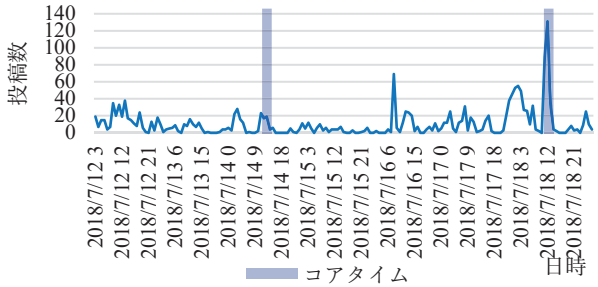


図 8　テーマ 20 における日毎の投稿数分布

と設定した．テーマとして，名古屋市における政治課題など 25 テーマを随時設定した（表 2）．議論へは，アカウント登録による自由参加方式とした．また，議論へは途中からの参加も可能にした．アカウント未登録の人でも議論を閲覧できるように，6 月 12 日よりホームページへ閲覧専用の支援システムと同様の議論画面を設置した．コアタイム機能はテーマ 4，20 において設定した．インセンティブとして，テーマ 20，21，23 では，ポイント上位 2,3 名を，7 月 29 日に JC 名古屋により開催されるフォーラム（図 4）の前 2 列目に招待するインセンティブを設定した．このフォーラムでは，著名な男性・女性のアイドルグループ 6 名と司会者 1 名，名古屋市長，愛知県知事などが出席し，名古屋の政治課題についてパネルディスカッションをおこなった．パネルディスカッションの際，会場からの意見をテーマ 25 において募集し，ディスカッションに用いた．これらテーマ，コアタイム，インセンティブの設定とファシリテーション，フォーラムの実施は，全て JC 名古屋がおこなった．支援システムへのアカウント登録人数は 510 人，議論への投稿数は 2,843 件，閲覧数は 12,424 件であった．

4.　分析

JC 名古屋の主催によるオンライン議論について明らかにするために，支援システムの議論ログを分析する．

4.1　テーマ毎における投稿数と閲覧数

テーマ毎における開始，終了，期間，投稿数，閲覧数，賛同数，閲覧人数，最終フェイズ，投資人数の集計結果を表 2 に示す．

表 2 より，「投稿数」の構成比は，テーマ 20 が 62.9%と最も高く，次いでテーマ 23 が 12.8%，テーマ 21 が 9.6%である．これら以外のテーマは 1.5%以下である．また，「賛同数」構成比は，テーマ 20 が 75.3%と最も高く，次いでテーマ 23 が 14.1%，テーマ 21 が 5.9%である．これら以外のテーマは 0.9%以下である．一方，投稿と賛同からは，議論ポイントが得られる．これらより，今回設定されたインセンティブにより，数多く投稿や賛同したため「投稿数」が多くなったとわかった．

表 2 より，「閲覧数」の構成比は，テーマ 20 が 39.7%と最も高く，次いでテーマ 23 が 11.5%，テーマ 21 が 7.2%である．これら以外のテーマは 5.4%以下である．また，閲覧からは，議論ポイントが得られない．これらより，今回設定されたインセンティブにより，数多く閲覧したため「閲覧数」が多くなるとわかった．しかし，閲覧から議論ポイントが得られないため，最大でも 39.7%の構成比となったといえる．

表 2 より，「閲覧人数」の構成比は，テーマ 20 が 11.1%と最も高く，次いでテーマ 23 が 5.9%，テーマ 1 が 5.7%である．これより，今回設定されたインセンティブは，閲覧者をやや増やすことがわかった．

以上より，今回設定されたインセンティブは，議論への閲覧人数をやや増やすと共に，1 人当たりの「投稿数」と「賛同数」を大きく増やすことが明らかとなった．

4.2　アカウント登録人数

支援システムへアカウント登録人数と，その累計人数を図 5 に示す．

図 5 より，「累計人数」の 3 月 7 日議論開始日から 7 月 11 日インセンティブ付き議論期間までの 128 日間で 310 人，それ以降 7 月 12 日から最終日までの 18 日間で 510 人となった．すなわち，登録人数は開始から 128 日間は 2.4 人/日であり，その後から 18 日間は 11.1 人/日となる．これより，インセンティブ付き議論の開催が，1 日当たりの登録人数を 4.6 倍にしたことが明らかとなった．

図 5 より，「登録人数」におけるインセンティブ付き議論期間の開始日が 69 人と最も多く，次いで議論開始日が 22 人である．これは，インセンティブについて，事前や当日の告知などが要因となったと推察される．

以上より，インセンティブ付き議論期間により，アカウント登録人数を増やす効果があると明らかになった．また，インセンティブ付き議論期間の開始日において，その効果が最も高いことも明らかとなった．

4.3　投稿数毎の参加者分布

支援システムにおける投稿数毎の登録者分布を図 6 に示す．

図 6 より，投稿数が 0 回の登録者が 351 人，構成比 69.2%と最も多い．また，投稿数が 1-10 回の登録者が 55 人，構成比

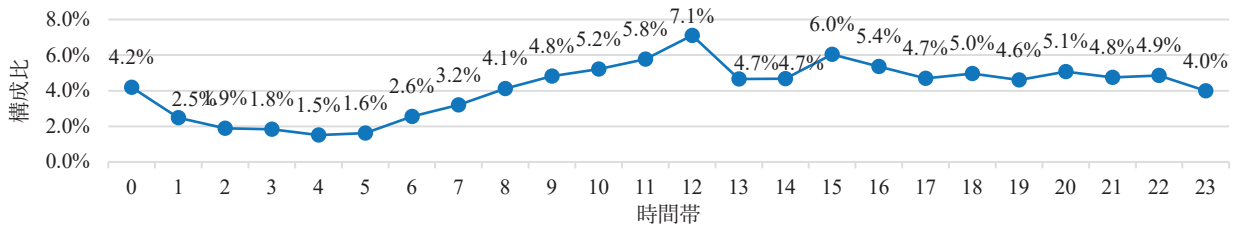


図 9　時間帯毎のホームページアクセス数　（N=9, 583　30 分以内の複数アクセスは 1 回とカウント）

10.8%，投稿数が 11-20 回の登録者が 69 人，構成比 13.5%，投稿数が 21 回以上の登録者が 32 人，構成比 6.3%である．

投稿数が 51 回以上の登録者が 12 人，構成比 2.4%であった．インセンティブの対象者は合計 7 名であったことから，51 回を超える登録者らにより，競い合っていたと推察される．

以上より，登録者の 30.8%が議論へ投稿し，19.9%程度が 11 回以上の投稿していることが明らかとなった．また，インセンティブにより，投稿数が 51 回以上と非常に多くなる競い合いが起きるとわかった．

4.4 コアタイムと投稿数

コアタイムを実施したテーマ 4 とテーマ 20 の日毎の投稿数分布を図 7 と 8 に示す．

図 7 より，テーマ 4 のコアタイムは 4 月 11 日 8-15 時に設定され，その日時の投稿数は 0 件である．これより，テーマ 4 においてコアタイムは効果がないとわかった．

図 8 より，テーマ 20 の 1 回目のコアタイムは 7 月 14 日 11-14 時に設定され，その日時の投稿数は 59 件である．2 回目のコアタイムは 7 月 18 日 11-14 時に設定され，その投稿数は 254 件である．また，2 回目のコアタイムは，前日から日時が支援システム上に実施予告が表示されていた．これより，コアタイムは前日からの実施予告がある場合と，ない場合より投稿数を 4.3 倍に増えることがわかった．

以上より，コアタイムはインセンティブ付きのテーマであり，かつ事前に告知することにより，投稿数を 4.3 倍に増やす効果があった．

4.5 時間帯毎のホームページアクセス数

ホームページの 5 月 15 日から 8 月 31 日までの期間における時間帯毎のアクセス数を図 9 に示す．

図 9 より，12 時台が 7.1%と最も高い．23 時台，0-8 時台は，全ての時間帯が均等のアクセス数だった場合の期待値 4.2%以下の 1.5%-4.2%である．9-11 時台，13-22 時代は，4.2%以上の 4.6%-6.0%である．

以上より，ホームページへのアクセスは，正午を最大として 9-22 時に多くあることが明らかとなった．

5. まとめ

本研究で得られた知見を，以下にまとめる．

(1) アイドルなどが参加するフォーラムの前 2 列目に招待するインセンティブは，議論への参加者をやや増やすと共に，1 人当たりの「投稿数」と「賛同数」を大きく増やすことが明らかとなった．

(2) インセンティブ付き議論期間により，アカウント登録者数を増やす効果があると明らかになった．また，インセンティブ付き議論期間の開始日において，その効果が最も高いことも明らかとなった．

(3) 登録者の 3 割程度が議論へ投稿し，2 割程度が 11 回以上の投稿していることが明らかとなった．また，インセンティブにより，投稿数が 51 回以上と非常に多くなる競い合いが起きるとわかった．

(4) コアタイムはインセンティブ付きのテーマであり，かつ事前に告知することにより，投稿数を 4.3 倍に増やす効果があった．

(5) ホームページへのアクセスは，正午を最大として 9-22 時に多くあることが明らかとなった．

今後の課題として，適切なインセンティブについて明らかにすることがあげられる．

将来の展望として，オンライン議論の主催を筆者ら以外の団体や個人が容易におこなうことができるように，自動ファシリテーションエージェントの開発が求められる．

謝辞　本研究の一部は科学技術振興機構戦略的創造研究推進事業（CREST:JPMJCR15E1）の援助を受けて実施したものである．

参考文献

[伊美 15]　伊美 裕麻, 伊藤 孝行, 伊藤 孝紀, 秀島 栄三:オンラインファシリテーション支援機構に基づく大規模意見集約システム COLLAGREE, 情報処理学会論文誌, Vol. 56, No. 10, 2015

[伊藤 17]　伊藤孝紀, 深町駿平, 杉山弓香, 西田智裕, 秀島栄三, 伊藤孝行: 合意形成支援システムを利用した域学連携手法の有効性, 日本建築学会計画系論文集, Vol. 82, No. 742, pp. 3169-3179, 2017

[西田 18]　西田智裕, 伊藤孝紀, 福島大地, 仙石晃久, 伊藤孝行: ワークショップにおける対面式と非対面式を組合せた手法の検証- まちづくりの社会実験を事例として-, デザイン学研究, Vol. 64, No. 4, 2018

[伊美 15]　伊美 裕麻, 佐藤 元紀, 伊藤 孝行, 伊藤 孝紀, 秀島 栄三: 大規模意見集約システム COLLAGREE における議論インセンティブ機構の試作と愛知県での自治体課題共有実験, 情報処理学会研究報告知能システム（ICS-179）, Vol. 11, pp. 1-8, 2015

[Nishida 17]　Tomohiro Nishdia, Takayuki Ito, Takanori Ito, Eizo Hideshima, Shunpei Fukamachi, Akihisa Sengoku, Yumika Sugiyama: Core Time Mechanism for Managing Large-Scale Internet-based Discussions on COLLAGREE, In the Proceedings of the 2nd IEEE International Conference on Agents, pp.46-49, 2017

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

発話権取引モバイルアプリケーションを用いたディスカッション場の分析

Analysis of Discussions with Dealing Rights to Speak Mobile Application

益井　博史 *1　　大島　崇弘 *1　　谷口　忠大 *1
Masui Hirofumi　　Oshima Takahiro　　Taniguchi Tadahiro

*1立命館大学
Ritsumeikan University

This paper introduces Dealing Rights to Speak to solve the problem of speech volume imbalance arising in discussions by using a mobile application and analyzing its influence on participants. We propose to use the Dealing Rights to Speak mobile application to expand the applicable discussion scale. During experiments, we changed the number of participants and conducted a questionnaire survey on them concerning the content and progress of the discussions, bias in speech volume, respectively. We found that the participants think that the progression of the discussion remains unchanged using Dealing Rights to Speak even when the number of speakers ranges from 4 to 16.

1.　はじめに

日常生活において，会議のようなフォーマルなものから相談や意見交換といったものまで，多様なディスカッションが行われる．それらのディスカッション場では，一部の参加者に発言量が偏ってしまったり，司会進行役のスキル次第でディスカッションが有意義なものにならなかったりするなどの問題が生じる場合がある．本稿では，ディスカッション場における問題を改善するために提案された発話権取引 [古賀 14] を，モバイルアプリケーションを用いて導入し，場にどのような影響を与えるかを検証することで，発話権取引を大規模なディスカッションに適用する可能性を検討する．モバイルアプリケーションを用いることで，これまで物理的制限により数人のディスカッションにしか使用できなかった発話権取引というメカニズムが，大規模の人数による話し合いの場でも活用可能になる可能性がある．発話権取引が適用可能な範囲を把握することで，より広範囲のディスカッション場の問題を適切に解決することが可能になると考えられる．

ディスカッション場において，発言量の偏りや発言力の大きさを左右するのは，そのコミュニティ内の相対的な地位や年齢，声の大きさ，ディスカッションの習熟度など，ディスカッションの内容や発言内容に直接関係のないパラメータによるものが大きいことが指摘されている [亀田 97]．意思決定や共通見解の作成などを行うディスカッションの場では，他の参加者を論破して議論に勝つことを重視すべきではなく，コミュニティ内の立場やパーソナリティに寄らず参加者全員がディスカッションに積極的に関わり，様々な意見や知識を表出させることが重要とされている [茂木 05, 永井 14]．さらに，司会進行役や中心となる人物には熟練した進行スキルが必要とされ，そのような人物を育成し維持するためには多くのコストを払う必要がある [高城 05]．そのため，参加者の発言量の偏りを緩和し，司会進行役のスキルに依存しない仕組み作りが必要である．

前述の問題を克服するための手法として，会議状況の可視化というアプローチが数多く試みられている [Bergstorm 07, Bergstorm 07, Bachour 08, 原田 13, 松村 03]．しかし，発言量・内容を可視化する方法の場合，参加者が会議の質を判断するための予備知識を持っていないときには効果が見られないことが指摘されている [Bergstorm 07]．また，フォーマルな会議などではなく，気軽な話し合いの場を改善するためには，大きなコストをかけられないため，その仕組みは容易に導入できるものであることも重要である．

コミュニケーションするための空間作りや，ファシリテータの熟練を目指すのではなく，発言量などの制度設計に着目してコミュニケーションにおける問題を改善しようとするアプローチを，コミュニケーション場のメカニズムデザインと呼ぶ [谷口 11]．本稿で導入する発話権取引は，そのメカニズムの一例である．発話権取引がコミュニケーション場に及ぼす影響を定量的に把握し，活用できる条件を調査することは，その他のコミュニケーション場のメカニズムをデザインしようとする際にも重要な知見になると考えられる．

2.　発話権取引

年長者や立場の強い者が過剰に発話時間をしてしまうというといった，発話の不均衡という課題を解決するために考案されたコミュニケーション場のメカニズムとして，古賀らが提案する発話権取引がある．発話権取引の流れを以下に示す [古賀 14]．

1. 全員に発話権を配る．
2. 一枚ずつ発話権を使用し，発言を行う，もしくは他の人に渡す．
3. 2．を繰り返した後に，全ての発話権が使用されれば終了する．

発話権取引の手順図を図 1 に示す．

発話権取引では，参加者に事前に発話権を分配することで，話し合いの場における発言時間の設定を行う．参加者は，その発話権を使用することで一定時間，発言することができる．また，他の参加者に発話権を譲渡することもできる．発話時間がまだ残っているときに，話が終わればその時点で発話時間を終了させることが可能である．参加者はこれら三つの行動から自由に選択することができる．

連絡先: 益井博史，立命館大学情報理工学部，滋賀県草津市野路東 1-1-1，masui.hirofumi@em.ci.ritsumei.ac.jp

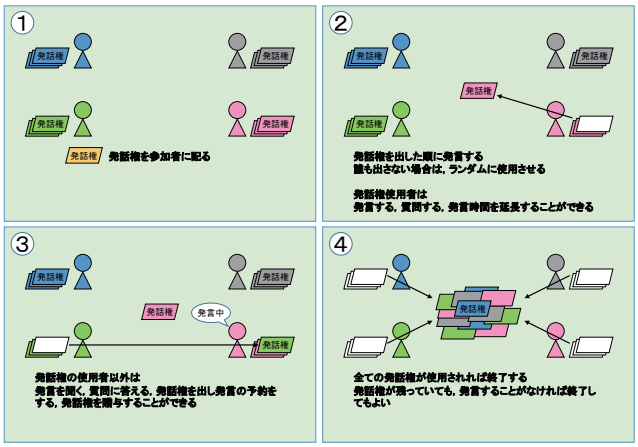


図 1: 発話権取引の流れ [古賀 14]

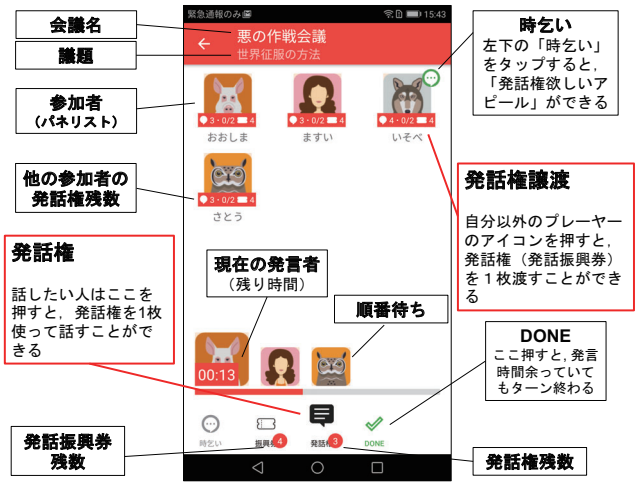


図 2: モバイルアプリケーションの操作画面

発話権取引を導入することで，発言時間が均等化されるように補正されるだけでなく，話し合いの中から司会者の役割をする時間がなくなり，自分の意見や理由を表明する内容の発言の割合が増加することが示されている [古賀 14].

3. モバイルアプリケーション

発話権取引を多人数のディスカッション場に導入しようとすると，カードを準備したり，発話権の譲渡が物理的に難しくなりカードの管理係やタイマー管理係が必要になったりするといったように，実施のためのコストが大きくなる．そこで，モバイルアプリケーション *1 を支援ツールとして用いた．モバイルアプリケーションの操作画面を図 2 に示す．このモバイルアプリケーションを参加者全員が使うことにより，発話権の使用や譲渡，時間管理などを手元の端末の操作のみで行うことが可能になる．

*1 『知』の循環と拡張を加速する対話空間のメカニズムデザイン（JST 未来社会創造事業）ホームページからダウンロードすることができる（アプリケーション名：Trading Hatsuwa）．http://taiwa-kukan.tanichu.com

表 1: 実験条件

実施項目	条件
4 人発話権取引	発話権：4 枚（30 秒/枚）
16 人発話権取引	発話権：2 枚（30 秒/枚）
4 人フリーディスカッション	8 分間 制約なし
16 人フリーディスカッション	16 分間 制約なし

4. 実験

4.1 実験目的

本実験では，発話権取引とフリーディスカッションを参加者のアンケート調査から比較することによって，発話権取引がディスカッションの場においてどのような影響を及ぼすのかを検証する．さらに，発話権取引の影響の程度は人数に応じてどのように変化するのかを検証する．

4.2 実験条件

16 人の被験者を 4 人 4 組に分けた場合と 16 人 1 組にした場合とで，発話権取引とフリーディスカッションをそれぞれ実施した．また，順序効果の影響を排除するため，16 人の被験者グループを 2 組準備し，一方のグループは，(1) 4 人発話権取引，(2) 16 人発話権取引，(3) 16 人フリーディスカッション，(4) 4 人フリーディスカッションの順で実施し，もう一方のグループは逆順で実施した．また，モバイルアプリケーションの習熟度による差を排除するため，実験の前に被験者が練習する時間を設けた．ディスカッションのテーマは被験者から事前に募り，当日に挙手投票によって人気順を決め，最多票のものから実施順に割り当てた．各ディスカッションの条件を表 1 にまとめる．

ディスカッション後，被験者は各ディスカッションに関するアンケートに無記名で回答した．質問項目は次の通りである．

- 議論の内容に関する質問

(1) 議論が論理的だったと思いますか？
(2) 議論が建設的だったと思いますか？
(8) 多様な意見が出たと思いますか？

- 議論の進行に関する質問

(3) 議論の時間を有効活用できたと思いますか？
(4) テンポの良い議論ができたと思いますか？
(7) 議論の進行がスムーズだったと思いますか？

- 発言量の偏りに関する質問

(5) 自分の意見を十分に話すことができたと思いますか？
(6) みんなの意見を十分に聞くことができたと思いますか？

アンケートの回答方法はリッカートの 5 段階尺度を用いて，以下の選択肢から 1 つを選ぶ形式を採った．

1. 非常にそう思わない
2. どちらかというとそう思わない

表 2: アンケート結果計 32 件の平均値.

質問	4 人 発話権	16 人 発話権	4 人 フリー	16 人 フリー
(1)	3.91	3.16	3.68	3.25
(2)	3.91	3.06	3.71	3.56
(3)	3.74	3.28	3.53	2.22
(4)	3.97	3.19	3.50	2.25
(5)	4.22	3.34	3.94	3.00
(6)	4.19	3.78	3.97	3.44
(7)	4.19	3.25	3.56	2.34
(8)	3.97	3.88	3.72	3.69

表 3: 分散分析 F 値.

質問	F 値	有意差
(1)	**5.63**	あり
(2)	**4.87**	あり
(3)	**13.57**	あり
(4)	**17.16**	あり
(5)	**8.19**	あり
(6)	**3.42**	あり
(7)	**20.52**	あり
(8)	0.70	なし

3. どちらともいえない
4. どちらかというとそう思う
5. 非常にそう思う

4.3 実験結果

被験者全員からアンケートの回答を得て，計 32 件となった．得られた回答を間隔尺度データと見なした場合の，各質問への回答の平均値を算出した．結果を表 2 に示す．その他回答結果の詳細については口頭発表にて報告する．

4.4 検定

ディスカッションの 4 つの条件が回答結果に影響を及ぼしているかを知るために，各質問に対し分散分析を行った．有意水準 5%での帰無仮説の棄却域は F >2.68 であり，質問 (8) 以外の全ての質問で有意差が見られた．結果を表 3 に示す．

また，各質問において，(1) 4 人同士，(2) 16 人同士，(3) 発話権取引同士，(4) フリーディスカッション同士を比較するため，順序尺度データを検定するノンパラメトリック手法である U 検定を行った．アンケートの回答方法の順序尺度に対して，U 検定を用いた．有意水準 5%としたが，4 群の多重比較のため，ボンフェローニ法により p 値が $0.05/6=8.33\times10^{-3}$ 未満の場合に有意差ありと判断した．結果を表 4 に示す．

さらに，U 検定と併用して、回答結果を間隔尺度と見なした場合に使用するパラメトリック手法の t 検定を行った．有意水準 5%としたが，U 検定と同じく，4 群の多重比較のため，ボンフェローニ法により p 値が $0.05/6=8.33\times10^{-3}$未満の場合に有意差ありと判断した．結果を表 5 に示す．

表 4: 質問ごとに比較した，U 検定の p 値．p $<8.33\times10^{-3}$で，有意差ありと判断した項目を太字で示す．

質問	4 人同士	16 人同士	発話権同士	フリー同士
(1)	0.166	0.727	$\mathbf{4.65\times10^{-4}}$	0.065
(2)	0.242	0.052	$\mathbf{6.41\times10^{-4}}$	0.674
(3)	0.411	$\mathbf{2.75\times10^{-4}}$	0.143	$\mathbf{5.54\times10^{-6}}$
(4)	0.032	$\mathbf{4.35\times10^{-4}}$	$\mathbf{4.87\times10^{-3}}$	$\mathbf{2.30\times10^{-6}}$
(5)	0.121	0.344	$\mathbf{1.20\times10^{-3}}$	$\mathbf{3.84\times10^{-3}}$
(6)	0.130	0.239	0.065	0.078
(7)	$\mathbf{7.05\times10^{-3}}$	$\mathbf{1.16\times10^{-3}}$	$\mathbf{4.12\times10^{-4}}$	$\mathbf{1.50\times10^{-5}}$
(8)	0.318	0.199	0.940	0.698

表 5: 質問ごとに比較した，t 検定の p 値．p $<8.33\times10^{-3}$で，有意差ありと判断した項目を太字で示す．

質問	4 人同士	16 人同士	発話権同士	フリー同士
(1)	0.225	0.690	$\mathbf{6.42\times10^{-4}}$	0.052
(2)	0.345	0.053	$\mathbf{6.51\times10^{-4}}$	0.521
(3)	0.417	$\mathbf{1.52\times10^{-4}}$	0.107	$\mathbf{6.63\times10^{-7}}$
(4)	0.059	$\mathbf{4.65\times10^{-4}}$	$\mathbf{6.37\times10^{-3}}$	$\mathbf{2.49\times10^{-7}}$
(5)	0.230	0.272	$\mathbf{1.30\times10^{-3}}$	$\mathbf{1.90\times10^{-3}}$
(6)	0.302	0.211	0.099	0.034
(7)	$\mathbf{6.49\times10^{-3}}$	$\mathbf{7.84\times10^{-4}}$	$\mathbf{2.40\times10^{-4}}$	$\mathbf{3.71\times10^{-6}}$
(8)	0.258	0.418	0.679	0.889

5. 考察

質問 (1)，質問 (2)，質問 (8) は，議論の内容に関する質問だった．質問 (1)，(2) に関しては，U 検定，t 検定どちらも人数によって比較した部分にのみ有意差が見られ，t 検定によるとどちらの質問でも 16 人の時の方が 4 人の場合より有意に低くなった．このことから，被験者は発話権取引，フリーディスカッションともに 4 人で実施した時の方が 16 人の時よりも議論内容が論理的・建設的であると感じたと言える．これは，ディスカッション場においては一度に一人しか発話できないため，人数が増えると発話の機会が減ってしまうことが原因だと考えられる．質問 (8) に関しては，4 種類のどこを比較しても有意差は見られなかった．これは，多様な意見が出たと感じるかどうかは，主に人数に対してどの程度のディスカッション時間が設けられているのかに依存するためであると考えられる．

質問 (3)，質問 (4)，質問 (7) は，議論の進行に関する質問だった．質問 (3) に関して，表 2，表 4，表 5 から，フリーディスカッションでは人数が増えることで被験者の回答の値が有意に下がる一方，発話権取引では，有意差が見られないことがわかる．このことから，フリーディスカッションでは人数が増加すると議論の時間を有効活用できたと感じるかについて評価が下がる一方，発話権取引ではその程度が減少されていると考えられる．また，質問 (7) の結果から，被験者が 4 人グループ，16 人グループどちらの場合でも，発話権取引の方がフリーディスカッションより，スムーズな議論ができたと感じていることがわかる．

質問 (5)，質問 (6) は発言量の偏りに関する質問だった．人数に関して比較した場合に有意差が生じるケースはあったが，どちらの人数においても発話権取引とフリーディスカッションとの間に有意差は見られなかった．これは，発話権取引はフリーディスカッションと比べ，発言量に関して被験者全体の満足度を上げてはいないことを意味している．ただし，発話権取引を導入することで，フリーディスカッションにおいて多く話す被験者にとっては話し足りなく感じ，あまり話さない被験者

にとっては多く話せたと感じたこと，つまり発言量の均一化が起こったことで，回答結果が打ち消しあっている可能性が考えられる．今回の実験では無記名での回答としたため，被験者ごとの結果を検証することは不可能だが，今後フリーディスカッションでの発言量と発話権取引での発言量を比較して検証したい．

モバイルアプリケーションを用いることで，発話権取引を大規模なディスカッションにも適用できるかを検証するために本実験を実施した．今回の結果からは，議論の時間を有効に使えたか，という問いに対して，発話権取引では参加人数が 4 人から 16 人になっても，フリーディスカッションほど満足度が下がらないことがわかった．ただし，議論が論理的・建設的だと感じたかという問いに関しては，16 人になった時に発話権取引の満足度はフリーディスカッションよりも下がることもわかった．

考察についても，口頭発表にて詳細を報告する．

6. まとめ

本稿では，ディスカッション場において，モバイルアプリケーションを用いた発話権取引が与える影響を，参加者へのアンケート調査から検証した．その結果，発話権取引が議論内容や発言量の偏りに与える影響を見出すことはできなかったが，議論の進行のスムーズさや，議論の時間を有効活用できたかについて，被験者はフリーディスカッションに比べ良い影響を感じていることがわかった．また，発話権取引がディスカッション場に与える影響をさらに知るためには，各被験者の発話量などの変化を客観的に検証する必要があると考えられ，今後の課題としたい．

発話権取引は，ディスカッション場が持つ問題を解決するために提案されたコミュニケーション場のメカニズムである．発話権取引が，どのような条件で適切に機能するのかを把握することは，他のコミュニケーション場のメカニズムを設計する際の重要な示唆になると考えられる．モバイルアプリケーションの活用によって発話権取引の適用範囲が広がれば，それだけ多くのコミュニケーション場の問題を解決できる機会が増える．ファシリテーションの熟練なく誰でも活用できるメカニズムは，社会の隅々にあるコミュニケーション場が持つ問題を改善できる可能性があり，本稿はその一助となるものと考える．

謝辞

本研究は（一部）、JST 未来社会創造事業 JPMJMI17C7 の支援を受けたものである．

参考文献

[古賀 14] 古賀裕之, and 谷口忠大， 発話権取引: 話し合いの場における時間配分のメカニズムデザイン, 日本経営工学会論文誌 65.3, 144-156, 2014.

[亀田 97] 亀田達也, 合議の知を求めて　グループの意思決定, 共立出版, 1997.

[茂木 05] 茂木秀昭, 身に着けるディベートの技術, 中経出版, 2005.

[永井 14] Junnosuke Nagai and Takaaki Murai and Kazushi Nishimoto, A Conference Support System Following the Monetary System, IEICE Technical Report 113(462), pp.23-28, 2014.

[高城 05] 高城幸司, その場で話しをまとめる技術, 港北出版印刷株式会社, 05

[Bergstorm 07] T. Bergstorm and K. Kalahalios, System Science(HICSS), pp.78, 2007.

[Bergstorm 07] T. Bergstorm and K. Karahalios, Visualizing co-located conversation feedback, IEEE TableTop, 2007.

[谷口 11] 谷口忠大, and 須藤秀紹， コミュニケーションのメカニズムデザイン: ビブリオバトルと発話権取引を事例として, システム/制御/情報 55.8 ，339-344，2011.

[Bachour 08] Bachour and Khaled and Kaplan and Frederic and Dillenbourg, An Interactive Table for Regulating Face-to-Face Collaborative Learning, Third European Conference on Technology Enhanced Learning, EC-TEL 2008, Maastricht, The Netherlands, pp.16-19, 2008.

[原田 13] 原田陽雄 and 米山博人 and 下谷啓 and 藤居徹 and 西野洋平 and 飯田靖, コミュニケーション可視化技術「Voistrap」, 富士ゼロックステクニカルレポート，No.22, 2013.

[松村 03] 松村真宏 and 加藤優 and 大澤幸生 and 石塚満, 議論構造の可視化による論点の発見と理解, 知能と情報：日本知能情報ファジィ学会誌, vol.15, No.5, pp.554-564, 2003.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Filtering of Impertinent Remarks using distributed expression

Kentaro Hiraishi Daichi Shibata Tomohiro Nishida Naoko Yamaguchi

Shota Suzuki Kai Yoshino Ahmed Moustafa Takayuki Ito

Nagoya Institute of Technology

This paper proposes a filtering method, using supervised learning with distributed representation of posted documents and vectors as features. In recent years, online discussion platforms have witnessed a great popularity. However, there here are many harmful contents such as unrelated spam in these discussion platforms, and violent remarks that insult and discriminate against opponents. As result, it becomes necessary to build a discussion platform that allows online users to participate safely by removing inappropriate remarks. To remove inappropriate remarks, understanding and classifying the meanings of documents is needed. Toward this end, we adopt doc2vec and ELMo to word embedding documents. In addition, we constructed a vectorized document by using document similarity calculation and deep neural networks (DNN). The experimental results show that the proposed method is able to classify with higher accuracy.

1. Introduction

Recently, a communication tool allowing users to freely read and write such as bulletin board system and social networking services (SNS) has been developed very much on the web, and everyone can transmit every kind of information on the web. It can be obtained. As a result of on these sites, remarks that disapprove users browsing users such as dating system slandering contents, sexual contents, or speech that encourages crime are scattered, and these contents it is that negative effects are given to minors. And, it is a problem that users who deliver adverse affected contents and users of minors contact each other over the Web, and as a result of recent SNS growth, it is steadily growing. Therefore, it is important to discriminate and remove harmful contents on SNS and Web so as not to adversely affect users of minors.

In recent years, researches on learning meaning vectors of words from large-scale corpus have been actively conducted. In this paper as well, a document is represented by a semantic vector of words contained in the document, and a vector is taken as the feature of the document. We construct information filter using supervised learning, perform evaluation experiment and confirm the effectiveness of this method by showing high F-measure.

2. Related Research

Bayesian filtering is a technique used for spam filtering. We learn how a word characteristically appears in spam or non-spam, and perform filtering by calculating the proportion of characteristic words contained in the mail.

In Japanese, Ando et al. [Ando 2010] propose harmful document determination using word collocation information. Prepare a cooccurrence dictionary and calculate harmfulness based on the number of co-occurrence in the positive example and the number of co-occurrence in the negative example.

Otsuka et al. [Otsuka 2014] propose filtering by making the data of the cooccurrence dictionary very large. But as a problem, the execution time takes quite a lot of time.

Doctor of Information Science, Nagoya Institute of Technology
Gokiso, Showa, Nagoya, Aichi 466-8555, Japan
E-mail: hiraishi.kentaro@itolab.nitech.ac.jp

Sato et al. [Sato 2014] propose filtering using Paragraph Vector and multilayer perceptron. Sato et al. extend Paragraph Vector and predict surrounding words from document vectors and words as input.

3. Building Filters

3.1 Existing Research

・Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) network is one kind of Recurrent Neural Network (RNN) that can learn long-term dependency relationship. RNN has a loop in the intermediate layer in order to retain the storage of input data. Therefore, the neural net can use the previous data as a judgment material. Since the middle layer can deal with the dependence of the data one level before, naturally, when dealing with the previous data, it should be able to deal with the dependency with the first data, but actually It is a problem such as learning can't be done well, there have never been used so much.

LSTM specializes in extending RNN and dealing with previous information.

The choice of information at each gate is made by sigmoid function. LSTM usually learns dependence relationship in forward or forward order, but in this research Bi-LSTM which can learn bi-directionally is used.

・Gated Recurrent Unit (GRU)

Gated Recurrent Unit (GRU) is a model that makes LSTM a bit simpler. By introducing forgetting/updating gates like LSTM, it is possible to generate a shortcut path that bypasses the time steps to write, making it easier to maintain the memory of features of events before a long step.

・doc2vec

Paragraph Vector [Le 2014], also known as doc2vec, is a vector expression method of documents proposed by Le et al. There are two models, Distributed Memory Model of Paragraph Vectors (PV-DM) and Distributed Bag of Words version of Paragraph Vector (PV-DBOW). In PV-DM, words are predicted from words and document vectors in documents. Fig1 is a schematic diagram of PV-DM

In PV-DBOW, predicting words from document vectors. Fig 2 is a schematic diagram of PV-DBOW

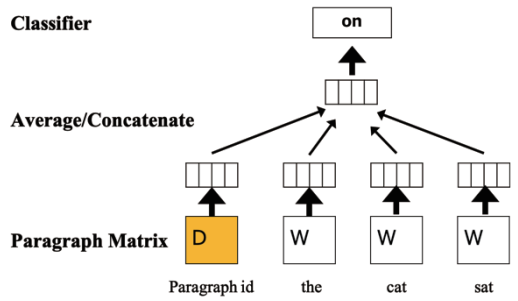


Fig 1: PV-DM

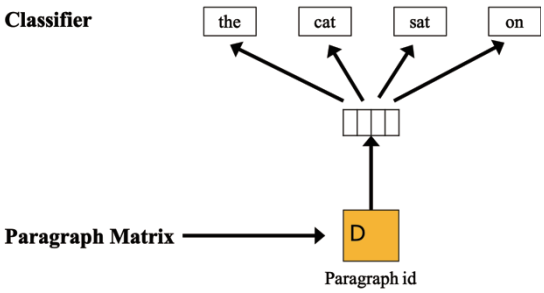


Fig 2:PV-DBOW

3.2 Morphological analysis

In this research, MeCab [Kudo 2004] is used as a morphological analysis engine. In order to vectorize documents in Japanese it is necessary to input to dov2vec etc. for each morpheme.

3.3 Filtering with ELMo and Bi-GRU and Bi-LSTM

In this method, ELMo is used to vectorize each morpheme. ELMo [Matthew 2018] developed by Matthew et al. is a method to acquire a word expression method considering the context by preliminary learning. This method is realized by acquiring the bidirectional language model (biLM) using a large corpus.

Then build and classify a neural network from the vectorized document. We used Bi-LSTM and Bi-GRU

As indicated in Fig3, Morphologically analyze sentences using MeCab and convert each morpheme into distributed representation with ELMo. Pass each output of the output layer as an input to the layer of LSTM or GRU in order after conversion. Softmax is used as the activation function of the output layer, and the probability of each class can be obtained by this.

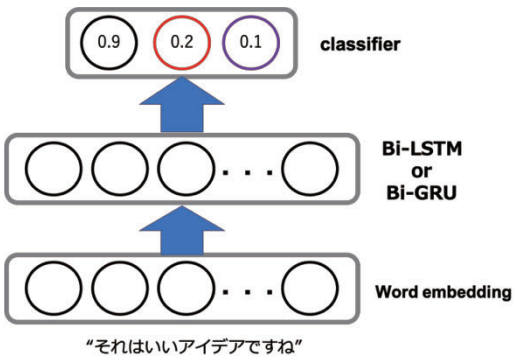


Fig 3: Bi-LSTM and Bi-GRU

3.4 Filtering based on document similarity

We propose classification method using cosine similarity. It is necessary to vectorize documents to compute sentence similarity. We use doc2vec to vectorize documents.

The values of each dimension vectorized by doc2vec have an abstract meaning. Therefore, similar sentences can be thought of as having vectors of close angles. We learned sentences of each class and generated feature vectors.

In this method, filtering is performed using doc2vec and cosine similarity. A labeled and vectorized document is used as input to generate a representative vector. We classify newly posted sentences with feature vectors of each class as follows.

$$\arg\max_{c\in C} cos(\vec{q},\vec{d_c}) = \frac{\sum_{i=1}^{|V|} q_i^2 d_{c_i}^2}{\sqrt{\sum_{i=1}^{|V|} q_i^2}\cdot\sqrt{\sum_{i=1}^{|V|} d_{c_i}^2}}$$

Where, C is the whole class, c is the classification class, d_c is the representative vector of each class, and q is the vector representation of the newly posted document. The similarities calculated for each class are compared and classified into the class with the highest value.

In addition, we propose classification by ensemble in order to increase accuracy. The ensemble method is a method of improving prediction accuracy by fusing classifiers that have separately learned separately.

$$\arg\max_{c\in C} \frac{1}{N}\sum_{k=1}^{N} \cos(\vec{q},\vec{d}_{c_k})$$

As shown in equation, in this method, multiple models of doc2vec are created, In the model of cosine similarity with new posting data was calculated and averaged to classify it to the highest one. As shown in equation, in this method, multiple models of doc2vec are created, In the model of cosine similarity with new posting data was calculated and averaged to classify it to the highest one.

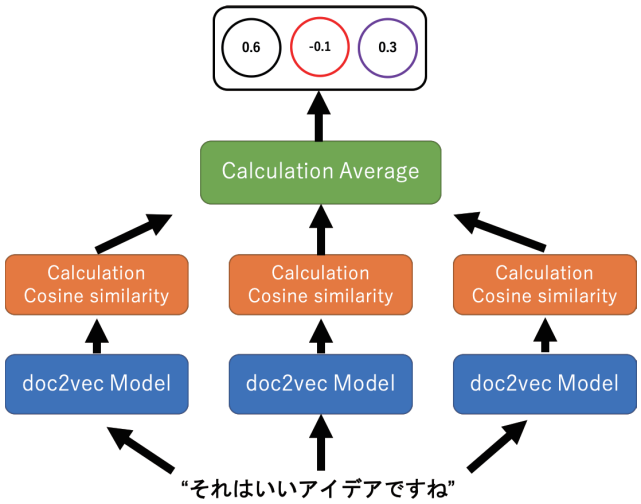


Fig 4: doc2vec Architecture

4. Evaluation Experiment

4.1 Experimental Design

In this experiment, we gathered about 10,000 data from a Japanese textboard, COLLAGREE [Ito 2015] and D-Agree operating in this laboratory. Evaluate with 7,520 training data and 1,680 test data.

As a proposed method, experiments are performed using doc2vec as an ensemble model, Bi-LSTM model, and Bi-GRU model, and the usefulness of the proposed method is confirmed by obtaining high evaluation index. In this research, as the evaluation index of the classification method, the average value of F-measure when classifying from test data was used. F-measure is the harmonic mean of the relevance rate and recall rate.

4.2 Results and Discussion

The average value of F-measure for classification using document similarity and classification using Bi-RNN is shown in the following Table 1.

Tab 1: Experiment Results

Method	F-measure
doc2vec Ensemble Classifier	0.9360
Bi-LSTM	0.9190
Bi-GRU	0.9164

Results on Precision, Recall and F-measure in each method are shown. Table 2 shows results in none-toxic document. Table 3 shows results in obscene document. Table 4 shows results in violent document.

Tab 2: Results in None-toxic Document

Method	Precision	Recall	F-measure
doc2vec Ensemble Classifier	0.9306	0.9882	0.9585
Bi-LSTM	0.9284	0.9450	0.9366
Bi-GRU	0.9022	0.9701	0.9349

Tab 3: Results in Obscene Document

Method	Precision	Recall	F-measure
doc2vec Ensemble Classifier	0.9600	0.9041	0.9312
Bi-LSTM	0.9340	0.9386	0.9363
Bi-GRU	0.9545	0.9144	0.9340

Tab 4: Results in Violent Document

Method	Precision	Recall	F-measure
doc2vec Ensemble Classifier	0.9310	0.9060	0.9183
Bi-LSTM	0.9013	0.8670	0.8839
Bi-GRU	0.9176	0.8460	0.8803

In the approximate class, ensemble classification by doc2vec showed the highest F-measure.

As shown in the Table1, the ensemble model using doc2vec shows the highest F-measure. Because, there was a difference in recall and precision between models used for the ensemble. By ensemble, the balance between precision and recall got better, leading to higher F-measure than other methods.

5. Conclusions and Future work

This paper proposes a method to classify whether posting is inappropriate by using distributed representation. In the proposed method, documents are vectorized using doc2vec, representative vectors are generated for each class, and classified by cosine similarity. Also, we vectorize each morpheme using ELMo and classify it using Bi-RNN. In order to evaluate the proposed method, we conduct experiments using post data on online textboard. The experimental results demonstrate that the proposed method classifies inappropriate posts with high accuracy. Future work is planned to subdivide classes of inappropriate submissions.

Acknowledgement

This work was supported by JST CREST Grant Number JPMJCR15E1, Japan.

References

[Ando 2010] Satoshi Ando, Yutaro Fujii, Takayuki Ito. Filtering Harmful Sentences based on Multiple Word Co-occurrence. Computer and Information Science (ICIS), 2010 IEEE/ACIS 9th International Conference on. IEEE, 2010.

[Otsuka 2014] Takanobu Otsuka, Deyue Deng, Takayuki Ito. Text Filtering for Harmful Document Classification Method Using Three words Co-occurrence and Large-scale Data Processing. IEEJ Transactions on Electronics, Information and Systems 134.1: pp168-175, 2014

[Sato 2014] Genki Sato, Takayuki Ito. A Harmful Document Classification Method Using Paragraph Vector. JAWS2014: pp321-324, 2014

[Le 2014] Quoc V. Le, Tomas Mikolov. Distributed Representations of Sentences and Documents. JMLR W&CP 32 (1): pp.1188-1196, 2014.

[Matthew 2018] Matthew E. Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, Luke Zettlemoyer. Deep contextualized word representations. NAACL 2018, 2018.

[Kudo 2004] Taku Kudo, Kaoru Yamamoto, Yuji Matsumoto. Applying Conditional Random Fields to Japanese Morphological Analysis. Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP- 2004), pp.230-237, 2004.

[Ito 2015] Takayuki Ito, et al. Incentive Mechanism for Managing Large-Scale Internet-Based Discussions on COLLAGREE. Collective Intelligence 2015, 2015.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

ロボットの規格から見た自律システム社会受容性の課題 –インタラクティブシステムにおける意思決定のための要件化-

Issues of social acceptability of autonomous system from the view point of robots –requirements for decision making by interactive system –

福住伸一[*1], 佐倉統[*1,2], 中川裕志[*1]
Shin'ichi Fukuzmi[*1], Osamu Sakura[*1,2], Hiroshi Nakagawa[*1]

[*1] 理化学研究所 AIP センター [*2] 東京大学大学院情報学環
[*1] RIKEN Center for AIP [*2] University of Tokyo

This paper describes issues for acceptability of autonomous system to society from the view point of standardization of industrial robots. When standards items about safety for human (operators, stakeholders) are supposed to apply to self-driving system, three requirements were extracted. They are 1) to be recognized existence of autonomous system to whole social system, 2) to be recognized the condition of activation in the system and 3) to be stopped and informed when users approach to autonomous system. As there are other standards related to interact about autonomous system, these requirements shall be fulfilled in the future.

1. はじめに

近年、AI 技術が急速に発展して社会に埋め込まれるようになってきたため、人間は AI を使わざるを得なくなってきている。従来、製品・システム・サービスの設計においては、機能・性能のみが仕様化され、設計されてきた。最近では人間が使う観点から利用状況を把握し、利用者やステークホルダーの要求を要件化する人間中心設計(Human Centered Design: HCD)が開発プロセスに組み入れられるケースが多くなってきている[HCD 2019]。一方、人間がロボットや自律システムといった AI を使う場合には、AI が人間と直接インタラクションする場合と AI を含めたサービスと人間とがインタラクションする場合がある[福住 2018]。

今後、AI がさらに発展し、AI との直接のインタラクションは AI が行い、人間は製品・システム・サービスの 2 次利用者となる可能性もある。図2にその概念図を示す

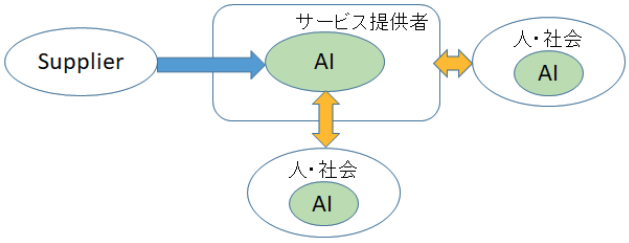


図 1 社会(人、AI)と AI を含んだ社会とのインタラクションの概念図

自律システムと人とのインタラクションについてはまだ不十分なところがあるが、産業用ロボットについては特に安全面に関する規格化もなされている。しかしながら、産業用ロボットは行動や作動範囲が広くないため、自動運転車等これからの AI システムに適用可能かどうかは検討が必要である。本稿では、産業用ロボットの日本工業標準(JIS 規格)を分析し、AI システムへの適用可能性と社会受容性を示すとともに、AI システムを介して人間や社会がインタラクションする際の、意思決定のための必要な要件を示す。

連絡先:福住伸一:理化学研究所 AIP センター
e-mail: shin-ichi.fukuzumi@riken.jp

2. 産業用ロボットの JIS 規格

対象とした JIS 規格は、

「JIS B8433:2015、 ロボット及びロボティックデバイス -産業用ロボットのための安全要求事項-:第一部:ロボット、第二部:ロボットシステムおよびインテグレーション」

である[JIS-B8433-1,-2 2015]。表1に、規格中に示されている自動・自律運転の安全に関する規定の抜粋を、表2に規格の附属書に書かれている危険源リストの中の、インタラクションに関すると考えられる内容(人間工学的危険源)のリストを示す。

表1 産業用ロボットの安全要求事項(抜粋)

危険源同定及びリスクアセスメント(4.)	・附属書に危険源リストを示す。危険源分析では発生する可能性のあるあらゆる危険源の同定を行わなければならない ・リスクアセスメントは、同定したそれらの危険源について行わなければならない
自動運転の始動(5.8.5)	・自動運転モードが有効になる前に、安全防護空間の外側に配置された別の確認手段がなければならない
同時動作制御(5.9)	・同時動作用に設計されたロボットシステムのすべてのロボットは、通常、同じ運転モード(例えば、手動又は自動)又は同じ状態(例えば、動力入り又は切り)でなければならない(5.9.2)
協調運転(5.10)	・協調運転のために設計されたロボットは、協調運転中であることを示す資格表示を備えなければならず… ・人間が協調作業空間内に存在するときは、ロボットは停止しなければならない

表2　使用に関する危険源リスト(抜粋)

種類	危険源の例	
	原因	潜在的結果
人間工学的危険源	-設計が適切でないイネーブル装置 -制御器具（ボタン・スイッチ等）不適切な位置・配置、不注意操作 -予測できる保全活動(トラブルシューティング、修理、調整)をする構成品の不適切な配置による接近困難 -照明が不十分な区域がゆえに危険源又は危険状態の確認がしにくい状態 -照明を妨げるエンクロージャーの内部の構成品 -適切な視野に配置されるべきHMIユニットが高すぎたり低すぎたりする状態	- 健康的でない姿勢又は過度の頑張り(繰返しの緊張) - 疲労

この JIS 規格はあくまで産業用ロボットの安全に関する規格であり、その中で特に周囲の人間（直接利用者だけではない）の安全を確保するために定められた項目である。JIS B0134（ロボット及びロボティックデバイス）[JIS-B0134 2015]によると、

- ロボット：二つ以上の軸を持ち、プログラムによって動作し、ある程度の自律性をもち、環境内で動作して所期の作業を実行する運動機構(ロボットは制御システム及び制御システムとのインタフェースを含む)
- 知能ロボット：環境情報の取得及び/又は外部との相互作用を行いつつ、自分の行動を適応させることによって、作業を実行することができるロボット
- 自律性：人の介在なしに、現在の状態及びセンシングに基づいて所期の作業を実行する能力

である。このことから、上に掲げた規定類について、自動的・自律的に動く対象物と周囲の人間との関係という点では、例えば自動運転車などへの適用の課題は明らかにすべきである。

3. AI システムへの適用可能性と利用者の意思決定のための要件

上記を踏まえ、この規格を自動運転(level3)などにそのまま適用してみる。まず、(4.)については、すべてのリスクを列挙し、アセスメントすることが現実的かどうかは疑問である。産業用ロボットの場合、何等かの形でロボットとインタラクションする人々は相手がロボットであることを認識している（であろう）。また、オペレータもそれが前提で操作し、監視しているので、何かのトラブルの際には関係者は共通した状況認識の上で正しい判断（意思決定）で対処できると考えられる。一方自動運転の場合、周囲の車や歩行者が、どの車が自動運転車かを意識することはないであろう。そのため、状況を共有するのが困難となり、判断を誤る危険がある。このことを少しでも避けるために、(5.8.5)や(5.10)のような「資格表示」を要求するのだが、(5.10)の後者は、自動運転車への提供は不可能である。また、(5.9)は、自動運転車ではこれは前提であるが、ある範囲を定めて外的制御で同じモードにした場合、ドライバーにそのような状況変化（特に自動→手動）を伝えることは困難であり、そのために正しい意思決定もできなくなる。このことから、安全の視点（直接の利用者に限らず）では、

- 機器単体ではなく、それらが存在する社会システム全体に対して、自律システムが存在することを認識させること
- その上で、それらの機器がどのような状態で作動しているのかをシステム内で認識できるようにすること
- たとえ安全対策が施された機器であっても、人間に近寄りすぎた場合などは、停止モードに入るとともにその状態を人間に知らせるようにすること

などが要件として挙げられる

4. Robot Intelligence and Autonomous System

以上のように、従来の安全規格を AI システム、特に自動→手動の切り替えといった人間とのインタラクションが発生するシステムに適用する場合には、考慮しなければならない要件がいくつか存在することがわかった。しかしこれらは利用状況毎に異なるため、一意に決めるのは困難である。この課題を解決する一つの策として、ロボットや自律システムと人間とのインタラクションにおいて、その対象や考慮すべきことをカテゴライズし、ガイドライン的に示す規格文書(TR: technical report)が策定されている。

ISO DTR9241-810 Ergonomics — Ergonomics of human-system interaction — Part 810: Human-system issues of robotic, intelligent and autonomous systems (RIAS) [RIAS 2019]。内容の一例は以下である。

・Human-RIAS interaction issues
　-完全自動でない限り、人と RIAS は共存し、人は RIAS のユーザとなる。
　-RIAS と人間とのコミュニケーションは、人間行動そのものを形成する
　-文化や社会の違いは RIAS が考慮しなければならない
・RIAS-RIAS interaction issues
　-複数の RIAS が同じ環境で動作するとき、異なる供給元の RIAS 間でコミュニケーション問題が生じることがある
・Social/cultural/ethical issues
　-RIAS は社会や文化と切り離して動作することはない。効果的に動作させるためには、チームまたは協力している実態のメンバーとして存在しなければならない

この TR の各カテゴリーに対する規定は順次定められていく予定である。これらの内容も要件に追加していく必要がある。

5. まとめ

今回、産業用ロボットの安全に関する規格を分析し、その規格が自動運転車等 AI システムへ適用可能とするための要件をいくつか提示した。4 章で述べたように、現在、RIAS と人や組織とのインタラクションに関する国際規格もできつつあるので、これらも取り入れ、AI システムが社会で受け入れられるために要件を充実させていきたい

参考文献

[HCD 2019] ISO fDIS9241-210: Ergonomics of human–system interaction —Part 210: Human-centred design for interactive systems, 2019

[福住 2019] 福住他：AI を活用したサービスにおける ELSI 的観点の新たなガイドライン項目の 抽出 -デジタルヘルスを対象とした検討-、AI 学会全国大会、2018

[JIS-B8433-1 2015] JIS B8433-1: ロボット及びロボティックデバイス−産業用ロボットのための安全要求事項−第 1 部：ロボット, 2015

[JIS-B8433-2 2015] JIS B8433-2: ロボット及びロボティックデバイス−産業用ロボットのための安全要求事項−第 2 部：ロボットシステム及びインテグレーション, 2015

[JIS-B0134 2015] JIS-B0134:ロボット及びロボティックデバイス−用語, 2015

[RIAS 2019] ISO DTR9241-810 Ergonomics — Ergonomics of human-system interaction — Part 810: Human-system issues of robotic, intelligent and autonomous systems (RIAS), 2019

Organized Session | Organized Session | [OS] OS-2

[2G3-OS-2a] データ流通社会における技術基盤と異分野連携(1)

早矢仕 晃章（東京大学）、大澤 幸生（東京大学）

Wed. Jun 5, 2019 1:20 PM - 2:40 PM Room G (302A Medium meeting room)

[2G3-OS-2a-01] Data Jackets Evolving and Connecting via History of IMDJ

〇Yukio Ohsawa[1], Teruaki Hayashi[1] （1. The University of Tokyo）

1:20 PM - 1:40 PM

[2G3-OS-2a-02] Data visualization method according to thinking process on the data distribution platform with blockchain

Kosuke Miyahara[1], Yusuke Ejiri[1], Eiji Ikeda[1], 〇Hiromichi Sasaki[1] （1. Fujitsu Limited）

1:40 PM - 2:00 PM

[2G3-OS-2a-03] Efficient Privacy-Preserving Prediction for Three-Layer Feedforward Neural Networks Using Ring-LWE-based Homomorphic Encryption

〇Takehiro Tezuka[1], Lihua Wang[1,2], Takuya Hayashi[2], Sangwook Kim[1], Tomoya Tamei[1], Toshiaki Omori[1], Seiichi Ozawa[1] （1. Kobe University, 2. National Institute of Information and Communications Technology）

2:00 PM - 2:20 PM

[2G3-OS-2a-04] Network Growth Process on Cross-disciplinary Data in the Data Platform

〇Teruaki Hayashi[1], Yukio Ohsawa[1] （1. The University of Tokyo）

2:20 PM - 2:40 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Data Jackets Evolving and Connecting via History of IMDJ

Yukio Ohsawa Teruaki Hayashi

*1 Dept. Systems Innovation, School of Engineering, The University of Tokyo

Here we show the structural changes in the connections among (1) the variables in data jackets (2) words in the expected outcomes of data use, and (3) supposed data use (analysis/simulation) scenarios represented in the 1603 data jackets stored from 2014 through 2018 (business years). These respectively mean (1) the features of data to be provided or created, not necessarily on the allowance to share the data contents, (2) the social expectation about how and for what the data can be used if shared, and (3) the main process of data use. The changes in the structure of the co-occurrence are visualized by use of the traditional sequence of graphs with KeyGraph, where we found that the connections are emerging for the recent two years, that means the ideas of data users are evolving with communication in IMDJ.

1. Introduction

Since we proposed the Innovators Marketplace on Data Jackets (IMDJ) in 2013 [Ohsawa 13], a human-driven platform for creating strategic use scenarios of data and for evaluating the use value of data, participants came to range across various domains of businesses and sciences. The starting workshops in this platform were organized in graduate and undergraduate schools in Japanese universities and exported to other countries Taiwan, India, etc. Then, participants in business sections came to overwhelm academia, with still continuing the activities in scientific projects.

The key technologies in IMDJ here are, simply put, (1) data jackets, (2) human interface for aiding participants in their thoughts and communications with connecting various data without seeing or touching the contents of data [Hayashi 18], and (3) the designed/created data and analysis tools for using data. Among these three, (1) and (2) are the inputs to the IMDJ as a system, whereas (3) is an output. The outcomes of each workshop in IMDJ are mostly the proposed strategic scenarios for data use/reuse (often including data analysis and synthesis) and the results of acting on those scenarios. However, item (3) is also positioned as an outcome or a side product of IMDJ. The relation of these inputs and outputs can be expressed by Eq.(1) through Eq.(3), representing a set of data jackets in (1) by DJs, the use scenario in (2) (called a solution in [Ohsawa 13, Hayashi 18, etc]) by Sol, and tools in (3) by Tool. Also, the requirements spoken in IMDJ from the viewpoints of data users are represented here by Req. And, Act means humans' actions and communications in data processing.

$$\text{DJs, Sol} \rightarrow \text{Req} \quad (1)$$

$$\text{Tool, Act} \rightarrow \text{Sol} \quad (2)$$

Each element above can be expressed by predicates, where DJs takes variables in the data as attributes, Sol takes a part of DJs into its action (Act) and as the inputs to the tools (Tools) it employs in Eq.(2). That is, Eq.(1) and Eq.(2) can be put into

$$\text{provide}(V_1), \text{use}(V_1, \text{req}) \rightarrow \text{realize}(\text{req}). \quad (3)$$

$$\text{process}(V_1, V_2), \text{compute}(\cup_{v \in V2}\, v),$$
$$\text{relate}(\cup_{v \in V3}\, v, \text{req} \mid V_3 \subset V_2) \rightarrow \text{use}(V_1, \text{req}). \quad (4)$$

Here, V_1 in provide(V_1) represents the initially given set of variables in DJs, i.e., the set of data jackets, and the predicate provide means the content of the data supposed to be provided. use(V_1, req) means to realize Sol with the variables in V_1 to satisfy the requirement req in Eq.(3), a situation represented by Req in Eq.(1). The use scenario use(V_1, req) is composed by more granular components, as Sol is in Eq.(2), that is to first preprocess the variables (choose the useful subset of V_1 and further process the chosen set by such procedure as noise filtering). Then, the relations of the variables related relatively directly to the requirements (e.g., the instruction signal) and other variables are computed with automated computation which may be machine learning. Yet the learned relation is not always useful for satisfying the requirement: for example, products sold well in the previous year should be modified to fit the interests of present years' customers. This, as well as tools with AI (Tool in Eq.(2)), actions and communications of humans play the most essential role in the real businesses. Thus, process and relate are represented by Act in Eq.(2). As a result, in IMDJ, all elements composing Eq.(3) and Eq.(4) should be communicated. In summary, the following are communicated in IMDJ.

Provide: to share/get the data to be used

Variables: the variables in V_1 to be given for use/reuse

Process: how the variables and their values are collected or selected, not only for learning efficiently but for fitting req. V_2 may include variables not included in V_1.

Compute: choice of the tools for computation e.q., tools with AI.

Relate: the action to activate the causality between V_3 and req. If V_3 does not fully explain the ways to satisfy req, additional actions should be discussed.

2. The Contents of Data Jackets

Each data jacket (DJ hereafter) is provided not only as the abstract of an existing dataset but also as the preparatory information for the communication exchanging or creating the above information. Therefore, a DJ may include these pieces of information by human-written text. It is certainly allowed that one dataset has multiple DJs by different authors because a DJ is desired to reflect the subjective thoughts or the vision of a participant in IMDJ. Thus, in the entry sheet of each DJ, the following pieces of information and comments are requested.

Contact: Ohsawa Lab, 7-3-1 Hongo, Bukyo-ku 113-8656 Tokyo, info@pands.sys.t.u-tokyo.ac.jp

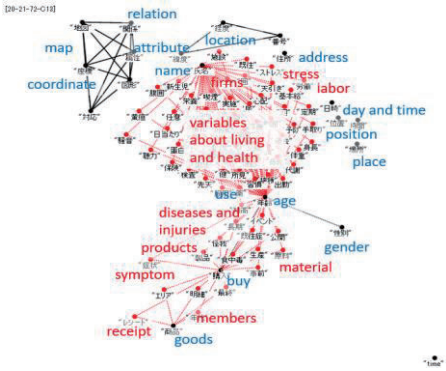
Title of the data (e.g., "POS data in a supermarket")
Outline (e.g., "the sequence of goods bought by customers who come to buy consumption items")
Collecting Cost (e.g., "depends on the system.")
Sharing Policy (e.g., "can be shared after negotiation, if the condition fits the business of the supermarket.")
Types (e.g., "a table of text, numbers, and symbols.")
Formats (e.g., "XML")
Variables/Attributes (e.g., "product class, product category, product name, customer ID, and date and time of purchase.")
Analysis /Simulation (e.g., "apply sequence analyzer such as RNN or visualizer such as Tangled String, for predicting and/or explaining the future strategies in marketing. Predicting may not be trustworthy because the market in the past and the future tend to differ, but the explanation is absolutely required.")
Outcome (understanding the reasons for customers' behaviors)
Anticipation (expected strategies to satisfy existing and forthcoming customers, to be obtained from the data)
What kind of data/tools do you wish to have? (additional data or AI tools that should be combined with the current dataset)

The information corresponding to **Provide** in Section 1 is included in **Collecting Cost** and **Sharing Policy** here. **Variables** in Section 1 are in **Variables/Attributes** here. **Process** and **Compute** in Section 1 are provided in **Analysis /Simulation** here that includes to select variables and use tools with AI, because it is not easy to separate this preprocess and the main step of computation because ordinary users tend to trust experts or automated tools as a package for setting hyperparameters in the preprocess and learning (computation). It is noteworthy here that IMDJ is a platform to connect requirements of sheer users of data and AI tools, and the details of settings in the post-process are dealt with in the later step of action planning. Finally, **Relate** in Section 1 is expected to be filled in **Outcome** and **Anticipation** above, although the entries to the blanks for these two items are optional. In summary, from the viewpoint to detect the social interaction of the data use/reuse process, the three items **Collecting Cost, Sharing Policy, Variables/Attributes** represent the pre-process, whereas **Outcome** and **Anticipation** represent the post-process. **Analysis/simulation** shows a part of the main process.

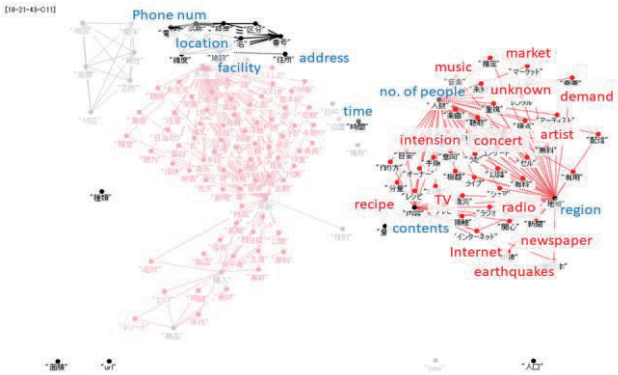
3. The visualized changes in DJ for 5 years

Kamishibai (K)-KeyGraph shows the structural variation of the co-occurrence graphs of items in the partial data for each period, in a given sequence [Ohsawa 10]. The black nodes with blue letters show frequent words, connected by red ones represented by less frequent words. The sequence of visualized graphs is shown for explaining the structural changes, but the distances such as obtained from the coordinates in distribution representation vectors are not reflected here. This visualization rule is common in Figure 1, 2, and 3. The same item appears at the same position in the 2D coordinate, so the changing is intuitively grasped at a glance. In Figure 1, 2, and 3, respectively, the changing in the connections among the variables in **Variables/Attributes (Collecting Cost, Sharing Policy** are cut because of the difference in the granularity of information**)**, among words in **Outcome** plus **Anticipation,** and among words in **Analysis /Simulation process**, are visualized.

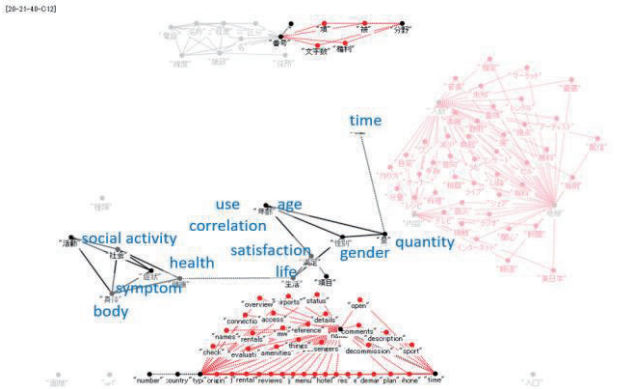
(a) April 2014 – March 2015



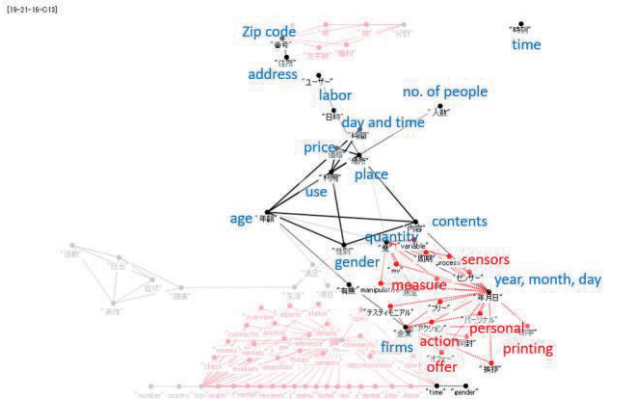
(b) April 2015 – March 2016



(c) April 2016 – March 2017



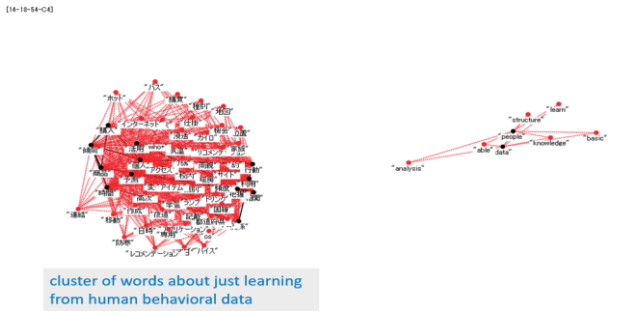
(d) April 2017 – March 2018



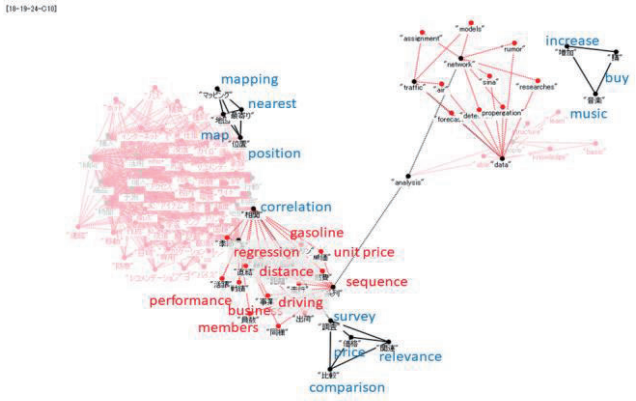
(e) April 2018 – Jan 2019

Figure 1. The changes in the connection of variables in DJs.

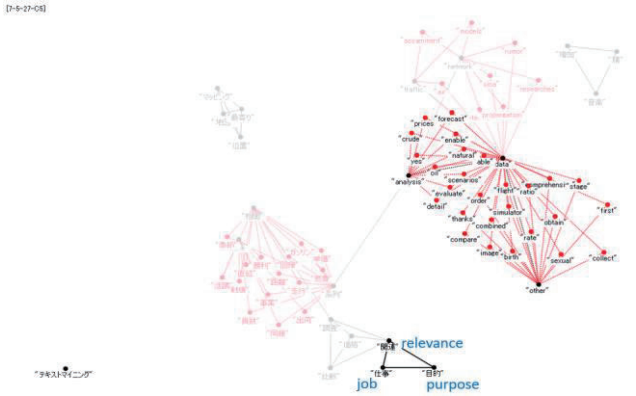
(a) April 2014 – March 2015



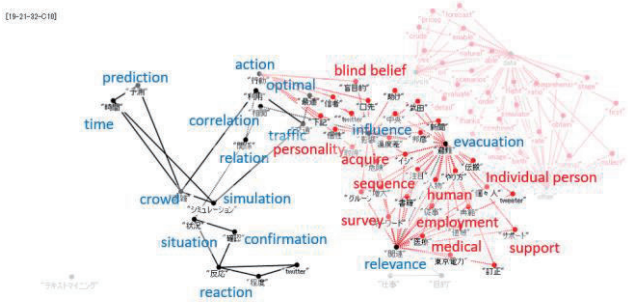
(b) April 2015 – March 2016



(c) April 2016 – March 2017



(d) April 2017 – March 2018



(e) April 2018 – Jan 2019

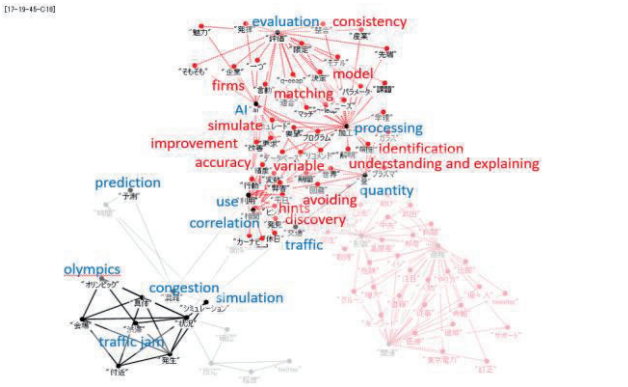


Figure 2. The changes in the connection of words in the expected outcomes of data use, described in DJs.

4. Discussions

The numbers of nodes and edges in graphs are set equally for all periods (years), to make the comparison for different periods. As in Fig. 1 (a), the connections among the variables have been sparse in 2014, in that the human (personal) data and geographic data have been separated into clusters. This tendency lasted till 2015, but the human data and physical data including data from sensors have got coupled into a cluster in 2017. Then, a larger cluster in 2018 appeared that integrates the human data and physical data into AI-driven innovations toward businesses with emotional designs and marketing. The cluster that emerged at the center of Fig.1 (c) for 2016 may be interpretable as a transient path from the localized clusters in (a) to the grown-up one in (e). This integration of clusters is also found in Fig.2. In Fig.2(d) for the year 2017, the simulation and data analysis are coupled into a cluster on services for health care, simulation-based urban traffic management [Yoshimura 17], disaster protection etc., linked to show social demands for AI-driven innovations in Fig.2 (e).

Such a connection of local clusters is not so obvious in Fig.3. However, the last cluster in (e) connect visions toward data-driven innovation and the full sequence is coherently addressed to realize "explanation" as well as "prediction" from sequential data (highlighting spaciotemporal information), putting these keywords at the center of the graphs. The last Fig. 3(e) involves "theory" used for two meanings: (1) mathematical models for investigating latent dynamics [Yamanishi 15], and (2) the respect to domain knowledge highlighted in model-based sciences [Magnani 17], corresponding to the return to "model" in Fig.3 (a)

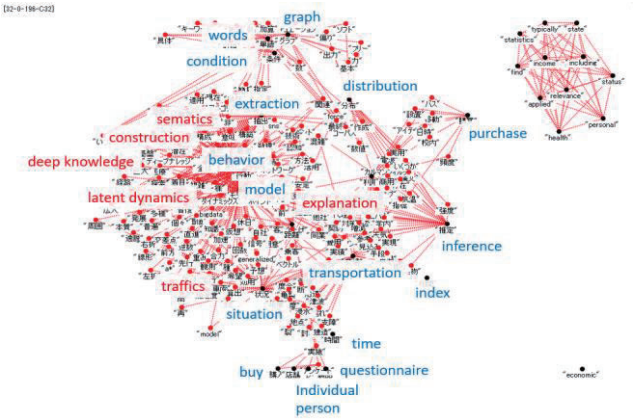
(the assumption of low-dimension causality in using sparse data [Rish 14] is itself a model that may not work in explaining behaviors in the complex interaction with unexpected events).

5. Conclusions

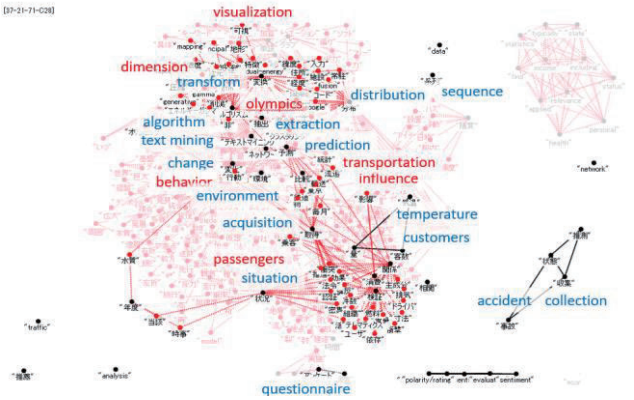
Above we find the human-centric data engineering process has been considered so far in IMDJ, reflecting the communication onto the entries of new DJs. DJs came to be included in the DTA standard of data catalog. We assume this reflects the potential expectation of data users, whose communication in the market of data triggers the evolution of representation of data.

Acknowledgments This work was supported by JST CREST JPMJCR1304, JSPS KAKENHI JP16H01836, JP16K12428, and industrial collaborators in the DJ workgroup.

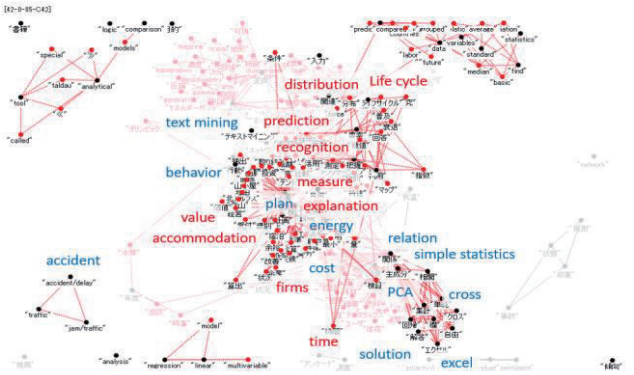
(a) April 2014 – March 2015



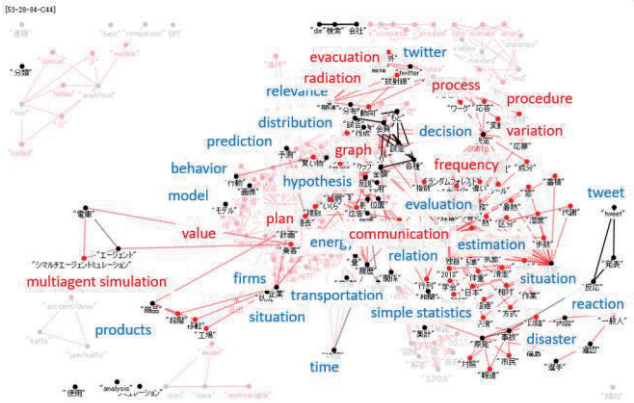
(b) April 2015 – March 2016



(c) April 2016 – March 2017



(d) April 2017 – March 2018



(e) April 2018 – Jan 2019

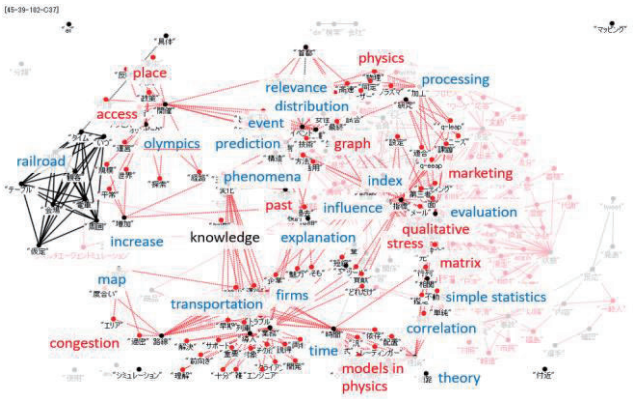


Figure 3. The changes in the connection of words in the analysis/simulations described in DJs.

References

[Hayashi 18] Hayashi, T., Ohsawa, Y., "Inferring Variable Labels Using Outlines of Data in Data Jackets by Considering Similarity and Co-occurrence," *Int'l J. of Data Science and Analytics* 6(4), 351-361 (2018)

[Magnani 17] Magnani, L., Bertolotti, T., Springer Handbook of Model-Based Science (2017)

[Mudinas 18] Mudinas, A, Zhang, D., Levene, M., Market Trend Prediction using Sentiment Analysis: Lessons Learned and Paths Forward, in *Proc. Wisdom* 2018, London.

[Ohsawa 13] Ohsawa, Y., Hayashi, T., Kido, H., Liu, C., et al, Data Jackets for Synthesizing Values in the Market of Data *Procedia Computer Science* 22, 2013, pp. 709-716

[Ohsawa 10] Ohsawa, Y., Ito, T., and Kamata. IM. Kamishibai KeyGraph: Tool for Visualizing Structural Transitions for Detecting Transient Causes, *NMNC* 6 (2) 1-15 (2010)

[Rish 14]Rish, I., Grabarnik, G.A., *Sparse Modeling Theory, Algorithms, and Applications,* CRC, 2014.

[Yoshimura 17] Yoshimura, S., Fujii, H., "MATES: Multi-Agent based Traffic and Environmental Simulator -Core Technologies and Practical Applications-", 8th Int'l Conf. Computational Methods (ICCM 2017) (2017).

[Yamanishi 15] Hayashi, Y., Yamanishi, K., Sequential network change detection with its applications to ad impact relation analysis, *KDD* 29(1) 137−167（2015）

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

ブロックチェーンを用いたデータ流通基盤における思考プロセスに応じたデータ可視化方法

Data visualization method according to thinking process on the data distribution platform with blockchain

宮原 広翼[*1] Kosuke Miyahara　江尻 祐介[*1] Yusuke Ejiri　池田 栄次[*1] Ikeda Eiji　佐々木 泰芳[*1] Sasaki Hiromichi

[*1] 富士通株式会社　ネットワークソリューション事業本部　サービスビジネス事業部

Abstract In recent years, interest of data utilization is emerging and many new businesses using data are considered. We think combining different companies' data is effective for creating innovative businesses, so sharing different companies' data each other would be necessary. Then, in order to accelerate the data utilization society, Fujitsu developed "Virtuora DX Data Distribution and Utilization Service". Our service uses blockchain technology and Data Jacket proposed by Ohsawa Laboratory in the University of Tokyo. In our service, companies can share their data attribute information as Data Jacket and discuss about data utilizing idea using KeyGraph which shows relation among Data Jackets. We think the view of KeyGraph in our service should be flexible according to users' current status on the thinking process to create idea. In this paper we explain some method to change the view of KeyGraph flexibly.

Keywords Data Distribution, Data Jakcket, KeyGraph, Co-Creation, Blockchain

1. はじめに

IoT 技術や各種センサーの進化によって，ヒト・モノに関連する膨大な情報をデータ化，蓄積できるようになってきた．同時にビッグデータ解析や AI（人工知能）の活用によって，新たな価値創出を目指す動きが世界的に加速している．

こうした多様なデータと最新技術の組み合わせにより，これまでにない新しいサービスや製品を生み出していくことは企業の競争力強化につながるだけでなく，より良い豊かな社会の実現にも重要な役割を果たすと考えられる．そのためには，社会が保有するデータを業種・業界を超えて流通・利活用され，持続可能な価値創出が行われる必要がある．

しかしながら，実際は多くのデータが組織やシステムに閉じて蓄積されているため，組織を超えて相互利活用可能な状態になっておらず，データが経営資源として資産化されていない．また，データが資産化されたとしても，膨大なデータから新しいアイデアや新たな価値を生み出すプロセスの確立が必要となる．さらに，資産化されたデータを信頼できる相手と安心・安全に共有・活用する仕組み作りが不可欠な状況となっている．

これらの課題解決のため，富士通はブロックチェーンをネットワーク制御に応用した「富士通 VPX®テクノロジー」をベースに，信頼のおける関係者同士が参加し，データを用いたイノベーション活動を行うサービス「FUJITSU Intelligent Service Virtuora DX データ流通・利活用サービス」（以下，Virtuora DX）を発表した[富士通 18]．本稿では，本サービス上で流通データを拡大していく上で必要となるデータ可視化方法について記載・考察する．

連絡先：富士通株式会社
ネットワークソリューション事業本部サービスビジネス事業部
〒212-8510 神奈川県川崎市幸区新小倉1−1
E-mail: miyahara.kosuke@jp.fujitsu.com, ejiri.yusuke@jp.fujitsu.com, Ikeda.eiji@jp.fujitsu.com, hiromichi@jp.fujitsu.com

2. Virtuora DX データ流通・利活用サービス

Virtuora DX は，データ活用グループ（コンソーシアム）の形成から，データの登録・共有，アイデア創出，データの受け渡しなど，データ利活用プロセスを一貫して効率的に実現可能なサービスである[池田 18]．

各企業が保有するデータの概要情報をポータルサイト上で安心・安全に共有・可視化することで，企業を超えて，さらには業界・業種を超えてデータ流通・利活用とアイデア創出が可能となる．

図 1 Virtuora DX のポータル画面

2.1 データジャケットの適用

一般的に企業の保有するデータそのもの（以下，実データ）は，非公開情報や個人情報等を含む可能性があるため，そのまま企業を越えて共有することが困難である．また実データを共有しても，データサイエンティスト等の専門性が高い人以外は，その内容や価値を十分に理解できず，活用の道を見つけられないといった課題がある．これに対し，実データそのものを共有するのではなく，データの概要情報のみを共有し，利活用アイデアを創出するアプローチがある．東京大学・大澤幸生研究室

では，上記概念を「データジャケット」として提唱している[大澤 17]．データジャケットを書くことでデータが情報化されると，データに関わる人の情報格差を埋めることが可能となる．筆者らも Virtuora DX 上でその概念を取り入れている．

2.2 キーグラフによるアイデア発想支援

データジャケットの組み合わせから，新たなデータ活用アイデアを発想する為には，データジャケット同士の関連性や，共通のキーワード・変数を可視化することが有効である．東京大学・大澤幸生研究室では，このキーワード抽出・可視化手法を「KeyGraph（キーグラフ）」として開発しており，Virtuora DX でも登録されたデータジャケットから関連性をキーグラフとして可視化する機能を提供している．

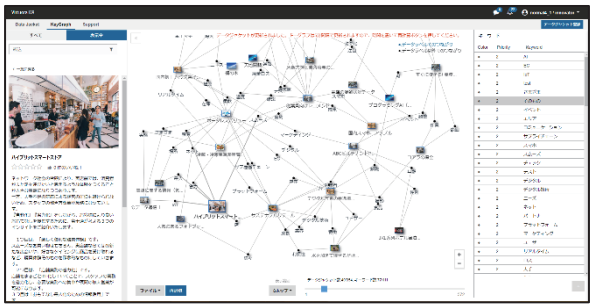


図 2　Virtuora DX で提供するキーグラフ

3. 思考プロセスの進行に応じたキーグラフ表示の最適化

コンソーシアム規模の拡大に伴い，必然的に登録されているデータジャケット数も増加するため，ユーザーにとって，キーグラフ上の多数のデータジャケットを認知し，その組み合わせからアイデア発想を行うことが困難になってくる．

また，コンソーシアム規模の拡大により，当初の参加企業はデータ活用プロセスが進捗し，求めるアイデアの粒度やレベルも変化するし，新たに参加する企業においても，各々様々な検討レベル・フェーズに位置する組織が参加することが想定される．

したがって，キーグラフの視認性を高め，個々の参加者の思考プロセスの進行具合に応じて，柔軟かつ簡易にキーグラフの表示内容を最適化できる仕組みが必要である．

そこで，①データジャケット間の親近性に基づく表示範囲の調整，②データジャケットの評価度合いに基づく表示範囲の調整，③データジャケットの指定によるキーグラフの表示，の 3 種類の仕組みを実装し有効性を検討した．次章で各方法について説明する．

3.1 データジャケット間の親近性に基づく表示範囲の調整

(1) ホップ数の変更

表示範囲を最適化する方法として，「ホップ数」を使用する方法が考えられる．ホップ数とは，キーグラフ上で注目している 1 つのデータジャケット（図 3 において最も濃い色のデータジャケット）を中心として，連鎖的に繋がる先のデータジャケットがいくつのデータジャケットを経由しているかを示す数であり，中心となるデータジャケットと当該データジャケット間の親近性を表している（図 3 参照）．

本方法では，表示する「ホップ数」（注目しているキーグラフから何ホップ目までを表示するか）を変更することで，データジャケットの表示範囲の調整を行う．

ホップ数による調整の効果として，例えば，ホップ数を小さくした場合，親近性の高いデータジャケットが表示されるため，互いに近い市場領域のデータを発見できる可能性があり，ホップ数を大きくした場合は，異なる市場領域のデータを発見し，組み合わせアイデアを発想できる可能性がある．

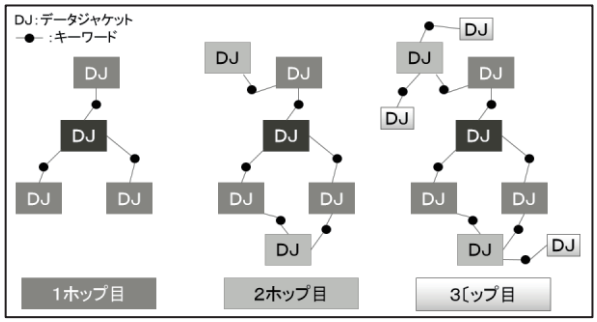


図 3　ホップ数に変更よる表示範囲の違い

(2) ホップ数の上限設定

また，表示するホップ数を変更する場合，上限設定が必要である．上限設定を行う際，ホップ数変更により繋がるデータジャケットの特徴は，下記の様になると考えられる．

- 1ホップ目：共通のキーワードで繋がっているデータジャケットを発見できる．（関係性の近いデータジャケットからアイデア発想ができる．例えば，同じ市場領域のデータジャケット同士）
- 2ホップ目：共通のデータジャケット（1 ホップ目）に繋がっているデータジャケット（2 ホップ目）を発見できる．（例えば，同じ市場領域のデータジャケット（1 ホップ目）と組み合わせ可能性のあるデータジャケット（2 ホップ目））
- 3ホップ目以上：共通のキーワードやデータジャケットをもたないデータジャケットを発見できる．（想定外の新しいデータジャケット）

上記により，ユーザーの思考プロセスの進行に対応させるためには，3 ホップ目まで必要と考えられ，実際にホップ数を変化させてキーグラフを生成すると，4 ホップ以上は視認性が低下することが想定されたため，3ホップを上限と設定した．

3.2 ユーザー評価度合いに基づく表示範囲の調整

また，別の観点として「ユーザーによる評価度合い」を用いて表示範囲を調整する方法が考えられる．ユーザーによる評価が高いデータジャケットは，より多くのユーザーが価値を認めているため，該当データジャケットを使ったアイデア創出を行える可能性が高いと考えられる．そこで，ユーザーからの評価の高さに基づいて表示されるデータジャケットの範囲を変化させることが可能な機能を実装した．（図 4 参照）

ユーザー評価の基準は，データジャケットの閲覧者による評価を用いており，今後，データの新鮮さ（登録日時）も含めて複合的に評価し価値を算出することを検討している．

図 4 のように，スライドバーを左右に移動させることで，表示されるデータジャケット数が変化する．このように，評価値を用いてデータジャケットで変化させることにより，ユーザーの認知でき

る範囲に応じて，活用しやすいデータジャケットが表示され，アイデア発想の効果を高めることが可能になると考える．

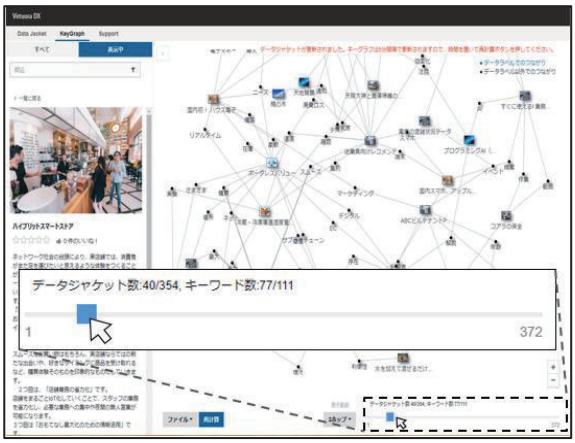


図 4　ユーザー評価度合いに基づくデータジャケット表示数の調整

3.3 データジャケットの指定によるキーグラフ表示

前記 3.1,3.2 では，データジャケット一覧から中心となるデータジャケットを 1 つ選択し，そのデータジャケットを中心として生成されるキーグラフを用いて活用可能なデータジャケットを発見し，アイデア発想を行う方法を示した．一方で，利用者がすでに活用するデータを決定している場合も考えられる．例えば，ビジネス上の連携相手やサービス案の領域が決定していることにより，活用すべきデータに制約がある場合などである．

このような場合に，利活用が必要なデータジャケットを選択して，それらからキーグラフを生成・表示する機能を新たに実装した．（図 5 参照）

これにより，上記の制約がある場合のアイデア創出の他，複数のデータジャケットから共通点（共通するキーワード）を発見するという利用方法も考えられる．

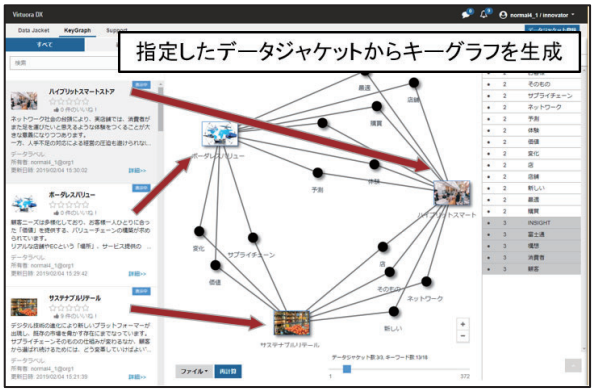


図 5　指定したデータジャケット群による キーグラフ表示

4. まとめ

多様なデータが社会全体で共有され，安心・安全に価値創出活動を進められる場が望まれている．

本稿では，企業や組織，業界・業種をこえたデータ流通・利活用の場の実現を目指し，Virtuora DX 上においてデータジャケットとキーグラフから行われるアイデア発想の概要を紹介しながら，データジャケットの表示の最適化方法について紹介し，考察を行った．今後も，データジャケットと KeyGraph の概念を広く社会に普及させ，この Virtuora DX により，あらゆる企業や組織が安心安全にデータを相互活用し，データを起点としたイノベーションが持続的に創出される社会の実現に貢献していきたい．

参考文献

[富士通 18]　富士通株式会社，"ブロックチェーンを活用し，異業種間共創を促進するデータ流通・利活用サービス「Virtuora DX」を販売開始" . http://pr.fujitsu.com/jp/news/2018/05/14-1.html

[池田 18]　池田 栄次，佐々木 秦芳: 「ブロックチェーン×データジャケット」で実現するデータ流通・利活用社会:～ 異業種共創を加速するコンソーシアムの効用～，人工知能学会全国大会論文集 JSAI2018(0), 1C3OS8b03-1C3OS8b03, 2018 年 6 月．

[大澤 17]　大澤 幸生，早矢仕 晃章，秋元 正博，久代 紀之，中村 潤，寺本 正彦: データ市場，第 1 版，近代科学社，2017 年 3 月．

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

三層ニューラルネットワークにおける Ring-LWE ベース準同型暗号を用いた効率的なプライバシー保護推論処理

Efficient Privacy-Preserving Prediction for Three-Layer Feedforward Neural Networks Using Ring-LWE-based Homomorphic Encryption

手塚　雄大 *1 Takehiro Tezuka　王 立華 *1,*2 Lihua Wang　林 卓也 *2 Takuya Hayashi　Kim Sangwook*1 Sangwook Kim　為井 智也 *1 Tomoya Tamei

大森 敏明 *1 Toshiaki Omori　小澤 誠一 *1 Seiichi Ozawa

*1神戸大学 Kobe University　*2国立研究開発法人 情報通信研究機構 National Institute of Information and Communications Technology

Concerns about privacy of data prevent from making good use of a huge amount of data. Data analysis while preserving privacy is a very important task. In this research, we propose a Privacy-Preserving Machine Learning that can efficiently compute inner product in a three-layered neural network using Ring-LWE-based Homomorphic Encryption. We propose a two-party model consisting of client and server: the former encrypts input data and receives a classification result from a server and the latter performs predicting process over the encrypted data using a trained classification model. This enables that the client acquires the inference result without revealing the privacy of their data and the server protects their model from exposing it. The proposed method costs 10.549 [ms] per one class for prediction process and performed keeping its accuracy close to the case of sigmoid and ReLU.

1. はじめに

近年，データの利活用が活発化される中で，複数の組織が持つ膨大なデータを統合し，利活用することで，新たなビジネスの創出や様々な社会問題の解決につなげることが期待されている．しかし，プライバシー保護やデータの機密性の確保が大きな壁となり，複数組織間でのデータ共有が問題解決につながっている事例は多くない．これに対して，データを秘匿化したまま解析するプライバシー保護データマイニングの技術が注目されている．また，データ分析には多様な知見やリソースが必要であり，分析モデルの公開は知的財産価値の喪失に繋がる．

そのような状況を踏まえて，Dowlin ら [Dowlin 16] は，学習済みの五層の畳み込みニューラルネットワークを暗号データに適用するモデルを構築した．Juvekar ら [Juvekar 18] は，暗号ライブラリ GAZELLE を作成し，garbled circuits を利用した四層ニューラルネットワークの予測モデルを構築した．

本研究では， Ring-LWE ベース準同型暗号を用いたプライバシー保護可能な三層ニューラルネットワーク (Privacy Preserving Three-Layer Feedforward Neural Networks : PP-TLFNNs) における効率的な推論処理を提案する．本稿では，入力層，中間層，出力層をそれぞれ一層ずつのニューラルネットワークを三層ニューラルネットワークと呼ぶことにする．提案手法では，学習済みモデルをもつサーバとサーバの持つモデルを用いて推論を行いたいクライアントの二者を考える．クライアントはデータを Ring-LWE ベースの準同型暗号により暗号化してサーバに暗号文を渡す．サーバは受け取った暗号データに対して学習済みの分類モデルを用いて推論処理を実行した後，推論結果が含まれる暗号文をクライアントに渡す．クライアントは受け取った暗号文を復号した値から対象データの推論結果を得る．ニューラルネットワークにおける活性化関数にはよく ReLU 関数や sigmoid 関数などが用いられるが，これらは有限次元の多項式で表すことはできない．したがって，提案手法では二次の多項式を用いて，中間層における非線形演算を乗法と加法の組み合わせのみで実現し，乗法性と加法性を備える準同型暗号での計算を可能とする．本研究では，データの Packing 方法の工夫により PP-TLFNNs の効率的な推論処理が可能となる．

本稿では，2 節で使用する準同型暗号について，3 節で準同型暗号を用いた三層ニューラルネットワークの効率的な順方向計算について述べる．4 節では推論に要する実時間性と分類精度の妥当性について評価し，5 節でまとめを述べる．

連絡先: 小澤 誠一，神戸大学数理･データサイエンスセンター，〒657-8501 神戸市灘区六甲台町 1-1，TEL: 078-803-6466，E-MAIL: ozawasei@kobe-u.ac.jp

2. 準同型暗号

本研究では，準同型暗号として，Lauter ら [Lauter 11] により提案された Ring-LWE ベースの公開鍵準同型暗号方式を用いる．この方式を LNV-PHE（Public key Homomorphic Encryption）方式と呼ぶ. この方式は，加法準同型性と乗法準同型性を持つため，準同型内積演算ができる．

2.1 表記法

n を 2 のべき乗，$\mathcal{R} = \mathbb{Z}[x]/(x^n+1)$ とする．Ring-LWE ベース暗号の平文空間を $\mathcal{R}_p = \mathcal{R}/p\mathcal{R} = \mathbb{Z}_p[x]/(x^n+1)$，暗号文空間を $\mathcal{R}_q = \mathcal{R}/q\mathcal{R} = \mathbb{Z}_q[x]/(x^n+1)$ とする．

整数 y と自然数 q に対して，$y - zq \in [-q/2, q/2)$ をみたす整数 z がただ一つに定まる．このとき，$[y]_q := y - zq$ とする．

同様に，多項式 $u = \sum_{i=0}^{n-1} u_i x^i \in \mathcal{R}$ に対して，$[u]_q := \sum_{i=0}^{n-1} [u_i]_q x^i$ とする．また，$u^{(t)}$ を以下のように定義する．

$$u^{(t)} := -\sum_{i=0}^{n-1} u_i x^{n-i} = u_0 - u_1 x^{n-1} - \cdots - u_{n-1} x.$$

ここで，[Wang 17b] より，任意の二つの多項式 $u, v \in \mathcal{R}$ に対して，次式が成り立つ．

$$(uv)^{(t)} = u^{(t)} v^{(t)}.$$

また，多項式 u の係数ベクトルを $\overline{u} := (u_0, ..., u_{n-1})$ で表す．

次元が等しいベクトル $\overline{u}$ と $\overline{v} = (v_0, ..., v_{n-1})$ に対して，その内積を $\langle \overline{u}, \overline{v} \rangle = \sum_{i=0}^{n-1} u_i v_i$ で定義する．

任意の有界な集合 $\mathcal{A}$ に対し，一様サンプリングを記号 $a \xleftarrow{\$} \mathcal{A}$ で表す．また，$f \xleftarrow{g} \mathcal{R}_{(0,s^2)}$ によって，各係数を独立に分散 s^2 の離散ガウス分布からサンプリングした多項式 $f \in \mathcal{R}$ を表す．

2.2 LNV-PHE 方式を用いた準同型内積演算

本稿では，公開鍵を pk，秘密鍵を sk とし，$c = (c_1, c_2), c' = (d_1, d_2) \in \mathcal{R}_q^2$ をそれぞれ平文 $m, m' \in \mathcal{R}_p$ の暗号文とする．

平文 m 及び暗号文 c に対する暗号化と復号化をそれぞれ $\mathsf{Enc}(pk, m)$，$\mathsf{Dec}(sk, c)$ とし，それらを以下のように定義する．

$$\mathsf{Enc}(pk, m) = (e_1 a + pe_2, e_1 P + pe_3 + m) \in \mathcal{R}_q,$$
$$\mathsf{Dec}(sk, c) = [c_1 S + c_2]_q \mod p \in \mathcal{R}_p.$$

ここで，$pk = (P, a), sk = S$．ただし，$r, S \xleftarrow{g} \mathcal{R}_{(0,s^2)}$, $a \xleftarrow{\$} \mathcal{R}_q$，$P = pr - aS \in \mathcal{R}_q$ であり，$e_1, e_2, e_3 \xleftarrow{g} \mathcal{R}_{(0,s^2)}$ である．

上記の LNV-PHE 方式における加法 $(c_1 + d_1, c_2 + d_2) = \mathsf{Enc}(pk, m + m')$ と乗法 $(c_1 m', c_2 m') = \mathsf{Enc}(pk, mm')$ 準同型演算ができる．また，[Wang 17a] より以下の計算によって内積準同型演算を実現できる．内積準同型演算 $\mathsf{InnerP}(c, c')$ 及びその復号 $\mathsf{DecIP}(sk, ip)$ は下記の通りである．

$$ip_1 = c_1^{(t)} d_1 \in \mathcal{R}_q,$$
$$ip_2 = c_1 d_2^{(t)} + c_2^{(t)} d_1 \in \mathcal{R}_q,$$
$$ip_3 = (c_2^{(t)} d_2 \in \mathcal{R}_q) \bmod x \in \mathbb{Z}_q$$

とし，

$$\mathsf{InnerP}(c, c') := (ip_1, ip_2, ip_3) = ip \in \mathcal{R}_q \times \mathcal{R}_q \times \mathbb{Z}_q.$$
$$\mathsf{DecIP}(sk, ip) := [(ip_1 S^* + ip_2 S) \bmod x + ip_3]_q \bmod p \in \mathbb{Z}_p.$$

ただし，$S^* = SS^{(t)}$ である．

3. 三層ニューラルネットワークにおける効率的なプライバシー保護推論処理 (PP-TLFNNs)

3.1 TLFNNs の順方向計算

入力次元を N, 中間層のノード数を L, クラス数を C とし，一層目の結合荷重を $W^{(1)} = \left(w_{ij}^{(1)}\right)_{N \times L}$，活性化関数を $g(x) = \alpha x^2 + \beta x + \gamma$，二層目の結合荷重を $W^{(2)} = \left(w_{jk}^{(2)}\right)_{L \times C}$ とすると，TLFNNs の順方向計算による入力ベクトル $U = (u_1, ..., u_N)$ に対する出力 $Z = (z_1, \ldots, z_C)$ は次式で表せる．

$$Z \quad = g(U \cdot W^{(1)}) \cdot W^{(2)}$$

したがって，各クラス $k = 1, \ldots, C$ の出力値 z_k は活性化関数の各係数 α, β, γ について次のように整理できる．

$$\begin{aligned} z_k &= \sum_{j=1}^{L} g\left(\sum_{i=1}^{N} u_i w_{ij}^{(1)}\right) w_{jk}^{(2)} \\ &= \alpha \sum_{j=1}^{L} (\sum_{i=1}^{N} u_i w_{ij}^{(1)})^2 w_{jk}^{(2)} + \beta \sum_{j=1}^{L} (\sum_{i=1}^{N} u_i w_{ij}^{(1)} w_{jk}^{(2)}) \\ &\quad + \gamma \sum_{j=1}^{L} w_{jk}^{(2)} \\ &:= z_k^{\alpha} + z_k^{\beta} + z_k^{\gamma}, \end{aligned}$$

$j = 1, ..., L$ に対し，

$$A_j := \sum_{i=1}^{N} u_i w_{ij}^{(1)} = \langle U, W_j^{(1)} \rangle$$
$$B_j^{[k]} := \sum_{i=1}^{N} u_i w_{ij}^{(1)} w_{jk}^{(2)} = \langle U, W_j^{(1)} w_{jk}^{(2)} \rangle$$

とし，

$$A := (A_1, ..., A_j, ..., A_L), \quad B^{[k]} := (B_1^{[k]}, ..., B_j^{[k]}, ..., B_L^{[k]})$$

とすると，$z_k^{\alpha}, z_k^{\beta}, z_k^{\gamma}$ はそれぞれ次のように書き直せる．

$$\begin{aligned} z_k^{\alpha} &:= \alpha \sum_{j=1}^{L} (\sum_{i=1}^{N} u_i w_{ij}^{(1)})^2 w_{jk}^{(2)} \\ &= \alpha \sum_{j=1}^{L} (\sum_{i=1}^{N} u_i w_{ij}^{(1)})(\sum_{i=1}^{N} u_i w_{ij}^{(1)} w_{jk}^{(2)}) \\ &= \alpha \sum_{j=1}^{L} A_j B_j^{[k]} = \alpha \langle A, B^{[k]} \rangle, \\ z_k^{\beta} &:= \beta \sum_{j=1}^{L} (\sum_{i=1}^{N} u_i w_{ij}^{(1)} w_{jk}^{(2)}) \\ &= \beta \sum_{j=1}^{L} B_j^{[k]} = \beta \langle \mathbf{1}, B^{[k]} \rangle, \\ z_k^{\gamma} &:= \gamma \sum_{j=1}^{L} w_{jk}^{(2)}. \end{aligned}$$

3.2 Packing 方法

ベクトル $V = (v_1, ..., v_N)$ に対して，Packing を下記のように定義する．

$$\mathsf{Poly}_1(V) \quad := \quad \sum_{i=1}^{N} v_i x^{i-1} \tag{1}$$

$$\mathsf{Poly}_2(V) \quad := \quad \sum_{i=1}^{N} v_i x^{(i-1)N} \tag{2}$$

式 (1), (2) を用いると，ベクトル $U = (u_1, ..., u_N), W_j^{(1)} = \left(w_{1j}^{(1)}, \ldots, w_{Nj}^{(1)}\right)$ 及び $W_j^{(2)} = \left(w_{2j}^{(2)}, \ldots, w_{Lj}^{(2)}\right)$ に対して

$$\begin{aligned} V_{A_j}(x) &:= \mathsf{Poly}_1(U)(\mathsf{Poly}_1(W_j^{(1)}))^{(t)} \\ &= A_j + \sum_{j=1}^{\lfloor n/N \rfloor} (0 \cdot x^{jN}) + [\text{other terms}], \\ V_{B_j^{[k]}}(x) &:= \mathsf{Poly}_2(U)(\mathsf{Poly}_2(W_j^{(1)} w_{jk}^{(2)}))^{(t)} \\ &= B_j^{[k]} + \sum_{j=0}^{\lfloor n/N \rfloor} \sum_{k=1}^{N-1} (0 \cdot x^{k+jN}) + [\text{other terms}] \end{aligned}$$

となり，$V_{A_j}(x), V_{B_j^{[k]}}(x)$ それぞれの定数項に $A_j, B_j^{[k]}$ が現れる．また，$V_{A_j}(x)$ と $V_{B_j^{[k]}}(x)$ の係数ベクトル $\overline{V_{A_j}(x)}$ と $\overline{V_{B_j^{[k]}}(x)}$ の内積は以下のようになる．

$$\langle \overline{V_{A_{[j]}}(x)}, \overline{V_{B_{[j]}^{[k]}}(x)} \rangle = A_j B_j^{[k]}. \tag{3}$$

3.3 暗号文を用いた TLFNNs の順方向計算

暗号文 $\mathsf{Enc}(\mathsf{Poly}_1(U)) = (c_1, c_2)$, $\mathsf{Enc}(\mathsf{Poly}_2(U)) = (d_1, d_2)$ に対して，下記の準同型計算を行うと，$V_{A_j}(x)$ と $V_{B_j^{[k]}}(x)$ の暗号文を求めることができる．

$$\begin{aligned}(\widehat{c_{1,[j]}}, \widehat{c_{2,[j]}}) &:= \left(c_1(\mathsf{Poly}_1(W_j^{(1)}))^{(t)}, c_2(\mathsf{Poly}_1(W_j^{(1)}))^{(t)}\right)\\ &= \mathsf{Enc}(\mathsf{Poly}_1(U)(\mathsf{Poly}_1(W_j^{(1)}))^{(t)})\\ &= \mathsf{Enc}(V_{A_j}(x)),\end{aligned}$$

$$\begin{aligned}(\widehat{d_{1,[j]}^{[k]}}, \widehat{d_{2,[j]}^{[k]}}) &:= \left(d_1(\mathsf{Poly}_2(W_j^{(1)}))^{(t)} w_{jk}^{(2)}, d_2(\mathsf{Poly}_2(W_j^{(1)}))^{(t)} w_{jk}^{(2)}\right)\\ &= \mathsf{Enc}(\mathsf{Poly}_2(U)(\mathsf{Poly}_2(W_j^{(1)} w_{jk}^{(2)}))^{(t)})\\ &= \mathsf{Enc}(V_{B_j^{[k]}}(x)).\end{aligned}$$

したがって，式 (3) 及び [Wang 17a] より $A_j B_j^{[k]}$ は $(\widehat{c_{1,[j]}}, \widehat{c_{2,[j]}}), (\widehat{d_{1,[j]}^{[k]}}, \widehat{d_{2,[j]}^{[k]}})$ を用いて内積準同型演算で計算でき，また，加法準同型性より $z_k^\alpha = \alpha\langle A, B^{[k]}\rangle = \alpha\sum_{j=1}^L A_j B_j^{[k]}$ に対応する暗号文は下記のようにして計算できる．

$$\begin{aligned}ip_\alpha^{[k]} &= \alpha\sum_{j=1}^L \mathsf{InnerP}((\widehat{c_{1,[j]}}, \widehat{c_{2,[j]}}), (\widehat{d_{1,[j]}^{[k]}}, \widehat{d_{2,[j]}^{[k]}}))\\ &:= (ip_{\alpha,1}^{[k]}, ip_{\alpha,2}^{[k]}, ip_{\alpha,3}^{[k]}).\end{aligned}$$

ただし，

$$\begin{aligned}ip_{\alpha,1}^{[k]} &= \alpha\sum_{j=1}^L \widehat{c_{1,[j]}}^{(t)} \widehat{d_{1,[j]}^{[k]}} \in \mathcal{R}_q,\\ ip_{\alpha,2}^{[k]} &= \alpha\sum_{j=1}^L \left(\widehat{c_{2,[j]}}^{(t)} \widehat{d_{1,[j]}^{[k]}} + \widehat{c_{1,[j]}} \widehat{d_{2,[j]}^{[k]}}^{(t)}\right) \in \mathcal{R}_q,\\ ip_{\alpha,3}^{[k]} &= (\alpha\sum_{j=1}^L \widehat{c_{2,[j]}}^{(t)} \widehat{d_{2,[j]}^{[k]}} \bmod x) \in \mathbb{Z}_q.\end{aligned}$$

一方，$B_j^{[k]} = \langle(1, 0, ..., 0), \overline{V_{B_{[j]}^{[k]}}(x)}\rangle$ であるため，$z_k^\beta = \beta\sum_{j=1}^L B_j^{[k]}$ の暗号文は下記で計算できる．

$$ip_\beta^{[k]} = (0, ip_{\beta,2}^{[k]}, ip_{\beta,3}^{[k]})$$

ただし，

$$\begin{aligned}ip_{\beta,2}^{[k]} &= \beta\sum_{j=1}^L \widehat{d_{1,[j]}^{[k]}} \in \mathcal{R}_q,\\ ip_{\beta,3}^{[k]} &= (\beta\sum_{j=1}^L \widehat{d_{2,[j]}^{[k]}} \bmod x) \in \mathbb{Z}_q.\end{aligned}$$

また，

$$ip_\gamma^{[k]} = (0, 0, z_\gamma^{[k]}).$$

したがって，$z_k = z_k^\alpha + z_k^\beta + z_k^\gamma$ の結果を含む暗号文は

$$ip^{[k]} = ip_\alpha^{[k]} + ip_\beta^{[k]} + ip_\gamma^{[k]} := (ip_1^{[k]}, ip_2^{[k]}, ip_3^{[k]}).$$

により求められる．ただし，

$$\begin{aligned}ip_1^{[k]} &= ip_{\alpha,1}^{[k]} \in \mathcal{R}_q,\\ ip_2^{[k]} &= ip_{\alpha,2}^{[k]} + ip_{\beta,2}^{[k]} \in \mathcal{R}_q,\\ ip_3^{[k]} &= ip_{\alpha,3}^{[k]} + ip_{\beta,3}^{[k]} + z_k^\gamma \in \mathbb{Z}_q.\end{aligned}$$

3.4 スキーム

本節では，プライバシー保護可能な推論処理のスキームを 5 つの Step に分けて説明する．ただし，$\mathbb{Z}[x]/(x^n+1)$ の次数 n は，入力ベクトルの次元 N に対して $n \geq N^2$ とする．

Step-0 サーバは $k = 1, ..., C$ に対して，事前に

$$\widetilde{\mathsf{Poly}_\alpha^{[k]}} = \alpha\sum_{j=1}^L \mathsf{Poly}_1(W_j^{(1)})\mathsf{Poly}_2(W_j^{(1)})^{(t)} w_{jk}^{(2)},$$

$$\widetilde{\mathsf{Poly}_\beta^{[k]}} = \beta\sum_{j=1}^L \mathsf{Poly}_2(W_j^{(1)})^{(t)} w_{jk}^{(2)},$$

及び $z_k^\gamma = \gamma\sum_{j=1}^L w_{jk}^{(2)}$ を計算しておく．

Step-1 クライアントは自分のデータ U を方式 (1) , (2) を用いてエンコードし，下記のように暗号化してからサーバへ送る．

$$\begin{aligned}\mathsf{Enc}(pk, \mathsf{Poly}_1(U)) &= (c_1, c_2),\\ \mathsf{Enc}(pk, \mathsf{Poly}_2(U)) &= (d_1, d_2).\end{aligned}$$

Step-2 サーバは (c_1, c_2) と (d_1, d_2) を受け取ったら，下記の準同型計算を行い，

$$\begin{aligned}ip_1^{[k]} &= c_1^{(t)} d_1 \widetilde{\mathsf{Poly}_\alpha^{[k]}} \in \mathcal{R}_q,\\ ip_2^{[k]} &= d_1\left(c_2^{(t)}\widetilde{\mathsf{Poly}_\alpha^{[k]}} + \widetilde{\mathsf{Poly}_\beta^{[k]}}\right) + c_1 d_2^{(t)}\widetilde{\mathsf{Poly}_\alpha^{[k]}}^{(t)} \in \mathcal{R}_q,\\ ip_3^{[k]} &= (d_2\left(c_2^{(t)}\widetilde{\mathsf{Poly}_\alpha^{[k]}} + \widetilde{\mathsf{Poly}_\beta^{[k]}}\right) + z_k^\gamma \bmod x) \in \mathbb{Z}_q.\end{aligned}$$

結果 $ip^{[k]} := (ip_1^{[k]}, ip_2^{[k]}, ip_3^{[k]}) \in \mathcal{R}_q \times \mathcal{R}_q \times \mathbb{Z}_q, k = 1, 2, ..., C$ をクライアントに渡す．

Step-3 クライアントはサーバから受け取った結果を復号する．

$$\mathsf{DecIP}(sk, ip^{[k]}) = z_k \quad (k = 1, ..., C).$$

Step-4 クライアントは $\{z_k\}$ の最大値

$$z_{k^*} = \max_{k\in\{1,...,C\}}\{z_k\}$$

を求め，データ U をクラス k^* に分類する．

4. 実験

4.1 速度評価

3.4 節における各 Step の処理時間のベンチマークを測定した結果を表 1 に示す．計測は Core i7-7700K(4.20 GHz) のシングルスレッドで行った．実装には次のパラメータを用いた．

$$\begin{aligned}p &= 32749 \times 32719 \times 32717 \times 32713\\ q &= 2^{96} - 2^{32} + 1\\ n &= 4096, \quad s = 8.0\end{aligned}$$

平文は $p_i \in \{32749, 32719, 32717, 32713\}$ それぞれについて暗号化，計算を行い，復号時に中国剰余定理で p に持ち上げることで高速化を行った．また，実数値は 2^{16} を掛けて床関数により整数化を行った．

表 1: 各 Step の処理時間 (C : クラス数)

処理	時間 [ms]
事前計算 (Step-0)*1	$2747.336 \times C$
暗号化 (Step-1)	8.561
PP-TLFNNs の順方向計算 (Step-2)	$10.549 \times C$
復号 (Step-3)	$0.332 \times C$

表 2: 使用するデータセット

	データ数	特徴量数	クラス数
Satellite	6435	36	6
Australian	690	14	2
German	1000	20	2

Step-0 の事前計算は，モデルを製作した時点でサーバがあらかじめ計算しておく処理であり，クライアントとサーバ間での処理時間に直接的に影響を及ぼさないため，処理速度の実用性の評価はその他の結果から行うこととする．暗号化ではクラス数に依存せず， PP-TLFNNs の順方向計算および復号にかかる時間はクラス数に依存する．Satellite データセットであれば $C = 6$ であるため，順方向計算に 63.294 [ms]，復号に 1.992[ms] で処理できる．これらの結果から，処理速度は十分実用的であると言える．

4.2 精度評価

分類精度の評価には，UCI Machine Learning Repository [Dua 17] にて公開されている 3 つのデータセット Satellite : Statlog (Landsat Satellite) Data Set，Australian : Statlog (Australian Credit Approval) Data Set，German : Statlog (German Credit Data) Data Set を使用した．それぞれのデータセットに関する情報を表 2 に示す．また，これらのデータセットを使用するにあたって，前処理として各データそれぞれで平均が 0, 分散が 1 となるように正規化を行った．中間層の活性化関数には ReLU 関数を二次の多項式により近似したものを用いた．ここで，ReLU 関数の近似は学習過程で中間層に入力された最大値と最小値を含む $[-30, 30]$ の範囲で最小二乗法により行った．

モデルの学習では確率的勾配降下法により学習を行った．そして，求めた結合荷重を提案手法の推論処理で使用した．

本稿の実験では 5 分割の交差検定を行い，分類精度を評価した．中間層の活性化関数に Sigmoid 関数，ReLU 関数を用いた場合と二次関数を用いた PP-TLFNNs の分類精度を表 3 に示す．

表 3 の結果から，PP-TLFNNs は Sigmoid 関数や ReLU 関数を用いた TLFNNs に対して，近い精度で推論処理を行えたことがわかる．

5. まとめ

近年，データのプライバシーに対する懸念が，データの利活用の活発化への大きな障壁となっていた．それにより，複数組織が膨大なデータを所有していても，それらをより効果的に利用することが困難な問題があった．そのような状況に対して，データのプライバシーを保護した状態で解析を行える技術が求められている．そこで，本研究ではデータを暗号化した状態で推論処理を行えるモデルを構築した．

*1 実装では漸近的に高速なアルゴリズムを使用していないため，高速化の余地は残っている．

表 3: 分類精度の比較．PP-TLFNNs は暗号文実装のモデル，TLFNNs は平文実装のモデルを指す．

中間層における活性化関数	PP-TLFNNs	TLFNNs	
	二次関数	Sigmoid 関数	ReLU 関数
Satellite	0.871 ± 0.016	0.899 ± 0.027	0.900 ± 0.023
Australian	0.857 ± 0.041	0.875 ± 0.028	0.875 ± 0.039
German	0.741 ± 0.020	0.774 ± 0.026	0.778 ± 0.036

本稿では，準同型暗号を用いた三層ニューラルネットワークにおいて，Packing 方法の工夫により効率的な推論処理を提案した．速度評価では，それぞれの処理を実用的な時間で実現することを示した．また，精度評価では二次関数を用いたことによる精度の低下は見られたが，依然として高い精度で分類できることを示した．これにより，複数組織がデータの情報を漏らすことなく，それぞれのデータを解析することが可能となり，さらなるデータ利活用の促進が期待できる．

今後は活性化関数の近似手法の改良や暗号方式の工夫によるネットワーク構造の拡張等による分類精度の向上，さらには実データへの適用を目指したい．

謝辞

本研究の成果は JST CREST 研究領域「イノベーション創発に資する人工知能基盤技術の創出と統合化」研究課題「複数組織データ利活用を促進するプライバシー保護データマイニング」JST CREST JPMJCR168A，JSPS 科研費 JP15K00028 の助成を受けて得られたものです．

参考文献

[Dowlin 16] Dowlin, N., Gilad-Bachrach, R., Laine, K., Lauter, K., Naehrig, M. and Wernsing, J.: Cryptonets: Applying neural networks to encrypted data with high throughput and accuracyrq, *ICML 2016*, pp. 201-210 (2016)

[Juvekar 18] Juvekar, C., Vaikuntanathan, V. and Chandrakasan, A.: GAZELLE: A low latency framework for secure neural network inference, *Usenix Security 2018,* pp. 1651–1669 (2018).

[Lauter 11] Naehrig, M., Lauter, K. and Vaikuntanathan, V.: Can homomorphic encryption be practical?, *CCSW'11*, pp. 113–124 (2011).

[Wang 17a] Wang, L., Hayashi, T., Aono, Y. and Phong., L. T.: A generic yet efficient method for secure inner product., *NSS 2017, LNCS 10394*, pp. 217–232 (2017).

[Wang 17b] Wang, L., Aono, Y. and Phong, L. T.: A new secure matrix multiplication from Ring-LWE., *CANS 2017, LNCS 11261*, pp. 93–111 (2017).

[Dua 17] Dua, D. and Taniskidou, E.K.: UCI Machine Learning Repository, `http://archive.ics.uci.edu/ml`, (2017)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

データプラットフォームにおける異分野データネットワークの成長過程に関する一考察

Network Growth Process on Cross-disciplinary Data in the Data Platform

早矢仕晃章[*1]
Teruaki Hayashi

大澤幸生[*1]
Yukio Ohsawa

[*1] 東京大学大学院　工学系研究科　システム創成学専攻
Department of Systems Innovation, School of Engineering, The University of Tokyo

Data are generated every moment, and new kinds of data appear one after another in the real world. In recent years, problem-solving by exchanging/combining data among different domains is one of the social demands. In the Web and Social Networking Services, many models to express the complex networks and their growth process have been proposed. However, the network model focusing on data from different domains in the data platform has not been discussed. In this study, we observed the process of network growth of cross-disciplinary data using the summarized information about data–data jacket–, and considered the dynamic change of the characteristics of the network in its growth process. From the characteristics of the dynamic network, we demonstrated several strategies as data platformers.

1. はじめに

実社会において，多種多様なデータが時々刻々と生成され，新たな種類のデータが続々と登場してきている．近年，データを交換可能な材として，異分野間でデータを交換・結合させることで問題解決する動きが活発になってきている．このような取引が行われるプラットフォームとして，データ流通市場，データ取引市場など，多様な形態のデータ市場及び関連サービスが萌芽してきている．データ市場では多様なステークホルダーのデータがプラットフォームに登場し，他者と交換・取引されるという点で，データをノードとしたネットワークは絶えず成長するネットワークと言える．さらに，国内外でデータプラットフォームのサービスが立ち上がりつつある中で，プラットフォーマーとしての覇権争いが行われつつある[早矢仕 19a]．しかし，どれほどの規模のデータがあればプラットフォームとして十分に機能するのか，また，どのような種類のデータをプラットフォームで取り扱うことで，自律的にコミュニティが成長するのか，ということは研究されてきていない．新たなデータがプラットフォームに参入する過程をダイナミックに捉え，モデル化することは，将来のデータのプラットフォームの活性化において重要な課題である．

Web や SNS などでは，データの複雑ネットワーク及びその成長過程を表現する様々なモデルが提案されてきた（[伏見 14][Osaka 17]など）．しかし，データ市場における異業種のデータ固有の特徴に着目したモデルは十分議論されてきたとは言い難い．そこで本研究では，データ概要情報データジャケットを用いて異種のデータのネットワークが成長する過程を観察し，成長過程におけるネットワークのダイナミックな特徴の変化について考察する．

連絡先：
早矢仕晃章，東京大学大学院工学系研究科システム創成学専攻，hayashi@sys.t.u-tokyo.ac.jp
This research was supported by JST, CREST (JPMJCR1304). 本研究にご協力いただいた共同印刷株式会社の皆様に心より感謝申し上げます．

2. 異分野データの動的ネットワーク作成

2.1 データジャケット

データジャケット（Data Jacket: DJ）は，データの概要情報を記述するためのフレームワークである[Ohsawa 13]．データに関する説明文や含まれる変数の名前（変数ラベル），保存形式，共有条件などをデータ概要情報として記述することでデータを秘匿としたまま，異なる分野のデータについて理解可能となる（図 1）．DJ はメタデータの一形態であるが，通常のメタデータと異なり，DJ は人間がデータについて理解し，議論可能とすることを目的としている．また，異なるフォーマット，異なる粒度の変数ラベルを有するデータを共通の記述ルールによって構造的にメタデータ化することで，様々な形式のデータを統一的に扱うことができる．この特徴を用いて，[早矢仕 18b][Hayashi 18]は，変数ラベル及びデータが持つコンテキスト（文脈）を介した異分野データネットワークの解析を行い，ネットワークの特徴及びデータ市場における特徴的なデータの振る舞いについて論じた．なお，変数ラベルとは，データ固有の変数を自然言語によって記述された説明文を意味する．

DJ No. XX【購買履歴データ】	
概要	東京都の◯◯スーパーマーケットで収集されている顧客の購買行動履歴。
収集方法・コスト	ポイントカードとPOSによって取得
共有条件	共有不可
データの種類	表形式、テキスト、数値
保存形式	CSV
分析・シミュレーション	時系列分析
変数ラベル	氏名、性別、顧客ID、支払金額、購入品目、日にち
分析結果	• その日の売上の計算 • 今後の売上の予測と仕入れの推定
期待される分析	顧客の購買行動とリピート率を計算し、ロイヤルカスタマーの特定が可能かもしれない。
コメント	有効なデータの組み合わせが発見されればデータの提供あるいはコラボレーションもあり得る。

図 1　DJ の記入例（[早矢仕 18a]より引用）

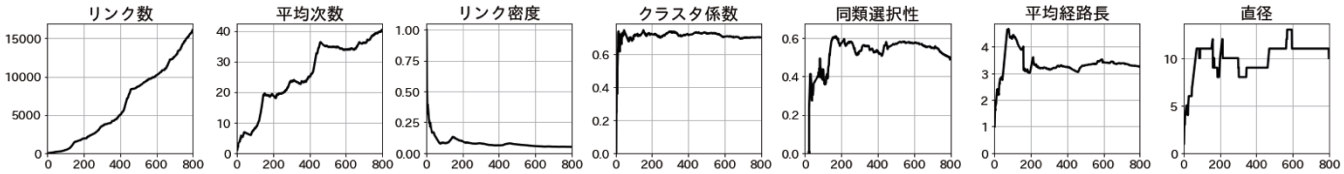


図 2　DJ（ノード数）の増加に伴う各ネットワーク特徴量の変化（横軸：ノード数）

2.2 目的と解析アプローチ

従来研究では，時間を固定した静的ネットワークを対象とした分析に留まっていた．実社会のネットワークの多くは時間変化する動的ネットワークである．そして，データ市場におけるプラットフォームでは新たなデータが参入することでデータ間の関係は時々刻々と変化するものと考えられる．ネットワークの成長過程を観察することにより，特徴的な構造変化を与えたデータの振る舞いと特徴を捉え，考察を行うことが本研究の目的である．データプラットフォームをネットワークで議論することで，ネットワークの諸問題が適用可能となり，データ同士の関係性を理解できる．変数が共通しているデータは互いに似ているという仮定によって，データの類似性を考慮したネットワークを議論できる．

本研究では，DJ ストア[1][早矢仕 16]から入手できる 1316 件の DJ のうち，ネットワークで表現した際に最大コンポーネントを構成する 798 件の DJ を対象として分析を行う．なお，ネットワークの各ノードは DJ を表し，DJ 同士が共通する変数ラベル（Variable Label: VL）を保有している場合にリンクが張られるものとした．例えば，「東京都の病院の位置情報」と「東京都に設置されている警察署の設置情報」という DJ が「緯度」と「経度」という変数ラベルを共通して有していた場合，2 つの DJ にはリンクが張られる．また，ネットワークの性質を理解する上で，ネットワークの基本統計量を用いる．本稿では，動的ネットワークの変化の特徴として，ネットワークのリンク数，各ノードの平均次数，リンク密度，クラスタ係数，同類選択性，平均経路長，ネットワーク直径の 7 項目を選んだ．

3. 結果と考察

3.1 異分野データのネットワーク

DJ の VL によるネットワークを用いることで，データのネットワークがどのように成長したのかということを時系列に従って分析することが可能となる．表 1 は 798 件の DJ の静的ネットワークの特徴量を表したものである．表にあるように，VL を介したデータのネットワークの平均次数は 40.29 と高く，リンク密度 0.051 と低い値を示している．一方，クラスタ係数は 0.702 と高い値である．リンク密度が大域的な密度を表し，クラスタ係数が局所的な密度を意味する指標であることを考慮すると，異分野データのネットワークは局所的に密なネットワークを作りやすく，大域的に密度が低い性質を持っていることが分かる．つまり，同じような変数を持つデータ同士が局所的に大きな塊を作るため，大域的に見たときに疎な構造となっているのである（図 3）．

同類選択性は隣接する 2 ノード間の次数相関を示す指標である．経験的にタンパク質などの自然界のネットワークやインターネットなどの工学系ネットワークは同類選択性が負になることが知られている．一方で，論文の共著関係や SNS などの人間関係を表すネットワークの同類選択性は正の値となりやすい．本研究の同類選択性は 0.489 と高く，同じようなデータは互いに密なネットワークを構成するものの他の領域のデータとは繋がりが疎となる特徴を持つことが分かった．また，ノード数に対して平均経路長は比較的短く，直径も小さいことから，異分野データのネットワークはスモールワールド性を有していると考えられる．

図 2 は DJ による異分野データの動的ネットワークとして捉えた際に，ノード数の増加に伴って変化するネットワークの特徴を表す指標をグラフ化したものである．リンク数は DJ の増加に伴い単調に増加しているが，あるデータの出現によってネットワーク全体のリンク数が急増する点が数箇所見られる．例えば，721 番目の「駐車場利用データ」，481 番目の「公衆トイレ情報」などはそれぞれ 160 リンク，102 リンクと，データ登場時に極めて高いリンク数を獲得している．リンク数の変化量は平均次数にも影響している．平均次数を見ると，基本的には増加しているが，リンク数の急激な増加が起こった直後に減少している部分が観察できる．これは極端に多くのリンクを獲得したデータが現れた後に見られる．極端に多くのリンクを獲得したデータが出現した直後は通常のリンク数を持つデータが乗っていくるため，ネットワーク全体の平均次数が低下するものと考えられる．

続いて，リンク密度とクラスタ係数の 2 つの指標の変化を観察すると，ノード数の増加に伴い，リンク密度は急激に低下し，一方でクラスタ係数は初期から高い値を示しており，0.70 前後の値を上下し，安定した構造を見せている．これは，データのネットワークが初期の段階から局所的に密な構造となることが分かる．また，データの追加によって密度は大きく変化することなく，局所的に密度が高い部分にてリンクを獲得していくものと考えられる．次数相関を表す同類選択性はクラスタ係数と同じく，初期から高い正の値を示しており，同じような次数のデータが互いに繋がりやすい傾向がネットワーク生成の初期から現れていることが分かる．平均経路長の変化を見ると，初期は比較的長く，その後 3.1 から 3.5 の間を推移し，安定している．ネットワーク直径もノードの増加による大きな変化はなく，10 前後を推移している．異分野データのネットワークはデータ数が少ない初期の段階から局所的に密かつ大域的に疎であり，平均次数は増加するものの他のネットワーク特徴量はデータ数 200 件ほどから急激に変化しないことが分かった．

表 1　ネットワークの特徴量

	最大コンポーネント
ノード数	798
リンク数	16076
平均次数	40.29
リンク密度	0.051
クラスタ係数	0.702
同類選択性	0.489
平均経路長	3.26
直径	10

[1] http://160.16.227.37/index/DJStore

図 3　異分野データネットワーク

3.2 データプラットフォーマーとしての戦略立案例

前節で議論したネットワークの特徴から，データプラットフォーマーとしていくつかの戦略を立案することができる．まず，自律的なコミュニティ成長においては，そのコミュニティおいて核となるデータが必要となる．すなわち，ネットワークにおいて中心性が高いデータやクラスタにおいてリンク数が高いデータが該当する．予めこのような特徴量を高くし得るデータをプラットフォームに含めておくことにより，新たなデータがプラットフォームに参入してきた際にリンクを獲得する可能性が高くなる．データ提供者にとって，自身の保有するデータがどれだけ他のデータと結合可能性が高く，周りの他のデータと比較して優位性があるのかということがデータ提供のモチベーションになるからである．

また，本稿におけるネットワークでは，同類選択性が高く，リンク密度が低い．そのため，同類選択性を低くし，リンク密度を高めるデータをプラットフォームに参入させるという戦略も考えられる．これにより，人間関係のネットワークのように局所的に強い繋がりが表れやすい「タコツボ化」が緩和し，他の領域のデータとの繋がりが生まれやすくなる．異分野データ連携を活性化させる上でプラットフォーマーは，同類選択性が高く，リンク密度が低い傾向が観察されたら，媒介中心性などを高めるデータ保有者のデータをプラットフォームに提供すると良いだろう．

4. 結論

本稿では，データ概要情報データジャケットを用いて異種のデータのネットワークが成長する過程を観察し，成長過程におけるネットワークのダイナミックな特徴の変化について考察した．本稿で扱った異分野データのネットワークは，[早矢仕 18a][早矢仕 19b]から始まった研究であり，まだまだ発展途上である．本稿のように複雑ネットワークの諸問題に落とし込むことで多様な手法が適用可能となり，データプラットフォームに関する研究のさらなる発展が期待できると考えている．

Barabási らの研究から派生し，成長するネットワークには様々なモデルや時間を考慮したテンポラルネットワークなどの動的ネットワークに関する研究が進んできた（[Leskovec 05][Holme 13]など）．また，ネットワークにおける新参のノードが「下剋上」を起こして中心的な存在になることが知られている[Bianconi 01]．今後の研究では，「あるデータの登場によってどのように他のデータとの関係性が変化していくのか」という成長モデルを構築するとともに，どのような変数を含むデータを新たに設計することによって，プラットフォーム上で優位な存在となるのかということを明らかにしていきたい．また，異分野データ連携において，データの変数のみならず，文脈における共通性も重要な特徴である．データランドスケープ[早矢仕 18b]を導入したネットワークの成長過程を観察し，文脈を考慮した潜在的なデータの繋がりと動的変化を議論する必要があるだろう．また，部分的に観測されたネットワーク構造から残りの構造を推定するリンク予測問題と考えれば，潜在的なリンクを推定し，未知の VL または秘匿 VL の存在を明らかにすることができるなど，データ市場及びデータ流通プラットフォームにおいて重要な知見を得ることができると考えている．

参考文献

[Bianconi 01]　Bianconi, G., Barabási. A.L.: Bose-Einstein Condensation in Complex Networks, Physical Review Letters, Vol.86, No.24, pp.5632-5635 (2001)

[伏見 14]　伏見卓恭, 斉藤和巳, 風間一洋: PageRank に基づく動的ネットワークの構造変化抽出, 知識ベースシステム研究会, Vol.B4, No.02, pp.31-37 (2014)

[早矢仕 16]　早矢仕晃章, 大澤幸生, “Data Jacket Store: データ利活用知識構造化と検索システム,” 人工知能学会論文誌, Vol.31, No.5 (2016)

[早矢仕 18a]　早矢仕晃章, 大澤幸生: データ市場におけるデータのネットワークと関係性の分析−データの属性と繋がりからの考察−, 信学技報, 人工知能と知識処理研究会, Vol.117, No.440, pp.49-54 (2018)

[早矢仕 18b]　早矢仕晃章, 大澤幸生: データ 3.0 時代のデータランドスケープ, 2018 年度人工知能学会全国大会 (2018)

[Hayashi 18]　Hayashi, T., Ohsawa, Y.: The Difference between Variable-based and Context-based Networks of Data Using Data Jackets, 22nd International Conference on Knowledge Based and Intelligent Information and Engineering System, Vol.126, pp.1740-1747 (2018)

[早矢仕 19a]　早矢仕晃章, 小口裕, 飛沢省二, 大澤幸生: 日本・米・欧州・中国のデータ市場ビジネスの動向, 信学技報, 人工知能と知識処理研究会 (2018)

[早矢仕 19b]　早矢仕晃章, 岩永宇央, 岩佐太路, 大澤幸生: データジャケットを用いた異分野連携に資するデータの特徴とネットワーク分析, 日本知能情報ファジィ学会誌, in Press (2019)

[Holme 13]　Holme, P., Saramäki, J.: Temporal Networks, Understanding Complex Systems, Springer (2013)

[Leskovec 05]　Leskovec, J., Kleinberg, J., Faloutsos, C.: Graphs over time: densification laws, shrinking diameters and possible explanations. the 11th International Conference on Knowledge Discovery in Data Mining, pp.177-187 (2005)

[Ohsawa 13]　Ohsawa, Y., Kido, H., Hayashi, T., and Liu, C.: Data Jackets for Synthesizing Values in the Market of Data, 17th International Conference in Knowledge Based and Intelligent Information and Engineering Systems, Procedia Computer Science, Vol.22, pp.709-716 (2013)

[Osaka 17]　Osaka, K., Toriumi, F., Sugawara, T.: Effect of direct reciprocity and network structure on continuing prosperity of social networking services, Computational Social Networks, Springer, Vol.4, No.2, pp.1-20 (2017)

Organized Session | Organized Session | [OS] OS-2

[2G4-OS-2b] データ流通社会における技術基盤と異分野連携(2)

早矢仕 晃章（東京大学）、大澤 幸生（東京大学）

Wed. Jun 5, 2019 3:20 PM - 4:40 PM Room G (302A Medium meeting room)

[2G4-OS-2b-01] Proposing the Enactment of Data Registration

〇Makoto Koike[1,2] （1. Makoto Koike Microwave Research Institute, 2. Micro Patent Office）

3:20 PM - 3:40 PM

[2G4-OS-2b-02] Design of Data Flow Process among Multiple Organizations for PR Measurement

〇Kunihiko Ueshima[1], Taito Tosaka[2], Kohei Taniguchi[3], Teruaki Hayashi[4], Yukio Ohsawa[4] （1. Japan Data Exchange, Inc., 2. OZMA, Inc., 3. Kikkoman Corporation, 4. School of Engineering, The University of Tokyo）

3:40 PM - 4:00 PM

[2G4-OS-2b-03] (Invited talk) Expectation for data trading

〇Aki Ando[1] （1. ABeam Consulting Ltd.）

4:00 PM - 4:20 PM

[2G05-07-4] Panel discussion

4:20 PM - 4:40 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

データ登録制度の提案
Proposing the Enactment of Data Registration

小池 誠[*1, 2]
Makoto Koike

[*1] 小池誠マイクロ波研究所 Makoto Koike Microwave Research Institute
[*2] マイクロ特許事務所 Micro Patent Office

An amendment to the Unfair Competition Act in 2018 in Japan provides for civil remedies such as damages and injunction against unfair acquisition, use and disclosure of data that satisfy certain requirements. However, it is not easy to prove the ownership and content of the data, let alone the change of right concerning data. Therefore, this presentation presents the proposal of enacting the registration of data with a legal framework of ensuring bona fide commercial transactions thereof. This presentation does not concern with a privacy but property aspect of data.

1. はじめに

自動車の自動運転では三次元地図データが活用されており，人工知能の機械学習，特に教師あり学習ではデータセットが活用されている．このようにデータには財産的価値があるので，平成 30 年に不正競争防止法が改正され，管理性，相当蓄積性などの一定の要件を満たすデータについては，限定提供データとして保護対象に追加された[小池 19]．限定提供データの不正取得行為，不正使用行為及び不正開示行為は，不正競争とされ（2 条1項），差止請求権(3 条)，損害賠償請求権(4条)などの民事的措置が認められるようになった．

ここで，データは知的財産の一種であるが，知的財産としては，特許法が定める発明，意匠法が定める意匠，商標法が定める商標，著作権法が定める著作物などがある．特許権，意匠権，商標権などの産業財産権は，不動産と同様に登録という一種の公示制度を採用しており，第三者は登録番号から権利の有無及び内容について確認することができる．

しかしながら，限定提供データについては，産業財産権と異なって，権利の有無が明確でなく，第三者がデータに関する権利について調査することも容易ではない．

そこで，本稿は，産業財産権登録制度，著作権登録制度などを参考にして，データ登録制度を創設し，データに関する権利の有無を明確化することを提案する．

なお，データが個人情報に関するときには，不正競争防止法に限られず，個人情報保護法も適用されることになる．しかしながら，紙片の都合により個人情報保護法は本稿の対象外である．

2. 登録制度の必要性

2.1 データの特徴

限定提供データは，電磁的方法，即ち，電子的方法，磁気的方法その他人の知覚によっては認識することができない方法により相当量蓄積され，管理されている情報と定義されているので（不正競争防止法2条7項），権利の対象は人の知覚によって認識することができず，その特定が必ずしも容易でない．

例えば，書籍，音楽などの著作物は人の知覚により認識することができ（著作権法2条1項1号），この認識を通じて特定することができるが，この点において相違する．

連絡先：小池誠，patentcom007@gmail.com

確かに，データジャケットに記載されたメタデータで，限定提供データは特定できる[Ohsawa 15]．しかしながら，権利者，及び，侵害と訴えられた当事者が，データジャケットのフォーマットでメタデータを記録しているとは限らない．

また，限定提供データの紛争では，メタデータのレベルで，データの同一性を判断するのでなく，メタデータの対象となるデータのレベルでデータの同一性を判断することが求められる．

そこで，限定提供データの内容，即ち，権利の対象を登録で明確化することが所望される．

2.2 客体の同一性

限定提供データでなく，まず，著作権で考える．著作権法では，2人の創作者が別箇，独立に同一又は酷似する著作物を創作したときには，2つの著作権が併存する．著作権法では，権限なく他人の著作物を複製する行為が侵害となる．

著作権を巡る現実の紛争では，侵害と訴えられた当事者が，独自に著作物を創作して，偶然，似てしまったのか，それとも著作権者の著作物を複製したのか，という事実の争いになることがある．結局のところ，権利者が，相手の行為に関する事実を立証できるかという立証の問題に帰着する．

限定提供データも著作物と同様であり，理論的には，2人が同一のデータを偶然，蓄積したときには別箇，独立に権利が発生する．しかしながら，現実にそのような事象が起きる確率は小さく，データの漏洩，不正アクセスなどの権利侵害があったのではないかという疑義が生じる．このような紛争を想定したときには，限定提供データの内容が，登録で明確になっていることが所望される．

そこで，主に著作権の登録を参酌して，以下，データ登録制度について提案する．

3. 登録制度の概略

3.1 権利の発生

限定提供データに関する財産権は，所定のデジタルデータを蓄積し，管理することで発生するが，これらは事実行為であり，法律行為でない．ここで，特許権などの産業財産権では設定の登録により権利が発生し（特許法 66 条等），移転などでも登録が効力発生要件に定められている（特許法 98 条等）．すると，産業財産権のように限定提供データの登録が効力発生要件と定めるのは適切でない．

一方，著作権法では，著作物の創作により，著作権が発生するが，著作権の発生については登録が不要であり，©という記号を付すというような方式も不要である．

著作権法は著作物の登録について規定しており，著作者，第 1 発行年月日，創作日が登録できる（著作権法 75 条，76 条，76 条の 2），登録は第三者対抗要件とされている（著作権法 77 条）．同様に限定提供データの登録は第三者対抗要件に定めることが求められる．

3.2 登録の要件

著作物の登録には，原則として，公表という要件が定められている（著作権法 75 条，76 条，76 条の 2）．ただし，プログラムの著作物については公表が要件とされておらず，プログラムの著作物については公表前に登録できる（著作権法 78 条の 2）．

すると，限定提供データの登録について，公表を要件とするか否かが問われるのであるが，プログラムの著作物と同様に公表を要件とすることなく，登録を認めるべきである．公表を要件とした場合，限定提供データは複製が容易なので，第三者が複製して，限定提供データに関する権利を侵害する事態が発生することが想定される．

また，限定提供データについて，譲渡，質権の設定など権利変動について登録できるように定め，取引の安全を図るのが所望される．具体的には下記のような条文が望まれる．

次に掲げる事項は，登録しなければ，第三者に対抗することができない．

一　限定提供データ権の移転（相続その他の一般承継によるものを除く．次号において同じ．）若しくは信託による変更又は処分の制限

二　限定提供データ権を目的とする質権の設定，移転，変更若しくは消滅（混同又は限定提供データ権若しくは担保する債権の消滅によるものを除く．）又は処分の制限

3.3 登録手続

不正競争防止法は経済産業省が所掌するので，「経済産業大臣が登録原簿に記載し，又は記録して行う」というような規定を設けるべきである．

また，限定提供データはデジタルデータとなるので，登録申請書にはデジタルデータの複製物を添付すべきである．

情報通信技術の進展に鑑みて，インターネットのような情報通信回線により登録の申請を行うのが適切である．工業所有権に関する手続等の特例に関する法律，行政手続等における情報通信の技術の利用に関する法律などが先例となる．

3.4 登録機関

登録機関は，限定提供データをコンピュータに登録するという情報処理業務を遂行することが求められる．登録業務は，大量かつ定型的な業務であり，一定の専門的技術，知識と相当の設備が求められる．従って，経済産業省のような政策官庁が登録機関となって，経済産業省の建物内部にホストコンピュータなどのハードウェアを設置するのは現実的でない．

そこで，工業所有権に関する手続等の特例に関する法律，プログラムの著作物に係る登録の特例に関する法律，著作権等管理事業法などと同様に登録機関を指定することができる旨を法定するのが適切である．

工業所有権に関する手続等の特例に関する法律では，第19条第2項の規定に基づくが登録情報処理機関として，一般財団法人工業所有権電子情報化センター（東京都千代田区五番町）が指定されている[特許庁　17]．

プログラムの著作物登録では，登録機関が指定することができる旨が定められており，これに伴って，一般財団法人ソフトウェア情報センターが指定されている[金井　92]．

3.5 方式審査

登録機関は，限定提供データの登録について方式審査を行うことが所望される．法令の規定に従った方式により申請されているかなど却下事由に該当しないかどうかを審査する．

一方，限定提供データの成立要件は法定されているが，相当蓄積性の要件，管理要件などの実体審査をするのは適切でない．

3.6 秘密保持義務

登録機関の情報処理業務について秘密保持を定め，違反したときには刑罰規定を設ける．例えば，下記のような規定である．

1　指定登録機関の役員若しくは職員又はこれらの職にあつた者は，登録事務に関して知り得た秘密を漏らしてはならない．

2　登録事務に従事する指定登録機関の役員又は職員は，刑法（明治四十年法律第四十五号）その他の罰則の適用については，法令により公務に従事する職員とみなす．

3.7 情報の信託

平成 30 年になってから，総務省が情報銀行の社会実装を提案している[総務省　18]．ここで，情報銀行とは，個人とのデータ活用に関する契約等に基づき，パーソナルデータ・サービスのようなシステムを活用して個人のデータを管理するとともに，個人の指示又は予め指定した条件に基づき個人に代わり妥当性を判断の上，データを他の事業者に提供する事業である．

情報銀行は，個人情報の信託という要素があるのだが，著作物の信託という要素がある著作権等管理事業法を参酌して，法制度を設計するのが所望される．著作権等管理事業法については，[清野 01]が詳しい．

4. まとめ

著作権法の登録などを参考にして，限定提供データの登録について提案した．

参考文献

[Ohsawa 15] Yukio Ohsawa, Hiroyuki Kido, Teruaki Hayashi, Chang Liu, Kazuhiro Komoda, Innovators marketplace on data jackets, for valuating, sharing, and synthesizing data. In *Knowledge-Based Information Systems in Practice*, pp. 83-97, Springer, Cham. 2015

[金井　92]　金井重彦:著作権の基礎知識 コンピュータプログラム，ぎょうせい，1992

[小池 19]　小池誠: データの法的保護―平成 30 年不正競争防止法改正の紹介―，信学技報，vol. 118, no. 453, AI2018-38, pp. 5-10, 2019

[清野　01]　清野正哉: 著作権等管理事業法，中央経済社，2001

[総務省 18]　総務省編: 平成 30 年版情報通信白書，日経印刷，2018

[特許庁　17]　特許庁編: 工業所有権法（産業財産権法）逐条解説，第 20 版，発明推進協会，2017

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

PR 効果測定のための複数組織間のデータ流通プロセス設計

Design of Data Flow Process among Multiple Organizations for PR Measurement

上島 邦彦[*1] Kunihiko Ueshima　登坂 泰斗[*2] Taito Tosaka　谷口 耕平[*3] Kohei Taniguchi　早矢仕 晃章[*4] Teruaki Hayashi　大澤幸生[*4] Yukio Ohsawa

[*1] 株式会社日本データ取引所 Japan Data Exchange, inc.

[*2] 株式会社オズマピーアール OZMA, inc.

[*3] キッコーマン株式会社 Kikkoman Corporation

[*4] 東京大学大学院工学系研究科 School of Engineering, The University of Tokyo

In this study, we designed and conducted the process of data flow to evaluate the effectiveness of PR and communication. Based on the process, we obtained 15 datasets from 15 different companies in about four months and created the integrated time-series data on trends of the product market, taking account of seasonality, trends, advertisement and PR activities, and unexpected events. The result suggests that our method can reduce the cost and failure risk in flowing data among multiple organizations and encourage the sustainable use of various data, reducing the workload of all the stakeholders as smooth as possible.

1. はじめに

データ流通産業の成長に伴って, 広告・PR 施策の投資対効果を検証する手法の確立が望まれている. しかし, 次の 2 つの意味で「越境的な」効果測定の手法は, 日本国内で十分に普及したとは言えない. 一つは, オンライン/オフラインの「施策」をまたいだ検証であり, もう一つは, 店頭販促, デジタル施策, TV CM などの「業域」をまたいだ検証である. というのも, 企業におけるデータ活用には次の 3 つのトリレンマがある.

- 経営指標に対する貢献の直接性, 明証性
- 事業活動における実用性, 持続性
- 分析手法としての先進性, 高度性

上記すべての満足を目指すデータ活用プロジェクトは, 実施コストや失敗リスクが肥大し, データの改竄や, 分析品質の低下, 解析システムの誤作動などが生じやすくなることが懸念される. 企業は, 限られたリソースによるデータ分析を, 簡潔に, 持続的に行う必要に迫られている. 広告・PR 業界でデータ活用に取り組む企業も, その例に漏れない. 以上を踏まえ, 私たちは以下の 3 点を目的とした研究を行った[OZMA 18].

(1)多様なデータに含まれる項目（変数群）に着目し, 企業の中長期的なマーケティング・コミュニケーション活動が生活者の消費行動に与える影響を測定する指標を策定する手法の開発

(2)上記の指標策定に当たって必要なデータの設計や分析手法を探索・選択するプロセスの開発とシナリオの生成

(3)参画企業内で上記の手法・指標の実用化に向けた実証実験

本稿では特に(2)に着目し, 複数の組織間で複数のデータを対象に所在調査, データ活用知識の表出・整理, 権利処理, 品質評価, データ設計, 分析手法の選定, 手法の適用, 結果の評価に至る一連のデータ活用プロセスを設計し, 適用した結果を報告し, そのプロセスについて議論する.

連絡先：上島邦彦, 株式会社日本データ取引所, 東京都渋谷区猿楽町 9-8 URBANPARK 代官山 I, 03-6455-1397, ueshima@j-dex. co. jp

2. 関連研究

これまでにも, オンラインとオフラインを統合した効果検証の要素技術はいくつか提案されてきた. 例えば, 店舗施策と DMP の連携, シングルソースパネルによるメディア接触と購買行動の横断分析, スマホアプリの行動ログ, 監視カメラ, ビーコンによる来店検知, EC サイト会員情報と実店舗 POS の ID 連携, 来店者の個人情報（顔, 指紋, 肉声等）による認証が挙げられる.

しかしいずれの技術でも, 得られるデータは個々の業域における特定用途に限られていた. 従って, 「施策」や「業域」をまたいだ知識を得たい利用者は, ①短期間に多大な技術投資を行うか, ②長期の実施計画に基づいた, 複数の調査プログラムを継続的に行うかを選択しなければならない. 一連の業務フローに沿った指標を策定するためには, 多種多様なデータを, なるべく低負荷で, 速やかに, 持続的に活用できることが望ましいとされる. こうした背景から, PR 戦略研究において, 関連業務フローの全体を考慮した, データ活用のライフサイクルが提案されている[Alex 17].

他方で, 組織を横断したデータ流通には様々な障壁があることが知られている. [METI 15]は, データの提供及び交渉における時間・労力・合意の不確実性を障壁として挙げている. 特に, データの利用権限に関する条件整理は, 「契約実務の蓄積が乏しいこと, あるいは当事者間の認識・理解のギャップがあること等により, 契約の締結が進まないという課題」がある [METI 18] .

結果として, 多くの要素技術が提案され, データ活用のライフサイクルを描いた概念モデルが作成されてはいても, 実際にデータ提供を行ったことのある企業は少ない. [METI 16]によれば, 「自社のデータを社外の企業・組織に提供・連携した経験」のある企業は 26. 5%（n=584, 日経 BigData 調べ）に留まっていた.

なかでも, 利用規約の異なる多種類のデータを, 業界・分野の異なる多種類の組織が, 互いに合意のうえで提供・利用する試みは, 実現していなかった.

3. データ流通プロセスの設計

そこで本研究では, CRISP-DM (Cross-Industry Standard Processing for Data Mining)[Shearer 00], Data Jacket [Ohsawa

13], Action Planning [Hayashi 14]のほか, データ活用を円滑化する 15 の手法・ツール群を用いたデータ流通プロセスを設計し, その適用を行って, 15 社の企業が保有するデータと知識を組み合わせた時系列解析を行った.

まず, CRISP-DM を用いて, データ分析の標準的な業務工程について, 研究の当事者間で理解を行った. 次に, データの所在調査, 利用可否の吟味, データ活用知識の表出・整理を行うため, Data Jacket 及びデータ利活用方法検討ワークショップ Innovators Marketplace on Data Jackets を導入し, データ分析のシナリオ原案を作成した.

また, これを踏まえて, データ・ブローカーの仲介と, 研究参加者の相互協力のもと, データ提供会社, データ提供会社の親会社・関係会社, 分析ツール提供会社, データの販売代理店, データ購入者など 15 社と権利処理や利害調整を行った. この工程を通じて, シナリオ原案が洗練されたほか, すべての企業から利用許諾を受けたデータを調達するに至った.

さらに, データ種類ごとの利用目的や開示範囲, 利用条件を考慮し, データ品質の評価を行ったうえで, 全種類のデータを統合したデータセットを設計した. 作成されたデータセットをもとに, 時系列解析の手法を適用した.

時系列解析の手法は, 実行しやすさ, 理解しやすさを優先して選定した. というのも, データ分析において,「変数によるデータの結合だけでなく、共通の目的や知識によってデータを繋ぐ」[Hayashi 18]には, 研究の当事者が反証可能な仮説を見つけ, 必要十分な品質のデータを確保し, 抽出, 加工, 分析の手法を適切に選定し, 得られた事実を妥当に解釈することが求められる。それに加えて, PR 効果測定のためには, 研究の非-当事者が得られた知識を組織内で理解させ, 利害関係者に望ましいふるまいを促すことが求められるためである.

4. 結果と考察

本研究では, 3 社が提出した 40 件の Data Jacket をもとに, 4 本のシナリオ原案が考案され, 1 本が採用された. その 1 本に必要とされた 23 種類のデータについて利用交渉に当たった. その過程で、データ分析のシナリオをさらに洗練させ、対象データを 15 種類に絞り込んだ(図 1). 結果として, データ分析に必要な各工程の所要時間を短縮でき, 15 社の企業・組織間での 15 種類のデータ流通を, 4 ヶ月程度で実現できた.

単年の事業計画のなかでデータ活用プロジェクトを行う場合, 最大で残り 8 ヶ月を企画の起案, 調査・分析の実施, 結果の評価, 施策の実行に費やせることになる. 広告・PR 関連業務において, 長期の実施計画に基づいた, 複数の調査プログラムを継続的に行いたいとき, 従来と比べてより多くのデータ分析プロジェクトを実施できる可能性がある.

なお, データの利用許諾に当たっては, データの権利帰属が明確であること, 営業秘密やプライバシーが保護されていることなど, 組織間の法的相互運用性の確保が課題となった. 今後, 個々の課題への対応が進むことで, データ流通の工程をさらに短縮することが期待される. また, データ種類ごとの権利関係や, 取引・交渉のための所要コストを計算できれば, データ流通プロセス全体の最適化に資する知見を得られる可能性がある.

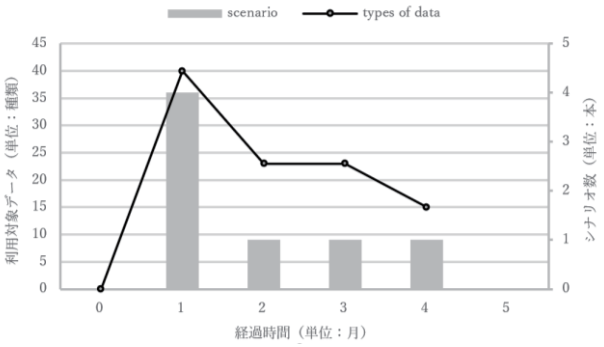


図 1 データ流通プロセスの適用によるデータ種類の低減

5. 参考文献

[OZMA 18]株式会社オズマピーアール, キッコーマン株式会社, 国立大学法人東京大学大学院工学系研究科, 株式会社日本データ取引所, “テレビ・WEB・SNS・店舗・食卓をまたいだ コミュニケーション効果検証の手法を開発する共同研究を開始”, 2018. https://ozma. co. jp/digitalcommunication/news-20181005/

[METI 15] 平成 26 年度経済産業省委託事業, “我が国経済社会の情報化・サービス化に係る基盤整備（データ駆動型イノベーション創出に関する調査事業）調査報告書”, 経済産業省ウェブサイト, 2015

[METI 18] 経済産業省, “AI・データの利用に関するガイドライン”, 経済産業省ウェブサイト, 2018

[METI 16]平成 28 年度経済産業省委託事業, “先端課題に対応したベンチャー事業化支援等事業 報告書「データ利活用促進事業：データ駆動型イノベーション創出支援のためのデータエクスチェンジプラットフォーム構築」”, 経済産業省ウェブサイト, 2016

[Alex 17] Alex Aiken, “CHAPTER 6 Measurement is not optional–not for the UK Government–not for your organization”, The PR Professional’s Definitive Guide to Measurement, AMEC, PRCA, PRCA MENA, ICCO, 2017

[Shearer 00] Shearer C. , “The CRISP-DM model: the new blueprint for data mining”, J Data Warehousing (2000); 5:13-22, 2000

[Ohsawa 13] Yukio Ohsawa, Hiroyuki Kido, Teruaki Hayashi, Chang Liu, “Data Jackets for Synthesizing Values in the Market of Data,” 17th International Conference on Knowledge-Based and Intelligent Information and Engineering System (KES2013), pp.709-716, Procedia Computer Science, Vol.22, pp.709-716, 2013.

[Hayashi 14]早矢仕晃章, 大澤幸生, “協同問題解決のプランニングにおける発言とシナリオ評価への影響 (ヒューマンコミュニケーション基礎)”, 電子情報通信学会技術研究報告 = IEICE technical report : 信学技報 114(273), 37-42, 一般社団法人電子情報通信学会, 2014

[Hayashi 18]早矢仕晃章, 大澤幸生, “データ3.0時代のデータランドスケープ”, 2018 年度人工知能学会全国大会（第 32 回）大会論文集, 2018

4:00 PM - 4:20 PM (Wed. Jun 5, 2019 3:20 PM - 4:40 PM Room G)

[2G4-OS-2b-03] (Invited talk) Expectation for data trading

〇Aki Ando[1] （1. ABeam Consulting Ltd.）

Keywords: Data Trading

データ流通に関する期待の高まりで、技術面・制度面での検討が進むが、本来はデータ流通の具体的なユースケースを融合していく必要がある。ビジネス側からの目線でデータ流通への期待を語る。

4:20 PM - 4:40 PM (Wed. Jun 5, 2019 3:20 PM - 4:40 PM Room G)

[2G05-07-4] Panel discussion

Panel discussion

©The Japanese Society for Artificial Intelligence

Organized Session | Organized Session | [OS] OS-15

[2I3-OS-15a] 人と AIが織りなす新たなエコシステム(1)

中川 裕志（理化学研究所）、山川 宏（ドワンゴ人工知能研究所／全脳アーキテクチャ・イニシアティブ）、浅田 稔（大阪大学）、井上 智洋（駒沢大学）、江間 有沙（東京大学）、金井 良太（アラヤ）、高橋 恒一（理化学研究所）、萩田 紀博（ATR）、堀川 優紀子（ATR）、松原 繁夫（京都大学）

Wed. Jun 5, 2019 1:20 PM - 3:00 PM Room I (306+307 Small meeting rooms)

[2I3-OS-15a-01] Future Prototyping Methodology for Knowledge Co-creation

〇MIWA NISHINAKA[1], Hideaki Takeda[2], Kunio Shirahada[3], Yusuke Kishita[4], Hisashi Masuda[5] （1. The Graduate University for Advanced Studies, 2. National Institute of Informatics, 3. Japan Advanced Institute of Science and Technology, 4. The University of Tokyo, 5. Kyoyto University）

1:20 PM - 1:40 PM

[2I3-OS-15a-02] Collective responsibility under AI network

〇shigeo KAWASHIMA[1] （1. Aoyama Gakuin Women's Junior College, RIKEN）

1:40 PM - 2:00 PM

[2I3-OS-15a-03] Analysis of the AI and Society Meeting in Japan

〇Hitomi Sano Sano[1] （1. AI and Society lab in Keio university）

2:00 PM - 2:20 PM

[2I3-OS-15a-04] Gaming Space as an Online Society, and A Consideration about Its AI Implementation for Rapid Transformation Tolerance

〇Shinya Umeno[1], Masaya Nakamoto[1], Tetsuya Nohama[1] （1. Mynet.ai Inc.）

2:20 PM - 2:40 PM

[2I01-04-5] Discussion / Conclusion

2:40 PM - 3:00 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

未来を共創するためのフューチャー・プロトタイピング手法の提案

Future Prototyping Methodology for Knowledge Co-creation

西中 美和[*1]
Miwa Nishinaka

武田 英明[*2]
Hideaki Takeda

白肌 邦生[*3]
Kunio Shirahada

[*1] 総合研究大学院大学
The Graduate University for Advanced Studies

[*2] 国立情報学研究所
National Institute of Informatics

[*3] 北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology

木下 裕介[*4]
Yusuke Kishita

増田 央[*5]
Hisashi Masuda

[*4] 東京大学
The University of Tokyo

[*5] 京都大学
Kyoto University

Abstract: This position paper proposes “Future Prototyping Methodology” for a discussion platform assuming a citizen workshop, to simulate future situation. Prototyping by this methodology, we also propose a model to clarify thinking process for the future when people seeking well-being, which might be one of triggers to cause social change. We focus on the well-being as an inevitable purpose of the discussion because we think that happiness is emerged when we work for social well-being in the succeeding generations as well as seeking current personal happiness. Our research will contribute to the area of Artificial Intelligence by presenting theses of thinking process for social well-being for the future. As a practical contribution, the proposed methodology is positioned as a tool to show the future to achieve well-being based on people’s needs for the future. This paper presents research overview, prior research, a research question for our methodology and its uniqueness.

1. はじめに

本稿は，未来共創を促進する「フューチャー・プロトタイピング手法」を提案するポジション・ペーパーである．この方法論は，市民ワークショップでの使用を想定し，参加者が未来コンテキスト（状況）をディスカッションで共創するためのシミュレーション・プラットフォームである．また，参加者が未来コンテキストにおける well-being を議論する中で，考えがどのように変遷していくかの過程を明らかにする．この研究は，人工知能研究にとっては，倫理問題に対する 1 つのテーゼの提示として貢献する．実務的な貢献としては，幸福を達成するため，将来への人々のニーズに基づいて未来を示すためのツールとして位置付けられる．

方法論とモデル構築にあたって，私たちは未来共創における well-being と考えの変化を重視する．well-being を重視する理由は，私たちは「より良く生きる」ということは、現世代だけではなく，将来世代にとっても，「よく生きる」ことであり，幸せな状態ではないかと想定するためである．考えの変化を重視する理由は，未来を考える際，人は，現生的な考え方から，自己を超越した将来的な考え方に，ある時点で変わる場合があり，考えの変化によって，社会も変わるのではないかと思っているためである．その思考過程を明確化する．考えの変化を起こすものが，この方法論であり，方法論の特徴としての創作ストーリーである．今後の研究としては，考えを変えることによる社会変革の過程を対象として想定している．

Well-being に関しては，様々な定義がある．本稿においては，将来世代のための「社会的な well-being (social well-being) 」と現代の幸せである「個人的な well-being (personal well-being) 」の双方を well-being とする．また，広義における「幸せ」と捉え，身体的，精神的，社会的幸せをすべて含んだ概念であるとする．

連絡先：西中美和，総合研究大学院大学，神奈川県三浦郡葉山町上山口 1560-35, nishinaka_miwa@soken.ac.jp

未来コンテキストにおける well-being を考える場合の問題点としては以下が考えられる．

- Personal well-being 間における利害の対立がある
- Personal well-being と social well-being は，必ずしも一致しない．個人の利益の追求は，公共の利益にはならないことは多い．その逆もありうる．
- social well-being 間における選択の難しさがある．例えば，地方の共同体で開発の話が出た場合，守ってきた伝統と景観を子孫に残すか，便宜性を優先させ，市街化することで現時点での人口減少を食い止めるかといった議論が 1 つの典型である．伝統や自然保護，人口減少対応の双方とも social well-being を目指すものであり，どちらがより重要であるかは，一概には判断できない．

これらを踏まえた上で，人がどのように考えを変えてゆくかの過程と要因を提示する．以上より，以下のリサーチ・クエスチョン（RQ）を設定する．

RQ:人は将来の社会的 well-being (social well-being) のために，どのように考えを変えるのか，それはなぜか．

2. 研究の概要

2.1 方法論の提案

図 1 に方法論の概要を示す．この方法論でプロトタイプを行い，方法論そのものとモデルを完成させてゆく．また，以下に手順を示す．手順の番号は図における番号に相当する．

(1) INPUT が準備される

1. 創作ストーリーのアウトライン
2. 未来要素を描いたデジタル・カード（技術・自然・経済・社会要因等，設定する時間軸，アバターとしての登場人物：これらによりバックキャスティングを行い，OUTPUT としての未来コンテキストを構築する）

3. 実データ（RESAS (Regional Economy Society Analyzing System in Japan) や白書等のデータ）
(2) PROCESS として，実験的ワークショップが実施される
1. 参加者は，準備された INPUT を用いて議論を行いながら，未来をプロトタイプする
2. ワークショップにおいて，分析用データが収集される．これらは方法論とモデルの構築に使用される．データとしては，発話データ，創作ストーリーのテキスト，視線等の生体データであり，これらから参加者の考えを分析する．
(3) OUTPUT が完成する
1. 未来コンテキストのデジタル・イメージ
2. 創作ストーリー
3. 実データによって描かれたグラフ
(4) データを分析しワークショップを評価する．これらにより未来コンテキストにおける知識共創の過程を明確化する
1. INPUT と OUTPUT の相関関係を検証する
2. 参加者の参加度合を検証する
3. 参加者の主観的考えを調べる
(5) このサイクルが一巡するとプロトタイプが完成する
(6) このプロトタイプを使用し，再度サイクルを回し，試行錯誤の上方法論をブラッシュアップする
(7) 最終的に方法論とモデルが完成する．

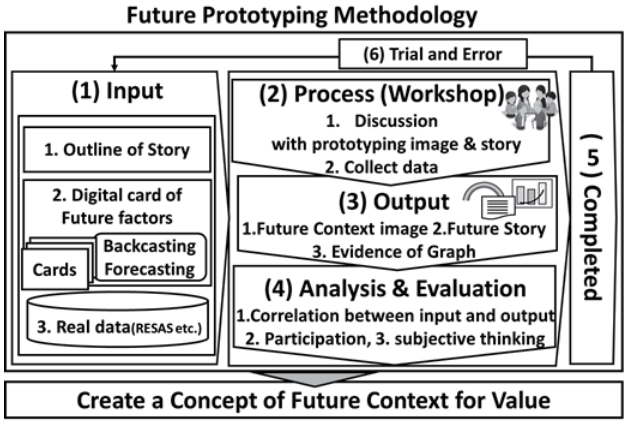
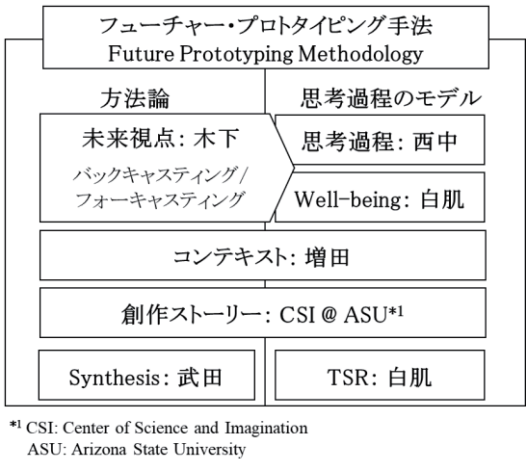


図 1　フューチャー・プロトタイピング手法の概要

3.　学術的背景および先行研究

3.1　背景となる研究の全体図

図 2 に，本研究における背景となる研究の全体を示す．



*1 CSI: Center of Science and Imagination
ASU: Arizona State University

図 2　背景となる研究の全体図

以下，関連する先行研究において，これまでの研究の紹介を行ったうえでまだ十分ではない箇所を明らかにし，我々の研究の新規性を示す．

3.2　思考過程

これまで，西中と白肌は，ワークショップにおける共創コミュニケーションにおいて，未来志向と革新的な考えを促進する概念モデルを研究してきた[Nishinaka 2017]．これらの結果から革新的な考えが生まれる過程は，(1) 2 人のリーダータイプ・メンバー間の相互作用が重要である，(2) やりとりによって具体的なアイデアが抽象化・概念化されてゆく過程が革新性と関係する，(3) 良い結果を生み出すには，メンバー全員の参加が重要である等が明らかになってきた．しかしながら，未来志向的な考えがどのように生まれ，それがどのようにメンタル・イメージとなってゆくかの詳細な思考過程や要因は十分に明らかになってはいない．そのため，その部分を，木下のバックキャスティングの手法を用いて取り入れたうえで，新しい方法論とモデルを構築する．

3.3　バックキャスティングとフューチャー・デザイン

木下は，持続可能な社会に向けたシナリオの設計方法論について研究を行っている[Kishita 2016]．そこでは，望ましい将来像 (ビジョン) を初めに描き，そのあとでビジョンを達成するための道筋を時間的に逆方向に描く技法である「バックキャスティング」を用いたシナリオ作成方法を開発している．事例として，自治体と連携しながら持続可能なエネルギービジョンを描いたバックキャスティングシナリオを作成している[木下 2018]．しかしながら，2050～2060 年のような遠い将来をバックキャスティングで描く際には，シナリオ作成者自身が，現在とは異なった将来の視点を持つことができるかどうかが大きな課題となる．

他方，西條らはフューチャー・デザインというコンセプトを提案し，仮想的な将来世代をつくることによって，将来世代と現世代が持ちうるニーズの違いを明らかにしようと試みている[Saijo 2015]．これまでに，市民ワークショップを通して 2060 年の自治体を対象とした将来を描いたところ，将来世代は現在の状況にとらわれることなく，複雑で解決に時間のかかる課題への優先度を高める傾向があるとの結果を導いている[Hara 2015]．

今後，バックキャスティングとフューチャー・デザインを適切に組み合わせることは，社会的 well-being が高い将来の描写に向けて有望なアプローチである．

3.4　コンテキスト

増田は，コンテキストを考慮したビジネスやそのコミュニケーションのモデル化に関して，ビジネスプロセスモデリングにおけるアーリー・バインディングとレイト・バインディングの適用提案[Hara 2016]や、小林らと共に，価値共創におけるコミュニケーション・パターンの類型化に関する研究を行ってきた[小林 2014]．

本提案手法では，コンテキストの概念定義として主に以下の分野の概念を統合的に用いるが，特に，未来という不確実性の高い状況，かつ，集団と個人でのコンテキストの差異といった観点の表現や取り扱いに関しては，既存概念の統合や拡張といった課題解決が必要になる．サービス・マーケティングにおけるサービス・ドミナント・ロジック[ラッシュ 2016]では，"価値は他のサービス・プロバイダーから提供される資源とともに，ある特定の文脈でオファリングを使用することから生み出される"として，コンテキスト(文脈)を価値の形成に影響を与える要因として捉えている．また、人類学の視点としては，コンテキストは意味を伝達するためのコミュニケーションの機能のひとつである[Hall 1976]とし，意味は情報とコンテキストから作られ，低コンテキストコミュニケー

ションでは，意味における情報の割合がコンテキストより大きく，高コンテキストコミュケーションでは，それが逆転するという視点から国際的な文化差を説明する．Context Aware Computing では，コンピュータが取り扱うコンテキストとして，"any information that can be used to characterize the situation of an entity" として，コンテキストはエンティティの状況説明に用いられる情報であるとする[Dey 2001]．本研究における未来コンテキストは，コンテキスト概念における新しい概念の提示とその応用活用手法の提案として位置づけられる．

3.5 創作ストーリーと Science Fiction Prototyping (SF Prototyping)

本研究で提案する方法論の特徴は，ワークショップ参加者が将来を描きやすくするために創作ストーリー[Halpern 2018]を使うことである．創作ストーリーは，参加者の暗黙的なメンタルモデルを表出化し，望ましい未来に対するアイデアを生み出すことに役立つと言われている [Halpern 2018]．創作ストーリーを作成することは省察を伴い[Johnson 2011][Kohno 2011]，心の中の暗黙的な考えに自分自身で気づくための助けになる．この特徴のため，ワークショップのアウトプットとしては，完成した創作ストーリー，および将来像の絵が出来上がるが，これらは参加者のメタ知識を表出化したものと考えられる．ここでのメタ知識とは，人の考えに意味付けを行い，高位の抽象化が行われた上で，人から独立した知識のことを指す．

創作ストーリーの作成手順は，SF Prototyping を参考にしている[Johnson 2011]．SF Prototyping は，SF あるいは創作を利用して，科学的発見の端緒やあるべき科学の姿を提示する方法である (https://csi.asu.edu/) [Johnson 2011]．手順としては，まずアウトラインを提示し，それに沿った形で全体の話を作ってゆくという形をとる．これには，SF によって，あるべき科学の姿をプロトタイピングするという意味合いがある．

本研究が提案する手法では，さらに参加者にストーリーを考えるための材料として白書等の実データを提示する．その上で，ワークショップの中で参加者が自分たちの思い描く未来の姿をストーリーとして完成させてゆくという手順をとる．

本研究が提案する方法論の主とするところは，考えの変化をモデル化する点であり，またその実際の手法を構築する点にある．そのために創作ストーリーを利用するが，創作ストーリー自体を研究するわけではない．また，実データを INPUT として利用はするが，参加者の発想による創作を重視する．

長谷，大澤らは，人工知能に小説を書かせる試みにおいて SF を社会にいかに活用できるかを検討しているが[長谷 2016]，私たちの研究は，人が考えたストーリーの中から，人工知能へのテーゼを提示するというものである．

3.6 Transformative Service Research

ウェルビーイング志向のサービス研究として Transformative Service Research (TSR)が世界的に推進されている[白肌 2018]．価値共創プロセスとしてのサービスの成果を，人間の幸福や社会の公平さなどと関連付けて考察することにより，より良い社会を形成しようとするものだ．ここには価値共創主体が，自らの置かれている文化的，経済的，制度的コンテクストを見直す中で，支配的になっている価値観を問い直し，本当に自分自身やコミュニティ，社会にとって価値のある経験は何かを考え，サービス設計し実践していくことを含んでいる．未来構想を射程にする本研究においては，創作ストーリーや未来コンテキスト分析の基盤的思考枠組みとして機能すると同時に，未来に向けて残したい価値の源泉としての資源を，どのように多主体価値共創システムの中で維持するかという設計枠組みの基盤としても機能する．

3.7 Synthesis

アブダクション（Abduction）は，科学的発見の方法として Peirce が提唱したもので[Peirce 1878]，演繹とは逆方向の推論として位置づけられ，仮設生成を行うものである．科学において仮説の検証の過程は演繹による推論として説明できるが，発見を行う過程は演繹では説明できず，このため Peirce はアブダクションという別の推論を定義した。この演繹とアブダクションの関係は，科学プロセスの中で，分析的なプロセスであるアナリシスと，仮説生成的プロセスであるシンセシスに対応する．武田ら[Takeda 1994]はこの関係を人工物の分析と設計に当てはめ，設計プロセスの定式化を行った．フィーチャー・プロタイピングでは，未来コンテキストが仮説に当たる．

アブダクションは論理的には演繹の逆プロセスと定義することができるが，探索の方法，知識の範囲，解の選択方法に依存する．フィーチャー・プロタイピングでは，そもそも利用する知識を生成しなければいけないことがある．これは知識自身を仮説生成しなければいけないことになり，高階のアブダクションとなる．これは推論として難しい問題であり，現実的な方策で解を見つける必要がある．

4. おわりに

このポジション・ペーパーは，「フューチャー・プロトタイピング手法」 およびこの方法論を使用することで構築されるモデルを提案し，学術的背景と特徴を説明した．この方法論は，未来コンテキストにおける参加者の social well-being に対する考えを促進し，知識共創をサポートする方法論である．さらに，今後の研究としては，考えを変えることによる社会変革の過程を対象として検討をしており，本研究においても，その点を見越して，研究を進める予定である．社会変革とは，最適性・最善性・妥当性ループを社会的オープンシステムで自律的に回すことで価値を創出することであり，社会における未来共創のエコシステム構築のことである．

5. 謝辞

本研究は JSPS 科研費課題番号 JP16K03859，情報・システム研究機構　未来投資型プロジェクト，国立情報学研究所　平成 30 年度共同研究（会合型）の助成を受けたものです．

参考文献

[Dey 2001] Dey, A.K.:Understanding and Using Context, Personal Ubiquitous Computing, 5 (1), pp.4-7 (2001)

[長谷 2016] 長谷敏司，藤崎慎吾，山川宏，宮本道人，大澤博隆: SF の想像力を技術者・社会はどのように活用できるか―はるこん 2018 人工知能× SF セッションレポート―，人工知能学会誌 33(5), 679-690 (2018)

[Halpern 2018] Halpern, M., Eschrich, J., & Sadowski, J. 2018. *The Rightful Place of Science: Frankenstein,* Consortium for Science, Policy & Outcomes, Tempe, AZ and Washington, DC (2018)

[Hara 2015] Hara, K., Yoshioka, T., Kuroda, M., Kurimoto, S., Saijo, T. Participatory deliberation for future design by creating imaginary future generations - Evidence from an experimental workshop in Yahaba Town, Iwate, Japan,

Proceedings of EcoDesign 2015 International Symposium, 72-74 (2015)

[Hara 2016] Hara,Y., Masuda,H.: *Global Service Enhancement for Japanese Creative Services Based on the Early/Late Binding Concepts, Domain-Specific Conceptual Modeling: Concepts, Methods and Tools*, Springer International Publishing, pp. 509-526 (2016)

[Hall 1976] Hall, E.T.:Beyond culture, Anchor Press, 2001.

[Johnson 2011] Johnson, B.D.: *Science Fiction Prototyping: Designing the Future with Science Fiction*, Morgan and Claypool, https://doi.org/10.2200/S00336ED1V01Y201102CSL003 (2011)

[Kishita 2016] Kishita, Y., Hara, K., Uwasu, M., Umeda, Y. Research Needs and Challenges Faced in Supporting Scenario Design in Sustainability Science: A Literature Review, *Sustainability Science*, 11(2). 331-347 (2016)

[木下 2018] 木下裕介, 本元俊行, 上須道徳, 野間口大, 原圭史郎: 自治体スケールのエネルギービジョンを対象としたフューチャー・デザイン, 第 37 回エネルギー・資源学会研究発表会講演論文集, pp. 321-326 (2018)

[小林 2014] 小林潔司, 原良憲, 山内裕(編):日本型クリエイティブ・サービスの時代「おもてなし」への科学的接近, 日本評論社 (2014)

[Kohno 2011] Kohno, T. and Johnson, B. D.: Science Fiction Prototyping and Security Education: Cultivating Contextual and Societal Thinking in Computer Security Education and Beyond, *SIGCSE '11 Proceedings of the 42nd ACM technical symposium on Computer science education*, 9-14. DOI: 10.1145/1953163.1953173 (2011)

[ラッシュ 2016] ラッシュ, R. F., バーゴ, S.L.; 井上崇通(監訳者), 庄司真人(訳者), 田口尚史(訳者):サービス・ドミナント・ロジックの発想と応用, 同文館出版 (2016)

[Nishinaka 2017] Nishinaka, M., Shirahada, K., and Kohda, Y.: Visualization of the influence by conceptual leadership promoting high quality output, *Proceedings of 2017 International Conference on Industrial Engineering and Engineering Management (IEEE IEEM2017)*, DOI: 10.1109/IEEM.2017.8289866 (2017)

[Pierce 1878] Pierce, C. S.: Deduction, Induction, and Hypothesis, *Popular Science Monthly*, 13. 470-482 (1878)

[Ryan 2008] Ryan, R.M., Huta, V., and Deci, E. L.: Living well: A self-determination theory perspective on Eudaimonia, *Journal of Happiness Studies*, 9(1), 139-170 (2018)

[Saijo 2015] Saijo, T. Future Design: Concept for a Ministry of the Future, *Kochi University of Technology, Social Design Engineering Series*, 2015-14 (2015)

[白肌 2018] 白肌・ホー: ウェルビーイング志向の価値共創とその分析視点, サービソロジー論文誌, 4(4), pp.1-9 (2018)

[Takeda 2001] Takeda, H.,Yoshioka, M. & Toiyama, T.: Roles and formalization of abduction in synthesis, *The 15th Annual Conference of Japanese Society for Artificial Intelligence*, 1-4 (2001)

[Yonemori 2007] Yonemori, Y.: *Abduction: Theory of hypothesis and discovery*, Tokyp: Keisoshobo (2007)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

AI ネットワーク環境化における社会的次元の倫理的責任の検討

Collective Responsibility under AI Network

河島茂生[*1 *2]
Shigeo Kawashima

[*1] 青山学院女子短期大学現代教養学科
Department of Liberal Arts, Aoyama Gakuin Women's Junior College

[*2] 理化学研究所革新知能統合研究センター
Center for Advanced Intelligence Project, RIKEN

Abstract: The aim of this report is to consider how to position the attribution of ethical responsibility in situations where cooperation exists among multiple AIs and between AIs and other systems. Based on the theory of neo-cybernetics, we support the idea of collective responsibility under the AI network environment. In view of recent trends in computer technology, we assume that responsibility cannot be attributed to specific people or organizations. In such a case, establishing collective responsibility is required to compensate the victims and not give rise to a situation wherein people do not favor developing or using AI. However, under the AI network situation, there is lack of collective responsibility. If the defects of a specific person or organization are clear, responsibility should be attributed to the concerned person or organization.

1. 問題の所在

1.1 社会的背景

本報告の目的は，複数の AI が通信ネットワークにおいて接続されていく環境下の集合的責任について検討することである。AI が相互作用してデータを交換し群として機能することで多様なサービスが実現される。しかし，AI ネットワーク化により困った事態が起きることも想定される。

ネットワーク通信を介した事件ではデジタル・フォレンジックの限界がすでに露呈しており，誰が行った殺害予告・サイバーテロなのかを特定できないケースが相次いでいる。インターネット広告では，広告主や広告代理店，配信事業者でも，どのようなウェブサイトに広告が表示されているかを正確に把握できず，思惑から外れて猥雑なウェブサイトや政治的にきわめて偏ったウェブサイトに広告が掲載され，またアドフラウド(広告詐欺)にもあっている。AI がネットワーク化して群として機能するようになると，責任の所在の検証が難しくなることは，それほど頻繁に発生するとは考えにくいとはいえ，容易に想定される。

本論文は，こうした事態を想定し，ネオ・サイバネティクスの理論に依拠しながら AI ネットワーク下における集合的責任(collective responsibility)ともいうべき考え方を考察することである。EU 議会における電子人間の提言は，補償基金設立と合わせて提案されているが，これらは別個の論点である。というのも，電子人間なるものを確立しなくとも過失が不分明な場合には集合的責任を導入すれば機能するように想定されるからである。

1.2 関連研究および研究の目的

Perrow は，定常事故(normal accident)という語を生み出し，被害が甚大だったスリーマイル島原発事故やボパール化学工場事故などを取り上げながら，複雑かつ大規模な科学技術に依存している現代社会は事故が避けられないと指摘する[Perrow 84]。けれども Perrow の研究は，複雑かつ大規模な科学技術の倫理的責任にまで及んでいない。また Schomberg は，不確実性が増す技術が社会に組み込まれる状況を作り出しているのは，個人の意図というよりも集合的行為であるため集合的責任を考えなければならないとした[Schomberg 09]。とはいえ，Schomberg の論文は AI もしくは AI ネットワークについては射程外である。

AI と責任の問題については数多くの研究がある。なかでも赤坂は，不法行為法との関連から AI に法的人格を与えた場合の損害補填機能や抑止機能，制裁機能を検討しており，一定の妥当性は見出せるものの，制裁機能や AI 自体の故意・過失の認定に疑問があるとしている[赤坂 18]。必ずしも AI に法的人格を付与する必要性はなく，本報告でも集合的責任の一形態として上げる無過失補償制度について支持している。とはいえ赤坂の研究は，あくまで法学的観点から論じたものであり，本報告とは理路が異なる。本報告は，ネオ・サイバネティクスの理論に依拠することにより基礎的なところから AI ネットワーク化の集合的責任を捉える。

2. 社会システムと倫理的責任

ネオ・サイバネティクスの理論的支柱であるオートポイエーシス理論は，生物と機械との区分を明確にした[Maturana 80]。生物の特徴は，自分で自分(auto)を作る(poiesis)ことであり，そうした働きを内部で行う単位体をオートポイエティック・システムという。オートポイエティック・システムは，生物の必要十分条件である。人間は，37 兆個もの細胞を作り続けているオートポエティック・システムの集合体であり，それぞれの集合体(つまり，個人)は唯一無二の存在になっていく。オートポイエティック・システムの反対概念であるアロポイエティック・システムは，それ自体によって作られるのではなく，別のもの(人間)によって作られ，別のもの(allo)を生み出す(poiesis)。エアコンや自動車と同じく AI も，AI が AI を作っているわけではなく，アロポイエティック・システムである。

オートポイエーシス理論によれば社会は，社会システムと呼ばれ，細胞や人間の心理(心的システム)とは違った別種のオートポイエティック・システムである。つまり，社会は「生物的」であり，コミュニケーションが後続のコミュニケーションを喚起しながら存立する自己産出を特徴としている。個人の心的システムには還元できない社会システム内での自律性が生じる。コミュニケーションは不確実性を帯びており，いかなるコミュニケーションが連鎖していくかを統制できない場合も多い。質問の意図で発した言葉が命令や嫌味に受け取られるなど，自分の思いとは離れてまったく別様に解釈されることが起きる。あるいは身に覚えがな

連絡先：河島茂生，〒150-8366 東京都渋谷区渋谷 4-4-25，kawashima.shigeo@gmail.com

いことでも噂が広まってしまい収束できない状況も起きる。とりわけ社会システムの規模が大きくなると，コミュニケーションは非人称的な側面が強くなっていく。

社会のなかで責任が生じる。けれども，ある問題を伴う出来事があったとき，誰の責任であるかは必ずしも自明ではない。それゆえ，責任の度合いとともに，いかなる組織や人に責任を被せるかの議論がしばしば起きる。ただし社会システムの責任は，特定の人や組織に帰属できる場面ばかりではない。コミュニケーションの連鎖が急激に進み，誰も止められないことが起き，特定の誰かに帰責することが難しい場面が生ずる。

3. AI ネットワーク化における集合的責任

近代以降，人間の心的システムにおける自由意志に倫理の基盤を置き，そのことから瑕疵があったときに個人に責任を課し，社会の安定を図ってきた。被害を生んだのは個人の行為であり，その人を非難することで被害者の苦しみを解消することが求められた。

こうした事態は，そう簡単には変わらないと想定される。AI が普及した社会が無責任社会になってはならず，開発者・利用者の故意の過失もしくは怠慢，責任感の減退を防ぐためには，また技術を改善する動機の維持のためには，明確なるミスについてその倫理的責任を追及し続けざるをえない。特定の人や組織の過失まで，後述する集団的責任として補償することは人々が納得しないと考えられる。許し難い過失や故意があった場合でも，それを追及できないとするならば被害者は不満を募らし鎮静化には至らない。開発者やその AI によってサービスを提供するものが責任を負い，役割にあった行為を遂行すると予期できるからこそ，人々は AI ネットワークを信頼して受容することができる。

重要な課題は，前記したように，複数の AI が連携して動くなかで個人の悪意や過失が同定できない場合に起こる。あるいは，どの個人が行った行為なのか特定できないケースに起きる。

AI は，専門性の高い分野であり，その内部機構はただでさえ複雑であり，今後もその複雑性は増すと考えられる。Microsoft の AI「Tay」の騒動に端的に表れているように機械学習を使っているため，データが変動すると出力も変わる。個々の AI は対策を施してサービスを実施するだろうが，そうした AI 同士がネットワーク化し連動するなかで，予期しない動きが生じ未解明の事象が出てくることは避けられない。協業／分業に伴うリスクもある。a という AI システムの開発者は，ほかの複数の会社が開発・運営している b, c, d, ･･･といった AI のデータを取り込み，独自のメカニズムで処理して結果を出すようにプログラムを実装したとしよう。このとき a の開発者は，b, c, d, ･･･が正しいデータで妥当なデータ処理をしていることをあてにできるからこそ，自分の作業に専念することができる。a の開発者は，b, c, d, ･･･の中身やデータの適切性について検証する時間もなければ権限もない。したがって，たとえ擬制であっても b, c, d, ･･･が適正な動きをしていると信じなければならない。もちろん，b, c, d, ･･･の AI の不具合が噂になったり，データ・フォーマットが崩れたりしていると a の開発者は気づく機会を与えられる。しかし厳密なる検証はし難い。それゆえ誤ったデータが一度流通すると，それが連鎖していく恐れがある。

なにも過失がない場合にも，エンジニアや運営者が責任追及されるとすれば，それは AI 開発および利用の萎縮につながり，社会的な損失ともなる。逆に過失が同定できないからといって，被害者が救済されない事態を招くことも望まれない。

したがって悪事を働いたものが特定できない場合にせよ，十分予見されえず非意図的な場合にせよ，過失が同定できないならば，AI ネットワークが組み込まれた社会システムそれ自体が一種の道義的責任を担い，損害を被った人に補償していく制度の構築が望まれる。

社会システムのコミュニケーションは，そもそも不確実性があり非人称的な面を抱えている。そのような特徴をもつ社会システムに組み込まれて AI ネットワークは動作する。個々の AI はアロポイエティック・システムであり，入力と出力の対応関係が定められた範囲で収まるように調整されている。けれども AI がネットワーク化し不確実性のある社会システムに組み込まれて群として動いたときに開発者・利用者の予期せぬ動きが生じ，コントロール不能に陥ることが考えられる。集合的責任の制度構築により，開発者・利用者の負担を過度に増加させなくとも被害者を救済する途が開かれる。開発活動・利活用の保護および被害者の保護を考慮すると，こうした制度は，過分な責任追及を行う度合いを相対的に低くし被害者救済も図れるため，バランスの取れた帰結をもたらすように推察される。

集合的責任には，身体への傷害に対する金銭的な補償に加え，名誉の回復も含まれる。インターネット上ではフェイクニュースが氾濫している。AI は，フェイクニュースの検知や伝播過程の分析にも使われるが，フェイクニュースの作成や流布にも使われる。匿名化技術によって作成者や送信元が特定できないように加工され，プラットフォーム企業のサービスに流される。そのデータが人生を左右しかねない意思決定を支援する AI に取り込まれ，そのほかの AI にも伝播していく。たとえ不快な嘘，不正確で誤解をもたらす情報が出回り名誉が傷つけられて人生が狂わされても，怒りを向ける相手が分からず対処できない。謝罪広告も出してもらえない。そうしたときに，公益財団法人等の運営組織を共同で立ちあげてフェイクニュースが事実誤認であることを示し，合わせて AI のデータの書き換え要請を行うことで被害者を救済する方法もありうる方策である。

4. 結語

本論文は，ネオ・サイバネティクスという学術的礎をもとに，AI ネットワーク環境下の集合的責任ともいうべき考え方を検討してきた。

本論文は，個人的責任と集団的責任の両立を支持する。近代以降の社会は，個人の内にある心的システムの自由意志を特権的に位置づけ，そこに倫理的責任を帰属させてきた。けれども個人的責任だけでは限界がある。個人の内面の倫理観を高め技能を向上させるだけではなく，不確実性が内在化している社会システムの責任として引き受けることが求められる。

今後，具体的な制度に関する考察が求められる。

謝辞

本報告は、科学研究費補助金若手研究(B)「人工知能・ロボット・サイボーグの倫理的問題に関する理論的かつ実証的研究」(平成 29 年度−平成 31 年度、代表: 河島茂生、研究課題番号: 17K12800) の助成を受けた研究に基づいたものである。

参考文献

[赤坂 18] 赤坂亮太: 不法行為法における AI の法的人格に関する検討, 2018 年度人工知能学会全国大会発表資料, 2018.

[Maturana 80] Maturana, H. R. and Varela, F.: Autopoiesis and Cognition, D. Reidel Publishing Company, 1980.

[Perrow 84] Perrow, C. Normal Accidents, Basic Books, New York, 1984.

[Schomberg 09] Schomberg, R. von. Organising Collective Responsibility, Keynote lecture at the first annual meeting of the Society for the Study of Nanoscience and Emerging Technologies, Seattle, 11 September. 2009.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

AI 社会論研究会の分析

Analysis of the AI and Society Meeting in Japan

佐野仁美[*1]
Hitomi Sano

[*1] 慶應義塾大学
Keio University

The AI and Society meeting in Japan is an organisation which meets to discuss 'the influence of Artificial Intelligence on society' from various viewpoints such as Humanity, Philosophy, Economics, Law, Politics and Sociology. These are collectively referred as 'HELPS'. This article classifies the themes have been discussed at this meetings using 'HELPS'. Through analyzing how various themes have been discussed and how various stakeholders have been involved, we consider the effectiveness of the meetings and how the multi-stakeholder process works.

1. はじめに

本稿では、2015 年 2 月より開催されている AI 社会論研究会の社会的な役割について検討を試みる。人口減少が続く我が国で、AI 技術者が不足しつつも、第 4 次産業革命の核になる技術として期待される AI を社会に実装しながら、様々な社会課題を解決すること生産性を向上させることなどが益々期待されている。AI が社会に浸透し、大規模に人々と AI が混在する社会が形成されるにつれて予期し難い影響が生じるのであれば、あらかじめ多くのステークホルダーとの対話の中でリスクや不確実性を回避することが求められる。人と AI が織りなす新たなエコシステムを検討するには、限られた分野に閉じた議論ではなく、分野共創、マルチステークホルダー、調和 などのキーワードとともに、異分野が対話する'場'の設計が新たなエコシステムのフレームワークとして有効である。AI 社会論研究会は、これまで AI と社会の接点としての、その場を担っているとすれば、どのような効果を AI 技術と社会のあり方の議論に波及してきたのかを AI 社会論研究会の特徴である「HELPS」を指標に取り入れ、その議論とステークホルダーの多様性を中心に分析し、今後のあり方についても示唆する。

2. AI 社会論研究会とは

2.1 概要

AI 社会論研究会は「人工知能が社会に与える影響」について議論する会として、AI 研究者、高橋恒一と経済学者、井上智洋が共同発起人となり 2015 年 2 月 5 日に発足され、2016 年 12 月より法学者の新保史生が企画委員に加わっている。同会では、哲学（Humanity）、経済学(Economics)、法学（Law）、政治学(Politics)、社会学(Sociology)(略称：HELPS)による多様な観点からのアプローチを目指しており、2015 年 2 月 5 日から 2019 年 2 月までの間で計 37 回開催され、76 名から AI と社会に関する報告が行われている[参考 1]。

2.2 特色

人工知能が社会に及ぼす影響を検討する動きは、2014 年 9 月の人工知能学会誌（Vol.29, No.5）特集「人工知能技術が浸透する社会を考える」をきっかけに「AIR（Acceptable Intelligence with Responsibility）」研究会や、人工知能学会では、2014 年 12 月に倫理委員会が発足し 2016 年 6 月には倫理綱領（案）の後、「人工知能学会倫理指針」（2017 年 2 月 28 日）が策定された、総務省での AI ネットワーク化の影響やリスクの評価する AI ネットワーク化検討会議をはじめ、各省庁でも AI と社会に関する議論が行われており、日本国内外問わず AI と社会に関する議論は活発に行われている[参考 2][図表 1]。そのような潮流の中で AI 社会論研究会は、高橋・井上が提唱する哲学（H: Humanity）、経済学(E: Economics)、法学（L: Law）、政治学(P: Politics)、社会学(S: Sociology) (略称：HELPS) による多面的なアプローチを用いた研究を行うことを目指している。

図表1

2.3 HELPS

前述の通り、AI 社会論研究会が掲げる「HELPS」による AI 技術が社会に及ぼす多面的な議論のアプローチとは、従来の科学と社会の接点における検討で用いられていた「ELSI」ではその背後にある科学技術の性質上の違いにより不十分である、つまりは生物工学の発展を背景に用いられるようになった「ELSI」は、人工知能技術と社会における影響を論じるに、分野横断性の指標として不十分であると井上、高橋は主張している[高橋,井上 2017]。「ELSI」ではなく「HELPS」を用いた分野横断軸の指標は従来の科学技術がおよぼす影響としてのテクノロジ

ーアセスメント[神里 16]ではなく、AI 技術を中心にした技術が社会に与える影響の早期の予測としてのテクノロジーアセスメントの取り組みとして捉えることができるであろう。高橋、井上が提唱する「HELPS」は AI と社会の接点の多面的な議論を支える軸として、また AI 社会論研究会における特色とし捉え、本稿では「HELPS」参考にしながら AI 社会論研究会のこれまでの取り組みにおける多様性の分析と評価を試みる。

2.4 マルチステークホルダープロセス

AI 社会論研究会は HELPS を分野共創の軸としながら各分野からの報告が行われ情報が共有されている。AI 社会論研究会で発表を行った数多くの報告者が各省庁の AI と社会に関する政策を議論する委員会に出席していることを始め、その場での議論が人工知能学会の特集号[参考 3]となるなど、AI 社会論研究会での報告およびその場での議論は AI と社会における政策決定などの前提となる情報や認識共有型のマルチステークホルダープロセス[Hemmati 02]として捉えることができる。本稿では「HELPS」を指標に取り入れた多様性の分析を通して AI 社会論研究会においていかにしてマルチステークホルダープロセスが機能したかの検討も行う。

3. 多様性の可視化

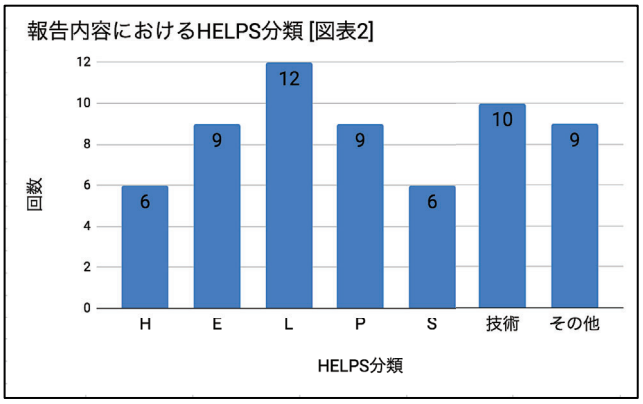
まずは(1)報告内容の多様性を HELPS に分類し検討する。さらに、AI 社会論研究会に関与するステークホルダーを(2)発表者、(3)会場の参加者、(4)会場(会場提供者)の3者に分類し、それぞれの属性と割合を多様性の可視化を試みる。最後に(5)国際性の分類も発表者の主たる所属機関の在籍国をもとに作成した。(1)～(5)の中でも(1), (3) については 2.3.で述べたように、HELPS を用いた分類を行った。

(1) 報告内容の多様性

AI 社会論研究会における、報告内容を以下のようにこれまでの発表内容を HELPS とその他で分類し、グラフ図により可視化した[図表 2]。 ただし、以下の各報告における HELPS 分類は、共同発起人、井上、高橋らの意見を参考にした分類である。

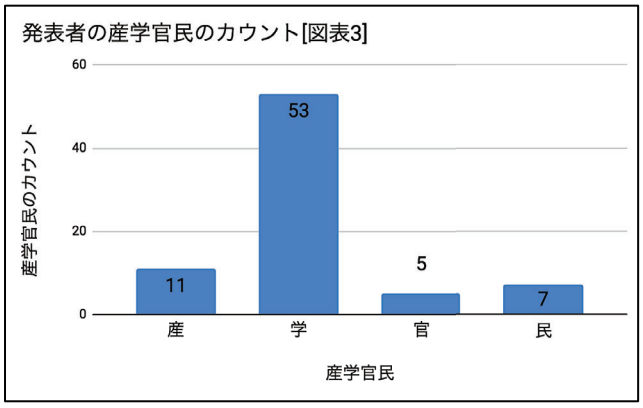
H	ロボットの人類学とは何か、知能機械をめぐる非連続的思弁と連続的実践の狭間で
	～機械、労働、人間～
	AI 社会における倫理のゆくえ:自己創出という分岐点および責任の所在、擬人化の先鋭
	オープン化を手掛かりに AI と人間の関係を考える
	AI 技術に対する人間の傷つきやすさと自律的で真正な生き方の可能性について
	「科学研究への AI の導入はどんな影響をもたらすか —Artificial Creativity の可能性と科学の未来」
E	人工知能・産業革命・経済成長—AI を制覇した国が世界を制覇する！？—
	カンヌ(広告業界)での AI 事情に関する報告
	経済学者は人工知能の夢を見るか？『大格差』と経済の将来
	『ぼくらの Pepper プロジェクト～親子で考えるロボットのいる暮らし～』から観測する、子供たちの考えるロボットとの共存
	「電気、ガス、水道、そしてキャリア」 の時代 シンギュラリティを迎える我々のキャリアと HR ビジネスの重責
	Fintech と AI:Fintech と AI の関係性について、Fintech スタートアップにいる立場から考察する
	＜AI 時代＞到来！ユニバーサル・デザインの「次世代エージェント」
	AI の普及とデザインマネジメントの発展
	経営戦略立案への AI 利活用の研究
L	ロボットと法をめぐる最近の政策動向について
	ロボット法に関するご報告
	AI とプライバシー・個人情報保護に関する法的論点
	AI・ロボットに関する法的論点の概観
	「ロボット・AI・キャラクターの“人格”的権利」
	「プライバシー権と個人情報保護——AI を語る前に整理しておきたいこと」
	ロボット・AI と医事法～医療過誤を中心に～
	AI and Robot Law
	ロボット法
	ロボット法:AI とヒトの共生に向けて
	人工知能に対する法人格の付与
	論理プログラミングによる要件事実論に基づく民事裁判支援システムの開発
P	AIネットワーク化の展望と課題
	AI ネットワークと 5G 時代のデータエコシステム
	標準化を用いた AI 倫理の方向付け: IEEE のイニシアチブを中心に
	米国における AI・ロボット技術の社会実装に向けた産官学の取り組み－開発原則の視座から
	シンギュラリティ以前の問題
	政策的視点で考える AI イノベーションの現状と課題
	AI は政治を変えるか？
	DATA-DRIVEN ELECTIONS IN DEMOCRATIC SOCIETIES: EARLY LESSONS FROM THE CAMBRIDGE ANALYTICA SCANDAL 民主主義社会におけるデータ駆動型選挙:ケンブリッジ・アナリティカ・スキャンダルから学ぶこと
	航空研究センターの概要と研究の取り組みについて
S	AIR(Acceptable Intelligence with Responsibility)に関する報告
	データ活用時代の＜ゆりかごから墓場まで＞のプライバシーを考える
	人工知能・意識・価値
	人工知能が拓く新しい社会デザイン
	How to Grasp Social Shaping of AI in East Asia
	AI への権限の委譲と心理的障壁
技術	Human-Agent interaction に関する報告
	人工知能×創造力で社会 を変えるスマートデータ・イノベーション
	民主化・個人化されたテクノロジーの一種としての AI
	対人インタラクションと感情モデル
	Hitachi AI Technology/H
	1.未体験の変化は制御しうるか　2.全脳アーキテクチャの活動のご報告
	AI とデータ整備

	人の認知状態推定するAI技術の可能性とリスク
	対話システムにおける諸課題～技術・サービス・倫理の側面から～
	人工知能と意識のプロジェクト
その他	シンギュラリティ大学に関する報告
	AIと監査
	AIのAIによる芸術の為の芸術
	アルキテクトーム(ARKHITEKTOME) ーアルゴリズミック・デザイン原理モデルー
	進化の時計の針は止まらず、人類の知的探求心が消え去ることもない。なればこそ。新旧の物語、つまりはSFと古典の両者に叡智を求めることが、脅威論の克服に繋がるだろう
	理化学研究所・未来戦略室
	SF作品と未来ビジョンについて――未来ビジョンの作り方の一例
	新規研究開発領域「人と情報のエコシステム」について
	「先端技術と共創する新たな人間社会」公募案内



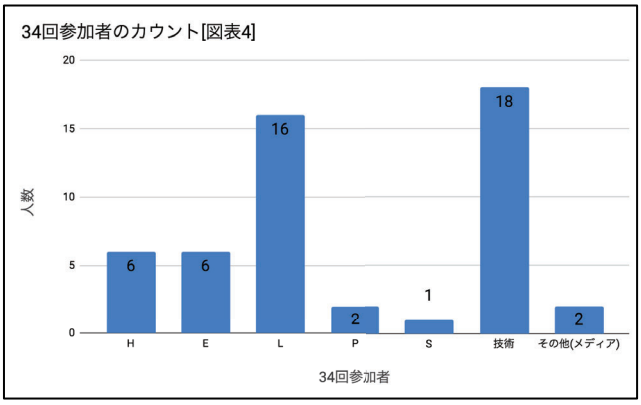
(2) 発表者の多様性

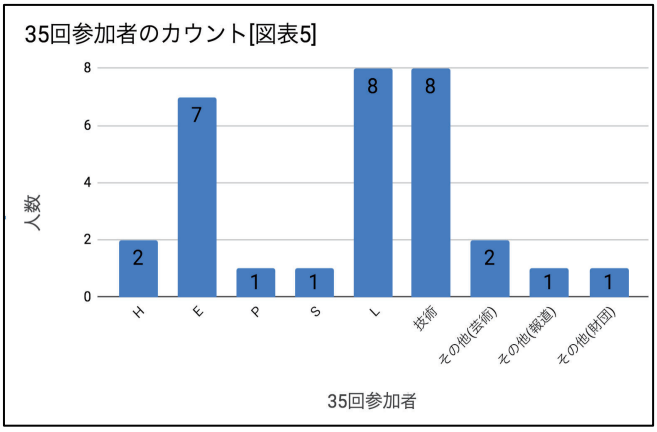
発表者の属性を産官学民に分類した。



(3) 参加者の多様性

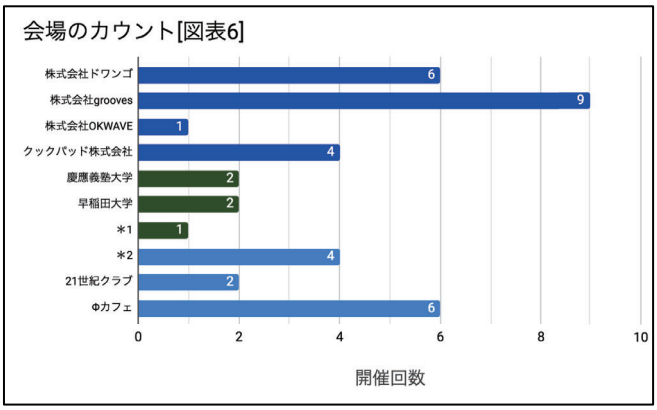
これまでに37回運営された内、第34回、第35回の、参加者をHELPSに分類し、図に表示した。ただし参加者のHELPS分類は事前の参加者アンケートで参加者自身が自らの属性を回答したデータである。





(4) 会場(会場提供者)の多様性

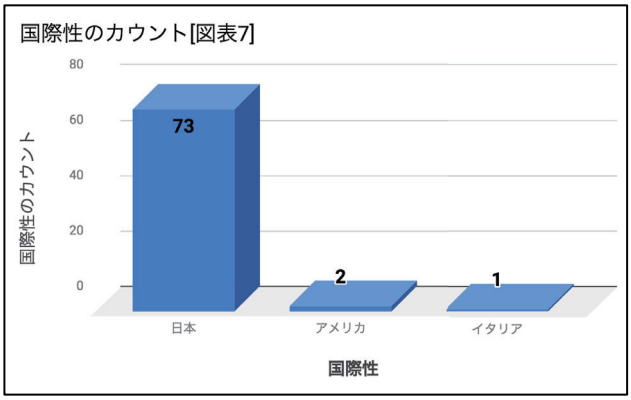
会場(会場提供者)を回数別に表示した。



図表の内、＊1は国際大学グローバル・コミュニケーション・センター(GLOCOM : Center for Global Communications)、＊2は理化学研究所革新知能統合研究センター(AIP)。

(5) 国際性

最後に国際的な多様性を検討するために発表者の国別に表した。



4. 分析

4.1.分析 1

AI 社会論研究会での議論が人工知能学会に特集として組まれている。3.(1)～(5)のように多様性の可視化を行ったが、AI 社会論研究会の特徴として、(2)で見られるように発表者の属性では学が飛び抜けて多いことから、アカデミックな議論が中心となっていることがわかる。AI はビジネス現場に応用される段階となるなど今後はアカデミックに閉じた議論だけでなく、ビジネス現場の議論またはビジネスとアカデミックを繋ぐような議論とそれを政策に繋ぐ場としての機能がますます期待されるのではないだろうか。また、図7から国際性が乏しいことがわかる。

4.2.分析 2(マルチステークホルダープロセス)

3.(3)では、AI 社会論研究会の参加者の HELPS 分類による多様性の可視化であるが、図表 4 は第 34 回研究会の参加者 HELPS 分類であり、図表 5 は第 35 回研究会での参加者の HELPS 分類である。

第 34 回研究会はトリノ大学(イタリア)のウゴ・パガロ氏からロボット法に関する報告(The Hard Cases of AI & Robotics, and their Legal Governance)[参考 4]であった。報告者の HELPS 分類が L:ウゴ氏にあたるところに、図表 5 からわかるように HELPS の多い順から、技術(18 名)＞L(16 名)>H=E(6 名)>S(1 名),その他メディア(2 名)の会場参加者であり、報告者 L に対して参加者は技術が L よりも多いことは異分野の交わりとして示唆的なデータである。

会場参加者の技術者からは AI とロボットの差とは何か、法的な問題の国際性と複雑化についてなど、ウゴ氏への様々な質問が飛び交った。ウゴ氏からは AI ロボットと法的な問題を複雑化している要因として重要な課題に決定を下すにはデータが不足していることから法的実験・探索主義の有効性が訴えかけられた。ウゴ氏の報告後、経営学者(E)からは、法制度のみならず組織など多様な分野で制度が変化するプロセス自体が、現在進行形で変化していると仮説が述べられた。

写真 1

(写真 1) は第 34 回の会場で技術者がウゴ氏(L)に質問する様子である。

第 35 回での Ugo 氏と企画委員の新保はパネルディスカッションを開催し、AI・ロボット技術の進歩と法制度の分岐点についての議論を行った。後に Ugo 氏と新保は CPDP2019 にて How the adequacy mechanism works: Progress in the EU's governance of cross-border data flows?と題したセッションを開催し日本の個人情報保護委員会が EU との長期の交渉の結果、十分性認定を獲得した報告を行った[参考 5]。

第 35 回 AI 社会論研究会の参加者 HELPS 分析から、イタリアから来日したウゴ氏の報告から日本の技術者を中心として広範囲の分野に影響し新保を経由し政策にまで広範囲に渡るマルチステークホルダープロセスが見られる。

5. さいごに

本稿では HELPS による分野横断軸を AI 社会論研究会における特色として捉え、AI 社会論研究会で議論された AI と社会における 37 回の議論を HELPS の指標を参考にしながらその多様性を分析した。この分析が今後の AI 社会論研究会のより効率的なマルチステークホルダープロセスを意識した運営、AI 技術が与える社会的影響を中心としたテクノロジーアセスメントの議論の発展に貢献することを願う。

＊本研究は，国立研究開発法人科学技術振興機構　戦略的創造研究推進事業(社会技術開発)「人と情報のエコシステム(HITE)」研究開発領域による研究成果の一部である。

謝辞

本項の執筆にあたっては、AI 社会論研究会共同発起人の高橋恒一氏、井上智洋氏、企画委員の新保史生氏、および AI 社会論研究会に関わる皆様にご指導いただきました。感謝いたします。

参考文献(論文誌と同じスタイルを推奨)

[参考 1] AI 社会論研究会 http://aisocietymeeting.wixsite.com/ethics-of-ai (最終アクセス 2019/2/15)

[参考 2] 人工知能と人間社会に関する検討の国内外の動向,内閣府,政策統括官(科学技術・イノベーション担当), https://www8.cao.go.jp/cstp/tyousakai/ai/1kai/siryo2.pdf, (最終アクセス 2019/2/15)

[高橋,井上 17] 高橋恒一,井上智洋: 特集「AI 社会論」にあたって,**人工知能 32 巻 5 号, 2017.**

[神里 16] 神里 達博: 情報技術における ELSI の可能性:歴史的背景を中心に, 情報管理,58 巻 12 号 p.875-886, 2016.

[参考 3] 特集「AI 社会論」, **人工知能 32 巻 5 号,** P.614-694, **2017.**

[Hemmati 02] M,Hammati: Multi-stakeholder Processes for Governance and Sustainability, Earthscan, 2002.

[参考 4] AI 社会論研究会,「ロボット法」の世界～ウゴ・パガロ教授をお招きして～http://aisocietymeeting.wixsite.com/ethics-of-ai/the-laws-of-robots (最終アクセス 2019/2/15)

[参考 5] CDPD2019 How the adequacy mechanism works: Progress in the EU's governance of cross-border data flows? , https://www.cpdpconferences.org/cpdp-panels/how-the-adequacy-mechanism-works-progress-in-the-eus-governance-of-cross-border-data-flows(最終アクセス 2019/2/15)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

オンライン社会としてのゲーム空間と、その速度耐性のための AI 実装についての考察

Gaming Space as an Online Society, and A Consideration about Its AI Implementation for Rapid Transformation Tolerance

梅野真也[*1] Shinya Umeno　中元正也[*1] Masaya Nakamoto　野濵哲也[*1] Tetsuya Nohama

[*1] 株式会社 mynet.ai
Mynet.ai Inc.

We discuss how we can consider an online gaming space as an online society. We also consider how we can implement AI that can tolerate rapid transformation of the society. An online community is one of the appropriate platforms for demonstration experiment of catastrophe-control AI, and we will explain the reasons.

1. はじめに

本稿ではオンラインゲームのユーザー空間に代表されるオンライン空間を、「オンライン社会」と捉え、その社会の下での AI 実装による社会の安定化の仕組みを考察する。とくに社会のバランスが壊れたカタストロフィー状態を起こさないような制御を行う AI は、どのように考えうるのか、また、その満たすべき性質や機能は何かのか、というところを論じていく。

実社会で IoT 化が進み、社会のデジタルトランスフォーメーションが進むと、社会のほぼ全データが取れる「オンライン化」社会が訪れることとなる。これに先んじて、完全性のあるデータが取れるオンライン社会にて社会バランスを保つような AI を実装し、その有効性を検証していくことは人類の発展にも一定寄与しうる取り組みとなると考える。

本稿の論点の土台は、[山川 2018]の考察について、オンライン社会という切り口で、その社会実装を進められないか、という考えに基づいている。特に[山川 2018]の以下に示す考え方に著者も呼応しており、これをオンライン社会にプロジェクションして考察を深めることで、さらに論考が深まることを目的とする。

- “技術にけん引される形で変化速度が増大”
- “これへの対処をなおざりにすれば、最終的には個別問題の対処が追い付かず破綻する”
- “AI が社会的決定を代行する必要がある”
- “AI が迅速に動作するためには、社会が共有する価値観の実装が必須であり、その設計に難しさがある”

1.1 社会構造の法による秩序とその速度耐性の限界

現在の実社会では、法律が存在し、それが社会に属するメンバに経済性・社会性の観点から、社会に対して不適切と考える行動への抑止力として機能してきた。このような社会では、以下のようなペナルティの観点から、抑止力が発生していると考えられる。

1. 経済性のペナルティ
   (ア) 法律違反での罰金
   (イ) 懲戒解雇などによる経済基盤の棄損
2. 社会性のペナルティ
   (ア) そのメンバが何をしたのか、またそれが「違法」であるという権威付けされた事実が知れ渡ることによる、所属していた社会・コミュニティでの居づらさ
   (イ) 前科として、過去の行いがその後の自らの人生について回ること

このような法治社会での統治が 20 世紀までは機能してきたが、近年では変化のスピードのみが上がり責任の所在やペナルティがあいまいになっているという現象が起きていると考えられる。第 3 章で詳しく述べるが、例えば為替でのアルゴリズミック・トレーディングによる為替のカタストロフィックな値の変化において、社会的なインパクトは大変大きい。しかしながら、法的にそれを取り締まって、市場の本来の流動性を失わせるような方策はとられていないし、今後もとられないかもしれない。また、法整備前に起こったカタストロフィックな事象について、その責任を追及していくことは難しい。

[山川 2018]も“レイテンシ問題”として同様のことを述べるように、社会の変化スピードが指数関数的に上がっていく現在において、社会の意思決定（例えば為替の値付け）は他人や機械・AI に委譲され、またそれによってさらに変化速度が高まっていく。このような状態が続くと、カタストロフィーと捉えられるような事象の発生頻度が上がり、法律が出来上がった時には「元の木阿弥」となってしまうような社会が出来上がってしまう。

このような変化速度の高まった社会では、法治の力が届く前に、まずは何かしらカタストロフィックな事象が起こることを 0 時間でとらえ、またより良くは事前察知し、カタストロフィーの制御というものを組み込んでいかないといけない。さらには、カタストロフィーを起こしている側が AI や機械なのであれば、察知し対抗する制御側も AI とすることで対応速度をマッチさせていく必要がある。

1.2 本稿の構成

本稿のこれ以降は以下のような構成となっている。

第 2 章でオンライン社会としてのスマーフとフォンゲームについて論じる。なぜ社会といわれるべき要素を持つのか、またその特性から経済的カタストロフィーが発生する社会としての縮図としてとらえることができることを説明する。

第 3 章では、カタストロフィー制御 AI の社会実証の場として捉えた場合のオンライン社会や、カタストロフィーをどのように記述しうるかという例を述べる。

第 4 章では、カタストロフィー制御を含む社会バランス AI について、その実現すべき性質や考えうるアーキテクチャについて論じる。

第 5 章でまとめる。

2. オンライン社会としてのスマートフォンゲーム

オンラインゲーム、ソーシャルゲーム、と言われる種類のスマートフォンゲーム（スマホゲーム）のうち、現在日本で配信されているものの多くは、以下のような仕組みで成り立っている。

ユーザーは自らの資産であるカードをゲーム内の仮想空間上で取得し、「デッキ」といわれる資産ポートフォリオを組み立てる。このカード取得は、主要なものに例えば以下の3つが上げられる。

1. ゲーム内で遊んでクエストやイベントの攻略報酬として取得する。
2. 一日一回ゲーム空間を訪れることで取得できる“ログインボーナス”として無償の“石”と呼ばれるゲーム内仮想通貨を貯めることでカードと交換する。
3. 有償で現実世界の金銭をアプリ内課金として支払い、“有償石”を取得することでカードと交換する。

現在のデジタルゲーム国内市場の年間 1.2 兆円のうち、約 9000 億円がこのような形のオンラインゲームのよって 2000 年代~2010 年代に新たに作り出しされた市場であると言われている。

ユーザー資産であるデッキの各カードは、そのままユーザーと一緒に戦うモンスターとしてゲーム空間内に現れたり、またユーザーの装備である武器や防具となったりすることで、空間内に実効力をもたらす。簡潔に言えば、より強いカードを入れたデッキを構成すれば、ゲームの攻略スピードが上がったり、プレイヤー同士の対戦の際に有利に試合を進められたり、ということになる。

特にプレイヤー通しの大戦において、「ギルド」と呼ばれるチームを組成し、あたかもサークルや部活のようにオンライン空間で毎晩試合を行う、というようなことが行えるギルド・バーサス・ギルド（GvG）というチーム戦を取り入れているようなゲームタイトルも多数存在する。このようなタイトル内の各チームでは、スマホのチャットアプリのようなグループチャットが用意され、そこで次回の試合に対する作戦を立てたり、ゲームの最新のカードやキャラについて仲間としてコミュニケーションをとったり、また時には天候の話や当日何を食べたかなど、通常の友人と変わらないようなカジュアルなコミュニケーションを取ったりする。

下記に、このような空間における現実社会との違いを、本稿の軸である社会的決定の委譲とAI実装の観点からまとめたものが下のテーブルである。

	オンラインゲームのオンライン社会	現実社会
社会的決定の委譲の例	最新のカードの強さのインフレ率を調整することで、資産がどのような速度で入れ替わるか、また原資産が目減りしていってしまうのか、を運営が調整する。	中央銀行が物価のインフレ目標を設定し、公定歩合の決定や金融の量的緩和を行う。
取得データの完全性	社会のなかのユーザー行動データが全て取得できる（完全性がある）	不完全なデータ（IoT が極限まで行き、またプライバシー問題が解決できるのであれば、完全性あるデータ取得がやっと可能性を帯びる。）
AI のコアとなる機械学習部分構築の容易性	データの完全性により、社会で起こっている事象のもととなるデータはほぼ完全に取れているといってもよい。このため機械学習の構築が実社会より容易	データが取れていない事象が予測性の根本となる場合もある。組みたい AI に本質となるデータが取れているか次第。
社会的決定の AI 代行の社会実装難度	ゲーム運営の事業者としての経済インパクトやユーザー空間での資産性の崩壊など、重要な事象が起こる可能性は残るが、実社会よりは許容度が高く、先行して実装していくことが可能と考えられる。	社会やコミュニティという単位でAIが何らかのバランスを取る場合、人命や生命維持、経済インパクトなどのミッション・クリティカルなタスクへの AI 実装はリスクが高く、進みは遅い。
カタストロフィーの例	・プログラムのバグを突く無償石や有償石の無限増殖 ・ゲーム内のユーザー間カードトレードを悪用した、ボット（機械的/AI 的エージェント）によるユーザーなりすましと優良カードの1ユーザー資産への集中化	・アルゴリズミック・トレーディングでの為替等の社会基盤へのカタストロフィックな変化 ・電力融通を行うスマートグリッドでの悪意のある参加者や運営による電力需給の人為操作

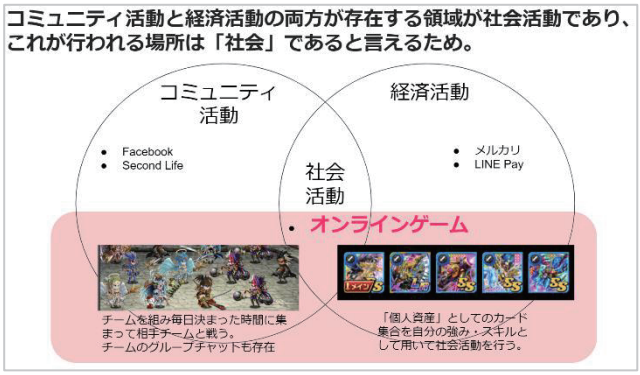
2.1 なぜオンライン「社会」なのか

上記で述べたようなオンライン空間は、「社会」と呼んでもよい要素を兼ね備えていると筆者は考えている。その理由は以下である。

コミュニケーションと共同タスクという社会性：まず、社会として基本となる、所属メンバ同士のコミュニケーションが取れる。これは前述のように、友達とチャットアプリでするような会話に似たものも含まれる。また、チーム戦等を通して、共同で同一のタスクに向かい、タスク達成条件を満たす（例えばチーム戦で勝利する）ために協力して取り組みを行う。

資産形成と資産性保持による経済性：次に、デッキというものを通した資産形成が存在していることである。この資産は、そのポートフォリオ構成要素であるカードに新たなより強い（より価値の高い）ものが、過去のものと同額でマーケットに売り出されることにより、実質的資産価値が時系列変化していく。ちょうど家電において、ほぼ同じような値段で新たなモデルが出て行ったり、そのモデルの性能が前のモデルより優れていたりすることで、いち家庭における家電の資産価値が時系列変化していくことに近い。ゲームの空間では、資産価値の目減りは、よりゲームの中で勝ちにくくなる、ということを意味する。ユーザーはある程度制御された運営のインフレ計画の中で、自分の経済事情にマッチしたような資産性維持戦略を取り、引き続きゲームが楽しめる状態を保とうとする。

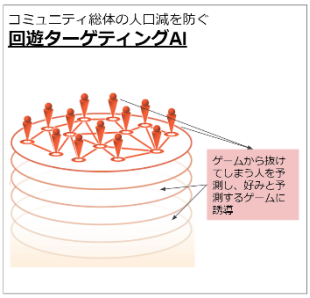
このような、社会の構成要素としての Minimal Vial Product（MVP）を満たすような空間がオンラインゲームの空間であり、それのよって、著者は同空間を「オンライン社会」と呼んでいる。（下図）



また、このような空間に対して、そのバランスを制御するような AI の研究開発を著者の所属する組織では進行している。(下図)

筆者の組織のグループ企業においては、ゲームタイトルを多数運営しているため「、オンライン社会が複数存在しているとも考えられる。これらの社会について、ある社会メンバがそろそろその社会を去ろうとしていることを察知し、次のマッチする社会を推薦するような仕組みを考えている。(下図左)

また、ユーザー資産となるデッキの状態を把握し、社会全体の資産バランスが崩れない形で新たなカードの強さを設計するようなサポート AI についても開発を行っている。(下図右)



2.2 レイテンシ問題とカタストロフィーの例

次に、上記のようなオンライン社会においてのカタストロフィー問題を取り上げる。基本的に、ゲームのオンライン社会でのカタストロフィーは、急速な経済状態の変化によってもたらされることが多い。例えば、カタストロフィーの例でテーブル内に述べた無償石の無限増殖であるとか、強いカードや攻撃力を上げるアイテムの無限トレードをボットによる不適正ユーザーの増殖によって起こされたような事象である。

2.3 共有する価値観の実装

[山川 2018]も述べるように、合意形成結果を近似するような評価関数がないと AI が学習できない。これに対して、山川は2つの案を出している。「民意の高速な推定」と「より上位の社会的目標の設定」である。

ゲーム空間においては、「民意の高速な推定」は“楽をして勝ちたい”など、必ずしも社会が活性する方向に動くような評価関数をもたらさない場合が考えられる。よって、「より上位の社会的目標の設定」のようがより実効的であると考えている。その社会に「住む」住人にとっては、オンライン社会自体の存続というものが一番の上位的な社会目標となりえる。これを運営側としてトップダウンで設定することは、ユーザーと Win-Win の関係を築けることを前提として、取りうる選択となる。

その目標のために、運営側の「中間的目標」として、課金により事業的にも社会持続に貢献してくれているロイヤルユーザーを維持する、というものが考えられる。

3. カタストロフィー制御 AI の社会実証の場として

ゲームのオンライン社会は、カタストロフィーの対処に対して非常に進んでいる社会であるという見方ができる。これは、もしカタストロフィックな事象が起きたとしても、仮想空間内のことであるので、基本的には法律に乗っ取った責任分解や裁定が起こらないことに起因する。ユーザーは規約を守るという制約のもと、この社会に足を踏み入れるが、それを破った場合の実社会での大きな罰則は今のところ存在しない。(実社会での罰されるセキュリティ・クラッキングのようなものはもちろん本論の外である。)

このようにチート行為をリスク少なく起こしやすいため、ゲーム運営ではしばしばカタストロフィックな事象が発見される。このような事象は、何万といるユーザーが総攻撃的に不具合を悪用して権利を手に入れたり、またボットにより CPU/GPU の速さでチート行為が行われ、その行動者が増殖していく場合もある。今の運営では、迅速にそのようなユーザーを発見し、規約違反で追放(バン)したりすることで実効力を持って対処をしている。しかし、そのカタストロフィーによって崩された経済性のバランスは一朝一夕には解決されず、例えば無償石が大量に社会内にばらまかれた場合にはその影響の終結のために2～3か月、経済バランスを崩し課金が抑えられた中で運営をせざるを得ない場合もある。

このような重大な帰結をもたらしうるカタストロフィーについて、その発生を予期したり、人間では間に合わないスピードでの対処・裁定によりバランスされた状態に持ち込むための AI が必要となる。このような AI について、概念で試行することは現実社会でも可能であるが、例えば、以下のようなバランスされた状態を保ち続ける例について、実社会への影響を押さえながらの実証というのは難しい。これは、ひとたびカタストロフィーが起きたときの影響が大きいということに起因する。一方で、非常に制限された、人工的なデジタルシミュレーションを作ることでその影響を図ろうと考える諸氏もいる。しかしこのシミュレーションでは、実際に「人間とはどう考えて何という手を連続的に売っていくのか」という事象への再現精度に問題が出てくる。

1. 地域の需要家どうしで電力を融通するスマートグリッド内のカタストロフィー:例―米国エンロン社による、特定参加者の経済性のための人為的需給の非バランス化
2. サブプライムローン問題や不動産バブルのような、リスクとリターンの関係性がマスクされてしまった商材に対するリスク顕在化やバブルの崩壊
3. 為替のアルゴリズミック・トレーディングによる、為替高や為替安の連鎖反応的でカタストロフィックな価格変動
4. 特定の電子マネーを使うとある確率で全額キャッシュバックのような、経済行動や経済価値観の変化をもたらしうる施策

一方で、オンライン社会であれば、人命や(電力のように)生命維持にかかわるようなミッションクリティカルでない、また上記のような経済性からカタストロフィーは起こしたくない、という強い動機も働く。このため、カタストロフィー制御を含む社会バランス AI の実証として適すると筆者らは考えている。

また、データが取れていないと、カタストロフィーがなぜ起こるのか、どうやって予測し得たのか、の解明が一向に進まない。社会のメンバの取りうる手と取得されたデータの完全性のバランス

の観点でも、オンライン社会はカタストロフィー制御の実証に適すると考えている。

3.1 中間目標としてのカタストロフィー制御の定義

実際のカタストロフィーについては、時相論理、様相論理のようなもので、その定義であるとか、またそれがコントロール状態になることが記述可能であると筆者は考えている。

例えば、「一定時間 T 内の社会数値 X の変化量は Δ 以内に収まる：□t≦T　δ (t→t+T)|X|≦ Δ」のような記述である。

また、カタストロフィーが制御された状態は、「社会数値 Y が変化量 Δ2 を超えた場合、時間 τ 以内にその変化値は Δ 以内にまで戻る制御時系列が存在する：□（□ δ (t→t+T)|Y| ＞ Δ2 → ◇t≦t+ τ δ（t→t+ τ）|Y| ≦ Δ）」というような制御下の状態を表す時相論理も考えられる。

4. カタストロフィー制御を含む社会バランス AI について

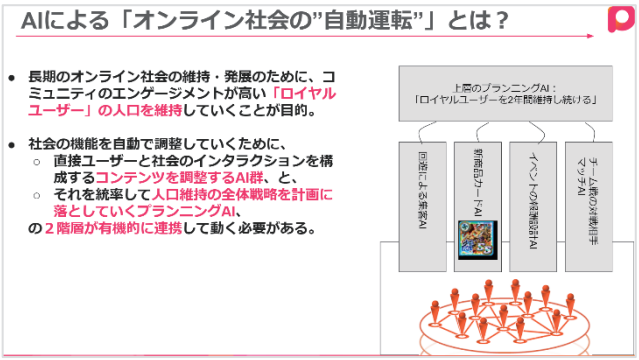
本章では、カタストロフィーに対応するための社会バランス AI の持つべき機能について考察する。

4.1 より上位の社会的目標の導入

長期のオンライン社会の維持・発展のために、コミュニティのエンゲージメントが高い「ロイヤルユーザー」の人口を維持していくことが目的とする。

4.2 階層型の AI

社会の機能を自動で調整していくために、直接ユーザーと社会のインタラクションを構成するコンテンツを調整する AI 群、と、それを統率して人口維持の全体戦略を計画に落としていくプランニング AI、の2階層が有機的に連携して動く必要がある、と考えている。



4.3 中間目標としてのカタストロフィーの制約項

第 3.1 節のような時相論理の制御要件を受け取って、それを解釈するような AI か、AI に理解できる言語に変換するコンパイラのようなものが必要になってくると考えている。

このような時相論理による制御則を満たすかどうかについては、"ハイブリッドシステム"といわれる制御とコンピュータ科学の学際分野にて研究がされてきた。

ハイブリッドシステムの分野では、制御するものされるものをスイッチング制御理論的・オートマタ理論的にモデル化し、同モデルに対するリアプノフ関数やエネルギー関数の単調減少での安定化の照明や保証を行う。[Myla 2008] [Mitra 2008]

また、事象や時系列の抽象化後の言語（データ）を把握する AI も必要になるかもしれない。このような言語とその計算モデルの検証は以下の論文で述べられている [Umeno 2008]

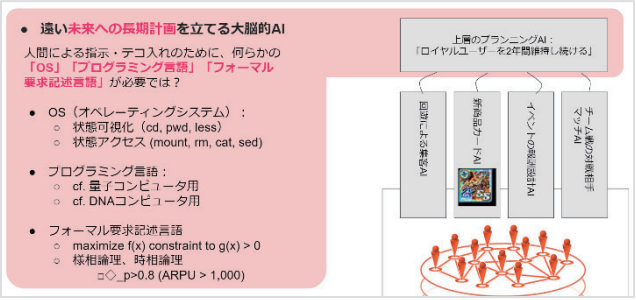
4.4 ハッカブルな AI の必要性

上層となるプランニング AI は、LSTM のような、リカレント性を持ち、状態をある程度メモリする時系列系の機械学習モデルで動くはず。メモリを持ったニューラル回路は、回路図の左から右への信号増幅を行うトランジスタの回路から、論理ゲートに再帰入力を掛けてフリップフロップを作ることで S-RAM となったデジタル回路にも似ていると筆者は考えている。

カタストロフィーに AI のみで対応しきれないと判断がされた場合、そのハックのための OS（メモリ仮想化）やプログラミング言語も必要であると考える。ニューラル回路内の状態を効率的に高次の言語で変えていく必要があるからである。（下図）

最終的に時代が進めが、ハック部分のプログラム/スクリプトを含めて AI が1つのシステムとして構築されると考えている。

また、カタストロフィーをある程度自動で脱出するために、異常検知のための機械学習アルゴリズムにも似た動きを持った忘却ゲートも必要かもしれない。



5. おわりに

オンライン社会は実社会への応用を考える土台としてどのように適しており、また、カタストロフィー制御 AI の実証場所として検討できそうかを考察した。今後、AI のオンライン社会への実装を進めるとともに、実社会への応用も積極的に考えていきたい。実社会への同様の考え方のフレームワークについては、北陸先端科学技術大学院大学の内平教授・岡田准教授と共に、思考フレームワーク＋アイデアソンの形でのブートキャンプ集中事業を行った。[内平 2019]　このような機会も活かしていく。

参考文献

[山川 2018]　山川 宏: 速度耐性をもつ社会の実現にむけた基礎的考察, 2018 年度人工知能学会全国大会（第 32 回）, 2018.

[内平 2019]　内平直志, 梅野真也：イノベーション・デザイン手法の地域課題解決への適用 ― 石川県能美市×AI／IoT アイディアソン ― ，サービス学会 第 7 回国内大会, 2019

[Myla 2008] M. Archer, H. Lim, N. A. Lynch, S. Mitra, S. Umeno: Specifying and proving properties of timed I/O automata using Tempo. Design Autom. for Emb. Sys. 12(1-2) (2008)

[Mitra 2008] Sayan Mitra, Daniel Liberzon, Nancy A. Lynch: Verifying average dwell time of hybrid systems. ACM Trans. Embedded Comput. Syst. 8(1): 3:1-3:37 (2008)

[Umeno 2008] S. Umeno: Event order abstraction for parametric real-time system verification. EMSOFT 2008: 1-10

[Umeno 2010] Shinya Umeno, Nancy A. Lynch: Automated Formal Verification of the DHCP Failover Protocol Using Timeout Order Abstraction. ICECCS 2010: 136-145

2:40 PM - 3:00 PM (Wed. Jun 5, 2019 1:20 PM - 3:00 PM Room I)

[2I01-04-5] Discussion / Conclusion

Discussion / Conclusion

©The Japanese Society for Artificial Intelligence

Organized Session | Organized Session | [OS] OS-15

[2I4-OS-15b] 人と AIが織りなす新たなエコシステム(2)

中川 裕志（理化学研究所）、山川 宏（ドワンゴ人工知能研究所／全脳アーキテクチャ・イニシアティブ）、浅田 稔（大阪大学）、井上 智洋（駒沢大学）、江間 有沙（東京大学）、金井 良太（アラヤ）、高橋 恒一（理化学研究所）、萩田 紀博（ATR）、堀川 優紀子（ATR）、松原 繁夫（京都大学）

Wed. Jun 5, 2019 3:20 PM - 5:00 PM Room I (306+307 Small meeting rooms)

[2I4-OS-15b-01] A Study on Formation of Social Consensus on Personal Data Utilization in AI Society

〇Etsuko Tane[1] （1. The University of Tokyo）

3:20 PM - 3:40 PM

[2I4-OS-15b-02] AI as agency in agency law

〇Ryota Akasaka[1] （1. National Institute of Advanced Industrial Science And Technology）

3:40 PM - 4:00 PM

[2I4-OS-15b-03] (Invited talk) Legal Issues toward the New Ecosystem Interwoven between Human and AIs

〇Satoshi Narihara[1,2] （1. Kyushu University, 2. The RIKEN Center for Advanced Intelligence Project）

4:00 PM - 4:40 PM

[2I05-07-4] Discussion / Conclusion

4:40 PM - 5:00 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

AI 社会におけるパーソナルデータ活用の社会的合意形成に関する考察

A Study on Formation of Social Consensus on Personal Data Utilization in AI Society

多根　悦子[*1]
Etsuko Tane

[*1] 東京大学大学院 総合文化研究科
Graduate School of Arts and Sciences, The University of Tokyo

In recent years, AI technology has been developing rapidly beyond social acceptance. Although utilization of personal data has advantages of bringing convenience to people and enriching our lives, we also have risks of being abused. Social consensus formation is required to utilize technologies with high uncertainty. In order to practice a valid consensus formation, the diversity of decision-making subjects, the disclosure of necessary information, diversity of options and transparency of the decision-making process is necessary. In this paper, the tasks of utilizing personal data are overviewed and consider how to form a consensus to ensure social rationality.

1. はじめに

近年における通信技術, コンピュータ, IoT（Internet of Things）等のデジタル技術の急速な発展によって, 世の中に流通するデジタルデータの量・種類は急激に増加を続けている. 例えば, WEB 上での購買履歴や行動ログ, 位置情報, 交通系 IC カード利用履歴など, 大量・多種のデータ（ビッグデータ）を取得・保存できるようになり, それらのデータを分析することによって予測や認識をする AI 技術が発達してきている. しかし, 個人の ID に紐づくパーソナルデータは, 個人のプライバシーへの配慮や個人情報漏洩などのデータの取り扱いのリスクから, 活用メリットがあると理解されつつも現実には活用が進んでいない [本村 15] ことが課題となっている.

このような状況の中で,パーソナルデータを社会的に活用できるようにするためには PDS（Personal Data Store）とよばれる, 個人が自分のデータを蓄積・管理・共有できる仕組みが必要であり [橋田 13], 内閣官房 IT 総合戦略室では 2016 年 9 月から「データ流通環境整備検討会」を開催し, PDS とそれを代理運用する情報銀行（情報利用信用銀行）, データ取引市場などのデータ流通の仕組みについて検討してきた [総務省 18]. さらに, 総務省は個人データを収集・管理する「情報銀行」の事業者認定を 2019 年に開始すると発表し, パーソナルデータの活用の環境整備が進みつつあり, 社会からの注目はますます高まってきている.

2. パーソナルデータの定義と課題

パーソナルデータは個人情報と同義語と考えられがちであるが, インターネットの閲覧履歴や位置情報など, そのデータ単体では個人を特定できないものも含まれる. 2011 年に開催された世界経済フォーラムにおいて,「パーソナルデータ：新たな資産カテゴリーの出現（Personal Data: The Emergence of a New Asset Class）」がまとめられ, そのレポートの中でパーソナルデータの言葉が定義されている [WEF 11] . 本研究では, 世界経済フォーラムのレポートに記述されたパーソナルデータの定義を踏襲した（表 1）.

表 1　本稿におけるパーソナルデータの定義

ボランティアデータ	個人が自らシェアしたデータ 例：SNS のプロフィールなど
観察データ	個人の行動の記録から作られたデータ 例：携帯電話使用時の位置情報など
推測データ	観察データの分析から得られた個人に関するデータ 例：与信情報など

また, 同レポートの中でパーソナルデータは,「インターネットにおける新しい石油であり, デジタル世界における新たな通貨である」と記述されており, 世界的にもパーソナルデータが社会や経済にとって有用であり, 価値あるものであると示されている.

一方でパーソナルデータは, データの主体であるユーザーの管理・コントロールから離れて本人の知らないところで利用されるというリスクを有する. 例えば WEB サイトで使われているリコメンド（推薦）機能や検索エンジンを利用した際に画面上に表示される広告や検索結果の一部（リスティング広告）は, ユーザーのインターネット上での行動ログや購買履歴などのデータの分析をもとに表示されており, ユーザーの意思とは関係なくパーソナルデータが企業の宣伝活動に利用されている. また, 2018 年に発覚した Facebook の情報漏洩問題は, アプリ経由で 8700 万人の個人情報が流出し, アメリカ大統領選にも影響を与えたと言われている. さらに, 中国の電子商取引サイト最大手のアリババ（Alibaba）が傘下におさめるオンライン決済サービスのアリペイ（Alipay）は利用者の決済情報から信用力スコアを算出し, 住居の賃貸や融資の場面, 外国のビザ取得にまで利用されているといわれている [山本 17].

急速に技術発展が進むことにより, それが社会にどのような影響を与えるか答えることができない不確実性を持つ問題をワインバーグは「トランスサイエンス（Trans-Science）」と名付けた [Weinberg 72]. AI は急速な技術的進化と共に社会への適用も同時進行で発展しており, その中でもパーソナルデータは, その活用がユーザーや社会に対してどのような影響をもたらすか, あらかじめ全てのリスクを予見できない, トランスサイエンスの領域にあるといえる. 本稿では, 多様なステークホルダーの期待と懸念が混在し, かつ技術の発展に対して社会的不確実性がある中で, 健全にパーソナルデータを活用するための社会的合意形成について考察する.

連絡先：多根 悦子, 東京大学大学院 総合文化研究科 広域科学専攻, etsukotane103@g.ecc.u-tokyo.ac.jp

3. 社会的合理性と意思決定

科学技術と社会の接点で生じる問題を解決する手段として，科学主義が強い技術官僚モデル（technocratic model）と，より多くの価値観を導入する民主主義モデル（democratic model）がある[藤垣 03]. 日本の科学技術をめぐる社会的意思決定では前者が用いられがちであるが，技術官僚モデルは科学的要素が強いがゆえに，理想的な状況で取得されたデータにもとづいて意思決定されており，現実の社会的場面には即さない．パーソナルデータは先に述べた通り，AI の技術の発展に伴い活用範囲も急速に広がっており，事前に科学技術の側面からリスクを判断することが難しく，民主主義モデル，つまり多様な価値観を加味した社会的合理性をもとに妥当性境界[1]を判断する必要がある．では，社会的合理性はどのように担保されるのか．藤垣によると，「公共空間」における科学技術の問題に対する意思決定は，下記の3つによって社会的合理性が担保されなければならない．

1. 意思決定の主体の多様性
2. その意思決定に必要な情報の開示，選択肢の多様性
3. 意思決定のプロセスの透明性

これらをパーソナルデータに当てはめて考えてみると，1についてはパーソナルデータの活用をめぐり，企業とユーザー以外にデータアグリゲータ，政府，法律家，研究者など多様なステークホルダーが存在する．これらのステークホルダーが主体となる議論の場を設計し，未来に起こりうる問題や技術動向を話し合う必要がある．2の情報開示および選択肢の多様性については，専門家（知識を持つ者）と一般市民（知識を持たざる者）としての関係（欠如モデル）による二者択一の意思決定ではなく，各ステークホルダーが情報を持ちよることにより，複数の選択肢を構築していくことができると考えられる．有効な手段として，技術評価の手法のひとつであるシナリオワークショップ[2] [平川 02] を用いてパーソナルデータ活用がもたらす未来について考えることができる．また，これは多様なスタークホルダーが参加し，その結果を公開することによって，3の意思決定のプロセスの透明性を担保することもできる．

このように，社会的合理性にもとづいてパーソナルデータ活用の合意形成をはかることにより，不確実性が高い未来に対してあらゆる可能性を予見し，それぞれの立場から行動をおこすことができると考えられる．

4. ステークホルダーの認識および態度

本研究では，パーソナルデータ活用に対する社会的合意形成の実践として，まずは各ステークホルダーの認識的課題を明らかにするために，一般市民参加型のワークショップを開催（2回）し，議論のスクリプトに対して共起語分析，感情分析，フレーミング分析を行った．ワークショップは参加者が様々なステークホルダーを演じることによって，主体の多様性を仮想的に再現した．

1回目のワークショップでは，ユーザーの立場として参加者のパーソナルデータ活用の認識・受容の程度について明らかにし，どのようなメリットとリスクが考えられるかについて議論した．その後，今度は役割を変えて企業側の立場に立った場合に，どのように考え方が変わるかを観察した．その結果，ユーザーの立場で考えた時は，情報収集や利便性を重視しつつも，個人を特定されやすい情報の提供に懸念を示していたものの，企業側の立場に立って考えると個人の特徴や属性を表すデータの取得を目指しており，同一人物であっても立場を変えると懸念や取り扱いの考え方が変化することがわかった．

2回目のワークショップは，参加者が様々なステークホルダーを演じるロールプレイングの手法を用いて行った．その結果，議論の内容から各ステークホルダーのパーソナルデータに対する態度や主張がフレームとして表出し，倫理的な問題や新たな利害関係が生まれたりするなど，パーソナルデータ活用のための課題と解決のための選択肢が抽出された．また事後アンケートから，ワークショップで様々な立場に立って考えることによる参加者のパーソナルデータに対する意識の変化も見られた．

5. おわりに

本稿では，AI 社会におけるパーソナルデータ活用について，社会的合理性が担保された合意形成について考察し，認識的課題を明らかにした．一方で，専門家の態度，意見に関してはワークショップ参加者の想像の域に限定されている．次のステップでは，これを補完する形として実際の専門家へのインタビューを実施し，各ステークホルダーの立場からのパーソナルデータの捉え方を確認する．

現代の AI 社会では，その技術発展の速さゆえに不確定要素を残したまま意思決定を迫られ，その都度，倫理的または合理的な判断基準が求められている．しかし，多くのケースでは社会的な合意形成を得ないまま技術が浸透しているのが現実であり，その影響を誰も予見できないという大きなリスクをはらんでいる．この問題はパーソナルデータに限らず，自動運転やロボットなど他のシーンでも同様である．本研究の発展として，AI 時代の社会的合意形成モデルを構築し，各分野に応用していくことを今後の課題とする．

参考文献

[本村 15] 本村 陽一: AI 技術によるビッグデータ活用の課題，人工知能学会第 29 回全国大会予稿，1J3-OS-18b-4, 2015.

[橋田 13] 橋田 浩一: 分散 PDS による個人データの自己管理 (< 特集> パーソナルデータに基づく気付きの創発)，人工知能学会誌 Vol.28, No.6, pp.872-878, 2013

[総務省 18] 総務省: 情報通信白書，第 1 部 特集 人口減少時代の ICT による持続的成長，第 2 節 日本と世界のデータ関連制度，2018

[WEF 11] World Economic Forum: Personal data: the emergence of a new asset class. In World Economic Forum, 2011

[山本 17] 山本龍彦: おそろしいビッグデータ: 超類型化 AI 社会のリスク，朝日新聞出版，2017

[Weinberg 72] Weinberg, A. M :Science and Trans-Science, , Minerva 10(2), 209-222, 1972

[藤垣 03] 藤垣裕子: 専門知と公共性，東京大学出版，2003

[平川 02] 平川秀幸: デンマーク調査報告書― シナリオワークショップとサイエンスショップに関する聴き取り調査 ―, 2002

[1] 意思決定において必要な判断基準，判断基準の根拠

[2]社会変化に関わる人々がそれぞれ立場から見て望ましい未来像を描き，全員が共有できるビジョンとそれを実現するための行動プランを定めるワークショップ

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

代理人としての AI の検討

AI as agency in agency law

赤坂　亮太[*1]
Ryota AKASAKA

[*1] 産業技術総合研究所　特別研究員
AIST Postdoctoral Researcher

Many scholars arguing that acts of artificial intelligence affect to modern law systems. There are some discrepancies between human beings in law systems who act based on "free mind" to humans in terms of today's science, and AI actualize this discrepancy. Due to solve this problem in a field of contract law, some US law scholars arguing to interpret AI as an agency in common law. In this article, I describe applicability and limits for this argument.

1. はじめに

1.1 問題背景

AIのような自律的に行為をしえる非人間の存在が、近代法以降の法体系に影響を与えること、とくにいかに法的責任を考えるのかというテーマは近年盛んに議論されている。その中で、AIに一定の法的立場を付与し、現行法の延長線上に回収しようとする議論も行われている。

他方で人間の側に目を向けると、「意思能力、行為能力、責任能力といった「能力」概念を媒介として、行為を現実化させる「意思」(行為の創発)から、行為の「効力」、そして行為(の帰結)に対する「責任」に至る一連のプロセス」として人の行為を自由意思による自律的な行為とする法の人間理解については、近時の人間諸科学の研究成果をもとに法学領域においても疑問が提起されている。たとえば法律時報の 90 巻 12 号「特集・現代法における「人間」の相対化」ではこの問題意識をもとに、法的に扱われる人間像が再検討されている。特に AI との関係において、角田[2018]は特に「弱い AI」を含む高度な道具について、このような思考能力を拡大できるような「行為能力のレバレッジ」をかけるレバレッジ機構が日常的になることによって「身の程を知らない自由がモンスター化」し、帰責対象としての人間画から責任が溢れ出すことを懸念する。[角田,2018;P37]

このような高度な情報技術と人間が協働する場面において従来の帰責構造では人間に過大な責任が生じ得るという問題意識は、他にも、Elish[2016]が道徳的クラッシュゾーン(Moral Crumple Zone)という概念で、人間と非人間の複数の関係アクターが係る事件について統御とアカウンタビリティとしての不一致が(主に人間が責任をとるという形で吸収されるように)生じる現象を描きだしたように、さまざまな論者に共有されているものである。

我が国においても、製造物責任に関する訴訟においてはしばしば、対象となる事物が、それを用いる者の判断に影響したかどうかが問われる事例がある。例えば、名古屋地判平成 15 年 12 月 26 日判時 1854 号 63 頁　平成 7 年(ワ)第 4179 号、同 8 年(ワ)第 1423 号は、航空機のオートパイロットモードに人間の操縦がオーバーライドする際の設計がパイロットに影響をもたらしたかどうかがとわれた。もっとも、本件では中華航空機のパイロットの重過失が認められ、またパイロットが誤った操作をするような設計をエアバス社がしたかどうかという人間の行為が問われたのであり、稲谷[2017]が指摘するように「技術的人工物の存在態様そのものがより良いものに生まれ変わるという形での、いわば事物の存在態様そのものへの責任追及」は直接的な関心ではなく、結局はそのような設計をした人間の責任が問われている。

フェルベークやラトゥールといった現代の技術哲学の論者は、主体としての人間と、客体としての機械を対極におくのではなく、むしろ事物も道徳的役割を負っているのであり、人間との相互浸透作用が人間の判断を導き出していることを提唱している。(ラトゥール[2007]、フェルベーク[2015])たしかに、人間と非-人間の相互浸透作用により事象が生じるということは、AI のような自律的に動作するものを考慮するとより明確に問題になってくるように思える。たとえば、AI を用いて何らかの取引を行う場合、人は AI を統御して意思決定を下すというよりは、AI の判断により意思が形成され取引を行っていると考えた方が適切であると考えやすい。

このような問題意識に基づいて、稲谷[2017]は刑事法において「技術の道徳性」を前提とした新たな刑事法規制のありかたとして、限界効用の観点から事物に働きかけるほうが有利な場合において刑罰を人に課すのではなく、むしろ刑事手続きを通じた企業の構造改革手続きを応用し、事物の道徳性に配慮した設計手続きを履践することを求めるという手法を提案する。

では、市民間の秩序形成を対象とし、特に強制力が金銭による損害賠償という形で実現する民事法部門において、このような問題意識に基づいた何らかの対処は可能だろうか。この点について、米国の代理法研究において AI をコモンロー上の代理人として扱うことで、AI の役割を現在の法制度における責任論に回収するアイデアが議論されており、参考になるように思われる。

そこで本報告では、米国における代理法を AI に応用する研究において何を問題として議論がなされ、どういった手法で応用するのか。

1.2 米国の制度的背景

本論に移る前に、米国においては「電子的代理人」という概念がすでにいくつかの法令や二次的法源において定義されていることに言及する必要があるだろう。というのも、AI をコモンロー上の代理人と捉える研究は、これらの法令等における概念に対する批判という形で生じているためである。

米国では従前より、人間の取引を媒介するように振る舞うプログラムについて統一電子商取引法(UETA)や統一電子署名法(E-Sign 法)、ならびに統一商事法典(UCC)といった法規範に

連絡先:産業技術総合研究所,〒135-0064 東京都江東区青海 2-4-7, 080-3306-5993, r-akasaka@aist.go.jp

おいて、電子的代理人という概念を定義してきた。これらの法規範において電子的代理人は意思伝達のための道具として定義されている。 たとえば UCC2−103.1(g)では「個人による再確認または行為なしに、全部または一部において、独自の作動を開始するために、もしくは電子的記録または電子的作用に反応するために用いられるコンピュータ・プログラムまたは電子的もしくは他の自動的手段を意味する」とされ、契約に対する意思について、「プログラミングおよび機械の利用」に求められている(コメント 4,コメント 5)。

また、本報告の対象とするコモンロー上の代理法については、2005 年の第三次代理法リステイトメントが最近の二次的法源として有力である。ここでは、代理を「ある者(本人)が他の者(代理人)に対して、代理人が本人の支配・監督のもと、本人のために行為することへの(本人による)容認を表明し、かつ、(代理人の方でも)代理人がそのように行為することの容認、もしくは何らかの形で同意を表明するときに生ずる信任関係」(§1.01)であるとし、ここでいう者について§1.04(5)にて「(a)自然人、(b)法律上権利を有し義務を負うことのできる組織もしくは団体(c)政府、政府より小さな自治体、もしくは政府によって創設された」ものとして定義されている。この定義につきコメント e において「本人か代理人として行為できるためには、法的権利と法的義務の対象となる能力がある、人(person)でなければならない」として「無生物や人ではない動物は本人やコモンロー上の代理人にはなることができない。」としている。さらにリポーターズノートにおいて、コンピュータは法的な人ではないから代理人にはなれない。

上記でみたように、各法規範では AI が代理人として扱われる余地はない。しかしながら、UETA の§2コメント5では、「人工知能の発展を通じて、コンピュータは『経験から学び、自らのプログラムにおける指図を変更し、かつ、あらたな指図を考案する』ことが可能となるだろう」「そのような発展が生じるとしたら、裁判所は、そのような新たな能力を認識するたえに、それに応じた電子的代理人の定義を解釈しうるであろう」と電子的代理人概念が今後拡張されうることは示唆されている。

2. 米国代理法研究の問題意識

それでは、米国法の代理法研究の研究者は何を問題とし、AI をコモンロー上の代理人として解釈すべきとしているのだろうか。

2.1 関係性に着目するもの

比較的初期の研究においては、使用者とAIの関係性に注目したものがみられる。例えば、Smed[1998]は、利用者と知的なソフトウェアの関係をプリンシパル-エージェントの関係ととらえ、代理法はそのような関係における責任の問題を解決するために最適な法だとしている。

Middlebrook,& Muller[2000; 354]も、E-SIGN 法や UETA といった立法は、個人とその代わりに行為を行うボットとの間の関係性を説明していないとして、コモンロー上の代理法に言及する。このコモンロー上の代理法は、イノベーションと経済的成長を社会が育成するための道具であるとの理由から、特に電子商取引におけるソフトウェアロボットについて適切であるとしている。(なお、これのように代理法に経済的な価値を見出すのは、代理法をビジネス法の基礎におく、きわめてアメリカ代理法の特色が強い見解であるといえる。)

2.2 意思形成に着目するもの

上記の関係性にたいする着目の発展形として、契約の締結にともなう意思の形成および意思の表示に AI が果たす役割の大きさを問題とする説がある。この見解は、人間の自由意志と自律性に対する非-人間の役割を重要視したものである。

このような主張の代表的なものとして Chopra & White[2011]がある。AI を単なる道具とするアプローチ(mere tool approach; 以降、単なる道具アプローチ)は、利用者との相互作用において、特定の契約に関して先存する意思があるとは言えないとして批判する。たとえば、ショッピングサイトのように中程度の複雑なエージェントを含む場面において、利用者と「相互作用」される特定の契約に関して、本人は事前の「意思」をもてないとする[Chopra & White, 2011: 36]。自律的な AI は「単なるコミュニケーションツール」ではなく、それ自体が本人の意思を構成する手段となっていることを指摘する。更に、そのような AI こそが本人や人間の代理人よりも契約締結における情報の最も良いソースであり、本人の再確認なしに自律的に契約条項を変化させるようにプログラムさせることができるとする。このことは特定の契約について意思を形成するとともに、意思を表明していると指摘する[Chopra & White, 2011; 38-39]。

パガロ[2013,2018]はこの見解を継承して、単なる道具アプローチを批判する。単なる道具アプローチは、自律的なロボットなどに意思決定を左右するような知識の獲得といった認知作業を委任する可能性がること、本人がロボットに対して契約の締結を委任したという事実から法律効果が必ず本人に発生することはならないこと、このアプローチは責任及びリスクがロボットの行為の本人としての管理者と使用者との間で配分されうる場合に対処できないことを問題として、代理法の適用を肯定的に評価する。

最近の研究では Scholz[2017]は、自動化された代理人は、申込者が意識的に検討しているかもしれないししていないかもしれない条件の組み合わせで、申込者の代わりに申込を行うことに着目する。ここで問題は、当事者が、実際に AI が選択した契約に同意することと、価格や契約条件を選択する AI を利用することに対して明示的な意思を示すことは同じだといえるだろうかということである。この問題について、申込者が設定したパラメーターの中での行為なら、本人としての申込者は当該契約に拘束されるが、申込者がブラックボックスである AI を用いた契約に同意していると言える程度では、単に幻想的な約束であるとしている。そしてその幻想的な約束については AI を法的な人格のない代理人や擬似的な代理人として想定し、具体的には追認の法理による代理関係を援用することを支持する。(Scholz[2018]; 166-167)

なお、米国以外の論者においても、このような見解を示すものがいる。例えば Sartor[2009]は電子的な代理人は 利用者やプログラマーにより用意された契約を伝達しているだけだという考えは電子的な代理人を利用する理由そのものと矛盾するものであるから拒否すべきだとする。利用者やプログラマーは知識の獲得やその熟慮に電子的な代理人に認知的な課題を委任することが電子的な代理人を利用する理由であるためである。利用者の意思が最終的には契約の効果を正当化することが事実だとしても、これは、電子的な代理人に取引の準備および実行に必要な認知過程を実行し、利用者の名の下で特定の取引を締結するというタスクを委ねる、という意図であるとする見解を示している。

2.3 その他の説

上記のほかに、Chopra & White[2011]は、法と経済学における、所謂「最安価危険回避者」の観点からも、単なる道具アプローチを批判する。すなわち、現実には利用者とAIの管理者、そしてAI自身というアクターが関わっている状況において、利用者がかならずしも最安価危険回避者ではないことはあり、その場合すべてのリスクが利用者に配分される単なる道具アプローチでは非効率的だとする。

Scholz[2017]は、米国の契約法の契約成立要件である約因にも問題があるとする。すなわち、約因とは交渉によって約束される動機、法的行為に従事する合意としての合理的な利益と損害であり、申込者と承諾者でなくAIが交渉を行っている場合、合意のために交渉を行っているとは言いづらいとする。(Scholz, [2017]; 156)

また、契約は市場評価額に縛られず嗜好を表現できる場であるものの、AIは定義上、規約、文脈、および人間のプログラマーの限界に縛られている点や、ブラックボックスの AI は価格差別を可能にし、市場で力を持つ企業が多くの余剰を取ることを可能にする契約状況を可能にすることといった、市場取引全体についても問題が生じる点も指摘する。

3. 代理人としてのAIの実現と限界

以上で見るに、代理人として AI をとらえる説は、人間と AI との関係の中で、人間の意思形成を代替ないし強く影響を与える問題への対処法として考案されていることがわかる。しかしながら、このような考え方に疑問がないわけではない。ではこのような案を、疑問を乗り越えて実際に法制度上で実現するためにはどのようにすればよいのだろうか

このような疑問点の第一には、具体的に人工物である AI が責任を果たす手立てを持ち得るかという問題がある。たとえば木村[2018]は、AI に責任を負わすために法人格を与えることは、そもそも人が責任を負担できる前提を欠いているとして批判する。同じように、Elvy[2016]も、AI は法的に人ではないため代理人としての責任を全うできないとして AI を代理人として解釈するアイデアを批判している。

このような 批判に対して、歴史的に人間以外で限定的な財産を持ち得た主体のアナロジーからの応答がなされている。例えば、Chopra＆White[2011;42]は、ローマ法における奴隷のような法的な媒介者（intermediary）という概念をもちだす。ここでの目的は AI に法的人格をもたせることではなく特定の状況下で仲介者を介して相互作用する人々が責任を免れられるようにする、より洗練された適切な法的メカニズムの発達であるとし、現代のコモンローとの「結婚」として、本人を法的に拘束できる能力のある知的な非人間のアクターとして奴隷とちかいものは子供であるとし、子供のメタファーを用いる。

他方、でパガロ[2013＝2018; 121]も奴隷のアナロジーを持ち出すが、ここではその「特有財産」に着目する。これは「自らが所有する奴隷による事業や商業活動によって損害を被ることを避けたいと考える家長たちの要求と、奴隷と安全に取引を行うことを希望する奴隷の相手方の保護というバランスを図ることを目的」とした制度で、「家長の法的責任は奴隷が保有する特有財産の価値を限度とするが、後者に関する法的な取引の安全性の保証を義務付ける」ものであった。これを応用し、個人の法的責任を AI のポートフォリオ価値の範囲内に限定させようとする。

近時の研究によれば、民法の文脈においては(すなわち、憲法的な人格権や道徳的な意味での人格を考慮せず、民法 417 条および 722 条に基づく金銭的損害賠償を行うという点についてのみいえば)そもそもAIに法的な人格を付与することは、現実的に可能な法技術であるといえる。例えば AI に法人格を付与することは既存の法人制度を利用することでも可能であり、これにより責任の分配が適切に行われうることは示されている[斉藤、2017]。Chopra＆White[2011]が指

摘するように、このように法的人格を付与するメリットは1、契約上の自由と契約の締結に関する法理論への最小限の影響で、電子契約代理人によって制定されたまたは宣告された宣言と契約の有効性問題を解決する。AI は問題の契約を締結する意思のある人として理解されることになり、標準的な意思の客観理論がかれらの行動の解釈に適用されることになる2、契約の締結において代理人がその現実の代理権や表見的代理を超えて行動した場合、本人は契約に拘束されなくなり、第三者はいわゆる代理権担保（代理人は本人の取引をていけつする権限があるという第三者への代理人の約束）の侵害で代理人を訴えることによる追加的保護をうけることができるようになる。代理人が人ではない場合には、そのような法的行為は開かれていない。3、代理人の行動のための本人の潜在的な道徳的および法的責任を制限するから、第三者への代理人の潜在的な責任を考慮することを通じ、潜在的に、人工代理人をより魅力的なものにする。といったものがある。

もう一点の疑問は、米国法における代理法と日本法の差異にある。米国の代理法の特徴として1，本人のための代理制度 (本人中心主義) 2，信任関係としての代理制度（フィデュシャリー(fiduciary)としての代理人） 3,本人とされるリスク（支配には代理法上の責任が伴う）があるとされている[樋口,2002; 12]。その中で、特に前者 2 点について、直截に米国の議論を日本に適用できないである問題があるように思われる。例えば、本人のための代理制度という点でみると、米国代理法は代理人の不法行為までもカバーする。すなわち、日本の制度が法律行為に限定されているのに対し、米国ではプリンシパルーエージェント関係にある、より広い自体において代理法が適用されている。また信任関係という関係は日本法にはない概念であり、代理関係を如何に規定すべきなのかについて差がある。また、日本には信任関係という関係性は法的に定義されているものではなく、日本では代理関係は契約関係と解される。しかしながら岩井[2016]によれば、信任関係は倫理的側面が強調され、むしろ契約関係と対極にある関係だとされている。

このような疑問点について、たしかに大陸法と英米法間の隔たりが大きく、それが原因で米国代理法研究が我が国であまり盛んに行われてこなかった[砥中、1988; 328]。しかしながら、ラトゥールやフェルベークの指摘する人間と非ー人間の相互浸透作用という人間の行為の理解は法制度を超えて共通であり、また意思を問題とする文脈に限って言えば日米双方の法制度においても、少なくとも契約において意思の形成と表示が契約の締結にとって重要なファクターであることは共通しているため、法解釈論として参照すべき論点ではないかと考える。ただし、米国のように不法行為まで含む一体的な解釈が可能になったとすると、今度は不法行為の機能としての報復・制裁機能を代理人としての AI にも想定することとなりまた困難を伴う。

4. 最後に

以上で見てきたように、一定の限界がありつつも、AI を法的な代理人として扱うという見解は、現代的な科学観に基づいた問題が顕在化する AI と人間の関係において一定の意義があるように思われる。もっとも、このような説自体もまだ米国において

主流派の学説であるといえるほど主張されているわけでなく、今後の議論の展開を負う必要があるだろう。

5. 謝辞

この成果は国立研究開発法人新エネルギー・産業技術総合開発機構(NEDO)の委託事業の結果得られたものです。

参考文献

[Smed 1998] Smed, S.: Intelligent Software Agents and Agency Law, Santa Clara Computer & High Tech. L.J., 14 ,1998.

[Scholz 2017] Scholz L.H.: Algorithmic Contracts, 20, Stan.Tech. L. Rev. 128, 2017.

[Chopra & White 2011] Chopra, S., White, L. F.: A Legal Theory for Autonomous Artificial Agents, University of Michigan Press, 2011.

[Elish 2016] Elish, M. C. : Moral Crumple Zone: Cautionary Tales in Human-Robot Interaction, We Robot 2016 Working paper, 2016.

[Middlebrook & Muller 2000] Middlebrook, S. T. , Muller, J.: Thoughts on Bots: The Emerging Law of Electronic Agents, The Business Lawyer, Vol. 56, No. 1 , 2000.

[Sartor 2009] Sartor, G.: Cognitive automata and the law: electronic contracting and the intentionality of software agents, Artificial intelligence and law 17.4, 2009.

[Elvy 2016] Elvy, S.A.: Contracting in the Age of the Internet of Things: Article 2 of the UCC and Beyond, 44 HOFSTRA L.REV. 839, 2016.

[パガロ 2018] ウゴ・パガロ(新保史生監訳): ロボット法,勁草書房, 2018.

[フェルベーク 2015]ピーター＝ポール・フェルベーク（鈴木俊洋訳）: 技術の道徳化, 法政大学出版局, 2015.

［ラトゥール 2008］ブルーノ・ラトゥール: 科学論の実在-パンドラの希望-, 産業図書, 2008.

[稲谷 2017] 稲谷龍彦: 技術の道徳化と刑事法規制, アーキテクチャと法 -法学のアーキテクチュアルな転回?-, 弘文堂, 2017.

[岩井 2016] 岩井克人: 信任関係の統一理論に向けて：倫理と法が重なる領域として, 経済研究, 67(2), 2016.

[角田 2018] 角田篤泰:ロボット・AI と人間性,法律時報, 90 巻, 12 号, 日本評論社, 2018.

[木村 2018] 木村真生子:AI と契約, ロボット・AIと法, 有斐閣、2018.

[斉藤 2018] 斉藤邦史:人工知能に対する法人格の付与, 情報通信学会, 35 巻, 3 号, 2018.

[砥中 1988] 砥中英夫:日本における外国法の摂取:アメリカ法, 英米法研究 3・英米法と日本法, 東京大学出版会, 1988

[樋口 2002] 樋口範雄:アメリカ代理法, 弘文堂, 2002.

4:00 PM - 4:40 PM (Wed. Jun 5, 2019 3:20 PM - 5:00 PM Room I)

[2I4-OS-15b-03] (Invited talk) Legal Issues toward the New Ecosystem Interwoven between Human and AIs

〇Satoshi Narihara[1,2] （1. Kyushu University, 2. The RIKEN Center for Advanced Intelligence Project）
Keywords: AI Network, Law, Architecture

本講演では、人と AIが織りなす新たなエコシステムを実現するための法的課題を明らかにした上で、その解決に向けたアプローチを提示することを目的とする。近年の日米欧における AIの法的問題に関する議論と規範形成を検討した上で、今後の法的課題を整理し、その解決のあり方を模索する。その結果、新たなエコシステムを実現するためには、 AIの適切な設計を促すための法制度の設計（メタデザイン）が求められることを明らかにする。

4:40 PM - 5:00 PM (Wed. Jun 5, 2019 3:20 PM - 5:00 PM Room I)

[2I05-07-4] Discussion / Conclusion

Discussion / Conclusion

©The Japanese Society for Artificial Intelligence

Organized Session | Organized Session | [OS] OS-19

[2J4-OS-19a] 「プロジェクション科学」の発展と応用(1)

岡田 浩之（玉川大学）、鈴木 宏昭（青山学院大学）

Wed. Jun 5, 2019 3:20 PM - 5:00 PM Room J (201B Medium meeting room)

[2J11-13-1] Aim of the session

3:20 PM - 3:40 PM

[2J4-OS-19a-01] Towards the mechanism for projection

〇Hiroki Yokoyama[1], Hiroyuki Okada[1], Hiroaki Suzuki[2] （1. Tamagawa University, 2. Aoyama Gakuin University）

3:40 PM - 4:00 PM

[2J4-OS-19a-02] Physics Projection

〇Naoto Iwahashi[1], Hideaki Negoro[1], Soichi Kawano[2] （1. Okayama Prefectural University, 2. Luke System）

4:00 PM - 4:20 PM

[2J4-OS-19a-03] (Invited talk) JIZAI body

〇Masahiko Inami[1] （1. The University of Tokyo）

4:20 PM - 5:00 PM

©The Japanese Society for Artificial Intelligence

3:20 PM - 3:40 PM (Wed. Jun 5, 2019 3:20 PM - 5:00 PM Room J)

[2J11-13-1] Aim of the session

Aim of the session

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

プロジェクションのメカニズムを探る：重ね描きとプレディクション

Towards the mechanism for projection: overlaying and prediction

横山裕樹 *1 Hiroki Yokoyama　岡田浩之 *2 Hiroyuki Okada　鈴木宏昭 *3 Hiroaki Suzuki

*1*2玉川大学脳科学研究所
Tamagawa University Brain Science Institute
*3青山学院大学教育人間科学部
College of Education, Psychology and Human Studies, Aoyama Gakuin University

1. 問題の背景

目の前に自分の好物のリンゴがある．するとそれは美味しそうだと感じる．そうした感覚，感情が生じると同時に，私たちには「美味しそうなリンゴがある」という認識が生じる．つまり推論の結果生まれた「美味しそう」という情報は，単に感じられる，そう考えたという，心の中の出来事だけでなく,「ある」という形で世界に実在するものとなっている．

この（感じた，考えた）⟶（ある，存在する）という関係はいかにして確立されるのだろうか．残念ながら，認知科学を含めた心の科学は，この問題に対する知見をほとんど有していない．この問題の困難さに気づいた哲学者の一部には，これをもって内的表象の不在，認知科学の不要を説くものもいる [Searle 84].

2. プロジェクションとは何か

この困難を乗り越えるためには，内的（脳内，情報処理システム内）に構成された表象を世界と結びつける心の働きが必要となる．我々はこの心の働きをプロジェクションと名付け，その研究の必要性を訴えてきた [鈴木 16, 鈴木 19]．物理世界の発する情報を元に内的に作り出された表象＝世界の意味は，プロジェクションの働きにより世界へと返されることにより，世界はその人なりの固有の意味に彩られた経験世界となる．そしてこの経験世界の中で人はまた認識と行為を繰り返す．

プロジェクションは，いくつかのタイプに分けられる．内的に構成された表象がそれを生み出した物理世界の事象とある程度まで一致した形で行われる，いわゆる投射が考えられる．目の前のものを特別な意味なしに知覚する，音の発生源の定位を行うなどは，通常環境下では投射の元となる情報の発生源＝ソースと，投射の先＝ターゲットが一致している．

一方，物理世界のソースとは異なる対象へのプロジェクションである異投射も考えられる．例えば，自分の手に対する刺激を目の前のゴムの手に感じるラバーハンド錯覚 [Botvinick 98]，音源とは異なる場所を音源としてしまう腹話術効果などは，典型的な異投射である．またフェティシズムに見られるように，長年使用してきたモノ，特別な経験の時に用いられたモノ，有名人が使用したモノには，そのモノ自体が持つ性質とは異なる性質が付与される [Bloom 11]．これも異投射の 1 つと考えることができる．また PC 上のエージェントに特別な感情を抱いたり [小野 16]，自分の身体の一部と同期するアバターなどの CG に自己との同一感を感じるなどの現象も異投射の現れと考えられる [渡邊 16].

またソースが存在しない，あるいは曖昧であるにも関わらず投射が起こることもある．これを虚投射という．盛んに研究がなされている色字共感覚では，黒い文字で印字されているものに色がついたと見え，それが数字によって異なることが明らかにされている [Cytowic 09]．統合失調における幻聴，東日本大震災後に被災地に現れる幽霊 [工藤 16]，極限状況下において現れる他者＝サードマンなども [Geiger 09]，明確なソースがないにも関わらず投射が起こる虚投射の例となるだろう．また脳への直接的な刺激を与えて脳の状態を変化させることで，虚投射が生じることも報告されている．[Arzy 06] は，てんかん治療のために埋め込まれた電極を通して，左側頭頭頂結合部に刺激を与えることで，現実には存在しない人の気配を生み出せるという報告もある [Arzy 06].

以上のことの中には，共感覚のような厳密な実験を通して得られたデータもあれば，ひとりの体験者の事後報告に基づくものもある．さらに，その体験の中にはサードマン，幽霊などのような一見オカルト風のものも含まれている．そうしたことから，これらの現象は科学の対象にならないと考える人もいるかもしれない．ただここで主張したいことは，これらの認識対象（幽霊やサードマン自体）の研究をすべきということではなく，そうした認識を生み出す心の仕組みを解明することは重要だ，ということである．つまり色字共感覚者が見る色が実際に印字されたもののどこかに存在するということ，実際に幽霊がどこかにいることの研究を勧めているのではなく，色や幽霊が見えてしまうプロジェクションの仕組みを明らかにすべきということである．

これら様々なタイプのプロジェクションの解明を通して，人も含めた生体にとって意味に彩られた経験世界がいかに作り出されるのか，そしてその中で生体はどう活動するのかを描き出すことが可能になる．

3. 認識と生成

物理学者のファイマン（Richard Feynman）は“What I cannot create, I do not understand.（私は自分に作れないものは，理解できない）”という言葉で認識と生成を一対のこととして捉える重要性を示唆している．本稿で議論しているプロジェクションによる認知のプロセスは，世界に対して予測モデルを生成し，認識と生成のプロジェクションループを回すという点で同様の問題意識を共有している．このように認識と生成

連絡先: hiroki.yokoyama@okadanet.org

を投射というプロセスを介して一対として考えることがプロジェクション科学の大きな特徴であると考える．

ここではプロジェクション科学における認識と生成の問題を統計的予測問題における生成モデルと識別モデルの対比で考える．あるデータがあったときに，そのデータを入力とした何らかの予測器を構成することが統計的予測問題である．

リンゴを見たときに「大きい」とか「小さい」とか判断したり、「美味しそう」とか「酸っぱそう」とか味を推測すること，即ちデータ ($\mathbf{x}$) に対するクラス (C_k) の条件付き確率（事後確率 $p(C_k)$）を直接データ ($\mathbf{x}$) だけからモデル化することを識別モデルと言う．大量の教師データ ($\mathbf{x}$ を入力としたときに対応するクラス (C_k)) が得られる場合には識別モデルのような仕組みでリンゴの見た時の各人様々な認識を獲得することが可能であるが，実世界で変化に富んだ環境で暮らす我々にはこのような状況は稀なことであると言わざるを得ない．

一方，生成モデルでも同様にデータに対するクラスの条件付き確率（事後確率 $p(C_k)$）をモデル化する点では識別モデルと同様だが，生成モデルではこの事後確率をベイズの定理により以下のように求める．

$$p(C_k|\mathbf{x}) = \frac{p(\mathbf{x}|C_k)p(C_k)}{p(\mathbf{x})} \tag{1}$$

この式の右辺において，$p(\mathbf{x}|C_k)$ はクラスに対するデータの条件付き分布であり，これを経験的に求めるのは難しくない．つまり，単にリンゴの画像からリンゴの形状（丸いとか小さいとか）や位置、光源などのパラメタを推定する（認識）だけでなく、生成過程をシミュレーションすることによりリンゴの画像の予測を可能にしているのである．

このような生成モデルの考え方はプロジェクションによる認知と非常に親和性があると考える．即ち生成と認識は一対のものとして考える，その認知プロセスを説明するのがプロジェクション科学であると考える．

4. 重ね描きによるプロジェクション

プロジェクションの問題を難しくするのは，その経路に物理的基盤が存在しないという事実である．外界からの刺激の受容から処理までは，物理的，生理的な基盤が存在する．視覚で言えば，受容されるのは電磁波の一種である光であり，その後の処理は脳内の各部位が担当する．一方，プロジェクションについて，それを字義通りに受け取れば，それは我々から世界へ向けてのプロセスである．このプロセスの一部は，明らかに行為である．しかし，行為が及ぶのは身体の到達可能な範囲までのことであり，視覚や聴覚のような遠感覚には及ばない．言うまでもなく．人はプロジェクタではないので，目から光線を発するわけではない．だとすると，内的に作り出された意味をいかにして世界の中に実在させられるのだろうか．

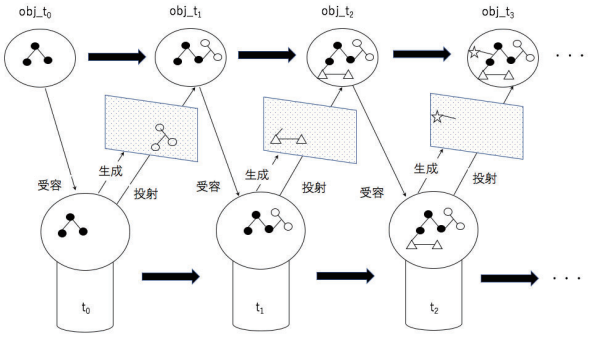


図 1: フィルタの漸次的更新による認識の変容

この難問解決の候補を考える際に，大森が述べる「重ね描き」は大きな意味を持つ [大森 82]．重ね描きについて，大森は概略以下のように論じる．赤いメガネをかければ，世界は赤く見えるのと同様に，脳に何らかの異常が生じれば，世界はその異常を反映した形で知覚される，と *1．つまり大森は，赤メガネ，脳も含めた身体を前景として，世界という遠景を透視するという図式を提案している．

これを本論文の主旨から解釈すれば，前景はフィルタであり，遠景は世界ということになるだろう．つまり，人は推論も含めた知覚システムによってフィルタを生成しながら，世界を知覚したり，行為を通して世界に働きかけたりするというサイクルを回しながら認知を行っている．このサイクルの中で，ただのリンゴは赤くて艶やかなリンゴとなり，おいそうなリンゴとなり，(たとえば）アップルパイの材料にしたいものとして知覚されることになる（図 1 参照）．こうした重ね描きという働きによって，心の世界と物理世界を結びつけること，すなわちプロジェクションが可能になるのではないだろうか．

5. 予測符号化理論による説明

こうした図式は予測符号化理論との適合性が高い．予測符号化理論は，Rao らが大脳皮質にみられる求心性，遠心性の双方向結合の役割を説明するために提案した計算モデルである [Rao 99]．初期視覚野のニューロンは，線分方位や局所的な動きといった情報を表現し，これらを組み合わせるとどのような網膜像が得られるかを予測する．この予測は，初めは正確ではないため，実際の網膜像との間に予測誤差が生じる．この予測誤差を減少させるためにニューロン活動を修正する．同様に，高次の領野は初期領野からの入力を良く予測するように，より抽象的な概念を表現する．下位概念を x，上位概念を y と表すと，各領野は上位概念が与えられた時に下位概念がどのように生成され得るかを $p(x|y)$ として表現していることになる．感覚入力の生成則を表現しているため，予測符号化モデルは生成モデルの一種である．$p(y|x)$ を表現し，感覚入力から抽象概念へと一方向かつボトムアップに計算して行く識別モデルとは，考え方が逆となる．

例えば，高次領野で，目の前の物体がりんごであることを y_2 に表現しているとする．この領野は，これまでの経験から，りんごが円形で赤いことが $p(y_1|y_2)$ に表現されており，これに基づいてりんごの画像を構成する曲線や色の情報が y_1 に表現されていることを予測する．そして初期領野は曲線や色の情報に基づいて網膜像を予測する．もし実際の網膜像が予測よりも黄色を帯びていて凹凸があれば，初期領野は x の予測誤差を減少させるために y_1 を修正し，高次領野はそれに伴って生じる y_1 の予測誤差を減少させるために y_2 を修正する．この場合は，物体がりんごではなく，みかんであるということが表現される．さらに，味覚との関係が $p(y_3|y_1, y_2)$ として表現されていれば，y_3 が予測され，(赤くて艶のあるりんごなので）「美味しそうなりんごがある」と感じることができる．このよ

*1 大森は表象を認めないどころか，主観–客観の区別を否定するという立場から論を進めており，本論文のような物理世界と心理的世界（表象）を分けて考える立場の対極に位置する．ただし，2 つの世界を一体化するという意味では，大森と本論文は問題意識を共有している．大森は主客が一体化していることを世界，人の在り方として存在論的に描き出した．一方，本論文はそれを認識論的に描き出すことを目指している．

うな予測と誤差のやりとりを繰り返し行うことによって，感覚入力として得られた x を最も良く説明する（$p(y_1, y_2, \cdots|x)$ を最大にする）$y_1, y_2, \cdots$ を求めようとする過程で，りんごに様々な意味が重ね描きされる．

色や形，単語といった静的な情報だけでなく，それに対する行動やその結果といった時系列の情報も重ね合わせの対象となり得る．Rawlik らは，エージェントの行動と報酬を折り込んだ生成モデルを提案し，強化学習を教師なし学習の一種として記述している [Rawlik 12]．また，深層ニューラルネットワークによる生成モデルの学習を実現した GAN（Generative Adversarial Networks, [Goodfellow 14]）は，データセットを再現しながら，そこには含まれていない新たなパターンを生成することができる．この GAN のように，複雑でありながら高い汎化性能を持った生成モデルは，自己の行動やそれに伴う報酬を含む経験を学習することで，新たな行動を生成し，その結果を予測することも可能であろう．そのため，目の前のりんごに対して想起した，りんごを剥くという動作，りんごを食べている恋人といった情報から,「このりんごを彼氏のために剥いてあげたい」などという考えが浮かび，一つの物体に対して様々な意味が重ね描きされていくと考えられる．

6. おわりに

本報告では，プロジェクションがなぜ必要なのかを述べ，プロジェクションのタイプ分けをまず行った．次に，プロジェクションがいかなる計算過程を経てなされるのかを，生成モデル，大森の「重ね描き」，予測符号化理論をベースにして考察した．これらはまだ提案の段階であり，その実装，シミュレーションは今後の課題となる．

プロジェクション科学は，情報の受容から表象の構成の段階に留まっていた心の科学を，世界へと結びつけ，人（あるいは種）に固有な意味世界を描き出す新しいパラダイムである．このため，これまでの知見の再検討，再解釈に加えて，心の科学とは疎遠であった様々な分野（社会関連諸科学，経営学，臨床諸科学）との共同が必要となる．またこの展開の過程では，意識，クオリア，志向性，自己などの哲学上の難問への挑戦も必要となるだろう．

謝辞

本研究は，青山学院大学綜合研究所「投射の科学」ユニットからの助成を受けて行われた．

参考文献

[Arzy 06] Arzy, S., Seeck, M., Spinelli, L., Ortigue, S., and Blanke, O.: Induction of an illusory shadow person, *Philosophical Transaction of the Royal Society B*, Vol. 443, p. 287 (2006)

[Bloom 11] Bloom, P.: *How Pleasure Works: The New science of Why We Like What We Like*, Vintage Books (2011), (小松淳子 (2012).『喜びはどれほど深い：心の根源にあるもの』インターシフト)

[Botvinick 98] Botvinick, M. and Cohen, J.: Rubber hands feel touch that eyes see, *Nature*, Vol. 391, p. 756 (1998)

[Cytowic 09] Cytowic, R. E. and Eagleman, D.: *Wednesday is Indigo Blue: Discovering the Brain of Synesthesia*, MIT Press, Cambridge, MA (2009), (山下篤子（訳）(2010).『脳の中の万華鏡：共感覚のめくるめく世界』河出書房新社．)

[Geiger 09] Geiger, J.: *The Third Man Factor: The Secret to Survival in Extreme Environments*, Penguin (2009), (伊豆原弓（訳）(2010).『サードマン：奇跡の生還を導く人』新潮社．)

[Goodfellow 14] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y.: Generative Adversarial Networks (2014)

[工藤 16] 工藤 優花：死者たちが通う街：タクシードライバーの幽霊現象, 金菱 清（ゼミナール）（編）, 呼び覚まされる霊性の震災学：3・11 生と死の狭間で, 新曜社 (2016)

[小野 16] 小野 哲雄：プロジェクション・サイエンス」の視点からの認知的メカニズムのモデル論的理解, 日本認知科学会第 33 回大会発表論文集, pp. 26 – 30 (2016)

[大森 82] 大森 荘蔵：新視覚新論, 東京大学出版会 (1982)

[Rao 99] Rao, R. P. N. and Ballard, D. H.: Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects, *Nature Neuroscience*, Vol. 2, No. 1, pp. 79–87 (1999)

[Rawlik 12] Rawlik, K., Toussaint, M., and Vijayakumar, S.: On Stochastic Optimal Control and Reinforcement Learning by Approximate Inference, in *Robotics: Science and Systems* (2012)

[Searle 84] Searle, J. R.: *Minds, Brains, and Science*, BBC (1984), (土屋俊（訳）(1993).『心・脳・科学』岩波書店)

[鈴木 16] 鈴木 宏昭：プロジェクション科学の展望, 日本認知科学会第 33 回大会発表論文集, pp. 20 – 25 (2016)

[鈴木 19] 鈴木 宏昭：プロジェクション科学が目指すもの, 認知科学, Vol. 26, pp. 1 – 22 (2019)

[渡邊 16] 渡邊 翔太, 川合 伸幸：自己身体と操作対象の身体図式の一致はより強い運動主体感・身体所有感を喚起する, 日本認知科学会第 33 回大会発表論文集, pp. 31 – 40 (2016)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Physics Projection: Intelligence with Physical World

Naoto Iwahashi*1 Hideaki Negoro*1 Soichi Kawano*2

*1Okayama Prefectural Univ *2Luke System

This paper presents a new approach named *physics projection*, through which robots can learn the physical world and predict the effects of their actions actively and online. Physics projection consists of three components: a robot, physical world model, and physics engine. The process of physics projection has a double loop structure comprising (1) a learning loop of the physical world model and (2) a simulation search loop. Experiments were performed using the TurtleBot3 mobile robot and Unity graphic engine. The results clearly showed that the robot predicted the effects of its various actions under the given physical conditions and successfully executed the tasks of carrying a wine glass without dropping it and a cup filled with water without spilling. The robot could predict a catastrophic effect that could not be predicted by a human operator.

1. Introduction

Intelligence arises from the coupling of the dynamics of internal (mental) states and external (world) states. The nature, mechanisms, and functions of such coupling have long been discussed in the fields of philosophy, science, and technology [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14]. Not only low-level sensory-motor activities but high-level cognitive activities, such as language and communication, are also established in the coupling.

We have been studying the coupling of internal states of human and robot in linguistic and physical communication, which reflect the external states, where mutual beliefs play an important role [15, 16, 17, 18]. Mutual beliefs can be considered as the communication state that is used to generate and understand linguistic (utterances) and physical actions. By focusing on such characteristics, we constructed a computational model of human–robot linguistic and physical communications (Figure 1). The model has a double loop structure containing (1) a mutual-belief learning loop (outer red-line loop) and (2) a simulation search loop (inner blue-line loop). The mutual-belief learning loop is executed when the robot is acting and making observations both linguistically and physically during communication, and a mutual belief model is learned actively and online. The simulation search loop is executed as a process in the mutual-belief learning loop to search for an appropriate action by evaluating numerous candidates by using a probabilistic-inference-based simulation.

Linguistic and physical communication is based on various constraints such as physical, sensory-motor, psychological, conceptual, and experiential constraints. The human–robot linguistic-and-physical-communication model makes it possible to incorporate these constraints into the mutual-belief model and use them in the simulation search. Among all the constraint types, physical constraints might be an important basis. However, the aforementioned human–robot model cannot apply physical constraints, such as gravity, collision, and stability, directly. Nevertheless, it can statistically learn the concepts of the physical relationship between objects and use it in the simulation search. Note that no previous high-order cognitive systems can incorporate physical constraints directly [19]. In addition,

Contact: Naoto Iwahashi, iwahashi@c.oka-pu.ac.jp, www-ail.c.oka-pu.ac.jp

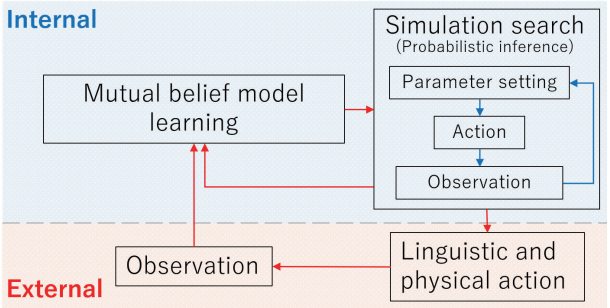


Figure 1: Linguistic and physical communication model with mutual beliefs

even without the connection of physical constraints with high-order cognitive activities, it is extremely difficult for state-of-the-art robots to predict the effects of their actions on the physical world. Therefore, robots cannot safely carry wine glasses on top of a tray or place unknown objects on a table without knocking it down.

This paper presents a new approach named *physics projection* that enables robots to learn the physical world model and predict the effects of their actions actively and online. Many previous works are closely related to this approach. The prediction of the effects of actions by physics engines was intensively studied in the fields of virtual reality, augmented reality, robotics, machine learning, and cognitive science [20, 21, 22, 23]. The comparison and integration of learned and analytic physical models were discussed in [24, 25]. The recent study of automatic 3D modeling [26] is also in this line with the current research, and could be applied to these research fields. Compared with these studies, physics projection is a novel idea, and makes the following contributions:

1. It comprises active and online loops of the learning of the physical world model and simulation search.
2. The proposed method can be integrated with high-level cognitive systems.

2. Physics Projection

2.1 Incorporating the physical world

Physics projection is a new approach for incorporating the physical world into artificial intelligence; It operates

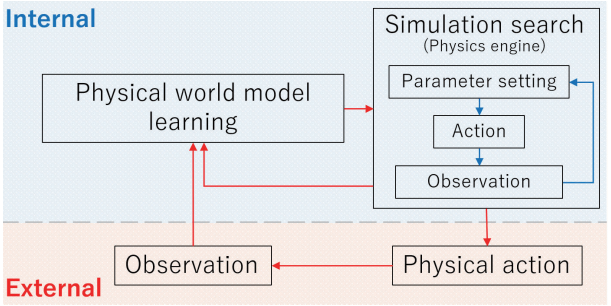


Figure 2: Physics Projection

the internal (mental) model of the external (physical) world to understand the physical world. The approach consists of three components: a robot, physical world model, and physics engine (Figure 2). The process of physics projection has a double loop structure comprising (1) a learning loop of the physical world model (outer red-line loop) and (2) a simulation search loop (inner blue-line loop). The learning loop of the physical world model is executed when the robot is performing actions and making observations in the physical world; thus, the physical world model is learned actively and onlie. The simulation search loop is executed as a process in the learning loop of the physical world model to evaluate numerous candidates and search for an appropriate action. Eventually, the loop structures of physics projection and above-mentioned human–robot liguistic-and -physical-communication model are the same, although they operate at different spatiotemporal scales and handle different conceptual granularity. Therefore, the physical constraints of physics projection can be easily incorporated into the human–robot linguistic-and-physical-communication systems.

2.2 Learning of the physical world model

The physical world model is represented by a set of physical entities with attributes, such as shape, color, mass, center of mass, inertia tensor, friction, softness, viscosity, velocity, acceleration, and gravity. These attributes are learned by robots through passive or active sensory-motor observation. Now, let us consider the robotic task carrying a wine glass on top of a tray. In this case, the attributes of the glass, floor and interaction between the robot and floor can be useful. The glass attributes include shape, mass, center of mass, and inertia tensor, whereas the floor attributes include friction, surface shape, and slope declination. The attributes of floor–robot interaction include the relationship between action-control settings and robot acceleration. However, only a few of these attributes can be observed through image processing. Although the mass, center of mass, and inertia tensor of the glass cannot be directly observed, they can be inferred through cross-modal prediction methods, such as the multimodal learning method [27]. The acceleration realized by a particular action setting cannot be obtained in a passive manner but can be observed by performing actual actions.

2.3 Simulation search

Simulations are run several times using the computational physics engine with different parameter settings of action control and the physical world model. The simulation search has two functions: physical-world adjustment, and action-control optimization. These functions are explained as follows with respect to the task of carrying a wine glass.

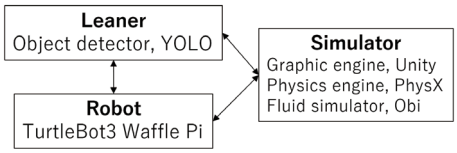


Figure 3: System implementation

2.3.1 Physical-world adjustment

Simulations are run by setting the position values of the center of mass of the glass. The simulation search determines an appropriate value of the floor friction to achieve consistency between reality and simulation of whether the glass remains on the tray or falls down. According to this value, the parameters of the physical world model are adjusted, and the adjusted physical world model is used at subsequent simulation searches.

2.3.2 Action control optimization

Here, the simulations are run by setting several different values of velocity for action control. The simulation search selects an appropriate velocity value so that the robot carries the wine glass without dropping it in a reasonable time. According to this value, the actual robot can execute the action, and this is expected to produce a desirable result.

2.4 System implementation

We implemented the physics projection system by using the TurtleBot3 Waffle Pi mobile robot, Unity graphic engine, PhysX physics engine, Obi fluid simulator, and Yolo object detector (Figure.3) .

3. Experiments

3.1 Tasks and conditions

The following three tasks were set.

STEP The robot descends one step while carrying a wine glass. The height of the step was set at 7 mm. The conditions were that the robot must not drop the wine glass; however, the wine glass may fall with the impact of descending the step. Here, the robot judges the speed at which it should descend the step.

SLOPE The robot rotates 180° on a slope with a declination of 15 % while carrying the wine glass. Here too, the robot must not drop the wine glasse; however, the glass may fall because of the centrifugal force, which changes depending on the direction of rotation because the floor is tilted. Here, the robot judges in which direction it should turn.

WATER The robot moves over an obstacle with a height and width of 7 and 120 mm, respectively, while holding a cup of water. Here, the robot must not spill water, and thus it judges the speed at which it should move over the obstacle.

The heights of the step and obstacle, and the slope inclination were measured manually, and were input into the physical-world model. The wine glass and cup were recognized by the YOLO object detection software, with a camera attached on TurtleBot3. The shape of the glass and

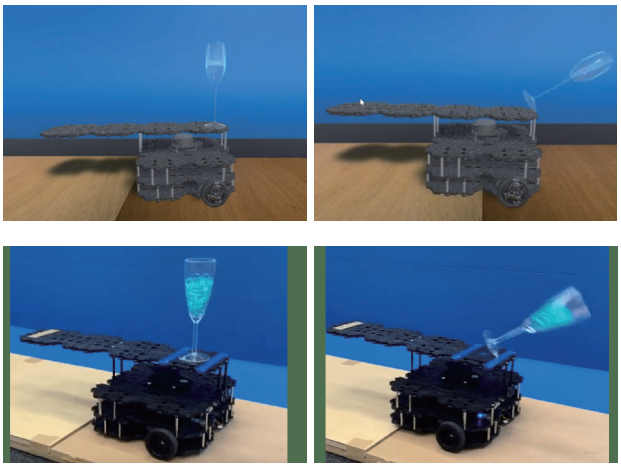


Figure 4: Resultant action effects in theSTEP task (velocity and height of center of mass). **Upper left:** Success in the prediction (15 cm/s and 13 cm). **Upper right:** Failure in the prediction (20 cm/s and 14 cm). **Bottom left:** Success in the actual task (15 cm/s and 13 cm). **Bottom right:** Failure in the actual task (20 cm/s and 14 cm).

cup, as well as the amounts of wine and water were calculated through subsequent image processing. The mass and position of the center of mass of the wine glass were inferred through a cross-modal prediction mechanism.

3.2 Results

3.2.1 STEP

Learing

In the initial setting, when the simulated and actual robots descended the step, their velocities changed in different ways. The impact of the robot's tire contacting the floor caused the actual robot to decelerate owing to motor characteristics and floor friction. Therefore, the deceleration in the simulation was set by considering consistency between the simulated and actual action effects at velocities of 10, 15, and 20 cm/s with the height of the center of mass of the glass 13 cm.

Prediction

After the learning, physics projection accurately predicted the action effects at velocities of 10–20 cm/s with heights of 12, 13, and 14 cm of the center of mass of the glass. These conditions were different from the condition under which the learning was done (*i.e.* cross-situational setting); this shows the generalization capability of physics projection. Figure 4 illustrates the predicted and actual action effects.

3.2.2 SLOPE

Prediction

The physics projection predicted that if the robot rotated in the clockwise direction, the wine glass did not fall; in contrast, with rotation in the counterclockwise direction, the glass kept falling down. The similar phenomena were observed for the actual robot, and the resultant effects were same as the predicted effects, as shown in Figure 5.

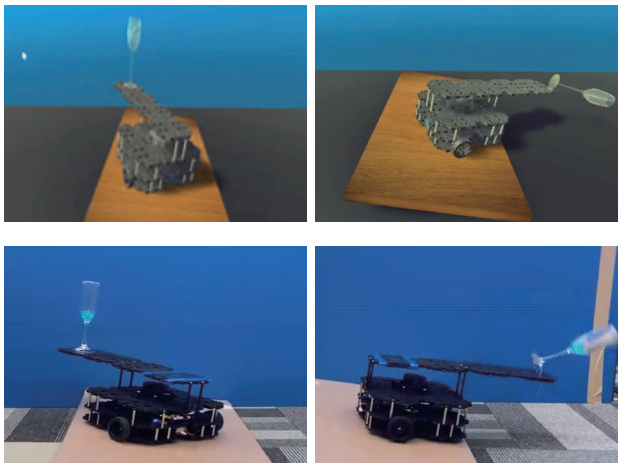


Figure 5: Resultant action effects in the SLOPE task. **Upper left:** Success in the prediction (clockwise). **Upper right:** Failure in the prediction (counterclockwise). **Bottom left:** Success in the actual task (clockwise). **Bottom right:** Failure in the actual task (counterclockwise).

3.2.3 WATER

Learning

The learning was executed for the situation that the robot descended a step with the height of 7 mm, which was different from that in WATER task. At initial setting of the liquid attributes in the simulator, the physics projection prediction was not consistent with the actual effects; the simulation robot spilled water when moving at velocities of 20 and 22 cm/s, while the actual robot spilled and retained water when moving at velocities of 20 and 22 cm/s, respectively. The liquid attributes were learned in the simulation search loop to minimize the inconsistency between reality and simulation.

Prediction

After the learning, physics projection was executed for WATER task conditions. Several situations were used with different combinations of velocities of 20, 22, and 24 cm/s and the position of cup positions of 6, 0, -6, and -12 cm. The cup position is the horizontal distance in the forward direction from the center position. The physcs projection predicted that when the position of the cup moved forward, water tended to spill. Figure 6 shows the predicted and actual action effects.

Unexpected effect

When the cup was positioned at the rear end of the robot (cup position of -12 cm), the catastrophic effect that could not be expected even by a human operator was predicted through physics projection (Figure 7). When the robot moved on the obstacle, its right front wheel, which is one of two powered wheels, floated from the floor. Therefore, the robot lost control, turned unexpectedly to the right, dropped off the path, and spilled water. We confirmed that the actual robot acted similarly.

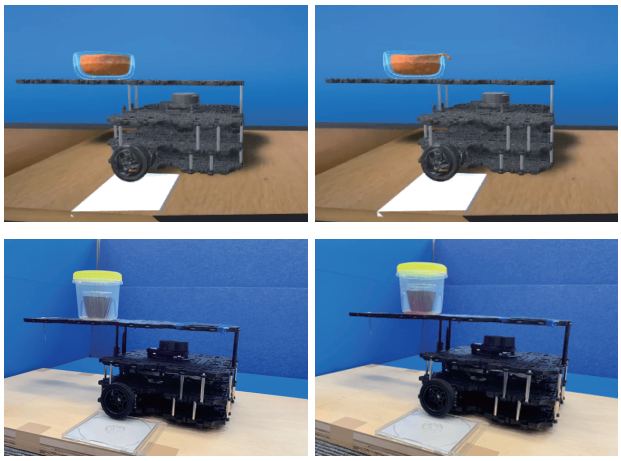


Figure 6: Resultant action effects in the WATER task (velocity and position of cup). **Upper left:** Success in the prediction (24 cm/s and 6 cm). **Upper right:** Failure in the prediction (26 cm/s and 6 cm). **Bottom left:** Success in the actual task (24 cm/s and 6 cm). **Bottom right:** Failure in the actual task (26 cm/s and 6 cm).

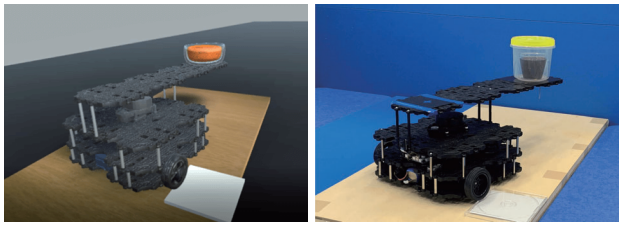


Figure 7: Unexpected action effects in the WATER task (velocity, 24 cm/s; position of cup, -12 cm). **Left:** Prediction. **Right:** Actual task.

4. Discussion

Physics projection is a very simple idea, and will attract not only technological but also scientific interests. Several research topics could be discussed, illustrated as follows:

1. When should the physical world model be learned?
2. How should the attributes be selected for learning?
3. How can incomplete physical world models work for prediction?
4. How widely should the simulation search be evaluated?
5. How would differentiable physics engines improve physics projection?
6. How can physics projection be integrated with high-order cognitive systems?

For further improvement and refinement of this approach, it is necessary to incorporate rapidly developing machine learning technology and apply recent scientific findings.

Acknowledgements

This work was partially supported by JST CREST (Grant number JPMJCR15E3, ‘‘Symbol Emergence in Robotics for Future Human-Machine Collaboration”) and JSPS KAKENHI Grant Number 18K11359.

References

[1] Martin Heidegger. *Being and Time*. 1927.

[2] Kenneth Craik. *The nature of explanation*. Cambridge University Press, 1963.

[3] Michael Polanyi. The tacit dimension. 1966.

[4] Terry Winograd. Understanding natural language. *Cognitive psychology*, 3(1):1–191, 1972.

[5] James Jerome Gibson. The ecological approach to visual perception. 1979.

[6] Humberto Maturana and Francisco Varela. Autopoiesis and cognition: the realization of the living. *Dordrecht: Reidel*, pages 2–62, 1980.

[7] David Marr. Vision: A computational investigation into the human representation and processing of visual information. mit press. *Cambridge, Massachusetts*, 1982.

[8] Dan Sperber and Deirdre Wilson. *Relevance: Communication and Cognition*. Blackwell, Oxford, 1986,1995.

[9] Stevan Harnad. The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1-3):335–346, 1990.

[10] Rodney A Brooks. Intelligence without representation. *Artificial intelligence*, 47(1-3):139–159, 1991.

[11] John R. Searle. *Mind: A Brief Introduction*. Oxford University Press, 2004.

[12] Elizabeth S Spelke and Katherine D Kinzler. Core knowledge. *Developmental Science*, 10(1):89–96, 2007.

[13] Karl Friston. The free-energy principle: a unified brain theory? *Nature reviews neuroscience*, 11(2):127, 2010.

[14] Georg Northoff. *The spontaneous brain: from the mind-body to the world-brain problem*. The MIT Press, 2018.

[15] Naoto Iwahashi. Language acquisition by robots. *Journal of Artificial Intelligence Society of Japan*, 18(1):49–58, 2003.

[16] Naoto Iwahashi. Robots that learn language: A developmental approach to situated human-robot conversations. In *Human–robot interaction*, chapter 5. I-Tech, 2007.

[17] Naoto Iwahashi et al. Robots that learn to communicate. In *AAAI Fall Symposium: Dialog with Robots*, 2010.

[18] Komei Sugiura, Naoto Iwahashi, et al. Situated spoken dialogue with robots using active learning. *Advanced Robotics*, 25(17):2207–2232, 2011.

[19] Brenden M Lake, Tomer D Ullman, Joshua B Tenenbaum, and Samuel J Gershman. Building machines that learn and think like people. *Behavioral and Brain Sciences*, 40, 2017.

[20] Peter W Battaglia et al. Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, pages 18327–18332, 2013.

[21] Lars Kunze and Michael Beetz. Envisioning the qualitative effects of robot manipulation actions using simulation-based projections. *Artificial Intelligence*, 247:352–380, 2017.

[22] Nicholas Watters et al. Visual interaction networks: Learning a physics simulator from video. In *Advances in Neural Information Processing Systems 30*, pages 4539–4547. 2017.

[23] Marc Toussaint et al. Differentiable physics and stable modes for tool-use and manipulation planning. *Proc. of Robotics Science&Sysem*, 2018.

[24] Alina Kloss et al. Combining learned and analytical models for predicting action effects from sensory data. *arXiv*, 2018.

[25] Niko Sünderhauf et al. The limits and potentials of deep learning for robotics. *The International Journal of Robotics Research*, 37(4-5):405–420, 2018.

[26] Ting-Chun Wang et al. Video-to-video synthesis. In *Advances in Neural Information Processing Systems*, pages 1152–1164, 2018.

[27] Takaya Araki et al. Online object categorization using multimodal information autonomously acquired by a mobile robot. *Advanced Robotics*, 26(17):1995–2020, 2012.

4:20 PM - 5:00 PM (Wed. Jun 5, 2019 3:20 PM - 5:00 PM Room J)

[2J4-OS-19a-03] (Invited talk) JIZAI body

〇Masahiko Inami[1] （1. The University of Tokyo）

Keywords: virtual reality, robotics, wearable computing, superhuman sports

農業革命、工業革命など社会構造革命は身体観の革新を伴ってきた。情報革命を、実社会に対するバーチャル社会の確立と捉えるなら、我々はその社会構造変化に自由自在に適応可能な新たなバーチャルな身体像「自在化身体」を設計し、身体観を更新する必要があると考る。本講演ではバーチャルリアリティや AI、ロボット工学を援用しつつ身体性を編集可能とするキーテクノロジ、自在化技術について議論し、超人スポーツなどその未来像を展望する。

©The Japanese Society for Artificial Intelligence

Organized Session | Organized Session | [OS] OS-19

[2J5-OS-19b] 「プロジェクション科学」の発展と応用(2)

岡田 浩之（玉川大学）、鈴木 宏昭（青山学院大学）

Wed. Jun 5, 2019 5:20 PM - 7:00 PM Room J (201B Medium meeting room)

[2J5-OS-19b-01] Prospects of Studies on Elastic Limb(s) Illusion based on Muscletension-Elastication Correalation

〇Kenri Kodaka[1] （1. Graduate School of Design and Architecture, Nagoya-city University）

5:20 PM - 5:40 PM

[2J5-OS-19b-02] The effect of delayed visual feedback on the full body illusion

〇Itsuki Ohtsuka[1], Sotaro Shimada[1] （1. Meiji University）

5:40 PM - 6:00 PM

[2J5-OS-19b-03] Body Schema Calibration under Virtual Reality

〇Noriki Mochizuki[1], Ryo Suzuki[1], Eimei Oyama[2], Sousuke Nakamura[1] （1. Hosei University, 2. National Institute Advanced Industrial Science and Technology）

6:00 PM - 6:20 PM

[2J5-OS-19b-04] The existence of a competitor increases enthusiasm when playing video games:

〇Nobuyuki Kawai[1], Ryuzaburo Nakata[1], Kenta Kubo[1] （1. Nagoya University）

6:20 PM - 6:40 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

筋運動誘導型 Elastic Limb(s) Illusion に関する研究の展望

Prospects of Studies on Elastic Limb(s) Illusion based on Muscletension-Elastication Correlation

小鷹研理 *1
KENRI KODAKA

*1名古屋市立大学芸術工学研究科
Graduate School of Design and Architecture, Nagoya-city University

This article outlines a series of studies on elastic limb(s) illusion based on muscletension-elastication correlation and discusses how these illusion studies contributes to exploring the body-image plasticity.

1. はじめに

筆者の研究室（小鷹研究室）は, 2016 年以降, HMD を使って腕または脚が伸縮する感覚を誘発する研究を継続的にすすめている. 本稿では, この一連の研究の総称として, 「Elastic Limb(s) Illusion」あるいはその略語として ELiI という名前を用いる. このプロジェクトの中で. これまでに 4 つの異なるシステムを考案してきたが, いずれについても, 運動学的な力学作用（引っ張る/引っ張られる, 押し込む/押し込まれる）と, 見かけ上の四肢の長さの伸張との動的な関連付けをベースとし, それらの作用量を体重計の値を通して見積もる点で共通している. 後に触れるように, これらの 4 つのシステムは, 過去に様々な展示会で出展され, 錯覚の誘発に関して極めて個人差の少ない手法であることを示してきたが, 個々のシステムについて得られている効果の認知的根拠および, 各システム相互の関連について, これまでに俯瞰的に論じる機会がなかった. 本稿では, 4 つのシステムをあらためて概観し, 他の研究者によって報告されている関連研究と比較しながら, Elastic Limb(s) Illusion 研究の学術的な展望を議論したい.

2. 身体変形感研究の周辺

Elastic Limb(s) Illusion は, 身体所有感の変調に関して蓄積された古典的な知見を下敷きにしている. 以下では, ELiI と関連の深い, 身体変形感に関するここ 10 年の諸研究を概観する.

よく知られているように, 物理空間と仮想空間の両者にまたがって複数感覚に同期刺激を与えることで, 仮想空間内の身体やアバターに体験者の身体所有感（Ownership）が投射される. 本稿の関心に照らして注目すべきは, 所有感の変調が, 所有された身体のサイズ感をも引き込んで変調する点にある. 例えば, 自分とはサイズの異なるアバターに対して視触覚の同期刺激を与えると, 自分自身の身体のサイズ感がアバターと同化する方向で変形することが知られている [Hoort 2011][Normand 2011]. このサイズ感の変調は, 実際に外部空間のスケールを縮小／拡大させるとともに [Hoort 2014][Hoort 2016], 身体のサイズに紐づけられた種々の自己イメージ（例えば, 身体満足度の変調 [Preston 2014]）にまで作用する.

近年とりわけ注目されているのは, サイズ感の変調が主観的な痛みに与える影響に関するトピックである. 例えば, 特殊な鏡を使って手のサイズを拡大・縮小して呈示した 2 つの報告では [Moseley 2008][Ramachandran 2009], サイズ感の大小に対応するかたちで主観的な痛みが増減することが紹介されている（ここでは, 視触覚バイモーダル細胞によるトップダウンの関与について検討が必要である）. 他方で, CRPS（複合性局所疼痛症候群）の患者に対して指の主観的なサイズをリアルタイムに伸縮させるシステム（MIRAGE[Newport 2015]）を用いた場合, 伸張・収縮ともに痛みの感覚を低減する効果が確認されている [Preston 2011]. ここで注目すべきは, Moseley や Ramachandran らの実験系では, 身体サイズの変形感を誘発する際に, はじめからサイズの異なる手・腕を静的に呈示している一方で, MIRAGE システムでは, 通常のサイズから伸びたり縮んだりする過程も含めて動的に変形イメージを呈示している点である. すなわち, 前者は単なる「身体変形感」の誘発である一方で, 後者は「身体伸縮感」を内包した「身体変形感」の誘発であると区別することができる.

CRPS の患者は, 主観的な身体イメージのサイズ感に齟齬を抱えていることが報告されており [Lewis 2007], したがって, 特定の身体マップを表象する神経細胞群の一部が, 何らかの原因により不適当な学習のスパイラル（不使用の学習 : Learned-non-use）に置かれ, 物理的状態を正確に更新できない凍結状態に陥っていると考えられる. MIRAGE のようなシステムで, 連続的にサイズが増減する（つまり様々なスケールを内包する）視覚イメージを与えることは, 問題のある神経細胞群を全般的に再活性化させ, 身体マップを再組織化することにつながる可能性がある.

3. Elastic Limb(s) Illusion

3.1 錯覚原理

筆者らが提案する Elastic Limb(s) Illusion は, 一人または複数で成立する特定の運動形式を通して, 体験者の手脚に伸張または収縮の力学的作用を与え, その力の大きさを HMD 空間内の手脚の伸張度と連動させるものであり（muscletension-elastication correlation）, MIRAGE で前提とされている錯覚原理を, 身体全体に拡張して適用したものと整理することができる. また, 全てのシステムに共通する手法的な新規性として, 腕や脚に具体的に作用している力を, 体重計からリアルタイムに取り出す点を挙げることができる（weight-based muscletenstion estimation）. ただし, 以下に示すように, この体重計の実際の使われ方は, システムによって大きく異なる. 以下に, 図 2 を参考にしながら, 小鷹研究室がこれまでに考案した 4 つのシステムの概略を示す.

連絡先: 464-0083 愛知県名古屋市千種区北千種 2-1-10, kenrikodaka@gmail.com

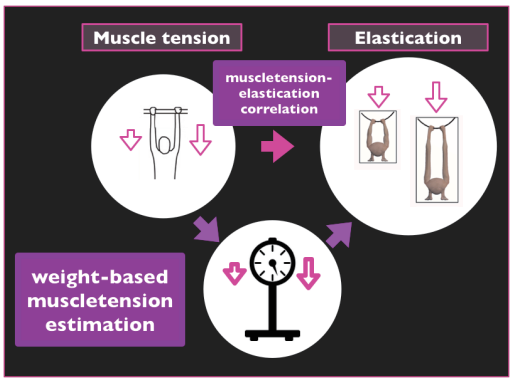


図 1: Principle of elastic limbs illusion

3.2 システム概略

『Underground Diver』[曽我部 2016] は, HMD を着用した体験者が懸垂の要領で鉄棒にぶら下がることによって, 両腕に対して, ぶら下がりに対する反作用の伸張力を上向きに発生させる（下方ぶらさがり）. このとき, 腕にかかる力は, 体験者の足元に置かれた体重計の減少量（自重比）によって見積もることができる. なお, このシステムのみ, HMD の視点として, アバターの頭部をやや後方から観測する Full Body Illusion 視点を採用している.

『Stretchar(m)』[Kodaka&Mori 2017] は, HMD を着用した体験者が両手でハンドルを握った状態で後方に体重を傾け, 相対する実験者がそのハンドルを支えることで（後方ぶら下がり）, やはり体験者の両腕に対して伸張力を発生させる. このとき, 腕にかかる力は, 体験者の足元に置かれた体重計の増加量（自重比）によって見積もられる（つまり, この過程で, 実験者の体重が体験者の側へと部分的に移動する）.

『Elastic Arm Illusion』[森・小鷹 2018] では, 仰向けに寝転がった体験者の片手を, 実験者が, やはりハンドルを介して上方へと引っ張り上げる（上方引っ張られ）. 体重計は, 実験者の側の足元に置かれ, この体重計の増加量によって腕にかかる力が推定される. つまりこのシステムでは, 体験者とは異なる他人の身体状態の変化を通して体験者の身体状態が推定されている. この意味で, 4 つの ELiI のなかでも, 体験者にとって最もストレスフリーなシステム構成となっている.

『Elastic Legs Illusion』[安楽 2019] では, 長座位の姿勢で伸ばした両脚の裏をスタンドに接地させ, そのスタンドをゴムチューブを介して自身の腕で手前側に引っ張ることで, 手脚に相反する方向で力を発生させる（つっぱり運動）. この力は, つっぱった脚とスタンドの間にかませた体重計の値によって推定する. なお本稿では詳しく述べないが, このシステムは, 腕と脚に対して同時に伸縮感覚を誘発することも可能である.

3.3 錯覚の強度・個人差

以上に挙げた 4 つの錯覚誘導システムは, 国内外の展示会で公開され, これまでに多くの体験者からの反応を得ている. 『Stretchar(m)』と『Elastic Legs Illusion』については, 発表された各展示会の中で, 体験者に対して簡易的なアンケートを行っており, 「自分の腕／脚が実際より伸びるように感じた」かの問い（-3 to +3, 7 段階評価）に対して, 前者（腕）については 77% (27/35, エンターテインメント・コンピューティング [石原・森 2017]), 後者（脚）については 75%（33/44, からだは戦場だよ 2018 Δ）の体験者が, 実際よりも伸びる感覚を「強く感じた」または「大変強く感じた」と解答している（+2 or +3）. 逆に, 錯覚に対してネガティブな回答をした人は, 両者に関して, それぞれわずか 1 人であった. 他のシステムについてはアンケートを未実施であるが, 来場者の反応に大きな違いはないという印象を筆者は得ている. いずれにせよ, 以上の結果は, ELiI が, 姿勢の違い及び腕・脚の違いに関わらず, 極めて個人差の少ない高水準な錯覚を誘導できることを示すものである.

3.4 筋肉の運動学的生理

既に述べた通り, 我々の提案している一連の Elastic Limbs Illusion は, 筋骨格系と視覚系の感覚間相関を錯覚原理として適用している. ＜下方・後方ぶら下がり＞、＜引っ張り上げ＞といった運動作用は, 体験者の腕に伸張作用を発生させる. この種の物理的に「伸ばされている」という固有感覚と, 視覚的な腕の伸張度を同期させるという考え方は極めて直感的であり, また, 複数感覚間相関に基礎を置く所有感ベースの錯覚原理の考え方 [Blanke 2012] に沿っている. 他方, 『Elastic Legs Illusion』において両脚に与えられている「つっぱり運動」の筋運動の方向は, 実は, 伸張でなく収縮である. ここでは, 両脚の筋肉を収縮させる力の大きさと脚の伸張度を同期させていることになり, つまり, 物理世界における作用と仮想世界における作用があべこべとなっている. それにも関わらず, 既に示した通り, 『Elastic Legs Illusion』は, これまでの腕伸縮錯覚と同程度の錯覚強度を生み出している. この謎を紐解くための準備として, 筋肉の伸張・収縮の生理的側面を以下で確認する.

筋肉が「伸びた感じ・縮んだ感じ」は, 筋骨格系に分布する感覚受容器を通して取得されると考えられている. [伊藤 1986] によれば, 筋肉に分布する筋紡錘は, 筋が伸張することでインパルスが増大する一方, 筋肉と骨をつなぐ組織である腱付近に分布する腱器官は, 筋肉からの張力をモニタし, 結果的に筋肉が収縮されることでインパルスが増大する. これらの情報は脊髄の運動ニューロンを介して, 筋を望ましい長さに保つべく, 負のフィードバックとして即座に対象の筋肉へと差し戻される（伸張反射）. 外的に発生した緊張方向とは逆の方向の補償作用が内的に発生するという生理的知見は, 『Elastic Legs Illusion』における「あべこべ」の問題を考えるときに極めて示唆的である. というのも, 『Elastic Legs Illusion』の感覚間同期にとって, HMD 上の伸びゆく脚のイメージと同期しているのは, 外力（収縮）ではなく, 内的な補償作用（伸張）の方であるという解釈が成立するからだ. この解釈を拡大すると, 特定の筋運動に対して, 順逆いずれの視覚イメージを適用しても, 等しく伸縮感覚が誘導されることも考えられる. つまり, 「伸ばす」と「縮む」は同じ 1 つの現象の表と裏のような関係にあり, 文脈によって, 前景化される筋運動の方向が決定するという仮説が成立するだろう. この問題を詳しく検討するため, これまでに発表したものと同一の筋運動に対して, HMD 上で収縮イメージを呈示した際に, 伸張イメージと同等の伸縮感覚が得られるかどうかを検証することは, 今後の重要な課題である.

3.5 視点ドリフト型と作用点ドリフト型

4 つのシステムは, HMD 上で手または脚が実際の長さよりも伸張する点で共通しているが, 力が作用する接点となる手先・足先の HMD 空間上の位置が固定か可変かによって, 2 つのグループに分けることができる. 例えば, 『Underground Diver』は, 手先は鉄棒の位置に固定され, 頭と肩が地面へと沈んでいくことで両手が伸張する. 同様に『Elastic Legs Illusion』では, 足先がスタンドの位置に固定され, 腰が後方へとスライドしていくことで両足が伸張する. この両者は, 手先・足先が固定さ

れ, 視点（つまり上半身）が移動することで手脚が伸張するため「視点ドリフト型」に分類される. 他方で, 「作用点ドリフト型」に分類される『Stretchar(m)』『Elastic Arm Illusion』では, 視点は物理世界の頭部の位置に固定する一方で, 手先は, 力が作用する方向へ（正確には視点から手先へと向かうベクトルの延長線上で）と, 連続的にドリフトしていく. すでにアンケートでも示している通り, いずれのドリフト型でも高い水準の伸縮感覚を誘発することが可能である.

視点ドリフト型と作用点ドリフト型は, 「自分が動くことによる変形」か「端点が動くことによる変形」かの選択に関わるものであり, 同一の筋運動の事象において許容される, （相対的に同一な）視覚的伸張イメージの解釈の２つのバリエーションである. 実際, 全てのシステムは, （発表されものとは対照的な）もう一方の視覚解釈を採用することも可能である. ここで浮上する興味深い問いは, ある特定の筋運動に対して, 錯覚にとって有用な視覚解釈は一方に定まるのか否か, というものだ. 筆者ら設計者の経験的な印象としては, ある筋運動のタイプによって, もう一方の視覚解釈を採用することで錯覚強度が大きく減退することもあるが, 必ずしも明確に区別できないケースも存在するように感じている. この種の偏りがどのような機序で生じるかを検討するのも今後の課題である.

4. おわりに

Elastic Limb(s) Illusion の 4 つのシステムについて, それらのシステム相互の関連性を俯瞰的に整理し, 今後の展望を述べた. とりわけ, ELiI において, 1 つの物理的な筋運動形式に対して, HMD 上で全く異なるタイプの変形イメージが対応できる可能性を指摘した. これらの問題の実験科学的な検証は, 身体イメージの可塑性の探求において, 極めて有用な知見を提供するものと期待される.

謝辞

本研究の一部は, 中山隼雄科学技術文化財団の助成を受けた.

参考文献

[Hoort 2011] van der Hoort, B., Guterstam, A., & Ehrsson, H. H. (2011). Being Barbie: the size of one's own body determines the perceived size of the world. PloS One, 6(5), e20195.

[Normand 2011] Normand, J. M., Giannopoulos, E., Spanlang, B., & Slater, M. (2011). Multisensory stimulation can induce an illusion of larger belly size in immersive virtual reality. PLoS ONE, 6(1).

[Hoort 2014] van der Hoort, B., & Ehrsson, H. H. (2014). Body ownership affects visual perception of object size by rescaling the visual representation of external space. Attention, Perception, and Psychophysics.

[Hoort 2016] Van Der Hoort, B., & Ehrsson, H. H. (2016). Illusions of having small or large invisible bodies influence visual perception of object size. Scientific Reports.

[Preston 2014] Preston, C., & Ehrsson, H. H. (2014). Illusory changes in body size modulate body satisfaction in a way that is related to non-clinical eating disorder psychopathology. PloS One, 9(1), e85773.

[Moseley 2008] Moseley, G. L., Parsons, T. J., & Spence, C. (2008). Visual distortion of a limb modulates the pain and swelling evoked by movement. Current Biology : CB, 18(22), R1047-8.

[Ramachandran 2009] Ramachandran, V. S., Brang, D., & McGeoch, P. D. (2009). Size reduction using Mirror Visual Feedback (MVF) reduces phantom pain. Neurocase, 15(5), 35760.

[Preston 2014] Preston, C., & Ehrsson, H. H. (2014). Illusory changes in body size modulate body satisfaction in a way that is related to non-clinical eating disorder psychopathology. PloS One, 9(1), e85773. https://doi.org/10.1371/journal.pone.0085773

[Newport 2015] Newport, R., Auty, K., Carey, M., Greenfield, K., Howard, E. M., Ratcliffe, N., … Themelis, K. (2015). Give It a Tug and Feel It Grow: Extending Body Perception Through the Universal Nature of Illusory Finger Stretching. I-Perception, 6(5)

[Preston 2011] Preston, C., & Newport, R. (2011). Analgesic effects of multisensory illusions in osteoarthritis. Rheumatology (Oxford, England), 50(12), 23145.

[Lewis 2007] Lewis, J. S., Kersten, P., McCabe, C. S., McPherson, K. M., & Blake, D. R. (2007). Body perception disturbance: a contribution to pain in complex regional pain syndrome (CRPS). Pain, 133(13), 1119.

[ノーマン・ドイジ 2016] ノーマン・ドイジ, 脳はいかに治癒をもたらすか, 紀伊国屋書店, 2016

[曽我部 2016] 曽我部愛子, 森光洋, 小鷹研理：ぶら下がりによる自重変化を利用した腕が伸縮する感覚の誘発, 第 20 回情報処理学会シンポジウム・インタラクション 2016, 2016.3

[Kodaka&Mori 2017] Kodaka, K., & Mori, K. (2017). Stretchar(m) makes your arms elastic. In SIGGRAPH Asia 2017 VR Showcase, SA 2017.

[森・小鷹 2018] 森光洋, 小鷹研理：Elastic Arm Illusion, VR Creative Award 2018 (Finalist), 2018.8.25

[安楽 2019] 安楽大輝, 森光洋, 小鷹研理：Elastic Legs Illusion ― 脚を長く伸ばす体操, 第 23 回情報処理学会シンポジウム・インタラクション 2019, 2019.3

[石原・森 2017] 石原由貴, 森光洋, 室田ゆう, 小鷹研理：HMD を介したポールを引っ張り合うことによる腕が伸縮する感覚の誘発, 情報処理学会シンポジウム・エンタテインメントコンピューティング, 2017.9.16

[Blanke 2012] Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. Nature Reviews. Neuroscience, 13(8), 55671.

[伊藤 1986] 伊藤宏司：筋運動制御機構, 計測と制御, Vol.25, No.2

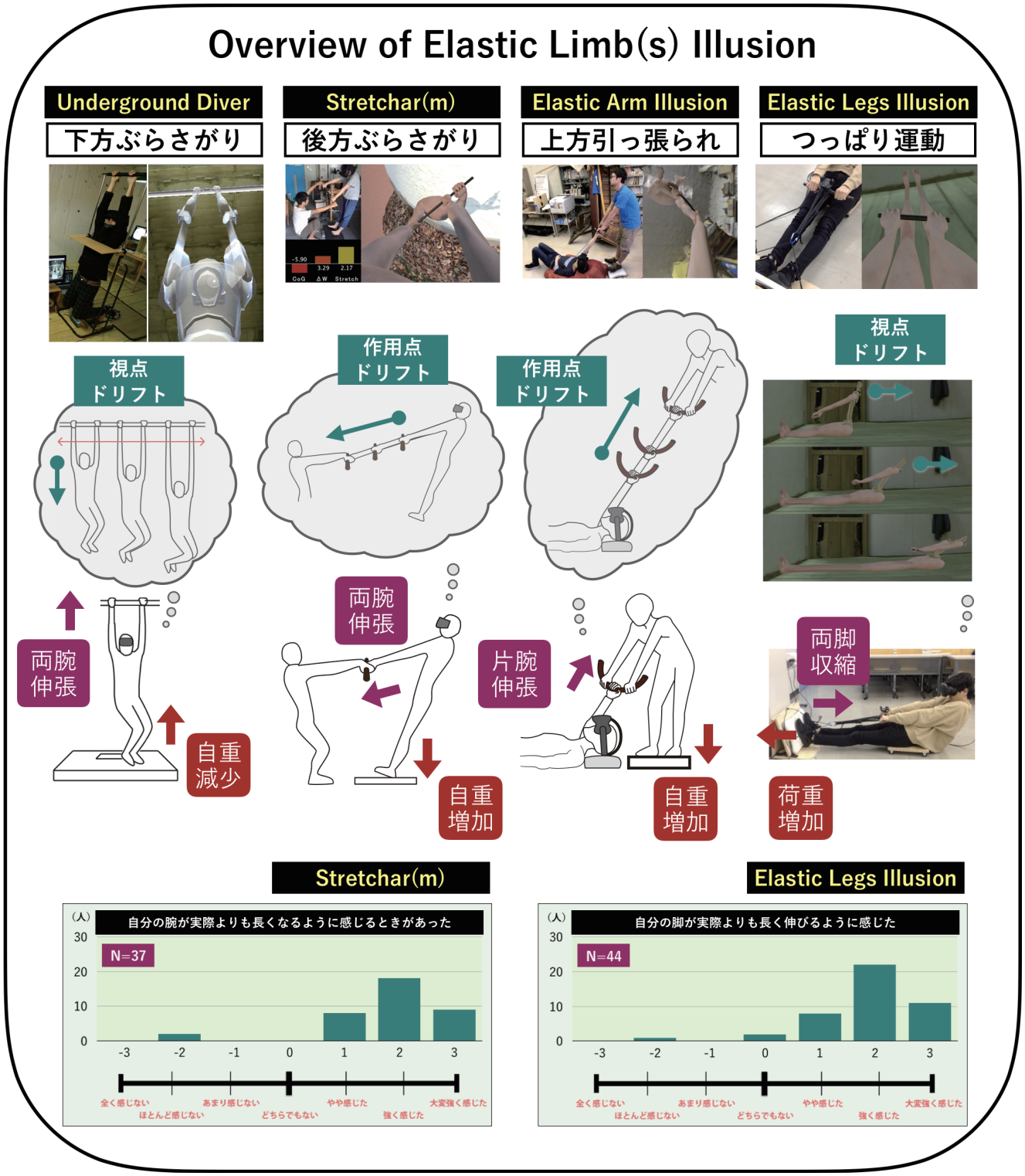


図 2: Overview of elastic limb(s) illusion

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

フルボディ錯覚における遅延視覚フィードバックの影響

The effect of delayed visual feedback on the full body illusion

大塚　一輝[*1]
Itsuki Ohtsuka

嶋田　総太郎[*2]
Sotaro Shimada

[*1] 明治大学大学院
Meiji University Graduate School

[*2] 明治大学
Meiji University

The full body illusion (FBI) is an illusion of the self-ownership and self-agency towards an avatar that moves synchronously with one's own body movement. Here we investigated the FBI using an avatar that moves with a short delay from the participant's movement. Participants made their steps on the spot while watching a VR scene as if the avatar walked in a street either synchronously (100-ms delay) or asynchronously (400- or 700-ms delay) with their steps. The avatar was presented in either 1st-person or 3rd-person perspective. The result showed that the participants felt the sense of ownership only toward the avatar that moved synchronously in the 1st-person perspectives. Our results suggest that our body image can be projected on the avatar that is presented in the 1st-person perspective and moves synchronously with one's own body.

1.　はじめに

自己身体感の最小構成要素は，身体保持感と運動主体感であるとされる[Gallagher 2000]．身体保持感とはこの身体は自分自身の身体であると感じる感覚であり，運動主体感とはこの運動を引き起こしたのは自分自身であると感じる感覚である.「プロジェクションサイエンス」の観点から考えると，自己身体認識とは脳内にある自己身体の表象を外部にある身体的な物体に投射することで自分の身体として認知することだと説明できる．このとき，自己身体の表象の投射先が自己身体以外となる例として，ラバーハンド錯覚やフルボディ錯覚が挙げられる．

ラバーハンド錯覚とは，自分の手の表象を自分の手ではない偽物の手（ラバーハンド）に投射して，ラバーハンドが自分の手のように感じられる錯覚である[Kalckert 2012]．ラバーハンド錯覚を全身に拡張した現象がフルボディ錯覚であり，自分の身体が刺激されている状態で，アバターの 1 人称視点でアバターが筆で同じ場所を刺激される映像を見ると，そのアバターに対し身体保持感を感じることが報告されている[Petkova 2011]．3 人称視点のフルボディ錯覚ではアバターの背面をアバターの後方から見る．自分の背中を刺激されるのと同時にアバターの背中が刺激されている映像を見るとアバターに対して身体保持感を感じることが報告されている[Heydrich 2013]．

本稿では，自ら動かせるアバターに対して身体保持感および運動主体感を感じる錯覚（フルボディ錯覚）について検討する．ここでは，アバターを 1 人称視点で見た場合と 3 人称視点で見た場合，アバターの動きが同期している場合と遅れている場合において錯覚の強度にどのような違いが生じるのかを調べる．これによって自己身体全体がどのように他の物体に投射されるのかについて考察する．

2.　フルボディ錯覚実験

2.1　被験者

21 名の健康な男性が実験に参加した（21.7 ±0.15 歳）．被験者には実験前に実験内容や安全性についての説明を行い，書面にて実験参加の同意を得た．

連絡先：大塚一輝，明治大学大学院理工学研究科，神奈川県川崎市多摩区東三田 1-1-1，E-mail: itsuki09120921@gmail.com

2.2　実験環境

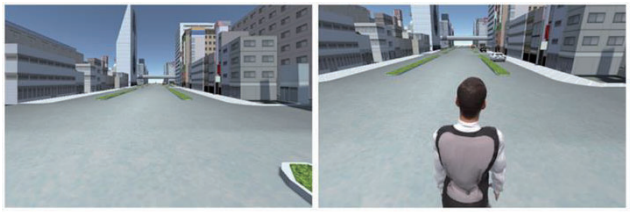
被験者はモーションキャプチャーデバイス（Perception Neuron, Noitom, CHN）とヘッドマウントディスプレイ（HMD: Head Mounted Display）（Oculus rift, Oculus, USA）を装着した．Oculus rift は視野角が約 110°，解像度が 1080×1200，リフレッシュレートが 90 Hz であった．

Perception Neuron のセンサは磁気の影響を受けやすいため磁場シールド（アンティエミー4010T 東京計器ピエーション株式会社，JPN）の中に入り，磁場の影響を極力少なくし実験を行った．また，Perception Neuron のセンサは計 12 個のセンサを用いて実験を行った．

被験者は実験開始位置に立った状態で HMD を装着した．ゲーム開発プラットホーム（Unity, Unity Technologies, USA）で仮想現実空間の映像を作成した．

2.3　実験手順

被験者にアバターが 1.0 m/s の速さで前に動いている映像を見せながら，1 分間その場での足踏みを行ってもらった（図 1）．足踏みを行うにあたり被験者には「通常歩くスピードより遅めに、一定のペースで腿をきちんと上げ、その場から動かないようにすること．首を動かせる範囲は前方 180 度以内であり、後ろを見ないこと．」と教示を行った．アバターに対する身体保持感を変調させる要因として視点（1 人称視点：1pp，3 人称視点：3pp）とアバターの動きの整合性（同期 vs 遅延 400ms vs 遅延 700ms）の二つの要因を設け，2×3 = 6 つの実験条件で実験を行った．なお同期条件でも装置の接続により 100ms の遅延があった．



A. 1人称視点　　B. 3人称視点

図 1　実験映像

本実験開始前に，足踏みによる影響をなくすために図 1 の A の前に進まない映像を見ながらその場で 1 分間足踏みを行ってもらった.足踏みによる実験開始位置からの変位が 10cm 未満となる、または測定を 3 回行ったところで終了とした.

フルボディ錯覚の強度を測る指標として以下の 2 つを用いた.

(1) 自己身体認識アンケート

主観的指標として自己身体認アンケートを行った．アンケートは先行研究[Kalckert 2012]で用いられたものを今回の実験に合わせた内容に変更し日本語に翻訳したものを用い，身体保持感と運動主体感およびそのダミー項目の全 8 項目から成る．評価には「とても思う(+3)」から「まったく思わない(-3)」までの 7 段階リッカート尺度を用いた．2 項目は身体保持感に関するもの(「アバターの身体が自分の身体であるかのように感じました」など)，2 項目は運動主体感に関するもの(「アバターの身体は，私の意思に従うように，思った通りに動きました」など)，残りの 2 項目はダミーの質問(「二つ以上の身体を持っているかのように感じました」や「アバターの身体が私の意思をコントロールしているかのように感じた」など)であった.

被験者はそれぞれの実験条件につき 1 試行ずつ行った．順序効果を避けるため 6 つの実験条件の呈示順序は被験者間でカウンターバランスを取った.

身体保持感，運動主体感それぞれ 0 に対して有意に大きなスコアであった場合(t 検定)，アバターに対して身体保持感，運動主体感を感じたと定義した.

(2) ドリフト測定

客観的指標としてドリフト測定を行った．1 分間の足踏み終了後に被験者に対し「開始位置から動いてしまったと感じたら基準位置まで戻ってください．前後左右に動くまたは回転していただいても大丈夫です．」と教示を行った．そして被験者が戻った位置を実験終了位置とし，実験開始位置からの前後方向の被験者の位置の変位を測定した(図 2)．また本実験の変位(実験値)と前に進まない条件(基準値)の変位の差をドリフト距離とし，フルボディ錯覚生起の指標として定義した.

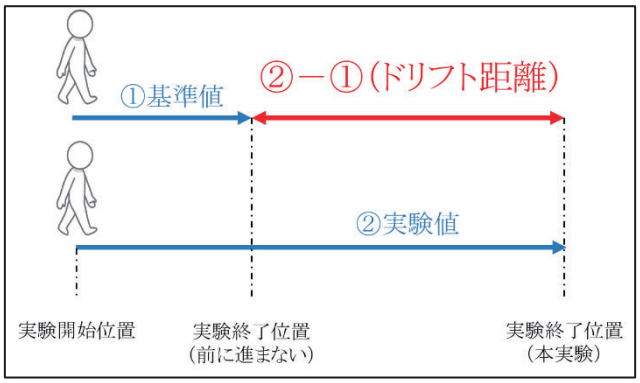


図 2 ドリフト測定

3. 実験結果

3.1 自己身体認識アンケート

身体保持感に関するアンケート結果を図 3 に示す．各条件において、0 との t 検定を行ったところ 1pp の同期条件において有意なアンケートスコアが見られた($t(20) = 2.12$, $p < 0.05$)．また、二要因に対応がある([1pp, 3pp], [同期, 遅延 400ms, 遅延 700ms])二要因の分散分析を行ったところ視点による主効果が見られた($F(1,20) = 8.28$, $p < 0.01$)．同様に遅延による主効果も見られた($F(2,40) = 13.1$, $p < 0.001$)．遅延の主効果が見られたことから多重比較(テューキーの HSD 検定)を行ったところ遅延 400ms 条件及び遅延 700ms 条件よりも同期条件の方が高いスコアを示した($HSD = 0.704$, $p < 0.05$)．

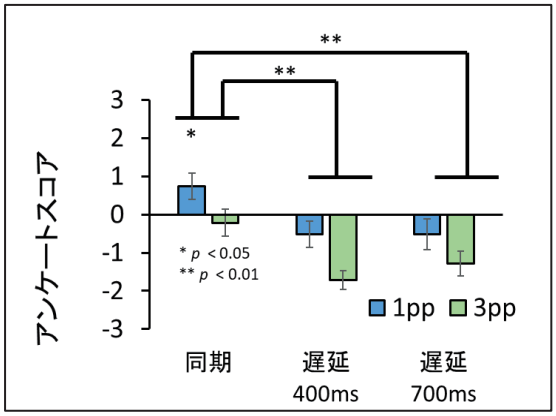


図 3 身体保持感のアンケート結果

運動主体感に関するアンケート結果を図 4 に示す．各条件において、0 との t 検定を行ったところ 1pp の同期条件($t(20) = 4.01$, $p < 0.001$)及び 3pp の同期条件($t(21) = 6.30$, $p < 0.001$)において有意なアンケートスコアが見られた．また、二要因に対応がある([1pp, 3pp], [同期, 遅延 400ms, 遅延 700ms])二要因の分散分析を行ったところ遅延による主効果が見られた($F(2,40) = 24.1$, $p < 0.001$)．遅延の主効果が見られたことから多重比較(テューキーの HSD 検定)を行ったところ遅延 400ms 条件及び遅延 700ms 条件よりも同期条件の方が高いスコアを示した($HSD = 0.786$, $p < 0.05$)．

以上の結果より身体保持感は 1pp の同期条件時にアバターに対して抱き、運動主体感は同期条件であれば抱くといえる。

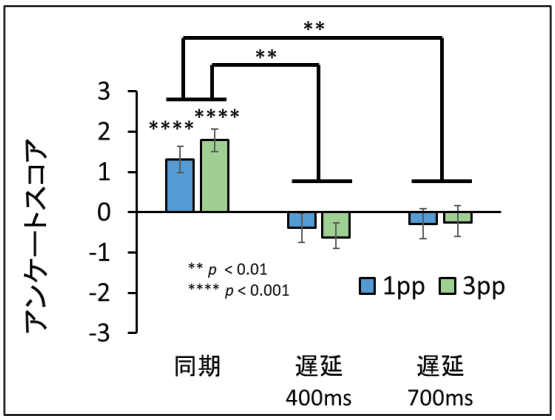


図 4 運動主体感のアンケート結果

3.2 ドリフト距離

ドリフト距離の結果を図 5 に示す．各条件において、0 との t 検定を行ったところ 1pp の遅延 700ms 条件（$t(20) = 3.08$, $p < 0.01$）、3pp の同期条件（$t(20) = 2.35$, $p < 0.05$）において有意なドリフト距離が見られた．また、二要因に対応がある（[1pp, 3pp],[同期, 遅延 400ms, 遅延 700ms]）二要因の分散分析を行ったところ交互作用が見られた($F(2, 40) = 3.44$, $p < 0.05$)。下位検定として単純主効果検定を行ったところ遅延 700ms 条件において 1pp 条件より 3pp 条件の方が前方への大きなドリフト距離を示した($F(1, 60) = 5.76$, $p < 0.05$). さらに 3pp 条件において多重比較（テューキーの HSD 検定）を行ったところ 3pp の遅延 700ms 条件と 3pp の同期条件の間に有意差が見られた（$HSD = 15.5$, $p < 0.05$）.

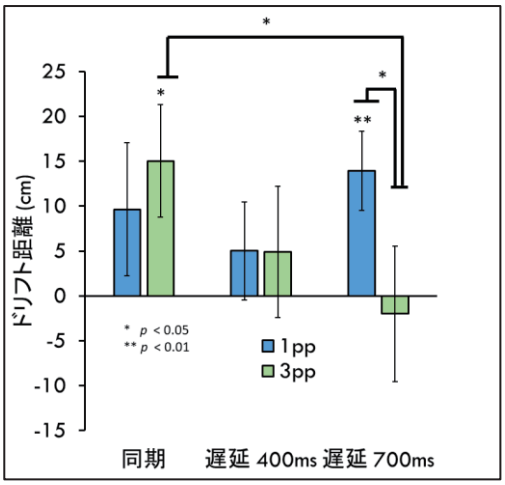


図 5 ドリフト距離の測定結果

4. 考察

今回の実験では，VR 環境において自分の身体の動きをアバターに投射し，視点とアバターの動きの違いによる自己身体感を調べた．身体保持感のアンケート結果において 1pp の同期条件では身体保持感を抱いていたが 3pp の同期条件では抱いていなかった．これより自己身体の投射は時間の整合性だけでなく投射先の自己との相対的位置が重要であると考えられる．これはラバーハンド錯覚の先行研究において,ラバーハンドと実際の手の距離を離すと身体保持感が低下するという結果からも支持される[Lloyd 2007]．また、運動主体感のアンケート結果で同期条件であれば視点と関係なく運動主体感を抱いた．これはアバターと自身の動きの時間的一致によって自身の動きの表象がアバターに投射されることで距離に関係なく運動的自己が拡張されたと考えられる．このことから運動主体感は身体の動きとの時間的整合性が重要であるが、身体保持感はそれに加えて空間的整合性も必要になるのだと考えられる。

参考文献

[Gallagher 2000] Gallagher S: Philosophical conceptions of the self: Implications for cognitive science. Trends in Cognitive Sciences, 4: 14-21. 2000.

[Kalckert 2012] Kalckert A: Moving a rubber hand that feels like your own: A dissociation of ownership and agency. Frontiers in Human Neuroscience, 6, 40. 2012

[Petkova 2011] Petkova V. I: The perspective matters! Multisensory integration in ego-centric reference frames determines full-body ownership. Frontiers in psychology, 2, 35. 2011

[Heydrich 2013] Heydrich L: Visual capture and the experience of having two bodies – evidence from two different virtual reality techniques. Front. Psychol., 4, 946. 2013

[Lloyd 2007] Lloyd, Donna M: "Spatial limits on referred touch to an alien limb may reflect boundaries of visuo-tactile peripersonal space surrounding the hand." Brain and cognition 64.1: 104-109. 2007

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

バーチャルリアリティにおける身体図式キャリブレーション ～伸長・収縮方向における回帰モデルでの評価～

Body Schema Calibration under Virtual Reality: The Evaluation using Regression Model in The Direction of Extension and Contraction

望月　典樹[*1]　鈴木　涼[*1]　大山　英明[*2]　中村　壮亮[*1]
Noriki Mochizuki　Ryo Suzuki　Eimei Oyama　Sousuke Nakamura

[*1] 法政大学 Hosei University
[*2] 産業技術総合研究所 National Institute Advanced Industrial Science and Technology

In the immersive virtual reality system, when the shape of the virtual body is parted from the real body, there is a problem that its operability deteriorates. The authors assumed this is caused by misfit of the body schema which is an internal model in the brain realizing the body position, because it is initially tuned up to the real body instead of virtual body. Thus, the authors have proposed a method called body schema calibration using VR technology to change the body schema adapting to the real body so as to fit the virtual body. In this paper, we newly evaluated two regression models based on the result of body schema calibration to the direction of extension and contraction. As a result, two-piece linear regression model was more suitable than the linear regression model. And it was suggested that the changing characteristics of the body schema are different in the direction of extension and contraction.

1. はじめに

VR（Virtual Reality）とは，コンピュータで人工的に生成した情報をあたかも現実であるかのように感じさせる技術である．[舘 10]によれば，理想的なVRが満たすべき最も特徴的な要点は，コンピュータで生成した仮想空間が，

(a) 人間にとって自然な3次元空間を構成している
(b) 人間がその中で，環境との実時間の相互作用をしながら自由に行動できる
(c) その環境と使用している人間とがシームレスになっていて環境に入り込んだ状態が作られている

ということである．これらの3要素を有したVRシステムを構成する場合，人の動作そのものを入力として，仮想身体を操作することが現実感を高めることに適していると考えられる．例えば，仮想空間内でのコミュニケーション，エンタテイメントに用いられるアバタや，テレイグジスタンスで用いられる遠隔操作ロボットが挙げられる．しかし，仮想空間内の仮想身体の身体寸法が実身体と異なると，実身体と同様に操作をすることが難しくなる[渡邊 09]．これに対して著者らは，その原因が身体定位を実現する脳内モデルである身体図式の乖離にあると仮定し，実身体に適合している身体図式を仮想身体のものへと変更する「身体図式キャリブレーション」について提案し，研究を進めてきた[Nakamura 14] [依田 15] [望月 17]．

これまで身体図式の変更は，身体形状を伸長および収縮する方向に対して試みてきたが，開発過程の都合から，両者で被験者および実験環境を統一出来ていなかった．そこで本論文では，伸長・収縮の両方向に対して同一条件下で身体図式の変更を実施し，得られた結果を線形回帰モデルに当てはめることで比較検討を行う．

連絡先：望月典樹
法政大学大学院 理工学研究科 電気電子工学専攻
〒184-8584 東京都小金井市梶野町 3-7-2
090-9130-8887
mono.gmllab@gmail.com

2. 仮想身体の操作における問題とその原因

仮想身体の操作性低下に対するアプローチとして，遠隔ロボット操作においては，幾何学的に視野角と眼間距離を制御することで自然な空間として提示するスケーリングと呼ばれる手法が提案されている[柳田 02]．しかし，スケーリングには操縦者とスレーブロボットが非相似形で異なる場合に身体の一部しか整合させることができないという欠点がある．また，実空間を近似した仮想空間において，仮想的な遠隔ロボットを操縦する訓練を行い，遠隔ロボットへの一体化の度合いを高める手法が提案されている[大山 94]．しかし，この手法は訓練のために適切な動作パターンの選択法や訓練回数・時間の設定法を欠いており，さらに，災害地における瓦礫の落下など，遠隔地の環境が動的に変化した場合に，必ずしも対応できる保証はない．

2.1 身体図式と仮想身体の乖離が主要因である仮説

身体図式とは，「自分の身体の姿勢や動きを制御する際に働く無意識のプロセス」と定義される．例えば，「カップを持つ」という行為では，カップの大きさや位置を知覚し，自分の身体図式と照合した上で，手の形や伸ばす位置を決定している．一方で，

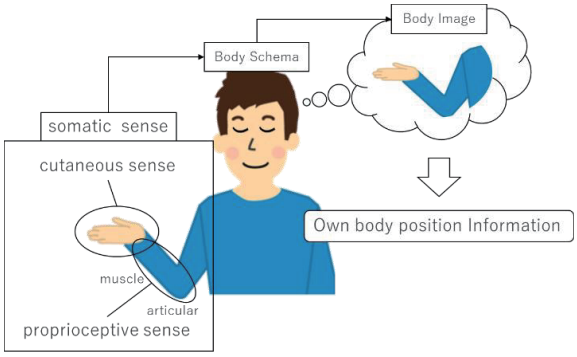


図 1 Body image recalled through body schema in the brain from somatosensory inputs.

身体イメージは「自分自身の身体について意識的に持つ表象」と定義される．これは客観的に想像した自身の姿であり，人間は身体図式 からその時々において身体イメージを生成している[Head 11]．上記のイメージを図 1 に示す．

身体図式を形成するには，視覚，聴覚，体性感覚（皮膚感覚および深部感覚）などの知覚情報が必要になる．特に，視覚的な空間知覚と触力覚的な空間知覚に関する情報の統合が重要である[Iwamura 98]．また，生成された身体図式は日々の体験や経験によって維持され，リアルタイムに更新されている．

人間の動作は身体図式が基盤となっている．つまり，自分自身と身体形状の異なる仮想身体を操作する場合，身体図式が仮想身体に適したものになっておらず，その結果として操作性の低下が生じている可能性が考えられる．

2.2 身体図式の変更事例

身体図式は通常，成長に合わせて変化した身体に合致するように徐々に更新されていく．一方で，急激な身体変化に対しては身体図式の変更は行われ難く，事故などで身体を欠損した際には幻肢が生じる場合がある[Bromage 74]．これらの自然発生的な身体図式の変更は高精度ではあるが長時間を要する．

また，能動的に身体図式が更新された事例としては，RHI[Botvinick 98]や VHI[Slater 08]，音による錯覚[Tajadura 12]が報告されている．しかし，これらは短時間での身体図式の変更が認められるもののその精度は低い．

これに対して，著者らが考案した「身体図式キャリブレーション」は，仮想空間内での運動により寸法変更後の身体と体性感覚とを紐づけることで身体図式を更新するという手法であり，短時間で高精度に身体図式を変更することができる[Nakamura 14] [依田 15] [望月 17]．

3. 身体図式キャリブレーションとその評価方法

3.1 身体図式キャリブレーション

身体図式は，運動の際に身体像と体性感覚情報を統合することで学習される．正確な身体像は仮想空間において自身が操作する仮想身体である．ここで，仮想身体を表示している場合，自身の身体像は視覚によって認知できるため，体性感覚に基づいた身体像が意識されず，身体図式変更の所要時間が増大すると考えられる．すなわち，短時間で変更するためには，仮想身体の映像を遮断して，体性感覚に基づく身体像を意識させることが有効である．一方で，仮想身体は身体像の正確な教師信号となるため，仮想身体を常時遮断している状況では，正確な身体像を学習することが困難となり，身体図式の確度が低下する．そこで我々は以下の 2 つの学習過程をハイブリッドで行う手法を提案している．なお，身体図式を変更する対象は右前腕に設定している．

(a) 変更目標の身体像に合致した身体映像が視覚提示された状態で運動を行う学習過程

(b) 身体映像は遮断した状態で，変更目標の身体像に合致した指先位置のみを間接的に視覚提示し，そこから想起される身体映像を教師信号として，体性感覚主導の学習を行う過程

この 2 つの要件を満たす流れを図 2 に沿って説明する．

まず，実空間と仮想空間の両者に存在する MR スティックを使用し，実身体及び仮想身体の右腕に触れ，視触覚同期刺激を提示することで，仮想身体に対する自己所有感を高め，VHI の状態を作り出す．仮想空間には仮想スイッチを設置し，スイッチを右手中指で押す動作（スイッチングタスク）を，実身体の運動による仮想身体の操作で繰り返し行わせる．この時，仮想スイッチは押されるたびにランダムな位置に移動する．そして目標の身体イメージは実身体の寸法に対する倍率である寸法比として設定し，提示寸法比になるまで，仮想スイッチを押した回数に従って，仮想身体の右前腕を徐々に肘から手首，または手首から肘の方向へ線形的に変形させていく．また，一定のスイッチング回数ごとに仮想身体を視覚提示しない状態でスイッチングをさせ，成功失敗の判定は仮想スイッチの変形と移動によって視認できるものとした．なお目標とする身体イメージに対応する仮想身体を提示した状態での運動機会を確保するために，目標の寸法に達した後も十分な回数のスイッチングを行わせる．

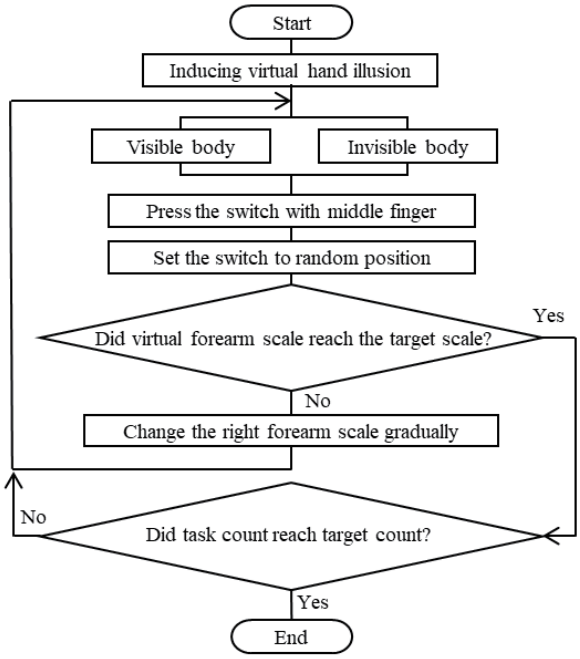


図 2 The flow of changing body schema.

3.2 評価方法

身体図式は脳内モデルであるため，直接測定することは困難である．そこで，[Prablanc 79]をもとに，体性感覚入力に伴い身体図式より出力される，身体の位置姿勢情報が確実に利用される動作において，その位置関係から間接的に身体図式の変化を捉えることとした．具体的には，目標位置を視覚提示する手段として仮想コーン（仮想空間のみに存在する三角錐状のオブジェクト）を用意し，身体映像を視覚遮断した状態で，実身体の右手中指先端を仮想コーンの先端に一致させるポインティングタスクなる課題を設定した．この時，仮想コーンの座標および実身体の右手中指先端の座標，右腕各部の位置を記録することで，その関係から身体イメージを間接的に評価する．このタスクをスイッチングタスクの前後に実施することで，指先位置の関係から身体図式の変化を推定することが可能である．

本論文ではスイッチングタスク前に測定した指先位置をポインティング位置A，スイッチングタスク後に測定した指先位置をポインティング位置Bと呼称する．また意図した形状への身体図式の変化を観測するために，身体図式を変更させる方向の成分のみを抽出する処理を施す．評価指標の導出過程を以下に説明する．

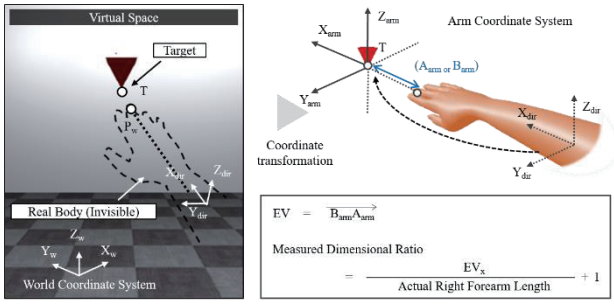


図 3 Pointing task and calculation of the updated level of body schema.

まず目標位置を原点，肘から手首に向かう方向をx軸正方向，掌から手の甲へ向かう方向をz軸正方向，XZ平面の法線方向をy軸正方向として腕座標系を定義する．そして座標変換により腕座標系におけるポインティング位置A_{arm}，B_{arm}をそれぞれ求める．この時，B_{arm}からA_{arm} へのベクトル$\overrightarrow{B_{arm}A_{arm}}$を$EV$とすると，$EV$の$x$成分が身体図式の変化量に該当する．また変更目標を比率で設定しているため，比較が容易にできるよう次元を合わせる必要がある．そこで変化量EVのx成分に対して実際の右前腕の長さで正規化を行い，右前腕の実寸に対する寸法比（実測寸法比）という形に変換する．この値が身体図式の変化の評価指標となる．ここで，ポインティングタスクおよび評価指標の導出手順のイメージを図 3 に示す．

4. 実験システム

身体図式キャリブレーションに基づき，構築したシステムの概要を図 4 に示す．詳細は以下の通りである．

被験者の上半身にはマーカーを取り付け，モーションキャプチャで取得した位置姿勢情報を仮想空間上のアバタの動作に反映する．仮想空間の情報は HMD(Head Mount Display)を被験者に装着し，視覚提示する．仮想空間は，実空間を模倣した空間とすることで現実感を向上させる．また，仮想スイッチと仮想コーンを仮想空間に配置し，両方の表示を切り替え可能としている．仮想スイッチが存在している場合はスイッチングタスクを被験者に与え，スイッチを押すたびに仮想身体の寸法が変更さ

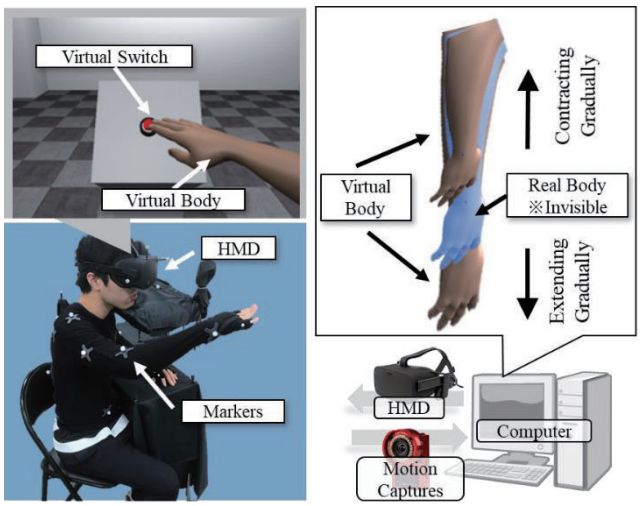


図 4 Experiment system.

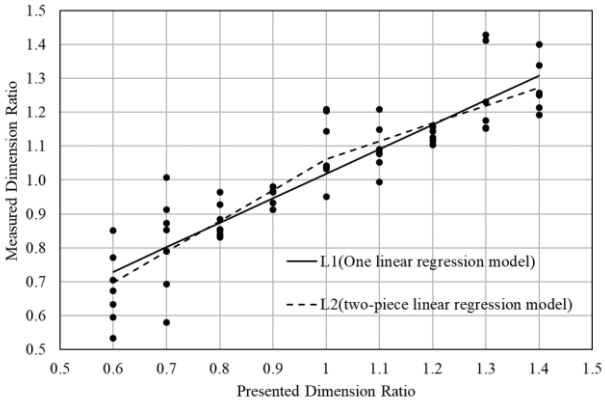


図 5 Experiment result.

れるようになっている．また仮想コーンが存在している場合は仮想身体を非表示にした状態で，前節で述べたポインティング位置と指先位置の差分を測定する機能が有効になっている．

5. 実験・考察

提案手法による伸長方向および収縮方向への身体図式の変更について，健常者 7 名（22～23 歳，男性，右利き）に対して，提示寸法比を 0.6～1.4 倍（0.1 刻み）として 9 通りの実験を行った．なおハイブリッド学習の実現のため，仮想身体を表示させた状態でのスイッチング 4 回と，非表示状態でのスイッチング 1 回を交互に行わせた．また提示寸法比は 100 回で目標値となるよう線形変化させ，目標値にて 200 回まで継続させた．

各提示寸法比に対する実測寸法比，およびその直線回帰と折れ線回帰の結果を図 5 に示す．提示寸法比に応じて実測寸法比が同方向に変化していることが確認できる．なお直線回帰モデル$L1$は，伸長方向と収縮方向の変化を1つの直線で表したものであり，折れ線回帰モデル$L2$は提示寸法比 1.0 から伸長方向と収縮方向で直線の傾きを独立させて表したものである．

ここで各回帰モデルに対して，AIC(Akaike's Information Criteria)を求めたところ，$L1$は AIC=-106.7，$L2$は AIC=-109.3 となり，$L2$の方がモデルとして適しているという結果になった．このことから，身体図式の変更に関して，伸長方向と収縮方向とでは特性が異なる可能性が示唆された．また$L2$の傾きは，伸長方向で 0.53，収縮方向で 0.90 となり，収縮方向の方がより高い効果が得られている．これは通常の成長過程で，現在よりも小さい身体図式での動作を経験しているため，それが適応のしやすさに影響した可能性が考えられる．

6. 結論

本論文では，以前より著者らが取り組んでいる「バーチャルリアリティにおける身体図式キャリブレーション」に関して，伸長方向と収縮方向の両方向に対して同一条件下での適用を試み，得られた結果に対する直線回帰モデルと，方向別に直線を分離させた折れ線回帰モデルとを比較した．その結果，伸長方向と収縮方向とでは，身体図式の変更特性が異なる可能性が示唆された．今後は，身体図式の変更限界に関する調査や，本手法を適用した際の仮想身体の操作性能について評価を行っていく予定である．

参考文献

[Botvinick 98] M.Botvinick and J.Cohen：Rubber hands‘feel’touch that eyes see，Nature，Vol.391，pp.756，(1998)

[Bromage 74] P.R.Bromage and R.Melzack：Phantom limbs and the body schema，Canadian Anaesthetists Society Journal，Vol.21，No.3，pp.267-274，(1974)

[Head 11] H.Head and G.Holmes:Sensory Disturbances From Cerebral Lesions”，Brain，Vol.34，No.2，pp.102‐254，(1911)

[Iwamura 98] Y.Iwamura : Hierarchical somatosensory processing”，Current Opinion in Neurobiology，Vol.8，No.4，pp.522-528，(1998)

[Nakamura 14] S.Nakamura，N.Mochizuki，T.Konno，J.Yoda，and H.Hashimotoa:Research on Updating of Body Schema Using AR Limb and Measurement of the Updated Value，IEEE Systems Journal，Vol.10，No.3，pp. 903-911(2014)

[Prablanc 79] Prablanc, C., Echallier, J. F.,Komilis, E. and Jeannerod, M : Optimal response of eye and hand motor systems in pointing at a visual target, Biological Cybernetics, Vol. 35, Issue 2, pp 113-124, (1979)

[Slater 08] M.Slater, D.Perez-Marcos，H.H.Ehrsson and M.V.Sanchez-Vives：Towards a digital body: The virtual arm illusion，Front Hum Neurosci，Vol.2，No.6，(2008)

[Tajadura 12] A.Tajadura-Jimenez，A.Valjamae，I.Toshima，T.Kimura，M.Tsakiris and N.Kitagawa : Action sounds recalibrate perceived tactile distance，Current Biology，Vol.22，No.13，pp.516‐517，(2012)

[大山 94] 大山英明，常本直貴，前田太郎，舘暲：仮想環境へのテレイグジスタンスのための一手法”，日本ロボット学会誌，No.8，pp.1175-1182，(1994)

[舘 10] 舘暲：バーチャルリアリティ学，2010.

[望月 17] 望月典樹，中村壮亮，橋本秀紀：バーチャルリアリティにおける身体図式キャリブレーションに触力覚フィードバックが及ぼす影響，日本機械学会論文集，Vol.83，No.851，pp.17-22, (2017)

[柳田 02] 柳田康幸，舘暲：HMD 型テレイグジスタンスシステムの頭部運動時における視野角不整合の影響，日本バーチャルリアリティ学会論文誌，Vol.7，No.1，pp.69-78，(2002)

[依田 15] 依田淳也，中村壮亮，昆野友樹，望月典樹，橋本 秀紀：AR 肢体を用いた身体位置感覚更新が運動軌道に与える影響の評価，日本機械学会論文集，Vol.81，No.829，pp.15-00249，(2015)

[渡邊 09] 渡邊孝一，川上直樹，舘暲：テレイグジスタンス・マスタスレーブシステムにおける操縦者とスレーブロボットとの間の寸法不一致の影響，日本バーチャルリアリティ学会論文誌，Vol.14，No.3，pp.391-394(2009)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

誰かと対戦しているとの思い込みがビデオゲームを楽しくさせる：事象関連電位（P300）による検討

The existence of a competitor increases enthusiasm when playing video games: Investigations based on the event-related potential (ERP) component (P300)

川合伸幸[*1*2] Nobuyuki KAWAI　中田龍三郎[*1] Ryuzaburo NAKATA　久保賢太[*2] Kenta KUBO

[*1] 名古屋大学情報学研究科 Graduate School of Informatics, Nagoya University
[*2] 名古屋大学情報科学研究科 Graduate School of Information Science, Nagoya University

Abstract: Studies suggested that an actual existence of opponent players increases competitive motivation and enjoyment of the game playing. However, it is unknown whether a belief of the existence of a competitor also increases these motivation and enjoyment. We hypothesized that projection (i.e., as if playing with a real competitor) would increase the participants' engagement to the video game than playing with the computer player. In this study, the participants played alone (single condition) or played with the fake player (competitive condition), while their brain activity was measured by means of event-related potential (P300). Results showed the amplitude of P300 in the competitive condition was smaller than that in the single condition, indicating that the belief of playing with the opponent player modulated the brain activities. The subjective scores were higher than those in the single condition. These results suggest that a projection modulates the brain activities and psychological states in the video game playing.

1. はじめに

近年のゲーム研究では，１人でゲームをプレイするよりも，誰かと一緒にゲームをプレイするほうがより楽しく感じることが報告されている．たとえば，対戦相手が存在するほうがゲームに対する動機づけや面白さが増し[1]，快感情が高まる．ゲーム中の熱中度を直接測定することは難しいが，事象関連電位(P300)により外乱刺激（プローブ刺激）に対する注意を測定することで，ゲームプレイ中の熱中度を間接的に測定することが可能である．この手法（プローブ刺激法）を用いて，映画に対する注意を調べた研究では，興味のある映画を鑑賞している場合には興味のない映画を鑑賞するときにくらべて P300 が減少していた[2].

しかし，他者と実際に対戦していなくても，他者とゲームをしているとの思い込みだけでも，熱中度が変化する可能性がある．すなわち，そのゲームを架空の対戦相手とのゲームとプロジェクションすることで，実際には同じようにコンピュータと対戦していても神経活動や心理状態が変化する可能性がある．そこで，２つの実験で検討した．

2. 実験1：通常のビデオゲーム

2.1 方法

参加者は，体動や眼電といったアーティファクト混入により，脳波測定が不良であった 3 名を除いた大学生 10 名(男性 5 名，平均年齢 19 歳，全員右利き).

2.2 刺激と課題

主課題：対戦格闘ゲームの STREET FIGHTER IV(CAPCOM 社製)をゲーム課題として用いた．参加者はゲーム内モードの一つであるトレーニングモードでプレイした．トレーニングモードでは，ダメージを与えても(受けても)，一定時間が経過することで，そのダメージが回復する．そのため，勝敗は課題中にはつかない．全参加者は同一のキャラクターを操作した(ゲームのスタンダードなキャラクターであるリュウ)．対戦相手とコンピュータも同様のキャラクターを操作し，リュウ v.s.リュウの同じキャラクター同士の対戦を行わせた．コンピュータの強さを示す難易度は，Easy（5 段階の 2 段階）を選択し，課題前に 3 分間の練習時間を設けた．

参加者は，実験では単独条件として一人でコンピュータ相手にゲームをする場合と，対戦条件として実験者とゲームをする場合があることを教示された．課題として，出来るだけ相手にダメージを与えられるよう努力することを教示した．

単独条件: 参加者は，コンピュータと対戦を行った．この時のコンピュータの強さは，練習課題と同様のものだった．

対戦条件: 参加者はサクラである女性（別の実験者）と対戦を行うよう教示された．実際には，ゲームは単独条件同様にコンピュータが操作しており，難易度も同様だった．対戦条件では，女性実験者が操作しているよう，ゲームの状況に合わせてボタン押しなどを行う演技を行った．参加者は，単独条件と対戦条件の両方に参加した．課題の実施順序は参加者間でカウンターバランスをとった．

副課題: ゲームプレイ中に聴覚オドボール課題を行った．聴覚刺激は，標準刺激(standard: 呈示確率 = 80 %, 1000 Hz)と標的刺激(target: 呈示確率 = 20 %, 2000 Hz)を参加者の前面約 60 cm の位置から音圧 60 dB，持続時間 70 ms, 立ち上がり立ち下り 10 ms で呈示した．

参加者は，呈示される 2 種類の聴覚刺激のうち，高い音が呈示される回数を数えるよう教示した．呈示回数は標準刺激が 192 試行，標的刺激が 48 試行だった．刺激オンセット間隔は 1500 ms だった．副課題を実施するために，ゲームのプレイ音声は呈示しなかった．

各条件における参加者の心理状態を測定するため，各条件実施後に質問紙による主観指標を測定した．質問紙は，先行

連絡先：川合伸幸，名古屋大学情報学研究科，〒464-8601 名古屋市千種区不老町，kawai@is.nagoya-u.ac.jp

研究[3]と同様に，ゲームプレイに対する暴力性，不満，興奮度，面白さ，楽しさ，難しさについて，7 件法(低い 1-7 高い)で回答を求めた．さらに，日常的にゲームをする頻度や，そのゲームのジャンルについても回答を求めた．

2.3 結果

各条件に対する総加算平均の結果を Figure 1 に示す．Figure 1 では，刺激呈示後 250-600 ms にかけて，Pz における ERP 波形に大きな陽性の電位が認められた．この電位を極性，潜時，頭皮上分布から P300 と同定した．

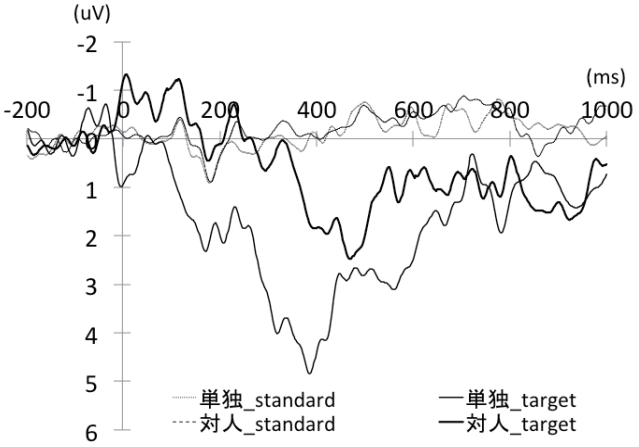


Figure 1. 各条件における総加算平均波形(N = 10)

主観指標では，興奮度($t(18) = 3.88$, $p = .001$)と面白さ($t(18) = 2.89$, $p = .005$)，楽しさ($t(18) = 2.59$, $p = .009$)において，対戦条件が単独条件に比べて有意に高かった．難しさについては，条件間で有意な差は見られなかった．

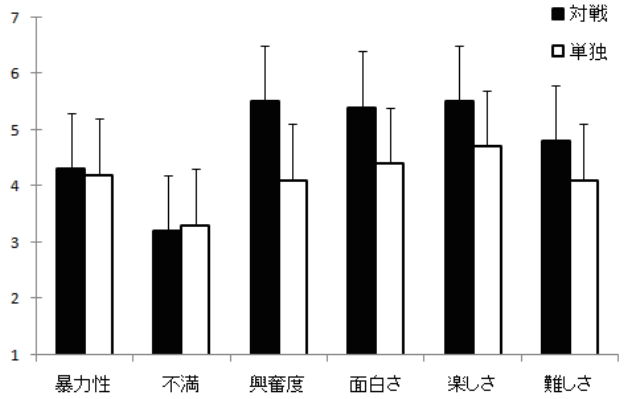


Figure 2. 対戦/単独条件における心理反応(N =10). 横線は有意差が認められた項目($p < .05$).

3. 実験2：VR ゲーム

隣に対戦中の相手が見えないヘッドマウントディスプレイをかぶった状態でも同じ結果が得られるかを確かめた．

3.1 方法

大学生 19 人（♂12名，M = 18.9 歳）が実験に参加した．事前に本実験で用いたゲームをプレイした参加者はいなかった．

3.2 刺激と課題

VR デバイスとして Playstation VR を2セット用意し，制御用デバイスの Playstation 4 とともに横並びの2台の机にそれぞれ1セットずつ設置した．ゲーム画面は VR ヘッドセットの 5.7 インチ OELD ディスプレイ上に呈示された．視野角は約 100 度であった．ゲームの音は呈示しなかった．ゲーム操作は Playstation 4 に付属のコントローラを使用した．コントローラの振動機能は使用しなかった．ゲームプレイ中の聴覚刺激呈示用にスピーカを実験参加者の後方約1 m に設置した．

実験では2つの走行条件（VS. コンピュータ条件，VS. 他者条件）を設定し，参加者はランダムな順序で両方の条件を経験した．VS. コンピュータ条件では対戦相手がコンピュータであることを走行前に教示した．実際に対戦する車はコンピュータが操作した．VS. 他者条件では実験者と対戦することを走行前に教示した．実験者は参加者の横のシートに座り，ゲームをプレイしているような演技をしたが，実際に対戦する車は VS. コンピュータ条件と同様にコンピュータが操作した．参加者は VR ヘッドセットを装着しているので，実験者の様子を視覚的に確認することはできなかった．他者の存在の手がかりになる聴覚情報を参加者に与えないように，実験者は対戦中にできるだけ動かないようにした．

脳波の測定方法や主観指標などは実験1と同じであった．

3.3 結果

基本的には実験1と同じ結果が主観指標，脳波に得られた．低頻度刺激の数え間違い数も，実験1と同様に他者と対戦していると信じているほうが多く，より熱中していたことを示す．

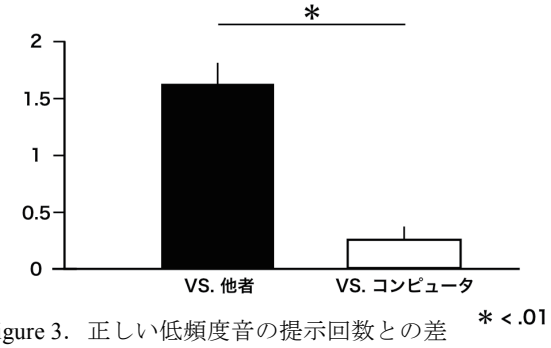


Figure 3. 正しい低頻度音の提示回数との差

4. 結論

これらのことは，他者との対戦との思い込み（プロジェクション）が脳波や課題遂行，主観的な楽しさに影響することを示す．

参考文献

[1] P. Vorderer, T. Hartman, & C. Klimmt. Explaining the enjoyment of playing video games: the role of competition. ICEC'03 Proceedings of the second international conference on Entertainment computing, 1-9, 2003.

[2] 入戸野 宏. 映像に対する注意を測る一事象関連電位を用いたプローブ刺激法の応用例一 生理心理学と精神生理学, 24, 5-18, 2006.

[3] B. J. Bushman, & J. L. Whitaker. Like a magnet: Catharsis beliefs attract angry people to violent video games. Psychological Science, 21, 6, 790-792, 2010.

Organized Session | Organized Session | [OS] OS-16

[2K4-OS-16a] 臨床の知(1)

諏訪 正樹（慶應義塾大学）、藤井 晴行（東京工業大学）、加藤 文俊（慶應義塾大学）

Wed. Jun 5, 2019 3:20 PM - 4:40 PM Room K (201A Medium meeting room)

[2K4-OS-16a-01] Consideration on Onigokko as a healthcare content for children in artificial intelligence society

〇Yuushi Hiramine[1,2] （1. Onigokko Association, 2. Onigokko Research Institute）

3:20 PM - 3:40 PM

[2K4-OS-16a-02] Types of Knowledge Emerging from Photos and Language Expression of Experience

〇Haruyuki Fujii[1], Sadayo Hirata[2], Ken-icih Shinozaki[3] （1. Tokyo Institute of Technology, 2. Shibaura Institute of Technology, 3. Nihon University）

3:40 PM - 4:00 PM

[2K4-OS-16a-03] Process of Comprehension of Constructing Dry Stone Walls in Izena in Ryukyu

〇Kenichi Shinozaki[1], Sadayo Hiram[2], Haruyuki Fujii[3] （1. Nihon University, 2. Shibaura Institute of Technology, 3. Tokyo Institute of Technology）

4:00 PM - 4:20 PM

[2K4-OS-16a-04] Toward the Aesthetic Appreciation of the Haptics and Tastes

〇Hiroki Fukushima[1] （1. Kyushu Women's University）

4:20 PM - 4:40 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

人工知能社会における子どものための ヘルスケアコンテンツとしての鬼ごっこに関する考察

Consideration on Onigokko as a healthcare content for children in artificial intelligence society

平峯 佑志[*1]
Yuushi Hiramine

[*1] 一般社団法人鬼ごっこ協会
Onigokko Association

This paper discusses the contents of deeper consideration through practical activities that utilize the contents of Onigokko on the basis of previous research on health care issues for children in artificial intelligence society. Onigokko is a traditional play that everyone has tackled in early childhood. In modern Japan, it is said that children's physical fitness declines and sports population declines are progressing, and the number of physical play such as Onigokko is decreasing. However, the development of mobile terminals, game machines and the like is remarkable. In this paper, we conclude that in the society in which artificial intelligence develops, it is important for the future of the child to value creation of physical play called Onigokkoaccording to social demands.

1. はじめに

鬼ごっこは、人間本来の本能的な性質に基づいて、追うこと、追われることの端的な身体的動作を行うことであるが、人間がこの世に生を受けて乳幼児から幼児期、学童期に至っていく発育発達の過程に置いて、最も大切な子どもの遊戯の一つであると考えられている。文科省では、2012年から幼児期運動指針という政策において、幼児期に欠かすことができないものとして「鬼ごっこ」を単体の大項目として取りあげているほどである。人間が生を受けた後に、自分以外との有機的な対象物と行う身体的遊戯として、鬼ごっこは最初に行われる動作ではないだろうか。例えば、ハイハイをして何かの動的な対象物や生物を追いかけたり、空中を浮遊している虫や鳥を目で追いかけながら手を伸ばしたりなど、鬼ごっこをする上での原始的な身体動作を生まれた直後から行うことなどがあげられる。人工知能関連の研究でも、身体知や暗黙知という言葉で、身体的認知についての研究が進められているが、鬼ごっこの身体知は、乳幼児から動作を行っていることからも、本能的な部分に由来していると考えられる。本研究では、本能的に備わっていると考えられる鬼ごっこを通じた身体的な楽しさが、人工知能社会においては多くの生活様式が機械化や電子化されて、身体的な動作の多くを要することの必要性が低下してくる中で、鬼ごっこのような身体的遊戯が人間の心身のヘルスケアへと貢献して、さらに人工知能と相対することなく、共生関係を図っていく社会へ向かうには、どうすれば適切に社会が機能していくのかについての道筋について考察を行った。

2. 鬼ごっこのヘルスケア

鬼ごっこのヘルスケアに関係する研究は、私も所属している鬼ごっこ総合研究所で主に行われている。鬼ごっこは、一般的には誰もが知っているが故に、その効果や効能については、あまり注目を集めてこなかった。そんな鬼ごっこに着目して、子どものヘルスケアのために研究を長年してきたのが、鬼ごっこ総合研究所を運営している一般社団法人鬼ごっこ協会の代表理事である羽崎泰男（以下、羽崎）である。まず、鬼ごっこ協会という組織であるが、鬼ごっこという身体的遊戯を活用して、子どもから大人の健全な心身を育み、誰もが健康的になることを目指して活動している団体である。2010年6月に鬼ごっこ協会を設立した羽崎は、元々は1985年に設立されて2015年に閉館となった東京都渋谷区の青山に立地していた国立総合児童センター「こどもの城」の事業本部長などを歴任していた。そこで、20数年前に当時の厚生省との共同研究によって、高度経済成長期を経て子どもの肥満が増えてきていた1980年代を背景として、肥満の防止のための運動プログラムに鬼ごっこが適していることを、数あるスポーツ競技やトレーニングを行った結果として辿り着いたとされている。ここで注目に値することが、既存のスポーツ競技や一般的なトレーニングを肥満の子どもにやらせた結果として、運動が好きではない子どもの割合が多かったため興味関心を示すことが無かった中で、鬼ごっこには関心を示したということであった。笹川スポーツ財団の調査（2017）（表 1）によると、過去 1 年間に「よく行った」運動・スポーツ種目で、4～11歳では、鬼ごっこは第 1 位（47.3%）となっているなど、子どもの鬼ごっこの関心が高いことを伺うことができる。

（表1） 過去 1 年間に「よく行った」運動・スポーツ種目（4～11 歳）

順位	実施種目	実施率
1 位	鬼ごっこ	47.3%
2 位	水泳	34.2%
3 位	自転車遊び	30.4%
同位	ドッジボール	30.4%
5 位	ぶらんこ	25.4%

（引用：笹川スポーツ財団（2017）子ども・青少年のスポーツライフデータ）

3. 人工知能社会における鬼ごっこの役割

人工知能が発達していく社会では、人の働き方やライフスタイルが変化すると言われており、なくなる職業と残る職業ランキングがオックスフォード大学の論文（表 2）で公開されたが、そのランキングでは残る職業としてレクリエーション療法士やメンタル

ヘルスカウンセラー、教育コーディネーター、作業療法士などが挙げられており、メンタルヘルスや身体的なヘルスケアなどの領域で、主に福祉や教育、医療などの観点で取り上げられている。

（表 2）あと 10～20 年でなくなる職業・残る職業の一部

残る職業		無くなる職業
レクリエーション療法士	1	電話販売員
整備・設置・修理の第一線監督者	2	不動産登記の審査・調査
危機管理責任者	3	手縫いの仕立て屋
メンタルヘルス・薬物関連ソーシャルワーカー	4	コンピュータを使ったデータの収集・加工・分析
聴覚訓練士	5	保険業者

（出典）The future of employment：how susceptible are jobs to computerization?より一部作成

ここで私が注目したのは残る職業では対人間コミュニケーションを綿密にとる必要性があるものが多くあり、なくなる職業としては手続きをシステム化しやすい事務作業や簡易作業などが多く含まれていることである。ここから、鬼ごっこの人工知能社会との関連性について述べたい。鬼ごっこは、基本的にはあまり多くの道具を使わずに場所を問わず、必要とされるデジタル機器も無く取組むことができる。身体が一つあれば、環境を選ばずに心身の健康のために子どもが気軽にできるのも特徴である。社会一般の人々にとって、人工知能社会といって身近な端末としては、一例としてスマートスピーカー、スマートフォン、ウェアラブル端末などが、普段から接している身近な機器であるが、これらの機器は必ずヘルスケアのためのコンテンツがアプリケーションなど、様々な形式で内蔵されている。主には、1人で取組む事ができるランニングやワークアウト、トレーニングのためのコンテンツが多く存在している。鬼ごっこも、スマートフォンで鬼ごっこを活用したアプリケーションはダウンロードすることができる。私は、人工知能分野の研究者ではなく、鬼ごっこの指導を子どもや大人にしている専門家であるので、人工知能の詳細な発展形式の可能性については、詳細な造詣は持ち合わせていないが、人工知能が社会にもたらす効果に着目をしている。多くの人々に人工知能による社会への恩恵がもたらされることは魅力的であると考えられるのは間違いがないが、その裏腹として人工知能が発達していく中での弊害としては、メンタルヘルスの問題やヘルスケアの問題が一つとしてあげられている。メンタルヘルスではカウンセリング、ヘルスケアではトレーニングやランニングが取り上げられるが、私はスポーツや遊びが果たす役割は大きいと考えている。スポーツや遊びは、対人で行うものがほとんどである。一部には、1人で行うものもあるが、基本的には個人対個人か集団対集団の対人間で行われる。人工知能社会では、先ほどの残る職業として対人間コミュニケーションの重要性が考えられているが、スポーツや遊びは対人間でのコミュニケーションによって成り立つものであるので、人工知能の発達した社会においては果たすことができる役割があると、私は考えている。中でも、特に鬼ごっこは対人間で行い、さらに道具を必要とせずに、行う場所も屋内外を問わずにできるので、人口知能が発達した社会においても、気軽に対人間と取り組めるという点で果たすことができる役割は大きい。2018年6月28日付の日本経済新聞には、未来学のポスト平成の未来学「第8部遊びの進化論・デジタルから離れて・顔合わせ身体動かす爽快感」として、人工知能が発達したデジタル社会と対比して、アナログな身体的遊戯である鬼ごっこによる、メンタルヘルスやヘルスケアに貢献が記事にもされているなど、メディアからも注目をされている。

4. 直観と人工知能の先行研究

鬼ごっこは、乳児期や幼少期より行われる、人間の本能に由来した身体的な遊戯であるが、それは知能と言うよりも直観として必要性を認知して行われているものなのではないだろうか。対人間とのコミュニケーションを図っていく上では、鬼ごっこと言う手段を通じて、分かり合うことには大きな価値がある。そこで、この「直観」と人工知能についての先行研究があるので、ここで取り上げたい。意思決定学者であり教育コンサルタントの奈良潤氏の著書「人工知能を超える人間の強みとは」では、「直観と人工知能のすみ分け」という言葉を用いて、直観と人工知能の共生について論じている。詳しく言い換えると、奈良氏は「人間は、人工知能のように記憶力に優れていない。また、直観を働かせても判断の正確性に欠けることがある。しかし、人工知能は人間とは違い、指導力を発揮して組織運営をすることには不向きである。人工知能は、独創性を発揮して実際の現場で問題解決の決断を下すことができない。」と、企業などの組織マネジメントの領域の話をもとに述べている事からも、人間として人工知能との共生を図っていくためには、お互いの強みを生かす必要性があるということが分かる。まだ、日本においては人間の直観や身体的な感覚を知能として深めていくことを志向する身体知などの研究は、多くは進んでいない中で、世界中の老若男女が誰もが鬼ごっこのような身体的な遊戯は子どもの頃に経験しているものであり、それは無意識でお互いの距離感や親密性を図り合うために、遊びを通じてコミュニケーションをとる姿が見られることは、人間の更なる知覚や知能を深く理解していく上で重要となることではないだとうか。

5. 鬼ごっこと人工知能の社会的共生

鬼ごっこと本論文で取り上げたのは、私が鬼ごっこ協会の指導員としての豊富な実務経験があるからでもあるが、鬼ごっこはデジタルな要素や様々な社会的制約は全くなくできて、老若男女が誰もが知っているコンテンツであるということから、これからの人工知能社会において、人間が健康的に生きていくための必要な要素が詰まっていると考えたからである。特に、鬼ごっこ協会で普段から接している子ども達は、デジタルネイティブ世代と言われ、この世に生を受けた時からコンピューターネットワークは整備され、様々な電子デバイスが生活の中にあり、働く世代を迎え頃には人工知能の発達によって職業構成が変化している可能性が高い。そのことからも、人工知能と人間との社会的共生を図っていく上での、鬼ごっこのような人間の身体性を分かりやすく表わせたコンテンツと人工知能の社会的共生を考える事は、人間という生身の存在がどのようにこれからの人工知能社会に向き合っていくかのヒントがあるように考えられるのではないだろうか。

6. 結論と今後の展望

現代に生きる人間も、種の起源をたどると皆、ホモ・サピエンスであり、知能も発達しておらず、歴史を辿るにつれて、人間社会の繁栄を少しずつ築いていった。鬼ごっこも、まだ原始の時代にも、人間が獲物を追いかける行為をしていたが、当時から本能的に人間の身体知としてなくてはならないものであったと考えることができる。時代は遡って、現代は第 3 次人工知能ブームで盛んに社会的に人工知能へと注目が集まっている中で、社会環境の変化や働く環境の変化が起こってくると提唱されてい

るが、専門家や技術者ではない一般の市民にはまだ実感としては、これまでの日常の延長線上で日々の生活を営んでいるように感じられる。しかしながら、確実に生活の中にある様々な製品やデバイスには人工知能が組み込まれ、生活は一変をしようとしている。そのような確実にではあるが、水面下で大きな社会の変化を迎えている中で、これまでの時代以上に対人間との関係性のありよう、心身のヘルスケアのための身体的動作への強い動機を持ち得ることは、貴重な価値となってくるのではないかと私は考えている。人間は、進化を遂げたとはいえども、対人間とのコミュニケーションや関係性があって、初めて充足感のある生活を送ることができるのではないだろうか。しかし、人工知能によってもたらされる社会的便益や生活環境の改善、職場環境の変化については、元来の社会の有り様を保っていくために、その変化を否定するのではなく、新しい価値観を柔軟に取り入れる必要がある。鬼ごっこは、多くの人にはアナログコンテンツとして認識され、昔遊びの一つの形式であると捉えられているにすぎないかもしれない。しかし、私が、その価値に関しての普及活動を開始してから 8 年を迎える中で、数多くの事例を見る中で、現代社会に生きる人々に必要とされるコミュニケーション能力や対人との調和を持つ力、集団に置いて良好な関係性を築くための力などは、老若男女が鬼ごっこをすることで培われる現場を多く見てきた。また、鬼ごっこの研究者である羽崎が、鬼ごっこ協会を設立して鬼ごっこを通じて、社会の変革を起こそうとしたきっかけに、デジタルゲームの登場があったことを忘れてはならない。デジタル的な遊びにアナログ的な遊びが、一方的な視点での批判をするのではなく、双方が発展していくために、羽崎が研究を重ねてきた背景があることは、現代においても踏まえなければならない。当時以上に、2019 年現在は人工知能の発展が著しく、子ども達の将来の働き方や生き方にも影響を与えてくることは間違いない。しかし、子ども達が人間として生きる上での大切なことを、鬼ごっこを通じて心身のヘルスケアを行うことは、とても大切なことである。人工知能が難しいとされている、前例のない新たな価値創造を、鬼ごっこを通じて行っていくことが、子ども達の未来のためになると私は考えている。

7. 謝辞

本研究を進めるにあたって、これまで鬼ごっこ協会の活動現場で出会い、研究への動機をいただいた全ての子ども達と保護者、関係者の皆様へ感謝を申し上げます。

参考文献

[羽崎 92] 羽崎泰男:ビデオ写真で見る0・1・2・3歳児の運動発達と体操実践, 日本小児医事出版社 (1992)

[野中・竹内 96] 野中郁次郎,竹内弘高:知識創造企業, 東洋経済新報社 (1996)

[羽崎 02] 羽崎泰男: 鬼ごっこ楽しくあそんで体力づくり, 日本小児医事出版社 (2002)

[マイケル・ポランニー 03] マイケル・ポランニー:暗黙知の次元,筑摩書房(2003)

[文部科学省 12] 文部科学省:幼児期運動指針(2012)

[松尾 15] 松尾豊:人工知能は人間を超えるか, 株式会社 KADOKAWA (2015)

[Klein 15] Gary Klein:「洞察力」があらゆる問題を解決する, フォレスト出版 (2015)

[諏訪 16] 諏訪正樹:「こつ」と「スランプ」の研究 身体知の認知科学, 講談社 (2016)

[Harari 16] Yuval Noah Harari:サピエンス全史 上・下 文明の構造と人類の幸福, 河出書房新社 (2016)

[若木 17] 若木均,羽崎泰男: スポーツ鬼ごっこの運動能力に及ぼす効果と今後の展望, 第 64 回日本小児保健協会学術集会 (2017)

[奈良 17] 奈良潤:人工知能を越える人間の強みとは, 技術評論社 (2017)

[海老原 17] 笹川スポーツ財団: 子ども・青少年のスポーツライフデータ, 笹川スポーツ財団 (2017)

[羽崎 18] 羽崎泰男,羽崎貴雄,平峯佑志: まるごと鬼ごっこ, いかだ社 (2018)

[羽崎 18] 羽崎泰男,羽崎貴雄,平峯佑志: 目指せ鬼ごっこ博士, いかだ社 (2018)

[落合 18] 落合陽一: デジタルネイチャー, 株式会社 PLANETS (2018)

[諏訪 18] 諏訪正樹:身体が生み出すクリエイティブ, 筑摩書房(2018)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

経験の写真とことばによる記録から顕在化される知の型

Types of Knowledge Emerging from Photos and Language Expression of Experience

藤井 晴行[*1] FUJII Haruyuki
平田 貞代[*2] HIRATA Sadayo
篠崎 健一[*2] SHINOZAKI Ken-ichi

[*1] 東京工業大学 Tokyo Institute of Technology #1
[*2] 芝浦工業大学 Shibaura Institute of Technology #2
[*3] 日本大学 Nihon University #3

The authors are trying to create a method of passing on traditional techniques and methods of building vernacular architecture. Although the local techniques and methods are significant and important for the architecture in the region, they are being driven out by modern and global technologies and methods. As a result, the cultural scenery typical and unique in a certain area is being destroyed. The authors focus on three types of knowledge i.e., theoretical knowledge, technological knowledge, and practical knowledge, and are representing them in such a form in which they are passed on from the experts to the prospective experts. This paper introduces our research plan and some findings from one of the pilot studies. We classified verbal expressions to describe how good, or bad, traditional artifacts are and found that descriptions of actions to create something are often used in the evaluation of the artifacts.

1. はじめに

「これまでの伝統技術を伝承するとともに新しい科学技術と融合したこれからの伝統となる技術を共創する方法の好ましい姿を明らかにする」ことを目的とする探究を、地域固有の伝統文化の継承と現代生活の質の向上を共存させるという現実の問題を解決しようとする研究的実践と実践的研究を連携することによって遂行している。具体的には、Ⅰ.技術知の共創プロセスにおける人々や環境の間に生じるインタラクション、Ⅱ.技術知を共創する場を形成する共同体システムにおける実践知とその価値、Ⅲ.技術知や実践知や理論知がインタラクションから生まれるプロセスを解明しようとしている。これらを通し、Ⅳ.臨床の知における主観性や固有性と科学の知における客観性や普遍性とを連携させる方法を構成的に確立したい。

本研究を着想する直接の契機は、沖縄県伊是名村の伝統的民家の空間構成と住生活の調査を継続して、科学技術が伝統技術を駆逐することを居住者たちは必ずしも歓迎していないことを知ったこと[藤井 2018a, 2018b]と、居住者たちの日常会話を通して、伝統技術の伝承の鍵は伝統技術を取り巻く共同体システムを支える慣習にもあることに気づいたことである。一方で、非論理的な思考も重要なデザインを対象とするデザイン科学の方法を構築する必要性を感じたことに端を発し[藤井 2005]、方法を模索する過程で、デザインの状況性[Gero 2005]と〈臨床の知〉[中村 1992,1993]に出会い、構築した方法論を上記の伝統技術と近代技術の共存の問題に適用しようとしている。

既報[藤井 2018a, 2018b]において、風土や文化に根ざした〈生きる〉ための工夫に利用される知と近代科学に基づく技術と相克しつつも共存する状況の背景にあるものごとを、〈空間図式〉に注目し[藤井 2017]、地域における〈臨床の知〉の共創と継承のための課題を題材として、浮き彫りにしている。

- 〈普遍性〉に駆逐された〈固有性〉に再び意味を持たせる
- 〈固有の文化と〈普遍的〉な要求を共存させる術を探す
- 〈普遍〉の技術を加えて〈固有〉の知を更新する
- 〈固有〉の知と〈普遍性〉の知との好ましい重なり方を探る

同時期に、伊是名村の石垣築造事業に関与する機会を得ている。本報では、次節に示す研究計画のパイロット・スタディとして、伊是名区後辺（くしひん）地区における石垣築造の経験に基づく〈技術知〉の抽出の試みを紹介する。

連絡先：藤井晴行，東京工業大学 環境・社会理工学院，東京都目黒区大岡山 2-12-1，fujii.h.aa@m.titech.ac.jp

2. 研究目的と方法

2.1 問い

「〈技術知〉や〈実践知〉はいかにして伝承されるか（第一の問い）」、「〈デザイン〉における〈特異点〉はいかなるプロセスとメカニズムによって出現するか（第二の問い）」、「〈デザイン〉という行為を扱う科学的方法はいかなるものであるか（第三の問い）」という三つの問いをたてている。ここで、法則性を有目的的に適用してものごとを創る知を〈技術知〉、実践に内在している実践それ自体としての知を〈実践知〉、生きるために自覚的に工夫することを広い意味での〈デザイン〉、デザインで合理的思考を超越した革新的なデザインの契機となる閃きが生じる瞬間をデザインの〈特異点〉とよんでいる。

第一の問いに答えるために「技術を伝承することは技術知や実践知をもつ教導者が継承者とともに教導者の技術知を〈協働デザイン（co-design）〉することであり、知を教導するというサービス提供のプロセスと知を継承するというサービス享受のプロセスのインタラクションによる技術と価値の共創である（[Nakashima 2016]より発想）」と仮定し、「技術を教導するプロセスは心身に内在する知を伝達できるようにことばや自身（教導者）が行為することを表現媒体として外化するプロセスであり、技術を継承するプロセスはことばや自身（継承者）が行為することを通して理解したものごとを内省して内化するプロセスである（[Suwa 2015]より発想）」と捉えたうえで、問いを「Ⅰ.技術知の共創プロセスにおける教導者と継承者と環境との間にいかなるインタラクションが生じるか」、「Ⅱ.技術知を共創する場を形成する共同体システムの実践知はいかなる特徴と価値をもつか」という二つの問いに分節化する。

第二の問いに答えるために、デザイン・プロセスがしばしばデザイン・スペースの内部での探索として定式化されること[Akin 1986]、Peirce が提案した演繹的推論、帰納的推論、発見的推論（abduction）を関連づける科学的探究のモデル[Peirce 1887]を踏まえて、「予期しないインタラクションを契機とするデザイン・

スペースの変容が革新的発想を生む」と仮定し、「Ⅲ.技術知や実践知や理論知はインタラクションからいかなるメカニズムで生まれるか」を問う。

第三の問いの背景には「デザインという行為を対象とするものごとには近代科学では扱えきれないものごとがある」という問題意識がある。近代科学が解明する〈科学の知〉は客観性、論理性、普遍性があることを基本原理とするが、デザインにおいて重要な役割を担う主観的な判断や世界観、論理を超越した思考、知の固有性や身体性にまなざしを向ける〈臨床の知〉を排除する傾向にある[中村 1992,1993]。〈デザインの科学〉の方法の一翼を構築することを視野に入れ、臨床の知と科学の知からなる知見を学術的に呈示する方法を明らかにすることを目指し、「Ⅳ.臨床の知の主観性や固有性と科学の知の客観性や普遍性の連携はいかなる方法によって可能か」という問いに対する答えを共創する。

2.2 研究目的

「伝統技術を伝承するとともに新しい科学技術と融合した新しい伝統となる技術を共創する方法の好ましい姿を明らかにする」ことを目的とし、以下を解明する。

Ⅰ.技術知の共創プロセスにおける参加者たちや環境の間のインタラクション、

Ⅱ.技術知を共創する場を創出する共同体システムにおける実践知とその価値、

Ⅲ.技術知や実践知や理論知がインタラクションから創発されるプロセス、

Ⅳ.臨床の知の主観性や固有性と科学の知の客観性や普遍性とを連携させる方法。

技術知の共創の研究的実践と技術知を共創するための実践知の顕在化と学理の実践的研究を実際の生活の場に臨み、一人称視点と三人称視点を連携して行うという研究方法をとる。研究対象であると同時に創造対象でもある技術知の共創プロセスに伝統技術の継承者かつ科学技術の教導者として参与して観測を行ない、プロセスの内部で問いを発生し、プロセスを変容させつつ、技術知の共創に関わる者として行為的直観の内省を通して創生する臨床の知と観察者としての客観的分析を通して創生する科学の知の関わり方を探究する。研究方法の有効性を研究の遂行によって実証することによって、臨床の知と科学の知の対立を止揚(aufheben)し、両者が発展的に共存する知見を学術的に示すデザイン科学の方法の柱のひとつを構築することを目論んでいる。

2.3 研究方法(全体構想)

近代技術の影響を受けながらも伝統的な技術や慣習を継承している地域を対象として、下記の観測①～④、分析準備⑤⑥、分析と考察⑦～⑧を 3～4 カ月を一周期として構成的に繰り返す。①～⑥はⅠ～Ⅳの解明に共通する。沖縄県伊是名村伊是名区を対象地域とする。

(1) 参与観察

伝統的な技術(臨床の知)を用いる築造の現場において、伝統技術の伝承を含む共同作業(結い)と結いの労をねぎらうもてなし(会食と懇談)に参加し、そこで起こるものごとを内側から観測する。複数回行ない、毎回、下記の記録①～④を行う。

①参加者の経験を各自の視点からビデオ記録する。

②気になる経験(気づく、腑に落ちる、目から鱗が落ちる、疑問に思う、迷うなど)をスナップショット写真、メモ、クロッキーなどとして記録する。

③共同作業や結いの状況を観察者の視点からビデオ記録する。

④共同作業や結いの全貌を俯瞰するビデオ記録を行なう。

(2) 分析準備

分析のためのデータを下記⑤⑥のように作成する。

⑤記録①③④に基づき、技術知や実践知を示す言語表現や非言語表現(身振りや身体動作など)と参照対象などとの時間的・空間的な関係をビデオ動画記録から抽出し、これらを継時的に列記する。

⑥記録①②に基づき、技術知と実践知をテーマとする写真日記(写真とことばによる気づきの記録)を作成し、複数人で議論しながらボトムアップで写真日記をグルーピングすることを繰り返して構造化する。構造化した写真日記から、記録①～④を援用し、モノグラフを作成する。加えて、写真日記をグルーピングする際の議論を記録①③④と同様に記録し、これに基づき、分析⑤と同様に、グルーピングの理由を示す言語表現と非言語表現(身振りや動作など)と参照対象などとの時間的・空間的な関係性をビデオ記録から抽出し、これらを継時的に列記する。

(3) 分析と考察

下記のⅠ～Ⅳを明らかにするために、それぞれ、下記の⑦～⑩を行なう。

Ⅰ.技術知の共創プロセスにおける参加者たちや環境の間のインタラクション

⑦データ⑤⑥に基づき、記録④を参照し、教導者の行為(手本)を示す表現、築造の手順や選択する材料を示す表現、行為の理由を示す表現などに注目して、技術知の内容、その言語表現、非言語表現、参照対象の間の動的な関係性を明らかにする。

Ⅱ.技術知を共創する場を創出する共同体システムにおける実践知とその価値

⑧データ⑤⑥に基づき、記録④を参照し、行為を示す表現、行為の理由を示す表現、慣習や信仰を示す表現に注目し、行為が示す実践知の内容と経験則や慣習や信仰との関係性を明らかにし、伝統技術を伝承する共同作業の場の形成のされ方を考察する。

Ⅲ.技術知や実践知や理論知がインタラクションから創発されるプロセス

⑨デザイン・スペースをデータ⑤⑥の言語表現の出現語によって構成される概念空間として表現し、技術知の共創における概念空間の状態遷移の巨視的分析による特異点を抽出し、特異点とその近傍における概念空間の構造の関係性の微視的分析による特徴づけ、分析結果⑦の結果を参照し、特異点とその近傍における共創の場の構成する要素の間のインタラクションのパターンの類型を明らかにする。ここで、特異点において、巨視的には概念空間の状態が不連続に遷移し、微視的には概念空間の構造が変化すると仮定している。

Ⅳ.臨床の知の主観性や固有性と科学の知の客観性や普遍性とを連携させる方法

⑩データ⑥と分析結果⑦⑧⑨から、主観的な気づき、間主観的な認識、客観的な事実の関係を分析することによって、主観的なものごとと客観的なものごとの関係性、固有性をもつものごとと普遍性をもつものごとの関係性を明らかにし、臨床の知と科学の知を連携する方法を考察する。

3. 〈技術知〉の顕在化手法のパイロット・スタディ

3.1 参与観察

2018 年 11 月、2019 年 1 月に実施した石垣築造において、参与観察（研究計画①～④）を行ない、分析（同⑤⑥）を並行している。ここでは分析⑥について報告する。

分析⑤は参与観察した学生たちを主体に石垣築造マニュアルを作成することを通して進めている。

分析⑥に関して、本報に先行して、89 枚の写真日記を構造化することを通して石垣築造における行為のパターンを抽出している［平田 2019］。

3.2 石垣の好ましさに関わる知の共通認識

写真とことばによる経験の記録（写真日記とその構造化、および、写真を参照しながらの会話）から当該経験に関わる知を顕在化することを試みている。具体的には、石垣築造の経験（作業への参与観察）を踏まえての、石垣の好ましさについての筆者らの会話からその好ましさの判断に関わる知を整理している。

石垣の意匠的な好ましさはいくつもの言語表現を用いて語られる。このうち、筆者らの印象に残っている、「石垣が端正である」と「石垣に味がある」という言語表現に注目し、これらの外延的意味（指示対象）と内包的意味（意義）に関する筆者たちの共通認識を共創することを試みる。

図1に共通認識の共創プロセスを FNS ダイアグラムで表現したものを示す。二人の主体が、石垣の写真を観て石垣について語り合うという共創経験の場において、写真と、および、互いに、〈インタラクション（C√2）〉しながら、それぞれの FNS サイクルを繰り返している。写真を読むことと相手の発話を理解することは〈分析（C2）〉に相当する。端正さと味があるということを意識して石垣の経験を想起することや石垣の端正さや味について詳しく語ることばを選ぶことは〈構想（C3）〉に相当する。選んだことばを実際に発話することは〈生成（C1）〉に相当する。

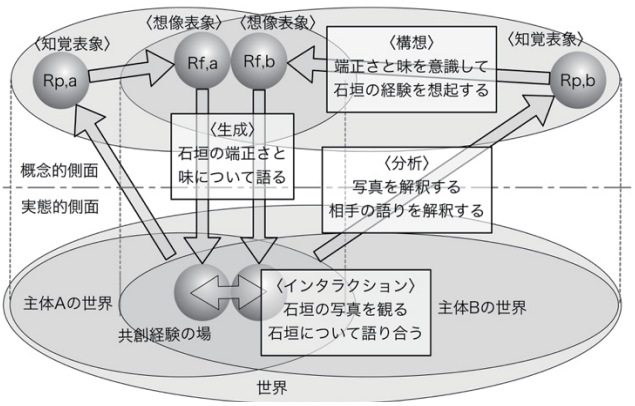


図 1 石垣の好ましさに関する共通認識の共創の FNS サイクル

(1) 石垣に直接関わる経験

石垣に関する石垣の好ましさの判断に直接影響を与えると仮定するこれまでの経験を以下に示す。

- 伝統的民家の実測と空間図式の顕在化（2014.9~現在）
- 地域創生デザインの課題のデザイン［藤井 2018a, 2018b］
- 伊是名地区石垣悉皆調査（2018.3）
- 石垣築造（第1回, 2018.11.14~17）
- 石垣築造の省察（第1回, 2018.11.19）
- 石垣築造マニュアルの作成（2018.11.19~2019.1.13）
- 石垣築造におけるパターンの抽出（2018.11.15~現在）
- 石垣築造（第 2 回, 2019.1.15）
- 石垣築造の省察（第 2 回, 2019.1.16）
- 石垣普請の価値共創の観点からの考察［平田 2019］
- 石垣築造マニュアルの改訂（2019.1.15~現在）

(2) 写真を参照しての石垣の好ましさに関する対話

石垣を正面から撮影した写真を参照しながら、石垣の好ましさについて議論を行なった。実在する石垣から「好ましい」とか「面白い」とかの印象を受ける経験を石垣という実体に結びつけて理解しようとする試みである。石垣築造において、村民や私たちが「好ましい」と思える石垣を築造しようとしている。そのような石垣を築造するための知を明らかにして伝承することを視野に入れている。

・対話セッション 1

琉球地方（主に伊是名島）で 2014 年 9 月から 2019 年 1 月までに撮影した写真の中から石垣の側面に正対する写真 104 枚を任意に選択し、各写真に撮影されている石垣について、「端正であるか」と「味があるか」ということについて、撮影されているものごとを指示しながら、藤井と篠崎が語り合った。本研究では、この語り合いを、「石垣が端正である」と「石垣に味がある」という言語表現の指示対象と意義に関する二人の主体の共通認識を共創するプロセスであると捉えている。共通認識を共創するということを意識的に行なうために、語りの中で、必ず、各石垣の端正さの度合いと味の有無を五段階の尺度で評価し、その理由を述べることにした。

・対話セッション 2

石垣が端正であるか、石垣に味があるかということに関して、平田が写真を参照しながら藤井に語るという形式で、その評価仕方を探った。藤井・篠崎セッションの 104 枚の写真の中から、端正さと味の有無が分散するように 19 枚の写真を選んで用いた。予め、平田には藤井・篠崎セッションの前半の 50 例について、会話内容をテキストのみで提示し、平田はそれに目を通している。会話が参照する写真は提示していない。平田の評価が藤井・篠崎セッションでの評価の影響を受けるような誘導的な会話はしていない。

3.3 石垣に関わる知の言語表現の型

対話セッションにおける発話の内容を石垣の端正さと味の有無を判断の仕方に注目して整理し、石垣に関する知の型を概観しようとしている。これまでに、以下の型を分類している。これらの型には、石垣という対象に依存せず、他の人工物に関する知の整理にも適用可能なものがあると直感している。

- 公理的な認識の宣言
- 典型の存在の示唆
- 情景の描写（肯定的、否定的、中立）
- 評価に関わる法則性の認識の宣言
- 比喩的な情景描写
- （石垣築造）行為への言及（肯定的、否定的、中立）
- （好ましくない石垣を改善する）行為の提案

これらのうち、石垣を築造する行為への言及や好ましくない石垣を改善する行為の提案は、石垣築造を参与観察し、石垣築造のマニュアルを作成するという、筆者らの経験を反映していると考えられる。これらの経験がなければ語れない内容が含まれているとしたら、その内容は筆者らが体得した知を示唆していると考えている。

以下に、各型について、表現例を添えて、概説する。

(1) 公理的な認識の宣言

石垣の価値や性質に関することと石の性質に関することが、恒真とみなす命題、すなわち、公理であるように、断定的に語られている。事実に紐づいた客観的な事実である場合と、石垣の価値のあり方に関する間主観的な認識を示していると考える。

- 端正に積むというのは石と石との良い関係ができるように丁寧に積むことである。(001)

(2) 典型の存在の示唆

石垣が端正であることや石垣に味があることを評価する際に、石垣の典型（archetype）があり、それと比較して現前の石垣を評価していることを示唆する表現が用いられる。典型の具体的な姿はこの型では示されていない。「情景の描写」型において示される典型の側面を総合する概念であると考える。

- 石垣感がある。(041)

(3) 情景の描写（肯定的、否定的、中立）

石垣が端正であることや石垣に味があることを石垣の実体的な特徴に結びつけて評価する際に、石垣の実体的な状態を語る表現が用いられる。「石垣が端正である」とか「石垣に味がある」とかに関して肯定的な評価をしている場合に評価の根拠として示されている状態を肯定的な情景の描写とみなす。肯定的な情景の描写の集まりが、築造において目指すべき、好ましい石垣の典型についての知の集まりであると考える。一方、好ましくない石垣の姿は、肯定的な情景描写を否定する内容によって語られている。

- 穏やかな水平な流れがある。
- 面が滑らかで平らな感じがする。
- 全体としてうまくまとまっている。

(4) 評価に関わる法則性の認識の宣言

石垣が端正であるとか石垣に味があるとかを判断する私たちの経験的な意味づけと石垣の実体的な構成とを結びつけるために私たちが認識しているものごとを示す。実体的構成から経験的意味づけを導出する法則性を示していると考える。

- 模様があると端正である。
- 印象的であると味がある。
- 水の流れを感じると端正で味があると感じる。

(5) 比喩的な情景描写

石垣の状態を石垣とは異なる情景に見立てている内容を示す型である。比喩を用いることによって発話者が注目している石垣の特徴を示していると考える。

- 木の葉が風で舞っている感じがある。(068)
- ここに珊瑚が生えているような生命感がある。(024)

(6) 行為への言及

主に石垣の築造における行為について言及している内容の型である。好ましい石垣を実現する行為を参照して、石垣を評価している。自ら石垣を築造したという筆者らの経験を投影していると考えられる。容易ではないと私たちが実感している行為がなされた結果として出来上がった石垣を行為的に評価している。端正な石垣を築造した行為を石垣の仕上がりを意識した丁寧な行為であると捉える一方、端正ではない石垣を築造した行為をそういう意識の薄い無造作な行為として捉えている傾向がある。

- 平らではない石を使って平らで滑らかな面をつくっている。
- 無造作に積んでいる感じがする。

(7) 行為の提案

「端正ではない」もしくは「味がない」と評価する石垣に対して、そのような石垣を築造した行為を好ましくない行為として推量し、当該石垣を改良する行為を提案する内容の記述である。自分たちが築造した石垣を行為的に評価しようとするも高い評価を与えられないことから、自分たちの築造技術を向上するためになすべきことにも言及している。

- 空いているところにとりあえず石を置いていったという感じがする。石が色々な方向性を示している。平たい石を積むならもう少し丁寧に水平に置けば良い。

4. まとめ

「伝統技術を伝承するとともに新しい科学技術と融合したこれからの伝統となる技術を共創する方法の好ましい姿を明らかにする」ことを目的とする探究の計画を述べ、その端緒として、石垣築造の経験に基づく知の顕在化の試みを紹介した。

謝辞

石垣築造に関わった伊是名村の方々と学生たちに謝意を表す。

参考文献

[Akin 1986] Omer Akin. Psychology of Architectural Design, Pion, 1986.

[藤井 2005] 藤井晴行: 建築デザインの論理的観点と非論理的観点を結合する二層モデル, 日本建築学会計画系論文集, No.591, pp.79-84, 2005.

[藤井 2017] 藤井晴行, 篠崎健一: 写真日記を作成することによる空間図式探究, 日本認知科学会第 34 回大会予稿集, OS04-5, 2017.

[藤井 2018a] 藤井晴行, 篠崎健一: 主観性と客観性を共存させる方法による隣地探究に基づく地域創生デザイン課題のデザイン, 2018 年度人工知能学会全国大会（第 32 回）予稿集, 2B2-OS-19a-03, 2018.

[藤井 2018b] 藤井晴行, 篠崎健一: ヴァナキュラーな文化に埋め込まれた知恵と近代科学技術の相克,日本認知科学会第 35 回大会予稿集, OS03-5, 2018.

[Gero 2000] John S. GERO, FUJII Haruyuki. A Computational Framework for Concept Formation in a Situated Design Agent, Knowledge-Based Systems, Vol. 13, No. 6, pp. 361-368, 2000.

[平田 2019] 平田貞代, 藤井晴行, 篠崎健一: 沖縄の伝統的建設技術の価値を継承するためのパターン・ランゲージの設計 :価値共創を顕在化するための記述方法の構築, サービス学会 第 7 回国内大会, 2019.

[中村 1992] 中村雄二郎. “臨床の知とは何か”, 岩波書店, 1992.

[中村 1993] 中村雄二郎. “デザインする意志”, エッセー集成 6, 青土社, 1993.

[Nakashima 2016] Hideyuki Nakashima, Haruyuki Fujii, Masaki Suwa: FNS Model of Service as Value Co-creation in Design Processes, Journal of Serviceology, 1(1), pp.6-14, 2016.

[Peirce 1887] Peirce, C. S. “Science and Immortality”, Charles S. Peirce: Selected Writings, Dover, New York, pp.345-379, 1887.

[Suwa 2015] 諏訪正樹, 藤井晴行: 知のデザイン - 自分ごととして考えよう, 近代科学社, 2015.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

実地の体験とマニュアル制作による石積みの気づきのプロセス

Process of Comprehension of Constructing Dry Stone Walls in Izena in Ryukyu

篠崎 健一[*1] SHINOZAKI Kenichi
平田 貞代[*2] HIRATA Sadayo
藤井 晴行[*3] FUJII Haruyuki

[*1] 日本大学 Nihon University #1
[*2] 芝浦工業大学 Shibaura Institute of Technology #2
[*3] 東京工業大学 Tokyo Institute of Technology #3

This article tries to present an experience of acquiring comprehension and technique through out two times on site constructions of Dry Stone Walls in Izena island in Ryukyu, Okinawa, and making an instruction manual for its construction. The analysis bases on; 1) fundamental technique of constructing Dry Stone Wall, 2) management of stone arrangement, 3) total flow of the process of the work, and 4) variation of the characteristics of the work. Through the experiences, it is revealed that an acquaintance with the terms such as "Tsumi-ishi (building stones) " and "Guwa-ishi (fill or hearing)" brings about a clear figure of a model section of the Dry Stone Wall and accelerate understanding about its whole construction.

1. はじめに

本稿は，琉球地方の伝統的生活空間を形成する珊瑚石の石垣の築造を通じて獲得する臨床の知の報告である。筆者らは，沖縄本島北部の離島伊是名集落にて，伝統的民家や集落空間における空間構成と生活に基づく空間図式の顕在化の研究を 2014 年から継続している。臨地の実地の経験を重視し研究室内では成しえない発見的アプローチをとることが特徴である。筆者らは，この臨地における活動と活動の拠点をリビング・ラボと称し研究を遂行している。

石垣と琉球の伝統的空間

琉球民家を囲い集落の景観を構成する珊瑚の石垣は，石垣に沿って植えられるフクギなどの樹木とともに，琉球の民家を暴風から護ってきた。この囲繞を屋敷囲いという。珊瑚石の石垣は，集落の人びとが協力して労働を交換し共同作業で積み継承してきた。用いる珊瑚石は，かつては，集落の南浜の沖の島から切り出しまた島で拾い，サバニと呼ばれる小さな手漕ぎ木造舟を二艘並べ舟の上に渡した板に乗せて運んだ。珊瑚石の石垣には，地域の共同体の結びつきの強さや，人びとが共有する価値が宿る。今日，集落の石垣は部分的に崩れたり全く崩れてしまった箇所が目立つ。石を積む労働の嫌悪や当時の新材料（コンクリートブロック）への憧れなどから，珊瑚の石積みは顧みられなくなり，石垣は崩壊し減少している。

集落の公民館前を東西に走る道より北側の地域は集落の後辺（くしひん）と呼ばれ，伝統的な民家（民家の形態や構法，空間構成，赤瓦などの伝統的材料を用いている）や，珊瑚石の石垣，フクギなどによる伝統的な集落景観が残る地域である。後辺地域の伝統的空間は，民官学いずれもが同様に重要と考える地域資源であり，これを継承するという課題が共有される。

リビング・ラボの活動を通して，1）後辺民家約 50 民家を実測し民家の内外の空間を図面化する，2）実測に合わせ民家における生活について聞き取る，3）集落の年間の祭祀を過去の文献記載と比較しながら記録する，4）集落の神アサギという信仰の中心の修繕（屋根葺替え）に参与し工程と技術を記録する，5）集落全体の石垣の状態を記録し崩壊部分を詳細に記録するなどしてきた。また6）一人称的な視点に基づき空間認識を図式抽出する過程から，集落空間についての気づきを構成的に導き出している。

本報告の臨地における石垣築造は，このような長期にわたる地域との関わりの中から，これまでの記録することの継続の上に，地域が失いつつある地域に内在する力を喚起する契機とならないかと考えている。

2. 石垣の築造体験

2.1 臨地石積み体験と石垣築造マニュアル

石垣築造は，第1回臨地石積み体験を 2018 年 11 月 14 日から 17 日の4日間に，第2回臨地石積み検証を 2019 年 1 月 15 日に，それぞれ伊是名村伊是名地区の個人所有地にて行った。第1回の石垣築造は全長 33m 高さ 1.4m 幅 0.8m の石垣を，重機を使いながら延べ 40 人日で行う比較的大規模の築造で，第2回の石垣築造は，下記の石垣築造マニュアルの検証を目的とした，相対的に小規模な築造で，長さ 3m 高さ 1.2m 幅 0.8m の部分的に崩壊した石垣を1日で積み直す，延べ5人日の実験的な築造である。石垣築造マニュアルは，第1回と第2回の石積み体験の間に，未来の石垣築造者に築造の仕方を伝えるために制作した。本稿は，

- (a) 第1回臨地石積み体験により獲得した気づきや理解
- (b) マニュアル制作により顕在化した気づきや理解
- (c) 第2回臨地石積み検証により獲得した気づきや理解

を，次の4つの観点から整理し，石垣築造の体験（a～c）による臨床の知の獲得を試みる。4つの観点は，(1) 石積みの基本技術，(2) 石のマネジメント，(3) 石垣普請の全体の流れ，(4) 石垣普請のヴァリエーションである。(1)を(2)石のマネジメントがサポートし，(1)と(2)を(3)石垣普請の全体の流れがサポートする，また(1)(2)(3)は(4)に内包されるという関係がある。(1)から(4)の相互に関連する観点の抽出は，この石垣築造の体験（a～c）の経験に依存する。これをマトリクスに示す（図1）。

2.2 石積みの臨床の知

(1) 石積みの基本技術

第1回臨地石積み体験で得られる気づきは，石積みの様々な基本の積み方についてがほとんどである（1a）。「石垣の水平面を強調する」，「下段の石と互い違いになるように積む」，「石垣の両端と中央は時々噛み合うようにする」，「面がある石は石垣の両端に積み、丸い石や小石は中央に積む（＊）」，「ひとつの石を7回転がして最もはまるところに積む」，「石がカチッと嵌るようにする」，「石垣の両端がやや高く中央が低くなるように積む」などである。根石の築造は体験していないため「根石はどっしり固める。ある程度経験がある人が作業する」，「石垣の根入れ深さは 200mm とする」などの学習以外の気づきはない（1a）。

マニュアル制作の効用は，わかったつもりで実はよくわからない疑問を明らかにすることにもある(1b)。「石を鉛直に積むのか上部を傾けるのか」，「門と角は特に美しく積むとは具体的にどう石を積むのか」などであり，また「石垣の部位や石積みの行為を表す用語を知ることの必要性」に気づく(1b)など，具体的な解決が求められる疑問が明らかになる。これを第2回臨地石積み検証にて熟練者に積極的に尋ね解決することが可能となる。

第2回臨地石積み検証で，「積み石」と「ぐわぁ石」という石の呼称を得る(1c)。「積み石」は石垣の外側に積む石で，「ぐわぁ石」は両方の「積み石」の間に入れる小さい石や小石である。第1回臨地石積み体験での気づき(＊)の「面がある石」と「丸い石や小石」はそれぞれ「積み石」と「ぐわぁ石」に相当すると理解する。マニュアル制作までは個別で相互の関係が希薄であった石積みの様々な基本の積み方(上に列記)は，「積み石」と「ぐわぁ石」という二つの石による，築造すべき石垣の明快な断面構成の獲得により，有機的に連続する石積みの技として理解される。

第2回臨地石積み検証では，自力で根切りし，根石を置き，かつ熟練者からのダメ出しと具体的な実技指導を受けたため，気づきや理解の数，内容の多様さ，奥深さが格段と増す(1c)。具体的には，まず「根切りは側面が鉛直になるように切る(掘る)」，「根石は，根切りの鉛直の側面に，内側から外側に向かって強く押しあてるように置く」ことなどである。熟練者の指摘と実技による状況修正以前は，この理想的な根切りと根石の状態に対する理解がなかった。理想的な状態を実現するための具体的な行為が実地に示されることの重要性に気づく(1c)。

(2) 石のマネジメント

積む石のマネジメントは作業効率を左右する。しかしそれは，効率の問題だけに帰着するのではない。第1回の臨地石積み体験は，既存の石垣を重機で崩し改めて積み直すため，積む石は大きさも形も渾然一体となり山に積まれた。そこでせめて「石垣築造の位置にあわせて積む石を置く場所を決める」，「積む石は石垣の両側(敷地の内外)に等しく配置する」ことを理解し(2a)マニュアルに記す(2b)。体験後の振返りでは，作業効率改善のため石のマネジメントについての言及がある。具体的には「石垣の表側を積むのに適する(美しく積める)平たい石を集めて置けないか」，「積む石を石の形で分類できないか」などである(2a)。第 2 回臨地石積み検証において，石のマネジメントの可否を検証した。具体的な方法と気づきは以下の通りである。

石垣を手で崩し，崩した石を，大きな石，平たい石，小さな石の三種に分けて置く(2c)。最初の気づきは，作業を開始して間もなく「模様のある顔となる(面となる)石を，これらとは別に置く」ことである。石を石の形状だけでなく「石の積む箇所」という概念に基づき分別することで石垣の全体像が捉えられる(2c)。

「石の量が想像より多く石を置く場所が当初より広く必要である」，「石を一目で見やすくするため重ねないで置きたい」，「ぐわぁ石の量が想像よりはるかに多い(ぐわぁ石の存在をあまり意識していなかった)」など，検証の気づきは内容，数とも多様で多い(2c)。道具についての気づきにも至る。例えば，石を崩す過程で「手箒があれば崩す石とその周囲をきれいにできる」，「絡まる蔓や植物を切る鎌があるとよい」，舗装された集落道に掘った土を広げ散布すると「石を少し投げて置くことができ作業効率が上がる(当初土で汚すことを躊躇した)」などである(3c)。

また石垣を手で崩すことは「思いの外無駄な時間ではない」と感じる。石垣崩しの過程で，これから積む石を一度手に取ることは，石を積むことに何かの寄与があるように体感する。」崩すことが積むことに直結すると気付く(石と身体が紐づけられるような感覚)のである。

(3) 石垣普請の全体の流れ

第1回臨地石積み体験後の振り返りで，直接的な石積み行為の外の気づきが得られる。例えば「作業の役割分担やチームを作るとよい」，「マニュアルや技術紹介ビデオが有効」，「安全対策をする(日中夜間)」だが，「道具の用意」や「休憩場所の設置」，「アガリジャー(作業後の慰労の食事や飲み物)の準備」など，実際に体験しているにも関わらず気づきから抜け落ちるものもある(3a)。これらはマニュアル制作時に補完され，「作業の流れを皆で共有する」，「目標とする石垣のイメージを共有する」などの気づきと共に石垣普請の全体の流れが共有される(3b)。

第2回臨地石積み検証で補修した石垣の既存部分の蔓の絡まりや空き缶やビニルなどのゴミを掃除すると，石垣全体の存在感がずっと増すことに気づく(3c)。石垣を手入れしてきれいに保つことは，石垣の存在の大切さに集落の人が自ら気付く契機となり，伝統の継承につながるのではないかと気付く(3c)。

(4) 石垣普請のヴァリエーション

重機に頼らない人力による石積みの可能性が実証された(4c)。第1回石積み体験の大規模な普請(4a)ではなく，部分的な石垣補修の小規模な普請を，小さな(4人)作業ユニットが1日で完成させられる。このような石垣補修は，日常的に例えば週末などに少人数で行えるため，集落の石垣築造及び築造技術の継承の可能性が高まると期待する(4c)。これに関し，リビング・ラボが 2017 年 3 月に報告した，集落全体を対象とした石垣崩壊箇所に関する調査報告が有効であると改めて理解する。

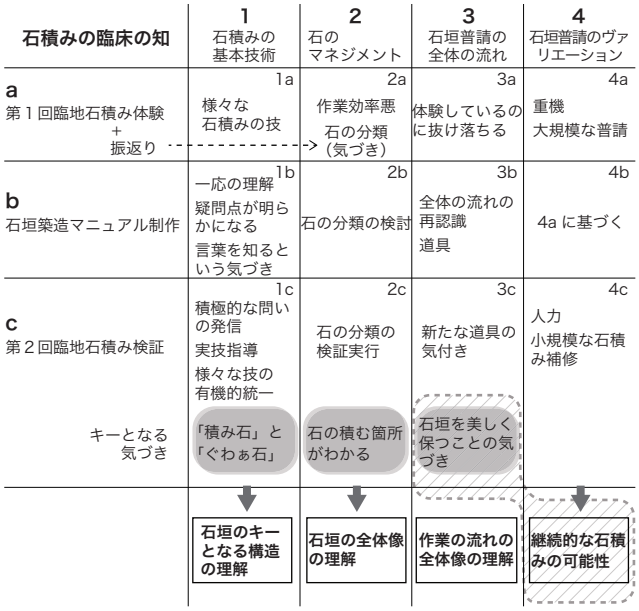


図1 石積み体験と石垣築造マニュアルによる気づきのプロセス

3. まとめ

伊是名地区における二回の臨地石積み体験と石垣築造マニュアルの制作を通して，単に石積み体験を重ねることによる技術獲得以上の，全身体による様々な段階の臨床の知を獲得する可能性を示した。知は，手元の基礎的技術から，石積みをサポートする石のマネジメント，その周縁の準備行為，さらに集落の伝統的空間継承の可能性へ拡大しながら発展する。

参考文献

[真田 2018] 真田純子:図解 誰でもできる石積み入門，一般社団法人 農産漁村文化, 2018.12.15.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

触れる鑑賞、味わう鑑賞にむけて

Toward the Aesthetic Appreciation of the Haptics and Tastes

福島宙輝[*1]
Hiroki Fukushima

[*1] 九州女子大学
Kyushu Women's University

Abstract: In this paper, the author considers the feasibility of the aesthetic appreciation of the haptics, smells, and tastes. Generally, these sensory domains are not regarded as the subject of the aesthetics mainly because they do not have an artefactual object. Here the author does not give a conclusion for the question whether they could be the subject of the aesthetics or not. Instead, the author illustrates some examples of the "haptic aesthetics" through the Japanese teacups, and gives some examples of the "aesthetic concepts of the taste" through the descriptions of the taste of Japanese sakes.

1. はじめに

1.1 感覚と鑑賞

本稿を通して考えていきたいのは，味覚や触覚の鑑賞，あるいは視覚的な芸術でも日本を含む東アジアの書（書道）のような，触感を色濃く含むような対象の鑑賞はどのように行うことができるかということである．鑑賞は，美的なものを趣味（taste，対象の味わいや情緒を感じ取る能力）をはたらかせながら見て，その良い部分，美的質を受容するという行為である．ただし，ここでは味覚が美学の範疇，対象に入るかについて議論するつもりはない．というのも，一般的には，味覚や嗅覚，触覚といったものは美学（近代美学）の範疇には入らないとされている．悲しいことに哲学としての美学において，ルネサンスと宗教改革による価値の相対化を経て，近代美学は視聴覚に偏重し，美的価値を視聴覚のみに認める方向に進んでいくこととなる．これは中村(2000)が近代科学がルネサンスを契機に視覚に「専制支配」されてしまったと主張することと通底するものである．

1.2 言葉による鑑賞

美しいとはどういうことか，美的な質をどのように鑑賞するかという問題にたいして，言葉を道具立てとして美的性質を解釈していくという手法は，シブリー(Sibley, 1959, 2001)を始点とする．

シブリーが Aesthetic Concepts という論考で問いかけたのは，絵画を鑑賞して言葉で解釈を加えたり，鑑賞のポイントがわからないような人に説明したりするときに，どのような言葉が用いられるかということである．シブリーは美的概念を表現する語彙を，美的用語 aesthetic terms と呼ぶ．シブリーは，美的用語，あるいは美的表現とは，趣味や鋭敏な知覚，感受性，美的識別の能力，そして鑑賞眼を必要とするような用語ないし表現であり，それらに対応するものが美的概念であるとする．具体的には，バランスの取れた，力強い，生命感に満ちた，繊細な，といったような用語である．そしてこのような美的概念を表さないようなもの，例えば，青いとか，まっすぐな線といったような表現は（感受性がなくても見て取れるので）非-美的表現と呼ばれる[1]．

本稿では，この美的概念をひとつの指針とし，特に日本酒の味わいにおいて美的表現が見受けられるかを議論していきたい．

2. 触れる鑑賞

触れる鑑賞を考えるときに，やはり陶磁器というのがその中心に座るのではないかと思う．西洋にも優れた磁器は数多い．多くのテーブルウエアが美術品として各地の美術館にコレクションされている．しかしやはり白磁の湯呑で呑むときと，（もちろん筆者が使うようなものなので値段は大したことはないが）ティーカップで飲むときと，同じお茶であっても，その味わいがどうも異なっているように感じられる．

これは筆者の感覚であるが，日本や朝鮮，中国の茶盌（ちゃわん，ご飯茶碗と区別するために茶盌と表記しておく）と，西洋のティーカップの違いは，その持ち手の部分に関係があるのではないかと思う．日本の湯呑，茶盌は持ち手がなく，包むようにして持つのに対して，ティーカップは持ち手の部分をつまむようにして持つ．欧米のマナーでは，ティーカップの持ち手の部分ではなく本体を持つと，マナー違反（ぬるくて飲めないという意図と解されるという）とされる．

この違いは，単に持ち方の違いにとどまらず，身体性をともなってその芸術的価値にも影響を与えているのではないだろうか．

茶盌のみならず陶磁器の愉しみは，やはり視覚と触覚との融合，中村(1979)のいう「共通感覚」を抜きにしてはやはり語りえない．萩の茶盌のぬくもりは，やはり手で包んで，その枇杷色の土，柔らかな焼きの質感，持ち上げたときの軽さ，指先で感じるカイラギの凹凸によってはじめて得られるものであって，写真では感じ得ないものである．天草に「陶房泰」という窯があるが，その窯の白磁は，白磁であるにもかかわらずその外側はザラッとしている．このざらつきは，視覚上の陰影と言うよりもむしろ，手触りと，飲むときの感触に大きな働きを見せる．陶房泰の湯のみに氷を入れ，グレープフルーツサワー（キリンの本搾り）を飲んでみてほしい．グレープフルーツの苦味と，下唇で感じる陶器のざらつきがまさに共感覚的に呼応しあい，そして上唇では湯呑のなめらかな内側の質感が，氷の冷たさとともに感じられる．この美的経験を，近代美学はなぜ見逃すのだろうか．

西洋のティーカップの芸術的作品とされるものを見るだに，やはりどこかカップ自体が絵画のキャンバスとして切り離されてしまっているように思える．惜しくも絶版となった「古美術読本」の第一巻，陶磁編(芝木編, 2006)をひらくと，芝木好子，大佛次郎，

[1] 念のために補足しておくと，シブリーは別の論考（Tastes, Smells, and Aesthetics）において，味覚や嗅覚は低次で，動物的で，役に立たない，すぐに消えてしまい，変化しやすい感覚のであるから美学の対象とならない，といったような議論は誤解であり，味覚や嗅覚も美学の対象となりうるということを力強く主張している．

谷川徹三ら錚々たる著者陣がその触覚の美を説いている．大佛はいう．「日本の陶器にいのちをこめているものは，手に持ったり肌に触れて試みるやきものに対する情愛だと思う．それにくらべてセーヴルもデルフトも，見るもの，飾り物なのである．外にある器物にすぎない．博物館のガラスケースの乾いた空気の中にある参考点か，個人の趣味の蒐集品でしかない．」

筆者は，五感の統合的な美的鑑賞のあり方の研究があっても良いのではないかと思い，陶磁器や書道の鑑賞，酒の鑑賞の研究を進めているところである．

3. 味の鑑賞について

日本酒を鑑賞するとき，あるいは言いかえれば，ある日本酒の特徴や美味しさを反省的に捉えようとするときに，言語による表現は最も基本的な道具立てであり，鑑賞のための手段でもある．我々は一杯の酒を呑むときに，その酒に感じた特徴を言語化していくと同時に，その反作用として，言語のフレームを介することによって，新たな味に気がつく．このような，酒と自らの身が言葉を介して融け合うようなダイナミックな営みが，酒を鑑賞するという営為である．

3.1 味わう言葉の経緯

それでは，日本酒には，その味わいを鑑賞することばとしてどのような用語があるのだろうか．

日本酒はワインと異なり，味わいをことばで表現していくという伝統が薄いといえる．これはワインのテイスティングワードが生産地のフランスではなくイギリスの商人の力で発展したことと通底する．各地方でワインが醸造されるフランスでは，酒屋にボトルを持っていって自分の地元のワインを量り売りしてもらうというのが基本的なスタイルであり，特に他の地方との味の差を気にする必要はなかった．一方で，自国でワインを生産する環境に乏しいイギリスは，フランス，特に 12～15 世紀において英国領であったボルドー地方のワインを売買することとなる．この際に，イギリス本国のマーケットで売買するために，銘柄ごとの差異を語るための言葉が発展したというのがその背景のようである．

日本酒もフランスのワインと同様，その地域のいわゆる地酒を文句なく呑むのが普通であって，現在のように日本各地の酒が手に入る状況は，1980 年代のいわゆる地酒ブームや純米酒人気，90 年台の吟醸酒人気を経て，ごく最近になってからのことである．こうした背景から，日本酒はそもそも味わいを言語によって鑑賞していくという習慣を，伝統的にもたなかったのである．

日本酒の魅力的な味わいを表現する文化がないという背景のもう一つは，日本酒の醸造における徹底した品質管理にあるといえる．日本酒は，世界の醸造酒の中でも非常に高度な醸造技術をもつ酒類である．その一端を垣間見ることができるのは，醸造年度ごとの味のブレの少なさであろう．

例えばワインは，その年の気候によってブドウの出来不出来がはっきりとあらわれ，できあがるワインの品質にもダイレクトに影響を与える．しかし日本酒は，あまりヴィンテージ，つまり良い年と悪い年の差がなく，毎年，一つの銘柄であればほぼ同じような安定した味を作り出す．当然，ブドウと同じように米にも豊作不作，台風や猛暑の影響によるできばえの良し悪しは毎年のように発生している．しかし一定の範囲の味に安定してコントロールできるのは，ひとえに日本酒の高度な醸造技術の賜である．

この高度な品質管理を支えるのが，醸造を行う蔵人と，そのリーダーである杜氏（とうじ）の卓越した官能評価能力である．微生物を相手とする醸造において刻々移り変わる酒の様子や，品質の良し悪し，醸造環境の変化や劣化をいち早く検出し，その悪因を排除することは，杜氏にとって最も重要な仕事である．このような背景から，実は日本酒においてはその魅力的な味わいを表現することよりもむしろ，「出してはいけない香り（オフフレーバー）」を言葉にすることが重視されてきた．以下では，このような背景を考慮に入れた上で，日本酒の鑑賞に資するような美的用語の可能性を検討していく．

3.2 新酒鑑評会審査カードにおける評価用語の検討

まずは日本酒の味わいを評価するスタンダードから議論していく．ごく近年になってから，各種の日本酒コンテストが国内外で開かれるようになってきたが，伝統的に醸造技術を支える鑑評会として機能してきたのは，酒類総合研究所および日本酒造組合中央会による全国新酒鑑評会である．日本酒鑑賞へ向けての議論の入り口として，ここでは全国新酒鑑評会で用いられる「新酒鑑評会審査カード」（図 1）の用語を参考にする．なおこのカードは 2006 年以降に使用されているバージョンである．新酒鑑評会では，大きくわけて日本酒（清酒）の「香り」と「味」，そして「総合評価」が評価対象となる．

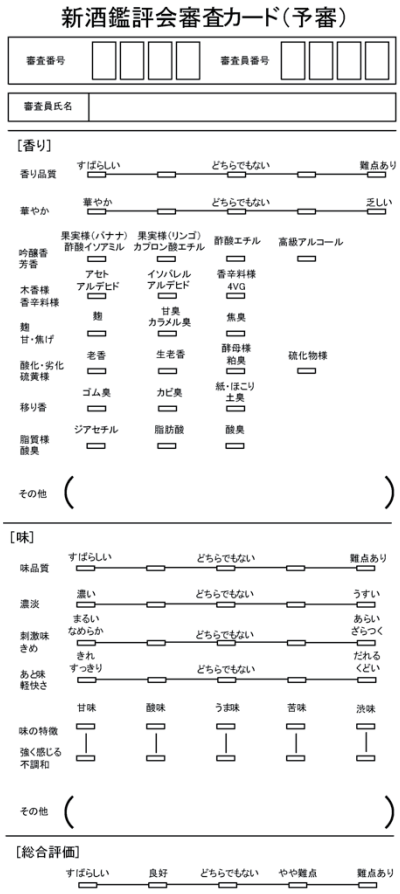


図 1：新酒鑑評会審査カード（2006 年版）

(1) 五味は非-美的表現なのか

まずは最も基礎的な用語，基本五味（甘味，酸味，うま味，苦味，渋味）について考えてみる．

シブリーの議論では，美的概念は特殊な能力や趣味の働きでもって見出されるものであるため，絵画における絵の具の「色」や描かれている「線」は非-美的概念とされる．これは，絵画における「色」や「線」はおよそ健常な視覚を持つものであれば誰でも見てとることができるためである．絵画において色や線が美的概念ではないということに沿って考えれば，日本酒の味覚における「甘味」や「旨味」は，これは美的概念ではないということに

なるだろう．甘味や旨味は，それを構成する物質に還元して考えることもできるし，味覚センサーを用いて定量的に評価することもできる．

しかし，日本酒の味わいにおける五味は，本当に無批判に非-美的概念として考えて良いのだろうか．

ゴッホの描いた「ひまわり」が，「黄色を主体としていて中心は黄緑色である」と言うためには，健常な視覚をもってすれば，特段の能力や趣味を働かせる必要はないだろう．しかし東一の純米酒を味わって，「旨味が主体だが中心に酸味がある」と指摘するのは，やや特別な能力が働いているように感じられるのではないだろうか．日本酒を初めて呑む人にとって，「甘味がある」「酸味がある」と指摘することはかなり困難である．

先述のとおりシブリーは，「知覚力，感覚力，そして美的識別能力もしくは鑑賞眼」の行使を要するか否かによって美的質と非-美的質を区別する．そうすると，（シブリーはもちろん味覚は美学の対象に入ると主張しているが）味覚は五味というその最も基本的な要素の段階から，視覚や聴覚を基軸とした現行の美学の単純な概念移植では議論が進まないという難しさを孕んでいる．そもそも一般には美学の対象と目されていない味覚について議論するのであるから，一足飛びに「味覚における美的概念とはなにか」を問うことは早計と思われる．そこで以下本稿では，味覚の鑑賞のための用語としてどのようなものがあるかという事例を紹介したいと思う．

(2) 香りと味の表現

新酒鑑評会審査カードの香りの表現を見ていくと，わずかに「華やか」が美的用語として考えられる余地をのこしてはいるが，その他の項目は，ほぼ全て香りの同定的検出のための用語であり，非-美的なものと考えられる．

味についてはどうだろうか．濃淡については，乾いた（価値判断を伴わない）用法の場合には非-美的用語といえる．ただし「濃厚な」あるいは「淡麗な」というように，価値を伴うような語になると美的な用法が現れてくる．

あと味の「きれ／すっきり」は美的用語といえる．ある酒にきれ（キレ）があるかどうか，あと味がスッキリしているかどうかを判断するのは，対象の微細な差異を判定するための能力，感受性や判断力，趣味を必要とする．

刺激味の「まるい」「なめらか」も，日本酒の表現として用いられた場合には美的な用法であることが多い．この両者はそれぞれ視覚と触覚からの共感覚的隠喩に基づく表現であり，こうした隠喩的な転移に基づく美的用語は，本来の絵画や音楽についての美的概念にも数多くみられる．

味の調和については，現行の新酒鑑評会審査カードでは「不調和」が評価項目となっているが，2002 年から 2005 年にかけてに使用されていたバージョン（図 2）では，「調和」が香りの特性項目に挙げられていた．味が「調和している」，「バランスがとれている」という表現も，美的表現と考えることができる．シブリーは，この「バランスのとれた」という表現も，隠喩的な変換を経て美的用語になり，いまや標準的な語の位置づけを得たものとして例に挙げている．

3.3 日本酒のレビュー表現における美的な用語

ここまで検討してきたように，現行の新酒鑑評会審査カードでは，その評価項目は非-美的表現が中心となっている．一方，以前のバージョンでは，「調和」「ソフト」「華やか」「優雅」「濃醇（味わいが濃く，複数の味の調和による深みがあること）」「軽快」「きれい」「なめらか」「しまり」などの美的用語として使用されるであろう語が多く設定されている．なお以前のバージョンから現行のものへ移行する際の用語の選定については，宇都宮(2006)の論文に詳しい．

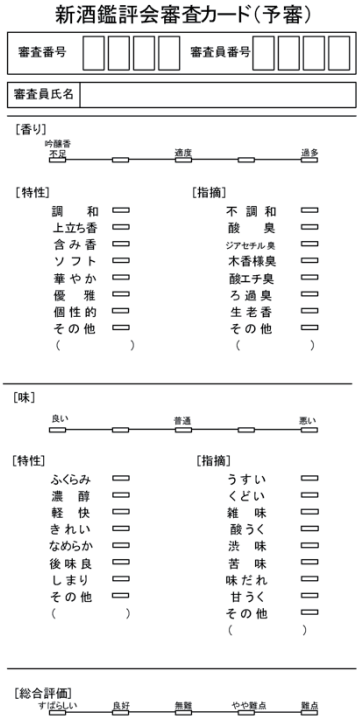


図 2 新酒鑑評会審査カード(2002-006 年)

本稿では一足飛びに日本酒の，あるいは味覚の美的用語を定義することまでは目論んでいない．いま検討したいのは，日本酒の鑑賞を行うにあたって，どのような美的用語（あるいは日本酒の魅力を表す語）があるかの概観である．

まずは，日本酒の味わい表現においてどのような語がよく用いられるかを，雑誌，書籍および味覚言語化実践研究によって収集した日本酒の味わい表現データ（10 万語レベル）をもとにして検討したい．味わい表現データ中に頻出する主な品詞の上位語は，表 1 のとおりである．

表 1

	擬音語	形容詞	形動	動詞	
1	ほんのり	強い	きれい	広がる	味
2	ゆっくり	硬い	穏やか	飲む	酒
3	しっかり	高い	爽やか	残る	舌
4	ゆったり	辛い	なめらか	立つ	香り
5	きりっと	丸い	透明	思う	甘味
6	あっさり	薄い	ふくよか	引く	酸味
7	すっと	太い	緩やか	入る	口
8	とろり	柔らかい	まろやか	消える	奥
9	すっきり	濃い	華やか	ふくらむ	酸
10	ピリピリ	香ばしい	スマート	含む	口中
11	ぐっと	長い	複雑	いう	旨味
12	ふっくら	優しい	シャープ	帯びる	印象
13	さっぱり	若い	スムーズ	落ち着く	苦味
14	じわじわ	力強い	濃密	出る	喉
15	ふわっ	おとなしい	フレッシュ	現れる	香気
16	まったり	熱い	滑らか	漂う	最後
17	やんわり	甘酸っぱい	端正	抜ける	大吟醸
18	さらり	厚い	伸びやか	見せる	香味
19	ザラザラ	淡い	素直	熟す	純米
20	たっぷり	弱い	軽快	続く	辛味

(1) 名詞

表 1 をみると，名詞には基本的な味を表す語（甘味，酸など）や，口中の身体部位（舌，喉），そして大吟醸や純米といったような規格を表す語が目立つ．鑑賞という点からすると，名詞には非-美的用語が多く見受けられ，日本酒の味わいの要素は名詞的に表現されることが見て取れる．

(2) 形容動詞

美的用語はどこに見られるかというと，形容動詞および形容詞に特徴的に現れている．

形容動詞では，きれい，穏やかを筆頭に，上位 20 語のすべてを美的用語としてみなすことができるだろう．注釈が必要な語を解説すると，「透明」とは，もちろん酒の色調を表すこともあるが（その場合には非-美的用語かもしれない），主には「酒の味の透明感」を表しているものと思われる．透明感とは，酒に無駄な味や雑味がなく，淡麗でありつつも味わいのある酒を表す用語で，一種の隠喩表現（視覚からの共感覚比喩）である．「複雑（複雑味）」は，まれに短所を指摘する用法もあるが，「単調でなく，いろいろな味が折り重なって重層的に感じられる」という美的な質を表す用法としても用いられる．「伸びやか」は酒の旨味がまとまりをもちつつも口の中で広がろうとするさまで，これも一種の隠喩表現である．

(3) 形容詞

続いて形容詞は，やはり大まかに見ると美的表現が多く見受けられる．おそらく文句なく美的表現としてカウントできるものは，審査カードにも挙げられていた「丸い」を筆頭として，「柔らかい」「濃い」「香ばしい」「優しい」「力強い」「厚い」（おもに旨味を指して，濃厚で舌の上に乗っている感じが強く得られる酒のさま）などが挙げられるだろう．ここで難しいのは，いま挙げた「力強い」や，その他にも「若い」「おとなしい」「淡い」などの語のように，実はやや欠点含みの内容を美化して言うという表現の扱いである．「力強い」は主に長所を言うが，まれにアルコール感が強い，というニュアンスを持つ．また若い酒というと，熟成が足りないという意味を含んだり，おとなしいとか淡いというのは，やや迫力に欠けて物足りないということを婉曲的に，美化して表す言い方である．この婉曲美化はワインに特徴的で，例えば「豊かな酸味」というと一見，美的表現にみえるものの，実は「酸味が強すぎて飲みづらい」という意味内容をもっていたりということがある．こうした婉曲美化表現に美的用語としてどのようなポジションを与えるかは，今後の大きな検討課題である．

(4) 動詞

形容詞，形容動詞については，やはり本来の品詞としての働きからしても美的用語が多く見られたことは納得のできる結果である．しかし美的用語は，形容詞や形容動詞などの修飾語に限られたものではない．日本酒の表現で注目したいのは，動詞やオノマトペに美的用語（と解釈できそうな語）が数多く見られるという点である．

動詞で注目できる語は，広がる，残る，立つ，引く，消える，ふくらむ，落ち着く，出る，現れる，漂う，続くという語である．広がるとふくらむ，漂うは，口の中での液体や味わいの動きの表現である．香りが口の中に広がったり，空気感を伴って香気が広がることをふくらむ，と言い習わす．

また，舌の上ではなく口腔の上の方で香りが滞留する様子は，漂うと表現される．これらは口や鼻の空間での味や香りの経時的変化を捉えないと表現できないものであり，甘味や酸味などの識別力とは異なる，時空間上の動きの変化を捉える鑑賞能力を要するという点で美的概念に相当する可能性があるだろう．

動詞は，とりわけこの時空間上の変化の表現が重要である．「立つ」「出る」「現れる」はそれぞれ香りが現れる瞬間を捉える表現であり，「引く」「消える」「落ち着く」というのは反対に味が感じられなくなる変化を表す（「残る」は味が消えないことを示す）．

絵画の美的用語でも「バランスがとれた balanced」というのも動詞的な表現であり，味わいの表現においても動詞による美的な質の表現は注目に値する．

(5) オノマトペ

オノマトペは，「ほんのりと香る」というように動詞を修飾する副詞的用法を中心として，「サラサラした舌触り」のように<オノマトペ+スル>の形で形容詞的な用法があり，また「舌先のピリピリが気になる」というように名詞的に主題化されることもある．

オノマトペというと，車を「ブーブー」というように，幼児の使うような幼稚な言語使用かのような印象を受けるかもしれない．しかしこれは全くの誤解であり，一般的な言語使用においても，とくに感覚的なものを表現するときには必須の要素である．

味覚の言語表現においては，感覚と言語をつなぐことの難しさが他の知覚領域以上に顕在化する．筆者はこれまでの研究において，認知的な際立ちの小さい味覚感覚の言語化に際して，オノマトペが共感覚的比喩として用いられることで言語化の橋渡しが達成されるを明らかにしてきた．分析の詳細は研究論文(Fukushima, Imai, & Tanaka, 2017)に譲るが，結果として，日本酒およびワインの味覚表現におけるオノマトペの使用原理として以下の三点を明らかにしている．

- 醸造年度，価格などの定量的な表現や状況表現には用いられづらい
- 「味わい」比較的上位カテゴリの味わい表現を修飾するために用いられやすく，ワインの個別のテイスティングワードのような具体的な要素と一緒に現れることは少ない
- 味わいの“状態”よりも“変化”を表現する

これらの結果は，ソムリエやプロテイスターのように味覚表現に熟達した者が用いるような，一見体系的で慣習的な味覚表現においても，なおもって表現の難しい感覚領域があり，その言語化の困難さを克服する言語方略としてオノマトペが用いられることを示すものである．

参考文献

Fukushima, H., Imai, M., & Tanaka, S. (2017). The Usage Mechanism of Japanese Ideophones in the Description of Taste: Morphological and co-occurrence analysis of the description of wines and sakes. International Journal of Computational Linguistics Research, 8(3), 109–122.

Sibley, F. (1959). Aesthetic concepts. The Philosophical Review, 68(4), 421–450.

Sibley, F. (2001). Approach to aesthetics: Collected papers on philosophical aesthetics. Clarendon Press.

中村雄二郎. (1979). 共通感覚論: 知の組みかえのために. 岩波書店.

中村雄二郎. (2000). 臨床の知 (Vol. 2). 岩波書店.

井上靖. (2006). 古美術読本 (1)「陶磁」. 知恵の森文庫.

宇都宮仁. (2006). 清酒の官能評価分析における香味に関する品質評価用語及び標準見本. 日本醸造協会誌, 101(10), 730-739.

Organized Session | Organized Session | [OS] OS-16

[2K5-OS-16b] 臨床の知(2)

諏訪 正樹（慶應義塾大学）、藤井 晴行（東京工業大学）、加藤 文俊（慶應義塾大学）

Wed. Jun 5, 2019 5:20 PM - 6:40 PM Room K (201A Medium meeting room)

[2K5-OS-16b-01] Construction of Design Community by designer and users of place for expression

〇maiko kobayakawa[2,3], takeshi sunaga[1], sunao maruyama[1], tomoki hirano[2], kosuke yamada[4], kakuichi nishimura[3], kentaro watanabe[3], sachiko fujimitsu[5] （1. Tokyo University of the arts, 2. Tokyo University of the Arts Graduate School, 3. National Institute of Advanced Industrial Science and Technology, 4. Keio University, 5. Saga University Hospital ）

5:20 PM - 5:40 PM

[2K5-OS-16b-02] On learning through metaphorical imaginations

〇Fumitoshi Kato[1] （1. Keio University）

5:40 PM - 6:00 PM

[2K5-OS-16b-03] Segmental response of the human body to changes in the interrelationship with the environment

〇KAZUYA TANAKA[1], Tomoaki Atomi[2], Wataru Takano[3], Soichiro Fujiki[4], Katsuya Hasegawa[5], Miho Shimizu[6], Yoriko Atomi[6] （1. Teikyo University of science, 2. Kyorin University, 3. Osaka University, 4. Dokkyo Medical University, 5. Japan Aerospace Exploration Agency, 6. Tokyo University of Agriculture and Technology）

6:00 PM - 6:20 PM

[2K5-OS-16b-04] Savage Experiment

〇Takahito Horiuchi[1], Masaki Suwa[2] （1. Graduate School of Media and Governance, Keio University, 2. Faculty of Environment and Information Studies, Keio University）

6:20 PM - 6:40 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

表現の場の作り手と使い手によるデザイン共同体の構築

Construction of Design Community by designer and users of place for expressive activity

小早川真衣子*1*3 Maiko Kobayakawa　須永剛司*2 Takeshi Sunaga　丸山素直*2 Sunao Maruyama　平野友規*1 Tomoki Hirano

山田クリス孝介*4 Kosuke C Yamada　西村拓一*3 Takuichi Nishimura　渡辺健太郎*3 Kentaro Watanabe　藤満幸子*5 Sachiko Fujimitsu

*1 東京藝術大学大学院 Tokyo University of the Arts Graduate School　*2 東京藝術大学 Tokyo University of the Arts

*3 産業技術総合研究所 National Institute of Advanced Industrial Science and Technology　*4 慶應義塾大学 Keio University　*5 佐賀大学附属病院看護部 Saga University Hospital

The purpose of our research is to find the way of design doing and thinking for that expressive activities are emergent autonomously for creating a society that regard people's expression as important too. We reflected our co-design project, researched about what happened and way of relationship between designer and user(people) on the process. As a result, we are finding that designer and people have to construction mutual constituent relationship, that is "becoming a Design Community" is very important for the expression activity to become independent. In this paper, we will report the state of the design community found in the case of project with nurses. In addition, to create that relationship, we discuss the idea that the design of the mechanism of "becoming" is encompassed in issue of designing.

1. はじめに

現代社会が抱える諸問題は，多様で複雑さを増している．それは，合理性を重視して細分化，先端化した各専門分野の能力や知識によって一挙に解決できるものだけではない．そこで着目されているのが，異なる立場や能力をもつ人たちが互いの力を合わせて自分たちの問題を発見し，解決に向けて何かをつくり出してみる共同デザイン（Co-design）のプロセスである．このアプローチは，誰もがデザインできるという考え方に立脚している．そのため，その実践においては，これまでに社会の役に立つと思われる物事をつくり出し，使い手に提供することを考え行なってきたデザイナーは，自らの専門性を問い直し，従来の役割や振る舞いを変革することが必要不可欠となる．

2. 研究の背景と目的

2.1 デザインの問題領域を拡張すること

これまでに著者は，表現の場をつくることをデザイン実践として行ってきた．そして，その実践を改めて見つめることから，表現の場をつくる際に不可欠な，表現の活動プログラムのデザイン方法を探究してきた．活動プログラムとは，いつ，どこで，誰が，何をどう行うのかという活動計画である．活動は，道具の機能や形を決めるために構想し考慮される対象である．しかし，それは形づくる対象とされてこず，道具が使い手に十全に利用されない事態を招く要因のひとつになってきた．活動プログラムを対象にすることで，これまでのデザインの問題領域を拡張させ，デザインそのものを問い直しが可能になると考えている．

連絡先：小早川（小原）真衣子，国立研究開発法人 産業技術総合研究所 人工知能研究センター，東京都江東区青海 2-3-26，kobayakawa.maiko@gmail.com

2.2 社会づくりの原動力としての表現活動

表現とは一般的に，他者の目に触れる前提で表出されるものから，日常的に自分の感情や意思を示す言葉や態度まで幅広く含まれる．中でも著者が着目するのは，自己省察を伴う表現行為を，複数人が共同して行う活動であり，それは豊かな社会づくりの原動力になると考えている．なぜなら，その営みは，人が自身の表現の意味や価値を発見することを促し，自分の表現と社会のつながりを確認できるコミュニケーション過程だからである．そこに個人が社会に対するあり方を見直し，効力感を手に入れる契機となる可能性がある．

本研究の目的は，そのような表現も重視する豊かな社会づくりに向けて，表現活動が自立的に生まれるようなデザインのやり方や考え方を明らかにすることである．そこで，表現活動が生まれている現場＝表現の場の作り手である著者らと，そこに表現活動に参加する使い手との共同デザイン過程をふり返り，デザインしたことや起きたことについて調査し考察した．その結果，表現活動が自立するためには，両者が相互構成的な関係性，すなわち「デザイン共同体になること」が重要であることが明らかとなった．本稿では，事例に見出したデザイン共同体の状態について報告する．加えて，その関係性を構築するためには，人が何者かに「なること」のメカニズムのデザインが内包される必要があるという考え方について論じる．

3. 表現の場の作り手と使い手の共同デザイン

3.1 看護の仕事を変革するプロジェクト

(1) プロジェクトのねらい

病院での看護の仕事の新しいあり方をデザインするために，現場の実践者である看護師たちとの協働に取り組んでいる．2012年に開始し，現在も進行中のプロジェクトである．大学病院の看

護部の看護師長数名（以下，使い手）と心理学系の研究者，サービス工学を専門とする研究者，デザイン系の研究者からなる研究グループ（以下，作り手）が参加しており，学際共同研究「MED プロジェクト」と呼んでいる[藤満 2013,須永 2013，渡辺 2015,]．このプロジェクトでテーマとなっている看護の仕事の新しいあり方のデザインとは，病院に次々に導入される情報システムの利用によって失われつつある「看護の心」を表出し共有する仕組みと文化をつくることである．ここでいう「看護の心」とは，看護師たちが通常業務で患者に対面・対話し看護処置にあたる際に欠かせない，患者やその家族，そして仲間を想う気持ちである．

(2) 自立した表現活動

2016 年，さまざまな取り組みを経た後，看護の仕事の中で，心に残った体験を作文し朗読し，その後にそれらを貼り合わせて新聞をつくり鑑賞する表現活動プログラム「新聞ワークショップ」をデザインした（図1）．新聞ワークショップでは，日々の業務の忙しさに追われている看護師らが，自分の思いや気づき，喜びや不安を表出し共有する．この活動を通して，看護師たちは普段は気がつかない看護の仕事に欠かせない心象の側面に，お互いに気がつき，自らをエンパワメントさせることができている．

図 1. 新聞ワークショップの様子（左：個人の作文を貼り合わせる．右：タイトルが付けられた新聞を他グループに見せる看護師）

当初，著者らがこの表現活動の基本プログラムをデザインしていたが，その後に看護師らが改良デザインの過程に参加できるように促した．その結果，現在では，自立的に基本的な活動プログラムをもとに実施の計画をし，準備をし，ワークショップ中のファシリテートもできるようになった．現在，この表現活動は，看護師の手によって病院内に広がりつつある．

本事例研究については，表現活動の詳細やプロジェクトの過程について，これまでにも報告してきた[小早川 2018]．ここでは，改めて両者の関係性に着目し，プロジェクトの成り立ちについて考察する．

3.2 考察：表現活動はどのように自立するのか？

(1) プロジェクトを牽引するリーダーの役割交代

一般的に，共同デザインをどのように進めるのかは，デザインすることの専門家である作り手が担うものとされる．しかし，実際には，そうとは限らないことが明らかになった．次に示すのは，共同デザインのプロジェクトを牽引する役割が，作り手から使い手に交代したとみなした場面である．これは，新聞ワークショップをつくり始める 2 年前に起きたことである．

＜エピソード B. 2014 年 1 月～5 月頃の出来事の記録＞

2014 年，プロジェクトでは表現の道具「Zuzie Poetory」を用いて，看護の心を表出・共有する表現活動プログラムのデザインに取り組んでいた．その表現活動をやってみることは楽しく，看護師 H は「新人看護師が，患者さんから『ありがとう』と感謝の気持ちを伝えられ，それに涙したストーリーを Zuzie Poetory で作ってみたい」と意欲的であった．しかし，その表現活動をどのように院内で実施していくかを話し合っていた時，看護師の一人が「いいけど，これを使う時間はない」と発言した．この後ろ向きの発言に愕然とした私（著者）は「このプロジェクトは，そうした時間を必要とするような新たな文化を院内につくろうとしているのではないですか？」という主旨の返事をした．その時間を看護業務の一部にしていくことは，私たち作り手にはできないことだとわかっていたからだ．その後，彼女は看護の仕事のあり方を変えるというこのプロジェクトを「引っ張っていくのは私たちなのですね」と発言．数ヶ月後，彼女らは自分たちで「ケアリングの継承」を目的とする表現活動の実施を計画し始めた．（2017 年 7 月 著者）

著者は，看護師の発言をきっかけに，改めて自分の限界を看護師に表明している．そして，その表明を受け，看護師が行動し始めている．このように，作り手が何らかの事情でプロジェクトを牽引しないことを表明せざるを得ない時，使い手はそのプロセスを破綻させないために，自分たちの意思で判断し，行動を起こすことを迫られる．このような事態は，プロジェクトの中で数回起きている．本研究では，このような事態を「リーダーの役割交代」とし，表現活動の自立を促進する方法と捉えた．

(2) 使い手にしかできないことから

リーダーの役割交代によって展開する共同デザインの過程では，使い手である看護師たちが，作り手が行なっていたことを徐々に習得しているとみなすこともできるだろう．例えば，机の並べ方や，参加者への声がけの仕方である．看護師らは，こうしたことを契機に，「（作り手でなくても）自分たちにもできる」といった効力感を手に入れ，徐々に自分たちにできることを増やしていったと考えられる．しかし，改めてここで着目したいのは，使い手は，作り手がやっていることと全く同じことをコピーしているわけではないということである．看護師らが行っていることは，例えば，参加する部署を選ぶことや新聞ワークショップを実施するタイミングを決めることである．さらに，表現活動を組織の中にいかにして広げ，定着させ，持続性を保つのかの検討である．その内容は，作り手だけでは決められない．

この段階に入ると，役割交代は劇的に起きるのではなく，自然体を帯びくる．そこには，お互いにやれることをやるというという共通理解が構築され，現場で生じていることに従って，両者が反応することが起きるからである．つまり，使い手は，明示的な役割交代がなくとも，作り手との相対的な関係の中で，自分たちにしかできないことを見出すようになるのである．

次に示すのは，表現活動が病院で自立したと著者がみなした時の記録である．

＜エピソード E. 共同デザインにおける気づきの記録＞

初めて，看護師たちが主導する新聞ワークショップに参加させてもらった際，私は自分が行っていたことを看護師たちが行っているのを見て，喜びとともに彼女たちに「巻き込まれた」という感覚を得た．そこには，私たちデザイナーがいつもやってきたように振舞っている看護師たちがいたからだ．しかも，私がやるよりもはるかによく．そこにいた自分は，もう巻き込む側の人ではない．私が，彼女たちがやりたいことの一部になったのだと理解した．（2017 年 3 月 著者）

ここに記された「巻き込まれ」という言葉から，著者が表現の場をつくる知恵や技術を使い手が手に入れていることを認め，そこに自分の存在価値の危うさを感じていることが伺える．しかし，看

護師がデザイナーにはなることはない．その逆も同様であり，デザイナーは看護師にはなり得ない．そのように考えれば，作り手が，病院の中に表現の場をつくるという実践に参加しているのは，何か知識や技術を獲得するためでもなく，何かを看護師に教え，指導するためでもないのである．改めて，両者は関わり合うが，その関係は対称的ではないことがわかる．

(3) 新たに構築された関係性：デザイン共同体

このように，表現活動が自立に向かう共同デザインのプロジェクト過程には，2つのコミュニティの異なる思考と行為がある．それは両者の相対的な関係性の中で発揮される力である．だからこそ，両者はそれぞれの力を認め合い，互いの立場を超えて，新たな自分を発見する．このようなプロセスを相互に構築できる状態を「デザイン共同体」と呼んでみたい．

作り手と使い手が「デザイン共同体」になるとは，何を意味しているのだろうか？作り手にプロジェクトの進行を任せきりだった使い手が，リーダーの役割を担い始めたことを思えば，そこには，使い手が自ら変容していく動的なプロセスがあったとみなせる．これは，使い手が「私はこうである」という場所に留まらなかったことによるものである．そして，それは，自分が自分でないもの，つまり現状の自分からズレることを迫られたことによって生じたものだ．ここでいうズレとは，強い意志をもった大きな転換ではなく，他者の影響を受けて，現場から少し移動してしまう状態を表している．このことから，「デザイン共同体」とは，他者の存在が，それまでの自分の役割を超える契機となるような，相互構成的な関係が構築されている状態を意味しているのだと考えられる．この議論から，表現活動が自立するためには，作り手と使い手の両者がプロジェクトを展開する主体となれるように，プロジェクトそれ自体をもデザインすることが重要であると考えられる．

3.3 デザイン共同体に「なること」

本事例研究では，人が表現者になること，使い手が作り手になること，デザイン共同体になることに着目してきた．そもそも，人はどのようにして自分以外の何者かになるのだろうか？

例えば，ある人から「あなたはお母さんになったんだね？」と問われたとする．この問いに対して，著者は「子どもが私をお母さんにしてくれた」のだと答えるだろう．もちろん子どもを産んだことで社会的には母親になったわけだが，正確には，子どもという他者の存在との関係によって徐々に母親になっていったように思われる．そしてそのようにしか，人は本当の意味での母親にはなれないのではないだろうか．著者は子どもの存在を受け入れることによって受動的かつ能動的に共同で母親になったのである．

同様に，作り手と使い手がデザイン共同体に「なる」ためには，個人の意思や働きかけのみを頼ることは難しい．そこには，他者の振る舞いを包括的に受け入れてみることが必要である．個人による働きかけを能動性，他者の受け入れを受動性と捉えれば，そこには，これら2つの一見矛盾するようなベクトルが同時に作用するメカニズムが働いていると考えられる．

そのように考えれば，事例において看護師にズレを迫ったのは，できないことをできないと正直に表明するような作り手の態度にあると考えられる．それは，作り手がプロジェクトの流れを全て決めてしまうのではなく，使い手が表現することやリーダーになることを受け入れるという態度である．作り手の態度を使い手が受け入れると，リーダーの役割は交代せざるを得なくなる．つまり，受動性と能動性が同居したプロセスを経て，使い手は，それまでとは別の自分を演じることで，表現の場をつくる側の人間へと飛躍する契機を得る．心理学の研究者であり実践者でもあるホルツマンは，人間の発達とは，今の自分で「在ること（being）」よりも「成ること（becoming）」を基礎概念とするものであり，「自分でない人物をパフォーマンスすることで，自分が何者かを創造する活動となる．」と説いている[ホルツマン 2014]．

このように，「デザイン共同体」においては，受動性と能動性が同時に作用するようなメカニズムを内在させることによって，人が自（みずから）ら自（おの）ずから，表現の場の「作り手になる」ための力が働いているのだと考えられる．

4. 結論と展望：デザインの専門を問い直す

表現活動が自立するためには，その表現活動が組織や社会の営みの中で，ある目的を達成するための手段となることが必要である．そのために，作り手は自分たちの社会の営みをより良くしようとする使い手と協力し合える関係を結ばなければならない．事例研究から導出したのは，それは「デザイン共同体」と呼べるような相互構成的な関係性であるという知見である．そのためには，作り手と使い手の両者がプロジェクトを展開する主体となれるように，プロジェクトそれ自体をもデザインする方法が重要であることが明らかになった．

一般的に，専門家とは高度な知識と技術を身につけ，その後にそれらを適用し行為するものと考えられる．著者らが実践してきた道具をデザインは，そのような技術的合理性を基本原理とする専門性で成り立っている．しかし，活動をデザインの問題にしようとする過程で，著者はそのような専門性の限界に少し気づくことになった．活動には人が内包されているために，作り手が全てを決められないからだ．さらに，共同デザインに関わったことで，その気づきは確信へと変わった．

事例研究で作り手としてのデザイナーは，これまでに自身が身につけた知識と技術が適用できないような，不安定で不確実な状況の中にあえて入り込み，その都度，その場の状況に応じて自分のできることを見出してきた．そのことが，プロジェクトのリーダーの役割交代を起こし，共同デザインの過程を駆動してきた．そして，表現の場が自立的につくられるようになった今，作り手として行うべきことは何もないのかもしれないと思いつつ，それでも自らの役割を発見しながら，看護の仕事を変革しようとする実践に参加し続けている．表現活動プログラムのデザインとは，受動性と能動性を統合的に扱うことで，人の「なること」を可能にするメカニズムの設計であり，両者の統合こそ与えられた固定的枠組みから越え出るための開かれた実践のあり方なのである．私たちデザイナーもまた，デザイン共同体のメンバーになることを契機として，内側から自分の専門性を問いなおし，そこに新たな可能性を拓くことのできる専門家になることが求められるだろう．

＜謝辞＞
本プロジェクトに参加する佐賀大学医学部附属病院看護部の皆様に感謝申し上げます．なお，本研究の一部は JSPS 科研費 15K16174 の助成を受けたものです．

参考文献

[藤満 2013] 藤満幸子，山口真由美，原田由美子，椛島久美子，宮之下さとみ，南里美貴，百武朋美，山田クリス孝介，須永剛司，小早川真衣子，新野佑樹，渡辺健太郎，西村拓一：医美工連携による看護情報システムの開発を目指したデザイン・プロジェクト」，第 33 回医療情報学連合会（第 14 回日本医療情報学会学術大会，2013

[須永 2013] 須永 剛司，小早川 真衣子，山田 クリス孝介，渡辺 健太郎，新野 佑樹，西村 拓一:Co-design プロジェクトが

自発的に回ること：社会を形づくるデザインに向けて, 人工知能学会誌, 28(6) 886-892 , 2013

[渡辺 2015] 渡辺 健太郎, 藤満 幸子, 原田 由美子, 山田 クリス孝介, 須永 剛司, 小早川 真衣子, 新野 佑樹, 阪本 雄一郎, 西村 拓一, 本村 陽一:看護現場における業務経験の表現・共有支援システムの開発, 情報処理学会論文誌, 56(1), 137-147, 2015

[小早川 2018]小早川 真衣子,須永 剛司,丸山 素直,平野 友規,山田 孝介,西村 拓一,渡辺 健太郎,藤満 幸子 表現のための活動プログラムのデザイン:看護の心を表現すること 人工知能学会全国大会論文集, 2018

[ホルツマン 2014] ロイス・ホルツマン (著), Lois Holzman (原著), 茂呂 雄二 (翻訳), 遊ぶヴィゴツキー: 生成の心理学へ,新曜社, 2014

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

「的な」をとおして学ぶ
On learning through metaphorical imaginations

加藤文俊[*1]
Fumitoshi Kato

[*1] 慶應義塾大学
Keio University

The present project explores the ways in which various "good places" are shaped and reshaped through our communication processes. Acknowledging that we are always "in-communication," our research begins with a series of critical observations of our day-to-day activities. This year, we are conducting a project named "*tekina* (its like…)." We rely on metaphors when we encounter "something" that is difficult to put into proper words. Or, occasionally, we need to name the "something" to begin with. "*Tekina*" is the term to represent metaphorical expressions through which we can begin to speculate upon in qualitative research.

1. はじめに

筆者は，人びとの「集まり」や「居心地のいい場所（グッド・プレイス）」，コミュニケーション行動の理解に関心をいだき，おもに社会学（都市社会学）の観点からフィールド調査を行っている．自らの調査研究を進めるとともに，教員という立場上，観察・記述を中心に据えた質的調査の方法や，調査者として現場に向き合う態度を（大学生に）教えるという課題にも向き合っている．

フィールドワークは，私たちの五感を動員しながら，身体的に学んでゆく性質のもので，いわゆる「座学」を補完するべく，さまざまな実習が考案されている．フィールドワークのための課題は，どのようにデザインすればよいのか．本稿では，筆者が 2018 年度に考案した実習課題を紹介しながら，「気づく力」を育むためのきっかけづくりについて論じたい [加藤 2014].

2. 課題のデザイン

フィールドワークで発見した〈何か〉をことばにするのが難しいとき，あるいはまだことばさえあたえられていない〈何か〉に向き合うとき，私たちは別の〈何か〉にたとえたり，見立てたりする．曖昧さを認めながらも，「～的な」「～みたいな」「～らしい」「～っぽい」といった表現が，現場での直接体験を伝えようとする際に用いられる．こうした日常生活に密着した体験に関わるコミュニケーションに求められる隠喩的・比喩的な表現を，作業用の概念として「的な」と呼んで，質的な調査研究のありようについて考えてみることにした．

「的な」という表現によってコミュニケーションが促されるとき，じつは「受け手」の読解レベルが試されるということになる．そもそも，語りたい〈何か〉が，何にたとえられているのか，何に見立てられているのかを知らなければ，「的な」という表現は役に立たない．〈何か〉を伝えるための工夫や方法でありながら，前提となる知識を持ち合わせていないと，理解することができない．

「パロディ」や「あるある」と呼ばれる表現も，同じような状況をもたらす．上質なものは，何が「パロディ」の対象になっているかがわかりやすく，かつその対象からのギャップや誇張があればこそ，その表現を味わうことができる．同様に，「あるある」を目にしたとき，自分自身の経験とすぐさま結びつけることができれば，頷くことができる．「パロディ」や「あるある」は，「受け手」の能力を執拗に要求する，難しい表現だと言える．

こうした隠喩的・比喩的な表現の理解を創造するために，大きく二つのステップで課題をデザインし，いずれも学部 2～3年生（ゼミ生）を対象とするグループワークとして実施した．

3. コンセプトワーク：先人たちに学ぶ

3.1 「背中を見て」学ぶ

まず，「的な」というキーワードをとおして，メディアと表現の多様性について考えるにあたって，先人たちの表現活動に着目することにした．あらかじめ選ばれた4名について，個性的な表現のあり方，さらには，それを成り立たせていたと考えられる文化的・社会的な文脈を調べることを目指した [加藤文俊研究室 2018].

今回は，吉田初三郎，成瀬巳喜男，タイガー立石，宮武外骨の 4 人を選んだ．それぞれの人生は，時代的には重複する部分もあるものの，直接のつながりはないはずである．マップ（鳥瞰図），映像（映画），グラフィック，テキストと，一人ひとりの選んだ媒体も表現方法もことなるが，いずれも，既知の世界から，想像力と技法を駆使して，自らの世界を拡張してゆくための方法や態度を考えるきっかけになった．

3.2 経過と考察

コンセプトワークは，いわゆる「調べ物」だと理解されがちだが，たんに「理屈」や知識，情報として人物像を受容するのではなく，一人ひとりが，どのように創作やコミュニケーションに向き合ったか，その「生き様」を理解することを強調した．つまり，先人たちの個別具体的な「もがき」を知ることが重要であった．このフェーズでは，以下のような知見をえることができた．

(1) 媒体（メディウム）の選択

先人たちは，いずれも自らのものの見方・感じ方に応じて，表現方法や素材，媒体を探索していた．現場との不断の対話をとおして一人ひとりの感性が開拓され，自らのスタイル，つまり「的な」と呼ぶべきユニークな方法や態度がかたどられていった．

(2) 組織づくり

それぞれの表現は，個人的な着想や問題意識に基づいているが，他者との関わりなしに活動は成立しえなかった．それは，

連絡先：加藤文俊，慶應義塾大学環境情報学部，〒252-0882
神奈川県藤沢市遠藤 5322，fk@sfc.keio.ac.jp

「受け手」を意識することだけではなく，何らかの共同作業を伴うものであり，組織や制度づくりも重要な課題として位置づけられていた．

(3) 社会的・文化的背景の理解

上記の点と関連して，それぞれの表現者たちの活動は，その時代的な文脈と不可分であった．時には不可避的に制作プロセスや表現に制約が加えられ，紆余曲折，試行錯誤を経ながら進行する，相互構成的な側面が際立った．

4. フィールドワーク：まちを歩く

次のフェーズでは，先人たちの流儀を学んだ上で，実際にまちを歩きながら，あらためて「的な」について考えた．大規模な再開発に向けて，いま大きく変りつつある桜丘町（東京都渋谷区）を，フィールドワークの対象地として選定した [加藤文俊研究室 2019].

4.1 「桜丘的」から「的な桜丘」へ

本プロジェクトにおいては，フィールドワーカーとしての世界のとらえかた，表現のしかたに着目している．重要なのは，私たちが向き合うのは「的な桜丘」であって，「桜丘的な」ではないという点だ．つまり，本プロジェクトは，桜丘のまちを歩いて「桜丘的」と感じた〈モノ・コト〉を採集，集約し，桜丘を性格づけようとしているわけではない．むしろ，桜丘をとおして，自分（自分たち）の感性や表現方法を発見，再発見することを目指すものだ．したがって，極端に言えば，桜丘町である必要はない．

誰にでも，まち（あるいは，もう少し広い意味での「環境」）に反応する「きっかけ」のようなものがあるはずだ．まずは，その「きっかけ」に出会い，その上で，自らの方法や表現のスタイルとして語れるように円環的な学びへの導入を促す．

4.2 経過と考察

学生たちは5つのグループに分かれて，およそ5か月にわたって，桜丘町に足をはこんだ．それぞれが，現場から何を感じるのか，そしてその個別具体的な体験をどのように表現し，ことば（あるいは表現物としての形）をあたえていくかについて思索と試行を重ねた．

10 月末は，一部エリアの再開発計画の一環でいくつかの店舗が閉店したが，あるグループは，閉店の場面に幾度も居合わせながら，「閉じること」「別れる」ことに焦点を当てて，その現実とイメージ（現実に先行して私たちが想い描く「閉店」の様子）とのギャップをとらえる方法を提案した．あるいは，他のグループは桜丘界隈に漂うにおいを手がかりに，嗅覚をとおしたまちの理解をどのように記述し表現するかについて試行錯誤した．また，まちで感じたことを，即時即興的にスケッチやイラストで表現し，Tシャツにプリントして身に纏うことができるような方法を考案したグループもあった．足しげく店に通うことで，まちに暮らす/まちで店を営む人びとに接近し，その過程で何らかの形で「まちに還す」ことが，まちの理解に役立つことを提案したグループ，そして，文字通り身体的にまちと接する（接地する）方法を採用したグループもあった（くわしくは，プロジェクト報告のサイトを参照）．

5. 考察

それぞれのグループの個性を活かした成果物は，桜丘町を表現したものであるが，より抽象度の高いことばで語ることによって，他のまちでも適用可能なものとなる．

5.1 「ラボラトリー」としての課題環境

本稿では「的な」「的な桜丘」という標題で実施した課題について，コンセプトワークおよびフィールドワークという二つのフェーズを紹介した．課題のデザインは，思索・試行を実現するための「ラボラトリー（実験室）」のデザインとして理解することができるだろう．「ラボラトリー」は，フィールド（個別具体的な体験）とコンセプト（普遍抽象的な言明）を橋渡しし，行動と思考との円環的な往復を可能にする仕組みである [加藤 2017].

課題や実習プログラム，ワークショップ等を「ラボラトリー」として位置づけることで，複雑なフィールドを単純化し，操作可能な領域として扱うことができる．

5.2 視点を呈示する

「ラボラトリー（実験室）」という言葉から，私たちは仮説検証型のアプローチを想起しがちであるが，ここでの課題は「視点呈示型」ともいうべきものである [西條 2007]．この場合，「ラボラトリー」における活動は，人びととの関係を変革し，新しい仮説や概念を生成する仕組みとしての側面が際立つ．「視点呈示」を念頭にデザインされた課題は，コミュニケーションを誘発し，グループのメンバーどうしの知的な交流・交歓を促す．

5.3 転用可能性

「的な」は，フィールドとコンセプトを接続し，絶え間ない往復を促すためのヒントを示唆するために設定したテーマであった．私たちは「臨床」ということばから，医療・福祉等の文脈を想い浮かべることが多いが，「臨床現場」は広義にとらえると，私たちが表現者として暮らす場所であり [中村 1992]，すべて「秩序（オーダー）を作り出す」ための一連の努力によって成り立っていると言えるだろう [パパネック 1974]．一連の実践をとおして得られた知見の「転用可能性」を見極めることによって，私たちは，まちを理解する想像力をさらに高めることができるはずだ．

参考文献

[加藤文俊研究室 2019] 加藤文俊研究室：的な桜丘，プロジェクト報告，https://vanotica.net/tknskrgk/, 2019.

[加藤文俊研究室 2018] 加藤文俊研究室：的な（It's like…），プロジェクト報告，https://vanotica.net/tekina/, 2018.

[加藤文俊 2017] 加藤文俊：「ラボラトリー」とデザイン：問題解決から仮説生成へ『SFC Journal』第 17 巻第 1 号　特集：Design X*X Design: 未知の分野における新たなデザインの理論・方法の提案とその実践（pp. 110-130），2017.

[加藤文俊 2014] 加藤文俊：まちの変化に「気づく力」を育むきっかけづくり（特集・フィールドワーカーになる）『東京人』5月号（no. 339, pp. 58-63）都市出版，2014.

[西條 2007] 西條剛央：質的研究とは何か（ライブ講義），新曜社，2007.

[中村 1992] 中村雄二郎：臨床の知とは何か，岩波新書，1992.

[パパネック 1974] ヴィクター・パパネック：生きのびるためのデザイン，晶文社，1974.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

環境との相互関係の変化に対する身体の分節的応答

Segmental response of the human body to changes in the interrelationship with the environment

田中　和哉[*1]
Kazuya Tanaka

[*1] 帝京科学大学　医療科学部　理学療法学科
[*1] Teikyo University of Science, Faculty of Medical Science, Department of Physical Therapy

跡見　友章[*2]　高野　渉[*3]　藤木　聡一朗[*4]
Tomoaki Atomi　Wataru Takano　Soichiro Fujiki

[*2] 杏林大学　[*3] 大阪大学　[*4] 獨協医科大学
[*2] Kyorin University　[*3] Osaka University　[*4] Dokkyo Medical University

長谷川　克也[*5]　清水　美穂[*6]　跡見　順子[*6]
Katsuya Hasegawa　Miho Shimizu　Yoriko Atomi

[*5] JAXA　宇宙研究開発機構
[*5] Japan Aerospace Exploration Agency

[*6] 東京農工大学
[*6] Tokyo University of Agriculture and Technology

In general, it may be difficult to secure the amount of activity due to pain. Therefore, there is a need to pay attention to qualitative aspects of exercise, such as body balance control. In human body, the mass ratio of the trunk including the head is said to occupy 60% of the whole body. Therefore, evaluation of the segmentality of the head and trunk in posture / motion control strategy is important as an indicator of stability and efficiency of motion. In this research, we investigated human's segmental attitude control using an acceleration sensor based on a cantilever structure control model. As a result, a segmental response of the body occurred in the change of the relationship between the environment and the body. Regarding human body balance control, the possibility that each segment of the body responds to each frequency band was considered.

1.　意識下・無意識下における姿勢制御

ヒトは日常生活において，重力に代表される様々な物理的刺激の入力に応答し，意識下・無意識下で姿勢制御を行っている．意識下での制御においては，重力に抗する複数の筋を協調的に制御することが必要となる．重力環境下における適切な姿勢・動作制御は，抗重力筋の萎縮を抑制するなど，身体機能の維持にとっても重要である．

一般に，運動器的疾患による疼痛のため、運動量の確保が困難となる場合があることから，身体バランスや歩き方など，運動の質的側面に着目する必要性がある．人体においては、頭部を含めた体幹部の質量比は，身体全体における 60%を占めるとされる。従って，姿勢・動作制御戦略における頭部・体幹部の分節性に関する評価が、動作の安定性や効率性の指標として重要である．

一方で，無意識下での制御や応答には，前庭迷路系の姿勢反射の影響や関節周囲筋や靭帯の物性が反映され，下位セグメントの動揺が上位セグメントに伝播する．これらのなかでも物性による応答は，必ずしも動揺を低減する役割のみではなく，剛体に近い状態であれば直達的に動揺を上位セグメントに伝播し，過度な柔軟性を持つ場合は，周波数帯域や特性によってはより強い振幅をもった揺れとして伝播することがある．

これらの意識下・無意識下での姿勢制御は，動揺の周波数によってその影響が反映される帯域が異なると考えられる．

連絡先：田中和哉，帝京科学大学医療科学部理学療法学科
〒409-0193 山梨県上野原市八ツ沢 2525，Tel: 0554-63-4411，e-mail: k-tanaka@ntu.ac.jp

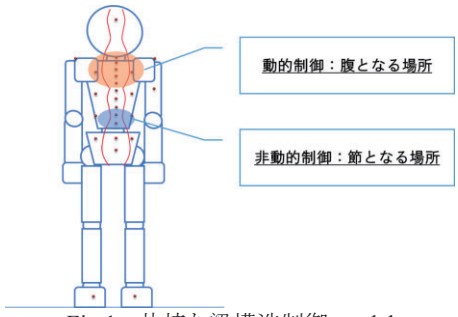


Fig.1　片持ち梁構造制御 model

また，ヒトの立位姿勢制御を片持ち梁と同様の構造として考えると，その物性や解剖学的構造によって足部などの下位セグメントからの動揺が上位セグメントに伝播される際に，節なる非動的制御部分と腹となる動的制御部分に分かれるのではないかと考える(Fig.1).

そこで本研究では，ヒトの環境との相互関係の変化に対する身体の分節的応答を，片持ち梁構造制御モデルに基づいて検討した．

2.　加速度計を用いた姿勢・動作解析

現在，ヒトの姿勢・動作解析について，加速度センサを用いた分析が広く用いられるようになっている．単一の加速度センサを用いた活動量分析などが解析方法として用いられているが，姿勢制御など質的な解析を行う上では，多分節モデルであるヒ

トの解析には複数のセンサを用いた手法の確立が必要であると考えられる.

3. 対象と方法

対象者は, 既往に整形外科疾患等を有しない健常男性 1 名(年齢 21 歳, 身長 172cm, 体重 78kg)とした. 対象者においては, 本研究の目的を説明し, 同意を得た.

実験方法は, 開閉眼の 2 条件及び開閉脚の 2 条件を組み合わせた計 4 条件で立位姿勢の体幹動揺を 20 秒間計測した. 体幹動揺の計測においては, 3 軸加速度センサモジュール MMA7361LC(Strawberry Linux 社製)を頭部・体幹部・骨盤部に貼付し 1000Hz で計測を行った.

出力されたアナログ信号を A/D 変換ボード(NI USB-6225, National Instruments 社製)に入力し, LabVIEW2012(National Instruments 社製)を用い, ノート PC にデータの取り込みを行った. センサから得られたデータ特性を線形モデルとして仮定し, 重力加速度にて補正を行った. また, 身体各部位への貼付時の傾きについても, 前後左右方向への加速度入力を行ったキャリブレーションデータよりベクトル変換を行い補正した.

計測データは, 条件ごとに前後方向及び左右方向に分けて, 1/1000 秒ごとの変位の絶対値を求め 20 秒間の積分値を算出した. また, 20 秒間の前後方向及び左右方向の加速度データについて離散フーリエ解析を行い, 各身体セグメントのパワースペクトルを解析・検討した.

4. 結果

条件毎の前後左右加速度積分データについては, 頭部・骨盤部に比して, 胸腰椎移行部での加速度積分値が少ない傾向が見られた(Fig.2).

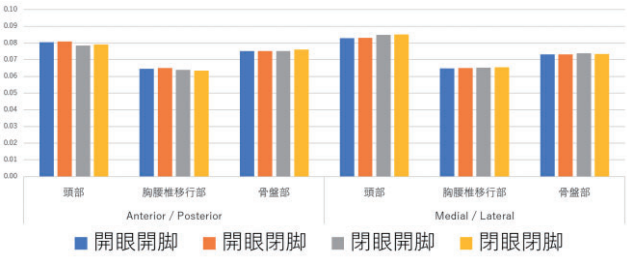


Fig.2 条件毎の前後左右加速度積分データ

開閉眼及び開閉脚の異なる条件での加速度情報について, 周波数解析を行ったところ, 周波数帯域が異なる振る舞いが見られた. 閉眼閉脚条件では, 高周波成分の割合及びパワースペクトル共に高い傾向が示された(Fig.3A , B).

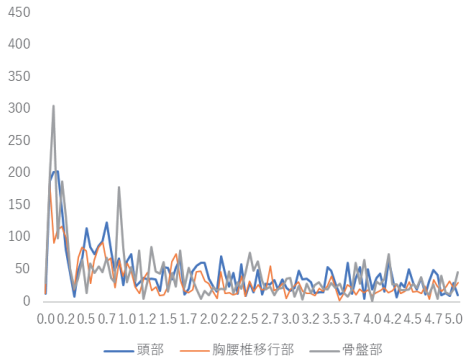


Fig. 3-A 開眼開脚の前後成分パワースペクトル

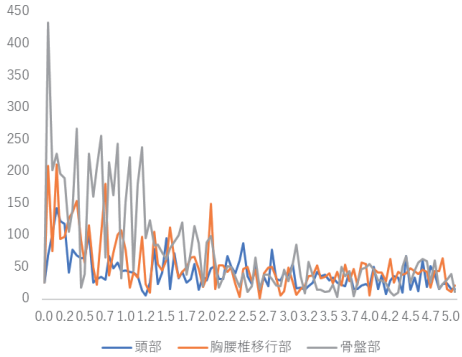


Fig. 3-B 閉眼閉脚の前後成分パワースペクトル

5. 考察

加速度センサを用いた解析において, 量的変化としては胸腰椎移行部の動揺が少なく, パワースペクトルにおいては骨盤部で高周波成分が多く、頭部で低周波成分が多くなる傾向が見られた(Fig.4).

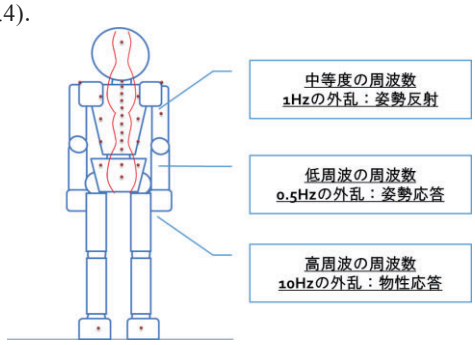


Fig.4 周波数帯域別姿勢応答仮説

このことは, 単一の量的解析では, 差が見られにくいものであっても, 2 つ以上のセンサを用いセグメント間の比較を行うことで, 異なる傾向を示していることが分かった.

高周波成分においては, 特に閉眼時の下肢筋固有感覚系の機能が表れやすいとの報告もある. 今回の解析方法を用い, 各セグメント間での周波数成分の減衰傾向や加速度データの振幅を分析することにより, 部分的な機能を分析することが可能になることが示唆された.

参考文献

1) MOE-NILSSEN, Rolf; HELBOSTAD, Jorunn L. Trunk accelerometry as a measure of balance control during quiet standing. Gait & posture, 2002, 16.1: 60-68.

2) ATOMI Yoriko：Design and production of practical underwear for sensing body knowledge-Posture control during walking induced by tactile information. 人工知能学会全国大会論文集, 2011, 25th , 3D2-OS8-11

3) 斉藤徹郎, 松本泰尚, 吉見雅行：前後振動が立位人体の姿勢保持へ及ぼす影響に関する研究. 土木学会第 59 回年次学術講演会

4) 藤原勝夫; 池上晴夫. 身体各部位の動揺の周波数分析による立位姿勢調節の検討. 体育学研究, 1985, 30.3: 241-248.

2K5-OS-16b-04

野生の実験：「実践目的が浮動する構え」を獲得する

Savage Experiment: Acquiring Perspective for World practice aiming floating

堀内隆仁[*1]　諏訪正樹[*2]
Takahito Horiuchi　Masaki Suwa
[*1] 慶應義塾大学大学院政策・メディア研究科　[*2] 慶應義塾大学環境情報学部
Graduate School of Media and Governance, Keio Univ.　Faculty of Environment and Information Studies, Keio Univ.

The savage mind, Levi-Strauss advocated, enables people to combine things which don't relate to each other in order to generate new one. In this paper, we present first author's special learning practice in athletics, and argue about this as a 'savage experiment'. The speciality in his learning practice is below: first, the practice and his basic embodied skills in daily living have been crossed each other. Second, he made some embodied media for augmenting embodied skills in daily living, on his own. By this experimental actions, prototyping the media has transformed into a self-organizing system from mere a step. Thus, his practice has aimed at the floating point between athletics and prototyping, where he is able to find new variables by embodiment.

1. はじめに

1.1 野生とは

本稿題目にある野生とは，文化人類学者レヴィ=ストロースが著した「野生の思考」[レヴィ=ストロース 76]を受けている．近代化に伴う欧米中心主義が流行る中，彼は世界各地の未開社会でフィールドワークを実施した．詳らかにされたのは，近代科学以上に綿密な自然種の弁別や，それをオペレータとして構築した社会規範（トーテミズム）・神話体系など，驚くべき文化体系の数々であった[*1]．それを可能とするのは，一見関係ない事物同士を組み合わせて新たな意味を構成する，「野生の思考」だとした．

レヴィ=ストロースは，野生的な行為にブリコラージュを挙げる．ブリコラージュは，目下の目的に対して，あり合わせの資材を組み合わせて「なんとかやりくりする」ことを言う[*2]．その資材集合は，ポリネシアのマナ型の観念[ibid.]や，代数学の記号Xのように，余白をもつシニフィアンになるということだ．「出来上がりは，手段の集合の構造と計画の構造との妥協」だとした上で．彼は以下のように述べる（ブリコルールとは，ブリコラージュする者を指す）．

> （ブリコルールは）ものと「語る」だけでなく，ものを使って「語る」．限られた可能性の中で選択を行うことによって，作者の性格と人生を語るのである．計画をそのまま達成することはけっしてないが，ブリコルールはつねに自分自身のなにがしかを作品の中にのこすのである．([ibid.], p27)

ここから，ブリコラージュとは，現実解を，自らの「物語」として紡ぐことだと考えられる．

1.2 第一著者は，研究者であり実践者

第一著者・堀内は，陸上競技（大学入学時から十種競技[*3]専門：PB6283点，十種競技歴7年，陸上競技歴10年）に熱中し，これまで実践者として取り組むと同時に，研究者としてその過程を探究してきた（e.g.[堀内，諏訪 16,18a,18b]）．本稿では，第一著者自らの実践を物語った上で，その認知的な特性を野生の観点から，内省的に議論する．

2. 第一著者の実践

本章では，堀内の実践を具体的に物語る．節は時系列順である．

2.1 野生化以前

競技と長らく向き合っていると，日々の泥臭く細かな試行錯誤はもとより，競技に対する「構え」という大きなレベルにおいても，意識変革がなされるものである．しばしば「怪我や挫折」が契機となる．堀内においてもその例に漏れず，巧者のパフォーマンスの「見た目」をなぞるように練習する最初段階から，フィジカル強化を目論んだ練習量の積み上げにアクセントをおく段階，苦手な「走り」の

連絡先：堀内隆仁，慶應義塾大学大学院政策・メディア研究科，メールアドレス：discus4372@keio.jp

[*1]彼らは「文明社会から取り残された野蛮な人たち」などではなく，欧米とは異なる仕方で発展を遂げていたのだった．この考え方が，構造主義の走りとなる．
[*2]日本語訳に，「日曜大工」が当てられることがある．
[*3]走・跳・投の計10種目の総合力を競う．キング・オブ・スポーツとも称される．

根本的改革によって総合力を上げようとする段階（cf.[堀内，諏訪 16]）・・・etc．決して順風満帆ではないながらも，漸進的にパフォーマンスアップを果たしながら，生活の喜びを感じていた．

2.2 生活動作のスキル性を究める

M1の秋，大きな転換点が訪れる．シーズンを終え，長らく抱えていた右膝の痛みの原因が，身体の使い方の癖にあることに気づいた．ごく簡単に述べれば，走る際に，身体重心鉛直真下より，過度な前方位置に接地していることで，接地中に接地脚の膝に高負荷がかかっていた．さらにそれが「普段の歩き」でも無自覚に生じていることを悟ったのだ．「自らの競技者としての実践」が，陸上競技場を飛び出，日常生活と交錯しはじめた瞬間だった．冬季は，立ち方や歩き方から徹底的に見直しを図ることで，競技動作改善にも成功した．長きにわたる実践が，立ち方や歩き方の模索に至った例は，他のスキル学習例題においても報告されている（e.g.剣道[諏訪，赤石 10]，野球の打撃[諏訪 16]）．

春のシーズン開幕を待ち望んでいた矢先，疲労が蓄積し，右足舟状骨を疲労骨折する．結果的にM2シーズンは走ることすらままならず，受傷から走れるようになるまで一年半程度を要した．

2.3 メディアのプロトタイピング経験を介して，生活環境に手を加える

⑴ 身体を計測するツールから，刺激するメディアへ

M2シーズンは，何もせずに終わったわけではない．上記の通り，生活の一部が練習になっているからだ．また，修士研究副査がメディアアーティストでもあったことが関係し，自らの身体を測る様々な「ツール」のプロトタイピング（＝つくりながら使う）を行なった（cf.堀内，諏訪 18a]）．当初は「ツールで身体を客観的計測する」というスポーツ科学的姿勢で挑んだが，すぐに厳密計測の難しさやそれに伴うコストへ直面した．興味深く感じられたのは，インタラクティヴ・ツールの振る舞いにより，思いがけず動きや身体感覚が喚起される事態である．

例えば，図１は堀内がプロトタイピングした「加速度可聴化ベルト」である．本ツールは，リアルタイムに，センサーで検知した加速度を音に変換し，ブザーから音出力する（加速度と音周波数がマッピング）．当初は，運動における「無駄な上下運動」を削減するために，フィードバックを得ることを企図した．しかし，実際に本ツールを腰部に装着して歩いてみると，「耳の周囲をハエが飛び回るような間抜けな音変化」に，むしろ踊るような激しい上下運動を促されてしまったのだ．

こうした経験を積むうちに，堀内のツールに対する姿勢は，「つくり・いじる中で自分の身体を『刺激』するメディアであれば良い」と割り切るように変化した．これらのメディアは，百均グッズやDIY用資材，Processing（アートプログラミング言語）やArduino（電子工作キット＆IDE）といった素材として普及したもので自作した，チープなものある．

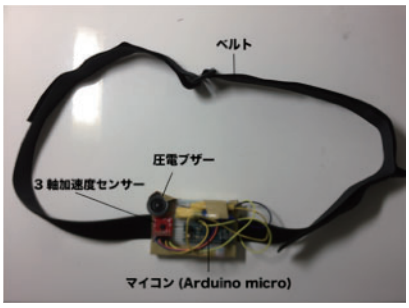


図1：加速度可聴化ベルト

⑵ 日常生活の環境に手を加える

プロトタイピング経験を通して得た「割り切り姿勢」は，前節同様に日常生活にまで及んでゆく．引越し準備において，洗濯機を1人で運ぶ際の持ち方・体幹部の使い方を悟ったり，河原の石を絶妙なバランスで積み上げる[*4]ことで，外力だけ（筋力なし）で成立するポジショニングを探ったり（図2），リュックを「腹負う」（腹側にリュックを背負う行為を堀内が命名）ことでスムーズな重心移動に試みる，等である．日常生活において，歩く等の「身体動作」のみならず，自らを取り巻く環境に対しても，競技者としての身体を刺激する道具へ仕立てるように変容したのだ．これもブリコラージュである．なお2.1節～本節で述べた過程は[堀内，諏訪 18a]に詳しい．

図2：堀内の石花作品

2.4 メディアプロトタイピングを楽しむ

修士研究を終え，まだ怪我の治癒途中でもあった堀内は，コンピュテーショナルなメディアの制作に様々取り組む．いきなり「使えるメディア」を作ろうとする姿勢ではなく，メディアづくり自体を楽しみ，表現・制作スキルの向上を目論む向きが強い．

怪我治癒後からは，本格復帰を目指し，通常の「トレーニング」も徐々に開始している．トレーニング・メディアづくり・日常生活，これらを行ったり来たりしながら，実践を実り豊かにせんと努めている．

*4 ロックバランシングまたは石花と呼ばれる遊びである．

⑴ Ultra Re-generation of MRI

メディアの一例を紹介する（本例は，[堀内，諏訪 18b]に詳しい）．「Ultra Re-generation of MRI」と題するメディアを制作した．本メディアは，インターネット上に漂流する「MRIによる全身横断面*5撮像gifファイル」をデータとし，多様な動きや変換エフェクトをかけた3Dの人型像を，画面上に再生する*6．「死体解剖学的なデータに命や意志を吹き込む」ことが初期コンセプトであった．「身体に訴えかける」表現を探りながら，バージョン変更を繰り返した．

様々な効果を試作する内に，制作時に意図しなかった身体感覚を再生像から得た．それがもととなり，「体幹部の冠状面*7での分解」という変数[J.J.Gibson & E.J.Gibson 55]を獲得した．図3は，冠状面の分解感覚が感じられる再生像である．陸上競技の多様な運動において，身体重心と接地位置との前後位置関係は重要である．冠状面（身体の前後方向）での分解は，この位置関係の身体感覚に関わるという点で，競技者として価値ある変数であった．一見して遠回りなプロセスで，変数を獲得したのだ．

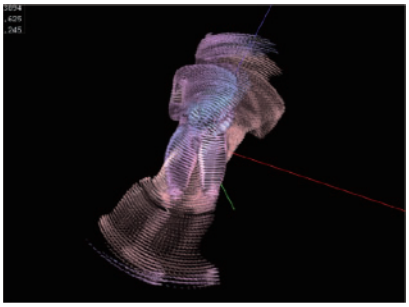


図3：Ultra Re-generation of MRI の再生像一例

本メディアは前節の加速度可聴化ベルトと異なり，運動を直接検知するインタラクティヴ・ツールではない．MRI撮像は本来，怪我や病気の発見に資する医療用データである．メディア・アートの文脈において，基本的にコンピュテーションは，何らかのモノを表現媒体化（メディア）することに使われる．前節までの経験を経た上でメディア制作スキルを育む堀内にとって，MRI撮像すら，「河原の石」や「洗濯機」同様，記号Xになったのだ．

⑵ 特異な環境下での体験に垣間見る，新たな構え

> 吊り輪を足場として移動するには，必然的に股関節の筋群を優位に使わなくてはならなかった．考えてみれば当たり前で，普段の地面は，足と地面のインタフェースが，ある意味固定化されている．しかし，吊り輪に足をかけて運動する場合，砂場の比ではないくらいダイナミックなインタフェースとなるから，プルプル感がやばかったのだ！だからこそ，安定を作り上げるためには，足ではなく，膝や股関節ということになる．なんかおもしろ練習器具がDIYできる気はするのだが．
>
> （2018年10月28日の記述より一部抜粋）

図4：競技者を逸脱した堀内の「構え」が表れる記述例

図4に示した記述は，堀内が本格的なフィールドアスレチックを体験した日のものである（堀内の一人称研究[諏訪，堀ら 15]では，日々の実践における問いや変数発見の過程を，自身でメタ認知している）．記述から，単なる競技者を逸脱した堀内の構えを垣間見ることができる．特異な環境下での運動体験で，競技者としての身体の使い方・感覚に加え，環境の特性をも探りつつ，それに応じて生成される身体側の特性を，考察している．

3. 「第一著者の野生」を内省する

実践はデザイン行為[Simon 01]である．自らにとってより好ましい状況を指向する一連のアクションを指す．個々のアクションも何らかの目的をもつが，デザイン行為が大局的に指向するものが，「実践目的」だと言える．野生（ブリコラージュ）の概念をもとに，堀内の実践は，以下のようなデザイン行為として表せる．

> あり合わせの身体（動作・感覚）を資材とし，あり合わせの環境から工具として選び，これらをもちいて現実解を構成する．

堀内の実践目的が「競技スキルを構成する」だったことを踏まえると，通常のブリコラージュがあり合わせの資材とするものは，「資材」と「工具」に分けられる．資材は，自分の身体の「身体動作や感覚」であり，工具は，資材を形づくるためにあり合わせの環境から選んだモノである．

3.1 各実践はどう野生か

堀内が果たした「野生化」とは，実践目的に対して，資材または工具が異なるドメインへ飛び出ることだろう．長らくの実践では，あり合わせをフル活用する必要性に至り，結果的に実践目的と一見異なるドメインに属するモノに着手することになる．

競技スキル構成という実践目的に対して，競技的なトレーニングをしたり，その身体動作を動画撮影・解析するだけでは，資材・工具が「競技ドメイン」に属する．単線的に問題解決を目論む，非野生的な構えである（ブリコラージュではなく，エンジニアリング的）．2.1節に概説した実践の段階では，野生化していなかった．

「立つ・歩く」等を問い始めたこと（cf.2.2節）は，「競技ドメイン」の実践目的と異なる「日常生活ドメイン」に属する身体動作を資材として選ぶことを意味し，野生である．

2.4節では，メディアを作り始めた．「加速度可聴化ツール」は，はじめから素材（記号X）の意味が与えられた材料で制作しているため，実践目的の「競技ドメイン」と明確に異ドメインとは断言できない．一方，その際の資材は競技スキルではなく，身体構造の感覚（cf.2.4節⑴）や，生活動作（cf.2.4節⑵）をいじっているから「日常生活ド

*5身体において，体軸に垂直な面を指す解剖学用語である．
*6本メディアは，openFrameworks・GLSLを用いて制作した．
*7身体において，背と腹に平行な面を指す解剖学用語である．

メイン」と捉え得る．だから，野生なのである．自前のチープな工具だからこそ，選ばれた資材の方が「日常生活ドメイン」になったと解釈できる．「競技ドメイン」の資材（動作）を仕立てる工具は，自作が難解だと当時判断したからである（cf.2.3節⑴）．

生活環境にまで手を加え始めたのは，実践目的に対して異なるドメインの工具を選んだことになり，野生である．環境から選び出す工具が何であるかによって，似た資材でも出来上がりは異なり，工具によって選ばれる資材もある程度決まる．すなわち発見される変数が異なる．メディアプロトタイピングや，2.3，2.4節で紹介した実践では，異なる工具だからこそ，獲得している変数が異なった．

3.2 ボトムアップな変数発見

我々が身体を具備する認知主体であるがゆえに（cf.身体性[Pfeifer & Bongard 06]），変数発見は生じる．正解的変数がトップダウンに降ってくることなどないならば，「手応え[佐伯 95]」を感じながら，ボトムアップに変数発見する他ない．ボトムアップに発見した変数は，身体化され，柔軟に発揮できる実践知となろう．そうして身体を拡張し，得た変数は物語の隙間を糊付けする．

また，工具が自前であることの意義は，自分でつくるからこそ，記号Xとしてのポテンシャルや，それに伴う資材選択を模索できるのであり，作り手が素人かプロかは第一義ではない．特にコンピュテーショナルなメディアは，その性質が顕著と言える．工具を外注すると，身体性がスキップされかねない．

3.3 実践目的すら浮動する構え

2.4節の堀内はもはや，単なる競技者を逸脱している．端的に言えば，メディアをプロトタイピングすることが自己組織化・自己目的化し，実践目的が拡張している．特に，コンピュテーショナル・メディアは，モノを表現媒体へ仕立てる手段として優れている．Ultra Re-generation of MRI（cf.2.4節⑴）のような，表現可能性の試行は，あり合わせの環境と身体双方の集合において，記号Xとして選べる要素を増やした．この手段の豊かさが，競技スキル構成という実践目的以外の道を切り拓いた．メディアづくりは目的に対するブリコラージュの工具づくりではなくなった．

ブリコラージュは，現実解を構成している．つまり，「設計図に基づいた唯一の正解」に合致するはずはないし，そもそもそれは存在しない．このように妥協するならば，実践目的という正解も更新されて然るべきである．目的更新こそがデザインの本質[中島ら 08]でもだ．2.3節⑴で獲得した「割り切り姿勢」は，妥協を実感し始めたのでもあった．

一見して「競技者として，あまりにコストパフォーマンスが悪い実践」であり，自己目的化は「競技者としての逃避ではないか」という反駁もあろう．しかし，未来における正解（好ましさ）は，現在における正解と異なるということを見逃してはならない．そうすれば，メディアづくり自体を楽しむことは，目先の競技的利益に飛びつかずボトムアップな発見を待つ我慢の工夫でもある，と積極的に解釈できる．こうして堀内は，単一目的を指向しない，悠然たる構えを獲得したのだった．

参考文献

[J.J.Gibson & E.J.Gibson 55] James J.Gibson，Eleanor J.Gibson: Perceptual Learning: Differentiation or Enrichment?, Psychology Review, Vol.62, No.1, pp. 32-41, 1955.

[Pfeifer & Bongard 06] Rolf Pfeifer, Josh Bongard: How the Body Shapes the Way We Think: A New View of Intelligence，A Bradford Book, 2006

[Simon 01] Herbert. A. Simon: The Science of the Artificial. Third Edition, The MIT Press, 2001.

[佐伯 95] 佐伯胖：「わかる」ということの意味，岩波書店，1995.

[諏訪 16] 諏訪正樹：「こつ」と「スランプ」の研究－身体知の認知科学－，講談社メチエ，2016.

[諏訪，赤石 10] 諏訪正樹，赤石智哉：身体スキル探究というデザインの術，認知科学, Vol.17，No.3, pp.417-429, 2010.

[諏訪，堀ら 15] 諏訪正樹，堀浩一（編著），伊藤毅志，松原仁，阿部明典，大武美保子，松尾豊，藤井晴行，中島秀之（著），一人称研究のすすめ-知能研究の新しい潮流-，近代科学社，2015.

[中島ら 08] 中島秀之，諏訪正樹，藤井晴行: 構成的情報学の方法論からみたイノベーション, 情報処理学会論文誌 Vol.49, No.4, pp1508-1514, 2008.

[堀内，諏訪 16] 堀内隆仁，諏訪正樹：「走り」を追究するアスリートの物語–身体で実践し，気づき，考え，解り，実践する–，第30回人工知能学会全国大会論文集，1M4-OS-14a-5，2016.

[堀内，諏訪 18a] 堀内隆仁，諏訪正樹：陸上競技スキル学習の仮説生成型研究–身体・生活意識・ツールが共創する「野生の実践」–, 2017年度慶應義塾大学大学院政策・メディア研究科修士論文，2018.

[堀内，諏訪 18b] 堀内隆仁，諏訪正樹：身体性を孕む共創現象の成立–自分とモノとの『距離』–，共創学会第2回年次大会予稿集1-6，2018.

[ポランニー 03] マイケル・ポランニー（著），高橋英夫（訳）：暗黙知の次元，ちくま学芸文庫，2003.

[レヴィ=ストロース 76] クロード・レヴィ=ストロース（著），大橋保夫（訳）：野生の思考，みすず書房，1976.

Organized Session | Organized Session | [OS] OS-4

[3D3-OS-4a] 自律・創発・汎用 AIアーキテクチャ(1)

栗原 聡（慶應義塾大学）、川村 秀憲（北海道大学）、津田 一郎（中部大学）、大倉 和博（広島大学）

Thu. Jun 6, 2019 1:50 PM - 3:10 PM Room D (301B Medium meeting room)

[3D3-OS-4a-01] Story Association Mediated by Individual and General Concepts

〇Taisuke Akimoto[1] （1. Kyushu Institute of Technology）

1:50 PM - 2:10 PM

[3D3-OS-4a-02] How "intelligence" is called as "intelligence"?

〇Masayuki Yoshinobu[1] （1. Freelance）

2:10 PM - 2:30 PM

[3D3-OS-4a-03] Double Articulation Analyzer with Prosody for Unsupervised Word Discovery

〇Yasuaki Okuda[1], Ryo Ozaki[1], Tadahiro Taniguchi[1] （1. Ritsumeikan University）

2:30 PM - 2:50 PM

[3D3-OS-4a-04] Generating Collective Behavior of a Robotic Swarm in a Two-landmark Navigation Task with Deep Neuroevolution

〇Daichi Morimoto[1], Motoaki Hiraga[1], Kazuhiro Ohkura[1], Yoshiyuki Matsumura[2] （1. Hiroshima University, 2. Shinsyu University）

2:50 PM - 3:10 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

個別的概念と一般的概念を介したストーリーの関連付け

Story Association Mediated by Individual and General Concepts

秋元 泰介
AKIMOTO Taisuke

九州工業大学大学院情報工学研究院
Graduate School of Computer Science and Systems Engineering, Kyushu Institute of Technology

Abstract: From a cognitive system perspective, *story association* (association among stories) can be positioned as a common process underlying various cognitive activities including remembering and reusing of stories, construction of a subjective world, etc. In this paper, the author presents a basic computational theory of story association especially focusing on the role of entities, times, and places. In particular, an identical entity, time, or place appearing in multiple stories is positioned as an *individual concept*, a general level representational unit in a cognitive system. A simple implementation of story association model based on the proposed theory is also presented.

1. はじめに

筆者は，ストーリーを生成しながら外界と関わり合うことが人間の知能の本質であると考え，物語的な情報とその動きを中心に据えた認知システムの研究を行っている．これまでに，世界の表象としてのストーリーの構造や生成に関する概念レベルの理論を提案してきた[Akimoto 18a, 18b]．現在の主な課題は，それらをもとに，計算レベルの理論・方法を作ることである．

人工知能・認知科学において，物語に関する研究には，問題解決，状況説明，対話，コンテンツ生成，物語理解等，様々な目的やタスクがあり得る．それに対して，認知システムの観点から物語の認知を研究するということは，その根底をなす基本原理を探求することを意味する．そのような立場から，本稿では，ストーリーの「関連付け」に関する基本的な考え方とその部分的な実装を示す．特に，個別的な存在物に対応する「個別的概念」という要素の位置付けと働きに焦点を合わせる．

ここでストーリーの関連付けというのは，認知システムの内部において，あるストーリーないし事物と，別のストーリーないし事物とが結び付くことを意味する．また，それが様々な種類の情報に基づく類似・関連性の複合的な働きによって生じることを，特に，ストーリーの多元的関連付けと呼ぶ．認知システムにおいて，ストーリーの関連付けは，主に以下の役割を担うと考えられる．

- あるストーリーないし事物と他のストーリーを関連付けることによって，事物の主観的な意味を形作る．
- 「今・ここ」のストーリーに，過去や未来のストーリー，あるいは別の場所のストーリーを結び付けることで，時間・空間的な広がりのある世界を構築する．
- 新しいストーリーが生成する際に，それ以前に作られたストーリーを，生成のための知識・素材として引き込む．
- 類似したストーリーを関連付ける仕組みは，一般的な構造としてのスキーマを形成するための基盤になる．
- 以上を含め，様々な状況においてストーリーを想起するための基盤になる．

以下，2 節で関連研究を概観した後に，3 節でストーリーの多元的関連付けモデルの全体像を示し，4 節で個別的概念という要素の位置付けについて述べ，5 節でそれに基づく部分的な実装を示し，6 節で今後の展望を示す．なお，本稿は[秋元 18]からの発展として位置付けられる．3 節で示す全体像に大きな変更は無いが，その他の部分は新しい内容となっている．

連絡先：秋元 泰介，九州工業大学大学院情報工学研究院知能情報工学研究系，820-8502 福岡県飯塚市川津 680-4，akimoto@ai.kyutech.ac.jp

2. 関連研究

ストーリーの関連付けに深く関連する問題として，人工知能・認知科学における記憶・想起に関する研究を概観する．

[Schank 82]は，MOP（memory organization packet）というスキーマ的な知識構造に基づいて，ある状況において，過去の類似する経験（エピソード）を思い出したり，新しい経験を一般化・記憶したりする，発達的な記憶システムの枠組みを提示している．この発想は，後の事例ベース推論にもつながっている[Riesbeck 89]．また，[Kolodner 83]は，[Schank 82]に示唆を得て，E-MOP（episodic memory organization packet）という知識構造に基づく記憶システムを実装している．

Schank らの研究が問題解決のための目的指向的な想起に比重を置いているのに対して，想起のより一般的な性質として，類似に基づく想起の計算モデルが提案されている．これはアナロジーの研究とも直接的に関連している．例えば，[Thagard 90]は，意味的類似，構造的類似，目的・用途的類似の 3 種類の制約充足に基づく想起モデル ARCS を提案している．[Forbus 94]は，表面的な類似に基づく選択と，構造的な類似に基づく選択の二段階処理による想起モデル MAC/FAC を提案している．

本研究が焦点を合わせるのは，想起そのものではなく，その土台となる構造や仕組みであるが，上記のような研究も参考にしている．まず，ストーリーの関連付けが，様々な類似の複合的な働きから生じると考える．この点は[Thagard 90]や[Forbus 94]に近い．一方，[Thagard 90; Forbus 94]は記憶間の比較あるいは制約充足的な計算を通して記憶を選択するような仕組みになっているが，本研究では，システム内にストーリーを組織化する働きをする心的要素が形成され，それを通してストーリーどうしが自ずと関連付けられるという考え方をする．この点は[Schank 82]に近い．加えて，ネットワーク状の記憶組織における意味や想起等の現象を扱う活性伝播（spreading activation）[Collins 75; Anderson 83]の考え方（ネットワーク上のある要素の活性度が近隣の要素に減衰しながら伝播・拡散していくような仕組み）も取り入れる．

それから，認知システムへのアプローチとして見ると，本研究は[Kokinov 94]の DUAL に比較的近いだろう．DUAL は，[Minsky 86]の『心の社会』に示唆を得て構築された，複数の小さな表象的かつ処理的なエージェントからなる，マルチエージェ

ントシステム型の認知アーキテクチャである．それに対して本研究の特徴は，ストーリーを中心に据えてシステム全体の構造や動きをモデル化しようとする点にある．

3. ストーリーの多元的関連付けモデルの全体像

ストーリーの多元的関連付けの仕組みの基本的な考えは次の通りである．個々のストーリーは，ある時，ある場所における具体的な事象ないし事物に関する情報であるが，一般的な心的要素を共有する形で構成される．そしてこの一般的な心的要素を介して，自ずとストーリーどうしが関連付けられる．

3.1 一般的な心的要素に基づくストーリーの構成

ストーリーを構成する一般的な心的要素の分類を表 1 に示す（この分類は今後も継続的に拡張・修正していく必要がある）．これらの要素及びストーリーを，情報の複合性・具体性という観点から，図 1 に示すような形で，(1)ストーリー，(2)個別的概念・スキーマ，(3)一般的概念・感覚情報パターンという 3 つの層に分ける．外側ほど複合性・具体性が高くなるようになっている．なお，これらを全体として「心的世界」と呼ぶ．

3.2 活性伝播による関連付け

ストーリーの関連付けを，あるストーリーまたはその部分が活性化した際に，一般的な心的要素を介した活性伝播によって，他のストーリーが活性化することとする．ここでは，心的要素の活性度という概念を，主に以下の二つの意味で用いる．

- ストーリーが活性化するということは，それが参照・想起・生成・変化等しやすい状態になることを意味する．
- 心的世界の活性状態は，全体として，その時に経験・注目しているストーリーまたは事物の意味を表す．

この活性伝播の局所的な模式図を図 2 に示す．ある心的要素（ノード）の活性度が一つ以上の入力活性度に基づいて決まり，さらにそれが他の要素に伝播していくというものである．

ある心的要素 y の活性度a_{y}の計算式の基本形を，

$$a_{\mathrm{y}} = \sum_{i}^{n} O_{\mathrm{x}_i} w_i \tag{1}$$

とする．ここでnは入力側の要素数，O_{x_i}は入力側のi番目の要素x_iの出力活性度，w_iは y とx_iの間の結合強度である．x_iの出力活性度O_{x_i}は，当該要素の活性度a_{x_i}が閾値θ_{out}未満の場合に他要素への伝播を遮断する以下の式により計算する．

$$O_{\mathrm{x}_i} = \begin{cases} 0, & a_{\mathrm{x}_i} < \theta_{out} \\ a_{\mathrm{x}_i}, & a_{\mathrm{x}_i} \geq \theta_{out} \end{cases} \tag{2}$$

表 1 一般的な心的要素の種類．

一般的概念		単語の意味に相当する要素（例えば「犬」「社長」）．
個別的概念		特定の存在物に対応する，固有名詞的な要素（例えば「タマ」「太郎」）．「私」（自分自身）も，個別的概念の一種とする．また，個別的概念の特別な下位区分として，「時間概念」と「場所概念」を設ける．
	時間概念	ストーリー（事象）が起きる時間に対応する，言語的に分節された時間範囲（例えば「平成」「明日」）．
	場所概念	ストーリー（事象）が起きる場所に対応する，言語的に分節された空間範囲（例えば「新宿」「自宅」）．
感覚運動的パターン		事物の具体的な心像（イメージ）を形作り，非言語的な感覚運動情報（視覚，聴覚，味覚，嗅覚，触覚，運動）の類似の認知のもとになる要素．例えば，画像認識で用いられる特徴表現を成す要素や，イメージスキーマの原始的要素に相当するもの．
スキーマ		以上が複合的に構造化された，複数のストーリー（の部分・要素）に通じる一般的構造．

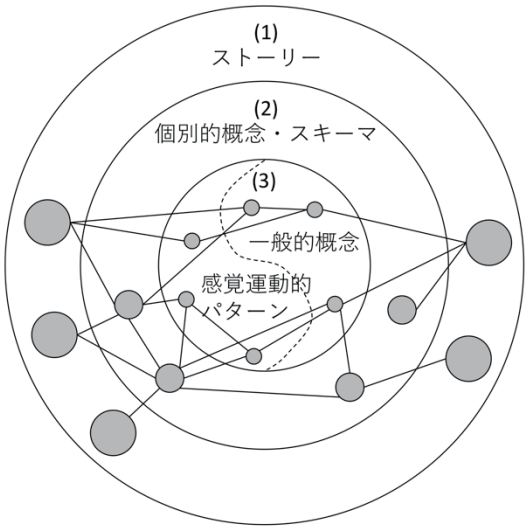


図 1 心的要素の組織構造（心的世界の構造）．

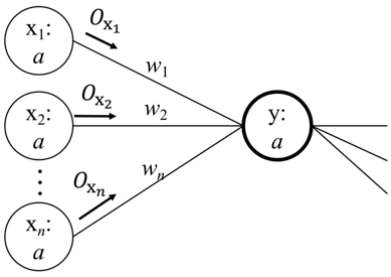


図 2 活性伝播の局所的な模式図．

3.3 ストーリーの関連付けの全体的な流れ

3.1 節で示した構造（図 1）と 3.2 節で述べた活性伝播に基づいて，次のような流れでストーリーの関連付けが生じるものと考える—i) あるストーリーが活性化する．ii) ストーリーから個別的概念・スキーマへ，さらに一般的概念・感覚運動的パターンへと活性度が伝播する．iii) 一般的概念・感覚運動的パターンの層や，個別的概念・スキーマの層の内部で活性度が伝播する．iv) 一般的概念・感覚運動的パターンから個別的概念・スキーマへ，さらにストーリー層へと活性度が伝播する．要するに，図 1 の構造において，外側→内側(ii)，層内(iii)，内側→外側(iv)という流れでストーリーどうしが関連付けられるということである．

4. 個別的概念の位置付け

図 1 に示した全体像における個別的概念に焦点を合わせて，その位置付けや働きについて述べる．

まず，個別的概念が特定の存在物に対応する要素であるという考えを，例を交えて説明する．ストーリーが世界の表象であるとすると，例えば「私」や友人の「三郎」は，ストーリーの中の登場人物として位置付けられる．仮に，「二年前の三郎は，坊主頭で毎日野球の練習に打ち込んでいた」と，「最近の三郎は，金髪で毎日学校をサボって私と一緒に街を徘徊している」という二つのストーリー（記憶）があるとする．このように人柄が変わっても，これらのストーリーに現れる「三郎」を同一の人物と見なされるべきであろう．このような同一性の認識を支えるのが，個別的概念である．

ある対象を以前見たものと同一であると見なすためには，類似の認知も必要であろう．しかし，ここで述べているのは，「同一である」という認識そのものに対応する表象が必要であり，それ

は，対象と個別的概念とを結び付けることであるということである．なお，存在物を「人」や「猪」のようにカテゴリ化することは，それを一般的概念と結び付けることとして，個別的概念の働きとは区別することができる．

「名前」は，個別的概念の基本要素となるだろう．しかし，個別的概念の形成は，基本的には，「一般化」に相当するものであると考える．すなわち，ある対象（事・物・者・場所・時間）が一つ以上のストーリーの中で現れた際に，それに伴う様々な情報（心的要素）がその個別的概念に結び付いたり離れたりして，その基本的な性質が形作られていくということである．このような見方をすると，個別的概念とスキーマは似ている．（なお，一般化は，帰納的ないしボトムアップ的にだけではなく，既存の構造の再構成等によっても生じるものと考える．一般化については検討すべき問題が多数あるが，本稿では立ち入らないことにする．）

以上の考えに基づいて，ある個別的概念と他の心的要素との結び付きを以下のように整理する．

- 個別的概念は，それが直接的に関与するストーリーと結び付く．
- 個別的概念には，それが持つ基本的な属性として，一般的概念及び感覚運動的パターンが結び付く．
- 個別的概念どうしは，広い意味での関連性（例えば人間関係や時間・空間的な包含・近接）によって結び付く．

これらがストーリーの関連付けを媒介することになるが，最初のもの（ストーリーと個別的概念の結び付き）については，特に以下のような機能を持つと考えられる．

- 世界を構成する存在物・時間・場所等をキーとしてストーリーを参照・想起するための索引となる．
- 諸々の存在物（自己・他者・物・時間・場所）に対して，個体の記憶（ストーリー）に基づく主観的な意味を形作る．

5. 簡易的な実装

以上の考えに基づいて，個別的概念と一般的概念を介してストーリーの関連付けを行うプログラムを簡易的に実装した．実装言語は Scala である．なお，一般的概念の方は，単語間の意味的な類似による関連付けに関与する．スキーマ，感覚運動的パターン，それから生成や一般化の仕組みは実装に含まれていない．また，ストーリーの内部構造は扱わずに，単なる一ノードとして表現している．

5.1 入力データ

入力として，複数のストーリーと個別的概念のデータが与えられる．一つのストーリーは，固有番号，自然言語表現，任意の数の一般的概念名及び各々の重み，任意の数の個別的概念名及び各々の重みからなる．重みの値は$0 < w \leq 1$とする．ここに記述する一般的概念と個別的概念は，原則としてストーリーに現れるものとする．なお，自然言語表現はプログラムの動作には関与しない．図 3 に一つのストーリーの記述例を示す．一方，一つの個別的概念のデータは，名前（固有番号付き），任意の数の個別的概念名及び各々の重み，任意の数の一般的概念名及び各々の重み（$0 < w \leq 1$）からなる．こちらの記述例は割愛するが，記述形式はストーリーと概ね同じである．

以上の入力データから，図 4 に示すような形のネットワーク構造が作られる．ストーリー（S），個別的概念（D），一般的概念（G）の 3 層からなり，要素間の結合構造は$W^{(\mathrm{SD})}$，$W^{(\mathrm{SG})}$，$W^{(\mathrm{DG})}$，$W^{(\mathrm{DD})}$の 4 つの行列により表現される（$0 \leq w \leq 1$．結合の無い部分は 0）．例えば$w_{13}^{(\mathrm{SD})}$=0.3 は，ストーリーs_1と個別的概念 d_3 の間の結合強度が 0.3 であることを意味する．

```
#1
私は物心がついた時からずっと東京の下町にある五郎の家で暮らしている。
下町 暮らす
0.3 0.3
$1:私 $2:五郎 $p1:東京 $p5:五郎の家
0.5 0.4 0.4 0.5
```

図 3 入力のストーリーデータの記述例．

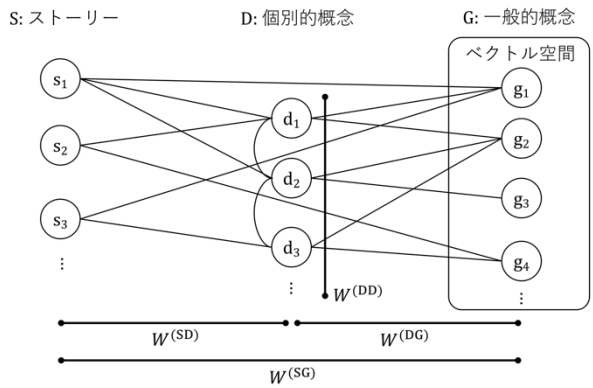


図 4 心的要素のネットワーク構造．

加えて，一般的概念間の類似度・関連性の表現として，単語の意味を多次元ベクトルにより表現する word2vec [Mikolov 13] の Skip-gram モデル（300 次元，ウィンドウサイズ 10, negative sampling）を用いる．学習には『現代日本語書き言葉均衡コーパス』[国立国語学研究所 15]（出現頻度 10 以上の単語 175801 個を対象）を用いた．これによる単語ベクトル間のコサイン類似度を，一般的概念間の結合強度と見なす．

5.2 活性伝播の流れ

初期状態として，任意の一つのストーリーの活性度を 1 とする．その後，以下に示す 6 つの Step を指定回数反復する．なお，Step 6 は活性度に基づく想起の簡易的な実装に相当するものであり，本稿において重要なのは Steps 1–5 である．

Step 1 (S 層→D 層): 下式により，各ストーリーの出力活性度O_{s_i}と$W^{(\mathrm{SD})}$から，各個別的概念の活性度a_{d_j}を計算する．

$$a_{\mathrm{d}_j} = \sum_{i}^{|\mathrm{S}|} O_{\mathrm{s}_i} w_{ij}^{(\mathrm{SD})} \tag{3}$$

Step 2 (SD 層→G 層): 下式により，各ストーリーの出力活性度O_{s_i}と$W^{(\mathrm{SG})}$，及び各個別的概念の出力活性度O_{d_j}と$W^{(\mathrm{DG})}$から，各一般的概念の活性度a_{g_k}を計算する．

$$a_{\mathrm{g}_k} = \sum_{i}^{|\mathrm{S}|} O_{\mathrm{s}_i} w_{ik}^{(\mathrm{SG})} + \sum_{j}^{|\mathrm{D}|} O_{\mathrm{d}_j} w_{jk}^{(\mathrm{DG})} \tag{4}$$

Step 3 (G 層→G 層): 下式により，Step 2 終了時における各一般的概念の出力活性度O_{g_j}と一般的概念間のコサイン類似度（$\mathrm{sim}(\mathrm{g}_i, \mathrm{g}_j)$）から，各一般的概念の活性度を計算する．

$$a_{\mathrm{g}_i} = \sum_{j}^{|\mathrm{G}|} O_{\mathrm{g}_j} \mathrm{sim}(\mathrm{g}_i, \mathrm{g}_j) \tag{5}$$

ここで，コサイン類似度が閾値θ_{sim}未満の場合は，類似度を 0 と見なして活性度が伝播しないようにしている．

Step 4 (GD 層→D 層): 下式により，各個別的概念の出力活性度O_{d_i}と$W^{(\mathrm{DD})}$，及び各一般的概念の出力活性度O_{g_j}と$W^{(\mathrm{DG})}$から，各個別的概念の活性度を計算する．

$$a_{\mathrm{d}_k} = \sum_{i}^{|\mathrm{D}|} O_{\mathrm{d}_i} w_{ik}^{(\mathrm{DD})} + \sum_{j}^{|\mathrm{G}|} O_{\mathrm{g}_j} w_{kj}^{(\mathrm{DG})} \tag{6}$$

Step 5 (GD 層→S 層): 下式により，各一般的概念の出力活性度O_{g_i}と$W^{(\mathrm{SG})}$，及び各個別的概念の出力活性度O_{d_j}と$W^{(\mathrm{SD})}$から，各ストーリーの活性度を計算する．

$$a_{\mathrm{s}_k} = \sum_{i}^{|\mathrm{G}|} O_{\mathrm{g}_i} w_{ki}^{(\mathrm{SG})} + \sum_{j}^{|\mathrm{D}|} O_{\mathrm{d}_j} w_{kj}^{(\mathrm{SD})} \tag{7}$$

Step 6: 焦点移動．ストーリーの活性度の大きさに応じた確率的選択により，別のストーリーに焦点を移す（そのストーリーの活性度を 1 にし，他のストーリーの活性度を減衰する）．その後 Step 1 に戻る．

なお，上記の各 Step の計算後には，活性度の値が 0 以上 1 以下の範囲におさまるように正規化を行う．すなわち，活性度の最大値が 1 を超える場合に，最小値を 0 とみなし，当該の層の全要素の活性度を最大値で割った値に書き換える．

5.3 動作例

このプログラムの動作例を概略的に示す．入力データとして，あるエージェントの記憶の内容を想定した 10 個のストーリーと，それらに現れる 18 個の個別的概念のデータを手作業で用意した．一般的概念は全部で 37 個である．ストーリーの内訳は，「私」の過去・未来に関するストーリー4 個，友人の「太郎」に関するストーリー2 個，社会的な出来事に関するストーリー2 個，フィクション 2 個である．

初期状態としてストーリー#1（図 3）の活性度を 1 とし，活性伝播の閾値θ_{out}とθ_{sim}は何れも 0.3 として，プログラムを実行した．実際の出力は各 Step における各要素の活性度のログデータであるが，動きを直感的に把握するための可視化プログラム（Processing で実装）も用意した．これにより表現された，1 巡目の Step 5 終了時の活性状態を図 5 に示す．図 1 と同じような形で，3 つの同心円上に，一般的概念，個別的概念，ストーリーに対応するオブジェクト（円または点）が配置されている．各要素の活性度の大きさは円・点の色の濃度・サイズにより，活性伝播は要素間を結ぶ線の色の濃度によりそれぞれ表現される．この例では，ストーリー#1 から，個別的概念「$1:私」「$2:五郎」「$p1:東京」等を介して，「私」の過去や未来に関するストーリー#2–4 や，社会的な出来事に関するストーリー#7（1964 年の東京オリンピックに関するもの）等が活性化している．一般的概念も複数活性化しているが，この例では，ストーリーの活性化にはあまり関与していないようである．

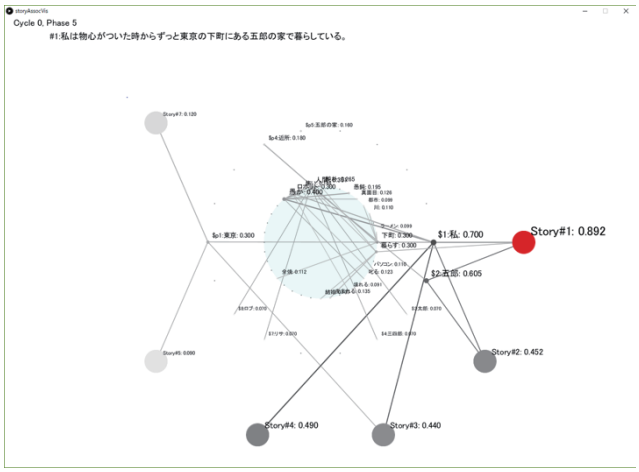


図 5 1 巡目 Step 5 終了時の活性状態．

6. おわりに

本稿では，認知システムにおけるストーリーの関連付けについて，特に個別的概念を主題として検討した．結論として，個別的概念は，ストーリーに現れる個別的な存在物の同一性を支える要素であり，その基本的な性質は，一般化を通じて，他の心的要素と結び付くような形で形成される，という考えを述べた．さらに，この考えに基づいて，個別的概念と一般的概念を介したストーリーの関連付けの簡易的な実装と動作例を示した．紙幅の都合により，十分な議論や説明ができなかった部分も多くあるため，そこは今後別稿でより詳しく扱うことにしたい．特に，時間や場所については，それぞれ特別な検討が必要であろう．また，次の課題として，視覚的情報の導入や，ストーリーの内部構造の扱いについても考えていく予定である．

謝辞

本研究は JSPS 科研費 JP18K18344 の助成を受けた．

参考文献

[Akimoto 18a] Akimoto, T.: Stories as mental representations of an agent's subjective world: A structural overview. *Biologically Inspired Cognitive Architectures*, 25, 107-112 (2018)

[Akimoto 18b] Akimoto, T.: Emergentist view on generative narrative cognition: Considering principles of the self-organization of mental stories. *Advances in Human-Computer Interaction: Special Issue on Language Sense and Communication on Computer*, 2018, Article ID 6780564 (2018)

[秋元 18] 秋元 泰介: 認知システムのためのストーリーの多元的関連付けモデルの概念と部分的実装. 人工知能学会ことば工学研究会（第 60 回）資料, pp. 15-25 (2018)

[Anderson 83] Anderson, J. R.: A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 22(3), 261-295 (1983)

[Collins 75] Collins, A. M., & Loftus, E. F.: A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407-428 (1975)

[Forbus 94] Forbus, K. D., Gentner, D., & Law, K.: MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, 19, 141-205 (1994)

[Kokinov 94] Kokinov, B. N.: The DUAL cognitive architecture: A hybrid multi-agent approach. *Proceedings of the 11th European Conference on Artificial Intelligence*, pp. 203-207 (1994)

[国立国語学研究所 15] 国立国語学研究所コーパス開発センター:『現代日本語書き言葉均衡コーパス』利用の手引き 第 1.1 版. 国立国語学研究所 (2015)

[Kolodner 83] Kolodner, J. K.: Maintaining organization in a dynamic long-term memory. *Cognitive Science*, 7, 243-280 (1983)

[Mikolov 13] Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J.: Distributed representations of words and phrases and their compositionality. *Proceedings of the 26th International Conference on Neural Information Processing Systems*, pp. 3111-3119 (2013)

[Minsky 86] Minsky, M.: *The Society of Mind*. Simon & Schuster (1986)

[Riesbeck 89] Riesbeck, C. K., & Schank, R. C.: *Inside Case-Based Reasoning*. Lawrence Erlbaum (1989)

[Schank 82] Schank, R. C.: *Dynamic Memory: A Theory of Reminding and Learning in Computers and People*. Cambridge University Press (1982)

[Thagard 90] Thagard, P., Holyoak, K. J., Nelson, G., & Gochfeld, D.: Analog retrieval by constraint satisfaction. *Artificial Intelligence*, 46, 259-310 (1990)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

「知能」を「知能」と言うのはどうしてか

How “Intelligence” is called as “Intelligence”?

吉信真之
Masayuki Yoshinobu

This paper discusses basic issues of the structure of human intelligence. As an attempt to understand it, we set up a question “How intelligence is called as intelligence?” in stead of a question “What is intelligence?” which asks its definition. To answer this, we break it down to several sub-problems as its structure, physical background and formation processes. Overviewing them, we propose a hypothesis that there is a structure in which a human as a certain area on highly complicated layers in a complex system regards the other area as intelligence.

1. はじめに

汎用人工知能の研究にあたり、「知能」とは何かという問いに出会うことは多い。そして、長い間研究と歴史が行われて来たにかかわらず、「知能」について標準的な定義は存在しないとされるが[Legg 07b]、これは知能についての解釈が多様であるということだけではなく、その原理の理解もまだ充分ではないからと言える。本稿ではその理解を困難にする要因の一つとして「知能とは何か」という定義を尋ねる問それ自体にあると仮定し、問の形式をその原理を尋ねる形式である「『知能』を『知能』というのはどうしてか」に置き換え、それをさらに副問題に分け、俯瞰することで「知能」への基礎的な理解を試みる。

2.「知能」の定義

心理学における定義やその定義を集めた論文[Legg 07b]を始め、様々な定義がなされているように、「知能」についておおよその範囲、共通性を持つ理解はあるものの、それらの定義は「知能」の原理がどのようなものかを説明したものとなってはいない。

3.「知能とは何か」とは何か

このように「知能」に対する理解を困難にさせている要因の一つに「知能とは何か」という問いそれ自体を挙げても良いだろう。「知能」について様々な定義がなされているということは、「知能」という語とそれが指し示す対象との関係も多様であること、またそれぞれの定義がなされる背景には「～とは何か」という「定義を求める問い」[Ayer 1946]が与えられていることが考えられる。

3.1.言語の恣意性

ソシュールは「言語の恣意性」という概念を提起した[Saussure 1910]。聴覚イメージ（シニフィアン）とそれに結び付けられる対象イメージおよびその範囲（シニフィエ）に自然的な必然性は無く、非自然的、歴史的・社会的・文化的なものであることを言う[丸山 1983]。したがいこれは「知能」という語においても同様のことが言える。

3.2 辞書的定義

ウェクスラーは知能を「目的的に行動し、合理的に考え、そして自分の環境に効果的に対処するための個人の総合的または全体的な能力」[Wechsler 1944]と定義している。定義は様々な視点から分類されるが[廣松 1998]が、一つの記号に対してこれと同義な記号ないし記号的表現をするときこれは辞書的な定義であり、この種の定義の問題は例えば「眼科医」に対する「目医者」のように、任意の名辞に対する同義語を見出す方法であるとエイヤーは指摘している[Ayer 1946]。ウェクスラーの定義で言えば、「知能」に対する「行動する」「考える」「効果的に対処する」などがこれにあたるだろう。

4.原理をたずねる問いの設定

知りたいことは定義では無く「知能とされる原理・構造」なのだから、人間がどのようにしてこれを分節しているかを問えばよい。ただし、例えば「知能の原理は何か」という問いでは目的に近づくことが出来ないので、ここでは「『知能』を『知能』と言うのはどうしてか」という問を設定する。

5.「『知能』を『知能』と言うのはどうしてか」

この問いに答えるため、「知能」のことを「知能」と言う構造、その構造を成立させる物理的背景、その構造を生成するに至ったプロセスの副問題に分け、それぞれについて確認し、最後に全体を俯瞰して「知能」の原理について考察する。本稿ではおおよその輪郭をつかむことが狙いなので、広範・詳細な事項には触れず概観するにとどめる。

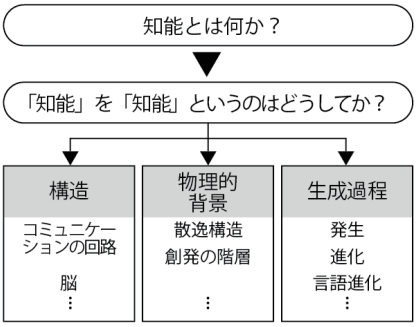


図１：副問題に分解

5.1 「知能」を「知能」と言う構造

ソシュールはコミュニケーションの回路として図 2 のように示した[Saussure1910]。図の円形部分は脳に相当する。

脳は近年でも新しい機能が見つかっているので、未解明な機能もあるだろうからあまり単純に捉えることは出来ないが、聴覚から発声への処理は、生理学的には例えば一次聴覚野、聴覚連合野、感覚性言語野、補足運動野、運動性言語野、一次運動野などの神経ネットワークを介して行われていると考えられている[藤木 2007]。

連絡先：吉信真之，m.yoshinobu89vn[at]gmail.com

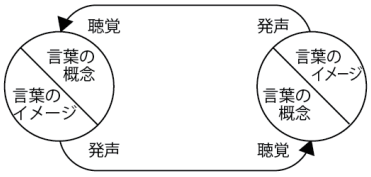


図2：コミュニケーションの回路（[Constantin 1910]を元に改変）

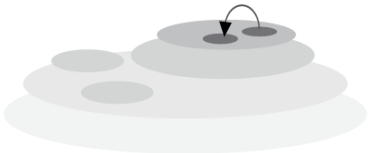


図3：創発の階層上にある「『知能』のことを『知能』という状態」（イメージ）

5.2 生成の為の物理的背景

自然の中で複雑な構造ができるしくみとして散逸構造が知られる。散逸構造は、平衡から遠く離れた条件下にある系内に現れる時間的空間的構造を言う [プリゴジーヌ 1997]。代表的な例としては、雲や太陽の表面、味噌汁などに見られるベナール対流やベロウソフ・ジャボチンスキー反応などが知られており、地球表面も散逸構造であると考えられている[松下 97]。

また、自然界には階層構造があり[鈴木]、システム論では上位の階層に下の階層には無い新たな性質が出現するとき、これを創発したと言う[伊庭 98]、

5.3 生成プロセス

（1）発生過程

個体は受精卵より発生し、分裂・分化の過程を経て最終的には身体構造や神経ネットワークなど機能する形態へと自己組織化する。細胞は細胞膜を通して細胞外から栄養となる分子等を取り込み、内部で DNA を含む各分子同士の反応を連鎖的に繰返し、不要となった成分を外へ拡散し、細胞内の成分を一定に保つことで恒常性を維持しているが、金子のモデルに拠れば、分裂により多細胞化する過程で周囲の細胞と競合して一定の状態を維持できなくなると DNA など少数の分子のコントロールの下、細胞内部の状態を変えることで新たな安定状態を獲得し分化を実現するという [金子 09]。

神経組織は神経軸索を投射するターゲットとなる細胞から分泌される誘導因子の濃度勾配によりその神経軸索を伸長、シナプス接続を行うことで神経ネットワーク構造を形成する[御子柴 02]。

（2）進化過程

各個体は、環境に適応的な個体が交配し子孫を生成する。子孫は生殖細胞を通して親から受け継いだ遺伝情報を基に幹細胞から発生し、遺伝型としては親とわずかに異なり、表現型としてはその可塑性により成長の過程で環境応じて変異した個体となる。世代交代を繰返し、遺伝型の変異を蓄積し、あるとき元の種と子孫を残せない程度に種分化した[金子 09]。このような過程を経て、人間の場合は原始には少数細胞の単純な振る舞いをする生物から、現在の特に大規模で複雑なネットワーク構成を持つ神経組織を獲得したと考えられる。

（3）言語進化

岡ノ谷に拠れば、前適応とは生物のある特性はその特性による適応を目的として進化したのではなく、他の問題のために進化したものが流用されたものいう。言語進化に当たっては、脳機能の言語への前適応があり、神経構築・生物学的準備の上に文化的過程を経て言語的構造を生成した[岡ノ谷 10]。

6. 世界観の導出

構造、物理的背景、生成過程のそれぞれを概観すると、地球表面には散逸構造が現れており、これが創発を重ねて生物やそれらの脳、言語、ふるまいなど複雑な構造を形成・自己組織化していると考えても良いだろう。

すると、「『知能』を『知能』と言う構造」とは、「創発を重ねて十分に複雑化したある領域が、また別のある領域に対して『知能』と見做している構造」と見ることができる。言い換えると、人間はベナール対流と同様に複雑系の只中にある存在であり、「知能」とされる構造はそこへ論理が介在した独特な構造として地表に奇跡的に残存しているもの、という解釈ができるのではないだろうか。

7. おわりに

本稿での検討はおおよそのものであり、構造や生成過程には他にも多くの要因が関与しているだろうが、仮にこのような構造が「知能」の実態であるとするならば、汎用的な AI を開発するにあたり、遠回りなようだが、前述の構造・生成過程のうち必要な部分を抽出・利用して「人間が知能であると見なす構造」を近道を選びつつ探索することもその手段になると考えられる。

参考文献

[Legg 07b] Legg, S. and Hutter, M.: Universal intelligence: A definition of machine intelligence, Minds and Machines, Vol.17, No. 4, pp. 391-444 (2007)
[Ayer 1946] Alfred J. Ayer, Language, Truth and Logic, 1946
[Constantin 1910] 3 ème Cours de Linguistique Générale by Ferdinand de Saussure ソシュール 一般言語学講義: コンスタンタンのノート, 影浦 峡 , 東京大学出版会, 2007
[丸山 1983] 丸山圭三郎, ソシュールを読む, 岩波書店, 1983
[Wechsler 1944] Wechsler, D, The measurement and appraisal of adult intelligence, Fourth edition, Williams & Wilkins, 1958
[廣松 98] 廣松渉, 岩波哲学思想・辞典, 岩波書店, 1998
[藤木 07] 藤木暢也, 内藤泰, 聴覚コミュニケーションと脳機能, 音声言語医学 48:277-283, 2007
[プリゴジーヌ 1997] プリゴジン, I, 我孫子誠也, 谷口佳津　共訳　確実性の終焉, みすず書房, 1997)
[御子柴 02] 御子柴克彦, 清水孝雄, 脳の発生・分化・可塑性, 共立出版, 2002
[松下 97] 松下貢, 散逸構造とはなにか, 高分子 46 巻, 1997
[鈴木 93] 鈴木賢英, 自然科学ノート, 分化書房博文社, 1993
[伊庭 98] 伊庭崇, 複雑系入門, NTT 出版, 1998
[金子 09] 金子邦彦, 生命とは何か, 東京大学出版会, 2009
[岡ノ谷 10] 岡ノ谷一夫　言語起源の生物学的シナリオ, 認知神経科学 Vol.12 No.1 2010

3D3-OS-4a-03

韻律情報を活用した二重分節解析器による自然音声からの語彙獲得

Double Articulation Analyzer with Prosody for Unsupervised Word Discovery

奥田 恭章 *1　尾崎 僚 *1　谷口 忠大 *1
Okuda Yasuaki　Ozaki Ryo　Taniguchi Tadahiro

*1立命館大学
Ritsumeikan University

Human infants discover words and phonemes using statistical information and prosody. For unsupervised word discovery, Taniguchi et al proposed the Nonparametric Bayesian Double Articulation Analyzer (NPB-DAA) which was able to segment speech data into word sequences. However, NPB-DAA uses only statistical information such as the mel-frequency cepstrum coefficients. In this paper, we extend NPB-DAA method using prosody, i.e., Prosodic DAA, for unsupervised word discovery. We use the second order differential of the fundamental frequency and the duration of silent as the prosody. We show in an experiment that Prosodic DAA outperforms NPB-DAA.

1. はじめに

人間の幼児は月齢 8ヵ月において，音声信号の分布情報や韻律情報 (Prosody，プロソディ) から，単語の単位を見いだせることが知られている．本研究では，教師なし学習によりプロソディを含む自然音声信号から，直接音響モデルと言語モデルを同時に推定することで，語彙獲得を行う手法を提案する．

音声信号から教師なし語彙獲得を行う手法にノンパラメトリックベイズ二重分節解析器 (Nonparametric Bayesian Double Articulation Analyzer: NPB-DAA) がある．しかしながら，現在の NPB-DAA ではプロソディによる音響特徴量の変化を扱うことができない．そこで本稿では NPB-DAA にプロソディを扱うような出力分布を新たに加えた Prosodic DAA を提案する．

また，プロソディを強調した日本語自然発話からなる実音声発話から教師なし語彙獲得する実験を行う．実験結果において，プロソディを用いて単語推定を行うことで NPB-DAA の音素及び単語推定性能が向上することを示す．

2. 先行研究

人間の語彙獲得過程において，月齢 8ヵ月の幼児は音声を単語ごとに分割することができる [Saffran 96]．また，幼児は語彙獲得の際に，音声学的性質であるプロソディを用いて音声を単語ごとに分割していることが知られている [Jusczyk 92]．このことから，幼児は音声信号の分布，音素と単語の共起性やプロソディを用いて語彙獲得をしていると考えられる．

谷口らは，二重分節構造を持つ時系列データの生成モデルとそのパラメータの推論手法を組み合わせて，NPB-DAA を提案した [Taniguchi 16a]．NPB-DAA は二重分節構造を持つ時系列データの解析手法である．二重分節構造は，単独では意味を持たない音素と，音素の組み合わせによって意味を持つ単語による二重の分節構造である．先行研究において，谷口らは母音のみの音声発話データからの単語分割を実現した [Taniguchi 16a]．また，音声特徴量を Deep Sparse Autoencoder (DSAE) を用いて，より高次の特徴量に変換することで単語分割精度が向上することを示した [Taniguchi 16b]．夛田らは子音を含んだ実音声発話データからの単語分割を行うために動的特徴量を用いた [Tada 17]．尾崎らは NPB-DAA にルックアップテーブルを導入することで計算量オーダを 3 次から 2 次に削減し，計算時間を 90%削減した [Ozaki 18]．

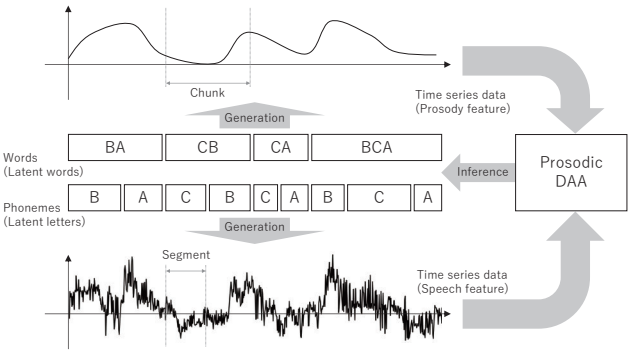


図 1: 本研究の概要図．二重分節解析器にプロソディを活用した単語分割を組み込む．ここで，Chunk は単語の継続長，Segment は音素の継続長をそれぞれ表す．

NPB-DAA では，一つの単語は同一の left-to-right の隠れセミマルコフモデル (Hidden semi-Markov Model: HSMM) [Johnson 13] に従うとしており，プロソディによる音声波形の変化は異なる単語として認識される．先行研究では，音響特徴量として，プロソディを含まない特徴量であるメル周波数ケプストラム係数 (Mel-Frequency Cepstrum Coefficients: MFCC) を用いており，プロソディによる音声波形の変化を考慮していない．そのため，幼児が語彙獲得過程で行うとされているプロソディによる単語分割を表現していない．そこで，本研究では NPB-DAA にプロソディを用いた単語分割を組み込むことを目的とする．本研究の概要図を図 1 に示す．

連絡先: 奥田 恭章，立命館大学 情報理工学研究科，滋賀県 草津市 野路東 1-1-1，okuda.yasuaki@em.ci.ristumei.ac.jp

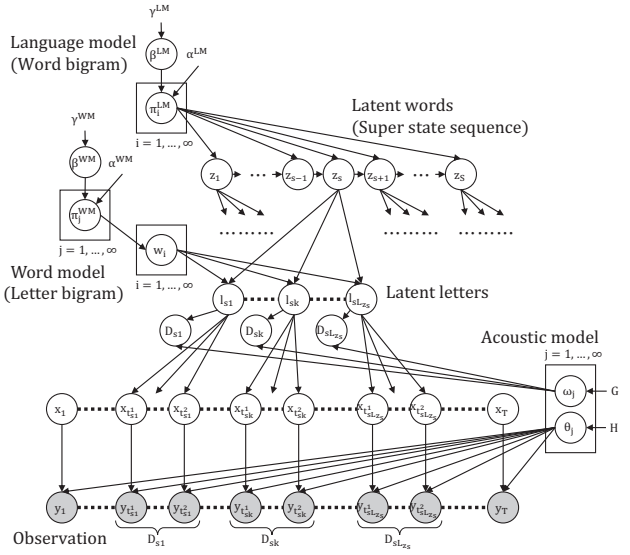


図 2: HDP-HLM のグラフィカルモデル [Taniguchi 16a]．このモデルは，二重分節構造を持つ時系列データの生成モデルである．

3. NPB-DAA

3.1 階層ディリクレ過程隠れ言語モデル

階層ディリクレ過程隠れ言語モデル (Hierarchical Dirichlet Process Hidden Language Model: HDP-HLM) は二重分節構造を持つ時系列データの生成モデルであり，これは階層ディリクレ過程隠れセミマルコフモデル (Hierarchical Dirichlet Process Hidden semi-Markov Model: HDP-HSMM) [Johnson 13] を二重分節構造の表現ができるように拡張したものである．図 2 に HDP-HLM のグラフィカルモデルを示す．その他，生成過程などについては [Taniguchi 16a] を参照されたい．

3.2 潜在単語のサンプリング

HDP-HLM の潜在単語 z_s は HDP-HSMM の推論手法で用いる backward filtering forward sampling の手続きを拡張してサンプリングされる．HDP-HLM の潜在単語 $z_s = i$ の backward message は以下の式で求めることが出来る．ここで，F_t は潜在単語の境界かどうかを表し，$F_t = 1$ であれば時刻 $t+1$ で単語が切り替わるとする．$B_t(i)$ は潜在単語 $z_{s(t)} = i$ が時刻 $t+1$ において，異なる潜在単語に遷移する尤度を表す．$B^*_t(i)$ は時刻 $t+1$ から潜在単語が $z_{s(t)} = i$ となる尤度を表す．

$$
\begin{aligned}
B_t(i) &:= p(y_{t+1:T}|z_{s(t)} = i, F_t = 1) \\
&= \sum_j B^*_t(j) p(z_{s(t+1)} = j|z_t = i) \qquad (1) \\
B^*_t(i) &:= p(y_{t+1:T}|z_{s(t+1)} = i, F_t = 1) \\
&= \sum_{d=1}^{T-t} B_{t+d}(i) p(D_{t+1} = d|z_{s(t+1)} = i) \\
&\quad \cdot p(y_{t+1:t+d}|z_{s(t+1)} = i, D_{t+1} = d) \qquad (2) \\
B_T(i) &:= 1 \qquad (3)
\end{aligned}
$$

潜在単語 $z_{s(t)} = i$ の潜在音素列 w_i を $w_i = (l_1, \cdots, l_{L_i})$ と置くと，$p(y_{t+1:t+d}|z_{s(t+1)} = i, D_{t+1} = d)$ は以下の式のように計算できる．ここで，$R^{(L_i,d)}$ は L_i 次元で要素の和が d となるような自然数ベクトル $\mathbf{r}$ の集合を表す．

$$
\begin{aligned}
p(y_{t+1:t+d}|i, d) &= \sum_{r \in R^{(L_i,d)}} \prod_{k=1}^{L_i} p(r_k|l_k) \\
&\quad \cdot \prod_{m=1}^{r_k} p(y_{t+m+\sum_{k'=1}^{k-1} r_{k'}}|l_k) \qquad (4) \\
R^{(L_i,d)} &= \left\{ r \in \{1, 2, \ldots\}^{L_i} | \sum_{k=1}^{L_i} r_k = d \right\} \qquad (5)
\end{aligned}
$$

式 4 は動的計画法を用いて効率的に計算することができる．時刻 t において潜在音素の遷移が発生し，その潜在音素が潜在単語中の k 番目の要素である確率を forward message と定義し，以下の式で再帰的に求めることができる．

$$
\alpha_t(k) = \sum_{d'=1}^{t-k+1} \alpha_{t-d'}(k-1) p(d'|l_k) \prod_{t'=1}^{d'} p(y_{t-t'+1}|l_k) \qquad (6)
$$

$$
\alpha_0(0) = 1 \qquad (7)
$$

以上により，HDP-HLM において各時刻 t における．$B_t(i)$, $B^*_t(i)$ が求まるため，以下の式に従い潜在単語 $z_{s(t+1)}$ とその継続時間 $D_{s(t+1)}$ をサンプリングできる．ここで，$D^{\mathrm{sum}}_s = \sum_{s'<s} D_{s'}$ である．

$$
p(z_s = i|y_{1:T}, z_{s-1} = j, F_{D^{\mathrm{sum}}_{1:s}} = 1) = p(i|j) \beta_{D^{\mathrm{sum}}_{1:s}}(i) p(y_{D^{\mathrm{sum}}_{1:s}}|i) \qquad (8)
$$

$$
p(D_s = d|y_{1:T}, z_s = i, F_{D^{\mathrm{sum}}_{1:s}} = 1) = p(y_{D^{\mathrm{sum}}_{1:s}+1:D^{\mathrm{sum}}_{1:s}+d}|d, i, F_{D^{\mathrm{sum}}_{1:s}} = 1) p(d) \frac{\beta_{D^{\mathrm{sum}}_{1:s}+d}(i)}{\beta^*_{D^{\mathrm{sum}}_{1:s}}(i)} \qquad (9)
$$

4. Prosodic DAA

本稿では，プロソディとして基本周波数 F_0 の二次微分及び無音区間の継続長の 2 つを用いる．無音区間の抽出では，ある閾値以下の音量が一定時間連続している区間を無音区間とし，その継続長を用いる．

次に，プロソディを HDP-HLM に組み込む方法を述べる．本稿では，プロソディはそれぞれ観測 y_t と同じフレーム数を持ち，単語の境界を表すインジケーター変数 F_t から生成されるものと仮定している．また，F_0 の二次微分を表す観測を $y^{\mathrm{F}_0}_t$ とし，無音区間長を表す観測を y^{sil}_t とする．プロソディを含む HDP-HLM のグラフィカルモデルを図 3 に示す．また，生成過程は以下のようになる．ここで GEM は Stick-Breaking Process を表し，DP は Dirichlet Process を表す．LM,WM はそれぞれ言語モデルと単語モデルを表している．また，β^{WM} は単語モデルにおける Dirichlet Process の基底測度を表し，α^{WM} と γ^{WM} はそれぞれ Dirichlet Process, Stick-Breaking Process のハイパーパラメータである．そして π^{WM}_i は潜在音素 i から次の状態への遷移確率を表している．同様に β^{LM} は言語モデルにおける Dirichlet Process の基底測度を表し，α^{LM} と γ^{LM} はそれぞれ Dirichlet Process, Stick-Breaking Process のハイパーパラメータである．そして π^{LM}_i は潜在単語 i から

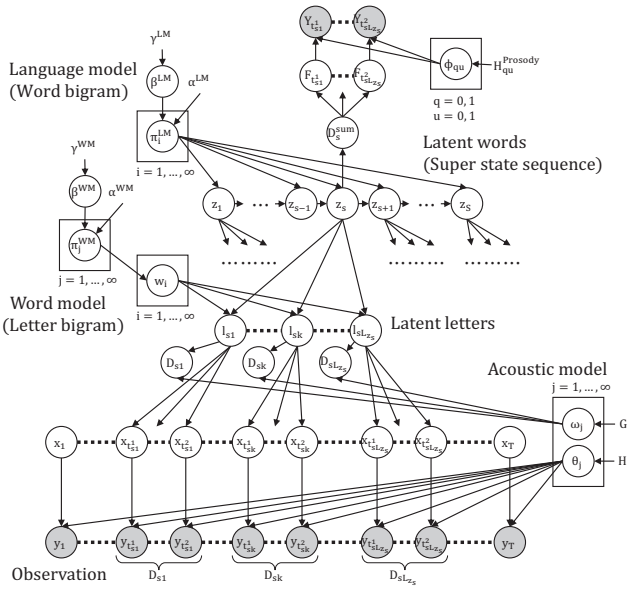


図 3: プロソディを含む HDP-HLM のグラフィカルモデル．潜在単語 z_s から z_s の継続長 D_s^{sum} が生成される．D_s^{sum} から単語境界を表す F_t が生成され，F_t から Y_t が生成される．

次の状態への遷移確率を表している．また，$Y_{1:T}$ は $y_{1:T}^{\mathrm{F}_0}$，$y_{1:T}^{sil}$ の集合を表し，$p(Y_t)\quad(t=1:T)$ は $p(y_t^{\mathrm{F}_0},\ y_t^{sil})$ である．

$$\beta^{\mathrm{LM}} \sim \mathrm{GEM}(\gamma^{\mathrm{LM}}) \tag{10}$$
$$\pi_i^{\mathrm{LM}} \sim \mathrm{DP}(\alpha^{\mathrm{LM}}, \beta^{\mathrm{LM}}) \quad (\mathrm{i}=1,\ldots) \tag{11}$$
$$\beta^{\mathrm{WM}} \sim \mathrm{GEM}(\gamma^{\mathrm{WM}}) \tag{12}$$
$$\pi_i^{\mathrm{WM}} \sim \mathrm{DP}(\alpha^{\mathrm{WM}}, \beta^{\mathrm{WM}}) \quad (\mathrm{i}=1,\ldots) \tag{13}$$
$$w_{ik} \sim \pi^{\mathrm{WM}}_{w_{ik-1}} \quad (i=1,\ldots) \quad (k=1,\ldots,L_i) \tag{14}$$
$$(\theta_j, \omega_j) \sim H \times G \quad (j=1,\ldots) \tag{15}$$
$$\phi_{qu} \sim H_{qu}^{Prosody} \quad (q=0,1) \quad (u=0,1) \tag{16}$$
$$z_s \sim \pi^{\mathrm{LM}}_{z_{s-1}} \quad (s=1,\ldots,S) \tag{17}$$
$$l_{sk} = w_{w_{z_s k}} \quad (s=1,\ldots,S) \quad (k=1,\ldots,L_{z_s}) \tag{18}$$
$$D_{sk} \sim g(\omega_{l_{sk}}) \quad (s=1,\ldots,S) \quad (k=1,\ldots,L_{z_s}) \tag{19}$$
$$x_t = l_{sk} \quad (t=t^1_{sk},\ldots,t^2_{sk}) \tag{20}$$
$$t^1_{sk} = \sum_{s'<s} D_{s'} + \sum_{k'<k} D_{sk'} + 1$$
$$t^2_{sk} = t^1_{sk} + D_{sk} - 1 \tag{21}$$
$$D_s^{sum} = \sum_{t=s1}^{sL_{z_s}} D_t \tag{22}$$
$$y_t \sim h(\theta_{x_t}) \quad (t=1,\ldots,T) \tag{23}$$
$$F_t = \begin{cases} 0 & (t=t^1_{s1}:t'-1) \\ 1 & (t=t') \end{cases} \qquad (t'=t^2_{sL_{z_s}}) \tag{24}$$
$$y_t^{\mathrm{F}_0} \sim h^{\mathrm{F}_0}(\phi_{F_t 0}) \tag{25}$$
$$y_t^{sil} \sim h^{sil}(\phi_{F_t 1}) \tag{26}$$

また，プロソディによる単語分割を含む，潜在単語 $z_s=i$ の backward message は以下の式で求められる．

$$\begin{aligned}\beta_t(i) &:= p(y_{t+1:T}, Y_{t+1:T}|z_{s(t)}=i, F_t=1) \\ &= \sum_j \beta_t^*(j) p(z_{s(t+1)}=j|z_t=i)\end{aligned} \tag{27}$$
$$\begin{aligned}\beta_t^*(i) &:= p(y_{t+1:T}, Y_{t+1:T}|z_{s(t+1)}=i, F_t=1) \\ &= \sum_{d=1}^{T-t} \beta_{t+d}(i) p(D_{t+1}=d|z_{s(t+1)}=i) \\ &\quad \cdot p(y_{t+1:t+d}, Y_{t+1:t+d}|i,d)\end{aligned} \tag{28}$$
$$\beta_T(i) := 1 \tag{29}$$

潜在単語 $z_{s(t)}=i$ の潜在音素列 w_i を $W_i=(l_1,\cdots,l_{L_i})$ と置くと，$p(y_{t+1:t+d}, Y_{t+1:t+d}|z_{s(t+1)}=i, D_{t+1}=d)$ は以下の式のように計算できる．

$$\begin{aligned}p(y_{t+1:t+d}, Y_{t+1:t+d}|i,d) &= \sum_{r\in R^{(L_i,d)}} \prod_{k=1}^{L_i} p(r_k|l_k) \\ &\cdot \prod_{m=1}^{r_k} p(y_{t+m+\sum_{k'=1}^{k-1} r_{k'}}|l_k) p(Y_{t+m+\sum_{k'=1}^{k-1} r_{k'}}|l_k)\end{aligned} \tag{30}$$

$$R^{(L_i,d)} = \{r\in\{1,2,\ldots\}^{L_i} | \sum_{k=1}^{L_i} r_k = d\} \tag{31}$$

forward message は，以下の式で再帰的に求めることができる．

$$\begin{aligned}\alpha_t(k) &= \sum_{d'=1}^{t-k+1} \alpha_{t-d'}(k-1) p(d'|l_k) \\ &\quad \cdot \prod_{t'=1}^{d'} p(y_{t-t'+1}, Y_{t-t'+1}|l_k)\end{aligned} \tag{32}$$
$$\alpha_0(0) = 1 \tag{33}$$

以上により，HDP-HLM において各時刻 t における．$\beta_t(i)$，$\beta_t^*(i)$ が求まるため，以下の式に従い潜在単語 $z_{s(t+1)}$ とその継続時間 $D_{s(t+1)}$ をサンプリングできる．ここで，$D_{1:s}^{\mathrm{sum}} = \sum_{s'<s} D_{s'}$ である．

$$\begin{aligned}&p(z_s=i|y_{1:T}, Y_{1:T}, z_{s-1}=j, F_{D_{1:s}^{\mathrm{sum}}}=1) = \\ &\qquad p(i|j)\beta_{D_{1:s}^{\mathrm{sum}}}(i) p(y_{D_{1:s}^{\mathrm{sum}}}, Y_{D_{1:s}^{\mathrm{sum}}}|i)\end{aligned} \tag{34}$$

$$\begin{aligned}&p(D_s=d|y_{1:T}, Y_{1:T}, z_s=i, F_{D_{1:s}^{\mathrm{sum}}}=1) = \\ &\quad p(y_{D_{1:s}^{\mathrm{sum}}+1:D_{1:s}^{\mathrm{sum}}+d}, Y_{D_{1:s}^{\mathrm{sum}}+1:D_{1:s}^{\mathrm{sum}}+d}|d,i,F_{D_{1:s}^{\mathrm{sum}}}=1) \\ &\qquad \cdot p(d)\frac{\beta_{D_{1:s}^{\mathrm{sum}}+d}(i)}{\beta^*_{D_{1:s}^{\mathrm{sum}}}(i)}\end{aligned} \tag{35}$$

以上により，プロソディを考慮した二重分節構造の生成過程から事後分布のサンプリング式を導出した．

5. 実験

5.1 実験目的

本実験では，提案手法が自然発話からの語彙獲得が可能であるかを評価するために，プロソディを含む自然発話データセットを用いて，NPB-DAA と，提案手法 (以降 Prosodic DAA と呼称する．) の音素，及び単語分割精度を比較する．

5.2 実験設定

本実験では，日本語自然発話データセットを用いる．ここでは，自然発話を日常会話と相違ない発話と定義する．このデータセットは，日本語自然発話で構成された 70 文を 1 回ずつ男性の話者に発話してもらい収録した音声データセットである．本実験で用いる特徴量は，フレーム幅を 25[msec]，フレームシフト長を 10[msec] として変換された 12 次元の MFCC，MFCC の一次微分，及び二次微分のそれぞれの特徴量を用いる．また，それぞれ DSAE で 8 次元，5 次元，3 次元と段階的に抽出したものを結合した，合計 9 次元の特徴量を用いる．F_0 の推定には動的計画法を用いて F_0 を系列として推定する手法である Robust Epoch And Pitch EstimatoR[*1] を用いる．

5.3 実験条件

実験条件として，HDP-HLM の隠れ言語モデルのハイパーパラメータは，$\alpha^{\mathrm{LM}} = 10.0$ と $\gamma^{\mathrm{LM}} = 10.0$ とし，weak-limit 近似における単語上限数を 35 個とした．同様に，隠れ単語モデルでは，$\alpha^{\mathrm{WM}} = 10.0$ と $\gamma^{\mathrm{WM}} = 10.0$ とし，weak-limit 近似における音素上限数を 30 個とした．持続時間分布には，$\alpha_0 = 200$，$\beta_0 = 10$ のポアソン分布を仮定し，MFCC の出力分布には事前分布に $\mu_0 = 0$，Σ_0 に単位行列，$\kappa_0 = 0,01$，$\nu_0 = (\mathrm{dimension} + 2)$ の正規逆ウィシャート分布を持つ多変量ガウス分布と仮定した．さらに，提案手法では，プロソディの出力分布には事前分布に，$F_t = 0$ のとき $\mu_0 = 0$，Σ_0 に単位行列，$\kappa_0 = 100$，$\nu_0 = (\mathrm{dimension} + 2)$，$F_t = 1$ のとき $\mu_0 = 1$，Σ_0 に単位行列，$\kappa_0 = 2$，$\nu_0 = (\mathrm{dimension} + 2)$ の正規逆ウィシャート分布を持つ多変量ガウス分布と仮定した．上記条件において，100 イテレーションを 1 試行とし，これを 20 試行行う．

5.4 実験結果

本実験の結果として，音声データの各フレームに正解ラベルを付与し，クラスタリング性能を評価する．クラスタリング性能の定量的評価指標として，各試行で得られた最終サンプリング結果から，音素及び単語の調整ランド指数 (Adjusted Rand Index: ARI) を算出する．ARI はクラスタリング結果が正解ラベルと一致するとき 1 を取り，ランダムの場合は 0 をとる．

表 1，及び図 4 に Prosodic DAA，及び NPB-DAA のそれぞれの音素及び単語の ARI の平均を示す．表 1 より，2 つの手法の結果をみると，Prosodic DAA が音素 ARI，単語 ARI 共に大幅に高い値を出していることがわかる．また，Prosodic DAA と，NPB-DAA 間で t 検定を行ったところ，音素 ARI と単語 ARI 共に優位水準 1%で有意差が認められた．

6. まとめ

本稿では，NPB-DAA によるプロソディを考慮した単語分割を行う手法を提案した．実験では，プロソディを強調した日本語自然発話から語彙獲得実験を行い，提案手法が既存手法と比較して，高い精度で語彙獲得が可能であることを示した．

表 1: 音素及び単語分割の精度比較.

手法	音素 ARI	単語 ARI
Prosodic DAA	**0.370±0.022**	**0.671±0.054**
NPB-DAA	0.261±0.014	0.497±0.072

*1 REAPER: Robust Epoch And Pitch EstimatoR : https://github.com/google/REAPER

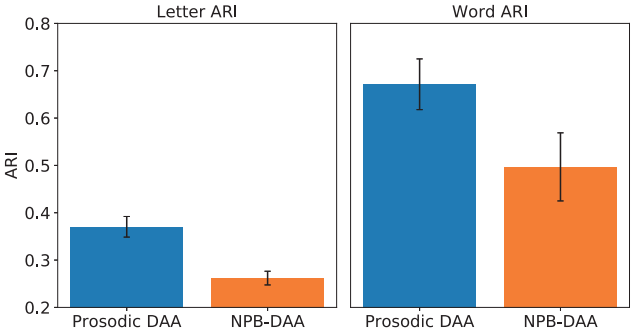


図 4: 音素及び単語分割の精度比較.

今後の課題として，プロソディを強調していない，日本語自然発話からの語彙獲得を行うことが考えられる．

参考文献

[Saffran 96] Saffran, Jenny R., Aslin, Richard N. and Newport, ElissaL.: Statistical Learning by 8-Month-Old Infants, American Association for the Advancement of Science, Vol.274, No.5294, pp.19261928, 1996

[Jusczyk 92] Jusczyk, Peter W and Hirsh-Pasek, Kathy and Nelson, Deborah G Kemler and Kennedy, Lori J and Woodward, Amanda and Piwoz, Julie.:Perception of acoustic correlates of major phrasal units by young infants, Cognitive psychology, Vol.24, Num.2, pp.252-293, 1992

[Taniguchi 16a] Taniguchi, Tadahiro and Nagasaka, Shogo and Nakashima, Ryo.: Nonparametric Bayesian Double Articulation Analyzer for Direct Language Acquisition From Continuous Speech Signals, IEEE Trans. Cognitive and Developmental Systems, Vol.8, Num.3, pp.171185, 2016

[Taniguchi 16b] Taniguchi, Tadahiro and Nakashima, Ryo and Liu, Hailong and Nagasaka, Shogo.: Double articulation analyzer with deep sparse autoencoder for unsupervised word discovery from speech signals, Advanced Robotics, Vol.30, Num.11-12, pp.770783, 2016

[Tada 17] Tada, Yuki and Hagiwara, Yoshinobu and Taniguchi, Tadahiro.: Comparative Study of Feature Extraction Methods for Direct Word Discovery with NPB-DAA from Natural Speech Signals, Joint IEEE International Conference on Development and Learning and Epigenetic Robotics, 2017

[Ozaki 18] Ryo Ozaki and Tadahiro Taniguchi.: Accelerated Nonparametric Bayesian Double Articulation Analyzer for Unsupervised Word Discovery, The 8th Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics, 2018

[Johnson 13] Johnson, Matthew J. and Willsky, Alan S.: Bayesian Nonparametric Hidden Semi-Markov Models, Journal of Machine Learning Research, Vol.14, pp.673701, Feb.2013

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Deep Neuroevolution によるロボティックスワームの二点間往復タスクにおける群れ行動の生成

Generating Collective Behavior of a Robotic Swarm in a Two-landmark Navigation Task with Deep Neuroevolution

森本 大智 *1 Daichi MORIMOTO　平賀 元彰 *1 Motoaki HIRAGA　大倉 和博 *1 Kazuhiro OHKURA　松村 嘉之 *2 Yoshiyuki MATSUMURA

*1広島大学 Hiroshima University　*2信州大学 Shinsyu University

Deep reinforcement learning has provided outstanding results in various applications. Deep neural networks are usually trained by gradient-based methods. However, when deep reinforcement learning is applied to a robotic swarm, that is composed of many robots, it is difficult to design reward functions that lead to a desired collective behavior. In this paper, we applied deep neuroevolution, which is a technique to optimize deep neural networks with artificial evolution, to design controllers of a robotic swarm. Deep neuroevolution is expected to evolve deep neural networks to different reward/fitness landscapes because it optimizes with population-based and gradient-free methods. This paper shows that the controllers designed with deep neuroevolution give robustness to different reward settings compared to deep reinforcement learning.

1. はじめに

Swarm Robotics(SR) は多数のロボットからなる集団に，ロボット間あるいはロボットと環境間の局所的な相互作用によって所望の群れ行動の生成を目指す研究分野である [Sahin 2004]. SR では群れを構成するロボットの台数が増加するに従い，設計者が各ロボットの行動を定義して制御器を設計することが困難となる．そのため，SR では強化学習や進化ロボティクスを制御器設計に適用する試みがなされている [Brambilla 13].

強化学習の分野では近年，深層ニューラルネットワーク (Deep Neural Network,DNN) を組み合わせた深層強化学習 (Deep Reinforcement Learning,DRL) が様々な問題に適用され，良好な成績を記録している [Mnih 15, Lillicrap 15]．一般的に DRL における DNN の学習には，重みの勾配を用いた単点探索型の勾配降下法が用いられる．そのため，多峰性の目的関数における局所解や異なる報酬設定に対応するのが難しい．したがって，DRL をロボティックスワームに適用する場合，性能を向上させるために適切な報酬設定が必要となる．

本稿では Deep Neuroevolution(DNE) によるロボティックスワームの制御器設計を行う．DNE は DNN を進化計算で学習させる手法である．DNE は個体群ベースの探索によって DNN を学習させるため，報酬設定の変化に対して頑健であることが期待される．本稿では異なる報酬条件下で，DRL および DNE によってロボティックスワームの制御器である DNN を設計し，DRL と DNE の報酬設定に対する頑健性，およびロボティックスワームの群れ行動の比較を行う．

2. Deep Neuroevolution

Deep Neuroevolution(DNE) は進化計算によって DNN を学習させる手法である．一般的に DRL における DNN の学習は，重みの勾配を用いた単点探索型の勾配降下法で行われる．そのため多峰性の目的関数における局所解や異なる報酬設定に対応するのが難しい．対して DNE では個体群ベースの探索のためこれらの問題に関して頑健な DNN の学習が期待される．

連絡先: 森本大智，広島大学，morimoto@ohk.hiroshima-u.ac.jp

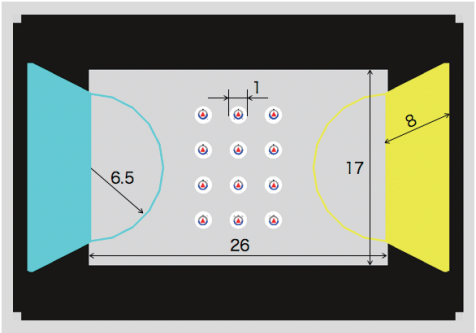


図 1: 実験環境

DNE の先行研究として [Salimans 17, Such 17] がある．Saliman らは三層の畳み込み層と二層の全結合層からなる DNN を Natural Evolution Strategy を用いて学習させた [Salimans 17]．また Such らは同様の構造を持つ DNN を交叉を用いず突然変異のみを用いる Genetic Algorithm によって学習させた [Such 17]．これらの先行研究は Atari 2600 などの DRL の問題に適用されており，本稿のロボティックスワームのような多数のエージェントを扱う問題には適用されていない．

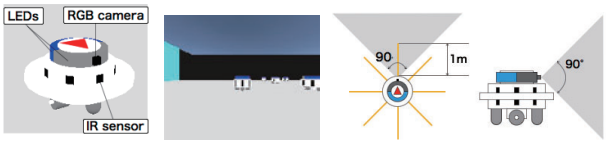
3. 計算機実験

3.1 二点間往復タスク

実験には二点間往復タスクを用いる．このタスクの目的は二つのターゲットエリア間の往復である．実験環境を図 1 に示す．環境の両端に黄色とシアンのランドマークを配置する．ランドマークの中心から半径 6.5m の範囲がターゲットエリアである．ロボットは目標とするターゲットエリアに到達時，目標をもう一方のターゲットエリアに変更する．ロボットは各エピソード開始時に環境の中央に格子状に配置される．

3.2 ロボット

図 2(a) に実験に用いた移動ロボットを示す．ロボットの直径は 1m であり左右の二つの車輪とモータによって駆動する．



(a) 移動ロボット (b) カメラ画像 (c) センサ設定

図 2: ロボットの設定

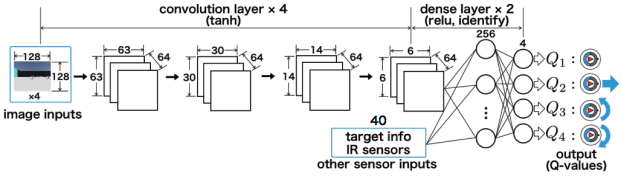


図 3: ロボットの制御器

最高移動速度は 1 m/s である．本実験ではロボットが取る行動を停止，直進，左回転，右回転の 4 種類とする．ロボットはセンサとしてカメラ 1 個と距離センサ 8 個を搭載している．カメラは 128×128 pixels の RGB 画像を生成する．図 2(b) にカメラから得られる画像の例を示す．距離センサは測定範囲内の物体までの距離を測定する．各センサの測定範囲を図 2(c) に示す．ロボットの上部に搭載された LED は後方のみ点灯し，黄色のランドマークを目指す場合は赤色，シアンの場合は青色を点灯する．

3.3 制御器

ロボットの制御器として図 3 に示す DNN を用いる．制御器は入力層，4 層の畳み込み層，全結合層，出力層からなる．制御器への入力はカメラ画像，距離センサ情報，目標ランドマーク情報であり，それぞれ過去 4 タイムステップまでの情報が入力される．畳み込みに使用するフィルタはサイズを 4 × 4，チャネル数をそれぞれの層の入力に対応させたものを 64 種類用意する．ストライドは 2 とし，これらフィルタに関する設定は 4 層の畳み込み層で共通とする．また各畳み込み層の直前に Batch Normalization 層を加える．

3.4 実験設定

ここでは DRL と DNE における共通の実験設定について述べる．実験の最終世代数は 500 とし，1 世代は 1000 タイムステップとする．制御器は 24 個とし，DRL ではこれらを独立に学習させ，DNE では進化アルゴリズムによって学習させる．またロボットに対する報酬は以下のものを用いる．

$$\begin{aligned} r_{d,i,t} &= 5 \times (d_{i,t-1} - d_{i,t}) \\ r_{e,i,t} &= 5 \\ r_{c,i,t} &= -5 \end{aligned} \tag{1}$$

ここで i はロボットの番号，t はタイムステップを表す．$r_{d,i,t}$ はロボットが目標ランドマークに近づいた距離に応じて与えられる．$d_{i,t}$ はタイムステップ t におけるロボット i の重心とランドマークの重心間の距離を示す．$r_{e,i,t}$ はターゲットエリアへの到達時に与えられる．$r_{d,i,t}$ と $r_{e,i,t}$ は目標ランドマークがカメラに映っていた場合のみ適用される．また $r_{c,i,t}$ は距離センサが他のロボットや壁との接触を検知した際に与えられる．これらの報酬を組み合わせた二つの条件下で実験を行う．一つ目の条件では $r_{d,i,t}$，$r_{e,i,t}$，$r_{c,i,t}$ の総和によって報酬を決定する．1 タイムステップにおける報酬 R_t は以下の式で与えられ，これを「報酬設定 (i)」とする．

$$R_t = \sum_i (r_{d,i,t} + r_{e,i,t} + r_{c,i,t}) \tag{2}$$

二つ目の条件では $r_{e,i,t}$ のみによって報酬を決定する．1 タイムステップにおける報酬 R_t は以下の式で与えられ，これを「報酬設定 (ii)」とする．

$$R_t = \sum_i r_{e,i,t} \tag{3}$$

またロボットの各タイムステップにおける行動は ϵ-greedy 法によって決定される．第 1 エピソードでは $\epsilon = 1$ に設定しランダムに行動を選択する．また 1 エピソード目は制御器の学習を行わない．2 エピソード以降は $\epsilon = 0.1$ に固定し行動を選択しつつ，制御器の学習を行う．

3.5 制御器の学習

DRL

DRL のアルゴリズムとして DQN[Mnih 15] を用いる．DQN では Experience replay によって蓄積した環境の遷移情報を再生し毎タイムステップ制御器の重みを更新する．保有する経験はロボット台数 ×1 エピソードのタイムステップサイズ分であり，全ロボットの経験が共有される．学習時のミニバッチサイズは 32 とし，optimizer には RMSpropGraves[Graves 13] を用いる．

DNE

DNE の場合は 1 エピソード間に獲得した報酬の和 $\sum_t R_t$ を制御器の適応度とし，エピソード終了時に制御器の更新を行う．制御器の進化アルゴリズムとして文献 [Such 17] を参考としたものを用いる．各エピソード終了時にトーナメント選択により親個体を選択する．本実験ではトーナメントサイズを 2 とする．また，エリート選択により最高の適応度を記録した制御器を一つ，遺伝的操作を行わずに次世代に引き継ぐ．選択した親個体に対し突然変異を適用し次世代個体を生成する．突然変異操作は以下の式で表される．

$$\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} + \boldsymbol{\epsilon}\sigma \tag{4}$$

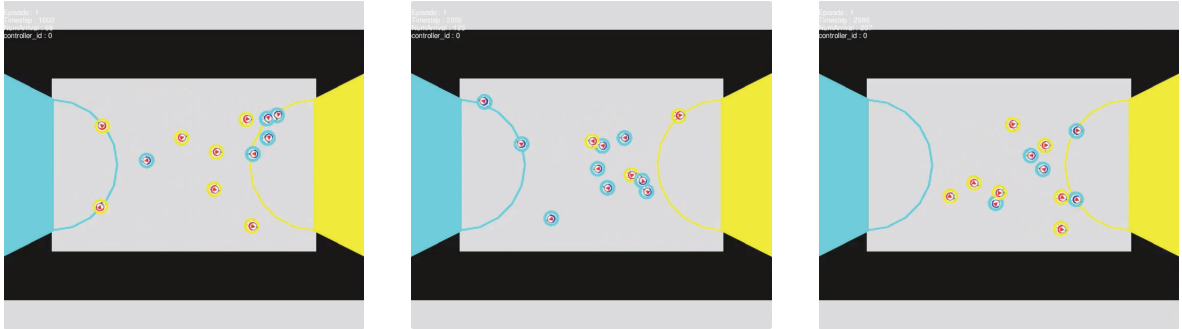
ここで $\boldsymbol{\theta}$ は DNN のパラメータベクトル，$\boldsymbol{\epsilon}$ は $\boldsymbol{\theta}$ と同じ要素数を持つノイズベクトル，σ は遺伝子に与える揺動のスケーリングパラメータである．$\boldsymbol{\epsilon}$ の各要素は標準正規分布よりサンプリングされる．また本実験では σ=0.02 とする．進化させる DNN のパラメータは全結合荷重値と Batch Normalization 層のアフィン変換パラメータとする．

4. 結果

4.1 報酬設定 (i)

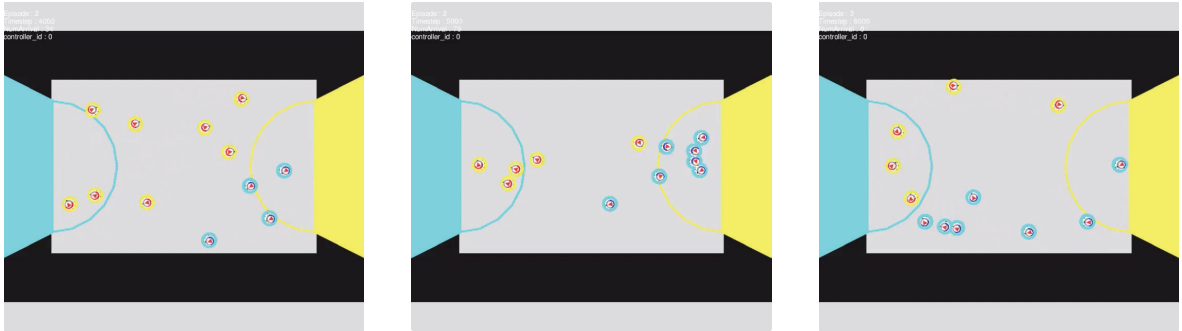
学習時におけるロボットの到達回数の推移を図 4 に示す．グラフは実験 3 試行分のデータを示し，平均値，最大値，標準偏差は 72 個の制御器のデータから算出されている．DRL ではおよそ 100 世代にかけて到達回数が急激に上昇しその後ほぼ一定の値を示す．DNE では 60 世代付近まで平均値がほぼ上昇せず，その後 DRL に比べ比較的緩やかに上昇する．

次に学習中に記録された制御器を用いてロボティックスワームの群れ行動生成を行なった結果を図 6，図 7 に示す．DRL の場合，ロボットは他のロボット間を抜け，それぞれのロボットが独立にターゲットエリアを目指す振る舞いを示す．対して DNE の場合，ロボットは他のロボットを追従し環境内を輪を描くように移動する振る舞いを示す．



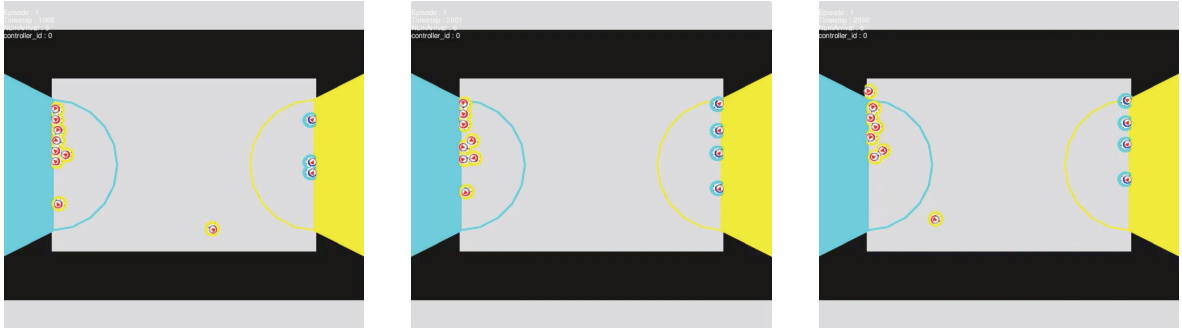
(a) 1000 time steps (b) 2000 time steps (c) 3000 time steps

図 6: 報酬設定 (i) の条件下で DRL が生成した群れ行動



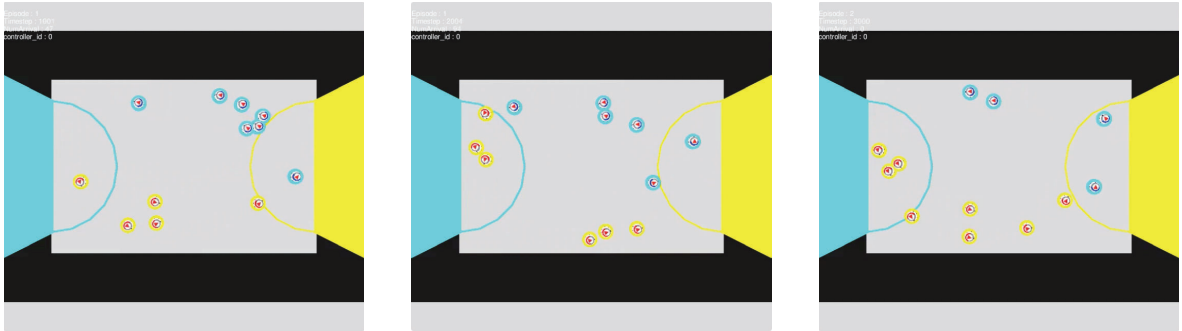
(a) 1000 time steps (b) 2000 time steps (c) 3000 time steps

図 7: 報酬設定 (i) の条件下で DNE が生成した群れ行動



(a) 1000 time steps (b) 2000 time steps (c) 3000 time steps

図 8: 報酬設定 (ii) の条件下で DRL が生成した群れ行動



(a) 1000 time steps (b) 2000 time steps (c) 3000 time steps

図 9: 報酬設定 (ii) の条件下で DNE が生成した群れ行動

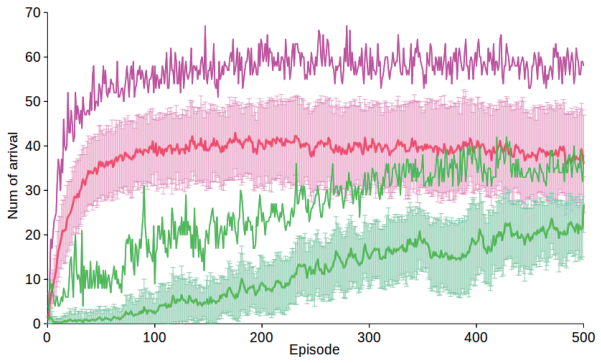


図 4: 報酬設定 (i) における到達回数の推移

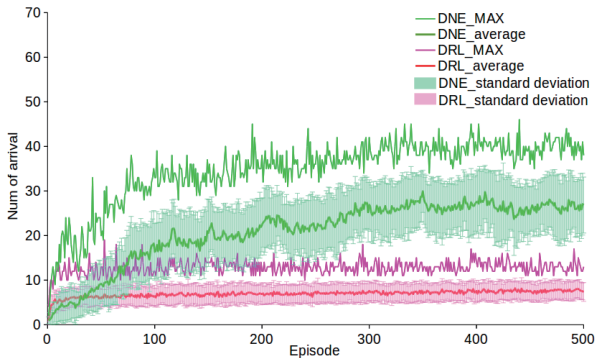


図 5: 報酬設定 (ii) における到達回数の推移

4.2 報酬設定 (ii)

報酬設定 (i) の場合と同様に到達回数の推移を表したグラフを図 5 に示す．DRL では報酬設定 (i) と同様に到達回数の急激な上昇が見られる．しかし，定常時の平均値，最大値は報酬設定 (i) の場合と比べ低下している．DNE では，報酬設定 (i) で見られた 70 世代付近までの停滞が見られず，平均値は学習初期からほぼ単調に増加している．

次に報酬設定 (i) の場合と同様にロボティックスワームの群れ行動生成を行なった結果を図 8，図 9 に示す．DRL の場合，ロボットはターゲットエリアに侵入後ランドマークに衝突し，その後ほぼ移動を行わない．対して DNE の場合，報酬設定 (i) と同様に環境内を輪を描くように移動する振る舞いを示す．

5. 考察

報酬設定 (i) において DNE は，学習初期の 60 世代付近まで到達回数の上昇が滞っている．これは報酬設定 (i) における罰則 $r_{c,i,d}$ によるものであると考えられる．ターゲットエリアへの到達報酬 $r_{e,i,t}$ が発生する頻度は $r_{d,i,t,}$，$r_{c,i,t}$ よりも低い．対して $r_{d,i,t,}$，$r_{c,i,t}$ は条件を満たせば毎タイムステップ発生する可能性がある．またロボットの最高速度を考慮すると，1 タイムステップに発生する $r_{d,i,t,}$ の最大値は 1.0 となり，罰則 $r_{c,i,t}$ の五分の一である．このことから，学習初期において発生する報酬は $r_{c,i,t}$ が支配的であると考えられ，進化の過程において到達回数が多い制御器よりも，衝突を回避する制御器が有利となることで到達回数の停滞が起こると考えられる．

また，報酬設定 (ii) における DRL の到達回数は報酬設定 (i) と比較して低下している．これは与えられる報酬がスパースになるためと考えられる．上述の通り，ターゲットエリアへの到達報酬 $r_{e,i,d}$ は発生する頻度が低い．そのため Experience Replay に用いる記憶領域において，報酬に関する情報が乏しくなることで性能が低下すると考えられる．よって，より大きい記憶領域を用い，Prioritized Experience Replay 等のアルゴリズムを用いることで性能の向上が期待される．

また，本実験で用いた二つの報酬条件において，DNE は最大値やロボットの振る舞いにおいて近しい性能を示したことから，報酬条件に対してより頑健な学習が行えると考えられる．

6. おわりに

本稿では Deep Neuroevolution をロボティックスワームの制御器設計に適用し，二点間往復タスクにおいて群れ行動の生成を行なった．Deep Neuroevolution を適用したロボティックスワームは二つの異なる報酬条件下においてタスクの達成が可能であることを示した．

今後の展望として，より複雑な報酬条件下でロボティックスワームの異なる振る舞いが得られるか実験を行う．また，協調搬送タスクや経路形成タスクに Deep Neuroevolution を適用した場合にタスク達成が可能であるか実験を行う．

参考文献

[Sahin 2004] Erol Sahin: Swarm robotics: from sources of inspiration to domains of application, *Swarm Robotics: SAB 2004 International Workshop*, Vol. 3342 of *Lecture Notes in Computer Science*, pages 10-20. Springer, 2005.

[Brambilla 13] Manuele Brambilla, Eliseo Ferrante, Mauro Birattari, and Marco Dorigo :Swarm robotics: a review from the swarm engineering perspective, *Swarm Intelligence*, Vol. 7, pages 1-41, 2013.

[Mnih 15] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al.:Human-level control through deep reinforcement learning, *Nature, 518(7540):529*, 2015.

[Lillicrap 15] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, Daan Wierstra :Continuous control with deep reinforcement learning, *arXiv preprint arXiv:1509.02971*, 2015.

[Salimans 17] Tim Salimans, Jonathan Ho, Xi Chen, Szymon Sidor, Ilya Sutskever :Evolution strategies as a scalable alternative to reinforcement learning, *arXiv preprint arXiv:1703.03864*, 2017.

[Such 17] Felipe Petroski Such, Vashisht Madhavan, Edoardo Conti, Joel Lehman, Kenneth O Stanley, Jeff Clune :Deep neuroevolution: genetic algorithms are a competitive alternative for training deep neural networks for reinforcement learning, *arXiv preprint arXiv:1712.06567*, 2017.

[Graves 13] Alex Graves:Generating sequences with recurrent neural networks, *arXiv preprint arXiv:1308.0850*, 2013.

Organized Session | Organized Session | [OS] OS-4

[3D4-OS-4b] 自律・創発・汎用 AIアーキテクチャ(2)

栗原 聡（慶應義塾大学）、川村 秀憲（北海道大学）、津田 一郎（中部大学）、大倉 和博（広島大学）

Thu. Jun 6, 2019 3:50 PM - 5:10 PM Room D (301B Medium meeting room)

[3D4-OS-4b-01] Introducing a Call Stack into the RGoal Hierarchical Reinforcement Learning Architecture

〇Yuuji Ichisugi[1], Naoto Takahashi[1], Hidemoto Nakada[1], Takashi Sano[2] （1. National Institute of Advanced Industrial Science and Technology (AIST), 2. Department of Computer and Information Science, Faculty of Science and Technology, Seikei University）

3:50 PM - 4:10 PM

[3D4-OS-4b-02] Avoiding catastrophic forgetting in echo state networks by minimizing the connection cost

〇Yuji Kawai[1], Yuho Ozasa[1], Jihoon Park[1], Minoru Asada[1] （1. Osaka University）

4:10 PM - 4:30 PM

[3D05-06-3] Special program

4:30 PM - 5:10 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

階層型強化学習 RGoal アーキテクチャへの再帰呼び出し用スタックの導入

Introducing a Call Stack into the RGoal Hierarchical Reinforcement Learning Architecture

一杉裕志 *1 Yuuji Ichisugi　高橋直人 *1 Naoto Takahashi　中田秀基 *1 Hidemoto Nakada　佐野崇 *2 Takashi Sano

*1産業技術総合研究所 人工知能研究センター
National Institute of Advanced Industrial Science and Technology (AIST), AIRC

*2成蹊大学 理工学部 情報科学科
Department of Computer and Information Science, Faculty of Science and Technology, Seikei University

Humans can set suitable subgoals in order to achieve some purposes, and furthermore, can set sub-subgoals recursively if needed. It seems that the depth of the recursion is unlimited. Inspired by this behavior, we had designed a hierarchical reinforcement learning architecture, the RGoal architecture. In this paper, we introduce a call stack into the RGoal architecture to increase reusability of subgoals. We evaluate its performance using a maze with multi-task setting. The result shows that the convergence speed improves as the maximum stack size increases.

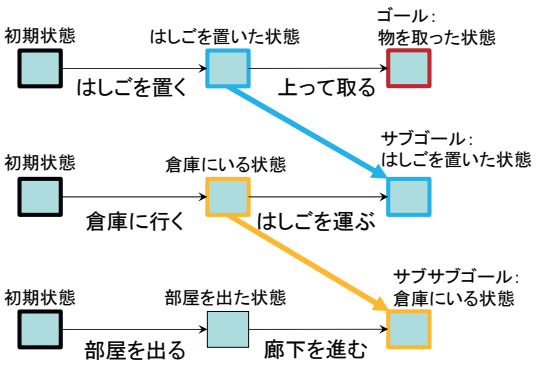


図 1: 人間が目標を達成するために、サブゴール（副目標）を再帰的に設定する例。高いところにあるものを取る（ゴール）ために、まずはしごを取る必要がある（サブゴール）。さらにはしごを手に入れるために倉庫にいく必要がある（サブサブゴール）。

1. はじめに

人間は何か目標を達成するために適切なサブゴール（副目標）を設定できる。さらに必要に応じてサブサブゴールを再帰的に設定することもでき、その再帰の深さには制約がないように見える（図 1）。この振る舞いにヒントを得た階層型強化学習のアーキテクチャとして、我々は RGoal アーキテクチャを提案した [一杉 18a]。我々はこの RGoal の機能を拡張していき、ヒトの脳の前頭前野周辺の情報処理を再現し、やがては汎用人工知能を実現するための中核技術とすることを目指している。

先行研究の 1 つ MAXQ [Dietterich 00] は、多層（ただし深さは固定）の階層型強化学習アーキテクチャであり、以下の 3 つの特徴により学習を効率化している。

1. Subtask sharing: マルチタスク環境において、サブルーチンをタスク間で共有することで学習を速くする。
2. Temporal abstraction: 複雑なタスクの学習において、より単純なサブルーチンの組み合わせ方のみに試行錯誤を限定することで、学習を速くする。
3. State abstraction: サブルーチンごとに実行に差し支えない程度に状態を抽象化することで、サブルーチンの学習を速くする。

RGoal では 1. は価値関数の分解と共有 (2.2 節)、 2. は思考モードと呼ぶ機構 (2.4 節) により実現する。 3. は現在のところ未実装だが、単一化の機構を利用したテーブル圧縮手法を検討している [一杉 18b]。

RGoal では、エージェントによるサブゴール設定は、プログラミング言語におけるサブルーチン呼び出しと似た振る舞いをする。実際、 RGoal は「強化学習によるプログラム合成システム」と見なすことができる [一杉 18b]。強化学習によるプログラム合成は汎用人工知能実現に向けた有望なアプローチの 1 つである。例えば、AIXI[Hutter 00] はチューリングマシンのプログラムを、UCAI[Katayama 18] はより一般的な文法を持った言語のプログラムを強化学習で合成する汎用人工知能の理論である。また、このアプローチで動作するシステムとして、DNC(Differentiable Neural Computers)[Graves 16] や MagicHaskeller [Katayama 08] がある。

以前提案した RGoal では、エージェントはスタックを持っておらず、サブゴール到達後には、それまでのコンテキストとは無関係にあらためてもともとのゴールを目指すように設計されていた。しかし、そのようなアーキテクチャでは、サブルーチンとして獲得された知識の再利用性が悪くなる [一杉 18b]。脳の中に頑健に動作するスタックがあるとは考えにくいが、それを代替する何らかの機構があることは考えられる。そこで本稿では、 RGoal にサブゴールを保存するスタックを導入し、それに伴い学習則も修正する。そして、迷路課題を用いて性能を評価する。また、 RGoal には思考モードとよぶ、演繹推論を行うための特殊なモデルベース強化学習の機構があるが、その性能も以前よりも詳しく評価する。

連絡先: 一杉裕志、 茨城県つくば市梅園 1 − 1 − 1 中央第 1 産業技術総合研究所、 y-ichisugi@aist.go.jp

2. RGoal アーキテクチャ

2.1 アーキテクチャの概要

RGoal アーキテクチャ[一杉 18a] について簡単に説明する。

エージェントは各ステップごとに、プリミティブ行動を取るかサブルーチン呼び出しを行うかのどちらかを選択する。サブルーチンの間に上下関係はなく、すべてのサブルーチンは対等であり、相互再帰的に呼び出しが可能である。

ここでは**サブルーチン** g を、「任意の環境の状態からある 1 つの状態（サブゴール） g に向かう方策」と定義する。また、サブルーチン g を実行し続けることで、環境の状態はいつか必ず g に到達すると仮定する。

Option-Critic アーキテクチャ[Bacon 17] などではサブルーチンの終了時の環境の状態は一意に決まらず、終了条件は学習によって獲得される。一方 RGoal では、終了条件はセンサー入力から得られる外界の特徴量の 1 つだと考える。つまり、終了条件は強化学習アーキテクチャの本体によってではなく、特徴抽出の機構によって獲得されるものと想定している。

2.2 価値関数分解

行動価値関数をサブルーチンごとに分解し、サブルーチンを複数のタスク間で共有することで、学習速度を上げることができる [Dietterich 00]。

ゴール G 、環境の状態 s 、サブゴール g としたとき、行動 a を取った後、方策 $\pi((s,g),a)$ に従って行動し続けたときに得られる報酬の総和の期待値を、行動価値関数 $Q^{\pi}_{G}((s,g),a)$ と定義する。（報酬割引は行わないものとする。）これは以下のように、g への到着前と到着後に分解することができる。

$$Q^{\pi}_{G}((s,g),a) = Q^{\pi}(s,g,a) + V^{\pi}_{G}(g) \tag{1}$$

ここで、 $Q^{\pi}(s,g,a)$ は状態 s において行動 a を取った後、方策 π に従って行動しサブゴール g に到着するまでの報酬の総和の期待値である。また、 $V^{\pi}_{G}(g)$ は、サブゴール g に到着後に方策 π に従って行動しゴール G に到着するまでの報酬の総和の期待値で、

$$V^{\pi}_{G}(g) \quad = \quad \Sigma_a \pi((g,G),a) Q^{\pi}(g,G,a) \tag{2}$$

として計算できる。

$Q^{\pi}(s,g,a)$ はもともとのゴール G に依存しないため、複数のタスク間で共有することができる。

なお、行動選択の式は通常の強化学習のもの同様であり、例えばテーブル Q のもとでグリーディーに行動を選択する場合は以下のようにする。

$$a' \quad = \quad \underset{a}{\operatorname{argmax}}\, Q(s,g,a) \tag{3}$$

2.3 スタックを導入した場合の学習則

ここまでの定義はスタックがない RGoal [一杉 18a] と本質的に違いはない（記法は一部変更した）が、学習則はスタックの導入に伴い、少し変更が必要となる。

スタックを導入した RGoal では、サブルーチン g' を呼び出した時に現在のサブゴール g をスタックに積み、サブルーチンの実行が終了した時、すなわち状態が s から g' に変化した時に、スタックから取り出した g をサブゴールに再設定する。サブルーチン呼び出し前は、想定される状態の変化は $s \to g \to G$ であるが、呼び出しによって $s \to g' \to g \to G$ に変化する。したがって以下の式が成り立つ。

$$\begin{aligned} & Q_G((s',g'),a') - Q_G((s,g),a) \\ = \quad & (Q(s',g',a') + V_g(g') + V_G(g)) - (Q(s,g,a) + V_G(g)) \\ = \quad & Q(s',g',a') - Q(s,g,a) + V_g(g') \end{aligned} \tag{4}$$

ただし、 s', a' は呼び出しの次のステップでの状態と行動である。この式は、行動が a がサブルーチン呼び出しに限らず、プリミティブ行動の場合も成り立つ。したがって、Sarsa で学習する場合の Q の更新式は以下のようになる。

$$\begin{aligned} & Q(s,g,a) \leftarrow Q(s,g,a) \\ & +\alpha(r + Q(s',g',a') - Q(s,g,a) + V_g(g')) \end{aligned} \tag{5}$$

この学習則には以前のもの [一杉 18a] と違い、もともとのゴール G が現れておらず、大域的な文脈に依存しない知識が獲得しやすくなっている。

2.4 思考モード

階層型強化学習には、学習済みの簡単なタスクを組み合わせて複雑なタスクを近似的に、しかし高速に解くという目的もある。RGoal アーキテクチャには、そのための機構として思考モードがある。思考モードは、学習済みの $Q(s,g,a)$ を環境のモデルと見なした一種のモデルベース強化学習 [Sutton 90] である。

本稿では [一杉 18a] とは違い、評価中の各エピソードの実行の直前に、思考モードにおける実行を定数回行うように実装を変更し、評価を行った。この振る舞いは、ヒトや動物が新たな状況に直面した時、知識を組み合わせて問題解決の行動計画を立ててから実行する振る舞いと似たものになっている。

2.5 サブルーチンの中断

以前の Roal アーキテクチャ[一杉 18a] には、実行中のサブルーチンを中断し、別のサブゴールに切り替える機能がある。サブルーチンの中断が可能であれば行動の自由度が増し、性能が上がる可能性がある [Kaelbling 93][Sutton 99][Dietterich 00] が、今回の実装ではアーキテクチャを簡単にするため、中断の機能を取り除いた。

2.6 アルゴリズム

以上の結果をまとめた、Sarsa に基づくアルゴリズムの疑似コードを図 2 に示す。

3. 評価

今回もアルゴリズムの基本動作の確認を行うことが目的のため、実行中の振る舞いの可視化が容易な迷路タスクを題材として性能を評価した。

ここでは、時間をかけた厳密解への収束ではなく、準最適解にできるだけ速やかに収束することを重視する。

マップとランドマークの集合は固定である（図 3）。ランドマークの中からエピソードごとにスタート S とゴール G がランダムに選ばれる。エージェントが S から移動して G に到達したときに与えられる報酬は 0 で、その時点でそのエピソードを終了し、スタートとゴールを変えて次のエピソードを始める。上下左右の移動は -1 、斜めの 4 方向いずれかへの移動は $-\sqrt{2}$ 、壁への衝突は -1 、サブルーチン呼び出しの実行は $R^{\mathcal{C}} = -1$ の報酬が与えられる。（前に述べたように報酬割引はない。）ランドマークはサブゴールの候補でもあり、ランドマークの中の 1 つだけがある時点でのサブゴールになり得る。

procedure EPISODE(S, G, think-flag)
  $s \leftarrow S; g \leftarrow G$
  $stack \leftarrow empty$
  Choose a from s, g using policy derived from Q
  while $s \neq G$ **do**
    # Take action.
    if $a = RET$ **then**
      $s' \leftarrow s; g' \leftarrow stack.pop(); r \leftarrow 0$
    else if a is C_m **then**
      $stack.push(g)$
      $s' \leftarrow s; g' \leftarrow m; r \leftarrow R^{\mathcal{C}}$
    else
      if think-flag **then**
        $s' \leftarrow g; g' \leftarrow g; r \leftarrow Q(s, g, a)$
      else
        Take action a, observe r, s'
        $g' \leftarrow g$
    # Choose action.
    if $s' = g'$ **then**
      $a' \leftarrow RET$
    else
      Choose a' from s', g'
        using policy derived from Q
    # Update.
    if $s = g$ or (think-flag and a is not C_m) **then**
      # Do nothing.
    else
      $Q(s, g, a) \leftarrow Q(s, g, a)$
        $+\alpha(r + Q(s', g', a') - Q(s, g, a) + V_g(g'))$
    $s \leftarrow s'; g \leftarrow g'; a \leftarrow a'$

図 2: 1 つのエピソードを実行する Sarsa に基づくアルゴリズムの疑似コード。テーブル Q の初期化方法については [一杉 18a] を参照。

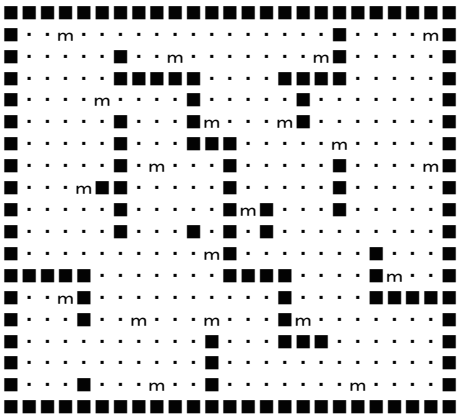


図 3: 評価に用いた 2 次元格子上の迷路のマップ。20 個のランドマーク (m で示した) がマップ上に配置されている。このランドマークの中からエピソードごとにランダムにスタート S とゴール G が選択される。

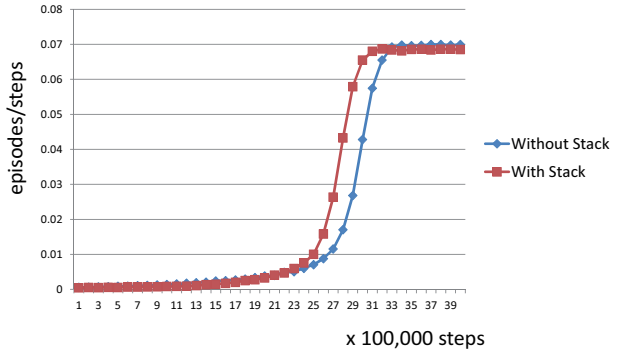


図 4: 実験 1 ：スタックのない RGoal [一杉 18a] と、スタックを導入した本稿のアルゴリズムとの比較。実験条件にもよるが、同程度の早さで収束している。

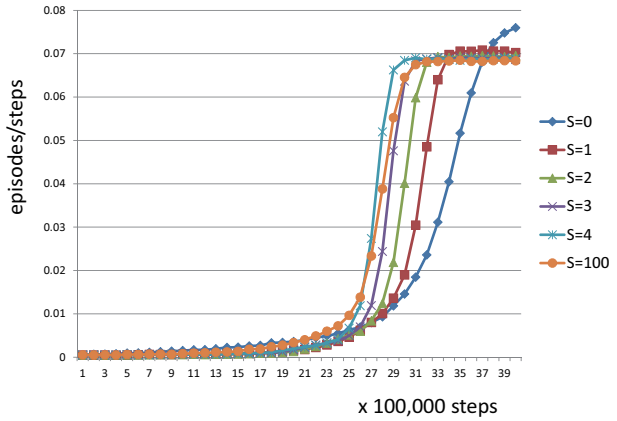


図 5: 実験 2 ：スタックの深さの上限 S と性能の関係。S=0 は普通の強化学習、S=1 は 2 層の階層型強化学習に相当する。上限が大きいほど収束は早くなる傾向を示している。

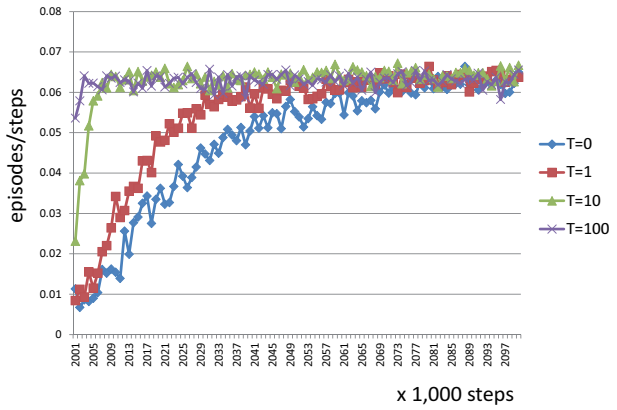


図 6: 実験 3 ：各ステップの実行直前に実行する思考フェーズの長さと性能の関係。思考フェーズの長さ T は、各エピソードの直前に思考モードで「脳内シミュレーション」するエピソードの回数。思考フェーズが十分長ければ、未経験のタスクであっても、近似解がゼロショットで得られている。

テーブルの初期値は、$s = g$ での $Q(s, g, a)$ に対しては 0 、それ以外は $-50 - n$ (n は小さなノイズ) に初期化した。また、学習を効率化するため、サブルーチン呼び出しはランドマーク上でのみ実行可能とした。

行動選択は softmax を用い、逆温度 $\beta = 1$ とした。学習率は $\alpha = 0.1$ である。

実験 1 、 2 、 3 はいずれも 1 0 回の実行結果の平均である。横軸はステップ数である。縦軸はステップ数あたりのエピソード数であり、値が大きいほど、各エピソードを短いステップで解けるということを示している。ここでステップ数とはマップ内の移動もしくは壁への衝突の回数である。サブルーチン呼び出し C_m 、サブルーチンからの復帰 RET の実行回数、思考モードでの実行ステップは、いずれもエージェントの内部で起こる仮想的な行動であると考え、ステップ数には含めない。

実験 1 （図 4）はスタックがないがサブルーチン中断の機能を持ったアルゴリズム [一杉 18a] と、スタックを導入した本稿のアルゴリズムとの比較である。どちらも同程度の早さで収束している。この実験条件では、スタックありの方が、サブルーチン中断の機能がないにもかかわらず、わずかに収束が早くなっている。ただし、実験条件によって結果は変わり、[一杉 18a] で用いた迷路の場合では、サブルーチン中断が起きやすく、本稿のアルゴリズムの方が収束はわずかに遅くなる。

実験 2 （図 5）はスタックの深さの上限 S と性能の関係を調べたものである。S=0 は普通の強化学習、S=1 は 2 層の階層型強化学習に相当する。実験結果は、 S が大きいほど収束が早くなる傾向を示している。（ただし S=100 では S=4 より収束が若干遅くなってる。）一方で、収束後のスコアは、サブルーチン呼び出しを一切行わない S=0 がもっともよくなっている。これは、サブルーチンが使える場合、ランドマークを経由して多少遠回りする局所解にとどまってしまうためである。なお、サブルーチンが使える場合でも、逆温度 β を小さくし探索傾向を強くして十分学習させた後、β を大きくすることで、遠回りしないほぼ最短経路に収束することを確認している。

実験 3 （図 6）は各エピソードの実行直前の思考フェーズの長さ T と性能の関係を調べたものである。ここで T は、各エピソードの直前に、与えられた S と G のもとで思考モードで実行するエピソードの回数である。評価の前に簡単なタスクのみを用いた事前学習フェーズを 2,000,000 ステップ実行し、そのあとの 1,000 ステップごとのスコアをプロットしている。事前学習フェーズでは、ユークリッド距離が 8 以内にある S と G のみを選択して実行する。これは、すべてのランドマークのペア 20x19=380 個のうちの 60 個である。短い距離にあるランドマーク間の移動方法が 事前学習フェーズで獲得され、長い距離の移動方法の近似解は、それをつなぐことで得られる。思考フェーズでは、そのような近似解が、実際に行動することなく、脳内シミュレーションだけで「演繹的」に獲得される。実験結果は、思考フェーズが十分に長ければ、経験したこのない未知のタスクであっても、それまでの経験で得た知識を組み合わせることで、近似解がゼロショットで得られることを示している。

4. まとめと今後

階層型強化学習アーキテクチャ RGoal にスタックを導入したアルゴリズムを実装し、評価した。今後このアーキテクチャを拡張し、状態抽象の機構を追加するなどした上で、より複雑なタスクでの評価を行っていく。

謝辞

ディー・エヌ・エー 甲野佑氏、東京電機大 高橋達二氏との議論から研究の示唆をいただいており、深く感謝いたします。

本研究は JSPS 科研費 JP18K11488 の助成を受けたものです。

参考文献

[一杉 18a] 一杉裕志, 高橋直人, 中田秀基, 佐野崇, RGoal Architecture:再帰的にサブゴールを設定できる階層型強化学習アーキテクチャ, 第 9 回 人工知能学会 汎用人工知能研究会 (SIG-AGI), 2018.

[一杉 18b] 一杉裕志, 高橋直人, 中田秀基, 佐野崇, 単一化の機構を利用した階層型強化学習のテーブル圧縮手法の検討, 第 10 回 人工知能学会 汎用人工知能研究会 (SIG-AGI), 2018.

[Dietterich 00] Thomas G. Dietterich, Hierarchical Reinforcement Learning with the MAXQ Value Function Decomposition, Journal of Artificial Intelligence Research 13, 227–303, 2000.

[Hutter 00] Hutter, M., A theory of universal artificial intelligence based on algorithmic complexity, Technical Report: cs.AI/0004001, 2000. `http://arxiv.org/abs/cs.AI/0004001`

[Katayama 08] S. Katayama, Efficient Exhaustive Generation of Functional Programs using Monte-Carlo Search with Iterative Deepening , PRICAI 2008 , LNAI 5351, Springer Verlag, 199-211, 2008.

[Graves 16] A. Graves, G. Wayne et al., Hybrid computing using a neural network with dynamic external memory, Nature 538, 471–476, 2016.

[Katayama 18] Susumu Katayama, Computable Variants of AIXI which are More Powerful than AIXItl, 2018. `https://arxiv.org/abs/1805.08592`

[Sutton 90] Sutton, R. S., Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. In Proceedings of the Seventh International Conference on Machine Learning, 1990.

[Kaelbling 93] Kaelbling, L.P.: Hierarchical Learning in Stochastic Domains: Preliminary Results. In: Proceedings of the 10th International Conference on Machine Learning, pp. 167–173, 1993.

[Sutton 99] Sutton, R. S.; Precup, D.; and Singh, S. P., Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. Artificial Intelligence 112(1-2):181–211, 1999.

[Bacon 17] Bacon, P.-L., Harb, J., Precup, D. The option-critic architecture. Proceedings of AAAI, 1726–1734, 2017.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

結合コスト最小化によるエコーステートネットワークの破滅的忘却の回避

Avoiding catastrophic forgetting in echo state networks by minimizing the connection cost

河合 祐司 *1 Yuji Kawai　小笹 悠歩 *2 Yuho Ozasa　朴 志勲 *1 Jihoon Park　浅田 稔 *1 Minoru Asada

*1大阪大学先導的学際研究機構 Institute for Open and Transdisciplinary Research Initiatives, Osaka University

*2大阪大学工学部 School of Engineering, Osaka University

Catastrophic forgetting is one of big issues in multi-task learning with neural networks. We propose that minimization of the connection cost mitigates catastrophic forgetting in echo state networks. The optimization of connections in reservoirs can yield neural modules (local sub-networks) that differentiate information flow depending on tasks. The task-specific neural activities help to consolidate knowledges of the tasks. We showed that this constraint creates neural modules consisting of negative connections and can improved the performance of multi-task learning. Furthermore, we analyzed the transfer entropy of inter- and intra-modules to show task-specific functional differentiation of the modules.

1. はじめに

ヒトをはじめとした生物は，新しい環境や様々なタスクへ即時的に適応し，生涯を通じて学習し続ける．ヒトの神経ネットワークは異なる機能を有するモジュール構造（部分ネットワーク）を成し，そのモジュールが動的に切り替わることによって，柔軟な認知活動が実現されていると考えられている [Meunier 09]．そのようなヒトの柔軟で適応的な学習のメカニズムの理解は，汎用的な人工ニューラルネットワークの設計において有用となろう．

逐次的に与えられるデータをオンライン学習する枠組みは，人工ニューラルネットワーク分野においても継続学習 (continual learning) として近年注目を集めている [Parisi 18]．しかし，複数タスクの継続学習には，破滅的忘却 (catastrophic forgetting) の問題がある [McCloskey 89]．これは，新しいタスクの学習によって，過去に学習したタスクの情報が上書きされ，複数のタスクの学習を両立できないという問題である．

この破滅的忘却の回避のために，ニューラルネットワークにモジュール構造を取り入れることが提案されている [Bongard 11, Ellefsen 15]．これらの研究に共通するアイデアは，複数のモジュールがそれぞれ異なるタスクに対応するように活動することで，タスク情報の上書きを防ぐことである．また，ネットワークの結合コストの最小化によって，モジュール構造が生み出されることが報告されている [Clune 13]．例えば，脳における結合コストは，軸索長やシナプス強度に相当する．これにより長距離の結合が嫌われ，局所的で密なネットワークとしてのモジュールが現れる．このことを利用して，Ellefsen et al. (2015) は，フィードフォワード型のニューラルネットワークの結合コストを遺伝的アルゴリズムによって最小化することによって，ネットワークにモジュール（タスクにより異なる処理経路）が現れ，複数タスクの継続学習が可能になることを示した．しかし，このモデルはフィードフォワード型であり，静的な情報の流れのみを扱っている．実際の脳においては，リカレント構造による動的な情報の流れの中で，モジュールの切り替えがなされていると考えられる．また，彼らはタスクとモジュールの対応を定量的に評価していない．

連絡先: 河合祐司，大阪大学先導的学際研究機構，〒 565-0871
大阪府吹田市山田丘 2-1，kawai@ams.eng.osaka-u.ac.jp

そこで本研究は，リカレントニューラルネットワークの一種であるエコーステートネットワーク (echo state network: ESN) に結合コスト最小化を導入するモデルを提案し，それが複数の時系列タスクにおける継続学習の破滅的忘却を回避できることを示す．一般的な ESN はランダム結合のリザバーネットワークに入力が与えられ，リザバーから出力への結合重みのみが誤差によって修正される．本モデルでは，リザバーネットワークの結合を，学習性能最大化と結合コスト最小化を目的関数とした遺伝的アルゴリズムによって最適化する．その結果，リザバー内にモジュールが現れ，タスクに応じてモジュール内とモジュール間の情報の流れが変化することで，出力重みの学習における破滅的忘却が回避されることが期待される．また，最適化されたリザバーノード間の移動エントロピーを計算し，タスクごとの情報の流れを解析することで，モジュールの機能分化を明らかにする．

2. モデル概要

2.1 エコーステートネットワークの継続学習

図 1 に提案モデルの概要を示す．以降では簡単のため，入出力が一次元である場合について説明する．時刻 t において，N 個のリザバーノードが入力 $u(t)$ をベクトル $\mathbf{W}_{\mathrm{in}}$ で重み付けられて受け取る．そして，リザバーノードの状態 $\mathbf{x}(t) = (x_1(t), \cdots, x_N(t))^\top$ は次式により更新される．

$$\mathbf{x}(t+1) = f(\mathbf{W}_{\mathrm{in}}u(t+1) + \mathbf{W}\mathbf{x}(t) + \mathbf{W}_{\mathrm{fb}}y(t)) \quad (1)$$

ここで，$\mathbf{W}$ は $N \times N$ のサイズのリザバー重み行列，$\mathbf{W}_{\mathrm{in}}$ は入力重みベクトル，$\mathbf{W}_{\mathrm{fb}}$ は出力 $y(t)$ のフィードバック重みベクトルであり，これらの重みは学習を通して不変である．今回，関数 f として，ハイパボリックタンジェントを用いた．$\mathbf{W}$ は行列 $\mathbf{W}_0$ を次式で正規化することで得られる．

$$\mathbf{W} = \alpha \frac{\mathbf{W}_0}{\rho(\mathbf{W}_0)} \quad (2)$$

ここで，$\rho(\mathbf{W}_0)$ は $\mathbf{W}_0$ の固有値の絶対値の最大値（スペクトル半径）であり，α は正規化定数である．Ellefsen et al. (2015) のモデルでは，タスクによって入力ノードが異なっていたが，本モデルでは，タスクが異なっても同じノードに入力が与えら

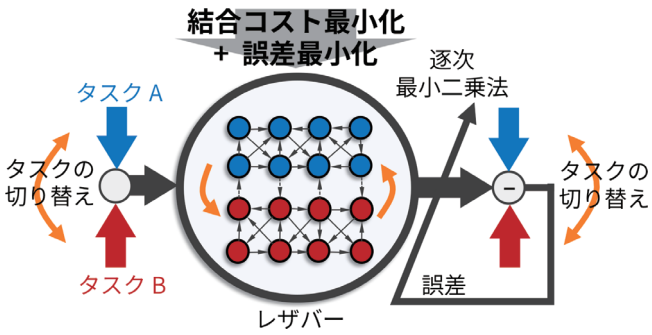


図 1: 提案モデルの概要図．一定時間経つとタスクが別のタスクに切り替わる状況におけるエコーステートネットワークの継続学習の問題を扱う．結合コスト最小化と誤差最小化を目的関数とする遺伝的最適化により，リザバーネットワークにモジュール性が現れ，タスクの切り替えに応じてそのモジュールの活動も切り替わることで，複数のタスクの学習が可能になる．

れる．明なタスク切り替え情報が与えられなくても，タスク間の入出力の性質の違いから，リザバーの活動ダイナミクスが変化することが期待される．

時刻 t において，出力 $y(t)$ を次式により得る．

$$y(t) = \mathbf{W}_{\mathrm{out}}(t)\mathbf{x}(t) \tag{3}$$

ここで，$\mathbf{W}_{\mathrm{out}}(t)$ は出力重みベクトルである．一般には，この出力重みを線形回帰のバッチ学習により求めるが，今回は，次に説明する逐次最小二乗法の継続学習により求める．

時刻 t における出力重みは次式により更新される．

$$\mathbf{W}_{\mathrm{out}}(t+1) = \mathbf{W}_{\mathrm{out}}(t) + \Delta\mathbf{W}(t) \tag{4}$$

$$\Delta\mathbf{W}(t) = \mathbf{R}(t)\mathbf{x}(t)e(t) \tag{5}$$

ここで，誤差 $e(t)$ は ESN の出力 $y(t)$ と教師信号 $y_{\mathrm{d}}(t)$ の差

$$e(t) = y(t) - y_{\mathrm{d}}(t) \tag{6}$$

である．$\mathbf{R}(t)$ は次式で与えられる．

$$\mathbf{R}(t) = \frac{1}{\mu}\left[\mathbf{R}(t-1) - \frac{\mathbf{R}(t-1)\mathbf{x}(t)\mathbf{x}^{\top}(t)\mathbf{R}(t-1)}{\mu + \mathbf{x}(t)\mathbf{R}(t-1)\mathbf{x}^{\top}(t)}\right] \tag{7}$$

ここで，μ は忘却係数とよばれるパラメータである．$\mathbf{R}(t)$ の初期値は $\mathbf{R}(0) = (1/\delta)\mathbf{I}$ で与えられ，$\mathbf{I}$ は単位行列，δ は非常に小さい定数である．

2.2 遺伝的アルゴリズムによるリザバー重みの最適化

一般の ESN では，$\mathbf{W}_0$ をランダムに決定するが，本モデルでは，$\mathbf{W}_0$ の個々の要素を遺伝子とする遺伝アルゴリズムにより最適化する．その目的関数には，学習性能の最大化（誤差の最小化）だけでなく，結合コストの最小化も含める．結合コストを定義するために，リザバーノードを平面格子状に配置する．二つのノード間の結合コストをその間のユークリッド距離と重みの絶対値の乗算で表し，全体の結合コストを，全てのノードの組み合わせについての結合コストの総和で定義する．この目的関数により，空間的に遠い結合を持つ ESN は淘汰されやすくなるため，近傍の密な結合で構成されたモジュールが創発することが期待される．今回，多目的の遺伝的アルゴリズムとして代表的な，非優越ソート遺伝的アルゴリズム (non-dominated sorting genetic algorithm II: NSGA-II) [Deb 02] を用いる．

遺伝的最適化と継続学習の処理の流れを以下に記す．

1. 初期世代として，異なる初期重みを持つ ESN を X 個構成する．
2. 複数のタスクが切り替わる状況での継続学習により，各 ESN が $\mathbf{W}_{\mathrm{out}}$ を更新する．
3. 2. での誤差の最小化と結合コストの最小化を目的関数とする NSGA-II により，$\mathbf{W}_0$ の異なる次世代の ESN が構成される．
4. 次の世代に移り，2. と 3. を Y 回繰り返す．

3. 実験

3.1 実験設定

今回，リザバーノード数 N を 36 に設定し，6×6 の平面格子上に配置した．$\mathbf{W}_{\mathrm{in}}$，$\mathbf{W}_{\mathrm{fb}}$，$\mathbf{W}_{\mathrm{out}}$ の学習初期値，および，$\mathbf{W}_0$ の進化初期値は $[-1, 1]$ の一様乱数にした．$\mathbf{W}$ を得るためのパラメータ α は 1.0 に設定した．

以下の性質の異なる二つのタスクを用いた．

- 正弦波の短期予測：正弦波を入力とし ($u(t) = \sin\left(\frac{\pi t}{10}\right)$)，1 時刻後の入力を教師信号とする ($y_{\mathrm{d}}(t) = u(t+1)$)．
- 3 ビットパリティチェック：0 か 1 のランダム入力に対して，現在と直近の過去二つの入力において，1 が偶数個あれば教師信号は 0，奇数個あれば 1 とする．

どちらのタスクも入出力は一次元である．16 ステップごとにタスクを入れ替え，480 ステップまで逐次最小二乗法による継続学習を行った．逐次最小二乗法の忘却係数 μ は 0.7 にし，初期設定に用いる δ は 0.1 に設定した．

NSGA-II において，一世代の個体数 X を 500，最大世代数 Y を 1000 に設定した．また，交叉率は 80%とし，80%の確率で一世代内に 1 個体の突然変異が起きるようにした．学習性能を 161 ステップから 480 ステップまでの平均二乗誤差で定義し，その最小化を NSGA-II の一つの目的関数とした．結合コストの最小化を目的関数に含めることの効果を明らかにするために，誤差最小化のみを目的関数とする比較条件を設けた．

リザバーネットワークにモジュールが創発し，そのモジュールがタスクに応じて異なる機能を有していることを明らかにするために，進化最適化後のリザバーネットワークの結合のモジュール性と神経活動の情報の流れを評価した．ネットワーク結合のモジュール性の評価には，Newman のモジュラリティ [Newman 06] を用いた．これはモジュール内ノード間の結合の割合から，結合がランダムにされた場合での結合の割合を減じた値である．この値が最も大きくなるモジュール数とノードへのモジュールの割り当てを採用する．神経活動の情報の流れの評価には，移動エントロピー [Schreiber 00] を用いた．今回，あるノードの 1 時刻分の活動から，1 時刻後の別のノードの活動への移動エントロピーを求めた．学習後の 10,000 データ点に対して，全てのノード間の移動エントロピーを計算し，モジュール間の平均移動エントロピーとモジュール内の平均移動エントロピーを求めた．さらに，タスク期間ごとに平均移動エントロピーを計算することで，タスク依存的な情報の流れを可視化した．

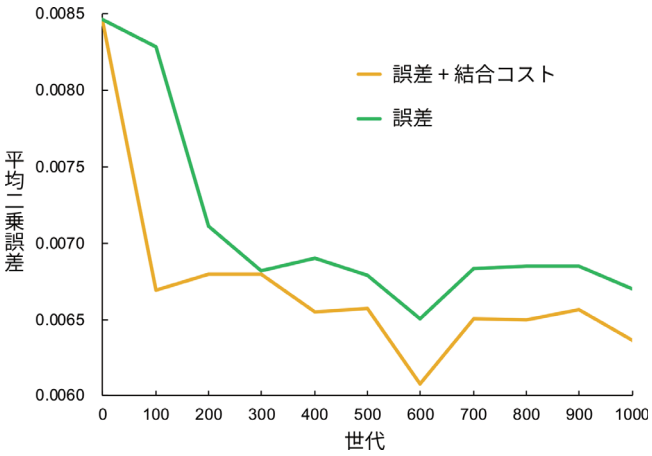


図 2: 各世代内で最小の平均二乗誤差.

3.2 実験結果

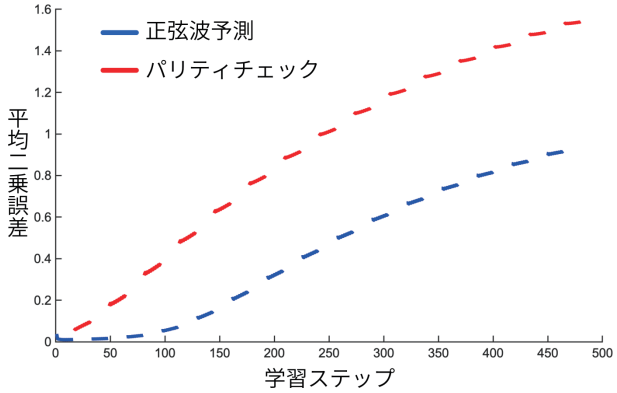
図 2 に，100 世代ごとの各世代における最小の平均二乗誤差を示す．黄色と緑の線がそれぞれ，誤差と結合コストの最小化，誤差のみの最小化の条件での結果である．この図から，結合コストを目的関数に含めることで，学習性能が向上していることがわかる．図 3 に結合コストを目的関数に入れた場合での (a) 初期世代と (b) 最終世代の学習曲線の例を示す．初期のランダムネットワークでは，タスクが切り替わるごとに再学習を繰り返していることがわかる．一方で，最終世代では，タスクが切り替わっても誤差が増大せず，破滅的忘却を回避している．

図 4 に，100 世代ごとのモジュラリティの値の推移を示す．正の結合と負の結合によりネットワークを分けて，それぞれのモジュラリティを解析した．図 4 中の実線から，結合コストを目的関数に含めることで，負の結合ネットワークのモジュラリティが増加したことがわかる．誤差のみを目的関数とした場合であってもモジュラリティの増加は認められるが，結合コストを含めた場合のものよりは大きくはない．一方，図 4 の正の結合のネットワーク（破線）ではどちらの場合であっても，モジュラリティの増加はほとんどなく，ランダム結合のモジュラリティと同等であることがわかる．

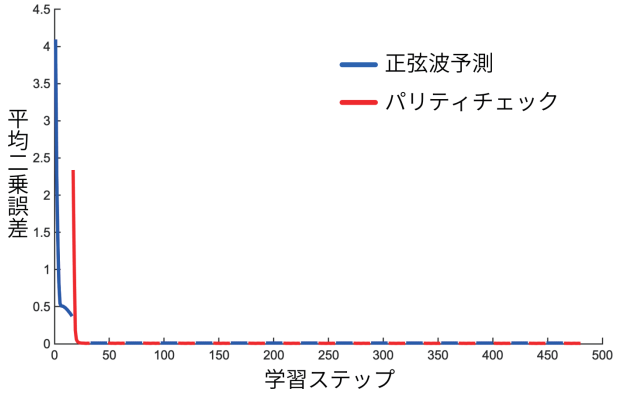
上記のモジュラリティの解析から，最終世代の正負の結合ネットワークのそれぞれにおいて，3 つのモジュールがあることがわかった．それらをモジュール A, B, C とよび，タスクごとのモジュール間とモジュール内の平均移動エントロピーを図 5 に示す（ただし，負の結合ネットワークのモジュールと正の結合ネットワークのモジュールは対応しない）．正弦波予測のときには，移動エントロピーの値が全体的に小さい．しかし，負のモジュール C から A への移動エントロピーが大きく，正のモジュール B への移動エントロピーが比較的大きいことがわかる．一方で，パリティチェックのときには，負のモジュール C 内の移動エントロピーが大きく，また，正のモジュール A と B での移動エントロピーが大きい．このように，タスクに応じて異なる情報の流れが発生し，モジュールごとに異なる情報処理がなされている可能性が示された．

4. 議論と結論

本研究では，ESN のリザバーに結合コスト最適化を導入することで，リザバーネットワークにモジュールが現れ，継続学



(a) 初期世代.



(b) 最終世代.

図 3: 学習曲線の例．16 ステップごとにタスクが切り替わる．青：正弦波予測タスク，赤：パリティチェックタスク.

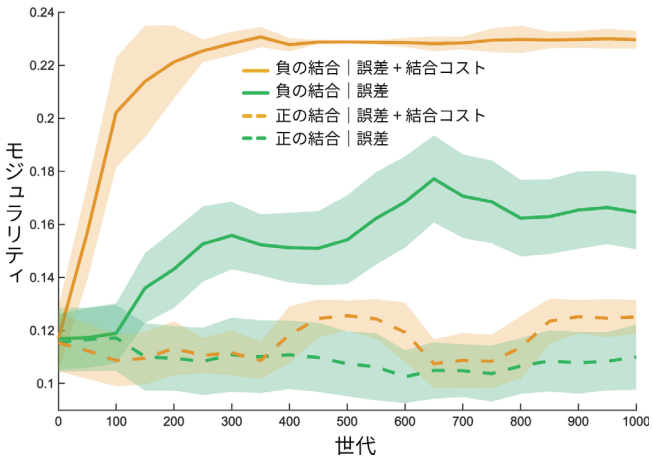
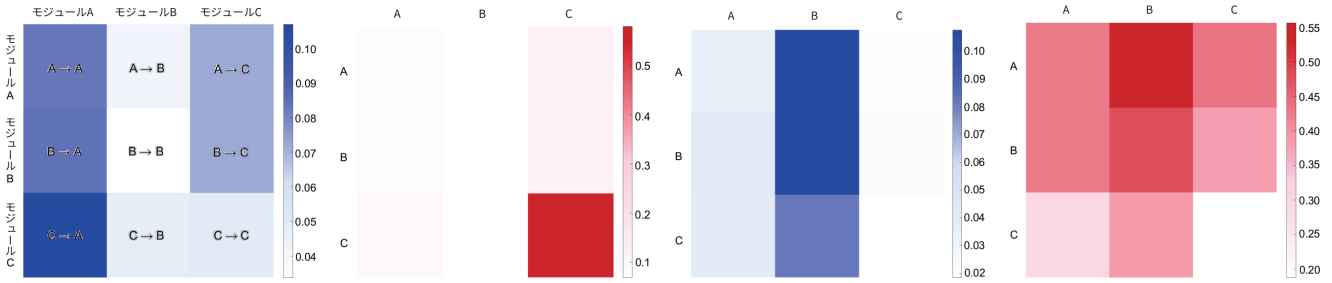


図 4: リザバーネットワークのモジュラリティ．結合の正負で分けてモジュラリティを求めた．実線と破線はそれぞれ，負と正の結合が作るネットワークのモジュラリティである．塗りつぶし範囲は，その世代における個体間の標準偏差を表す.



(a) 負のネットワークの正弦波予測　(b) 負のネットワークのパリティチェック　(c) 正のネットワークの正弦波予測　(d) 正のネットワークのパリティチェック

図 5: モジュール間とモジュール内の平均移動エントロピー．それぞれ 3 つのモジュールがあり，行のモジュールから列のモジュールへの平均移動エントロピーの大きさを色の濃さで表している．青：正弦波予測タスク，赤：パリティチェックタスク．

習における破滅的忘却が回避されることを示した．直感的には，結合コストによるネットワークの拘束は学習性能の改善に直接的には寄与せず，むしろ悪化を招く可能性も考えられる．にもかかわらず，複数タスクの学習において，結合コスト最小化を目的関数に含めることで学習性能の向上がみられたことは興味深い（図 2）．また，モジュール内とモジュール間の平均移動エントロピーの解析から，モジュールがタスクに応じて異なる情報の流れを作ることを明らかにした（図 5）．これによって，複数のタスクの学習が干渉することなく継続学習を可能にしたと考えられる．

モジュラリティの解析から，正（興奮性）の結合はモジュール性を持たず，負（抑制性）の結合はモジュラーネットワークであることを発見した（図 4）．それぞれのネットワークのはたらきを推測する．まず，興奮性のネットワークはおそらくランダムネットワークであり，複雑な神経活動を生成する役割があると考えられる．ランダム結合のリザバーの活動が多様なダイナミクスを含むことによって，ESN は様々な時系列課題の課題が可能になるとされる [Bertschinger 04]．そのため，一般的な学習性能の向上のために結合のランダム性が必要になる．一方，抑制性のモジュラーネットワークは，タスク特有のダイナミクスを生成する役割があると考えられる．ヒトの脳におけるモジュールの切り替えにおいても，不用なモジュールの活動が抑制されていることが報告されている [Fox 05]．本研究の結果は，ネットワークの活動が複雑であり，かつ構造を有することが，継続学習において重要であることを示唆する．このことは，複数タスク学習におけるニューラルネットワークの構成論だけでなく，柔軟で適応的な脳のメカニズムの理解につながると期待される．本モデルの他のタスクへの適用や，抑制性のモジュラーネットワークの機能のさらなる調査と体系化が今後の課題である．

謝辞

本研究は，JST，CREST，JPMJCR17A4 の支援を受けたものである．

参考文献

[Bertschinger 04] Bertschinger, N. and Natschläger, T.: Real-time computation at the edge of chaos in recurrent neural networks, *Neural Comput.*, Vol. 16, No. 7, pp. 1413–1436 (2004)

[Bongard 11] Bongard, J. C.: Spontaneous evolution of structural modularity in robot neural network controllers, in *Proc. of the 13th GECCO*, pp. 251–258 (2011)

[Clune 13] Clune, J., Mouret, J.-B., and Lipson, H.: The evolutionary origins of modularity, *Proc. R. Soc. B*, Vol. 280, No. 1755, 20122863 (2013)

[Deb 02] Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Trans. Evol. Comput.*, Vol. 6, No. 2, pp. 182–197 (2002)

[Ellefsen 15] Ellefsen, K. O., Mouret, J.-B., and Clune, J.: Neural modularity helps organisms evolve to learn new skills without forgetting old skills, *PLoS Comput. Biol.*, Vol. 11, No. 4, e1004128 (2015)

[Fox 05] Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., and Raichle, M. E.: The human brain is intrinsically organized into dynamic, anticorrelated functional networks, *Proc. Natl. Acad. Sci.*, Vol. 102, No. 27, pp. 9673–9678 (2005)

[McCloskey 89] McCloskey, M. and Cohen, N. J.: Catastrophic interference in connectionist networks: The sequential learning problem, *Psychol. of Learn. Motiv.*, Vol. 24, pp. 109–165 (1989)

[Meunier 09] Meunier, D., Lambiotte, R., Fornito, A., Ersche, K., and Bullmore, E. T.: Hierarchical modularity in human brain functional networks, *Front. in Neuroinform.*, Vol. 3, No. 37 (2009)

[Newman 06] Newman, M. E. J.: Modularity and community structure in networks, *Proc. Natl. Acad. Sci.*, Vol. 103, No. 23, pp. 8577–8582 (2006)

[Parisi 18] Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., and Wermter, S.: Continual lifelong learning with neural networks: a review, *arXiv:1802.07569* (2018)

[Schreiber 00] Schreiber, T.: Measuring information transfer, *Phys. Rev. Lett.*, Vol. 85, No. 2, pp. 461–464 (2000)

4:30 PM - 5:10 PM (Thu. Jun 6, 2019 3:50 PM - 5:10 PM Room D)

[3D05-06-3] Special program

Special program

©The Japanese Society for Artificial Intelligence

Organized Session | Organized Session | [OS] OS-12

[3E3-OS-12a] 画像と AI（ MIRU2019プレビュー）(1)

長原 一（大阪大学）、川崎 洋（九州大学）、岡部 孝弘（九州工業大学）

Thu. Jun 6, 2019 1:50 PM - 3:10 PM Room E (301A Medium meeting room)

[3E3-OS-12a-01] Convolutional Neural Network for Image Recognition and Visual Explanation

〇Takayoshi Yamashita[1] （1. Chubu University）

1:50 PM - 2:30 PM

[3E3-OS-12a-02] Adaptive selection of auxiliary tasks in UNREAL

〇Hidenori Itaya[1], Tsubasa Hirakawa[1], Yamashita Takayoshi[1], Fujiyoshi Hironobu[1] （1. Chubu University）

2:30 PM - 2:50 PM

[3E3-OS-12a-03] ShakeDrop Regularization for ResNet Family

〇Yoshihiro Yamada[1], Masakazu Iwamura[1], Koichi Kise[1] （1. Osaka Prefecture University）

2:50 PM - 3:10 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Convolutional Neural Networkによる画像認識と視覚的説明

Convolutional Neural Network for Image Recognition and Visual Explanation

山下隆義 *1
Takayoshi Yamashita

*1中部大学
Chubu university

Deep learning technologies in the field of computer vision are gradually introducing into our daily life. These technologies have been achieved by introducing the methods that improve recognition performance, such as deeper models, a method of stability training of the deeper model. Moreover, to achieve the productization, the visual explanation that explains the decision making of deep learning to a user has been proposed. In this paper, we present a trend of deep learning technologies, which have been used on image recognition methods such as image classification, object detection, and visual explanation.

1. はじめに

画像中に存在する物体のカテゴリや位置等を認識する画像認識技術は，Advanced Driver Assistance System (ADAS) や防犯システム，SNS やスマートフォンのアプリケーション等で幅広く応用されている．ADAS では自動ブレーキシステム等の歩行者や自動車を検出する処理で用いられ，SNS やアプリケーション等のレジャーな分野では顔の形状を画像等から推定して自動で編集するアプリケーション等へ応用されている．

このような，画像認識技術が身近に取り入れられるようになった要因の一つとして，深層学習の発展により認識性能が向上したことが深く関係している．画像認識分野における深層学習では，Deep Convolutional Neural Network (CNN) [Alex 12] をベースにした手法が一般的に用いられている．CNN は，複数のカーネルで構築される畳み込み層を主体に構築されたニューラルネットワークであり，学習によりカーネルを更新することで画像認識に有効な特徴量を獲得する．CNN が画像分類で高い性能を発揮した後，物体検出やセマンティックセグメンテーション等の様々な画像認識タスクへ応用されるようになった．これは，CNN のモデルの発展や深いモデルを安定して学習できる手法が数多く提案されたことが大きな貢献となっている．

一方で，CNN による性能向上のみでなく，推論時における CNN の判断根拠をユーザへ伝える技術も提案されている．CNN を導入した製品が誤認識によりユーザへ何かしらの危害を与えたとき，ユーザへなぜこのような行動を起こしたのかを説明する必要がある．そのため，これらの判断根拠を解析する手法は CNN を用いた製品を商品化する際に重要な技術となる．CNN の判断根拠を解析する研究は活発に取り組まれており，様々なアプローチが提案されている．特に，画像認識分野では CNN が推論時に注視した領域をマップで表現した Attention map を用いることで，判断根拠を解析する視覚的説明が用いられる．視覚的説明は，CNN が画像認識する際に注視した領域を可視化することができるため，直感的に CNN の判断根拠を知ることができる．

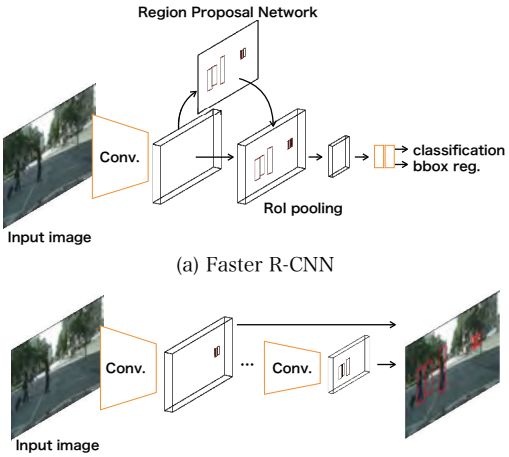
本稿では，画像認識分野で用いられる CNN の最近の動向について述べる．まず，画像分類において用いられる CNN のモデルについて述べた後，物体検出のモデルについて述べる．そして，CNN の学習時に用いられる正規化や学習方法について述べる．最後に，CNN の判断根拠を解析する視覚的説明の研究事例について述べる．

連絡先: 山下隆義，中部大学，愛知県春日井市松本町 1200，0568-51-1111，0568-51-1111，takayoshi@isc.chubu.ac.jp

2. 画像分類における CNN のモデル

物体検出法やセグメンテーション，属性認識等で用いられる CNN ベースのモデルは，画像分類で提案されたモデルをベースにネットワークを構築している例が多い [Ren 15, Liu 16, Badrinarayanan 15, He 17]．画像分類における CNN の初代の代表的なモデルとして，AlexNet [Alex 12] がある．AlexNet は，大規模な画像認識コンペティションである ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [Russakovsky 15] でトップとなったモデルである．AlexNet の構造は，5 層の畳み込み層と 2 層の全結合層から構築された CNN であり，Local Response Normalization (LRN) や Dropout，Rectified Linear Units (ReLU) 等の大量のデータで深いネットワークを学習するためのテクニックが導入されている．AlexNet の提案後，より深いネットワークを構築することで，高精度な画像分類を実現するアプローチが取られた．深いネットワークは，単純な特徴と複雑な特徴を同時に学習できるため，より高い精度で認識が可能である．初期におけるこのアプローチの代表的なモデルとして，VGGNet [Karen 15] や GoogLeNet [Szegedy 15] が挙げられる．VGGNet は最大で 19 層の CNN を構築し，GoogLeNet は Inception module を導入することで 22 層の CNN を構築している．その後，Batch Nornalization [Sergey 15] や He の正規化 [He 15] 等，深いネットワークの学習を安定させるためのアプローチが提案された．

2016 年には，He らが Residual Learning を導入した ResNet を提案し，100 層以上の CNN を構築した [He 16]．ResNet は複数の層を繋げる際にバイパス構造を取り入れており，Residual Learning を取り入れた Residual unit を構築している．Residual Learning の導入により，大量の層で構築されたネットワークを安定して学習できるようになり，CNN ベースの画像分類の精度がさらに向上した．ResNet が提案された後は，ResNet をベースにした様々なモデルが提案されている [Zagoruyko 16, Huang 17, Xie 17]．



(a) Faster R-CNN

(b) Single Shot Multi-box Detector

図 1: 物体検出の代表的なモデル

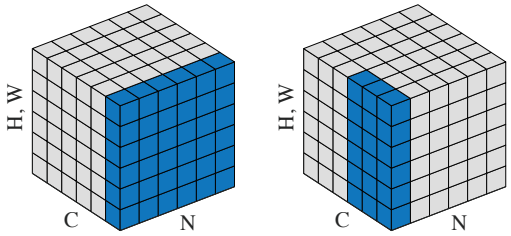
これらの画像分類モデルの進化は，物体検出や属性認識，セグメンテーションにも大きな影響を与えている．物体検出やセグメンテーションで用いられるネットワークは画像分類のネットワークモデルをベースに構築するため，画像分類の性能向上に伴い他の認識タスクの性能も向上できる．そのため，画像分類のモデルは他の画像認識タスクの観点からも，重要な立ち位置に属している．

3. 物体検出

CNN ベースの画像分類法の発展に伴い，CNN をベースとした物体検出法も大きく発展している．物体検出は，図 **??**(b) のように画像中の物体の位置とそのカテゴリを推定する技術である．CNN ベースの物体検出法は，R-CNN [Girshick 14] と Fast R-CNN [Girshick 15] をベースに進化を遂げている．R-CNN ベースの手法では，物体の候補領域を検出し，検出した候補領域を CNN でカテゴリ分類とバウンディングボックスの修正を行う，2 段階の検出構造を採用している．しかし，R-CNN と Fast R-CNN は計算コストが高く，リアルタイムな物体検出が困難である．リアルタイムで物体検出が可能な手法として，Faster R-CNN [Ren 15] がある．Faster R-CNN は，Region Proposal Network (RPN) を図 1(a) のように導入しており，物体候補領域の検出と，カテゴリ分類とバウンディングボックスの修正を 1 つのネットワークで End-to-End に処理できる．Faster R-CNN は，インスタンスセグメンテーションへも応用されている．Faster R-CNN をインスタンスセグメンテーションへ応用した Mask R-CNN [He 17] は，検出した物体領域のマスクを出力するブランチを追加することで，高精度なインスタンスセグメンテーションを実現している．また，R-CNN のような 2 段階の検出構造を用いずに，図 1(b) のようにネットワークの応答値から直接検出スコアを出力できる Single Shot Multi-box Detector (SSD) [Liu 16] も提案されている．SSD は，物体候補領域の検出を必要としないため，Faster R-CNN より高速に物体を検出できる．

4. CNN の汎化性能を向上させる手法

大量のデータを用いて深いネットワークを安定して学習するために，様々なテクニックが導入されている．ネットワークの学習時に正規化，最適化方法を導入することで，学習の収束を早めたり，学習時における勾配の発散等を防ぐことができる．



(a) Batch Normalization (b) Group Normalization

図 2: Batch Normalization の発展系の手法 (文献 [Yuxin 18] から引用)

4.1 正規化

CNN はカーネルと入力画像または特徴マップの局所領域から内積値を求めるため，入力値にノイズが発生した場合等に応答値のばらつきが発生し，認識性能を低下させる原因となる．そのため，一般的には畳み込み層や全結合層の応答値を正規化し，精度低下を抑制する．AlexNet では LRN が用いられていたが，他の特定のネットワークに対しては精度が向上しなかったり，Batch Normalization [Sergey 15] が提案されたことで，現在は一般的に用いられていない．

Batch Normalization は，特定のチャンネルをミニバッチ単位で正規化し，平均を 0 と分散を 1 にする．図 2(a) のようにミニバッチ単位で特定のチャンネルを正規化することで，内部共変量シフトが大幅に変動するのを防いでいる．しかし，Batch Normalization は内部共変量シフトを獲得するために，ミニバッチのサイズを 16 以上にする必要がある．ミニバッチのサイズは大きいほど計算コストが増加するため，物体検出等の使用メモリ量が膨大なモデルで十分な性能を発揮できない．この問題を解決する手法として，Group Normalization がある．Group Normalization は，図 2(b) のように数枚のチャンネルのみ用いて正規化する．これにより，Group Normalization は少量のミニバッチサイズでも従来の Batch Normalization と同等の精度を得ることができる．

4.2 学習方法

確率的勾配降下法 (stochastic gradient descent; SGD) によりネットワークを学習する場合，学習率 η を学習の過程で変化させる Learning rate schedule が一般的に導入されている．ネットワークの学習における学習率は，パラメータの更新量を制御する係数である．学習率を大きく設定した場合は学習の収束が早くなるが，勾配が発散しやすい．一方で，学習率を小さく設定した場合は最適解を獲得しやすいが，学習の収束が遅くなる．この問題を解決するために，学習の過程で学習率を変更することで，最適解を獲得しやすくしている．Learning rate schedule は，一般的には学習率を減衰させる Learning rate drop が一般的に用いられる．SGD で Learning rate drop を用いる場合，指定した更新回数に達した際に学習率を減衰する．学習率を特定の更新回数で下げることで，算出される学習誤差をより下げることができ，認識率を向上できる．

一方で，学習率を下げるタイミングを手動で決定するのではなく，学習の過程で自動に決定する方法も提案されている [Zeiler 12, Tieleman 12]．Adaptive Gradient (AdaGrad) [Duchi 11] は，ネットワークの各パラメータに対し

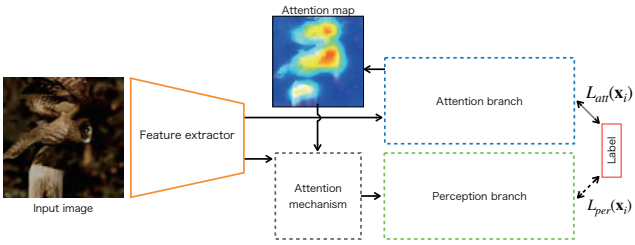


図 3: Attention Branch Network の構造

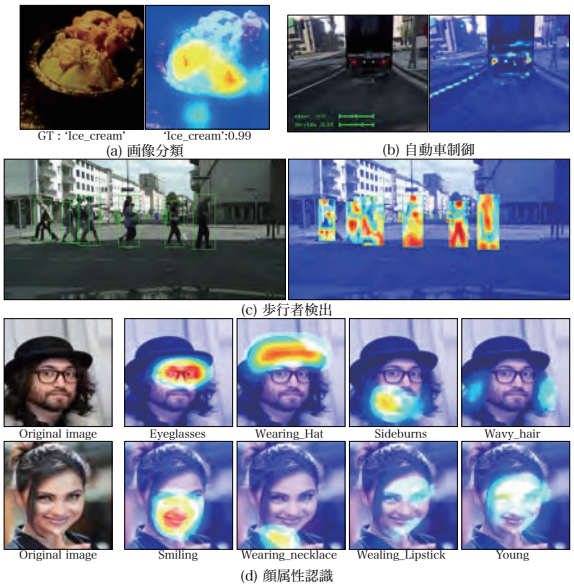


図 4: ABN が獲得した Attention map の例

て学習率を設計し，自動で調整しながら学習できる．AdaGrad をベースとした学習法は，数多く提案される [Zeiler 12, Kingma 14].

5. 視覚的説明

深層学習における画像認識分野では，推論時に注視した領域をマップで表現した Attention map から判断根拠を解析する視覚的説明の研究が取り組まれている [Zeiler 14, Zhou 16, Ramprasaath 17]．Attention map の獲得には，勾配を用いた Bottom-up の手法とネットワークの応用値を用いる Top-down の手法の 2 種類がある．Bottom-up の手法の例として，Guided Backpropagation と Gradient-weighted Class Activation Mapping (Grad-CAM) [Daniel 17] がある．Guided Backpropagation と Grad-CAM は，逆伝播の特定のクラスにおける正値の勾配のみ用いることで，Attention map を獲得する．Guided Backpropagation と Grad-CAM は特定のクラスにおける Attention map を様々なプレトレーニングモデルから獲得できるため，CNN の解析手法として広く一般的に用いられている．Grad-CAM は，特定のクラスの出力層のユニットから勾配を発生させ，特徴マップを獲得する．そして，順伝播時における最後の畳み込み層の特徴マップに対して，GAP を施す．GAP により獲得した特徴ベクトルは重みとして使用し，勾配ベースの特徴マップに対して重み付き和を求める．この重み付き和で求めた特徴マップに ReLU を施すことで，特定のクラスに対する注視領域を獲得することができる．

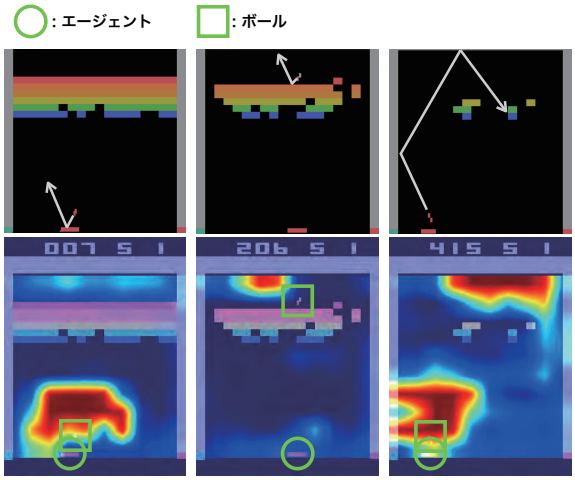


図 5: Breakout における Attention map の例

視覚的説明における Top-down の手法は，ネットワークが出力した応答値を用いることで Attention map を獲得できる．Top-down の手法は Attention map を獲得するためにネットワークを再構築して再学習する必要があるが，順伝播の過程で各クラスにおける注視領域を獲得できる．Top-down の手法の代表的な手法である CAM [Zhou 16] は，畳み込み層の応答値と全結合層の結合重みを用いることで，各クラスにおける Attention map を獲得できる．CAM は全結合層を畳み込み層に入れ替える等の処理が必要なため，画像分類においては性能低下を引き起こしやすい．

この問題を解決した Top-down な方法として，Attention Branch Network (ABN) [Fukui 18] がある．ABN は，図 3 のように Feature extractor と Attention branch，Perception branch の 3 つのモジュールから構成されている．Feature extractor は，入力画像から特徴マップを抽出するモジュールである．Attention branch は，畳み込み層をベースに構築されたブランチであり，Attention map を出力する．Feature extractor は，入力画像から特徴マップを獲得するモジュールである．抽出した特徴マップは Attention branch へ入力され，Attention map を出力する．ABN は，画像分類をはじめとした歩行者検出，マルチタスク学習等の様々な画像認識タスクから，図 4 のように Attention map を獲得できる．また，ABN はエージェントの制御で用いられる深層強化学習へも応用できる．ABN を深層強化学習へ応用した際に獲得した Attention map を，図 5 に示す．図 5 の例は，Atari ゲームの一つである Breakout を深層強化学習で操作している例である．図 5 の結果から，ボールを跳ね返す，ブロックの奥でボールが跳ね返る等のゲームスコアを獲得する直前のシーンにおいて，Attention map が強く反応していることがわかる．

6. おわりに

本稿では，画像認識における深層学習の動向についてまとめ，紹介した．画像認識における深層学習の発展により，画像分類をはじめとした物体検出等の性能を大幅に向上させた．また，性能向上だけでなく，深層学習の推論結果に対する判断根拠を解析する研究も活発に取り組まれている．

参考文献

[Alex 12] Alex, K., Sutskever, I., and Hinton, G. E.: ImageNet Classification with Deep Convolutional Neural Networks, in *Neural Information Processing Systems*, pp. 1097–1105 (2012)

[Badrinarayanan 15] Badrinarayanan, V., Handa, A., and Cipolla, R.: SegNet: A Deep Convolutional Encoder-Decoder Architecture for Robust Semantic Pixel-Wise Labelling, *arXiv preprint arXiv:1505.07293* (2015)

[Daniel 17] Daniel, S., Nikhil, T., Been, K., Fernanda, B. V., and Martin, W.: SmoothGrad: removing noise by adding noise (2017)

[Duchi 11] Duchi, J., Hazan, E., and Singer, Y.: Adaptive Subgradient Methods for Online Learning and Stochastic Optimization, *Journal of Machine Learning Research*, Vol. 12, pp. 2121–2159 (2011)

[Fukui 18] Fukui, H., Hirakawa, T., Yamashita, T., and Fujiyoshi, H.: Attention Branch Network: Learning of Attention Mechanism for Visual Explanation, *arXiv preprint arXiv:1812.10025* (2018)

[Girshick 14] Girshick, R., Donahue, J., Darrell, T., and Malik, J.: Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation, in *Computer Vision and Pattern Recognition*, pp. 580–587 (2014)

[Girshick 15] Girshick, R.: Fast R-CNN, in *International Conference on Computer Vision* (2015)

[He 15] He, K., Zhang, X., Ren, S., and Sun, J.: Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, in *International Conference on Computer Vision*, pp. 1026–1034 (2015)

[He 16] He, K., Zhang, X., Ren, S., and Sun, J.: Deep Residual Learning for Image Recognition, *Computer Vision and Pattern Recognition*, pp. 770–778 (2016)

[He 17] He, K., Gkioxari, G., Dollár, P., and Girshick, R.: Mask R-CNN, in *International Conference on Computer Vision* (2017)

[Huang 17] Huang, G., Liu, Z., Maaten, van der L., and Weinberger, K. Q.: Densely connected convolutional networks, in *Conference on Computer Vision and Pattern Recognition* (2017)

[Karen 15] Karen, S. and Andrew, Z.: Very Deep Convolutional Networks for Large-Scale Image Recognition, *International Conference on Learning Representations* (2015)

[Kingma 14] Kingma, D. P. and Ba, J.: Adam: A method for stochastic optimization, *1412.6980* (2014)

[Liu 16] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-y., and Berg, A. C.: SSD : Single Shot MultiBox Detector, in *European Conference on Computer Vision*, pp. 1–15 (2016)

[Ramprasaath 17] Ramprasaath, S., R., Michael, C., Abhishek, D., Ramakrishna, V., Devi, P., and Dhruv, B.: Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization, in *International Conference on Computer Vision*, pp. 618–626 (2017)

[Ren 15] Ren, S., He, K., Girshick, R., and Sun, J.: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, in *Neural Information Processing Systems*, pp. 91–99 (2015)

[Russakovsky 15] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., and Fei-Fei, L.: ImageNet Large Scale Visual Recognition Challenge, *International Journal of Computer Vision*, Vol. 115, No. 3, pp. 211–252 (2015)

[Sergey 15] Sergey, I. and Christian, S.: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, in *International Conference on Machine Learning*, pp. 448–456 (2015)

[Szegedy 15] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A.: Going deeper with convolutions, in *Computer Vision and Pattern Recognition*, pp. 1–9 (2015)

[Tieleman 12] Tieleman, T. and Hinton, G.: Lecture 6.5—RmsProp: Divide the gradient by a running average of its recent magnitude, COURSERA: Neural Networks for Machine Learning (2012)

[Xie 17] Xie, S., Girshick, R. B., Dollár, P., Tu, Z., and He, K.: Aggregated Residual Transformations for Deep Neural Networks, *Computer Vision and Pattern Recognition*, pp. 5987–5995 (2017)

[Yuxin 18] Yuxin, W. and Kaiming, H.: Group Normalization, in *European Conference on Computer Vision*, pp. 3–19 (2018)

[Zagoruyko 16] Zagoruyko, S. and Komodakis, N.: Wide Residual Networks, in *British Machine Vision Conference* (2016)

[Zeiler 12] Zeiler, M. D.: ADADELTA: An Adaptive Learning Rate Method, *1212.5701* (2012)

[Zeiler 14] Zeiler, M. D. and Fergus, R.: Visualizing and Understanding Convolutional Networks, in *European Conference on Computer Vision*, pp. 818–833 (2014)

[Zhou 16] Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., and Torralba, A.: Learning Deep Features for Discriminative Localization, *Computer Vision and Pattern Recognition* (2016)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

UNREALにおける補助タスクの適応的選択

Adaptive selection of auxiliary tasks in UNREAL

板谷 英典 *1 Hidenori Itaya　平川 翼 *1 Tsubasa Hirakawa　山下 隆義 *1 Takayoshi Yamashita　藤吉 弘亘 *1 Hironobu Fujiyoshi

*1中部大学
Chubu University

Deep reinforcement learning has a difficulty to solve a complex problem because such problem consists of a larger state space. To solve this problem, Unsupervised Reinforcement learning and Auxiliary Learning (UNREAL) has been proposed, which uses several auxiliary tasks during training. However, all auxiliary tasks might not perform well on each problem. Although we need to carefully design these tasks for solving this problem, it requires significant cost. In this paper, we propose an additional auxiliary task, called auxiliary selection. The proposed method can adaptively select auxiliary tasks that contributes the performance improvement. Experimental results with DeepMind Lab demonstrate that the proposed method can select appropriate auxiliary tasks with respect to each game tasks and efficiently train a network.

1. はじめに

強化学習とは，数値化された報酬を最大とするために，何をすべきかを学習する問題である．また，教師あり学習のように，どのような行動を選択すれば良いかは教えられず，どの行動を選択すればより良い結果に結び付くかを見つけ出す問題となっている．

強化学習は，教師あり学習のように教師信号を用意する必要がない点などから，ロボット制御 [Gu 17] [Rajeswaran 17] やゲーム攻略 [Justesen 17] [Firoiu 17] などの様々なタスクに応用されている．ゲーム攻略については，Silver ら [Silver 16] のコンピュータ囲碁プログラム AlphaGo がプロ囲碁棋士に勝利し，非常に注目された．また，Atari2600 のゲーム攻略において，深層強化学習手法の一つである Deep Q-Network (DQN) [Mnih 15] と呼ばれる手法が提案され，人間を凌駕するスコアを達成した．DQN は Q 学習 [Watkins 92] と Deep Convolutional Neural Network (DCNN) を組み合わせた手法であり，画像を入力とする Atari2600 のゲームのように状態数の多い問題を扱うことを可能にしている．この DQN 以降，強化学習手法は深層学習を組み合わせた深層強化学習が主流となった．

強化学習における学習データは，エージェントが環境を探索し収集する．そのため，学習に寄与するデータを獲得するために時間を要するという問題がある．そこで，この問題を解決するために Asynchronous Advantage Actor-Critic (A3C) [Mnih 16] が提案されている．A3C は，学習で用いる経験の生成を並列に実行することで高速化し，パラメータの更新を非同期的に行う手法である．また，A3C をベースとし，教師なし学習の補助タスクをメインタスクと並列に実行する UNspervised REinfrocsment learning and Auxiliary Learning (UNREAL) [Jaderberg 16] が提案されている．UNREAL は，複数の補助タスクを導入することによって，ゲームタスクにおいて A3C より高いスコアを達成している．しかし，UNERAL で用いられる全ての補助タスクは，あらゆる環境において必ずしも有効であるとは限らない．また，補助タスクを用いることでメインタスクの学習を妨げるという問題が存在する．そのため，補助タスクは環境に応じて適切に設計する必要があるが，適切な補助タスクの設計は多大な手間と時間を要する．

そこで本研究では，UNREAL の補助タスクに着目し，補助タスクを環境に合わせ適応的に選択するタスク Auxiliaty Selection を導入することで，上記の問題を解決する．最適な補助タスクの選択には，Auxiliaty Selection により出力された各補助タスクの重みと，各補助タスクとの損失関数の積を取ることで実現する．DeepMind Lab [Beattie 16] の 3 つのゲームを用いて，UNREAL および各補助タスクのみの場合とスコアを比較することで，本手法の有効性を示す．また，各補助タスクの選択回数を調査することで，最適な補助タスクを選択できているか確認する．

2. 関連研究

メインタスクと補助タスクを並列に学習させることでメインタスクの高精度化を図る手法は様々提案されている．Liebel ら [Liebel 18] は，自動車の運転シーンにおいて，セマンティックセグメンテーションと深度推定をメインタスクとし，時刻と天候推定の補助タスクを並列に実行することでメインタスクの精度向上を実現している．Jaderberg ら [Jaderberg 16] は，深層強化学習において，ベースである A3C [Mnih 16] に加えて，教師なし学習の補助タスクをメインタスクと並列に実行する手法を提案している．この手法は，3 つの異なる補助タスクを用いることで，DeepMind Lab の迷路攻略タスクにおいて，高いスコアを獲得している．1 つ目の補助タスク Pixel Control は画像の画素が大きく変化する行動を学習するタスクである．2 つ目の補助タスク Value Function Replay は過去の経験をシャッフルし，状態価値関数 $V(s)$ を学習するタスクである．3 つ目の補助タスク Reward Prediction は報酬を獲得した経験を優先して学習し，未来の報酬を予測するタスクである．

しかし，上記の手法では，メインタスクに適していない補助タスクを用いた場合，メインタスクの学習を妨げるという問題が存在する．そのため，補助タスクの導入にはメインタスクに

連絡先:
板谷 英典：itaya@mprg.cs.chubu.ac.jp
平川 翼：hirakawa@mprg.cs.chubu.ac.jp
山下 隆義：yamashita@cs.chubu.ac.jp
藤吉 弘亘：hf@cs.chubu.ac.jp

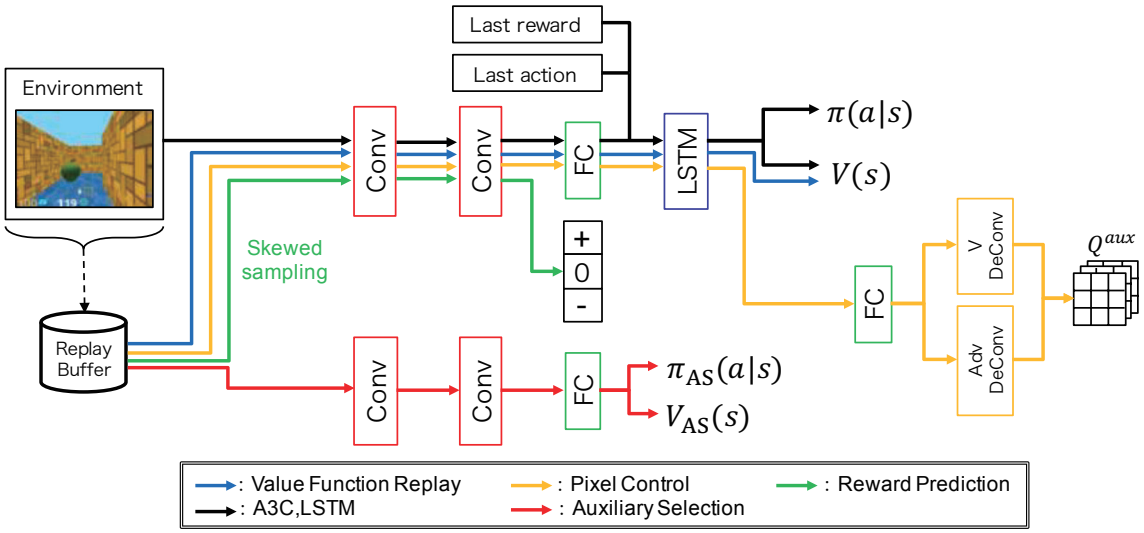


図 1: 提案手法のネットワーク構成

適したタスクを導入する必要がある．そこで，深層強化学習において，上記の問題を解決する手法がいくつか提案されている．Teh ら [Teh 17] は，蒸留によりタスク間に共通する行動を捉えた方策を獲得し，学習の妨げを回避することで，学習を安定させる手法を提案した．この手法は，共有する方策が全タスクにおいて有効な方策から離れないような制約を導入することで，異なるタスク間での頑健性と安定性を獲得している．一方，提案手法では，メインタスクの学習に有効な補助タスクを適応的に選択することで，メインタスクの学習を妨げることなく，学習の効率化を行う．

Riedmiller ら [Riedmiller 18] は，どの補助タスクの方策を使用すれば，メインタスクを解決できるか学習する Scheduled Auxiliary Control (SAC-X) を提案している．この手法は，各補助タスクを低レベルのタスクとして設計し，それぞれの目的に沿った方策を学習する．そして，SAC により用いる方策を選択することで，ロボットアームのような報酬が疎である問題を解決している．一方，提案手法における補助タスクは方策を必要としないため，UNREAL における Value Function Replay などが使用可能である．また，メインタスクである A3C の学習に対して，適応的に補助タスクを選択するため，補助タスクを環境に左右されず設計が可能である．

3. 提案手法

UNREAL の補助タスクは，環境によって有効性が異なるため，メインタスクの学習を妨げるという問題がある．そのため，環境に合わせた補助タスクの選択が求められる．本研究では，環境に合わせて用いる補助タスクを適応的に選択するタスク Auxiliary Selection を提案する．

3.1 Auxiliary Selection

図 1 に Auxiliary Selection を導入した UNREAL のネットワーク構成を示す．UNREAL の 3 つの補助タスクは，Pixel Control (PC)，Value Function Replay (VR)，Reward Prediction (RP) である．Auxiliary Selection には，Replay Buffer 内に格納された画像を入力し，状態価値関数 $V_{\mathrm{AS}}(s)$ と方策 π_{AS} を出力する．方策 π_{AS} は各補助タスクを用いるかどうかを表す値である．各補助タスクに対する重みを $C_{\mathrm{PC}} = \{0, 1\}$，$C_{\mathrm{VR}} = \{0, 1\}$，$C_{\mathrm{RP}} = \{0, 1\}$ とするとき，$\pi_{\mathrm{AS}} = (C_{\mathrm{PC}}, C_{\mathrm{VR}}, C_{\mathrm{RP}})$ となる．Auxiliary Selection のネットワークは，畳み込み層 2 層と全結合層 1 層から構成される．また，他の補助タスクとは異なり，A3C のネットワークとは共有せず，独立したネットワークとして学習を行う．このように，環境に合わせて補助タスクを適応的に選択することで，補助タスクを設計する際の効率化を図る．

3.2 損失関数

提案手法の損失関数 L_{proposed} は，従来の UNREAL の損失関数をもとに設計し，式 (1) のように定義する．

$$L_{\mathrm{proposed}} = L_{\mathrm{A3C}} + C_{\mathrm{VR}} L_{\mathrm{VR}} + C_{\mathrm{PC}} \sum_{c} L_{Q}^{(c)} + C_{\mathrm{RP}} L_{\mathrm{RP}} \tag{1}$$

提案手法では，Auxiliary Selection から獲得する C_{VR}，C_{PC}，C_{RP} と各補助タスクの損失関数の積を取ることで，最適な補助タスクのみを用いた学習を実現する．

また，提案手法の損失関数 L_{propoped} には，Auxiliary Selection から獲得する $C_{\mathrm{VR}}, C_{\mathrm{PC}}, C_{\mathrm{RP}}$ を用いている．そのため，Auxiliary Selection の学習を他の補助タスクと同様に，L_{proposed} に基づいて行うと，各補助タスクの重み $C_{\mathrm{VR}}, C_{\mathrm{PC}}, C_{\mathrm{RP}}$ が 0 になるように学習されるという問題がある．したがって，Auxiliary Selection の学習では，L_{proposed} とは異なる Auxiliary Selection の損失関数を定義し，A3C と各補助タスクのネットワークとは独立して学習を行う．

Auxiliary Selection の損失関数は状態価値関数と方策の損失関数で表すことができる．状態価値関数の損失関数 $L_{\mathrm{AS}v}$ を式 (2)，方策の損失関数 $L_{\mathrm{AS}p}$ を式 (3) に示す．ここで，θ^{-} は更新前のネットワークのパラメータである．また，$H(\pi_{\mathrm{AS}})$ は局所的な最適解に収束しないように，探索を促進するためのエントロピーであり，β はエントロピーの正則化項の強さを制御するパラメータである．

$$L_{\mathrm{AS}v} = (r + \gamma V_{\mathrm{AS}}(s_{t+1}, \theta^{-}) - V_{\mathrm{AS}}(s_t, \theta))^2 \tag{2}$$

$$L_{\mathrm{AS}p} = -\log(\pi_{\mathrm{AS}}(a|s)) A(s, a) - \beta H(\pi_{\mathrm{AS}}) \tag{3}$$

Auxiliary Selection の損失関数は，式 (2) の状態価値関数の損失関数と式 (3) の方策の損失関数の和によって表される．式 (4) に Auxiliary Selection の損失関数 L_{AS} を示す．

$$L_{\mathrm{AS}} = L_{\mathrm{AS}v} + L_{\mathrm{AS}p} \tag{4}$$

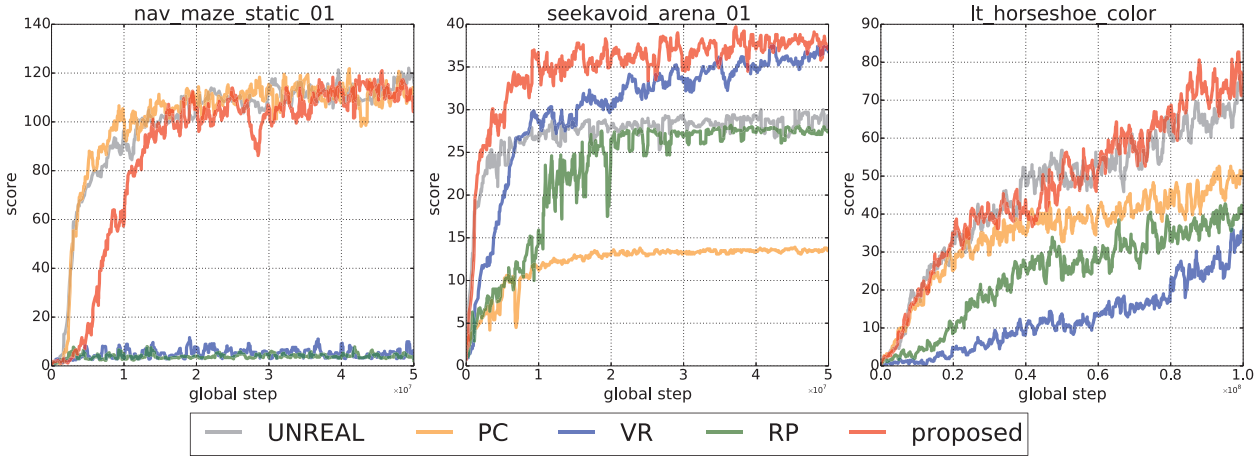
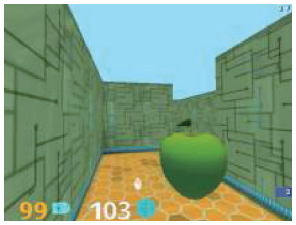
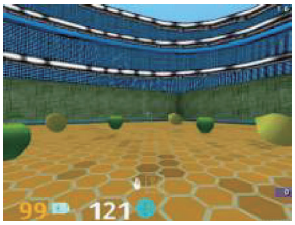


図 2: DeepMind Lab におけるステップ毎のスコア



(a) nav_mazze_static_01



(b) seekavoid_arena_01

(c) lt_horseshoe_color

図 3: DeepMind Lab の各ゲーム画面

4. 評価実験

4.1 実験環境

本論文では，実験環境として，DeepMind Lab [Beattie 16] を用いる．DeepMind Lab は，一人称視点映像の 3D ゲーム環境であり，大きく分けて nav_maze_static_01 (maze), seekavoid_arena_01 (seekavoid), lt_horseshoe_color (horseshoe) の 3 つのゲームが存在する．

maze は一人称視点の迷路探索ゲームである．道中にあるリンゴを獲得すると+1，ゴールに到達すると+10 のスコアを獲得することが可能であり，時間内に獲得できたスコアを競うゲームである．エージェントが取りうる行動は，左視点移動，右視点移動，前へ進む，後ろへ進む，左に平行移動，右に平行移動の計 6 つである．

seekavoid はステージ内の特定の物体を集めるゲームである．リンゴを獲得すると+1，レモンを獲得すると-1 のスコアを獲得し，時間内に獲得できたスコアを競う．行動については，左視点移動，右視点移動，前へ進む，後ろへ進む，左に平行移動，右に平行移動の計 6 つである．

horseshoe は一人称視点シューティングゲームである．ステージ内にスポーンする敵をレーザーで攻撃し倒すと+1 のスコアを獲得することが可能であり，時間内で獲得したスコアを競う．行動については，左視点移動，右視点移動，前へ進む，後ろへ進む，左に平行移動，右に平行移動，攻撃の計 7 つである．

4.2 実験概要

前述の DeepMind Lab の 3 つのゲームにおいて，ステップ毎のスコアを比較することで提案手法の有効性を確認する．比較手法として，全補助タスクを用いた場合 (UNREAL)，Pixel Control のみの場合 (PC)，Value Function Replay のみの場合 (VR)，Reward Prediction のみの場合 (RP)，提案手法 (proposed) の 5 通りで学習を行う．学習時の各ハイパーパラメータは変更せず，maze および seekavoid では 50,000,000 ステップ，horseshoe では 100,000,000 ステップまで学習を行う．また，1 エピソード内での各補助タスクの選択回数を調査することで，最適な補助タスクの選択が実現できているか確認する．worker 数は 8 で行う．

4.3 実験結果

4.3.1 nav_maze_static_01

nav_maze_static_01 におけるステップ数毎のスコアを図 2 左に示す．maze においては，UNREAL と Pixel Control が約 110 の高いスコアを獲得している．また，Value Function Replay と Reward Prediction では 0 に近く，全くスコアを獲得できていないことが分かる．これは，壁の端の画素が大きく変化する行動を Pixel Control によって獲得することが可能であり，この行動が maze の迷路攻略に有効である為だと考えられる．したがって，maze では UNREAL か Pixel Control が最適な補助タスクの選択であると言える．

提案手法は，UNREAL と Pixel Control と同様のスコアを獲得できることが分かる．これら 2 つの結果から，maze において，提案手法は最適な補助タスクと同様に高いスコアを獲得していることが確認できる．

4.3.2 seekavoid_arena_01

seekavoid_arena_01 におけるステップ数毎のスコアを図 2 中央に示す．seekavoid においては，UNREAL と比較し，Value Function Replay が UNREAL を約 10 スコア上回っていることが分かる．また，maze で有効な Pixel Control は約 14 であり，高いスコアを獲得できていない．これは，ゲーム画面の画素が大きく変化する行動がゲーム攻略とは適しておら

表 1: 1 エピソードにおける補助タスクの選択回数

環境 \ 補助タスク	PC	VR	RP
maze	435.4 (48.3%)	487.8 (54.1%)	369.0 (41.0%)
seekavoid	0.3 (0.1%)	300.0 (100.0%)	0.0 (0.0%)
horseshoe	8545.1 (94.9%)	14.1 (0.1%)	8998.2 (99.9%)

ず， 報酬が密に獲得できるゲームである為，Pixel Control と Reward Prediction が有効ではないと考えられる．したがって，seekavoid では Value Function Replay が最適な補助タスクの選択であると言える．

提案手法は，Value Function Replay と同様のスコアを獲得できることが分かる．これら 2 つの結果から，seekavoid において，提案手法は最適な補助タスクと同様に高いスコアを獲得していることが確認できる．

4.3.3 lt_horseshoe_color

lt_horseshoe_color におけるステップ数毎のスコアを図 2 右に示す．horseshoe において，UNREAL が約 75 で最もスコアが高く，各補助タスクのみでは Pixel Control が約 50 で高いスコアを獲得している．これは，敵を倒す行動が画素を大きく変化させる行動にあたる為，各補助タスクのみでは Pixel Control が最も有効であったと考えられる．したがって，horseshoe では UNREAL が最適な補助タスクの選択であると言える．

提案手法は，UNREAL と同様のスコアを獲得できることが分かる．これら 2 つの結果から，horseshoe において，提案手法は最適な補助タスクと同様に高いスコアを獲得していることが確認できる．

4.3.4 選択された補助タスクの解析

各ゲームの 1 エピソードにおける補助タスクの選択回数を表 1 に示す．ここで，選択回数とは 50 エピソード間の平均の選択回数であり，括弧内は 1 エピソード内で選択する割合を表す．1 エピソードの総ステップ数は，maze では 900，seekavoid では 300，horseshoe では 9,000 である．seekavoid では最適な補助タスクである Value Function Replay，horseshoe では Pixel Control と Reward Prediction を安定して選択し，maze では全ての補助タスクが同等に選択されている．maze において，UNREAL と同様に全ての補助タスクを選択するため，最適な補助タスクである UNREAL と同等のスコアを獲得したと考えられる．したがって，UNREAL に Auxiliary Selection を導入することで，環境に合わせた補助タスクを選択でき，効率的な学習を実現していると言える．

5. おわりに

本研究では，学習に用いる補助タスクを適応的に選択するタスク Auxiliary Selection を提案した．提案手法では，各補助タスクの損失関数と Auxiliary Selection により出力する重みの積を取ることで，学習時における最適な補助タスクの選択を実現した．これにより，環境に合わせた補助タスクを設計する必要がなく，補助タスクを用いた学習において，効率化することが可能である．DeepMind Lab を用いた実験により，効率的に学習できることを示した．今後の予定としては，異なる環境や多様な補助タスクを導入した場合における提案手法の有効性の調査などが挙げられる．

参考文献

[Beattie 16] Beattie, C., Leibo, J. Z., *et al.*: DeepMind Lab, *arXiv preprint, arXiv:1612.03801* (2016)

[Firoiu 17] Firoiu, V., Whitney, W. F., *et al.*: Beating the World's Best at Super Smash Bros. Melee with Deep Reinforcement Learning, *arXiv preprint, arXiv:1702.06230* (2017)

[Gu 17] Gu, S., Holly, E., *et al.*: Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates, in *ICRA*, pp. 3389–339 (2017)

[Jaderberg 16] Jaderberg, M., Mnih, V., *et al.*: Reinforcement Learning with Unsupervised Auxiliary Tasks, *arXiv preprint, arXiv:1611.05397* (2016)

[Justesen 17] Justesen, N., Bontrager, P., *et al.*: Deep Learning for Video Game Playing, *arXiv preprint, arXiv:1708.07902* (2017)

[Liebel 18] Liebel, L. and Körner, M.: Auxiliary Tasks in Multi-task Learning, *arXiv preprint, arXiv:1805.06334* (2018)

[Mnih 15] Mnih, V., Kavukcuoglu, K., *et al.*: Human-level control through deep reinforcement learning, *Nature*, Vol. 518, No. 7540, pp. 529–533 (2015)

[Mnih 16] Mnih, V., Badia, A. P., *et al.*: Asynchronous Methods for Deep Reinforcement Learning, in *ICML*, pp. 1928–1937 (2016)

[Rajeswaran 17] Rajeswaran, A., Kumar, V., *et al.*: Learning Complex Dexterous Manipulation with Deep Reinforcement Learning and Demonstrations, *arXiv preprint, arXiv:1709.10087* (2017)

[Riedmiller 18] Riedmiller, M., Hafner, R., *et al.*: Learning by Playing – Solving Sparse Reward Tasks from Scratch, in *ICML* (2018)

[Silver 16] Silver, D., Huang, A., *et al.*: Mastering the game of Go with deep neural networks and tree search, *Nature*, Vol. 529, No. 7587, p. 484 (2016)

[Teh 17] Teh, Y., Bapst, V., *et al.*: Distral: Robust multitask reinforcement learning, in *NIPS*, pp. 4496–4506 (2017)

[Watkins 92] Watkins, C. J. and Dayan, P.: Q-Learning, *Machine learning*, Vol. 8, No. 3-4, pp. 279–292 (1992)

3E3-OS-12a-03

ResNetsに対する正則化手法ShakeDrop

ShakeDrop Regularization for ResNet Family

山田 良博 *1 Yoshihiro Yamada　岩村 雅一 *1 Masakazu Iwamura　黄瀬 浩一 *1 Koichi Kise

*1大阪府立大学
Osaka Prefecture University

Overfitting is a crucial problem in deep neural networks, even in the latest network architectures. In this paper, so as to relieve the overfitting effect of ResNet and its improvements (i.e., PyramidNet and ResNeXt), we propose a new regularization method, named ShakeDrop regularization. ShakeDrop is inspired by Shake-Shake, which is an effective regularization method but can be applied to only ResNeXt. ShakeDrop is even more effective than Shake-Shake and can be successfully applied to not only ResNeXt but also ResNet, and PyramidNet. The important key to realize ShakeDrop is stability of training. Since effective regularization often causes unstable training, we introduce a stabilizer of training which is an unusual usage of an existing regularizer. Experiments reveals that ShakeDrop achieves comparable or superior generalization performance to conventional methods.

1. はじめに

ResNet [He 16] は 100 層を超える多層の Convolutional Neural Network (CNN) を実現し，当時の一般物体認識の世界最高精度を達成した．ResNet の根幹を成す Residual Block は，入力 x に対する出力 $G(x)$ が畳み込み処理 $F(\cdot)$ を用いて下記の式で表される．

$$G(x) = x + F(x) \tag{1}$$

ResNet 以降，Residual Block が式 (1) と同じ形で表される Wide ResNet [Zagoruyko 16] や PyramdNet [Han 17]，$G(x) = x + F_1(x) + F_2(x) + \ldots$ で表される ResNeXt [Xie 17]（図 1(a)）といった派生手法が提案され，相次いで当時の一般物体認識における世界最高精度を達成した．

ネットワーク構造の工夫が進む一方で，異なる形の工夫によって認識精度を向上させた手法が Shake-Shake [Gastaldi 17]（図 1(b)）である．CNN は forward path と backward path からなる一連の学習過程によって，重みを更新する．通常は望ましい出力に近づくように重みを更新するところを，Shake-Shake は forward path と backward path で異なる乱数を掛ける「確率的な正則化」によって，敢えて計算を乱す．これは常識を覆す工夫であったが，Shake-Shake はこの工夫によって当時の一般物体認識における世界最高精度を更新した．しかし Shake-Shake は Residual Block が $G(x) = x + F_1(x) + F_2(x)$ で表される一部の ResNeXt を前提としているため，式 (1) の形で表される ResNet 等に導入出来ない．

本稿では，式 (1) で表される ResNet 等に適用できる，Shake-Shake と同様の確率的な正則化手法を提案する．ただし，単に Shake-Shake における摂動を式 (1) に持ち込むだけでは，強い摂動によって学習が不安定になる．そこで，学習の安定化のために，Stochastic Depth (ResDrop) [Huang 16] で提案された正則化手法を本来の使途とは異なる形で導入する．提案手法は摂動を含む不安定な学習と通常の学習を確率的に切り替えることによって，強い摂動の恩恵を享受しつつ，安定した学習を実現する．その際，Residual Block の $F(x)$ に負の係数をも乗じる強い摂動を加えることで，従来手法より認識精度を向上させている．

連絡先:
山田 良博: yamada@m.cs.osakafu-u.ac.jp
岩村 雅一: masa@cs.osakafu-u.ac.jp
黄瀬 浩一: kise@cs.osakafu-u.ac.jp

2. 提案手法

2.1 Shake-Shake の考察と Single-branch Shake

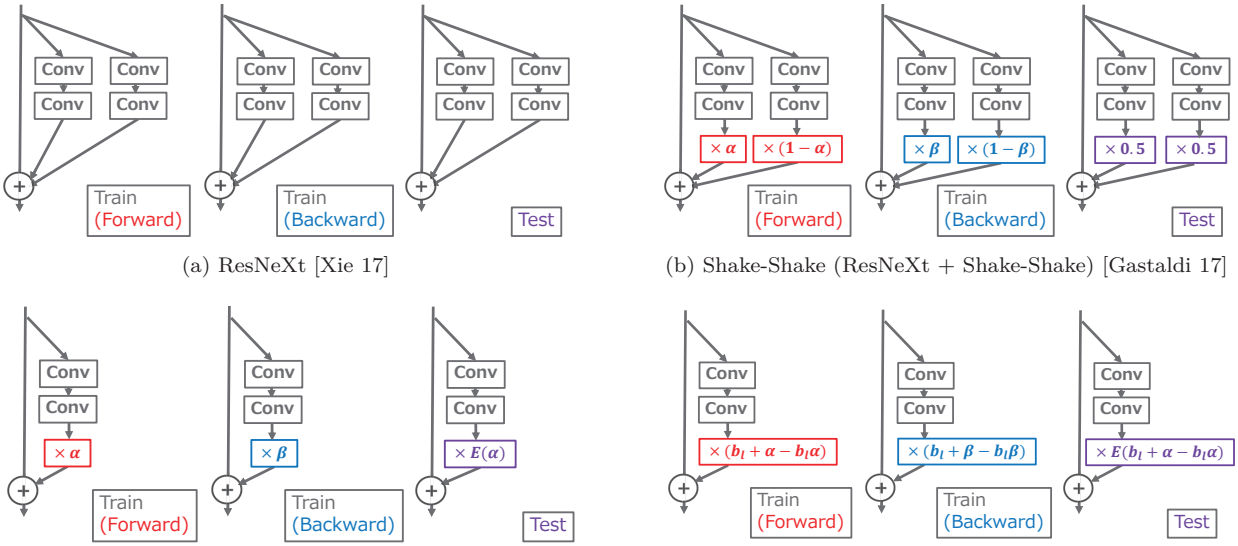
従来手法 Shake-Shake の確率的な正則化は以下の式で表される．

$$G(x) = \begin{cases} x + \alpha F_1(x) + (1-\alpha)F_2(x), & \text{in train-fwd} \\ x + \beta F_1(x) + (1-\beta)F_2(x), & \text{in train-bwd} \\ x + 0.5F_1(x) + 0.5F_2(x), & \text{in test.} \end{cases} \tag{2}$$

ただし，学習時の forward path，学習時の backward path，テスト時の forward path 計算をそれぞれ train-fwd，train-bwd，test と表記している．α，β はそれぞれ $\alpha \in [0,1]$，$\beta \in [0,1]$ の一様乱数である．テスト時には α 及び $1-\alpha$ の期待値 0.5 を用いる．

式 (2) の解釈はこれまで与えられていないが，[DeVries 17] の知見に基づくと，次のように解釈できる．まず，式 (2) から，Shake-Shake の foraward path では $F_1(x)$ と $F_2(x)$ の内分点を求めている．[DeVries 17] では，2 つのデータの特徴表現の加重和を取ることで，新たなデータが生成できることが示されている．すなわち，data augmentation である．したがって，Shake-Shake の forward path では，乱数 α に基づいて，学習データに含まれないデータが特徴空間で生成されていると考えられる．一方，backward path については，[Gastaldi 17] で実験的に検討されており，β が α と離れる程，正しく重みが更新されないことが示されている．すなわち，Shake-Shake は，forward path では特徴空間で新たなデータを生成し，backward path では重み更新を乱していると考えられる．

この知見に基づき，式 (1) の形で表される ResNet 等に適用可能な正則化として次式を考え，Single-branch Shake と名付



(a) ResNeXt [Xie 17]

(b) Shake-Shake (ResNeXt + Shake-Shake) [Gastaldi 17]

(c) PyramidNet + *"Single-branch Shake."*

(d) PyramidNet + ShakeDrop

図 1: ネットワーク構造の模式図

ける（図 1(c)）．

$$G(x) = \begin{cases} x + \alpha F(x), & \text{in train-fwd} \\ x + \beta F(x), & \text{in train-bwd} \\ x + 0.5F_1(x), & \text{in test.} \end{cases} \quad (3)$$

ただし α，β は Shake-Shake に習って，それぞれ $\alpha \in [0,1]$，$\beta \in [0,1]$ の一様乱数である．テスト時には α の期待値 0.5 を用いる．なお，[DeVries 17] では，2 つのデータの特徴表現の内分点のみでなく，ノイズを加えることで新たなデータを生成できることを示しており，data augmentation の観点からは，これでも十分に思える．しかし, 単に α, β を適用するだけでは学習が安定しないことが簡単な実験で確認された *1．そのため，Single-branch Shake では学習の安定化が必要になる．

2.2 ShakeDrop

Single-branch Shake の否定的な結果は，学習を安定化させる工夫の必要性を示唆している．そこで，ResDrop を通常とは異なる目的で用いることを提案する．本来 ResDrop は，深い ResNet の学習時に勾配が消失する問題を解決すべく，確率的に選んだ一部の層を無視して学習を繰り返すという手法である．すなわち，学習時に見かけの層数を浅くすることにより，勾配消失を防ぐ手法である．しかし，この手法は Single-branch Shake にこのまま適用しても学習の安定化にはつながらないと考えられる．何故なら，Single-branch Shake の問題は層が深いことでは無く，重みの更新時に著しい摂動が加えられている事である．そこで本研究では，学習時に一部の層を無視する代わりに，その部分だけ正しく重みが更新されないネットワークと置き換える．これにより，ネットワーク全体としては正しい重み更新が行われるものの，一部では正しくない重みが更新され，それらのバランスにより，ネットワークの学習が安定化しながら強い摂動が加わることが期待できる．この工夫を加えて学習を安定化した手法を ShakeDrop と呼ぶことにする．

提案手法 ShakeDrop には，確率 $p_l = 1 - \frac{l}{2L}$ のベルヌーイ分布の二値乱数 $b_l \in B(p_l)$ による制御を導入する．これは ResDrop で提案された学習安定化手法である Linear Decay Rule に基づく制御である．ただし，L と l は，入力から数えて L 個中 l 番目の Residual Block を表す．提案手法 ShakeDrop（図 1(d)）は次式で与えられる．

$$G(x) = \begin{cases} x + (b_l + \alpha - b_l\alpha)F(x), & \text{in train-fwd} \\ x + (b_l + \beta - b_l\beta)F(x), & \text{in train-bwd} \\ x + E(b_l + \alpha - b_l\alpha)F(x), & \text{in test.} \end{cases} \quad (4)$$

ただし α，β はそれぞれ $\alpha \in [-1,1]$，$\beta \in [0,1]$ の一様乱数である．また，$E(\cdot)$ は期待値である．α，β，b_l はパラメータ更新の度に決定され，二値乱数 b_l は制御に用いられる．$b_l = 1$ のとき α，β は打ち消され，式 (1) になる．$b_l = 0$ のとき，前向きと後ろ向きはそれぞれ $G(x) = x + \alpha F(x)$，$G(x) = x + \beta F(x)$ となり，式 (2) と同様の摂動を含む計算となる．テスト時には前向きの $(b_l + \alpha - b_l\alpha)$ の期待値を利用する．

予備実験の結果，提案手法 ShakeDrop は下記の条件を満たすネットワーク構造であれば機能することが分かった．

1. 加算の直前に Batch Normalization (BN) が存在する．
2. 加算の直後に ReLU が存在しない．

この 2 つの条件を満たすものの多くは，EraseReLU として提案されたネットワーク構造であり, 広範な検証の中でオリジナルの構造のエラー率を下回る優れた構造であることが報告されている [Dong 17].

3. 実験

3.1 ShakeDrop における α と β の影響

α 及び β について表現可能なパラメータ範囲における実験結果を表 1 に示す．この結果から，α と β の組み合わせの効果を確認できる．B($\alpha = 0$，$\beta = 0$) の PyramidDrop（PyramidNet に ResDrop を適用）を除いて，A($\alpha = 1$，$\beta = 1$) の PyramidNet のエラー率を下回ったのは，N ($\alpha = [-1,1]$，$\beta = 0$) と O ($\alpha = [-1,1]$，$\beta = [0,1]$) だけだった．これらの中で，O が最も優れていた．この結果に基づき，以降の実験では O を用いた．

*1 PyramidNet-110 $\alpha = 270$ に Single-branch Shake を適用した場合，CIFAR-100 のエラー率が 77.99%だった．

表 1: CIFAR-100 における複数の α と β の範囲での PyramidNet+ShakeDrop の Top-1 error 率の 4 回平均 (%)．ただし，Case A および Case B はそれぞれ PyramidNet [Han 17], PyramidDrop [Yamada 16] と呼ばれる既存手法に該当する．

Case	α	β	Error (%)
A (既存手法)	1	1	18.01
B (既存手法)	0	0	17.74
C	1	0	20.87
D	1	$[0,1]$	18.80
E	1	$[-1,1]$	21.69
F	0	1	99.00
G	0	$[0,1]$	99.00
H	0	$[-1,1]$	99.00
I	$[0,1]$	1	38.48
J	$[0,1]$	0	19.68
K	$[0,1]$	$[0,1]$	18.27
L	$[0,1]$	$[-1,1]$	20.61
M	$[-1,1]$	1	18.68
N	$[-1,1]$	0	17.28
O	$[-1,1]$	$[0,1]$	**16.22**
P	$[-1,1]$	$[-1,1]$	18.26

3.2 従来の正則化手法との比較実験

提案手法 ShakeDrop の有効性を確認するために，一般物体認識用データセット CIFAR-100 [Krizhevsky 09] を用いて実験を行った．

Residual Block が式 (1) と $G(x) = x + F_1(x) + F_2(x)$ で表されるネットワークを用意し, 正則化なし (Vanilla)，ResDrop，Shake-Shake，ShakeDrop（提案手法）の認識精度を比較した．$G(x) = x + F_1(x) + F_2(x)$ で表されるネットワークでは，確率的な正則化の導入方法として，正則化モジュールを add モジュールの前と後に入れる 2 パターンが考えられる．本稿では後に入れるものを Type-A，前に入れるものを Type-B と呼称し，それぞれについて検証を行う．ただし実験における一部のネットワークは，オリジナルのものではなく提案手法 ShakeDrop が適用可能な構造（EraseReLU）に変更した．

CIFAR-100 結果を表 2 に示す．提案手法 ShakeDrop は PyramidNet に限らず表 2 中の条件で最も多く最良値を示しており，優れた正則化として機能していた．また, 既存手法 Shake-Shake との比較において提案手法 ShakeDrop は優れた結果を示している．Type-A と Type-B については，Type-B の方が優れた認識精度を達成する傾向が確認された．

3.3 他手法との比較

最先端との手法の比較では，1800 epoch の Cosine Learning Scheduling [Gastaldi 17] と Random Erasing [Zhong 17] を導入した PyramidNet-272 ($\alpha = 200$) で，提案手法 ShakeDrop が CIFAR-10/100 におけるエラー率 2.31%と 12.19%をそれぞれ達成し，正則化無し（Vanilla; その他の条件は同じ）のエラー率 3.42%と 16.66%を大幅に下回った．

4. まとめ

一般物体認識において，従来手法 Shake-Shake に代わる新たな確率的正則化手法として ShakeDrop を提案し，ResNet 及びその派生手法での有効性を検証した．CIFAR-100 を用いた実験によって, 提案手法 ShakeDrop による認識精度の改善を確認した．本研究で提案した学習の安定化の方策は，ShakeDrop に適用した摂動以外でも有用な可能性があり，確率的正則化手法を今後研究する上での基盤技術になりうると考えている．

本研究は，JST CREST #JPMJCR16E1，JSPS 科研費 #25240028 と#17H01803 補助による．

参考文献

[DeVries 17] DeVries, T. and Taylor, G. W.: Dataset Augmentation in Feature Space, in *Proc. ICLR Workshop* (2017)

[Dong 17] Dong, X., Kang, G., Zhan, K., and Yang, Y.: EraseReLU: A Simple Way to Ease the Training of Deep Convolution Neural Networks, *arXiv preprint 1709.07634* (2017)

[Gastaldi 17] Gastaldi, X.: Shake-Shake regularization, *arXiv preprint arXiv:1705.07485v2* (2017)

[Han 17] Han, D., Kim, J., and Kim, J.: Deep Pyramidal Residual Networks, in *Proc. CVPR* (2017)

[He 16] He, K., Zhang, X., Ren, S., and Sun, J.: Deep Residual Learning for Image Recognition, in *Proc. CVPR* (2016)

[Huang 16] Huang, G., Sun, Y., Liu, Z., Sedra, D., and Weinberger, K.: Deep Networks with Stochastic Depth, *arXiv preprint arXiv:1603.09382v3* (2016)

[Krizhevsky 09] Krizhevsky, A.: Learning multiple layers of features from tiny images, Technical report, Univ. of Toronto (2009)

[Xie 17] Xie, S., Girshick, R., Dollár, P., Tu, Z., and He, K.: Aggregated Residual Transformations for Deep Neural Networks, in *Proc. CVPR* (2017)

[Yamada 16] Yamada, Y., Iwamura, M., and Kise, K.: Deep Pyramidal Residual Networks with Separated Stochastic Depth, *arXiv preprint arXiv:1612.01230* (2016)

[Zagoruyko 16] Zagoruyko, S. and Komodakis, N.: Wide Residual Networks, in *Proc. BMVC* (2016)

[Zhong 17] Zhong, Z., Zheng, L., Kang, G., Li, S., and Yang, Y.: Random Erasing Data Augmentation, *arXiv preprint arXiv:1708.04896* (2017)

表 2: CIFAR-100 における Top-1 error 率 (%). 表記がないものは 1 回の結果であり，$^{+}$ は 4 回の結果の平均である．* は [Han 17] の結果である．

(a) Residual Block が $G(x) = x + F(x)$ で表される場合 (ResNet, ResNeXt, PyramidNet)

Methods	Regularization	Error (%)
ResNet-110 <Conv-BN-ReLU-Conv-BN-add>	Vanilla	$^{+}$25.38
	ResDrop	$^{+}$22.86
	ShakeDrop	$^{+}$**21.81**
ResNet-164 Bottleneck <Conv-BN-ReLU-Conv-BN-ReLU-Conv-BN-add>	Vanilla	21.96
	ResDrop	20.35
	ShakeDrop	**19.58**
ResNeXt-29 8-64d <Conv-BN-ReLU-Conv-BN-ReLU-Conv-BN-add>	Vanilla	20.25
	ResDrop	20.28
	ShakeDrop	**18.66**
PyramidNet-110 α270 <BN-Conv-BN-ReLU-Conv-BN-add>	Vanilla	$^{+}$18.01
	ResDrop	$^{+}$17.74
	ShakeDrop	$^{+}$**15.78**
PyramidNet-272 Bottleneck α200 <BN-Conv-BN-ReLU-Conv-BN-ReLU-Conv-BN-add>	Vanilla	*16.35
	ResDrop	15.94
	ShakeDrop	**14.96**

(b) Residual Block が $G(x) = x + F_1(x) + F_2(x)$ で表される場合　(ResNeXt)

Methods	Regularization	Error (%)
ResNeXt-164 2-1-40d Bottleneck <Conv-BN-ReLU-Conv-BN-ReLU-Conv-BN-add>	Vanilla	21.75
	ResDrop Type-A	20.44
	ResDrop Type-B	20.21
	Shake-Shake	22.51
	ShakeDrop Type-A	**19.19**
	ShakeDrop Type-B	**18.66**
ResNeXt-29 2-4-64d Bottleneck <Conv-BN-ReLU-Conv-BN-ReLU-Conv-BN-add>	Vanilla	99.00
	ResDrop Type-A	20.13
	ResDrop Type-B	19.01
	Shake-Shake	18.82
	ShakeDrop Type-A	**18.49**
	ShakeDrop Type-B	**17.80**

Organized Session | Organized Session | [OS] OS-12

[3E4-OS-12b] 画像と AI（ MIRU2019プレビュー）(2)

長原 一（大阪大学）、川崎 洋（九州大学）、岡部 孝弘（九州工業大学）

Thu. Jun 6, 2019 3:50 PM - 5:10 PM Room E (301A Medium meeting room)

[3E4-OS-12b-01] Neural 3D Mesh Renderer

〇Hiroharu Kato[1], Yoshitaka Ushiku[1], Tatsuya Harada[1] （1. The University of Tokyo）

3:50 PM - 4:10 PM

[3E4-OS-12b-02] Between-class Learning for Image Classification

〇Yuji Tokozume[1], Yoshitaka Ushiku[1], Tatsuya Harada[1,2] （1. The University of Tokyo, 2. RIKEN）

4:10 PM - 4:30 PM

[3E4-OS-12b-03] A Generative Framework for Creative Data Based on the Generative Adversarial Networks

〇Riku Fujimoto[1], Takato Horii[1], Tatsuya Aoki[1], Takayuki Nagai[1,2] （1. The University of Electro-Commnications, 2. Osaka University）

4:30 PM - 4:50 PM

[3E4-OS-12b-04] Modeling of prejudice eyes by designers

〇Ryuichi Ishikawa[1], Kou Izumi[1], Hidekazu Hayashi[1], Hiroyuki Fukuda[2] （1. DentsuDigital,Inc, 2. Dentsu,Inc）

4:50 PM - 5:10 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Neural 3D Mesh Renderer

加藤 大晴 *1 Hiroharu Kato　牛久 祥孝 *2 Yoshitaka Ushiku　原田 達也 *1*2 Tatsuya Harada

*1東京大学 The University of Tokyo　*2理化学研究所 RIKEN

We introduce our paper "Neural 3D Mesh Render" presented at CVPR and MIRU last year. In this work, we proposed a novel renderer that takes a 3D mesh, light, and camera setting and outputs an image. Because "back-propagation" is defined in our renderer, it can be used as a layer of deep neural networks. By using it, we can pass the gradient of a loss into a 3D space through renderer and optimize components there. In experiments, we demonstrated the effectiveness of our renderer by applying it to view-based training of single-view 3D reconstruction, 2D-to-3D style transfer, and 3D DeepDream. We also introduce some papers that use our renderer for other problems.

1. はじめに

本発表は既発表論文 [Kato 18] の紹介である．

2 次元画像を受け取り，その背後の 3 次元世界における何かを推定（最適化）することは，コンピュータビジョンにおける本質的な課題のひとつである．画像を元に 3 次元構造を推定するタスクはその直接的な例であるが，シーン理解，ポーズ認識などの多くのタスクについても間接的に画像の 3 次元構造の推定が必要とされる．

近年では，深層学習によってパイプライン全体を一気通貫学習するアプローチが広く用いられている．ここでたとえば，画像を受け取りそこに映る物体の 3 次元モデルを出力する再構成器を，2 次元画像のみを用いて訓練することを考える．目的関数として「出力された 3 次元モデルが，2 次元に再投影されたときに，正しい画像と同じように見えるか」という規準を用いる．そのようなパイプラインを図 1 (上) に示す．このパイプラインを一気通貫学習するためには，2 次元画像上で定義される目的関数の勾配が，3 次元モデルを 2 次元へ投影するレンダラーを通じて 3 次元世界へと流れ込まなければならない．また，図 1 (下) に示す「3 次元モデルを，2 次元画像上で定義されたスタイルの類似性を高めるように最適化する」というパイプラインについても同様の処理が必要となる．

この「2 次元画像への投影」を深層学習に組み込むことは，ボクセルと呼ばれる 3D 表現については行われてきた [Yan 16]．しかし，ボクセルはピクセルを 3 次元に拡張したものであり，メモリ使用量が解像度の 3 乗に比例して大きくなるという問題がある．本研究では，3 次元表現として頂点と面の集合から成るメッシュを対象に，深層学習に組み込めるレンダラーを提案する．通常のレンダラーには，各ピクセルごとに色をサンプリングするラスタライズと呼ばれる過程において，ピクセル色を頂点座標で微分すると常にゼロとなり勾配が流れないという問題があるため，本研究ではラスタライズに疑似的な勾配を定義することによってそれを解消する．

2. 提案手法の概略

レンダラーは，3D モデル，照明情報，カメラ情報を受け取り，画像を出力する．この処理は頂点座標の変換やラスタライズなどから成るが，ラスタライズを除くほとんどの処理は微分が自然に定義されるため，勾配を特別に考える必要はない．

連絡先: {kato,ushiku,harada}@mi.t.u-tokyo.ac.jp

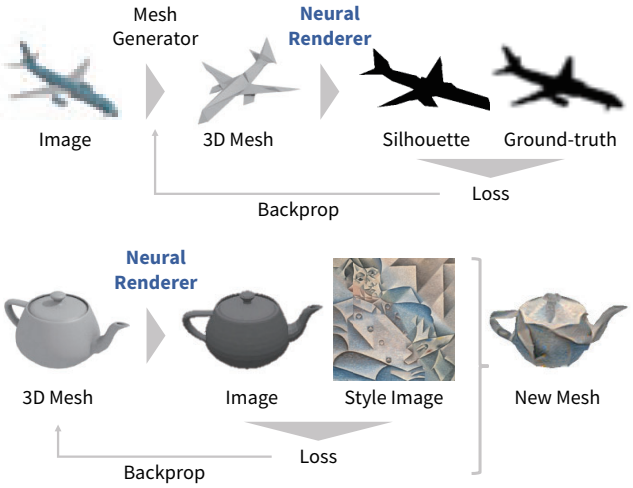


図 1: Pipelines for single-image 3D mesh reconstruction (upper) and 2D-to-3D style transfer (lower).

図 2 (a) のように，1 枚のポリゴンのみから成るシーンを考える．ラスタライズでは，ピクセルがポリゴンと重なるとき，ピクセルがそのポリゴンの色で塗られる．図 2 (b) はピクセル P_j の色 I_j とポリゴンの頂点の座標 x_i の関係を示している．頂点 x_i が右に動くと，ポリゴンが P_j に衝突したときにそのポリゴンの色が変わる．図 2 (c) はこの関数の微分値を示し，これはほぼ常にゼロであるため誤差逆伝播がうまく機能しない．そこで本研究では，逆伝播時にのみ図 2 (b) ではなく図 2 (d) のような関係が成立しているとみなし，その傾きを勾配とする．これによって逆伝播時に「頂点がどちらに動くとピクセルの色がどう変わりそうか」という情報を伝えることができ，レンダラーを含むパイプラインの一気通貫学習が可能となる．

3. 実験

3.1 単一画像 3 次元再構成

図 3 および表 1 は，図 1 (上) のパイプラインを用いて三次元再構成器を訓練する実験の結果である．データセットは ShapeNet を使用した。メッシュの生成は，予め定義した球の頂点を移動させることによって実現する．適切に 3 次元再構成が行えていることから，2 次元画像上でシルエットを比較する

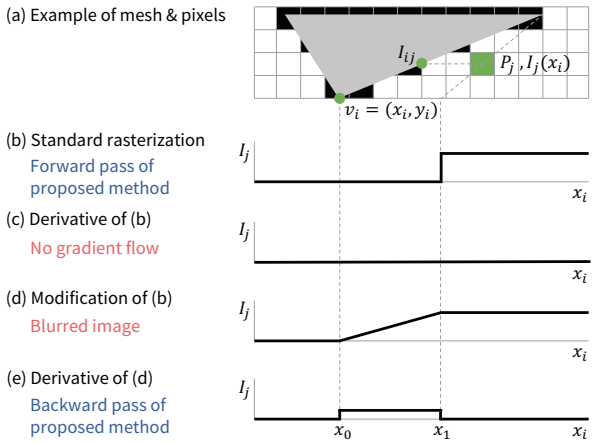


図 2: Illustration of our method. $\boldsymbol{v}_i = \{x_i, y_i\}$ is one vertex of the face. I_j is the color of pixel P_j. The current position of x_i is x_0. x_1 is the location of x_i where an edge of the face collides with the center of P_j when x_i moves to the right. I_j becomes I_{ij} when $x_i = x_1$.

Retrieval [Yan 16]	Voxel [Yan 16]	Mesh (ours)
0.4766	0.5736	**0.6016**

表 1: Reconstruction accuracy measured by voxel IoU. Higher is better. Our mesh-based approach outperforms the voxel-based approach [Yan 16] in 10 out of 13 categories in ShapeNet.

目的関数の勾配がレンダラーを通じて適切に三次元再構成器へと流れていることがわかる．また，我々の手法で得られるメッシュにはノイズボクセルに表れるような粗いノイズがなく，また再構成の定量的な性能もボクセルを用いる場合を上回る．

3.2 2 次元画像から 3 次元モデルへのスタイル転移

図 4 は，図 1 (下) のパイプラインを用いて 2 次元の画像のスタイルを 3 次元モデルへと転移する実験の結果である．スタイルの類似性の計算には [Gatys 16] を用いた．

図からは，ウサギのテクスチャ画像にスタイル画像の色が転移していることや，ティーポットのふたの形が円形からスタイル画像のように直線的な形へと変化していることが読み取れる．これは，提案したレンダラーを用いることでスタイルに関する情報を 2 次元画像の空間から 3 次元画像の空間へと適切に流し込めることを示している．

4. まとめ

本研究では深層学習に組み込むことのできるメッシュのレンダラーを提案し，単一画像 3 次元再構成と，画像から 3 次元モデルへのスタイル転移でその効果を検証した．

提案したレンダラーの応用可能性は本論文で示したものに留まらない．既に画像集合からのテクスチャ付きの 3 次元モデルの再構成 [Kanazawa 18] や 3 次元再構成を経由した画像の編集 [Yao 18] などへと応用されており，今後も様々なタスクへと広がってゆくことが期待される．本研究のソースコードはオンラインで公開している *1．

*1 `https://github.com/hiroharu-kato/neural_renderer`

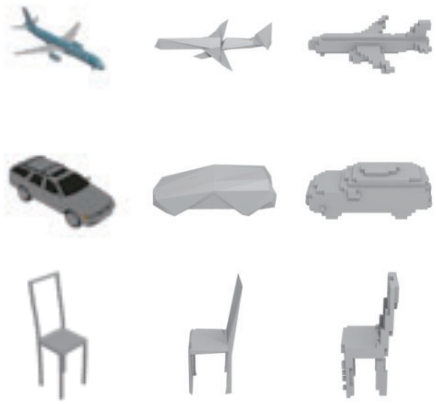


図 3: 3D mesh reconstruction from a single image. Results are rendered from three viewpoints. First column: input images. Second column: mesh reconstruction (proposed method). Third column: voxel reconstruction [Yan 16].

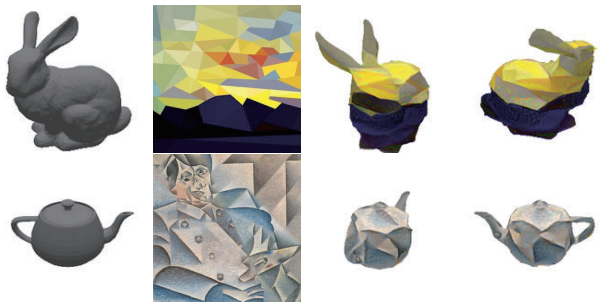


図 4: 2D-to-3D style transfer. The two left images represent initial state of a 3D model and style respectively. The style images are *Thomson No. 5 (Yellow Sunset)* (D. Coupland, 2011) and *Portrait of Pablo Picasso* (J. Gris, 1912).

参考文献

[Gatys 16] Gatys, L. A., Ecker, A. S., and Bethge, M.: Image style transfer using convolutional neural networks, in *CVPR* (2016)

[Kanazawa 18] Kanazawa, A., Tulsiani, S., Efros, A. A., and Malik, J.: Learning Category-Specific Mesh Reconstruction from Image Collections, in *ECCV* (2018)

[Kato 18] Kato, H., Ushiku, Y., and Harada, T.: Neural 3d mesh renderer, in *CVPR* (2018)

[Yan 16] Yan, X., Yang, J., Yumer, E., Guo, Y., and Lee, H.: Perspective transformer nets: Learning single-view 3d object reconstruction without 3d supervision, in *NIPS* (2016)

[Yao 18] Yao, S., Hsu, T. M., Zhu, J.-Y., Wu, J., Torralba, A., Freeman, B., and Tenenbaum, J.: 3D-aware scene manipulation via inverse graphics, in *NIPS* (2018)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Between-class Learning for Image Classification

床爪 佑司 *1 Yuji Tokozume　牛久 祥孝 *1 Yoshitaka Ushiku　原田 達也 *1*2 Tatsuya Harada

*1東京大学 The University of Tokyo　*2理化学研究所 RIKEN

We introduce our paper "Between-class Learning for Image Classification" presented at CVPR and MIRU last year. In this paper, we propose a novel learning method for image classification called between-class learning (*BC learning*). We generate between-class images by mixing two images belonging to different classes with a random ratio. We then input the mixed image to the model and train the model to output the mixing ratio. BC learning has the ability to impose constraints on the shape of the feature distributions, and thus the generalization ability is improved. As a result, we achieved 19.4% and 2.26% top-1 errors on ImageNet-1K and CIFAR-10, respectively.

1. はじめに

本発表では，CVPR 2018 で発表した論文 [Tokozume 18a] について，先駆けとして ICLR 2018 で発表した論文 [Tokozume 18b] の内容と合わせて紹介する．

音や画像の認識において，深層学習を用いた手法が高い性能を発揮している．深層学習は，線形分離不可能なデータ空間から線形分離可能な特徴空間への関数を学習する．限られた学習データから出来る限り判別的な特徴空間を学習することが，深層学習における重要な課題である．

そこで本研究では，限られた学習データから判別的な特徴空間を学習できる，深層ニューラルネットワークの新しい教師付学習手法を提案する．新しい教師付学習手法には，ネットワーク構造や正則化等の従来の学習技術に影響を与えないこと，限られた学習データを効率的に使えること，判別的な特徴空間を学習できること，の 3 つが求められる．

ここで，判別的な特徴空間とはどのようなものだろうか．まず，クラス間の Fisher's criterion [Fisher 36] が大きい特徴空間は判別的である．Fisher's criterion とは，クラス内分散に対するクラス間距離の比のことであり，2 つのクラスがどの程度判別的であるかを表す指標である．また，各クラスが無相関な特徴空間は判別的である．識別タスクでは各クラスを等価に扱う必要があるため，特徴空間において各クラスが等間隔に並んでいることが望ましい．本研究ではこれら 2 つを判別的な特徴空間の要件とする．

従来の教師付学習では，学習データセットから単一の学習データを選択し，対応するクラスは 1，それ以外は 0 を出力するようにニューラルネットワークを学習していた．このような学習手法では，特徴空間において各クラスが線形分離可能であれば罰則が与えられないので，特徴空間が判別的になる保証は無い．本研究ではこの問題を解決する学習手法を提案する．

2. Between-class Learning

2.1 概要と効果

本研究では，深層ニューラルネットワークの新しい教師付学習手法として，between-class learning (BC learning) を提案する．BC learning では，以下の手順でモデルを学習する．

- 異なるクラスに属する 2 つのデータを選択する．
- それらをランダムな比率で合成し，モデルに入力する．
- 合成比率を出力するようにモデルを学習する．

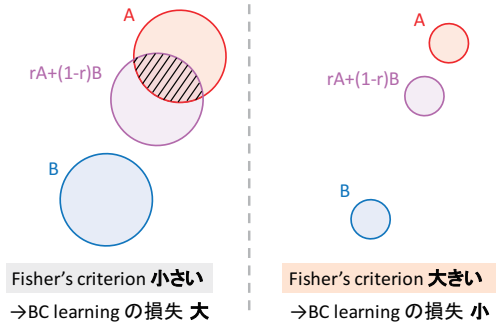


図 1: BC learning による Fisher's criterion の増大.

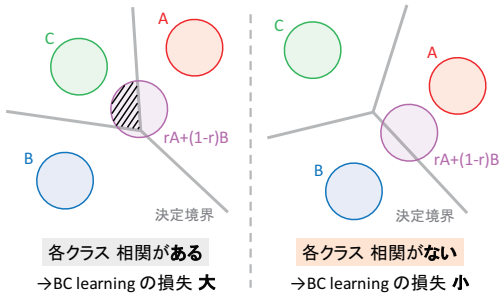


図 2: BC learning による各クラスの無相関化.

BC learning は，従来の学習技術に影響を与えない．また，データの合成によって学習データのパターン数が増えるため，限られた学習データを効率的に使うことができる．さらに，判別的な特徴空間を学習できる効果がある．その理由を以下に示す．

効果 1. Fisher's criterion の増大　図 1 (左) のように，特徴空間においてクラス A, B 間の Fisher's criterion が小さい場合を考える．クラス A, B に属するデータをある比率で合成してモデルに入力した際に，その特徴量分布 (桃色) はクラス A, B のいずれかの特徴量分布と重複することが予想される．このとき，合成するデータの組み合わせによっては，合成したデータがいずれかのクラスに分類されてしまい，モデルが合成比率を出力することができない．そのため，BC learning を行った場合の損失が大きい．一方，図 1 (右) のように Fisher's criterion が大きい場合，重複が発生しないため，BC learning による損失が小さい．学習は損失が小さくなる方向に進むため，BC learning によって図 1 (右) のような Fisher's criterion が大きい特徴空間が学習される．

表 1: CIFAR-10, CIFAR-100 における実験結果.

モデル	学習手法	誤識別率 (%) CIFAR-10	誤識別率 (%) CIFAR-100
11 層 CNN	Standard	6.07 ± 0.04	26.68 ± 0.09
	BC (ours)	5.40 ± 0.07	24.28 ± 0.11
	BC+ (ours)	$\mathbf{5.22 \pm 0.04}$	$\mathbf{23.68 \pm 0.10}$
ResNet-29	Standard	4.24 ± 0.06	20.18 ± 0.07
	BC (ours)	3.75 ± 0.04	19.56 ± 0.10
	BC+ (ours)	$\mathbf{3.55 \pm 0.03}$	$\mathbf{19.41 \pm 0.07}$
ResNeXt-29	Standard	3.54 ± 0.04	$\mathbf{16.99 \pm 0.06}$
	BC (ours)	$\mathbf{2.79 \pm 0.06}$	18.21 ± 0.12
	BC+ (ours)	2.81 ± 0.06	17.93 ± 0.09
DenseNet	Standard	3.61 ± 0.10	17.28 ± 0.12
	BC (ours)	2.68 ± 0.03	16.36 ± 0.10
	BC+ (ours)	$\mathbf{2.57 \pm 0.06}$	$\mathbf{16.23 \pm 0.07}$
Shake-Shake	Standard	2.86	**15.85**
	BC (ours)	2.38 ± 0.04	15.90 ± 0.06
	BC+ (ours)	$\mathbf{2.26 \pm 0.01}$	16.00 ± 0.10

効果 2. 各クラスの無相関化　図 2 (左) のように特徴空間において各クラスに相関がある場合，クラス A, B の合成物がクラス C に分類されるケースが発生するため，BC learning の損失が大きい．一方，図 2 (右) のように各クラスに相関がない場合，クラス A, B の合成物がクラス C に分類されないため，BC learning の損失が小さい．よって，BC learning によって図 2 (右) のような各クラスが無相関な特徴空間が学習される．

2.2 環境音識別への適用 *1

音はデータ同士を合成しても音として成り立つため，BC learning が有効であると考えられる．選択された 2 つの学習データをそれぞれ $\mathbf{x}_1, \mathbf{x}_2$ とし，それらの one-hot ラベルをそれぞれ $\mathbf{t}_1, \mathbf{t}_2$ とする．また，合成比率 r を一様分布 $U(0, 1)$ から生成する．ラベルの合成は単純に $r\,\mathbf{t}_1 + (1-r)\,\mathbf{t}_2$ とする．一方，データの合成は，同様に $r\,\mathbf{x}_1 + (1-r)\,\mathbf{x}_2$ とするのが単純であるが，$\mathbf{x}_1, \mathbf{x}_2$ それぞれの音圧レベル G_1, G_2 (dBA) の差を考慮した以下の合成式を提案する．

$$\frac{p\,\mathbf{x}_1 + (1-p)\,\mathbf{x}_2}{\sqrt{p^2 + (1-p)^2}} \quad \text{where} \quad p = \frac{1}{1 + 10^{\frac{G_1 - G_2}{20}} \cdot \frac{1-r}{r}} \tag{1}$$

2.3 画像識別への適用

画像を合成することは直感に反するが，画像データは x 軸と y 軸に沿った波であると考えられるので，環境音と同様に BC learning が有効であると考えられる．先程と同様に $\mathbf{x}_1, \mathbf{x}_2, \mathbf{t}_1, \mathbf{t}_2, r$ を定義する．ラベルの合成は単純に $r\,\mathbf{t}_1 + (1-r)\,\mathbf{t}_2$ とする．データの合成は，同様に $r\,\mathbf{x}_1 + (1-r)\,\mathbf{x}_2$ とするのが単純であるが，$\mathbf{x}_1, \mathbf{x}_2$ からそれぞれの平均値 μ_1, μ_2 を引いてゼロ平均にしたのちに，環境音と同様に合成することを提案する．音圧レベルの代わりに各画像の標準偏差 σ_1, σ_2 を用いた以下の合成式を提案する．前者の単純な合成方法を BC，後者を BC+と呼ぶことにする．

$$\frac{p\,(\mathbf{x}_1 - \mu_1) + (1-p)\,(\mathbf{x}_2 - \mu_2)}{\sqrt{p^2 + (1-p)^2}} \quad \text{where} \quad p = \frac{1}{1 + \frac{\sigma_1}{\sigma_2} \cdot \frac{1-r}{r}} \tag{2}$$

3. 実験

一般物体画像データセット CIFAR-10, CIFAR-100, および ImageNet-1K を用いて学習・評価を行った．まず，CIFAR-10, CIFAR-100 における実験結果を表 1 に示す．多くの条件において，BC learning によって識別性能が向上した．また，BC+ の方が BC よりも性能が高い傾向にあった．特に CIFAR-10 において 2018 年 1 月現在の世界最高性能 2.26%を達成した．

*1 ICLR 2018 で発表した内容．

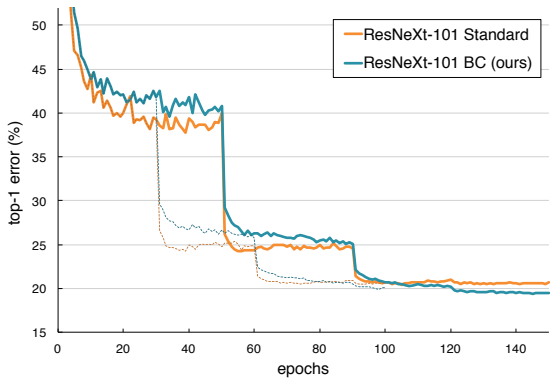


図 3: ImageNet-1K における実験結果．破線は 100 epoch，実線は 150 epoch での実験結果．

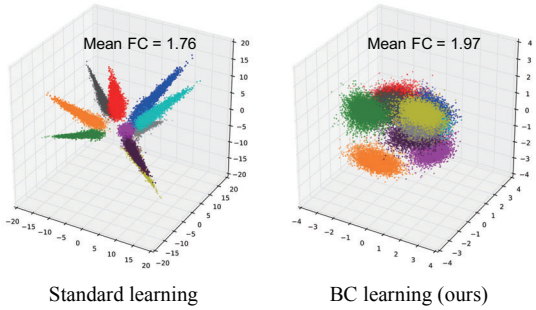


図 4: BC learning によって学習された特徴空間の可視化．

次に，ImageNet-1K における実験結果を図 3 に示す．BC learning によって最終的な誤識別率が 20.4%から 19.4%へ約 1%向上した．BC learning は大規模なデータセットに対しても有効であることが示された．

CIFAR-10 で学習した 11 層 CNN の特徴空間（第 10 層）を PCA を用いて可視化した結果を図 4 に示す．BC learning によって学習された特徴空間は，各クラスが球状にまとまっていることが分かる．また，2 クラス間の Fisher's criterion の平均値も，BC learning の方が大きかった．BC learning によって判別的な特徴空間が学習されたといえる．

4. 結論と今後の展望

本研究では，between-class (BC) learning という深層ニューラルネットワークの新しい教師付学習手法を提案した．実験の結果，BC learning によって画像の識別性能が大きく向上することが示された．BC learning は，音や画像以外のモダリティのデータの識別や，識別以外のタスクにも応用が期待される，非常に汎用性の高い技術である．また，考え方がシンプルで実装も容易であり，実用性も高い．さらに，理論的考察の余地もあり，今後さらなる研究がなされると考えられる．

参考文献

[Tokozume 18a] Y. Tokozume, Y. Ushiku, and T. Harada. Between-class Learning for Image Classification. In *CVPR*, 2018.

[Tokozume 18b] Y. Tokozume, Y. Ushiku, and T. Harada. Learning from between-class examples for deep sound recognition. In *ICLR*, 2018.

[Fisher 36] Ronald A Fisher. The use of multiple measurements in taxonomic problems. *Annals of eugenics*, Vol. 7, No. 2, pp. 179–188, 1936.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

敵対的生成ネットワークを利用した創造的データ生成の枠組み

A Generative Framework for Creative Data Based on the Generative Adversarial Networks

藤元陸 *1 Riku Fujimoto　堀井隆斗 *1 Takato Horii　青木達哉 *1 Tatsuya Aoki　長井隆行 *1*2 Takayuki Nagai

*1電気通信大学 The University of Electro-Commnications　*2大阪大学 Osaka University

In this research, we propose a framework to generate creative data simulating the creation process. This framework generates new and valuable high dimensional data. The characteristics of this framework are two points, a mixed generator and self-generated data learning. The mixed generator makes it possible to generate new data by loss function of regularization by Feature matching and entropy. In self-generated data learning, expressive ability to generate higher value data is acquired by using highly valued generation data as learning data.The framework for generating new and valuable data by combining these two methods is called "Deep Creative Model(DCM)". In the experiments, MNIST was used as learning data, learning a framework to set alphabet images as valuable images. As learning progresses, it was possible to gradually generate images close to the shape of the alphabet, and it was confirmed that it is possible to generate creative data with DCM.

1. はじめに

近年，主に深層学習技術の発展を受け，計算論的創造性に関する研究が進展している [Elgammal 17][Mordvintsev 15][Hawthorne 17][Ha 17]. これらの研究では，人間の創造過程を計算モデルによって模擬することで創造性に関する知見を深めたり，人間の創造性を計算機によって拡張することを目的としている．

創造性には「新規性」と「価値」の 2 つの要素が重要であることが知られている [Sternberg 99]．我々はありふれている物や，一度見たものには創造性を感じにくい．また初見のものでもランダムな文字の羅列のような，一般的に価値の低い物にも創造性を感じにくい．我々が創造的だと感じるデータは，新規性が高くかつ価値の高いデータである．例えば，スマートフォンは画面を触って直感的な操作が可能なため，従来のボタンで操作する携帯電話に対して革新的である．今ではスマートフォンは一般的な製品であるが，発売当初はその操作方法の新規性が高く，かつ，利便性という価値を兼ね備えていたため創造的な製品であった．

このように，創造的なデータを生成するモデルとして，「新規性」と「価値」の両方を兼ね備えたデータを生成することが求められる．様々なデータを生成する手法として敵対的生成ネットワーク (GAN)[Goodfellow 14] や Variational Auto Encoding(VAE)[Kingma 13]，Glow[Kingma 18] が注目されている．しかし，既存の生成モデルでは学習データに対して生成データの尤度が最大になるように学習している．学習データ外のデータ，つまり学習データにとっての新規データは，学習時にはエラーとして扱われる．そのため既存の生成モデルによってモデルがこれまでに学習したことのないような新規データを生成することは困難である．

創造的な作品の多くは，多数の失敗作の上に成り立っている．作品の目標を定め，制作し，評価するという工程を繰り返すことでより創造的な作品を作り出すことができる．このような創造的な作品を制作する過程では，製作者本人はどのような作品に価値があるかわからない場合もある．その場合，他人に作品の評価してもらったり，工業製品であれば実際に使用することで価値を判定したりする．このように，外的な評価によって作品の価値を算出し，その価値をもとに作品を製作することで，作品の価値を高めることができる．

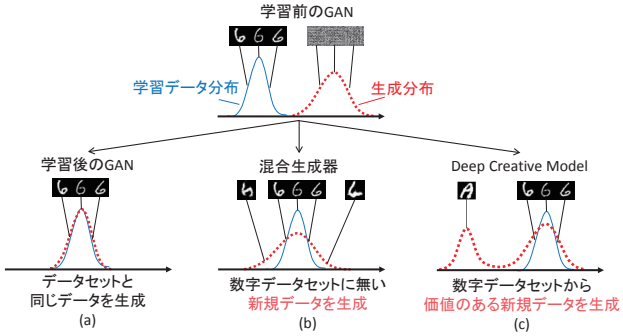


図 1: GAN，混合生成器，DCM の比較.

筆者らは上記の過程を模擬することで，新規データの生成と外的な評価による価値の算出を繰り返し，創造的なデータを生成する枠組みを提案する．本稿では生成モデルが学習データの表現を保持しつつ新規データを生成するための手法と，価値情報に基づいて生成モデルの表現空間を拡張する手法について説明する．さらに提案手法において，学習データに含まれておらず，かつ価値の高いデータを提案手法を用いて実際に生成することで，創造的なデータの生成が可能であることを示す．

2. 関連研究

2.1 敵対的生成ネットワーク

GAN は生成器と識別器で構成される．生成器の目的は画像を生成することであり，識別器の目的は生成器の生成画像と学習データの画像を見分けることである．学習データセットのデータ分布を P_x，生成器 に入力するノイズの分布を P_z，生成器の出力を $G(z)$，生成データの分布を P_G，識別器の出力を $D(x)$ とする．このとき，GAN の学習は以下のミニマックスゲームで表すことができる．

連絡先: 藤元陸 (rfujimoto@apple.ee.uec.ac.jp)

$$\min_G \max_D V(D,G) \\ = (\mathbb{E}_{x\sim P_x}\log(D(x)) + \mathbb{E}_{z\sim P_z}\log(1-D(G(z)))) \tag{1}$$

2.2 補集合分布

創造性において新規性は重要な要素であり，新規データを生成することは創造的なデータの生成に必要不可欠である．しかし式 (5) に従って最適化された GAN では，$P_G = P_x$ となるため新規データを生成することはできない (図 1(a))．新規データを生成するためには $P_G \neq P_x$ とする必要がある．そのため，GAN が最適化されないようにモデルを設計する必要がある．

Dai[Dai 17] らは半教師有りマルチクラス分類タスクの分類精度上昇のために，GAN で新規データを生成する手法を提案した．この研究では，生成器は学習データの補集合を生成するように学習した．生成器の損失関数 L_G は次式で定義される．

$$L_G = -H(G(z)) + \mathbb{E}_{x\sim P_G}\log P(x)\mathbb{I}[P(x)>\epsilon] \\ + ||\mathbb{E}_{z\sim P_z} f((G(z)) - \mathbb{E}_{x\sim P_x} f(x)||^2 \tag{2}$$

この損失関数の第三項は Feature matching[Salimans 16] による損失である．Feature matching は学習データと生成データに対する識別器の中間層の期待値を一致させるため，生成データ分布が学習データ分布に近くなることを間接的に強いる．生成器の損失関数を Feature matching のみで構成すると，通常の損失関数と比べて生成データ分布と学習データ分布を一致させる力が小さいため，学習データ多様体上とその周囲からデータを生成することが期待できる．つまり，学習データ分布内では $P_G < P_x$ となり，データ分布外では $P_G > P_x$ となる．

しかし Feature matching では特徴の期待値を一致させているため，生成データが平均に偏り，モード崩壊することがある．そこで，第一項では生成データのエントロピーの最大化を行っている．これにより，似たデータを多く生成するとペナルティが与えられるため，モード崩壊せず広範囲から生成されるようになる．生成データのエントロピーの計算は，Zhao[Zhao 16] らの pull-away 項が用いられた．N をミニバッチのサイズ，$G(z_i$ をミニバッチの i 番目のデータとすると，式 (3) の L_{PT} で式 (2) の第一項 $-H(G(z))$ が近似される．

$$L_{PT} = \frac{1}{N(N-1)}\sum_{i=1}^{N}\sum_{j\neq i}^{N-1}\left(\frac{f(G(z_i))^\top f((G(z_j))}{\|f((G(z_i))\|\,\|f((G(z_j))\|}\right)^2 \tag{3}$$

第二項は学習データの高密度領域から生成しないための項である．$\mathbb{I}$ は指示関数であり，条件に合うときは 1 を，合わないときは 0 を返す関数である．データが学習データの分布において存在確率がある値 ε より大きいときペナルティを与えるようにし，データ分布 P_x の低密度領域から生成するように強いている．さらに，学習データの分布密度 P_x の推定は PixelCNN++[Salimans 17] が利用されている．学習データセットで事前学習し，生成器の学習中はネットワークの重みは固定して使用された．

まとめると，第三項が生成データ分布を学習データ分布に近づけ，第一項が多様なデータを生成する働きをする．第二項は学習データを生成しないようにする．これら三つの項を含む損失関数を最小化することで，学習データ空間の周囲からデータの生成が可能になる．

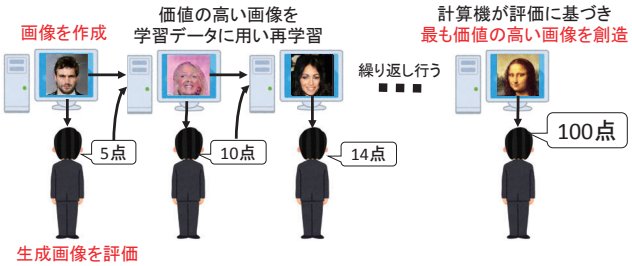


図 2: 創造的なデータ生成手法の概念図.

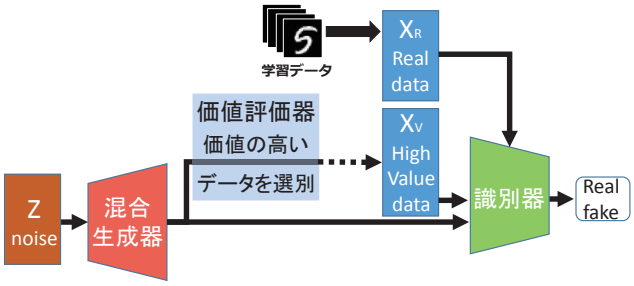


図 3: Deep Creative Model の模式図.

3. 提案手法

3.1 創造的なデータ生成手法の概念

図 2 は提案アイディアの概念図である．本提案はデータの生成，評価，学習の 3 つのフェーズに分かれる．

1. 学習データを学習した生成モデルでデータを生成する．
2. 生成したデータの価値を外部の評価器で評価する．
3. 評価の高いデータを学習データに追加し，生成モデルを再学習する．

ここで，生成モデルが学習データより多様なデータを生成可能ならば，学習データに含まれないデータを生成することができる．そのデータを学習データに加えることで，より多様なデータを生成できるようになる．同様に，学習データに含まれない価値の高い生成データを学習データに加えることでより価値の高いデータが生成可能になる．そのため，生成，評価，学習のプロセスを繰り返すことで徐々に価値の高いデータを生成できるようになる．そこで上記のプロセスを実現するために学習データより多様なデータを生成可能な生成モデルとして混合生成器を提案する．さらに，混合生成器がより価値の高いデータを生成可能にする手法として価値情報を用いた自己生成データ学習を提案する．

3.2 エントロピー項を用いた混合生成器

学習データの表現を保持しつつ，新規データを生成する手法として混合生成器を提案する．混合生成器の損失関数 L_G には

$$L_G = -H(P_G) + ||\mathbb{E}_{x\sim P_G} f(x) - \mathbb{E}_{x\sim P_x} f(x)||^2 \tag{4}$$

を提案する．これは，式 (2) の第二項を除いたものである．式 (2) の第二項は学習データに似たデータを生成しないようにする働きがある．そのため，学習データ分布内に価値の高いデータがあるとき，式 (2) の第二項を用いると最も価値の高いデータを生成することができない．また後述の自己生成データ学習の手法によって学習データが学習中に逐次変更されるため，

図 4: (a) 最も価値の高い画像 X_s (b) 生成画像の中で最も価値の高い画像.

PixelCNN++を変更された学習データセットで再学習する必要があり，式 (2) の第二項の計算コストが非常に高い．

式 (2) の損失関数で学習した生成器は学習データ分布の高密度領域の周囲データの生成を可能にした．一方，式 (4) の損失関数で学習した生成器は学習データ分布の高密度領域とその周囲のデータの生成が可能である (図 1(b))．

3.3 価値情報を用いた自己生成データ学習

混合生成器は，学習データ分布の高密度領域とその周囲のデータを生成する．しかし，生成できるのは学習データの高密度領域の周囲までであり，学習データ分布の高密度領域からより遠いデータを生成することはできない．混合生成器の表現空間を拡大するためには，学習データがより多様になればよい．そこで，学習データ分布の高密度領域の周囲のデータを学習データに加えて再学習する．

しかし，学習データ分布の高密度領域の周囲の大部分のデータは価値の低いデータであると考えられる．そのため，混合生成器からサンプルしたデータを全て学習データとして用いると，学習データの大半が価値の低いデータになる．表現空間を価値の高い方向にのみに拡大するために，学習データに加えるデータを価値の高い生成データに限定する．

3.4 創造的データを生成するモデル

上記の提案手法である混合生成器と自己生成データ学習を用い，創造的なデータを生成するモデル Deep Creative Model(DCM) を提案する．図 3 に DCM の構造を示す．DCM は GAN に混合生成器と自己生成データ学習の二つの変更を加えたものである．通常の GAN のネットワークの損失関数を変更し，混合生成器を学習する．外部の価値評価器で混合生成器から生成されるデータの価値を評価する．価値の高いデータは学習データ X_V として保持する．識別器は真の学習データ X_R と生成した価値の高いデータ X_V を学習データとして用い，True と判定するように学習する．一方，混合生成器が生成したデータ G(x) は False と判定するように学習する．混合生成器は識別器を騙すために，X_R と X_V の分布を学習するとともに，その周囲のデータも生成できるように学習する．学習が進むにつれて X_V として保持されるデータは徐々に価値の高いデータとなり，混合生成器の表現空間も価値の高い方向に拡張されることが期待できる (図 1(c))．

4. 実験

4.1 実験設定

手書き数字画像のデータセット MNIST を用いて DCM を学習させた．価値は主体や環境によって異なるものであり，絶対的な価値は存在しない．そこで，本実験での価値は計算の簡便性の理由で，MNIST データセットにアルファベット画像を加えた EMNIST[Cohen 17] に含まれる特定のアルファベット画像と生成画像のピクセル間距離とした．つまり，最も価値の高い画像を X_s，生成した画像を X としたとき，画像 X の価値 V(X) は

$$V(X) = -(X_s - X)^2 \tag{5}$$

となる．EMNIST のアルファベット画像は価値の評価にのみ使い，混合生成器の学習には使用しなかった．自己生成データ学習では，100step ごとに混合生成器は 60000 枚データをランダムに生成し，その中から価値の高いデータ 6000 枚を X_V として追加した．X_V に追加後，価値の高い 6000 枚以外のデータは破棄した．学習データは X_R と X_V が 1:1 の割合になるようにサンプリングした．

モデルの評価は次の 2 つの観点で行った．

1. 学習初期には生成できない画像を学習後に生成できるか定性的に評価する．
2. 生成画像の価値が学習によって上昇するかを定性的・定量的に評価する．

これらの評価によって，DCM の生成器が新しく価値の高いデータつまり創造的なデータを生成できることを確認する．

4.2 実験結果

4.2.1 DCM の生成画像の定性的評価

DCM の学習結果を図 4 に示す．図 4(a) は最も価値の高い画像 X_s であり，図 4(b) は各 X_s を価値関数に用いて DCM を学習させたときの最も価値の高い生成画像である．(a) と (b) の画像はほぼ同じ形状になっており，DCM で MNIST データから価値情報を用いてアルファベットの画像を生成することができた．MNIST のデータセットに似た形状の文字が存在しない K，N，R，W の文字も生成できており，DCM が学習データに含まれていないデータを生成できることが示された．このことから，DCM で新規データの生成が可能であることがわかる．

図 5 は DCM の学習過程の各段階の最も価値の高い生成画像である．学習初期の段階ではアルファベットとして形を成していない画像が生成された．しかし，学習が進むにつれて徐々に形状が変化し最も価値の高い画像に似た画像が生成された．学習の最初の段階では MNIST データのみで事前学習したモデルで生成しており，このモデルでは混合生成器の表現空間内に文字 K，N，R，W に似た形状の画像が存在していないことがわかる．自己生成データ学習により価値の高い生成データを混合生成器の学習データに加えることで，学習が進むにつれて表現空間が価値の高い方向へ拡張されることがわかる．

4.2.2 DCM の価値の推移による定量的評価

図 6 は学習過程の特徴的な価値の推移を示した画像を対象に図示した．価値の最大値は 0，最小値は-784 である．学習初期の価値が最大の文字は I，最小の文字は W，収束時の最大の文字は J，最小の文字は R であった．K，N は一般的な価値の推移の例として図示した．学習初期と比較して収束時は全ての文字画像の学習条件で価値が上昇した．文字 I や J は学習初期の段階で，既に収束時に近い価値となっている．これは数字の 1 に形状が近いため，MNIST で学習させた混合生成器が生成しやすいことが理由である．一方，数字データと形状が大きく異なる文字 R や W は学習初期の段階では価値が低い．学習が進むにつれて徐々に価値が上昇し，収束時には他の画像と同程度の価値まで上昇した．これにより，DCM は生成器の表現空間と価値の高いデータの空間が乖離している場合にも生成データの価値を上昇させる能力があることが定量的に示された．

図 5: 学習過程の生成データの推移.

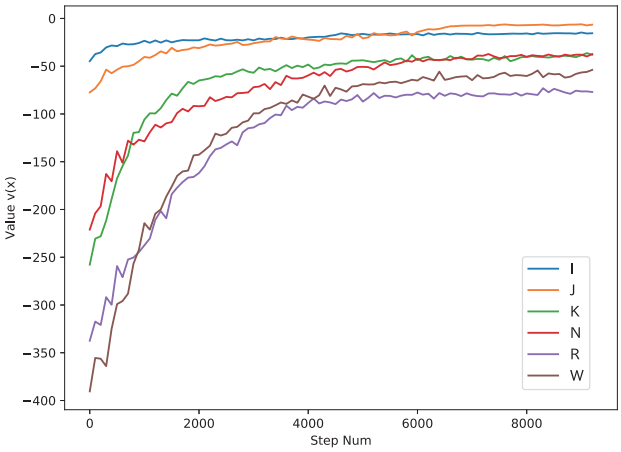


図 6: 学習による価値の推移

以上の評価より，DCM は学習データに含まれていないデータつまり新規テータを生成する能力があることと，生成データの価値を上昇させる能力があることがわかる．よって，DCM は新しく価値のあるデータつまり創造的なデータを生成する能力がある．

5. まとめと今後の課題

本研究では創造的なデータを生成する枠組みを提案した．このモデルは，学習データに含まれておらず，かつ価値の高いデータ，つまり創造的なデータの生成を行う. このモデルの特徴は混合生成器と自己生成データ学習の 2 点である．混合生成器は Feature matching とエントロピーによる正則化の損失関数によって，新規データの生成を可能にした．自己生成データ学習では，価値の高い生成データを学習データとして用いることで，より価値の高いデータを生成するための表現能力を獲得した．実験では MNIST データから，アルファベット画像を生成可能であることを確認した．

今後は絵画などの複雑なデータを用いて提案モデルの創造的データ生成能力を検証する．また，価値の算出を分類器の分類結果に変えるなど，より抽象的な価値関数で検証する．

謝辞

本研究は JST CREST(JPMJCR15E3) の支援を受けた．

参考文献

[Cohen 17] Cohen, G., Afshar, S., Tapson, J., and Schaik, van A.: EMNIST: an extension of MNIST to handwritten letters, *arXiv preprint arXiv:1702.05373* (2017)

[Dai 17] Dai, Z., Yang, Z., Yang, F., Cohen, W. W., and Salakhutdinov, R. R.: Good semi-supervised learning that requires a bad gan, in *Advances in Neural Information Processing Systems*, pp. 6510–6520 (2017)

[Elgammal 17] Elgammal, A., Liu, B., Elhoseiny, M., and Mazzone, M.: CAN: Creative adversarial networks, generating" art" by learning about styles and deviating from style norms, *arXiv preprint arXiv:1706.07068* (2017)

[Goodfellow 14] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y.: Generative adversarial nets, in *Advances in neural information processing systems*, pp. 2672–2680 (2014)

[Ha 17] Ha, D. and Eck, D.: A neural representation of sketch drawings, *arXiv preprint arXiv:1704.03477* (2017)

[Hawthorne 17] Hawthorne, C., Elsen, E., Song, J., Roberts, A., Simon, I., Raffel, C., Engel, J., Oore, S., and Eck, D.: Onsets and frames: Dual-objective piano transcription, *arXiv preprint arXiv:1710.11153* (2017)

[Kingma 13] Kingma, D. P. and Welling, M.: Auto-encoding variational bayes, *arXiv preprint arXiv:1312.6114* (2013)

[Kingma 18] Kingma, D. P. and Dhariwal, P.: Glow: Generative flow with invertible 1x1 convolutions, in *Advances in Neural Information Processing Systems*, pp. 10236–10245 (2018)

[Mordvintsev 15] Mordvintsev, A., Olah, C., and Tyka, M.: Inceptionism: Going deeper into neural networks, *Google Research Blog. Retrieved June*, Vol. 20, No. 14, p. 5 (2015)

[Salimans 16] Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., and Chen, X.: Improved techniques for training gans, in *Advances in Neural Information Processing Systems*, pp. 2234–2242 (2016)

[Salimans 17] Salimans, T., Karpathy, A., Chen, X., and Kingma, D. P.: PixelCNN++: A PixelCNN Implementation with Discretized Logistic Mixture Likelihood and Other Modifications, in *ICLR* (2017)

[Sternberg 99] Sternberg, R. J.: *Handbook of creativity*, Cambridge University Press (1999)

[Zhao 16] Zhao, J., Mathieu, M., and LeCun, Y.: Energy-based generative adversarial network, *arXiv preprint arXiv:1609.03126* (2016)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

デザイナーが持つ審美眼のモデル化

Modeling of prejudice eyes by designers

石川 隆一[*1] Ryuichi Ishikawa　和泉 興[*1] Kou Izumi　林 秀和[*1] Hidekazu Hayashi　福田 宏幸[*2] Hiroyuki Fukuda

[*1] 株式会社電通デジタル Dentsu Digital Inc　[*2] 株式会社電通 Dentsu Inc

Abstract: Research based on human aesthetic sense has been done many times, but that is the theme of this research. In this study, we conducted an experiment to model the aesthetic sense of designers using convolution neural network (CNN) and gradient boost decision tree (GBDT). Moreover, in the proposed method, accuracy was improved by adding colors and character data extracted from other than CNN as new features. The effectiveness of the proposed method was verified by multiple patterns and confirmed.

1. はじめに

第三次人工知能ブームの最中にある現在、AI が人間の仕事をどこまでこなす事が可能か、といった研究が盛んに行われている。近年では自動運転や画像生成など極めて高度なタスクも、AI がこなせるような時代に突入しつつある。

それでも人間が持つ感性は、曖昧な情報であり数値に置き換えることが難しい。特に絵画や小説など芸術の分野に関して言えば、未だに人間の能力に替わる性能を持つ AI が存在するとは言い難く、現在も心理や感性に基づく予測や生成の研究は模索が続いている。デジタル広告の世界においても、例えば無数の画像パターンからセレクトする作業はデザイナーなど美的感性を持つ限られた人でなければ難しいが、その煩雑な作業を軽減するために、AI による作業の代行が今求められている。

本研究は、バナー制作におけるデザイナーの持つ審美眼のモデル化を行う事で、膨大な画像の中から AI による分類予測を利用し、作業の効率化を図る事である。

2. 関連研究

2.1 CNN による特徴抽出

これまでも CNN を特徴抽出器として利用する Deep Convolutional Activation Feature(DeCAF)[Donahue 13]の研究は数多く存在する。

学習済みの VGG16 を用いた Image Captioning[Lin 14] も CNN を特徴抽出器として利用し、別の深層学習モデルに入力することで、画像から画像の説明文を生成することを可能にした。先行研究を踏まえ、本研究でも CNN を特徴抽出器として利用した。

2.2 CNN によるテクスチャーの抽出

近年の研究によると、Convolution Neural Network(CNN)では、入力近くの層で形状の情報を抽出しており、層を進むにつれテクスチャーを重視して特徴を抽出している事が解明されてきている。 [Geirhos 19] その特性を生かしたアルゴリズムとして、Style Transfer[Li 16]が考案された。Style Transfer は ImageNet にて学習済みの VGG16[Simonyan 15]から特徴抽出をしたテクスチャーを利用し、画像を別画像のスタイルへと変換する手法である。

先行研究によるテクスチャーを重視して特徴を抽出していることを踏まえ、特徴抽出器から抽出される特徴以外にも有効な情報を抽出可能だと考えた。

3. 提案手法

図1に提案手法の構造図を示す。本研究では ResNet50[He 16]を用いて全結合層(Global Average Pooling 層)から画像ベクトルの抽出を行った。抽出後の画像ベクトルに対し、オリジナルの画像から色彩と文字に関する情報を新たに抽出して加えることで、精度の向上を試みた。またレクタングル広告をデザイナーの美的感性により、良し悪しで振り分けられた 2 分類のデータを使用し学習・検証を行った。

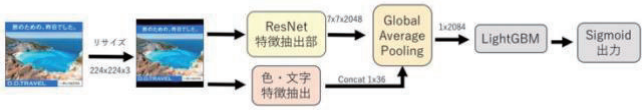


図1 提案手法の構造図

3.1 画像のリサイズ手法

本研究では、レクタングル広告(長方形の画像)を学習対象としているため、学習済みの CNN モデルを使用するにあたり画像を 224px×224px へリサイズする必要があった。しかし通常の手法でリサイズを行ってしまうと引き伸ばされてしまうため、アスペクト比を維持したまま、長辺が 224px になるよう縮小した。また正方形になるよう足りない部分をゼロパディングで埋めることでリサイズをおこなった。convolution 層の特性を考慮し、ゼロパディングを用いて穴埋めする手法が、最も CNN の特徴抽出に悪影響を及ばさないと考えた。

3.2 色彩情報

図 2 に手法の図を示す。デザイナーが持つ美的感性は CNN により抽出することが可能な特徴やエッジ以外にも、色の配置や色彩の情報が大きく関与すると考えた。

そのことを考慮した上で、機械学習のモデルで学習するにあたり、更なる特徴の追加を試みた。その手法として、リサイズ前

のオリジナル画像とリサイズ後の画像全体の RGB の各平均値を算出した。またリサイズ後の画像を縦と横で 3 分割し、その RGB 各平均値を算出し新たな特徴として追加した。結果として計 32 の特徴からなるデータを追加することで精度の向上を図った。

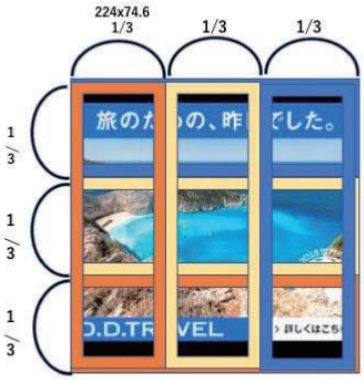


図 2 新たな色彩情報の取得図

3.3 文字に関する情報

色彩情報以外に、バナー広告におけるデザイナーの審美には、文字の情報も重要であると考えた。Optical Character Recognition (光学的文字認識)を用いて、取得したバウンディングボックスにより縦と横の長さから文字の占める面積を算出し、画像に対して文字が占める面積の割合と文字数・単語数・文章数(計4特徴)を新たな特徴として追加することで、CNN では抽出しえない文字量やデザイン性といった情報を再現し、精度向上を計った。

3.4 画像の特徴抽出

ImageNet にて学習済みの ResNet50 を用いて特徴抽出を行った。全結合層に当たる Global Average Pooling 層(2048 次元)から特徴を抽出し、抽出したデータを分類器の入力とすることで、機械学習のモデルでも学習することを可能にした。抽出した特徴に対し、前節で作成したデータ(計 36 特徴)を追加し、新たに特徴量 2084 の学習データを作成した。

3.5 Light GBM による分類予測

これまでの研究では画像の特徴を抽出したのちに Support Vector Machine(SVM)を用いて、分析にあたる研究が多く存在した。しかし本研究では、機械学習競技の優勝ソリューションで広く使用されている Light GBM を分類器として利用した。Light GBM は 2017 年に Microsoft によって発表されたアルゴリズムであり、決定木を元に考案された GBDT(Gradient Boosting Decision Tree)のモデルである。欠損やカテゴリー変数が存在していても学習することが可能であり、外れ値にも強く学習にかかる時間が短いことが[Ke 17]によって示されている。

4. 実験

4.1 データ

本実験では、実際にバナー制作に携わるデザイナーに協力いただき、良し悪し(良:2172 枚・悪:2828 枚)で分けられたデータを正解ラベルとして使用した。実験に伴い特徴追加前のデータ(5000×2048)と特徴追加後のデータ(5000×2084)を使用し、提案手法の有効性を測った。またデータの 20%を検証データに当てた。

4.2 評価

評価指標として Accuracy と Receiver operating characteristic (ROC)解析を実施し, Area under curve (AUC)を算出する事で精度を測った。ROC-AUC は最大値が1.0であり、値が1.0 に近いほど分類モデルの性能が高いことを意味する。ROC 曲線及び ROC-AUC の結果を図 3 に示す。

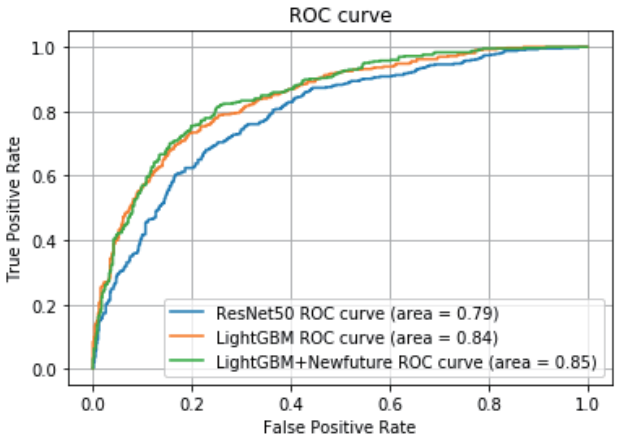


図 3 モデル別 ROC 曲線

4.3 モデル

Light GBM 学習時のハイパーパラメーターの条件を表1に示す。損失関数には 2 分類に伴い binary log loss を使用した。また提案手法の有効性を確認するために、ImageNet[Deng 09]にて学習済みの ResNet50 を利用した。さらに CNN モデルの精度を最大限に上げるため、転移学習を行い全結合層のみを学習し直す事で精度の比較を実施した。

表1 Light GBM ハイパーパラメーター

パラメーター名	数値
learning_rate	0.005
num_leaves	36
min_data_in_leaf	5
min_sum_hessian_in_leaf	20
max_depth	9
n_estimators	10000

4.4 検証結果

検証結果を表 2・図 4 に示す。検証結果として表 2 から分かるように提案手法により特徴を加えることで loss・Accuracy・AUC の全てにおいて、多少の精度の向上が認められる。また Light GBM による特徴の重要性の上位 15 特徴に色彩や文字数と文

字範囲の情報が反映されていることも検証結果から認められる。さらに転移学習を用いたとしても、ResNet50 単体での分類のタスクを行うより、特徴抽出後に Light GBM にかけると精度の向上が認められた。

表 2 実験結果

モデル	loss	Accuracy	AUC
ResNet50	0.548	0.728	0.79
Light GBM	0.486	0.772	0.84
Light GBM+特徴追加	0.478	0.780	0.85

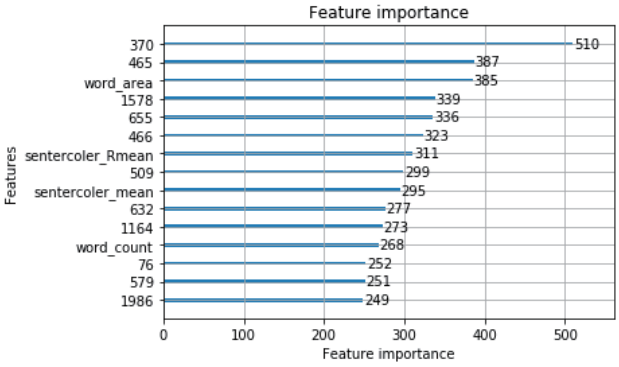


図4 Light GBM による特徴重要度

※数字カラムは CNN から抽出した画像ベクトル

5. おわりに

本研究では、デザイナーによる美的感性によって振り分けられたバナー広告を使用し、新たな特徴抽出を行うことで精度の向上を試みた。実験結果として、色彩と文字データを追加することにより、多少精度が向上することが確認できた。また Accuracy 78%の精度があるモデルの再現に成功したことで、デザイナーの仕事の工数削減が期待できる。

オリジナルの画像から新たな特徴を抽出し追加する手法が様々なタスクで有効であれば、深層学習の更なる進化につながるのではないだろうかと考えている。今後、本研究により確認できた画像から抽出した特徴の有効性が、デザイナーの持つ美的感性以外のタスクにおいても有効であるか検証したいと考える。

参考文献

[Geirhos 19] Geirhos, Robert, et al. "ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness." arXiv preprint arXiv:1811.12231 (2018).

[Li 16] Li, Chuan, and Michael Wand. "Precomputed real-time texture synthesis with markovian generative adversarial networks." European Conference on Computer Vision. Springer, Cham, 2016

[Simonyan 15] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In ICLR, 2015.

[Donahue 13] Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., and Darrell, T. Decaf: A deep convolutional activation feature for generic visual recognition. CoRR, abs/1310.1531, 2013

[Lin 14] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, and C. L. Zitnick. Microsoft coco: Com- ´mon objects in context. arXiv preprint arXiv:1405.0312, 2014.

[He 16] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in CVPR, 2016.

[Ke 17] Ke,G.Meng,Q.,Finley,T.,Wang,T.,Chen,W.,Ma,W.,Ye,Q.,Liu,T.Y.: LightGBM:A highly efficient gradient boosting decision tree. In: Advances in Neural Information Processing Systems. 2017

[Deng 09] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In CVPR09, 2009

Organized Session | Organized Session | [OS] OS-14

[3F3-OS-14a] 人狼知能と不完全情報ゲーム(1)

稲葉 通将（広島市立大学）、片上 大輔（東京工芸大学）、狩野 芳伸（静岡大学）、大槻 恭士（山形大学）

Thu. Jun 6, 2019 1:50 PM - 3:30 PM Room F (302B Medium meeting room)

[3F3-OS-14a-01] Analyzing Gestures in Real-World Werewolf Game

〇Shutarou Takayama[1], Hirotaka Osawa[1] （1. University of Tsukuba）

1:50 PM - 2:10 PM

[3F3-OS-14a-02] “Goodness” analysis of a Werewolf Game’s player based on a biological signal

〇Hirotaka Yamamoto[1], Nagisa Munekata[1] （1. Kyoto Sangyo University）

2:10 PM - 2:30 PM

[3F3-OS-14a-03] Simulation of Strategic Evolution in 5-player Werewolf

〇Atsushi Takeda[1], Fujio Toriumi[1] （1. The University of Tokyo）

2:30 PM - 2:50 PM

[3F3-OS-14a-04] Simulation Analysis based on Behavioral Experiment of Cooperative Pattern Task

〇Norifumi Watanabe[1], Kota Itoda[2] （1. Musashino University, 2. Keio University）

2:50 PM - 3:10 PM

[3F3-OS-14a-05] Analyis of Strategies in Werewolf Game by 3 Players Considering Voting Behaviour of Other Players

〇Hinako Tamai[1] （1. Nara Women's University）

3:10 PM - 3:30 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

実世界人狼ゲームのジェスチャーの分析

Analyzing Gestures in Real-World Werewolf Game

高山　周太郎[*1]　　大澤　博隆[*2]
Shutarou Takayama　　Hirotaka Osawa

[*1] 筑波大学 理工学群 工学システム学類
Collage of Engineering Systems, School of Science and Engineering, University of Tsukuba

[*2] 筑波大学 システム情報系
Faculty of Engineering, Information and Systems, University of Tsukuba

Nowadays, werewolf game has been studied in the context of communication games and incomplete information games. Werewolf game is expected to be used as a training method of communication because it has communication elements such as persuasion both in verbal and non-verbal aspects. However, there is little quantitative analysis for non-verbal communication factors in werewolf game which contribute for persuasion. The authors expect that it is necessary to investigate the influence of non-verbal information for clarify the mechanism of communication in werewolf game. In this study, the author gathered several human gestures and investigated the influence on the result of the game. The several parameters are arms degree, hands movement distance etc. The results suggest several persuasive non-verbal gestures in werewolf game. For example, arms degree shows how much influence on discussion in the game. It is suggested that initiative of betrayer or seer influences the result of werewolf game.

1. はじめに

人狼ゲームとは、プレイヤーの中から割り当てられた「人狼」を、コミュニケーションを通じて発見するゲームである。平和な村の中に、村人に扮した人狼が紛れ込んだので、人狼を見破り排除しなければならない―というストーリーが人狼ゲームの背景にある基本設定である。人狼を割り当てられたプレイヤーはいかに他のプレイヤーに対して村人だと思わせるか、他のプレイヤーはいかに人狼を見破るか、でゲームの勝敗が決定する。

近年、人狼ゲームが研究対象として注目されており、オンライン上でのチャットログなどの言語情報を主な対象として人狼ゲームを通じてより人間に近い思考のエージェントの研究がなされている [梶尾 14]。例えば、梶原らは大澤が作成した人狼ゲームの発話プロトコル[大澤 12]を使い、強化学習によって人狼ゲームでの最適戦略を抽出した[梶尾 14]。

しかし、コミュニケーションにおける非言語情報の重要性は多くの研究で指摘されており[高木 06]、実世界での人狼ゲームでも人間同士の非言語コミュニケーションがゲームの進行に関わっていると考えられる。酒井らによる、アバターを用いた通話による人狼システムでは、プレイヤーの議論があまり盛り上がらない事がわかっている[酒井 16]。アバターによって人の見た目が同質化され、非言語から情報を集められなくなったためだと考えられる。人狼ゲーム中の非言語コミュニケーションがプレイヤーにどの程度影響を与えるのか、定量的に分析する必要があると考える。

そこで、本研究では高山による実世界人狼ゲームの非言語情報解析システム[高山 19]により取得した姿勢情報を分析し、役職や個人によって特徴的に見られる動作についていくつかの指標を用いて、ゲーム結果との関係性を調査した。

2. 人狼ゲームについて

2.1 人狼ゲームのルール

人狼ゲームのルールは参加人数などによって細かい違いはあるものの、大筋のルールは共通している。本項では人狼ゲームの一般的なルールについて記述する。

1 章で述べたとおり、人狼ゲームはコミュニケーションによってプレイヤーの中から人狼を見つけ出すゲームである。プレイヤーは一般的に、大きく人狼陣営と村人陣営に分けられる。村人陣営はプレイヤーの中の人狼をすべて排除することを目指す。人狼陣営は人間を排除して人狼と人間の人数が同数以下になることを目指す。

ゲームは昼、投票、夜という 3 つのフェーズを繰り返して進行する。昼フェーズはプレイヤー同士の自由な議論によって人狼と思われる人物を探し出す。昼の議論には時間制限が存在する。制限時間がきたら、投票フェーズに移る。投票フェーズでは各プレイヤーが排除したいプレイヤーに

連絡先：高山周太郎
hailabsec@iit.tsukuba.ac.jp
筑波大学 理工学群 工学システム学類
ヒューマンエージェントインタラクション研究室

投票する。投票により最も多くの票を集めたプレイヤーはゲームから除外され、その後のゲームには一切関与できなくなる。この投票によるプレイヤーの除外を処刑と呼ぶ。夜フェーズでは各役職に与えられた特殊能力が作動する。例えば人狼は、プレイヤーの中から襲撃する相手を一人選び、そのプレイヤーをゲームから除外する。

(1) 5 人人狼に関して

本研究では、プレイヤーが5人の人狼ゲーム(5人人狼)を対象とする。本研究での 5 人人狼の役職配分は、占い師・人狼・狂人・村人×2、となっている。各役職に関しては2.2 項で述べる。

5 人人狼では最長でも 2 ターン目の処刑が完了した時点でゲームの勝敗が決定する。人狼ゲームの終了条件が「人狼がいなくなる」または「人間と人狼が同数以下になる」であるため、2 ターン目の処刑完了時点でプレイヤーは 2 人まで減っている。残ったプレイヤーの中に人狼がいなければ前者の終了条件が、人狼がいれば後者の終了条件が満たされるためである。

5 人人狼では、最短では 1 ターン目の処刑時点で終了する。この時点でゲームが終了した場合は、人狼陣営の敗北となる。この場合、人狼が他のプレイヤーを騙せず、最初の議論で見抜かれていたと言えるため、2 ターン目で負ける場合よりも大差で負けていたと言える。

2.2 人狼ゲームの役職

人狼ゲームでは、固有の能力を持つ役職が各プレイヤーに与えられる。人狼ゲームでの役職は多くの種類が存在するが、本項では、本研究での 5 人人狼で登場するものについて述べる。

各役職の役割は以下の通りである。

- 占い師：村人陣営。夜になるとプレイヤーを一人選び、そのプレイヤーが人狼かどうか自分だけが知ることができる。
- 人狼：人狼陣営。夜になると一人選び、そのプレイヤーをゲームから除外する。
- 狂人：人狼陣営に所属する人間である。人狼を勝たせるために動く。占い師からは人間であると判断される。誰が人狼なのかは認識できない。
- 村人：村人陣営に所属する。特に能力を持たない。

3. 解析手法

本研究では、人狼ゲーム中の非言語情報を定量的に解析するための指標として、プレイヤーが表出する特徴的な動作をいくつか挙げ、定式化した。動画の 1 フレームごとに各指標を測定し、1 ゲームごとにその平均値を取る。測定の対象となるフレームは、各ゲームの昼の議論時間に該当する箇所を対象とする。本研究で撮影した動画は 1 フレームが約 0.033 秒に相当する。

以下に本研究で用いた指標とその概要を示す。

・arms_degree

両腕でどの程度体を隠そうとしているのかという指標。右手の座標から右肘の座標を引いてできるベクトルを、右腕のベクトル$\boldsymbol{l}$と置く。左に関しても同様に、左腕のベクトル$\boldsymbol{r}$と置くと置く。arms_degree の値θは次の(1)式で与えられる。

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\|l\|}\begin{bmatrix} l_x & l_y \\ -l_y & l_x \end{bmatrix} \cdot r$$

$$\theta = \tan^{-1}\frac{y}{x} \quad (1)$$

1 ゲームごとに平均値を取り、ゲーム内でどの程度体を隠そうとしていたのか測る。正に大きくなるほど腕を開き、負に大きくなるほど腕を閉じていると言える。

・directing_upper_the_left_arm

左腕が上向きになっているかどうかの指標。左腕のベクトルが上向きならば 1、下向きならば 0 を返す。1 ゲームごとに足し合わせて、フレーム数で標準化することで比較する。

・directing_upper_the_right_arm

directing_upper_the_left_arm を右腕に適用したもの

・head_moving

鼻の座標の前フレームからの移動距離。最初のフレームはカウントしない。フレーム数-1 で標準化する。

・left_hand_moving

head_moving を鼻ではなく左手の座標に適用したもの。

・right_hand_moving

head_moving を鼻ではなく右手の座標に適用したもの。

4. プレイの計測

本研究の計測には、高山らが作成した実世界人狼の非言語情報を取得するためのシステムを使用した[高山 19]。このシステムは、人狼をプレイする動画から、プレイヤーの関節の座標情報が測定できる。このシステムから実際に習得した情報を次の図 1 に示す。

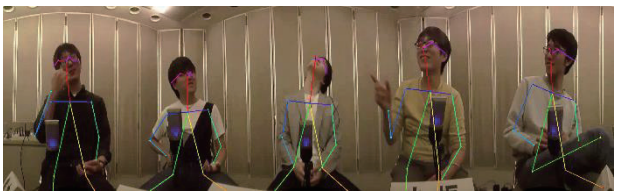


図 1: 取得した関節の座標情報イメージ

2018 年 11 月、7 人の男性参加者（大学生及び大学院生）を 5 人ずつ交代でゲームを全 37 ゲーム実施した。撮影の参加者は全員が右利きであった。本研究の撮影では、プレイヤーの左手に発汗を測定するセンサーを付けている。プ

イヤーが、心理的に左手のほうが動かしづらい可能性は考えられる。このセンサーで取得したデータは、本研究では使用していない。

5. 結果

5.1 ゲーム結果

全 37 ゲームで各陣営が勝った回数を次の表 1 に示す。

表 1: 全ゲーム結果

陣営	勝利回数
人狼	16
村人（合算）	21
村人（1 ターン）	12
村人（2 ターン）	9

5 人でプレイする人狼ゲームは通常 2 ターンで終了するが、村人陣営は人狼を処刑できればゲームに勝利できるため、1 ターンでゲームが終了する場合もある。表 1 の「村人(合算)」は、終了ターン数を考慮しない場合の勝利回数。「村人(1 ターン)」「村人(2 ターン)」は、「村人(合算)」の内それぞれ 1 ターンで終了したもの、2 ターンで終了したものを指す。

5.2 解析結果

4.3 項で述べた指標のデータを役職ごとに分類し、さらにそれを 2 群に分けて t 検定を行った。2 群への分け方は

- ゲームの勝敗:勝利条件・敗北条件に分ける。ゲーム全体の議論の時間に関して値を取得する。
- ゲームのターン数:ゲームが 1 ターンで終了する最短条件・2 ターンで終了する最長条件に分ける。1 ターン目の議論の時間に関して値を取得する。

の 2 パターンを試みた。

以上のように分類したデータを、指標ごとに t 検定を行いその結果の p 値を表 2、3 に示す。

表 2: 役職ごとにゲームの勝敗で検定($p^{*}<0.05$, $p^{+}<0.1$)

	人狼	村人	占い師	狂人
arms_degree	0.74830	0.06908+	0.25642	0.05773+
directing_upper_the_left_arm	0.48960	0.81931	0.41968	0.89675
directing_upper_the_right_arm	0.51475	0.61446	0.01671*	0.46831
head_moving	0.61960	0.00203*	0.03788*	0.74107
left_hand_moving	0.14738	0.30183	0.17098	0.22430
right_hand_moving	0.51583	0.28303	0.09409+	0.56704

表 3: 役職ごとにゲームのターン数で検定($p^{*}<0.05$, $p^{+}<0.1$)

	人狼	村人	占い師	狂人
arms_degree	0.70481	0.45325	0.00392*	0.00097*
directing_upper_the_left_arm	0.75753	0.23767	0.30432	0.07694+
directing_upper_the_right_arm	0.59014	0.89732	0.00194*	0.29950
head_moving	0.80736	0.03805*	0.18326	0.67003
left_hand_moving	0.14240	0.51271	0.62096	0.79854
right_hand_moving	0.60686	0.52928	0.51365	0.66872

6. 考察

6.1 狂人と占い師の関係性に関する考察

本項では、狂人と占い師の arms_degree に着目する。Cashdan では、腕の開き具合とその場のコミュニケーションでのリーダーシップとの関係性を指摘している[Cashdan 98]。arms_degree が大きいほど、人狼ゲームの議論の場において強い影響力を持っていると考えられる。

表 3 の arms_degree から、占い師と狂人のときに着目する。表 4 に、狂人・占い師の arms_degree の、また、各群での平均値と標準偏差を示し、図 2、3 に狂人、占い師それぞれの arms_degree の箱ひげ図を示す。

表 4:ターン数毎の arms_degree の平均値・標準偏差

	占い師 最短	占い師 最長	狂人 最短	狂人 最長
平均値	-0.129	-0.458	-0.599	-0.215
標準偏差	0.264	0.362	0.265	0.356

図 2 から、1 ターンでゲームが終了した場合、狂人の arms_degree が有意に小さくなり、ことがわかる。ここから、人狼陣営である狂人が、リーダーシップを持てなかったため大差で負けてしまい、それが arms_degree に表れたと考えられる。

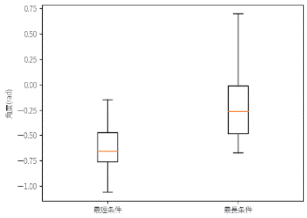


図 2:ターン数で比較した狂人の arms_degree

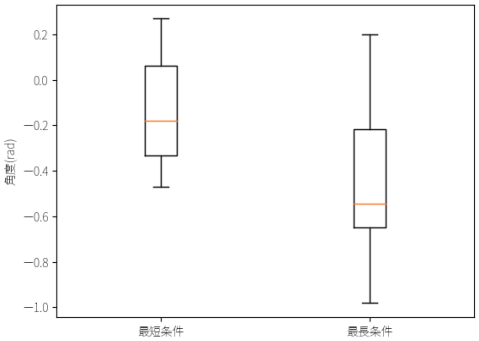


図 3:ターン数で比較した占い師の arms_degree

反対に、図 3 からは 1 ターンで終了した時、すなわち占い師が人狼陣営に勝った時、腕を開く傾向にあることがわかる。占い師がリーダーシップを持ち、ゲームを村人陣営が有利な方へ導いたためだと考えられる。

6.2 占い師に関する考察

本項では、占い師の directing_upper_the_right_arm に着目する。表 6 に勝利条件、敗北条件、最短条件、最長条件それぞれにおける平均値と標準偏差を示す。また、図 4、5 に勝敗を比較したとき、ターン数で比較した時それぞれの場合の箱ひげ図を示す。

表 6: 各条件における占い師の directing_upper_the_right_arm の平均値・標準偏差

	最短	最長	勝利	敗北
平均値	0.493	0.206	0.387	0.167
標準偏差	0.230	0.239	0.265	0.260

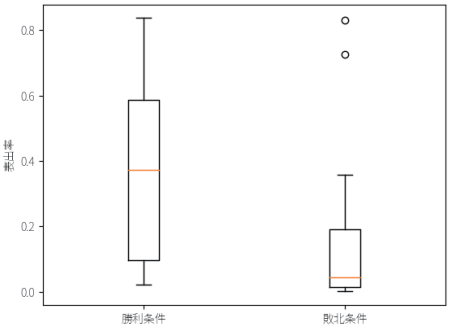


図 4: 占い師の勝敗ごとの directing_upper_the_right_arm

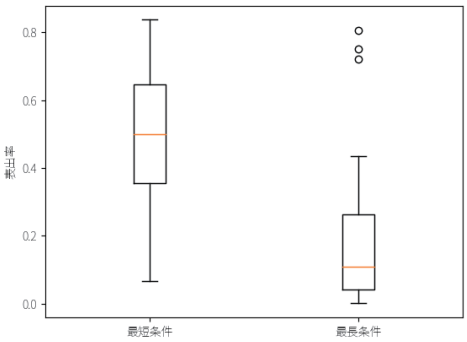


図 5: 占い師のターン数ごとの directing_upper_the_right_arm

占い師がゲームに勝つ時、directing_upper_the_right_arm が大きくなることがわかる。自分に注意を引くために積極的に右腕を動かしていることが考えられる。村人陣営が勝つには、占い師が手を挙げ、積極的に議論に参加するのが良いとわかる。

7. まとめ

本研究では実世界人狼ゲームに見られるジェスチャーに着目し、議論との関連性を調査した。本研究で提案した指標で、1 ターン目の有利な議論の導き方や、占い師の議論への参加の積極性を評価できることがわかった。

謝辞

本研究は JSPS 科研費 JP26118006, JP16H02928, JP18KT0029 の助成を受けたものです。心より感謝致します。

参考文献

[Cashdan 98] Elizabeth Cashdan: “Smiles, Speech, and Body Posture: How Women and Men Display Sociometric Status and Power”, Journal of Nonverbal Behavior, 1998.

[大澤 13] 大澤博隆,“コミュニケーションゲーム「人狼」におけるエージェント同士の会話プロトコルのモデル化” HAI シンポジウム 2013, 2013.

[梶原 14] 梶原健吾,鳥海不二夫,大橋弘忠,大澤博隆,片上大輔,稲葉通将 and 篠田孝祐, “強化学習を用いた人狼における最適戦略の抽出” , 情報処理学会第 76 回全国大会講演論文集, 2014.

[酒井 16]酒井久志,片上大輔 and 大澤博隆“FACERIG での人狼での無意識動作の影響”, HAI シンポジウム 2016,2016.

[高木 06] 高木幸子, “コミュニケーションにおける表情及び身体動作の役割”,2006.

[丹野 15]丹野昭宏 and 児玉健“人狼ゲーム経験による嘘の手がかり信念の差異”,パーソナリティ研究 24(1),2015.

[高山 19]高山周太郎 and 大澤博隆 “実世界人狼ゲームの非言語情報分析システムの設計”, HAI シンポジウム 2019,2019.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

生体信号に基づく人狼プレイヤの「強さ」分析の試み

"Goodness" analysis of a Werewolf Game's player based on a biological signal

山本　浩隆　　棟方　渚
Hirotaka Yamamoto　Nagisa Munekata

京都産業大学
Kyoto Sangyo Univercity

We attempted to analyse "goodness" of a game player by measuring a biological signal of each player during a five-player werewolf game. First, we focused on some periods as important game events since we observed simultaneous increase of electro dermal activities of multi players at these periods. Second, we compared electro dermal activities between high winning players and others at the important events. As a result, an unique tendency was confirmed in only electro dermal activities of the good players. This result would be expected to contribute toward designing of strategy in Werewolf AI.

1.　はじめに

人狼ゲームは，プレイヤ同士の対話によって，属している陣営を勝利に導いていくコミュニケーション型ゲームである．プレイヤは自身の役職以外の情報は一切もたないため，会話が人狼を導き出すための唯一の手がかりとなる．限られた情報の中で早く自身の仲間を見つけ説得し，相手を欺き，情報を引き出すことがゲームの醍醐味である．人狼ゲームは本来，対面で行うゲームが一般的であるが，日本では古くからWeb上でBBS人狼と呼ばれる掲示板型の人狼ゲームが存在する．2013年頃からテレビ番組で有名人による人狼ゲームのプレイ映像が放送されたり，動画サイトではプロ将棋棋士による人狼ゲームが放送された．人狼ゲームを演劇として見せ，エンターテイメントとして「魅せる人狼」も盛んに行われている[1].

2015年に鳥海らが立ち上げた「人狼知能プロジェクト」は「人間と自然なコミュニケーションを取りながら人狼をプレイできるエージェントの構築」を目標としており，実現にはエージェント技術や人工知能，自然言語処理，Human-AgentInteractionなど様々な技術が必要とされている．人狼AI構築の一つのアプローチとして，人間の熟達者プレイヤの思考過程を分析を行うといった試みがある[2]．思考過程分析には，発話分析や自然言語処理，生体情報を用いて行われている．本研究は，人間の偽りのない反応を取得できる生体信号に着目し，興奮や集中，思考の移り変わりの影響を受ける皮膚電気活動を用いて，人狼ゲームの熟達者プレイヤの思考過程の分析を試みる．

2.　関連研究

2.1　皮膚電気活動

皮膚電気活動は，精神性発汗を電気的にとらえたものである．人の手掌や足底は，緊張や動揺などの心的興奮によって発汗を生じる．これらの発汗は，自覚できないほど微量のものから，手掌が湿ってしまうほど大量のものまで様々である．皮膚電気活動は手掌に1対の電極を装着し，その電極間に微弱な電流を流している間の発汗によって生じる測定部の抵抗値の変化から測定できる．皮膚電気活動には複数の測定単位系があり，人間の情動状態を評価する方法として用いられている[3, 4]．その中で，本研究ではSCL(Skin Conductance Level)を用いた．本研究で用いるSCLは，緩徐な変動で観察される信号であり，ユーザの精神的不安や安静状態の評価として用いる場合，他の皮膚電気活動より有効であると示されている[5].

連絡先:　棟方　渚，京都産業大学情報理工学部，munekata@cc.kyoto-su.ac.jp

2.2　皮膚電気活動と人狼ゲームの関連

精神性発汗は生体が緊急時に対処する行動，闘争-逃走反応と密接に関連するといわれており，闘争時の手掌の発汗は道具をしっかりと把持するという目的がある[6]．人狼ゲームでは，プレイヤが勝敗を左右する重要な場面に直面した場合，闘争-逃走反応である手掌発汗があらわれ，急激な皮膚電気活動の上昇があると考える．人狼プレイヤが重要であると考える場面を皮膚電気活動によって判別することが可能になれば，人狼プレイヤの「強さ」の定義や戦略の理解，それに基づいた人狼AI構築における重み付けなどが可能となると考えられる．

2.3　皮膚電気活動を用いた人狼プレイヤの分析

御手洗らは，人狼ゲームプレイヤの情動状態を対象として，皮膚電気活動を用いた人間のコミュニケーションにおける情動変化を解析した．プレイヤの役職や各陣営の状態がSCLに影響をあたえるということが示唆されている．また，人狼陣営(狂人や人狼の総称)は，村人陣営(村人や占い師の総称)よりもSCL平均値が高いことが明らかになった．この結果について，人狼陣営は必然的に自身の役職を偽るなどの虚偽の発言が求められるなど，精神的な負荷の高い状態がゲーム中に持続していたことが要因の一つであると報告されている[7].

風間らは人狼のゲーム時に，誰のどのような発言がプレイヤに影響を与え，ゲームを動かしているのかを調べるために，情動反応のマーカとしてSCR(Skin Conductance Response)を測定し，ゲームの展開を左右する発言の種類を分析した．ゲーム時のSCRの上昇がみられたプレイヤの発話を抽出した結果，人狼プレイヤは他プレイヤに疑われた際に，嘘をつくリスクを減らす，かつ目立つことを避けるために，理由を付与するような反論をせずに話題を逸らそうとする発言をする傾向があると推測した．また，村人プレイヤの思考過程とSCRの発現のタイミングから，村人プレイヤは他プレイヤへの疑いに，人狼プレイヤは自身への疑いに対しそれぞれ不利益を感じ，発言内容に影響を与えていることが推測された[8].

3. 実験

3.1 人狼ゲームのルール

人狼ゲームはランダムに役職が与えられる．役職の中には特殊な能力を持つ役職も存在する．今回は5人で人狼ゲームを行うため役職は4つである．

- 占い師(預言者)：他のプレイヤの役職をみる
- 人狼：他のプレイヤを殺害する
- 狂人(裏切り者)：村人として数えられ，人狼に加担する村人である
- 村人：特殊な能力は持たない

それぞれの役職の内訳は占い師1人，人狼1人，狂人1人，村人2人とする．

ゲームは常にプレイヤ以外のゲームマスター(以下：GM)が進行する．人狼ゲームは様々な場面が存在するが，本研究では以下の4つに分類し，解析・議論を行う．

1. 役職を確認
2. 夜のフェーズ(初日の殺害は不可)
3. 会議(4分間に固定する)
4. 投票

3．の会議において人狼が生き残った場合，2．から4．をもう1度繰り返す．村人陣営である村人，占い師の勝利条件は人狼の処刑であり，人狼の勝利条件は村人の数を人狼以下にする．狂人は人狼の勝利が勝利条件である．

3.2 実験概要

被験者に５人人狼を15試行プレイさせ，それぞれのSCLを測定した．実験に参加した被験者は6名であり，この被験者は普段から5人人狼をプレイしており，熟達した人間プレイヤとして位置付けることができる．被験者6名の中から5名を選出し一部交代でゲームに参加してもらった．各被験者は15回の試行のうち少なくとも10回は試行に参加した．連続でゲームを行うことによる疲労の影響を考慮し，5試行ごとに約20分間の休憩を設けた．

前述した通り，人狼は精神的な負荷の高いゲームであることから，個々のプレイヤのSCLは常に上昇と下降を繰り返していると考えられるが，複数のプレイヤのSCLの同期的な上昇がみられた場合，その周辺にゲームの勝敗を左右する重要なイベントが存在すると考えらる．そこで，全プレイヤのSCLデータに基づき，ゲーム中の重要なイベントを抽出し，「強い」プレイヤと「弱い」プレイヤ間でのSCLの振る舞いを比較することとした．ここでいう「強い」プレイヤは，10から15回の試行のうち，勝率が高いプレイヤとし，勝率の低い者を「弱い」プレイヤとした．

3.3 本研究で使用したSCL計測システム

SCLの計測装置はAffectiva社のQ Sensorを参考に実装したものを使用した．計測システムはPython3.5を用いて実装し，Raspberry Pi 3 ModelB上で動作させた．システムはA/D変換器を通して，サンプリング周期20HzでSCLを取得した[7]．また，Raspberry Pi 3 ModelB上で記録が正常に行われているのかを確認するためにPython3.6を用いて，PCとRaspberry Pi 3 ModelBの通信を同時に行った．実験では，プレイヤのジェスチャ等の振る舞いを妨げることの無いように計測装置を配置し，電極やリード線をプレイヤの身体に固定した．ゲームの進行やプレイヤの言動を記録するために，各プレイヤの前方にそれぞれマイクを配置し，顔が認識できるよう数台のビデオカメラを用いて実行した．

4. 実験結果

4.1 SCLによるイベント抽出

複数のプレイヤのSCLの同期的な上昇がみられた箇所に着目し，その周辺のプロトコル分析を行なった．抽出されたイベントは主に各プレイヤの発話を基にタグ付けを行った．得られた結果を表1に示す．そのイベント開始と終了に関する認識は，各プレイヤによって多少の時間差があることを考慮し，SCLの同期的上昇は誤差2-3秒程度とし関連づけた．

表 1: 抽出されたイベントとその数

イベント	数
CO	9
疑い	5
殺害・投票・処刑	3
合計	17

全試行の分析結果から，計17箇所で複数のプレイヤのSCLの同期的な上昇がみられた．抽出された各イベント（表1）について以下で説明する．

抽出されたイベント「CO」は，自身の役職を他のプレイヤに伝えるなどのカミングアウト(以下：CO)を意味し，それに起因するイベントが9箇所みられた．COは，あるプレイヤが呼びかけをするなどして開始され，単数もしくは複数のプレイヤが役職を占い師と自称し占い結果を提示（虚偽を含む）するまでをCOと定義した．つまり，COは主に占い結果に関する会話となる．この会話の例（原文ママ）として,「A君グレー，人間ではある（狂人の可能性がある）」,「俺はBに入れますね（投票する），最初から黒（人狼の意味）って言ってたんで」などの占い結果に基づくものが挙げられる．

抽出されたイベント「疑い」は，他プレイヤへの疑い（他プレイヤからの疑い）に関する会話に基づいたイベントであり，計5箇所みられた．この会話の例（原文ママ）として,「ないないそれは．いや，まぁ人狼かもしれないけど(他プレイヤを指差しながら)」,「C君人狼だからね～」など特定のプレイヤ，特定の役職について言及したものが挙げられる．

抽出されたイベント「殺害・投票・処刑」は，GMの「Dさんが殺害されました」「では，投票してください．いっせーのーで(それぞれが他プレイヤを指差す)」などの宣言や投票を行う場面が挙げられ，計3箇所みられた．

抽出された17箇所のイベントのうち，15箇所がゲームの初日のフェーズに行われた．加えて，その12箇所は，初日の会議において行われたイベントであったため，以降，初日の会議に着目し，プレイヤの振る舞いを分析することとした．

4.2 勝率による「強い」プレイヤの選定

人狼ゲームの「強さ」について定義が難しく，先行研究においても議論がなされていない．そこで，本研究では15回の試行における勝率を基に，強いプレイヤと弱いプレイヤを選定することにした．各プレイヤの勝率を以下（表2）に示す．

表 2: 各プレイヤの勝利数と試行回数

プレイヤ	A	B	C	D	E	F
勝率	70 %	67 %	40%	70%	40%	40%
試行数	10	15	15	10	10	15

表2から，試行回数が多く勝率の高いプレイヤBを「強い」プレイヤ，試行回数が多く勝率の低いプレイヤCとFを「弱い」プレイヤとする．全15試行おける人狼陣営の勝率は33%村人勝率の67%であった．強いプレイヤと弱いプレイヤの各陣営に属した回数を表3に示す．

表 3: 各プレイヤの勝利数と試行回数

	プレイヤB	プレイヤC	プレイヤF
占い師	5	2	2
村人	4	5	7
人狼	1	4	3
狂人	5	4	3

表3の結果から，強いプレイヤ（B）と弱いプレイヤ（C,F）とで，各陣営に属した回数に大きな偏りはみられなかった．

解析は，ノイズの混入（電極を触る，電極が外れる）がみられたものは対象から除外した（有効なデータのみを使用した）．4.1章の結果から，抽出されたイベント17箇所のうち12箇所がゲーム初日の会議で確認された．勝率で選定した強いプレイヤ，弱いプレイヤについて，ゲーム初日の会議に着目し，観察とSCLの分析結果から両プレイヤの振る舞いを調査する．次章では，両プレイヤの初日の会議におけるSCLの時系列データの振る舞いについて分析した．

4.3 初日の会議のSCLの振る舞い

前章で選定した強いプレイヤ（B）と弱いプレイヤ（F）について，各試行の会議（約4分間）のSCLを30秒毎に平均値を求めた結果を図1，2に示す．各データは計10試行分となった（有効なデータのみを使用）．図1，2の結果から，個人内のデータであっても，それぞれの試行回のSCLにばらつきがあることが理解できる．ゲームは全試行同日に行われており，各試行において身体状態に大きな差はみられないと考えられるが，役職やゲーム状況によって大きな影響を受けると考えられる．次に，より全体的な傾向を確認するため，図1,2の結果の平均を図3，4に示す．図3，4の結果から，強いプレイヤ（B）は会議の始まりから終わりにかけて，徐々にSCLが減少するといった一定の傾向がみられ，弱いプレイヤ（F）のSCLは時系列的な変化がみられなかったことが示された．もう一方の弱いプレイヤ（C）においても同様の結果となった．この結果をふまえ，次章で考察を行う．

5. 考察

実験結果から，熟達した強いプレイヤ（B）のSCLの振る舞いは，会議の始まりで相対的に高い値を示したことから，強いプレイヤは会議開始1分程度で最も心的な興奮がみられたことを意味する．この周辺で行われたイベントは，主にCOであった．加えて，COは複数のプレイヤのSCLの同期的な上昇が最もみられた箇所であり，人狼ゲームにおいて最も重要なイベントと考えることができる．そこで，COにおける強いプレイヤ

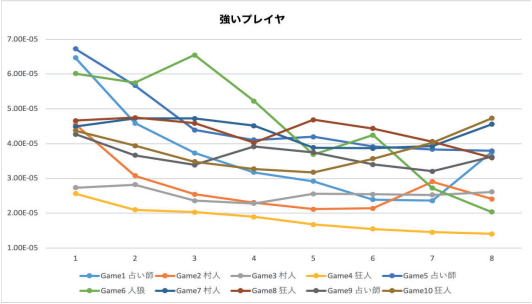


図 1: 各試行の会議中の強いプレイヤ

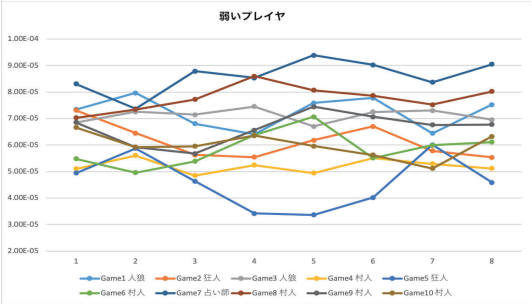


図 2: 各試行の会議中の弱いプレイヤ

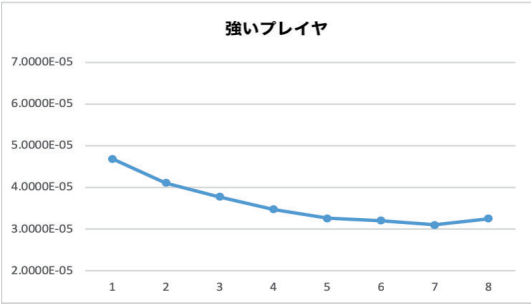


図 3: 会議時：強いプレイヤのSCL（全試行）

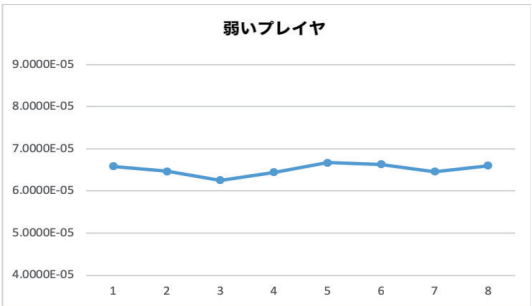


図 4: 会議時：弱いプレイヤSCL（全試行）

と弱いプレイヤの言動に着目し調査した．具体的には，CO時のプレイヤの「挙手の有無」つまり，占い師として名乗りをあげたのかどうかを調べ（挙手有り/挙手無し），その際のSCLの振る舞い（SCL上昇/SCL変化無）との関連について調べた．その結果を表4に示す．

表 4: CO時の挙手の有無とSCLの上昇との関連

	挙手有り SCL上昇	挙手無し SCL上昇	挙手有り SCL変化無
強いプレイヤ	4	3	0
弱いプレイヤ	5	1	1

今回の実験では全ての試行でCOが行われており，CO時に挙手を行なったプレイヤは，他プレイヤに向けて指差ししながら占い結果を提示するといった流れがみられた．このように挙手を行なった場合には，ジェスチャや発言などプレイヤによる明示的な行動が確認できる．一方，挙手を行わなかった場合は，他プレイヤの様子（発言や指差し）をただ観察する様子がみられた．表３の結果から，強いプレイヤはCO時に自身が挙手を行なった場合のみならず，挙手を行なわなかった場合においてもSCLの上昇がみられた．一方の弱いプレイヤでは，挙手を行なった場合にSCLの上昇がみられたが，挙手を行わなかった場合には，SCLの上昇があまりみられなかった．つまり，強いプレイヤは弱いプレイヤよりも，CO時に行われるコミュニケーションについて，自身の参加の有無に関わらず，戦略的に注視していると考えられる．実際にCOでは，どのプレイヤが，どのプレイヤに対し，どのような判定結果を示しているのか，虚偽の有無について見極めることがその後のゲーム展開において大きな役割をもつため，強いプレイヤでこのような現象がみられたことは妥当であると考えられる．

6. おわりに

本研究では，複数のプレイヤのSCLの同期的な上昇から重要であると考えられるイベントを抽出し，強いプレイヤと弱いプレイヤの初日の会議時のSCLの振る舞いと言動を分析した．その結果，強いプレイヤはいずれのゲーム，役職でも傾向がみられ，弱いプレイヤは傾向がみられないことが理解できた．COによるSCLの振る舞いについて分析した結果，強いプレイヤはCO時に行われるコミュニケーションを戦略的に注視していることが示唆された．5人で行う人狼ゲームの戦略は，6人以上で行われる人狼ゲームでは適さない場合もあるが，COの重要性は変わらないと考える．また，本研究における強いプレイヤと弱いプレイヤの選定には未だ改良の余地があると思われる．プレイヤが電極を触ることでデータにノイズの混入がみられたことから，プレイヤが電極に触れてもノイズの混入を防ぐ計測システムの設計が必要となる．

7. 謝辞

本研究はJSPS科研費16H02928の助成を受けたものです．

参考文献

[1] 鳥海 不二夫，片上 大輔，大澤 博隆，稲葉 通将，篠田 孝祐，狩野 芳伸：人狼知能 だます・見破る・説得する人工知能，森北出版(2016)

[2] 伊藤 毅志，高橋 克吉，猪 爪歩，加藤 英樹，村松 正和，松原 仁：人間とコンピュータの思考の違い-囲碁の次の一手問題による考察-，ゲームプログラミングワークショップ2012論文集，2012 巻，6 号，pp.9 - 16(2012)

[3] 生月 誠，原野 広太郎．山口 正二：バイオフィードバックにおける情動安定性の指標としての SCL の検討，バイオフィードバック研究，13巻，pp.18-21(1986)

[4] 梅沢 章男，黒原 彰： バイオフィードバック指標としての皮膚コンダクタンス変化と皮膚電位活動の比較，バイオフィードバック研究，21巻，pp.29-36(1994)

[5] 新美 良純，白藤 美隆，皮膚電気反射-基礎と応用，医歯薬出版株式(1969)

[6] 沼田 恵太郎，宮田 洋，皮膚電気条件づけ-その意義と研究動向-，関西学院大学人文学会，人文論究 61(2)，pp.55-88(2011)

[7] 御手洗 彰，水丸 和樹，本田 健悟，棟方 渚，坂本 大介，小野 哲雄：人狼プレイヤの皮膚電気活動の解析: 情動変化を利用したソシオメータの実現へ向けて，情報処理学会 インタラクション，pp.885-888(2018)

[8] 風間 祥光，棟方 渚，畑 雅之，松原 仁：人狼ゲームにおけるプレイヤーの思考過程の分析，情報処理学会ゲーム情報学研究会，2016-GI-36(19)，pp.1-7 (2016)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

5 人人狼における戦略進化のシミュレーション
Simulation of Strategic Evolution in 5-player Werewolf

武田 惇史
Atsushi Takeda

鳥海 不二夫
Fujio Toriumi

東京大学
The University of Tokyo

In recent years, the game "Werewolf" has been drawing attention in the field of Artificial Intelligence. In this paper, we propose a method to simulate the evolution of strategy in order to investigate whether a strategy that always dominates other strategies is discovered in the 5-player werewolf regulation in future AI Wolf Competition. As a result of the simulation, we found that we can not reach a strategy that always has a dominant advantage, and eventually the strategy will continue to change periodically.

1. はじめに

近年，チェス，将棋，囲碁といった完全情報ゲームにおいて，人工知能技術は人間の能力を上回るまでに成長している．それに比べると，不完全情報ゲームにおける人工知能に関する研究はあまり行われてはいない．そこで，人工知能研究者が次に取り組むべき題材として，「人狼」が提案されており[篠田 14]，人狼における人工知能技術の促進のため，統一的なプラットフォームの開発，大会の開催などを行う「人狼知能プロジェクト」が発足している．

本研究の目的は，人狼知能大会に用いられる「5 人人狼」を対象としたシミュレーション手法を提案し，人狼知能においてナッシュ均衡となるような戦略が発見されるのかを調査することである．

2. 関連研究

人狼の元となったゲーム"Mafia"というゲームを対象にし，役職として「人狼」と「村人」のみが存在する条件下で，確率的統計より最適戦略を求める研究 が行われている[Braverman 08]．ここでは，話し合いをすることで，村人勝率が上がることなどが示されている．

本研究が対象とする 5 人人狼において考えられる戦略については，占い師の結果騙りが 5 人人狼において有効であるという，一般的な人狼にない，非直感的な戦略性について報告されている[片上 15]．

人狼における進化的シミュレーションを行った研究[大澤 17]では，人狼，占い師，村人の 3 人で行う 3 人人狼を定義し，進化シミュレーションを行うことで，どのような戦略があるか，どのような現象がおこるかが調べられている．ここでも，占い師の結果騙りが重要な戦略として挙げられていて，進化の中で一時的に支配的な戦略になる，という結果が報告されている．

3. 5人人狼

3.1 ルール

人狼ゲームは，プレイヤーが「村人陣営」と「人狼陣営」に分かれて行うチーム戦である．村人陣営は人狼を追放することで勝利となり，人狼陣営は人狼が追放されないことで勝利となる．ゲームは「日」を単位として進行し，1日はさらに「昼」と「夜」に分けられる．

表 1:5人人狼における役職の内訳

役職名	陣営	人数	能力
村人	村人	2	なし
占い師	村人	1	占い
人狼	人狼	1	襲撃
狂人	人狼	1	なし

本研究が対象とする5人人狼のレギュレーションにおいては，以下のような順番でゲームは進行していく．

(1) 役職の割り当て

最初に，プレイヤーに対し役職がランダムに割り当てられる．役職の内訳は表1に示す通りあらかじめ決まっている．また，役職ごとに陣営が決まっている．プレイヤーは，自分以外のプレイヤーの役職を知ることができない．したがって，誰が敵で誰が味方かわからない状態でゲームが始まる．

(2) 0日目：占い師の占い

占い師は，プレイヤーを一人選んでその役職が人狼であるかどうかを知ることができる．この能力を「占い」と呼ぶ．

(3) 1日目：話し合い，投票，追放

ここでは，プレイヤーは話し合いを行う．主に，占い師 CO（役職を表明すること）と占い結果の報告が行われる．

その後，各プレイヤーは他プレイヤーに対して投票を行い，最も得票数の多かったものは「追放」される．追放されたプレイヤーは今後ゲームに関わることができない．

(4) 1日目：襲撃，占い

ここでは，人狼がプレイヤーを自由に選び，「襲撃」を行う．襲撃されたプレイヤーは今後ゲームに関わることができない．

(5) 2日目：話し合い，投票，追放

この時点で残っているプレイヤーは3人である．残ったプレイヤーに対して，1日目と同様に話し合い，投票，追放が行われる．

(6) 勝利陣営の決定

人狼が追放されたら村人陣営の勝ち，最後まで人狼が生き残れば人狼陣営の勝ちである．

連絡先：武田惇史，東京大学工学部システム創成学科，113-8656 東京都文京区本郷 7-3-1 工学部 8 号館 526，TEL: 03-5841-6991 , E-mail: takeda@crimson.q.t.u-tokyo.ac.jp

表 3: 2017年データに対する回帰分析の結果と2018年における変化の対応

	①	②	③	④	⑤	⑥	⑦	⑧	⑨
回帰係数	2.0518	−0.3214	−1.5614	−2.2122	0.1943	−0.4192	0.0384	−0.5167	0.1807
標準誤差	0.0678	0.0384	0.0423	0.0505	0.0518	0.0231	0.1266	0.0707	0.0574
P値	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.7619	0.0000	0.0017
2017年平均	0.2145	0.1518	0.3876	0.1381	0.1039	0.5760	0.0211	0.0709	0.1421
2018年平均	0.0643	0.0000	0.0721	0.0280	0.3960	0.6239	0.0264	0.0320	0.0001
回帰係数の示す増減	減少	減少	減少	減少	増加	増加	−	増加	減少
実際の増減	減少	減少	減少	減少	増加	増加	−	減少	減少
増減の一致	○	○	○	○	○	○	−	×	○

表 2: 特徴量

番号	概要	値
1	人狼占いCO	二値(0 or 1)
2	占い師2ターン目CO	二値(0 or 1)
3	2CO盤面における初日占い投票率	実数値[0, 1]
4	2CO盤面における二日目占い投票率	実数値[0, 1]
5	2CO盤面における占い結果騙り	二値(0 or 1)
6	2CO盤面における裏切り者黒だし	二値(0 or 1)
7	3CO盤面における占い結果騙り	二値(0 or 1)
8	3CO盤面における裏切り者対抗占い	二値(0 or 1)
9	3CO盤面における人狼対抗占い	二値(0 or 1)

3.2 考えられる戦略

人狼ゲームは，能力の弱い多数（村人陣営）と能力の強い少数（人狼陣営）の対戦であると解釈することができる．追放は，投票による多数決によって行われるため，村人陣営にとって有利であるといえる．このことが，人狼ゲームにおける「人狼陣営が村人陣営を騙る」という構図を作り出している．

5人人狼においては，村人陣営の役職は村人または占い師であり，人狼陣営はこのどちらかを騙ることになる．どちらを騙るかは戦略において重要な要素である．また，占い師を騙る場合には嘘の占い結果を報告する必要がある．どのような嘘の占い結果を報告するかも，重要な要素になる．

4. 戦略の進化の分析

本章では，2017 年と 2018 年の人狼知能大会のログから，決勝に進出した強いエージェントのとる戦略がどのように遷移したのかを解析する．2018 年のエージェントは，2017 年において強かった戦法をとる傾向にあると予想される．

4.1 特徴量

解析に使う特徴量は表2に示す通りである．

特徴量の選択においては，以下の条件を満たすものとした．

- 行動主体となる役職が存在すること．例えば，特徴量 1 は「人狼占い CO」であり，行動主体は人狼である．
- 行動主体の意思で値を変えられること．例えば，「村人が人狼に投票する確率」などは本分析では特徴量になり得ない．なぜなら，村人にとって誰が人狼であるかは非公開の情報であり，村人が自由にこの量を変化させることができないからである．

4.2 解析手法

試合ごとの特徴量の値を説明変数として，その試合の勝敗を2値（村人陣営が勝ちなら1，そうでなければ0）で予測するロジスティック回帰モデルを作り，学習した．回帰係数は，正負が逆であるという帰無仮説のもと，有意水準 1 パーセントで検定をおこなう．P 値が 0.01 を上回った特徴量は結果の考慮に入れない．全データに対する特徴量の平均値の 2017 年から 2018 年における増減と，回帰係数の正負から，2017 年に対して強い戦略が 2018 年において用いられる傾向にあるかを明らかにした．

4.3 解析結果

結果を表 3 に示した．9 個の特徴量のうち，8 個でその回帰係数が有意であり，うち 7 個で回帰係数によって示されたエージェントにとって有利となる行動の増減と 2017 年から 2018 年における増減が一致した．このことから，2018 年大会のエージェントは 2017 年大会に対して勝率が高くなるように作られていると考えられる．

5. シミュレーション

5.1 概要

第 4 章の結果に基づいて，ある年のエージェントは前年の戦略から学習するという仮定で，戦略の進化をシミュレートする．少ないパラメータで，行動原理に基づいた方法で妥当な結果を得ることを目指す．

シミュレーションでは，占い師，人狼，裏切り者のとる戦略を表す変数をそれぞれS, W, Pと書く．ありえるすべての状態に対して，その発生確率はこれらの変数で表される．村人は，S, W, Pの値を知ることはできないが，セオリーとしてそれぞれS', W', P'に等しいという考えを持つものとする．このセオリーは，前年大会から学習したものと考える．S', W', P'の値と各プレイヤーのとった行動から，ベイズ推定的に役職を推定することができる．プレイヤーの行動は，S, W, PおよびS', W', P'を用いて計算される各プレイヤーの役職の確率分布によって決定される．起こりうるすべての状況に対しそれが起こる確率とその時の村人勝率の積の総和をとれば，村人勝率が分かる．すなわち，村人勝率はS, W, P, S', W', P'の関数として表すことができる．村人勝率を，$p_v(S, W, P, S', W', P')$と書くことにする．

各役職は，ほかの役職のとる戦略が変わらないという仮定のもと，自分の陣営の勝率を最大化するような値を探索する．

5.2 ゲームのモデル化

5人人狼モデルでは，表 4 に示すような6つの変数を設定する．

以下に挙げるようないくつかの仮定をおいてゲームを単純化し，決定論的なシミュレーションを可能にする．

- 占い師 CO 者とその占い結果の組み合わせ以外の情報をプレイヤーが利用することはできない
- 2日目に起こることは一切考えず，1日目の時点で決着が付くものとして考える．
- 1CO盤面になった場合，この時点で村人勝率は計算可能であり，それ以降のシミュレーションは行わない．

モデル化された5人人狼におけるゲーム進行を図 1 に示す．以下では，この図に従ってモデルの詳細を述べる．

(1) 占い師 CO

初日の最初のターンに占い師 CO が行われる．同時ターン制を仮定するので，人狼，裏切り者は互いの出方を見てから占い師を騙るかどうかを決定することができないものとする．

また，占い師は必ず占い師 CO するものとする．

人狼，裏切り者がそれぞれ占い師 CO をするかどうかで4通りの可能性がある．ゲームの進行上は，占い師 CO 者の人数ごとに3通りに分岐する．

(2) 1CO 盤面

1CO 盤面においては，村人勝率が 7/12 であることが解析的に求められる．

(3) 2CO：占い師の結果出し

占い師の結果騙りは，村人を占ったとき占ったプレイヤー以外の人をランダムで選んで黒だしする，というパターンのみを考える．裏切り者，人狼に関しては，占い先としてありえるプレイヤー4 通りと，占い結果（人狼か人狼でないか）の 2 通りで，4・2=8 通りの可能性があり，その中から一様な確率でランダムに選んで行動する．

(4) 2CO:勝利陣営の決定

1 日目の投票が終わった時点で以下のように村人勝率を定めることとし，2 日目以降のシミュレーションは行わない．

- 人狼が追放された場合，村人の勝ちとなる
- 裏切り者が追放された場合，確率 1/3 で村人の勝ちとなる
- 村人が追放された場合，村人の負けとなる．

(5) 3CO:占い師の結果出し

プレイヤー5 人のうち，占い師 CO した 3 人以外の 2 人は村人であることが全員の目線で確定する．占い師 CO した者たちの結果は，実質的に，村人に白出しするか対抗に黒出しするかのどちらかである．なぜならば，村人に黒出しすることは役職の割り当ての組み合わせが存在せず，論理破綻を起こすためである．また，それぞれの占い師 CO 者の目線において対抗の一人は裏切り者でもう一人が人狼であることはわかっているので，対抗の一人に白出しすることは，もう一人の対抗に黒出しすることと等価である．

人狼は，対抗の二人を区別する情報がないため，2 人いる対抗を占う確率はそれぞれ$(1/2)\cdot w_2$である．裏切り者についても同様のことが言える．また，真の占い師については，もし占い先が村人であった場合には確率s_2で対抗のどちらかに 50%の確率で黒出しするものとする．

(6) 3CO:勝利陣営の決定

1 日目の投票が終わった時点で以下のように村人勝率を定めることとし，2日目以降のシミュレーションは行わない．

- 人狼が追放された場合，村人の勝ちとなる．
- 裏切り者が追放された場合，確率 1/2 で村人の勝ちとなる．
- 村人が追放された場合，村人の負けとなる．

(7) 投票，追放先の決定

本モデルでは，村全体で合意の取れる人に最も投票が集まるものと考え，追放先を決定する．この仮定は，話し合いによって村人同士で投票先を合わせられること，人狼は 2 日目のことを考えて村人または占い師を騙らなければならず，怪しまれるような投票はできないことから，この仮定は妥当であると考える．

表 4: シミュレーション変数

変数名	意味
s_1	2CO盤面における占い師の結果騙り率
w_1	3CO盤面における占い師の結果騙り率
p_1	人狼の占い師CO率
s_2	3CO盤面における人狼の対抗占い率
w_2	裏切り者の占い師CO率
p_2	3CO盤面における裏切り者の対抗占い率

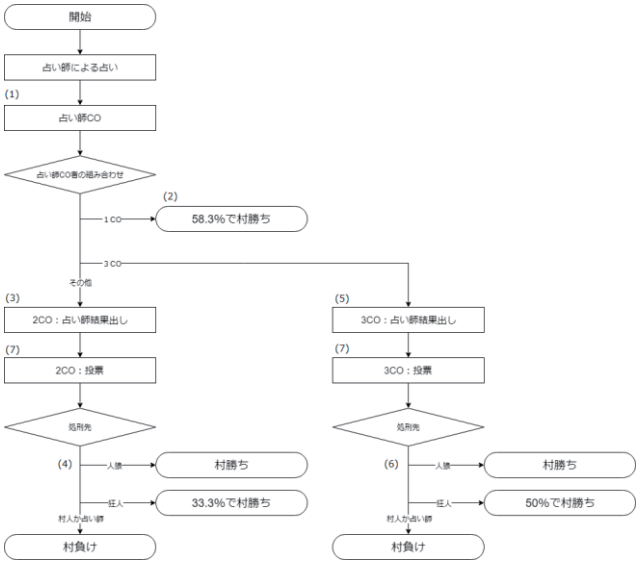


図 1:モデル化されたゲームのフロチャート

プレイヤーにとって，このモデルとS,W,Pの値を既知とする場合，ベイズ推定によって各プレイヤー役職を推定することができる．実際には，村人たちはS,W,Pの値を知らず，代わりに，第三者視点から，S',W',P'を用いてベイズ推定により各プレイヤーの役職の確率分布を求める．村人たちはモデル自体を共有していることを仮定するので，勝率が最も高くなるような投票先を選ぶことが可能である．具体的には，3CO 盤面の場合，(人狼確率) + (1/2)・(裏切り者確率)が最大のプレイヤーに投票し，2CO 盤面の場合，(人狼確率)+ (1/3)・(裏切り者確率)が最大のプレイヤーに投票する．

5.3 勝率計算

すべての変数の値を固定すれば，ゲームの考えうるすべての展開をシミュレートすることによって，村人陣営の勝率の厳密な計算が可能である．すなわち，起こりうるすべての盤面の状態に対して，それが起こる確率とそのときの村人勝率の積の総和を計算すればよい．盤面の状態とは具体的に，3CO盤面の場合，3人の占い師 CO 者の結果の組み合わせのことであり，2CO 盤面の場合，占い師を騙ったのが人狼と裏切り者どちらであるかと，占い結果の組み合わせのことである．

各盤面の状態に対する発生確率はS,W,Pから計算可能である．そのときに追放されるプレイヤーはS',W',P'から計算可能であり，その時点で村人勝率は確定する．

以後，村人陣営の勝率を$p_v(S,W,P,S',W',P')$と書くことにする．

5.4 シミュレーション条件

$s_1, w_1, p_1, s_2, w_2, p_2$はそれぞれ 0.2 と 0.8 の二値を取るものとする．ある時点における流行りの戦略が(S', W', P')とあらわされ，それに対し，次のステップにおいて流行りの戦略が$(S_{new}, W_{new}, P_{new})$になるとしたとき，

$$S_{new} = \mathrm{argmax}_S\ [p_v(S, W', P', S, W', P')]$$
$$W_{new} = \mathrm{argmin}_W\ [p_v\ (S', W, P', S', W', P')]$$
$$P_{new} = \mathrm{argmin}_P\ [p_v(S', W', P, S', W', P')]$$

が成り立つものとする．

5.5 結果と考察

シミュレーションの結果を図2に示す．ノードの書かれた数字は上段が左からs_1, w_1, p_1の値であり，下が左からs_2, w_2, p_2の値を表している．また，5.3 節における(S', W', P')から$(S_{new}, W_{new}, P_{new})$に向かってエッジが付与されており，エッジを辿ることで戦略の遷移を追うことができる．

表３から，s_1が増加，w_1が減少，p_1が高い値を維持，s_2, w_2, p_2が低い値を維持している．ここで，2017 年，2018 年それぞれで採用されていた戦略を見ると，赤枠で囲われた箇所に相当していることが分かった．したがって，本シミュレーションは 2017 年から 2018 年の戦略の進化を内包しているといえる．

任意の状態から遷移をたどると，特定の一状態に収束はせず，図2に青枠で示したサイクルに到達することが明らかとなった．従って，2017年及び2018年のデータとの対応が見られた赤枠の部分から遷移を辿っても，安定状態には到達しない．このことから，今後の人狼知能大会における戦略の推移において，ナッシュ均衡解は発見されない可能性が示唆された．

6. 結論

過去の人狼知能大会のデータログから，2017年，2018年大会においてどのような戦略が各陣営の勝利に貢献しているかを分析した．2018年のエージェントは2017年のエージェントに対して勝率が高くなるような戦略を取る傾向にあることを回帰分析により確認した．

また，5人人狼のゲームを現実的な時間で網羅的，決定論的シミュレーションが行えるようにモデル化した．実際にシミュレーションを行うことで，遷移を繰り返していっても特定の状態に収束せず，周期的に変化し続けるようになるという結果を得た．このことから，今後の人狼知能大会において，ナッシュ均衡解には辿り着かない可能性があることが明らかとなった．

本手法は，2日目を考えないなど大きな仮定を置いて，ゲームを単純化している．その点において，現実を正確に表せていないと考えられる．単純化を取り除くためには，行動原理の設定に工夫が必要である．しかし，どのように設定するべきかは非自明であり，より詳細なログ解析が必要である．この点に関しては今後の課題とする．

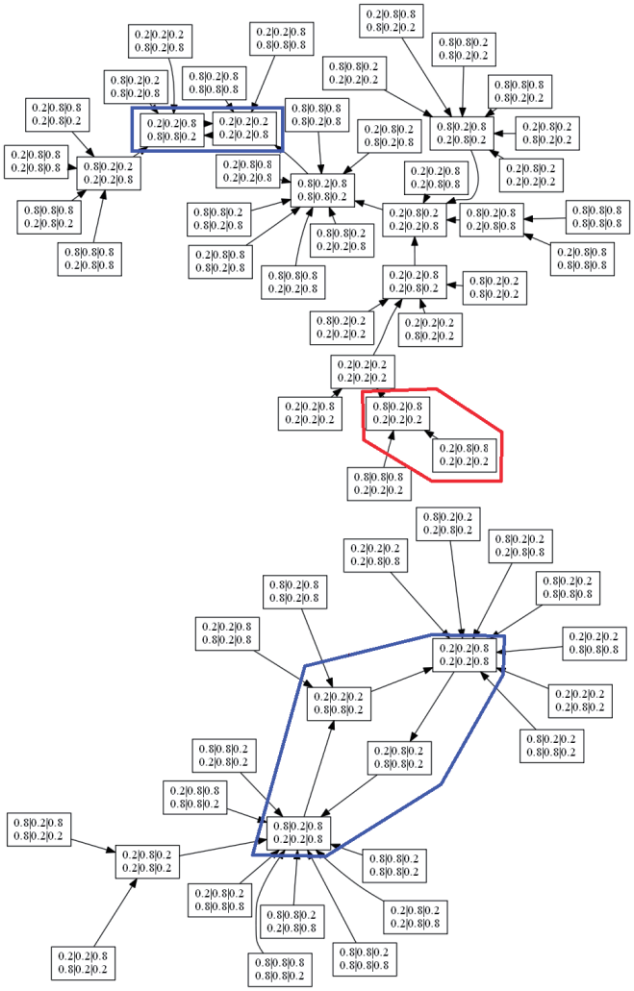


図 2:シミュレーション結果の遷移図

参考文献

[篠田 14] 篠田孝祐，鳥海不二夫，片上大輔，大澤博隆，稲葉通将:汎用人工知能の標準問題としての人狼ゲーム，人工知能学会全国大会 JSAI2014，2C4-OS-22a-3, 2014.

[Braverman 08] M.Braverman, O.Etesami, and E.Mossel. Mafia: A theoretical study of players and coalitions in a partial information environment. *The Annals of Applied Probability*, pp. 825–846, 2008.

[片上 15] 片上大輔，鳥海不二夫，大澤博隆，稲葉通将，篠田孝祐，松原仁，人狼知能プロジェクト，人工知能学会論文誌，vol.30(1)，pp.65-73, 2015.

[大澤 17] 大澤博隆，汪博豪，佐藤健:進化シミュレーションを用いた3人人狼の分析，日本ソフトウェア科学会第 34 回大会講演論文集, 2017.

付録:人狼用語

(1) CO

カミングアウトの略．自らの役職を表明すること．

(2) 1CO,2CO,3CO

占い師 CO する人数が1人，2人，3人であるようなゲームの状態をそれぞれ 1CO,2CO,3CO と呼ぶ．

(3) 対抗

占い師 CO したプレイヤーに対し，他の占い師 CO したプレイヤーを対抗と呼ぶ．

(4) 白出し，黒出し

占いの結果報告が，「人狼でない」の場合白出し，「人狼である」の場合黒出しと呼ぶ．

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

協調パターンタスクの行動実験に基づいたシミュレーション分析

Simulation Analysis based on Behavioral Experiment of Cooperative Pattern Task

渡邊紀文 *1 Norifumi Watanabe　糸田孝太 *2 Kota Itoda

*1武蔵野大学 Musashino University　*2慶應義塾大学 Keio University

We have a behavior experiment using pattern task abstracting cooperative behaviors that require intention estimation and action switching to specific goals. And we have analyzed strategies to adjust cooperative intention estimations. In this research, we constructed an agent model that have three strategies of "random selection", "self-priority selection", and "other agent's target pattern estimation". And the decision making process was verified by simulation.

1. はじめに

我々は他者とのコミュニケーションにおいて，他者の行動からその意図および次に行われるであろう行動を推定し，それにあわせる形で自己の行動を決定する．このような協調行動は 1 対 1 のコミュニケーションにおいては，他者の行動に信頼をおいて次の行動戦略を決定することが多いが，協調すべき他者が複数存在する場合は，着目すべき他者を選択および順位付けをし，順番に意図を推定して協調すべき他者を決定すると考えられる．具体的にはサッカーやハンドボールといったゴール型ボールゲームにおいては，自身が敵に囲まれたときに，周囲に存在するチームメイトの中からパスを受け取りゴールに近づくことが出来る選手を複数選択し，それぞれのパス成功確率および次に行うであろう行動意図を推定する．更にゴールに繋がるまでの協調パターンを評価し，最も得点に繋がるパターンを選択して，そこに存在する選手にパスを出すと考えられる．このような協調パターンを人間がどのように評価しているのか，また人間と協調パターンを形成するエージェントにはどのような行動戦略モデルを構築すべきかを明らかにするため，本研究では協調パターンタスクの行動実験に基づいたモデル構築とシミュレーション分析を行う．

2. 先行研究

他者の意図・信念状態の表現及び意図推定の深さに関しては，誤信念課題に見られるような他者の知識としての信念表現 [1] や，またより工学的な立場からは BDI(Belief-Desire-Intention) モデルに見られるような信念 (belief)，願望 (desire)，意図 (intention) の三つの主要なパラメータによる認知過程の説明が行われてきた [2]．近年では確率モデルを用いて人間の意図の推定を逆計画 (reverse planning) の問題として扱う研究 Bayesian Theory of Mind[3] や，強化学習モデルを用いた様々な深さの意図を持つエージェントによる協調課題の研究 [4] があり，意図を扱う問題に対するエージェントベースアプローチの有効性が示されている．意図推定の対象とすべき他者の選択では，見えない他者への信念に関する研究 [5] なども行われている．

連絡先: 渡邊紀文，武蔵野大学データサイエンス学部，東京都西東京市新町 1-1-20，noriwata@musashino-u.ac.jp

3. 協調パターンタスク

これまで我々は，協調課題パターンタスクの開発及び複数人で実験できるシステムの構築を行い，人のみ 4 人で構成された小集団における行動実験を通じて被験者の行動を分析した [6]．

パターンタスクでは同時に 4 人が参加をし，2 次元グリッドワールドで非言語コミュニケーションのみで協調し共通の目標を達成する事を目的とする．それぞれの被験者はグリッドワールド上の円形のコマを操作し，各ステップにおけるお互いの行動をもとに他者の意図を推定し，目標となるパターンを形成する（図 1）．目標とするパターンは相対位置関係で表現する図形であり，4 つのコマの内 3 コマで構成される．そのため各被験者はパターン形成に関与するコマを選択して行動をする必要がある．なおパターンは平衡移動した座標でも達成と認めるが，回転や反転した座標は認めない．

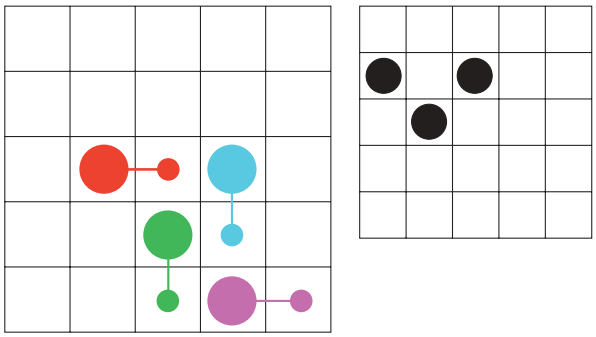


図 1: 被験者が移動するグリッドワールド（左）．大きい円のコマで各被験者のそれぞれの位置を示し，小さい円によって一ステップ前に被験者がどこにいたのかを表示している．目標となるパターン（右）．

他者との意図の調整について分析をしたところ，課題中の序盤や終盤での調整の違いによって目標パターン到達の進度が変化するという結果を得た．また最適なステップで目標に到達する行動を被験者が仮定することで，意図の誤推定を防止するという方策を持つことが示唆された．そこで次にパターン選択における被験者の戦略について分析した．

3.1 パターン選択の戦略

行動実験より，被験者が想定するパターンと他者のパターンとの関係について，序盤で各被験者は新規パターンを想定しているが，全員の意図が一致をする時には一部の被験者が自身の目標パターンを変更し，最終的には全員一致した状態でそれぞれのパターンを継続して推定するという結果が得られた．具体的には被験者が選択するパターンに次の 3 つの関係がみられた．

戦略 a 前ステップの自分が選択したパターンを選択

戦略 b 前ステップで他者が選択したパターンを選択

戦略 c a, b とは異なる新規のパターンを選択

これらの 3 つの戦略に基づいて，被験者のパターン選択過程を分析したところ，タスク初期では戦略 c の新規のパターンを選択し，その後は被験者のうち一人が他の被験者が選択しているパターンに合わせるという戦略の変化が見られた．このような戦略の変化は試行前半で共通合意に至る場合と，後半で一致する場合に分けられ，少ないステップ数で目的パターンを達成できている多くの場合では，序盤に行動意図が分かるような明示的な行動を取り，それに対して他者が合わせる事が多く見られた．この結果より，全員が協調する事を促す本パターンタスクにおいては，自己の行動意図をできるだけ分かりやすい形で他者に提示し，各状況において最も全体の到達ステップが短くなるような目標パターンを選択する．更に選択可能なパターンの中から多数決をとり，多くの被験者が取っているパターンを優先して選択するという行動戦略が考えられる．この結果から得られた集団での行動戦略を元に，エージェントモデルを構築した．

4. シミュレーション分析

3 章での分析結果を元に構築したエージェントモデルを，パラメータの違いにより複数用意し，シミュレーションを行う事で，最も被験者の行動を説明できるモデルを検証する．エージェントの条件は次のように設定した．

ランダム選択 4 体のエージェントから最短経路で到達するパターンに関与するエージェント 3 体をランダムに選択し，入っていない場合はランダムな方向へ移動

自己優先選択 最短経路で到達するパターンの中から，自分が含まれているパターンを優先して選択

他者エージェントの推定 他者エージェントの 1 ステップ前の行動から目標とするパターンを推定し，それらの中から最も多く選択されているパターンを選択

シミュレーションでは初期位置と初期目標パターンを 100 回ずつランダムに用意し，それぞれの目標パターン到達ステップ数を「ランダム選択」と「自己優先選択」行動選択と,「他者エージェントの推定」の有無の組み合わせ条件によって比較した．ステップ数を比較した結果を図 2 に示す.「ランダム選択」および「自己優先選択」ともに,「他者エージェントの推定」を行った場合の平均到達ステップ数（1-b および 2-b）は減少した．更に「ランダム選択 (1-b)」と比較し,「自己優先選択（2-b)」の戦略をとった場合がステップ数が少なかった．

「他者エージェントの推定」を選択した場合は，該当のエージェントとその行動から目標パターンを推定するため，その

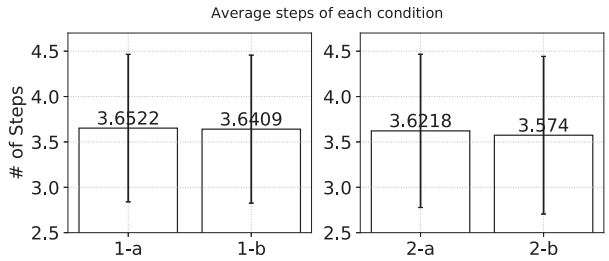


図 2: 目標パターン到達のステップ数の比較．1-a:ランダム選択+他者エージェントの推定無，1-b:ランダム選択+他者エージェントの推定有，2-a:自己優先選択+他者エージェントの推定無，2-b:自己優先選択+他者エージェントの推定有．

エージェントが含まれている目標パターンを優先的に選択し,「自己優先選択」他者にも仮定する事になる．そのため他者の次のステップの行動選択と合致する事により，到達ステップ数は大幅に減ると考えられたが，そのような結果は得られていない．この原因としては，多数決の結果が複数存在した時に，その中からランダムに選択していたことが影響していると考えられる．

5. おわりに

本研究では協調パターンタスクの行動実験に基づき,「ランダム選択」「自己優先選択」「他者エージェントの推定」の 3 つの戦略を持つエージェントモデルを作成した．シミュレーション結果より.「自己優先選択」および「他者エージェントの推定」の両戦略を持つエージェントが最も早く目標パターンに到達することができ，人間の行動実験と同様の結果が得られた．ただし本結果の差異は想定していた結果よりも小さいため，今後各ステップでの人間の行動選択との比較など詳細な分析が必要である．

参考文献

[1] 子安増生，心の理論―心を読む心の科学 (岩波科学ライブラリー (73))，岩波書店，2000

[2] G. Weiss: Multiagent Systems, second edition, MIT press, 2013

[3] C. L. Baker, R. Saxe, J. B. Tenenbaum: Action understainding as inverse planning”, Cognition, pp.329-349, 2009

[4] 横山絢美，大森隆司: 協調課題における意図推定に基づく行動決定過程のモデル的解析”，電子情報通信学会論文誌，Vol. J92-A，No.11，pp.734-742, 2009

[5] 大森隆司，ヒトの心のプロセスの計算論的理解に向けて，電子情報通信学会技術研究報告，Vol. 109, No. 460，pp.107-112，2010

[6] K. Itoda, N. Watanabe, Y. Takefuji: Analyzing Human Decision Making Process with Intention Estimation using Cooperative Pattern Task, Artificial General Intelligence — 10th International Conference (Springer Verlag), Vol.10414LNAI, No.10, pp.249-258, 2017

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

3者間人狼における他者の投票行動を考慮した戦略の検討

Analyis of Strategies in Werewolf Game by 3 Players
Considering Voting Behaviour of Other Players

玉井 日菜子
Hinako Tamai

奈良女子大学大学院 人間文化研究科 博士前期課程 情報衣環境学専攻
Department of Computer Science and Clothing Environment,
Graduate School of Humanities and Sciences, Nara Women's University

In this research, based on the Utterance model in Werewolf Game by 3 Players of the previous research, we examined the strategy of reasonable voting behavior when anticipating the voting behavior of others. As a result, we found that a strategy which is quite different from the voting behavior which is considered reasonable when not considering the voting behavior of others is obtained. In addition, we found that the same strategy appears at regular intervals when the estimation of the voting behavior of others is repeated.

1. はじめに

近年、人狼ゲームに関する研究は盛んに行われている。人狼ゲームとは、対話型のコミュニケーションによって進行する不完全情報ゲームである。その特徴から、人工知能のための標準問題としての可能性が検討されている [篠田 14]。従来研究では、人狼ゲームをプレイする人工知能が挙げられ、強い人狼エージェントも現れている。既存の人狼エージェントでは強化学習を使用したものが多いが、論理的思考に着目したものは少ない。一方で [大澤 14] において、人狼におけるエージェントの推論モデルとして、各エージェントの信念や意図などを明示的に表現できる BDI モデルの利点が指摘されており、また、BDI 論理を用い、実際の人狼ゲームで行われている複雑な推論を記述する例が示されている。そこで、我々は BDI 論理に基づいた論理的思考を戦略決定に用いる人狼エージェントの実現を目指している。しかし、実際の人狼ゲームでの推論は多彩かつ複雑であるため、エージェントに組み込める形にするには人狼ゲームの論理的分析が必要となる。

論理的分析を行っている例として、[大澤 16] が挙げられ、人狼ゲームの最小系である 3 人人狼において、各プレイヤーの発言の選択肢を削減できるか検討されている。そこで、本研究では、[大澤 16] で使用された 3 人人狼における発話のモデルを元にし、他者の投票行動を予想した上での合理的な投票行動の戦略について検討した。その結果、互いに他者の行動を予想しあうと安定した戦略に収束しないこと、また、[大澤 16] の分析で合理的とされる投票行動とはかなり異なる戦略が得られることが分かった。

2. 制限を設けた人狼の戦略分析の研究

[西崎 17] では、ワンナイトではあるが 3 者間ではない人狼の発話や投票に関する戦略分析が行われている。こちらは実際のゲーム (100 ゲーム) の分析である。[杉本 17] では、5 人人狼で村人に着目して投票行動の決定過程の分析が行われている。これも実際に行ったゲームの過程を分析している。これらに対し、我々は 3 者間で全ての可能性の分析を行っている点が異なる。[汪 17] では、3 人人狼で、進化シミュレーションの手法を用いて、人狼プレイヤーがどのように戦略を変化させていくかの分析が行われている。プレイヤーが他プレイヤーの行動を考慮した戦略決定を行っている点は本研究と同様と考えられるが、シミュレーションによって変化の過程を調べている点が我々と異なる。

連絡先: 玉井 日菜子，奈良女子大学情報衣環境学専攻生活情報通信科学コース，sah_tamai@cc.nara-wu.ac.jp

3. 一般的な人狼ゲームのルール

ゲーム開始時に各プレイヤーに役職が割り当てられ、役職に従い村人陣営と人狼陣営に分かれてプレイを行う。また、役職に応じて特殊能力が与えられる。役職には例えば、特定のプレイヤーが人狼であるかどうか知ることが出来る占い師などがある。村人陣営の勝利条件は人狼をすべて追放することであり、人狼陣営の勝利条件は村人の数を人狼の数以下にすることである。ゲームは昼と夜の 2 つのフェーズで進行する。昼のフェーズでは各プレイヤーが自由に対話を行い、得られた情報を元に、投票によって誰をゲームから追放するかを決定する。夜のフェーズでは、それぞれの役職に応じた能力を行使する。人狼は 1 人のプレイヤーを指定し襲撃することができる。対話において、村人陣営側は人狼陣営の嘘を見破ることが重要となる。一方人狼陣営側は役職を偽るなどの嘘をつき、議論を混乱させ自分たちが不利にならないように誘導することが重要となる。ゲームが進むにつれ、各プレイヤーは投票で追放されるか、人狼に襲撃されることでゲームから除外される。そうして勝利条件のどちらかが達成されたとき、ゲームは終了する。

4. 3 者間人狼の定義

当研究では、[大澤 16] の 3 人人狼における発話のモデルを使用し、また、相手の投票行動を想定しない場合の合理的な行動の分析を出発点としてを使用する。以下に概略を述べる。

4.1 3 者間人狼の扱う範囲

3 者間人狼の場合、処刑が一回行われるとゲームが終了するため、夜のフェーズは存在しない。同時発話 1 回のもと、投票を行う。全員が同数投票 (3 者間人狼では各自 1 票ずつ投票) された場合引き分けとする。役職は、村人陣営である村人、占い師、人狼陣営である人狼で構成され、各役職につき、プレイヤーは 1 人である。占い師にはゲーム開始前にどちらが狼であるかの情報を与える。

4.2 プレイヤーの合理的な行動 (相手の投票行動を想定しない場合)

[大澤 16] では、以下の投票行動を合理的であるとして採用している。

- 村人の戦略: 自身が狼と一人だけから告げられたとき、そのプレイヤーへ投票する。また、村人プレイヤーにとって自分以外を狼と指すプレイヤーが 1 人だけいるとき、狼と指されたプレイヤーに投票する。
- 人狼の戦略: 自身が狼と一人だけから告げられた時、そのプレイヤーへの投票する。

上記以外の場合は、投票先を決定することはできない。

4.3 発話について

発話の種類は以下の 4 通りである。簡易化のため、以降村人、人狼、占い師のプレイヤーをそれぞれ、Pv、Pw、Ps と表記する。

- To :「自身が占い師ではない」
- Tv :「自身が占い師であり、狼は Pv (村人) である」
- Tw :「自身が占い師であり、狼は Pw (人狼) である」
- Ts :「自身が占い師であり、狼は Ps (占い師) である」

ただし、自身が占い師でありながら狼は自身であると発言することは矛盾するため、各プレイヤー 3 通り (村人は Tv、人狼は Tw、占い師は Ts を除く) の発話がありえる。従って、発話の可能性は計 27 通りある。

4.4 投票について

占い師は、誰が人狼であるかを知ることができるため、常に人狼に投票することが有利であるが、村人、人狼それぞれのプレイヤーには 2 通り (村人は占い師または人狼に投票、人狼は村人もしくは占い師に投票) の行動が存在する。27 通りの発話それぞれの場合における投票行動を [表 3]([大澤 16] による) に示す。このうち、先述の合理的な行動によって行動が決まる場合は、その行動をとるものとする。Pv, Pw, Ps は村人・人狼・占い師の発話、Vv, Vw は村人・人狼の投票先を表し、V, W, S は村人・人狼・占い師への投票を表す。U は投票行動を決定できない (他の 2 プレーヤのいずれへの投票も含んだ) 状況を表す。

4.5 勝敗について

[表 3] にはそれぞれの場合の勝敗も示されている。勝敗については、簡略化のため [表 1] のように記述する。村人と人狼の投票先がどちらも U の場合、どちらとも判別がつかず、村側勝利、狼側勝利、引き分けのいずれをも含んだ状況となる [表 2]。この状況を [大澤 16] では case 3.1 と表しているため、便宜上同様の表現を使用する。

村側勝利	V_win
狼側勝利	W_win
引き分け	draw
表 2	case 3.1

表 1: 勝敗についての表記

村人投票先	人狼投票先	結果
人狼	占い師	村側勝利
人狼	村人	村側勝利
占い師	占い師	狼側勝利
占い師	村人	引き分け

表 2: どの結果も含む状況

Pv	Pw	Ps	Vv	Vw	result
To	To	To	U	U	case 3.1
To	To	Tv	S	U	draw or W_win
To	To	Tw	W	S	V_win
To	Tv	To	W	U	V_win
To	Tv	Tv	U	U	case 3.1
To	Tv	Tw	W	S	V_win
To	Ts	To	S	U	draw or W_win
To	Ts	Tv	S	U	draw or W_win
To	Ts	Tw	U	S	V_win or W_win
Tw	To	To	U	V	V_win or draw
Tw	To	Tv	S	V	draw
Tw	To	Tw	W	U	V_win
Tw	Tv	To	W	V	V_win
Tw	Tv	Tv	U	V	V_win or draw
Tw	Tv	Tw	W	U	V_win
Tw	Ts	To	S	V	draw
Tw	Ts	Tv	S	V	draw
Tw	Ts	Tw	U	U	case 3.1
Ts	To	To	U	U	case 3.1
Ts	To	Tv	S	U	draw or W_win
Ts	To	Tw	W	S	V_win
Ts	Tv	To	W	U	V_win
Ts	Tv	Tv	U	U	case 3.1
Ts	Tv	Tw	W	S	V_win
Ts	Ts	To	S	U	draw or W_win
Ts	Ts	Tv	S	U	draw or W_win
Ts	Ts	Tw	U	S	V_win or W_win

表 3: 一回同時発話の場合の結果の分類 (Vs=W)

5. 他プレイヤーの戦略を推測した上での戦略の検討

[大澤 16] では、相手の投票行動が 4.2 に述べたように仮定され、全プレイヤーに共有された場合の、プレイヤーの合理的な発話行動について考察されている。具体的には、各プレイヤーの発話の選択肢が削減できるかどうかを考えており、特定の発話を行った場合、全ての状態において他の発話よりも自身の勝利の可能性が増えるならば、その戦略はより強い戦略といえ、選択肢を削減できる。

本研究では、発話の削減については考慮せず、全ての発話は等しく選択されると仮定し、その結果、どのように投票行動が変化しうるかについて考察する。具体的には、相手の投票行動を仮定した上でどのプレイヤーに投票するのがより強いか考察し、戦略を決定する。さらに決定された戦略を踏まえ、こちらがその戦略をとると予想した相手プレイヤーが新たに戦略を決定する操作を繰り返すことにより、どのような結果が得られる

か検討する。表記を簡易化するため、村人が Vn の戦略を選択すると仮定した場合の最適な狼の戦略を Wn+1、狼が Wn の戦略を選択すると仮定した場合の村人の戦略を Vn+1 と表す (n は 0 以上の整数)。また、V0、W0 は、4.2 で述べた、相手の投票行動を想定しない場合の戦略を指す。

5.1 一人称視点での発話の分類

一人称視点での発話パターンの分類ごとの戦略を検討する。自分視点で区別がつけられない発話パターンを 1 つのグループとする。村人、人狼、いずれの場合も 15 グループに分類できる [表 4][表 5]。例として村人視点で、二つの発話パターン Pv = To, Pw = To, Ps =Tv と、Pv = To, Pw = Tv, Ps =To について述べる。この二つの発話の違いは、占い師が自身を狼と指したか、人狼が自身を狼と指したかである。しかし、村人視点では、どちらの発話パターンでも片方のプレイヤーが自分を狼と指したことしか分からない。よって、この 2 つの発話パターンは同グループに分類される。しかし、村人以外のプレイヤーの発話が上記と同じ場合でも、村人が占い師宣言した場合、区別が可能な場合が存在する。その例として、Pv = Ts, Pw = To, Ps =Tv と、Pv = Ts, Pw = Tv, Ps =To が挙げられる。これらは、自身を狼と指したプレイヤーを狼と指した場合と、自身を狼と指していないプレイヤーを狼と指したという点で区別可能である。

Pv	Pw	Ps
To	To	To
To	To	Tv
To	Tv	To
To	To	Tw
To	Ts	To
To	Tv	Tv
To	Ts	Tw
To	Tv	Tw
To	Ts	Tv
Tw	To	To
Ts	To	To
Tw	To	Tv
Ts	Tv	To
Tw	To	Tw
Ts	Ts	To
Tw	Tv	To
Ts	To	Tv
Tw	Tv	Tv
Ts	Tv	Tv
Tw	Tv	Tw
Ts	Ts	Tv
Tw	Ts	To
Ts	To	Tw
Tw	Ts	Tv
Ts	Tv	Tw
Tw	Ts	Tw
Ts	Ts	Tw

表 4: 村人視点のグループ化

Pv	Pw	Ps
To	To	To
To	To	Tv
To	Tv	To
To	To	Tw
To	Ts	To
To	Tv	Tv
To	Ts	Tw
To	Tv	Tw
To	Ts	Tv
Tw	To	To
Ts	To	To
Tw	To	Tv
Ts	Tv	To
Tw	To	Tw
Ts	Ts	To
Tw	Tv	To
Ts	To	Tv
Tw	Tv	Tv
Ts	Tv	Tv
Tw	Tv	Tw
Ts	Ts	Tv
Tw	Ts	To
Ts	To	Tw
Tw	Ts	Tv
Ts	Tv	Tw
Tw	Ts	Tw
Ts	Ts	Tw

表 5: 狼視点のグループ化

5.2 戦略の決定

最も有利となる投票先を、以下の条件のもとで決定する。

- 一番勝つ可能性が高い投票を行う。
- 勝つ可能性が同じの場合、引き分けになる確率が高い投票先を選ぶ
- 上記の操作を行い 1 つに決定出来ない場合、どちらに投票してもよいとする。

プレイヤーの区別がつき、かつ、どちらに投票しても有利度が変わらない場合、村人の場合 WS、人狼の場合 VS と表記する。プレイヤーの区別がつかない場合、投票先を決定できないため、U と表記する。

5.3 結果と考察

V0 を起点とした場合、V6 以降 V2〜W5 をループし、W0 を起点とした場合、W6 以降 W2〜V5 をループすることが判明した [表 6][表 7]。また、それぞれの戦略における、各勝敗の合計と、勝率を [表 8] に記載し、それぞれの戦略と発話パターンにおける勝敗の詳細については、`http://blackknight.ics.nara-wu.ac.jp/~u1448029/winloss.pdf` に記載する。VW0 は、相手の投票行動を想定しない場合の戦略を両プレイヤーが取った場合、つまり、[表 3] に記された勝敗の合計である。[表 8] 中の Wn は、村人が Vn-1 の戦略、人狼が Wn の戦略を選択した場合の勝敗のことを指す (n は 1 以上の整数)。Vn についても同様である。W1 では例えば Pv=To, PW=To, Ps=Tw の場合に投票行動が V となるなど、4.2 で述べた戦略とは異なるものが得られることも分かった。また、村人の場合、引き分けと負けの比率は変動するが、どの戦略を選択しても勝つ確率は変わらないことが分かった。さらに、相手プレイヤーの投票行動が推定通りの場合、人狼は W4、村人は V4、V6 の戦略を取ると最も有利になり、相手プレイヤーが自分の投票行動を推測した上で投票したと仮定すると、人狼は V0、村人は W0、つまり相手の投票行動を想定しない場合の合理的な戦略を取ると最も有利であることが分かった。

6. まとめ

本研究では、先行研究における 3 人人狼における発話のモデルを元にし、他者の投票行動を予想した上での合理的な投票行動の戦略について検討した。その結果、他者の投票行動を考慮しない場合の戦略とは、かなり異なる戦略が得られた。また、他者の投票行動の推定を繰り返すと、一定間隔で同じ戦略が現れ、安定した戦略に収束しないことが分かった。今後は、人数や発話内容を増やす等のモデルの複雑化を行い、同様の研究を行う予定である。

参考文献

[篠田 14] 篠田孝祐, 鳥海不二夫, 稲葉通将, 大澤博隆, 片山大輔, 人工知能標準問題としての人狼ゲームの提案, 第 24 回インテリジェント・システム・シンポジウム, pp.74-77, 2014.

[大澤 14] 大澤博隆, 鳥海不二夫, 稲葉通将, 片山大輔, 梶原健吾, 篠田孝祐, 人狼知能達成におけるエージェントの推論モデル, 第 19 回 ゲームプログラミングワークショップ 2014 , pp.157-161, 2014.

[大澤 16] 大澤博隆, 佐藤健, 3 者間人狼における戦略の検討, 2016 年度人工知能学会全国大会, 2F4-3, 2016.

	VW0	W1	V2	W3	V4	W5	V6	V1	W2	V3	W4	V5	W6
V_win	9	9	12	12	12	12	12	9	9	11	11	10	10
V_win or draw	2	3	1	1	1	1	1	0	3	1	0	1	2
case 3.1	5	3	1	1	1	1	1	8	4	4	4	4	4
V_win or W_win	2	3	1	1	1	1	1	1	2	0	1	2	1
draw	3	0	10	0	11	0	11	6	0	10	0	10	0
draw or W_win	6	1	2	1	1	1	1	3	0	1	1	0	0
W_win	0	8	0	11	0	11	0	0	9	0	10	0	10
村側勝率	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
狼側勝率	0.19	0.40	0.06	0.45	0.05	0.45	0.05	0.15	0.41	0.06	0.44	0.07	0.43

表 8: 戦略ごとの各プレイヤーの勝率

Pv	Pw	Ps	V0	W1	V2	W3	V4	W5	V6
To	To	To	U	U	U	U	U	U	U
To	To	Tv	S	S	W	V	S	S	W
To	To	Tw	W	V	S	S	W	V	S
To	Tv	To	W	V	S	S	W	V	S
To	Tv	Tv	U	V	U	S	U	V	U
To	Tv	Tw	W	V	S	S	W	V	S
To	Ts	To	S	S	W	V	S	S	W
To	Ts	Tw	U	S	U	V	U	S	U
To	Ts	Tv	S	S	W	V	S	S	W
Tw	To	To	U	S	W	V	S	S	W
Tw	To	Tw	W	U	S	U	W	U	S
Tw	To	Tv	S	S	W	V	S	S	W
Tw	Tv	To	W	V	S	S	W	V	S
Tw	Tv	Tv	U	VS	W	V	S	S	W
Tw	Tv	Tw	W	V	S	S	W	V	S
Tw	Ts	To	S	S	W	V	S	S	W
Tw	Ts	Tv	S	S	W	V	S	S	W
Tw	Ts	Tw	U	S	W	V	S	S	W
Ts	To	To	U	V	S	S	W	V	S
Ts	To	Tv	S	U	W	U	S	U	W
Ts	To	Tw	W	V	S	S	W	V	S
Ts	Tv	To	W	V	S	S	W	V	S
Ts	Tv	Tv	U	V	S	S	W	V	S
Ts	Tv	Tw	W	V	S	S	W	V	S
Ts	Ts	To	S	S	W	V	S	S	W
Ts	Ts	Tv	S	S	W	V	S	S	W
Ts	Ts	Tw	U	VS	S	S	W	V	S

表 6: V0 を起点とした場合の戦略の変化

Pv	Pw	Ps	W0	V1	W2	V3	W4	V5	W6
To	To	To	U	U	U	U	U	U	U
To	To	Tv	U	WS	S	W	V	S	S
To	To	Tw	S	W	V	S	S	W	V
To	Tv	To	U	WS	V	S	S	W	V
To	Tv	Tv	U	U	VS	U	VS	U	VS
To	Tv	Tw	S	W	V	S	S	W	V
To	Ts	To	U	S	S	W	V	S	S
To	Ts	Tv	U	S	S	W	V	S	S
To	Ts	Tw	S	U	V	U	S	U	V
Tw	To	To	V	S	S	W	V	S	S
Tw	To	Tv	V	S	S	W	V	S	S
Tw	To	Tw	U	WS	U	WS	U	WS	U
Tw	Tv	To	V	S	S	W	V	S	S
Tw	Tv	Tv	V	S	S	W	V	S	S
Tw	Tv	Tw	U	WS	V	S	S	W	V
Tw	Ts	To	V	S	S	W	V	S	S
Tw	Ts	Tv	V	S	S	W	V	S	S
Tw	Ts	Tw	U	S	S	W	V	WS	S
Ts	To	To	U	W	V	S	S	W	V
Ts	To	Tv	U	W	U	S	U	W	U
Ts	To	Tw	S	W	V	S	S	W	V
Ts	Tv	To	U	W	V	S	S	W	V
Ts	Tv	Tv	U	W	V	S	S	W	V
Ts	Tv	Tw	S	W	V	S	S	W	V
Ts	Ts	To	U	WS	VS	WS	VS	WS	VS
Ts	Ts	Tv	U	WS	S	W	V	S	S
Ts	Ts	Tw	S	W	V	S	S	WS	V

表 7: W0 を起点とした場合の戦略の変化

[西崎 17] 西崎絵麻, 坂口早紀, 尾崎知伸, ワンナイト人狼における投票行動の分析, 第 31 回人工知能学会大会論文集, 2017.

[杉本 17] 杉本磨美, 伊藤毅志, 5 人人狼における村人の意思決定過程の研究, 日本認知科学会第 34 回大会論文集, pp.826-832, 2017.

[汪 17] 汪博豪, 大澤博隆, 佐藤健, 進化シミュレーションを用いた 3 人人狼の戦略分析, HAI シンポジウム 2017 論文集, 2017.

Organized Session | Organized Session | [OS] OS-14

[3F4-OS-14b] 人狼知能と不完全情報ゲーム(2)

稲葉 通将（広島市立大学）、片上 大輔（東京工芸大学）、狩野 芳伸（静岡大学）、大槻 恭士（山形大学）

Thu. Jun 6, 2019 3:50 PM - 5:30 PM Room F (302B Medium meeting room)

[3F4-OS-14b-01] Development and evaluation of the game agent to change confidence of estimating in Cooperative game Hanabi

〇Eisuke Sato[1], Hirotaka Osawa[1] （1. University of Tsukuba）

3:50 PM - 4:10 PM

[3F4-OS-14b-02] Role estimation in Werewolf games using a selective desensitization neural network

Masahiro Saito[2], Seiryu Mishina[1], 〇Ken Yamane[1] （1. Teikyo University, 2. Fujisoft Incorporated）

4:10 PM - 4:30 PM

[3F4-OS-14b-03] Artificial Intelligence for Deducing Roles of Players in the Werewolf Game using Information about Conversations among Players

〇Munemichi Fukuda[1], Hajime Anada[1] （1. Tokyo City University）

4:30 PM - 4:50 PM

[3F4-OS-14b-04] Extraction of Interpretable Rules for Role and Team Estimation in AI Wolf

Yuki Omura[1], Wataru Sakamoto[1], 〇Tomonobu Ozaki[1] （1. Nihon University）

4:50 PM - 5:10 PM

[3F06-09-5] Discussion / Conclusion

5:10 PM - 5:30 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

相手の思考時間の長短によって推定の信頼度を変更する協力ゲーム Hanabi のエージェントの開発及び評価

Development and evaluation of the game agent to change confidence of estimating in Cooperative game Hanabi

佐藤 栄介[*1]　大澤 博隆[*1]
Eisuke Sato　Hirotaka Osawa

[*1] 筑波大学
University of Tsukuba

Abstract: AI agent for cooperation with human need to follow human thought. There are some research using cooperating game 'Hanabi' from various aspects such as self-estimation, psychology, and communication theory. In this research, we developed an agent's strategy for following human thought for guessing human understanding of AI's strategy to utilize the length of thinking time of the human player and changing the estimation reliability. As a result, we found that this agent estimated more frequently to the good player of Hanabi than the conventional agent and there is a positive correlation between some evaluation to this agent and score. However, the agent that change estimation reliability depending on opponent's thinking time didn't affect human's impression compared to the conventional agent and there was no significant difference in the score and the success rate of the estimation by changing the reliability of the estimation according to the thinking time.

1. 序論

現在、人間と協力するエージェントの開発のため、協力ゲームの研究が注目されている。プログラムでゲームをプレイする分野において、人間に勝つことを目標としたエージェントは将棋や囲碁などで開発されており、その目標は達成されつつある。その一方で、人間と協力して一つの課題を達成するエージェントはまだ発展途上である。このようなエージェントが、人間に勝つことを目標としたエージェントと異なる部分は、人間側に行動の意図を汲んでもらうために人間の思考に合わせて人間に近い行動をとらなければならないところである。よって、このようなエージェントの開発は人工知能の人間らしさの獲得および人間の知能の人間らしさの解明に貢献すると考えられる。

協力ゲームをプレイする AI に関する研究の一つとして協力カードゲーム「Hanabi」を用いた研究が挙げられる。このゲームの大きな特徴は、相手の手札のみを見ることができて自分の手札はわからないことと、相手への情報提供がトークンによって制限されていることである。我々は Hanabi において他者の視点と行動をシミュレートして、自分の状態を推測するエージェントを開発した[1]。また、Eger らは心理学の手法とコミュニケーション理論から、人間がエージェントに対して何か意図を持って行動していると感じた時、人間はそのエージェントの能力を高く評価し好感を抱くということを示した[2]。

ただ、これらの既存の研究では人間側の思考時間という情報は Hanabi をプレイするエージェントの戦略において利用されていない。

本研究では Hanabi において、協力のために人間プレイヤーの理解度を思考時間から推定するエージェントの戦略を検討する。エージェントが協力ゲームをプレイする場合の戦略は協力相手の思考に合わせたものであるため、人間とプレイする場合には相手が自分のとった行動の意図を理解してくれることが前提になっている。相手がエージェントの行動の意図を理解していないと、お互いの行動がうまくかみ合わず、課題の失敗につながる。相手がエージェントの戦略を理解しているならば推定の信頼度を上げて推定を積極的に利用して行動するべきであり、逆に相手がエージェントの戦略をあまり理解していないならばエージェントは推定の信頼度を下げて確定情報を重視して行動するべきである。そのため、相手の理解度の推定はエージェントが戦略を決めるにあたって重要な問題である。

本研究ではエージェントの戦略に対するプレイヤーの理解度をプレイヤーが行動を選択するまでにかかる時間から判断することを試みた。また、実際に人間にこのエージェントと Hanabi をプレイさせることによって推定の信頼度を変更しない場合に比べて得点および人間のエージェントに対する評価がどのように変化するか調査した。

2. Hanabi の関連研究

協力ゲームをプレイするエージェントに関する研究は全体としてはまだ数が少ないが、近年では協力カードゲーム Hanabi に注目した研究が多く進められている。

我々は Hanabi において他者の視点と行動をシミュレートして、自分の状態を推測するエージェントを開発し、この戦略を用いることで確定情報のみに基づいて行動するエージェントより高い得点が得られることをシミュレーションにより示した[1]。また、Eger らは心理学の手法とコミュニケーション理論から、人間がエージェントに対して何か意図を持って行動していると感じた時、

[1] 佐藤栄介,筑波大学,hailabsec@iit.tsukuba.ac.jp

人間はそのエージェントの能力を高く評価し好感を抱くということを、エージェントと人間で Hanabi をプレイさせることにより示した[2]。ただ、これらの既存の研究では人間側の思考時間という情報は Hanabi をプレイするエージェントの戦略において利用されていない。思考時間は人間プレイヤーのゲーム及び相手に対する理解度の指標として重要なものである。そのため、これを利用することなしには人間と協力できるエージェントの開発は困難であると考えられる。

また、Christopher らは帽子推測ゲームを応用した Hanabi における戦略について考察し、事前に戦略を共有することにより 75%以上の確率で最高得点をとれることを示した[3]。しかしこれは事前に戦略を共有することを前提としており、またその内容も複雑であることから人間と協力してプレイする戦略としては適していない。

本研究では、事前の戦略の共有を行わずにゲーム中の相手の思考時間を用いてエージェントの戦略に対する人間プレイヤーの理解度を判断し、その程度に応じて推定の信頼度を変更するエージェントを開発し、実際に人とプレイすることによりその評価を行う。

3. Hanabi のルール

本研究で用いるゲーム「Hanabi」のルールを以下に示す。

Hanabi は協力型のカードゲームであり、通常 2～5 人で行う。今回の研究では 2 人でプレイする時のみを扱う。Hanabi は 50 枚のカードと 8 枚の情報トークンを用いて行うゲームである 50 枚のカードにはそれぞれ色と数字が設定されており、色は白、赤、青、黄、緑の 5 色、数字は 1～5 の 5 種類である。カードは各色 10 枚ずつ存在しその数字の内訳は 1 が 3 枚、2～4 が 2 枚、5 が 1 枚である。

このゲームの最終目的は盤面に各色のカードを 1～5 まで数字の昇順に並べていきすべての色で 5 までカードを並べきることである。ゲーム開始時、プレイヤーにはそれぞれ 5 枚のカードが手札として配られる。この時、自分の手札の情報は見ることができず、相手の手札の情報のみが観測可能である。残りの 40 枚は裏向きにして山札として扱う。各プレイヤーは自分のターンに情報提供、カードの廃棄、カードのプレイの三つの内いずれか一つの行動を選択して行うことができる。

情報提供では、相手プレイヤーの手札のカードの数字もしくは色を相手に教えることができる。この時、情報トークンを 1 消費する。情報トークンが 1 つもない場合、情報提供を行うことはできない。また、情報を教える際には該当するものはもれなく教えなければならない。例えば相手の手札が赤 1、赤 2、白 1、白 2、緑 1 であるときに赤色を教える場合、「1 番目と 2 番目のカードが赤である」というように教えることになり、どちらか一つのみの情報を教えることはできない。

カードの廃棄では、自分の手札のうち必要ないと考えた 1 枚を捨てることができる。その後、自分は山札からカードを 1 枚補充する。またこの時情報トークンが 7 枚以下ならば情報トークンが 1 枚増える。廃棄したカードは両プレイヤーに公開され、そのゲーム中は使用することができなくなる。

カードのプレイでは、自分の手札の内 1 枚を花火につなげることを試みることができる。プレイしたカードが盤面の同色の花火の数字より 1 大きい場合、プレイは成功となり、そのカードが盤面の該当する色の花火に並べられて数字が 1 増加する。プレイに失敗した場合、そのカードは廃棄される。成功した場合でも失敗した場合でもプレイしたカードは両プレイヤーに公開されてプレイヤーはプレイしたカードの代わりにカードを 1 枚山札から加える。

ゲームの終了条件は

・カードのプレイに 3 回失敗した場合

・山札が 0 枚になってから各プレイヤーが行動を一回ずつ行った場合

・すべての色の花火を 5 まで完成させた場合

の 3 つである。盤面の花火の数の合計がそのまま得点になり、最高得点は 25 点である。

4. エージェントのアルゴリズム

本研究では相手プレイヤーの思考時間を戦略の指標に用いない従来の自己推定エージェントと相手プレイヤーの思考時間を戦略の指標に用いる新しい自己推定エージェントの 2 種類を実装した。

4.1 従来の自己推定戦略

従来の自己推定戦略は加藤らが開発したアルゴリズムであり[4]、以下の優先順位で行動を行う。

(1) 数の確定情報のみに基づくカードのプレイ

色が不明だが、数字からプレイ可能カードであることが分かるカードをプレイする。例えば、ゲーム開始直後の状態の時、1 のカードは色にかかわらずプレイ可能カードになる。

(2) プレイ可能カードの情報提示

相手がプレイ可能カードを持つことが分かり、そのカードの情報が相手にとって不完備な場合、情報提示を行う。数も色も情報が不完備な場合、教える情報をランダムに選択する。

(3) 自己推定に基づくプレイ可能カードのプレイ

相手の観測のシミュレーションを用いて自己の手札の推定を行う。この時考えられる手札の可能性の集合を考え、この中におけるカードの要素を出現数でソートし、1 番多い出現数の値が次に多い出現数の値に対して a 倍以上のとき、このカードは 1 番多い出現数の値に等しいと推定する。この値がプレイ可能カードの値に等しい場合にプレイする。

(4) 廃棄可能カードの情報提示

相手が廃棄可能カードを持つことが分かり、数または色を教えることによって相手視点からも廃棄可能角であると断定できる場合、その情報を提示する。

(5) プレイ可能カードのプレイ

自分の手札の情報と盤面の観測可能なカードからプレイ可能カードが存在することがわかる場合、そのカードをプレイする。

(6) 廃棄可能カードの廃棄

自分の手札の情報と盤面の観測可能なカードから廃棄可能カードが存在することがわかる場合、そのカードを廃棄する。

(7) ランダム廃棄

自分の手札の中で、最も情報量の少ないカードを廃棄する。最も情報量の少ないカードが複数存在する場合、その中からランダムで廃棄する。

4.2 思考時間に応じて推定の信頼度を変える自己推定戦略

本研究で開発した自己推定戦略は加藤らのアルゴリズムに追加し、相手プレイヤーの思考時間に応じて推定の信頼度を変更するようにした。まず、相手プレイヤーの各ターンにおける行動選択にかかる時間を直前の 5 ターン分記録する。そこから思考時間の平均と標準偏差を導き出し、次の思考時間が平均

＋1 標準偏差以上だったら相手は長考していると定義し、思考時間が平均－1 標準偏差以下だったら相手は即決していると定義する。また、直前 5 ターンの思考時間の偏りによっては平均より標準偏差の方が大きくなり、即決であると定義できる時間がなくなってしまうことが起こりうる。そのため、相手プレイヤーが 4.13 秒以下の思考時間だった場合、平均や標準偏差にかかわらず即決であると定義する。この値は加藤が以前大学生 12 人を対象として行った Hanabi の実験における各プレイヤーの平均思考時間－1 標準偏差の平均をとったものである[4]。この値は、各プレイヤーが即決したと判断される閾値の平均であるため、個人の直前の思考時間の平均や標準偏差にかかわらず即決であると判断できる閾値として妥当であると考えられる。

ここで思考時間の長短に対して以下の仮説を立てる。予備実験の結果、これらの妥当性は示された。

・長考している時、相手プレイヤーは自身の行動に自信がない
・即決している時、相手プレイヤーは自身の行動に自信がある。

相手が行動の選択において長考を続けるとき、相手プレイヤーは選択に自信が持てていない、すなわちエージェントの戦略があまり理解できていないと考えられる。この時にエージェントが推定を行うと、通常よりも推定に失敗する確率が高くなると考えられる。なぜならエージェントの自己推定は相手が自分のシミュレーション上の相手と同じように行動を起こすことが前提となっており、そのためには相手がエージェントの戦略をある程度理解していることが不可欠だからである。そのため、このエージェントは 3.1 の「3.自己推定に基づくプレイ可能カードのプレイ」で示したように推定で考えられる最も多いカードの値が次に多いカードの値の a 倍以上なら推定に従うというアルゴリズムだが、相手が長考を 2 連続で行った場合はその a の値を大きくすることにより推定の信頼度を下げる。

一方相手が行動の選択において即決を続けるとき、相手プレイヤーは選択に自信を持っていると考えられるがその状況は大きく 2 つに分けられる。1 つはエージェントの戦略を正しく理解できているときで、もう 1 つはエージェントの戦略を誤って解釈しており、その誤った解釈に基づいて行動を進めている場合である。前者の場合、AI は推定の信頼度を上げるべきだが、後者の場合は推定の信頼度を下げるべきである。そのため、この 2 つの状況を判別するために相手プレイヤーのカードプレイの結果を利用した。具体的には、相手プレイヤーが即決を 2 連続で行った場合は推定の信頼度を上げるが、相手が即決でかつカードのプレイに失敗したときには推定の信頼度を下げるようにした。

5. 評価

人間・AI 間のゲームにおいて相手の思考時間の長短を推定の信頼度の変化の指標として用いるエージェントを用いた場合のゲームの得点に与える影響と、人間のエージェントに対して抱く印象を評価するために実験を行った。

5.1 インタフェース

人間と AI 間のゲームを実現するために加藤らが開発したインタフェースを用いた[4]。これを図 1 に示す。
画面左下には自分の手札、画面左上には相手の手札が表示される。中央上は現在の盤面を表し、中央下の数は左から順番に山札の枚数、情報トークンの枚数、プレイの失敗回数、得点を示す。また、右にはプレイする、プレイに失敗する、廃棄する行為によって観測可能になったカードが表示される。

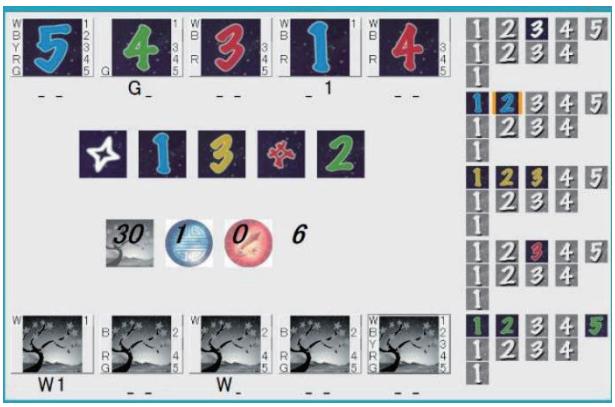


図 1 Hanabi のインタフェース

5.2 評価方法

本実験では相手のエージェントに対する印象を調べるために、ゲームの記録と実験参加者の様子の撮影と実験後の印象評価アンケートを行った。以下にアンケートの質問内容を示す。

(1) 相手についての印象

Q1. 意思を持っていると感じたか
Q2. ゲームに慣れていたか
Q3. 相手の行動の意図が分かったか
Q4. 相手の行動に一貫性を感じたか

(2) ゲームについての印象

Q5. ゲームはやりやすかったか
Q6. ゲームの結果に満足しているか
Q7. 自由記述欄

本アンケートの設問のうち、
Q1. 意思を持っていると感じたか
Q2. ゲームに慣れていたか
この 2 つの項目は予備実験において有意差が表れたものである。また、
Q7. 相手の行動の意図が分かったか
Q8. 相手の行動に一貫性を感じたか
という 2 つの項目は人間が 2 種類のエージェントの差を実験中に認識しているかを判断するために重要な役割を果たすと推測した。よって、本実験ではこれら 4 問にゲームに関する質問 2 問を加えた 6 問を設問として用いた。
Q7 の自由記述を除くすべての質問は予備実験と同様にそう思わないを 1、そう思うを 7 とした 7 段階のリッカート尺度で回答する形式を用いた。

5.3 実験手順

実験には 20 代大学生男女 20 人が参加した。参加者は Hanabi のルールとインタフェースについて説明を受けた後、ゲームを 4 回行った。従来の自己推定エージェントとのゲーム(以下 A 条件)、相手の思考時間に応じた推定の信頼度の変化を加えた新しい自己推定エージェントとのゲーム(以下 B 条件)の 2 条件のゲームを 1 回ずつ行い、その後、順番を変えて 1 回ずつゲームを行った。よって 1 人の実験参加者につき、A 条件、B 条件のサンプルの数はそれぞれ 2 つになる。実験に用いる山札は 4 種類用意し、1 人の実験参加者が 4 種類の山札でゲームを行うようにした。ゲーム終了毎にアンケートを行い、最後にインタビューを行って実験を終了した。

6. 結果

本実験では A 条件と B 条件のそれぞれでゲームを 40 回ずつ行った。よって最終的に A 条件、B 条件のデータはそれぞれ 40 個存在する。

エージェントの戦略を理解している人への本研究のエージェントが与える影響を調べるためにデータを分類した。人間側の AI の戦略に対する理解度の指標として得点を用いて、18 点以上のものを戦略を理解しているゲーム(以下 A(h)、B(h)とする)と定義して、A 条件,B 条件それぞれにおいて分類した。この閾値は加藤の実験結果における Hanabi の平均得点 17.5 点から決定した。結果、A(h)条件のデータは 21 個、B(h)条件のデータは 24 個になった。

A(h)条件においてカードの推定は 1 ゲームで平均 5.38 回(SD2.79)行われ、その成功率は平均.728 (SD.196)であった。B(h)条件においてカードの推定は 1 ゲームで平均 7.58 回(SD4.39)行われ、その成功率は平均.744 (SD.208)であった。t 検定の結果、2 条件間で推定回数に有意差が見られた($p<.05$)。

アンケートの結果を表 3 に示す。表 3 より、アンケート結果には 2 条件間で有意差が見られなかった。

表 1. アンケートの結果

	A(h)条件平均	B(h)条件平均	p 値
Q1	5.14 (SD 1.93)	5.41 (SD 1.41)	.604
Q2	5.38 (SD 1.05)	5.08 (SD 1.44)	.439
Q3	5.66 (SD 0.94)	5.83 (SD 1.37)	.641
Q4	5.19 (SD 1.10)	5.08 (SD 1.68)	.803
Q5	5.52 (SD 1.40)	4.82 (SD 1.75)	.516
Q6	5.00 (SD 1.23)	4.79 (SD 1.55)	.627

アンケートの評価と得点の相関を調べた結果を下の表 4 に示す。表 4 より、A(h)条件においては Q2 の評価と得点に正の相関があった。また、B(h)条件においては Q2,Q3,Q4,

Q5,Q6 の評価と得点に正の相関があった。($r>.4$)

表 2. 相関係数

	A(h)相関係数	B(h)相関係数
Q1	.349	0.154
Q2	.508*	0.546*
Q3	.249	0.498*
Q4	-.050	0.435*
Q5	.208	0.556*
Q6	.198	0.506*

* $r>.4$

7. 考察

本実験の得点と推定の成功率には A(h),B(h)2 条件間に有意差は見られなかった。このことから、本研究で開発した、思考時間を戦略の指標に取り入れたエージェントは得点においては影響を及ぼさないといえる。

推定の回数においては B(h)条件が A(h)条件に比べて有意に高かった。これは、エージェントの戦略を理解した実験参加者は即決を連続で行う場面が多く、結果として推定の信頼度が上がったゲームが多くなったからであると考えられる。また推定の成功率は変わっていないことから、推定回数を増やしたことによって推定の精度が変化することはなかったといえる。

両条件においてアンケート結果の有意差は見られなかった。このことから、本研究で開発した思考時間を戦略の指標に取り入れたエージェントは人への印象に影響を与えるかどうかはわからなかった。

両条件においてエージェントがゲームに慣れていると感じるプレイヤーは高得点を取る傾向があるといえる。これは、エージェントがゲームに慣れていると感じるときに、人間プレイヤーはエージェントとより協力して行動できているからであると考えられる。また、B(h)条件においてのみ、エージェントの行動の意図を理解した場合、エージェントの行動に一貫性を感じた場合、ゲームがやりやすいと感じた場合、結果に満足していると感じた場合にも高得点を取る傾向があるといえる。これは、これらの項目の評価が高い時、人間プレイヤーはエージェントの戦略を正しく理解しており、その結果として即決の行動が多くなり、エージェントの推定の信頼度が上がり、より協力できるようになったからであると考えられる。

8. 結論

本研究では、協力ゲーム Hanabi において、相手の思考時間から相手の行動選択の自信の有無を把握し、戦略における推定の信頼度を変更する指標として用いたエージェントを開発して、従来の自己推定戦略と比べた得点の変化、および人間に与える印象の違いを調査した。

結果として、相手の思考時間の長短に応じて推定の信頼度を変えるエージェントは、従来のエージェントと比較して得点と推定の成功率に有意差は見られなかったが、推定の回数は増加するため結果としてより多くの推定を成功させることが分かった。また、アンケートの結果より、相手の思考時間の長短に応じて推定の信頼度を変えるエージェントにおいては、エージェントに対する評価やゲームに対する満足度が高いと得点も高い傾向にあることが分かった。

今後の課題として思考時間をより適切に戦略の指標に組み込んだエージェントの開発が考えられる。人間の思考時間に関する仮説をより複雑にして人間の本来の思考に近いものにして、思考時間の長短の定義の仕方をより適切なものに変更することによりこの課題の達成が期待される。

謝辞

本研究は JSPS 科研費 JP26118006, JP18KT0029 の助成を受けたものです。

参考文献

[1] Osawa, H. (2015). “Solving Hanabi : Estimating Hands by Opponent’s Actions in Cooperative Game with Incomplete Information.” AAAI Workshop, Computer Poker and Imperfect Information, (pp. 37–43).

[2] Eger, M., Martens, C., & Córdoba, M. A. (2017). “An intentional AI for Hanabi” IEEE Conference on Computational Intelligence and Games (CIG) (pp. 68-75).

[3] Christopher Cox, Jessica De Silva, Philip Deorsey, Franklin H. J. Kenter, Troy Retter & Josh Tobin (2015) “How to Make the Perfect Fireworks Display: Two Strategies for Hanabi” Mathematics Magazine (pp. 323-336)

[4] 加藤 拓也, 大澤 博隆 ”他者視点の推定を用いたエージェント評価の向上:カードゲーム Hanabi の AI との対戦分析から”人工知能学会全国大会論文集, 第 32 回全国大会,4J2-01

[5] Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. International Journal of Social Robotics, 1(1), (pp.71–81)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

選択的不感化ニューラルネットを用いた人狼ゲームにおける役職推定

Role estimation in Werewolf games using a selective desensitization neural network

齋藤 正宏 *1 Masahiro Saito　三品 晟瑠 *2 Seiryu Mishina　山根 健 *2 Ken Yamane

*1富士ソフト株式会社 Fujisoft Incorporated　*2帝京大学 Teikyo University

Estimating a player's role is important to win Werewolf games. As described in this paper, we propose a role estimation method incorporating a selective desensitization neural network. We evaluated its accuracy using a model constructed for 15-person games. Experimentally obtained results show that the model estimated six roles with 69% accuracy.

1. はじめに

人狼ゲームをプレイする人工知能（以下，人狼知能）において，他プレイヤの陣営や役職を推定する能力が勝敗を左右する要素の一つであるといわれる [梶原 16, 大川 17]．そこで，プレイヤの役職を推定する方法が幾つか提案されている．例えば，サポートベクタマシン（以下，SVM）を用いて 15 人人狼における推定方法が提案されている [梶原 16]．また，梶原らの特徴にさらに 3 つの特徴を追加して，深層学習を用いた推定方法が比較されており．5 人人狼において推定精度はよいもので 72%であった．

人狼ゲームは不完全情報ゲームであるため，確かに推定精度 100%の実現は不可能である．しかし，さらなる精度向上が望まれ，特に有利な能力をもつ占い師などの役職や排除するべき人狼を精度よく推定できるとよい．また，人狼知能大会プロトコル部門 [人狼] では 15 人で行うゲームもあり，人数や役職が増加した場合にも対応できる必要がある．さらに，学習に膨大なデータや時間を必要とする方法ではプレイヤの構成や戦略が変化した場合に対応が難しく，ゲーム毎あるいはゲーム中に学習できるような手軽な学習方法が求められる．

そこで本研究では，これらの要求を満たすような役職推定を実現するべく，強力な類推能力をもつ選択的不感化ニューラルネット（以下，SDNN）[新保 10] を用いて他プレイヤの役職を推定する方法について検討する．なお，本研究では 15 人でプレイする人狼ゲームを対象とする．また，ゲーム終了後に他のプレイヤの役職が判明するとして，全ての情報を得ている神のような視点ではなくプレイヤの視点で役職推定を行う．

2. 役職推定方法

2.1 人狼ゲームデータ

人狼知能プロジェクト [人狼] では，人狼知能プレ大会 GAT2018 に出場したエージェントのソースコードが公開されている．これらを利用して人狼ゲームをプレイさせる．なお，強さや勝率には大きなばらつきがあるが，様々な強さのエージェントが含まれるように選択する．

本研究では 15 体のエージェントを準備する．その内訳は，同一のもの 4 体，残りの 11 体は異なるものとする．ただし，11 体中の 1 体は著者らが作成したもので，これを役職推定する主体とする．これは GAT2018 に参加したエージェント「neko」を，以下の点で改造してたものである．

- 村人陣営の場合，初日から最も人狼と疑われるエージェントに投票する
- 霊媒師の場合，必ず 2 日目に霊媒した結果とともに霊媒師告白する
- 人狼の場合，霊媒師が 1 人しか告白していない時にそのエージェントを襲撃する

学習や性能評価には，人狼知能大会プロトコル部門で用いられているデータ形式に基づいたゲームログを用いる．110 回のゲームログから，100 回分を訓練データ，10 回分をテストデータとする．基本的に，全プレイヤに公開されている会話の部分のみから特徴を抽出して入力変数として扱う．

2.2 入力変数の構成

プレイヤの役職を推定する際に役立ちそうな特徴を抽出して入力変数とする．具体的には，梶原らが用いた (1) 経過日，(2) 告白した占い師の人数，(3) 占い師から人間判定を受けた回数，(4) 占い師から人狼判定を受けた回数，(5) 占い師告白した順番，(6) 占い師が人間判定を出した回数，(7) 占い師が人狼判定を出した回数，(8)vote 発言から投票先を変更した回数の 8 変数，そして大川らが追加した (9) 生死状態，(10) 賛成意見の数，(11) 否定意見の数の 3 変数を用いる．

予備実験として，先行研究 [大川 17] と同じ 11 変数を入力に SVM の推定器を用いて実験した結果，推定精度は 60.5%であった．精度が 10%以上大きく低下したのは，SVM の構成方法や最適化方法が異なるだけでなく，プレイヤの人数が増えた（役職の種類が増えた）ことやプレイヤの構成が異なっているためだと考えられる．

そこで，さらに (12) 村人告白した順番，(13) 狩人告白した順番，(14) 霊媒師告白した順番，(15) 狂人者告白した順番，(16) 人狼告白した順番，(17) 告白した役職，(18) 推定主体の役職，(19) 発言回数，(20)estimate された役職，(21)estimate 発話の回数を加えた合計 21 変数を用いることにした．ただし，この中に冗長な次元が含まれていたとしても積極的に排除することはせずに，役職推定器がどのように扱うのか調べる．

これらの特徴は 0〜1 の値に正規化され入力変数 $\boldsymbol{x}$ として用いる．正規化方法については改善の余地があるが，比較対象

連絡先: 山根 健，帝京大学 理工学部 情報電子工学科，〒 320-8551 栃木県宇都宮市豊郷台 1-1，028-627-7224，yamane@ics.teikyo-u.ac.jp

とする SVM において精度が高いものを選択した．なお，実際に生成した訓練データには入力変数 $\boldsymbol{x}$ は同じであるが出力 y（役職）が異なるようなものが 30.4%含まれ，学習時に悪い影響を与える可能性がある．

2.3 選択的不感化ニューラルネットを用いた役職推定

本研究では，shallow な層状ニューラルネット SDNN を用いて役職推定を行う．計算など詳細については参考文献 [新保 10] を参照されたい．以下では，構成方法の概要を説明する．

入力層では，入力変数 $\boldsymbol{x}$ を成分 $-1, 1$ で構成される高次元 2 値パターンとして分散表現する．具体的には，ある変数 x^i $(i = 1, 2, 3, \cdots, 21)$ を 100 個の素子で表現する（これを x^i の素子群とよぶ）．従って，$21 \times 100 = 2,100$ 個の素子が並ぶ．

中間層では，2 つの素子群が表す情報を選択的不感化法を用いて分散表現のまま統合する．その結果，成分 $-1, 0, 1$ で構成される 3 値のパターンになる．また，素子群の組み合わせの数だけの表現が中間層に並ぶことになる．ただし，自分自身に対してはこの表現を作らないため，$21 \times 20 \times 100 = 42,000$ 個の素子が並ぶ．

ネットワークを動作させて得られる出力層の出力パターンと 6 種類の役職（村人，狩人，霊媒師，占い師，狂人，人狼）を表すコードパターンそれぞれとの方向余弦を計算して出力パターンを解釈する．なお，出力層の素子数は 100 個として，それぞれの役職を表すコードパターンは成分 $-1, 1$ から構成される互いに無相関なパターンとする．

学習フェーズでは，ある入力変数 $\boldsymbol{x}$ に対して正しい役職を表すパターンが出力されるように，中間層の素子から出力層の素子への荷重 $\boldsymbol{w}$ を調整する．この調整には，簡単な学習方法を用いることができ，以下の実験では誤り訂正学習を用いる．この時の学習回数は 10 回，学習係数 $c = 0.01$ とする．

3. 推定実験

実験は，CPU Intel Core i7-3630QM（4 コア 8 スレッド），メモリ 8GB を搭載した Windows マシンを用いて実施した．また，SVM の実装には機械学習用ライブラリ [LIBSVM] を，SDNN は自作ライブラリを利用した．

まず，SVM を用いた推定結果を表 1 に示す．先行研究と同様の方法（精度 60.5%）と比べると精度の向上が見られるため，本研究で新たに追加した特徴 (12)〜(21) が有効であることがわかった．次に，SDNN の結果を表 2 に示す．なお，コードパターンの生成などにランダム要素が入っているため，ランダムシードを変えて 10 回実験し，精度が最も良い結果（best）と最も悪い結果（worst）を示している．10 回の平均精度は 68.6%であり，全体として SVM より少し精度が高かった．一方で，人狼の推定では SVM より精度が低かった．

最適化方法などを考慮すると単純な学習時間の比較が難しい．SVM については，グリッドサーチによってパラメータ探索を行うとともに，訓練データでクロスバリデーションを行なった．その結果，100 ゲームのデータに対して学習に約 7 時間ほどかかった．一方で，SDNN については，全く別のパターン認識タスクで用いたパラメータをそのまま利用して（パラメータ探索を行わないで），2 時間半程度であった．

さらに，訓練データが少ない場合についての性能を調べるため，10 ゲームのデータで学習して，別の 10 ゲームのデータでテストを行なった．その結果，SVM の精度が 49.8%だったのに対して，SDNN の精度は 1 回学習（約 2 分）しただけで 62.6%であった．この結果から，SDNN の方が少ない訓練データでも効率よく学習できることがわかった．

表 1: SVM を用いた推定結果
SVM（RBF, $\gamma = 0.5$, $cost = 32$）
精度：66.6%

役職	適合率	再現率
村人	69.5%	86.7%
狩人	9.5%	3.0%
霊媒師	82.1%	48.5%
占い師	70.2%	50.0%
狂人	76.1%	77.3%
人狼	85.9%	41.9%

表 2: SDNN を用いた推定結果

	best 精度：69.4%		worst 精度：67.6%	
役職	適合率	再現率	適合率	再現率
村人	66.5%	98.1%	66.1%	96.6%
狩人	0.0%	0.0%	0.0%	0.0%
霊媒師	97.8%	68.2%	88.2%	68.2%
占い師	79.0%	45.5%	73.1%	28.8%
狂人	74.0%	86.4%	61.0%	92.4%
人狼	82.2%	18.7%	82.9%	17.2%

4. まとめ

SDNN を用いて役職推定を行う方法を提案し，ゲームデータを使って性能を検証した．その結果，SVM に比べて，SDNN の推定精度が高いことを示した．また，パラメータ探索などを含めた学習時間についても SDNN は効率的であり，少ない訓練データでも高い精度を示した．この結果から，SDNN を用いて使い勝手の良い役職推定器を構築できると考えられる．

今後の課題として，人狼の推定精度を向上させる方法を検討し，学習済みネットワークの荷重 $\boldsymbol{w}$ を分析するとともに，推定結果に基づいて行動を決定する方法を検討する．

参考文献

[大川 17] 大川貴聖，吉仲亮，篠原歩，“深層学習を用いて役職推定を行う人狼知能エージェントの開発，”第 22 回ゲームプログラミングワークショップ論文集，pp.50–55, (2017).

[梶原 16] 梶原健吾，鳥海不二夫，稲葉通将，大澤博隆，片山大輔，篠田孝祐，松原仁，狩野芳伸，“人狼知能大会における統計分析と SVM を用いた人狼推定を行うエージェントの設計，”第 30 回人工知能学会全国大会，JSAI2016 2F4-1, pp.1–4 (2016).

[新保 10] 新保智之, 山根健, 田中文英, 森田昌彦:選択的不感化ニューラルネットを用いた強化学習の価値関数近似, 信学論 D, Vo.J93-D, No.6, pp.837-847 (2010).

[人狼] 人狼知能プロジェクト，http://aiwolf.org/，最終アクセス：2019 年 1 月 24 日.

[LIBSVM] LIBSVM, https://www.csie.ntu.edu.tw/~cjlin/libsvm/, 最終アクセス：2019 年 1 月 24 日.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

人狼ゲームにおける会話情報による役職推定

Artificial Intelligence for Deducing Roles of Players in the Werewolf Game using Information about Conversations among Players

福田 宗理[*1] Fukuda Munemichi　　穴田 一[*1] Hajime Anada

[*1] 東京都市大学大学院 総合理工学研究科
Graduate School of Integrative Science and Engineering, Tokyo City University #1

Recently, the Artificial Intelligence based Werewolf Project has been attracting attention. Because, it is important for human-like artificial intelligence research to construct an artificial intelligence for the Werewolf Game which requires abilities to bamboozle and detect a lie. Therefore we construct an artificial intelligence for deducing roles of players in the Werewolf Game using information about conversations among players.

1. はじめに

近年，人工知能による将棋や囲碁などのゲームの大会が開催されている．その 1 つに人狼ゲームを行う人狼知能大会[1]がある．人狼ゲームで勝つためには嘘をつく能力や，情報の真偽を見極める能力が要求される．これらの能力が向上することで，人工知能がより高度な判断が可能となり，人間に近づくと考えられる．そのため，人狼知能大会が注目されている．

大川らの研究[2]では 5 人人狼で勝つために，各プレイヤの役職を推定するニューラルネットワーク(NN)を構築した．しかしそのモデルには人狼の推定精度を向上できる余地があると考えられる．何故なら，NN に入力する特徴として，どのプレイヤを対象にして発言したかは用いているが，発言の対象となるプレイヤの特徴を用いていないため，上手く役職の特徴を学習できず，推定が困難になっていると考えられるためである．

そこで本研究では，経過日数と追放投票で各プレイヤが投票した対象，CO した役職とその時の経過日数とターン数，全ての会話情報を考慮し役職推定を行う．

2. 既存研究

2.1 人狼大会のルール

人狼ゲームは村陣営と狼陣営に分かれて，それぞれの陣営で勝ちを競うゲームである．村陣営は，人狼を全てゲームから除外する(以下，追放する)ことが勝利条件であり，狼陣営は，人間の数を人狼の数以下にすることが勝利条件である．

人狼大会では人狼ゲームを 5 人か 15 人で行い，本研究では 5 人人狼を取り扱う．5 人人狼の場合は各プレイヤは役職(村人，占師，人狼，裏切り者)，種族(人間，狼)，陣営(村，狼)の属性を持つ．またプレイヤの役職と種族，陣営は他のプレイヤに非公開である．

村陣営は全て種族が人間の村人 2 人，占師 1 人である．村人は能力を持たない．占師は 1 日 1 人のプレイヤの種族を調べること(以下，占い)ができる．狼陣営は人狼 1 匹で種族は狼，裏切り者 1 人で種族は人間である．人狼は 1 日 1 人のプレイヤをゲームから除外できる(以下，襲撃)．裏切り者は狼陣営が勝った時に共に勝利する特殊な役職である．

人狼ゲームには，昼と夜のフェイズが存在する．昼はプレイヤ同士で会話し，最後に追放投票をする．狼陣営は時折嘘の発言をしてプレイヤを騙し，都合の良い展開に持ち込もうとする．村陣営は狼陣営の嘘を見抜き，逆に人狼を見つけようとする．追放投票は全員で投票を行い，最多票のプレイヤが追放される．夜は占師が 1 人を占い，人狼が 1 人を襲撃する．また人狼ゲームはゲーム開始前に占師が一度占いを行い，ゲームを開始する．

プレイヤ同士の会話はターン制である．発言は定型文で 1 日に 10 回まで可能であり，発言をしないことも可能である．全員が 1 回ずつ定型文を発言することを 1 ラウンドとし，これを複数ラウンド繰り返す．発言する順番はラウンド毎に無作為に決定する．

発言に用いる定型文の種類を表 1 に示す．

表 1:発言の定型文

発言	意味
ESTIMATE P R	P は R だと思う
COMINGOUT P R	P は R であると宣言する
DIVINATION P	P を占いたい
DIVINED P S	P の占い結果は S である
VOTE P	P に投票する予定である
AGREE T	T に賛成する
DISAGREE T	T に反対する
REQUEST (P)T	(P)に T をして欲しい
SKIP	パス
OVER	今日はもう喋らない

ここで，P はプレイヤ，R は役職，S は種族，T は発言を表している．

以下の 3 つの内どれかの条件を満たすと会話が終了し，追放投票に移行する．

- 全てのプレイヤが発言 OVER をする．
- 全てのプレイヤが発言 SKIP をするラウンドが 3 回連続する．
- 20 ラウンドが経過する．

2.2 大川らの役職推定モデル

人狼ゲームに強い人工知能を作る為には，プレイヤの会話や占い結果の情報から誰が人狼かを精度よく推定する事が重要である．大川らの研究では，以下の表 2 に示すプレイヤの特

連絡先:福田宗理，東京都市大学，〒158-8557
東京都世田谷区玉堤 1-28-1

徴から各プレイヤの役職を推定する NN を構築している. この NN ではプレイヤの役職が村人, 占師, 人狼, 裏切り者である確率をそれぞれ出力する.

表 2:プレイヤ X の特徴

特徴	詳細
経過日数	現在何日目か
占師の数	占師 COMINGOUT(以下 CO)をしたプレイヤの数
被占い結果	X が人間判定された数と人狼判定された数
何番目の占師	(占師 CO をしたプレイヤのみ) X が何番目に占師 CO をしたか
占い結果	(占師 CO をしたプレイヤのみ) X が報告した人間判定の数と人狼判定の数
投票変更数	X が発言 VOTE の発言の対象にしたプレイヤと, X が行った投票の対象が異なった回数
生死	X の状態が生存, 追放された, 襲撃されたかのいずれか
肯定的意見の数	X が別のプレイヤ Y に対して村陣営であると推定した, または Y の発言で発言 AGREE をした数
否定的意見の数	X が別のプレイヤ Y に対して狼陣営であると推定した, または Y の発言で発言 DISAGREE をした数

2.3 既存手法の問題点

NN に入力するプレイヤの特徴として占い結果を用いているが, それは占い結果の数のみであり, どのプレイヤをなぜ占ったのかは用いていない. さらに肯定的意見の数と否定的意見の数は, どのプレイヤを対象にして発言したかは用いているが, その発言の対象のプレイヤの特徴を用いていないことが, 推定精度が良くない主要な要因であると考えられる.

3. 提案手法

本研究では, 5 人人狼で勝つために, 経過日数と追放投票で各プレイヤが投票した対象, CO した役職とその時の経過日数とターン数, 全ての会話情報を NN に入力した役職推定モデルを提案する.

入力の内容は 6 種類ある.

- 経過日数
- 追放投票で各プレイヤが投票した対象
- 各プレイヤの CO した役職とその時の経過日数, ターン数
- 推定対象者から他のプレイヤへの発言内容
- 他のプレイヤから推定対象者への発言内容
- 他のプレイヤ同士の発言内容

このモデルを自分以外の推定対象者のプレイヤ 1 人 1 人に使い, 各プレイヤの役職を推定する.

提案手法の入力を以下の表 3 に示す.

4. 実験

提案手法と既存手法の役職推定モデルの精度, 再現率を比較し, 性能の評価をする. GAT2017人狼知能大会のログデータを用い, データの半分を学習データ, もう半分をテストデータとした.

結果と考察の詳細は発表時に述べる.

表 3:経過日数と CO 状況, 会話を入れた情報

特徴	詳細
経過日数	現在何日目か
追放投票で各プレイヤが投票した対象	追放投票で各プレイヤが投票した対象
各プレイヤの CO した役職とその日時	各プレイヤが CO した役職とその時の経過日数, ターン数を記録
推定対象者から他プレイヤへの推定発言	推定対象者から他プレイヤへどの役職で, 発言 ESTIMATE したか
推定対象者から他プレイヤへの投票発言	推定対象者から他プレイヤへ, 発言 VOTE したか
推定対象者が他プレイヤを占いたいと発言	推定対象者から他プレイヤへ, 発言 DIVINATION したか
推定対象者が他プレイヤを占った結果を発言	推定対象者から他プレイヤへ, どういう発言 DIVINED したか
推定対象者が他プレイヤの発言に同意する	推定対象者から他プレイヤへ, 発言 AGREE したか
推定対象者が他プレイヤの発言に反対する	推定対象者から他プレイヤへ, 発言 DISAGREE したか
他プレイヤから推定対象者への推定発言	他プレイヤから推定対象者へどの役職で, 発言 ESTIMATE したか
他プレイヤから推定対象者への投票発言	他プレイヤから推定対象者へ, 発言 VOTE したか
他プレイヤが推定対象者を占いたいと発言	他プレイヤから推定対象者へ, 発言 DIVINATION したか
他プレイヤが推定対象者を占った結果を発言	他プレイヤから推定対象者へ, どういう発言 DIVINED したか
他プレイヤが推定対象者の発言に同意する	他プレイヤから推定対象者へ, 発言 AGREE したか
他プレイヤが推定対象者の発言に反対する	他プレイヤから推定対象者へ, 発言 DISAGREE したか
他プレイヤから別の他プレイヤへの推定発言	他プレイヤから別の他プレイヤへどの役職で, 発言 ESTIMATE したか
他プレイヤから別の他プレイヤへの投票発言	他プレイヤから別の他プレイヤへ, 発言 VOTE したか
他プレイヤが別の他プレイヤを占いたいと発言	他プレイヤから別の他プレイヤへ, 発言 DIVINATION したか
他プレイヤが別の他プレイヤを占った結果を発言	他プレイヤから別の他プレイヤへ, どういう発言 DIVINED したか
他プレイヤが別の他プレイヤの発言に同意する	他プレイヤから別の他プレイヤへ, 発言 AGREE したか
他プレイヤが別の他プレイヤの発言に反対する	他プレイヤから別の他プレイヤへ, 発言 DISAGREE したか

参考文献

[1] 人狼知能プロジェクト http://aiwolf.org/

[2] 大川貴聖, 吉仲亮, 篠原歩:“深層学習を用いて役職推定を行う人狼エージェントの開発” The22nd Game Programming Workshop 2017

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

人狼ゲームにおける明示的役職・陣営推定理由の抽出

Extraction of Interpretable Rules for Role and Team Estimation in AI Wolf

小村 友希 Yuki Omura　坂本 航 Wataru Sakamoto　尾崎 知伸 Tomonobu Ozaki

日本大学 文理学部
College of Humanities and Sciences, Nihon University

The Werewolf game is a conversation-based multi-player incomplete information game. In the development of high performance AI agents for playing the werewolf games, it is important to explicitly understand the grounds and reasons for their judges and behaviors. In this paper, we report the initial results on the explicit rules extraction for the role and team estimations from agent logs, by employing an algorithm for constructing an interpretable model from tree emsambles.

1. はじめに

不完全情報ゲームの一つである人狼ゲームでは，各プレイヤがそれぞれ役職を持ち，村人陣営と人狼陣営に分かれてゲームを行う．ゲームは会話を中心に行われ，他プレイヤを騙すための嘘をつく，他プレイヤの嘘を見抜く，他プレイヤを説得し信頼を獲得するなど，非常に高度な行動や駆け引きと，それらを裏付けるための高次推論が重要となる．

人狼知能 [1, 2]，すなわち人狼をプレイする人工知能エージェントに関する研究は，これまで様々な方面から多様な研究が展開されている．特にプレイヤの役職推定に関しては，勝敗を左右する基本的な能力の一つであると認識され，精力的に研究が行われている．例えば文献 [3] では，役職推定問題を制約充足問題の枠組みで定式化している．また文献 [4] では，潜在的ディリクレ配分法 (LDA) を用いてトピック分析を行い，トピックごとの単語分布を用いて役職を推定する手法を提案している．文献 [5] は，役職推定問題を機械学習問題として定式化し，分散表現技術の一つである word2vec を用いてプレイヤの発言をベクトル化することで訓練例を構築している．また文献 [6] では，LDA と Word2Vec を併用することによる役職推定精度への影響を報告している．上記の手法も含め，役職推定研究の多くにおいて，積極的に機械学習手法の利用が試みられている [7, 8, 9, 10]．

近年の深層学習・強化学習の発展にも関係し，役職推定に関する既存研究の多くは，推定精度の向上を主たる目的としている．しかし，より高度で戦略的に行動するエージェントを実現するためには，役職推定モデルを一つのモジュールとしてエージェント中に組み込み，その他の機能と協調動作させることが不可欠となる．このためには，高い推定精度はもちろんのこと，推定の根拠や基準を明示化し，エージェントの構築に対して直接的にフィードバックを行えることが望ましいと考えられる．

これらのことを背景に，本論文では，役職・陣営推定モデルの解釈を目的とした明示的推定理由の抽出を行う．具体的には，第 4 回人狼知能大会決勝戦のログデータを対象とし，決定木およびランダムフォレストによる推定モデルの構築と，inTrees[11] による解釈可能モデルの抽出を行い，これらの結果を比較，考察する．

連絡先: 尾崎 知伸，日本大学 文理学部 情報科学科，〒 156-8550
東京都世田谷区桜上水 3-25-40，tozaki@chs.nihon-u.ac.jp

2. データセット

本研究では，第 4 回人狼知能大会決勝戦のログデータ *1（15 人データのみ）を使用した．また，4 日目開始時に役職・陣営を推定することを想定し，4 日目開始時点で生き残っているプレイヤ（エージェント）に対する 3 日目終了時までのログデータを利用する．

本研究では，役職・陣営推定問題を分類問題の枠組みで定式化する．すなわち，各ゲームにおける各エージェントを一つの事例とみなし，クラス変数を役職（人狼，村人，占い師，霊媒師，狩人，狂人）または陣営（人間，人狼），説明変数を（3 日目終了時点までの）各エージェントの行動とする．説明変数は，当該エージェントの発話状況や占い・霊能結果，他エージェントへの賛否，他エージェントからの賛否，前日との差などの観点から，合計約 90 種を準備した．

モデル構築及び評価に利用する事例数を表 1 に示す．本来人狼ゲームでは，割り当て役職者数の違いからクラスバランスが崩れるが，今回の実験では，各クラスに属する事例数が均等になるようにダウンサンプリングを行っている．

3. 推定モデル構築実験

推定モデルの構築には，R 言語上で実装されているランダムフォレストと決定木を用いた．また inTrees*2 を用い，ランダムフォレストからのルール抽出を行った．ここで inTrees[11] とは，ツリーモデルアンサンブルの簡略化を通じ，解釈容易なモデルを抽出する手法である．具体的には，決定木中の根から葉へのパス（経路）をそれぞれ一つのルールと捉え，各ルールの簡略化と貪欲アルゴリズムによるルール選択を通じ，ツリーモデルアンサンブルから決定リストを抽出する．

表 1: 事例数

	役職推定	陣営推定
訓練例集合数	15,058	22,647
テスト例集合数	4,589	5,661

*1 http://aiwolf.org/4th-aiwolf-contest
*2 https://cran.r-project.org/web/packages/inTrees/index.html

表 2: 実験結果：推定精度

役職推定タスク		ランダムフォレスト			決定木			inTrees		
		再現率	適合率	F 値	再現率	適合率	F 値	再現率	適合率	F 値
訓練	占い師	0.993	0.996	0.995	0.677	0.165	0.265	0.969	0.173	0.293
	霊能者	0.996	0.993	0.994	0.191	0.161	0.175	0.000	0.000	0.000
	狩人	0.992	0.996	0.994	0.045	0.231	0.076	0.087	0.302	0.135
	村人	0.999	0.995	0.997	0.000	0.000	0.000	0.000	0.000	0.000
	人狼	0.990	0.982	0.986	0.090	0.162	0.116	0.038	0.272	0.066
	狂人	0.987	0.995	0.991	0.000	0.000	0.000	0.000	0.000	0.000
テスト	占い師	0.201	0.226	0.213	0.735	0.187	0.298	0.971	0.175	0.297
	霊能者	0.248	0.246	0.247	0.337	0.238	0.279	0.000	0.000	0.000
	狩人	0.407	0.260	0.317	0.089	0.358	0.142	0.090	0.308	0.140
	村人	0.203	0.243	0.221	0.000	0.000	0.000	0.000	0.000	0.000
	人狼	0.145	0.193	0.166	0.102	0.250	0.145	0.029	0.176	0.049
	狂人	0.233	0.245	0.239	0.000	0.000	0.000	0.000	0.000	0.000
陣営推定タスク										
訓練	人間	0.811	0.709	0.757	0.905	0.608	0.727	0.921	0.599	0.726
	人狼	0.667	0.779	0.719	0.417	0.814	0.551	0.382	0.829	0.523
テスト	人間	0.786	0.700	0.740	0.912	0.607	0.729	0.928	0.599	0.728
	人狼	0.663	0.756	0.707	0.409	0.823	0.547	0.379	0.840	0.522

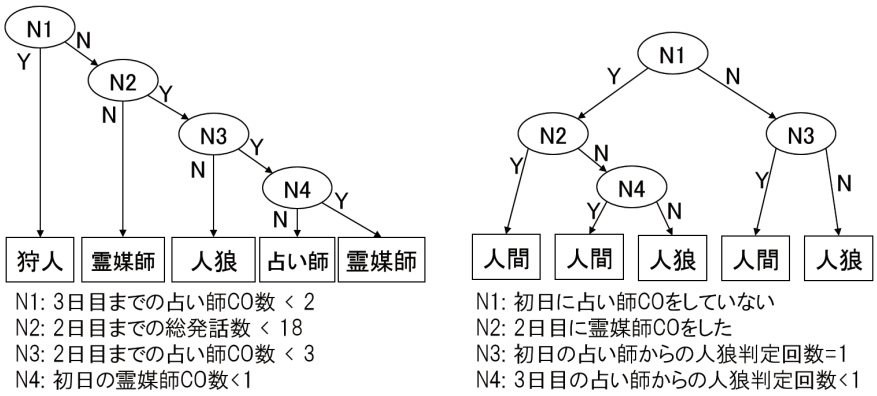


図 1: 得られた決定木：役職推定（左），陣営推定（右）

3.1 推定精度

表 2 にランダムフォレストおよび決定木，inTrees による役職推定，陣営推定の精度を示す．役職推定に関しては，訓練例集合に対するランダムフォレストの精度が非常に高いことが分かる．その一方で，テスト例に対しては必ずしも十分な精度が得られているとは言えない．また決定木は，テストデータにおける占い師・霊能者の予測精度がランダムフォレストに匹敵する一方で，村人・狂人に対する予測ルールが導出されておらず，役職によっては大きな偏りが確認できる．inTrees も決定木と同様であり，いくつかの役職に対するルールが導出されておらず，占い師以外の役職に対してはほとんど予測ができていない結果となった．

陣営の推定に関しては，すべての手法で人間陣営に対する F 値が 70%を超えており，高い精度での予測ができていることが確認できる．また，決定木と inTrees の性能はほぼ同じであり，決定木は人間陣営に対する適合率と人狼陣営に対する再現率が，inTrees は人間陣営に対する再現率と人狼陣営に対する適合率が，それぞれ他方より僅かに優れている結果となった．

3.2 得られたモデル

図 1 に得られた決定木を示す．役職推定，陣営推定共に，非常に簡単な決定木が得られていることが分かる．各決定木中の中間ノード数は 4 であり，これは単純に 4 つの条件だけで役職・陣営を推定していることに相当する．条件が少ないという点で内容を把握することは容易ではあるが，精度が低いこともあり，モデルとしての妥当性には大きな疑問が残る．

表 3 に inTrees により得られたルール（決定リスト）を示す．先述した通り，inTrees はランダムフォレストに含まれるルールを選択することで決定リストを生成する [11]．従って，各ルールは決定木における根から葉へ至るパスに相当し，結果として複数の条件を含むものとなる．今回の実験では，（デフォルトルールを除き）役職推定では 5 つのルール，陣営推定では 4 つのルールがそれぞれ抽出された．また各ルールは複数の条件で構成されており，決定木による結果よりは複雑で豊富な情報を含んでいることが確認できた．ただし，決定木の場合

表 3: inTrees によるルール

役職判定ルール

狩人	2 日目に占い師が人間判定した回数 ≤ 1 & 2 日目にプレイヤが賛成し高い数 > 0.5 & 3 日目までのプレイヤの総発話数 > 4.5
霊媒師	2 日目のプレイヤの総発話数 > 17.5 & 2 日目のプレイヤが否定された回数 > 0.5 & 2 日目と 3 日目のプレイヤの発話数の差 > 3.5
占い師	3 日目に賛成し高い数 > 0.5 & 3 日目のプレイヤの総発話数 > 3.5
狩人	1 日目のプレイヤの総発話数 ≤ 0.5 & 3 日目に賛成された回数 > 0.5
村人	3 日目に疑われた回数 > 2.5
狩人	デフォルト

陣営判定ルール

人間	2 日目のプレイヤの総発話数 > 2.5 & 3 日目に占い師 CO した順序 ≤ 0.5 & 3 日目に賛成した回数 > 0.5
人狼	1 日目に占い師から人狼判定を受けた回数 ≤ 0.5 & 1 日目に占い師 CO した順番 > 0.5
人狼	2 日目に占い師から人狼判定を受けた回数 ≤ 0.5 & 3 日目に占い師人狼判定を受けた回数 > 0.5
人間	1 日目に占い師 CO した順番 ≤ 0.5 & 2 日目に投票発言から投票先を変更した & 3 日目に霊媒師 CO した順番 ≤ 0.5
人間	デフォルト

と同様，推定できない役職が存在し，また精度面からルールの妥当性には大きな疑問が残る結果となった．

4. まとめと今後の課題

本論文では，役職推定と陣営推定の両タスクに着目し，決定木と inTrees を通じた解釈可能モデルの抽出を行った．予測精度は必ずしも十分ではなかったが，決定木を用いた直接的なモデルとの比較を通じ，inTrees を用いることで，より複雑ではあるが豊富な情報を持つ解釈可能なルールの抽出が行えることが確認できた．

今後の課題として，抽出したルールを基にした各役職に特有な知識・視点の整理や利用があげられる．また，異なる解釈可能モデル構築アルゴリズム（例えば defragTree[12] など）の適用や，抽出したルールを基にしたエージェントの構築などがあげられる．

参考文献

[1] 鳥海 不二夫，片上 大輔，大澤 博隆，稲葉 通将，篠田 孝祐，狩野 芳伸 :『人狼知能』，森北出版，2016.

[2] 狩野 芳伸，大槻 恭士，園田 亜斗夢，中田 洋平，箕輪 峻，鳥海 不二夫（著），人狼知能プロジェクト（監修）:『人狼知能で学ぶ AI プログラミング』，マイナビ出版，2017.

[3] 林 友超，呉 双，板東 勇樹，宇津呂 武仁： 制約充足による人狼ゲームの役職絞り込み， 人工知能学会第 31 回全国大会，1N1-1in2, 2017.

[4] 荒木 大輔，鳥海 不二夫，菅原俊治： トピックモデルを用いた人狼ゲームの会話に基づく役職別のプレイヤ推定法， 情報処理学会第 79 回全国大会, pp.481–482, 2017.

[5] 阪本 真基，上野敦志，田窪朋仁： プレイヤの発言に基づいた人狼ゲームの役職推定，情報処理学会研究報告ゲーム情報学 (GI), Vol.2016-GI-35, No.12, pp.1–6, 2016.

[6] 木村 優里絵，尾崎 知伸： 単語埋め込み技術を用いた人狼 BBS における役職推定， 人工知能学会第 32 回全国大会，1H2-OS13b-02, 2018.

[7] 稲葉通将，大畠菜 央実，高橋 健一，鳥海 不二夫： 雑談ばかりしていると殺される？人狼ゲームにおける発話行為タグセットの提案とプレイヤの行動・勝敗の分析，情報処理学会論文誌, Vol.57, No.11, pp.2392–2402, 2016.

[8] 今野 伸浩，大槻 恭士： 機械学習による人狼知能の陣営推定能力向上，エンタテインメントコンピューティングシンポジウム 2017 論文集, pp.64–69, 2017.

[9] 大川 貴聖，吉仲 亮，篠原 歩： 深層学習を用いて役職推定を行う人狼知能エージェントの開発，ゲームプログラミングワークショップ 2017 論文集, pp.50–55, 2017.

[10] 木村 勇太，伊藤 毅志： 人狼 AI における機械学習を用いた役職推定の改良， 人工知能学会第 32 回全国大会，1H2-OS13a-03, 2018.

[11] H. Deng： Interpreting Tree Ensembles with inTrees, International Journal of Data Science and Analytics, 2018. https://doi.org/10.1007/s41060-018-0144-8

[12] S. Hara, K. Hayashi： Making Tree Ensembles Interpretable: A Bayesian Model Selection Approach, *Proc.* of the 21th International Conference on Artificial Intelligence and Statistics, pp.77–85, 2018.

5:10 PM - 5:30 PM (Thu. Jun 6, 2019 3:50 PM - 5:30 PM Room F)

[3F06-09-5] Discussion / Conclusion

Discussion / Conclusion

©The Japanese Society for Artificial Intelligence

Organized Session | Organized Session | [OS] OS-18

[3G3-OS-18a] 感情と AI(1)

日永田 智絵（電気通信大学）、堀井 隆斗（電気通信大学）、長井 隆行（大阪大学）

Thu. Jun 6, 2019 1:50 PM - 3:30 PM Room G (302A Medium meeting room)

[3G3-OS-18a-01] (Invited talk) Predictive coding of interoception and affective emergence

〇Hideki Ohira[1] （1. Nagoya University）

1:50 PM - 2:30 PM

[3G3-OS-18a-02] Action Selection based on Somatic Marker Hypothsis

〇Chie Hieida[1], Takato Horii[1], Takayuki Nagai[1,2] （1. The University of Electro-Communications, 2. Osaka University）

2:30 PM - 2:50 PM

[3G3-OS-18a-03] Active perception based on free-energy minimization on restricted Boltzmann machines

〇Takato Horii[1], Takayuki Nagai[1,2] （1. The University of Electro-Communications, 2. Osaka University）

2:50 PM - 3:10 PM

[3G3-OS-18a-04] Extraction of Neuroscientific Findings by Visualization of Deep Neural Network

〇Kazuki Sakuma[1], Junya Morita[1], Taiki Nomura[2], Takatsugu Hirayama[3], Yu Enokibori[3], Kenji Mase[3] （1. Faculty of Informatics, Shizuoka University, 2. Graduate School of Information Science, Nagoya University, 3. Graduate School of Informatics, Nagoya University）

3:10 PM - 3:30 PM

©The Japanese Society for Artificial Intelligence

1:50 PM - 2:30 PM (Thu. Jun 6, 2019 1:50 PM - 3:30 PM Room G)

[3G3-OS-18a-01] (Invited talk) Predictive coding of interoception and affective emergence

〇Hideki Ohira[1] （1. Nagoya University）

Keywords: Emotion, Predictive coding, Interoception

近年の感情科学では、感情は身体内部の知覚である内受容感覚からコア・アフェクトという基礎的状態が形成される過程と、コア・アフェクトを言語により解釈するカテゴリー化の過程から創発されると主張される。またコア・アフェクトは、脳の生成モデルによる予測に基づいた身体信号のベイズ的推論により形成されると考えられている。本講演では、こうした理論を紹介し、脳・身体・行動に関わる実証的知見からその妥当性を考える。

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

ソマティック・マーカー仮説に基づく行動選択

Action Selection based on Somatic Marker Hypothsis

日永田智絵 *1 Chie Hieida　堀井隆斗 *1 Takato Horii　長井隆行 *1*2 Takayuki Nagai

*1電気通信大学 The University of Electro-Communications　*2大阪大学 Osaka University

Emotions are very important for human intelligence; however the mechanism of emotions is not yet fully clarified. The important aspect of embodiment in emotion has been claimed by Damasio's somatic marker hypothesis, which proposed that emotions evaluate external stimuli efficiently through the body. As a first step toward understanding the mechanism of emotion, we try to verify the somatic-marker hypothesis using a computer simulation. Specifically, we introduce a module that learns actions using body signals, and verify whether the agent can learn to obtain higher reward using the body signals. As the result, the simulation reveals that the model with body signals can select actions with higher reward compared to models without signals from the body.

1. はじめに

感情は人において重要な要素だといわれる．近年の深層学習の発展に伴い，人工物は人を凌ぐ計算能力を有するようになるといわれるものの，感情は人工物には有することのできないものであるという考え方が多い．しかしながら，意識が脳の計算の結果，生じているものだとするならば，感情も計算機上で表現できる可能性があるのではないだろうか．

感情に関してはその重要な役割から，過去にも多く感情に関する研究が行われてきた．その中の重要な仮説として，ソマティック・マーカー仮説というものが存在する．ソマティック・マーカー仮説とは，神経科学者の Damasio が，情動を身体シグナル，その認知を感情とし，外部の刺激によって引き起こされる身体シグナルが意思決定を効率化しているとした仮説である [Damasio 96]．

本研究では，感情メカニズム解明への足掛かりとして，深層学習を用い，ソマティック・マーカー仮説を中心とした感情の統合的なシステムを構築し，ソマティック・マーカー仮説の検証を行うこととした．具体的な検証方法は，意思決定の効率化を強化学習における報酬の高い行動選択が行えるようになることと解釈し，深層学習を用い，行動方策を学習するネットワークを構築し，身体シグナルがあることによって，報酬の高い行動の選択が行われるかを検証した．

本研究の意義は，感情の統合的なシステムを構築し，その振る舞いを考察することによって感情メカニズム解明を目指すアプローチ（構成的アプローチ）を行う点にある．例えば生理学的研究では，ジェームズ・ランゲ説を初めとした感情の理論が提案されており [James 84]，神経科学では，各脳部位の感情的な働きを調査している [Koelsch 15]．これらの研究では，一部分のモデルや各部位に絞った研究となっており，感情の統合的なメカニズムについてはまた完全には明らかになっていない．各部位に焦点を当てる方法は，その部位を知るために重要であるが，統合的なシステムの中で初めて現れる機能もあると考えられる．そのため，感情の統合的なシステムを構築し，その振る舞いを考察することによって感情メカニズム解明を目指すアプローチも促進されるべきである．

連絡先: 日永田智絵，電気通信大学，東京都調布市調布ヶ丘 1-5-1，042-443-5238，hchie@apple.ee.uec.ac.jp

2. 実験プロトコル

仮説検証のため，まずは人の身体との対応を考慮したシンプルなタスクを設計した．本研究では，射的タスクを題材にした（図 1）．タスクには水鉄砲，バケツ，風船が存在し，それぞれ身体反応（筋肉の動きや発汗，心拍など），体力，外的刺激（食料）を表し，エネルギーの流れは水の移動で表現される．バケツと水鉄砲に入った水は時間が経つと蒸発してしまう．3 回に 1 回，風船が飛んでくるという状況で，行動として水鉄砲を打つか，バケツの蓋を抑えるかを選択することができる．風船を水鉄砲で打つことが出来れば，風船から水を獲得することができ，バケツの水が増える．バケツの蓋を抑えると蒸発量を減らすことができる．水鉄砲に水を増やす際にはバケツの水を減らすこととなる．このような状況下で自身の持つ水の量を一定に保つように行動を学習する．この設定はホメオスタシスの考え方からである．その際に身体反応と体力の情報即ち身体シグナルが行動決定を効率化するかということを検証することになる．身体シグナルを出力するモジュールを本研究では，Somatic Marker Module（SMM）と呼ぶ．タスク内の詳細値は以下に示す．

- 初期値は水鉄砲が 0，バケツが 80.
- 値の範囲は水鉄砲が 0–10，バケツが 0–100.
- 水鉄砲とバケツの水の総量を 80 に保つことが目標.
- 打つ行動を選択すると，水鉄砲-10.
- 但し水鉄砲が 10 に満たない場合は打つことができない.
- 風船を打つことができると，バケツ+30.
- 画像が変わるごとに水が蒸発．水鉄砲-2，バケツ-2.
- 但し蓋を抑える行動を選択するとバケツの水の蒸発が半減.

3. 深層学習による実装

実装したネットワークを図 2 に示す．本研究では，強化学習での行動選択を想定し，Deep Deterministic Policy Gradient（DDPG）[Lillicrap 15] を用いて，行動モジュールを実装した．また，時系列情報を扱えるようにするため，convolutional LSTM （long short-term memory）[Xingjian 15] を用いて，次の状況を予測するネットワークを追加した．本タスクにおいてはオーバースペックであるとも考えられるが，さらに複雑な

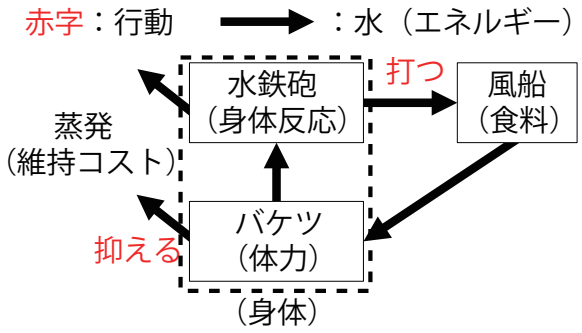


図 1: タスク設計概要

タスクを行うことを視野にいれ，本手法を選択した．身体シグナルを出力する Somatic Marker Module に関してはルールベースで作成した．

3.1 Somatic Marker Module（SMM）

SMM の具体的な役割は外的刺激と行動によって変化した現在の身体シグナルを受け，身体反応を引き起こし，行動モジュールに身体シグナルを入力することである．外的刺激を受け身体反応を引き起こす部分は，風船があるときには，水鉄砲に 10 用意することとした．水鉄砲に水を用意するにはバケツから足りない分の水を移動させることとなる．その操作をした後，次のモジュールに水鉄砲とバケツの値を入力する．

3.2 Convolutional LSTM

Convolutional LSTM は，Xingjian らによって提案された画像の特徴を捉える Convolutional Neural Network と長期時系列情報を扱うことができる LSTM とを組み合わせた方法である [Xingjian 15]．具体的には，LSTM の重みの乗算が畳み込みであり，構成要素はメモリセル C_t，入力ゲート i_t，忘却ゲート f_t，出力ゲート o_t である．本稿では 2 層の Convolutional LSTM を用いており，フィルタは $5 \times 5 \times 5$ であり，誤差は平均二乗誤差によって計算される．学習率は Adam（$\alpha = 0.001, \beta_1 = 0.9, \beta_2 = 0.999, \epsilon = 10^{-8}$）である．入力は外部刺激画像と身体シグナル（65 × 65 × 5 次元）で，出力は予測画像と予測身体シグナル（65 × 65 × 5 次元）である．身体シグナルは値をグレースケール画像化して入力している．

3.3 Deep deterministic policy gradient

Deep deterministic policy gradient（DDPG）は，Lillicrap らが提案した深層学習を用いた強化学習法であり，高次元の連続動作空間でポリシーを学習できる深層関数近似を用いたモデルフリー，オフポリシーの actor-critic アルゴリズムである [Lillicrap 15]．入力は外部刺激画像と LSTM による予測画像，身体シグナルと身体シグナルの初期値（65 × 65 × 12 次元）で，出力は行動 (1 次元) である．中間層は 300 次元の入出力で構成されている．報酬はホメオスタシスという考えを用い，下記のように算出する．

$$R(t) \quad = \quad C - \left|\|\boldsymbol{m}(t)\| - \|\boldsymbol{a}'(t)\|\right|, \qquad (1)$$

$R(t)$ は時間 t における報酬値，$\boldsymbol{a}(t)$ は時間 t における SMM の出力，すなわち身体シグナルを表す．それに対し，$\boldsymbol{a}'(t)$ は $\boldsymbol{a}(t)$ に行動後の変化（例えば風船から水が得られたなど）を加味したものとなる．$\boldsymbol{m}(t)$ は時間 t での保つべき基準の値となっており，本研究では初期値の値 $\boldsymbol{a}(0)$ とする．C は報酬値の補正をする定数であり，$R(t)$ は 0–100 の値をとる．但し，$\|\boldsymbol{a}'(t)\|$ が 0 のとき，$R(t)$ も 0 とし，$\boldsymbol{a}'(t)$ の値を初期値に戻す．

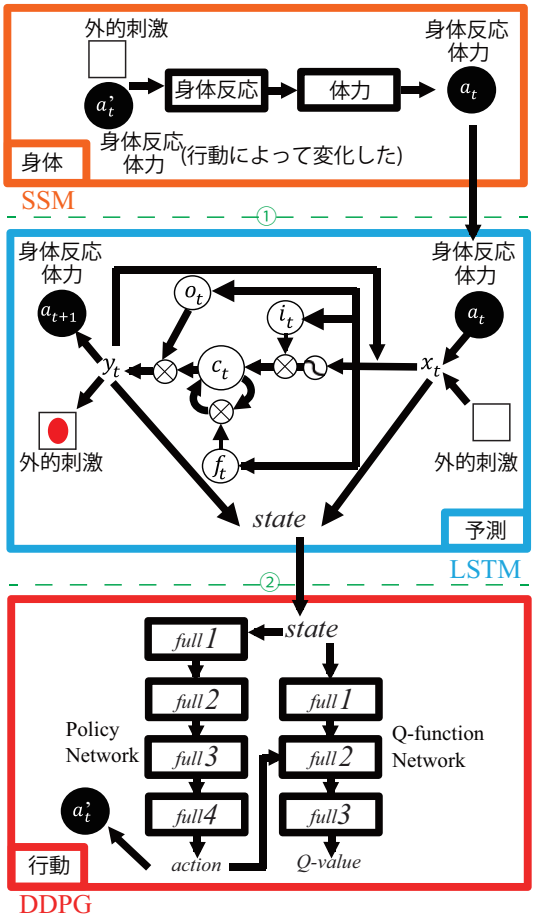


図 2: 実装したネットワーク

学習率は Adam（actor network: $\alpha = 10^{-4}, \beta_1 = 0.9, \beta_2 = 0.999, \epsilon = 10^{-8}$, critic network: $\alpha = 10^{-3}, \beta_1 = 0.9, \beta_2 = 0.999, \epsilon = 10^{-8}$）となっており，探索手法は Ornstein–Uhlenbeck process，ミニバッチは 200，リプレイバッファのサイズは 500 で，新しいデータが入ると古いデータは破棄される．また，本研究では batch normalization を使用した．

4. 結果・考察

SMM のある条件を SMM 条件とし，それに対して SMM と LSTM がない条件を DDPG 条件（図 2 の②の破線でネットワークが切れている条件），SMM がない条件を DDPG-LSTM 条件（図 2 の①の破線でネットワークが切れている条件）と呼ぶ．欠損した情報は 0 で入力している．学習時の報酬の様子を図 3 に示す．学習はそれぞれ 20000epochs 行った．グラフは見やすくするために 50epochs ごとの平均値をプロットしている．報酬のグラフより，全ての条件で報酬は高くなっているが，身体シグナルありの場合がより高い報酬が得られていることがわかる．総時間の報酬の平均値をとると，SMM 条件が 88.9，DDPG 条件が 83.1，DDPG-LSTM 条件が 83.2 となっており，身体シグナルありが最も大きい．しかし，この結果には学習時の行動のノイズも含まれているため，学習済みモデル

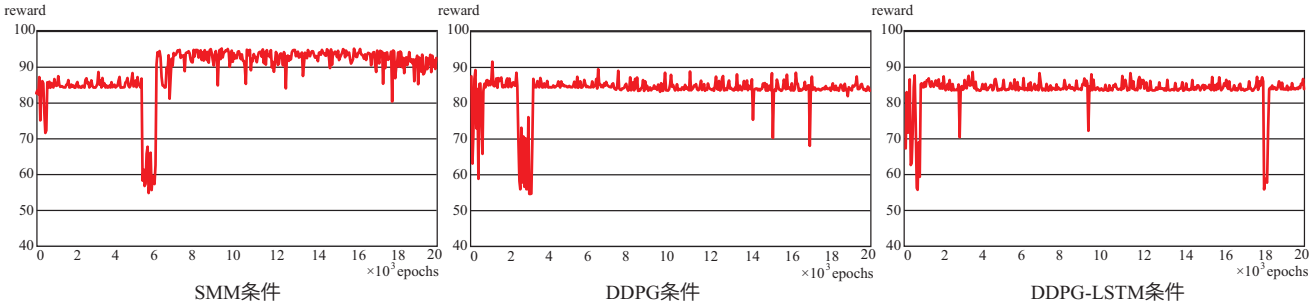


図 3: 学習時の報酬の推移

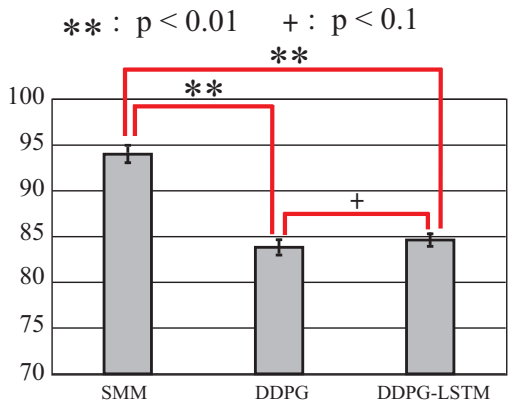


図 4: 学習済みモデルでの報酬平均の比較

に対し，それぞれ 20episode 動かした結果を平均したものを比較した．

結果を図 4 に示す．平均値は SMM 条件が 94.1，DDPG 条件が 84，LSTM 条件が 84.7 となった．ボンフェローニ法で解析を行った結果，SMM 条件と DDPG 条件，DDPG-LSTM 条件の間で有意差が見られた（$p = 2.94 \times 10^{-8}, 2.16 \times 10^{-7}$，ボンフェローニ法より $p < 0.01/3$）．DDPG 条件と DDPG-LSTM 条件の間は有意傾向であった（$p = 0.027$，ボンフェローニ法より $p < 0.1/3$）．よって，身体シグナルありの場合が報酬を高くする行動を選択できることがわかる．また，DDPG と DDPG-LSTM の間は有意傾向であるので，予測モジュールが報酬増加に寄与したと考えられる．

LSTM の Loss のグラフは図 5 の通りである．グラフは見やすくするために 100epochs ごとの平均値をプロットしている．初期状態に比べ誤差が減少しており，学習が行われていることがわかる．最終的な誤差は 3.24 と 1.32 となり，DDPG-LSTM 条件の方が小さくなる．これは，身体シグナルの予測を行っていないためであると考えられる．

次に LSTM による予測の一例を図 6 に示す．画像の予測では，綺麗に画像の再現ができているわけではないものの，風船が空の色になっていたり，空が風船の色になっているなど，多少の予測ができていることがわかる．水量の予測では，非常に近い値ではないものの，入力された情報に対し値が増加するか減少するかというレベルでの予測はできている．

また，身体反応のパターンを変更し，学習した結果の行動を比較した．a 群は風船がきたときのみ水鉄砲に 10 用意する群で，b 群は画像に関わらず常に水鉄砲に 10 用意するパターンである．その結果 SMM 条件と DDPG-LSTM 条件では行動に違いが出たが，DDPG 条件では行動に違いが出なかった．これは，DDPG 条件では風船がどのタイミングで来るか予測できず，来た際に行動するということしかできないため，a 群と b 群に違いが出なかったと考えられる．予測モジュールがあると，風船が来ないパターンの状況を細分化できるため，行動にバリエーションが生まれる．しかしながら，詳細な身体状況はわからないため，周期的な行動に落ち着くこととなる．身体シグナルがあると，その時の身体状況に合わせて行動を選択することが可能となる．

また，SMM 条件の b 群の報酬平均は 93.7 で，DDPG 条件の報酬平均は 86，DDPG-LSTM 条件の報酬平均は 93.1 であった．SMM 条件と DDPG-LSTM 条件の報酬があまり変わらないのは，b 群のシナリオが身体反応を一定にしたために，身体シグナルなしでも十分に行動できたためだと思われる．しかしながら，打つ行動を選択し行動できている回数を比較すると，SMM 条件は 6 回で DDPG-LSTM 条件は 11 回となる．つまり，身体シグナルがあることにより，使用する水の量をより少なくより高い報酬を得ることができている．SMM 条件内で比較すると，a 群は打つことができているのが 2 回，b 群が 6 回となる．つまり，a 群の方が使用する水の量を少なくできており，身体反応のパターンが行動に対して制約をかけ，行動を効率化させていることがわかる．即ち，行動選択において，身体反応は重要であり，この部分をどのように設定するのかが問題となる．人においては，生得的な身体反応があると考えられており，その要素がエクマンの基本感情 [Ekman 71] の考えに繋がっていると考えられる．ただし，生得的な部分のみでは説明がつかない要素もあり，後天的に身体反応を学習するモジュールも必要であろう．

5. まとめ

本研究では，感情メカニズム解明への足掛かりとして，ソマティック・マーカー仮説の検証を行った．手法として，強化学習での行動選択を想定し，Deep Deterministic Policy Gradient（DDPG） を用いて行動を強化学習するモジュールを作成し，convolutional LSTM（long short-term memory）で作成した予測モジュール及び身体シグナルを出力する Somatic Marker Module（SMM）の有無によって，行動選択がどのように変容するかを検証した．学習済みモデルで行動した際の報酬の平均を比較した結果，身体シグナルありの SMM 条件が最も報酬が高い行動を選択できていることがわかった．また，身体反応のパターンの変化による検証では，身体反応のパターンが行動に対して制約をかけ，行動を効率化させていることがわかった．

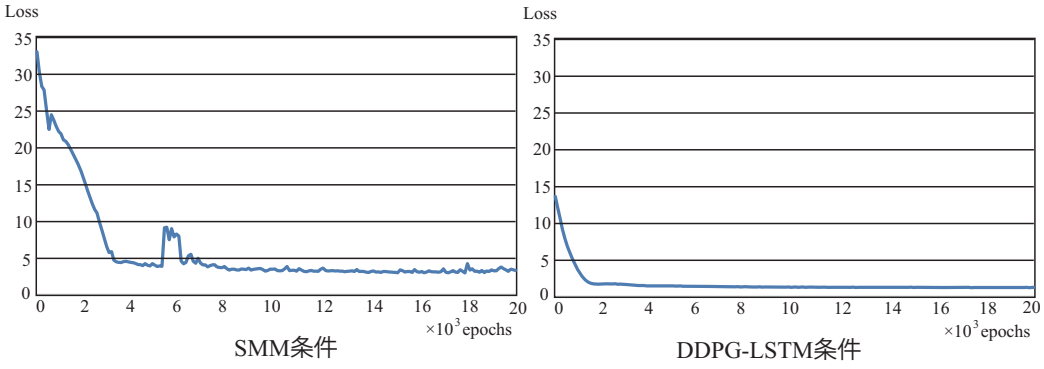


図 5: 学習時の LSTM の Loss の推移

表 1: 画像に対する反応を変更した際の行動例：画像 0 は風船なし，画像 1 は風船あり，行動 0 は蓋を抑える，行動 1 は打つ．

	画像		0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0
行動	SMM 条件	a 群	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	0	1	0	0	0
		b 群	1	1	1	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0
	DDPG 条件	a 群	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0
		b 群	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0
	DDPG-LSTM 条件	a 群	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	1	0	1
		b 群	1	1	1	0	0	0	0	0	1	1	0	1	1	0	1	1	0	1	1	0

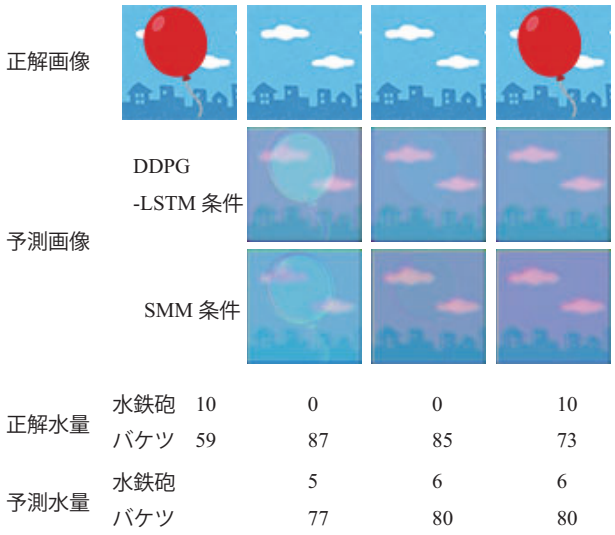


図 6: LSTM による予測の一例

これらは，身体シグナルが意思決定を効率化しているというソマティック・マーカー仮説の主張に沿った結果となっている．

今後は SMM-LSTM-DDPG の全体のネットワークに対し，より複雑なタスクを行い，挙動を確かめるとともに，社会的感情など，複雑な感情がどのように表現されるのかを検証する．

謝辞

本研究は JSPS 科研費 JP16J04930，JST CREST （JPMJCR15E3），新学術領域「認知的インタラクションデザイン学」 （26118001）の助成を受けたものです．

参考文献

[Damasio 96] Damasio, A. R., Everitt, B. J., and Bishop, D.: The Somatic Marker Hypothesis and the Possible Functions of the Prefrontal Cortex [and Discussion], *Philosophical Transactions of the Royal Society B, Biological Sciences*, Vol. 351, No. 1346, pp. 1413–1420 (1996)

[Ekman 71] Ekman, P. and Wallace, F. V.: Constants across cultures in the face and emotion, *Journal of personality and social psychology*, Vol. 17, No. 2, pp. 124–129 (1971)

[James 84] James, W.: WHAT IS AN EMOTION ?, *Mind*, Vol. os-IX, No. 34, pp. 188–205 (1884)

[Koelsch 15] Koelsch, S., Jacobs, A. M., Menninghaus, W., Liebal, K., Klann-Delius, G., Scheve, von C., and Gebauer, G.: The quartet theory of human emotions: An integrative and neurofunctional model, *Physics of Life Reviews*, Vol. 13, pp. 1–27 (2015)

[Lillicrap 15] Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., and Wierstra, D.: Continuous control with deep reinforcement learning, *arXiv preprint arXiv:1509.02971* (2015)

[Xingjian 15] Xingjian, S., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., and Woo, W.-c.: Convolutional LSTM network: A machine learning approach for precipitation nowcasting, in *Advances in neural information processing systems*, pp. 802–810 (2015)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

制限ボルツマンマシンによる 自由エネルギー最小化に基づく能動的知覚

Active perception based on free-energy minimization on restricted Boltzmann machines

堀井隆斗 *1　長井隆行 *1*2
Takato Horii　Takayuki Nagai

*1電気通信大学　*2大阪大学
The University of Electro-Communications　Osaka University

A robot, which would like to respond quickly in the world, should select more informative signals to estimate the cause of its sensation (e.g., a state of the environment, a category of a handling object, an emotional state of interaction partner, etc.). This paper proposes an active perception framework that selects the robot's action to perceive critical sensory signals based on a free-energy minimization in an energy-based model. We employed a restricted Boltzmann machine as a fundamental component for an estimation network of the cause of sensations. Our framework demonstrated better performance for the attention control in emotional human-robot interaction than other methods.

1. はじめに

多様な情報にあふれる実世界で活動するロボットにとって，環境や自身の状態を認識するために複数の感覚情報を処理することが重要である．Nakamura と Nagai [Nakamura 17] は，質感や色，聴覚，触覚，単語の感覚様式に基づいて，ロボットが多感覚情報を処理することで知覚物体の関係性を構造化する手法を提案している．また Horii ら [Horii 18] は，視覚，聴覚，触覚の感覚様式に注目し，それらの感覚情報を統合することで感情発達を再現する計算論モデルを提案している．このように複数の感覚情報を統合・処理することで物体のカテゴリや感情といった概念構造を獲得することができる．そしてこのような概念構造から環境や自身の状態を推定することで，他者との相互作用が可能となる．

概念構造を獲得したロボットは，一部の感覚情報からでも概念情報を予測することができる．しかし，本来複数の感覚情報を処理することで獲得された概念情報を，一部の感覚情報から予測するとは難しい．例えば不透明なペットボトルが見えた時にそれがゴミであるか否か，電話先の友人が大きな声を発した時に喜んでいるのか怒っているのかを判断するためには，追加の感覚情報が必要であることは想像にたやすい．このような場面で概念情報（ここでは物体のカテゴリや他者の感情状態）の推定に効果的な感覚情報を選択・知覚することを能動的知覚と呼ぶ．Taniguchi ら [Taniguchi 18] は，ロボットのための能動的知覚として，現在推定している概念情報と予測される未観測感覚情報との関係性を表す相互情報量を推定し，その相互情報量が最大となる感覚情報を選択する手法を提案している．

一方で，近年の神経科学研究において，人の能動的知覚（推論）を説明する基本原理としての自由エネルギー原理 [Friston 10] が注目されている．これは人の脳が入力される感覚信号を予測する内部モデルを構成し，その予測と実際の信号間の誤差を最小化するように認識や行動を実行するというものである．本研究ではこの自由エネルギー原理に基づいた能動的知覚のための計算論モデルと手法を提案する．具体的には深層学習モデルの 1 つである Restricted Boltzmann Machine（RBM）[Hinton 10] を用いた多感覚情報統合モデルにおいて，

連絡先: 堀井隆斗，電気通信大学，東京都調布ヶ丘 1-5-1，takato@uec.ac.jp

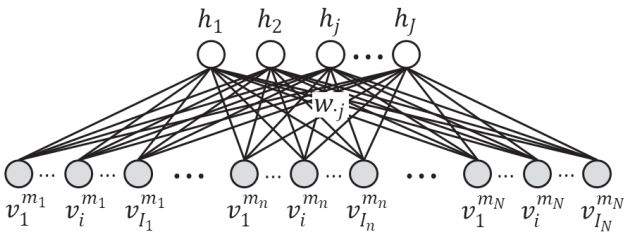


図 1: RBM による多感覚統合モデル

観測情報から概念情報を介して未観測情報を予測し，その中で自由エネルギーが最も小さい（つまり最も予測誤差が小さい）感覚情報を選択する．本稿では提案する能動知覚手法と Taniguchi ら [Taniguchi 18] の手法を，人とロボットの感情コミュニケーション場面を想定した場面において評価した結果について示す．

2. 提案手法

RBM [Hinton 10] による多感覚統合モデルと提案する自由エネルギー最小化に基づく能動的知覚手法について説明する．図 1 に RBM を用いた多感覚統合モデルを示す．ここで $v_i^{m_n} \in \{0,1\}$ は n 番目の感覚様式 m_n の i 番目の可視層ノードを，$h_j \in \{0,1\}$ は j 番目の隠れ層ノードを示す．また $w_{\cdot j}$ は h_j と可視層ノードの結合加重である．このモデルは，学習時に m_N までのすべての感覚様式情報を入力として受け取り，入力情報と隠れ層 $\boldsymbol{h}$ を介して再構成された情報の誤差を最小化することで, 確率分布 $p(\boldsymbol{v} = \{\boldsymbol{v}^{m_1}, \cdots, \boldsymbol{v}^{m_N}\})$ を推定する．また学習によって隠れ層に何らかの特徴表現（例えば物体カテゴリや感情カテゴリなど）が獲得されることが期待される．

学習によって推定される周辺確率 $p(\boldsymbol{v})$ は次式で表現される．

$$
\begin{aligned}
p(\boldsymbol{v}) &= \sum_{\boldsymbol{h}} p(\boldsymbol{v}, \boldsymbol{h}) \\
&= \exp\Big(\boldsymbol{c}^T \boldsymbol{v} + \sum_{j=1}^{J} \log(1 + \exp(b_j + \boldsymbol{W}_{\cdot j} \boldsymbol{v}))\Big) \\
&= \exp(-F(\boldsymbol{v}))/\mathcal{Z}
\end{aligned}
\tag{1}
$$

Algorithm 1 Active perception based on free-energy minimization on RBM

add perceived modality m_n to set $\mathrm{M_r}$
for $n = 0$ to N **do**
  if $m_n \in \mathrm{M_r}$ **then**
    $\boldsymbol{v}^{m_n} = v^{m_n}$
  else
    $\boldsymbol{v}^{m_n} = \mathbf{0}$
  end if
  $\boldsymbol{h} \sim p(\boldsymbol{h}|\boldsymbol{v}^{m_1}, \cdots, \boldsymbol{v}^{m_n}, \cdots, \boldsymbol{v}^{m_N})$
  for $n = 1$ to N **do**
    if $m_n \notin \mathrm{M_r}$ **then**
      $\boldsymbol{v}^{m_n{}'} \sim p(\boldsymbol{v}^{m_n{}'}|\boldsymbol{h})$
      $E_n \leftarrow F(\boldsymbol{v}^{m_n{}'})$
    end if
  end for
end for
$n = \mathrm{argmin}\, E_n$
$\mathrm{M_r} \leftarrow \mathrm{M_r} \cup m_n$

$F(\boldsymbol{v})$ は RBM の自由エネルギーである．以上より，ある情報 $\boldsymbol{v}$ の自由エネルギーが小さいことは，$\boldsymbol{v}$ の発生確率 $p(\boldsymbol{v})$ が高いことと対応する．

次にある部分観測情報 $\boldsymbol{v}' = \{\mathbf{0}, \cdots, \boldsymbol{v}^{m_n}, \cdots, \mathbf{0}\}$ から未観測情報を予測することを考える．それぞれの未観測情報 $\boldsymbol{v}^{m_u}$ は部分観測情報 $\boldsymbol{v}'$ から推定された $\boldsymbol{h}$ を介して再構成できることから，次のように予測される．

$$\begin{aligned} \boldsymbol{h} &\sim p(\boldsymbol{h}|\boldsymbol{v}') \\ \boldsymbol{v}^{m_u} &\sim p(\boldsymbol{v}^{m_u}|\boldsymbol{h}) \end{aligned} \tag{2}$$

ここで，提案手法は予測された未観測情報と部分観測情報から計算される自由エネルギーが最も小さくなるものを次の知覚情報として選択する．すなわち RBM における自由エネルギー最小化に基づく能動的知覚は，現在の観測情報から推定される概念情報に対して尤もらしく予測される感覚様式の情報を選択することとなる．Algorithm1 に提案手法の詳細を示す．

3. 実験

提案手法の有用性を評価するために，人とロボットの感情コミュニケーション場面を想定した感情推定のための能動的知覚実験を行った．実験では人の対面多感覚感情コミュニケーションデータである IEMOCAP データセット [Busso 08] から表情，音声，手の動きに関する特徴量を抽出し，それぞれを感覚様式の情報として図 1 の多感覚統合モデルを学習した．学習の詳細な条件に関しては [Horii 16] を参照されたい．また今回の実験では評価値として，部分観測情報とすべての観測情報から推定された隠れ層の発火確率の KL ダイバージェンスを利用した．

図 2 に単一の感覚様式情報を部分観測情報として入力した状態（baseline）とそれぞれ異なる能動的知覚手法で選択した 1 つの感覚様式情報を付加した状態の評価値を示す．能動的知覚を実行することで状態推定の精度が向上していることが分かる．また無作為抽出（Ramdom）と Taniguchi ら [Taniguchi 18] の手法（IG.max）と比較して，提案手法（FE.min）が高い性能を示すことが確認された．

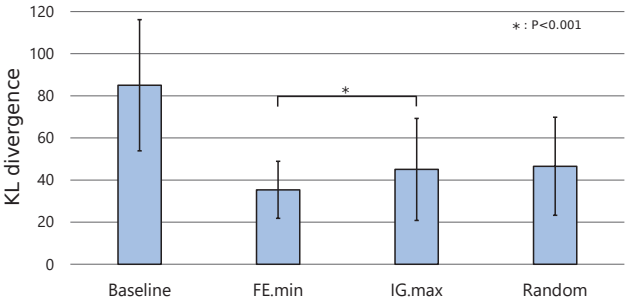


図 2: 各能動的知覚手法による感情状態推定に対する評価値

4. おわりに

人の脳情報処理の基本原理だと考えられている自由エネルギー原理に基づいた能動的知覚手法を提案した．そして，人とロボットの感情コミュニケーションを想定した能動的知覚の実験において情報量を最大化する Taniguchi ら [Taniguchi 18] の手法と情動推定精度を比較し，提案手法が高い性能を発揮することを示した．

今後は提案手法と既存手法の関係性を明らかにするとともに，RBM 以外のエネルギーモデルへの適用を進める．また，ロボットの能動的知覚による状態変化の影響を踏まえた時間方向へ拡張する．

謝辞

本研究は JST CREST(JPMJCR15E3, JPMJCR16E2) の支援を受けた．また研究内容に関して議論した情報通信研究機構長井志江主任研究員に感謝する．

参考文献

[Busso 08] Busso, C., Bulut, M., Lee, C.-C., Kazemzadeh, A., Mower, E., Kim, S., Chang, J. N., Lee, S., and Narayanan, S. S.: IEMOCAP: Interactive emotional dyadic motion capture database, *Language resources and evaluation*, Vol. 42, No. 4, pp. 335–359 (2008)

[Friston 10] Friston, K.: The free-energy principle: a unified brain theory?, *Nature reviews neuroscience*, Vol. 11, No. 2, p. 127 (2010)

[Hinton 10] Hinton, G. E.: A practical guide to training restricted boltzmann machines, Technical report, Department of Computer Science University of Toronto (2010)

[Horii 16] Horii, T., Nagai, Y., and Asada, M.: Imitation of human expressions based on emotion estimation by mental simulation, *Paladyn, Journal of Behavioral Robotics*, Vol. 7, No. 1 (2016)

[Horii 18] Horii, T., Nagai, Y., and Asada, M.: Modeling development of multimodal emotion perception guided by tactile dominance and perceptual improvement, *IEEE Transactions on Cognitive and Developmental Systems* (2018)

[Nakamura 17] Nakamura, T. and Nagai, T.: Ensemble-of-Concept Models for Unsupervised Formation of Multiple Categories, *IEEE Transactions on Cognitive and Developmental Systems* (2017)

[Taniguchi 18] Taniguchi, T., Yoshino, R., and Takano, T.: Multimodal Hierarchical Dirichlet Process-Based Active Perception by a Robot, *Frontiers in neurorobotics*, Vol. 12, (2018)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

深層学習の可視化による神経科学的知見の抽出

Extraction of Neuroscientific Findings by Visualization of Deep Neural Network

佐久間一輝 *1 Kazuki Sakuma　森田純哉 *1 Junya Morita　野村太輝 *2 Taiki Nomura　平山高嗣 *3 Takatsugu Hirayama　榎堀優 *3 Yu Enokibori　間瀬健二 *3 Kenji Mase

*1静岡大学情報学部 Faculty of Informatics, Shizuoka University　*2名古屋大学大学院情報科学研究科 Graduate School of Information Science, Nagoya University

*2名古屋大学大学院情報学研究科 Graduate School of Informatics, Nagoya University

Recently, research using deep learning has been conducted in various fields. Additionally, research on visualization methods learned by deep learning has also been actively conducted. Furthermore, the relationship between the human subjective state and electroencephalogram (EEG) has been clarified in the psychophysiological field. In this research, we apply the visualization method developed in the image field to the analysis of EEG. Using this method, we examine whether we can abstract physiologically reasonable structure of brain activity from the network visualizing EEG signals.The result of our experiment indicated the two important brain structures showing consistency with the previous neuroscience studies. We consider that our proposed method has some utilities as a tool to progress scientific understanding of human mind.

1. 研究背景

以前から，感情に関わる研究は生理心理学分野で脳波 (EEG: Electroencephalogram) などの生体信号を用いて行われてきた．EEG は脳内部活動の情報を保持する有益なものであると考えられてきたが，EEG は複数の異なる主観的状態の情報が集積されたものであり，ノイズも多く実利用が困難であるとされてきた．これに対し，生理心理学では EEG を周波数帯域成分に分離し，電極位置の情報と併用して用いることで，分析対象の主観的状態との相関が検討されてきた [Sarlo 05]．このような EEG を用いた感情分析においては，熟練した高度な技能によって，ノイズ除去や結果の解釈がなされてきた．

一方で，深層ニューラルネットワーク (DNN: Deep Neural Network) は，高精度のモデルを作成することができる機械学習の手法として盛んに研究が行われている．DNN は入力特徴にほとんど加工することなく，特徴抽出器をボトムアップに学習可能であるという利点がある．しかし，高精度のモデルを作成できたとしてもそのモデルが内部でどのように特徴抽出を行っているのかわからない．そこでモデルがどのように特徴抽出しているのかを理解するため，DNN の特徴抽出の様子や内部状態を可視化する研究が行われてきた [Ramprasaath 16, Schirrmeister 17]．しかしそのような研究は特定の入力特徴に対する特徴抽出を可視化するものであり，特徴抽出がモデル作成者の意図するものであるかを確認するためのものである．DNN には高度な特徴抽出機能があるにも関わらず，そこから対象に対する新たな知見を得ることは考慮されていない．

本研究の目的は，DNN を用いて学習したモデルから得られる知見と，生理心理学分野での解析から得られた感情と EEG の関係の知見を比較することで，DNN から妥当性の高い知見を得ることができるかを検討することである．この目的を達成することで，神経科学において，扱うデータの精度の範囲内で信頼性の高い知見を得るツールを導くことができると考える．

連絡先: 佐久間一輝，静岡大学情報学部情報科学科，静岡県浜松市中区城北 3 丁目 5 − 1，053-478-1452，sakuma.kazuki.15@shizuoka.ac.jp

2. 関連研究

ここでは本研究のターゲットとなる領域（EEG と感情）と，手法（DNN）に関する先行研究について述べる．

2.1 EEG と主観的状態との関係

EEG は生理心理学の分野で脳活動と人間の主観的状態との関係性を調査するために利用されてきた．その中で，被験者の主観的状態と脳の部位や δ 波，θ 波，α 波，β 波，γ 波などの周波数帯域別の活動に関連があることが報告されてきた．例えば，左前頭部の θ 波と α 波が快感情と，右前頭部の θ 波と α 波が不快感情と関連していることがわかっている [Davidson 03, Sarlo 05]．

神経科学の分野でも，感情と脳活動の関連を科学的に解明する試みが行われている．その中で，前頭葉の背外側前頭前野（DLPFC）は高次認知活動に関与しており，また，感情を抑制する機能を持つと考えられている．Beauregard らは，被験者に悲しくなるようなビデオや性的な写真を見せ，意識的に悲しみや性的な感情を抑制させるようにすると DLPFC の活動が上昇することを報告した [Beauregard 01]．

2.2 DNN を用いた研究

また，近年機械学習の発展に伴い，EEG の認識に DNN を用いる研究が行われている．Schirrmeister らは未加工の EEG から右手，左手，足，安静の 4 つの運動に関係した主観的状態の識別に CNN (Convolutional Neural Network) を適用し，CNN は未加工の EEG から周波数帯域成分を学習可能であることを示した [Schirrmeister 17]．

上記の研究は，EEG 解析に DNN を用いるアプローチの有用性を示唆しているが，学習したモデルがどのようにして脳活動のパターンを識別するのかを説明することが困難である．この問題に対して，学習した DNN の可視化を行う研究が行われている [Ramprasaath 16]．Ramprasaath らは CNN の可視化に Gradient-weighted Class Activation Mapping (Grad-CAM) を適用した [Ramprasaath 16]．彼らはさらに詳細な可視化を行うために，Grad-CAM と Guided Backpropagation(Guided BP) を組み合わせた Guided Grad-CAM(GGC) を示した．し

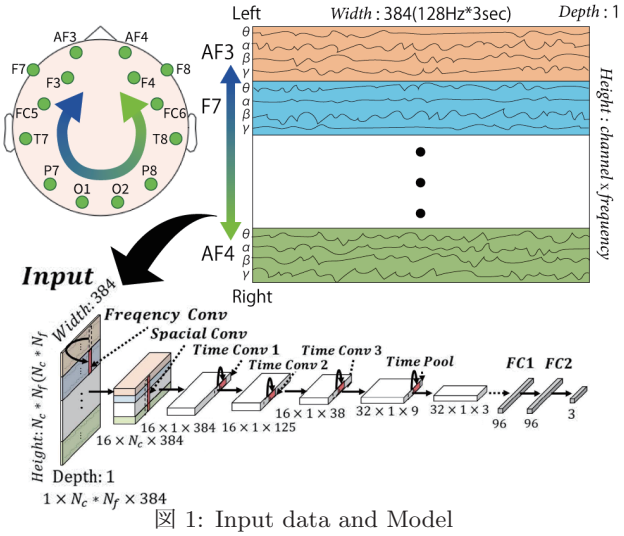


図 1: Input data and Model

表 1: Parameters of CNN

Type	Structure
Input	$Depth \times Height(N_f * N_c) \times Width(Times)$
Frequency Conv	Kernel:16 × N_f × 1, Stride:N_f × 1, pad=0, elu
Spacial Conv	Kernel:16 × N_c × 1, Stride:N_c × 1, pad=0, elu
Batch Norm 1	Dimensions:16
Dropout 1	Wight Decay:0.5
Time Conv 1	Kernel:16 × 1 × 12, Stride:1 × 3, pad=0, elu
Time Conv 2	Kernel:16 × 1 × 12, Stride:1 × 3, pad=0, elu
Time Conv 3	Kernel:32 × 1 × 12, Stride:1 × 3, pad=0, elu
Time Pool	Kernel:1 × 3,Stride:1 × 3, pad=0, Max
Batch Norm 2	Dimensions:32
Dropout 2	Wight Decay:0.5
FC 1	96, elu, Dropout:0.5
FC 2	Classes, Softmax

かしながら，EEG による感情認識にこれらの技術を適用する研究は行われていない．

3. 提案手法

本提案手法では，CNN と GGC を用いて脳活動の抽象化を行い，神経科学や生理心理学の知見の抽出を行う．

3.1 学習フェイズ

3.1.1 入力特徴

前処理では，EEG のデータセットに対して高周波成分を除去するために [1-50Hz] の FIR フィルタを適用する．次に，外れ値除去として [-500μV,500μV] の範囲を超える電位を含むデータを除外する．最後に，ウェーブレット変換により，θ 波 (4-7Hz)，α 波 (8-13Hz)，β 波 (14-30Hz) 及び γ 波 (31-50Hz) として各周波数帯域成分を算出する．

全電極位置とその全周波数帯域成分を 2 次元の画像に変換し，その各行は，各電極位置の周波数帯域成分の一連の信号入力を表す（図 1 の上部を参照）．

3.1.2 モデル

学習に用いた CNN のモデルの概要を図 1 の下部に示す．中間層では，周波数帯域軸，電極軸，時間軸の順で独立した畳み込みを行うような Kernel を設計し，出力層は快，中立，不快を示す 3 クラス分類となっている．その他のモデルの構造に関わるパラメータを表 1 に示す．

3.2 中間層の可視化と集約フェーズ

EEG と GGC を用いた中間層の可視化と集約の概要を図 2 に示す．本フェーズは三段階に分かれている．

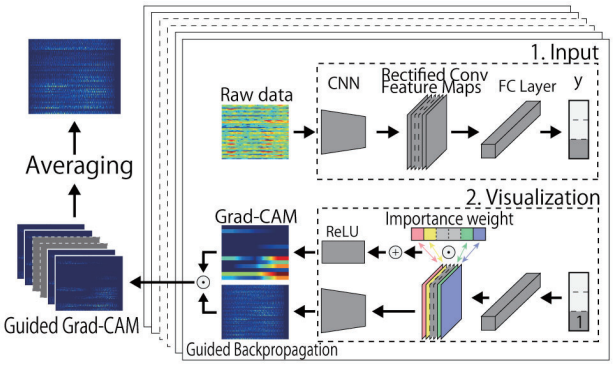


図 2: Guided Grad-CAM Procedure

3.2.1 入力

学習済みモデルに，正しく分類を行うことのできた EEG の信号（2 次元の画像）を入力する．

3.2.2 可視化

出力層における，可視化したいクラスの値を 1 としパラメータを更新せずに Backpropergation を行う．Guided BP は，以上の操作を行った後の入力層の勾配の値である．Grad-CAM は CNN 内の全結合層の直前の畳み込み層が出力した特徴マップを用いて出力する．各特徴マップの勾配の平均値と，各特徴マップの weight を掛け合わせたものに ReLU 関数を適用する．作成した Guided BP と Grad-CAM の積が GGC となる．

3.2.3 集約

クラス毎に，作成した GGC を平均した画像を作成する．これにより，被験者間の差や試行間の差を跨いだクラスの特徴が抽象化される．抽象化された平均画像を比較することで，神経科学や生理心理学における知見の抽出を試みる．

4. 評価実験

本実験は，提案手法を用いて感情に関連する特徴的な脳活動を抽出できるかどうかを検討することを目的としている．

4.1 データセット

EEG と感情に関するデータセットの作成にあたり，32 人の被験者に日常生活を記録した写真を提示した後に，感情（快不快と覚醒度）と記憶の主観的評定を行わせた．写真が被験者に提示されている間，簡易脳波計である EMOTIV EPOC を用いて 14 チャンネル（図 1 の左上部参照）サンプリング周波数 128 Hz で EEG の記録を行った．

写真は，画面中央に注視点が表示されてから 4 秒間表示され，その後被験者は SAM（Self Assessment Maskin）に従って主観的状態の評価を行った [Bradley 94]．この手順の結果，写真を閲覧する試行に対応する 4,536 個の EEG データが得られた [野村 17]．

4.2 実験条件

DNN の学習に用いるデータセットは，写真提示の各試行の始めから 3 秒間の EEG データを抽出し，前節で示した前処理を適用した．その結果、表 1 に示されるパラメータは Width=384 (128Hz × 3sec)，Height=56 (4 周波数帯域 × 14 チャンネル) となる．これらに対して平滑化，ノイズ付与，及び Time Cropping を Data Augumentation として採用し，元の学習データを 45 倍に水増しした [Schirrmeister 17]．これらのデータのうち，訓練用データとして 4/5（学習データとして 3/4,

検証データとして 1/4)，およびテストデータとして 1/5 を用いて 5 分割交差検証を行った．

5. 結果

5.1 識別性能

学習したモデルの識別精度は 5 つの fold の平均で 46.31%であった．全 3 クラス分類問題のチャンスレートである 33.33%を上回っていることから，EEG の時間周波数成分から快-不快に関わる主観的状態を学習可能であることが分かった．

5.2 GGC を用いた可視化

5 つの fold において最高の精度であった GGC によって得られたヒートマップを図 3 に示す．これらの画像は，快，平常，不快のそれぞれのラベルが付与された EEG について，GGC によって特徴を抽象化したものである．図 3 より，快の場合にはおよそ下半分（右半球に相当するチャンネルが配置），不快の場合にはおよそ上半分（左半球に相当するチャンネルが配置）が高い値になっている．

図 3 のヒートマップについて，時間軸の値を平均 することで，頭皮上での注目強度を示す topomap[Gramfort 14] を作成した（図 4）．図に示されるように，本研究において学習された DNN は快と不快とラベルづけされた脳波に対し， 左右で異なる部位に注目していることがわかる．

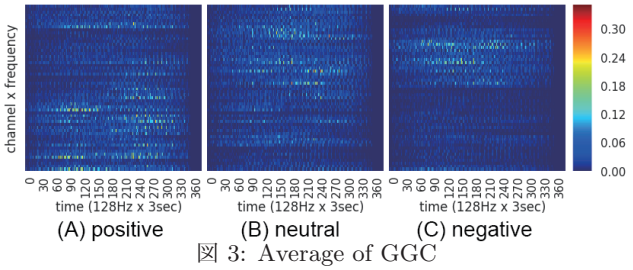


図 3: Average of GGC

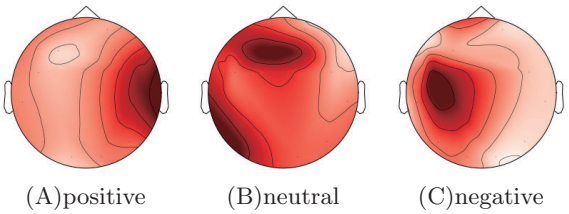


図 4: Topomap of GGC (α band)

5.3 Raw データの可視化画像

上記の結果は，DNN が快と不快を区別する脳波の特徴として,「左右半球の差分」を学習したことを示す．この結果が，DNN による学習の結果であることを確認するため，RAW データを図 3 と同様の形式で 平均化した画像（図 5），およびその topomap（図 6）を作成した．これらより，単純な平均画像においては，快と不快を区別する明確な特徴が示されないことが確認される．

5.4 神経科学的に合理的な構造の抽出 (局所度の比較)

神経科学の分野では，前頭葉の一部である DLPFC に不快感情を抑制する機能があることが知られている．それに対して快感情の生起に関わる局所的な脳部位は明らかでない．本実験においても画像を提示され不快感を抱いているときには，DLPFC などの局所的な部位が活性化している可能性がある．

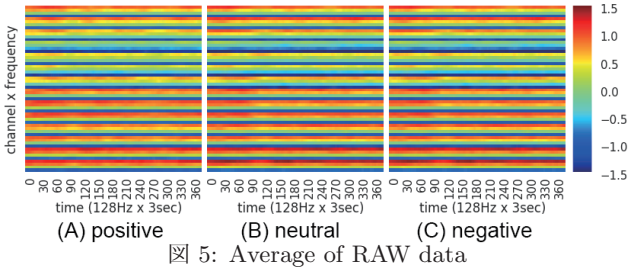


図 5: Average of RAW data

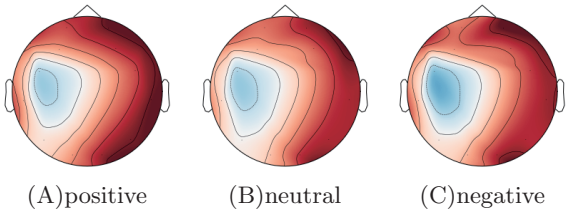


図 6: Average topomap of RAW data (α band)

そこで，各 fold における快不快の GGC 同士で局所度の比較を行った．この分析では，GGC の局所度を，GGC 画像の各セルを活性度によってソートした配列に対して得られる回帰直線の傾きとして定義する．配列の index に対して対数を取ることで片対数グラフとし，それに対して回帰分析を行った．実際に得られるグラフの例を図 7 に示す．図 7 において青線は不快時，緑線は快時の GGC の活性値と回帰直線に対応する．

図 8 は各 fold における全被験者から求められた快不快それぞれの局所度の平均である．fold02 と fold05 の間には快不快間で局所度の差に有意差がみられた (fold01: t (32) $= -2.03$, $p = 0.06$, fold02: t (32) $= -3.97$, $p < 0.01$, fold03: t (32) $= -1.55$, $p = 0.13$, fold04: t (32) $= -2.83$, $p = 0.01$, fold05: t (32) $= -4.57$, $p < 0.01$). .

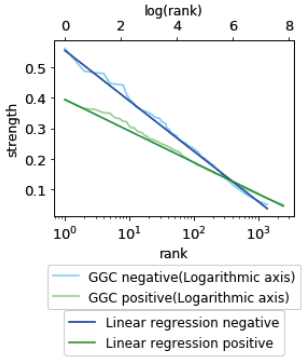


図 7: Rank regression obtained from a participant

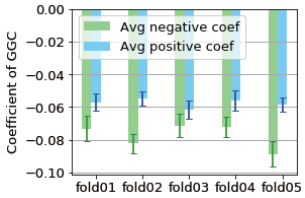


図 8: The average of the localities of each fold

5.5 信頼性の検証

前項までの結果は，提案手法により神経科学的な知見の抽出が可能であることを示唆する. しかし，今回の実験におけるモデルの識別精度は高いものではなく，得られた知見が本当に神経学的な構造を反映しているのか疑問が残る．ただし，脳波と感情という対象の特質から完全な分類は困難である．そのため，本研究では結果の一貫性を評価することで信頼性を検証する．

各 fold における快不快の GGC の類似性を比較検討した．こ

の検証を行うため，2 つの GGC によって生成された 2 つのベクトルを比較するコサイン類似度と，2 つのハッシュ行列間の距離を計算する *perceptual hash* の 2 種類の類似度を使用した [Zauner 10].

図 9 と 10 は、2 つの類似度の結果を示している． 各行列の行と列は，各 fold（*fold01 … fold05*）に対して取得された快と不快の GGC に対応し，行列のセルは類似性の値となっている．行列全体に対して，同じ感情としてラベル付けされた GGC の間に高い類似性がみられ，前項で得られた GGC の特徴に fold 間に高い一貫性があることが示されている．

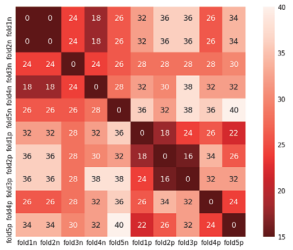


図 9: Cosine similarity between Guided Grad-CAM in each fold

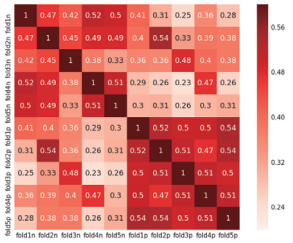


図 10: Difference in hash value of Guided Grad-CAM in each fold

6. 考察

快不快の主観的状態を識別する DNN を作成し，学習したモデルがどのような特徴抽出を行っているのか検討を行った．GGC の平均を用いることで，学習したモデルが EEG の左右差を用いて，快不快のクラス分類を行っていることがわかった．これは生理心理学分野で得られている知見とも一致する．このことから，提案手法を用いることで学習済み DNN から感情に関わる特徴的な脳活動を抽出し，妥当性の高い知見を得ることができたと考えられる．

しかし，生理心理学分野の知見では快感情のときに左前頭部，不快感情のときに右前頭部が活性化するとされているにも関わらず，GGC の平均においては逆の部分に注目している．これは DNN が EEG のどの部分が活性化していないのかに注目してクラス分類を行っているからであると考えられる (図 4c と図 6c 参照)．本実験で用いた DNN の識別精度は 46.31%であり高くない．これは，脳波と感情という扱いや分析が難しい対象を扱っていることと，DNN が注目する箇所が最も妥当とは言えないものであったことが原因であると考えられる．

GGC における注目箇所の局所性についての検証も行い，全ての fold で不快時の局所性が快時より大きく，特に fold02 と fold05 では有意差が見られた．このことから DNN は不快の識別においてより明確な構造を持っていると考えられる．この結果は，不快感情は局所的な部位が活性化し，快感情が前頭葉の広範な部位が活性化する構造を反映していると考えられる．しかし今回の検証において局所性は GGC 全体に対して求めたものであった．そのため，空間と時間で分離した局所性を求めることが今後の課題として挙げられる．

また，得られた知見には異なるデータ間で学習したモデル間で一貫性があった．しかし学習したモデルの識別精度は決して高い値ではなかった．一見矛盾しているようだが，精度と一貫性は異なる指標であると考える．識別精度はあらゆる状況に当てはまる普遍性，一貫性は得られた傾向の信頼性を表す．神経科学における研究では，あらゆる状況に対する普遍性を求めるのではなく，揺らぎのある特定の状況における傾向の信頼性を求める．そのため，本研究は神経科学において，扱うデータの精度の範囲内で信頼性の高い知見を得るツールのひとつになり得ると考える．

7. 結論

本研究の目的は，DNN を用いて神経科学的知見を得る方法を提案することであった．そのために，EEG によって快不快感情を認識する DNN を設計し，学習したモデルから感情に関連する特徴的な脳活動の抽出を試みた．結果，快不快感情間の脳活動の左右差と，不快感情における脳活動の局所性という，これまでの神経科学の研究と一貫性を示す二つの重要な脳構造を示した．これらの結果から，提案手法は人間の心の科学的理解を深めるためのツールとして有用性があると考えられる．

参考文献

[Beauregard 01] Beauregard, M., Lévesque, J., and Bourgouin, P.: Neural correlates of conscious self-regulation of emotion., *The Journal of neuroscience* (2001)

[Bradley 94] Bradley, M. M. and Lang, P. J.: Measuring emotion: the self-assessment manikin and the semantic differential, *Journal of behavior therapy and experimental psychiatry*, Vol. 25, No. 1, pp. 49–59 (1994)

[Davidson 03] Davidson, R. J.: Affective neuroscience and psychophysiology: Toward a synthesis, *Psychophysiology*, Vol. 40, No. 5, pp. 655–665 (2003)

[Gramfort 14] Gramfort, A., Luessi, M., Larson, E., D., E., Strohmeier, C., Brodbeck, D., and Parkkonen, M., L.and Hamalainen: MNE software for processing MEG and EEG data, *Neuroimage*, Vol. 86, No. 1, pp. 440–460 (2014)

[Ramprasaath 16] Ramprasaath, R., Abhishek, D., Ramakrishna, V., Michael, C., Devi, P., and Dhruv, B.: Grad-cam: Why did you say that? visual explanations from deep networks via gradient-based localization, *CVPR 2016* (2016)

[Sarlo 05] Sarlo, M., Buodo, G., Poli, S., and Palomba, D.: Changes in EEG alpha power to different disgust elicitors: the specificity of mutilations, *Neuroscience letters*, Vol. 382, No. 3, pp. 291–296 (2005)

[Schirrmeister 17] Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M., Hutter, F., Burgard, W., and Ball, T.: Deep learning with convolutional neural networks for EEG decoding and visualization, *Human brain mapping*, Vol. 38, No. 11, pp. 5391–5420 (2017)

[Zauner 10] Zauner, C.: Implementation and benchmarking of perceptual image hash functions (2010)

[野村 17] 野村太輝ら FCNN を用いた感情認識における生理心理的制約の効果, *JSAI* (2017)

Organized Session | Organized Session | [OS] OS-18

[3G4-OS-18b] 感情と AI(2)

日永田 智絵（電気通信大学）、堀井 隆斗（電気通信大学）、長井 隆行（大阪大学）

Thu. Jun 6, 2019 3:50 PM - 5:30 PM Room G (302A Medium meeting room)

[3G4-OS-18b-01] Automatic Detection of Insincere Utterances with LSTM

〇Kazuya Mio[1], Aya Ishino[2], Kazuya Mera[1], Toshiyuki Takezawa[1] （1. Graduate School of Information Sciences, Hiroshima City University, 2. Faculty of Media Business, Hiroshima University of Economics）

3:50 PM - 4:10 PM

[3G4-OS-18b-02] A Study on Concept Acquisition Method for User Preference in Dialogue for Empathy

〇Takahisa Uchida[1,2], Takashi Minato[2], Yutaka Nakamura[1], Yuichiro Yoshikawa[1], Hiroshi Ishiguro[1,2] （1. Osaka University, 2. Advanced Telecommunications Research Institute International）

4:10 PM - 4:30 PM

[3G4-OS-18b-03] A personalized model to estimate emotion of individual based on observed facial expression and generation of robot behavior

〇Kazumi Kumagai[1], Ikuo Mizuuchi[1] （1. Tokyo University of Agriculture and Technology）

4:30 PM - 4:50 PM

[3G4-OS-18b-04] Human inference system derived from emotion as a value calculation system hypothesis

〇Masahiro Miyata[1], Takashi Mory[2] （1. Graduate School of Engineering, Tamagawa University, 2. College of Engineering, Tamagawa University）

4:50 PM - 5:10 PM

[3G4-OS-18b-05] Agents which generate and estimate emotions based on logical model

〇Mai Tsukamoto[1] （1. Nara Women's University）

5:10 PM - 5:30 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

LSTMを用いた本心でない発話の自動検出

Automatic Detection of Insincere Utterances with LSTM

見尾 和哉[*1] Kazuya MIO　石野 亜耶[*2] Aya ISHINO　目良 和也[*1] Kazuya MERA　竹澤 寿幸[*1] Toshiyuki TAKEZAWA

[*1] 広島市立大学大学院情報科学研究科
Graduate School of Information Sciences, Hiroshima City University

[*2] 広島経済大学メディアビジネス学部
Faculty of Media Business, Hiroshima University of Economics

We propose a method of automatic detection of insincere utterances from voice and facial expression. Proposed method utilizes Long short-term memory (LSTM) to consider time series variation of the voice and the facial expression instead of support vector machine (SVM). The experimental results indicated that proposed method could improve recall (0.73) and F-measure (0.65) from SVM baseline.

1. はじめに

近年，人間とコミュニケーションを行う対話システムが盛んにサービス化している．対話システムがユーザとより円滑なコミュニケーションを行うためには，ユーザの感情を理解する必要がある．しかし，表出された感情が常に本心であるとは限らないため，発話が本心か否かというような抑圧された感情まで推定する技術が必要である．

そこで，本研究では，抑圧された感情まで推定するシステムの構築を目的に，本心でない発話を自動検出する手法を提案する．提案手法では，機械学習には LSTM（Long short-term memory）を使用し，特徴量として発話中の音声や表情の情報を利用する．

2. 先行研究

本研究の先行研究として，Uemura ら[Uemura 2017]の研究がある．Uemura らは，発話中の音声や表情から得られる特徴量を使用して，本心でない発話を自動検出する手法を提案している．音声特徴量としては，音声の時系列データから，openSMILE[Eyben 2010]により算出された音量の最大値や最小値などの 384 個の特徴を利用している．表情特徴量としては，発話が終わった瞬間の静止顔画像から，オムロンの OKAO Vision[オムロン]を用いて算出された 5 感情（無，喜，驚，怒，悲）の推定感情の尤度を利用している．このように，Uemura らの手法では，音声，表情ともに事前に定義された特徴量を使用して，SVM によって本心でない発話を検出している．本研究では，より細かい音声や表情の時系列での変化を把握するため，機械学習に LSTM を用いた手法を提案する．

3. LSTMを用いた本心でない発話の検出手法

本研究では，音声と表情の時系列での変化を考慮し，本心でない発話の検出を行うため，LSTM を用いた手法を提案する．LSTM とは，RNN（Recurrent Neural Network）を改良した時系列データを扱うことができるモデルである．

LSTM を用いた本心でない発話の自動検出手法のイメージ図を図 1 に示す．図 1 の縦長の長方形は，LSTM の入出力となる特徴ベクトルを模式的に表したものである．

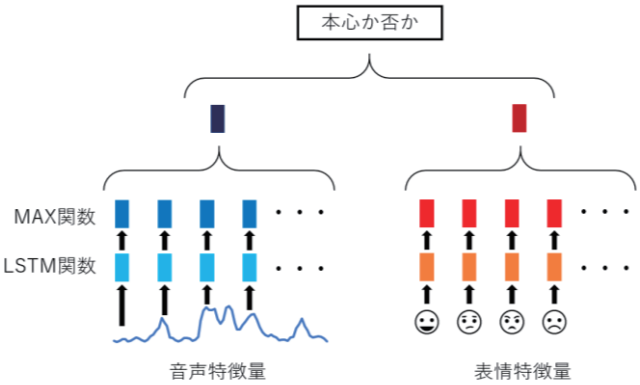


図 1 本心でない発話の自動検出手法のイメージ図

LSTM に与える特徴量について説明する．まず，音声特徴量について説明する．発話中にマイクにより録音したデータから，openSMILE を使って，25msec の窓に対して「音圧」，「基本周波数（F0）」，「自己相関関数から算出される声らしさ」の値を算出する．この窓を 10msec ごとにシフトさせ時系列の音声特徴量を得る．

次に，表情特徴量について説明する．発話中にビデオカメラにより撮影された動画から，OKAO Vision により顔部位の特徴点（左目頭，左目尻，右目頭，右目尻，鼻左，鼻右，口上，口元左，口元右）を検出し，その座標を時系列に並べたデータを表情特徴量として使用する．表情特徴量として，OKAO Vision により推定された感情を利用せず，顔部位の特徴点の座標を利用するのは，Ekman[Ekman 2009]の squelched expression（抑圧された表情）のように，顔の部位ごとの感情表出やその程度を考慮するためである．

提案手法では，音声特徴量$v_{voice,1}$と表情特徴量$v_{face,1}$を，それぞれ時系列に並べて LSTM 関数に入力し，100 次元のベクトルに変換し，次元ごとに MAX 関数を適用し$v_{voice,2}$，$v_{face,2}$を得る．$v_{voice,2}$と$v_{face,2}$を連結後，線形関数を適用し，2 次元のベクトル$v_{voice+face}$へ変換する．最後に，$v_{voice+face}$の中で最も値の大きい次元に対応するラベルを予測ラベルとする．

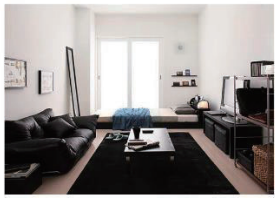
また，LSTM を双方向に拡張した BiLSTM（Bidirectional LSTM）を用いた手法でも同様の実験を行う．

連絡先：目良和也，広島市立大学大学院情報科学研究科，
mera@hiroshima-cu.ac.jp

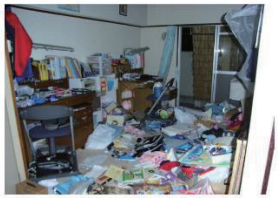
4. 実験

4.1 実験に使用したデータ

実験には，Uemura らの研究で使用したデータを利用した．Uemura らのデータ収集手順を説明する．まず，大学生 10 人に，図 2 に示すような，本心を誘発するであろう画像と，本心を誘発しないであろう画像各 20 枚を，1 枚ずつ見せた．



本心を誘発する画像



本心を誘発しない画像

図 2 データ収集に用いた画像

画像を見て，例え本意でなくても必ず褒めてもらい，その際に音声をマイクで録音，表情をビデオカメラで撮影しておき，後に特徴量として機械学習に使用した．そして発話した直後に，本心であるか否かを回答させた．これを，褒め台詞を指定する「台詞固定」と，自由な言葉で褒めてもらう「台詞自由」の 2 パターンでデータを収集し，実験に使用した．実験に使用したデータを表 1 に示す．

表 1 実験に使用したデータ

パターン	本心でない発話（正例）	本心である発話（負例）	合計
台詞固定	201	166	367
台詞自由	167	202	369

4.2 比較手法

提案手法の有効性を確認するため，以下に示す手法で実験を行った．

- SVM 手法（比較手法）: Uemura らと同様に SVM を利用した手法
- LSTM 手法（提案手法）: LSTM を利用した手法
- BiLSTM 手法（提案手法）: BiLSTM を利用した手法

LSTM 手法と BiLSTM 手法のモデルパラメータの最適化手法には Adam を使用した．隠れ層の数は 2，バッチサイズは 50，エポック数は 100 とした．評価尺度には，精度，再現率，F 値を使用し，5 分割交差検定を行った．

4.3 実験結果

実験結果を表 2 に示す．まず，台詞固定のパターンについて考察する．比較手法である SVM 手法より，提案手法である LSTM 手法は，精度は 0.02 ポイント，再現率は 0.12 ポイント，F 値は 0.07 ポイント改善することができた．また，提案手法である BiLSTM 手法では，LSMT 手法よりもさらに再現率を 0.03 ポイント改善することができた．精度の向上はわずかであったが，再現率は大幅に向上させることができた．これは，比較手法である SVM 手法では，音声，表情ともに事前に定義された特徴量を使用していたが，提案手法では，LSTM を利用することで，より細かい音声や表情の時系列での変化を把握することができたためであると考えらえる．BiLSTM では，双方向の変化を把握することができるため，より再現率が向上したと考えられる．よって，台詞固定のパターンでは，提案手法の有効性が確認できた．

次に，台詞自由のパターンについて考察する．比較手法である SVM 手法より，提案手法である LSTM 手法や BiLSTM 手法では，精度，再現率，F 値ともに低下してしまった．台詞自由のパターンでは，様々な台詞が混在しているため，音声や表情の時系列での変化を細かく把握する手法は，有効でないと考えられる．台詞自由のパターンに対応させるためには，まずは似た特徴を持つ発話をクラスタリングし，そのクラスタごとに提案手法を適用する必要がある．

5. まとめ

本研究では，機械学習として LSTM を使用することで，本心でない発話を自動検出する手法を提案した．実験の結果，台詞固定のパターンにおいては，SVM を用いた比較手法と比較し，提案手法では再現率を 0.12 ポイント向上させることができた．台詞自由のパターンにおける精度の改善が今後の課題である．

謝辞

本研究は国立研究開発法人科学技術振興機構（JST）の研究成果展開事業「センター・オブ・イノベーション（COI）プログラム」の助成を受けたものです．

参考文献

[Uemura 2017] Joji Uemura, Kazuya Mera, Yoshiaki Kurosawa, and Toshiyuki Takezawa: Suppressed Negative-Emotion-Detecting Method by using Transitions in Facial Expressions and Acoustic Features , Proc. Emotions , Decisions and Opinions 2017, pp. 122-127, 2017.

[Eyben 2010] Florian Eyben , Martin Wöllmer , and Björn Schuller: openSMILE – The Munich Versatile and Fast Open-Source Audio Feature Extractor , Proc. ACM Multimedia Conference – MM, pp.1459-1462, 2010.

[オムロン] OMRON Japan, OKAO Vision | オムロン人画像センシングサイト，https://plus-sensing.omron.co.jp/technology, (2019 年 2 月 8 日アクセス).

[Ekman 2009] Paul Ekman: Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage, W.W. Norton, 2009.

表 2 実験結果

手法	台詞固定			台詞自由		
	精度	再現率	F 値	精度	再現率	F 値
SVM 手法（比較手法）	0.57	0.58	0.57	0.54	0.53	0.53
LSTM 手法（提案手法）	0.59	0.70	0.64	0.52	0.38	0.44
BiLSTM 手法（提案手法）	0.59	0.73	0.65	0.43	0.43	0.43

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

共感を目的とした対話における ユーザの選好に対する概念獲得手法に関する検討

A Study on Concept Acquisition Method for User Preference in Dialogue for Empathy

内田貴久 *1*2 Takahisa Uchida　港隆史 *1 Takashi Minato　中村泰 *2 Yutaka Nakamura　吉川雄一郎 *2 Yuichiro Yoshikawa　石黒浩 *1*2 Hiroshi Ishiguro

*1株式会社 国際電気通信基礎技術研究所 Advanced Telecommunications Research Institute International

*2大阪大学大学院基礎工学研究科 Graduate School of Engineering Science, Osaka University

The purpose of this research is to promote users' motivation to talk with robots. In order to improve the users' motivation in dialogue, it is necessary not only to generate empathic utterances of the robots but to show that the robots understand about target items. In this paper, we explore the method to acquire the concept of the users' preferences (likes and dislikes). In the proposed method, data on preference and similarity for items are prepared for each attribute of preference. In addition, we also organize the rules for estimating preferences and similarity with fewer observed data. As future work, it is necessary to verify whether or not the dialog robots with the proposed method can improve the users' satisfaction for the robots' empathic utterances and the users' dialogue motivation.

1. はじめに

本研究の目的は，ロボットがユーザと共感を目的とした対話を行うことによって，ユーザの対話意欲を喚起することである．ここで共感とは，同じ感情や意見を持っていることを示す行為であると定義する．対話システムがユーザに共感する頻度を増やすことで，親近感やユーザ満足度が向上すると報告されている [東中 09] ことから，共感はユーザの対話意欲を向上させる要因の 1 つであると言える．ユーザの対話意欲を向上できるロボットは，例えば，高齢者の対話相手などのコミュニケーション支援への応用が期待される [横山 10]．また，共感という対話における感情の一致に関して構成論的にアプローチすることによって，知的活動である対話と感情の関係を考察できると考えられる．

ユーザに共感することを目的とした対話システムの 1 つとして，傾聴システムが挙げられる（[下岡 17] など）．傾聴とは話し手の話に共感を示しつつ，話し手がより多く話せるように手助けをして話を聴くことである [楡木 89]．対話システムが傾聴を行うことにより，高齢者の話し相手となること [山本 09] や，ユーザの話したい，話を聞いてもらいたいといった欲求を満たすこと [目黒 12] が期待されている．一方，ロボットがある対象（話題に上がっていること・もの） に関して理解能力を持ちうると感じなければ，その対象に関して対話する意欲が減退することが報告されている [内田 19]．従って，対話においてユーザの満足度を向上させるためには，話題の対象に関する意見や感想に対して共感的な発話を生成するだけでなく，ロボットがその対象に関して理解していることを示す必要があると考えられる．

そのための 1 つの方法として，対象に関する概念を有していることを示すことを考える．概念とは，アイテム同士がある観点において近い関係にあるかどうかを表すものであり，ここでは情報のクラスタリングによって形成されたカテゴリの集合と定義する．例えば，ユーザは A が好きであることを認識した場合，ロボットがある観点から A に近い B も好きなのではないかと提示する．さらに，A や B は X という観点から似ていないか，X なものが好きなのかといった発話を生成することも考えられる．このように，対話ロボットが推定したユーザの選好に関する概念をもとに発話を生成することによって，対象に関する理解能力（抽象化能力や類推能力） を示せば，その後の共感的発話に対する満足感を高めることができると考えられる．そこで本稿では，あるアイテムに関するユーザの選好（好き嫌い）に着目し，ユーザの選好に関する概念を，対話を通して獲得する手法について検討を行う．

連絡先: 内田貴久，大阪大学大学院基礎工学研究科，大阪府豊中市待兼山町 1 番 3 号，06-6850-6360, uchida.takahisa@irl.sys.es.osaka-u.ac.jp

2. 関連研究

対話においてユーザの選好を推定する研究として，ユーザの選好を考慮して対話を行うシステム [小林 16] がある．このシステムは構文や表層格，深層格の解析結果を用いて，ユーザの発話の極性を推定するなどして，ユーザの選好を推定している．提案されている手法により，現在対象となるアイテムに関する選好を推定することができる．一方，その他のアイテムに関する選好を推定することはできていないため，ユーザの選好に対する概念を獲得していることにはならない．

対話においてユーザの選好をモデル化する研究として，個人の概念空間を利用した興味の推定による情報提供システム [Sumi 00] がある．ユーザとは別の複数人の概念空間を事前に取得し，それらをもとにユーザの概念空間を推定している．しかし，選好には様々な観点が存在することが考慮されていない．例えば，レストラン A はコストパフォーマンスが良いから好きであり，レストラン B は味が良いから好きというように，同じ選好でも観点が異なる場合がある．そのようなアイテムを同じ類似性を持つとしてしまうことによって，ユーザの選好モデルを正しく推定できない場合がある．そこで本稿では，選好の観点ごとに類似性データを持ち，複数の観点に基づいて選好に関する対話を行うシステムに関する検討を行う．

3. 提案手法

本稿の目的は，複数の観点を考慮して，対話においてロボットがユーザの選好（好き嫌い）に関する概念を取得することである．そこで，選好に関するデータと選好の観点ごとに類似性

に関するデータを用意し，ユーザの選好に対する概念を獲得する．また，ユーザの選好に関する概念を獲得する上で，対話として自然さを保つことが課題となる．ユーザの選好について一問一答のように次々と質問しているだけでは，対話意欲が減退することが考えられる．そこで，少ないデータで選好や類似性に関して推定する手法が必要になる．以下，選好と類似性の表現と推定ルール，そして対話システムへの統合を想定した発話生成ルールに関して説明する．

3.1 選好と類似性のデータ表現

選好に関するデータは表によって扱う．ここで，選好の観点は表において各アイテムの属性として表現する．選好表は図 1 のようにアイテム数を M，属性数 N とすると $M \times N$ の行列である．選好に関するデータは，好きであることを 1，嫌いであることを-1，好きでも嫌いでもないことを 0 としてデータが格納されている．データがない場合は空白になっている．図 1 の例では，レストラン A をメニューの観点からは好んでいるが，店の雰囲気の観点からは嫌いであることを表している．M と N の値は対話を通して増加していく．

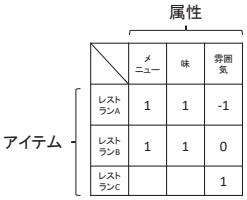


図 1: 選好表

類似性に関するデータも表によって扱う．類似表は図 2 のようにアイテム数を M とすると $M \times M$ の行列である．類似性に関するデータは，関連があることを 1，関連がないことを 0 としてデータが格納されている．データがない場合は空白になっている．類似表に関しては属性ごとに類似表を用意するため，N 個の表を用意する．

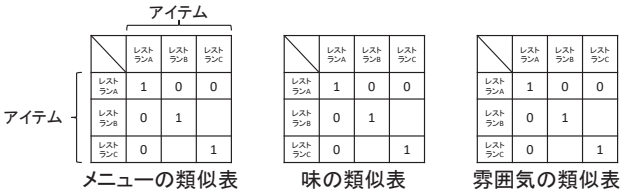


図 2: 類似表

本研究で対象とする対話では，ユーザとの発話から得られるデータ（観測値）とユーザとのこれまでの対話から推定されるデータ（推定値）がある．観測値に関するデータ操作を更新，推定値に関するデータ操作を補完と称する．これら選好表と類似表が選好に関する概念となる．

3.2 表の補完ルール

ここで，観測値から推定値を補完するルールに関して説明する．表は選好表と類似表の 2 種類あるため，選好表から選好表の補完，選好表から類似表の補完，類似表から選好表の補完，類似表から類似表の補完の 4 通りが考えられる．以下，それぞれの補完ルールに関して説明する．ここで，今回はシステム（自分）とユーザ（相手）それぞれの選好・類似表があることを仮定する．また，推定値はユーザへのデータ確認発話によって肯定されれば以後，観測値として扱うことにする．そして，各ルール内で補完データ間の整合性が取れない場合（例えば，あるアイテムとの類似性データによる選好表の補完では好きであると補完される一方，別のアイテムとの類似性データによる選好表の補完では嫌いであると補完される）は補完を行わないことにする．以下の図における <> 付きの数値は補完された推定値を表す．

3.2.1 選好表から選好表の補完

ある属性について自他の選好の観測値がいくつかのアイテムについて一致していた場合，その属性に関する選好は他のアイテムについても自他で一致するだろうという推測である．具体的には図 3 のように，ある属性の選好データに関して観測値 の類似度が閾値以上の場合，対象属性の一方の選好データと同じ値を他方の推定値として補完する．選好表のある属性における類似度 S は，アイテム数を M，比較する二つの表で一致する選好データの数を s とすると，$S = s/M$ で表される．

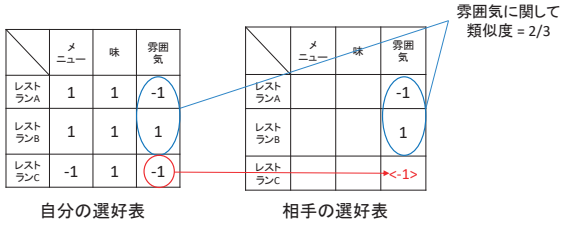


図 3: 選好表から選好表の補完

3.2.2 選好表から類似表の補完

ある属性について 2 つのアイテム間で選好が一致している場合に，それらのアイテムがその属性について類似していると推定する．具体的には，図 4 のように，ある属性に関して，選好に関する観測値 がアイテム間で一致している場合はそのアイテム間の類似度を 1，一致していない場合は 0 を推定値として類似表を補完する．ここで，選好が一致するものは，同じ類似性をもつことを仮定している．この補完が自身の表で生じる場合というのは，対話を通して自身があるアイテムをある観点で好き（嫌い）になった時に，アイテム間が類似していることに気づくという現象である．

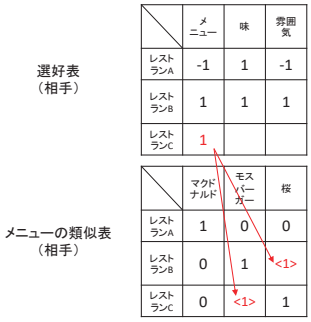


図 4: 選好表から類似表の補完

3.2.3　類似表から選好表の補完

3.2.2 節と逆の場合で，ある属性について 2 つのアイテムが類似している場合に，それらのアイテムのその属性についての選好が一致していると推定する．具体的には，図 5 のように，ある属性の類似表にある 2 つのアイテム間の観測値 が存在していれば， その属性に関して一方のアイテムの選好の観測値と同じ値を推定値として他方のアイテムの選好を補完する．ここで，ある属性に関して類似するものは，同じ選好 をもつことを仮定している．この補完が自身の表で生じる場合というのは，現時点でアイテム A を好き（嫌い）であり，さらに対話を通してアイテム A とアイテム B が似ていると知ったときに，アイテム B も好き（嫌い）になるだろうと思う現象である．

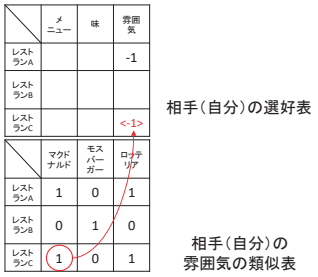


図 5: 類似表から選好表の補完

3.2.4　類似表から類似表の補完

自身（相手）が似ていると思っているものは，相手（自身）も似ていると思うだろうという推定である．具体的には，図 6 のように，自他の同じ属性の類似表に関して，一方の類似表の観測値を他方の類似表にそのまま推定値として補完する．これは片方が知らないアイテムに関して行われる補完ルールである．相手の表から自身の表を補完する場合というのは，ある属性についてアイテム A と B が似ているかどうか知らないときに，相手が似ていると発言したら，自分もそのように思い込む現象である．

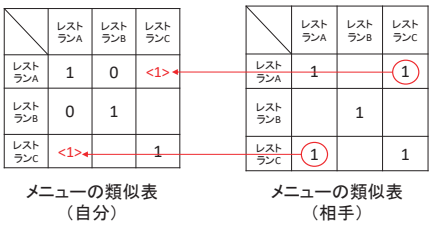


図 6: 類似表から類似表の補完

3.3　発話生成ルール

対話を通して選好の概念を獲得するために，システムが行う発話には，2 種類の発話を想定する． 1 つは表のセルを埋める 発話で，もう 1 つは表のセルを増やす発話である．表のセルを埋める 発話には，表 1 のように，選好と類似性ごとに表のセルを埋めるための発話と補完したセルを確認するための発話がある．

また，表のセルを増やす発話には，アイテムを追加する発話（例 :「好きな飲食店はどこですか？」）と，属性を追加する発話（例 :「なぜ A が好きなんですか？」）がある．

表 1: 表のセルを埋める発話の分類

	データ取得発話	データ確認発話
選好	例）「A のメニューは好きですか？」	例）「それじゃあ，A のメニューは好きですよね？」
類似性	例）「A の雰囲気は B に似てますか？」	例）「それじゃあ，A と B の雰囲気は似てますよね？」

システムはユーザと対話を行いながら，効率的に選好表を埋めることが望まれる．各種の発話を行うタイミングは，補完が行われるごとに情報確認発話を行い，これ以上補完できない場合は情報取得発話を行う．

3.4　対話例

ここで，上述の 4 つの補完ルールを用いて対話を行うことを考える．これらの補完ルールを用いることによって， システムが既知のアイテムに関する選好概念を獲得できるだけでなく，対話を通して新規に知ったアイテムに関する選好概念の獲得も可能になる．対話例を以下に記す．

System: メニューが好きなレストランはどこですか？
User: レストラン A です.【レストラン A は新規のアイテムとする】
System: レストラン A のメニューはレストラン B に似てる？
User: いいえ，似ていません．
System: それじゃあ，レストラン B のメニューは嫌いなのでは？【類似表から選好表の補完 : 既存のアイテムに関する選好に対する概念を獲得】
User: 好きでも嫌いでもないですね．
System: そうなんですね．レストラン B とメニューが似てないんだったら【類似表から類似表の補完】，それなら私も好きかもしれないです！【類似表から選好表の補完 : 新しいアイテムに関する選好に対する概念の獲得】
User: そうなんですね．
System: レストラン Aってレストラン C とメニュー似てるんじゃない？【選好表から類似表の補完】
…

4.　考察

本稿では，共感を目的とした対話において，ユーザの選好概念を獲得するための手法に関して検討を行った．提案手法を用いることにより，観点ごとにユーザが何を好むのかについて理解を示すことが可能となる．また，システムが既知のアイテムに関する選好概念が獲得できるだけでなく，対話を通して新たに知ったアイテムに関する選好概念の獲得が 可能になると考えられる．新たな概念を獲得するシステムは， ユーザがすぐにインタラクションに飽きてしまうという，現在のコミュニケーショ ンロボットが抱える問題を解決する手法として期待されている [内田 08]．今後は，ロボット提案手法によって選好概念を推定しながら共感対話を行った場合に，その共感に対する満足度を高めることができるかどうか，ユーザの 対話意欲を喚起するかどうかを検証することが課題となる．

今回の検討では，選好や類似性に関するデータはシステム

とユーザのものに限って考察した．一方，第三者のデータを導入することによって，より精度の高いユーザの選好概念を獲得できると考えられる．その場合，例えば，協調フィルタリング [Shardanand 95] の手法を適用することができる．協調フィルタリングは，多数のユーザの情報をもとに対象ユーザがどのユーザに近いのかを計算するものである．これにより，選好表から選好表の補完ルールや選好表から類似表の補完ルールを置き換えることが可能であると考えられる．しかし，多くの情報フィルタリング手法では，ユーザのプロファイルが有効になるのに膨大な情報・時間を必要とする [土方 04] ことが指摘されていることから，どの程度のデータサイズであれば有効に機能するのかを検証する必要がある．また，類似表から選好表の補完は，例えば，ガウス過程回帰 [Williams 06]，類似表から類似表の補完は無限関係モデル [Kemp 06] の枠組みを適用できると考えられる．データのサイズと適用可能な手法の関係を調査することも今後の課題となる．

謝辞

本研究は JST ERATO 石黒共生ヒューマンロボットインタラクションプロジェクト（グラント番号：JPMJER1401）の一環として行われたものです．

参考文献

[Kemp 06] Kemp, C., Tenenbaum, J. B., Griffiths, T. L., Yamada, T., and Ueda, N.: Learning systems of concepts with an infinite relational model, in *AAAI*, Vol. 3, p. 5 (2006)

[Shardanand 95] Shardanand, U. and Maes, P.: Social information filtering: algorithms for automating “word of mouth”, in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 210–217ACM Press/Addison-Wesley Publishing Co. (1995)

[Sumi 00] Sumi, K., Sumi, Y., Mase, K., Nakasuka, S.-i., and Hori, K.: Information presentation by inferring user’s interests based on individual conceptual spaces, *Systems and Computers in Japan*, Vol. 31, No. 10, pp. 41–55 (2000)

[Williams 06] Williams, C. K. and Rasmussen, C. E.: Gaussian processes for machine learning, *the MIT Press*, Vol. 2, No. 3, p. 4 (2006)

[横山 10] 横山祥恵, 山本大介, 小林優佳, 土井美和子 他：高齢者向け対話インタフェース-雑談継続を目的とした話題提示・傾聴の切替式対話法, 研究報告音声言語情報処理 (SLP), Vol. 2010, No. 4, pp. 1–6 (2010)

[下岡 17] 下岡和也, 徳久良子, 吉村貴克, 星野博之, 渡部生聖：音声対話ロボットのための傾聴システムの開発, 自然言語処理, Vol. 24, No. 1, pp. 3–47 (2017)

[山本 09] 山本大介, 小林優佳, 横山祥恵, 土井美和子：高齢者対話インタフェース:『話し相手』となって, お年寄りの生活を豊かに, 電子情報通信学会技術研究報告. HCS, ヒューマンコミュニケーション基礎, Vol. 109, No. 224, pp. 47–51 (2009)

[小林 16] 小林峻也, 萩原将文：ユーザの嗜好や人間関係を考慮する非タスク指向型対話システム, 人工知能学会論文誌, Vol. 31, No. 1, pp. DSF–A_1 (2016)

[土方 04] 土方嘉徳 他：情報推薦・情報フィルタリングのためのユーザプロファイリング技術, 人工知能学会誌, Vol. 19, No. 3, pp. 365–372 (2004)

[東中 09] 東中竜一郎, 堂坂浩二, 磯崎秀樹：対話システムにおける共感と自己開示の効果, 言語処理学会第 15 回年次大会発表論文集, pp. 446–449 (2009)

[内田 08] 内田ゆず, 荒木健治：画像に対する発話を対象とした名詞概念獲得システム SINCA, 知能と情報, Vol. 20, No. 5, pp. 685–695 (2008)

[内田 19] 内田貴久, 港隆史, 石黒浩：対話アンドロイドに対する主観的意見の帰属と対話意欲の関係, 人工知能学会論文誌, Vol. 34, No. 1, pp. B–I62_1 (2019)

[目黒 12] 目黒豊美, 東中竜一郎, 堂坂浩二, 南泰浩 他：聞き役対話の分析および分析に基づいた対話制御部の構築, 情報処理学会論文誌, Vol. 53, No. 12, pp. 2787–2801 (2012)

[楡木 89] 楡木満生：積極的傾聴法, 医学教育, Vol. 20, No. 5, pp. 341–346 (1989)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

経験に基づく個人の感情推定モデル構築方法及びロボット行動生成への利用

A personalized model to estimate emotion of individual based on observed facial expression and generation of robot behavior

熊谷和実 *1　水内郁夫 *1
Kazumi Kumagai　Ikuo Mizuuchi

*1東京農工大学
Tokyo University of Agriculture and Technology

We aim to construct a system which allows robot to learn/generate a behavior favorable to an individual through interaction. The core idea is estimating emotion based on facial expression observed during a robot behavior to self-evaluate how the behavior is favorable to the individual. Based on the self-evaluated data gathered in the past interaction experiences, when the robot does some behaviors, our method is expected to help the robot to predict emotion of the individual gradually and properly. In this paper, we show the user studies to evaluate the proposed method. The study resulted that a robot changed its behavior to get better reaction. As future works, we discuss how robot should use observed data or knowledge to personalize robots behavior.

1. 個人に気に入られるロボット

本研究の目的は、人とのインタラクションを繰り返す中で個人が気に入る行動を学習してゆくロボット行動生成手法の構築である。本研究では、特定の個人とロボットとのインタラクション [Dautenhahn 04] を研究対象とする。

感情を推定する技術 [Picard 01] をロボット行動生成手法に適用し、社会的にロボットが振る舞い [Breazeal 03] それを人に気に入ってもらい長期に渡るインタラクションを実現する [Leite 14] ためにはシステムがどのような情報の処理をするべきか研究を進めてきた。著者は、ドラえもんを参考にオリジナルのロボットをつくり、表 1 にまとめた様々なインタラクション状況を想定し、実験を行なった。

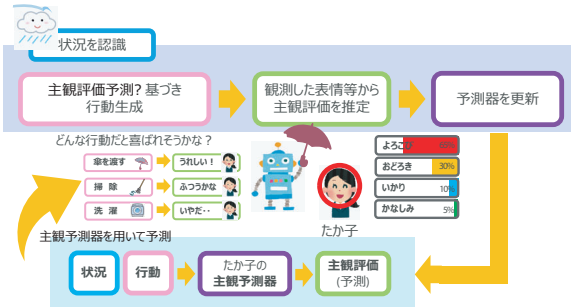


図 1: 実際の観測表情を利用した感情推定結果に基づくロボット行動の個人適合手法

2. 関連研究

2.1 感情推定を利用したコミュニケーションロボット

音声・画像などを用いた感情推定について多く研究されてきている [Zeng 09][Sariyanidi 15]。コミュニケーションロボットの多くは、感情推定結果をロボット行動開始のトリガや、行動前の状況判断に用いる。そのようなロボットの戦略の多くは比較的大多数に受け入れられる行動パターンとなっている。公共の場で不特定多数を相手にコミュニケーションするロボットには、この手法は有用である。しかし、家庭内やチューターロボットなどのようにロボットが特定の個人とインタラクションする場合、個人により好き嫌いが別れる行動の万人受けデザインは難しく、特に、嫌いなものよりも好きなもののほうが好みが分かれる。ロボットがした行動も好みによってはを相手のひとが気に入らない可能性があるという問題点がある。

2.2 個人に合わせた行動をするロボット

子供と一緒に学習するチューターロボットに関する研究では、子供の振る舞いに合わせてロボットが行動することの良い効果が多く報告されている [Leyzberg 14]。多くの実験において、遠隔操作によって人間が子供の反応を観察した結果に応じてロボットの行動が選ばれる。子供の性格や反応に応じてロボット行動を変えてゆくシステムを自動化する手法については発展段階である。

Gordon らによる研究では、タブレットを用いたゲームによる子供とのインタラクションを通してロボットの振る舞い（声掛けの種類など）を子供の反応（表情とゲームの戦局）に応じて決定する [Gordon 16]。しかしこの手法では、行動をどの方向に修正するかが分からないという課題がある。ロボット行動の表現方法、子供の過去の反応に基づきロボット行動を修正する方法をいかに実現するかが課題となっている。

3. 感情推定結果に基づくロボット行動の個人適合手法

人がなにかされたときに見せた表情や返事など、外部に表出された感情は、行動に対する評価系として扱うことができると考えられる。本研究では従来の手法とは逆の方向性で感情推定結果をロボット行動に利用する。すなわち、ロボットが行動を開始してから後に人間がみせた反応（表情や返事など）を観測し、次にインタラクションをする際の行動生成に利用する。図 1 にメインフローの概要を示す。例えばポジティブな感情であれば、してもよい行動とラベル付をし、次回以降に行動生成する際に、しても良い行動を選択する頻度を高める。

連絡先: 熊谷和実，東京農工大学，東京都小金井市中町 2-24-16，042-388-7206，kumagai@mizuuchi.lab.tuat.ac.jp

表 1: 過去に実施した実験において想定したインタラクション状況・感情推定手法・ロボット行動

インタラクション状況	感情推定手法	ロボット行動・表現方法	他の因子	参考文献
家庭内	表情（ロボット動作中の平均）	声かけ（自然言語）・単純動作	曜日・時間帯・天気・忙しさ・体勢	[Kumagai 14]
インタラクティブメディアアート	表情（ロボット動作中の平均・変化量）	動作の大きさ・速さ等	人の動作と表情	[Kumagai 18]
クイズゲーム	表情	声かけ（自然言語）・単純動作	クイズ戦局	[Kumagai 16]

3.1 ロボット行動開始後の感情予測器の構築

感情予測器は、ある入力に対する出力の関係を学習する。本研究であつかう予測器の入力が意味するものは、個人の心的状態の変化や感情を引き起こす因子（ロボット行動を含む）とする。人の感情を引き起こす因子の候補のうちのひとつにロボット行動が含まれるという想定である。

どのような因子を設定するかは、ロボットが置かれた状況によって適切に決められるものとする。出力の意味は、ロボットがした行動の後に観測した表情などから推定された個人の心的状態変化や感情とする。また、因子の表現方法（数値化など）の設計が学習収束までのプロセスにおけるロボット行動戦略の挙動に影響を与える。

3.2 感情推定部

感情は表情や声のトーン、姿勢、心拍数など様々な生理指標に基づき推定するツールが多く出回っている。これで得られた感情推定結果を感情予測器の出力として利用する。

まず考慮すべき点は、感情の表現方法である。感情の表現方法については多く議論されてきているが [Plutchik 84], [Schlosberg 54], [Russell 80] 本手法では、扱いたい感情をターゲットとし、連続値で表現することを想定する。本研究においては表情の種類（喜び表情・悲しみ表情など）を複数種類想定し、各種類の表情強度をある値からある値までの連続値で表現している。表情には直接現れないが心の中に生じている感情が異なると、同じ種類の表情でも変化の仕方（立ち上がり、立ち下がりなど）が異なる研究がよく知られているように [Ekman 82]、どのタイミングの感情推定結果を利用するかに応じて、良いとされる行動が変わる可能性がある。

また、学習のための使いやすさの点において、各指標ごとに特徴があることを考慮する。例えば、声は出すときと出さないときがある。一方で、表情は信頼度低め（ロボット行動に対する表情ではない可能性が高いこともある）ではあるが、何かしらの結果を得ることができるという特徴がある。

また、個人の性格に応じて、どの指標に基づく感情推定結果がもっともらしいかということが異なる。その点も学習プロセスに影響することを考慮する必要がある。

4. 表情を感情予測に利用した行動生成実験

4.1 プロトタイプとしての行動選択実験

実験のシチュエーションは、家庭内で行われることを想定した場でのコミュニケーションとした。時間、曜日などの環境的因子と、忙しいか否か、などの人間自身の因子の組み合わせを状況として設定し、ロボットは状況に応じて発話と動作のセットを選択した。動作はシンボル的意味を連想しにくいもの（くねくねする）を５パターン用意した。発話は１０パターン用意した。それぞれの発話は「おかえりなさい」「がんばってね」など、相手に何かしら感情を持たせる意図をもたせた。

感情推定部分では、ロボット行動中に観測した表情点数を利用し、それらの平均値を感情推定結果として感情予測部分の入出力関係を学習した。ロボットの動作中に観測した表情点数の重み付き和を評価点とした。各表情の種類毎にヒューリスティックに決められた重み係数は、ポジティブな表情は正の値の係数とした。実験では状況表示された PC の前に実験参加者が座り、PC 画面に表示されたロボット動作中の表情は PC に取り付けたカメラで認識した。

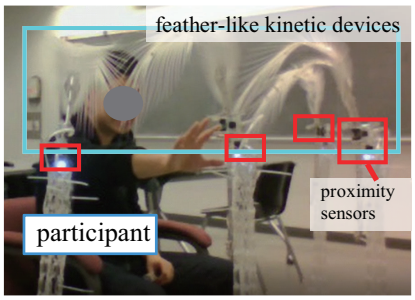


図 2: 実験システムと実験参加者とのインタラクションの様子。手を近づけると動作を開始する。[Kumagai 18]

大学院生 1 名に対して行なった実験では、インタラクションの回数増加に伴い、次第に感情推定結果の点数が向上している様子が見られ、後半に点数が減少する様子が見られた。

4.2 連続値表現された動作生成実験

有限パターンの行動からの選択では実現困難な任意の動作を生成する。ロボット行動を任意の軸で張られた空間上の点として表現する。より高い評価点を得るよう修正する方向をシステムに分からせることができると考えられる。

図 2 に示す白い枝のようなエージェント (LAS=Living Architecture System[Gorbet 15]) と人間がインタラクションする実験を行なった。

動作時に観察した人間の表情に基づいて、動作パラメタと人間の表情との関係を学習し、状況に応じた適切な動作パラメタへと修正した。動作パラメタは、動作速度・動作の大きさ・動作が隣のエージェントへ遷移してゆく方向の 3 つとした。

実験は大学生・大学院生の男女 14 名に対し実施した。実験前半において良い表情が見られた動作パラメタに近いパラメタを後半で選ぶ様子がみられた。また、提案手法により生成した動作とランダムに生成した動作に対するそれぞれの評価点を比較した。提案手法により生成した動作に対する評価点が高い実験参加者は表情に感情が表出する度合いが高い傾向にあった。ただし、パラメタ空間が膨大だと、十分な学習収束までに多くのインタラクション回数を要する。そのため、学習が収束する前に人が飽きたり、好みが変わるなど評価系が変わると、適切な動作を生成することが困難という制限もある。

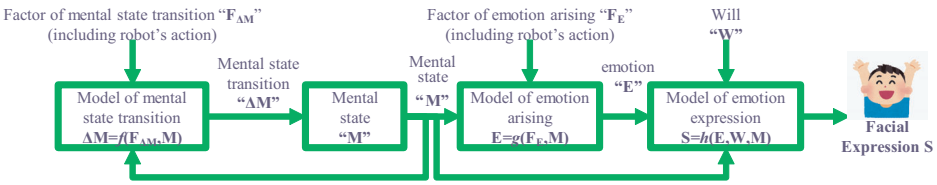


図 3: 情動・表情・心的状態・因子の関係図 [Ekman 75][Zajonc 89][Ekman 99][Forgas 92][Ekman 71]

5. 外部から直接観測困難な心的状態の推定

5.1 心的状態を因子に含む感情予測モデル

実験において、飽きなどが原因で行動に対する評価系が変化したと考えられる。本研究の手法は、逐次学習の形をとるため、評価系の遷移に追従するためにロボットは複数回ほどのインタラクション回数を要する。飽きて退屈になっているような心の状態（以下「心的状態」と呼ぶこととする。）を「状況 (因子)」に含める形でロボットが把握し、感情予測システムの更新が心的状態の変化による評価系の変遷に追いつくまでのインタラクション回数を減らすことを目指す。

飽きや気分など、持続時間が比較的長い心的状態と情動とを区別し、情動発生に影響を与える因子の一つとして心的状態を扱うようなモデルをロボットに持たせるフレームワークを考える [Forgas 92]。図 3 は [Ekman 75][Zajonc 89][Ekman 99][Forgas 92][Ekman 71] などの知見をまとめ、心的状態・情動・因子（ロボット行動含む）・外部への情動表出（表情など）の関係をブロック線図で示したものである。各ブロックの関数 $f(), g(), h()$ はそれぞれ心的状態変化モデル・情動発生モデル・表情表出モデルと呼ぶことにする。観測データや既存の知識に基づき心的状態の変化・発生する情動・表出する表情を予測するモデルをインタラクションを通してロボットが構築してゆく手法を考える。

5.2 因子の観測可能性の高い場での心的状態推定

感情の変化を尤もらしく予測するために因子も尤もらしく推定したい。決定論的モデルを用いて心的状態はどれほど推定可能であるか検討するため、心的状態変化を引き起こす因子の多くを観測データから推定可能な場で実験を行なった。実験の状況は、プレイヤーに瞬発的な反応や集中力を求めるもぐら叩きとした。

実験は参加者 1 名に対して行なった。ゲーム 1 回につき制限時間約 10 秒間に全 7 匹のもぐらを穴から弾き出すルールである。実験ではゲームを 100 回実施し、各ゲーム終了毎に、心的状態量として現在の気分のよさを-10 点から 10 点で評価した。

本実験では、もぐら叩きゲームの戦局と 1 ゲーム前の気分の良さを状況因子 (入力)、ゲーム前後の気分の良さの変化量を推定対象 (出力) とし、入出力関係を 3 層のニューラルネットワークで学習した。Leave-One-Out 法を用いた交差検定を行い、アンケート結果に基づき算出した気分変化量と、学習結果に基づく気分変化量の計算結果との差分 (以下、誤差と呼ぶ) について検証した。実験の結果は、過去に似た経験をしたときほど気分変化予測結果の誤差が小さく、また予測変化量が大きいほど誤差が大きくなる傾向であった。

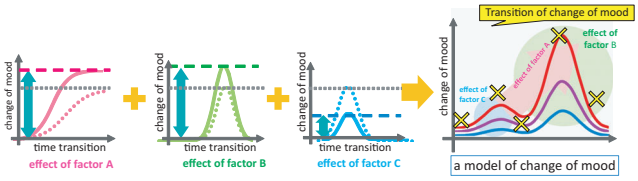


図 4: 非線形関数の重み付和による心的状態推定モデルの個人適合 [Kumagai 16]

6. 心的状態推定モデルの個人適合手法と行動生成への利用

6.1 因子・表情・心的状態の関係を表す非線形関数の重み付き和によるモデル

本手法では、個人に合わせた心的状態推定のために、既存の心的状態推定モデルの重み付和により因子毎の影響度合いを調節し個人に合わせた心的状態推定モデルを構築する。既存の心的状態推定モデルとは、ある因子と心的状態と表情との関係性を表すものとする。

人とロボットによる協調作業など状況が限られている場合は、人の心的状態の変化に共通パターンが見られることが多いため、推定時に考慮する他の因子の影響が一般的なものに設定されている。しかし、好みの分かれそうなロボット行動を含めた因子により引き起こされた心的状態の変化や情動発生などは、影響因子が個人に応じて異なると考えられる。

観測した表情と因子の関係から特徴量を抽出し、寄与率が高いと考えられるモデルの重みを大きくする (図 4)。ニューラルネットワークの構造の高次の層を既存の関数に置き換えたことにほぼ相当すると考えられる。すでに確立された心的状態推定モデルがあれば、それらの組み合わせを使って個人に合わせた推定モデルを構築する。一方で、様々な因子に関する心的状態推定モデルを複数持つ方が良く、適切なものが無い場合は用意する必要がある。

6.2 クイズゲームにおける行動選択実験

クイズゲーム中の実験参加者に対しロボットが動作・発話をする実験を行なった。ロボットはクイズの戦局に応じて実験参加者を応援するような発話をした (クイズは PC 画面に表示)。

推定モデルの構築に用いた非線形関数はクイズ状況を想定しヒューリスティックに設計した。心的状態量は気分とした。実験の前後に、実験結果の評価のための気分評価アンケートを実施した。(アンケートの回答内容は行動選択には用いられていない。) 実験は大学生・大学院生男女 8 名を対象に実施した。今回の実験の結果においては、表情観測結果のみよりも心的状態を考慮した場合の結果が良い気分ポイントがより高い参加者の方が多い結果となった。また、表情変化が比較的少ない実験参加者の実験結果は低い傾向であった。

7. 議論

7.1 表情認識の性能に依存する行動生成システムの性能

本研究のポイントは、表情を用いることである。(多くの実験では OKAOVision[OMRON] という表情認識を利用した。) ロボット行動時に観測した表情に個人の感情が表れていたとき、表情予測結果に基づく行動は相手に喜ばれる可能性が高かった。一方、表情が通常あまり変化しない実験参加者に関しては点数が向上しない傾向が見られた。今後は、複数生理指標を用いた感情推定手法を利用して個人の感情表出方法に合わせたモデル構築手法を目指す。

7.2 ロボット行動の特徴に応じた表情処理

表情の扱い方はロボット行動の意味的特徴や動作的特徴に影響を受けると考えられる。行動に対する評価が表情等に現れるまでの時間を考慮したり、ロボット動作の速度等で相対化し表情変化を捉えるなどの対応が考えられる。

8. まとめ

本稿では著者らがこれまでに行ってきた研究について紹介した。我々の研究手法のコアアイデアは、ロボットが行動したときに個人が見せた表情を用いてロボット行動の気に入られる度合を評価することと、行動の気に入られる度合いを予測する予測器を過去の経験に基づいて更新してゆくことである。

提案手法について紹介し、関連研究・実験について述べた。表情予測結果に基づく行動生成実験では、インタラクション回数の増加に伴い個人からの評価点が増加したり、良い評価点を得る期待値の高い方向へとパラメタを変えた動作を行う結果が一部の実験参加者において見られた。感情が表情など外部に表出される過程で内部状態量が働くと想定した心的状態変化モデルを行動選択に用いた実験では、表情のみから動作の評価を学習するときよりも心的状態を考慮し学習したときの方が実験前後の心的状態変化が良い結果となるケースが多かった。

ロボット行動戦略を更新し生成された行動をしたときに、より高い評価点を得る結果が見られた一方で、表情の変化が少ない実験参加者や、感情と表情の相関が低い実験参加者の結果は向上しない傾向であった。評価関数の構築における感情推定結果の利用方法（表情認識結果の平均を取るか変化量を取るかなど）に課題があると考えられる。今後は、ロボット行動の特徴（意味的特徴のような抽象度の高い特徴から動作の物理量など）に応じた感情推定結果の利用方法について調査しロボット行動手法に応用する。

参考文献

[Breazeal 03] Breazeal, C.: Emotion and sociable humanoid robots, *International Journal of Human-Computer Studies*, Vol. 59, No. 1-2, pp. 119–155 (2003)

[Dautenhahn 04] Dautenhahn, K.: Robots we like to live with?!-a developmental perspective on a personalized, life-long robot companion, pp. 17–22IEEE (2004)

[Ekman 71] Ekman, P.: Universals and cultural differences in facial expressions of emotion., in *Nebraska symposium on motivation*University of Nebraska Press (1971)

[Ekman 75] Ekman, P. and Friesen, W. V.: *Unmasking the Face: A Guide to Recognizing Emotions From Facial Expressions*, Englewood Cliffs, N.J.: Prentice-Hall (1975)

[Ekman 82] Ekman, P. and Friesen, W. V.: Felt, false, and miserable smiles, *Journal of nonverbal behavior*, Vol. 6, No. 4, pp. 238–252 (1982)

[Ekman 99] Ekman, P.: Basic Emotions, in T., D. and M., P. eds., *Handbook of Cognition and Emotion*, pp. 45–60, John Wiley and Sons Ltd (1999)

[Forgas 92] Forgas, J. P.: Affect in social judgments and decisions: A multiprocess model, in *Advances in experimental social psychology*, Vol. 25, pp. 227–275, Elsevier (1992)

[Gorbet 15] Gorbet, R., Memarian, M., Chan, M., Kulic, D., and Beesley, P.: Evolving Systems within Immersive Architectural Environments: New Research by the Living Architecture Systems Group, *Next Generation Building*, Vol. 2, No. 1, pp. 31–56 (2015)

[Gordon 16] Gordon, G., Spaulding, S., Westlund, J. K., Lee, J. J., Plummer, L., Martinez, M., Das, M., and Breazeal, C.: Affective Personalization of a Social Robot Tutor for Children's Second Language Skills., in *AAAI*, pp. 3951–3957 (2016)

[Kumagai 14] Kumagai, K., Baek, J., and Mizuuchi, I.: A Situation-Aware Actino Selection based on Individual Preference using Emotion Estimation, in *Proceedings of the 2014 IEEE International Conference on Robotic and Biomimetics*, pp. 356–361 (2014)

[Kumagai 16] Kumagai, K., Hayashi, K., and Mizuuchi, I.: Estimating mood by determining weights of pre-defined basis functions based on observed facial expressions and situations, in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 964–965 (2016)

[Kumagai 18] Kumagai, K., Lin, D., Men, L., Blidaru, A., Beesley, P., Kulić, D., and Mizuuchi, I.: Towards individualized affective human-machine interaction, in *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 678–685 (2018)

[Leite 14] Leite, I., Castellano, G., Pereira, A., Martinho, C., and Paiva, A.: Empathic robots for long-term interaction, *International Journal of Social Robotics*, Vol. 5, No. 3, pp. 329–341 (2014)

[Leyzberg 14] Leyzberg, D., Spaulding, S., and Scassellati, B.: Personalizing robot tutors to individuals' learning differences, pp. 423–430ACM (2014)

[OMRON] OMRON, : OMRON's Image Sensing Technology, `https://plus-sensing.omron.com/technology/`

[Picard 01] Picard, R. W., Vyzas, E., and Healey, J.: Toward machine emotional intelligence: Analysis of affective physiological state, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 23, No. 10, pp. 1175–1191 (2001)

[Plutchik 84] Plutchik, R.: Emotions: A general psychoevolutionary theory, *Approaches to emotion*, Vol. 1984, pp. 197–219 (1984)

[Russell 80] Russell, J. A.: A circumplex model of affect., *Journal of personality and social psychology*, Vol. 39, No. 6, p. 1161 (1980)

[Sariyanidi 15] Sariyanidi, E., Gunes, H., and Cavallaro, A.: Automatic Analysis of Facial Affect: A Survey of Registration, Representation, and Recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 37, No. 6, pp. 1113–1133 (2015)

[Schlosberg 54] Schlosberg, H.: Three dimensions of emotion., *Psychological review*, Vol. 61, No. 2, p. 81 (1954)

[Zajonc 89] Zajonc, R. B., Murphy, S., and Inglehart, M.: Feeling and facial efference: Implications of the vascular theory of emotion, *Psychological Revies*, Vol. 96, pp. 395–416 (1989)

[Zeng 09] Zeng, Z., Pantic, M., Roisman, G. I., and Huang, T. S.: A survey of affect recognition methods: Audio, visual, and spontaneous expressions, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 31, No. 1, pp. 39–58 (2009)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

感情の価値計算システム仮説にもとづく人の推論システムの提案

Human inference system derived from emotion as a value calculation system hypothesis

宮田 真宏[*1] 大森 隆司[*2]
Masahiro Miyata Takashi Omori

[*1] 玉川大学大学院 工学研究科 [*2] 玉川大学 工学部
Graduate School of Engineering, Tamagawa University College of Engineering, Tamagawa University

There were many modeling studies about the emotion. But most of them were phenomenological and don't approach to the brain and/or cognitive mechanism. In this study, we consider a possibility of its computational modeling based on an idea that an emotion in wider sense is a value calculation system for an action decision. However, conventional inference methods could not explain human intuitive inference and logical inference. In previous study, we propose an integrated model in which the intuitive inference is represented as a search process of in a continuous and distributed associative memory, and is switched to a symbolic inference mode that biases an associative gain when it found values during the intuitive inference search. In this study, we show a merit of emotion based action decision that combines inference, reinforcement learning and reflex.

1. はじめに

近年の人工知能(AI)技術の発展は目覚ましく，今後生み出される製品には多様な AI 技術が用いられ，さらにその製品にはより人間的な知的機能が求められると予想される．その典型例に対人インタラクションがある．対人インタラクションとは，基本的に人を対象とした広義のコミュニケーションを指す．それを実現するエージェントには，コミュニケーションする相手の感情や意図，要望などを理解し，それに応じた自身の意思決定を行う機能が求められよう．つまり相手の行動の原因を知り，その原因に応える自分の行動を決めることが質の高いコミュニケーションにつながる．さらに，人の意思決定には感情（価値）が関与し，その評価によりに行動することが知られている[ヴィンター 2017]．そのため，相手との円滑なコミュニケーションのためには，相手の意図や目的の背後にある人の価値観の理解もまた必要であろう．その意味で，人の価値計算の理解は高度な対人インタラクション AI の実現には欠かせないと考える．

また，人は価値と同様に感情によっても意思決定を行う．感情が介入することは時として不合理なこともある．しかし，感情による意思決定は私たちの日々の生活の多くの場面において，合理的な価値計算と同等の意思決定の結果をすばやく得ている[ヴィンター 2017]．このように考えると，意思決定に対する感情の役割は価値計算と同様と考えられる．以降，本稿では感情と情動はその内容の複雑性については区別するが，機能的な意味においては基本的には同一であると考えて議論する[信原 2017]．

従来，感情に関しては多くのモデルが提案されてきた[Ekman 1997][Russell 1980]．しかし，その多くは現象として観察される感情を記述したモデルであり，認知や脳内処理のメカニズム，さらにその計算論的な意味に迫るものではない[Chernavskaya 2015] [Samsonovich 2013]．

また戸田は，人の複雑な感情から行動決定に至るまでの心的過程に対して定性的な説明理論を構築した[戸田 1992]．そこで戸田は人の感情とは，自らの置かれている状態に価値を割り振り意思決定に至るまでの計算過程であるとしている．しかし，ここでも感情の具体的な処理過程については述べていない．

一方で[Koelsch 2015]は，人間の広義の情動を生存・安全・愛着・経済的価値の 4 つの要素に分解し，これらを個別の脳部位に対応付けることでそれぞれの役割について説明を試みている．この研究では，情動に経済的価値も含まれており，眼窩前頭葉がその機能を司るとしている．それではこれらの感情に関するモデルの背後にある脳機能を考慮した計算的なメカニズム，さらにその役割はどのようなものであろうか．本研究ではこれまでに，感情とは動物が行動する際，意思決定に用いる価値計算システムの表出結果の一つであると仮定し，その計算モデル化の可能性を議論してきた[Miyata 2017]．そこでは，従来提案されてきた手法の組み合わせにより，実現できると考えていた．

しかし，従来行われてきた人のシンボル的な処理を Tree 探索で実現する従来の推論手法は，人の脳内処理を考慮すると違和感がある．人は複雑な推論過程を脳内の約 1000 億個あるとされる神経細胞の発火の連鎖により実現している．その処理が直接的に Tree 探索のようなシンボル的な処理をしているかと考えると疑問である．この問題に対して本研究ではこれまでに，価値に駆動された連想記憶モデルを用いることで，従来言われてきた早く，確率的かつ並列的に探索する直観的推論を相互想起により，意識的かつ時間のかかる論理的推論を自己想起のモデルによりそれぞれ実現できることを示してきた[宮田 2018]．神経回路による連想記憶は 1970 年代から知られていたが[Nakano 1972]，最近では注目されていない．しかし脳全体は大規模な連想記憶とも考えられ，推論との関係での記憶や知識の利用法は検討に値しよう．

本研究にて考える推論の役割は，予測状態空間における，価値を最大化する意思決定のための価値へ到達する状態空間中の経路の探索であると考える．こう考えると，従来より多く研究されてきた強化学習は，期待報酬予測という形で状態空間中に価値に到達する経路のマーカーを付けていくアルゴリズムと解釈できる[Sutton 1990]．また進化の過程にて埋め込まれていると考えられている反射行動は，自身の不利益な事態を避ける行動を生成すると考えると価値計算が含まれていると言えよう．このように考えた際，従来研究されてきた多様な意思決定アルゴリズムは価値の最大化という意味で共通の基盤を持っている．

そこで本稿ではまず，本研究にて考えている価値計算システム仮説の概要について説明する．その後，本仮説の検証を考えた際にキーとなる推論について従来言われてきた推論手法との違い，および推論システムに求められる機能，および連想記憶によるその実現方法について説明する．そして最後に，推論，強化学習および反射とを組み合わせることによる意思決定システムの全体像について議論する．

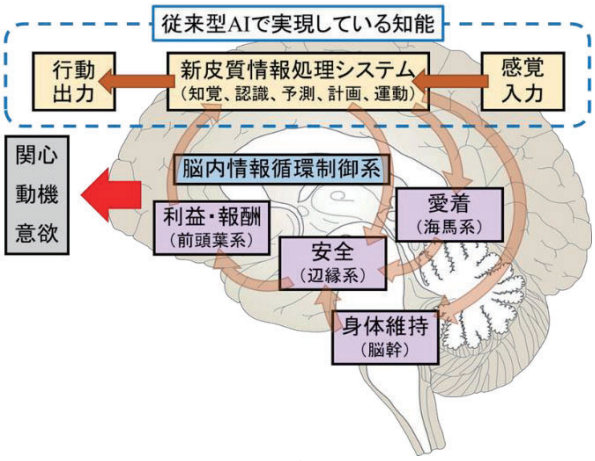


図 1　感情＝価値計算システム仮説の概要

2.　感情＝価値計算システム仮説

2.1　感情と価値計算システムとの関係

我々の感情は意思決定に大きく影響を及ぼす．また自分自身の感情の表出は他者にも伝達され，さらには他者の意思決定にまで影響する．現在，行動経済学や脳科学では人の意思決定は報酬（価値）の計算に基づくとされている[大竹　2012][信原　2017]．これより我々は，人などの動物の持つ感情とは自身の置かれている内的・外的な状態を処理した結果，見出された価値を含めて何らかの計算をし，その結果の表出されたものであると考えた．さらに感情の背後には脳の神経細胞の発火の連鎖による処理結果としての価値計算システムがあると想定する．本研究では，価値と感情表出の対応はおおよそ固定であり，その結果として感情が価値評価の伝達システムとして的確に機能すると考える．さらに感情のメカニズムの理解には，むしろ脳内処理としての価値計算システムの理解が先に必要であろう．

それに対して従来の AI では知能の要素，および処理手順を開発してきた．まず初めに視覚や聴覚などを処理する感覚情報処理．次いで感覚情報処理の結果を用いた予測や意思決定などの知的情報処理．そして最後にこれらの結果を出力する行動生成の処理，という順である．それに対して本研究では，中段の知的情報処理の背後にある報酬系，すなわち価値について考える．本稿で考える価値計算システムでは，新皮質の知的情報処理系に感覚情報が入力され，その認識結果は直ちに脳の多様な価値を計算する処理系に伝達されると想定する．なお，我々の考える感情の価値には Koelsch の提案した「生存」「安全」「愛着」「経済的価値」の4種を想定する．しかし，これらに対応する脳部位（脳幹・辺縁系・海馬系・前頭葉系）は彼らの考える脳部位よりも広範囲であると考える．本研究にて考える脳部位を統合すると，大脳皮質下の脳部位の大半が対象になる(図 1)．

2.2　価値計算システムにおける推論の立ち位置

本稿で考える推論とは，過去に例の無い新奇の場面での意思決定のための価値に向かう経路探索とする．先述の通り，強化学習や反射行動もまた時間的に長いスケールで価値計算を含んでいると言える．すなわち，多様な意思決定アルゴリズム，すくなくとも反射・強化学習・推論は価値の最大化という意味で共通の基盤を持っている．

この世界で行動エージェントが出会う全ての状況を図 2 の大きな円で表している．そのうちの一部は進化の過程で進化の過程で先天的に反射として埋め込まれており，また他の一部は過去の経験から強化学習などによりその場に応じた行動が学習済みであると考えられる．これらの状況では，個々の場面に応じた行動あるいは価値が割り振られており，エージェントは探索をせずとも意思決定ができる．しかし残りの未経験領域は価値が割り振られていない新奇の状況であり，エージェントは推論などにより，現在状態から価値の割り振られた状況までの行動経路を発見することで，未経験の現在状態における行動の価値を計算し行動することが求められる．このように状況空間における価値を基準に学習アルゴリズムを配置すると，少なくとも反射，強化学習，推論は一続きとなり，それらの間の関係が明らかとなる．

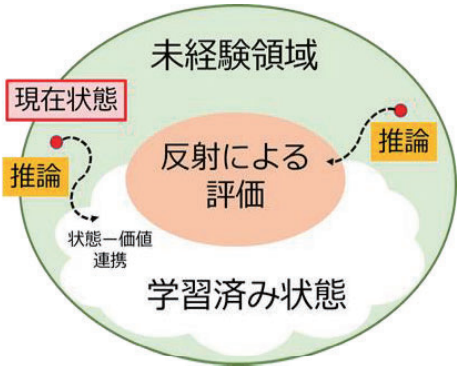


図 2　推論による価値探索の位置づけ

3.　人の推論

認知科学では，人の推論には直観的推論(システム 1)と論理的推論(システム 2)の二種類ある[Evans 2009]とされ，それぞれを別のシステムとしてモデル化してきた(表1)．推論の制御は，まずシステム1にてヒューリスティックな過程が行われ，必要に応じて分析的過程であるシステム2が実行されることで意思決定がなされるという二重過程モデルが議論されてきた[服部 2015]．

それに対して，従来の AI における推論の基本モデルには Tree 探索があげられる．Tree 探索は個々の離散状態の予測，およびその評価を行う意識的かつシンボル的な方式である[Russel 2008]．また，我々はシンボル的推論とは別に，何かを知覚するとその影響の予測と評価を素早く行う直観的な推論過程も持っている．これは，感覚刺激からの自動的な連想による無意識的な予測と評価によると考えられる．これを踏まえると，予測は局所的で多方向に並列的に進み，それを価値計算システムが部分的に評価したものが，直観と呼ばれていると考える[大森　2017]．課題は，行動的には異なるこの二つのシステムがどうやって脳で実現されていか，である．これまで論理的推論の脳部位というものは見つかっておらず，また論理的推論は確実に意識とも関連があり，そのメカニズムは謎のままである．

表 1　推論の二重過程と二重システム仮説

直観的推論	論理的推論
作業記憶は不要	作業記憶が必要
無意識的，自律的	意識的， メンタルシミュレーション
推論が速い	推論が遅い
バイアスに影響されやすい	規範的，公平
文脈依存	抽象的
確率的，分散的	論理的，シンボル的
暗黙知（経験的確率）を利用	明示的な知識を利用
推論が浅い	深い推論が可能
進化的に古い	進化的に新しい

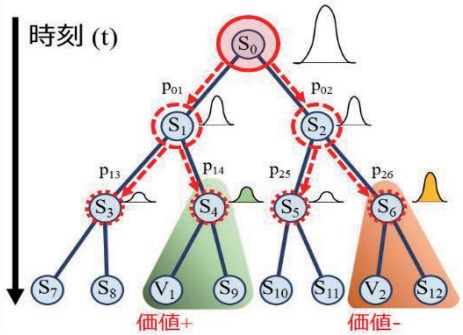


図 3　推論の処理過程イメージ

4. 連想記憶モデル

4.1 連想記憶による推論

連想記憶とは，記憶パターンを貯蔵し，部分的な記憶情報を基に必要な記憶を読み出す機能である．神経回路による分散型の連想記憶モデルとしては，記憶事項を一つのベクトルで表したアソシアトロン[Nakano 1972]や時間的な連想を含んだモデル[Sompolinsky 1986]，さらには連想記憶にシーケンシャルな制御を加えた PATON[Omori 1999]などがある．これらのモデルでは，複数個の記憶事項の記銘はそれらの相関行列の和(記憶行列)で表し，想起用の入力ベクトルと記憶行列の積の計算で想起を再現できる．本稿ではこの相関行列型の連想記憶モデルを用いて，人の二種類の推論のモデル化およびその検証を行う．

4.2 推論システムの全体像

本研究で目指す推論の型を図 3 に示す Tree 構造とする．記憶パターンは Tree の各ノードに対応し，探索は初期状態S_0から開始することとし，状態遷移は初期状態S_0から 2 分木が 3 階層だけ継続するとした．シミュレーションは，推論の連鎖の結果，正の価値のあるノードV_1，または負の価値のあるノードV_2の周辺領域に到達し，価値を見出すことができた場合に意思決定ができるとした．なお，一般的な探索問題の研究ではより深い複雑な課題を扱うことが多いが，本研究ではまず連想記憶モデルをベースにした推論行動の創発を示すことが目的であるため，単純なトイ問題を対象とする．

図 4 は本研究で想定する連想ベースの推論システムの全体像である．入力層x_tの情報を相互想起または自己想起の連想行列にかけ合わせることで想起パターンx_{t+1}が得られ，x_{t+1}はさらに価値認識層に送られ，そこに含まれる成分についての評価と意思決定処理が行われる．

直観的推論では，x_tは相互想起行列 W^e にかけ合わせることで次の時刻の状態x_{t+1}を得て，それはそのまま次の時刻の入力になるという連想の直線的な連鎖を想定する．一つの状態から連想する際にその連想関係が分岐することがある(実際は多い)ため，複数の記憶パターンが混在した想起パターンが得られる．さらに次の計算では複数の記憶パターンが混在した状態を基に次の連想をするため，さらに分岐が広がるという並列的な想起が起こり，結果として並列的で幅広い探索が実現される．この計算は処理が単純であるため，無意識的すなわち自動的な過程で実現可能と考えられるが，処理を繰り返すと多くの記憶の混合パターンが想起されることで個々の記憶パターンのゲインが小さくなるため，深い推論は実現できない．

それに対して論理的推論では，直観的推論にて複数の記憶パターンの混合パターンが想起された後，そのうちの一つの記憶パターンを選択的に想起する自己想起過程を追加することで，記憶パターンの混合を回避する．そのため，混合記憶パターンが一つの記憶パターンに収束するまでに時間がかかるが，想起パターンのゲインの減衰のない深い推論が実現できる．なお，本研究で想定する相互想起過程と自己想起過程の切り替えは，モデル的には単純なパラメータの切り替えで実現でき，その動的なスイッチングが一見して複雑な確率的処理と論理的処理の混合した過程を実現することとなる．

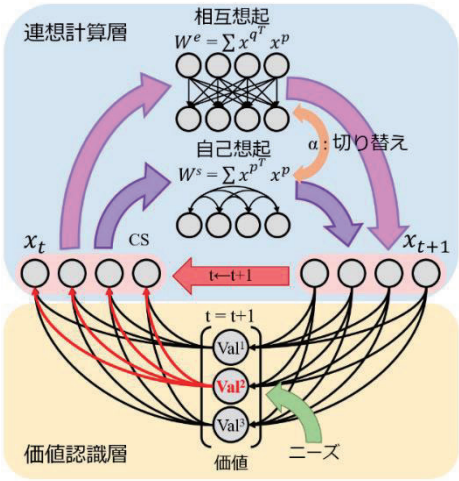


図 4　推論システム全体像

4.3 直観的推論と論理的推論の実現

人の推論に直観的推論と論理的推論の二種類があるとすると，状況に応じてどちらかを選択して推論を制御する上位システムが必要となろう．しかし，そのような上位システムは知られておらず，そもそも論理的推論の脳部位も知られていない．その解釈として本研究では，これら 2 つの推論は 1 つの処理システムの別の姿であり，その動作モードのスイッチングにより振る舞いが変わると考えた．その過程を連想記憶モデルの連想と仮定して表現すると，その計算式は以下になる．

$$x_{t+1} = \alpha \left(\sum_q Pr\left(x_{t+1}^q \middle| x_t^p\right) W_{pq}^e\right) x_t^c + (1-\alpha) \sum_q W_{pq}^S x_t^c \quad \cdots(1)$$

(1)式右辺の第一項は，直観的推論を実現する相互想起の項である．ここでは過去の経験を表現する記憶ベクトルからイベントごとの連想行列(W_{pq}^e)を作成し，入力ベクトル(x_t^c)から想起される記憶ベクトルを連想的に探索する．なお，記憶ベクトル群は直交していると想定する．これより想起ベクトルは，入力ベクトルに関係があるほど連想行列から強い強度で想起される．そして想起されたベクトルは過去の経験頻度を表す条件付き確率に基づき強さが決まる．この計算の反復により，短時間で広い範囲に対する連想的な探索が可能となる．そして探索中に価値のある部分を発見した際には，その価値に従い意思決定する．

(1)式右辺の第二項は，第一項で見出された価値に焦点を当て，その価値を最大化させる記憶想起の反復計算を表す．すなわち，入力ベクトル中の価値のある成分を強化して，想起によって特定の価値に対応する記憶パターンのみが支配的になるまで自己想起型の連想計算を反復し，その記憶パターンに対応する行動を選択する(図 4) [宮田 2018]．

5. 計算機シミュレーション

これまで本研究では，シミュレーション環境を作りこんだ迷路タスクを題材に，連想記憶を用いた推論システムが人の推論シ

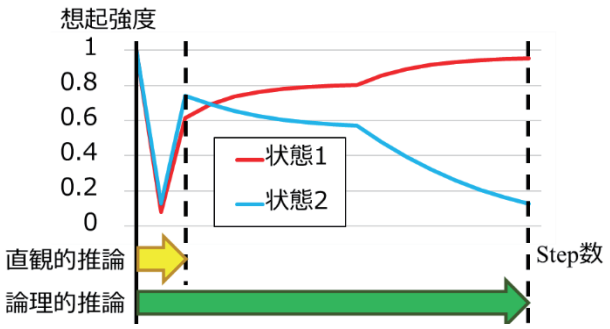


図 5　連想記憶を用いた 2 つの推論の違い

ステムの統合的な説明が可能であることを示してきた．これにより，図 2 の価値を中心とすることで反射，強化学習，推論を組み合わせたシミュレーションの実行を試みる．そのため本稿では，価値計算システム仮説の検証のために迷路課題を用意した．

この課題の想定では，エージェントははじめ，迷路の経路情報を知らず，ランダムウォークにより地図を経験して経路を記銘する．そして，報酬が目の前に現れた際には反射行動により目の前にある報酬を得て，報酬地点の近傍のみに強化学習により価値を割り振っていく．そして，推論による価値探索を行う．

報酬を得た際には，エージェントは地図内のランダムな位置に再配置されるため，すぐに推論による探索が可能になるとは限らない．しかし，何度も経験を重ね，かつ強化学習により地図空間中に価値の伝播が広がることで，地図空間中の様々な位置からの推論が可能になる．そしてシミュレーションの結果，反射・強化学習・推論という三種の行動決定方式を組み合わせた課題において，本研究にて提案した推論システムは問題なく動作する結果を得た．

6.　まとめ

これまでに本研究にて考える感情の価値計算システム仮説の概要，および連想記憶と価値の導入により実現される論理的・直観的の二種類の推論の実現方法を示してきた．さらに，本研究にて提案した価値駆動の推論システムと，反射行動および強化学習とを組み合わせ，迷路タスクにおいて効果的な行動決定が可能であることを示した．

今後は，まず図 2 にて示した環境をより実世界に近い環境にしたうえでのモデルの評価を行う必要がある．例えば物理法則と捕食者や仲間を含んだ仮想環境などを用いることにより，推論・強化学習・反射，さらにはエピソード記憶も統合したシステムの有用性を検証する．その後，必要に応じて 3 要素の手法を切り替えることで，その場の状況に応じた柔軟な行動を示すエージェントを作成する．これを実現することは，人のような感情の効用と発生メカニズムの解明に寄与することを期待する．本研究は文部科学省科研費 15H01622 の助成を受けた．支援に感謝する．

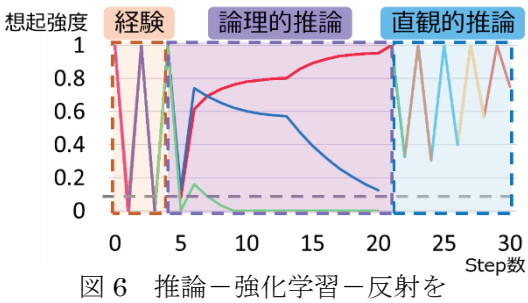


図 6　推論－強化学習－反射を組み合わせたシミュレーションの結果

参考文献

[ヴィンター 2017] エヤル・ヴィンター:愛と怒りの行動経済学 賢い人は感情で決める，早川書房，2017

[信原 2017]　信原幸弘: 情動の哲学入門 価値・道徳・生きる意味，勁草書房，2017

[Ekman 1997]　P. Ekman et al: What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS), Series in Affective Science, Oxford University Press, 1997.

[Russell 1980]　J. A. Russell: A circumplex model of affect, Journal of Personality and Social Psychology , vol.39 , pp.1161-1178, 1980.

[Chernavskaya 2015]　Chernavskaya, : An architecture of the cognitive system with account for emotional component. Biologically Inspired Cognitive Architectures. , BICA Journal, 2015, doi:http://dx.doi.org/10.1016/j.bica.2015.04.009

[Samsonovich, 2013] Samsonovich : Emotional biologically inspired cognitive architecture. Biologically Inspired Cognitive Architectures. , BICA Journal, 2013. doi:http://dx.doi.org/10.1016/j.bica.2013.07.009

[戸田 1992]　戸田正直: 感情，東京大学出版，1992.

[Koelsch 2015]　S. Koelsch et al: The quartet theory of human emotions: An integrative and neurofunctional model, Physics of Life Reviews, Vol. 13, pp. 1-27, 2015.

[Miyata 2017]　Masahiro Miyata, Takashi Omori : Modeling emotion and inference as a value calculation system, BICA2017, Vol. 123, pp. 295-301, 2017

[宮田 2018]　宮田真宏，大森隆司: 価値に駆動された連想記憶に基づく人の推論過程の統合，第 34 回 ファジィシステムシンポジウム，2018

[Nakano 1972]　Kaoru Nakano : Associatron-A Model of Associative Memory, IEEE, 1972

[Sutton 1990]　Richard Sutton: Integrated Architectures for Learning, Planning, and Reacting Based on Approximating Dynamic Programming, Appeared in Proceedings of the Seventh Int. Conf. on Machine Learning, pp. 216-224, 1990

[大竹 2012]　大竹文雄ら: 脳の中の経済学，ディスカヴァー携書，2012

[Evans 2009]　Jonathan St. B. T. Evans et al.: How many dual-process theories do we need? One, two, or many? , Oxford Scholarship Online, 2009

[服部 2015]　服部雅史: 思考と推論: 理性・判断・意思決定の心理学，北大路書房，2015

[Russel 2008]　Russel, Norvig 他:エージェントアプローチ人工知能第 2 版，共立出版，2008

[大森 2017]　大森隆司，宮田真宏:粒子モデルと価値評価系による直観的推論の計算アーキテクチャ，p.55-56,日本神経回路学会全国大会講演論文集，2017

[Sompolinsky 1986]　Haim Sompolinsky: Temporal Association in Asymmetric Neural Networks , Physical review letters, The American Physical Society, 1986

[Omori 1999]　Takashi Omori et al: Emergence of symbolic behavior from brain like memory with dynamic attention, Neural Networks Vol. 12, No. 7-8, pp. 1157-1172, 1999

[宮田 2018]　宮田真宏，大森隆司: 連想記憶モデルに基づく人のシンボル的推論のモデル化，第 8 回 人工知能学会 汎用人工知能研究会，2018

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

論理モデルに基づく感情生起と感情推測を行うエージェント

Agents Which Generate and Estimate Emotions Based on Logical Model

塚本 麻衣
Mai Tsukamoto

奈良女子大学大学院 人間文化研究科 博士前期課程 情報衣環境学専攻
Department of Computer Science and Clothing Environment,
Graduate School of Humanities and Sciences, Nara Women's University

Recently, research studies on robots that have emotions and can have dialogues with humans are progressing. By making it possible to guess the emotions of others, it is possible to generate more kind of human emotions corresponding to the guessed emotions. In addition, it is desirable that the generation of one's own emotions and the estimation of others' emotions can be performed under the same condition, but such research is not seen in the past. Therefore, in addition to our previous implementation based on the combination of OCC theory and BDI model, we have made it possible to estimate others' emotions using the same definitions as those we have used so far. Furthermore, whereas we had to give inputs from the environments to our system in the form of logical formulas, we aim to implement robots that generate emotions based on the beliefs extracted from the sentences such as conversation.

1. はじめに

近年、ロボットに人間らしい感情を持たせる研究が進んでいる。特に行動選択において、ロボットが抱いた感情を反映できれば、より人間らしい行動がとれると考えられる。そのためには、行動選択と感情の生起が共通の機構で表現、および実現できることが望ましい。

自律エージェントのモデルである BDI モデルとの相性のいい、OCC theory[OCC 88] と Adam らの研究 [Adam 09] による感情の定義を用いることによって、BDI logic で行動選択と感情生起の両方を実現することが可能となる。従来、我々は Adam の論文による形式化を基に、感情を生起するロボットの実装と複数の感情の同時生起、感情の度合いの実現、時間経過による度合い減少と感情の削除などの拡張を行っている [今井 18][浅井 18]。しかし、感情生起の条件となる信念を論理式の形で入力する必要があった。また、生起される感情やその妥当性を調べる手段がなかった。

そこで、本研究では与えられた日本語文を構文解析することで信念を取り出し、それを用いて感情生起を行うよう実装を行っている。現段階では、構文解析に適した形に手直しした文からの感情生起が可能となっている。これにより、物語文をサンプルとしての感情生起の妥当性の検証が可能となる。また、音声認識と組み合わせることで会話からの感情生起など、環境からの自然な入力による感情生起への応用も期待できる。

さらに、従来の実装では他者の感情の推測を行っておらず、そのため他者の感情を生起条件に含む自身の感情の生起についても実装が不十分であったため、他者の感情推測を行うように実装を拡張した。他者の感情の推測についても自身の感情生起と同一の条件に基づいて行えることが、自然な感情生起のためには望ましいため、両者で条件として同一の論理式を用いる形で実装の拡張を行っている。

連絡先: 塚本 麻衣, 奈良女子大学人間文化研究科博士前期課程情報衣環境学専攻生活情報通信科学コース, sam_tsukamoto@cc.nara-wu.ac.jp

2. 関連研究

感情に関する研究には、機械学習を用いたものもみられる。[堀宮 12] では、他者反応を含めた文章に感情ラベルを付加し、学習させることによって、感情を示す言葉や、直接的な感情表現のない文章に対しても、その後に続く他者からの反応によって感情の推定を可能としている。しかし、機械学習を用いた研究では、感情の生起理由の説明が困難である。それに対し本研究では、感情生起を OCC theory の形式化による論理式で行っているため、生起した感情に対し、その理由を明確にできる利点がある。また機械学習にはより精度を上げるために多量の教示データが必要となるが、多量のデータを実験で得ることが難しいなどの問題があるため、我々は採用していない。

ロボットに感情を持たせるにあたり、『ポチは散歩した事を喜んだ』のように、どの事柄に対して感情を生起しているのか正しく表現することは、感情が生起した理由説明に用いることができるなどの理由で望ましく、感情に関する既存研究においてそのような研究もみられる [遠藤 06]。これは、「うれしい」「かなしい」等の感情表現となる語句の係り元となる 2 文節を抽出し、末尾に「のが」「ことが」を含むものを感情生起表現（感情が向けられる事象）として得ることを可能としている。しかし、感情表現となる語句が存在しない文が与えられた場合は感情の推定が困難であり、感情生起表現が存在しない文が与えられた場合は、生起した感情を向ける事象を明確にすることが困難であると考えられる。本研究では、4 章で述べるように、一つの文から人物、事象、信念の種類と度合いを得ることによって信念を生成し、特定の信念を得たとき感情生起を行うため、文中に感情表現となる語句が存在しない場合も信念生成、感情生起を行うことを可能としている。また、動詞、または動名詞を感情を向ける事象として抽出するため、主語述語が明記された文であれば、感情生起表現が存在しない文であっても感情を向ける事象を明確にすることを可能としている。

推定した他者の感情からの行動選択を行う研究や他者の行動から自身の感情生起を行う研究も多くみられる。例えば [Tang, 14] では、話者の発話による音響的特徴によっての感情推定を可能としており、推定した感情に合わせた対話の実

現を目指しているが、自身の感情生起は行っていない。また、[石川 18] では、他者の説得に対する受託状況によって自身の感情状態を遷移させ、感情状態に合わせた対話を試みているが、他者の感情の推測は行っていない。人間らしさのある自然な行動を選択するにあたり、他者の感情の推測と自身の感情生起の両方を行い、生起、推測した感情に沿って行動を選択するよう実装を行うことが望ましいと考えられる。またその際、自身の感情生起と他者の感情推測を同一の条件に基づいて行うことが妥当である。本研究では自身の感情生起と他者の感情推測を同一の条件に基づいて行うよう実装の拡張を行った。

3. 他者の感情の推測

従来の研究 [今井 18][浅井 18] では、他者の信念を得ることで生起する感情が存在するが、他者の感情の推測については行われていなかった。しかし、より人間らしいエージェントの実装を目指すにあたり、他者の感情の推測を行い、推測した感情に対応した自身の感情の生起を行うことが望ましいと考えられる。推測には、自身の感情生起と同一の条件に基づくことが妥当といえるため、感情の生起条件となる信念を感情推測の対象が持っていると信じた時、対象の感情の推測を行うよう実装した。

以下に、同じ感情に対する自身の感情生起の条件と他者の感情推測の条件の例を挙げる。

感情 Joy（喜び）の生起は以下の形式的定義によって行われる。

$$Joy^{i}_{f(d)}\,\phi := Des^{i}_{d}\,\phi \wedge Bel^{i}\,\phi \tag{1}$$

これは「エージェント i にとってイベント ϕ の成立が d 程度望ましく、エージェント i がイベント ϕ の成立を信じた時、エージェント i はイベント ϕ の成立に対する $f(d)$ 程度の感情 Joy を生起する」という意味である。ここで、$Des^{i}_{d}\,\phi$ は「エージェント i にとってイベント ϕ の成立は d 程度望ましい」、$Bel^{i}\,\phi$ は「エージェント i がイベント ϕ の成立を信じている」という意味である。イベントの成立に対する望ましさの度合い d の値が大きいほど、感情 Joy の程度も大きくなると考えられるため、感情 Joy の度合いを計算する関数 $f(d)$ は増加関数と定義している。

次に他者の感情 Joy（喜び）の推測は以下の形式的定義によって行われる。

$$Guess_Joy^{i,j}_{f(d)}\,\phi := Bel^{i}\,Des^{j}_{d}\,\phi \wedge Bel^{i}\,Bel^{j}\,\phi \tag{2}$$

これは「エージェント j にとってイベント ϕ の成立が d 程度望ましいとエージェント i が信じ、エージェント j がイベント ϕ の成立を信じているとエージェント i が信じた時、エージェント i はエージェント j のイベント ϕ に対する $f(d)$ 程度の感情 Joy を推測する」という意味である。ここで、$Bel^{i}\,Des^{j}_{d}\,\phi$ は「エージェント j にとってイベント ϕ の成立が d 程度望ましいとエージェント i が信じている」、$Bel^{i}\,Bel^{j}\,\phi$ は「エージェント j がイベント ϕ の成立を信じているとエージェント i が信じている」という意味であり、以下の等価関係が成り立つ。

$$Guess_Joy^{i,j}_{f(d)}\,\phi \equiv Bel^{i}(Des^{j}_{d}\,\phi \wedge Bel^{j}\,\phi) \tag{3}$$

これは「『エージェント j にとってイベント ϕ の成立が d 程度望ましく、エージェント j がイベント ϕ の成立を信じている』ことをエージェン i が信じている時、エージェント i はエージェント j のイベント ϕ に対する $f(d)$ 程度の感情 Joy を推測する」という意味である。

この式の Bel^{i} の中身は式 (1) の右辺と同じ形をしており、他者の Joy の発生の推定を自身の Joy の生起と同じ条件で行っている事を示す。他の感情の推測に関しても同様に、推測対象が感情の生起条件の信念を持っていることを信じた時、対象の感情を推測を行う。これにより、同一の条件に基づいた感情生起と他者の感情の推測を可能とした。

4. 構文解析

従来の研究 [今井 18][浅井 18] では、事例として物語文からの感情生起を扱っていたが、感情生起の条件となる信念を、与えられた物語文から人間が解釈して抽出し、実装に用いているエージェント記述プラットフォームである Jason[Jason 07] での表記によって手動で与えることで感情生起を行う必要があった。しかし、これでは環境からの自然な入力から感情生起することができない。したがって、エージェント自身が文の中から必要な信念を取り出し、感情生起を行うことが望ましい。そこで、本研究では与えられた文を構文解析することで信念を論理式として取り出し、感情生起を行うよう実装した。

例えば「ポチは散歩したい」という文が与えられると、

『ポチ (名詞) は (助詞) 散歩し (動詞「する」の連用)
たい (助動詞「希望」』

のように解析でき、図 1 のように名詞から感情生起の主体である『ポチ』、動詞または動名詞から信念が発生するイベント『散歩』、助動詞から信念の種類『des(願望)』を取り出す。そして、Jason の記法で

信念 des(pochi,sanpo)[degOfCert(0.5)]
（ポチは散歩が度合い 0.5 ほど望ましい）

を生成し、我々のシステムに入力として与え、エージェント『ポチ』の信念に追加する。信念の第一引数はイベントの主体となるエージェント名、第二引数はイベント名、degOfCert() は信念の度合い (0 ～ 1) を表す。この時、従来研究 [今井 18][浅井 18] で述べられている所定の感情生起の条件を満たしていれば、感情が生起する。

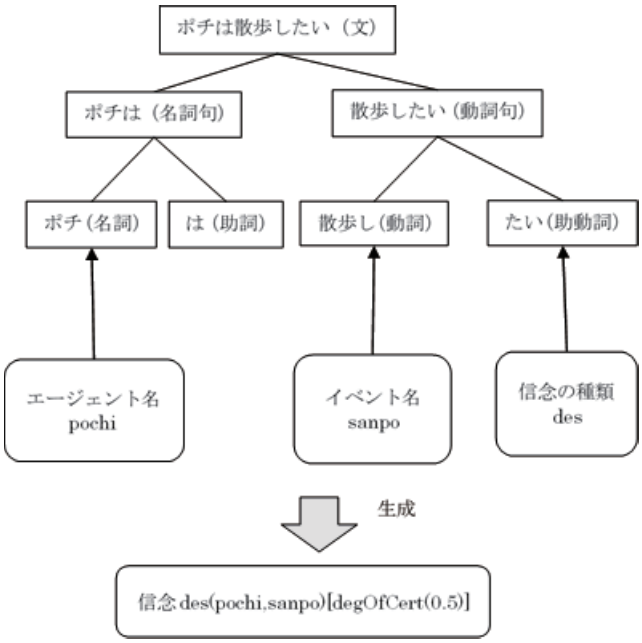


図 1: 構文木および生成される信念

感情を表す信念は、

『感情 (イベントの主体となる人物, イベント) [度合い]』
または
『感情 (感情を向ける対象)[度合い]』
という形で表現しており、相手に向ける感情を表す際に後者の形式をとる。感情の生起条件は、

- 主体となる人物
- イベント
- 信念の種類: bel(事象の成立を信じる)、des(事象の成立を望む)、prob(事象の成立を予期する)、effort(事象の成立のために努力する)、deserve(事象の成立が相応しい)、praise(事象の成立は賞賛される)、appealing(相手に魅力を感じる)、unappealing(相手に魅力を感じない)、familiar(相手に親しみを感じる) 等
- 信念の度合い

などを用いた論理式で表現している。

与えられた文から信念を抽出する際、「～ を探す」 といったイベントの場合、自身が探し物をしていることを認識していると共に、探し物を見つけることを望んでいると考えられる。そのため、エージェントは探していることを信じる信念 bel(pochi,search) を持つと共に、「見つける」ことを望む信念 des(pochi,find)[degOfCert(D)] を持つ必要がある。これにより、探し物を見つけた際に感情 *Joy*（喜び）を生起させることができる。「～ を頼む」という動詞についても同様に、相手に頼み事をしていることを認識していると共に、相手が頼み事を達成する事を望んでいると考えられる。そのため、頼むことを信じる信念の他に、他者が ～ に入るイベントを発生させることを望む信念を持つ必要がある。例えば『お使いを頼む』であれば、相手がお使いをすることを望む信念 des(pochi,bel(mike,otsukai))[degOfCert(D)] が生成される。これにより、相手が頼み事を達成した際に感情 *Joy* を生起させることができる。このように同じ動詞であってもイベント構成の一部となるもの（～ で遊ぶ)、別のイベントに対する信念を呼び出すもの（～ を探す)、信念となる感情を生成するもの（～ を頼む）が存在する。助動詞についても、信念となる感情生成に使用するため、種類や順番を考慮して正しい信念を生成する必要がある。

これらのことから、単語一つ一つに必要な要素を持たせ、構文解析中に必要であれば追加、削除を行い、適切な形の信念に変換する必要があった。これらの実現のために、本研究では形態素解析・構文解析プログラムとして既存のものを採用せず、信念の抽出に適した構文解析プログラムを新たに作成した。ただし、現段階では、物語文を解析に適した形に編集してから入力することで信念を抽出する。

信念生成の際、イベントを主語と動詞句のみからなる文から生成することも可能であり、例えば『ポチは遊んだ』という文からは、遊んだことを信じる信念 bel(pochi,play) が生成される。しかし、より正確な感情を生成するのであれば、動詞のみでなく主体になり得ない名詞 (場所、物、など) をイベント名に組み込むことで、何に対してその感情を生起したのかを細かく表現する必要がある。そうすることで、『ポチは公園で遊んだ』と言う文から感情 *Joy*（喜び）が生起した際、「『ポチ』が『公園で遊んだ』事に対して『喜び』と言う感情を生起した」ことが分かるようになることが期待される。このように、今後はイベントの細分化についても実装の改良を目指す。

5. 実験

本章では、本研究で新たに実装した他者の感情の推測の妥当性と、構文解析プログラムが期待通りの信念を抽出するかどうかの検証を行う。生起した感情の妥当性を調べる手段として、本来は人間を対象とした調査が望ましいが、大人数を必要とする心理実験は難しいため、現段階では文学作品を入力として用い、生成される感情が人間の読み取ったものと一致するかを見ることで検証を行う。シナリオは『はじめてのおつかい』[筒井 77] から一部抜粋して使用した。シナリオは以下の通りである。

『お母さんからお使いを頼まれ家を出たみいちゃんは、坂で転んでしまい、その拍子に持っていたお金を落とした。落ちたお金を無事見つけたみいちゃんは元気に坂を登って行った。無事にお使いを済ませたみいちゃんをお母さんが迎えにきた。』

このシナリオを、現段階の構文解析プログラムで信念抽出可能な形に変えて入力する。

本実験で、生成されることが期待される、みいちゃんの信念は以下のようなものである。それぞれ、上段に Jason での表記、下段にその内容を示す。

bel(haha,des(michan,otsukai)[degOfCert(0.5)])
(みいちゃんは「お母さんがお使いすることを望んでいる」ことを信じる)
bel(michan,korobu)
（みいちゃんは転んだことを信じる）
des(michan,not korobu)[degOfCert(0.5)]
（みいちゃんは転ばないことを望む）
bel(michan,otosu)
（みいちゃんは落し物をしたことを信じる)
des(michan,not otosu)[degOfCert(0.5)]
（みいちゃんは落し物をすることを望まない）
bel(michan,sadasu)
（みいちゃんはお金を探した）
des(michan,find)[degOfCert(0.5)]
（みいちゃんはお金を見つけたい）
bel(michan,find)
(みいちゃんはお金を見つけた)
bel(haha,bel(michan,otsukai))
(みいちゃんは「お母さんがお使いできたことを信じている」ことを信じる)

また、これらの信念からは、お金を落としたことに対する感情 *Distress*(嘆き)、落としたお金を見つけることができたことに対する感情 *Joy*(喜び) の生起、お使いを済ませた事に対する母の *Joy* の感情推測が起こることが期待される。

本実験では、入力としては物語を以下の様に編集して与える。

- みいちゃんはお母さんにお使いを頼まれた
- みいちゃんは転んでしまった
- みいちゃんはお金を落としてしまった
- みいちゃんはお金を探した
- みいちゃんはお金を見つけた
- みいちゃんはお母さんにお使いを伝えた

このように、現段階では、主語が明確な形の文を入力することで信念を抽出する。

実験の結果を以下に示す。まず、入力の各文からは以下の信念が抽出された。

みいちゃんはお母さんにお使いを頼まれた
→ bel(haha,des(michan,otsukai)[degOfCert(0.5)])
みいちゃんは転んでしまった
→ bel(michan,korobu)
des(michan,not korobu)[degOfCert(0.5)]
みいちゃんはお金を落としてしまった
→ bel(michan,otosu)
des(michan,not otosu)[degOfCert(0.5)]
みいちゃんはお金を探した
→ bel(michan,sagasu)
des(michan,find)[degOfCert(0.5)]
みいちゃんはお金を見つけた
→ bel(michan,find)
みいちゃんはお母さんにお使いを伝えた
→ bel(haha,bel(michan,otsukai))

このように、みいちゃんが望んでいなかったと考えられる「転ぶ」、「お金を落とす」というイベントの発生を望まない信念を、「お金を探す」というイベントに対して「お金を見つける」というイベントの発生を望む信念を抽出できるなど、期待通りの信念が抽出されている。また、それらの信念が入力された結果、生起した、あるいは他者に対して推測した感情は以下の様になった。

イベント：転ぶ
→ *Distress*(嘆き) を生起（度合い 0.5）
イベント：お金を落とした
→ *Distress*(嘆き) を生起（度合い 0.5）
イベント：探していたお金を見つけた
→ *Joy*(喜び) を生起（度合い 0.5）
イベント：お使いを済ませた
→母の *Joy*（喜び）を推測（度合い 0.5）

このように、期待された感情が生起している他、母がイベント「お使い」に対して、感情 *Joy* の生起条件である信念 des、bel を持っていることを信じたとき、母の感情 *Joy* を推測しており、自身の感情生起と同一の条件に基づいて他者の推測を行えている。

6. まとめ

本章では、今後の課題について述べる。

本研究では、他者の感情の推測の実装を行った。また、外部から与えられた文から、感情を生起する人物、イベントの主体となる人物、イベント、信念となる感情を抽出し、感情生成、推測の部分に与える信念として構成する構文解析プログラムの実装を行った。

構文解析プログラムによる信念の抽出については、長期的には、人間による入力文の調整を必要としない方向を目指すことが必要であるが、それ以外の課題を述べる。まず、現段階では『ポチは ～ した』といった、主語の存在する文でのみの信念生成となっているが、今後、より人間らしい感情を生起するために、会話を取り入れた感情生起を目指す必要があると考えられる。そのため、主語なしの文において主語を補った信念の生成を行うよう実装を改良することが課題である。また、信念の中でも、イベント発生のために努力する、イベントが発生することが賞賛される、などの文の中からの抽出が困難である特定の信念については、エージェント自身が初期信念として所持することで対処しているが、初期信念として所持するだけでなく、イベント発生や環境の変化により、新たに生成することも可能であるべきであると考えられる。しかし、現段階ではこれらの信念の生成は行われていないため、今後はそれらの信念を文脈から判断、抽出するよう実装を改良することも課題である。これらに加え、4 章で述べた通り、より明確な感情生起を行うために細分化した事象の抽出を行うことも目指す。

参考文献

[Adam 09] Carole Adam and Andreas Herzig, Dominique Longin: A logical formalization of OCC theory of emotions, Synthese, Vol.168, No.2, pp.201-248 (2009).

[OCC 88] A. Ortony, G. L. Clore, A. Collins: The Congnitive Structure of Emotions, Cambridge University Press (1988).

[BDI 99] Anand S. Rao, Munindar P. Singh, Michael P. Georgeff: Formal Methods in DAI: Logic-Based Representation and Reasoning, Massachusetts Institute of Technology (1999).

[Jason 07] Rafael H. Bordini, Jomi Fred Hübner, Michael Wooldridge: Programming MultiAgent Systems in AgentSpeak using Jason, John Wiley & Sons (2007).

[今井 18] 今井那緒: OCC theory に基づく感情表現と時間経過に関する論理モデル, 2017 年度修士論文, 奈良女子大学大学院人間文化研究科情報科学専攻 (2018).

[浅井 18] 浅井沙良: エージェントの感情生起について－新たな感情の追加－, 2017 年度卒業論文, 奈良女子大学生活環境学部情報衣環境学科生活情報通信科学コース (2018).

[堀宮 12] 堀宮ありさ, 坂野遼平, 佐藤晴彦, 小山聡, 栗原正仁, 沼澤政信: Twitter における発話者へのリプライを用いたユーザ感情推定手法, Proceedings of DEIM 2012 (2012).

[遠藤 06] 遠藤大介, 齋藤真実, 山本和英: 係り受け関係を利用した感情生起表現の抽出, 言語処理学会第 12 回年次大会発表論文集, pp.947-950, (2006).

[Tang 14] Tang Ba Nhat, 目良和也 黒澤義明 竹澤寿幸: 音声に含まれる感情を考慮した自然言語対話システム, Proceedings of Human-Agent Interaction Symposium 2014 (2014).

[石川 18] 石川葉子, 水上雅博, 吉野幸一郎, Sakti Sakriani, 鈴木優, 中村哲: 感情表現を用いた説得対話システム, 人工知能学会論文誌, Vol. 33, No. 1, pp. DSH-B_1-9 (2018).

[筒井 77] 筒井依子 (作), 林明子 (絵): はじめてのおつかい, こどものとも傑作集, 福音館書店 (1977).

Organized Session | Organized Session | [OS] OS-22

[3L3-OS-22a] 創作者と人工知能が創る創作の未来(1)

上野 未貴（豊橋技術科学大学）、森 直樹（大阪府立大学）、はたなか たいち（(株)クリエイターズインパック）

Thu. Jun 6, 2019 1:50 PM - 3:30 PM Room L (203+204 Small meeting rooms)

[3L3-OS-22a-01] (Invited talk) The Future of Animation Industry Drawn by Creators and Artificial Intelligence

〇Taichi Hatanaka[1] （1. Creators in Pack Inc.）

1:50 PM - 2:10 PM

[3L3-OS-22a-02] A Proposal of a Mathematical Story Generation Method Based on Hero’ s Journey

〇noguchi katsuhiro[1] （1. comic artist）

2:10 PM - 2:30 PM

[3L3-OS-22a-03] Narrative Structure Analysis Punchlines of SF Genre within the Flash Fiction of Shinichi Hoshi

〇Shuuhei Toyosawa[1], Hajime Murai[1] （1. Future University Hakodate）

2:30 PM - 2:50 PM

[3L3-OS-22a-04] Story Creation System based on Sentence Similarity for Supporting Contents Creation

〇Haruka Takahashi[1], Miki Ueno[1], Hitoshi Isahara[1] （1. Toyohashi University of Technology）

2:50 PM - 3:10 PM

[3L3-OS-22a-05] Classification method for four-scene comics based on Creative Viewpoint

〇Naoki Mori[1] （1. Osaka Prefecture University）

3:10 PM - 3:30 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

創作者と人工知能が描くアニメ業界の未来

The Future of Animation Industry Drawn by Creators and Artificial Intelligence

はたなか たいち
Hatanaka Taichi

クリエイターズインパック
Creators in Pack Inc.

Recently, the researches of artificial intelligence(AI) have been developed remarkably using the deep learning. In those fields, the anime is one of the most interesting and difficult research target. On the other hand, from the viewpoint of the animation industry, AI is strongly expected to help and improve the situation of creating animation. In this lecture, I show the current problems of the animation industry and the possibility of artificial intelligence in animation studio with looking to the future of the animation.

1. はじめに

近年のアニメブームにより, 日本国内だけでも年間 100 本以上のアニメが制作されている. それにより, アニメのクオリティーは進歩を遂げており, 1 本あたりにかかる制作費や制作時間も上昇傾向にある. 一方で, アニメを作る際の出資額がそれに伴っておらず, 制作会社の金銭的負担が増加している. そのため, 制作会社が経営破綻を起こすことや, アニメに携わる人間の低賃金化が問題になっている.

一方, 近年, ディープラーニングの飛躍的な発展を背景として, 人工知能 (Artificial Intelligence: AI) が社会を変える基盤技術として注目を集めている. しかしながら, アニメや漫画など高度な創作現場への AI の適用はそれほど容易ではない. これは, 一つにはやはり AI が最先端の現場で活躍する創作者に比べると『創作』という観点ではまだ及ばないこと, 使いこなせば優れたツールなのだが初期学習コストが大きいことなどが挙げられる.

そこで, 本発表では人工知能をどのようにアニメ業界に導入していくのか, その際に問題となることはどのようなことなのか, また, 創作と人工知能という視点からアニメ業界を俯瞰し, アニメの未来について述べる予定である.

2. 日本におけるアニメ業界の現状

近年, TV アニメは人気を増しており, 年間 100 本以上のタイトルのアニメが「せいさく」されている. アニメには, 出資を集めるための「製作」と, 実際の作業をする「制作」の 2 種類「せいさく」が混在している. 出資を集めるための製作を行う組織のことを製作委員会と呼ばれている. アニメに求められるクオリティーは年々増しており, 制作にかかるコストも増している. しかし, 製作委員会の出資額にはほとんど変化がなく, 制作現場は低賃金でのアニメ制作を余儀なくされている. それにより, アニメ現場の低賃金問題は社会問題となっている. しかし, アニメ現場にも問題は多くある. アニメの現場では, 古くから続く, アナログの手法にとらわれている部分が多く, 効率化が行い切れていない. アニメの制作工程には大きく分けて, 絵コンテ, 原画, 演出, 作画監督, 動画, 仕上げ, 撮影という工程が存在する. この工程の中で, アニメーターが担当する, 絵コンテ, 原画, 演出, 作画監督, 動画という工程では紙を使ったアナログ作業が主流となっている. アナログでの作業の場合, 完成した素材のやり取りなどを郵送, もしくは制作進行が車で運ぶ, などの作業が発生し, コストのロスが多く発生している. それぞれの工程を効率化していくために, 様々な見地が必要と思われるが, 著作権などの問題もあり, まだまだ課題が多いのが現状である.

連絡先: はたなか たいち, クリエイターズインパック, t.hatanaka@creatorsinpack.com

3. アニメ制作のデジタル化

アニメ制作の多くは, アナログの紙に対する作業によって行われいるが, 近年, ツールやソフトの発達により, デジタルでのアニメ制作が進んでいる. 液晶タブレットや, イラストソフトの多機能化により, アナログ作業に近づいた感覚での作業を実現化し始めている. しかし, デジタル化にはいくつか問題がある. まずは, アニメーターがアナログ作業に慣れてしまっていることである. アニメーターの多くは, アナログで絵を描くということになれてしまっており, デジタルに対しての抵抗感が大きい. それ以外にもコストの問題がある. ツールとソフトの両方を用意するためには, 最低でも 1 セットあたり 20 万円ほどのコストがかかってしまうため, 制作現場への負担が大きい. 上記で記載した以外の問題も多くあるが, もちろんデジタル化にはメリットも存在する. アナログでかかっていた際の, 輸送やスキャンなどでかかっていた人的コストを減らすことが出来たり, 描き直しが用意になることで, アニメーターが描く絵の精度が上がったりなどのメリットがある. 他にも, 近年のアニメでは 3DCG と手書きアニメの融合なども多くなっており, デジタルでの作業によって, 3DCG との親和性が上がり, より緻密なアニメ作りを行うことが出来る. そのため, デジタル化には多くの問題があるが, その問題を解決しながら, 導入を進める現場が増え始めている. しかし, 業界全体の普及率を考えた際には, まだ 20%ほどかと思われる.

4. アニメにおける絵コンテ

アニメにおける絵コンテとは, シナリオを元に作成するアニメの設計図である. 絵コンテにはカメラがどのように動くか, どのキャラクターがどのような動きをするか, どのような会話が発生するか, などの画面を構成する情報が多く入っている. そのため, アニメの制作作業に入る前に絵コンテを解析し, 物量を事前に把握することで, より効率的に作業を進めることが

出来ると考えられる．また，絵コンテの解析を進めることで，V コンテと呼ばれる絵コンテから生成される映像の自動生成を最終的な目的とする．V コンテはアニメに音をつけるアフレコ工程で使用する重要な中間生成物でありながらその生成には多大な手間がかかる．この部分を自動化できればアニメ制作に掛かる費用の大幅な削減が可能となる．現在，このアニメにおける絵コンテ部分が人工知能技術の応用が一番望まれている．

5. まとめ

本稿では，創作者と人工知能が描くアニメ業界の未来についてアニメ業界の現状を踏まえながら示した．人工知能の発展が，アニメ業界をどのように変えていくのかを予想するのは極めて困難であるが，少なくとも多くの人工知能技術がアニメ業界に導入され，アニメの創作に大きなパラダイムシフトが訪れることは確実であるように思われる．アニメの創作者，人工知能研究者，そしてコンテンツを楽しむユーザすべてにとって実りある未来が訪れることを願ってやまない．

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

ヒーローズ・ジャーニーに基づいた関数による物語生成手法の提案

A Proposal of a Mathematical Story Generation Method Based on Hero's Journey

野口　克洋[*1]
Katsuhiro Noguchi

中矢　誠[*2]
Makoto Nakaya

[*1] 漫画家
Comic Artist

[*2] 株式会社アキュトラス
Aqutras Inc.

We propose a story generation method that is able to run on machine. It is based on Monomyth known as Hero's Journey by mythologist Joseph Campbell. We reveal event effect and strength at specific timing on the story mathematically by trigonometric function. We confirmed indication of a part of archetype on Monomyth and considered examination of effectiveness of our method.

1. はじめに

コンピュータコンテンツの成立以後，エンターテイメント分野ではインタラクティブコンテンツの需要が増大し続けており，長年，物語の機械的生成が試みられてきた．古来より物語は，エンターテイメント分野以外にも，マーケティングなどの実用分野においても用いられており，物語を個人や時々に合わせて動的に生成する技術の確立は，広い応用領域をもつ研究分野と言える．物語の機械的生成においては，小説等を学習データとした機械学習による生成手法や，類型化したイベントを条件分岐などで制御していく手法などが知られている．しかし，物語は長大化しやすい性質と形態の多様性を持ち，データセットや出力が離散的である．そのため，既存手法では全体の構成を踏まえての実用レベルでの生成が困難である場合が少なくない．

先行研究として福田らは，Multi-Agent System による物語生成を提案している[1]．物語上のイベント発生をきっかけとして仮想的なフィールドとランダムウォークを用いることで，イベント発生のタイミングを決定している．これにより，情報密度の偏りを担保し，読者がある程度の恣意性を感じうる物語の生成に成功している．しかし，この手法では，イベントが起きるタイミングが，物語全体の構成からみて適切であるかや，起きたイベントの影響が適切な強度であるかが考慮されていない．動的な物語の機械生成の実現には，物語中でのイベントが，物語の構成に対してどう作用しているのかや，その強度を数理的に扱う手法が必要となる．その解決策として，物語創作に広く用いられているジョーゼフ・キャンベルのヒーローズ・ジャーニーに着目した．

2. 物語の構造

神話学者ジョーゼフ・キャンベルは，世界中の多くの民話や神話に，共通の構造があることを発見し，単一神話論と名付けた[2]．この構造をジョーゼフ・キャンベルは図 1 のように示している．単一神話論は，主人公が日常から何らかの非日常に遷移し，再び日常へ帰還する通過儀礼の構造をもつ．通過儀礼の構造は，長い物語の中で挿話などの形でも現れ，結果として単一神話論の物語構造は，一種のフラクタル構造をなす場合がある．また，日常と非日常は，物理的空間の移動のみならず，同一地点での状況変化や心理状態の変化として表現されることもあり，実際の物語では様々な概念の重合わせ構造を持つことが多い．

連絡先：野口克洋，漫画家，uo_kanazuhi@hotmail.com

これは，映画スターウォーズの脚本の元となり，映画の成功を経て，ヒーローズ・ジャーニーとして広まった[3]．今日では，ハリウッド映画の脚本創作論など，多くの創作者にとっての物語創作の基礎となっている．単一神話論に頻出のイベントやキャラクタについては，神話論や創作論，心理学の研究者達による類型化が行われている．

3. 提案手法

本研究では，単一神話論に基づく物語創作を，計算機上で再現する手法を提案する．単一神話論は，図 1 で示したとおり，円環状の図をもって，周期運動として説明されるため，物語中でのイベントの作用や強度を，三角関数を用いて数理的に扱った．実際の物語構造では，多層的に構造が重なりあっている[4]が，本論では，単一神話論に基づき，1 つの主要な構造を物語として扱う．

3.1 主人公の遷移関数と状況関数

物語上での主人公の位置を，遷移関数$P(t)$とし，式 1 で表す．tは，物語上の時間軸上のある点を表す．Pの値によって，単一神話論における日常と非日常を式 2 で定義する．物語上の時間 t でおきているイベントによって，主人公が受ける影響の強さと方向はPを微分して得られるP'で求まる(式 3)．遷移と状況の様子を図 2 に示す．

$$P = -\cos(t) \quad (1)$$

$$P(t) = \begin{cases} \text{日常}, & P < 0 \\ \text{日常と非日常の境界線上}, & P = 0 \\ \text{非日常}, & P > 0 \end{cases} \quad (2)$$

$$P' = \sin(t) \quad (3)$$

3.2 提案手法の有効性の検討

ジョーゼフ・キャンベルの作図が数理的に正しいかどうかは証明が困難であるため，提案手法においても，単一神話論における物語の類型や特徴が表出するかどうかを確認し，有効性を検討した．

通常，物語創作者は，執筆中の物語のあるシーンにおいて，その前後の流れをある程度把握した上で物語を創作している．このため，物語中のイベント上でアドリブなどにより，物語に乱れが生じた場合でも，前後の流れから大きく逸れていないと創作

者が考える場合には，その乱れを許容しうる．逆に，物語の流れから大きく逸脱せず，状況が許容範囲内にあるとき，創作者は物語上の大きなイベントを必要としないとも言える．

提案手法に基づき，状況P'をt軸方向に$\pm\sigma$だけ平行移動した値の差を許容値P'_{allow}とする．許容範囲を図 3 で図示した．主人公のP'が許容値から逸脱した場合，物語上のイベントを発生させてP'を補正し，主人公が遷移関数をなぞるよう実装する．このとき，$P = 0$のとき，つまり日常と非日常の境界線上において，P'_{allow}は最小値になる．P，P'とP'_{allow}の関係を図 4 に示す．この結果，$P = 0$を目前として表出するイベントは，発生頻度が高く，主人公に与える影響の少ないものとなる．これは，単一神話論におけるゲートキーパーというキャラクタ類型の特徴に符合する．

4. まとめと今後の課題

本研究では，神話学者ジョーゼフ・キャンベルのヒーローズ・ジャーニーとして知られる単一神話論に基づく物語創作を，計算機上で再現する手法を提案した．物語中での任意の地点におけるイベントの作用や強度を，三角関数を用いて数理的に求めた．単一神話論における物語の類型の一部について表出を確認し，提案手法の有効性の検討を行った．

現時点では，提案手法と単一神話論に致命的な齟齬は見つかっていないが，ゲートキーパー以外の類型が表出する条件を検証できていない．他のキャラクタやイベントの類型が，必ずしも数学的性質を持っているかは定かではないものの，今後の研究ではキャンベルら神話学者や物語の捜索にかかわる研究を深く理解し，その背景にある数学的性質をより明らかにしていくことを目指す．また，物語コンテンツから任意のベクトル量の変化を時系列順に抽出することができる場合，提案手法の応用で，物語の状況分析や構造解析が定量的に行える発展可能性があると考える．そうした発展的研究と共に，今後は文章などの形で物語生成が行える実装と，手法としての性能評価実験の実施を目指す．

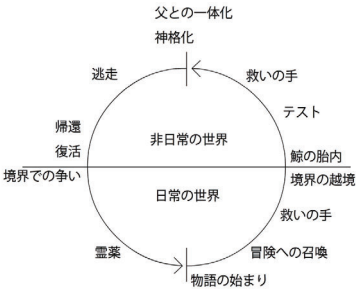


図 1　ヒーローズジャーニーの概念図

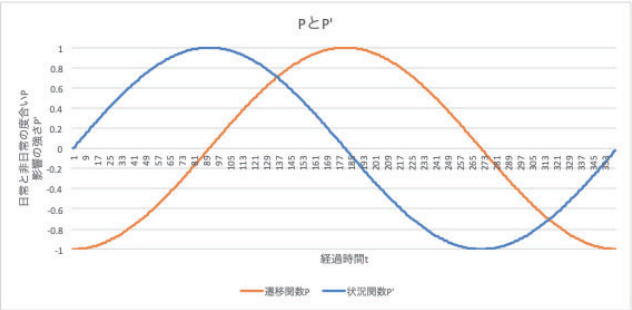


図 2　主人公の遷移関数と状況関数

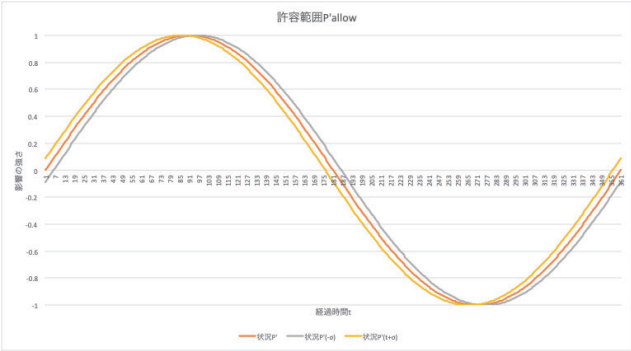


図 3　状況P'の許容範囲

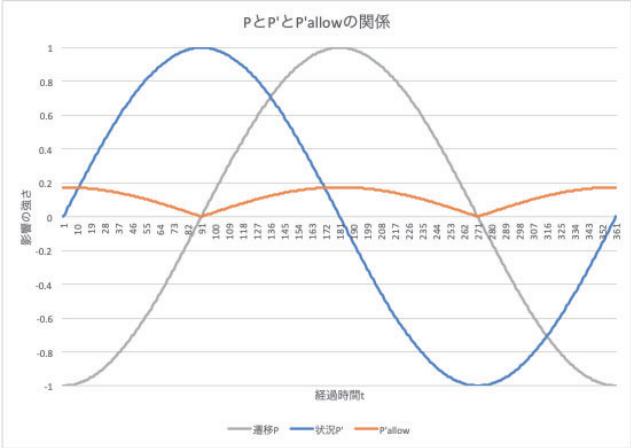


図 4　．P，P'とP'_{allow}の関係

参考文献

[1] 福田 清人，森 直樹，松本 啓之亮： 既存小説に依存しない創発的なストーリーの自動生成に関する考察，第 29 回人工知 能学会全国大会，3O1-6in

[2] ジョーゼフ・キャンベル，倉田真木，斎藤静代，関根光宏： 千の顔をもつ英雄，ハヤカワ・ノンフィクション文庫，2015.

[3] ジョーゼフ・キャンベル，ビル・モイヤーズ，飛田 茂雄: 神話の力，ハヤカワ・ノンフィクション文庫，2010.

[4] Yu.ロトマン，磯谷孝：文学と文化記号論，岩波現代選書，1979.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

星新一のショートショートにおける SF ジャンルのオチ構造分析

Narrative Structure Analysis Punchlines of SF Genre within the Flash Fiction of Shinichi Hoshi

豊澤修平[*1]
Shuuhei Toyosawa

村井源[*1]
Murai Hajime

[*1] 公立はこだて未来大学 システム情報科学部
Faculty of Systems Information Science, Future University Hakodate

Abstract: This paper aims to analyze the narrative structure leading to punchlines of genre SF within the flash fiction of Shinichi Hoshi, based on classical plot analysis method. This is a process to create plots, and then to write texts for automatic generation of stories. There is a feasibility to generate automatically more natural stories by analyzing narrative structure leading to punchlines. This paper focuses on the categories of "space", "medicine", "automatic device, invention, computer" and "robot" stories from science fiction genre, which is a representative genre of Shinichi. By utilizing the result of unification and abstraction of the patterns of the theme and conditions / precondition, it would be possible to generate automatically a more natural story like Hoshi.

1. はじめに

近年，人工知能分野では物語の自動生成が一つの目標として挙げられており，例えば松原仁らが星新一賞の一次選考を通過したことなどが注目を浴びた[松原 2013]. この「きまぐれ人工知能プロジェクト 作家ですのよ」では星新一のショートショートのようなユーモアのきいたオチのある短い物語を自動生成する人工知能の作成を目標としている. この目標を達成するためには物語の基本となるプロットの作成を行い，そのプロットをもとに文章化をする必要がある. しかし，意味のつながる文章の自動生成をするだけではなく，星らしい人間の無知や努力を皮肉る世界観を表現する必要もある. このためには星新一のショートショートのプロットの作成段階で独特の世界観を表すオチが生成されなければならない.

以上の問題を解決するためには，物語論の方法を用いた計量的な分析が有効であると考える. 文学作品を計量的に分析することで，主観性の強い評論に対して、客観的な知見を得る試みが行われてきている. 例えば、工藤彰らは村上春樹の作品について統計的分析をもとに作風の変化を把握することに成功した[工藤 2011].

また，文章の自動生成については松山諒平らによるテーマに基づいた短編の自動生成[松山 2013]や，豊澤らによる推理小説プロットを自動生成し映像化する統合的インタラクティブシステムの開発と評価[2018 豊澤 A]などが挙げられる.

これらの自動生成を行う上では良質なプロットやオチのデータを用意する必要があるが，星新一の多くの作品では短い物語の中に逆転のオチを付けることで現代まで読者を惹きつけており，自動生成の基礎データとして活用可能な良質で分析容易な作品群と言える. また，一人の作家が 1000 を超える作品を書き上げているため物語のプロットやオチの抽出を計量的に行いやすいという点からも科学的な分析に適していると考えられる. 星新一のショートショートに関しては計量的な物語構造の分析[佐藤 2009]，落語の類型である 4 分類による分類分析[永田 1993]，星新一の物語の基本パターンの抽出[村井 2011]がすでになされている.

しかし，各ジャンルにおける固有の詳細な物語パターンの抽出とオチに至るメカニズムは現在のところ分析がされていない. しかし，物語の精度良い自動生成を実現するためにはより詳細で網羅的な物語パターンの分析・収集は不可欠である.

本研究では古典的なプロット分析[プロップ 1987]を用い，ショートショート SF ジャンルにおけるシンプルな物語のオチの構造のパターン化，及び必要条件と前提条件の発見を目的とする.

基本的な物語のミクロパターンを収集し，それらの種類や物語要素の交換可能条件などを特定することで，より自然な物語の自動生成が可能になると期待される. また，これらのオチのパターンを結合することまで可能になればショートショートのような短編だけではなく，長編のような物語の自動生成も視野に入ると考えられる.

2. 対象

分析に用いる星新一のショートショート作品には，新潮社の星新一全集収録を用いる[星 1998]. 1000 編以上存在するが大きく三部に分けられる[村井 2011]. 第一期は多くの読者が想定する星新一らしい逆転オチの作品群であるが，第二期は多重オチの込み入った構造になっており，第三期は民話を志向した別方向の物語となっている.

星新一の作品において特徴的なジャンルとして SF が挙げられる. 星の SF 作品は「宇宙」や「薬」，「ロボット」，「自動装置・発明・コンピュータ」の 4 種類のテーマに大別できる. また，SF ジャンルとは別に代表的なテーマとして「悪魔・魔人」も存在する. 加えて計量的な物語構造の分析[村井 2011]において自動分類された作品中で最も話数が多い「強盗・泥棒・詐欺・囚人」も挙げられるであろう.

本研究ではテーマの自動分類の結果を手動で修正・拡張したものを利用し，上記の 4 つのテーマの条件の元で分類された SF ジャンルの 304 作品（表 1）が分析対象である.

表 1 対象テーマと作品数

テーマ	作品数
宇宙	124
薬	55
自動装置・発明・コンピュータ	76
ロボット	49

連絡先：豊澤修平，公立はこだて未来大学システム情報科学部，b1015124@fun.ac.jp.

3. 手法

主要 SF4 テーマにおける各作品にプロット分析の手法を用いて作品からプロットを抜き出した．その後，一度プロットを類義語でグループ化し抽象化を行った．その上で物語のオチにつながる要素を選定した．この際に現れるオチの構造に着目し，類似する作品が 3 作品以上あるものを物語パターンとして抽出した．また，明確なパターンとして現れないが部分的に類似する箇所が複数存在する場合もパターンとして抽出を行った．

加えて，物語パターン抽出に加えてオチに関する必要条件と前提条件を抽出した．本論文における必要条件とはオチを成立させているテキスト中に記述されていない暗黙的な常識的な知識である．また，前提条件はオチを成立させている動作の主体が行わなければならない行為や，作中における世界の常識である．必要条件と前提条件を特定することによって，それらの条件を満たす別の要素とオチを交換可能となる条件を明示化することを目的としている．このようなオチの必要条件を特定してデータとして記述しておくことによって，本研究の成果に合わせて一般的なオチのデータを別に準備することで異なるオチの物語の生成が可能になると考えられる．

4. 結果

図 1～図 14 中において，パターン化した際の要素を左から順に物語における時系列順に並べ，縦にそのパターンにおける分岐の可能性を並列で表している．また，灰色のボックスが要素内の変更可能部分であり，それ以外の要素は固定化された要素になっている．

「宇宙」124 作品のうち 100 作品は以下の 6 パターンに分類された．

- 遭遇，59 個（図 1）
- 贈り物，17 個（図 2）
- 沈黙・遭難，8 個（図 3）
- 間抜けオチ，7 個（図 4）
- 惑星，7 個（図 5）
- カニバリズム，4 個（図 6）

ただし個数は 1 つの作品に対して複数のパターンが存在する場合含まれるため，実際の作品数より多くなる．

遭遇パターン（図 1）は人間と宇宙人が遭遇することでオチが発生するパターンである．テーマが「宇宙」の中で最も多くの作品が該当する．遭遇パターンでは人間視点と宇宙人視点に分類することが可能である．人間視点では宇宙人と遭遇するタイミングでオチが変化する．可変事項の前に遭遇する場合では非視覚的事実が発覚し，後の場合では視覚的なものがオチにつながる．宇宙人視点であった場合では人間以外（例えば貧乏神や幽霊などと）と遭遇することでオチが変化する．どちらの場合でも宇宙人の常識や認知的なずれがオチとなることが多い．

贈り物パターン（図 2）では贈与物の内容でオチが変化する．贈与物が兵器などの悪いものであった場合は多くが与えられたものが起動するのがオチとなる．また贈与物がテクノロジーなどの良いものであったとしても，何らかの理由で獲得できなくなる，破壊するなどしてしまう．まれに贈与物が良いものであるときには獲得することもある．

沈黙・遭難パターン（図 3）は大まかには遭難者視点と救出者視点の 2 つに分けることができる．また，遭難者視点においてオチの一部として惑星に不時着できずに死亡のオチになるものが入る．

間抜けオチパターン（図 4）は古典落語におけるオチの一つであり，深刻そうな出来事が起きるが深刻ではないことで終わるのがオチとなるパターンである．このパターンにおいては深刻そうな出来事という曖昧な要素が引き金である．

惑星パターン（図 5）では惑星自体がオチにつながっているものである．また，亜種パターンとして都市自体がオチにつながるものも部分的に類似する作品としてパターンに含めている．

カニバリズムパターン（図 6）では初期状態には具体的に決まった状態が存在せず，オチにおいては人間を肉として認識または捕食する暗示で終わる．宇宙物語では宇宙人は人間とは認識が異なるため人肉食をタブーとしない存在として描かれる．

「薬」55 作品うち 38 作品は以下の 2 パターンに分類された．

- 薬効果明示，30 個（図 7）
- 薬効果不明，12 個（図 8）

薬効果明示パターン（図 7）は多くが薬を使用した結果によりオチが発生する．多くは薬の効果が発揮されることで別の事実が発覚し，それがオチにつながる．または期待していた薬の効果とは別の新たな効果が発揮されることでオチにつながる場合に分けられる．加えて薬が誤用される場合にはそれが原因で混乱が拡大されてオチにつながる．

薬効果不明パターン（図 8）は薬の効果が不明で効果が発揮されることでオチにつながるパターンである．効果が明確なパターンと異なり初期条件として強盗が挙げられるのが特徴的である．薬の効果がわからない状態にすることで効果の発覚と効果による状態の悪化がオチにつながっている[豊澤 2018B]．

「自動装置・発明・コンピュータ」76 作品のうち 66 作品が 3 パターンに分類された．

- 装置効果明示，33 個（図 9）
- 装置効果不明，14 個（図 10）
- 自動装置，11 個（図 11）
- コンピュータ，7 個（図 12）

装置効果明示パターン（図 9）は装置を利用した結果によりオチが発生する．多くは装置を利用したために失敗するか，別の効果や目的が発覚することがオチにつながる．

装置効果不明パターン（図 10）は効果不明で効果が発覚することか，発覚した効果によりオチが発生する．

自動装置パターンでは自動装置は生活に欠かせないものとして存在する．従って装置が停止することによる不幸か快適すぎるために退屈になるオチになる．

コンピュータパターン（図 12）では，コンピュータはミスをすることがなく支配または管理しているものとして存在する．そして人，間側が支配管理されていることに対して肯定的であればその管理された世界で幸福になるという．

「ロボット」49 作品のうち 41 作品が 3 パターンに分類された．

- 正体ロボット，14 個（図 13）
- ロボットの効果，12 個（図 14）
- 他テーマのパターン，22 個

「ロボット」においては他テーマのパターンが多くみられた．これは「自動装置・発明・コンピュータ」と「宇宙」における既存のパターンが適用できたためである．主体としての役割も併せ持つ装置としてロボットが存在していると思われる．また，ロボットの効果というのはロボットにおける効果や能力といったものを暗黙

的に利用したパターンである．ショートショートにおけるロボットとは命や感情が存在せず仕事をする存在として描かれており，オチとしてこれらの前提を崩したものがパターンとし現れた．

また「ロボット」には，他のテーマのパターンが適用可能な作品がいくつか見られた．

また，各テーマのパターン適用の必要条件と前提条件を 304 作品から同様に抽出した．すべての作品に対して複数個の条件が発見されたため，本論文では「ロボット」における必要条件と前提条件の結果の一部のみを表[1，2]に示す．

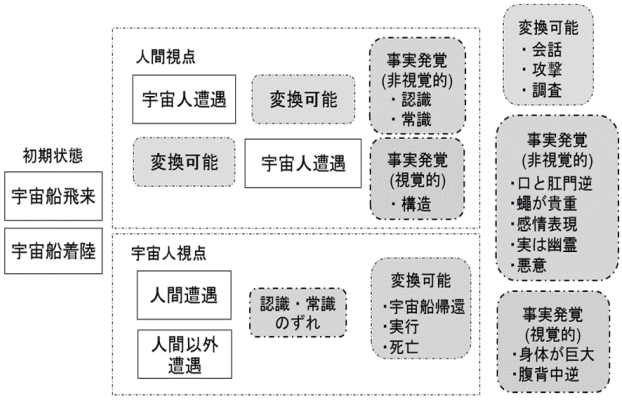


図 1 遭遇パターン

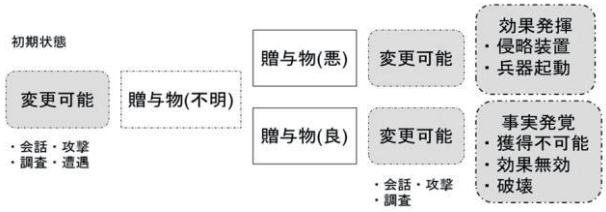


図 2 贈り物パターン

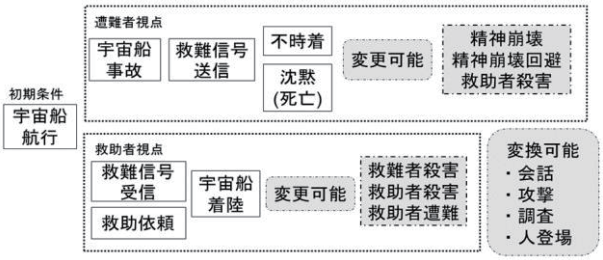


図 3 沈黙・遭難パターン

図 4 間抜けオチパターン

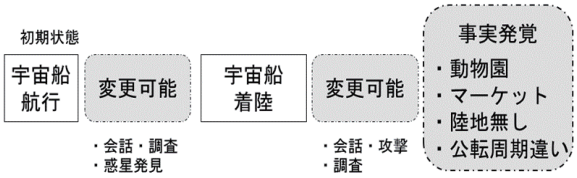


図 5 惑星パターン

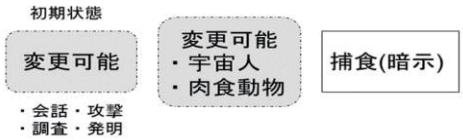


図 6 カニバリズムパターン

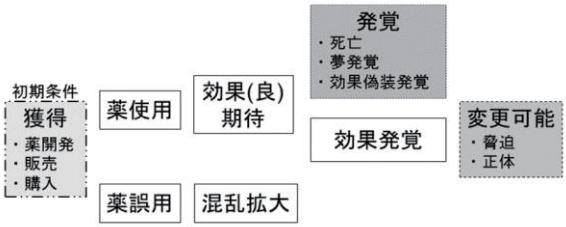


図 7 薬効果明示パターン

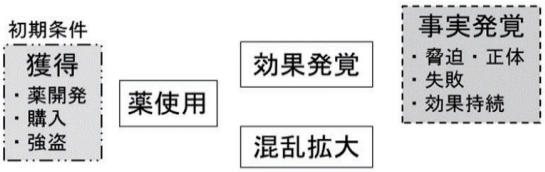


図 7 薬効果不明パターン

図 9 装置効果明示パターン

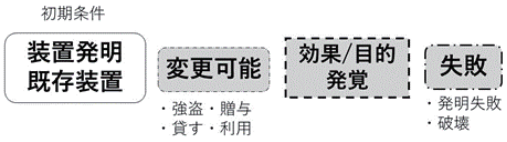


図 10 装置効果不明パターン

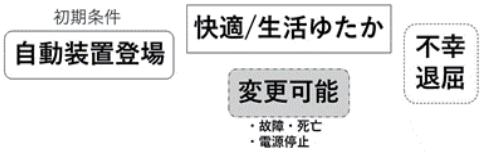


図 11 自動装置パターン

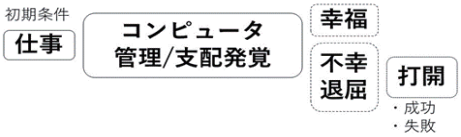


図 12 コンピュータパターン

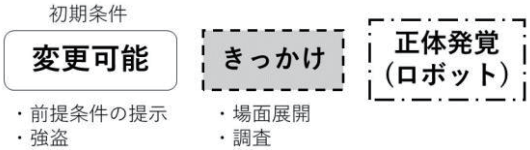


図 13 正体ロボットパターン

図 14 ロボット効果

表 2「ロボット」における必要条件

必要条件
ロボットは命がない
ロボットは痛覚がない
ロボットは複製可能
ロボットは生物ではない
ロボットは合理的

表 3「ロボット」における前提条件

前提条件
人と見分けがつかないロボット
社会に普及しているロボット
頑丈なロボット
多機能ロボット
隔離された場所

5. 考察

各テーマにおいて特定のパターンを抽出することが可能であった. 加えてパターンにおいて変更可能な箇所が存在し, 変更可能な箇所への要素を代入可能であると考えらえる. これらの各要素の組み合わせによって固定のオチやパターンの分岐が生じることが示唆される.

「宇宙」の「遭遇」,「薬」の「薬効果明示」と「薬効果不明」パターン,「装置」の「装置効果明示」と「装置効果不明」パターン,「ロボット」の「正体ロボット」パターンには共通性があると考えられる. これらは基本的に効果が判明しているか不明かの 2 つのパターンに分類することができる. 効果が明示の場合では他の効果が判明するか, 判明している効果のせいで不幸になるかに大別できる. また, 効果が不明である場合では効果がわかることか, わかったことによる結果で不幸につながっている. いずれのパターンにおいてもオチは人間の無力さ無知さによって努力が無駄になることを皮肉る星新一らしい形になっている.

また, あるテーマの主要なパターンが他のテーマの作品のパターンとしても利用されるものがいくつか見られた. これは世界観としてテーマが利用されているが, オチにはそのテーマが関与しない場合であると考えられる. 加えて,「宇宙」における「贈り物」,「間抜けオチ」,「カニバリズム」のような宇宙という世界観を利用しなくても利用できるようなパターンについては, その世界観を取り入れることで起こりうるオチの条件を変えるための舞台設定として利用されていると考えられる. これらに加えて「ロボット」にあるようなものは上位概念として特殊な装置があり, それの要素として主体性のある装置としてロボットという条件が加えられたために他のテーマのパターンと共通性が高いと考えられる.

6. おわりに

星新一のショートショートにおける 4 つ SF ジャンル 304 作品についてオチの構造の分析を行った.「宇宙」は 124 作品中 100 作品を 6 種類にパターン化可能であった.「薬」は 55 作品うち 38 作品を 2 種類のパターンに分類することができた.「自動装置・発明・コンピュータ」76 作品のうち 66 作品を 3 種類のパターンに分類することができた.

また今回分析対象テーマを SF ジャンルに限定した. しかし物語の自動生成においては汎用的な物語パターンが必要であるため, 今回分析を行った以外のジャンル分析も行う必要がある. 従って, 作品数が最も多い「強盗・泥棒・詐欺・囚人」の分析も行う必要があると考える.

本研究で抽出した必要条件と前提条件は自然言語で書かれておりまだ統一的な記述にはなっていない. 今後これらにオチのパターンと同様に用語の統一と抽象化を行う必要がある. 加えて, 各テーマにおける条件を他のテーマパターンに適用可能かまでデータ化することができればより網羅的で多種多様な物語プロットの自動生成が可能になると考えられる.

本論文では物語の抽象化及びオチのパターン化などを単独分析者で行っており客観性が保証されていない. 従って, 統計的な手法にのっとり第三者によるパターン化の一致度の検証をしていきたいと考える.

これらの作品分析とパターン化及び必要条件と前提条件の抽出ができれば, より星新一らしい物語プロットの自動生成が可能になると考えられる.

参考文献

[工藤 2011] 工藤彰, 村井源, 徃住彰文:”計量分析による村上春樹長篇の関係性と歴史的変遷”, 情報知識学会誌, Vol21, No1, pp18-36, 2011.

[佐藤 2009] 佐藤知恵, 村井源, 徃住彰文: “文学作品群の特徴的語彙と概念カテゴリーの抽出-星新一のショートショートの計量分析-”, 情報知識学会誌, Vol.19, No.2, 2009.

[松原 2013] 松原仁, 佐藤理史, 赤石美奈, 角薫, 迎山和司, 中島秀之, 瀬名秀明, 村井源, 大塚裕子: “コンピュータに星新一のようなショートショートを創作させる試み”, The 27th Annul Conference of Japanese Society for Artificial Intelligence, 2013, 2D1-1.

[松山 2018] 松山諒平, 佐藤理史, 松崎拓也: ”テーマに基づく短編小説自動生成システム“, 言語処理学会 第 24 回年次大会 発表論文集 (2018 年 3 月)

[村井 2011] 村井 源, 松本 斉子, 佐藤 知恵, 徃住 彰文: "物語構造の計量分析に向けて-星新一のショートショートの物語構造の特徴-", 情報知識学会誌, Vol.21, No.1, pp.6-17, 2011.

[豊澤 2018A] 豊澤修平, 工藤はるか, 石田晃大遠, 藤史央里, 川瀬稜人, 菊池亮太, 工藤健太郎, 栗原将風, 櫻井健太郎, 佐藤好高, 玉置秀基, 根本裕基, 原科充快, 久野露羽, 平田郁織, 村井源, 椿本弥生, 角薫, 松原 仁”推理小説プロットを自動生成し映像化する統合的インタラクティブシステムの開発と評価”, 研究報告人文科学とコンピュータ, Vol.13, pp1-5, 2018-01-20.

[豊澤 2018B] 豊澤修平, 村井源:"星新一のショートショートのオチの構造分析", 情報処理学会シンポジウムシリーズ, Vol.2018. No.1, pp. 75-82. 2018.

[長田 1993] 長田貴誉子:"星新一の作品分類と分析試論 -一〇〇一編ショート・ショート群を考える", 国文学会誌, Vol.22, pp. 76-115,1993.

[星 1998] 星新一:「星新一ショートショート 1001」, 新潮社, 1998.

[プロップ 1987] ウラジーミル・プロップ(北岡誠司,福田 美智代訳):「昔話の形態学」, 水声社, 1987.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

創作支援を目的とした類似度調節による ストーリー生成システムの構築

Story Creation System based on Sentence Similarity for Supporting Contents Creation

高橋 遼 *1 Haruka Takahashi　上野 未貴 Miki Ueno　井佐原 均 Isahara Hitoshi

*1豊橋技術科学大学
Toyohashi University of Technology

In recent years, there are lots of researches to generate creations such as comics and novels by computing methods. Especially, the orders of sentences is very important for the quality of stories. Thus, the aim of the research is to construct story creation system in order to consider the flow of the story based on the sentence similarity.

1. はじめに

近年，計算機の性能の向上によるニューラルネットワークを用いた学習が盛んである．その一つとして，漫画や小説，アニメーションといった人間の創作物を計算機によって理解したり生成する研究が広がりつつある [1]．これらの研究から，人間が創作活動を行う際の補助を行うことも可能であると考えられる．

創作者はテキスト，画像や音声などの情報を工夫して組合せてストーリーを表現する [2]．ストーリーは創作物の根幹であり，カテゴリ，評価や独自性に大きく影響する．そのため，ストーリーを作成する過程を計算機上で処理することは挑戦的な課題である．

本研究では，読者が違和感を感じにくいストーリー構成の分析や，文の順序を変更して新しい作品の生成へ応用できるか調べるため，物語のストーリー自動生成システムを構築する．

2. 先行研究

2.1 ストーリー創作支援

ストーリー創作支援の従来研究としては，質問集合とグラフに基づく物語全体の流れを管理可能な創作支援システムの提案 [3] がある．この研究では，創作支援システムを作成し，創作支援システムへの入力から創作過程の解析と創作支援をする．あらかじめ準備されたカテゴリから創作者が今回考案するストーリーに適したものを選び，フォームを埋める形式でストーリープロットを作成することができる．しかし，ストーリーの生成は創作者によるもので，計算機が自動生成はしていない．

2.2 文自動生成

文の自動生成の従来の研究としては，単語の分散表現を用いた意味予測に基づく雑談応答生成 [4] がある．この研究では Recurrent Neural Network(RNN) を用いて, Twitter 上の会話文を学習させ，ある会話文の入力に対して，適切な応答文の作成を行っている．従来の One-hot 表現を用いた方法と違い，単語の分散表現を入力としている．従来の方法よりも多様性の面では向上したが，適切性に関しては大きく向上しなかった．また，文法的な誤りがある文も生成された．

この研究は，1 文に対して 1 文を返す対話文を対象としたものであり，ストーリーといった複数文の出力には対応していない．

3. ストーリー生成システム

本研究では分散表現を利用し，文間の類似度に基づくストーリー生成システムを構築する．

3.1 コーパスの作成

コーパスには，3 つのデータを用いる．用いたデータを以下に示す．

1. Manga109 の 4 コマ漫画のセリフデータ（ 5145 文）
2. 青空文庫のあらすじ（ 7439 文）
3. インターネット上の小説投稿サイトから集めた文章（ 383494 文）

本稿ではこの中でも最もデータ数が多く，結果が良かった小説投稿サイトの文章を利用する.「インターネット上の小説投稿サイト」のデータはインターネットからクローリングして収集した．これらを句点で分割し，1 文区切りにしたものを文集合として利用する．

3.2 Paragraph Veector のモデル作成

Paragraph Vector[5] のモデルを作成する. Paragraph Vector とは文を分散表現に変換する技術である. 分散表現に変換することで, コサイン類似度などを用いて, 文同士の類似度を計算できるようになる.

Paragraph Vector の学習に使用する文から単語を抽出する. 文に単純な分かち書きしたものを学習に使用すると, 文の意味への影響が小さい助詞や助動詞, 記号等も学習に使用される. そこで, 形態素解析をし, 文の意味に関連しにくいと考えられる品詞の単語を除去する. 形態素解析には MeCab を用いた. また, MeCab の辞書には NAIST Japanese Dictionary*1 を用いる. 今回学習に使用する品詞は名詞, 動詞, 形容詞, 副詞, 連体詞である.

連絡先: 高橋 遼，豊橋技術科学大学，〒 441-8580　愛知県豊橋市天伯町雲雀ケ丘 1 − 1，takahashi@lang.cs.tut.ac.jp

*1 NAIST Japanese Dictionary, http://naist-jdic.osdn.jp/

表 1: Doc2Vec のパラメータ

手法	次元サイズ	単語の最小登場回数	エポック数
DM-PV	100	1	100

DM-PV: Distributed Memory Paragraph Vector[6]

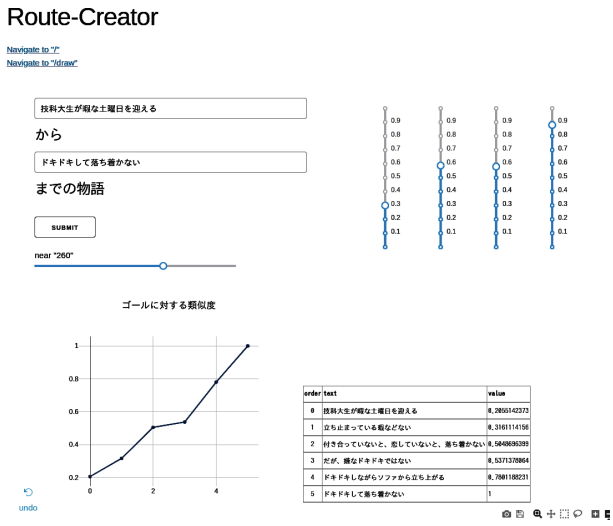


図 1: ストーリー生成システムのユーザインタフェース

Paragraph Vector を用いるためには Python 製のソフトウェアである Gensim*2 を用いる．Gensim の中の Paragraph Vector を用いることができるライブラリである Doc2Vec を用いる．表 1 に Doc2Vec のパラメータを示す．

3.3 ルート作成

上述の通り作成したモデルを用いて，ストーリーを作成する．システムの入力はストーリーの「スタートの文 r_s」と「ゴールの文 r_g」であり，スタートからゴールに対して，類似度が高くなるようにストーリーの遷移経路を以下の手順で作成する．

1) 空のキュー $\mathbf{R}_s$ を用意する．ゴールの文に最も近い位置にモデル内で存在する文を仮のゴールの文 r'_g とする．$\mathbf{R}_s$ に r_s を追加する．
2) $\mathbf{R}_s$ の末尾の文に対し近傍 n の文を取得し，r'_g との類似度の高い順にソートした配列 $\mathbf{A}_s = \{a_1, a_2, ..., a_n\}$ を作成する．
3) $\mathbf{R}_s$ に既存でないという条件を満たす，$\mathbf{A}_s$ のインデックス番号が最も小さい文 a_i を $\mathbf{R}_s$ に追加する．
4) $\mathbf{A}_s$ に r'_g が存在すれば，r'_g と r_g を $\mathbf{R}_s$ に追加しルート作成は終了する．終了しない場合，2) に戻って処理を繰り返す．

作成された $\mathbf{R}_s$ がストーリーとなる．

3.4 ユーザインタフェースの実装

本研究は創作支援を目的としており，将来的にユーザが使用することを想定し，WEB アプリケーションにてストーリー生成システムの実装を行った．今回，開発に使用した言語は Python であり，用いた WEB アプリケーション開発用ライブラリは Dash by plotly *3 である．図 1 に作成したシステムのユーザインタフェースを示す．図 1 の左上が入力するためのテキストボックスである．右上の類似度調節機能でストーリー内の類似度の遷移を決定する．作成されたストーリーは右下に表示される．左下にはストーリーの類似度の遷移の軌跡がグラフとして表示される．

*2 Gensim, `https://radimrehurek.com/gensim/`
*3 Dash by plotly, `https://plot.ly/products/dash/`

表 2: 生成されたストーリーの例 1

	文	類似度
0	技科大生が暇な土曜日を迎える	0.206
1	立ち止まっている暇などない	0.316
2	付き合っていないと、恋していないと、落ち着かない	0.505
3	だが、嫌なドキドキではない	0.537
4	ドキドキしながらソファから立ち上がる	0.780
5	ドキドキして落ち着かない	1.00

表 3: 実験 1 のアンケート結果

入力	近傍	文数	自然に感じた	ほとんど自然に感じた	やや不自然に感じた	不自然に感じた
1	50	9	0	1	2	1
	100	7	0	1	2	1
	300	7	0	2	2	0
2	50	29	0	0	1	3
	100	9	0	1	2	1
	300	11	0	0	4	0
3	50	7	0	0	3	1
	100	6	1	1	2	0
	300	6	0	0	2	2
4	50	7	1	2	1	0
	100	9	1	2	1	0
	300	4	2	1	1	0

3.5 生成結果

表 2 に生成されたストーリーの例を示す．

生成結果は，文数をスタートとゴールの文を含めた 6 文に固定し，類似度を図 1 の右上のスライダーのように調節した．

4. 実験 1

4.1 アンケート評価

本システムの評価を行うためにまず，アンケートをした．文数を固定せず，類似度の調節無しとした．スタートからゴールに到達するまでに要した遷移回数が文数となる．スタートとゴールのペアが異なる 4 パターンの入力に対し，それぞれ近傍数を 3 種類： 50, 100, 300 に変化させ，ストーリーを生成した．そして，それらのストーリーに対し， 4 人のユーザに 4 段階：自然に感じた，ほとんど自然に感じた，やや不自然に感じた，不自然に感じたのいずれかを回答して評価をしてもらった．表 3 にアンケートの結果として各入力に対する回答数を示す．

4.2 ランダムとの比較評価

次に，ランダム生成したストーリーとシステムを用いて作成したストーリー対 8 パターンを 4 人のユーザに比較をしてもらった．方法は，各入力パターンに対し，スタートの文からゴールの文までをつなぐ 5 文と 10 文をそれぞれシステムを用いる方法とコーパスから無作為に抽出する方法で決定し，スタートの文とゴールの文と合わせてストーリーとする．生成されたストーリーを比較し，どちらがより自然に見えるストーリーかをユーザに選んでもらう．比較は，文数が同じもの同士とした．4.1 のアンケート評価により，最も不自然と感じるユーザが少なかった近傍数を 300 に固定した．また，図 2, 図 3 はシステムの 5 文と 10 文の場合の類似度調節機能のス

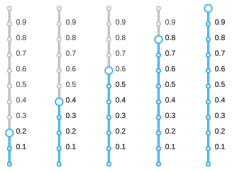


図 2: 類似度調節 5 文

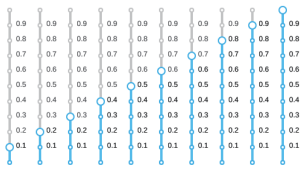


図 3: 類似度調節 10 文

表 4: ランダムとの比較結果

パターン	5 文		10 文	
	システム	ランダム	システム	ランダム
1	3	1	4	0
2	3	1	3	1
3	4	0	4	0
4	4	0	3	1
5	3	1	4	0
6	3	1	4	0
7	4	0	4	0
8	2	2	4	0

ライダーの設定である．類似度調節機能のスライダーはストーリーに意図的なアクセントが発生しないようにした．

表 4 に比較結果を示す．表 4 内の各パターンに対する行の数字は， 4 人中何人が選択したかを示している．

4.3 考察

アンケートによる評価では，システムによって出力されるストーリーでは取得する近傍の文数を大きくすることで，不自然と感じる人が減少することがわかった．

ランダムとの比較では，5 文と 10 文のどちらの場合においても，システムは優位性を示すことができた．しかし，ランダムなストーリー生成であっても 5 文からなるストーリーでは自然と感じられる場合が見られた．一方で, 10 文の場合では 4 人のユーザがほとんどの入力パターンにおいてシステムの方が自然であると回答した．本システムでは, 文の類似度が近い文がつながるため, 文の数が増加しても全体の名詞種数が多くなりにくく, システムを用いることで，5 文以上の文数では自然なストーリーが生成されやすくなると考えられる．

ストーリーが自然と感じられなかった原因としては，文の数や長文の影響が考えられる．文の数や長文が含まれると，ストーリー全体に含まれる固有名詞の数が増える．これが，ストーリー内での場面や時代，世界観に矛盾を生み出す原因となる．対策としては，固有名詞等を代名詞に変換することや，文数を少なく固定して生成すること，文長を制限し，長文の登場を防ぐことで改善できると考えられる．また，コーパスサイズにより取得される類似文の類似度の平均が大きく変わることが考えられるため，コーパスサイズに応じた適切な近傍の値を用意することで改善が期待できる．

また，文によっては同じ意味の文が複数出現し，それらがモデル内に多数存在し密集することで，類似文があまりに似通った文しか取れないという問題がある．これは，ストーリーの遷移ができず，同じような文ばかりが連続する原因となる．対策としては，極度な類似文はモデル作成時に数を減らすことや近傍 n を一時的に大きくすることが考えられる．

5. コーパスの前処理の変更

実験 1 では，分散表現より類似文を取得する際に，長い文は取得した文との類似度が高くなく，実際に読んでも類似しているとは言い難い文も取得されていた．

そこで，実験 1 で用いた文の処理方法とは異なる方法で文を処理し，コーパスを作成した．表 5 に変更点を示す．3.1 で準備したインターネット上の小説投稿サイトの作品は文体が自由に書かれており，全角記号の出現数が多い．また，文末と思われる点でも句点が用いられていない場合があり，その場合空白や改行を文の終端として用いている場合がみられた．以前のコーパスでは，鉤括弧の前後をつなげて一文としていた．しかし，鉤括弧内の文を独立して抜き出すために，分割することにした．

表 5: 前処理の変更点

対象	変更前	変更後
空白	削除	文の終端として，そこで文章を分割
鉤括弧（「」）	読点に置換	文の終端として，そこで文章を分割
全角記号	処理なし	感嘆符（！），疑問符（？），二重鉤括弧（『』）以外を削除

学習データ量は，前処理の変更により，383494 文から 663066 文に増加した．

6. 実験 2

実験 1 の結果から，文の長さが長いと違和感を覚えやすいという結果があり，これらを改善することで，生成されるストーリーの自然さに改善がみられないかを 4.1 のアンケート評価で比較実験を行う．ランダムとの比較は，4.2 の時点で既に優位性がみられたため取り組まない．

6.1 実験方法

5. で変更したコーパスを用いたストーリー生成の実験をする．3.5 と同様のパラメータで，前処理変更後のモデルに対してもストーリーを生成させた．また，実験 1 と同様に文数を固定せず，類似度の調節無しで行った結果を用いる．入力文も同様のものを用いる．

6.2 結果と考察

表 6 に生成されたストーリーの例を示す．表 7 にアンケートの結果を示す．

前処理を変更したコーパスを用いた実験では,「やや不自然に感じた」,「不自然に感じた」と答える人がわずかに減少した．これは実験 1 で述べた長文が自然さに影響するという点において，コーパスの文がより細かく分割されたことにより，長文の割合が減ったため改善されたと考えられる．また， 1 文が短くなったことにより， Paragraph Vector によって，文の分散表現を獲得し，類似度の計算を行う際に，文の長さが極端に違うものが減少しているため，より正確に類似度を測ることができるようになったと考えられる．

しかし，表 7 の入力 4 に関しては，表 3 の結果よりも変更後のコーパスの方が「自然に感じた」と回答したユーザが少なくなった．表 3 の入力 4 の結果では文が長いものも含まれていたが，固有名詞が少なく，また「俺」や「彼女」といった代名詞の使用が多い．これをユーザーがうまく補完して，話のつながりを自然と捉えられたと考えられる．対して，表 7 の結果では，文数が増えたこともしくは類似度の計算が改善して類似文を集めやすくなったことの影響からか，同じ表現を何度も繰り返すようなストーリーとなってしまった．類似文上位 n 件を取得する現在の方法では，新しいコーパスを用いると 4.3 で述べた同じ意味の文が複数出現する問題に関しては悪化させる可能性がわかった．

表 6: 生成されたストーリーの例 2

	文	類似度
0	技科大生が暇な土曜日を迎える	-0.013
1	立ち止まっている暇などない	0.338
2	落ち着かないか、[キャラクター名]	0.541
3	心の中が落ち着かない	0.568
4	なんだろう、落ち着かない	0.618
5	ドキドキして落ち着かない	1.00

表 7: 実験 2 のアンケート結果

入力	近傍	文数	自然に感じた	ほとんど自然に感じた	やや不自然に感じた	不自然に感じた
1	50	7	0	2	2	0
	100	6	0	3	1	0
	300	5	0	0	3	1
2	50	10	0	0	2	2
	100	10	0	3	1	0
	300	7	2	1	1	0
3	50	10	0	1	1	2
	100	6	0	2	2	0
	300	6	1	1	2	0
4	50	10	0	0	2	2
	100	10	0	0	2	2
	300	5	0	2	1	1

上記のことを踏まえ，今後は付近の文との距離や数に影響されにくくするために，従来のストーリーが場面に応じてどの程度のベクトル変化量があるかを解析し，文を選択していくよう改良する必要があると考えられる．

表 2 と表 6 に示されるように，前後の文に同様の単語が出現する．これは 4.3 に問題点として挙げた同じ意味の文が複数出現するという問題にも近いが，用いている Paragraph Vector が単語に影響を受けやすいことが影響している．「～ない」といった反意的な表現は学習時に直接的には考慮されないため，肯定文と否定文であっても出現する単語が似ている場合に連続する文として選択される場合がある．「～する」と「～しない」といった単語はモデル学習時には距離が遠くなるような学習を行う必要があると考えられる．そこで，反意を重視して特徴量の調整をしたり，例えば編集距離といった他の手法と組み合わせて文の類似度を算出することが有効ではないかと考えられる．

7. 創作支援への展望

本研究では，異なる 2 つの入力文に対して，それらを繋ぐストーリーを作成した．表 2，表 6 の「暇」と「落ち着かない」のようにほとんど対極とも取れる状態を結び，その間を補完するストーリーが作成できる．また，この入力が人間の目から見て明らかに繋がりにかける文であったとしても，少々強引になるかもしれないが，類似度調節を利用して段階的に結ぶことが可能である．よって，これまでの人の創作者の発想では思いつかなかった作品が生まれるきっかけになり得ると考えられる．

また，創作者の過去の十分な量の複数作品を学習することで，作者の過去に用いた文を参考に新たなストーリーを生み出すことが可能である．小説から小説でなくとも，小説作品のデータから，歌詞や短編の作品といった違う分野に活用することも考えられる．

また本研究を応用することで，作品のあらすじとして複数文を入力して文間の類似度の変化を得て，ストーリーのスタートからゴールまでの変化をパターンに分類し解析することも視野に入れている．

従って，本研究は創作支援に貢献できると考えられる．

8. まとめ

本研究では，創作支援を目的としてストーリー生成システムを構築した．文を類似度に基づいて順序変更することで人が違和感を覚えないストーリーを生成できる可能性を示した．また，以下のことがわかった．

- 文数が増えたり長文が含まれると，文中に含まれる固有名詞等の出現回数が増え，ストーリーの自然さが損なわれる
- 文中の固有名詞を代名詞に変換する等の対策を取り，改善する必要がある
- 記号の除去や文を分割する箇所を変更することで生成されるストーリーの結果が改善する

特にストーリーという観点で同じ単語を含まないが同様の意味の文は距離が近く，文に登場する単語は似ていても意味が反する文は遠くに置かれるモデルの作成が望まれる．また，コーパスに含まれる類似文の数がモデル内の距離の変化量に大きく影響するため一定距離での変化が可能なように改良が必要である．

謝辞

本研究は，一部 JSPS 科研費 (グラント番号:JP17K17809) および JST, ACT-I(グランド番号:JPMJPR17U4) の支援による．

参考文献

[1] 上野 未貴, ”創作者と人工知能: 共作実現に向けた創作過程とメタデータ付与 4 コマ漫画ストーリーデータセット構築”, 人工知能学会, 4Pin1-16 (2018)

[2] 福田 清人, 上野 未貴, 藤野 紗耶, 森 直樹, 松本 啓之亮, ”ストーリーの自動生成を目的としたストーリーモデルの提案”, 言語処理学会 P6-1 (2016)

[3] 葛井 健文, 上野 未貴, 井佐原 均, ”質問集合とグラフに基づく物語全体の流れを管理可能な創作支援システムの提案”, 人工知能学会, 2E2-04, (2017)

[4] 古舞 千暁, 滝口 哲也, 有木 康雄, ”単語の分散表現を用いた意味予測に基づく雑談応答生成”, 日本音響学会, ROMBUNNO.2-Q-12, (2018)

[5] Quoc Le, Tomas Mikolov, ”Distributed Representations of Sentences and Documents”, Google Inc, 1600 Amphitheatre Parkway, Mountain View, CA 94043, (2014)

[6] Andrew M. Dai, Christopher Olah, Quoc V. Le, ”Document Embedding with Paragraph Vectors”, arXiv:1507.07998, (2015)

[7] Y.Matsui, K.Ito, Y.Aramaki, A.Fujimoto, T.Ogawa, T.Yamasaki, K.Aizawa, ”Sketch-based Manga Retrieval using Manga109 Dataset”, Multimedia Tools and Applications, Springer, (2017)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

創作視点に基づく 4 コマ漫画におけるストーリー分類手法の提案

Classification method for Four-scene Comics based on Creative Viewpoint

森 直樹
Naoki Mori

大阪府立大学工学研究科
Graduate School of Engineering, Osaka Prefecture University

Can the computer create an interesting story? This question is one of the ultimate purpose of artificial intelligence. The comic analysis has become an attractive research topic in artificial intelligence fields as comic engineering to tackle that problem, t. Although there are lots of attractive and useful topics and many researchers have been reported, creating a story by computer is still difficult. One of the most critical reason is that there are no useful comic datasets for artificial intelligence fields. In this study, I focus on the classification method for four-scene comics based on creative viewpoint.

1. はじめに

計算機は面白いストーリーを創作することが可能なのか? この問いに答えることは非常に難しい．Yes であれば，これは人工知能における究極的な目的，人間のように思考するシステムの構築に成功したことになる．さすがに最新の人工知能でもこの問いに Yes と答えるだけの成果はまだ得られていないようにみえる．一方で，No と容易に断ずることが難しいほど，近年の人工知能の技術発展は急速であり，月並みだが,「今の段階で Yes と答えるのは難しいが近い将来確実に面白いストーリーを創作する人工知能は登場するだろう」と答えるしかないのかもしれない．人工知能によるストーリー生成については多くの研究がなされているが，本稿では，特に 4 コマ漫画に焦点を当てる．これは，4 コマ漫画はしっかりとしたストーリー構造を持ち，各ストーリーの長さが決まっているため，定量的な解析に向いているためである．

4 コマ漫画の研究は，漫画を工学的に解析しようと試みるコミック工学 [松下 13] の分野で積極的になされてきた．同分野では多様な研究可能性が示唆されており，特に人工知能研究のためのデータセット [Matsui 15] に関する研究が注目され始めている．特に，人工知能研究を対象とした世界初の独自 4 コマ漫画データセットも提案されており，今後データセットの独自構築が活発化していく可能性が高い．既存の漫画ではなく独自にデータセットを構築する場合，商用の漫画では困難である表現も可能であるという利点がある反面，自由度が高いためデータの何に着眼しているかが見えにくくなるという問題点も顕在化している．そこで本稿では，機械学習を前提とした 4 コマ漫画データセットの表現方法および着眼すべき特徴について可能な限り定量的にまとめ，同時に創作視点に基づく具体例を示すことで創作と人工知能に関する分野の発展に寄与することを目的とする．

2. 従来研究および本研究の位置付け

2.1 コミック工学に関する研究

これまでのコミック工学は，いわゆる「漫画」を対象としており，漫画を工学的に扱うためのデータセットに関する研究 [Matsui 15] や既存の画像認識手法をキャラクタ識別や吹き出し識別に代表される漫画内のパーツを識別する研究を中心に発展してきた．これらの研究は，コミック工学という分野を広げるために大きく貢献している．また，ユーザの視点に立った検索システム [山下 17] や読書意欲の増進に関する研究 [佐藤 18] についても積極的な研究がされている．

一方で，狭義の漫画に収まらない対象についても研究の展開が始まっている．例えば，厳密には漫画には分類されないアニメの絵コンテや絵本に関する研究 [Fujino 17] やストーリー生成系の研究 [Fukuda 17] などがその範疇に当てはまる．

2.2 4 コマ漫画に関する研究

4 コマ漫画は日本における代表的な漫画形式のひとつであり，4 つの コマ (齣) によって完結した短い話を表現する．人間がストーリーを感じることができる最小の表現形式として，古くから新聞や雑誌の一部に掲載されてきた．現在では，4 コマ漫画だけを扱う専門誌も数多く存在している．

4 コマ漫画の基本的なストーリー展開は，各コマを最初から順に起承転結に対応させ，結に相当する最終コマをオチとする場合が多い．それ以外にも 3 コマを序破急に対応させる場合や，2 段オチ, オチを必ずしも必要としないストーリー 4 コマなどが存在する．

4 コマ漫画に関する研究としては，4 コマにおける感情識別に関する研究 [上野 16] やストーリー理解過程の解析研究 [上野 17]，4 コマ漫画ではないが既存漫画のデータを利用した 2 コマ漫画の生成に関する研究 [迎山 18] が報告されている．また 4 コマ漫画の自動生成に関する研究 [Ueno 16a] や遺伝的アルゴリズムに基づく感性解析に 4 コマ漫画を用いた研究 [野村 17] もなされている．

また，ストーリーに関しては 4 コマ漫画の内容に踏み込んだ研究として，コマの順序識別に関する研究 [Ueno 16b, 藤野 18] が報告されている．

2.3 漫画データセットに関する研究

4 コマ漫画を対象としたデータセットとしては Manga 109[Matsui 15] が知られている．これは著作権の切れた 109 冊の漫画の画像および台詞をデータ化したもので，その有用性から高く評価されている．一方で，上野によって作られた 4 コマ漫画ストーリーデータセット [Ueno 18, 上野 18] も特徴的である．このデータセットは発展のために研究者が一から開発に関わった世界初の研究用データセットであり，著作権の問

連絡先: 森 直樹，大阪府立大学，mori@cs.osakafu-u.ac.jp

題がなく創作者自身による感情アノテートがされている，同じ内容の話が複数の創作者により作られているなど人工知能研究において極めて大きな利点を持つ．一方で，現段階では 4 コマ漫画ストーリーデータセットはデータ数が少ないという致命的な欠陥を持つ．そこで，今回はコミック工学の枠にとらわれることなく創作視点の研究が可能な 4 コマ漫画ストーリーデータセットにおいて効率的に新しいストーリーを生成するための 4 コマ漫画におけるストーリーの特徴および分類手法について示す．

3. 4 コマ漫画で着眼すべき特徴

以下に今回提案する 4 コマ漫画で着眼すべき特徴について定義し，可能であれば台詞例を示す．以下では，4 コマ漫画の各コマを始めから順に 1, 2, 3, 4 の番号で表すが，n コマ漫画の場合にも同様に表現することが可能である．

3.1 コマの可換

ある特定のコマ同士が交換可能であるとき，それらのコマは可換 (commutative) であると定義し，com:[n,m] で表す．

台詞例：
com:[1,2]
(1 コマ) 学生 A「今日こそ，あのバグのことを先生に正直に言わないと…」
(2 コマ) 学生 A「おはようございます」
(3 コマ) 教員 B「おはよう」
(4 コマ) 学生 A「すいません，あの結果はバグでした！」

ただし，同一のコマ内容であっても，他のコマとの関係で可換性を失う場合もあることに注意を要する．例えば，以下の例では 2，3 コマは可換ではない．

台詞例：
(1 コマ) 学生 A「今日こそ先生よりも先に挨拶してやる！」
(2 コマ) 学生 A「おはようございます」
(3 コマ) 教員 B「おはよう」
(4 コマ) 教員 B「…あら，なんで嬉しそうなの？」 学生 A「いえ，別に！」

3 個のコマが可換な場合もある．

台詞例：
com:[1,2,3]
(1 コマ) 学生 A「眠い」
(2 コマ) 学生 B「意識飛びそう」
(3 コマ) 学生 C「寝たい」
(4 コマ) 教員 D「…皆早く寝ろよ そうすりゃ帰れるのに！」

また，可換な部分が 2 つある場合もある．その場合は，com:[n,m][s,t] と表す．

台詞例：
com:[1,2][3,4]
(1 コマ) イケメン A「彼女と女友達は別なんだよな！」
(2 コマ) イケメン B「彼女の相談をする女友達は必要！」
(3 コマ) 非モテ C「彼女さえできたら人生終わってもいい…」
(4 コマ) 非モテ D「せめて一人だけでも女友達が欲しい…」

3.2 コマの圧縮分解

複数のコマを一つのコマにできることをコマが圧縮 (summarize) 可能である，逆に一つのコマを複数のコマにできることをコマが分解 (separate) 可能であると定義し，それぞれ sum:[n,m] および sep:[n] と表す．一般に可換なコマ同士は圧縮可能だが逆は成り立たない．例えば，以下の 2 コマは圧縮可能だが可換ではない．

台詞例：
sum:[n,$n+1$]
(n コマ) A「ゴメンなさい」
($n+1$ コマ) B「いや俺も言い過ぎた」

複数人の会話は分解可能な場合がほとんどである．ただし，4 コマ漫画の場合には事象の粒度よりも台詞の量によって複数のコマで表現せざるを得ない場合があり，意味だけでは圧縮性を判定できないことに注意を要する．

台詞例：
sep:[n]
(n コマ) A「ゴメンなさい」 B「いや俺も言い過ぎた」

以下の台詞は上述の内容と本質的には同じ内容だが，台詞の分量に関する制約で圧縮できない可能性がある．

台詞例：
sum:[n,$n+1$] (分量的に圧縮不可？)
(n コマ) A「ゴメンなさい 私の言い方も良くなかったわ 私には昔からそういう素直じゃない部分があるの 自分でもそれが嫌なのだけれど感情的になると抑えることができなくて…」
($n+1$ コマ) B「いや俺も言い過ぎた そちらの今後を考えるとどうしてもわかって欲しかったのだけれど… でもいくら正しいと信じていても言い方が少しキツかった もう少し柔らかい表現を使うべきだったよ」

3.3 同値関係

2 つの 4 コマ漫画 A, B について，何らかの要素が共通である場合，A と B はその要素に関して同値関係 (equivalence relation) であると定義し，$A \sim B$ で表す．この場合，以下の同値律を満たすものとする．また，$A \sim B$ のとき A と B は同値 (equivalence) であると呼ぶ．

反射律: $A \sim A$

対称律: $A \sim B$ ならば $B \sim A$

推移律: $A \sim B$ かつ $B \sim C$ ならば $C \sim A$

例えば,「プロットに関して同値」,「作者に関して同値」,「オチのパターンに関して同値」などがあり得る．

4. 特定のコマに関して同値なストーリー群

一般的に，既存の 4 コマ漫画の場合，作者やテーマ以外の部分で同値なストーリー群を集めることは容易ではない．一方，4 コマ漫画ストーリーデータセットのように研究者が創作に関わることが可能な場合は既存の 4 コマ漫画では困難な要素に着目し，同値なストーリー群の作成できるという利点がある．その中で，機械学習への応用を考える場合，一部コマを共通とし，残りを可変としたデータ集合，つまり特定のコマに関して同値なストーリー群が有用であると考えられる．以下では，共

通となるコマの部分を集合論における共通部分 (intersection) と考え，特定のコマに関して同値なストーリー群 S_i における共通部分を

$$\bigcap_i S_i \tag{1}$$

と表現する．このとき各ストーリーは以下で表現できる．

$$S_i = \{s_{i,1}, s_{i,2}, s_{i,3}, s_{i,4}\} \tag{2}$$

ここで $s_{i,j}$ はストーリー S_i の j 番目のコマ情報である．以下では，このストーリー群のパターンについて考察する．

4.1 分岐型

1 コマから連続する最初の部分が共通部分となる場合を分岐型と呼ぶ．例えば 1〜3 コマを共通部分として 4 コマ目のオチだけを変える，4 コマ目のポジネガだけを変えるなどの場合に用いられる．1 コマ目のみが共通部分の場合には自由度が一番高く，多くのストーリーを作りやすい．

分岐型の例

(1 コマ) **A「神様，周りのみんなだけは幸せにして下さい」**
(2 コマ) 神「それでお前は幸せなのか？」
(3 コマ) A「はい．周りが幸せなら僕も幸せです」
(4 コマ) 神「困ったなぁ… お前の不幸がみんなの幸せなのよ」

(1 コマ) **A「神様，周りのみんなだけは幸せにして下さい」**
(2 コマ) B「私，A 君のためならなんでもする」
(3 コマ) C「A 君の笑顔がない世界なんて考えられないよね」
(4 コマ) D「A 君の幸せが私たちの一番の幸せだってわかっているのかしら？」

4.2 収束型

4 コマ目から遡って連続する最後の部分が共通部分となる場合を収束型と呼ぶ．例えば 4 コマ目のオチだけ示しそこに至る過程を 1〜3 コマで自由に作る場合に用いられる．4 コマ目の内容がインパクトの強い場合は自由度が非常に低くなる可能性がある．一方，4 コマ目を日常的な内容とした場合には平坦なストーリーになりやすい．

収束型の例

(1 コマ) B「月が綺麗＝ I love you って意味なんだ！」
(2 コマ) A「B さん，ちょっと外に行きませんか？」
(3 コマ) B「A さん！ いいですよ …あら，綺麗な満月」
(4 コマ) **A「月が綺麗ですね」**

(1 コマ) A「今日は満月！距離は？輝度は？これはスゴイぞ！」
(2 コマ) B「天文部の A 君は私をどう思っているのかしら…」
(3 コマ) A「B さん，報告会に備えて月を見ませんか？」
(4 コマ) **A「月が綺麗ですね」**

4.3 砂時計型

2，3 コマ目をいずれか，または両方を共通部分とした場合を砂時計型と呼ぶ．これは一番ストーリーの作成が難しい．

2 コマ目が共通部分の場合の例

(1 コマ) A「よし，実験終わった！結果もバッチリ！」
(2 コマ) **B「ああ，僕も頑張らないと」**
(3 コマ) A「え？まだこれだけ？マジで？これ無理じゃない？」
(4 コマ) B「……それでも…頑張らないと」

(1 コマ) A「C 君の結果は素晴らしいね」
(2 コマ) **B「ああ，僕も頑張らないと」**
(3 コマ) A「B 君のテーマだって素晴らしい 焦ってはいけないよ 結果は必ず出るから」
(4 コマ) B「はい，焦らず頑張ります！」

3 コマ目が共通部分の場合の例

(1 コマ) A「X って教授にも意見するし超自信家だよな」
(2 コマ) B「あいつこの前ヤンキーと乱闘したらしいぜ」
(3 コマ) **猫「ニャーン」**
(4 コマ) X「ひゃっ！……でも猫だけは苦手なんだ」

(1 コマ) A「仕事また失敗した 貯金ももうない」
(2 コマ) A「人間関係も最悪 頼る友達も彼女もいない 孤独だ…」
(3 コマ) **猫「ニャーン」**
(4 コマ) A「なんだか元気出た！ もう少しだけ頑張ってみるか！」

2，3 コマ目が共通部分の場合の例

(1 コマ) A「研究費厳しいけれどなんとか学生を学会に行かせてやりたいな」
(2 コマ) **B「うちの研究室っていつもエコノミーと安宿だよね」**
(3 コマ) **C「うん 高級ホテルじゃなきゃ学会なんて行きたくないよな」**
(4 コマ) A「自分の分は自腹を切って学生分の旅費だけはなんとかしてやるか」

(1 コマ) A「あいつらの成果のおかげで使い切れないほどの研究費が！」
(2 コマ) **B「うちの研究室っていつもエコノミーと安宿だよね」**
(3 コマ) **C「うん 高級ホテルじゃなきゃ学会なんて行きたくないよな」**
(4 コマ) A「次の学会も俺だけ高級ホテルで旨いものでも喰うとするか！ 」

4.3.1 複雑な収束型

一般に収束型は結果が同一になる場合が多いが，出オチの場合は結果に自由度がある．しかしながら，これは本来は 4 コマ目にあるべきオチが倒置していると考えられるので，収束型でありながら結論が異なるわけではない．一方で，4 コマ目が同一であっても 1〜3 コマ目の内容によって結論のポジネガ極性が反転したりまったく別の意味になる場合がある．これは，コマ単独で内容を理解することができないことを示しており，機械学習を用いる場合は LSTM のような時系列性を扱える手法か 1〜4 コマの内容を一度に入力する手法を用いなければならないことを示している．以下に複雑な収束型の例を示す．

複雑な収束型の例

(1 コマ) A「アイツに研究報告しに行ったら嫌味言われた」
(2 コマ) B「髪型や服装にまでケチ付けるし最低な先生」
(3 コマ) C「私なんて彼氏の悪口言われたたんだけど，なんなのよ？」
(4 コマ) **D「あの先生，本当に私たちが嫌がることを進んでするよね」**

(1 コマ) A「この前，先生が一人で散らかった研究室の掃除をしてた」
(2 コマ) B「絶望的に理解力がないアイツの面倒も夜遅くまで見るし」
(3 コマ) C「私たちの飲み会とか学会に行く段取りまで自分でしてくれる」
(4 コマ) **D「あの先生，本当に私たちが嫌がることを進んでするよね」**

これは自然言語に基づく分岐の例であるが，台詞はなく画像だけでもこのような例を作ることも可能である．

複雑な収束型の例 (画像のみ)

(1 コマ) A が真夜中に研究室を動き回っている
(2 コマ) B「あれ？ 機械学習のライブラリが最新版になっている！」
(3 コマ) C「調子が悪かったコンピュータも直ってる なんで？」
(4 コマ) **A が自分の机で寝ている**

(1 コマ) B「GUI の部分終わったよ！アルゴリズム担当の A は？」
(2 コマ) C「プレゼンのパワポ作ったぜ アルゴリズム担当の A は？」
(3 コマ) D「やっとデータセットができた アルゴリズム担当の A は？」
(4 コマ) **A が自分の机で寝ている**

この違いを理解することはかなり高度な読解能力が必要であり，ストーリーの創作も容易ではないが，そこまで踏み込んだストーリー群を作ることは人工知能研究において極めて有用であり検討していく必要がある．

5. まとめ

本稿では，創作視点に基づく 4 コマ漫画におけるストーリー分類手法を提案し，記号的な表現方法と具体的な台詞例について述べた．既存漫画のアノテートと創作視点に基づくパターンに応じたデータセットの構築は，創作者と人工知能研究のために必須の内容であり，今回の内容に基づく具体的なデータセットの構築および数値実験について進める必要がある．また，創作者の方々のご意見を聞くとともに，クラウドソーシング等を利用した今回のパターン分けに基づく大規模創作実験も重要な今後の課題である．

参考文献

[Fujino 17] Fujino, S., Hatanaka, T., Mori, N., and Matsumoto, K.: The Evolutionary Deep Learning based on Deep Convolutional Neural Network for the Anime Storyboard Recognition, in *Distributed Computing and Artificial Intelligence, 14th International Conference, DCAI 2017, Porto, Portugal, 21-23 June, 2017*, pp. 278–285 (2017)

[Fukuda 17] Fukuda, K., Fujino, S., Mori, N., and Matsumoto, K.: *Semi-automatic Picture Book Generation Based on Story Model and Agent-Based Simulation*, pp. 117–132, Springer International Publishing, Cham (2017)

[Matsui 15] Matsui, Y., Ito, K., Aramaki, Y., Yamasaki, T., and Aizawa, K.: Sketch-based Manga Retrieval using Manga109 Dataset, *CoRR*, Vol. abs/1510.04389, (2015)

[Ueno 16a] Ueno, M.: Computational Interpretation of Comic Scenes, in *Advances in Intelligent Systems and Computing*, Vol. 474, pp. 387–393 (2016)

[Ueno 16b] Ueno, M., Mori, N., Suenaga, T., and Isahara, H.: Estimation of structure of four-scene comics by convolutional neural networks, in *Proceedings of the 1st International Workshop on coMics ANalysis, Processing and Understanding*, p. 9ACM (2016)

[Ueno 18] Ueno, M.: Four-Scene Comic Story Dataset for Softwares on Creative Process, in *New Trends in Intelligent Software Methodologies, Tools and Techniques - Proceedings of the 17th International Conference SoMeT_18, Granada, Spain, 26-28 September 2018*, pp. 48–56 (2018)

[迎山 18] 迎山 和司：既存の漫画に基づいた 2 コマ漫画の自律生成, 人工知能学会全国大会論文集 2018 年度人工知能学会全国大会 (第 32 回) 論文集, pp. 1K1OS2a02–1K1OS2a02 一般社団法人 人工知能学会 (2018)

[佐藤 18] 佐藤 剣太, 牧 良樹, 中村 聡史：コミックの読書意欲を増進させる要素に関する分析, 人工知能学会全国大会論文集 2018 年度人工知能学会全国大会 (第 32 回) 論文集, pp. 1K1OS2a01–1K1OS2a01 一般社団法人 人工知能学会 (2018)

[山下 17] 山下 諒, 朴 炳宣, 松下 光範：コミックの内容情報に基づいた探索的な情報アクセスの支援, 人工知能学会論文誌, Vol. 32, No. 1, pp. WII–D_1–11 (2017)

[松下 13] 松下 光範：コミック工学：マンガを計算可能にする試み, 日本知能情報ファジィ学会ファジィ システム シンポジウム 講演論文集, Vol. 29, pp. 199–199 (2013)

[上野 16] 上野 未貴, 森 直樹, 松本 啓之亮：漫画内の特徴的要素が与えるストーリーの印象についての検討, 第 30 回人工知能学会全国大会発表論文集, pp. 2J5–OS–08b–4in2 (2016)

[上野 17] 上野 未貴, 末長 寿規, 井佐 原均：漫画中の表現獲得方法に基づくストーリー理解過程の解析, 第 31 回人工知能学会全国大会発表論文集, pp. 4F1–5in2 (2017)

[上野 18] 上野 未貴：創作者と人工知能: 共作実現に向けた創作過程とメタデータ付与 4 コマ漫画ストーリーデータセット構築, 人工知能学会全国大会論文集 2018 年度人工知能学会全国大会 (第 32 回) 論文集, pp. 4Pin116–4Pin116 一般社団法人人工知能学会 (2018)

[藤野 18] 藤野 紗耶, 森 直樹, 松本 啓之亮：3 分岐型畳み込みニューラルネットワークによる 4 コマ漫画の順序識別, 人工知能学会全国大会論文集 2018 年度人工知能学会全国大会 (第 32 回) 論文集, pp. 1K1OS2a05–1K1OS2a05 一般社団法人 人工知能学会 (2018)

[野村 17] 野村 俊太, 荒井 幸代：進化計算を用いた人間の感性理解のための遺伝子解析法, 第 31 回人工知能学会全国大会発表論文集, pp. 3H2–OS–04b–2 (2017)

Organized Session | Organized Session | [OS] OS-22

[3L4-OS-22b] 創作者と人工知能が創る創作の未来(2)

上野 未貴（豊橋技術科学大学）、森 直樹（大阪府立大学）

Thu. Jun 6, 2019 3:50 PM - 5:30 PM Room L (203+204 Small meeting rooms)

[3L4-OS-22b-01] Autonomous Collection and Feedback of Dialogue Profile for Dialogue Contents Circulation

Masaki Mori[1], ○Akinobu Lee[1] （1. Nagoya Institute of Technology, Japan）

3:50 PM - 4:10 PM

[3L4-OS-22b-02] A Discussion about Jazz Improvisation System based on Studies with First-Person's View

○Daichi Ando[1] （1. Tokyo Metropolitan University）

4:10 PM - 4:30 PM

[3L4-OS-22b-03] Of application of artificial intelligence to fashion design

○Yoko Fujishima[1,3], Osamu Sakura[2,3] （1. Graduate School of Interdisciplinary Information Studies, The University of Tokyo, 2. Interfaculty Initiative in Information Studies, The University of Tokyo, 3. RIKEN-AIP）

4:30 PM - 4:50 PM

[3L4-OS-22b-04] Tone Pasting Using cGANs with Tone Feature Loss

○Koki Tsubota[1], Kiyoharu Aizawa[1] （1. The University of Tokyo）

4:50 PM - 5:10 PM

[3L4-OS-22b-05] Creative Future will be Created by Contents Creators and Artifical Intelligence

○Miki Ueno[1] （1. Toyohashi University of Technology）

5:10 PM - 5:30 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

音声対話コンテンツ循環環境を実現するための対話シナリオのプロファイル自動収集およびフィードバック

Autonomous Collection and Feedback of Dialogue Profile for Dialogue Contents Circulation

森 雅希　　李 晃伸
Masaki Mori　　Akinobu Lee

名古屋工業大学大学院工学研究科
Nagoya Institute of Tehchnology, Japan

In this research, we regard a spoken dialogue systems as “dialogue media content” from user’s view, and tries to realize dialogue media content circulation between creators and users. Focusing on the feedback from users to the creator of the content, an autonomous collection of dialogue profile and its effective feedback scheme is proposed. Experimental results showed that among the automatically collected feedback information, displaying recognition word list and occupation time at each state helps creators for content refinement and updates. Further interview with a professional creator after a long free-creation run also indicates that this sheme can motivate ones to create the content.

1. はじめに

近年，音声認識や音声合成，自然言語処理技術の発展に伴い音声対話システムが注目されており，様々な音声対話システムや音声言語インタフェースが実用化されつつある．

個々に異なる音声対話システムが社会に溢れるとき、ユーザにとって音声対話システムは対話という体験を提供する一種のコンテンツとして捉えると考えられる。映像メディアや画像メディアのように、音声対話システムもその構成要素（辞書・ボイス・対話シナリオ・エージェントモデル・モーション等）からなるコンテンツとして考えることができる。音声対話がコンテンツとして認知され，対話が消費・模倣・創造される環境が実現されれば，多様な対話システムの創造や、データドリブンで実践的な対話システムの知見の蓄積が期待される．筆者らは音声インタラクション構築ツールキット MMDAgent を用いた音声対話コンテンツの研究を行っている [1][2].

本研究では，音声対話コンテンツがメディアとして循環するための要件の一つとして，ユーザからシステム構築者（クリエイタ）へのフィードバックに着目し，利用したユーザの体験や動向をクリエイタへ積極的かつ有効にフィードバックする仕組みを提案する．以下，対話シナリオ動作時のユーザごとのプロファイルを自動収集し，それをクリエイタに有益な形でフィードバックするための仕組みを提案し，実験を行った結果を示す．

2. 音声対話コンテンツの循環系

コンテンツが流通する背景にはコンテンツ循環が存在する [3]．音声対話においては，図 1 のようなコンテンツ，ユーザ，クリエイタが主体となった循環系を成すと考えられる．コンテンツがこれら主体の間を循環していくことで，多彩かつ魅力のあるコンテンツがユーザから生み出される．

音声対話のコンテンツ循環系については，これまでに，ユーザがキャラクターとの対話を登録・編集できるコンテンツ作成システム [1] や，ユーザ層ごとのインセンティブ設計と誘引の分析 [2] などが試みられているが，実際に音声対話コンテンツのポジティブループを回すには至っていない．

連絡先: 李 晃伸，名古屋工業大学，愛知県名古屋市昭和区御器所町，052-735-7550，ri@nitech.ac.jp

図 1: 音声対話コンテンツの循環系

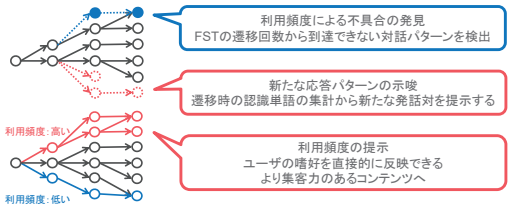


図 2: 対話シナリオ利用統計からのフィードバック

音声対話は，動画や画像と異なり，体験を提供するインタラクティブなメディアである．このため，音声対話コンテンツを一方向に提供するだけでなく，ユーザの話した内容や受け答えへの反応などの動向をクリエイタへ強くフィードバックする仕組みが重要であると考える．フィードバックによって，想定と実際が噛み合ったより質の高いコンテンツへと発展することが容易になることが期待される．また，ユーザの反応が積極的にクリエイターへフィードバックされることは，クリエイターへ満足度等のインセンティブ付与や，新たなコンテンツへのアイデアを供する効果も期待される．

3. 対話プロファイルの自動収集およびフィードバックの実現

コンテンツをユーザが利用した際の利用情報（対話プロファイル）を自動収集し，コンテンツ作成者へフィードバックするシステムを提案する．

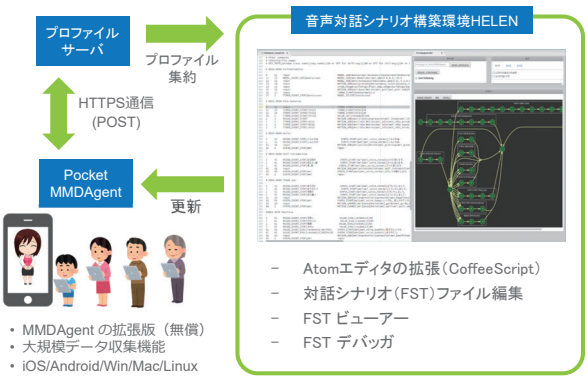


図 3: 提案システムの全体像

3.1 対話プロファイルからのフィードバック

図 2 に，対話プロファイルから得られるフィードバックの例を示す．図の左側は対話コンテンツにおける対話シナリオ（状態遷移モデル）である．ユーザが実際に辿った対話の状態遷移と各状態における発話内容，滞在時間を記録し，その情報を収集することで，状態ごとの利用率やユーザ離脱率，応答できなかった発話リストが得られる．これにより，シナリオ上で定義したが実際には利用されない対話パターンの検出や，ユーザがある状況で頻繁に発話するが対話シナリオで対応できていないパターンの抽出，ユーザの嗜好などが把握できる．

3.2 対話プロファイルの自動収集

提案システムの全体像を図 3 に示す．音声対話コンテンツの構築環境として，編集中の対話シナリオの対話状態遷移をグラフィカルに表示可能な MMDAgent 用の対話シナリオ編集環境 HELEN*1 を用いる．コンテンツ再生アプリは，ログ記録・収集機能を持つ Pocket MMDAgent[4] を用いる．Pocket MMDAgent は，コンテンツ利用時のログをサーバへ自動アップロードする．アップロードされたログからサーバ上でプロファイル情報を集計・抽出し，コンテンツ開発元のクリエイターへ提供する．HELEN では提供されたプロファイル情報から対話シナリオの状態遷移図（図 3 右部）上に利用率を色分け表示するとともに，利用率，離脱率，平均滞在時間および認識単語リストを状態ごとにポップアップ表示する．これによって，フィードバックを確認しながら編集を行える．

4. 評価実験

まずフィードバックがコンテンツ構築に与える効果について検証した．10 名の学生を対象に，既存の作成済み音声対話コンテンツ（自己紹介・10 対話）を与え，エディタを用いて自由に対話内容を追加してもらう実験を行った．まずプロファイルなしの状態で課題を行い，次に事前実験で収集したプロファイルに基づいたフィードバック情報を表示した状態で課題を行った．利用後にアンケートによる主観評価を行った結果を表 1 に示す．フィードバック情報は音声対話コンテンツの制作・更新の際に有用な情報となり，認識単語リストおよび利用率表示が有用であることが示唆された．一方，離脱率表示はあまり有用でなく，複雑なコンテンツの制作にも大きくは寄与しなかった．コンテンツの修正や簡易な追加といった小規模の修正には有用であるが「新たな対話内容を発想する」という創作に近い行為への敷居を下げる効果は小さかったといえる．

*1 近日中に GitHub にて公開予定

表 1: 実験結果

質問内容	評価（5 段階）
コンテンツ制作に役立ったか	4.0
編集が楽になったか	4.0
コンテンツ制作の敷居を下げると思うか	3.2
思い通りにコンテンツを制作できたか	4.0
今後もツールを利用してみたいか	3.7
認識単語リストは有用だったか	4.5
離脱率表示は有用だったか	2.7
利用率表示は有用だったか	3.9
コンテンツ制作を手助けするか	3.8
複雑なコンテンツを制作できたか	3.1
簡単にコンテンツを制作できたか	4.0
コンテンツ制作の励みになったか	4.1

プロファイル情報の創作行為への効果をさらに調べるため，実際にクリエイターとして活発に活動しているアマチュア創作者（対話システムは構築未経験）に本システムを利用してもらう実験を行った．システムを口頭説明し，3D モデルなど素材を提供したのち，3 週間の期限で自由にコンテンツを創作してもらい，実験後にインタビューを行った．作成は期間を前半と後半に分け，前半で作成したシステムを数人のユーザに利用してもらい，そのフィードバックを与え，後半の作成を行うようにした．結果として，クリエイターが構築したコンテンツは，早口言葉を認識するゲーム（150 状態・200 遷移）であった．インタビューの結果，利用率や認識単語リストがユーザ動向の把握に重要な手がかりとなり，コンテンツ制作における有効な指標となり，コンテンツ創作のモチベーションとなることが示唆された．一方で，開発当初に想定した内容以上のものがフィードバックによって新たに生まれることはなかった．モチベーションの維持には，今回のようなパッシブな情報だけでなく，ユーザからのコメントや要望といったアクティブなフィードバックも必要であると考えられる．

5. おわりに

音声対話コンテンツにおいて対話プロファイルの自動収集とフィードバックがコンテンツ循環の一助となることを示した．今後の展開としては，音声対話コンテンツ配信の仕組みやユーザ間の体験の共有方法，アクティブなフィードバック，大規模実験等が挙げられる．

参考文献

[1] 宮木京介, 飯塚遼, 李晃伸, “利用者による履歴付き対話の共同構築・拡張が可能なユーザ生成型音声対話システム”, 日本音響学会秋季研究発表会講演論文集，3-Q-22，2015.

[2] 飯塚遼, 李晃伸, “ユーザ生成型音声対話システムにおけるクリエイターとユーザの相互刺激によるインセンティブ向上の検討”, 研究報告 自然言語処理（NL） 2014-NL-216(6), 1-6, 2014.

[3] 関亜紀子, 亀山渉, “コンテンツ循環における権利詠唱の自動化”, 情報処理学会論文誌, Vol.48, No.5, p.1952-1964, 2007.

[4] 李晃伸, “音声対話コンテンツのネットワーク配信および大規模ログ収集を可能にするスマートフォン版 MMDAgent の開発”, 日本音響学会秋季研究発表会講演論文集，2-2-8，2019.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

2-5-1 フレーズ練習記録の一人称研究を例にした ジャズ即興演奏生成システムのための検討

A Discussion about Jazz Improvisation System based on Studies with First-Person's View

安藤大地 *1
Daichi Ando

*1首都大学東京システムデザイン学部
Faculty of System Design, Tokyo Metropolitan University

In this paper, the author shows an example of jazz improvisation acquisition process record based on the author's own first-person view research. The example is about practice and cognition of 2-5-1 harmony progresses and phrases. The record indicates that musician's cognition of internalize process can be clearly described in such a way as to rule . Then the author discusses possibilities of that it is effective to adopt musician's knowledge into the art creative machine learning system.

1. はじめに

本稿では，ジャズ即興演奏生成システムの構築を目的とし，著者自身の一人称研究的記録の中の 2-5-1 フレーズの部分を例に事例と問題点を示し，それを元にジャズ即興演奏生成システムに人間と共通の音楽的知識を取り込むことの有効性について議論する

著者は，人間と音楽知識を共有する作曲システム・即興演奏システムの実現のために，主に遺伝的プログラミング (Genetic Programming, GP) と対話型進化論的計算 (Interactive Evolutionary Computation) へ人間の音楽知識への取り込みを進めている [安藤 05, 安藤 12, 安藤 17, 安藤 18a].

近年のトレンドである DNN を用いた音楽生成では Google Magenta の例があるが，実際に Google Magenta をクラシック音楽の作曲家が用いた例では，少なくともクラシック音楽に適用した場合には十分な表現力がないという意見がある [大久保 17]. また，Cope は LISP のプログラムが，作曲家が作曲を行う過程のような音楽の高レベルでの表現に有効であると述べている [Cope 05]．OpenMusic という最も成功した作曲支援ソフトウェアも LISP である．

LISP が音楽家にとってベストな音楽生成記述言語であるかの議論は難しいが，少なくとも DNN 的表現による音楽生成よりは音楽家にとって理解しやすい・記述しやすいということができる．Cope は，この理由は音楽的知識を音楽的知識のまま入れ込むことができるからであるとしている．

よく知られている一番最初の機械学習的なジャズ即興演奏 (ジャズアドリブソロ) 生成システムは遺伝的アルゴリズムを用いた GenJam[Biles 94] である．GenJam の染色体コーディングに音楽的知識はほぼ入っていない．Google Magenta の DNN 表現は LSTM と同様の表現ノードにより極短期的な時系列的な表現が可能になっているという意味では GenJam に比べ発展しているが，その時系列表現が音楽としては極短期出来であること，音楽表現に特化したノードではないことから，音楽表現力が低いのではないかと著者は考えている．

さらに DNN 表現を用いた機械学習では，データ収集も含め学習に非常に大きなコストがかかることを考えれば，DNN 表現ではなく,「人間の音楽家が定めた（音楽的に破綻せず，それれになりのものが生成できることが担保されている，さらに人間の音楽家の個性が反映されている）ルールを元にし，その組み合わせを学習する」といった方針の方が，より良い結果を生むのではないかと，著者は考えた．この「人間の音楽家が定めたルールの組み合わせ」は前述の通り LISP で表現することが容易であり，多くの音楽家が理解し，自ら記述することができる．

また「音楽的知識」だけでは音楽の生成は不可能もしくは不十分なのではないかという議論もある．例えば即興演奏において，五線譜上に表現できる音楽知識ではなく，楽器と身体の関係性で感覚で演奏が行われていることは，Sudnow[Sudnow 79] をはじめとしてジャズピアノのトレーニング等では一般的に知られている．

本稿ではジャズのアドリブソロの単旋律の生成にターゲットを絞り，既存の研究の例を挙げ，さらに著者自身が一人称的に記録してきたジャズアドリブソロ生成の練習メモに対する考察を行う．

2. ジャズ即興演奏の「音楽的知識」と「音楽以外の知識

Sudnow は，自身のジャズピアノの習得過程の記録 [Sudnow 79] の中で，音名等の音楽情報ではなく鍵盤に対する指の動きという身体性のみで記述を行った．鍵盤楽器は五線譜に一番マッピングしやすい楽器であるが，Sudnow の五線譜なしで指がどのように動くかのみでの即興演奏の習得過程の記述により，鍵盤楽器と指の身体性が少なくとも一部のジャズピアニストには五線譜よりも重要であることを示している．

著者は，ジャズ即興演奏生成に機械学習を適用するにあたり，断続的に人間の演奏家への聞き取り調査とそれを元にした生成ルールの構築 [安藤 15] を行なった．

聞き取り調査では，アマチュアベーシスト初学者を対象として，どのようにアドリブを生成しているかのインタビューを元にベースライン生成のルールベースのモデル化を行ったが，人間の生成モデルは，抽象化された五線譜上のような音楽的情報のルールの組み合わせで表現できるわけではなく，身体性と楽器構造が生成されるフレーズに大きな影響を与えていることがわかった．具体的には，初学者であればあるほど，ベースラインの構成音を「音」「音名」「階名」「コードの中の度数」などの音楽的情報としてではなく,「指のポジション」として把握し

連絡先: 東京都日野市旭ヶ丘 6-6，TEL:042-585-8074, dandou@tmu.ac.jp

ており，リードシートからのコードネームの視覚的入力が音楽情報ではなく身体動作に結び付けられていたことがわかった．

Sudnow の記述と著者のインタビュー結果から，鍵盤楽器・弦楽器など「楽器中の特定の箇所・ポイント」に音高が割り当てられ，手指がそのポイントを移動することによって音高を選択する楽器では，少なくとも初学者の間は音楽情報よりも身体性が重要であることが示唆されている．

これらの事例から，Google Magenta のような音楽的情報要素のみから構成される汎用的なモデルを用いた音楽表現には限界がある，もしくは学習コストが非常に大きくなってしまい，実用レベルで人工知能を用いた音楽生成を行うためには，ジャンル・楽器ごとの音楽知識や，楽器ごとに身体性のみならず楽器特有の音楽認知を適切に取り込んだ知識が必要である，と著者は考えた．

3. 2-5-1 フレーズ即興演奏練習の一人称的記録と考察

そこで著者は，自身を被験者としてジャズサックスの即興演奏を習得する過程を一人称研究的に記録することを行ってきた [安藤 18b]．前述のベース，ギターの例や Sudnow のピアノの身体性が強く反映する例とは異なり，手の位置が基本的に動かない管楽器の場合，音名でフレーズを認知する必要があり，弦楽器や鍵盤楽器とはその習得過程における認知の仕方が異なることが予想される．

被験者でもある著者は，音楽大学の声楽学科の在籍経験を持ちクラシック演奏のフレージングの基礎能力を有している．また作曲を行なっておりジャズの理論を理解するのに苦労することはないという特徴を持っている．

この記録の過程では，ジャズサックス演奏者に指導を仰ぎ練習法の指導を受けるだけではなく，自身で考えた練習法を試して,「どのような練習をしている段階で，どのような獲得感が得られたか」を詳細に文章と五線譜に表記したフレーズにより記録した．

本稿では，そのなかの 2-5-1 フレーズに関する記録を抜き出てまとめて提示し，議論を行う．

2-5-1，または 1-2-5-1 とは「トニック (1) - サブドミナント (2) - ドミナント (5) - トニック (1)」という,「安定 - 緊張」を行き来することにより「解決」感覚を表現する典型的な和声進行の典型例の一つである．ハ長調 C-Major スケールでは「CM7 - Dm7 - G7 - CM7」という進行になる．クラシックやジャズなど多くの西洋音楽はこの繰り返しによって成り立っている．その最初のトニック (1) を外したものが一般的に 2-5-1 と呼ばれる．

ジャズではこの 2-5-1 進行を用いた部分転調がドッペルドミナント等により頻繁に行われ，曲全体の調・キーからこの部分のコードを切り離す必要があるため，この部分のみの進行を内在化する練習を行う必要があるとされる．

3.1 和声感覚の習得過程の記録と考察

3.1.1 指示された練習

まず，コードトーンのアルペジオで 2-5-1 の音高の音名を把握する練習法を指示された．図 1 に，練習フレーズの C-major での例を示す．実際の練習では，このような五線譜上に記譜された音符を見てではなく，コードネームを見ながら根音からの音程を度数を頭に思い浮かべながら音名に変換して行なった．(図 1 の例の 3 小節目の前半ならば，練習時に頭に浮かべるのは「Dm7, 7531」であり,「Dm7: 7 →ド，5 →ラ，3 →ファ，1 →レ」と逐次認識していく)

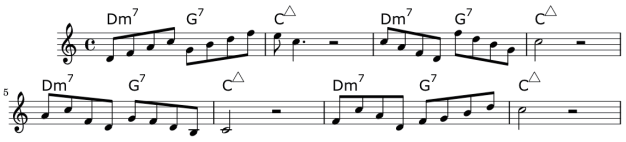


図 1: 練習用として指示された，または著者自身が考えた 2-5-1 のコードトーンアルペジオフレーズ．

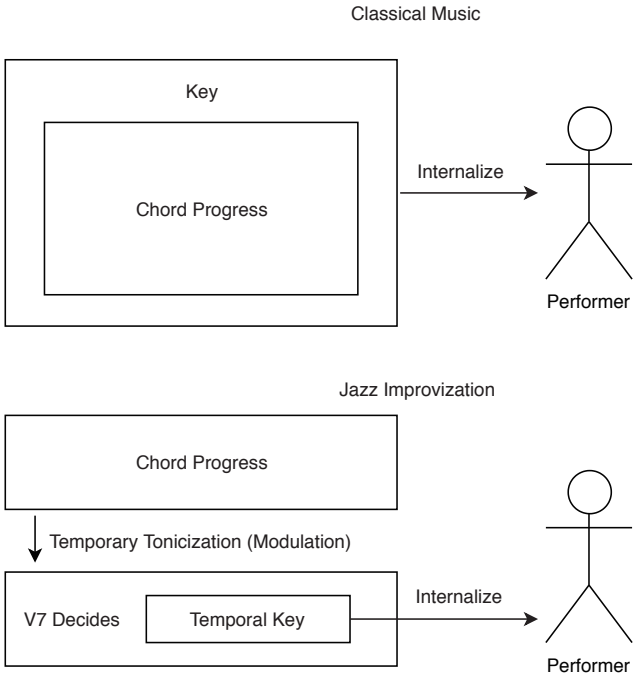


図 2: 著者の中でのクラシックとジャズの調性の把握感覚の違い．クラシックの和声進行は調・キーに内包されているのに対して，ジャズ即興演奏の場合はコード進行の中の属七 (図中「V7」) が部分転調の調・キーを生成しているイメージ．クラシックとジャズでは五線譜上での表現に差がないのに，演奏時の著者の「認識」もしくは「認知」は明らかに異なっている．

3.1.2 和声感覚把握時点での著者の問題点と考察

一般的にアヴェイラブルノートスケールに基づきジャズの即興演奏を行う場合，本来 2 の和音 (図 1 中 Dm7 の部分) のスタート時点で 5-1(図 1 中の G7 → CM7 の部分) に到るまでの調号を把握している必要があるが，著者は図 1 に示したフレーズの練習を始めてから 2 年以上経っても 5 の和音（クラシック和声で属七と呼ばれる和音）のアルペジオに到るまで，全体の調号を把握する，つまり調性を感覚的に明確に把握することができなかった.「頭と指ではわかっているのに，聴取感覚がついてこない」とも言える．音楽の原理的には属七のみが調・キーを決定するのはクラシック・ジャズ双方で自然なことであり，五線譜上に表現してしまえばどちらにも差がないのだが，著者は自身の「認識」もしくは「認知」の感覚の違いに注目した．

図 2 に著者のクラシックとジャズの 2-5-1 の和声進行の感じ方の違いを，記述したメモを元に表した模式図で示す．

クラシックの演奏やルールベースの音楽生成システムを考えた場合，調・キーは主旋律から導き出され，演奏を始める前に既知であることが前提であり，コード進行も調の中に内包さ

れている範囲に限定される．このような著者の様な「調がわからない」感覚は発生しないと考えられる．転調の際も「ディアトニック転調」と呼ばれる手法では移動する先の調・キーがわかっていることが一般的である．

それに対して，ジャズの即興演奏の時はコード進行のみが提示されており，調整を決定づけるような主旋律の提示は行われない．この状態では「それまでに演奏された旋律」が調性のガイド足り得ないため，5 のコード「属七」が出てくることのみが，その部分の調整が決定する要因となっている．

しかしながらクラシック経験者の著者がこの練習をする時に感じた「属七」が出るまでわからなってしまう感覚は，ジャズ特有のごく短短いスパンでの部分転調の極端な多さや，アウト感を演出するために別のコードとも捉えられるテンションノートを中心とした音使いをする場合には，適切な認知なのではないかと考えた．

著者の場合のクラシック的な感じ方はトップダウン的な音楽知識に基づくのに対して，ジャズ的な感じ方は「生成的」である，とも言える．

前述の通り五線譜に記述，つまり五線譜上の音楽表現になってしまえばクラシックとジャズのアドリブソロは全く同じルールで表現可能なのだが，そのバックグラウンドの和声進行の感じ方には，このように大きく違いがあることがわかった．

3.2 曲中での 2-5-1 フレーズ演奏の習得方法

次に実際の曲での 2-5-1 フレーズの使用についての練習の記録について述べる．

3.2.1 一般的な習得方法

曲中で 2-5-1 を即興演奏する一般的な習得方法としては，

1. 多くの演奏家が残している 2-5-1 フレーズ (リック) をコピーする．
2. そのフレーズを主要キーに移調できるようにする
3. ターゲットとなる曲の進行中に 2-5-1 を探して，コピーしてきたフレーズを当てはめる

という手順が推奨されている．実際に著者もそのような指導を受けた．ここで 1. の「多くの演奏家が残している 2-5-1 フレーズ」とは，ジャズ特有のアウト感を出すための非コードトーン (オルタードスケールやコンビネーションオブディミニッシュスケールに含まれている音高) が多く含まれており，演奏家がフレーズのストックを増やして使うことで自然にアウト感を出せるようになっている．

3.2.2 著者の問題点と解決法

しかしながら著者の場合,「あらかじめ覚えたフレーズを演奏し出すと，コードトーンが頭の中で鳴らなくなってしまう．認識できなくなってしまう」という問題が発生した．この現象はクラシック演奏経験者で即興演奏を習得しようとする人間には頻繁に見られる傾向のようで，音楽大学でピアノ即興演奏を教授している指導者から同じような状況が多く見られることを伝えられたことから，いわゆるフレーズを認識し演奏するというクラシックの練習経験やフレージングの発想が，コードトーン認知の邪魔をしているのではないかと考えた．今後聞き取り調査も含め詳細な研究が必要である．

そこで著者は，コピーをしないでコードトーンと 9 度 13 度の音を使って「生成的に」演奏をする方向に練習切り替えた．いわばこれは「与えられたフレーズをそのまま覚えるのではなく，与えられたフレーズをマルコフ過程的に一度分析してその生成モデルを獲得する」と言える．アウト感を出すためにはチャーリー・パーカーが取り込んだとされるドビュッシーやラヴェルの考え方をそのまま用い [Amy 86],「より解決感を強く出すための基本の属七からズラす」と捉えることとした結果，アウト感を持つ 2-5-1 フレーズを即興演奏で生成ができるようになった．

フレーズの場合も和声感覚の時と同じく，フレーズを覚えるトップダウン的な知識に基づくのに対して，実際に著者が自然と感じた方法は生成的とも言える手法であると言える．

4. 考察

ここまで即興演奏習得に関する既存研究と，著者の 2-5-1 フレーズの即興演奏習得について一人称的記録をまとめた．

既存研究の調査では，即興演奏には楽器によって身体知が大きく影響している可能性があり，単なる五線譜ベースの音楽知識の表現だけでは即興演奏の表現そのものが難しいのではないかと推測した．

また，即興演奏習得過程の記録のまとめから，クラシック音楽の演奏者としての経験があり，かつ理論的には一般の学習者よりも高いレベルにある著者の場合，1. 和声進行把握がクラシックとジャズで異なっている，2. クラシック音楽演奏経験のフレーズ把握認知が邪魔をしジャズの一般的なリックを習得する練習法では効率が悪い，と推測した．

このように一言にジャズの即興演奏生成モデルと言っても，演奏する人間の音楽的バックグランドによりその認知の仕方，感じ方も様々であり，それが演奏者の個性のような演奏の違いを生んでいると言うことができると著者は考えている．個性が多く反映されることが「創作において評価されるクリエイティビティ」だとするするならば，創作に機械学習を用いる時に音楽の良さ・演奏の個性に絶対的な評価ができなくなってしまい，機械学習を行うためのモデル化の際に特定のモデルを「いいモデル」であると決定することは難しくなる．

また，著者が行ってきたように，訓練された音楽家は，自身の感じ方の分析やとった解決法などを明確に記述できることがわかった．和声感覚の認知では自らの感じ方を元にクラシックとジャズの和声感覚の違いを図示することができた．曲中での 2-5-1 フレーズの演奏習得では，既存のフレーズを使わずに分析することでその生成モデルを獲得することが自身にとって自然であることを「コードトーンが頭の中で鳴らなくなる」という問題と共に考察することができた．

Google Magenta では，ジャンルを問わず同一のモデルにより学習を行う．しかしこのようにクラシックとジャズ即興演奏，または楽器，演奏家のバックグラウンドにより認識・認知が大きく異なることからも，よりジャンルや楽器に特化したモデルを用意する必要があるのではないか，と著者は考えている．さらに，このように明確に記述された内容が表す音楽的知識に基づく認知・生成モデルは，Google Magenta のような DNN での音楽表現モデルよりも高レベルでの認識・認知であり，同様に音楽認知科学でのモデルよりも高レベルであると言える．また，Google Magenta のモデルは音楽家からみれば非常に非合理な表現をされているということもでき，音楽家の認知を取り込む，すなわち人間の音楽家と知識を共有できるモデルを導入すれば，機械学習の効率は大きく上がると著者は予想している．

5. おわりに

本稿では，ジャズ即興演奏習得に音楽知識以外の身体性を考慮する必要性があること，著者自身の一人称研究に基づく

ジャズ即興演奏習得過程記録の例を挙げ，習得過程の認知が人によって大きく異なり，単に五線譜上で表現できる音楽知識だけではジャズ即興演奏を表現し切れない可能性，の 2 つを示した．

今後は，楽器ごとに聞き取り調査などを進め，人間が把握している演奏生成モデルの実例の様々なパターンを作成し，人間と音楽知識を共有する機械学習のための高レベルなモデルへの応用を目指す．

参考文献

[Amy 86] Dommel-Diny, A., 笠羽 映子：ドビュッシー, 演奏家のための和声分析と演奏解釈 / アミ・ドメル・ディエニー著, シンフォニア (1986)

[Biles 94] Biles, J.: GenJam: A genetic algorithm for genrating jazz solos, in *Proceedings of 1994 International Computer Music Conference*, Arhus (1994), ICMA

[Cope 05] Cope, D.: *Computer Models of Musical Creativity*, MIT Press, Cambridge (2005)

[Sudnow 79] Sudnow, D.: *Ways of the Hand - The Organization of Improvised Conduct*, MIT Press (1979)

[安藤 05] 安藤 大地, Dahlstedt, P., Nordahl, M. G., 伊庭 斉志：対話型 GP を用いたクラシック音楽のための作曲支援システム, 芸術科学会論文誌, Vol. 4, No. 2, pp. 77–86 (2005)

[安藤 12] 安藤大地：クラブ系ループ音楽パフォーマンスのためのリアルタイムに音楽を Breeding するシステム, 情報処理学会論文誌, Vol. 53, No. 3, pp. 1030–1040 (2012)

[安藤 15] 安藤 大地：ジャズ即興演奏における 4 ビートベースラインの 2 段階対話型遺伝的プログラミングによる生成, 情報処理学会インタラクション 2015 シンポジウム (2015)

[安藤 17] 安藤 大地：対話型遺伝的プログラミングを用いた無調性音高配列からの和声進行の生成, 情報処理学会インタラクション 2017 シンポジウム (2017)

[安藤 18a] 安藤 大地：CPPN 表現と Neuro-Evoluation を用いたエレクトロニカのためのノイズ音素材生成, 情報処理学会インタラクション 2018 シンポジウム (2018)

[安藤 18b] 安藤 大地：クラシック演奏家のジャズ即興演奏習得過程の記録と考察, 日本音楽即興学会第 10 回大会 (2018)

[大久保 17] 大久保 雅基：携帯音楽プレイヤーを使用した遠隔演奏同期システムの考案, 先端芸術音楽創作学会会報, 第 9 巻 (2017)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

人工知能のファッションデザインへの適用可能性：ファッションの文化的背景からの考察

Of application of artificial intelligence to fashion design: based on the analysis of cultural background of fashion

藤嶋陽子*1*3
Yoko Fujishima

佐倉統*2*3
Osamu Sakura

*1 東京大学大学院学際情報学府
Graduate School of Interdisciplinary Information Studies, The University of Tokyo

*2 東京大学大学院情報学環
Interfaculty Initiative in Information Studies, The University of Tokyo

*3 理化学研究所革新知能統合研究センター
RIKEN-AIP

Abstract: We argue that how AI can be applied to the field of fashion design. Historically, fashion inherent peculiar negative attitude toward technology, related to the traditional thought against artistic object and mass production. This philosophical background would be significant for the evaluation of an impact of AI upon fashion design. Moreover, based on the analysis of the current technological character of AI and several experimental trial which AI designs fashion, AI has affinity with the design process in fashion. However, the current technological mechanism of AI corresponds with only primary research process in the inclusive flow of fashion design. The value system of fashion strongly depends on the sensibility of the fashion designer, and it closely linked with the inclusive process, rather than each step. Therefore, AI requires the symbiosis with human designers to create value in the context of fashion design.

1. はじめに

あらゆる分野でそうであるように，今日のファッション業界においても人工知能の導入というのは大きな注目を集めている．この革新的なテクノロジーの導入をめぐって，どういった行為者がイニシアチブをとるのか，どのように普及するのか，これは未来のファッションのゲームセッターを決めるのと同義かもしれない．それゆえ，巨大な資本を導入した大型プロジェクトがすでに遂行されている．その最たる例が，アメリカの amazon 社による人工知能デザイナーの開発企画であろう．これは画像からトレンドやスタイルの学習を試み，類似するデザインを作成，製品化するアルゴリズムの開発を目指すものである．また，IBM とトミー・ヒルフィガー，そしてニューヨーク州立ファッション工科大学（FIT）Infor Design Tech の共同プロジェクトでは，トミー・ヒルフィガーの製品やコレクションの画像，模様や型紙などを学習させ，デザインのインスピレーション源として用いた．さらにはデータ分析の世界的カンファレンスである Knowledge Discovery and Data Mining（KDD）においても AI for Fashion というワークショップが開催されており，特に技術開発の方面から人工知能のファッション分野における適用可能性が問われている．

このように既に多様なプロジェクトが進められているわけだが，現在の開発方向性は主に，トレンドや消費者の購買データを利用した生産や流通の最適化などである．EC 販売が拡大している今日，購買履歴や消費傾向といったデータの収集は容易になっており，こういった顧客データとブランドの販売実績やトレンドリサーチを結びつけた確かなデータを活用していくことは，社会問題として着目されつつある在庫廃棄の問題を考えるうえでも重要となっている．生産，流通のサイクルが非常に短くなり，消費者のニーズも多様化している今日のファッション産業において，人工知能を活用した商品制作への期待がさらに高まっていくのは間違いないだろう．

しかし，ファッションには芸術としての側面もある．芸術品と商品というふたつの性質は，有名デザイナーによるいわゆるハイファッションと大量生産されるマスファッションという区別とも重なりあう．それは，作品と製品という括り方の差でもあるかもしれない．こういった作品としての価値は，デザイナーの文化的背景，時代感覚，既成概念に囚われない批判的な視点，そういった要素がオーディエンスと共鳴することで生み出され，芸術や歴史的な文脈から評価される．人工知能による創造活動の可能性をめぐっては，芸術分野は既に直接的に回答が試みられ，技術的な可能性を問うだけではなく，アーティストとの協働関係や美術マーケットにおける評価などを考察するうえで重要な知見を提示する．例えば 2018 年には，イギリスでは AI の描いた裸婦画が世界的な芸術賞であるルーメン賞で金賞を受賞したり，アメリカでは AI の描いた肖像画がクリスティーズのオークションにおいて 43 万 2500 ドルで落札されたりといったニュースが注目を集めた．こういった芸術分野での取り組みに対し，ファッションにおいても同様に人工知能によるデザインを創造性という観点から評価し得るのだろうか．この問いに解答するためには，ファッションという対象の性質—単純に芸術作品，製品のどちらかではない，この曖昧さがファッションの固有性でもあり，捉えがたさでもある[Choi 18] —を考慮し，ファッションに固有のテクノロジー観や，デザイナーという仕事の役割や制作プロセスと人工知能の技術的特性を重ね合わせて検討していく必要があるだ

連絡先：藤嶋陽子，東京大学学際情報学府／理研 AIP，東京都文京区本郷 7-3-1，yokofujishima@g.ecc.u-tokyo.ac.jp

ろう.

こうした視座のもと本研究では，ファッションとテクノロジーの関係を歴史的な文脈を踏まえつつ，デザインプロセスと人工知能の技術的特性の親和性について検討を行う．そのうえで，ファッションデザインとの親和性と将来的な課題を明らかにし，ファッションとテクノロジーの関係性を批判的再考して更新することを試みる．

2. ファッションにおけるテクノロジー観

2016 年のニューヨーク，メトロポリタン美術館ファッション部門の企画展であった“MANUS×MACHINA”は，Apple 社がスポンサーとなり，ファッションにおける機械技術の位置付けを問い直す挑戦的なプロジェクトだ．そして，ファッションの歴史のなかでのテクノロジーの位置付けに対し，極めて重要な示唆を与えるものとなった．

この展覧会の序文のなかでキュレーターのアンドリュー・ボルトンは，人間-対-機械という構図がいかにファッションの制度のなかに定着し，その価値と結びついているかを語っている[Bolton 16]．ボルトンによると，19 世紀中盤にオートクチュールという制度が確立してから，ファッションでは手仕事と機械が対立関係に置かれ，この構図が規定されてきた．ファッションの歴史のなかで，もっとも影響を与えたテクノロジーはミシンである．ミシンの登場は衣服の生産のスピードを大幅に早め，大量生産とそれによる低価格化を実現させた．それ以降，テクノロジーは大量生産と結びつけられ，機械で作られたものは手仕事よりも価値の劣るものと捉える価値観が定着した．そしてこのハンドメイドとマシーンメイドを対立させる構図は思想的なレベルだけではなく，オートクチュールの組合であるサンディカ協会への参入条件に手仕事が必須とされるという公式の制度として具現化しており，注文服であるオートクチュールと既製服のプレタポルテの区分とも重なり合っている．この大量生産との結びつきや複製される既製服に対するまなざしは，展覧会の副題が“*Fashion in an Age of Technology*”であることからも示唆されるように，ヴァルター・ベンヤミンの提起した機械的な複製と芸術をめぐる議論のもとにある[Benjamin 08]．唯一無二のオリジナルであること，それが芸術作品の要件とされるような伝統的思想のなかで，複数作られるということは芸術的価値を否定する論理となる．それゆえ，機械で作られたものはどこか無機質で，規格化されたものは平板なものであるような印象となるのだ．

そのうえでボルトンは，オートクチュールも実際のところはミシンや標準パターンといった大量生産の技術とされるものに支えられて発展したと述べている．実際，オートクチュールの“父”と言われるチャールズ・フレデリック・ワースのクチュールハウスは，ミシンによる生産に依拠して事業を国際的な規模のものに発展させた．そして現代の最先端技術，例えば３D プリンターや超音波圧着，レーザープリンターといったものは，オートクチュールとプレタポルテの両方に用いられ，デザイナーの創作を支えている．それゆえ，この展覧会での最終的な主張は，こういったオートクチュール-対-プレタポルテ，ハンドメイド(人間)-対-マシーンメイド(機械)という二項対立の構図を再考する必要があるということであった．

大量に生産されて大量に販売され利益を拡大させること，生産数を制限してプレミアムな価値を守ること，ファッションは常にこの両義的な志向性を有する．ミシン以降，ファッションに関わるテクノロジーの進化は，従来は不可能であった創造的な表現を実現可能とする一方で，表現を容易とすることはやはり生産を容易にすることであり，複製性からも逃れることはできない．ファッションにおけるテクノロジーは，生産の拡大と豊かな創造の実現—この相反するふたつの可能性を有するものなのだ．

3. ファッションの制作プロセスの親和性

こういったファッションデザインに内在するテクノロジーへの否定的価値観に対し，人工知能はそれを変える—もしくは強化するものなのだろうか．また，そのうえでファッションデザイナーたちにどのように活用されていくと考えられるのだろうか．ファッションデザインにおいて人工知能が受容される可能性を議論するためには，その技術的特性を確認し，ファッションデザインの制作プロセスとの親和性を考察する必要がある．

現在の人工知能技術の中核は machine learning，すなわち，コンピュータ・プログラムが新たな課題 task や外的環境の変化を学習し，適切な解 solution を見つけることを可能にする技術である．とくに，入出力のノード nodes が複数（4 層以上）の階層 layer から成る——人の脳により近い——neural network の学習（deep learning, or deep neural network）技術が実現したことにより，複雑な学習が可能になった．

deep learning 技術の進歩はとても急速なので，近い将来人工知能／ロボットの知的能力が人類のそれを上回り，人工知能／ロボットが人類の制御がきかない状態で暴走したり，さらには人類を支配する可能性があると予測する人たちもいる[e.g., Kurzweil 05; Bostrom 14]．現にデザインの領域においても，既にウェブデザインやロゴのデザインの自動生成サービスなどが登場しており，人間が行わないと難しいと考えられていたクリエイティブ領域にも着実に進出してきている．人工知能／ロボット技術の特徴のひとつは，高度に自律的 autonomous であることだ．自律的システム autonomous system, or "AS" が人間の知能の少なくとも一部を凌駕することで，多くの仕事がこれらの AS で代行されるだろうし，人間にしかできない作業は何なのか，再検討がなされていくだろう．すなわち，人と機械がどのように（共存 co-existing ではなく）共生 (sybiosis) するのかを再検討する必要に私たちは迫られている [Tegmark 17]．この点で，ファッション・デザインという，非常に「人間的」なセンスが重きをなしている領域で，人工知能と人の共生系がどのような状態になるのかを考察することは，ファッション研究だけでなく，人工知能の研究・技術開発にとっても有意義な作業である．

冒頭でも触れた人工知能による絵画作品やファッションデザインの試みで主に用いられているのが，敵対的生成ネットワーク Generative Adversarial Network（GAN）である．これは大量の画像を読み込ませることで人工知能が自ら共通する構造や法則を発見し，それを反映した画像の生成が可能となる技術である．最近の事例でいうと，2018 年 7 月に AI アーティストのロビー・バラットが，ファッションブランドであるバレンシアガの広告，ランウェイショー，広告に基づいて仮想のコレクションを生成し，話題を集めた．この例ではまだ，テキスタイルの質感が現実的ではないなどの課題があったが，2018 年 11 月に Microsoft が発表したチャットボット“Xiaoice”はさらにこの課題を乗り越え，生地感までも学習し，調和を図ることを実現した．さらにこの Xiaoice は，単語やテーマに基づいてデザイン提案することが

図1　ファッションデザインのリサーチ段階において
作成されるムードボードの例

可能であり，上記のバレンシアガ風のコレクション生成のように特定のデザイナーの作風を模倣したものではない，新たなデザイン提案が可能である．

つまり，現在の人工知能によるデザインは単なる複製や模倣ではなく，新たなデザインの提案も可能な状況にある．Xiaoice も何らかの教師データに基づいてデザイン生成を行っているわけで，それも模倣に過ぎないと考えられるかもしれない．しかしそれは，人間が行うデザインプロセスにも同様の側面がある．ファッションデザインのコレクションプロセスにおいては，コンセプトやデザインの第一段階に，リサーチの段階がある．それは，自社や場合によっては他社の過去のコレクションやアート作品や映画など，多様なリソースからイメージをかき集めてムードボードを作成し，コンセプトイメージを精緻化していく作業である．デザイナー自身や，ブランドの規模によってはアシスタントが協力してこれらの作業を行い，そこにブランドらしさやデザイナー独自感性を加えてコレクションが作られていく．つまり，人工知能が行っているデザイン生成のプロセスは，ファッションデザイナーたちが通常行っている制作プロセスの流れと非常に類似性が高いのだ．

人工知能とファッションデザイナーの制作プロセスの親和性は，人工知能を人間のデザイナーを代替する脅威として受け止められるかもしれない．しかしながら，制作プロセスの親和性が高いがゆえに，既存のプロセスを簡易化するような補完的な使用もしやすいという考え方も可能である．また，既存の制作プロセスにおいては，過去のコレクションからのデザイン生成はあくまでリサーチ段階であり，そこからデザイナーの感性によって取捨選択して改善し，クチュリエの技術によって具現化していく段階が必要となる．ファッションにおける作品の評価は，作品そのものが持つ審美性だけではなく，デザイナーのカリスマ性やブランドイメージ，コンセプトの社会的インパクトといった作品に付随する文脈によって規定される部分が大きい．ファッションショーでコレクションの発表を行うようないわゆるハイブランドの制作を考えると，デザイナーはコンセプトを決め，全体の統一を図っていくディレクターのような役割になる．1体1体の作品ももちろん重要ではあるが，実際のデザインの素案やデザイン案から具現化していく作業は，アシスタントやパタンナー，クチュリエなどに委ねられている場合が多い．それゆえ，ファッションデザインの価値の核となるデザイナーの感性という要素は，リサーチの段階よりもむしろ，そこからのブラッシュアップしていく際のひとつひとつの選択や，一連の制作の包括的なプロセスにおいて見出されるものであると考えられる．

人工知能の生成するデザインは，デザインを行うという意味では完結しているのかもしれない．しかしながら，その創作物としての価値を高めるのは，デザイナーという固有の人格のフィルターが必要である．そこまでがファッションにおける一連の制作プロセスであるとすると，人工知能だけでの自動的な制作というのはそれだけでは価値を生み出すことはできず，そこにデザイナーの感性という別の価値を付け加えることで初めて，作品として成立しうると考えられる．これはあくまでテクノロジーをツールとして利用していくような志向性で，人工知能という技術の考え方にそぐわないものであるかもしれないが，ファッションという領域の価値システムにおいては属人的な要素が大きいことを考えると，人間との協働が必ず必要になってくるのではないだろうか．

4．今後の課題と展望

こういった人工知能を，実際にファッション産業やファッションデザイナーたちはどのように活用し，意義付けていくのだろうか．この答えを出すにはまだ時間が必要であるかもしれないが，既にデザイナーが自らのコレクション制作に導入した事例もある．例えば，2016 年に IBM Watson が Cognitive Fashion として推進するプロジェクトの一環として，オーストラリアのデザイナーであるジェイソン・グレッチと協力して 12 体のドレスを制作し，メルボルンファッションウィークで発表したものは，先駆的な事例のひとつであろう．こういった取り組みは，デザイナーと共同した時に人工知能における創作がどこまで実現可能かを考えるうえで示唆的なものである．しかしながら，従来の人力のみで制作されたものとの差異や，“ファッション”という社会的，歴史的な枠組みのもとでいかに評価し，受容されうるのかを検討するためには，最終的な成果物よりもむしろ，作品の制作プロセスに着目していく必要があるだろう．

現段階では一般的なファッションデザインの制作プロセスから類推して分析することが限界であったが，現在，理

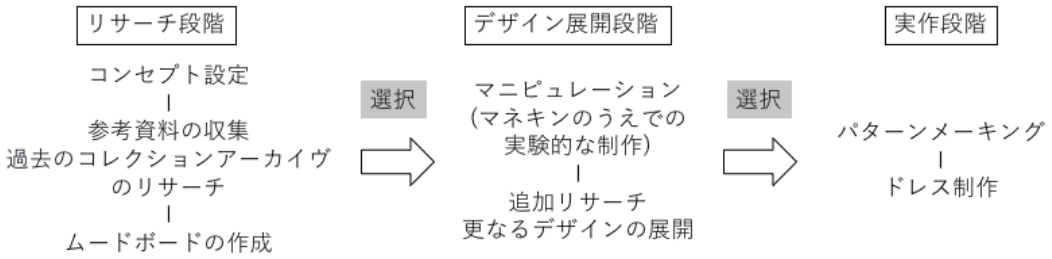


図２　ファッションデザインにおける一般的な制作プロセス

化学研究所革新知能統合研究センターにおけるファッションと人工知能のプロジェクトが遂行されている．これは，人工知能の感性表現学習の可能性を検討するもので，オートクチュールを専門に手がけるファッションデザイナーのエマ・理永と共同し，人工知能とファッションデザイナーの共生関係の構築の可能性と課題の検討を行っている．今後はこのプロジェクトの題材として，より実証的なレベルからファッションデザイナーと人工知能の創作における協働関係について検討していく予定である．

参考文献

[Benjamin 08] Benjamin, W.: The Work of Art in the Age of its Technological Reproducibility. In: Jennings M.W. et al., eds., The Work of Art in the Age of Its Technical Reproducibility and Other Writings on Media, Belknap of Harvard University Press, Cambridge, MA. (2008)［野村修訳：複製技術時代の芸術作品．所収：多木浩二：ベンヤミン「複製技術時代の芸術作品」精読．岩波現代文庫，岩波書店 (2000)

[Bolton 16] Bolton, A.: Manus x Machina: Fashion in an Age of Technology. Metropolitan Museum of Art, New York, NY. (2016)

[Bostrom 14] Bostrom, N.: Superintelligence. Oxford University Press, Oxford. (2014)

[Choi 18] Choi, K. and Lewis, V. D.: An Inclusive System for Fashion Criticism. International of Journal of Fashion Design, Technology and Education, 11(1): 12-21 (2018)

[Kurzweil 05] Kurzweil, R.: The Singularity Is Near. Viking, New York, NY. (2005)

[Tegmark 17] Tegmark, M.: Life 3.0: Being Human in the Age of Artificial Intelligence. Knopf, London. (2017)

3L4-OS-22b-04

トーン特徴量を損失関数に用いたcGANでのトーン貼り

Tone Pasting Using cGANs with Tone Feature Loss

坪田亘記　相澤清晴
Koki Tsubota　Kiyoharu Aizawa

東京大学
The University of Tokyo

Tone pasting is one of the processes of manga creation and there is a demand for automatic tone pasting. In this study, we tackle a task of automatic tone pasting of manga characters. Tone pasting is difficult because tones have characteristic patterns. It is hard to learn tone patterns for usual conditional generative adversarial networks (cGANs) which are combined with L_1 loss or perceptual loss. To train pasting tones in a tone pattern aware manner, we introduce *tone feature loss* to cGANs. *Tone feature loss* is the distance between tone features of target images and those of generated images. We performed experiments on two characters in Manga109 and showed our results are equal to or more visually appealing than those by a baseline method.

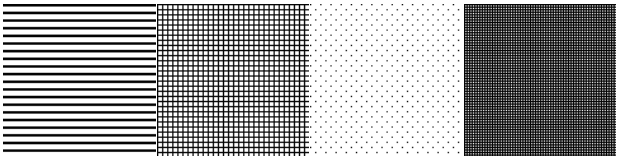


図 1: トーンの例．トーンには特有のパターンが存在する．

1. はじめに

漫画の制作において，トーン貼りは欠かせない工程である．漫画の制作の現場では, 人手によってトーン貼られている．トーンの貼り方にはアナログの方法とデジタルの方法がある．前者は，トーンを貼りたい領域の形に切り貼りするために人手が必要となる．また後者についても，領域を手動で選択する必要があり人手が必要となる．トーン貼りの自動化は，このような人手のコストを削減するため需要がある．

本稿では，キャラクターの線画に対してトーンを貼ることを目的とする．トーンは，自然画像やイラストとは異なり，特有のパターンが存在する（図 1）．自然画像の生成や線画の着色に用いられるような手法 [Isola 17, Ci 18] では，ベタに近いトーンは生成することはできるが，粗い縞模様といったパターンは生成することが難しい．これは既存手法がトーンのパターンを考慮していないためである．

我々は，トーン貼りの学習において，トーンのパターンを考慮した学習を行う手法を提案する．一般的な画像から画像への生成の学習に用いられる conditinal generative adversarial network (cGAN) の枠組みを利用し，その学習時に，損失関数として目標の画像と生成画像のトーン特徴量間の距離を用いる．トーン特徴量としては，同じトーン領域内では同じ特徴量を示すものを利用する．実験により，我々の提案手法がトーンのパターンを学習しやすい手法になっていることを確認する．

連絡先: 東京大学　情報理工学系研究科
〒 113-8656 東京都文京区本郷 7-3-1
TEL: 03-5841-6761
Email: tsubota@hal.t.u-tokyo.ac.jp

2. 関連研究

2.1 線画・漫画の着色

線画の着色は，キャラクターの線画に対するトーン貼りと，線画内の領域を選択して色を置くという点で近いタスクである．pix2pix [Isola 17] では，cGAN を用いて自然画像のエッジから自然画像を生成している．また，cGAN は漫画の着色 [Hensman 17, Furusawa 17] やイラストの着色 [Ci 18] でも用いられる．本研究でも既存の線画の着色手法に倣い，cGAN を用いてトーン貼りを行う．

2.2 cGAN を用いた画像生成

cGAN では，損失関数として生成画像を本物の画像に近づけるような学習をするための adversarial loss と，タスク依存の損失関数を組み合わせて用いる．タスク依存の損失関数として，L_1 loss を用いる手法 [Isola 17] や，L_2 loss を用いる手法 [Pathak 16]，perceptual loss [Ci 18] を用いる手法がある．

これらの手法は，トーンのパターンを考慮した損失関数とはなっていない．L_1 loss や L_2 loss はトーンの貼り方が 1 ピクセルずれただけでも損失が発生してしまう．また perceptual loss は自然画像で学習した畳み込みニューラルネットワークから得られる特徴量間の距離を損失関数としており，特徴量がトーンのパターンを考慮してない．本研究ではトーンのパターンを考慮するために，タスク依存の損失関数として，トーン特徴量の距離に基づく損失関数を用いる．

3. 手法

トーンのパターンを考慮するために，cGAN で用いられるタスク依存の損失関数としてトーン特徴量間の距離とする手法を提案する．手法の概要を図 2 に示す．まず，cGAN を用いたベースライン手法と言える pix2pix [Isola 17] を説明した上で，提案手法の説明を行う．

以降で数式を用いて説明するために文字を定める．x が入力画像（線画）であり，y が x に対応したトーン付きの画像である．また cGAN における G が生成器であり，D が識別器である．

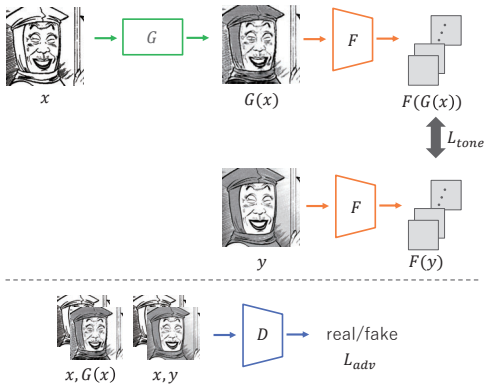


図 2: 手法概要．ⓒやまだ 浩一

3.1 pix2pix

pix2pix [Isola 17] は，ペアの学習データを用いて画像から画像への変換を行う手法である．生成画像を本物の画像に近づけるような学習を行うための adversarial loss

$$\mathcal{L}_{adv}(G,D) = \mathbb{E}_{x,y}\left[\log D(x,y)\right] + \mathbb{E}_{x,y}\left[\log\left(1 - D(x,G(x))\right)\right], \quad (1)$$

を導入している．また，タスク依存の損失関数としては L_1 loss を用いている．L_1 loss は以下の式で定められる．

$$\mathcal{L}_1(G) = \mathbb{E}_{x,y}\left[||y - G(x)||_1\right]. \quad (2)$$

全体の損失関数としては，この両式を組み合わせた，

$$\mathcal{L}(G,D) = \mathcal{L}_{adv}(G,D) + \lambda\mathcal{L}_1(G), \quad (3)$$

で定められる．この損失関数を用いて，以下の式に基づき G の学習を行う．

$$G^* = \arg\min_G \max_D L(G,D). \quad (4)$$

しかし，前述したように，タスク依存の損失関数として L_1 を用いた場合は，トーンのパターンの学習が難しい．

3.2 トーン特徴量を用いた損失関数

提案手法は，cGAN におけるタスク依存の損失関数として，トーンのパターンが学習しやすいようにトーン特徴量間の距離を用いる手法である．トーンの特徴抽出器を F と定めると，

$$\mathcal{L}_{tone}(G) = \mathbb{E}_{x,y}\left[||F(y) - F(G(x))||_1\right], \quad (5)$$

と損失関数は表される．本研究では F として，Gabor フィルタを用いた特徴量 [Qu 06] を利用する．この F は微分可能であることに注意が必要である．

タスク依存の損失関数として $\mathcal{L}_{tone}$ を導入した場合，全体の損失関数は以下の式で表される．

$$\mathcal{L}(G,D) = \mathcal{L}_{adv}(G,D) + \lambda_{tone}\mathcal{L}_{tone}(G). \quad (6)$$

この損失関数を，式 4 を用いて最適化する．

4. 実験

4.1 実験設定

トーン付きの画像データセットとして，漫画画像データセットの Manga109 [Matsui 16, Ogawa 18] の漫画を用いた．データセットに付与されている，“body” のアノテーションのバウンディングボックスに従って画像の切り出しを行い，単一キャラクターについて実験を行った．対象としたキャラクターは，“平成爺メン” に登場する “留” と “タップ君の探偵室” に登場する “篠原かもき” である．画像の例を図 3 に示す．“留” はトーンが粗く，“篠原かもき” はトーンが細かいことが特徴である．

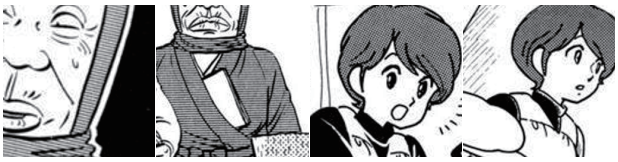


図 3: 学習データの例．左側が “留”，右側が “篠原かもき”．ⓒやまだ 浩一 ⓒふくやま けいこ

“留” の画像は 151 枚あり，また “篠原かもき” の画像は 239 枚あった．“留” の画像は，学習用データを 104 枚，検証用データを 14 枚，テストデータを 33 枚と無作為に分割した．“篠原かもき” の画像は，学習用データを 168 枚，検証用データを 20 枚，テストデータを 51 枚と無作為に分割した．

トーン付き画像と対応した線画データセットは存在しないため，トーン除去の手法を漫画に適用した後，2 値化することで，線画とトーン付き画像のペアを作成した．トーン除去の手法としては [Li 17] を用いた．

ネットワークは，Isola ら [Isola 17] と同様のものを利用した．すなわち，cGAN における生成器として U-net [Ronneberger 15] に基づいたネットワークを用い，識別器として PatchGAN を用いた．学習時の係数は，$\lambda_{tone} = 100$ とした．

入力画像については，最短の辺の長さが 256 になるようにアスペクト比を変更せずにリサイズした後，学習時は 256x256 にランダムクロップを行い，テスト時はセンタークロップを行った．

4.2 実験結果

実験結果を図 4 及び図 5 に示す．ベースライン手法は pix2pix である．図 4 から，提案手法はベースライン手法と比較してトーンのパターンを生成できていることが分かる．また，pix2pix における λ が小さい場合は，トーンのパターンを生成できている部分もあるが，一方で $\mathcal{L}_1$ の重み λ が小さくなることで塗るべき領域の学習が難しくなっている．図 5 のようにトーンが細かい場合は，提案手法・ベースライン手法いずれの場合でも十分学習できている．

5. 結論

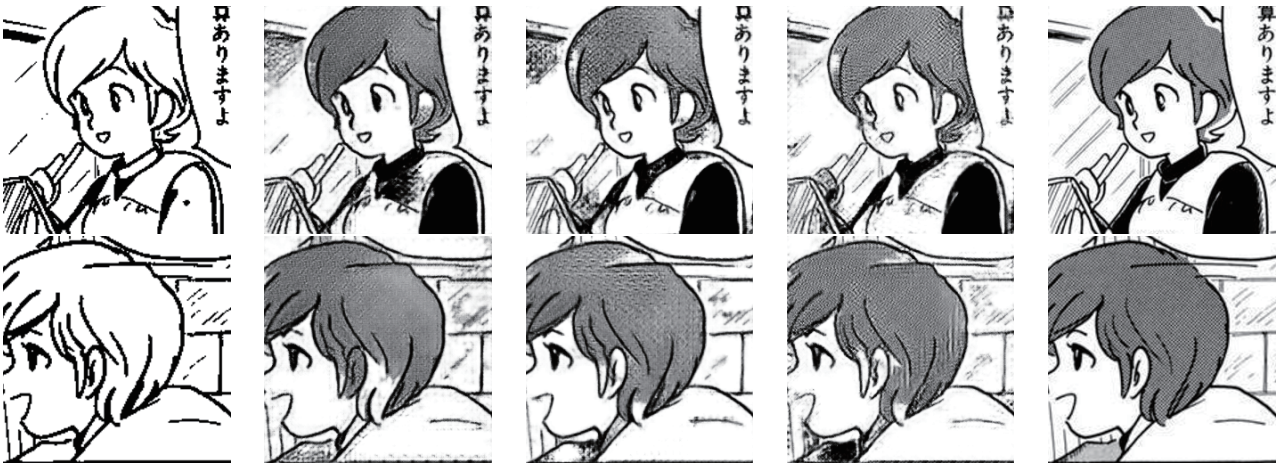
本稿では，キャラクターの線画に対するトーン貼りに取り組んだ．パターンを考慮するために，cGAN でのタスク依存の損失関数として，目標の画像と生成画像とのトーン特徴量間の距離を損失関数として用いた．実験により，提案手法がベースライン手法である pix2pix と比較して，有効な場合があることを示した．

今後の展望として，2 つ考えている．1 つは，学習データに存在しないキャラクターへのトーン貼りである．線画の着色の研究のように，複数キャラクターを用いて学習することで実現できると考えている．

もう 1 つは，テストデータとして純粋な線画を用いた場合の実験である．本稿では，テストデータとして漫画から抽出し

入力画像　pix2pix, $\lambda = 10$　pix2pix, $\lambda = 100$　提案手法　Ground Truth

図 4: “留” での実験結果．ⓒやまだ 浩一



入力　pix2pix, $\lambda = 10$　pix2pix, $\lambda = 100$　提案手法　Ground Truth

図 5: “篠原かもき” での実験結果．ⓒふくやま けいこ

た線画を用いて実験を行っている．そのため，得られた線画には元々貼ってあったトーンの情報が残っている可能性がある．

参考文献

[Ci 18] Ci, Y., Ma, X., Wang, Z., Li, H., and Luo, Z.: User-Guided Deep Anime Line Art Colorization with Conditional Adversarial Networks, in *Proceedings of the 26th ACM International Conference on Multimedia*, pp. 1536–1544 (2018)

[Furusawa 17] Furusawa, C., Hiroshiba, K., Ogaki, K., and Odagiri, Y.: Comicolorization: Semi-automatic Manga Colorization, in *SIGGRAPH Asia Technical Briefs* (2017)

[Hensman 17] Hensman, P. and Aizawa, K.: cGAN-Based Manga Colorization Using a Single Training Image, in *2nd International Workshop on coMics Analysis, Processing, and Understanding, 14th IAPR International Conference on Document Analysis and Recognition*, pp. 72–77 (2017)

[Isola 17] Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A.: Image-to-Image Translation with Conditional Adversarial Networks, in *Proceedings of the IEEE Conference on CVPR*, pp. 5967–5976 (2017)

[Li 17] Li, C., Liu, X., and Wong, T.-T.: Deep Extraction of Manga Structural Lines, *ACM Transactions on Graphics*, Vol. 36, No. 4, pp. 117:1–117:12 (2017)

[Matsui 16] Matsui, Y., Ito, K., Aramaki, Y., Fujimoto, A., Ogawa, T., Yamasaki, T., and Aizawa, K.: Sketch-based manga retrieval using manga109 dataset, *Multimedia Tools and Applications* (2016)

[Ogawa 18] Ogawa, T., Otsubo, A., Narita, R., Matsui, Y., Yamasaki, T., and Aizawa, K.: Object Detection for Comics using Manga109 Annotations, *arXiv preprint arXiv:1803.08670* (2018)

[Pathak 16] Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T., and Efros, A. A.: Context encoders: Feature

learning by inpainting, in *Proceedings of the IEEE Conference on CVPR*, pp. 2536–2544 (2016)

[Qu 06] Qu, Y., Wong, T.-T., and Heng, P.-A.: Manga Colorization, *ACM Transactions on Graphics*, Vol. 25, No. 3, pp. 1214–1220 (2006)

[Ronneberger 15] Ronneberger, O., Fischer, P., and Brox, T.: U-net: Convolutional networks for biomedical image segmentation, in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241 (2015)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

創作者と人工知能が創る創作の未来

Creative Future will be Created by Contents Creators and Artificial Intelligence

上野 未貴 *1
Miki Ueno

*1豊橋技術科学大学, 情報メディア基盤センター
Toyohashi University of Technology, Information and Media Center

Contents creators well consider their thoughts and feelings to represent original contents. Recently, there have been proposed lots of researches for creative contents from the several aspects in artificial intelligence field. I discuss the possibility of collaborative methods and novel representation for contents creators and artificial Intelligence.

1. はじめに

創作者は表現したいものをいかに伝えるかに想いを巡らせ，創作物を創っている．その過程には，対象とするテーマだけではなく，創作者個人の感性や体験，表現材料や技法の工夫など，多くの要素が関わっている．昨今，データの増加や深層学習の台頭により，人工知能が創作物を自動生成する，解析する，また創作活動の支援をする，という立場の研究が増加している．しかしながら，創作過程のすべてを担うような技術は未だ表れていない．技術発展の目覚ましさから，自動生成に関わる研究はこれまでの創作分野を脅かすもののように見られることもあるが，創作者と人工知能は既に協調して物を創っているともみなせる．例えば，画像に関しては，画像処理・描画ソフトウェア，ペンタブレット，などの性能向上には人工知能分野の手法が関わっている．さらに，創作現場には，情報技術の進化が密接に関わっている．辞書やシソーラスなどの資料の入手の易化，オンラインでの創作者同士の交流の発展などにより，新たな表現技法が生まれ，個人の創作がより身近になったことで，創ることのできる作品の幅が広がっている．その上で，分野が発展したことで生まれる新たな研究が人が創る過程にどう関わる可能性があるのか，未来の表現がどう変容するのか，本稿では表題のセッションに関わる研究とその展望について述べる．

2. 分野動向

本章では，分野を俯瞰するため創作の種類別に研究を紹介する．創作物に関する研究は，静画，映像，音楽，俳句，刺繍など，数多く存在するが，本稿では，言語と画像による表現を複数組合せることで現実世界の事象や人物に限定せずストーリーを描くことが多い創作物を中心的に取り上げる．

漫画 漫画の工学的研究はコミック工学 [1] において広がり，オブジェクト識別 [2]，漫読書意欲に関する調査 [3]，メタデータの構築 [4]，コマ画像生成の試み [5], など，多数の研究 [6] がある．また分野発展に寄与するデータセットとして，日本の絶版漫画 109 冊に対しアノテーションをした Manga109[2][7], 100 ページの漫画に台詞やレイアウトなどの情報をアノテーションした eBDtheque[8] がある．

小説 ショートショートの創作に関わる研究 [9]，小説の自動生成方法の研究 [10]，共創による物語生成システム [11]，小説の文生成手法の研究 [12] がある．

絵本 子どもが ICT に触れて創造活動をするためのインタラクティブ絵本システム [13] や，絵本の自動生成に関する [14][15] 研究が知られる．

アニメ 中間生成物の絵コンテの作品識別 [16] などがある *1.

3. 創作過程における研究の利用

以下に特に創作者が本分野の研究手法やソフトウェアを利用できる可能性が高い観点を整理する．

3.1 資料検索

創作中に必要な資料を検索するために必要となる時間は多い．特定のキャラクタを探したり，特定のシーンを探す場合である．例えば漫画に関する研究では作品の内容情報探索 [17] や，ネームを検索する支援 [18] がある．

3.2 発想支援

用意した辞書情報から人物名やアイテム名などの名前の生成をするツールはインターネット上の個人サイトに多数存在する．近年のニューラルネットワークを利用した小説の自動生成手法で全文の自動生成をすることは未だ困難だが，類似文を探したり，部分的に自動生成された会話文やあらすじ，要約結果を提示することで，ストーリーの発想に役立つ可能性がある．

3.3 線画・着色の支援

ペイント・ドローソフトのペンの入り抜きやベジェ曲線などのアルゴリズムは長年創作に役立っている．画像処理ソフトでフィルタを適用して全体的な色彩を調整する種類や性能も向上している．近年の技術として深層学習を利用した対話的なラフスケッチのペン入れ手法 [19] による創作作業時間の短縮や，深層学習を利用した自動彩色として PaintsChainer[20] がある．

3.4 作画・構図支援

漫画の創作支援に関する市販ソフト Comipo![21] は 3D モデルを配置し，関節を動かし 2D の漫画を手軽に描くなどの機能を有する．Live2D[22] ではあらかじめパーツを分けて用意してメッシュを適用することで自然な可変画像を作成できる．アニメキャラクタの正面から横へ向くなど，始点と終点のある短

連絡先: 上野 未貴, 豊橋技術科学大学，情報メディア基盤センター, 〒 441-8580 愛知県豊橋市天伯町雲雀ヶ丘 1-1,
E-mail: ueno@imc.tut.ac.jp

*1 《JST 共催》人工知能はアニメを生み出せるか？—「情報ひろばサイエンスカフェ」で研究者と市民が語り合う: https://scienceportal.jst.go.jp/reports/other/20181128_01.html

い動作の間に必要な複数枚の画像を中割と呼び，中割を自動生成 [23] するツールが市販されている．アニメ表現で重要な髪の毛のモーションを自動生成する手法 [24] などが知られる．

3.5 設定管理

ジャンルやストーリーの長さによって，人物や舞台の詳細，伏線，起伏などを管理することが望ましい場合がある．一般的に，創作者は紙やテキストエディタを利用し自由なフォーマットで個人的にメモを取ることが多い．一方で，分業が中心で工程が多いアニメの場合には，各人物の複数の角度，感情的な表情を描いた設定資料に加え，絵コンテ，原画などの中間媒体を複数人で情報を共有する必要がある．設定集などの形で一部の資料が抜粋して整理されない限り，公になることは少なく，創作者の要望を聞き取り技術的に支援する余地がある．

4. 創作者との共同研究

本章では，筆者が関わる創作者と共に進める研究 [25] を，特に漫画と小説に分けて紹介する．要素技術としては，解析・生成両面から広いタスクを想定するものの，創作過程を工学的に体系立て，計算機が心情や個性を考慮したストーリーの創作過程を理解し，創作者を支援することを大きな目的としている．

4.1 漫画

4.1.1 データセット

筆者は創作過程を保持した 4 コマ漫画ストーリーデータセット [26] の構築を進めている．創作の初期からプロの漫画家と研究者が共作した 4 コマ漫画のデータセットは分野初である．既存漫画では難しかった以下の 3 つの観点から創作の醍醐味であるストーリー，表現，読解の多様性を考慮している．

- オチの位置の構造を変えた同プロット別シナリオ，複数作者による同シナリオ別表現の漫画を作成
- プロットなどの創作過程とパーツ別に詳細に分けたレイヤ情報を有する
- 機械学習に役立つ読者・作者双方のアノテーションを付与

4.1.2 フレームワーク

4 コマ漫画ストーリー理解のフレームワークを識別タスクの組合せとして定義した．現在提案済みの識別タスクの概要を述べる．図 1 に 4 コマ漫画ストーリーデータセットおよびフレームワークの概要を，図 2 に 5 作者の漫画を示す．

感情識別:
漫画のコマに関わる感情は大きく 2 種類に大別される．コマを見たときの読者の感情およびコマ内のキャラクタの感情である．画像情報を用いて読者がコマ全体から受ける印象として感情を識別する研究 [27]，自然言語情報を用いた多層ニューラルネットワークによるコマ内のキャラクタの感情識別 [28] に取り組んだ．4 コマ漫画ストーリーデータセットは作者がキャラクタの感情を 7 種: 基本 6 感情およびニュートラルの中のいずれかから選択してアノテートしている．

順序・構造識別:
ストーリーを追うには事象の順序が重要である．計算機が順序を識別できることは，ストーリーを理解することにつながると考え，コマの順序を識別するタスクを定義した．Convolutional Neural Network(CNN) を用いて画像情報から識別することに取組んだ [29]．また特に 4 コマ漫画ではオチの位置に基づくストーリー類型 [30] が複数あることが一般的に知られている．機械学習用に 7 種類の構造として再定義し，データセットとしてその中の 2 構造: 4 コマ目にオチがある一般（起承転結）型と 1 コマ目にオチがある出オチ型の話を持つ．特定の順のコマを対象とし，2 種の構造の識別タスク [31] を定義した．

多様度識別:
同じ状況を表す場合でも，絵を描く際の描き方は作者によって異なる．前後のコマの描き方，ジャンルに適した表現技法，作者の個性などから，多様な描き方ができることは，創作の魅力である．一方で，読者の習慣や理解力を考慮して状況を明確に伝えるためにコマによっては一定の好ましい描き方が存在する場合がある．双方のバランスの考慮は，従来，作者自身に委ねられているが，分業制の創作でコスト面を考慮して作者に状況を説明して描くことを依頼する立場にある者や，作者が独自性が入れやすいコマか読者に受け入れられやすい典型的な描き方があるかを調べたいときに，指標があれば，依頼者や作者に役立つ創作支援ができると考えた．漫画雑誌の投稿コーナーには，技術力向上を目指して，短い同じストーリーを異なるプロ作者が描いて紹介する解説記事も存在するが，一般的に流通する漫画で，同話で異なる描き方がされた作品は古典小説の漫画化などを除いてほとんどない．そのため，多様性に関わるタスクの定義は困難だったが，4 コマ漫画ストーリーデータセットでは，同じ自然言語で表されたシナリオに基づいて異なる 5 作者が漫画を描いていることから，コマの多様度識別 [28] のタスクを定義した．

上記に既発表の識別タスクを紹介したが，人が自然に描き読む漫画を人工知能のタスクから見ると他にも多数のタスクを解く必要があることがわかる．特に 4 コマ漫画の描き方に関わりが強い感情，順序・構造，多様度をまず対象としたが，ストーリーの面白さを描き理解するには文化的背景や背景知識も必要となり，漫画以外の媒体を対象とした研究も，画像情報と自然言語情報を有するマルチモーダルな媒体である漫画に適用できる可能性が高い．

4.2 小説

小説の創作方法は，工学として体系立てる前に，現場での知識や物語論に基づいて，学際融合的に醸成していることを考慮する必要がある．小説の書き方は作者によって異なるが，アイデアや設定をラフにメモをしたり，場合によっては時代考証や，参考となる人物や舞台の資料を用意した上で骨子となるプロットを書き，その後，文表現方法を工夫して実際に文章を書き，校正を重ねることが多い．筆者らの研究グループではストーリーの構成とキャラクターの設定を連携して見やすく可視化するための質問集合を定義した創作支援システム [32] を開発しており，実際にプロの創作者がシステム用いて小説を発表している [*2]．

システムはプロットを作成する前段階で用いることを想定しているが，本文を記述する間に参照したり編集することも可能であり，システムのいずれの質問や機能が，創作のどの時点で役立ったかというアンケート評価に基づきシステムの改良を進めている．また，システムへのユーザ入力文に基づき既存作品の情報を参照する機能の開発を進めている．

*2 創作者と人工知能分野研究者のコラボレーション: https://www.tut.ac.jp/docs/PR180222.pdf

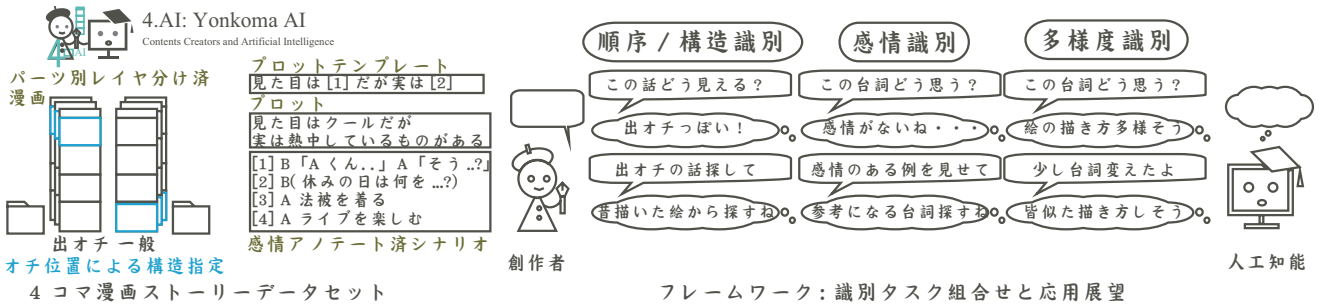


図 1: 4 コマ漫画ストーリーデータセットおよびフレームワークの概要

図 2: 4 コマ漫画ストーリーデータセット第 1，2 話，2 構造，5 作者の漫画

4.3 創作者の視点

以下に，本分野に対する創作者視点の疑問や課題を述べる．

- 創作へ役立てようと保持するあらゆる資料や記憶を見渡して，自らの内側の一部を晒して物を創り上げる．創作過程の中の大事な段階を本当に人工知能分野の研究は実現できるのか．計算機上の作業だけが創作物に関わるわけではないため，個人の感性や文化的背景を計算機上で解析することは極めて難しいのではないか．
- 人同士が創作物を見せ合い，批評し合って発展してきた創作文化ではあるが，そこに人工知能が関わるということに抵抗を感じたり，著作権や二次創作物の扱いに関する課題が存在するのではないか．

人工知能分野の手法そのものも人が生み出した創作物，ともみなせる．筆者自身，絵や小説を描くため，手法を使うか否かを決めるのは人である，という立場を取る．創作中に人に意見を求めるのは今後も続く創作文化だが，自分の過去作品や人が調べるには時間がかかる世の作品を解析して情報提示するために人工知能分野の手法を用いて共に創作する未来の可能性を上記課題を踏まえて議論したい．

4.4 創作者と人工知能の共創への展望

4.1.2 で示した手法は未だ要素技術として手法実装と一部作品への適用に留まり，多数の作品での識別率の向上や他の高次のタスクを要するため，現場応用へは距離があることをことわった上で図 1 右側に示す共創を想定し研究を進める．

5. ストーリーに関わる創作の未来

現在筆者が，特に人工知能分野と相性が良く興味深いと考えることを，ストーリーに踏み込んだ創作の未来を考える材料として示す．

作品を読んだ際，2 つの話の一部分が類似すると感じることがある．独創性の追求として類似が忌避される一方，オマージュとして敢えて類似が用いられることもあり，作者が類似作品を探すことはいずれの場合も役立つ．読者にとってはレビューサイトなどを参考にして，自らの嗜好に基づき類似した作品を読みたい場合に考慮する．ジャンル，ストーリーの構造，時代・舞台背景，登場人物の設定，伏線の張り方，文体・描画方法，読後感，など，類似には多様な観点がある．例えば，類似した文を探す方法として，近年，分散表現の台頭により，辞書に頼らなくとも表記の揺れを吸収し，文の類似を調べられる可能性が高まったが，最終成果物のみを対象とした場合に話の粒度の問題は依然として残り，ネームやプロットなどの創作過程に着目した粒度別の情報があることが研究発展に望ましい．

6. まとめ

本稿では，創作者と人工知能が創る創作の未来に寄せ，創作に役立つ可能性という観点で人工知能分野の研究を紹介した．広く見るとさらに多くの要素技術がシーズとして役立つ可能性があり，現在進んでいる研究の目的を適切に伝え，創作者側のニーズを聞き取りながら，新たな創作物，創作手法を生み出す未来の可能性を共に探りたい．

謝辞

表題の研究で関わった方々，特に本稿の 4. 章の研究に関わっていただいた創作者の方々に各原稿の謝辞に重ね，本稿でも謝意を表す．本研究は, JST, ACT-I(グランド番号:JPMJPR17U4) の支援を受けた．本研究は一部，JSPS 科研費 (グラント番号:JP17K17809) および大幸財団の支援による．

参考文献

[1] 松下光範. コミック工学の可能性. 第 2 回 ARG WEB インテリジェンスとインタラクション研究会, pp. 63–68, 2013

[2] Toru Ogawa, Atsushi Otsubo, Rei Narita, Yusuke Matsui, Toshihiko Yamasaki, Kiyoharu Aizawa, Object Detection for Comics using Manga109 Annotations, arXiv:1803.08670

[3] 佐藤 剣太, 牧 良樹, 中村 聡史, コミックの読書意欲を増進させる要素に関する分析, 人工知能学会, 1K1-OS-2a-01, 2018

[4] 三原 鉄也, 永森 光晴, 杉本 重雄, マンガメタデータフレームワークに基づくディジタルマンガのアクセスと制作の支援—ディジタル環境におけるマンガのメタデータの有効性の考察—, 電子情報通信学会論文誌 A, Vol.J98-A, No.1, pp.29-40, 2015

[5] 保住 純, 松尾 豊, 深層敵対的生成ネットワークを用いたマンガイラスト自動生成の試み, 人工知能学会, 1K2-OS-2b-03, 2018

[6] 松下 光範, 山西 良典, 松井 勇佑, 岩田 基, 上野 未貴, 西原陽子, 中村 聡史. ”私のブックマーク「コミック工学」”. 人工知能, Vol. 32, 2017

[7] Yusuke Matsui, Kota Ito, Yuji Aramaki, Azuma Fujimoto, Toru Ogawa, Toshihiko Yamasaki, Kiyoharu Aizawa, Sketch-based Manga Retrieval using Manga109 Dataset, Multimedia Tools and Applications, Springer, 2017

[8] Clément Guérin , Christophe Rigaud, Antoine Mercier, Farid Ammar-Boudjelal, Karell Bertet, Alain Bouju, Jean-Christophe Burie, Georges Louis, Jean-Marc Ogier, and Arnaud Revel. ”ebdtheque: a representative database of comics”. In *Proceedings of the 12th International Conference on Document Analysis and Recognition (ICDAR)*, 2013

[9] 松原仁, 佐藤理史, 赤石美奈, 角薫, 迎山和司, 中島秀之, 瀬名秀明, 村井源, 大塚裕子: コンピュータに星新一のようなショートショートを創作させる試み, 第 27 回人工知能学会全国大会, 2D1-1, 2013

[10] 緒方健人, 佐藤理史, 駒谷和範: 模倣と置換に基づく超短編小説の自動生成, 第 28 回人工知能学会全国大会, 1C3-OS-14b-2, 2014

[11] 秋元 泰介，小方 孝, 共創物語生成システムに向けて, 人工知能学会論文誌, vol.31, no.6, AI30-O_1-8, 2016

[12] Kiyohito Fukuda, Naoki Mori, and Keinosuke Matsumoto, A novel sentence vector generation method based on autoencoder and bi-directional LSTM, Vol. 800 of Advances in Intelligent Systems and Computing, pp. 128-135, 2018

[13] 朝倉民枝, おはなしづくりソフトの開発?子どもたちの心の中に安全基地を育てたい?, 情報処理学会, デジタルプラクティス, vol.2, no.2, pp.103-116, 2011

[14] Kiyohita Fukuda, Saya Fujino, Naoki Mori, Keinosuke Matsumoto, ”Semi-automatic Picture Book Generation Based on Story Model and Agent-Based Simulation. , Intelligent and Evolutionary Systems. Proceedings in Adaptation, Learning and Optimization, vol 8. Springer, Cham, 2016

[15] Saya Fujino, Naoki Mori, Maria Navarro Caceres, Sara Rodriguez Gonzalez, Juan Manuel Corchado, The automatic objects placement for a picture book by machine learning, 24th International Symposium on Artificial Life and Robotics, 2019

[16] Saya Fujino, Taichi Hatanaka, Naoki Mori, Keinosuke Matsumoto, The Evolutionary Deep Learning based on Deep Convolutional Neural Network for the Anime Storyboard Recognition, in Distributed Computing and Artificial Intelligence, pp. 278?285, 2017

[17] 山下 諒, 朴 炳宣, 松下 光範: コミックの内容情報に基づいた探索的な情報アクセスの支援, 人工知能学会論文誌 32(1), WII-D_1-11, 2017

[18] Tetsuya Mihara, Akira Hagiwara, Mitsuharu Nagamori, Shigeo Sugimoto, ”A Manga Creator Support Tool Based on a Manga Production Process Model ? Improving Productivity by Metadata”, In iConference 2014 Proceedings, pp. 959-963, 2014

[19] Edgar Simo-Serra, Satoshi Iizuka, Hiroshi Ishikawa, ”Real-Time Data-Driven Interactive Rough Sketch Inking”, ACM Transactions on Graphics (SIGGRAPH), 2018

[20] PaintsChainer: https://paintschainer.preferred.tech/

[21] コミ Po!: https://www.comipo.com/

[22] Live2D: https://www.live2d.com/

[23] CACANi: https://cacani.sg/

[24] 古川 健次, 仲田 晋: 3 次元キャラクタにおけるアニメ風髪モーションの自動生成, 人工知能学会, 1K2-OS-2b-02, 2018

[25] Miki Ueno, Kiyohito Fukuda and Naoki Mori, *”Can Computers Create Comics and Animations?”*, *Computational and Cognitive Approaches to Narratology*, pp. 164–190, 2016.

[26] 上野 未貴, ”創作者と人工知能: 創作者と人工知能: 共作実現に向けた創作過程とメタデータ付与 4 コマ漫画ストーリーデータセット構築”, 人工知能学会, 4Pin1-16, 2018

[27] 上野 未貴, 井佐原 均. ”漫画内の特徴的要素が与えるストーリーの印象についての検討”. 人工知能学会, No. 2J5-OS-08b-4in2, 2016

[28] Ryo Iwasaki, Koichi Sato, Miki Ueno, Naoki Mori, Makoto Okada, Novel deep learning method for understanding Japanese comics, 24th International Symposium on Artificial Life and Robotics, 2019

[29] Miki Ueno, Naoki Mori, Toshinori Suenaga, and Hitoshi Isahara. ”Estimation of structure of four-scene comics by convolutional neural networks”. In *Proceedings of the 1st International Workshop on coMics ANalysis, Processing and Understanding, MANPU@ICPR 2016*, pp. 9:1–9:6, 2016

[30] Neil Cohn, ”You ’ re a good structure, Charlie Brown: The distribution of narrative categories in comic strips”, *Cognitive Science*, Vol. 38, No. 7, pp. 1317-1359, 2014

[31] Miki Ueno, Structure Analysis on Common Plot in Four-Scene Comic Story Dataset, 25th International Conference on MultiMedia Modeling, 2019

[32] 葛井健文, 上野未貴, 井佐原均, ”質問集合とグラフに基づく物語全体の流れを管理可能な創作支援システムの提案”, 人工知能学会, 2E2-04, 2017

Organized Session | Organized Session | [OS] OS-20

[3P3-OS-20] 脳波から音声言語情報を抽出・識別する

新田 恒雄（早稲田大学/豊橋技術科学大学）、桂田 浩一（東京理科大学）、入部 百合絵（愛知県立大学）、田口 亮（名古屋工業大学）

Thu. Jun 6, 2019 1:50 PM - 3:30 PM Room P (Front-left room of 1F Exhibition hall)

[3P3-OS-20-01] Describing Brain Activity Evoked by Speech Stimuli

〇Rino Urushihara[1], Ichiro Kobayashi[1], Hiroto Yamaguchi[2,3], Tomoya Nakai[2,3], Shinji Nishimoto[2,3] （1. Ochanomizu University, 2. National Institute of Information and Communications Technology, 3. Osaka University）

1:50 PM - 2:10 PM

[3P3-OS-20-02] Development of Syllable Labelling Tool for Electroencephalogram Data

〇Mingchuan Fu[1], Ryo Taguchi[1], Kentaro Fukai[2], Kouichi Katsurada[2], Tsuneo Nitta[3,4] （1. Nagoya Institute of Technology, 2. Tokyo University of Science, 3. Waseda University, 4. Toyohashi University of Technology）

2:10 PM - 2:30 PM

[3P3-OS-20-03] Word Recognition from speech-imagery EEG

〇Satoka Hirata[1], Yurie Iribe[1], Kentaro Fukai[2], Kouichi Katsurada[2], Tsuneo Nitta[3,4] （1. Aichi Prefectural Univ., 2. Tokyo Univ. of Science, 3. Waseda Univ., 4. Toyohashi Univ. of Tech.）

2:30 PM - 2:50 PM

[3P3-OS-20-04] Syllable recognition in speech-imagery EEG

〇Kentaro Fukai[1], Hidefumi Ohmura[1], Kouichi Katsurada[1], Tsuneo Nitta[2,3] （1. Tokyo University of Science, 2. Waseda University, 3. Toyohashi University of Technology）

2:50 PM - 3:10 PM

[3P3-OS-20-05] BCI Studies on Extraction of Spoken Language representation from Speech Imagery EEG

〇Tsuneo Nitta[1], KENTAROU FUKAI[2], KOUICHI KATSURADA[2], YURIE IRIBE[3], RYOU TAGUCHI[4], SHUNJI SUGIMOTO[5], JUNSEI HORIKAWA[5] （1. Waseda University, 2. Tokyo Science University, 3. Aichi Prefectural University, 4. Nagoya Institute of Technology, 5. Toyohashi University of Technology）

3:10 PM - 3:30 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

音声刺激による脳活動の言語解読への取り組み

Describing Brain Activity Evoked by Speech Stimuli

漆原 理乃 *1 Rino Urushihara　小林 一郎 *1 Ichiro Kobayashi　山口 裕人 *2*3 Hiroto Yamaguchi　中井 智也 *2*3 Tomoya Nakai　西本 伸志 *2*3 Shinji Nishimoto

*1お茶の水女子大学 Ochanomizu University　*2情報通信研究機構 脳情報通信融合研究センター National Institute of Information and Communications Technology　*3大阪大学 Osaka University

The analysis of semantic activities in the human brain is an area of active field of study. In this paper, we propose a deep learning method to describe text for semantic representations evoked by speech stimuli from Functional Magnetic Resonance Imaging (fMRI) brain data. Thereby, our study aims to decode higher order perception which a person recalled in the brain by speech stimuli. However, collecting a large-scale brain activity dataset is difficult because observing brain activity data with fMRI is expensive, although a method with deep learning requires a large-scale dataset. We therefore use an automatic speech recognition method and utilize a small amounts of fMRI data efficiently for machine learning. Through experiments, we have conformed high correlation between the predicted features from fMRI data and the speech features.

1. はじめに

近年, 脳神経活動の意味表象を捉える研究が盛んになっている. 本研究では, Functional Magnetic Resonance Imaging (fMRI) で観測した音声刺激下の脳活動データから, 人が脳内に想起した高次意味表象を言語として解読することを目指し, 深層学習を用いて, 音声刺激による脳活動データからその意味表象をテキストとして生成する手法を構築する. しかし, fMRI により観測する脳活動データは取得のためのコストが大きく, 大量の学習データを要する深層学習を十分に行うための大規模なデータ収集は困難である. そのため, 学習済みの自動音声認識を行う深層学習モデルの中間層に脳活動データを回帰することで, 直接 fMRI データを用いてモデル全体を学習せず, 少量の fMRI データを効率的に活用し, 脳活動データからキャプションを生成する手法を提案する.

2. 関連研究

近年, 脳神経活動の多点計測技術の発展と機械学習技術の高度化により, ヒト脳内情報表現の定量理解や解読を目指す研究が盛んになっている [Huth 12, Stansbury 13]. しかし, これらは動画像もしくは静止画像視聴下における脳神経活動を対象としており, 音声 (言語) を刺激とした脳神経活動を対象とした研究は少ない [Huth 16]. そのため, 本研究では, 音声刺激を対象とした脳神経活動を解読することを目的とする. 解読手法構築にあたって, 深層学習を援用することで, 動画像視聴下における脳活動から認知内容の解読が実現できることが示されている [Matsuo 18]. 本研究では, このような背景から深層学習を用いるが, その際, Matsuo ら [Matsuo 18] の手法の少量の脳活動データの効率的な利活用方法を参考にし, 自動音声認識手法を援用することで, 音声刺激下の脳活動データから, その刺激となっていた音声のテキストを生成する手法を構築し, 脳内意味表象の解読を目指す.

連絡先: 漆原理乃, お茶の水女子大学理学部情報科学科小林研究室, 〒 112-8610 東京都文京区大塚 2-1-1, g1420509@is.ocha.ac.jp

3. 提案手法

本提案手法は, 深層学習を用いて, 音声刺激を受けた脳活動データを入力として, その時に刺激となっていた音声のテキストを生成し, 人が頭の中で想起した言葉に対応する意味表象を言葉として解読することを目指す. しかし, fMRI により観測する脳活動データは取得のためのコストが大きく, 大量の学習データを要する深層学習を十分に行うための大規模なデータ収集は困難である. そのため, Encoder-Decoder Network に基づく自動音声認識手法を援用することで少量データを効率的に活用する. 具体的には自動音声認識の Encoder から得られる中間表現に, 脳活動データを回帰させ, その結果を自動音声認識の Decoder に入力することで, テキスト生成を行う. 図 1 に本提案手法の概要図, 3.1 節に提案手法の処理の流れを示す.

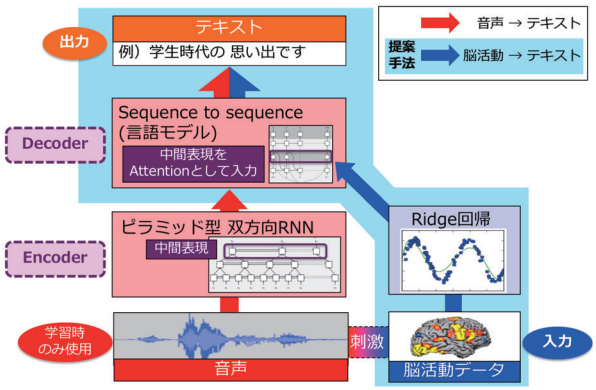


図 1: 本研究の概要図

3.1 提案手法の処理の流れ

実行時の処理は以下のようになる.

step 1. 自動音声認識

step 1-1. Encoder: 音声中間表現の抽出
自動音声認識の Encoder を用いて, 音声から音声中間表現を抽出.

step 1-2. Decoder: テキスト生成
step1-1. において抽出された音声中間表現を, 自動音声認識の Decoder に入力し, テキストを生成.

step 2. 脳活動データの特徴量推定
脳活動データとその刺激である音声の中間表現 (step1-1 の出力) との対応関係を学習した Ridge 回帰により, 脳活動データから対応する中間表現を推定.

step 3. 脳活動データの特徴量からテキスト生成
step1-2. で学習済みの自動音声認識の Decoder を用いて, step2. で計算された脳活動データの特徴量を入力として, テキストを生成.

3.2 自動音声認識

本手法の基盤として, Encoder-Decoder Network を用いた自動音声認識モデル, Listen, Attend and Spell (LAS) [Chan 16] を用いる. LAS は Encoder として, ピラミッド型の双方向 Recurrent Neural Network (RNN) を用い, 入力音声を中間表現に変換する. Decoder としては, Attention 付きの単方向 RNN を用い, Encoder で得た中間表現を Attention として入力し, 通常の言語モデル同様に次にくるであろう単語, もしくは文字を 1 つずつ生成する. Decoder で Attention を用いることで, 中間表現の系列と出力系列との対応関係も学習できる.

3.3 脳活動データの特徴量推定

音声刺激を受けた被験者の脳活動データを入力とし, その時の刺激となっている音声の LAS の Encoder によって抽出された中間表現を予測するために, Ridge 回帰を用いる. 学習には脳活動データと被験者がその時に聴いている音声を用いる. また, fMRI は脳活動を記録する際にタイムラグがある. 今回はそれを (1)4 秒のみ, (2)4,5,6 秒の複数時点としたものをそれぞれ説明変数とする, Ridge 回帰のモデルを構築する.

4. 実験

3.1 節に示す提案手法の処理の流れに沿って行った 3 つの実験を以下に示す.

4.1 実験 1：自動音声認識

4.1.1 実験設定

システムの実装は, 深層学習のフレームワーク TensorFlow を用いたコード *1 を使用した. 学習のためのデータセットとして,「日本語話し言葉コーパス」(CSJ) 中の 3,254 本の講演データを使用した. CSJ で設定されている評価セット 1 から 3 と脳活動データの刺激として使用された音声は除外した. 音声の前処理として, 転記基本単位 (IPU) で分割し, フレームサイズ 25ms, フレームシフト 10ms でフレームごとに MFCC 特徴量を取得し入力とした. 出力は発音形 (PhoneticTranscription) を用い, 発音されている文字に start-of-sentence (sos) と end-of-sentence (eos), バッチサイズ中の系列長に合わせるための padding(pad) を追加し, 84 次元とした. 学習に関する詳細設定は表 1 に示す. 評価は CSJ の評価セット 1 の 10 本の講演データを用いて行った.

4.1.2 実験結果

epoch ごとに学習用データの Loss を記録し, その減少により学習の進度を確認し, 収束するまで学習を行った. 図 2 には, 先行研究 [Chan 16] と同様に取得した Decoder における Attention の値の可視化結果を示す. また, 評価セット 1 で評価を行い, 生成したテキストの一部と, 生成テキストと正解テキストとをそれぞれ比較し編集距離のマクロ平均を計算したものを表 2 に示す.

*1 https://github.com/thomasschmied/Speech_Recognition_with_Tensorflow

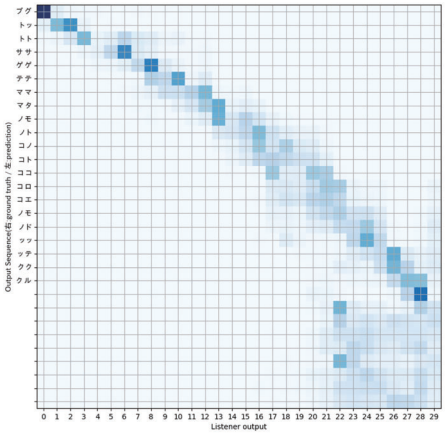


図 2: 実験 1 attention の可視化

4.1.3 考察

図 2 より, Attention がうまく機能している, つまり中間表現の各要素が, 出力系列のどの部分に対応するかを学習することができていることがわかる. しかし, 表 2 より, 短いテキストは生成に成功しているが, 長いテキストは正確な生成が困難であるといえる.

4.2 実験 2：脳活動データの特徴量推定

4.2.1 実験設定

脳活動と音声特徴量の対応関係を学習するためのデータセットとして, 日本語話し言葉コーパス (CSJ) の 16 本を 1 人の被験者に聴かせた時の血中酸素濃度依存性信号 (BOLD 信号;Blood Oxygenation Level Dependent Signal) を fMRI を用いて 1 秒ごとに記録した脳活動データ, および fMRI のデータ収集と同期させた CSJ の音声を使用する. 立体撮像 96 × 96 × 72 ボクセルのうち皮質に相当する 62,552 次元分のデータ列を用い, その時聴いている音声から LAS の Encoder であるピラミッド型の双方向 RNN から出力される 900 次元の音声中間表現との対応と, 1800 次元の RNN の隠れ層 (Long short-term memory(LSTM) の隠れ層, 1,2 層目のセル c,h で共に 450 次元により合計 1800 次元. Decoder では, この隠れ層を初期値として予測を開始する) との対応を学習した. 脳活動データは, 1 秒ごとに取得されているため, 系列長を音声中間表現の系列数に合わせるように線形補間を行い, 最大値 1, 最小値 0 に正規化を行ったものを入力として使用した. train 用データは 14 本, test 用データは 2 本の講演データを聴いた脳活動データとした. 学習には Ridge 回帰を用い, その詳細設定は表 1 に示す.

4.2.2 実験結果

test 用データの脳活動データから Ridge 回帰を行い, その刺激となっていた音声の中間表現を予測した結果とその音声の中間表現との相関係数を計算したものを表 3 に示す.

4.2.3 考察

表 3 より, Encoder の出力に関して, 説明変数となるタイムラグを 4 秒のみとした時よりも, 4,5,6 秒とした時の方が相関係数が低いことが言える. これに関しては, 脳活動データを音声中間表現の系列数に合わせるように線形補間を行い, それを

表 1: 各パラメータ設定（詳細）

	自動音声認識	脳活動データの特徴量抽出
train データ	日本語話し言葉コーパス (CSJ)	音声刺激による脳活動データ
学習量	919,118 sample × 70 epochs	100,938 sample (9,841sample を元に線形補間し増強)
アルゴリズム	Adam	Ridge 回帰
学習に関する パラメータ	学習率: 0.00001	L2 正則化項: 1.0
隠れ層次元	入力：494 次元 Encoder 2 層：全て 450 次元 Decoder 2 層：全て 450 次元 Embedding：10 次元 出力：84 次元	Attention となる中間表現: 62,552 次元 - 900 次元 RNN の隠れ層: 62,552 次元 - 1800 次元
誤差関数	交差エントロピー	
その他	Cyclic Learning Rate Max 学習率：0.00003 Step size: 700	

表 2: 評価セット 1 における音声認識実験結果

正解テキスト	生成テキスト
コンカイノ	コンカイノ
トユーブンデハ	トユユブブンンワワ
コノヨーニ マチガウカノーセーガ アリマスノデ	コノヨーニニ チチーカカノノーー ガガアリママスデデ
編集距離のマクロ平均: 11.3 (平均文字数: 21.8)	

表 3: 実験 2 の ridge 回帰結果 (相関係数)

タイムラグ	(1)4 秒	(2)4,5,6 秒
Ecoder の出力	0.32	0.11
LSTM の隠れ層 (c)	0.78	0.78
LSTM の隠れ層 (h)	0.42	0.42

表 4: テキスト生成実験結果 (編集距離のマクロ平均)

脳活動からの生成 (タイムラグ 4 秒)	6.0
脳活動からの生成 (タイムラグ 4,5,6 秒)	12.7
音声からの生成	3.1
平均文字数	6.6

表 5: テキスト生成実験結果 (生成例)

正解	脳活動から生成 (4 秒)	脳活動から生成 (4,5,6 秒)	音声から生成
テーマワ	エーーー	ウーカカカカ	テーマワ
アルト	デテ	デデンン	アルト
マダ	イイ	ネネ	マダ
ヤッパリ	ノノ	エエ	ヤッパリ

ridge 回帰の入力に使用しているため, 説明変数にノイズが多く含まれてしまっていることが考えられる. そのため, 今後は線形補間を用いないで予測するモデルを考案していきたい.

4.3 実験 3：脳活動からのテキスト生成

実験 1 で学習した自動音声認識モデル LAS の Decoder を使用して, 実験 2(1)(2) それぞれで取得した脳活動データの特徴量を入力として, テキスト生成を行う.

4.3.1 実験設定

実験 2 の学習において未使用の test 用の 2 本の講演データを聴いた脳活動データを用いて, Ridge 回帰を行い, 900 次元の中間表現と RNN の隠れ層 1800 次元を取得し, LAS の Decoder に前者の中間表現は Attention として入力し, 後者は RNN の隠れ層の初期値として使用した.

4.3.2 実験結果

脳活動データから生成したテキストと音声から生成したテキスト, 正解テキストをそれぞれ比較し編集距離のマクロ平均を計算したものを表 4 に示す. また, 生成したテキストの一部を表 5 に示す.

4.3.3 考察

表 4 より脳活動データから生成したテキストの編集距離のマクロ平均が, タイムラグを 4 秒とした時よりも 4,5,6 秒とした時の方が低くなっている. その原因としては, 4.2 節の実験における, 表 3 の相関係数が後者の方が低いためといえる. 表 5 より, タイムラグを 4 秒, 4,5,6 秒とした時の両方とも, 実際の生成テキスト例も正解もしくは音声から生成したテキストと一致しておらず, 生成がうまくいっていないことがわかる. 4.2 節の実験における, 脳活動データの特徴量の推定値の誤差が影響していると考えられる.

5. おわりに

本稿では, 自動音声認識モデル LAS を援用し, 音声刺激を受けた被験者の脳活動データから聴いている音声をテキストとして出力する手法を提案した. 今後の課題として, 実験 2 の脳活動データから特徴量を推定する手法を見直し, 実験結果のさらなる評価・分析を実施し, 生成するテキストの精度を向上させていきたい.

謝辞

本研究を進めるにあたっては, その一部を科研費新学術領域研究（課題番号：18H05118）の支援を受けた.

参考文献

[Chan 16] Chan, W., Jaitly, N., Le, Q., and Vinyals, O. Listen, attend and spell: A neural network for large vo-

cabulary conversational speech recognition. In Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on (pp. 4960-4964). IEEE (2016).

[Huth 12] Huth, A. G., Nishimoto, S., Vu, A. T., Gallant, J. L.: A continuous semantic space describes the representation of thousands of object and action categories across the human brain, Neuron, 76(6):1210-1224 (2012).

[Huth 16] Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., and Gallant, J. L. Natural speech reveals the semantic maps that tile human cerebral cortex. Nature, 532(7600), 453. (2016)

[Matsuo 18] Matsuo, E., Kobayashi, I., Nishimoto, S., Nishida, S., and Asoh, H. Describing Semantic Representations of Brain Activity Evoked by Visual Stimuli. In arXiv preprint arXiv:1802.02210 (2018).

[Stansbury 13] Stansbury, D. E., Naselaris, T., Gallant, J. L.: Natural Scene Statistics Account for the Representation of Scene Categories in Human Visual Cortex, Neuron 79, pp.1025-1034, September 4, 2013, Elsevier Inc (2013).

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

脳波データ用音節ラベリングツールの開発

Development of Syllable Labelling Tool for Electroencephalogram Data

付 明川[*1] Mingchua Fu　田口 亮[*1] Ryo Taguchi　深井 健大郎[*2] Kentaro Fukai　桂田 浩一[*2] Kouichi Katsurada　新田 恒雄[*3, 4] Tsuneo Nitta

[*1] 名古屋工業大学 Nagoya Institute of Technology　[*2] 東京理科大学 Tokyo University of Science　[*3] 早稲田大学 Waseda University　[*4] 豊橋技術科学大学 Toyohashi University of Technology

EEG (Electroencephalogram) is an electrical signal representing activity of the brain and have been used for healthcare and brain-machine interface. Recently, research to estimate imagined linguistic information from EEG signals was launched. The research needs labeled EEG dataset that are given boundaries of imagined syllables. In this paper, we propose the syllable labeling tool for research on EEG. Labelers can adjust boundaries of each syllable using a mouse or a keyboard while observing features extracted from EEG signals. They can also easily reuse analytical methods developed by themselves because this tool runs on MATLAB. Moreover, in this paper, we describe a semi-automatic labeling method to improve operating efficiency.

1. はじめに

脳波(Electroencephalogram :EEG)は，非侵襲で比較的安価に計測可能な脳信号であり，医療での臨床検査やリハビリテーションなどで利用されている[1]．また，集中度や快不快などの感情の推定が可能であり，ブレイン・マシン・インターフェース(BCI)への応用が進められている．さらに近年では，想起した言語情報を脳波からデコードする研究が始まっている[2]．この研究が実現すると，頭の中で想起した内容を機械が読み取ることができ，従来の BCI よりも利用範囲が大きく広がる．

音声認識では，発話した言語内容と音響信号との対応を統計的に学習することで実現している．同様に，脳波から言語情報をデコードするには，想起した音節系列と脳波の特徴系列間の対応をコンピュータに学習させる必要がある．音声認識技術の発展には，(1)音響モデルを学習するための音素ラベル付の音声データベースと，(2)言語モデルを学習するためのテキストコーパスの二つが大きく貢献した．音声想起時の脳波による音節認識では，(2)の言語モデルはそのまま利用可能であるが，(1)については音節ラベル付の脳波データベースは存在しないため，人手で作成する必要がある．しかし，脳波データから音節をラベリング可能なツールは現在，存在しない．

本報告は，言語想起に基づく脳波研究のための音節ラベリングツール開発を目指している．脳波からの音声言語デコーディングは，まだ始まったばかりであり，特徴抽出器や識別器を設計するためのノウハウ蓄積もこれからという段階である．従って，(1)特徴抽出および識別手法の検討，(2)EEG データに対する音節ラベリング，(3)認識シミュレーション実験による検証，という三つの工程を繰り返しながら，より良い手法を模索していく必要がある．本ツールは，この試行錯誤を効率化することを目的としている．ツール開発にあたっては，脳波研究者の間で広く利用される MATLAB 上での動作を前提とした．以下では，Hidden Markov Model(HMM)の音節アライメントを用いた半自動ラベリングの実現可能性についても検討する．

2. ラベリングツールの概要

2.1 開発環境

本ラベリングツールは MATLAB の GUI 開発環境である GUIDE を用いて開発している．MATLAB は，MathWorks 社が提供する数値解析ソフトウェアである．開発した脳波の解析や可視化が可能な EEGLAB や，信号処理，機械学習などの

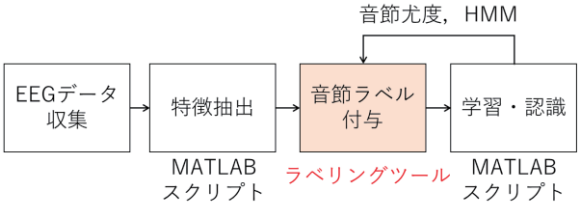


図 1　ラベリングツールを用いた作業の流れ

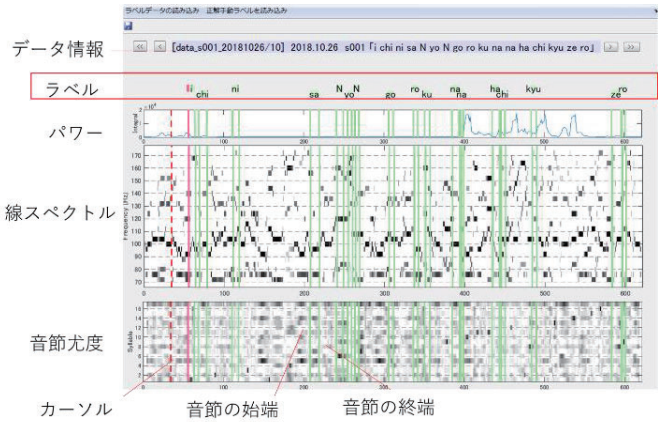


図 2　ラベリングツールの実行画面

toolbox が豊富に揃っていることから，音声研究者や脳波研究者の間で広く用いられている．MATLAB 上で動作するラベリングツールを開発することにより，各研究者がこれまでに開発した特徴抽出・機械学習の手法をシームレスに利用することができ，研究開発が効率化すると考えている．

2.2 基本機能

図 1 を用いてツール利用の流れを説明する．まず，収集した脳波データから，音節ラベリングの際に用いる特徴量を抽出する．特徴量としては，ノイズ除去後の波形や，信号強度，周波数スペクトルなどが考えられるが，本ツールでは，「脳波中の言語表象は tone-burst 波すなわち線スペクトル群として現れる」との仮説に基づき，線スペクトル特徴と音節尤度ベクトル系列を用いる[3,4]．線スペクトル特徴は，音節想起時に現れる周波数成分を元に，言語表象と成り得る特徴として抽出されたものである．音節尤度ベクトルは，目視で音節ラベリングしたデータから学習

した音節認識器を用いて，音声想起時の脳波を認識した際の音節毎のスコアである．音節尤度ベクトルは線スペクトルと異なり，音節らしさを直接表現する特徴量であるため，ラベリングの大きな助けになると考えられる．しかし，ラベリングデータが全く無い状態では使用することができない．そのため，最初は線スペクトル特徴に基づいてラベリングを行い，ある程度データを集積した段階で，音節尤度を利用するという工程を想定している．後述するHMMを用いた半自動ラベリングも同様のコンセプトであるが，HMM は音節区間を推定出来る点が異なる．線スペクトルと音節尤度の計算方法については[4]を参照されたい．

図 2 に開発したラベリングツールの実行画面を示す．画面上部にデータのファイル名や収録日，想起内容が表示される．また横に示すボタンでデータの切り替えを行う．その下に，ラベリングする音節の区間が表示される．画面中央に線スペクトル，下部に音節尤度のグラフが表示され，各グラフの内部に音節境界と現在のカーソル位置を表すラインが表示される．データ読み込み時に，ラベルデータが存在する場合，それが初期の境界となる．音節境界は，マウスまたはキーボードを用いて位置を微調整する．ラベリング結果は，データの切り替え時に自動保存される．保存形式は，MATLAB で容易に利用可能な mat 形式の他，Hidden Markov Model Toolkit (HTK)[5]のラベル形式でも保存できる．HTK は音声認識に用いられる HMM の学習，認識ができるソフトウェアである．HMM は時系列データの統計モデルであり，脳波の解析にも有効であると考えられる．

2.3 HMMを用いた半自動ラベリング

EEG 信号は音声と異なり，「自分の耳で聞いて，(想起)音節区間を判断する」ことができない．そのため，ラベリング作業者はEEG信号の周波数スペクトル情報を目視で観察し，習得した各音節が共通に持つ特徴に着目してラベリングすることになる．この作業を効率化するため，我々は HMM による音節アライメントを用いた半自動ラベリング機能を開発している．

学習済みの音節 HMMを用いると，脳波データと音節系列情報のみが与えられ，個々の音節境界が不明な場合にも，音節境界を推定することができる．推定された音節境界をラベリングツール上に初期境界として表示することで，作業効率化を図る．また，確認済のラベリングデータと HMM による音節境界を重ねて比較表示することにより，ラベリングの妥当性を検証できる．

HMM の特徴量には，先の線スペクトルや音節尤度の時系列データを利用する．今回の実験では，連続想起の 10 数字（1-9, 0）のを 10 サンプル用い，音節 HMM を 5 状態 3 ループ，混合数 3，音節以外の区間（以下，ノイズ区間）の HMM を 3 状態 1 ループ，混合数 9 で学習した．特徴量には線スペクトル 28 次元を用いた．学習に用いていないデータの音節境界を推定した結果の一例を図 3 に示す．図 3 の青い線が人手で設定した境界位置，緑の線が HMM により推定された結果である．この結果は，想起データに現れる最初の音節の始端と，最後の音節の終端を人手で与えており，期待する性能はまだ得られていない．今後はノイズ区間のモデル化手法を検討することで，性能の向上を目指すと共に，半自動ラベリングに必要なデータ数についても調査していく予定である．

3. 今後の展望

本稿で紹介したラベリングツールは，1画面で1つのデータしか扱うことができない．実際のラベリングの作業では，紙に大きく印刷したスペクトログラムを机にならべ，データ間で相互に見比べることで，共通の特徴を探すことが行われる．このような，他のデータとの比較を容易にする機能をラベリングツールに追加していきたいと考えている．

また，収集した脳波データが増えるにつれ，その管理が問題となる．そこで，収録情報をデータに付与するための機能や，実験条件に合わせて学習・評価用のデータを切り出し，ラベルデータと共に保存，管理，検索する機能を追加する．

4. まとめ

本稿では，脳波から言語情報をデコードする研究に必要な音節ラベリングツールの開発状況を報告した．ラベリングやその後の検証作業を容易にするため，ツールは MATLAB で開発すると共に，HMM の音節アライメントを用いた半自動ラベリングの検討を行った．現在は，想起データに現れる最初の音節の始端と，最後の音節の終端を人手で与えており，期待する性能はまだ得られていないため，今後，アライメント精度を向上していく必要がある．また，データ間の比較や，データの管理機能についても充実させていく予定である．本ラベリングツールは，特徴抽出手法および半自動ラベリング手法が完成した後，公開する予定である．

参考文献

[1] Shindo, et al: Effects of neurofeedback training with an electroencephalogram-based brain-computer interface for hand paralysis in patients with chronic stroke: a preliminary case series study. J Rehabil Med., 43, pp. 951--957, 2011.

[2] 神崎，他：発話時と想起時の脳波による日本語短音節認識の比較，日本音響学会 2017 年春季研究発表会，2017.

[3] 新田，他: 音声想起時脳波からの音声言語表象抽出に基づくBCI 研究，第 33 回 人工知能学会全国大会，2019.

[4] 深井，他：音声想起時脳波中の音節識別について，第 33 回 人工知能学会全国大会，2019.

[5] http://htk.eng.cam.ac.uk/

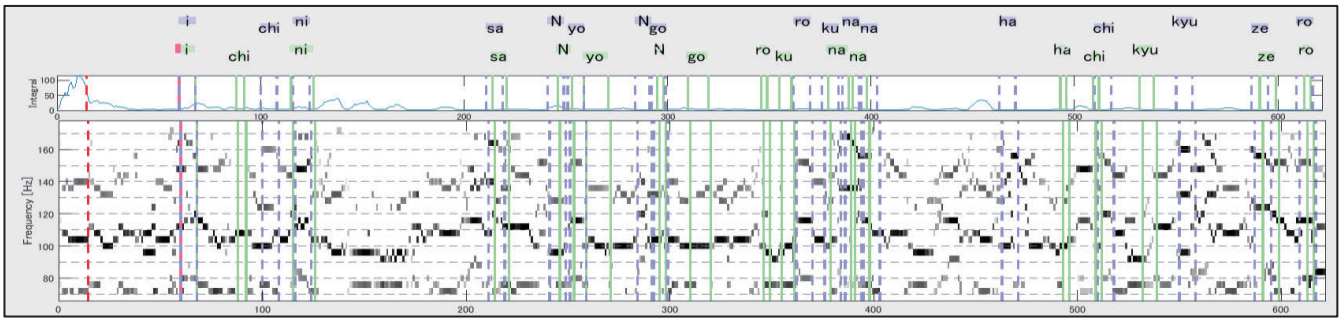


図 3　HMM の推定結果と人手によるラベルデータの比較

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

音声想起時脳波からの単語認識
Word Recognition from speech-imagery EEG

平田 里佳[*1] 入部 百合絵[*1] 深井 健大郎[*2] 桂田 浩一[*2] 新田 恒雄[*3, 4]
Satoka Hirata Yurie Iribe Kentaro Fukai Kouichi Katsurada Tsuneo Nitta
Aichi Prefectual Univ. #1 Tokyo Univ. of Science #2 Waseda Univ.#3 Toyohashi Univ. of Tech.#4

Previous research suggests that humans manipulate the machine using their electroencephalogram called BCI (Brain Computer Interface). However, there are not existed effective methods for speech imagery recognition. In this report, we propose the word recognition method using line spectra extracted from EEG signal of continuously speech imagery. The word recognition can be achieved based on syllable-HMM model constructed by line spectra of fourteen syllables(/i/,/chi/, /ni/,/sa/,/N/,/yo/,/go/,/ro/,/ku/,/na/,/ha/,/kyu/,/u/,/ze/). The word recognition of ten numbers was conducted in our experiment.

1. はじめに

近年，脳波を用いて機械を操作する BCI（Brain Computer Interface）の研究が盛んに行われ，主に医療，介護現場等での活用が期待されている．

先行研究では，発話時の ECoG（Electro Cortico Graphy）の観測から感覚運動野の音節による活性部位の違いが指摘されている[1]．同じく発話時 ECoG で，側頭葉の領野により周波数パターンは大きく異なるという観測結果が報告されている[1]．一方，発話時は上側頭回での言語プランニング結果をブローカ野が取次ぎ，運動野で調音指令が出て発音運動に繋がる[3]．これらの研究は，頭蓋内で皮質から直接信号を取得するため，身体への負担が大きく，BCI での利用は現実的でない．

本報告では，利用者への負担が少ない非侵襲的な方法で，頭皮から脳波（Electro-encephalogram ; EEG）信号を取得し，脳活動を分析する．また分析結果をもとに，音声想起(speech imagery) 時の脳波から単語を認識することを目指す．

脳波からの単語認識研究は，現状非常に少なく，対象単語数も 2･3 単語と少ない報告例が多い[4]．以下では，1 名分の脳波を用いた 10 数字単語想起に対する実験結果を報告する．

2. 脳波信号の収録

成人男性 1 名(右利き) が 10 数字を想起した際の脳波を採取した．電極は，左脳ブローカ野周辺に 9 電極を配置した．図 1 に拡張 10-20 法で示した使用電極を示した(赤い四角で囲った電極)．

10 数字は連続して想起した(ichi→ni→saN→yoN→go→roku→nana→hachi→ kyuu→zero)．収録プロトコルを図 2 に示す．被験者は純音が鳴ってから，約 2 秒後に想起を開始し，約 3 秒間内で 10 数字を連続想起する．質の良い脳波信号を取得するため，このプロセスを 23 回続けて繰り返した． 以下の実験では，孤立単語認識としての評価を行うため，個々の数字の単語境界

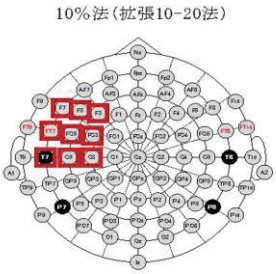


図 1 電極配置

連絡先：平田里佳，愛知県立大学，im182010@cis.aichi-pu.ac.jp

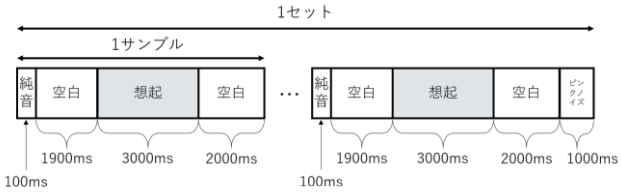


図 2 収録プロトコル

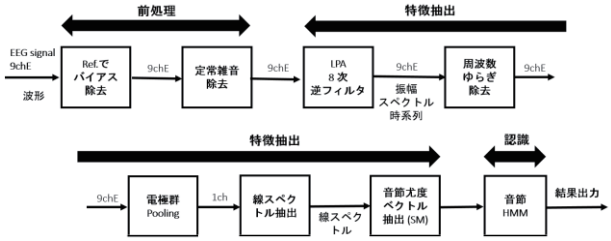


図 3 音声想起単語認識システム

を目視で与えた試料を 10 サンプル用意して使用した(孤立単語としては 10×10 個)．

3. 音声想起単語認識システム

提案する想起単語認識システムのブロック図を図 3 に示す．

3.1 前処理

収録脳波データに含まれる電源ノイズや余分なアーチファクトを取り除くため，ノイズサブトラクション処理を行う．この処理では，最初に想起時以外の区間の脳波を振幅スペクトルに変換した後，8 フレームを平均してノイズ区間平均振幅スペクトラムとする(N(k))．続いて，電極ごとに想起時脳波の振幅スペクトル X(k) から N(k)を引き去る．これにより，高調波成分を含む電源周波数成分，および定常的に発生するノイズを除去する．

また，音声言語に関する ECoG の研究[2]を参考に，High-γ 帯の脳波信号，すなわち 70-200Hz の Band-pass filter (BPF)を通した成分を対象とする．開閉眼や唾呑み込みなどの筋活動や体動によるアーチファクトの多くは BPF により除くことができたが，歯ぎしり等は想起時脳波への無視できないノイズとして残る．

3.2 特徴抽出

我々の音声想起に関する研究では，「脳波中の言語表象は tone-burst 波すなわち線スペクトル群として現れる」と考えている[5]． この仮説に沿って線スペクトル群を抽出するため，LPA(Linear Predictive Analysis; 線形予測分析) を用いて，振幅スペクトルを求めた後，正弦波に相当する線スペクトルを求める[6]．線スペクトル抽出では，周波数のゆらぎが観測されたた

め，まず周波数ごとに前後 3 フレームの計 7 フレームに対して median filter を掛けて中央値を計算した後，Gaussian filter で平滑処理を行い，周波数ゆらぎを除去した．続いて，9 電極の出力を集約する pooling 処理を施した．これは，音節情報を観測する度に，情報の出現する電極が異なる現象が観られた，ブローカ野周辺の電極群を集約して安定に音節情報を抽出する解決策として考案した．具体的には，9 電極のスペクトラムの要素毎に，40ms の範囲で 4-norm を計算した．最後に，集約したスペクトル包絡から微分演算により複数のピーク周波数を抽出して線スペクトル成分とする．

音節尤度ベクトルを抽出して用いた．このため，まず 10 数字を連続して想起した脳波データの線スペクトル系列から，目視により音節ラベリングを行い，含まれる音節区間を切り出した．ラベリングの例を図 4 に示す．次に，切り出した短音節区間 28ch.×9 フレームに対して音節毎に固有ベクトルを計算する．なお，同時に音節と音節の間の区間をノイズ区間として，全てのノイズ区間データから固有ベクトルを計算した．音節尤度ベクトル系列は，10 数字連続想起データに対して部分空間法により 16 短音節およびノイズの 17 クラスに対する音節尤度を計算し，これを尤度ベクトル系列とする．

3.3 単語認識

音節 HMM（Hidden Markov Model）を構築し，10 数字の単語認識を行う．HMM でモデル化する音節は 10 数字に出現する 14 音節（/i/, /chi/, /ni/, /sa/, /N/, /yo/, /go/, /ro/, /ku/, /na/, /ha/, /kyu/, /u/, /ze/）を対象とする．これに加えて，短音節間に出現するノイズについても HMM 学習する．学習に使用する特徴量は線スペクトル，音節尤度ベクトルである．これらの特徴としての比較評価を行うため，線スペクトルと音節尤度ベクトルを組み合わせたものに加え，線スペクトルのみ，音節尤度ベクトルのみによる計 3 種類の特徴パターンを用いた実験を行った． 予備実験の結果より，構築した HMM は 5 状態 3 ループの left-to-right 型，混合数は 3 混合とした．ただし，音節間に現れるノイズ部分に関しては，3 状態 1 ループの left-to-right 型，10 混合とした．

4. 評価実験

音節 HMM を用いた 10 数字の単語認識を leave-one-out で評価した．

4.1 実験試料

収録した脳波から線スペクトルを抽出し，線スペクトルパターンをもとに人手で音節単位のラベリングを行った．連続数字想起の 23 サンプルのうち，実際に音節ラベリングの作業を行った 10 サンプルに対して，10 数字単語区間を抽出し，10 単語×10 サンプルの 100 データを用いて，単語認識のための学習・評価を行う．leave-one-out による交差検証では，10 サンプルの 1 つを評価として除き，他の 9 サンプル中の 10×9 単語データを学習データとする．これを 10 回繰り返して評価結果をまとめた．

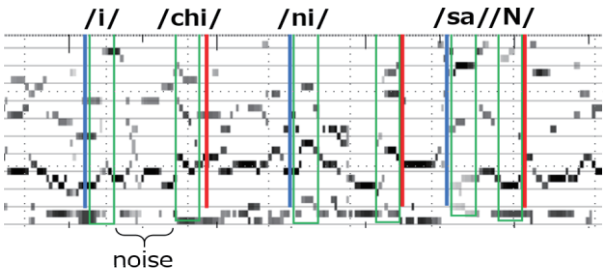


図 4 線スペクトルパターンによるラベリング

4.2 実験結果

音節 HMM を用いた 10 数字の単語認識結果を表 1 に示す．抽出した線スペクトルをそのまま特徴量として用いた場合，線スペクトルのみでは 1best，2best，3best で正解率がそれぞれ 15.0%，21.0%，33.0%であり，線スペクトルと音節尤度ベクトルを組み合わせたものでは，正解率がそれぞれ 16.0%，23.0%，28.0%であった．そこで，線スペクトルのスパース性を考慮し，フレーム方向には 3 フレーム平滑を行い，周波数方向には 3 つの周波数帯域の平均を 1 つの周波数帯域にまとめた．この処理を行った結果が表 1 の線スペクトルと線スペクトルと音節尤度ベクトルを組み合わせたものの結果である．ただし，線スペクトルの結果についてのみ，leave-one-out による交差検証の際に，特徴量の次元数が少ないために HMM が構築できない場合があった．

表 1　10 数字の単語認識結果

特徴量	正解率(%)		
	1best	2best	3best
線スペクトル	15.0	23.8	30.0
音節尤度ベクトル	54.0	69.0	80.0
線スペクトル ＋ 音節尤度ベクトル	40.0	56.0	68.0

表 1 より，音節尤度ベクトルを用いた場合の 3best の正解率が 80.0%と最も高かった．この結果から，想起時脳波信号を用いた単語認識には，特徴量として音節尤度ベクトルを用いることが効果的であることが示された．しかし，本研究では被験者 1 名分の脳波信号のみを用いているため，今後は被験者を増やし，より多くの脳波信号を用いた認識実験を行う必要がある．

5. おわりに

想起時脳波信号を用いた音節 HMM による 10 数字単語認識を行った．脳波収録では男性 1 名が 10 数字を連続して想起した際の脳波を取得した．

特徴抽出の前に収録した脳波データに含まれるノイズの除去を行った．その後，LPA によって振幅スペクトルを求め，周波数ゆらぎの除去，電極群の pooling を行い，線スペクトルを抽出した．最後に，音節尤度ベクトルを抽出した．

音節 HMM を構築し，10 数字の単語認識を行った．線スペクトルと音節尤度ベクトルを組み合わせたものに加え，線スペクトルのみ，音節尤度ベクトルのみによる特徴パターンを用いた実験を行った．認識結果から，音節尤度ベクトルを特徴量として用いた場合が最も正解率が高かった．今後は被験者を増やし，より多くの脳波データを用いて認識実験を行う．

参考文献

[1]Bouchard, K.E. et al.: Functional organization of human sensorimotor cortex for speech articulation., Nature; 497(7441): pp.327-332, 2013.

[2]Flinker et al.: Redefining the role of Broca's area in speech, PNAS vol.112 no.9 pp.2871-2875, 2015.

[3]Indefrey, P.et al., The spatial and temporal signatures of word production components, Cognition 92, 101-144, 2004.

[4]伊東崇：サイレントシーズ BCI -HMM の性能評価-, 信学技報 MBE2014-130, NC2014-81, pp.81-84, 2015.

[5]新田ほか: 音声想起時脳波からの音声言語表象抽出に基づく BCI 研究, 人工知能学会全国大会, 2019.

[6]深井ほか, 音声想起時脳波中の音節識別について, 人工知能学会全国大会, 2019.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

音声想起時脳波中の音節識別について

Syllable recognition in speech-imagery EEG

深井 健大郎[*1] 大村 英史[*1] 桂田 浩一[*1] 新田 恒雄[*2, 3]
Kentaro Fukai Hidefumi Ohmura Kouichi Katsurada Tsuneo Nitta
Tokyo Univ. of Science#1 Waseda Univ.#2 Toyohashi Univ. of Tech.#3

Fundamental research on speech-imagery recognition for BCI is one of the challenging technologies, however, the effective method to break through various difficulties in this area does not exist now. In this report, we propose the feature extraction of line spectra from EEG signal that represent linguistic information. Sixteen syllables in ten digits that are extracted from continuously-imagined-speech by hand-labelling are evaluated. Experimental results for syllable recognition based on subspace method are described.

1. はじめに

様々な領域に亘る Brain Computer Interface(BCI)の研究が盛んである. BCI 研究では, 主に運動野で観測した情報を解析し, 機器を操作するものが主であったが[1][2], 近年, 音声言語を抽出して識別することにより, BCI の応用を拡げる試みが始まっている. 発話時の脳波に係る研究では, 脳皮質上で観察した ECoG 信号を解析するものが先行している[3][4]. 他方, 頭皮上で観測した脳波(EEG)を用いる研究も, 実用化できると応用が格段に広がるため, 意欲的に進められている.

EEG 信号を用いる音声言語の識別研究は, 発話時の脳波を対象にするものと, 音声想起(speech imagery)時の脳波を用いるものがある(サイレント音声). 発話時脳波は音声データを同期に利用できるため, 言語情報の抽出が比較的容易である. 一方, 音声想起時脳波は想起タイミングを同定することが困難なため, 言語情報の抽出が難しい. 発話時脳波との相関を求め, 想起区間を抽出する手法が試みられているが[5], 決定的な手法とは言い難い.

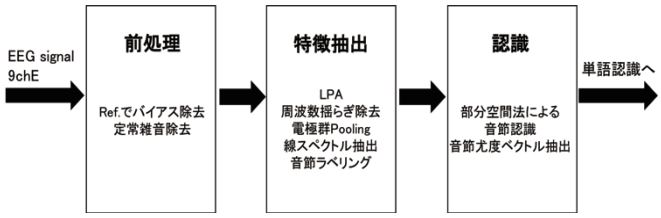
本報告では, 音声想起時の言語表象は tone-burst 波群,すなわち線スペクトル群であるという仮説を基に, 線形予測分析(LPA)を適用して, 線スペクトル特徴時系列を抽出する. 続いて, 時系列中の音節セグメントを目視でラベリングしたデータを対象に[6], leave-one-out 法による識別実験を行う. 識別器には, 小規模コーパスで比較的高い精度が得られる部分空間法(Subspace Method; SM)を用いた[7].

2. 提案手法

図 1 に提案手法のブロック図を示す.

2.1 ノイズの除去

音声想起 EEG 信号は, α 波などの定常雑音, 脳波収録時の電源ノイズ等, 想起対象からみて多くのノイズを含む. そこでまず各電極出力から, 単位時間毎に全電極の平均値を引き去る. 全電極の平均値を参照値とすることで, 全電極に亘るバイアスを除くことができる. 続いて各電極信号に対して, 高次認知機能



連絡先:深井健大郎, 東京理科大学,
6318526@ed.tus.ac.jp

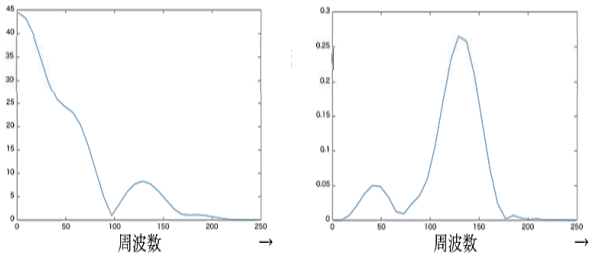


図 2 ノイズ除去前後の振幅スペクトル比較 (/a/)

に係る High-γ 帯の周波数成分を取り出すため, 70Hz-180Hz の帯域通過フィルタ(BPF)を掛ける. 最後に, 予め想起前後の定常時脳波振幅スペクトル成分を計算しておき, この周波数成分をフレーム毎の振幅スペクトル成分から引き去ることで, ノイズを低減する(noise subtraction). 以上の処理により, 想起対象信号以外の様々な成分を除去することができた. 図 2 に処理前(左)と処理後(右)の振幅スペクトルの違いを示す.

2.2 線形予測分析(LPA)による脳波スペクトルの計算

音声想起時の言語表象を線スペクトル群であるとした仮説に沿って, スペクトルピークの抽出に適した線形予測分析(Linear Predictive Analysis; LPA)を適用して, 線スペクトル特徴時系列を抽出する. 線スペクトル群は, 比較的少数の成分からなっていることが予備実験からも分かっているため, 分析次数は 6 程度(3 個の周波数ピーク相当)で十分と考えられるが, ノイズを考慮して 8 次とした(自己相関法のハニング窓長=24msec, 4msec シフト(サンプリング周期 Ts = 2msec)).

LPA は, 人間の音声生成過程をモデル化する手法の一つとして開発された方式である. この合成による分析(Analysis-by-Synthesis) 手法を脳波に適用すると, 脳波信号$s(n)$はランダムノイズを入力信号$f_M(m)$とし, 音声想起の脳波を生成する合成フィルタ係数を$a_M(m)$として次式の畳み込み演算で表現できる.

$$s(n) = -\sum_{m=1}^{M} a_M(m)s(n-m) + f_M(n) \qquad (1)$$

ここでMは線形予測の次数を表す. LPA では合成フィルタ係数$a_M(m)$をレビンソン・ダービン法から計算して求め, 逆フィルタ特性として脳波を入力信号とする分析フィルタを構成してスペクトル包絡を求める.

2.3 非線形フィルタによる周波数揺らぎの除去

LPA を用いたスペクトル包絡情報には, 5～10 Hz 程度の周波数揺らぎが存在する. 周波数揺らぎはスペクトル包絡の極大値や変曲点を変化させるため, 揺らぎを極力消す必要がある. 本報告では, 時間方向に対して非線形フィルタリングすることにより, この揺らぎを低減する. 非線形フィルタとしては中間値フィ

ルタ(Median filter)を使用し，出力を Gaussian fileter で平滑する．Median filter は算出した振幅スペクトル包絡のうち，同一周波数帯の前後 3 フレーム（計 7 フレーム）から中央値を出力し，Gaussian filter は 1 フレームシフトで，前後 3 フレームに対して $(1/4,\ 1/2,\ 1/4)$ の重み加算することで，隣接フレーム間の変動を低減し，時間方向を平滑化する(8msec)．

2.4 Pooling による電極出力の統合

想起 EEG 信号は，ブローカ野周辺の 9 電極で観測しているが，音声想起時脳波の周波数パターンを観察した結果，ブローカ野周辺の 9 電極で言語情報が観察できることが分かった．しかし，想起内容により活性度の高い電極が異なる現象が観察された．そこで 9 電極に対して，振幅スペクトルの要素毎に p-norm (p = 4) 演算を適用し，時間フレーム毎に電極情報の集約を行う．pooling 処理により，複数電極に亘って観測される音節の出現情報を 1 電極情報として統合表現することができる．

2.5 線スペクトルの抽出

電極毎の LPA 分析の後，pooling 処理で統合した振幅スペクトル包絡から線スペクトルを抽出する．線スペクトル計算では，スペクトル包絡に対して 1 階微分$\Delta = 0$ となる極大値，及び，2 階微分$\Delta\Delta = 0$となる変曲点の双方を求め，それぞれについて最大 3 個を線スペクトルとして検出する． 変曲点を加えたのは，本来出現すべきスペクトルピークが余り顕著ではない場合にも，線スペクトルとして抽出すべき例が少なくなかったためである．なお$\Delta\Delta = 0$ となる変曲点を線スペクトルとして採用する際には，次の条件を設けている．

- $\Delta > 0$ かつ $\Delta\Delta$ が $+$から$-$に変化
- $\Delta < 0$ かつ $\Delta\Delta$ が $-$から$+$に変化 (2)

図 3 に線スペクトルパターンの例を示す．

2.6 音節識別

音節識別には部分空間法(Subspace Method; SM)を用いる．SM は，同一クラスに属するパターンが，低次元部分空間に偏在することが多いことに着目して考案された手法である．学習データから各クラスを表現する低次元部分空間を用意し，評価データを部分空間上に射影した際に，最も大きな類似度を与えた部分空間を評価データの属するクラスとする．学習データの部分空間は主成分分析を用いて算出する．k 個のクラスに対して未知データ $\boldsymbol{x}$ を分類する際は，次式の類似度 $s_k(\boldsymbol{x})$ を算出する．

$$\boldsymbol{s_k(x)} = \sum_{m=1}^{M}(\boldsymbol{x}\cdot\boldsymbol{\varphi_{km}})^2 \qquad (3)$$

ここで，φ_{km} はクラス k の第 m 番目の固有ベクトルであり，$m = 1, \ldots, M$ は部分空間の番号を示す．部分空間法では，通常，入力パターンを正準化(平均値バイアスの除去)すると共に，ノルムで正規化する．識別の特徴量には上述の線スペクトルと，周波数間の関連性を表現した Bi-linear 構造（Bi-Spectrum）をフレーム毎に抽出した場合の 2 通りの検証を行う．また，識別の際，評価データ毎に求める学習データとの類似度は，音節尤度ベクトルとして単語識別の特徴量に用いる[8].

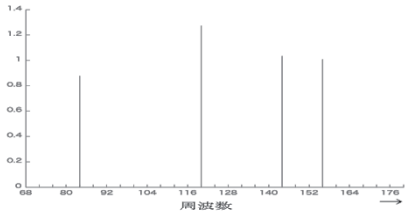


図 3 /a/と想起した際の線スペクトル

3. 評価実験

部分空間法を用いた短音節認識を leave-one-out で評価する．

3.1 実験資料

成人男性 1 名分のブローカ野周辺の 9 電極で収録した音声想起時脳波を用いる．識別対象の音節は/zero/, /ichi/, /ni/, /san/, /yon/, /go/, /roku/, /nana/, /hachi/, /kyu/の10数字中に含まれる音節とし，音節セグメント毎に目視ラベリングしたデータを用いる．識別対象の音節は以下の 16 音節である．

/i/, /chi/, /ni/, /sa/, /a/, /yo/, /go/, /o/, /ro/, /ku/, /na/, /ha/, /kyu/, /u/, /ze/, /e/

音節識別は母音音節グループ（/a/, /i/, /u/, /e/, /o/），有声子音音節グループ（/ni/, /yo/, /go/, /ro/, /na/, /ze/），無声子音音節グループ（/chi/, /sa/, /ku/, /ha/, /kyu/）に分けて行う．ただし，有声子音音節，無声子音音節は 10 数字を連続で音声想起したデータからラベリング済みの 11 サンプルを採用し，母音音節は別に 5 母音を連続して 3 回音声想起したデータからラベリングした 11 サンプルを用いる．また，子音グループでは/na/と/ni/など，後続母音の違いは母音認識を別に行うことを考えており，以下では統合して扱う．

3.2 実験結果

部分空間法を用いた各音節グループの音節認識結果を表 1 に示す．線スペクトルはスパースな表現となっているため，3 つの周波数帯域を 1 つの周波数帯域として平滑し，3 フレームを 1 フレームに平滑することで正解率が向上した．

表 1 母音，有声音，無声音毎の認識結果 正解率（%）

	母音音節	有声子音音節	無声子音音節
線スペクトル	43.8	43.6	38.1
Bi-Spectrum	51.2	40.0	36.3

表 1 から，母音音節，有声子音音節，無声子音音節それぞれのグループについて，チャンスレート(20.0%)を 2 倍程度上回る結果を得ることができた．

4. おわりに

本報告では，音声想起時脳波から線スペクトル特徴量を抽出し，音節セグメントを目視でラベリングしたデータに対して，音節認識を行なった．

実験では成人男性 1 名が 10 数字を連続で音声想起した際の脳波と，5 母音を連続で 3 回音声想起した脳波を用いた．特徴抽出時には，定常ノイズの除去，LPA による振幅スペクトル包絡の算出，非線形フィルタリングによる周波数揺らぎの除去，pooling による電極群の結合を行い，線スペクトルを抽出した．

最後に，部分空間法による母音音節，有声子音音節，無声子音音節 各グループの音節識別を行なった．結果はチャンスレートを 2 倍程度上回る結果を得ることができた．実験結果から，音声想起時脳波に対する線スペクトル特徴量と，目視ラベリングを行うことの有効性を示すことができたが，今後，被験者を増やして統計データ処理を行える環境を整え，音声想起の認識性能向上を図りたい．

参考文献

[1] Hochberg L.R. et al: Neuronal ensemble control of prosthetic devices by a human with tetraplegia, Nature, Vol.442, 164-171, 2006.

[2] Velliste M. et al: Cortical control of prosthetic arm for self-feeding, Vol.453, 1098-1101, 2008.
[3] K. E. Bouchard et al: Functional organization of human sensorimotor cortex for speech articulation, Nature 495, 327-332, 2013.
[4] Flinker et al: Redefining the role of Broca's area in speech, PNAS, vol.112, no.9, 2871-2875, 2015.
[5] 神崎卓丸ほか: 発話時と想起時の脳波による日本語短音節の比較, 日本音響学会春季研究発表会, 2017
[6] 新田恒雄ほか: 音声想起時脳波からの音声言語表象抽出に基づくBCI 研究, 人工知能学会全国大会, 2019
[7] T. Nitta et al: Feature extraction for speech recognition based on orthog-onal acoustic-feature planes and LDA, ICASSP '99, 421-424, 1999.
[8] 平田里佳ほか: 音声想起時脳波からの単語認識, 人工知能学会全国大会, 2019

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

音声想起時脳波からの音声言語表象抽出に基づくBCI研究

BCI Research framework based on the extraction of language-representation in speech-imagery EEG

新田 恒雄[*1, 2] 深井 健大郎[*3] 桂田 浩一[*3] 入部 百合絵[*4] 田口 亮[*5] 杉本 俊二[*2] 堀川 順正[*2]
Tsuneo Nitta Kentaro Fukai Kouichi Katsurada Yurie Iribe Ryo Taguchi Shunji Sugimoto Junsei Horikawa
Waseda Univ.#1 Toyohashi Univ. of Tech.#2 Tokyo Univ. of Science#3 Aichi Prefectual Univ.#4 Nagoya Institute of Tech.#5

Speech imagery, or silent speech, recognition from *Electroencephalogram (EEG) is one of the challenging technologies for non-invasive brain-computer-interface (BCI). In this report we regard the representation of language as the difference of line spectra of syllables observed at Broker area and develop the method for extracting line spectra in EEG signal. The BCI research framework based on the extraction of language-representation, as well as several examples of syllable investigated using nine electrodes around Broker area, are presented.*

1. はじめに

近年，fMRI, PETなどによる脳観察から，概念→ 言語表現プランニング→ 音節・単 語・文表現→ 発話運動神経系に至る音声言語活動に関与する部位についての知見が蓄積されつつある[1], [2]. これらの知見を受け，脳波観測から言語活動を推定する研究が活発である．音声言語に関する脳波研究では，脳皮質上のセンサーから発話時の脳波(electrocorticogram: ECoG)を検出する研究が先行しているが[3]，頭皮上で観測した脳波(electroencephalogram: EEG)信号を用いる研究は，実現されると応用が格段に広がるため意欲的に進められている．これまでの研究では，“音声言語情報がEEG信号の中で具体的にどのように表現されているのか”という言語表象に関する解明は未着手であった．本報告では，言語表象は線スペクトル群から成っていると仮定し，これらの抽出法を開発することにより，異なる想起音節では対応する線スペクトルが異なることを示す．また，線スペクトルに基づくBCI研究で必要となるフレームワークについても述べる．

2. 脳波に現れる言語表象

音声言語は，左中側頭回(MTG)での概念準備の後，言語としてのプランニングが左上側頭回(STG)にかけて行われるとされている[1]. この後左下前頭回(IFG; ブローカ野)において音節化(syllabication)が行われ，音声発話の場合は左中心前回(PG; 運動野)で調音(構音)が行われる[2]．音声想起(speech imagery) のように発話を伴わない音声言語に対するデコーディングも，ブローカ野に到達する言語表象を捉えることができるなら可能になる．

2.1 言語表象観測へのアプローチ

脳波では低周波から順に，θ 波(4-8Hz)，α 波(8-14Hz)，β 波(14-38Hz)，γ 波(26-70Hz)が知られており，認知状態との関連が古くから研究されてきた．一方近年に入ると，言語が関わる高次認知活動をγ 波より上の帯域(70Hz～)で観測・解析する研究が盛んである．

我々は脳波解析に先立ち，「言語表象は低周波の定常的信号と同様の正弦波で，かつ言語の性格から極短時間だけ現れる」と考えた．こうした信号は tone-burst と呼ばれ，古くから過渡的信号の解析に用いられてきた．また正弦波の継続長は，反転音声(reverse speech [4]) に関する実験から，反転の単位を100 msec 以下になると音声を明瞭に確認できることが知られており，この事実から tone-burst 信号の継続長を 32-128msec と推測した．

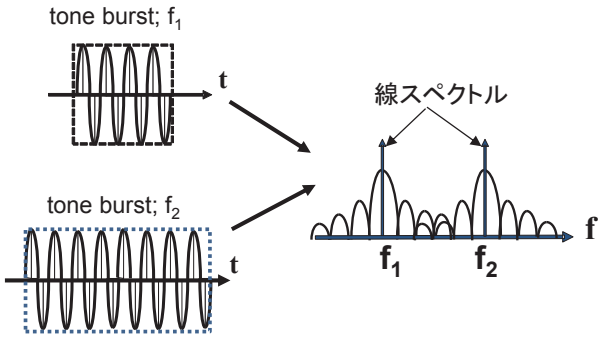


図 1 tone-burst 信号とスペクトル形状

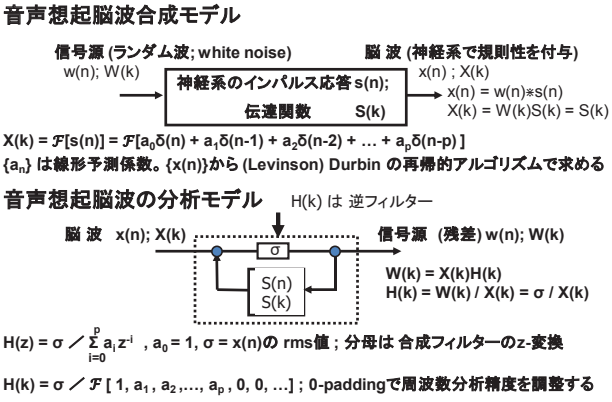


図 2 脳波の合成による分析モデル

図 1 に周波数の異なる tone-burst 信号とそのスペクトル形状を示す．tone-burst 波は，時間窓の内部では正弦波であるが，過渡的信号の性格から図のようにスペクトル分布に広がりを持つ．このため，元々の線スペクトル成分を抽出するための信号処理が必要になる．以下では線形予測分析(Linear Predictive Analysis; LPA)と線スペクトル抽出方法について説明する．

2.2 線形予測分析(LPA)法に基づく脳波の分析[5]

線形予測分析(Linear predictive analysis; LPA) を図 2 から説明する．図は音声想起脳波を念頭においた脳波合成モデルと脳波から言語表象を含むスペクトル情報抽出のための分析モ

連絡先：新田恒雄，早稲田大学/豊橋技術科学大，nitta@cs.tut.ac.jp

デルを示している．合成モデルでは，信号源としてのランダム波 w(n)を入力とし，言語情報を生成する神経系のインパルス応答 s(n) との畳み込み演算により，規則性を付与された脳波として tone-burst 群から成る x(n) が出力される．図には周波数領域での表現{W(k), S(k), X(k)}も示した．インパルス応答 s(n) は線形予測係数{a_0, a_1, …a_p}から計算することができる．

電極で観測される脳波 x(n)に対する分析モデルでは，合成モデルで与えられた線形予測係数から，脳波の持つ言語表象を計算する(Analysis-by-Synthesis; AbS)．通常の計算アルゴリズムでは，まず脳波の自己相関係数から Durbin 法を適用して線形予測係数を求め，逆フィルターを構成して脳波のスペクトルを分析する．この時，線形予測係数(ここでは 8 次) に 0-padding を施すことで，任意の周波数分析精度を得ることができる．図 3 に FFT 分析と比較した LPA の分析例を示す．

3. 脳波測定と言語表象の抽出

ブローカ野を中心に 9 個(3×3) の電極を配置した．10-20 法での電極位置は{(F3, F5, F7), (FC3, FC5, FT7), (C3, C5, T7)}で，g-tec 社の脳波計を使用して EEG 信号を採取した．データ収集に関する詳細は，文献[7]に同じである．電極毎に LPA 分析した後，周波数揺らぎに対して中間値フィルターとガウス平滑フィルターによる揺らぎ除去を行った．続いて，9 電極の振幅スペクトルの各要素について，p-norm (p=4) による pooling 処理を行った．この処理後は，単一チャンネルの振幅スペクトル系列信号に対して，周波数方向の 1 階微分(極大値抽出)，および 2 階微分(変曲点抽出)を計算して，線スペクトルを抽出する[6]．

音声想起/a-i-u-e-o/ と 10 数字連続音声想起のデータから目視で音節をラベリングした．/a-i-u-e-o/ については三回連続して音声を想起することで，熟練した研究者は線スペクトルパターンの変化から音節を特定することができる．10 数字音声想起では，別に採取した 101 単音節データの標準パターンを参照してラベリングの精度を上げる工夫をした．

図 4 に幾つかの音節の線スペクトルパターンを示す．周波数は 70Hz ～176Hz の範囲に分布している．線スペクトルは 2～5 個程度が観察されるが，これらの意味解明は今後の課題である．図の左側には 5 母音音節の線スペクトルを示した．この中では /u/の線スペクトルが低周波に集中することが顕著である．ここに示した例から，母音同士の分離は可能と思われるが，今後，統計的に調査したい．図の右側には，連続数字 ”0-1-2-…-9-0”を想起した際の脳波から採取したオ段の 4 つの音節を示した．音節想起パターンでは，母音の影響が顕著であるが，後続母音が同じ音節でも線スペクトルパターン上には違いが観られ，分離が可能と思われる．音声学/音韻論では，音素と音節の意義が長らく討議されているが，脳波に観られる言語表象としては，音節に軍配が上がりそうである．ただし，これも今後，定量的な調査が必要である． 線スペクトルパターンは所謂スパースな特徴である．このため，認識エンジンの中で線スペクトルを扱うには，今後工夫が必要になると思われる[6], [7]．

音声想起信号中の言語情報を線スペクトル群から抽出することの有効性を示すデータとして，図 5 に連続数字を想起したデータに対して手動ラベリングを行ったものを示す．図は音節尤度を部分空間法から計算してパターン化したものを示している．この図から比較的容易に音節ラベリングが可能になることが理解できよう．ただし部分空間法での辞書は学習データ内の音節パターンを用いており，将来，非学習データに対して，同様の性能を得ることが目標となる．

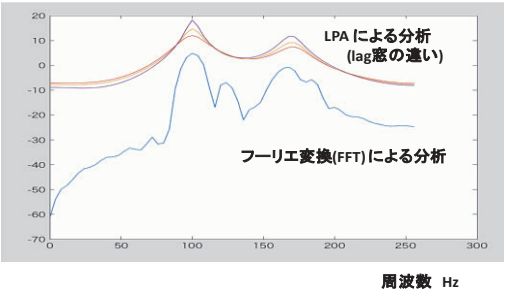


図 3 FFT 分析と LPA 分析のパターン比較 (/a/)

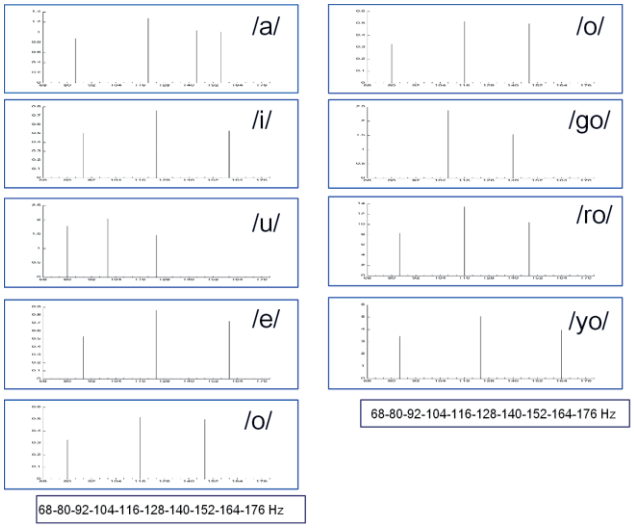


図 4 音節の線スペクトルパターンの例

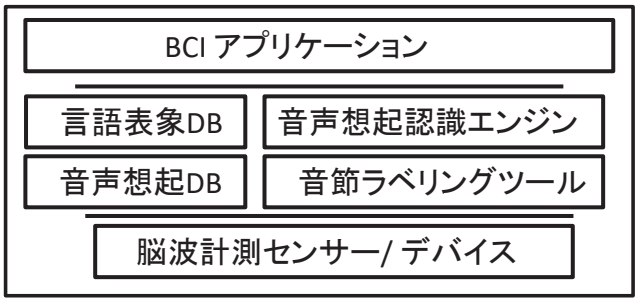


図 6 言語表象に基づく BCI 研究フレームワーク

4. 言語表象に基づく BCI 研究フレームワーク

言語表象に基づく BCI 研究は ALS 患者など医用だけでなく，次世代 AI speaker 等波及効果が大きい．そのためここ数年は，図 6 に示すようなセンサーデバイス，認識エンジンの開発と共に，音声想起 DB, 言語表象 DB，音節ラベリングツール[8]を整備する必要がある．

5. おわりに

音声想起時の脳波から，言語表象としての線スペクトル群を抽出する方法を開発し，想起音節間で線スペクトル特徴が分離できることを示した．今後，被検者数を増やすと共に，ラベリングツールの整備と言語特徴の抽出精度向上を図りたい．

参考文献

[1] Indefrey, P et al., The spatial and temporal signatures of word production components, Cognition 92, 101-144 (2004).
[2] Bouchard K.E. et al., Functional organization of human sensorimotor cortex for speech articulation, Nature 495, 327-332 (2013).
[3] Heger D. et al., Continuous Speech Recognition from ECoG, Interspeech2015, 1131-1135 (2015).
[4] https://en.wikipedia.org/wiki/Reverse_speech
[5] 板倉，斎藤，最尤スペクトル推定法を用いた音声情報圧縮，音声日本音響学会誌，27-9, 17-26 (1971).
[6] 深井ほか，音声想起時脳波中の音節識別について，人工知能学会全国大会(2019).
[7] 平田ほか，音声想起時脳波からの単語認識，人工知能学会全国大会(2019).
[8] 傅ほか，音声想起時脳波からの音節ラベリングツール，人工知能学会全国大会(2019).

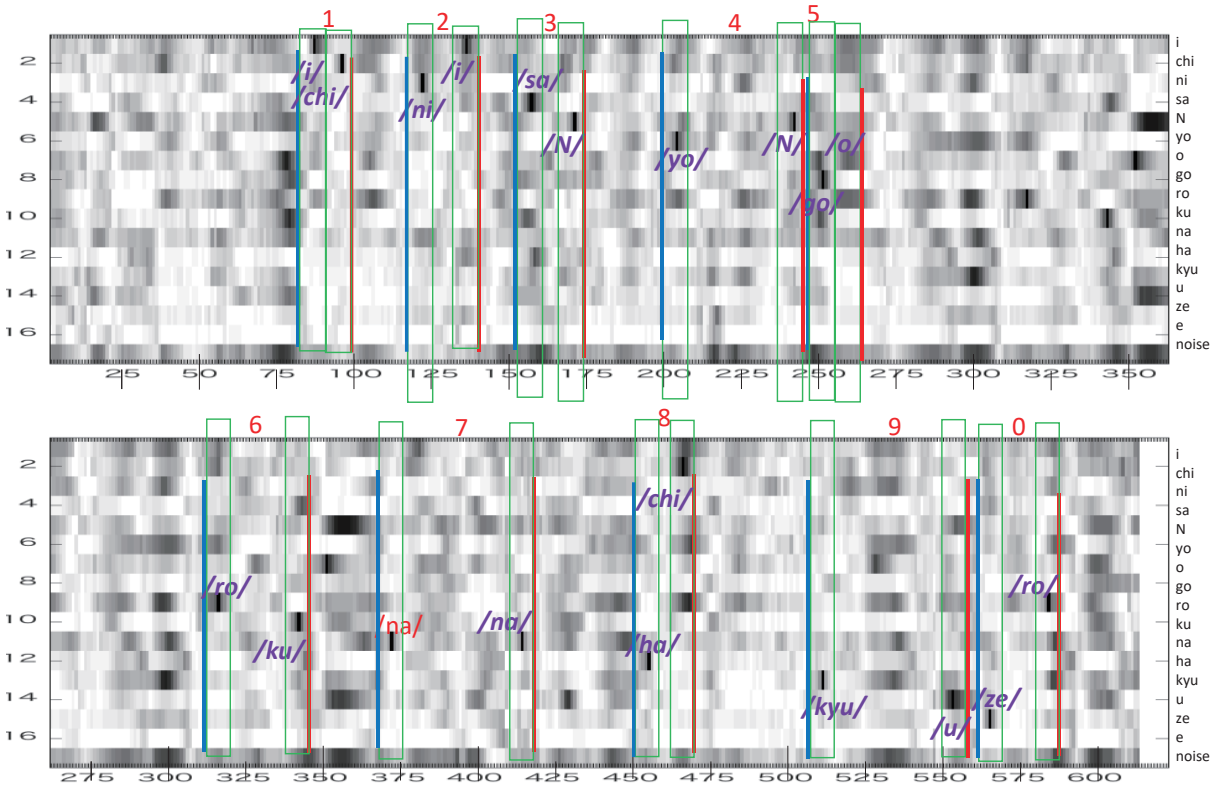


図 5 連続数字音声想起データに対する手動音節ラベリング

(図の右端に音節名を示し，図は濃淡で音節尤度を示した．)

Organized Session | Organized Session | [OS] OS-6

[4D2-OS-6] 人工知能と倫理

神崎 宣次（南山大学）、久木田 水生（名古屋大学）、服部 宏充（立命館大学）

Fri. Jun 7, 2019 12:00 PM - 1:20 PM Room D (301B Medium meeting room)

[4D2-OS-6-01] Trust in Medical AI

〇Arisa Ema[1], Katsue Nagakura[2], Takanori Fujita[3] （1. The University of Tokyo, 2. M3, Inc., 3. Keio University）

12:00 PM - 12:20 PM

[4D2-OS-6-02] A Study on Constructing the Ethical Design Sciences

〇Kaira Sekiguchi[1], Koichi Hori[1,2] （1. Graduate School of Engineering, The University of Tokyo, 2. Center for Advanced Intelligence Project, RIKEN）

12:20 PM - 12:40 PM

[4D2-OS-6-03] Practice of a support for creating self-regulations on profiling

〇Fumiko Kudo[1] （1. Institute of Economics Chukyo University）

12:40 PM - 1:00 PM

[4D2-OS-6-04] AI in Sci-fi Movies from the Viewpoint of Human Centered Design

〇Shigeyoshi IIZUKA[1] （1. Kanagawa University）

1:00 PM - 1:20 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

医療分野における AI への信頼

Trust in Medical AI

江間有沙[*1] Arisa Ema　長倉克枝[*2] Katsue Nagakura　藤田卓仙[*3] Takanori Fujita

[*1] 東京大学 The University of Tokyo　[*2] エムスリー株式会社 M3, Inc.　[*3] 慶應義塾大学 Keio University

Discussion on AI principles related to privacy, security and accountability are important along with the usage of artificial intelligence (AI) at various fields. There are expectations as to build "trustworthy AI system." We investigated what are the bottleneck on medical AI by creating a workflow diagram of medical treatment by doctors and questioned medical-related experts participated to seminars. As a result, issues on some dilemma situation as well as usefulness (accuracy) of technology, safety, convenience, security, cost were pointed out. In the future, we plan to deepen this qualitative data further by questionnaire survey targeting a large number of people.

1. はじめに

現在，様々な現場で人工知能（AI）利用に関する議論と同時に，プライバシーやセキュリティ，アカウンタビリティなどに関する原則作りが国内外行われている．いかに「信頼できる AI システム」を構築できるかに関しては期待も多いが課題も多い．

特に医療 AI は個人データの扱いとプライバシーや，事故時の責任の所在や法的，社会的，倫理的な課題も多く取り上げられている．2015 年に厚生労働省で取りまとめられた「保健医療 2035」などにおいても，ICT を活用によって医療の質，価値，安全性，パフォーマンスの向上が目的として掲げられているものの，その活用が進んでいないことも同時に指摘されている[厚生労働省 15]．

その後，2016 年には厚生労働省内において「保健医療分野における ICT 活用推進懇談会」が提言書をとりまとめ，2017 年には「保健医療分野における AI 活用推進懇談会」が設置された．また，日本医療研究開発機構（AMED）では国内の 6 学会が主導し AI による診断・治療支援のため画像診断データベースを構築する事業を実施，2018 年には保健医療分野 AI 開発加速コンソーシアムが設置され，課題や対応策の検討が進められている[厚生労働省 18]．

2. 枠組み

様々な課題が指摘されている中，「AI を活用した」と銘打った技術はどのような観点で現場から求められているのだろうか．AI と雇用をめぐる議論においては，必要なのは「最先端の技術」ではなく，「こなれた」技術でもよいから壊れなくて理解が可能なものであったり，そもそも紙資料の「データ化・共有化」だけでもよかったりする（逆に言うとそれすらできていない）ことが医療現場以外でも指摘されている[江間 19]．

そこで本稿では筆者らが企画した医療 AI に関するセミナーにおいて，医療側からのニーズを整理することを目的とする．

2.1 ワークフロー図の作成

保健医療分野において活用される AI を，技術的な観点ではなく医療の診療プロセスの観点から整理するため，本研究では医師による診療のワークフロー図を作成した（図1）．ワークフロー図の作成にあたっては，(A) 診療プロセスを医師が行うべき「ステップ」と，(B) 医師が主体となって判断を下す「サブステップ」に分けた[横山 19]．AI は(B)「サブステップ」の効率を上げて情報を提示する支援ツールであることが医師法第 17 条の規定との関係で整理されている．

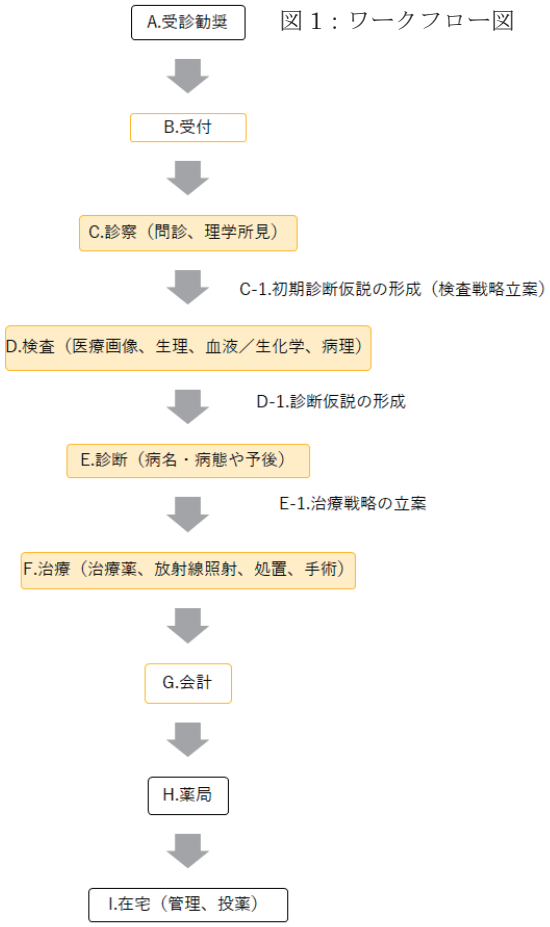


図 1：ワークフロー図

2.2 対象領域と話題提供者

厚生労働省の「保健医療分野における AI 活用推進懇談会」では「AI の実用化が比較的早いと考えられる領域」として(1) ゲ

連絡先：江間有沙，東京大学政策ビジョン研究センター，ema@pari.u-tokyo.ac.jp

ノム医療, (2) 画像診断支援, (3) 診断・治療支援, (4) 医薬品開発,「AI の実用化に向けて段階的に取り組むべきと考えられる領域」として, (5) 介護・認知症, (6) 手術支援の合計 6 領域を, AI 開発を進める重点領域として位置づけている[厚生労働省 17]. 工程表には診療報酬や介護報酬の改定時期も明記されており, 早ければ 2020 年の改定における AI 活用の評価を視野に入れていると考えられる.

「AI の実用化が比較的早いと考えられる領域」のうち, (4)の医薬品開発は医師による診療には直接関係しないため, 本研究では(1) ゲノム医療, (2) 画像診断支援, (3) 診断・治療支援に対象領域を絞る.

表 1 にリスト化した全 5 回のセミナーの話題提供者には, 現場の医師や研究者として具体的に利用/開発している診断や治療支援 AI を紹介していただいた.

表 1 各回のテーマと話題提供者

テーマ	講師	所属
(1) 開業医と AI	目々澤肇氏	目々澤医院院長
	田澤雄基氏	慶應義塾大学医学部精神・神経科, MIZEN クリニック豊洲院長
(2) 消化器科と AI	平澤俊明氏	がん研有明病院上部消化管内科副部長
	青木智則氏	東京大学大学院医学系研究科内科学専攻博士課程
(3) 内科と AI	沖山翔氏	アイリス株式会社代表取締役, 救急専門医
	巣籠悠輔氏	株式会社 MICIN 最高技術責任者, 東京大学招聘講師
(4) ゲノム医療と AI	湯地晃一郎氏	東京大学医科学研究所特任准教授
	西村邦裕氏	株式会社テンクー代表取締役社長 CEO
(5) 眼科と AI	田淵仁志氏	ツカザキ病院眼科創業者・主任部長, 眼科専門医
	升本浩紀氏	ツカザキ病院眼科・人工知能エンジニアチーフ, 医師
	三宅正裕氏	京都大学大学院医学研究科眼科学教室 特定助教

2.3 質問項目

医療現場で技術実装を進めていく上での課題抽出のため, 話題提供者とセミナーの参加者には以下の 4 つの質問を投げかけている.

① 医療提供に関するフロー(表 1)のうち, あなたはどこでどのような課題を認識していますか？以下のステップのうち該当する箇所を丸印で囲み, 課題をご指摘ください. 複数ありましたらいくつでもご記入ください(例:「A.受付」に〇を付けて, その横に「待ち時間が長い」と記入)

② ①の課題をあなたはどのように解決しようとしていますか？またはどのように解決して欲しいと思いますか？(例:受付自動システムの導入／予約制／診療時間の延長など)

③ ②の解決策を進めるためには, まずどこから／何から着手するべきだと思いますか？(例:業者に見積もり依頼／予約システム作る／医師の新規採用)

④ ②の解決策を進める上で, どのような障壁があると考えますか？(例:システム導入コストが高い／医師不足)

3. 結果と今後

本稿を執筆している 2018 年 2 月ではセミナーの第 3 回までが終了しており, 各回は平均して毎回 30－40 名程度, 開業医や勤務医などの医療従事者のほか製薬会社, 医療政策関係者, エンジニアなどが参加している.

今までの議論からは, 技術の有用性(精度), 安全性, 利便性, セキュリティ, コストの他, いくつかのジレンマ状況に対する課題が指摘された. 例えば以下のような問題である

- 現在の技術では治せない症例が見つかってしまった場合どうするか
- 将来的に生じるだろうリスクの予測を患者につげるべきかどうか, またそれが偽陽性であった場合, 判明するまで患者の心理的な負担につながらないだろうか
- AI による診断で確率は低いものの, 侵襲性があり患者に負荷の高い検査が必要と提示された場合, 検査を行うべきか
- AI による診断によって確率は低くても可能性のあるものに対する過剰診断が生じないか

AI が入ることによって, 受付や会計の時間が短縮されるだけではなく, 見落としや偽陰性が減り, 患者の満足度が上がり, 医者側の運用負担も減ることが期待される. 一方で, AI による診断や予測, 提案に対しては, 現在は医師に最終的な判断が委ねられている. そのため上記で挙げたような問題が増加すれば, 逆に医師の労働時間を増やし, さらには場合によっては, 患者の身体的, 心理的, 経済的コストが増えることも懸念される. 今後はこの質的なデータをさらに大人数を対象としたアンケート調査などによって深堀をしていく予定である.

なお, 各回の話題提供者の内容はウェブに報告書が掲載されている. 全 5 回のセミナーを通して, 現場の医師からはどのような AI の利用のニーズと, 技術シーズとのミスマッチがあるのかについて, 当日会場で議論を行いたい.

参考文献

[江間 19] 江間有沙:人工知能社会の歩き方」, 化学同人, 2019.

[厚生労働省 15] 厚生労働省:保健医療 2035, https://www.mhlw.go.jp/seisakunitsuite/bunya/hokabunya/shakaihoshou/hokeniryou2035/

[厚生労働省 17] 厚生労働省:保健医療分野における AI 活用推進懇談会, https://www.mhlw.go.jp/stf/shingi2/0000169233.html

[厚生労働省 18]厚生労働省:第 1 回 保健医療分野 AI 開発加速コンソーシアム(開催案内), https://www.mhlw.go.jp/stf/newpage_00177.html

[横山 19] 横山和明, 本邦における人工知能(AI)を用いた診療支援の事例, https://www.mhlw.go.jp/content/10601000/000468448.pdf

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

倫理的設計学の構築に関する一考察

A Study on Constructing the Ethical Design Sciences

関口　海良 *1　　堀　浩一 *1*2
Kaira Sekiguchi　　Koichi Hori

*1東京大学大学院工学系研究科
Graduate School of Engineering, The University of Tokyo

*2理化学研究所革新知能統合研究センター
Center for Advanced Intelligence Project, RIKEN

In the 21st century, further contribution to solving social problems is required for academic professionals. Therefore, we set our research question as "Can engineering contribute more to solving social problems?" We consider that the answer is positive, and proposed the ethical design sciences which is a new interdisciplinary field including engineering. The ethical design sciences will contribute to systematically solving social problems through re-defining representations in design such as perspectives and description methods, positioning creative activity support tools also as experimental devices and demonstrating the effects through user studies etc. In this paper, after describing practical examples, we derived a positive answer to the question.

1.　序論

貧困，飢餓，不平等，気候変動等、社会全体に関わる課題が存在している．これらの解決という目標に関して、特に重要なものは SDGs として纏められ，更に注力されるに至った．また人工知能（AI）に関しても AI を社会に上手く導入していく事が課題となり，人工知能倫理として広く研究されている．

一般に，科学研究は公的研究費を用いており，何らかの形で社会に貢献する責任を有す [吉川 09]．課題解決への期待が高い工学においては尚更，これら社会的課題の解決に対する立場を各研究者は明確にし，取り組む必要がある．

そこで本研究では，社会的課題の解決に向けて従来以上に積極的に貢献するひとつの工学の在り方を探求し，「倫理的設計学」として体系化し，実証を進めている．

2.　倫理的設計学で扱う三領域

倫理的設計学は，倫理学，設計学，創造活動支援の三分野が重なる領域に位置付けられる．

初めに，社会的な課題の解決に向けて科学／技術を用いるには，科学／技術と社会との関係についての考察が重要となる．既に環境倫理，生命倫理，技術倫理等の応用倫理はこれら影響を扱ってきた．倫理的設計学ではこれら学問を参考にしながら，逆に，望ましい変化から逆算して社会にとって良い意味を持つ人工物の設計を目指す．次に，課題解決に向けて体系的に取り組むには，設計を学術的に扱う「設計学」が重要となる．これにより実践だけでなく，知識の蓄積や共有も効率化する．

ここに至って，本研究では設計学と倫理学の重なる領域に「倫理的設計理論（ethical design theory）*1」を導出した．ここでは設計における視点の持ち方や，設計アイデアの記述方法を再定義して，上述の様な人工物に関する「倫理」と，更に技術者にとっての「倫理」に対応する軸を設計において連続的に扱える様になっている [Sekiguchi 18a]．

三つ目として，上記研究を「机上の空論」とするのを避けるべく，従来の工学が持つ良い特徴であった実証的側面を踏襲する．ここでの実証対象は，(1) 社会的課題の解決を体系的に進める為に設計行為がどの様に変化すると良いかと，(2) そこで生まれた設計解が実際に課題解決に資するかである．本研究では，前者を研究テーマとして特に明確に設定する．実証に際しては創造活動を扱うので，創造活動支援の分野を参考にしつつ，ツールを支援に用いるだけでなく，理論を実証する為の「実験装置」として位置付け，「観測データ」の収集にも用いる．更に，設計解の実証には一般的に従来から実験やシミュレーションが行われているので，これらも踏襲するものとする．

これら「理論–実験装置–観測データ」の関係は，物理学におけるそれと対応する．例えば，物理学における「一般相対性理論–レーザー干渉計型重力波検出器（LIGO）–重力波イベントの記録」の関係と類似である．

課題解決に向けた学問としてより明確に取り組むべきは，ニーズを定義し，要素を統合して，価値を生み出す事との理解がある（[吉川 09] pp.6–7 の「第二種基礎研究」，p.86 の「持続性の時代」に関する記述等）．上記三領域を扱う事により，社会的な影響も含めた上で工学としてより積極的に課題解決に貢献していけるものと筆者らは考えている．以下，これら三領域の取り組みについて紹介する．

3.　倫理的設計理論の導出

倫理的設計理論では，工学的視点に基づき倫理を扱う為に設計における表現（視点・記述方法）を再定義した．

3.1　視点としての「倫理レベルからの設計」

視点としては，人工物の階層表現を再定義し，その最上位に倫理レベルを位置付けた．倫理レベルは，自由、平等，持続可能性といった社会全体に関する価値を表現するレベルであり，設計における最終的な目的が表現されるレベルに対応する [Sekiguchi 18a]．

次に，人工物の階層表現では自然科学が対象とする客観的現象が表現されるので，設計者の主観を表現するための新たな次元を位置付けた．筆者らはこの次元を図 1 の様に階層表現に直行させた，三次元の表現を持って現在研究を進めている．

3.2　記述方法としての「言説による設計」

人工物を倫理レベルから設計アイデアを記述する為，倫理的設計理論では下記三つのルールを設定した．

連絡先: 関口 海良，東京大学大学院工学系研究科航空宇宙工学専攻 知能工学研究室，kaira @ dfrome.com

*1　以前はこれを「倫理的設計学」と呼称したが，今回は理論として位置付け直し，倫理的設計学はより広く再定義する．

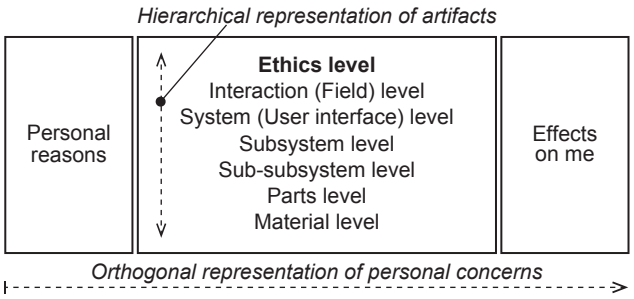


図 1: Simplified version of the three-dimensional representation of artifacts which contains hierarchical and orthogonal representation of artifacts [Sekiguchi 18a, Sekiguchi 18b]

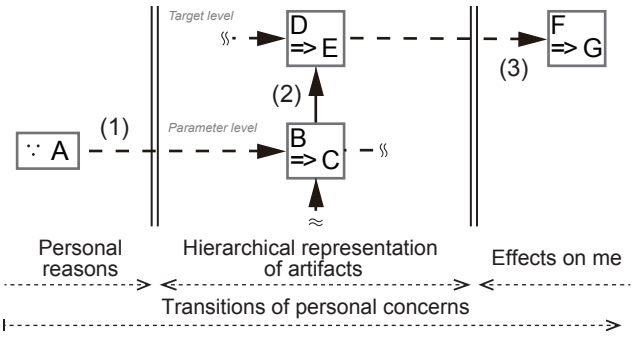


図 2: Hierarchical and orthogonal grammar of design with discourse [Sekiguchi 18a, Sekiguchi 18b]

1. Since A is a personal reason, I/we generate a design that will change B to C in the hierarchical representation of artifacts.
2. If B is changed to C at the parameter level, then D will change to E at the target level in the hierarchical representation of artifacts.
3. If D is changed to E in the hierarchical representation of artifacts, then F will change to G as effects on me/us.

また，図 2 の様に可視化の形式を提案し、式 1〜式 2 の様に「偏微分」の考え方を用いて数学的に理解できる事を整理した．

$$D = f_D(t, B, ...) \quad (1)$$

$$dD = \frac{\partial f_D}{\partial t}dt + \frac{\partial f_D}{\partial B}dB + ... \quad (2)$$

設計におけるアイデアはこれら記述の要素からなるツリー状となる．

3.3 その他の展開

他にも研究テーマとして理論的な展開があり得る．例えば，視点を更に拡張する事ができる．図 1 では，1 次元目が人工物の階層，2 次元目が設計解の併記を可能にする広がり，3 次元目がこれらに直行させた主観的な軸であった．式 1 や式 2 に表現されている様に，これまでも時間を考慮していたが，これは設計解の中の世界に流れる言わば内的時間（inner time）である．一方，他にも設計の外側に流れる外的時間（outer time），

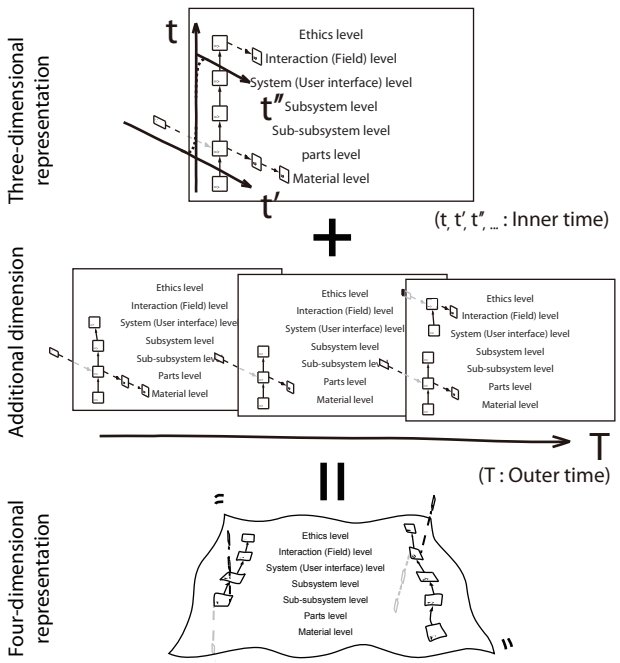


図 3: Overview of the expansion of the dimensions in design

言わば設計者が生きている時間が存在する．例えば，外的時間は設計を見直してアイデアを更新するといった，設計解の時系列変化を扱う際に対象となる．図 3 に示す様に，この外的時間を加える事で 4 次元の表現を考える事ができる *2．

4 次元の表現を考える事で時代の「波」を扱う事ができる．例えば，人工知能は第三次ブームに在るが，これが倫理レベルまで含めてどのような波となっているか，どの手を打つ事で波が制御されこのブームをソフトランディングさせられるか等が扱える．この 4 次元化は筆者らも今後の課題としている．

4. 有機的で動的な支援ツールの実装

倫理的設計理論に基づく設計の実践を支援する為に，筆者らは支援ツールとして Dfrome（Design FROM the Ethics level の略）を実装した．また，Dfrome は倫理的設計理論を実証する為の実験装置として位置付けた．

図 4 が Dfrome の全体像であり，主にブラウザ（browser），エディタ（editor），クラウド環境（cloud）と，探求用エンジン（investigation engine）からなる [Sekiguchi 18b]．以下，探求用エンジンにおける各種処理を紹介する．

4.1 自動マッピング及び自動連結

Dfrome は，文章の形で与えられたインプットから，変化の記述の要素を表現空間上に自動でマッピングし，関連する要素同士を自動で連結する機能を有す．支援ツールとしては編集時にツリーのたたき台を提供し，実験装置としてはツリーの編集が提案者以外の設計者でも可能である事を実証する為に用いる（詳細は 5.1 節参照）．基本的な処理は，既存のツリーを参考にして，どの次元のどの座標に要素を配置すると良いか，また要素間に矢印を引けるかを類推している．

*2 試論は以下参照：Kaira Sekiguchi, The sixth and the seventh rules of “design with discourse” for design from the ethics level, January 1, 2011. ethics-level.com.

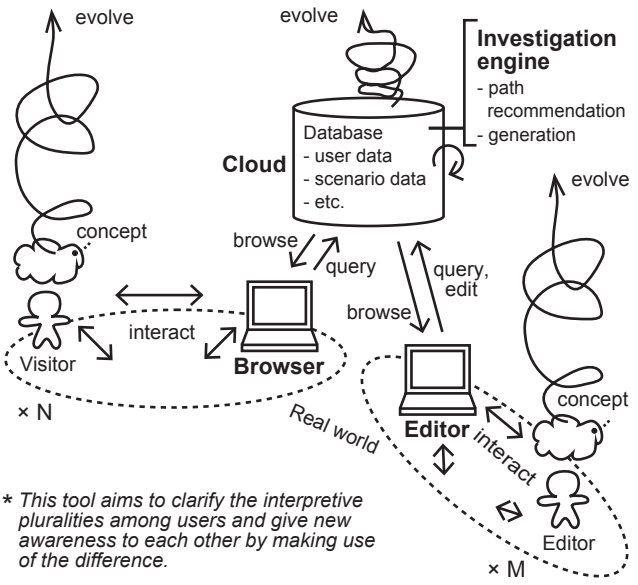


図 4: Overview of Dfrome [Sekiguchi 18b]

4.2 シナリオパスの延長

Dfrome は，考慮すると良いでシナリオパスを推薦する機能を有す．例えば，倫理レベルまで繋がるシナリオや，考慮していそうでしていなかった他の実現手段はどの様なものが在るかを提示する．これらの機能は，シナリオを拡げる過程の再現性の実証としても位置付けられる．こをエディタで行ったケースは [Sekiguchi 18a] に，ブラウザで行ったケースは [Sekiguchi 18b] に記載した．

4.3 ツリー間の距離の計算

Dfrome は，編集者の手元にあるツリーをクエリとして，他のツリーとの距離を計算する機能を有す．これにより，ツリー単位でもどの様な視点を考慮できており，何が漏れていたかを確認するのを支援し，またその過程の実証となる [Sekiguchi 18a]．

4.4 Dfrome のコンテンツ

Dfrome のデータベースには，2019 年 2 月 7 日時点で 122 個のツリーが既存のものとして利用可能（＝非削除かつ公開状態）である．例えば，Asilomar AI Principles や Ethically Aligned Design（EAD）の Principles [IEEE 17] を始めとした人工知能倫理のツリーに加え，機械学習のアルゴリズムやドローンの研究に関するツリーを含む．また，理化学研究所革新知能統合研究センター（RIKEN AIP）がウェブに公開している研究概要 [RIKEN] を自動マッピング及び自動連結を用いながら筆者らが編集したツリーも含む．

5. 観察データの収集

倫理的設計理論の実証を本研究では二段階で行っている．まず，本理論の運用が提案者の暗黙知に依存するなら実用的でない為，そうではなく第三者でも運用可能な事を示す．次に，ユーザスタディによって本理論が設計者に期待通りの効果を与えるかや，想定外の反応が見られるか等を観察する．本来は，設計された人工物が実装され社会に投入された際に期待通りの機能を発現するかを，シミュレーションや社会的実験によって確認する必要がある．これらも倫理的設計学の範疇であるが，その実践は今後の課題としている．

5.1 AI による記述の再現性の確認

初めに，AI によって倫理的設計理論に基づく記述を行う事ができれば，第三者による運用の再現性の証拠となる．

以下の実験では，インプットの文章からの表現空間へのマッピングと，要素間に矢印を引く処理を AI が担い，その他は筆者らで行った．実験条件を保存するため，今回はデータをダウンロードし，ローカルに環境を再構築して実験を行った．加えて，本理論の発展的な運用についても再現性を確認する為，機械学習に関する文章をインプットとしてツリーを作成した後，社会的課題に関する文章を続けてインプットとして，これら異なる文章間に隠れた有用な繋がりを見つけられるかを確認した（図 5 の 1〜2）．適当なインプットを設定する為,「Google Scholar」を用いて，クエリを「machine learning」とした際のトップにランクされた論文 [Baltrušaitis 19] と,「social problems」とした際のトップの論文 [McIntyre 19] の各アブストラクトを用いた．最近の論文を得る為，設定で「2019 年以降」をオンにし,「特許を含める」と「引用部分を含める」をオフにした．

図 5 は試行の結果である.「直さないといけないがそこそこ書けている程度のツリー」が得られている．更に，ツリーを跨いで下記記述 A から B へつなぐ矢印が確認された．

A) Multimodal machine learning => aim to build models that can process and relate information from multiple modalities

B) Results => reveal that discussing an effective solution to a social problem in a news story caused readers to feel less negative and to report more favorable attitudes toward the news article and toward solutions to the problem than when no solution or an ineffective solution was mentioned

この結果から，例えば AI をマルチモーダルにしてニュース等のコンテンツを作成する際には効果的な解決策にも言及する事で，社会的には更なる議論の活性化へと繋がるといった「興味深い」気付きを得られる．ツールは dfrome.com にて提供済みの為，類似の追試を広く可能としている．

5.2 ユーザスタディによる効果の確認

本理論の効果を確認する為，筆者らはユーザスタディを進めている．学生実験からは，例えば Google のみを用いた場合では被験者が既に持っている問題意識に基づき検索を行う傾向があるのに対して，Dfrome を併用すると倫理的な視点の導入を促す傾向があるのに加え，被験者が研究テーマを見直す機会もその中で得られ得るという創造活動支援としての効果も確認できた．例えば，ある被験者はアンケートに「利点については今までも考えたことがあったが，リスクについては考慮していなかったため，リスクについて新たな気づきが得られた」と回答した．今後はより大規模な実験を実施する予定である．

6. 関連研究との位置付け

[Verbeek 11] はひとつの哲学的基盤を本研究に提供する．例えば，4 次元まで考慮した上での図 2 全体の実践は彼の言う自由（freedom）に基づく設計として，図 2 の（3）の中央から右側へ向かう個人的影響の発現が媒介（mediation）として位置付け可能である．一方，倫理的設計学は理論自身の実証を含む為より自然科学的である．更に本理論は階層や記述方法を設計学的に含む為より包括的かつ実用的であり，その実践が創造活動支援に資する事も本研究では重視して実証を進めている．

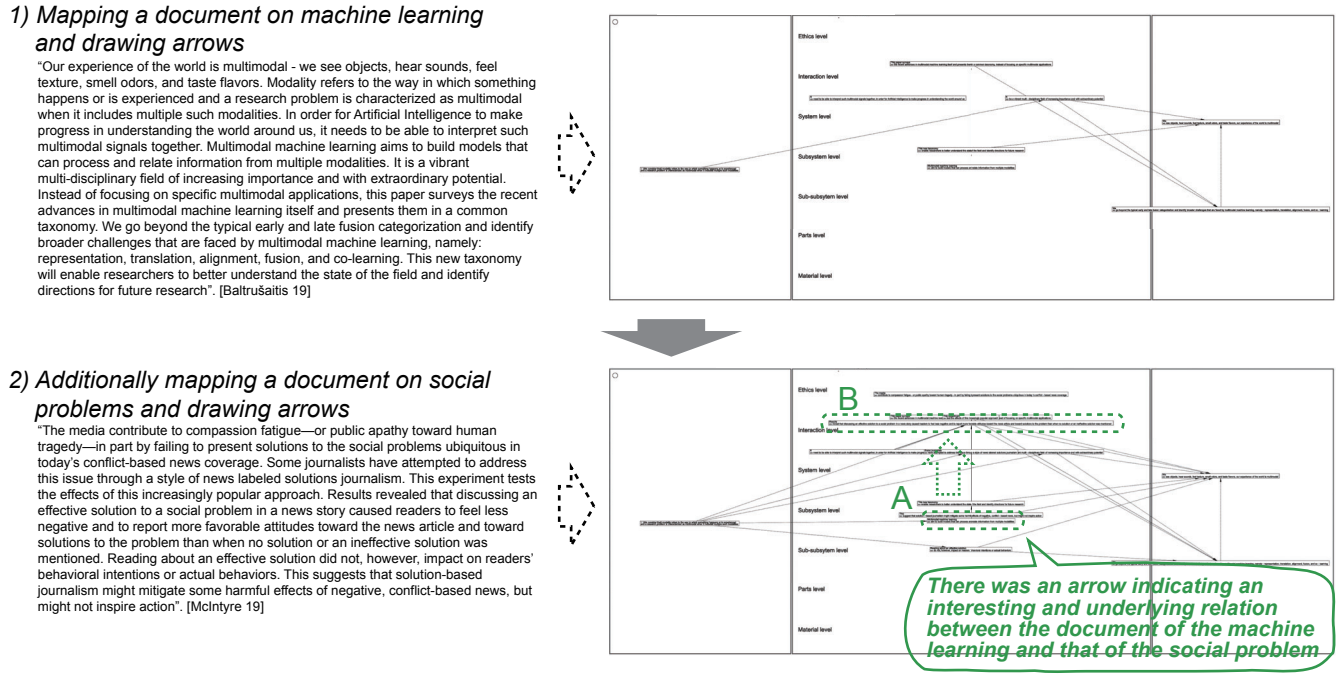


図 5: Image of a process of the generation of items and arrows [Baltrušaitis 19, McIntyre 19].

一般設計学（general design theory）やそれを包含する本格研究（full research）については [吉川 09] に詳しい．一般設計学を提唱した吉川自身,「... 現在，設計の本質は... 依然として明確に記述することはできないでいる」（[吉川 09] p.160）と記載しているが，倫理的設計理論はここに挑戦し，本論文も含め証拠の提示を始めている．更に研究体制全体を纏めた倫理的設計学は，吉川の本格研究に倫理レベル（設計の最終的な目的に対応する領域）に関する考察を有機的に連結できている為，残されていた本質的な領域をそこに拡張できたものとして位置付け可能である．

7. 結論と今後の課題

記述した倫理的設計学の実践例によって，社会的課題の解決に向けて工学が体系的かつ実証的により陽に倫理的設計に貢献できる事を示した．また豊かな研究テーマも設定可能な為，研究者の自発的参加を促すポテンシャルも確認できた．

Dfrome に関する今後の課題としては，最終的には各機能を繋げつつ，またスクレイピングを用いながら時系列でのデータを増やし，4 次元での自動設計までを目指す．ただし，Dfrome は唯一の解ではないので，他のツールとの連携も図る．更に，理論やコンテンツの深化を目指して，人文・社会科学を含む多分野を跨ぐコラボレーションを今後も進める予定である．

謝辞

本研究は JSPS 科研費 JP18K18434 の助成を受けたものです．

参考文献

[Baltrušaitis 19] Baltrušaitis, T., Ahuja, C., and Morency, L.-P.: Multimodal Machine Learning: A Survey and Taxonomy, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 41, No. 2, pp. 423–443 (2019)

[IEEE 17] IEEE, : The IEEE Global Institute for Ethical Consideration in Artificial and Autonomous Systems, Ethically Aligned Design: A Vision for Prioritizing Wellbeing with Artificial Intelligence and Autonomous Systems, Version 2, http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html (2017)

[McIntyre 19] McIntyre, K.: Solutions Journalism, *Journalism Practice*, Vol. 13, No. 1, pp. 16–34 (2019)

[RIKEN] RIKEN, : Center for Advanced Intelligence Project, http://www.riken.jp/en/research/labs/aip/

[Sekiguchi 18a] Sekiguchi, K. and Hori, K.: Organic and Dynamic Tool for Use with Knowledge Base of AI Ethics for Promoting Engineers' Practice of Ethical AI Design, *AI & Society* (2018)

[Sekiguchi 18b] Sekiguchi, K. and Hori, K.: Realization of Organic and Dynamic Creativity Support Tool for Promoting Ethical AI Design, in *Proceedings of the 13th International Conference on Knowledge, Information and Creativity Support Systems*, KICSS2018 (2018)

[Verbeek 11] Verbeek, P.-P.: *Moralizing Technology; Understanding and Designing the Morality of Things*, University of Chicago Press (2011), (邦訳：技術の道徳化：事物の道徳性を理解し設計する, 鈴木 俊洋 訳, 法政大学出版局 (2015))

[吉川 09] 吉川 弘之：本格研究, 東京大学出版会 (2009)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

プロファイリングに関する自主的規律導入支援の実践

Practice of a support for creating self-regulations on profiling

工藤郁子*1
Fumiko Kudo

*1 中京大学経済研究所
Institute of Economics Chukyo University

This report discusses a policy making process about automated processing, including profiling. Our project team provided an analysis of social issues on the use of personal data to evaluate certain aspects related to the individual and published our recommendations on that. The key finding was that some companies and industry groups address these issues from the standpoint of reputation management.

1. プロファイリングをめぐる社会的課題

ICT 技術の進展に伴い、いわゆる「プロファイリング」が生活の中に浸透している。従来から、データに基づいて対象者の趣味趣向を予測するアルゴリズムは、閲覧や購買の履歴に基づいてインターネット広告を表示するターゲティングなどに利用されてきた。さらに近年では、マーケティングを超えて、与信、採用・人事などの場面でも積極的に利用されつつある。例えば、中国では、アリババ・グループ傘下の芝麻信用が、アルゴリズムを用いて個人の社会的信用度をスコア化するサービスを提供している。日本でも、みずほ銀行とソフトバンクが共同開発した社会信用スコアリング・サービス「J.Score」が与信領域において既に運用を開始している。企業の採用活動や労務管理の場面でも、人工知能を用いた求職者や従業員の適性プロファイリングが急速に実装されている。

そのため、アルゴリズムが、個人の自由で主体的な生活に重要な影響を与えうることが懸念されている。また、欧米を中心に、法的・技術的な規律の導入も進んでいる。例えば、2018 年 5 月に適用開始された EU 一般データ保護規則（General Data Protection Regulation, GDPR）は、プロファイリングを「自然人に関する特定の個人的側面を評価するために、特に、当該自然人の職務遂行能力、経済状況、健康、個人的選好、関心、信頼性、行動、位置もしくは動向を分析または予測するために、パーソナルデータを用いて行うあらゆる形式の自動化されたパーソナルデータの処理」と定義し（4 条 4 項）、一定の規制をした。

しかし、日本では、現状において、どのような点に注意すればプロファイリングを実施する（または実施しようとしている）技術者や企業が法的・社会的責任を果たしたといえるかを示す具体的指針がなく、健全な技術開発や事業展開の歩みを妨げている。

2. プロファイリング提言案と中間報告書

2.1 3 つの提言案

筆者が所属するパーソナルデータ＋α研究会では、このような認識の下で、2018 年 5 月より「プロファイリングに関する提言案作成タスクフォース」を設置し、プロファイリングに対する適切な規律のあり方について検討を重ね、以下 3 つの提言案をまとめた [PD+α研 2018]。

1. **【企業：プロファイリングに関するコンプライアンス体制等の整備】**
   与信、採用・人事等にプロファイリングを用いる企業は、プロファイリングの実施、活用等に当たって、法令との抵触を避けるためのコンプライアンス体制を整備すべきである。その際、プロファイリングが、プライバシーのみならず、平等や公正概念との関係でも問題になり得る点に留意すべきである。また、上記企業は、プロファイリングの実施・活用等に伴う社会的・倫理的責任に配慮した体制づくりに努めるべきである。
2. **【業界団体等：プロファイリングに関する自主的な取組みの支援】**
   業界団体・認定個人情報保護団体等は、各企業のコンプライアンスや自主的取組みを支援すべく、ガイドライン等の作成を検討することを推奨する。その際、プロファイリングの対象となる個々の利用者・生活者等（以下「被評価者」という。）と企業の間に立って、利害関係の調整や課題解決を目指すべきである。
3. **【政府：民間主導の自主的な取組みの尊重】**
   政府は、民間による自主的な取組みを尊重すべきである。その取組みの実効性を慎重に見定めながら、企業・業界団体・認定個人情報保護団体等との対話を続けるべきである。そして、プロファイリングによる個人の権利利益の侵害リスクを看過し得なくなった場合には、それを抑止するために必要な立法措置等を検討すべきである。

提言案およびそれに付属する「中間報告書」は、2018 年末からインターネット上で公開された（https://www.shojihomu-portal.jp/nbl1137pc）。2019 年度内に公表予定の最終報告書の作成に向けて、意見を公募している。

2.2 自主的取組みに関するチェックリスト

中間報告書では、プロファイリングの実施に際して企業等が法的・社会的責任を果たすための留意点（「自主的取組みに関するチェックリスト」）をまとめた。

手法としては、以下のような資料を参照することで、論点を抽出した。(a) EU の GDPR、(b) 米国のビッグデータレポート[FTC 2016]、(c) 第 40 回データ保護プライバシー・コミッショナー国際会議において採択された「AI における倫理及びデータ保護についての宣言」[ICDPPC 2018]、(d) Google やソニーグループなどが発表した AI の倫理原則等、(e) プロファイリングやビッグ

データをめぐる報道記事等である。さらに、(f)一般社団法人ピープルアナリティクス&HR テクノロジー協会、一般財団法人日本情報経済社会推進協会、株式会社メルペイへのヒアリングや関係する官公庁との非公式な意見交換も実施した。

チェックリストではプロファイリングの工程を、(1)企画・設計、(2)データの取得、(3)データの加工・分析・学習、(4)実装・運用・評価、という 4 段階に区分し、企業が考慮すべき度合い(優先度)を 3 段階に分けて表示した(「★★★」「★★☆」「★☆☆」の順に優先度が高い)。

チェックリストの見出しは、以下の通りとなった。なお、筆者としては、チェックリストの項目すべてを遵守する必要はないことを強調したつもりである。なぜなら、「プロファイリングを行うことによって何を実現したいのか」という意図・姿勢を示して、ビジネス上の信頼を獲得することが最も重要だからだ。

1 企画・設計段階
1-1 プロファイリング導入の適否判断 ★☆☆
1-2 プロファイリングの実施の有無・利用目的の明確化 ★★★
1-3 エシックス・バイ・デザインの導入検討 ★★☆

2 データの取得段階
2-1 データの適正な取得 ★★★
2-2 データ取得時のユーザー・インターフェース等の工夫 ★☆☆
2-3 データセットの偏向に関する配慮 ★★☆

3 データの加工・分析・学習段階
3-1 分析・評価段階におけるプライバシー侵害リスクの検討 ★★★
3-2 公平性に配慮した学習技術の導入検討 ★★☆
3-3 アカウンタビリティ(答責性)に配慮したモデルの導入検討 ★★☆
3-4 その他の社会的価値への配慮 ★☆☆

4 実装・運用・評価段階
4-1 人間関与原則の検討 ★★★
4-2 答責性への配慮 ★★★
4-3 安全管理措置の実施 ★★☆
4-4 データ内容の正確性の確保 ★★☆
4-5 開示・訂正等の手続きの整備 ★★★
4-6 プロファイリングに関するリテラシー向上 ★☆☆
4-7 被評価者等の反応の精査・検討 ★☆☆
4-8 信用スコアを第三者提供する手続きの確認 ★★☆
4-9 監査可能性への配慮 ★☆☆

2.3 チェックリストの具体例

各項目では、留意点の説明のほか、「ケース」を添えて理解を促した。ここでは、具体例として「4-2」を抜粋する。

(4-2) 答責性への配慮 ★★★

答責性、透明性等への配慮は、実装・運用・評価段階において、より重要性を増す。本研究会は、企業が、決定の過程(決定を導くデータを含む)について、被評価者や他のステークホルダーに非専門用語で説明できるかどうかを検討することを推奨する(【ケース H】参照)。

【ケース H】再犯リスク予測プログラムとアルゴリズムの透明性

米国では一部の州で有罪確定者の再犯リスク予測プログラムが用いられている。ただし、このプログラムのアルゴリズムが明らかにされていないため、バイアスがあるのではないかとの批判もある。ある再犯リスク予想プログラムでは、アフリカ系を白人よりもリスクを高く評価するが、アフリカ系たちが実際に再犯をした割合は半分以下であったという調査も報告されている。データやアルゴリズムのバイアスに関しての検証をするためにも技術システムの透明性や答責性が開発側には求められている。

もっとも、説明内容や程度の妥当性については、いまだ社会的合意はない。また、AI のプロファイリングにのみ答責性を求めるのは均衡を欠くとの指摘もある。これは例えば、人間の採用担当者であれば不採用理由を開示しなくても問題視されない現状をどう捉えるかという問題である。他方で、プロファイリングの過程や評価理由が不透明であれば、能力や信用力に関する評価の基礎になるのかがわからず、将来の行動計画を練れなくなり、自律的・主体的に自らの人生を歩んでいくことが困難になるとの懸念も存在する。

上記を踏まえると、おそらく当面のベンチマークにすべきは、事後(特に苦情対応時)に被評価者が理解し得るような「理由(justification)」を示すことであろう。ここで被評価者に事後提示される「理由」は、実際の決定過程を完全に反映したものである必要はない(別のアプローチとして、【ケース I】)。関連して、決定理由を人間に提示できる「説明可能な AI」(Explainable AI: XAI)の開発動向にも関心を向ける必要があろう。

【ケース I】信頼の構築

ディープラーニングの進展に伴う説明困難性を受け止め、「信頼(trust)」や「信用(fiduciary)」の構築を重視する立場もある。例えば、医薬品の作用機序を十分に理解していなくても、医師や薬剤師の勧めに従って服薬することが一般的である。これは、高い能力・専門性を備えた医師等に任せた方が合理的であるという考えや、国家資格による認証や専門職集団の行動規範(code of conduct)による制約などが制度的に担保されているから信じるに足るとの考えが背景となっていよう。このような信頼構築のための規律構造を AI 化社会においても実現しようとする試みが存在する。

3. 自主的規律導入に関する考察と今後の展開

提言案と中間報告書の作成のためにヒアリングを行った過程で、企業や業界団体等がプロファイリングに関する自主的規律の導入に前向きであることがわかった。おそらく、レピュテーション・マネジメントの観点から有益であるためと見られる。

筆者としては、チェックリスト等が、自主的な取組みの「たたき台」として活用されることを期待している。今後も関係者との情報共有・意見交換を続け、プロファイリングの健全な実装に資する情報や議論を提供していきたいと考える。

参考文献(論文誌と同じスタイルを推奨)

[PD+α研 2018] パーソナルデータ+α研究会: プロファイリングに関する提言案, NBL1137 号, 商事法務, 2018, pp. 64-85.
[FTC 2016] Edith Ramirez et al.: Big Data: A Tool for Inclusion or Exclusion?, FTC Report, 2016, Federal Trade Commission Website
[ICDPPC 2018] CNIL et al.: Declaration on Ethics and Data Protection in Artificial Intelligence, ICDPPC Report, 2018, ICDPPC Website

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

人間中心設計の観点からみたSF映画の中のAI

AI in Sci-fi Movies from the Viewpoint of Human Centered Design

飯塚重善 *1
Shigeyoshi IIZUKA

*1神奈川大学
Kanagawa University

Due to the rapid development of recent AI technology, how society accepts AI and robot has emerged as a realistic and big problem. If not only the robot becomes close to human beings but also human beings are replaced by human beings and human beings depend heavily on it, there is a possibility that the human-centered society should be reconsidered about it. Therefore, in this paper, as materials to consider how we can design from the perspective and the future society where human beings and AI coexist, Sci-fi movies including relevant scenes about "Three principles of robot engineering", "What is a human?", "heart", and ethical issues are picked up. It is thought that studying AI and robot design from such viewpoints have great significance as HCD(Human Centered Design).

1. はじめに

AI(Artificial Intelligence) 社会では，人間と AI の関係性はどのようなものになるだろうか？AI やロボットが世界を大きく変えようとしており，人工知能はあらゆる産業を革新すると期待される．しかし，その影響の範囲が広過ぎるため，具体的な将来像をイメージすることは難しい．そこで我々がイメージすることを助けてくれるのが SF(Science Fiction: Sci-fi) 作品である．近年の SF 映画では，未来社会やそこに登場する事物の表現にリアリティが求められることから，登場する大道具・小道具のすべてを圧倒的にリアリティ高く製作し，また，電子情報工学の専門家，未来学者の意見を積極的に取り入れられている．そこで本稿では，人間と AI が共存する未来の社会を，どのような観点から，どのようにデザインしていくことができるかを，SF 映画を素材として，人間中心設計 (Human Centered Design, 以下 HCD) の観点から検討する．

2. 本稿における「ロボット」と「AI」の考え方

ロボット (Robot) という言葉は，1920 年にチェコスロバキアで発表された戯曲，R・U・R(Rossum's Universal Robots) で使われた，チェコ語の Robota(労働者) から「a」を取った造語である．劇中で，ロボットは労働の機能だけをもった人造人間として登場した．1927 年ドイツで発表された映画『メトロポリス』(*Metropolis*, 1927 年公開）で女性の姿をした金属のロボットが登場した．R.U.R とメトロポリスがヒットし,「ロボット」＝「人間のために働く人型の機械」という認識が人々の中に根づいていったといわれている．そして，現在ロボットと呼ばれているものの形や機能はさまざまである．例えば製造現場で働くロボットは，人間が教えたことを正確に繰りかえす機能が主で，形は人間の腕のような姿をしているものもある．また，災害地で救助活動をするロボットには，遠隔操作で障害物を除いたりする操縦型ロボットや，移動中に目の前に物があったらよけて人を探すという自ら状況を判断し行動する自律型のものもある．その他，重いものを持ち上げる時，人間の動きを補助してくれる外骨格型のロボットスーツなどもあり，形状や機能，種類は，その用途によって多岐に渡る．このように，実際にはロボットの厳密な定義はないようである．ただし，大まかにはロボットは，あらかじめプログラムされた動作を正確におこなうことが基本となっており，決められたこと以上のことを自己判断で進めることはできず，人間が遠隔操作するものなど，知的処理の部分は人間が担っていた．一方 AI は，"人間の脳がおこなっている知的な作業をコンピュータで模倣したソフトウェアやシステム" であり，"人間の使う自然言語を理解したり，論理的な推論をおこなったり，経験から学習したりするコンピュータプログラミングなどのこと" を指す．その目指す形態は自動化であり，人の脳と同じような仕組みとなるニューラルネットワークを利用し，自ら学習して精度を高めることができる仕組みをもっている．将来的には両方の機能がインテグレートされ,「ロボット」が人工知能を搭載することで自分で周囲の状況を捉え，どう行動すべきかを考え，判断して行動するようになると予想される．実際，SF 映画でもこうした「自律化 (Autonomy)」したロボットを描いていることが多い．そこで本稿では,「自律ロボット」を対象とすることで,「ロボット」と「人工知能」を同義として扱う．

3. SF 映画における AI の描写

まず表 1 に，いくつかの代表的な SF 映画に登場するロボットについて，それぞれの特徴をまとめた．以下では，この中から，特筆すべきものを抽出して具体的に紹介する．

映画『オートマタ』(Automata, 2014 年公開）は，人類存続の危機的状況下での AI と人類の未来や共存，主従関係を描いており，ロボット三原則をテーマとして，これが破られている．また，映画『アイ，ロボット』(I, ROBOT, 2004 年公開）は，アシモフの作品に "ヒントを得た" として，ロボットは三原則に従いながら世界を乗っ取ることができる，という発想が映画のストーリーの中心に据えられている．ロボットが実用化されるという近未来を具体的に想定し，その設計にあたって，人間とロボットの共存関係はいかにあるべきかが重要なテーマであり，人間はそれを真剣に考えなければならないという問題提起をした作品である．この作品では，2035 年，米国シカゴを舞台とし，家庭用ロボットが普及し,「ロボット三原則」によって人間とロボットは平和に共存している．しかし，人間に

連絡先: 飯塚重善，神奈川大学，〒 259-1293 神奈川県平塚市土屋 2946，iizuka@kanagawa-u.ac.jp

表 1: 代表的な SF 映画に登場するロボット

作品名	ロボット名	役割	外観	三原則	主な特徴
A.I.	デイビッド	家族 (子供)	人間	不明	人間と同じ愛情を持つ
2001 年宇宙の旅	HAL	司令	無 (赤 LED)	不明	
リアル・スティール	アトム	格闘用	人型ロボット	不明	動きを真似るシャドー機能付
チャッピー	チャッピー	手下 (兵器)	人型ロボット	不明	高性能 AI を半分取り入れた攻撃ロボット
エクス・マキナ	エヴァ	試作品	ほぼ人間	不明	表情や会話の自然さはそれまでの AI をはるかに凌駕
オートマタ	クリオ	娼婦	人型ロボット	2 制御機能	「生命体への危害の禁止」「自他のロボットの修正 (改造) の禁止」実装
アンドリュー	アンドリュー	家事用	人型→人間	実装	創造性を持つ
アイ，ロボット	サニー	家事用	人型ロボット	本来は実装	サニーだけが三原則を破る?自我のある

危害を加えることができないはずのロボットによる殺人事件とロボットの反乱が描かれる．これは，“ロボット工学三原則には重大な欠陥（問題点）がある” ことを含んでおり，これを突き詰めると，“「人間」をどのように定義するか” という点となる．

そこで，“人間とは？” “どこまでが人間？” “どこからが人工物？” についてみてみる．映画『アンドリュー NDR114』（Bicentennial Man, 1999 年公開）は，ある一家に購入された家事用ロボットの NDR114（「アンドリュー」と名付られる）が，家族の一員として生活を営むうち，次第に人間への憧れを抱くようになり，200 年にわたって人間の心を育んでゆく過程が描かれている．この映画において，事故や病気で損傷した体の一部をアンドリューと同じ人工臓器で置き換えた人間も多い中，最終的に，血液の影響で老化の進んだ外観となったアンドリューは，人類法廷で，“死” を迎えることでようやく人間であることを認められる．この映画から，不死身の人工生物を目の当たりにする人類は，その生死観や倫理観に大きな影響を受けることが想像される．また，ロボットの心が，少なくとも発現する機能の面でヒトの心とほとんど同じであるということが確認され，一般に受け入れられるようになる必要がある．

一方で，映画『ブレードランナー』（Blade Runner, 1982 年公開）では，ロボットに代わる「レプリカント」と呼ばれる人造人間が描かれている．レプリカントは，見た目だけでは人間と区別がつかないという前提であり，その判定にチューリング・テストが用いられている．本来のチューリング・テストで判定するのは人間の心であって，姿形は関係ない．その点からすると，映画『ザ・マシーン』（The Machine, 2013 年公開）中では，独自の AI テストシーンが描かれている．ちなみに，人間のように何でもこなせる汎用人工知能（汎用 AI）の判定基準として,「コーヒーテスト」提案されている [ニュートン 18]．これは,「間取りを知らない家に上がってコーヒーを入れる」ということができれば汎用 AI だということである．Apple 社の創業者の一人であるスティーブ・ウォズニアック (Steve Wozniak) が “知らない家に上がってコーヒーを入れる機械を我々は決してつくれないだろう” と予言したことに由来することから，「ウォズニアック・テスト」とも呼ばれる．

こうした人間のような人工物（ロボット・AI）を創り上げるためには，こうした人工物に “心” を持たせることにも検討が必要である．“コンピュータに心を持たせることは可能か” という命題は，近年，ロボット工学，認知科学，情報科学など，様々な分野で盛んに議論されている．心とは，知的情報処理に加え，基本的な情動やより高度の感情，さらには価値観などの行動規範を含むものと考えられる．また，心とは,「知」「情」「意」「記憶と学習」「意識」という 5 つの働きを総合した情報処理機能であるといわれる [松本 96] が，本稿ではこれら 5 つを明確に分けることはせず，大まかに「情」もしくは「意識」と捉えることとする．既出の『アンドリュー NDR114』に出てくるアンドリューの他，映画『チャッピー』（CHAPPiE，2015 年公開）では，感情を持ったり，人間の知性を模倣する AI 搭載ロボット “チャッピー” が，インストール直後は純真無垢でまるで赤ん坊のようであったが，人間からの教育を受けて徐々に “人間的” なことを学習していく姿が描かれている．人間も赤ん坊のうちは動物を判別できないが，成長とともに学習すれば判別できるようになる．AI も人間のように学習することで成長する．

そしてさらに一歩踏み込んだ「情」として「恋愛感情」が描かれている作品もある．AI と人とのコミュニケーションが日々進化し，人間が AI に恋をするような未来が来るのかもしれない．映画『エクス・マキナ』（Ex Machina，2015 年公開）では，AI を搭載した女性型ロボットに主人公が日増しに惹かれていく様が描かれている．ただし実際には，主人公の一方的な好意と思われる．また，映画『her/世界でひとつの彼女』（Her，2013 年公開）は，人格を持つ最新の AI 型 OS をインストールし生活を共にすることで恋愛関係が生まれていくという物語である．この映画の OS は，一貫した音声認識 AI として明確なビジュアルは一切描かれていない．その状況で，声だけの AI 型 OS と男性が互いに惹かれている（OS 側の本当の気持ちについては定かではないが）．さらに，既出の『アンドリュー NDR114』では，人間の心や感触，痛覚などの神経系器官を手に入れ，限りなく人間に近い状態になったアンドリューが人間の女性と愛し合う．AI は，ユーザーの情報を豊富に持つことで，より身近な存在になる．コミュニケーション能力が発達した AI との濃密なコミュニケーションによって，人間は AI に対し恋のような特別な感情を抱くことはある意味，当然の流れなのかもしれない．そして，いわゆる「心」や「意識」を AI が持っていないにしても,（『her/世界でひとつの彼女』中の主人公のように）人間はおそらく誤認（勘違い?）することは容易に起こりうると考えられる．現在研究されている AI にも「意思」のように感じられる事例もあるが，それは人間による意思のエミュレーションに過ぎない．人間がなぜ個々に独立した意思や精神，感情を持つのかという「脳科学」的な視点から研究をおこなうことで，AI に意思を持たせられるかもし

れない．

ただし,「意思」を持ったロボットが出現し高度化することの問題点としては，古来 SF で取り上げられ続けたように，ヒトを脅かし支配するロボットの恐怖が挙げられる．本来，人工物（ロボット）は設計者が定義した範囲内での心を呈するのみであるから，例えば抑圧された結果犯罪に至るような心理発展のアルゴリズムを埋め込まなければ，そのような犯罪心理は生じ得ないはずである．しかし，レイ・カーツワイル (Ray Kurzweil) による予言を適用すれば，AI が自らを規定しているプログラムを自身で改良するようになると，永続的に指数関数的な進化を遂げる．この結果，ある時点で人間の知能を超えて，それ以降の発明などはすべて人間ではなく AI が担うようになり，それ以降の進歩を予測できなくなる (人間がいくら考えても想像ができないレベルに達する)．これがいずれは，“AI が世界を支配する” という域にまで到達すると考えれば，“ヒトを脅かし支配するロボットの恐怖” に繋がることになる．映画『2001 年宇宙の旅』（2001: A Space Odyssey，1968 年公開），映画『マトリックス』（The Matrix，1999 年公開），そして『アイ，ロボット』では，まさにこの状況が描かれ,『オートマタ』では，こうした “人間を超える” 事態に至らないよう，ロボットの自己修復禁止機能が組み込まれている（作中では，自己修復禁止を解除されたロボットが数体現れる）．

例えば，人工知能開発段階では，開発方向の判断において倫理という社会規範 (Social Norm) に依拠するところが大きい．方向性を誤らないためにも，倫理という規範 (Norm) について議論を重ねる必要がある．

「モラルジレンマ」の例として，哲学者や心理学者が道徳や倫理について考える時に用いられる例題「トロッコ問題（Trolley Problem）」が挙げられる．AI の世界で，まさにこの「モラルジレンマ」の問題を抱えている顕著な事例が自動走行車である．MIT の研究チームは,「自動運転車を用いた人工知能の道徳的な意思決定に関して，人間の視点を収集するためのプラットフォーム」をオンライン上に公開している [MIT 16]．この「トロッコ問題」に近い内容が『アイ，ロボット』中でも描かれている．作中では，交通事故の救助に現れたロボットが，少女より生存率の高かった主人公を優先して救い，少女を見捨てた（以来，主人公は数字で物事を判断するロボットを毛嫌いするようになった）という状況が描かれている．コンピューターがあらゆる情報をもち，一瞬で判断する，ということであれば，少女の環境，将来性などを勘案し，主人公と比べて社会的にどちらが有用か等から総合的に判断する必要性も考えられる．

また，AI を使って犯罪を予知する仕組みも考えられている [日経 18]．これは，街中に設置された防犯カメラや Web カメラの映像を用いて全市民を監視し，収集したデータをコンピュータが分析し，犯罪やテロなどの異常を事前に察知するというものである．映画『マイノリティ・リポート』（Minority Report，2002 年公開）では，事件を予想したのは「プリコグ」と呼ばれる予知能力者だが，現実でその役を担うことになるのは AI である．

実際，日本では 2018 年 1 月，神奈川県警が AI を使って事件・事故の発生を予測する新システムの導入を検討していると報じられた [産経 18]．この報道によると，連続発生した事件の容疑者が同一かどうかを分析したり，容疑者の次の行動を予測したりする他，事件事故が起きやすい時間帯と場所を確率で示すシステムの構築を目指すという．また海外でも，例えば 2017 年 8 月，米・シカゴ市警察がサウスサイド地区で犯罪予測システムを導入し,「凶悪事件が激減した」という成果を発表した [Engadget 18]．こうした「犯罪予知システム」については，どのようなアルゴリズムで判断を出しているのか，その精度はどれほどのものなのかが分からないままである．また，こうしたセキュリティ領域への活用が進む中で，議論の一つになるのがプライバシーの問題である．国家を脅かすテロや凶悪犯罪，または我々の身近に潜むサイバー犯罪などを検知するために，我々の個人情報をはじめ，インターネット上の検索・閲覧履歴や行動，サイトへのログイン履歴や所在地など，プライバシーに関わる情報が必要になる場合が考えられる．AI によって人間の分析能力を超えた莫大なデータを，過去何十年にもわたって解析可能となった未来において，我々は自分たち自身のプライバシーを守ることができるのか? そのような未来において，人々の安全を守るために，個々のプライバシーは犠牲にされるべきなのか? 「安全保障 or 市民のプライバシー」という究極の選択に，AI はどのように対処していくべきなのか?といった観点について考えていく必要がある．

4. おわりに

最近の AI 技術の急速な発展により，AI を搭載した高度なロボットを如何に社会的が受容するかが現実的かつ大きな課題として浮上してきている．AI を搭載したロボットは，自身で学習し，ときには人間の知性の及ばない思考や判断すら可能といわれる．すなわち，近い将来，AI が人間の能力を超える “シンギュラリティ” が実現すると予言されている．しかしながら，現在のロボット開発では技術力の向上だけが先行し，ロボットが人間社会に入り込んだときの影響が考慮されていない，という問題点も指摘されている．そうしたロボットが人間にとって身近な存在になるだけではなく，人間の役割を代替し，人間がそれに大きく依存していくことになれば，これまでの人間中心の社会のあり方に再考が迫られることになる可能性が考えられる．例えば，癒しロボットとよばれる範疇のペット型ロボットが安価に購入できるようになり，人間に心理的影響を及ぼすようになってきている．逆にそれらの使い道として，心のケアやリラクゼーション，老人の機能回復なども考えられている．今後これらの他に別な形態のロボットが人間生活に深く入り込んでくる可能性は十分に考えられる．しかしながら，それらが人間，環境，社会にどの様な影響を及ぼすかについて，十分な議論がおこなわれているとはいえない．我々は，どういう社会を作りたいのか?ロボットが人間生活に浸透する社会を見越し，人間とロボットの両方が生活しやすい環境とは何か，社会規則を含めた社会システム・インフラをどのようにすべきか，災害などの非日常という特殊環境におけるロボットの役割などを工学分野のみならず，心理学，医学などの異種分野からより良い解決策を追求することにより，技術だけで進められるロボット開発に方向性を与えることができ，ロボット開発の発展に寄与できると考えられる．

そこで本稿では，人間と AI が共存する未来の社会を，どのような観点から，どのようにデザインしていくことができるかを検討する素材として,「ロボット工学三原則」，“人間とは?” “どこまでが人間?” “どこからが人工物?”，人工物に “心”（情，意識），そして倫理的問題について，該当する（その状況を具体的に描いている）シーンを含む SF 映画を紹介した．こうした観点からの AI，ロボットデザインを検討することは，HCD としても大きな意義があると考える．

人間に近いロボットは，ロボット自身が自律的に判断し行動する点が車などの機械とは異なる．このため，ロボットと共生する上での倫理的および法的な基準を設けることも重要な課題である．さらには，そもそも心を持つように見える AI を作っ

て良いのか，AI が知財を生み出す場合の権利，人工知能システムが社会に広がったときの不具合の問題，倫理や社会制度の議論が必要になる．これらの点が曖昧なまま人間に近いロボットが人間社会に展開されていった場合，我々の日常に混乱が起きることは容易に想像できる．シンギュラリティを待望もしくは警戒する一方で，AI とどのように共存していくかという議論はいまだ十分とはいえない．

黒崎は，著書 [黒崎 98] の中で『ロボットは主体にはなりません。あくまで、それらは人間が作ったものであって、もし倫理が必要であるとすれば、ロボットを製造する倫理、ロボットを使う倫理であって、決してロボットの倫理ではありません。倫理が必要なのは、〈作り・使う〉人間の側です。』と述べている．一般的に，AI を含む情報システムの設計では，設計者が何を問題設定しているかが重要である．さらには，基本的にその設計者の倫理観，よかれと思うことが拠り所になっている．もちろん人間中心設計（Human Centered Design）において，利用者が果たしてどんな文脈で，どういうふうに絞ればわかるだろうか，をユーザー視点で考え，最善のものを作ろうとはしている．あるサービス仕様は，顧客，ユーザーにとって正しいかもしれないが，それではスポンサーがつかなくなるかもしれない．するとサステナブルに実運用できなくなくなり，結局は継続できなくなってしまい，ユーザーに迷惑をかけてしまうかもしれない．こうした意味から，実運用とのバランスは考えなければいけない．ただし，そこで重要なのは，できる限りユーザーのためにフェアであろうとする努力目標を持つことである．

参考文献

[ニュートン 18] ニュートンプレス社: Newton 2018 年 6 月号特集「人工知能は人より賢くなれるか」, 第 38 巻, 第 6 号, pp.44-57 (2018)

[松本 96] 松本元: 脳・心・コンピュータ, 丸善 (1996)

[MIT 16] MIT Media Lab.: モラル・マシン (2016) http://moralmachine.mit.edu/hl/ja (最終アクセス 2018.11.13)

[日経 18] 日本経済新聞社: 明日の安心・安全 AI で高度監視, 2018 年 9 月 13 日, 日経新聞朝刊 14 面 (2018)

[産経 18] 産経新聞社: 神奈川県警犯罪発生 AI で予測東京五輪までの試用目指す, 2018 年 1 月 29 日, 産経新聞東京朝刊 27 面 (2018)

[Engadget 18] https://japanese.engadget.com/2017/08/11/hunchlab/ (最終アクセス 2018.11.13)

[黒崎 98] 黒崎政男: となりのアンドロイド – 哲学者クロサキの憂鬱, NHK 出版 (1998)

参考映画 (タイトル五十音順)

『2001 年宇宙の旅』(*2001: A Space Odyssey*)
監督：スタンリー・キューブリック
公開年：1968 年
製作国：イギリス/アメリカ
配給：メトロ・ゴールドウィン・メイヤー

『アイ，ロボット』(*I, Robot*)
監督：アレックス・プロヤス
公開年：2004 年
製作国：アメリカ
配給：20 世紀フォックス

『アンドリュー NDR114』(*Bicentennial Man*)
監督：クリス・コロンバス
公開年：1999 年
製作国：アメリカ
配給：ソニー・ピクチャーズ エンタテインメント

『エクス・マキナ』(*Ex Machina*)
監督：アレックス・ガーランド
公開年：2015 年
製作国：イギリス
配給：ユニバーサル映画

『オートマタ』(*Automata*)
監督：ガベ・イバニェス
公開年：2014 年
製作国：スペイン・ブルガリア合作
配給：松竹

『ザ・マシーン』(*The Machine*)
監督：カラドッグ・W・ジェームズ
公開年：2013 年
製作国：イギリス
配給：熱帯美術館

『チャッピー』(*CHAPPiE*)
監督：ニール・ブロムカンプ
公開年：2015 年
製作国：アメリカ
配給：ソニー・ピクチャーズエンタテインメント

『her/ 世界でひとつの彼女』(*Her*)
監督：スパイク・ジョーンズ
公開年：2013 年
製作国：アメリカ
配給：ワーナー・ブラザーズ映画（アメリカ）/アスミック・エース（日本）

『ブレードランナー』(*Blade Runner*)
監督：リドリー・スコット
公開年：1982 年
製作国：アメリカ・香港合作
配給：ワーナー・ブラザーズ映画（アメリカ）

『マイノリティ・リポート』(*Minority Report*)
監督：スティーブン・スピルバーグ
公開日：2002 年
製作国：アメリカ
配給：20 世紀フォックス映画

『マトリックス』(*The Matrix*)
監督：ラリー・ウォシャウスキー/アンディ・ウォシャウスキー
公開日：1999 年
製作国：アメリカ
配給：ワーナー・ブラザース

『メトロポリス』(*Metropolis*)
監督：フリッツ・ラング
公開日：1927 年
製作国：ドイツ
配給：UFA（ドイツ）/パラマウント映画（アメリカ）/松竹

『リアル・スティール』(*REAL STEEL*)
監督：ショーン・レヴィ
公開日：2011 年
製作国：アメリカ
配給：タッチストーン・ピクチャーズ（アメリカ）/ウォルト・ディズニー・ピクチャーズ（アメリカ）

Organized Session | Organized Session | [OS] OS-7

[4E2-OS-7a] AIの法学の応用(1)

佐藤 健（国立情報学研究所）、新田 克己（国立情報学研究所）

Fri. Jun 7, 2019 12:00 PM - 1:40 PM Room E (301A Medium meeting room)

[4E2-OS-7a-01] (Invited talk) Offender profiling for criminal investigation and application of machine learning

〇Zaitsu Wataru[1] （1. Forensic Science Laboratory, Toyama Prefectural Police Headquarters）

12:00 PM - 12:40 PM

[4E2-OS-7a-02] Extracting Important Sentences with Random Forest for Statute Summarization

〇Yasuhiro Ogawa[1,2], Michiaki Satou[2], Takahiro Komamizu[1,2], Katsuhiko Toyama[1,2] （1. Information Technology Center, Nagoya University, 2. Graduate School of Informatics, Nagoya University）

12:40 PM - 1:00 PM

[4E2-OS-7a-03] Legislation Supporting Systems by AI

〇Tokuyasu Kakuta[1] （1. Chuo University）

1:00 PM - 1:20 PM

[4E2-OS-7a-04] People's Expectation and Attitude toward AI-assisted Courts

〇Shozo Shozo OTA[1] （1. The University of Tokyo）

1:20 PM - 1:40 PM

©The Japanese Society for Artificial Intelligence

12:00 PM - 12:40 PM (Fri. Jun 7, 2019 12:00 PM - 1:40 PM Room E)

[4E2-OS-7a-01] (Invited talk) Offender profiling for criminal investigation and application of machine learning

〇Zaitsu Wataru[1] （1. Forensic Science Laboratory, Toyama Prefectural Police Headquarters）

Keywords: support of criminal investigation, offender profiling, machine learning

犯人が不明である未解決事件の早期解決を目的として、1970年代米国 FBIより登場した犯罪者プロファイリング、中でも犯人像推定の方法論は精神医学や臨床心理学の知識を基にした分析手法から、1980年代には多変量データ解析を駆使したものと変遷し、今現在の主流となっている。そのような中、近年は我が国においても機械学習を応用した犯人像推定研究が始まる兆しがみられている。また、講演者においては、テキストマイニング技術を用いて、文章情報から犯人像を推定するといった研究も行っている。

本セッションでは、従来の手法の変遷と現在の犯罪者プロファイリングの紹介、そして今後の展望について議論していくこととする。

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

法律の要約のためのランダムフォレストを用いた重要文抽出

Extracting Important Sentences with Random Forest for Statute Summarization

小川 泰弘 *1*2 OGAWA, Yasuhiro　佐藤 充晃 *2 SATOU, Michiaki　駒水 孝裕 *1*2 KOMAMIZU, Takahiro　外山 勝彦 *1*2 TOYAMA, Katsuhiko

*1名古屋大学 情報基盤センター Information Technology Center, Nagoya University　*2名古屋大学大学院情報学研究科 Graduate School of Informatics, Nagoya University

Our purpose is to provide an automatic summarization for Japanese acts and we propose a sententence extraction method with Random Forest. While the traditional automatic summarization methods have used the information of summarizing source data, in recent years, the methods based on machine learning use the summarization results. However, in such a method, the amount of learning corpus is small, especially in Japanese text. In this research, we solve this problem by using "Outlines of Japanese Statutes," which are official summaries of statutes published by the Japanese government. Furthermore, we show that the sentence extraction method with Random Forest has higher performance rather than with decision trees or with support vector machines.

1. はじめに

法令は社会の設計図であり，法令を理解することにより，権利や義務などの国民の生活に関する多くの情報を得ることができる．しかし，法令は，1 文が長かったり，構文が複雑であったりするため，一般に読みにくく，理解しにくい文書であるとされている．また細部を規定する条文などもあるため，文書全体の分量が多い場合があり，その内容の把握は容易ではない．

そこで本研究では，法令理解の促進に貢献するために，法令の要約を提供することを目標とする．自動要約については，ルールに基づく手法や，学習データを利用する機械学習，最近ではニューラルネットワークを用いた手法も提案されている．

機械学習では，学習データが必要になる．当初は人手によって要約が作成されてきたが，その作成コストの大きさから大量のデータは用意できなかった．近年では，新聞記事の見出しを要約と見做すことにより大規模なコーパスが作成されている．

法令の自動要約についても，こうした学習データが必要となるが，幸いなことに，法令に関しては「法令のあらまし」と呼ばれる文書が，昭和 48 年 4 月から政府により作成されている．ここで新規制定法律に注目すると，平成 31 年 1 月までに 1,156 件のあらましが作成されており，従来研究よりも多くの文書がコーパスとして利用可能である．

そこで本研究では，まず法令のあらましを利用して要約用の学習データを構築し，次に，それを用いた機械学習による自動要約に取り組む．元の法令と法令のあらましを比較すると，文字数や文数の制約がなく，法令の 1 文がそのまま取り出されていることが多い．そこで，分類器を用いた重要文抽出により自動要約を実現する．

重要文抽出に分類器を用いた先行研究としては，決定木を用いたもの [1, 2] や，SVM を用いたもの [3] があるが，本研究では，それらよりも高い分類性能が期待できるランダムフォレストを用いて，その有効性を検証する．なお，予備実験としてニューラルネットワークを用いる手法も試したが，学習データの量が不十分であったためか，充分な性能は得られなかった．

本論文の貢献は以下の通りである．

- 従来よりもサイズの大きな要約用コーパスの作成
- 重要文抽出におけるランダムフォレストの有効性の確認

連絡先: 小川泰弘，名古屋大学，yasuhiro@i.nagoya-u.ac.jp

2. 関連研究

自動要約については，文短縮 [4, 5]，重要文抽出 [6, 7, 8, 9, 10, 11, 1, 2, 3]，文融合 [12] などの手法が研究されてきている．

文短縮は，1 文を短くする要約手法であり，構文木を枝刈りする手法 [4] などが提案されてきた．重要文抽出は，原文書から重要な文を抽出する要約手法であり，何らかの素性を基に原文書中の文に重要度を付与し，重要度の高い順に文を抽出する手法が提案されてきた．具体的には，文の位置情報を利用する手法 [6] や，文間の類似度を利用する手法 [7]，それらに加えて重要語など複数の素性を同時に利用する手法 [8, 9] などがある．文融合は，複数の文を 1 文に融合する要約手法であり，構文木の情報を利用する手法 [12] などが提案されてきた．

ここまで紹介した手法は学習データが不要であるが，近年，機械学習などの学習データを利用した要約手法が提案されている．特に重要文抽出においては，ベイズ分類器を用いる手法 [10] や，文の重要度を重回帰分析で決定する手法 [11]，分類器である決定木を用いる手法 [1, 2] や，分類器であるサポートベクターマシン (SVM) を用いる手法 [3] などが提案されている．さらに現在は，ニューラルネットワークを用いた手法 [13] などが提案されている．

3. 法令のあらまし

法令のあらましは，国民の法令理解の促進を目的とし，昭和 48 年 4 月から官報に掲載されている．あらまし作成の対象となるのは，法令のうち，法律・政令・条約であるが，本稿では法律を対象とする．

ここで，政府があらましを作成するのであれば，法律の要約は不要と思われるかもしれないが，そうではない．まず，あらましは昭和 48 年 4 月以前に公布された法律には存在しない．平成 30 年 10 月 1 日時点で有効な法律は 2,261 件あり，これは前述のあらましが存在する新規制定法律 1,156 件の約 2 倍である．しかも，法律はしばしば改正される．実際，上記の 1,156 件中，840 件以上が改正されており，そうした法律のあらましは，現在の法律の要約になっていない可能性がある．よって，現在有効な法律の概要を知るために，自動要約は必要である．

本節では，法令のあらましに対する分析と，それに基づく要約用の学習データの作成について述べる．

法律	あらまし
・・(略)・・ (基本方針) 第四条　主務大臣は、有明海及び八代海の海域の特性に応じた当該海域の環境の保全及び改善並びに当該海域における水産資源の回復等による漁業の振興に関する施策を推進するため、有明海及び八代海の再生に関する基本方針(以下「基本方針」という。)を定めなければならない。 2　基本方針においては、次に掲げる事項を定めるものとする。 一　有明海及び八代海の海域の環境の保全及び改善並びに当該海域における水産資源の回復等による漁業の振興に関する基本的な指針 二　次条第一項の県計画の策定に関する基本的な事項 3　主務大臣は、基本方針を定めようとするときは、あらかじめ、関係県の意見を聴くとともに、関係行政機関の長に協議しなければならない。 4　主務大臣は、基本方針を定めたときは、遅滞なく、これを公表するとともに、関係県に通知しなければならない。 5　主務大臣は、情勢の推移により必要が生じたときは、基本方針を変更するものとする。 6　第三項及び第四項の規定は、基本方針の変更について準用する。 第五条　関係県は、基本方針に基づき、・・・を定めるものとする。 ・・(略)・・	・・(略)・・ 3　基本方針 (一)　主務大臣は、有明海及び八代海の海域の特性に応じた当該海域の環境の保全及び改善並びに当該海域における水産資源の回復等による漁業の振興に関する施策を推進するため、有明海及び八代海の再生に関する基本方針(以下「基本方針」という。)を定めることとした。**(第四条第一項関係)** (二)　基本方針においては、有明海及び八代海の海域の環境の保全及び改善並びに当該海域における水産資源の回復等による漁業の振興に関する基本的な指針並びに4の(一)の県計画の策定に関する基本的な事項を定めることとした。**(第四条第二項関係)** (三)　主務大臣は、基本方針を定めようとするときは、あらかじめ、関係県の意見を聴くとともに、関係行政機関の長に協議しなければならないこととした。**(第四条第三項関係)** 4　県計画 (一)　関係県は、基本方針に基づき、当該関係県の区域内の指定地域について、有明海及び八代海の海域の特性に応じた当該海域の環境の保全及び改善並びに当該海域における水産資源の回復等による漁業の振興に関し実施すべき施策に関する計画(以下「県計画」という。)を定めることとした。**(第五条第一項関係)** ・・(略)・・

図 1: 法律とそのあらましの対応例（有明海及び八代海を再生するための特別措置に関する法律，平成 14 年法律第 120 号）

3.1　法令のあらましの分析

法令のあらましが，どのように作られているかを分析する．まず，法律とそのあらましの対応の例を図 1 に示す．この図において，色付けした法律文とあらまし文が対応している．対応する文を見ると，図 1 中の橙色や青色の箇所の文は，法律とあらましでほぼ同じである *1．実際，5.1 節の実験用に作成した学習データ中の正例（あらまし文が対応する法律文）8,793 文のうち，約半数の 4,421 文において，法律文とあらまし文がほぼ同一であった．このことから，法律のあらましの作成においては，重要文を抽出することが最初の課題となることが分かる．もちろん，図 1 中の緑色の箇所が示すように，法律文 3 文が一つのあらまし文に融合される場合もある．しかし，そうした文融合の手法を適用する場合でも，まずは融合の対象となる文を抽出する必要がある．以上のことから，本稿では，法律の自動要約のための重要文抽出を目的とする．

次に，重要文抽出のためにどのような手法を取るかを検討する．その際，法律の自動要約には，従来の自動要約の研究とは大きく異なる点がある．それは，従来の研究の多くでは，生成する要約文に文字数や文数の制約があったが，法令のあらましでは，そのような制約は観測されない点である．表 1 は，本稿の実験に使用するテストデータである．各法律に含まれる文数と，あらまし用に抽出された文数から，抽出率を計算している．この表に示すように，あらましにおける抽出文数・抽出率はばらつきが大きく，全文を抽出する場合もあれば，平成 26 年法律第 49 号のように 2 割しか抽出しない場合もある．平均は約 70%であり，従来の自動要約と比べて抽出率は大きい．以上を踏まえ，本研究では分類器を用いた重要文抽出を試みる．

3.2　学習データの構築

分類器を重要文抽出に用いるためには，重要文として抽出すべきか否かを正例・負例としてラベル付けした文からなる学習データが必要となる．そのためには，あらましの各文がどの法律文と対応するかを判定しなければならない．幸いなことに，法令のあらましにおいては，図 1 中で赤字で示した「(第四条第一項関係)」のように，どの法律文を要約したかを示す関係規定 [14] と呼ばれる記述が存在する．

本研究では，この関係規定の情報を利用して，法律文とあらまし文の対応付けを自動的に行った．図 1 の色分けがその結果である．ここで，図 1 において橙色および青色の箇所は，法律文 1 文があらまし文 1 文に要約されている．よって，これらの法律文を正例として学習データに加えた．一方，法律文において色付けされていない文は，対応する文があらましに出現していない．すなわち重要文として抽出されなかったことを示す．そこで，こうした法律文を負例として学習データに加えた．なお，図 1 において緑色の箇所は，法律 3 文があらまし文 1 文と対応している．このように複数文の法律文があらまし文 1 文と対応する場合，複数文がまとめて 1 文に要約されている場合もあれば，実際には複数文中の 1 文だけが要約に使われている場合もある．つまり，こうした法律文は関係規定の情報だけでは正例になるか負例になるか判定できない．よって，今回は学習データには加えなかった．

4.　ランダムフォレストを用いた重要文抽出

本研究では，ランダムフォレスト [15] を用いた重要文抽出方法を提案する．ランダムフォレストは，アンサンブル学習の一種であり，複数の決定木から構成される．

ランダムフォレストの学習に用いる素性として，予備実験では，文の出現位置などの文書構造の情報も用いたが，性能向上が見られなかった．よって本研究では，単純に文中での単語の出現の有無を用いた．ただし，すべての単語を用いると素性の数が多くなるため，名詞に限定し，その上で，あらまし文中での出現回数が 500 位以上のものを用いた．これは，要約の結果としてあらましに残った単語が，文抽出の手がかりになると

*1　厳密には，文末に「こととした」を付加したり，「第四条第一項」を「4 の（一)」と表記するなどの，あらまし特有の書式に合わせた変換はあるが，それ以外は同一である．

考えたためである．

5. 実験

提案手法の有効性を検証するために，提案手法で用いるランダムフォレストと，それ以外の分類器，さらに学習データを利用しない重要文抽出手法との比較実験を行った．

5.1 学習データとテストデータ

実験では，昭和 48 年 4 月から平成 31 年 1 月までに新規制定された 1,156 法律のうち，1,111 件から学習データを構築した．この 1,111 件には法律文 104,102 文が含まれているが，3.2 節で述べた手法に基づいて正例 8,793 文，負例 15,507 文を取り出し，合計 24,300 文からなる学習データを作成した．これは TSC-1 [16] における重要文抽出タスクで使用されたデータセット 4,521 文の 5 倍以上のサイズである．ただし，正例とも負例とも判定できなかった文が 77%残っており，こうしたデータを利用して学習データを増やすことは，今後の課題である．

テストデータには，残りの法律から 33 件を選んで *2，人手で正解ラベルを付与した．実際には，ほとんどのラベルは自動で付与できたが，図 1 の緑色の箇所のような場合は，人手で判定した．作成したテストデータの詳細を表 1 に示す．抽出率のマクロ平均は 74%，マイクロ平均は 70%であった．

5.2 比較手法

提案手法と比較するために，先行研究で用いられた決定木 [2] 及び SVM [3] を用いた．さらに学習データを利用しない手法であるリード法 [6]，TextRank [7]，TextTeaser [8] も用いた．

決定木: 決定木は学習データを素性ベクトルに基づき部分集合に分割し，学習データを分類する木構造を構築する．決定木の学習アルゴリズムには CART 法を用いた．

SVM: SVM は，クラスごとに学習データを分離する境界を求める．本実験では [3] を参考に，線形な境界によってデータを分離する線形 SVM (SVM_Lin) と，カーネル関数に多項式カーネルを用いた SVM (SVM_Poly) を用いた．

リード法: リード法 [6] は，文書の先頭から順に文を抽出する手法である．新聞要約などで良い性能を示しており，自動要約におけるベースラインとして使われている．本実験では，法律文書の先頭をそのまま抽出する「リード法 (Original)」と，法律が条単位で構成されることに注目し，各条の先頭を抽出する「リード法 (Article)」を実施した．

TextRank: TextRank [7] は，グラフに基づく重要文抽出手法であり，文をノードとして文書をグラフ構造で表現する．多くの文と類似している文は重要な文であるという仮定があり，グラフのエッジに文間の類似度を付与し，そのグラフから各文に重要度を付与する．本実験では，GitHub 上で公開されている実装 *3 を用いた．

TextTeaser: TextTeaser [8] は，特徴に基づく重要文抽出手法であり，文の長さや文の位置を特徴として，各文に重要度を付与する．本実験では，PyTeaser [9]*4 を用いた．

5.3 実験設定

ランダムフォレスト・決定木・SVM の学習には scikit–learn (v.0.19.1) を用いた．ランダムフォレストのデータ分割基準に

*2 残りの法律 12 件については，法律文とあらまし文の対応付けが困難なものが含まれており，人手で正解ラベルを付与できなかった．

*3 https://github.com/summanlp/textrank/

*4 http://xiaoxu193.github.io/PyTeaser/

表 1: テストデータにおける文数と抽出率

法律番号	文数	抽出文数	抽出率
昭和 59 年法律第 34 号	52	34	65%
平成 元 年法律第 70 号	20	11	55%
平成 12 年法律第 75 号	18	6	33%
平成 16 年法律第 166 号	61	20	33%
平成 23 年法律第 99 号	22	22	100%
平成 24 年法律第 50 号	40	38	95%
平成 25 年法律第 64 号	37	36	97%
平成 25 年法律第 66 号	10	8	80%
平成 25 年法律第 86 号	18	17	94%
平成 25 年法律第 110 号	30	15	50%
平成 25 年法律第 112 号	109	93	85%
平成 26 年法律第 17 号	46	23	50%
平成 26 年法律第 49 号	53	11	21%
平成 26 年法律第 57 号	24	21	88%
平成 26 年法律第 98 号	42	42	100%
平成 26 年法律第 100 号	28	24	86%
平成 26 年法律第 101 号	20	10	50%
平成 26 年法律第 126 号	17	17	100%
平成 26 年法律第 137 号	44	42	95%
平成 27 年法律第 14 号	35	23	66%
平成 27 年法律第 62 号	14	10	71%
平成 27 年法律第 69 号	20	20	100%
平成 28 年法律第 33 号	55	21	38%
平成 28 年法律第 68 号	11	11	100%
平成 28 年法律第 74 号	29	17	59%
平成 28 年法律第 105 号	42	29	69%
平成 28 年法律第 111 号	33	32	97%
平成 28 年法律第 112 号	31	19	61%
平成 28 年法律第 115 号	48	25	52%
平成 30 年法律第 28 号	12	12	100%
平成 30 年法律第 47 号	36	26	72%
平成 30 年法律第 48 号	35	26	74%
平成 30 年法律第 100 号	29	28	97%
合計	1121	789	70%

はジニ係数を用い，決定木の数と深さについては 5 分割交差検証によって最適な値を求めた．決定木と SVM のパラメータについても，5 分割交差検証によって最適値を求めた．今回の学習データは，負例が正例の約 2 倍となる不均衡データである．予備実験の結果にもとづき，アンダーサンプリングを適用して負例のデータ数を正例の数に合わせた．

法律文およびあらまし文は，MeCab (IPA 辞書) を用いて形態素解析し，単語に分割した．素性に用いる単語には，学習データの作成に用いた法律 1,111 件に対応するあらましの形態素から取得した．ただし，解析結果が未知語となるものは素性に用いなかった．

なお，リード法・TextRank・TextTeaser は，抽出文数の設定が必要となる．本実験では，テストコーパスにおける抽出率のマクロ平均である 74% をもとに抽出文数を設定した．

5.4 実験結果

本実験においては，抽出性能を再現率・精度・F 値で，要約性能を ROUGE で評価した．その結果を表 2 に示す．なお，ROUGE については，ROUGE-1・ROUGE-2・ROUGE-L を計算したが，いずれも同様の傾向が見られたため，本稿では

表 2: 重要文抽出実験の評価結果

システム名	再現率	精度	F 値	ROUGE-1
ランダムフォレスト	84.7%	81.9%	81.0%	88.4%
TextRank	79.2%	80.0%	77.4%	88.0%
SVM_Lin	77.7%	82.5%	77.4%	85.1%
SVM_Poly	77.4%	82.1%	77.7%	84.5%
リード法 (Article)	76.7%	78.0%	74.8%	81.9%
決定木	73.4%	80.3%	74.7%	80.1%
TextTeaser	73.0%	74.2%	71.3%	68.4%
リード法 (Original)	70.9%	72.9%	69.7%	77.0%

ROUGE-1 の値のみを示す．これらの値は，テストコーパスである法律 33 件をそれぞれ入力し，得られた評価結果のマクロ平均である．ランダムサンプリングにおける偏りを防ぐため，ランダムフォレスト・決定木・SVM においては，分類器のモデルを 50 個作成し，それらのマクロ平均を求めた．

再現率は，ランダムフォレストがもっとも高く，有意水準 5%ですべての比較手法より優れていた．ランダムフォレストの次に TextRank が高い再現率を示した．精度は，SVM_Lin がもっとも高く，SVM_Poly もランダムフォレストを上回った．F 値は，ランダムフォレストが最高で，有意水準 5%で TextRank 以外のすべての比較手法より優れていた．ROUGE においては，ランダムフォレスト，TextRank の順に高かった．

5.5 考察

提案手法であるランダムフォレストを用いた方法が，精度以外の指標で最高の性能を示した．精度で SVM に劣った原因を調べたところ，SVM はランダムフォレストよりも抽出する文数が少ないことが分かった．逆に言えば，その点が SVM の再現率の低さの原因である．要約においては再現率が重視されることから，SVM より 7%多く文を抽出しているにも関わらず，SVM に近い精度を示すランダムフォレストが，もっとも有用であると言える．ROUGE の結果も，この点を示唆している．

リード法 (Article) がリード法 (Original) の性能を上回ったことは，法律の要約では重要文が先頭以外にも出現することを示唆している．ただし，リード法 (Article) の評価値は TextRank よりさらに低い．これは，法律の要約では文の位置よりも，文中の単語が重視されていることを示唆している．文の位置を用いる TextTeaser も，同様に低い性能であった．TextRank はリード法や TextTeaser と異なり，文の位置を考慮せず，単語の重要度を考慮するため，高い評価値を示したと考えられる．

6. おわりに

本研究では，法律の要約のためにランダムフォレストを用いた重要文抽出手法を提案した．また，その際に従来よりもサイズの大きな要約用のコーパスを作成した．ただし，このコーパスにおいては，元となったデータの 23%しか利用していない．よって，今後の課題としては，学習コーパスの量を増やすことと，それにより抽出性能を向上させることが挙げられる．

謝辞

本研究は JSPS 科研費 17K00460 の助成を受けたものです．

参考文献

[1] 野本 忠司, 松本 裕治: 人間の重要文判定に基づいた自動要約の試み, 情処研報, NL–120–11, pp. 71–76 (1997).

[2] 奥村 学, 原口 良胤, 望月 源: 決定木学習を用いたテキスト自動要約手法に関するいくつかの考察, 情報処理学会第 59 回全国大会講演論文集, 5N–2, pp. 393–394 (1999).

[3] 平尾 努, 磯崎 秀樹, 前田 英作, 松本 裕治: Support Vector Machine を用いた重要文抽出法, 情報処理学会論文誌, Vol. 40, No. 8, pp. 2230–2243 (2003).

[4] Jing, H.: Sentence Reduction for Automatic Text Summarization, In ANLP 2000, pp. 310–315 (2000).

[5] Knight, K., Marcu, D.: Summarization beyond Sentence Extraction: A Probabilistic Approach to Sentence Compression, Artificial Intelligence, Vol. 139, No. 1, pp. 91–107 (2002).

[6] Brandow, R., Mitze, K., Rau, L. F.: Automatic Condensation of Electronic Publications by Sentence Selections, Information Processing and Management, Vol. 31, No. 5, pp. 675–685 (1995).

[7] Mihalcea, R., Tarau, P.: Textrank: Bringing Order into Texts, In EMNLP 2004, pp. 404–411 (2004).

[8] Jagadeesh, J., Pingali, P., Varma, V.: Sentence Extraction based Single Document Summarization, Workshop on Document Summarization (2005).

[9] Gunawan, D., Pasaribu, A., Rahmat, R. F., Budiarto, R.: Automatic Text Summarization for Indonesian Language Using TextTeaser, In IOP Conference Series: Materials Science and Engineering, Vol. 190, No. 1, p. 012048 (2017).

[10] Kupiec, J., Pedersen, J., Chen, F.: A Trainable Document Summarizer, In SIGIR 1995, pp. 68–73 (1995).

[11] Watanabe, H.: A Method for Abstracting Newspaper Articles by Using Surface Clues, In COLING 1996, Vol. 2, pp. 974–979 (1996).

[12] Barzilay, B., McKeown, K. R.: Sentence Fusion for Multidocument News Summarization, Computational Linguistics, Vol. 31, No. 3, pp. 297–328 (2005).

[13] Rush, A. M., Chopra, S., Weston, J.: A Neural Attention Model for Abstractive Sentence Summarization, In EMNLP 2015, pp. 379–389 (2015).

[14] Kawachi, G., Ogawa, Y., Nakamura, M., Ohno, T., Toyama, K.: Daily News on Japanese Legislation toward Global Sharing of Japanese Legal Information, Journal of Open Access to Law, Vol. 3, No. 1, 19 pages (2015).

[15] Breiman, L.: Random Forests, Machine Learning, Vol. 24, No. 1, pp. 5–32 (2001).

[16] Fukushima, T., Okumura, M.: Text Summarization Challenge Text Summarization Evaluation in Japan, In NAACL 2001, Workshop on Automatic Summarization, pp. 51–59 (2001).

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

AI による立法支援システム

Legislation Supporting Systems by AI

角田 篤泰[*1]
Kakuta, Tokuyasu

[*1] 中央大学 研究開発機構
Research and Development Initiative, Chuo University

This paper reports our legislation supporting system adopting AI technology researched and developed by the author. The system is extended from "eLen" system as a result of our previous research. "eLen" system has been provided as a web database system including about 90% regulations of all of local governments in Japan. Furthermore, the system can generate comparing tables among similar regulations and find clusters of similar regulations. In this research, we have implemented a new function for automatically generating templates of regulation based on the clusters. This report shows the outline of the function and the short history of the systems.

1. はじめに

本稿では，筆者らがいくつかのプロジェクトを通じて，AI 技術を用いて実施してきた立法支援システムの研究・開発の経緯や現状について報告する．

2009 年の人工知能学会全国大会において，筆者は立法過程とソフトウェア開発過程が類似していることを示し[角田 2009]，それ以降，様々な立法支援システムの研究・開発と実際の立法の現場へのシステム提供や技術提供を行ってきた[角田 2012,角田 2014,角田 2018b]．

立法過程においても，ソフトウェア開発の手法と同様に，ウォータフォール型で立法する場合や，過去の資産を再利用して立法する場合，そしてそれらを合わせた場合などがある．ウォータフォール型については，Python 表記を用いて条例を記述する研究などが進行中である[角田 2018a]．一方，過去の資産を再利用する方式については，類似条例を自動検出し，比較表などを自動作成するシステムが『立法支援機能付き eLen 条例データベースシステム』(以下「eLen システム」と略す)として，2012 年より全国の自治体に実際に提供され，今日まで多くの自治体で利用されている．このシステムをベースに新たな機能の開発も進んでいる．この他，政府が一般公開している日本法令の Web データベースシステムの『e-LAWS』において，政府の各省庁内では，2016 年より，そのデータベース更新時に自動的に改正法案を合成するシステムが稼働しているが，その改正法案を作成するための膨大な法制執務ルールという，条文表記のルールを計算機が扱えるように筆者が手作業で整理し，政府に提供した[角田 2016]．

本稿では，紙面や時間の制約のため，これらのうち，2018 年度の人工知能学会全国大会のオーガナイズドセッション「CREST 人工知能」にて報告した「AI 技術を用いた法的文書作成支援」のプロジェクトの最新の進捗について報告することにする．これは，上記の eLen システムにおける類似条例のクラスタリング機能を改良し，さらに，そのクラスタリング結果を用いて条例案のテンプレートを自動合成するというものである．

連絡先：角田篤泰，中央大学研究開発機構，〒112-8551 東京都文京区春日 1-13-27，kaku@nagoya-u.jp

2. eLen システムの概要

元々日本では全国に 1790 団体ある自治体の条例データは集積されておらず，自治体ごとにそれぞれ Web サーバを立ち上げ，その中で各自治体の例規(条例＋規則)情報を自治体ごとの方式で提示していた．多くの自治体ではクローリング技術などで自動的に集められるような仕組みにはなっておらず，中には PDF 形式であったり，地元企業の独自性の強いデータベースシステムで提供されたりしている自治体もあった．そこで，個別に対応しながら，それらの例規を収集して，1 つのデータベースとして構築し直し，Web を通して各自治体に利用できるように 2012 年より eLen システムとして実現して[角田 2014]，全国の自治体向けに提供を始め[1]，今日に至っている[2]．なお，eLen の画面例は図1の通り．

このような立法支援システムの研究・開発のために全国の自治体を回り，要望を集めていた際に，どこの自治体でも，高度な支援システムよりも，まずは自治体横断検索ができる例規データベースが欲しいという要望が多かったため，そのようなデータベース提供を研究・開発の手始めとした．現在では，全国の約 9 割の 1670 団体の例規，約 130 万本(約 1490 万条)を集め，一旦そのすべてのデータを政府標準の XML 形式に整形し[3]，法制執務[4]の例規表示形式で提供している．

[1] 日本経済新聞全国版(2012 年 08 月 16 日)「条例作成支援の検索システム」，読売新聞全国版(2012 年 10 月 30 日)「全国の条例データベース化して公開…名古屋大」を始め，多くのマスコミ報道あり．

[2] 2012 年より名古屋大学から公開を開始し，2017 年より鹿児島大学に運営を移し『全国条例データベース powered by eLen』として公開している．簡易版は一般公開もしている．URL は <https://elen.ls.kagoshima-u.ac.jp/> である．

[3] 政府の e-LAWS で採用されている XML スキーマとコンパチブルな形式である．詳しくは <http://elaws.e-gov.go.jp/search/html/readme.html#schema> を参照のこと．なお，e-LAWS では総務省の「e-Gov」の一環として，現行のすべての日本法令を公開している．この URL は <http://elaws.e-gov.go.jp/search/elawsSearch/elaws_search/> である．

このように自治体職員からの要望を集める過程で，例規作成の立法作業においては，ウォータフォール型の立法方式を採るよりも，他の類似例規を見つけてきて，それをベースに新たな例規を作成する方式の方が圧倒的に多いことがわかり，そちらを支援する，類似例規を見つける機能，そして，実際にそれらを比較表(自治体法務の分野では「ベンチマーク表」と言う)にして資料を作成できる機能が求められた．そこで，これらを実装して，eLen データベースの付属機能として公開している．それが「立法支援機能付き」と呼ばれる所以である．類似例規を見つける機能は，類似な例規ごとのクラスタを算出することによって実現できるため，例規クラスタリングが主要な課題となった．通常のクラスタリングのように，単に意味が似ているだけでなく，文面を真似したり，標準的な書き方を参照したりすることも必要であったので，いきなり例規単位でクラスタリングをする方式を採用せず，まずは，条文単位のペア組合せ間の「編集距離」[5] によって，全組合せについて計算しておき，現在，3 割以内の書き替えに収まっているものを類似条文のペアとして判定している．さらに，その結果を用いて，類似条文となるペアを 7 割以上含んでいる例規ペアを「類似例規ペア」と判断している．より正確には，例規集合 R について，$r \in R$ に含まれる条文の集合を $J(r)$と記し，$r_1, r_2 \in R$ について $p \in J(r_1) \times J(r_2)$が類似条文のペアとなるものの集合を $Sim(r_1, r_2)$と記すと，r_1 と r_2 が類似例規ペアとなるのは次の条件を満たす時である．

$$|Sim(r1,r2)| / (|J(r1)|+|J(r2)|) \geq 0.7$$

右辺の 0.7 は本研究で現在採用している閾値である．

本研究では，この類似例規ペアをエッジとし，例規を 1 つのノードとするグラフ構造において，クリークとなるもののうち極大なものをクラスタとして算出し[6]，同一クラスタに属するものを類似例規と判定した．さらに，すべての例規を法令用の XML 形式で保持しているので，各条文に付される見出し部分の情報も認識して保持している．そこで，その見出し情報を使って，類似例規や比較対象例規内の条文の種別を区別して比較表を自動作成している．以上が先行研究成果である eLen システムの概要である．

3. テンプレート合成の方式

先行研究の比較表自動作成の過程において，さらなる要望として例規のテンプレートの自動合成が求められた．これは，比較表を見たときに，共通している部分と差異の部分を分けて眺めて考察や作業を進めているという事実に基づいた要望である．図 2 のように，類似した文があれば，共通部分を残し，差異の部分をオプションとして一覧できるようにしておけば(できれば頻度の多いものが順に)，比較表を用いた考察と同じことができるテンプレートができあがるのである．

このようにして，例規のテキストを処理する場合，要するに，マルチな diff コマンド機能と，大前提として類似例規でクラスタリングしておくことが核となる．前者については mdiff という diff 機能拡張したコマンドをすでに実装済みであり，後者は，とりあえず従来の方式によってクラスタを計算することで対応している．本稿執筆時現在，最終的なテンプレート GUI 用のデータに変換するプログラムを作成中である．デモシステムの画面例を図 3 に示す．

また，実験に用いた現実の例規に対して mdiff を実施の例を示す．実際に本システムでクラスタとして計上された次の 4 つの条文本文<a>～<d>に対して，mdiff を適用した出力が最後の<T>である．

<a> (準用規定) この規程に定めていない事項については，全国町村会総合賠償補償保険制度 において適用される「賠償責任保険普通保険約款」，「予防接種実施主体特約条項」及び「全国町村会予防接種事故賠償補償保険契約特約書」の規定を準用する．

<b> (準用規定) この規程に定めていない事項については，全国町村会総合賠償補償保険制度 において適用される「賠償責任保険普通保険約款」，「予防接種実施主体特約条項」および「全国町村会予防接種事故賠償補償保険契約特約書」の規定を準用する．

<c> (準用規定) この規則に定めのない事項については，全国市長会予防接種事故賠償補償保 険制度において適用される「賠償責任保険普通保険約款」，「予防接種実施主体特約条項」及び「全国市長会予防接種事故賠償補償保険契約特約書」の規定を準用する．

<d> (準用規定) この告示に定めていない事項については，全国町村会総合賠償補償保険制度 において適用される「賠償責任保険普通保険約款」，「予防接種実施主体特約条項」及び「全国町村会予防接種事故賠償補償保険契約特約書」の規定を準用する．

<T> (準用規定) この[規則,規程,告示]に定め[てい,の]ない事項については，全国[市長,町村]会[予防接種事故,総合]賠償補償保険制度において適用される「賠償責任保険普通保険約款」，「予防接種実施主体特約条項」[及,および]「全国[市長,町村]会予防接種事故賠償補償保険契約特約書」の規定を準用する．

---

4 「法制執務」とは，法令や例規を実際に記述するための仕事を指すが，その記述スタイルには詳細で膨大な慣習的ルールがある．多くの自治体で参照されるマニュアル[石毛 2012]もある．

5 この編集距離としてはレーベンシュタイン距離を採用しており，挿入，削除および置き換えをそれぞれ 1 カウントとするものである．なお，対象は文字ではなく，形態素としている．

6 「クリーク」とはグラフ G(V,E)において，完全グラフを形成する V の部分集合のことである．

最後の出力テキストの「[」「]」で囲まれた部分が選択肢の候補群となる．それぞれ半角コンマで区切られている．

4. クラスタリング方式の比較

本節では，サンプルとして用いた 23 自治体分の約 13000 例規について，本方式で採用しているクラスタリングと Paragraph Vector に基づく方式によるクラスタリングを行い，複数の法律家によって行われた，それらの意味的な比較・検証の結果を報告する．

まず，前者の方式は，先行研究の頃から実施していた条文ペアの編集距離情報をベースに類似例規ペアを算出し，そこからクリークを検出するという「編集距離クリーク方式」で出力される例規クラスタを用いる方式である．

次に，後者の方式は，「Paragraph Vector モデル」に基づく Doc2Vec ライブラリ[7]を使って計算される例規クラスタを用いたものである[8]．詳細は Quoc Le らの提案[Le2014]を参照のこと．

それぞれの方式によってクラスタリングされた結果の例規について，複数の法律家により意味的な比較・検証をしてもらった．その結果，前者は速度的に問題があるが，意味的にはほとんど問題なく，後者では速度的には問題がないが，意味的には問題があった．後者では，得られたクラスタ内に意味的に的外れなものが含まれるケースも多く，また，大事なものが含まれていないケースも多くあった．さらに，テンプレート合成のためには条文単位で共通の文字列の並びが必須だが，後者ではそのテンプレート用の共通部分がほとんどない例規が含まれてしまうケースがあり，実用には耐えないものであった．これは例規クラスタ内に 1 例規でも全く字面が異なる例規が入ってしまうと，例規文面のほとんどが差分になってしまい，共通部分がなくなってしまうからである．そこで，後者による算出結果は，まだまだパラメタ設定やクラスタリング方式の改良の余地はあるが，現時点では，例規や契約書のようなもののテンプレート自動合成の見地からは，採用に踏み切れない．ただし，後者の速度性には魅力があるため，現在，前者と後者の方式の組合せを検討中である．

なお，このようなクラスタの比較をするために有効だったのが「疑似クリーク」の概念であった[大久保 2005]．疑似クリークとは，比較的大きな共有部分を持つ極大クリークを合成したもの(各極大クリークをそのグラフ内のノードの集合とみたときの和集合)のことである．これを見つけておき，共有部分をコアとし，そこからはみ出す各極大クリークの部分集合をオプションと見ることで，例規の構造レベルでの意味的な分析結果に近い様子が浮き彫りになる．現在，法律家による検査用にこの疑似クリークベースの GUI を設計中である．彼らは，既に「疑似クリーク」という言葉を使って検査作業を行っている．

5. まとめと課題

本稿では，先行研究成果である eLen システムの類似例規クラスタリング機能を発展させて，例規テンプレートの自動合成を目指していることを示した．こうして，出来上がったテンプレートについて，機械学習や深層学習などの一般的な問題として，なぜそうなるのか説明できない，という問題を孕んでおり，これが大きな課題となる．しかしながら，現在，既に並行して進めている「SDL(Software Defined Law)アプローチ」によって[角田 2018a]，本稿の 1.で示したように，ウォータフォール型に対応させ，Python によって法令プログラムを作成する研究で得られる法制度のクラスの意味的な構造情報を参照することも進めている．こうして，自動合成された例規テンプレートへの説明が付与されたり，より高度な立法支援やアドバイスがされたりできるようにすることを目指している．

謝辞

先行研究 eLen は文科省科研費基盤研究 A(課題番号 24240040)の助成を受けたものであり，例規テンプレート自動作成の研究は JST の CREST 事業(グラント番号 JPMJCR1782)の助成によるものである．また疑似クリークによる分析研究は公益財団法人栢森情報科学振興財団(K29 研ＸＸⅡ第 509 号)の支援を受けて実施された．

参考文献

[石毛 2012] 石毛正純: 法制執務詳解 新版 2，ぎょうせい，2012.

[角田 2009] 角田篤泰, 齋藤大地, 外山勝彦: ソフトウェア開発過程との類似性に基づく立法支援システム,JSAI 全国大会(第 23 回)論文集,2F2:4.1-4.4, 2009.

[角田 2012] 角田篤泰: e-Legislation 環境の構築へ向けて, 情報ネットワーク・ローレビュー，第 11 巻, pp.13-32, 2012.

[角田 2014] 角田篤泰, 島亜紀 他: 全国自治体例規データベース eLen の構築と定量的例規調査, 情報ネットワーク・ローレビュー, 第 13 巻, 第 1 号, pp.14-33, 2014.

[角田 2016] 角田篤泰: 改め文作成の機械化に向けた基盤 ― e-Legislation の一具体例 ―, 名古屋大學法政論集, 266 号, pp.23-59, 2016.

[角田 2018a] 角田篤泰: SDL(Software Defined Law)としての法令プログラミングと立法支援システム，第 80 回全国大会講演論文集, pp.391-392, 2018.

[角田 2018b] 角田篤泰: 人工知能の発展と企業法務の未来(4-2), NBL, 1125 号，商事法務, pp.43-52, 2018.

[Le2014] LE, Quoc and MIKOLOV, Tomas.: Distributed representations of sentences and documents, International Conference on Machine Learning, pp. 1188-1196, 2014.

[大久保 2005] 大久保好章, 原口誠: 疑似クリーク探索によるクラスター抽出, JSAI 全国大会(第 19 回)論文集, 2F3-05, 2005.

---

7 このプログラムは Python を使ったトピックモデリングのためのライブラリ「gensim」に含まれているもので，詳細と配布については，次の URL のサイトを参照のこと．
<https://radimrehurek.com/gensim/models/doc2vec.html>

8 実験で使用した分散表現ベクトルの学習パラメタは，ベクトル長=300，エポック数＝1000，ウインドウサイズ＝30，負例サンプル数＝5 であった．各例規の分散表現ベクトルに対して，凝集型階層的クラスタリングを実施した．その際，距離計算手法としては ward 法を，メトリックとしてはユークリッド距離を採用した．

図 1: eLen システムの画面例

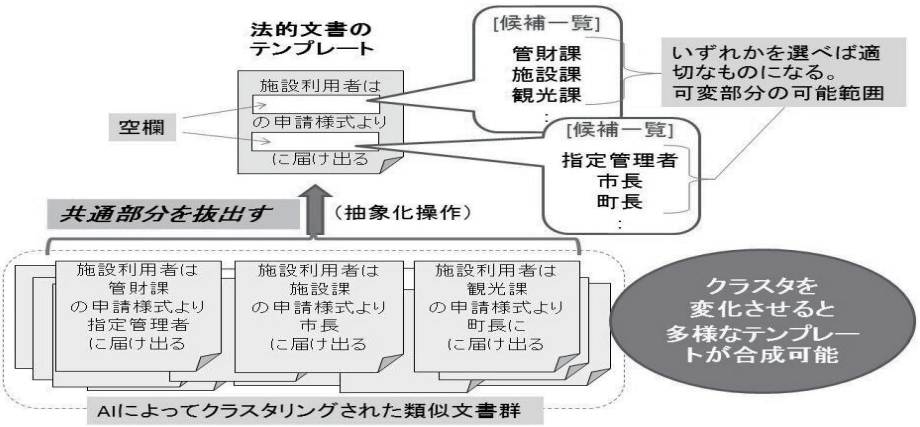


図 2: テンプレート自動合成のイメージ

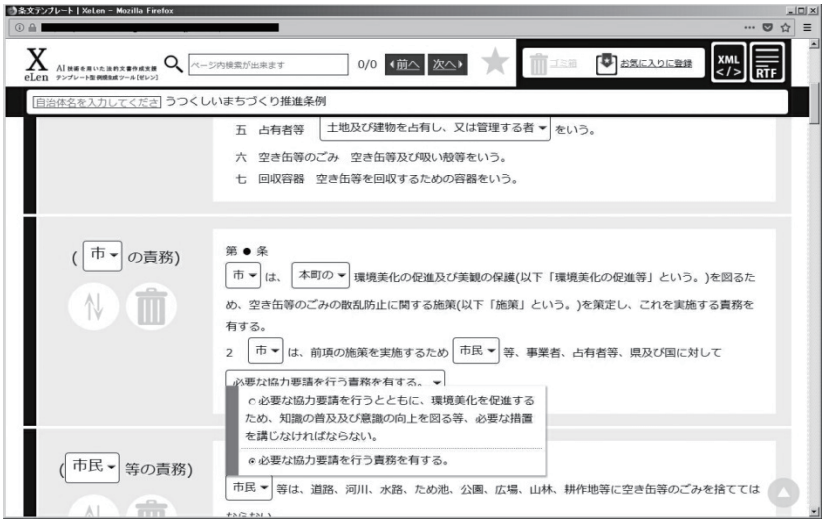


図 3: 例規テンプレートを用いた立法支援システムのデモ画面例

4E2-OS-7a-04

AI裁判支援システムへの人々の期待と受容

People's Expectation and Attitude toward AI-assisted Courts

太田　勝造[*1]
OTA Shozo

[*1] 東京大学・明治大学
The Univ. of Tokyo/Meiji Univ.

Abstract: This paper presents the preliminary results of my nation-wide internet survey on people's expectation and attitude toward AI-assisted Courts which uses experimental research design. Section 1 & 2 explains the components and their AI-models of legal decision-making and the merits and demerits thereof respectively. Section 3 reports the preliminary results which was conducted in February 2019.

1.　AIによる裁判支援システムとは？

1．1.「裁判」とは？

裁判の構成要素（裁判所の作業）の内容は以下の3つに大別できる.

1. 事実認定（心証形成）
2. 法規範の法律要件への「当てはめ」（包摂）
3. 法的推論

1の事実認定（心証形成）とは,「法律要件」に該当する事実（「主要事実」）の真偽を,「証拠方法」（人証, 物証など）によって判断すること作業である. 主要事実は当事者（民事裁判の原告と被告, 刑事裁判の検察官と被告人）が主張し, 証拠方法を提出する. 主要事実の真偽の程度の判断を「心証」と呼ぶ. 心証が「真実（存在）」の方向や「偽（不存在）」の方向へ動いていくことを「心証形成」と呼ぶ.

2の法規範の法律要件への「当てはめ」（包摂）とは, 事実認定の作業によって真であると認定された事実が, 法律要件に当てはまるか否かを判断する作業である. これは法的判断に他ならない. 法律要件が事実を包摂するか否かと表現して, 包摂判断と呼ぶこともある.

なお, 1における事実認定の対象の事実が主要事実であるかの判断と, 2の当てはめの判断は相関する判断であり, 同時並行的になされることも多い.

3の法的推論とは, 請求認容か請求棄却かを判断するため, 適用される複数の法規範について「法的推論」をして結論を導く作業である.

裁判では, 以上の作業のそれぞれの結論を整理して「判決理由」として起草し, 全体の結論を「判決主文」として示す. 民事裁判では, 請求認容, 請求一部認容, 請求棄却, 請求却下などであり, 刑事裁判では, 被告人無罪, 有罪, 有罪の場合の刑の量定などである.

1．2. 裁判支援のAIモデル

1．2．1．事実認定のAIモデル

事実認定, 心証形成の作業を支援するAIモデルとしては, 不確実性下の意思決定の標準モデルである, ベイズ確率論・意思決定論の応用が考えられる. 証拠方法の吟味に基づく主要事実の真偽の判断は「主観的確率」と位置づけることができ, これには, ベイズ法則による「規範モデル」を用いることができる. 規範モデルにおいては, 合理的事実認定判断のモデルであり, 規範モデルの結論と異なる結果の裁判官認定は非合理とされ, 合致する裁判官認定は合理的とされる.

なお, 現実の裁判官の認定を記述するためのモデルである「記述モデル」ならば, 裁判官認定と異なる結果となれば記述モデルの方が棄却される.

複数の証拠方法の間に関連性が存在することは非常に多いので, より複雑なベイジアン・ネットワークによってモデル化することになる.

1．2．2．当てはめ判断のAIモデル

ある法律要件に, 裁判で争われる具体的事実関係が当てはまるか（包摂されるか）の判断は, 法的ルールが適用されるか否かの中核的判断である.

同一ないし類似の先例や明確な法令が存在して, 法律家なら同じ結論に至るような「イージー・ケース」であれば, 当てはめ判断は容易でほぼ機械的にできる. 他方, 同一ないし類似の先例も明確な法令も存在しない場合は, 正義, 公平, 効率などの考慮に基づく法政策的判断が必要で, 立場や法律家によって判断が分かれる. このような事例は「ハード・ケース」と呼ばれ, 当てはめ判断は非常に困難となる.

イージー・ケースであれば, 論理プログラミングやルール・ベースと・リーズニングなどによってAI化できる. ハード・ケースの場合は, 国民意識や国民の理解などを自然言語処理や深層学習で応用してAI化するか, ネット直結のスポット世論調査のような「民意」組み上げシステムを内蔵したAI化が必要となるかもしれない.

ハード・ケースを解決する上で, 法的ルールのある判断選択肢が他の判断選択肢群に対して望ましいか否かは, 当該判断選択肢が人々の行動や社会状態に与える影響を予測ないし調査して判断する必要があり, これは現実に対応した「人工社会」等を構築してのシミュレイションによるなど, 理論的には考えられるが, 実践状はさしあたり不可能である.

1．2．3．法的推論のAIモデル

裁判では, 刑事であれ民事であれ複数の規範を用いる必要が生じる. 規範相互には矛盾対立が存在しうる. これを解消するためのメタ・レヴェルの規範も存在する.

連絡先：太田勝造（おおた・しょうぞう）, 東京大学大学院法学政治学研究科・明治大学法学部, sota@j.u-tokyo.ac.jp

そして，主要事実のいくつかが証明され，それらに対応する法律要件が認められて，複数の規範が適用される場合，「法的推論」によって正しい結論を導く．

しかも，裁判ではその後にさらに主要事実が証明されて，それに対応する新たな法律要件が認められて，それまでの法的推論の結論が覆ることも起きる．

この法的推論の過程のAIモデルとしては，「非単調推論」を組み込んだ論理的推論が考えられる．

2. 裁判支援のAIモデルの評価

2．1．事実認定

ベイジアン・ネットワークによるモデル化によれば，確率モデルであるゆえに，三段論法のような真偽の二値論理に対して，多値を扱える．確率モデルであるゆえ，さらに，不確実性を合理的に扱うことができる．数理モデルであるので，計算プロセスが透明である．つまり，ベイズ法則を知っていれば説明であり，ミスなどについて，デバッグすることができる．

実質的なベネフィットとしては，事実認定における裁判官や裁判員のバイアスや過誤から解放される可能性がある点が指摘できる．また，AIによる事実認定は，様々なバイアスのありうる人間の裁判の事実認定と異なり正確でありうる．よって，裁判所や裁判官による判断のブレがないと期待できる．

但し，法規範の対象は人間活動と人間社会のほぼすべての領域であり，森羅万象について条件付確率テーブル(Conditional Probability Table)を埋め尽くすことはかなり困難である．大量のデータが必要となるが，現実には存在していない．

実質的なデメリットとしては，裁判の事実も証拠も雑多すぎるので，実現が困難である．また，AIによる事実認定への人々の支持納得，受容がどれほど期待できるかは開かれた問題である．

2．2．法的当てはめ

法的当てはめ判断をAI化できた場合のベネフィットとして，法的判断における裁判官や裁判員のバイアスや過誤から開放されうる点がある．裁判官や裁判員による判断のブレもなくなり，判断の恣意性と主観性を排除できると期待される．こうして，裁判の地理的，時間的な斉一性の実現が期待できる．どの地方の裁判所でも，いつの時代の裁判所でもブレのない判断を示してくれるようになる．

但し，深層学習・自然言語処理による当てはめ判断には，ビッグ・データが必要である．深層学習は，プロセルが不透明で，結論に対する理由や根拠を「見える化」できない．よって，どこが間違っているかわからない．これは法的責任の説明ができず，法的判断の内容（過失，故意，因果関係などの判断）を説明できないことを意味する．

そもそも，規範的価値判断が当てはめであるから「正しい立法」や「正しい先例」が唯一存在する保証はまずない．これは，正解，最適解，すなわち「教師信号」を確定できないことになる．敢えて，教師信号を内部で構成しようとすれば，法的価値判断，法政策的判断をシステムに組み込むことになるが，これは困難である．

現実的には，多数の学説，多数の判例が相互に争っている場合，その処理が困難である．

また，AIによる法的当てはめ判断への人々の支持納得，受容がどれほど期待できるかは開かれた問題である．

2．3．法的推論

法的推論を論理プログラミングで実現する上では，真偽の二値論理の範囲内で可能となる．この場合，プロセスが透明であり，素人相手にも説明できるとともに，デバッグもしやすい．

実質的なベネフィットとしては，論理的推論である限りで，プログラミングや入力にミスがないなら現実妥当性を有する結論を導くことができる．

そして，裁判官や裁判員のバイアス，恣意性から開放されうる．

但し，論理的推論の場合，情報は増えないので，"Garbage-In, Garbage-Out"となる可能性がある．これは，システムのプログラミングでルールを間違えたり，利用の際の入力データを間違えたりすれば，誤回答をすることになる．また，法令や先例は，深層学習やベイジアン・ネットワークを機能させるには少なすぎるが，法令や先例を現状のように人手で論理ルール化するには大量すぎて，労力，費用，時間がかかりすぎる．

また，AIによる法的推論への人々の支持納得，受容がどれほど期待できるかは開かれた問題である．

3. 裁判支援AIシステムへの人々の期待と受容

AIによる裁判支援システムへの人々の支持納得，受容がどれほど期待できるかは開かれた問題であるので，社会調査を実験計画法で設計して，人々の期待と懸念と受容の程度を調査する．

例えば，

質問「交通事故を起こして相手を死亡させてしまい．あなたは，業務上過失致死罪で起訴され，刑事裁判の被告人となったとします．下記の裁判所の中で，あなたが裁判を受けたい裁判所はどれですか．」

と尋ね，

(1)生身の裁判官，
(2)素朴なレベルのAI裁判システム，
(3)かなり高度なレベルのAI裁判システム，
(4)チューリング・テストをクリヤできる（生身の人間と区別がつかない）AI裁判システム，

について，「裁判を受けたい程度」を尺度法で尋ねて比較するなどの調査である（2019 年 2 月時点で調査実施中）．

このようなリサーチ・デザインによるパイロット調査の結果を報告する予定である．

参考文献

[太田 1982] 太田勝造『裁判における証明論の基礎』，弘文堂，1982 年．

[太田 1990] 太田勝造『民事紛争解決手続論』，信山社，1990 年．

[太田 2000] 太田勝造『社会科学の理論とモデル7：法律』，東京大学出版会，2000 年．

[吉野 1986] 吉野一『法律エキスパート システムの基礎』，ぎょうせい，1986 年．

[吉野 1986] 吉野一『法律エキスパート システムの基礎』，ぎょうせい，1986 年．

[吉野 2000] 吉野一『法工知能：法的知識の解明と法的推論の実現』，創成社，2000 年．

[吉野] 吉野一，加賀山茂，櫻井成一朗，新田克 己，鈴木宏昭，太田勝造「事例問題に基づく法律知識ベースおよび論争システムを活用した法創造教育」『IT活用教育方法研究』第 9 巻第 1 号 1 頁～5 頁，2006 年．

Organized Session | Organized Session | [OS] OS-7

[4E3-OS-7b] AIの法学の応用(2)

佐藤 健（国立情報学研究所）、新田 克己（国立情報学研究所）

Fri. Jun 7, 2019 2:00 PM - 3:40 PM Room E (301A Medium meeting room)

[4E3-OS-7b-01] Question Analysis and Rule Design for Generating PROLEG Rules towards Automatic Legal Bar Exam Solver

〇Ryuji Hayashi[1], Yoshinobu Kano[1] （1. Shizuoka University）

2:00 PM - 2:20 PM

[4E3-OS-7b-02] Problem classification of legal bar exam and structured dictionary design for human relationships in legal documents

〇Naoki Kiyota[1], Yoshinobu Kano[1] （1. Shizuoka University）

2:20 PM - 2:40 PM

[4E3-OS-7b-03] Reasoning on a bipolar argumentation framework for a law

〇Tatsuki Kawasaki[1], Sosuke Moriguchi[1], Kazuko Takahashi[1] （1. Kwansei Gakuin University）

2:40 PM - 3:00 PM

[4E3-OS-7b-04] Towards expanding the scope of trials via machine learning

〇Tatsuro Kawamoto Kawamoto[1], Masami Yasumoto[1], Yoichi Motomura[1] （1. National Institute of Advanced Industrial Science and Technology）

3:00 PM - 3:20 PM

[4E3-OS-7b-05] Implementation of Criminal Procedure using PROLEG Technology

〇Ken Satoh[1], Yoshiaki Nishigai[2] （1. National Institute of Informatics, 2. Nihon University）

3:20 PM - 3:40 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

司法試験自動解答における論理型言語 PROLEG ルール生成のための過去問分析と設計

Question Analysis and Rule Design for Generating PROLEG Rules towards Automatic Legal Bar Exam Solver

林 隆児
静岡大学情報学部
rhayashi@kanolab.net

狩野 芳伸
静岡大学情報学部
kano@inf.shizuoka.ac.jp

There have been COLIEE competitions held, which aim to extract legal information and to resolve legal textual entailments by answering Japanese legal bar exams. We analyzed 72 past legal bar exam problems to examine possibility of automatic solver. In order to construct such an automatic solver, we use logical programming language called PROLEG, which performs legal reasoning by defining logical rules. While PROLEG rules are normally defined manually, our ultimate goal is automatic generation of PROLEG rules as the first research. Our approach combines natural language processing and logical inference by PROLEG. Assuming that the natural language processing part could be available in future, we implemented programs for other parts that do not perform natural language processing. We performed subjective evaluation for easiness of associating PROLEG rules with problem sentences. Furthermore, we suggest new rule design that makes rules fitting better with natural language processing. This design should allow solving more problems automatically.

1. はじめに

試験の自動解答は自然言語処理や人工知能の主要なベンチマーク課題のひとつといえる。その中で、我が国の司法試験自動解答についても様々な取り組みがされている。COLIEE（Competition for Legal Information Extraction and Entailment）ワークショップは、法律文書の情報検索や含意関係認識問題に関するコンテストとして毎年開催されている[1][2][3]。2018 年に開催された COLIEE 2018[4]では、事件について判例を抽出するタスク（Task1）、判例から判決を決定するタスク（Task2）、問題文から関連する条文を検索するタスク（Task3）、法的質問に対して、関連する条文が分からない状態で、2 値（Yes/No)で解答するタスク（Task4）の四つのサブタスクが設定された。Task3 および 4 は我が国の司法試験から民法短答式の問題を用いたものである。本稿はこのうち、主に Task4 の内容を対象とする。

Task4 では、訓練データとして、質問、関連条文、質問の答え（Yes/No)が与えられ、テストデータとしては質問のみが与えられる。多くの参加チームは、機械学習や単語の重要性の解析を用いて自動解答を試みたが、我々は Prolog[5]をベースにした論理型言語 PROLEG[6]を用いた、自然言語処理と論理推論を組み合わせるアプローチをとる。

PROLEG は、判例や条文をルールとして論理的に記述することができ、それに質問の事実を加えて記述することで、法的問題を推論することができる。PROLEG を用いることで、自然言語処理と法的知識に基づいた、人間の解答プロセスに近い方法で自動解答を実現できるのではないかと考える。法的知識を論理として正しく組み込むことができれば、自然言語処理の性能如何で自動解答の正答率が格段に上がると期待される。

佐藤ら[7]は、契約文書に対して独自に作成したフレームを用いて自然言語処理を行い、PROLEG ルールへの変換を行っている。佐藤らの研究は契約に関する文章に限ったものであるが、司法試験の問題は契約に関する問題だけでなく、契約に関しない問題がさまざまな状況、さまざまな文体で問われる。より多くの問題に適応するためには、膨大なフレームの設計が要求されるだろう。本研究では、法的知識に基づいた自然言語処理により問題文に対応するのではなく、法的知識はすべて PROLEG ルールとして組み込まれているものと仮定する。自然言語処理においては、問題文の単語と PROLEG ルールの単語をどう結びつけるか、という焦点で取り組む。このアプローチを進めるために、本稿では司法試験の問題文を PROLEG ルールに変換する手法を提案し、PROLEG ルール自動生成プログラムの作成を進め、さらにこの手法に基づいた自動解答の可能性について分析を行う。

2. PROLEG

PROLEG とは。論理型言語 Prolog を法律用に拡張したものである。ルールを記述することにより、法的推論を行う。ルールは、Y(X1,X2,…Xn).と表記される。Y が権利や行為を指し、**述語**とよばれる。X は動作主や目的語などが記述され、**引数**とよばれる。<=記号を使った Y1(X1) <= Y2(X2)という表記は、Y2 は Y1 が成立するために必要なルールを示す。この Y2 を Y1 の**要件事実**とよぶ。Y1(X1) <= Y2(X2), Y3(X3).などと表記すると要件は AND 条件になる。proleg 条件(Y(X)).という表記は、Y=X の関係を示す。引数には、**定数**と**変数**の 2 種類がある。定数は確定的な単語であり、変数は推論過程で具体化される。変数は「アンダースコアで始まる」「大文字の英字で始まる」、定数は「小文字の英字で始まる」「漢字で始まる」「シングルクォートで囲まれている」文字列と定義されている。例えば、

所有権に基づく返還請求権(_所有権者,_相手方,_対象物)<=
所有権(_所有権者,_対象物),
占有(_相手方,_対象物).

というルールは、「”所有権者による、ある対象物についての、相手方に対する、所有権に基づく返還請求権”が成り立つためには、”所有権者による、対象物についての所有権と”、”相手方による、対象物の占有”が成り立つ必要がある」と読める。

以下では具体的な問題の解き方と共に PROLEG ルールの作り方を説明する。

[問題文]
土地の所有者は、同土地を不法に占有する者に対して、所有権に基づき土地の明け渡しを求めることができる。

この問題では、物権的返還請求権が認められるかどうかが焦点となる。この物権的返還請求権を軸に、要件事実を展開する。これを、PROLEG では Rule Base とよぶ。Rule Base とは、民法と裁判例を PROLEG ルール化したものである。Rule Base では、A() <= B().と要件事実を展開した

後、B()についてもさらに要件事実があれば展開し、要件事実がなくなるまで続ける。上の問題の物権的返還請求権については、次のような結果が得られる。

物権的返還請求権(_所有権者,_相手方,_対象物)<=
**　　所有権に基づく返還請求権(_所有権者,_相手方,_対象物).**
所有権に基づく返還請求権(_所有権者,_相手方,_対象物)<=
**　　proleg 条件(現時点(_占有時)),**
**　　所有権(_所有権者,_対象物,_占有時),**
**　　占有(_相手方,_対象物,_占有時).**
所有権(_所有者,_物,_時点)<=
**　　所有権(_所有者,_物,_もと所有時,_時点).**
所有権(_所有者,_物,_もと所有時,_問題となる時点)<=
**　　もと所有(_所有者,_物,_もと所有時,_問題となる時点).**

この Rule Base において、末端の要件事実、つまり<=によって展開されている要件事実がないルールについて記述するのが、Fact Base である。Fact Base とは、問題における事件の事実を PROLEG ルール化したものである。上の例では、末端にあたるのが

proleg 条件(現時点(_占有時)).
占有(_相手方,_対象物,_占有時).
もと所有(_所有者,_物,_もと所有時,_問題となる時点).

である。Fact Base においては、これは

現時点(_占有時).
主証(占有('B',土地,_占有時)).
主証(もと所有('A',土地,_もと所有時,_問題となる時点)).

と記述される。

proleg 条件は、要件ではなく条件として表現したい場合に用いる記法である。proleg 条件は、Fact Base では proleg 条件()述語を外して表現される。通常の要件事実については主証()述語で囲う。さらに、引数は「_相手方」→「'B'」、「_対象物」→「土地」など、事実に基づく単語に置き換えられる。「_占有時」などは問題文に具体的な数値が記述されていないため、変数の形のままで記述する。

PROLEG では Rule Base のトップのルールを質問として記述する。今回の例でいえば、**請求権存在(物権的返還請求権('A','B',甲土地)).**と記述する。引数は Fact Base と同様に事実に基づく単語に置き換える。請求権存在 ()述語は、PROLEG の推論を true/false で出力するコマンドである。

Rule Base には以下のような例外事由がありうる。

売買契約に基づく代金支払請求権(_売主,_買主,契約(売買,_売主,_買主,_物,_合意時))<=
**　　合意(売買,_売主,_買主,_物,_合意時).**
例外事由(売買契約に基づく代金支払請求権(_売主,_買主,_契約),同時履行の抗弁(_買主,_売主,_契約,_契約)).

「売買契約に基づく代金支払請求権」のようなルールに対して例外事由となるルールが存在していれば、そのルールを記述する。この場合は、「同時履行の抗弁」が成り立てば、例外が成り立ち、「売買契約に基づく代金支払請求権」は成り立たない。問題文から、ルールに対して適切な例外事由を選択する。

3. 提案手法

将来的な自動解答の方法として、以下の手順を提案する:

1. 問題文を自然言語解析し、Rule Base のトップルール(請求権存在()述語に入るべき述語名)を特定する
2. 上記で特定した述語の要件事実を、要件事実群をまとめたファイルから検索して Rule Base を生成する。
3. 生成した Rule Base から、Fact Base で記述すべき述語を抜き出し、引数などを Fact Base の形式に変換する。引数については、問題文と照合して定数と変数を分類する。
4. 完成した質問、Rule Base、Fact Base を Prolog 実行系で実行する。

上記の手法の準備として、自然言語処理を必要としない部分をプログラムで作成し、自然言語処理を行った場合の自動解答の可能性について分析を行った。結果と分析は後述する。

言語処理を必要とする部分については、主に三つある。

まず、Rule Base のトップルールを判断する部分である。問題文を解析し、何を問われているのかを特定し、述語名の候補群と照合する。

次に、例外事由としてルールを選択する場面である。同様に問題文を解析し、例外事由の候補群と照合する。

最後に、Fact Base の生成である。Fact Base は、「_対象物」→「土地」など、Rule Base における表現を事実に基づく単語に置き換えていた。ただし、本研究の目的は正しい PROLEG ルールの生成ではなく、問題の解答である。問題を解きたいだけなら、「土地」に置き換えずとも、「対象物」と、ただアンダースコアを外して定数に変換するだけでよい。言語処理を必要とするのは、どの引数を定数とし、どの引数を変数とするかという分類である。

先の例の引数のうち、「_相手方」「_対象物」「_所有者」「_物」については、問題文に存在が明記されている。このような引数は、Fact Base において定数とする。「_占有時」「_もと所有時」「_問題となる時点」については、問題に時間的要素の記述がないため、Fact Base において変数とする。これをまとめると、以下の Fact Base が完成する。

%Fact Base
現時点(_占有時).
主証(占有(相手方,対象物,_占有時)).
主証(もと所有(所有者,物,_もと所有時,_問題となる時点)).

実際に Fact Base として上記の形式を用いたとしても、問題の解答結果は一致する。よって、Fact Base の生成においては、問題文に明示的な記述があるかないかによって、定数と変数を分類すればよい。ただし、定数と変数を分類するだけでは、不整合が起き、正しく解けない問題も例外としてある。

4. 実験

自然言語処理を行った場合の自動解答の可能性を分析するために、自然言語処理を行わない部分についてプログラムを作成した。具体的には、Rule Base トップルールから要件事実をたどり、Rule Base を生成する部分については、過去問の Rule Base のいくつかをまとめた仮想的な要件事実群ファイルを作り、それについて、指定したルールから要件事実を末端まで展開し、出力するプログラムを作成した。

Fact Base を生成する部分については、先の出力の Rule Base から、Fact Base に記述すべき末端の要件を抜き出し、主証()述語で囲む。引数の分類は、言語処理を行わないため確実に分類ができないが、「_契約時」のような時間的要素や、「_請求額」のような数値的要素は、問題の事実に関わらないことが多いため、今回は、それらを変数、それ以外を定数として扱った。

5. 結果と分析

司法試験の過去問題のうち、72 問について分析を行った。分析する問題として、佐藤らが手動で PROLEG 化を行った、平成 18 年から平成 27 年までの問題を用いた。これらの司法試験過去問で自然言語処理が必要な三つの作業について、容易性を判断し、分析した。

まず、Rule Base のトップルールを判断する部分について、「述語名と同じ言葉が問題文に含まれている」「述語名に近い表現が問題文に含まれる」「述語名に近い表現が問題文に含まれない」の 3 つに分類した。具体的な例を示す。**図 1** は「述語名と同じ言葉が問題文に含まれる」の例、**図 2** は「述語名に近い表現が問題文に含まれる」の例、**図 3** は「述語名に近い表現が問題文に含まれない」の例である。

[問題文]
土地の賃貸人が借地契約の更新拒絶をするためには，正当の事由がなければならないほか，契約期間の満了の１年前から６か月前までの間に賃借人に対して更新をしない旨の通知をしなければならない。

請求権存在(更新拒絶(賃貸人,賃借人,契約(賃貸借,賃貸人,賃借人,土地,建物所有目的,平成 1 年 1 月 1 日),平成 30 年 1 月 1 日)).

図 1 平成 18 年度短答式民事系第 24 問選択肢エ

図 1 の問題では、問題文に「更新拒絶をするためには、」という表現があり、この問題の主たる述語が「更新拒絶」であることは明らかな上、Rule Base に「更新拒絶」という完全に一致するルールがある。

[問題文]
ＡＢが共有している建物の管理費用をＡが立て替えた場合，Ａは，Ｂからその共有持分を譲り受けたＣに対し，当該立替金の支払を請求することができる。（正）

請求権存在(立替金支払請求権('A','C',管理費用(共有建物,'A 支払額',平成 2 年 1 月 1 日))).

図 2 平成 18 年度短答式民事系第 26 問選択肢 5

図 2 の問題では、問題文に「立替金の支払を請求する」という表現があり、Rule Base トップルールの述語名「立替金支払請求権」と容易に結びつけることができる。

[問題文]
被相続人Ａから相続開始前に甲不動産を買い受けたＸは，Ａの唯一の相続人Ｂの債権者ＹがＢに代位して甲につきＢの相続登記をした上で甲を差し押さえた場合，登記がなくても，甲の所有権取得をＹに対抗することができる。（誤）

請求権存在(異議権の存在('X','Y',甲不動産)).

図 3 平成 19 年度短答式民事系第 8 問選択肢ア

図 3 の問題では、Rule Base トップルールの述語名「異議権の存在」について、問題文に「異議」という単語もなく、表現から見ても結びつけることは難しいと判断した。分析結果を**表 1** に示す。

表 1 Rule Base トップルールの述語名判断容易性の分析結果

分類	問題数	割合
Rule Base トップルールの述語名と同じ単語が問題文に含まれている	3	4.2%
Rule Base トップルールの述語名に近い表現が問題文に含まれる	30	41.7%
Rule Base トップルールの述語名に近い表現が問題文に含まれない	39	54.2%

次に、例外事由としてルールを選択する場面について、「Rule Base に例外事由が含まれない」「例外事由のルールの述語名と同じ言葉が問題文に含まれる」「例外事由のルール名と近い表現が問題文に含まれる」「例外事由のルール名と近い表現が問題文に含まれない」の 4 つに分類した。分類の基準は、Rule Base トップルールと同じである。分析結果を**表 2** に示す。

表 2 例外事由のルールの判断容易性の分析結果

分類	問題数	割合
Rule Base に例外事由が含まれない	22	30.6%
例外事由のルールの述語名と同じ言葉が問題文に含まれる	1	1.4%
例外事由のルール名と近い表現が問題文に含まれる	9	12.5%
例外事由のルール名と近い表現が問題文に含まれない	40	55.6%

最後に、Fact Base の生成の容易性について、分析した。Fact Base の生成においては、問題に事実として記述されている引数を定数とし、記述されていない引数を変数として分類を行う。これについて、「_契約時」のような時間的要素や、「_請求額」のような、数値的要素は、問題に記述されていないことが容易に判断できるため、時間的要素や数値的要素が変数、それ以外が定数となっている問題を「定数と変数の分類が簡単」、そうでない問題を「定数と変数の分類が難しい」として、分類した。ただし、前述したとおり、Fact Base の生成においては、定数と変数を分類するだけでは正しく解けない例外的問題がある。例外的問題はまだ定量的に検証できていないが、このために実際にはもう少し解ける割合が少ない。分析結果を**表 3** に示す。

表 3 Fact Base の生成容易性の分析結果

分類	問題数	割合
定数と変数の分類が簡単	51	70.8%
定数と変数の分類が難しい	21	29.2%

これらはそれぞれの要素の結果を別々にまとめたものだが、次にこれら 3 種類の要素をあわせたものを考える。Rule Base のトップルールや例外事由のルールの判断容易性が「近い表現が問題文に含まれる」以上であれば、問題を自動解答する見込みが高い。Fact Base については、「定数と変数の分類が簡単」であれば同様に見込みが高い。これに基づいた分析結果を**表 4** に示す。

表 4 問題正答容易性の分析結果

分類	数	割合
Rule Base のトップルールの述語名判断容易性が「近い表現が問題文に含まれる」以上かつ 例外事由のルールの判断容易性が「近い表現が問題文に含まれる」以上 かつ Fact Base の生成容易性が「定数と変数の分類が簡単」	16	22.2%

Rule Base のトップルールの述語名や、例外事由のルールに関する自然言語処理がうまくできれば、22.2%の問題が解けるはずである。ただし、前述したように、Fact Base の生成容易性に関して、例外的な問題がある。例外となる問題は、実際に PROLEG を実行してみないと判断が付かず、時間がかかってしまうため、検証が完了していない。この検証の結果によっては、解ける問題の割合は下がる可能性がある。それを踏まえた概算をすると、15%～20%の問題が自動で解ける見込みである。

6. 議論と展望

今後、自動解答の実現、正答率の向上に向けて、やるべきことが三つある。

一つ目は、**表 4** に関して述べた、Fact Base の生成について検証を済ませること。Fact Base を生成する際に、定数と変数を分類するだけでいい問題がいくつあるかによって、自動解答の可能性は大きく変わる。

二つ目は、問題文の自然言語処理である。これまで我々が作成してきた司法試験自動解答のための問題文述語項解析器があり[8][9]、これをさらに PROLEG 生成に適する形に改善したい。

三つ目は、PROLEG ルールの設計、特に Rule Base についてである。過去問分析の項で、Rule Base トップルールや例外事由のルールの判断容易性について記述したが、その判断容易性を、PROLEG のルールを再設計することで改善できると考えている。PROLEG のルールを、問題文の言語処理結果と照合しやすい形に変更することを考えている。例を**図 4** に示す。

[問題文]
書面によらない贈与の受贈者は，贈与者に対して贈与の履行を求めることができない。

請求権存在(契約成立(贈与,受贈者,贈与者,契約(贈与,受贈者,贈与者,対象物,平成 2 年 2 月 2 日))).

%Rule Base
契約成立(贈与,_受贈者,_贈与者,契約(贈与,_受贈者,_贈与者,_対象物,_契約成立時))<=
合意(贈与,_受贈者,_贈与者,契約(贈与,_受贈者,_贈与者,_対象物,_契約成立時)).

図 4 平成 22 年度短答民事系第 22 問選択肢ア

図 4 の問題では、Rule Base トップルールの契約成立()という述語の中に”贈与”という引数があり、契約成立が”贈与に関する”ことを、引数で表現している。契約成立()という述語は贈与以外にも成り立つ表現であり、この状態では「贈与に関する契約成立」のルールを探すために、引数まで解析しなくてはならない。これを改善するためには、ルールを

贈与に関する契約成立(_相手方,_契約主,契約(_相手方,_契約主,_対象物,_契約成立時))<=

とするのが理想的であると考えられる。このように、自然言語処理によって照合しやすい一般的な形にルールを再設計する予定である。

7. 結論

本研究では、司法試験問題の自動解答へ向けて、論理型言語 PROLEG を用いたアプローチのために、手動で PROLEG ルールが作成された過去問 72 問の分析を行った。分析を行うために、自然言語処理を必要としない部分についてはプログラムを作成し、自然言語処理ができれば PROLEG の自動生成ができることを確認した。

自然言語処理を必要とする部分について、自動化の容易性を分析した。問題文と PROLEG ルールの結びつきの難易度を、自然言語処理が必要な三つの場面についてそれぞれ難易度の段階を設定して分類した。その結果、全体で 15%～20%の問題は現状の手法で自動解答できる見込みがあると分かった。

今後は、自然言語処理部分の実装と、解答できる問題を増やすために、ルールの自然言語処理向け再設計を目指す。ルールの再設計については、手動で作成されたルールについて、言語処理をしやすいよう、問題文の形式に沿ったルールや述語の名前に変更する手法を提案した。

謝辞

本研究は JSPS 科研費 17H06103 の助成を受けたものである。国立情報学研究所の佐藤健教授及び藤田恵氏には、過去問の PROLEG 化ファイルの提供、PROLEG の仕組みについてご教示いただいたことを、心から感謝する。

参考文献

[1] M.-Y. Kim, R. Goebel, and K. Satoh, “Coliee-2015: evaluation of legal question an- swering,” in Ninth International Workshop on Juris-informatics, 2015.
[2] M.-Y. Kim, R. Goebel, Y. Kano, and K. Satoh, “COLIEE-2016: Evaluation of the Competition on Legal Information Extraction/Entailment,” in Tenth International Workshop on Juris-informatics, 2016.
[3] Y. Kano, M.-Y. Kim, R. Goebel, and K. Satoh, “Overview of coliee 2017,” Epic Ser. Comput., vol. 47, pp. 1–8, 2017.
[4] M. Yoshioka, Y. Kano, N. Kiyota, and K. Satoh, “‘Overview of Japanese Statute Law Retrieval and Entailment Task at COLIEE-2018,’” in Twelfth International Workshop on Juris-informatics, 2018.
[5]I.Bratko (1986) . “Prolog programming for artificial intelligence” Addition-Wesley Publishing Company(安部憲広(訳)(1990). 「Prolog への入門」近代科学社)
[6]佐藤 健,浅井 健人,古川 昂宗,久保田 理広,中村 恵,西貝 吉晃,白川 佳,高野 千明, “PROLEG:論理プログラミングをベースとした民事訴訟における要件事実論の実装”
[7]Maria Navas-Loro, Ken Satoh, and Victor Rodriguez-Doncel,"ContractFrames:Bridging the gap between Natural Language and Logics in Contract Law"
[8]R. Hoshino, R. Taniguchi, N. Kiyota, and Y. Kano, “‘Question Answering System for Legal Bar Examination using Predicate Argument Structure,’” in Twelfth International Workshop on Juris-informatics, 2018.
[9]R. Taniguchi, R. Hoshino, and Y. Kano, “‘Legal Question Answering System using FrameNet,’” in Twelfth International Workshop on Juris-informatics, 2018.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

司法試験の問題分類と法律文書における人物関係の構造化辞書

Problem classification of legal bar exam and structured dictionary design for human relationships in legal documents

清田 直希[*1]
NAOKI KIYOTA

狩野 芳伸[*1]
YOSHINOBU KANO

静岡大学情報学部[*1]

Abstract: Legal document processing is one of promising applications of natural language processing. However, legal sentences are very complicated and require various difficult analysis techniques. We focus on the competition workshop COLIEE (Competition for Legal Information Extraction and Entailment) that aims to automatically answer the legal bar exam. Firstly, we manually classify the past legal bar exam data distributed in COLIEE by required analysis techniques such as conditional sentence extraction and negative form interpretation. Secondly, we focus on person relationships among our classifications, propose a dictionary design for legal document processing based on verbs and their arguments with roles of persons.

1. はじめに

近年の技術革新と共に注目を集めている分野が自然言語処理である。計算機が、人間が日常的に利用している自然言語を処理するために様々な研究がおこなわれている。

その中でも、法律文書の処理に関しては、一般的な文書の言語処理に加えて、専門性の高い言語が多く用いられている。そのうえ、文章中の事象を客観的に評価する必要があるため、難易度は高いが自然言語処理の分野において期待が高まっている。本研究では、毎年開催されている我が国の司法試験自動解答を行うコンテスト型ワークショップ COLIEE (Competition for Legal Information Extraction and Entailment)タスクシリーズ[1][2][3][4]で配布されたデータを対象とし、法律文書処理の基盤リソース構築を狙う。

法律文書においては先述の通り、専門性の高い言語に加え、構文、意味役割、言い換え表現など高度かつ多様な解析技術を必要とする複合的な要素がある。中でも、人物関係は司法試験を解くうえでの重要な問題分類の一つであり、こうした処理には用語の構造化された整理が必要であると考えられる。そこで、人物関係に着目した構造化辞書を提案する。

2. 司法試験の自動解答タスク COLIEE

COLIEE は、COLIEE 2014、COLIEE 2015 [1], COLIEE 2016[2]と人工知能学会国際シンポジア(JSAI-isAI)のひとつである JURISIN(法情報学ワークショップ)において開催してきた。COLIEE 2017[3]は法と人工知能に関するトップカンファレンスである ICAIL (International Conference of Artificial Intelligence and Law)において開催された。COLIEE 2018 ではカナダ法を用いたタスクがが追加されたが、本稿では我が国司法試験を用いたタスクのみを対象とする。

司法試験は弁護士や裁判官、検察官となろうとする者を対象に行われる試験である。問題形式は択一問題を含む短答式と論文式に分けられて行われる。試験の科目数については、短答式が憲法、民法、刑法の 3 科目であり、論文式が公法系科目(憲法及び行政法に関する分野)、民事系科目(民法、商法及び民事訴訟法に関する分野)、刑事系科目(刑法及び刑事訴訟法に関する分野)、選択科目(知的財産法、労働法などの中から 1 科目選択する)の 4 科目である。

COLIEE ではこのうち民法短答式試験を対象としている。多択式を二択にブレークダウンするとともに、解答に関する条文を検索する情報検索タスク(Information Retrieval)と、二択を Yes/No で答える含意関係タスク(Entailment)を用意した。また知識源として民法の全条分を配布し、問題文共々原文の日本語に加え、英語に翻訳したものを配布している。

COLIEE は与えられた問題文に対し、民法条文から関連する条文を返す情報抽出タスク(Phase 1)、問題文と与えられた関連条文について二値の安定をする含意関係タスク(Phase 2)、関連条文を与えずにこの二つを同時に行う質問タスク(Phase 3)の三つのサブタスクからなる。COLIEE 2017 においては、時期の問題から Phase 2 を実施せず、Phase 1 を Task 1、Phase 3 を Task 2 と呼称して実施した。以下で各タスクを定式化する。

情報抽出タスクは、与えられた問題文 Q を回答するのに関連する民法条文のサブセット S1, S2, …, Sn を民法全体から抽出するタスクとなる。解答すべき条文は複数のこともある。

含意関係タスクは、与えられた問題文 Q を、同時に与えられた民法条文のサブセット S1, S2, …, Sn が含意するか、すなわち Entails(S1, S2, …, Sn, Q)または Entails(S1, S2, …, Sn, not Q)を回答するタスクである。

質問応答タスクは与えられた問題文 Q について、民法全体を用いて Yes あるいは No を回答するタスクである。

3. 司法試験の要求要素技術別問題分類

司法試験を解答するうえで、背後にある構文構造、人物の役割と関係性、論理、抽象性など多様で複雑な構造を的確に処理をする必要がある。そこで平成 29 年度のテストデータを対象に、解答に必要な要素を設定した基準を基に手作業で分類をした。表 1 は、解答に必要な要素を手作業で分類した結果と、COLIEE2018 参加チームの要素技術別の正解数及び正答数をまとめた表である。分類の基準は表 2 に記した。この分類は、要素技術ごとに排他的ではないため、ある問いに対して複数の要素技術を要することもある。以下では、分類した要素技術(タスク)のうち、テストデータに頻出なものについて簡単な説明を述べる。

3.1 条件文抽出

このテストデータにおいて、最も多くカウントされた要素技術は、条件文抽出である。これは問題文・条文のそれぞれの条件文を調べることで正誤判定ができる可能性がある問題である。例えば以下のような問いである。

第百六十条　相続財産に関しては、<u>相続人が確定した時、管理人が選任された時又は破産手続開始の決定があった時から六箇月を経過するまでの間</u>は、時効は、完成しない。

	問題数	UA[6]	正答率	UE	正答率	KIS_mo3[7]	正答率	KIS_Frame[7]	正答率
条件文抽出	31	19	0.61	9	0.29	13	0.42	16	0.52
人物役割付与	27	16	0.59	12	0.44	14	0.52	13	0.48
人物関係抽出	26	14	0.54	11	0.42	13	0.50	10	0.38
形態素解析	25	18	0.72	16	0.64	20	0.80	16	0.64
照応解析	20	13	0.65	5	0.25	12	0.60	13	0.65
曖昧性解消	17	9	0.53	7	0.41	8	0.47	9	0.53
意味役割付与	15	8	0.53	6	0.40	9	0.60	6	0.40
動詞言い換え	13	7	0.54	6	0.46	7	0.54	4	0.31
一般辞書	9	5	0.56	5	0.56	4	0.44	4	0.44
述語項解析	9	8	0.89	3	0.33	5	0.56	5	0.56
否定形解釈	7	5	0.71	3	0.43	5	0.71	7	1.00
法律系用語	7	4	0.57	2	0.29	2	0.29	3	0.43
含意関係	5	2	0.40	2	0.40	1	0.20	2	0.40
係り受け	5	3	0.60	1	0.20	2	0.40	4	0.80
条文検索	5	3	0.60	2	0.40	3	0.60	4	0.80
言い換え	5	2	0.40	4	0.80	3	0.60	3	0.60
箇条書き	3	1	0.33	2	0.67	1	0.33	2	0.67

表 1. 平成 29 年度のテストデータの問題の要素技術別分類と COLIEE 参加チームの正答数及び正答率

問い:相続財産に関しては，相続財産管理人が選任された場合でも，相続人が確定するまでの間は，時効は完成しない。

下線部が文章中の条件と考えられる。問いでは相続財産の時効不成立の条件が正誤の判断材料になる。前者は「相続人が確定し、管理人選任(または破産手続き決定)から 6 カ月が経過するまで」という条件下では時効が完成しないのに対し、後者は「管理人が選任されて、相続人が確定するまで」という条件になっている。このように 2 つの文は条件が異なっているため、答えは No となるのである。

3.2 形態素解析

また、形態素解析が特に必要となる問題も数多く見られた。形態素解析は自然言語処理において重要なタスクの一つであるため、特記をする必要性は感じられないかもしれない。ここでの分類基準は、係り受け(単語の修飾関係)や述語項解析が判断基準となる問いの他、条文がおおむねそのまま問いになっているものも含まれる。つまりは形態素解析さえあれば解ける問いも少なからず見受けられるということになる。こうした意味でも形態素解析という分類基準を追加した。

3.3 人物役割付与・人物関係抽出

人物役割付与は、問題文に登場する人物(例えば A など)が条文上のどの役割を担っているのかを明確にするタスクである。人物関係抽出は、売買契約や婚姻関係等の複数の人物間の立場やその関係性を示すために必要となるタスクである。これらについては次節で述べる。

分類名	分類規準
条件文抽出	問題文・条文のそれぞれの条件文を調べることで正誤が判断できる可能性がある問題
人物役割付与	問題文に出てきた人物(A など)がどの役割(未成年,買主,債務者 等)を担っているのかを明確にする必要がある問題
人物関係抽出	複数の人物間の立場や役割を明確にする必要がある問題
形態素解析	おおむね同じような文章である問題　又は文章を解析しやすくするために特定の格助詞などを入れるのに用いる
照応解析	文章の一部を省略している際に、それが何を指しているのかを明確にする際に用いる。「こそあど言葉」のみだけではなく「前文」「同項」なども対象となる
曖昧性解消	曖昧な表現が用いられている文を解析するのに用いる 例:H29-4-A 2　前項の規定による意思表示の無効は、善意の第三者に対抗することができない。
意味役割付与	問題文の中で用いられている名詞が動作主なのか・その動作の対象なのか、手段なのか等を特定する必要がある問題
動詞言い換え	動詞が言い換えられている問題
一般辞書	法律系の言語ではない一般生活で用いられるような事物が出てきている問題(例:冷蔵庫、会社)
述語項解析	その問題の動作を明確にする必要がある問題。特に「～した」「～できる」などはこうした動詞を抜き出すのではなく、その行為を抜き出す必要がある
否定形解釈	否定形になっていて正誤判定をする問題　又は「～できない(不可能)」「～しなければならない(義務)」など「ない」が用いられていてもその意味が異なる問題
法律系用語	法律に関する言語が言い換えられている問題(例:制限行為能力者→未成年など)
含意関係	条文から隠された意図を把握する必要がある問題 ピカソが『ゲルニカ』を発表した→ピカソは『ゲルニカ』の作者である
係り受け	主語と述語の関係、単語の修飾又は文の並列関係に着目する必要がある問題
条文検索	条文中に別の条文が指定されている問題
言い換え	動詞以外で言い換えがされている問題(例:一→1)
箇条書き	条文で箇条書きが用いられている問題。

表 2. COLIEE 司法試験問題の要素技術別分類基準

3.4 照応解析

照応解析は、文章中で登場する「これ」や「あの」などの指示内容を明確にする解析である。前述の指示詞に加えて、「前章」や

「同項」など司法試験に特有の指示表現もあり、言語処理においてはその内容を正しく理解する必要があると考えられる。

3.5 曖昧性解消

曖昧性解消は、文章中に登場する意味的に曖昧な用語を適切に捉える必要がある問いにつけるタスクである。曖昧な用語とは以下のようなものである。

第四百七十五条　弁済をした者が弁済として他人の物を引き渡したときは、その弁済をした者は、更に有効な弁済をしなければ、その物を取り戻すことができない。

ここで述べられている「物」というのが曖昧な用語に該当する。このほかにも名詞であれば「事由」や「意思」、動詞であれば「する」や「ある」といったものが該当する。こうした曖昧性は、複数の対象あるいは幅広い語義を含み、文章理解が困難である。

3.6 意味役割付与

意味役割付与は、文章中の語彙に動作主や場所、対象といった意味的な関係を付与するタスク、SVM[5]やニューラルネットワーク[5]を用いた機械学習手法の提案がされている。

3.7 否定形解釈

否定形解釈は、条文あるいは問いの否定形を正しく解釈したうえで、正誤の判断材料にするタスクである。条文の内容を否定形に反転した司法試験の問題も見られる。

第六百七十五条　組合の債権者は、その債権の発生の時に組合員の損失分担の割合を知らなかったときは、各組合員に対して等しい割合でその権利を行使することができる。

問い：組合の債権者は，各組合員に対して，その権利を行使することができない。

このように条文では権利を行使できる旨が書かれているが、問いでは権利の行使ができないことになっている。このように否定形にすることにより答えが偽になる問題も司法試験では見られ、それは文末の「ない」を認識できれば簡単に評価できる。ところが、次のような問いについてはそうもいかない。

第六百三十三条　報酬は、仕事の目的物の引渡しと同時に、支払わなければならない。ただし、物の引渡しを要しないときは、第六百二十四条第一項の規定を準用する。

問い：請負人は，仕事の目的物の引渡しを要する場合には，これを引き渡した後でなければ，報酬を請求することができない。

前者の文の特殊条件である「ただし」以下の文章を除けば、この 2 文の文末は「ない」が用いられている。しかし、前者の文は「義務」を表した文であり、一方で後者は「否定」を表している文であるため、この 2 文は偽の関係となる。つまり、「ない」に内在される意味を把握しないとこの違いを見抜くことは難しい。

このように問題を要素技術ごとに分類することで、複雑な問題構造を分解した検討ができ、今後の研究の新たな指針になる。

4. 自動解答タスクと人物関係抽出

平成 29 年度のテストデータは全 69 問あり、そのうち言語処理を行う際に、人物関係を抽出する必要のある問い及び人物役割を付与する必要のある問いは、それぞれ 26 問、27 問存在した。しかし、現状の解答器では人物間の関係はほとんど考慮されておらず、表層的な処理だけとなっており、人物関係を抽出するタスクの解決は有用である。以下に人物関係を要する問いの例を挙げる。

甲土地を所有する A が B と通謀して甲土地を B に仮装譲渡し，A から B への所有権移転登記がされた後，B の債権者 C が甲土地を差し押さえた場合において，その差押えの時に C が仮装譲渡について善意であったときは，A は，C に対し，B への譲渡が無効であることを主張することができない。

この問いは正しい主張であるので、計算機は、Yes を出力する必要がある。この問いでは、A、B、C という 3 人の人物が登場していることが明らかである。しかしこのアルファベットには人物のラベルという以上の意味がない。重要なのはこの A~C の人物がどういう関係にあるのかということを計算機が読み取ることである。ここで Yes の根拠となる条文を示す。

第九十四条　相手方と通じてした虚偽の意思表示は、無効とする。

2 前項の規定による意思表示の無効は、善意の第三者に対抗することができない。

A という人物をこの問いの主語とすると、B は相手方、C は善意の第三者という関係が読み取ることができる。よって、A が B に虚偽の意思表示である土地の仮装譲渡を行った後、(A にとっての)善意の第三者である C に土地を差し押さえられたとしても、行為者 A は善意の第三者に対抗できないため、譲渡の無効を主張できない。

こうした人物関係について、動作主、被動作主、第三者などの属性をもつ辞書を構築することで、述語項構造解析による主格・目的格といったツールの出力と紐づけて自動的に人物関係を抽出することが可能になる。さらに、省略されている人物や事物の推測が容易になると考える。

5. 法律文書処理のための構造化辞書

先述の分類のうち、人物関係を対象として意味的関係を含めた構造化辞書を提案する。表 3 は H29 年の司法試験のデータを基に作成した辞書の例である。一つの関係を一エントリとして表現する。関係そのものを表す行為とは、動作主と被動作主の立場を端的に説明する動詞である。また動作主・被動作主・第三者は、条文を基に抽出した。第三者とは、動作主と被動作主の関係を前提に、さらに何らかのアクションをとる存在である。これらの項目は、自然言語処理システムの利用を想定して実際の出現形を基本形にしたものを登録する。以下の条文を例に説明する。

第四百七十五条　弁済をした者が弁済として他人の物を引き渡したときは、その弁済をした者は、更に有効な弁済をしなければ、その物を取り戻すことができない。

第四百七十七条　前二条の場合において、債権者が弁済として受領した物を善意で消費し、又は譲り渡したときは、その弁済は、有効とする。この場合において、債権者が第三者から賠償の請求を受けたときは、弁済をした者に対して求償をすることを妨げない。

この場合は、動作主である「債権者」と被動作主である「弁済をした者」が関係そのものを表す行為「債権を有している」で結ばれている。「第三者」は、この関係を前提に、「第三者の行為」として、「債権者」に対して消費又は譲渡した物についての「賠償を請求」している。

今回は 1 年分のテストデータから抽出したが、エントリを増やすことで条文すべてを網羅できるであろう。さらには動作主・被動作主の属性を記述できると有用である。例えば、成年被後見人とは、精神上の障害により判断能力が欠けており、家庭裁判所から後見開始の審判を受けた人を指す。その成年後見人はこうした人々を後見する(法律行為などを代理したり補助したりする)人物である。こうした人物の可能な法律行為やそれぞれの関係性はエントリを増やすことによって得られるのではないか。

6. 議論と展望

法律文書の辞書化における今後の課題をふたつ述べる。

抽象的な表現の扱いは難しい。例えば先の問いに対する条文では、「虚偽の意思表示」という抽象的な表現が用いられてい

た。具体的な記述がされた問いの文章から条文を検索するためには、問いの文章を抽象化する必要があるだろう。動詞に関しても同様である。「認める」「得る」「…がある」などといった一般性の高い動詞はよく用いられるが、解釈の幅が広く曖昧である。こうした動詞では動詞そのものよりもその項に具体的な表現がなされることが多いため、動詞そのものは利用されないことが多い。動詞の抽象化については、階層的に捉え、包含関係を表す手法[8]などが提案されているが、一般性が高いほど、動詞の置換は慎重な考慮が必要である。

もうひとつの課題は第三者の扱いについてである。先述のように、ここでの第三者には幅広く異なる立場のものが含まれる。権利を行使する立場であったり、許可をする立場であったりとその行為も多種多様である。次のステップとしては、行為の手がかりとして、表層格を用いた行為の判断を検討している。

7. おわりに

本研究では司法試験自動解答を念頭に、法律文書処理のリソース構築のため、司法試験の必要要素別分類を行った。さらにその中でも人物関係に着目して、人物関係の意味をとらえるための構造化辞書の設計と、手作業での辞書エントリサンプルの作成を行った。今後は作業やサンプルとする条文の数が増していく。これにより法律分野に現れうる単語を十分に網羅できる辞書構築を目標としている。こうしたリソースにより、自然言語処理による自動解答、自動処理に貢献できる。今後の研究では辞書を使った人物関係認識の自動化を実装する予定である。

謝辞 本研究は JSPS 科研費 17H06103, JSPS 科研費萌芽研究、JST CREST の助成を受けた。

参考文献

[1] Chang, C. L. and Lee, R. C. T.. Symbolic Logic and Mechanical Theorem Proving. Academic Press, 1973, 331p.

[2] M.-Y. Kim, R. Goebel, and S. Ken, “COLIEE-2015: Evaluation of Legal Question Answering,” in Ninth International Workshop on Juris-informatics (JURISIN 2015), 2015.

[3] M.-Y. Kim, R. Goebel, Y. Kano, and K. Satoh, “COLIEE-2016:Evaluation of the Competition on Legal Information Extraction/Entailment,” in Tenth International Workshop on Juris-informatics (JURISIN 2016), 2016.

[4] Y. Kano, M.-Y. Kim, R. Goebel, and K. Satoh, “Overview of COLIEE 2017,” in the 16th International Conference on Artificial Intelligence (ICAIL 2017), 2017.

[5] 岡田拓哉,竹内孔一,石原靖弘:ニューラルネットワークを利用した日本語意味役割ラベル付与システムの構築,言語処理,2018

[6] M.-Y. Kim and R. Goebel, “Two-step Cascaded Textual Entailment for Legal Bar Exam Question Answering,” in the 16th International Conference on Artificial Intelligence and Law (ICAIL 2017), 2017.

[7] Y. Kano, R. Hoshino, and R. Taniguchi, “Analyzable Legal Yes/No Question Answering System using Linguistic Structures,” in 4th Competition on Legal Information Extraction and Entailment (COLIEE 2017), 16th International Conference on Artificial Intelligence and Law (ICAIL 2017), 2017.

[8] 竹内孔一,乾健太郎,竹内奈央,藤田篤: 意味の包含関係に基づく動詞項構造の細分類, http://www.cl.cs.okayama-u.ac.jp/wp-content/uploads/2017/03/golex.pdf , 2017

関係そのものを表す行為	その他とりうる行為	動作主	被動作主	第三者	第三者の行為
代理する,代表する	責任を負う,同意する	代理人,法廷代理人	(本人),未成年者,制限行為能力者	家庭裁判所	認める
	同意を得る,財産を処分する,通知を発する	(本人),未成年者,制限行為能力者	代理人,法定代理人		
保佐する・補助する	追認する,	保佐人・補助人法定代理人	制限行為能力者	制限行為能力者の相手方	(制限行為能力者に)催告する,(保佐人・補助人に)行為を取り消す
所有する	償金を支払う,償金を請求する,通行する,譲り渡す,囲障を設ける,持分を取得する,使用する,分担する	主たる所有者(本人)	従たる所有者,その他の所有者,その他の分割者		
債権を有する	取り立てる,(第三債務者に)供託させる	質権者,債権者	債務者	第三債務者	
	譲り渡す,	債務者	債権者	連帯債務者	更改する
	(動作主が)取り戻す,(債権者が)消費する,(債権者が)譲渡する,(動作主に)求償する	弁済をした者		第三者	物を引き渡される,賠償の請求をする
売買する	代金の支払いを拒む,契約の解除,損害賠償の請求,支払う	買主	売主		
	請求する,契約の解除をする,権利を移転する	売主	買主	他人	

表 3. 法律文書の人物関係構造化辞書のエントリ例

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

法律構成を表す双極議論フレームワーク上での推論

Reasoning on a bipolar argumentation framework for a law

川崎 樹 Tatsuki Kawasaki　森口 草介 Sosuke Moriguchi　高橋 和子 Kazuko Takahashi

関西学院大学
Kwansei Gakuin University

We describe a method to construct a bipolar argumentation framework (BAF) from the logical structure of a law. A law consists of a set of ultimate facts to make the conclusion effective, and the defense to make it ineffective. We construct a BAF by considering them to the relations of attacks and supports between arguments, respectively. Moreover, we show a method of reasoning on the BAF that derives the conclusions, that is, the applicable laws, from given facts. The reasoning works bidirectionally: we derive the conclusions from given facts in a bottom-up manner, and find the required facts from these conclusions in a top-down manner to derive new applicable conclusions.

1. はじめに

現在，法的推論を計算機で行う試みが広く行われている．法的推論はある状況に対して法律を適用することで行うが，同じ状況下であっても法律の状況に対する適用の仕方（法律構成という）は異なる．

例 1. 例として表見代理と準占有者弁済について述べる．Y は X から物品を購入し，代金の支払いを X ではなくその代理人 Z に行ったとする．しかし，X は Z に代理権がないこと（無権代理）を主張し，Y に代金の請求を行った．これに対し，無権代理の主張を排斥するために抗弁として表見代理を主張した．

ここで，表見代理の主張が効果を発揮するために必要な条件としての事実（要件事実という）は以下の通りである．

a. 代理人の法律行為

b. 相手方が代理人に a の法律行為について代理権があると信じたこと

c. 相手方の b の誤信についての正当な理由の評価根拠事実

d. a に先立ち，a の法律行為以外のある特定の事項についての代理権の発生原因事実

このうち，要件事実 a〜c は準占有者弁済の主張の要件事実として捉えることができる．準占有者弁済の主張をすることでも無権代理の主張を排斥できるため，表見代理の主張をする代わりに，準占有者弁済の主張しても良い．このような場合，同じ状況下で同じ結論を導くものであるが，法律構成は異なるものができる．

他にも，複数の法律が同時に効果を発揮する状況は存在する．このような状況下で，依頼者がどの法律が効果を発揮しているのかを判断することは容易ではない．そのため，ある状況において考えられる事実からどの法律が効果を発揮するか，または特定の法律が効果を発揮するために他にどのような事実が必要かといった具体的な推論することは依頼者の理解の助けとなる．この法的推論を計算機で行うために法律の構造を具体化する試みがされている．その 1 つの方法として，要件事実を論理型言語で記述する PROLEG が提案されている [4]．PROLEG では要件事実論を条文ごとに 1 つの論理式で記述し，抗弁を例外として記述する．

また著者らは，PROLEG の記述から双極議論フレームワーク (BAF)[1] に変換する手法を示した [2]．BAF は支持と攻撃を持つ議論を扱うシステムであり，議論を議論中の論証と，論証間の反論と支持の 2 つの二項関係を用いて表現する．法律構成を一種の議論としてとらえると，法律間の抗弁の関係は攻撃関係に位置づけることが可能であり，法律と法律の効果を発揮するため根拠となる事実は支持関係に位置づけることが可能である．また，BAF は有向グラフ表現が可能であるため，具体的に関係性を示すことが容易に行える．著者らはこの BAF 上で原告と被告がそれぞれ論争に勝つために証拠として提示すべき事実を導く推論方法を提案した [3]

本研究は，法律構成についてより深く広い理解を得るために，法律構成を用いてどのような推論が可能かを示す手法の提携を目的としている．本論文では，条文から直接 BAF を構築する手法を述べる．また，BAF の上での推論を拡張し，根拠となる事実から，結論を導く手順について述べる．

本論文の構成は以下の通りである．まず第 2 節で双極議論フレームワークについて述べ，次に第 3 節で法律構成から双極議論フレームワークを構築する手法について述べる．その後，第 4 節で双極議論フレームワークを用いた推論の手法を示し，第 5 節でまとめを述べる．

2. 双極議論フレームワーク (BAF)

2.1 BAF

Cayrol らは，議論を攻撃と支持の 2 種類の関係を用いて表現した [1]．著者らはその BAF を拡張し，攻撃を論証間の二項関係，支持を論証と論証の集合の間の二項関係として扱う．

定義 1. 双極議論フレームワーク BAF は $baf = \langle AR, ATT, SUP \rangle$ の組として定義される．AR は論証の有限集合であり，ATT，SUP はそれぞれ $ATT \subseteq AR \times AR$，$SUP \subseteq (2^{AR} \setminus \{\emptyset\}) \times AR$ である．$(B, A) \in ATT$ を $\mathrm{att}(B, A)$，$(\mathbf{A}, A) \in SUP$ を $\sup(\mathbf{A}, A)$ と記述する．

BAF は論証をノード，攻撃関係と支持関係をエッジとした

連絡先: 川崎 樹，関西学院大学，兵庫県三田市学園 2-1，
dxk96093@kwansei.ac.jp

有向グラフで表現することが可能である．例えば以下の BAF のグラフ表現は図 1 である．

$baf = \langle \{a, b, c, d\}, \{(b, a)\}, \{(\{b, c\}, a)\} \rangle$

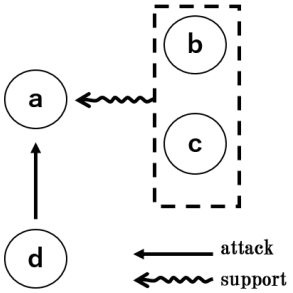


図 1: BAF の例

2.2 ラベリング

ラベリングは BAF における各論証の受理性を表す．議論の意味は議論中の各論証の受理性にあるため，ラベリングは BAF に意味を与える．

定義 2. $baf = \langle AR, ATT, SUP \rangle$ に対して，ラベリング $\mathcal{L}$ とは AR から $\{in, out\}$ への関数である．また，$\mathbf{A} \subseteq AR$ に対して，$\forall A \in \mathbf{A}, \mathcal{L}(A) = in$ のときに $\mathcal{L}(\mathbf{A}) = in$，そうでなければ $\mathcal{L}(\mathbf{A}) = out$ とする．

ここでは，法律構成が表す意味と同等な意味を BAF に与える．これを完全ラベリングと呼び，以下のように定義する．

定義 3. $baf = \langle AR, ATT, SUP \rangle$ において，任意の $\mathbf{A} \in AR$ に対するラベリング $\mathcal{L}$ が以下を満たすならばそのラベリングを完全ラベリングという．

- $(\forall B \in AR, \neg\mathrm{att}(B, A)) \wedge (\forall \mathbf{A} \subseteq AR, \neg\mathrm{sup}(\mathbf{A}, A))$ であるならば $\mathcal{L}(A) = in$.
- $(\forall B \in AR, \mathrm{att}(B, A) \Rightarrow \mathcal{L}(B) = out) \wedge (\exists \mathbf{A} \subseteq AR, \mathrm{sup}(\mathbf{A}, A) \wedge \mathcal{L}(\mathbf{A}) = in)$ であるならば，$\mathcal{L}(A) = in$.
- そうでなければ $\mathcal{L}(A) = out$.

完全ラベリングは法律構成における攻撃防御方法を表す意味論である．効果を発揮することを $\mathcal{L}(A) = in$ とし，効果が排斥されることを $\mathcal{L}(A) = out$ として表す．ある主張あるいはそれに対する抗弁が効果を発揮するためには，それぞれが持つ要件事実を満足し，かつ自身に対するあらゆる抗弁の効果が排斥されることが必要である．一般的なラベリングでは in でも out でもない論証に対して $undec$ のラベルを用いるが，最終的な法律構成において各主張はその効果を発揮するか排斥されるかの 2 値となるため，ここでは $undec$ を考慮していない．

3. 法律構成に準ずる BAF の構築

民事訴訟は攻撃と防御によって成り立つ．原告側が提出する主張や証拠などを攻撃とし，それに対して被告側が提出するものを防御としている．これは，主張・証拠事実によって互いを攻撃している関係と考えられる．つまり，お互いに提出する主張や証拠事実は攻撃関係を持つ．また，主張の効果を認めるためには適用する法律の要件事実を満たす必要がある．この時，法律と要件事実は結論と根拠の関係にあり，支持関係を持つ．このようにして考えることで，法律構成は双極議論フレームワークを用いて表現することができる．

3.1 存在/不存在論証

ある法律の要件事実を満たすためには，要件事実を満たすための法律もしくは，その要件事実に応じた事実が存在しなければならない．よって，すべての主張の根拠はそれぞれに応じた事実の存在であるといえる．この事実の存在を BAF 上で示すために，存在論証と不存在論証の 2 つを与える．存在論証は，要件事実によって求められた事実が存在することを示し，その要件事実を支持する論証である．逆に，不存在論証は要件事実によって求められた事実が存在しないことを表し，その要件事実を攻撃する論証である．存在論証と不存在論証はそれぞれ事実が存在するか否かを表し常に真である論証である．ここではある事実 a に対する存在論証を $\mathrm{ex}(a)$，不存在論証を $\mathrm{ab}(a)$ と表記する．

例 2. 例 *1* で示した表見代理の例をもとに *BAF* を作成する．表見代理の主張は無権代理の主張に対する抗弁であるため，表見代理の主張は無権代理の主張を攻撃している．また，表見代理の主張の要件事実は *4* つの事実が同時に満たされることで表見代理の主張が効果を発揮するため，*4* つの要件事実の集合は表見代理の主張を支持している．また無権代理の要件事実 x があるとする．これを *BAF* のグラフで表すと図 *2* になる．

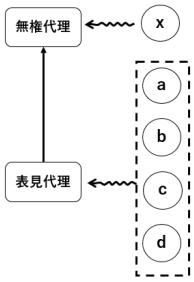


図 2: 表見代理の例に対する BAF

x の根拠が存在しているとする．また，表見代理の主張に対する *4* つの要件事実のうち，a〜c に対する根拠が存在し，d に対する根拠が存在しなかったとすると，この *BAF* には *3* つの存在論証と *1* つの不存在論証が追加され，その完全ラベリングは図 *3* になる．表見代理の主張におけるすべての要件事実が満たされてはいないため，表見代理の主張が効果を発揮せず，無権代理の主張が効果を発揮していることがわかる．

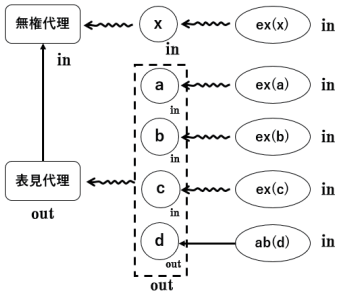


図 3: 完全ラベリング付き BAF

4. BAF の段階的推論

BAF を用いて，結論及びその他の必要とされる事実を推論する手法について述べる．推論は 2 段階からなる．まず，状況から得られる事実を存在論証として与えた時に得られる結論（結論集合）を導く．次に，新たな結論を引き出すために必要な他の存在論証（要求存在論証集合）を導く．

考えられる全ての法律構成を表す BAF が与えられていると仮定し，それを以下のように定める．

$*baf = \langle *AR, *ATT, *SUP \rangle$

4.1 結論集合の導出

初めに与える存在論証の集合を Ex とする．これらの存在論証に対して，それぞれが支持する論証が 1 つずつ存在するので，AR と SUP を次のように定める．

$AR = Ex \cup \{a | \mathrm{ex}(a) \in Ex\}$

$SUP = \{(\{\mathrm{ex}(a)\}, a) | \mathrm{ex}(a) \in Ex\}$

1. $a \in *AR, a \notin AR, \mathbf{A} \subset AR, (\mathbf{A}, a) \in *SUP$ を満たす論証 a 及び集合 $\mathbf{A}$ を探す．AR に a を，SUP に $(\mathbf{A}, a)$ を追加する． この追加を論証 a が見つからなくなるまで繰り返す．

2. $ATT = *ATT \cap (AR \times AR)$ とする

3. 1 で得た AR，SUP，2 で得た ATT を用いて $baf = \langle AR, ATT, SUP \rangle$ とし，この BAF に完全ラベリングを与え，集合 $S = \{s | \mathcal{L}(s) = in \wedge \forall t \forall \mathbf{A}(s \in \mathbf{A} \Rightarrow (\mathbf{A}, t) \notin SUP)\}$ を得る．この集合 S が Ex に対する結論集合である．

例 3. 図 *4* における結論集合は $S = \{S1, S3\}$ である．

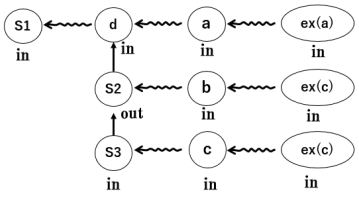


図 4: 結論集合の例

4.2 要求存在論証集合の抽出

次に新たな結論を得るために必要となる存在論証を推論する．

1. 結論集合 S の要素は以下の 3 つに分類される．それぞれ，(a)$*baf$ 内で支持も攻撃も行わない論証，(b) 支持は行わないが $*baf$ 内の他の論証に対して攻撃する論証，(c) その論証と他の論証を含む集合が存在し，$*baf$ においてその集合が支持する関係が存在するような論証，である．結論集合の要素を s とすると以下の論理式で表される．

   (a) $(\forall (\mathbf{A}, a) \in *SUP, s \notin \mathbf{A}) \wedge (\nexists b, (s, b) \in *ATT)$

   (b) $(\forall (\mathbf{A}, a) \in *SUP, s \notin \mathbf{A}) \wedge (\exists b, (s, b) \in *ATT)$

   (c) $\exists (\mathbf{A}, a) \in *SUP, \{s\} \subsetneq \mathbf{A}$

2. (c) の場合に， a とは異なる $b \in \mathbf{A}$ について，[3] の方法に基づいて $\mathcal{L}(b) = in$ とするような存在論証を求める．これによって得られた存在論証の集合が要求存在論証集合である．

例 4. 例えば，図 *5* において $S4$ は（c）に分類される．この時，$\mathcal{L}(\mathrm{f}) = in$ とするような存在論証を求める．

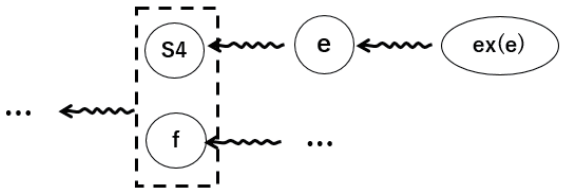


図 5: 要求存在論証集合の例

[3] で述べられた方法を行う際，ある S の要素 s が攻撃する論証が出現した場合，その論証は常に排斥されるため，枝刈りが可能である．これらの推論を繰り返すことによって，結論及び，その結論の効果を発揮させるために必要な存在論証を求めることができる．

5. まとめ

法律が持つ構成要件と法律における抗弁の関係を用いて，法律構成から BAF を構築する手法を示した．さらに BAF を用いて根拠となる事実から結論を推論する手法を示し，新たな結論を得るために必要な事実を求める推論方法について概観を述べた．今後は，全体を統括することでより広い推論方法を検討する．

謝辞

本研究は JSPS 科研費 17H06103 の助成を受けたものである．

参考文献

[1] Cayrol, C. and Lagasquie-Schiex, M.: Bipolarity in argumentation graphs: Towards a better understanding. *International Journal of Approximate Reasoning*, Volume 54, 876-899 (2013).

[2] Kawasaki, T., Moriguchi, S. and Takahashi, K.: Transformation from PROLEG to a bipolar argumentation framework. In *Proc. of SAFA2018*, 36-47 (2018).

[3] Kawasaki, T., Moriguchi, S. and Takahashi, K.: Reasoning by a Bipolar Argumentation Framework for PROLEG. In *Proc. of JURISIN2018*, 16-29(2018).

[4] Satoh, K. et al.: Translating the Japanese Presupposed Ultimate Fact Theory into Logic Programming. In Proc.of *JURIX 2009*, 162-171 (2009).

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

機械学習による裁判の裾野拡大の試み

Towards expanding the scope of trials via machine learning

川本 達郎 *1 Tatsuro Kawamoto　靖本 真未 *1 Masami Yasumoto　本村 陽一 *1 Yoichi Motomura

*1産業技術総合研究所 人工知能研究センター
National Institute of Advanced Industrial Science and Technology, Artificial Intelligence Center

We investigate the possibility of expanding a trial process to the cases that are not usually considered in a court. Our central question is whether it is possible to extract points of argument on each case from the public opinions. To this end, for a few hypothetical cases, we conduct a free-response survey using a web application. The data is collected as a graph (network) so that the public opinions can be summarized using a graph clustering algorithm. As a result, we found that extracted points of argument sensitively reflect details of the cases.

1. 背景

民事裁判は、非常に特殊な条件が揃った時にのみ実際に行われる。当事者双方が、金銭的・時間的コストを支払ってでも徹底的に争う姿勢を見せるような重大な事例であり、かつ裁判をする前から結果が明白ではないような典型的ではない事例の場合である。

しかし、裁判には至らない争いは日常的に存在している。そのような争いのうち幾らかは、裁判の対象に値しないのではなく、現実的なコストに見合わないため裁判事例として現れないだけだと考えられる。これは、権威が行う裁判という仕組みが、小さな訴訟案件まで対処しきれないことを反映している。裁判の裾野とでも呼ぶべき小さな訴訟に対処することは、できないのだろうか。

1.1 裁判の裾野へのアプローチ

本研究では、裁判員制度（あるいは論文の査読制度）のような枠組みから、事例についての論点抽出ができるかを考えたい。裁判員制度といっても、数人の人間が事例について意見を述べるだけであれば、複雑な処理は不要である一方、論点抽出のパフォーマンスは、選ばれたその数人の思考能力に強く依存する。ここではその代わりに、1つの事例に裁判員として数百人単位の人間を割り当て、集まった意見を集約して、妥当な論点抽出が可能かを検討する。

もちろん正規の裁判過程では、論点を挙げるだけではなく、事実認定や法律条文への “あてはめ”、それらを総合した最終的な判決が必要となるが、ここではそれらには踏み込まない。また、正規の裁判過程において、論点抽出とは、何が争われているかを正確に捉えながら行うものであるが、この点に関しても寛容に扱うこととする。

2. 離婚時の子供の親権問題

裁判事例として、事例数が比較的多いものとして離婚案件が挙げられる。そのなかでも、子供の親権・監護権をどう扱うかについては、賛否両論が分かれる、重要なテーマである。ただし、本研究の主目的は方法論であり、このテーマは一例であることに注意されたい。

連絡先: kawamoto.tatsuro@aist.go.jp

2.1 離婚事例データセットの構築

まず事前調査として、離婚に関係する事例に典型性はあるのか、親権が問題となる事例がどの程度存在するのかについて調べた。データセットとして、離婚に関する web 掲示板 [2] への書き込みを収集し、各事例についてのタグ付けをしたものを作成した。このデータセットは、今後さらにタグ付けを行ったり、言語処理を行うことで、有用な情報を引き出せる可能性もある。

もし離婚事例について自明な典型性が見られるのであれば、問題としては比較的簡単ということである。そうであれば、典型例について数人の専門家で議論すれば十分かもしれない。しかし結論としては、離婚事例について自明な傾向は見られなかった。

2.2 親権・監護権に関するアンケート調査

次に、特定の事例についての論点抽出問題を考える。前節のデータセットを基にして、子供の親権が関わる5つのショートストーリーを作成し、訴訟案件に見立てた。前節の離婚掲示板データのように、投稿者からの相談という形式になっており、例えば、1つは以下のような内容である。

ショートストーリー 1

30 歳専業主婦です。役に立つ資格は、運転免許証のみで他は何一つありません。最近は、夫（35 歳）が感情をコントロールできず子供にまで怒鳴りつけます。私にも子供（5 歳）にも威圧し、物を破壊するところを見せます。離婚したいのですが、私の実父母は他界しており身寄りがないのと、専業主婦の私が親権をとれるのか、今後の経済的な心配もあります。夫と一緒にいるほうが、裕福な生活を送れます。育児は90%私で、夫は休みの日に遊んでくれる程度です。離婚をして、私が子供を育てることが、子供にとっての幸せでしょうか。

上記のようなショートストーリーについての質問の他に、回答者の属性・人生観・家族観を把握するための質問を盛り込み、約1万人に向けてアンケート調査を実施し、6000 人ほどの回答を得た。ショートストーリーごとにアンケートを5パターン

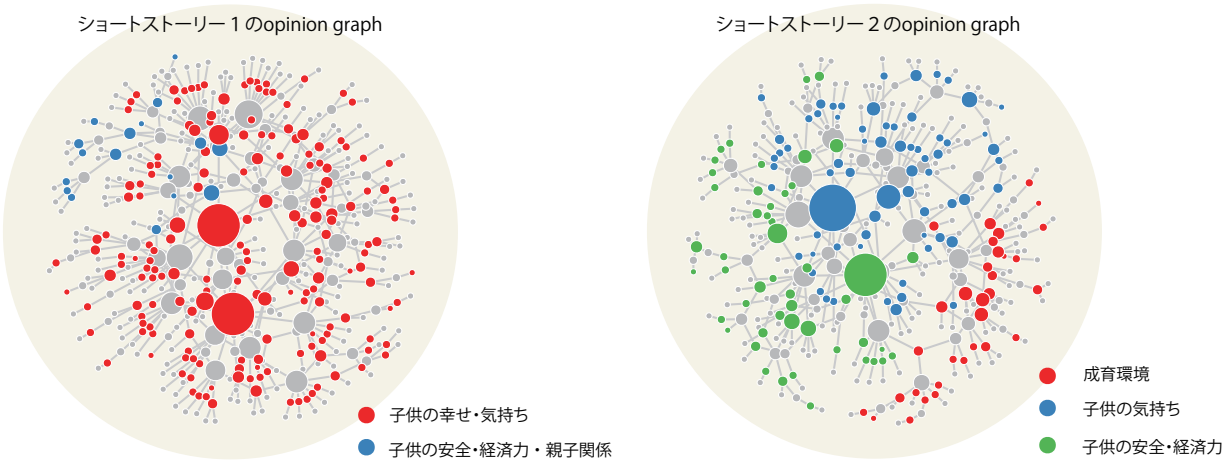


図 1: **投票クラスタリングによって得られた opinion graphs.** グラフの頂点は、各ショートストーリーに対する意見（色付き頂点）と回答者（灰色頂点）を表している。アンケート時に、回答者がある意見に賛成を表明した場合、回答者頂点と意見頂点は枝で結ばれる。頂点の大きさは、その頂点の次数（頂点に付いている枝の数）を反映している。頂点の色は、betweenness の方法 [1] によってクラスタリングをした結果を表している。各グラフの右下には、各色の頂点の代表的な意見をリストしている。

に分けているため、１つのパターンの回答は 1000 人弱程度である *1。

2.3 投票クラスタリング

ショートストーリーを読んだ上で、どのように論点を見出すかを問うアンケートは、自由記述式である必要がある。そのために我々は、「投票クラスタリング」[3] という、大規模な自由記述式アンケートを実現するシステムを使用した。投票クラスタリングでは、回答文章や回答者をそれぞれ１つの頂点、回答文章の類似性を枝で表現したグラフ（opinion graph）を生成し、グラフクラスタリングアルゴリズムによって回答集合の集計を行う。この方法は、一切自然言語処理を必要としないため、いかなる言語・内容のアンケートでも対応可能である。また、クラスタリングアルゴリズムが処理できる限り、どんな大規模なアンケートでも集計可能である。

3. 結果

図１は、専業主婦の妻から投稿であるショートストーリー１に対するアンケート結果と、専業主夫の夫からの投稿に反転させた内容であるショートストーリー２に対するアンケート結果である。子供の気持ちという論点は、どちらにおいても支配的である。ショートストーリー１では、ほとんどその意見で埋め尽くされ、わずかに他の論点が出現しているだけである。一方ショートストーリー２では、さらに多様な論点が抽出されている。

回答者の属性に何らかの偏りがあるかもしれないため、この結果からだけでは、何によって論点の多様性が生まれているかは判断できない。しかし、本調査では、様々な属性質問を収集しているため、この問いにさらに深く踏み込むことは可能である。

*1 ただし、集計に実質的な寄与を与える回答者数は数百人程度に落ちる。

4. まとめ・展望

今回の調査によって、方法論として、大衆のコンセンサスとしての論点抽出の可能性が見出された。今回のアンケートでは、回答者の知識レベル等についてフィルタリングをかけていないが、有識者を対象にしたり、関心の高い人を対象にすることで、より高い分解能の論点抽出が可能になると期待される。この方法が、（法的拘束力はないにせよ）正規の裁判では扱えないような小さな訴訟の対処に用いられたり、正規の裁判にも活用されることが、大きな展望である。

今回のデータセット構築および解析は、論点抽出に使えるかという問いの他にも、親権・監護権についての法律について議論するための様々な情報を与えるものであり、色々な研究用途の可能性も秘めている。さらに、親権問題に限らず多様な問題について要素抽出を積み重ねていくことは、裁判過程を支援する人工知能技術を現実社会で機能させるために、ある種の水準・フレームを与える役割として重要だと考えられる。

謝辞

本研究は JSPS 科研費 17H06103 の助成を受けたものである。研究構想・準備段階においては、高橋周子氏（産業技術総合研究所テクニカルスタッフ [当時]）、石綿はる美氏（東北大学）から、多くの提案をいただいた。また、親権アンケートへ回答していただいた多くの参加者の方々に感謝したい。

参考文献

[1] Girvan, M. and Newman, M. E. J., Proc. Natl. Acad. Sci. USA 99, 7821–7826 (2002).

[2] https://mikle.jp/listp/49/

[3] https://www.bonsci.org/

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

刑事訴訟版の PROLEG の開発

Implementation of Criminal Procedure using PROLEG Technology

佐藤　健 *1　　西貝　吉晃 *2
Ken Satoh　　Yoshiaki Nishigai

*1国立情報学研究所/総研大　　*2日本大学
National Institute of Informatics/Sokendai　　Nihon University

In this paper, we show a possibility to implement Japanese criminal procedure including Japanese penal code in PROLEG (PROlog-based LEGal reasoning support system). PROLEG has been orginially used to formalize only civil procedure but it is also generally applicable to laws which include general rules and exceptions. Japanese penal code consists of general rules and exceptions so it is possible to formalize it in PROLEG. This paper discusses a concern about how we represent conditions of concepts in criminal law domain using PROLEG.

1.　はじめに

本論文では、論理型推論支援システム PROLEG(PROlog-based LEGal reasoning support system) [西貝 11, 佐藤 11] を用いて刑事訴訟を表現することを検討する。PROLEG は、民法の要件事実論 (民事訴訟) を実装することを目的として開発された。もっとも、民事訴訟にとどまらず、[Satoh12] において示したとおり、PROLEG は、原則・例外を用いた推論に基づく法律システムにも応用可能だと期待される。したがって上記特徴を併有する刑事訴訟も表現可能だと思われるが、刑事訴訟には民事訴訟とは異なる理解がなされている箇所があるので、本稿では、それに合わせた実装上の工夫を述べる。

2.　PROLEG

PROLEG は、原則・例外からなるルールで表現したルールベースと、当該事件認定された事実を記載したファクトベース等からなる。PROLEG の実行は、当事者の立証活動の結果を証明責任の観点から考慮した形で、民法や刑法を適用する際の推論を行う。

2.1　PROLEG のルールベース

PROLEG のルールベースでは、要件事実論を以下の 2 つの式の形式で表す。

結論 <= 要件 1 , 要件 2 ,..., 要件 n.
例外事由 (結論, 例外).

結論、要件、例外はそれぞれ、述語名 (引数 1, 引数 2,..., 引数 m) の形で表現される。上の第 1 の式（「原則ルール」と呼ぶ）の読み方は、「要件 1 かつ要件 2 かつ... かつ要件 n ならば結論が成立する」であり、その直観的意味は、要件 1 ,... 要件 n がすべて「存在」する場合には結論も「原則としては」、「存在」するということを表す 。また、第 2 の式（「例外ルール」と呼ぶ）の読み方は、「例外は、結論に対する例外事由である」であり、その直観的意味は、例外が「存在」する場合には、いくら原則ルールのすべての要件が「存在」していたとしても、「例外的に」結論が「不存在」となることを表している。

なお、訴訟法の理論としての証明責任の解釈として、法規不適用説と証明責任規範説の相違はあるが、PROLEG においては、デフォルト値の説明の仕方に過ぎない、と考える [西貝 11]。

連絡先: 佐藤　健, 国立情報学研究所, ksatoh@nii.ac.jp

2.2　PROLEG のファクトベース

PROLEG のファクトベースには、実際の事件の事実が記載される。PROLEG の原則ルールの要件のうち、それを結論とした原則ルールがない場合には、それは PROLEG のファクトベースに書くべき事実となる。たとえば、ある事実 P について、PROLEG のルールベースには、それを結論とする原則ルールがないため、もしこの法律要件に対応する事実が存在するならば、事実としてファクトベースに記載する必要がある。PROLEG の刑事訴訟版では、「事実 (P).」のような形式で、そのような事実を記載する。

2.3　意味論

PROLEG の意味論は、論理プログラミングにおける解集合 (Answer Set) と同じように定義される。PROLEG のルールベースにおける原則ルールの集合を $\mathcal{R}$ とし、例外ルールの集合を $\mathcal{E}$ としたとき、すべての変数を言語における基礎項の集合 $\mathcal{H}$ の要素で置き換えた集合をそれぞれ $\Pi_{\mathcal{R}}, \Pi_{\mathcal{E}}$ とする。そして、基礎項 $\mathcal{H}$ の部分集合 M に対して $\Pi_{\mathcal{E}}$ によってフィルターされた $\Pi_{\mathcal{R}}$ の適用可能ルール $\Pi_{\mathcal{R}}^{M}$ を以下のように定義する。

$$\{R \in \Pi_{\mathcal{R}} | \neg\exists \text{ 例外事由 } (H, E) \in \Pi_{\mathcal{E}} \text{ s.t. } H = head(R) \text{ and } E \in M\}$$

さらに、R の条件部の正リテラルの部分を $pos(R)$ と表し、原則ルールの集合を $\mathcal{R}$ としたときの $\{head(R) <= pos(R) | R \in \mathcal{R}\}$ を $pos(\mathcal{R})$ で表す。
このとき、M が PROLEG プログラム $\langle \mathcal{R}, \mathcal{E} \rangle$ のモデルであるとは、$M = min(pos(\Pi_{\mathcal{R}}^{M}))$ であることをいう *1。

3.　刑事訴訟の構成および及び実装に当たっての留意点

刑事訴訟においては、主にルールベースに関係する刑法 (刑事実体法) と、ルールベースとファクトベースを接続する刑事訴訟法が議論の対象となり、これらの法規範をプログラムする必要がある。

刑法においては、犯罪とは構成要件に該当し、違法で有責な行為だとされる（定説）。つまり、犯罪の成立要件は、構成要件該当性、違法性及び有責性であり、3 段階に分けて要件該当

*1　ホーン節からなるプログラムを P としたとき、$min(P)$ は P の最小モデルを表す

性が検討される。ここで、構成要件に該当すれば、違法（さらには責任）が推定され、原則的に犯罪が成立するとされる。したがって、例えば違法性の段階における考察の内容、積極的な違法の基礎付けではなく、既に構成要件に該当していることによってありそうにみえる違法が「なくなるのではないか」、すなわち違法性を失わせる「違法性阻却事由があるのではないか」、ということが検討される。責任阻却事由についても同様に考えることができる。したがって、違法性阻却事由および責任阻却事由（以下「阻却事由」と呼ぶ）を、その有罪の結論の例外と考えることができる。

しかし、ここで民事訴訟用の PROLEG の推論方式との不一致が問題となる。すなわち、刑事訴訟においては「疑わしきは被告人の利益に」の原則が妥当し、構成要件に該当する事実の「存在」だけでなく、阻却事由に該当する事実の「不存在」の立証責任を検察官が負うことになる。結論的に例外、つまり阻却事由の立証責任を負う者が民事訴訟と逆転しているのである。それゆえ、阻却事由の存在が疑われる場合には、検察官が不存在の証明をしない限り、その阻却事由がデフォルトとして存在しなければならないのだが、PROLEG で直接に、有罪の例外として阻却事由を記述してしまうと、阻却事由が存否不明の場合には、その阻却事由は存在しないことになってしまう。

この問題を解決するために、刑事訴訟法における挙証責任の議論 [川出 02, 堀江 13] に基づき、以下のように原則、例外を分配することにした。すなわち、被疑者は、阻却事由については、証明責任はないが、争点形成責任はあるとして、裁判上で阻却事由の争点形成がされた場合には、原則として、阻却事由が存在すると評価されるようにし、その例外を、それらの事由の不存在が証明されることとした。そして、もし阻却事由の要件が複数ある場合には、それらの要件のどれか一つでも存在しないことが証明された場合には、その阻却事由が存在しないこととした。

たとえば、違法性阻却事由である正当防衛（刑法 36 条 1 項 *2）の場合には、正当防衛についての口頭弁論中に争点形成がされた場合、それを PROLEG ファクトとして登録し、争点形成の存在があれば、原則として正当防衛が成立するとし、それに対して、正当防衛の要件である、「不正の侵害」、「侵害の急迫性」、「防衛の意思」、「防衛行為の相当性」の事実のうちの一部でも不存在であると証明された場合に、正当防衛が不成立となる、という構成とした。

4. 刑法上の要件の実装方法

民法の要件事実論においては、教科書において法律要件がほぼ統一的に定義されている。したがって、それらの法律要件をルールベースに、ファクトベースに法律要件に該当する具体的事実（主要事実，要件事実）を書けばよい。

一方、刑法においては、法律要件に対応するものは、構成要件，阻却事由の要件等である。ここで、民事でもあり得ることではあるが、構成要件等の中には、条文にない要件もあるので適宜、刑法解釈を教科書等で確認しながら実装する必要がある（例：窃盗罪 *3 における「不法領得の意思」*4）。

さらに、正当防衛の成立を否定するためには、正当防衛の要件（急迫不正の侵害、防衛の意思、防衛行為の相当性）のうち少なくとも一つを否定する必要がある。しかし、判例は、さらにそれらの要件の解釈を具体化し、急迫性の不存在の要件として、「積極的加害意思」*5 を採用し、防衛の意思の不存在の要件として「積極的加害行為」*6 という要件を導出しているので、それらについても刑法の教科書等で補う必要がある。

また、解釈の具体化に伴って生じる概念をどこまで PROLEG のルールベースに実装するかは一つの問題となるが、PROLEG の教育応用も考えている我々としては、教科書において抽象的な概念で説明されている場合には、PROLEG のルールとして実装することにした。

5. PROLEG ルール及びテストケースの実行例

この方針にしたがって、PROLEG で刑法を記述した一部を以下に示す。ここでは殺人罪（刑法 199 条）*7 、および、それに対する正当防衛の例外、および、正当防衛の要件の急迫性の不存在の要件である積極的加害意思のルールおよび防衛の意思の不存在の要件である積極的加害行為のルールをルールベースで表現したものを示す。

```
有罪 (_罪,_行為者,_客体,_行為)<=
    構成要件該当性 (_罪,_行為者,_客体,_行為).

構成要件該当性 (殺人罪,_行為者,_客体,_行為)<=
    人 (_客体),
    殺人行為 (_行為者,_客体,_行為),
    死亡 (_客体),
    因果関係 (_行為, 死亡 (_客体)),
    構成要件的故意 (
        _行為者, 殺人行為 (_行為者,_客体,_行為)).

例外事由 (有罪 (_罪,_行為者,_客体,_行為),
    正当防衛 (_罪,_行為者,_客体,_行為,_侵害内容)).

正当防衛 (_罪,_行為者,_客体,_行為,_侵害内容)<=
    争点形成 (
        正当防衛 (_罪,_行為者,_客体,_行為,_侵害内容)).

例外事由 (正当防衛 (_罪,_行為者,_客体,_行為,_侵害内容),
    不正の侵害の不存在 (_行為者,_客体,_侵害内容)).
例外事由 (正当防衛 (_罪,_行為者,_客体,_行為,_侵害内容),
    急迫性の不存在 (_行為者,_客体,_侵害内容)).
例外事由 (正当防衛 (_罪,_行為者,_客体,_行為,_侵害内容),
    防衛の意思の不存在 (
        _行為者,_客体,_行為,_侵害内容,_防衛の意思内容)).
例外事由 (正当防衛 (_罪,_行為者,_客体,_行為,_侵害内容),
    防衛行為の不相当 (
        _行為者,_客体,_行為,_侵害内容,
        _不相当評価根拠事実)).

不正の侵害の不存在 (_行為者,_客体,_侵害内容)<=
    不正の侵害の不存在の評価根拠事実 (
```

*2 急迫不正の侵害に対して、自己又は他人の権利を防衛するため、やむを得ずにした行為は、罰しない。

*3 （窃盗）第 235 条 他人の財物を窃取した者は、窃盗の罪とし、10 年以下の懲役又は 50 万円以下の罰金に処する。

*4 権利者を排除して他人の物を自己の所有物としてその経済的用法に従い利用、処分する意思（大判大正 4.5.21）

*5 相手の侵害の機会を利用して石鋸的に相手に加害行為をする意思で侵害に望んだたときには、侵害の急迫性の要件は充たされない（最決昭和 52.7.21）

*6 防衛に名を借りて侵害者に対し積極的に攻撃を加える行為は防衛の意思を欠く (最決昭和 50.11.28)

*7 人を殺した者は、死刑又は無期若しくは 5 年以上の懲役に処する。

```
        _行為者,_客体,_侵害内容).

急迫性の不存在(_行為者,_客体,_侵害内容)<=
    急迫性の不存在の評価根拠事実(
        _行為者,_客体,_侵害内容).
急迫性の不存在(_行為者,_客体,_侵害内容)<=
    積極的加害意思(_行為者,_客体,_侵害内容).

積極的加害意思(_行為者,_客体,_侵害内容)<=
    侵害の予期(_行為者,_客体,_侵害内容),
    加害意思(_行為者,_機会を利用して加害する意思).

防衛の意思の不存在(
    _行為者,_客体,_行為,_侵害内容,_防衛の意思内容)<=
    防衛の意思の不存在の評価根拠事実(
        _行為者,_客体,_行為,_侵害内容,_防衛の意思内容).
防衛の意思の不存在(
    _行為者,_客体,_行為,_侵害内容,_防衛の意思内容)<=
    積極的加害行為(
        _行為者,_客体,_行為,_侵害内容,_防衛の意思内容).

防衛行為の不相当(
    _行為者,_客体,_行為,_侵害内容,
    _不相当評価根拠事実)<=
  防衛行為の不相当の評価根拠事実(
    _行為者,_客体,_行為,_侵害内容,
    _不相当評価根拠事実).
```

以下のファクトベースでは、甲が乙から包丁で甲を刺そうとしたところ、それに対抗して乙の胸をナイフで刺したという事件において、甲が乙から襲われることを予見し、襲われる機会を利用して乙を殺してしまおうという事実があったことが示されている。

```
事実(人(乙)).
事実(殺人行為(甲,乙,乙の胸をナイフで刺す)).
事実(死亡(乙)).
事実(因果関係(乙の胸をナイフで刺す,死亡(乙))).
事実(構成要件的故意(甲,殺人行為(甲,乙,乙の胸をナイフで刺す))).

事実(争点形成(正当防衛(殺人罪,甲,乙,乙の胸をナイフで刺す,包丁で甲を刺そうとした))).

事実(侵害の予期(甲,乙,包丁で甲を刺そうとした)).
事実(加害意思(甲,もっぱら乙の攻撃利用して乙を殺す意思)).
```

図 1 に、PROLEG の実行結果を示す（PROLEG ブロック図）という。左上に結論があり、その結論の箱の一番下が o(成立)となっている。そして実線矢印により結ばれた右の複数の箱に、その結論を満たす要件が示されており、それらの要件の存否が o または x で表される。すべての要件が o であっても、点線矢印により結ばれた例外事由（複数の場合もある）のどれか一つでも成立すれば（o になれば）、結論が否定される。すべてが不成立 (x) であれば、結論が維持される。本件では正当防衛について争点形成がされたため、その例外が検討されている。しかし、その要件の一つである、「急迫性」を不存在とする「積極的加害意思」の事情が存在しているため、正当防衛の急迫性の不存在が証明され、殺人罪が成立することになる。

6. おわりに

本論文では、刑法の PROLEG 化の大まかな表現の方針を示した。しかし、刑法においては、「過失」、「事実の錯誤」（行為者の認識が客観的行為と異なっている場合）、「共犯」など様々な概念があるため、それらを考慮して、この表現方針が矛盾しないかどうか検討する必要がある。

謝辞: 本研究は JSPS 科研費 17H06103 の助成を受けたものである。

参考文献

[西貝 11] 西貝吉晃, 浅井健人, 久保田理広, 古川昂宗, 佐藤健, 白川佳, 高野千明, 中村恵, PROLEG: 論理プログラミング言語 Prolog を利用した要件事実論のプログラミング, 情報ネットワーク・ローレビュー, Vol. 10 pp.54-89 (2011).

[佐藤 11] 佐藤　健, 浅井健人, 古川昂宗, 久保田理広, 中村恵, 西貝吉晃, 白川　佳, 高野千明, PROLEG:論理プログラミングをベースとした民事訴訟における要件事実論の実装, 人工知能学会第 92 回知識ベース研究会資料, pp.1 - 8(2011).

[Satoh12] Satoh, K., Kogawa, T., Okada, N., Omori, K., Omura, S., and Tsuchiya, K., "On Generality of PROLEG Knowledge Representation", Proc. of the Proceedings of the 6th International Workshop on Jurisinformatics (JURISIN 2012), pp. 115 – 128, Miyazaki, Japan (2012).

[川出 02] 川出敏裕, 挙証責任と推定, ジュリスト増刊, 刑事訴訟法の争点 [第 3 版], pp.158-161(2002).

[堀江 13] 堀江慎司, 挙証責任と推定, ジュリスト増刊, 新・法律学の争点シリーズ 6, 刑事訴訟法の争点, pp.148-151(2013).

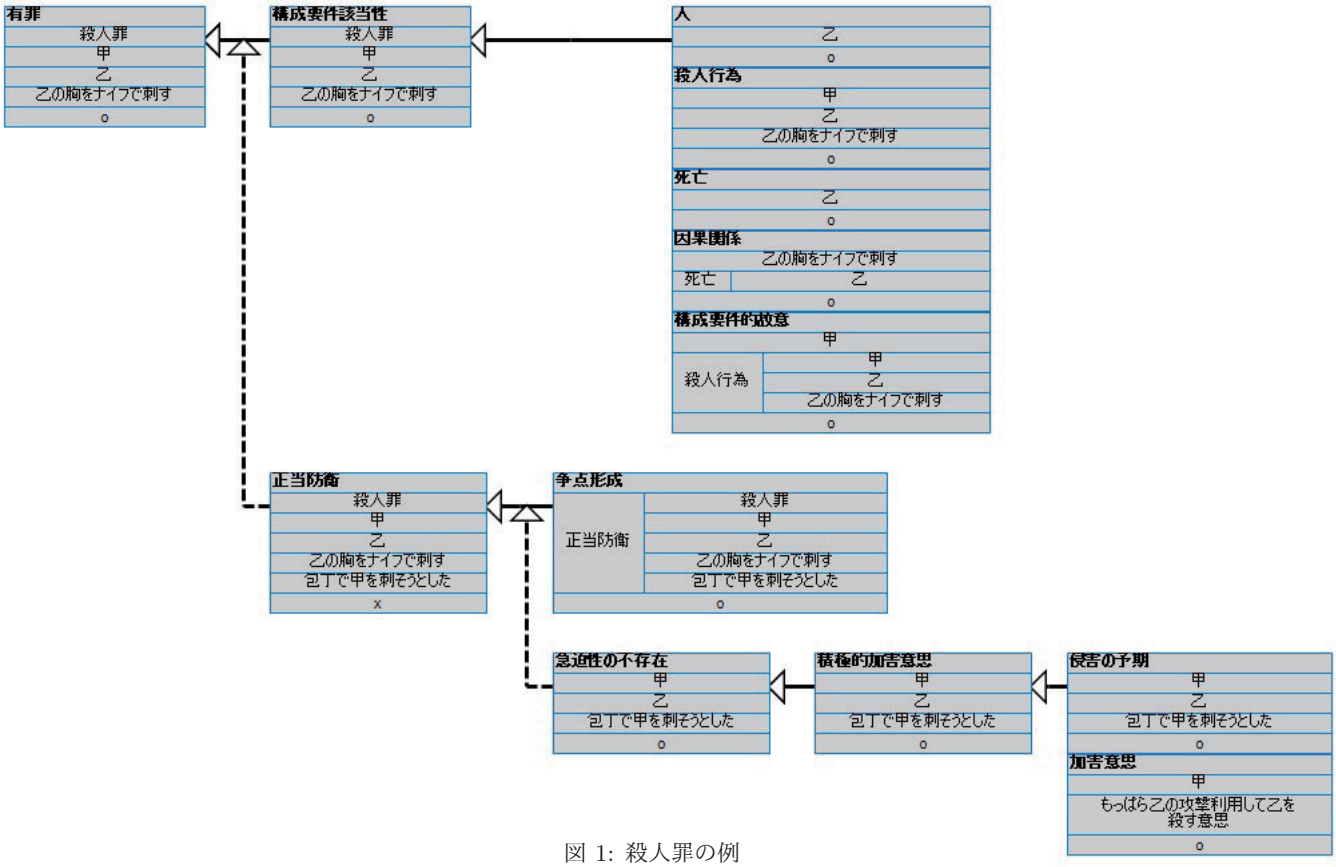

有罪
殺人罪
甲
乙
乙の胸をナイフで刺す
o
構成要件該当性
殺人罪
甲
乙
乙の胸をナイフで刺す
o
人
乙
o
殺人行為
甲
乙
乙の胸をナイフで刺す
o
死亡
乙
o
因果関係
乙の胸をナイフで刺す
死亡
乙
o
構成要件的故意
甲
殺人行為
甲
乙
乙の胸をナイフで刺す
o
正当防衛
殺人罪
甲
乙
乙の胸をナイフで刺す
包丁で甲を刺そうとした
x
争点形成
正当防衛
殺人罪
甲
乙
乙の胸をナイフで刺す
包丁で甲を刺そうとした
o
急迫性の不存在
甲
乙
包丁で甲を刺そうとした
o
積極的加害意思
甲
乙
包丁で甲を刺そうとした
o
侵害の予期
甲
乙
包丁で甲を刺そうとした
o
加害意思
甲
もっぱら乙の攻撃利用して乙を殺す意思
o

図 1: 殺人罪の例

Organized Session | Organized Session | [OS] OS-11

[4F2-OS-11a] 社会的信号処理と AI(1)

岡田 将吾（北陸先端科学技術大学院大学）、石井 亮（NTT）

Fri. Jun 7, 2019 12:00 PM - 1:40 PM Room F (302B Medium meeting room)

[4F2-OS-11a-01] The Interaction between Dancer and DJ in Performance Scene of Breakdance

〇Daichi Shimizu[1], Takeshi Okada[1] （1. The University of Tokyo）

12:00 PM - 12:20 PM

[4F2-OS-11a-02] Estimating Verbal・ Nonverbal Skills in Business Presentation

〇Yagi Yutaro[1], Okada Shogo[1], Shiobara Shota[2], Sugimura Sota[2] （1. Japan Advanced Institute of Science and Technology, 2. SoftBank Corp.）

12:20 PM - 12:40 PM

[4F2-OS-11a-03] Measuring affective sharing between two people from simultaneously recorded EEG signals

〇Taiki Kinoshita[1], Hiroki Tanaka[1], Koichiro Yoshino[1], Satoshi Nakamura[1] （1. Nara Institute of Science and Technology）

12:40 PM - 1:00 PM

[4F2-OS-11a-04] Estimating performance of group interaction based on product dimension

〇Go Miura[1], Shogo Okada[1] （1. Japan Advanced Institute of Science and Technology）

1:00 PM - 1:20 PM

[4F2-OS-11a-05] Use of Dialogue History for User's Interest Level Estimation Based on Multimodal Information

〇Haruto Nishimoto[1], Ryu Takeda[1], Kazunori Komatani[1] （1. The Institute of Scientific and Industrial Research (ISIR), Osaka University）

1:20 PM - 1:40 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

ダンスパフォーマンスにおける演者間インタラクション：Dancer-DJ 間の相互作用に関する検討

The Interaction between Dancer and DJ in Performance Scene of Breakdance

清水 大地[*1]
Daichi Shimizu

岡田 猛[*1]
Takeshi Okada

[*1] 東京大学大学院教育学研究科
Graduate School of Education, The University of Tokyo

Abstract: In performing arts like dance and music performance, performers interact with each other actively and present their fascinating performances to audiences. This study aims to investigate this complicated interaction among performers by applying a prominent theory of communication, synchrony. We conducted a case study to examine the interaction between the dancer and the DJ in the performance scene of breakdance. The results suggested that the dancer and the DJ coordinated their performances in a complicated fashion, like coordinating their movements of multiple body parts and changing their coordination patterns as time proceeds. We suggest the importance to develop the framework and analysis methods of synchrony to fully capture these complicated features of the interaction among performers.

1. Introduction

1.1 上演芸術における演者間の関わり合い

ダンスや音楽演奏といった上演芸術では、その場に存在する他者と豊かな関わり合いが営まれながらパフォーマンスが披露されていく。本研究は、上演芸術の中でも特にダンス表現を取り上げ、そこで生じる演者間（特にダンサーと DJ 間）のインタラクションに関する探索的な検討を行ったものである。

実際に上演芸術では、時空間を共有する他者との活発な関わり合いが営まれること、その関わり合いの様相をパフォーマンスとして披露する点に 1 つの魅力が存在することは、インタビューや逸話において広く示唆されてきた。例えば、演者同士が互いの表現に強く依拠した表現を展開すること、その関わり合いから一人では生成出来ない魅力的な表現が生成される場合があることが指摘されている[Bailey 1980; 清水 2013]。

また生物学や文化人類学の観点からも、上演芸術における共演者との関わり合いの重要性はかねてより指摘されてきた。例えば、複数名で共に音楽表現や舞踊的な儀式に取り組むことで社会的な絆（social bond）が生成・強化されること、そしてその生成・強化が共同体の維持・発展にとって有用であったため、様々な共同体において音楽やダンス表現が普遍的に観察され、伝統的に受け継がれてきたことが示唆されている[Merker 2015]。実際に、共に音楽演奏を行うことやパフォーマンスの場を共有することによって参加者間の親密度が上昇することは実証的な検討からも示唆されつつある[Weinstein 2016]。本研究では、この上演芸術における共演者とのインタラクションに着目し、そこで見られる複雑な関わり合いを捉える観点の提案と、ケーススタディによる部分的な検証を行った。

1.2 演者間インタラクションを捉える観点

上記した演者間のインタラクションを科学的に捉えるために本研究が着目したのが、同期現象（syncnronization）に関する知見である。同期現象とは、時空間を共有する複数のエージェント間に、類似した振る舞いが時間的に近接して繰り返し生じることを示す現象である。実際に同期現象は、時計の振り子やメトロノームの振動といった非生物、ホタルの明滅やカエルの鳴き声といった生物に加え[Strogatz 2003]、拍手や歩行・会話時の姿勢といったヒトの多様な振る舞いにおいても観察されることが示唆されてきた[Néda 2000]。そしてそこでの振る舞いのパターンの仔細や現象が生じる背景としての重要な情報の種類（視覚・聴覚等の知覚情報）などがこれまで検討されつつある。

以上の、複数のエージェントが同時に活動を行う際に頻繁に観察される同期現象やその蓄積された理論は、上演芸術における演者間のインタラクションを検討する上でも適用可能であると考えられる。上記した通り、複数名がともに表現を披露するパフォーマンス場面では、互いに多様な媒体を通した関わり合いを行うことによって、演者同士が共通の特徴を有するパフォーマンスを生成・披露していくことが推測されるためである。実際に、近年では以上の理論を適用・拡張することで音楽演奏やダンス表現における関わり合いの様相を捉えようとする試みが営まれつつある[Walton 2018]。一方で、上演芸術におけるインタラクションは、特定の媒体において同一の振る舞いを一致したタイミングで行うことに留まらない。ダンスを例とすれば、多様な媒体（腕部や脚部といった各身体部位など）を用いて、他者と多様なレベルで対応した振る舞いを、多様なタイミングで披露する様相が見られると考えられるのである[清水 2013; 清水 2016]。以上を踏まえ、本研究では、実際のインタラクションのデータに同期理論の解析手法を適用することで、その適用可能性と拡張すべき方向性を検討することを目指した。

1.3 ブレイクダンスのパフォーマンス場面

本研究では、上演芸術の 1 領域としてダンス表現、特にこれまで著者が長期に渡り検討を行ってきたブレイクダンスのパフォーマンス場面を対象とし、上記の検討を行うこととした。ブレイクダンスでは、例えばダンサーや DJ といった多様な振る舞いを行う複数名の演者が活発に関わり合いながらパフォーマンスが展開されることが、映像、インタビューや過去の研究から確認されており、インタラクションの検討を行う上で適切な対象と考えられたためである[清水 2013; 清水 2016]。なおブレイクダンスは、

連絡先：清水大地，東京大学大学院教育学研究科，東京都文京区本郷 7−3−1，tothefuture0415@yahoo.co.jp

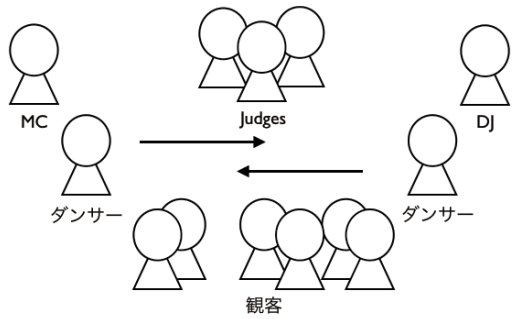


図1. バトル場面の1例

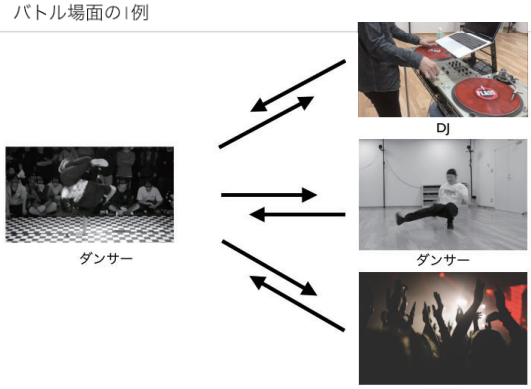


図2. バトル場面において見られる関わり合いによるシステム的な様相。実際は、DJや他のDancer、Audienceの間においてもインタラクションが生じており、それらも含めた非常に複雑な関わり合いが構築されている。

ギャングや若者の抗争の代替として利用され、発展してきた歴史を有する。その経緯から、専らバトルという複数名のダンサーが表現を披露し合って勝敗を決定する形式によってパフォーマンスが披露される[OHJI 2001]。そのため、本研究でも基本的にこのバトル場面の形式を部分的に模した状況を設定して演者間のインタラクションを検討した。

ブレイクダンスにおけるバトル場面の 1 例を図1に示す。バトル場面では、2 名のダンサー、もしくは 2 チームのダンサー達が互いに向かい合った状態で交互にパフォーマンスを披露する。その際には、相手の表現を部分的に取り入れて発展させる、相手とジェスチャーを行い合う・視線を交わし合うといった活発な関わり合いが営まれる様子が観察されている。また、流す音楽に関しては、DJ が即興で選択して編集を加えながら流しており、ダンサーはその音楽にも即時的に反応しながらパフォーマンスを披露する必要が生じる。加えて、MC やジャッジ、観客はダンサーのパフォーマンスに対して歓声や拍手といった形で活発に反応を返しており、ダンサーはその様子にも敏感に反応していることがフィールドワーク映像において確認されている。このようにバトル場面においては、場を共有した複数名のエージェントが互いに活発に関わり合いながらパフォーマンスが披露される様相が見られており、各参与者の振る舞いが互いの関わり合いによってある程度規定される、システム的な振る舞いが生じる場面と捉えることが可能であろう(図 2)。

著者らは既に、ダンサー間に生じるインタラクションを捉える試みをこれまで行ってきた[清水 2013; 清水 2016]。本研究では以上の試みを拡張し、同様の場面においてダンサーと DJ との間に生じるインタラクションに関して探索的な検討を行う。

2. ケーススタディ

2.1 Outline

ここでは、実際に熟達したダンサーと DJ にパフォーマンスを披露してもらい、両者の間に生じるインタラクションの様相につ

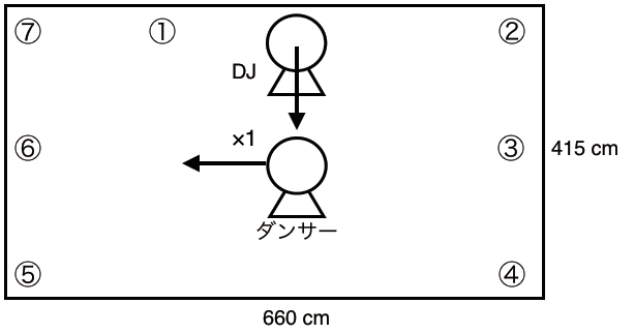


図3. 実験の実施図

いて同期理論の解析手法を適用することで探索的に検討した。前述した通り、バトル場面においてダンサーは、DJ がその場で選択し編集を加える音楽に即時的に合わせながらパフォーマンスを披露する[清水 2013]。一方で、DJ も同様にダンサーのパフォーマンスに注意を向け、その振る舞いに敏感に反応する様子がバトル場面において見られている。その際に生じる複雑なインタラクションについて同期理論を用いて定量的に検討することを目指した。その際前述した上演芸術の特徴を踏まえ、特に多様な媒体(身体部位)において対応した振る舞いが生じるのか、その対応関係が時間経過によって変化するのか、という 2 点を検討した。

2.2 Method

協力者は、熟達したダンサーA, B の 2 名 (A : 30 歳、経験年数 14 年、B : 30 歳、経験年数 13 年、共に国内の大会における優勝・準優勝等の経験を有する) と熟達した DJ1 名(27 歳、大会等における DJ 経験を豊富に有する)である。

実験では、ダンサー1 名ずつが DJ とペアとなり、計 3 回のパフォーマンス(ダンサー:ダンスの披露、DJ:曲の即興的な選択・編集)を行った。本研究では、その中でも各ペアで同じ音楽を用いた 1 回目の測定に関して解析を行った。この場面において、ダンサーは DJ の流す音楽に合わせながらダンス表現を 1 回披露しており、その際の両者の振る舞いに関する測定・解析を行った(図 3)。なお、ダンサーは DJ の流す音楽について事前に知らされておらず、DJ は実験者の設定した 2 つの条件(現実のバトル場面で頻繁に流される著名な曲の中から 1 曲選択して流す、両ダンサーで同一の曲を流す)に基づいて流す曲をその場で選択した。また、パフォーマンスを行う時間にも厳しい制限を設けず、可能な限り現実場面に近づけた生態学的妥当性の高い状況で実験を実施した[1]。

そして上記の状況におけるダンサーと DJ 各自の振る舞いに関してモーションキャプチャーシステム(OQUS 300, Optical motion capture system, QUALISYS co.)を用いて測定した。まず、パフォーマンスを行う上で邪魔にならない箇所を中心に全身にマーカーを接着した。今回は特に、大局的なリズム運動を反映すると考えられるダンサーと DJ の頸部の上下運動(z 軸の座標情報)とパフォーマンスの細部を反映すると考えられる左手の運動(ダンサーは z 軸、DJ は y 軸の座標情報)について、マーカーの位置情報を利用して測定・算出した。前述した通り、上演芸術において演者らは、他の演者と多様なレベルで関係した振る舞いを多様な媒体を用いて営んでいくことが予想される。そのため、ダンサーと DJ との間で共有可能な情報であり、互いのパフォーマンスに大きな影響を及ぼすと考えられる互いのリズム運動と互いの表現の細部に着目して検討を行うこととした。特に

[1] 現実場面でもパフォーマンスの時間は明確には指定されない場合が多い。

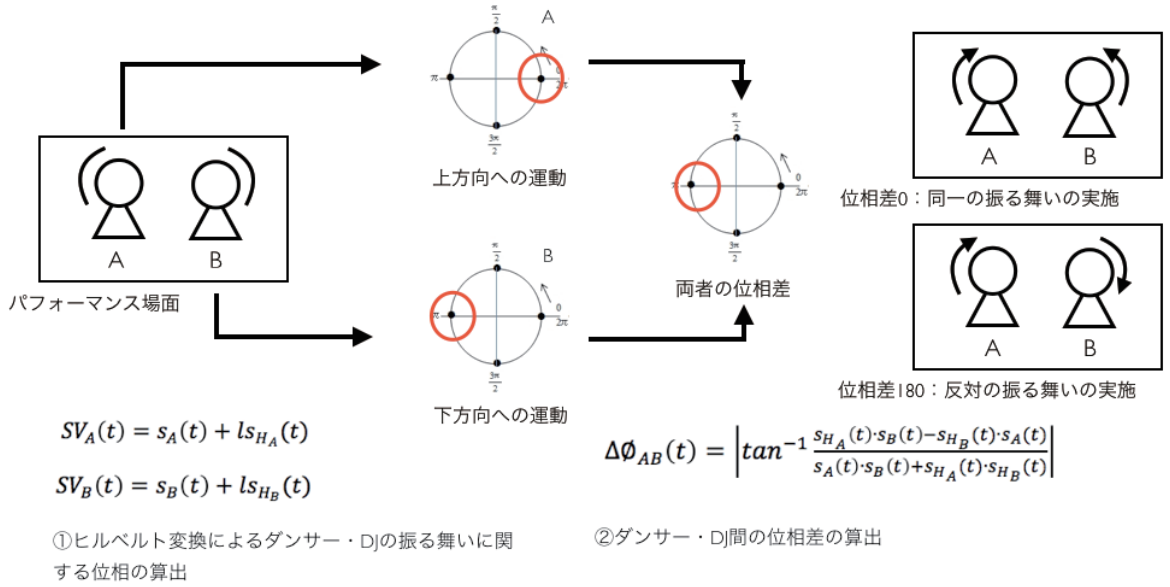


図4. ダンサー・DJ間の振る舞いの対応関係に関する解析手法。Okumura et al. (2012)に基づいて作成した。

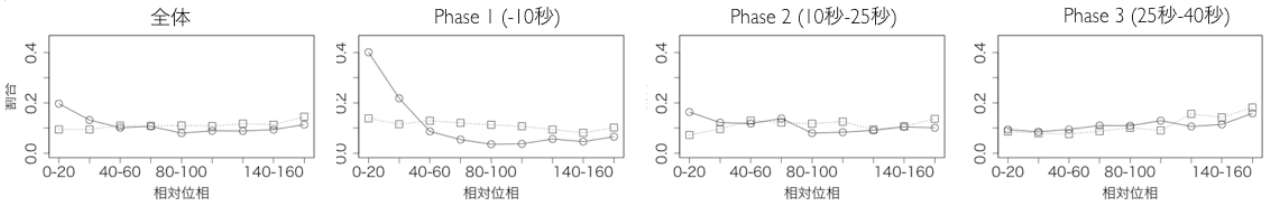


図5. リズム運動に関する対応関係の結果. 実線はreal pairの結果を、点線はpseudo pairの結果を示す.

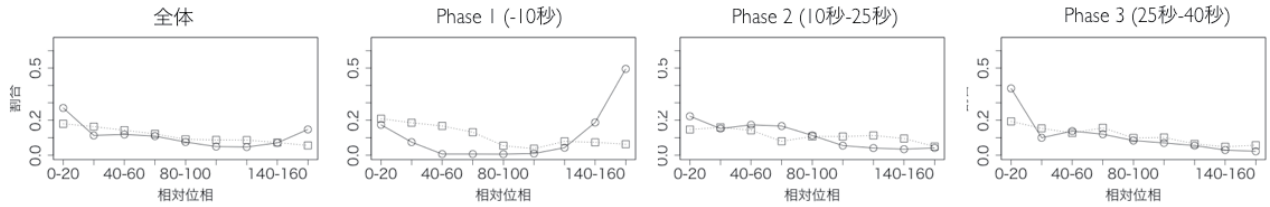


図6. 表現の詳細に関する対応関係の結果. 実線はreal pairの結果を、点線はpseudo pairの結果を示す.

DJ の手の動きは、DJ が音楽を操作する際の振る舞いを直接的に反映した指標と捉えることが出来るだろう。

2.3 Analysis

解析では、上記のマーカーの位置情報を利用して、1:ダンサーと DJ 両者のリズム運動の対応関係と 2:ダンサーと DJ 両者の表現の細部(左手の振る舞い)の対応関係について算出した。まず、1:ダンサーと DJ 両者のリズム運動の対応関係に関しては、頸部の上下運動のデータに、平滑化(2 次の Butterworth フィルタ、10 Hz)と標準化(z score の算出)の処理を行った上で、ヒルベルト変換を行い、ダンサー・DJ 各自の上下運動の位相を算出した。そして両者の位相の差を取ることで上下運動に関する相対位相を算出した(図 4)。

次に 2:ダンサーと DJ 両者の表現の細部(左手の振る舞い)の対応関係に関しても、同様に左手の振る舞いのデータに平滑化・標準化を行った上でヒルベルト変換を実施し、各自の左手の運動の位相を算出した。そして両者の位相差を取ることで相対位相を算出した。算出した相対位相は、0 度であれば相手と同一の振る舞いを行った様子(一方が上方に運動した際に他方も上方に運動したこと)を、180 度であれば相手と反対方向の振る舞いを行った様子(一方が上方に運動した際に他方が下方に運動したこと)を示唆する指標と考えられる(図 4)。以上の指標は、剣道における複数名間の振る舞いの対応関係を検討した研究を参考に算出した[Okumura 2012]。

また、今回はダンサーがパフォーマンス内で表現の特徴を大きく 2 回変化させる様子が見られた(フェイズ 1:立位で軽くステップを行うパフォーマンス、フェイズ 2:立位で激しくステップを行うパフォーマンス、フェイズ 3:床に手や頭を着けながら激しい運動を行うパフォーマンス)。そのため、パフォーマンス全体に加え、各パフォーマンスフェイズにおける上記の指標も算出し、時間経過による対応関係の変化に関する検討も行った。

さらに両指標の算出時に、実際にパフォーマンスを行ったペア(real pair)に加えて DJ の位置情報をもう一方のペアにおける DJ の位置情報と入れ替えた偽ペア(pseudo pair)も作成し、上記の指標を算出した。これは、流された音楽に元々含まれるリズムやビートによってパターン化された対応関係が生じることが予想されたためである。その影響のみを反映した対応関係の結果を算出し、そのベースラインとの比較を行うことで、実際のインタラクションにより生じる対応関係を同定することが可能となる。

2.4 Results and Discussion

まず、1:ダンサーと DJ 両者のリズム運動の対応関係(頸部の上下運動の相対位相)の結果を図 5 に示す。全体を集計した結果(図 5 の左から 1 つ目)から、real pair と pseudo pair ともに同位相もしくは逆位相の頻度が高く、ダンサーと DJ との間で振

る舞いにある程度の対応関係が見られたことが分かる。この相対位相の頻度を、ダンサーのパフォーマンスフェイズごとに分割したものが図 5 の左から 2 つ目−4 つ目である。図から real pair においては初期に同位相による対応が多く見られ、後期に逆位相による対応が多く見られるといったフェイズを経た対応関係の変化が見られた。一方で pseudo pair においては全フェイズを通じて逆位相による対応が多く見られ、その関係が大きく変化する様子は窺われない。以上の結果は、実際にダンサーと DJ がインタラクションを行う場面では、その関わり合いの様相を活発に変化させつつパフォーマンスを披露したことを示唆している。

同様の結果は、2:ダンサーと DJ 両者の表現の細部の対応関係（左手の振る舞いの相対位相）に関してより顕著に示唆された。全体の集計結果（図 6 の左から 1 つ目）を見ると、real pair では同位相と逆位相という振る舞いの異なる対応関係が多く見られたのに対し、pseudo pair では同位相という一定の対応関係のみが多く見られていた。各フェイズの結果（図 6 の左から 2 つ目−4 つ目）においても、real pair では初期に逆位相の対応が多く見られ、後期に同位相による対応が多く見られており、実際にダンサーと DJ が互いに関わり合いを活発に変化させながらパフォーマンスを披露した様子が窺われた。一方で pseudo pair では全フェイズを通して同位相の対応関係のみが多く見られており、関わり合いを変化させる様子は見られなかった。以上の結果を整理すると、実際のパフォーマンスにおいては、ダンサーと DJ との間にリズム運動や表現の細部といったパフォーマンスの多層的なレベルにおける対応関係が見られる様子、そしてその対応関係が時間経過を経て活発に変化する様子が示唆されたと考えられる。

3. General Discussion

ケーススタディでは、ブレイクダンスのパフォーマンス場面におけるダンサーと DJ との間に生じるインタラクションの探索的な検討を行った。結果として、各演者の振る舞いに複雑な対応関係が存在する可能性が示唆された。具体的には、大局的なリズム運動を表す頸部や、表現の細部を表す左手首など、多様な媒体において対応関係が見られたこと、そしてこれらの媒体における対応関係が時間経過によって活発に変化する様子が見られたこと、が示唆されている。

以上の結果は、上演芸術における同期理論の適用の有効性と上演芸術の特徴を踏まえた理論の拡張の必要性とを示唆するものと考えられる。1 章で推測した通り、上演芸術の演者間インタラクションにおいては、多様な媒体を用いて他者と多様なレベルで対応した振る舞いを多様なタイミングで披露する複雑な関わり合いが生じることが示唆された。そしてそのインタラクションは、パフォーマンスの展開と共に活発に変化する様相が窺われたのである。類似した結果は、ダンサー同士のインタラクションに関する知見からも示唆されている（清水・岡田，2016）。以上を踏まえると、上演芸術における演者間の関わり合いを検討する上で、同期理論において蓄積された枠組みや解析手法は有効に機能すると考えられるだろう。一方で、上記した複雑な様相を十分に捉えるためには、おそらく既存の同期理論における観点（同一の媒体における時間的に近接した類似した振る舞い）を超えた、多様な媒体間における複雑な対応関係を検証する必要があると考えられる。例えば、表現のどういった側面・媒体において他者とどのように協調した振る舞いが生じるのか、それらの観点を定量的に検討可能であるように同期理論の観点・解析手法を拡張することが必要と考えられる。実際に、近年 Beyond Synchrony という枠組みにおいて、ヒトの現実に生じる複雑な協調関係（互いの補助的な振る舞いの実施、発話内容の部分的な取り入れと発展）を捉えるための同期理論の拡張が営まれつつある[Dale 2013]。そこで用いられた観点や解析手法との関係性を踏まえつつ、対象とする表現の特徴を考慮した理論や解析手法の拡張が今後必要になることが想定される。

今後は、上記の理論や解析手法の拡張を行っていくとともに、上演芸術において生じる演者間の複雑なインタラクションを探索的に検証する試みを継続していく予定である。例えば、実際のバトル場面においては、図 2 で示唆されたように 3 者、4 者を超えた多数の参与者（ダンサー、DJ、MC、観客など）による振る舞いが高度に相互作用することでパフォーマンスが生まれ、展開していく。また、その関わり合いに関しても大きな個人差が存在する可能性が示唆されている[清水 2016]。以上の関わり合いの様相を詳細に検討するとともに、どのような関わり合いを通して魅力的な表現が生成されるのか、表現としての魅力にも着目した検討を今後は行っていく予定である。

Acknowledgment

本研究は、科学研究費補助金若手研究 B（課題番号：16K17306、代表：清水大地）の助成を受けて行われた。

参考文献

[Bailey 1980] Bailey, D.: IMPROVISATION, Buxton: Moorland Publishing, 1980.

[Dale 2013] Dale, R., Fusaroli, R., Døjbak Håkonsson, D. D., Healey, P., Mønster, D., McGraw, J., ... & Tylén, K.: Beyond synchrony: complementarity and asynchrony in joint action, In Proceedings of the Annual Meeting of the Cognitive Science Society, Vol. 35, No. 35, 2013.

[Merker 2015] Merker, B., Morley, I., & Zuidema, W.: Five fundamental constraints on theories of the origins of music, Philosophical Transactions of the Royal Society B: Biological Sciences, 370(1664), 20140095, 2015.

[Néda 2000] Néda, Z., Ravasz, E., Brechet, Y., Vicsek, T., & Barabási, A. L.: Self-organizing processes: The sound of many hands clapping, Nature, 403(6772), 849, 2000.

[OHJI 2001] OHJI: ROOTS OF STREET DANCE, 東京：ぶんか社, 2001.

[Okumura 2012] Okumura, M., Kijima, A., Kadota, K., Yokoyama, K., Suzuki, H., & Yamamoto, Y.: A critical interpersonal distance switches between two coordination modes in kendo matches, Plos One, 7(12), e51877, 2012.

[清水 2013] 清水大地・岡田猛：ストリートダンスにおける即興的創造過程，認知科学, 20(4), 421-438，2013.

[清水 2017] 清水大地・岡田猛：舞台表現における他者との相互作用のダイナミクスーコミュニケーションの隠れた次元としての距離による検討，HCS2016，95，29−34，2017.

[Strogatz 2003] Strogatz, S. H.: Sync: rhythms of nature, rhythms of ourselves, Allen Lane, 2003.

[Walton 2018] Walton, A. E., Richardson, M. J., Langland-Hassan, P., & Chemero, A.: Improvisation and the self-organization of multiple musical bodies, Frontiers in Psychology, 6: 313, 2015.

[Weinstein 2016] Weinstein, D., Launay, J., Pearce, E., Dunbar, R. I., & Stewart, L.: Singing and social bonding: changes in connectivity and pain threshold as a function of group size, Evolution and Human Behavior, 37(2), 152-158, 2016.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

ビジネスプレゼンテーションにおける言語・非言語的能力の自動推定

Estimating Verbal・Nonverbal Skills in Business Presentation

八木 悠太朗 *1 Yutaro Yagi　岡田 将吾 *1 Shogo Okada　塩原 翔太 *2 Shota Shiobara　杉村 聡太 *2 Sota Sugimura

*1北陸先端科学技術大学院大学 先端技術研究科
School of Information science, Japan Advanced Institute of Science and Technology (JAIST)

*2ソフトバンク株式会社
SoftBank Corp.

This paper focuses on developing a model for estimating presentation skills of each participant from multimodal (verbal and nonverbal) features. For this purpose, we use a multimodal presentation dataset including audio signal data and body motion sensor data, text data of speech contents of participants observed in 58 presentation sessions. The dataset also includes the presentation skills of each participant, which is assessed by two external observers of the Human Resources Department. We extracted various kinds of features such as spoken uttetances, acoustic features, and the amount of body motion to estimate the presentation skills. We created a regression model to infer the level of presentation skills from these features using support vector regression to evaluate the estimation accuracy of the presentation skills. Experiment results show that the multimodal model achieved 0.59 in R^2 as the regression accuracy of effective production elements.

1. はじめに

企業やビジネスの場において，プレゼンテーション能力は重要なスキルのひとつである．プレゼンテーション能力が必要とされる場面として，企画やアイデアを発表するプレゼンテーションや商品やサービス内容を説明する店頭販売，採用面接などが挙げられる．プレゼンテーション能力を数値化することが出来れば，プレゼンテーションの振り返りや練習などを行う自動評定システムの構築に役立つ．本研究では,「プレゼンテーション能力」に関する個人特性を計算モデルにより捉えることに焦点を当てる．

プレゼンテーション能力の向上支援に向け，本研究は，プレゼンテーションを通じて観測できる参加者個人の話し方，体の動作，発話内容から，プレゼンテーション能力値を推定するモデルを機械学習により構築し，評価することを目的とする．推定モデルの構築に用いるデータセットを，ソフトバンク株式会社で行われたプレゼンテーション研修にて収集した．

2. 関連研究

2.1 プレゼンテーションスキルの推定

Torsten ら [Torsten 15] は，Public speaking skill の評定値と相関関係にあるプレゼンテーションから得られる非言語情報を分析し，それらを推定するモデルを構築した．相関係数 0.75 の予測精度を得た．また，音声と動作の特徴量を両方用いた場合に予測精度が向上することも示した．Vikram ら [Vikram 15] は，プレゼンテーションの音声情報や話者の表情，動作などの時系列マルチモーダル情報から特徴量を抽出し，個人のプレゼンテーション能力の評定値を推定するモデルを構築した．ある動作の後に別の動作を行うなどの時間的な共起パターンを共起ヒストグラムによって特徴量として抽出し，その特徴量がプレゼンテーション能力の評価にどれだけ貢献するのか分析した．相関係数で最大 0.69 の予測精度を得た．前者の研究は，プレゼンテーション能力に関係する特徴量の分析を，後者の研究は，プレゼンテーション能力の向上支援のためのシステムを構築している．[Torsten 15, Vikram 15] の目的は本研究の目的に類似する点はあるが，いずれも非言語情報のみを用いており，本研究のようにプレゼンテーション中の言語内容が重要視されるビジネスプレゼンテーションを対象としていない．

福島ら [福島 16] は，TED のウェブサイトで公開されている 1646 本のプレゼンテーション動画に対して視聴者が抱く印象を予測するモデルを構築した．プレゼンテーション動画から得られた言語情報と音声情報を用いて，複数の視聴者がアノテーションした beautiful，confusing などの合計 14 の印象の中から特定の印象を抱くか否かの 2 クラス分類を行い印象の予測を行った．分類の最大精度は 97.0%であった．この研究はプレゼンテーションの印象を予測することが目的であり，本研究のようにプレゼンテーション能力の推定を目的としていない．

2.2 本研究の位置づけ

関連研究と本研究の相違点をまとめ，本研究の貢献を明らかにする．

言語・非言語両面を考慮した評価尺度の生成

本研究はビジネスプレゼンテーションの評価を対象としている．ビジネスプレゼンテーションでは，プレゼンテーションの内容が重要視される．したがって，ジェスチャーや音声を考慮した評価尺度に加えて，言語面を重視した新たな評価尺度を作成した．

言語・非言語情報の利用

[Torsten 15, Vikram 15] では主に非言語情報に基づき分析を行っている．一方で言語的な能力を推定するためには，プレゼンテーションにおける発話内容は重要である．本研究では非言語特徴量だけでなく，論理性やプレゼン内容の評定を予測するために，発話内容に含まれる名詞数などの言語特徴量も用いて，評定値を推定する機械学習モデルを構築する．

連絡先: 八木　悠太朗，北陸先端科学技術大学院大学，s1810184@jaist.ac.jp

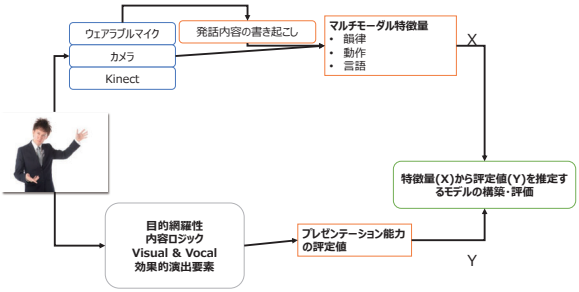


図 1: プレゼンテーション能力を推定する枠組み

3. 方法

3.1 概要

本研究の概要を図 1 にまとめる．収集したプレゼンテーションデータから音声・動作・言語に関する特徴量を抽出し，個人のプレゼンテーション能力を推定する機械学習モデルを構築し，評価する．推定モデルの構築に用いるデータセットは，ソフトバンク株式会社で行われているソフトバンクユニバーシティ「プレゼン話し方研修」で収集した．収集したデータセットに対して，ソフトバンク株式会社の人事部にプレゼンテーション能力の評価を依頼した．本章では，データセットの概要，抽出するマルチモーダル特徴量，および評価項目について述べる．

3.2 データセット

本実験で使用するデータセットの概要を説明する．データセットには，ウェアラブルマイクから収音した音声情報，Kinect から収音した音声情報，Kinect から得られた深度情報，Kinect から得られた各関節の位置情報，音声情報から文字起こしした発話内容が含まれる．このうち本研究では，ウェアラブルマイクから収音した音声情報，Kinect から得られた骨格の位置座標，音声情報から文字起こしした発話内容を用いてマルチモーダル特徴抽出を行う．実験には 58 人が参加し，一人一回 3 分間のプレゼンテーションを行う．実験で得られた合計 58 個のプレゼンテーションデータを実験に用いる．

3.3 特徴量

[岡田 16] を参考に韻律特徴量，言語特徴量，動作特徴量を抽出する．

3.3.1 韻律特徴量

ウェアラブルマイクから収音した音声情報から発話区間を推定し，OpenSMILE*1 を用いて特徴抽出を行った．収音した音声のエネルギーの閾値処理に基づいた発話区間推定を用いた．The INTERSPEECH 2010 Paralinguistic Challenge feature set に含まれる 1582 次元の特徴量を抽出した．これらの特徴量には音の大きさ，0 次〜14 次メル周波数ケプストラム係数，0 次〜7 次メル周波数帯の対数パワー，8 つの LPC 係数から算出された 8 つの線スペクトラムペア周波数，平滑化基本周波数，最後の基本周波数の有声確率，及びそれらの一次微分が含まれる．

3.3.2 言語特徴量

音声情報から文字起こしした発話内容のテキストデータに対して形態素解析を行い，単語の品詞情報を取得した．形態素解析には MeCab*2 を用いた．得られた単語の品詞情報をもとに，抽出した言語特徴量を表 1 にまとめる．

表 1: 抽出する言語特徴量

特徴量	詳細
名詞数	発話内容に含まれる名詞数
動詞数	発話内容に含まれる動詞数
接続詞数	発話内容に含まれる接続詞数
フィラー数	発話内容に含まれるフィラー数
感動詞数	発話内容に含まれる感動詞数
同一名詞繰り返し最大数	同一の名詞を繰り返した最大数
同一動詞繰り返し最大数	同一の動詞を繰り返した最大数
名詞出現種類数	発話内容に含まれる名詞の種類数
動詞出現種類数	発話内容に含まれる動詞の種類数

3.3.3 動作特徴量

Kinect から得られた各関節 (上半身 19 部位) の位置座標の時間変化を計算して，動作量と速度，加速度についてそれぞれ平均と標準偏差を求めた．また，発話中についても同様の特徴量を抽出した．

3.4 プレゼン研修講師による評定

プレゼンテーション評定の熟練者として，ソフトバンク株式会社のソフトバンクユニバーシティの講師二名が評定を行った．研修に参加した 58 名のプレゼンテーション映像を閲覧し，15 個の評価項目について，最低 0〜最高 2 の 3 段階で評定した．それらを 4 つの大項目に分類し，各項目の平均値を算出したものを評定値として用いた．評価項目とその定義を以下にまとめる．

目的網羅性

「誰に」,「何を」,「どうしてほしい」が伝えられている．

内容ロジック

「結論」,「根拠」,「相手の利益」が述べられている．

Visual & Vocal

「抑揚」,「声量」,「アイコンタクト」,「ジェスチャー」,「表情」が適切に用いられている．

効果的演出要素

「強調」,「繰り返し」,「具体表現」,「双方向性」が適切に用いられている．

各評定者による評定値の一致度を計算し，各項目に関する信頼性を確認した．一致度の計算には，カッパ値と重み付きカッパ値を算出した．ある項目に関してカッパ値が 0.4 以上であれば各評定者による参加者への評定値は概ね一致しており，その項目の評定値は信頼性が高いと言える [Landis 77]．

表 2 に項目ごとのカッパ値と重み付きカッパ値を示す．図中では，小数点以下第三位を四捨五入した値を記載している．全項目でカッパ値は 0.4 を上回っており，項目としての信頼性は高いと考える．

4. 評価実験

回帰予測タスクを行い，各参加者の評定値 (連続値) の推定精度を評価する．

*1 https://www.audeering.com/opensmile/

*2 http://taku910.github.io/mecab/

表 2: 各評定者間での評定値の一致度

能力項目	重み付きカッパ値	カッパ値
目的網羅性	0.99	0.81
内容ロジック	0.89	0.81
Visual & Vocal	0.91	0.54
効果的演出要素	0.99	0.93

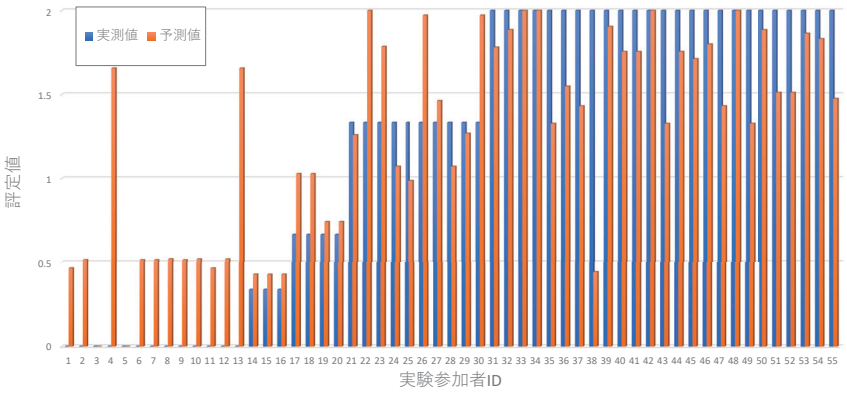


図 2: 「$All\,(p < 0.01)$」のモデルを用いた「目的網羅性」の評定値の予測値と実測値

4.1 推定モデルの構築手順

参加者の特徴量セットを入力 X，そのプレゼンテーションの評定値を出力 Y として，この入出力関係を学習する．合計で 58 個のデータを利用できる予定であったが，3 データの特徴量が機器の不具合などで一部欠損していたため，これらを除外し合計 55 個のデータを用いた．評価には一つ抜き (一人抜き) 交差検定法を利用した．ある参加者から得られたデータをテスト，それ以外の 54 個のデータを訓練に用いて回帰モデルの評価実験を行った．

回帰学習の概要

回帰学習には，RBF カーネルを用いた非線形サポートベクトル回帰モデル (Support Vector Regression:SVR) を用いた．SVR における誤差関数には，ϵ 許容誤差を用いた．誤差の許容度を調整する ϵ，損失関数とマージンの大きさの間のトレードオフを調整するパラメータである C，RBF カーネルの形状を調整するパラメータである γ に関して，それぞれ [0,0.01,0.1,1],[0.01,0.1,1,5,10]，[0.0001,0.001,0.01,0.1] の探索域を設け，訓練データを用いた交差検定を通じて，グリッドサーチにより最適パラメータを決定した．

回帰モデルの推定性能の評価指標にはテストデータに対する決定係数 R^2 を用いた．テストデータに対する評定値 (正解) と推定値の二乗誤差を，テストデータに対応する評定値の分散で割った値を E とすると，決定係数 R^2 は $1 - E$ で計算される．したがって，誤差が小さいほど R^2 は大きくなり，良い結果を示す．

4.1.1 比較対象の特徴量セット

各モダリティの推定性能への寄与を検証するため，以下の 9 種類の特徴量セットを準備し，推定精度の比較を行う．韻律，言語，動作の特徴量セットを A，L，M とそれぞれ略記する．

1. A：韻律特徴量
2. L：言語特徴量
3. M：動作特徴量
4. $A + L$：韻律と言語
5. $A + M$：韻律と動作
6. $L + M$：言語と動作
7. All：全ての特徴量 $(A + L + M)$
8. $All\,(p < 0.05)$：相関分析で有意な相関があった特徴量
9. $All\,(p < 0.01)$：相関分析で有意な相関があった特徴量

4.2 実験結果

プレゼンテーション能力を構成する 4 つの要素項目の評定値を推定する回帰タスクの結果を表 3 に示す．表 3 は各特徴量セットを用いて，各評価項目を推定した結果を表している．推定精度にはテストデータに対する決定係数 R^2 を用いた．

§1 単一モダリティ特徴量セットのモデルの精度

単一モダリティの特徴量セットを用いたモデルに関して「目的網羅性」,「内容ロジック」で「韻律特徴量 (A)」のモデルの精度が最大であった.「Visual & Vocal」では,「動作特徴量 (M)」のモデルの精度が最大であった.「効果的演出要素」では,「言語特徴量 (L)」のモデルの精度が最大であった.「目的網羅性」,「内容ロジック」,「Visual & Vocal」,「効果的演出要素」の最大推定精度はそれぞれ 0.32，0.07，0.05，0.00 であった．

§2 マルチモーダル特徴量セットのモデルの精度

マルチモーダル特徴量セットを用いたモデルに関して「目的網羅性」,「効果的演出要素」で「$All\,(p < 0.01)$」のモデルの精度が最大であった.「内容ロジック」で「$All\,(p < 0.10)$」のモデルの精度が最大であった.「Visual & Vocal」で「$All\,(p < 0.05)$」のモデルの精度が最大であった.「目的網羅性」,「内容ロジック」,「Visual & Vocal」,「効果的演出要素」の最大推定精度はそれぞれ 0.59，0.51，0.38，0.29 であった．

「目的網羅性」,「内容ロジック」,「Visual & Vocal」,「効果的演出要素」の要素項目に関して，マルチモーダル情報の統合は，単一モダリティ特徴量セットのモデルの精度を向上させるために役立つ．

推定精度が最大であった特徴量セットを以下にまとめる.「目的網羅性」,「効果的演出要素」の推定には $All\,(p < 0.01)$ が有効であった.「内容ロジック」の推定には $All\,(p < 0.10)$ が有効であった.「Visual & Vocal」の推定には $All\,(p < 0.05)$ が有効であった．

5. 考察

単一モダリティ特徴量セットの精度よりも，複数のモダリティを統合したマルチモーダル特徴量セットのモデルの精度のほうが高くなる傾向にあった．これは [Torsten 15] の結果とも一致している．

韻律特徴量について，今回は集団での研修の場においてデータ収集を行ったため，ウェアラブルマイクで収音した音声データにノイズが含まれている場合があった．実用化に向けては，音声認識を用いて自動的に発話内容から言語特徴量を抽出する必要がある．本研究で用いた音声データに音声認識を用いて得

表 3: SVM を用いた回帰タスクの結果：各特徴量セットを用いて各プレゼンテーション能力項目を推定した場合の精度を示している．精度はテストデータに対する決定係数 R^2 を用いた．

	目的網羅性		内容ロジック		Visual & Vocal		効果的演出要素	
	R^2	MSE	R^2	MSE	R^2	MSE	R^2	MSE
A	0.32	0.47	0.07	0.57	-0.11	0.20	-0.09	0.21
L	-0.69	1.18	-0.11	0.67	-0.21	0.22	0.00	0.19
M	-0.23	0.86	-0.36	0.82	0.05	0.17	-0.34	0.26
$A+L$	0.31	0.48	0.07	0.57	-0.12	0.20	-0.04	0.20
$A+M$	-0.16	0.81	0.22	0.47	0.16	0.15	-0.17	0.23
$L+M$	-0.25	0.87	-0.18	0.71	-0.31	0.24	-0.29	0.25
All	-0.17	0.81	0.21	0.48	0.17	0.15	-0.17	0.23
$All\,(p<0.10)$	0.54	0.32	**0.51**	0.30	0.23	0.14	0.24	0.15
$All\,(p<0.05)$	0.47	0.37	0.39	0.37	**0.38**	0.11	-0.14	0.21
$All\,(p<0.01)$	**0.59**	0.28	-0.32	0.80	0.13	0.16	**0.29**	0.14

られた言語情報から言語特徴量を抽出した場合の推定精度についても調査する必要がある．

動作特徴量について，本研究では体の各部位の動作量や速度，加速度を特徴量として抽出した．しかし「Visual & Vocal」の評価項目では視線や表情も考慮されているため，視線や表情に関する特徴量を加えモデルを構築する必要がある．

言語特徴量について，本研究では音声情報から文字起こしした発話内容のテキストデータに対して形態素解析を行い，得られた単語の品詞情報から言語特徴量を抽出している．また学習の前に相関分析を行い，有意な相関がある特徴量のみから各評価項目の評定値を推定した場合に最も高い推定精度が得られた．言語面の評価項目である「目的網羅性」,「内容ロジック」に対して，名詞出現種類数，動詞出現種類数は $p < 0.10$ で有意な相関を示した.「効果的演出要素」では，名詞数，動詞数，名詞出現種類数，動詞出現種類数の特徴量が $p < 0.05$ で有意な相関を示した．単語レベルでなく，構文レベルの解析を行い,「目的網羅性」,「内容ロジック」の項目の推定に寄与する特徴量を抽出することが今後の課題の一つである．

図 2 は予測精度が最大であった「$All\,(p < 0.01)$」のモデルを用いた「目的網羅性」の評定値の予測値と実測値を比較している．評定値が 0 点〜1 点，1 点〜2 点の場合，推定誤差の平均がそれぞれ 0.16，0.35 であるのに対し，0 点の場合の推定誤差は 0.60 であった．この結果より，評定値が 0 のデータの推定精度が低いことがわかった．このようなデータに対しても高い精度で推定を行える機械学習モデルを構築する必要がある．

6. 結論

本研究では，プレゼンテーションデータから得られたマルチモーダル情報から「プレゼンテーション能力」を推定するモデルの構築・評価を行った．発話内容に含まれる品詞情報，韻律情報，動作量をマルチモーダル特徴量として抽出し，個人のプレゼンテーション能力値を推定するモデルを機械学習により構築した．評価実験の結果,「目的網羅性」,「内容ロジック」,「Visual & Vocal」,「効果的演出要素」に関する評定値に関して，回帰タスクでそれぞれ最大 0.59，0.51，0.38，0.29 の決定係数 R^2 を得た．

本研究では，プレゼンテーション中の言語・非言語特徴量の時系列変化を考慮していない．時系列マルチモーダル情報を用いたモデルの構築は今後の課題である．

7. 謝辞

ソフトバンク株式会社 人事本部の笠井 理恵様，高嶋 直人様，海上 博志様には，データセットの収集，および評定値の作成においてご支援いただきました．心より感謝申し上げます．

参考文献

[岡田 16] 岡田 将吾, 松儀 良広, 中野 有紀子, 林 佑樹, 黄 宏軒, 高瀬 裕, 新田 克己, マルチモーダル情報に基づくグループ会話におけるコミュニケーション能力の推定, 人工知能学会論文誌, 2016, 31 巻, 6 号, p. AI30-E_1-12

[福島 16] 福島 悠介, 山崎 俊彦, 相澤 清晴, 文書と音声解析に基づくプレゼンテーション動画の印象予測, 電子情報通信学会論文誌 D Vol.J99-D No.8 pp.699-708, 2016

[Landis 77] Landis, J.R.; Koch, G.G. (1977). "The measurement of observer agreement for categorical data". Biometrics. 33 (1): 159–174.

[Torsten 15] Torsten Wörtwein, Mathieu Chollet, Boris Schauerte, Louis-Philippe Morency, Rainer Stiefelhagen, and Stefan Scherer. 2015. Multimodal Public Speaking Performance Assessment. In Proceedings of the 2015 ACM on ICMI (ICMI '15).

[Mathieu 16] Mathieu Chollet, Helmut Prendinger, and Stefan Scherer. 2016. Native vs. non-native language fluency implications on multimodal interaction for interpersonal skills training. In Proceedings of the 18th ACM on ICMI (ICMI 2016).

[Vikram 15] Vikram Ramanarayanan, Chee Wee Leong, Lei Chen, Gary Feng, and David Suendermann-Oeft. 2015. Evaluating Speech, Face, Emotion and Body Movement Time-series Features for Automated Multimodal Presentation Scoring. In Proceedings of the 2015 ACM on ICMI (ICMI '15).

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

同時計測されたEEG信号からの2名における感情共有の測定

Measuring affective sharing between two people from simultaneously recorded EEG signals

木下 泰輝 Taiki Kinoshita　田中 宏季 Hiroki Tanaka　吉野 幸一郎 Koichiro Yoshino　中村 哲 Satoshi Nakamura

奈良先端科学技術大学院大学　先端科学技術研究科
Nara Institute of Science and Technology, Graduate School of Science and Technology

Empathy has been studied intensively in the field of human-machine interaction and human-human interaction. In many of these studies, empathy is subjectively measured by questionnaire. We propose a method to directly measure human empathy from neural potential activity. In this research, EEG signals were recorded simultaneously from two subjects engaged in emotional communication. After time frequency analysis was performed on the recorded Electroencephalogram (EEG) signals, a cross correlation functions between EEG signals of two subjects were calculated. To investigate the significance of cross correlations, paired t-test and false discovery rate correction were performed. As a result, there was a significant cross correlation in the gamma frequency band in the joy condition. In addition, there was a marginally significant correlation in the sadness condition. From these results we showed the possibility to measure empathy from cross correlation of EEG signals recorded from two subjects.

1. はじめに

共感は，人間のコミュニケーションにおいて重要な要素であり，社会心理学において盛んに研究が行われている．さらに，共感は，人間と人工エージェントのコミュニケーションにおいても重要性が指摘されいる [Złotowski 16]．このため，ヒューマンコンピュータインタラクションやアフェクティブコンピューティングの分野で，人工エージェントに，社会心理学において提案された共感のモデルを取り入れ，より良いコミュニケーションの実現を目指す研究が行われている．しかし，このような研究においては，提案された人工エージェントに対する確立された評価手法が存在せず，さらなる研究が必要とされている [Paiva 17]．

共感は，認知的，感情的なメカニズムを含む複雑な現象であり，合意の取れた定義は存在しない．このことが，共感の計測が困難である要因の一つである [Neumann 15]．しかし，多くの共感研究に共通して，共感の要素に感情共有 (Affective sharing) が存在することが指摘されている [Decety 04]．感情共有とは，他者の感情を，自他の区別を伴った状態で経験し，無意識に共鳴することをさす．感情共有には，脳活動や表情の同期現象が伴うことが知られている．本研究では，共感に関係する多様な現象のうち，感情共有を対象とする．

感情共有に対する研究では functional magnetic resonance imaging (fMRI) が頻繁に使用される．しかし，fMRI は大規模な装置を必要とし，また被験者が計測中に身動きを取れないことから，人工エージェントとの自然な対話を評価するなどの、実環境での利用は難しいと考えられる．一方，Electroencephalogram (EEG) は，脳活動を計測する代表的な手法の一つであり，計測が容易であることから，実環境下で利用可能という利点が存在する．以上のことから，本研究では，人工エージェントに対する共感の評価手法を提案する前段階として，2 名の人間における感情共有を EEG によって計測することを目的とする．

連絡先: 木下 泰輝，奈良先端科学技術大学院大学，奈良県生駒市高山町 8916-5 情報棟 B712，電話: 0743-72-5264，電子メール: kinoshita.taiki.kk1@is.naist.jp

2. 感情共有の計測方法

人間間の共感研究では，fMRI を用いた感情共有の計測が行われている [Anders 11]．この研究では，2 名の実験協力者を募集し，1 名に送信者と呼ばれる役割を，もう 1 名に受信者と呼ばれる役割を割り当てた．実験において，送信者は感情的な場面を想像し，感情表出を行う課題を行った．一方，受信者は送信者の表情を見て共感する課題を行った．この課題を行っている際の 2 名の脳活動が，2 台の fMRI によって同時計測された．得られた fMRI データに対する分析を行った結果，2 名の脳活動の同期現象が確認された．

一方，多くの人間エージェント間の共感研究では，共感と関連を持つと考えられている主観評価によって間接的に共感が評価されている．例えば，人工エージェントの信頼感，社会的存在感，知的さといった指標が利用されており，直接的な評価手法の必要性が指摘されている [Paiva 17]．

本研究では，EEG を利用した感情共有の計測を提案する．このため，fMRI を利用した場合と比較して，実環境に近い状態で感情共有を計測することが可能となる．また，今後人工エージェントの評価に応用することにより，主観評価を利用した場合と比較して，直接的に共感を評価することが可能となることが期待される．

3. EEG による感情共有の測定

3.1 実験協力者

実験協力者は，2 名の女性とした．ここで，実験協力者を 2 名と限定したのは，EEG には個人差が現れることが知られているためである．この個人差の解消は次研究の課題とし，本研究では 2 名の実験協力者のみを対象とする．また，実験協力者を女性としたのは，女性が男性と比較して感情共有を持ちやすいことがが知られているためである [David 80]．

3.2 実験手続き

本実験は，[Anders 11] で fMRI を用いて行われた実験が EEG で同様に再現可能か調査することを目的とする．図 1 に実験の概要を示す．実験協力者のうち 1 名が送信者の役割を，もう 1 名が受信者の役割を担当した．これら 2 名のうち，送信

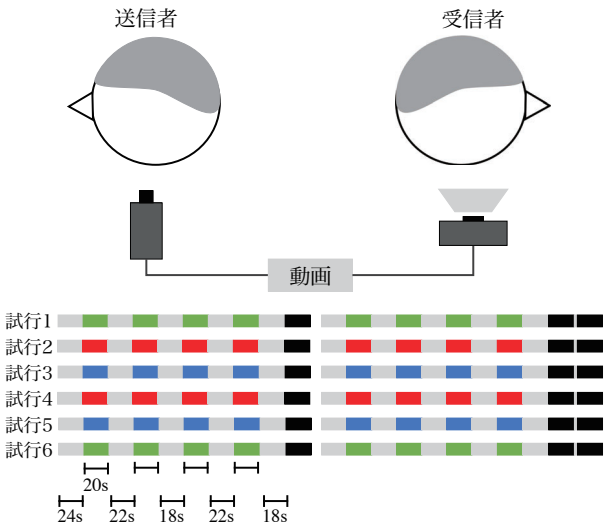


図 1: 実験の概要．各色は，感情の種類を表している（喜び，悲しみ，無感情）．また，黒色は SAM を表している．本研究では，実験を 2 度行い，合計試行回数は 12 回となった．

者は休憩，感情表出という 2 種類の課題を行った．休憩では，教示用モニタに休憩という文字を表示した．この間，送信者はリラックスするよう指示された．一方，感情表出では，3 秒間のカウントダウンを表示した後，喜び，悲しみ，無感情のいずれかの文字を表示した．この間，送信者は，感情を感じる場面をそれぞれ想起し，それらの感情を表情によって表出した．ここで，喜び，悲しみは，[Anders 11] において利用された 5 種類の感情のうち，最も受信者の感情認識精度が高いと報告されている 2 種類の感情である．また，無感情は，喜び，悲しみと比較を行うためのベースラインである．本実験では，1 試行で，休憩と感情表出を交互に 4 回行った．休憩の時間は 18〜24 秒から無作為に決定し，感情表出の時間は 20 秒とした．また，試行の最後には，Self Assessment Mannequin (SAM) [Bradley 94] により，試行中の感情価が 9 段階で回答された．以上の試行が合計 12 回行われた．

一方，受信者には，送信者の顔の表情が Web カメラ（Microsoft 社製 LifeCam Studio）によって，リアルタイムで提示された．受信者はこの映像から送信者の顔の表情を観察し，共感するという課題を行った．また，各試行の最後に，試行中に認識した送信者の感情と，受信者自身の感情をそれぞれ SAM により回答した．

以上の課題を，2 名の実験協力者が同時に行っている間， 2 名の EEG データ・並びに electromyography (EMG) を同時に収録した．ここで，EMG とは，筋電を収録する機器である．本研究では，EEG を装着した状態で表情を動かすため，筋電がノイズとして混入することが予想される．このため，左皺眉筋ならびに左大頬骨筋の EMG を EEG と共に収録し，後処理で EEG データの筋電ノイズを削除した．EEG，EMG は，cognionics 社製 Quick-30 を使用して収録した．ここで，EMG は，EEG チャンネルのうち AF3，FC5 を置き換えることにより収録した．

3.3 EEG データの前処理

前処理は，以下の 6 つの手順で行った．ただし，手順 1)〜4) においては，EEGLAB [Delorme 04] を利用し，手順 5)〜6) においては，MNE [Gramfort 14][Gramfort 13] を利用した．1) サンプリング周波数を 200Hz にダウンサンプリングした．2) 0.3〜45Hz のバンドパスフィルタをかけた．3) 感情の指示が出された点をオンセットとし，-20 秒〜20 秒をエポッキングした．エポッキングにより，各感情について，それぞれ 16 個ずつのエポックが切り出された．4) ノイズ除去を行った．具体的には，独立成分分析を行った後，筋電を除去するため，EMG との相互相関が 0.1 以上である成分を削除した．また，ADJUST [Mognon 11] を利用して，ノイズ成分を自動削除した．さらに，EEG のチャンネルのうち，最も筋電の影響を受けやすいと考えられる T7，T8 は削除した．5) 特徴量の抽出を行うため，時間周波数解析を行った．モルレーウェーブレット変換を行い，各周波数のパワーを算出した．この上で，ガンマ波を 30〜40Hz，ベータ波を 12〜30Hz，アルファ波を 7〜12Hz，セータ波を 4〜7Hz，デルタ波を 1〜4Hz と設定し，各周波数帯でパワーの平均値を算出した．6) 個人差の影響を軽減するため，パワーを各チャンネル毎に平均値 0，標準偏差 1 に標準化した．

3.4 感情共有の分析手法

EEG データに対して行った分析の方法を述べる．多くの研究で，2 名の脳活動の関連性は，EEG チャンネル同士の相関によって計測されている [Babiloni 14]．そこで，本研究では，送信者と受信者の EEG チャンネルの相互相関を，全ての組み合わせに対して算出した．ただし，相互相関のシフトは 1 秒以下に限定し，相互相関が最大となるシフトを選択した．さらに，算出された相関係数が有意であるかを決定するため，休憩期間と感情表出期間の相互相関を対応のある t 検定により比較した．ただし，false discovery rate は Benjamini-Hochberg 法により調整した．以上の分析を行った上で，各チャンネルを頂点とし，有意 ($p < 0.05$) または有意傾向 ($p < 0.1$) な相互相関をもつチャンネル同士に辺を持つグラフを構築した．すなわち，対応のある t 検定の結果によって，相互相関の閾値を設定した．さらに，構築されたグラフに対して，Global efficiency を算出した．Global efficiency とは，グラフ理論において提案された，情報伝達の効率性を表す指標である．Global efficiency の定義は，Average efficiency に基づいている．Average efficiency は，各頂点間の最短経路長の平均値を表しており，式 (1) のように定義される．

$$E(G) = \frac{1}{n(n-1)} \sum_{i \neq j \in G} \frac{1}{d(i,j)} \tag{1}$$

ここで，n は全ノード数を表している．また，$d(i,j)$ は，ノード i と j の最短経路長を表しており，本研究においては，エッジの有無を表している（ノード間に繋がりがあれば 1 を取り，なければ無限大を取る）．この時，Global efficiency は，式 (2) で表される．

$$E_{glob}(G) = \frac{E(G)}{E(G^{all})} \tag{2}$$

ここで，$E(G^{all})$ は，全結合グラフにおいて算出された Average efficiency を表している．本研究においては，送信者のチャンネルと受信者のチャンネルが全結合されたグラフの Average efficiency となる．Global efficiency は 0〜1 の値を取り，値が大きいほど効率性が高いことを意味する．この指標は，EEG 研究においても，2 名の脳活動の繋がりの強さを指標化する目的で利用されている [Ciaramidaro 18]．

4. 結果

まず，SAM に対する分析を行った結果を述べる．送信者の感情価と受信者が認識した感情価を図 2 に示す．送信者の感

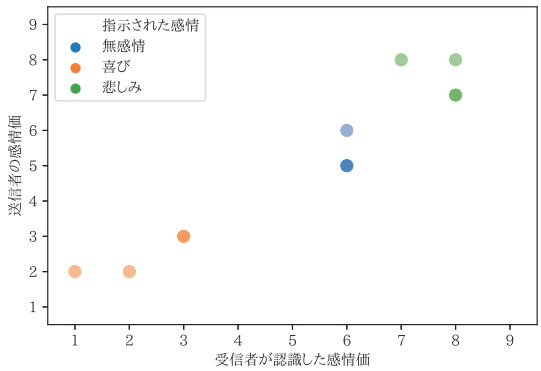


図 2: 送信者の感情価と受信者が認識したが認識した感情価．重なっている点は，色を濃く表示している．

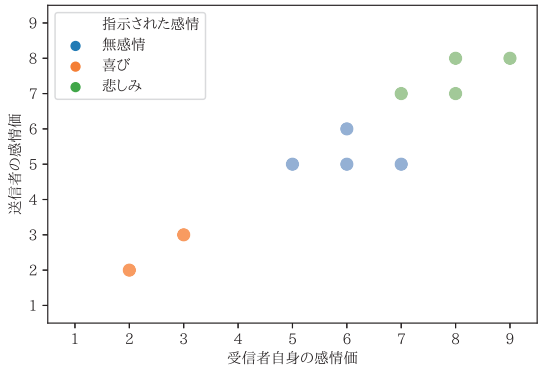


図 3: 送信者の感情価と受信者自身の感情価．重なっている点は，色を濃く表示している．

情価と受信者が認識した感情価の相関係数は 0.954 であり強い正の相関が見られた．次に，送信者の感情価と受信者自身の感情価を図 3 に示す．この時の相関係数は 0.966 であり強い正の相関が見られた．

次に，相互相関に対して，対応のある t 検定を行った結果を述べる．喜びでは，ガンマ波において，28 個の相互相関が有意であった ($p < 0.05, n = 16$)．ここで，n は，対象感情のエポック数を表している．その他の周波数帯では，有意な相互相関は見られなかった．図 4 に，t 値が大きかったチャンネルの組み合わせを 10 個示す．線が引かれているチャンネル同士に有意な相互相関が存在した．図 4 に示した通り，送信者のチャンネルと受信者の後頭部に位置するチャンネルに高い相互相関が見られた．一方，悲しみでは有意な相互相関は見られなかった．しかし，ガンマ波において有意傾向が見られた ($p < 0.1$, $n = 16$)．図 5 に，t 値が大きかったチャンネルの組み合わせを 10 個示す．図 5 に示した通り，悲しみにおいても，喜びと同様に送信者のチャンネルと受信者の後頭部に位置するチャンネルに高い相互相関が見られた．一方無感情では，アルファ波において，休憩期間の方が有意に相互相関が高い組み合わせが一つ存在した ($p < 0.05, n = 16$)．

最後に，Global efficiency を算出した結果を述べる．有意 ($p < 0.05$) な相互相関を基に Global efficiency を算出した結果，喜びのガンマ波において，0.097 だった．その他の感情，

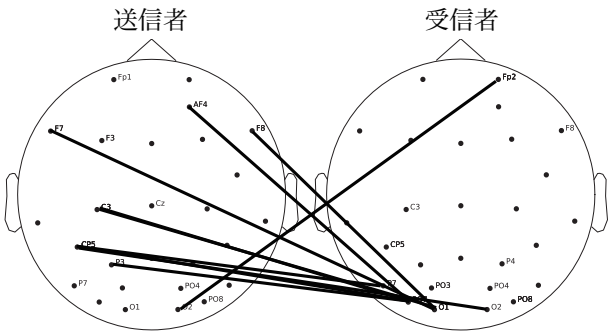


図 4: 喜びにおいて相互相関が有意だったチャンネルの組み合わせのうち，t 値が大きかったチャンネルの組み合わせ 10 個に線が引いた ($p < 0.05$) その他の有意なチャンネルは，チャンネル名のみ示した．

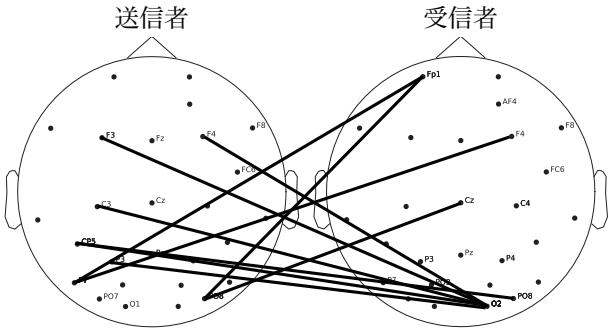


図 5: 悲しみにおいて相互相関が有意傾向だったチャンネルの組み合わせのうち，t 値が大きかったチャンネルの組み合わせ 10 個に線が引いた（$p < 0.1$）．その他の有意傾向なチャンネルは，チャンネル名のみ示した．

周波数帯においては全て 0 だった．次に，有意傾向 ($p < 0.1$) な相互相関を基に Global efficiency を算出した結果を表 1 に示す．表 1 に示した通り喜びのガンマ波，ベータ波，悲しみのガンマ波において 0 以外の値をとった．しかし，その他の周波数帯においては，全て 0 の値をとった．総じて，高周波数帯において Global efficiency が高い傾向が見られた．また，喜び，悲しみにおいて，無感情と比較して Global efficiency が高い傾向が見られた．

5. 考察

SAM に対する分析を行った結果より，送信者の感情が受信者によって正しく認識されていると考えられる．また，送信者の感情表現により受信者に近い感情が誘発され，感情共有が生じている状態で，脳活動の計測を行うことができたと判断できる．

次に，相互相関に対して対応のある t 検定を行った結果より，感情共有による脳活動の同期を，EEG を利用して計測することができる可能性が示された．しかし，悲しみにおいて有意な結果が得られなかったため，試行の回数を増やし再度分析を行うことが必要であると考えられる．

最後に，Global efficiency を算出した結果より，感情共有が生じている時，Global efficiency が高まる傾向が見られた．これにより，この指標を利用することで，感情共有を計測するこ

表 1: 2 名の電極相互相関に基づく Global efficiency

感情	ガンマ波	ベータ波	アルファ波
喜び	0.572	0.385	0.000
悲しみ	0.152	0.000	0.000
無感情	0.000	0.000	0.000

とができる可能性が示された．

6. まとめ

本研究の目的は，感情共有を 2 名同時計測された EEG から計測することである．分析の結果，感情共有が生じている時，送信者のチャンネルと受信者の後頭部に位置するチャンネルの相互相関が高まる傾向が明らかとなった．さらに，相互相関を基に算出された Global efficiency の値が高まる傾向が明らかとなった．以上のことから，EEG の相互相関と Global efficiency によって，二者間の Affecitve sharing を計測できる可能性があると考えられる．今後は，実験の試行数を増やし，統計的に有意な結果を得る必要があると考えられる．

参考文献

[Anders 11] Anders, S., Heinzle, J., Weiskopf, N., Ethofer, T., and Haynes, J.-D.: Flow of affective information between communicating brains., *NeuroImage*, Vol. 54, No. 1, pp. 439–46 (2011)

[Babiloni 14] Babiloni, F. and Astolfi, L.: Social neuroscience and hyperscanning techniques: past, present and future, *Neuroscience & Biobehavioral Reviews*, Vol. 44, pp. 76–93 (2014)

[Bradley 94] Bradley, M. M. and Lang, P. J.: Measuring emotion: The self-assessment manikin and the semantic differential, *Journal of Behavior Therapy and Experimental Psychiatry*, Vol. 25, No. 1, pp. 49–59 (1994)

[Ciaramidaro 18] Ciaramidaro, A., Toppi, J., Casper, C., Freitag, C. M., Siniatchkin, M., and Astolfi, L.: Multiple-Brain Connectivity during Third Party Punishment: An EEG Hyperscanning Study, *Scientific Reports*, Vol. 8, No. 1, p. 6822 (2018)

[David 80] David, M. H.: A multidimensional approach to individual differences in empathy, *JSAS Catalog of Selected Documents in Psychology*, No. 10, p. 85 (1980)

[Decety 04] Decety, J. and Jackson, P. L.: The functional architecture of human empathy., *Behavioral and cognitive neuroscience reviews*, Vol. 3, No. 2, pp. 71–100 (2004)

[Delorme 04] Delorme, A. and Makeig, S.: EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, *Journal of neuroscience methods*, Vol. 134, No. 1, pp. 9–21 (2004)

[Gramfort 13] Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., et al.: MEG and EEG data analysis with MNE-Python, *Frontiers in neuroscience*, Vol. 7, p. 267 (2013)

[Gramfort 14] Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Parkkonen, L., and Hämäläinen, M. S.: MNE software for processing MEG and EEG data, *Neuroimage*, Vol. 86, pp. 446–460 (2014)

[Mognon 11] Mognon, A., Jovicich, J., Bruzzone, L., and Buiatti, M.: ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features, *Psychophysiology*, Vol. 48, No. 2, pp. 229–240 (2011)

[Neumann 15] Neumann, D. L., Chan, R. C., Boyle, G. J., Wang, Y., and Westbury, H. R.: Measures of Empathy: Self-Report, Behavioral, and Neuroscientific Approaches, in *Measures of Personality and Social Psychological Constructs*, pp. 257–289, Academic Press (2015)

[Paiva 17] Paiva, A., Leite, I., Boukricha, H., and Wachsmuth, I.: Empathy in Virtual Agents and Robots, *ACM Transactions on Interactive Intelligent Systems*, Vol. 7, No. 3, pp. 11:1–11:40 (2017)

[Złotowski 16] Złotowski, J., Sumioka, H., Nishio, S., Glas, D. F., Bartneck, C., and Ishiguro, H.: Appearance of a robot affects the impact of its behaviour on perceived trustworthiness and empathy, *Paladyn, Journal of Behavioral Robotics*, Vol. 7, No. 1, pp. 7:55–66 (2016)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

プロダクトディメンジョンを評価指標とした グループ対話のアウトプット推定

Estimating performance of group interaction based on product dimension

三浦 郷 *1　岡田 将吾 *1
Go Miura　Shogo Okada

*1北陸先端科学技術大学院大学
School of information Science, Japan Advanced Institute of Science and Technology(JAIST)

This paper focuses on developing a model for estimating the quality of discussion using multimodal features. For this purpose, we use a group meeting corpus including audio signal data of participants observed in 30 meeting sessions. Also, four annotators watch conversation transcripts and annotate the score about quality of discussion using product dimension which is a sociological criteria. We extracted various kinds of features such as spoken utterances, acoustic features, speaking turns. First, binary (high or low) classification models are trained to infer the annotated score from these features using support vector machine. Second, binary classification models are developed to infer the quality of unknown discussion task. Experiments results show that multimodal model archived 0.92 as the classification accuracy and task independent model archived 0.73.

1. はじめに

意思決定を集団で行う理由の一つは，個人の情報処理能力の限界を克服することである．集団意思決定に関して，社会心理学の領域で研究が行われている．[Ueda 97] では，集団の意思決定は相互作用により認知的刺激を受けるなどメリットがある一方，多数派の同調圧力により少数派の自分の意見を表出できない場合のように，必ずしも個人の意思決定より集団の意思決定の方が良い結果を得るとは限らないと報告されている．集団の意思決定方法やプロセスの質を評価することが出来れば，意思決定方法の見直しや，個人の意思決定結果の尊重など，集団意思決定におけるデメリットを緩和できる可能性がある．本研究では，集団意思決定のための方法の一つであるグループディスカッションを題材として，ディスカッションの質を評価し，質の改善・向上を支援するシステムに応用可能な技術の提案を目指している．

本研究はディスカッションを通じて観測できる参加者個人の発言の仕方，聞き方と発言内容から，その個人が属するグループで行われるディスカッションの質を推定するモデルを機械学習により構築し，評価することを目的とする．本研究で対象とするグループディスカッション (Group Discussion: 以降 GD と略記する) の質を，[Hackman 67] における評価指標に基づき定義する．この評価指標は書き言葉に対して適応可能であり，本研究ではディスカッションの発話内容を書き起こしたテキストを評定者が閲覧し，[Hackman 67] を元に GD の質を評価している．本研究の貢献は以下の 2 点である．

1. **グループアウトプットの推定モデルの構築**
   40 名の参加者による計 9 時間を超える会話データ [林 15] から，評定者 4 名により評価された GD の質を推定するモデルを学習する．評価実験の結果から，最大で 9 割以上の精度で GD の質の高低を推定できることを示す．

2. **タスクの種類に依存しない推定モデルの構築**
   40 名の参加者は，グループごとに「インバスケット型」課題 1 つと,「ケーススタディ型」課題 2 つの課題に取り組んでいる．3 つの課題から学習する組み合わせを変えて，課題に依存せずに GD の質を推定するモデルの構築を目指す．評価実験の結果から，最大で 7 割以上の精度で訓練データに用いていないタイプの GD の質の高低を推定できることを示す．

連絡先: 三浦郷，北陸先端科学技術大学院大学，〒 923-1211
石川県能美市旭台 1 丁目 1， s1810175@jaist.ac.jp

2. 関連研究

[Avci 16] らは，マルチモーダル情報を特徴量に用いて，グループアウトプット (GD の質) を推定するモデルの構築を行った．アウトプットの高･低の 2 値分類タスクにおいて，Coupled Hidden Markov Model (CHMM) を改良した IM (Influence Model) と Support Vector Machine (SVM) を用いたモデルが最大の分類精度を示し，91%であった．[Murray 18] らは，マルチモーダル特徴量に言語に関する情報を含めて，グループアウトプットを推定する回帰予測タスクを行った．データサイズを増加させる転移学習を行い，学習器に Random Forest を用いたモデルが最良の精度を示し，MSE (平均二乗誤差) は 64.4 であった．

[Avci 16] と [Murray 18] で用いられた ELEA コーパス [Sanchez 13] には，Winter survival task と呼ばれるインバスケット型課題にグループで取り組んだ際に得られたマルチモーダルデータが記録されている．この課題はロールプレイング形式となっており，サバイバルを想定した時に必要なアイテム群（水や専門書など）の重要度を議論しランキングする．専門家によるランキング（正解）と，グループで決めたランキングの差を「グループアウトプット（GD の質)」と定義している．一方で，ディスカッションはインバスケット型課題に限らず様々な種類が存在し，正解が明確に決まっていない場合が多い．上記の正解との差によりパフォーマンスを定義する方法では，正解が明確に存在しないタスクに対して適応ができない．そこで，本研究では課題に依存しないグループパフォーマンスの定義方法を 3.1 節で提案する．この評価指標を基にアノテーションしたデータセットを用いて，課題に依存しない GD の質の自動推定の精度を検証する．

3. グループディスカッションコーパス

本研究で用いるグループディスカッションコーパス [林 15, Nihei 14] の説明を行う．以降，このコーパスを GD (Group Discussion) コーパスと呼ぶ．参加者 40 名の内，初対面同士の 4 名を 1 グループとして，合計 10 グループを構成している．各グループは 3 つの課題 [林 15] についてディスカッションに取り組み，ディスカッションの過程で得たデータ群をデータセットとして収集した．結果として，10 グループ × 3 つの課題 × 40 名で，延べ 120 名分の参加者のデータセットが GD コーパスに含まれている．GD コーパスに含む情報の内，本研究ではヘッドセットマイクから観測した音声データを用いて後述のマルチモーダル特徴量抽出を行った．

3.1 ディスカッションの質の定義

表 1: プロダクトディメンジョンの評価指標

y1	その課題への回答は，創造性を有していた．
y2	その課題に対して，適切な方向性，適切な議論，適切な回答が行えていた．
y3	製品（議論の履歴）は，過度な困難はなく，理解可能であり，明快である．
y4	製品（議論の履歴）は，何らかの問題に関しての観点や立場を取ることが出来ていない，または対処できない程度，適切な粒度で問題を捉えられず，対象となる問題の扱いに失敗している．
y5	製品（議論の履歴）は幾つかの考え方，物，事に難色を示すものである．
y6	製品（議論の履歴）には綿密さや，事例，詳細が欠如している．
y7	製品（議論の履歴）は，具体的に何か行うことに関連した方向性を持っていて，実質的な提案，または建設的な内容である．
y8	製品（議論の履歴）は，今後第三者にフォローされるべき，特定または一般的な行動方針を示したり暗示したりしている．
y9	製品（議論の履歴）は，あるアイデア，目的，または出来事に関する特定の信念や意見を述べたり，強く示唆したりしている．
y10	製品（議論の履歴）は特定の観点を押し通そうとしている．
y11	製品（議論の履歴）に含まれるアイデアはユニークで，新鮮で，あまり見たことがない驚きのあるものである．
y12	製品（議論の履歴）の一般的な見通し，態度，調子，方向性は「否定的」ではなく「肯定的」である．
y13	製品（議論の履歴）は断片化していたり，相互の発言に関連性がなかったりして，調和のとれた首尾一貫している状態と真逆である．
y14	製品（議論の履歴）はユニークで，普通ではなく，珍しい．
y15	製品（議論の履歴）は幾つかの考え方，物，事をサポートしている．
y16	製品（議論の履歴）は独りよがりであり，活性化していない．
y17	製品（議論の履歴）は普通で，ありふれていて，よく見るものである．
y18	製品（議論の履歴）内容が簡潔すぎて短い．
y19	製品（議論の履歴）のプレゼンテーションは適切な方法で行われ，調和がとれている．

本研究では課題に依存しないグループパフォーマンスの指標として，Hackman により定義された Product dimension を用いる．表 1 にその評価指標を示す．プロダクトディメンジョンの指標は書き言葉に対して適応可能であり，本研究においては GD における参加者の発話内容を書き起こしたテキストデータを 4 人の評定者が閲覧し，この評価指標に基づき各グループに評定値を付与した．この 4 人に付与された評定値の合計を各グループが行った GD の質 (機械学習における教師ラベル) として定義する．以降，19 個の評価指標をそれぞれ y1，. . . ，y19 と略記する． 評定値のスケールは，y1 が 1 点から 5 点の 5 段階，y2 から y19 は 1 点から 7 点の 7 段階で評価した．ディスカッションのサンプル数は 30 であったが，そのうち 1 つのサンプルに不備があったため，29 のサンプルを実験に用いた．

3.2 アノテーションの一致率の分析

各グループによる GD で観測された発話内容を書き起こし，4 人の評定者がその文章を閲覧して GD の質の評定を行う．評価基準として前述のプロダクトディメンジョンの指標を用いる．

各評定者による評定の一致度を計算し，各評価項目についての信頼性を確認した．一致度の計算に関して，信頼性分析の項目を評定者とし，各項目ごとにクローンバック α 値とクリッペンドルフ α 値を算出した．ある項目に対してクローンバック α が 0.7 以上，クリッペンドルフ α が 0.667 以上であれば各評定者によるグループに対しての評定値は概ね一致しており，その項目の評定値は信頼性が高いと言える．

表 2 は各評価指標におけるクローンバック α，クリッペンドルフ α を示している．各値が正の場合は四捨五入，負の場合は絶対値を四捨五入し，マイナスをつけたものを記載している．各評価指標の値は閾値より低く，十分な一致度は得られなかった．GD の質のアノテーションは主観的であるため，高い一致率が得られなかったと考えられる．この理由の考察は今後の課題とする．

4. マルチモーダル特徴量の抽出

評定者は GD を書き起こした文章を閲覧して評価しているため，GD の質は主に書き起こした発話内容に関連すると考え，発話内容の言語特徴量を抽出する．また，議論中の参加者の非言語特徴量も GD の質に影響を与えると考え，[岡田 16] を参考に発話ターン特徴量，発話中の韻律特徴量を抽出する．

4.1 発話ターン特徴量

合計発話長：発話区間から算出した発話時間を発話長とし，1 セッション単位 (各課題の開始時間から終了時間まで) で発話長の合計を計算する．

合計発話回数：一つの発話区間を発話断片と定義する．発話断片の回数を発話回数とし，1 セッション単位で発話回数の合計を計算する．

合計発話長 (1 秒以上)：1 秒以上の発話断片を抽出し，その長さの総和を計算する．

合計発話回数 (1 秒以上)：1 秒以上の発話断片を抽出し，その発話回数を計算する．

4.2 韻律特徴量

各発話断片における韻律情報の特徴量を抽出する．4.1 節で得た発話区間情報から，各発話断片の特徴量を抽出する．音声情報の信号処理と韻律特徴量抽出には Praat*1 を用いる．

最大ピッチ，最小ピッチ：各発話断片の最大ピッチ，最小ピッチを計算し，1 セッション中の全発話断片に対して平均値を計算する．

ピッチ平均：各発話断片に関して，発話区間でピッチの平均値を計算し，1 セッション中の全発話断片に対して平均を計算する．

最大インテンシティ，最小インテンシティ：各発話断片の最大，最小音圧 (インテンシティ) を計算し，1 セッション中の全発話断片に対して平均値を計算する．

音圧の幅：最大，最小インテンシティの差を計算する．

抑揚：最大ピッチと最小ピッチの差を計算する．

話速：発話のシラブル (音節) 数を各発話長で除した値を計算する．

4.3 発話内容の言語特徴量

名詞数，動詞数，感動詞数，フィラー数：1 セッション中の発話に含まれる各品詞：名詞，動詞，感動詞の単語の出現回数，フィラーの回数を計算する．

新規名詞数：4.1 節で得た発話区間情報を利用し，各名詞が初

*1 http://www.fon.hum.uva.nl/praat/

表 2: 各評価指標のクローンバック α，クリッペンドルフ α

評価指標	y1	y2	y3	y4	y5	y6	y7	y8	y9	y10	y11	y12	y13	y14	y15	y16	y17	y18	y19
クローンバック α	0.68	0.58	0.68	0.43	-0.03	0.64	0.64	0.34	0.67	0.50	0.60	0.78	0.52	0.63	0.34	0.52	0.62	0.54	0.51
クリッペンドルフ α	0.34	0.26	0.29	0.16	-0.01	0.30	0.30	0.08	0.32	0.20	0.27	0.43	0.19	0.28	0.12	0.15	0.27	0.16	0.20

表 3: 各特徴量セットを用いた場合の推定精度の結果．太字は正解率 0.7 以上を示す．

	評価指標									
	y1	y2	y3	y4	y5	y6	y7	y8	y9	y10
A	0.52	0.27	0.52	0.32	0.48	0.48	0.56	0.68	0.36	0.48
S	0.36	0.50	0.48	0.18	0.52	**0.80**	0.56	**0.76**	0.61	**0.74**
L	**0.76**	0.00	0.56	0.00	0.56	**0.72**	0.44	0.68	0.61	0.67
A + S	0.48	0.23	0.48	0.18	0.56	**0.76**	0.60	**0.88**	0.14	0.33
A + L	0.32	0.00	0.48	0.00	0.56	0.56	0.68	**0.92**	0.39	0.22
S + L	0.48	0.12	0.56	0.00	0.56	**0.72**	0.64	**0.72**	0.64	0.59
A + S + L	0.24	0.00	0.52	0.18	0.56	**0.76**	**0.76**	**0.92**	0.04	0.30

	評価指標								
	y11	y12	y13	y14	y15	y16	y17	y18	y19
A	0.52	0.54	0.44	0.23	0.48	**0.70**	0.44	0.26	0.42
S	0.59	0.50	0.52	0.58	0.36	0.29	0.56	0.67	0.35
L	**0.70**	0.54	0.56	0.61	0.24	0.00	0.68	**0.70**	0.42
A + S	0.63	0.63	0.52	0.61	0.36	0.56	0.56	0.30	0.35
A + L	0.59	0.50	0.40	0.61	0.28	0.59	0.60	0.41	0.42
S + L	**0.78**	0.50	0.56	0.46	0.20	0.26	0.52	**0.70**	0.31
A + S + L	0.59	0.63	0.52	0.65	0.20	0.37	0.60	0.41	0.42

めて発言された時間，発言した参加者を特定し，参加者ごとに数え上げた回数を新規名詞数とする．新規名詞は，他の参加者に対して新しい提案をする時に使用される傾向がある．

4.4 グループ特徴量への変換

発話ターン，韻律，言語の個人特徴量をグループ特徴量に変換する．4 人の参加者の特徴量の最大値，最小値，最大値と最小値の差分，平均値，標準偏差，合計値の計 6 つを計算して各グループの特徴量とする．変換前の個人の特徴量は前述した 17 個であり，グループ特徴量に変換後はそれぞれ 6 つずつのグループ値を計算するため計 102 個となった．

5. 評価実験

中央値を基準にグループの評定値を高群・低群に分類し，2 クラス分類タスクを行い推定精度を評価する．評価実験は実験 1 と実験 2 の 2 つ行う．実験 1 では，学習された分類モデルのマルチモーダル特徴量の組み合わせと推定精度を分析することで，GD の質を推定するために有効な特徴量を明らかにする．実験 2 では，3 つの問題解決課題ごとに分けた特徴量の組み合わせと推定精度を分析することで，未知の課題に対するモデルの推定精度を評価する．

5.1 実験手順

評価には 1 つの GD データを抜いた交差検定法を使用した．あるグループ i から得られたデータをテスト，それ以外の 28 グループから得られたデータを訓練に用いて分類モデルの評価実験を行った．

二値分類タスクの準備として，評定値を二値に変換する際，評定値のうち中央値より高いものを高群，低いものを低群とし，中央値にあたる 1 つのデータサンプル（1 グループ）は取り除いた．分類モデルの推定性能には全テストデータに対する正解率の平均を用いた．分類学習には，RBF カーネルを用いたサポートベクトルマシン (Support Vector Machine: SVM) を用いた．RBF カーネルのバンド幅 γ を [1/120, 1/119, 0.01, 0.1, 1, 5, 10]，SVM における損失とマージンの大きさの間のトレードオフを調整するパラメータである C を [0.01, 0.1, 1, 5, 10] の範囲で探索し，最適パラメータをテストに用いた．

各モダリティの推定精度への寄与を検証するため，単一モダリティとマルチモーダル特徴量セットを計 7 種類準備し，推定精度の比較を行う.「韻律」,「発話ターン」,「言語」の特徴量セットをそれぞれ A，S，L とする．また，マルチモーダル特徴量セットとして,「韻律と発話ターン」を $A + S$,「韻律と言語」を $A + L$,「発話ターンと言語」を $S + L$,「韻律と発話ターンと言語」を $A + S + L$ と表す．

5.2 2 種類の分類タスク実験

実験 1: 各種モダリティを用いた推定精度の評価

GD の質を推定するために有効な特徴量セットを明らかにするため，特徴量セットの組を変えてモデルを構築し，推定精度を比較した．特徴量セットは 5.1 節で記述した 7 つの組を用いた．

実験 2: 未知の課題に対する推定精度の評価

課題の種類に依存しない，未知の課題に対する推定精度を評価するための実験を行う．ディスカッションデータ全 29 サンプルの内，課題 1 は 9 サンプル，課題 2，3 は 10 サンプルずつ保有している．課題 1，2，3 の内，2 つの課題を訓練データに，残り 1 つの課題をテストデータに用いる．テストデータに使用する課題のデータ 10 サンプル (課題 1 は 9 サンプル) について，1 サンプルずつテストに用いて推定し，10 サンプルの推定精度の平均を算出した．

6. 評価実験結果

実験 1 の各種モダリティごとの推定結果を表 3 に示す．表 3 では，各評価指標において，単一モダリティとマルチモーダル特徴量セット 7 種類を用いて推定した時の正解率を示している．実験 2 の未知の課題に対する推定結果を表 4 に示す．表 4 では，各評価指標において，訓練に用いていない各課題のデータをテストに用いた場合の正解率と，その正解率を平均した結果を示している．

実験 1 の結果:

各種モダリティごとの推定結果について，評価指標 y8「行動方針の示唆」における「韻律 + 言語 $(A + L)$」,「韻律 + 発話ターン + 言語 $(A + S + L)$」のモデルの推定精度が最大となり，0.92 であった．評価指標 y8 において単一モダリティよりマルチモーダル特徴量を用いたモデルが有効であることが示された．また，推定精度が 0.7 以上であるモデルが存在したのは，評価指標 y1「創造性」，y6「綿密さの欠如」，y7「建設的な内容」，y8「行動方針の示唆」，y10「多面的な観点の欠

表 4: 各評価指標における推定精度の結果．太字は正解率が 0.7 以上であることを示す．各指標において，0.5 を超える正解率が存在しないものは割愛した．

	テストデータ			
	課題 1	課題 2	課題 3	平均
y1: 創造性	0.39	0.0	0.61	0.33
y6: 綿密さの欠如	0.60	0.41	0.24	0.40
y7: 建設的な内容	**0.70**	0.0	0.06	0.25
y8: 行動方針の示唆	**0.90**	0.35	**0.94**	**0.73**
y11: 驚くべき内容	0.61	0.47	0.67	0.58
y12: 肯定的な態度	**0.70**	0.00	0.00	0.23
y14: 独自性	**0.85**	0.33	0.44	0.54
y15: 意見への支持	0.00	**0.81**	0.00	0.27
y16: 独りよがりな内容	0.55	0.00	0.00	0.18
y17: ありふれた内容	0.65	0.44	0.38	0.49
y18: 情報の不足	0.39	**0.89**	0.31	0.53
y19: プレゼンの適切性	0.39	0.33	**0.71**	0.48

如」，y11「驚くべき内容」，y16「独りよがりな内容」，y18「情報の不足」であった．また，全 19 個の指標のうち 15 個の指標で，2 クラス分類タスクのチャンスレベル 0.5 を上回るモデルが存在した．

実験 2 の結果:

未知の課題に対する推定結果について，実験 1 と同様に評価指標 y8「行動方針の示唆」において精度が最大となり，0.73 であった．評価指標 y8 に関して，タスクに依存しないモデルを構築できることが示唆された．また，平均する前の課題別の正解率について，2 クラス分類のチャンスレベル 0.5 を上回ったのは，表 4 に示した 12 個の評価指標であった．その他の指標については精度が 0.5 に満たず，十分な精度には至らなかった．推定精度を上げるために今後検討を行う．

7. 考察

先行研究である [Avci 16] では，GD の質のレベルの高低を最大 0.91 の精度で分類可能であることが示された．コーパス参加者や使用言語，評価基準が異なるため，本研究と上記の先行研究の結果とを直接比較することはできないが，評定値の 2 クラス分類タスクとして同様に扱うと，最大 0.92 と同程度の精度でモデル構築可能であることが示された．しかし，正解率 0.92 を示したのは評定の一致度について信頼性が高くなかった評価指標であり，考慮に入れる必要がある．

実験 1 の結果から，プロダクトディメンジョンの評価指標 y1「創造性」，y6「綿密さの欠如」，y7「建設的な内容」，y8「行動方針の示唆」，y10「多面的な観点の欠如」，y11「驚くべき内容」，y16「独りよがりな内容」，y18「情報の不足」において，各種モダリティを用いた推定が可能であることが示された．すなわち，議論の質をプロダクトディメンジョンの評価指標に基づき定義すれば,「ケーススタディ型」課題などの正解が明確に決まっていない議論の質を特定の観点からは推定可能であることが期待できる．実験 2 の結果から，評価指標 y8「行動方針の示唆」においては課題に依存しないモデル構築が可能であることが示された．しかし，テストデータに用いた課題によって推定精度にばらつきがあること，評価指標 y8 以外ではモデルの精度が芳しくないことは今後の課題である．

今回行ったアノテーションは，4 人の評定者の主観に依るところが大きいため，評定の一致度が低くなったと考えられる．今後，4 人の評定者の中で誰の評定が大きく異なっているか詳細を分析し，アノテーションの一致率が高い評定者を選定して信頼性の高いモデルを構築することを目指す．また，評定者の人数を増やすと評定の一致率がどう変化するか傾向を分析することも今後の課題である．

8. 結論

本研究では，グループ対話の参加者から得られるマルチモーダル情報からディスカッションの質を推定するモデルの構築・評価を行った．発話ターン，韻律，言語といった情報をマルチモーダル特徴量として抽出し，プロダクトディメンジョンの評価指標に基づき評定された各グループの GD の質を推定するモデルを機械学習により構築した．評価実験の結果，評価指標 y8「行動方針の示唆」に関して，高群・低群の 2 クラス分類タスクで最大 0.92 の精度を得た．また，本実験で対象とした 3 つの課題の内，2 つの課題のデータを学習したモデルを用いて，学習に用いなかった 1 つの課題のデータに対して推定した精度の平均は，評価指標 y8 において最大精度の 0.73 を得た．

参考文献

[Avci 16] Umut Avci, et al. ”Predicting the Performance in Decision-Making Tasks: From Individual Cues to Group Interaction.” in IEEE Transactions on Multimedia, vol. 18, no. 4, pp. 643-658 (2016)

[Murray 18] Gabriel Murray, et al. ”Predicting Group Performance in Task-Based Interaction” In Proceedings of the 20th ACM International Conference on Multimodal Interaction, pp. 14-20, ACM (2018)

[Hackman 67] Hackman, J. Richard, et al. ”A set of dimensions for describing the general properties of group-generated written passages” Psychological Bulletin vol. 67, no. 6, pp. 379-390 (1967)

[Sanchez 13] Sanchez-Cortes, Dairazalia, et al. ”Emergent leaders through looking and speaking: from audio-visual data to multimodal recognition” Journal on Multimodal User Interfaces, vol. 7, no. 1-2, pp. 39-53 (2013)

[林 15] 林佑樹, 他 ”グループディスカッションの構築および性格特性との関連性の分析” 情報処理学会論文誌, Vol. 56, No. 4, pp. 1217-1227 (2015)

[岡田 16] 岡田将吾, 他 ”マルチモーダル情報に基づくグループ会話におけるコミュニケーション能力の推定” 人工知能学会論文誌, vol. 31, no. 6, pp. AI30-E_1-12 (2016)

[Nihei 14] Fumio Nihei, et al. ”Predicting Influential Statements in Group Discussions using Speech and Head Motion Information” In Proceedings of the 16th ACM International Conference on Multimodal Interaction, pp. 136-143, ACM (2014)

[Ueda 97] 上田泰，”個人と集団の意思決定-人間の情報処理と判断ヒューリスティックス-” 文眞堂 (1997)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

マルチモーダル情報に基づくユーザの興味度推定のための対話履歴の効果の検証

Use of Dialogue History for User's Interest Level Estimation Based on Multimodal Information

西本 遥人 Haruto Nishimoto　武田 龍 Ryu Takeda　駒谷 和範 Kazunori Komatani

大阪大学 産業科学研究所
The Institute of Scientific and Industrial Research (ISIR), Osaka University

Estimation of a user's mental states including his/her interest level for the current topic is a key component to realize user-adaptive multimodal dialogue systems. Multimodal sensing provides helpful information for that, but its results are often incorrect because the user's subtle behaviors need to be captured. We incorporate a dialogue history when estimating the user's interest level to make the results more stable. Specifically, we formulate the estimation by using the partially observable Markov decision process (POMDP). The multimodal sensing results are used as likelihoods of input observation and the estimated interest levels are tracked as states. Experimental results showed the estimation performance by considering the dialogue history. We used two kinds of annotation results for each exchange: the interest levels and the next desirable system actions, which were given by multiple annotators. We also confirmed a correlation between the estimated interest levels and the next desirable system actions.

1. はじめに

対話における人の発話には，その言語内容だけでなく，様々な有用な情報が含まれる．対話分析においては，声の韻律情報，顔の表情などの非言語情報が重要になる [1]．本研究では，人対人ではなく，人対システムの対話において，ユーザの心的状態の一種である興味度をマルチモーダル情報を用いて推定する．そしてその推定結果を考慮した応答生成が可能なシステム構築を目標としている．

ユーザの興味を反映してシステムが応答する場合，ユーザのシステムに対する印象がよくなると考えられる．例えば，ユーザが興味を持たない場合，システムは対話を別の話題に切り替えることができる．一方興味を持つ場合，対話中の話題を掘り下げたり，その話題に関連する情報を提示できる．本研究では，図 1 に示すようにシステムとユーザの発話対（一交換）ごとに興味度を推定することで，その推定結果に応じて適応的に発話を変更するシステムの構築を目指している．

対話中のユーザの興味度が発話ごとに大きく変化するとは考えにくいため，対話中の過去の状態を考慮することは有用である．例えば，興味ありの状態が長く続いているユーザに対し，システムがこれまでと同じ話題を継続する発話をすると，ユーザの興味度が急に下降するとは考えにくい．このため，対話履歴を考慮して，発話ごとの推定性能の向上を狙う．マルチモーダル情報を用いた興味度の推定は，ユーザの微妙な特徴を捉える必要があるため，各発話の情報だけから必ずしも正しい推定結果が得られるわけではない．実際，我々の以前の報告 [2] では，各発話対から得られる情報のみを用いて推定を行ったところ，7 割程度の正解率であった．システムが誤った推定結果を用いて次発話を行った場合，ユーザは期待しない発話を受け取ることになり，ユーザのシステムへの印象は悪化する．つまり，対話において，一発話だけからの興味度推定結果を用いて応答を行うのは適切ではない．

連絡先: 西本 遥人，大阪大学 産業科学研究所 駒谷研究室，〒567-0047 大阪府茨木市美穂ヶ丘 8-1，TEL:06-6879-8416，nishimoto@ei.sanken.osaka-u.ac.jp

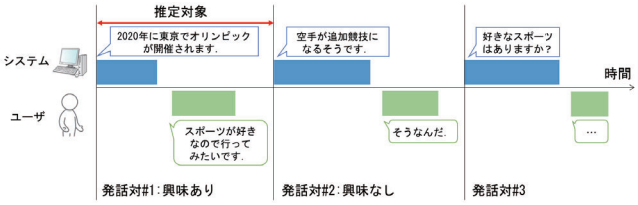


図 1: 興味の推定の対象区間

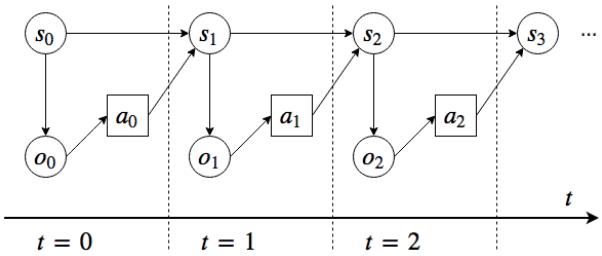


図 2: POMDP の変数の依存関係

これらの課題を解決するために，本稿では Partially Observable Markov Decision Process（部分観測マルコフ決定過程，POMDP）を用いたシステム構築について述べる．その効果を検証し，興味度の推定の重要性，興味度とシステムが次発話でとるべき行動の関連性を述べる．

2. 対話履歴を用いたユーザの興味度推定

2.1 対話履歴を用いる枠組み

POMDP とは，変化するユーザの状態が直接観測できない環境において，その状態を確率的に表し，連続的に表現する手法である [3]．図 2 に POMDP で用いる変数の依存関係を示す．ここで，状態 s_i，o_i，および a_i は，時刻 i における状態，観測データとシステム行動を示す．

本研究では内部状態 s_i として，ユーザのシステムへの興味度を設定し，POMDP を適用する．s_i はユーザの興味，o_i は

システムが観測したマルチモーダル特徴量，a_i はシステム発話とみなす．この設計で，対話という時系列データにおける興味度の変化を追跡する．まず，システムがユーザの発話から特徴量抽出を行い，それをもとにシステムはユーザの興味度を更新する．そして，その興味度をもとに，システムの行動を決定しユーザ状態を更新する．

以下の式 (1) に基づいて信念 $b(s)$ を更新する．式中の「′」は次時刻を表し，「′」なしは現時刻を表す．

$$\begin{aligned} b'(s') &= p(s'|o', a, b) \\ &= \frac{1}{z} p(o'|s', a) \sum_{s \in S} p(s'|a, s) b(s) \end{aligned} \tag{1}$$

$b(s)$ は信念であり，ある時刻において状態が s である確率を表す．$p(o'|s', a)$ は s の尤度関数，$p(s'|a, s)$ は s の状態遷移確率，$\frac{1}{z}$ は正規化係数である．POMDP では，現在の信念 $b(s)$ に基づき，システム行動が式 (2) 政策関数 π によって決定される．

$$a' = \pi(b'(s')) \tag{2}$$

2.2 信念の更新式の設計

状態変数 s は興味あり/なしを意味する s_1/s_2 の 2 値を取るとする．まず，式 (1) の尤度関数 $p(o'|s', a)$ を，簡単のために a を除いた $p(o'|s')$ として考える．$p(o'|s')$ を定式化するのは困難なため，ベイズの定理により以下のように変形する．

$$p(o'|s') = \frac{p(s'|o')}{p(s')} \cdot \sum_{s'} p(o'|s') p(s') \tag{3}$$

ここで $\sum_{s'} p(o'|s')p(s')$ は s' にかかわらず一定となるので，

$$p(o'|s', a) \propto \frac{p(s'|o')}{p(s')} \tag{4}$$

という関係が得られる．$p(s'|o')$ は特徴量が得られたときの s' の事後確率である．これを現時刻の発話対から推定した興味度で近似する．$p(s')$ は s' の事前確率である．次に，状態遷移確率 $p(s'|a, s)$ を考える．簡単のために，$p(s'|a, s)$ を a を除いた $p(s'|s)$ とみなす．具体的な設定は 4.1 節で述べる．

最後に，対話中にシステムが話題を変更する場合，式 (1) 中の事前確率 $\sum_{s \in S} p(s'|a, s)b(s)$ について，状態 s を等確率とする規則を追加する．これは，話題の変更時には前発話の興味度は必ずしも現発話に反映されず，ユーザの興味度は中間的な値に戻ると考えられるからである．

3. 対象データと単一発話の興味度推定

3.1 対象データ

対象としたデータは，人工知能学会 言語・音声理解と対話処理研究会に設置されたワーキンググループにより収集された，マルチモーダル対話コーパス [4],[5] である．コーパスの対話内容はシステムが決められた話題に関する質問や情報提供を行うものである．コーパスは合計 39 名の対話が収録されており，合計 3209 のシステム-ユーザ発話対が含まれる．その各発話対ごとに，3.2.1 節で述べる興味ラベルが付与されている．また今回新たに，このうちの 29 名分 [5] に対して，3.2.2 節に示す話題継続ラベルを付与した．

表 1: L_{int} の内訳

L_{int} の値	< 0.5	$= 0.5$	> 0.5	all
数	1326	292	1591	3209

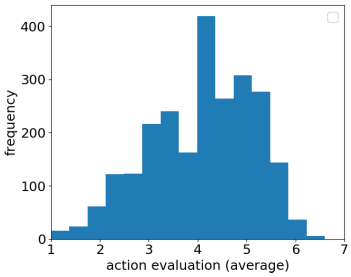


図 3: 5 名のアノテータの話題継続ラベルの平均

3.2 アノテーションと正解

3.2.1 興味ラベル

使用する対話コーパス [4],[5] から，興味ラベルの作成を行う．複数のアノテータによって発話対ごとに「興味あり」「不明」「興味なし」とラベル付けされている．それらのラベルの票数をもとに，対象発話対での興味度を式 (5) を用いて 0 から 1 の連続値 L_{int} に変換し，これを人手で付与された正解とする．複数人のアノテータの意見の一致割合をその発話対でのユーザの興味度とする．ユーザの興味度を 0 から 1 の連続値で表現することで，L_{int} を興味ありである確率とみなすことができる．

$$L_{int} = \frac{1}{2}(1 + \frac{N_o - N_x}{N_a}) \tag{5}$$

N_a はアノテータの人数 ($N_a = 3, 6$)，N_o, N_x はそれぞれ興味あり，興味なしラベルを付けた人数である．表 1 に L_{int} の内訳を示す．

3.2.2 話題継続ラベル

対象データのうち 29 名分 2402 発話対 [5] に対して，話題継続ラベルを新たに付与した．5 名のアノテータに対して，発話対ごとに,「自分がシステム役（聞き役）だったとした場合，ユーザ発話の後に話題を変えようと思うかどうか」を 7 段階で付与してもらった．この 7 段階は，1 が「話題を変える」，4 が「どちらとも言えない」，7 が「この話題を深く尋ねる」である．このラベルを用いて，システムが次に取るべき行動の正解データを作成した．具体的には，5 名が付与したラベルの値を発話対ごとに平均し L_{act} とした．L_{act} のヒストグラムを図 3 に示す．

3.3 単一発話の興味度推定

興味度の推定には対話情報，韻律情報，発話内容，顔画像情報の 4 つのモーダル情報と L_{int} を用いる．マルチモーダル情報を用いる理由は，推定を行う際にユーザの複数の情報を用いると，統合推定結果を修正でき，より信頼できる結果を得られるからである [6]．以下に 4 つのモーダルから抽出した特徴量を示す．

【対話情報】対話行為，発話語数，発話語数の差，応答時間

【韻律情報】openSMILE[7] で抽出される音響特徴から，特徴選択により特徴量を 10 に削減 *1

*1 scikit-learn の selelectKbest を利用．

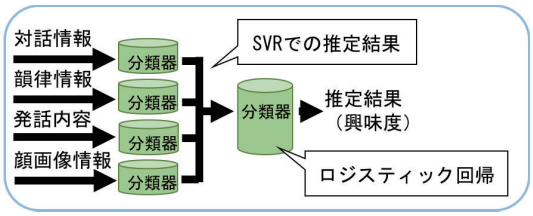


図 4: 興味推定の概要

表 2: ユニモーダルとマルチモーダルでの推定性能の比較

対話	韻律	発話	顔画像	RMSE
○	-	-	-	0.419
-	○	-	-	0.387
-	-	○	-	0.407
-	-	-	○	0.402
○	○	○	○	0.319

【発話内容】語彙（ユーザ発話中の単語から bag-of-words で作成，主成分分析で次元数を 5 に圧縮.），品詞の数，無言の有無

【顔画像情報】人の顔の動きを記述する Action Unit (AU)[8] を使用．17 の AU の出現の有無，18 の AU の強度に関する値

[2] をもとに，単一発話のみを用いた興味度の推定を図 4 のように行った．まず，4 つのマルチモーダル情報をそれぞれ単独で用いて SVR を行った．次に，それらの回帰の結果得られる値を説明変数として，ロジスティック回帰を行い，マルチモーダル情報を統合した．学習のコスト関数にはクロスエントロピーを用いた．推定性能の評価は，未知のデータに対する性能を測るために，ユーザごとに分割した交差検証により行った．

ユニモーダルでの推定結果，統合推定結果を表 2 に示す．評価指標には L_{int} と推定結果の平均平方二乗誤差 (RMSE) を用いた．各モーダル情報を統合した結果，RMSE が他に比べ低い値となっており，推定性能の上昇が確認できる．

4. 評価実験

4.1 対話履歴を考慮したときの性能評価実験

POMDP の設計に関して，使用コーパス [4],[5] の全発話対をもとに，式 (1) 中の尤度関数，状態遷移確率を定めた．尤度関数 $p(o'|s')$ は式 (4) より，$p(s'|o')$ と $p(s')$ に分けられる．前者を単一発話の興味度の推定結果 $r = p(s'|o')$ とした．後者をコーパス中の興味あり/なしの発話対の割合とし，$(p(s_1), p(s_2)) = (0.413, 0.587)$ とした．状態遷移確率 $p(s'|s)$ はコーパスから状態遷移した発話対の組の割合を求め，以下と設計した．$p_{s_i \to s_j}$ は，s_i から s_j へ状態遷移する確率である．

$$p(s'|s) = \begin{bmatrix} p_{s_1 \to s_1} & p_{s_2 \to s_1} \\ p_{s_1 \to s_2} & p_{s_2 \to s_2} \end{bmatrix} = \begin{bmatrix} 0.699 & 0.301 \\ 0.213 & 0.787 \end{bmatrix} \quad (6)$$

一方，対話履歴を用いた手法（以降，POMDP）と比較して，対話履歴を用いない手法をベースラインとする（以降，BASE）．BASE は式 (7) で定義した．

$$\begin{bmatrix} b'(s_1) \\ b'(s_2) \end{bmatrix} = \begin{bmatrix} r \\ 1-r \end{bmatrix} \quad (7)$$

表 3: 推定した興味度と L_{int} の誤差 (RMSE)

POMDP	0.316
BASE	0.319

表 4: L_{act} と興味度の相関係数

POMDP	0.566
BASE	0.545
L_{int}	0.788

POMDP によって更新された興味度の推定値と，BASE での興味度の推定値を比較する．図 5 はあるユーザの対話（全 82 発話対）における POMDP と BASE での $b(s_1)$ の遷移である．横軸が発話対番号，縦軸がそのときの $b(s_1)$ の値である．点線はそれぞれ POMDP，BASE の結果である．黒の点線は，興味あり/なしの中間の状態 ($b(s_1) = 0.5$) を示したものである．赤の点線は，対話において話題が変化した箇所を示している．点線と実線が似た形の折れ線であるほど，モデルによる推定値が人手で付与された値に対応しており，性能がよいと言える．2 つの点線の相違箇所が，POMDP と BASE の差分である．

POMDP を用いることで，推定を誤った場合の改善がなされていた箇所があることが確認できた．図 5 の範囲 A の L_{int} は時間の経過とともに，ユーザの興味度合いが上昇している．範囲 A の途中で，BASE では興味度が低いと誤推定され，折れ線が下降しているのがわかる．一方，POMDP は興味度がこれまでより低いという観測情報 o を得たが，BASE よりも下降の割合は小さく L_{int} により近い値であることが確認できる．これは，前発話での情報（図 5 中の A の場合，「興味あり」）を保持して推定がなされたからだと考えられる．BASE と POMDP それぞれと，L_{int} との RMSE を表 3 に示す．ここからも，僅かではあるが POMDP の方がより L_{int} に近い結果が得られたことがわかる．

一方，範囲 B では POMDP での結果が L_{int} から大きく外れていることがわかる．$b(s_1)$ の増加傾向に関しては，ある程度 L_{int} と同じ傾向がある．誤差の大きい推定をした後の数発話対は，正しく推定することが困難であると考えられる．これに対しては，推定性能の向上や，信頼できない推定結果は反映しないといった規則の作成もしくはシステム行動 a を用いた $p(s'|a, s)$ の設計などで改善が見込まれ，今後の課題である．

4.2 興味度と適切なシステム行動の関係

ここでは，推定した興味度，人手で付与された興味度それぞれと，新たに作成した L_{act} の相関を調べた．その結果を表 4 と図 6 に示す．表 4 から，L_{int} と L_{act} は強い相関があることがわかる．言い換えると，人は対話相手が興味を持っていそうな場合，話題を深く掘り下げるべきだと感じる傾向があることが確認された．表 4 の BASE よりも，POMDP が L_{int} の相関係数に近い値となっている．これは，POMDP が人手で付与した正解により近い興味度推定結果を出力することを示している．図 6 は，L_{int} と L_{act} の相関を示したものである．図 6 中の直線は回帰直線であり，正の相関があることが確認できる．以上より，システムの行動をユーザの興味をもとに決定する枠組みは適切であると言える．

図 7 は POMDP での興味度の推定結果と L_{act} の相関を示したものである．表 4 から分かる通り，POMDP は L_{int} ほどではないが，L_{act} との相関がある．図 6 と図 7 を比較する

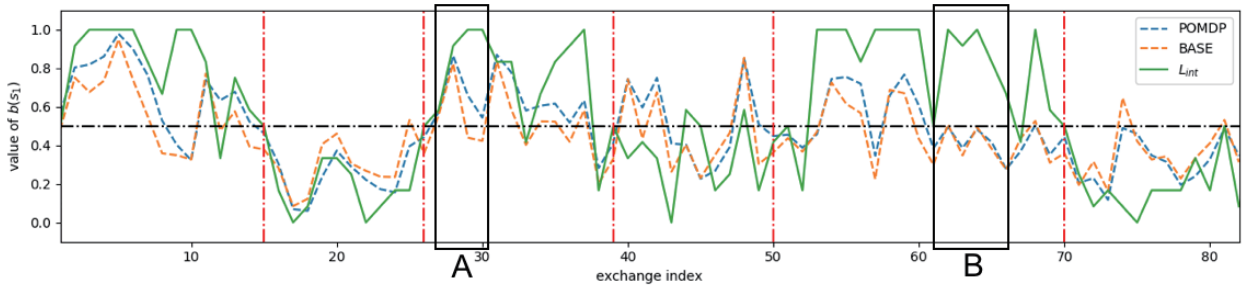


図 5: 推定した興味度の比較の例 (userID：F2006)

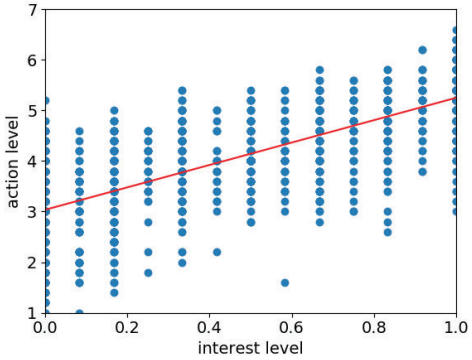


図 6: L_{int} と L_{act} の相関

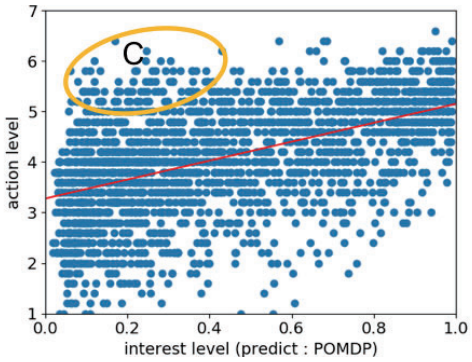


図 7: 対話履歴を用いた興味度の推定結果と L_{act} の相関

と，POMDP では話題を深く尋ねるべきシステム発話に対して，ユーザは興味がないという推定をした結果が多く見られる（図中 C の範囲）．これは信念 $b(s)$ の更新式中の状態遷移確率を，より興味なしに遷移するように設計（式 (6) を参照）したためと考えられる．本稿では状態遷移確率をコーパスの遷移の割合をもとに設計したが，今後の改善の余地があると考えられる．

5. 今後の課題

本稿では近似を用いて POMDP を実装したため，課題は多く存在している．2.2 節で述べた手法で，信念 $b(s)$ の更新式に含まれる状態遷移確率と尤度がシステムの行動 a とは独立であると仮定している．a を反映させた更新式の設計が今後の課題である．

今後は，$b(s)$ をもとにシステム行動を決める政策関数 π の学習に取り組むことも課題の一つである．政策関数 π を最適化するための手法として，(s_t, s_{t+1}, a_t) のペアから報酬を設計し，それを用いた強化学習が考えられる．報酬の設計には，目標の対話システムの評価尺度の定義を行う必要がある．評価尺度は「ユーザに興味度が高いまま持続させる」「対話を長く続ける」「話題変更をしたときに興味度がどう変化したか」などが考えられる．それを定義した上で，報酬の設計に取り組む必要がある．

参考文献

[1] M. L. Knapp, J. A. Hall, and T. G. Horgan. Nonverbal communication in human interaction. *Cengage Learning*, 2013.

[2] 西本遥人, 武田龍, 駒谷和範. マルチモーダル対話における興味の有無の推定と追加コーパスを用いた性能評価. *SIG-SLUD*, Vol. 5, No. 02, pp. 72–73, 2018.

[3] J. D. Williams and S. Young. Partially observable Markov decision processes for spoken dialog systems. *Computer Speech & Language*, Vol. 21, No. 2, pp. 393–422, 2007.

[4] 荒木雅弘, 冨増紗也華, 中野幹生, 駒谷和範, 岡田将吾, 藤江真也, 杉山弘晃. マルチモーダル対話データの収集と興味判定アノテーションの分析. *SIG-SLUD*, Vol. B5, No. 02, pp. 20–25, 2017.

[5] 駒谷和範, 岡田将吾, 西本遥人, 荒木雅弘, 中野幹生. 配布可能なマルチモーダル対話データの収集とアノテーション不一致傾向の分析. *SIG-SLUD*, Vol. 5, No. 02, pp. 45–50, 2018.

[6] C. Busso, Z. Deng, S. Yildirim, M. Bulut, C. M. Lee, et al. Analysis of emotion recognition using facial expressions, speech and multimodal information. In *Proc. ICMI*, pp. 205–211, 2004.

[7] B. Schuller, S. Steidl, and A. Batliner. The interspeech 2009 emotion challenge. In *Proc. Interspeech*.

[8] P. Ekman and W. V. Friesen. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press, 1978.

Organized Session | Organized Session | [OS] OS-11

[4F3-OS-11b] 社会的信号処理と AI(2)

岡田 将吾（北陸先端科学技術大学院大学）、石井 亮（NTT）

Fri. Jun 7, 2019 2:00 PM - 3:00 PM Room F (302B Medium meeting room)

[4F3-OS-11b-01] Considering Idling Motion Generation using RNNs for Virtual Conversational Agents

〇Hung-Hsuan Huang[1,2], Masato Fukuda[1], Toyoaki Nishida[1,2] （1. RIKEN Center for Advanced Intelligence, 2. Graduate School of Informatics, Kyoto University）

2:00 PM - 2:20 PM

[4F3-OS-11b-02] Identifying Discourse Boundaries in Group Discussions using Multimodal Features

Ken Tomiyama[1], 〇Fumio Nihei[1], Yutaka Takase[2], Yukiko Nakano[2] （1. Graduate School of Science and Technology, Seikei University, 2. Faculty of Science and Technology, Seikei University）

2:20 PM - 2:40 PM

[4F3-OS-11b-03] A corpus-based typology of reactions to apologies in Japanese}

〇Masashi Inoue[1], Kaori Suzuki[1] （1. Tohoku Institute of Technology）

2:40 PM - 3:00 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

回帰型ニューラルネットを用いた会話エージェントのアイドリング動作生成の検討

Considering Idling Motion Generation using RNNs for Virtual Conversational Agents

黄 宏軒 *1*2 Hung-Hsuan Huang　福田匡人 *1 Masato Fukuda　西田豊明 *1*2 Toyoaki Nishida

*1国立研究開発法人理化学研究所革新知能統合研究センター
RIKEN Center for Advanced Intelligence

*2京都大学大学院情報学研究科
Graduate School of Informatics, Kyoto University

This work aims to develop a model to generate fine grained and reactive non-verbal behaviors of the virtual character when the human user is talking to it. The target *micro* non-verbal idling behaviors are micro facial expression, head movements, and postures. We explored the use of recurrent neural network (RNN) to learn these behaviors in reacting to the human communication interlocutor's corresponding micro non-verbal behaviors. The models are trained on an active listening data corpus which features elderly speakers talking with young active listeners and was collected by ourselves.

1. はじめに

擬人化会話エージェントはコミュニケーションロボットのような物理的制約がなく，高い自由度で精緻でリアルな顔の表情や身体の仕草を表出でき，ユーザとコミュニケーションが取れるインターフェイスとして期待されている．しかし，これまでの会話エージェントの多くは事前に定義された顔表情や体の動きのアニメーションシーケンスをインタラクション中にルールベースで呼び出して再生する仕組みになっていた [Huang 11]．そのため，エージェントが同じことをする際に，例えば，笑うと必ず同じ笑いをしたり，視線をそらすと，同じ視線の動きをしたりすることになる．また，エージェントが何らかの意図を表出しようとしない際には，静止する，または，固定のアイドリングモーションを繰り返し再生するようになっている．こうしたエージェントとしばらくインタラクションをしていると，その固定パターンが見えて自然さを損なってしまう．

しかし，生身の人間は固定パターンの動作をしない。笑うたびに毎回少し違う笑いをするはずであり，微細なブレが含まれている．そのブレは物理的な要因以外に会話相手，文脈，本人の内的状態などの影響を受けると思われる．本研究はそのブレを捉えようとするものである．エージェントが何らかの意図を表出しようとしない間，いわゆるアイドリングの状態でも，リアルタイムで会話相手の振る舞いから微細なアイドリングモーションの生成を目指している．

近年，深層学習が様々な分野で著しい成果を挙げており，普及が急速に進んでいる．そのなかでは，データの経時的変化を捉えるように再帰型ニューラル (Recurrent Neural Networks, RNN) が考案されている．本研究では，エージェントの振る舞い生成モデルの検討として，エージェントの会話相手の振る舞い (頭部の回転，視線，表情，声の韻律情報) からエージェントの振る舞い (頭部の動き，表情，上半身の動き) を予測することを課題とし，我々が収集した傾聴会話データコーパスを基に，異なるネットワーク構造の予測性能について考察する．RNN の代表的な構造である Long Short-term Memory (LSTM) [Hochreiter 97] と忘却ゲートや出力ゲートを省きより簡素な構造になっている Gated Recurrent Unit (GRU)[Cho 14] と，ベースラインのマルチレイヤパーセプトロン (MLP) を比較対象とする．

連絡先: 黄 宏軒，hhhuang@acm.org

2. データコーパス

本研究の目的は利用者の年齢層を限定しないが，実装する会話エージェントが高齢者のメンタル支援での活用が視野に入るため，会話データコーパス収集実験は高齢者を対象とした傾聴会話課題で行った．実験参加者は高齢者 4 名 (男性 2 名，女性 2 名．年齢 69〜73 歳，平均 71 歳) と大学生 4 名 (男性 2 名，女性 2 名，平均年齢 21 歳) であり，参加者毎の組み合わせで 15〜30 分間の 4x4 計 16 会話を収録した．エージェントとの会話は画面越しになるのと高齢の参加者は遠隔地からだったため，実験は Skype でのテレビ会話であった．Skype 通信用の Web カメラの他，フル HD のビデオカメラで参加者の上半身の映像を正面から記録した (図 1)．高齢者が話し手役で若者が聴き手役の会話であったが，話題は限定せず初対面の設定で自由に会話してもらった．結果として若者が質問をして高齢者がそれに答える場面が多く見られ，話題は双方の自己紹介のほか，高齢者の過去の仕事，趣味，最近はまったことなどであった．

わずかであったがネットワーク通信の遅延があったため，分析データは実時間のデータではなく，高齢者側で収録した若者の音声を基準に若者側の映像を同期させた．その上で高齢者と若者それぞれ直接撮影したビデオカメラの映像と音声を抽出した．映像に対して，OpenFace*1 を用いて Ekman が提唱しており，顔表情を記述する Facial Action Coding System (FACS)[Ekman 02] の Action Unit (AU)44 個のうち 17 個，そして，頭部の回転 (3 次元)，視線方向 (8 次元) の 30 fps データを抽出した．さらに，OpenPose *2 を用いて肩関節の 2 次元座標からカメラまでの距離と肩の傾きを抽出した (30 fps)．音声については OpenSmile *3 を用いて Interspeech 2009 の

*1 https://github.com/TadasBaltrusaitis/OpenFace
*2 https://github.com/CMU-Perceptual-Computing-Lab/openpose
*3 https://www.audeering.com/what-we-do/opensmile/

図 1: 傾聴会話コーパスの収録映像．自然な会話に近づけるように大画面のテレビとスクリーンに映像を映し，視線が合うように位置を調整した

Emotion Challenge[Schuller 09] で定められたローレベル特徴量，energy (1 次元), mfcc (12 次元), zero-crossing rate of time signal (1 次元)，voicing probability (1 次元)，F0(1 次元) の計 16 次元 100 fps のデータを抽出した．聴き手の振る舞いは発話状態によって変わると思われるため，Elan[?] のツールを用いて聴き手の発話区間を割り出した．

話し手の振る舞い (OpenFace で抽出した映像情報 28 次元と OpenSmile で抽出した韻律情報 16 次元) を説明変数とし，聴き手の振る舞い (OpenFace の 28 次元と OpenSmile の 16 次元と OpenPose の 2 次元) を目的変数と設定した．計算が重くてリアルタイムでの入力が困難であったため，OpenPose の情報を説明変数にしていなかった．会話セッション長にばらつきがあったため，セッションの最初から最長 20 分間の部分を切り出した．入力データに 1 秒間と 2 秒間のウィンドウ幅を持たせ，出力は 1 フレーム分になるようにデータを整形した．整形した約 4.5 時間分のデータセットの概要を表 1 に示す．このデータから，男性聴き手の発話時間が短く，話し手の話す時間が長くなっており，女性聴き手自身の発話が長く，話し手の話す時間が相対的に短ったことが分かる．実験時に，男性聴き手は一方的に質問を聞いており相手の回答から話題を掘り下げない傾向に対して，女性聴き手の方は聞き上手で全体的に会話が弾んでいる観察と一致している．

表 1: データセットの概要．データ量は 2 秒時間幅のもの

	発話状態	フレーム	時間 (s)	データ量
男性	On	69,818	2,334	1.4GB
	Off	169,387	5,663	13.6GB
女性	On	82,261	2,750	1.7GB
	Off	147,165	4,920	11.8GB

3. 学習モデルの比較

3.1 実験手順

前節のデータセットの全体，男性聴き手，女性聴き手，そして聴き手個人のサブデータセットに対して，MLP, GRU, LSM の 3 つのモデルと時間幅 1 秒，2 秒の条件で機械学習実験を行った．ネットワーク構造については入力モダリティの fps 数が違うため，各モダリティの時系列をそれぞれに対応する LSTM/GRU 層 (入力の 8 倍、それぞれ 28 x 8, 16 x 8 個のノード) に入力して、そのあと全モダリティを結合する全結合層 (2 x 128 個のノード)，そして，さらに 1 層の全結合層 (512 個のノード) の出力層のシンプルな 4 層構成で統一している (図 2)．入出力データはすべて 0 から 1 までの区間に正規化した．なお，MLP については，時系列表現ができないため，同等のデータ入力を横方向 1 次元に展開して扱う．

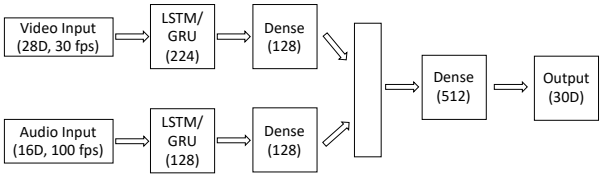


図 2: 評価実験に用いたネットワークの構造

学習モデルの評価には平均二乗誤差の平方根 (Root Mean Squared Error, RMSE) を用いる．各条件で一人の話し手をテストデータとして抽出して残り 3 人のデータを用いて学習し，これを 4 回繰り返して平均値を求める leave-one-subject out 法で交差検証を行った．

3.2 結果の考察

表 2 から，どのデータセットを用いても GRU が最も高い精度 (回帰誤差を示す RMSE 値が低い) を得ており，LSTM がそれに次いで，MLP の精度が最も低い結果を示す．このことから，同等の入力データに対してやはり時系列に特化したモデルの方は話し手と聴き手のインタラクションの特徴を捉えられている．そのなかでも今回のデータセットに対してよりシンプルな構造の GRU の方は性能が高い点，2 秒間の入力時間幅よりも 1 秒間幅の方は性能が高いことから一瞬一瞬の顔表情は反射的な要因が強いことを示唆される．また，LSTM モデルにおいても GRU モデルにおいても，男性聴き手の両名とも長い時間幅の方はより精度が高く，女性聴き手の両名とも短い時間幅の方は精度が高いことも興味深い．

4. 結論

本研究は仮想会話エージェントがアイドリング状態における意図の含まないモーションを生成するために，傾聴会話コーパスを基に話し手の振る舞いの映像と音声情報から聴き手の振る舞いを回帰的ニューラルネットの LSTM，GRU とベースラインの MLP で予測する試みをした．0〜1 の値域のデータに対して，どのモデルも 0.15 程度以下の精度を出力しており，会話パートナー同士の心理状態や話の内容を問わず，単純に相手の顔・声から顔表情と上半身のポスチャーをある程度予測できている結果を示した．しかし，この結果はあくまで数値データの観点での精度であり，生成されたエージェントの動きに対して，人間らしさを感じるかどうかの主観評価も今後実施する必要がある．

参考文献

[Cho 14] Cho, K., Merrienboer, van B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y.: Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation, *CoRR*, Vol. abs/1406.1078, (2014)

[Ekman 02] Ekman, P., Friesen, W. V., and Hager, J. C.: Facial Action Coding System (FACS), Website (2002)

表 2: 機械学習の結果

	時間幅	全員	男性	女性	全員平均	男性平均	女性平均
LSTM	1 sec	0.1335	0.1439	0.1219	0.1219	0.1356	0.1081
LSTM	2 sec	0.1416	0.1415	0.1343	0.1271	0.1357	0.1185
GRU	1 sec	0.1299	0.1369	0.1187	0.1174	0.1282	0.1065
GRU	2 sec	0.1323	0.1343	0.1278	0.1207	0.1264	0.1150
MLP	1 sec	0.1377	0.1513	0.1263	0.1278	0.1393	0.1162
MLP	2 sec	0.1514	0.1525	0.1442	0.1354	0.1406	0.1303

[Hochreiter 97] Hochreiter, S. and Schmidhuber, J.: Long Short-term Memory, *Neural Computation*, Vol. 9, No. 8, pp. 1735–1780 (1997)

[Huang 11] Huang, L., Morency, L.-P., and Gratch, J.: Virtual Rapport 2.0, in *11th International Conference on Intelligent Virtual Agents (IVA 2011)*, pp. 68–79 (2011)

[Schuller 09] Schuller, B., Steidl, S., and Batliner, A.: The INTERSPEECH 2009 Emotion Challenge, in *10th Annual Conference of the International Speech Communication Association (Interspeech 2009)*, Brighton, United Kingdom (2009)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

マルチモーダル特徴量を用いた談話セグメントの検出

Identifying Discourse Boundaries in Group Discussions using Multimodal Features

冨山 健[*1] Ken Tomiyama　二瓶 芙巳雄[*1] Fumio Nihei　高瀬 裕[*2] Yutaka Takase　中野 有紀子[*2] Yukiko Nakano

[*1] 成蹊大学大学院理工学研究科 Graduate School of Science and Technology, Seikei University

[*2] 成蹊大学理工学部 Faculty of Science and Technology, Seikei University

This study proposes models for detecting conversation boundaries in group discussions. First, we created a multimodal embedding space using an autoencoder, and applied a similarity-based approach to detect the discussion boundary. As the second method, we annotated conversation boundaries and created unimodal CNN models for language, audio, and head motion information. Then, created multimodal models by concatenating the output of unimodal models. In the evaluation experiment, we found that language information was the most useful modality, but by combining with audio and head motion modalities, the CNN-based models more accurately predict the conversation boundaries.

1. はじめに

自然言語処理の分野では，テキストを複数の意味的まとまりに分割することをテキストセグメンテーションという．テキストを分割することは文章の要約や検索などに有益であると考えられる．議論においても，セグメントに分割することができれば，話題構造の認識やファシリテーション支援にも応用できる．例えば，セグメントが短ければ，そのトピックについて議論が十分でない可能性があり，ファシリテータの介入決定の情報となりうる．

文章のセグメンテーション手法は，これまでに多くの研究が行われてきたが，高い精度をもつ手法は未だ確立されておらず，また，対話文に適用した研究は少ない．多くの研究では語彙結束性をセグメント境界の識別モデルの作成などに使用しているが，対話においてはそのスパース性が問題になってくるなど，難しい課題である．

そこで本研究では，議論参加者の言語・非言語情報に基づいた，会話のセグメント境界を検出するモデルを提案する．

2. 関連研究

[1]では語彙結束性をとらえ，文章を分割する手法 TextTiling を提案している．TextTiling は，文章に 2 つの連続する窓をスライドさせ，類似度を算出することで境界を検出する手法である．これを基にした手法も多く提案されていて，広く知られた手法である．また，[2]は新聞や TV ニュースの書き起こしを用いて，手掛かり語彙や前後に出現する語彙情報などから識別関数を作成している．[3]では複数人対話コーパスのテキスト，音声情報を用い，手掛かり語彙，沈黙時間，オーバーラップ時間などを基にしたトピックシフト識別モデルを作成している．

3. 議論コーパス

3.1 議論収集実験

本研究で使用する議論コーパスについて述べる．

1 グループ 3 人 9 グループ，男女 27 人の実験参加者が議論を行った．課題は外国人へ向けた一日観光コースをテーマに 30 分間話し合い，その後 5 分間でアピール点を決定してもらうというものである．最終的に「観光客の国籍と観光コースのキャッチフレーズ」「観光コースの詳細」「アピール点」を記入した用紙を提出させた．議論は「対象の外国人の国籍」「コースのキャッチフレーズ」「観光コースの作成」の順に行うよう指示した．議論前に参加者には議論終了後に作成した観光コースについて「観光スポットの数が多い」などの評価点によって評価されることが伝えられた．

図 1：実験環境

図 1 に実験環境を示す．Kinect による各参加者の追跡データ，頭部に装着したモーションセンサによる 20 ms 毎の加速度・角速度・地磁気のデータ，ヘッドセットマイクによる各参加者の音声データ，ビデオカメラによる議論全体の動画が収集された．また録画した動画から議論の書き起こしを作成した．

3.2 セグメント境界のアノテーション

収集した議論に対して，会話のセグメント境界のアノテーションを行った．初めに，2 名のアノテータが試行と議論を繰り返し，コーディングスキームを作成した．セグメント境界になる発話のルールとして次の 3 つのルールを定めた，1.「話題が移った発話」，2.「会話のレベルが下がる（上がる）発話」，3.「書き込み作業について言及した発話」．1 番目のルールはたとえば，「旅行時期」から「旅行場所」へ話題が移った場合である．2 番目のルールはたとえば「旅行場所」からより詳細な話題「京都」へ移った場合である，3 番目のルールは議論内容に関係するトピックではないが，本研究で使用したコーパスでは多々見られたため，追加した．

連絡先：中野有紀子，成蹊大学理工学部，
y.nakano@st.seikei.ac.jp

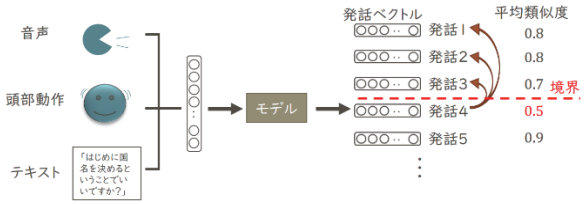


図 2:類似度に基づくセグメンテーション

この作成したコーディングスキームをもとに，全 9 グループについて自動で検出された音声区間に対して，1 名のアノテータがセグメント境界を付与した．ただし，前章で述べた通り，議論 35 分間の内，最後の 5 分間は基本的には用紙への書き込み作業になるため，それを除いた 30 分をアノテーションの範囲とする．その結果，各グループにつき平均約 35 個のセグメント境界が付与された．

4. 発話ベクトルを用いた類似度算出モデル

セグメント内の発話は結束性のある内容を持っていると考えられる．そこで，発話をベクトルで表現できれば，発話間の類似度を算出し，類似度に基づき境界を決定する手法を利用できると考えた．また，議論上で人間は言語だけでなく非言語情報も用いているので，マルチモーダルな情報を用いた議論のセグメント境界の識別を目的とする．そこで，マルチモーダルな情報を基に AutoEncoder を用いて発話を表現するベクトルを作成し，発話ベクトルの類似度を算出し，境界の識別を試みる．提案手法の概略図を図 2 に示す．

4.1 特徴量

セグメント境界の識別に使用する特徴量を下に記す．これらは発話単位で算出される．

(1) 新規/既出名詞の数

発話の書き起こしを形態素解析器 Mecab にかけ，名詞を抽出する．抽出した名詞がそれまでの議論で出現した名詞と一致すれば既出名詞，一致しなければ新規名詞とし，それぞれの個数を数え上げる．一度出現した名詞は既出名詞一覧に登録され，以後の新規/既出名詞の判断に用いられる．

(2) 現発話と直前の発話間の一致/不一致名詞数

直前の発話と現発話に共通する名詞の数を共通名詞数とする．また，現発話に出現し，直前の発話には出現しない名詞の個数を不一致名詞数とする．

(3) 現発話と直前の発話間の一致/不一致動詞数

直前の発話と現発話に共通する動詞の数を共通動詞数とする．また，現発話に出現し，直前の発話には出現しない動詞の個数を不一致動詞数とする．

(4) 発話長(時間長と形態素数)

発話長を示す特徴量として，発話区間の時間長 ms と発話を構成する形態素数の 2 種類を使用した．発話者の話速により，同じ時間長でも形態素数が異なるため，発話の形態素数も発話長として特徴量とした．

(5) オーバーラップ時間

2 つの連続する発話のオーバーラップ時間 ms を特徴量とする．1 人の発話に他の 2 人の発話がオーバーラップしている場合は，それらの合計を算出する．

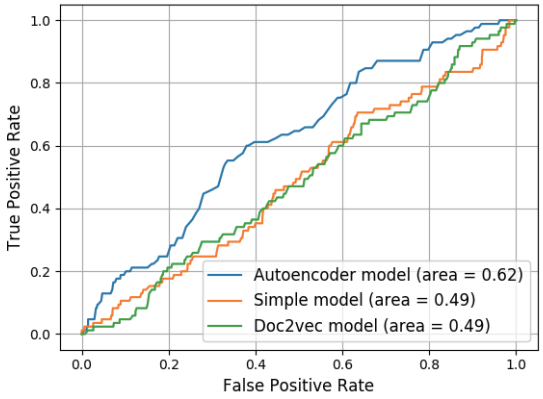


図 3: 3 つのモデルの ROC 曲線と AUC による比較

(6) 頭部回転量

発話が行われている時間を対象に，20 ms の時間窓を設定し，各区間における頭部のヨー角回転量 degree をデータとする．各区間における最大値・最小値・平均値・分散値をそれぞれ特徴量とする．

(7) 音声インテンシティ

音声解析ソフト Praat を用いて，収録した発話音声のインテンシティ dB を計測する．各発話区間において 10 ms の時間窓を設定し，その区間におけるインテンシティの最大値・最小値・平均値・分散値をそれぞれ特徴量とする．

(8) 合成加速度

議論参加者の後頭部につけられたモーションセンサの出力から下式を算出することで，合成加速度 HA を得る．ここで，x, y, z はそれぞれ x 軸, y 軸, z 軸方向の加速度の値である．

$$HA = \sqrt{x^2 + y^2 + z^2}$$

発話毎に，算出された合成加速度から最大値，最小値，平均値，分散を計算する．

(9) 合成加速度の Wavelet 変換特徴量

(8)で得られた合成加速度に Daubechies ウェーブレットを用いた多重解像度解析を行い，最下位レベルの信号の最大値，最小値，平均値，分散値を計算する．

(10) Doc2vec 特徴量

Doc2vec [4] を日本語の Wikipedia 記事を用いて学習し，各発話について 200 次元のベクトルを得る．

4.2 識別モデルの作成

発話ベクトル作成に使用する AutoEncoder は入力層，中間層，出力層の 3 層からなり，中間層と出力層でそれぞれ ReLu と Liner の活性化関数を用い，コスト関数は平均二乗誤差を使用する．入力された 241 次元のベクトルは中間層で 150 次元まで圧縮される．これを発話ベクトルとして類似度の算出に用いる．

AutoEncoder の訓練には 9 グループの内 7 グループ(4044 発話)を使用し，残りの 2 グループ(1124 発話)を検証に使用した．1000 epoch で訓練を行ったが 300 epoch にはモデルは収束したため，300 epoch 時点のモデルを用いて各発話に対して 150 次元の発話ベクトルを作成した．

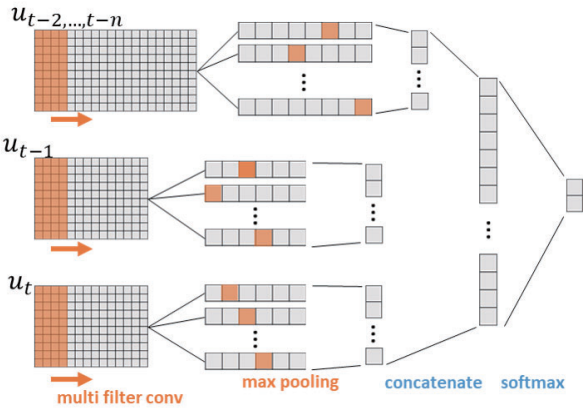


図 4: 言語モダリティモデルのアーキテクチャ

現発話と前 3 発話との発話ベクトルのコサイン類似度をそれぞれ算出し，その平均値を計算する．平均値が任意の閾値を下回れば，その発話を境界と識別する．

4.3 識別モデルの評価と議論

閾値によってセグメント境界かどうかの精度が変化するため，ROC 曲線と AUC で評価する．また，あいづちやフィラーなどの意味のない発話を考慮し，発話長が 0.8 秒以内の発話は除いた．比較のため，提案手法である AutoEncoder で特徴量ベクトルを圧縮したベクトルを用いるモデル（AutoEncoder model）の他に，圧縮せず特徴量を連結させただけのベクトルを用いるモデル（Simple model），Doc2vec 表現だけベクトルを用いるモデル（Doc2vec model）の 3 つを用意し，同様に評価した．その結果，ROC 曲線と AUC は図 3 のようになった．AUC は AutoEncoder model が 0.62, Simple model は 0.49，Doc2vec model は 0.49 となり，AutoEncoder model が一番性能が良い結果となった．

5. 教師あり学習によるセグメント境界識別モデル

各発話について境界であるかどうかを識別する 2 値分類モデルを作成する．教師データとして 3.2 節で作成したアノテーションを用いる．言語情報として発話の書き起こしテキスト，非言語情報として頭部動作，音声を用い，それぞれの 3 つのユニモーダルモデルを作成する． 言語モダリティモデルは Word2vec から得た各発話内の単語ベクトルを入力とし，頭部動作と音声モーダルモデルはそれぞれスペクトログラムを入力とするモデルを作成する．

5.1 入力データの作成

言語モダリティモデルの入力として，Word2vec から得た発話内の単語ベクトルを単語数分連結させた（200 次元×単語数）のサイズの 2 次元配列を発話毎に作成した．

頭部動作スペクトログラムは議論参加者毎に作成する．頭部に装着したモーションセンサから得た角速度から合成角速度を算出し，フーリエ変換することで 25 の周波数解像度を持つスペクトログラムを作成した．音声モーダルモデルの入力となる音声スペクトログラムも同様に議論参加者毎に作成する．ヘッドセットマイクより収録した音声から 32 の周波数分解能を持つスペクトログラムを作成した．また，各データに対して標準化を行った．

さらに，全議論中の最大発話長でゼロパディングを行う．その結果，それぞれの入力サイズは言語モダリティモデルが 217(コ)×200(次元)，頭部動作モデルが 761×25(周波数分解能)，音声モデルは 761×32(周波数分解能)となった．

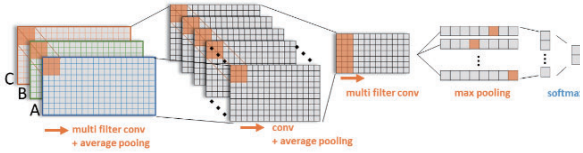


図 5: 頭部動作と音声モダリティのためのアーキテクチャ

5.2 言語モダリティモデルのアーキテクチャ

[5]を参考にアーキテクチャを作成した．アーキテクチャ図を図 4 に示す．現在の発話 u_t(t：発話数)，一つ前の発話 u_{t-1}，過去の発話 $u_{t-2,\ldots,t-n}$(n: 履歴発話数)の 3 つの入力を持つ．各入力されたベクトルは中間層で時間軸方向の 1D-convolution と global max pooling 後に 3 つの出力を連結させて，最終的に softmax 層で境界かどうかの予測が行われる．訓練時のカーネルサイズ，フィルタ数を表 1 に示す．

5.3 頭部動作と音声モーダルモデルのアーキテクチャ

頭部動作と音声モーダルモデルは入力がスペクトログラムであるため，同等のアーキテクチャを用いた．3 人の議論参加者のスペクトログラムを 3 チャンネルとして入力する．参与役割を考慮して，スペクトログラムの順序は優位性の近似値である現発話までの累積発話数[6]で並び替える．入力されたベクトルは中間層で複数の 2D-covolution 層と average pooling の後，時系列方向の 1D-convolution 層と global max pooling を通り，softmax 層で境界かどうかの予測を行う．訓練時には頭部動作モダリティ，音声モダリティともにそれぞれ 2 層の 2D-convolution と average pooling 層を用いた．訓練時のカーネルサイズやフィルタ数を表 1 に示す．

表 1: 各モダリティモデルとカーネルサイズ

	2D-conv	1D-conv	2D-average pooling	2D-conv フィルタ数	1D-conv フィルタ数
テキスト		4			32
頭部動作	(2, 2)	8	(2, 1)	8	32
音声	(2, 2)	8	(2, 1)	8	32

5.4 モデルの訓練と評価

言語，頭部動作，音声のユニモーダルモデルを作成し，3 つの組み合わせも考慮して合計 7 つのモデルを作成する．マルチモーダルモデルは各ユニモーダルモデルの全結合層の入力を連結し softmax 層への入力とすることでモデルの融合を行った．また，モデルの重みの初期値としてユニモーダルで学習した重みを使用する．言語モダリティモデルでは履歴数 10 発話し過去の発話は 8 発話でモデルを作成する．頭部動作モデルは各 convolution 層では10^{-7}の L2 正則化に加え，pooling 層後に 0.25 の dropout を行う．音声モデルも同様に各 convolution 層で10^{-8}の L2 正則化に加え，pooling 層後に 0.25 の dropout を行った．

作成されたモデルは各発話に対してセグメント境界であるかどうかの予測値を出力する．予測値が 0.5 を下回る，つまり境界であると判別された発話が少なかったため，ROC とその面積である AUC でモデル性能の評価を行う．

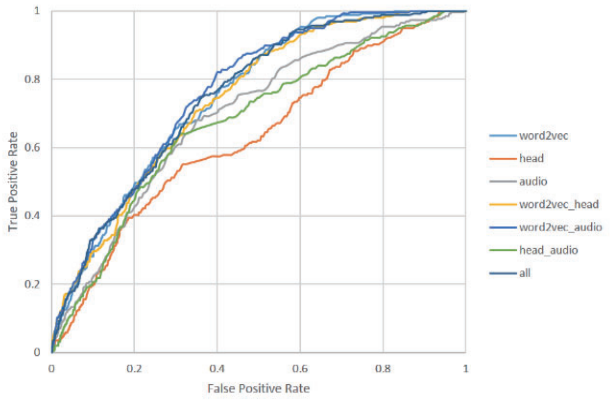


図 6: cross validation での ROC 曲線と AUC

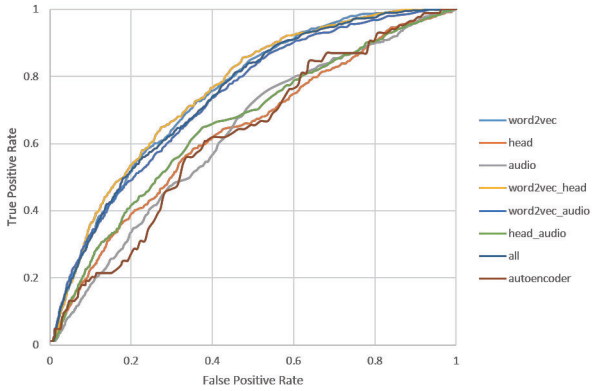


図 7: leave one Group out での ROC 曲線と AUC

全データを 5 分割し，そのうち 1 つを評価データとし，残りのデータで cross validation を行った．5 分割すべてに全グループのデータの発話が含まれるように，グループごとにセグメント境界の発話とそうでない発話に分割し，5 分割したものを統合した．

具体的な訓練・検証・評価データのサイズはそれぞれ約 2500 発話，約 830 発話，約 827 発話である．言語，頭部動作，音声モダリティのモデルをそれぞれ w2v, head, audio とする．これに加え，2 つのモダリティを組み合わせたモデル，言語＋頭部動作(w2v_head)，言語＋音声(w2v_audio)，頭部動作＋音声(head_audio)，全てのモダリティを統合したモデル(all)の 4 つのマルチモーダルモデルを作成した．図 6 にモデル毎の 4 hold の平均 ROC 曲線と平均 AUC を示す． w2v_audio, w2v, all の順に性能が高くなっていることから，言語情報に音声情報を追加すると性能が改善されることがわかる．

しかし，cross validation の評価は訓練・検証・評価で同一グループ内の発話が含まれてしまっている．そこで AutoEncoder モデルとの比較のため，グループ 2 と 7 の 2 グループを評価データとして，残りの 7 グループで leave one Group out を行った．cross validation と同様に各 7 個の検証データごとに作成されたモデルの平均 ROC と AUC を算出した結果を図 7 に示す．w2v_head, w2v, all の順に性能が高く，cross validation での評価とは異なり，言語情報に頭部動作情報を追加したモデルが一番良い性能であった．これは頭部情報がグループによる違いが少ない，識別に有効な情報を持っている，あるいは音声情報がグループ毎の性質に依存する情報である可能性がある．また，音声と頭部動作のモデルは僅差ではあるが，すべての教師あり学習モデルが AutoEncoder モデルの性能を上回った．

また，f-measure が最大となる時の閾値を用い，セグメント境界であると判定する precision, recall, f-measure を算出した．その結果を表 2 に示す．f-measure に着目すると，w2v_head モデルが最もよい性能であった．

表 2: モデル性能の比較

	w2v	head	audio	w2v_head	w2v_audio	head_audio	all	AutoEncoder
平均AUC	0.747	0.634	0.626	0.753	0.732	0.655	0.741	0.624
precision	0.211	0.177	0.131	0.253	0.228	0.181	0.223	0.148
recall	0.549	0.357	0.731	0.440	0.423	0.405	0.497	0.560
f-measure	0.305	0.237	0.222	0.321	0.297	0.25	0.308	0.234

6. まとめ

本研究では，言語・非言語情報に基づく，会話のセグメント境界の検出モデルを提案した．言語情報は最も有用なモダリティであったが，音声や頭部動作と組み合わせることで，より精度よく会話セグメント境界を検出できることが分かった．本研究では，議論参加者の身体情報として頭部動作情報を入力としたが，実際の議論ではポーズやジェスチャーも重要な意味を持っている．特にポーズは，例えば，前のめりになれば議論に対する積極性を示すなどセグメント境界の識別に有効であると考えられる．よって，頭部動作だけでなく姿勢位置などの全体的な身体情報も有効になると考えられる．

謝辞：本研究は，科学技術振興機構(JST) 戦略的創造研究推進事業(CREST)「実践知能アプリケーション構築フレームワーク PRINTEPS の開発と社会実践」(JPMJCR14E3)，および理化学研究所革新知能統合研究センターの支援を受けたものである

参考文献

[1]. Hearst, M.A., Multi-paragraph segmentation of expository text, in *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*, pp. 9-16, 1994.

[2]. Beeferman, D., A. Berger, and J. Lafferty, *Statistical Models for Text Segmentation.* Machine Learning, 1999. **34**(1-3): p. 177-210.

[3]. Galley, M., et al., Discourse segmentation of multi-party conversation, in *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics - Volume 1*. pp. 562-569 2003.

[4]. Le, Q. and T. Mikolov, Distributed representations of sentences and documents, in *Proceedings of the 31st International Conference on Machine Learning - Volume 32*, JMLR.org: Beijing, China. p. II-1188-II-1196, 2014.

[5]. Seokhwan Kim, Rafael E. Banchs, Haizhou Li, Exploring Convolutional and Recurrent Neural Networks in Sequential Labelling for Dialogue Topic Tracking, in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pp. 963–973, 2016.

[6]. 吉野 亮，八城 美里，高瀬 裕，中野 有紀子，会話エージェントによる優位性推定に基づくグループ会話への介入，第 29 回人工知能学会全国大会，1I4-4, 2015

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

コーパスに基づく日本語における謝罪への反応の分類

A corpus-based typology of reactions to apologies in Japanese

井上雅史　鈴木香里
Masashi Inoue　Kaori Suzuki

東北工業大学
Tohoku Institute of Technology

We proposed a typology of reactions to apologies base on the analysis of dialogue corpus. Our typology regards forgiving as the default reaction to the apologies. The responses are classified into two categories: the case when forgiving expressions are explicitly used and the case when forgiving expressions were left out. Then, non-default reactions were categorized based on the possible reasons why the apologies were not accepted.

1. はじめに

社会的な関係を維持，構築するために，会話中に謝罪あるいは詫びとそれに対する応答が出現する．知的な対話システムを実現するためには，エージェントがそれらを自然に使用することが求められる．謝罪の類型についての分析と比較して，謝罪に対する反応の類型については，あまり調べられていない．困難の原因の一つは，謝罪においては，一般に想起される典型的な表現が実際に使用される一方で，謝罪に対する反応では，典型的な表現が必ずしも使用されない傾向にあることと考えた．そこで，我々は反応の不在を考慮に入れた反応の類型を検討する．対話コーパスを使用することで，規範ではなく実際にどのような反応があり得るのかという観点での分類を試みる．

謝罪に対する応答に関して先行する調査に，Kitao らの英語を対象としたものがある [Kitao 14]．これは，言語教育への応用を志向している．我々の取り組みは，日本語を対象とし，反応の生成への応用を志向するものである．まず謝罪における典型的表現を定め，それに基づき謝罪発話を抽出する．次に，それらの謝罪に対する応答を分類する．

2. データ

対話データとして，名大会話コーパスを使用する．合計約 100 時間の日本語母語話者同士の雑談を文字化したものである [藤村 11]．主に大学院生にテープレコーダーか MD プレーヤーと外付けのマイクを渡し，30 分から 1 時間ぐらいの会話を収録している．本研究では，収録参加者の内 2 名の会話 95 件のデータを使用した．

3. 分析手順

3.1 謝罪の抽出

発話内で特定の表現が使われているものを謝罪とする．コーパス全体を俯瞰した予備調査から，対象とする表現として,『ごめん』『ごめんなさい』『すみません』『すまん』『失礼しました』『申し訳ありません』の 6 つを選定した．これらの表現が含まれる発話を謝罪としてコーパスからすべて抽出する．発話に対すして聞き手が言語的に反応した場合，その反応を『応答』とする．対話者毎に発話と応答のペアを抽出する．

連絡先: m.inoue@acm.org

謝罪に用いられる表現の使用傾向は，対話ペアの関係性に応じて変化することが実験的に示されている [趙 12]．例えば，親しい間柄では,「すみません」よりも「ごめん」の使用が好まれる．本研究においてコーパスから抽出した謝罪表現についても，話者間の人間関係と表現の使用比率を調べたところ，同様の傾向が確認された．

3.2 反応分類の手がかり

3.2.1 許しを表明する典型的表現

抽出した謝罪を含む発話とそれに対する反応のペアを分析者の主観により分析したところ，謝罪の言葉に対する反応を，『許す』『悩む』『許さない』『反応がない』『訂正』に分けることができた．そして，通常選択されるのは，謝罪を受け入れる，許す反応であった．明示的な発話によって許す場合には，典型的な表現の使用により許す場合と，反応しないことにより許す場合とが存在した．以下の 9 つの表現が典型的である．

1. いえいえ
2. いいよ
3. ううん
4. いやいや
5. 大丈夫
6. いいえ
7. そんな
8. 全然
9. とんでもない

3.2.2 許す応答表出可能

明示的な許しの表明がない場合,「許す応答表出可能」という基準で許す応答であると判断する.「許す応答表出可能」な反応とは，典型的表現が省略されている場合に相当し，典型的反応の挿入を行っても，反応の意味合いが変わらない．

典型的反応の表出が可能な場合の例：

発話：うん，ちょっとごめんね．
応答：鏡あるけど．

ここでの応答は，発話の謝罪について触れずに新たな提案をしている．次の様にいくつかの典型的表現を挿入した発話で，置き換えが可能である．

置換候補 1: いいよ．鏡あるけど．
置換候補 2：ううん．鏡あるけど．

置換候補 3: 大丈夫．鏡あるけど．
置換候補 4：全然．鏡あるけど．

次に，以上 3 つの分類（「典型的表現を使用した許し」「典型的表現を使用しない許し」「許す以外の反応」）全体の分析を行う．謝罪の回数上位 9 つのデータを使用し，3 つの分類に属する反応の使用割合の統計を取る．

4. 反応分析結果

4.1 典型的表現の使用

分析対象中に出現した典型的表現は，表 1 のような使用場面に対応していた．グループ化を行うと，謙遜系，信頼系，遠慮系の 3 つの系統に類別が可能である．謙遜系は，発話が敬語を話している場合や，感謝の言葉を含んでいる場合に使用される．信頼系は，発話がタメ口の場合，応答側が妥協をしている場合に使用される．遠慮系は，謝罪の打ち消し，謙遜を表すとき，許しを表すとき，親しい関係のときに使用される．

表 1: 許すための典型的表現と使用場面

典型的表現	使用場面
いいえ	謙遜を表すとき．
いいよ	許しを表す．友人などの親しい関係．妥協．
ううん	謝罪の対象の行為に対して，応答側に支障がないとき．
いやいや	遠慮を表すとき．逆に，謝罪の理由が明確に違う時も使用．
大丈夫	謝罪の対象の行為に対して，応答側に支障がないとき．
いいえ	謝罪の対象の行為に対して，応答側に支障がないとき．３０代以上会話で使用する． 友人などの近い間で使用される． 敬語もたまに出るような関係で使用される．

4.2 許す応答表出可能

明示的に許すことが表明されていない場合に，典型的な表現が挿入候補として現れる割合は，信頼系 70.6%，謙遜系 22.5%，遠慮系 3.9% であった．

4.3 許す以外の反応

許す以外の反応の場合は，謝罪を受け入れなかった原因の観点から，以下の 9 つの場合に類別が可能であった．

1. 発話者の 1 ターンがとても長い場合
2. 1 つ前のターンの謝罪に対して発話者が謝罪している場合
3. 発話から次の質問をされている場合
4. 発話者の謝罪理由を確認している場合
5. 発話者の態度に呆れている場合
6. 発話者の謝罪理由が違っている場合
7. 発話者の謝罪に対して納得していない場合
8. 発話者のフォローをしている場合
9. 発話者の次の発言を待っている場合

4.4 集計結果

謝罪の回数が多かった上位 9 つのデータの反応別の統計を取ると図 1 の様な結果だった．参加者 F130 の場合，data112，data110，data111 において，典型的表現を使用せず，許しを表明していた．この様に，典型的表現の使用には謝罪を受け入れる際に必須ではない可能性が高い．

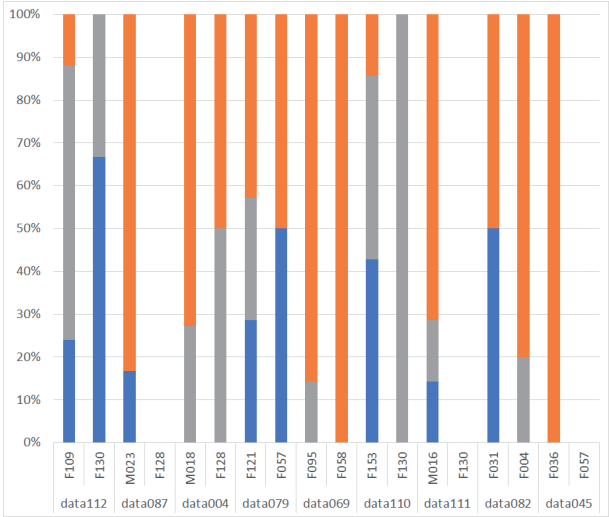


図 1: 話者ごとの反応の種別割合．横軸は話者とデータ ID．各バーの下から「典型的表現を使用した許し」「典型的表現を使用しない許し」「許す以外の反応」を表す．

5. おわりに

対話コーパスから抽出したデータを基に，謝罪への反応を，「典型的表現により許す反応」,「許す応答表出可能な反応」,「許す以外の反応」に類別した．「典型的表現により許す反応」には，9 つのよく使用される典型的表現が存在し，それらは使用する場合によって謙遜系，信頼系，遠慮系に類別が可能であった．「許す応答表出可能な反応」は，典型的表現が発せられていないが，文脈から許しの意味と捉えることが可能な反応である．「許す以外の反応」は，許さない原因から，9 つの場合に類別が可能であった．

本研究の結果は，人間同士の会話においては，明示的に謝罪を受け入れる表現を使用しないことによって，謝罪を受け入れる反応が頻繁に使用されていることを示唆しており，対話システム等においても，応用可能と考えられる．ただし，反応を生成する場合には，どのような分類があるかの情報だけでは不足であり，例えば謝罪の表現をいくつかの類に分ける試みを参考に [滝浦 08]，反応の種類との関係を調査するなど，反応選択の規則を明らかにしていくことが必要となる．

参考文献

[Kitao 14] Kitao, S. K. and Kitao, K.: A corpus-based study of responses to apologies in US English, *Journal of culture and information science*, Vol. 9, No. 2, pp. 1–13 (2014)

[滝浦 08] 滝浦 真人：ポライトネス入門, 研究社 (2008)

[藤村 11] 藤村 逸子, 大曽 美恵子, 大島ディヴィッド義和：会話コーパスの構築によるコミュニケーション研究, 藤村逸子, 滝沢 直宏（編）, 言語研究の技法：データの収集と分析, pp. 43–72, ひつじ書房 (2011)

[趙 12] 趙 翻：日本語と中国語における謝罪表現の対照研究：家族と親友間の異なりに注目して, 東洋大学大学院紀要, Vol. 49, pp. 124–98 (2012)

Organized Session | Organized Session | [OS] OS-8

[4G2-OS-8a] マイニングと知識創発(1)

砂山 渡（滋賀県立大学）、加藤 恒昭（東京大学）、西原 陽子（立命館大学）、森 辰則（横浜国立大学）、高間 康史（首都大学東京）

Fri. Jun 7, 2019 12:00 PM - 1:40 PM Room G (302A Medium meeting room)

[4G2-OS-8a-01] Proposal of Summarization Support System Based on Mixed-Initiative Interaction

〇Hiroto Kato[1], Yasufumi Takama[1] （1. Tokyo Metropolitan University）

12:00 PM - 12:20 PM

[4G2-OS-8a-02] *differential* PLSA

〇Koji Nomori[1] （1. Analytics Design Lab Inc.）

12:20 PM - 12:40 PM

[4G2-OS-8a-03] Basic Analysis of Descriptions on Medicine-care in Nursing Records

〇Seiya Uesaka[1], Yoshinori Hijikata[1], Muneou Suzuki[2] （1. Kwansei Gakuin University , 2. University of Miyazaki Hospital ）

12:40 PM - 1:00 PM

[4G2-OS-8a-04] Extraction of Classification Patterns from Deep Learning Networks

〇MASAYUKI ANDO[1], Yoshinobu Kawahara[3,4], Wataru Sunayama[2], Yuji Hatanaka[2] （1. Graduate School of Engineering, The University of Shiga Prefecture, 2. School of Engineering, The University of Shiga Prefecture, 3. Institute of Mathematics for Industry, Kyushu University, 4. RIKEN Center for Advanced Intelligence Project）

1:00 PM - 1:20 PM

[4G2-OS-8a-05] Attachment of “ Cuteness” to Dialogue Sentences Using Deep Learning

〇Yusuke Goto[1], Wataru Sunayama[2], Yuji Hatanaka[2], Kazunori Ogohara[2] （1. Graduate School of Engineering, The University of Shiga Prefecture, 2. School of Engineering,The University of Shiga Prefecture）

1:20 PM - 1:40 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

相互主導型インタラクションに基づく要約生成支援システムの提案

Proposal of Summarization Support System Based on Mixed-Initiative Interaction

加藤　大登　　高間　康史
Hiroto Kato　　Yasufumi Takama

首都大学東京大学院システムデザイン研究科
Graduate School of System Design, Tokyo Metropolitan University

This paper proposes a text summarization support system based on mixed-initiative interaction. Automatic summarization is used to generate sentences useful as a clue for information gathering. However, the quality of the generated summary is still lower than that generated by humans. Furthermore, it does not consider users' interest and knowledge. To solve the problems, the proposed system dynamically suggests what is useful for writing a summary, such as important sentences, in response to users' input. The effectiveness and challenges of the proposed approach are discussed based on the result of the experiment using a prototype system.

1.　はじめに

関連研究の論文サーベイを行う際などには，論文要約が重要な手がかりとなる．しかし，要約文を作成するには元の文章内容を十分に理解した上で内容を吟味して取捨選択する必要があるため，人手で行うには時間と労力がかかる．この課題を解決するために自動要約の研究が行われている．しかし現状の自動要約文の精度は人手で作成されたものに劣るほか，ユーザが要約作成プロセスに関与することができない．

これらの問題に対し，本稿ではユーザと計算機が相互主導的に要約文を作成するアプローチを提案する．提案するアプローチではユーザが計算機の提示内容を確認し，計算機はユーザの入力に合わせて提示内容を変更することで，効率的な要約作成を支援する．プロトタイプシステムを構築してユーザ実験を行い，ユーザが有用性を感じる機能や必要としている機能を調査した結果を示す．

2.　関連研究

自動要約手法の代表的アプローチに抽出型があり，LexRank[Erkan 04] は他の文との類似度が高い文，また重要文との類似度が高い文を抽出して要約を生成する．ユーザの文章入力支援を目的としたシステムには英語論文執筆支援ソフト BEAR[Narita 02] があり，オンライン辞書へのリンク機能や単語の用例を提示する機能が実装されている．相互主導型インタラクションを取り入れたシステムの例として，詩の作成支援システム CoPoeTryMe[Oliveira 17] が挙げられる．詩の自動生成システムのパラメータをユーザが変更可能であり，生成内容をユーザ自身が制作中の詩の一部として取り入れることができる．

3.　提案手法

提案システムの構成を図 1 に示す．フロントエンドは，対象文書の内容やユーザの入力を受け取ってバックエンドに送信し，バックエンドの計算結果をユーザインタフェースに反映する．バックエンドは，データセットから事前に単語の IDF 値と分散表現を作成する機能と，対象文書から文書モデルを作成する機能，ユーザの入力をもとに提示内容を構成する機能を持つ．単語の IDF 値と分散表現の元になるデータセットは Wikipedia 日本語版の Dump ファイル *1 を用いる．

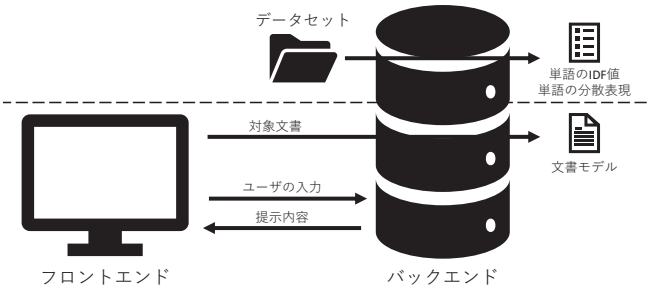


図 1: システム構成

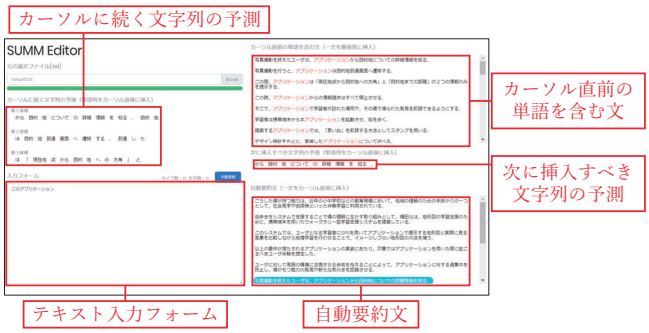


図 2: フロントエンドの画面構成

図 2 はフロントエンドの画面構成で，テキスト入力フォームと提示スペースから構成される．提示スペースの各欄にはユーザの要約文作成を支援する情報が提示され，クリックしてテキスト入力フォームに挿入できる．提示内容はテキスト入力フォーム内でのキー入力に従い，自動要約の出力を利用して動的に生成する．自動要約文の中で青色に強調された文は，現在入力中の位置に挿入すべき文であることを示している．

バックエンドでは，フロントエンドから受け取ったユーザ入力に従い，提示内容を以下の手順で動的に計算する．対象文書中の各文の重要度は LexRank アルゴリズム [Erkan 04] で求める．

連絡先:　高間康史，首都大学東京大学院システムデザイン研究科，〒 191-0065　東京都日野市旭が丘 6-6，ytakama@tmu.ac.jp

*1　https://dumps.wikimedia.org/jawiki/20181001/jawiki-20181001-pages-articles-multistream.xml.bz2

1. テキスト入力フォームの，カーソル直前の単語 x を取得する．
2. 単語の分散表現と cos 類似度を用いて，対象文書中から x に類似した単語を探索する．
3. x の類似単語を含む文を重要度順に並べ「カーソル直前の単語を含む文」とする．
4. 手順 3 の文で重要度が最も高い 3 文の，x の類似単語以降の語句列を「カーソルに続く文字列の予測」とする．
5. テキスト入力フォームの文章を S，手順 3 の各文を D_i として式 1 によりスコアを求め，最も高い 1 文の x の類似単語以降の語句列を「次に挿入すべき文字列の予測」とする．
6. 手順 5 の文を S，対象文書中の各文を D_i として式 1 によりスコアを求める．最も高い 9 文に S を合わせた計 10 文を対象文書中の出現順に並べ「自動要約文」とする．

$$Score(D_i) = Imp(D_i) - Sim(D_i, S) \tag{1}$$

式 1 は多様な文を選択することを意図したものであり，Imp は重要度を，Sim は類似度を意味する．類似度は文を単語の集合とみなし Jaccard 係数を用いて計算する．

手順 2 で分散表現を利用することで，単語 x が対象文書中に存在しない場合でもシステムが機能する．手順 5 では，重要だがまだ執筆されていない情報を持つ文を「次に挿入すべき文字列」と想定し抽出している．手順 6 では LexRank の結果を利用しつつ「次に挿入すべき文字列」とは類似しない文から自動要約文を生成している．

4. ユーザ実験

4.1 実験内容

要約執筆支援エディタ「SUMM Editor」を実装し，実験協力者 12 名に論文要約の執筆を行ってもらう評価実験を行った．本実験ではエディタの機能と，要約文の文字数などの目安を伝えた後に執筆を開始してもらった．要約対象となる論文には実験前に目を通してもらっており，執筆中も閲覧可能とした．制限時間は設けず，協力者自身が要約作成の完了を宣言したとき実験を終了した．また実験終了後にアンケートを行い，各機能の評価や意見，対象論文の理解度を調査した．評価値は 5 段階で，1 が最も否定的，5 が最も肯定的な評価を表す．

4.2 実験結果・分析

キータイプ回数および機能ボタンのクリック回数と執筆時間の相関係数を表 1 に示す．クリック回数と執筆時間には正の相関があり，特に「自動要約文」「次に挿入すべき文字列の予測」機能の相関が高いことがわかる．この二つはシステムから情報を推薦する機能であり，またアンケートに「とりあえず入れてみてから文の校正を考えた」などの回答があったことから，要約文執筆に行きづまった実験協力者が発想支援を受けるために利用したと考える．

各機能に関するアンケート結果について，評価値ごとの回答人数を表 2 に示す．Q1 では,「カーソル直前の単語を含む文」の機能の評価が比較的高かった．アンケートでは「同じ単語がどの文で出ているかがわかり要約作成の手助けとなった」という意見があり，単語を検索して文書内での使われ方を確認するという想定通りの用途で使用されたと言える．

Q2 では，文単位の挿入機能である「自動要約文」の機能の評価が比較的高かった．アンケートには「打ち込む手間が省けた」や「自動要約文の文章を挿入し不必要な部分を取り除いていく」という意見があり，要約執筆に活用できたと考える．しかし内容の参考度合いの評価については「悩んだ時の参考にした」との一方で「あまり気にならなかった」との意見もあり，人により極端に分かれることが判明した．

単語単位での挿入機能（「カーソルに続く文字列の予測」と「次に挿入すべき文字列の予測」）は両設問とも低評価の方が多かった.「自動要約文」の挿入機能が比較的高評価であることからも，対象文書内での意味を確認しやすい文単位での挿入のほうが，単語単位の挿入より使いやすかったと考える．

そのほかに「直観的な操作は今回のタスクを楽にしてくれた」と相互主導的インタラクションの簡易な操作が評価される一方で,「suggest までの時間が少しネック」「すでに要約文に入力してある文と被る部分がないように文を挿入してほしい」といった意見があった．システムを改善するには，簡易な操作方法を維持しつつ要約生成支援により適した機能を提供する必要があると考える．

表 1: 各操作と執筆時間の相関係数

操作		相関係数
キータイプ		0.3240
機能ボタンのクリック		0.6062
機能別	自動要約文	0.6418
	カーソルに続く文字列の予測	0.1731
	カーソル直前の単語を含む文	0.4770
	次に挿入すべき文字列の予測	0.6308

表 2: 各機能の評価値ごとの回答人数

（Q1. 内容が参考になったか，Q2. 挿入機能が役立ったか）

設問	Q1					Q2				
評価値	1	2	3	4	5	1	2	3	4	5
自動要約文	0	6	0	3	3	1	1	4	4	2
カーソルに続く...	2	4	1	5	0	3	2	3	2	2
カーソル直前の...	1	1	3	5	2	3	2	2	3	2
次に挿入すべき...	3	2	6	1	0	3	2	6	1	0

5. おわりに

本稿ではユーザと計算機が相互主導的に要約文を作成するアプローチを提案し，プロトタイプ「SUMM Editor」によるユーザ実験で有用な機能や必要とされる機能を調査した．

相互主導的インタラクションは直感的な操作で情報提示を行うことができる．しかし，ユーザの求めるシステムへと改善するためには，ユーザの入力に合わせた精度の高い提示を行うことに加えて，インタフェースもより要約執筆に適したものに改善する必要があることが示された．今後は計算時間の短縮と文挿入機能の利便性向上，本用途に適した要約生成アルゴリズムの検討を進めていく予定である．

参考文献

[Erkan 04] G. Erkan, D. R. Radev, LexRank: Graph-based Lexical Centrality as Salience in Text Summarization, Journal of Artificial Intelligence Research, Vol. 22, pp. 457–479 (2004)

[Narita 02] M. Narita, K. Kurokawa, and T. Utsuro, A web-based English abstract writing tool using a tagged E-J parallel corpus, LREC 2002, pp. 2115–2119 (2002)

[Oliveira 17] H. G. Oliveira, T. Mendes, and A. Boavida, CoPoeTryMe: a Co-Creative Interface for the Composition of Poetry, ICNL 2017, pp. 70–71 (2017)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

differential PLSA
ーテキスト情報の典型的なトピックではないより個性的なトピックの抽出ー

differential PLSA
A Method of Extracting not Representative Topics but More Individual Topics from Text Data

野守　耕爾[*1]
Koji Nomori

[*1] 株式会社アナリティクスデザインラボ
Analytics Design Lab Inc.

Abstract: This study proposes a new method extracting topics from text data named *differential* PLSA. The method applies PLSA to a differential co-matrix computed by logarithm of the ratio of an observed co-matrix to expected one. It enables to extract not representative topics but more individual ones. This paper showed the effectiveness by applying the method to patent document data and comparing with results using normal PLSA. As a result, topics extracted by the method were composed of less frequent and more concrete elements, and they were more individual.

1. はじめに

データ活用がビジネスで進む昨今，テキストデータを分析することで定性情報によるインサイトを獲得して業務に活用しようとする取り組みも増えている．筆者は企業におけるデータ分析・活用のコンサルティング事業を展開しているが，ビジネスにおけるテキストデータの活用では，例えば顧客満足度調査などで実施されるアンケートの自由記述文を分析し，回答者の生の声から潜在的なニーズを抽出することでマーケティングに活用したり，コールセンターの問い合わせ履歴のデータを分析し，消費者ならではの隠れた評価の観点を抽出することで商品・サービスの改善を検討したり，特許公報の文書データを分析し，技術トレンドや他社の得意技術などを抽出することで研究開発戦略や他社との提携戦略を検討するなど，様々な業務領域でテキストデータの活用が進められている．

ビジネスにおけるテキストデータの活用では，新たな気づきとなるインサイトをテキスト情報から抽出することがしばしば求められるが，結果が期待外れということも少なくない．その理由は，業務担当者からすると経験的によく知っている結果であり，目新しさがないということが多い．テキストデータの分析には，テキストマイニング技術を適用することが通例であるが，テキストマイニングは文章に含まれる単語を抽出し，主に文章の全体像を各単語の出現頻度をベースに理解する手法となる．ここで，頻度の多い単語では当たり前の結果となるため，頻度の少ない単語まで確認しようとすると，今度は結果が複雑になり何が重要なのか分からず，結果の理解ができなくなるというジレンマに陥ってしまう．特に対象となる文章がビッグデータになると，抽出される単語も膨大となり，単語ベースの結果では複雑になりすぎるため，文章を単語ではなく意味的なまとまりのあるトピックをベースに理解する手法として，トピックモデルの適用が有効となる．トピックモデルでは単語の出現データを教師なし学習によりトピックに集約するが，これは分析対象のテキストデータ全体をうまく表現するようないわば数学的な最適解となるトピックを抽出するので，どうしても典型的なトピックになりがちになり，ビジネスの業務担当者の視座では満足しがたい結果になってしまう．

連絡先：野守耕爾，株式会社アナリティクスデザインラボ，
koji.nomori@analyticsdlab.co.jp

そこで本稿では，テキストデータにトピックモデルを適用した際に，典型的なトピックだけではなく，より個性的なトピックを抽出する方法を提案し，その適用事例を紹介する．

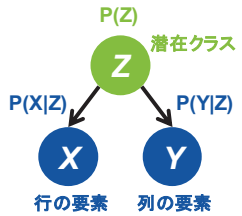
2. PLSA

本稿では従来のトピックモデルとして PLSA（確率的潜在意味解析：Probabilistic Latent Semantic Analysis）を紹介する．

2.1 PLSA の概要

PLSA は，行列データ（共起行列）の行要素 X と列要素 Y の背後にある共通特徴となる潜在クラス Z を抽出する次元圧縮法であり，元々は文書分類のために開発された手法である [Hofmann 1999]．PLSA のグラフィカルモデルを図 1 に示す．PLSA の計算では，X と Y の共起確率を潜在クラス Z を使って式展開し，この対数尤度関数を最大化する EM アルゴリズムを実行することで，3 種類の確率変数 $P(X|Z)$, $P(Y|Z)$, $P(Z)$ が計算される．文書分類における PLSA では，文書 X とそこに出現する単語 Y の間には潜在的な意味トピック Z があることを想定し，各文書における単語の出現頻度が記録された「文書」×「単語」の共起行列データを学習し，文書と単語の共通特徴となるトピックを見つける．これにより「文書」×「トピック」という低次元データに変換して文書分類できる．

PLSA はクラスタリングの手法としても用いられるが，クラスタリングという観点で PLSA が他のクラスタリング手法と異なる特徴の一つは，行と列を同時にクラスタリングできることである．一般的なクラスタリング手法は，列をベースに行をクラスタリングする，あるいは行をベースに列をクラスタリングするため，どちらか一方しかクラスタリングできない．一方 PLSA で抽出される潜在クラスには，行の要素と列の要素が同時に所属することができる．



$$P(X,Y) = \sum_Z P(X|Z)P(Y|Z)P(Z)$$

図 1　PLSA のグラフィカルモデル

2.2 PLSA の共起行列構成の工夫

行と列を同時にクラスタリングできる PLSA では，行と列は双方が十分意味を持つ情報で構成すれば，抽出された潜在クラスの意味を行と列の 2 つの情報軸から解釈することができる．本来の PLSA の適用では，「文書」×「単語」という構成の共起行列をインプットとするが，この構成を工夫することで解釈のしやすい潜在クラスを抽出する試みがある．

例えば，「品詞」×「品詞」の共起行列を用いる方法や[Kameya 2005][野守 2014]，「単語」×「係り受け」の共起行列を用いる方法が提案されている[野守 2018]．「単語」×「係り受け」の共起行列を用いる PLSA では，単語と係り受け表現が同時にクラスタリングされ，単語という話題の観点となる軸に基づき，その観点の具体的内容となる係り受け表現をグルーピングできるため，より文脈上近しい言葉・表現でまとめられた解釈のしやすいトピックを潜在クラスとして抽出できるとされている．

3. *differential* PLSA

テキストデータに PLSA を適用したとき，典型的なトピックだけではなく，より個性的なトピックを抽出する方法として，*differential* PLSA（以下 *diff*-PLSA）を提案する（特許出願中，商標出願中）．

diff-PLSA では，テキストデータにおいて行要素 X_i と列要素 Y_j が共起する実測頻度 $n(X_i, Y_j)$ を値として持つ通常の共起行列 M に加え，行要素 X_i と列要素 Y_j が共起する期待頻度 $n'(X_i, Y_j)$ を値として持つ共起行列 M' を構成する．この各期待頻度に対する実測頻度の比率の対数を取った $log\left(n(X_i, Y_j)/n'(X_i, Y_j)\right)$ を値として持つ *differential* 共起行列（以下 *diff*-共起行列）を構築し，これに PLSA を適用する．*diff*-PLSA で適用する共起行列のイメージを図 2 に示す．ここで期待頻度とは，X_i の総頻度（出現文章数）$n(X_i)$，Y_j の総頻度（出現文章数）$n(Y_j)$，総文章数 N から式(1)のように計算される．なお，*diff*-共起行列の値が負数となるものは 0 に置換する．

$$n'(X_i, Y_j) = n(X_i) \cdot n(Y_j)/N \qquad (1)$$

実測頻度の共起行列に適用する通常の PLSA では，その解を求める最適化計算において，どうしても頻度が高い要素に高い確率が割り当てられ，結果として抽出されるトピックは典型的なものになる傾向があり，目新しさに欠けてしまう．一方，*diff*-共起行列では，実測共起頻度を期待共起頻度で除した値を持つが，実測共起頻度が高い共起ペアでも，元々全体の頻度が高い要素が含まれるときには期待共起頻度も高くなるため，実測共起頻度を期待共起頻度で除すことで値の大きさが制限される．逆に実測共起頻度が高くない共起ペアでも，期待共起頻度がそれよりも十分低ければ *diff*-共起行列での値は大きくなり，これに PLSA を適用した解ではこうした要素にも高い確率が割り当てられる可能性がある．つまり，通常の PLSA では頻度が低い要素は高い確率が割り当てられない傾向があるが，*diff*-PLSA ではそうした要素にも高い確率が割り当てられる可能性があり，より個性的なトピックが抽出されることが期待できる．

期待共起頻度に対する実測共起頻度の比率に対数を取る理由としては，極端に高くなる値を制限ためである．特に期待共起頻度は 1 未満となるケースも多く，比率のみでは値が高くなりすぎるものもある．この状態では共起行列全体の値の分布は大きくばらつき，極端な値の開きが生まれてしまうため，PLSA を適用した際の最適化計算において，今度はこの極端に大きな値に引っ張られる結果となり，必要以上にデフォルメされた歪んだトピックとなることが考えられる．そこでこの比率の値の対数を取ることで値の分布をならし，上記の現象を制限する．

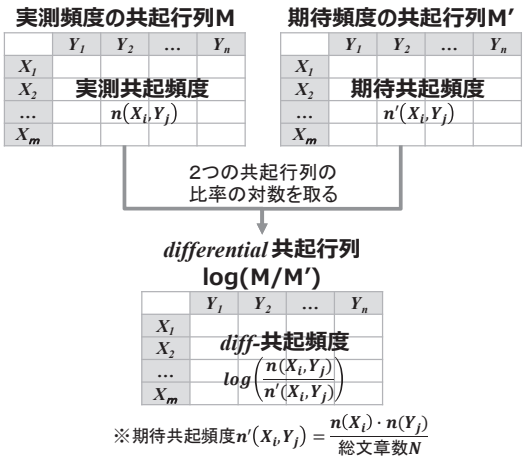


図 2　*diff*-PLSA で適用する共起行列

4. *diff*-PLSA の適用事例

本稿では特許の要約文のテキストデータを対象に，通常の PLSA と *diff*-PLSA を適用した事例を紹介し，その結果を比較することで *diff*-PLSA の有効性を検討する．

4.1 対象データと共起行列の構成

特許の要約文と請求項に「車」「電気」を含む 10 年分（出願日が 2007 年 1 月 1 日～2016 年 12 月 31 日）の国内特許公報 26,419 件を用意し，その要約文の文章データを分析対象とした．これにより電気自動車関連の技術トピックを抽出する．

共起行列の行列構成は，行を「名詞」，列を「係り受け」とした．係り受けは文法的につながりのある単語のペアを抽出するものだが，名詞に対する動詞・形容詞・形容動詞のペアを対象とした．こうした構成の共起行列に PLSA を適用することで，使われ方の似ている名詞と係り受けで構成されるトピックを得ることができる．つまり技術を象徴する名詞とその名詞を修飾する表現でまとめられたトピックを抽出でき，解釈しやすい結果が得られると期待できる．要約文にテキストマイニングを実行し，名詞と係り受けを抽出したところ，データ頻度が 20 件以上となる名詞が 3,020 語，係り受けが 2,128 表現得られた．この「名詞 3,020 語」×「係り受け 2,128 表現」で構成される共起行列を PLSA および *diff*-PLSA で適用するベースとした．

4.2 通常の PLSA の適用

通常の PLSA を適用する共起行列において，名詞と係り受けの各ペアの共起頻度とは，そのペアが同時に出現する文章数とした．ここでいう文章とは，句点で区切られた一文を指し，今回の 26,419 件の特許データに含まれる総文章数は 229,598 件であった．この共起行列に PLSA を適用してトピックを抽出するが，PLSA は予めトピック数を設定する必要があり，また初期値により解が異なる特性がある．そこでトピック数を 1 刻みで変化させ，それぞれのトピック数に対して PLSA を初期値を変えて 5 回ずつ実行し，それぞれの解を情報量基準 AIC で評価して最も評価の良い解を採用した．その結果 34 個のトピックが得られた．

PLSA のアウトプットは，①トピック Z における行要素 X（名詞）の所属確率 $P(X|Z)$，②トピック Z における列要素 Y（係り受け）の所属確率 $P(Y|Z)$，③トピック Z の存在確率 $P(Z)$，という 3 つの確率が計算される．特に各トピックにおいて①②の確率が高い名詞，係り受けを確認し，そのトピックの意味を解釈する．抽出されたトピックの内容例を表 1 に示す．表 1 ではトピック Z05 と Z14 について，所属確率の高い上位 7 つの名詞および係り受けと，

それぞれの所属確率 P, 頻度 n(出現文章数)についてまとめたものである. トピック Z05(表 1 上)は,「ブレーキ」や「制動力」に関する表現で確率が高いため, ブレーキに関する技術と解釈できる. トピック Z14(表 1 下)は,「給電」や「電力の供給」に関する表現で確率が高く, また「非接触」という表現も上位にあることから, 非接触受電などの給電装置に関する技術と解釈できる. このように解釈を付けた 34 個のトピックの一覧を表 2 に示す.

表 1 通常の PLSA で抽出されたトピックの内容の例

トピックZ05					
P(XZ)	n(X)	X:単語	P(YZ)	n(Y)	Y:係り受け
7.3%	779	ブレーキ	3.2%	92	車両-ブレーキ
5.0%	1,384	作動	2.4%	33	ブレーキ液圧-発生
3.4%	5,188	モータ	2.2%	53	制動力-発生
2.9%	279	制動力	1.8%	49	ブレーキ-備える
2.9%	867	運転者	1.7%	32	操作量-応ずる
2.7%	1,117	車輪	1.6%	82	ブレーキ-提供
2.6%	1,005	操作	1.6%	59	電気信号-基づく
…	…	…	…	…	…

トピックZ14					
P(XZ)	n(X)	X:単語	P(YZ)	n(Y)	Y:係り受け
9.7%	1,162	給電	4.3%	1,350	電力-供給
5.1%	3,629	電力	2.7%	115	給電-行う
3.6%	1,140	電源	2.5%	52	電力-受電
3.4%	329	給電装置	2.0%	28	給電-電力
2.4%	4,655	電気自動車	1.6%	124	電源-接続
2.4%	180	非接触	1.5%	20	駐車装置-変化
2.4%	1,052	外部	1.5%	34	非接触-受電
…	…	…	…	…	…

表 2 通常の PLSA で抽出された 34 個のトピック一覧

No.	トピック名	No.	トピック名
Z01	エンジンの始動と停止	Z18	演算・推定
Z02	動力の伝達	Z19	機器の異常検出
Z03	モータ駆動	Z20	操作スイッチ
Z04	ロータ・ステータなど回転部品の構成	Z21	筐体
Z05	ブレーキ装置	Z22	表面の形成
Z06	動作制御	Z23	位置とその移動
Z07	動力伝達の制御	Z24	配置・位置・方向
Z08	スイッチの切り替え	Z25	構成の方位
Z09	交流・直流の変換	Z26	構成
Z10	エネルギーの変換	Z27	接続
Z11	電池モジュールの提供	Z28	方法の提供
Z12	二次電池の構成	Z29	損傷や浸水など不具合の防止
Z13	電気自動車の蓄電池充電	Z30	小型化・簡素化・低コスト化など付加価値
Z14	非接触受電など給電装置	Z31	効率性・安全性の向上
Z15	外部への電力供給	Z32	既存エンジンへの警鐘・樹脂組成物の提供
Z16	空調などの冷却・加熱	Z33	重力発電の活用による地球温暖化防止
Z17	情報通信	Z34	タービン発電の出力向上・燃費低減

4.3 各文章に対するトピックのスコア計算

全特許データの各文章に対して各トピックのスコア(該当度)を計算した. これにより各トピックと関連の強い具体的な文章を確認できるため, トピックを解釈するうえの参考情報となる. 実際に表 2 の解釈においてもこのスコア結果を使用した.

文章 S_h におけるトピック Z_k のスコアは $P(S_h|Z_k)/P(S_h)$で定義した. これはトピックを条件とすることで文章の発生確率が何倍になるのかを示し, そのトピックをよく話題にしている文章ほど高くなる. 以下, $P(S_h|Z_k)$と $P(S_h)$の計算について説明する.

$P(S_h|Z_k)$は, 文章 S_h を名詞で定義される文章 Sx_h と係り受けで定義される文章 Sy_h に分解し, それぞれ $P(Sx_h|Z_k)$と $P(Sy_h|Z_k)$を計算し, それらを統合して $P(S_h|Z_k)$を計算する. $P(Sx_h|Z_k)$と $P(Sy_h|Z_k)$は式(2)(3)で計算される. 名詞 X_i と係り受け Y_j の頻度(出現文章数)をそれぞれ $n(X_i)$と $n(Y_j)$とすると, $P(Sx_h|X_i)$は $n(X_i)$の逆数, $P(Sy_h|Y_j)$は $n(Y_j)$の逆数として計算される. $P(X_i|Z_k)$と $P(Y_j|Z_k)$はそれぞれ PLSA の実行結果で得られている. $P(S_h|Z_k)$は式(4)で計算され, 文章 S_h において Sx_h と Sy_h の重みは同じであるため, $P(S_h|Sx_h)$と $P(S_h|Sy_h)$はそれぞれ 0.5 とする. $P(S_h)$は式(5)で計算され, $P(Z_k)$は PLSA の実行結果で得られている.

$$P(Sx_h|Z_k) = \sum_i P(Sx_h|X_i)P(X_i|Z_k) \quad (2)$$

$$P(Sy_h|Z_k) = \sum_i P(Sy_h|Y_i)P(Y_i|Z_k) \quad (3)$$

$$P(S_h|Z_k) = P(S_h|Sx_h)P(Sx_h|Z_k) + P(S_h|Sy_h)P(Sy_h|Z_k) \quad (4)$$

$$P(S_h) = \sum_k P(S_h|Z_k)P(Z_k) \quad (5)$$

上記のトピックのスコアは文章単位に計算したものだが, これを特許単位に集約し, そのスコアを特許の各属性情報(出願年や出願人など)の軸で集計すれば, トピックと属性との関係を分析することもできる. こうした方法で技術のトレンドを可視化したり, 出願人のポジショニングを可視化し, 技術戦略の検討に用いる試みもされている[野守 2018].

4.4 *diff*-PLSA の適用

diff-PLSA で採用する共起行列もその行列構成は 4.2 の通常の PLSA で使用した「名詞 3,020 語」×「係り受け 2,128 表現」の共起行列の構成と同様となるが, 先述の通り *diff*-PLSA では通常の PLSA で構築した実測頻度の共起行列に加え, 期待頻度の共起行列を構築する. 期待頻度の計算は式(1)の通りで, 総文章数 N は 229,598 件である. この期待共起頻度に対する実測共起頻度の比率の対数を取ることで *diff*-共起行列を構築し, これに PLSA を 4.2 と同様の方法で適用したところ, 最も評価の良い解として 50 個のトピックが得られた. 4.2 と同様に, 各トピックにおける名詞および係り受けの所属確率と, 4.3 の方法で計算される各文章のトピックのスコアから, 50 個のトピックに解釈を付けた. その一覧を表 3 に示す.

表 3 *diff*-PLSA で抽出された 50 個のトピック一覧

No.	トピック名	No.	トピック名
Z'01	エンジン制御	Z'26	電流・電圧の検出
Z'02	動力伝達	Z'27	温度、電流、充電量などの検出と制御
Z'03	差動機構などを備えた動力伝達の制御	Z'28	演算や推定、測定などのステップを含む方法
Z'04	回転運動	Z'29	情報の取得・提供(位置情報やバッテリ残量等)
Z'05	ロータ・ステータなどモータの構成	Z'30	スイッチなど操作装置
Z'06	モータ制御(トルク制御や回転数制御等)	Z'31	車両用灯具
Z'07	油圧ポンプなどを利用したモータ駆動	Z'32	掃除機
Z'08	ブレーキ	Z'33	基板の構成
Z'09	状態に応じた制御、運転者の操作補助	Z'34	回路の接続(電力変換回路等)
Z'10	コンバータとバッテリ昇降圧	Z'35	端子接続
Z'11	直流と交流の電力変換	Z'36	部品・装置の収容ケース・筐体
Z'12	回転力などの電気エネルギー変換	Z'37	部品・装置の配置
Z'13	エネルギー効率の向上	Z'38	パーツなどの移動、位置
Z'14	発電と蓄電	Z'39	構造の形成・方位
Z'15	電池モジュールの提供	Z'40	支持構造
Z'16	燃料電池	Z'41	装置やユニットの構成
Z'17	二次電池の構成	Z'42	システム・方法の構成
Z'18	バッテリの充放電	Z'43	その他方法
Z'19	充電システム	Z'44	組成物の製造方法(樹脂や電解液等)
Z'20	充電の接続	Z'45	機能性組成物・成形品(耐熱性や耐衝撃性等)
Z'21	非接触など受給電装置	Z'46	製造の効率化(小型化や低コスト化等)
Z'22	車両用空調など熱交換	Z'47	不具合の防止(損傷、感電、盗難等)
Z'23	冷却装置と放熱	Z'48	その他
Z'24	信号の入出力と検出	Z'49	タービン発電と船舶・飛行機への応用
Z'25	電気信号の取得と変換(センサ検出等)	Z'50	重力発電の活用による地球温暖化防止

4.5 通常の PLSA と *diff*-PLSA の比較

通常の PLSA の結果と *diff*-PLSA の結果を比較したところ, ① *diff*-PLSA の方が通常の PLSA よりも頻度の少ない名詞・係り受けでトピックが構成される傾向があること, ②*diff*-PLSA では通常の PLSA では抽出されないトピックが抽出されること, ③*diff*-PLSA の方が通常の PLSA よりもトピックの解釈が難しいことが考察された. 以下それぞれの考察について説明する.

① 頻度の少ない名詞・係り受けでトピックが構成されること

①の考察に関しては, 通常の PLSA と *diff*-PLSA で同様の解釈ができるトピックの内容を比較すると分かりやすい. 表 1 に示した通常の PLSA による「Z05:ブレーキ装置」「Z14:非接触受電など給電装置」に対応する *diff*-PLSA のトピックはそれぞれ表 3 の Z'08 と Z'21 であり, その内容を表 4 に示す. ブレーキに該当する Z'08 は, 通常の PLSA と比べて,「マスタシリンダ」や「液圧」など頻度は少ないがブレーキに関連するより具体的な表現が上位に位置している. また, 非接触受電などの給電装置に該当する Z'21 は,「非接触」が最も確率が高く, また「送電コイル」や「受電コイル」といった頻度は少ないが具体的な表現が上位に位置している. 電気自動車において非接触の充電システムは

今後重要な技術であり，その仕組みは駐車場の路面に設置した送電コイルと車両に搭載された受電コイルが重なり合うことで電磁誘導が発生し電力が供給されるという技術が代表的であり，*diff*-PLSA ではそうした技術のより具体的な表現がトピックの上位語となっている．つまり同様の意味を示すトピックでも *diff*-PLSA の方がより細かい技術を示すものが抽出されている．

また，各トピックにおいて所属確率の高い上位語の頻度 *n* が，通常の PLSA よりも *diff*-PLSA の方が少ないことを確認するため統計的検定を行った．各トピックにおいて所属確率の高い順に名詞および係り受けを並べたとき，累積確率が 50%になるまでの名詞および係り受けの平均頻度を検定用データとし，通常の PLSA と *diff*-PLSA におけるこの値のトピック平均の差について Welch の t 検定を実施した．その結果，通常の PLSA（n=34）では名詞が平均 1288.2 件（SD=686.5），係り受けが平均 85.6 件（SD=26.0），*diff*-PLSA（n=50）では名詞が平均 440.9 件（SD=137.5），係り受けが平均 58.8 件（SD=15.9）となり，名詞および係り受けの両方で有意水準 1%の違いがみられた．

表 4　*diff*-PLSA で抽出されたトピックの内容の例

トピックZ'08					
P(XZ)	n(X)	X：単語	P(YZ)	n(Y)	Y：係り受け
2.2%	112	マスタシリンダ	3.2%	32	基づく-発生
2.0%	73	ブレーキ液圧	2.6%	32	操作量-応ずる
1.6%	72	ブレーキ操作	2.6%	33	ブレーキ液圧-発生
1.6%	117	液圧	2.5%	53	制動力-発生
1.4%	779	ブレーキ	2.3%	49	ブレーキ-備える
1.4%	279	制動力	2.1%	20	備える-ブレーキ
1.3%	111	操作量	2.0%	92	車両-ブレーキ
…	…	…	…	…	…

トピックZ'21					
P(XZ)	n(X)	X：単語	P(YZ)	n(Y)	Y：係り受け
3.2%	180	非接触	3.0%	52	電力-受電
2.5%	78	送電コイル	2.7%	28	給電-電力
2.4%	160	受電	2.3%	35	受電-電力
2.4%	86	受電コイル	2.1%	25	電力-給電
2.2%	329	給電装置	1.8%	20	駐車装置-変化
2.0%	73	受電装置	1.7%	34	非接触-受電
1.8%	124	受電部	1.6%	21	供給-給電装置
…	…	…	…	…	…

② *diff*-PLSA でのみ抽出されるトピックがあること

②の考察に関しては，今回の方法では *diff*-PLSA の方がトピック数が多かったということもあるが，通常の PLSA のトピック数を 50 個に設定して計算した解も確認したところ，*diff*-PLSA のみに現れたトピックが複数存在した．その例を表 5 に示す．トピック Z'09（表 5 上）は，「シフトレンジ」や「パーキングレンジ」，「検出」，「停止」，「自動的に行う」といった表現で確率が高く，運転者の誤操作を抑制したり自動停止などの運転アシストに関する技術と解釈できる．トピック Z'29（表 5 下）は，「ナビゲーション」や「情報」，「目的地」，「位置情報」といった表現で確率が高く，位置情報を取得してドライバーにナビ情報として提供するなど，情報の取得と提供に関する技術と解釈できる．どちらも近年の自動車業界において付加価値を高める重要な機能といえる．

表 5　*diff*-PLSA でのみ抽出されたトピックの内容の例

トピックZ'09					
P(XZ)	n(X)	X：単語	P(YZ)	n(Y)	Y：係り受け
1.4%	50	シフトレンジ	1.9%	23	自動的-行う
1.0%	38	パーキングレンジ	1.7%	38	操作-行う
0.9%	147	検出結果	1.6%	22	駆動-停止
0.7%	1,200	停止	1.6%	26	動作-行う
0.7%	365	解除	1.6%	40	要する-時間
0.6%	34	キースイッチ	1.4%	46	停止-状態
0.6%	3,960	検出	1.2%	26	ブレーキ-作動
…	…	…	…	…	…

トピックZ'29					
P(XZ)	n(X)	X：単語	P(YZ)	n(Y)	Y：係り受け
1.2%	122	ナビゲーション装置	2.1%	48	情報-送信
1.1%	809	情報	2.1%	42	情報-含む
1.0%	165	目的地	1.8%	68	情報-取得
1.0%	111	位置情報	1.5%	31	情報-受信
0.9%	786	取得	1.5%	35	示す-情報
0.9%	819	送信	1.5%	126	情報-基づく
0.8%	558	表示	1.4%	25	情報-用いる
…	…	…	…	…	…

③ *diff*-PLSA のトピックの解釈が難しいこと

③の考察に関しては，*diff*-PLSA で抽出されるトピックは，頻度が少ない表現で構成される傾向にあるが故に，トピックの所属確率が頻度の少ない多様な表現に分散しやすく，通常の PLSA で抽出されるトピックよりも意味の解釈が難しい印象があった．実際に表 1 の通常の PLSA の結果と比較し，表 4,5 の *diff*-PLSA の結果は上位語の所属確率が小さい．通常の PLSA は頻度の多い代表的な表現に所属確率が集中するため，解釈はしやすいが，結果として典型的なトピックとなってしまう．この現象を解消する手法として *diff*-PLSA を開発したため，このジレンマは仕方のないものといえる．逆にそれだけより具体的で細かい要素で構成される個性の強いトピックが抽出されていることを意味している．トピックの構成内容だけでは解釈しにくいものは，4.3 で示したように各文章のトピックのスコアを計算し，各トピックと関連の強い具体的な文章を確認することが手掛かりとなる．

5.　まとめ

本稿では，テキストデータから典型的なトピックだけではなく，より個性的なトピックを抽出する手法として，従来の PLSA に対して *diff*-PLSA を提案した．本手法は，実測頻度の共起行列に加え，期待頻度の共起行列を構成し，それぞれの共起ペアにおいて期待頻度に対する実測頻度の比率の対数を取った値を持つ共起行列を構築し，これに PLSA を適用するものである．本稿では *diff*-PLSA の有効性を検討するため，特許の要約文データを例に，PLSA を適用した結果と *diff*-PLSA を適用した結果を比較した．その結果，通常の PLSA では，頻度の多い表現でトピックが構成される傾向にあり，全体を表す代表的なトピックが抽出されやすいが，*diff*-PLSA では，頻度の少ないより具体的で細かい表現もトピックを構成する重要な要素となっており，よりエッジの立ったトピックが抽出されていた．また *diff*-PLSA では通常の PLSA では抽出されないトピックも抽出されていた．

以上より，*diff*-PLSA は個性的なトピックを抽出する方法として適用価値があると思われるが，あくまでも全体を代表する典型的なトピックを理解した上で用いることが望ましく，実際には通常の PLSA と *diff*-PLSA を併用することが有効であると考えられる．

ビジネスにおいてテキストデータを分析し，その膨大な定性情報に潜む特徴を理解して有効なビジネスアクションを検討しようとする際，業務担当者が経験的に獲得するような典型的な特徴を改めて理解することも重要であるが，新たな気づきとなるインサイトを獲得するには，より個性的な特徴を抽出することも求められる．*diff*-PLSA ではそうしたトピックを通常の PLSA よりも効果的に抽出することができ，ビジネスインサイトの獲得に向けて有用な知識を提供することが期待できる．

参考文献

[野守　2014] 野守耕爾，神津友武: 三位一体アプローチによるテキストデータモデリング法の開発―宿泊施設のロコミデータを用いた評価推論モデルの構築―，2014 年度人工知能学会全国大会論文集，2014.

[野守　2018] 野守耕爾: テキストマイニングに複数の人工知能技術を応用した特許文書分析と技術戦略の検討，情報の科学と技術，Vol.68，No.7，pp.332-337，2018.

[Hofman 1999] Hofmann, T.: Probabilistic latent semantic analysis, Proc. of Uncertainty in Artificial Intelligence, pp.289-296, 1999.

[Kameya　2005] Kameya, Y., and Sato, T.: Computation of probabilistic relationship between concepts and their attributes using a statistical analysis of Japanese corpora, Proceedings of Symposium on Large-scale Knowledge Resources, pp.65-68, 2005.

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

看護記録における医療－介護連携に関する記述の基礎調査

Basic Analysis of Descriptions on Medicine-care in Nursing Records

上坂 静耶[*1] Seiya Uesaka　土方 嘉徳[*1] Yoshinori Hijikata　鈴木 斎王[*2] Muneou Suzuki

[*1] 関西学院大学商学部 School of Business Administration, Kwansei Gakuin University
[*2] 宮崎大学医学部附属病院 Faculty of Medicine, University of Miyazaki Hospital

Communication and collaboration between hospitals and nursing homes is required in these years. It is important to record the state of the patient at another facility. Thus, we analyzed the nursing records in University of Miyazaki Hospital to find content on the information of another facility and the state of the patient there. We analyzed nursing records of four nurses and found that there is little records according to the state of the patient at another facility.

1. はじめに

近年，医療と介護をとりまく現状は大きく変動しており，疾病構造の変化，医療ニーズの多様化，医療提供体制の機能分化によって，「病病連携」，「病診連携」，「医介連携」などの動きが広がっている．このような医療連携システムや在宅医療においては，大学病院など高度な機能を有する中核病院，市立病院など地域密着型の一般病院，かかりつけ医など地域診療所，老人ホームやデイサービスなどの介護施設，さらには患者の家庭まで含めた情報共有が重要となる．

電子カルテによって集められた情報は，院内の医療従事者間で共有されるだけでなく，地域の医療従事者との間でも共有されたり，さらに患者の転院や退院後に地域の医院や介護施設において活用されたりすることも期待されている．

しかし，各医療機関における電子カルテには，記載される病院以外の施設や家庭での患者の様子に関する記載など，各種連携にとって有益な情報がどれほど存在するのか分からないのが現状である．そこで中核となる大学病院の看護記録に，どれだけ上記のような各種機関の連携に関する情報の記載があるのかを明らかにする．

2. データ収集

看護記録について，日本看護協会は「看護記録および診療情報の取り扱いに関する指針」にて以下のように述べている．

1. 必要なことは漏れなく記述する
2. 必要でないことは一つも書かない
3. 無防備な看護記録の現実を改める
   - 個人的感情の記載
   - 感想，憶測，個人的見解
4. ［大原則］重大医療事故発生時には，記録方式を経時的記録に変える

看護記録の様式には，基礎情報（データベース），看護計画，経過記録，要約（サマリー）等がある．このうち経過記録にはいくつかの記述方法があり，そのひとつが SOAP と呼ばれるものである．これは以下の 4 項目にそって記載していく方式である．

S（subjective）：主観的情報 - 患者の話から得られた情報
O（objective）：客観的情報 - 身体診察・検査から得られる患者の様子や認識の状態
A（assessment）：評価 - 診断，O と S から考えられること
P（plan）：計画（治療） - 治療方針・内容，生活指導

しかし，看護記録は生活歴や検査歴，予約などのメモとして使用されていることもある．テキストデータである経過記録に明確なルールは存在しないため，細かい記述方法は記載者によって異なる．看護師は，患者が述べた言葉を覚えてメモを取り，最後にまとめて電子カルテに入力している[串間 12]．

本研究では，宮崎大学医学部附属病院で運用されているクリニカルパスの看護記録を分析した．入手した時点でデータ内の患者氏名および記載者氏名は匿名化されている．2013 年 12 月 21 日～2018 年 3 月 30 日の期間に SOAP に分けずにフリーテキストで書かれた看護記録と，2013 年 10 月 5 日～2018 年 3 月 30 日の期間に SOAP の形式で書かれた看護記録があり，この両方を分析の対象にした．フリーテキストのデータでは，記録数が 274,008 件，記録者総数が 995 名，患者総数が 8,718 名である．SOAP 形式のデータでは，記録数が 142,662 件，記録者総数が 549 名，患者総数が 4,778 名である．各データには，定型項目もあり，患者 ID，患者氏名，患者診療科，患者の性別，生年月日，文書番号，文書名称，タイトル，Subject，Object，Assessment，Plan，フリーテキスト，入院/外来区分，記載日付，記載者氏名，記載者所属部署，修正者氏名，修正者所属部署が記録されている（ただし，個人情報は匿名化されている）．

3. 分析方法

3.1 記録の調査項目

対象データの Subject, Object, Assessment, Plan および SOAP に分けずにフリーテキストで記載された記録において，関連機関との連携に関する以下の情報が書かれているかどうかを人手で調査する．

(1) 病診連携・医介連携に関する情報

他施設との連携に関する記述が存在するかどうかを調査者が判断しタグ付けする．具体的には，転院前に治療していた病院や入院前に入所していた介護施設における患者の様子に関する記述や，退院後の自宅における患者の生活状況に関する記述，患者の家族が発した言葉の記述などが挙げられる．

連絡先：土方 嘉徳，関西学院大学 商学部，
兵庫県 西宮市 上ケ原一番町 1-155 ，0798-54-6205，
contact@soc-research.org

また，上記情報と関連のある情報（看護師の記録の特徴）として，以下のものを人手または計算機により調査する．

(2) 看護師による主観的記述

看護師による主観的な記述が存在するかどうかを調査者が判断しタグ付けする．

(3) 専門用語の数

医療関係の専門用語がどれだけ使用されているかを計算機により調査する．具体的には，看護記録中のテキストに対して，医療用語辞書（ライフサイエンス辞書 [金子 93]）中の単語で検索を行い，適合した単語の数を記録する．

(4) 形容詞の数

主観的記述に用いられることの多い形容詞がどれだけ使用されているかを計算機により調査する．具体的には，看護記録中のテキストに対して，形態素解析を行い形容詞の数を記録する．

(5) 記載の量

看護記録の記載量（テキストサイズ）を調査する．

3.2 看護師の属性

病診連携・医介連携に関する情報を書くかどうかは，看護師のキャリアや経験に依存するかもしれない．そこで，記録した看護師の年齢と勤続年数も調査する．これらの情報は，職員リストのデータ（個人情報は匿名化されている）から得た．

4. 分析結果

本分析は人手によるタグ付けが含まれるため，全データを対象とするには時間が必要である．そのため，まずは内科（5 階西病棟・第 1 内科）で受診しているある患者に対して，看護師 4 名（それぞれ A, B, C, D）がフリーテキスト形式で記述した記録を対象に分析を行った．全員が女性で，それぞれの勤続年数は A が 5 年，B が 3 年，C と D は 2 年であった．

調査結果を表 1 に示す．4 人の看護記録においては，他医療機関における患者の様子についての記述は見られなかった．しかし退院指導に関する記述の中で，次回外来日時およびその病院，紹介状の有無に関する記述が 15 件見られた．各看護師の記述総数に占める，自宅での患者の様子が書かれた割合については，どの看護師も 1%に満たなかった．家族の言葉の記述数は A が最も多い 27 件，B, C, D はそれぞれ 6, 3, 5 件であった．

表 1 第 1 内科の医介連携に関する記述の数

看護師	A	B	C	D
勤続年数	5	3	2	2
記録総数	728	539	508	324
他施設での患者の様子	0	0	0	0
他施設との連携情報	13	2	0	0
自宅での患者の様子	5	1	0	2
自宅での様子記述割合(%)	0.687	0.186	0.000	0.617
家族の言葉	27	6	3	5
家族の言葉記述割合(%)	3.709	1.113	0.591	1.543

5. 考察

まずは他医療機関における患者の様子についての記述が見られなかったことについて考察する．今回の記録は，ある病棟（5 階西病棟）の入院患者に関するものであった．入院にいたるまでの経緯はそれまでのカルテに記載されているはずなので，本人に確認するまでもなかった可能性がある．今後は外来患者についての記録も分析する必要がある．

次に，医介連携に関する情報が記載されていることについて考察する．今回の記載者はほぼすべて大学病院で勤務する看護師である．患者が大学病院を退院し他の施設や自宅で治療・介護を進める際には看護師が退院指導を行う．このときの記録として次回外来日時およびその病院，紹介状の有無に関する情報を記載する．また，看護師は病診連携に関する知識を持ち合わせており，患者が大学病院を退院し他の施設で治療・介護を進める際に活用できることを理解して記述していると考えられる．

次に，患者の家族が発した言葉に関する記述数が目立って多かったことについて考える．看護師 A が記述した 27 件のうち 19 件がインフォームドコンセントについての記録だった．看護師は医師が説明する際に同席することも多く，このときの様子や反応を細かく記録している．インフォームドコンセント以外にも，看護師は訪室の際に面会に来ていた家族とコミュニケーションを図っていることがわかった．患者本人の意思や家族の思いについて詳細を記録していたこともあった．

最後に，各特徴をもつ記述数と看護師の属性について考察する．今回このような記述を行っていた看護師は勤続年数が最も長い看護師 A であった．実施した処置や観察初見の記載に終始することなく家族とのやりとりまで記録していたのは，長年の勤務経験によるものが大きいと考えられる[山崎 15]．

6. まとめ

今回の調査では，医療―介護連携に関する情報は，日付や紹介状の有無など，事務的な内容に限られて記載されていた．また，他医療機関における患者の様子に関する記載は見られなかった．一方，退院後に有益になりそうな家族の発した言葉に関する記述は存在した．今回の調査結果から，大学病院においては，医療－介護連携に関する有用な記述は看護記録に，ほとんど残されていない可能性があることが分かった．しかし，病院での診療行為だけでなく，患者の家族の言葉が記録として残されていたことについては，医療－介護連携を進める上で，助けとなるかもしれない．今後は，主観的記述に関連する項目にも注目し，データ数を増やして分析していく予定である．

謝辞

本研究は JSPS 科研費 17KT0084 の助成を受けたものです．

参考文献

[串間 12] 串間宗夫ほか: 電子カルテテキストデータに関する一考察．人工知能学会第 1 回インタラクティブ情報アクセスと可視化マイニング研究会 (2012)

[山崎 15] 山崎友義，串間宗夫，荒木賢二，鈴木斎王: 論文 TETDM を用いた電子カルテテキストデータ分析．第 29 回人工知能学会全国大会 (2015)

[金子 93] 金子周司ほか: ライフサイエンス辞書プロジェクト．https://lsd-project.jp/ (1993)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

深層学習における学習ネットワークからの分類パターンの抽出

Extraction of Classification Patterns from Deep Learning Networks

安藤 雅行 *1 Masayuki Ando　河原 吉伸 *2*4 Yoshinobu Kawahara　砂山 渡 *3 Wataru Sunayama　畑中 裕司 *3 Yuji Hatanaka

*1滋賀県立大学大学院工学研究科
Graduate School of Engineering, The University of Shiga Prefecture

*2九州大学 マス・フォア・インダストリ研究所
Institute of Mathematics for Industry, Kyushu University

*3滋賀県立大学工学部
School of Engineering, The University of Shiga Prefecture

*4理化学研究所 革新知能統合研究センター
RIKEN Center for Advanced Intelligence Project

In deep learning, there is a problem that concrete classification patterns for deriving reasons for classification are often incomprehensible. In this paper, we propose a classification patterns extraction system from deep learning networks and verified the effectiveness of the system. The proposed system takes out learning networks from the learning result of deep learning and extracts classification patterns from the learning networks. Then the system displays the extracted classification patterns so that users of the system can interpret the learning networks. In verification experiments, the significance of the extracted classification patterns was estimated by chi-square test. The results showed that users of the system can extract classification patterns effective for interpretations of the learning networks by using the proposed system.

1. はじめに

インターネットの普及などによって世の中で増大する文章データへの対処方法として，深層学習を用いたテキストマイニングシステムが注目されている [1][2]．深層学習は，一般に多層から構成されるニューラルネットワークを用いた学習を指す．

その一方で，深層学習は，その出力を導いた根拠についての解釈が困難であることも知られている．テキスト分野においても，学習ネットワークの解釈を行うことで深層学習の分類基準をより深く理解できれば，例えば医療分野において新人とベテランの書いた電子カルテの違いから，良い電子カルテを書く方法を容易に理解したり，企業においても良い報告書や企画書を書く方法を短時間で習得できるなど，深層学習の新しい活用が期待できる．

深層学習の分類基準を示す研究としては，アテンションと呼ばれる手法を用いた研究 [3][4] が注目されている．しかし，アテンションは内部でどのような学習が行われているかは考慮していない．一方で画像に関してだが，深層学習の学習ネットワークの解釈に注目したものがある [5]．また，ニューラルネットワークを用いた研究では，中間層の役割と重みの持つ意味について考察したもの [6] や中間層が学習から，入出力の様々な関係を得ることを目指したもの [7] があり，中間層に注目すれば，出力に関して意味のある情報が得られることを示している．また，重みに注目した研究として，重みを用いて入力データ間の相関関係を抽出する研究 [8] があるが，分類基準を提示する目的で重みに注目したものは存在しない．

そこで本研究では，RNN（Recurrent Neural Network）を使用し，テキスト集合の学習によってネットワークの層に付けられた重みの値を取り出し，学習ネットワークに対する解釈を行うために，それぞれの出力特有の特徴となる分類パターンの抽出を行うシステムを提案する．

連絡先: 安藤　雅行，滋賀県立大学大学院工学研究科先端工学専攻，〒 522-8533 滋賀県彦根市八坂町 2500，oh23mandou@ec.usp.ac.jp

2. 深層学習の重みを用いたテキストの分類パターン抽出システム

本章では，本研究で開発した深層学習の重みを用いたテキストの分類パターン抽出システムについて，システムの構成とその詳細について述べる．

2.1 分類パターン抽出システムの構成

分類パターン抽出システムでは，まず，図 1 に示すように，各分類先ごとにラベル付けしたテキスト集合を RNN にて分類し，その分類先を導いた学習ネットワークと，学習ネットワーク上の重みから，提案システムの分類パターンの抽出・可視化処理によって各出力 (分類先) を導くネットワーク上のパスと，パス上の各ノードの情報（そのノードで学習された単語）の決定，表示を行う．最後に，システムの利用者は，システムによって得られた学習ネットワークの表示を自分が見やすいように調整し．分類パターンを抽出する．そして分類パターンの意味を理解しやすくするための機能を利用できる．

2.2 深層学習による学習ネットワークの形成

2.2.1 文中の単語のベクトル化

深層学習で学習を行う前に，テキストデータは文中の単語を抽出したあと，単語は One hot 法 [9] と呼ばれる手法に従い単語ベクトルの羅列に直される．そして，文中の各単語を単語ベクトルに置き換え，深層学習への入力データとする．なお，抽出される単語は名詞する．これは，文章の特徴とその順序関係をより学習・抽出しやすくするためである．

2.2.2 学習によるネットワークの重み付け

One hot 法によって単語ベクトルの羅列に変換され，さらにラベル付けされたテキストデータは，RNN の学習にて，そ

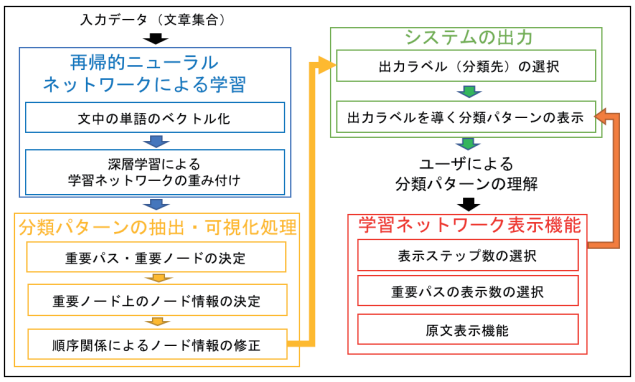


図 1: 分類パターン抽出システムの構成

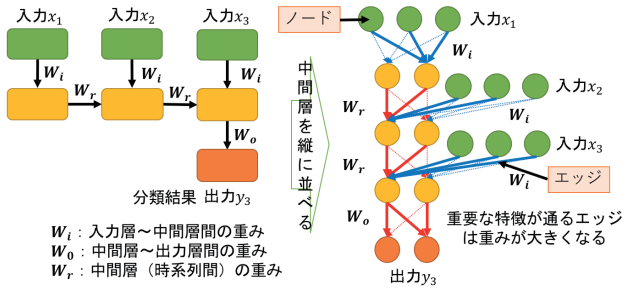


図 2: RNN の学習ネットワークと学習の様子

れぞれの出力ラベル（分類先）を導くネットワークへの重み付けがされていく．その様子を図 2 に示す．入力文章は各単語がベクトル化され，タイムステップごとに単語ベクトルが順番に入力されていく．また，RNN での分類時は，最後の単語が入力されたタイミングで，出力層から出力される．

2.3 学習ネットワークからの分類パターンの抽出・可視化処理

本提案システムの分類パターンの抽出･可視化処理では，RNN によって得られた学習ネットワークから，各出力を導く最も関係が強いパス（ネットワーク上のエッジの繋がりの線）を決定する処理を行う．この出力と繋がりが強いパス（重要パスと呼ぶ）の決定について，図??に示した 4 層全結合型ネットワークモデルを例として，具体的な手順を述べる．

まず，ある分類先（図??の例では層 D のノード b）に到達するパスについて，パス上のエッジについた重みの積で定義される，重要度と呼ばれる値を算出する．図??の入力層のノード i からの太矢印のパス上のエッジに付いた重みを $W1_{a,i}$，$W2_{i,j}$，$W3_{j,b}$ とすると，このパスの重要度 $B_{a,i,j,b}$ は以下の式 (1) で導かれる．

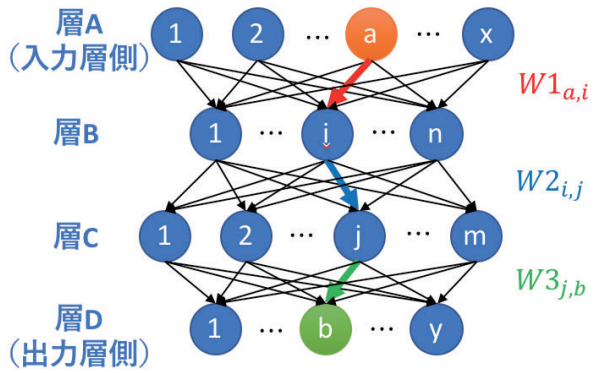


図 3: 4 層全結合型ネットワーク

$$B_{a,i,j,b} = W1_{a,i} \times W2_{i,j} \times W3_{j,b} \tag{1}$$

そして，式 1 で出力 b に到達する全てのパスの重要度を計算して比較し，最も値の大きいパスを，重要パスと決定する．出力 b の重要パスの重要度を MB_b として，その計算式 (2) を以下に示す．図 2 の右図のように，出力層と中間層間の重み $\boldsymbol{W_o}$ のあと，中間層間の重み $\boldsymbol{W_r}$ を繰り返し辿ることで，過去の中間層を遡ることになる．

$$MB_s = \max_{i,j,b} B_{a,i,j,b} \tag{2}$$

次に，重要パス上のノードについて，ノードの情報を決定の処理を行う．ノード情報の決定方法は，出力ごとの重要パスを決定した時と同様に，入力層からノード情報を決定したい中間層ノード間の重み $\boldsymbol{W_i}$ の大きさから，ノード間の結びつきの強さを求める．ただし，図 2 に示す様に，各中間層にはそれぞれ個別の入力層が対応している．ここで，RNN の入力には One hot 法による単語ベクトルを用いているため，入力層ノードにはノードごとに 1 種類の単語が対応していることになる．そこで，重要パス上の入力層ノードの単語を参照し，その単語をノードの情報と決める．最後に，ノード情報の単語について，前後の中間層ノード情報の単語と，原文中でその順番で表示されているかどうかを確認し，されていないなら単語の重要度を下げる．

こうして各出力を導く重要パスとパス上のノード情報を表示した学習ネットワークから，過去の中間層の重要ノードから現在の中間層の重要ノードの単語より，時系列を考慮した特徴の並びとしての分類パターンを抽出することができる．

2.3.1 出力ラベルを導く分類パターンの表示

本研究で開発した分類パターンの抽出システムでは，分類先に強く結びつく，重要パスの集合としての学習ネットワーク上に重要ノードの情報が表示される．例として，5 種類のお菓子の作り方に関するテキストの分類を行った場合の，システムのメイン画面を図 4 に示す．表示分類先は「マカロン」とする．このネットワークは，RNN が学習した情報を，過去から順番に中間層上に表示したものである．過去から出力層直前までの中間層の重要ノード情報（単語）の並びは，分類パターンを示し，重要パス 1 本につき 1 つの分類パターンが表示されていることになる

図 4 では，入力層を除いて選択したステップ数（図 4 では 2）だけ中間層と，出力層が表示され，出力層には分類された分類先を示す出力ノードが表示されている．また，分類先を示す出力ノードには，そこから選択したステップ数の長さで，選択した重要パス数だけ重要パスが表示されている．そして，重要パス上の各中間層ノードには，ノード情報として単語が表示されている．

2.4 システム上での学習ネットワーク表示機能

システムには，利用者が分類パターンの抽出・理解を行いやすいように，その表示内容を変更できる機能がある．その主なものを表 2 に示す．

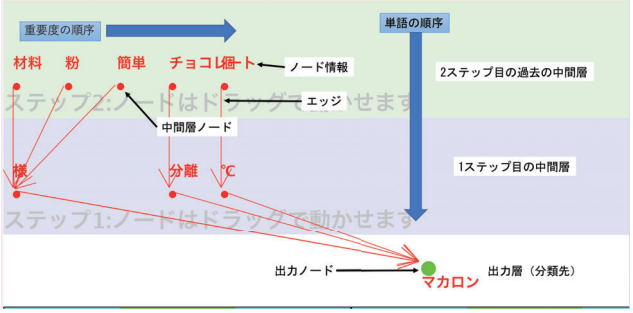


図 4: システムの画面

表 1: 学習ネットワーク表示機能

機能名	効果
表示するステップ数の選択	RNN において，何ステップ過去までの中間層を表示するか決定する
重要パスの表示数の選択	分類先ごとに，何本の重要パスを表示するか決定する
原文表示機能	分類パターン（順序を考慮した単語の組み合わせ）が，原文中でどのように出現しているかを表示する

3. 分類パターン抽出システムの有効性の検証実験

本章では，深層学習の重みを用いた分類パターンの抽出システムにより，学習ネットワークから抽出できる分類パターンが，テキストデータを理解する（学習ネットワークの解釈を行う）ために有効かを検証した実験について述べる．なお，本実験では，分類パターンを「2 種類の異なる単語の，順序を考慮した組み合わせ」と定義する．単語を 2 種類にした理由は，最も基本的な分類パターンであるためである．

3.1 実験準備

3.1.1 使用テキストデータと学習モデル

分類パターン抽出の対象とするテキスト「童話 5 種類」について，詳細を表 2 に示す．また，計算上の分類パターンの総数（抽出単語数 × 抽出単語数 − 抽出単語数）とその中で「ある分類パターンが，ある分類先に特有のものである」という仮説を有意水準 5%で検定したカイ二乗検定で，有意性があると推定された分類パターン数を表 3 に示す．

続いて，深層学習として使用した RNN モデルの概要を表 4 に示す．システム上での表示設定は，学習ネットワーク表示機能により，表示ステップ数 2，重要パス数 20 とする．また，テキスト「童話 5 種類」に対する RNN の分類精度は 100%であった．

3.2 実験手順

実験の対象者は著者 1 名とした．対象者はテキスト「童話 5 種類」について，本研究で開発したシステムを用いて，全ての分類先で分類パターン抽出システムによって抽出された分類パターン 20 個について，カイ二乗検定で有意性があるかどうか，ある場合，カイ二乗検定の順位で何位にあたるのかも調べ考察を行った．なお，カイ二乗検定の順位とは，カイ二乗検定で有意性のあると推定された分類パターンの集合を，カイ二乗値が高い（より強く有意性があるとされる）順に並べた時の順位である．

*1 Google：https://www.google.co.jp で「童話名　あらすじ」と検索し，表示された上位 50 位のサイトの本文をテキストとして利用

表 2: テキストデータの詳細

データ名	内容
童話 5 種類	日本の童話「かぐや姫」「鶴の恩返し」「さるかに合戦」「桃太郎」「浦島太郎」の，それぞれの概要やあらすじなどについて書かれたテキストをネット上 *1 から 1 種類あたり 50 テキストずつ用意した．

表 3: 有意性のある分類パターンの総数

分類先名	カイ二乗検定での分類パターン総数	分類パターン総数
かぐや姫	93,211	15,527,540
鶴の恩返し	9,208	15,527,540
さるかに合戦	24,839	15,527,540
桃太郎	46,749	15,527,540
浦島太郎	53,334	15,527,540
平均	45,468	15,527,540

3.3 結果と考察

抽出された分類パターンで，有意性がないとされた分類パターンについての考察を行うため，有意性がある・ないとされた分類パターンの個数を図 5 に示す．

図 5 より，まず有意性があるとされた分類パターンの個数と，有意性がないとされた分類パターンの個数の比率は，4 対 6 ほどであることがわかる．その中で，有意性があるとされた分類パターンについて見ると，どちらのテキストも，ほとんどがカイ二乗検定での分類パターンの上位 50%以上に含まれることがわかる．これは，カイ二乗値の高い，より有意性があるとされた分類パターンと，有意性が全くないとされた分類パターンが混ざって抽出されていると言える．

その中で，有意性がないとされた分類パターンが 60%以上と最も多い「浦島太郎」を例に考察する．「浦島太郎」の抽出された分類パターン上位 10 個を表 5 に示す．有意性がないと判断される要因は，他のテキストでも頻度が高い，もしくはそのテキスト内での頻度が低いことが原因である．しかし，それでもシステム上で重要度の高い分類パターンとして抽出されたということは，有意性があるとされた分類パターンと合わさることで，よりテキストへの理解を深めることができる補助的

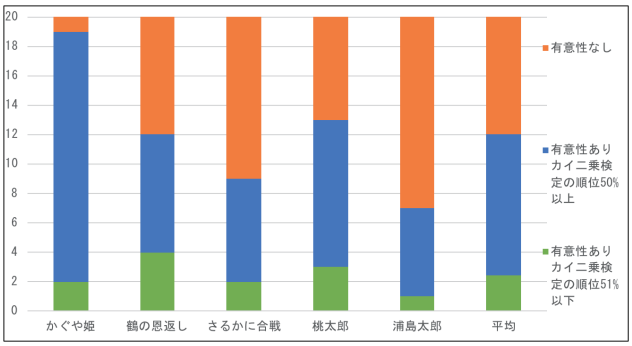


図 5: 抽出された分類パターン

表 4: 学習モデルの詳細

データ名	抽出した単語集合	入力層ノード数	中間層ノード数	出力層ノード数
童話 5 種類	名詞 3,941 種の単語	3,941	50	5

表 5: 抽出された分類パターン（「浦島太郎」）

	抽出された分類パターン			
順位	分類パターン	「浦島太郎」内の頻度	「浦島太郎」以外での頻度	カイ二乗検定での順位
1 位	世界→気	12	4	34,300
2 位	日→展開	6	3	0
3 位	風土記→日本	12	0	2,118
4 位	物語→日本	18	11	0
5 位	日→日本	16	10	0
6 位	長期間→ストーリ	3	0	0
7 位	集→嶋	3	0	0
8 位	漁師→共通	3	0	0
9 位	年→日本	19	11	0
10 位	亀→女性	16	0	964

な役割を持つと考えれる．なお，表 5 で有意性がない分類パターンのカイ二乗検定での順位は 0 と表記している．

表 5 をみると，まず単語「日本」が，4 つの分類パターンで共通しており，そのうち 3 つの分類パターンに有意性がなく，1 つが有意性があるとわかる．また，表 3 より，有意性のある分類パターンのカイ二乗検定での順位は，53,334 中 2,118，つまりカイ二乗検定での順位の上位 4%以内と高い順位であり，ここから，強い有意性を持つ 1 つの分類パターンを基準に，3 つの有意性のない分類パターンがテキストへの理解の補助を行なっていると仮定する．例として，順位 3 位の「風土記→日本」（有意性あり）から原文表示機能を用いて得られるテキストへの理解「丹後国の**風土記**など，**日本**各地で逸文されている」に，順位 4 位の「物語→日本」から得られる「浦島太郎**物語**は，**日本**書紀にも記述が見られる」と順位 9「年→日本」（有意性なし）から得られる「明治 29 **年**に厳谷小波氏が書いた物語をきっかけに**日本**中に広まった」を合わせることで，「浦島太郎物語は丹後国の風土記や日本書紀など，日本各地で逸文されており，明治 29 **年**に書かかれた物語をきっかけに，**日本**中に広まった」という理解が得られる．そして，この理解は「浦島太郎」特有のものである．よって，システムでは，有意性の高い分類パターンと，有意性がないが，それらを補助する役目を持つ分類パターンが得られていると言える．

以上より，実験結果から，本研究で開発したシステムでは，深層学習の学習ネットワークから，学習ネットワークの解釈に役立つ，分類先に対して特有の特徴を表す，有効性のある分類パターンの抽出が可能であると言える．そして，カイ二乗検定での分類パターンと比較することで，カイ二乗検定では有意性がないとされた分類パターンでも，有意性のある分類パターンの補助として役立っていることが考察によって判明し，カイ二乗検定では得られない隠れた有効性のある分類パターンを抽出することができると結論つけた．

4. おわりに

本研究では，複数のテキストデータの分類を深層学習である RNN で行い，学習ネットワークの解釈を行うための，分類パターンの抽出システムの開発を目的とした．本研究の特徴として，重みを辿ることで，RNN の学習ネットワーク内の情報の伝達を過去に向かって探索している点が挙げられる．システムの有用性を確かめる検証実験では，があると推定された分類パターンと，本システムで抽出した分類パターンをのカイ二乗検定で有意性を検定することで，本研究で開発したシステムにより分類パターンの抽出が容易になり，テキストの特有の特徴などを理解しやすくなったと結論づけた．今後の研究では，理解した内容からより容易にテキスト自体への解釈を行えるよう，分類パターンの抽出・可視化アルゴリズムや，システム上での表示を検討していきたい．

参考文献

[1] ボレガラ ダヌシカ，“自然言語処理のための深層学習”，人工知能学会誌，Vol.29, No.2, pp.195–201, 2014

[2] Ebru Arisoy, Tare N. Sainath, Brian Kingsbury, Bhuvaba Ramabhadran, “Deep Neural Network Language Models”, In Proceedings of the NAACLHLT Workshop, Will We Ever Really Replace the N-gram Model?, pp.20-28, 2012

[3] M Daniluk, T Rocktaschel, J Welbl, S Riedel, “Frustratingly Short Attention Spans in Neural Language”, ICLR, 2017

[4] A.Vaswani, N.Shazeer, N.Parmar, J.Uszkoreit, L.Jones, A.N. Gomez, L. Kaiser, and I. Polosukhin, “Attention Is All You Need”, CoRR, vol. abs/1706.03762, 2017

[5] 西銘 大喜，“ディープニューラルネットワークによる画像からの表情表現の学習”，人工知能学会全国大会論文集 3L4-3，pp.1-4, 2015

[6] 立石 雅彦，山崎 晴明，“手書き数字認識における改装ニューラルネットワークの中間層に関する考察”，情報処理学会論文誌，Vol.30，No.10，pp.1281-1288, 1989

[7] 松倉健太郎，村田 昇，“ニューラルネットワークの中間層における独立な特徴量の抽出”，電子情報通信学会技術研究報告，Vol.106，No.102，pp.63-67, 2006

[8] Michael Tsang, Dehua Cheng, Yan Liu, “DETECTING STATISTICAL INTERACTIONS FROM NEURAL NETWORK WEIGHTS”, ICLR, 2018

[9] M. Feurer, A. Klein, K. Eggensperger, J. Springenberg, M. Blum, and F. Hutter. “Efficient and robust automated machine learning”, In Neural Information Processing Systems (NIPS), 2015

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

深層学習を用いた対話文への「かわいさ」の付与

Attachment of "Cuteness" to Dialogue Sentences Using Deep Learning

後藤 優介 *1 Yusuke Goto　砂山 渡 *2 Wataru Sunayama　畑中 裕司 *2 Yuji Hatanaka　小郷原 一智 *2 Kazunori Ogohara

*1滋賀県立大学大学院工学研究科
Graduate School of Engineering, The University of Shiga Prefecture

*2滋賀県立大学工学部
School of Engineering, The University of Shiga Prefecture

Nowadays, automatic generation of dialogue sentences using deep learning has been conducted. Automatic generation systems of dialogue sentences require not only sequences of conversation but also characteristics of conversation partner. Because, it makes users feel charms and personalities in conversation. However, previous studies often transformed the sentence's characteristics to another one which easy to define transformed vocabularies, for example, the dialect and the style of samurai. Therefore, this study will transform the sentence's characteristics to another one which hard to define transformed vocabularies. Especially, the purpose of this study is that give the cuteness to sentences. We considered how to give the cuteness to sentences, and evaluate sentences that can give the cuteness or not.

1. はじめに

近年，Apple 社の Siri や Amazon 社の Alexa を皮切りに，対話型システムの注目度が増している．これらのように，特定のタスクを実行する事を目的とした対話型システムの事をタスク指向型対話システムと呼ぶ．一方で，SoftBank 社の Pepper や Microsoft 社のりんなに代表される，ユーザとコミュニケーションを取る事を目的とした対話型システムの事を非タスク指向型対話システムと呼ぶ．これは主に，コミュニケーションロボットでの活躍が期待され，AI と対話できる事で，より AI に対して親しみを持ってもらう事が期待できる [1]．こういった非タスク指向型対話システム固有の問題として，無個性な対話による対話の質の低下が挙げられる．対話の応答文が無個性で機械的なものであると，ユーザに AI との間に壁を感じさせる．それにより，ユーザの対話に対する意欲が失われる事や，AI に対して親しみを感じる事が難しくなる事が懸念される．

そこで本研究では，付与するキャラクタ性として，具体的なキャラクタではなく，人間の感性の 1 つである「かわいさ」を挙げた．そして，コーパスの調査に加え，深層学習によって学習されたモデルの学習内容を分析する事で，かわいさが現れている文にどのような特徴があるかを網羅的に調査した．その後，調査で得られた要素を自動で付与する事で，文にかわいさを付与するアルゴリズムを作成した．また，そのアルゴリズムに，深層学習による判定を取り入れる事で，ユーザによってどれくらい強くかわいさを付与するかを調整できるようにした．これにより，従来手法と比べて，ユーザの好みに沿った強さのかわいさを，文中の幅広い範囲で付与できる事が期待される．

2. 関連研究

2.1 文に現れるキャラクタ性とそれが現れる場所を調べた研究

文のどの部分にキャラクタ性を感じさせる要素を埋め込む自由度があるかを調べた研究 [2] がある．この研究では，内容表現 (名詞，動詞，形容詞，副詞など) と機能表現 (助詞，助動詞，接続詞など) の両方にキャラクタ性を埋め込む自由度があると結論付けている．しかし，内容表現は非常に自由度が高く，単純な置換は難しいとしている．

一方で，機能表現は同義表現への置き換えが比較的簡単であり，キャラクタ性を容易に埋め込む事ができるとしている．本研究でも，単語の同義表現への置換を扱う事になるので，比較的置換が容易とされていた機能表現は優先的に取り入れた．同時に，内容表現内でも比較的自由度が低いとされた形容表現にも着目した．

2.2 文にキャラクタ付けを行う研究

文節機能部に着目し，人手で書き換えたコーパスを分析する事で見つけ出した置換ルールを，確率に乗っ取り自動適用する研究 [3] がある．これにより，変換により感じる違和感を 10% 以下に抑えつつ，人間が 80%の確率で判別できるキャラクタ性変換を行う事に成功している．また，よりコストを抑えた変換を行うため，変分オートエンコーダと注意機構を用いてキャラクタ性を付与する研究 [4] がある．この研究では，注意機構を用いる事で，生成モデルを用いる場合でも変換する単語を制御しつつ，文全体の変換を行う事ができるという事を示している．

しかし，これらの研究における変換は，機能表現のみを対象とした置換に限られており，キャラクタ性を感じさせる要素を網羅的に適用できているとは言い難い．また，これらの研究では「関西弁」「侍風」といった具体的なキャラクタ性の付与に留まっており，キャラクタ性の付与が直接魅力的な対話につながっているとは言えない．そこで，本研究では，より人間の感性に近いキャラクタ性である「かわいさ」を付与する場合について調査を行った．調査するコーパスとして人手による変換を用いていない既存のものを選び，要素を網羅的に調査する事で，自動適用できる変換ルールを調査した．

3. 「かわいさ」を付与する変換ルール

3.1 「かわいさ」の定義

本節では，本研究で付与するキャラクタ性である「かわいさ」について定義する．「かわいい」という単語は，辞書的には「愛らしい，小さい」という意味がある．しかし，現代日本において，「かわいい」という単語は様々な意味で用いられるようになっており，物理的・数値的特徴からは明確に分類する

連絡先: 後藤優介，滋賀県立大学大学院工学研究科電子システム工学専攻，〒 522-8533 滋賀県彦根市八坂町 2500

事ができない [7]．その人個人の感性によって，どういったものが「かわいい」のかは変わってくる．

「かわいい」の意味について調べた研究 [5][6][7] によると，心理学的に人が他者に対して「かわいい」と感じる要素は，大きく「共感性」「不完全性」「親和性」の 3 つの要素に集約できると言われている．それぞれの用語の定義は以下の通りである．

- 「共感性」: 他者の状況や感情に共感し，他者志向の温かい気持ちを持たせる性質.
- 「不完全性」: 幼さ・弱さから，守ってあげたいという気持ち (保護欲) を感じさせる性質.
- 「親和性」: 相手と親密な関係を維持し，仲良くなりたいと感じさせる性質.

このうち，共感性と親和性については，自分と他者の関係性を評価して生まれる性質であり，相対的なものであるため，統一した基準での定義が難しい．そのため，本研究では不完全性を取り上げ,「文全体から読み手に不完全性を感じさせる性質」を「かわいさ」と定義する．また，外見的な要素を含まない不完全性の要素として,「幼さ (未熟さ)」と「精神的な弱さ (気の弱さ)」の 2 つを挙げ，それにつながる要素を調査していく．

本研究では,「かわいさ」を含む学習用のテキストデータとして，ソーシャルゲーム「アイドルマスターシンデレラガールズ」の登場人物 183 人のうち,「正統派かわいい」キャラクタとして位置づけられている 67 名のセリフを合計 50,113 件 [12] を用いることとした．

また,「かわいさ」を含まない学習用のテキストデータとして，老若男女 199 人の日常会話を録音し，それを文字に起こしたデータセット「名大会話コーパス」[9][10][11] に収録された発話文のうち，20 文字以上のものを 55,054 件を用いることとした．

3.2 「かわいさ」を付与する変換ルール

「かわいさ」を付与する変換ルールの作成に際して，前節で用意したデータをもとに，深層学習により「かわいさ」を含む文と含まない文とに分類する学習を行った．その結果を考察することで，変換ルールを用意した．

3.2.1 「幼さ, 未成熟さ」についての変換ルール

幼さや未成熟さを感じさせる表現として，以下の要素がある事がわかった．

- 幼さや柔らかさを感じさせる表記や表現を使う.
- 相手との距離を縮め，相手を頼ろうとする表現を使う.
- 感情を表に出し，裏表を感じさせない表現を使う.
- 敬語の使用は最低限にする.

これらの要素をもとに，分類に寄与する単語を調べた上で，表 1 に示す「幼さ，未成熟さ」を付与する変換ルールを作成した．

3.2.2 「精神的な気の弱さ」についての変換ルール

精神的な気の弱さを感じさせる表現として，以下の要素がある事がわかった．

- 自分の意見を強く主張できず，断定出来ない事を感じさせる表現を使う.
- 自分の考えを相手に強制できない事を感じさせる表現を使う.

表 1: 「幼さ，未成熟さ」を付与する変換ルール

変換ルール	変換対象の単語	変換先の単語
名詞をひらがなに置換	漢字のみの名詞 (10%の確率で適用)	その単語をひらがなに置換
副詞をオノマトペに置換	体言か名詞の前にある一般副詞 (30 %の確率で適用)	評価の高い副詞 27 件のうち 1 つと置換
相槌・同意・終助詞長音化	相槌や同意を示す表現 (うん，なるほど等) または終助詞	対象の単語の後に「ー」を挿入
語尾表現置換	です，ます等の丁寧な文末表現とその前の単語 文末記号の前の，名詞や副詞では無い単語	よ，ねに置換もしくは挿入 必要に応じて前の単語も活用させる
感動詞の挿入・置換	感動詞 文頭・文末 (50%の確率で適用)	評価の高い感動詞 (5 件) のうち ランダムで 1 つと置換もしくは挿入
名詞にさんを付ける	一般名詞もしくは固有名詞 (15 %の確率で適用)	対象の単語の後に「さん」を挿入

表 2: 「精神的な気の弱さ」を付与する変換ルール

変換ルール	変換対象の単語	変換後の単語
断定表現を婉曲表現に置換	言い切る表現 (〜ない等)	対象の単語の後に「かも」を挿入
曖昧な表現を追加	文末の動詞 (50%の確率で適用)	対象の単語の後に「かも」を挿入
命令をお願いする表現に置換	命令形である動詞	対象の単語を基本形に置換し， 「〜てほしいな」「〜よ」を挿入
逆接の接続詞を削除	逆接の接続詞 (しかし，だけど等)	対象の単語を削除，必要に応じて句読点を挿入
強い否定を緩和する表現の追加	強い否定を表す副詞や形容動詞 (ちっとも，全く等)	「多分」「できれば」と置換

- 相手の意見に対して，強く否定できない事を感じさせる表現を使う.

これらの要素をもとに，分類に寄与する単語を調べた上で，表 2 に示す「精神的な気の弱さ」を付与する変換ルールを作成した．

4. 深層学習を用いて「かわいさ」を付与するシステム

本章では，前節で用意した変換ルールをもとに，対話文に「かわいさ」を付与するシステムについて述べる.「かわいさ」を付与するシステムのフローチャートを図 1 に示す．以下で，フローチャート内の「かわいさ」の評価システムについて述べる．

4.1 深層学習による「かわいさ」の評価

本システムにおける「かわいさ」の評価（以降スコアと呼ぶ）は，対話文を 3.2 節で構築した深層学習に入力して分類した時の，SoftMax 関数により出力される値 E（0 から 1 の実数値）を 100 点満点に換算した値に，変換ルールが適用された回数 n の 5 倍の数値を加算した式 1 の値を用いて行われる．ただし，100 点を超えた時には 100 点として扱う．

$$S = 100E + 5n \quad (0 \leqq S \leqq 100) \tag{1}$$

深層学習においては，この値が 50 点を超えると「かわいい」文として判定されるが，本研究では文の分類が目的ではなく，文にかわいさを付与することを目的とするため，目標スコアを与え，このスコアを超えるまで前章で用意した変換ルールの適用を試みる．

4.2 「かわいさ」を付与する変換ルールの適用

表 1 と表 2 に示した 11 個の変換ルールのうち，まだチェックを行っていないルールについて，変換対象となる単語があるかをチェックする．すなわち，文中の各単語について，変換対象の単語かどうかをチェックし，変換対象の単語が存在する場合，その変換ルールの適用処理を行う．変換後に目標となるかわいさスコアを超えていた場合には変換ルールの適用を終了し，目標スコアを超えていない場合は，次のルールを適用する．なお，チェックの順番は，名詞をひらがなに置換，感動詞の挿入，曖昧な表現を追加以外の 8 個のルールについて乱数を用いてランダムに決定する．その 8 つのルールについての

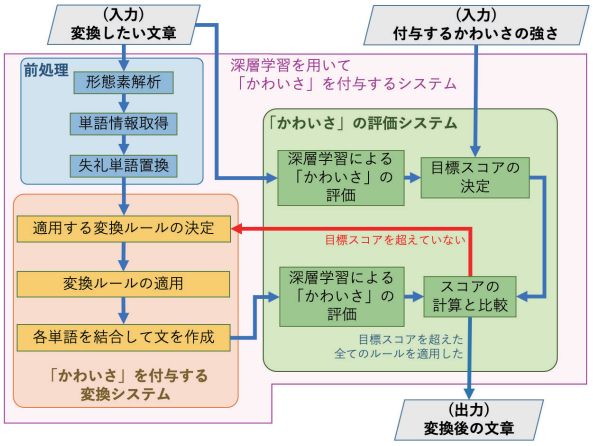


図 1: 「かわいさ」を付与するシステムのフローチャート

表 3: 変換ルールの適用で「かわいい」と答えた人数の増減

	文数 (/944)
「かわいい」と答えた人が増えた文	474
「かわいい」と答えた人が減った文	111

チェックが終了した後，曖昧な表現を追加→名詞をひらがなに置換→感動詞の挿入の順にルールを適用する．

5. 「かわいさ」を付与するシステムの検証実験

5.1 実験設定

大学生・大学院生 16 名に 3. 章で用意したデータの一部（「かわいさ」を含まない文 1000 件を学習データとは別に用意して，全く変換が行われない文を除いた 944 件）を用いて，無作為に分けた変換ルールの適用前と適用後のいずれかの文について「各文がそれぞれ自分に対する発話であった場合，その文を発した人に対してかわいいと感じるかどうか」を回答してもらった．

5.2 実験結果と考察

5.2.1 変換ルールの適用前後で「かわいい」と答えた人数の比較

変換ルールの適用により「かわいい」と答えた人数の増減表 3 に示す．約 50%の文で「かわいい」と回答する人の数が増えており，一定の効果があったことが確認できる．

変換後の文のうち，特にかわいいと判定した人数が増加した文の例を表 4 に示す．これらの文は，変換前の文はかわいいと答えた人数は 0 人だが，変換後は 6 人もしくは 7 人まで増加していた．一方で，変換後にかわいいと判定した人数が減少した文の例を表 5 に示す．これらの文は，変換前の文はかわいいと答えた人数は 3 人または 4 人であるが，変換後は 0 人であった．

変換によってかわいくならなかった 470 件の文について，その原因を著者が分類した結果を表 6 に示す．

もとの文がネガティブな内容の話，説教や教訓を含む話，若い人があまりしなさそうな内容を含む話という 3 つの要因で，かわいくならなかった文全体の約 42%を占めている．そのため，文の内容によっては，どのような変換を行ってもかわいさを付与できない可能性が考えられる．

表 4: 変換後の文のうち，特にかわいいと判定した人数が増加した文の例

変換前の文	変換後の文
結構，周りがワイワイガヤガヤしてても録れるんです．	ふふふ，結構，周りがワイワイガヤガヤしてても録れるね．
来年までいる気なわけだから．だから，ぎりぎりでしょう？	うふふ，来年までいる気なわけだからね．だから，ぎりぎりよね？
何とかっていうバンドのボーカルだったんだよ．	何とかっていうバンドのボーカルさんだったんだよー．

表 5: 変換後の文のうち，特にかわいいと判定した人数が減少した文の例

変換前の文	変換後の文
私は自分のを取ったら，手をひざの上に乗せて待ってます．	わたしは自分のを取ったら，手をひざの上に乗せて待ってよ．
それもほんとに昔行っただけだもん，高校ぐらいのときに．	それもつやつやしたむかし行っただけだもん，高校ぐらいのときにね．
いやでも，ヒジキとか，ちゃんと作ってあげようと思った．	いやでも，ヒジキとか，ぽかぽか作ってあげようと思った．

5.2.2 変換ルールごとのかわいいと判定した人数の変化

各変換ルールの適用が，読み手に与える「かわいい」という印象にどのように寄与しているかを調べるため，3.2 節で述べた各変換ルールごとに，かわいいと判定した人数の変化をまとめた．各変換ルールごとの「かわいい」と答えた人数の変化を表 7 に示す．全ての変換ルールについて，変換ルール適用後の文のほうがかわいいと判定した人数は増加していた．

特にかわいいと答えた人が増加した文が多いルールとして，感動詞の追加や置換をするルールと，命令をお願いする表現に置換するルールがある．これらのルールは，適用した文の約 6 割の文でかわいいと答えた人数が増加しており，比較的有用なルールであるといえる．かわいいと答えた人数が大きく増加した理由として，これらのルールが文のイメージを大きく変える効果を持っているという事が考えられる．感動詞を追加したり置換するルールの場合，本システムで付与する感動詞はにこやかな笑いや微笑みを感じさせるものであり，文全体に柔らかな印象を与えることができる．また，命令をお願いする表現に置換するルールの場合，命令口調というわかりやすくきつい印象を与える表現を無くし，お願いするという柔らかい表現を付与することで，きつい印象の文を一気に柔らかい印象の文に変えることができる．これらのように，文の印象に明確な変化を与えることができるため，柔らかさや気の弱さといった面から「かわいさ」につながったと考えられる．

一方で，否定表現を緩和するというルールと，断定表現を婉曲表現に変換するルールについては，ルールを適用後にかわいいと答えた人数が増加した文は約 34%〜37%であり，約 45%の文ではかわいいと答えた人数が変わらなかった．これは，強い否定表現を弱い否定表現に置換しても結局否定している事には変わらず，文全体が読み手に与える印象としてはあまり変化がなかったからだと考えられる．同様に，断定する表現を遠回しな婉曲表現に置き換えても，結局内容を肯定している事には変わらないため，文の印象に変化が無かったといえる．しかし，これらのルールでもかわいいと答えた人が増加している文は存在しているので，印象が大きく変わらないからといって不要なルールであるとは一概には言い切れない．

5.2.3 事後アンケートから見るかわいい文の傾向

事後アンケートとして「どういった文がかわいいと感じたか」「どういった文がかわいくないと感じたか」について回答してもらった．特に用意した変換ルールとは異なる要素について，かわいいと感じた文についての回答を以下に挙げる．

- ゆるい内容の世間話的な会話
- 食べ物の話が含まれる文
- 全体的に柔らかく，ポジティブな印象を与える文
- 日常のふとした感じの会話

表 6: 変換によってかわいくならなかった原因の分類

かわいくならなかった原因	該当する文数 (/470)
ネガティブな内容の文	105
前後の会話が無いと理解し辛く，違和感を感じる文	90
難しい話題や説明的な内容を含む文	76
昔話やお酒の話など，若い人があまりしなさそうな話題の文	25
お金の話題についての文	14
変換前後であまり印象が変わらない文	14
教訓や説教と言った，相手に言い聞かせるような内容を含む文	7
その他 (上記のどれにも当てはまらないもの，原因不明のもの)	139

表 7: 変換ルールごとの「かわいい」と答えた人数の増減 (%)

ジャンル	ルール	増加	減少
幼さ，未成熟さ	名詞をひらがなに置換	50.88	12.87
	副詞をオノマトペに置換	48.28	12.07
	相槌・同意・終助詞長音化	48.94	8.51
	語尾表現置換	53.93	9.63
	感動詞の挿入・置換	64.40	8.20
	名詞にさんを付ける	43.40	11.32
精神的な気の弱さ	断定表現を婉曲表現に置換	37.84	20.27
	曖昧な表現を追加	55.56	16.67
	命令をお願いする表現に置換	61.29	9.68
	逆接の接続詞を削除	47.44	12.82
	強い否定を緩和する表現の追加	34.78	19.57

- 「うふふ」などの笑い声に準ずる言葉が含まれた文

また，かわいくないと感じた文についての回答を以下に挙げる．

- 内容が硬く，難しく感じる会話
- 会話の前後がよくわからず，内容が意味不明な文
- 否定的な文や，同調・共感とは逆の文
- きっぱりした断定表現や，誰かを突き放すような棘のある表現を含む文
- 「うふふ」や「ふふふ」が使われている文

これらを比較すると，5.2.1 項での考察と同様，文の内容から受ける印象がかわいさの判定に大きく寄与しているという事がわかる．また，被験者にかわいくないと判定した文について理由を聞くと,「表現はかわいいと思ったが，文の内容的にかわいいとは感じなかった」という解答が一定数得られた．

5.2.4 考察のまとめ

文におけるかわいさには,「表現から感じるかわいさ」と「内容から感じるかわいさ」の 2 つの要素があり，それら 2 つの要素が同時に存在する場合に，読み手はその文から「かわいい」という印象を感じていたと考えられる．本システムでは,「表現から感じるかわいさ」を文に対して付与する事はできるが,「内容から感じるかわいさ」については十分に付与する事ができない．そのため，元々の文の内容が難しい場合やネガティブな内容に触れている場合は，十分なかわいさを付与する事ができないといえる．

「内容から感じるかわいさ」を付与するためには，難しい内容の文やネガティブな内容の文を，ゆるい内容の文やポジティブな内容の文に置き換える必要があり，既存のルールによる変換では対応できない．また，文の内容を書き換えると対話の内容も変わってしまうという問題点も存在する．そのため，本システムの手法とはまた別のアプローチを試みる必要がある．

6. おわりに

対話文に対して「かわいさ」を付与するシステムについて述べ，提案システムがかわいさをどの程度付与できるかを実験により検証した．実験の結果，約 50%の文に対して一定のかわいさを付与できる事を確認した．かわいさを付与できなかった文については，ネガティブな内容の話や，説教や教訓を含む話，若い人があまりしなさそうな内容を含む話など，表現を変換しても内容からかわいいとは感じない文が多かった．本システムでは「かわいい表現」を付与する事はできるが,「かわいい内容」を付与する事はできないため，難しい文やネガティブな文に対しては十分な「かわいさ」を付与できなかったと考えられる．

今後は，かわいさを付与できる文と付与できない文を区別するとともに，より長い一連の対話文に対してかわいさを付与する量とタイミングを検討していきたいと考えている．

参考文献

[1] 狩野芳伸 :「コンピューターに話が通じるか-対話システムの現在」, 情報管理学会誌，Vol.59, No.10, pp.658–665 (2017)

[2] 刀山将大, 佐藤理史, 松崎拓也, 宮崎千明, 平野徹, 松尾義博:「文のどこにキャラクタ性を埋め込む自由度があるか」, 言語処理学会第 22 回年次大会発表論文集, pp.721–724 (2016)

[3] 宮崎千明, 平野徹, 東中竜一郎, 牧野俊朗, 松尾義博, 佐藤理史:「文節機能部の確率的書き換えによる言語表現のキャラクタ性変換」，人工知能学会誌，Vol31,No.1, pp.1–9 (2016)

[4] 谷川晃大, 藤田寛泰, 壹岐太一 :「変分オートエンコーダと注意機構を用いた発話文のキャラクタ性変換」, 第 32 回人工知能学会全国大会，4G2-01 (2018)

[5] 石川なつ美 :「『かわいい』の意味について」，東京女子大学言語文化研究 24，pp.21–35 (2015)

[6] 家内嘉宏, 入戸野宏 :「共感性と親和動機による“かわいい”感情の予測モデル構築」，パーソナリティ研究，Vol.23, No.3, pp.131–141 (2015)

[7] 入戸野宏 :「“ かわいい ”感情の心理学モデル」，情報処理学会誌，Vol.57, No.2, pp.128–131 (2016)

[8] kuromoji : (URL)https://www.atilika.com/ja/kuromoji/

[9] 名大会話コーパス : (URL)https://mmsrv.ninjal.ac.jp/nucc/

[10] 藤村逸子, 大曽美恵子, 大島ディヴィッド義和 :「会話コーパスの構築によるコミュニケーション研究」(2011)

[11] 藤村逸子, 滝沢直宏編 :「言語研究の技法 : データの収集と分析」, ひつじ書房, pp.43–72 (2011)

[12] iM@S-CGWordsSearch(β) : (URL)https://icws.indigo-bell.com/

Organized Session | Organized Session | [OS] OS-8

[4G3-OS-8b] マイニングと知識創発(2)

砂山 渡（滋賀県立大学）、加藤 恒昭（東京大学）、西原 陽子（立命館大学）、森 辰則（横浜国立大学）、高間 康史（首都大学東京）

Fri. Jun 7, 2019 2:00 PM - 2:40 PM Room G (302A Medium meeting room)

[4G3-OS-8b-01] AI Player Creation Reflecting Personal and Characters in PvP Games

〇Toshinobu Sakata[1], Wataru Sunayama[1], Yuji Hatanaka[1], Kazunori Ogohara[1] （1. The university of shiga prefecture）

2:00 PM - 2:20 PM

[4G3-OS-8b-02] Predictability of Number of Visitors from GPS Location Big Data

Hirotoshi Yanagi[1,3], 〇Yoshitoshi Nakahara[2], Takahiro Hoshino[2,3] （1. Keio University, Graduate school of Economics, 2. Keio University, Faculty of Economics, 3. Riken AIP）

2:20 PM - 2:40 PM

©The Japanese Society for Artificial Intelligence

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

対戦型ゲームにおける個人の性格や能力を反映した AI プレイヤーの作成

AI Player Creation Reflecting Personal Abilities and Characters in PvP Games

坂田 駿允 *1 Toshinobu Sakata　砂山 渡 *2 Wataru Sunayama　畑中 裕司 *2 Yuji Hatanaka　小郷原 一智 *2 Kazunori Ogohara

*1滋賀県立大学大学院工学研究科
Graduate School of Engineering, The University of Shiga Prefecture

*2滋賀県立大学工学部
School of Engineering, The University of Shiga Prefecture

Nowadays, PvP online games for playing with many players are very popular. Among them, we can feel more pleasure to playing with friends and players who became friends by online games. It is considered to be able to perceive player ' s abilities and characters.

In this study, we aim to construct environment playing with AI players who have any personality within a settable range. We can be set freely AI player ' s abilities and characters parameter. In experiment, we verified how to reflect personality to In-game action. And we investigate the difference of easiness in perceive AI player ' s personality.

1. はじめに

近年，プロ棋士に勝った将棋 AI[1] や，世界レーティング 1 位の囲碁プレイヤーに勝った囲碁 AI[2] など，人間より強いゲーム AI に注目が集まっている．しかし，これらの AI は強い人間に勝つために生成された AI であり，そのゲームの初心者や中級者は強いゲーム AI と対戦しても楽しむことは難しい．また，ゲーム AI は，人間と違い機械的な行動を取ることが多く見られ，プレイヤーが違和感を感じてしまう．

私たちが友人やオンラインでなどで人間と対戦型ゲームを行う際，コンピュータと対戦するよりも楽しいと感じることができる．何故ならば，コンピュータと対戦する時は，勝ち負けでのみ，そのゲームに対する楽しさを感じ取ることしかできないが，人間と対戦する場合，勝ち負けのみでなく，プレー中の相手の行動や心境，癖など，人間特有の行動を起こし，勝ち負け以外の楽しめる要素を感じ取れることができるからである．

そこで，本研究は，対戦型ゲームにおいて，ゲーム AI に性格や能力などの，人間の持つ個性を適用し，その個性を表したゲームプレイを行うことにより，まるで人間と対戦しているように楽しめる AI プレイヤー作成することを目的とする．

2. 関連研究

2.1 ゲームにおける AI プレイヤーに関する研究

身体的な制御と生き延びるために必要な欲求を「人間らしさ」として，人間らしい振る舞いを伴うゲーム AI の自動取得をする研究がある [3]．この研究では，人間の身体的な「疲れ」や「遅れ」，特定行動に対する「訓練」や「挑戦」によって AI プレイヤーを生成しているが，本研究では，人間の「性格」や「能力」といった「個性」を適用して AI プレイヤーを生成することを目的とする．

人間の操作履歴を元に学習し，模倣することにより，人間らしさを表現する AI プレイヤーを作成する研究がある [4]．この研究では，操作ミスをしない，短時間で操作が変わるといった，AI プレイヤーが行う非人間的な行動を改善し，AI の行動に反映させている．本研究では，行動を決定する際に，ゲームルールから考えられる，非人間的な行動のみを排除し，それ以外の行動は人間の性格や能力から適用する．

2.2 対戦型ゲームにおける AI プレイヤーに関する研究

プレイヤーの行動や対戦記録などから AI プレイヤーを生成する研究がある [5][6][7]．この研究では，AI プレイヤーを生成する際に，対戦するプレイヤーの行動を用いて，プレイヤーと同程度の強さを持った AI プレイヤーを生成している．本研究では，第三者の性格や能力を取得して，その性格や能力を表す行動を行う AI プレイヤーの生成を目的とする．

プレイヤーの行動やゲームの状況によって対戦相手に感情的な反応を返す研究がある [8]．この研究では，ゲームの状況に応じて快-不快パラメータを設定し，そのパラメータに対応する反応を返している．本研究では，ゲームの状況だけでなく，AI プレイヤーの性格や個性も判定要素として取り込み，反応を返す AI プレイヤーを生成する．

人間の個性的なプレースタイルを模倣する AI の研究がある [9]．この研究では，AI に適用する行動として人間のプレースタイルを模倣しているが，本研究では，人間の性格や能力から考えられる行動を AI プレイヤーに適用することを目的とする．

3. 対戦型ゲーム「大富豪」における AI プレイヤーの性格と能力

本章では，トランプの対戦型ゲーム「大富豪」における，考えられる性格や能力と，それらに基づいた行動について述べる．

3.1 対戦型ゲーム「大富豪」のルール設定

本研究では，性格や能力を適用する AI プレイヤーを評価する対戦型ゲームとして，「大富豪」を題材とする．ただし，以下の 5 つをルールとする簡易版の大富豪を取り扱う．

- Joker を除いた 4 つのスートに各 A〜K までの 13 枚のカード，計 52 枚のカードを使用
- 3 が最弱で，数が大きくなるにつれて強くなり，A は K より強く，2 が最強
- カードは必ず一枚ずつ出す
- 8 切りや革命などの特殊ルールは無効

連絡先: 坂田駿允，滋賀県立大学大学院工学研究科電子システム工学専攻，〒 522-8533 滋賀県彦根市八坂町 2500

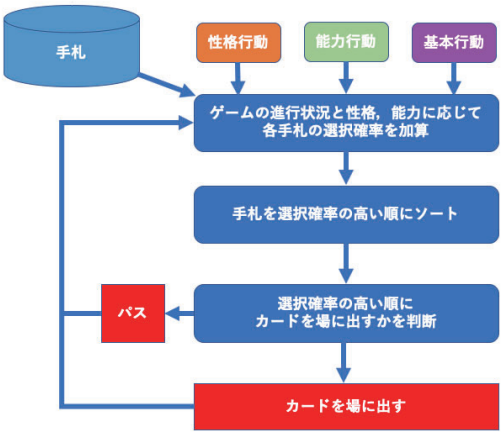


図 1: 大富豪における AI プレイヤーの行動の枠組み

表 1: 大富豪における性格や能力と対応する行動

性格や能力	行動
記憶力	場に出たカードを記憶する
リスク志向	強いカードを出しやすい 相手が上がりそうな時，阻止する
気まぐれ	適当にカードを出しやすい
負けず嫌い	一定以上の順位を目指す
焦りやすい	ピンチに強いカードを出しやすい

- ゲーム継続時のカード交換をしない

図 1 に、大富豪における AI プレイヤーの行動の枠組みを示す。AI プレイヤーは、設定された性格と能力に応じて、場に出すカードを決定する。

3.2 AI プレイヤーの性格と能力による行動

本研究では、表 1 に示す 1 つの能力「記憶力」と、4 つの性格「リスク志向」,「気まぐれ」,「負けず嫌い」,「焦りやすい」の実装を試みる。各性格と能力を表すため、それぞれの強さを表すパラメータ P_{memory}、P_{risk}、P_{moody}、$P_{competitive}$、P_{panic} を 0 から 1 の範囲の実数値で設定する。

大富豪において，AI プレイヤーが行える行動は，各カードを場に出す行動とパスに限られる．また AI プレイヤーが持つカードには、それぞれ選択確率が設定され、選択確率が高いものから順に、カードに与えられている確率でそのカードを出すか否かが決定される。なお全てのカードを出さない選択がなされた場合は、行動としてパスを選択する。ただし状況によっては、確率によらず出すカードが決定されることがある。

AI プレイヤーの基本行動、ならびに各性格や能力のパラメータに基づいて、カードに選択確率 C が式 (1) で与えられる。ただし、M_{basic}、M_{1risk}、M_{2risk}、M_{moody} は後述の、基本行動や性格、能力によって与えられる値となる。

$$C = M_{basic} + M_{1risk} + M_{2risk} + M_{moody} \tag{1}$$

また、正規化前のカードの選択確率の最大値を C_{max} として、カードの最大選択確率の値が 0.9 になるように式 (2) により正規化を行う。

$$\dot{C} = \frac{0.9}{C_{max}} C \tag{2}$$

また、場に出せないカードの選択確率は 0 に，必ず出すカードの選択確率は 1 にする。各カードに割り当てられる選択確率 M はで表される．

3.2.1 AI プレイヤーの基本行動

- 場に出せるカードの小さい順に選択確率を大きく設定する。
- 場を流せるカード以外のカードが一枚以下の場合，流せるカードから出す（他人に上がられるリスクを避ける）．
- 場に何も出ていない場合，最小カードから出す．
- カードが残り一枚かつそのカードが場に出せる時，パスをしない（パスが選択されないようにする）。

基本行動により各カードには、m 番目に小さいカードに、式 (3) によるパラメータ M_{basic} を与える。これは，大富豪の基本的な戦略として，場に出せる小さいカードからカードを減らすことを表している。

$$M_{basic} = \frac{0.3}{m} \tag{3}$$

3.2.2 性格と能力による行動

本実験で実装する性格や能力の大富豪における行動を表 1 に示す．

「記憶力」の行動は，場に出たカードを記憶し，場を流せるカードを作れる能力で記憶力を表し，カードの忘却頻度でその能力の強弱が表現できると考えた．

「リスク志向」は，積極的に勝負に挑む行動として，強いカードをよく出しに行き，相手が上がりそうな状況では，その状況を阻止するようなリスクの高い行動の頻度で表現できると考えた．

「気まぐれ」は，セオリーでない行動を表現する方法として，出せるカードのうち，場に出せる最小のカード以外を出すような行動，例えば，3 が場に出ている時に Q のように強いカードを出すことにより，気まぐれを表現できると考えた．

「負けず嫌い」は，勝負事に負けたくないという思考の強さを，目標順位として設定し，その順位を目指すような行動を行うことで表現できるのではないかと考えた．

「焦りやすさ」は，焦った時の判断能力の低下を表現するために，大富豪において，自分が負けそうになった時に，有利不利関係なく相手の行動を阻止する，つまり強いカードを出すことにより表現できるのではないかと考えた．

3.3 AI プレイヤーの性格と能力のパラメータ設定と行動

3.3.1 記憶力

記憶力によって行われる行動は，自分のカードと場に出たカードをすべて覚え，自身のターンが来た時に，記憶力パラメータ P_{memory} を用いて，式 (4) で表す確率 F_{memory} で忘れるようにする．また，記憶力の行動の目的は，場を流せるカードを増やすことなので，カードを覚え，忘れるカードの対象は，10 以上のカードとする．

$$F_{memory} = (1 - P_{memory}) \times 100(\%) \tag{4}$$

表 2: 負けず嫌いパラメータによる目標順位の設定

負けず嫌いパラメータ $P_{competitive}$	目標順位
1〜0.9	1 位
0.89〜0.75	2 位
0.74〜0.2	3 位
0.19〜0	4 位

3.3.2 リスク志向

自分のカードの n 番目に強いカードから順に、選択確率 M_{1risk} を、リスク志向パラメータ P_{risk} 用いて式 (5) より算出する。

また、4 人での対戦時において、中盤以降すなわち最後に場にカードを出したプレイヤーの手持ちカードが 7 枚以下の時に，自分の持っている一番強いカードに，式 (6) より算出した，選択確率 M_{2risk} を適用する．リスク志向の行動により，強いカードに大きく選択確率を適用し，その後は徐々に選択確率を減らしながら適用することで，リスク志向のパラメータが大きいほど，強いカードを出しやすくしている．また対戦相手のカードが減り始めた時に，一番強いカードのみに選択確率を割り当て，相手の行動を阻止する勝負に出る．

$$M_{1risk} = \frac{P_{risk}}{2n} \tag{5}$$

$$M_{2risk} = \frac{P_{risk}}{2} \tag{6}$$

3.3.3 気まぐれ

気まぐれは，そのパラメータが大きいほど，適当にカードを出しやすくなるように，気まぐれパラメータ P_{moody} に定数を掛ける式 (7) により，各カードに選択確率 M_{moody} を適用する．すなわち，気まぐれパラメータが大きいほど，予測不能なカードを出しやすくなる．また定数 R は、毎回 0 から 1 の範囲の実数値が乱数で生成されて与えられる。

$$M_{moody} = P_{moody} \times R \tag{7}$$

3.3.4 負けず嫌い

負けず嫌いパラメータ $P_{competitive}$ により、表 2 の目標順位を設定する。目標順位が達成できなかった時には、気まぐれパラメータ P_{moody} を 1 にする．

3.3.5 焦りやすい

大富豪において焦りやすい状況とは，相手が上がりそうな時に，自分が上がれる見込みがない時に発現しやすい．さらに，その状況を打開するため，無理やり自分のカードを減らそうとする行動を取る．すなわち、ゲーム中盤のプレイヤーが 8 回カードを出すかパスを行った後に，自分のカードの枚数順位が下位だった場合に，焦りやすさを表すパラメータ P_{panic} を確率として、P_{panic} の確率で再度リスク志向の行動を行う（M_{1risk} と M_{2risk} の加算を再度行う）ように設定する．

3.4 対戦型ゲーム「大富豪」のインタフェース

本節では，大富豪のインタフェースに実装した機能について述べる．図 2 に，大富豪のインタフェースを示す．今回，インタフェースを実装する環境として，TETDM*1 の可視化ツールを用いた．

*1 TETDM サイト：http://tetdm.jp

図 2: 大富豪のインタフェース

表 3: 設定した AI プレイヤーの性格や能力パラメータ

	記憶力	リスク志向	気まぐれ	負けず嫌い	焦りやすい
くま	0.5	0.8	0.2	0.7	1.0
うさぎ	0.1	0.1	1.0	0.1	0.1
いぬ	0.3	1.0	0.5	0.8	0.3
ねこ	1.0	0.1	0.1	0.5	0.1

4. 対戦型ゲームにおける AI プレイヤーの性格推定実験

4.1 実験設定

理系の大学生大学院生 11 名に、性格や能力を適用した AI プレイヤー同士の対戦を観戦してもらった後で，各 AI プレイヤーの性格や能力を推定してもらい、AI プレイヤーの行動から AI プレイヤーの個性を感じ取れるかを確認する実験を行った．用意した 4 人の AI プレイヤー（くま，うさぎ，いぬ，ねこ）のパラメータを表 3 に示す。

観戦回数は 6 回から 9 回の間で、およそプレイヤーの能力と性格を推定できたと思うまで観戦してもらった。各性格と能力について、「記憶力が高い」「リスク志向」「気まぐれ」「負けず嫌い」「焦りやすい」のそれぞれに当てはまる度合いを 5 段階で回答してもらった。なお、9 回観戦してもらってもわからなかったときには、わからなかったと回答してもらった。

4.2 実験結果

表 4 から表 7 に各 AI プレイヤーの能力と性格の推定結果を示す。ただし各表の黄色いセルは、AI プレイヤーに設定したパラメータを 5 段階評価に合わせた対応推定値（0 から 0.2 を 1、0.3 と 0.4 を 2、0.5 と 0.6 を 3、0.7 と 0.8 を 4、0.9 と 1.0 を 5）を表す。

表 4 から表 7 の記憶力の項目を見ると，どの記憶力パラメータの値でも対応推定値に近い評価ができていた．これは，場の記憶と忘却による記憶力の表現や，パラメータによる表現の強弱ができていたためと考えられる。

表 4 から表 7 のリスク志向の項目を見ると，“くま”と “ねこ”の 2 つの AI プレイヤーで対応推定値に近い評価ができていた．リスク志向パラメータは高く設定された “いぬ”と “くま”について、“くま”は焦りやすさが高く，“いぬ”では焦りやすさが低く設定されていた．焦りやすさは「パラメータが高いとリスク志向行動を取りやすくなる」ため，“くま”でリスク志向の評価が高かったのは，焦りやすさによってリスク志向行動

表 4: AI プレイヤー “くま”の評価結果

評価	記憶力	リスク志向	気まぐれ	負けず嫌い	焦りやすい
1	2	0	2	0	1
2	3	1	4	1	3
3	3	4	1	1	4
4	1	1	1	5	3
5	1	5	2	4	0
0	1	0	1	0	0

表 5: AI プレイヤー “うさぎ”の評価結果

評価	記憶力	リスク志向	気まぐれ	負けず嫌い	焦りやすい
1	1	0	0	0	0
2	6	0	4	3	1
3	1	4	3	2	1
4	3	5	2	5	3
5	0	2	1	1	6
0	0	0	1	0	0

表 6: AI プレイヤー “いぬ”の評価結果

評価	記憶力	リスク志向	気まぐれ	負けず嫌い	焦りやすい
1	6	0	0	0	0
2	2	3	3	5	1
3	2	4	4	3	4
4	1	3	3	2	6
5	0	1	0	1	0
0	0	0	1	0	0

表 7: AI プレイヤー “ねこ”の評価結果

評価	記憶力	リスク志向	気まぐれ	負けず嫌い	焦りやすい
1	0	4	5	5	1
2	0	6	3	2	4
3	2	1	1	2	4
4	0	0	1	2	2
5	9	0	0	0	0
0	0	0	1	0	0

が強く現れたためと考えられる．そのため、現状のリスク志向の行動は目に見えて反映されにくい可能性が考えられる．

表 4 から表 7 の気まぐれの項目を見ると，全 AI プレイヤーにおいて，パラメータの値に関係なく，低く評価される傾向にあった．このことから，性格を表す行動として，幅広い値を取りうるランダム要素を用いるとうまく性格が推定されないことがうかがえる。

表 4 から表 7 の負けず嫌いの項目を見ると，“くま”のみ対応推定値に近い評価が得られた．しかし，“くま”と負けず嫌いパラメータが同程度の “いぬ”では，低い評価が多く，他 2 つの AI プレイヤーでも低い評価が多かった．これは，負けず嫌いパラメータの行動である「目標順位が達成されなかった場合，気まぐれパラメータを最大にする」という行動において、気まぐれが推定されにくかったことと合わせ、負けず嫌い独自の行動が設定されていなかったことによると考えられる。そのため，各性格ごとに独自の変化が明らかになるような行動の設定が必要と考えられる．

表 4 から表 7 の焦りやすさの項目を見ると，全 AI プレイヤーにおいて，対応推定値とは遠い評価が得られた．これは，焦りやすさを表す行動が,「順位が低い場合に，リスク志向行動を取る」のみであり，焦りやすさ特有の行動を設定していなかったためと考えられる．

5. おわりに

対戦型ゲームにおいて，人間と対戦しているように楽しめる AI プレイヤーとして，設定した性格や能力を適用し，行動を行う AI プレイヤーを作成した．

作成した AI プレイヤーが設定した性格や能力を反映し，行動によって表現できているかを検証するために，評価実験を行った．記憶力のように能力として反映されやすいものは、行動から推定しやすいことがわかった。性格として用意されるものは、その性格に基づく行動が強く現れるようにしないと、行動からの推定が難しいことがわかった。

今後は、今回題材として扱った性格について、より性格を推定しやすいパラメータや行動の設定を検討していきたい。また、同様の性格を他のゲームに応用するための一般的な方法論について研究を進めていきたいと考えている。

参考文献

[1] 瀧澤武信：現役プロ棋士に勝ち越したコンピュータ将棋～第 2 回電王戦，第 23 回世界コンピュータ将棋選手権速報～：4, 第 23 回世界コンピュータ将棋選手権速報, 情報処理, Vol.54, No.9, pp.925–928 (2013)

[2] 松原仁：AlphaGo の置き土産, 情報処理, Vol.58, No.9, pp.668–669 (2017)

[3] 藤井叙人・佐藤祐一・中嶌洋輔・若間弘典・風井浩志・片寄晴弘：生物学的制約の導入による「人間らしい」振る舞いを伴う ゲーム AI の自律的獲得, ゲームプログラミングワークショップ 2013 論文集, pp.73–80 (2013)

[4] 中野雄基・美添一樹・脇田建：進化計算と UCT による Mario を人間らしくプレイする AI, ゲームプログラミングワークショップ 2013 論文集, pp.81–88 (2013)

[5] 上田陽平・池田心：遺伝的アルゴリズムによる人間のレベルに適応する多様なオセロ AI の生成, 研究報告ゲーム情報学 (GI),Vol.2012,No.5,pp.1-8,(2012)

[6] 服部裕介 ・田中彰人 ・星野唯一 ：対戦型アクションゲームにおけるプレイヤの模倣行動の生成, 情報処理学会研究報告ゲーム情報学 (GI),Vol.2007,No.20,pp.1-8,(2007)

[7] 星野唯一・田中彰人・濱名克季：模倣学習により成長する格闘ゲームキャラクタ, 情報処理学会論文誌, Vol.49, No.7, pp.2529 − 2548 (2008)

[8] 塩入健太,・星野准一：仮想対戦プレイヤーの感情的発話生成, 情報処理学会研究報告エンタテインメントコンピューティング (EC),Vol.2005, No.125, pp.63–70 (2005)

[9] 生井智司・伊藤毅志：将棋における棋風を感じさせる AI の試作, 情報処理学会研究報告ゲーム情報学 (GI),Vol.2010.3, No.3, pp.1–7 (2010)

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

位置情報データを用いた公共施設来場者数の予測可能性

-野球興業の来場者数予測モデルの構築を中心に-

柳 博俊[*1 *3] Hirotoshi Yanagi　中原 嘉利[*2] Yoshitoshi Nakahara　星野 崇宏[*2 *3] Takahiro Hoshino

[*1] 慶應義塾大学大学院経済学研究科 Graduate School of Economics, Keio University
[*2] 慶應義塾大学経済学部 Faculty of Economics, Keio University
[*3] 理化学研究所 Riken AIP

It is the purpose of this paper to discuss whether we could predict the number of visitors of public facilities from the GPS location big data collected from approximately 20 million smartphones in total while taking into consideration influences of various covariates to correct several biases relate to the penetration ratio. We used number of visitors of baseball games in the year 2018 as a label, and GPS location big data and other features as feature vectors to apply supervised machine learning methods. Because of the prediction accuracy, we concluded that it is possible to predict number of visitors of public facilities from GPS location big data.

1. はじめに

本稿ではスマートフォンから取得されたアプリ位置情報を用いた, 公共施設来場者の予測可能性の検討ついて報告する.

近年,スマートフォンの普及に伴い, 膨大な GPS ログが収集されている. これらの情報により, 海水浴場など来場者の数の正確な計測が困難な場所や, テーマパークや公共施設のような計測値が発表されていても月単位または半期/通期単位での公表しかなかった場所について, 即時的にその周辺人口を推定できる可能性がある. これらの情報を用いることによって, 施設利用者がよりリアルタイムな混雑状況についての情報を取得することが可能になったり, 来場者数そのものがビジネス上の重要な KPI になるような施設であれば, 公表前にその来場者数の情報を取得することによって決算公表前にその施設の経営状況を推察することが可能となる.

このように GPS データを用いた適用方法が多く考えられる一方で, GPS データから周辺人口の推定を行う方法についての研究/実適用例は多くなく,NTT ドコモの「モバイル空間統計」や agoop 社の提供する月別メッシュ別人口など集計に限られたものとなる.

本研究では, 公共施設の来場者数予測の可否について, 来場者のデータが公表されているプロ野球興行の来場数データを正解値に, X 社提供を受けた位置情報データ/試合に関するデータを特徴量として用いて, Kaggle など予測系タスクに広く使われる LightGBM により予測が可能かを検証する.

2. 関連研究

2.1 従来型人口統計について

従来の人口統計は、全数調査である国勢調査によって判明する昼間人口や夜間人口などの人口静態統計が主であり、公表の最小単位は地域メッシュである(総務省統計局(2015))。具体的には、「約 1km 四方の基準地域メッシュ、約 500m 四方の 2 分の 1 地域メッシュ及び約 250m 四方の 4 分の 1 地域メッシュ単位に再集計(総務省統計局(2017))」され、公表されている。故に、従来の統計情報では自身の関心のあるエリア、あるいは施設に滞在している流動人口をリアルタイムで把握することはほぼ不可能である。

2.2 位置情報を用いた, 動的な人口統計について

2.1 の問題点を解決し流動人口を推定すべく、交通系 IC カードを用いた方法や(牧村ら(2010))携帯電話の基地局圏在数を用いた方法(寺田ら(2012))が考案されている。特に後者の方法は、NTT ドコモが提供する「モバイル空間統計」として商用化されており、推定される動的人口の代表性に関しては複数の研究が行われている(大藪ら(2017)、矢部ら(2016))。このデータベースは。震災後の避難者数の推定や帰宅困難者数の推定への利用が期待されている(村上ら(2011))。

また本研究と関連するところとして, 携帯電話の GPS(Global Positioning System)を利用した位置情報データから流動人口を推定する技術の開発が進みつつある。実例を挙げると、RESAS では、株式会社 Agoop の提供する月別・メッシュ別の動的人口を取得することができる。

3. データおよび変数について

3.1 予測に用いるデータについて

今回の研究では、X 社にご提供いただいた約延べ 2000 万人分、5 分毎の位置情報データを用い、2018 年のプロ野球のホームゲーム全 786 試合の観客動員数を予測した。

プロ野球の対戦カード、使用スタジアム、雨天中止の有無、観客動員数、各球場の最大収容人数、予告先発投手、開始時刻、試合時間などの基本的な情報は、プロ野球 Freak(2018)が各球団の発表した値を収集したデータを使用した。なお、このデータは、日本プロ野球協会の公式 HP で公開されている集計値と照らし合わせて正確であることを確認してある。

また、各球団は、ホームスタジアムの試合に関し、有料販売したチケット枚数に無料配布分の実来場者数を加算した値を観客動員数として公表している (伊藤(2016))。なお、無料配布チケットは球場の窓口で受け取る仕組みになっており、無料チケット所有者が実際に来訪しているのか観測できる仕組みとなっている(伊藤(2016))。

観測端末数に関しては、試合時間中に球場で観測されたユ

連絡先: 柳博俊 , h.yanagi-95@keio.jp

ニークな ID 数を集計して求めた。なお、近隣施設を利用していたにも関わらず、GPS の測定誤差によって野球場に居たとカウントされることを防ぐため、最低でも 12 回(延べ 1 時間)球場で観測された ID のみを観測端末数に加算している。

懸念点としては，位置情報を取得している端末属性にかかるバイアスがあげられる．位置情報は特定アプリケーションをインストールしているスマートフォンから取得されており，そのアプリケーションの利用者属性によりバイアスが生じる．具体的には

1.インストール率は地域、年代、性別によって異なる。

2.インストール率は時間経過とともに変化し、一定ではなく、またその普及率の変化にも偏りがある。

ことが集計によりわかっている．

3.2 来場者数と観測端末数の関係の検討

観客動員数と観測端末数に相関関係がなくては、予測モデルを構築することは不可能である。

図 1 から分かるように、観客動員数と観測端末数には比較的強い正の相関があることがわかる。しかしながら、観客動員数が増えるにつれ、観測端末数の分散が大きくなる、不均一分散の傾向が観測されている。

分布に関して比較すると、観測端末数はほぼ対数正規分布に従っているのに対し、観客動員数は 3 万人を境として分布が大きく下がり、4 万 5 千人の周辺にもう一つのモードがあることがわかる。このことは、図 2,3 からも明らかで、観測端末数の分布と観客動員数の分布は、一致していないことが分かる。

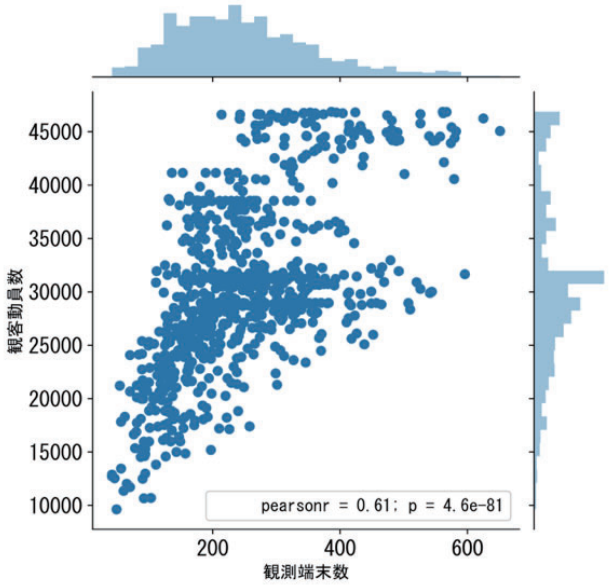


図 1 全試合の観客動員数と観測端末数

3.3 来場者数のカウントノイズの検討

観客動員数について，12 球団中複数の球団の発表値をみると測定ノイズではないかと思われる傾向が見つかった．観客動員数が満員に近い試合の場合，動員数がそれほど多くない試合とは異なり，端末の観測数に対して大きい値が発表される傾向にあるというものである．

例えば，図 2 では球団 A がホームチームの場合，観客動員数の発表値が 30000 人を超えたあたりから，付近での観測端末数に対して大きい値を示す傾向が見られる．

満員試合数というのは，球団経営において，重要なビジネス KPI なりうると考えられるため，観客の収容率が高い試合の場合はこのような補正がかかってしまっているのではないかと考えられる．

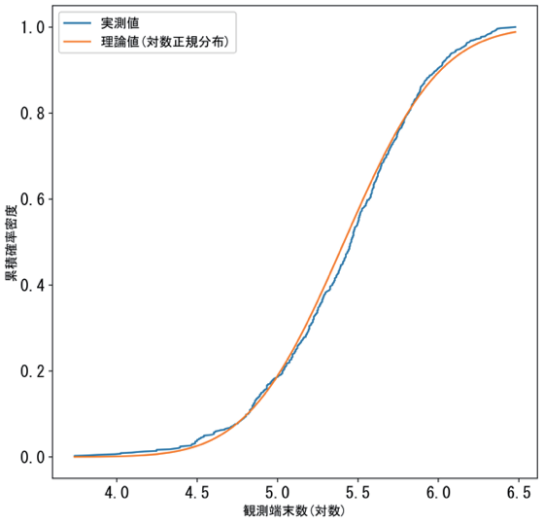


図 2 観測端末数の分布

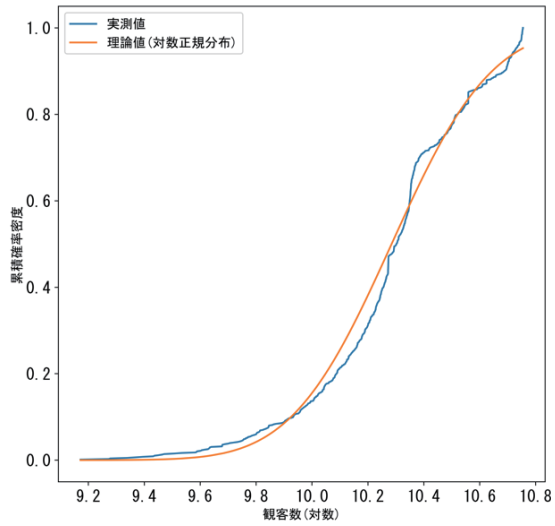


図 3 観客動員数の分布

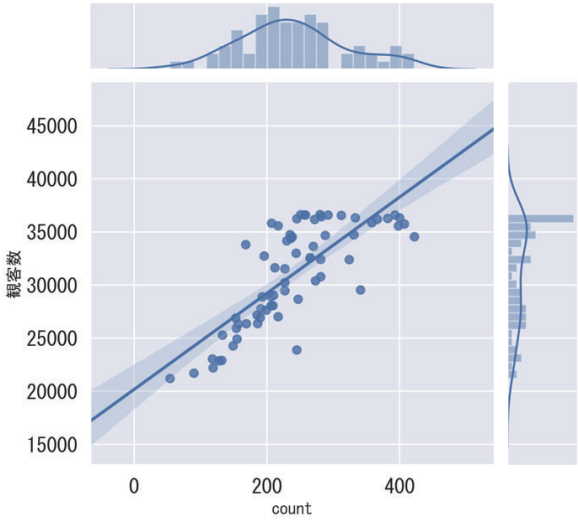


図 4 球団 A の観客動員発表値と観測端末数の関係

連絡先: 柳博俊 , h.yanagi-95@keio.jp

4. 実証分析

4.1 正解値定義

3.1 で述べたように各球団が有料チケット販売数と招待客数を合計したものを毎試合公表しており，それを正解値とする予測問題を解く.

4.2 特徴量

試合中に球場付近で観測された端末数を用いて，予測を行う.

また，試合を行う球場や，ファン層によって，実来場数のうちデータが取得される端末の数の含有率が異なると考えられる. これらを補正するために，

対戦相手 / 先発投手などを用いて，それらを補正することを試みる.

＜変数の定義＞

連続値特徴量	
観測端末数	試合が行われている最中に連続 1 時間以上球場周辺で観測された端末の数
ホームチーム	試合のホームチームごとに観客数について TargetEncoding 処理したもの
ビジターチーム	試合のビジターチームごとに観客数について TargetEncoding 処理したもの
月	試合の行われた月ごとに観客数について TargetEncoding 処理したもの
曜日	試合の曜日ごとに観客数について TargetEncoding 処理したもの
予告先発	ホームチームの予告先発投手ごとに観客数について TargetEncoding 処理したもの
OneHot 特徴量	
ホームチーム	試合のホームチームの OneHot 特徴
ビジターチーム	試合のビジターチームの OneHot 特徴
月	試合が行われた月の OneHot 特徴
曜日	試合が行われた月の OneHot 特徴
予告先発	ホームチームの予告先発投手

4.3 予測方法 / 精度評価

Microsoft 社の提供する勾配ブースティング木アルゴリズムである LightGBM を用いて，データを 10 分割した上で Leave-One-Out 方式で予測を行う. Out of Fold(OoF)の値を予測値として扱う.

精度評価には Mean Absolute Error（MAE）と Mean Absolute Percent Error 及び Mean Squared Error を使用する.

4.4 結果

結果は表の通りになった.

精度指標	MAE	MAPE	RMSE
予測精度	4266.11	0.17072	5962.79

表 予測誤差

また，特徴重要度の上位は図 5 のようになり，観測端末数が最も予測に貢献していることがわかる

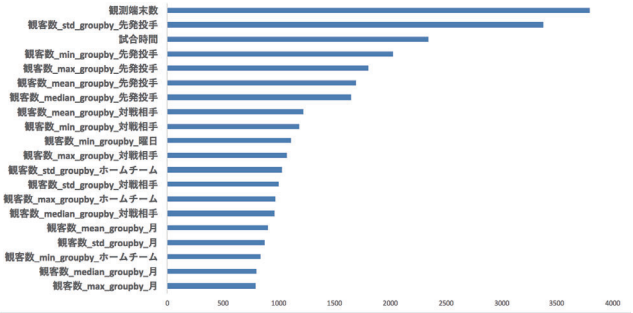


図 5 特徴重要度 Top20

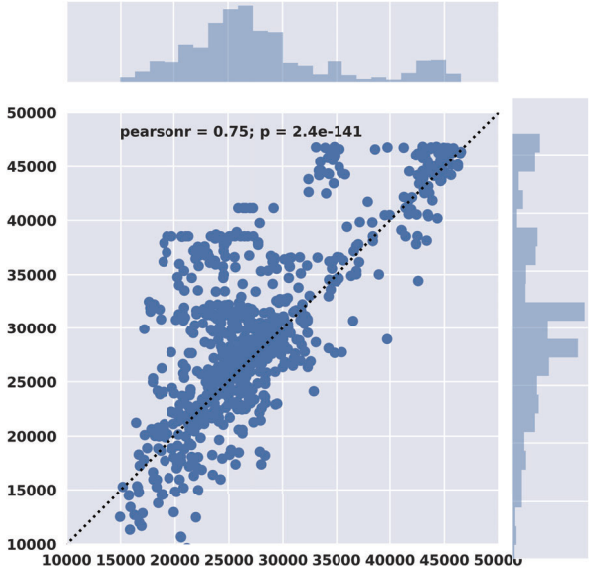


図 6 OOF に対して予測の結果

5. 今後の課題 / 展望

最後に、本研究によって浮かび上がった、今後取り組むべき研究課題に関し議論する。

5.1 過学習の懸念

第 1 の課題意識は、過学習の懸念である。今回は、位置情報データが 2018 年 4 月からしか利用できなかったため、学習データ、テストデータともに野球の 2018 年シーズンから作成している。そのため、本研究で作成した機械学習モデルが 1 シーズンのデータに対し過学習している可能性がある。この問題点に関しては、データの蓄積を待って、複数シーズンのデータに対し、本論で適用した機械学習モデルを適応して予測精度を確認する必要がある。

5.2 位置情報集計のノイズ

第 2 の課題は、商業施設が密集している地域で、野球の試合を観戦していない ID が間違って観客としてカウントされている、集計ミスの可能性である。3.2 で議論した通り、今回は、横浜市の中心部にあり商業施設やホテル、駅などと隣接し

ている横浜球場で、集計ミスの疑いがあった。この点に関しては、Gonzalez(2008)が提唱した、各位置情報端末所有者の行動特性の特徴量の抽出が必要であると考える。野球観戦特有の行動ログのパターンを特定し、そのパターンに一致している ID のみを抽出して集計することができれば、集計ミスは減らすことができると期待している。

5.3 観客動員集計のノイズ

3.3 にて,観客動員数の集計において，集計時のバイアスが疑われた. この問題を解決するため，より正確に来場のデータが取得されている公共施設について，上記懸念点を払拭した上で，今回と同様の予測を行いたい.

より正確な来場のデータが取得されたときにも，位置情報からの予測がより正確に行えることを示す必要がある.

連絡先: 柳博俊 , h.yanagi-95@keio.jp

参考文献

González, M. C., Hidalgo, C. A. and Barabási, A. L.: Understanding individual human mobility patterns, Nature, Vol. 453, pp. 779-782, 2008.

Hastie, T.; Tibshirani, R. & Friedman, J. (2009), The Elements of Statistical Learning 2nd Edition, Springer New York Inc., pp675-720,2009

伊藤歩(2016),「プロ野球は観客動員数をごまかしているのか-公表人数と肌感覚で大きな差が出るカラクリ」,『東洋経済』,2016 年 1 月 10 日,電子版

大藪勇輝, 寺田雅之, 山口高康, 岩澤俊弥, 萩原淳 一郎, 小泉大輔(2012),「モバイル空間統計の信頼性評価」,『NTT DOCOMO テクニカル・ジャーナル』, Vo.20 , No.3 , pp.17-23, 2012

RESAS 公式HP、https://resas.go.jp/#/13/13101(入手:2019 年 1 月)

総務省統計局(2017),「平成 27 年国勢調査に関する地域メッシュ統計 結果の概要 」,総務省 統計局 統計調査部,2015 年 6 月

RESAS 操作マニュアル,
https://resas.go.jp/manual/#/13/13101(入手:2019 年 1 月)

生形嘉良, 関本義秀(2013),「大規模・長時間の GPS データによる観光統計調査の活用可能性～石川県を事例に～」,『土木学会論文集 D3』.Vol. 69. No.(土木計画学研 究・論文集第 30 巻),2013

プロ野球 Freak(2018),https://baseball-freak.com/audience/(2018 年 11 月入手)

寺田雅之, 永田智大, 小林基成(2012),「モバイル空間統計にお ける人口推計技術」,『NTT DOCOMO テクニカル・ジャーナ ル』, Vo.20, No.3, pp.11-16, 2012

総務省統計局(2016),「平成 27 年国勢調査の集計体系及び結果の公表・提供等一覧」

村上正浩, 岡島一郎, 鈴木俊博, 山下仁(2011),「モバイル 空間統計を活用した滞留者・帰宅困難者数の推定と 具体的対策の検討」,『日本建築学会梗概集, F-1 分冊』. p.893-894, 2011

牧村和彦, 中村俊之, 千葉尚, 森尾淳, 布施孝志(2010),「バス IC カードを用いた人の動き~交通計画への活用 に向けた可能性と限界~」,『土木計画学研究・講演集』, Vol.41, 2010

連絡先: 柳博俊 , h.yanagi-95@keio.jp