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

○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

○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

○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

○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

○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

○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

○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

○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], ○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

○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
○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
○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
○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
○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
○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
○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
○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
○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
○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
○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)

[1F4-OS-17b-01] A Comparative Analysis of Labor Force Survey Data among Vegetable Cultivation Systems based on Agriculture Activity Ontology

○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

○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

○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

○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

○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

○Tetsuo Ono[1] （1. Hokkaido University）

3:20 PM - 3:40 PM

[1G3-OS-13a-02] (Invited talk) Cognitive science of choice and manipulation

○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?

○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

○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

○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

○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

○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

○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

○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

○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

○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

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

○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

○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

○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

○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

○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 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　北垣 亮馬 *2　鳥海 哲史, 林 幹久 *3
Toshihiko Yamasaki, Yuki Obuchi, Yuan Lin　Ryoma Kitagaki　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/

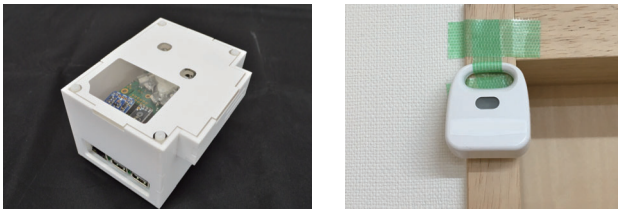


図 1: 独自開発した IoT センサと補助的に用いるオムロンの環境センサ

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。

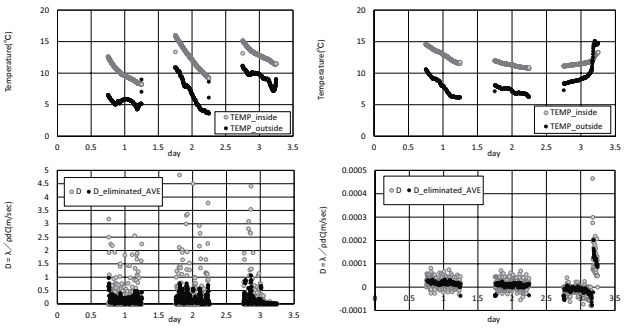
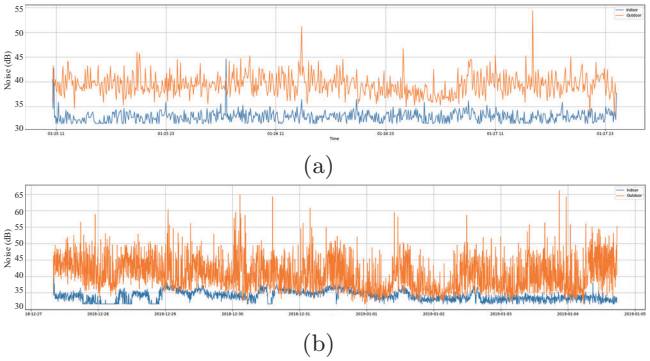


図 2: 断熱性能の簡易指標 D の計算結果。



(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，電子メイルアドレスなど

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

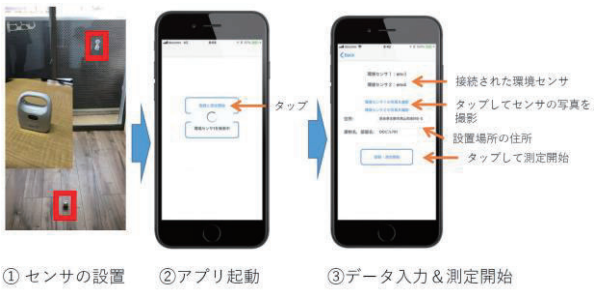


図 2 センサ設置プロセス

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 lo cation, 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/fr data/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

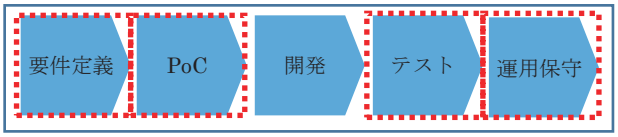
営業スタッフの作業時間が 1 物件あたり約 70%短縮され、1 ヶ月あたりに換算すると約 3,000 時間の作業時間削減に寄与することが見込まれる。

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.

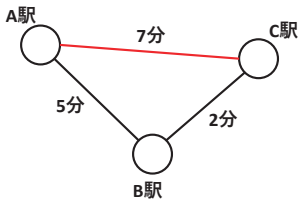


図 1: 路線ネットワーク例

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

○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

○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

○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], ○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

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

○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 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．

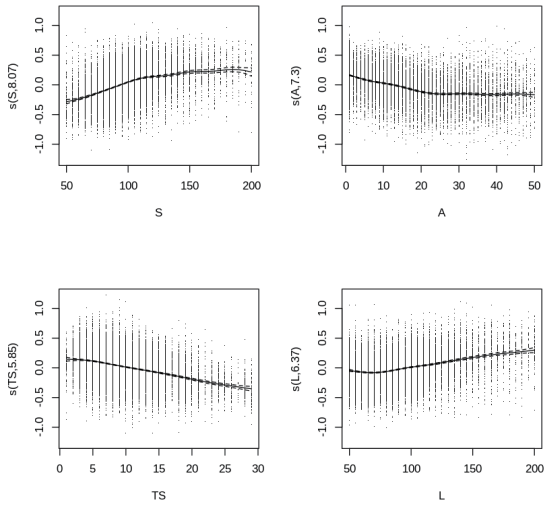


図 1: 各変数と取引価格の関係:GAM

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．

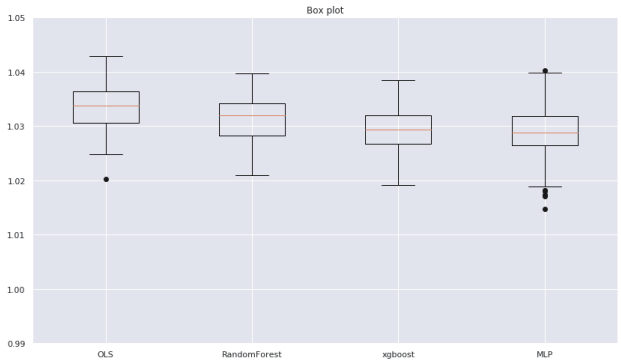


図 2: 誤差率の平均値の分布

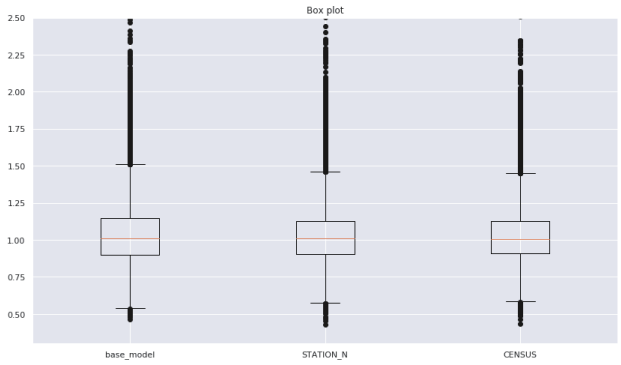


図 3: ダミー変数の効果

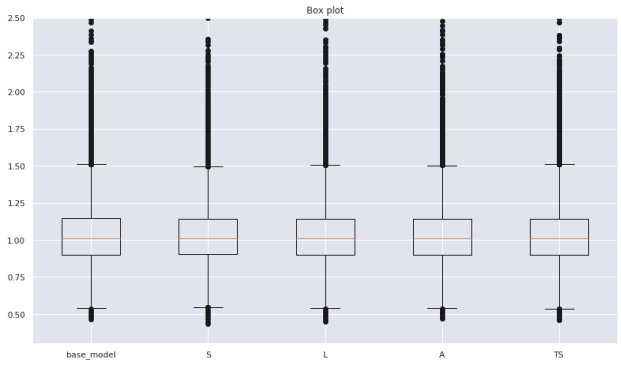


図 4: 説明変数ごとの効果

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６７３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○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．

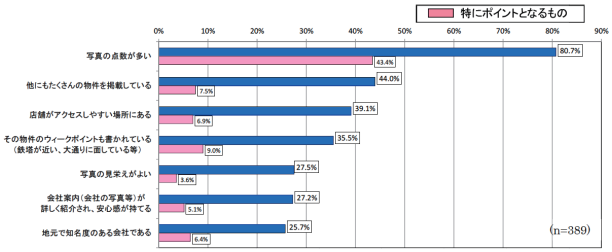
本稿では，不動産・住宅情報サイト LIFULL HOME'S における AI を活用した UX 向上の取り組み事例として，不動産物件画像への深層学習の適用による 2 つのアプリケーションと，不動産価格推定アルゴリズムを用いた参考価格算出システムを紹介する．その上で，不動産業界全体の UX を向上していくために何が必要かについて論じる．

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

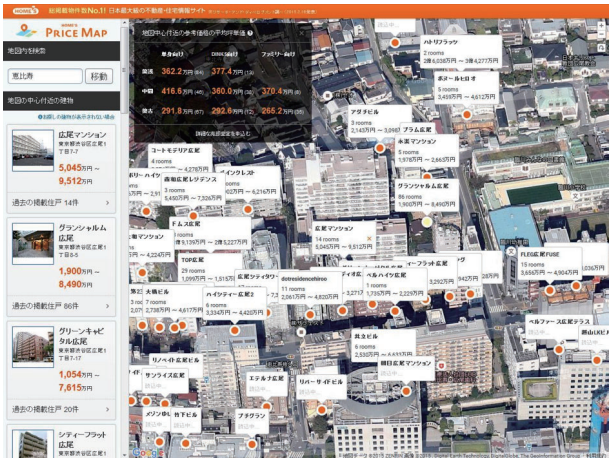


図 4: 「LIFULL HOME'S プライスマップ」の画面

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．

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

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

*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

○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

○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

○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

○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

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

○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 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 Σ $(= \Sigma_1 \cup \Sigma_2)$ 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]

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

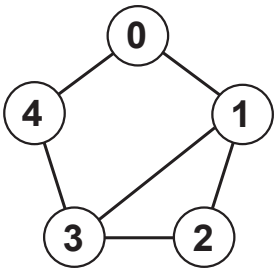
$$n_0 \neq n_1 \qquad n_0 \neq n_4 \qquad n_1 \neq n_2$$
$$n_1 \neq n_3 \qquad n_2 \neq n_3 \qquad n_3 \neq n_4$$

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.

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` (¬), `and` (∧), `or` (∨), `xor` (⊕), `iff` (⇔), `imp` (⇒), etc. can be used. As comparison operators, `le` (≤), `lt` (<), `ge` (≥), `gt` (>), `ne` (≠), `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.




```
<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

```
<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

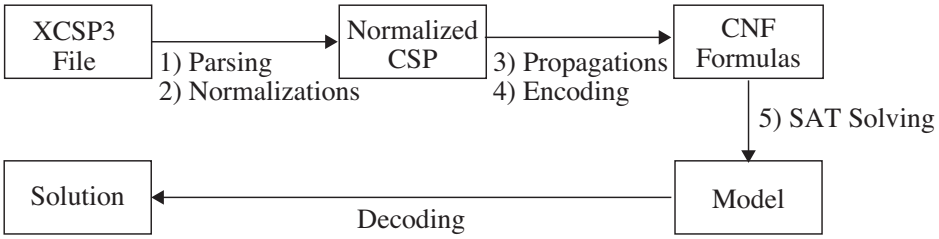


Figure 2: Framework of sCOP

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

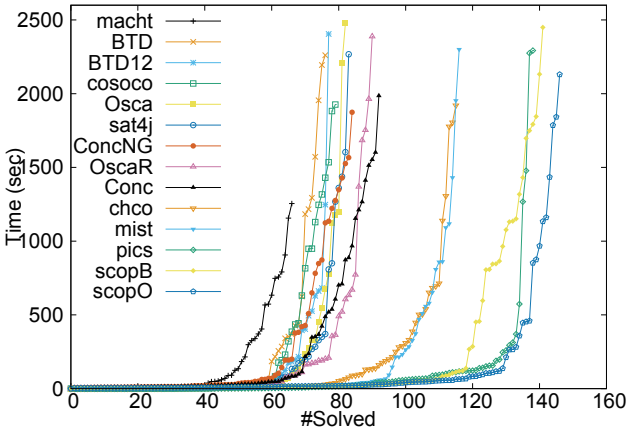


Figure 3: Cactus Plot of CSP-Standard-Sequential Track

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

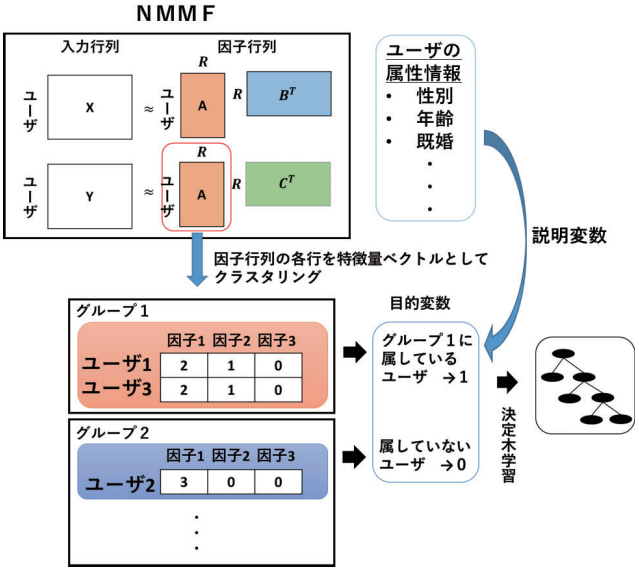


図 1: 提案手法の流れ

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

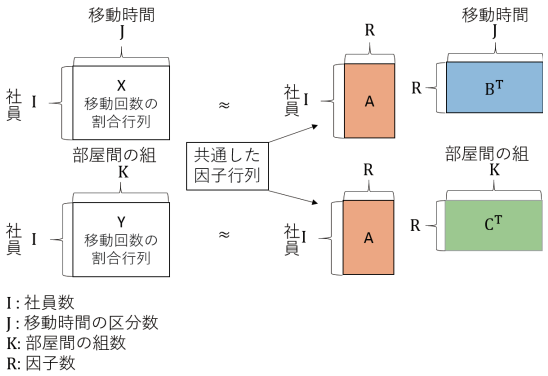


図 2: 行列分解形

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。

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

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

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

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。

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

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

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

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。

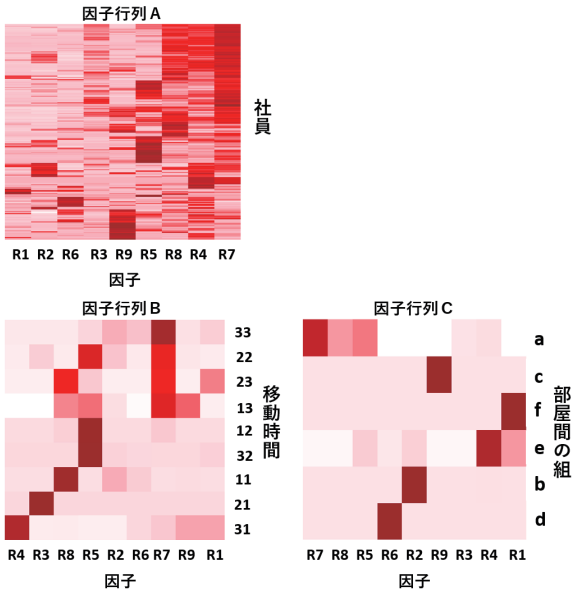


図 3: NMMF による因子分解の結果

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。

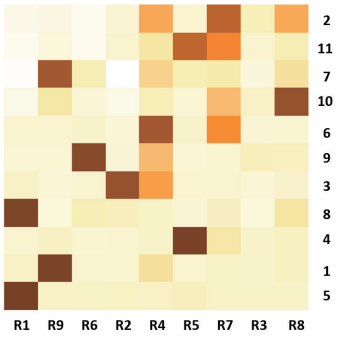


図 4: 因子行列 A をクラスタリングした結果のクラスタ中心のヒートマップ、行がクラスタ番号、列が因子

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。

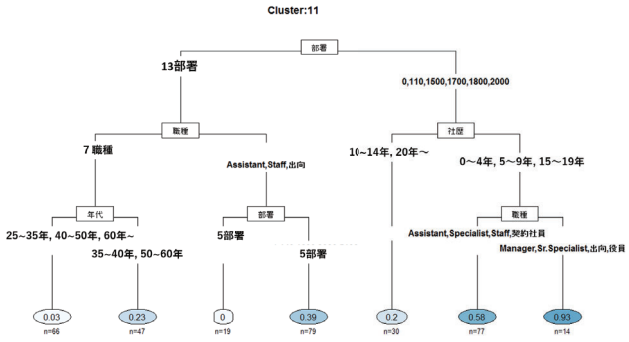


図 5: クラスタ 11 の分類木

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, 石樽隼人, 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

○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

○Takafumi Horiuchi[1], Kazunori Ueda[1] （1. Waseda University）

3:40 PM - 4:00 PM

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

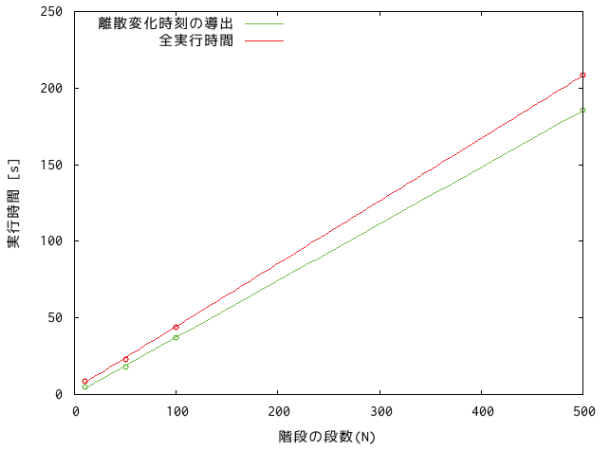


図 2: 階段を跳ねる質点のプログラムにおける段数と実行時間の関係

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(⟨0, G, P, false⟩)
5:     repeat
6:         ⟨T_curt, G_curt, P_curt, tf_curt⟩ := PQ.pop()
7:         if P_curt = false then
8:             continue
9:         end if
10:        if tf_curt = true then
11:            Sol := Sol ∪ {⟨T_curt, P_curt⟩}
12:            PQ := ⋃_{⟨T,g,p,tf⟩∈PQ} {⟨T, g, ¬P_curt ∧ p, tf⟩}
13:            continue
14:        end if
15:        H := CalculateRelaxedGuards(G_curt)
16:        for h ∈ H do
17:            MinResult := FindMinTime(Subst(h, S), P_curt)
18:            for ⟨T_min, p⟩ ∈ MinResult do
19:                CGSResult :=
20:                CheckGuardSat(Subst(S, t = T_min), G_curt, p)
21:                for ⟨tf, p⟩ ∈ CGSResult do
22:                    PQ.push(
23:                        ⟨T_min, {g | g ∈ G_curt ∧ g ⇒ h}, p, tf⟩)
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.

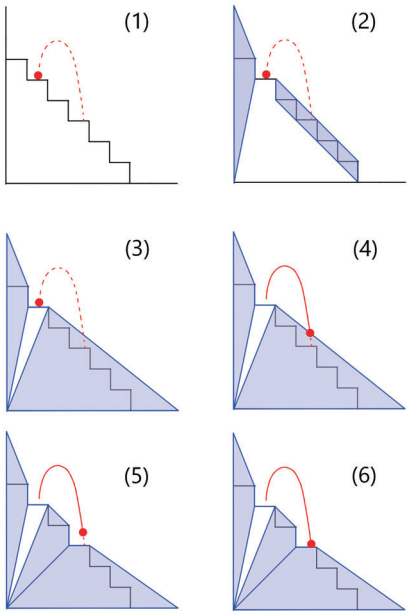


図 3: 階段を跳ねる質点のプログラムに提案アルゴリズムを適用した解探索の流れ

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

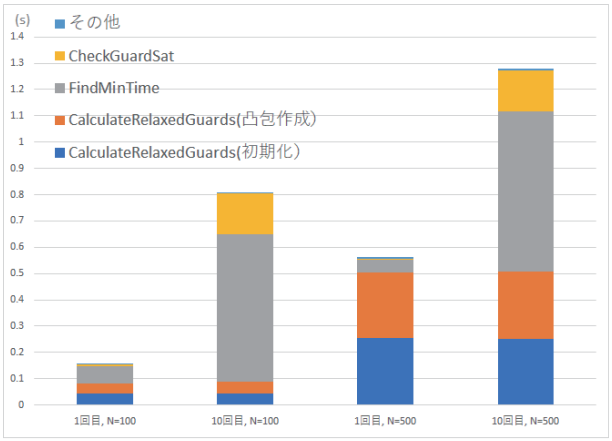


図 4: 実験結果の内訳

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.

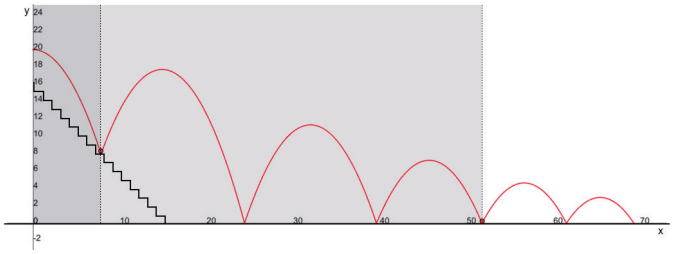


Figure 2: Concept of Omitting Verbose Guard Evaluations

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.

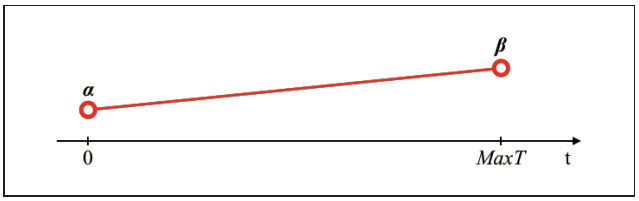


Figure 3: Concept of Approach to Uniform Monotonicity

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:

$$SimulationTime = 0.6083\,N^2 + 0.8763\,N + 1.691 \quad (4)$$

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.

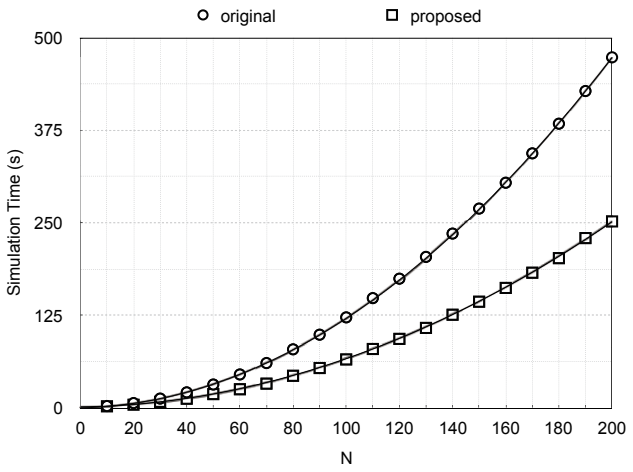


Figure 4: Simulation Time of Split Surface Model

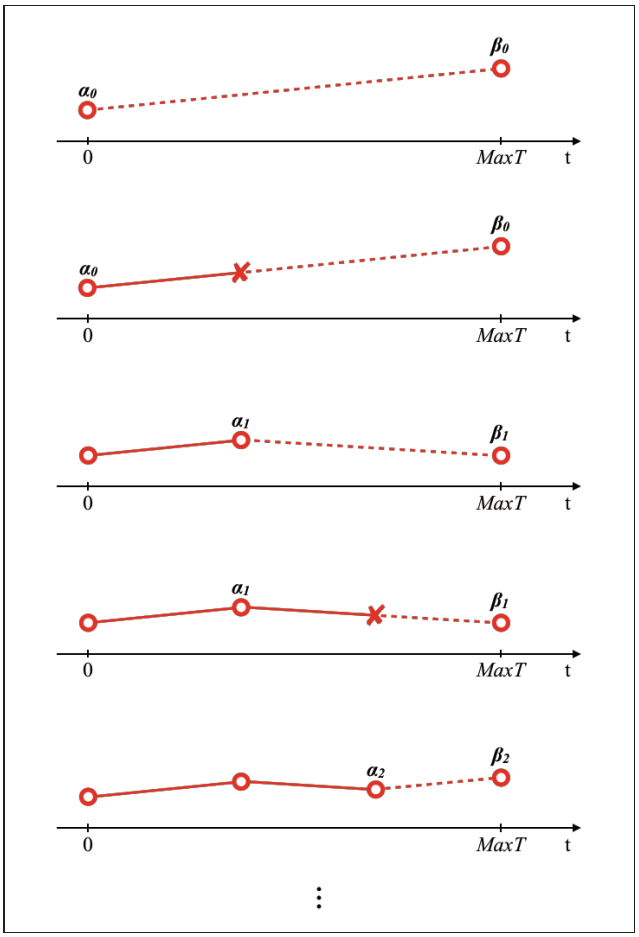


Figure 5: Concept of Approach to Uniform Monotonicity

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

○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

○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

○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

○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

○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

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

○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 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

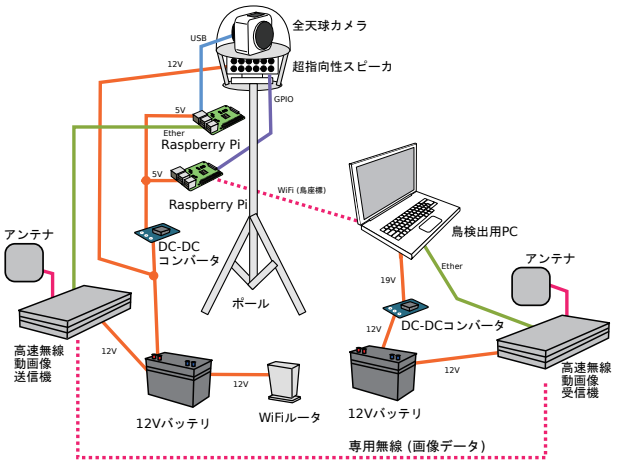


図 1: 追い払いシステムの構成

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

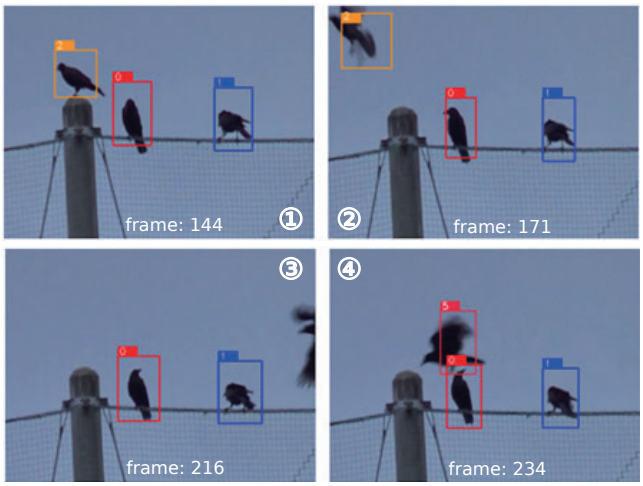


図 2: 複数個体の鳥を追跡する様子

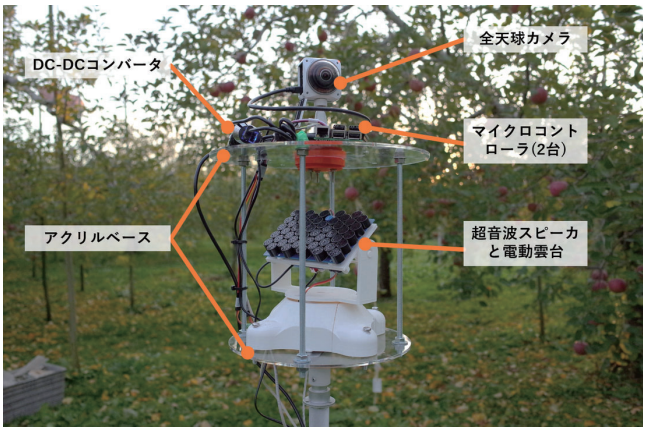


図 3: 開発した鳥追い払いシステム

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．

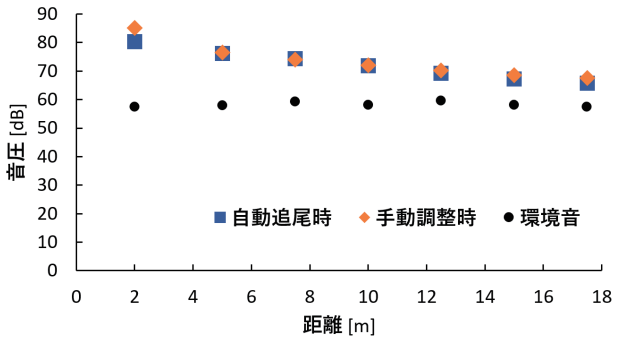


図 4: 音圧測定結果

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

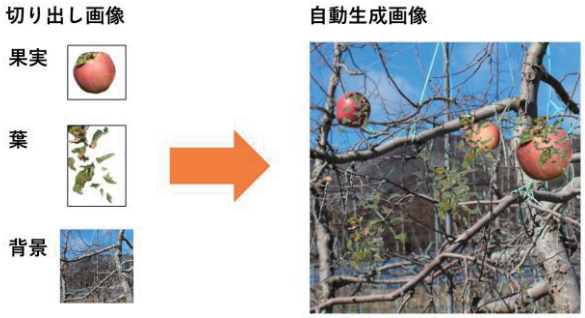


図 1: 合成画像の生成方法

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．

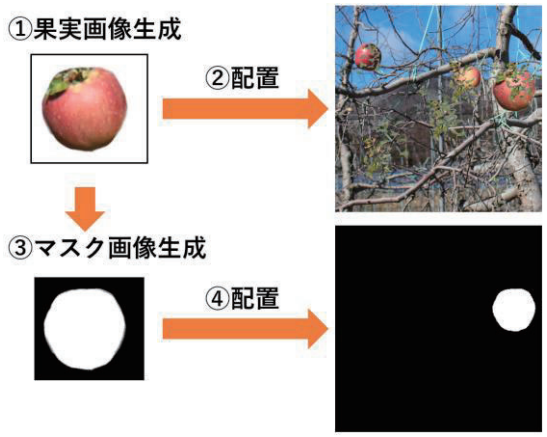


図 2: 果実ごとのバイナリマスクの生成

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．

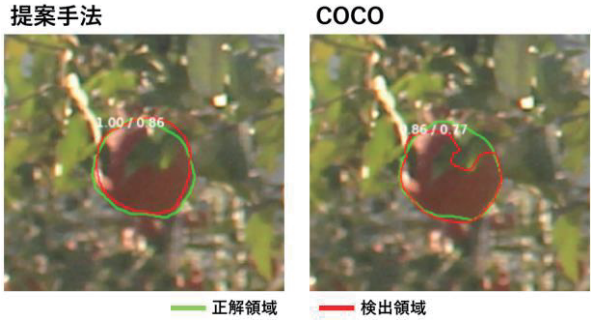


図 3: 隠れ果実領域の検出例(緑 : 真値，赤 : 推定値)

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．

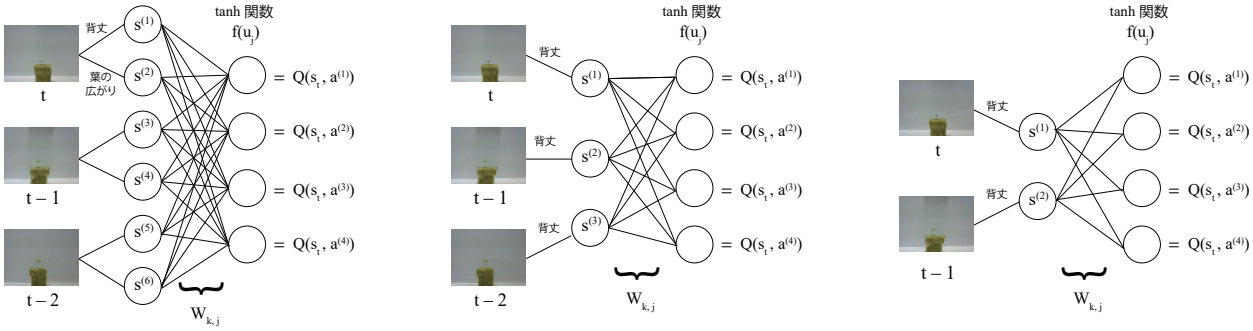
連絡先: 難波脩人，神奈川大学大学院工学研究科電気電子情報工学専攻　能登研究室，〒 221-8686 神奈川県横浜市神奈川区六角橋 3-27-1，電話：045-481-5661

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，



(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．

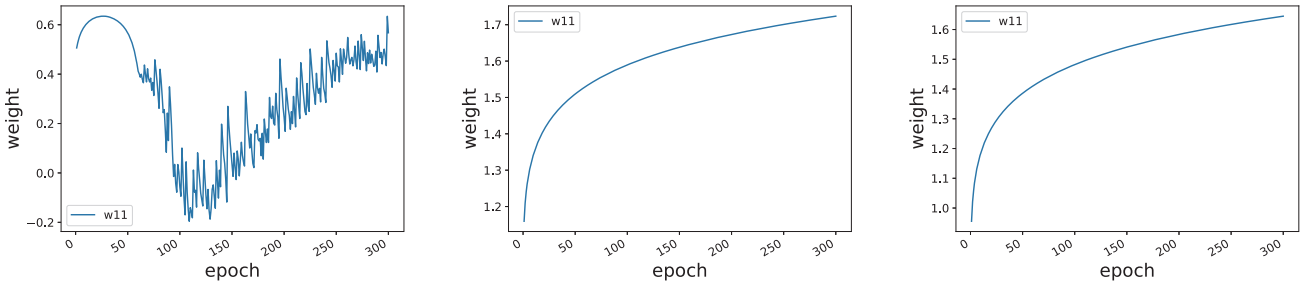
2.2 experience replay を用いた学習

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．

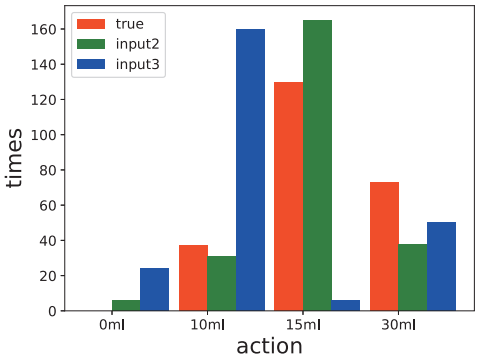


図 3: action の出現頻度

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．

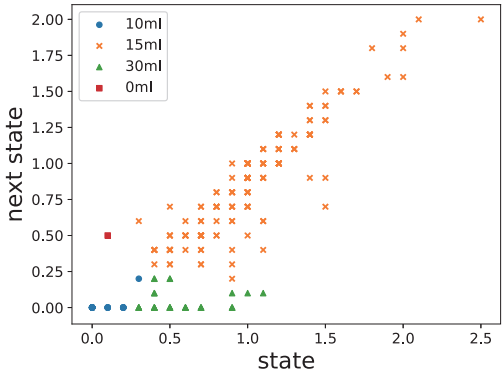


図 4: action の選択分布（ER による結果）

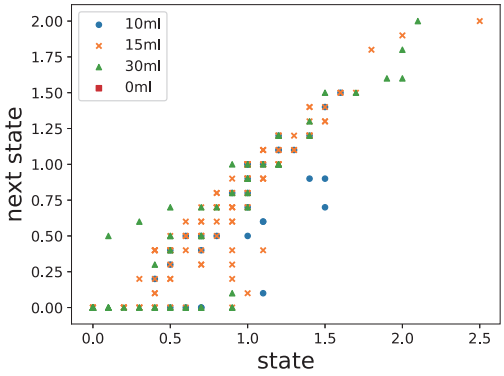


図 5: action の選択分布（人の手による結果）

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

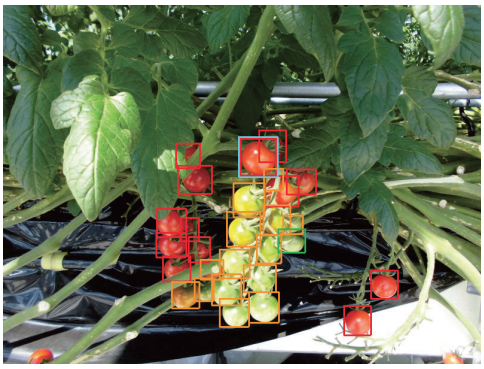


図 1. タグ付けされた撮影画像

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．

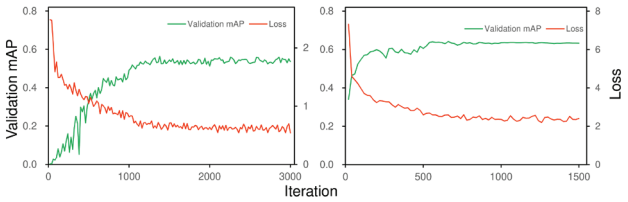


図 2.　学習過程の損失，およびに精度の変化

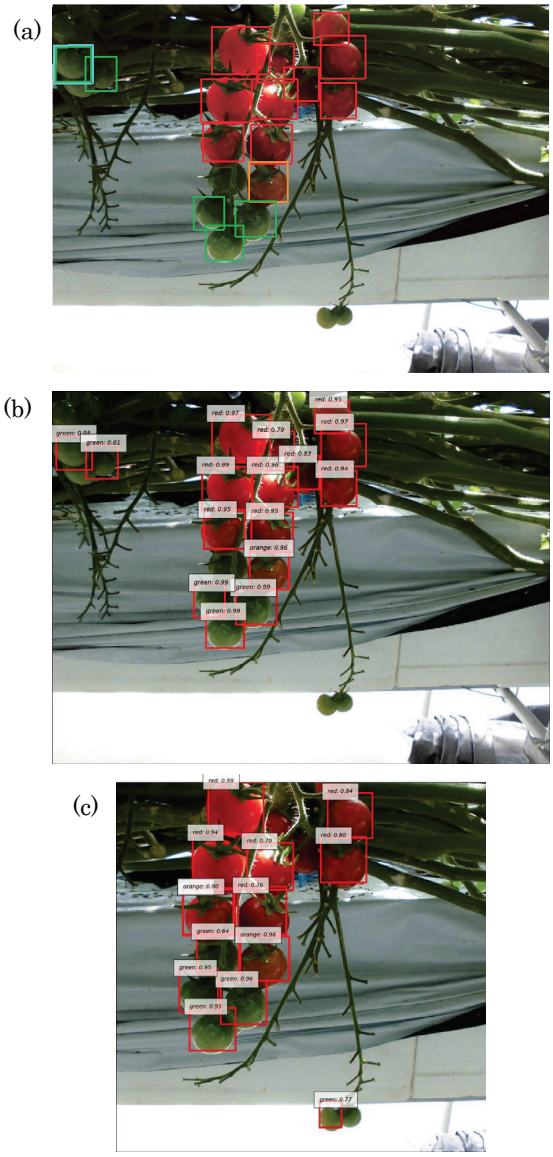


図 3.　学習済みモデルによる画像検出の結果．(a) 正解データ，(b) F-RCNN，(c)SSD

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

○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

○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

○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

○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

○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 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．

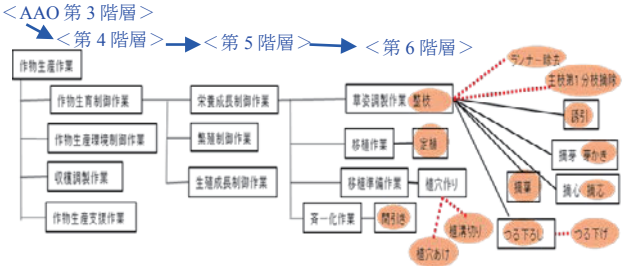


図1．農業経営指標データ名の AAO への対応付け例
□AAO 農作業名，同じ概念の用語は空白区切りで表示，●農業経営指標データ名
■AAO 農作業名と農業経営指標データ名が一致

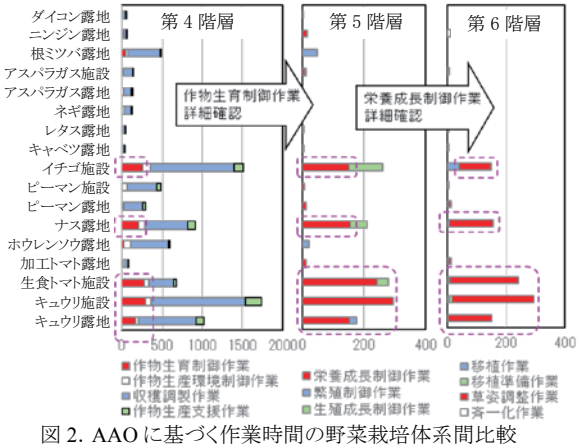


図 2．AAO に基づく作業時間の野菜栽培体系間比較
図1の対応付け結果に基づき，AAO 第 4 階層，第 5 階層（“作物生育制御作業”の下位階層），第 6 階層（“栄養成長制御作業”の下位階層）で作業時間を集計し比較．
横軸は 10a 当たりの作業時間を示す．

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．

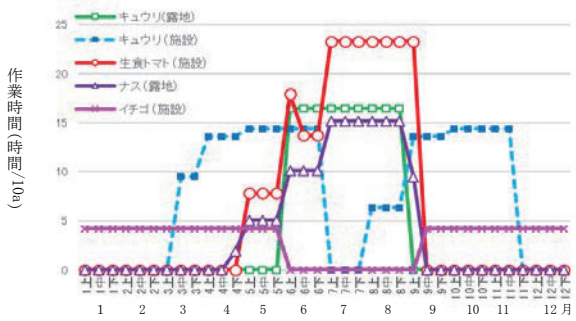


図 3．AAO に基づく旬別作業時間の野菜栽培体系間比較
図1の対応付け結果に基づき，“草姿調整作業”の作業時間を旬別に集計し比較．

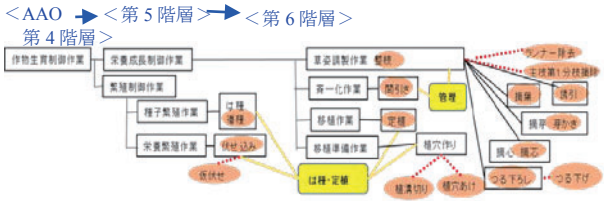


図 4．統計データ名の AAO への対応付け，および農業経営指標データとの関係例
□AAO 農作業名，同じ概念の用語は空白区切りで表示，●農業経営指標データ名
■AAO 農作業名と農業経営指標データ名が一致，■統計データ名

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

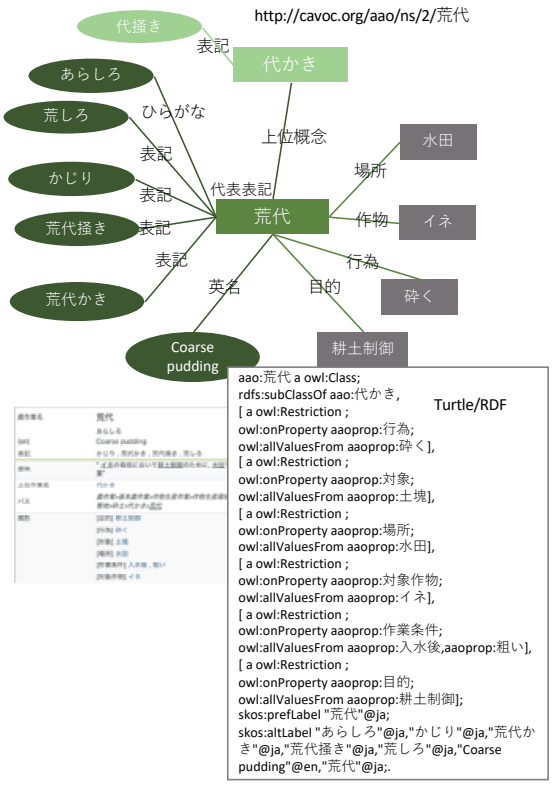


図 1: 「荒代」における情報の意味関係.

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/

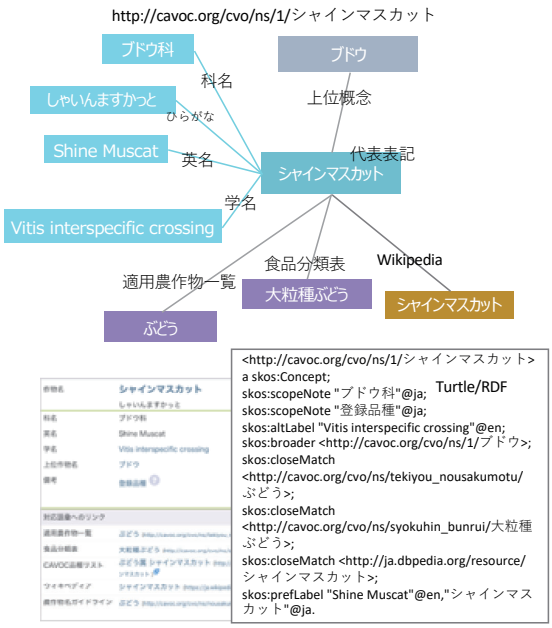


図 2: 「シャインマスカット」における情報の意味関係.

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

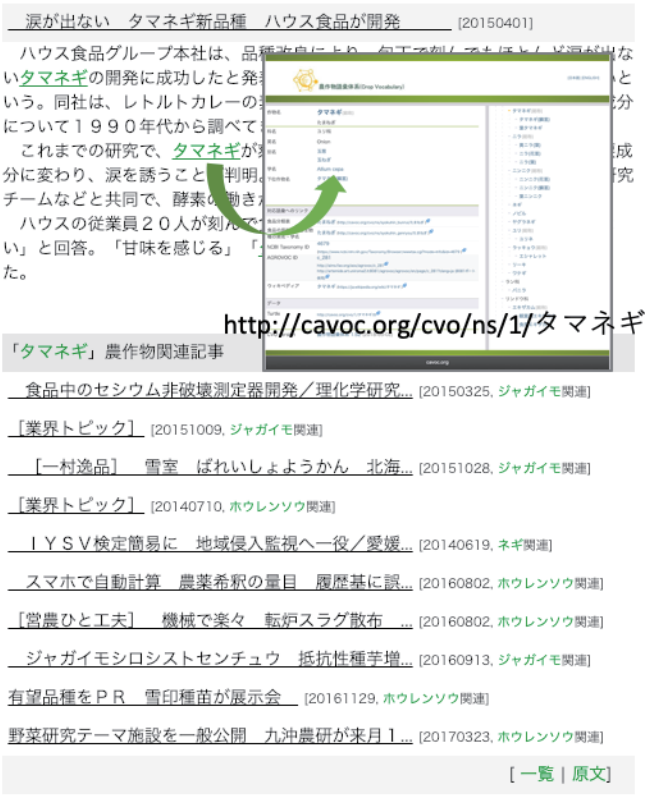


図 3: 農作物間の類似性による関連記事リスト.

記事と共に関連記事として表示されている．

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，

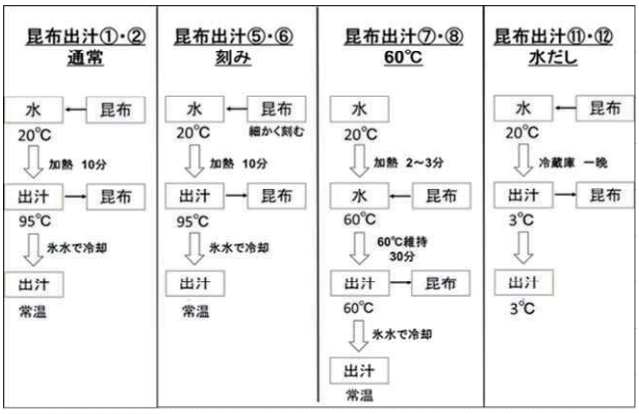


図 1 昆布出汁の調理方法

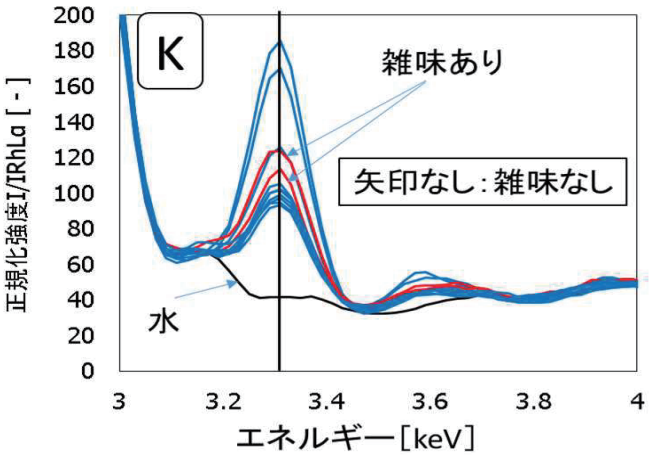


図2 各出汁の蛍光 X 線分光スペクトル（3.4keV）

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.

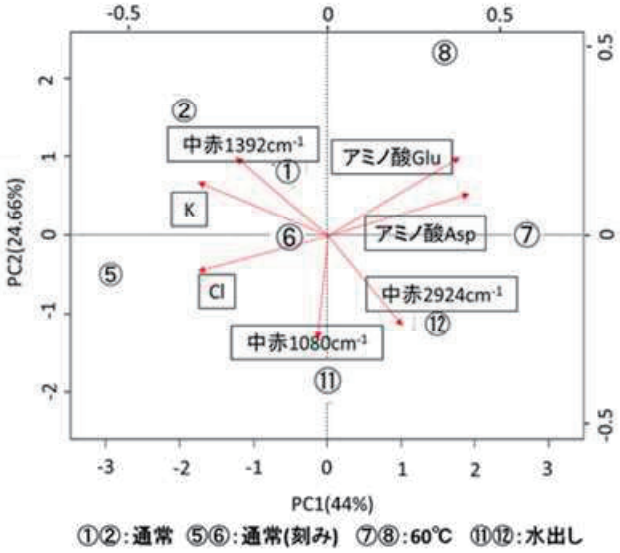


図 3 昆布出汁の主成分分析結果

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

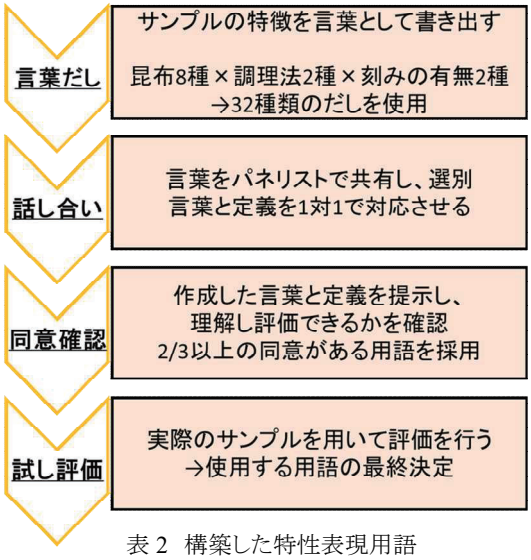


表 2 構築した特性表現用語

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

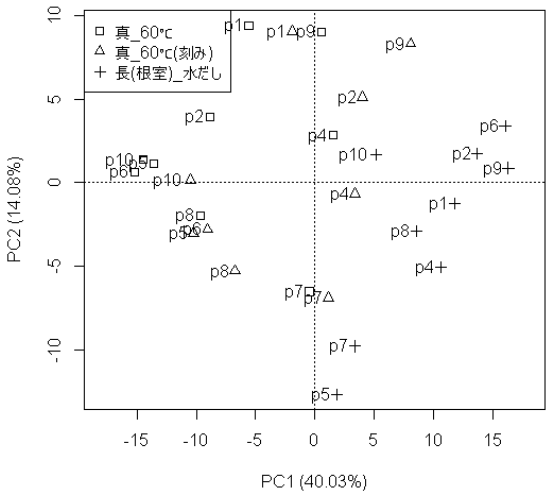


図 4 評価値の主成分分析結果

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．

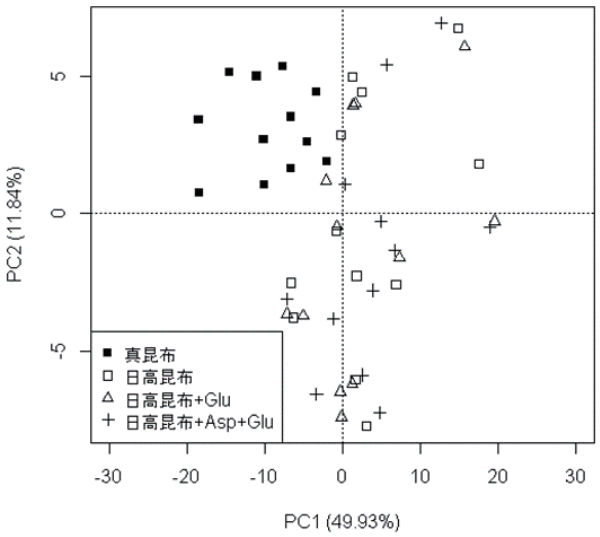


図 5　QDA 法による各出汁の評価値の主成分分析結果

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] 曽我 真子[*1] 兒嶋 朋貴[*2] 増田 達明[*2]
Yurie Iribe Mako Soga Tomoki Kojima 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．

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

図 1 実験風景

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

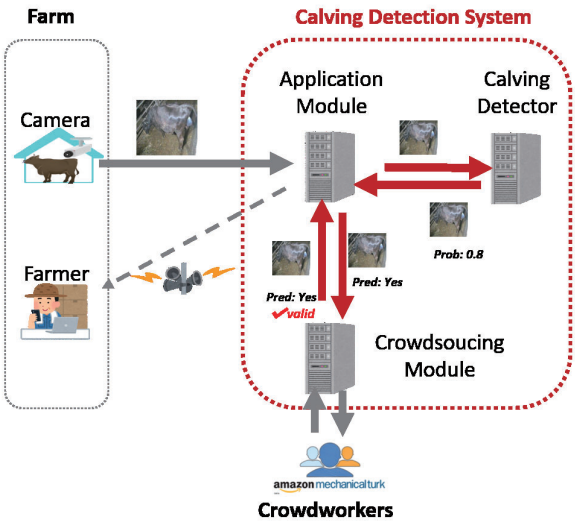


図 1: 映像による分娩検知システムの構成．アプリケーションモジュール（Application Module），分娩検知器（Calving Detector），クラウドソーシングモジュール（Crowdsourcing Module）の 3 つの部分からなる．

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．

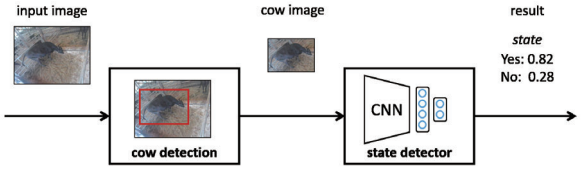


図 3: ニューラルネットワークによる分娩検知器の構成．まず牛の領域を検出し，次に尿膜・羊膜が露出しているか識別する．

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．

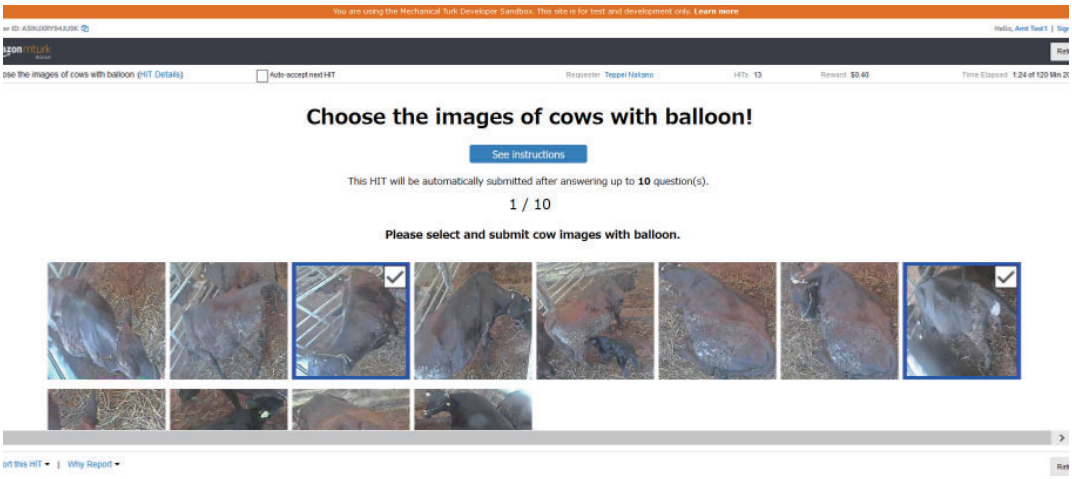


図 4: クラウドワーカーによる尿膜・羊膜の露出の検証タスク画面．複数提示された画像から，尿膜・羊膜が露出している画像を選択する．あらかじめ正解ラベルが付与された画像も提示され，それらの画像の回答をチェックし，ワーカーの信頼性を確認する．

- 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

○Tetsuo Ono[1] （1. Hokkaido University）

3:20 PM - 3:40 PM

[1G3-OS-13a-02] (Invited talk) Cognitive science of choice and manipulation

○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?

○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

○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 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．

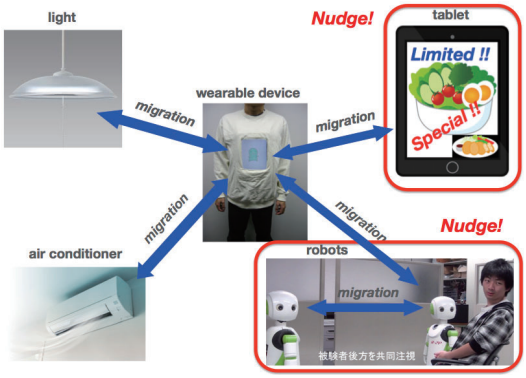


図 1 エージェントの「移動」による“ナッジ”の実装

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．

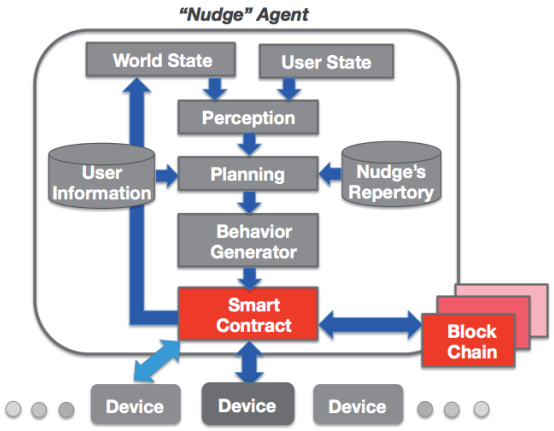


図 2 “ナッジ”エージェントのシステムの概要

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．

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

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

○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 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

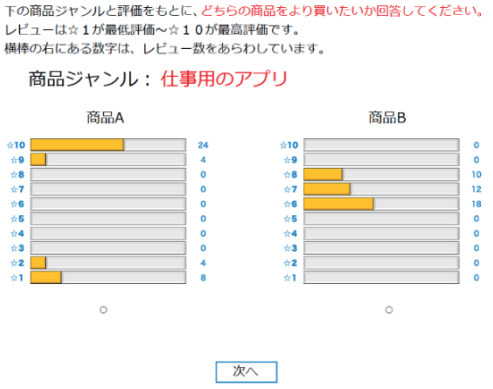


図 1. 実験画面． 商品ジャンルの性質は実用性．商品 A は高分散，商品 B は低分散．レビュー数は 40 のものを例示している．

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．

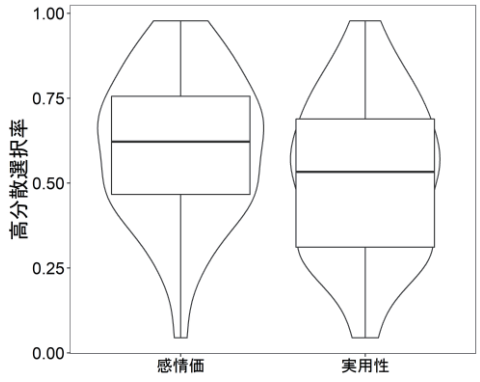


図 2. 実験結果

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．

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.

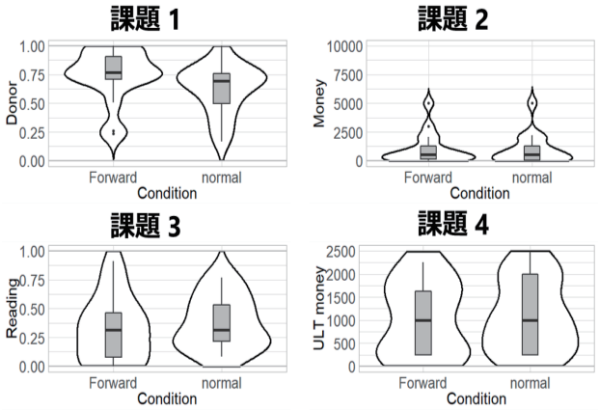


図 2: 実験 1 の結果. 左が前傾群(実験 1 の参加者), 右が通常群([白砂 18]における通常体勢群)の回答値の分布を示す.

[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.

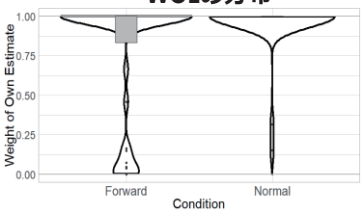
この実験課題は, 2 つのフェーズ[3]から構成された. 第 1 フェーズは実験冒頭に実施され, 参加者は 9 つの各出来事の西暦年および範囲について, 何も参照せず自身で推定を行った(このときの回答を「最初の推定値」と表記する). 第 2 フェーズは実験末尾に実施され, 参加者は第 1 フェーズと同じ課題に回答した. ただしこのとき, 参加者には「他人の推定値」が呈示されており, 参加者はこの値を参照しながら再度, 推定を行うことができた(このときの回答を「2 回目の推定値」と表記する). なお, この「他人の推定値」は, 実際には実験者により生成された乱数であった[4]. 仮説として, 前傾群の方が, 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

WOEの分布



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.

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.

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

○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

○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

○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 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．

$$\begin{aligned} p(\theta|s, a) &\propto p(a|s, \theta) \\ p(a|s, \theta) &= \frac{\exp(\beta Q(s, a; \theta))}{\sum_a \exp(\beta Q(s, a; \theta))} \end{aligned} \quad (1)$$

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

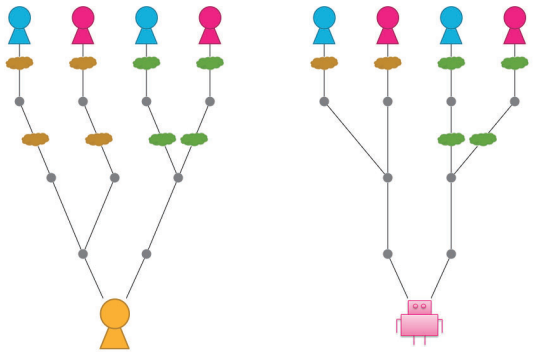


図 1: 参加者実験で用いた HAT タスク

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．

$$p(\theta|s,a) \quad = \quad \frac{\exp(\beta Q(s,a;\theta))}{\sum_a \exp(\beta Q(s,a;\theta))} \qquad (2)$$

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．

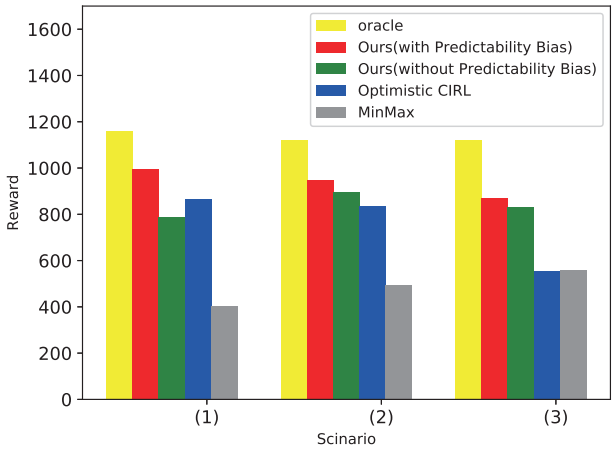


図 2: 各シナリオにおける各手法の性能評価

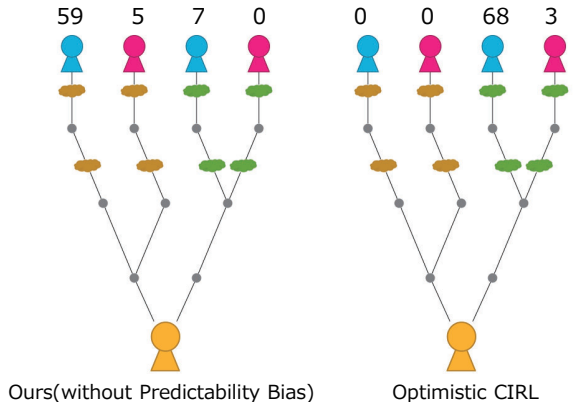


図 3: シナリオ 3 の例における参加者の最終目的地

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.

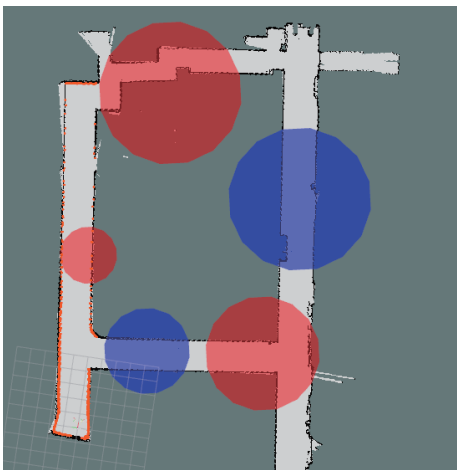


図 2: ゲインマップ

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

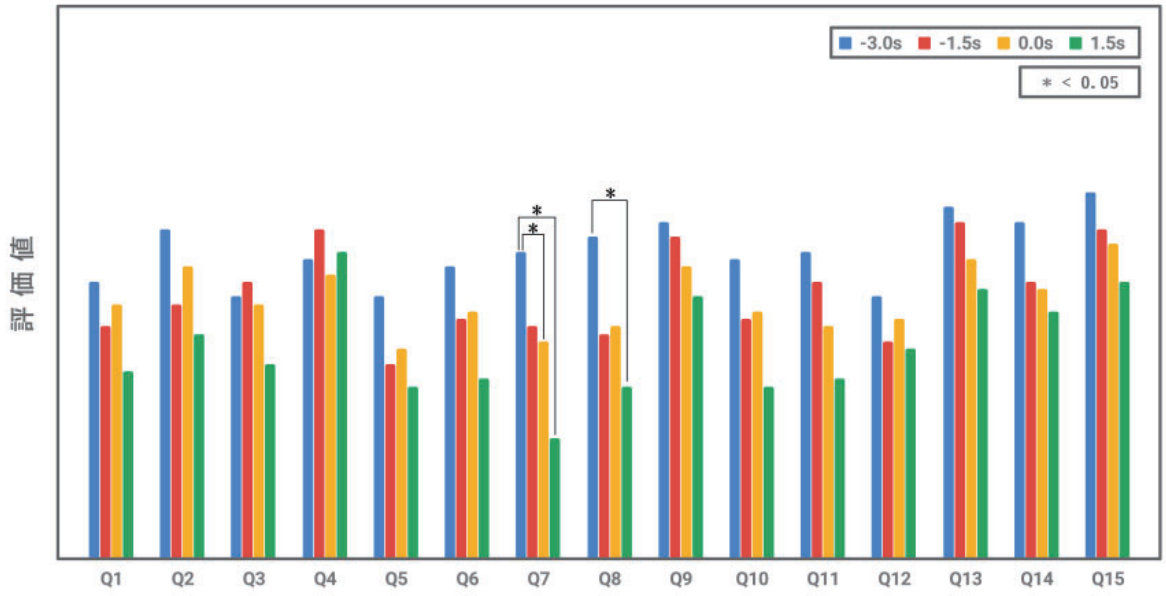


図 3: Mizusaki システム試乗後のアンケート結果

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

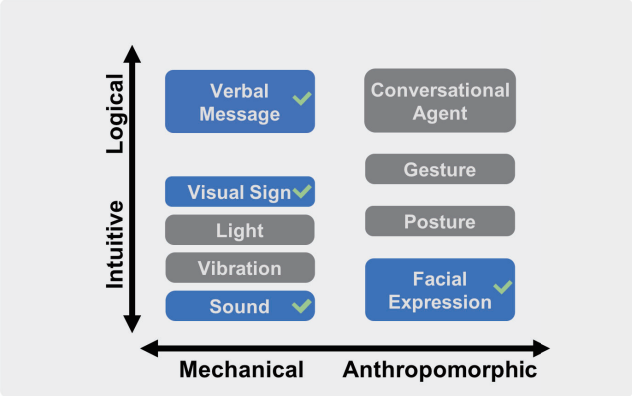


図 1: 信頼較正 CUE の分類

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

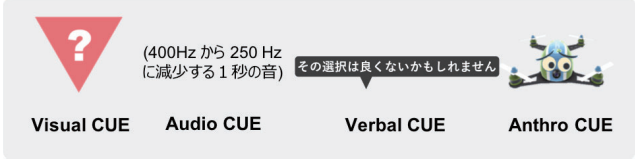


図 4: 実験で評価した 4 種類の信頼較正 CUE

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 ($F(4, 112) = 5.29, p = 0.0006, \eta_p^2 = 0.13$)．

*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

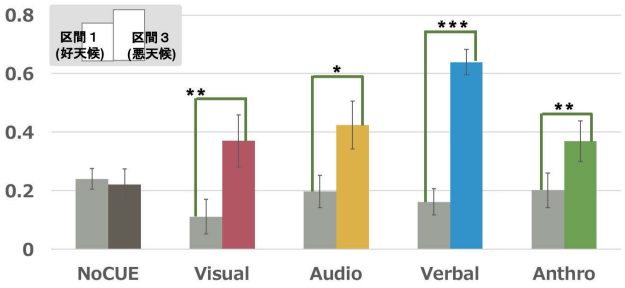


図 5: 信頼較正 CUE 毎の手動選択率（区間 1 と区間 3）

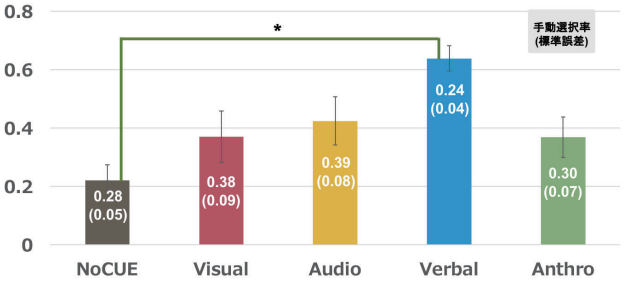


図 6: 信頼較正 CUE 毎の手動選択率（区間 3 での多重比較）

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

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

○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

○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

○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

○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 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

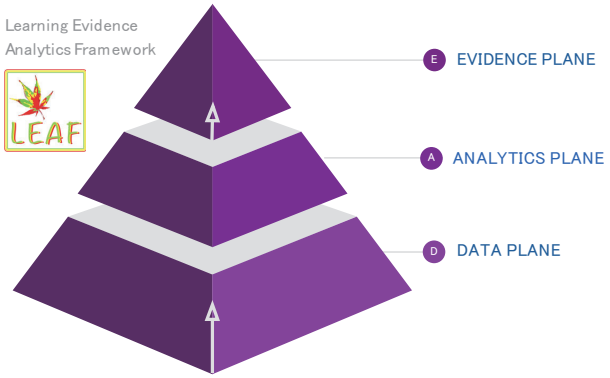


Figure 1. Extraction of evidence from teaching-learning data

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.

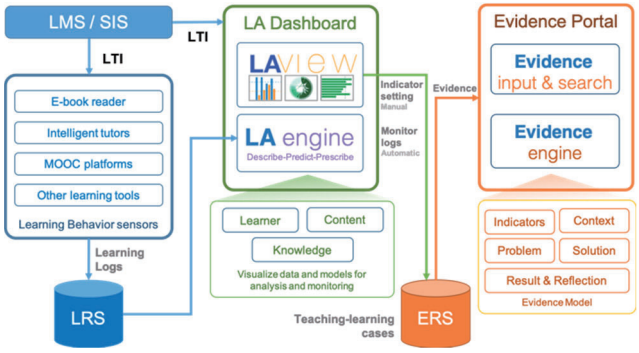


Figure 2. Components of Learning Evidence Analytics Framework (LEAF)

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　畠中 利治 *2　松田 岳士 *1
Nobuhiko Kondo　Toshiharu Hatanaka　Takeshi Matsuda

*1首都大学東京　*2大阪大学
Tokyo Metropolitan University　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．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．

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

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.

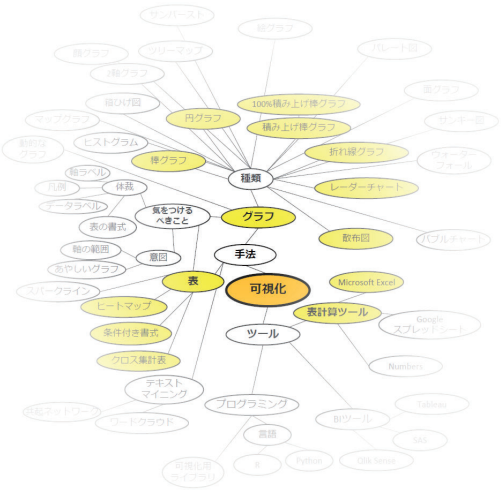


図 3: ヒントとしての概念マップの例（テーマ 1）

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．

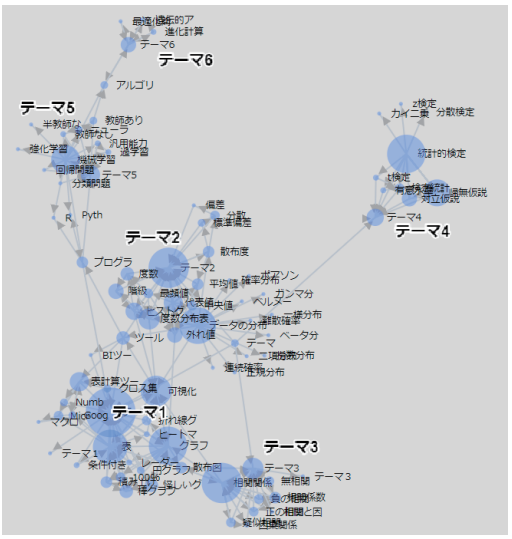


図 4: 学生 A のオンラインノートのリンク構造

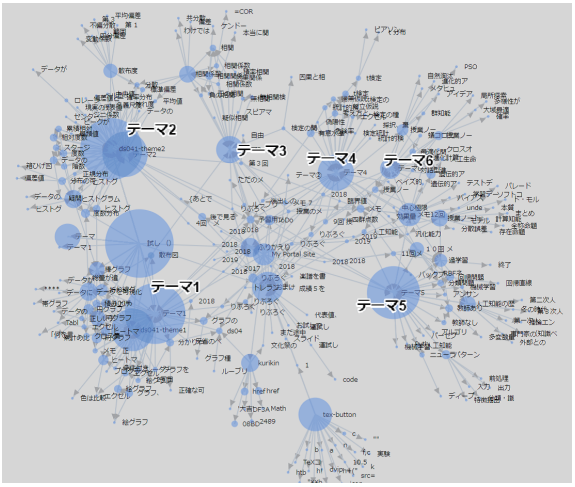


図 5: 学生 B のオンラインノートのリンク構造

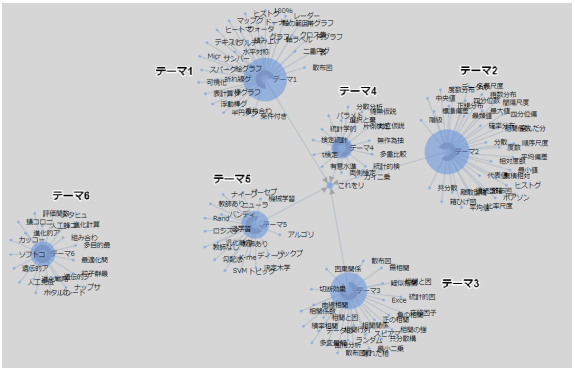


図 6: 学生 C のオンラインノートのリンク構造

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．

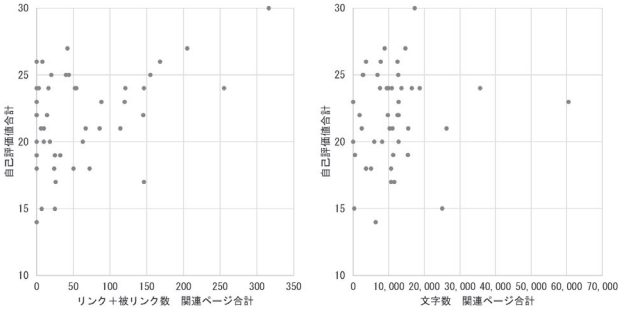


図 7: オンラインノート作成状態指標値と自己評価値の散布図

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

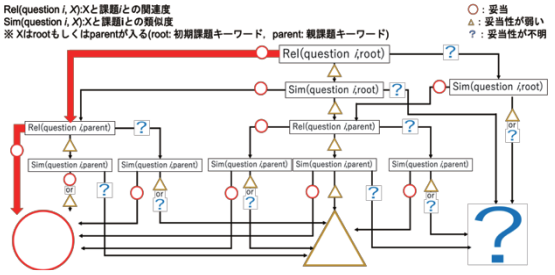


図 2：診断アルゴリズム

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

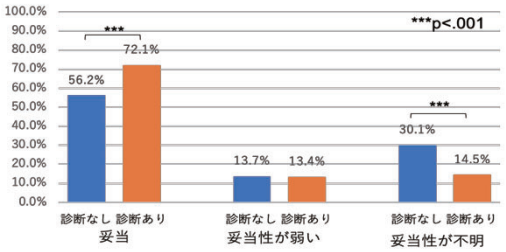


図 3：システムの提示結果

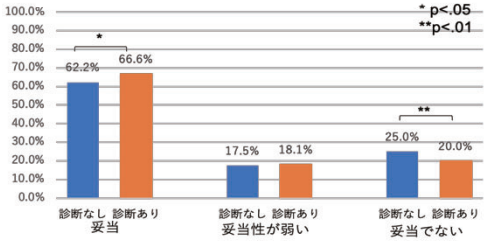


図 4：人手の評価結果

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

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

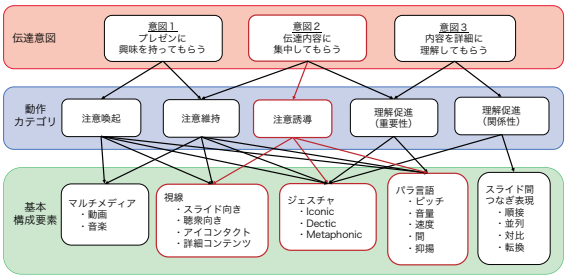


図 1 プレゼンテーション動作モデル

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．

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

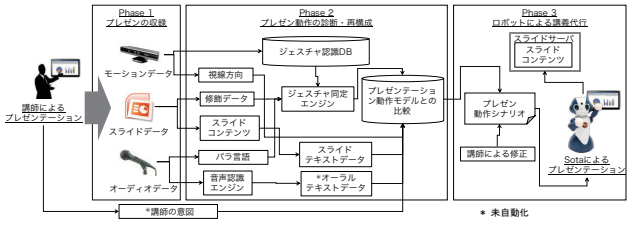


図 2 講義代行システムの全体構成図

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．

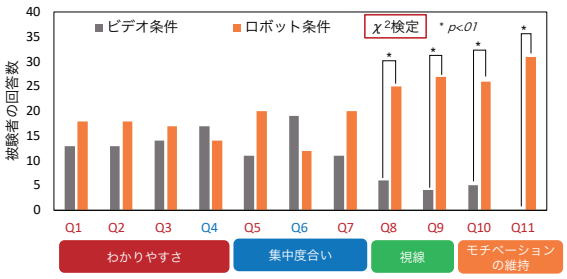


図 3 各質問項目の結果

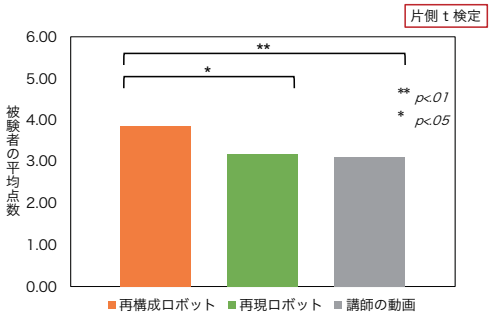


図 4 理解度テストの平均点数

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