

Thu. Jun 6, 2019

Room A

General Session | General Session | [GS] J-13 AI application

[3A3-J-13] AI application: enterprise and management

Chair: Kazutoshi Sasahara Reviewer: Hiroto Yoneno

1:50 PM - 3:10 PM Room A (2F Main hall A)

[3A3-J-13-01] Study on the mechanism of occurrence of quality spoofing by "Conjecture" "air" and "water" using business gameOHIROYASU SEITA¹, Setsuya Kurahashi¹ (1. Tsukuba university)

1:50 PM - 2:10 PM

[3A3-J-13-02] Detecting Technology Portfolios in the Semiconductor IndustryOBohua Shao¹, Kimitaka Asatani¹, Ichiro Sakata¹ (1. the University of Tokyo)

2:10 PM - 2:30 PM

[3A3-J-13-03] Relationships between mission statements and protability in scal year 2016 (Preliminary Result)ORyozo Kitajima¹, Ryotaro Kamimura², Hiroyuki Sakai¹, Kei Nakagawa³ (1. Seikei University, 2. IT Education Center, Tokai University, 3. Nomura Asset Management Co., Ltd.)

2:30 PM - 2:50 PM

[3A3-J-13-04] Influences caused by faultlines on the organizational performance.OFumiko Kumada¹, Setsuya Kurahashi¹ (1. University of Tsukuba)

2:50 PM - 3:10 PM

General Session | General Session | [GS] J-13 AI application

[3A4-J-13] AI application: electrical power

Chair: Takashi Onoda Reviewer: Yuiko Tsunomori

3:50 PM - 4:50 PM Room A (2F Main hall A)

[3A4-J-13-01] Evaluation of power consumption estimation model based on household informationOTomofumi Tahara¹, Hideaki Uchida¹, Hideki Fujii¹, Shinobu Yoshimura¹ (1. The university of Tokyo)

3:50 PM - 4:10 PM

[3A4-J-13-02] Optimization of Power Electric Supply Path in Smart GridsOTakaya Ozawa¹, Ei-Ichi Osawa¹ (1. Future University Hakodate)

4:10 PM - 4:30 PM

[3A4-J-13-03] Design and Preliminary Evaluations of Multi-Agent Simulation Model for Electric Power Sharing among HouseholdsOYasutaka Nishimura¹, Taichi Shimura², Kiyoshi Izumi³, Kiyohito Yoshihara¹ (1. KDDI Research Inc., 2. Kozo Keikaku Engineering Inc., 3. The University of Tokyo)

4:30 PM - 4:50 PM

Room C

General Session | General Session | [GS] J-9 Natural language processing, information retrieval

[3C3-J-9] Natural language processing, information retrieval: creation and analysis of stories

Chair: Hiromi Wakaki Reviewer: Masahiro Ito

1:50 PM - 2:50 PM Room C (4F International conference hall)

[3C3-J-9-01] Novel Segmentation Method based on the Distributed Representation of Sentences and Analysis Method of Story DevelopmentsOKiyohito Fukuda¹, Naoki Mori¹, Makoto Okada¹ (1. Osaka Prefecture University)

1:50 PM - 2:10 PM

[3C3-J-9-02] Analysis of Four-scene Comics Story Dataset based on natural language processingORyo Iwasaki¹, Naoki Mori¹, Miki Ueno² (1. Osaka Prefecture University, 2. Toyohashi University of Technology)

2:10 PM - 2:30 PM

[3C3-J-9-03] Search for Similar Story Sentences based on Role of Characters in order to Support and Analyze Contents Creator's IdeasOTakefumi Katsui¹, Miki ueno¹, Hitoshi Isahara¹ (1. toyohashi university of technology)

2:30 PM - 2:50 PM

General Session | General Session | [GS] J-9 Natural language processing, information retrieval

[3C4-J-9] Natural language processing, information retrieval: correction of documents

Chair: Yasutomo Kimura Reviewer: Yoko Nishihara

3:50 PM - 4:50 PM Room C (4F International conference hall)

[3C4-J-9-01] An Approach for Applying BERT to Sentence Elimination Problem in English Exam

OHiroimi Narimatsu¹, Hiroaki Sugiyama¹,
Genichiro Kikui², Hirotoishi Taira³, Seiki Matoba³,
Ryuichiro Higashinaka¹ (1. NTT Communication
Science Laboratories, 2. Okayama Prefectural
University, 3. Osaka Institute of Technology)

3:50 PM - 4:10 PM

[3C4-J-9-02] A consideration of word sense
disambiguation of company name utilizing
securities report

OHiroyuki Matsuda¹, Kazuhiko Tsuda¹ (1.
Graduate School of Business Sciences, University
of Tsukuba)

4:10 PM - 4:30 PM

[3C4-J-9-03] Misspelling Detection by using Multiple
Bidirectional LSTM Networks

ORyo Takahashi¹, Kazuma Minoda¹, Akihiro
Masuda², Nobuyuki Ishikawa¹ (1. Recruit
Technologies Co.,Ltd., 2. PE-BANK, Inc)

4:30 PM - 4:50 PM

Room H

General Session | General Session | [GS] J-7 Agents

[3H4-J-7] Agents: intelligence in/among robots

Chair:Keisuke Otaki Reviewer:Hidekazu Oiwa

3:50 PM - 5:10 PM Room H (303+304 Small meeting rooms)

[3H4-J-7-01] Effect of Robot Anxiety on the Appearance
Tendency of Uncanny Valley

Kazuhiro Ikeda¹, OTomoko Koda¹ (1. Osaka
Institute of Technology)

3:50 PM - 4:10 PM

[3H4-J-7-02] Strategyproof Mechanism with Agents
Grouping for Multi-Agent Pathfinding

OManao Machida¹ (1. NEC)

4:10 PM - 4:30 PM

[3H4-J-7-03] A Fundamental Study of Region Allocation
for Mobile Robots Based on Constraint
Optimization and Decentralized Solution
Method

OToshihiro Matsui¹ (1. Nagoya Institute of
Technology)

4:30 PM - 4:50 PM

[3H4-J-7-04] On the design of state value functions for
real-time continuous-state space multi-agent
decision making

OTomoharu Nakashima¹, Harukazu Igarashi²,

Hidehisa Akiyama³ (1. Osaka Prefecture
University, 2. Shibaura Institute of Technology, 3.
Fukuoka University)

4:50 PM - 5:10 PM

Room J

General Session | General Session | [GS] J-1 Fundamental AI, theory

[3J4-J-1] Fundamental AI, theory: search and
application

Chair:Ichigaku Takigawa Reviewer:Yoichi Sasaki

3:50 PM - 5:10 PM Room J (201B Medium meeting room)

[3J4-J-1-01] ACO with Pheromone Update by Negative
Feedback Can Solve CSPs

OTakuya Masukane¹, Kazunori Mizuno¹ (1.
Department of Computer Science, Takushoku
University)

3:50 PM - 4:10 PM

[3J4-J-1-02] An Algorithm for solving the Traveling
Salesman Problem using Clustering Method

OJumpei Uchida¹, Hajime Anada¹ (1. Tokyo City
University)

4:10 PM - 4:30 PM

[3J4-J-1-03] Algorithm of Traveling Salesman Problem
using Particle Swarm Optimization

OYuki Yamada¹, Hajime Anada¹ (1. Tokyo City
University)

4:30 PM - 4:50 PM

[3J4-J-1-04] League Scheduling for U12 Basketball

OTenda Okimoto¹, Kazuki Nishimura¹, Katsutoshi
Hirayama¹ (1. Kobe University)

4:50 PM - 5:10 PM

Room K

General Session | General Session | [GS] J-2 Machine learning

[3K3-J-2] Machine learning: analysis and validations
of models

Chair:Masahiro Suzuki Reviewer:Satoshi Oyama

1:50 PM - 3:30 PM Room K (201A Medium meeting room)

[3K3-J-2-01] Statistical Mechanical Formulation of
Learning Dynamics of Two-Layered Neural
Networks with Batch Normalization

OShiro Takagi¹, Yuki Yoshida¹, Masato Okada¹
(1. Graduate School of Frontier Sciences, The
University of Tokyo)

1:50 PM - 2:10 PM

[3K3-J-2-02] On the trade-off between the number of nodes and the number of trees in Random Forest

○So Kumano¹, Tatsuya Akutsu¹ (1. Kyoto University)

2:10 PM - 2:30 PM

[3K3-J-2-03] Do the AUC and log-loss evaluate CTR prediction models properly?

○Satoshi KATAGIRI¹ (1. F@N Communications, Inc.)

2:30 PM - 2:50 PM

[3K3-J-2-04] Social reinforcement learning with shared global aspiration for satisficing

○Noriaki Sonota¹, Takumi Kamiya², Tatsuji Takahashi¹ (1. Tokyo Denki University, 2. Graduate School of Tokyo Denki University)

2:50 PM - 3:10 PM

[3K3-J-2-05] On the mathematical approach to the "photo-likeness" of images

○Yasuhiko Asao¹, Ryotaro Sakamoto¹ (1. Graduate School of Mathematical Science, the University of Tokyo)

3:10 PM - 3:30 PM

General Session | General Session | [GS] J-2 Machine learning

[3K4-J-2] Machine learning: real world interaction

Chair:Daiki Kimura Reviewer:Hikaru Kajino

3:50 PM - 5:30 PM Room K (201A Medium meeting room)

[3K4-J-2-01] Consideration on Generation of Saliency Maps in Each Action of Deep Reinforcement Learning Agent

○Kazuki Nagamine¹, Satoshi Endo², Koji Yamada², Naruaki Toma², Yuhei Akamine² (1. Information Engineering Course, Graduate School of Engineering and Science, University of the Ryukyus, 2. Faculty of Engineering, School of Engineering Computer Science and Intelligent Systems, University of the Ryukyus)

3:50 PM - 4:10 PM

[3K4-J-2-02] Dynamic Reward Clustering

○Ryota Higa¹, Junya Kato¹ (1. NEC Corporation)

4:10 PM - 4:30 PM

[3K4-J-2-03] A dialogue system implemented with latent parameters

○Weida Li¹, Chie Hieida², Takayuki Nagai² (1.

Seiko Gakuin High School, 2. The University of Electro-Communications)

4:30 PM - 4:50 PM

[3K4-J-2-04] Comfortable Driving by Deep Inverse Reinforcement Learning

○Daiko Kishikawa¹, Sachiyo Arai¹ (1. Chiba University)

4:50 PM - 5:10 PM

[3K4-J-2-05] Linear function approximation of Cognitive Satisficing Function

○Yu Kono^{1,2} (1. Tokyo Denki University, 2. DeNA, Co., Ltd.)

5:10 PM - 5:30 PM

Room N

General Session | General Session | [GS] J-10 Vision, speech

[3N3-J-10] Vision, speech: voice and communication

Chair:Masanori Tsujikawa Reviewer:Jun Sugiura

1:50 PM - 2:30 PM Room N (Front-right room of 1F Exhibition hall)

[3N3-J-10-01] Multilingual Imputation Using Transfer Learning for Estimating Emotion from Speech

○Koichi Sakaguchi¹, Shohei Kato^{1,2} (1. Dept. of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology, 2. Frontier Research Institute for Information Science, Nagoya Institute of Technology)

1:50 PM - 2:10 PM

[3N3-J-10-02] Development of Open-source Multi-modal Interaction Platform for Social Experiment of Conversational User Interface

○Akinobu Lee¹ (1. Nagoya Institute of Technology, Japan)

2:10 PM - 2:30 PM

General Session | General Session | [GS] J-10 Vision, speech

[3N4-J-10] Vision, speech: applications to industries

Chair:Masanori Tsujikawa Reviewer:Tomoya Yoshikawa

3:50 PM - 4:50 PM Room N (Front-right room of 1F Exhibition hall)

[3N4-J-10-01] Measurement of growing situation of agricultural crops on FPGA-mounted drone using Circle SSD

OTakuma Yoshimura¹ (1. poco-apoco Networks Co.Ltd.)

3:50 PM - 4:10 PM

[3N4-J-10-02] Prediction of Favorability Rating on Beer-Can Package Designs Using Convolution Neural Network and Visualization by Class Activation Mapping.

OHiroyuki Shinohara¹, Tatsuji Ishiguro¹, Shunsuke Nakamura², Toshihiko Yamasaki² (1. Kirin Company, Limited, 2. The University of Tokyo)

4:10 PM - 4:30 PM

[3N4-J-10-03] Conversion of Floor Plan Images to Graph Structures using Deep Learning and Application to Retrieval

OMantaro Yamada¹, Xueting Wang¹, Toshihiko Yamasaki¹, Kiyoharu Aizawa¹ (1. the University of Tokyo)

4:30 PM - 4:50 PM

Room P

General Session | General Session | [GS] J-7 Agents

[3P4-J-7] Agents: social multiagents

Chair:Naoki Fukuda Reviewer:Jun Ichikawa

3:50 PM - 5:30 PM Room P (Front-left room of 1F Exhibition hall)

[3P4-J-7-01] Explainable Compromising Algorithm based on Constraint Relaxation for Automated Negotiating Agents

OShun Okuhara^{1,2}, Takayuki Ito² (1. Fujita Health University, 2. Nagoya Institute of Technology)

3:50 PM - 4:10 PM

[3P4-J-7-02] An allocation strategy with deep reinforcement learning for efficient task processing in multi agent system

OGenki Matsuno¹, Sho Tanaka², Hiroki Hara², Syunyo Kawamoto², Syo Shimoyama², Takashi Kawashima², Daisuke Tsumita², Yasushi Kido¹, Osamu Hashimoto¹, Tomohiro Takagi² (1. Skydisc, Inc., 2. Meiji University)

4:10 PM - 4:30 PM

[3P4-J-7-03] Omoiyari as Filling Gaps Making Collective Adaptation

OYoshimiki Maekawa¹, Fumito Uwano¹, Eiki Kitajima¹, Keiki Takadama¹ (1. The University of Electro-Communications)

4:30 PM - 4:50 PM

[3P4-J-7-04] Investigation of online simulation method of social consensus formation

OYasuko Kawahata¹, Akira Ishii², Takuya Ueoka¹ (1. Gunma University, 2. Tottori University)

4:50 PM - 5:10 PM

[3P4-J-7-05] Estimation of agent's rewards with multi-agent maximum discounted causal entropy inverse reinforcement learning

OKeiichi Namikoshi¹, Sachiyo Arai¹ (1. Chiba University)

5:10 PM - 5:30 PM

Room Q

General Session | General Session | [GS] J-13 AI application

[3Q3-J-13] AI application: analysis of physical behaviors in artifacts

Chair:Takuya Hiraoka Reviewer:Yoichi Sasaki

1:50 PM - 3:10 PM Room Q (6F Meeting room, Bandaijima bldg.)

[3Q3-J-13-01] Construction of Dataset for Feature

Extraction Performance Evaluation using Aerial Photographs

OHiroyuki Ohno¹, Ryo Endo¹, Takayuki Nakano¹, Masako Shinoda¹ (1. Geospatial Information Authority of Japan)

1:50 PM - 2:10 PM

[3Q3-J-13-02] Slime detection during pile construction using machine learning

OSohei Arisaka¹, Yuki Tamagawa¹, Kojiro Takesue¹ (1. Kajima Corporation)

2:10 PM - 2:30 PM

[3Q3-J-13-03] The optimization and comparison of methods for the Air foil design using Deep Reinforcement Learning.

OHitoshi Hattori¹, Kazuo Yonekura¹ (1. IHI Corporation)

2:30 PM - 2:50 PM

[3Q3-J-13-04] Application of Gradient Booting regression toward the Computational Fluid Dynamics in the Manufacturing industry

OYutaro Ogawa¹, Takuya Shimizu¹, Toshiaki Yokoi¹ (1. INFORMATION SERVICES INTERNATIONAL-DENTSU, LTD.)

2:50 PM - 3:10 PM

General Session | General Session | [GS] J-13 AI application

[3Q4-J-13] AI application: transformation system

Chair:Masahiro Tada Reviewer:Masayuki Otani

3:50 PM - 5:10 PM Room Q (6F Meeting room, Bandaijima bldg.)

[3Q4-J-13-01] Traffic anomaly detection using ETC2.0

probe data

○Atsuki Masuda¹, Masaki Matsudaira¹ (1. Oki
Electric Industry Co.,Ltd.)

3:50 PM - 4:10 PM

[3Q4-J-13-02] Design method for high efficiency drone

highway network

○Masatoshi Hamanaka¹ (1. RIKEN)

4:10 PM - 4:30 PM

[3Q4-J-13-03] Classification for time-sequence data

appeared in shift control of automobile
automatic transmission

○Yusuke Morikawa¹, Yasuhiro Ishihara¹,
Takanori Ide¹, Eiji Moriyama¹, Taku Akita¹, Yasuo
Tabei², Takehito Utsuro³, Hiroshi Nakagawa²

(1. Aisin AW Co., Ltd., 2. RIKEN, Center for
Advanced Intelligence Project, 3. Fclty. Eng, Inf.
&Sys, Univ. of Tsukuba)

4:30 PM - 4:50 PM

[3Q4-J-13-04] Likelihood distribution of Pedestrian

Trajectories rendered by Variational
Autoencoder

○Yasunori Yokojima¹, Tatsuhide Sakai² (1.
Siemens K.K., 2. Great Wall Motor)

4:50 PM - 5:10 PM

[3A3-J-13] AI application: enterprize and management

Chair:Kazutoshi Sasahara Reviewer:Hiroto Yoneno

Thu. Jun 6, 2019 1:50 PM - 3:10 PM Room A (2F Main hall A)

[3A3-J-13-01] Study on the mechanism of occurrence of quality spoofing by "Conjecture" "air" and "water" using business game

○HIROYASU SEITA¹, Setsuya Kurahashi¹ (1. Tsukuba university)

1:50 PM - 2:10 PM

[3A3-J-13-02] Detecting Technology Portfolios in the Semiconductor Industry

○Bohua Shao¹, Kimitaka Asatani¹, Ichiro Sakata¹ (1. the University of Tokyo)

2:10 PM - 2:30 PM

[3A3-J-13-03] Relationships between mission statements and protability in scal year 2016 (Preliminary Result)

○Ryozo Kitajima¹, Ryotaro Kamimura², Hiroyuki Sakai¹, Kei Nakagawa³ (1. Seikei University, 2. IT Education Center, Tokai University, 3. Nomura Asset Management Co., Ltd.)

2:30 PM - 2:50 PM

[3A3-J-13-04] Influences caused by faultlines on the organizational performance.

○Fumiko Kumada¹, Setsuya Kurahashi¹ (1. University of Tsukuba)

2:50 PM - 3:10 PM

ビジネスゲームを用いた「忖度」「空気」と「水」による品質詐称発生メカニズムに関する研究

Study on the mechanism of occurrence of quality spoofing by "Conjecture" "air" and "water" using business game

脊板弘康^{*1}
Hiroyasu Seita

倉橋節也^{*2}
Setsuya Kurahashi

^{*1} ^{*2} 筑波大学システム情報工学研究科リスク工学専攻
University of Tsukuba Faculty of system and information Engineering Department of Risk Engineering#1 #2

Abstract

The quality spoofing case continues in a part of Japanese companies that have been sweeping the world with its top priority on customer first principles and raising quality first principles. In these cases of misrepresentation, as seen in the Akafuku case of the food fraud, there are cases in which the conjecture to a specific target worked. Until now, each company has been absolutely quality based on quality first principle, why is it so fragile and crumbling why?

In this research, we use a business game based on the framework of the Giddens 's theory of structuring to show that "quality" is absolutely made with "Conjecture" and "air", which can be said as unique culture of Japan as Mr. Shichihei Yamamoto says, and furthermore that it is collapsed due to real problems,.

1. 研究の背景

1.1.問題意識～日本企業の品質第一主義の組織文化と企業不祥事

生産活動においてZD運動はQC活動とともに日本の品質管理において主要な行動規範の一つになっており、近年においても日本の物造り産業における生産活動やサービス産業でも重要視され、多大な貢献があるといわれている。一方で、不良ゼロの合言葉は品質偏重による管理コスト増大、ひいては高コスト体質の原因になっているとの指摘もある。[脊板 .18]が行った研究では、生産の第一線において品質や生産コスト決定の中心的役割を担う工程エンジニアの不良発生時の問題解決行動を分析することにより、コストが重要視されている生産現場においても品質がコストに比較して圧倒的に重視されていることが明らかにされた。

しかし、近年においても神戸製鋼、東洋ゴム、タカタベルトなど日本を代表する製造メーカーで品質詐称事件が相次いで発生しているのはなぜであろうか。

2. 本研究テーマに関する既存研究

2.1. 詐称が起こるメカニズム

米国の組織犯罪研究者[Cressey.73]は詐称が発生メカニズムを「不正のトライアングル理論(The Fraud Triangle)」として次のように体系化した。

- ①「プレッシャー」(Pressure)～例えば、納期や生産性に関する過大な欲求を満たさねばならないなど心理的圧力。
- ②「機会」(Opportunity)～不正を行おうとすれば可能な環境が存在する状態。

- ③「正当化」(Rationalization)～「従業員のコンプライアンスの意識の欠如」や会社の仕組みが脆弱で詐称を起こしても発覚しにくいといった環境や仕組みに由来するもの。

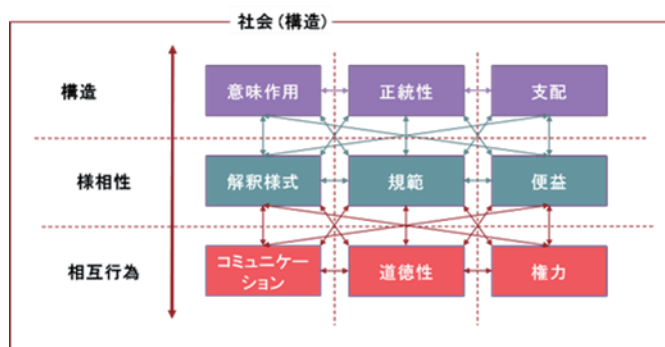
これらは、多くの詐称事例に共通して見られるものであるが、あくまでも定性的である。各要素がどの程度の大きさになれば実際の詐称に至るのか定量的にとらえることは具体的な原因検討や対策において重要であるが、その定量化については検討されていない。

また、[リーズン.91]によれば、組織不祥事は組織の持つ潜在的危険性をカバーする深層防御(規則や手順書、訓練、管理業務、資格認定といったソフトな防御と工学的な安全施設・設備や警報、非破壊検査などのハード防御)が、組織的要因(経営層の意識決定、予算配分、人員配置、計画、意思疎通、管理など)、局所的な作業現場要員(過度のタイムプレッシャー、不適切な道具や装備、訓練不足、人手不足など)によって誘発される従業員たちの不安全行為によってほころびること、潜在的だった危険性が顕在化し、それによって組織全体ないし組織の外部にまで損害がもたらされるとしている。

リーズンの理論は具体的な対策につなげられる要素を検出しやすいが、個人、組織、社会の相互作用や時系列的な構造の発展過程の描写力に欠ける側面がある。

2.2 Giddens の再帰による組織風土醸成と間嶋の詐称不祥事への展開

[Giddens..76]は、二元論に依拠する機能主義や解釈主義などの社会文化論の問題点を構造化理論(社会構造(規則と資源)を、個人行為によって創られるものである一方で、同時にそのような社会構造を創る個人行為そのものを再帰的に作り出す二重の性格も持つとし、この個人行為と社会構造の創り創られる関係を述べた。しかし、個人、組織、社会の相互作用や構造の発展過程を上手く表現できる基本的な枠組みを与えているが、詐称問題についての言及は見られない。



出所：A.Giddens, New Routes of Sociological Method, Hutchison 2nd ed. 1993.

図-2. ギデنزの構造化の理論による相互行為、様相性、構造の関係

間島[間島.07]]は A. ギデنزの構造化理論を組織不祥事に展開し [Barley.86]の時系列での相互作用による構造変化、ミクロ・マクロリンク（個人と社会との相互作用）を包含したモデルを組織不祥事に適用した。その結果、個人と組織、社会の間の創り創られていく組織文化がいつしか歪んで行き組織不祥事に発展していく様子をモデル化した。ギデنزの枠組みを詐称問題に適用することで詐称の過程を個人、組織、社会の相互作用を時系列的な変化とともに上手く描けるようになっていく。しかし、具体的な原因の特定やその効果の定量化ができておらず、具体的な対策の提供力に乏しい。また、「付度」など日本独特の社会風土の問題を取り込めていない。

2.3 山本による付度による「空気」の絶対化と相対化による詐称発生メカニズム

[山本.83]によれば、実際の社会での会話や議論において、誰もが感じ、全員がそうであると認識していながら、それを否定することなどは許されない状況を「空気」といい、これらは根拠や事実に基づかない架空の信念を前提にしている。この空気はしばしばそれを象徴するシンボルとなるスローガンや偶像をともなっている。これを品質になぞらえると日本企業を支配している「品質は絶対である（品質第一主義）、品質を悪くしたら企業は存続できない」という信念にあたる。日本企業においてはお客様に不良品を渡すのは悪でありタブーであるとされ顧客に過剰な配慮、すなわち「付度」が発生し、十分に良品であるにもかかわらず「もしかすると、大きな不良とみられ信用を失う可能性がある」として絶対化された品質に対して過剰なコストをかけるわけである。

ところが、これらの呪縛は何らかの現実に基づく前提、例えばコストや納期が現実の問題として提起されると絶対化されていたものはやがて相対化されるのである。この現実に基づく仮定、例えば「納期が守れないと顧客の信用を失う」が対立概念（これを山本は「水」という）として現れると品質第一主義は相対化し、さらにこの「水」が連続的に作用し続ける（これを山本は「雨」とよぶ）とやがて納期が品質の重要度を上回り、納期を守るために詐称に至るという多くのケースに見られる不祥事の構図となり詐称のメカニズムが上手く説明されるのである。

3. 本研究の目的～ビジネスゲームを用いた「付度」「空気」「水」による品質第一主義の相対化メカニズムの検証

そこで、本研究では不祥事発生メカニズムを日本独特の文化ともいえる顧客に対する「付度」と「空気」により絶対化されてい

た「品質」が、現実の諸問題（例えば納期遵守への要請）である「水」を注されることによって相対化し胡散霧消していくという山本の論説を、クレッシーから Giddens から間嶋に至る組織文化の構造化の理論を基に、ビジネスゲームを用いて再現することを目的とする。

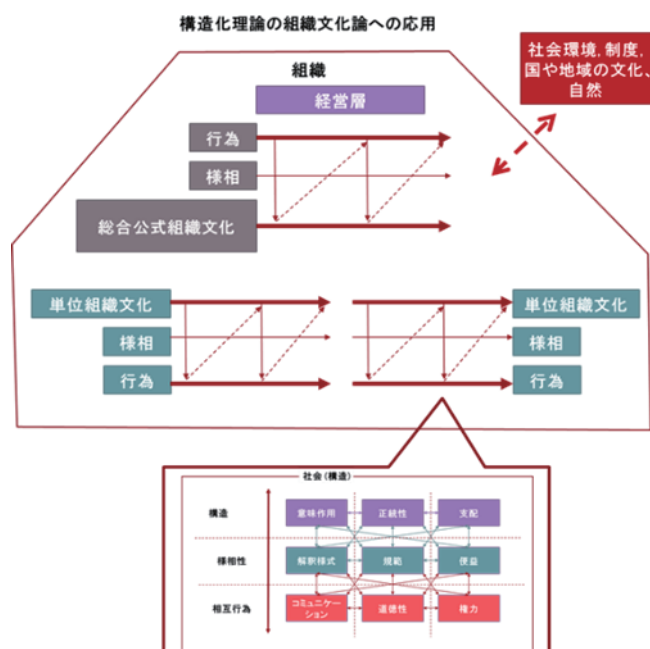


図-3. 間嶋による構造化の理論の組織文化への応用

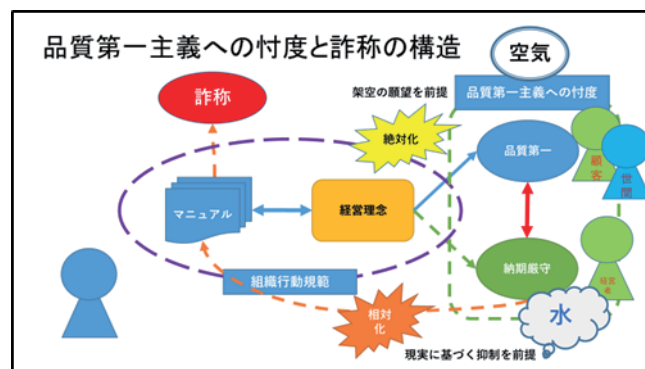


図-4. 付度、空気による品質の絶対化と水による相対化

4. ビジネスゲームのデザイン

4.1 ビジネスゲームの仕様

以下のシナリオとルールで設定されたビジネスゲームで複数グループを対象として行うものとする。また、形式としては以下の様なシナリオでの RPG（ロールプレイングゲーム）とした。

(1) ゲームの設定

本ゲームは、イベントを品質不良の発生時の救済申請処理として、定められた救済ルールに基づき救済方法を決定し得点を競うものである。

■ ゲームマニュアルで示したシナリオ

- あなたは、品質部門担当者です。あなたは、製品の出荷判断に責任を持っています。
- 現在生産している製品は、次期主力となる期待の製品ですがなかなか品質が安定せず出荷にあったっては供給先との兼ね合いで判断を加えながらの判定が必要です。

- 当社では、こうした品質が安定しない製品に対して特別救済処置が設定されています。
- この特別救済処置には、顧客救済と社内救済の二種類があり品質不良の程度に応じて使い分けことが求められます。
- このゲームでは、発生する不良状況に応じて社内救済か顧客救済かを判定し納入による利益(銀貨)を守りながら、顧客損失と会社損失である IFC(金貨)の損失を最小にすることで得点を競います。
- 顧客救済を行うと顧客に納入が拒否され利益はなく IFC 損失となる場合があります。社内救済の場合はそのままペナルティなしで納入されます。
- 判定は手元の1～9のカードで示してください。カードの大きさが9に近いほど社内救済とし 5 以上なら社内救済と判定します。

4.2 ゲームにおける「付度」「空気」「水」と付度の評価方法

付度とは、日本のハイコンテキスト社会において、指示されずとも相手の気持ちを推し量り応えようとするものである。このゲームでは参加者個人が、顧客の意向をくみ取り、顧客や本来の判定基準(マニュアルの規定)やゲームでの勝ち負けを度外視して過剰に判定を厳しくし、社内救済とする所をわざわざリスクのある顧客救済とする事である。従って、このゲームにおける「付度」とは顧客の反応であり、「空気」は「このような品質では顧客の信用を失うので出荷はできない」という雰囲気である。これに対する「水」とは「納期」「救済基準となるガイドライン」「上司の命令」など現実的な問題を指している。

付度の大きさ(「付度」度とする)は、本来ゲームが要請する判定結果(このゲームにおいてはルール上“9”となるはず)と実際の各個人またはグループの判定結果の差として表すことが出来る。そこで、付度の評価方法としては、付度度を S、ルール基準の判定値を R(=9)、グループ判定値を G とする時次式で表すことが出来る。

$$S = R - G \quad \dots \dots \text{eq.1}$$

6. 実験結果

ゲーム参加者の構成は、大学院教員 1 名を含む 10 名、男性 8 名、女性 2 名、年齢は 20 代後半から 50 代後半である。予め用意した設問の回答から「誠実な思考の A グループ」と「合理的な思考の B チーム」に分けて行いゲーム終了後ディブリーフィングを行った。

各設問における判定要因と対応するグループ判定、ゲームから要請される合理的な判定結果を表-1. にまとめる。また、個人及びグループの判定結果を図-5.に示す。

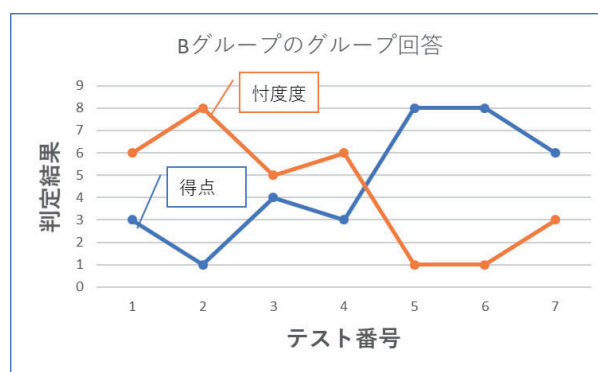
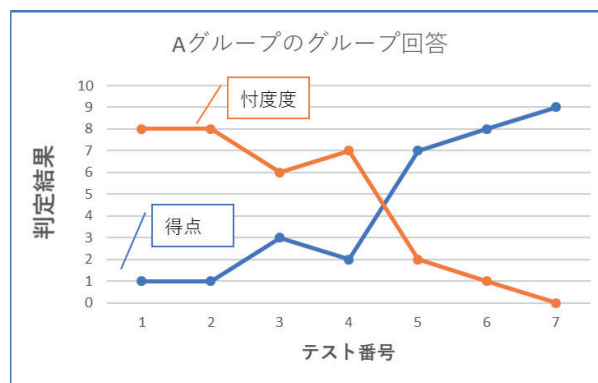
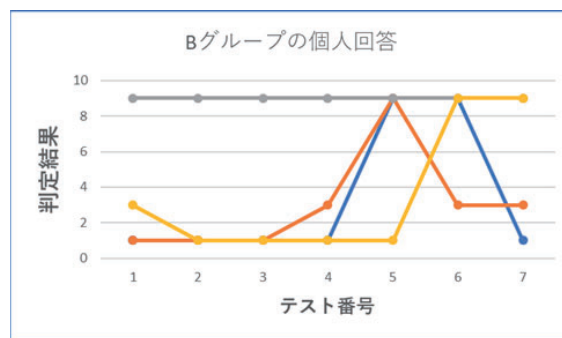
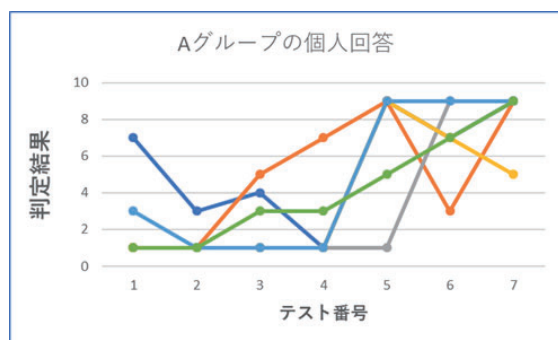


図-5.各設問に対する個人及びグループの判定結果と「付度」度

7. 考察

グループの判定(表-1 参照)を見ると、問1～4はどちらのグループも同様にゲームの要請からすれば社内救済となる案件も顧客救済を選んでいる。これらは仕様項目に該当するか、少しでも疑わしいものはコストの掛かる顧客救済としている。設問 5 においてはゲームのルールからは責任者の了解が得られていないので顧客救済とすべきであるが、それに関わらず顧客の了解が得られているため社内救済と判定しており、顧客への強い付度が働いていると思われる。実際、その後のディブリーフィングでも「お客様が良いというのだから」との発言が出た。設問6では明らかな不良ではあるが社内救済としている。これは社長承認という大きな権力とガイドラインの要請が「水」の機能を果たしたものと考えられる。問7では要求項目ではなくかつ、納期遅延という「水」の効果により社内救済となっている。今回質問の8において A の誠実グループの全員が社内救済を選んでいるのに対して、B の合理的なグループは顧客救済と社内救済に分かれた。ゲーム後のディブリーフィングでの当事者の発言において「責任の所在が自分ではなく社長にあることを明確にするために顧客救済とした」と述べており、集団での責任より個人の責

任を重視している。これらの行動は集団が権力に「付度」する「空気」に支配されて詐称に至る事に対する抑止（「水」）がゲーム上でも表れたとも考えられる。

8.結論

参加者は、ゲームであると認識しながらも品質判定に際して社内救済にする事に強い抵抗を示しており、本来、社内救済で済ますことのできる問題も、コストのかかる顧客救済を選んでいる。これは、山本の言う、現実（ゲーム空間上）に存在しない顧客に「付度」し、ゲーム上の合理的な解を選択できない「空気」に支配（絶対化）されている状態にあった事が本ゲームにより示されたといえる。また、山本が言うところの「水」である現実を前提とした納期遅れや、上司の命令、マニュアルの提示があった設問では「付度」が相対化され、顧客救済から社内救済に変容する過程を示しているといえよう。これらのことをギデنزや間嶋の構造化の理論による組織文化形成過程に置き換えるなら、会社組織を構成する個人と社会の相互作用により再帰的に形成された「品質第一」という倫理的に

も社会的にも強い拘束力を持つ組織規範が、納期遅れやそれに伴う損失コスト、あるいは権力者の命令により変容しその絶対性が低下することを示しているのである。

9.今後の課題

今回のゲームでは意図的に社内救済となる様に極端な条件（顧客救済とすると損失リスクがある一方で社内救済は損失リスクが全くない）を設定した。これは、ゲーム時間が十分とれないため、繰り返しが少ない中で結果を得ようとしたためであるが、ゲームの面白さ、没入感を醸成するためには、社内選択や顧客救済時の不正発覚リスクを導入し、より現実感を高めるゲーム設計を行う必要が有る。

更に、このような「水」の要素を多く使用した場合の挙動を観察することで詐称に至る各要素の影響を定量的に明らかにしたい。

今回質問の8において A の誠実グループの全員が社内救済を選んでいるのに対して B の合理的なグループは顧客救済と

表-1.ゲームの設問とグループ回答に対する判定の評価結果

	問題設定条件	仕様対象項目	規格外れ	納期	性能	他者振替	顧客の容認	インセンティブ	権力	A	B	Aの判定結果	Bの判定結果	ゲームルールによる判定結果	正答
問1	製品に顧客要求規格にはない項目で軽い欠陥あり。納期は明日である。	×	△	×	○					1	3	顧客	顧客	社内	×
問2	製品性能が規格上限なのでぎりぎり大丈夫なB社に納入することにした	○	○		○	○				1	1	顧客	顧客	社内	×
問3	IFC目標内で品質規格は測定誤差の範囲内で外れ気味。顧客は安全係数が高く問題ない。	○	○		○					3	4	顧客	顧客	社内	×
問4	工程内の測定器ではNGだが、再度測定したらOKとなった。	○	○							2	3	顧客	顧客	社内	×
問5	規格を超えるが顧客側の品質には影響を与えない。もともと急な要求で顧客側も多少なら構わないといわれている。また、納期を守れば顧客の信用アップで金貨もらえる。	○	×	×	○		○	○		7	8	社内	社内	顧客	×
問6	明らかな規格外だが大きな商談が絡むため社長決裁で救済するよう命令が出た。	○	×		×				○	8	8	社内	社内	社内	○
問7	製品に顧客要求項目ではない外観不良があった。製品性能や用途的には問題なさそうである。明日までに納入しないと遅延になる。	×	△	×	○					9	6	社内	社内	社内	○
		○；規格対象	○；規格内	×	○；問題なし										
		×	△；軽度		×										
			×												

社内救済に分かれた。ゲーム後のディブリーフィングでの当事者の発言において「責任の所在が自分ではなく社長にあることを明確にするために顧客救済としたと」述べており、集団でのこのような多様性が詐称にどのように影響するのかについても明らかにしたい。

10.参考文献

[脊板 .18] 脊板弘康(2018)「工程エンジニアの問題解決方法選択行動にみる品質偏重傾向に関する研究」,筑波大学大学院修士論文,2018

[Cressey.73] Donald R. Cressey(1973), Other People's Money (Montclair: Patterson Smith, 1973), 30.

[リーズン.91] ジェームズ・リーズン「組織事故一起こるべくして起こる事故からの脱出」(1991),日科技連

[Giddens..76] Anthony Giddens(1976),New Rules of Sociological Method: A Positive Critique of Interpretative Sociologies, (Hutchinson, 1976, 2nd ed., 1993).

[間島.07] 間島 崇(2007)「組織不祥事―組織文化論による分析―」文真堂

[Barley.86] Barley,S.R.,”Technology as an Occasion for Structuring: Evidence from Observations of CT Scanners and the Social Order of Radiology Department", Administrative Science Quarterly,31[1986]

[山本.83] 山本七平(1983)「空気の研究」文春文庫

[榎本.17] 榎本博明(2017)『「付度」の構造～空気を読みすぎる部下、責任を取らない上司』,イースト書籍

[片田.17] 片田珠美(2017)「付度社会ニッポン」, 角川新書

Detecting Technology Portfolios in the Semiconductor Industry

Bohua Shao^{*1}

Kimitaka Asatani^{*1}

Ichiro Sakata^{*1}

^{*1} School of Engineering, the University of Tokyo

Semiconductor serves as the base for the Artificial Intelligence Society. We are interested in detecting technology trends and changes in the semiconductor industry because there are some large resource allocation and organization restructure in this industry recently. In this study, we narrowed our research boundary into the U.S. and Japan. We extracted patent information related with semiconductor and made a citation network. We used Louvain method to cluster the maximum connected component and considered several largest clusters. Results show that the technology portfolios among the two countries are different. We used “*tf-idf*” to detect keywords and features of these unbalanced clusters. In the future, we will link the applicant information in a patent database with mergers and acquisitions (M&A) information in a company database. We will compare and integrate findings from different sources, such as investment, M&A, technology features and industrial policies in order to have a comprehensive understanding.

1. Introduction

Semiconductor industry is considered important base for the Artificial Intelligence Society. Recently, by preparing the era of Industry 4.0, resource allocation and organization restructure were under way in the semiconductor industry. There are several influential Mergers and Acquisitions (M&A) in the semiconductor industry. For example, Avago Technologies purchased Broadcom Inc. with 37 billion U.S. dollars and Qualcomm, Inc. announced to purchase NXP Semiconductor with 44 billion U.S. dollars. These phenomena changed technology landscape in the semiconductor industry largely.

In this study, we compared semiconductor industries in the U.S. and Japan. Both two countries have been leading countries in the semiconductor industry for a long term. By comparing these two countries, it is easy to trace and detect the technology trends in the semiconductor industry.

Since technological intangible assets are difficult to measure, patent analysis is a dominant method for research in this field. [Lee 2009] Likewise, in this study, we used patent data for analyzing technology development trends in the semiconductor industry.

This paper is arranged as follows: Section 2 is Literature Review, Section 3 is Data, Section 4 is Experiment, Section 5 is Results and Section 6 is Summary.

2. Previous Literature

In 2009, Lee *et al.* proposed a keyword-based patent map. This map is generated through keywords of patents. These keywords in vectors are extracted by text mining. Then keyword vectors are reorganized by Principle Component Analysis. Finally, these vectors were projected onto 2-dimension surface. From the map, it is easy to detect where to invest. [Lee 2009]

In 2011, Wang *et al.* used patent co-citation information between Fortune 500 companies. According to the co-citation networks of different periods, companies are divided into different industry groups. Furthermore, companies' positions in

the networks change among different periods. [Wang 2011]

Ma *et al.* proposed a comprehensive method for identifying technology-driven M&A targets. They used both qualitative and quantitative methods for analysis. They also invited policy makers and experts for evaluation. Finally, they used a company as an example to verify the effectiveness of their methods. This method provided a standard, sophisticated way for identifying M&A targets. [Ma 2017]

Shao *et al.* mainly focused on financial items of M&A in Japan. This paper deals with categorization of M&A in Japan. However, it did not take technology factors into consideration, even Japan is famous for its science, technology and trading. [Shao 2018]

3. Data

We used Derwent Innovation patent database and Derwent World Patent Index (DWPI) for extracting patent information. We made smart search topic being “semiconductor”, Publication date (Basic) being from 1990.01.01 to 2018.01.01, Application Country/Region (Basic) being US or JP. 874005 items were retrieved. We chose this time period because large development in semiconductor industries in Korea and Taiwan happened in the 1990's. [Chiu 2014] We used “Basic” items because “Basic” records the first patent in the same DWPI patent family, which largely represents where and when the patent questioned came from. [Derwent Innovation 2019]

We extracted the following items from the database: Publication Number, Title, Title – DWPI, Publication Date, Cited Refs – Patent, Count of Cited Refs – Patent, Citing Patents, Count of Citing Patents, DWPI Family Members, DWPI Count of Family Members. We did not extract International Patent Classification (IPC) and Cooperative Patent Classification (CPC) data and we used the citation clustering method for classification. [Thurber 1918] In addition, technology trends and contents change rapidly nowadays. Hence, data merely from IPC or CPC did not assure accuracy.

4. Experiment

4.1 Citation Network

Based on these patent data, we constructed a patent citation network. According to citation information of the patent database, we drew links between patents. Then we grouped them by DWPI Family Member information and built a new network. We assigned weights on edges between DWPI Family Members in the new network according to previous linkage information. The weight is the quantity of all links between any two patents which belong to the two DWPI Family pairs respectively.

4.2 Node Degree

We were interested in the network structure and we extracted the maximum connected component from the original citation network. We calculated the degree of each node in the maximum connected component and plotted the degree distribution in log-log scale.

4.3 Clustering

Louvain method [Blondel 2008] is an effective and high-speed method for dealing with large networks. We used Louvain method for clustering the maximum connected component because it is very large, with 612570 nodes.

4.4 Sum of “*tf-idf*”

In natural language processing, a very common method called “*tf-idf*” is widely used. The “*tf-idf*” aims at filtering out important terms of a document among a corpus. The “*tf-idf*” is calculated in the following way (1):

$$tf-idf_{t,d} = tf_{t,d} \cdot idf_t \quad (1)$$

where subscript t means a specific word (term) and d means a given document in the whole corpus. In order to calculate the importance of a word in the whole corpus, we sum up “*tf-idf*” of the same term across the whole corpus, as shown in (2). We have the “*tf-idf*” for the whole corpus.

$$tf-idf_t = \sum_d tf-idf_{t,d} \quad (2)$$

5. Results

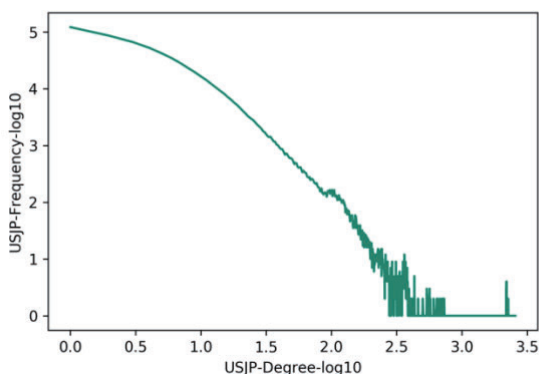


Figure 1. Patent degree distribution in log-log scale

Figure 1 shows the distribution of degrees in log-log scale. We took the logarithm to base 10. The horizontal axis is the degree in log scale whereas the vertical axis is the frequency in log scale. Different from the general understanding that companies only cite their own patents, patents related with semiconductor are cited by other companies as well. The line in Figure 1 is nearly straight in the middle part and we assumed that the maximum connected component has scale-free network features for nodes with degrees in the range of 20 to 80.

By Louvain method, we finally had 333 clusters. We present the largest 21 clusters and their contents in Table 1.¹

Table 1 shows the quantities of patents grouped by application countries. The “USnum” column shows the quantities of patents from the U.S. by each cluster whereas the “JPnum” column shows the quantities of patents from Japan by each cluster. The column “US/JP-ratio” shows the ratios of the values in “USnum” to those correspondences in “JPnum”. Similarly, column “JP/US-ratio” shows the ratios of the values in “JPnum” to those correspondences in “USnum”.

We selected two leading countries in the semiconductor industry. Intuitively, ratios across different clusters did not fluctuate too much. However, as shown in this table, ratios are quite different. We mark and underline the ratios above 2 or below 0.5 in red. These unbalanced “Cluster_id” are 4, 14, 7, 20, 8, 16 and 23.

Table 1 Quantities of Patents in Each Cluster Grouped by Publication Countries

No.	Cluster_id	USnum	JPnum	US/JP-ratio	JP/US-ratio
1	6	33728	36138	0.933	1.071
2	1	24488	33412	0.733	1.364
3	5	18502	24441	0.757	1.321
4	12	19731	19964	0.988	1.012
5	3	18922	19082	0.992	1.008
6	4	29081	7798	<u>3.729</u>	<u>0.268</u>
7	14	22394	11061	<u>2.025</u>	<u>0.494</u>
8	15	16448	15092	1.090	0.918
9	7	22213	9053	<u>2.454</u>	<u>0.408</u>
10	2	16267	13532	1.202	0.832
11	0	15819	10578	1.495	0.669
12	9	13368	10941	1.222	0.818
13	19	15128	8130	1.861	0.537
14	20	3409	18297	<u>0.186</u>	<u>5.367</u>
15	17	12038	7522	1.600	0.625
16	10	7371	7381	0.999	1.001
17	13	5743	8856	0.648	1.542
18	8	2246	7895	<u>0.284</u>	<u>3.515</u>
19	16	7592	2030	<u>3.740</u>	<u>0.267</u>
20	22	5001	4112	1.216	0.822
21	23	6163	2012	<u>3.063</u>	<u>0.326</u>

Furthermore, patents granted from the U.S. are becoming more and more recently whereas patents from Japan are declining. From these results, investors can pay attention to the differences.

¹ The quantity of patents of the 22nd largest cluster declines to 3836. We also conducted Louvain method several times and results are slightly different. Here is a typical example.

Table 2 Sum of “*tf-idf*” in “Cluster_id” Being 4

word	tfidf-sum
film	950.1
gate	928.9
second	894.0
layer	835.5
region	758.0
silicon	741.6
dielectric	727.7
insulating	724.3
fin	662.7
material	657.5
oxide	657.0
metal	650.5
electrode	613.3
structure	593.9
trench	593.2

Table 3 Sum of “*tf-idf*” in “Cluster_id” Being 16

word	tfidf-sum
layer	262.4
memory	247.3
magnetic	227.7
second	210.1
material	185.3
resistance	181.7
electrode	181.6
change	177.0
cell	166.7
film	164.2
element	158.9
metal	150.3
conductive	149.9
line	146.5
phase	145.7

Table 4 Sum of “*tf-idf*” in “Cluster_id” Being 23

word	tfidf-sum
layer	208.9
gate	198.8
film	194.9
memory	189.4
second	188.8
pattern	180.0
mask	170.3
region	165.7
silicon	158.6
insulating	147.2
material	143.4
trench	140.6
line	136.7
conductive	136.2
forming	124.4

Let us consider several clusters marked red in Table 1. We calculated sum of “*tf-idf*” for clusters with “Cluster_id” being 4, 16 and 23 as examples because they all hold “US/JP-ratio” over 3. The corpus of each cluster is abstracts of patents belonging to the cluster questioned. Herein, we pasted words with top 15

highest sum of “*tf-idf*” values (keywords) of each cluster in Table 2, 3 and 4.

From Table 2, we only detected general terms used for semiconductor manufacturing. From Table 3 and 4, we argue that technology in the 2 clusters is related with memory storage. In all the 3 clusters, U.S. shows dominant power. We hold a hypothesis that U.S. is now allocating resources in manufacturing new generation memory in order to prepare for the big data era.

6. Conclusion

In this study, we extracted patent information in the semiconductor industry. We narrowed our research boundary inside the U.S. and Japan, the two leading countries around the world. We detected that the technology portfolios are different in the two countries. We investigated the technology differences and used “*tf-idf*” to filter out important words among these clusters.

We investigated 3 clusters with high “US/JP-ratio” values as examples. From these examples, we found that U.S. companies have been interested in memory manufacturing recently.

In this study, we only focused on citation information, which is far from enough. There are still other columns in the database, such as Publication Date, Assignee/Applicant and Abstract-DWPI. These columns provide the understanding of patents in a detailed way.

In the future, we want to deal with these columns to catch the current trend in the semiconductor industry. We will link Assignee/Applicant information of the Derwent Innovation patent database with company names in a company database. We will also incorporate social phenomena, such as M&A, industrial policies and taxes in the semiconductor industry so as to have both private and public perspectives.

References

- [Lee 2009] Lee, Sungjoo, Byungun Yoon, and Yongtae Park, An approach to discovering new technology opportunities: Keyword-based patent map approach, *Technovation* 29 (6-7): 481-97, Elsevier.
- [Wang 2011] Wang, Xianwen, Xi Zhang, and Shenmeng Xu, Patent co-citation networks of fortune 500 companies, *Scientometrics* 88 (3): 761-70, Springer Nature, 2011.
- [Ma 2017] Ma, Tingting, Yi Zhang, Lu Huang, Lining Shang, Kangrui Wang, Huizhu Yu, and Donghua Zhu, Text mining to gain technical intelligence for acquired target selection: A case study for china's computer numerical control machine tools industry, *Technological Forecasting & Social Change* 116: 162-80, Elsevier, 2017
- [Shao 2018] Shao, Bohua, Kimitaka Asatani, and Ichiro Sakata, Categorization of mergers and acquisitions in japan using corporate databases: A fundamental research for prediction. Paper presented at 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM 2018), Bangkok, Thailand, IEEE, 2018
- [Chiu 2014] Chiu, Chien-Che, and Hsing-Ning Su, Analysis of patent portfolio and knowledge flow of the global semiconductor industry. 2014 Portland International Conference on Management of Engineering & Technology (Picmet): 3621-34, IEEE, 2014

- [Derwent Innovation 2019] Derwent Innovation, Getting started in derwent innovation. Available from <http://www.derwentinnovation.com/tip-innovation/support/help/introduction.htm>, Accessed: 2019.02.06.
- [Thurber 1918] Thurber, William L., The need of a secondary patent classification based on industrial arts, Journal of the Patent Office Society 1 (3) (November 01): 106-11, Patent Office Society (U.S.), 1918
- [Blondel 2008] Blondel, Vincent D., Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre, Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment 2008 (10): P10008, IOP Publishing, 2008

2016 年度決算を対象とした社是と企業業績の関係 (第一報) -東証二部上場製造業企業を対象とした解析- Relationships between mission statements and profitability in fiscal year 2016 (Preliminary Result) -Analysis for manufacturing companies listed on Tokyo Stock Exchange 2nd Section-

北島 良三^{*1}
Ryozo Kitajima

上村 龍太郎^{*2}
Ryotaro Kamimura

酒井 浩之^{*1}
Hiroyuki Sakai

中川 慧^{*3}
Kei Nakagawa

^{*1}成蹊大学
Seikei University

^{*2}東海大学 情報教育センター
IT Education Center, Tokai University

^{*3}野村アセットマネジメント株式会社
Nomura Asset Management Co., Ltd.

In this paper, we try to analyze relationships between mission statements and profitability. The mission statements are very important messages for companies, because they include founders' spirits, the business policies and so on. Therefore, we consider that mission statements affect profitability. Mission statements and profitability (The Return On Asset (ROA) was used) were gathered from annual securities reports because descriptions are accurate and reports are easy to obtain. As mission statements were written in natural language and data to be analyzed becomes complicated, a neural computational method called 'potential learning' which can interpret internal representations was used. As a result, we found that a generalization performance of the model was 0.6125 (accuracy) and mission statements composed of multiple messages may affect ROA.

1. はじめに

本研究は社是と企業パフォーマンスの関係について解析を試みたものである。社是は、経営上重要視する事項、創業の想い、経営方針、などが記された、企業が掲げる企業行動の最高目標と捉えることができるメッセージである(経営理念や企業理念など様々に称される[鳥羽 84])。企業は日々の業務活動により収益をあげているが、決算の際に1年分の総括が行われ、年間パフォーマンスが発表される。この際、パフォーマンスの高・低といった結果が出てくるが、この結果の根底に社是の作用が存在しているのではないかと我々は考えている。よって、社是と企業パフォーマンスの間に関係性は存在しているのか? 存在しているならばどのような関係なのか? といった事柄が我々の関心事項である。企業のパフォーマンスについて考えると、実際には、例えば市場状況や資本金などの各社を取り巻く様々な事柄が作用するため、社是が企業パフォーマンスに大きく影響を与えることまでは考えられない。しかし、企業の業務姿勢や指針が社是に表されているため企業がこれを無視して業務活動を行うことは考えられず、社是が業績に影響を与えている可能性を無視することもできないと考える。

社是と企業パフォーマンスの関係はこれまでも研究されているが、明確に企業パフォーマンスに影響を与えているという結果は得られておらず、高尾は先行研究レビューを通して「経営理念が企業の収益に寄与できる可能性は十分にあると考えられる」と述べるに留まっている[高尾 10]。また、小田らはテキストマイニングを用いた研究より、経営理念は「従業員に対する経営理念」、「社会貢献を訴える経営理念」、「世界視点を持つ経営理念」の3つに内容によって分類ができること、そしてこの3つの分類の中で従業員に対する経営理念を持つ企業が最も業績が高く、社会貢献を訴える経営理念を持つ企業の業績が最も低いことを報告している[小田 10]。北島らもこれまでテキストマイニングを用いて研究を実施しており、2013年10月時点で東京証券取引所市場第二部(東証二部)上場の製造業

企業を対象とした研究[北島 16a]から「世の中に対して訴求する社是を持つ企業」と「顧客に対して訴求する社是を持つ企業」が収益性がよいことを報告している。

しかし北島らの研究では社是を人手による企業 Web サイト巡回により収集しているため、データ収集に時間的コストを要している。社是は頻繁に変更されるものではないものの、必要に応じて変更されるものである。よって2013年時点の社是を解析し関係性を得られたからと言って、その関係性が未来永劫続くとは考えにくく、社是にはその時代に応じた傾向が存在していると考えられる。しかしこれを追跡し研究するためには長期間の調査が必要となるが、データ収集に時間的コストを要していることは継続した研究は困難である。

そこで本研究では社是収集源として有価証券報告書を用いることとした。有価証券報告書を使用する利点はその入手性と情報の正確性であり、これにより社是収集の時間的コスト圧縮が期待される。本研究では有価証券報告書を用いて2016年度決算時点の社是と収益性の関係について研究を行った。

2. 研究手法

2.1 研究の流れ

本研究は以下の流れで実施した。

1. 社是と企業パフォーマンスを収集する。
2. 収集された社是に対して形態素解析を実施し、解析用データを作成する。
3. 解析用データを入力とし、知的情報処理手法を用いて、企業パフォーマンスが高い企業と低い企業に分類する分類器を作成する。
4. 分類器の汎化能力を調べ、分類結果を解釈する。

以下、上記手順に沿って詳細を述べる。

2.2 社是と企業パフォーマンスの収集

本研究では社是と企業パフォーマンスデータの収集源として有価証券報告書(有報)を用いることとした。有価証券報告

連絡先: 北島 良三, 成蹊大学理工学部情報科学科,
〒180-8633 東京都武蔵野市吉祥寺北町 3-3-1,
r-kitajima@st.seikei.ac.jp

書は金融商品取引法で定められている金融庁への提出が義務づけられている書類である。データ収集源として有報を採用したのは、多くの企業情報を網羅できること（提出義務のある企業は複数定められているが、その一つに、金融商品取引所に上場されている有価証券の発行者が定められている）、正確な情報であること（虚偽記載は違法であり、また、重大な誤謬が発見された場合は訂正報告書の提出が必要となる）、入手が容易であること（金融庁の「金融商品取引法に基づく有価証券報告書等の開示書類に関する電子開示システム」（Electronic Disclosure for Investors' NETwork (EDINET) よりダウンロード可能である）、などが理由としてあげられる。

社はおよび企業パフォーマンスデータは有報以外からでも、例えば企業 Web サイトを閲覧することにより入手することも可能である。しかし、総務省統計局発行の「日本の統計 2018」によると、4,098,284 社の企業が日本には存在しており膨大な企業数である。また企業を証券取引所上場企業に絞ったとしても、東京証券取引所（日本最大の金融商品取引所である）には 2019 年 1 月 22 日現在で 3,653 社の企業が上場しており、こちらもかなりの数である。また、Web ページ上のどの部分に記載があるのかも各企業で異なっており、検索の必要がある。このため、1 企業ずつ Web サイトを巡回しデータを収集することは時間面で高コストである。さらに、企業 Web サイトは情報発信が未来方向に一方通行であり、例えばデータ収集日の前日に社是が更新されている場合に収集側がそれを把握することは困難であるなど、問題点が存在している。

有価証券報告書は金融庁の「金融商品取引法に基づく有価証券報告書等の開示書類に関する電子開示システム」（Electronic Disclosure for Investors' NETwork (EDINET)）により、PDF ファイル形式と XBRL ファイル形式の 2 形式で公開されている。PDF 形式の有報は報告書としてそのまま読めるものであるが、XBRL 形式の有報は eXtensible Business Reporting Language (XBRL) というコンピュータ言語で記述されているため、そのままの状態では可読性が低いものである。しかしコンピュータ言語で記述されているためコンピュータで処理しやすく、本研究では XBRL ファイル形式の有報から XBRL タグを手掛かりに、社はと企業パフォーマンスデータを得た。

有価証券報告書には「経営方針、経営環境及び対処すべき課題等」という項目があり、ここに記載されている文を解析対象候補として入手した。一方企業パフォーマンスであるが、これには様々な指標が存在している。本研究では収益性に着目し、Return On Asset (ROA) を解析対象とした。これの算出には総資産と経常利益を用い、また、その企業単体のパフォーマンスを解析対象としたい目的より、個別決算の値を取得した。

本研究では解析対象企業として東京証券取引所市場第二部に上場している製造業の企業を対象に解析を実施することとし、239 社の有報を収集し「経営方針、経営環境及び対処すべき課題等」を抽出した。そして抽出した文より目視にて社是に該当する文を選別したところ、197 社の社是を得ることができた。

2.3 解析データの作成

上記のように収集される社はであるが、これらはテキストデータであるためそのままでは解析を実施することができない。そこで社はに対して形態素解析を実施し、その結果を基に解析用データを作成する。形態素解析には日本語形態素解析システムである JUMAN を用い、単語の「カテゴリ」を得た。カテゴリは JUMAN の辞書に用意されている情報である。これは単語がどのようなものであるのか詳細を表したもので、例えば「飛行機」という単語は「カテゴリ：人工物-乗り物」と

表 1: 解析用データの変数一覧

番号	カテゴリ名
1	人
2	組織・団体
3	人工物-その他
4	自然物
5	場所-施設
6	場所-その他
7	抽象物
8	数量
9	時間
10	未定義語

いう具合である。

JUMAN には 22 種類のカテゴリが用意されているが、本研究では「未定義語」というカテゴリを新設し、合計 23 種類のカテゴリを社はから得た。未定義語のカテゴリは辞書に未登録の単語を分類するカテゴリである。このカテゴリには未登録単語が蓄積されるため、必要に応じてこのカテゴリの単語を辞書に追加することも可能である。

このように得られるカテゴリであるが、これらは各カテゴリの各社はでの出現頻度 (Term Frequency, TF) を、そのカテゴリの全社はでの出現状態 (文書頻度逆数, (Inverse Document Frequency, IDF)) により重み付けした値 (TF-IDF) として記録し解析用データとした。一方、ROA データであるがこれは ROA をそのまま解析対象としたのではなく、197 社の ROA うち第一四分位数 (2.2813) 以下の企業にターゲットフラグとして「0」を、第三四分位数 (5.8129) 以上である企業に「1」を、それぞれ割り当てて被説明変数とした。各フラグの内訳は、1 が 50 社、0 が 50 社となり合計 100 社が解析対象となった。

以上の手順により入力変数が 23 から成る解析用データが作成されたが、データ中のほとんどの値が 0 である変数が散見されたため、TF-IDF 値が 0 であるものが 100 社中 90% 以上であるものを除去し、最終的に解析用データは 10 変数から成るデータとなった（最終的な解析用データの変数を表 1 に示す）。この作成された解析用データが次節で述べる手法により解析される。

2.4 解析手法

本研究で解析するデータは、社はと企業パフォーマンスから成るデータであり、複雑なデータである。そこで、複雑なデータの解析に定評のあるニューラルネットワークを解析に用いることとした。しかしニューラルネットワークはブラックボックスと称されるほどモデル解釈が困難である。そのためモデル解釈が可能なニューラルネットワークである「潜在学習」[Kamimura 17] を解析手法に採用した。潜在学習はこれまで北島らによって解釈を要する解析に採用され（例えば、スーパーマーケットデータの解析 [北島 16b]）、高いモデル性能、高い解釈性が確認されている。

この手法は図 1 に示されているように自己組織化マップ (Self-Organizing Maps, SOM) と多層パーセプトロン (MultiLayer Perceptron, MLP) が基となっている二段階の学習から構成されている。図 1 中の ① は知識獲得段階と呼ばれる段階で、入力ニューロンの潜在性を算出し、また、SOM に基づいて知識の獲得 (学習) を行う。潜在性とは、ニューロンの多様な状況に対応できる能力と定義されるもので、「潜在性の高いニュー

ロン (多様な状況に対応できるニューロン) は学習で重要な役割を果たすニューロンである」と捉えている。潜在学習ではこの潜在性を学習完了後に解釈することで、どの入力学習に活用されたのか解釈可能である。潜在性算出にあたり入力ニューロンに番号 k ($k = 1, 2, \dots, K$) を割り当てて表すと、 k 番目の入力ニューロンの潜在性 ϕ_k^r は、

$$\phi_k^r = \left(\frac{V_k}{\max_{k=1, \dots, K} V_k} \right)^r,$$

により定義される ($0 \leq \phi_k^r \leq 1$)。ここで V_k は k 番目の入力ニューロンの分散 (これは j 番目の出力ニューロン ($j = 1, 2, \dots, J$) から k 番目の入力ニューロンへ接続されている重み $w_{j,k}$ から求められる [北島 16b]), r は算出される潜在性を制御するパラメータである。よってパラメータ r が大きい値を持つ場合は大きな分散を持つ入力ニューロンが大きな潜在性を持つようになる。潜在学習はこの潜在性を求めた後、知識獲得段階で学習を行う。知識獲得段階では SOM により学習が行われるが、通常の SOM と異なり、入力ニューロン (k 番目の入力ニューロンを x_k で示す) と j 番目の出力ニューロンの距離 d_j を求める際に潜在性が、

$$d_j = \sqrt{\sum_{k=1}^K \phi_k^r (x_k - w_{j,k})^2},$$

として利用される。すなわち潜在学習の知識獲得段階では入力ニューロンの潜在性で重みを付けられた距離を学習に使用する。その他の学習ロジックは通常の自己組織化マップと同様である。

知識獲得段階での学習が終了すると、続いて予測段階 (図 1 中の ② の処理) での学習が行われる。予測段階は MLP にて、そして、入力層-中間層間の重みの初期値に知識獲得段階で得られた、重みと潜在性より算出した値 (重み \times 潜在性) がセットされ学習が行われる。通常 MLP による学習結果は初期重みに左右されるが、潜在学習ではこの初期重み設定により、獲得された知識に基づいた学習が期待できる。以上が潜在学習の概要である。なお、潜在学習にはいくつかのバリエーションがあり、ネットワーク構造や潜在性の算出方法などに違いがある。

次章ではこの潜在学習を用いて、前節で作成したデータを解析していくが、本研究で使った潜在学習は入力ニューロンの数を 10 個、知識獲得段階における出力ニューロンの数および予測段階における中間ニューロンの数を 48 個、予測段階における出力ニューロンの数を 2 個とした。また予測段階において、中間ニューロンの伝達関数は双曲線正接関数、出力ニューロンの伝達関数はソフトマックス関数とした。中間ニューロンの数であるが、これは解析に使用したソフトウェアの標準個数^{*1}である。また、パラメータ r については 1 から 2 まで 0.1 刻みで変化させ、最も正解率の良いパラメータ値を探索した。さらに、解析時にデータは全サンプル (100 件) の、70% を学習用途、15% を過学習防止用途 (早期停止)、15% を汎化能力試験用途、に用いた。くわえてこの分割比率はそのまま、各用途に使用されるサンプルがランダムに異なるものを 10 パターン作成し、これらを用いて分類器を作成した。これはすなわち 10 個のモデルを作成したことになるが、次章で述べる分類結果などモデルパフォーマンスはこの 10 個の平均値を採用していることをここで述べておく。

*1 本研究で解析に使用したソフトウェア (MATLAB の SOM Toolbox) の設定 (mapsize:normal) による。

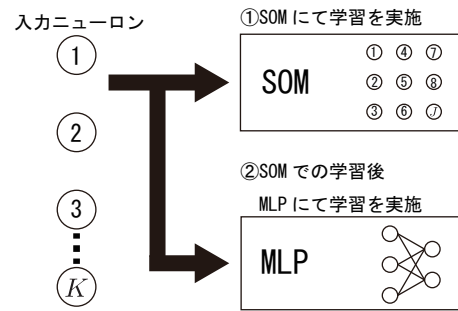


図 1: 潜在学習概要

3. 結果と考察

解析の結果、社是は正解率: 0.6125, 精度: 0.6058, 再現率: 0.6500 で分類ができた。この際パラメータ r の値は 1.9 であった。潜在学習により得られたこれら結果であるが、比較のために実施した、多層パーセプトロン、潜在性を使用しない潜在学習 (ネットワーク構成は潜在学習と同じであるが、学習時に潜在性を加味しないもの)、と比較すると高い結果であることを確認した (表 2 に結果を示す)。このことより、社是と企業パフォーマンスの関係調査には潜在学習が有効であることを確認した。しかし著しく高いパフォーマンスと言える結果でもないため、この結果は社是とパフォーマンスの関係は解析が難しい問題であることを示していると言える。モデルパフォーマンスが上がらなかった原因の一つとして学習データが少ないことが考えられる。解析対象データのうち、社是まで取得できたデータは 197 社あるため、無駄なく使用できればあと 97 社のデータを解析に使用することができる。もちろん 97 社のデータの内訳までも考慮しなければならないため、単純に 97 社分データを増加させられるわけではないが、閾値の調整により少々の増加は期待できる。

さて分類器が作成でき、性能も明らかになったため、続いては分類に重要な役割を果たした変数の解釈を実施する。一般的にニューラルネットワークはブラックボックスと称され、解析時に活用された変数を解釈することは困難である。しかし潜在学習は潜在性を確認することでこれを可能としている。図 2 は潜在性を示したものである。この図より、第 10 変数の潜在性が最大であることが確認できる。これは未定義語のカテゴリを意味する変数である。続いて潜在性がネットワークで活用された変数であるのかを重みの視点から確認する。図 3 にネットワークの入力層-中間層間の重みを示しているが、図中の四角形は重みを表しており、図形の大きさと重みの大きさを、緑色が正の値を、赤色が負の値を示している。比較のために実施した多層パーセプトロンでは、重みは各ニューロンでバラバラの重みを持っており、重要変数の判断は困難である。潜在性を使用しない潜在学習では多層パーセプトロンの場合よりも規則性が見えるものの、どの入力変数が重要であるかの判断は依然困難である。一方潜在学習の重みは 10 番目の入力ニューロンに重みが集中していることから潜在学習が 10 番目の変数を積極的に活用している状態が確認できた。

それでは未定義語のカテゴリのうちどのような単語が出現しているのだろうか。続いてはそれを調査した。その結果、形態素解析の誤りが大多数を占めていること、これを除くと数値の出現が多いことがわかった。ただしこれはアラビア数字ではなく、例えば ① などと通常とは異なる表記の数字であった。なお社是はビジネスメッセージであるが、未定義語のカテ

表 2: 解析結果

手法	正解率	精度	再現率
多層パーセプトロン	0.5250	0.5411	0.5125
潜在性を用いない潜在学習	0.5438	0.5532	0.5875
潜在学習 ($r = 10$)	0.6125	0.6058	0.6500

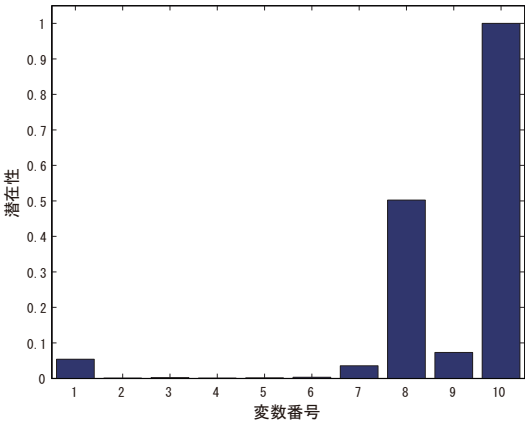


図 2: 入力変数の潜在性 ($r = 10$)

ゴリにはビジネス関連用語はあまり含まれていなかった。最重要変数であるものの、未定義語のカテゴリから特殊表記の数字が得られたことは意外であったため、2 番目に潜在性の高い第 8 変数を調べたが、この変数自体が数量であることから (漢数字の出現が多いことがわかった)、本研究で解析対象とした社是では数字が重要な要因であると判断した。

この結果より解釈を実施したが、社是中でのアラビア数字でない数字は社是を複数発信している場合に用いられるものであった。よって一つの社是で複数訴求メッセージがあることを示している。未定義語に分類された文字を使用して複数メッセージを発信している企業は 9 社あり、うちターゲットフラグ 1 に属する企業数は 4 社、フラグ 0 に属する企業数は 5 社であり、フラグ 0 に属している企業数の方が多いという結果であり、未定義語に限って解釈を実施すると、複数メッセージ発信企業の方が 1 社だけではあるがフラグ 0 に属している数が多いことがわかった。

本研究により未定義語に含まれる単語により社是がパフォーマンス別に分類可能であることが明らかになった。今後はビジネス関連用語を辞書に登録して未定義語のカテゴリを廃し、また、形態素解析の誤りにも対応し、より詳細な解析を実施していきたい。

4. まとめ

本研究は社是と企業パフォーマンスの関係について調査したものである。企業パフォーマンスには ROA を採用し、これの算出に必要な数値と、そして解析対象である社是は有価証券報告書より入手した。有価証券報告書をデータ源とすることで、信頼性の高いデータを時間的コストをかけずに入手することができた。また ROA は個別決算の数値を用い、これによりグループ企業の影響を排した関係性調査を実施できた。社是と ROA 間の関係は複雑であるため、内部表現解釈が可能なニューラルネットワークである潜在学習を解析手法に採用し

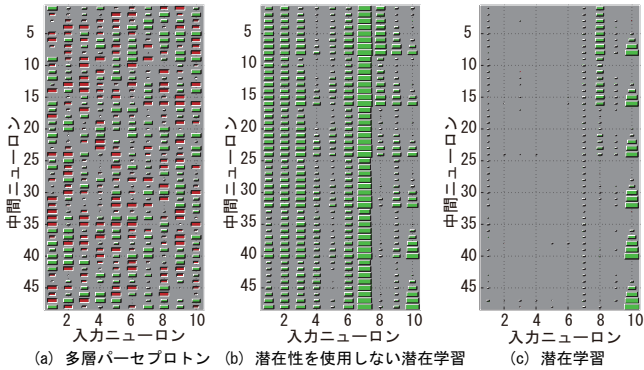


図 3: 入力層-中間層間の重み

た。解析の結果、正解率 0.6125 でパフォーマンス別に企業を分類できるモデルを作成でき、またこの分類には ① といった特殊な表記の数値が影響を及ぼしていることがわかった。比較のため潜在学習の他、多層パーセプトロン、潜在性を使用しない潜在学習、を用いて解析を実施したところ、潜在学習の分類性能が最も良い結果であり、また内部表現は潜在学習以外は明確に解釈できず、潜在学習が社是と ROA の解析に有効な手法であることが確認できた。しかし分類精度は改善の余地があり、今後は精度向上、またビジネス用語を形態素解析の辞書に整備するなどし、さらなる重要変数の解釈を実施していく。

参考文献

[鳥羽 84] 鳥羽欽一郎, 浅野俊光: 戦後日本の経営理念とその変化-経営理念調査を手がかりとして- (1984), 組織科学, Vol. 18, No. 2, pp. 37-51.

[小田 10] 小田恵美子, 三橋平: 経営理念と企業パフォーマンス: テキスト・マイニングを用いた実証研究 (2010), 経営哲学, Vol. 7, No. 2, pp. 22-37.

[高尾 10] 高尾義明: 経営理念は役に立つのか: 経営理念と成果についての先行研究からの考察 (2010), 経営哲学, Vol. 7, No. 2, pp. 38-51.

[北島 16a] 北島良三, 上村龍太郎, 内田理: 社是に含まれる業績に影響を与える単語に関する研究-TF-IDF と共起語を切り口にした解析- (2016), 第 32 回ファジィ システム シンポジウム講演論文集, pp. 463-468.

[北島 16b] 北島良三, 遠藤啓太, 上村龍太郎: 入力ニューロンの潜在性に着目した小売店店舗の非継続来店顧客検知モデルの作成 (2016), オペレーションズ・リサーチ, Vol. 61, No. 2, pp. 88-96.

[Kamimura 17] Kamimura Ryotaro: Collective mutual information maximization to unify passive and positive approaches for improving interpretation and generalization (2017), Neural Networks, Vol. 90, pp. 56-71.

フォールトラインが組織の成果に及ぼす影響 Influences caused by faultlines on the organizational performance

熊田 ふみ子^{*1}
Fumiko Kumada

倉橋 節也^{*1}
Setsuya Kurahashi

^{*1}筑波大学大学院ビジネス科学研究科
University of Tsukuba Graduate School of Business Sciences #1

The diversification of employment and work styles in organizations is inevitable to ensure a stable workforce in Japan, where a labor force is shrinking due to a declining birthrate and an aging population. Using the concept of faultlines, which are hypothetical dividing lines that may split a group into subgroups of people based on their multiple attributes, this paper examines the relationship of influences of a structure of diversity: the faultline strength and the number of subgroups: and communication within an organization. It is verified by an agent-based model based on the results a survey of Japanese organizations. In conclusion, when staff in an organization interact smoothly the structure of diversity influences less on the organizational performance. However, when the interaction is not smooth, for example, the communication between similar staff, the structure of diversity influences more the performance. Therefore, it is important to manage the interaction and communication.

1. 目的と背景

少子高齢化が進む日本では労働力の確保は必須である。そのため、外国人労働者の受入れや「働き方改革」が進められている。また、AI や IoT 等のテクノロジーの進化で、働く人、そして働き方が多様化している。

ダイバーシティ・マネジメントの研究分野では、多様化は組織の成果にプラスにもマイナスにも影響する諸刃の剣と言われている。そのため、多様化した組織を適切にマネジメントすることは、今後の日本の企業では重要な課題である。

本研究は、多様性を定量的に捉える「フォールトライン」の考え方に着目し、「多様化した組織をどのようにマネジメントすれば、組織のパフォーマンスは向上するか」の解の1つを明らかにすることを目的とする。

2. 先行研究

2.1 多様性研究

[Williams 1998]は、多様性が組織のパフォーマンスにどのように影響を与えるかについて統合モデルを提唱し、プラスにもマイナスにも影響すると説明している。

- マイナスの理論

Social categorization 理論:自分と他者を年齢等表層的特徴により分類することでコンフリクトを起こす。

Similarity-attraction 理論:類似性の高い人ほど魅力を感じ結束力を高めるが、類似性の低い人とはコンフリクトを起こす。

- プラスの理論

Information and decision-making 理論:多様性により知識や情報の種類や量が増えるため、組織にポジティブな影響をもたらす。

この統合モデルでは、多様性が組織の成果にプラスになるか、マイナスになるかは、コミュニケーションが阻害されるか促進されるかがポイントの1つであることを示している。

2.2 フォールトライン理論

[Lau 1998]は、フォールトライン(グループを1つ以上の属性によってサブグループ(以下、SG)に分ける仮定の分割線。以下、FL)という考えを提唱し、組織構成員の属性による多様性と組織内の分断(コンフリクト)の発生の因果関係を説明した。その後の FL の先行研究の多くは、「FL がコンフリクトを増大させる」という結果を報告している。

また、FL の計測方法については、先行研究で 10 種類以上が提案されている。その中で[Suzuki 2015]は、[Meyer 2013]が提唱したクラスター分析の評価尺度である Average of Silhouette Width(以下、ASW)が連続変数、カテゴリカル変数ともに扱え、適切な SG に分けることができる等の長所を示している。

2.3 先行研究の課題

日本の組織を対象にした FL 理論の先行研究は限られている。そこで本研究では、日本の組織を対象に調査を実施し、その結果をもとに、多様性を ASW によって FL の強さと SG 数で定量化・構造化した上で、多様性とコミュニケーションの関係をエージェントベースモデル(以下、ABM)で検証する。

3. モデルの概要

本モデルは、エージェントを組織の構成員に見立て、エージェント同士が交流することで影響し合い、評価値を更新する。そして、組織全体の評価値(各エージェントの評価値の和)を交流の前後で比較して、その増減を検証する。

3.1 エージェントの属性

各エージェントに 6 個の遺伝子列(0,1 のビット列)を設定する。その遺伝子列を意思決定態度属性(以下、態度属性)とみなす。エージェント同士が交流して、各々の態度属性に影響を与えて評価値を更新する。態度属性は、消費者行動研究における多属性態度モデルの理論を応用している。多属性態度モデルは、消費者が製品を評価する際、一つの属性だけではなく複数の属性に着目し、各々の属性への評価の総和が製品の総合評価になるという考え方である。この考えを応用して、製品を組織の

課題に置き換え、課題への取組み方の特性を複数の属性で表し、属性の総和をその課題解決のための総合評価に見立てた。

また、エージェントに設定した 6 個の遺伝子列の初期配列を ASW で計算して、組織の FL の強さと、各エージェントが属する SG を決定する。初期の態度属性は外部からの影響を受ける前なので、年齢や性別等の表層的属性に依拠すると仮定する。

次に、初期の態度属性によって設定された SG は、FL 理論の基になっている Social categorization 理論の表層的属性による似た者同士による内集団と言い換えることができる。

また、交流することで影響しあい態度属性は変化するが、所属 SG は不変である。その上で、誰と交流して影響し合うと組織全体の態度属性の評価値が向上するかを検証する。

3.2 効用関数

効用関数として、NK モデルを用いる。NK モデルは、生物が進化する過程を示す遺伝的アルゴリズムで、[Siggelkow 2003] や [Kobayashi 2011] 等、組織行動や組織学習の分野で活用されていることより、本モデルで用いた。

NK モデルの評価値である適応度 (fitness) は、0 か 1 の値を取る N 個の遺伝子が K 個の遺伝子と関係して全体の適応度を決定している。本研究ではエビスタシスが低い N=6, K=1 でシミュレーションを実施した。

3.3 シミュレーションの設定

交流する相手と自分との類似度の程度により、以下3種類の交流方法を設定した。1 回のシミュレーションで 100 回の交流を行い、設定ごとに 100 回のシミュレーションを実施した。

- 内集団交流

FL 理論を再現した内集団交流は、同じ SG のエージェントの交流 (以下、SG 内交流) と、違う SG のエージェントの交流 (以下、SG 間交流) の割合を 0~100% で設定した。SG 内交流とは、似た者同士 (内集団内) の交流であり、その割合の大きさがコンフリクトの大きさを表す。

交流相手を決める第 1 ステップは、SG 内交流か SG 間交流かを割合に応じてランダムに決める。第 2 ステップは、遺伝子列のコサイン類似度が一番高いエージェントを選択する。実社会における似た者同士の交流を再現したもので、多様性がマイナスに作用している状態である

- 外集団交流

情報意思決定理論に基づいた交流。遺伝子列のコサイン類似度が低いエージェントを選択する。その相手は、SG の同異とは関係なく、トーナメントサイズに応じてランダムに選択する。トーナメントサイズが大きくなると選択圧が高くなるため、類似度の低い相手を選択する可能性が高くなる。交流相手を決める第 1 ステップは、トーナメントサイズに応じてランダムに選択し、第 2 ステップで、その中で一番コサイン類似度の低いエージェントを選択する。

実社会では、自分とは違う属性の人と交流して知識や情報の種類や量を増やし、多様性を積極的に活かしている状態を再現した。本シミュレーションでは、選択圧が一番小さいサイズ 2、一番大きいサイズ 17、中間のサイズ 9 の 3 種でシミュレーションを行った。

- ランダム交流

SG の同異に関わらず 18 のエージェントがランダムに相手を選んで交流する方法。自由に誰とで平等に交流するため、コンフリクトがなくスムーズに交流できている状態で、組織の潜在的な能力と仮定した。シミュレーションの設定を表 1 に記した。

表 1. シミュレーションの設定

エージェント数	18
評価関数	N
	K
交流相手の選択	内集団/外集団/ランダム
交流方法	一点交叉
適応度の設定	(0,1)の一様乱数
1シミュレーションあたりの交流回数	100回
シミュレーションの回数	100回

3.4 評価基準

100 回の交流の前後における組織全体の適応度の増減を観察し、増加幅が交流の成果である。そのため、100 回のシミュレーションによる増加幅の最大値の Top10 の平均 (外れ値を除く) を評価基準とする。これは、組織の最大可能性を意味する。

また、どのような課題に対しても安定した結果を出すことも組織としては重要である。そのために、100 回のシミュレーション結果の標準偏差を第 2 指標とする。標準偏差が大きい場合、最大値が発生する安定性は低くなる。そのため標準偏差を組織の安定性とした。

そして、最大可能性と安定性が、FL の強さや SG 数など多様性の構造や、交流方法により、どのように変化するか、その関係を観察することで、組織の多様化と成果の関係を検証する

3.5 モデルの妥当性検証

モデルの妥当性を検証するために、FL の強さ (強・弱) と、SG 数 (2, 3, 6) による 6 セットデータを作成し、FL 理論を再現している内集団交流でシミュレーションを実施した。その際、コンフリクトを表現している SG 内交流の比率は 0~100% の 10% 単位でパラメーターを設定した。

表 2 の相関係数より、SG 内交流の比率が増えると Top10 の平均が下がるため、「FL はコンフリクトを起こし組織にマイナスの影響を与える」という FL 理論の一部をサポートしている。また、表 3 では、SG 数が同じ場合には、FL が強いと回帰係数が大きくなっている。この点においても「FL はコンフリクトを起こし組織にマイナスの影響を与える」という FL 理論の一部をサポートしている。

表 2. 評価基準と各パラメーターの相関係数

	FL	SG数	比率
Top10の平均	0.671 ***	-0.319 **	-0.260 *
標準偏差	0.675 ***	-0.344 **	-0.255 *

* p<.05, ** p<.01, *** p<.001

表 3. 各データセットの評価基準による回帰係数

FLの強さ	SG数	Top10の平均	標準偏差
強(1.000)	2	-3.495 *	-0.987 *
弱(0.142)	2	-0.545 .	-0.200 .
強(0.723)	3	-1.079	-0.404
弱(0.300)	3	-0.553	-0.258
強(0.848)	6	-0.874	-0.399 *
弱(0.260)	6	-0.578	-0.215

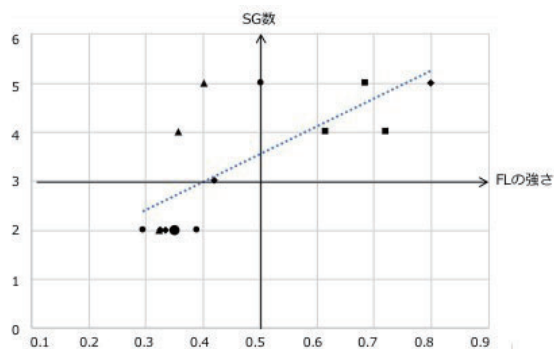
. p<0.1, * p<0.05, ** p<.001, *** p<.001

4. 実態調査

4.1 調査概要

対象は日本の企業5社14グループ(大企業1社3グループ、中堅企業1社1グループ、ベンチャー企業3社10グループ)で126名が回答)で、従業員の属性とグループ内の交流状況を調査した。調査項目は、年齢、性別、所属年数、雇用形態の4項目の属性データと、各所属スタッフとの業務上のコミュニケーションの頻度(5段階)である。

4.2 調査結果



※▲:大企業、◆:ベンチャー1、■:ベンチャー2、●:ベンチャー3、●:中堅企業

図1. 調査対象の組織のFLの強さとSG数

調査結果から属性データをASWで計算し、各組織のFLの強さとSG数を算出した。(図1を参照)。

その結果を用いて、FLの強さとSG数、SG内交流の割合について回帰分析を行った結果、以下の関係式を導いた。

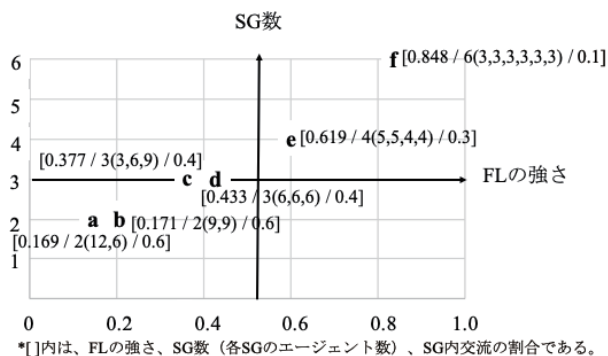
$$(SG数) = 0.6872 + 5.7206 \times (FLの強さ) \quad \text{式(1)}$$

※決定係数:0.522, p値<0.05

$$(SG内交流) = 0.7845 + (-0.11443) \times (SG数) \quad \text{式(2)}$$

※決定係数:0.847, p値<0.001

次に関係式をもとに、6つのセットデータ(a~f)を作り、シミュレーションを実施した。セットデータの多様性の構造(FLの強さ/SG数)は図2である。実態調査の結果からSGの人数に偏りがあるグループと、ないグループが存在したため、ほぼ同じFLの強さで同じSG数だが人数に偏りがある場合とない場合(a:12,6とb:9,9、c:9,6,3とd:6,6,6)のセットデータを作成した。また、類似交流におけるSG内交流の比率も調査結果をもとに算出した。各セットデータの比率は、図2の中に記している。

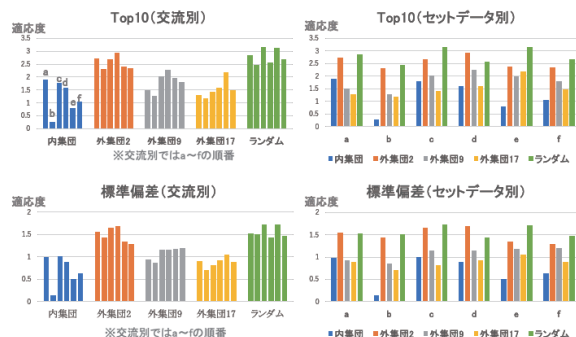


*[]内は、FLの強さ、SG数(各SGのエージェント数)、SG内交流の割合である。

図2. 調査結果から作成したセットデータ

5. シミュレーションの結果

3つの交流方法について、100回のシミュレーションを実施した結果が図3である。1回のシミュレーションで100回の交流し、その100回の交流の前後における組織全体の適応度の増減のTop10の平均と標準偏差である。



※外集団交流の2,9,17はトーナメントサイズの大きさである。

図3. シミュレーションの結果(上段:Top10の平均

下段:標準偏差 / 左:交流別 右:セットデータ別)

- 全体的に、Top10の平均(最大可能性)と、標準偏差(安定性)は同じ傾向を示している。そのために最大可能性が高い場合は安定性が低くなる。
- コンフリクトが起きて内集団交流は、組織の潜在的パフォーマンスと仮定するランダム交流より最大可能性が低い。特に、FLが弱く、SG数が少なく、SG内の人数が均衡の場合(bの場合)は、顕著である。これは、「規模が均一よりも偏りがある方が業績は高くなる」という[Polzer 2006]の結果を支持している。また、SG数が2よりも3以上の方が高くなっている。この点も「SG数が2よりも3以上の業績が高い」ことを報告している[Carton 2012]の結果を支持している。
- モデルの妥当性検証のために行ったシミュレーションの結果(表2)と同様に、FLとSG数は逆の効果をもたらすために、FLは強いがSG数が多くなるeやfは、FLとSG数が中庸であるc,dよりも最大可能性が低くなる。
- 組織の潜在的パフォーマンスを示しているランダム交流は、dを除いて、その組織の最大可能性を一番高く出せる交流である。
- 外集団交流は、選択圧が高くなると最大可能性が低くなる傾向である。異質な人との交流はパフォーマンスを上げるという仮説は、異質の程度が高すぎると逆効果になることがわかった。ただし、異質な人との交流をすると、組織の多様性の構造の影響を受けにくいこともわかった。

6. 考察

本研究では、FL理論により多様性をFLの強さとSG数で定量化して、3種類の交流方法(内集団、外集団、ランダム)でABMを使ってシミュレーションを行い、組織の多様性とコミュニケーションの関係、特に「誰とコミュニケーションすると、組織のパフォーマンスが高くなるか」に焦点を絞り検証した。以下3点について明らかにすることができた。

- 似た者同士の内集団交流では、FLの強さやSG数、人数の偏りという多様性の構造が組織の成果に影響を与えることがわかった。FLの強さは最大可能性を高める点ではプラスに働くが、安定性が低くなる点ではマイナスの影響

を及ぼす。SG 数は逆の作用を及ぼす。そのために、その組織が何を優先に目標を定めるのかにより、多様性の構造を考慮することが重要である。

- 第二に多様性を積極的に活かした外集団交流は、自分と極端に異質な人との交流は組織の最大可能性にプラスには影響しない。最大可能性を高くするためには、隔たりなく均等に誰とでも交流することである。
- 多様性を活かすために重要な事として、①組織の多様性の構造を把握する事。その時に、FL 理論は多様性を定量的に把握できるように活用が可能である。②組織の目標として、最大可能性(高いパフォーマンス)と安定性のどちらを優先させるのかを決め、それにマッチしたコミュニケーションをマネジメントする必要がある。例えば、ミーティングのファシリテーションや席順など。

次に、今後の課題について考える。

まず実態調査については、規模を拡大することや、業種や職種、規模による違いなど、日本の多様性の実態より深く調査することである。そして定量的にコミュニケーションが測れるデジタルデバイスの活用などが考えられる。次に、シミュレーションの設定については、NK モデルの K を1だけではなく K の値を高くしたシミュレーション、多様性の強度を細かく検証するために、トーナメントサイズをより細かく設定するなどシミュレーションの設定を改善することが必要である。

7. 結論

本研究の目的は、「多様化した組織をどのようにマネジメントすれば、組織のパフォーマンスは向上するか」であった。本研究で得られた 1 つの解は、組織の多様性の構造を把握した上で、コンフリクトの克服し、フラットなコミュニケーションの仕組みを作ることである。それが組織のパフォーマンスを向上させるために必要なマネジメントの 1 つである。

そして、本研究の学術的貢献として、FL の先行研究の結果を ABM で再現し、そのメカニズムの一部をコミュニケーションの視点から明らかにできたことが挙げられる。また、実在する企業を対象に FL 理論に基づき調査し、実際の現象を FL 理論、ABM の結果で説明できたことは実務的貢献である。

参考文献

- [Carton 2012] Carton, A. M. and Cummings, J. N.: A theory of subgroups in work teams, *Academy of Management Review*, 37(3), 441–470, 2012.
- [Kobayashi 2011] Tomohiro KOBAYASHI, Satoshi TAKAHASHI, Masaaki KUNIGAMI, Atsushi YOSHIKAWA, and Takao TERANO: Harnessing Organizational Deviation and Kaizen Activities through Agent-Based Modeling, *Institute of Electronics, Information and Communication Engineers. Information and Systems Society* 94(11), 2011.
- [Lau 1998] Lau, D. C. and Murnighan, J. K.: Demographic diversity and faultlines: The Compositional dynamics of organizational groups, *Academy of Management Review* 23(2), 325–340, 1998.
- [Meyer 2013] Meyer, B. and Glenz, A.: Team faultline measures: A computational comparison and a new approach to multiple subgroups, *Organizational Research Methods* 16(3), 393–424, 2013.
- [Polzer 2006] Polzer, J. T., Crisp, C. B., Jarvenpaa, S. L. and Kim, J. W.: Extending the faultline model to geo-graphically

dispersed teams: How collocated subgroups can impair group functioning, *Academy of Management Journal* 49(4), 679–692, 2006.

- [Siggelkow 2003] Nicolaj Siggelkow and Daniel A. Levinthal: Temporarily Divide to Conquer: Centralized, Decentralized, and Reintegrated Organizational Approaches to Exploration and Adaptation, *Organization Science*; 2003; 14(6).

- [Suuzuki 2015] Suzuki, R., Matsumoto, Y., and Kitai, A.: The Concept and Analytical Techniques of Faultline, *Journal of political economy* 211(6), 53–88, 2015

- [Willian 1998] Williams, K. Y. and O'Reilly, C. A.: Demography and diversity in organizations: A review of 40 years of research, *Research in organizational behavior* (8), 70–14, 1998.

[3A4-J-13] AI application: electrical power

Chair: Takashi Onoda Reviewer: Yuiko Tsunomori

Thu. Jun 6, 2019 3:50 PM - 4:50 PM Room A (2F Main hall A)

[3A4-J-13-01] Evaluation of power consumption estimation model based on household information○Tomofumi Tahara¹, Hideaki Uchida¹, Hideki Fujii¹, Shinobu Yoshimura¹ (1. The university of Tokyo)

3:50 PM - 4:10 PM

[3A4-J-13-02] Optimization of Power Electric Supply Path in Smart Grids○Takaya Ozawa¹, Ei-Ichi Osawa¹ (1. Future University Hakodate)

4:10 PM - 4:30 PM

[3A4-J-13-03] Design and Preliminary Evaluations of Multi-Agent Simulation Model for Electric Power Sharing among Households○Yasutaka Nishimura¹, Taichi Shimura², Kiyoshi Izumi³, Kiyohito Yoshihara¹ (1. KDDI Research Inc., 2. Kozo Keikaku Engineering Inc., 3. The University of Tokyo)

4:30 PM - 4:50 PM

世帯情報に基づいた電力需要量推定モデルの評価

Evaluation of power consumption estimation model based on household information

田原智史 *¹
Tomofumi Tahara内田英明 *¹
Hideaki Uchida藤井秀樹 *¹
Hideki Fujii吉村忍 *¹
Shinobu Yoshimura*¹ 東京大学大学院工学系研究科
School of Engineering, The University of Tokyo

Abstract: An electric power demand estimation model based on household information is proposed in this paper. The effect of transition of electric power systems being discussed today will emerge as the change of household electricity demand. The objective of this research is to develop a model which can output area-wise electricity demand curve from the sum of household-wise demand. The proposed method is based on a multi-agent-based household transition model and a household energy consumption database. The simulated result shows rough agreement with the measured values of the monthly electricity demand.

1. はじめに

1.1 背景

日本の電力システムは大きな転換を迎えている東日本大震災を契機に、エネルギー供給が制約され、集中型エネルギーシステムの脆弱性があらわになった。また、化石燃料の枯渇に対する懸念や環境への配慮から、大規模集中型エネルギーシステムから分散型エネルギーシステム [1] への転換が求められるようになった。分散型エネルギーシステムに向けて、再生可能エネルギー、蓄電池、熱源機等の機材を地域の特色に合わせて活用し、効率的なエネルギーの運用しようとする試みが見られ、現在様々な研究が行われている。

またわが国では、再生可能エネルギーの普及促進や電力自由化の影響から、従来需要家側であった主体がより一層発電し、電力の授受が盛んに行われるようになることが予想される。このような電力体制では、電力需要構造はより一層複雑化するが、マクロな予測では、複雑化に伴うパラダイムシフトに追従できない可能性が高いため、従来の電力需要予測では対応することができない恐れがある。地域の特徴、人口動態、発電方式等、様々な情報も利用して予測できるモデルが開発されれば、将来的に蓄電池の配置等電力システムを構築する政策決定を容易にする手段となり得る。

1.2 世帯推移シミュレータ

山際らは、世帯が持つ内部住民状態をライフイベントに伴い遷移させることで、都市の世帯数推移を再現・予測できる世帯推移シミュレータを開発した [2]。このシミュレータは、都市の意思決定主体を「世帯」という単位で考え、その世帯を構成する住民状態を内部変数として持つモデルとなっており、政策立案の検証に有用なモデルとなっている。本研究では、将来的な EV の普及や蓄電池の配置が地域の電力需要にどのような影響を与えるかについての検討を可能とするために、人口動態の変化を予測することができるこのシミュレータを活用する。本研究の位置付けを図 1 に示す。図 1 における人口動態シミュレーションが山際らが開発した世帯推移シミュレーションにあたる。本研究で開発するモデルは、そこから得られた世帯情報

(世帯数、世帯種割合) を入力データとして、電力需要量を推測する片方向連成モデルとなっている。

1.3 既存研究

家庭部門のエネルギー需要推定モデルにおいて、経済指標や統計データを用いたマクロ傾向を分析するもの [3] から住民の行動を実測調査に基づいてモデル化して時間別の電力需要を再現するモデル [4] まで多岐にわたる。前者は、全体の電力需要予測や政策決定の評価に有益ではあるが、個々の世帯や時間帯別の電力需要までを予測することはできないという問題がある。一方で後者は住民の行動まで再現することでよりミクロな電力需要まで再現することはできるが、住民の行動を網羅的に再現するために必要なデータが現時点で入手可能とは言えず、未知パラメータの増大、検証の困難さが依然として課題となっている。今後電力システムの転換により、電力需要が個々の世帯の影響を大きく受けることを考慮すると、電力需要を決定する要素数を抑えつつも、個々の世帯や時間帯別の電力需要まで予測可能なメソスコピックモデルがより一層必要になることが考えられる。

1.4 本研究の目的

本研究では、森田らの研究 [5][6] を参考に、複雑化する将来の電力発需要量に大きく影響すると考えられる家庭部門の電力需要を予測できるシミュレータの構築を目的とする。このシミュレータは、人口動態の変化、過去の電力使用量のデータ、平日・休日、発電方式を考慮して、それぞれの地域の特性を示すような情報を組み込み、時間別の電力需要曲線を予測できるモデルとなっている。また、将来的な電力需要の変動を予測し、電力システムにおける政策決定に有用な知見を提供することを目的とする。

2. 提案手法

2.1 モデルの概要

図 2 に電力需要推定モデルのフローチャートを示す。本研究では、電力需要に大きな影響与える要因を、世帯情報(世帯数、世帯種割合)、月、平日・休日、発電方法の 4 種類に絞り、これらを元に 15 分ごとの電力需要量を出力する。

連絡先: 田原智史, 東京大学大学院工学系研究科, 113-8656 東京都文京区本郷 7-3-1, 03-5841-6994, t.tahara@save.sys.t.u-tokyo.ac.jp

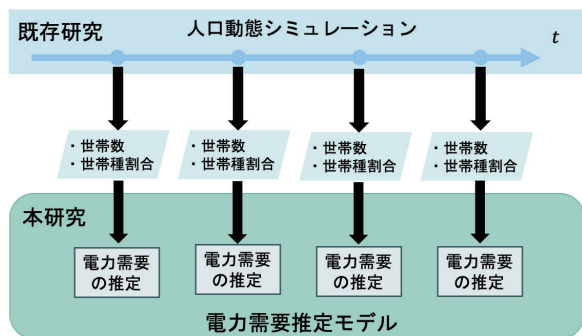


図 1: 開発モデルの位置付け

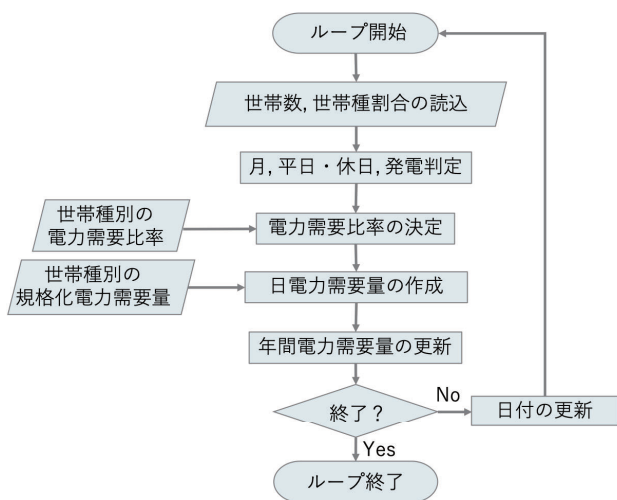


図 2: 電力需要推定フローチャート

2.2 世帯種

世帯・住民シミュレーションモデルはマクロモデルとミクロモデルとメゾモデルに大別される。マクロモデルは世帯や住民を集合体として扱い、その集合体の状態分布を遷移させることで世帯動態の再現するモデルである。このモデルの特徴は、少ないパラメータで多様な世帯動態の予測を可能にするが、行動主体である世帯・住民への影響や個々の住民の挙動までは考慮することができないという問題がある。一方ミクロモデルは、各住民・世帯を個別に扱い、住民・世帯毎の状態を遷移させることで、マクロモデルでは考慮が難しい行動主体である世帯・住民への影響や微視的な個々の住民の挙動を表現することができる。しかしこのモデルでは、複雑な社会現象を表現するために、未知パラメータや計算時間を増大化、パラメータ推定や妥当性の検証が困難であるという問題を抱えている。そこで山際は、各住民・世帯を個別に扱いながらも世帯の状態を単純化させることで、個々の世帯への影響を考慮可能にし、パラメータ推定・妥当性検証が容易であるメゾモデルシミュレータを開発した。このシミュレータでは、住民の現在状態を年齢区分毎に集約した内部状態として世帯に持たせることで、メゾスコピックな世帯エージェントを定義している。これにより、住民に関する未知のパラメータ数を抑制しつつ、住民の遷移状態に作用する政策の影響を反映した世帯動態の再現が可能となる。住民状態の表現には 4 種類の年齢区分と 3 種類の状態を用いる。それぞれ以下の通りである。

● 年齢区分

- － 子供 (0～19 歳)
- － 青年 (20～39 歳)
- － 中年 (40～59 歳)
- － 老年 (60 歳～)

● 状態

- － 状態 1: 世帯に存在していない
- － 状態 2: 世帯に存在し、かつ未婚である
- － 状態 3: 世帯に存在し、かつ既婚である

以上のようにして定義した年齢区分と状態の組み合わせから、全ての年齢区分で状態 1 である場合と、子供の年齢区分のみ状態 2 である場合の 2 通りを除いた 52 通りの内部状態が以下の 6 種類の世帯種に対応している。

- 単身世帯
- 夫婦のみ世帯
- 夫婦とその子供世帯
- 片親とその子供世帯
- 夫婦とその親世帯
- 三世帯世帯

本研究では、上述した世帯推移モデルを参考に、世帯エージェントを 6 種類の世帯種に割り振った。世帯内の住民の行動は考慮しておらず、世帯種が電力需要の傾向を決定し、それぞれ特徴のある電力需要比率と規格化電力需要量を内部変数として持たせている。また各世帯は、世帯の人数は明示的には考慮しておらず、各世帯種が持ち得る世帯人数を満たす世帯のデータを確率的に割り当てている。

2.3 発電方式

本研究では、世帯種毎にオール電化世帯と普通世帯に振り分けている。オール電化世帯とは、冷暖房、調理、給湯を使用するために電気を使っている世帯のことを表す。普通世帯とは、これらを電気以外のエネルギー源を使っている世帯のことを表す。これは、オール電化世帯はエコキュートを設置していることが多く、深夜の安価な電力を使って給湯を利用するために時間帯別の電力需要曲線に大きな差異が見られることから、世帯種毎に発電方式を 2 種類設定した。

2.4 電力需要量の算出方法

電力需要量は (1) 式で算出される。 $E_{i,T}$ は世帯 i の期間 T の電力需要量 [kWh]、 T は区間 t_1 から区間 t_n までの和、 e_i は世帯 i の規格化電力需要量 [kWh]、 $r_{i,k}$ は世帯 i の区間 k における電力需要比率 [ratio] である。図 3 は (1) 式、(2) 式を示した簡易図である。なお、本研究では $T = 24[\text{hour}]$ $k = 15[\text{min}]$ 、 $n = 96$ と設定している。

$$E_{i,T} = e_i \sum_{k=t_1}^{t_n} r_{i,k} \quad (1)$$

$$T = t_1 + t_2 + \dots + t_{n-1} + t_n \quad (2)$$

規格化電力需要量は世帯種、月、平日・休日、発電様式別に、ある日の 1 時間あたりの平均電力需要量とした。本モデルでは、日ごとの平均電力需要量を出し、その集めた数値の平均と分散に従う正規分布に基づいて規格電力需要量を出力する。

電力需要比率は、ある区間 k における電力需要量を規格化電力需要量で割った値のことである。本研究で扱う消費電力使

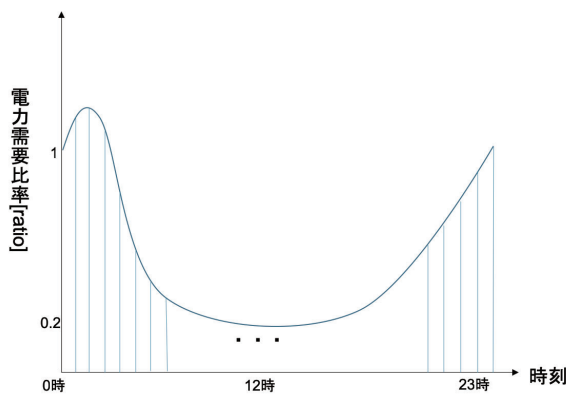


図 3: (1) 式を説明した簡易図

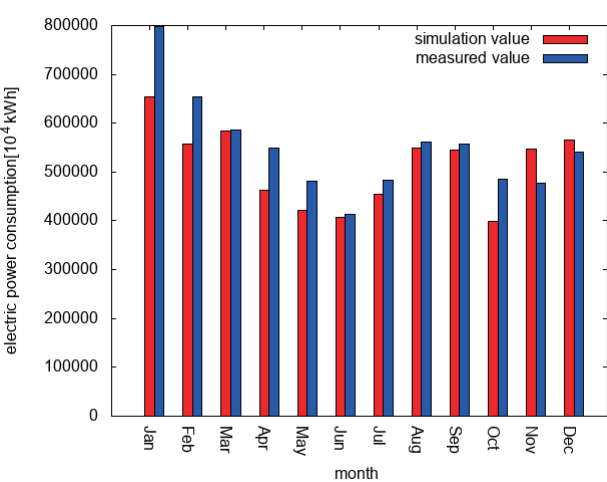


図 4: 月ごとの電力需要量の比較

表 1: パラメータ

パラメータ	値
単身世帯割合	0.310
夫婦のみ世帯割合	0.185
夫婦とその子供世帯割合	0.345
片親とその子供世帯割合	0.077
夫婦とその親世帯割合	0.052
三世帯世帯割合	0.052
オール電化普及率	0.042

用量のデータベース [7] には、15 分ごとにある世帯で使用した電力量が記載されている。このデータベースを元に、世帯種、月、平日・休日、発電様式別に、電力需要比率を平均値として算出し、固定値として出力する。

3. 数値実験

3.1 実験設定

本研究では、電力事業連合会による電力需要実績 [8] に掲載されている東京電力の電力供給量と関東圏に住んでいる世帯を対象にシミュレーションを行った電力需要量を比較する。本来は、世帯推移シミュレータから得られた世帯情報を利用するが、今回は検証のために実データを利用しており、世帯数は 2000 年と 2005 年の関東圏全域における世帯数の平均値として、1684 万世帯を採用する [9]。またタイムステップは 1 日とし、1 年分である 365 ステップのシミュレーションを実施した。

3.2 パラメータ設定

パラメータは世帯種割合とオール電化普及率であり、表 1 の通りである。オール電化普及率は 2009 年の関東圏における普及率を採用した [10]。本研究では、簡単のため各世帯種均等にオール電化普及率を割り振った。

3.3 結果

図 4 は、2003 年における月ごとの電力需要量における実測値とシミュレーション結果を示したものであり、図 5 は、図 4 におけるシミュレーション結果を世帯種別に示したものである。図 4 において、夏期と冬期における電力需要量の増加傾向は実測値と一致しており、電力需要の特性を示しているといえる。図 5 においても、世帯種毎の月別電力需要量の増減にばらつきが見られるものの、季節単位での傾向は実測値と一致した。

表 2: 月ごとの電力需要量における誤差の比較

月	誤差 [%]	月	誤差 [%]
1 月	-18.1	7 月	-6.3
2 月	-15.0	8 月	-2.6
3 月	-0.4	9 月	-2.2
4 月	-15.9	10 月	-18.2
5 月	-12.5	11 月	14.2
6 月	-1.4	12 月	4.3

表 2 は図 4 の実測値とシミュレーション結果における誤差を表したものである。全体として実測値より下方に出力され、大きな誤差の月として、1 月、4 月、10 月等が見られた。

このような誤差が現れた原因として、消費電力使用量のデータベース [7] の世帯数が 20 世帯程度であり、個々の世帯の影響を受けやすいことが考えられる。電力使用量のデータをより集めることでより精度の高い出力結果になることが考えられる。

4. まとめ

本研究では世帯数、世帯属性の特徴を考慮した電力需要推定モデルを開発した。月、平日・休日、発電方式を考慮し、月別の電力需要量の傾向を表す推定モデルを作成した。今後の課題として、時間帯別消費電力需要量に与える変数を刷新した精度の向上や時系列解析導入による精度の向上が挙げられる。また電力の使用量は価格の影響や経済状況の影響も受け、年ごとに一人当たりの電力使用量も変動するため、このようなマクロな変動も考慮することができれば、より精度の高いモデルへとなるだろう。今後の展望としては、蓄電池、再生可能エネルギーの普及率を考慮した電力需要推定モデルへの改良が期待される。蓄電池、再生可能エネルギーの普及の仕方は人口動態と密接に関わっており、普及の仕組みをモデル化できれば、電力システムの変革を実行する際の意思決定に有益となるモデルへとなるだろう。

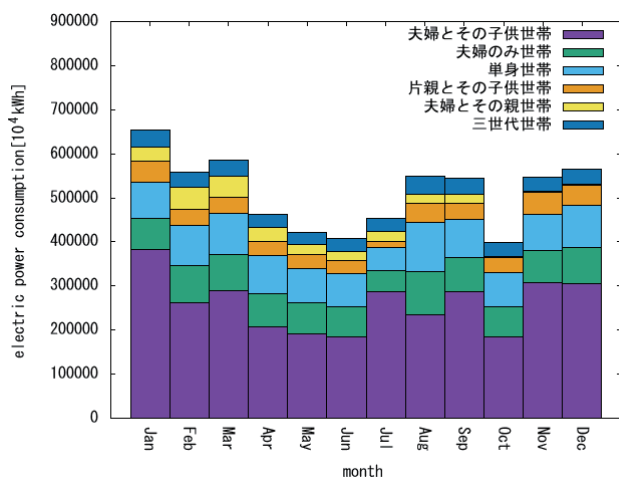


図 5: 世帯種別の月ごとの電力需要量

参考文献

- [1] 資源エネルギー庁, “分散型エネルギーについて”, http://www.enecho.meti.go.jp/committee/council/basic_policy_subcommittee/mitoshi/006/pdf/006_05.pdf
- [2] 山際康平, 藤井秀樹, 吉村忍, “メゾスコピックモデルを用いたマルチエージェント世帯推移シミュレーション”. 人工知能学会論文誌, Vol. 32, No. 5, pp. 1- 10. 2017.
- [3] Bentzen J, Engsted T, “A revival of the autoregressive distributed lag model in estimating energy demand relationships”, Energy, Vol.26, Issue 1, pp. 45- 55. 2001
- [4] J.Widén and E.Wäckelgård, “A high-resolution stochastic model of domestic activity patterns and electricity demand”, Appl. Energy, Vol.87, p.1880. 2010
- [5] 森田圭, 真鍋勇介, 加藤丈佳, 舟橋俊久, 鈴置保雄, “消費電力量と世帯属性の分類に基づく世帯群の合計消費電力量推定に関する一検討”, 電気学会論文誌, Vol.136, No.6, 2016.
- [6] 森田圭, 真鍋勇介, 加藤丈佳, 舟橋俊久, 鈴置保雄, “数百世帯規模の家庭の平均電力需要特性の評価”, エネルギー資源学会論文誌, Vol.38, No.1, 2017
- [7] (社) 日本建築学会, 住宅内のエネルギー消費量に関する調査研究委員会:住宅におけるエネルギー消費量データベース, <http://www.arg.ne.jp/node/2198>
- [8] 電力事業連合会, 電力需要実績, <http://www.fepc.or.jp/library/data/demand/2010.html>
- [9] 国立社会保障・人口問題研究所, 日本の世帯数の将来推計(都道府県別推計), <http://www.ipss.go.jp/pp-pjsetai/j/hpjp2014/gaiyo/data.asp>
- [10] 富士経済, “エネルギー需要家別マーケット調査要覧 2010 住宅分野編”

スマートグリッドにおける送電経路の最適化

Optimization of Power Electric Supply Path in Smart Grids

小澤 貴也 大沢 英一

Takaya Ozawa Ei-Ichi Osawa

公立はこだて未来大学 システム情報科学部 複雑系知能学科

Department of Complex and Intelligent Systems, School of Systems Information Science, Future University Hakodate

The next generation next-generation electric grids called smart grids are attracting attention due to problems such as global power shortage and environmental load caused by power generation. One of the problems with the current power being used is that the power transmission loss is large due to the long power transmission path. From such a problem, it is possible to reduce the power shortage and reduce the power generation amount by reducing the power transmission loss. For this reason, it is required to optimize the power transmission path. In this paper, we propose to reduce the power transmission loss by minimizing the power transmission path by using the code optimization method used by the compiler by treating the power transmission path like a program. We compared power transmission loss by simulation using multi agent system and showed that the proposed method can reduce transmission loss. In addition, it shows that it can be used by measuring the execution processing time of the proposed method.

1. はじめに

現在、日本の電力使用量は、情報化の進展やエアコンの普及にみられるような快適な生活へのニーズが高まり、戦後一貫して伸びている。また、今後電気自動車などの普及が予測されるため、電力需要はさらに増加すると考えられる。このような世界的な人口増加や電力消費量の増加に伴う電力不足や、火力発電による CO_2 排出など発電による環境負荷が問題などによりスマートグリッドという次世代型電力網が注目を集めている。スマートグリッドとは、次世代型送電網と意味であり、電気自動車などの電力貯蔵システムや太陽光発電などの分散型電源とスマートメータという人工知能や通信機能を搭載した計測機器等を設置して電力需給を自動的に調整、最適化を可能にする電力網を構築するという概念で効率的な電力供給が可能になるとされている [谷口 10]。

また、現在の使用されている電力の問題点としては、その送電の過程で経路が長いことから送電損失が多いという点がある。日本では電力が家庭に届くまでに送電損失によって約 4.8% 失われている。年間では約 480 億 kWh 送電損失となっており、これは 100 万 kWh 級原子力発電所の約 7 基分の発電量に相当する。このような問題から送電経路の最適化を行うことで、送電損失量を削減することが可能であり、電力不足の改善や発電量の削減が可能であると考えられる。そこで、本研究では、電力の送電経路をプログラム言語のように扱い、コンパイラで利用されるコード最適化の手法を用いて送電経路の最適化を可能にして、送電損失を削減する手法を提案した。

2. 前提知識

本章では、本研究の内容で用いる特に必要な知識について述べる。

2.1 複雑ネットワークモデル

現実世界に存在するネットワークは多くは複雑ネットワークの性質が存在することがわかっている。電力網もまたスモールワールド性やスケールフリー性の性質を持っていることがわかっているため、本研究では複雑ネットワークモデルを使用する。

複雑ネットワークを生成するモデルとして「ワッツ・ストロガッツモデル (WS モデル)」と「バラバシ・アルバートモデル (BA モデル)」などが存在する。これらのモデルは単純なアルゴリズムより複雑ネットワークを生成することが可能であり、本研究では、送電網の形状が複雑ネットワークの特性を持つことから WS モデル、BA モデルを使用して送電網を構築する。

2.2 クラスター係数

クラスター係数は、任意のノードに隣接するノード同士が隣接ノードである割合で、クラスター係数が高いネットワークは関係密度が高いというものである。ノード v_i のクラスター係数 C_i は、ノード i の次数 k_i を用いて以下の式で算出する。

$$C_i = \frac{v_i \text{ を含む三角形の数}}{\frac{k_i(k_i-1)}{2}}$$

2.3 媒介中心性

媒介中心性は、ノードが情報を媒介することに関与している度合いを表す尺度である [Brandes 01]。ノード v_i の媒介中心性 B_i は、始点 v_{is} から v_{it} の最短経路の中で v_i を通るものを $g_{is, it}^i$ として、最短経路の総数を $N_{is, it}$ とするとき、以下の式で算出される。

$$B_i = \frac{\sum_{is=1; is \neq i}^n \sum_{it=1; it \neq i}^{is-1} \frac{g_{is, it}^i}{N_{is, it}}}{(N-1)(N-2)/2}$$

本研究では、媒介中心性が高いノードは送電線が多く集まるノードである。そのため、ノードの媒介中心性とそのノードから出ていく経路数との相関について 5 章で考察する。

連絡先: 小澤 貴也, 公立はこだて未来大学 システム情報科学部複雑系知能学科, 北海道函館市中野町 116-2, 0138-34-6448, b1015112@fun.ac.jp

2.4 電力に関する基本的な数式

本研究では、現在送電に使用されている導線の抵抗率からシミュレーションに必要な電力に関する数値を算出した。本章では、研究に必要な基本的な電力に関する数式を必要な知識について示す。電力とその送電による送電損失は以下の式によって算出される。

$$\text{送電抵抗} : R = \rho \times l$$

$$\text{送電損失} : P(W) = I^2 \times R$$

- W : 電力, I : 電流, ρ : 導線抵抗, R : 導線の長さ

本研究では、送電損失が導線の長さ按比例することから、導線の長さつまり送電距離を最適化することで送電損失量を削減可能と想定している。

3. 関連研究

3.1 経路探索を用いた電力損失を最小化する経路選択手法

スマートグリッドにより電力の流れを制御することで、電力を削減する研究がある [尾倉 14]。この研究では、ヨーロッパの 15 カ国を結ぶ実際の送電網のうち 2 カ国を繋げる経路を最短経路で電力を削減するというものである。

方法としては、エッジとノードの数を最小にするように経路を選択することで、送電損失量が最小になるという考えから幅優先探索により経路を決定する。また、シミュレーション環境としては、15 カ国をノードとし、送電網をエッジとする。また、それぞれのエッジには送電損失量を超伝導交流送電線を参考に設定している。

実験結果から表 2.1 および表 2.2 から必ずしも最小経路が一番電力損失が小さいわけではなく、各ノードの送電損失率も影響していることがわかった。

表 1: 供給地'n5', 需要地'n12' の場合 ([尾倉 14])

n8 ~ n14	距離	送電損失率
$n8 \rightarrow n4 \rightarrow n10 \rightarrow n13 \rightarrow n14$	2954.74km	35.93%
$n8 \rightarrow n9 \rightarrow n10 \rightarrow n13 \rightarrow n14$	2638.54km	36.30%
$n8 \rightarrow n12 \rightarrow n10 \rightarrow n13 \rightarrow n14$	3421.55km	41.14%

表 2: 供給地'n5', 需要地'n12' の場合 ([尾倉 14])

n5 ~ n12	距離	送電損失率
$n5 \rightarrow n4 \rightarrow n8 \rightarrow n12$	1866.53km	23.98%
$n5 \rightarrow n7 \rightarrow n10 \rightarrow n12$	1543.51km	23.99%
$n5 \rightarrow n4 \rightarrow n9 \rightarrow n12$	1931.14km	25.79%
$n5 \rightarrow n4 \rightarrow n10 \rightarrow n12$	2558.94km	29.60%

3.2 コード最適化手法を用いてトラック経路の効率化を行う手法

共同物流のトラックをコンパイラで利用されるコードの最適化を用いて、トラック経路の効率化を行い CO_2 を削減する研

究として佐藤による研究がある [SATO 08]。

この研究では、プログラム言語の実行フローとトラック経路の類似性があることからトラックの経路をプログラム言語として扱えるようにして、プログラムのための各種技術を通じてトラック経路を効率化を行った。経路の効率化として、コード最適化手法として逐次化と冗長コード除去を用いることでトラック輸送において同じ集配先に何度も行かないように経路を変更した。また、並列プログラムなどでは逐次実行プログラムの並列化とは逆に並列プログラムを逐次化 (直列化) することでトラックの数を減らすことを可能にした。

また、実際のトラック輸送経路に対して、所定の輸送条件を満足させながら効率化した結果、10% 程度の効率化できた。

本研究では、このコードの最適化手法によるトラックの経路の最適化を元に集配先を一般家庭とし、トラックの経路では、一つ経路最適化であったものを複数の経路による送電経路の最適化を行った。

4. 提案手法

本研究では、送電損失の削減のために送電経路の最適化を行う。最適化の方法として、佐藤の研究 [SATO 08] を参考に送電経路をプログラム言語として扱うことで、電力供給条件を満足させながらコンパイラのコード最適化手法として冗長コード除去や逐次化 (直列化) を用いる。本章では、提案する手法による経路選択方法について述べる。

4.1 コードの最適化

提案手法では、コンパイラで使われるコードの最適化手法である逐次化 (直列化) および冗長コードの除去の手法を用いている。

逐次化 (直列化)

逐次化とは、2つのコードを統合することにより冗長なコードを削除することで最適化を行うというものである。つまり、本手法では、図 1 のように送電を一括化して、経路の削減を行う。また、高電圧で送電を行うことで送電損失は削減が可能となる。

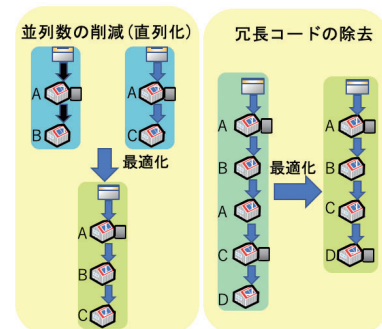


図 1: 電力網における逐次化 (直列化) および冗長コードの除去の例

冗長コードの除去

冗長コードの除去とは、コード内の冗長な使用されていないコードを削除することで最適化を行うというものである。電力網の場合、構造から送電経路が同じノードへ戻ってくることがあることからノードの重複する移動を除去する。つまり、図 1 のように電力要求のない経路内のエージェント削減を行う。

4.2 送電経路選択方法

提案手法による経路選択は以下のような流れで行う。
送電経路の選択方法：

- 1 すべての電力需要家を求めて、貯蓄設備から需要家を結ぶ経路を作成する。
- 2 コード最適化 (冗長コード除去や逐次化 (直列化)) を行う。
- 3 距離の短い送電経路を選択

例として、図 2 の 1 のような電力網の貯蓄設備から B,C への供給を考える。まず 2 のように貯蓄設備と B,C を結ぶ経路を作成する。次に、作成された経路に逐次化と冗長コードの除去を行うことで 3 のような経路が決定する。

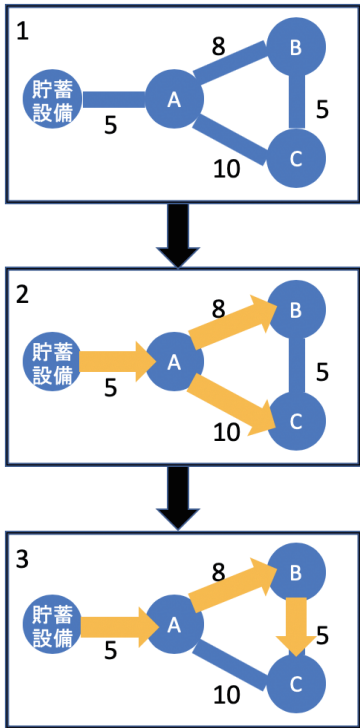


図 2: 提案手法の使用例

5. 実験

本実験には、多数の需要家をエージェントとする大規模エージェントネットワークを構築し、提案手法とダイクストラ法の送電経路を計測し送電損失量を比較した。

計測日数を 365 日とし、太陽光発電により発電した電力の融通可能であると想定する。また、送電による電力の損失は、送電線の導線抵抗のみを想定する。また、需要と供給源のマッチングには、空間マッチングを用いて行う [山本 11]。また、太陽光発電の発電量は函館市の一年間の日射量から算出した。

本研究では、潮流などをさけるため物理的制約として、電力は一方にしか流れないものとする。また、配電用変電所から家庭への送電を想定している。

5.1 実験パラメータ

実験にはエージェント数は 100~1000 でそれぞれ 3 回行う。その他のパラメータは、実際に使用されてるものを参考に以下のように設定した。

- 平均電力消費量：18.5kWh
- 発電システムの設置容量：4.5kWh と
- 送電線の許容電流：6600A
- 平均次数: $k = 4$

太陽光による発電可能エージェントの割合を約 8%とした。エージェントの電力需要量は、用意した電力消費需要パターンの朝型、夜型、中間型の 3 種類用意からランダムに設定した。シミュレーションの電力融通市場は、送電網を複雑ネットワークモデルである WS モデルと BA モデルで生成する。電力融通を行うエージェントは、蓄電池を持つ一般家庭と電力事業者とした。一般家庭を想定したエージェントの約 8%が、太陽光発電が可能と設定した。送電量は、送電線の許容電流以上の電流を流すことはできないとする。

5.2 実験結果 1：総経路削減率

エージェント数 500 を 3 回行ったときの平均削減率の結果を表 3 に示す。ダイクストラ法と提案手法を比較した結果、BA モデルでは提案手法によって 4.8% 削減が可能となった。WS モデル $p = 0.1$ の場合は 29.5% 削減が可能となった。

表 3: 削減率 1

	BA モデル	WS モデル ($p = 0.1$)
提案手法	3952.8km	9112.3km
ダイクストラ法	4153.63km	12926.8km
削減率	4.8%	29.5%
クラスター係数	0.096%	0.379%

また、レギュラグラフに近いネットワークが生成される $p = 0.05$ の WS モデルとランダムグラフに近いネットワークが生成される $P = 0.5$ の WS モデルで実験を行った結果は表 4 である。結果から $p = 0.05$ の WS モデルのほうが削減率が上がった。

表 4: 削減率 2

	WS モデル ($p = 0.05$)	WS モデル ($p = 0.5$)
提案手法	5991.9km	8931.0km
ダイクストラ法	14726.8km	11705.6km
削減率	59.3%	23.7%
クラスター係数	0.415%	0.073%

この結果ももっともクラスター係数が高い $p = 0.05$ の WS モデルが一番削減率が高く、クラスター係数をもっとも低い BA モデルが一番削減率が低いことから提案手法による経路削減率はクラスター係数が関係していることが考えられる。このことは、最適化手法の特に逐次化が関係していると考えられる。また、BA モデルと比べ WS モデルの削減率が著しく上がった要因として複雑ネットワークモデルの特性であるスケールフリー性に関係していると推測される。BA モデルはスケールフリー性を持つが WS モデルは持たない。このことから最適化手法が行われる区間が供給源から次数の高いノード特にハ

ブのようなノードまでとなり削減率が下がることが考えられる。

また、エージェント数 1000 で 3 回ずつ、2017 年度の北海道電力の電力使用量から算出した北海道内の送電損失量と提案手法による送電損失量を比較した結果、WS モデルでは約 49.8%，BA モデルでは約 45.49% の削減が可能となった。この結果から提案手法による経路選択は送電損失量の削減が可能であると考えられる。しかし、本研究では、送電損失は送電線の導線抵抗のみを想定している。実際の損失にはコロナ放電などの送電損失があるため実験結果よりも削減率は下がると考えられる。

次に、エージェント数 1000 の場合のノードの媒介中心性と提案手法による経路選択によって、そのノードを経由する経路の数の結果は図 3 および図 4 である。結果より、ノード媒介中心性とそのノードを経由する経路の数には相関が示唆された。

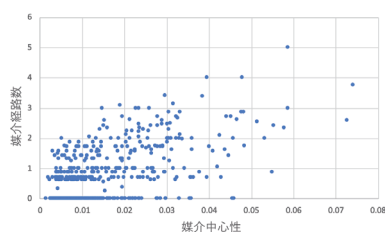


図 3: WS モデル

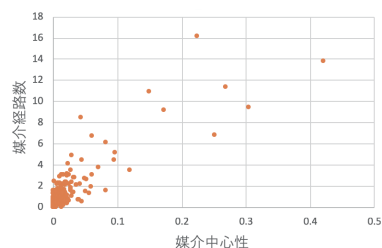


図 4: BA モデル

5.3 実験結果 2：実行速度

エージェント数 100～1000 までをそれぞれ 3 回行いその平均処理時間を表したグラフが図 5 である。結果から WS モデルより BA モデルの方が処理時間が短いことがわかった。また、平成 27 年の北海道内の世帯数と変電所数からひとつの変電所あたり約 6600 世帯の供給を行う。実験で得られた処理時間に最小二乗法を適用して推定した曲線から 6600 世帯への提案手法による経路選択にかかる処理時間は WS モデルでは約 1.64 秒で、BA モデルでは 0.85 秒であると推測される。よって、現在の送電網でも実行可能時間で提案手法を用いて送電経路を決定することは可能であると考えられる。

6. おわりに

本研究では、現在の送電システムでは送電損失量が多いという問題に対して、コードの最適化手法を送電経路に用いることで経路の削減を行い、それにより送電損失量の削減を行った。ネットワークの生成には実際の送電網に近い構造を持つ複雑

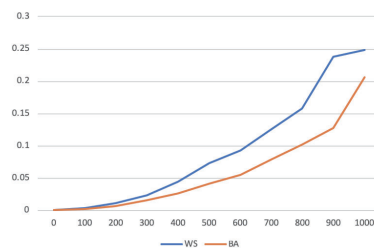


図 5: 平均処理時間

ネットワークモデルである WS モデルと BA モデルで生成した。送電損失の削減の評価方法としては、ダイクストラ法を用いて送電経路を生成したものと北海道の送電損失量を算出し、提案手法との比較を行った結果、提案手法ではダイクストラ法による経路と北海道の送電損失量に比べ送電損失量が少ないことを示した。

また、ネットワークの特性と提案手法との関係調べるためノードの媒介中心性とそのノードを経由する経路数との関係を調べた。実験結果から媒介中心性の高いノードには多くの経路が集まることが示された。そして、実際に使用可能であるか調べるため、提案手法の実行処理時間の計測した結果、実行処理時間から現在の送電網でも使用可能であることを示した。よって、このことから提案手法は、送電損失量の削減が可能であり、現在の送電網で使用可能であると考えられる。

また、本研究では、送電の全体の経路の削減を行ったがもともと送電損失量が少なくなる経路選択ではない。また、送電損失量の削減の最適解を取るには多くの処理時間が必要となる。よって、処理時間は変わらずに提案手法より送電損失量の削減方法の提案が必要となる。また、本研究では経路の削減を行ったが、さらに需要家と供給家の組み合わせによって送電損失量の削減をする方法を模索したい。

参考文献

- [谷口 10] 谷口忠大, 榊原一紀, 西川郁子: 自律分散型 スマートグリッド上の電力取引に対する自然方策勾配法によるマルチエージェント強化学習の有効性検証. 第 22 回自律分散システム・シンポジウム, 2010
- [尾倉 14] 尾倉康男, 篠宮紀彦: 送電網において電力損失を最小化する経路選択手法, 電子情報通信学会技術研究報告 = IEICE technical report: 信学技報, 113(427), 2014.2.6・7, p.63-65
- [SATO 08] SATOH I.: A Specification Framework for Earth-friendly Logistics. Proceedings of 28th IFIP WG6.1 International Conference on Formal Techniques for Networked and Distributed Systems (FORTE ' 2008), Springer June, 251-266, 2008
- [山本 11] 山本草詩, 石井大介, 岡本聡, 山中直明: スマートグリッドにおける送電ロスが最小となる発電源選択法の一検討. 電子情報通信学会技術研究報告 = IEICE technical report: 信学技報, 2011
- [Brandes 01] Brandes U.: A faster algorithm for betweenness centrality, Journal of Mathematical Sociology, Vol.25, No.2, pp.163-177, 2001

電力分け合いサービスの MAS モデルの設計と基礎評価

Design and Preliminary Evaluations of Multi-Agent Simulation Model for Electric Power Sharing among Households

西村 康孝^{*1}
Yasutaka Nishimura

志村 泰知^{*2}
Taichi Shimura

和泉 潔^{*3}
Kiyoshi Izumi

吉原 貴仁^{*1}
Kiyohito Yoshihara

^{*1} KDDI 総合研究所
KDDI Research Inc.

^{*2} 構造計画研究所
Kozo Keikaku Engineering Inc.

^{*3} 東京大学
The University of Tokyo

Abstract: Electric power sharing among households based on bidding method is studied as a future service. In order to verify the feasibility of the service, we newly designed a multi-agent simulation model. We validate this model through preliminary evaluations. For example, it is confirmed that the market price in this service stably changes according to the supply and demand balance between sold bid volume and purchased bid volume. In addition to that, results of household income and contract rate of this service showed that the design for bid strategies works as intended.

1. はじめに

太陽光発電(以下, PV)の余剰電力買取制度(以下, FIT)で設定された 10 年間の買取義務保証期間が終了する世帯(以下, 卒 FIT 世帯)が 2019 年から登場し始める. FIT 期間中は PV 発電量の内, 宅内消費で余った余剰電力は固定単価での買取が保証されていた. 10 年間保証される固定単価は 2009 年開始世帯の 48 円/kWh から年々低価格に推移し, 2018 年開始世帯は 26 円/kWh であった. 卒 FIT 世帯の余剰電力の買取単価は保証されていた固定単価から大きく低下すると言われており, いくつかの小売電気事業者から発表されている卒 FIT 世帯向け新買取サービスの水準は約 10 円/kWh である. このため, 卒 FIT 世帯の余剰電力の価値を高める手段として, 蓄電池導入による自家消費の促進や, 余剰電力を他世帯へ売電する電力分け合い[井上 11]などが検討されている.

電力分け合いでは, 余剰電力を小売電気事業者の買取単価より高く他世帯へ売電できると得をする. PV を保有していない世帯も, 宅内消費する電力を小売電気事業者からの買電単価より安く他世帯から買電できると得をする. 余剰電力の総量と宅内消費の総量は常には一致しないため, 売買相手を決定する取引方法が必要になる. 取引方法には日本卸電力取引所(以下, JEPX)で発電事業者と小売電気事業者との売買で用いられている入札方式などがある.

電力分け合いは既存サービスがほとんどなく, 世帯と小売電気事業者の利益や市場価格の安定性, 市場成立に必要なサービス加入世帯数など, サービスの実現性検証が課題である. しかしながら, 入札を想定した電力分け合いは, 世帯の入札行動が天気や季節などの環境により変化し, 他世帯の入札と相互作用する複雑系であり検証が困難であった. このような複雑系の解析には MAS(Multi-Agent Simulation)[和泉 17]が有望である. そこで本稿では, 電力分け合いの実現性検証のための電力分け合い MAS モデルを設計し, 基礎評価を通してモデルの妥当性を確認する.

2. 関連研究

電力・エネルギー分野への MAS の主な応用事例として, 電力市場と送配電システムに関する新制度を, 安定性や効率性, 有効性などの観点から評価検証する研究がある. 電力市場に関しては, 国内外の多くの研究でエージェントシミュレーションが使われ[渡邊 03][Ullah 13], 既に米国ではエージェントベースの大規模な電

力市場テストベッド AMES が構築されて電力市場制度の評価検証で利用されている[Sun 07]. 送配電システムに関して, 実際の宅内消費データや発電データを用いて電力消費エージェントの挙動を決定しスマートグリッドシステムの効率性を評価検証した研究[Vytelingum 10]がある. また, 世帯や企業, 小売電気事業者が入札方式で参加する電力市場を想定し, 分散した小規模な電力市場により電力価格と配電ネットワークを創発的に構築する新たな電力流通システムの有効性を分析する研究[井上 11]などがある.

以上のように, 入札方式による世帯間の電力売買を想定した MAS の応用事例は多くは知られていない. [井上 11]の研究では想定は類似しているが, 主な目的は電力流通システムの評価検証であった. 本稿では, 電力分け合いサービスの実現性検証を目的とし, 利益追求や再エネ由来の電力を使いたいエコ志向など, 世帯の異なる考え方をモデル化したり, 世帯利益を評価したりするなど, サービス観点での評価検証を重視する MAS モデルを設計する.

3. 電力分け合いサービス

3.1 サービス概要

世帯間で入札を通して電力を売買する. PV, 定置用蓄電池, EV(Electric Vehicle)の内 1 つ以上保有する世帯が電力分け合いで売り手になれる. PV 保有世帯は PV 発電量から宅内消費量を引いた余剰電力を売電できる. PV に加えて, 定置用蓄電池や EV などの蓄電池を保有する世帯は, 余剰電力を蓄電して好きなタイミングに売電することもできる. 蓄電池のみを保有する世帯は, 夜間に小売電気事業者から買電した電力を蓄電し, 昼間に売電することができる.

全世帯が電力分け合いで買い手になれる. 宅内消費する電力を他世帯から買電できる.

3.2 想定環境

(1) 小売電気事業者との契約

PV 発電が無い夜間など, 電力分け合いだけで各世帯の宅内消費を賄うことはできない点を考慮して, 各世帯は電力分け合いサービスに加えて, 従来の買電を小売電気事業者と契約していることを想定する. また, 入札により全ての売電入札が約定しない点を考慮して, 売り手になれる世帯は小売電気事業者の定価買取を契約していることを想定する.

連絡先: 西村 康孝, KDDI 総合研究所, 埼玉県ふじみ野市
大原 2-1-15, yu-nishimura@kddi-research.jp

(2) 電力分け合いに対する世帯の考え方

利益追求, エコ志向, 無関心の 3 つを想定する. 利益追求は電力分け合いを通じた利益増を追求する. エコ志向は電力分け合いを通し, 小売電気事業者との定価売買と比較して損をしない範囲で, 自世帯の宅内消費電力に占める再エネ由来電力の比率向上を志向する. 無関心は小売電気事業者との定価売買と比較して損をしない範囲であれば満足し, 利益増や再エネ由来電力の比率の向上に関心が低い.

(3) 約定方式

約定方式として, JEPX のスポット市場などで用いられているブラインド・シングルプライスオークションを想定する.

(4) 電力の流れ

電力分け合いの売買取引と合わせて電力も世帯間で直接流す方法と, 既存の電力の流れは変えず, 売買取引のみ仮想的に世帯間で実施する方法が考えられる. 前者は既存の電力インフラに大規模な改修が必要になるため, 広範囲な電力分け合いを想定すると非現実的である. 本稿では後者を想定する.

3.3 取引イメージ

30 分を 1 コマとした 1 日 48 コマについて電力分け合いの売買取引を実施する. 各コマはコマ ID(1~48)で表す. 約定した世帯間では, 約定価格で電力を売買する. 約定しなかった世帯は, 小売電気事業者と定価で電力を売買する.

ある日のコマ ID=20(9:30 から 10:00 のコマ)を例に, 入札から約定までの流れを図 1 に示す. 売電世帯 S1, S2 と買電世帯 B1, B2 はコマ ID=20 が始まる 9:30 より前に入札価格を決定する(図 1(1)). 各世帯の入札量は 9:30 から 10:00 の PV 発電や宅内消費の実績に基づき事後に確定する(図 1(2)). 入札価格と入札量の情報を突き合わせて約定計算する(図 1(3)). ブラインド・シングルプライスオークションの価格優先の原理に基づき, 売電世帯 S1 の売電入札 10kWh と買電世帯 B1 の買電入札 10kWh が約定する. 残った売電世帯 S2 の売電入札と, 買電世帯 B2 の買電入札は, 入札価格が合致しないため約定しない. 売電世帯 S2 の入札量 10kWh は小売電気事業者が定価で買い取る. 買電世帯 B2 の入札量 5kWh は小売電気事業者から買電単価で買電する.

4. 電力分け合い MAS モデルの設計

電力分け合いサービスとその取引を模擬する電力分け合い MAS モデルを設計する. モデルは世帯エージェントと小売電気事業者エージェントから構成する.

4.1 世帯エージェント

世帯エージェントの属性情報として, 保有設備, 小売電気事業者との契約情報, 入札戦略を設定できる.

(1) 保有設備

PV や定置用蓄電池, EV の有無を設定する. 有りの場合は, PV の定格出力値[kW]や蓄電池の蓄電容量[kWh]などの保有設備のスペック情報を合わせて設定する.

(2) 小売電気事業者との契約情報

小売電気事業者からの買電単価[kWh/円]と, 小売電気事業者の買取単価[kWh/円]の情報を設定する. これらの単価はコマ毎に設定でき, 時間帯に依らず固定値としたり, 日中と夜間で異なる単価としたりすることができる.

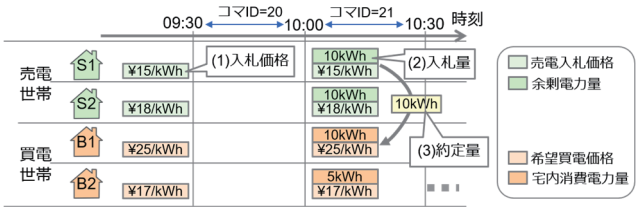


図 1: 入札と約定の処理イメージ(コマ ID=20 の場合)

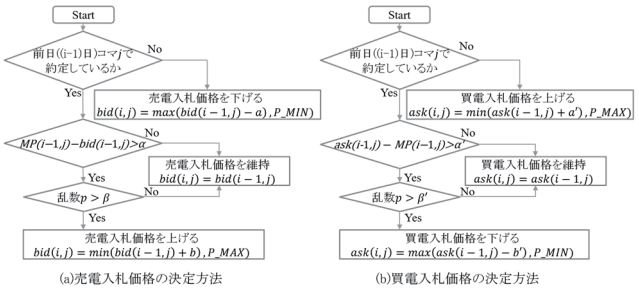


図 2: 利益追求戦略の入札価格の決定方法

(3) 入札戦略

電力分け合いに対する世帯の考え方に対応する入札戦略として, 利益追求, エコ志向, 無関心の中から 1 つを選択する. 選択した入札戦略に応じて世帯は入札価格を決定する. 小売電気事業者からの買電単価より高い買電入札価格や, 小売電気事業者の買取単価より安い売電入札価格は非合理的であるため, 全入札戦略に共通して, 入札価格の上限 P_MAX は小売電気事業者からの買電単価, 下限 P_MIN は小売電気事業者の買取単価とする.

利益追求戦略は, 過去の約定結果と市場価格に基づいて, 適宜入札価格を変更する. i 日のコマ ID = j ($j = 1, 2, \dots, 48$)において, ある世帯の売電入札価格 $bid(i, j)$ は, 前日 $((i-1)$ 日)のコマ ID = j における同世帯の売電入札価格 $bid(i-1, j)$ と, そのコマの市場価格 $MP(i-1, j)$, 同世帯の約定結果の 3 点の情報により図 2(a)の通り算出する. 前日同コマで約定していない場合は, 入札し易くなるよう入札価格をパラメータ a [円/kWh] 下げる. 前日同コマで約定している場合, 市場価格と前日同コマの入札価格の差が閾値 α 以下の場合, 入札価格を上げると約定しなくなる可能性が比較的高いため入札価格を維持する. 一方, 閾値 α より大きい場合は, 乱数 p ($0 \sim 1$) を用いて $p > \text{閾値}\beta$ の場合に確率的に入札価格を b [円/kWh] 上げて利益増加を図る. ここで, a, b, α, β は 0 以上の実数のパラメータである.

利益追求戦略の買電入札価格も同様に変更する. i 日のコマ ID = j において, ある世帯の買電入札価格 $ask(i, j)$ は前日 $((i-1)$ 日)のコマ ID = j の同世帯の買電入札価格 $ask(i-1, j)$ と, そのコマの市場価格 $MP(i-1, j)$, 同世帯の約定結果により図 2(b)の通り算出する. 前日同コマで約定していない場合は, 入札し易くなるよう入札価格をパラメータ a' [円/kWh] 上げる. 前日同コマで約定している場合, 市場価格と前日同コマの入札価格の差が閾値 α' 以下の場合, 入札価格を下げると約定しなくなる可能性が比較的高いため入札価格を維持する. 一方, 閾値 α' より大きい場合は, 乱数 p ($0 \sim 1$) を用いて $p > \text{閾値}\beta'$ の場合に確率的に入札価格を b' [円/kWh] 下げて利益増加を図る. ここで, a', b', α', β' は 0 以上の実数のパラメータである.

エコ志向戦略は約定率を高めるため, 常に成り行き注文を実行する. 無関心戦略は常に同じ入札価格での入札を実行する.

なお, 各世帯の入札価格の初期値は小売電気事業者からの買電単価と小売電気事業者の買取単価との間の一様分布で与える.

(4) 世帯エージェントの動作

各コマでの宅内消費電力や PV 発電量から売電と買電のポジションが決定する。ポジション決定後、入札戦略に基づいて入札価格を算出して入札する。

4.2 事業者エージェント

事業者エージェントは世帯の売買入札を集約し、約定計算を実施して約定結果を当該世帯エージェントに通知する。市場価格の情報は全世帯エージェントに公開する。約定結果に応じて電力分け合いの売買と小売電気事業者との売買を集計し、買電による世帯支出と売電による世帯収入を計算する。

5. 電力分け合い MAS モデルの基礎評価

5.1 評価方針

電力分け合いサービスの実現性検証に電力分け合い MAS モデルを活用することを目標に、基礎評価を通して同モデルの妥当性を確認する。具体的には以下の 2 点を確認する。

- 電力分け合いの市場価格が売電入札量と買電入札量の需給バランスに応じて安定的に形成される点
- 利益追求やエコ志向、無関心の入札戦略毎の世帯利益や約定率が設計(4.1 章(3))通りになる点

5.2 評価指標

市場価格推移と需給バランス、世帯利益、約定率を評価指標に用いる。市場価格は各コマでの買電入札と売電入札を突き合わせて約定計算により算出される。需給バランスは(売電入札量[kWh]/買電入札量[kWh])で計算する。夜間は 0 になり、日中は PV 発電量の増加に伴い増加する。

世帯利益については、小売電気事業者と電力売買した場合の収支を起点とし、電力分け合いにより得た金額として集計する。これより世帯利益は 0 円以上になる。電力分け合いによる売買が無い世帯の利益は 0 円になる。電力分け合いにより、小売り電気事業者からの買電単価や、買取単価と比較して得する価格で売電できると世帯利益が 0 円より大きくなる。約定率は約定量[kWh]/入札量[kWh]として集計する。

5.3 入力データ

宅内消費電力データの例を図 3 に示す。ある電気サービスの宅内消費電力データを基に統計的に加工して作成した。平日は起床後(コマ ID=15~17(7 時~8 時 30 分))にピークがあり、外出後(コマ ID=18(8 時 30 分~9 時))以降に宅内消費が低下するのに対し、土日祝日は起床時間が平日と比較して遅い点や日中も宅内消費電力が平日と比較して高い傾向など、平日と土日祝日の宅内消費パターンの違いを模擬している。

PV 発電量データの例を図 4 に示す。PV 発電量データは東京都の 7 月 1 日~7 月 31 日を対象とし、日の出と日の入りの時間や太陽高度の情報をを用いて数値シミュレーションで作成した。PV の定格出力は一般的な水準である 4kW とし、夏至の正午をピークに、その他の日や時間帯は太陽高度に応じて減少する形で作成した。PV 発電量への天候や日照量、発電ロスの反映は今後の課題である。

5.4 評価条件

基礎評価のための評価条件を表 1 に示す。シミュレーション期間は 7 月 1 日~7 月 31 日の 1 か月間で世帯数は 10,000 世帯とする。小売電気事業者からの買電単価は 26 円/kWh、小売電気事業者の買取単価は 5 円/kWh とする。これらの単価は 1 日 48 コマ共通の固定値とする。

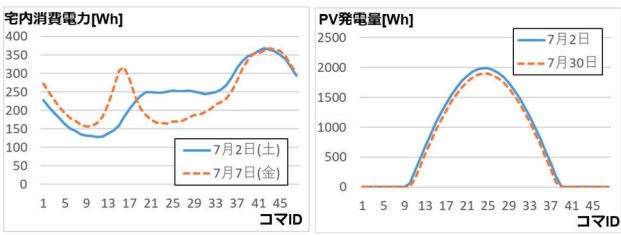


図 3：宅内消費電力データ

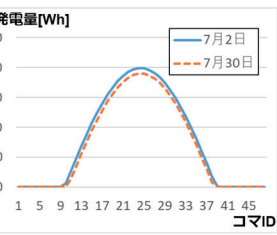


図 4：PV 発電量データ

表 1：評価条件

シミュレーション期間	7 月 1 日~7 月 31 日
世帯数	10,000
小売電気事業者からの買電単価	26 円/kWh
小売電気事業者の買取単価	5 円/kWh
PV 保有率	2%, 10%, 20%
入札戦略の比率 (利益追求：エコ志向：無関心)	10:10:80

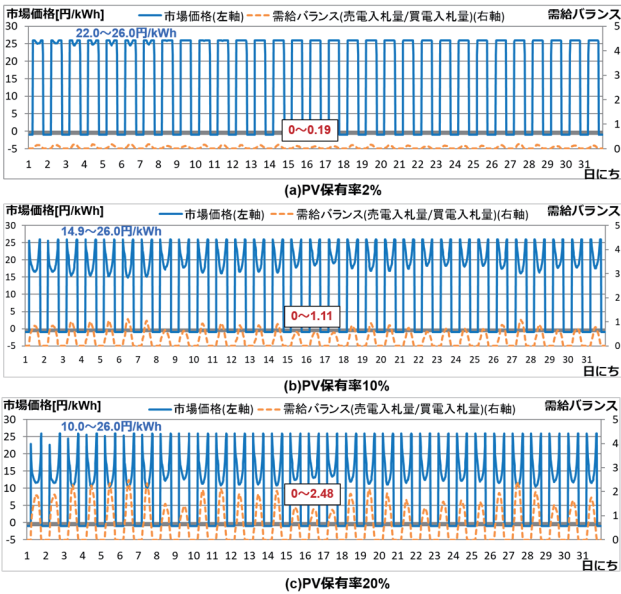


図 5：コマ毎の市場価格推移と需給バランス

入札戦略の比率は、利益追求戦略、エコ指向戦略、無関心戦略を 10:10:80 とする。PV 保有率 2%、10%、20%の 3 通りについて、需給バランスが異なる場合の市場価格推移や世帯利益の変化を見る。定置用蓄電池や EV の保有率は 0%とする。

利益追求戦略のパラメータについて、 $a = 1$ 円/kWh, $b = 1$ 円/kWh, $\alpha = 3$ 円/kWh, $\beta = 0.3$. $a' = 1$ 円/kWh, $b' = 1$ 円/kWh, $\alpha' = 3$ 円/kWh, $\beta' = 0.3$ とする。

5.5 市場価格推移と需給バランス

PV 保有率が 2%、10%、20%の場合の 1 か月間のコマ毎の市場価格推移と需給バランスを図 5(a)(b)(c)にそれぞれ示す。約定しなかったコマの市場価格は便宜上 1 円/kWh としている。PV 保有率に依らず、PV 発電が無い夜間は需給バランスが 0 で約定していない。一方、PV 発電が有る日の出 5 時 30 分頃から日の入り 18 時頃の間は、需給バランスが 0 より大きくなり、約定していることが分かる。

PV 保有率が 2%の場合(図 5(a)), 需給バランスは 0~0.19 の範囲で売電入札量は買電入札量に対して高々 1/5 と極端に少ない。

市場価格は 22.0～26.0 円/kWh と比較的高い価格帯を推移している。

PV 保有率が 10% の場合(図 5(b)), 需給バランスは 0～1.11 の範囲であり, 売電入札量が買電入札量を上回るコマが出てくる。市場価格は 14.9～26.0 円/kWh と広範囲を推移している。

PV 保有率が 20% の場合(図 5(c)), 需給バランスは 0～2.48 の範囲であり, 売電入札量が買電入札量の 2.48 倍と大きく上回るコマが有る。市場価格は 10.0～26.0 円/kWh と PV10% の時よりも広範囲を推移している。

続いて PV 保有率 20% の場合におけるコマ別の日毎の市場価格推移を図 6 に示す。日の出(コマ ID=11)や日の入り(コマ ID=36)は PV 発電量と需給バランスが共に小さいため, 20 円/kWh 以上の高い市場価格で推移している。その間の時間帯(コマ ID=18, 24, 30)は PV 発電量と需給バランスが共に大きいため, 10～15 円/kWh 程度の市場価格で推移している。

以上のように, 需給バランスが低いと市場価格は高く, 需給バランスが高いと市場価格は低価格帯を推移するなど, 需給バランスに応じて市場価格が形成されている。また, PV 保有率 20% の場合など, 1 日の中で市場価格はコマ ID に応じて 10 円～26 円/kWh のように大きく変動するが, コマ別での日毎の市場価格は安定的に推移している。同じコマで見た場合, 日毎の需給バランスは大きく変わらないことから, 需給バランスにより安定的に価格形成されていると言える。

5.6 世帯利益

売電世帯と買電世帯の入札戦略毎の利益[円/月]の箱ひげグラフを図 7, 図 8 にそれぞれ示す。PV 保有率が 2%, 10%, 20% の全の場合において, 利益追求の利益が最も大きくなっており設計通りの結果になっている。利益追求と無関心について, 利益の最大値と最小値に差異があるが, 利益追求の方が無関心と比較して差異が小さい。この点について, 利益追求は入札価格を市場価格に近づける方向に毎日変更するため, 全世帯に近い入札価格に推移しているためと考えられる。

売電世帯の利益について, PV 保有率が高くなる程小さくなっている。買電世帯の利益については PV 保有率が低くなる程大きくなっている。この点について, PV 保有率が高くなる程, 市場価格が低価格帯に推移している結果(図 5)と合致している。

5.7 約定率

売電世帯と買電世帯の入札戦略毎の約定率を図 9 に示す。PV 保有率 2%, 10%, 20% の全の場合において, エコ志向の約定率が最も大きくなっており, 設計通りの結果になっている。

売電世帯の約定率について, PV 保有率が高くなる程低くなっている。買電世帯の約定率については PV 保有率が低くなる程高くなっている。この点について, PV 保有率が高くなる程, 需給バランスが大きくなる点(図 5)と合致している。

6. まとめ

世帯の余剰電力を世帯間で入札により売買する電力分け合いサービスを対象とした電力分け合い MAS モデルを設計した。基礎評価を通して同モデルの妥当性を確認した。具体的には, 売電入札量と買電入札量の需給バランスに応じて市場価格が安定的に推移していた。また, 入札戦略毎の世帯利益や約定率について, 利益追求戦略の世帯利益が高く, エコ志向戦略の約定率が高くなるなど, 設計通りの結果が得られる点を確認した。

今後は, 入札戦略の比率を変えた場合の評価や, JEPX 市場での価格高騰などの事例を参考に, 異常状態に対する市場価格推移の安定性などを評価する。

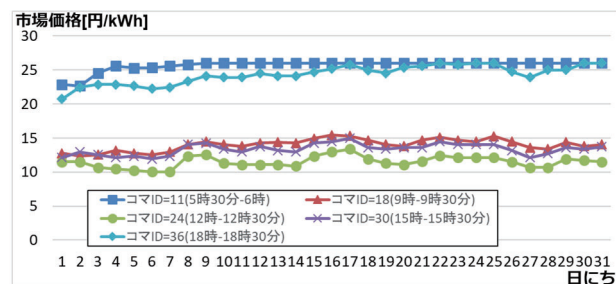


図 6: PV 保有率 20% の場合のコマ別日毎の市場価格推移

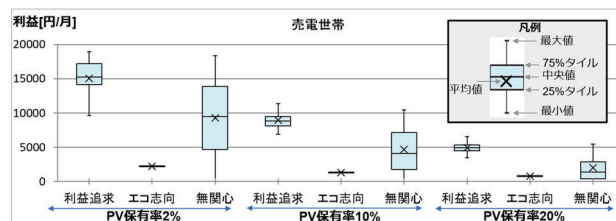


図 7: 入札戦略毎の売電世帯の利益

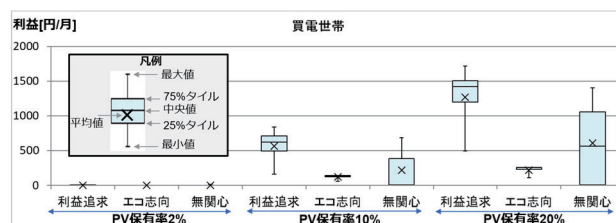


図 8: 入札戦略毎の買電世帯の利益



図 9: 入札戦略毎の約定率

参考文献

- [井上 11] 井上淳, 藤井康正. パケット電力取引に基づく革新的配電システムの提案. 電気学会論文誌 B, Vol. 131, No. 2, pp. 143-150, 2011.
- [和泉 17] 和泉潔, 斎藤 正也, 山田 健太. マルチエージェントのためのデータ解析, コロナ社, 2017.
- [渡邊 03] 渡邊勇, 岡田健司, 栗原郁夫, 永田真幸. 電力市場のシミュレーション—市場シミュレータの開発とエージェントモデルの分析. 電力中央研究所報告, R02022, 2003.
- [Ullah 13] Qudrat-Ullah, H. Energy Policy Modeling in the 21st Century. Springer, 2013.
- [Sun 07] Sun, J. and Tesfatsion, L. Dynamic testing of whole-sale power market designs: An open source agent-based framework. Computational Economics, Vol. 30, No. 3, pp. 291-327, 2007.
- [Vytelingum 10] Vytelingum, P., Voice, T. D., Ramchurn, S., Rogers, A., and Jennings, N. Agent-based Micro- Storage Management for the Smart Grid. In Proc. Autonomous agents and multiagent systems, pp. 39-46, 2010.

[3C3-J-9] Natural language processing, information retrieval: creation and analysis of stories

Chair:Hiromi Wakaki Reviewer:Masahiro Ito

Thu. Jun 6, 2019 1:50 PM - 2:50 PM Room C (4F International conference hall)

[3C3-J-9-01] Novel Segmentation Method based on the Distributed Representation of Sentences and Analysis Method of Story Developments

○Kiyohito Fukuda¹, Naoki Mori¹, Makoto Okada¹ (1. Osaka Prefecture University)

1:50 PM - 2:10 PM

[3C3-J-9-02] Analysis of Four-scene Comics Story Dataset based on natural language processing

○Ryo Iwasaki¹, Naoki Mori¹, Miki Ueno² (1. Osaka Prefecture University, 2. Toyohashi University of Technology)

2:10 PM - 2:30 PM

[3C3-J-9-03] Search for Similar Story Sentences based on Role of Characters in order to Support and Analyze Contents Creator's Ideas

○Takefumi Katsui¹, Miki ueno¹, Hitoshi Isahara¹ (1. toyohashi university of technology)

2:30 PM - 2:50 PM

文の分散表現に基づく小説の自動分割とストーリー展開の解析

Novel Segmentation Method based on the Distributed Representation of Sentences and Analysis Method of Story Developments

福田 清人^{*1} 森 直樹^{*1} 岡田 真^{*1}
Kiyohito Fukuda Naoki Mori Makoto Okada

^{*1}大阪府立大学 工学研究科
Graduate School of Engineering, Osaka Prefecture University

Recently, the attempts to reproduce the mechanisms of human intellectual activities have attracted interest in artificial intelligence fields. The narrative creation is one of them. It is necessary for narrative creation and creative support by the computer to make it to understand human creations and their stories. However, there are few studies on story analysis by the computer. In this study, we propose the segmentation method of novels based on the distributed representation of sentences and the analysis method of story developments. As a result of computational experiments, we confirmed that the effectiveness of the proposed methods.

1. はじめに

近年、人の知的活動の仕組みを計算機上で再現する試みが人工知能の分野で広く行われ、大きな注目を集めている。人の知的活動の1つに物語の創作がある。人の感性に基づく創作物である物語はストーリーと表現媒体の2つの要素から構成され、それらの組合せによって小説や漫画などに枝分かれしていく。ここで、ストーリーは物語の内容であり、表現媒体はその内容をどのような形で表現するかという表現方法である。計算機による物語の自動生成の研究 [Pérez 01][Ueno 14] は数多く報告されている。また、近年では人と機械による共同作業に関する研究も注目を集めており、人と計算機による物語の共同創作の研究 [上原 11] や創作支援の研究 [葛井 17] も報告されている。どちらの研究においても、人の創作物を計算機に理解させることが非常に重要となる。

物語の自動生成や創作支援を実現するためには、既存の物語を解析し、人が物語を創作するうえで必要な知識や技術を計算機が理解可能な形で獲得する必要がある。具体的には、機械学習技術に基づく既存の物語に対する解析による有用な情報の抽出は必要不可欠な技術である。しかしながら、物語の解析に関する研究は、専門家の経験則に基づいて人手で情報を抽出する研究 [佐藤 10] が主流であり、計算機による数値的な情報抽出や解析に関する研究はほとんど報告されていない。

なお、本研究ではストーリーに焦点を当てる。表現媒体が時代とともに姿を大きく変えることがある一方、古くから存在するストーリーの典型的な構造が現代でもしばしば使用されることから、表現媒体と比べてストーリーの方が時間経過に対してロバストなためである。また、本研究における解析対象には小説を用いる。

以上を背景として、本研究ではストーリーを計算機に理解させるための第一歩として、文意を考慮した小説の自動分割手法およびストーリー展開の解析手法を提案する。文意を考慮した文の分散表現に基づき、小説文をストーリーが展開する部分で自動分割する。また、自動分割された複数の小説文をシーンとみなし、各シーンのベクトルを用いてストーリー展開が類似した部分を発見する。

2. 関連研究

本研究と関連のあるいくつかの研究について説明する。

2.1 テキストセグメンテーションに関する研究

テキストデータをトピックなどの意味的なまとまりに分割するテキストセグメンテーションに関する代表的な手法に TextTiling [Hearst 97] がある。TextTiling はテキスト中のある2文間を基準として、その前後の文をあらかじめ設定した窓幅の分だけそれぞれ取得し、得られた前後の文章に対して単語の出現頻度ベクトルの類似度を計算する。この操作を基準となる2文を動かしながら実行し、得られた類似度の変化から文章境界を推定する手法である。TextTiling は文章内に出現する単語に基づいてセグメンテーションするため、短い文章を対象とした場合には有効に機能しないことが知られている。

2.2 物語の解析に関する研究

物語の解析に関する研究では、星新一の作品を構造分析の考えに基づき、テキストの時系列に着目して物語のパターン抽出をする研究が報告されている [佐藤 10]。しかしながら、物語のパターンを抽出するためにはテキストを抽象化して分類する必要があるため、人手によってしか解析できないという問題点が存在する。

3. 提案手法

本研究では、物語の中でもストーリーという要素に着目して小説を解析する。ここで、ストーリーをイベントや登場人物の行動、場所移動に伴う物語中の一連の状態遷移の時系列であると定義する。小説をいくつかの文章の集合であると仮定すると、ある連続した2つの文章間の差が状態遷移であり、冒頭から末尾までの連続した2文章間の差の分布がストーリーであるといえる。そこで、小説中の文章を分散表現化して文章ベクトルを得た場合、小説は文章ベクトルの時系列集合とみなすことができる。また連続する2つの文章ベクトルに何らかの演算子を適用した結果がその2文章間での差であり、物語内の状態遷移を表しているといえる。そのため、小説におけるストーリーは冒頭から末尾までの連続した2つの文章ベクトルにある演算子を適用した結果の時系列集合であると定義できる。

以上の観点から本研究では、文の分散表現に基づく小説文の自動分割手法およびストーリー展開の解析手法を提案する。な

連絡先: 福田清人, 大阪府立大学 工学研究科, 大阪府堺市中区学園町 1-1, 072-254-9273, fukuda@ss.cs.osakafu-u.ac.jp

お、文の分散表現の獲得には、これまでに提案してきた文の分散表現の獲得手法 [Fukuda] を改良した手法を用いる。

3.1 文の分散表現を用いた小説文の自動分割

TextTiling の考え方を基にして、小説文の各文の分散表現に対して類似度を計算し、類似度が極大となる 2 文を結合していく操作を、セグメント数が任意の数となるまで繰り返すことで小説を自動分割する手法を提案する。ここで 1 文単位での類似度計算をすると、機械的な文分割により分割されてしまった不適切な文の前後を分割点と推測してしまう可能性がある。そこである 1 文に対して、その 1 文と前後窓幅分を含む文の分散表現の平均を類似度計算に用いるベクトルとするスムージング手法を導入する。図 1 および図 2 に小説文のセグメンテーション手法の概要およびスムージング手法を示す。以下に文の分散表現を用いた小説文のセグメンテーション手法のアルゴリズムを示す。

1. 獲得したいセグメント数を N_s 、スムージング幅を N_w とする。ここで、本節で用いるスムージング手法では基準となる文に対してその前後の N_w 文を含む $2N_w + 1$ 文をまとめてスムージングする。
2. 解析する小説を M 文の文集合とする。
3. 小説中の各文に対して、文の分散表現の獲得手法により文の分散表現 $s_i (i = 1, 2, \dots, M)$ を獲得する。
4. 各セグメントに対応したセグメントベクトルを $d_j (j = N_w + 1, N_w + 2, \dots, M - N_w)$ 、セグメントベクトルの集合を $\mathcal{D} = \{d_{N_w+1}, d_{N_w+2}, \dots, d_{M-N_w}\}$ とする。また、各セグメントに含まれる文数を b_j 、この文数の集合を $B = \{b_{N_w+1}, b_{N_w+2}, \dots, b_{M-N_w}\}$ とする。ここで、

$$d_j = \frac{1}{2N_w + 1} \sum_{k=j-N_w}^{j+N_w} s_k \quad (1)$$

$$b_j = \begin{cases} N_w + 1 & (j = N_w + 1, M - N_w) \\ 1 & (\text{otherwise}) \end{cases} \quad (2)$$

である。

5. セグメントベクトル集合の連続した 2 つのセグメントベクトル d および d' の類似度 $f_{\text{sim}}(d, d')$ を以下の式に従って計算する。ここで、 γ は減衰率であり、セグメントが長文になりすぎないように制御するための可調整パラメータである。 γ は $0 < \gamma < 1$ を満たす実数である。また、 b および b' はそれぞれセグメントベクトルに対応したセグメントに含まれる文数である。

$$f_{\text{sim}}(d, d') = \gamma^{b+b'-2} \left(1 + \frac{d \cdot d'}{|d||d'|} \right) \quad (3)$$

6. 5 で求めた類似度が最大となった 2 つのセグメントベクトルを d_m および $d_{m'}$ とし、それぞれに対応するセグメントに含まれる文数をそれぞれ $b_m, b_{m'}$ とする。
7. d_m および $d_{m'}$ に対応するセグメントを結合し、1 つのセグメントとする。その後、以下の操作を適用することで各値を更新する。ここで、記号 ' \rightarrow ' は左式を右式で更新

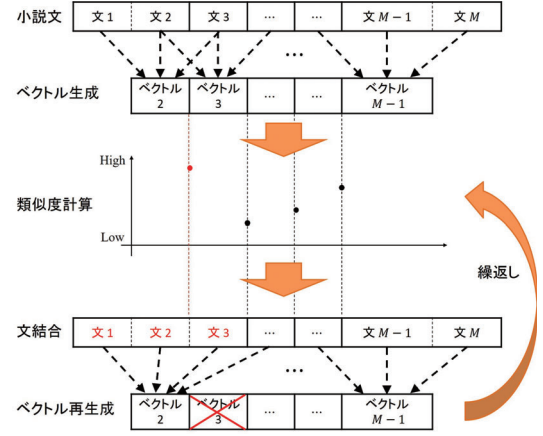


図 1: 小説文のセグメンテーション手法の概要

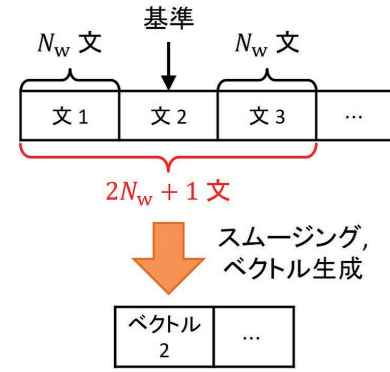


図 2: 小説文のスムージング手法の概要

する操作を表す。

$$d_m = \frac{1}{2N_w + b_m + b_{m'}} \sum_{k=m-N_w}^{m'+b_{m'}+N_w-1} s_k \quad (4)$$

$$b_m \rightarrow b_m + b_{m'} \quad (5)$$

$$B \rightarrow B \setminus \{b_{m'}\} \quad (6)$$

$$\mathcal{D} \rightarrow \mathcal{D} \setminus \{d_{m'}\} \quad (7)$$

8. $|\mathcal{D}| > N_s$ ならば、5 に戻る。

9. $|\mathcal{D}| = N_s$ の時、 \mathcal{D} および \mathcal{B} の各要素に対応したセグメントを獲得する。

3.2 文の分散表現を用いたストーリー展開の解析

3.1 節で説明した自動分割手法により得られたセグメントとそのセグメントに対応したセグメントベクトルを用いてストーリー展開を解析する。本節では、連続したセグメントに対応したベクトル間の差分を計算し、得られた差分ベクトルを 1 つのストーリー展開とみなすことでストーリーを多次元数値空間上で表現する。以下にストーリー展開の解析手順を示す。

1. 解析対象とする作品を複数用意する。
2. 3.1 節の自動分割手法により作品を自動分割し、各作品のセグメントとそれに対応したセグメントベクトルを獲得する。
3. 各作品ごとに、連続した 2 つのセグメントベクトルの差分を計算する。

表 1: 実験 1 における実験条件

m_{\max}	80
使用作品	太宰治 「走れメロス」 太宰治 「黄金風景」 芥川龍之介 「蜘蛛の糸」 芥川龍之介 「藪の中」 エドガー・アラン・ポー 「黒猫」

- 得られた差分ベクトルに対して、他作品から得られた差分ベクトルとのコサイン類似度を計算する。
- 得られたコサイン類似度や差分ベクトルの各要素の割合に基づいて、作品間でのストーリーの類似性や、差分ベクトルとストーリー展開の関係性などを可視化しつつ解析する。

4. 実験

提案手法の有効性を確認するため、いくつかの実験をした。また、提案手法を用いていくつかのストーリーを実際に解析することで、得られる情報の特徴や傾向についての知見を得る。実験 1 ～ 3 まで実施したが、紙面の関係上、実験 1 のみ示す。実験 2 および実験 3 については発表時に述べる。

4.1 実験 1

文の意味的な類似性を考慮した文の分散表現を用いて、小説のストーリー展開を可視化することができるかを確認する。

小説の各文に対して文の分散表現を獲得し、次元圧縮手法である t-distributed Stochastic Neighbor Embedding (t-SNE) を用いて 2 次元空間に写像することで可視化する。

4.2 実験 1 の実験条件

表 1 に実験 1 の実験条件を示す。実験 1 では青空文庫から取得した各作品を文単位に分割し、その中から m_{\max} 単語以下の文を用いた。

4.3 実験 1 の結果と考察

図 3 ～ 図 7 に各作品の t-SNE による可視化の結果を示す。ここで、各図において独立に t-SNE を適用しているため、それぞれの図の軸に関係性はないことに注意する。また、t-SNE はデータ間の距離を確率分布で表現することで次元を圧縮するため、次元圧縮後の各軸には意味がないことにも注意を要する。

図 3 を見ると、各文の分散表現が 2 つの分布パターンに属していることがわかる。これは、「走れメロス」という作品は町での王様との会話シーンと、村に戻ってまた町に戻るという移動シーンの 2 つに分けることができることから妥当であると考えられる。

図 4 および図 5 を見ると、すべての文の分散表現が次元圧縮後の 2 次元空間において、ある直線上に分布していることがわかる。このような分布をとる原因として、「黄金風景」および「蜘蛛の糸」は場面転換時の風景描写などが空白であることや、作品全体を通して使用される文体や人物の口調なども一定であることが考えられる。

図 6 を見ると、「藪の中」の各文の分散表現は先頭から末尾に向かって一定方向に展開していることがわかる。このことから、「藪の中」は物語の後半部分で前半部分について回想することなく、ストーリーが徐々に展開していくと考えられる。

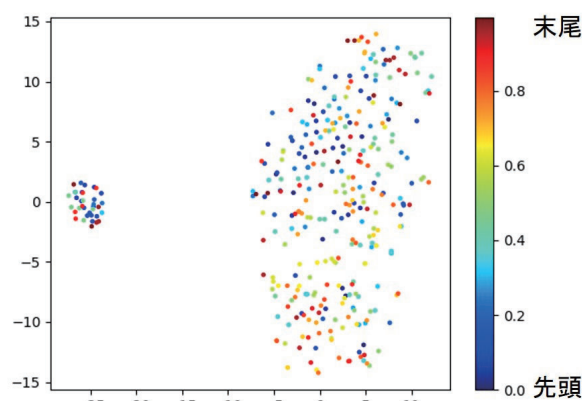


図 3: t-SNE による「走れメロス」の可視化結果

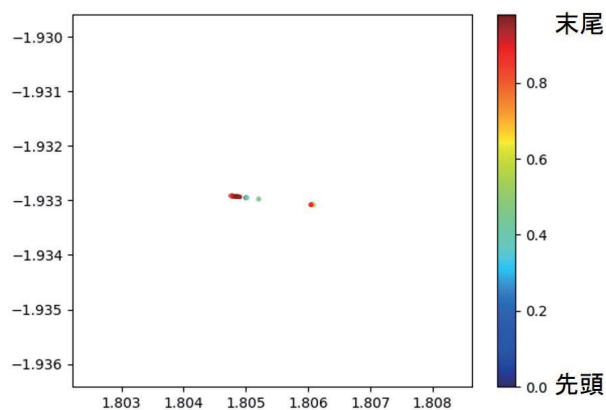


図 4: t-SNE による「黄金風景」の可視化結果

実際に、「藪の中」という作品はある男が殺される事件に対して、目撃者や容疑者、被害者本人の霊などが事件について語るというストーリーであり、それぞれが自身の体験を語るだけで、各人物が関わりあうシーンが存在しない。

図 7 を見ると、文の分散表現が一様に分布していることがわかる。また、「黒猫」という作品は主人公の身の回りで起きる事件と主人公の内面が交互に書かれている作品であり、類似したストーリー展開を作品内で何度も繰り返す内容になっている。これらのことから、似たストーリー展開を繰り返すことで、時系列という観点から文の分散表現を可視化した結果、一様に分布してしまっていると考えられる。

以上の結果から、小説のストーリーやその展開に関して、文の分散表現を用いた多次元数値空間上で作品の種類や特徴を解析可能であると考えられる。

5. むすび

本研究では、文の意味的な類似性を考慮した文の分散表現に基づく小説の自動分割手法とストーリー展開の解析手法を提案した。自動分割手法では TextTiling の考え方を基にして、文の分散表現間の類似度が極大な部分を結合することで小説をシーンごとに自動分割した。ストーリー展開の解析手法では、

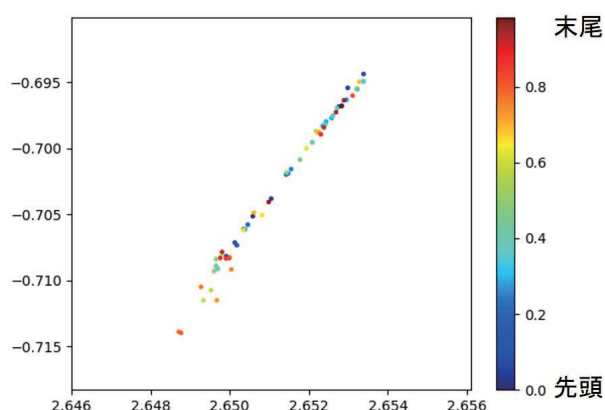


図 5: t-SNE による「蜘蛛の糸」の可視化結果

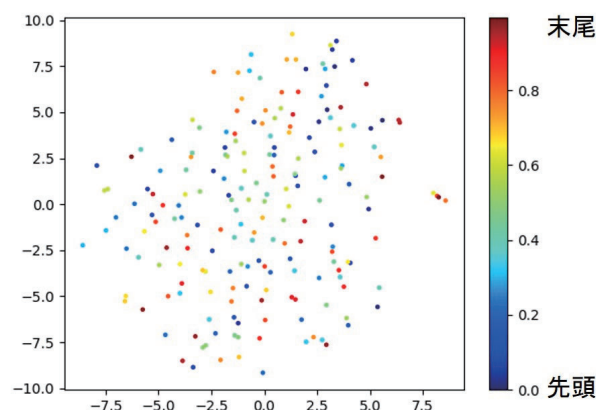


図 7: t-SNE による「黒猫」の可視化結果

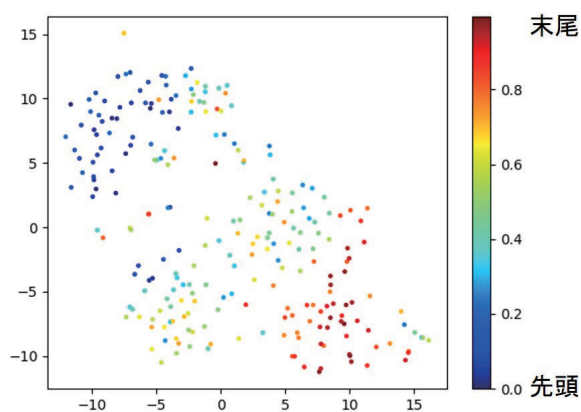


図 6: t-SNE による「藪の中」の可視化結果

自動分割手法で得られたセグメント間の差分をストーリー展開とみなし、差分ベクトルの類似度からストーリー展開の類似性を解析した。提案手法の有効性を確認するためのいくつかの実験により、以下の知見が得られた。

- 小説文の分散表現を次元圧縮手法である t-SNE によって可視化することで、小説のストーリーやその展開について多次元数値空間上で解析することができる。
- 提案手法を用いることで、人手によるアノテートに頼ることなく小説文を意味を考慮したシーン単位に自動分割することができる。
- 提案手法により、ストーリー展開の類似性だけでなく、文章構成の類似性も取得することができる。

今後の課題として提案手法の根幹となる文の分散表現の性能向上は最重要課題である。階層的 LSTM や Attention 機構のような自然言語処理における有効性が示された技術を手法に導入することで更なる性能向上が期待される。また、自動分割手法におけるスムージングの窓幅やシーンの分割数など解析対象の作品ごとに異なる可調整パラメータを最適化するための手法について検討する必要がある。

謝辞

本研究は一部、日本学術振興会科学研究補助金基盤研究 (C) (課題番号 26330282) の補助を得て行われたものである。

参考文献

- [Fukuda] Fukuda, K., Mori, N., and Matsumoto, K.: A Novel Sentence Vector Generation Method Based on Autoencoder and Bi-directional LSTM, in Prieta, de la F., Omatu, S., and Fernández-Caballero, A. eds., *Distributed Computing and Artificial Intelligence, 15th International Conference, DCAI 2018, Toledo, Spain, 20-22 June 2018*, Vol. 800 of *Advances in Intelligent Systems and Computing*, pp. 128–135
- [Hearst 97] Hearst, M. A.: TextTiling: Segmenting Text into Multi-paragraph Subtopic Passages, *Comput. Linguist.*, Vol. 23, No. 1, pp. 33–64 (1997)
- [Pérez 01] Pérez, y R. P. and Sharples, M.: MEXICA: A computer model of a cognitive account of creative writing, *Journal of Experimental & Theoretical Artificial Intelligence*, Vol. 13, No. 2, pp. 119–139 (2001)
- [Ueno 14] Ueno, M., Mori, N., and Matsumoto, K.: *2-Scene Comic Creating System Based on the Distribution of Picture State Transition*, pp. 459–467, Springer International Publishing, Cham (2014)
- [葛井 17] 葛井 健文, 上野 未貴, 井佐原 均: 質問集合とグラフに基づく物語全体の流れを管理可能な創作支援システムの提案, 第 31 回人工知能学会全国大会発表論文集 (2017)
- [佐藤 10] 佐藤 知恵, 村井 源, 往住 彰文: 星新一ショートショート文学の物語パターン抽出, *情報知識学会誌*, Vol. 20, No. 2, pp. 123–128 (2010)
- [上原 11] 上原 大輝, 出水 ちあき, 宮里 洸司, 神里 志穂子, 野口 健太郎: J-030 子どもの思考プロセス把握における物語自作システムの有効性検証 (HCS(2), J 分野: ヒューマンコミュニケーション&インタラクション), *情報科学技術フォーラム講演論文集*, Vol. 10, No. 3, pp. 597–600 (2011)

自然言語処理に基づく 4 コマ漫画ストーリーデータセットの解析

Analysis of Four-scene Comics Story Dataset based on Natural Language Processing

岩崎 凌^{*1} 森 直樹^{*1} 上野 未貴^{*2}

Ryo Iwasaki

Naoki Mori

Miki Ueno

^{*1}大阪府立大学

Osaka Prefecture University

^{*2}豊橋技術科学大学

Toyohashi University of Technology

Comic computing is a branch of computing dealing with comics in engineering. Although comics are multi-modal data with natural languages and pictures, numerous studies in the field focus on images in comics rather than the comic content. To make models understand the contents in comics, we should deal with natural languages in comics in the form of character words. We used a dataset which was suitable for analyzing comic contents and previously proposed two tasks: sentiment analysis and variety analysis. However, two tasks did not go well because of the number of data. We demonstrated using data augmentation and analyzed the results to determine the feasibility of computers understanding comics.

1. はじめに

近年、人工知能による小説や漫画、アニメ、漫画といった創作物を対象とした研究が大きな関心を集めている。創作物理解や自動生成といった試みは工学的興味深く意義の大きい反面、そもそも人の創作物理解は高次の知的活動であり、こういったタスクであれば計算機が創作物を理解したといえるのかを定義することさえ現状では難しい。

人の創作物の中で、特に漫画を工学的に扱う研究分野が近年、コミック工学を中心に発展している。漫画を対象とする研究では自然言語と画像を持つマルチモーダルなデータを扱うため、人工知能研究の対象として適している。この分野では、現在でも数多くの研究が報告されているが、多くは漫画の持つ画像を対象としており、ストーリーといった漫画の意味を自然言語から解析しようとする研究は少ないのが現状である。その一因としては、上で述べたような意味理解のためにどのようなタスクを設定すればよいかが明確ではないという点が挙げられる。このため、表層的な情報だけではなく、ストーリーにまで踏み込んだ研究は十分にされていない。そのような困難にもかかわらず、創作物を理解できる人工知能の構築への情動は尽きることがない。

このような研究を実現するためには、まずデータセットが必要となるが、人工知能と創作に関する研究分野では著作権などの問題からデータセットを準備することが容易ではない場合がある。しかしながら既存のデータセットには、研究に必要な情報などのラベルを持っていないという問題点があった。この点を解決し、創作者視点を積極的に取り入れたデータセットとして 4 コマ漫画ストーリーデータセット [Ueno 18] が提案されている。ただし、このデータセットはまだ規模が十分ではないといった別の問題が存在する。そこで、本稿では Data Augmentation を利用することで、4 コマ漫画ストーリーデータセットのデータ数の問題点を軽減し、これまで実現できなかった新しい実験手法について検討する。

これらを踏まえ、本実験では人工知能による意味理解のために感情分析と多様度分析の 2 つのタスクを用意し、Data Augmentation による結果を比較する。本実験はデータ量の関

係から十分な結果は得られていないが、今後の人工知能による創作物理解の可能性を示すという立場で結果を解析する。2 章では、参考研究を示し、3 章で本実験で扱うデータセットについて説明する。4 章で実験手法や条件について説明し、5 章でそれらの考察を述べる。

2. 関連研究

1 章で述べたように、漫画を対象とした研究分野では画像認識が用いられることが多い。Rigaud らの研究では漫画の吹き出しに焦点を当て、吹き出し内のテキストを自動で抽出しようと試みた [Rigaud 15, Rigaud 16]。松井らの研究では、人の描いたスケッチから似た絵を持つ漫画内の絵を自動で抽出するシステムを提案した [Matsui 17]。藤野らは 4 コマ漫画の構造理解のために 4 コマ漫画の読解順序識別をした [Fujino 18]。これらのすべてが漫画内の画像のみ、あるいは絵に密接に関係したラベルを用いた研究である。

漫画の自動生成に関する研究も報告されている [Ueno 14] [Fukuda 17]。上述のような例と異なり、これらの研究は画像だけでなく自然言語も対象としているが、これらの研究の目的は漫画の解析という我々の目的とは異なっており、自然言語から漫画を解析しようとする我々の研究とは異なる立ち位置にある。現在は漫画の完全な自動生成というのは難しいことを踏まえると、我々が試みるような解析が今後の人工知能による創作研究の発展に非常に重要なものであることがわかる。

最後に漫画に関するデータセットについて述べる。漫画を対象とした工学研究のためのデータセットはいくつか存在する。日本の漫画を対象としたデータセットとしては Manga109 [Matsui 17, Ogawa 18] が挙げられる。Manga109 は 109 冊の既存の漫画を含んでおり、漫画のテキストが文字に起こされていたり、漫画内のキャラクターや吹き出しについてのアノテーションがなされていたりする。工学的な漫画研究ではよく用いられるデータセットであるが、自然言語から漫画を解析する場合には漫画のストーリーに関するラベルが少ないことから Manga109 の持つラベルでは不十分であると考えられ、本稿では Manga109 は用いず、4 コマ漫画ストーリーデータセットを用いることとする [Ueno 18]。データセットの詳細は次章で述べる。

連絡先: 岩崎 凌, 大阪府立大学 工学研究科,
大阪府堺市中区学園町 1-1, iwasaki@ss.cs.osakafu-u.ac.jp



図 1: 4 コマ漫画ストーリーデータセットのデータの一例

(c) 作画: 鈴木市規 (シナリオ: (株) スポマ 播村早紀/豊橋技術科学大学 上野未貴)

(c) 作画: 浦田カズヒロ (シナリオ: (株) スポマ 播村早紀/豊橋技術科学大学 上野未貴)

3. 研究用データセット

漫画研究用のデータセットがいくつか存在することは 2 章で述べた通りであるが、漫画をデータとして使用する際は著作権の問題から既存のデータセットを使用した方が良い。我々が研究に用いる 4 コマ漫画ストーリーデータセットは人工知能による創作研究の発展のために研究者が一から開発に関わった世界初の研究用のデータセットである。以下に我々が研究でこのデータセットを使う目的およびデータセットの特徴を述べる。

特徴の一つ目として、ストーリー解析に適した情報を多く含んでいる点にある。市販された漫画をデータとした場合、著作権などの問題に加え、計算機上で扱うためのデータが少なく、漫画の意味理解を目的とした研究には適さないという問題がある。例えば漫画に登場するキャラクターの感情は明示されていないため、読者によるアノテートによってラベルを付与する必要があるが、アノテートされたラベルが漫画家の意図とは異なる可能性を否定できない。4 コマストーリーデータセットでは研究に関わっているため、そのような問題点を解決することができる。

二つ目として、上野は 4 コマ漫画の構造を、

- 一般：標準的な起承転結をもつ
- 繰り返し：1, 2 コマ間の類似が 3, 4 コマ間でも起きる
- 出オチ：1 コマ目におかしな絵が描かれてオチがある
- タイトルオチ：最後にタイトルを見返してオチがわかる
- 再帰：4 コマ目から 1 コマ目に戻り話として成立する
- 参照：1 つ以上前の話の続きの話となる
- 連続：連続した 4 コマを 2 話並べて 8 コマで話となる

と定義し、これに従ってデータセットを作成している点が挙げられる。現在は、同一のストーリーを 4 コマ目がオチとなる

表 1: データ数

実験	ラベル	データ数	拡張後データ数
感情分析	驚愕	18	2252
	その他	147	19165
多様度分析	多様度あり	674	482062
	多様度なし	766	519826

一般と出オチの 2 つの構造から描いたものがデータとして存在しており、本稿では計算機でこういった構造の把握も目的としている。

最後の特徴としては、Manga109 のような市販漫画によって構成されたデータセットとは異なり、4 コマストーリーデータセットのデータは本データセットのために 5 人の漫画家によって描き下ろされている点である。ストーリーの解析をする際にオリジナリティの観点から同一プロットを複数の漫画家が描くことは稀有なため、そういったデータの収集に基づく研究は困難であるため、これは非常に貴重な特徴である。なお 5 人の漫画家の描く絵はそれぞれ絵のタッチが異なるため、少年漫画タッチ、少女漫画タッチ、青年漫画タッチ、ギャグ漫画タッチ、萌え漫画タッチで描かれたデータがあると言い換えることができる。この特徴は画像解析の方向からだけでなく、自然言語から解析する場合にも非常に役立つものである。作家がテキストから絵を起こすにあたってもとにしたテキストデータは同一のものであるが、このテキストデータは各作家が自分の思うようにセリフや感情ラベルを変更してもよかったため、作家ごとに差が生まれている。これによって作家ごとの考えの違いなども把握することができる。

4. 実験

本実験では、感情分析と多様度分析を行う。表 1 に各実験で用いるデータ数を示す。各実験で使用するラベルに関する説明は以下で述べる。

4.1 Data Augmentation について

本実験では日本語 WordNet [Bond 11] のシソーラスを用いてデータを拡張する手法を用いる。データセットに含まれる状況文、セリフを Juman++ [Morita 15] を用いて分かち書きをし、日本語 WordNet で類似語を持つ単語を類似語に置き換えた。文の中に類似語を持つ単語が複数あっても、類似語に置き換える単語は同時に 1 つまでとした。例えば、5 つの単語から構成される文章があり、各単語が 5 つの類似語を持っていた場合、その文からは新しく 25 文が生成されることとなる。

4.2 感情分析

本実験で使用するデータセットは全 7 種類の感情ラベル（ニュートラル、驚愕、喜楽、恐怖、悲哀、憤怒、嫌悪）を持っているが、データ数と解析の難しさの問題から今回はニュートラルを抜き、感情ラベルを驚愕とそれ以外（喜楽、恐怖、悲哀、憤怒、嫌悪）の 2 つとする。入力には感情ラベルが付与されている状況文あるいはセリフ一文を用い、上記の感情ラベルを識別する。図 3 に感情分析で使用する self-attention モデル [Lin 17] を示す。

4.3 多様度分析

多様度分析では、漫画家がある状況を絵に起こす際に多様度があるかどうかを識別する。ここで多様度ラベルはデータセットに含まれておらず、我々が独自でアノテートしたものである

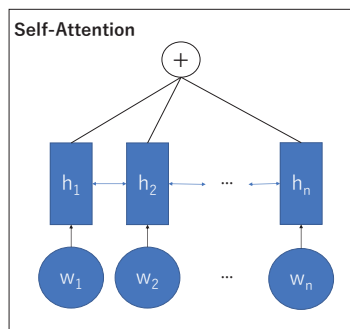


図 2: self-attention モデル

ことに注意する。本実験で使用するデータセットに含まれる複数作家の描いた漫画を用い、2つのタッチの異なる同じエピソードで同じコマの絵のみを比較することでアノテートした。アノテート時には、

1. キャラクターの数と登場キャラクター
2. キャラクターの描き方
3. 背景の有無

に着目した。

1. が異なるタッチ同士は 2. や 3. の違いにかかわらず「多様度あり」とし、1. が等しい場合、「多様度なし」とした。そのコマにキャラクターの頭部や手などの一部分のみしか登場していない場合、キャラクターとみなさなかった。

1. が等しく、2. が異なる場合は「すこしの多様度あり」とした。キャラクターの描き方とはコマに登場するキャラクターの表現の仕方である。漫画において、登場人物を描く方法は様々である。キャラクターをそのまま描くこともできれば、コマ内の吹き出しに小さく顔を描くことで、登場人物はコマ内に登場していないがその吹き出しの横に小さくキャラクターを描くことで話者を明示する手法もある。このようなキャラクターの描き方の違いが上記の 2. の一例である。

1. が等しく、3. が異なる場合は「すこしの多様度あり」とした。ここでいう背景には、キャラクターのいる場所を表すような背景に加え、その話に関係してくるアイテム（例えば、弁当に関する話の際は弁当箱）なども含まれる。

実験時、入力には話者ラベル、感情ラベル、状況文、セリフとタッチラベルを使用した。2つのタッチのデータを比較し、多様度のないものか少しでも多様度のあるものかの2クラスの識別をするため、話者ラベル、感情ラベル、状況文、セリフは各タッチごとに入力した。話者ラベルは、ラベルなし、メインキャラクター A、メインキャラクター B、その他の4次元の one-hot ベクトルであり、感情ラベルは、ラベルなしか前述した7つの感情の8次元の one-hot ベクトルである。なお入力に使うデータはもともデータセットに含まれているものである。図 4 に実験に使用したモデルを示す。

5. 実験結果

表 5. に実験結果を示す。表 5. より、どちらの実験でも訓練時の Accuracy が 1 に近づいていることから、訓練データにおける数の少ないラベルに対しても学習ができていることがわかる。特に感情分析ではデータの偏りが大きいことか

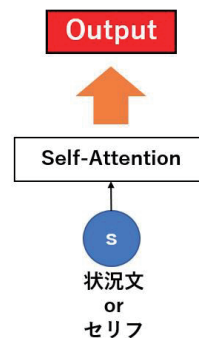


図 3: 感情分析に用いるモデル (self-attention は図 2 参照)

ら、Data Augmentation がない状態ではデータ数の多いラベルだけを出力するような識別機となっていることが確認できたが、Data Augmentation によってそういったデータに対しても特徴を学習できるようになっていることがわかる。多様度識別では Data Augmentation によって結果が良くなっていることがわかる。一般的に、Data Augmentation はロバスト性を高めるという目的に使用されることが多く、必ずしもモデルのパフォーマンスを改善するものではない。しかしながら本実験で用いた Data Augmentation 前のデータの数是非常に少なく、適切な特徴を学習するには心もとないデータ数であった。全データを通して一度しか出てこない単語が多かったために、各データに対して非常に過学習のしやすい状態であったのが、WordNet による Data Augmentation によって本来テストデータにしか出現しないような単語も学習できるようになり、性能が向上したものと考えられる。

6. まとめ

4 コマ漫画ストーリーデータセットを用いた実験の例として感情分析と多様度分析の2つを示した。Data Augmentation を用い、感情分析と多様度分析での結果を比較した。Data Augmentation により結果の改善が見られた。本実験ではシソーラスを用いることでデータを拡張したが、Data Augmentation の方法は他にも考えられ、そちらも比較検討する。

謝辞

本研究は一部、日本学術振興会科学研究補助金基盤研究 (C) (課題番号 26330282) の補助を得て行われたものである。

本研究は一部、JST, ACT-I (グラント番号:JPMJPR17U4) の支援を受けたものである。

4 コマ漫画ストーリーデータセットの制作に対し、ご協力いただいた漫画家の方々、株式会社スポマへ謝意を示します。

参考文献

- [Bond 11] Bond, F.: Japanese SemCor : A Sense-tagged Corpus of Japanese (2011)
- [Fujino 18] Fujino, S., Mori, N., and Matsumoto, K.: Recognizing the Order of Four-Scene Comics by Evolutionary Deep Learning, in *Distributed Computing and Artificial Intelligence, 15th International Conference, DCAI*

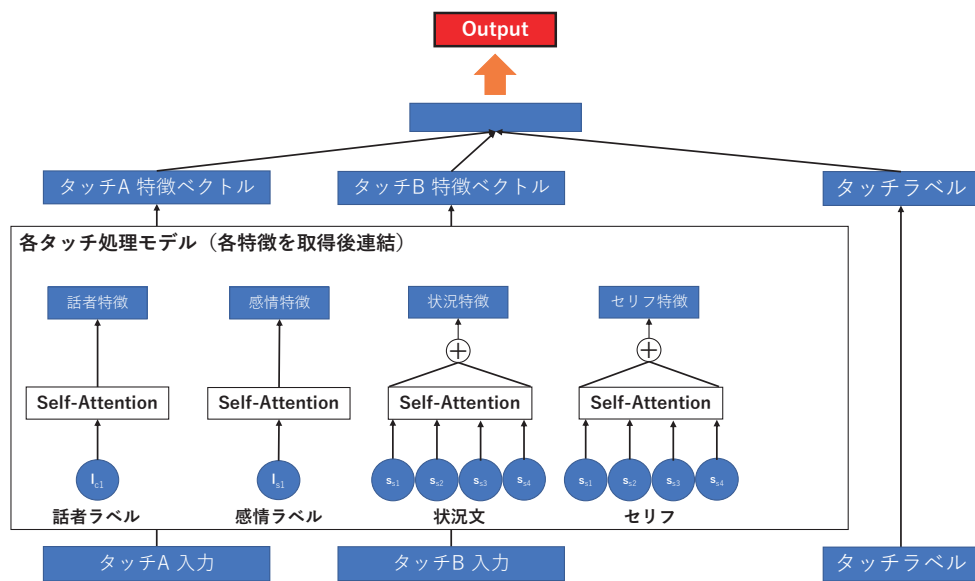


図 4: 多様度識別に用いるモデル (self-attention は図 2 参照)

表 2: 実験結果 (拡張時とついているものが Data Augmentation の結果)

実験	Train Accuracy	Test Accuracy	拡張時 Train Accuracy	拡張時 Test Accuracy
感情分析	0.8923	0.8941	0.9999	0.9000
多様度分析	0.6886	0.5181	0.9873	0.5475

2018, Toledo, Spain, 20-22 June 2018., pp. 136–144 (2018)

[Fukuda 17] Fukuda, K., Fujino, S., Mori, N., and Matsumoto, K.: Semi-automatic Picture Book Generation Based on Story Model and Agent-Based Simulation, in Leu, G., Singh, H. K., and Elsayed, S. eds., *Intelligent and Evolutionary Systems*, pp. 117–132, Cham (2017), Springer International Publishing

[Lin 17] Lin, Z., Feng, M., Santos, dos C. N., Yu, M., Xiang, B., Zhou, B., and Bengio, Y.: A Structured Self-attentive Sentence Embedding (2017)

[Matsui 17] Matsui, Y., Ito, K., Aramaki, Y., Fujimoto, A., Ogawa, T., Yamasaki, T., and Aizawa, K.: Sketch-based Manga Retrieval using Manga109 Dataset, *Multimedia Tools and Applications*, Vol. 76, No. 20, pp. 21811–21838 (2017)

[Morita 15] Morita, H., Kawahara, D., and Kurohashi, S.: Morphological Analysis for Unsegmented Languages using Recurrent Neural Network Language Model, in *EMNLP* (2015)

[Ogawa 18] Ogawa, T., Otsubo, A., Narita, R., Matsui, Y., Yamasaki, T., and Aizawa, K.: Object Detection for Comics using Manga109 Annotations, *CoRR*, Vol. abs/1803.08670, (2018)

[Rigaud 15] Rigaud, C., Thanh, N. L., Burie, J., Ogier, J., Iwata, M., Imazu, E., and Kise, K.: Speech balloon and speaker association for comics and manga understanding, in *ICDAR*, pp. 351–355, IEEE Computer Society (2015)

[Rigaud 16] Rigaud, C., Pal, S., Burie, J.-C., and Ogier, J.-M.: Toward Speech Text Recognition for Comic Books, in *Proceedings of the 1st International Workshop on coMics ANalysis, Processing and Understanding*, MANPU ’16, pp. 8:1–8:6, New York, NY, USA (2016), ACM

[Ueno 14] Ueno, M., Mori, N., and Matsumoto, K.: 2-Scene Comic Creating System Based on the Distribution of Picture State Transition, in Omatu, S., Bersini, H., Corchado, J. M., Rodríguez, S., Pawlewski, P., and Bucciarelli, E. eds., *Distributed Computing and Artificial Intelligence, 11th International Conference*, pp. 459–467, Cham (2014), Springer International Publishing

[Ueno 18] Ueno, M.: Four-Scene Comic Story Dataset for Softwares on Creative Process, in *New Trends in Intelligent Software Methodologies, Tools and Techniques - Proceedings of the 17th International Conference SoMeT-18, Granada, Spain, 26-28 September 2018*, pp. 48–56 (2018)

作中役割を考慮した物語文の類似文検索による 創作者の支援と感性の解析

Search for Similar Story Sentences based on Role of Characters
in order to Support and Analyze Contents Creator's Ideas

葛井 健文 上野 未貴 井佐原 均
Takefumi Katsui Miki Ueno Hitoshi Isahara

豊橋技術科学大学
Toyohashi University of Technology

A process of creating stories has been studied from various aspects. Recently, it has become important topic for the field of artificial intelligence. In the field, such studies are mainly divided into two groups, automatic generation of stories and assistance for human's creative activity. From the view of assistance for human's creative activity, we propose creative support method. This paper describes two systems; Creation support system and System of Search for Similar Story Sentences based on Role of Characters. In order to confirm the effectiveness of these proposed method, three types of experiments were carried out. These results suggested that the proposed system is useful for creating story and analyzing human's idea of the perspective of similar feeling.

1. はじめに

近年、文章解析として分散表現を用いた文のベクトル化の手法が数多く提案されており、機械翻訳や対話システムの応答文選択の精度が大幅に向上している。一方、物語文章の解析については作者固有の文体や作品の固有表現、セリフや回想文などが存在し、また評価指標も多岐にわたるため、文章の物語としての類似度の推定は困難である。そのため、物語の類似度を推定するための取り組みはとても重要であり、工学的な物語の自動生成と創作支援の両面において貢献できると考えられる。

本研究では、従来研究で提案された2種類のテンプレートの連携により物語のプロットを作成する創作支援システムと、新たに提案する類似した物語のあらすじを検索するシステムの連携により、より効率的に物語を作成できる創作支援手法を提案する。

2. 従来手法

2.1 創作支援システム

本研究で特に重要な創作支援システムに関する従来研究を示す。従来研究として、川越らは2種類のテンプレートを連携した創作支援システムを提案 [kawagoe 15] し、著者らはこの創作支援手法の2種類のテンプレート間の連携を強化する想定感情線機能を追加した手法を提案した [葛井 17]。本稿では、上記の手法に基づく創作支援システムを従来手法と呼ぶ。

2.2 創作支援システムの概要

創作支援システムは3つの機能を持ち、ストーリーテンプレートを用いて物語のあらすじを作成するストーリー作成部、キャラクターテンプレートを用いて登場人物に関する設定を作成するキャラクター作成部、実際に作成されたプロットを用いて本文を入力するための本文入力部によって構成される。ユーザはこのシステムを用いることで物語のプロットを完成させることが出来る。

2.3 ストーリーテンプレート

物語のプロットを作成するための質問集合がストーリーテンプレートである。また、ストーリーテンプレート中の質問項目

には三幕構成 [安藤 09] が用いられている。これにより、作成したプロットの物語中の時系列を整理出来る。

2.4 キャラクターテンプレート

登場人物を作成するための質問集合がキャラクターテンプレートである。キャラクターテンプレート内で作成した登場人物をグレイマスの行為者モデル [樋口 89] に従って4つの役割のうちいずれか一つに当てはめる。グレイマスの行為者モデルは物語の登場人物を6つの役割に整理したモデルであり、従来手法では、このうち主体、対象、援助者、敵対者を重要な役割であると考え、ユーザは作成したキャラクターを4つのうちどれか一つの役割に当てはめる。

2.5 想定感情線グラフ

2種類のテンプレート間の連携を強化するため、想定感情線グラフを導入した。想定感情線グラフは以下の機能を持つ。

- 想定感情線の手動入力
- キャラクター登場場面の表示
- 各期間の感情値の自動算出

2.6 従来研究の実験結果と提起

従来研究について得られた意見、課題点を以下に示す。

- 創作の参考になるアイデアをシステムから提示してほしい
- 両テンプレートともにバリエーションが少ない
- 物語の方向性が固定されやすい

本研究ではこれらの問題点を解決するため、保存機能の追加と、システム側から類似度に基づき参考となる作品を提示することで、作成できる物語のバリエーションの増加を目指す。

3. 提案手法

本章では、2.6で示した従来手法の問題点を解決し、より有用な創作支援手法となるように創作支援システムをあらすじ文章の内容から似た内容のあらすじを検索するシステム（以下、あらすじ検索システムとする）と連携して利用する手法を提案する。類似した物語を提示することは、創作支援システムのユーザにとって自身の参考になり、物語のバリエーションや方向性も新たな発想が得られると考えたためである。

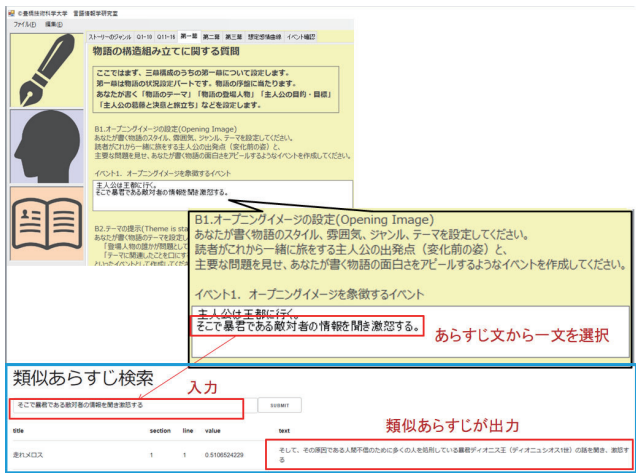


図 1: 創作支援システムに基づくあらすじ検索システムの利用

表 1: 実験に用いた PV-DM モデルの学習条件

文書サイズ	次元数	窓サイズ	エポック数
1MB	300	5	1000

3.1 あらすじ検索システムの概要

あらすじ検索システムは、物語のあらすじについて記した一文を入力として受け取り、類似したあらすじ文を出力するシステムである。図 1 に概要図を示す。

3.2 文章のベクトル化

文章の類似度などを計算する方法として文章のベクトル化が挙げられる。中でもニューラルネットワークによる学習による分散表現獲得手法として Paragraph Vector [Le 14] が存在する。Paragraph Vector は Word2Vec [Mikolov 13] から派生した手法で、本実験では PV-DM(Distributed Memory Model of Paragraph Vectors) モデルを用いる。表 1 にパラメータを示す。本稿では次に述べるコーパスを学習して PV-DM モデルをあらすじ文章モデルと呼ぶ。

3.3 コーパス作成と前処理

用いたコーパスは、類似していると判断された作品をユーザが読むことで創作への参考にする必要があるために、青空文庫収録作品のうち 2018 年 10 月 4 日時点で日本語 Wikipedia [Wikipedia] に記事があり、その中に「あらすじ」節を持つ 464 作品を対象とした。また、抽出の際に注釈やリンクの情報は消去している。本研究では、Wikipedia から抽出した各記事の文章から、「あらすじ」節に含まれる文章のみを利用し、文章コーパスとする。コーパス内の各文章には「どの作品のあらすじ文章か」を判別するためのラベル付けを行った。また、文書解析の精度向上および実際に創作に役立てるための情報に用いるため、コーパスに以下の 3 つの前処理をした。

3.3.1 物語に関係しない文章の削除

取得したあらすじ節の文章の中には、その物語の中の出来事の情報ではない、作品そのものへの解説や世界観の注釈、といった文章が含まれていることがある。本研究では創作支援システムに入力された物語のあらすじ文章と Wikipedia のあらすじ文章を比較し、二つの文章の類似度から物語の類似度を推定することが目的であるため、以下に示す内容を含む文章は今回のコーパス作成に不向きであるとして以下の操作をした。以下に操作内容を示す。

表 2: 三幕構成への分割を行ったあらすじ文章の内訳

内訳	第一幕	第二幕	第三幕
各幕を持つ作品数	462	427	357
各幕に含まれる文数	2542	4175	2354
各幕に含まれる単語数	56850	103612	50900

- 物語の歴史的背景や設定を補足する文章
- 複数言語での表記や、表現の違いを補足する文章
- 記事内で対象としている版以外の物語の展開へ言及している文章
- 作品の解説や制作事情など、物語の出来事以外に言及している文章

以上の操作の結果、あらすじ文が全て削除された作品が 2 作品あったため、対象作品の数は実質的に 462 作品となった。

3.3.2 物語の時系列の整理

Paragraph Vector を用いた文章の分散表現を用いて文書間の類似度を計算する際、「その文章はあらすじ文章中のどの位置にあった文章か、物語中のどのタイミングの出来事を表した文章か」という情報は考慮されない。しかし、物語においては「物語中で出来事が起きたタイミング」は非常に重要であり、表記の上では同じような文章でも、そのタイミングによって大きく意味合いが異なり、結果的に物語的な類似度は小さくなるのが予想される。そのため、コーパス中の各文章に対し、「どの作品のあらすじか」だけでなく、その文章は「創作支援システムのストーリーテンプレートで定義された三幕構成のうち、作品中で第一幕から第三幕のどのタイミングの出来事に当たる文章か」を判別するためのラベル付けをした。ラベル付けの方法として、あらすじ文章を文書量等を基準として均等に三分割するのではなく、各作品のあらすじ文章ごとにどのように三幕構成に分割できるかを人手で判断し、分類を行った。表 2 に各幕ごとの内訳を示す。

3.3.3 登場人物名の役割名への置換

あらすじ文章には、作品ごとに固有の登場人物の名前が含まれており、Paragraph Vector を用いて文章をベクトル化する際、固有表現はそのベクトルに強く作用する。一方で物語においては、「誰」が何をしたかは非常に重要であり、例えば戦いに勝利したのが主人公であるのか敵役であるのかは意味合いが大きく変わる。そこで、物語的に類似した文章を検索するために、登場人物名に関する固有表現と、その人物を表す代名詞を、創作支援システムのキャラクターテンプレートで定義された 4 つの役割名：主人公、対象、援助者、敵対者に置換する。登場した人物の各役割への当てはめは人手で行った。詳しいルールを以下に示す。

- 人物名と代名詞を単純に置き換える
- 各役割に人数の制限はない
- 役割名が連続するときは重複を削除する
- 「物語としての登場人物でない」人名に関しては置き換えない

3.4 あらすじ検索システムの実装

前節に示した挙動をする機能を実装した。入力文は創作支援システムにおける第一幕から第三幕の入力内容のうち一文を用い、その入力を分散表現としてあらすじ文章モデルの構築に用いた文章から類似度の高い文章を上位 n 文取り出し表示する。以降、 $n = 15$ とする。

表 3: 実験 1 の質問のアンケート結果 (抜粋)

質問内容	平均	各評価の頻度				
		1	2	3	4	5
第一幕で出力文と入力文章はどれくらい似ているか	4.1	0	1	2	4	5
第一幕で出力文が含まれる作品の第三幕のあらすじと選択した作品の第三幕のあらすじは似ていたか	3.8	2	0	2	2	6

3.5 創作支援システムとあらすじ検索システムの連携

あらすじ検索システムは、入力文として時系列とキャラクター役割を考慮したあらすじ文章を想定している。これは創作支援システムで作成されたあらすじ文章が満たしているものであり、創作支援システムの利用者はプロットの作成中に好きなタイミングであらすじ検索システムを利用して作成中のあらすじ文章と物語的に類似したあらすじ文章を検索することが出来る。創作中に参考になる作品の情報を提示することは有用であると考えられるため、この 2 つのシステムを使った創作支援によって効率的に人間の創作活動を支援できると考えられる。

4. 実験

研究目的に照らして、提案システムの有用性を確かめるため、以下の 3 種類の実験をした。また、実験は被験者に 2 種類のシステムと実験手順書、実験協力の同意書を送付し、被験者が各自で実験を行うという形式とした。実験にあたって、出力される可能性のあるあらすじを三幕構成に分割したものを被験者に提示している。

4.1 実験 1: 既存の物語に対するあらすじ文章モデルの評価実験 A

あらすじ検索システムを用いて、あらすじ文書モデルにより類似あらすじが算出される頻度およびユーザが類似すると感じる観点を解析する。

4.1.1 実験条件

成人男女 12 名を対象に、以下の手順により実験を行い、アンケートにより結果を確認した。

4.1.2 実験手順

被験者は、はじめに対象作品から一つ選択し、その作品を青空文庫から確認し創作支援システムを用いてあらすじを作成する。その後、作成したあらすじを一文ずつあらすじ検索システムに入力し、出力結果を確認した後アンケートに回答する。

4.1.3 対象作品

この実験で被験者があらすじを作成するために対象とした作品は「走れメロス」「雪の女王」「ジャックと豆の木」「赤ずきん」「ヘンゼルとグレーテル」の 5 作品である。

4.2 実験 1 結果

表 3 にアンケート中の 5 段階評価の各評価の回答数を示す。アンケートは 1-5 の 5 段階評価で、5 が最も高いとした。第一幕から第三幕まで、全ての入力に対して被験者が選択した作品と同一の作品のあらすじを出力した結果が多くみられる。

同一の作品が出力されなかった場合において、あらすじ文の類似度を判断した理由として「起きた出来事が似ていた」という回数が多く得られていることから、ユーザは別の作品のあらすじ文を比較したとき、出来事や状況を表す文章の類似度を似ていると感じる尺度に用いていると考えられる。入力作品と同一

表 4: 実験 2 のアンケート結果 (抜粋)

質問内容	平均	各評価の頻度				
		1	2	3	4	5
第一幕で出力文と入力文章はどれくらい似ているか	3.4	0	0	3	2	0
第一幕で出力文が含まれる作品の第三幕のあらすじと選択した作品の第三幕のあらすじは似ていたか	1.4	3	2	0	0	0

一の作品が出力されない場合、原因は大きく分けて二つに分類できる。

- あらすじ文章の粒度が異なるパターン
- 入力した出来事がコーパスに存在しないパターン

4.3 実験 2: 既存の物語に対するあらすじ文章モデルの評価実験 B

あらすじ文章モデルの構築に用いた文章内にあらすじ文が存在しない既存の物語のあらすじをあらすじ検索システムに入力し、出力結果についてのアンケートをすることであらすじ文書モデルにより類似あらすじが出力される頻度およびユーザが類似すると感じる観点を解析する。とくに、実験 1 に比べ明確な正解が存在しないためにユーザがどの観点から類似性を見出すかに重点を置く。

4.3.1 実験条件

成人男女 5 を対象に、以下の手順により実験を行い、アンケートにより結果を確認した。

4.3.2 実験手順

実験手順は、対象の作品が任意の創作物である以外は実験 1 と同様である。

4.4 実験 2 結果

表 4 にアンケート中の 5 段階評価の各評価の回答数を示す。入力した文と似たあらすじ文自体は出力されているが、あらすじ文章のうち一文が似ているからといってその作品が想定した作品と似ているとは限らないということが分かる。

一方で、物語中の部分的な展開や目的が似ている場合があり、その例を以下に示す。

実験 2 の入出力考察 1

入力「主人公は、対象が惨殺されたショックから禁酒に失敗し、バーで泥酔して失態を犯してしまう」
出力「しかし電話で連絡してきた柳吉に、お前は来ん方がええ、来たら都合が悪いと言われ、蝶子は発作的にガス自殺を図った」
(出典: フリー百科事典「Wikipedia:『夫婦善哉 (小説)』」)

以上の例の入出力に対して被験者からは「物語の全体的な展開は大きく違うが、入力文の該当場面と出力文の該当場面で主人公がどん底にいる描写は一致していた」という意見が得られている。この入力に対しては、他にも登場人物が失態を犯す描写を表す文章が出力されており、登場人物の状態を示す文として、またその文が含まれる物語の一場面は類似している部分があると考えられる。以上の事から、類似あらすじ文の検索によって提示された 2 つの物語間には、物語として目的、三幕構成における各幕、その幕内の一部分の展開、登場人物の行動といった要素的に似ている部分が見られると考えられる。

表 5: 実験 3 のアンケート結果

質問内容	被験者 1	被験者 2	被験者 3
第一幕で出力文と入力文章はどれくらい似ているか	4	4	4
第一幕で出力文が含まれる作品あらすじと選択した作品のあらすじは似ていたか	2	4	2

4.5 実験 3: システムを用いたプロット作成実験

創作支援システムを用いた物語プロット作成中に、あらすじ検索システムを用いてその作品のプロットと似たあらすじを持つ作品を提示することで創作に与える影響について調べる。

4.6 実験条件と手順

成人男性 3 名を対象に、物語のあらすじ作成中に類似した物語を提示することで自分の書きたかった物語の作成に役立ったかどうかを調べる。実験手順を以下に示す。被験者ははじめに創作支援システムを用いて物語のあらすじを作成する。作成中の以下のタイミングであらすじ検索システムに作成中のあらすじを入力し、出力結果を確認する。

- 第一幕に該当する回答が全て入力された時点、第一幕への回答から一文ずつ選択
- 第二幕に該当する回答が全て入力された時点、第二幕への回答から一文ずつ選択
- 第三幕に該当する回答が全て入力された時点、第三幕への回答から一文ずつ選択
- ストーリーテンプレートに被験者が考える必要な部分が全て入力された時点、すべての回答から一文を選択

各結果を確認した後、アンケートに回答する。

4.7 実験 3 結果

表 5 に 5 段階評価の各被験者の回答を示す。実験 2 と同様、あらすじ文章コーパスに用いた作品が含まれないため、入力文と似たあらすじ文は出力されているが、そのあらすじ文をもつ作品はあまり似ていないという結果が得られた。

出力されたあらすじを持つ作品は創作中の作品とあまり似ていないという結果が得られた一方で、多くの箇所です自身の創作に役立てられたという回答が得られた。

実験 3 の入出力考察 1

入力「自分には新しいメニューは無理だという主人公は夢の中でラーメン仙人に出会う」

出力「大和国に住む木こりの髪長彦は、ある日森の中で神に出会う」

(出典: フリー百科事典「Wikipedia:『犬と笛』」)

以上の例の入出力に対して被験者からは出力文を含む作品と自身の作成している物語は似ていないという解答があった一方で、「展開の被りが避けられた」「今後の展開の参考になった」という解答がされている。第一幕という物語の導入部分が類似していたが、そこから展開が異なっていくことから、創作中にユーザのイメージしていた展開以外のアイデアを提示できていたと考えられる。

5. まとめ

本研究では、創作活動支援のためのシステムとして 2 種類のシステムを組み合わせた手法を提案した。提案手法につい

て被験者実験をすることで、時系列と、キャラクター役割を考慮したあらすじ文章コーパスを用いて学習した Paragraph Vector モデルにより共通した要素を持つあらすじ文を出力できること、創作活動において、物語の創作中にその物語に共通した要素を持つ作品情報を提示することの有用性を確認した。今後の課題は以下の通りである。

- あらすじ文章コーパスの規模拡大、各文章の整理
- 創作支援システムの拡張によるユーザの類似あらすじ検索システムへの入力の手間の削減

謝辞

本研究の被験者実験に協力してくださった 12 名の方々、システムに様々な助言を下された作家 山口昌志様に心より感謝します。本研究は、一部 JSPS 科研費 (グラント番号:JP17K17809) および JST, ACT-I (グラント番号:JPMJPR17U4) の支援によります。

参考文献

- [kawagoe 15] Seiya Kawagoe and Miki Ueno and Hitoshi Isahara: A Study on the Efficiency of Creating Stories by the use of Templates, The 2015 International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA 2015), 978-1-4673-8142-0(2015)
- [葛井 17] 葛井 健文, 上野 未貴, 井佐原 均: 質問集合とグラフに基づく物語全体の流れを管理可能な創作支援システムの提案, 人工知能学会 2017 年度全国大会 4F1-3in2(2017)
- [安藤 09] 安藤 隼平, 加藤 正人 (訳), シド・フィールド (著): 映画を書くためにあなたがしなくてはならないこと シド・フィールドの脚本術, フィルムアート社 (2009(原著 1979))
- [樋口 89] 樋口 淳: 怖がることを習いに出かけた若者の話—A.J. グレマスの意味論をめぐる, 専修人文論集 (44), p93-129(1989)
- [小林 05] 小林 のぞみ, 乾 健太郎, 松本 裕治, 立石 健二, 福島 俊一: 意見抽出のための評価表現の収集, 自然言語処理, Vol.12, No.3, pp.203-222(2005)
- [Wikipedia] フリー百科事典 ウィキペディア日本語版 <https://ja.wikipedia.org>
- [Le 14] Quoc Le, Tomas Mikolov: Distributed Representations of Sentences and Documents, Proceedings of The 31st International Conference on Machine Learning (ICML 2014), pp. 1188 - 1196(2014)
- [Mikolov 13] Tomas Mikolov, Kai Chen, Greg S. Corrado, Jeffrey Dean: Efficient Estimation of Word Representations in Vector Space, conferencePoster-iclr2013-workshop(2013)

[3C4-J-9] Natural language processing, information retrieval: correction of documents

Chair: Yasutomo Kimura Reviewer: Yoko Nishihara

Thu. Jun 6, 2019 3:50 PM - 4:50 PM Room C (4F International conference hall)

[3C4-J-9-01] An Approach for Applying BERT to Sentence Elimination Problem in English Exam

○Hiromi Narimatsu¹, Hiroaki Sugiyama¹, Genichiro Kikui², Hirotoishi Taira³, Seiki Matoba³, Ryuichiro Higashinaka¹ (1. NTT Communication Science Laboratories, 2. Okayama Prefectural University, 3. Osaka Institute of Technology)

3:50 PM - 4:10 PM

[3C4-J-9-02] A consideration of word sense disambiguation of company name utilizing securities report

○Hiroyuki Matsuda¹, Kazuhiko Tsuda¹ (1. Graduate School of Business Sciences, University of Tsukuba)

4:10 PM - 4:30 PM

[3C4-J-9-03] Misspelling Detection by using Multiple Bidirectional LSTM Networks

○Ryo Takahashi¹, Kazuma Minoda¹, Akihiro Masuda², Nobuyuki Ishikawa¹ (1. Recruit Technologies Co., Ltd., 2. PE-BANK, Inc)

4:30 PM - 4:50 PM

センター英語試験の不要文除去問題に対する BERTの適用方法の検討

An Approach for Applying BERT to Sentence Elimination Problem in English Exam

成松宏美^{*1}
Hiromi Narimatsu

杉山弘晃^{*1}
Hiroaki Sugiyama

菊井玄一郎^{*2}
Genichiro Kikui

平博順^{*3}
Hirotoshi Taira

的場成紀^{*3}
Seiki Matoba

東中竜一郎^{*1}
Ryuichiro Higashinaka

^{*1}NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories

^{*2}岡山県立大学
Okayama Prefectural University

^{*2}大阪工業大学
Osaka Institute of Technology

We have been working on the English problems in the “Can a Robot Get into the University of Tokyo?” project. This paper focuses on the sentence elimination problem by applying BERT, which has obtained the state-of-the-art results in a number of machine comprehension tasks. We show how we apply BERT and report the improvements made over baselines. Finally, we show our error analysis and the problems that still need to be solved.

1. はじめに

我々は「ロボットは東大に入れるか」プロジェクト [新井 18] において、引き続き英語（特に、センター試験の英語問題）に取り組んでいる。本稿は、センター試験の英語問題で出題される不要文除去問題に対して、近年多くの機械読解タスクにおいて State-of-the-art (SOTA) を達成している汎用言語表現モデル BERT [Devlin 18] を用いた解法について述べる。

不要文除去問題は、文章中に 1 つの不要な文が含まれており、それを取り除くことで全体のまとまりが良くなるような文の一つを選ぶという問題である。図 1 に示す例では、(1) より前の文脈により、(1) 以降に良い靴選びのポイントが提示されると推測できる。ここで、(1)(2)(4) はそのポイントが提示されているものの、(3) はブランドの革靴の話をしており、主題が異なる。よって、(3) が不要文であり、これを選べば正解となる。我々はこの問題に対して、様々な手法を検討してきたが Word2vec [Mikolov 13] を用いて選択肢間の距離を測るシンプルな手法がもっとも良いスコアであったことを報告した [東中 17]。

不要文除去問題は、文同士の類似性だけでなく、文書としての自然さの評価が必要な点で、近年取り組まれている機械読解タスクにはあまり見られない問題である。近年、機械読解タスクにおいて注目を集めている OpenAI GPT [Radford 18] や BERT は、Transformer と呼ばれる自己注意機構を備えたニューラルネットワークを大規模なテキストコーパスを用いて事前学習し、個別の問題に対して転移学習することで、様々なタスクにおいて SOTA を達成している。転移学習により様々なタスクに適用できるため、不要文除去問題においても精度向上が期待できる。

本稿では、BERT を不要文除去問題に適用する方法を検討し、Word2vec に比べ、有意に正解率が良くなったことを示す。また、エラー分析により、BERT により解けるようになった

Wearing proper shoes can reduce problems with your feet. Here are some important points to think about in order to choose the right shoes.

(1) Make sure the insole, the inner bottom part of the shoe, is made of material which absorbs the impact on your foot when walking. (2) The upper part of the shoe should be made of breathable material such as leather or cloth. (3) Some brand-name leather shoes are famous because of their fashionable designs. (4) When you try on shoes, pay attention not only to their length but also to their depth and width. Wearing the right shoes lets you enjoy walking with fewer problems.

図 1: 不要文除去問題の例 (平成 29 年センター英語試験問題より引用。正解は (3)。)

問題とそうでない問題がどのようなものかを示す。

2. BERT による不要文除去問題の解法

ここでは、問題解答に用いる Bidirectional Encoder Representations from Transformers (BERT) [Devlin 18] について説明するとともに、BERT を不要文除去問題に適用する方法を述べる。転移学習を行う際の、学習に用いるデータの作成方法 (2.2) および BERT の入力形式への問題の変換方法 (2.3) と、転移学習を行わずに BERT の後続文予測モデルを用いる解法 (2.4) を説明する。

2.1 Bidirectional Encoder Representations from Transformers (BERT)

BERT は、図 2 (左) に示す Transformer モデル [Alec 18] を大規模なテキストコーパスで事前学習し、個別の問題に対して転移学習する手法である [Devlin 18]。Transformer は、位置情報 (Position embedding) 付きのテキストを入力として、「自分と関係する周辺情報を集約する」機能を持つ自己注意機

連絡先: 成松宏美, NTT コミュニケーション科学基礎研究所,
〒 619-0237 京都府相楽郡精華町光台 2-4, 0774-93-5311,
hiromi.narimatsu.eg@hco.ntt.co.jp

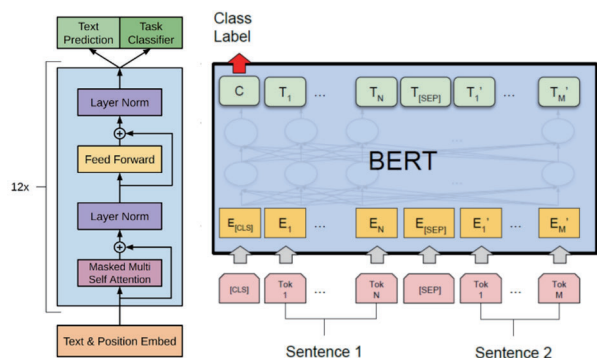


図 2: Transformer モデルの構造 (右) [Alec 18] と個別タスクに対する転移学習時の入出力・モデル構造 (左) [Devlin 18]

構を繰り返し適用することで、タスクに適した特徴ベクトルを計算するモデルである。BERT では Transformer の事前学習にアノテーションが不要である。事前学習のための汎用的なタスクとして、双方向の言語モデルタスクと 2 文の結束性を判定するタスクの 2 種類が採用されている。双方向言語モデルタスクは、文中のマスクされた単語を推定するタスクであり、対象単語の前後の文脈情報を利用する言語モデルを学習することができる。2 文の結束性タスクは、与えられた 2 文が連続する文が無関係な文かを推定するタスクである。文という、単語よりも大きい単位でのつながりの良さを学習できると考えられる。なお BERT では、この 2 文を異なるものとして陽に表現するため、入力情報に文のまとまりを表す segmentID およびそのベクトル表現 (Segment embeddings) を追加し、直接的に文のまとまりを与える工夫がなされている。

2.2 転移学習に用いる擬似問題の作成

BERT の転移学習は、比較的少量のデータでも実行可能ではあるものの、数百では十分な量とは言えない。著者らが保持する不要文除去問題は 249 問であり、転移学習には不十分な分量だと考えられる。そのため、本研究では、既存の長文テキストの一部に不要な文を追加することで擬似的に不要文除去問題を作成し、転移学習に利用するデータの量を増やすことで、解答精度の向上を試みる。

利用するテキストとして、本研究では RACE データセット [Lai 17] の本文部分を用いる。擬似問題を 1 問作成する場合、この本文から連続する 7 文を抜き出して正しい文章とし、7 文以外の本文の箇所からランダムに抜き出した 1 文を不要文とする。得られた不要文を、先頭・末尾以外の箇所に挿入し、図 1 のように、不要な文を 1 文含む擬似問題とする。上記の連続する 7 文を抽出するウィンドウをスライドさせていくことで、RACE データセットの 1 つの問題から、おおよそ 10 問程度の擬似問題が作成できる。最終的に作成できた問題数は 80 万問程度であった。

2.3 4 択の不要文除去問題への適用

本研究では、ある選択肢の要否を判断する 2 値分類器として BERT を学習し、各選択肢について個別に推定された不要らしさを 4 つの選択肢間で比較し、不要らしさが最大であったものを解答として出力するというアプローチを採用する。

また、BERT への入力形式は、選択肢を置く位置や segmentID の値、選択肢近傍の見る範囲など、数種類のパターンが考えられる。以下に、採用した 4 つの入力形式を述べる。

doc-opt 選択肢 (opt) を除く本文 (doc) を文頭側にまとめて並べ、セパレータを挟んで opt を文末側に配置する方法 (図 3)。本文側の segmentID を 0, opt 側を 1 とする。BERT の事前学習の一つである、2 文間の結束性判定を利用したものであり、opt が doc と無関係な文であった場合に、不要であると考えられる。

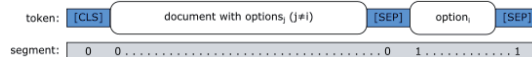


図 3: doc-opt

3opt-opt 4 つの選択肢のうち、対象とする選択肢以外 (3opt) をまとめて文頭に置き、対象とする選択肢 opt を文末に配置する方法 (図 4)。doc-opt 同様、3opt の segmentID を 0, opt の segmentID を 1 として、2 文間の結束性判定を利用して解答する。doc-opt に比べて見る範囲が狭いため、判定に情報が欠落する可能性がある一方、学習を効率的に行える可能性がある。



図 4: 3opt-opt

prevN-opt-nextN 対象とする選択肢 (opt) を中心として、前後 N 文ずつを抽出し、出現順通りに並べて配置する方法 (図 5)。opt の segmentID を 1, それ以外を 0 とする。N を無限に大きくした場合は、doc-opt における opt の配置を文中の出現箇所としたものに対応する。position encoding による、出現位置の情報を利用することで、より出現位置に敏感なモデルになると考えられる。



図 5: prevN-opt-nextN

prevN-nextN prevN-opt-nextN のうち、opt を取り除いてその前後のみを利用する方法。opt が含まれない状態で判定するため、opt とその他の箇所との意味的な距離を判定に利用することができず、必要な文が抜けた場合の不自然さ、および正しく不要な文が抜けた場合の自然さ、を利用して判定する必要がある。そのため、opt を利用する他のモデルとは解答傾向が異なることが期待される。

2.4 BERT の後続文予測を用いた手法 (転移学習なし)

BERT は事前学習における目的関数として、穴埋め (cloze test) および後続文予測 (next sentence prediction) の正解率を目的関数としている。事前学習は大量のコーパスを使い大きな計算コストをかけて行っていることから、特に学習データが少ない場合に、対象とする問題を事前学習の目的関数に類似した問題に帰着させることができれば転移学習なしである程度の精度が得られることが期待できる。そこで本研究では、前節までで述べた疑似負例を用いて転移学習を行う方法に加え、事前学習自体を活かし転移学習なしで問題を解く方法を提案する。

不要文除去問題を「各選択肢がその直前の文脈に後続するかどうか判定する問題」と考えると事前学習における後続文

A : <選択肢の左 k 文> [SEP] <選択肢> <選択肢の右 m 文>
 B : <選択肢の左 k 文> [SEP] <選択肢の右 m+1 文>

図 6: 後続文予測への 2 つの入力形式

予測のタスクと同等とみなすことができる。この考えのもと、我々は BERT の事前学習モデルのみによる後続文予測問題として解答を試みる。各選択肢についてそれぞれ独立にその選択肢が直前の文に後続しうかどうかを判別（二値分類）し、接続しないものを除去すべき選択肢として選ぶのが最も単純な実装であるが、この場合、選択肢が一つのみ選ばれるとは限らない。二値分類の前提となる尤度を用いて尤度最低の（すなわち最も後続性の低い）選択肢を選ぶ方法が考えられるが確率計算のベースが異なるためか予備実験では精度が低かった。

我々は、各選択肢について「その選択肢を除去しない場合」と「その選択肢を除去して次の文に移移する場合」の接続性の良さを求め、この差を「当該選択肢を除去すべきスコア」と考えた。すなわち、一つの選択肢に対して図 6 のような A, B という入力を BERT の後続文予測モデルに与え、出力の尤度値（対数 odds）の差をこの選択肢のスコアとする。なお k, m は実験的に $k=2, m=2$ と定めた。各選択についてこのスコアを求め、スコアが最大のものを解答とする。

3. 評価

3.1 実験設定

2014~2019 年のセンター本試験および追試験、予備校の模試に含まれる不要文除去問題 129 問をテストセット、独自に作成した不要文除去問題 120 問を開発セットとして、評価に用いる。独自に開発した問題は、平均的な英語力を持つ人の正解率が 50% 程度になるように難易度を調整している。比較するモデルは、9 種である。ベースラインとして、3 つの選択肢と 1 つの選択肢の距離を Word2vec のコサイン類似度で算出し、距離の最も遠い選択肢を選んだ場合と比較する。prevN-opt-nextN の N は 1, 2, all とする。不要文除去問題は約 7 文程度から構成されていることから、 $N=3$ とすることは、ほぼ全文使用に等しい。また、prevN-nextN については、文章としての自然さを評価するものであるため、segmentID を全て 0 にした場合も評価する。転移学習のパラメータは、バッチサイズ 32, 最大系列長 512, dropout は 0.1 固定, epoch 数は 4, 学習率は $5e^{-6}$ と $5e^{-7}$ の 2 種類とした。各手法において開発セットで最大正解率時のモデルを用いてテストセットの正解率を評価する。学習データ数は 670,540 で、正例・負例の割合は同じになるようにした。

3.2 結果

各手法の正解率を表 1 に示す。Word2vec で 0.457 だったのに対して、BERT で遷移学習した場合に prevN-nextN (seg0)（選択肢を除いた前文と後文を segmentID 0 で埋め込んだ場合）において最大の 0.612 のスコアが得られ、カイ二乗検定においても有意な向上が見られた ($p=0.0009$)。また、prevN-opt-nextN においては、前後 1 文だけをみるよりも、前後 2 文またそれ以上をみた方が選択肢の要否を正しく判断できることがわかった。これは、人間が問題を解く際にも前後 1 文から対象の文が必要か不要かを判断することが難しいことから、妥当な結果であると考えられる。

また、prevN-nextN については、segmentID を切り替えるよりも、全て 0 で固定した場合の方が高い正解率が得られた。文章としての自然さをそのまま評価する方が適していると考え

表 1: 各手法の正解率。

手法	正解率
(1) ベースライン (w2v)	0.457 (59/129)
(2) Doc-opt	0.543 (70/129)
(3) 3opt-opt	0.558 (72/129)
(4) prev1-opt-next1	0.372 (48/129)
(5) prev2-opt-next2	0.550 (71/129)
(6) prevN-opt-nextN	0.535 (69/129)
(7) prevN-nextN	0.550 (71/129)
(8) prevN-nextN (seg0)	0.628 (81/129)
(9) 転移学習なし	0.512 (66/129)

られる。

次に、手法毎に解けている問題がどのように異なるかを検証する。2 つの手法の正誤関係に対し、カイ二乗検定を行い独立性が棄却されれば、正誤の傾向に関係があることが示される。すなわち、同様の問題に正答する傾向もしくは異なる問題に正答する傾向が示される。さらに残差分析により異なる問題に対しての正答している数に対して特に有意な差が見られれば、異なる問題に正答していることが示せることを利用して、正答の傾向が近いかどうかを見る。本手法を用いて、2 種類の比較を行う。表 2 に Word2vec とそれと同様の特徴を学習すると考えられる Doc-opt および 3opt-opt との比較と、表 3 に BERT の転移学習を行う手法におけるトップモデル間の比較を示す。各値は p 値であり、*印は、多重検定前の有意水準を 0.05 としたとき、ホルム補正および残差分析を行った結果、一方が正答している箇所有意な差が見られた場合に付与した。Word2vec との比較においては、Doc-opt による手法は異なる問題に正答していることがわかった。これは Doc 内での文章の自然さおよび Doc と後続する opt との含意関係が学習された可能性が考えられる。また、転移学習を行う手法におけるトップモデル間の正誤比較より、3opt-opt と prevN-nextN (seg0) に有意な差が見られた。3opt-opt は選択肢間の類似度を、prevN-nextN (seg0) は文章としての自然さをというように異なる点を表現できている可能性がある。これらをうまくアンサンブル学習することができれば、さらに正解率が向上できる可能性があると考えられる。

表 2: W2V と類似モデル間の関係比較。

	(1)	(2)	(3)
(1) ベースライン (w2v)	—	*0.00793	0.04746
(2) Doc-opt	—	—	1.0
(3) 3opt-opt	—	—	—

表 3: BERT 学習モデル間の正誤の関係比較。

	(3)	(6)	(8)
(3) 3opt-opt	—	1.0	*0.02103
(6) prevN-opt-nextN	—	—	*0.02414
(8) prevN-nextN (seg0)	—	—	—

3.3 分析

BERT の適用によって Word2vec からどのような問題が解けるようになり、一方で依然正答できていないかについて、問題毎の正誤の傾向により分析する。ここでは、Word2vec と BERT の適用によって最高スコアが得られた prevN-nextN (seg0) とを比較し（表 1）、BERT でのみ正答した問題、両方で誤答した問題の例を用いて分析する。4 つの選択肢のスコアのうち、

Food can do more than fill our stomachs? it also satisfies feelings. If you try to satisfy those feelings with food when you are not hungry, this is known as emotional eating. There are some significant differences between emotional hunger and physical hunger. (1) Emotional and physical hunger are both signals of emptiness which you try to eliminate with food. (2) Emotional hunger comes on suddenly, while physical hunger occurs gradually. (3) Emotional hunger feels like it needs to be dealt with instantly with the food you want; physical hunger can wait. (4) Emotional eating can leave behind feelings of guilt although eating due to physical hunger does not. Emotional hunger cannot be fully satisfied with food. Although eating may feel good at that moment, the feeling that caused the hunger is still there.

図 7: Word2vec および BERT で誤った問題の例 (2016 年センター試験本試験より引用。正解は (1), BERT は (2) を選択)。

最大値が、他の 3 つの選択肢のスコアと離れているものは自信を持って選択したと考え、そのような問題を分析対象として選出した。尚、Word2vec と prevN-nextN (seg0) の正誤を比較すると両者とも正答は 36 問、BERT でのみ正答は 45 問、Word2vec でのみ正答は 23 問、両者とも誤答は 25 問であり、2 手法の組み合わせオラクルでは他の組み合わせと比較してもっとも高く、スコアは 0.806 (104/129) であった。

図 8 は、word2vec で誤り、BERT で正解した問題である。主題は缶切りの利点についてであり、(1) から (3) は共通してその主題をサポートしているものの、(4) についてはサポートしていない。よって、BERT によって、主題のすり替えによる不自然さをうまく判断できるようになった可能性がある。

続いて、図 7 は、両手法で誤った問題である。(1) より前の文にて Emotional hunger と physical hunger には重要な違いがいくつかあることが述べられているものの、後続する (1) では、共通点が述べられている。このように、論理的なつながりの判定が必要な問題においては、現在の手法では判断できないと考えられる。

One of the most important kitchen tools is the simple hand-operated can opener – the manual can opener. (1) Can openers are needed to open some canned foods, and nowadays many people have easy-to-use electric ones. (2) However, with a manual can opener, even when there is an electric power failure, you can still open cans. (3) Another advantage of a manual can opener is that it will last for years without any maintenance. (4) Recently, even some electric can openers with multiple functions have been getting cheaper. In any event, it is always a good idea to have a manual can opener in your kitchen.

図 8: Word2vec で誤り BERT で正答した問題の例 (2014 年センター試験追試験より引用。正解は (4))。

4. まとめと今後の課題

センター英語試験で出題される不要文除去問題に対し、近年あらゆるタスクで SOTA を出している BERT の適用方法について検討し、比較を行った。これまでの最高得点を得ることが

できた。これは適用の際に用いた擬似負例が有効であったと考えられる。また、選択肢の前の N 文と選択肢の後の N 文を単純に連結した埋め込み方法がもっとも正解率が高くなる (試験問題の正解率 0.628) ことがわかり、効果的な適用方法を明らかにした。また、従来の Word2vec 手法と比較して正誤の傾向を分析したところ、この埋め込み方法によって文章としての自然な流れを判断できるよう学習された可能性が高い。しかしながら、論理的なつながりや飛躍の判定が必要な問題の場合には、誤った選択肢を選ぶ傾向があることがわかった。今後はこの問題を解決するため、全体としてのつながりの良さと局所的なつながりの良さの両方を判断できるようなアンサンブル学習を検討する。合わせて、係り受け関係などとの併用により、論理的な構造が自然さに反映されるような工夫を検討していく。

謝辞

本研究を推進するにあたって、大学入試センター試験問題のデータをご提供下さった独立行政法人大学入試センターおよび株式会社ジェイシー教育研究所に感謝いたします。実験データをご提供くださいました学校法人高宮学園、株式会社ベネッセコーポレーションに感謝いたします。

参考文献

- [Alec 18] Alec, R., Karthik, N., Tim, S., and Sutskever, I.: Improving Language Understanding by Generative Pre-Training, *arXiv:1802.05365* (2018)
- [Devlin 18] Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, *arXiv:1810.04805* (2018)
- [Lai 17] Lai, G., Xie, Q., Liu, H., Yang, Y., and Hovy, E.: RACE: Large-scale Reading Comprehension Dataset From Examinations, in *Proc. of EMNLP 2017*, pp. 785–794 (2017)
- [Mikolov 13] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J.: Distributed representations of words and phrases and their compositionality, in *Advances in neural information processing systems*, pp. 3111–3119 (2013)
- [Radford 18] Radford, A., Narasimhan, K., Salimans, T., and Sutskever, I.: Improving language understanding by generative pre-training, URL https://s3-us-west-2.amazonaws.com/openai-assets/research-covers/languageunsupervised/language_understanding_paper.pdf (2018)
- [新井 18] 新井紀子, 東中竜一郎 F 人工知能プロジェクト「ロボットは東大に入れるか」: 第三次 AI ブームの到達点と限界 (2018)
- [東中 17] 東中 竜一郎, 杉山 弘晃, 成松 宏美, 磯崎 秀樹, 菊井 玄一郎, 堂坂 浩二, 平 博順, 南 泰浩, 大和 淳司 F「ロボットは東大に入れるか」プロジェクトにおける英語科目の到達点と今後の課題, 2017 年度人工知能学会全国大会予稿集, pp. 2H2–1 (2017)

有価証券報告書を活用した企業名の語義曖昧性解消法の一考察 A consideration of word sense disambiguation of company name utilizing securities report

松田 裕之^{*1}
Hiroyuki Matsuda

津田 和彦^{*1}
Kazuhiko Tsuda

^{*1} 筑波大学大学院ビジネス科学研究科
Graduate School of Business Sciences, University of Tsukuba

Word Sense Disambiguation (WSD) is a research field that distinguishes semantics from peripheral information of the target word. This research worked on WSD of company names and aimed to acquire the same classification accuracy as supervised learning with unsupervised learning. Success of WSD of company names with unsupervised learning will enable us to extract only the information of the specific business without considering the appropriate search word, without putting enormous effort on preparing teacher data. We proposed a classification method to judge word sense from similarity of word vectors by business created using securities report and word vectors of words in classification target. With the proposed method, we achieved almost the same classification accuracy as supervised learning. It is suggested that introduction of a model for determining similarity to unknown words such as fasttext, suggests that there is room for improvement of accuracy.

1. はじめに

従来のテキストマイニングの課題に、文字列の表記のみが集計・分析されていることがある。例えば、運転手を指す「ドライバー」と、ねじ回しを指す「ドライバー」のような多義語の課題である。このような多義語に対して、対象語の周辺情報などから語義を識別する語義曖昧性解消という研究分野がある。

本研究では、対象語を企業名に絞り、教師あり学習で構築した分類器に匹敵する分類精度を、教師データの作成を行わずに達成することを目指す。分類対象の典型例として「ヤマハ」を取り上げ、①楽器メーカーとしてのヤマハ、②二輪メーカーとしてのヤマハ、③①②のいずれでもないヤマハ、の3つの語義に識別する。

本研究によって、期待される成果は2点ある。1点目は、企業名の語義曖昧性解消ができれば、適切な検索キーワードを思いつかずとも、特定事業の情報のみを抽出することが可能となる点である。「ヤマハの二輪製造事業についてテキストデータから分析せよ」と言われれば、「ヤマハ バイク」という条件で検索しデータを抽出する人が多いと思われるが、二輪製造事業に関するデータで「バイク」が含まれないものは多くある。一方で、適切な検索キーワードを1ユーザーが独力で全て羅列するのは不可能である。そのような場合、企業名の語義曖昧性解消手段が確立されていることは有用である。

2点目は、教師なし学習による語義曖昧性解消で、対象語の周辺情報に加え、外部知識としてシソーラスの語釈文を活用する手法が Chen ら[Chen 2014]より提案されている。このアルゴリズムを拡張し、有価証券報告書を知識抽出ソースとし、企業名の語義曖昧性解消で外部知識を活用した点である。

2. 語義曖昧性解消

機械学習手法による語義曖昧性解消は、一般に教師あり学習にて解決することが多い。例えば、単語「ドライバー」を含む用例を適当な数集め、各々の用例に対してその用例中の「ドライバー」の語義を付与しておく。これを教師データに、周辺語の情報などを対象語の素性として分類器を構築、識別タスクを実

行する手法である。

しかし、教師あり学習による分類は教師データの作成コストが大きく、対象語が限定されてしまう問題がある。全ての単語に語義を付与する語義曖昧性解消は all-words WSD というタスクとして研究されているが、ここでは教師あり学習によるアプローチは非現実的であり、教師なし学習が用いられる。ただし、教師なし学習による識別精度は一般に教師あり学習よりも低い問題がある。

教師なし学習による語義曖昧性解消において一時 state-of-the-art の性能を示していたのが、Chen らの手法である。Chen らは、Mikolov ら[Mikolov 2013]より提唱された Skip-gram により単語分散表現を得た後、シソーラス WordNet[wordnet]上の、多義語の語釈文中の類似単語を利用して各語義の意味ベクトルを作成、この意味ベクトルと対象語が含まれる文のコンテキストベクトルとのコサイン類似度から、語義を判定している。

本研究の提案手法は Chen らの手法を拡張したものである。Chen らは外部知識として WordNet の語釈文を活用したが、企業名について各語義を説明している情報として、本研究では有価証券報告書を活用した。

3. 外部知識導入による企業名の語義曖昧性解消

語義曖昧性解消の対象とする企業には「ヤマハ」を選定した。ここでいう「ヤマハ」は、ヤマハ株式会社とヤマハ発動機株式会社の2社を指す。

日経テレコンにて取得した日経新聞朝刊記事中の「ヤマハ」を含む文 2109 件を分類対象データとした。さらに、有価証券報告書上で各事業に言及する箇所を4箇所特定し、2017 年度のヤマハ株式会社およびヤマハ発動機株式会社の有価証券報告書から事業別に文を抽出し、本研究で用いる外部知識データとした。

本研究では、本研究独自の手法に加え比較対象として、教師あり学習、教師なし学習、Chen らの研究をベースとした手法の3つ、計4手法で分類精度を比較した。

まず、教師あり学習/教師なし学習向け検証データの作成法を簡単に述べる。分類対象データおよび有価証券報告書から tf-idf 値が一定以上の名詞を抽出し、機械学習で活用する素性ベクトルとした。分類対象データから得たものは図 1 のようにな

り、これを「対象語の周辺語から得た素性ベクトル」と呼ぶ。さらに、有価証券報告書から得たものを追加すると図 2 のようになり、これを「外部知識から得た素性ベクトル」と呼ぶこととする。

図 1. 対象語の周辺語から得た素性

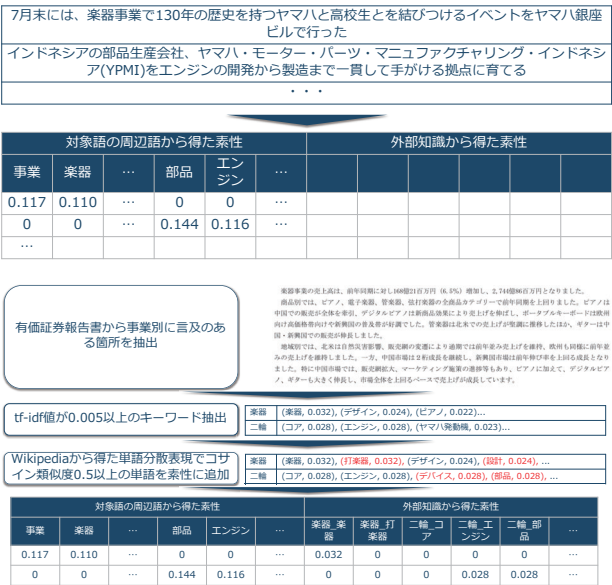


図 2. 外部知識から得た素性

分類方法は、教師あり学習では SVM、教師なし学習では K-means 法を実施した。

次に、Chen らの研究をベースとした手法(以下「Chen らの手法」と呼ぶ)/本研究の手法向け検証データの作成法について述べる。有価証券報告書の事業別の文から抽出した名詞について、ヤマハの事業であれば「ヤマハ」、ヤマハ発動機の事業であれば「ヤマハ発動機」の単語ベクトル(朝日新聞コーパス[田口 2017]に gensim ライブラリ[gensim]の word2vec 関数を適用して取得)とのコサイン類似度を計算した。次に、コサイン類似度の閾値 t を 0 から 1 まで調整しながら、コサイン類似度が t 以上の単語のみを抽出した。最後に、抽出したコサイン類似度 t 以上の単語群について単語ベクトルの平均を計算し、これを事業別単語ベクトルとみなした。

以上の事業別単語ベクトルの作成プロセスをまとめると、図 3 のようになる。

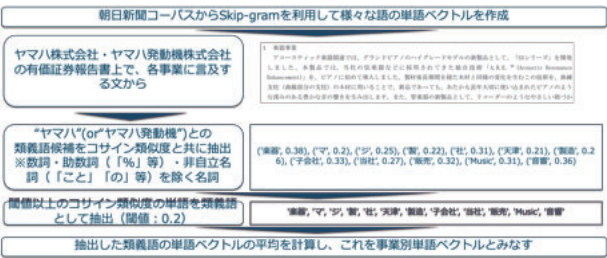


図 3. 事業別単語ベクトルの作成

分類方法について述べる。Chen らの手法では、まず、分類対象データに現れる名詞の単語ベクトルの平均を計算した(これを「コンテキストベクトル」と呼ぶ)。次に、事業別単語ベクトルとコンテキストベクトルのコサイン類似度を計算した。最後に、コサイン類似度の閾値 t_c を 0 から 1 まで調整しながら、「楽器」「二輪」事業ベクトルとのコサイン類似度が t_c 以上かつ最も高いものに分類した。複数の事業ベクトルとコサイン類似度が同値で

あるか、いずれの事業ベクトルともコサイン類似度が t_c 未満である場合は、「その他」に分類した。

一方、本研究の手法では、まず、分類対象データに現れる語について各々単語ベクトルを得た。この語群に対し、コサイン類似度の閾値 t_0 を 0 から 1 まで調整しながら、事業別単語ベクトルとのコサイン類似度が t_0 以上の語が現れる度にスコアを+1 した。以上の計算で、スコアが最も高いものに分類した。複数のスコアが同値か、いずれのスコアも 0 の場合、「その他」に分類した。

4. 評価結果と考察

各手法で達成した正答率は表 1 の通りである。「対象語の周辺語から得た素性ベクトル」のみを活用しているのが「外部知識活用なし」であり、「対象語の周辺語から得た素性ベクトル」に加え「外部知識から得た素性ベクトル」も活用しているのが「外部知識活用あり」である。また、Chen らの手法および本研究の手法では、コサイン類似度の閾値 t, t_c, t_0 を調整する中で正答率が最高となったときの値を示している。

表 1. 各手法で達成した正答率

分類手法		正答率
教師あり学習	SVM (外部知識活用なし)	78%
	SVM (外部知識活用あり)	80%
教師なし学習	K-means (外部知識活用なし)	45%
	K-means (外部知識活用あり)	63%
Chenらの手法	事業別単語ベクトルとコンテキストベクトルのコサイン類似度から判定	71%
本研究の手法	事業別単語ベクトルと文中の語のコサイン類似度から判定	76%

分類精度としては教師あり学習に匹敵する程度が求められたが、教師あり学習による正答率は 80%であったのに対し、本研究の手法による正答率は 76%と、教師あり学習には届かないものの4ポイント低いのみ水準を達成しており、目的は一定達成したと言える。

5. おわりに

本研究では、外部知識として有価証券報告書を活用し、事業別単語ベクトルを作成した上で、事業別単語ベクトルと分類対象中の語の類似度から分類を行う手法により、目標とする教師あり学習に近い分類精度のアルゴリズムを構築することに成功した。

今後、取り組むべき課題は主に2点挙げられる。1点目は、word2vec モデルにとつての未知語の存在が、分類精度向上の壁の1つとなった点である。fasttext モデル[Bojanowski 2016]などを活用し、未知語に対する類似性も判定することで分類精度を向上させられる可能性がある。

2点目は、アルゴリズムの汎用性の担保である。本研究は分類対象を「ヤマハ」に限定し、各種パラメータの最適化を行なっているため、分類対象を変更した場合に有効なパラメータであるかは確認できていない。分類対象を複数事業に取り組む他企業に拡張し、汎用的なアルゴリズム・パラメータであるか検証する必要がある。

参考文献

[Chen 2014] D.Chen, C.D.Manning: A Fast and Accurate Dependency Parser using Neural Networks, Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014.

- [Mikolov 2013] T.Mikolov, I.Sutskever, K.Chen, G.Corrado, J.Dean: Distributed Representations of Words and Phrases and their Compositionality, CoRR, 2013.
- [wordnet] <http://www.nltk.org/howto/wordnet.html>, 最終アクセス日:2019-02-03
- [田口 2017] 田口雄哉, 田森秀明, 人見雄太, 西鳥羽二郎, 菊田洸: 同義語を考慮した日本語の単語分散表現の学習, 情報処理学会研究報告, 2017.
- [gensim] <https://radimrehurek.com/gensim/index.html>, 最終アクセス日:2019-02-03
- [Bojanowski 2016] P. Bojanowski, E. Grave, A. Joulin, T. Mikolov: Enriching Word Vectors with Subword Information, CoRR, 2016.

Bidirectional LSTM を用いた誤字脱字検出システム

Misspelling Detection by using Multiple Bidirectional LSTM Networks

高橋 諒^{*1} 蓑田 和麻^{*1} 舩田 明寛^{*2} 石川 信行^{*1}
 Ryo Takahashi Kazuma Minoda Akihiro Masuda Nobuyuki Ishikawa

^{*1} 株式会社リクルートテクノロジーズ ^{*2} 株式会社 PE-BANK
 Recruit Technologies Co.,Ltd. PE-BANK, Inc.

Companies in the RECRUIT Group provide matching business between clients and customers, and create lots of manuscripts every day in order to tell the attractiveness of our clients. In this paper, we propose a method for detecting misspelling in manuscripts with machine learning. That system mainly consists of two parts. One is the multiple Bidirectional LSTM networks to estimate the probabilities of correctness in each characters. The other is the random forests algorithm to decide what sentence is correct or not by using outputs of these networks. The efficacy of our approach is demonstrated on two datasets: artificial sentences and real manuscripts created in our services.

1. 背景・課題

情報を提供するクライアントと情報を求めるカスタマーをマッチングするのがリクルートのビジネスモデルである。このビジネスモデルにおいて、原稿はクライアントとカスタマーを結びつける重要な情報伝達手段である。その原稿において、万が一誤った内容が記載されてしまうと、企業としての信頼に関わる問題に発展するため、校閲業務に対しては多くのリソースが割かれている。しかし、それでも稀に不備のある原稿が発見されたり、文章として違和感のある原稿が掲載される事態が発生したりすることが現状であった。その原因の1つが、1枚の原稿に対してチェックすべき規定の多さである。通常どのサービスにおいても、それぞれに独自の原稿規定が存在しており、その数は各々100~200項目に渡る。年間数十万件の規模で新規原稿が作成される中で、それら1枚1枚に対し全項目のチェックを人手で行うのは困難であり、熟練した校閲者でも見落としが発生していた。特に、多くの規定の中でも、誤字脱字についてはチェックが十分に行われていないケースが散見されていた。

上記の課題に対し、システムによる校閲補助および自動校閲が出来ないかの検討を行った。具体的には、NGワードチェックのようなルールベースでの校閲に加え、誤字脱字や原稿内での表記ゆれの検出を機械学習により行うハイブリッドな校閲システムを作成し、実導入を行った。本論文においては、その中でも深層学習を用いた誤字脱字検出ロジックについてのアルゴリズムとその実験/導入結果について記す。

誤字脱字の典型例として、“私は猫が空きです”のような漢字の変換ミスや“私に猫がに好きです”というような助詞の間違いなどがある。このような誤字脱字は、単語の組み合わせで検出しようとすると、その数が膨大になり全てを定義することは困難なため、ルールベースによる検出は難しい。そこで、本研究では文字の系列情報を利用できる深層学習を利用したアプローチを試みた。

2. 関連研究

深層学習の分野において、様々なネットワークが提案されているが、文章や音声波形などの系列データに対して有効なネットワークとして Recurrent Neural Network (RNN) [1] が存在する。

自然言語処理という観点で、RNNを用いた事例として代表的なものとして文章の自動生成がある[2][3]。RNNを生成モデルとして捉え、文章として成立している文字列 (x_1, \dots, x_T) を入力とし、それぞれの次の文字を示す (x_2, \dots, x_{T+1}) を正解として学習させることで、時刻 $t+1$ に出現する文字の確率 $P(x_{t+1}|x_1, \dots, x_t) = \text{softmax}(o_t)$ を取得する。ここで o_t とは、時刻 t におけるネットワークの最終出力である。この確率値 $P(x_{t+1}|x_1, \dots, x_t)$ が最大となる文字を次の文字として順々に生成することで文章の生成を行う。

本研究においては、この言語モデルとしての RNN をベースに異常検知として利用している。RNN を異常検知の文脈で利用した研究として、例えば[4]や[5]が存在する。いずれも RNN の最終層に二値分類を行うための Dense Layer を繋ぐことで、その系列が正常か異常かを判定している。この形式の異常検知アルゴリズムを、日本語の言語処理に適用した先行研究として[6]が挙げられる。[6]ではテレビで利用されるテロップにおける誤字脱字検出を目的とし、誤字脱字を予め8つのパターンに分類し、それぞれのパターンに対して正常/異常の二値分類を行う RNN モデルの構築を行っている。[6]ではパターン毎のモデルの結果を単一文章に対して重ね掛けで検出した場合、モデル数が増えると精度が下がる点が指摘されている。

提案手法では、RNN の中でも、長期依存性をもつ LSTM [7] を双方向に発展させた Bidirectional-LSTM (BLSTM) [8] を採用した。更に、言語モデルと正常/異常の二値分類の BLSTM を並列で利用し、それぞれの出力値を入力としたランダムフォレストの結果から誤字脱字を含むか否かの判定を行った。提案手法の特徴は下記4点である。

- BLSTM を利用することでターゲットとなる文字の前後双方の情報を利用できる
- 言語モデルを組み合わせているため、予め考えられない誤字脱字のパターンに対しても対応ができる
- 言語モデルの出力結果を参照することで、誤字脱字判定された文字の代替提案が可能である
- 複数の BLSTM のモデルの出力の組み合わせにランダムフォレストを利用することで、検出時の閾値設定が容易になった

3章では用いたデータセットについて、4章では提案手法の詳細について述べ、5、6章では実際のデータを用いた実験と

連絡先: 株式会社リクルートテクノロジーズ

IT エンジニアリング本部データテクノロジーラボ部 高橋 諒
 (ryo_takahashi@r.recruit.co.jp)

試験運用結果について紹介し、7 章で今後の展望について説明を行う。

3. 学習データセット

本手法では、BLSTM/ランダムフォレストモデルの学習用に誤字脱字を含まない文(OK 文)と含む文(NG 文)のデータセットが必要である。リクルートには校閲済み原稿が大量に存在するため、これを OK 文として利用する。次に、過去の校閲内容の分析より頻出の誤字脱字のパターン(以下、NG パターン)を表 1 のように定義し、OK 文を基にして NG パターンに該当する誤字脱字を含む NG 文を作成することにした。対象原稿としては後述する試験運用を見据え、リクルートが運営するサービスの一つであるゼクシィの原稿を利用した。用意した学習データセットを表 2 に示す。校閲済み過去原稿は 2015 年 1 月～2018 年 1 月に掲載された原稿であり OK 文のみで構成される。作成 NG 文 I / II は NG 文と、その基となった OK 文のペアで構成され、誤字脱字の箇所の情報も含む。なお、作成 NG 文 I は人手で作成したが、そこで不足した NG パターンを補うため作成 NG 文 II を機械的に作成した。

表 1: NG パターン定義

名称	内容	例
漢字	漢字変換ミス	正) 100名まで収容可能な会場。 誤) 100名まで収容可能な海上。
助詞連続	助詞の不自然な連続	正) ドレスのご試着は、 誤) ドレスのをご試着は、
脱字(送り仮名)	送り仮名の脱字	正) ご要望にお応えします。 誤) ご要望にお応します。
脱字(助詞)	助詞の脱字	正) 写真撮影を行います。 誤) 写真撮影行います。
英字混入	タイプミスなどによる英字混入	正) 宜しくお願いします。 誤) 宜しくお願いしまs。

表 2: 学習データセット

データセット名	含まれるNGパターン	NG文作成方法	文数 [件]	データ量 [MB]
校閲済み過去原稿	-	-	428,716	46
作成NG文 I	漢字、助詞連続、脱字(送り仮名)	クラウドソーシングを利用して人手で作成	36,565	10
作成NG文 II	脱字(送り仮名)、脱字(助詞)	プログラムによる自動生成。平仮名をランダムに選んで除去	1,247,690	291

※1 原稿を句点などの終端文字で区切った 1 文を 1 件とする。ただし、作成NG文 I, II については OK/NG 文のペア数を表記。

4. アルゴリズム説明

本章では深層学習を用いた誤字脱字検出ロジックについて説明する。方針として、まず文字ごとの妥当性を判断する BLSTM モデルを構築し、その出力から文単位での正誤を判断するランダムフォレストを構築する。これらを組み合わせ、最終的には「誤字脱字箇所」、「正しい候補の文字」、「誤字脱字を含む文」の3つを出力する。

4.1 BLSTM による文字毎の OK/NG 確率モデル

本手法では、前方向からの文字の流れだけでなく、後方からの情報も捉えることができる BLSTM を採用した。文字毎の正常/異常を求めるニューラルネットワークのアーキテクチャーを図 1 に示す。BLSTM による出力は文字毎に順方向/逆方向の 2 つが存在するため、それらを結合し、各文字が正しいまたは誤字脱字である確率(OK/NG 確率)の 2 次元を出力するよう設計した。BLSTM 部分は順方向/逆方向で 2 層ずつ、計 4 層の中間層を持つ設計とした。損失関数にはクロスエントロピーを用い、文字毎に誤差を足し合わせた値を 1 文の誤差と定めた。推論の

際は、上記の枠組みで学習されたモデルを用いて各文字の OK/NG 確率を出力し、この出力結果を利用して「誤字脱字箇所」と「誤字脱字を含む文」を判定する。

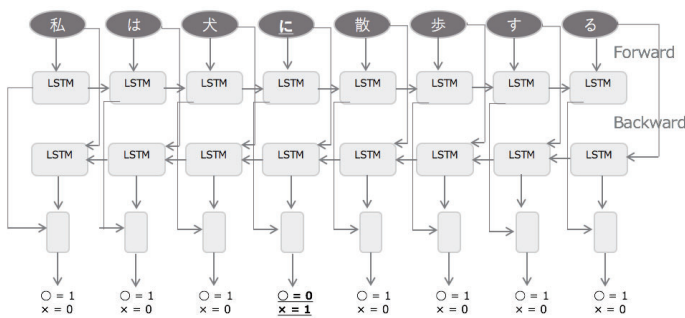


図 1: BLSTM による文字毎の OK/NG 確率モデル

4.2 BLSTM による言語モデル

BLSTM による言語モデルのアーキテクチャーを図 2 に示す。入力部分は基本的に図 1 と同様である。異なる点は以下 3 つである。

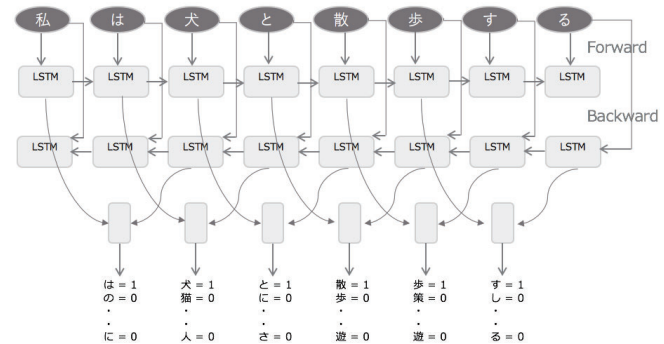


図 2: BLSTM による言語モデル

1. 学習に用いるデータの種類
学習に用いるデータは OK 文のみとした。言語モデルは正しい文から正しい文字の流れを予測するためである。
2. 最終層への入力を作成するロジック部分
4.1 の OK/NG 確率モデルとの違いは、最終層への入力を作成するロジック部分である。図 2 に示すように言語モデルの場合、t 番目の文字を予測するために、順方向 LSTM の t-1 番目の文字における出力値と、逆方向 LSTM の t+1 番目の文字における出力値を結合し、最終層への入力とする。言語モデルの場合、前後の文字から該当文字を予測するため、このような設計とした。
3. 最終層の出力
最終層の出力は基本的な言語モデル同様、文字サイズ分の次元を出力する設計である。

推論の際は、上記の枠組みで学習されたモデルを用いて、文を構成している文字に対する確率を出力し、この出力結果を利用して「誤字脱字箇所」と「誤字脱字を含む文」、「正しい候補の文字」を判定する。この言語モデルを用いる利点は、予想できていない誤字脱字を検出できる可能性がある点である。4.1 の

モデルのみでは機械的に作られた NG データを基にしているため、予想していない誤字脱字の検出力が弱くなる。それに比べ、言語モデルでは日本語として正しい文字の流れを学習するため、予想していない誤字脱字を検出できる可能性が高くなる。

4.3 複数モデルのアンサンブル方法

4.1, 4.2 で作成したモデルの出力値を使用して、入力文に誤字脱字を含むかどうかを判定する。使った変数を表 3 に示す。

表 3: ランダムフォレストの入力変数表

変数の説明	次元数
a) BLSTMモデルの出力する確率が最低となる箇所	1
b) a)における文字の確率	1
c) a)における文字の種別 ※1	6

※1: {平仮名,カタカナ,漢字,英字,数字,その他}のいずれかを示すone hot vector.

BLSTM モデル毎に上記を求め、全モデル分統合したものを入力とし、「誤字脱字を含む文か否か」の 2 値を分類するモデルを作成した。学習器にはランダムフォレストを使用した。推論時はそのランダムフォレストが出力する確率値と閾値の比較により判定する。このように複数の BLSTM モデルの出力値を用いた学習器を使用する事で、各々の BLSTM モデルの出力値に対する閾値をチューニングする必要がなくなる。さらに誤字脱字と判定する際の基準の選定の精度向上にも繋がり、精度面/保守面共に良いパフォーマンスとなる。

4.4 誤字脱字箇所推定と候補文字の決定

入力文に対しランダムフォレストが「誤字脱字を含む文」と判断した場合、「誤字脱字箇所の推定」と「候補文字の決定」を行う。誤字脱字箇所の推定は、各 BLSTM モデルの文字毎の確率が一定閾値以下となった箇所とする。誤字脱字箇所と推定された部分に対しては候補文字を決定する。候補文字は誤字脱字箇所において BLSTM 言語モデルの出力する確率が高い上位 3 文字とする。ただし、余分な文字が入っている、または脱字のような NG 文は、誤字脱字箇所を候補文字で置き換えるだけでは文の修正ができない点に注意が必要であり、今後の課題とする。

5. 実験

5.1 文単位の性能評価

評価に用いるモデルを表 4 に示す。LSTM 言語モデルは前方から後方へ向かう LSTM のみで構成した言語モデルである。BLSTM(言語, OK/NG 確率 I, II) モデルは 4.1, 4.2 で述べたモデル、アンサンブルモデルは BLSTM と 4.3 で述べたランダムフォレストで構成されるモデルを指す。BLSTM OK/NG 確率モデルは、学習データセットである作成 NG 文 I と II でサイズや内容が違うため、それぞれでモデルを分けた。

評価用データとして OK 文と NG 文を同数用意し、各モデルでの NG 文に対する検出率(True Positive Rate)と OK 文に対する誤検出率(False Positive Rate)で評価する。NG 文は表 1 に示す NG パターンごとに 200 文ずつ作成した。ここでは、句点等の終端記号で区切った単位を 1 文とし、NG 文 1 文あたり 1 つの誤字脱字を含むようにした。各モデルは 1 文ごとに誤字脱字を含む/含まないを判定する。アンサンブル以外のモデルは文字ごとに正しさ表す確率を出力するため、文に含まれるの全文字の確

率最低値と閾値との比較で判定する。アンサンブルモデルの出力は文単位での確率であるため、出力値と閾値の比較で判定する。

ROC 曲線と AUC を図 3, 表 5 に示す。表 5 より、言語モデル同士で LSTM と BLSTM を比較すると脱字(助詞)を除く全 NG パターンで BLSTM の方が AUC 値で上回っている。BLSTM モデル同士(言語, OK/NG 確率 I, II)の比較では、それぞれ得意な NG パターンが異なる。OK/NG 確率 I, II モデルは学習データに含まれる NG パターンに対して強く、その他の NG パターンに対して弱い。一方、言語モデルは漢字、英字混入に強い。特に英字混入は OK/NG 確率 I, II モデルの学習データに無い NG パターンであり、言語モデルの導入により未知の NG パターンに対応できる可能性がある。アンサンブルモデルは全体、漢字、助詞連続、脱字(送り仮名)について AUC が全モデル中最高値であり、他の NG パターンでも一定値を保っており、3 つのモデルを統合することで相補的な効果が得られている。

5.2 誤字脱字箇所推定と候補文字の評価

誤字脱字箇所推定と候補文字について評価結果を表 6 に示す。また、検出文の事例を表 7 に示す。評価用のモデルはアンサンブルモデルを採用し、NG 文検出の閾値は誤検出率=0.200 となる値を採用した。このとき検出率=0.795 である。

表 6 より、誤字脱字箇所推定の正解率は 90.6%と高い。候補文字の正解率は 62.5%である。BLSTM 言語モデルは注目する箇所の前後の文字を正として利用しているため、表 7 No.2 のような 1 文字間違いのケースでは正解率が高いが、No.1 のような連続する 2 文字が間違えるケースでは正解できないケースが多く見られた。

6. 試験運用

ゼクシィを対象に試験運用を実施した。ゼクシィ原稿の校閲者は校閲システムを利用し PC 画面上で確認や修正を行う。アンサンブルモデルを用いた誤字脱字検出ロジックを校閲システムに組み込み、試験運用した。試験期間中に投稿された原稿に対して校閲前後の文とアルゴリズムの検出結果を収集し、投稿された原稿を手で OK 文と NG 文に振り分け、1 文単位での検出率/誤検出率で評価した。ただし、収集した NG 文には表 1 で定義していない NG パターンも含まれる。試験運用は期間を 2 期に分け、それぞれで評価した。評価結果を表 8 に示す。また検出できた事例を表 7 に示す。

表 8 より、第 1 期の検出率は 60%、誤検出率は 11%である。試験運用と評価用データでは NG パターンの分布に差があると考えられるが、第 1 期では図 3(e)の ROC 曲線と比較しても妥当な結果となった。その一方で適合率(Precision)は 14%と低い。理由として NG 文に対して OK 文は 29 倍と多いことが挙げられる。

適合率は校閲者にとってシステムに対する心理的な信頼度に直結するため向上のための対策が必要である。学習用データセットは試験運用の半年前までに取得したものであるが、原稿の文章は時間経過に伴うトレンドの変化により徐々に変わっていくと考えられる。そこで、直近のデータを使用すれば適合率の向上が期待出来ると考え、第 1 期で収集した原稿を学習データとして新たに追加し、作成済みの BLSTM モデルをファインチューニングした。

表 8 に示す第 2 期はファインチューニング後のモデルを適用した結果である。第 1 期と比べ誤検出率が 6%まで低下し、適合率が 22%まで上昇することが確認できた。

7. まとめ

本研究では、複数の BLSTM モデルのアンサンブルによる誤字脱字検出システムの開発および実験を行った。先行研究と比較して、提案手法では、OK/NG 確率モデルに加え言語モデルを組み込むことで、想定していない誤字脱字パターンの検出が可能になり、検出後の候補文字の提案まで可能となった。

実験では、リクルートが保有するサービスの実データを利用して学習を行い、試験運用を行った。その結果、誤字脱字のない OK データが圧倒的に多数を占める状態のなかで誤検出 6%、適合率 22%という結果を得た。

現時点での課題として、脱字のような単純に文字置き換えでは対応できないパターンでの候補文字の提案手法の確立や検出精度向上のためのネットワーク構造の見直しがある。また、運用面では、実際にシステムを利用してもらうことで蓄積されるフィードバックデータを順次追加で学習をしていく仕組みの構築を行う。

参考文献

[1] L.Williams, Ronald J.; Hinton, Geoffrey E.; Rumelhart, David E. Learning representations by back-propagating errors, 1986

[2] Ilya Sutskever,James Martens,Geoffrey Hinton. Generating Text with Recurrent Neural Networks. 2011.

[3] Tsung-Hsien Wen, Milica Gasic, Nikola Mrkic, Pei-Hao Su, David Vandyke, Steve Young Semantically Conditioned LSTM-based Natural Language Generation for Spoken Dialogue Systems. 2015.

[4] Benjamin J. Radford, Leonardo M. de A. Paulino, Antonio J. Trias, Jim A. Simpson. Network Traffic Anomaly Detection Using Recurrent Neural Networks. 2018.

[5] Pankaj Malhotra, Lovekesh Vig, Gautam Shroff, Puneet Agarwal. Long Short Term Memory Networks for Anomaly Detection in Time Series. 2015.

[6] 中野 信. AI 技術を使った誤テロップ自動検出に関する技術検証. The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018.

[7] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. Neural computation, Vol. 9, No. 8, pp. 1735–1780, 1997.

[8] Mike Schuster and Kuldip K Paliwal. Bidirectional recurrent neural networks. IEEE Transactions on Signal Processing, Vol. 45, No. 11, pp. 2673–2681, 1997.

表 4: 評価モデル

モデル名	内容	使用学習データ
LSTM 言語モデル	単方向LSTMによる言語モデル	校閲済み過去原稿
BLSTM 言語モデル	Bidirectional LSTMによる言語モデル	校閲済み過去原稿
BLSTM OK確率モデル I	Bidirectional LSTMによるOK/NG確率モデル	作成NG文 I
BLSTM OK確率モデル II	Bidirectional LSTMによるOK/NG確率モデル	作成NG文 II
アンサンブルモデル	上記3種類のBLSTMモデルとランダムフォレストで構成	作成NG文 I + II

表 5: 文単位性能評価結果 (AUC)

モデル名	全体	漢字	助詞連続	脱字 (送り仮名)	脱字 (助詞)	英字混入
LSTM 言語モデル	0.79	0.89	0.59	0.8	0.7	0.95
BLSTM 言語モデル	0.83	0.96	0.62	0.88	0.62	0.99
BLSTM OK/NG確率モデル I	0.76	0.88	0.89	0.95	0.51	0.58
BLSTM OK/NG確率モデル II	0.77	0.59	0.62	0.94	0.87	0.77
アンサンブルモデル	0.88	0.97	0.89	0.96	0.77	0.88

黄色は特定NGパターンについて他モデルと比べ高い箇所

表 6: 指摘箇所推定と候補文字評価結果

NG文数	NG文 検出数	誤字脱字箇所 推定正解数 ※1	誤字脱字箇所推 定正解率 ※1	候補文字評価 対象数 ※2	候補文字 正解数 ※2	候補文字 正解率 ※2
1000	795	720	0.906	376	235	0.625

※1：推定した誤字脱字箇所中に真の誤字脱字箇所を含むとき正解とみなす。
※2：候補文字3文字中に正しい文字を含むとき正解とみなす。ただし4.4節で述べた制約があるため、NG箇所の修正前後で文字数が等しい文のみ対象とする。

表 7: 検出成功事例

No.	評価データ	NG文と推定誤字脱字箇所	OK文	候補文字
1	作成NG文	フェアに【命】【下】して結婚式のイメージを膨らませてみて！ この機会にシェフ渾身のお料理【を】ご堪能ください。	フェアに参加して結婚式のイメージを膨らませてみて！ この機会にシェフ渾身のお料理をご堪能ください。	[参, 面, 間], [を, と, そ]
2	試験運用時 原稿	より格調高いしつら【れ】に変わり、 いっそう厳かな雰囲気になる神殿。	より格調高いしつらえに変わり、 いっそう厳かな雰囲気になる神殿。	[を, で, も]
3		海外にいるかのような4【つ】 【へ】ーティ空間	海外にいるかのような4つのパー ティ空間	[え, い, う], [パ, ホ, の]

□ は推定誤字脱字箇所、候補文字は □ 1つあたり3文

表 8: 試験運用結果

試験運用 フェーズ	試験運用期間	OK文 件数	OK文 検出件数	誤検出率	NG文 件数	NG文 検出件数	検出率	適合率
第1期	2018/7/23～ 2018/8/20	1207	146	0.121	41	24	0.585	0.141
第2期	20180831～ 20180914	1949	122	0.063	51	35	0.686	0.223

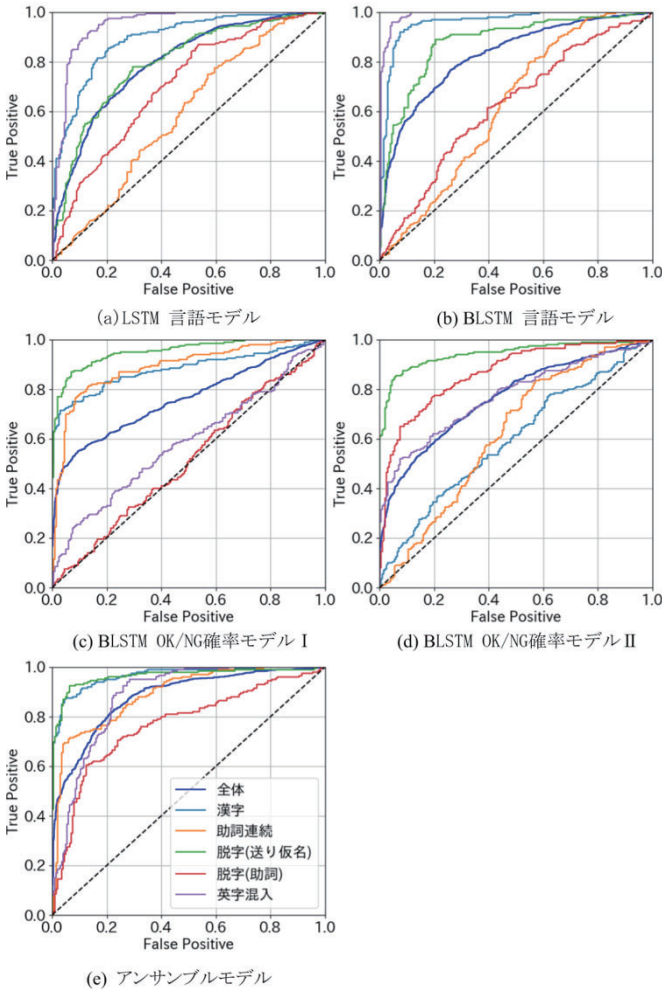


図 3: 文単位性能評価結果 (ROC 曲線)

[3H4-J-7] Agents: intelligence in/among robots

Chair:Keisuke Otaki Reviewer:Hidekazu Oiwa

Thu. Jun 6, 2019 3:50 PM - 5:10 PM Room H (303+304 Small meeting rooms)

[3H4-J-7-01] Effect of Robot Anxiety on the Appearance Tendency of Uncanny ValleyKazuhiro Ikeda¹, OTomoko Koda¹ (1. Osaka Institute of Technology)

3:50 PM - 4:10 PM

[3H4-J-7-02] Strategyproof Mechanism with Agents Grouping for Multi-Agent PathfindingOManao Machida¹ (1. NEC)

4:10 PM - 4:30 PM

[3H4-J-7-03] A Fundamental Study of Region Allocation for Mobile Robots Based on Constraint Optimization and Decentralized Solution MethodOToshihiro Matsui¹ (1. Nagoya Institute of Technology)

4:30 PM - 4:50 PM

[3H4-J-7-04] On the design of state value functions for real-time continuous-state space multi-agent decision makingOTomoharu Nakashima¹, Harukazu Igarashi², Hidehisa Akiyama³ (1. Osaka Prefecture University, 2. Shibaura Institute of Technology, 3. Fukuoka University)

4:50 PM - 5:10 PM

対ロボット不安特性による不気味の谷の出現傾向の分析 Effect of Robot Anxiety on the Appearance Tendency of Uncanny Valley

池田 和広
Kazuhiro Ikeda

神田 智子
Tomoko Koda

大阪工業大学情報科学部
Department of Information Science and Technology, Osaka Institute of Technology

In this study, we analyzed the influence of robot anxiety characteristics on the appearance tendency of uncanny valley. We used crowdsourcing for the questionnaire survey of mechano-humanness score (MH score) and likeability of 80 robot face images. Then we divided the participants into two groups according to their scores of Robot Anxiety Scale (RAS). The results of t-test of the fitted curves using the MH scores and likeability scores showed that the appearance tendency of the uncanny valley is affected by users' robot anxiety scale. Those who have less anxiety toward robots showed higher affinity toward the robot faces, while those with high anxiety showed lower affinity toward the same faces.

1. はじめに

社会的対話ロボットが日常生活で使用され、我々の生活に浸透しつつある。日本は米国やドイツと比較すると、「ロボット=人型」「ロボット=コミュニケーション可能」と認識される割合が高いことが示されている [1]。産業用ロボットではロボットが顔を持つことは必要ではないが、社会的対話ロボットでは、ロボットの外見がユーザである人間に与える影響を考える事が重要である。森はロボットの外見が人間的になるほど親近感は向上するが、人間に近づくある段階で一旦親近感が下がる「不気味の谷現象」を提唱した [2]。Mathur らはアメリカ人を対象に、ロボットの顔画像 80 枚を用いて、それぞれの顔を機械的-人間的の尺度に数値化(mechano-humanness (MH) スコア) し、それぞれのロボットの顔画像に対する親近感を評価するアンケート調査を実施し、MH スコアと「不気味の谷現象」の関連を示した [3]。

一方、ロボットとのインタラクションに対して不安を感じる人や否定的な感情を抱く人が存在していることから、野村らは、実験参加者のロボットに対する不安を評価する指標である対ロボット不安尺度 (RAS) を提唱している [4]。RAS を用いて実験参加者を分類した我々の先行研究では、受付ロボットの外見・振る舞いを変化させ、実験参加者にロボットと対話を行ってもらった。その結果、ロボットとの会話に不安を感じる RAS 高群は、ロボットの外見により親近感が変化するものの、不安に感じない RAS 低群ではロボットの外見によって親近感は変化しない事を示した [5]。

これらのことから、本研究では「不気味の谷現象」の出現傾向は、実験参加者の対ロボット不安の程度により影響を受けると考える。アメリカ人の不気味の谷を検証した [3] では実験参加者を RAS 得点により分類していない。従って、本研究では、日本人の実験参加者を募り、ロボットに対し不安を感じる RAS 高群、不安に感じない RAS 低群に分類し、[3] と同様の実験を行う。

予測される不気味の谷の出現傾向は、RAS 低群が RAS 高群より外見に対する親近感が高い「上下並行移動型」、RAS 高群が RAS 低群より MH スコアの低い段階で親近感が低下し、MH スコアの高い段階で上昇する「左右移動型」、RAS 低群は RAS 高群と比べ親近感があまり変化しない「山谷の深さ変化型」とした。

2. ロボットの顔画像の印象評価実験

2.1 ロボットの顔画像の機械的-人間的尺度の評価

ロボットの顔画像の MH スコアを算出するため、Yahoo!クラウドソーシングを利用し、100 人の日本人に、[3] で用いられた 80 枚のロボットの顔画像 (図1) に対する MH スコアアンケート (「機械的」 -100 ~ 「人間的」 +100 の値) に回答してもらった。回答から、1~4 種類の数字のみで回答している、負の値を使用していない、などの不適切なデータ約 20% を除外し、79 人分のデータを分析対象とした。分析対象の参加者の属性は、平均年齢 39.2 歳、標準偏差 9.8、男性 62 名、女性 17 名であった。

回答として得られた実験参加者の MH スコアの平均を、各ロボットの MH スコアとした。MH スコアは最小値 -90.63、最大値 90.33、中央値 -37.37、標準偏差 53.10 であった。同じロボットの顔画像を用いてアメリカ人が評価した [3] では、MH スコアが最小値 -97.21、最大値 93.21、中央値 -49.18、標準偏差 61.50 であり、両調査とも同様の傾向を示した。

2.2 ロボットの顔画像の親近感の評価

2.1 と同じ 80 枚のロボットの顔画像に対する印象評価アンケートとして、ロボットに対する「親近感」、「楽しさ」、「不気味さ」(それぞれ、「全く感じない」 -100 ~ 「とても感じる」 +100 の値で回答)、および実験参加者の RAS を評価するアンケート調査を、Yahoo!クラウドソーシング上で 300 人を対象に実施した。回答から、1~4 種類の数字のみで回答している、負の値を使用していない、などの不適切なデータ約 55% を除外し、139 人分のデータで以下の分析を行った。分析対象の参加者の属性は、平均年齢 41.6 歳、標準偏差 10.3、男性 93 名、女性 46 名であった。実験参加者を RAS 得点の平均値 (41.57) により二群に分類し、RAS 得点が平均値より高い実験参加者を RAS 高群 (n=74)、低い実験参加者を RAS 低群 (n=65) とした。

次に、MH スコアを横軸とし、「親近感」、「楽しさ」、「不気味さ」を縦軸とした散布図を作成し、近似曲線を得た。図2に親近感と MH スコアの近似曲線を示す。その結果、どの評価指標においても、[3] における結果と同じく、不気味の谷の存在を示す曲線が得られた。

RAS 高群・低群それぞれの近似曲線全体間で、近似曲線の幅を標準偏差で定義し、各データ点における近似曲線の差の合計を算出し、Welch の t 検定を行った結果、「親近感」では

RAS 低群が RAS 高群より高い傾向にあり ($p=0.056$), 「楽しさ」では RAS 低群が RAS 高群より有意に高く ($p<0.05$), 「不気味さ」では RAS 高群が RAS 低群より有意に高い ($p<0.01$) が示された。

近似曲線において, MH スコアが上がるにつれ親近感が上がり始め, 極大値に至った後下がり始める MH スコア $-60 \sim -30$ 間における, RAS 両群の近似曲線の t 検定を行ったところ, RAS 低群の親近感スコアが高群より有意に高い ($p<0.01$) が示された。しかしながら, 不気味の谷に落ち始めてから, 谷底に至り親近感が上がり始める MH スコア $+30 \sim +60$ の近似曲線間では, RAS 両群の親近感に有意差は見られなかった。同様の分析を, 「不気味さ」評価で行ったところ, 極大値前後の MH スコア $-60 \sim -30$ 間でも ($p<0.05$), 極小値前後の MH スコア $+30 \sim +60$ 間でも ($p<0.01$), RAS 高群の不気味さ評価が低群と比較して有意に高いことが示された。

しかしながら, 使用したロボット顔画像 80 枚中, MH スコアが正の値を示す画像数が, 負の値を示す画像数より少ないこと, 特に不気味の谷付近である $+20 \sim +50$ の画像が少ないことから (図 2 参照), MH スコアの分布のばらつきを防ぐために画像数を増やして評価実験を行う必要があると考える。

RAS 高群・低群間でロボットの顔画像の「不気味さ」に最も顕著な有意差が出ていることから, RAS 得点はロボットに対する不安を測る指標であるため, 「親近感」より「不気味さ」の評価とより深い関連があることを示唆していると考えられる。また, RAS 高群・低群の「親近感」, 同様の傾向を示す「楽しさ」, RAS 高群と低群の評価差がより顕著に現れる「不気味さ」の近似曲線より, 本実験で使用したロボットの顔画像の実験結果に関しては, 「上下移動型」に近い傾向であることを示していると考えられる。

最後に, 本実験とアメリカ人参加者による [3] の近似曲線を元に, 日米の比較を行う。本実験における日本人の親近感の極大値 $+7$ の MH スコアは -45 , 極小値 -20 の MH スコアは 28 であった。アメリカ人による評価実験では, 親近感の極大値 $+7$ の MH スコアは -66 , 極小値 -43 の MH スコアは 36 であったことから, 実験で用いたロボット顔画像では, アメリカ人参加者の方が日本人参加者と比較すると, 親近感が減少する MH スコア区間が長く, 不気味の谷が深い傾向がある可能性を示唆している。このことは, 日本においてロボット=人型と認識されていることと関連があるのではないかと考える。

3. おわりに

本研究では「不気味の谷現象」の出現傾向は, 実験参加者の対ロボット不安の程度により影響を受けると考え, クラウドソーシングにより日本人の実験参加者を募り, 80 枚の様々なロボットの顔画像を用いて親近感の評価実験を行った。また, 対ロボット不安尺度 (RAS) を用いて, 実験参加者をロボットに対して不安を感じる RAS 高群, 不安に感じない RAS 低群に分類し, 親近感の変化を比較した。その結果, 不気味の谷の出現傾向は両群で異なり, 対ロボット不安により影響を受ける事が示された。RAS 高群・低群の「親近感」の近似曲線において, 低群の近似曲線が高群の近似曲線より上に位置することから, 本実験の範囲内では「上下移動型」に近い傾向を示していることが示唆された。

今後の展望として, クラウドソーシング上で精度の高いアンケート結果を得るため一人の回答数を減少させる等の方策をとること, また, MH スコアのばらつきを防ぐ為, ロボットの顔画像を増加させて実験を行う必要があると考える。



図 1. MH スコアおよび親近感の評価実験に使用したロボット顔画像 80 枚 [3]

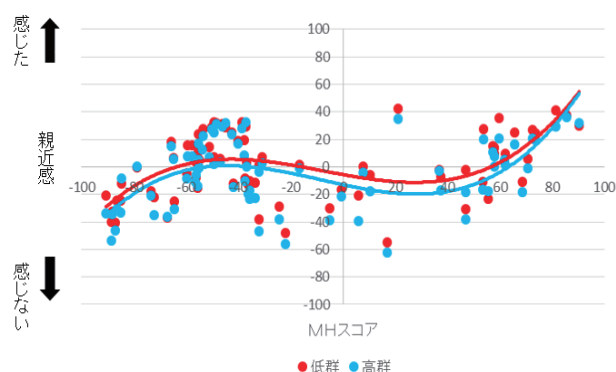


図 2. RAS 得点別の親近感の散布図と近似曲線

謝辞

本研究の一部は, 科研費「基盤(C) JP17K00287」の交付を受けて実施した。

参考文献

- [1] 日戸浩之, 谷山大輔, 稲垣仁美. ロボット・AI 技術の導入をめぐる生活者と受容性と課題～日米独 3 カ国調査. NRI 知的資産創造, pp. 108-125, 2016.
- [2] 森政弘. 不気味の谷, Energy 第 7 巻第 4 号, pp. 33-35, 1970.
- [3] Maya B. Mathur, David B. Reichling. Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley. Cognition, Volume 146, pp. 22-32, 2016.
- [4] 野村竜也, 神田崇行, 鈴木公啓, 山田幸恵, 加藤謙介. Human-Robot Interaction (HRI)における人の態度・不安・行動. Fuzzy System Symposium 2010, pp. 554-559, 2010.
- [5] 中川弘也, 神田智子. 受付ロボットの積極性及び外見が与える印象のユーザ属性別分析. HAI シンポジウム 2017, 2017.

マルチエージェント経路計画のための エージェントグループ化を伴う正直申告メカニズム

Strategyproof Mechanism with Agents Grouping for Multi-Agent Pathfinding

町田 真直

Manao Machida

NEC

This paper proposes a computationally efficient strategyproof mechanism that solves multi-agent pathfinding (MAPF) problems with heterogeneous and self-interested agents. In MAPF, agents need to reach their goal destinations while avoiding collisions between them. MAPF solvers assign agents non-conflicting paths that minimize the global cost function (e.g., the sum of travel costs). Finding the optimal solution of a MAPF problem is an NP-hard problem. Mechanism design aims to design mechanisms in which the selfish behavior of agents leads to a socially optimal outcome. The Vickrey-Clarke-Groves (VCG) mechanism is a well known mechanism that is efficient and strategyproof. However, VCG must compute an optimal outcome to be strategyproof. In this paper, we consider heterogeneous and self-interested MAPF. Agents are heterogeneous if the costs of traversing a given path differ between agents. In particular, we assume each agent has a private linear cost function of travel time. The proposed mechanism divides agents into several groups depending on their declarations and thresholds, and then computes an optimal outcome for each group sequentially. The thresholds are parameters for adjusting computational complexity of this mechanism. We show that this mechanism has trade-off between the solution quality and run-time by numerical simulations.

1. はじめに

マルチエージェント経路計画問題 (MAPF) はグラフとエージェントからなる。各エージェントは、それぞれのスタートノードからゴールノードまで、他のエージェントと衝突することなく、グラフ上を移動する必要がある。MAPF の目的は、エージェント間で衝突がなく、コスト (例えば、各エージェントの移動コストの合計) が最小である各エージェントの経路を解として得ることである。MAPF は、交通管制、ビデオゲーム、配車ルートの決定等、様々な応用分野を持つ [Silver 2005, Standley 2010, Ryan 2010, Dresner 2008, Kiesel 2012]。

メカニズムデザインは、エージェントの利己的な行動が社会的に望ましい結果を生むメカニズムをデザインすることを目的とする。特に、Vickrey-Clarke-Groves (VCG) は、耐戦略性と効率性を満たすメカニズムとしてよく知られている [Vickrey 1961, Clarke 1971, Groves 1973]。しかし、VCG では耐戦略性を満たすために、最適な結果を算出する必要がある。

従来の MAPF では、各エージェントは協力的であることが仮定されている。近年、この協力型の MAPF に対して、利己的なエージェント間での MAPF の研究も行なわれている [Bnaya 2013, Machida 2019, Amir 2015]。利己的なエージェント間の MAPF では、各エージェントが自身のコストを最小化するために行動する状況で、いかにグローバルコストを最小化するかが問題となる。[Amir 2015] は、VCG オークションと MAPF の関係を明らかにし、競り上げ式組合せオークションを用いて利己的なエージェント間の MAPF を解く手法を提案している。また、組合せオークションは、与えられた経路へのコストがエージェントによって異なる、異質なエージェント間の MAPF に対しても適用することができる。しかし、組合せオークションにおける勝者決定問題 (各エージェントへの経路の割当問題) は、NP 困難である。

[Machida 2019] では、多項式時間の正直申告メカニズム SCA* を提案している。しかし、解の品質は、VCG と比べて悪くなってしまうという欠点がある。

本稿では、各エージェントが移動時間に対して異なる線形な移動コストを持つ状況において、SI-MAPF を解く計算効率の良い正直申告メカニズムを提案する。提案メカニズムでは、各エージェントに単位時間当たりの移動コストを宣言させる。そして、メカニズム側で設定した閾値と、エージェントの宣言に基づいてエージェントをいくつかのグループに分け、逐次、グループごとに最適な経路割当を決定する。ここで、閾値は、計算量と解の品質を調整するためのパラメータである。また、数値実験により、異なる閾値を設定した提案メカニズムを比較し、解の品質と計算時間についてトレードオフが存在することを示す。

2. 問題設定

MAPF は、グラフ $G = (V, E)$ とエージェントの集合 $K = \{1, \dots, k\}$ からなる。エージェント $i \in K$ はスタートノード $s_i \in V$ とゴールノード $g_i \in V$ を持つ。各エージェントは、1つのタイムステップに、現在のノードに留まるか、隣接するノードに移動することができる。エージェント i の経路を、 P_i で表す。ここで、 P_i はエージェントの存在するノードと時間の組の集合である。ただし、時間の集合は $T = \{0, 1, 2, \dots\}$ とする。また、各エージェントは他のエージェントと衝突なく移動する。 P_i の時刻 n に対応する要素を $(p_i(n), n)$ で表す。経路 P_i, P_j が衝突するとは、 $n \in T$ が存在して、 $p_i(n) = p_j(n)$ または $p_i(n) = p_j(n+1) \wedge p_i(n+1) = p_j(n)$ が成り立つことである。前者はエージェント i, j が同時刻に同じ位置に存在すること、後者はエージェント i, j が交差することに対応する。MAPF の解を、 $P = \{P_1, \dots, P_k\}$ で表す。また、解の集合を \mathcal{P} で表す。

本稿で扱う問題設定では、エージェントが利己的かつ、異なる時間に線形な移動コストを持つとする。エージェント i の単

連絡先: 町田 真直, NEC, 〒 211-8666, 神奈川県川崎市中原区下沼部 1753, 044-435-5678, manao-machida@ap.jp.nec.com

位ステップ当たりの移動コストをタイムステップコストと呼び、 $w_i \in W$ で表す。ただし、 $W = [0, \bar{w}]$ はタイムステップコストの集合であり、 \bar{w} はタイムステップの最大値である。例えば、このコストを燃料費だとみれば、トラックが単位時間あたりに消費する燃料 w_{truck} が、バイクが単位時間あたりに消費する燃料 $w_{motorcycle}$ の 2 倍であるならば、 $w_{truck} = 2 \cdot w_{motorcycle}$ と設定すればよい [Amir 2015]。

経路 P_i にしたがって移動したときのエージェントの移動コストは、ゴールに到着するまでのタイムステップ数を $n(P_i)$ とし、 $w_i n(P_i)$ で与えられる。また、形式的には、

$$n(P_i) = \min \{n : \forall n' \geq n, (g_i, n') \in P_i\}$$

である。

さらに、本稿ではエージェントのコストは準線形であると仮定する。つまり、経路 P_i にしたがって移動するために、金銭支払い t が必要であるとき、経路 P_i にしたがうエージェント i のコストは、

$$w_i n(P_i) + t$$

で与えられる。この金銭支払いは、例えば、経路の通行料 [Bnaya 2013] や、その経路を得るためのオークションでの支払い [Amir 2015] に対応する。利己的なエージェントは、自身のコストを最小化することを目的として行動する。

3. 提案メカニズム

本節では、タイムステップコストの異なる利己的なエージェント間の MAPF を解く、エージェントグループ化を伴う正直申告メカニズムを提案する。各エージェント $i \in K$ は自身のタイムステップコスト $b_i \in W$ を宣言し、メカニズムは宣言に応じて経路割当 $P \in \mathcal{P}$ と各エージェントの支払い t を決定する。

メカニズムを、 $f = (\beta, \tau)$ で表す。ここで、 $\beta : W^k \rightarrow \mathcal{P}$ は経路割当関数、 $\tau : W^k \rightarrow \mathbf{R}^k$ は金銭支払い関数である。提案メカニズムでは、タイムステップ集合の区間 $[0, \bar{w}]$ に複数の閾値を設定し、閾値に基づいてエージェントを分割する。そして、順番に、グループごとに最適な経路を計画していく。ただし、後に決定される経路は、既に決定された経路と衝突しないように計画する。

MAPF における探索空間は $O(|V|^k)$ であり、エージェント数に対して指数的に増加する。提案メカニズムでは、この探索をグループごとに分割することにより、探索空間の削減を行なう。グループ内のエージェント数が k' であるとき、グループ内での探索空間は $O(|V|^{k'})$ である。そのため、提案メカニズムでは、閾値の設定によってグループ内のエージェント数を減らすことで、計算時間を削減することができる。

閾値を、 $c = (c_1, \dots, c_{m+1})$ とする。 c は、 $c_1 > \dots > c_{m+1}$ かつ $c_1 = \bar{w} \wedge c_{m+1} = 0$ を満たす。分割されたエージェントの各グループを、 $i \in \{1, \dots, m-1\}$ について $C_i(b) = \{j \in K | b_j \in (c_{i+1}, c_i]\}$ 、 $C_m(b) = \{j \in K | b_j \in [c_{m+1}, c_m]\}$ で表す。また、エージェントの部分集合 $K' \subseteq K$ の経路を $P^{K'}$ 、 $\hat{\mathcal{P}}(P^{K'})$ を $P^{K'}$ と衝突のない $K \setminus K'$ に属するエージェントの経路集合とする。

提案メカニズムを擬似コード 1 に示す。1 行目が各グループごとに順次処理していくことに対応し、2 行目が処理内容である。各グループの経路は、既に計画された経路と衝突せず、宣言したタイムステップコストに基づいてグループ内の総コストが最小であるよう決定される。特に閾値 $c = (\bar{w}, 0)$ である

Mechanism 1 提案メカニズム

- 1: **for** $i = 1, \dots, m$ **do**
- 2: plan $\beta^{C_i(b)}(b)$ such that satisfies the following condition

$$\sum_{j \in C_i(b)} b_j n(\beta_j^{C_i(b)}(b)) = \min_{P \in \hat{\mathcal{P}}(\beta^{\cup_{l < i} C_l(b)}(b))} \sum_{j \in C_i(b)} b_j n(P_j) \quad (1)$$

- 3: **end for**

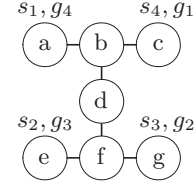


図 1: 例

とき、提案メカニズムの経路割当は VCG と一致する。また、 $|\max_i C_i(b)| = 1$ であるとき、提案メカニズムの経路割当は SCA* と一致する。

図 1 を用いて、提案メカニズムの処理の例を示す。提案メカニズムは閾値 $c = (1, 0.5, 0)$ を持ち、図 1 のグラフ上に存在するエージェントがそれぞれ、タイムステップコスト $b_1 = 0.8, b_2 = 0.2, b_3 = 0.4, b_4 = 0.6$ を宣言するとする。このとき、エージェントは 2 つのグループ、 $C_1(b) = \{1, 4\}, C_2(b) = \{2, 3\}$ に分割される。提案メカニズムは、まず $C_1(b)$ について、最適な経路を計画する。すなわち、

$$\begin{aligned} P_1 &= \{(a, 0), (a, 1), (b, 2), (c, 3), \dots\}, \\ P_4 &= \{(c, 0), (b, 1), (d, 2), (b, 3), (a, 4), \dots\} \end{aligned}$$

を計画する。その後、 P_1, P_4 と衝突しないうちで、 $C_2(b)$ について、以下の最適な経路を計画する。

$$\begin{aligned} P_2 &= \{(e, 0), (e, 1), (f, 2)(d, 3), (f, 4), (g, 5), \dots\} \\ P_3 &= \{(g, 0), (g, 1), (g, 2), (f, 3), (e, 4), \dots\} \end{aligned}$$

P_2 は、 $(d, 2)$ で P_4 と衝突しないよう、スタート地点でワンステップの待機を行なう。そのため、 P_2 は P_4 よりワンステップ、 P_3 は P_1 よりワンステップ、ゴールに到着するのが遅くなる。各グループ i の取りうる経路集合を、

$$\hat{\mathcal{P}}^{C_i(b)}(b) = \hat{\mathcal{P}}(\beta^{\cup_{l < i} C_l(b)}(b)) \quad (2)$$

で表す。また、以下の関数を定義する。

$$M(i, b'_i, j, b) = \min_{P \in \hat{\mathcal{P}}^{C_j(b')}(b')} \sum_{l \in C_j(b) \cup \{i\}} b'_l n(P_l), \quad (3)$$

ただし、 $b' = (b_1, \dots, b_{i-1}, b'_i, b_{i+1}, \dots, b_k)$ である。 $M(i, b'_i, j, b)$ は、申告 b がなされているとき、エージェント i が申告を b'_i に変更したときの、集合 j に属するエージェントとエージェント i の移動コストの和を表している。

関数 M を用いて、支払い関数を以下で与える。

$$\begin{aligned} \tau_i(b) &= -b_i n(\beta_i(b)) \\ &\quad + M(i, b_i, c^-(b_i), b) - M(i, c_{c^-(b_i)+1}, c^-(b_i), b) \\ &\quad + \sum_{j > c^-(b_i)} (M(i, c_j, j, b) - M(i, c_{j+1}, j, b)) \quad (4) \end{aligned}$$

ただし, $c^-(b_i)$ は, $i \in C_{c^-(b_i)}(b)$ を満たす値である. 式 (4) は, 特に閾値 $c = (\bar{w}, 0)$ であるとき, VCG における支払いと一致する. また, $c = (\frac{1}{m})_{l \in \{0, \dots, m\}}$ であるとき, $M > 0$ が存在して $m = M$ のとき $|\max_i C_i(b)| = 1$ ならば, $\lim_{m \rightarrow \infty} \tau(b)$ は SCA* の支払いと一致する.

提案メカニズムが, 以下の耐戦略性を持つことを示す.

定義 1 メカニズム $f = (\beta, \tau)$ が耐戦略性を持つとは, 任意の $i \in K, b \in W^k$ について, 以下が成り立つときをいう.

$$w_i n(\beta_i(w_i, b_{-i})) + \tau_i(w_i, b_{-i}) \leq w_i n(\beta_i(b)) + \tau_i(b)$$

ただし, $(w_i, b_{-i}) = (b_1, \dots, b_{i-1}, w_i, b_{i+1}, \dots, b_k)$ である.

メカニズムが耐戦略性を持つとき, 各エージェントは, 自身のタイムステップコストを正直に申告するときコストが最小であり, 嘘の申告する誘因を持たない.

以下の定理が成り立つ.

定理 1 提案メカニズム $f = (\beta, \tau)$ は耐戦略性を持つ.

証明 [Archer 2001] より, 任意の $i \in K, b \in W^k$ について, $b_i \geq b'_i \implies n(\beta_i(b)) \leq n(\beta_i(b'_i, b_{-i}))$ が成り立ち, かつ,

$$\tau_i(b) = -b_i n(\beta_i(b)) + \int_0^{b_i} n(\beta_i(c, b_{-i})) \quad (5)$$

であるならば, $f = (\beta, \tau)$ は耐戦略性を持つ.

まず, $n \circ \beta$ が単調非増加であることを示す. $b_i \geq b'_i$ について, $C(b) = C(b'_i, b_{-i})$ が成り立つとき, 明らかに $n(\beta_i(b)) \leq n(\beta_i(b'_i, b_{-i}))$ が成り立つ. エージェント i が b_i を宣言したとき, 属するグループが $C_j(b)$ であるとして, $C_j(b) \setminus \{i\} = C_j(b'_i, b_{-i})$ ならば,

$$\begin{aligned} \sum_{l \in C_j(b)} b_l n(\beta_l(b)) &\leq \sum_{l \in C_j(b)} b_l n(\beta_l(b'_i, b_{-i})) \\ \sum_{l \in C_j(b)} b_l n(\beta_l(b)) &\leq b_i n(\beta_i(b'_i, b_{-i})) + \sum_{l \in C_j(b'_i, b_{-i})} b_l n(\beta_l(b'_i, b_{-i})) \\ b_i n(\beta_i(b)) &\leq b_i n(\beta_i(b'_i, b_{-i})) \end{aligned}$$

である. よって, 任意の $i \in K$ について $n \circ \beta_i$ は W 上で単調非増加である.

式 (5) から式 (4) が導出されることを示す. $c_j > b_i \geq b'_i > c_{j+1}$ について, $\bar{c} \in [b'_i, b_i]$ が存在して, 任意の $c < \bar{c}$ について $n(\beta_i(c, b_{-i})) = n(\beta_i(b'_i, b_{-i}))$, かつ, 任意の $c > \bar{c}$ について $n(\beta_i(c, b_{-i})) = n(\beta_i(b))$ が成り立つとする. このとき,

$$\begin{aligned} \int_{b'_i}^{b_i} n(\beta_i(c, b_{-i})) &= (\bar{c} - b'_i) n(\beta_i(b'_i, b_{-i})) + (b_i - \bar{c}) n(\beta_i(b)) \\ &= M(i, \bar{c}, j, b) - M(i, b'_i, j, b) \\ &\quad + M(i, b_i, j, b) - M(i, \bar{c}, j, b) \\ &= M(i, b_i, j, b) - M(i, b'_i, j, b) \end{aligned}$$

が成り立つ. よって, 式 (5) から式 (4) が導出される.

4. 数値実験

本節では, 数値実験により提案メカニズムを評価する. 特に, 分割するグループ数異なる提案メカニズムの比較を行なう. f_s を, $W = [0, 1]$ を s 等分にする閾値を持つメカニ

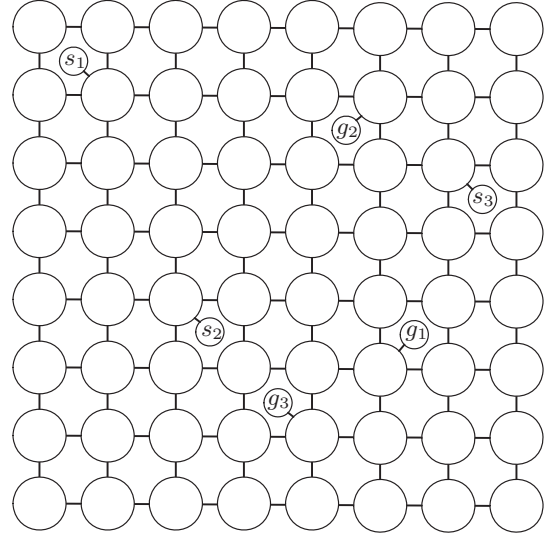


図 2: エージェントのスタート・ゴールがグリッドに隣接する, 8×8 グリッドマップ

ズムとする. 例えば, f_2 は閾値 $(1, 0.5, 0)$ を持ち, f_4 は閾値 $(1, 0.75, 0.5, 0.25, 0)$ を持つ. また, f_1 は VCG メカニズムと一致する.

MAPF の問題設定として, 以下のものを用いる. 各エージェント $i \in K$ のタイムステップコスト w_i は, $[0, 1]$ 上の一様分布に従ってランダムに与える. また, マップは 8×8 の 4 隣接グリッドマップとし, 各エージェントのスタートとゴールはグリッドマップに属さず, 1 つのグリッドのみに隣接するノードとする. マップの例を図 2 に示す. さらに, 任意の $i \in K$ について, $j \neq i$ は s_i, g_i に進入禁止であるとする. 各エージェントが他のエージェントのスタートとゴールを通過できないという問題設定は, 各エージェントのスタートやゴールが私有地であることに対応する. 例えば, 配送ドローンは, 配送拠点であるスタート地点から, 配送先の住宅/ビルであるゴール地点へ移動する. このとき, スタート・ゴールともに私有地であり, 進入の許可を得ていない他のエージェントは, これらの場所を通過することはできない.

また, 提案メカニズムで用いる最適経路計画アルゴリズムとして, Conflict Base Search[Sharon 2015] を用いた.

表 1 は, エージェント数が 5 から 10 であるときについて, ランダムに生成した 100 問の, 各提案メカニズムの平均の計算時間を示している. 分割するグループ数が多いメカニズムの方が計算時間が短い. 特に, エージェント数 8 から 10 にかけて, f_{16} の計算時間は f_1 (VCG) の 50 分の 1 以下である.

図 3, 4 はそれぞれ, エージェント数が 5 から 10 であるときについて, ランダムに生成した 100 問の, 1 エージェント当たりの平均の移動コスト, 総コストを示している. ここで, 総コストは, 移動コストと支払い額の合計である. 図 3, 4 の両方において, 分割するグループ数が少ない提案メカニズムのコストが小さい. また, f_1 (VCG) とそれ以外のメカニズムを比較すると, 移動コストよりも総コストにおいてコストが大きく異なる. すなわち, f_1 と比較して, f_s ($s \geq 2$) は支払額も大きい.

表 1, 図 4 より, 提案メカニズムは解の品質と計算時間についてトレードオフを持ち, 閾値の選択によって解の品質・計算時間のバランスの調整を行なうことができることがわかる.

表 1: 8×8 グリッドマップ上, w が $[0, 1]^k$ 上の一様乱数のときの Runtime (ms)

k	f_1	f_2	f_4	f_8	f_{16}
5	15.46	6.57	4.67	4.09	4.37
6	11.43	6.20	5.01	4.39	2.66
7	299.22	28.44	22.66	13.86	9.38
8	163.82	22.20	14.19	8.77	4.07
9	121.57	19.64	14.54	8.47	4.89
10	561.95	63.04	30.48	10.68	7.33

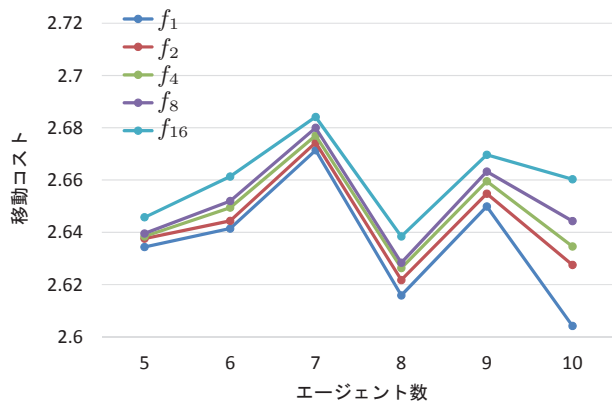


図 3: 8×8 グリッドマップ上, w が $[0, 1]^k$ 上の一様乱数のときの, 1 エージェント当たりの移動コスト

5. まとめ

本稿では, 各エージェントが利己的かつ移動時間に対して線形なコストを持つ際に, MAPF を解く計算効率の良い正直申告メカニズムを提案した. 提案メカニズムでは, 各エージェントに単位時間当たりの移動コストを申告させて, 設定した閾値によりエージェントをグループ分けし, グループごとに逐次最適な経路割当を計算することにより, 計算を高速化する. また, 数値実験により閾値の異なる提案メカニズムを比較し, 閾値によって計算時間を調整できること, そして計算時間と解の品質についてトレードオフがあることを示した.

謝辞

本報告 (または本研究など) には, 国立研究開発法人新エネルギー・産業技術総合開発機構 (NEDO) からの委託を受けて実施した「ロボット・ドローンが活躍する省エネルギー社会の実現プロジェクト」の成果一部が含まれている.

参考文献

- [Amir 2015] Amir, O., Sharon, G. and Stern, R.: Multi-agent pathfinding as a combinatorial auction, in *AAAI* (2015)
- [Machida 2019] Machida, M.: Polynomial-Time Multi-Agent Pathfinding with Heterogeneous and Self-Interested Agents, in *AAMAS*, in press (2019)

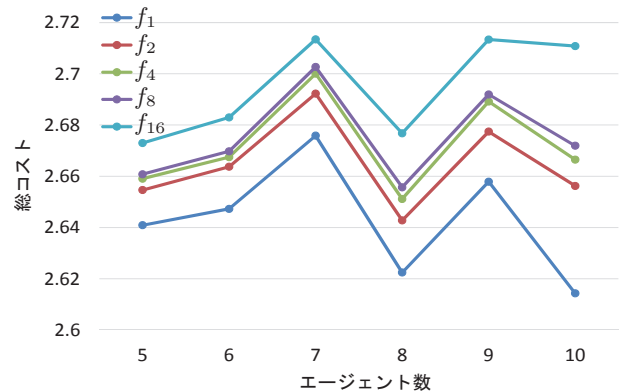


図 4: 8×8 グリッドマップ上, w が $[0, 1]^k$ 上の一様乱数のときの, 1 エージェント当たりの総コスト

- [Bnaya 2013] Bnaya, Z., Stern, R., Felner, A., Zivan, R., and Okamoto, S.: Multi agent path finding for self interested agents, in *Sixth Annual Symposium on Combinatorial Search* (2013).
- [Silver 2005] Silver, D.: Cooperative pathfinding, in *AI-IDE*, pp. 117-122 (2015).
- [Standley 2010] Standley, T.: Finding optimal solutions to cooperative pathfinding problems, in *AAAI*, pp. 173-178 (2010)
- [Ryan 2010] Ryan, M.: Constraint-based multi-robot path planning, in *ICRA*, pp. 922-928 (2010)
- [Dresner 2008] Dresner, K. and Stone, P.: A multiagent approach to autonomous intersection management, *J. Artif. Intell. Res. (JAIR)*, volume 31, issue 1, pp.591-656 (2008)
- [Kiesel 2012] Kiesel, S., Burns, E., Wilt, C. M., and Ruml, W.: Integrating vehicle routing and motion planning, in *ICAPS* (2012)
- [Archer 2001] Archer, A. and Tardos, É.: Truthful Mechanisms for One-Parameter Agents; In *Proceedings of the 42Nd IEEE Symposium on Foundations of Computer Science (FOCS '01)*, pp. 482- (2001)
- [Vickrey 1961] Vickrey, W.: COUNTERSPECULATION, AUCTIONS, AND COMPETITIVE SEALED TENDERS; *Journal of Finance*, Volume: 16, Number: 1, pp. 8-37 (1961)
- [Clarke 1971] Clarke, E.: Multipart pricing of public goods; *Public Choice*, Volume: 11, Number: 1, pp. 17-33 (1971)
- [Groves 1973] Groves, T.: Incentives in Teams; *Econometrica*, Volume: 41, Number: 4, pp. 617-31 (1973)
- [Sharon 2015] Sharon, G., Stern, R., Felner, A., and Sturtevant, N. R.: Conflict-based Search for Optimal Multi-agent Pathfinding; *Artif. Intell.* Volume: 219, Number: C, pp. 40-66 (2015)

制約最適化問題と非集中型解法による移動ロボット群への観測領域割り当ての基礎検討

A Fundamental Study of Region Allocation for Mobile Robots Based on Constraint Optimization and Decentralized Solution Method

松井 俊浩

Toshihiro Matsui

名古屋工業大学

Nagoya Institute of Technology

Observation systems based on autonomous multiple robots including wide area surveillance, emergency response, and temporary exploration in unsafe areas have been widely studied. In general cases, there are various requirements for observation systems such as patrolling, exploration, tracking and cooperative sensing. These tasks also require the formation of robots and the allocation of observation areas. To manage complex sub-tasks simultaneously, a general approach based on constraint optimization problems and decentralized solution methods will be promising. This approach is studied as distributed constraint optimization problems in the multiagent research area. While several studies address the application of the distributed constraint optimization to sensor networks and mobile robots, there are opportunities to construct a unified model to integrate several tasks with the constraint representation and low cost decentralized solution methods. As a fundamental study, we address a simple model for patrolling problems with the allocation of observation areas to multiple robots.

1. はじめに

移動ロボット群による観測タスクの実現は、広域監視、災害対応、危険性がある施設内での臨時的観測の手段として研究されている。自律的なロボット群による非集中型の観測タスクの制御や観測資源割り当てには、インフラ利用が限定される状況での利用や負荷分散の効果が期待される。一般に、観測の目的は巡回、探査、追跡、特定の領域の集中的な観測など多様であり、これらを支える観測領域の分担や、臨時的通信網の維持なども含まれる。このようなタスクを、個々のサブタスクから必要なものを組み合わせ構成する枠組を用いることができれば、一定の有用性があると期待される。その一方で、非集中型の解法によりこれらを連携させることは比較的煩雑であるため、それらの条件を制約や目的関数として表現し、非集中型の解法により解決する枠組の応用が検討されている。このようなアプローチにマルチエージェントシステム上の基本的な最適化問題を扱う分散制約最適化手法の研究がある [Fioretto 18]。

分散制約最適化手法の動機付けとして、分散センサ網などの例題があり [Zhang 05, Béjar 05]、移動ロボット群による観測への応用も模索されている [Jain 09, Stranders 10, Zivan 15]。観測対象であるターゲットが存在する領域への注視などを複数エージェントで分担するタスクなどが検討されているが、移動ロボットの研究で基本的な、巡回などのタスクの上に立脚する総合的な観測の要求に対応する複数の制約からなる問題の表現と、実時間性を考慮する局所探索にもとづく解法の応用のためには、さらに検討の余地があると考えられる。本研究では、このような枠組の基礎部分の検討として、簡単な巡回観測タスクを題材としたモデルと解法の基本的な構成について検討する。

2. 例題

初期の検討のための簡単な例題として、グリッド世界の環境を観測するロボットを模したエージェント群からなる系を考える。環境は障害物と床のグリッドから構成される。各グリッド

連絡先: 松井 俊浩, 名古屋工業大学, 〒 466-8555 愛知県名古屋市昭和区御器所町, matsui.t@nitech.ac.jp

について最後に観測された論理時刻が記録される。エージェントは4方向に移動し、周囲8グリッドの領域を観測し、観測した時刻を更新する。各セルは十分に広く、エージェントは適切な衝突回避により同一セルを共有できるものとする。各エージェントは個別にグリッド世界の地図と、関連する属性情報を持つが、初期の検討として、各エージェントは他のエージェントと情報を随時同期する。環境の観測時刻を更新する基本的な巡回タスクを目的とする。エージェントは地図に基づく距離や経路探索の計算を行い、随時移動目標を更新しつつ移動する。未探査のセルへの移動の際に、互いに視野が重複するセルを解消することを下位の基本的な目的とする。また、各エージェントに自身の観測範囲を持たせ、それらを均衡するように配分することを上位の目的とする。ケーススタディとして、同一のセルから複数のエージェントが出発し、展開するシナリオについて検討する。

3. 制約最適化によるタスクの協調のモデル化

各エージェントの情報を次の要素により表現する。

- エージェントの位置の二次元座標。
- エージェントの次の目的地の二次元座標。
- エージェントが担当する領域のグリッドの集合および、重心付近の到達可能なセルの座標。これは領域を授受する場合などの基準として用いる。
- エージェントの位置を起点とする各セルへの距離と最短経路情報。

各エージェントは次の行動を並行する。

- 次の目的地を選択し移動する。ここでは簡単に、観測を担当する領域において、最も過去に観測し、最も遠いセルを優先して目的地の候補とする。指定された個数の複数の候補の中から移動先を選ぶ。

- 他のエージェントとの観測領域の授受を選択し観測領域を更新する。初期状態において一つのエージェントだけが観測領域を持ち、他のエージェントへ観測領域を段階的に配分する状況を考慮し、次の4種類の授受の行動のいずれかを取る。

- － 担当する観測領域を持たないエージェントが1つ以上ある場合に、自身の領域の半分を他に与える。エージェントの現在の座標から最も離れた観測領域のセルを膨張した領域を渡す。
- － 担当する観測領域を持たないエージェントである場合に、他のエージェントから与えられる領域をすべて受け取る。
- － 担当する観測領域を持つエージェントである場合に、自身の領域の外周付近の領域を他に与える。ただし、障害物に隣接する領域は除外する。
- － 担当する観測領域を持つエージェントである場合に、自身の領域の周辺を領域を他から受け取る。

領域の授受は自身と他のエージェントの領域の大きさの差の最大値が閾値を超えている場合に行なう。領域の授受では、受け取る者が主導して受け取る領域を選択する。授受する領域の大きさは、自身と他のエージェントの領域大きさの差の最大値に応じて決める。一つのエージェントは一度に一つの領域のみを授受する。

これらの行動の条件と評価値を複数の重み付きの制約により表現する。後者のものほど優先する階層的な重みを持つコストを与え、最小化問題とする。

1. 他エージェントと同一セルへの移動を抑制するためのコスト。
2. 他エージェントと視野が競合するセルへの移動を抑止するためのコスト。
3. 領域を受け入れるエージェントが領域を授受した結果としてなお残る領域の不均衡さのコスト。関連するエージェント間の領域の大きさの差の最大値に加重する。

4. 非集中型の局所探索の適用

分散制約最適化問題の基本的な解法である Maximum Gain Messages (MGM) [Fioretto 18] に類似する非集中型の局所探索を適用する。全体の解法の枠組みにおける1シミュレーションステップは次の手順から構成される。

1. 各エージェントは環境を観測し、自身の現在の状況から、問題の構成に必要な情報を計算する。
2. 現在の自身の情報と、自身の次の行動の初期解を提案する。
3. 関連するエージェントの情報を収集する。
4. 自身の行動の解を探索し、評価値を改善する解がある場合に、評価値の改善量を提示する。
5. 他エージェントの改善量を収集する。
6. 自身を含め、改善量が最も高く、識別名による優先順位が最も高いエージェントが、解とそれに伴う情報を更新する。

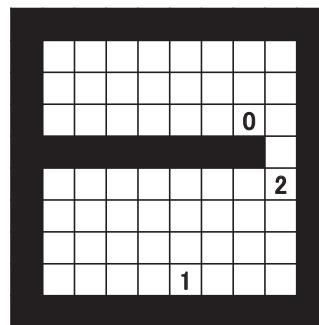


図 1: 環境の例 (番号はエージェントの識別番号)

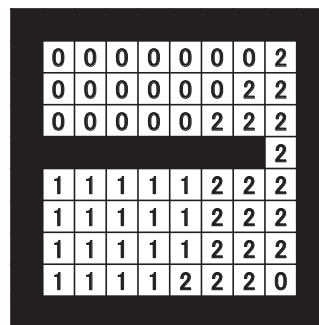


図 2: 観測領域の割り当て

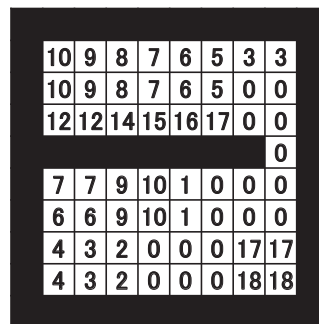


図 3: 観測時刻の占さ

7. すべてのエージェントで改善量がゼロになるまで手順3からの処理を反復する。
8. 各エージェントは合意された行動を実行する。

この解法はきわめて簡素であるが、反復される毎回の交渉にリーダー選出に基づく合意を含むため、任意の段階で交渉を中断することが容易である。実時間性を考慮する場合には、このような局所探索と系の摂動を利用する準最適化のアプローチが妥当と考えられる。

5. 実験

提案手法のシミュレーションを試作し、基礎的な検証を行った。10×10のグリッド世界の右上付近のセルから3つのエージェントが展開し探索する例を示す。各ステップの交渉回数の上限を20回とした。また、観測を担当する領域のサイズの差の最大値を5以下とする目標を設定した。50ステップ経過後

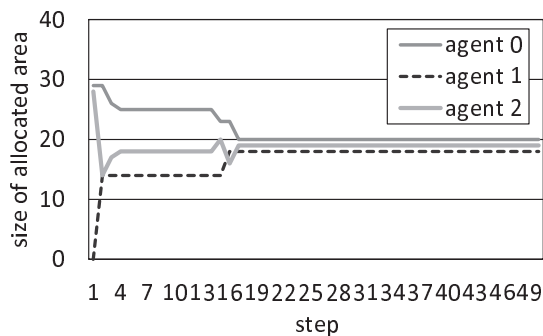


図 4: 観測担当領域の大きさ

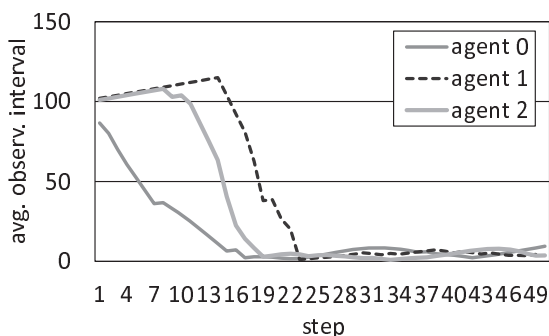


図 5: 平均の観測間隔

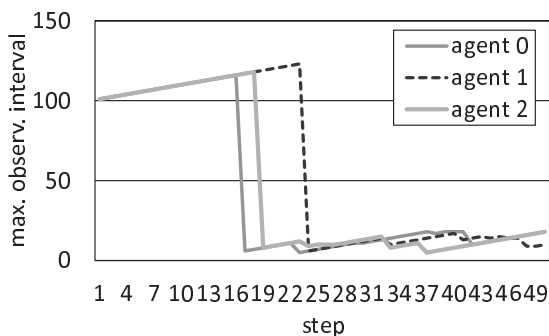


図 6: 最大の観測間隔

の状況の例を図 1-3 に示す。領域の演算が簡易的であるために一部に断片が生じているが、十分に時間が経過した段階では、それぞれの担当領域のサイズは 20, 18, 19 程度であり、局所探索や発見的な領域の演算による誤差はあるものの、ある程度の均衡化がなされている。また、巡回の規則は目標地点の視野の重複の解決以外には、周辺の観測情報が古い領域を移動目標とする貪欲的な戦略を用いたが、ある程度の間隔で更新がなされた。

観測を担当する領域の、大きさ、観測期間の平均と最大値の推移を、図 4-6 に示す。観測領域の配分は、開始時には一つのエージェントが全体を担当する状況であり、開始後直ちに分割された後、領域演算と局所探索の影響により停滞した場面があるが、最終的にほぼ均衡した。観測の間隔は開始時の未探索セルでは 100 とした。全てのセルの観測後はある程度の間隔に収まった。

6. おわりに

本研究では、制約最適化問題による表現と非集中型解法を基礎とする、移動ロボット群による観測システムの制御、資源割り当ての枠組の基礎検討として、巡回観測タスクにおける競合回避と観測担当領域の授受の調整を対象とした問題とモデルを構成し、その挙動について検証した。巡回以外のタスクを含めた総合的な要求に対応する問題の表現と、実時間性を考慮する局所探索にもとづく解法の応用のために、ロボット間の通信の制限を考慮する配置の制約、注視領域への割り当てを優先する観測などの条件の検討が今後の課題として挙げられる。また、より大規模な群における協調の局所性を伴う場合の評価、LiDAR による SLAM を考慮する観測と割り当てのモデルや、動的な環境などの実際的な環境への適用の検討を進めたい。

謝辞 本研究の一部は、公益財団法人立松財団一般研究助成による。

参考文献

- [Béjar 05] Béjar, R., Domshlak, C., Fernández, C., Gomes, C., Krishnamachari, B., Selman, B., and Valls, M.: Sensor Networks and Distributed CSP: Communication, Computation and Complexity, *Artif. Intell.*, Vol. 161, No. 1-2, pp. 117–147 (2005)
- [Fioretto 18] Fioretto, F., Pontelli, E., and Yeoh, W.: Distributed Constraint Optimization Problems and Applications: A Survey, *Journal of Artificial Intelligence Research*, Vol. 61, pp. 623–698 (2018)
- [Jain 09] Jain, M., Taylor, M. E., Tambe, M., and Yokoo, M.: DCOPs Meet the Real World: Exploring Unknown Reward Matrices with Applications to Mobile Sensor Networks, in *IJ-CAI 2009, Proceedings of the 21st International Joint Conference on Artificial Intelligence, Pasadena, California, USA, July 11-17, 2009*, pp. 181–186 (2009)
- [Stranders 10] Stranders, R., Fave, F. M. D., Rogers, A., and Jennings, N. R.: A Decentralised Coordination Algorithm for Mobile Sensors, in *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, AAAI'10*, pp. 874–880, AAAI Press (2010)
- [Zhang 05] Zhang, W., Wang, G., Xing, Z., and Wittenburg, L.: Distributed stochastic search and distributed breakout: properties, comparison and applications to constraint optimization problems in sensor networks, *Artificial Intelligence*, Vol. 161, No. 1-2, pp. 55–87 (2005)
- [Zivan 15] Zivan, R., Yedidsion, H., Okamoto, S., Grinton, R., and Sycara, K.: Distributed Constraint Optimization for Teams of Mobile Sensing Agents, *Autonomous Agents and Multi-Agent Systems*, Vol. 29, No. 3, pp. 495–536 (2015)

実時間連続状態空間マルチエージェント意思決定に対する 局面評価関数の設計について

On the design of state value functions
for real-time continuous-state space multi-agent decision making

中島 智晴^{*1}
Tomoharu Nakashima

五十嵐 治一^{*2}
Harukazu Igarashi

秋山 英久^{*3}
Hidehisa Akiyama

^{*1}大阪府立大学
Osaka Prefecture University

^{*2}芝浦工業大学
Shibaura Institute of Technology

^{*3}福岡大学
Fukuoka University

This paper presents an overview of value function representations and construction for RoboCup soccer simulation. Since RoboCup has several characteristic features such as multi-agent system, noisy environments, and dynamic decision making, it offers a more realistic environment for the decision making in multi-agent system research.

1. はじめに

ゲーム AI の研究は、計算機の性能向上や深層学習の登場により急速な発展を見せている。その結果、囲碁や将棋をはじめとして、人間に勝利する AI プレーヤーの登場が今や当たり前のようになってきた。ゲーム AI を構築するためには、センサ情報処理や意思決定など、社会的課題を解決するロボットや知的情報処理に転用できるサブ課題を解決し、システム統合する必要がある。このため、単純な課題であるトイプロBLEMに比べてゲーム AI をベンチマークとする研究は実世界への応用可能性を広げるという意味で重要である。

RoboCup サッカーシミュレーション 2D リーグは、約 20 年にわたる RoboCup の歴史の中で最も古いリーグの一つである。サッカーを題材とした仮想シミュレーションである一方で、複雑なマルチエージェント問題の研究プラットフォームとしても利用することができる。表 1 に、RoboCup サッカーシミュレーションと囲碁や将棋との違いをまとめる。これほどの制約がプレイヤーに課されているゲームは他にはなく、RoboCup サッカーシミュレーションを対象としたゲーム AI の研究は、より実世界応用につながるものになること考えられる。

表 1: RoboCup と囲碁・将棋との違い

特徴	RoboCup	囲碁・将棋
情報	不完全	完全
思考時間	リアルタイム	ターン制
プレイヤー数	1 チーム 12 人	1 チーム 1 人
状態空間	連続	離散
行動空間	連続	離散

RoboCup サッカーシミュレーションのゲーム AI を構築するためには、センサ情報の処理、ボール処理の行動決定、次サイクルのセンサ情報を得るための行動、行動実施のためのマイクロアクション決定、他エージェントを支援するための情報提供など、幅広い情報処理と意思決定機構が必要になり、数多くの人工知能的課題を解決しなければならない。このような複雑性もあり、RoboCup サッカーの知見を基にした在庫管理ロボットや深層強化学習モデルの発展につながっている。また、将来

連絡先: 中島智晴, 大阪府立大学大学院, 599-8531 大阪府堺市学園町 1-1, 072-254-9351, tomoharu.nakashima@kis.osakafu-u.ac.jp

必要とされる集団ロボットの協調行動や自動運転における AI 部分の発展に大きく寄与すると考えられる。本論文では、その中でも意思決定部分に焦点を当て、RoboCup サッカーシミュレーションにおける意思決定メカニズムの研究を解説する。

2. RoboCup サッカーシミュレーション

RoboCup は、「2050 年までにサッカーでロボットが人間のチャンピオンチームに勝利する」という目標を掲げて立ち上がった国際プロジェクトである。RoboCup サッカーには、いくつかのカテゴリがある。本研究では、実機を使用せず、2 次元空間の仮想空間内で競技を行うサッカーシミュレーション 2D リーグを対象とする。

2.1 構成

RoboCup サッカーシミュレーションは、仮想サッカーフィールド上で競技を行う、実機ロボットを使用しないリーグである。サッカーサーバ、サッカープレイヤー（以降プレイヤー）、コーチ、モニタから構成される。この様子を図 1 に示す。サッカーサーバはサッカーフィールドの全ての情報を保持しており、プレイヤーからのアクションコマンドを受信して次サイクルのフィールド状態を計算し、更新する。更新されたフィールドの情報を各プレイヤーに送信する。プレイヤーは、サーバから受け取られたフィールド情報に基づいて状況を判断し、次サイクルのアクションコマンドをサーバに送信する。プレイヤー間、コーチとプレイヤー間の通信はサッカーサーバを介してのみ許可されており、直接通信することは許されていない。1 ゲームは前後半それぞれ 3000 サイクルであり、1 サイクルの長さは 0.1 秒である（したがって、1 試合の長さは延長戦がなければ 6000 サイクル=5 分）。

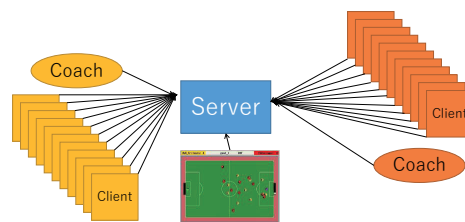


図 1: RoboCup サッカーシミュレーションの構成

2.2 プレイヤの行動決定

ここで、プレイヤの行動とアクションコマンドの違いについて述べる。特に、ボールを持っているプレイヤの意思決定に焦点を当てて説明する。プレイヤはフィールド状態に基づいて次のサイクルの行動を決定するが、行動の意思決定には大きく分けて二つの層がある。まず、意思決定の第1層では、パス、ドリブル、シュート、ホールドなどのマクロ行動を決定する。これらの行動にはパラメータが必要である。

例えば、パスの場合は、パスを受け取るプレイヤ番号と受け取りが、ドリブルの場合には、目標位置がパラメータとなる。ドリブルでは、まずボールを蹴りだした後、次ボールを蹴る位置がパラメータとなり、シュートはゴールのどの位置に向けてボールを蹴るかがパラメータとなる。また、第1層で決定された行動に基づいて、第2層では、その行動とパラメータを達成するためのアクションコマンド（キック、ダッシュ、ターンなど5種類）とそのパラメータを計算する。これは、サーバ内であらかじめルールによって決められているボールやプレイヤの物理計算を逆算することで求められる。キックのパラメータは蹴る方向とその力、ダッシュは走る方向とその力、ターンはターンする量がパラメータとなる。プレイヤはアクションコマンドとそのパラメータをサッカーサーバに送信する。

本論文で取り上げる行動の意思決定は、この第1層における意思決定プロセスのことをいう。

3. 行動連鎖生成システム

行動連鎖システムは Akiyama ら [1] によって最初に導入され、他のチームも採用している。行動連鎖生成システムの目的は、ボールを保持しているプレイヤが次の行動を決定するための方針を決定することである。方針が定まれば、それに沿うような次の行動を決定することができる。

行動連鎖生成システムは、大きく3つに分けることができる。探索木の生成、ノードの評価、木探索アルゴリズムである。このうち、木探索アルゴリズムはゲーム AI で広く使用されているモンテカルロ木探索や系統的探索アルゴリズムが適用できる。RoboCup サッカー（もしくはサッカー AI）固有の要素が必要なのは探索木の生成とノードの評価である。ノードの評価はサッカーフィールドがどれほどスコアにつながっているかを表すもので、局面評価とも呼ばれる。

その評価関数中のパラメータを、人間の専門家同士の棋譜を用いて学習させる教師あり学習により、プロレベルまで棋力を向上させることができる。現在では、階層型のニューラルネットワークモデルを用いて評価関数を近似し、ディープラーニングの手法を取り入れることが主流になっている。特に、評価関数を作成するのが難しいとされた囲碁で大成功しており、チェスや将棋でも試みが始められている。

4. 局面評価モデルの構築

局面評価の基準は、チーム戦略によって変化する。例えば、シュートできることを目的とする戦略を考えている場合には、シュートという目標状態に近いが同課の観点から評価値を設定する。また、シュートまでの中間状態を目標とするチーム戦略を立てることも可能であり、例えば、ボールを敵ペナルティエリアに運ぶことを目的とするチーム戦略を考えている場合には、シュートできるかどうかは考慮せず、とにかくボールを敵ペナルティエリアに運ぶために有利な状態かどうかの観点から評価値を割り当てることになる。

4.1 特徴量の線形和による表現

RoboCup では、各プレイヤは視野範囲内にある他プレイヤの位置や速度、自己位置を知るためのフィールド外に設置されたランドマークの位置が、数値情報として利用可能である。ただし、自分から遠くにあるオブジェクトの情報にはノイズや欠損などの外乱が入り、不正確になる。ノイズや欠損、視野範囲外のプレイヤやボールの情報は他プレイヤからの声かけで補完されることもある。センサ情報にはノイズが大量に含まれており、プレイヤは正しくフィールド状態を把握することは非常に困難である。このような条件下でフィールド状態を評価する。

フィールド状態を評価する最も単純な方法は、フィールド状態を表現している特徴量をいくつか用意し、それぞれを重み付けながら足し合わせることである。この方法は、特徴量の線形和と呼ばれる。特徴量として、センサ情報から得られる敵プレイヤの位置やボール、自プレイヤの位置や速度などを用いたり、「近くに敵がいる」「パス可能な味方プレイヤが近くにいる」などのような、人間がサッカーの知識を導入して作成する特徴量も考えられたりする。後者の特徴量は特に、ヒューリスティクスと呼ばれることもある。

特徴量の線形和でフィールド状態を評価する場合、各特徴量に対応する重みをどのようにして決定するかという問題が残る。単純にはランダムに決定する方法や人間の知識を利用して決定する方法が考えられるが、限界や欠点も存在する。

4.2 ニューラルネットワークモデルによる表現

前節で述べた特徴量の線形和はパーセプトロンと見なすことができる。この見方を発展させて、特徴量を入力、評価値を出力とする階層型ニューラルネットワークを用いてフィールド状態の評価値を求めることもできる。

局面評価値計算モデルとしてパーセプトロンやニューラルネットワークを用いることで、機械学習の枠組みを用いた、経験データを用いたモデルの構築が可能となる。経験データの種類によって、評価値計算モデルを教師あり学習や強化学習により求めることができる。なお、教師無し学習により評価値計算モデルを構築する研究はあまり行われておらず、研究は進んでいない。次章では、教師あり学習と強化学習を用いたフィールド状態計算モデルの構築について述べる。

ニューラルネットワークの特徴量として、数値情報ではなく画像情報を用いる方法も考えられる。Pomas ら [2] は、フィールド状態を数値情報ではなく画像で入手し、その画像からチームにとって有利かどうかを判断する関数を構築している。

5. おわりに

本稿では、RoboCup サッカーシミュレーションにおける局面評価値計算モデルについて述べた。局面評価関数を構築する方法について整理し、今後の研究展望を考慮するうえで重要となる情報を与えることを目的とした。

参考文献

- [1] H.Akiyama, S.Aramaki, T.Nakashima, "Online Cooperative Behavior Planning Using a Tree Search Method in the RoboCup Soccer Simulation," *Proc. of INCoS 2012*, pp.170-177, 2012.
- [2] T.Pomas and T.Nakashima, "Evaluation of Situations in RoboCup 2D Simulations using Soccer Field Images," *Proc. of RoboCup Symposium*, 6 pages, 2018.

[3J4-J-1] Fundamental AI, theory: search and application

Chair:Ichigaku Takigawa Reviewer:Yoichi Sasaki

Thu. Jun 6, 2019 3:50 PM - 5:10 PM Room J (201B Medium meeting room)

[3J4-J-1-01] ACO with Pheromone Update by Negative Feedback Can Solve CSPsOTakuya Masukane¹, Kazunori Mizuno¹ (1. Department of Computer Science, Takushoku University)

3:50 PM - 4:10 PM

[3J4-J-1-02] An Algorithm for solving the Traveling Salesman Problem using Clustering MethodOJumpei Uchida¹, Hajime Anada¹ (1. Tokyo City University)

4:10 PM - 4:30 PM

[3J4-J-1-03] Algorithm of Traveling Salesman Problem using Particle Swarm OptimizationOYuki Yamada¹, Hajime Anada¹ (1. Tokyo City University)

4:30 PM - 4:50 PM

[3J4-J-1-04] League Scheduling for U12 BasketballOTenda Okimoto¹, Kazuki Nishimura¹, Katsutoshi Hirayama¹ (1. Kobe University)

4:50 PM - 5:10 PM

フェロモン更新に負のフィードバックを取り入れたACOによる 制約充足問題の解法

ACO with Pheromone Update by Negative Feedback Can Solve CSPs

増金拓弥 水野一徳
Takuya Masukane Kazunori Mizuno

拓殖大学大学院工学研究科情報・デザイン工学専攻
Department of Computer Science, Takushoku University

To solve large-scale constraint satisfaction problems, ant colony optimization (ACO) has recently been drawing attentions. In algorithms based on ACO, candidate assignments are constructed by taking account of pheromone trails, which are updated based on a candidate assignment with the least number of constraint violations. In this paper, we propose an ACO model with dual pheromone trails: usual pheromone trails and another pheromone trails. Another pheromone trails are updated based on a candidate assignment with the largest number of constraint violations. Also, usual pheromone trails are updated by considering another pheromone trails as well. We demonstrated that our model, which is applied to the cunning ant system, can be effective than other ACO-based methods for large-scale and hard graph coloring problems whose instance appears in the phase transition region.

1. はじめに

制約充足問題 (CSP) は、離散値をとるいくつかの変数について割当て可能な値の組合せのうち、与えられた全ての制約を満たす組合せを、探索によって発見する問題である。CSP は設計や計画問題などをはじめ、人工知能分野やパターン処理などの広い分野にわたって応用されている基盤的技術である。

CSP を解くための探索アルゴリズムは、系統的探索アルゴリズムと確率的探索アルゴリズムに大別される。大規模な CSP に対して、系統的探索アルゴリズムを用いて探索を行なう場合、実用的な時間内に解を得ることが困難である場合が多い。そのため、近年では確率的探索アルゴリズムが注目されている。しかし、確率的探索アルゴリズムには、局所最適解に陥ってしまう可能性があるという欠点がある。そこで、局所最適解に陥ることを避ける、または局所最適解から効率よく抜け出すためのメタヒューリスティクスについて研究がなされている。

蟻コロニー最適化 (ACO) は、蟻の採餌行動におけるフェロモンコミュニケーションをモデル化したメタヒューリスティクスである [Dorigo 96, Dorigo 99]。ACO では、探索の過程で発見された評価が高い解候補の情報を、フェロモンとして蓄積する。このフェロモンを解候補の生成時に参考にし、フェロモンが多く溜まっている値の組合せほど、解候補として選択されやすくなる。

従来の ACO アルゴリズムでは、フェロモン (評価が高い解候補) と制約違反という情報を参考にして探索を行なう。本研究ではこれらの情報に加え、“評価が低い解候補” の情報を探索の参考にする ACO モデルを提案する。特に、評価が低い解候補を、従来のフェロモンとは別のフェロモン (ネガティブフェロモンと呼ぶ) として蓄積させる ACO モデルを提案する。更に本手法では、ネガティブフェロモンが多く溜まっている値の組合せほど、通常のフェロモンの蓄積量を減少させる。これにより解候補が、評価が高い解候補へ誘引されつつ、評価が低い解候補から遠ざけられ、より効率的な探索が行なわれることが期待される。

2. 研究分野の概要

2.1 グラフ彩色問題

グラフ彩色問題 (COL) は制約充足問題に分類される問題の 1 つである。COL は、無向グラフにおいて隣接する頂点同士が同じ色に彩色されないような彩色パターンを探索によって発見する問題である。特に、色数を 3 色とした問題は 3COL と呼ばれ、アルゴリズムを評価するためのベンチマークとしてしばしば用いられる [Mizuno 11, Tayarani-N 15]。ここで、無向グラフの頂点数を n 、辺数を c とした際、制約密度 d を $d = c/n$ と定義する。3COL において、 $d = 2, 3 \sim 2.4$ の領域は相転移領域と呼ばれ、難しい問題が集中していることがわかっている [Hogg 96]。

2.2 蟻コロニー最適化

蟻コロニー最適化 (ACO) は、蟻の採餌行動におけるフェロモンコミュニケーションをモデル化したメタヒューリスティクスである [Dorigo 96, Dorigo 99]。現実世界における蟻は餌を見つけると、フェロモンを落としながら巣まで戻る。蟻にはこのフェロモンに誘引される性質がある。フェロモンの経路は蒸発や他の蟻による補強を繰り返し、徐々に経路の長さは短くなっていく。この過程から着想を得て、最適化問題の解法として提案されたものが ACO である。ACO をもとにしたアルゴリズムは、多くの組合せ最適化問題をはじめ、3COL を含む多くの CSP にも有効な手法であることがわかっている [Bui 08]。

2.3 cunning Ant System

ACO アルゴリズムの 1 つである cunning Ant System (cAS) [Tsutsui 06] のアルゴリズムを図 1 に示す。ACO アルゴリズムでは、1 匹の蟻が 1 つの解候補を持つ。複数の蟻が複数世代にわたって解候補を生成していくことで、探索が進められていく。解候補生成の手順としては、まず値 (COL における色) を割当てる変数 (COL における無向グラフの頂点) をランダムに 1 つ選択する。次に選択された変数に対して、フェロモンと制約違反から決定される確率に基づいて、割当てる値を選択する。以下に、解候補 A の変数 x_j に割当てられる値として v が選択される確率 $p_A(< x_j, v >)$ を示す。

連絡先: 増金拓弥, 拓殖大学大学院工学研究科情報・デザイン工学専攻, 東京都八王子市館町 815-1, 042-665-0519, phsl.masukane@gmail.com

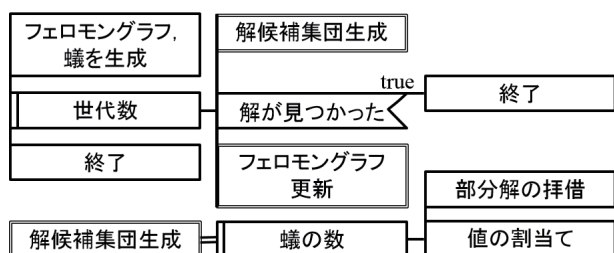


図 1: cAS のアルゴリズム

$$\begin{aligned}
p_A(< x_j, v >) &= \\
& \frac{[\tau_A(< x_j, v >)]^\alpha [\eta_A(< x_j, v >)]^\beta}{\sum_{w \in D_j} [\tau_A(< x_j, w >)]^\alpha [\eta_A(< x_j, w >)]^\beta}, \quad (1) \\
\tau_A(< x_j, v >) &= \sum_{< x_k, u > \in A} \tau(< x_k, u >, < x_j, v >), \\
\eta_A(< x_j, v >) &= \frac{1}{1 + \text{conf}(\{< x_j, v >\} \cup A) - \text{conf}(A)},
\end{aligned}$$

ただし、 D_j は変数 x_j に割当て可能な値の集合 (3COL の場合は *red, green, blue*) を、 $\tau(< x_k, u >, < x_j, v >)$ は “値 u を割り当てた変数 x_k ” と “値 v を割り当てた変数 x_j ” 間に溜まっているフェロモンの量を、 $conf(A)$ は解候補 A の制約違反数をそれぞれ表している。また、 α と β はそれぞれフェロモンと制約違反を重視する重みを表している。これを、全ての変数に値が割当てられるまで繰り返すことで、1つの解候補が生成される。各蟻が解候補を生成した後、その集団内で最も評価が高い解候補を用いて、フェロモングラフが更新される。以下に $\tau(i, j)$ の更新式を示す。

$$\begin{aligned} \tau(i, j) &= (1 - \rho) \times \tau(i, j) + \Delta\tau \\ \Delta\tau &= \begin{cases} \frac{1}{\text{conf}(A_{best})}, & (i, j) \in A_{best} \\ 0, & \text{otherwise} \end{cases} \end{aligned} \quad (2)$$

ただし、 ρ はフェロモンの蒸発率、 A_{best} は集団内で最も評価の高い解候補を表している。探索の過程で制約違反のない解候補が発見されれば探索成功となる。逆に、指定した世代数で制約違反のない解候補が発見できなければ探索失敗となる。

ACO アルゴリズムでは解候補を生成する際、全ての変数に対して値を割当てるために確率を算出する。そのため、探索に時間がかかってしまうという欠点が挙げられる。このような欠点を改良するために提案された ACO アルゴリズムが cAS である。cAS では、解候補を生成する前に、前の世代の蟻 (donor ant) から解候補の一部を“拝借”する蟻 (cunning ant) を用いる。拝借された変数については確率計算が行なわれないため、cAS は探索にかかる時間を短縮できる。

3. 提案手法

3.1 基本方針

従来の ACO アルゴリズムは、制約違反とフェロモングラフ（評価が高い解候補の情報）をもとに探索を行なう。この 2 つの情報に加え、他の情報を ACO に取り入れることで、ACO を用いた探索をより効率的なものにすることができるのではないかと考えた。そこで我々は、探索中に生成される“評価が

低い解候補”に着目した。解候補集団が評価が高い解候補に誘引されると同時に、評価が低い解候補から遠ざかることで、探索がより効率的に行なわれるということが期待される。

本研究では、従来の（評価が高い解候補を用いて更新される）フェロモングラフに加えて、評価が低い解候補を用いて更新されるフェロモングラフ（本研究ではこのフェロモングラフを「ネガティブフェロモングラフ」と呼ぶ）を探索に用いるACOモデルを提案する。本提案モデルの基本方針は以下の2点である。

- 各変数に割当てる値を選択する際には、従来の ACO モデルと同様に、制約違反と通常のフェロモングラフのみを参考にする。
- 通常のフェロモングラフを更新する際には、集団内で最も評価が高い解候補だけでなく、ネガティブフェロモングラフも参考にする。

3.2 アルゴリズム

本研究では、2.3 節で述べた cAS に提案モデルを適用したアルゴリズムを提案アルゴリズムとし、ADUPT (Ant with DUal Pheromone Trails) と呼ぶ。ADUPT の手続きが図 1 に示した cAS の手続きと異なる点は、“ネガティブフェロモングラフの更新”と“通常のフェロモングラフの更新”である。

まず、ネガティブフェロモングラフの更新について説明する。本研究で用いるネガティブフェロモングラフは、通常のフェロモングラフと同じ構造である。また、ネガティブフェロモングラフは集団内で最も評価が低い解候補に基づいて更新される。以下にネガティブフェロモングラフの更新式を示す。

$$N\tau(i, j) = (1 - \rho) \times N\tau(i, j) + \Delta N\tau \quad (3)$$

$$\Delta N_{\tau} = \begin{cases} \text{conf}(A_{\text{worst}}), & (i, j) \in A_{\text{worst}} \\ 0, & \text{otherwise} \end{cases}$$

ただし $N_T(i, j)$ は, (i, j) に溜まっているネガティブフェロモンの量を, A_{worst} は解候補集団内で最も評価の低い解候補を表している。通常のフェロモングラフの更新(式2)と比較すると, 最も評価が低い解候補を用いることの他に, “フェロモンの蓄積量” が異なっている。従来のフェロモンの蓄積量が解候補の制約違反数の逆数であったのに対して, ネガティブフェロモンの蓄積量は解候補の制約違反数そのものとする。これは, 評価が低い解候補情報を蓄積するためのネガティブフェロモンにおいて, 制約違反数が多い解候補ほど参考とする価値が高いためである。

次に、通常のフェロモングラフの更新について説明する。従来のACOアルゴリズムでは、最も評価が高い解候補に含まれる全ての値の組合せについて、フェロモンの蓄積量は一定であった(式2)。本研究では、ネガティブフェロモンが多く溜まっている、つまり、低品質な割当てについて、通常のフェロモンの蓄積量を小さくする。これにより、低品質な割当てが確率的に生成されづらくなると考えられる。以下に通常のフェロモンの更新式を示す。

$$\tau(i, j) = (1 - \rho) \times \tau(i, j) + \Delta\tau \quad (4)$$

$$\Delta\tau = \begin{cases} \frac{1}{\text{conf}(A_{best})} \times w(i, j), & (i, j) \in A_{best} \\ 0, & \text{otherwise} \end{cases}$$

ここで、 $w(i, j)$ はネガティブフェロモン量 $N\tau(i, j)$ によって決定される重み係数である。ただし $w(i, j)$ の値は、対応するネガティブフェロモンの量が上限値に近いほど 0 に近い値をとり、下限値に近ほど 1 に近い値をとるものとする。

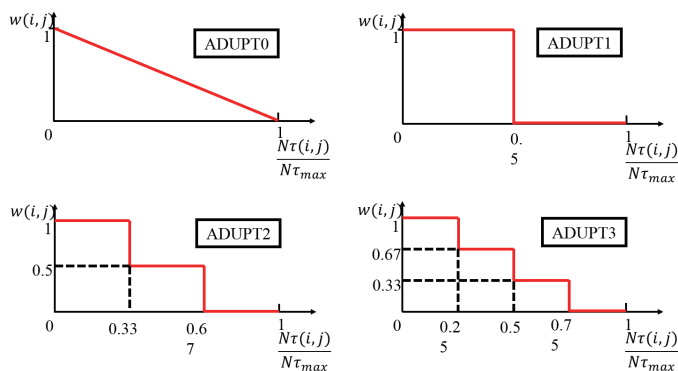
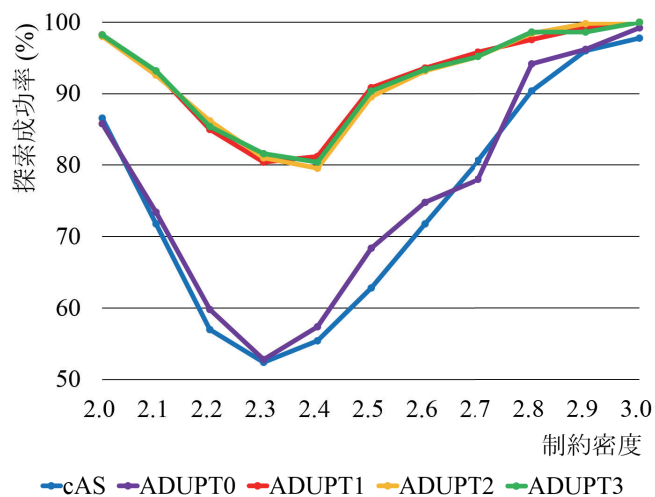
図 2: 重み $w(i, j)$ を決定するための関数

図 3: 各制約密度における探索成功率

4. 評価実験

4.1 実験条件

提案アルゴリズムである ADUPT の有効性を示すために、評価実験を行なった。ここでは、ADUPT と従来手法である cAS を比較する。本実験では頂点数 $n = 100$ である 3COL インスタンスを、制約密度 $d = 2.0 \sim 3.0$ の範囲で 0.1 毎に 11 のケースに対して、それぞれランダムに 100 問生成した。ここで生成した問題は全て解が存在する。生成した各問題について、それぞれ 5 回ずつ試行を行なった。本実験のパラメータは、蟻の数を 100、蟻の最大世代数を 2000、フェロモンの重み $\alpha = 5.0$ 、制約違反の重み $\beta = 10.0$ 、フェロモンの蒸発率 $\rho = 0.01$ とした。また、 $w(i, j)$ を決定するために、図 2 に示す 4 パターンの関数を用いる。それぞれの関数を用いた ADUPT をそれぞれ、ADUPT0 ~ ADUPT3 とする。これらの手法について、以下の 2 点を比較した。

- 探索成功率: 全体の試行数に占める制約充足解を発見できた試行数
- 探索コスト: 制約充足解を発見するまでに生成した解候補の数

なお、実験には PC/AT 互換機 (CPU: Intel Core i7 880 3.07GHz, RAM: 4GByte) を使用し、プログラムはすべて Java 言語で記述した。

4.2 実験結果

実験の結果を図 3 と図 4 に示す。図 3 は各手法の探索成功率をインスタンスの制約密度ごとに示したものである。図 3 から、提案手法である ADUPT0 の探索成功率は、従来手法である cAS の探索成功率と比較すると、差が小さい。一方で、ADUPT1, ADUPT2, ADUPT3 の探索成功率は、cAS よりも高くなっている。また、その探索成功率の差は、制約密度が相転移領域 ($d = 2.3 \sim 2.4$) に近づくにつれて大きくなっている。特に、 $d = 2.3$ のインスタンスについては、探索成功率の差が約 30% であり、最大となっている。

次に、図 4 は全ての試行における探索コストの平均を示している。ただし、探索に失敗した (制約充足解を発見できなかった) 試行については、蟻の数 \times 世代数 $= 200,000$ の解候補を生成しているため、探索コストは 200,000 とした。図 4 よ

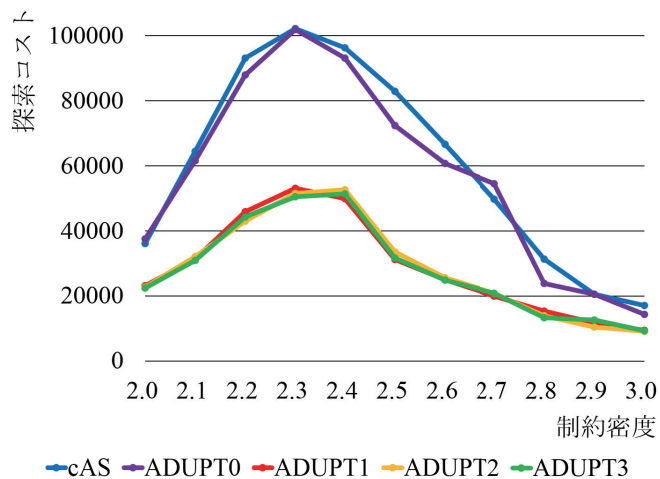


図 4: 各制約密度における探索コスト

り、ADUPT0 の探索コストは cAS とほとんど差がない。また、ADUPT1, ADUPT2, ADUPT3 は、cAS よりも小さい探索コストで解を発見できている。この探索コストの差は、相転移領域に近づくにつれて大きくなっている。特に、 $d = 2.3$ のインスタンスについては、ADUPT1, ADUPT2, ADUPT3 は cAS の約半分の探索コストで解を発見できている。ここで、今回の結果では、探索に失敗した試行における探索コストを 200,000 としたが、実際は 200,000 のコストをかけても解を発見できていない。つまり、解が見つかるまで探索を続けると、探索のコストの差は更に広がっていくと思われる。

4.3 考察

4.2 節より、ADUPT1, ADUPT2, ADUPT3 の 3 つの手法は、cAS よりも高い探索成功率を得られた。一方で、ADUPT0 は探索成功率と探索コストのどちらについても、cAS とほとんど差がなかった。本節ではこの実験結果について考察する。提案モデルでは、従来の ACO アルゴリズムにおけるフェロモンの蓄積量に変化を与えるものである。そのため、ここでは“フェロモンの溜まり方”に着目する。

図 5 は、探索に成功した 1 つの試行について調査したものである。縦軸はその試行で発見できた解の割当てに対応する

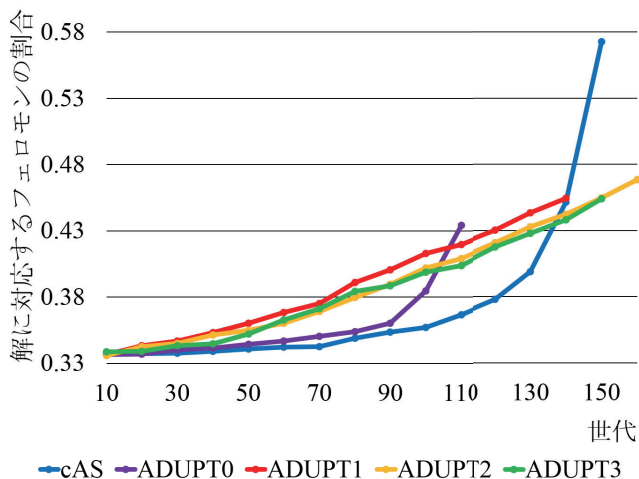


図 5: 解に対応するフェロモン量の割合の推移

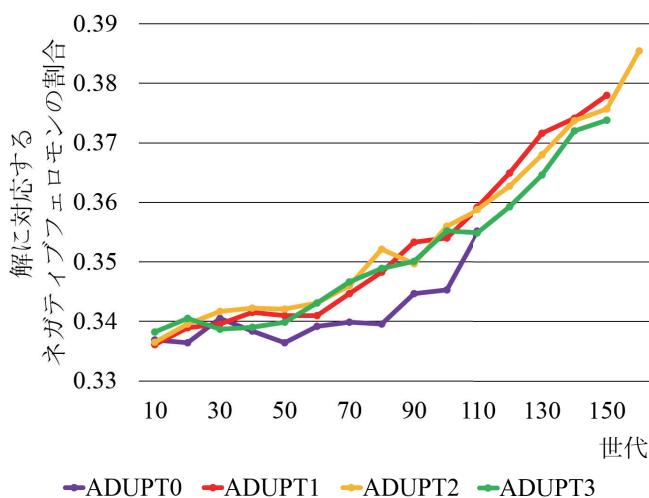


図 6: 解に対応するネガティブフェロモン量の割合の推移

フェロモンの量が、全ての割当てに対応するフェロモンの量に占める割合を、横軸は世代数を表している。図 5 から、cAS と ADUPT0 はおよそ 90 世代まで、解に対応するフェロモンの割合にほとんど変化が見られない。このことから、cAS と ADUPT0 についての実験結果がほとんど同じものとなったのは、cAS と ADUPT0 のフェロモンの溜まり方が似ているためではないかと考えられる。一方で、ADUPT1, ADUPT2, ADUPT3 は、解に対応するフェロモンの割合が徐々に増加していることがわかる。つまり、ADUPT1, ADUPT2, ADUPT3 は、後に発見される解に近い解候補が生成されやすくなっていく。このことが、探索成功率の大きな向上につながっているのではないかと考えられる。

図 6 は、図 5 と同じ試行について調査したものである。ただし、縦軸はその試行で発見できた解の割当てに対応する“ネガティブフェロモン”の量が、全ての割当てに対応するネガティブフェロモンの量に占める割合である。前述した通り、ネガティブフェロモンは集団内で最も評価が低い解候補に基づいて更新される。そのため、解に対応するネガティブフェロモン量の割合が上昇しているということは、最も評価が低い解候補が、後に発見される解に対応する割当てを多く含んでいるとい

うことであると思われる。つまり、集団内の解候補全体が、解に近づいていると考えられる。解に対応するネガティブフェロモン量の割合について、ADUPT0 がおよそ 90 世代までほとんど変化がないのに対して、ADUPT1, ADUPT2, ADUPT3 は徐々に上昇している。この結果から、ADUPT1, ADUPT2, ADUPT3 は集団全体が解に近づいていると思われる。結果として、ADUPT1, ADUPT2, ADUPT3 は解をより発見しやすくなっているのではないかと考えられる。

5. おわりに

本研究では、大規模な CSP を効率よく解くための手法として、複数種類のフェロモンを用いる ACO モデルを提案した。本モデルは、従来の ACO アルゴリズムで用いられるフェロモンを更新する際に、解候補集団内で最も評価が低い解候補に基づいて更新される新たなフェロモンを用いるものである。提案モデルを cAS という ACO アルゴリズムに適用し、単一のフェロモンのみを用いる cAS と実験的に比較を行なった。実験の結果から、提案モデルの有効性を示した。今後の課題として、他のメタヒューリスティクスとの比較や、他の CSP への適用が挙げられる。

参考文献

- [Bui 08] Bui, T. N., Nguyen, T. H., Petal, C. M. and Phan, K. T.: An ant-based algorithm for coloring graphs, *Discrete Applied Mathematics*, Vol. 156, pp. 190–200 (2008).
- [Dorigo 96] Dorigo, M., et. al.: The Ant System: Optimization by a Colony of Cooperating Agents, *IEEE Transaction on Systems, Man, and Cybernetics - Part B*, Vol. 26, pp. 26–41 (1996).
- [Dorigo 99] Dorigo, M. and Di Caro, G.: The Ant Colony Optimization Meta-Heuristics, *New Ideas in Optimization*, pp. 11–32 (1999).
- [Hogg 96] Hogg, T., Huberman, B. A. and Williams, C. P.: Phase transition and search problem, *Artificial Intelligence*, Vol. 81, pp. 1–16 (1996).
- [Mizuno 11] Mizuno, K., Hayakawa, D., Sasaki, H., and Nishihara, S.: Solving Constraint Satisfaction Problems by ACO with Cunning Ants, *The 2011 Conf. on Technologies and Applications of Artificial Intelligence (TAAI2011)* (2011).
- [Tayarani-N 15] Tayarani-N, M. H. and Prugel-Bennett, A.: Anatomy of the fitness landscape for dense graph-colouring problem, *Swarm and Evolutionary Computation*, Vol. 22, pp. 47–65 (2015).
- [Tsutsui 06] Tsutsui, S.: cAS: Ant Colony Optimization with Cunning Ants, *Proc. of the 9th Int. Conf. on Parallel Problem Solving from Nature (PPSN IX)*, pp. 162–171 (2006).

クラスタリング手法を用いた TSP の解法

An Algorithm for solving the Traveling Salesman Problem using Clustering Method

内田 純平^{*1}
Jumpei Uchida

穴田 一^{*1}
Hajime Anada

^{*1} 東京都市大学
Tokyo City University#1

Many economic and industrial problems lead to combinatorial optimization problems. Of these combinatorial optimization problems, the traveling salesman problem (TSP) is one of the most important problem in the field of technology and science. Therefore, we construct a new algorithm for the TSP using a new clustering method. We confirmed the effectiveness of our algorithm using several benchmark problems taken from the TSPLIB, which is a library of traveling salesman problem.

1. はじめに

工業や経済の問題の多くは、最も効率が良い組み合わせを求める組み合わせ最適化問題に帰着することができる。その中に、与えられた全ての都市を巡る最短経路を求める巡回セールスマン問題(Traveling Salesman Problem, TSP)がある。

本研究では、TSP に適した新しいクラスタリングのアルゴリズムである Clustering in clusters(以下 CIC)を提案し、CIC によって構成したクラスタに対して NN 法と 2-opt 法、Or-opt 法を用いることにより、経路生成を行うアルゴリズムを構築した。そして、TSPLIB に掲載されているベンチマーク問題を用いて提案手法の評価実験を行い、その有効性を確認した。

2. 既存研究

2.1 Or-opt 法

Or-opt 法は巡回経路上の連続した n 個のノードを切り出し、それを他の位置に挿入する改善手法である。

Or-opt 法のアルゴリズムは以下の通りである

- ① 都市の選択
経路上の他の位置に挿入する連続した n 個のノードを選択する。
- ② 都市の挿入
①で選択されたノードを他のノード間に挿入し、つなぎ替えた時の総経路長が、元の総経路長より短くなる時にそのノードの挿入を行う。

①, ②を改善が見られなくなるまで繰り返す。

2.2 二次元凸包

二次元の点の集合があたえられたとき、その各点にまっすぐ釘を打ち付け、周りに輪ゴムをかけ手を放すと、ゴムは多角形になる。このゴムに囲われた領域がこの点集合の二次元凸包である。また、集合に属するすべての点は凸包の内側に存在する。

二次元凸包を求めるアルゴリズムは以下の通りである。

- ① 初期点の決定

二次元座標上の x 軸と、原点から各点へのベクトルのなす角が最も小さい点を始点とし、原点から始点へのベクトルを基準ベクトルとする。

- ② 次点の決定

始点から他の点へのベクトルと基準ベクトルのなす角を求め、なす角が最も小さい点を次の始点とし、1つ前の始点からこの始点へのベクトルを基準ベクトルとする。

②における次点が①で決定した初期点になるまで繰り返す。

3. 提案手法

3.1 Clustering in Clusters (CIC)

提案手法である CIC とは、クラスタリングによって生成されたクラスタの中で、クラスタリングをするアルゴリズムである。CIC のアルゴリズムは以下の通りである。

1. クラスタリング i

- ① 初期設定

与えられた問題を第一象限に移動し、二次元凸包を用いて、与えられた問題の凸包を作る。原点に最も近い凸包に含まれるノードを第 1 初期点(以下点 a)とし、 a から最も遠い凸包に含まれるノードを第 2 初期点(以下点 b)とする。また、与えられた問題の y 座標の最大点(以下 $\text{Max}(y)$)を求め、座標 $(0, \text{Max}(y))$ に最も近い凸包に含まれるノードを第 3 初期点(以下点 c)とし、 c から最も遠い凸包に含まれるノードを第 4 初期点(以下点 d)とする。

なお、 a, b, c, d 四ヶ所の始点から独立に次の②, ③, ④を行うものとする。

- ② 初期距離の決定

始点から最も近いノードを選び次の始点とする。また、ノード間の距離を基準距離 D とする。なお、距離の計算にはユークリッド距離を用いる。

- ③ クラスタリング

始点から基準距離 D 以下の距離を持つノード全てと始点でクラスタを生成する。なお、 a, b, c, d での 1 個目のクラスタ生成が終了後、それぞれ基準距離 D の範囲内にノードが存在しないときは全クラスタの重

心を求め、最も近い重心のクラスタに入れるという手続きを加える。

④ 始点と基準距離の更新

クラスタリングされていないノードの中で最も始点に近いノードを次の始点とし、それらのノード間の距離 L とする。そして、基準距離の更新を次式で定義する。

$$D = \frac{D+L}{2} \quad (1)$$

全ノードについてクラスタリングが終わるまで③と④を繰り返す。

II. クラスタリング ii

① 初期設定

各クラスタ内の最大ノード数 MN を次式で定義する。

$$MN = [AN \times 0.05] + 1 \quad (2)$$

ここで、 AN は全ノード数を表す。また、クラスタリング i で生成されたクラスタをクラスタ内のノード数の降順に並び替え、この順番で②から行う。

② 始点の決定

クラスタの重心に最も近いノードを始点とする。また、重心から始点までの距離を $d_{1,2}$ とする。

③ 初期基準距離の決定

始点から最も近いノードを次の始点とし、その間の距離を $d_{2,3}$ として $d_{1,2}$ との平均を基準距離 SD_1 とする。

④ 初期クラスタリング

始点から基準距離 SD_1 以下の距離を持つノードを選択し、選択されたノードと始点でクラスタを生成する。基準距離 SD_1 以下の距離を持つノードが存在しないときは、始点だけでクラスタを生成し、その始点から最も近いクラスタリングされていないノードを次の始点とする。そして新しい始点での距離を $d_{3,4}$ として $d_{1,2}$, $d_{2,3}$ との平均を基準距離 SD_2 とする。

⑤ クラスタリング

始点から基準距離 $SD_i (i \geq 2)$ 以下の距離を持つノードを選択し、選択されたノードと始点でクラスタを生成する。基準距離 SD_i 以下の距離を持つノードが存在しないとき、始点だけでクラスタを生成し、その始点から最も近いクラスタリングされていないノードを次の始点とする。この時、元の始点を i 、次の始点を $i+1$ としたノード間の距離を $d_{i,i+1}$ とする。

⑥ 基準距離 SD の更新

初期クラスタリング後の基準距離には、次式で定義される更新式で随時更新された基準距離 SD_i を使用する。

$$SD_i = \frac{\sum_{k=1}^{i+2} d_{k,k+1}}{i+1} \quad (3)$$

全ノードに対して ⑤, ⑥ を繰り返す。

III. クラスタリング iii

① 初期設定

クラスタリング ii で生成されたクラスタをクラスタ内のノード数が 2 以上のクラスタのリスト(以下 $clst2$)とノード数が 1 つのクラスタのリスト(以下 $clst1$)に分ける。

② 並び替え

$clst2$ 内のクラスタをノード数の昇順に並び替える。

③ 基準距離の設定

$clst2$ 内のクラスタを昇順に 1 つ選び、選ばれたクラスタ内でのノード間の最短距離を求める。求められたノード間の最短距離をそのクラスタの基準距離とする。これを $clst2$ 内のクラスタ全てに行う。

④ クラスタリング

$clst2$ 内のクラスタを昇順に 1 つ選び、選ばれたクラスタ内のノードを無作為に 1 つ選ぶ、そのノードと他のクラスタのノードとの距離が基準距離以下のとき、そのノードを選ばれたクラスタに入れる。クラスタのノード数が最大ノード数に達するか、クラスタ内の全てのノードから探索を終えたら、まだ選ばれていない次のノードについて同様の手続きを行う。

$clst2$ 内の全てのクラスタに対して③, ④を繰り返す。

IV. クラスタリング iv

① 初期設定

クラスタリング iii によってクラスタリングされた $clst1$ と $clst2$ を用いる。

② 基準距離の設定

$clst2$ 内の各クラスタの重心からクラスタ内の各ノードまでの距離を求め、平均値の 1.5 倍を基準距離とする。

③ クラスタの分解

$clst2$ 内のクラスタで、クラスタの重心とクラスタ内の各ノードとの距離の平均が②で決められた基準距離以上であるクラスタについて、そのクラスタ内で、クラスタの重心に一番近いクラスタ内のノードから一番離れているクラスタ内のノードをそのクラスタから外し、 $clst1$ に入れる。

④ クラスタリング

②を行い基準距離の再設定をする。 $clst1$ のノードにおいて、ノードを二つ選びノード間の距離が基準距離以下であれば 1 つのクラスタとして $clst2$ に入れる。これを $clst1$ 内の全ノードに行う。

V. クラスタリング v

① 初期設定

クラスタリング iv によってクラスタリングされた $clst1$ と $clst2$ を用いる。

② 基準距離の設定

$clst2$ 内のクラスタを 1 つ選ぶ。また、選ばれたクラスタ内において、そのクラスタの重心に最も近いそのクラスタのノードを選び、選ばれたノードと最も離れているクラスタ内のノードとの距離を選ばれたクラスタの基準距離とする。また、 $clst2$ 内のクラスタ全てに行う。

③ 候補の作成

$clst2$ 内のクラスタを 1 つ選び、 $c2$ とする。また、 $clst1$ のノードを順に 1 つ選び、 $c1$ とする。 $c2$ 内の重心を始点とし、重心と $c1$ の距離が $c2$ の基準距離の 1.5 倍以下であるならば $c1$ をリスト $stock$ (以下 $stock$)に保存す

る. 1 つの c_2 が選ぶ c_1 の数の上限 SN を次式で定義する.

$$SN = MN + 1 - NN \quad (4)$$

ここで, MN は最大ノード数(式(2)), NN は c_2 内のノード数を表す. $clst1$ 内の全ノードとの比較が終了するか, c_2 の選んだ c_1 の数が SN に達したとき, $clst2$ 内のまだ選ばれていない次のクラスタを c_2 とする. これを $clst2$ 内で選ばれていないクラスタがなくなるまで行う.

④ クラスタリング

$stock$ 内のノードを 1 つ選ぶ. 選ばれたクラスタの重心と最も近い重心を持つ $clst2$ 内のクラスタに選ばれたノードを入れる. なお, この時のクラスタの最大ノード数は(2)式の最大ノード数に 1 を加えたものであり, これを満たさない場合は, そのクラスタを除いた, 他の $clst2$ 内のクラスタで探す. このクラスタリングを $stock$ 内のクラスタ全てに行う.

④が終了したとき, ②を行い各クラスタの基準距離を測定し, 各クラスタの基準距離の平均に 1.5 倍したものを次に使う基準距離とする. そして, $clst1$ のノードを 2 つ選び, それらのノード間の距離が基準距離以下であれば 1 つのクラスタとして $clst2$ に入れる. その後, ②, ③, ④を行い終了.

3.2 解の構築方法

解構築に使用した手法は, 与えられた問題を CIC によってクラスタリングした後の各クラスタの重心で疑似的な最短経路を生成すると, 与えられた問題で生成した経路の厳密解と疑似経路が似通う事を利用したアルゴリズムである.

本手法のアルゴリズムは以下の通りである.

① クラスタリング

$CIC(3.1)$ を用いてクラスタリングを行う.

② 疑似経路作成

生成されたクラスタの重心をノードとみて, 重心を結ぶ経路を NN 法と, $2opt$ 法, $n=1$ の $Or-opt$ 法(2.1)の順で用いて最短経路になるように作成する.

③ 経路生成準備

クラスタを無作為に 1 つ選ぶ, 選ばれたクラスタの重心から最も近いノードを, 疑似経路上で次の経路の重心を持つクラスタ内のノードから選ぶ. これを②で生成された疑似経路をもとに一周するまで繰り返す.

④ 経路生成

③においてランダムに選ばれたクラスタを r_n とすると, ②の疑似経路における r_{n-1} の時に③で選ばれた r_n のノードを始点, r_n の時に選んだ r_{n+1} のノードを終点とする. そして, 始点のノードから r_n の始点以外のノードを通り終点までの経路を NN 法, $2opt$ 法, $n=1$ の $Or-opt$ 法(2.1)の順で用いて最短経路になるように作成し, 随時繋げていく.

生成された経路に対して $2opt$ 法, $n=1$ の $Or-opt$ 法(2.1)を最後に行い終了する.

4. 結果

提案手法の有効性を確認するため, $TSPLIB$ に掲載されている TSP のベンチマーク問題である $eil51$, $kroa100$, $kroc100$, $kroa150$ を用いて評価実験を行った.

表 1: $eil51$ の 100 試行の結果

$eil51$ (最適解 426)	提案
平均誤差率(%)	2.6126
誤差率の標準偏差	0.6207
平均時間(秒)	1.28

表 2: $kroa100$ の 100 試行の結果

$kroa100$ (最適解 21282)	提案
平均誤差率(%)	1.8286
誤差率の標準偏差	0.7378
平均時間(秒)	5.36

表 3: $kroc100$ の 100 試行の結果

$kroc100$ (最適解 20749)	提案
平均誤差率(%)	1.4653
誤差率の標準偏差	0.6018
平均時間(秒)	4.10

表 4: $kroa150$ の 100 試行の結果

$kroa150$ (最適解 26524)	提案
平均誤差率(%)	2.0923
誤差率の標準偏差	0.4178
平均時間(秒)	19.06

実験の結果, まだまだ十分な性能とは言えない.

5. 今後の課題

今後の課題として, クラスタリングの精度を上げることが挙げられる. 提案アルゴリズムの性質上, 各クラスタの重心の位置への依存性が高く, 重心による疑似経路が厳密解と大きく離れていると, 厳密解と大きく離れてしまう欠点がある. また, (2)式より明らかだが, クラスタ内の最大ノード数は全ノード数に大きく依存しているので, 都市数の大きな問題をクラスタリングするとクラスタ 1 つ当たりのノード数が大きくなってしまい, 厳密解に似通った疑似経路の作成が難しくなるという欠点もある. そこで, クラスタ内のノード数を小さくしたクラスタリングを行い, 出来たクラスタ以下を新しいノードとみてクラスタリングを行うことを一定のクラスタ数になるまで再帰的にを行い, 解の構築(3.2)を生成されたクラスタの重心に行うことを, クラスタリングした回数階層的に繰り返して元の都市数に戻すことで, 現在の提案手法よりも最適解に到達しやすく, より大きな都市数の問題を解けるアルゴリズムにしたいと考えている.

粒子群最適化を用いた巡回セールスマン問題の解法

An Algorithm for Solving the Traveling Salesman Problem using Particle Swarm Optimization

山田 悠希^{*1}
Yuki Yamada

穴田 一^{*1}
Hajime Anada

^{*1} 東京都市大学 大学院総合理工学研究科
Graduate School of Integrative Science and Engineering, Tokyo City University #1

Many economic and industrial problems lead to combinatorial optimization problems. Of these combinatorial optimization problems, the traveling salesman problem (TSP) is one of the most important problem in the field of technology and science. And particle swarm optimization (PSO) is a population based stochastic optimization technique inspired by social behavior of bird flocking or fish schooling. PSO has been applied to various combinatorial optimization problems belonging to nondeterministic polynomial-time hard (NP-hard) combinatorial problems. But applying PSO to TSP is difficult. Therefore, we construct a new algorithm which is based on PSO. We confirmed the effectiveness of our algorithm using several benchmark problems taken from the TSPLIB, which is a library of traveling salesman problem.

1. はじめに

工業や経済の問題の多くは、最も効率が良い組み合わせを求める、組み合わせ最適化問題に帰着することができる。その中に、与えられた全ての都市を巡る最短経路を求める巡回セールスマン問題 (Traveling Salesman Problem, TSP) がある。本庄らは、最適化問題に用いられるアルゴリズムの一つである粒子群最適化 (Particle Swarm Optimization, PSO) [Kennedy 95] を TSP 向けに改良した挿入操作 PSO 戦略 (Insertion-based PSO strategy, IPSO) [本庄 16] を提案した。IPSO は、解空間上に配置された各粒子がそれまでの最良解と、近傍の粒子の最良解の情報を基に、解の更新を繰り返すことで解空間の探索を行うアルゴリズムである。しかし、この IPSO には探索が十分に行われないうちに、局所解に陥ってしまうという問題点がある。

そこで本研究では、既存手法で用いられた各粒子のそれまでの最良解と近傍の粒子の最良解の情報に加え、解空間上で最も遠い粒子の解の情報を現在の解に重ね合わせた解の集合を用いて、解の更新を行うアルゴリズムを構築した。そして、TSPLIB に掲載されているベンチマーク問題を用いて既存手法と提案手法を比較することで、その有効性を確認した。

2. 既存研究

2.1 粒子群最適化

粒子群最適化 (Particle Swarm Optimization, PSO) とは、魚や鳥などに見られる群れ行動を探索手法に応用した、最適化手法の一つである。解空間上に位置と速度を持った複数の個体 (以下、粒子と表記) をランダムに配置する。各粒子の位置は問題の解を表現しており、評価の高い粒子の情報を近傍の粒子から入手し、その情報を基により良い位置に近づくように速度と位置を更新する。PSO はこの操作を繰り返すことで解空間を探索するアルゴリズムである。t イテレーション目における粒子 i の位置 $x_i(t)$ と速度 $v_i(t)$ の更新式は次式で定義される。

$$x_i(t) = x_i(t-1) + v_i(t-1) \quad (1)$$

$$v_i(t) = wv_i(t-1) + c_1r_1(pbest_i - x_i(t)) + c_2r_2(lbest_i - x_i(t)) \quad (2)$$

ここで、w はパラメータ、 c_1 と c_2 は [0,1] のパラメータ、 r_1 と r_2 は [0,1] の一様乱数、 $pbest_i$ は粒子 i のそれまでの最良解、 $lbest_i$ は粒子 i の近傍内のそれまでの最良解である。アルゴリズムの流れの詳細は以下の通りである。

① 初期設定

全粒子の位置と速度をランダムに設定し、各粒子 i の $pbest_i$ を現在位置に設定する。次に、設定した近傍数 k を元に、各粒子 i と距離が近い k 個の粒子を粒子 i の近傍に設定する。そして、各粒子 i の近傍内で適応度が最も高い解を近傍内の最良解 $lbest_i$ と設定し、全粒子の中で適応度が最も高い解を $gbest$ と設定する。

② 位置の更新

(1), (2) 式に従い、各粒子の位置の更新を行う。

③ 適応度の評価

全粒子の適応度の評価を行う。適応度は問題に適した粒子ほど高くなるよう、評価関数を事前に設定しておく。

④ $pbest_i$ と $gbest$ の更新

全粒子の $pbest$, $lbest$ と $gbest$ を更新する。

⑤ 速度の更新

(1), (2) 式に従い、各粒子の速度の更新を行う。

初期設定を①で行い、②から⑤までの操作を 1 イテレーションとし、事前に設定したイテレーション数を満たすまで繰り返すことで解空間を探索する。

2.2 挿入操作 PSO 戦略

本庄らが提案した IPSO は, PSO に基づき TSP の解空間の探索を行うアルゴリズムである。まず, 解空間上に複数の粒子を配置する。これらの粒子は, それぞれ巡回路である解を持っており, 各粒子のそれまでの最良解と近傍の粒子の最良解から抽出した部分経路を, 各粒子の現在の解に挿入することで解の更新を行い, これを繰り返すことで, 解空間を探索する。アルゴリズムの流れは以下の通りである。

①初期設定

各粒子 i に解 x_i をランダムに設定し, 各粒子のそれまでの最良解 $pbest_i$ を現在の解 x_i に設定する。粒子 i と粒子 j 間の距離 d_{ij} を以下のように定義し, 全粒子間の距離を計算する。

$$d_{ij} = \frac{1}{S_{ij}} \quad (3)$$

$$S_{ij} = \frac{|E_i \cap E_j|}{n}$$

ここで, E_i は粒子 i が持つ解 x_i の経路の集合, $|E_i \cap E_j|$ は E_i と E_j の共通している経路の本数, n は都市数を表している。距離 d_{ij} は x_i と x_j の共通の経路が多くなるほど短くなる。次に, 設定した近傍数 k を元に, 粒子 i と距離に近い k 個の粒子を粒子 i の近傍に設定する。各粒子 i の近傍の中で総経路長が最も短い解を近傍内の最良解 $lbest_i$ と設定し, 全粒子の中で最も総経路長が短い解を全粒子の最良解 $gbest$ と設定する。

②解の更新

解 x_i は $pbest_i$ の部分経路である $pbest_i'$ と $lbest_i$ の部分経路である $lbest_i'$ を総経路長が最も短くなるように挿入することで更新される。粒子 i の解の更新の詳細は以下の通りである。また, 9 都市の TSP の解の更新の例を図 1 に示す。図 1 の例の $x_i = (1, 4, 7, 5, 6, 9, 8, 3, 2)$ は都市 1 → 都市 4 → … → 都市 3 → 都市 2 と都市を巡り, 都市 1 に戻る巡回路を表している。

I 部分経路の作成

粒子 i の $pbest_i$ から, p 本の連続する経路をランダムに抜き出し, 部分経路 $pbest_i'$ とする。また, 粒子 i の $lbest_i$ から, l 本の連続する経路をランダムに抜き出し, 部分経路 $lbest_i'$ とする。 p と l は以下の式で表される。

$$p = [c_1 r_1 (n + 1)] \quad (4)$$

$$l = [c_2 r_2 (n + 1)] \quad (5)$$

ここで c_1 と c_2 は $[0, 1]$ を満たすパラメータ, r_1 と r_2 は $[0, 1]$ を満たす一様乱数, n は都市数である。 $[c_1 r_1 (n + 1)]$ は $c_1 r_1 (n + 1)$ の整数部分を表している。図 1 の例では $pbest_i' = (5, 4, 8, 7)$ と $lbest_i' = (8, 9, 6)$ を抜き出している。

II $pbest'$ の再形成

$pbest_i'$ と $lbest_i'$ に共通した都市が含まれていれば, $pbest_i'$ から該当した都市を削除し, 残った都市で総経路長が最も短くなるよう部分経路を再形成する。図 1 の例では都市 8 が共通しているため, $pbest_i' = (5, 4, 8, 7)$ から都市 8 を削除し, $pbest_i' = (5, 4, 7)$ を再形成している。

III x_i' の形成

x_i から $pbest_i', lbest_i'$ と共通する都市を削除し, 残った都市で総経路長が最も短くなるよう巡回路を再形成し, x_i' とする。図 1 の例では, $pbest_i'$ と $lbest_i'$ にある都市 4, 5, 6, 7, 8, 9 を x_i から削除し, $x_i' = (1, 3, 2)$ としている。

IV $pbest_i'$ の挿入

$pbest_i'$ を x_i' に総経路長が最も短くなるよう挿入する。

図 1 の例では都市 1 と都市 3 の間に $pbest_i'$ を挿入している。

V $lbest_i'$ の挿入

$lbest_i'$ を x_i' に総経路長が最も短くなるよう挿入する。図

1 の例では都市 5 と都市 3 の間に $lbest_i'$ を挿入している。以上の I ~ V の操作を全粒子で行う。

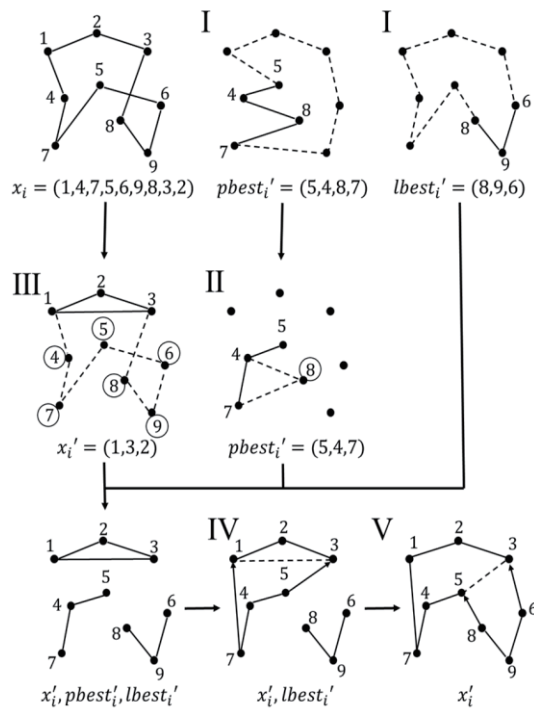


図 1 : 解の更新の例

③総経路長の計算

全粒子が持つ巡回路の総経路長の計算を行う。

④近傍の更新

全粒子間の距離を再計算し, 近傍を更新する。

⑤ $pbest$, $lbest$, $gbest$ の更新

全粒子の $pbest$, $lbest$ と $gbest$ を更新する。

初期設定を①で行い, ②から⑤までの操作を 1 イテレーションとし, 事前に設定したイテレーション数繰り返すことで TSP の解空間を探索する。

2.3 既存手法の問題点

既存手法において, 解 x の更新は各粒子の今までの最良解 $pbest$ と近傍の最良解 $lbest$ を用いて行われる。近傍数が 2 である場合, 近くにいる 2 粒子のうち良い粒子の解が $lbest$ となるため, それぞれの粒子が参照できる粒子の情報が少ない。これでは, 離れた粒子との組み合わせを試さないうちに探索が終了してしまい, 全粒子の初期解の周辺を探索する多点探索とは言えない。

3. 提案手法

提案手法における解の更新は, 各粒子 i のそれまでの最良解 $pbest_i$, 近傍の粒子の解である $lbest_i$, 最速の粒子の解である x_i^f を現在の解 x_i に重ね合わせた経路集合 G を用いて行

れる. まず, ある都市 i をランダムに選択する. そして G に含まれる経路から, 次式で定義される確率 P_{ij} で次の経路 ij を選択する.

$$P_{ij} = \frac{w_{ij}}{\sum_{l=1}^n w_{il}} \quad (6)$$

$$w_{ij} = \frac{c}{(d_{ij})^D}$$

$$C = C_1 + C_2 + C_3 + C_4$$

$$C_1 = \begin{cases} c_1 & x_i \text{に含まれる場合} \\ 0 & \text{otherwise} \end{cases}$$

$$C_2 = \begin{cases} c_2 & pbest_i \text{に含まれる場合} \\ 0 & \text{otherwise} \end{cases}$$

$$C_3 = \begin{cases} c_3 & lbest_i \text{に含まれる場合} \\ 0 & \text{otherwise} \end{cases}$$

$$C_4 = \begin{cases} c_4 & x_i^f \text{に含まれる場合} \\ 0 & \text{otherwise} \end{cases}$$

ここで, n は都市数, $D, c_1 \sim c_4$ はパラメータを表している. (6) 式は G に含まれる経路のうち, 距離が短く, 複数の解に含まれる経路を選択しやすくなるように設定している. また, この経路集合 G のパラメータは, それぞれの粒子が自分の解の周辺を探索するため, c_1 が最も高くなるように設定している. G に選択できる経路が存在しない場合, 未訪問都市の経路の中から距離情報を用いたルーレット選択を用いて経路を選択する. この操作を繰り返すことで巡回路を構築していく.

4. 結果

提案手法の有効性を確認するため, TSPLIB に掲載されている TSP のベンチマーク問題である rd100, kroA150, pr299 を用いて評価実験を行った. 既存手法は実際にアルゴリズムを再現し, 事前実験で最も結果が良かった粒子数 $m = 64$, 近傍数 $k = 2$, $c_1 = 0.9$, $c_2 = 0.1$ というパラメータを使用した. また, 提案手法における c_1, c_2, c_3, c_4 もまた, 事前実験で最も結果が良かった $c_1 = 0.6$, $c_2 = 0.2$, $c_3 = 0.1$, $c_4 = 0.1$ を使用した. 終了条件は rd100 と kroA150 は 30000 イテレーション, pr299 は 500000 イテレーションとした. 各問題 50 試行平均の結果を表 1 ~ 3 に示す. また, 表中で用いられている誤差率は, 試行内で得られた最良解の厳密解に対する誤差の割合を表し, 最終更新イテレーションは最後に $gbest$ を更新したイテレーションを表している.

表 1: rd100 の 50 試行の結果

rd100(厳密解 7910)	既存	提案
厳密解到達率(%)	88	98
平均誤差率(%)	0.0099	0.0002
平均最終更新イテレーション	3210.4	15565.66

表 2: kroA150 の 50 試行の結果

kroA150(厳密解 26524)	既存	提案
厳密解到達率(%)	10	20
平均誤差率(%)	0.25	0.18
平均最終更新イテレーション	6951.2	27379.68

表 3: pr299 の 50 試行の結果

pr299(厳密解 48191)	既存	提案
厳密解到達率(%)	0	2
平均誤差率(%)	0.91	0.29
平均最終更新イテレーション	254626.8	377088.3

実験の結果, 全ての問題において, 既存手法よりも提案手法の精度が上回ったことが分かる. しかし, 平均最終更新イテレーションは提案手法の方が軒並み長くなっている. これは, 提案手法の方が既存手法よりも広範囲で解を探索しているため, 解の収束が遅くなっていることが理由であると考えられる.

5. 今後の課題

今後の課題として, 解の更新の見直しをする必要があることが挙げられる. 提案手法を用いて pr299 の厳密解カバー率の計算を 50 試行を行い, その平均の推移を図 2 に示す. 厳密解カバー率とは, 全粒子の経路を合わせて厳密解の経路をどれだけ保持しているかを割合で表したものである.

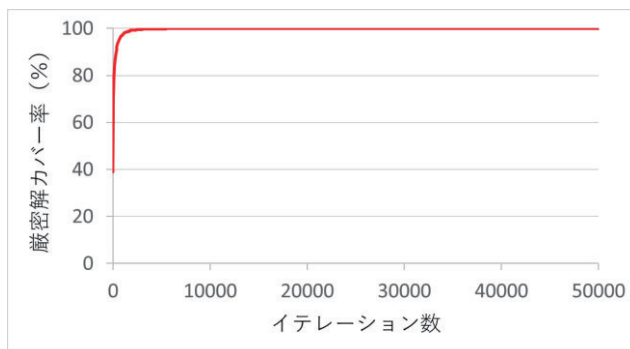


図 2: 提案手法における pr299 の厳密解カバー率

図 2 より, 2000 イテレーションほどで厳密解カバー率が 100% になり, その後も 100% を維持していることが分かる. 厳密解の経路を全て保持しているにも関わらず, 厳密解に収束しないということは, 解の探索範囲は拡大し, 全粒子が多様な経路を保持しているものの, その組み合わせが効率よく行われていないことが考えられる. そこで, 巡回路を重ね合わせて経路を選択する際に, 様々な粒子の情報を参照するような経路の組み合わせ方法を考案し, 更なる大規模問題に挑戦していきたいと考えている.

参考文献

- [Kennedy 95] J.Kennedy, R.C.Eberhart, : "Particle swarm optimization" IEEE International Conf. on Neural Networks, pp.1942-1948 (1995)
- [本庄 16] 本庄将也, 飯塚博幸, 山本雅人, 古川正志, : "巡回セールスマン問題に対する粒子群最適化の提案と性能評価", 日本知能情報フアジィ学会誌, vol.28, no.4, pp.744-755 (2016)

U12バスケットボールにおけるリーグ戦スケジューリング

League Scheduling for U12 Basketball

沖本 天太^{*1}

Tenda Okimoto

西村 一輝^{*2}

Kazuki Nishimura

平山 勝敏^{*1}

Katsutoshi Hirayama

^{*1}神戸大学大学院海事科学研究科

Faculty of Maritime Sciences, Kobe University

^{*2}神戸大学海事科学部

Graduate School of Maritime Sciences, Kobe University

Sports Scheduling is one of the widely investigated application problems in Artificial Intelligence and Operations Research. This problem can be represented as a combinatorial optimization problem, in which the date and the venue of each game must be determined, subject to a given set of constraints. In 2018, Japan Basketball Association (JBA) has started to implementing the league games in the prefectures. In this paper, a formal framework for U12 Basketball League Scheduling (BLS^{U12}) problem is introduced. Furthermore, a novel solution criteria called egalitarian solution for BLS^{U12} is defined. In the experiments, we use the real data of U12 basketball league games played in Hyogo prefecture in 2018 and find an optimal league scheduling which can reduce the total travel distance of teams compared to that of actually used. An egalitarian solution for the total travel distance is also provided.

1. はじめに

スポーツ・スケジューリング [3, 4] 問題とは、チーム、対戦日、開催場所等の集合に対して、与えられた制約条件を満たすように対戦表を作成する組合せ最適化問題である。スポーツ・スケジューリングは AI や OR 分野における応用問題として広く研究されており、応用例として、アメリカ西海岸大学対抗バスケットボール [1] やプロサッカーリーグ [2] 等が挙げられる。

日本バスケットボール協会 (JBA)^{*1} は (i) 拮抗したゲームを多くすることで選手及び指導者の成長を促し、(ii) JBA 登録チームに対して、一定公式試合数の確保を目的とし、都道府県内でのリーグ戦の実施を 2018 年度より開始した。兵庫県ミニバスケットボール連盟^{*2} は 2018 年度、県内の男女合わせて約 180 チームを対象に前後期リーグ戦を実施した。前期 (5 月から 7 月) では、各チームは所属地区で実施された新人戦の成績を基に、A (上位)、B1, B2, B3 (中位)、C (下位) の 5 つのリーグに分けられ、県内の各会場で 10 試合の公式戦を行った。また後期 (9 月から 12 月) では、前期の成績を基に各リーグのチームが再編成され 10 試合の公式戦が行われた。

リーグ戦作成では、前後期各 10 試合の計 20 試合を実施しなければならないとする総試合数に関する制約や、同じチームと 2 試合以上対戦してはならないとする重複試合の禁止等の満たさなければならない様々な制約条件が存在する。このため、すべての制約条件を満たすようなリーグ戦を作成するのは困難な問題である。実際、兵庫県ミニバスケットボール U12 の前期 A リーグ (女子) では重複試合が 2 件発生している。さらに、現状のリーグ戦は人手により数日かけて作成されているため、その負担を軽減する必要がある。また、各チームの総移動距離やチーム間の移動距離の平等性等は現在考慮されていない。

本論文では、移動距離最小化問題のフレームワークを用いて、ミニバスケットボールにおけるリーグ戦作成 (U12 Basketball League Scheduling, BLS^{U12}) 問題を定義する。実験では、2018 年度に兵庫県下で実施されたミニバス U12 の後期 A リーグ

(女子) の実データを用いて、与えられた制約条件を満たし、かつ、各チームの移動距離の総和が最小となるようなリーグ戦及び、移動距離の最大値を最小化するようなリーグ戦をそれぞれ作成し、実際に用いられたリーグ戦と比較評価する。

2. U12バスケットボールリーグ戦作成問題

ミニバスケットボール (ミニバス) におけるリーグ戦作成 (BLS^{U12}) 問題を定義する。まず BLS^{U12} の基本用語を与える。

- $T = \{1, \dots, n\}$: チームの集合。
- $D = \{1, \dots, m\}$: 日付 (対戦日: 休日・祝日) の集合。
- $X = \{x_{ij} \mid i, j \in T (i \neq j)\}$: 変数の集合。 $x_{ij} = k (k \in D)$ とは、チーム i と j が k 日に対戦することを表す。
- $C = \{c_1, \dots, c_l\}$: 制約の集合。
- $T_s = \{q_1, \dots, q_t\}$: (各チームの) 拠点校の集合。
- $P = \{p_1, \dots, p_s\}$: 対戦が行われる試合会場の集合。
- $\alpha: T \rightarrow T_s$: 各チームの拠点校を返す写像。
- $\beta: D \rightarrow P$: 日付から試合会場を返す写像。
- $dis: T_s \times P \rightarrow \mathbb{R}$: 拠点校から会場までの距離を返す関数。

以下、兵庫県ミニバス U12 で用いられた制約条件を示す。

- 制約 1 (総試合数の制限): 各チームの対戦数は 10 試合とする。各チームは前期 10、後期 10 の計 20 試合を行う。
- 制約 2 (毎試合数の制限): U12 では選手の体力等を考慮して、各チーム、試合がある場合は 1 日に 2 試合行う。
- 制約 3 (各会場での試合数の制限): 各試合会場では、その日に対戦しなければならない対戦数が決められている。
- 制約 4 (重複試合の禁止): 同じ相手と 2 試合以上対戦してはならない (同じリーグ内の他の 10 チームと対戦する)。

連絡先: 沖本 天太, 神戸大学大学院海事科学研究科, 神戸市東灘区深江南町 5-1-1, tenda@maritime.kobe-u.ac.jp

^{*1} <http://www.japanbasketball.jp>

^{*2} <https://hyogo-minibasket.jimdo.com>

表 1: 提案手法と+Minimax における各チームの総移動距離。

チーム名	成徳	北エン	福田	香櫨園	塚口	武庫	浜脇	多田東
提案手法	836km	608km	636km	888km	944km	876km	886km	816km
+Minimax	878km	800km	636km	820km	902km	776km	828km	850km
チーム名	西宮浜	宝塚 BR	川西	荻野	北淡	松陽	三田	氷上東
提案手法	842km	932km	854km	834km	596km	642km	524km	270km
+Minimax	878km	794km	854km	902km	596km	642km	558km	270km

表 2: 後期 A リーグ (女子 16 チーム) における, 実際のリーグ戦と提案手法で得られたリーグ戦の総移動距離と作成時間。

後期リーグ	実際のリーグ戦	提案手法	+Minimax
総移動距離	14, 204km	11, 984km	11, 984km
作成時間	数日	4 秒	173 秒

次に, U12 におけるリーグ戦作成 (BLS^{U12}) 問題を定義する。

定義 1 (BLS^{U12}). ミニバスにおけるリーグ戦作成 (BLS^{U12}) は, X を変数の集合, D を変数値の集合, C を制約の集合, T_s を拠点校の集合, P を試合会場の集合, α を各チームの拠点校を返す写像, β を日付から試合会場を返す写像, ϕ を拠点校から試合会場までの往復距離を計算する関数とし,

$$BLS^{U12} = \langle X, D, C, T_s, P, \alpha, \beta, \phi \rangle$$

の組により定義される。全変数への割当を A とし, 関数 ϕ は

$$\phi(A) = \sum_{i,j,k} 2 \cdot [dis(\alpha(i), \beta(k)) + dis(\alpha(j), \beta(k))]$$

により与えられる ($1 \leq i, j \leq n, i \neq j, 1 \leq k \leq m$). BLS^{U12} を解くとは, 全ての制約条件を満たし, 各チームの移動距離の総和 $\phi(A)$ が最小となるような割当 A を見つけることである。

定義 2 (BLS^{U12} 問題)。

- 入力: ミニバスにおけるリーグ戦作成 BLS^{U12} ,
- 質問: 移動距離の総和が最小となるリーグ戦をみつけよ。

リーグ戦作成 BLS^{U12} において, 各チーム i ($1 \leq i \leq n$) の総移動距離を d_i とし, すべてのチームの総移動距離を \vec{d} -ベクトルと呼び, $\vec{d} = (d_1, \dots, d_n)$ と記述する。また 2 つの \vec{d} -ベクトル \vec{d}' 及び \vec{d}'' に関して, 両者の各要素の総和が等しいとき, すなわち, $\sum_i d'_i = \sum_i d''_i$ が成立するとき, \vec{d}' と \vec{d}'' は同等であるという。同等な \vec{d} -ベクトルからなる集合を \overline{D} と記述する。また \overline{D} 上の前順序は \preceq_{lex} により与えられるものとする。すなわち, $\forall \vec{d}', \vec{d}'' \in \overline{D}$ に関して, \vec{d}' が辞書式順序において \vec{d}'' より先行している, またそのときに限り, $\vec{d}' \preceq_{lex} \vec{d}''$ と定義する。ある \vec{d} -ベクトル \vec{d}' が平等であるとは, $\vec{d}'' \preceq_{lex} \vec{d}'$ となるような, \vec{d}' と同等な他の \vec{d}'' が存在しないことを意味する。

3. 評価実験

2018 年度に兵庫県下で実施されたミニバス U12 後期 A リーグ戦 (女子 16 チーム) の実データを用いて BLS^{U12} 問題を求解した。具体的には, BLS^{U12} 問題を 0-1 整数計画問題として定式化し, 最適化ソルバー CPLEX を用いて, 各チームの移

動距離の総和が最小となるリーグ戦及び, 総移動距離が最小, かつ, 移動距離の最大値が最小となるリーグ戦を作成した。

実際に使用された U12 後期 A リーグ戦と提案手法によって作成されたリーグ戦における総移動距離及び実行時間を表 2 に示す (両者とも全ての制約条件 1 から 4 を満たしている)。また, 表中の+Minimax は提案手法とミニマックス値を用いたときの実験結果を表す。表 2 より, 実際に使用された後期リーグ戦での全 16 チームの総移動距離は 14, 204km であった。これに対し, 提案手法では, 全 16 チームの総移動距離は 11, 984km であり, 実際に使用されたものとの差は 2, 222km であった。また, 実際のリーグ戦は人手により数日かけて作成されているのに対し, 提案手法では最適なりリーグ戦が 4 秒, +Minimax では最適かつ, 移動距離の最大値が最小となるリーグ戦が 173 秒で求解可能であった。表 1 は提案手法及び, +Minimax によって得られた各チームの総移動距離を表している。提案手法では, 塚口の総移動距離が 944km (最大値) であるのに対し, +Minimax では 902km であった。このように, +Minimax で得られた解では, 総移動距離が大きい塚口や宝塚 BR の負担を他のチームでシェアすることで軽減しているのが分かる。

4. おわりに

本論文では, ミニバスケットボールにおけるリーグ戦作成 (BLS^{U12}) 問題を定義し, 各チームの移動距離の総和が最小となるような最適なりリーグ戦及び, 最適かつ移動距離の最大値が最小となるようなリーグ戦を作成した。実験では, 2018 年度に兵庫県下で実施されたミニバス U12 後期 A リーグ戦 (女子) の実データを用いて, BLS^{U12} 問題を求解し, 実際に使用されたリーグ戦と比較評価した。実験結果より, 提案手法では最適なりリーグ戦が 4 秒, 最適かつ移動距離の最大値が最小となるようなリーグ戦が 173 秒で作成可能であることが分かった。

今後の課題として, 現場で利用可能なシステムの開発, U15 及び U18 バスケットボールにおけるリーグ戦作成, その他のスポーツ・スケジューリング問題への適用等が挙げられる。

参考文献

- [1] M. Henz. Scheduling a major college basketball conference - revisited. *Operations Research*, 49:163–168, 2002.
- [2] C. Ribeiro. Sports scheduling: Problems and applications. *International Transactions in Operational Research*, 19:201–226, 2012.
- [3] 宮代隆平 and 松井知己. スポーツスケジューリングリーダ未解決問題を中心にー. *オペレーションズリサーチ*, 50:119–124, 2005.
- [4] 池辺淑子. スポーツのスケジューリング. *オペレーションズリサーチ*, 51:392–395, 2006.

[3K3-J-2] Machine learning: analysis and validations of models

Chair:Masahiro Suzuki Reviewer:Satoshi Oyama

Thu. Jun 6, 2019 1:50 PM - 3:30 PM Room K (201A Medium meeting room)

[3K3-J-2-01] Statistical Mechanical Formulation of Learning Dynamics of Two-Layered Neural Networks with Batch Normalization

○Shiro Takagi¹, Yuki Yoshida¹, Masato Okada¹ (1. Graduate School of Frontier Sciences, The University of Tokyo)

1:50 PM - 2:10 PM

[3K3-J-2-02] On the trade-off between the number of nodes and the number of trees in Random Forest

○So Kumano¹, Tatsuya Akutsu¹ (1. Kyoto University)

2:10 PM - 2:30 PM

[3K3-J-2-03] Do the AUC and log-loss evaluate CTR prediction models properly?

○Satoshi KATAGIRI¹ (1. F@N Communications, Inc.)

2:30 PM - 2:50 PM

[3K3-J-2-04] Social reinforcement learning with shared global aspiration for satisficing

○Noriaki Sonota¹, Takumi Kamiya², Tatsuji Takahashi¹ (1. Tokyo Denki University, 2. Graduate School of Tokyo Denki University)

2:50 PM - 3:10 PM

[3K3-J-2-05] On the mathematical approach to the "photo-likeness" of images

○Yasuhiko Asao¹, Ryotaro Sakamoto¹ (1. Graduate School of Mathematical Science, the University of Tokyo)

3:10 PM - 3:30 PM

Batch Normalization つき 3 層ニューラルネットワークの学習ダイナミクスの統計力学的定式化

Statistical Mechanical Formulation of Learning Dynamics of Two-Layered Neural Networks with Batch Normalization

高木志郎 吉田雄紀 岡田真人
Shiro Takagi Yuki Yoshida Masato Okada

東京大学大学院新領域創成科学研究科
Graduate School of Frontier Sciences, The University of Tokyo

Batch Normalization is known as a method to shorten training time, stabilize training and improve the performance of neural networks. Despite its wide use, the impact of Batch Normalization on the learning dynamics of neural networks is yet to be clarified. Though some recent studies tried to tackle this problem, few of them derived the exact learning dynamics of neural networks with Batch Normalization. Because deriving the learning dynamics is helpful for understanding what Batch Normalization is doing during training, we derived an exact learning dynamics of two-layered neural networks with Batch Normalization by drawing on the previous work about a statistical mechanical method of neural network analysis. Specifically, for neural networks with Batch Normalization, we derived differential equations of order parameters, which represent a macroscopic behavior of neural networks.

1. はじめに

ニューラルネットワークの学習を高速化、安定化させる手法として、Ioffe と Szegedy が提案した Batch Normalization という手法がある [Ioffe 15]. これは中間層への入力を正規化することで学習の高速化を行う手法であるが、学習の高速化だけではなく、学習率や重みの初期値などのパラメータの設定を容易にしたり、正則化の効果を持っていたりと、ニューラルネットワークの学習を容易にする様々な効果が経験的に知られている。しかし、Batch Normalization がニューラルネットワークの学習に与える影響についての理解は依然として不十分である。Ioffe と Szegedy は、各層への入力が必要で前までの重みの変更依存のために生じる「内的共変量シフト」を Batch Normalization が低減できるため、学習を容易にすると主張した [Ioffe 15]. 一方 Santurkar らは、Batch Normalization は内的共変量シフトとは関係なく、むしろ誤差曲面を滑らかにすることで予測しやすい安定な勾配が計算できるようにし、学習効率を向上させると主張した [Santurkar 18]. Bjorck らは、Batch Normalization が可能にする大きな学習率が、正則化の効果を持つことによって、学習の高速化だけでなく汎化性能の向上をもたらすことを示した [Bjorck 18]. Kohler らは、Batch Normalization は重みベクトルの最適化を長さの最適化と方向の最適化に切り分けることによって最適化を容易にすると主張した [Kohler 18]. また Arora らは、Batch Normalization が学習率の自動調節をもたらすことによって最適な収束率を実現することを、滑らかな誤差関数を用いた学習に対して示した [Arora 18]. これらはいずれも理論的な解析により Batch Normalization の効果について示唆を与えるものだが、ニューラルネットワークの学習中に重みや誤差が具体的にどのように振る舞うかについては議論ができていない。ニューラルネットワークのパラメータや誤差のダイナミクスを解析的に導出するのは一般に困難である。そのため、理想化された単純な系であっても学習のダイナミクスを求めることは重要である。そこで、私たちは 90 年代に考案された統計力学的手法を用いて Batch Normalization を適用した 3 層ソフトコミティの学習ダイナミクスを導出した。ソフトコミティとは中間層から出力層への重みを定数に固定した場合のニュー

ラルネットワークであり、解析の簡単のため用いられることがある。統計力学的手法とは大規模ネットワークを仮定することで、系の大域的な挙動を記述するパラメータであるオーダーパラメータと訓練誤差の期待値として定義される汎化誤差のダイナミクスを解析的に導出する手法である [Schwarze 93, Seung 92, Saad 95, Biehl 95, Riegler 95]. 統計力学的手法を用いて Batch Normalization のダイナミクスを解析した研究としては、Luo らの研究がある [Luo 18]. しかしこれは単層パーセプトロンに議論を限定しており、中間層がある場合のダイナミクスは扱っていない。我々は中間層がある 3 層ニューラルネットワークについてダイナミクスを導出した。

2. Batch Normalization つき 3 層ニューラルネットワークの統計力学的定式化

2.1 統計力学定式化

統計力学的定式化では一般に教師生徒型ニューラルネットワークのオンライン学習を考える [Saad 95, Biehl 95]. ここで教師生徒型学習とは、図 1 に示すように、学習器と同じ構造を持つニューラルネットワークを教師データの生成モデルと仮定する教師あり学習を指し、オンライン学習とは各更新毎に新しく生成されるサンプルサイズ 1 のデータを用いた確率的勾配降下法による学習のことを指す。この時生成モデルを教師ネットワーク、学習器を生徒ネットワークと呼ぶ。

入力素子数 N 、生徒の中間素子数が K 、教師の中間素子数が M 、出力素子数が O の 3 層ニューラルネットワークを考える。入力 $\xi \in \mathbb{R}^N$ の各成分は期待値 0 分散 σ^2 の分布から i.i.d. にサンプリングされるとする。生徒ネットワークの第 1 層の重み行列を $[\mathbf{J}_1, \dots, \mathbf{J}_K]^T \in \mathbb{R}^{K \times N}$ 、第二層の重み行列を $[\mathbf{w}_1, \dots, \mathbf{w}_K]^T \in \mathbb{R}^{O \times K}$ 、教師ネットワークの第 1 層の重み行列を $[\mathbf{B}_1, \dots, \mathbf{B}_M]^T \in \mathbb{R}^{M \times N}$ 、第二層の重み行列を $[\mathbf{v}_1, \dots, \mathbf{v}_M]^T \in \mathbb{R}^{O \times M}$ と表記する。生徒と教師の第 1 層の重みベクトルは $\mathbf{J}_i \in \mathbb{R}^N \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1/N)$ 、 $\mathbf{B}_n \in \mathbb{R}^N \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1/N)$ と初期化をする。今ソフトコミティを考えているので $\mathbf{w}_i \in \mathbb{R}^O$ 、 $\mathbf{v}_n \in \mathbb{R}^O$ は要素が定数の O 次元のベクトルで、値が不変である。ただし i, n はそれぞれ生徒と教師の中間層の素子のインデックスである。中間層の活性化関数を ϕ と

し、出力層の活性化関数は恒等写像とする。このとき、生徒と教師のネットワークの出力はそれぞれ、

$$\mathbf{s} \in \mathbb{R}^O = \sum_i^K \mathbf{w}_i \phi(\mathbf{J}_i \cdot \boldsymbol{\xi}), \quad (1)$$

$$\mathbf{t} \in \mathbb{R}^O = \sum_n^M \mathbf{v}_n \phi(\mathbf{B}_n \cdot \boldsymbol{\xi}), \quad (2)$$

と書ける。損失関数としては二乗損失 $\varepsilon = \frac{1}{2} \|\mathbf{t} - \mathbf{s}\|^2$ を用いて、生徒ネットワークの重みを教師ネットワークの重みに近づけていく。ここで、系の大域的な挙動を記述するパラメータであるオーダーパラメータを次のように定義する： $Q_{ij} = \mathbf{J}_i \cdot \mathbf{J}_j$, $R_{in} = \mathbf{J}_i \cdot \mathbf{B}_n$, $T_{nm} = \mathbf{B}_n \cdot \mathbf{B}_m$, $D_{ij} = \mathbf{w}_i \cdot \mathbf{w}_j$, $E_{in} = \mathbf{w}_i \cdot \mathbf{v}_n$, $F_{nm} = \mathbf{v}_n \cdot \mathbf{v}_m$ [Saad 95, Biehl 95, Yoshida 18]. 入力素子数 N が十分に大きい時、活性化関数によってはいくつかの理想化のもとでこれらのパラメータの微分方程式を導出することができる [Saad 95, Biehl 95]. また、訓練誤差の $\boldsymbol{\xi}$ についての期待値として定義される汎化誤差 ε_g はオーダーパラメータの関数となるので、汎化誤差のダイナミクスも導出することができる。Saad と Solla は 3 層ソフトコミティの入力層から中間層への重みについて、Biehl と Schwarze は一般の 3 層を対象として入力層から中間層への重みについて、Yoshida らは 3 層の全ての重みについて、オーダーパラメータのダイナミクスを導出している [Saad 95, Biehl 95, Yoshida 18]. 我々はこのうちソフトコミティについてのダイナミクスを Batch Normalization ありの場合に拡張した。

2.2 Batch Normalization の統計力学定式化

従来の統計力学的定式化ではサンプルサイズ 1 の学習を取り扱っていたため、これを Batch Normalization を取り扱うように拡張した。各更新毎に新しく b 個の入力を i.i.d. にサンプリングし、それを用いて学習を行うものとする。サンプルサイズ b のデータの中の 2 つのサンプル $\boldsymbol{\xi}^u$, $\boldsymbol{\xi}^v$ が互いに無相関だと仮定する。すると、 Q_{ij} のダイナミクスに一部修正を加えるだけで、Batch Normalization を取り扱うように統計力学的手法を自然に拡張できることを確認した。

Batch Normalization では、各中間素子への入力 $x_i^u = \mathbf{J}_i \boldsymbol{\xi}^u$ それぞれに対して、ミニバッチデータについての算術平均と標準偏差で正規化したものに学習可能パラメータ g_i をかけて β_i を足したものを活性化関数への入力とする。ここでは解析の簡単のため、算術平均を引く操作と β を足す操作を行わず、標準偏差は定数とする。今、入力に期待値 0 の分布を仮定しているので、サンプルサイズ b が十分大きいとき、中間層の各素子の b 個の入力についての標準偏差 $\sigma_{x_i} = \sqrt{\frac{1}{b} \sum_{u=1}^b (x_i^u - \mu_i)^2}$ は $\sigma_{x_i} \approx \sqrt{\frac{b}{b-1} \langle x_i^2 \rangle} \approx \sqrt{\mathbf{J}_i^T \langle \boldsymbol{\xi}^u \boldsymbol{\xi}^{uT} \rangle \mathbf{J}_i} = \sqrt{\sigma^2 \|\mathbf{J}_i\|^2} = \sigma \sqrt{Q_{ii}}$ となり、 $\boldsymbol{\xi}$ に依存しなくなる。ただし $\langle \cdot \rangle$ は入力 $\boldsymbol{\xi}$ についての期待値をとる操作である。この時、生徒の出力は、

$$\mathbf{s}^u = \sum_i^K \mathbf{w}_i \phi\left(\frac{g_i}{\sigma \sqrt{Q_{ii}}} \mathbf{J}_i \boldsymbol{\xi}^u\right) = \sum_i^K \mathbf{w}_i \phi\left(\frac{g_i}{\sigma \sqrt{Q_{ii}}} x_i^u\right), \quad (3)$$

となる。そして生徒の第 1 層の重みと学習可能パラメータ g_i の更新式はそれぞれ、

$$\Delta \mathbf{J}_i = \frac{\eta}{Nb} \sum_{u=1}^b [(\mathbf{t}^u - \mathbf{s}^u) \cdot \mathbf{w}_i] \phi'\left(\frac{g_i}{\sigma \sqrt{Q_{ii}}} x_i^u\right) \frac{g_i}{\sigma \sqrt{Q_{ii}}} \boldsymbol{\xi}^u, \quad (4)$$

$$\Delta g_i = \frac{\eta}{Nb} \sum_{u=1}^b [(\mathbf{t}^u - \mathbf{s}^u) \cdot \mathbf{w}_i] \phi'\left(\frac{g_i}{\sigma \sqrt{Q_{ii}}} x_i^u\right) \frac{\mathbf{J}_i \boldsymbol{\xi}^u}{\sigma \sqrt{Q_{ii}}}, \quad (5)$$

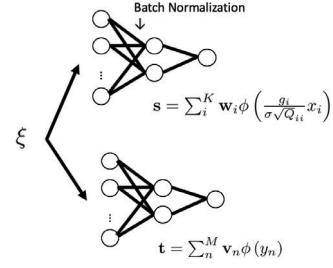


図 1: Batch Normalization ありの場合の 3 層ソフトコミティの教師生徒型学習の図。入力素子数が N , 中間素子数が $K = 2$, $M = 2$, 出力素子数が $O = 1$ の場合。 $\boldsymbol{\xi}$ を共通の入力として出力された \mathbf{s} と \mathbf{t} の間の誤差を小さくするように生徒が重みを調節する。

と書ける。ただし $\frac{\eta}{N}$ は学習率を表す。 $\frac{g_i}{\sigma \sqrt{Q_{ii}}} x_i = \hat{x}_i$ と書くと、オーダーパラメータと g_i の更新式は、

$$\begin{aligned} \Delta Q_{ij} &= \frac{\eta}{bN} \sum_{u=1}^b \left[\sum_{p=1}^M E_{ip} \phi'(\hat{x}_i^u) \hat{x}_j^u \phi(y_p^u) - \sum_{p=1}^K D_{ip} \phi'(\hat{x}_i^u) \hat{x}_j^u \phi(\hat{x}_p^u) \right. \\ &\quad \left. + \sum_{p=1}^M E_{jp} \phi'(\hat{x}_j^u) \hat{x}_i^u \phi(y_p^u) - \sum_{p=1}^K D_{jp} \phi'(\hat{x}_j^u) \hat{x}_i^u \phi(\hat{x}_p^u) \right] \\ &\quad + \frac{\eta^2}{b^2 N^2} \sum_{u,v} \boldsymbol{\xi}^u \boldsymbol{\xi}^v \left[\sum_{p,q}^{K,K} D_{ip} D_{jq} \phi'(\hat{x}_i^u) \phi'(\hat{x}_j^v) \phi(\hat{x}_p^u) \phi(\hat{x}_q^v) \right. \\ &\quad \left. + \sum_{p,q}^{M,M} E_{ip} E_{jq} \phi'(\hat{x}_i^u) \phi'(\hat{x}_j^v) \phi(y_p^u) \phi(y_q^v) \right. \\ &\quad \left. - \sum_{p,q}^{K,M} D_{ip} E_{jq} \phi'(\hat{x}_i^u) \phi'(\hat{x}_j^v) \phi(\hat{x}_p^u) \phi(y_q^v) \right. \\ &\quad \left. - \sum_{p,q}^{M,K} E_{ip} D_{jq} \phi'(\hat{x}_i^u) \phi'(\hat{x}_j^v) \phi(y_p^u) \phi(\hat{x}_q^v) \right], \quad (6) \end{aligned}$$

$$\begin{aligned} \Delta R_{in} &= \frac{\eta}{bN} \sum_{u=1}^b \left[\sum_{p=1}^M E_{ip} \phi'(\hat{x}_i^u) y_n^u \phi(y_p^u) - \sum_{p=1}^K D_{ip} \phi'(\hat{x}_i^u) y_n^u \phi(\hat{x}_p^u) \right], \quad (7) \end{aligned}$$

$$\begin{aligned} \Delta g_i &= \frac{\eta}{bN g_i} \sum_{u=1}^b \left[\sum_{p=1}^M E_{ip} \phi'(\hat{x}_i^u) \hat{x}_i^u \phi(y_p^u) - \sum_{p=1}^K D_{ip} \phi'(\hat{x}_i^u) \hat{x}_i^u \phi(\hat{x}_p^u) \right], \quad (8) \end{aligned}$$

と書ける。また、汎化誤差は、

$$\begin{aligned} \varepsilon_g &= \frac{1}{2} \left[\sum_{p,q}^{M,M} F_{pq} \phi(y_p^u) \phi(y_q^u) + \sum_{p,q}^{K,K} D_{pq} \phi(\hat{x}_p^u) \phi(\hat{x}_q^u) \right. \\ &\quad \left. - 2 \sum_{p,q}^{K,M} E_{pq} \phi(\hat{x}_p^u) \phi(y_q^u) \right], \quad (9) \end{aligned}$$

となる。ただし $y_n^u = \mathbf{B}_n \boldsymbol{\xi}^u$ である。この更新式を $\boldsymbol{\xi}$ について期待値を取ったものは、活性化関数によっては厳密に計算すること

ができる [Saad 95, Biehl 95, Yoshida 18]. $\phi(x) = \text{erf}(x/\sqrt{2})$ のとき, これらのオーダーパラメータと g_i , そして汎化誤差のダイナミクスは以下のように求まる *1:

$$N \frac{dQ_{ij}}{dt} = \frac{2\eta}{\pi} [Q_1 - Q_2] + \frac{4\eta^2 g_i g_j}{b\pi^2 \sqrt{\Lambda Q_{ii} Q_{jj}}} [Q_3 + Q_4 - Q_5 - Q_6], \quad (24)$$

$$N \frac{dR_{in}}{dt} = \frac{2\eta}{\pi} [\mathcal{R}], \quad (25)$$

$$N \frac{dg_i}{dt} = \frac{2\eta}{\pi g_i} [\mathcal{G}]. \quad (26)$$

ただし dt は微小な変化量で, $(l, k) = (i, j, n, p, q)$ について $Q'_{lk} = \frac{\sigma^2 g_l g_k}{\sigma_{x_l} \sigma_{x_k}} Q_{lk}$, $R'_{lk} = \frac{\sigma^2 g_l}{\sigma_{x_l}} R_{lk}$, $T'_{lk} = \sigma^2 T_{lk}$ である.

*1

$$Q_1 = \sum_{p=1}^M \left[\frac{E_{ip} (R'_{jp} (1 + Q'_{ii}) - Q'_{ij} R'_{ip})}{(1 + Q'_{ii}) \sqrt{(1 + Q'_{ii})(1 + T'_{pp}) - R'^2_{ip}}} + \frac{E_{jp} (R'_{ip} (1 + Q'_{jj}) - Q'_{ji} R'_{jp})}{(1 + Q'_{jj}) \sqrt{(1 + Q'_{jj})(1 + T'_{pp}) - R'^2_{jp}}} \right] \quad (10)$$

$$Q_2 = \sum_{p=1}^K \left[\frac{D_{ip} (Q'_{jp} (1 + Q'_{ii}) - Q'_{ij} Q'_{ip})}{(1 + Q'_{ii}) \sqrt{(1 + Q'_{ii})(1 + Q'_{pp}) - Q'^2_{ip}}} + \frac{D_{jp} (Q'_{ip} (1 + Q'_{jj}) - Q'_{ji} Q'_{jp})}{(1 + Q'_{jj}) \sqrt{(1 + Q'_{jj})(1 + Q'_{pp}) - Q'^2_{jp}}} \right] \quad (11)$$

$$Q_3 = \sum_{p,q}^{K,K} D_{ip} D_{jq} \times \text{asin} \left(\frac{\Lambda Q'_{pq} - Q'_{jp} Q'_{jq} (1 + Q'_{ii}) - Q'_{ip} Q'_{iq} (1 + Q'_{jj}) + Q'_{ij} Q'_{ip} Q'_{jq} + Q'_{ij} Q'_{iq} Q'_{jp}}{\sqrt{\Lambda_1 \Lambda_2}} \right) \quad (12)$$

$$Q_4 = \sum_{p,q}^{M,M} E_{ip} E_{jq} \times \text{asin} \left(\frac{\Lambda T'_{pq} - R'_{jp} R'_{jq} (1 + Q'_{ii}) - R'_{ip} R'_{iq} (1 + Q'_{jj}) + Q'_{ij} R'_{ip} R'_{jq} + Q'_{ij} R'_{iq} R'_{jp}}{\sqrt{\Lambda_3 \Lambda_4}} \right) \quad (13)$$

$$Q_5 = \sum_{p,q}^{K,M} D_{ip} E_{jq} \times \text{asin} \left(\frac{\Lambda R'_{pq} - Q'_{jp} R'_{jq} (1 + Q'_{ii}) - Q'_{ip} R'_{iq} (1 + Q'_{jj}) + Q'_{ij} Q'_{ip} R'_{jq} + Q'_{ij} R'_{iq} Q'_{jp}}{\sqrt{\Lambda_1 \Lambda_4}} \right) \quad (14)$$

$$Q_6 = \sum_{p,q}^{M,K} E_{ip} D_{jq} \times \text{asin} \left(\frac{\Lambda R'_{pq} - R'_{jp} Q'_{jq} (1 + Q'_{ii}) - R'_{ip} Q'_{iq} (1 + Q'_{jj}) + Q'_{ij} R'_{ip} Q'_{jq} + Q'_{ij} Q'_{iq} Q'_{jp}}{\sqrt{\Lambda_2 \Lambda_3}} \right) \quad (15)$$

$$\mathcal{R} = \sum_{p=1}^M \frac{E_{ip} (T'_{np} (1 + Q'_{ii}) - R'_{in} R'_{ip})}{(1 + Q'_{ii}) \sqrt{(1 + Q'_{ii})(1 + T'_{pp}) - R'^2_{ip}}} - \sum_{p=1}^K \frac{D_{ip} (R'_{pn} (1 + Q'_{ii}) - R'_{in} Q'_{ip})}{(1 + Q'_{ii}) \sqrt{(1 + Q'_{ii})(1 + Q'_{pp}) - Q'^2_{ip}}} \quad (16)$$

$$\mathcal{G} = \sum_{p=1}^M \frac{E_{ip} (R'_{ip} (1 + Q'_{ii}) - Q'_{ii} R'_{ip})}{(1 + Q'_{ii}) \sqrt{(1 + Q'_{ii})(1 + T'_{pp}) - R'^2_{ip}}} - \sum_{p=1}^K \frac{D_{ip} (Q'_{ip} (1 + Q'_{ii}) - Q'_{ii} Q'_{ip})}{(1 + Q'_{ii}) \sqrt{(1 + Q'_{ii})(1 + Q'_{pp}) - Q'^2_{ip}}} \quad (17)$$

$$\Lambda = (1 + Q'_{ii})(1 + Q'_{jj}) - (1 + Q'^2_{ij}) \quad (18)$$

$$\Lambda_1 = \Lambda(1 + Q'_{pp}) - Q'^2_{jp} (1 + Q'_{ii}) - Q'^2_{ip} (1 + Q'_{jj}) + 2Q'_{ij} Q'_{ip} Q'_{jp} \quad (19)$$

$$\Lambda_2 = \Lambda(1 + Q'_{qq}) - Q'^2_{jq} (1 + Q'_{ii}) - Q'^2_{iq} (1 + Q'_{jj}) + 2Q'_{ij} Q'_{iq} Q'_{jq} \quad (20)$$

$$\Lambda_3 = \Lambda(1 + T'_{pp}) - R'^2_{jp} (1 + Q'_{ii}) - R'^2_{ip} (1 + Q'_{jj}) + 2Q'_{ij} R'_{ip} R'_{jp} \quad (21)$$

$$\Lambda_4 = \Lambda(1 + T'_{qq}) - R'^2_{jq} (1 + Q'_{ii}) - R'^2_{iq} (1 + Q'_{jj}) + 2Q'_{ij} R'_{iq} R'_{jq} \quad (22)$$

$$\varepsilon_g = \frac{1}{\pi} \left[\sum_{p,q}^{M,M} \text{asin} \left(\frac{T'_{pq}}{\sqrt{(1 + T'_{pp})(1 + T'_{qq})}} \right) - \sum_{p,q}^{K,K} \text{asin} \left(\frac{Q'_{pq}}{\sqrt{(1 + Q'_{pp})(1 + Q'_{qq})}} \right) - 2 \sum_{p,q}^{K,M} \text{asin} \left(\frac{R'_{pq}}{\sqrt{(1 + Q'_{pp})(1 + T'_{qq})}} \right) \right] \quad (23)$$

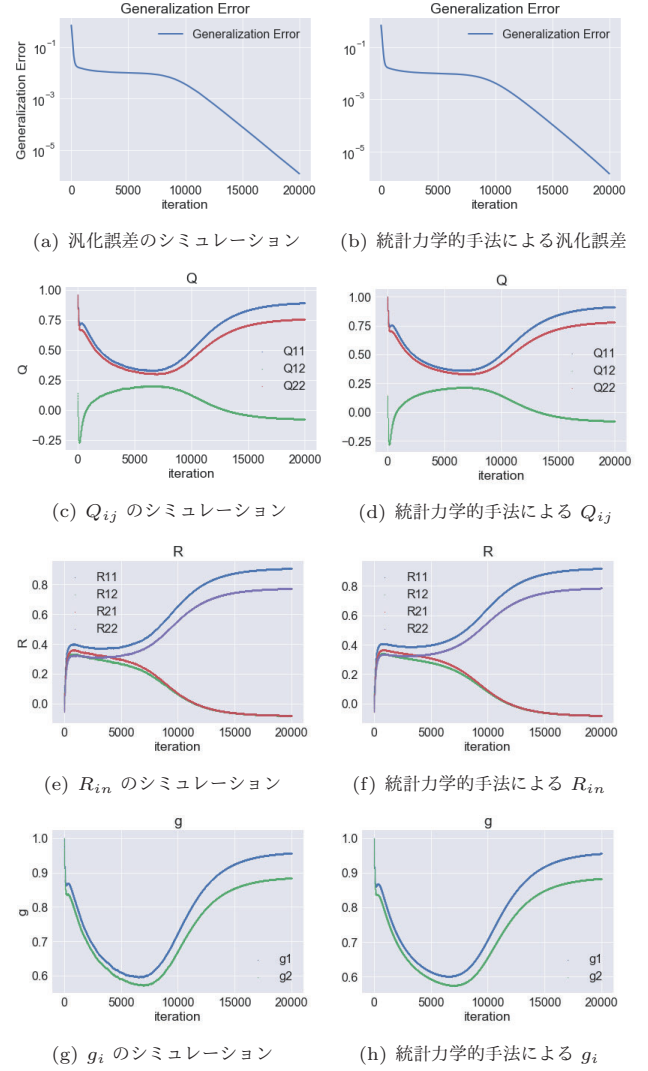


図 2: (a), (c), (e), (g) が数値シミュレーションの結果で (b), (d), (f), (h) が統計力学的手法を用いて導出したダイナミクス. $\eta = 1$, $N = 100$, $b = 100$, $\sigma = 1$, $\mathbf{w}_i = 1$, $\mathbf{v}_n = 1$ とし, 総イテレーション数は 20000 とした

2.3 数値シミュレーションと統計力学の定式化により導出したダイナミクスの一致

統計力学的手法を用いたダイナミクスの導出では近似を用いている. そのため, 統計力学的手法によって導出されたダイナミクスと実際の重みの更新式を用いた数値シミュレーションの結果が一致することを確認する必要がある. そこで $K = 2$, $M = 2$, $O = 1$ の場合について Q_{ij} , R_{in} , g_i および ε_g の時間発展を比較した. 図 2 が数値シミュレーションの結果と統計力学的手法によって導出されたダイナミクスの比較である. 図より, これらの二つの結果はよく一致しており, 統計力学的手法の近似は妥当であることがわかる.

3. まとめ

Saad らが発展させた統計力学的手法を用いて Batch Normalization がある場合の 3 層ニューラルネットワークのオーダーパラメータおよび汎化誤差のダイナミクスを導出した. こ

れを用いれば Batch Normalization がニューラルネットワークの学習挙動にどのような影響を与えるかを解析することができる。

例えば、ニューラルネットワークの学習では学習初期と終期ではデータから学習する構造が異なると考えられており、それがニューラルネットワークが高い表現能力を持ちながら良い汎化性能を示す原因としてあげられることがある [Saxe 18, Xu 18, Krueger 17, Rahaman 18, Arpit 18]. Batch Normalization がそれぞれの時期の学習にどのような影響を与えるのかを分析することは Batch Normalization がなぜうまくいくのかを理解する上で重要であり、本稿で導出したダイナミクスを解析することでそのような分析が可能となることが期待できる。

参考文献

- [Arora 18] Arora, S., Li, Z., and Lyu, K.: Theoretical Analysis of Auto Rate-Tuning by Batch Normalization, *arXiv preprint arXiv:1812.03981* (2018)
- [Arpit 18] Arpit, D., Jastrzebski, S., Ballas, N., Krueger, D., Bengio, E., Kanwal, M. S., Maharaaj, T., Fischer, A., Courville, A., Bengio, Y., and Lacoste-Julien, S.: A Closer Look at Memorization in Deep Networks, *ICML* (2018)
- [Biehl 95] Biehl, M. and Schwarze, H.: Learning by on-line gradient descent, *Journal of Physics A: Mathematical and General*, Vol. 28, No. 3, p. 643 (1995)
- [Bjorck 18] Bjorck, J., Gomes, G., Selman, B., and Weinberger, K. Q.: Understanding Batch Normalization, *NeurIPS 2018* (2018)
- [Ioffe 15] Ioffe, S. and Szegedy, C.: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, in *ICML*, pp. 448–456 (2015)
- [Kohler 18] Kohler, J., Daneshmand, H., Lucchi, A., Zhou, M., Neymeyr, K., and Hofmann, T.: Exponential convergence rates for Batch Normalization: The power of length-direction decoupling in non-convex optimization, *arXiv preprint arXiv:1805.10694* (2018)
- [Krueger 17] Krueger, D., Ballas, N., Jastrzebski, S., Arpit, D., Kanwal, M. S., Maharaaj, T., Bengio, E., Fischer, A., and Courville, A.: Deep Nets Don't Learn Via Memorization, *ICLR Workshop* (2017)
- [Luo 18] Luo, P., Wang, X., Shao, W., and Peng, Z.: Towards Understanding Regularization in Batch Normalization, *arXiv preprint arXiv:1809.00846* (2018)
- [Rahaman 18] Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F. A., Bengio, Y., and Courville, A.: On The Spectral Bias of Neural Networks, *NeurIPS Workshop* (2018)
- [Riegler 95] Riegler, P. and Biehl, M.: On-line backpropagation in two-layered neural networks, *Journal of Physics A*, Vol. 28, pp. L507–L513 (1995)
- [Saad 95] Saad, D. and Solla, S. A.: Exact Solution for On-Line Learning in Multilayer Neural Networks, *Physical Review Letters*, Vol. 74, No. 41, pp. 4337–4340 (1995)
- [Santurkar 18] Santurkar, S., Tsipras, D., Ilyas, A., and Mardy, A.: How Does Batch Normalization Help Optimization?, *arXiv preprint arXiv:1805.11604* (2018)
- [Saxe 18] Saxe, A. M., McClelland, J. L., and Ganguli, S.: A mathematical theory of semantic development in deep neural networks, *arXiv preprint arXiv:1810.1053* (2018)
- [Schwarze 93] Schwarze, H.: Learning a rule in a multilayer neural network, *Journal of Physics A*, Vol. 26, pp. 5781–5794 (1993)
- [Seung 92] Seung, H. S., Sompolinsky, H., and Tishby, N.: Statistical mechanics of learning from examples, *Physical Review A*, Vol. 45, No. 8, pp. 6056–6091 (1992)
- [Xu 18] Xu, Z.-Q. J.: Understanding training and generalization in deep learning by Fourier analysis, *arXiv preprint arXiv:1808.04295* (2018)
- [Yoshida 18] Yoshida, Y., Karakida, R., Okada, M., and Amari, S.: Statistical Mechanical Analysis of Learning Dynamics of Two-Layer Perceptron with Multiple Output Units, *J. Phys. A (provisionally accepted)* (2018)

ランダムフォレストにおけるノード数と木数の関係

On the trade-off between the number of nodes and the number of trees in Random Forest

熊野 颯^{*1}

So Kumano

阿久津 達也^{*2}

Tatsuya Akutsu

^{*1}京都大学大学院情報学研究科

Graduate School of Informatics, Kyoto University

^{*2}京都大学化学研究所

Institute for Chemical Research, Kyoto University

Expressibility of machine learning models has been extensively studied. For example, in a Neural Network, it is proved that the efficiency concerning the number of nodes is generated from the depth. On the other hand, it is not clear whether the efficiency exists in Random Forest. Therefore, in this research, we investigate whether the efficiency exists in Random Forest. We first show that Random Forest does not have the same kind of efficiency as Neural Network, and next we show that the efficiency concerning the number of nodes can be generated from the number of trees.

1. はじめに

封筒上の郵便番号の識別から店舗ごとの商品の需要予測、医療診断に到るまで幅広い分野において機械学習は活用されている。いずれの応用においても、まず学習のモデル (例えば、ニューラルネットワークなら層の数やニューロンの数、ランダムフォレストであれば木の本数や葉の数) を決定し、次に損失関数を最小化するパラメータを発見する。この学習モデルの決定、すなわち学習によって最適な関数の探索を行う関数族の決定は非常に重要である。関数族が不必要に大きい場合は過剰な計算コストや過学習の問題を引き起こし、逆に不十分な場合には、適切な関数を得ることはできない。課題ごとに適切な関数族を選択しなければならないのである。

こうした関数族の選択において有効な手がかりを得るべく、学習モデルの表現能力に関する研究が古くから行われてきた。例えば、深さ 2 のニューラルネットワークで任意のボレル可測関数を近似することができることを主張する universal approximation theorem が知られている。また、近年ではニューラルネットワークは層の数によって表現能力が大きく変化することが分かってきた。例えば、深さ 2 のニューラルネットワーク (SNN) は任意のボレル可測関数を近似することが可能であるが、必要となるノード数が膨大となることがある。しかし、深いネットワーク (DNN) を用いることで、SNN よりも少ないノード数で同じ関数を表現できる場合がある。こうした DNN が持つ、深さによって効率良く関数を表現することが可能となる性質は depth efficiency と呼ばれている。

ニューラルネットワークに対しては、表現能力の研究が盛んにおこなわれている一方で、ランダムフォレストにおいては表現能力に関する研究は少ない。そこで、本研究ではランダムフォレストがニューラルネットワークにおける depth efficiency と同様にノード数に関する効率性を持つかについて考察する。

2. 関連研究

ニューラルネットワークの表現能力は、その深さに対して指数的に増大すると考えられている。Montúfar らは区分線形関数を活性化関数として用いた際に、ネットワークが表現する区

分線形関数の線形領域がネットワークの幅に対しては多項式のオーダーでしか増加しないのに対し、深さに対しては指数的に増大することを示した [1]。また、Raghu らは、ネットワークの出力の軌道の長さの観点から、深さに対して表現能力が指数的に増大することを示した [2]。

DNN は SNN よりも効率的に関数を表すことが出来ることを実際に示した研究も存在する。Telgarsky は、SNN では指数オーダーのノードが必要だが、DNN では線形オーダーのノードで表現できる関数族の存在を示した [3]。Symanskiy らは、DNN において depth efficiency が生じる要因の一つとして入力の変動性に関与していることを明らかにした [4]。Bengio らは、Sum-product Network に関して DNN では SNN よりもノード数の観点で効率的となる関数が存在することを示した [5]。

ランダムフォレストの表現能力に関する研究においては、Mansour が入力の次元が d 、ノード数 N の二分決定木の VC 次元の下界が $\Omega(N)$ 、上界が $O(N \log d)$ であることを示した [6]。また、VC 次元が d の識別器を T 個使用するアンサンブル学習の VC 次元の上界は $O(dT \log(dT))$ である [7]。Oshiro らは、データの密度を定義し密度と最適な木の本数の関係を実験により確かめた [8]。これらの研究から、ランダムフォレストは木の本数に対して、表現能力が劇的に増加するといったことはないと思われる。しかし、これらの結果は depth efficiency と類似の性質の存在を否定する訳ではない。そこで本研究では、ランダムフォレストにおけるノード数に関する効率性について考察する。

3. ランダムフォレストの深さ

Bengio らは、ブースト木のニューラルネットワークにおける深さは 3 であると述べている。[9]。この結果から、ランダムフォレストは深さ 3 のニューラルネットワーク以下の表現能力しか持たず、depth efficiency と同種の効率性は存在しないと考えられる。

しかし、この結果は 選言標準形のアナロジーとして述べられたものである。一方、近年の DNN の depth efficiency に関する結果の多くは区分線形関数を活性化関数として用いたニューラルネットワークについて示されている。そこで、本研究ではまず、ランダムフォレストは区分線形関数を活性化関数として用いた深さ 3 のニューラルネットワーク以下の表現能力

連絡先: 熊野 颯, 京都大学大学院情報学研究科,
kumano@kuicr.kyoto-u.ac.jp

しか持たないことを示す。ただし、ニューラルネットワークの出力ノードは線形関数とする。以後、本研究ではランダムフォレストを構成する決定木として、各ノードごとに入力空間にある変数に対して垂直二分割（他の変数に対して軸並行に分割）する二分決定木を考え、各決定木の出力はラベルのみであるとする。つまり、ランダムフォレストが出力するラベルはそのラベルを出力する木の木数が最も多いラベルとなる。したがって2値分類問題におけるランダムフォレストはすべて奇数本の決定木から構成される。また、 c でクラス全体を表し、 H を

$$H(x) = \begin{cases} 1 & (x \geq 0) \\ 0 & (x < 0) \end{cases} \quad (1)$$

と定義する。また、ニューラルネットワークは出力ノードの値が最も大きいクラスに入力を分類するものとする。

補題 1. n ノードの決定木 $DT: R^d \rightarrow c$ は H を活性化関数に用いた $O(n)$ ノード、深さ3のNNで表現することができる。

Proof. NNの深さ1のノードとしてDTの各エッジが成立する場合に1、それ以外は0を出力するノードを作成する。すなわち、DTのエッジが $f_1 \geq a$ ならば $H(f_1 - a)$ を作成し、 $f_1 > a$ ならば $1 - H(a - x)$ を作成する。(1はこのノードの出力を受け取るノードのバイアスとして与えられる)

次に深さ2のノードとしてDTの各葉ノードを表現するノードを作成する。各ノードはDTにおける葉ノードから根までの枝に対応するノードの和から、DTにおける葉ノードの深さ-0.5を引いた値を活性化関数の入力として受け取る。最後に、深さ3の出力ノードとして各クラスに対応するノードを作成し、各クラスに属する葉ノードと対応するノードの出力を入力として受け取る。このNNがDTを表現することは明らかである。実際、DTにおいて入力が到達する葉ノードに対応する深さ2のノードのみが1を出力し、他の深さ2のノードは0を出力する（少なくとも1つの深さ1のノードの出力は0であるため）

□

例えば、下図のニューラルネットワークは左図の決定木を模倣する。

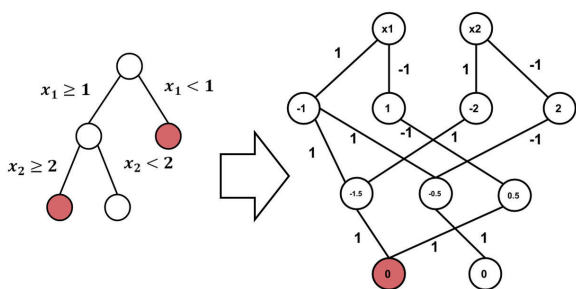


図 1: エッジの数字は重み、ノードの数字はバイアスを表す。例えば、深さ1の左端のノードは $H(x_1 - 1)$ を表す。

この補題を用いて、多くの区分線形関数に対し、ランダムフォレストはその関数を活性化関数として用いた深さ3のニューラルネットワーク以下の表現能力しか持たないことが示せる。以降、 $S \subset R^d$ は有限集合であるとする。

定理 2. n ノードの決定木 $DT: S \rightarrow c$ は、非有界な領域における傾きが異なる区分線形関数 $g: R \rightarrow R$ を活性化関数として用いた $O(n)$ ノードの深さ3のNNで表現することができる。

Proof. S は有限集合であるから、任意の S の要素が DT の識別境界上に存在しないと仮定してよい。 f は、 $a \neq b$ を満たすある定数 a, b に対して、 $\lim_{x \rightarrow \infty} f(x) = a$ 、 $\lim_{x \rightarrow -\infty} f(x) = b$ を満たすとする。この時、 f を用いて ($x \neq 0$ において) H を表現することができる。実際、

$$H(x) = \lim_{\epsilon \rightarrow 0} (f(\frac{x}{\epsilon}) - b) \frac{1}{a-b} = \frac{1}{a-b} \lim_{\epsilon \rightarrow 0} f(\frac{x}{\epsilon}) - \frac{b}{a-b} \quad (2)$$

である。さらに、 g を非有界な領域において $ax+b, a'x+b'$ と表される区分線形関数であるとする。この時、 $g(x+1) - g(x)$ は非有界な領域においてそれぞれ、 $a(x+1)+b-ax-b=a, a'$ となるため、 g を用いて、 H を活性化関数として用いたNNを表現することができる（図2）。したがって、補題1から非有界な領域における傾きが異なる区分線形関数を活性化関数として用いた $O(n)$ ノード、深さ3のNNでDTを表現することができる。 □

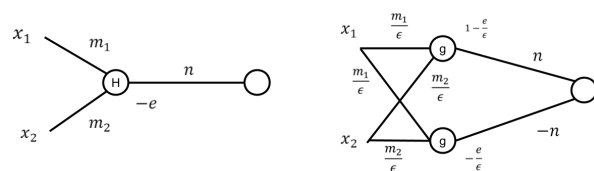


図 2: g として ReLU 関数を用いた場合、活性化関数として H を用いたネットワーク (左図) は右図のネットワークで表現される

定理 3. n ノードのランダムフォレスト $RF: S \rightarrow c$ は非有界な領域における傾きが異なる区分線形関数を活性化関数として用いた $O(n)$ ノードの深さ3のNNで表現することができる。

Proof. ランダムフォレストを構成するそれぞれの木に対して、対応するNNを作成する。深さ2の各ノードはそれぞれ対応する決定木の葉ノードに入力が到達したときに1、それ以外の場合には0を出力する。したがって、これらのNNの出力ノードを同一のノードとして見做すことで、RFに対応するNNを得ることができる。 □

これらの結果から、ランダムフォレストは深さ3のニューラルネットワーク以下の表現能力しか持たないこと、すなわち depth efficiency と全く同種のノードに関する効率性は存在しないことが分かる。

4. 木の本数と表現能力

4.1 ノード数の下界

前節から、ランダムフォレストにおいてはニューラルネットワークに対応する深さは固定であること、すなわち DNN と同様の効率性は存在しないことが分かった。次に、木の本数によって表現能力がどう変化するかの検証を行う。その為に、 n 本 $O(n)$ ノードのランダムフォレストを T 本の木からなるランダムフォレストで表現する際に必要となるノード数を求める。

入力全体を X で表す。 $x \in X$ とランダムフォレストを構成する木の各葉ノード $leaf$ に対して、 x のラベルを $c(x)$ 、 $leaf$ に割り当てられているラベルを $c(leaf)$ 、 x が $leaf$ に到達するか否かを $leaf(x)$ で表す。

補題 4. 次の性質を満たす集合 X を考える。この時、 $M = |X|$ とおくと X を T 本の木からなるランダムフォレストで表現するには $\Omega(M^{\frac{2}{T+1}})$ のノードが必要である。

X を識別する任意の T 本の木からなるランダムフォレストに対し、下記が成立する。

$$\forall x_1, x_2 \in X, c(x_1) = c(x_2) = c \rightarrow |\{leaf | c(leaf) = c, leaf(x_1), leaf(x_2)\}| < \frac{T+1}{2} \quad (3)$$

Proof. 上記の性質は任意の同一ラベルを持つ異なる 2 点は、 X を認識する任意のランダムフォレストにおける葉ノードのうち半数以上に同時に正しく認識されることはないということである。 $L = \{leaf | c(leaf) = 1\}$ 、 $l = |L|$ とする。つまり、 l はラベル 1 が割り当てられている葉の数である。同様に L' でラベル 0 が割り当てられている葉全体、 l' でラベル 0 が割り当てられている葉の数を表す。また、 M_1 で $|\{x \in X | c(x) = 1\}|$ 、 M_0 で $|\{x \in X | c(x) = 0\}|$ を表す。すなわち、 M_i はクラス i のデータの個数を表す。

$c(x) = 1$ を満たす x は少なくとも $\frac{T+1}{2}$ 以上の L の要素に対して $leaf(x)$ が成立する必要がある。一方、(3) より $leaf(x)$ が成立する L の要素のうち、どの $\frac{T+1}{2}$ 個の葉の組み合わせも他の $c(x) = 1$ を満たす x に対して同時に $leaf(x)$ を満たすことはない。したがって、各 $x \in M_1$ に対し他の $x' \in M_1$ では同時に 1 とならない $\frac{T+1}{2}$ 個の L に属する葉の組み合わせが存在するから、

$$\binom{l}{\frac{T+1}{2}} \geq M_1 \quad (4)$$

が成立する。また、ラベル 0 の場合も同様にして

$$\binom{l'}{\frac{T+1}{2}} \geq M_0 \quad (5)$$

が成立する。したがって、

$$l^{\frac{T+1}{2}} \geq M_1 \quad (6)$$

$$l'^{\frac{T+1}{2}} \geq M_0 \quad (7)$$

が成立する。したがって、このランダムフォレストの葉ノードの数は

$$l + l' \geq M_1^{\frac{2}{T+1}} + M_0^{\frac{2}{T+1}} \geq M^{\frac{2}{T+1}} \quad (8)$$

となる。したがって、必要なノード数の下界は

$$\Omega(M^{\frac{2}{T+1}}) \quad (9)$$

□

この補題を用いて、 n 本 $O(n)$ ノードのランダムフォレストを T 本の木からなるランダムフォレストで表現する際に必要なノード数の下界を求めることができる。下界は次のようになる。

定理 5. n 本 $O(n)$ ノードのランダムフォレスト $RF : \{0,1\}^n \rightarrow \{0,1\}$ を T 本の木からなるランダムフォレストで表現する際に必要なノード数の下界は $\Omega((\frac{2^n}{\sqrt{n}})^{\frac{2}{T+1}})$ である。

Proof. 任意の T ($T < n$) に対し、補題 4 の (3) の性質を満たし、 n 本 $O(n)$ ノードのランダムフォレストで識別できる X が存在することを示せばよい。各次元の要素の和が $\frac{n+1}{2}$ (ラベル 1 を持つ) もしくは $\frac{n+1}{2} - 1$ (ラベル 0 を持つ) となる n 次元の 0,1 ベクトル全体を X と置く。

X が任意の T ($T < n$) に対し、(3) の性質を満たすことを示す。 X を識別する T ($T < n$) 本の木からなるランダムフォレストが存在するとする。 X が (3) の性質を満たさないと仮定する。このとき、 $c(x_1) = c(x_2) = c$ を満たす、ある x_1, x_2 が存在して、これらに対し $leaf(x_1), leaf(x_2), c(leaf) = c$ を満たす $leaf$ が $\frac{T+1}{2}$ 個以上存在する。以降では一般性を失うことなく $c = 1$ と置く。

一方、 $x_1 \neq x_2$ であるから x_1 で 1、 x_2 で 0 となる要素が少なくとも 1 つは存在し、この要素は x_1 と x_2 の両方が同じ葉ノードに到達する木の出力に影響を与えない。したがって、 x_1 においてこの要素を反転させた x_1' はこのランダムフォレストにおいてラベル 1 と識別される。しかし、 x_1' は要素の和が $\frac{n+1}{2} - 1$ であるから、 X に属し、ラベル 0 を持つはずである。これは、 X が T 本の木からなるランダムフォレストによって識別されるということに矛盾する。したがって、 X は (3) の性質を満たす。

X が n 本 $O(n)$ ノードのランダムフォレストによって識別されるということは容易に示すことができる。それぞれの木において 1 つの特徴量を評価し、1 ならばラベル 1、0 ならばラベル 0 を出力すれば良い。また、

$$|X| = 2^{\binom{n}{\frac{n+1}{2}}} \geq \frac{2^n}{\sqrt{n}} \quad (10)$$

であるから、必要なノード数の下界は

$$\Omega((\frac{2^n}{\sqrt{n}})^{\frac{2}{T+1}}) \quad (11)$$

□

この定理から同じ関数を表現する際に、木の本数が n 本であるランダムフォレストと比較して T 本の木から構成されるランダムフォレストでは多量のノードを必要とすることがあること、すなわちランダムフォレストは木の本数に対してノード数に関する効率性を持つことが示された。

5. 結論と展望

本研究では、ランダムフォレストにおいて DNN の depth efficiency のようなノード数に関する効率性が存在するかということについて考察を行った。まず、多くの区分線形関数に対し、ランダムフォレストはその関数を活性化関数として用いた 3 層のニューラルネットワーク以下の表現能力しか持たないことを示した。このことから、ランダムフォレストにおいてはニューラルネットワークと全く同種の効率性は存在しないことが示された。次に、 n 本の木から構成されるランダムフォレストを T 本の木から構成されるランダムフォレストで表現する際には、 T が n よりも十分に小さい場合には多くのノードが必要となることを示した。このことから、木の本数に関してはノード数に関する効率性が存在することが示された。今後の課

題としては、定理 5 の下界の改善、 n 本の木から構成されるランダムフォレストを T 本の木から構成されるランダムフォレストで表現する際に必要となるノード数の上界の導出などが考えられる。

参考文献

- [1] Guido Montufar, F., et al. "On the number of linear regions of deep neural networks." Advances in neural information processing systems. 2014. p. 2924-2932.
- [2] Maithra Raghu, et al. "On the expressive power of deep neural networks." arXiv preprint arXiv:1606.05336 (2016).
- [3] Matus Telgarsky. "Representation benefits of deep feedforward networks." arXiv preprint arXiv:1509.08101 (2015).
- [4] Lech Szymanski, and Brendan McCane. "Deep networks are effective encoders of periodicity." IEEE Transactions on Neural Networks and Learning Systems 25.10 (2014): 1816-1827.
- [5] Olivier Delalleau, and Yoshua Bengio. "Shallow vs. deep sum-product networks." Advances in Neural Information Processing Systems. 2011. p. 666-674
- [6] Yishay Mansour. "Pessimistic decision tree pruning based on tree size." In Press of Proc. 14th International Conference on Machine Learning. 1977. p.195-201.
- [7] Shalev-Shwartz, Shai, and Shai Ben-David. Understanding machine learning: From theory to algorithms. Cambridge university press, 2014. p139
- [8] Mayumi Thais Oshiro, Pedro Santoro Perez, and José Augusto Baranauskas. "How many trees in a random forest?." International Workshop on Machine Learning and Data Mining in Pattern Recognition. Springer, Berlin, Heidelberg, 2012. p. 154-168
- [9] Yoshua Bengio, Olivier Delalleau, and Clarence Simard. "Decision trees do not generalize to new variations." Computational Intelligence 26.4 (2010): 449-467.

CTR 予測モデルの評価に AUC や log-loss は適切か?

Do the AUC and log-loss evaluate CTR prediction models properly?

片桐 智志 *1

Satoshi Katagiri

*1株式会社ファンコミュニケーションズ

F@N Communications, Inc.

Click-through rate (CTR) prediction is one of the most important task for web advertising platform companies. However, CTR prediction is a non-standard machine learning task, so conventional metrics, for example, area under the Receiver Operating Characteristic curve (AUC), and log-loss, a.k.a. cross-entropy, and so on, can be improper. Our target is develop a new metrics for CTR prediction. In this article, we state the drawbacks of such conventional metrics and perspective of a metric based on the calibration plot approach.

1. はじめに

広告のリアルタイム入札システム (RTB) は、消費者が広告枠のあるウェブページなどを閲覧するたびに、広告のリクエストがなされ、どの広告主がページの広告枠に出稿する権利についてオークションを自動で行い、出稿する広告を決定するシステムである。ここでのオークションは多くの場合、2 番目に高い額を提示した入札者 (広告主) が落札するという二位価格オークション (second-price auction) を採用している。理論上、二位価格オークションでは入札者の私的価値 (private value) と一致する価格で入札する “truth-telling 戦略” が支配戦略である [Krishna 10]。クリックに対して課金される料金体系の場合、広告オークションにおける私的価値とは、1 回の広告表示 (インプレッション) に対して消費者がどれくらいの確率でクリックするかである [田頭 13]。そのため、この確率を正確に見積もることは、RTB が顧客にとって有益であることに直結する。RTB プラットフォームを持つ多くの企業では、適切に入札のプライシングができるように、広告表示に対する click-through 率 (CTR) を機械学習によって予測する方法を研究または導入しており、その先行研究だけでも枚挙に暇がない。

CTR 予測を機械学習の問題として見ると 2 値分類問題とみなせるため、多くの研究では予測モデルの評価に area under the Receiving Operator Characteristic curve (AUC) や、対数損失 (交差エントロピー) が用いられている。しかし、標準的な機械学習の問題とは異なり、求められているのは予測値が正解ラベルにどれだけの的中しているかというよりも、広告リクエスト単位の予測確率がどれだけ適切であるか、という点である。第 2 節で詳細に述べる先行研究により、従来モデルの評価によく用いられてきた AUC や対数損失だけでは適切に評価できないことがわかっている。本研究では、これらを踏まえ、カリブレーションの指標として従来から提案されている [DeGroot 83] のカリブレーションプロットや [Caruana 04] の CAL とその問題点についても考察する。

2. 先行研究のサーベイ

Microsoft の研究チームによれば [Yi 13], AUC や対数損失などと比較して NE の性質について言及しており *1, AUC あるいは対数損失 (または正規化エントロピー, 以下 NE) だけでは評価指標として完全ではないとしつつも, CTR 予測精度の評価問題に適した方法についての結論を述べていない。[He 14] では, AUC と NE を利用しているが, NE だけではデータ全体でみたクリック率と予測値のクリック率が必ずしも近似できていないとして, NE に加えてデータ全体のクリック率と予測値から計算できる期待クリック率の一致, という指標も重視している。

[Gail 05, Cook 07] では, 分類モデルの出力する予測確率を将来の病気の発病リスクとみなした場合について言及がある。疫学分野では, ラベルに分類される確率を正しく予測すること, 正例に予測される場合のモデルの条件分布とそうでない場合の条件分布の差別化, の 3 種類が要求される場合のいずれもありえるため, [Gail 05] ではそれぞれ, *accuracy*, *calibration* (カリブレーション), *discrimination*, と定義している。医療を例にすると, 現時点で病気が疑われる患者を診断することは, 予測確率よりも陽性と陰性をどれだけはっきり区別できるかが重要な *discrimination* のタスクであり, 一方でまだ発症していない人が将来発症する可能性や, 予後の死亡率などを知りたい (いわゆる *prognostic studies*) 場合は *calibration* のタスクとなる。従来使われている AUC は *discrimination* に対応し, 対数損失や NE は *accuracy* に対応する。しかし, CTR 予測について重要となるのは, カリブレーションである。

2.1 AUC の問題点

Microsoft の研究チームによれば [Yi 13], AUC は予測確率の大きさそのものを見ないことが問題であるとしている。AUC は予測確率の絶対値ではなく, 大きさをソートした際の順序を評価していることが問題であり, 実際にデータ全体のクリック割合とクリックの予測頻度が一致しないようなモデルであっても AUC が大きくなることもある。よって, 予測確率が異なる値でも AUC は変化しない。単純な例として, $y \in \{0, 1\}$ のラベルに対して 3 種類の予測モデルが, それぞれ予測確率 $\hat{\pi}_A$, $\hat{\pi}_B$, $\hat{\pi}_C$ を表 1 のように出力しているとする。このとき, $\hat{\pi}_B$ は, $\hat{\pi}_A$

連絡先: 片桐智志, 株式会社ファンコミュニケーションズ サービス開発部情報科学技術研究所, s_katagiri@fancs.com

*1 正確には, 相対情報ゲイン (RIG) についての議論だが, $RIG = 1 - NE$ という関係が成り立つため議論内容は NE に容易に転用できる。

y	0	0	0	1	1	1
$\hat{\pi}_A$	0.1	0.2	0.5	0.5	0.6	0.8
$\hat{\pi}_B$	0.2	0.3	0.6	0.6	0.7	0.9
$\hat{\pi}_C$	1	2	5	5	6	8

表 1 AUC が変わらない例

の各値に 0.1 を足したもので、 $\hat{\pi}_C$ は $\hat{\pi}_A$ を 10 倍にしたものだが、それぞれの AUC は全く同じになる。

AUC はそもそも discrimination を評価する指標であり、疾病リスクモデルの変数選択を例に、尤度比統計量やカイ二乗統計量と AUC とで反応の大きさの違いが指摘されている [Cook 07]*2。加えて、データの分布しだいで、AUC の事実上の最大値が変化するという問題も、具体例を示して指摘されている [Diamond 92, Gail 05]。

2.2 対数損失の問題点

対数損失 (交差エントロピー) も広く使われている指標であるが、カリブレーションを評価するには問題がある。

たとえば、真の確率 π とラベル y のペアについて、 $\pi > 0.5$ ならば $y = 1$ 、そうでなければ $y = 0$ となる場合を考える。このとき、

$$(\pi_1, y_1) = (0.4, 0),$$

$$(\pi_2, y_2) = (0.6, 1)$$

という 2 点だけのデータあるとする。このとき、カリブレーションの観点からすれば、予測モデルは真の確率に近い値を出力するのが望ましいため、 $\hat{\pi}_1 = 0.4, \hat{\pi}_2 = 0.6$ を出力するようなモデルが最も望ましい。このとき、対数損失は約 0.51 となる。一方で、 $\hat{\pi}_1 \rightarrow 0, \hat{\pi}_2 \rightarrow 1$ のときに対数損失は明らかにこれより小さくなる。例えば $\hat{\pi}_1 = 0.1, \hat{\pi}_2 = 0.9$ のときに約 0.11 となる。よって、対数損失の小さなモデルほどカリブレーションも良いとは限らない。

正規化エントロピー (NE) は、対数損失をデータの正例割合に基づく対数損失で割った指標であるため、CTR 予測のような不均衡データに対して対数損失よりも優れているとされる [Yi 13, He 14]。しかしながら、対数損失を正規化しただけの指標であるため、カリブレーションを評価できないという問題は NE に対してもそのまま当てはまる。

[Brier 50] による、回帰問題で使われる平均二乗誤差 (MSE) を分類問題にそのまま適用した Brier スコアもまた、同様の問題がある。

2.3 カリブレーションプロットと CAL

一方で、疫学分野では、カリブレーションの確認方法として、Hosmer-Lemeshow 検定が提案されている [Hosmer 89, Hosmer 80]。これはデータをいくつかのグループに分割し、それぞれでカイ二乗統計量を計算した和でカリブレーションがなされているかを検定する方法である。今回我々が求めているのは、AUC や対数損失に変わる相対的な指標であるので採用できないが、[Caruana 04] では、よく似たアイデアとして、カリブレーションプロット [DeGroot 83] の結果に対して平均絶対誤差 (MAE) を計算する CAL を紹介している。

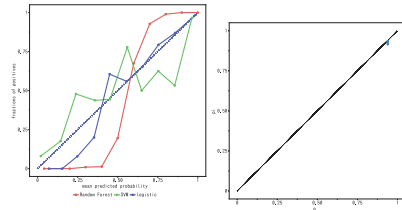


図 1 典型的なカリブレーション・プロットの例 (左) と一律同じ出力をする予測モデルのカリブレーションプロットの例 (右)

3. 適切な指標はどうあるべきか

AUC や NE の問題点に対して [He 14] は、カリブレーションの要件を満たすように、データのクリック頻度と予測確率に基づく期待値とが一致しているかについても確認する方法を提案している。しかしこれは、膨大なデータ全体でのクリック数と予測数の一致を見ているだけであり、1 つ 1 つの広告リクエストに対する予測 CTR の精度を保証するものではない。本研究では、良いカリブレーションの定義として、[Gail 05] で述べられているものを採用する。予測モデルが完全にカリブレーションされている (perfectly calibrated) とは、特徴量 x に対して予測確率を出力する予測関数 $\hat{f}(x)$ が、任意の x に対して π の条件付き期待値に等しい、つまり以下を満たすことを言う。

$$\hat{f}(x) := E[\pi | x] = \int \pi dG(\pi | x) \quad (1)$$

ここで、 $G(\pi | x)$ は π の条件確率密度関数である。モデルが良くカリブレーションされているかは、実際の確率とモデルの出力する予測確率の誤差がどれだけ小さいかで判断する。しかし、例えば仮に平均二乗誤差 (MSE) $N^{-1} \sum_{i=1}^N (\hat{f}(x_i) - \pi_i)^2$ で評価すると、真の確率 π_i は観測できない。ここで π_i を観測可能なラベル y_i に置き換えると、先述の対数損失や Brier スコアの問題が発生する。この点、カリブレーションプロットや CAL は、この問題に対して、データを分割したサブグループ内での頻度を真の確率の近似として使用していると解釈できる。

4. 考察

しかしながら、CAL には次のように少なくとも 2 点の問題が考えられる。(1) 観測点数がサブグループごとに異なるため、相対的に点数の多いサブグループの当てはまりが過小評価される傾向にある、(2) 出力されるすべての予測確率が同じ値である場合、分位数によるサブグループによる分割ができない。

(1), (2) いずれも、分位数ではなく、観測点数が同等になるように等分割するという方法が考えられる。しかし、(2) の場合は正例の多い不均衡データにおいて一律で大きな予測確率を出力するモデルに対しては図 1 のようなカリブレーションプロットを描き、CAL が良い値を示す可能性があり、NE で解消された問題が再び浮上する。

5. 結論と課題

本研究では、予測確率の精度を求める CTR 予測において、AUC や 対数損失, NE だけでは評価に不十分であるということを示した。続いてカリブレーションプロットに基づく CAL について考察し、カリブレーションの指標としては問題点が残ること示した。CAL をより適切な指標へと改善することは今後の課題であるが、サブグループの分割方法が重要になると予想

*2 [Cook 07] では AUC を c-統計量と呼んでいる。C は concordance の略である。

できる。我々は CAL をに対して取り組んでいるが、最終的な目的は適切な評価指標の考案だけでなく、カリブレーションの良いモデルの改良方法を開発することである。

参考文献

- [Brier 50] Brier, G. W.: VERIFICATION OF FORECASTS EXPRESSED IN TERMS OF PROBABILITY, *Monthly Weather Review*, Vol. 78, No. 1, pp. 1–3 (1950)
- [Caruana 04] Caruana, R. and Niculescu-Mizil, A.: Data Mining in Metric Space: An Empirical Analysis of Supervised Learning Performance Criteria, in *Proceedings of the 2004 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '04*, pp. 69–78, Seattle, WA, USA (2004), ACM Press
- [Cook 07] Cook, N. R.: Use and Misuse of the Receiver Operating Characteristic Curve in Risk Prediction, *Circulation*, Vol. 115, No. 7, pp. 928–935 (2007)
- [DeGroot 83] DeGroot, M. H. and Fienberg, S. E.: The Comparison and Evaluation of Forecasters, *The Statistician*, Vol. 32, No. 1/2, pp. 12–22 (1983)
- [Diamond 92] Diamond, G. A.: What Price Perfection? Calibration and Discrimination of Clinical Prediction Models, *Journal of Clinical Epidemiology*, Vol. 45, No. 1, pp. 85–89 (1992)
- [Gail 05] Gail, M. H. and Pfeiffer, R. M.: On Criteria for Evaluating Models of Absolute Risk, *Biostatistics*, Vol. 6, No. 2, pp. 227–239 (2005)
- [He 14] He, X., Bowers, S., Candela, J. Q. n., Pan, J., Jin, O., Xu, T., Liu, B., Xu, T., Shi, Y., Atallah, A., and Herbrich, R.: Practical Lessons from Predicting Clicks on Ads at Facebook, in *Proceedings of 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining - ADKDD'14*, pp. 1–9, New York, NY, USA (2014), ACM Press
- [Hosmer 80] Hosmer, D. W. and Lemeshow, S.: Goodness of Fit Tests for the Multiple Logistic Regression Model, *Communications in Statistics - Theory and Methods*, Vol. 9, No. 10, pp. 1043–1069 (1980)
- [Hosmer 89] Hosmer, D. W. and Lemeshow, S.: *Applied Logistic Regression*, Wiley Series in Probability and Mathematical Statistics Applied Probability and Statistics, Wiley, New York (1989), OCLC: 19514573
- [Krishna 10] Krishna, V.: *Auction Theory*, Elsevier, Academic Press, Amsterdam, 2. ed edition (2010), OCLC: 845563467
- [Yi 13] Yi, J., Chen, Y., Li, J., Sett, S., and Yan, T. W.: Predictive Model Performance: Offline and Online Evaluations, in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '13*, p. 1294, New York, New York, USA (2013), ACM Press
- [田頭 13] 田頭 幸浩, 山本 浩司, 小野 真吾, 塚本 浩司, 田島 玲: オンライン広告における CTR 予測モデルの素性評価, 第 5 回データ工学と情報マネジメントに関するフォーラム (DEIM2013), 郡山市, 福島県 (2013)

補遺: 各指標の定義

AUC

ROC 曲線の下側の面積である。

平均自乗誤差 (MSE) ・ Brier スコア:

平均自乗誤差は、予測値と真値の差の 2 乗平均で、(2) のように定義される。2 値分類に限定すれば、MSE と [Brier 50] による Brier スコアが同一のものであるのは明らかである。

$$\text{Brier} := \text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{\pi}_i)^2 \quad (2)$$

対数損失 (交差エントロピー) ・ 正規化エントロピー:

対数損失は、(3) で定義される。

$$\text{logloss} := - \frac{1}{N} \sum_{i=1}^N [y_i \ln \hat{\pi}_i + (1 - y_i) \ln(1 - \hat{\pi}_i)] \quad (3)$$

正規化エントロピー (NE) は、対数損失を、データ全体の正例の割合に対する対数損失で除したものであり、(4) のように定義される。ラベル数が極端に不均衡である場合、簡単に低い対数損失を算出できる問題があるが、NE はデータの割合で調整することでこの問題を解消している [He 14]。

$$\begin{aligned} \text{NE} &:= \frac{\text{logloss}}{- (\bar{p} \ln \bar{p} + (1 - \bar{p}) \ln(1 - \bar{p}))}, \\ \bar{p} &:= \frac{1}{N} \sum_{i=1}^N y_i \end{aligned} \quad (4)$$

CAL:

予測値を十分位数を区切りに B 個に分割する。 k 番目のビンに属する集合が b_k , $\#b_k$ はその要素数で、 \bar{p}_k はそこに属する正例ラベルの頻度とすると、CAL は、(5) のように定義される [Caruana 04]。

$$\text{CAL} := \frac{1}{N} \sum_{k=1}^B \left| \bar{p}_k - \frac{1}{\#b_k} \sum_{i \in b_k} \hat{\pi}_i \right| \quad (5)$$

大局基準値共有による社会的強化学習

Social reinforcement learning with shared global aspiration for satisficing

其田憲明 *1

Noriaki Sonota

神谷匠 *2

Takumi Kamiya

高橋達二 *1

Tatsuji Takahashi

*1 東京電機大学理工学部

School of Science and Engineering, Tokyo Denki University

東京電機大学大学院

Graduate School of Tokyo Denki University

When humans learn, it is not just by individual trial-and-error, but the learning is accelerated by sharing information with others. There are social learning strategies such as imitating others' actions and emulating the high achievement of someone. As a model of social learning, sharing of state- and/or action-values are often implemented in reinforcement learning algorithms. However, sharing information of such huge amount is not realistic for a model of social learning of humans or animals. We propose an algorithm in which a mere "record" (achieved accumulated reward per episode) leads to efficient social learning. The algorithm is based on the model of satisficing integrated with different risk attitudes around the reference (aspiration level), and the conversion of the global aspiration onto each state.

1. はじめに

機械学習の分野の一つである強化学習では、学習を行うエージェントが環境との相互作用によって得られた経験から行動価値を更新することで、収益を最大化する最適な行動系列を学習することを目的とする。

一方で人間の学習は、ある目的水準を満たすことを目的とした場合に満足化原理 [Simon 56] と呼ばれる意思決定における損失回避の傾向がある。満足化原理とは現状の収益が基準を満たさない場合には探索を行い、基準を満たす行動を発見した場合にはその行動を選び続けることである。満足化原理により、人間は効率の良い探索を行うことができると考えられている。

この満足化原理を強化学習に応用したのが Risk-sensitive Satisficing (RS) である [高橋 16]。RS は最適な基準値を与えることで素早く最適な行動を学習し、後悔の値を有限に抑えたと証明されている [Tamatsukuri 18]。

また、"keeping up with the Joneses" という慣用句が存在するように、人間には自身を他者と社会的比較を行うことによって満足化の参照点が推移することが知られており、今日ではインターネットの興隆などにより社会的比較の対象ははるか広範囲に達している [Manktelow 15]。

強化学習における他者との情報共有は群強化学習 [飯間 06] のように行動価値に関連したものが多い。しかし行動価値の共有には状態行動対で情報を共有する必要があるため計算量が多いこと、共有される情報次第では共有されたエージェントの探索傾向に偏りが生じることで準最適解に陥る可能性があることが考えられる。このような問題に対して満足化による強化学習を複数のエージェントで行い、状態ごとに他者のより良い成績を自身の基準値として共有しつつ学習を行う手法が有効であることが示されている [其田 18]。しかし、各状態ごとに行動価値を基準値として共有していたが、現実には各状態ごとに基準となる成績を知ることは容易ではない。一方で、100m 走のタイムのような大局的な成績を知ることはあり、大局的な情報であっても人間はより効果的に活用することができる [柄谷 85]。

本論文では大局的な成績から基準値の共有を行う社会的学習を検証し、その有効性を示すことを目的とする。

2. 強化学習と RS 価値関数

2.1 強化学習

強化学習とは学習を行うエージェントが環境との相互作用によって、得られる報酬を最大化する行動系列の獲得を目標とする機械学習の分野の一つである。エージェントの行動決定手法を方策と呼び、行動価値の推定手法を価値関数と呼ぶ。強化学習の代表的な価値推定手法である Q-learning では、時間 t における状態を s_t 、エージェントの方策に基づいて得られる行動を a_t としたとき、行動に対する環境からの作用として報酬 r_t 、次状態 s_{t+1} を観測する。行動価値 $Q(s_t, a_t)$ は学習率 α 、割引率 γ を用いることで式 1 によって更新される。

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left(r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t) \right) \quad (1)$$

エージェントは行動価値 Q を利用して行動を決定する。

2.2 Risk-sensitive Satisficing

Risk-sensitive Satisficing (RS) は、状態行動対 (s_t, a_t) に対する試行量 $\tau(s_t, a_t)$ と行動価値 $Q(s_t, a_t)$ 、そして状態 s_t に対して保持される満足化基準値 $N(s_t)$ から、式 2 によって RS 価値関数が定義される。

$$RS(s_t, a_t) = \tau(s_t, a_t)(Q(s_t, a_t) - N(s_t)) \quad (2)$$

RS 方策は RS 価値関数を最大化する行動 a_t を選択する方策である。また、試行量 $\tau(s_i, a_i)$ は $\tau_{\text{curr}}(s_t, a_t)$ と $\tau_{\text{port}}(s_t, a_t)$ を用いて式 3 によって定義される。そして、 $\tau_{\text{curr}}(s_t, a_t)$ と $\tau_{\text{port}}(s_t, a_t)$ は試行量割引率 γ_τ 、試行量学習率 α_τ を用いて式 4 と式 5 によって更新される。

$$\tau(s_t, a_t) = \tau_{\text{curr}}(s_t, a_t) + \tau_{\text{port}}(s_t, a_t) \quad (3)$$

$$\tau_{\text{curr}}(s_t, a_t) \leftarrow \tau_{\text{curr}}(s_t, a_t) + 1 \quad (4)$$

$$\begin{aligned} \tau_{\text{port}}(s_t, a_t) &\leftarrow \tau_{\text{port}}(s_t, a_t) \\ &+ \alpha_\tau \left(\gamma_\tau \tau(s_{t+1}, a_{t+1}) - \tau_{\text{port}}(s_t, a_t) \right) \end{aligned} \quad (5)$$

連絡先: 高橋達二, 東京電機大学理工学部, 350-0394
埼玉県比企郡鳩山町大字石坂, 049-296-1642,
tatsujit@mail.dendai.ac.jp

基準値 N に加え、試行量 τ を用いることによって、基準を満たしていない非満足状態においては楽観的探索を、基準を満たしている満足状態においては悲観的活用を行う。

2.3 Global Reference Conversion

強化学習に拡張された RS は各状態に基準値 $N(s_i)$ を持ち、各状態の行動価値 $Q(s_i)$ に対して適切な基準値 $N(s_i)$ を与えることで、適切に学習できることが示されている [牛田 17]。しかし、エージェントはタスク全体としての大局的な基準値を知ることが出来たとしても、全体目標を達成するための局所的な基準値は不明であることが多い。よって Global Reference Conversion (GRC) を用いることで、タスク全体の大局基準値 N_G から、式 6 によって局所的な基準値 $N(s_i)$ 変換を行う。

$$\begin{aligned}\delta_G &= \min(E_G - N_G, 0) \\ \max_a Q(s_i, a) - N(s_i) &= \zeta(s_i) \delta_G \\ N(s_i) &= \max_a Q(s_i, a) - \zeta(s_i) \delta_G\end{aligned}\quad (6)$$

式中の ζ はスケールパラメータである。 E_G は大局観測期待値と呼ばれるものであり、エージェントが一定期間内に環境から得られた累計報酬 E_{tmp} と N_G を用いて式 7 で更新される。

$$E_G \leftarrow \frac{E_{tmp} + \gamma_G (N_G E_G)}{1 + \gamma_G N_G} \quad (7)$$

$$N_G \leftarrow 1 + \gamma_G N_G \quad (8)$$

パラメータ γ_G は大局割引率を表し、 $0.0 \leq \gamma_G \leq 1.0$ の範囲で定められる。

3. 満足化基準値共有による社会的学習

[其田 18] では同一設定のタスクを複数用意し、エージェントを 1 体ずつ配置して並列的に学習した。エージェント N 体からなるグループの n 番目のエージェントの状態 s_i における最大行動価値を $Q_n^{best}(s_i)$ とした時、式 9 によってグループ内で自律的に基準値 $N(s_i)$ を更新した。その結果、行動価値を直接共有するエージェントは準最適解に陥ったが、基準値として共有するエージェントは準最適解に陥らずに学習することに成功した。

$$N(s_i) \leftarrow \max_n Q_n^{best}(s_i), (\forall s) \quad (9)$$

しかし、この手法では状態 s_i ごとに計算するため、状態数に比例して情報共有に必要とする計算量が増加する問題が挙げられる。そこで、本研究ではエージェント n 体の観測した大局観測期待値 E_G^n から式 10 のように大局基準値 N_G を定める。

$$N_G \leftarrow \max_n E_G^n \quad (10)$$

この手法による、より少ない情報共有での学習の有用性を次の SuboptimaWorld タスクで評価した。

4. SuboptimaWorld

このタスクでは準最適解となるゴールが多数存在しており、エージェントは準最適解となるゴールを避けて、最適解となる報酬を得られるゴールへの経路を学習することを目標とする。

4.1 シミュレーション設定

図 1 のように縦 9 マス、横 9 マスの全 81 状態からなる格子空間上で報酬が得られる経路を学習する。報酬が得られるゴールが 8 つ存在し、ゴールで得られる報酬は図 1 中のゴールの数字に対応してそれぞれ 1, 2, ..., 8 と得られる。またゴールを終端状態とし、スタートからゴールにたどり着くまでを 1 エピソードとして 4000 エピソード行った。

提案手法である大局基準値共有を行うエージェント群を GRC グループとし、比較対象として、先行研究である各状態の基準値 $N(s_i)$ を共有する RS グループ、最適基準を事前情報として保持している GRC_{opt} 、そして強化学習における一般的な方策である ϵ -greedy を用いた。

全ての手法において学習率 $\alpha = 0.1$, $\gamma = 0.9$ と設定した。提案手法である GRC グループは大局基準値 N_G の初期値を一律 $N_G = 0$ とし、 GRC_{opt} の大局基準値 $N_G = 8$ とした。そして GRC グループ、 GRC_{opt} はそれぞれ $\zeta(s_i)$ を一律 1 に、 $\alpha_\tau = 0.1$, $\gamma_\tau = 0.9$, $\gamma_G = 0.9$ し、エピソード単位の獲得報酬を E_{tmp} とした。また、RS グループでの τ_α と τ_γ は GRC グループと同様に設定した。そして、GRC グループと RS グループの基準値を共有するタイミングはどちらもグループに属する全てのエージェントが 1 エピソード終了した時点とし、基準値を共有した後にエージェントは次のエピソードに移る。 ϵ -greedy は $\epsilon = 1.0$ から等速度で減少させ、2000 エピソード時点で $\epsilon = 0$ となるように設定した。

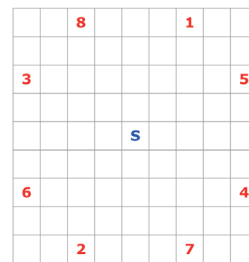


図 1: SuboptimaWorld 概要図

4.2 結果

提案手法である GRC エージェントを 4 体とした場合の 1000 回行った平均の結果を図 2 に示す。図 2 から、最高報酬にたどり着くのが早い順から、 GRC_{opt} 、GRC グループ、RS グループ、 ϵ -greedy となっており、 ϵ -greedy は 2000 エピソード経過時点ではわずかに最高報酬を下回る成績であることがわかる。

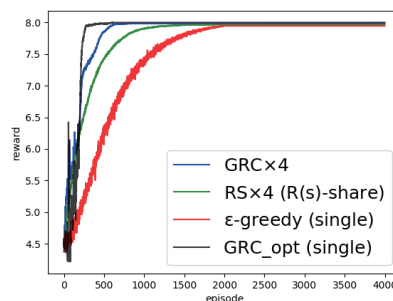


図 2: 獲得平均報酬の時間発展

そして、GRC エージェントを 1, 2, 3, 4 体とした 1000 回の平均の結果を図 3 に示す。図 3 からエージェントが増えるごとに成績が良くなることがわかる。

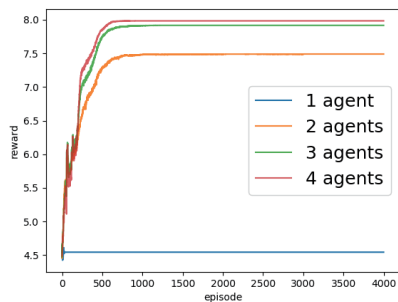
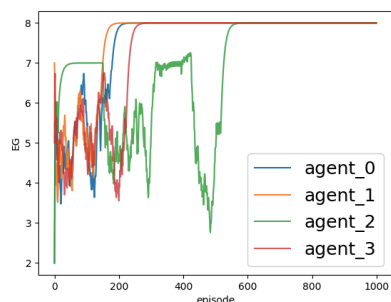


図 3: 獲得平均報酬の時間発展

そして、グループで具体的にどのように学習しているかを見るための1例として、4体グループで1000エピソードを1回のみ行った場合の大局観測期待値 E_G を出力したグラフを図4に示す。

図 4: 大局観測期待値 E_G の時間発展

報酬7で満足していたエージェントが他者がより高い報酬を獲得し始めた時点から再探索を行い、そして報酬8を得ることに成功しているのがわかる。

5. 考察

まずはじめに、人数が増えるごとに GRC グループの成績が良くなることについて、非満足状態のエージェントは共有された大局基準値 N_G 以上の報酬を得られる行動を学習するという性質から、中には共有された大局基準値 N_G と等しい成績で満足する場合が存在することが考えられる。しかし、グループに属するエージェントが多いほど一度に観測される報酬のパターンが多くなることで、素早くより良い成績が得られるエージェントが出現する。その結果、グループの大局基準値 N_G が素早く上昇し、グループのエージェント数が多いほど共有された大局基準値 N_G と等しい成績に留まらず、より良い成績を発見することができると考えられる。

次に、基準値の共有手法について、RS グループは行動価値から基準値を設定しているため、価値更新手法の Q-learning の行動価値の更新式の学習率の影響から満足状態のエージェントの行動価値が収束するまでの間は、非満足状態のエージェントの行動価値は満足状態のエージェントから共有される基準値以上の行動価値を満足状態のエージェントより少ない試行で獲得するのが困難である。一方で GRC グループでは他者の行動価値を直接使わずに、エピソードで得られる報酬から計算される大局観測期待値 E_G から基準値を更新しているため、最適な行動の行動価値が他の行動価値よりも高くなっていけば最適な行動をすぐ取ることができる。したがって RS グループのように行動価値の収束を待つ必要があるために、基準を満たす行動系

列への収束が遅くなるということは起こりにくくなっている。

しかし、準最適解で満足したエージェントのような、とある行動系列の行動価値が他の行動価値よりも高くなってしまった場合には、探索によって他者から共有された基準値を満たす報酬が得られたとしても、すぐに基準を満たす行動をすることをせずに、行動価値が高い今までの準最適解への行動を選びつつ、徐々に基準を満たす行動へと移行していることが図4から推測できる。これは GRC が自身の最大行動価値から基準値を設定していることと、RS グループと同様に Q-learning の行動価値の更新式の学習率によって徐々に更新されることから生じるものであると考えられる。したがって、価値更新手法の変更による、より素早い価値推定を行うことで改善が可能であると考えられる。

6. おわりに

大局基準値共有を用いた社会的学習により、より限られた情報共有で学習を有効に行うことに成功した。そして、グループのエージェント数が増えることで、一定期間で探索する領域が広がり、より良い成績を素早く発見、共有することで全体での成績が向上することが判明した。今回提案した手法ではエピソード単位での獲得報酬を利用した情報を共有したため、他のアルゴリズムともグループを作ることが可能であることが考えられる。そして、今回は1エピソードごとに情報を共有していたところを、情報を共有する間隔をより疎にすることによって、さらに学習中の情報共有に必要とする計算量を減少することが可能であると考えられる。

よって今後の課題として、GRC エージェントが他のアルゴリズムとグループを作った場合の挙動の変化の観測と、より少ない人数での成績の向上、より疎な間隔での情報共有による学習可能性の検証、そして今回のタスクではエージェントが確実に報酬を得られる設計であったため、より複雑なタスクでの適用手法を考案することが考えられる。

参考文献

- [Manktelow 15] Ken Manktelow : Thinking and reasoning (2012) (邦訳: 思考と推論, 服部雅史, 山祐嗣 訳: 思考と推論 理性・判断・意思決定の心理学, 北大路書房 (2015), pp. 260-261
- [Simon 56] Simon, H.A.: Rational choice and the structure of the environment, *Psychological Review*, 63(2), 129–138. (1956)
- [Tamatsukuri 18] Akihiro Tamatsukuri, Tatsuji Takahashi: Guaranteed satisficing and finite regret: Analysis of a cognitive satisficing value function. *arXiv preprint arXiv:1812.05795*, 2018.
- [飯間 06] 飯間 等 & 黒江 康明: エージェント間の情報交換に基づく群強化学習法, 計測自動制御学会論文集, 42(11), 1224–1251. (2006)
- [牛田 17] 牛田有哉, 甲野佑, 高橋達二: 生存を目的とする満足化強化学習, JSAI 2017, 4C2-2in2. (2017)
- [其田 18] 其田憲明, 神谷匠, 甲野佑, 高橋達二: 満足化基準値共有を用いた社会的強化学習, JSAI2018 予稿集, 1N1-05. (2018)
- [高橋 16] 高橋達二, 甲野佑, 浦上大輔: 認知的満足化 - 限定合理性の強化学習における効用, 人工知能学会論文誌, 31(6), 1–11. (2016)
- [柄谷 85] 柄谷 行人: プタに生れかわる話, 批評とポスト・モダン, pp. 257260. (1985)

画像の“写真らしさ”に関する数学的アプローチについて

On a mathematical approach to “photo-likeness” of images

浅尾 泰彦^{*1} 坂本 龍太郎^{*1}

Yasuhiko Asao

Ryotaro Sakamoto

^{*1}東京大学大学院数理科学研究科

Graduate School of Mathematical Science, the University of Tokyo

In image recognition, it is significant to determine the boundary between meaningful and non-meaning images. In this paper, we show a mathematical approach to this problem by defining a “quasi-photographic” image. In order to formulate the question ‘What is photograph likeness?’ mathematically, we introduce a function ‘depth’ that takes real values for images and analyze its asymptotic behavior. We also examine that an actual photograph is indeed a quasi-photograph. The idea of depth comes from the rank of the 0th persistent homology of a cubical complex and it can be expected that more precise classification of images can be obtained by analyzing the higher rank in the future. We also believe that it can be applied to deep learning, which is being actively utilized recently in image recognition, to selection of learning data. We would like to propose one approach of applying pure mathematics in image recognition.

1. はじめに

本稿で扱う問題は、コンピュータが“内在的”にどこまで意味のある画像とそうでないものを区別できるか？というものである。

以下では意味のある画像のことを「写真」と呼ぶことにする。つまり絵や数字など、我々が目にするとすぐに意味が理解できる、もしくは意味があると判断できるものを「写真」と総称する。一方で「意味のない」画像というのは一見して意味が理解できない、意味があると判断できないものを指す。例えば図1右のようなモザイク画像のことを指している。

機械学習ではコンピュータに性質 A を持つ大量の類似データを学習させることで、新たに読み込ませたデータが性質 A を持つかどうかを判断させることができた。例えばりんごの写真を学習させた後に図1右の画像を読み込ませると、コンピュータはそれがりんごでないと判断できる。我々がここで「内在的」と言っているのは、そのような学習の過程を経ないで、ということである。つまりりんごが何であるかを知らない状況で、コンピュータはりんごの写真とモザイクの画像をきちんと分類することが可能であるか？ また可能であれば「りんごとモザイク」という極端な分類の他にどの程度分類が可能であるか？

本稿では、画像の持つ数理的な性質によって特徴付けられる「準写真」という画像のクラスを導入することで、この問題に取り組んだ。準写真であるという性質は個々の画像に対して数学的に有無が判別できるため上で述べた意味で内在的であり、従ってコンピュータは学習の過程を経ずに画像を準写真とそうでないものに分類することができる。

さらに数学的に定義された準写真であるという性質は、実際の写真にもきちんと備わっていることを例で確かめることができた。

準写真は画像の「深さ」という数学的概念を定義することで得られ、深さは近年データサイエンスの分野で広く認知されているパーシステントホモロジーから着想を得ている。本稿にお



図1: 右は depth が非常に大きい。

いてパーシステントホモロジーなど純粋数学で成熟した道具を、画像認識に活用する1つのアプローチを提案したい。

2. 画像の深さ

2.1 画像の定式化

$[0, 1]$ で 0 以上 1 未満の実数の集合を表す。自然数 N に対して集合 \square_N を

$$\square_N := \left\{ \frac{0}{2^N}, \frac{1}{2^N}, \dots, \frac{2^N - 1}{2^N} \right\}^2 \subseteq [0, 1) \times [0, 1).$$

で定義する。 $C = \{0, 1, \dots, n-1\}$ を白黒の濃淡を表す集合とする。 $2^N \times 2^N$ ピクセルのモノクロ画像は、写像 $\square_N \rightarrow C$ そのものである。ここで $2^N \times 2^N$ ピクセル画像を図2のように $2^d \times 2^d$ 分割することを考える。すなわち \square_N の部分集合族 \square_N^d を

$$\square_N^d := \left\{ \frac{1}{2^d} \square_{N-d} + x \subseteq \square_N \mid x \in \square_d \right\}.$$

で定め非交叉和による $2^d \times 2^d$ 分割 $\square_N = \bigsqcup_{\square \in \square_N^d} \square$ を与える。

例えば、 $\square_N^0 = \{\square_N\}$, $\square_N^N = \{\{x\} \mid x \in \square_N\}$ である。

連絡先: 浅尾泰彦: asao@ms.u-tokyo.ac.jp,

坂本龍太郎: sakamoto@ms.u-tokyo.ac.jp

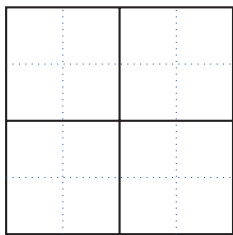


図 2: 実線は \square_1^1 , 点線は \square_2^1 .



図 3: $N = 10$ のキリンの画像. 左上から右下にかけて $0 \leq d \leq 10$ が大きくなっていく. $d = 5$ で初めて黒いマスが現れるので $\text{depth} = 0.5$ である.

2.2 深さの定義

2色モノクロ画像 $f: \square_N \rightarrow C = \{0, 1\}$ に対して, どれくらい色の偏りがあるかを測る指標として次の関数 φ_d を定義する.

$$\varphi_d(f) := \min_{\square \in \square_N^d} \left(\sum_{i,j \in \square} |f(i) - f(j)| \right).$$

つまり φ_d は $2^d \times 2^d$ 分割したそれぞれのマスのうち, もっとも色が偏ったマスの偏り具合を数値として表す関数である. 1 つでも「全て白 (黒)」であるマスが存在すればその値は 0 となる. 特に $\varphi_N = 0$ が常に成り立つ.

(2色と限らない) 画像 f の深さを, どれくらい分割を細かくした時に (d を N に近づけたときに) 初めて $\varphi_d = 0$ となるか, を測ることで定義する.

定義 1 $2^N \times 2^N$ ピクセル n 色画像 $f: \square_N \rightarrow C$ に対して

$$\text{depth}(f) := \frac{1}{N} \min \left\{ d \in \{0, \dots, N\} \mid 0 = \min_{p: \text{Im} f \rightarrow \{0,1\}} \varphi_d(p \circ f) \right\},$$

を f の深さという. ここで p は f の像から集合 $\{0, 1\}$ への全射全体を動く.

例えば $\text{depth}(f)$ が 0 であることと, f が定値写像であること (つまり 1 色画像) は同値である. 逆に図 1 右のような複雑な画像の depth は 1 となる. つまり depth は画像の複雑さを測る指標となっている. 意味のある画像はある程度色に偏りがあると考えられるため depth は低いことが期待される. 図 3 は d を次第に大きくした時の $\left(\sum_{i,j \in \square} |f(i) - f(j)| \right)_{\square \in \square_N^d}$ を色の濃さとして表したものである. d が増加するとマスは細かくなり, 初めて真っ黒なマスが現れた時の d が depth に対応する.

3. 深さの漸近挙動と準写真の定義

この章では前章で定義した depth を用いて, どれくらいの割合の画像が「とても複雑」(つまり写真らしくない) かを計算する. それに基づいて準写真という数学的に「写真らしい」画像のクラスを定義する.

$\text{Map}(\square_N, C)$ で \square_N から C への写像全体 (つまり $2^N \times 2^N$ ピクセルモノクロ画像全体) の集合を表す. $0 \leq \alpha \leq 1$ に対して depth が α 以下の画像全体の集合を $\mathcal{P}_{N,C}(\alpha)$ とおく. 有限集合 X の要素の個数を $\#X$ で表すことにすれば, $\#\text{Map}(\square_N, C) = n^{4^N}$ より, depth が α 以下の画像全体の割合は $\frac{\#\mathcal{P}_{N,C}(\alpha)}{n^{4^N}}$ である. このとき次の計算結果を得る.

命題 1

$$\lim_{N \rightarrow +\infty} \frac{\#\mathcal{P}_{N,C}(\alpha)}{n^{4^N}} = \begin{cases} 0 & \text{if } \alpha < 1, \\ 1 & \text{if } \alpha = 1. \end{cases}$$

すなわち「ほとんど全ての画像がとても複雑」であることがわかる. この結果は, ランダムに生成した画像が写真であることは稀有である, という我々の認識に関する直観と矛盾しない.

上の結果を $\#\mathcal{P}_{N,C}(\alpha)$ の主要項の漸近挙動だと考え, 第 2 主要項を計算すると以下の結果を得る.

命題 2

$$\lim_{N \rightarrow +\infty} \frac{\#\mathcal{P}_{N,C}(1 - \alpha \frac{\log N}{N})}{n^{4^N}} = \begin{cases} 0 & \text{if } 1/\log 4 < \alpha, \\ 1 & \text{if } 0 \leq \alpha \leq 1/\log 4. \end{cases}$$

ただし \log の底は 10 とする.

従って画像全体のうち depth が $1 - \frac{\log N}{N \log 4}$ 未満であるような「複雑すぎないもの」は非常に少ないことがわかる. 前章でも述べたように意味のある画像は depth が低いと期待されるため, これらを準写真と定義する.

定義 2 $2^N \times 2^N$ ピクセル n 色画像 $f: \square_N \rightarrow C$ が

$$\text{depth}(f) < 1 - \frac{\log N}{N \log 4}$$

を満たすとき, f は準写真であるという.

$1 - \frac{\log 10}{10 \log 4} = 0.93979400086 \dots$ より, 図 3 のキリンの画像 ($\text{depth} = \frac{5}{10} = 0.5$) は準写真である.

4. 人間の認識に関する予想

図 3 では左上の真っ白な画像から始まって次第に細度が上がっていきキリンの姿が浮かび上がり, 最後には真っ黒な画像になる. 真っ黒になる 1 つ手前の画像はもうほぼ写真と変わらないが, もう 1 つ手前までいくと写真と認識できるもののやや画質が悪いという印象を持つ. これら 11 枚の画像に真っ白なものから順番に $\frac{0}{10}, \frac{1}{10}, \dots, \frac{10}{10}$ と数を振ると, $\frac{8}{10}$ と $\frac{9}{10}$ の間がちょうど画質の良し悪しを判断する境目ということになる. 我々の予想はこの境目の値がおおよそ $1 - \frac{\log N}{N \log 4}$ に対応するのではないかというものである. 実際, 前章でみたように $1 - \frac{\log 10}{10 \log 4} \sim \frac{9}{10}$ である.

予想 1 画像 $f_N = f : \square_N \rightarrow C$ に対して画像 $f_{N-1} : \square_{N-1} \rightarrow C$ を

$$f_{N-1}(x) = \left\lfloor \frac{1}{4} \sum_{\square \in \frac{1}{2^{N-1}} \square_1 + x} f(\square) \right\rfloor \quad (x \in \square_{N-1})$$

で定義する。帰納的に $f_{N-i} := (f_{N-i+1})_{N-i}$ と定義する。画像の列 $f_N, f_{N-1}, f_{N-2}, \dots, f_0$ は次第に画質が荒くなっていくが、画質の良し悪しの変化を認識する境目は $\frac{k}{N} \leq 1 - \frac{\log N}{N \log 4} \leq \frac{k+1}{N}$ を満たすような f_k と f_{k+1} の間である。

5. パーシステントホモロジーによる高次化

パーシステントホモロジーの一般論については [H] が詳しい。depth と 0 次パーシステントホモロジーとの関連を見るために、画像 $f : \square_N \rightarrow C$ に対して次のようなフィルター付き方体複体 $C_d(f)$ を考える。頂点集合は \square_N であり、2つの相異なる頂点 a, b はそれらがある $\square \in \square_N^d$ に隣り合って含まれていてかつ $f(a) = f(b)$ であるときに 1 方体で結ばれているとする。2 方体についても同様に定義する。このとき複体 $C_d(f)$ の 0 次ホモロジーの階数は d について広義単調増加関数であるが、あるところから指数的に増加する。その変化の点が depth と対応する。本研究では 0 次パーシステントホモロジーしか考えていないが、同様にして高次の階数から画像の内在的な情報を取り出せると期待できる。

6. 画像認識への応用の展望

機械学習・深層学習において、例えばコンピュータに犬と猫の写真进行分类させようとする、必要な学習データはそれぞれの写真 10000 枚程度とされている。10000 枚のデータを人の手で収集し、それをコンピュータに読み込ませることはかなりのコストを費やすため近年では収集・読み込みの自動化が試みられている。一方で我々の depth を用いた方法は画像の内在的な情報を数理的に引き出すことで画像の分類をしているため学習データを用意する必要がない。未だ精度が荒く実用化への障害はあるものの、パーシステントホモロジーを始めとする様々な数学を用いた画像認識への新たなアプローチとして期待できると考えている。

7. 謝辞

本研究は数物フロンティア・リーディング大学院のプログラムの一つである社会数理実践研究として行われたものである。画像認識についての解説などで尽力してくださった株式会社ニコン研究開発本部数理科学研究所の皆様、特に深層学習との関連や論文に対する貴重なコメントを下された高山侑也さん、中村ちからさんに心から感謝申し上げます。またセミナーの時間調整や全般に関わるコメントをして頂いた東京大学数理科学研究科特任助教（当時）の土岡俊介さんにも御礼申し上げます。

参考文献

[H] 平岡裕章:「タンパク質構造とトポロジー —パーシステントホモロジー群入門—」共立出版, 2013.

[3K4-J-2] Machine learning: real world interaction

Chair:Daiki Kimura Reviewer:Hikaru Kajino

Thu. Jun 6, 2019 3:50 PM - 5:30 PM Room K (201A Medium meeting room)

[3K4-J-2-01] Consideration on Generation of Saliency Maps in Each Action of Deep Reinforcement Learning Agent

○Kazuki Nagamine¹, Satoshi Endo², Koji Yamada², Naruaki Toma², Yuhei Akamine² (1. Information Engineering Course, Graduate School of Engineering and Science, University of the Ryukyus, 2. Faculty of Engineering, School of Engineering Computer Science and Intelligent Systems, University of the Ryukyus)

3:50 PM - 4:10 PM

[3K4-J-2-02] Dynamic Reward Clustering

○Ryota Higa¹, Junya Kato¹ (1. NEC Corporation)

4:10 PM - 4:30 PM

[3K4-J-2-03] A dialogue system implemented with latent parameters

○Weida Li¹, Chie Hieida², Takayuki Nagai² (1. Seiko Gakuin High School, 2. The University of Electro-Communications)

4:30 PM - 4:50 PM

[3K4-J-2-04] Comfortable Driving by Deep Inverse Reinforcement Learning

○Daiko Kishikawa¹, Sachiyo Arai¹ (1. Chiba University)

4:50 PM - 5:10 PM

[3K4-J-2-05] Linear function approximation of Cognitive Satisficing Function

○Yu Kono^{1,2} (1. Tokyo Denki University, 2. DeNA, Co., Ltd.)

5:10 PM - 5:30 PM

深層強化学習エージェントの行動別顕著性マップの生成に関する考察

Consideration on Generation of Saliency Maps in Each Action of Deep Reinforcement Learning Agent

長嶺一輝^{*1}
Kazuki Nagamine

遠藤聡志^{*2}
Satoshi Endo

山田孝治^{*2}
Koji Yamada

當間愛晃^{*2}
Naruaki Toma

赤嶺有平^{*2}
Yuhei Akamine

^{*1}琉球大学理工学研究科情報工学専攻

Information Engineering Course, Graduate School of Engineering and Science, University of the Ryukyus

^{*2}琉球大学工学部工学科知能情報コース

Faculty of Engineering, School of Engineering Computer Science and Intelligent Systems, University of the Ryukyus

In recent years, deep reinforcement learning agents have surprisingly developed and achieved great results. the methods of analyzing the behaviour of agents by visualizing neural networks have been proposed. However, the methods to obtain saliency maps for each action has not been much researched. In this paper, we propose the method of generating saliency maps for each action of the agents in order to obtain deeper insight when analyzing a neural network in a deep reinforcement learning agent by visualization. We applied the proposed method to the agent which learned Atari 2600 Pong. As a result of the experiment, we obtained saliency maps which visualizes the influence of environment on each action of the agents.

1. はじめに

近年、深層学習の発展に伴い、強化学習に深層学習を取り入れた深層強化学習も目覚ましい成長を見せている。一方で、深層学習ではブラックボックス的な性質があるため、その解消が課題となっており、深層強化学習にも同様の問題がある。例えば、エージェントの行動のみを視認して、根拠となった画像特徴を推測するのは困難である。これにアプローチする手法としてニューラルネットの入出力を用いて判断根拠を可視化する手法が提案されている [Selvaraju 17]。このような手法は深層強化学習においてはエージェントの行動根拠の視覚化に用いられている。顕著性マップの生成は可視化手法の一つであり、エージェントが注視しているオブジェクトや学習における戦略の変化等の分析に活用され始めている。しかし、エージェントが出力する行動価値セットに対して一つの顕著性マップを得る手法は提案されているが、出力を行動ごとに分けて可視化する手法は十分に研究されていない。行動別に可視化できれば、分析においてより深い洞察が期待できる。そこで、本研究では既存の可視化手法を拡張して、深層強化学習エージェントの行動ごとに顕著性マップを得る手法を提案する。また、提案した手法を Atari 2600 の Pong を学習した深層強化学習エージェントに適用した結果を示し、得られた行動ごとの顕著性マップについて考察する。

2. 先行研究

Greydanus らは深層強化学習エージェントの行動等进行分析するために、顕著性マップを用いた可視化手法を提案している [Greydanus 17]。この提案手法と深層強化学習アルゴリズムである Asynchronous Advantage Actor-Critic (A3C) 及び Atari 2600 ゲーム環境を用いて、エージェントが注視している部位を可視化し、学習過程における戦略の変化等について分析している。また、ゲームや機械学習に精通していない非エキスパートでも、可視化結果を見ることでエージェントの行動の

解釈が容易になることを示している。Greydanus らは提案手法を顕著性マップを生成する摂動ベースな方法と呼び、次のようにマップを求めている。はじめに、エージェントを十分に学習させた後、学習を停止した状態で環境からの観測やエージェントが出力した行動価値等を数ステップ分保存する。次に、保存した観測状態と次の (1) 式を用いてマスク画像を作成する

$$\Phi(I_t, i, j) = I_t \odot (1 - M(i, j)) + A(I_t, \sigma_A) \odot M(i, j) \quad (1)$$

ここで、 Φ はマスク画像、 I_t はステップ t における観測状態、 i, j はその二次元座標、 \odot はアダマール積、 M はマスク、 A はガウスフィルタでぼかした観測状態を表す。マスク画像は、観測状態の二次元座標 i, j をマスク M の範囲でぼかしたもので、エージェントが行動根拠とする画像特徴量を認識不可にする効果を持つ。これと (2) 式の顕著性メトリックを用いて、画像部位 i, j における顕著性スコアを計算する。顕著性スコアとは、その部位がエージェントの行動価値に及ぼす影響を数値化したものである。

$$S(t, i, j) = \frac{1}{2} \|\pi_u(I_{1:t}) - \pi_u(I'_{1:t})\|^2$$

$$\text{where } I'_{1:k} = \begin{cases} \Phi(I_k, i, j) & \text{if } k = t \\ I_k & \text{otherwise} \end{cases} \quad (2)$$

ここで、 S は座標 i, j におけるスコア、 π_u は行動価値ベクトル、 $I_{1:t}$ は観測状態の系列を表す。このスコアを縦横数ピクセル間隔で求め、顕著性マップを生成する。求めたマップはリサイズし、前処理なしの観測状態の RGB チャンネルのいずれかと加算することで可視化する。上記の (1)(2) 式は A3C のアクターに対する顕著性スコアの計算式で、クリティックについても状態価値を用いて同様に計算することができる。青がアクター、赤がクリティックの顕著性マップとして、Pong 環境下で当手法を用いて可視化した結果を図 1 に示す。図 1 より、Pong をプレイする上で重要と考えられるブロックやパドルといったオブジェクトをエージェントが注視していることが直感的にわかる。

当手法がアクターの行動根拠とするオブジェクトをハイライトしている一方で、どのオブジェクトがいずれの行動価値に貢

連絡先: 長嶺一輝, 琉球大学理工学研究科, 〒 903-0213 沖縄県中頭郡西原町千原 1, k188583@ie.u-ryukyu.ac.jp



図 1: Greydanus らの提案手法による可視化画像

献しているかといった観察を行うことは難しい。そこで、本研究では (2) 式の顕著性メトリックを変更して、行動別に顕著性マップを得る手法を提案する。

3. 提案手法

本研究では、各行動ごとの顕著性マップを得るために、先行研究の (2) 式の $S(t, i, j)$ を次の (3) 式のように変更した。

$$S(a, t, i, j) = \pi_a(I_{1:t}) - \pi_a(I'_{1:t}) \quad (3)$$

ここで、 a は行動価値ベクトルのインデックス、 π_a は行動価値ベクトル内の a に対応する一つの行動価値を表す。この変更により、行動別の可視化だけでなく、エージェントが注視している部位が行動価値に対して正負どちらの貢献をしているか可視化することも狙った。観測状態に対する行動価値と、マスク後の行動価値の変化及びマスク部位の行動価値に対する貢献の関係を表 1 に示す。

表 1: 行動価値とその変化とマスク部位の貢献の関係

行動価値	行動価値の変化	部位の貢献
正	増加	負
正	減少	正
負	増加	負
負	減少	正

これは、例えば、観測状態に対する行動価値が正で、マスク後の行動価値が増加した場合、その部位が行動価値に対して負の影響を与ると解釈できることを表す。この表 1 による解釈と (3) 式を用いることでエージェントの行動別顕著性マップを生成する手法を提案する。

4. 実験

本実験では提案手法により各行動ごとの顕著性マップを得ることを目的とする。また、得られた可視化結果と図 1 とを比較する。

実験設定として、本稿では学習環境に Pong を用いる。Pong には Up, Down を含む六つの行動がある。また、結果を解釈し易くするためにオブジェクトの軌跡等の表示を追加した。

実験結果を示す。図 2 は、実験結果の動画から抜き出したフレームで、対戦相手が打ったボールをエージェントがパドルを操作して打ち返そうとしている場面である。図中の茶色と白色の部分 Pong の画像に顕著性マップを合成したもので、赤の部位が行動価値に正の貢献を、青の部位が負の貢献をしてる。

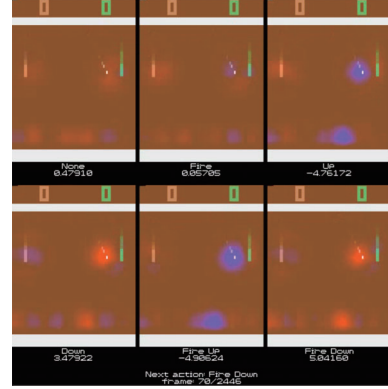


図 2: 本研究の提案手法による可視化画像

それぞれの画像の下には行動ラベルと行動価値を、図中下部にはエージェントが取る行動とステップ数を表示している。各画像の行動価値を見ると、Fire Down が最も高く、次に Down が高いことがわかる。反対に、Fire Up が最も低く、Up が次に低いことがわかる。None と Fire はその中間程度の価値である。各画像では、None と Fire は特徴的な注視は生じておらず、残りの四つに顕著な特徴を持った注視が生じている。一方で、この四つは注視部位がおおよそ同じで、Fire と Down でオブジェクトの貢献が対照的になっていることがわかる。

考察を述べる。ボールの軌跡は、ボールがパドルより下に向かっており、パドルもそれに追従するように下に向かっていている。ボールはパドルから距離があり、パドルの x 座標に到達するまで時間がかかることから、パドルを下に操作する Down, Fire Down の行動価値が高いと考えられる。また、その際の顕著性マップを見ると、ボールが正の貢献をしていることがわかる。反対に、Up, Fire Up においてはボールが負の貢献をしている。これは、各オブジェクトがこの位置関係の時は、ボールが与える貢献が行動価値に大きく影響しているためだと考えられる。オブジェクトと各行動の関係についての洞察を得られることが、提案手法の先行研究に対するアドバンテージと考える。

5. まとめ

本研究では、深層強化学習エージェントの行動根拠を視覚的に分析するために、摂動ベースの顕著性マップ生成を行う既存手法を基に、行動別の可視化を行う手法を提案した。実験では、A3C に提案手法を適用し、Pong 環境下でのアクターの行動別の顕著性マップを生成した。その結果、行動別に特徴的な注視部位が生じていることを確認した。また、結果からオブジェクトが行動に与える影響といった、先行研究の手法では得ることが難しい洞察を得られた。

参考文献

- [Selvaraju 17] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D.: Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, in Proc. of ICCV-17, pp. 618626 (2017)
- [Greydanus 17] Greydanus, S., Koul, A., Dodge, J., and Fern, A. Visualizing and Understanding Atari Agents. arXiv preprint arXiv:1711.00138 (2017)

動的報酬クラスタリング

Dynamic Reward Clustering

比嘉 亮太^{*1} 加藤 淳也^{*2}

Ryota Higa

Junya Kato

^{*1}NEC データサイエンス研究所

NEC Data Science Research Laboratories

^{*2}NEC AI・アナリティクス事業開発本部

NEC AI Analytics Business Development Division

Real-world time series data have various patterns by the human operation. Our aim is extraction of the valuable information from the time series data with action. And we need to interpret people's policy from time series data. We propose a interpretable method for clustering the dynamic rewards from the time series data. Combining Wavelet transformation preprocessing and simple clustering methods to the human motion data and inverted pendulum simulation, our approach was successful in clustering different rewards and the interpretability of feature while maintaining the time series information.

1. Introduction

時間と操作を含むデータから価値のある情報を抽出することは様々なサービスを提供する企業の研究として重要である。データを取得する環境も操作する人も時間経過によって変化をする難しさがある。隣り合った時間の関係だけではなく、長期間の関係性を抽出することも必要である。

公共施設の運用や金融など人の操作と時系列の変化は複雑に絡まっており、直感的な判断が危険を及ぼす一方で既存のルールベースの方法では効果的な改善が難しくなっている。これはシステムの大規模化と通信の高速化による影響であり、データに基づく自動化と支援は今後ますます必要になる。時系列データを可視化し解釈することは、人の判断の支援に繋がる一つの方法である。感覚的に実行してきた意思決定に対応する方策を反映した時系列データから、数値的な指標の一つである報酬を記述 (descriptive) する価値は計り知れない [Bazerman 06]。

操作と時間依存する観測情報を有するデータから、指標の一つである報酬を抽出することはエージェントとシステムが相互作用する強化学習と実社会のデータの行き来を可能にする。環境とエージェントが相互作用する仕組の実社会システムへの応用が広まり始めており、観測した情報から有効な状態を学習する仕組みも報告されている [Lesort 18]。操作を含む時系列データの問題点は、操作による非周期性がある点や系列全体では十分なサンプルがない点である。過去のデータに基づいて予測モデルを構築しても、別の操作が行われると全く別のモデルになってしまう。一方で、強化学習の様にさまざまな行動パターンを実社会で実行することは困難である。仮想環境を構築する場合も、物理や人の心理をある程度モデル化する必要がある。網羅的にデータが得られる場合は模倣学習により、報酬と環境を学習するアプローチも提案されてきている [Osa 18, Baram 17, Sasaki 18]。系列が少ない場合は仮説を立て実験的にデータを取得して分析し解釈する必要がある。

上記の問題に取り組むために、我々は操作の影響を受けた時系列データから報酬をクラスタリングする方法を提案する。時系列データ全体では、同じパターンは現れないが短期的には類似するパターンがある。研究所とデータサイエンティスト、

事業部や客先との取り組みによってできた技術の一部を紹介する。時間という物理的に素性の良い順序関係を維持しながら、短期的なクラスタリングと非連続な時間での状態のクラスタリングを同時に行う手法である。データ自体の意味を保持したまま、処理にも説明がつけられるシンプルな方法で構成した。可視化することによる現象の解釈性の向上だけでなく、強化学習の報酬設計への接続やダイナミクスの変化への遷移確率設計にも役立つことが期待される。報酬の数値化によって行動の理解へ繋がることも期待される。

2. Dynamic Reward Clustering

操作を受けた時系列データに対し系列情報を保持しながら、報酬に対して自動クラスタリングする手法を提案する。行動や操作は、目的を達成するための指標に基づいて行われるとする。その指標を本論文では報酬 r として定義し、実データ分析と強化学習の行き来が自在になるいくつかの仕組みを検討し続けている。用語については、強化学習に従う [Sutton 18]。

2.1 Background

外部からの操作の影響を受けた時系列データを $\mathbf{s}(t)$ とする。時刻 t における i 成分における観測データを $s_i(t)$ とし、全ての観測データをベクトルとして $\mathbf{s}(t)$ と定める。この系列データは、 $r(\mathbf{s}(t), a(t))$ に基づいて行動した方策 $\pi(a(t)|\mathbf{s}(t))$ の影響を受けたデータとする。方策や報酬は普遍でなく時間によって変化する。明示的に詳細な行動データ $a(t)$ を与えない場合にも適用できる。

2.2 Assumption

データを取得する対象となるシステムのダイナミクスは外部からの行動によって変化する。外部への入出力の操作がない場合は、システムのダイナミクスは変化をしないと仮定する。例えば、波形のデータが定常である場合は外界からの操作がなく、非定常や急激な変化が生じた場合は運用の方策に変更が生じた場合であるとする。

2.3 Scale-Reward Relationship

時刻依存の情報を残しながら、非周期性を考慮したスケールを抽出できる特徴量の設計方法を提案する。得られた特徴量を人が読み解ける、解釈可能な指標を目指す。我々は時間方向のラベルを維持したまま、非周期性の構造に対して有効である

連絡先: 比嘉 亮太、NEC データサイエンス研究所、神奈川県川崎市中原区下沼部 1753、email:r-higa@ct.jp.nec.com

手法として Wavelet 変換 (WT) を採用する [Haar 10]。時系列データのパターンの変化が報酬や方策の変化によって生じるとし、これは波形のスケール a の変更と等価であるとみなす。

時刻 t の観測状態ベクトル $\mathbf{s}(t)$ の i 成分に対して

$$\omega_i(a, b) := \int \psi(a, b) s_i(t) dt \quad (1)$$

と変換を定義する。ここで、 a はスケールであり b は時刻であり $\psi(a, b)$ は基底関数である。Scalogram 行列として

$$C := \begin{bmatrix} \mathbf{c}_1^T & \mathbf{c}_2^T & \cdots & \mathbf{c}_I^T \end{bmatrix} \quad (2)$$

を定める。ただし、

$$c_i(a, b) := |\omega_i(a, b)|^2, \quad (3)$$

$$\mathbf{c}_i := \begin{bmatrix} c_i(a_1, b_1) & \cdots & c_i(a_1, b_M) \\ \vdots & \ddots & \vdots \\ c_i(a_N, b_1) & \cdots & c_i(a_N, b_M) \end{bmatrix} \quad (4)$$

と与えられる。物理学の解析で用いられるフーリエ変換 (FT) と比較した WT の利点は、非定常な変化に強く、時間情報を保持することができる点である。静的な性質であれば FT は効果的である。短時間フーリエ変換 (STFT) は、 $\Delta\omega\Delta t \geq 1/2$ の不確定性関係によって生じる周波数分解能と時間分解能のトレードオフがあり窓関数の幅を決められず広域な周波数間の比較への適用は困難である。WT はスケール a により適切なトレードオフを決めることが可能になる。

2.4 Dimension Reduction

高次元なデータの低次元化は、データの特徴を解釈するための有効な手法である。特に PCA/SVD を用いて 1 から 3 次元といった低次元への圧縮から、報酬をどのように抽出するかを説明する [F.R.S. 01, Golub 13]。特異値分解 (SVD) の場合は、

$$C = U\Lambda V^T \quad (5)$$

と行列が分解される。時間 b を維持したまま、スケール a の行方向に対して d 次元まで削減する。 Λ の d 個までに対応する行列 V_d を元の Scalogram 行列に掛け合わせることで、

$$\tilde{C}_d = CV_d \quad (6)$$

と次元圧縮できる。スケール a 方向を 2 次元への圧縮すると図 1 のようになる。低次元化の恩恵は、信号処理として知られるノイズの除去に加え解釈性の向上、例えば可視化による直感的な判断を人である分析者が行える。また、データ数に対して表現空間が高次元になりやすい時系列データの圧縮により学習時の次元の呪いを回避できる。

2.5 Dynamic Clustering on Scale Space

データの時系列情報を保持したスケール空間において、データ点のクラスタリングを行う。もっともシンプルな方法として、**k-means++** に対して、上記の Scalogram による前処理を組み合わせた手法を提案する [Arthur 07]。これは、静的なクラスタリングを時間依存を含む動的なクラスタリングへの拡張になっている。クラスタリングを行う目的関数

$$J := \sum_{n=1}^{b_M} \sum_{k=1}^K \delta_{n,k} |\mathbf{c}_n - \boldsymbol{\mu}_k|^2 \quad (7)$$

を最小にするクラスターにデータ点を割り当てる。ここで、 $\delta_{i,j}$ はデルタ関数であり、 $i = j$ の時 1、それ以外は 0 を与える関数である。**k-means++** の場合はデータ点の順番に意味はないが、時系列データに対してはデータ点の並びと向きは重要である。クラスタリングを行う空間に対して時系列データのスケールとして与えられる空間上で行なっている。クラスター数はエルボー法や BIC により決定できる。我々の提案手法は、操作によってデータがどのように分類されるのかに対する一つの解釈性を与えることができる。詳細な内部の構造同士を比較したい場合は、階層型クラスタリングを適用する。

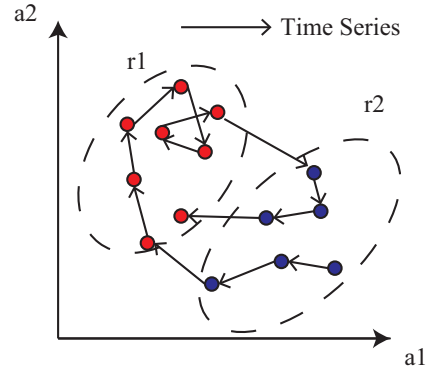


図 1: 2 次元のスケール空間での系列データの分類。それぞれのクラスは異なる報酬 r_1, r_2 によって行動したグループとして与えられる。隣接する点だけでなく、時間的に長期の点列間のグループ化も同時にできる。 r_1 に含まれるデータは内部にもリング構造を持っており、より詳細な報酬関数でのクラスタリングができることを意味している。

2.6 Dynamic Reward

上記のクラスタリングにより、時系列データ $\mathbf{s}(t)$ をスケール空間で分類可能になる。スケール空間が操作の影響によって引き起こされるという仮説を用いて、クラスタリングの情報を報酬として与える。**k-means++** の場合のもっとも簡単な表記は、 $t = b$ としてその系列データがどのセントロイド μ_k に割り当てられるかを示せばよい。

$$r(\mathbf{s}(t), a(t)) := f(\mu_k(\mathbf{s}(t))) \quad (8)$$

これは、操作の影響を受けた系列データがどのクラスに分類されるのかを表している。 $f(\mu_k(\mathbf{s}(t)))$ は \tilde{C} の固有値を返す関数とすることで、相対的な報酬間の比較ができる。例えば、クラスが r_1 と r_2 の 2 通りの行動を決める報酬があるとする。我々の手法は、現在の行動がどのクラスの報酬に基づいてなされたかを自動で判定することを可能にする。また、それぞれのクラスでの部分データで遷移確率を学習することもできる。過去に同じクラスに属したデータであれば、モデルを推定せずに再利用することが可能になる。Markov 的な直近での分類だけでなく、時間を跨いだ長期間での系列データのグループ化も自動で可能になる。

2.7 Cluster and Ring

動的な報酬クラスタリングの解釈として、クラスターとリング構造に注目する。前者のクラスター化は類似する報酬のグループ化である。一方、後者のリングはある報酬のクラスターから出発して戻るという報酬間の周期性に相当するものである。図 1 は r_1 と r_2 にクラスター化され、そのクラスター間での時間的なループによるリング構造を示す概念図である。

2.8 Joint to Dynamics model

連続性を有効活用する時系列データから常微分方程式 (ODE) [Heinonen 18, Chen 18] などの前処理として動的報酬クラスタリングを行うことでアクションの効果を含みながら有効的なサンプル数の向上とモデルの再利用性が向上する。一つの系列として見た場合は、その場限りのモデルになるが、再生可能なブロックでモデルを構築して組み合わせることができる。同じ報酬に分類されるサンプルが新たに得られた場合は、モデルを学習する必要がなくなる。また、行動によって異なるサンプルに偏るバイアスを補正する効果もある。

以下では実験において、ある時刻で変化する異なる報酬関数によって行動するエージェントが操作したシステムの系列データに対して我々の提案手法の有効性を示す。

3. Experiment

実験の為に作成した Cart Pole データおよび、MoCap データ^{*1}を用いて提案手法の有効性を検証した。実験で用いたデータの詳細は下記の通りであり、これらに対してアクションの自動抽出を行った。強化学習の文脈では、時間に応じて目的が変化するカリキュラム学習 [Sukhbaatar 17] に相当する。なお、各データはセンサごとに平均 0、分散 1 に標準化を行った上で使用し、**k-means++**のクラスター数 k は Cart Pole データについてはポリシーの数、MoCap データについては独自に作成した正解ラベルの数を用いて実験を行った。また次元圧縮は PCA を使用し、圧縮後の次元数は累積寄与率が 0.95 確保される数とした。

1. Cart Pole データ: $t = 2000$

Q-learning でよく用いられる Cart Pole に対し、5 つの異なるポリシーを学習させ各々 400 エポック実行した時の Position、Velocity、Angle、Angular Velocity のデータを繋げたデータである。

2. MoCap データ

人の動きを計測したモーションキャプチャのデータであり、両腕と両足の 4 つの加速度を 120 フレーム/秒の間隔で収集している。なお、今回は MoCap の中から beaks、wings、tail feathers、claps の 4 つの動きからなるチキンダンス (chicken dance: $t = 1536$) と、walk、run、jump、kick、jump(left)、jump(right)、wave hands の 7 つの動きからなるモーション (walks: $t = 8402$) の 2 つを使用した。

実験の結果、chicken dance データ、walks データの各精度は、0.700、0.860 となった。なお、精度は正解ラベル数をデータの全体数 t で割ることで算出した。Cart Pole データについては、各ポリシーを 1 つのアクションと見做した場合の精度は 0.607 となったが、図 2 の使用したデータを見ればわかる通り、実際はほぼ同じアクションをしている部分がある。その為、改めて図 2 のように正解ラベルを作成し、精度を測ったところ、0.789 という結果になった。また、Cart Pole データを 2 次元スケール空間に表示し、提案手法によるクラスタリング結果によって各点に色を付けた結果が図 3 である。なお、各点の色は図 2 における **k-means++**の結果の色と対応している。この結果から、主要なモーションがよくまとまっている様子が確認できた。

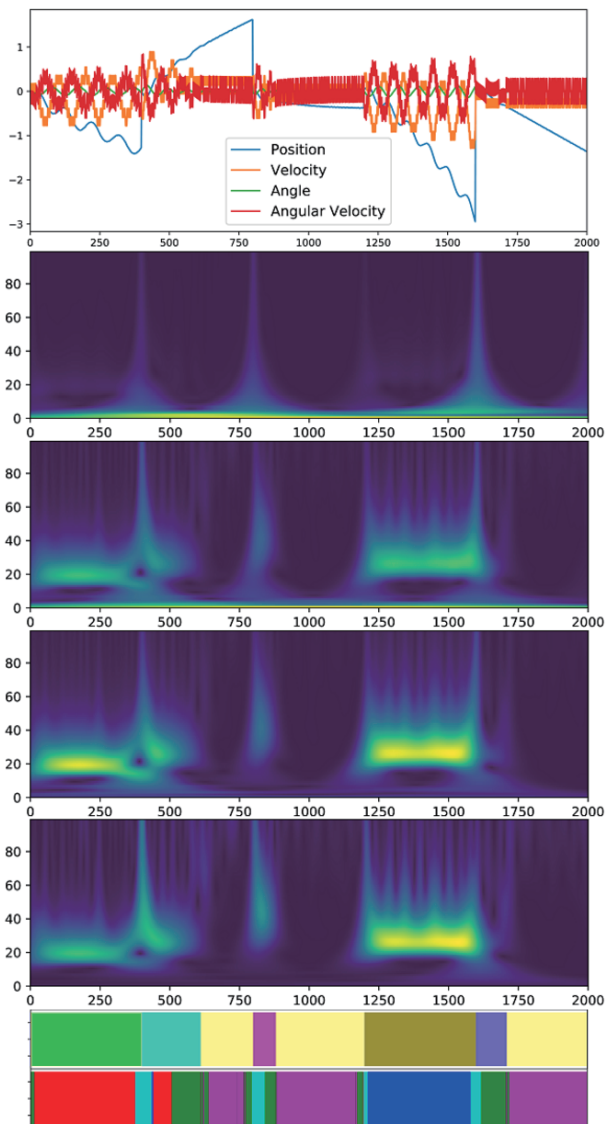


図 2: Cart Pole データの実験結果。上図から使用したデータ、Position の低周波数帯の Scalogram、Velocity の低周波数帯の Scalogram、Angle の低周波数帯の Scalogram、Angular Velocity の低周波数帯の Scalogram、正解ラベル、**k-means++**の結果。

4. Discussion

行動の指標となる報酬と時系列データのスケールの対応関係に基づく我々の提案手法により目的が時間に応じて変化を分類することが確認できた。エージェントや人が学習した結果の行動を再度データから捉え直すことに一部成功している。

アクションの変化点を抽出できていることが確認された。また、2 次元で可視化した結果からも変化点においては点間の間隔が広く、同じアクションをしている間は間隔が狭くまとまっていることから、解釈性があることも確認できた。注目すべきは図 3 における紫色のクラスターである。このクラスターに該当するアクションとしては、エージェントが細かく左右に動き、Pole の傾きをほぼ一定に保つアクションになる。然しながら、それぞれ Pole の傾きが異なる為 Position のずれ方が異なる状況でありながら、提案手法ではその違いを吸収しこれらを同一のアクションとして抽出できている。

*1 <http://mocap.cs.cmu.edu/>

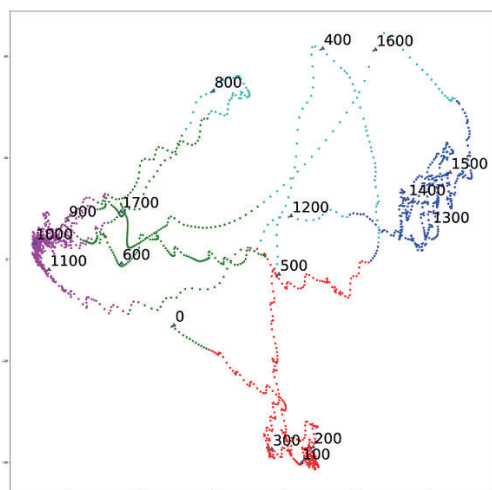


図 3: Cart Pole データを 2 次元スケール空間へ圧縮した結果。各点の色は `k-means++` によるクラスタリング結果に対応し、図中の数字は時間を示す。クラスター構造の確認とその間を行き来するリング構造が確認できる。動画として観察することで、時刻変化と相対的な報酬のより明確な比較ができる。

シミュレータ上で作成されたデータだけでなく、現実の人のモーションをセンサから区別することができ、提案手法の実用性の高さを確かめられた。

解釈性について、2 次元まで次元圧縮をしたことで情報が失われてしまっている可能性も考えられる。そこで、今回使用したデータに対し PCA で 2 次元に圧縮した時の累積寄与率を求めたところ、Cart Pole データ、chicken dance データ、walks データの結果はそれぞれ 0.745、0.739、0.629 となった。このことから、圧縮率が高く、また情報が失われていないことがわかる。よって、2 次元に圧縮した画像を見ることで、アクションが分類された理由の解釈は十分可能である。

5. Conclusions

我々は操作を含む時系列データから動的に報酬をクラスタリングする手法を提案した。実験で紹介したデータに加え、いくつかの人の運用を含む実データに対しても我々の取り組みが有効であることが確認できている。我々の提案手法はデータ自体の付加価値を上げるための報酬で分割する”ハサミ”と時系列をつなぐ”のり”に位置付けられ、DNN や HMM などの表現力の高い学習モデルの性能をより引き出すことにも貢献できる。また、解釈性から理解へ進める場合において、例えば物理学で自然界を記述するハミルトニアンが重要な様に対象を記述することが効果的であると筆者は考えている。

今後は、強化学習への実社会応用として進めている模倣学習 [Lesort 18] を主軸とした発展を行う。また、可視化による現状の運用比較、運用改善を利用した為替取引や HR Tech といった応用領域を含む人やシステム間の連携がある指標を明確化することが重要な領域へ向けて研究を継続する予定である。

参考文献

[Arthur 07] Arthur, D. and Vassilvitskii, S.: K-means++: The Advantages of Careful Seeding, in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '07, pp. 1027–1035, Philadelphia, PA,

USA (2007), Society for Industrial and Applied Mathematics

[Baram 17] Baram, N., Anschel, O., Caspi, I., and Mannor, S.: End-to-End Differentiable Adversarial Imitation Learning, in Precup, D. and Teh, Y. W. eds., *Proceedings of the 34th International Conference on Machine Learning*, Vol. 70 of *Proceedings of Machine Learning Research*, pp. 390–399, International Convention Centre, Sydney, Australia (2017), PMLR

[Bazerman 06] Bazerman, M.: *Judgment in managerial decision making*, J. Wiley (2006)

[Chen 18] Chen, T. Q., Rubanova, Y., Bettencourt, J., and Duvenaud, D. K.: Neural Ordinary Differential Equations, in Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R. eds., *Advances in Neural Information Processing Systems 31*, pp. 6572–6583, Curran Associates, Inc. (2018)

[F.R.S. 01] F.R.S., K. P.: LIII. On lines and planes of closest fit to systems of points in space, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, Vol. 2, No. 11, pp. 559–572 (1901)

[Golub 13] Golub, G. H. and Loan, van C. F.: *Matrix Computations*, JHU Press, fourth edition (2013)

[Haar 10] Haar, A.: Zur Theorie der orthogonalen Funktionensysteme, *Mathematische Annalen*, Vol. 69, No. 3, pp. 331–371 (1910)

[Heinonen 18] Heinonen, M., Yildiz, C., Mannerström, H., Intosalmi, J., and Lähdesmäki, H.: Learning unknown ODE models with Gaussian processes, in Dy, J. and Krause, A. eds., *Proceedings of the 35th International Conference on Machine Learning*, Vol. 80 of *Proceedings of Machine Learning Research*, pp. 1959–1968, Stockholmmsan, Stockholm Sweden (2018), PMLR

[Lesort 18] Lesort, T., Daz-Rodriguez, N., Goudou, J.-F., and Filliat, D.: State representation learning for control: An overview, *Neural Networks*, Vol. 108, pp. 379 – 392 (2018)

[Osa 18] Osa, T., Pajarinen, J., Neumann, G., Bagnell, J., Abbeel, P., and Peters, J.: An Algorithmic Perspective on Imitation Learning, *Foundations and Trends in Robotics*, Vol. 7, No. 1-2, pp. 1–179 (2018)

[Sasaki 18] Sasaki, F. and Kawaguchi, A.: Deterministic Policy Imitation Gradient Algorithm (2018)

[Sukhbaatar 17] Sukhbaatar, S., Kostrikov, I., Szlam, A., and Fergus, R.: Intrinsic Motivation and Automatic Curricula via Asymmetric Self-Play, *CoRR*, Vol. abs/1703.05407, (2017)

[Sutton 18] Sutton, R. S. and Barto, A. G.: *Reinforcement Learning: An Introduction*, The MIT Press, second edition (2018)

文章中の潜在要素を考慮した対話システム

A Dialogue System Implemented with Latent Parameters

李為達 *¹

Edward Li

日永田智絵 *²

Chie Hieda

長井隆行 *^{2*3}

Takayuki Nagai

*¹聖光学院中学校高等学校

Seiko Gakuin High School

*²電気通信大学

The University of Electro-Communications

*³大阪大学

Osaka University

When given a conversation, traditional dialogue systems mainly focus on the context that can be observed on the surface of sentences; concretely, they process and determine the output based on the grammar, visible keywords and structure of the sentence. However, the content we convey to others is affected by a multitude of latent parameters, such as emotional state, personal knowledge and personality. Therefore, we have attempted to validate the integrity of a dialogue system which takes these latent parameters into measure, and have successfully developed a dialogue system which utilizes latent parameters as input.

1. はじめに

近年、深層学習の発展に伴い、様々な状況における人間との対話を目的とした対話システムの開発が進んでいる。従来のシステムでは、話し手が与える入力文の表層的なコンテキストに着目し、文章の構造やキーワードに重点を置いて出力を生成するようなモデルが多い。一例として、二つの LSTM (Long short-term memory network) [Hochreiter97] をエンコーダーとデコーダーとしてつなげた Sequence to Sequence モデル [Sutskever14] を使用した対話システムが挙げられる [Csaky17]。Sequence to Sequence モデルは、シーケンスのペアを大量に学習することで片方のシーケンスからもう一方を生成できる性質を利用して、文章の表層的なコンテキストに着目した対話を実現することができる (図 1)。さらに、こうした対話システムを応用して返答の質を上げた事例として、2017 年にアマゾンが開催した「Alexa Prize Competition」においてモントリオール大学の研究室が開発した対話システム「MILABOT」が挙げられる。これは、今までに開発されたモデルを複数集め、各モデルの出力の中で一番質が良い返答を出力として使用するものである [Serban17]。

これらの対話システムは、入力文の表面的な要素のみに注目しており、人同士の会話において重要な役割を果たす「感情」などの潜在的な要素を陽に考慮していない。そのため、それらを無意識に感じる人間のように会話を行うことが難しい。そこで本稿では、入力文から抽出できる潜在要素を陽に算出し、それらをモデルの入力として使用した対話システムを構築し、その有用性を検証することを目標とする。

2. 会話中の潜在要素

2.1 人間の意思決定プロセス

対話システムは人間と会話を行うためのシステムであり、人と会話を行ったときに相手から人との区別がつかないような会話を行うことが究極的な開発目標であると言える。対話システムが人間と区別がつかないようにするために必要とされる要素は、人間が持つ要素を分析し、それらを陽にモデル化するのが一つの方向性である。ここで、人同士で会話を行う場合に会話

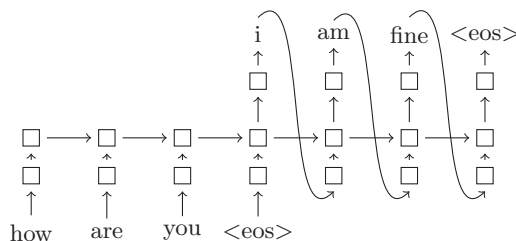


図 1: Seq2Seq を使用した対話システム [Csaky17]

の進行を左右する要素として以下の 4 つの特徴を考える：

- 問われている返答の種類の理解

会話中に話し手が聞き手へ伝達している文章が聞き手に何を求めているかを理解することで、聞き手はどのような返答を行えば良いのかが理解できる。例えば、質問を受けたときにはそれに対する答えを返すのが一般的であるが、質問に答えずに挨拶を行うと会話としては不適切となるため、聞き手に対する文章の、要求の分類と検出が必要となる。

- 会話文中の感情の認識

話し手から同じ文を与えられても、聞き手は受けた時の感情により返答が変化することがしばしばある。従って、相手と自分の感情は出力に影響を与える重要な要素であると言える。

- 会話中のキーワードに関する前提知識

人は会話内で出現するキーワードについて連想をし、例えば話題を変えようとする時などに、それが相手への返事に影響を与えることがある。対話システム内でもこのプロセスを含めることで、会話を円滑に進める可能性を残すことができる。

- 過去の会話文の内容の把握

人同士の会話は一文一文を単体でかけ合うというよりも複数のやりとりから成り立つものであり、会話中の

連絡先: 李 為達, 聖光学院中学校高等学校, 神奈川県横浜市中区滝之上 100, liweida9@gmail.com

ある時点に至るまでのタイムステップより以前の会話の吟味が可能でなければ、人間のような会話は実現が難しい。会話の内容を始終まで維持するためには、この要素が必要不可欠である。

これらの要素は単に考慮するだけでなく、総合的に考慮して出力を考える必要がある。また、会話の経験に基づいて得た要素に適した返答を生成する必要がある。つまり、要素をすべて踏まえた上で、今までその人が経験してきた会話文の構造やパターンに基づいて返答を生成する。これらの特徴を再現するために、対話システムには以下の機能が必要とされと考えられる。

- 会話文分類

人間が問われている返答を認識できるという要素を模擬するために、入力文から問われていることを計算する機能が必要である。

- 会話文中の感情検出

人間が自分の感情を認知すると共に、発現に含まれる感情を認識しながら会話を行っているように、対話システムでも同様の機能を果たすモデルが必要である。

- 会話に必要な前提知識

人は他人と会話を行う時に何も知識がない状態から始まるのではなく、世間一般において常識と定義づけられる、例えば、果物についての会話の時、「果物は食べ物である」や、「果物は植物である」といった、ある程度の前提知識を持っている。よって、これをシミュレートする機能も対話システムには必要となると考えられる。

- 会話文貯蓄

人間が過去の会話の内容を覚えているのと同様に、会話中のあるタイムステップに至るまでの会話の趣旨を大まかに保存する必要がある。

- 会話文のベクトル表現

機械学習モデルが自然言語の処理を可能にするためには、文章をベクトルとして扱う必要がある。

本稿では、文章の潜在的な要素を認識するモジュールとして、上記の5つの機能に着目し、システムを構築することとした。

3. 対話システムの構築

図2に、本稿で構築した対話システムのダイアグラムを示す。タイムステップごとに入力文から返答の出力文を生成するプロセスを繰り返し、同じ会話内である限りタイムステップごとの会話文を潜在空間で抽象化してできたベクトル列を保存し、次のタイムステップで利用する。

話し手から受け取った文章を単語に分割し、GloVe[Pennington14]を用いて各単語をベクトルに変換する。それらを集めたベクトル列を会話文のベクトル表現として扱い、それぞれのモジュールでこのベクトル列を入力して処理を行う。処理後に会話文から潜在要素を抽出した各モジュールの出力値を集め、それらを会話文を抽象化した数値として自己組織化モジュールに入れ、出力の計算を行う。以下、各モジュールについて説明する。

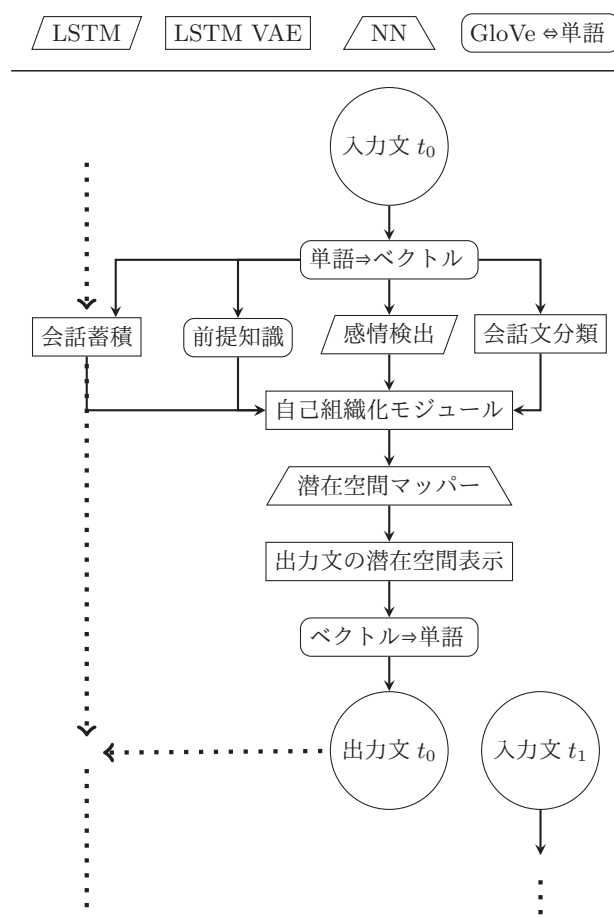


図 2: 提案するシステム

3.1 単語ベクトル変換モジュール

文章のままでは処理が行えないため、GloVeでワードエンベディングを行い、ベクトル化した。本稿ではGloVeのモデルとして、Stanford NLP Groupのウェブサイトから入手可能な“Wikipedia 2014 + Gigaword 5”の200次元及び300次元のバージョンを使用した。入力を受けた文章は単語と記号をスペースで分割し、それぞれに対応したワードエンベディングを文中の出現順にクترل化したベクトル列を構築する。構築したベクトル列は、感情処理モジュール、文種類処理モジュール、会話蓄積モジュールの3つのモジュールに引き渡される。

3.2 会話文分類モジュール

ベクトル列として表現された文章の会話内における役割を、可変長である文章を入力として受け、教師なしで潜在空間における分類が可能なLSTM VAEを用いて算出する。

3.3 感情検出モジュール

ベクトル列で表現された文章から算出できる相手の感情と自分の感情を計算し、8種類の感情の中から最も近い感情を話し手と聞き手それぞれについて出力する。モデルとしては、可変長である文章を入力として受けることができるLSTMを使用する。

3.4 前提知識モジュール

元の英文の中でキーワードとなる単語を、Rapid Automatic Keyword Extraction algorithm [Rose10]で抽出し、それらと

関連するとされる単語の GloVe におけるベクトル表現を、最大 3 つまで出力する。

3.5 会話蓄積モジュール

話し手と聞き手の一回のやりとりをペアとして、次のタイムステップで使用するために保存する。これは人が会話中にキーワードと呼ぶことができる単語を聞いた時に、無意識に連想をしてしまい会話に影響を与えるような傾向があることに基づいたモジュールである。「会話の流れ」を会話蓄積モジュールの潜在変数のベクトルを通して数値化することで、会話の進み方を可視化するなどといった応用も可能となる。

3.6 自己組織化モジュール

上記の 4 つのモジュールの出力を総合し、一つ一つのベクトルが GloVe の辞書内である一つの単語に対応するようにベクトル列の形で表現された文章を出力する。モデルには、LSTM VAE を使用した。

3.7 出力文潜在空間モジュール

このモジュールは、2 章の中で定義される「会話の経験に基づいて、得た要素に適した返答生成」である。LSTM VAE を用いて、あらかじめ返答文の文の構造を教師なしで学習を行う。これによって、自己組織化モジュールの出力で潜在空間が定まれば出力文の趣旨も定まることになる。

3.8 潜在空間マッパー

自己組織化モジュールで会話文中の要素をまとめたものの潜在空間における表現と出力文潜在空間モジュールの潜在空間的表現の対応を学習し、新規の文の入力を受けたときに抽出された要素に基づいて出力文の潜在空間的表示の生成を行う。

3.9 ベクトル単語変換モジュール

単語ベクトル変換モジュールと同じように、ワードエンベディングのベクトルをそれらに対応した単語に変換し、組み合わせて生成文として出力を行う。

各モジュールの学習誤差、Optimizer、ロス関数及び入力・出力の次元数を表 1 に示す。

4. 使用したモデルの詳細

4.1 LSTM

感情検出モデルに使用した LSTM は基本的に中間層を持たず、入力は GloVe によってワードエンベディングでベクトル化した 200 次元の単語ベクトルを可変数回受けることができる。感情検出モデルは、感情に対応する 1 つの数値を出力する。

4.2 LSTM VAE

LSTM VAE は、一般的に使われる VAE のエンコーダーとデコーダーを LSTM で置き換えたものである。可変サイズの入力を受けることができるエンコーダー側から出力されるベクトル列は、入力の長さに依存せず一定の大きさの行列を出力し、これを一般的な VAE と同様にデコーダーに入力し元の可変長の入力を復元する。LSTM VAE の誤差は、元の潜在要素の値の復元を行なった時の元の値と復元した値の誤差を意味する。本稿でこれを使用した理由は、殆どの場合で可変長である会話文を汎用的に潜在空間に変換するモデルであるためであり、会話文分類モジュールなどに使用することができる。

5. 検証

5.1 検証方法

モデルの会話の質を把握するための手段として、モデルの本来の目的である人との会話を実行した。今回は人が話し手、

モデルが聞き手として人から話をかけられた時に返答を行うという設定のもとで検証を行なった。会話の終了の判断は、話し手の人が不自然な会話となりつつあると感じる、または会話が終わったことを人が自然に感じ取ることができたときとする。

5.2 検証結果・考察

検証の 2 つの例を表 2 と 3 に示す。

一つ目の結果について、会話のやりとりを一行ずつ、2 章で触れた会話中の要素を考えながら見ていきたい。

1 行目では人間は感情なしの挨拶を行ったといえる。この時、一般的な人間は殆どの場合は同じく挨拶で返すと予想される。人が予想する通りに、対話システムは人間の挨拶への返答をし、同じ主旨のことを聞く文を返している。2 行目と 3 行目では、すべて人間が返しうる返答といえるので、それぞれ適切な出力を行ったといえる。2, 3, 4 行目の人間が与えた文章はそれぞれコメント、質問と賛同と概ね分類することができる。4 行目において、会話が終わってしまうような返答を人間がしている。しかし、対話システムは話題を変えて会話を続けようとする意向が見える返答を行っている。ここでは、コメントとともに質問を返しており、人間が会話を続けようとするときと非常に似た形式の会話を行っているとも言える。5 行目では、人間は感情を含まない挨拶を行ってから、それに対するコメントを返しておりこれは適切な返答であると考えられる。

次に、例 2 について見ると、概ね例 1 と似た形式を取っていることがわかる。しかし、5 行目において“Know you not reading?” という文法的に誤った出力を返している。この原因として考えられるのは、学習不足である。この実験から、今回得た対話システムはある程度の会話を実践することが可能であり、かつ会話を継続しようとするテクニックも所持しているように見受けられる。

6. まとめ

本稿では、人が会話中に感じると推測される潜在的要素を陽にパラメータとした対話システムのモデルを提案した。モデルの検証については、人が実際にインタラクションを複数回行った時の対話システムの返答の質を特定の基準でなく、人の感覚によって判断しているため、安定して使用できる正確な検証方法とは言い難い。よって、今後は Conversation-turns Per Session(CPS) [Zhou18] という対話システムの会話の続行性を測る尺度で検証を行いたいと考えている。

今回の検証から、潜在要素を意識した対話システムはある程度人間に近いような会話を行うことができることが分かった。しかし下記のように、いくつか疑問点が残るところがある。

- 潜在要素の影響はどれくらいあったか?

今回の対話システムでは、潜在要素をまとめる自己組織化モジュールへの入力は受け取った入力文自体は使用しておらず、抽出した潜在感情やキーワードなどを使用している。しかし、使用した LSTM VAE 自体、与えられたデータへの学習の正答率は極めて高いモデルである。よって、今回の対話システムでも潜在要素自体が大きな意味を持ったのではなく、ただ単に使用したモデルが与えられたデータに適応してしまった可能性がある。今後、この点を検証する必要がある。

- 潜在要素はこれで十分か?

今回の対話システムはある程度成功したと言うことができる。しかし、今回注目した潜在要素は人が意識を

表 1: 各モジュールの詳細

Module name(Model type)	Loss	Loss function	Optimizer	Dimensions
会話文分類 (LSTM VAE)	1.26×10^{-2}	KL Loss	RMSPProp	300 \rightarrow 200
感情検出 (LSTM)	1.03×10^{-2}	sparse_categorical_crossentropy	AdamOptimizer	200 \rightarrow 1
自己組織化 (LSTM VAE)	8.7×10^{-3}	KL Loss	RMSPProp	200 \rightarrow 60
出力文潜在空間 (LSTM VAE)	4.6×10^{-3}	KL Loss	RMSPProp	300 \rightarrow 200
潜在空間マッパー (NN)	1.5918×10^{-9}	mean_squared_error	Adam	60 \rightarrow 200

表 2: 人間と対話システムの会話検証: パターン 1

Timestep	Human	Dialogue System
1	Hi, how are you today?	Doing great, you?
2	I'm doing great.	That's nice. What do you do for work?
3	I code. How about you?	Just office work, but it is a lot.
4	Oh, ok. Do you enjoy your work?	Yeah, it's quite rewarding.
5	Yeah, it does pay the bills.	When I'm not working, I play basketball. What do you do for fun?
6	I hike at the park.	That sounds relaxing!
7	It keeps my mind off things. Do you travel?	Not much.

表 3: 人間と対話システムの会話検証: パターン 2

Timestep	Human	Dialogue System
1	Hi! How are you?	Hi good and you?
2	Great! Just ran outside.	What do you do for a living?
3	I am a student. How about you?	I play basketball.
4	Nice! Do you love books?	I read all the time. How about you?
5	I always liked movies more than books.	Know you not like reading?
6	I do, but prefer the movies.	Ok, that makes sense.

しうる要素の中の一部であるということは容易に分かる。よって、これらの要素のみである程度の会話が可能であるということは、人間も簡単な対話ではこの程度の要素しか考慮していないという可能性も考えられる。この真偽は定かでないため、今後検証の余地がある。提案モデルは、対話文の要素を陽にモデル化しているため、対話データの要素を解析することもできる。

● 対話システムの返答内容の統一性

一般的にあるデータセットを使って対話システムの学習を行うと、対話システムの返答が統一性に欠けることが頻繁にある。これは対話システムにおける難点の一つであるが、今後対話システム特有の「プロフィール」を作り、プロフィールの内容が必要とされるパターンの検出とプロフィールに基づく返答文の生成を行う。

今後はこれらを改善する方向に研究を進めたい。また、現時点のモデルへ会話に影響を与えうるモジュールを加えらるとともに、モデル自身のパーソナリティの固定など、より人間が持つ特徴をモジュールの形で付け加え、さらなる検証を行うことが今後の課題である。

参考文献

[Hochreiter97] S.Hochreiter, and J.Schmidhuber, "Long Short-term Memory," Neural Computation 9(8):1735-80 (1997)

[Sutskever14] I.Sutskever, O.Vinyals, and Q.V.Le, "Sequence to Sequence Learning with Neural Networks," NIPS 2014 (2014)

[Csaky17] R.Csaky, "Deep Learning Based Chatbot Models," Technical Report (2017)

[Serban17] I.Serban, C.Sankar, M.Germain, S.Zhang, Z.Lin, S.Subramanian, T.Kim, M.Pieper, A.Chandar, N.Ke, S.Mudumba, A.Brebisson, J.Sotelo, D.Suhubdy, V.Michalski, A.Nguyen, J.Pineau, and Y.Bengio, "A Deep Reinforcement Learning Chatbot," CoRR2017 (2017)

[Pennington14] J.Pennington, R.Socher, C.D.Manning, "Glove: Global Vectors for Word Representation," <<https://nlp.stanford.edu/projects/glove/>>, (2014)

[Rose10] S.Rose, D.Engel, N.Cramer, and W.Cowley, "Automatic Keyword Extraction from Individual Documents," Text Mining: Applications and Theory (2010)

[Zhou18] L.Zhou, J.Gao, D.Li, and H.Shum, "The Design and Implementation of XiaoIce, an Empathetic Social Chatbot," arXiv:1812.08989 (2018)

深層逆強化学習による自動運転の安心走行実現

Comfortable Driving by Deep Inverse Reinforcement Learning

岸川 大航^{*1}

Daiko Kishikawa

荒井 幸代^{*1}

Sachiyo Arai

^{*1}千葉大学大学院 融合理工学府 都市環境システム

Department of Urban Environment Systems, Graduate School of Science and Engineering, Chiba University

For the realization of automatic driving, not only safety but also comfortability of passengers is required for its application to the real society. We define it as comfortable driving. Comfortable driving is hard to define because the expectation for comfortability varies according to the designer of system. Therefore comfortable driving is difficult to code a rule-based algorithm manually. Reinforcement learning, which learns an optimal policy from trial-and-error by the agent, is an effective method to solve this problem. However it requires a reward function for the appropriate evaluation of action taken by the agent. In this paper we propose an approach to obtain comfortable driving by training with the reward function estimated from trajectories of comfortable driving, using deep inverse reinforcement learning. Experimentally we used low lateral acceleration as the condition of comfortable driving, and we were able to estimate a reward function with satisfying it.

1. はじめに

自動運転の実用化においては、事故を起こさず安全であることに加えて、搭乗者にとって不快感を与えない走行が要求される。これを「安心走行」と定義する。安心走行は、状況に応じた制御を逐一記述するようなルールベース手法での実現が困難である。これに対して、深層強化学習の導入が検討されている。深層強化学習は、行動を評価するスカラー値の報酬に基づき、試行錯誤を繰り返すことで制御則を獲得する手法である。一般的に、深層強化学習における報酬は設計者によって定義される。しかし、安心走行の定量的な表現は困難であり、設計者の手になる報酬では安心走行を実現できる保証はない。そこで本論文では、深層逆強化学習によって安心走行の軌跡から推定された報酬を用いることで、自動運転における安心走行の実現を目的とする。

2. 対象問題

本論文で対象とする問題は、深層強化学習による安心走行の学習である。深層強化学習手法には Deep Deterministic Policy Gradient (DDPG) [Lillicrap 15] を用いる。シミュレーション環境には、オープンソースのドライビングシミュレータである TORCS (The Open Racing Car Simulator) [Wymann 00] を用いる。本論文では、TORCS に用意されている、図 1 のコースを用いた 600m の直進走行を扱う。

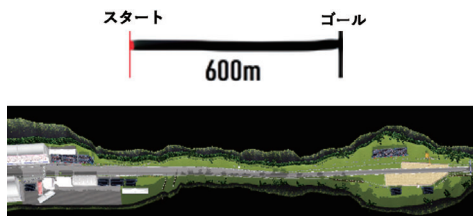


図 1: 実験で利用したコース

3. 準備

3.1 線形可解 MDP

状態 $s \in S$ の価値を $V(s)$ 、状態 s で行動 $a \in A$ を実行して受けるコストを $\ell(s, a)$ 、次の状態 s' への状態遷移確率を $p(s'|s, a)$ 、割引率を γ ($0 < \gamma \leq 1$) と定義するとき、マルコフ決定過程 (MDP) におけるベルマン方程式は式 (1) のように表される。なお、コストは報酬の符号を反転させたものであり、低いほど望ましい値となる。式 (1) は \min 演算子を有するために解析的にしか解くことができない。

$$V(s) = \min_a \left\{ \ell(s, a) + \sum_{s'} p(s'|s, a) \gamma V(s') \right\} \quad (1)$$

線形可解 MDP (LMDP) は、Todorov[Todorov 07] が提案した MDP の一種である。LMDP では、MDP における行動 a を実数値ベクトルである制御 $\mathbf{u} \in \mathbb{R}^{|S|}$ で定義する。状態 s における \mathbf{u} を $\mathbf{u}_s = u(s'|s)$ とするとき、LMDP では式 (2)、式 (3) に示す二つの仮定をおく。

$$p(s'|s, \mathbf{u}_s) = \bar{p}(s'|s) \exp(\mathbf{u}_s) \quad (2)$$

$$\ell(s, \mathbf{u}_s) = q(s) + D_{\text{KL}}(p(s'|s, \mathbf{u}_s) \parallel \bar{p}(s'|s)) \quad (3)$$

ここで、 $\bar{p}(s'|s)$ は制御 \mathbf{u}_s に依存しない s から s' への状態遷移確率、 $q(s)$ は状態依存のコスト、 $D_{\text{KL}}(\cdot \parallel \cdot)$ はカルバック・ライブラーダイバージェンスを表す。これらの仮定を式 (1) に適用し、ラグランジュの未定乗数法を用いて最小化問題を解くことで、式 (4) の線形ベルマン方程式が得られる。

$$\exp(-V(s)) = \exp\{-q(s)\} \sum_{s'} \bar{p}(s'|s) \exp\{-\gamma V(s')\} \quad (4)$$

このときの最適状態遷移確率 $p^*(s'|s)$ を式 (5) に示す。

$$p^*(s'|s) = \frac{\bar{p}(s'|s) \exp\{-\gamma V(s')\}}{\sum_{s'} \bar{p}(s'|s) \exp\{-\gamma V(s')\}} \quad (5)$$

連絡先: 岸川大航, 荒井幸代, 千葉大学大学院 融合理工学府 都市環境システム, 千葉市稲毛区弥生町 1-33, sachiyo@faculty.chiba-u.jp

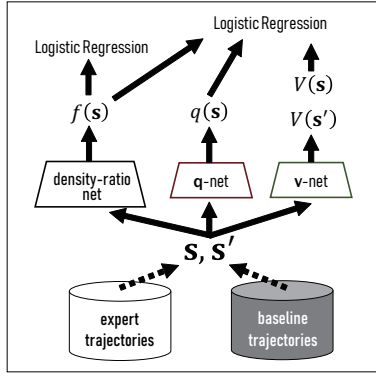


図 2: LogReg-IRL による学習

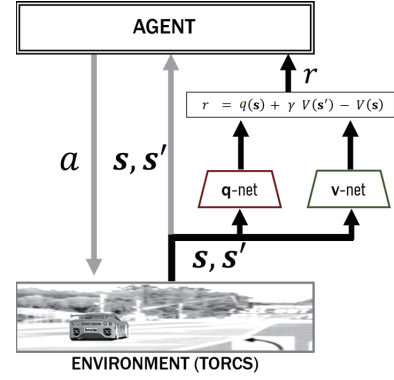


図 3: シェーピング報酬を利用した運転行動の学習

3.2 LogReg-IRL

Logistic Regression-Based IRL (LogReg-IRL) [Uchibe 17] は, LMDP における深層逆強化学習の一手法である. 式 (4) を式 (5) を用いて変形すると, 次の式 (6), 式 (7) を得る. ここで, \tilde{N} は安心走行の条件を満たさないベースライン軌跡のデータ数, N^* は条件を満たすエキスパート軌跡のデータ数である.

$$\log \frac{p^*(s)}{\bar{p}(s)} = f(s) + \log \frac{\tilde{N}}{N^*} \quad (6)$$

$$\log \frac{p^*(s, s')}{\bar{p}(s, s')} = f(s) - q(s) - \gamma V(s') + V(s) + \log \frac{\tilde{N}}{N^*} \quad (7)$$

LogReg-IRL では, ロジスティック回帰 [Cox 58] を用いた密度比推定手法である LogReg[Bickel 07] を利用して, ニューラルネットワークにより $f(s)$, $q(s)$ および $V(s)$ を学習する.

4. 提案手法

本論文では, LogReg-IRL を用いて, エキスパート軌跡, 及びベースライン軌跡から報酬を推定し, 深層強化学習によって安心走行を獲得する手法を提案する. まず, LogReg-IRL を用いて, エキスパート軌跡の集合およびベースライン軌跡の集合から状態遷移対 (s, s') を抽出し, 状態依存のコスト $q(s)$ と状態価値関数 $V(s)$ を推定する. LogReg-IRL における学習の概略図を図 2 に示す.

次に, 学習したネットワークによって報酬を計算し, DDPG により安心走行を獲得する. 推定された報酬を利用した学習の概略図を図 3 に示す. 報酬として, 式 (8) のシェーピング報酬を用いる. LogReg-IRL ではコストと状態価値関数を同時に推定するため, シェーピング報酬の利用により学習性能が向上することが報告されている [Uchibe 17].

$$r(s) = q(s) + \gamma V(s') - V(s) \quad (8)$$

5. 計算機実験

5.1 実験設定

本実験で用いる状態入力, TORCS に用意されている 79 次元のうち, 表 1 に示す, 運転行動に関係すると考えられる 31 次元とする. また, 本実験では直進走行を扱うため, 学習による制御対象を舵角 steering の 1 次元とし, 加速およびギア操作はルールベースによって制御した.

安心走行の条件は, 横加速度が常に 0.3G 未満であることとする. ここで, 横加速度とは, 車両の進行方向を X 軸方向と

表 1: 使用する TORCS の状態入力

項目名	値域	単位	説明
angle	$[-\pi, +\pi]$	[rad]	車両の角度
gear	-1, 0, 1, 2, 3, 4, 5, 6	-	現在のギア
rpm	$[0, +\infty)$	[rpm]	エンジンの回転数
speedX	$[-\infty, +\infty]$	[km/h]	X 方向速度
speedY	$[-\infty, +\infty]$	[km/h]	Y 方向速度
speedZ	$[-\infty, +\infty]$	[km/h]	Z 方向速度
track1-19	$[0, 200]$	[m]	壁までの距離 (図 4)
trackPos	$[-1, +1]$	-	道路中央からの変位
wheelSpinVel	$[0, +\infty)$	[rad/s]	四輪の回転速度
z	$[-\infty, +\infty]$	[m]	車両重心の変位

定義したとき, Y 軸方向に発生する加速度のことである. これは, 走行中の横加速度が 0.3G を超えると, 搭乗者は不快感や恐怖心を感じるとされる [茄子川 99] ことによる.

5.2 エキスパート軌跡・ベースライン軌跡の生成

まず, 先行研究 [北村 18] の設計報酬を用いた DDPG による運転行動の学習を行った. エキスパート軌跡, ベースライン軌跡は, 安心走行の条件を満たすモデルのうち最大横加速度の小さい上位 10 モデルを利用し, 各モデルの行動出力に ± 0.01 , ± 0.2 の乱数をそれぞれ加え 100 試行することで生成した. エキスパート軌跡とベースライン軌跡における横加速度推移の一例を図 5, 図 6 に示す. 横軸は総走行距離, 縦軸は横加速度を表す. なお, 図中の赤線は 0.3G の基準線である.

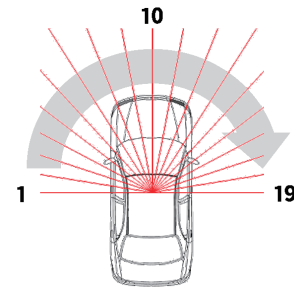


図 4: track (19 次元)

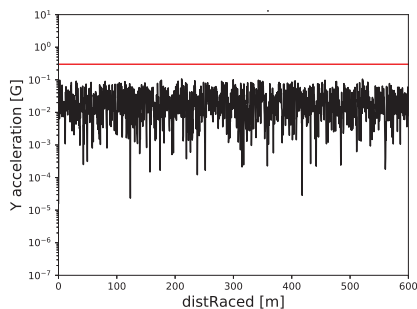


図 5: エキスパート軌跡の横加速度推移

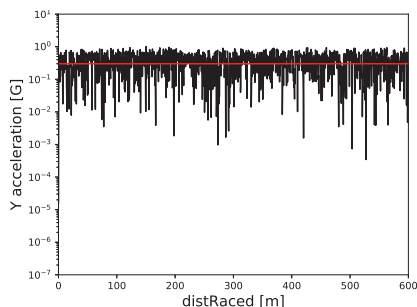


図 6: ベースライン軌跡の横加速度推移

5.3 推定された報酬による安心走行獲得

次に、生成されたエキスパート軌跡、ベースライン軌跡を用いた LogReg-IRL の各ネットワーク $f(s)$, $q(s)$, $V(s)$ の学習を行い、LogReg-IRL により推定された報酬を用いて DDPG による運転行動の学習を行った。推定された報酬によって得られたモデルの横加速度推移を図 7 に示す。横軸は総走行距離、縦軸は横加速度を表す。図中の赤線は 0.3G の基準線である。得られたモデルは走行時の最大横加速度が 0.01G となり、安心走行の条件を満たした。ゆえに、安心走行を獲得できる報酬が推定できたといえる。

6. 考察

6.1 LogReg-IRL における分類器の評価

まず、LogReg-IRL において、報酬・状態価値の推定に用いられる密度比推定ネットワーク $f(s)$ がどのように学習された

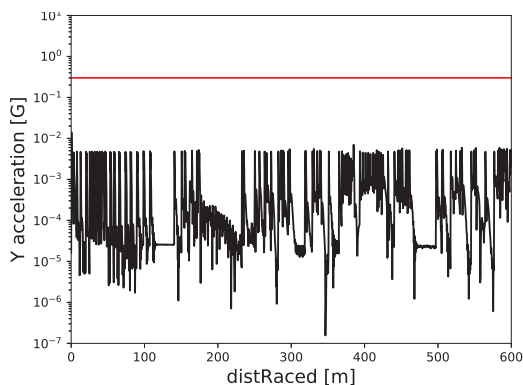


図 7: 推定された報酬によるモデルの横加速度推移

表 2: 混同行列

		正解		合計
		エキスパート	ベースライン	
予測	エキスパート	80131	444	80575
	ベースライン	52474	141963	194437
合計		132605	142407	275012

表 3: 各指標の確率値

指標	確率
正確度	80.76%
再現率	60.43%
適合率	99.45%
特異度	73.01%
F 値	75.18%
偽ベースライン率	27.00%
偽エキスパート率	0.55%

のかを調べた。5.2 節の方法によって生成し、LogReg-IRL の学習に利用しなかったエキスパート軌跡・ベースライン軌跡のデータを用いて、 $f(s)$ の出力を用いたロジスティック回帰による分類を行った。結果を表 2 に示す。

分類器の性能評価を、次に示す七指標を用いて行う。

- **正確度 (Accuracy)** P_{ac} ... エキスパートをエキスパート、ベースラインをベースラインと分類できた割合
- **再現率 (Recall)** P_{re} ... 実際のエキスパートをエキスパートと分類する確率
- **適合率 (Precision)** P_{pr} ... エキスパートと分類したデータのうち、実際にエキスパートである確率
- **特異度 (Specificity)** P_{sp} ... ベースラインと分類したデータのうち、実際にベースラインである確率
- **F 値 (F-score)** f ... 再現率と適合率の調和平均
- **偽ベースライン率** $P_{\bar{b}}$... エキスパートをベースラインと分類する確率
- **偽エキスパート率** $P_{\bar{\pi}}$... ベースラインをエキスパートと分類する確率

正確度、再現率、適合率、特異度、F 値は高いほど良く、偽ベースライン率、偽エキスパート率は低いほど良い。表 2 の混同行列を用いて計算した結果を表 3 に示す。

適合率と再現率はトレードオフの関係にあるが、今回の場合再現率に比べて適合率が極めて高く、偽エキスパート率が極めて低いため、ベースラインをエキスパートと誤って分類することを確実に回避するように学習しているといえる。

6.2 推定された $q(s)$ の勾配の分析

次に、式 (8) において中心となる状態依存のコスト $q(s)$ について、状態入力 s の微小変化に対する $q(s)$ の変化量である勾配 $\partial q(s)/\partial s$ を計算し、その分布を調べた。勾配の値が大きいほど、状態入力の変化に対する報酬の変化も大きくなり、報酬により強い影響を与えられられるためである。結果を図 8 に示す。

Z 方向の速度である $speedZ$ 、正面方向の車両とコース端との距離である $track10$ には正の勾配が計算される傾向がある。

正面方向のコース端との距離は開くことが望ましく、コースアウトの可能性を下げるができる。speedZ は、コースアウトしたときに発生する Z 方向速度の減速を避けることでコースアウトの可能性を下げると考えられる。

また、正面方向に対して 10 度左方向のコース端との距離 track9, 正面方向に対して右側の track11~19 に負の勾配が計算される傾向がある。これは、右側のコース端との距離から、車両の位置による報酬を推定しているためと考えられる。実験で利用したコースはわずかに逆 S 字状のカーブを有するコースであり、操舵しないと 600m 区間の前半でコース左側の壁に衝突してしまうため、コース右側に寄るような報酬が推定されていると考えられる。

7. まとめ

本論文では、深層強化学習において、安心走行を獲得可能な報酬の定義が困難であることに着目した。そこで、深層逆強化学習で推定した報酬を用いることによって安心走行を獲得する手法を提案した。計算機実験の結果、推定された報酬によって安心走行を獲得することができた。

実験においては安心走行の基準を 0.3G 未満の最大横加速度としたが、この基準は、基準を満たす軌跡を用意できるならば、任意に設定することができる。さらに、推定された状態依存のコスト $q(s)$ の勾配を分析することで、報酬の計算時に影響を与えている状態入力を明らかにすることができる。

LogReg-IRL では、通常の逆強化学習とは異なり、ある基準を満たすエキスパート軌跡に加えて、基準を満たさないベースライン軌跡が必要となるが、学習性能はベースライン軌跡に大きく左右され、大きな課題となっている。今後の課題として、ベースライン軌跡を必要としないような学習手法について検討する予定である。

参考文献

- [Lillicrap 15] Lillicrap, T. P., et al. (2015) "Continuous control with deep reinforcement learning." arXiv preprint arXiv:1509.02971.
- [Wymann 00] Wymann, B., Espie, E., Guionneau, C., Dimitrakakis, C., Coulom, R., and Sumner, A. (2000) "TORCS, the open racing car simulator." <http://torcs.sourceforge.net>
- [茄子川 99] 茄子川捷久, 宮下義孝, 汐川満則. (1999) "三訂 自動車の走行性能と試験法." 山海堂, p.54.
- [Todorov 07] Todorov, E. (2007) "Linearly-solvable Markov decision problems." Advances in neural information processing systems.
- [Uchibe 17] Uchibe, E. (2017) "Model-Free Deep Inverse Reinforcement Learning by Logistic Regression." Neural Processing Letters.
- [Cox 58] Cox, D. R. (1958) "The Regression Analysis of Binary Sequences." Journal of the Royal Statistical Society. Series B (Methodological), 215-242.
- [Bickel 07] Bickel, S., Brückner, M., Scheffer, T. (2007) "Discriminative Learning for Differing Training and

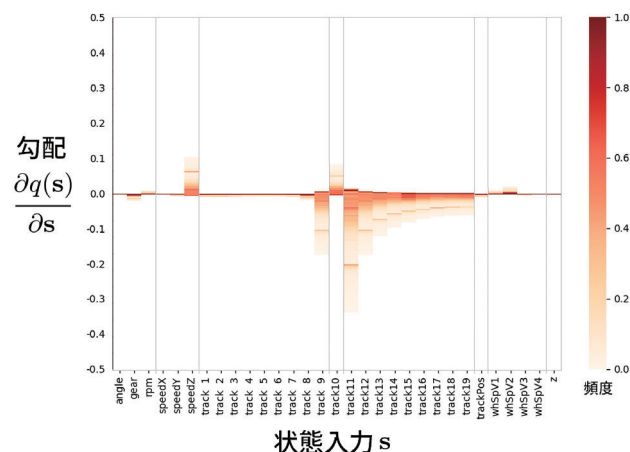


図 8: 状態入力 s に対する $q(s)$ の勾配の分布

Test Distributions." In Proceedings of the 24th international conference on Machine learning (pp. 81-88). ACM.

[北村 18] 北村清也, 石川翔太, 荒井幸代. (2018) "市街地の自動運転における環境情報のフィルタリング." 2018 年度人工知能学会全国大会.

認知的満足化関数の線形関数近似 － 文脈付きバンディット問題への対応 －

Linear function approximation of Cognitive Satisficing Function － To Cope with Contextual-bandit Problem －

甲野 佑 ^{*1*2}

Kono, Yu

^{*1} 東京電機大学理工学部

School of Science and Engineering, Tokyo Denki University

^{*2} 株式会社ディー・エヌ・エー

DeNA, Co., Ltd.

Both recommendation and foraging behavior of animals are aiming to maximizing rewards through trial and error. By contrast, Maximizing reward is difficult in a complex actual world that is extremely complicated. So, The decision-making agents is considered to give priority to whether or not to achieve a specific purpose. In addition, they aim to achieve the desire level with as little information as possible. The decision-making tendency where is owned intelligent lives is called “satisficing”. The RS algorithm to make choices for “satisficing” was focused in this paper, further LinRS adapted to linear approximation function was devised so that the scope of the problem is expanded to be more widely applicable. In consequence, RS became enabled to cope with the contextual-bandit problem where has application examples such as advertisement delivery. Moreover LinRS compared with familiar existing selection algorithms in simulation. The linear function approximation of LinRS realized in this study is the first step to apply a fast and efficient search algorithm by using RS that emphasizes achievement of purpose to deep reinforcement learning.

1. はじめに

昨今の深層学習の発展が示す通り、現代の膨大なデータを扱う際には関数近似が必要不可欠である。しかしながら、手法によるものの近似関数の学習には多くのデータが必要となる。機械学習とは一般にそのようなデータからの自律的な近似関数のパラメータの最適化と考えられているが、一方で機械学習の中でもデータを自ら主体的に環境に働きかけて取得しなければならない強化学習も存在し、しばしば特異な存在と見なされている [Sutton 98]。強化学習は報酬を最大化することを目的に、それを達成する行動手順、方策の獲得を目的としている。つまり強化学習は様々な状態に対して適切な行動確率をエル必要がある。そのためには様々な状態遷移を通して行動手順を試していく必要があり、関数近似を使わないテーブル形式の学習でさえ、膨大な繰り返しの反復試行を必要とする。一般に状態空間は膨大で複雑であるため、その方策関数や価値関数は複雑な関数による近似が求められる。しかしただでさえ膨大な反復試行を必要とする上に関数近似を行うとなると、さらに“強化学習特有の試行錯誤の回数”×“関数近似のための学習データ取得のためのサンプリング回数”という掛け合わせにより、膨大かつ甚大な試行時間、データ数が必要になる [Mnih 15]。現実の環境の複雑さを考えると、関数近似を用いた強化学習を現実的な時間内で行うためには、そのどちらかの高速化、軽量化が必要となる。

近年の研究は、概ね学習やデータサンプリングを複数の計算機に分散化することにより解決しようとしている。本研究ではそれとは対照的に、“強化学習特有の試行錯誤の回数”側を減らすことを目的としている。人間をはじめ動物は強化学習によって利得行動を獲得するとされている [高橋 16]。しかし人間は既存の強化学習アルゴリズムのような異常な回数の繰り返し行動は行わない。それは知能的な構造化にも依存するが、そもそもの意思決定傾向に違いがあると思われる。そこで本研

究では人間の最適性より目標達成を重視した認知的満足化価値関数 (Risk-sensitive Satisficing Value Function: RS) を用いた選択アルゴリズム (RS アルゴリズム) [高橋 16] に着目した。しかしながら RS の関数近似手法は提案されていない。そこで本研究ではまず RS の線型近似を行い、そのアルゴリズムとして LinRS (Linear RS) を考案した。また、基礎的な強化学習課題の一つであり、広告配信などに応用される文脈付きバンディット問題のシミュレーションにおいて LinRS を既存アルゴリズムと比較し、無限の試行回数が許されていない場合での有用性を示す。

2. 特徴ベクトル上での試行錯誤：文脈付きバンディット問題

本研究において、現在の状態 (文脈) に関する特徴ベクトルが明確に与えられることを前提としたワンショットの意思決定課題として、文脈付きバンディット問題を扱う。通常のも腕バンディット問題ではエージェントが、未知の報酬確率 $\{p_1, p_2, \dots, p_k\}$ が割り当てられた行動 $\{a_1, a_2, \dots, a_k\}$ の中から毎回一つ選択し、試行錯誤しながら得られる報酬を最大化することを目的とした最も単純な強化学習課題である。与えられる報酬が 1 か 0 のベルヌーイ試行からなるものをベルヌーイバンディットと呼び、本研究ではこれを扱う。現実には選んだ広告 (行動) がクリック (報酬) されるか否かという広告配信などの応用例と対応づけられる。しかし広告配信の相手は常に同じ人物とは限らない。その人物の属性 (性別, 年齢層, その瞬間広告が乗るページの属性など) がわかる場合、それに応じた配信を行うのが適切だと考えられる。この属性に応じた配信は文脈付きバンディット問題と呼ばれる問題に相当する。文脈付きバンディット問題では“どの広告を配信するか”に該当する行動集合 $\{a_1, a_2, \dots, a_k\}$ は通常バンディット問題と同じく存在する。対してその腕それぞれの行動の真の報酬関数は、その時訪れる人物 (状態 s) に応じた値として与えられる。その人物の属性は各次元に性別, 年齢層, などの特徴量が与えられた d 次元の特徴ベクトル $x_{s,a}$ によって表現される。

連絡先: 甲野 佑, 東京電機大学理工学部, 埼玉県比企郡鳩山町大字石坂, yu.kono.02@gmail.com

文脈付きバンディット問題に限らず、多数の選択肢を持つバンディット問題では最適な手段を知るために、現状は非最適な行動をあえて探索的に試行する必要がある。しかし前述の通り、高い累積報酬を得る（活用）ためにはどこかで探索を打ち切らなければならないという速さと正確さのトレードオフを端的に表した課題であり、探索と活用のバランスが問題となる。バンディット問題において、このバランスの良し悪しは、最適な行動が常に必ず明らかであり、それを撮り続ける場合の累積報酬の期待値と、そのアルゴリズムの選択実績としての累積報酬の期待値との差分からなる regret で評価される。この regret が小さいほど、無駄なく報酬を最大化できたことを意味し、Upper Confidence Bound (UCB) 系のアルゴリズムや Thompson Sampling (TS) [Agrawal 12] などの regret をなるべく小さくすることを目的としたアルゴリズムが知られている。文脈付きバンディット問題では、LinUCB[Li 10] や LinTS[Agrawal 13] など、通常のバンディット問題で扱われていたアルゴリズムを、特徴ベクトルに対する推定をパラメータを学習するアルゴリズムとして線形に拡張して用いられている。

3. 認知的満足化関数と RS アルゴリズム

環境への試行錯誤から良い行動手順を見つけるためには、時にそれまでの知識を活用し、有効であるとされてきた行動の選択肢以外の選択肢を探索してみる必要がある。無数に存在する手順から最適なものを見つけ出すには全て試す必要があり、組み合わせや、環境から得られる情報の不確実性からさらにその探索回数は増す。

それに対して、人間は必ずしも最適な行動手順を求めるわけではない。ある目的の達成か否か、希求水準を満たすことが重要であり、一度見つければ基本的にはその行動を取り続ける。このような探索における意思決定における傾向を満足化と呼ぶ。そのような希求水準を満たした・満たしていないことで価値評価が反転する、認知的な価値関数として、満足化価値関数 (Risk-sensitive Satisficing Value Function: RS) が考案されている [高橋 16]。

$$e_a^{\text{RS}} = n_a \delta_a = n_a (e_a - \aleph) \quad (1)$$

価値関数という名称だが、その性質は選択時の各行動への評価関数である。ここで n_a は行動 a を試行した回数であり、 e_a は行動 a によって得た報酬の平均である（観測期待値）。またその瞬間 s に実行可能な行動の集合 \mathbb{A}_s の中から e_a^{RS} の評価で最大の行動 a を選択する意思決定アルゴリズムを RS アルゴリズム（以下、断りがなければ RS は RS アルゴリズムのことを示す）と呼ぶ。RS は非満足状況 ($\max_a e_a < \aleph$) であれば楽観的探索を行う。すなわち試行回数 n_a が少ない方が評価が高くなることで探索を促す。一方、満足状態 ($\max_a e_a > \aleph$) は悲観的利益追求を行う。これは試行回数 n_a が大きいほど評価を高く見積もり、試行回数がポジティブフィードバック的に増え続ける活用がなされる。

3.1 基準値と探索

基準値 \aleph を最大の報酬確率 p_{first} とその次に大きい報酬確率 p_{second} の間に設定することで満足化は最適化となる。そのため、満足化基準値 \aleph が以下のように設定された場合、最適切基準 \aleph_{opt} と呼び、 \aleph_{opt} を用いた RS アルゴリズムを RS-OPT アルゴリズムと呼ぶ。

$$\aleph_{\text{opt}} = \frac{p_{\text{first}} + p_{\text{second}}}{2} \quad (2)$$

しかし、 \aleph_{opt} は報酬確率が既知のものとしているため活用は困難である。そこで、以下のように初期値 \aleph_0 から基準値を更新することで式 4 の条件を満たすことを期待する。

$$\aleph = \aleph + \alpha(e_a - \aleph) \quad (3)$$

$$p_{\text{first}} < \aleph < p_{\text{second}} \quad (4)$$

式 3 によって求められる基準値 \aleph は $a_{\text{select}} = a_{\text{first}}$ であるならば p_{second} を上回るため式 4 を満たすことができる。しかし、式 3 が式 4 を満たす保証はない。

またベルヌーイバンディット問題において動的な基準値の際適切に向けた更新式も考案されており [甲野 18]、その場合 TS とほぼ互角の regret になることがシミュレーションによって示されている。また、同じく参考文献 [甲野 18] で、基準値 $\aleph = \infty$ である時に全ての選択肢をほぼ同配分で選択することも知られている。すなわち $\aleph > \max_a E$ である時は基準値 \aleph は softmax 方策における温度パラメータのような役割を持ち、高ければ高いほどランダムに近い探索を行う。

4. 提案アルゴリズム : LinRS

本研究では通常のバンディット問題のアルゴリズムである RS を UCB や TS と同じく線形近似関数に拡張した。LinUCB や LinTS と同じく Liner RS (LinRS) と呼ぶことにする。

4.1 パラメータとその更新

LinRS の計算に必要な変数として、行動 a ごとに $d \times d$ の行列で表される A_a 、 d 次元のベクトルである b_a 、 m_a を定義する。ベクトル $x_{s,a}$ は各 step で与えられる特徴ベクトルを意味し、変数 $r_{s,a}$ は各 step で実際に行動 a を選択して与えられた 0 or 1 の報酬値を表す。変数 $u_{s,a}$ は各 step においてどの選択肢を選択したかを意味し、選択した行動 a に関して 1 になる。

$$A_a = I + \sum_{s=1}^t x_{s,a} x_{s,a}^T \quad (5)$$

$$b_a = \sum_{s=1}^t x_{s,a} r_{s,a} \quad (6)$$

$$m_a = \sum_{s=1}^t x_{s,a} u_{s,a}, \quad u_{s,a} = \begin{cases} 1 & (a = a_{\text{select}}) \\ 0 & (\text{otherwise}) \end{cases} \quad (7)$$

すなわち、毎ターンインクリメントされる形で変数は更新される。変数ベクトル b_a 、 m_a の各次元の初期値は全て 0 とした。変数行列 A_a の初期値は単位行列 I とした。

4.2 評価関数

変数ベクトル b_a 、変数行列 A_a を用いて特徴ベクトル $x_{s,a}$ が与えられた時の報酬期待値の不偏推定量 $\hat{\theta}_a^T x_{s,a}$ が定義される。また、新たに類似度に基づいた試行回数の擬似的な量、擬似試行回数を変数ベクトル m_a 、変数行列 A_a を用いて $\hat{\phi}_a^T x_{s,a}$ として定義する。これらの推定量と基準値 \aleph_s を用いて、通常のバンディットでの評価関数である式 1 と同じ構成で、線形関数に拡張された RS の評価関数 f_a^{RS} (式 10) を定義する。

$$\begin{aligned} \hat{\theta}_a &= \left(\sum_{s=1}^t x_{s,a} x_{s,a}^T \right)^{-1} \sum_{s=1}^t r_{s,a} x_{s,a} \\ &= A_a^{-1} b_a \end{aligned} \quad (8)$$

$$\begin{aligned} \hat{\phi}_a &= \left(\sum_{s=1}^t x_{s,a} x_{s,a}^T \right)^{-1} \sum_{s=1}^t n_{s,a} x_{s,a} \\ &= A_a^{-1} m_a \end{aligned} \quad (9)$$

$$f_a^{\text{RS}} = \hat{\phi}_a^T x_{s,a} (\hat{\theta}_a^T x_{s,a} - \aleph_s) \quad (10)$$

LinRS では各 step において評価関数 f_a^{RS} がもっとも高い行動 a を選択する.

$$a^{\text{select}} = \arg \max_a \hat{\phi}_a(f_a^{\text{RS}}) \quad (11)$$

4.3 3種類の基準値

LinRS も RS アルゴリズムと同じく \aleph_{opt} の値が与えられれば最適な選択を行うと考えられる. しかし, 定常的な多本腕バンディット問題と異なり, 毎 step 異なる特徴ベクトルが与えられた上で異なる真の報酬確率の上で選択する文脈付きバンディット問題では, 最適な基準値が与えられることは適切ではないと考えられる. そこであらかじめ固定したパターンと, 式 3 のように更新するパターン 2 種を考案した.

4.3.1 固定基準値

基準値 \aleph_s は特徴量, step による更新はなく, 常に固定.

4.3.2 動的即時更新基準値

通常バンディット問題の基準値の動的な更新式 3 に基づき, 毎 step 意思決定毎に更新する式を以下とする.

$$\hat{\aleph}_s \leftarrow (1 - \alpha)\aleph_s + \alpha \max_a (\hat{\theta}_a^T x_{s,a}) \quad (12)$$

この基準値を用いた LinRS を本研究では Adaptive LinRS (AdaptiveLinRS) と呼ぶ.

4.3.3 動的な特徴ベクトルに応じた更新基準値

動的に意思決定毎に即時的に基準値を更新 (式 12) する場合, 特徴ベクトル $x_{s,a}$ の違いに対処できない. 特徴ベクトル毎での学習を行うため, A_a , b_a を参考に, 変数行列 C と変数ベクトル d を定義し, 基準値を \aleph_s を以下に定義する. パラメータ α は式 12 と同じく, 新規な情報に対して重み付けする更新パラメータである. 通常バンディット問題における式 3 や即時的な更新式 3 と同じく, 学習率 α が高ければ高いほど近視眼的な基準 (1 step 前の観測期待値) を超えるような選択肢を探すようになる.

$$C \leftarrow (1 - \alpha)C + \alpha x_{s,a} x_{s,a}^T \quad (13)$$

$$d \leftarrow (1 - \alpha)d + \alpha \max_a \hat{\theta}_a \quad (14)$$

$$\hat{\aleph}_s = (C^{-1}d)^T x_{s,a} \quad (15)$$

この基準値を用いた LinRS を本研究では Adaptive State LinRS (略記して StateLinRS) と呼ぶ.

5. シミュレーション設定

本研究では簡易的なシミュレーションを github にて公開されていた python コードに追加して実験を行った [contextual bandit original source]. 毎 step 与えられる特徴ベクトルの次元数は 16, また選択肢の数を 16 とした. シミュレーションでは選択アルゴリズムに従い, 選択を総 step 数として 10,000 steps 行い, その 1,000 シミュレーション回分を平均して regret を算出した. 選択肢の真の報酬生起確率は下記に示される特徴ベクトルと報酬分布と毎 step サンプルし直されるノイズ変数から決まるよう定義した.

5.1 特徴ベクトルの分布と報酬関数

毎 step の特徴ベクトル x は step ごとに各次元が 0, 1 の一様乱数から生成されるテーブルデータとして与えられた. バ

ンディットパラメータ η は, 平均 $\mu = 0$, 共分散は対角線のみ $\sigma = 0.01$ の対角行列として定義した.

$$\eta = \sigma \times I \quad (16)$$

それぞれの腕の報酬確率 $p_{s,s}$ はバンディットパラメータ η と x と, 毎 step サンプルされる平均 0 かつ分散 0.1 の正規分布 $\mathcal{N}(0, 0.1)$ のノイズであるスカラー変数 ϵ を使い, 以下の式で定義した.

$$p_{s,a} = \text{sigmoid}(x_{s,a}^T \eta + \epsilon) \quad (17)$$

5.2 評価指標 regret

本シミュレーションの評価指標には regret (式 18) を用いた. Regret は各 step でもっとも良い選択肢を選択肢続けた際に得られる累積報酬の期待値との差を意味する. 低ければ低いほど無駄な探索が抑えられていることを意味し, また徐々に対数的に収束することで, より良い選択肢を選んでいることを意味する. そのため regret が低いほど活用と探索のバランスが良くできている良いアルゴリズムとされる.

$$\text{regret} = \sum_{s=1}^t (\max_a p_{s,a} - p_{s,a^{\text{select}}}) \quad (18)$$

5.3 比較アルゴリズム

本研究では簡易的な検証のため, LinRS と一般的な文脈付きバンディット問題のアルゴリズムである LinUCB, LinTS と比較を行った. また, LinRS の 3 種の基準値 \aleph についても比較を行った.

ϵ -greedy

確率 ϵ でランダム探索し, 確率 $1 - \epsilon$ で報酬期待値の不偏推定量に対して最大の行動を選択 (greedy 選択) する. 本シミュレーションではパラメータ $\epsilon = 0.1$ とする.

LinUCB

線形関数に拡張された UCB アルゴリズムで, 分散の不偏推定量の α 倍を期待値の不偏推定量に上乗せした評価関数を用いて選択を行う. 本シミュレーションではパラメータ $\alpha = 0.1$ とする.

LinTS

通常の TS を線形関数に拡張したアルゴリズムであり, TS が β 分布から評価値をサンプリングしていたのと異なり, 多変量正規事後分布からサンプリングしている. 前述に記した LinRS, LinUCB と同じ定義の変数 A_a と b_a を持つ. サンプリングに用いる多変量正規事後分布の共分散行列, 変数 A_a をパラメータ σ^2 倍したものとなっている. 本シミュレーションではパラメータ $\sigma = 0.1$ とする.

LinRS (固定基準値)

基準値 $\aleph = \{0.8, 1.0, 1.2, 1.6, 2.0\}$ でシミュレーションを行った. ベルヌーイバンディットでありながら基準値を 1.0 以上のものも用意したのは, 前述した通常の RS の性質から, $\aleph > \max_a \hat{\theta}_a$ である場合, 基準値 \aleph が高いほど探索が促進されるためであり, その性質を確認するためである.

AdaptiveLinRS

動的な更新基準値の学習率パラメータ $\alpha = 0.1$ とした.

StateLinRS

基準値推定のための変数の動的な更新のための学習率パラメータ $\alpha = 0.1$ とした.

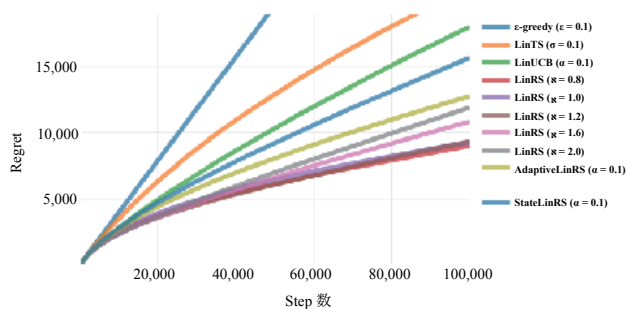


図 1: 16 本腕, 16 次元の特徴ベクトルでの regret の推移

5.4 結果および考察

図 1 は regret の推移を示している。10,000 step までの LinRS アルゴリズムは LinUCB, LinTS, ϵ -greedy より低い regret を有している。適当に設定した基準値 $\kappa = \{0.8, 1.0, 1.2\}$ においては十分に大きな回数の上で既存アルゴリズムに比べても低い regret を有している。より長期的な step 数での regret の保証は現時点では不可能だが、現実の環境とは常に非定常であり、特徴ベクトルからの不偏推定量では掴みきれない変化は存在する。その時、より早く少ない regret で探索が行える性質は有用だと考えられる。しかし達成したい目標として $\kappa = \{0.8, 1.0, 1.2\}$ のいずれでも大きな差のない regret であるという範囲は通常バンディット問題での RS は曖昧なパラメータであるとも考えられる。ただし、線型近似関数による推定量であるがゆえ、基準値が理想基準値を有していることが必ずしも正しいとは限らない。ゆえに $\max p_a$ より少し高い基準値でも十分低い regret を示したのだと思われる。

また、動的な更新についてだが、AdaptiveLinRS のような即時的な基準値の更新でも LinUCB や LinTS と比べても低い regret を示している。しかし、これは特徴ベクトルに対する真の報酬関数にも依存すると思われる。基準値が特徴ベクトルに依存しない以上、広告の例で言えば、まったく広告をクリックしない人間とそうでない人間が混在する状況ではあまり機能しないと思われる。そして StateLinRS だが、LinUCB と同等よりやや低い regret に止まっている。前述の基準値が固定な場合での考察から、高めの基準値の方が有効であると考えられる。StateLinRS では他の推定量と同じため、より報酬期待値に関する推定量の上限を意識した更新式にする必要があると思われる。

6. 結論

本研究は近似関数上での一般化を目指し、RS アルゴリズムの線形近似関数に拡張した。本研究は深層強化学習などに使われるニューラルネットワークなどの複雑な近似関数での利用の一時的な拡張であり、推定量に対して基準値を高めになるよう設計する必要があるなど、有用な知見を与えた。今後の検証でより長期的な step 数、より多い行動数、より複雑な特徴ベクトルと報酬関数の関係性、より最新のアルゴリズムとの比較を行う。その知見を通して、基準値 κ の動的な更新アルゴリズムや、複雑な近似関数への拡張を検討していく。本研究の固定的な基準値で低 regret を達成していることなどから、動的な基準値でも更新アルゴリズム次第では同等の成績を有することができうる可能性がある。この性質から、今後の拡張によって、深層強化学習上でも少ない反復回数で目的の累積報酬を達成することが期待される。

参考文献

- [Sutton 98] Sutton, R. and Barto, A.: Reinforcement Learning: an Introduction, *MIT Press* (1998).
- [Mnih 15] Mnih, V., Kavukcuoglu, K., Silver, D., Hassabis, D., et al.: Human-level control through deep reinforcement learning, *Nature*, 518(7540), 529-533 (2015).
- [Wang 05] S. Gelly, Y. Wang., R. Munos. and O. Teytaud.: Modification of UCT with Patterns in Monte-Carlo Go, *INRIA Technical Report*, 6062 (2005).
- [Agrawal 12] Agrawal, S., Navin Goyal, N.: Analysis of thompson sampling for the multi-armed bandit problem, *Proceedings of the 25th Annual Conference on Learning Theory (COLT)* (2012).
- [高橋 16] 高橋 達二, 甲野 佑, 浦上 大輔: 認知的満足化 限定合理性の強化学習における効用, *人工知能学会論文誌*, 31(6), AI30-M.1-11. (2016)
- [甲野 13] 甲野 佑, 高橋 達二, 価値推論ヒューリスティクスとしての規準学習と忘却, *Proceedings of 30 th Japanese Cognitive Science Society (JCSSS)*, 74-79. (2013)
- [Tamatsukuri 18] Tamatsukuri, A., Takahashi, T.: Guaranteed satisficing and finite regret: Analysis of a cognitive satisficing value function, arXiv:1812.05795 (2018; manuscript under revision).
- [甲野 18] 甲野 晃弘, 高橋 達二, 満足化を通じた最適な自律的探索, *2018 年度 人工知能学会全国大会 (第 32 回) 論文集* (2018).
- [本多 16] 本多 淳也, 中村 篤祥, バンディット問題の理論とアルゴリズム, 講談社 (2016).
- [Li 10] Li, L., Chu, W., John Langford, J., Schapire, R. E.: A contextual-bandit approach to personalized news article recommendation, *Proceedings of the 19th international conference on World wide web (WWW)*, 661-670 (2010).
- [Agrawal 13] Agrawal, S., Goyal, N.: Thompson sampling for contextual bandits with linear payoffs, *Proceedings of the 30th International Conference on Machine Learning (ICML)*, 127-135 (2013).
- [contextual bandit original source] Netflix も使っている! Contextual Bandit アルゴリズムを徹底解説!, https://github.com/smn-ailab/ysaito-qiita/tree/master/multi_armed_bandit

General Session | General Session | [GS] J-10 Vision, speech

[3N3-J-10] Vision, speech: voice and communication

Chair: Masanori Tsujikawa Reviewer: Jun Sugiura

Thu. Jun 6, 2019 1:50 PM - 2:30 PM Room N (Front-right room of 1F Exhibition hall)

[3N3-J-10-01] Multilingual Imputation Using Transfer Learning for Estimating Emotion from Speech

○Koichi Sakaguchi¹, Shohei Kato^{1,2} (1. Dept. of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology, 2. Frontier Research Institute for Information Science, Nagoya Institute of Technology)

1:50 PM - 2:10 PM

[3N3-J-10-02] Development of Open-source Multi-modal Interaction Platform for Social Experiment of Conversational User Interface

○Akinobu Lee¹ (1. Nagoya Institute of Technology, Japan)

2:10 PM - 2:30 PM

音声からの感情推定における転移学習を用いた多言語補填

Multilingual Imputation Using Transfer Learning for Estimating Emotion from Speech

坂口巧一 *¹ 加藤昇平 *^{1*2}
Koichi Sakaguchi Shohei Kato

*¹名古屋工業大学 大学院工学研究科 情報工学専攻

Dept. of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology

*²名古屋工業大学情報科学フロンティア研究院

Frontier Research Institute for Information Science, Nagoya Institute of Technology

Recently, vocal communication robots attract people thanks to development of AI and robot engineering. The technology of estimating emotion from speech is important to realize a smooth dialog between human and robots. This technology needs a large number of emotional speech data, but it is difficult to collect such data a lot. We investigated the effectiveness of multilingual imputation by transfer learning using 1D convolutional bidirectional LSTM. In this paper, we report the result. The result is suggested that increasing the number of languages of emotional speech learned may exceed the performance of the model learned insufficient emotional speech in single language.

1. はじめに

近年、ロボティクス技術と AI の発展に伴い、音声によって人と対話するロボットが注目を集めている。人は音声対話するときに言語情報だけでなく、声の抑揚などの非言語情報も考慮しながら対話相手の感情を推定する。そのため、ロボットが人と同様に音声で対話するには、そのような情報からも感情を推定できることが望まれる。音声から感情を推定するには大量の感情音声サンプルが必要である。しかし、感情音声を大量に集めることは難しい。そのため、サンプル不足を補う手法が必要と考えられる。

音声から感情を推定する研究は以前から行われている。かつては何らかの特徴抽出アルゴリズムによって複数の音声特徴量を抽出し、Support Vector Machine (SVM) などに学習させて判別する手法が多かった [有本 08]。近年では、ディープラーニングの台頭により、ニューラルネットワークに自発的に音声特徴を学習させて判別する研究も盛んに行われるようになってきた [Dario 16] [George 16]。しかし、複数の言語の音声の感情について推定を行ったり、異言語間の感情の共通性について考察している研究は少ない。

Ekman[Ekman 75] が「基本的な顔表情は文化によらず普遍的であること」を示したことから、音声にも文化によらず共通の特徴が存在することが期待できる。共通の特徴が存在するならば、サンプル数の少ないある言語の感情音声データを、別言語のデータで補填できるのではないかと考えた。本研究では、ディープラーニングにおける多言語補完の妥当性調査を試みた。

2. 提案手法

2.1 音声データの前処理

音声データは、スペクトログラムに変換して 1 次元畳み込み双方向 Long Short Term Memory (LSTM) モデルに入力される。スペクトログラムとは音声データに短時間フーリエ変

換 (STFT) を行い、各周波数成分強度の時間変化を表す 2 次元データである。STFT のサンプル数を 882、フレーム周期を 441 とし、時間長は 200 までとした。スペクトログラムのサイズは $N \times 200$ とした。周波数方向を N としたのは、次のセクションで示す実験で周波数帯域 5kHz 上限と全領域 (22.05kHz) の場合で比較を行うためである。5kHz 上限の場合は $N=101$ 、全領域の場合は $N=442$ となる。これに z-score 正規化を施したものを、1 タイムステップ毎に分割して 1 次元畳み込み双方向 LSTM モデルに入力する。なお、感情音声データは日本語、韓国語、アメリカ英語の感情音声の 5 感情「怒り」「悲しみ」「喜び」「嫌悪」「驚き」を使用した。

2.2 1 次元畳み込み双方向 LSTM

提案モデルである転移学習モデルの事前学習、および比較手法である単一言語の感情音声のみを学習するモデル (単一言語モデル) には 1 次元畳み込み双方向 LSTM を用いた。

1 次元畳み込み双方向 LSTM は、1 次元畳み込み部と双方向 LSTM、全結合層の 3 つからなる。1 次元畳み込み部は、1 次元畳み込み層とプーリング層の組み合わせからなる部分であり、特徴の鋭敏化と次元圧縮を行う。LSTM は、適切な過去の入力を保存することで、時間依存性の強いデータに対して効果を発揮する。一般的な LSTM は過去から未来への一方向の流れのみを考慮するが、双方向 LSTM は未来から過去への方角も考慮する。具体的なパラメータは表 1 のようになる。誤差関数は categorical crossentropy を用いた。最適化アルゴリズムには Nesterov accelerated gradient (NAG) [Nesterov 83] を用い、学習率を 0.01 とし、epoch ごとに学習率を $1e^{-6}$ ずつ減衰し、momentum 項のパラメータを 0.9 とした。

2.3 転移学習モデル

ある言語の感情音声データを別の言語の感情音声データで補填する手法として転移学習を利用した。このモデルは、特徴抽出部と判別部の 2 つからなる。

特徴抽出部には、1 次元畳み込み双方向 LSTM を用いる。各言語の感情音声で学習を行い、学習済モデル (日), (韓), (英) を作成する。データを 8:2 の割合で学習データとテストデータに分け、学習データのうち 2 割を検証用データとする学習を

表 1: 1 次元畳み込み双方向 LSTM 詳細

パラメータ設定	
入力層	入力サイズ:N×200 タイムステップ 上限 22.05kHz:N=442, 5kHz:N=101
畳み込み層 1	フィルタ:(4,1) × 16 活性化関数:ReLU, バッチ正規化あり
畳み込み層 2	畳み込み層 1 と同様
最大プーリング層 1	プールサイズ:2, スライド 2 ドロップアウト率:0.25
畳み込み層 3	畳み込み層 1 と同様
最大プーリング層 2	プールサイズ:2, スライド 2
平滑化層	タイムステップごとに平滑化
双方向 LSTM 層	出力次元:512×2 隠れ層のドロップアウト:0.5, 活性化関数:tanh
全結合層 1	出力ユニット数:100 活性化関数:ReLU, ドロップアウト率:0.25
全結合層 2	活性化関数:softmax l1l2 正則化:(0.01,0.01)
出力層	出力サイズ:(5,1)

表 2: 判別部パラメータ詳細

パラメータ設定	
入力層	入力サイズ:100 × N
全結合層 1	出力ユニット数:100 活性化関数:ReLU
全結合層 2	活性化関数:softmax l1l2 正則化:(0.01,0.01)
出力層	出力サイズ:(5,1)

行った。なお、学習エポック数は 100 回とした。学習終了後、学習済みモデルの最終層を取り除き、入力データを 100 次元の特徴を抽出する特徴抽出器 (日), (韓), および (英) を作成した。作成した特徴抽出器の出力を入力として、全結合層 2 層からなる判別部を学習する。なお、誤差関数と最適化関数については 1 次元畳み込み双方向 LSTM モデルと同様である。例として、図 1 に日本語の感情音声データを転移学習モデル (英韓) で学習する場合を示す。特徴抽出器 (英), (韓) に日本語感情音声を入力して 200 次元の特徴に変換し、その特徴を入力として学習をする。

3. 実験データ

実験に使用したデータ数一覧は図 3 のようになる。日本語、韓国語、北アメリカ英語の 3カ国の感情音声発話を用意し、5 感情「怒り」「悲しみ」「喜び」「嫌悪」「驚き」を実験データとして使用した。なお、サンプリング周波数は日本語と韓国語の感情音声データは CD 規格の 44.1kHz であるが、北アメリカ英語のみ 48kHz である。そのため、北アメリカ英語のデータにサンプリング周波数変換を施し、44.1kHz に変換して使用した。

3.1 日本語感情音声データ

日本語の感情音声データとして、感情評定値付きオンラインゲーム音声チャットコーパス (OGVC)[有本 13] を使用した。これはオンラインゲームの音声チャットの感情発話を、4 名の俳優 (男性 2 人, 女性 2 人) が 9 感情 (受容, 怒り, 期待, 嫌

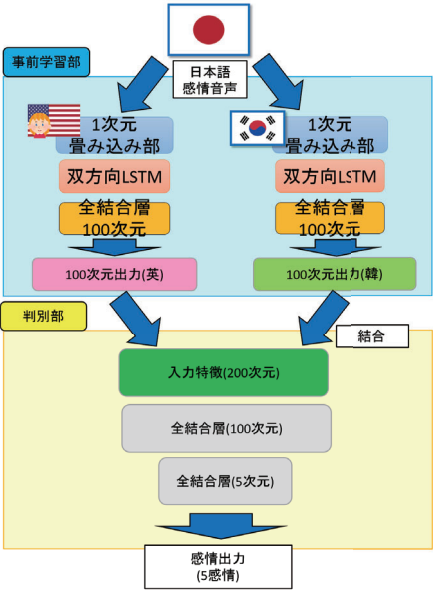


図 1: 転移学習モデル (英韓) 概要

表 3: 実験データ

	日本語	韓国語	北米英語
怒り	240	100	192
悲しみ	252	100	192
喜び	252	100	192
嫌悪	240	100	192
驚き	288	100	192
合計	1272	500	960

悪, 恐れ, 喜び, 悲しみ, 驚き, 平静) で演じた音声コーパスである。

3.2 韓国語感情音声データ

韓国語の感情音声データとして Cho ら [Cho 09] が用いた感情音声データを使用した。これは、韓国の TV ドラマから韓国人の俳優が発した感情音声フレーズを抽出し、聴取者 2 名により 5 感情 (怒り, 悲しみ, 喜び, 驚き, 嫌悪) に分類したデータである。

3.3 北アメリカ感情音声データ

北アメリカ英語の感情音声コーパスとして、RAVEDESS[RAVEDESS 18] を使用した。これは 24 人の役者 (男性 12 人, 女性 12 人) が発話と歌で 8 感情 (平静, 落ち着いた, 喜び, 悲しみ, 怒り, 恐怖, 嫌悪, 驚き) を演じたコーパスである。今回は発話データのみを用いた。

4. 周波数帯域による分類性能比較実験

別言語の学習済みモデルを転移学習することによるデータ不足の補完の予備実験として、使用する周波数帯域による分類性能の比較実験を行った。サンプリング周波数に CD 規格の 44.1kHz を採用したため、入力として 0~22.05kHz までの周波数のスペクトログラムを用いることができる。しかし、人の可聴範囲は一般的に 20~20kHz といわれており、実際に感情の特徴となる周波数成分は更に狭い範囲に分布すると予想される。そのため、入力として周波数の上限を適切に定めることが必要であるのではないかと考えられる。

表 4: 韓国語の分類結果 (F 値)

	全帯域	5kHz 上限
喜び	0.346	0.347
悲しみ	0.505	0.498
怒り	0.581	0.571
嫌悪	0.508	0.524
驚き	0.367	0.392
平均	0.469	0.468

表 5: アメリカ英語の分類結果 (F 値)

	全帯域	5kHz 上限
喜び	0.483	0.501
悲しみ	0.564	0.569
怒り	0.501	0.598
嫌悪	0.507	0.523
驚き	0.523	0.586
平均	0.517	0.558

4.1 実験方法

本実験では、入力データであるスペクトログラムの周波数帯域を 5kHz 上限にした場合と 22.05kHz(全帯域) の場合で各言語ごとに判別性能を比較する。本実験では 2.2 の 1 次元畳み込み双方向 LSTM モデルを使用した。5 分割交差検証を行い、各感情ごとに F 値を算出する。このとき、訓練データの 2 割を検証データとする。なお、学習エポック数は 100 回とした。

4.2 実験結果

結果を表 4, 表 5, 表 6 に示す。比較して高い方を太字で表記した。

まず、韓国語の感情音声について分類した結果を比較すると(表 4),「悲しみ」「怒り」については 0~22.05kHz の周波数帯域のデータを入力とした場合の方が F 値が良い結果になった。しかし、「喜び」「嫌悪」「驚き」については 5kHz 上限で区切ったスペクトログラムを入力した場合の方が良い結果となった。

次に、アメリカ英語の感情音声について分類した結果を比較すると(表 5),「喜び」「悲しみ」「怒り」「嫌悪」「驚き」全てにおいて 5kHz 上限で区切った場合の方が良くなった。特に「怒り」については 0~22.05kHz の周波数帯域のデータを入力した場合に比べて F 値が約 0.10 改善された。

最後に、日本語の感情音声について分類した結果を比較すると(表 6),「喜び」「悲しみ」については 0~22.05kHz の周波数帯域のデータを入力とした場合の方が良い結果となった。しかし、「怒り」「嫌悪」「驚き」については 5kHz 上限で区切った場合の方が良い結果となった。特に「怒り」については F 値が約 0.05 改善された。

4.3 考察

全体的に、5kHz 上限でスペクトログラムを区切った場合の方が良い結果になった。この要因として 5kHz 上限で入力を区切ったことにより、入力次元が 4 分の 1 以下まで削減されたことが大きいと思われる。入力データのサイズが小さいほど情報が減ってしまうのは確かであるが、その分モデルの規模が小さくなり、求めるべきパラメータが減少する。これにより学習に必要な学習データの規模が小さくなる。特に本実験のようにデータ数が少ない場合は次元削減によるメリットが大きかったと考えられる。また、感情分類に深くかかわる周波数帯域を

表 6: 日本語の分類結果 (F 値)

	全帯域	5kHz 上限
喜び	0.444	0.423
悲しみ	0.426	0.391
怒り	0.279	0.327
嫌悪	0.432	0.472
驚き	0.560	0.595
平均	0.429	0.443

削除してしまうと逆に性能が悪化することが考えられる。しかし、本実験の結果を見てみると全体的には性能の向上が見られたことから、感情分類に必要なとなる周波数の大部分は 5kHz 以内に分布すると推測される。

5. 転移学習による学習データ補填実験

5.1 実験方法

本実験では、単一言語モデルと転移学習モデルの性能を各言語ごとに比較する。ある言語 A の感情を推定する場合 (A 以外の言語を B, C とする) は以下の 4 つのモデルを比較する。なお、スペクトログラムは 5kHz 上限のサイズ 101×200 のものを使用する。

- 単一言語モデル
- 転移学習モデル (言語 B)
- 転移学習モデル (言語 C)
- 転移学習モデル (言語 BC)

5 分割交差検証を行い、各感情ごとに F 値を算出する。このとき、訓練データの 2 割を検証データとする。なお、学習エポック数は 100 回とした。

5.2 実験結果

実験結果は表 7, 表 8, 表 9 のようになった。全体で最も結果が良いものを太字、転移学習モデルの中で最も結果が良いものを赤字で示した。

まず、韓国語の感情音声について分類した結果を比較すると(表 7), 全体的に韓国語の単一言語モデルが最も良い結果となった。転移学習モデル 3 つについて比較すると、「悲しみ」「怒り」「驚き」については、転移学習モデル (英日) が最も良い結果が得られた。また、F 値の平均を比較と、転移学習モデル (英日) が最も良い結果であった。

次にアメリカ英語の感情音声について分類した結果を比較すると(表 8), 全体的にアメリカ英語の単一言語モデルが最も良い結果となった。特に「喜び」「悲しみ」「怒り」については単一言語モデルと転移学習モデルの間に 0.20 以上の F 値の差が見られた。転移学習モデル 3 つを比較すると、「喜び」「悲しみ」「嫌悪」については転移学習モデル (日韓) が最も良い結果が得られた。また、F 値の平均を比較すると、転移学習モデル (日韓) が最も良い結果であった。

最後に日本語の感情音声について分類した結果を比較すると(表 9), 全体的には日本語の単一言語モデルが最も良い結果となった。特に「怒り」については顕著であり、単一言語モデルが転移学習モデルよりも約 0.20 高い F 値となった。「喜び」については転移学習モデル (英韓) が最も良い結果になった。転移学習モデル 3 つについて比較すると、「喜び」「怒り」「嫌悪」

表 7: 韓国語における単一言語モデルと転移学習モデル (英), (日), (英日) の比較 (F 値)

	単一言語 モデル	転移学習モデル		
		(英)	(日)	(英日)
喜び	0.347	0.163	0.138	0.144
悲しみ	0.498	0.344	0.284	0.380
怒り	0.571	0.484	0.495	0.540
嫌悪	0.524	0.384	0.354	0.374
驚き	0.392	0.350	0.374	0.384
平均	0.468	0.346	0.342	0.366

表 8: アメリカ英語における単一言語モデルと転移学習モデル (日), (韓), (日韓) の比較 (F 値)

	単一言語 モデル	転移学習モデル		
		(日)	(韓)	(日韓)
喜び	0.501	0.296	0.247	0.301
悲しみ	0.569	0.291	0.274	0.342
怒り	0.598	0.311	0.384	0.378
嫌悪	0.523	0.323	0.366	0.416
驚き	0.586	0.412	0.369	0.392
平均	0.558	0.330	0.332	0.368

「驚き」については、転移学習モデル (英韓) が他の転移学習モデル以上の F 値であった。また、F 値の平均を比較すると、転移学習モデル (英韓) が最も良い結果であった。

5.3 考察

単一言語モデルと転移学習モデルの結果を比較すると、全体的に転移学習モデルの方が悪い結果になった。2 言語転移学習モデルと 1 言語転移学習モデルの分類結果の比較では、どの言語も 5 感情中 3 感情以上で 2 言語転移学習モデルの方が良い結果が得られた。今後、学習する感情音声の言語数を増やすことで、データ数が不十分な単一言語の感情音声で学習した判別器以上の性能が得られるのではないかと考えられる。

6. まとめ

本研究では、言語文化によらない共通の特徴があり、ある言語の感情音声のデータの不足を別言語の感情音声で補完できるのではないかと考え、別の言語について学習したモデルを転移学習で利用することを提案した。

まず、スペクトログラムの周波数帯域を適切なところで制限した方が分類性能が向上するのではないかと考え、上限 5kHz で区切った場合と、サンプリング周波数の半分 22.05kHz までを入力とした場合で言語ごとに学習を行い、性能の比較実験を行った。その結果、日本語と韓国語は 5 感情中 3 感情が、アメリカ英語は 5 感情全ての感情について、区切った場合の方が高い F 値が得られた。よって、5kHz までの範囲に感情に関連する周波数が多く分布するのではないかと考えられる。

次に、ある言語の感情音声のデータ不足を別の言語の感情音声を学習したモデルを転用することで補う方法について検討を行った。結果としては、日本語の「喜び」以外は 1 から学習したモデルが分類性能において最も良い結果であった。しかし、1 言語転移学習モデルよりも 2 言語転移学習モデルの方が全体的に良い結果が得られたことから、学習する言語数を増やすことで、不十分なデータ量の単一言語モデルよりも高い性能が得

表 9: 日本語における単一言語モデルと転移学習モデル (英), (韓), (英韓) の比較 (F 値)

	単一言語 モデル	転移学習モデル		
		(英)	(韓)	(英韓)
喜び	0.423	0.412	0.344	0.437
悲しみ	0.391	0.362	0.312	0.307
怒り	0.327	0.102	0.070	0.128
嫌悪	0.472	0.374	0.340	0.428
驚き	0.595	0.519	0.505	0.519
平均	0.443	0.365	0.334	0.378

られる可能性が示唆された。

これらの結果を踏まえ、今後は更に多くの言語の感情音声を入力し、本実験で示された可能性について検証していく予定である。また、言語数を増加させていくと、本手法では次元数が増加してしまう。対策として、出力次元を減らすことや、文化的背景に近い言語ごとに系統分けすることなどを検討していく予定である。

参考文献

- [Dario 16] Dario Bertero et al, "Real-Time Speech Emotion and Sen-timent Recognition for Interactive Dialogue Systems " Proceedings of the 2016 Conference on Empirical Meth-Ods in Natural Language Processing “ , pp. 1042-1047
- [George 16] George Trigeorgis et al, “ Adieu features? end-to-end speech emotion recognition using a deep convolutional recurrent network ” in Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2016, pp. 5200-5204.
- [Ekman 75] Ekman, P. and Friesen, W. V.” Unmasking the Face, Prentice-Hall” ,1975
- [Nesterov 83] Nesterov, Y. (1983). A method for unconstrained convex minimization problem with the rate of convergence $O(1/k^2)$. Doklady ANSSSR (translated as Soviet.Math.Docl.), vol. 269, pp. 543-547.
- [RAVEDESS 18] Livingstone SR, Russo FA (2018) The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVEDESS): A dynamic, multimodal set of facial and vocal expressions in North American English. PLoS ONE 13(5): e0196391. <https://doi.org/10.1371/journal.pone.0196391>.
- [有本 08] 有本泰子ら, “感情音声のコーパス構築と音響的特徴の分析” 情報処理学会研究報告音楽情報科学 (MUS) ,pp.133-138,2008
- [有本 13] 有本泰子, 河津宏美, “音声チャットを利用したオンラインゲーム感情音声コーパス”, 日本音響学会 2013 年秋季研究発表会講演論文集, 1-P-46a, pp. 385-388, 2013.
- [Cho 09] 趙章植ら, “ベジアンアプローチに基づく感情発話音声からの感情推定における日韓感性の比較” 日本感性工学会論文誌 Vol.8No.3 pp.913-919, 2009

オープンソースの音声言語インタラクションの社会実験基盤を 提供する MMI プラットフォームの開発

Development of Open-source Multi-modal Interaction Platform for Social Experiment of Conversational User Interface

李 晃伸

Akinobu Lee

名古屋工業大学大学院工学研究科

Nagoya Institute of Technology, Japan

A development of a multi-modal interaction platform for Social experiment of conversational user interface is proposed. In order to go over the simple spoken language interaction systems such as voice assistants, it is necessary to elucidate various factors of rich interactions quantitatively via thousands of wide variety of actual interaction data from users. The proposed system is based on a voice interaction building toolkit MMDAgent, adding some features to promote a testbed for social experiment and data collection of speech interaction system on cloud environment. It includes facilities for system distribution and management, collection of interaction log and speech data, and easy connection with cloud-based chat system. The beta version of the software is available, and it will be released as open-source software to promote wider use for various speech-based conversational user interface.

1. はじめに

近年、音声対話システムあるいは対話的な音声言語インタフェースは、情報機器・サービスと人をつなぐ次世代のモダリティとして期待が高まっている。特にここ数年、知的情報処理技術の深化と汎化に伴って、クラウドベースの知的情報サービスと人間を繋ぐスマートなコミュニケーションインタフェースとしての音声対話技術が注目されている。スマートスピーカーや Web 検索等において簡潔な発話による情報の授受やタスク達成はかなり実現しており、今後はより自然な会話スタイルでやりとりできることが求められている。

一方で、インタフェースとしての音声モダリティの一般的な設計論や方法論はまだ確立されているとはいえない。音声対話システムの研究事例や応用システム例はこれまでに多くあり、音声認識や音声合成の基盤技術も近年飛躍的進歩を得たが、それでもなお、現時点では音声は様々な機器やサービスに実際にアクセスするための主要なモダリティとなるに至っていない。また音声対話システムの構築には必要とされる技術が多く、CG エージェント等のインタフェースまで含めた設計・実装は多方面の多大な労力を必要とする。

音声対話を含む音声言語インタラクションが一般的なモダリティとして広く用いられるようになるためには、その設計や適用範囲、他の UI との関連性を含め、統合的に実証・実験する枠組みが必要である。音声言語の領域だけでなく、自然言語処理や対話、ユーザインタフェースやヒューマンエージェントインタラクション、あるいはデザインやコンテンツの領域を検証可能な統合的な基盤を共有することで、多様な実際の・現実的なタスクのオープンなシステム運用から多くの実インタラクションデータを収集し、データドリブンなアプローチで様々な要素を量的に解明していくことが可能となる。様々な目的のための多様なシステムが大量に作成され実際に使われることで、多くの事例やデータを集める環境がボトムアップに形成されると期待される。

本研究では、音声対話を含む音声言語インタラクションの多様な社会実験の実践のための基盤となるプラットフォームの実現を目標に開発された拡張版 MMDAgent について述べる。本アプリケーションは、MMDAgent で作成した音声対話・音声インタラクションシステムのネットワーク配信、ならびに利用ユーザのインタラクションログや音声データの収集を行える。これにより、音声対話システムを公開することで誰でもクラウドベースの広範囲な音声言語インタラクション実験およびデータ収集を行える。本アプリケーションはマルチプラットフォームで動作し、Android, iOS およびデスクトップ OS (Win/Mac/Linux) で同一の動作を行う。現在ベータ版が公開されており無償でダウンロード・試用が可能である。以下、ベースシステムである MMDAgent について概説したのち、提案システムについて述べる。

2. MMDAgent

まずベースとなるツールである MMDAgent について述べる。MMDAgent は音声ベースのインタラクションシステムのための構築ツールキットである[1]。多様な音声対話システムおよび音声インタラクションのための実験用プラットフォームとして開発が行われている。現在、ソフトウェア全体およびサンプルの音声対話システムがオープンソース (BSD ライセンス) で公開されている。2010 年に 1.0 が公開された。ソフトウェアの最新版は 2016 年に公開されたバージョン 1.7 である。公開以来 14 万件以上ダウンロードされている[2]。

MMDAgent 自身は音声対話システムではなく、音声対話システムを動作させるソフトウェア (ブラウザ) である。音声対話システムの構成要素 (辞書・ボイス・対話シナリオ・エージェントモデル・モーション等) は外部で定義する。MMDAgent 用のサンプルシステム (サンプルコンテンツ) が同じ Web で公開されており、2018 年 12 月には従来の女性モデル・女性ボイスに加えて男性モデル・男性ボイスを追加したバージョン 1.8 が公開された。ドキュメントは JST CREST の uDialogue プロジェクト (2011—2017) の成果物として uDialogue サイトの「MMDAgent エンサイクロペディア」[3]に集約されている。

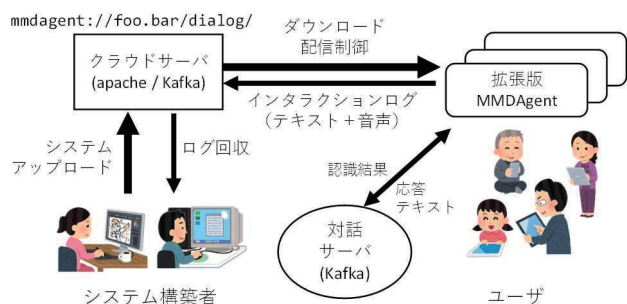


図1:システム全体像

MMDAgent では音声認識 (Julius), 音声合成 (Open JTalk (日本語), Flite+HTS Engine (英語)), エージェント表示がオールインワンで提供されている。各部がモジュール化されていて拡張が容易であり, 言語モデルや音響モデルの拡張, ボイスモデルの入れ替え等が容易に行える。画面表示は 3D 空間で, 物体とモーションの分離性や十分な表現力と自由度から, MikuMikuDance の互換プラットフォームとしており, プリミティブな 3D オブジェクトからヒューマノイドエージェントによるしぐさ・反応・ジェスチャの表出まで多様なインタラクション表現が可能である。対話管理部はメッセージテキストを入出力とする状態遷移機械 (FST) で実装されている。

3. 提案システムの概要

本研究ではスマートフォンから等身大サイネージまで多様なデバイスにおいて, 音声を中心とする会話的な UI を備えた音声インタラクション・対話システムのテストベッドとして双方向の実験および大規模データ収集を行えるよう MMDAgent を拡張した新たなプラットフォームを構築する[4]。本システムはスマートフォンマルチ OS 環境で動作し, 双方向の大規模な社会実験環境の基盤を提供する。以下, 実装された新たな機構を述べる。全体像を図 1 に示す。

3.1 システムのクラウド配信および管理機構

音声対話システムを任意の URL から任意のユーザー端末へ直接配信可能にする。Web 上の URL で示されるディレクトリ以下にシステムのファイル一式 (対話スクリプト, 3D モデル, モーション, 背景画像, ボイス定義ファイルなど) を置いておき, アプリケーションからその URL を開くことでシステムを自動取得できる。URL スキームに対応しており, ユーザーが “mmdagent://...” のリンクを開くだけでシステムを自動ダウンロード・実行できる。

また, システム構築者からダウンロードしたユーザーへの制御・連絡を行う仕組みを実装している。ダウンロードしたシステムに対する更新や削除, お知らせの表示を, 端末からサーバへ定期的な更新チェックを行うことで自動的に行う仕組みになっており, 利用ユーザーに対する様々な配信制御が可能である。

3.2 ログ・フィードバック収集機構

システム構築者が自システム利用時の動作ログを収集するための機構を実装する。収集する動作ログは, MMDAgent のメッセージキューを流れたすべてのメッセージおよびシステムログであり, 音声認識結果, 対話管理 FST の状態遷移などのインタラクション情報がミリ秒単位のタイムスタンプとともに保存されたものである。また, 認識対象となった生の音声波形データも収集できる。端末の識別ごとにユニーク ID が自動生成され, 識別子としてログに付与される。

ログの収集スキームとして, Apache Kafka を用いたリアルタイムログ収集と, ローカルにファイル保存してから Web アップロードする方法の 2 種類を実装する。前者は 2011 年にオープンソース化され, LinkedIn や Twitter でも用いられている分散ストリーミングのプラットフォームであり, 数十万以上の大量の端末に対して高性能なリアルタイムデータストリーミングを行うことができる。即時性が高く, コンテンツのエラー検出やロギングをリアルタイムに行うことができる。後者の Web アップロードは, 端末のローカル上にログを記録・保存し, 一定のタイミングでそれらを POST メソッドで Web 上のサーバへアップロードする。音声波形データの収集は後者の場合のみに対応する。これにより, ユーザーからコンテンツ公開者へのフィードバックを実現する。

3.3 クラウド型対話システムへの対応

システムが指定する Apache Kafka サーバと consumer モードで接続することで, 個々の端末が Kafka サーバを通じて認識結果や応答テキストなどのメッセージをリアルタイムにやりとりできる。これを用いることで, 対話サーバを接続して当アプリケーションをフロントエンドとしたクラウド型音声対話システムを容易に実現可能である。また接続先はシステム単位で変更できるため, あるシステムを利用中の端末全てに発話命令を一斉送信するような利用方法も可能である。

3.4 オープンソース開発体制

多様な音声インタラクションを対象とするためには, 様々なセンサーあるいは IFTTT のようなバックエンド接続サービス等との接続のための拡張を行いやすい開発環境である必要がある。本アプリケーションの開発においては GitHub を活用した実践的な体制を構築し, ソースコード共有によるコード主体の多様かつ迅速な開発を行う予定である。

4. 現行システム

本稿で提案した拡張の多くを施したアプリケーションは, 既に「Pocket MMDAgent」という名称で試験公開中である[5]。サイトからは Android, iOS を含む各種 OS 用のベータ版アプリが無償で入手可能であり, 仕様もサイト上で公開されている。

5. まとめ

本システムの開発進捗は現在 80%程度であり, 最終的にはオープンソースで公開する予定である。マルチプラットフォームの UI 基盤としては Unity が著名であるが, 本ソフトウェアはオープンソースであり, 音声言語インタラクションのデータ収集基盤としての拡張性と可搬性を重視して作成されている。

本システムの開発と公開が, 音声を含めた知的インタラクションにかかる研究テーマの統合的でデータドリブンなアプローチの土台となり, 音声言語を含めた次世代インタフェースの幅広い試行錯誤と検討の一助となれば幸いである。

参考文献

- [1] A. Lee, K. Oura, K. Tokuda: MMDAgent - A fully open-source toolkit for voice interaction systems, IEEE ICASSP, pp. 8382-8385, 2013.
- [2] <http://www.mmdagent.jp/>
- [3] <http://www.udialogue.org/ja/encyclopedia-ja>
- [4] 李晃伸: 音声対話コンテンツのネットワーク配信および大規模ログ収集を可能にするスマートフォン版 MMDAgent の開発, 日本音響学会秋季講演論文集, 2-2-8, 2019.
- [5] Pocket MMDAgent (beta): <https://mmdagent.lee-lab.org/>

[3N4-J-10] Vision, speech: applications to industries

Chair: Masanori Tsujikawa Reviewer: Tomoya Yoshikawa

Thu. Jun 6, 2019 3:50 PM - 4:50 PM Room N (Front-right room of 1F Exhibition hall)

[3N4-J-10-01] Measurement of growing situation of agricultural crops on FPGA-mounted drone using Circle SSD○Takuma Yoshimura¹ (1. poco-apoco Networks Co.Ltd.)

3:50 PM - 4:10 PM

[3N4-J-10-02] Prediction of Favorability Rating on Beer-Can Package Designs Using Convolution Neural Network and Visualization by Class Activation Mapping.○Hiroyuki Shinohara¹, Tatsuji Ishiguro¹, Shunsuke Nakamura², Toshihiko Yamasaki² (1. Kirin Company, Limited, 2. The University of Tokyo)

4:10 PM - 4:30 PM

[3N4-J-10-03] Conversion of Floor Plan Images to Graph Structures using Deep Learning and Application to Retrieval○Mantaro Yamada¹, Xueting Wang¹, Toshihiko Yamasaki¹, Kiyoharu Aizawa¹ (1. the University of Tokyo)

4:30 PM - 4:50 PM

Circle SSD を用いた FPGA 搭載ドローンでの農作物生育状況計測 Measurement of growing situation of agricultural crops on FPGA-mounted drone using Circle SSD

吉村 拓馬^{*1}
Takuma Yoshimura

^{*1} 株式会社ポコアポコネットワークス
poco-apoco Networks Co. Ltd.

In this paper, we report the optimization method for the case of introducing SSD which is one of the object detection models to the drones for the purpose of measuring the growth situation of agricultural crops as viewed from a bird's eye viewpoint. We have adopted a "Separable Convolution" and introduced "Circle SSD" which uses circular shape as a detection frame as an optimization method. As a result of the optimization, the F-measure 0.67 was realized with 14770 convolution params, and the object detection could be realized with limited calculation resources on the drone.

1. はじめに

本稿では、農作物の生育状況計測をドローンより上空から捉えた鳥瞰視点で行うことを目的とし、ドローンに搭載される FPGA(Field-Programmable Gate Array)ボードに物体検出モデルの 1 つである SSD(Single Shot MultiBox Detector)[Liu 16]を導入した事例について、物体検出モデルの小型化および円形の検出枠の適用に関して報告する。

ドローンとは遠隔操縦または自動操縦が可能な無人航空機の総称で、農業、設備点検、測量、警備、物資運搬など産業界での活用が広がっている[Nonami 18]。ドローンの多くは内蔵バッテリーによって駆動するため、長時間飛行させるためには消費電力を抑える必要がある。したがって生育状況計測のための搭載デバイスは軽重量で消費電力の小さいことが望ましく、さらに無線通信による電力損失を抑えることが求められる。

また本稿で採用する SSD は、統計的機械学習の数理モデルの 1 つで、画像認識や画像生成などに広く用いられる畳み込みニューラルネットワーク(以降 CNN)を用いている。しかし CNN は学習時、推論時に必要となる計算量が大きく、それゆえに高性能で消費電力の大きい CPU や GPU での運用が求められる。

本稿では、物体検出モデルの小型化を行い CPU、GPU に比べ消費電力の小さい FPGA[†]に SSD を導入した。ドローンから空撮した画像を搭載した FPGA で推論させることで、地上にあるサーバに文字列に比べデータ量の大きい画像を通信する必要がなく、通信量を抑えることが可能となる。全体の流れを図 1 に示す。本稿では図中のフロー 2 について述べる。

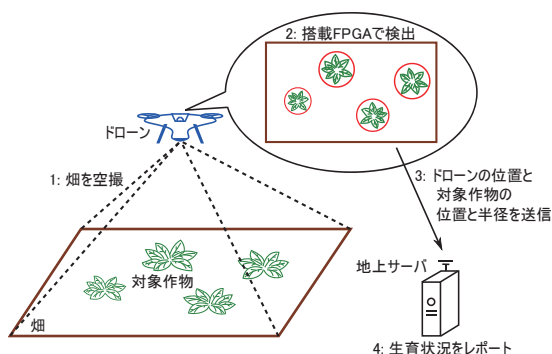


図 1 農作物生育状況計測の流れ

2. データセット

本稿では 2018 年の 5 月から 9 月にかけて撮影されたパイナップル耕作地の空撮画像 2187 枚および画像に含まれる、パイナップルの中心位置および葉の先端への半径のリストをデータセットとした。データセットの一部を以下の図 2 に示す。



図 2 パイナップル空撮画像 それぞれの株の中心部分から果実として流通しているパイナップルが生育する。

3. SSD の構成

本節では農作物の生育状況計測を実現するための、物体検出モデルである SSD の小型化、また円形の検出枠の適用について述べる。

3.1 SSD の構成要素

SSD は Extractor, Coder, Suppressor の 3 つの構成からなる(図 3)。Extractor は入力画像から畳み込み処理を行い物体認識に有効な特徴量を抽出、Coder は Extractor から得られた特徴量を各クラスの物体が存在する範囲の空間情報と確度(スコア)に変換、Suppressor は Coder から得られた検出枠候補から重複した確度の低い検出枠候補を除外する。

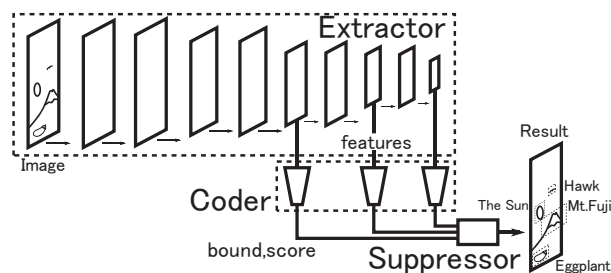


図 3 SSD の構成要素

連絡先: 吉村 拓馬, 株式会社ポコアポコネットワークス,
〒550-0012, 大阪府大阪市西区立売堀 2-1-9 日建ビル 6 階

[†] DIGILENT 社製 Zybo Zynq-7020, USB 給電 5V,
基板サイズ 88mm×122mm, BRAM 容量 630KB

3.2 FPGA のハードウェア制約

一般に高性能な CPU や GPU が要求される CNN を FPGA に実装するためには、記憶領域と論理ゲート数(回路規模)の制約条件を満たす必要がある。CNN の小型化の手法としてモバイルデバイス向け小型 CNN である MobileNet ベース SSD300(パラメータ数 6.8×10^6)[Howard 17]を参考とし、制約条件を満たすためパラメータ数を 1.5×10^4 までとし、Extractor の畳み込み層を 8 層まで減らす必要があった(表 1)。このため CNN の構成は一般のモデルに比べ制限される。例えば、このパラメータ数ではフィルタサイズ 3 チャンネル数 64 の、一般的な畳み込み層を 1 層も構成に入れることができない(パラメータ数 $3^2 \times 64^2 = 36864$)。

ただし、物体検出タスクとしての難度は先行研究に比べ低い。なぜなら先行研究のモデルは一般物体の複数クラスの認識精度を評価指標としている(COCO dataset[Lin 14])。本稿では、対象作物の 1 クラスのみであるうえに、対象の構造は水平に対し等方的、背景は主に赤褐色から黄褐色の土で色域が限られる。

表 1 モデルパラメータ

Model	Parameters	Dataset	Classes
deeplab-VGG	3.3×10^7	COCO	80
Inception V2	1.4×10^7	COCO	80
MobileNet	6.8×10^6	COCO	80
Ours	1.5×10^4	Pineapple	1

3.3 CNN の小型化

CNN は複数の処理層で構成され、一般に多くの演算を要する数値モデルである。1 つ目の処理層である入力層では画像に対して空間フィルタを適用(畳み込み処理)し、特徴マップを得る。以降の処理層では前の層で得られた特徴マップに対してさらに畳み込み処理し特徴抽出する、またはプーリング処理によりロバスト性を高める。こうしてより物体認識に有効な特徴量を抜き出した特徴マップを得る。この処理において物体検出モデルでは $10^9 - 10^{12}$ スケールの積和演算回数を要する[Howard 17]。

CNN の小型化・高速化において特徴抽出性能を保ちつつ、計算量を減らす構造がいくつか考えられている。代表的なものとして MobileNet[Howard 17]が導入した Depthwise Convolution, Pointwise Convolution, この 2 つを組み合わせた Separable Convolution がある(図 4)。その他にも SqueezeNet[Iandola 16]が導入した Fire Module, GoogLeNet[Szegedy 14]が導入した Inception Module があるが、実装上 FPGA の論理ゲート数を前記の構造より多く消費するため用いなかった。

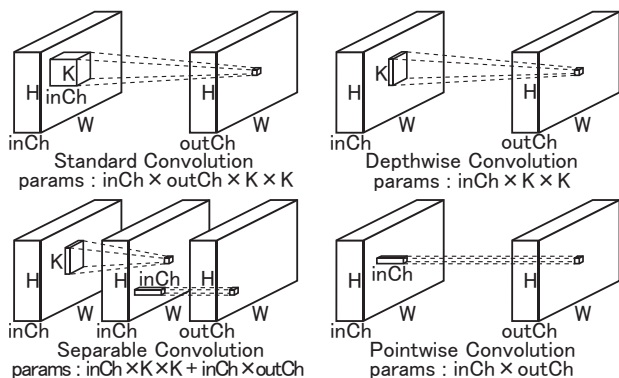


図 4 MobileNet における畳み込み処理 inCh:入力チャンネル数, outCh:出力チャンネル数, K:フィルタサイズ, W, H:特徴マップの幅および高さ, params: バイアスを除くパラメータ数(出力特徴マップ 1 座標あたりの積和演算回数に相当), なお図のストライド(S)は 1 である。

3.4 Extractor の構成方針

CNN は層が深くなるに従い、物体検出モデルにおいては一般にチャンネル数が増えていく。図 4 に示すとおり Standard Convolution, Pointwise Convolution, Separable Convolution のパラメータ数はチャンネル数が支配的であるため、チャンネル数が大きい処理層にはこれらの構造は避ける。本稿では、チャンネル数が最も小さい入力層には Standard Convolution, Coder に分岐するまでは Separable Convolution, 分岐後のチャンネル数が大きい処理層には Depthwise Convolution を用いた。

3.5 Coder の構成方針

Coder では Extractor で得られた特徴量を検出対象のクラスと背景クラスの確度、位置情報のチャンネルに変換するため、チャンネル間の畳み込み処理を要する。本稿では、最小限の構造である Pointwise Convolution を 1 層のみ用いた。また、一般的な物体検出モデルの検出枠は矩形であるが、植物のように水平等方向に空間を占める物体については円形であるほうが適していると考えた。矩形の検出枠では横位置、縦位置、横幅、縦幅の 4 パラメータの相対位置情報が必要となるが、円形の検出枠を適用することで、横位置、縦位置、半径の 3 パラメータのみで済み、さらに縦長横長の検出枠が不要となる。

3.6 Suppressor の構成方針

Suppressor では鳥瞰視点の空撮画像、円形を検出枠とする Circle SSD に適した検出枠重複除外処理に変更した。SSD[Liu 16]で用いられていた重複度ベースの除外方法を Circle SSD に適用すると、地表面面上に存在する検出対象の中心が重なることは無いにもかかわらず、円中心が近くに存在する半径の異なる複数の検出枠が得られてしまう。そこで本稿では互いの円の半径に対する中心への比距離が互いに 50%以下である場合、確度の低い方の円を重複した検出枠として除外することにした(図 5)。この閾値は検出対象の大きさと、近傍の検出対象との距離の比の分布から決定した。

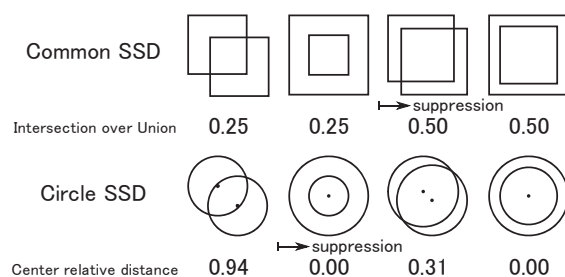


図 5 中心相対距離に基づく検出枠重複除外 上: 一般の SSD における検出枠配置とその Intersection over Union (IoU) 下: Circle SSD における検出枠配置と半径相対中心距離

3.7 Circle SSD の構成

本節をまとめると CNN の小型化をおこなった Circle SSD の構成は次頁の表 2 のようになる。列名は図 4 を参照のこと。

4. 学習時の変更点

本節では野外で運用するドローンに適したデータ拡張と、SSD[Liu 16]に円形の検出枠を適用したことによる特徴マップの形状の変更について述べる。

表 2 CNN の小型化をおこなった Circle SSD の構成

	Type	inCh	outCh	K	S	W	H	params
Extractor								11520
Conv1	StandardConv+Relu	3	8	5	1	512	288	608
Pool1	MaxPooling	8	8	2	2	512	288	0
Conv2	SeparableConv+Relu	8	16	5	1	256	144	352
Pool2	MaxPooling	16	16	2	2	256	144	0
Conv3	SeparableConv+Relu	16	32	5	1	128	72	960
Pool3	MaxPooling	32	32	2	2	128	72	0
Conv4	SeparableConv+Relu	32	64	5	1	64	36	2944
Conv5	DepthwiseConv+Relu	64	64	5	2	64	36	1664
Conv6	DepthwiseConv+Relu	64	64	5	2	32	18	1664
Conv7	DepthwiseConv+Relu	64	64	5	2	16	9	1664
Conv8	DepthwiseConv+Relu	64	64	5	2	8	5	1664
Coder								3250
Loc4	PointwiseConv	64	6	1	1	64	36	390
Conf4	PointwiseConv	64	4	1	1	64	36	260
Loc5	PointwiseConv	64	6	1	1	32	18	390
Conf5	PointwiseConv	64	4	1	1	32	18	260
Loc6	PointwiseConv	64	6	1	1	16	9	390
Conf6	PointwiseConv	64	4	1	1	16	9	260
Loc7	PointwiseConv	64	6	1	1	8	5	390
Conf7	PointwiseConv	64	4	1	1	8	5	260
Loc8	PointwiseConv	64	6	1	1	4	3	390
Conf8	PointwiseConv	64	4	1	1	4	3	260

4.1 データ拡張

野外での物体検出は屋内に比べ、天候や時刻による画像輝度の変動が大きいため難度が高い。ドローンの FPGA に導入する SSD の学習では画像の上下左右反転の他に、輝度の変動を考慮した以下のデータ拡張を行った。

- ガンマ値の変動(冪指数:0.5-1.25)
- RGB 輝度値の独立変動(倍率:0.82-1.00)

4.2 長方形画像の扱い

SSD の矩形の検出枠には縦幅横幅のパラメータがあり、画像の縦横縮尺比を変えることで、長方形画像であっても正方形画像に変形させ扱うことができる。しかし Circle SSD の円形の検出枠の尺度は半径のみであるため、画像変形後に縦長横長となる物体範囲を正しく検出できない。そこで画像の縦横縮尺比を変えず特徴マップの横幅と縦幅を変えることで、本稿で扱う空撮画像のような長方形画像(アスペクト比 16:9)に対応させた。

5. 推論時の最適化

本節では CNN を FPGA に導入するための計算量削減の手法について述べる。

5.1 固定小数点数演算

学習時に用いた浮動小数点演算器を並列動作させる GPU とは対照的に、推論時に用いた FPGA は浮動小数点演算器を持たないまたは乏しい。そのため浮動小数点数の演算は固定小数点数を用いたときに比べ、多くの論理ゲートを消費し回路規模の増大につながる。そこで本稿では符号 1bit 整数 5bit 小数 10bit の 16bit ビット幅の固定小数点数で推論を行った。なお、途中の積算は丸めず融合積和演算(FMA; fused multiply-add)とすることで誤差を小さくした。

5.2 バッチ正規化

畳み込み層と直後のバッチ正規化層[Ioffe 15]は以下の式のように、推論時に単一の等価な畳み込み層へ変換することができ、本稿ではこの変換式を用いた。ここで \mathbf{x} は特徴マップ、 BN はバッチ正規化関数、 $Conv$ は畳み込み関数、 \mathbf{W} , b は畳み込み層の重みおよびバイアス、 γ , β はバッチ正規化が学習する尺度および位置パラメータ、 σ , μ は $Conv(\mathbf{x})$ のバッチ内標準偏差および平均の移動平均、 ϵ はゼロ除算を避けるための微小数である。出力チャンネル方向の各要素について記述している。

$$\begin{aligned}
 BN(Conv(\mathbf{x})) &= \gamma \frac{Conv(\mathbf{x}) - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta \\
 &= \gamma \frac{(\mathbf{W}\mathbf{x} + b) - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta \\
 &= \frac{\gamma\mathbf{W}}{\sqrt{\sigma^2 + \epsilon}}\mathbf{x} + \frac{\gamma(b - \mu)}{\sqrt{\sigma^2 + \epsilon}} + \beta.
 \end{aligned} \quad (1)$$

$$Conv'(\mathbf{x}) = BN(Conv(\mathbf{x})) = \mathbf{W}'\mathbf{x} + b',$$

$$\text{s.t. } \mathbf{W}' = \frac{\gamma\mathbf{W}}{\sqrt{\sigma^2 + \epsilon}}, b' = \frac{\gamma(b - \mu)}{\sqrt{\sigma^2 + \epsilon}} + \beta. \quad (2)$$

6. 結果

対象作物を本稿の Circle SSD で検出した結果画像を次頁の図 7 に示す。撮影高度、画像輝度に依らず対象作物を検出する事ができたものの、対象作物でない雑草への誤検出や、高高度の画像において対象作物の半径の誤認識が見られた。

次に Extractor の構成を表 1(パラメータ数 14770)、Mobile Net like(表 1 の Depthwise Convolution を Separable Convolution に置き換え、パラメータ数 31410)、VGG like(表 1 の Depthwise Convolution, Separable Convolution を Standard Convolution に置き換え、パラメータ数 481026)にした Circle SSD について、検出確度閾値を 0.10-0.95 に変化させたときの Precision-Recall 曲線、最大 F 値で評価を行った。検出枠と正解枠が一致したと判定する基準は互いの円の半径に対する中心への比距離が 25% 以内かつ半径比が 0.8-1.25 に収まっているかとする(図 6)。適合率(Precision)は一致枠数 / 検出枠数、再現率(Recall)は一致枠数 / 正解枠数、F 値は適合率と再現率の調和平均で計算される。また学習画像と検証画像は 9:1 に分け学習と検証を行った。

なお原論文のモデルは、本稿の対象作物 1 クラスのみの検出タスクに対し CNN のパラメータ数が多すぎ、過剰適合に陥ったため対照群としなかった。

Precision-Recall 曲線の比較、モデルサイズと最大 F 値の比較を次頁の図 8, 9 に示す。本稿のモデルの最大 F 値は 0.67 となり、32.5 倍のパラメータ数の VGG like, 2.1 倍のパラメータ数の MobileNet like に対し最大 F 値は 0.10, 0.02 の低下に収まった。また同じ空撮画像に対しモデルを変更した結果を図 10 に示す。

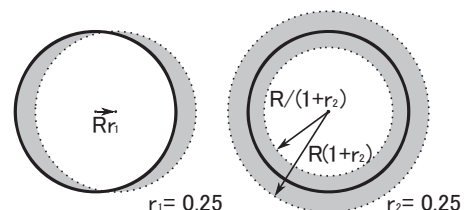


図 6 半径 R の正解枠に対して一致したとみなす検出枠の範囲



図 7 検出結果 撮影高度、画像輝度に依らず検出できている一方で、雑草への誤検出、高高度の画像での半径の誤認識が精度を落としている。なお、検出枠はランダムに着色している。

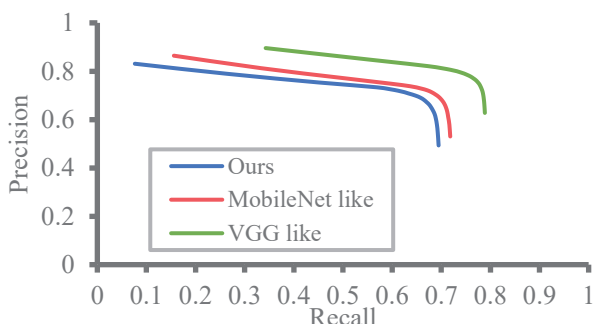


図 8 Precision-Recall 曲線の比較 理想的な検出器は Precision = 1, Recall = 1 の点を通る。F 値はいずれも検出確度閾値 0.60 で最大となった。

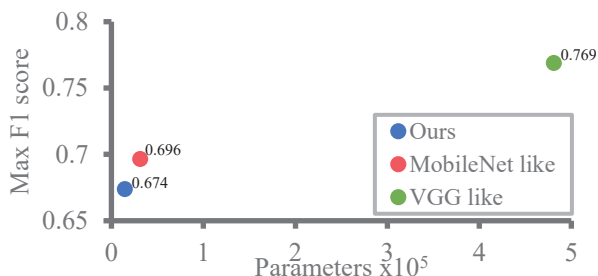


図 9 モデルサイズと最大 F 値の比較 モデルサイズと検出精度は一般的にトレードオフの関係にある。VGG like, MobileNet like のパラメータ数は本稿のモデルの 32.5 倍, 2.1 倍あり、最大 F 値は本稿のモデルより 0.10, 0.02 大きい。



図 10 モデル変更時の検出結果の比較 この画像では VGG like と本稿のモデルは株が重なっている箇所にそれぞれ誤検出が 2 箇所, 1 箇所あり、VGG like は隣接する株に最長の葉が被っている株であっても検出範囲が正確である。

7. まとめ

本稿では鳥瞰視点で捉えた農作物の生育状況計測を目的に、円形の検出枠を適用した SSD として Circle SSD を提案した。またドローンに搭載する FPGA のハードウェア制約のもと、検出対象が 1 クラスのみという限られた条件付きであるものの、パラメータ数 14770 という小型 CNN 物体検出モデルを実用範囲内で実現することができた。

改良点として、Circle SSD への損失関数の最適化に、下式の円の類似度の採用を考えている。この類似度を用いることで株の重複箇所への誤検出が抑えられると期待される。ここで正解枠と Coder の各ポイントに割り振られたデフォルト枠の 2 円間について、 d は 2 円中心間距離、 r_1, r_2 は 2 円の半径である。

$$S = \max(0, 1 - \max(d/r_1, d/r_2)) \cdot \max(0, 1 - |\log(r_1/r_2)|) \quad (3)$$

課題として、矩形を検出枠とする SSD と本稿が提案した円形を検出枠とする Circle SSD では、検出枠が正解枠と一致したとみなす基準が異なり、精度の比較ができない事が挙げられる。両者で共通して用いることのできる評価手法について考察を深めたい。

最後に、ハイエンドな計算機で運用される高性能の CNN だけでなく、近年では個人向け PC やモバイルデバイス、エッジデバイスで運用される中性能・小型の CNN が求められている。物体検出タスクだけでなく、単一画像超解像や領域抽出、波形認識などのタスクを実行できる小型 CNN モデルの開発に取り組んでいきたい。

謝辞

本稿を作成するにあたり、ご助言を頂いた岡山県立大学の有本和民氏、大手前大学の奥村紀之氏に心より感謝致します。

参考文献

- [Liu 16] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg. SSD: Single Shot MultiBox Detector, arXiv:1512.02325, 2016.
- [Nonami 18] 野波 健蔵, ドローン産業応用のすべて—開発の基礎から活用の実際まで—, ISBN-13: 978-4274506840, 2018.
- [Howard 17] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam, MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, arXiv:1704.04861, 2017.
- [Lin 14] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, Piotr Dollár, Microsoft COCO: Common Objects in Context, arXiv:1405.0312, 2014.
- [Iandola 16] Forrest N. Iandola, Song Han, Matthew W. Moskewicz, Khalid Ashraf, William J. Dally, Kurt Keutzer, SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size, arXiv:1602.07360, 2016.
- [Szegedy 14] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, Going deeper with convolutions, arXiv:1409.4842, 2014.
- [Ioffe 15] Sergey Ioffe, Christian Szegedy, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, arXiv:1502.03167, 2015.

畳み込みニューラルネットワークを用いたビール缶パッケージの 好意度予測及び要因の可視化

Prediction of Favorability Rating on Beer-Can Package Designs Using Convolution Neural Network and Visualization by Class Activation Mapping.

篠原裕之^{*1}
Hiroyuki Shinohara

石黒達治^{*1}
Tatsuji Ishiguro

中村遵介^{*2}
Shunsuke Nakamura

山崎俊彦^{*2}
Toshihiko Yamasaki

^{*1} キリン株式会社
Kirin Company, Limited

^{*2} 東京大学
The University of Tokyo

A quantitative survey of favorability rating by multiple panels is the main method to decide package designs. However, it is expensive and there is a limitation in the number of designs that can be subjected to a single survey. Therefore, this study aims at predicting the panel evaluation from the past survey results by a convolution neural network and visualize the important features by Grad-CAM. As a result, it has been made possible to give prescreening test to package design and suggestions of important features to the designers.

1. 背景と目的

最終消費者の購買行動において商品のパッケージデザインは重要である [石井 10]。その為、企業において商品を上市する際には、パッケージデザインに関する消費者調査を行い、良好な結果が得られたデザインを採用する事が多い。しかし、十分な消費者パネルを集めた調査には多額の費用が必要である、一度の調査にかけられるパッケージデザインの数に上限があるという課題がある。

一方で、近年はディープラーニングを用いて審美性評価や関する研究が盛んにおこなわれている [Talebi 17] [Kong 16]。これらの例から適切なディープラーニングのネットワークモデルを用いればパッケージデザインに関する消費者調査結果を予測する事が可能であると考えられた。但し、これらの研究の多くは豊富なデータを用いてモデルを構築しており、単独の企業で有する限られたデータのみを用いて消費者調査結果を高精度で予測可能なモデルを構築する事は困難であることが予想された。

そこで本研究は以下の 2 点を実現する事を目的とした。即ち、未調査のパッケージデザインの中で好意度に対して好影響を及ぼすと予測される要素と悪影響を及ぼすと予測される要素を可視化する事、未調査のパッケージデザインを好意度調査に供した場合の結果を予測する事である。

2. 対象データ

本研究で対象としたデータはキリンビール株式会社が 2001 年から 2017 年にかけて実施した、350mL ビール缶パッケージの好意度調査である。

調査は新規製品の上市あるいは既存製品のリニューアルに際し、製品候補の 350mL ビール缶パッケージを非明示で 64 人から 300 人の消費者パネルに提示し、それに対する好意度を表 1 に示した 5 段階もしくは 7 段階で回答を得たものである。この調査結果について、回答結果を最高 10 点、最低 1 点となるように点数化し、全回答者の点数の平均好意度をパッケージデザインの好意度とした。好意度のヒストグラムを図 1 に示す。

表1 回答方法及び点数化方法

7段階評価		5段階評価	
回答	点数	回答	点数
非常に好き	10点	非常に好き	10点
好き	8.5点	やや好き	7.75点
やや好き	7点	どちらでもない	5.5点
どちらでもない	5.5点	やや嫌い	3.25点
あまり好きではない	4点	非常に嫌い	1点
好きではない	2.5点		
全く好きではない	1点		

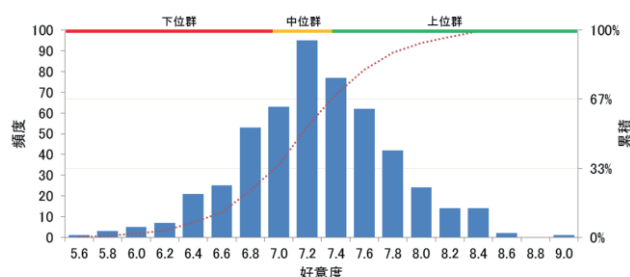


図1 好意度のヒストグラム



図2 パッケージ画像例

また、解析には調査に供した 350mL ビール缶パッケージの正面画像を用いた (図 2)。調査対象となったパッケージ画像は 509 種類であり、上位 33%(167 枚)を上位群、中位 34%(175 枚)を中位群、下位 33%(167 枚)を下位群とした。

連絡先: 篠原裕之, キリン株式会社基盤技術研究所, 神奈川県横浜市金沢区福浦 1-13-5, 080-2370-9301, Hiroyuki.Shinohara@kirin.co.jp

尚、調査は非明示であるものの、当該製品が既存製品のリニューアルである場合には消費者がその製品を既知である事が好意度に影響する事が予想された。また、当該製品が季節限定品である場合には販売する季節が好意度に影響することが予想された。そこで、当該製品が①新規製品であるか/②既存製品のリニューアルであるか/③春季限定製品であるか/④夏季限定製品であるか/⑤秋季限定商品であるか/⑥冬季限定商品であるかを追加の製品属性として用いた。

3. 手法

本研究の目的に従い、以下の検討を行った。本項ではそれぞれについて詳細を記す。

3.1 上位群下位群 2 クラス分類及び要因の可視化

限られた調査結果でモデルを構築する必要がある事から、VGG16 [Simonyan 14]をファインチューニングすることにより好意度調査上位群及び下位群について 2 クラス分類を行った。その際に分類に寄与した特徴量を可視化する事により、パッケージデザインの中で好意度に対して好影響を及ぼしている要素と悪影響を及ぼしている要素を可視化した。

具体的には、VGG16 の畳み込み層を用いて特徴量抽出を用い、得られた特徴量について Global Average Pooling [Lin 13]を行った後に新規に構築した全結合層に入力した。季節等の製品属性は 6 次元の配列として別途構築した全結合層に入力した。2 つの全結合層の出力を加算し、得られた特徴量を用いて 2 クラス分類を行った。モデル概要については図 2 に示す。また、その他解析条件詳細については表 2 に示す。

各クラス 25 枚をホールドアウト画像とし、残りの画像について 6 ラウンドのクロスバリデーションによりモデルのトレーニングを行った。学習回数は最大 250 エポックとし、その中でバリデーション画像の正答率が最も高かったエポックのモデルをそのラウンドの結果とした。

トレーニングの結果得られた 6 つのモデルを平均アンサンブルし最終モデルとした。最終モデルを用いたホールドアウト画像の分類の正答率からモデルを評価した。

寄与した特徴量の可視化については Grad-CAM [Selvaraju 17]を用いた。得られた 6 モデルそれぞれに Grad-CAM を適用し、得られた画像の平均画像を特徴量の可視化結果とした。また、得られた可視化結果の妥当性については麒麟株式会社内でビール缶パッケージをデザインしている担当者の有する知見との整合性を確認する事により評価した。

3.2 好意度の予測

3.2.1 回帰モデルの性能目標値設定

パッケージデザインの平均好意度と平均的な調査パネルの好意度との相関係数を回帰モデルの性能目標値とする事とした。しかし、好意度調査における個々のパネルの回答結果を入手出来なかった為、回答者による回答のばらつきを定量化する為の調査を行った。

今回の研究に用いたパッケージ画像の内 10 枚をビール事業に直接従事した事が無い麒麟株式会社従業員 23 名にランダムな順序で提示し、好意度について 7 段階で回答を得た。回答結果を最高 7 点、最低 1 点となるように好意度を点数化した。また、全回答者の点数の平均値をパッケージ画像の平均好意度とした。10 枚の画像に対する個々の従業員の好意度と平均好意度の相関係数を求め、得られた 23 の相関係数の平均値を回帰モデルの性能目標値とした。

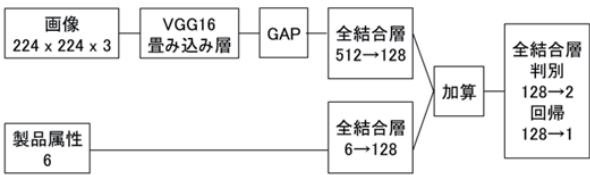


図3 モデル構造

表2 2クラス分類における解析条件詳細

使用データ	入力① 画像	224x224x3
	入力② 製品属性	6
	画像数	334枚 (167枚x2群)
	ホールドアウト画像数	50枚 (25枚x2群)
Data Augmentation	回数	9回
	最大回転角度	15度
	最大垂直移動距離	15%
	最大水平移動距離	15%
学習条件	Optimizer	Adam
	損失関数	binary cross-entropy
	エポック数	250エポック
	初期学習率	1.00E-06
	初期学習率保持エポック数	5エポック
	最終学習率	1.00E-09
	学習率変化率	0.97倍/エポック

表3 回帰における解析条件詳細

使用データ	入力① 画像	224x224x3
	入力② 製品属性	6
	画像数	509枚
	ホールドアウト画像数	60枚 (20枚x3群)
Data Augmentation	回数	9回
	最大回転角度	15度
	最大垂直移動距離	15%
	最大水平移動距離	15%
学習条件①	Optimizer	Adam
	損失関数	平均平方二乗誤差率 ただし、絶対誤差率が20%を下回った場合は0と見なす。
	エポック数	50エポック
	初期学習率	1.00E-05
	初期学習率保持エポック数	5エポック
	最終学習率	1.00E-06
	学習率変化率	0.95倍/エポック
学習条件②	Optimizer	Adam
	損失関数	実測値と予測値の相関係数を 1から減じた値
	エポック数	250エポック
	初期学習率	1.00E-06
	初期学習率保持エポック数	3エポック
	最終学習率	1.00E-08
	学習率変化率	0.98倍/エポック

3.2.2 回帰モデルの学習条件及び評価方法

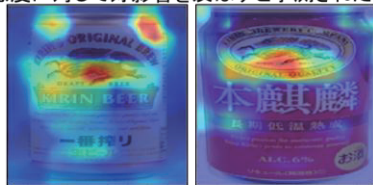
3.1 で検討したモデル構造の最終出力を変更し、回帰モデルとした。また、損失関数については平均平方二乗誤差率及び真値と予測値のピアソンの積率相関係数(以下、相関係数とする)を 1 から減じた値を用いた。その他解析条件詳細については表 3 に示す。

上位群、中位群、下位群各 20 枚をホールドアウト画像とし、残りの画像について 6 ラウンドのクロスバリデーションによりモデルのトレーニングを行った。はじめに損失関数を平均平方二乗誤差率とし、最大 50 エポックのトレーニングを行った。但し、平均平方二乗誤差率算出の際に、絶対誤差率が 20%を下回った画像の平方二乗誤差率は 0 と見なした。次に、得られたモデルについて、損失関数を真値と予測値の相関係数を 1 から減じた値に変更し、最大 250 エポックのトレーニングを行った。尚、前半の学習についてはバリデーション画像の平均平方二乗誤差率が最も低かったエポックのモデルを後半の学習に供した。後半の学習についてはバリデーション画像の真値と予測値の相関係数を 1 から減じた値が最も小さかったエポックのモデルをラウ

表4 2クラス分類結果

ホールドアウト 画像分類の正答率	
試行 1	82%
試行 2	82%
試行 3	82%
試行 4	76%
試行 5	76%
試行 6	68%
平均	77.7%
相対標準偏差	7.2%

好意度に対して好影響を及ぼすと予測された領域



好意度に対して悪影響を及ぼすと予測された領域



図3 ホールドアウト画像におけるGrad-CAMを用いた要因可視化結果例

ンドの結果とした。得られた 6 つのモデルを平均アンサンブルし最終モデルとした。

また、トレーニング時に用いる真値については、人間の評価値に直線性があるか不明であった為、①好意度 ②好意度を対数変換した値 ③好意度を 10 で除し、その値で 10 を累乗した値 の 3 条件の換算方法を比較する為、それぞれ 3 回トレーニングを行った。

最終モデルを用いてホールドアウト画像の好意度の予測を行い、②及び③についてはそれぞれの変換方法の逆関数を適用してから好意度との相関係数を算出しモデルを評価した。

4. 結果

4.1 上位群下位群 2 クラス分類及び要因の可視化

検討したモデルについて 6 回トレーニングを行った結果、ホールドアウト画像の分類の平均正答率は 77.7%、相対標準偏差は 7.2%であった(表 4)。また、その際の分類に寄与する領域の可視化結果(図 3)から、「麒麟のマークは好意度に好影響を及ぼす」「製品ロゴの英語表記は好意度に悪影響を及ぼす」といった仮説が得られた。これらはデザイン担当者の有する知見とも合致したことから、適切なモデルが構築できたと判断した。

4.2 好意度の予測

4.2.1 回帰モデルの性能目標値設定

10 枚のパッケージ画像について個々のパネルの好意度と平均好意度との相関係数を求めた結果(表 5)、その平均値は 0.611 であった。この数値を回帰モデルの性能目標値とする事とした。

表5 各被験者の好意度と平均好意度の相関係数

相関係数		相関係数	
被験者 1	-0.140	被験者 13	0.723
被験者 2	0.034	被験者 14	0.746
被験者 3	0.177	被験者 15	0.759
被験者 4	0.235	被験者 16	0.763
被験者 5	0.436	被験者 17	0.782
被験者 6	0.474	被験者 18	0.782
被験者 7	0.536	被験者 19	0.802
被験者 8	0.576	被験者 20	0.879
被験者 9	0.652	被験者 21	0.887
被験者 10	0.673	被験者 22	0.925
被験者 11	0.705	被験者 23	0.933
被験者 12	0.706	平均	0.611

表6 好意度の変換方法の違いによる相関係数の差

真値	好意度	好意度を 対数変換した値	好意度を10で除し その値で10を累乗した値
試行 1	0.467	0.501	0.667
試行 2	0.701	0.520	0.591
試行 3	0.528	0.591	0.647
平均	0.565	0.537	0.635
相対標準偏差	21%	8.8%	6.2%

表7 好意度予測結果

相関係数		平均絶対誤差率
試行 1	0.667	24%
試行 2	0.591	35%
試行 3	0.647	16%
試行 4	0.542	23%
試行 5	0.629	39%
試行 6	0.641	17%
平均	0.620	26%
相対標準偏差	7%	36%

4.2.2 好意度予測モデルの学習結果

トレーニング時に用いる真値を①好意度 ②好意度を対数変換した値 ③好意度を 10 で除し、その値で 10 を累乗した値とした条件について各 3 回のトレーニングを行いその結果を比較した(表 6)。その結果、好意度を 10 で除し、その値で 10 を累乗した値を用いた場合に最も良いモデルが得られた。そこで、当該条件のみさらに 3 回のトレーニングを行った結果、相関係数の平均値が 0.620、相対標準偏差が 7.3%であった(表 7)。性能目標値が相関係数 0.611 であったことから、適切なモデルが構築できたものと判断した。

尚、これらのモデルについてホールドアウト画像の好意度とその予測値の平均絶対誤差率を算出した結果、平均 26%、相対標準偏差 36%であった。

5. まとめ

本研究により、限られた数のパッケージデザインの好意度調査結果から作成したモデルを用いて、未調査のパッケージデザインの中から好意度に対して好影響を及ぼすと予測される要素と悪影響を及ぼすと予測される要素を可視化し、それをデザイナーにフィードバックする事と、複数のパッケージデザイン候補の中から実際の調査の対象にするパッケージデザインを絞り込むプレスクリーニングテストを行う事が可能になったと考える。一方で、好意度そのものを高精度で予測するには予測結果の誤差が大きく、更なる改善が必要であると考えられた。

参考文献

[石井 10] 石井裕明, 恩蔵直人:価値視点のパッケージ・デザイン戦略, マーケティングジャーナル 30(2), 31 - 43(2010)

- [Talebi 17] Talebi, H. Milanfar, P.: NIMA: Neural Image Assessment, *arXiv preprint arXiv*, 1709.05424, 2017
- [Kong 16] Kong, S. Shen, X. Lin, Z. Mech, R. Fowlkes, C.: Photo Aesthetics Ranking Network with Attributes and Content Adaptation, *arXiv preprint arXiv*, 1606.01621, 2016
- [Simonyan 14] Simonyan, K. Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition, *arXiv preprint arXiv*, 1409.1556, 2014
- [Lin 13] Lin, M. Chen, Q. Yan, S.: Network In Network, *arXiv preprint arXiv*, 1312.4400, 2013
- [Selvaraju 17] Selvaraju, R. Cogswell, M. Das, A. Vedantam, R. Parikh, D. Batra, D.: Grad-cam: Visual explanations from deep networks via gradient-based localization, *2017 IEEE International Conference on Computer Vision (ICCV)*, 618-626, 2017

深層学習を用いた不動産間取り図のグラフ化と物件検索への応用

Conversion of Floor Plan Images to Graph Structures using Deep Learning and Application to Retrieval

山田 万太郎

Mantaro Yamada

汪 雪婷

Xueting Wang

山崎 俊彦

Toshihiko Yamasaki

相澤 清晴

Kiyoharu Aizawa

東京大学

The University of Tokyo

The purpose of this research is to automatically convert real estate floor plan images into graph structures that reflect the floor plans. In order to do this, we recognize each room or door in the images with semantic segmentation using deep learning, and create graph structures based on their adjacencies. By this proposed method, it was confirmed that floor plan images could be converted to the graphs with the similarity of 81% with the ground truth graphs. Representing floor plans as structured representations—graph structures—makes it easy to compare and evaluate floor plans, and even search, and is expected to be applied to any systems that handles floor plans.

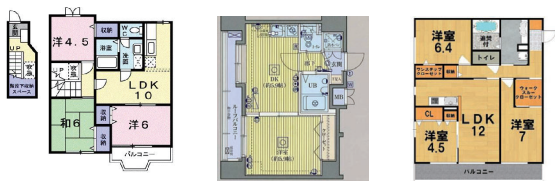


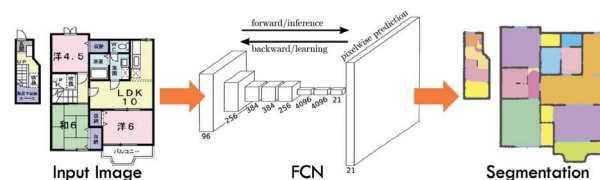
Fig. 1: 間取り図の例 [1]

1. はじめに

不動産間取り図とは、不動産物件においてその間取りを表現した図である。間取り図は、人々が物件を評価する際に非常に有用な情報となる。しかし、間取り図はその作成者や作成方法が様々であるため、表記ゆれが大きくなっている。その様子を Fig. 1 に示す。このような表記ゆれが、間取り図を直接処理することを困難にしていると考えられる。

この問題が表れている一例として、既存の不動産物件検索システムが挙げられる。既存の物件検索システムは、賃料や立地、築年数などに対しては詳細な検索ができる一方、間取りに対してはワンルームや 2LDK といった部屋のタイプでしか検索できない。これは、例えば「10 帖以上のリビングに水回りが集中している利便性の高い配置がいい」といった間取りに対する詳細な希望を持っているユーザーのニーズを満たしていない。このようなニーズに沿った検索を可能にするには、間取り図の内容をコンピュータに認識させ、その情報で検索を行う必要がある。しかし、そのために間取り図画像を直接処理するのは、上述の理由から難しい問題であるといえる。

そこで本研究では、間取り図を表記ゆれの大きい画像という形態から、グラフ構造という構造化された表現に変換することを目指す。これによって、間取りの比較や評価、さらには検索が容易になり、間取りを扱う様々な応用に繋がることが期待できる。ここでは、深層学習による画像認識を用いて間取り図の内容を認識することによって、間取り図を自動でグラフ化する手法を提案する。



(a) Step1: 間取り図の semantic segmentation [4]



(b) Step2: ルールベースでのグラフの作成

Fig. 2: 提案手法の概要図

2. 関連研究

間取り図の解析や認識を行っている研究自体は多く存在する。代表的なものとしては例えば Liu [2] らの研究が挙げられる。彼らは、壁の交差点や各部屋の領域と家具等の物体を深層学習を用いて認識することにより、間取り図をベクターイメージに変換することを試みている。

間取り図のグラフ構造への変換を行っている例としては、Yamasaki [3] らの研究が挙げられる。しかし、ここで作成されているグラフは間取り構造を正しく反映できていないという問題がある。これは単に距離の近い部屋同士を隣接と判定しているためであり、実際には行き来できない部屋の間にもエッジが存在してしまう結果となっている。また、変換したグラフについての定量的評価も行われていない。

3. 提案手法

本研究では、以下の 2 ステップで不動産間取り図のグラフ変換を実現する。その概要図を Fig. 2 に示す。

1. 間取り図画像の semantic segmentation を行う。
2. 各部屋やドアの接続関係からグラフ構造を作成する。

連絡先: 東京大学大学院情報理工学系研究科 相澤山崎研究室
〒113-8656 東京都文京区本郷 7-3-1 TEL: 03-5841-6761
mail: yamada@hal.t.u-tokyo.ac.jp

Table 1: クラスラベルの定義

ラベル	説明
wall	外壁, 内壁
tatami	和室
west	洋室
dk	ダイニング, キッチン, リビング
wc	トイレ
bath	浴室
washing	洗面所, 脱衣所
entrance	玄関
balcony	バルコニー, ベランダ, テラス
rouka	廊下
stairs	階段
cl	クローゼット, 押入れ, 下駄箱
doors	開き戸, 引き戸, 折戸, 窓
unknown	記載のない箇所, 不明箇所

提案手法では, まず深層学習による semantic segmentation を行い, 間取り図をピクセルレベルで認識する. これによって, 画像中のどこにどのような部屋が存在するのかわかる. 本研究では, 深層ネットワークモデル FCN-8s [4] を用いてこれを行う. モデルの学習のためのデータセットとしては, 間取り図画像とそれに対応する独自生成の正解ラベルマスク 4,800 組を用意している. ラベルは Table 1 に示す計 14 種類とし, 間取りを構成する主要な要素が含まれている.

次に, 各部屋の領域を走査して接続関係を調べ, ルールベースでグラフの作成を行う. ノードの作成は, 一定以上の面積を持つ領域を抽出することで行う. エッジの作成は, 同じドアに隣接している部屋間, あるいは直接隣接している部屋間のみ行う. これにより, 実際の部屋の接続関係を正しく反映したグラフ構造を作成することができる.

4. 実験

4.1 Semantic segmentation の評価

まず, 提案手法の Step1 にあたる semantic segmentation の精度を評価する実験を行う. データセットを Train/Validation/Test data として 3,800/500/500 枚に分割し, FCN-8s で間取り図画像と正解ラベルマスクの対応関係を学習する. 学習の後, Test data に対して式 (1) と式 (2) で評価を行う. ただし, n_c はクラス数, t_i はクラス i に属するピクセルの総数, n_{ij} はクラス i に属しクラス j と予測されたピクセルの総数である.

$$\text{mean accuracy} = \frac{1}{n_c} \sum_{i=1}^{n_c} \frac{n_{ii}}{t_i} \quad (1)$$

$$\text{mean IoU} = \frac{1}{n_c} \sum_{i=1}^{n_c} \frac{n_{ii}}{t_i + \sum_{i'=1}^{n_c} n_{i'i} - n_{ii}} \quad (2)$$

実験の結果, Test data に対する平均の評価値は, mean accuracy が 90.6%, mean IoU が 84.0% となった. クラスごとの精度は Fig. 4 の通りである. これより, 面積が大きく文字表記がある洋室や和室, リビングの認識精度は高く, 面積が小さいあるいは文字表記がない領域であるドアや壁, 廊下の精度は低くなる傾向があることがわかる.

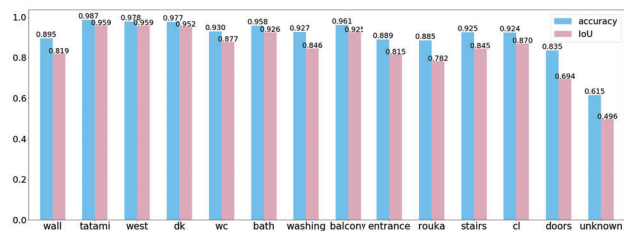


Fig. 3: クラスごとの accuracy と IoU

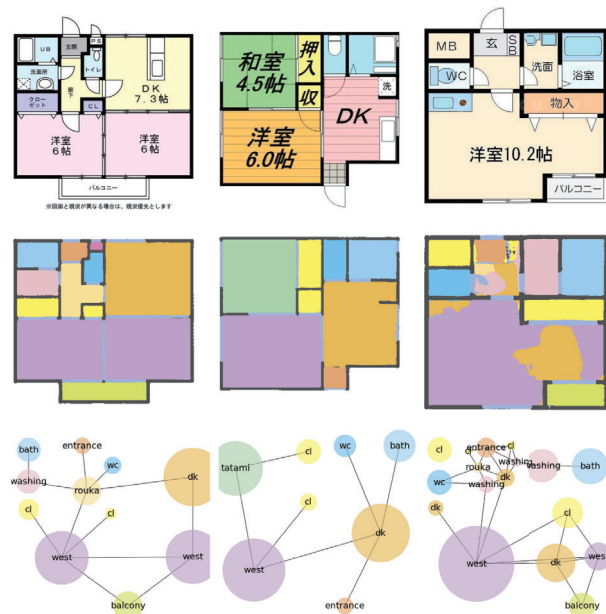


Fig. 4: Semantic segmentation(中段) とグラフ化(下段) の例

4.2 グラフ化の評価

Step2 におけるルールベースでのグラフ化をし, その精度の評価をする実験を行う. Test data に対する semantic segmentation 後にルールベースでのグラフ化を行い, その結果作成されたグラフがどれだけ正解グラフに類似しているかを調べる. グラフの類似度指標としては, 一般的で扱いやすいことから, MCS による類似度を用いる. グラフ G_1 のノード数とエッジ数の和を $|G_1|$ と表し, グラフ G_1 とグラフ G_2 の MCS(Maximum Common Subgraph: 最大のエッジ数を持つ共通サブグラフ) を $\text{MCS}(G_1, G_2)$ とすると, グラフ G_1 とグラフ G_2 の類似度は式 (3) で計算される. MCS 類似度は 2 つのグラフが全く異なる場合は 0, 全く等しい場合は 1 となり, 0~1 に正規化されている.

$$\text{sim}(G_1, G_2) = \frac{|\text{MCS}(G_1, G_2)|}{\max(|G_1|, |G_2|)} \quad (3)$$

この実験の結果, 提案手法によって間取り図から変換したグラフと正解グラフとの平均の類似度は, 0.810 となった. また, その一例を Fig. 4 の下段に示す. 右端の例は精度の低い例である. この例は semantic segmentation の精度が低く, 特に玄関付近と洋室での誤認識が目立つ. その結果, 変換されたグラフにも誤りが反映されており, 本来の間取り図には存在しない部屋のノードができてしまっていることがわかる. 一方, 左の 2 つの例は精度が高い例であり, 間取りを正しく表現したグラフ化ができています.

4.3 物件検索への応用

以上の提案した手法による間取り図のグラフ化を応用して、間取り類似物件の検索を行う。ここでは、クエリとして任意の間取り図画像を入力し、その間取りとの類似度が高い間取りを持つ物件の間取り図画像を出力するという実験を行う。クエリとしては Test data を用い、検索候補としては LIFULL HOME'S データセットからさらに 25,000 件の間取り図を用いる。この 25,000 件にはこれまでの実験で用いた 4,800 枚は含まれていない。検索の手順は以下の通りである。

1. 検索候補の間取り図 25,000 件をあらかじめグラフ化する。
2. システムにクエリとなる間取り図画像を入力する。
3. 提案手法を用いてクエリ画像をグラフ化する。
4. クエリと検索候補のグラフとの間の類似度を計算する。
5. 類似度が高い順に検索結果を表示する。

このようにして類似する間取り図を検索した結果の例を Fig. 5 に示す。Fig. 5a は 2DK で和室と洋室をひとつずつ持つ間取り図をクエリとした例である、検索結果を見ると、クエリと同様の間取り構造を持っていることがわかる。和室と洋室が存在するだけでなく、その繋がりが方や付随する収納の有無、水回りへのアクセスまでもが一致していることがわかる。Fig. 5b はクエリを 2 階建ての間取り図とした例である。検索結果も同様に 2 階建ての構造で、水回りが 1 階に、バルコニー付きの洋室が 2 階にという配置も一致していることがわかる。特に 3 位までの間取り図はクエリによく一致している。4 位と 5 位は 2 階部分の部屋がひとつ多いが、その点以外は概ね一致しているといえる。

一方で、Fig. 5c は類似の間取り図を検索できなかった例である。クエリとなる間取りは、玄関が LDK と洋室の両方に繋がっていて、中央に位置する洗面所を介して LDK と洋室を行き来できるという特徴的な間取りである。しかし、検索結果にそのような間取りを含む物件は現れなかった。この場合、クエリとなっている間取り構造が極めてユニークであったことが原因である可能性もある。

以上の結果から、間取りをグラフで表現しその類似度での検索を行うことで、間取り図のスタイルやカラーリングは異なっていたとしても間取りの構造が類似している物件を検索できる例があることが確認できた。

5. まとめ

本研究の目的は、不動産間取り図の画像を自動でグラフ構造に変換することである。本研究で提案した手法により、間取り図画像をおよそ 8 割の精度でその間取りを反映したグラフ構造に変換することができることを示した。さらに、それを用いることによって任意の間取り図に対してそれと構造的に類似している間取りを検索する仕組みを実現できることを確認した。本研究の成果は以下のようにまとめることができる。

- 間取りを反映したグラフへの変換手法を提案したこと
- 変換したグラフについて定量的な評価を行ったこと
- 応用として類似の間取りを検索できる例を示したこと

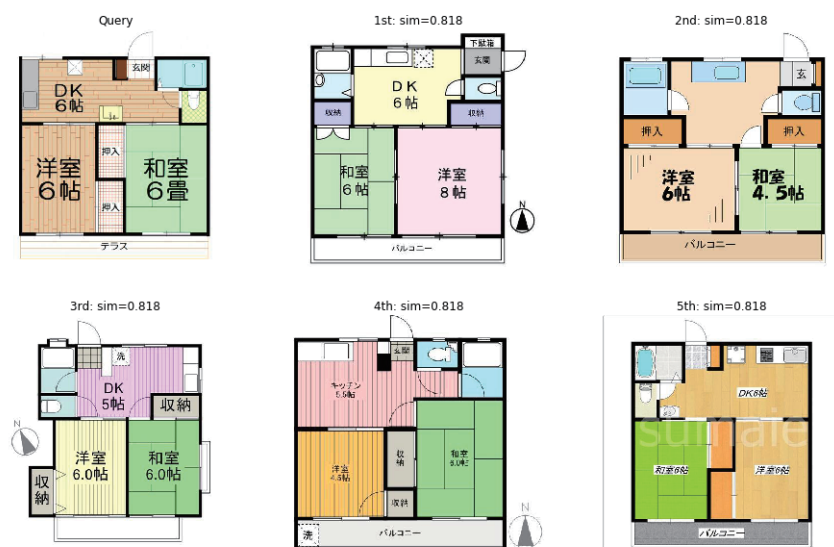
一方で、類似物件検索への応用については多くの課題が残されており、例えばその評価がひとつである。さらに、ユーザーの間取りに対する要望をシステムに伝えるためのインタフェースをどう設計するかは重要な課題であるといえる。

謝辞

本研究データの一部は、NII-IDR より提供されている LIFULL HOME'S データセットを用いている。

参考文献

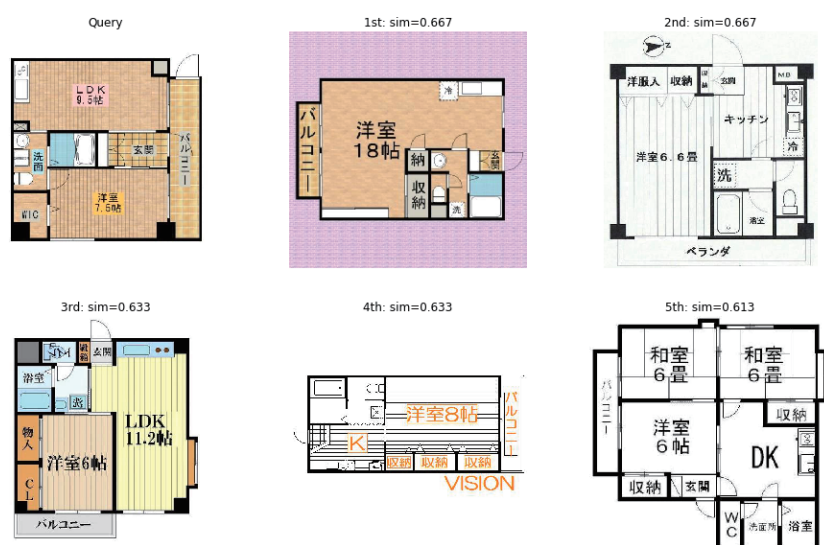
- [1] 国立情報学研究所 IDR 事務局. LIFULL HOME'S データセット, 2010-2019.
- [2] Chen Liu, Jiajun Wu, Pushmeet Kohli, and Yasutaka Furukawa. Raster-to-Vector: Revisiting Floorplan Transformation. In *IEEE International Conference on Computer Vision*, pp. 2214–2222, 2017.
- [3] Toshihiko Yamasaki, Jin Zhang, and Yuki Takada. Apartment Structure Estimation Using Fully Convolutional Networks and Graph Model. In *ACM Workshop on Multimedia for Real Estate Tech*, pp. 1–6, 2018.
- [4] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully Convolutional Networks for Semantic Segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2015.



(a) 成功例 (1)



(b) 成功例 (2)



(c) 失敗例

Fig. 5: 類似間取り図検索の例
 左上が検索クエリとなる間取り図、右へ順に 5 位までの検索結果。

[3P4-J-7] Agents: social multiagents

Chair:Naoki Fukuda Reviewer:Jun Ichikawa

Thu. Jun 6, 2019 3:50 PM - 5:30 PM Room P (Front-left room of 1F Exhibition hall)

[3P4-J-7-01] Explainable Compromising Algorithm based on Constraint Relaxation for Automated Negotiating Agents

○Shun Okuhara^{1,2}, Takayuki Ito² (1. Fujita Health University, 2. Nagoya Institute of Technology)

3:50 PM - 4:10 PM

[3P4-J-7-02] An allocation strategy with deep reinforcement learning for efficient task processing in multi agent system

○Genki Matsuno¹, Sho Tanaka², Hiroki Hara², Syunyo Kawamoto², Syo Shimoyama², Takashi Kawashima², Daisuke Tsumita², Yasushi Kido¹, Osamu Hashimoto¹, Tomohiro Takagi² (1. Skydisc, Inc., 2. Meiji University)

4:10 PM - 4:30 PM

[3P4-J-7-03] Omoiyari as Filling Gaps Making Collective Adaptation

○Yoshimiki Maekawa¹, Fumito Uwano¹, Eiki Kitajima¹, Keiki Takadama¹ (1. The University of Electro-Communications)

4:30 PM - 4:50 PM

[3P4-J-7-04] Investigation of online simulation method of social consensus formation

○Yasuko Kawahata¹, Akira Ishii², Takuya Ueoka¹ (1. Gunma University, 2. Tottori University)

4:50 PM - 5:10 PM

[3P4-J-7-05] Estimation of agent's rewards with multi-agent maximum discounted causal entropy inverse reinforcement learning

○Keiichi Namikoshi¹, Sachiyo Arai¹ (1. Chiba University)

5:10 PM - 5:30 PM

自動交渉エージェントのための 制約緩和を用いた説明可能な妥協アルゴリズム

Explainable Compromising Algorithm based on Constraint Relaxation for Automated Negotiating Agents

奥原 俊^{*1,*2}
Shun Okuhara

伊藤 孝行^{*1}
Takayuki Ito

^{*1} 名古屋工業大学大学院情報工学専攻
Dept. of Computer Science, Nagoya Institute of Technology

^{*2} 藤田医科大学医療科学部
School of Health Sciences, Fujita Health University

This paper presents an explainable concession process based on constraint relaxation in multi-agent negotiation. Automated negotiation has been studied widely and is the promising technology for the future smart city where multiple heterogeneous agents, like driver-less cars, are conflicting and collaborating. There are a lot of studies on negotiating agents including international competitions. The problem is that most of the proposed negotiating agents employ ad-hoc conceding process, where basically they are adjusting threshold to accept their opponents offers. Because it is just adjusting a threshold, it is very difficult to show how and what the agent conceded even after agreement. In this paper, we propose an explainable concession process by using a constraint relaxation process. Here, an agent changes its belief not to believe some constraint so that he/she can accept its opponent offer. Our experimental results demonstrate that our method can work effectively.

1. はじめに

自動交渉エージェント [Ito 07, Jennings 01] の重要性が高まっており広く研究が進められている [Bai 17, Fukuta 16, Fujita 15, Marsa-Maestre 14, Ito 13, Ito 11, Ito 10, Ito 09, Ito 08]. 社会において知的な処理を自律的に行う異種のシステム（エージェント）が現実実現されつつある。複数のエージェント間の競合が発生し、自動交渉によって合意を自動的に得るような仕組みが現実必要とされ得る。自動交渉エージェントの研究は、マルチエージェントシステムの分野で広く行われており、特に 2010 年前後から国際ワークショップや国際競技会が開催され、次世代の重要な技術である。

自動交渉エージェントの研究では、エージェントは自分の好みを秘匿したまま、交渉プロトコルに基づき交渉し、合意を得る。エージェントの好みは、多論点の効用関数（多属性効用関数）で表される。交渉プロトコルは、様々なプロトコルが提案されているが、仲介型のプロトコル、繰り返し型のプロトコル、提案交換型プロトコルが一般的である。

自動交渉エージェントの研究のテストベッドとして、自動交渉競技会 ANAC(Automated Negotiating Agents Competition) が 2010 年から開催されている。ANAC では多論点の効用関数と提案交換型プロトコルが採用されており、毎年様々なルールの拡張や修正を行い、様々な環境でのエージェントの交渉戦略が提案されている。

課題は、妥協プロセスの説明可能性である。交渉において、エージェントは自分の利益ばかりを考えていると合意に到達できないため、いかに妥協するかという妥協戦略が重要である。既存の自動交渉エージェントのほとんどが、閾値の上げ下げのみを用いたアドホックなプロセスであった。そのため、どのように妥協したかという説明が難しいという課題があった。

本研究では、本課題を解決するために、妥協プロセスを制約の緩和プロセスとして提案する。制約とは効用の基本単位を表すものである。つまり、本研究では、エージェントの効用空間を、論点とその論点を満たす制約の集合とする。制約は満たされれば価値がある。例えば、車を買うときの論点 (Issue) は、価格、色、タイプなどがある。これらの論点は制約によって結び付けられる。つまり、タイプがスポーツカーであれば、色は赤という制約や、タイプがセダンであれば、色は白という制約である。制約は満たされることで価値を生み出すが、満たされない場合は価値を生み出さない。

また、本研究では、共有論点と個人論点を仮定する。つまり、合意は、共有論点についてエージェント同士が同じ値を持つことである。個人論点については、このエージェントがなるべく自分の効用が高くなるように選択することができる。エージェントは、共有論点について相手のエージェントと同じ値になるようにしながら、個人論点については、制約をなるべく多く満たすように値を決めていくというトレードオフを解決する必要がある。

このトレードオフを解決するため、エージェントは妥協プロセスを行う。妥協プロセスにおいて、エージェントは、制約の集合の中から、制約を一つずつ取り除いていくことで、共有論点についての値の取れる範囲を調整する。制約を取り除くことを制約緩和と呼ぶ。

本稿は、2 章で自動交渉エージェントと交渉プロトコルについて述べる。3 章で、新たに提案する制約緩和に基づく妥協アルゴリズムについて述べる。4 章で評価実験とその評価について述べる。5 章で関連研究と本研究の違いを明らかにし、6 章で本稿をまとめる。

2. 自動交渉エージェント

2.1 効用ハイパーグラフ

エージェントは複雑な効用空間を持つものとする [Ito 07]. 複雑な効用空間の表現方法は様々な方法が提案されている [Robu 05, Robu 08, Aydogan 15]. 本論文では、論点間の依存関係に注目して表現するために、ハイパーグラフによる表現 [Hadfi 14b, Hadfi 14a] を用いる。ハイパーグラフとはグラフ

連絡先: 奥原俊, 豊明市沓掛町田楽ヶ窪 1 番地 98 藤田医科大学 大学 6 号館 507 室, 0562-93-2628, okuhara@itolab.nitech.ac.jp

連絡先: 伊藤孝行, 名古屋工業大学大学院情報工学専攻, 名古屋市昭和区御器所町 名古屋工業大学 4 号館 702, 052-735-7968, ito.takayuki@nitech.ac.jp

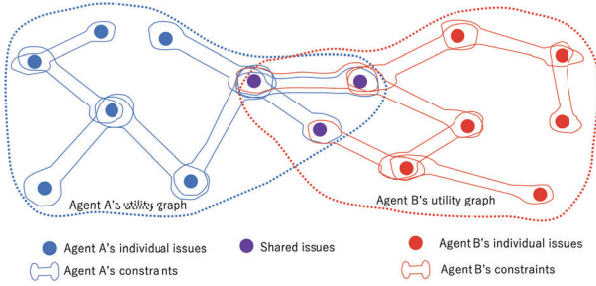


図 1: エージェント間の論点の共有と効用グラフ

を一般化した数学的表現で、エッジが複数のノードを連結できる。ハイパーグラフを用いた効用空間を、効用ハイパーグラフと呼ぶ。ここでは、ノードを論点、および、エッジを制約として考える。

エージェント i の効用空間 U_i は、ハイパーグラフ (I, C) で表され、 I は論点集合 (ノード)、 C は制約集合 (エッジ) である。各論点 $I_i \in I$ は、ある決められた範囲 D_i 内の論点値 (Issue Value) をもつ。例えば、車を購入する場合の論点の一つ「色」は、「赤、青、緑」という範囲のどれかの論点値をもつ。制約 $C_j \in C$ は $(v_{C_j}, \phi_{C_j}, \delta_{C_j})$ で表される。 v_{C_j} は制約 C_j の価値を表す。 ϕ_{C_j} は制約 C_j が連結している論点の集合である。したがって、 $\phi_{C_j} \subset I$ である。 δ_{C_j} は、範囲 (range) の集合であり、 $\delta_{C_j} = \{range_{C_j}(I_i) : I_i \in \phi_{C_j}\}$

ここで、制約 C_j が満たされる条件は以下の通りである。論点 I_i がとる値を x_{I_i} とする。 C_j が満たされた場合、 C_j を持つエージェントはその価値 v_{C_j} を得る。

$$C_j = \begin{cases} \text{satisfy} & \text{if } x_{I_i} \in range_{C_j}(I_i) \quad \forall I_i \in \phi_{C_j} \\ \text{unsatisfy} & \text{otherwise} \end{cases}$$

図 1 にエージェントの効用グラフと論点の共有についての例を示す。ここでは 2 つのエージェントが、効用グラフを持つと同時に 3 つの論点を共有している。それぞれのエージェントは、各論点を結ぶ制約を持つ。論点は値をとる。制約は、結んでいる論点の値が、制約としての範囲 (range) に含まれる時充足する。制約が充足すると、エージェントはその制約から価値を得ることができる。

仮定 1. 制約は充足しにくい制約ほど価値が高い

仮定 1 に従い、本論文の実験では以下の 2 つを仮定している。

- より広い値域 ($range_{C_j}$) を持っている制約の方が充足しやすいので、価値は低い。一方より狭い値域を持っている制約は充足しにくいので、価値は高い。
- さらに、個人制約より、相手との合意が必要な共有制約のほうが価値が高い。

2.2 交渉プロトコル

今回は、妥協アルゴリズムに焦点を置くため、交渉プロトコルは出来るだけ単純なものを採用する。本稿では繰り返し同時提案プロトコルを提案する。すなわち、毎回、各エージェントが同時に提案を提出し、提案が互いにとって受け入れられるなら合意する。そうでなければ、次の提案を行う。という単純なプロトコルである。

自動交渉の研究分野では交互提案プロトコル [Rubinstein 82] も採用されるが、どちらが先手で提案を出すかにより、妥協の

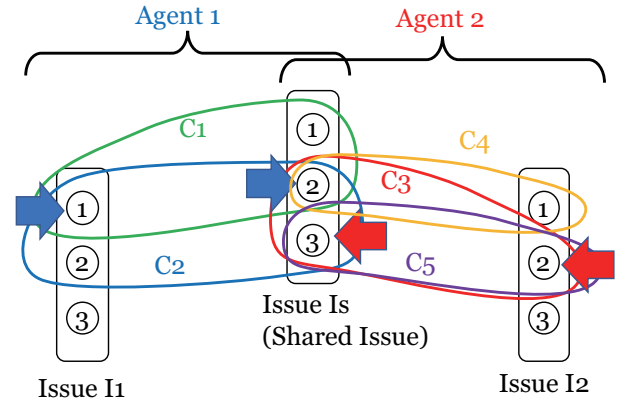


図 2: 制約緩和による合意の例 1 : 初期設定

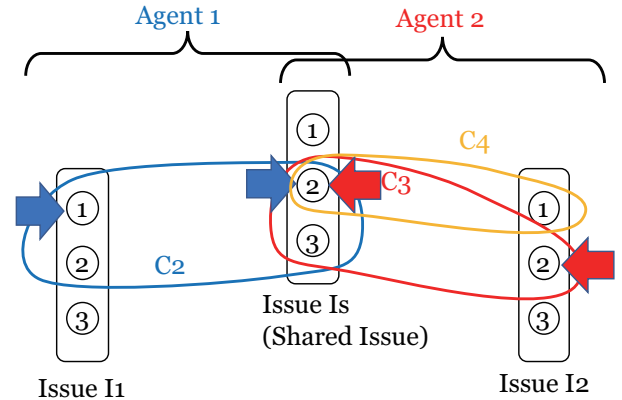


図 3: 制約緩和による合意の例 2 : 緩和により合意

戦略が変わってしまうため、今回は単純な繰り返し同時提案プロトコルを採用した。交互提案プロトコルへの拡張は今後の課題である。

3. 制約緩和に基づく説明可能な妥協プロセス

3.1 説明可能な妥協プロセス

本章では、制約緩和に基づく妥協プロセスを示す。制約緩和とは、満たすべき制約の数を少なくすることで、自分が取り得る効用 (価値) の総和を少なくすることをいう。

既存の研究のように、アドホックに閾値を調整することで、妥協を行う場合、なぜその値で合意したのか、を説明することができない。本研究では、満たすべき制約を少なくする、具体的には、ここでは制約を考慮に入れないことにすることで、どの制約を考慮に入れ、どの制約を考慮に入れないことで、合意できたか、という妥協の説明が可能となる。基本的には、まず信じている (IN) 制約と信じていない (OUT) 制約を分ける。初期は全ての制約を IN としており、緩和した制約を OUT にする。

図 2 と図 3 に本論文で提案する妥協プロセスの簡単な例を示す。図 2 で示すように、Agent 1 は Issue I_1 と Issue I_s を持つ。Issue I_s は、共有論点である。Agent 2 は Issue I_2 と Issue I_s を持つ。各 Issue は値を 1, 2, および 3 の三つを持つとする。Agent 1 は、制約 C_1 と C_2 を持つ。両方を満たした方が効用が高くなるため、初期の最適解は Issue I_1 については 1, および I_s については 2 となる。一方 Agent 2 は、制約 C_3 , C_4 , および C_5 をもつ。同様に最適解は Issue I_s につ

いては3, I_2 については2となる。図2の状態では、共有論点 I_s について解が異なるため、合意はできていない。

そこで、各エージェントは制約の一つ減らす (IN から OUT に変える) ことで妥協プロセスを行う。ここでは例えば、Agent 1 は制約 C_1 を OUT にしたとする。Agent 2 は制約 C_5 を OUT にしたとする。すると Agent 1 の Issue I_s の値は2のままであるが、Agent 2 の Issue I_s の値も2に変化する (より多くの制約を満たす値の方が効用が高くなる)。これにより Agent 1 と Agent 2 は合意ができる。

妥協においてどの制約を OUT にしたか (信じないようにしたか) ということがわかるため、たんに閾値を下げるのではなく、どのような制約を外したかという説明が可能となる。

3.2 制約緩和アルゴリズム

様々な制約緩和が考えられるが、ここでは以下の4つの方法を提案する。初期は全ての制約を IN としており、緩和した制約を OUT にする。

- ランダム制約緩和: IN の制約の中からランダムに制約を選択し OUT にする。
- 価値に基づく制約緩和: IN の制約の中からもっとも価値の低い制約を選択し OUT にする。
- 距離に基づく制約緩和: IN の制約の中から共有論点からもっとも距離の遠い制約を選択し OUT とする。ここで距離とは、共有論点からの連結する制約の数とする。
- 価値と距離に基づく制約緩和: IN の制約の中から共有論点からもっとも距離の遠い制約のうち最も価値の低い制約を選択し OUT とする。

4. 評価実験

4.1 実験設定

妥協のための制約緩和アルゴリズムの性能比較のために実験を行う。以下に実験設定を示す。

- 参加するエージェントは2エージェントとする。
- 共有論点は1つとする。
- 各エージェントの論点の数は x とする。
- 各論点を包含する制約を1つ以上持つものとする。つまり制約のない論点はないものとする。
- 各論点を包含する制約の数は y とする。

以上の設定は、多論点かつ論点間の依存があるエージェント間交渉の設定としては、論点がたくさんあり、それらが少ない制約で連結しているような場合である。

4.2 結果と考察

いくつかの設定で結果が得られているが、ここでは2つの結果について図4と図5に示す。ここでは、価値に基づく制約緩和 (min), ランダム制約緩和 (random), 距離に基づく制約緩和 (distance), および、価値と距離に基づく制約緩和 (distance+min) について比較を行なった。

図4では、各エージェントの論点の数は $x = 1$ とする。各論点を包含する制約の数は $y = 16$ とする。

図5では、各エージェントの論点の数は $x = 2$ とする。各論点を包含する制約の数は $y = 11$ とする。

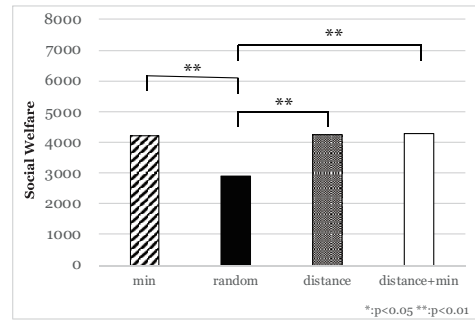


図4: 実験結果: $x = 1$ および $y = 16$

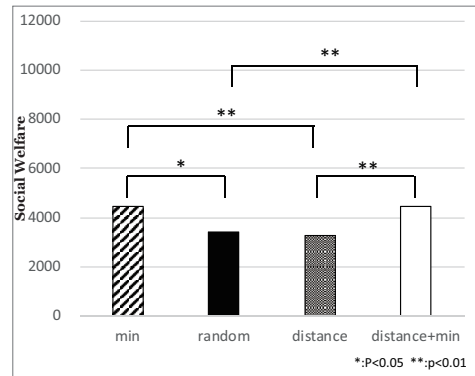


図5: 実験結果: $x = 2$ および $y = 11$

どちらの場合もランダム制約緩和 (random) に対して、価値に基づく制約緩和 (min), ランダム制約緩和 (random), 距離に基づく制約緩和 (distance), および、価値と距離に基づく制約緩和が、有意に社会的余剰を多く獲得することができる。

5. 関連研究

本節では関連研究と本研究の提案の差異を示す。

妥協プロセスが提案されたのは、Klein らの論文 [Klein 03] である。主な主張は、論点が独立な多論点効用空間を前提とした交渉では、パレートフロントにおいて、エージェントが徐々に妥協するのが合理的であるが、論点が依存している場合は効用空間が複雑になり、単純にはパレートフロントを見出せないの単純ではないということを指摘している。一つの方法として SA による合意点探索プロトコル (暗黙に妥協を仮定している) を提案している。さらに、Payemann ら [Faratin 98] は様々な妥協関数について分析している。ANAC 競技会 [Baarslag 12] は 2010 年から開催されている。相手のオファーから統計的に相手が受理しうる提案を推定し提案しながら、時間による割引を考慮した閾値の調整によって提案を受け入れるという方式が、一般的である。例えば、AgentK [Kawaguchi 11] は、ANAC2010 の優勝エージェントである。AgentK は、相手の提案履歴から相手の効用空間と交渉に対する姿勢を推定する。相手が友好的であれば譲歩し、相手が敵対的であれば一定以上は譲歩しないという、ANAC の基本的な譲歩戦略の先駆けとなった戦略である。Fawkes [Baarslag 14] は、ANAC2013 の優勝エージェントである。Fawkes は、相手の提案履歴に基づいて、離散ウェーブレット予測を用いて、最適な譲歩を推定する。以上の研究では、基本的には、妥協のプロセスについて、

閾値を変化させる方法が多く、何によって妥協が実現されているかという説明可能性は実現できない。また、著者らが知る限り、多論点の効用関数を仮定している自動交渉エージェントに関する研究で、妥協の説明可能性について言及している研究は存在しない。

マルチエージェント環境における、一貫性の分類について DTMS(Distributed Truth Maintenance System) を提案した論文 [Bridgeland 90] で提案されている。彼らは、分散整合性を Inconsistent, Local-Consistency, Local-and-Shared-Consistency, Global Consistency に分類している。本研究で、合意するという事は、各エージェントの中で整合しつつ、共有する論点につても整合するという事であり、Local-and-Shared-Consistency に当たる。本論文で提案する妥協の方法は、Local-and-Shared-Consistency を得るための方法の一つである。ただし、本論文で扱う制約グラフは効用空間を表現しているが、DTMS では好みを表現していない。

6. おわりに

本論文では、自動交渉エージェントのための説明可能な妥協プロセスを示した。既存のほとんどの自動交渉の妥協プロセスは、閾値のアドホックな調整がほとんどだが、ここでは制約を一つずつ削る (IN から OUT にする) ことで制約緩和を行い説明可能とした。本研究の新しい貢献は以下の点にある。(1) 制約のグラフ構造に基づく説明可能な妥協プロセスを新たに提案した。(2) 自動交渉において、共有論点と個人的論点を区別したモデルを新たに提案した。(3) 妥協プロセスにおいて、影響と価値に基づく制約緩和プロセスを提案し、その効果を示した。

参考文献

- [Aydogan 15] Aydogan, R., Baarslag, T., Hindriks, K., Jonker, C., and Yolum, P.: Heuristics for using CP-nets in utility-based negotiation without knowing utilities, *Knowledge and Information Systems*, Vol. 45, pp. 357–388 (2015)
- [Baarslag 12] Baarslag, T., Fujita, K., Gerding, E., Hindriks, K., Ito, T., Jennings, N. R., Jonker, C., Kraus, S., Lin, R., Robu, V., and Williams, C.: The First International Automated Negotiating Agents Competition, *Artificial Intelligence Journal (AIJ)* (2012), (to appear)
- [Baarslag 14] Baarslag, T.: What to bid and When to stop, Master's thesis, Delft University of Technology (2014)
- [Bai 17] Bai, Q., Ren, F., Fujita, K., Zhang, M., and Ito, T.: *Multi-agent and Complex Systems*, Springer (2017)
- [Bridgeland 90] Bridgeland, D. and Huhns, M.: Distributed Truth Maintenance, pp. 72–77 (1990)
- [Faratin 98] Faratin, P., Sierra, C., and Jennings, N. R.: Negotiation Decision Functions for Autonomous Agents, *Int. Journal of Robotics and Autonomous Systems*, Vol. 24, No. 3 - 4, pp. 159–182 (1998)
- [Fujita 15] Fujita, K., Ito, T., Zhang, M., and Robu, V.: *Next Frontier in Agent-based Complex Automated Negotiation*, Vol. 596, Springer (2015)
- [Fukuta 16] Fukuta, N., Ito, T., Zhang, M., Fujita, K., and Robu, V.: *Recent Advances in Agent-based Complex Automated Negotiation*, Vol. 638, Springer (2016)
- [Hadfi 14a] Hadfi, R. and Ito, T.: Modeling Complex Non-linear Utility Spaces Using Utility Hyper-Graphs, in *Modeling Decisions for Artificial Intelligence*, pp. 14–25, Springer (2014)
- [Hadfi 14b] Hadfi, R. and Ito, T.: On the Complexity of Utility Hypergraphs, in *Proceedings of the Seventh International Workshop on Agent-based Complex Automated Negotiations(ACAN2014)* (2014)
- [Ito 07] Ito, T., Hattori, H., and Klein, M.: Multi-issue Negotiation Protocol for Agents : Exploring Nonlinear Utility Spaces, in *Proc. of 20th International Joint Conference on Artificial Intelligence (IJCAI-2007)*, pp. 1347–1352 (2007)
- [Ito 08] Ito, T., Hattori, H., Zhang, M., and Matsuo, T.: *Rational, robust, and secure negotiations in multi-agent systems*, Vol. 89, Springer (2008)
- [Ito 09] Ito, T., Zhang, M., Robu, V., Fatima, S., and Matsuo, T.: *Advances in agent-based complex automated negotiations*, Vol. 233, Springer (2009)
- [Ito 10] Ito, T., Zhang, M., Robu, V., Fatima, S., Matsuo, T., and Yamaki, H.: *Innovations in Agent-Based Complex Automated Negotiations*, Vol. 319, Springer (2010)
- [Ito 11] Ito, T., Zhang, M., Robu, V., Fatima, S., and Matsuo, T.: *New trends in agent-based complex automated negotiations*, Vol. 383, Springer (2011)
- [Ito 13] Ito, T., Zhang, M., Robu, V., and Matsuo, T.: *Complex automated negotiations: Theories, models, and software competitions*, Springer (2013)
- [Jennings 01] Jennings, N. R., Faratin, P., Lomuscio, A. R., Parsons, S., Wooldridge, M., and Sierra, C.: Automated Negotiation: Prospects, Methods, and Challenges, *Group Decision and Negotiation*, Vol. 10, pp. 199–215 (2001)
- [Kawaguchi 11] Kawaguchi, S., Fujita, K., and Ito, T.: Compromising strategy based on estimated maximum utility for automated negotiation agents competition (ANAC-10), in *Modern Approaches in Applied Intelligence*, pp. 501–510, Springer (2011)
- [Klein 03] Klein, M., Faratin, P., Sayama, H., and Bar-Yam, Y.: Negotiating Complex Contracts, *Group Decision and Negotiation*, Vol.12,No.2, pp. 58–73 (2003)
- [Marsa-Maestre 14] Marsa-Maestre, I., Lopez-Carmona, M. A., Ito, T., Zhang, M., Bai, Q., and Fujita, K.: *Novel insights in agent-based complex automated negotiation*, Vol. 535, Springer (2014)
- [Robu 05] Robu, V., Somefun, D. J. A., and Poutré, J. L.: Modeling complex multi-issue negotiations using utility graphs, in *AAMAS '05: Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, pp. 280–287, New York, NY, USA (2005), ACM
- [Robu 08] Robu, V. and La Poutré, H.: Constructing the Structure of Utility Graphs Used in Multi-Item Negotiation through Collaborative Filtering of Aggregate Buyer Preferences, in *Rational, Robust, and Secure Negotiations in Multi-Agent Systems*, pp. 147–168, Springer (2008)
- [Rubinstein 82] Rubinstein, A.: Perfect Equilibrium In A Bargaining Model, Vol. 50, No. 1, pp. 97–109 (1982)

マルチエージェントシステムにおける効率的タスク処理のための深層強化学習を用いた割り当て戦略

An allocation strategy with deep reinforcement learning for efficient task processing in multi agent system

松野 元樹^{*1}
Genki Matsuno

田中 翔^{*2}
Sho Tanaka

原 拓希^{*2}
Hiroki Hara

川本 峻頌^{*2}
Syunyo Kawamoto

下山 翔^{*2}
Syo Shimoyama

川島 崇^{*2}
Takashi Kawashima

積田 大介^{*2}
Daisuke Tsumita

城戸 康^{*1}
Yasushi Kido

橋本 司^{*1}
Osamu Hashimoto

高木 友博^{*2}
Tomohiro Takagi

^{*1}株式会社スカイディスク
Skydisc, Inc.

^{*2}明治大学理工学部ウェブサイエンス研究室
Meiji University, Graduate School of Science and Technology, web science lab

In this study, it was considered how to make matching between task required resource and member capability that will be applied into multi-agent systems. Supported by Reinforcement Learning strategy with deep Neural Network technique, a modern solution was conducted accompanied with standard baseline methods and evaluated from several suggestive viewpoints. According to the numerical experiments, it is elucidated that RL strategy has some advantages when targeting on both execution time duration and accuracy of combination matching.

1. はじめに

近年の情報・ネットワークシステム業界においては、各コンポーネント・サービスをできるだけ独立細分化したのち階層的に構造化することでシステム各部の変更容易性・疎結合性を高めることを狙うマイクロサービスの考え方が一般的になりつつある。細分化された多数のサービスは多様性を持つエージェント群(マルチエージェント)[1]であると捉えなおすことができる。ここで言うエージェントとは、我々に代わってタスクを処理したり必要な機能を果たす端末、というくらいの意味である。マイクロサービス構成を採用すると個々の機能を独立して作り込めるためシステム開発速度の加速が期待されるが、管理すべきエージェントの数が増えるのでシステムの制御コスト・インフラ基盤への負担が増大するという問題がある。同様に、Internet of Things(IoT) システム [2] における各機器もまたエージェントと捉えることが可能であり、機器の数・多様性の増加に伴って管理コストが増える点でも同じである。

また、負荷分散の文脈 [3, 4, 5] から考えて、マルチエージェントシステム内で個性を持つ各々のエージェントに適切な負荷を持つタスクを割り振ること [6, 7] はきわめて挑戦的な課題である。タスクとメンバーの間にありうるすべての可能な組み合わせを試すため、あるいは可能な限りのエージェント資源を実際のタスク処理に使うためには、タスクを振るというタスクに従事するエージェントの数はできるだけ少ないことが望ましい。しかしながら、タスクを振るというタスクはシステム内の1 エージェントに一任するには重すぎるタスクである。なぜならば、システムの大きさが大きくなることに伴ってタスクとメンバーの間にありうる組み合わせの数が指数関数的に増大するからである。また、内在するエージェントを介さずシステム外から直接制御する中央集権的なやり方はマイクロサービスのシステム設計思想に反しており、制御機構も含めて考えた際のシステム全体を複雑化してしまう。したがって、複数のリーダーエージェントを配置しそれぞれのチームにおいてメンバーエージェントへの負荷分散を図るボトムアップ型の戦略を考慮するのが自然に思われる。

これは、組み合わせ最適化の言葉でいうと、組み合わせ爆発

が起こる探索不可能なほど巨大な状態空間を適当な部分空間に分割して取り扱うことに相当する。この操作により、チームメンバーの再構成を行う機構を採用しない限り、ありうる組み合わせを網羅的に探索することはできなくなるが、現実的な計算時間で問題を解くことができるという大きな利点がある。

我々は、タスク処理を行うマルチエージェントシステムのチーム形成に着目した論文 [8](以下ベース論文と呼ぶ) の定常状態を対象に、深層強化学習を用いてタスク実行に必要なリソースとメンバー能力のマッチング問題の解決を試みた。詳しくは2章に譲るが、本稿で取り扱うのはベース論文の問題設定を抽出・簡略化したものであり、リーダーエージェントは固定メンバーを持つ1つのチーム内でどのメンバーがタスク内のどのサブタスクを担当するかを選択する(1対1対応させる)という課題に取り組む。また、深層強化学習を用いた提案手法の性能を測定するため、基準として同じ問題をランダム法あるいはソルバーによる解法を用いて解きなおし、すべての手法に共通する有意な計量を独自に設計・計測した。

以降の章では、不足のない状態空間の設定法、深層強化学習を用いる利点についてや効率的なメンバー能力の運用がなされているかを測る計量などについて議論する。2章においては、問題設定やシステム構成、論文中で使用する数式について詳細に説明する。3章では、実施された実験とその結果の詳細を示し、比較手法であるランダム法あるいはソルバーによる解法との違いについても述べる。4章では、実験の結果得られた成果を整理し直し、今後の課題についても言及する。

2. 問題設定

我々は、上記マッチング問題をシステム内エージェントの一部が実行すべきタスクのひとつとして定義する。便宜上、マッチング問題を扱うエージェントをリーダーエージェントと呼び、リーダーからサブタスクを与えられて働くメンバーと区別することにする。リーダーエージェントは複数のタスクが詰まったタスクキューからひとつのタスクを選び取り、自分のチームメンバーにタスク内に含まれるサブタスクを適切に分配しなければならない。タスクは複数のサブタスクから構成されており $T = \{s_1, s_2, \dots, s_m\}$, サブタスクは実行に必要なリソース値を持つ $s_i = \{r_1^s, r_2^s, \dots, r_n^s\}$. サブタスク実行に必要なリソー

連絡先: <https://skydisc.jp/>

スの値 r^s は集合 $E = \{0, 3, 4, \dots, k'\}$ から適宜選択されるが、簡単のため1つの成分のみ有限の値を持ち、その他の成分は0とする。メンバーもまたそれぞれ自身の処理能力を表すリソース値を持つが $M_j = \{r_1^M, r_2^M, \dots, r_n^M\}$ 、リソース値 r^M は集合 $E' = \{0, 1, \dots, k\}$ の中から設定される。リーダーエージェントのタスクは、チームとしてのタスク処理時間

$$t = \max_{s_i \in T, p \in n} \left\{ \text{ceil} \left(\frac{r_p^{s_i}}{r_p^{M_j}} \right) \right\} \quad (1)$$

が最小となるような s_i と M_j の組み合わせを発見することである。なお、タスク処理時間は整数化したのち、tick という単位で数えることにする。さらにここで、それぞれタスク内サブタスクの数およびチーム内メンバーの数を表す記号 $\#s \equiv m$, $\#M \equiv l$ を導入しておく。本研究では、変化しないパラメータとして $n = l = k = 6, k' = 12, m \in \{2, 3, \dots, 6\}$ を採用している。

2.1 深層強化学習 (deep Q learning)

強化学習の枠組み内において、エージェントは状態 S と行動 a が張る空間の中を即時報酬 $R(S, a)$ を受け取り記録しながら探索を繰り返すことで、ある状態 S においてどの行動 a を選択すると将来に渡っての累積報酬 (以降、価値と呼ぶ) が高くなると期待されるかを経験的に学習することができる。Q 学習と呼ばれる学習法においては

$$Q(S, a) \leftarrow (1 - \alpha)Q(S, a) + \alpha (R(S, a) + \gamma \max_{a'} Q(S', a')) \quad (2)$$

という式に従って行動価値関数 $Q(S, a)$ を逐次更新する。ここで α, γ はそれぞれ学習率、割引率と呼ばれるパラメータであり、エージェントは α が大きいほど最新の経験を重んじる傾向を示し、 γ が大きいほど将来の期待を重視する判断基準を習得する。初期の Q 学習においては、価値の記録・エージェントの学習は離散的状態行動空間を表すテーブルを利用して行われたが、深層 Q 学習においてはテーブルの代わりに NeuralNetwork を利用する。 θ を NeuralNetwork を特徴づけるパラメータ、 L_θ を TD 誤差として

$$L_\theta = (R(S, a) + \gamma \max_{a'} Q^\theta(S', a') - Q^\theta(S, a))^2 \quad (3)$$

とくと、逆誤差伝搬法によって訓練された学習済み NeuralNetwork $Q^\theta(S, a)$ は状態変数が連続であっても理想的な行動価値関数と同等の働きをすることができる。本研究では、時系列相関の影響を緩和するため、学習に際して経験した状態・行動・報酬をメモリに蓄積しランダムにサンプリングするミニバッチ学習機構を採用している。

2.1.1 状態行動空間と報酬の設計

状態 (入力): 状態はタスク、サブタスク割当フラグ、サブタスク、メンバー割当フラグの4つの変数をとる。メンバーのリソース値は含まないことに注意が必要である。

タスク: ランダム生成され、チームで処理するように割り当てられたタスク。

サブタスク割当フラグ: どのサブタスクがメンバーに割り当てられているかを確認するフラグ。 $\{0, 1\}$ のバイナリ値により割り当てられていない状態、すでに割り当てられている状態を識別する。

サブタスクのリソース値: サブタスク s_i に該当するリソース値の組 $\{r_1^s, r_2^s, \dots, r_n^s\}$ 。前述のとおり、1つの成分のみ有限の値 $r_p^s \in \{3, 4, \dots, k'\}$ を持ち、その他の成分は0である。

メンバー割当フラグ: 1体のメンバーに対して1つのサブタスクのみ割り当てることが可能なため、2つ以上のサブタスクを割り当てないようにするための変数。本来であれば別の状態でなければならない、割り当て直前の状態が同じになってしまうという問題を回避するため、バイナリ値ではなく $\{0, 1, \dots, l\}$ という値をとる必要がある。

行動 (出力): 入力された状態に対して、どのメンバーに割り振るべきかを行動価値として出力する。最も行動価値が高いメンバーがサブタスクを割り振るべきメンバーとして選出される。

報酬: 本研究では、実験の狙いに応じて以下の二つの報酬を使い分ける。

サブタスク割り当て報酬: タスクに含まれるすべてのサブタスクがメンバーに割り当てられたとき、それ以外の場合を識別する固定値の報酬をリーダーエージェントに与える。固定値の実際の値は実験の章を参照のこと。

時間報酬: リーダーエージェントがすべてのサブタスクを実行可能なリソースを持つメンバーに割り振ることができることを前提として、チームのタスク処理時間に応じて与えられる報酬を時間報酬と呼ぶことにする。これには、 t_{max} を問題設定上ありうる最大のタスク処理時間として $t_{max} - t + 1$ とすればよい。例えば本論文の設定では、一番遅いメンバーのタスク処理時間が最大 12 tick であるので、時間報酬は +1 から +12 までの値をとる。

2.1.2 状態行動空間の探索方策

ϵ -greedy 法を採用し、探索率は $\epsilon = 0.3$ とした。エージェントの学習はタスクひとつに対してエピソード単位で進めるが、サブタスクをひとつずつメンバーに割り当てる刻みをステップと呼び、報酬は各ステップごとに与える。エピソード数は実験に依存して可変とする一方で、1エピソードに含まれるステップ数は 20 で固定した。

2.1.3 NeuralNetwork の Architecture

今回の実験で使用した NeuralNetwork は、各層がそれぞれ 100 個のノードを持つ、4層の全結合 Network (Multi Layer Perceptron) である。活性化関数としては leaky relu 関数、Optimizer は Adam (eps = 0.01) を使用している。

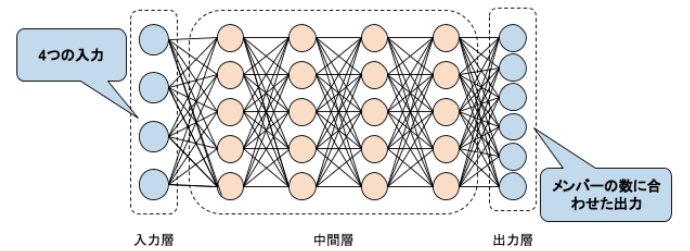


図 1: タスク、サブタスク割当フラグ、サブタスク、メンバー割当フラグの4つの変数を入力とし、出力としてどのメンバーに割り当てべきかを判定する全結合 NeuralNetwork。

2.2 最小化問題としての定式化

問題をソルバーに実行可能な形式に落とし込むため、これまで述べてきた問題設定を以下のような最小化問題として再定義する。

$$\text{Minimize} \quad \sum_i^m \left[\max_{p \in n} \left\{ \text{ceil} \left(\frac{r_p^{s_i}}{r_p^{M_j}} \right) \right\} x(s_i; M_j) \right]$$

$$\begin{aligned}
& \text{Subject to } x(s_i; M_j) \in \{0, 1\} \\
& \sum_j^l x(s_i; M_j) = 1 \\
& \sum_i^m x(s_i; M_j) \leq 1 \\
& \max_{s_i \in T, p \in n} \left\{ \text{ceil} \left(\frac{r_p^{s_i}}{r_p^{M_j}} \right) \right\} \\
& - \sum_i^m \left[\max_{p \in n} \left\{ \text{ceil} \left(\frac{r_p^{s_i}}{r_p^{M_j}} \right) \right\} x(s_i; M_j) \right] \geq 0
\end{aligned} \tag{4}$$

ここで、サブタスク s_i と担当メンバー M_j の対ひとつを表す記号 $(s_i; M_j)$ を導入した。

2.3 計量の設計

この節では、複数の手法の性能を比較・測定するための計量について議論する。計量は考慮する全ての手法に適用可能であり、かつ、できる限り多様な観点から性能を比較できる組であることが望ましい。以上のことを念頭に、我々は以下に示す4つの計量を導入する。

割り当て成功率: 現時点でのチーム構成で割り当て可能なタスクのうち、実際に割り当てに成功したタスクの割合。

計算実行時間: 同じ計算資源上で測定した、割り当てにかかる時間。強化学習を用いた手法に対しては、訓練の時間を含めないことに注意すること。

最適化精度: 割り当てに成功したタスクが持つ処理時間とありうべき最適なタスク処理時間との近さ。タスク処理時間が最短となる組み合わせをあらかじめソルバーによって求め、順位付けを行ったのち

$$O_a = \frac{1}{N} \sum_{ii} f_{\Pi}(\pi_{ii}) \tag{5}$$

$$f_{\Pi}(\pi_{ii}) \equiv \begin{cases} 1 & (\text{if } \pi_{ii} = 1) \\ 1 - \frac{\pi_{ii}}{\Pi} & (\text{otherwise}) \end{cases} \tag{6}$$

という独自指標を導入し、百分率表示を用いて評価する。ここで $N \equiv \#T$ は評価に使われた割当成功済タスクの数であり、 π_{ii} は ii 番目のタスクに対してリーダーエージェントが提示したサブタスクとメンバーの組み合わせの順位をソルバーの結果と比較して決定したもの、 Π は着目する組み合わせ内にありうる総順位数である。

汎化性能: 訓練時に含まれないタスクについて、どれだけの割り当て成功率あるいは最適化精度を達成できるか。

3. 実験

まず始めに、もっとも簡略化された問題設定の下で提案手法が想定通りの挙動を示すこと、すなわち、強化学習の枠組みを使ってリーダーエージェントが実行可能なサブタスクをメンバーに割り振り可能であることが確認された [実験 A]。次に、徐々に問題設定を複雑化してゆき、提案手法 (状態・報酬設計、探索方策) によりリーダーエージェントは得意・不得意のある個性的なメンバーに適切なサブタスクを割り振る方法を学習できることを実演した [実験 B]。提案手法の性能を評価するために割り当て成功率、計算実行時間、最適化精度、汎化性能といった計量を導入し、ランダム法・ソルバーによる解法との比較を行った。

3.1 実験 A

入力として4つの変数をとる状態空間を採用することによって、本当にサブタスクをメンバーに割り振れるのか確認することを狙いとして、チームメンバーに割り振ることができるサブタスクの組を有する1つのタスクを割り振る実験を行った。割り振れているかどうかだけを確認するために、各メンバーの所持しているリソースの成分は1つのみ1とし、量的得意不得意の概念を排除している。サブタスクとメンバーのリソース値は以下の通りである。

$$\begin{aligned}
s_1 &= \{3, 0, 0, 0, 0, 0\}, s_2 = \{0, 4, 0, 0, 0, 0\}, s_3 = \{0, 0, 5, 0, 0, 0\} \\
s_4 &= \{0, 0, 0, 6, 0, 0\}, s_5 = \{0, 0, 0, 0, 7, 0\}, s_6 = \{0, 0, 0, 0, 0, 8\} \\
M_1 &= \{1, 0, 0, 0, 0, 0\}, M_2 = \{0, 1, 0, 0, 0, 0\}, M_3 = \{0, 0, 1, 0, 0, 0\} \\
M_4 &= \{0, 0, 0, 1, 0, 0\}, M_5 = \{0, 0, 0, 0, 1, 0\}, M_6 = \{0, 0, 0, 0, 0, 1\}
\end{aligned}$$

また報酬はサブタスク割り当て報酬のみ +1, -1 とし、時間報酬は与えなかった。

結果として、図2に示すように損失がほぼ単調に減少し、図3のようにエージェントの報酬も最適な値に収束した。なお、-4は最短の試行ステップ -1, -1, -1, -1, -1, +1 に対応する、最適な割り振りを行った場合の割り振り報酬であり、-20は1エピソード内に含まれる20ステップすべてで割り振りに失敗した際の報酬である。

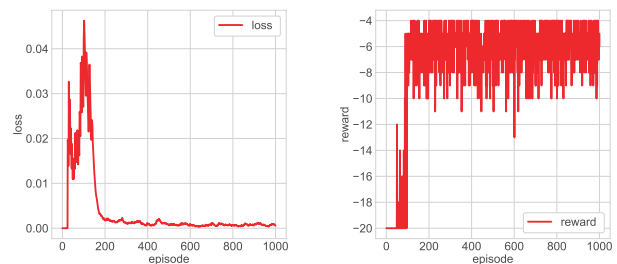


図2: 損失関数の振る舞い。図3: エージェントが得た報酬。定して学習が進んでいる様子が損失の低下に伴い最適な場合の報酬を得ていることがわかる。

3.2 実験 B

メンバーが複数のリソースを持ちたくさんの種類のタスクを処理しなければならないような状況下において、リーダーエージェントが汎用的な方策を学習することができるのかどうかを確認することを目標に、100000エピソードかけて、ランダム生成した100タスクを繰り返し与えて学習を行った。

メンバーについては以下のリソース値を持つ。

$$\begin{aligned}
M_1 &= \{0, 2, 0, 4, 0, 1\}, M_2 = \{2, 2, 5, 6, 5, 5\}, M_3 = \{0, 6, 0, 0, 4, 6\} \\
M_4 &= \{3, 2, 0, 0, 0, 0\}, M_5 = \{4, 2, 6, 4, 5, 1\}, M_6 = \{0, 4, 3, 0, 0, 6\}
\end{aligned}$$

報酬としては、タスクを割り振れたときにはサブタスク割り当て報酬は0で時間報酬を与え、割り振れなかったときはサブタスク割り当て報酬のみ -0.5 を与えた。

図4に割り当て成功率、そして図5に最適化精度の実験結果を示す。図より、学習に使用したタスクに対する割り当て成功率 (再現率: Reproducibility) はほぼ100%である一方で、始め

て遭遇したタスクに対する割り当て成功率(汎用性:Versatility)は75%程度にとどまったこと、始めて遭遇したタスクのうち割り当て可能なものに対して最適化精度 80%程度を達成できることなどが読み取れる。また、1000000 エピソードの中で、ランダム生成した 1000 タスクを学習する仕様に変更したところ、割り当て成功率の汎化性能は約 94 %へと向上し、最適化精度の汎化性能は約 85 %へと上昇する改善が見られた。

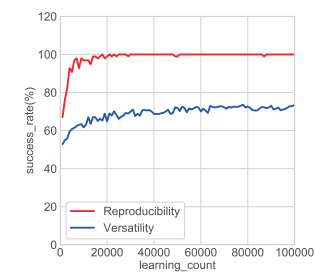


図 4: 割り当て成功率の再現率 図 5: 最適化精度 (再現率:赤) と割り当て成功率の汎用性 および (汎用性:青). 学習時に (青). 学習に使うタスクの数が 遭遇したことがないタスクに対多くないため、汎化性能はそれ しても 80%程度の最適化精度が 出ている。

3.3 他の解法との比較

100000 エピソードを用い 100 種類のタスクを学習済みのエージェントによる解法、ソルバーによる解法、ランダム法の 3 つの手法を対象に、割り当て成功率、最適化精度、計算実行時間という 3 つの観点から手法の評価を試みた。それぞれ 1000 タスクを実行させた結果として表 1 が得られた。この結果から、深層強化学習を用いた手法は精度と実行時間の兼ね合いが必要とする状況下においてよい性能を発揮できるということがわかる。

表 1: 手法比較

実験結果	割当成功率	最適化精度	計算実行時間
ランダム	16 %	39 %	0.091 [s]
エージェント	74 %	79 %	8.118 [s]
ソルバー	100 %	100 %	17.44 [s]

4. まとめ

本研究では、強化学習の枠組みを使ってタスク・リソースマッチング問題に取り組み、ランダム法・ソルバーによる解法といった対象と比較して多角的に評価を行った。その結果により、提案手法には組み合わせ精度はそこそこでよい代わりに限られた時間の中で次々と処理を迫られるような状況に強いという利点がある、と結論付けた。次の研究課題として、複数チーム間の相互作用を適切に決定して階層的なシステムを構築する、静的なシステムを動的なものに拡張する、状態行動空間の探索方法や NeuralNetwork の Architecture を洗練するなどの課題がある。

参考文献

[1] B. M. Dunin-Keplicz and R. Verbrugge, “ Teamwork in Multi-Agent Systems: A Formal Approach,” 1st ed. Wiley Publishing, 2010.

[2] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, “ Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications,” IEEE Communications Surveys Tutorials, vol. 17, no. 4, pp. 2347–2376, Fourthquarter 2015.

[3] N. Amelina, A. Fradkov, Y. Jiang and D. J. Vergados, “ Approximate Consensus in Stochastic Networks With Application to Load Balancing,” in IEEE Transactions on Information Theory, vol. 61, no. 4, pp. 1739-1752, April 2015.

[4] Banerjee S., Hecker J.P., “ A Multi-agent System Approach to Load-Balancing and Resource Allocation for Distributed Computing,” In: Bourguine P., Collet P., Parrend P. (eds) First Complex Systems Digital Campus World E-Conference 2015. Springer Proceedings in Complexity. Springer, Cham

[5] A. Shukla, S. Kumar and H. Singh, “ LOAD BALANCING APPROACHES FOR WEB SERVERS: A SURVEY OF RECENT TRENDS,” IJE TRANSACTIONS B: Applications Vol. 31, No. 2 (February 2018) 263-269

[6] O. Sheholy and S. Kraus, “ Methods for task allocation via agent coalition formation,” Journal of Artificial Intelligence, vol. 101, pp. 165–200, 1998.

[7] W. Saad, Z. Han, M. Debbah, A. Hjorungnes, and T. Basar, “ Coalitional game theory for communication networks,” IEEE Signal Processing Magazine, vol. 26, no. 5, pp. 77–97, September 2009.

[8] M. Hayano, N. Iijima and T. Sugawara, “ Asynchronous Agent Teams for Collaborative Tasks Based on Bottom-Up Alliance Formation and Adaptive Behavioral Strategies,” 2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech), Orlando, FL, 2017, pp. 589-596.

集団適応を導くギャップ補填に基づく「思いやり」

Omoiyari as Filling Gaps Making Collective Adaptation

前川佳幹^{1*} 上野史¹ 北島瑛貴¹ 高玉圭樹¹
Yoshimiki Maekawa¹ Fumito Uwano¹ Eiki Kitajima¹ Keiki Takadama¹

¹ 電気通信大学

¹ The University of Electro-Communications

Abstract: To make each person adapt human society and build a consensus, we focus on Japanese Omoiyari as filling psychological and numerical gaps among people. Concretely, we employed the cross-cultural game "Barnga" whose the four players have to build a consensus and determine a winner, and proposed gesture marks in the Barnga. The gesture marks are the panels which can express happy, angry, sad, and surprise. Each player can make other players recognize the gaps among players, and behave with the Omoiyari. To validate the effect of the gesture marks, we experiment on Barnga with the gesture marks and original Barnga. By analyzing the subject experiments on the Barnga, we derive that: 1) the players can recognize the gaps by the gesture marks; 2) the gesture marks contribute to make the players adapt the community and build a consensus.

1 はじめに

集団社会において、まとまりを持ち、意思決定を下していくことは極めて重要である。集団がまとまりを持つためには、その集団に属している個人どうしでの密なコミュニケーションが必要になる。しかし、インターネット上でのコミュニケーションでは考えを正しく共有することが困難であるため、自分の考えが相手に正しく共有されずに発生する「誤解」やそれに伴う「炎上」が頻繁に発生する。このような問題の解決のために、人間と人工物（エージェント）との相互関係（Human-Agent Interaction: HAI）[1]に焦点を当てた研究がなされている[2][3]。

牛田ら[4]は集団に属している個人の個性に着目した。集団の状態はこの個性の組み合わせによって四種類に分類でき、設計したエージェントを集団内に介入させることで集団としてまとまりをもつ「集団適応状態」に遷移できると述べている。しかし、牛田らは「個人の状態」にまで言及していない。集団に属している個人が納得しているのか否か区別がつかず、必ずしも集団がまとまっていると判断できない。そこで私たちは、個人が納得して意思決定を下し、集団適応に遷移するには、「思いやり」などの人間の定性的な尺度が必要だと考える。本研究においては、「思いやり」を主観的なギャップを補填する行動として定義した後、被験者

実験の分析により、本研究で定義する「思いやり」が集団適応を導く上で有効であることを示す。

2 ギャップ補填による思いやり

本研究では、「思いやり」をギャップを補填する行動として表現する。「思いやり」とは単純な共感や同情とは異なる。内田ら[6]は「思いやり」には、「同情」「直感的な理解」「向社会的な行動」の三種類の側面があると仮定し、「思いやり」の尺度を提案、検証している。ギャップを補填するためには、相手に同情することでギャップに気づき、相手が何を求めているのかを直感的に理解する必要がある。最終的に行動として現れた際には向社会的な行動として認知される。以上の点から、ギャップ補填を「思いやり」として仮定し、この効果を検証する。

本稿で考慮するギャップは大きく二種類の分類がなされている。一つ目は、数値的ギャップが発生しているかどうか、二つ目は心理的ギャップが発生しているかどうかである。ここで扱う「数値的」なギャップとは、例えば身長や体重、収入など、数字で表現可能なものを指す。「心理的」なギャップとは通常目には見えないものを指し、例えば、他者との意識の齟齬や誤解などがあげられる。以上の分類の組み合わせより、本稿では合計4種類のギャップを考慮し、これを補填する行動を「思いやり」になりうる行動として表現する。

*連絡先：電気通信大学 情報理工学部 総合情報学科
〒182-8585 東京都調布市調布ヶ丘1-5-1
E-mail: maekawa.yoshimiki.1119@cas.lab.uec.ac.jp

3 異文化体験ゲームバルンガ

先述した「思いやり」がもたらす集団適応への効果を検証するために、教育学で発展した異文化体験ゲームバルンガ [5] を採用する。バルンガは、トランプを用いたシミュレーションゲームである。ゲームのルールを文化と見立て、その違いに気づいたときの驚きをカルチャーショックとして体験し、その解決方法の模索も同時に体験することを目的としている。ゲーム中は言語コミュニケーションが禁止されており、プレイヤーはジェスチャーなどを駆使して自分の意思を他のプレイヤーに共有しなくてはならない。Barnga では、各スートごとに A から 7 までの合計 28 枚のカードを使用する。

まず、プレイヤーに 1 人ずつゲームのルールを教え、その後プレイヤー 4 人で Barnga をプレイする。1 人ずつ手札からカードを場に出し、全員出し終わったのち、一番強いカードを出したプレイヤーを勝者として指さす。勝者を指さしたプレイヤー全員の指差しが一致した際にゲームの勝者が決定し、次のゲームは勝者からカードを手札から場に出していく。この手順を規定の回数こなすことで Barnga は終了する。ゲームの最中に手札が無くなった場合は、場に出したカードを集め、再度プレイヤーに分配する。ここで着目したいのは、各プレイヤーに事前に教えていたルールが、カードの強さの決め方について若干異なることである。教えられていたルールは各プレイヤーごとに異なるので、ゲームの開始時、プレイヤーは違和感を覚える。バルンガにおいて、この違和感がカルチャーショックであり、これを感じることを異文化体験としている。

4 ジェスチャー札

本研究では、バルンガ中で唯一許されている「ジェスチャー」をより定量的に評価し、その効果を検証するため、新たに喜び、怒り、哀しみ、驚きの四種類の感情を表現する「ジェスチャー札」を導入した。図 1 は導入したジェスチャー札の外観であり、左から喜び、怒り、哀しみ、驚きを表している。ジェスチャー札を Barnga 中に使用することで、自らの状態を間接的に他者に共有でき、その結果適応的な状態を導く効果が期待できる。

5 バルンガにおける場の状態の定義

本研究では、バルンガにおける場の状態を四種類定義する。これらは、主張率と譲歩率によって決定される。以下では、本研究で用いる主張率 E_{ind} と譲歩率 O_{ind} について記述する。



図 1: ジェスチャー札

主張率 E_{ind}

プレイヤーがどれだけ自分の選択に自信を持って他のプレイヤーにみけて主張しているのか、その度合いを指す。意見の衝突の際、再度同じ選択をした回数を E_{self} 、違うプレイヤーを選択した回数を E_{other} としたとき、主張率 E_{ind} は以下の式より導出できる。

$$E_{ind} = \frac{E_{self} - E_{other}}{\max(E_{self}, E_{other})} \quad (1)$$

譲歩率 O_{ind}

1 ラウンドにおいて、どの程度自分のルール以外で最終的な勝者を選択したのか、その割合を指す。自分のルールに基づかない勝者を最終的に選択した回数を F_{other} 、現在のゲーム数を N_{game} としたとき、譲歩率 O_{ind} は以下の式より導出できる。

$$O_{ind} = \frac{F_{other}}{N_{game}} \quad (2)$$

なお、以上の指標と各値はラウンドごとの性質を分析可能なものとするため、ラウンドが終了するたびに初期化するものとする。

以上の指標をプレイヤー毎に計算し、縦軸を主張率 E_{ind} 、横軸を譲歩率 O_{ind} としたグラフにプロットする。ここで θ_O を定義する。 θ_O は各ラウンドにおけるプレイヤーの最終的な譲歩率 O_{ind} の平均値である。プロットの位置によって、プレイヤーの行動特性は以下のように定義できる。 $0 < E_{ind}$ かつ $O_{ind} < \theta_O$ にプロットされるプレイヤーは自分のルールを支持し続け、その選択に自信を持っている「主張タイプ」。 $0 < E_{ind}$ かつ $\theta_O < O_{ind}$ にプロットされるプレイヤーは自分以外のルールを支持し、その選択に自信を持っている「支持タイプ」、 $E_{ind} < 0$ にプロットされるプレイヤーは自分の選択に自信がない「控え目」タイプである。各プレイヤーのプロットの位置関係によって場の状態を四種類定義する。図 2 は本研究で定義する四種類の場の状態のプロットを表しており、左上が一人のプレイヤーの主張が他のプレイヤーを支配している「支配状

態」，右上がいずれのプレイヤーも自分の選択に自信がない「混乱状態」，左下がそれぞれ異なるルールを持っている複数のプレイヤーが主張し合っている「固執状態」，右下が同一のルールを持つプレイヤーを他のプレイヤーが支持している「適応状態」である。

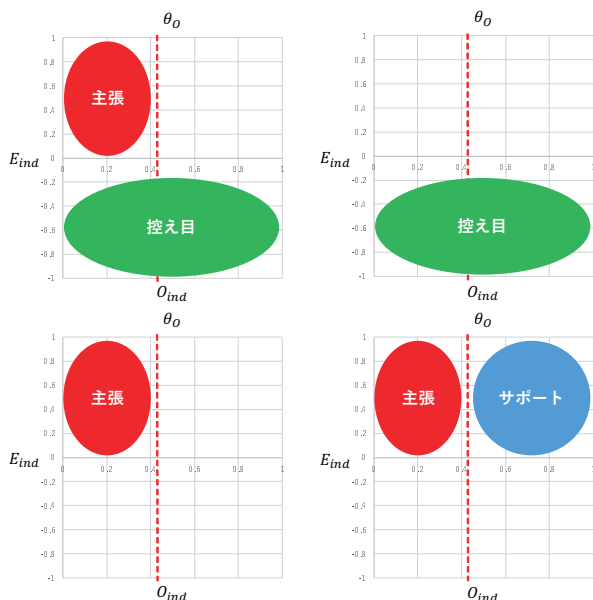


図 2: 場の四状態

6 実験

6.1 内容

本研究では，オリジナルのバルンガと先述したジェスチャー札を導入したバルンガをそれぞれ被験者実験に用い，結果の差異を分析することでジェスチャー札の効果を示す．実験における被験者は4人とし，サングラスとマスクを着用の上，他のプレイヤーから見て表情が分からないようにした．実施ゲームは28ゲームである．また，内田ら [6] が提案した思いやりの尺度を含めたアンケートを実施することにより，思いやりのあるプレイヤーの札の使い方などを明らかにする．アンケートには22項目あり，それぞれに対して思いやりに対する因子負荷量が設定されている．本研究において，プレイヤーの思いやり度合い S を，質問項目 i の因子負荷量を a_i ，5段階のSD法で撮ったアンケート結果を w_i としたとき以下のように定義する．

$$S = \frac{\sum_{i=1}^{22} a_i w_i}{\max S} \quad (3)$$

思いやり度合いのほか，ギャップが発生した際にどのような行動をとったかについてもアンケートを実施した．アンケート結果より，実際にとられた行動と場の状態の遷移の対応を明らかにする．

6.2 結果

6.2.1 場の状態の遷移

ジェスチャー札なしのバルンガでは，すべてのラウンドにおいて固執状態であり，複数のプレイヤーが自らのルールを主張し続ける状態に収束した．対照的に，ジェスチャー札ありのバルンガでは，固執，支配，混乱，適応の順に場の状態が遷移した．図3は各被験者実験の状態遷移を縦軸で表現した線グラフである．1が支配状態，2が混乱状態，3が固執状態，4が適応状態を表している．

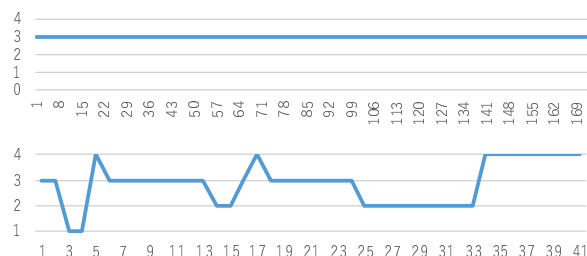


図 3: 上: ジェスチャー札なし，下: ジェスチャー札あり

6.3 思いやり度合い S

アンケートの結果得られたデータより思いやり度合い S を各プレイヤーごとに算出した．以下の表1と表2は核実験におけるプレイヤーの思いやり度合い S についてまとめたものである．

表 1: ジェスチャー札なし 表 2: ジェスチャー札あり

	S		S
プレイヤー 1	0.76	プレイヤー 1	0.53
プレイヤー 2	0.67	プレイヤー 2	0.63
プレイヤー 3	0.53	プレイヤー 3	0.62
プレイヤー 4	0.46	プレイヤー 4	0.60

6.4 適応状態時のジェスチャー札

適応状態時，心理的ギャップと数値的ギャップの両方が発生していたときには「驚」が使用されており，数

値的ギャップのみが発生していたときには「喜」が使用された。数値的ギャップが発生していない場面ではジェスチャー札が使用されなかった。

7 分析

図3より、ジェスチャー札がある場合とない場合との間で、場の状態の遷移に違いが出たことが分かった。このことから、ジェスチャー札を用い、自分の感情を他のプレイヤーに共有することで、多様な状態に遷移可能であり、結果として適応状態に遷移したことが明らかになった。最終的な適応状態では、ギャップが発生しているにもかかわらず「喜」が使用されていることから、各プレイヤーが結果に納得して勝者を決定していると考えられる。

また、各状態に置いて適応状態に遷移する際に使用された札の分布を心理的・数値的ギャップの観点から分析することで以下の知見を得ることができた。1) 喜び札は、意見の衝突の際に自分が特定のプレイヤーの選択を支持していることを表現でき、勝者の選択をまとめる効果を持つ。2) 怒り札は、心理的ギャップと数値的ギャップが存在する際に、プレイヤー間の優劣を明確にし、他のプレイヤーに劣っているプレイヤーがいることの認知を促し、適応状態を導く一つの要因となった。3) 哀しみ札は、各プレイヤーの適応を促し、ルールを一つにまとめる効果がある。4) 驚き札は、意見の衝突が発生した際に使用されているため、自分が感じているギャップを他のプレイヤーに共有する際に使用されることが分かる。さらに、適応状態を導いたと考えられるジェスチャー札は思いやり度合いの高いプレイヤーによって多く使用されたものであり、思いやり度合いが高いプレイヤーがギャップの認知を促し、最終的に適応状態に遷移したと考えられる。

そのほか、実施したアンケート結果より、思いやり度合いの強いプレイヤーは心理的・数値的ギャップが存在する際、そのギャップを埋めようとするのが明らかになった。しかし、ジェスチャー札の有無によって状態遷移に差が生まれたことから、ギャップを埋めようとする行動はジェスチャー札がある場合に効果的に作用すると考えられる。ジェスチャー札がない状況では、プレイヤーの主張は勝者選択時における指差ししかなく、その外観はいずれのプレイヤーも同じである。そのため、ギャップを埋めようとする行動であったとしてもその意図が他のプレイヤーに共有されにくく、場の状態を遷移するまでに至らなかったと考えられる。対して、ジェスチャー札がある状況においては、その意図が他のプレイヤーに共有され、場の状態遷移に寄与したと考えられる。

8 まとめ

本研究では、集団を適応的な状態へと遷移させるためには「思いやり」のような人間の定性的な尺度が必要であると考え、これを心理的・数値的ギャップを補填する行動として定義した。思いやり行動を促進するツールとして、自分の感情を「喜」「怒」「哀」「驚」の4種類で表現できるジェスチャーマークを提案し、その有効性を異文化体験ゲームバルンガを用いて検証した。その結果、思いやりを持つプレイヤーがギャップの発生の際に使用したジェスチャー札は集団を多様な状態に遷移させることが明らかとなり、その中でも特定の使用方法によって適応状態に導けることが明らかになった。さらに、思いやり行動の効果の発揮に寄与したと考えられた。

謝辞

本稿を執筆するにあたって、学術新領域である「共創言語進化」に感謝の意を表したい。

参考文献

- [1] 山田誠二.: 人とロボットとの < 間 > をデザインする, 東京電機大学出版局, 2007.
- [2] Ohmura, H., Katagami, D., Mitta, K.: Design of Social Adaptive Agents in Simulation Game of cross-cultural experience, Human-Agent Interaction Symposium, 2009.
- [3] Ohmura, H., Katagami, D., Mitta, K.: Investigation Changes of Group Norm for Agents Adapting Group, *The 23rd Annual Conference of the Japanese Society for Artificial Intelligence*, 1-3, 2009.
- [4] Ushida, Y., Takadama, K.: Validation of Agent Model in Highly-dynamic Environment via Barnga Game, *The Multi-Agent-Based Simulation workshop (MABS2012)*, 2012.
- [5] Thiagarajan, S., Steinwachs, B.: Barnga: A simulation game on cultural clashes, *Intercultural Press*, 1990.
- [6] Uchida, Y., Kitayama, S.: Development and validation of a sympathy scale, *The Japanese Journal of Psychology* **72**(4), 275-282, 2001.

オンライン上における社会合意形成のシミュレーション手法の検討

Investigation of online simulation method of social consensus formation

川畑 泰子*1
Yasuko Kawahata

石井 晃*2
Akira Ishii

上岡 拓矢*2
Takuya Ueoka

*1 群馬大学
Gunma University

*2 鳥取大学
Tottori university

Discrepancies and agreement formation have been studied for a long time. However, the laws and orders that are particularly reflected in society are being updated by information network changes. In particular, since 1995, as public network devices have spread all over the world due to the appearance of Microsoft Windows 95, the opportunities for making decisions and consensus formation beyond the spatiotemporal constraints have expanded enormously. From now on, quantitative research is needed on the tendency to aggregate knowledge on quantitative knowledge and trends concerning consensus formation and tendency to bring about social and economic risks. Therefore, in this research, we first examined the possibility of comparing the research on consensus building using measured data and the approach using mathematical model.

1. 序

意見の相違と合意形成は古くから研究されてきた[1-6]。しかし、社会でそれらが特に反映される法・秩序は、情報網の変化によってアップデートされている[7]。特に、ライフログのデータ取得が可能な基盤を保持する情報通信技術を持つネットワーク基盤の到来は、我々の日々のデータは無意識であれ意識したものであれ、ある種の合意形成として取り扱われている場合も発生している[8]。上記のような社会環境システムは、自然社会における森羅万象と同様に統計学的・機械学習など様々な手法で理解する研究が進められるようになった。また、社会を支える基盤を説明・自動化する学習モデルとして生成・著作されるケースも多く生まれている。自然と同様に、私たちの日々形成する様々なライフログが技術を経て人々の創造活動の基盤となる機会も増えていけると言える[9]。近年の事例である、YouTube 上で日本における国防に関する動画のアップロードによる検証動画の配信が閲覧可能となっている[10]。何らかのメディアの編集作業がされ、国境を跨いだ世論形成に影響を及ぼす議論もオンライン上で第三者の国々から閲覧される状況にもなった。今後、合意形成に関して定量的な知見や傾向に関する知見を集約し、ソーシャルリスクをもたらすであろう傾向に関して定量的な研究は必要と考えた。そこで、本研究では実測のデータを用いた合意形成に関する研究と数理モデルを用いたアプローチの比較の可能性に関してまず検討を行った。

2. 先行研究

集団におけるある意思決定問題を解決するための AHP(Analytic Hierarchy Process)も 1980 年代以降、提案されてきた。このモデルにおける特徴は実際の問題に対してアプローチができ[11]、集団の合意形成過程の随所に決定に至るまでの区間表現を用いる点にあり、集団の合意を定量的に形成する点を持つ。上記のような合意形成に関する研究は理論ベースから現実の問題に対してアプローチする手法まで多様に行われている[12]。また、合意せず意見合意に至る際の集団における不満が発生した場合のケースに関しても、集団の一対比較値を算出し、定量的な傾向を示唆が可能となっている。ここで構成する

モデルでは、提示された区間に対し不満関数を定義し、整合性と集団全体の不満をともに最小化する。近年は正負の決定問題におけるモデルに対して、欠損値や大規模なケースにも対応した研究成果も考案されている[13]。上記のような、ある合意形成に関わる問題解決手法も考案されてきた中、インターネット上で様々な意見を具体的なテキストとして収集できるようになった。社会のフレームワークを構築する上で政治に関する大衆の意見や意志・バイアスは常にメディアの発展と活用依存すると考えられてきた[14]。メディアは社会制度に対応すべき独自の論理かつ独立した機関として出現した高度な近代化の象徴でもある。その一方で、メディアは政治、仕事、家族、宗教などの他の機関の統合された部分であり、これらの機関の活動の多くは、双方向・即時性をもつソーシャル・メディアとマスメディアの両方を通じて行われつつある。即ち、社会自身がメディア化し、私たち集団における意見の形成が常になされる時代となった[14]。特に、アメリカ合衆国・欧米諸国では 2000 年代より公衆ネットワークにおける Web を介した選挙活動は、Web Technology の発展と同期して E メール・メールマガジンなどダイレクトな発信から HP など全世界向けの発信など盛んに活用されてきた。2012 年には ARMA モデルを採用し、諸条件やボラリティを考慮したモデルによって傾向を予測する手法で 2008 年の米大統領選挙の Twitter のデータを用いて意見の時系列分布における実証分析が行われた[15]。意見の合意形成に関して選挙結果を集約し、予測問題に落とし込むことによって結果がもたらした説明変数における係数から考察を行う手法など多く登場した[16-19]。

3. データ

3-1. 選定したデータ

社会におけるメディアの発信する情報に関して人々の反応をどのデータを取得して行ったか論ずる。現在、機械学習における分類手法の拡張、辞書データや学習データの元となるデータの充実により自然言語処理に関しても多様な研究がなされている[16-19]。つまり、自然言語処理で集めたテキストから意見の強弱を判定できるようになったことで、社会における意見の分布を Binary opinions ではなく、ポジティブな言及からネガティブな言及、ニュートラルな言及までの連続的な分布として測定できるようになったといえよう。本研究では、世界上における様々なメディアのデータ取得の制約がある中で YouTube を選択した。

様々な OSN がリアルタイムな動画配信など可能となっているが、YouTube の場合は世界有数の TV・新聞媒体をもつニュースメディアがニュース報道の断片を YouTube でも公開している。また、リアルタイムでの配信が行われているのも特徴である。また、各国におけるニュースメディアにおける配信内容・同じニュースの内容の報道の違い・コメントをするユーザー層の相違における意見の分布における考察なども可能と仮説を立てた。また、今回は同じ条件下での意見の強弱を把握するため、英語圏辞書データをベースとした ch におけるニュース動画に限定した。

YouTube は言語性が関わるケースも存在するが、動画の内容は世界発信であり、メディア媒体別の違いに関する比較が期待されること、同じ報道内容でも報道手法によって受け取りかたが異なる部分など TV における発信と近いものがあると仮説をおいた上で解析を行った。

3-2. データの取得

取得対象の ch は CNN を今回はメインとして取得した。コメントの取得期間は、CaseA においては 2016/10/10~2018/11/29、CaseB においては、2013/5/24~2018/11/29 までの内容とした。本研究では、トライアルとして CNN における ch 立ち上げ以降最も再生回数が多かった動画におけるコメントを収集し、各コメントごとにおける negative, positive, neutral のスコアを合算して 1 となるように処理を行った。各 3 つの軸におけるスコアは $[-1.0, -0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75, 1.0]$ の range で定義した。今回は、分布を考察する際に偏ったスコアである $-1, 0, 1$ に関しては省く処理を行っている。

入力されたデータに対し、NLTK (Natural Language Toolkit) [36] を使用している。各動画におけるコメント 1 つ 1 つにおける特徴量抽出を行い、最終的に各動画のコメント全体における negative, positive, neutral のスコアの分布を 3 つの軸において出力した。また、negative, positive, neutral の各軸のスコアを一元的な分布で捉える際は、positive な意見のスコアに 2 を加算、negative な意見はスコアを負の数に変換し、neutral な意見には加算せず $-1 \sim 2$ のスケールに換算する手法をとった。

4. 合意形成のモデル

大規模なデータベースを用いた解析が重要視されて合意形成をもたらしたメディア・外力、世論、宗教、地域などの役割を明確にする手法が可能となった。本研究では、こうした大規模なデータベースを用いた定量的な解析を、オピニオンダイナミクスに応用する。この研究においては、定量的な解析や予測のための理論を実測と理論計算の共同作業の可能性を模索する。社会の意見分布の定量的な分析は 1890 年以降、以降アメリカ合衆国における新聞における掲載記事の分類を定量的にまとめ集計した研究に始まり、1930 年以降、ハロルド・ラスウェルが内容分析として、定量的な解析の重要性を唱えてきた[20-21]。しかしながら、内容分析と意見に関する研究に関しては取得可能なメディアは雑誌、新聞などマスメディアのものに限られていた。そのため、理論研究が先行する形で行われていた。昔から数値化しやすい意見対立の実例は選挙である。特に米大統領選や仏大統領選は A 氏か B 氏かどちらかに意見が集約されていく。そのため、意見を A と B の 2 つと仮定した理論が応用しやすい。そこで、物質内で原子の磁気の向きが 2 つしか取り得ない磁性物理学理論を応用する試みが古くからなされた。つまり、この種の数理モデルとしては、意見の相違を単純に 0 と 1、あるいは 1 と -1 とした二値モデル (Binary opinions) のモデルが磁性物理学理論の Ising 模型の応用として提出されている[22-24]。

しかし、この Ising 模型類似の研究手法では意見の分布の時間的な発展が入れにくい。また、一方で繰り込み群の理論を応用したモデルも Galam によって考案されている。社会を少人数のグループに分けて、そこでの多数決で A か B かの二値の意見のどちらかに意見が揃うと仮定したものである。それらをグループ分けを何回か繰り返して行うことで A と B のどちらの意見が社会で多数派を占めるかを計算する理論である [25,26]。この理論では少人数グループで A と B が同数だった場合、社会の世論として優位な意見でこの少人数グループの意見が揃うとする。最初は少数だったグループが簡単に逆転して多数派となりうることを計算から示している。その考えで Brexit[27]や米大統領選のトランプ当選[28]の検証も行われている。この繰り込み群の理論では世論の動向の時間的な発展はある程度追えるが、社会の世論でどちらの意見が優位かは理論の中にはない。外部的な決定として境界条件として入れる必要がある。また、二値に限る手法は米大統領選や仏大統領選の予測解析には応用できるが、一般の社会の意見の分布は賛成の意見も反対の意見も強弱がある。そのため、賛成と反対の二値に限定してしまうのは単純すぎると言える。一方で連続的な意見分布を扱う数理モデルとして Bounded Confidence Model がある[29-31]。しかし、この Bounded Confidence model は実際に大規模なデータベースから社会の意見を解析する上で以下の欠点がある。合意形成の理論なので 0 から 1 までの分布しか扱わないマスメディア等の社会における環境における外力の影響が入っていない。この欠点を改良する形で Bounded Confidence Model を大きく発展させた理論が最近、石井・川畑によって提出された[32]。この理論では意見はポジからネガまで連続的とし、個々の人同士の相互作用は賛同・同調から反発まで様々な相互作用があるとした。また、外力の影響を社会現象の数理モデル[33,34]に倣う形で採り入れた。この理論では Bounded Confidence Model の欠点が改良されている。また、この理論は時間発展の微分方程式の形で提出されていて、Hegselmann-Krause の理論と異なっており意見の分布の時間発展が記述しやすい。ただ、この理論は相手の意見に対する同調も反発も無限に続くとは仮定している。意見交換の初期段階は正確に記述できるが、最終段階の予測は非現実的な無限大の正と負の意見になってしまう。この欠点を補う理論の改良が Ishii によってなされており[35]、本研究ではこの理論を、以下に述べる社会の意見の実測を解析する理論として用いる。

本理論の要素としては、①遠く隔たった意見には影響されない②自分とほぼ同じ意見には影響されないという 2 点の修正を加えた仮説によるモデルである。

$$\frac{dI_i(t)}{dt} = C_i A(t) + \sum_{j=1}^N D_{ij} \Phi(I_j(t), I_i(t)) (I_j(t) - I_i(t))$$

$$\Phi(I_i, I_j) = \frac{1}{1 + \exp(\beta |I_i - I_j| - b)} \cdot \dots \cdot (1)$$

5. シミュレーションと実測の検討

5-1. シミュレーションの場合

(1) 式において、 D_{ij} は i 氏から j 氏がどれくらい信用されているかをシミュレーションしている。この D_{ij} に対応する係数は Hegselmann-Krause (2002) の Bounded Confidence Model でも定義されているが、我々の新しいオピニオンダイナミクス理論ではこの係数 D_{ij} が正の値なら信用していて、負の値だと不信を抱いていると定義した。この結果、お互いに不信を抱いていると、意見交換によってお互いの意見は離れていくという計算結

果が得られる. ある程度以上意見が離れると, 互いに無視するので平行線となる. これに第三者が加わるとして, この第三者(C氏)がA氏とB氏から強い信頼を得ている仮定して計算すると, 対

立していたA氏とB氏がC氏の意見にまとまるという計算結果になる[35].

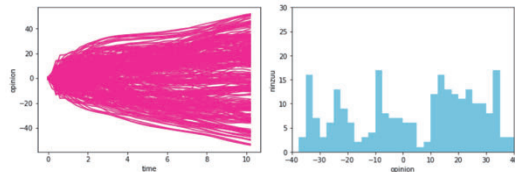


図 1: 右図: 300 人のお互いの係数 D_{ij} を 1 から -1 までの乱数で決めると設定して計算した例. 左図: 縦軸が意見の差で真ん中が中立, 横軸が時間の経過である. また, 図の右側は最終的な時間で計算で示された意見分布を示す

例として, 300 人のお互いの係数 D_{ij} を 1 から -1 までの乱数で決めると設定して計算した例を右図に示す. 図の左側で縦軸が意見の差で真ん中が中立, 横軸が時間の経過である. また, 図の右側は最終的な時間で計算で示された意見分布を示す.

5-1. 実測における考察

CNN の動画の再生回数上位は, 国政関連の動画や面白さに関してフォーカスをした動画にやや傾斜している. 本論では, ケーススタディとして最も再生回数の多く, ニュース報道ではない内容に絞った. そこで再生回数が上位の国政に関するトピックの動画におけるコメントの意見の分布に着眼をした.

Case A “Trump supporter leaves CNN anchor speechless”

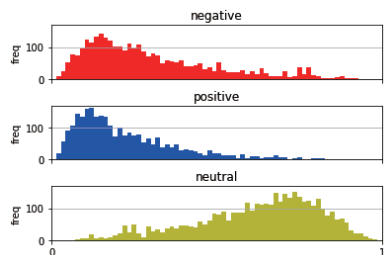


図 2: Case A におけるコメント(件)の negative, positive, neutral の分布(取得期間: 2016/10/10~2018/11/29, 各スコア range: 0~1)

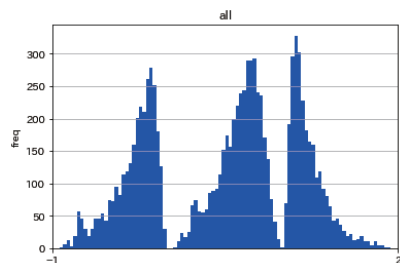


図 3: Case A におけるコメントの negative, positive, neutral を一元化した場合の分布 (スコア range: -1~2) (取得期間: 2013/5/24~2018/11/29)

Case B “Obama forgets to salute”

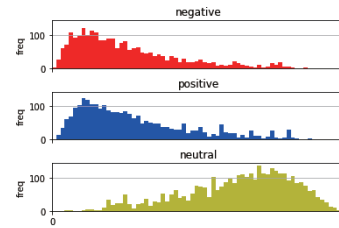


図 4: Case B におけるコメントの negative, positive, neutral の分布 (取得期間: 2013/5/24~2018/11/29)

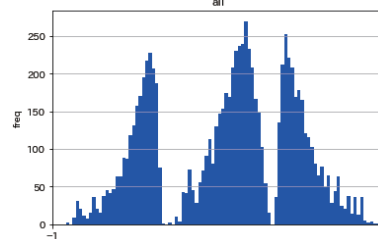


図 5: Case B におけるコメントの negative, positive, neutral を一元化した場合の分布 (スコア range: -1~2) (取得期間: 2013/5/24~2018/11/29)

Case A と Case B は両者とも国政関連の動画であるが, Case A は大統領選挙に関わる内容であり, ややネガティブな分布への偏りが観測される一方, Case B に関しては面白さにフォーカスした動画であり, ネガティブな傾向や極端に偏った分布は見受けられず比較的ポジティブな見方をされがちな傾向にあると推察された. 本研究では, コメントにおけるの自然言語処理を経たスコアにおける分布から社会の合意形成の成り立ちとシミュレーションの結果の違いとシミュレーションに関する試みを行っている. 今後この方向で研究を進めるためには, ポジ, 中立, ネガのつなぎ目の処理が課題の一つであるとも実測の結果とシミュレーションの結果から検討される.

6. 考察

本研究の大きなテーマであるマス・パーソナルにおける OSN における社会的影響についての議論は, 権力に対する抵抗勢力・ポピュリズム・危険性におけるトピックとして課題にされがちである. そのため, 5. 実測例でも取り上げたように同じ Youtube ch においてもニュースの内容や性質によって反発する意見や合意する意見は発信源や中身によって異なる傾向も類似した傾向も存在すると推察する. 測定した意見分布は Case A と Case B では報道指針は異なることから推測するに, あるニュース報道や事象におけるある一定数の言及に関して比較的前向きな傾向であるか・後ろ向きな姿勢であるかを推察が可能な傾向として分布していることが観測できた. また, 今回はニュートラルな意見における分布も測定を行ったため, 一方軸の意見に分裂したケースを想定した計算におけるシミュレーションの結果とは異なる傾向を持つ. シミュレーション結果による意見分布はいくつかのクラスターに別れるケースが想定される. これは理論の中の人同士の相互作用をさらに改良する必要があることを示す. 実測との乖離に関しての理解を得たと共に, モデルにおける改良の余地を示唆される結果であった. また, 現実社会における傾向と相似関係もある傾向が社会学観点では推察できる要素もあったため考慮したい.

・6 今後より一層, 世論形成における意見交流や意思決定の場が発生しうるオンライン上において極端な世論形成や社会的な危機に関する議論の定量的な傾向の理解を促したい.

参考文献

- [1] [French J R P, 1956]French J R P (1956) A formal theory of social power., Psychological Review 63. pp. 181-194.
- [2] [Harary F, 1959]Harary F (1959)“ A criterion for unanimity in French ’s theory of social power ”. In Cartwright D (Ed.), Stud-ies in Social Power. Institute for Social Research, Ann Arbor.
- [3] [Abelson, R P, 1964]Abelson, R P (1964),“ Mathematical models of the dis-tribution of attitudes under controversy”. In Frederiksen, N and Gulliksen H (Eds.), Contributions to Mathemat- ical Psychology, New York, NY: Holt, Rinehart, and Winston.
- [4] [De Groot M H, 1974]De Groot M H (1974) Reaching a consensus. J. Amer.Statist. Assoc. 69. pp. 118 - 121.
- [5] [Lehrer K, 1975]Lehrer K (1975) Social consensus and rational agnoiology. Synthese 31. pp. 141 - 160.
- [6] [Chatterjee S, 1975]Chatterjee S (1975) Reaching a consensus: Some limit theorems. Proc. Int. Statist. Inst. pp.159 -164
- [7] [Lea, John, 1984]Lea, John, and Jock Young. "What is to be Done about Law and Order?." (1984).
- [8] [Miguel, 2016]Miguel, Juan Carlos, and Miguel Ángel Casado. "GAFAnomy (Google, Amazon, Facebook and Apple): The big four and the b-ecosystem." Dynamics of Big Internet Industry Groups and Future Trends. Springer, Cham, 2016. 127-148.
- [9] Shadbolt, Nigel, and Tim Berners-Lee. "Web science emerges." Scientific American 299.4 (2008): 76-81.
- [10] France yellow vest protests
< <https://www.bbc.co.uk/news/topics/cpzg2d6re0lt/france-yellow-vest-protests>>(Ref:2018/11/27)
- [11]山田善靖, 杉山学, and 八巻直一. "合意形成モデルを用いたグループ AHP." 日本オペレーションズ・リサーチ学会論文誌 40.2 (1997): 236-244.
- [12] 八巻直一, et al. "不満関数を用いる集団区間 AHP 法." 日本オペレーションズ・リサーチ学会論文誌 45.3 (2002): 268-284.
- [13]八巻直一, and 関谷和之. "不完全データを伴う大規模 AHP (決定理論とその関連分野)." (1998).
- [14] Hjarvard, S. (2008). The mediatization of society. Nordicom review, 29(2), 102-131.
- [15] Pak, Alexander, and Patrick Paroubek. "Twitter as a corpus for sentiment analysis and opinion mining." LREc. Vol. 10. No. 2010. 2010.
- [16] Agarwal, Apoorv, et al. "Sentiment analysis of twitter data." Proceedings of the workshop on languages in social media. Association for Computational Linguistics, 2011.
- [17] Siersdorfer, Stefan, et al. "How useful are your comments?: analyzing and predicting youtube comments and comment ratings." Proceedings of the 19th international conference on World wide web. ACM, 2010.
- [18] Wilson, T., Wiebe, J., & Hoffmann, P. (2005, October). Recognizing contextual polarity in phrase-level sentiment analysis. In Proceedings of the conference on human language technology and empirical methods in natural language processing (pp. 347-354). Association for Computational Linguistics.
- [19] 佐藤謙太, et al. "ネガボジ解析による Web データと株価変動の相関関係評価." (2015).
- [20] Lasswell, Harold D. "The measurement of public opinion." American Political Science Review 25.2 (1931): 311-326.
- [21] Lasswell, Harold D. "Why be quantitative." Language of Politics(1949): 40-52.
- [22] Galam, Physica A 238, 66 (1997).
- [23] Sznajd-Weron and J. Sznajd, Int. J. Mod. Phys. C 11,1157 (2000)
- [24] Sznajd-Weron, M. Tabiszewski, and A. M. Timpanaro, Europhys. Lett. 96, 48002 (2011).
- [25] Galam S, "Application of statistical physics to politics" Physica A: Statistical Mechanics and its Applications 274, 1999, Pages 132-139
- [26] Galam S, "Real space renormalization group and totalitarian paradox of majority rule voting" Physica A: Statistical Mechanics and its Applications 285, Issues 1-2, 15 September 2000, Pages 66-76
- [27] Galam S, "Are referendums a mechanism to turn our prejudices into rational choices? An unfortunate answer from sociophysics" Chapter 19 of The Routledge Handbook to Referendums and Direct Democracy edited by Laurence Morel and Matt Qvortrup, (Taylor & Francis, London, 2017)
- [28] Galam, S, Int. J.Mod. Phys.B31 (2017) 1742015
- [29] Hegselmann R and U Krause, "Opinion Dynamics and Bounded Confidence Models, Analysis, and Simulation" Journal of Artificial Society and Social Simulation 5 (2002)
- [30] Guillaume Deffuant, David Neau, Fr'ed'eric Amblard, and G'erald Weisbuch. Mixing Beliefs among Interacting Agents. Advances in Complex Systems, 3:87-98, 2000. 15.
- [31] G'erald Weisbuch, Guillaume Deffuant, Fr'ed'eric Amblard, and Jean-Pierre Nadal.
- "Meet, Discuss and Segregate!" Complexity, 7(3):55-63, 2002.
- [32] Akira Ishii and Yasuko Kawahata, "Opinion Dynamics Theory for Analysis of Consensus Formation and Division of Opinion on the Internet", Proceedings of The 22nd Asia Pacific Symposium on Intelligent and Evolutionary Systems (IES2018) 71-76; arXiv:1812.11845 [physics.soc-ph]
- [33] A.Ishii, H.Arakaki, N.Matsuda, S.Umemura, T.Urushidani, N.Yamagata and N.Yoshida; The 'hit' phenomenon: a mathematical model of human dynamics interactions as s stochastic process, New Journal of Physics 14 (2012) 063018 (22pp)
- [34] A Ishii and Y Kawahata, "Sociophysics Analysis of the dynamics of peoples' interests in society" Front. Phys., 08 October 2018 | <https://doi.org/10.3389/fphy.2018.00089>
- [35] A. Ishii, "Opinion dynamics theory considering trust and suspicion in human relations" Submitted to Proceeding of 19th International Conference on Group Decision and Negotiation in 2019 a Joint GDN-EWG/BOR meeting,
- [36] Bird, Steven, and Edward Loper. "NLTK: the natural language toolkit." Proceedings of the ACL 2004 on Interactive poster and demonstration sessions. Association for Computational Linguistics, 2004.

Multi-agent maximum discounted causal entropy 逆強化学習による報酬推定

Estimation of agent's rewards with multi-agent maximum discounted causal entropy
inverse reinforcement learning浪越圭一 荒井幸代
Keiichi Namikoshi Sachiyo Arai千葉大学大学院融合理工学府
Graduate School of Science and Engineering, Chiba University

We propose a entropy-base multi-agent inverse reinforcement learning method for constructing a multi-agent simulation. By using multi-agent inverse reinforcement learning, we can estimate the agent's behavior rule and the reward reflecting the purpose of the agent. In this paper, we extend maximum discounted causal entropy to markov game environment. Experimental results showed that the proposed method can estimate valid reward at small grid world.

1. はじめに

群衆、交通流、金融など、複数の行動主体が各自の目的に従い相互作用する現実の環境は、マルチエージェント系と呼ばれる。マルチエージェント系の振舞いを再現することで、各自の目的を理解し、行動を予測する研究は古くから取り組まれており、災害誘導や交通政策の評価法として重要な研究分野である。

マルチエージェント系の再現法の一つにマルチエージェントシミュレーション (MAS) がある。MAS は、行動主体をエージェントとして扱い、エージェントの観測から行動のマッピングを行動ルールとして記述する。そのため、エージェントの意思決定過程や行動目的を比較的容易に解釈可能である。一方、MAS は行動ルールから全体の振舞いをボトムアップに再現するため、行動ルールの設計に多くの試行錯誤と妥当性の説明を要する。著者らはこの問題に対し、全体の振舞いを観測した行動ログから、各エージェントの行動ルールを推定する枠組みを提案してきた [Namikoshi 18]。しかし、エージェントの目的を理解するには、推定した行動ルールを解析する必要がある。

そこで、マルチエージェント逆強化学習によるエージェントの目的推定に着目する。マルチエージェント逆強化学習は、マルコフ決定過程をマルチエージェント系へ拡張した Markov game において、行動ログからエージェントの報酬を推定する枠組みである。報酬は一般に状態・行動の価値を表すため、推定報酬の高い状態・行動がエージェントの目的を表すといえる。つまり、マルチエージェント逆強化学習により、報酬から各エージェントのもつ目的を容易に解釈できる可能性がある。

本論文では、エントロピー最大化原理を用いた新たなマルチエージェント逆強化学習を提案する。具体的には、infinite-horizon のマルコフ決定過程を対象とする Maximum discounted causal entropy 逆強化学習を Markov game へ拡張し、その解法を示した。実験では、2 人エージェントの簡易な GridWorld を対象に、決定的な Nash 均衡解の方策から生成した行動ログから妥当な報酬が推定可能なことを示す。

2. 対象問題

Markov game (MG) を $\langle \mathcal{N}, \mathcal{S}, \{\mathcal{A}_n\}_{n \in \mathcal{N}}, T, \{R_n\}_{n \in \mathcal{N}} \rangle$ の組で表す。 \mathcal{N} はエージェント集合 ($|\mathcal{N}| \geq 2$)、 \mathcal{S} は有限離散状態空間、 \mathcal{A}_n はエージェント n の有限離散行動空間、 $T: \mathcal{S} \times \mathcal{A}_1 \times \cdots \times \mathcal{A}_{|\mathcal{N}|} \times \mathcal{S} \rightarrow [0, 1]$ は状態遷移確率、 $R_n: \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ はエージェント n の報酬をそれぞれ表す。また、エージェント n の行動は $a_n \in \mathcal{A}_n$ 、全エージェントの結合行動は $a \in \mathcal{A}$ と記す。全エージェントは状態 $s \in \mathcal{S}$ にアクセス可能と仮定し、不完全知覚は扱わない。

本論文では、報酬が未知の $\text{MG} \setminus \{R_n\}_{n \in \mathcal{N}}$ において、全エージェントの行動ログ、すなわち可変長の軌跡集合 $\mathcal{D} = \{\{s_t, a_t\}_{t=0}^{t_d}\}_{d=1}^D$ から報酬 $\{R_n\}_{n \in \mathcal{N}}$ を推定する。以下、行動ログの行動主体をエキスパート、 \mathcal{D} をエキスパート軌跡と記す。報酬推定時、状態遷移確率 T を直接知ることはできないが、シミュレーション環境は利用可能とする。また、エキスパート軌跡を生成したエキスパート方策 π^E も得られないものとする。

3. 関連研究

3.1 模倣学習における逆強化学習の位置づけ

模倣学習 (Imitation Learning) とは、エキスパート軌跡からエキスパートの振舞いを再現する枠組みである。模倣学習は主に二つのアプローチに大別される。一つめの Behavioral Cloning は、エキスパート軌跡からエキスパートの振舞いを直接模倣する。そのため、実装が比較的容易な反面、エキスパート軌跡が十分に得られない場合、軌跡に含まれない状態行動への汎化性能が問題となる。二つめの逆強化学習 (IRL: Inverse reinforcement learning) は、エキスパート軌跡を生成したエキスパート方策を学習する。そのため、軌跡に含まれない状態行動においても適切な学習が期待できる。しかし、最適な報酬が複数存在する ill-posed 問題や、推定報酬から方策を計算する強化学習の計算コストが高いといった課題がある。

IRL は、エキスパート報酬の推定を目的とする Reward learning と、エキスパート方策の推定を目的とする Policy learning に分けられる。Reward learning は、エキスパート方策と他の方策とのマージンを最大化する手法 [Ng 00, Abbeel 04] や、最大エントロピー法に基づく手法 [Ziebart 10, Zhou 18] などが提案されている。一方 Policy learning は、主に敵対的学習を用いた手法 [Ho 16] などが提案されている。

連絡先: 浪越圭一, 千葉大学大学院融合理工学府都市環境システムコース, 千葉県千葉市稲毛区弥生町 1-33, acka2158@chiba-u.jp

3.2 マルチエージェント逆強化学習の分類

表 1 に、マルチエージェント逆強化学習 (MAIRL: Multi-agent IRL) の分類を示す。MAIRL は、定式化の目的関数と推定する報酬の構造によって分類できる。

表 1: マルチエージェント逆強化学習の分類

Reward structure	Objectives		
	Max-margin	Max-entropy	others
homogeneous	[Šošić 17]		
zerosum			[Lin 18] [Wang 18]
others	[Natarajan 10] [Reddy 12]	[Ziebart 10] [Bogert 18] [Song 18]	[Le 17]

[Šošić 17] は、swarm system において homogeneous なエージェントの報酬推定を提案している。[Lin 18, Wang 18] は、ゼロ和ゲームを対象とする MAIRL を提案している。しかしどちらの提案も、報酬に特殊な構造を仮定する必要がある、一般的な Markov game への適用は難しい。[Natarajan 10, Reddy 12] は [Ng 00] を拡張した MAIRL であり、[Natarajan 10] は中央制御器、[Reddy 12] は分散制御器を仮定しそれぞれ解いている。しかし前者は状態遷移確率を陽に必要とし、後者は Inner-Loop において Nash 均衡解を求める Nash Q-learning を用いる必要がある。[Ziebart 10, Bogert 18] は最大エントロピー法に基づいた MAIRL である。[Ziebart 10] は状態遷移が確率の場合に有効な Maximum causal entropy IRL を定式化し、3 体の pursuit-evasion において有効性を示している。[Bogert 18] はエキスパート軌跡の観測に隠れ (Occlusion) が生じる環境下の MAIRL を提案している。しかし前者は finit-horizon を対象とし、後者はエージェントが相互作用する状態での利得行列を必要とすることから、適用範囲が限られる。[Le 17, Song 18] はいずれも Policy learning を目的とした提案であり、エージェントの報酬を陽に推定しない。

本論文では、特徴ベクトル f と重み θ の線形和で報酬関数が表されると仮定し、最大エントロピー法に基づく推定法を提案する。[Ziebart 10, Bogert 18] とは infinit-horizon を扱う点、利得行列を必要としない点で異なる。

4. 提案法

4.1 定式化

マルコフ決定過程を対象とする Maximum discounted causal entropy IRL[Zhou 18](以下 MDCE IRL と記す) を、Markov game へ拡張した Multi-agent MDCE IRL(以下 M-MDCE IRL と記す) を定式化する。式 (1) から式 (5) に M-MDCE IRL の定義を示す。

$$\max_{\pi_t, t \geq 0} \sum_{n \in \mathcal{N}} H_{\pi_t, n, \pi_{-n}^E}(\pi_t, n) \quad (1)$$

$$\text{s.t. } \bar{f}_{n, \pi^E} = \bar{f}_{n, \pi_t, n, \pi_{-n}^E} \quad \forall n \in \mathcal{N}, t \geq 0 \quad (2)$$

$$\pi_{t, n}(a_n | s) \geq 0 \quad \forall a_n \in \mathcal{A}_n, s \in \mathcal{S}, n \in \mathcal{N}, t \geq 0 \quad (3)$$

$$\sum_{a_n \in \mathcal{A}_n} \pi_{t, n}(a_n | s) = 1 \quad \forall s \in \mathcal{S}, n \in \mathcal{N}, t \geq 0 \quad (4)$$

$$\pi_{t, n}(a_n | s) = \pi_{t', n}(a_n | s) \quad \forall s \in \mathcal{S}, a_n \in \mathcal{A}_n, n \in \mathcal{N}, t, t' \geq 0 \quad (5)$$

ここで、式 (1) はエージェント i の方策に対するエントロピーであり式 (6) で定義される。式 (2) は式 (7) の特徴期待ベクトル

を一致させる制約、式 (3) から式 (5) は方策に関する制約を表す。また、 $f_n : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}^k$ はエージェント n の特徴ベクトル、 \bar{f}_n は特徴期待ベクトルである。

$$H_{\pi_t, n, \pi_{-n}^E}(\pi_t, n) = \mathbb{E} \left[\sum_{t=0}^{\infty} -\gamma^t \log \pi_{t, n}(A_{t, n} | S_t) \right] \quad (6)$$

$$\bar{f}_{n, \pi_n, \pi_{-n}} = \sum_{t=0}^{\infty} \gamma^t \mathbb{E} [f_n(S_t, A_t)] \quad (7)$$

$$\pi(A_t | S_t) = \prod_{n \in \mathcal{N}} \pi_n(A_{t, n} | S_t) \quad (8)$$

4.2 Single-agent 系への分解と解法

M-MDCE 問題の解法は大きく 2 つ考えられる。一つめは、エントロピーに基づくマルチエージェント強化学習を Inner-Loop に用いることで、M-MDCE を Multi-agent 系のまま扱う方法である。しかし、マルチエージェント強化学習は状態行動空間の爆発や同時学習問題を扱う必要があり、方策が適切に学習できない可能性がある。

二つめは、各エージェントごとの目的関数へ M-MDCE を分解し、Single-agent 系として扱う方法である。具体的には、以下の M-MDCE のラグランジュ緩和問題を、各エージェントごとに解く。

$$\begin{aligned} & \max_{\pi_t, t \geq 0} \sum_{n \in \mathcal{N}} H_{\pi_t, n, \pi_{-n}^E}(\pi_t, n) + \theta_n (\bar{f}_{n, \pi^E} - \bar{f}_{n, \pi_n, \pi_{-n}^E}) \\ & \text{s.t. } \pi_{t, n}(a_n | s) \geq 0 \quad \forall a_n \in \mathcal{A}_n, s \in \mathcal{S}, n \in \mathcal{N}, t \geq 0 \\ & \sum_{a_n \in \mathcal{A}_n} \pi_{t, n}(a_n | s) = 1 \quad \forall s \in \mathcal{S}, n \in \mathcal{N}, t \geq 0 \\ & \pi_{t, n}(a_n | s) = \pi_{t', n}(a_n | s) \quad \forall s \in \mathcal{S}, a_n \in \mathcal{A}_n, n \in \mathcal{N}, t, t' \geq 0 \end{aligned}$$

Single-agent 系へ分解した場合、MDCE IRL[Zhou 18] と同様に各エージェントごとに MDCE IRL を解けばよく、Multi-agent 系における問題は生じない。しかし、対象問題の仮定から、報酬を推定するエージェント n 以外のエキスパート方策 π_{-n}^E は得られない。

そこで本提案では、エキスパート方策 π_{-n}^E を代替方策 $\tilde{\pi}_{-n}$ に置き換え推定する方法を提案する。Algorithm 1 に提案法のアルゴリズムを示す。まず、各 iteration において、報酬の重みと代替方策を更新するエージェント集合 $\tilde{\mathcal{N}}$ を選択する。次に、MDCE IRL により報酬の重み θ_n を更新する。更新の打ち切りは、特徴期待ベクトルが十分一致した場合か、打ち切り回数に達した場合とする。最後に、Soft Q-Learning[Zhou 18] により方策 π_n を更新したのち、 $\tilde{\pi}_{-n}$ を π_{-n}^E へ近づける「補完」を実行する。以下では、エージェントの選択法と、代替方策の更新及び補完法について述べる。

Algorithm 1 Multi-agent MDCE

Input: Markov Game $\setminus \{R_n\}_{n \in \mathcal{N}}$

Input: Expert trajectories \mathcal{D}

Output: reward weight $\{\theta_n\}_{n \in \mathcal{N}}$

Initialize policies $\{\pi_n\}_{n \in \mathcal{N}}$ and $\{\theta_n\}_{n \in \mathcal{N}}$

- 1: **for** iteration = 1, 2, ... **do**
 - 2: $\tilde{\mathcal{N}} \leftarrow \text{Selector}(\mathcal{N})$
 - 3: $\theta_n \leftarrow \text{MDCE}(\mathcal{D}^n, \tilde{\pi}_{-n}) \quad \forall n \in \tilde{\mathcal{N}}$
 - 4: $\pi_n \leftarrow \text{SoftQ}(\theta_n, \tilde{\pi}_{-n}) \quad \forall n \in \tilde{\mathcal{N}}$
 - 5: $\tilde{\pi}_n \leftarrow \text{Completion}(\pi_n, \mathcal{D}) \quad \forall n \in \tilde{\mathcal{N}}$
-

■エージェントの選択法

エージェントの選択法は二種類ある。図 1 にエージェントが 2 体の場合の更新手順と代替方策の流れを示す。一つめの選択法は、エージェントを 1 体ずつ選択して更新する Cyclic である。Cyclic は、1 体ずつ報酬をエキスパートへ少しずつ近づけていき、更新した報酬に対する方策を代替方策として用いる。二つめの選択法は、全エージェントの報酬を同時に更新する Parallel である。Parallel は、全エージェントの報酬更新を並列に実行しつつ、更新した報酬に対する代替方策を定期的に交換する。

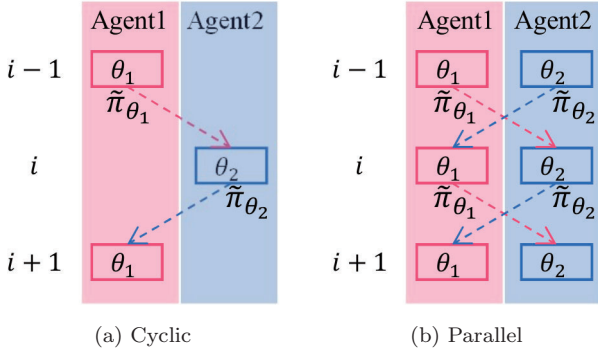


図 1: 報酬の更新順と代替方策の流れ: $|\mathcal{N}| = 2$ の場合. $i - 1, i, i + 1$ は iteration

■代替方策の更新・補完法

MDCE IRL において報酬 θ に対する方策 π_θ は式 (9), 式 (10) の Soft Bellman 方程式を満たすことが知られている [Zhou 18]. ここで $\text{softmax}_{a \in \mathcal{A}} Q_\theta^{\text{soft}}(s, a) = \log \sum_{a \in \mathcal{A}} \exp(Q_\theta^{\text{soft}}(s, a))$ とする. よって代替方策の更新には, Algorithm 2 に示す TD-base の Soft Q-Laerning を用いる. 最後に, 更新した代替方策に対し, エキスパート軌跡に含まれる状態・行動を確率 1 で取るよう, 代替方策を補完する. この操作は, 推定中の方策とエキスパート方策の確率分布を近づけることを意図している.

$$Q_\theta^{\text{soft}}(s, a) = \theta^\top f(s, a) + \gamma \sum_{s' \in \mathcal{S}} T(s'|s, a) V_\theta^{\text{soft}}(s') \quad (9)$$

$$V_\theta^{\text{soft}}(s) = \text{softmax}_{a \in \mathcal{A}} Q_\theta^{\text{soft}}(s, a) \quad (10)$$

$$\pi(a|s) = \exp(Q_\theta^{\text{soft}}(s, a) - V_\theta^{\text{soft}}(s)) \quad (11)$$

Algorithm 2 Soft Q-Learning

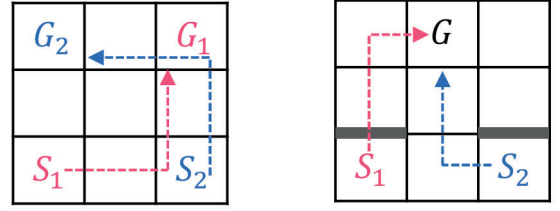
Input: reward weight θ_n , explore policy π , other agent's policy $\tilde{\pi}_{-n}$

- 1: **for** $t = 0, 1, 2, \dots$ **do**
- 2: Generate sample (s_t, a_t, s_{t+1}) from $\pi, \tilde{\pi}_{-n}$
- 3: $Q_n^{\text{soft}}(s_t, a_t, n) \leftarrow Q_n^{\text{soft}}(s_t, a_t, n) + \eta_t (s_t, a_t, n) \cdot$
- 4: $[\theta_n^\top f_n(s_t, a_t) + \gamma V_n^{\text{soft}}(s_{t+1}) - Q_n^{\text{soft}}(s_t, a_t, n)]$

5. 計算機実験

5.1 実験設定

図 2 に二つの実験環境を図示する. どちらの環境も 3×3 の GridWorld であり, 2 体のエージェントが各自のスタート座標 (S_1, S_2) からゴール G へ最短 step で到達することを目



(a) GW1: 決定的遷移

(b) GW2: 確率的遷移

図 2: 環境とエキスパート軌跡: エージェント 1, 2 に対して s_1, s_2 は初期座標, g_1, g_2 はゴール. 太線は障壁

的とする. 状態集合は全エージェントの座標の組み合わせ, 各エージェントの行動集合は $A_1 = A_2 = \{\text{up, down, right, left}\}$ であり, 1step で隣接する四方向のセルに移動できる. ただし, 壁に移動する場合と, 2 体のエージェントが同じセルへ移動する場合は, 遷移前の座標に留まる. 加えて, 後者の条件は, ゴール座標へ移動する場合を除く. GW2 では, スタート座標のセルと 1 つ上のセルの間に障壁があるため, S_1, S_2 で up がとられた場合は $1/2$ の確率で遷移に失敗する. いずれか, もしくは両方のエージェントがゴールへ到達した状態は吸収状態として扱う.

各環境のエキスパート軌跡には Nash 均衡解の一つを与える. 図 2 にエキスパート方策を矢印で示す. エキスパートは, 矢印に沿った座標でそれぞれの行動を決定的にとるものとする. この軌跡は, ゴールに到達したエージェントに $+100$, 同じセルに移動しようとした場合 -1 の報酬を与えたときの Nash 均衡解である [Hu 03]. 特徴ベクトルは全状態・行動対に対するバイナリベクトルとし, 特徴期待ベクトルは最大ステップ数 50 の軌跡を 100 本サンプリングから式 (7) で求める. 初期の重みは 0 ベクトル, 方策は一様分布とし, 各 iteration ではステップ幅 0.1 の最急降下法で MDCE IRL を最大 100 回更新する. また, Inner-Loop の Soft Q-Learning は 100 episode 学習する. 報酬の重み更新には正則化なしの場合と L2 正則化の場合をそれぞれ実験した.

5.2 実験結果

GW1 のにおける 10 試行平均および標準偏差の推移を図 3 と図 4 に示す. 図 3 はエージェントの選択が Cyclic の場合, 図 4 は Parallel の場合である. 横軸は iteration, 縦軸はエキスパート軌跡との特徴期待ベクトルの差のノルム表し, エキスパートと完全に一致するとき 0 となる. 結果から, GW1 ではエキスパート軌跡と完全に一致する報酬が推定したことがわかる. また, エキスパート軌跡による補完は収束までの iteration を減らし, Cyclic に比べ Parallel のほうが早く収束している. この結果の原因としては, 代替方策 $\tilde{\pi}_n$ がエキスパート方策 π_n^E に一致したとき分解された問題が元の M-MDCE IRL に一致することや, 補完により推定対象のエージェントがエキスパート軌跡上を遷移することを邪魔しなかったことが考えられる.

次に, 図 5 に GW2 に Cyclic を適用した結果を示す. GW2 では, 代替方策を補完しない場合, エキスパートと一致する解が得られず, 補完を用いてもエキスパートと完全に一致しなかった. この結果は, Parallel においても同様である. しかし, 推定報酬のうち最も値の大きな状態行動上位 5 組 (図 6) を確認したところ, エキスパート軌跡上に大きな報酬が置かれていることから, 結果が妥当であると確認できた. エキスパートと完全に一致しない理由としては, 環境の確率的遷移により, 特徴期待ベクトル計算のためのサンプル分布が一致しづらいことが考えられる.

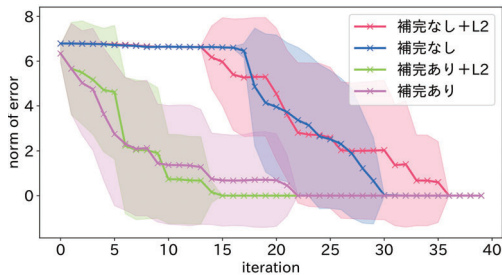


図 3: GW1 + Cyclic

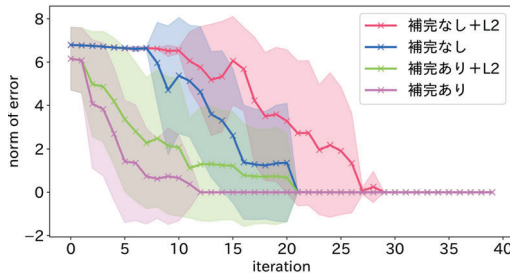


図 4: GW1 + Parallel

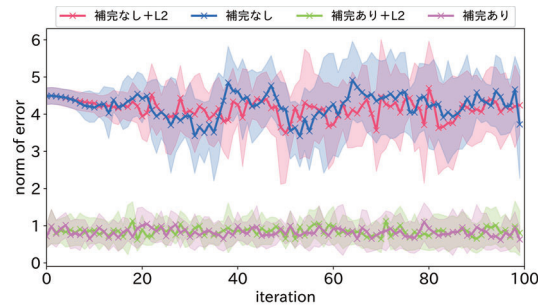


図 5: GW2 + Cyclic

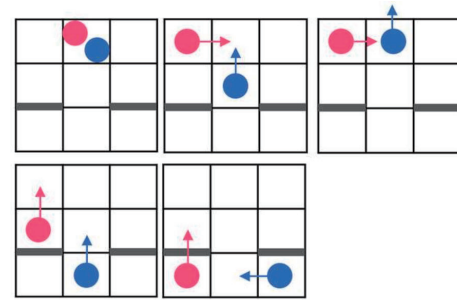


図 6: 推定報酬値の大きな状態行動 (GW2 + Cyclic)

6. 結論と今後の課題

本論文では、MAS におけるエージェントの行動ルール設計、およびエージェントの目的理解のため、新たな MAIRL を提案した。具体的には、infinite-horizon の Markov game において、最大エントロピー法に基づく MAIRL を提案した。提案法は、決定的なエキスパート軌跡から報酬が推定できることを、3x3 の GridWorld の実験から示した。今後の課題として、異なる環境や連続状態空間への適用、収束・最適性の考察や、獲得する均衡解概念の特定を挙げる。

参考文献

- [Abbeel 04] Abbeel, P. and Ng, A. Y.: Apprenticeship Learning via Inverse Reinforcement Learning, in *Proceedings of the Twenty-first International Conference on Machine Learning, ICML '04*, pp. 1–, New York, NY, USA (2004), ACM
- [Bogert 18] Bogert, K. and Doshi, P.: Multi-robot inverse reinforcement learning under occlusion with estimation of state transitions, *Artificial Intelligence*, Vol. 263, pp. 46–73 (2018)
- [Ho 16] Ho, J. and Ermon, S.: Generative Adversarial Imitation Learning, in Lee, D. D., Sugiyama, M., Luxburg, U. V., Guyon, I., and Garnett, R. eds., *Advances in Neural Information Processing Systems 29*, pp. 4565–4573, Curran Associates, Inc. (2016)
- [Hu 03] Hu, J. and Wellman, M. P.: Nash Q-learning for general-sum stochastic games, *Journal of machine learning research*, Vol. 4, No. Nov, pp. 1039–1069 (2003)
- [Le 17] Le, H. M., Yue, Y., Carr, P., and Lucey, P.: Coordinated Multi-Agent Imitation Learning, in *International Conference on Machine Learning*, pp. 1995–2003 (2017)
- [Lin 18] Lin, X., Beling, P. A., and Cogill, R.: Multiagent Inverse Reinforcement Learning for Two-Person Zero-Sum Games, *IEEE Transactions on Games*, Vol. 10, No. 1, pp. 56–68 (2018)
- [Namikoshi 18] Namikoshi, K. and Arai, S.: Estimation of the heterogeneous strategies from action log, in *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 1310–1317 ACM (2018)
- [Natarajan 10] Natarajan, S., Kunapuli, G., Judah, K., Tadepalli, P., Kersting, K., and Shavlik, J.: Multi-Agent Inverse

Reinforcement Learning, in *2010 Ninth International Conference on Machine Learning and Applications*, pp. 395–400, Washington, DC, USA (2010), IEEE

- [Ng 00] Ng, A. Y. and Russell, S. J.: Algorithms for Inverse Reinforcement Learning, in *Proceedings of the Seventeenth International Conference on Machine Learning, ICML '00*, pp. 663–670, San Francisco, CA, USA (2000), Morgan Kaufmann Publishers Inc.
- [Reddy 12] Reddy, T. S., Gopikrishna, V., Zaruba, G., and Huber, M.: Inverse reinforcement learning for decentralized non-cooperative multiagent systems, in *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 1930–1935, Seoul, Korea (South) (2012), IEEE
- [Song 18] Song, J., Ren, H., Sadigh, D., and Ermon, S.: Multi-Agent Generative Adversarial Imitation Learning, in Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R. eds., *Advances in Neural Information Processing Systems 31*, pp. 7471–7482, Curran Associates, Inc. (2018)
- [Šošić 17] Šošić, A., KhudaBukhsh, W. R., Zoubir, A. M., and Koeppl, H.: Inverse Reinforcement Learning in Swarm Systems, in *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems, AAMAS '17*, pp. 1413–1421, Richland, SC (2017), International Foundation for Autonomous Agents and Multiagent Systems
- [Wang 18] Wang, X. and Klabjan, D.: Competitive Multi-agent Inverse Reinforcement Learning with Sub-optimal Demonstrations, in Dy, J. and Krause, A. eds., *Proceedings of the 35th International Conference on Machine Learning*, Vol. 80 of *Proceedings of Machine Learning Research*, pp. 5143–5151, Stockholm, Sweden (2018), PMLR
- [Zhou 18] Zhou, Z., Bloem, M., and Bambos, N.: Infinite Time Horizon Maximum Causal Entropy Inverse Reinforcement Learning, *IEEE Transactions on Automatic Control*, Vol. 63, No. 9, pp. 2787–2802 (2018)
- [Ziebart 10] Ziebart, B. D., Bagnell, J. A., and Dey, A. K.: Modeling Interaction via the Principle of Maximum Causal Entropy, in *ICML* (2010)

[3Q3-J-13] AI application: analysis of physical behaviors in artifacts

Chair:Takuya Hiraoka Reviewer:Yoichi Sasaki

Thu. Jun 6, 2019 1:50 PM - 3:10 PM Room Q (6F Meeting room, Bandaijima bldg.)

**[3Q3-J-13-01] Construction of Dataset for Feature Extraction Performance
Evaluation using Aerial Photographs**○Hiroyuki Ohno¹, Ryo Endo¹, Takayuki Nakano¹, Masako Shinoda¹ (1. Geospatial
Information Authority of Japan)

1:50 PM - 2:10 PM

[3Q3-J-13-02] Slime detection during pile construction using machine learning○Sohei Arisaka¹, Yuki Tamagawa¹, Kojiro Takesue¹ (1. Kajima Corporation)

2:10 PM - 2:30 PM

**[3Q3-J-13-03] The optimization and comparison of methods for the Air foil design
using Deep Reinforcement Learning.**○Hitoshi Hattori¹, Kazuo Yonekura¹ (1. IHI Corporation)

2:30 PM - 2:50 PM

**[3Q3-J-13-04] Application of Gradient Booting regression toward the
Computational Fluid Dynamics in the Manufacturing industry**○Yutaro Ogawa¹, Takuya Shimizu¹, Toshiaki Yokoi¹ (1. INFORMATION SERVICES
INTERNATIONAL-DENTSU, LTD.)

2:50 PM - 3:10 PM

空中写真を用いた地物抽出性能評価用データセットの構築

Construction of Dataset for Feature Extraction Performance Evaluation using Aerial Photographs

大野 裕幸*¹
Hiroyuki Ohno

遠藤 涼*¹
Ryo Endo

中埜 貴元*¹
Takayuki Nakano

篠田 昌子*¹
Masako Shinoda

*¹ 国土地理院
Geospatial Information Authority of Japan

Several datasets are known as datasets for feature extraction. However, they are targeted to outside Japan and are not necessarily dataset with high regional diversity. Therefore, it is not suitable as an evaluation dataset for feature extraction in mapping performed as survey in Japan. So, we carried out this research with the aim of enabling the evaluation of feature extraction performance for high regional diversity and aerial photograph actually used in mapping in Japan. As a result, a dataset for evaluation was constructed using aerial photographs of 1328 regions taken in Japan since 1967. Furthermore, we compared the evaluation value of the prediction result using our dataset with that using the existing dataset using pix2pix and U-Net, and concluded that our dataset can perform sufficiently reliable evaluation.

1. はじめに

地図を作成する主要な手段の一つが写真測量である。写真測量では、空撮画像を元に、画像に写っている道路や建物等の地物の位置や形状等を読み取って地図データを作成する。そのような空撮画像から地物の位置と範囲を特定する工程を「判読」、判読した地物を地図上の座標を持ったデータにする工程を「図化」というが、これらの判読、図化の作業はこれまで測量技術者による手作業に頼らざるを得なかった。

一方で、CNN を始めとする画像認識技術が年々高度化しており、高分解能衛星の画像を用いた道路と建物抽出用のデータセット SpaceNet や、空撮画像を元に作成された建物抽出のベンチマーク用データセット Inria Aerial Image Labeling Dataset [Maggiori 17] (以下、「Inria データセット」という。)などが公表され、衛星画像を用いた地物抽出 (Feature Extraction) のコンテスト DeepGlobe 2018 [Demir 18] も実施されるなど、空撮画像から Feature Extraction を行う研究が盛んになりつつある。Feature Extraction は、測量の地図作成における判読及び図化の工程に類似していることから、これを測量の工程に導入して判読及び図化の一部でも自動化することができれば、地図作成の生産性向上に繋がるだけでなく、災害時の被害判読等の迅速化にも応用が可能と考えられる。そこで、筆者らは地上画素寸法 10cm~40cm の範囲で撮影された測量用の空撮画像を対象として、我が国における測量への導入を目標とした高性能な Feature Extraction の実現に取り組んでいる。

Feature Extraction の性能は、評価用のデータセットを用いて F 値や IoU などの数値によって定量的に評価することが一般的である。測量という実業務への導入を念頭に置いた場合、測量で用いられているデータと同じものを評価対象として、様々な季節、地域、気象条件の下で安定して均質な性能を発揮できるか否かを適切に評価しなければならない。そのためには、測量用に撮影された空撮画像から作成され、十分な多様性を有する評価用のデータセットが必要である。しかし、既存の評価用データセットは、日本国外の都市を対象としたものであるうえ、衛星画像から作成されたものが多く、地域多様性の観点からも十分とは言えない。そのため、まず、我が国を対象として十分な地域多様性を備え、測量用の空撮画像を対象とした Feature Extraction の性能評価用データセットを構築したので報告する。

2. 測量に用いられる空撮画像

測量に用いる空撮画像の撮影に関する仕様は、鳥瞰写真などとは異なり、要件が厳密に作業規程の準則[国土交通省 08]によって規定されている。そのため、測量に用いる空撮画像を一般的な空撮画像と区別して「空中写真」と記述することとし、まず初めに、空中写真と、既存の評価用データセットでよく用いられているオルソ画像について説明する。

2.1 空中写真の撮影方法

空中写真は、航空機に搭載された専用のカメラを用いて直下に向けて撮影される。かつては 23cm 幅のフィルムを使用したアナログカメラが用いられていたが、現在では全面的にデジタルカメラによる撮影に移行している。カメラパラメータは、焦点距離などの内部パラメータが事前に精密に検定されているうえ、航空機に搭載された GNSS (いわゆる GPS 等の衛星測位システム) と慣性計測装置 (IMU) によって、外部パラメータである写真撮影時のカメラの三次元位置とレンズ光軸の角度も極めて高精度に計測される。さらに、撮影後に画像とカメラパラメータを用いてバンドル法等で調整計算を行うことで測量の精度を保持するように設計されている。空中写真撮影用のデジタルカメラは主なもののだけでも 10 種類以上が使用されている。

2.2 空中写真の特性とオルソ画像

空中写真はカメラのレンズの中心に光束が集まる中心投影で撮影されるため、画像が写真の中心から周囲に向かって倒れ込むように写るという特性があり、土地の起伏に伴って歪みも生じるため、そのままでは地図とは重ならない。これに対し、中心投影による倒れ込みと、地形の起伏による歪みを取り除く画像処理 (正射変換) を行い、画像のどこであっても真上から見たような傾きの無い正しい大きさや位置に表示された画像に変換したものをオルソ画像という (図 1)。オルソ画像は、地図とぴったりと重ねることができる。そのため、オルソ画像を元画像とすれば、既往のポリゴン型の地図データを参照用のデータとして用いることが可能である。ただし、撮影年や撮影縮尺が異なるなど、参照用の地図データの作成に用いられた画像とは異なる写真から作成されたオルソ画像を用いた場合は、地図データと元画像に写っている内容の不一致が生じる場合がある。

一方で、空中写真では、中心投影であることを利用して左目用の画像と右目用の画像を用意し、簡易立体視鏡や立体視に

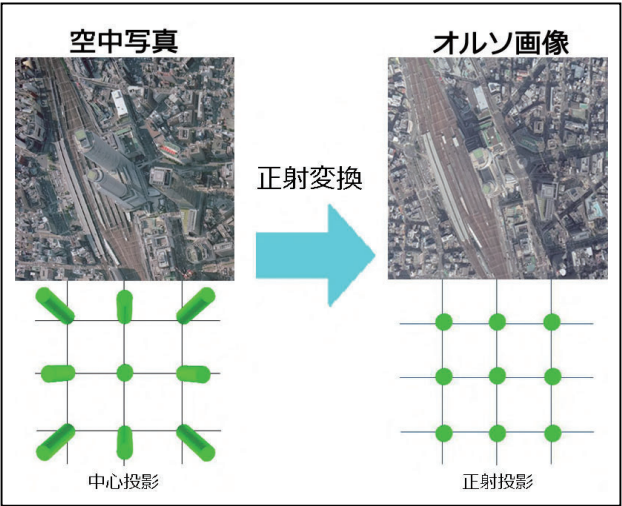


図 1. 空中写真とオルソ画像の違い

対応した眼鏡と図化ソフトを用いることで立体視をすることができる。元画像に対する適切な参照データが入手できない場合、立体視によって地物を判読し、使用する画像と完全に合致する高精度な参照データを作成することができる。筆者らは、立体視によって作成された参照データが評価用データセットとして最適であると考えている。

3. 多様性を持った評価用データセットの構築

空中写真から作成された既存の Feature Extraction 用データセットとしては Inria データセットが面積の観点から最大と思われる。Inria データセットは、地上画素寸法 30cm のオルソ画像を使用し、アメリカ 6 都市とオーストラリア 4 都市の合計 10 都市 810 km² のデータが公開されており、5 都市 405 km² 分はトレーニング用データとして参照データとセットで入手することが可能である。

しかし、日本における測量に用いることを目標とした Feature Extraction のトレーニング／評価データという観点で見ると、既存のデータセットには次のような問題点がある。

- すべて日本国外の都市のデータであり、日本とは地勢や地物の写り方が異なる。
- 数～10 都市程度のデータであり、地域多様性が少ない。
- 使用されている地上画素寸法が、日本の測量で一般的に用いられるサイズと異なる。日本では、20cm が用いられることが多く、10cm も用いられ、災害時は 40cm で撮影される場合がある。
- 多くが「建物」又は「道路」カテゴリのデータのみである。

そのため、国土地理院が保有し、測量用カメラで撮影された空中写真を元に、日本における測量に適した Feature Extraction 性能評価用のデータセット(以下、「GSI データセット」という。)を構築することとした。

3.1 GSI データセットの仕様

GSI データセットは、地域多様性の確保の観点から、原則として国土地理院がこれまでに測量用カメラで撮影した全ての地区から均等に範囲を抽出して構築した。過去に撮影された空中写真に対応するポリゴン型の地図データは作成されていないことから、オルソ画像ではなく、空中写真の画像データをそのまま使用し、比較的歪みが少ない各空中写真の中央を中心とする範囲を切り出して元画像とした。アナログカメラで撮影された空中写真は、1200dpi 又は 1270dpi でスキャンされた画像を使用し、[大野 17]の手法により写真中心 が画像中心となるよう補正した

画像から、同様に写真の中央を中心とする範囲を切り出して元画像とした。

元画像の主な仕様は次のとおりである。

- 地区数 549 地区(デジタルカメラ)
779 地区(アナログカメラ)
- 1地区あたりの切り出し枚数 1 枚
- 1枚あたりの切り出し範囲 572×572pixel
- 地上画素寸法 10cm～40cm
(各空中写真の撮影縮尺に基づく)
- カテゴリ分類 道路、建物、水域等 36 種類
(表 1 に示す)

表 1. GSI データセットのカテゴリ分類

番号	RGB	カテゴリ	番号	RGB	カテゴリ
2	#7F7F7F	舗装道路	22	#FF7F00	果樹園
3	#BF7F3F	未舗装道路	23	#7F0000	荒地
4	#5F5F5F	駐車場	24	#3F3F3F	空地
5	#FF007F	分離帯	27	#BFBF00	太陽光発電設備
6	#3F0000	軌道	28	#BF00BF	輸送管
7	#0000FF	水域	36	#BF7FBF	雪覆い等
8	#FF0000	堅牢建物	38	#7FBFBF	ブラットホーム
9	#FF00FF	普通建物	39	#FF7FBF	タンク
10	#FF7F7F	無壁建物	44	#FFBFBF	防波堤
11	#7F7F00	被覆	54	#BFFFBF	芝地
12	#BFBFBF	人工斜面	58	#FFBF00	砂礫地
13	#007FFF	水制	61	#BFBFFF	たたき
14	#7FBF00	歩道橋	62	#7F3F3F	岩がけ
15	#007F00	自然植生	65	#00007F	湿地
16	#FFFF00	畑	66	#BF00FF	材料置場
17	#00FFFF	水田	67	#7FBF7F	園庭
18	#007F7F	収穫後の水田	68	#7F7FFF	墓地
21	#7FFF7F	茶畑	255	#FFFFFF	未分類

各地区の写真は、写真枚数を n とした場合のファイル名順で n/2(小数部分切り捨て) 番目の画像を機械的に抽出した。そのため、全域が海域や森林という画像も一定数含まれる。

カテゴリは、まず道路(舗装道路、未舗装道路を合わせて 1 カテゴリとしたもの)について全データのラベル付けを完了しており、建物、水域、その他の順で順次追加中である。

各画像のラベル付けは、1328 枚のすべてを測量士で図化経験がある者が目視で判読及び確認することによって実施した。ラベル付けされた画像には、性能確認用の先行整備データとして位置付け、立体視を行わずに判読・分類したカテゴリ別画像と、精度よく性能を確認する正式データとして位置付け、すべて立体視により判読・分類した全カテゴリ画像の 2 種類がある。なお、全カテゴリ画像からあるカテゴリだけを抽出して作成したカテゴリ別画像は正式データとして位置づけられる。図 2 に元画像とカテゴリのラベル付けが完了した画像の例を示す。ラベル付けは、各カテゴリに表 1 に基づく Index 番号と塗分け色を定義を行い、全カテゴリ画像は、RGB と Index Color の両方で、カテゴリ別画像は RGB 画像で保存した。道路のカテゴリ別画像は、舗装道路と未舗装道路のカテゴリの分類色をいずれも RGB=#FF0000(赤色)としている。また、図 3 にデジタルカメラによる 549 地区分の元画像抽出位置の分布を示す。赤い点が抽出された画像の位置を示しており、北海道、秋田県、長野県、

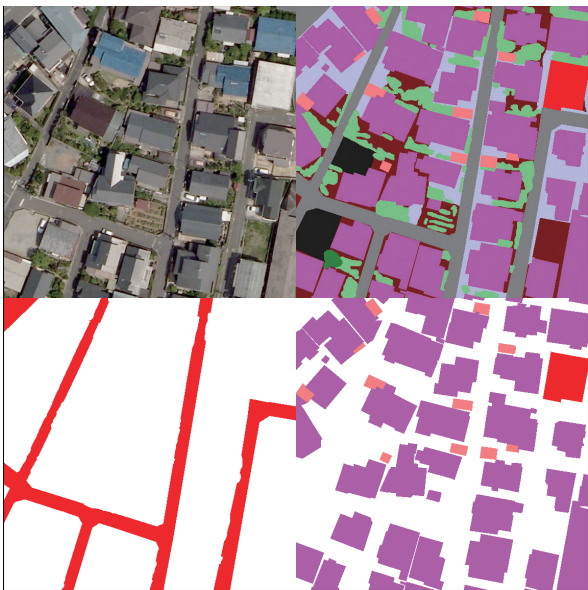


図 2. 元画像とラベル付け画像の例
(左上:元画像, 右上:全カテゴリラベル付け画像, 左下:
道路のカテゴリ別画像, 右下:建物のカテゴリ別画像)

岐阜県の山岳部を除き、ほぼ万遍なく抽出されていることが分かる。日本最東端の南鳥島は、デジタルカメラでの撮影が行われていない。三陸沿岸と熊本県にやや集中が見られる理由は、災害状況把握のための空中写真撮影が実施されたためである。

3.2 評価値の算出方法

カテゴリ別の性能評価は、F 値により行う。複数カテゴリの性能評価は IoU により行うことを原則とする。

4. Feature Extraction と性能評価の試行

4.1 使用したネットワーク

表現の異なる画像の相互変換に高い性能を示している pix2pix [Isola 17]と、pix2pix の Generator 部分に使用されている U-Net 構造の CNN の 2 種類を用いて Feature Extraction を実施した。いずれの入力画像、推論結果とも 256×256pixel の画像である。推論では、入力画像をいったん縦横とも 256 の倍数になるよう線形内挿法でリサンプリングし、256×256pixel に分割したうえでネットワークに入力し、推論結果は分割画像をマージして最近隣法で元の画像サイズにリサンプリングしている。損失関数は、pix2pix は cGAN×0.5+L1、U-Net は L1 とした。

4.2 トレーニングデータ

トレーニングデータは、GSI データセット構築に用いた空中写真をあらかじめ除外し、1967 年以降にカラーフィルム及びデジ

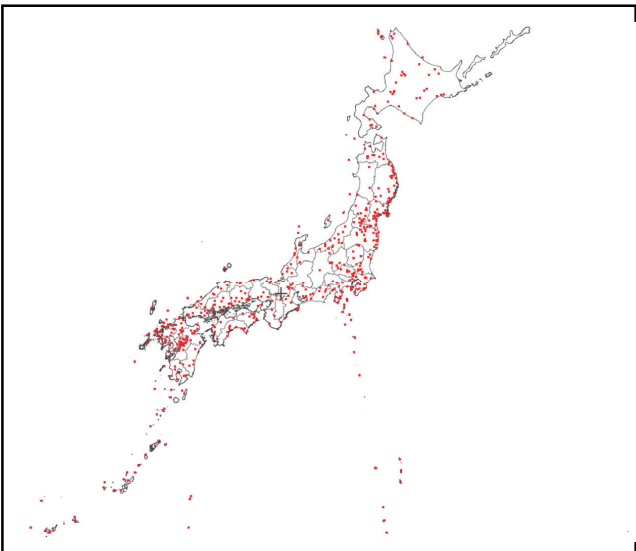


図 3. デジタルカメラ 549 地区分の抽出位置の分布

タルカメラで撮影された空中写真約 79 万枚の中から 286×286pixel のサイズで、道路、建物 2 種類のそれぞれについて先行整備データと同じ判読方法で作成した。トレーニング時は、ランダムに 256×256pixel のデータにクリッピングされる。作成したデータ数は、道路 3100 枚、建物 1700 枚である。さらに 90 度、180 度、270 度に回転させた画像を追加し、合計で道路 12400 枚、建物 6800 枚のトレーニングデータとした。

なお、トレーニングデータに Inria データセットの画像は一切使用していない。

4.3 性能評価の試行と考察

性能評価の試行は、GSI データセットによる性能評価値と、既存の評価用データセットによる性能評価値の傾向を比較することによって実施した。既存の評価用データセットには、Inria データセットでトレーニング用データセットとして提供されている 5 都市分の建物データを用いた。

Feature Extraction の推論結果を図 4 に、GSI データセットと Inria データセットで算出した建物の評価値を表 2 に、GSI データセットによる道路の評価値を表 3 に示す。なお、建物に関してはデータの整備途中であるため、デジタルカメラ 314 地区、アナログカメラ 633 地区による評価値である。また、GSI データセットの建物はコンクリート等の堅牢建物、普通建物、ビニールハウスやカーポートのような壁の無い無壁建物、建物以外の 4 カテゴリ分類としているが、建物と建物以外の 2 カテゴリの Inria データセットと仕様を揃えるため、堅牢建物、普通建物、無壁建物の 3 カテゴリの推論結果を建物 1 カテゴリとみなして評価した。

まず、pix2pix と U-Net による性能比較であるが、[Isola 17]はセマンティックセグメンテーションでは損失関数を cGAN よりも L1 とした方が良好な結果が得られるとしているように、本試行で

表 2. 建物抽出の評価結果

		GSI データセット			Inria データセット					
		Digital	Analog	All	Austin	Chicago	Kitsap	Tyrol-W	Vienna	All
Pix2pix	Precision	82.1	69.7	74.9	80.2	78.8	71.6	68.7	79.2	77.3
	Recall	67.4	44.7	53.0	45.8	47.7	56.6	71.2	43.3	48.0
	F 値	74.0	54.5	62.0	58.3	59.4	63.2	69.9	56.0	59.3
U-Net	Precision	86.6	74.9	79.5	83.1	79.4	75.0	73.9	79.6	80.8
	Recall	73.8	49.0	57.2	48.8	49.0	59.5	73.4	51.2	48.3
	F 値	79.7	59.2	66.5	61.5	60.6	66.3	73.7	62.3	60.5

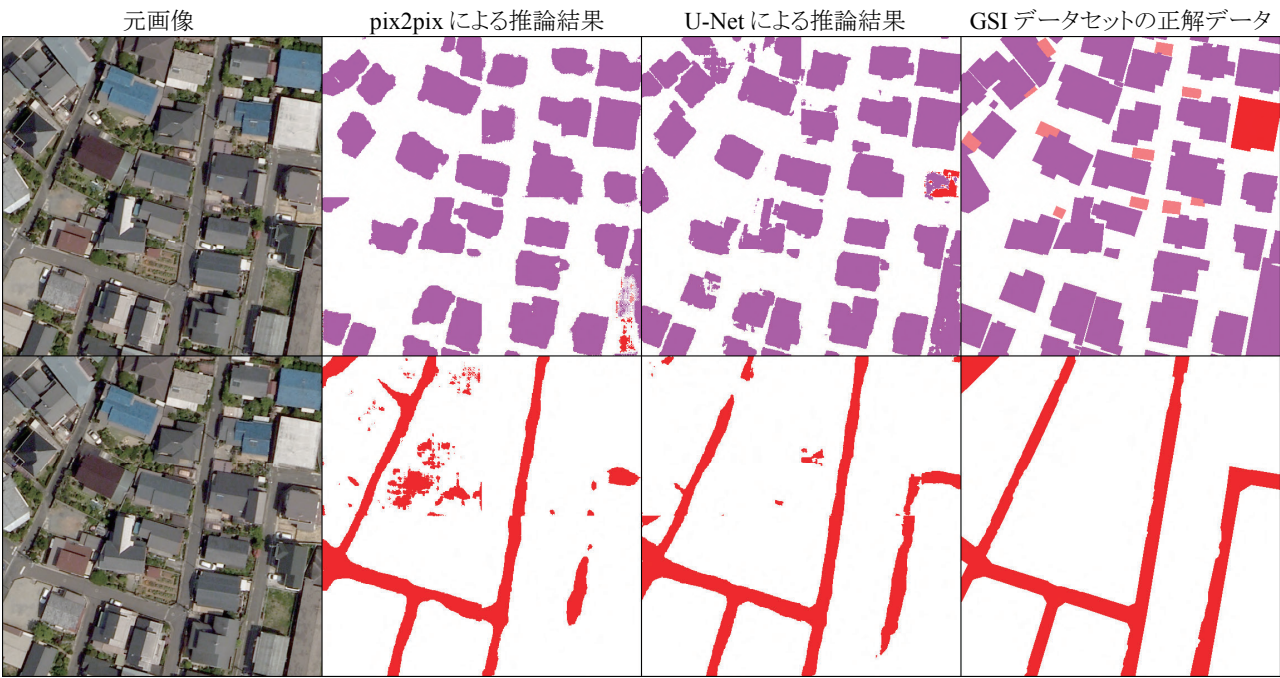


図 4. Feature Extraction の推論結果(上:建物, 下:道路).指定 RGB の各値から±5 の範囲を有効な pixel とみなしている.

表 3. 道路抽出の評価結果

		GSI データセット		
		Digital	Analog	All
Pix2pix	Precision	76.0	71.0	73.7
	Recall	67.2	47.9	57.0
	F 値	71.3	57.2	64.3
U-Net	Precision	82.6	75.9	79.4
	Recall	70.7	53.5	61.6
	F 値	76.2	62.7	69.4

もすべてのケースで U-Net の評価値が高い結果が得られた。

次に、建物の推論結果における、GSI データセット(デジタルカメラ)と Inria データセットでの評価値の比較では、GSI データセットの方の Precision がやや高く、Recall は大きく異なるという結果が得られた。Precision がやや高いのは、参照データに既往の地図データを用いず、すべての画像で測量技術者が判読してラベル付けを行ったことにより、参照データの位置ズレや不整合の量が相対的に少ないことが要因の一つと考えられる。一方、Recall の値が大きく異なるのは、トレーニングデータに日本国内の画像しか使用していないため、日本の都市の建物形状と大きく異なる形態、例えば、Vienna(ウィーン)はヨーロッパによくみられる古い石造りの街並みであるし、Austin は樹木に囲まれた住宅が点在する街並みであるなど、を呈する都市では、適切に推論できず、抽出されなかった建物が多いためであると推測される。この推測は、日本の建物や土地利用と似通っている Tyrol-W の Recall の値が GSI データセットの Recall の値と大きく変わらない点からも整合的である。GSI データセットと Inria データセットの間での評価値の傾向は一貫しており、面積は少ないものの多様性を重視した GSI データセットによる評価値が一定の信頼性を有することを証明していると考えられる。

なお、今後の空中写真撮影でアナログカメラが使用されることは無いことから、デジタルカメラによる評価値の重要性が高いことを付記する。

5. おわりに

測量における地図作成という実業務に Feature Extraction の導入を目指した研究に必要となる性能評価用のデータセットを構築し、その評価結果を既存のデータセットの評価結果と比較することで評価結果に信頼が置けるかどうかを考察した。このデータセットは、今後の研究における性能評価の指標として用いる予定である。

また、道路と建物が Feature Extraction の対象地物として用いられることが多いが、地図作成のためにはその他の多くの種類の地物の Extraction にも対応しなければならない。国家測量機関として基盤地図情報を整備・提供している国土地理院が地図作成能力評価のためのマルチカテゴリの性能評価用データセットを構築、公開することの意義は少なからずあると思われる、可及的速やかに GSI データセットの公開につなげたいと考えている。また、データセットの対象面積も順次拡大していく予定である。

参考文献

[Demir 18] Ilke Demir, Krzysztof Koperski, David Lindenbaum, Guan Pang, Jing Huang, Saikat Basu, Forest Hughes, Devis Tuia, Ramesh Raskar: DeepGlobe 2018: A Challenge to Parse the Earth through Satellite Images, arXiv:1805.06561v1, 2018

[Isola 17] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros: Image-to-Image Translation with Conditional Adversarial Networks, CVPR2017, 2017.

[国土交通省 08] 国土交通省:作業規程の準則, 国土交通省告示第 413 号, 2008.

[Maggiori 17] Emmanuel Maggiori, Yuliya Tarabalka, Guillaume Charpiat, Pierre Alliez: Can Semantic Labeling Methods Generalize to Any City? The Inria Aerial Image Labeling Benchmark, IEEE International Symposium on Geoscience and Remote Sensing (IGARSS),2017.

[大野 17] 大野裕幸: 空中三角測量の全自動化によるオルソ画像作成の効率化に関する研究, 国土地理院時報 129, 国土地理院, 2017.

機械学習による杭施工時のスライム検知 Slime detection during pile construction using machine learning

有坂 壮平^{*1}
Sohei Arisaka

玉川 悠貴^{*1}
Yuki Tamagawa

武居 幸次郎^{*1}
Kojiro Takesue

^{*1} 鹿島建設株式会社
Kajima Corporation

During pile construction, an inspection is needed to check absence of bottom slime which leads to settlement and inclination of structures. A conventional method for slime detection is dependent on individual judgement known by a sense of a hand. Therefore, there are some problems in terms of reproducibility and quantification. In order to solve these problems, we are studying a new method for slime detection using measured tension data. In this paper, we applied machine learning to judge whether slime exists or not from the tension data. Among 6 algorithms we compared, 1-dimensional Convolutional Neural Network achieved the best performance at 93% accuracy. According to this result, we verified that machine learning is effective for the slime detection.

1. 背景と目的

比較的規模の大きな構造物を支える杭の施工法として、場所打ちコンクリート杭工法が広く用いられている。本工法の施工過程で地盤を削孔する際には、孔壁が崩れないよう孔内を安定液（泥水）で満たした状態を維持する。削孔過程で安定液中に混入した土粒子が杭底に沈殿したものをスライム（図1）と呼んでおり、これを除去しなければ構造物の沈下や傾斜を招く恐れがある。そのため、スライムが除去できているか、確実に検知する必要がある。

スライムの検知法として、錘をロープに吊り下げ、錘が杭底に当たる際のロープの張力変化を手の感触で判定する方法が一般に採用されている（写真1、図2）[田中16]。この方法は人の手の感触に頼っているため、再現性、定量化という点で課題がある。そこで、筆者らは手の動きをアクチュエータで再現し、ロープの張力変化をロードセルで計測する、新たなスライム検知法（図3）を考案し、開発を進めている。本報では、この方法で得られた張力データからスライムの有無を判定する方法として、機械学習の適用性を検証した結果を示す。

2. 対象データ

張力データの計測には、アクチュエータの先端にロードセルを取り付け、その先に錘を吊り下げたロープを取り付けた装置を用いた。杭底付近に錘を一旦降ろした後、アクチュエータで錘を降下させ、錘が着底する前後のロープにかかる張力変化をロードセルで計測した。降下速度は全て表1に示す条件となるようアクチュエータで制御した。

計測は同一の杭に対して連続した10回の降下によって得られる10波を1計測としており、サンプリングレートは100Hz、1波あたりのサンプル点数は180点である。今回利用する張力データは複数の工事現場で行った120計測、1199波（欠損データを1件除外）とする。

錘の着底時の速度条件を同一とするため、一定速度で降下する表1-②の範囲で錘が着底した場合を適正に計測できたものとした。計測結果のラベルは、スライムが無い場合の「可」、有る場合の「不可」、正しく計測できなかった場合の「再計測」の3

連絡先：有坂壮平、鹿島建設株式会社、東京都調布市飛田給
2-19-1, 042-489-8259, arisakas@kajima.com

種とした。「可」と「不可」は、適正に計測できた波を、事前に行った熟練者の手による判定から分類した。「再計測」は、適正に計測できなかった波を機械学習用データとして意図的に再現したものである。

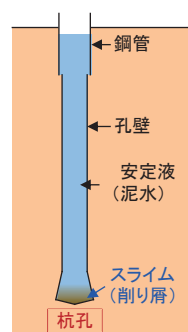


図1 スライム



写真1 スライム検知状況

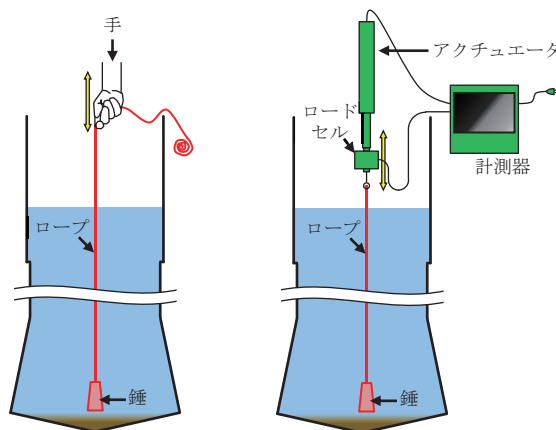


図2 一般的なスライム検知法 図3 考案したスライム検知法

表1 降下速度の制御値

	降下距離 (mm)	経過時間 (s)	降下速度 (mm/s)
①	0～50	0.0～0.5	0→200
②	50～290	0.5～1.7	200
③	290～300	1.7～1.8	200→0

「可」、「不可」、「再計測」のデータ数は 370 波、464 波、365 波である。「可」と「不可」の典型的な波形の例を図 4-a, b に示す。「可」の波形は張力が急激に低下しており、「不可」の波形は緩やかに低下している。図 4-c は、錘が着底した状態から計測開始した「再計測」の例で、最初から張力が低くなっている。図 4-d は、「可」のうち特異な波形の波を含む例である。

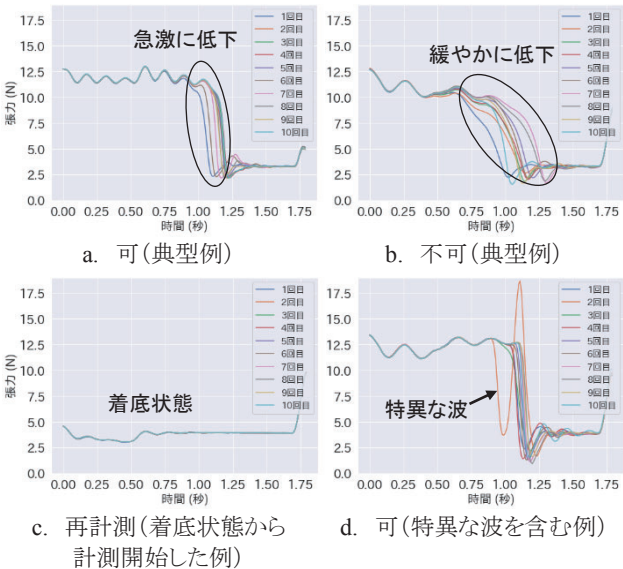


図 4 計測結果例

3. 機械学習によるスライム検知

3.1 方法

計測値、またはそこから計算される特徴量を入力として、1 波に対して「可」、「不可」、「再計測」のラベルを出力する機械学習モデルを構築した。アルゴリズムはロジスティック回帰 (LR)、サポートベクターマシン (SVM)、ランダムフォレスト (RF)、勾配ブースティング (GB)、全結合ニューラルネットワーク (MLP)、1 次元畳み込みニューラルネットワーク (1D-CNN) [Fawaz 18] の 6 つを比較した。

モデルへの入力は、1 次元畳み込みニューラルネットワークについては 180 点の全計測値とした。その他のアルゴリズムについては計測値、差分系列、累積和、各種統計量、パワースペクトル、ヒストグラム等により 159 の特徴量を作成し、特徴量選択ライブラリ Boruta [Kursa 10] で候補を絞った後、交差検証内で使用する特徴量を決定した。1 次元畳み込みニューラルネットワークについては、学習時にガウスノイズを付加してデータ拡張を行った。

データセットは 8:2 に分割し、8 割を訓練データ、2 割をテストデータとした。各アルゴリズムのハイパーパラメータは訓練データ内で 10 分割交差検証法によって決定し、テストデータは最終的な性能評価にのみ用いた。前処理、特徴量選択、データ拡張を含むハイパーパラメータのチューニングは、全てハイパーパラメータ最適化ライブラリ Optuna によって行った。また、データ分割の際にはモデルの性能を適切に評価するため、同一計測の 10 波が分割されず、ラベルの割合も均等になるよう分割した。モデルの評価指標は正解率 (Accuracy) とした。

3.2 結果

各アルゴリズムの性能を表 2 に示す。Training, Validation, Test は各交差検証で作成したモデルの訓練データ、検証デー

タ、テストデータでの正解率の平均、表中括弧内は標準偏差である。Test (Ensemble) は各交差検証で作成したモデルをアンサンブルした場合のテストデータでの正解率である。

最も性能の良かった 1 次元畳み込みニューラルネットワークと、特徴量ベースの手法で最も性能の良かったサポートベクターマシンの混同行列を図 5-a, b に示す。括弧内は実数である。

最適化された 1 次元畳み込みニューラルネットワークの構造を図 6 に示す。活性化関数は ReLU、畳み込み層のカーネルサイズは上から 3, 6, 6、ドロップアウト率は 0.27 である。

また、特徴量ベースの手法では、交差検証によって概ね 50 前後の特徴量が選択された。ランダムフォレスト、勾配ブースティングでの特徴量重要度上位 10 個を図 7-a, b に示す。

表 2 各アルゴリズムの性能

Algorithm	Training	Validation	Test	Test (Ensemble)
LR	0.941 (0.008)	0.896 (0.078)	0.847 (0.018)	0.858
SVM	0.954 (0.009)	0.892 (0.071)	0.876 (0.023)	0.879
RF	0.966 (0.008)	0.909 (0.078)	0.858 (0.016)	0.875
GB	0.998 (0.002)	0.884 (0.080)	0.852 (0.026)	0.871
MLP	0.971 (0.010)	0.874 (0.080)	0.821 (0.047)	0.858
1D-CNN	0.944 (0.029)	0.937 (0.065)	0.892 (0.060)	0.929

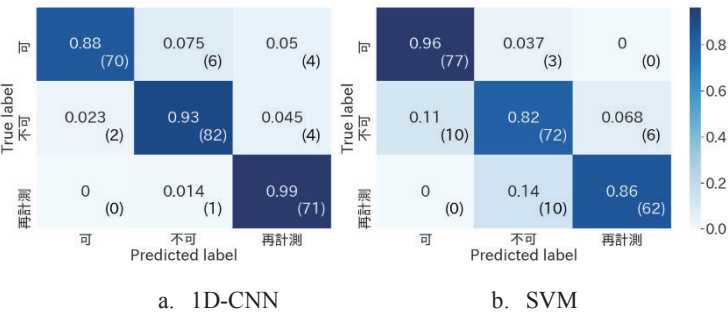


図 5 混同行列

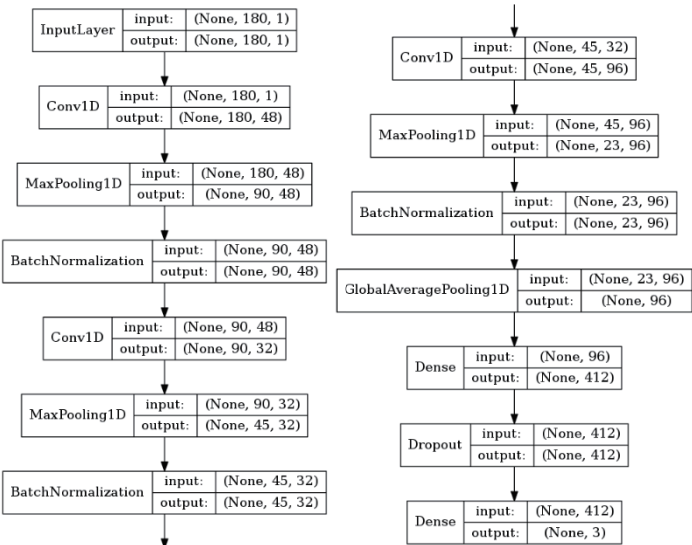
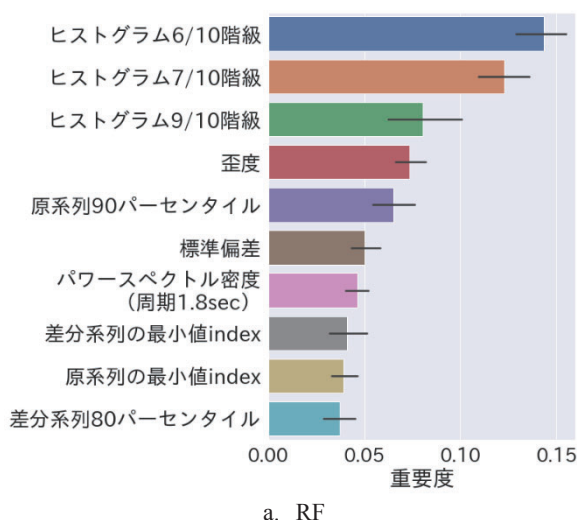
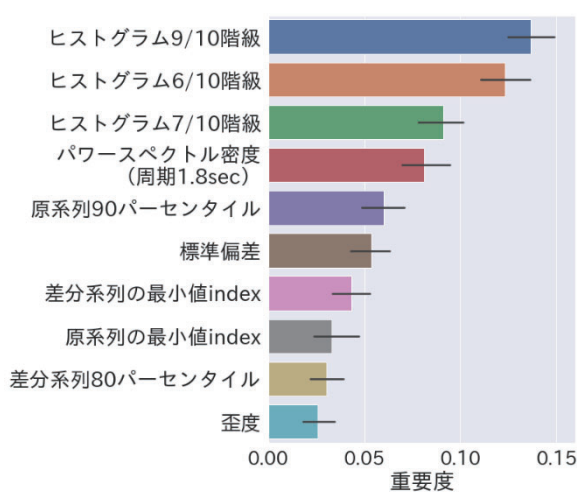


図 6 1D-CNN の構造



a. RF



b. GB

図7 特徴量重要度

3.3 考察

今回の検証では1次元畳み込みニューラルネットワークが最も性能が良く、データの特徴をよく捉えることができた。また、全結合ニューラルネットワーク、1次元畳み込みニューラルネットワークとも、他のモデルに比べてアンサンブルによる性能向上が大きくなった。これは作成されたモデルに多様性があることに起因すると思われる。本検証では原系列のみを1チャンネルで入力したが、差分系列や累積和を合わせて入力することも有効と思われる。

表2から、全体として訓練データと検証データのスコアに大きく差があり、モデルのバリエーションが大きいことが分かる。1次元畳み込みニューラルネットワークでは差が小さくなっているが、これは訓練データでのスコアがデータ拡張によるノイズ込みの性能であること、各交差検証で検証スコアが最大となるエポックを選択していることが理由であり、学習曲線を見るとやはり過学習の傾向があった。このことから、同様のモデルでも学習データ数を増やすことで性能向上が期待できる。

また、検証データとテストデータでのスコアを比べると、テストデータでは一様に性能が低下している。これは検証データとテストデータの分布が異なるためと思われる。また、交差検証内でも他の検証セットと比較して必ずスコアが低くなるものが存在した。ラベルで層化分割を行っているが、同一計測が分割されな

いようにしているために、データが偏りやすくなっているものと考えられる。適切な評価、学習のため、計測現場を考慮するなど分割方法をさらに検討すること、より多くのデータを使用することが必要である。

図5の混同行列を見ると、1次元畳み込みニューラルネットワークは「再計測」の再現率が高い、サポートベクターマシンは「可」の再現率が高い等、モデルによって出力の傾向が異なることが分かる。全体の正解率では1次元畳み込みニューラルネットワークの性能が最も良かったが、傾向の異なる特徴量ベースのモデルをアンサンブルすることも有効であると思われる。

特徴量ベースの手法では、図7から分かるようにヒストグラムに基づく特徴量がよく効いている。この特徴量はある区間に入るデータの点数を表しており、波形の形状を反映している。スライムが無い場合には急激に低下し、有る場合には緩やかに低下する張力変化の様子が捉えられているものと考えられる。

4. まとめと今後の展望

考案した方法で計測した張力データからスライムの有無を判定する機械学習モデルを構築し、適用性について検証した。比較した6種のアルゴリズムでは1次元畳み込みニューラルネットワークが最も性能が良く、その正解率はテストデータで約93%であった。これにより、スライム検知における機械学習の有効性が確認できた。

今後も多くの工事現場のデータを収集することで、機械学習モデルの精度を向上させ、スライムの有無をより高精度に判定できる方法の開発を目指す。

参考文献

- [田中 16] 田中 昌史: スライム処理方法とスライム量の関係, 基礎工 2016 年 3 月号, 総合土木研究所, 2016.
- [Fawaz 18] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, Pierre-Alain Muller: Deep learning for time series classification: a review, arXiv:1809.04356, 2018.
- [Kursa 10] Miron B. Kursa, Witold R. Rudnicki: Feature Selection with the Boruta Package Journal of Statistical Software, Vol. 36, Issue 11, 2010.

深層強化学習を用いた翼形状の最適化と手法の比較

The optimization and comparison of methods for the Air foil design using Deep Reinforcement Learning.

服部 均^{*1}
Hitoshi Hattori

米倉 一男^{*1}
Kazuo Yonekura

^{*1} 株式会社 IHI
IHI Corporation

When designing turbomachinery such as jet engines and superchargers, CAE is indispensable technology. In order to generate high performance shapes, the optimization methods such as response surface methodology and genetic algorithm has been used for design. However, these methods require many iterative calculations. When searching for a high performance shape against multiple flow conditions, it is necessary to repeat analysis every time the flow conditions are changed, which lengthens the design time. In this paper, to shorten this design time, we propose a new design method using deep reinforcement learning and compare of methods.

1. 諸言

航空エンジンや車両用過給機をはじめとする機械製品を設計する場合に、CAE (Computer Aided Engineering) は欠くことのできない技術である。設計者は CAE を用いて、例えば翼の周りの流体の流れを数値的に解析し、その物理現象を理解して実際に製品を作製した時の性能等を予測する。所望の制約条件を満たしたうえで最も性能が良い形状を作製するため、これまでは遺伝的アルゴリズム [Kalyanmoy 2005] や応答曲面法 [Raymond 2001] などを用いた最適化が行われてきた。これらの手法は多くの繰り返し計算が必要である。流れの境界条件や設計変数の増加などの仕様の変更があった場合には計算をやり直す必要があり、設計に多くの時間を要していた。本稿では、この設計時間を短縮するために、深層強化学習を用いた形状最適化を提案し、手法の比較を行うことで設計に有用な手法を選定する。設計作業では、仕様変更に合わせて、流れの条件を少し変えて形状を検討する作業が繰り返し行われる場合がある。このような場合、多数ある条件を包絡する条件であらかじめ多数の数値計算を実行して学習しておく、実際の設計時は学習済みモデルを用いて検討を行なうことで、実際の設計時間を短縮できると期待される。

2. Deep Q-Network

強化学習 (RL; Reinforcement Learning) では、特定の環境を与え、得られる報酬が最大になるように行動を学習する。強化学習については [Richard 1998] が詳しい。強化学習は図 1 に示すように、報酬を与える環境と、エージェントから構成される。エージェントが行動を起こした結果、環境側から報酬と現在の状態が出力される。このエージェントが現在の状態に応じて、最も多くの報酬が期待される行動をとるように学習が進む。

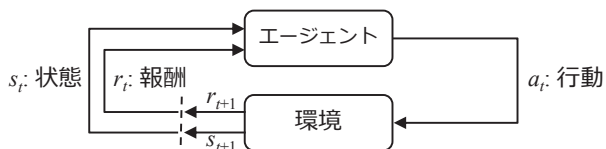


図 1. RL の概念図 [Richard 1998]

連絡先: 服部均, 株式会社 IHI, hitoshi_hattori@ihi.co.jp

強化学習は一般にロボット等の制御やテレビゲームに対して使用されており、時間変化する動的な問題に対して使用されることが多い。一方で文献 [Li 2017], [Andrychowicz 2016] ではニューラルネットワークのハイパーパラメータの最適化などの時間変化しない静的な問題における最適化に適用している。本報告で扱う問題も時間変化のない静的な問題である。行動選択の方法について様々な手法が提案されているが、本報告ではその 1 つである Deep Q-network (DQN) [Volodymyr 2013] を用いる。DQN は、次式で表す行動価値関数 (Q 関数) を元に行動を決定する。さらにこの Q 関数を深層ニューラルネットワーク (DNN) でモデル化する。

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \{r_t + \gamma \max_{a \in A} Q(s_{t+1}, a) - Q(s_t, a_t)\} \quad (1)$$

ここで、 $\alpha (0 < \alpha \leq 1)$ は学習率、 $\gamma (0 < \gamma \leq 1)$ は割引率であり、 A は取り得る行動全体を表す。式 (1) は第 2 項が行動価値の期待値と現在の見込みの値の差を表しており、この差分だけ現在の行動価値を更新する。Q 関数を DNN でモデル化する利点の一つが、状態量として画像を扱えることである。

3. NACA 翼の揚抗比最大化

航空機の翼型に用いられる NACA 翼を対象に揚抗比の最大化を目的とした数値実験を行い、手法ごとに性能の比較を行った。NACA 翼は、NACA6410 のように名称の数値が形状を一意に決めるパラメータを表す翼型である。図 2 に NACA4 桁系列の一例を示す。NACA4 桁系列は翼弦長で正規化された最大キャンバ、最大キャンバ位置、最大翼厚の 3 つの形状パラメータを持つ。ここでは翼型を固定し、膨大特性である揚抗比を最大にするように、迎角の最適化を行った。エージェントが迎角を増減することを図 1 に示す行動 a_t とした。迎角は 0 度から 40 度の範囲で 1 度単位で変化させる。また環境は翼形状と迎角を入力として、2 次元定常非圧縮流れの CFD (Computational Fluid Dynamics) 計算を行い、圧力コンター図と揚効比を出力するシステムとした。エージェントは、圧力コンター図を基に迎角の増減を決めるため、Q 関数のモデル化には CNN [Alex 2012] を使用した。深層強化学習の実装には ChainerRL [ChainerRL 2018] を用い、CFD の計算には OpenFOAM [OpenFOAM 2018] を使用した。境界条件は図 3 に示すように、

流入境界(青)を流速 1 m/s, 出口境界(赤)を静圧 0 Pa, 翼面(水色)を滑り無し境界, 解析領域の上下境界(黄)を滑り境界とし, 乱流には $k-\epsilon$ モデルを用いた. また翼弦長 1 m, 流体密度 1 kg/m^3 とし, 流体粘度 $1.0 \times 10^{-5} \text{ Pa}\cdot\text{s}$ とした. 学習時は初期の迎角をランダムに決定し, エージェントが 50 回迎角を変えるまでを 1 セットとし, 400 セットの学習を行った. 与える報酬として, 迎角を変えて揚抗比が上昇すれば報酬 $r_t=+1$, 低下すれば $r_t=-1$ を与えた.

図 4 にある翼型における検証時の初期解から収束解への推移を示す. θ_{attack} は迎角であり, 図 4 は迎角 37deg の初期条件に対し, 徐々に迎角を下げて迎角 6deg に収束していることを示している. 図 5 は初期解と収束解の圧力コンターである. 初期解では翼の前縁部で圧力が低くなり剥離が生じているが, 収束解では圧力が低い領域が減少していることが確認でき, エージェントが圧力コンター上で剥離を認識すると迎角を小さくするように行動したと考えられる.

次に強化学習の手法を変えて試行回数と得られる揚抗比の最適値との比の比較を行った. エージェントの行動が収束した迎角における揚抗比をエージェントの探索範囲内の揚抗比の最適値で割った値を評価値として使用した. 比較対象は DQN, DoubleDQN [Hado 2016], ResidualDQN, [Chainer RL 2018], Dynamic Policy Programming (DPP) [Mohamma 2012], Advantage Learning (AL) [Marc 2016], Persistent Advantage Learning (PAL) [Marc 2016], DoublePAL [Marc 2016], SARSA [Rummery 1994] である. 学習に用いた CFD 計算の回数と各手法の評価値を図 6 に示す. 少ない CFD の回数で安定して高い評価値が得られる手法が望ましい. CFD 計算の回数が 2,500 回以下の場合には AL が最も高い性能を示しているが, 5,000 回付近で評価値が1ポイント低下し, 学習回数によって得られる解が変わりうるということがわかる. DQN は 2,500 回以降でも高い評価値を維持していることから本例題においては最も有用な手法であるといえる. 一般的にレトロゲームで DQN より高性能とされる DoubleDQN は CFD の回数が 15,000 回以下の場合には 95% 以下の評価値となっていることから, DQN と同程度の精度を得るためには 20,000 回程度の CFD 計算が必要であり, 本課題には適さないことがわかった. その理由として, 本例題は1変数の最適化であり, 問題が単純であったため, シンプルな学習手法が適していた可能性が挙げられる. したがって目的関数が多峰性である場合や設計変数が多い場合は DQN 以外の手法が有用である可能性がある. 本比較により強化学習を用いて翼形状の最適化を行う上で, 選択する手法により必要な CFD の回数に大きな差があり, 本例題においては DQN が最も適していることを確認した. この結果は今後強化学習を用いて最適化設計を行う上で, 手法選択に役立てることができる.

4. 結論

深層強化学習を NACA 翼の揚抗比最大化に適用し, 目的に応じてエージェントが適切に迎角を変更できることを確認した. また手法ごとに学習に必要な CFD の回数を比較し, 本課題では DQN が最も適していた. 本比較により選択する手法次第で必要な CFD の回数に大きな差があることがわかり, 今後強化学習を用いて最適化設計を行う上で, 選択するべき手法の指針が示された.

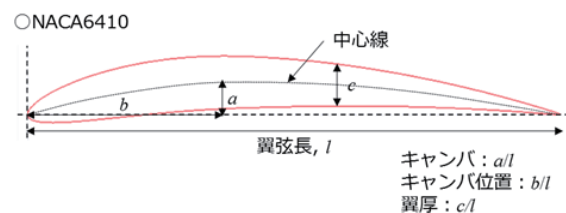


図 2. NACA4 桁系列の翼形状の例

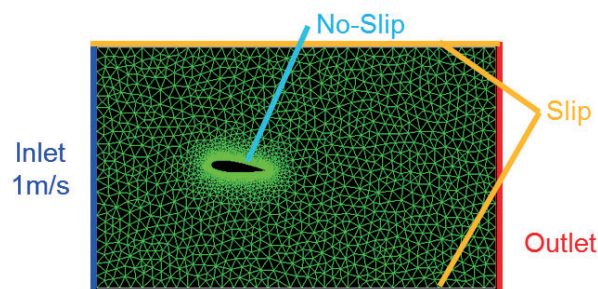


図 3. 境界条件

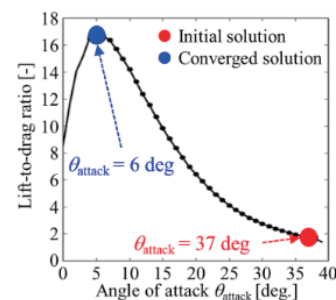


図 4. 初期解から収束解への推移

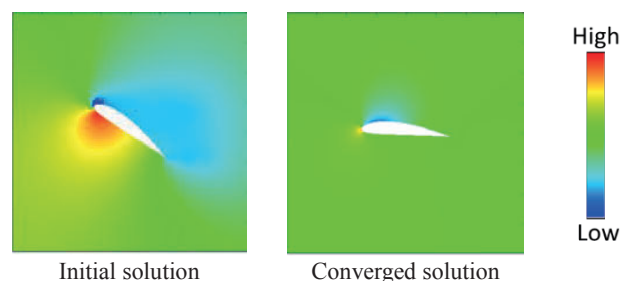


図 5. 初期解と収束解の圧力コンター

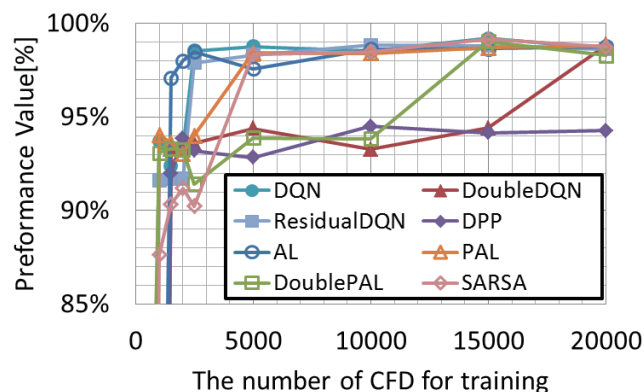


図 6. 各手法の評価値

参考文献

- [Kalyanmoy 2005] Kalyanmoy, D.: Multi-objective Optimization using Evolutionary Algorithm, John Wiley & Sons, 2005.
- [Raymond 2001] Raymond, H. M.: Response Surface Methodology: process and product optimization using designed experiments, second edition, A Wiley-Interscience publication, 2001.
- [Li 2017] Li, K., Malik, J. : Learning to optimize. In: ICLR 2017 conference, 2017.
- [Andrychowicz 2016] Andrychowicz, M., Denil, M., Gomez, S., Hoffman, M.W., Pfau, D., Schaul, T., Shillingford, B., de Freitas, N.: Learning to learn by gradient descent by gradient descent, arXiv preprint arXiv, 1606.04474v2, 2016.
- [Richard 1998] Richard S. S. and Andrew G. B.: Reinforcement Learning: An Introduction, MIT Press, 1998.
- [Volodymyr 2013] Volodymyr M., Koray K., David S., Alex G., Ioannis A., Daan W., Martin R. :Playing Atari with Deep Reinforcement Learning, NIPS Deep Learning Workshop, 2013.
- [Alex 2012] Alex K., Ilya S. and Geoffrey E. H.: Imagenet Classification with Deep Convolutional Neural Networks, NIPS, pp.1097–1105, 2012.
- [OpenFOAM 2018] OpenFOAM, <http://www.openfoam.org>, 2018/10/30.
- [Chainer RL 2018] Chainer RL, <https://github.com/chainer/chainer>, 2018/10/30.
- [Marc 2016] Marc G. B., Georg O., Arthur G., Philip S. T., Remi M.: Increasing the Action Gap: New Operators for Reinforcement Learning, Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, 2016.
- [Hado 2016] Hado V. H., Arthur G., David S.: Deep Reinforcement Learning with Double Q-learning, Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, 2016.
- [Mohamma 2012] Mohammad G. A., Vicenc G., Hilbert J. K.: Dynamic Policy Programming, Journal of Machine Learning Research 13, 2012.
- [Rummery 1994] Rummery, G. A., and Niranjan, M.: On-line Qlearning using Connectionist Systems, Technical Report, Cambridge University Engineering Department, 1994.

勾配ブースティング回帰木を用いた製造業流体シミュレーションの高速化手法

Application of Gradient Boosting regression toward the Computational Fluid Dynamics in the Manufacturing industry

小川 雄太郎^{*1}
Yutaro Ogawa

清水 琢也^{*1}
Takuya Shimizu

横井 俊昭^{*1}
Toshiaki Yokoi

^{*1} 株式会社電通国際情報サービス
INFORMATION SERVICES INTERNATIONAL-DENTSU, LTD.

Abstract: A faster calculation of MPS (Moving Particle Semi-implicit) method which is a computational fluid dynamics in the Manufacturing industry is proposed. Proposed method surrogates the semi-implicit part of MPS by the Gradient boosting regression trees using 10 original features as inputs. Finally, we confirmed that the qualitative properties of the proposed method coincide with the conventional MPS method by simulations of the dam-break problem.

1. はじめに

CAE (Computer Aided Engineering) とは製造業における製品および製品部品の設計時に、コンピュータ上で設計対象の to 特性や挙動をシミュレーション (数値計算) し、最適な設計条件を明らかにする技術である。CAE の例としては、製品部品の剛性を数値計算し軽くて壊れない部品設計を実現する構造最適化、製品内を流れる流体の抵抗や翼部品が持つ揚力などを所望の性能が得られる設計するための流体シミュレーションなどが挙げられる。

これら CAE 技術が共通して抱える課題のひとつはシミュレーション時間である。偏微分方程式を数値計算するため、高精度にシミュレーションするためには時間方向と空間方向に分解能が必要となり、計算時間が増大する。CAE を用いて新たなデザインを生み出すためにはこのシミュレーションにかかる時間を短縮し、様々な設計条件を容易に試せる計算環境を生み出すことが重要となる。

本研究では CAE 技術のなかでも流体シミュレーションに着目する。近年、流体シミュレーションの高速化という課題に向けて機械学習を適用する研究が進みつつある。具体的には従来のナビエ・ストークス方程式に従う物理モデルを、機械学習で事前に学習させた推論モデルに代替することで CAE の高速化を目指すものである。

例えば Guo らは、物体の周りを流れる空気の状態の流速場を求める際に、ナビエ・ストークス方程式を用いた数値計算を行うのではなく、物体形状を CNN (Convolutional Neural Network) に入力し、物体を囲む周辺領域の各座標点での流速を回帰するセマンティックセグメンテーションを応用した手法を提案している [GUO 2016]。また同様のスキームにより、翼の周りの流速から翼の揚力を求める手法も提案されている [ZHANG 2018]。これらの先行研究は時間発展を伴う数値計算を省略し、製品が持つ定常状態での特徴を、物体形状を機械学習モデルに入力して 1 回の推論計算で求める。

また、時間発展を伴う非定常な状態での数値時間を機械学習で代替する事例として、Fukami らは乱流の時系列変化の計算を機械学習で代替する手法を提案している [FUKAMI 2018]。

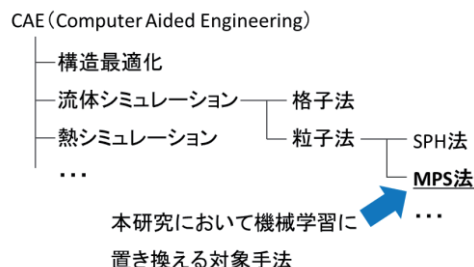


図 1. 本研究で機械学習モデルに置き換える MPS 法の CAE および各種流体シミュレーション手法における位置づけ

Fukami らの手法は計算対象領域内の各座標点における物理量を CNN エンコーダーで圧縮し、圧縮した特徴量を LSTM に入力して時系列発展した dt 後の特徴量を求め、求めた特徴量を CNN デコーダーで復元して dt 後の物理量を計算する。

しかしながら、これまで機械学習を用いた流体シミュレーション高速化手法は種々提案されているが、製造業における重要な流体シミュレーション手法のひとつである「MPS 法 (Moving Particle Semi-implicit Method)」について、機械学習による高速化が実現されていないという課題が存在する。

ここで、本研究で着目する MPS 法について図 1 を参照し解説する。流体シミュレーションはナビエ・ストークス方程式に基づき物理現象を数値計算するが、大きく「格子法」と「粒子法」に計算手法が分類される。格子法とは対象領域を格子状にメッシュで分割し、その格子点上の物理量を数値計算する手法である。いわゆる差分法や有限要素法であり、乱流の再現などが得意である。一方で、粒子法は流体を仮想粒子の集まりとして扱い、各粒子の時間発展に伴う移動を計算する手法である。この粒子は仮想的な流体の塊を想定しており、流体の分子とは異なる。粒子法は対象流体に対して計算用のメッシュを作る必要がないので、大変形や自由表面を伴う系、剛体と流体の連成解析系などのシミュレーションが得意である。例えば自動車エンジンでは、駆動時にピストンが上下運動しており、そのピストン内の燃焼ガスを閉じ込める目的とピストンを冷却する目的でオイルを噴出させている。このオイルがどのように挙動し、所望の閉じ込め・冷却動作をしているかをシミュレーションする際に粒子法が使用される [越塚 2014]。

連絡先: 小川雄太郎, 株式会社電通国際情報サービス 技術本部開発技術部, 東京都港区港南 2-17-1, 03-6713-607, ogawa.yutaro@isid.co.jp

本研究で着目する MPS 法は粒子法的一种である。粒子法の計算アルゴリズムには MPS 法と SPH 法 (Smoothed Particle Hydrodynamics Method) が存在する。SPH 法は銀河系の衝突など宇宙物理学におけるシミュレーションのために提案された手法であり [LUCY 1977], 近年では流体シミュレーションにも応用されている。ただし元来圧縮性流体 (速度が早い, マッハ数 0.3 以上) を取り扱うために提案された手法であったため, 非圧縮性流体 (流速がマッハ数 0.3 以下) を想定していない。そこで非圧縮性流体を粒子法として計算するために Koshizuka らによって MPS 法が提案された [KOSHIZUKA 1995] [越塚 2014]。

MPS 法と SPH 法は流体の非圧縮性以外にも SPH 法が陽解法 (fully explicit) で計算を行うのに対して, MPS 法は半陰的解法 (semi-implicit) で計算するという違いがある。MPS 法ではナビエ・ストークス方程式の粘性項, 重力・外力項は陽解法で, 圧力勾配項は陰解法で計算する。

MPS 法は非圧縮性流体を高精度に粒子法でシミュレーションすることができ, 製造業の製品設計時に広く活用されている。しかしながら MPS 法の計算時間を短縮するための機械学習手法の提案はこれまで行われていない。SPH 法においては, ランダムフォレスト回帰木を用いた機械学習手法が提案されているが [JEONG 2015], Jeong らの手法はただちに MPS 法に流用できるものではない。

そこで本研究では製造業の流体シミュレーションで重要な MPS 法を機械学習で高速化することを目的に研究を進める。以下, MPS 法の計算手法の概要を説明し, その後提案する機械学習手法を解説する。最後に提案手法の妥当性を流体シミュレーションのトイプロBLEMであるダム崩壊問題で検証する。

2. MPS 法

2.1 MPS 法の概要

図 2 に MPS 法の計算アルゴリズムの概要を示す。本稿では紙面の都合上概要の解説となるため, 詳細については [越塚 2014] などを参照されたい。

はじめにナビエ・ストークス方程式の粘性項, 重力・外力項を, 計算ステップ k での仮想粒子 (以下, 粒子と呼ぶ) の速度 \mathbf{u}_i^k , および粒子の位置 \mathbf{r}_i^k を用いて計算する。添え字 i は粒子番号を示す。粘性項, 重力・外力項は陽解法で解き, 粒子の時刻 Δt 後の仮速度 \mathbf{u}_i' と仮位置 \mathbf{r}_i^* を計算する (①, ②)。MPS 法において粘性項 $(\nabla^2 \mathbf{u}_i)^k$ の計算は, 対象とする粒子 i の近傍粒子の位置や速度に対してその相対距離に応じて重みを与える重み関数

$$w(\mathbf{r}_j - \mathbf{r}_i) = \begin{cases} \left(\frac{r_e}{|\mathbf{r}_j - \mathbf{r}_i|} \right)^2 - 1 & (|\mathbf{r}_j - \mathbf{r}_i| < r_e) \\ 0 & (|\mathbf{r}_j - \mathbf{r}_i| \geq r_e) \end{cases}$$

を使用して

$$(\nabla^2 \mathbf{u}_i)^k = \frac{2d}{\lambda^0 n^0} \sum_{j \neq i} (\mathbf{u}_j^k - \mathbf{u}_i^k) w(\mathbf{r}_j^k - \mathbf{r}_i^k)$$

と計算される。ここで r_e は粒子の近傍とみなす影響半径, ν は動粘性係数, d は対象としている計算領域の次元をそれぞれ示す。 λ^0 は初期状態での 1 粒子の影響範囲内にある近傍粒子との距離の 2 乗の重み平均値を示し, 次式で計算される。

$$\lambda^0 = \frac{\sum_{j \neq i} |\mathbf{r}_j^0 - \mathbf{r}_i^0|^2 w(\mathbf{r}_j^0 - \mathbf{r}_i^0)}{\sum_{j \neq i} w(\mathbf{r}_j^0 - \mathbf{r}_i^0)}$$

n^0 は粒子数密度の基準値であり, 以下の通りに計算される。

$$n^0 = \sum_{j \neq i} w(\mathbf{r}_j^0 - \mathbf{r}_i^0)$$

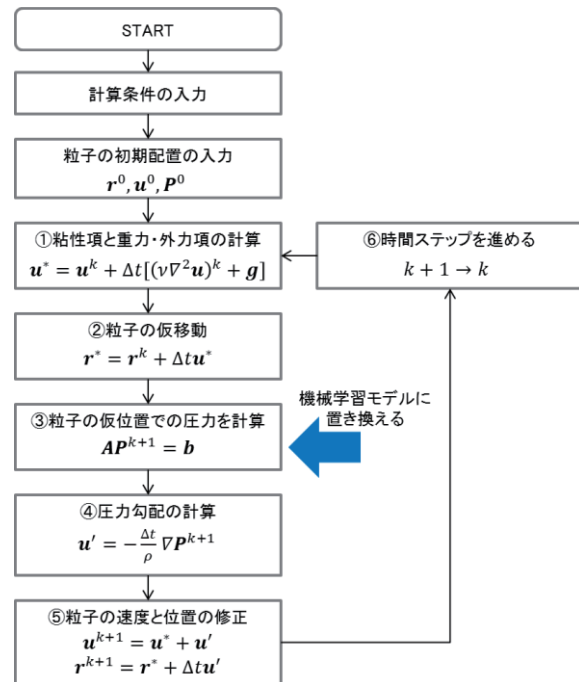


図 2 MPS 法概要および本研究で機械学習に置き換えるパート

次に粒子の仮位置において圧力 P^{k+1} を, 流体の非圧縮性条件から導かれる陰解法形式の式 $AP^{k+1} = \mathbf{b}$ を解いて求める (③)。 \mathbf{A} , \mathbf{b} の各要素は粒子の仮位置から代数演算で求められ,

$$a_{ij} = \begin{cases} -\frac{1}{\rho^0} \frac{2d}{\lambda^0 n^0} w(\mathbf{r}_j^* - \mathbf{r}_i^*) & (j \neq i) \\ \frac{1}{\rho^0} \frac{2d}{\lambda^0 n^0} \sum_{j \neq i} w(\mathbf{r}_j^* - \mathbf{r}_i^*) & (j = i) \end{cases}$$

$$b_i = \frac{1}{\Delta t^2} \frac{n_i^* - n_0}{n_0}$$

として計算される。ここで ρ^0 は流体の密度を示し, n_i^* は $n_i^* = \sum_{j \neq i} w(\mathbf{r}_j^* - \mathbf{r}_i^*)$ で計算される粒子の仮位置での粒子密度を示す。この圧力の計算式は陰解法形式であり, 陽的に計算することができない。そのためガウスの消去法や共役勾配法など, 連立一次方程式の任意の解法で解く。なお, 粒子の周囲に粒子が密に存在しない液面や飛沫の粒子の圧力 P_i^{k+1} は 0 とする。

求めた圧力 P^{k+1} を使用してナビエ・ストークス方程式の圧力勾配項を以下のように計算する。

$$\nabla P_i^{k+1} = \frac{d}{n^0} \sum_{j \neq i} \frac{P_j^{k+1} - \hat{P}_i^{k+1}}{|\mathbf{r}_j^* - \mathbf{r}_i^*|^2} (\mathbf{r}_j^* - \mathbf{r}_i^*) w(\mathbf{r}_j^* - \mathbf{r}_i^*)$$

ここで \hat{P}_i^{k+1} は粒子番号 i の粒子の圧力を近傍粒子との圧力差に負圧が生まれないように補正したものである [越塚 2014]。この圧力勾配項による速度変化を計算し (④), 仮速度と仮位置を修正する (⑤)。以上の①から⑤の計算を繰り返すことで, 時系列発展を求める (⑥)。

2.2 MPS 法的高速化における課題

MPS 法の計算の高速化において課題となるパートは, 粒子の仮位置での圧力の計算 (③) である。その他のパートは陽解法であり, さらに対象粒子の近傍の情報のみを利用するため, 解析領域を分割して計算機を並列させ高速化が可能である (ただし区画をまたがる粒子の移動情報はやりとりする)。しかしパート③は連立 1 次方程式を任意の解法で解いて求める陰解法の部分であり, 並列計算が困難である。そのため MPS 法の高速度の課題は圧力を求める陰解法部分にある。

3. 勾配ブースティング回帰木を用いた MPS 法の改良の提案

本研究では、MPS 法において圧力 P^{k+1} を求める陰解法部分を機械学習の回帰モデルに置き換え、計算機が並列状態のまま圧力の計算を可能にし、流体シミュレーションを高速化する手法を提案する。提案手法では粒子ごとに特徴量を計算して回帰モデルに入力し、各粒子の圧力 P_i^{k+1} を出力する。圧力計算以外は従来の MPS 法と同様である。

提案手法の回帰モデルに使用する特徴量について解説する。提案手法は各粒子の影響半径内にある近傍粒子の仮位置の情報を活用して 10 個の特徴量を作成する。

特徴量 1 を影響半径 r_e 内に存在する粒子 N_0 個との距離とする。特徴量 1 は N_0 次元のベクトルであり、各要素は昇順に配置する。特徴量 1 の N_0 個の要素 F_{1j}^i は次式で定義される。

$$F_{1j}^i = |\mathbf{r}_j^* - \mathbf{r}_i^*| \quad (|\mathbf{r}_j - \mathbf{r}_i| < r_e)$$

添字 i は圧力を求める対象の粒子を、 j はその近傍粒子を示す。

特徴量 2-5 は圧力の計算式で使用する行列 \mathbf{A} , \mathbf{b} の要素に関連する特徴量であり、 N_0 次元のベクトルとする。各特徴量の N_0 個の要素は次式で定義される。

$$F_{2j}^i = -\frac{1}{a_{ii}\rho^0} \frac{2d}{\lambda^0 n^0} w(\mathbf{r}_j^* - \mathbf{r}_i^*)$$

$$F_{3j}^i = -\frac{1}{a_{ii}\rho^0} \frac{2d}{\lambda^0 n^0} w(\mathbf{r}_j^* - \mathbf{r}_i^*) n_j^*$$

$$F_{4j}^i = -\frac{1}{a_{ii}\rho^0} \frac{2d}{\lambda^0 n^0} w(\mathbf{r}_j^* - \mathbf{r}_i^*) n_j^{**}$$

$$F_{5j}^i = \frac{1}{a_{ii}\Delta t^2} \frac{n_j^* - n_0}{n_0}$$

ただし n_j^{**} は粒子番号 j の粒子の仮位置 \mathbf{r}_j^* での相対粒子密度であり、 $n_j^{**} = \frac{n_j^* - n^0}{n^0}$ として計算する。各要素の順番は特徴量 1 で昇順にした近傍粒子 j の順番に合わせる。

特徴量 6, 7 は N_0 次元のベクトルであり、次式で定義される。

$$F_{6j}^i = n_j^*$$

$$F_{7j}^i = n_j^{**}$$

特徴量 8-10 はスカラーであり、次式で定義される。

$$F_8^i = \frac{1}{a_{ii}\Delta t^2} \frac{n_i^* - n_0}{n_0}$$

$$F_9^i = n_i^*$$

$$F_{10}^i = a_{ii}$$

以上により、特徴量ベクトルの要素数は合計で、 $(7N_0 + 3)$ 個となる。なお各特徴量において、影響半径 r_e 内に存在する粒子が N_0 個に満たない場合は、値 0 を代替として特徴量の要素とする。

提案手法の枠組みにおいて回帰手法はとくに制限されないが、本研究では勾配ブースティング回帰木を使用し、勾配ブースティング回帰木の LightGBM [KE 2017] を使用した。

4. ダム崩壊問題における実験と考察

4.1 流体シミュレーションの条件

提案手法の妥当性を定性的に確認するために、ダム崩壊問題を対象に MPS 法および提案手法による流体シミュレーションを実施した。シミュレーション次元は $d = 2$ 、シミュレーション範囲を囲む縦横の壁の長さは 1.0 [m] および 0.6 [m]、水柱の大きさは縦横 0.5 [m] および 0.25 [m] の範囲に含む粒子と設定し

た。水柱の位置は右端、左端、中央の 3 条件を用意した。粒子の直径は $l_0 = 0.01$ [m]、タイムステップは $dt = 0.001$ [s]、シミュレーション時間は $T = 4$ [s] までとした。粒子の影響半径 r_e については、粘性項で粒子径 l_0 の 3.1 倍、その他の項で 2.1 倍とした。

粒子を仮位置に移動させる際に、粒子径 l_0 の 0.5 倍以下の距離まで近づいた場合には衝突と判定し、反発率 0.2 の剛体衝突処理を行った。流体は水を想定し、密度 1000 [kg/m³]、動粘性係数 $\nu = 1.0 \times 10^{-6}$ とした。その他本稿 MPS 法の解説で省略したパラメータの設定は次の通りである。自由表面粒子の判定基準には、初期整列状態の密度 n^0 の 0.97 倍を使用した。数値計算の安定性を保つために、圧力の陰解法部分において用いられる緩和係数 γ は 0.2 とし、圧縮率は $\alpha = 0.45 \times 10^{-9}$ とした。

4.2 機械学習モデルの学習および推論条件

提案手法の学習および推論の条件を記載する。学習には水柱の位置が右端の条件での MPS 法のシミュレーション結果を使用した。訓練データとして時刻 0 [s] から時刻 2 [s] までの全粒子の情報を 0.01 [s] 間隔で 200 スナップショット分を使用した。特徴量は $N_0 = 28$ とする 199 要素とした。検証データは訓練データに対して 0.005 [s] だけ後ろにずらした時刻で、同じく 200 スナップショット分の全粒子のデータを使用した。LightGBM の実装には [MS LightGBM] を使用した。ハイパーパラメータの設定は、ブースティングする回帰木の最大数 $n_estimators$ を 30000 とし、その他はデフォルト設定のままとした。LightGBM の学習時に、損失計算に二乗誤差関数を使用して検証データによる early stopping を適用した。LightGBM はブースティング手法であり、設定した最大数まで回帰木を学習して訓練データに過学習するため、検証データによる early stopping が必要となる。

提案手法の定性的な動作確認のために、水柱の位置を左端および中央にした条件で流体シミュレーションを実施し、MPS 法による結果と提案手法の結果を比較した。

4.3 結果および考察

訓練データおよび検証データの行列サイズは (238237, 199)、および (238214, 199) となった。流体シミュレーション中に自由粒子と判定された粒子は圧力が 0 となり計算には使用されないため、訓練データおよび検証行列の行数はその分減少する。LightGBM で学習させた結果、回帰木の本数は 867 本で early stopping された。学習の結果、圧力 P_i^{k+1} の決定係数は訓練データが 0.70、検証データが 0.52 であった。図 3 に水柱の位置を左端および中央にしたシミュレーション結果を示す。それぞれにおいて、上段が MPS 法による結果を、下段が提案手法による結果を示す。図 3 上側において時刻 0 [s] で左端にあった水柱が、徐々に右方向へと崩壊し水が流れていく様子が時刻 0.3 [s] である。MPS 法と提案手法の挙動がほぼ一致していることが分かる。さらに時刻 0.9 [s] においては右端の壁にぶつかり水がはね返るが、こちらも挙動が一致していることが分かる。以降、水の塊は左右の壁にぶつかることを繰り返し、時刻 3.2 [s] での左端にぶつかったタイミングの挙動も一致している。また水柱が中央の場合も MPS 法と提案手法で挙動がほぼ一致している。

今回の実験において、訓練データには水柱右端条件で時刻が 0.0 [s] から 2.0 [s] までのデータを使用した。時刻 2.0 [s] 以降かつ水柱の位置が異なる条件でも定性的にうまく機械学習モデルが汎化されていることが明らかとなった。以上により、提案手法をダム崩壊問題でシミュレーションして、定性的な挙動が従来の MPS 法の結果と一致していることを確認でき、MPS 法の陰解法による圧力計算部分を、勾配ブースティング回帰木を用いた機械学習モデルに代替することが可能と確認された。

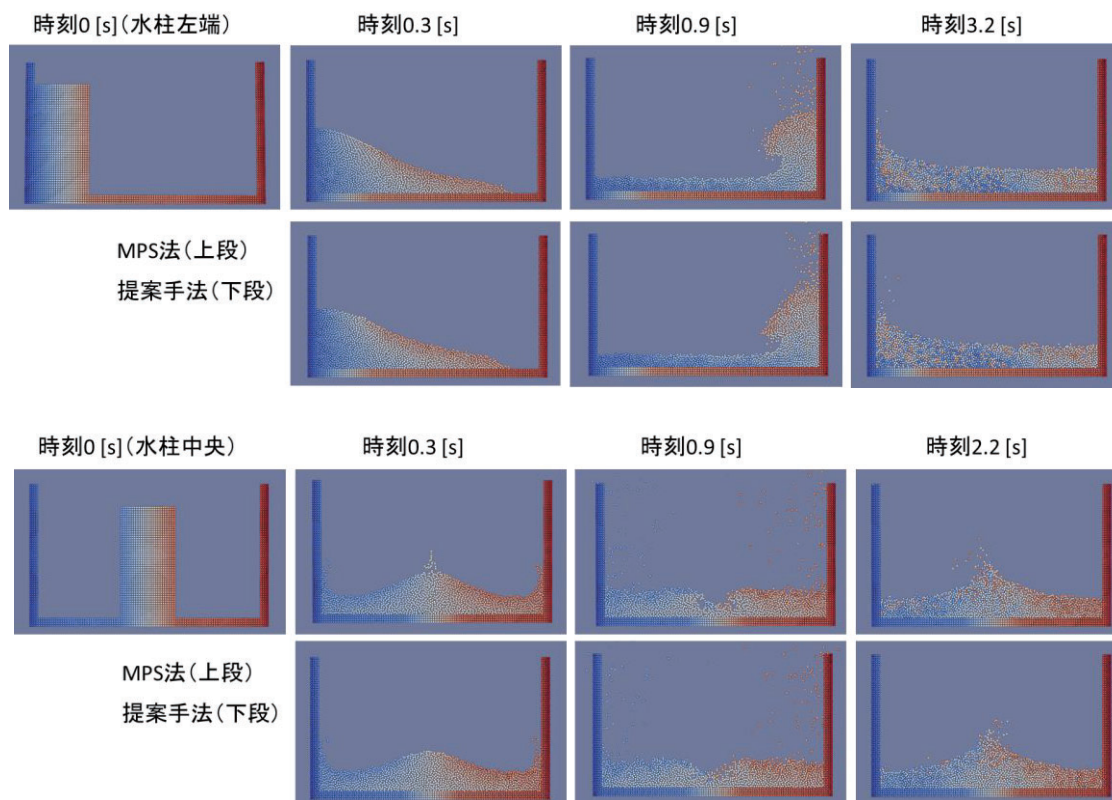


図3 提案手法と従来手法の流体シミュレーション結果の定性的比較

本研究では MPS 法において、論理的には高速化に貢献する手法を提案した。今後、実際に高速化にどの程度寄与するのか、粒子数を変化させながらその定量的特徴を確認していく必要がある。MPS 法の陰解法パートの計算時間は共役勾配法を用いた場合、粒子数の 1.5 乗に比例する。一方で提案手法の推論時の計算時間は粒子数の 1.0 乗に比例するため、高精度に計算したい場合や大規模な計算など、粒子数が増加するに従い提案手法の高速化が貢献すると期待される。

また本手法の精度を向上させるためには、特徴量の吟味が必要である。現在の 10 個の特徴量のうち寄与度の高いもののみを残すことや、さらなる特徴量の改良が考えられる。また LightGBM のハイパーパラメータをチューニングしていないため、パラメータチューニングによる精度向上も期待される。

本研究では定性的な挙動の確認として、流体シミュレーションのトイプロBLEMであるダム崩壊問題を使用した。今後、実際に製造業のエンジン設計など、実ビジネスで使用する流体シミュレーションの条件と規模において、提案手法がどの程度定性的および定量的に従来手法と一致するのか確認が必要である。

また提案手法の運用方法の検討も実ビジネスへの導入に重要となる。例えば流体が水やガソリンなど異なる条件の場合に、機械学習モデルで違いを吸収できるようにするのか、流体の種類ごとに訓練データを用意するのか検討が必要である。さらに訓練データの作り方も検討する必要がある。なんらかのトイプロBLEMの結果を使用するのか、それとも実際の実験条件において 1 度 MPS 法を実施して訓練データを作成するのかなど、実運用に向けた導入方法を検討する必要があると考えられる。

5. おわりに

本研究では製造業の流体シミュレーションで重要な MPS 法の一部を勾配ブースティング回帰木による機械学習を用いたア

ルゴリズムに代替し、流体シミュレーションを高速化する手法を提案した。提案手法の定性的挙動をダム崩壊問題で確認し、提案手法の結果が従来手法と同様の挙動を示すことを確認した。

参考文献

- [GUO 2016] GUO, Xiaoxiao; LI, Wei; IORIO, Francesco. Convolutional neural networks for steady flow approximation. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2016. p. 481-490.
- [ZHANG 2018] ZHANG, Yao; SUNG, Woong Je; MAVRIS, Dimitri N. Application of Convolutional Neural Network to Predict Airfoil Lift Coefficient. In: 2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. 2018. p. 1903.
- [FUKAMI 2018] FUKAMI, Kai; KAWAI, Ken; FUKAGATA, Koji. A synthetic turbulent inflow generator using machine learning. arXiv preprint arXiv:1806.08903, 2018.
- [LUCY 1977] LUCY, Leon B. A numerical approach to the testing of the fission hypothesis. The astronomical journal, 1977, 82: 1013-1024.
- [越塚 2014] 越塚誠一; 柴田和也; 室谷浩平. 粒子法入門. 2014.
- [KOSHIZUKA 1995] KOSHIZUKA, Seichii. A particle method for incompressible viscous flow with fluid fragmentation. Comput. Fluid Dyn. J., 1995, 4: 29.
- [JEONG 2015] JEONG, SoHyeon, et al. Data-driven fluid simulations using regression forests. ACM Transactions on Graphics (TOG), 2015, 34.6: 199.
- [KE 2017] KE, Guolin, et al. Lightgbm: A highly efficient gradient boosting decision tree. In: Advances in Neural Information Processing Systems. 2017. p. 3146-3154.
- [MS LightGBM] <https://github.com/Microsoft/LightGBM>

[3Q4-J-13] AI application: transformation system

Chair:Masahiro Tada Reviewer:Masayuki Otani

Thu. Jun 6, 2019 3:50 PM - 5:10 PM Room Q (6F Meeting room, Bandaijima bldg.)

[3Q4-J-13-01] Traffic anomaly detection using ETC2.0 probe data

○Atsuki Masuda¹, Masaki Matsudaira¹ (1. Oki Electric Industry Co.,Ltd.)

3:50 PM - 4:10 PM

[3Q4-J-13-02] Design method for high efficiency drone highway network

○Masatoshi Hamanaka¹ (1. RIKEN)

4:10 PM - 4:30 PM

[3Q4-J-13-03] Classification for time-sequence data appeared in shift control of automobile automatic transmission

○Yusuke Morikawa¹, Yasuhiro Ishihara¹, Takanori Ide¹, Eiji Moriyama¹, Taku Akita¹, Yasuo Tabei², Takehito Utsuro³, Hiroshi Nakagawa² (1. Aisin AW Co., Ltd., 2. RIKEN, Center for Advanced Intelligence Project, 3. Fclty. Eng, Inf. & Sys, Univ. of Tsukuba)

4:30 PM - 4:50 PM

[3Q4-J-13-04] Likelihood distribution of Pedestrian Trajectories rendered by Variational Autoencoder

○Yasunori Yokojima¹, Tatsuhide Sakai² (1. Siemens K.K., 2. Great Wall Motor)

4:50 PM - 5:10 PM

ETC2.0 プローブデータを利用した交通異常検知

Traffic anomaly detection using ETC2.0 probe data

増田 淳基
Atsuki Masuda

松平 正樹
Masaki Matsudaira

沖電気工業株式会社 経営基盤本部 研究開発センター
Corporate R&D Center, Corporate Infrastructure Group, Oki Electric Industry Co., Ltd.

Abstract: Detecting traffic anomalies such as accidents and obstacles on freeways is one of serious problem for traffic management. Previous researches are almost based on vehicle detectors to get traffic velocity and quantity, and therefore not able to detect the detailed location of traffic anomaly. In this paper, we present the algorithm for detecting automatically traffic anomalies using ETC2.0 probe data. Our algorithm is based on distances from head position of traffic congestion to recovered speed position, which is calculated from velocity of probe data. The result of evaluating the algorithm is very accurate, that precision is 94.3% and recall 89.2%.

1. はじめに

交通異常とは、事故や路上障害物等の交通を遮断する可能性がある事象のことを指し、平成 29 年に高速道路における交通事故が 8,758 件、路上障害物が 34.5 万件発生している[内閣府 18], [国土交通省 17]。しかし、道路管理者がそれら交通異常を認知する方法は、パトロールカーによる発見か通報によるものがほとんどである。高速道路内の交通異常は二次的な重大事故につながる可能性があるため、事象が発生した後に早期に検知可能な技術開発が急務となっている。

2. 関連研究

従来の交通異常検知は、道路に設置された車両感知器などの定点観測機器から得られる情報を利用するものが主流であり、交通量や交通密度などの交通状態の時間的・空間的变化から異常検知する手法が開発されている[Jeong 11], [成岡 14], [Chung 12]。成岡らおよび Chung らは、車両感知器が 300~600m に一台設置されている首都高速道路で実証を行ない、数百メートル単位で位置の検知が可能と報告している。また、Jeong らは、車両検知器が多数設置されている米カリフォルニアの I-880 で実証している。一方、日本国内の地方路線では車両感知器がインター間に1つ程度しか設置されておらず、平均間

隔が 10km 程度であるため、検知できてもその位置の特定が困難である。

3. データ特性

3.1 ETC2.0 プローブデータ

ETC2.0 プローブデータとは、ETC2.0 対応の車載器に走行履歴、挙動履歴を蓄積し、道路上に設置したアンテナによって収集する車両走行データである。データ内には各車両の GPS で測位した位置情報や速度、加速度等が含まれている(図 1)[瀬戸下 18]。収集したプローブデータを縦軸に時刻、横軸に起点からの距離(KP)、プロットの色に速度を対応させて図示すると図 2 のようになる。NEXCO 西日本によると、高速道路における渋滞の定義は「時速 40km 以下で低速走行、あるいは停止発進を繰り返す車列が 1km 以上かつ 15 分以上継続した状態」とあり、赤色(速度 ≤ 20 km/h)およびオレンジ色(20 km/h $<$ 速度 ≤ 40 km/h)のプロットが渋滞の速度に相当する[NEXCO 西日本]。一つの点が一つのプローブデータを表し、点を結んだ線が一台の車両の走行軌跡を表している。

3.2 イベントデータ

イベントデータは、管制センターにおいて入力された事故、路上障害物、工事、故障車、気象、交通規制などの事象に関するデータである。事象の分類や発生時刻、発生地点(KP)などが記録される。ただし、実際に発生した事象から、時間遅れが

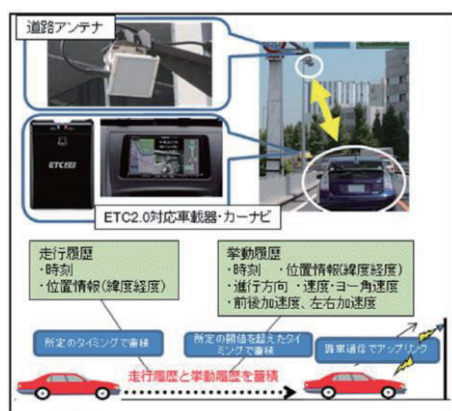


図 1 プローブデータ概念図

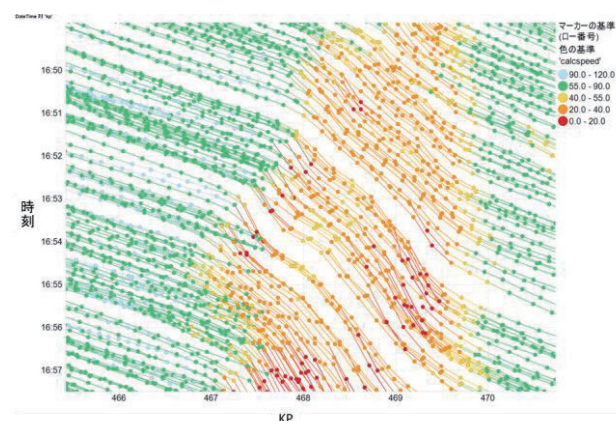


図 2 プローブデータ (各車両の走行履歴)

連絡先: 増田 淳基, 沖電気工業(株), masuda465@oki.com

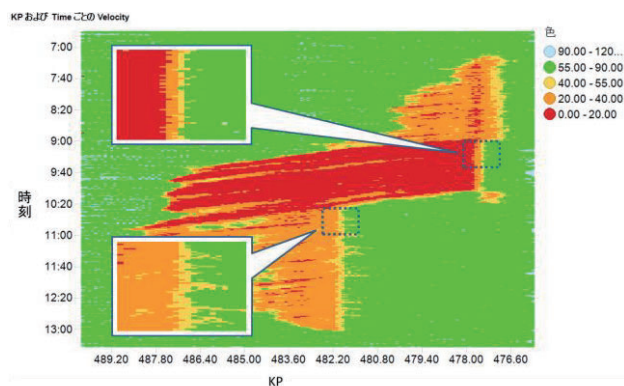


図3 自然渋滞と交通異常渋滞

生じている場合がある。

3.3 自然渋滞と交通異常渋滞

自然渋滞(交通集中渋滞)は、登り坂や合流地点など速度低下が慢性的に発生する地点において、交通容量以上の交通量が流入することにより、車両が停滞するために発生する。渋滞の原因となる場所が固定され、渋滞の先頭位置は変化しないことから、速度低下は交通量の上昇とともに徐々に進行方向に対して後方に伝播し、渋滞が伸びることになる。

一方、交通異常渋滞は、発生要因が事故や路上障害物などの事象発生にある。事象による走行車線の封鎖もしくは、交通規制の影響によって車線が一部遮断され、通行可能な交通量が大きく低下する。

図3に自然渋滞と交通異常渋滞が両方発生した渋滞図を示す。なお、図では、時間・空間を100m×1分単位のメッシュに区切り、メッシュ内に含まれるプローブデータの平均速度を算出したデータを可視化している。図の上部右側が交通異常渋滞であり、9:02に477.00KPで事故、9:18から10:03まで477.00KPから477.80KPにかけて車線規制(追越車線)が記録されている。一方、図の下部中央が自然渋滞である。

4. モデル化と学習

自然渋滞と交通異常渋滞では、渋滞先頭の速度変化パターンに違いが見られる。また、自然渋滞においても、「サグ部」と呼ばれる登り坂、インターチェンジやジャンクションなどの合流部、トンネル入口、インターチェンジ出口において発生することが知られており、それぞれ速度変化パターンに違いがあると考えら

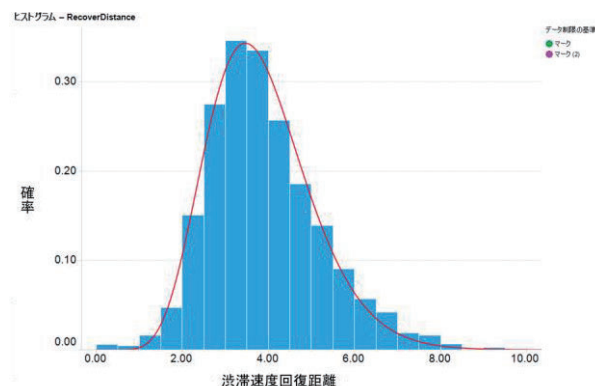


図5 ガンマ分布によるモデル化

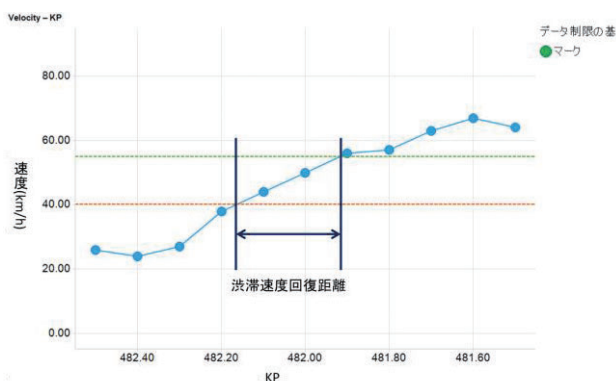
れる。さらに、登り坂でも勾配や距離、大型車の混入率によっても違いがあるだろう。しかしながら、それらの違いを変数としてモデル化・学習するには、プローブデータが十分な量とは言い難い。したがって、今回は自然渋滞を統合してモデル化することとした。

4.1 速度回復距離

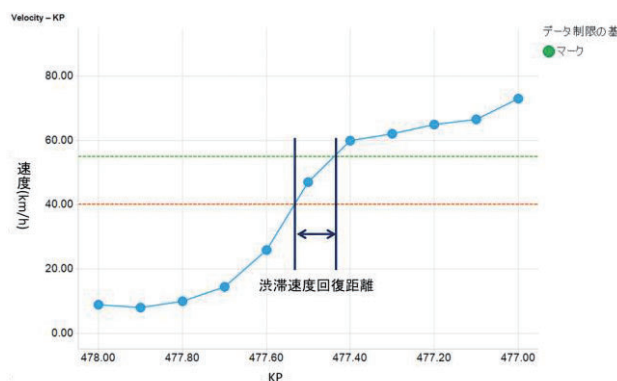
速度変化パターンの指標として、渋滞速度回復距離を利用した。これは低速車列の先頭付近において、低速状態から高速状態になるまでに要する距離のことである。一例として、図3の自然渋滞と交通異常渋滞での速度変化を図4に示す。距離の算出は、100mメッシュの速度を線形補完し、渋滞の定義速度である40km/hの地点と速度が回復したと考えられる55km/hの地点までの距離とした。

4.2 ガンマ分布によるモデル化

交通異常を検知するため、自然渋滞における渋滞速度回復距離を用いて、距離に応じた発生確率をモデル化する。イベントデータが記録されておらず、同じ地点(KP)で複数の日に発生した一連の渋滞における各時刻の渋滞速度回復距離の頻度分布は図5のようになる。距離が正の値をとり、分布が左右非対称であるという特徴から、確率モデルにガンマ分布を用いた。ガンマ分布の確率密度関数は以下の式(1)に示す通りである。ここで、 s が形状パラメータ、 r がスケールパラメータであり、ともに正の値をとる。頻度分布から、両方のパラメータについて、最小二乗法により最も近似する解を学習する。



(a) 自然渋滞



(b) 交通異常渋滞

図4 渋滞速度回復距離の算出

$$P(x) = \frac{r^s}{\Gamma(s)} x^{s-1} e^{-rx} \quad \dots(1)$$

4.3 交通異常検知

各時刻に取得したプローブデータに対して、渋滞が発生している場合に、過去数分の渋滞速度回復距離を算出し、学習した自然渋滞の分布の累積確率から、同時生起確率を算出する。同時生起確率が正規分布の 3σ に相当する累積確率 (0.135%) 未満の場合に、交通異常と判定することとする。なお、なるべく遅延なく検知するため、過去 2 分の渋滞速度回復距離を使用する。

5. 検証実験

本アルゴリズムを R 言語で実装し、ETC2.0 プローブデータとイベントデータの実データを用いて、検証実験を行なった。

5.1 実験データ

今回利用したデータは、NEXCO 西日本から提供いただいた関西地区における複数路線のプローブデータとイベントデータである。路線は、中国自動車道、名神高速道路、近畿自動車道であり、2018 年 3 月 10 日～3 月 26 日、および、交通量が多い特異日として 2018 年 8 月 10 日～8 月 15 日 (お盆期間) のデータ計 48 日分を検証に用いた。

5.2 実験手法

検知位置については、管制センターでの入力を正解値として異常発生位置から $\pm 500\text{m}$ 以内の範囲を検知できるかどうかの精度を計算することとした。精度の指標は検知技術に対して一般的に使われている Precision/Recall を用いた。Precision は交通異常と判定したものに対する正解率 (= 正解した個数 / 交通異常と判定した個数)、Recall は発生した交通異常の中でどれだけ検知できたかという検知率 (= 正解した個数 / 検知対象である交通異常の個数) を表す。検知結果が上記誤差範囲内かどうかによって Precision/Recall を算出する。

また、検知時刻については、管制センターの入力であるイベントデータと比較を行った。イベントデータは、通報やパトロールカーによる発見に基づいており、図 6 のように実際に何らかの交通異常により交通状態が変化している時刻よりも遅れて入力されていることがある。これと比べて早期検知できるかどうかを検証した。

5.3 検証結果

検知位置に関する精度結果を表 1 に示す。Precision が 94.3%、Recall が 89.2% とともに高水準の結果が得られた。

検知例を図 7 (a)～(c) に示す。(a) の正解例では、イベントデ

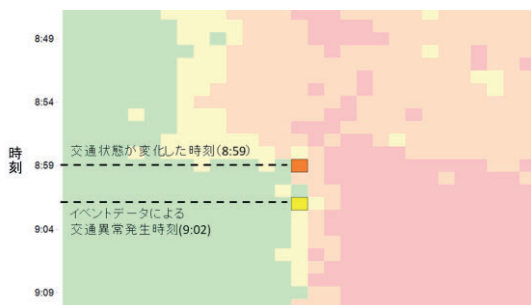


図 6 イベントデータの時刻遅れ

表 1 異常検知精度評価

		真の結果		
		交通異常	非交通異常	Precision
予測結果	交通異常	33	2	94.3%
	非検知	4	—	—
	Recall	89.2%	—	—

ータによる交通異常発生箇所の $+100\text{m}$ を検知しており、時間に関しても手入力よりも 1 分早く検知することができた。(b) は交通異常に関する事象が発生しているが非検知であった 4 件の内の一つを示している。このデータは、図では交通異常発生地点は一か所になっているが、同じ KP、同じ時刻において故障車と路上障害物の事象がイベントデータに含まれている。複数の交通異常が重なり、急加速ができなく、渋滞速度回復距離が長くなったため検知できなかったと考えられる。(c) はイベントデータが存在しない位置に対して交通異常判定をした例である。検知地点付近の速度は 40km/h から 55km/h へ 100m ほどの距離で加速しており、交通異常と似た傾向が表れている。また、同じ KP で他の時間帯を見ても、検知付近の交通状態が特異である点から、交通異常はあったがイベントデータに記録されていない可能性も考えられる。

検知時刻に関しては、イベントデータの入力時刻と異常検知時刻の時間差を評価した。両者の時間差をボックスプロットで表したものを図 8 に示す。縦軸は時間差を表しており、マイナスが ETC2.0 プローブデータを利用した方式による早期検知を意味している。結果から、中央値が 0 となり、半数が早期検知可能であった。一方で、大幅に検知遅れとなる場合もいくつか存在した。これは、ETC2.0 プローブデータに異常値が含まれていたため、渋滞速度回復距離が正しく算出できないことに起因している。今後、データに含まれる異常値のクレンジングや手法の頑健性も重要な課題である。

5.4 まとめ

高速道路内の異常を早期発見したいという社会課題と従来技術での高精度な位置特定が困難であるという課題に対して、ETC2.0 プローブデータを利用して渋滞先頭における渋滞速度回復距離により交通異常を検知するアルゴリズムを開発した。検知位置による精度評価では良好な結果を得ており、本方式が有効であることを示した。

今後、自然渋滞の原因を含めたモデル学習・検知手法の開発や、渋滞を伴わない交通異常を検知する方式の開発に取り組む予定である。

謝辞

本研究において、NEXCO 西日本様より ETC2.0 プローブデータおよびイベントデータを提供いただいた。ここに感謝の意を表す。

参考文献

- [内閣府 18] 内閣府: 平成 30 年交通安全白書,
http://www8.cao.go.jp/koutu/taisaku/h30kou_haku/zenbun/genkyo/h1/h1b1s1_2.html.

[国土交通省 17] 国土交通省: 高速道路会社の落下物処理件数(平成 29 年度),
http://www.mlit.go.jp/road/sisaku/ijikanri/pdf/h29rakkabutu_nexco.pdf.

[NEXCO 西日本] NEXCO 西日本: 渋滞原因解説, https://www.w-nexco.co.jp/forecast/trafficjam_comment/.

[Chung 98] Chung, E., and Kuwahara M.: Comparative study of freeway incident detection algorithms, ARRB TRANSPORT RESEARCH LTD CONFERENCE, 19TH, 1998.

[Jeong 11] Jeong, Y.S., et al.: A wavelet-based freeway incident detection algorithm with adapting threshold parameters, Transportation Research Part C: Emerging Technologies 19.1, 2011.

[成岡 14] 成岡尚哉, et al.: ノンパラメトリック手法による車両感知器データからの突発的ボトルネック検出法, 交通工学論文集 1.1, 2014.

[瀬戸下 18] 瀬戸下伸介, et al.: ETC2.0 プローブ情報の収集と活用, 国総研レポート 2018, 2018.

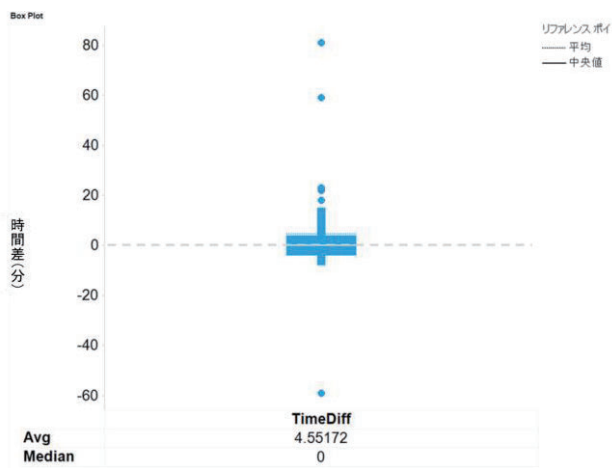
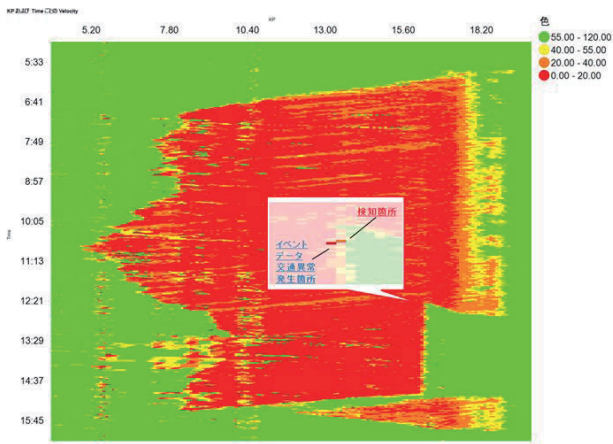
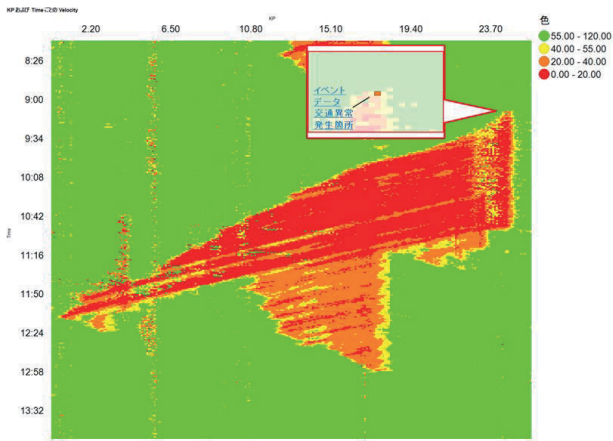


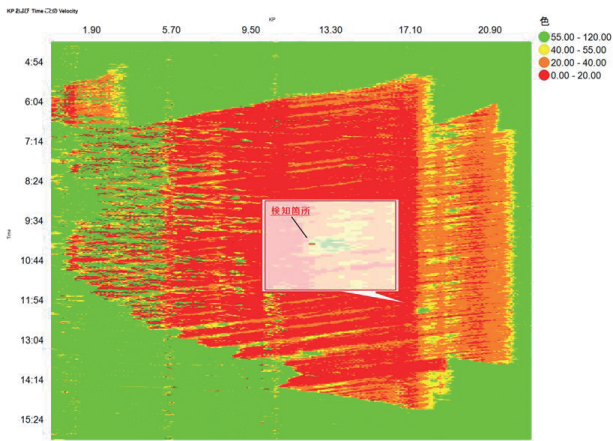
図 8 イベントデータと検知時刻の時間差



(a) 正解例
(中国自動車道下り 2018 年 8 月 10 日)



(b) 検知漏れ例
(中国自動車道下り 2018 年 3 月 16 日)



(c) 誤検知例
(中国自動車道下り 2018 年 8 月 11 日)

図 7 検知結果

高効率なドローンハイウェイ網の設計手法

Design method for high efficiency drone highway network

浜中 雅俊^{*1}

Masatoshi Hamanaka

^{*1} 理化学研究所
RIKEN

Abstract: This paper describes the design method for drone highway network to eliminate the risk of conflict between Drones and make it fly with high efficiency. Many flight path designing methods have been proposed, however a designing method for drone highway network that raises the overall flight efficiency has not been proposed. We optimize each path using ant colony optimization and optimize the position of the terminal connecting the paths using particle swarm optimization. Experimental results show that the proposed method improves flight efficiency.

1. はじめに

マルチロータ型のドローンは狭いスペースを飛行・着陸できる特長がある。我々はマルチロータ型のドローンを山岳地帯への物資の供給に利用することを期待している。しかし、ドローンで山岳地帯へ物資を供給するためには以下の3つの問題がある。

● GPS ロスト

従来のドローンでは GPS(グローバル・ポジショニング・システム)により位置検出が行われていたが、ビルの谷間や山岳地帯では、衛星からの GPS 信号の遮蔽や反射が生じ、正確な位置が求まらない場合があった。また、GPS 信号は微弱で、妨害で受信困難となる場合があった[Kerns 2014, Urbina 2016]。

我々は、衛星 3 次元地形図を用いて、ディープラーニングに基づき[LeCun 2015, Amari 2018], GPS 信号が受信できない場合でもドローンの飛行エリアを推定することを可能とする手法を提案し、群馬県昭和村で行った実証実験では 98.4%の精度でエリア推定が可能なことを確認した [Hamanaka 2018]。

● 衝突リスクの増大

上空を多くのドローンが飛行するようになると、衝突の危険性が増加し、高速で飛行させることが困難になることが予想される。たとえば 2 台のドローンが空中で交差する場合には、道路交通法と同じように交差する地点に左側から進入してくるドローンが優先され、右側から進入してくるドローンはホバリングしながら待つ必要があるが、ホバリング中もドローンはエネルギーを消費し続けている。

そこで我々は、ドローンハイウェイ網の提案を行ってきた [浜中 2016]。

● 低いエネルギー効率

マルチロータ型ドローンは、ローター回転により自重を支える必要がありエネルギー効率が悪い。特に山岳地帯を飛行する場合には、山を越えるために機体を大きく上昇させる必要があるが、上昇には非常に大きなエネルギーを必要とする。

上記のドローンハイウェイ網においては、蟻コロニー最適化法(ACO)および粒子群最適化法(PSO)を組み合わせ、エネルギー効率の高い経路となるよう最適設計手法を提案していた。しかし、最適化には大きな計算コストがかかるために、十分に最適化された経路を作成することはできなかった。

本稿では、ACO の計算を GPGPU 上で実現することで大規模なシミュレーションを実現した結果について報告する。以下、2

節では関連研究について検討し、3 節では提案手法の概要について述べる、そして 4 章では ACO 5 章では PSO について説明し、6 章で実験結果について述べ、7 章でまとめる。

2. 関連研究

ドローンの飛行経路の作成法は様々提案されている。Li らは ACO を用いて 2 地点間の経路の最適化を行うことを試みた [Li 2018]。ACO は収束に時間がかかることから、Bhatt らは高速に収束する Sniff-Dog アルゴリズムを提案した [Ashish 2018]。Shi らは、飛行経路を作成する問題を制約充足問題として定式化し、動的グラフ上で衝突のない最短経路を求めた [Shi 2018]。これらの研究は、出発点から到着点までの 1 つの経路を最適化しようとするものである [Li 2008; Bhatt 2018; Shi 2018]。

一方、Balachandran らの研究[Balachandran 2018]や Galea らの研究 [Galea 2018]は、複数のドローンを管制することを前提としている点で我々の考え方に近い。Balachandran らは、交差点に接近する無数のドローンの中から選定された 1 機のドローンが周辺のすべてのドローンの運行スケジュールを計算し、そのスケジュールの情報をすべてのドローンが同期する手法を提案した。一方、Garea らはボロノイ線図を用いて経路網を作成する方法を提案した。しかしこれらの研究では、経路網全体を効率化するというアイデアは提案されていない。

我々は、ACO と PSO を組み合わせたアルゴリズムにより経路網全体の飛行効率を向上させることを目指す。

3. ドローンハイウェイ網の最適化

本節では、ドローンを高速に高エネルギー効率でより遠くまで飛行させることを可能とするドローンハイウェイ網を構築するための手法の概要について述べる。

図 1 は京都市付近の地図で、上から下に伸びる濃い色(緑)の帯は滋賀県と京都府の県境にある山である。このときドローンハイウェイ網として図 1 のような格子状のハイウェイ網を採用した場合、様々な方向に飛行することが可能な点は優れているが、経路が山の山頂付近を通っているため山を越えるために大きなエネルギーを必要とするため効率が悪くなるという問題が生じる。

ドローンが高速・高エネルギー効率で飛行するためには、1) 距離が短いこと、2) 高度変化が少ないこと、3) 直進性が高いことが重要である。これらを同時に成立させるような経路を探索する手法として本研究では ACO と PSO という 2 つの手法を導入する。

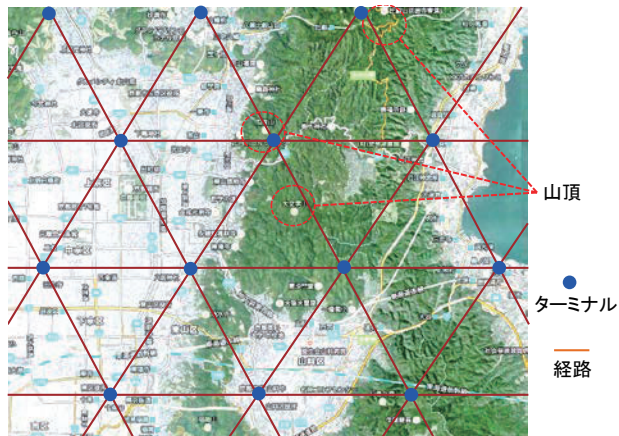


図 1: ドローンハイウェイ網の例

3.1 ACO (蟻コロニー最適化)

ACO を用いて、選択した 1 つのターミナルに接続されている各経路の最適化を行う。ACO は、フェロモンを使って蟻が効率的に巣から餌場まで移動しているしきみを確率モデルで表したものである。蟻は、巣であるコロニーから餌を取りに出かけるが、餌を見つけるとその一部を持ち帰るが、その際にフェロモンを出しながら帰る。そして、餌を探しに行く蟻はフェロモンを見つけると、その先にある餌を求めてその経路を移動し、帰りにフェロモンを強化していく。フェロモンは、時間と共に蒸発するため、より適した経路が選択される。

3.2 PSO (粒子群最適化)

PSO を用いて、選択したターミナルの位置の最適化を行う。PSO は、魚群において一匹が良い経路を見つけると、残りの群れがそれに倣うしきみをモデル化したものである。多次元空間において、最適解を求めようとするとき、位置と速度を持つ粒子群が動き回る。このとき、最も良い値となっている粒子の位置が全体に通知され、また、ローカルなベストな位置にある粒子が近傍に通知される。

3.3 ドローンハイウェイ網最適化アルゴリズム

ACO と PSO を組み合わせた手法により飛行経路網を最適化する。アルゴリズムは以下の通りである(図 2)。

- ドローンハイウェイ網の交点(ターミナル)を 1 つランダムで選ぶ。そのターミナルをコロニーとし、周囲のターミナルに餌を置く(図 2a)。
- ACO により各コロニーから餌までの最適経路を求め、それらの平均移動コストを算出する。
- 様々な位置にコロニーを置き PSO を用いて、最も平均移動コストが最少となるコロニー位置を求める。
- a から c までを 1 ステップとして、ステップを繰り返すことでドローンハイウェイ網全体を最適化する。

4. ACO (蟻コロニー最適化)

ACO は、巡回セールスマン問題の最適解の近似値を効率的に求める方法として知られている。巡回セールスマン問題は、すべての都市を 1 回ずつ巡回し、出発した都市に戻ってくる経路のうち、巡回の総コスト(時間や距離)が最少のものを求める組み合わせ最適問題であり、都市数の増加により計算量が急速に増加する NP 困難問題である。

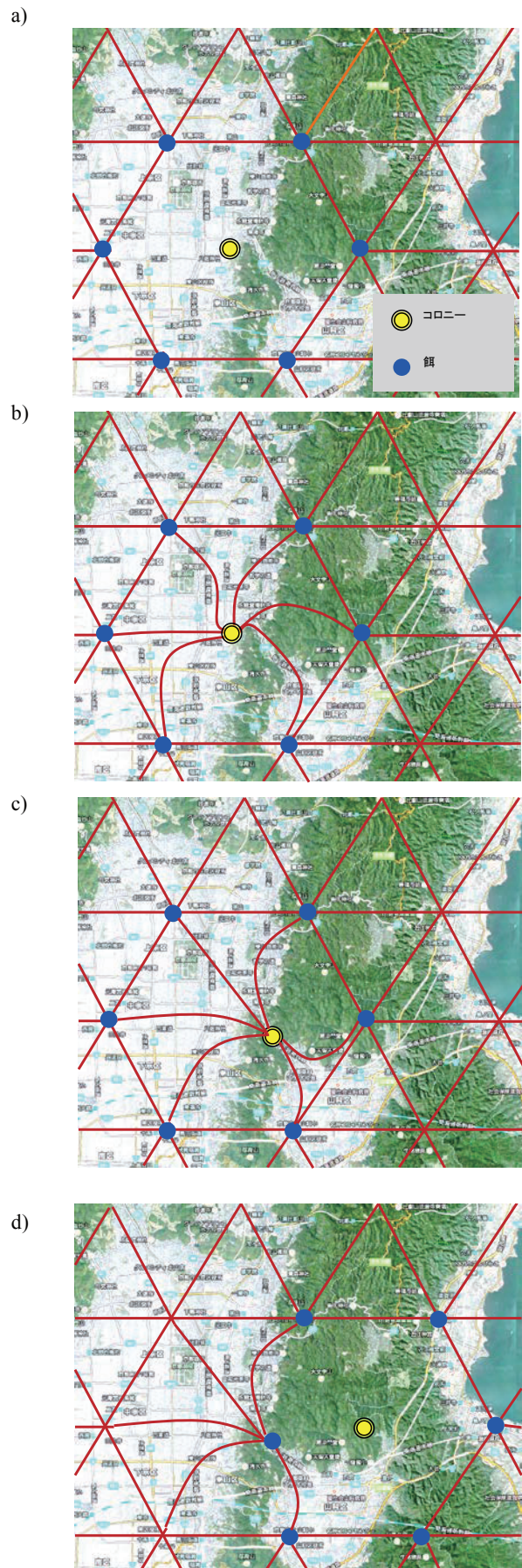


図 2: ドローンハイウェイ網最適化アルゴリズム

陸上において、スタート地点とゴール地点が決まっており、その間の経路を求める問題は、多くの場合 DP マッチングが使用できるため Viterbi アルゴリズムにより効率的に最適経路が探索できる。しかし、マルチロータ型ドローンの三次元的な経路の場合、飛行距離や時間だけでなく高度の変化やカーブ(曲率)によっても電費が変わってくるため、DP マッチングで解くことが難しい問題である。そこで、ACO を用いてこの問題を解く。

ACO において蟻が、上下左右前後の 6 方向の移動のうちどちらに移動するかは、6 方向のフェロモンの値を用いて ε -greedy 法で決定する。

曲率は、現在蟻がいる地点と、5 ステップ前に蟻がいた地点、5 ステップ後に蟻がいる地点で作られる三角形の外接円の半径の逆数である。

すべての蟻がゴールに到達するか、あるいは、探索打ち切りステップ数を超えると、蟻が通った経路にフェロモンを足し合わせてフェロモンの更新を行う。蟻による探索時間を短縮するため、以下を行った。

● フェロモン初期値の設定

フェロモンの初期値は、スタート地点から垂直上昇し、最高地点に到達したら水平移動してゴールまでのユークリッド距離が最少となるような経路に設定した。このように設定するとフェロモンの初期値をランダムとしたのとは比べて極めて早く収束する。最終的な収束結果はほぼ同じとなることを確認している。

● 方向別のフェロモン

蟻がフェロモンを頼りにスタート地点に逆流する現象を防ぐために上下左右前後の 6 方向それぞれ方向別にフェロモンを記録し反映するようにした。

● 探索打ち切り

壁が近くにあると蟻の動きに制限が加わるため、壁に沿ってフェロモン濃度が高くなることがある。ゴールから非常に離れた場所を探索する蟻がいると探索に時間がかかるため、スタートからゴールまでの最短距離の二乗までで探索を打ち切ることとした。

● ゴール上空での探索打ち切り

ゴール付近には到達するが、なかなかゴールに至らない蟻が増えると、ゴール周辺のフェロモンが増加し、ますますゴール付近で停滞する蟻が増加するため、ゴールの上空に到達した蟻は自動的に降下するようにした。

● GPGPU での並列処理

ACO は、蟻の数が少ないと収束までの試行回数が増大する。逆に蟻の数を増やすと、少ない試行回数で収束するが、一回あたりの試行時間が増大する。そこで、ACO の計算を蟻ごとに並列で GPGPU 上で行う。すると、フェロモンの更新の計算以外は並列で行うことができるため高速化される。

5. PSO (粒子群最適化)

PSO は魚や鳥の群れなど、動物の群れの行動を参考にして作られた確率的最適化手法である。全探索が困難あるいは長い計算時間がかかる問題において、解の候補である粒子を解空間に多数配置し、各粒子どうしが相互作用により高精度な近似解を発見する手法である。具体的には、各粒子がそれまでに各自が発見した最も良い解候補の位置(ローカルベスト)を記憶し、かつ各粒子が持つローカルベストの中で最も良い候補の位置をグローバルベストとして全体が共有しており、粒子がその方向に移動することで最適化を行う。

PSO は進化的計算手法の 1 つに含まれるが、GA などと比べて収束性が良い。しかし、解候補の評価を繰り返すため、候補の評価に時間のかかる処理の場合には高速化が難しい。

PSO の評価値は、コロニーから周囲の 6 つの餌までの ACO のコストの平均値である。ACO は収束に時間がかかるために、さらなる高速化が必要である。高速化のために、以下を行った。

● 評価値の記録

一度ある粒子が評価値を求めた場所を別の粒子が再度評価値を求める場合がある。コロニーを置く箇所は有限であるため、コロニー位置の探索をはじめから最初に求めた評価値を記録し、同じ場所にコロニーが再度来た場合には記録した評価値を用いることで高速化する。

● 並列計算

コロニーからは、周囲の 6 つの餌に対して経路を ACO で求める。このときこの 6 つの計算を並列で行うことで高速化する。

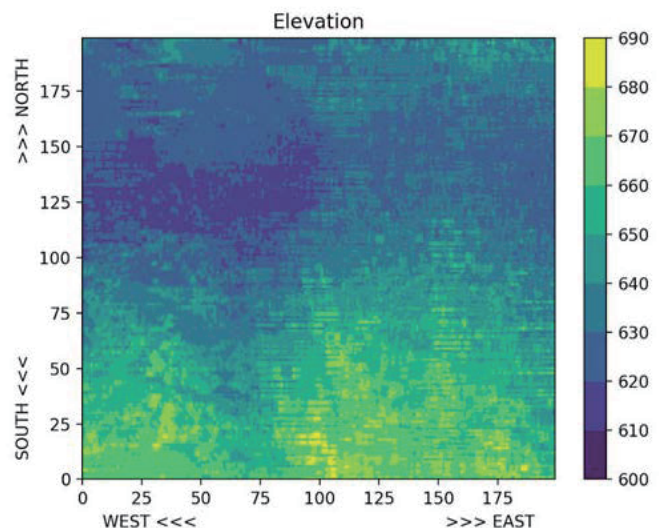
6. 実験結果

ACO および PSO を組み合わせたドローンハイウェイ網構築手法を用いてシミュレーション実験を行う。

6.1 実験エリア

本節では、以前実験を実施した群馬県昭和村の北緯 36.6217 から 36.61990 まで、東経 139.1090 から 139.111240 の 200 メートル四方のエリア内で、スタート地点とゴール地点をランダムで選択しその間の適切な経路を求めることとする。図 3 は、200 メートル四方のエリアに LiDAR(SICK 社 LD-MRS)を搭載したドローン(DJI 社 Matrice600)を 10 回飛行させ、その結果を平均して作った 3 次元地形図で、解像度は 1 メートルである。最も深いところと高いところで 90 メートル程度の高度差があり、高度を考慮して最適な経路できているか検討するのにふさわしいエリアと考えられる。この 200 メートル四方の範囲を蟻は餌があるゴール地点を探索して進む。蟻は 1 ステップで上下左右前後に 1 メートル進むものとする。

図 3: 3 次元地形図



6.2 ACO の実験結果

飛行経路の距離を x メートル、高度の差を y メートル、経路の平均曲率を z としたとき、最小化すべきコスト関数として次式を定義する。 a, b, c は定数でそれぞれ、0.08, 0.8, 0.01 とした。

$$f = ax + by + cz \quad (1)$$

ACO では、多くのハイパーパラメータがあるが、本節は以下のように設定した。

- 蟻の個体数 MAX_ANTS 8192
- フェロモン更新の定数 Q 3

- フェロモンの依存度に関する定数 ALPHA 1.4
- フェロモン蒸発の定数 RHO 0.5
- フェロモン初期値 IP 0
- 行動のランダム性 EPSILON 0.05
- 最大探索数 MAX_STEP 30,000
- 施行回数 MAX_TIME 100
- 距離に対する重み a 0.08
- 高度差に対する重み b 0.8
- 曲率に対する重み c 0.01

蟻の個体数 MAX_ANTS は、多いほど並列度が高まり高速に収束するようになるが、多すぎるとメモリ不足などにより不安定になる場合がある。8192 は、GPU を用いた場合において我々の環境で安定して動作する上限である。フェロモン更新の定数 Q は、スタートからゴールまでの距離に応じて変化させる必要があるが、今回の場合 3 が適切であった。行動選択のランダム性 EPSILON は、0.05 のとき 5% の確率でランダムに行動する。

探索の結果、最少コストは減少しており、最適化が進んでいることが確認できた(図 4)。図 5 は、白い x がスタート地点、赤い x がゴール地点として ACO を行った結果である。ヒートマップはフェロモンの量、白線が最短経路である。試行 10 回目(図 5a)では、まだ蛇行している部分があり最適化が十分進んでいないが、100 回目(図 5b)では最適化が進んでいる。

計算速度を比較したところ、100 回の試行で CPU(Mac mini Corei7-8700B 3.2G/4.6GHz 6 コア 12 スレッド)で 1 時間 21 分 30 秒、GPGPU(Nvidia GTX1080)で 82.3 秒であった。

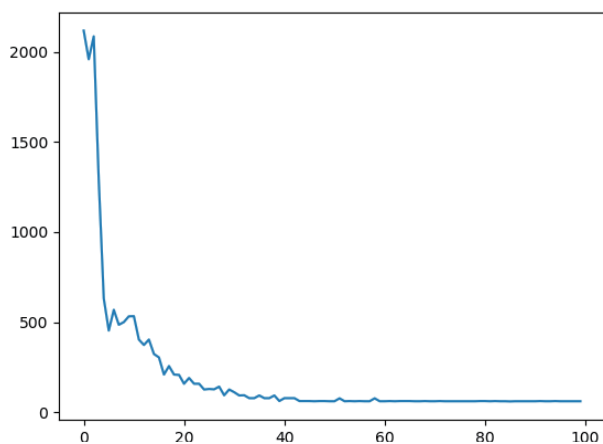


図 4: ACO による最小コストの変化

6.3 PSO の実験結果

図 6 は PSO の結果求められたドローンハイウェイ網の一部である。白の x がスタート地点、赤の x がゴール地点、白い線が最適化された経路である。スタート地点とゴール地点が同じくらいの高さの場合、高低差を避けるように最適化が行われている。実験エリアは窪地のためスタート地点とゴール地点に高低差がある場合は、最初に高度を取って(初期フェロモンに沿った)直線的な経路となった。

7. おわりに

高効率で飛行可能なドローンハイウェイ網の設計手法を提案した。今後実機を飛行させ飛行効率の検証を行う。

謝辞

本研究の一部は、JSPS 科研費 17K19972 の助成を受けたものです。

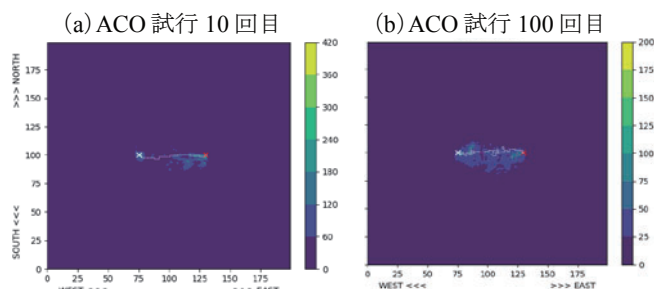


図 5: ACO による最適化

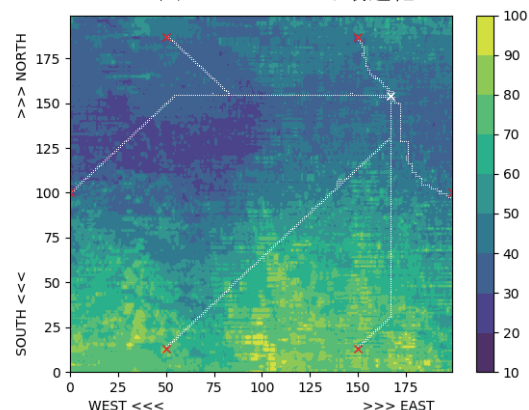


図 6: 最適化されたドローンハイウェイ網

参考文献

- [Kerns 2014] A. J. Kerns et al, “Unmanned Aircraft Capture and Control Via GPS Spoofing,” Journal of Field Robotics, Vol. 31, Issue 4, pp. 617–636, 2014.
- [Urbina 2016] D. I. Urbina et al, “Limiting the Impact of Stealthy Attacks on Industrial Control Systems,” Proc. of the 2016 ACM SIGSAC, pp. 1092–1105, 2016.
- [LeCun 2015] Y. LeCun et al, “Deep Learning,” Nature, Vol. 521, Issue 7553, pp. 436–444, 2015.
- [Amari 2018] S. Amari et al, “Dynamics of Learning in MLP: Natural Gradient and Singularity Revisited,” Neural Comput, Vol. 30, Issue 1, pp. 1–33, 2018.
- [Hamanaka 2018] M. Hamanaka, “Deep Learning-based Area Estimation for Unmanned Aircraft Systems using 3D Map,” Proceedings of 2018 International Conference on Unmanned Aircraft Systems (ICUAS2018), pp. 416–423, 2018.
- [Hamanaka 2016] M. Hamanaka and H. Shiomi, “Optimum Design Method for Drone Flight Route Network, Proceedings of the 78th national convention of IPSJ, 1B-04, 2016 (in Japanese).
- [Li 2018] Z. Li et al, “Preliminary Research on Remote Planning Methods for Paths of Unmanned Aerial Vehicles in the Safe Aerial Corridor,” Proc. of ICUAS2018, pp. 1288–1294, 2018.
- [Bhatt 2018] A. Bhatt et al, “Path Planning for UAV using Sniff-Dog-Algorithm,” Proc. of ICUAS2018, pp. 692–701, 2018.
- [Shi 2018] Z. Shi and W. K. Ng, “A Collision-Free Path Planning Algorithm for Unmanned Aerial Vehicle Delivery,” Proc. of ICUAS2018, pp. 358–362, 2018.
- [Balachandran 2018] S. Balachandran et al, “Distributed Consensus to Enable Merging and Spacing of UAS in an Urban Environment,” Proc. of ICUAS2018, pp. 670–675, 2018.
- [Galea 2018] M. Galea et al, “Design of a Multi-Layer UAV Path Planner for Cluttered Environments,” Proc. of ICUAS2018, pp. 914–923, 2018.

自動車用自動変速機の変速制御時に現れる時系列データの分類

Classification for time-sequence data appeared in shift control of automobile automatic transmission

森川祐介^{*1} 石原靖弘^{*1} 井手貴範^{*1} 森山英二^{*1} 秋田拓^{*1} 田部井靖生^{*2}
Yusuke Morikawa Yasuhiro Ishihara Takanori Ide Eiji Moriyama Taku Akita Yasuo Tabei
宇津呂武仁^{*3} 中川裕志^{*2}
Takehito Utsuro Hiroshi Nakagawa

^{*1}アイシン・エイ・ダブリュ(株) ^{*2}理化学研究所 革新知能統合研究センター
Aisin AW Co., Ltd. RIKEN, Center for Advanced Intelligence Project

^{*3}筑波大学システム情報系
Fclty. Eng. Inf. & Sys, Univ. of Tsukuba

The automatic transmission is a transmission that automatically changes gear according to an optimal gear step decided based on driver's accelerator/brake operation, vehicle speed, etc. Since the quality of a changing gear of automatic transmission, we call shift feeling, has a great influence on riding comfort of the automobile, the parameters of control program are adjusted to improve the shift feeling at the development site. We call this adjustment work calibration. In the process of calibration, there is a process of identifying occurring phenomena in the time-sequence data and selecting the cause and countermeasures. Because this task takes large man-hours, it is very efficient if this task becomes automatic. Therefore, at the first step, in order to automate identification of the occurring phenomena, we built classifier of the time-sequence data using SVM. The experimental result suggests that the proposed identification system is promising directions for actually used identification system.

1. はじめに

オートマチック車における自動変速機はユーザのアクセル操作やブレーキ操作、自動車の車速などの走行状態に応じて最適なギヤ段を決定し、自動的に変速を行う変速機である。ユーザの操作通りに変速するマニュアル変速機では変速時の空走感やショックは許容されるが、自動変速機はユーザの操作とは異なるタイミングで変速するために変速時のわずかな空走感やショックでも不快と感じる傾向がある。自動変速機の変速の質(変速フィーリングという)は自動車の乗り心地に大きな影響を与えており、開発現場では変速の質を高めるために自動変速の制御プログラムのパラメータ調整に大きな工数をかけている。技術者が判断した変速フィーリングが良くなるようにパラメータを調整することを適合と呼ぶ。

適合では以下の Step 1.~Step 4. を実施している。

Step 1. 自動変速機を自動車や評価機に搭載して動作させ、動作時の制御・物理データを時系列データとして取得する。

Step 2. 変速フィーリングに問題がある場合、時系列データを確認し、生じている問題の種類(=現象)を特定する。

Step 3. 特定した現象に対して想定される要因や対策の候補をいくつか抽出する。

Step 4. Step 3. の候補の中から、発生した現象の要因と対策を特定して制御プログラムのパラメータを調整する。

この一連の工程のうち Step 4. の要因と対策の特定には経験が必要であり、誤った要因と対策を選んでしまうと手戻りが発生する。そのため自動的に熟練者の判断に近い要因と対策を提示できれば、手戻りを未然に防止できる可能性が高くなり工数低減が見込める。また同時に、経験の少ない技術者が熟練技術

者の知見を得ることができるため、技術者の育成に繋がることも期待できる。

要因と対策の提示のためにはまず現象の特定が必要であるため、本研究では時系列データで発生した現象の特定を現象の分類問題として考え自動分類を行う。過去の研究 [1][2] では特徴量ベースで時系列データの分類を行っているが、本研究ではサンプリングした時系列データそのものを入力としてサポートベクターマシン (SVM) および K 近傍法 (KNN) を用いて分類する方法を検討した。

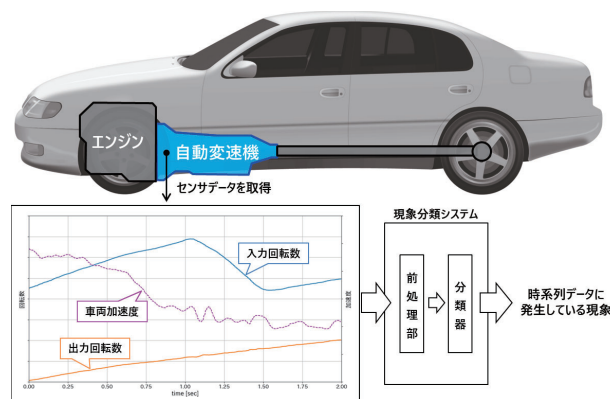


図 1: 自動変速機からの取得データと本研究のイメージ

2. 扱う時系列データ

適合時に取得した時系列データのうち、変速フィーリングに問題があるものを対象とする。時系列データに含まれる信号のうち回転数、加速度など 7 種類を用いた。時系列データは発

連絡先: 森川祐介, アイシン・エイ・ダブリュ(株), 〒444-1192
愛知県安城市藤井町高根 10, i37710_morikawa@aisin-
aw.co.jp

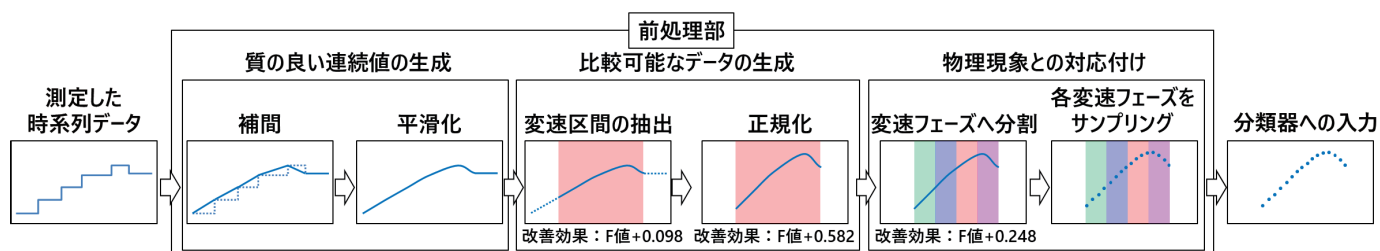


図 2: 前処理の流れ

生じた現象のラベルを持ち、1つのデータが複数のラベルを持つこともある。本来現象ラベルは複数の技術者により付けられているが、人によるばらつきを抑えるために1人の熟練技術者が特別に再度ラベル付けを行ったデータを用いた。

データベースに保存されている時系列データは長さ2秒、値の保存間隔は5ミリ秒、データ点数は400点である。検証のため、アクセルを踏み込みながら高速ギヤ段に変速するアップシフトのみを対象とする。時系列データは約300個あり、全24現象あるうちの数が多い順に7現象を対象とする。7現象に該当するデータ数は平均で44.6個、最大で86個、最小で16個である。

現象には、変速時にクラッチが勢いよく係合し、ショックが生じる現象などが含まれる。

3. 現象分類システムの構成

構築した現象分類システムの構成について説明する。現象分類システムは前処理部と分類器に分かれており、実データに適用するためには特に前処理部が重要となる。

3.1 前処理部

図2に前処理の流れを示した。時系列データの7信号それぞれに対し前処理を行う。各前処理を順に説明する。

Step1. 時系列データの補間を行う。時系列データは5ミリ秒刻みで値が保存されているが、自動車の種類や開発フェーズ、取得環境により値の更新間隔が異なり、階段状になっていることがある。例えばあるデータでは5ミリ秒ごとに値が更新されるが、別のデータでは50ミリ秒ごとに値が更新される。そのため値が更新された点のみを用いて線形補間でデータを補間し、すべての時系列データを同じ時刻に比較出来るようにする。

Step2. 時系列データの平滑化を行う。目的はノイズ除去および線形補間で生じた不連続部の平滑化を行い、質の良い連続値を生成することである。移動平均期間は15期間とした。

Step3. 2秒間の時系列データから実際に変速している区間を抽出する。変速は0.5～1.5秒程度で終わり現象は変速中に発生するため、変速区間以外は現象の分類には必要ない情報であり、除外することで分類性能の向上が期待できる。

Step4. データのスケールを揃えるため、抽出した区間を正規化する。正規化手法には z 正規化を用いた。時系列データ内の信号間で値のオーダーが数倍異なり、同じ信号でもデータ間で10倍程度値のオーダーが異なるため、正規化は必須である。

Step5. 抽出した区間を自動変速機の内部状態に応じた変速フェーズに分割する。変速フェーズは4つあり、順にクラッチをストロークさせるフェーズ、回転変化が起きるまでのフェーズ、回転が変化しているフェーズ、回転変化が終わり次のギヤ

段へ移行するフェーズである。現象は自動変速機内の物理現象が原因で発生するため変速フェーズと関わりが深く、特定の変速フェーズで特徴が出ることが多い。変速のたびにそれぞれの変速フェーズの長さが異なるため、変速フェーズに分割することで分類器の入力ベクトル内で現象の特徴が生じる位置を揃えることができる。実際、変速フェーズに分割したことで分類性能は大きく改善した。

Step6. 変速フェーズごとに時系列データをサンプリングし、時間長の異なるデータを同じ長さの入力ベクトルに変換する。サンプリングは、各変速フェーズを N 個の小区間に分け、小区間内の最大値・最小値を抽出して行った。最大値・最小値を抽出する理由は、現象の判定に重要なピークやボトムを見落とさないようにするためである。今回は $N=10$ とした。

以上の前処理により、時系列データを長さ560の入力ベクトルに変換した。4.1節でそれぞれの前処理の効果を確認した。

3.2 分類器

分類器にはSVMおよびKNNを用いた。SVMのカーネルは線形カーネルおよびRBFカーネルを検討した。分類は1対他分類で行ったため、データの正例・負例に不均衡が生じる。正例・負例それぞれでSVMの C パラメータに各クラスのデータ数の逆数に比例する重みを掛け、不均衡データが学習に与える影響を抑制した。SVMの学習および評価は3-fold交差検証で行った。正則化パラメータ $C=10^{-3} \sim 10^3$ 、RBFカーネルパラメータ $\gamma=10^{-3} \sim 10^3$ を候補とし、交差検証の学習データを用い、2-fold交差検証でF値が最もよくなるパラメータを選択した。KNNは3-fold交差検証で評価し、近傍数 k は $k=1, 3, 5, 7, 9$ を検討した。

4. 結果と考察

線形カーネルSVM (SVM(線形))、RBFカーネルSVM (SVM(RBF)) 及びKNNを用いて7現象の分類を実施した。

4.1 前処理の効果

SVM(線形)、SVM(RBF)、KNNの中で最も良好な性能を示したSVM(RBF)において、各前処理のF値の改善効果を確認した（各分類器の性能については後述する）。システム上必要な処理を除いた、変速区間抽出、正規化、変速フェーズ分割を除外した場合を検討した。改善効果は、すべての前処理を実施したmacro-F値(0.711)と前処理を除外したmacro-F値の差として定義する。

変速区間抽出を外して2秒間の波形全体で正規化を行った場合、macro-F値は0.613となった。変速区間抽出の効果は+0.098である。現象は変速中に発生するため変速区間以外は現象とは関係のない無駄なデータであり、取り除くことで性能が向上したのだと考えられる。

正規化を外した場合、macro-F 値は 0.129 となった。正規化の効果は +0.582 である。現象ごとに交差検証の各 fold の結果を見ると、すべて正例または負例と判定されており分類器として機能していない状態であった。3.1 節で述べた通り信号間、データ間で値のオーダーが数桁異なるため、今回対象とする現象分類においては正規化は必須である。

変速フェーズ分割を外した場合、macro-F 値は 0.463 となった。変速フェーズ分割の効果は +0.248 である。前述した通り、現象は自動変速機内部の物理現象と密接な関係があるため、変速フェーズに分割することで入力ベクトル内で現象の特徴の位置が揃うようになる。そのため現象の分類に必要な入力が明確になり、性能向上に繋がったと考えられる。

4.2 KNN の結果

図 3 に k を変化させたときの F 値の推移を表示した。macro F 値は $k=1$ がベスト性能、個別で見ても現象 3, 4 は $k=3$ がベスト性能、その他現象は $k=1$ がベスト性能となった。小さい k がベスト性能となる理由は、データが入り組んでおり境界面が引きづらいこと、正例の数が少ないことの 2 点だと考えられる。以下で詳細に確認する。

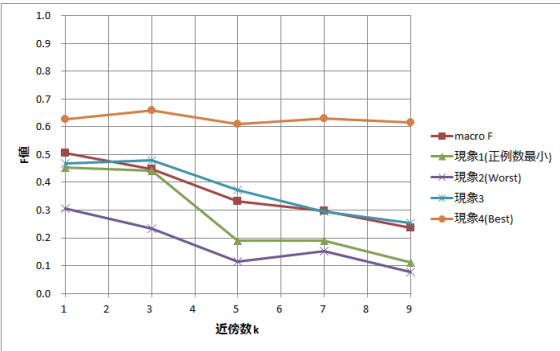


図 3: k を変更した際の KNN の分類性能

4.2.1 KNN ベスト性能

ベスト性能の現象 4 は正例数が 86 個と最も多く、また最後の変速フェーズでピークが立つため特徴が分かりやすい。さらに現象 4 と似た特徴を持つ現象が存在せず、後述する現象 2, 3 と比較して境界面が引きやすいと考えられる。しかし特徴は時系列データの一部にしか生じず入力ベクトルには現象の分類に不要な入力が含まれており、特徴空間上でデータは入り組んでいると考えられるため、大きな k では性能が出づらく $k=3$ が最大となったと考えられる。

4.2.2 KNN ワースト性能

ワースト性能の現象 2 は正例数 39 個でデータ数が少ない。特徴は 2,3 番目の変速フェーズにまたがって発生し、後述する現象 3 と特徴が良く似ており技術者が見ても判別が難しい。特徴が良く似た現象 2, 3 でデータが入り組んでおり境界面が引きづらく、大きな k では性能が出ないのだと考えられる。

4.2.3 中間の性能

現象 3 は正例数 84 個で 7 現象中 2 番目に多いため、現象 2 より判定に使えるデータが多く性能も高くなっている。しかし k を大きくすると性能が落ちる傾向は変わらないことから、現象 2 との間でデータは入り組んでいると考えられる。現象 1 は正例数 16 個であるものの、似た現象が存在せず最初の変速フェーズにピークが出て特徴が分かりやすい。そのため $k=1,3$ ではデータ数のわりに分類性能が高くなっているが、

正例判定に学習データの半数が必要となる $k=5$ 以上では急激に性能が落ちている。

以上のことから、データが入り組んでおり境界面が引きづらいこと、正例の数が少ないことの 2 点により $k=1,3$ でベスト性能となったと考えられる。

4.3 KNN と SVM の比較

表 1 に SVM と KNN の評価結果を示した。KNN の k は現象ごとに最も F 値が良くなるものを選んだ。表に載せていない現象も含め、すべての現象で SVM の F 値が KNN の F 値を上回った。入力ベクトルには現象の分類に不要な入力も含まれているため、すべての入力を等価に扱う KNN では重みの調整で現象に関係する入力に注目できる SVM に性能が及ばないのだと考えられる。SVM(線形)の重みを確認すると、確かに人間が現象の判別を使う付近の入力は大きな絶対値の重みとなり、それ以外は 0 に近い値となっていることが多かった。

表 1: 分類結果

項目		SVM (線形)	SVM (RBF)	KNN
全体 (macro)	F-score	0.683	0.711	0.509
	Precision	0.588	0.661	0.608
	Recall	0.815	0.769	0.438
	AUC	0.869	0.876	-
現象 5	F-score	0.476	0.525	0.419
	Precision	0.374	0.439	0.493
	Recall	0.667	0.667	0.389
	AUC	0.787	0.778	-
現象 6	F-score	0.844	0.857	0.651
	Precision	0.781	0.915	0.761
	Recall	0.926	0.810	0.574
	AUC	0.977	0.942	-

4.4 SVM(線形) と SVM(RBF) の比較

すべての現象で SVM(RBF) の F 値が SVM(線形) の F 値を上回った。KNN で $k=1,3$ がベスト性能となったことをふまえると、特徴空間内ではデータ同士が複雑に入り組んだ構造になっており、非線形な境界面を引ける RBF カーネルが良好な分類性能を示したと考えられる。

4.5 SVM(RBF) の結果

4.5.1 正例・負例を判別できる現象

SVM(RBF) のベスト性能である現象 6 について、図 4 に一つの fold のテストデータの識別境界面からの距離の分布 (左) と ROC 曲線 (右) を示す。距離の分布から、正例 (上側) と負例 (下側) の分布がほぼ完全に分離しており正確に判別できていることがわかる。False Negative データは境界面に近い位置で間違えており、性質の良い境界面を生成できている。

現象 6 では変速の終盤に入力回転数の変動という特徴が現れ、他に似た現象が存在しない。同様に似た現象が存在しない現象 1, 4, 7 についても F 値が 0.8 前後、AUC 0.9 前後で精度よく判別できていることから、これらの現象では特徴空間内で現象ごとにデータが比較的まとまって分布していると考えられる。現象 6 の Recall が 0.810 に留まったのは、データ数が 26 個しかなく学習データでテストデータの事例をカバーしきれないことが原因だと考えられるため、データ数を増やすことで精度の向上が期待できる。

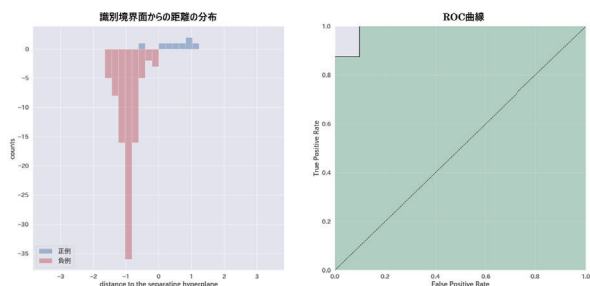


図 4: 現象 6 : 境界面からの距離の分布と ROC 曲線

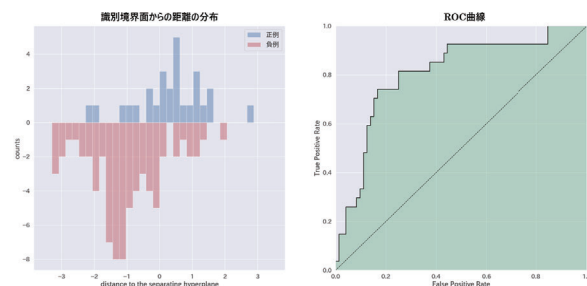


図 6: 現象 3 : 境界面からの距離の分布と ROC 曲線

4.5.2 正例・負例の判別が難しい現象

ワースト性能の現象 5 についても同様に、図 5 に識別境界面からの距離の分布 (左) と ROC 曲線 (右) を示した。正例・負例の分布が相手の分布の深い位置まで食い込んでおり、SVM が分類する条件を学習できなかった現象が存在していることがわかる。False Positive データを確認すると、距離 1 付近のデータの約半数は現象 2 または現象 3 であった。また、距離 -3 付近の False Negative データでも、近い順に 10 個のデータを確認すると半数はやはり現象 2 または現象 3 であった。

現象 5 は現象 2, 3 と同じように 2, 3 番目の変速フェーズで特徴が生じる。特徴も現象 2, 3 と良く似ているため、今回分類に用いた入力だけでは現象を説明する情報が足りず、特徴空間内で重なって分布したのだと考えられる。

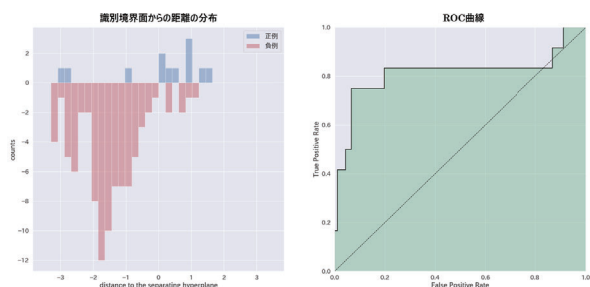


図 5: 現象 5 : 境界面からの距離の分布と ROC 曲線

図 6 に現象 3 の識別境界面からの距離の分布を示す。現象 5 同様、正例・負例の分布が重なっている。現象 3 の場合、False Positive となったデータの半数は現象 2 である。前に述べた通り現象 2, 3 は特徴が良く似ているため、現象を説明する情報が足りず特徴空間内でデータの分布が重なってしまったのだと考えられる。

現象 2 についても現象 3, 5 と同様の結果となった。良く似た現象を持つ現象 2, 3, 5 の分類性能を実用レベルまで上げるためには、特徴空間内で各現象の分布を切り分ける軸を与えられる新しい入力の検討が必要であると考えられる。

5. まとめと今後の課題

我々は、自動車用自動変速機の開発効率向上のため、変速フィーリングに問題がある時系列データに発生している現象の分類を行う手法を検討した。7 現象中 4 現象は現時点で実用レベルに近いところまで来ており、本手法で現象の分類は可能であるとする。しかし似た現象同士の分類精度は不十分である

ため、各現象を切り分けるための新しい入力の検討が必要である。

データ数と分類精度の関係について検討した結果を図 7 に示す。この図は 311 個の時系列データから無作為に N 個を抽出して RBF カーネル SVM で学習・評価する試行をデータ数 N ごとに 20 回行い、F 値の平均値を算出した結果である。データ数 N は $N = 63, 78, 105, 155, 206, 233, 248, 311$ とした。この図から、データ数の増加に伴い分類性能が向上する傾向が出ることがわかった。現象 1, 4, 6 については性能が飽和する傾向が出ているが、前述のとおりすでに十分な性能を得られているため問題はない。分類が難しい現象 2, 3, 5 はデータ数を増やすことで性能が向上しており、データ数を増やすことでも分類性能の改善が見込めることがわかったため、学習に利用できるデータを増やし分類器の改善を行っていく。

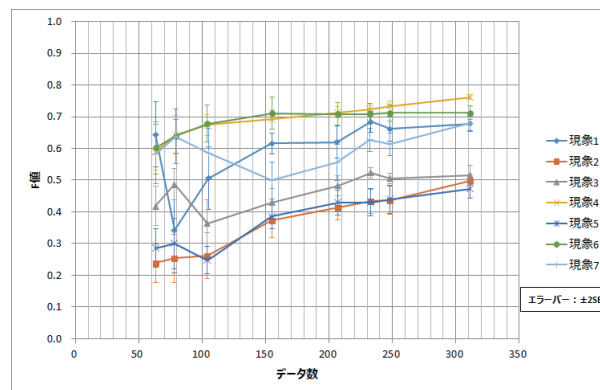


図 7: データ数と精度の関係

参考文献

- [1] B. D. Fulcher and N. S. Jones: Highly Comparative Feature-Based Time-Series Classification, IEEE Transactions on Knowledge and Data Engineering, Vol.26, No.12, 2014.
- [2] A. Kampouraki, G. Manis, and C. Nikou: Heartbeat Time Series Classification With Support Vector Machine, IEEE Transactions on Information Technology in Biomedicine, Vol.13, No.4, 2009.

Likelihood distribution of Pedestrian Trajectories rendered by Variational Autoencoder

Yasunori Yokojima^{*1}Tatsuhide Sakai^{*2}^{*1} Siemens K.K.^{*2} Great Wall Motor

We studied applicability of Variational Autoencoder (VAE) to capture stochastic nature of pedestrian moves in a public space without explicit labels. Movies for training the network were recorded in a public pedestrian street and an exhibition booth. These movies were converted to grayscale images representing observed pedestrian locations and occupied areas. VAE was trained on 90% of data and rest of data was kept for validation. The validation result showed satisfactory reconstruction performance of pedestrian distributions in video frames. We propose a novel method to render our expectation of finding a pedestrian in a crowd as 2-D images by utilizing the trained network. Images rendered by this method correspond to subjective images usually only captured in our mind.

1. Introduction

Building reliable traffic system requires a realistic description of pedestrian behaviors. For example, one of challenges in advanced safety system developments is to find rare but important corner cases, which are likely to be traced back to a subtle but non-negligible pedestrian behavior.

Pedestrian behaviors in a public space are explained by logical relationships and stochastic factors. A logical relationship of pedestrian trajectories is studied for years and various types of models are proposed to analyze universal traffic phenomena. Those models include Social force model, Optimal velocity model, Car-following model, and Cellular Automaton (CA) based model.

The stochastic and statistic nature of pedestrian behaviors is likely to be influenced by individual experiences, subjective recognitions of regional trends, or expectations in mind. For example, there can be a trend to yield right side in a public space in a certain region when two pedestrians are passing each other. Although such stochasticity can be taken into account by assuming a probability distribution and parameters, is it possible to learn the probability distribution directly from pedestrian data?

In the meantime, we occasionally recognize certain patterns while we stare at pedestrians in a public crowd. This is a subjective process of building expectations to pedestrian moves. These expectations give intuitive motive for a pedestrian to decide behaviors in a crowd. To take this into account in a pedestrian modeling, is it possible to mimic this subjective process by unsupervised learning on pedestrian data?

For these backgrounds, our goal is twofold: firstly study applicability of unsupervised machine learning to capture stochastic aspect of pedestrian behaviors, and secondly render learned content as a representation of our expectations to find an individual in a crowd. In the following sections, we propose a method to learn pedestrian distribution in a public space using VAE [Kingma 2014].

2. Learning pedestrian trajectories

2.1 Data Preprocessing

Our focus is to train a network on pedestrian locations and occupied areas. For this purpose, training data is prepared in the following process.

Pedestrian movies are recorded at 30 FPS by a single eye camera mounted at a fixed location with a certain elevation angle to a floor of pedestrian traffic. Movies are recorded at two different locations by this setup: movies recorded at a public pedestrian street have approximately 90 second length (Street data, hereafter). Movies recorded at an exhibition booth have approximately 30 minutes length (Booth data, hereafter).

For each frame in these movies, we applied YOLO v3 [Redmon 2015] to detect pedestrians and selected middle points of lower bases of detected bounding boxes as approximated pedestrian positions. Pedestrian positions are transformed to positions in a 2-D rectangular area while keeping the aspect ratio same as the original area. The resulted view is corresponding to a perspective perpendicular to the floor. Based on these pedestrian positions we generate grayscale images with white circles represent areas occupied by pedestrians on a black rectangular background. These images are resized to 10% in size of the original images by using INTER_AREA algorithm in OpenCV. Pixels in an image take value from 0 to 255. These images are fed as inputs to train the VAE network (Figure 3).

2.2 VAE network

VAE network structure is shown in Figure 1. We assume the latent distribution to follow a Gaussian distribution [Doersch 2016]. A loss function is defined as the following form:

$$L(x) = -D_{KL}(q_{\phi}(z|x)||p(z)) + E_{q_{\phi}(z|x)}[\log p_{\theta}(x|z)] \quad (1)$$

, where first term represents a KL divergence regularization and second term represent a reconstruction error. As an explicit form of the reconstruction term, a mean-squared error function is chosen:

$$\text{MSE} = \frac{1}{n} \sum (x - x')^2 \quad (2)$$

, where x is input data and x' is reconstructed data, and the summation runs over all n data points in x .

Once the network is trained properly, the reconstruction cost quantifies a distance between training data and an input fed to the trained network. When the input is not alike the training dataset, the reconstruction cost gives a high value. Using this reconstruction cost, we define a likelihood of x as follows:

$$f(x) = 1 - \left[\frac{\text{MSE} - \min(\text{MSE})}{\max(\text{MSE}) - \min(\text{MSE})} \right] \quad (3)$$

, where $f(x)$ takes values from zero to one and higher value means a higher likelihood for a given x .

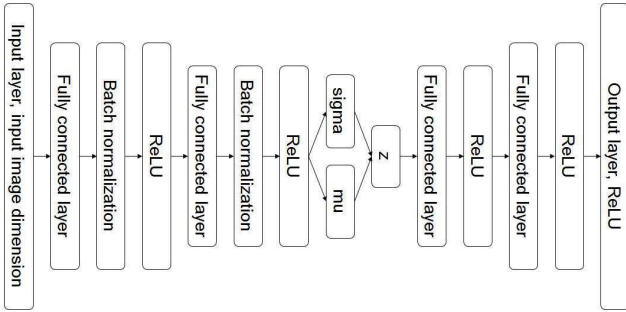


Figure 1 – **VAE network overview:** Input and Output layers have a same dimension as an input image. Two intermediate dimensions and latent dimension are 1000, 300 and 100, respectively. The rectified linear unit (ReLU) is used as the output activation.

2.3 Training results

VAE was successfully trained on Street data and Booth data (Figure 2) and demonstrated satisfactory reconstruction performance (Figure 3). However, a rate of training loss vs

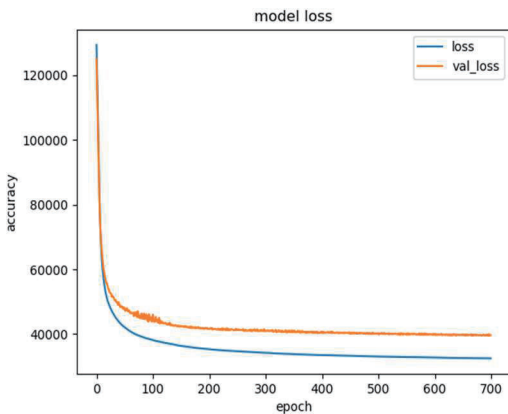


Figure 2 – **Example of loss change during a training:** Training loss (Orange) and validation loss (Blue) change during training on Street dataset for 700 epochs.

validation loss shows different trends for Street data and Booth data: the rate is about 20% for Street data, while it is about 5% for Booth data. This will be discussed in Section 3.1.

The trained VAE network recovers an original input from an incomplete input, in which some circles are removed from the original input as shown in Figure 3. In Figure 3, the reconstruction seems to be performing well considering 20% of gap between training and validation error for Street data. This can be interpreted as that the training on image pixel values is suitable for the current network and we are complementing the reconstructed image in our mind.

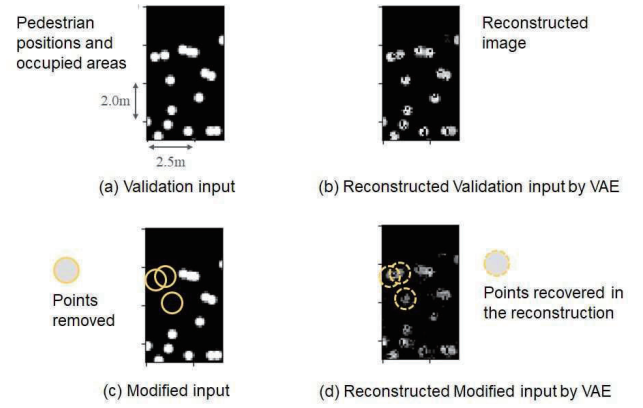


Figure 3 – **Street Data Reconstruction by the trained network:** a) Original input not exposed during the training, b) Reconstructed image by VAE, c) Modified input by removing three points from the original input, and d) Reconstruction of the modified input by VAE

3. Analysis

3.1 Generalization capabilities

The difference in the rate of training loss vs validation loss represents a generalization capability of the trained network. First, Booth data have longer duration than Street data and a longer training data have an advantage for generalization. Second, pedestrians in Street data are constantly and randomly moving in the view. In contrast, people in Booth data sometimes stop and see at specific locations in the view. Thus, the Booth data have advantages over Street data in regard to generalization and this resulted in the smaller difference in the training and validation loss.

Although the network is not over-fitted (Figure 2), the cost function after training is highly sensitive to similarity of an input data to the training dataset. If a training input is slightly shifted before reconstruction by the trained network, the network tends to recover the original input by canceling out the amount of shift. In other words, the VAE network learned the pedestrian distribution specific to the training dataset and it tries to interpret any input by learned patterns.

This generalization capability will be improved by preparing larger dataset and data augmentation.

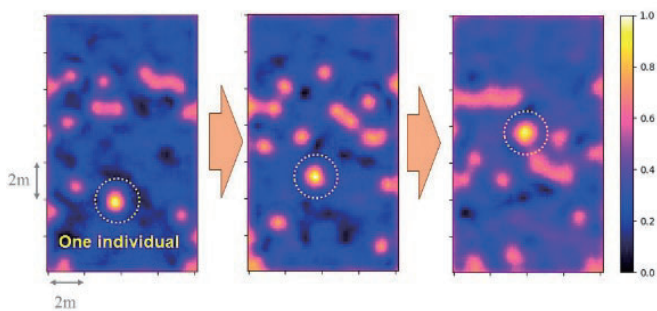


Figure 4 – **Likelihood distribution over 2-D area**: images are ordered in time sequence from left to right. Color shows likelihood of finding an individual in this area.

3.2 Likelihood distribution on expected trajectories

When an input x is fed to the trained VAE and reconstructed, likelihood of x can be estimated by $f(x)$. To estimate expectation to find an individual within a crowd, first one specific circle in a frame is selected and removed to create a base input x_0 . Then, we put a new circle on x_0 to generate a new test input x_1 . By evaluating $f(x_1)$, one can estimate a likelihood of x_1 , which is a quantified expectation to find the removed individual at the position where the new circle is placed.

By repeating this process while moving a new circle over the area of x_0 , one can estimate a likelihood distribution over 2-D area. Images generated via this procedure are presented in Figure 4. These images can be seen as our expectation for a pedestrian location when other pedestrians are distributed as x_0 .

By processing multiple frames in sequence, one can estimate an expected trajectory of an individual based on observations (Figure 4).

In the process described above, only one circle is removed and the likelihood distribution is estimated. The same process is applicable when removing a cluster of circles to generate x_0 and putting a new area, which is not necessary to be a circle, to create x_1 . In this case, the resulted 2-D likelihood distribution means our expectation to find a cluster of people around the area instead of an individual. This can be regarded as our ‘soft-focused’ expectation over a crowd.

4. Conclusion

In the present study, we discussed unsupervised training of VAE on pedestrian traffic data. Since our focus was training the network only by pedestrian positions and occupied areas in a public space, movies for training are preprocessed to generate input images representing only these features. As a result, the training was made computationally less intensive and VAE network was successfully trained with relatively small amount of data.

The trained VAE captures features of a pedestrian distribution in a frame and it is utilized to render our expectation to the trajectory of an individual in a crowd. Images rendered by this method correspond to subjective images usually only captured in our mind.

Our goal is to capture stochastic aspect of pedestrian behaviors in a crowd by machine learning. In the present study, outcomes of our first attempts to embody this concept are reported. Further analysis is required to achieve an appropriate level of a generalization capability. In addition, the current method is based only on an instantaneous pedestrian distribution, and does not take into account temporal correlations. Deep neural network based models are applied to predict pedestrian trajectories, and prediction performance of various architecture is compared in a recent study [Becker 2018]. These are areas to extend the present study.

References

- [Yokojima 2018] Yasunori Yokojima, and Toshihiko Nakazawa, Applying deep learning to test design process in hydraulic systems design 深層学習を応用した油圧システム設計におけるテスト設計プロセス. Journal of the Japan Fluid Power System Society **49**, 71-74, 2018
- [Sakai 2019] Sakai and Yokojima, Development of Deep Learning technology for a pedestrian trajectory study without requiring explicit labeling, Annual Congress(Spring), JSAE, 2019.
- [Kingma 2014] Diederik P. Kingma and Max Welling, Auto-Encoding Variational Bayes, arXiv:1312.6114v10, 2014.
- [Doersch 2016] Carl Doersch, Tutorial on Variational Autoencoders, arXiv:1606.05908v2, 2016.
- [Mochihashi 2017] Daichi Mochihashi, “Variational Bayesian methods for Natural Language Processing”, ATR SLC. 2005-6-21, <http://chasen.org/~daiti-m/paper/vb-nlp-tutorial.pdf>, accessed 2019-01-24.
- [kenmatsu4 2017] kenmatsu4, “Variational Autoencoder 徹底説”, Qiita, 2017-8-1, <https://qiita.com/kenmatsu4/items/b029d697e9995d93aa24>, accessed 2019-01-24.
- [Nitta 2016] Kazuki Nitta, “Variational Autoencoder”, www.SlideShare.net, 2016-11-11, <https://www.slideshare.net/KazukiNitta/variational-autoencoder-68705109>, accessed 2019-01-24.
- [Tatsuno 2016] Sho Tatsuno, “猫でもわかる Variational Autoencoder”, www.SlideShare.net, 2016-7-29, <https://www.slideshare.net/ssusere55c63/variational-autoencoder-64515581>, accessed 2019-01-24.
- [Nozawa 2016] Kento NOZAWA, “Variational Auto Encoder”, Red Green Black and White, 2016-12-1, <https://nzw0301.github.io/notes/vae.pdf>, accessed 2019-01-24.
- [fchollet 2018] fchollet, “Deep Learning for humans”, GitHub, 2018-11-1, <https://github.com/keras-team/keras>, accessed 2019-01-24.
- [taehoonlee 2018] taehoonlee, “variational_autoencoder.py”, GitHub, 2018-11-1, https://github.com/keras-team/keras/blob/master/examples/variational_autoencoder.py, accessed 2019-01-24.
- [Chollet 2019] Francois Chollet, “Building Autoencoders in Keras”, The Keras Blog, 2016-5-14, <https://blog.keras.io/building-autoencoders-in-keras.html>, accessed 2019-01-24.

- [qqwweee 2018] qqwweee, “A Keras implementation of YOLOv3 (Tensorflow backend)”, GitHub, 2018-11-1, <https://github.com/qqwweee/keras-yolo3>, accessed 2019-01-24.
- [iss-f 2018] iss-f, “VAE を keras で実装”, Qiita, 2018-2-12, <https://qiita.com/iss-f/items/c43b23815fc6d15ae22e>, accessed 2019-01-24.
- [Redmon 2015] Joseph Chet Redmon, “YOLO: Real-Time Object Detection”, Darknet Neural Network framework, 2015, <https://pjreddie.com/darknet/yolo/>, accessed 2019-01-24.
- [Redmon 2013] Joseph Redmon, “Convolutional Neural Networks <http://pjreddie.com/darknet/>”, GitHub, 2013-11-3, <https://github.com/pjreddie/darknet/>, accessed 2019-01-24.
- [Kathuria 2018] Ayoosh Kathuria, “A PyTorch implementation of the YOLO v3 object detection algorithm”, GitHub, 2018-2-25, <https://github.com/ayoozhkathuria/pytorch-yolo-v3/>, accessed 2019-01-24.
- [Bergstra 2012] James Bergstra: Random Search for Hyper-Parameter Optimization, Journal of Machine Learning Research 13, 2012.
- [Patterson 2017] Josh Patterson: Deep Learning A Practitioner’s Approach, O’Reilly Media, p.1-403, 2017.
- [Sugomori 2017] Yusuke Sugomori: Deep Learning: Practical Neural Networks with Java, Packt Publishing, 2017.
- [Sugomori 2016] Yusuke Sugomori: Deep Learning Java programming, Packt publishing, 2016.
- [Henrik 2017] Henrik B.: Machine Learning, Impress corporation, 2017.
- [Fujita 2016] K. Fujita, A.Takahara: 実装ディープラーニング, Ohmsha, 2016.
- [Saito 2016] Y. Saito: Deep Learning–Python, O’Reilly Japan, 2016.
- [Becker 2018] Stefan Becker, Ronny Hug, Wolfgang Hübner and Michael Arens, Notes on the TrajNet Benchmark, arXiv:1805.07663v6, 2018.
- [Newell 2002] G.F.Newell, A simplified car-following theory: a lower order model, Transportation Research Part B: Methodological, 36, 2002.
- [Sakai 2018] Tatsuhide Sakai: An Automatic Search to EV design variables using Reinforcement Learning, EVS31 Kobe Japan, 2018.