Fri. Jun 7, 2019

Room D

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

[4D2-OS-6] 人工知能と倫理

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

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

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

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

12:00 PM - 12:20 PM

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

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

12:20 PM - 12:40 PM

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

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

12:40 PM - 1:00 PM

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

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

1:00 PM - 1:20 PM

Room E

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

[4E2-OS-7a] AIの法学の応用(1)

佐藤 健（国立情報学研究所）、新田 克己（国立情報学研究所）

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

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

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

12:00 PM - 12:40 PM

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

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

12:40 PM - 1:00 PM

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

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

1:00 PM - 1:20 PM

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

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

1:20 PM - 1:40 PM

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

[4E3-OS-7b] AIの法学の応用(2)

佐藤 健（国立情報学研究所）、新田 克己（国立情報学研究所）

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

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

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

2:00 PM - 2:20 PM

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

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

2:20 PM - 2:40 PM

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

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

2:40 PM - 3:00 PM

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

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

3:00 PM - 3:20 PM

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

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

©The Japanese Society for Artificial Intelligence

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

Room F

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

[4F2-OS-11a] 社会的信号処理と AI(1)

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

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

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

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

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

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

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

[4F3-OS-11b] 社会的信号処理と AI(2)

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

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

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

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

Room G

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

[4G2-OS-8a] マイニングと知識創発(1)

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

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

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

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

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

©The Japanese Society for Artificial Intelligence

Deep Learning Networks

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

1:00 PM - 1:20 PM

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

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

1:20 PM - 1:40 PM

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

[4G3-OS-8b] マイニングと知識創発(2)

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

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

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

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

2:00 PM - 2:20 PM

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

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

2:20 PM - 2:40 PM

©The Japanese Society for Artificial Intelligence

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

[4D2-OS-6] 人工知能と倫理

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

Fri. Jun 7, 2019 12:00 PM - 1:20 PM Room D (301B Medium meeting room)

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

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

12:00 PM - 12:20 PM

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

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

12:20 PM - 12:40 PM

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

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

12:40 PM - 1:00 PM

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

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

1:00 PM - 1:20 PM

©The Japanese Society for Artificial Intelligence

医療分野における AI への信頼

Trust in Medical AI

江間有沙[*1] Arisa Ema　長倉克枝[*2] Katsue Nagakura　藤田卓仙[*3] Takanori Fujita

[*1] 東京大学 The University of Tokyo　[*2] エムスリー株式会社 M3, Inc.　[*3] 慶應義塾大学 Keio University

Discussion on AI principles related to privacy, security and accountability are important along with the usage of artificial intelligence (AI) at various fields. There are expectations as to build "trustworthy AI system." We investigated what are the bottleneck on medical AI by creating a workflow diagram of medical treatment by doctors and questioned medical-related experts participated to seminars. As a result, issues on some dilemma situation as well as usefulness (accuracy) of technology, safety, convenience, security, cost were pointed out. In the future, we plan to deepen this qualitative data further by questionnaire survey targeting a large number of people.

1. はじめに

現在，様々な現場で人工知能（AI）利用に関する議論と同時に，プライバシーやセキュリティ，アカウンタビリティなどに関する原則作りが国内外行われている．いかに「信頼できる AI システム」を構築できるかに関しては期待も多いが課題も多い．

特に医療 AI は個人データの扱いとプライバシーや，事故時の責任の所在や法的，社会的，倫理的な課題も多く取り上げられている．2015 年に厚生労働省で取りまとめられた「保健医療 2035」などにおいても，ICT を活用によって医療の質，価値，安全性，パフォーマンスの向上が目的として掲げられているものの，その活用が進んでいないことも同時に指摘されている[厚生労働省 15]．

その後，2016 年には厚生労働省内において「保健医療分野における ICT 活用推進懇談会」が提言書をとりまとめ，2017 年には「保健医療分野における AI 活用推進懇談会」が設置された．また，日本医療研究開発機構（AMED）では国内の 6 学会が主導し AI による診断・治療支援のため画像診断データベースを構築する事業を実施，2018 年には保健医療分野 AI 開発加速コンソーシアムが設置され，課題や対応策の検討が進められている[厚生労働省 18]．

2. 枠組み

様々な課題が指摘されている中，「AI を活用した」と銘打った技術はどのような観点で現場から求められているのだろうか．AI と雇用をめぐる議論においては，必要なのは「最先端の技術」ではなく，「こなれた」技術でもよいから壊れなくて理解が可能なものであったり，そもそも紙資料の「データ化・共有化」だけでもよかったりする（逆に言うとそれすらできていない）ことが医療現場以外でも指摘されている[江間 19]．

そこで本稿では筆者らが企画した医療 AI に関するセミナーにおいて，医療側からのニーズを整理することを目的とする．

2.1 ワークフロー図の作成

保健医療分野において活用される AI を，技術的な観点ではなく医療の診療プロセスの観点から整理するため，本研究では医師による診療のワークフロー図を作成した（図1）．ワークフロー図の作成にあたっては，(A) 診療プロセスを医師が行うべき「ステップ」と，(B) 医師が主体となって判断を下す「サブステップ」に分けた[横山 19]．AI は(B)「サブステップ」の効率を上げて情報を提示する支援ツールであることが医師法第 17 条の規定との関係で整理されている．

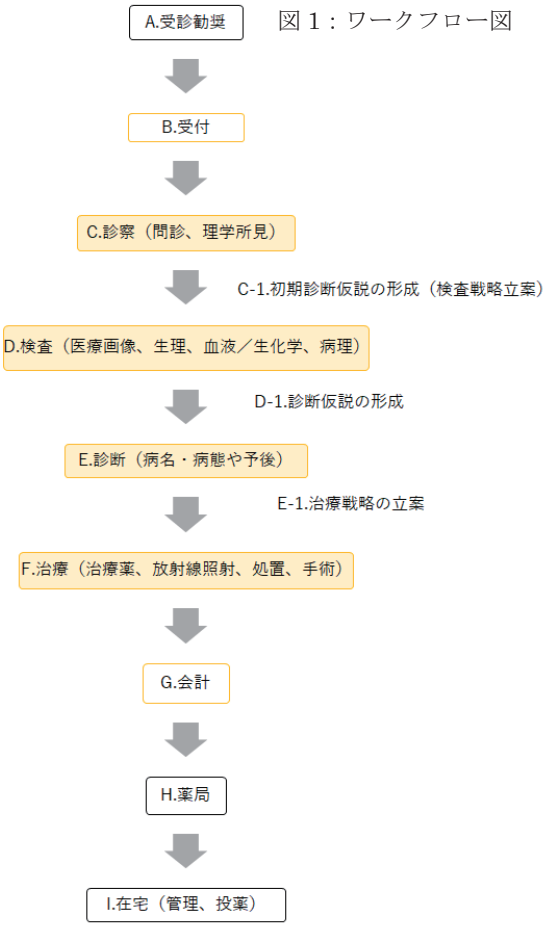


図 1：ワークフロー図

2.2 対象領域と話題提供者

厚生労働省の「保健医療分野における AI 活用推進懇談会」では「AI の実用化が比較的早いと考えられる領域」として(1) ゲ

連絡先：江間有沙，東京大学政策ビジョン研究センター，ema@pari.u-tokyo.ac.jp

ノム医療, (2) 画像診断支援, (3) 診断・治療支援, (4) 医薬品開発,「AI の実用化に向けて段階的に取り組むべきと考えられる領域」として, (5) 介護・認知症, (6) 手術支援の合計 6 領域を, AI 開発を進める重点領域として位置づけている[厚生労働省 17]. 工程表には診療報酬や介護報酬の改定時期も明記されており, 早ければ 2020 年の改定における AI 活用の評価を視野に入れていると考えられる.

「AI の実用化が比較的早いと考えられる領域」のうち, (4)の医薬品開発は医師による診療には直接関係しないため, 本研究では(1) ゲノム医療, (2) 画像診断支援, (3) 診断・治療支援に対象領域を絞る.

表 1 にリスト化した全 5 回のセミナーの話題提供者には, 現場の医師や研究者として具体的に利用/開発している診断や治療支援 AI を紹介していただいた.

表 1 各回のテーマと話題提供者

テーマ	講師	所属
(1) 開業医と AI	目々澤肇氏	目々澤医院院長
	田澤雄基氏	慶應義塾大学医学部精神・神経科, MIZEN クリニック豊洲院長
(2) 消化器科と AI	平澤俊明氏	がん研有明病院上部消化管内科副部長
	青木智則氏	東京大学大学院医学系研究科内科学専攻博士課程
(3) 内科と AI	沖山翔氏	アイリス株式会社代表取締役, 救急専門医
	巣籠悠輔氏	株式会社 MICIN 最高技術責任者, 東京大学招聘講師
(4) ゲノム医療と AI	湯地晃一郎氏	東京大学医科学研究所特任准教授
	西村邦裕氏	株式会社テンクー代表取締役社長 CEO
(5) 眼科と AI	田淵仁志氏	ツカザキ病院眼科創業者・主任部長, 眼科専門医
	升本浩紀氏	ツカザキ病院眼科・人工知能エンジニアチーフ, 医師
	三宅正裕氏	京都大学大学院医学研究科眼科学教室 特定助教

2.3 質問項目

医療現場で技術実装を進めていく上での課題抽出のため, 話題提供者とセミナーの参加者には以下の 4 つの質問を投げかけている.

① 医療提供に関するフロー(表 1)のうち, あなたはどこでどのような課題を認識していますか？以下のステップのうち該当する箇所を丸印で囲み, 課題をご指摘ください. 複数ありましたらいくつでもご記入ください(例:「A.受付」に〇を付けて, その横に「待ち時間が長い」と記入)

② **①の課題**をあなたはどのように解決しようとしていますか？またはどのように解決して欲しいと思いますか？(例:受付自動システムの導入／予約制／診療時間の延長など)

③ **②の解決策**を進めるためには, まずどこから／何から着手するべきだと思いますか？(例:業者に見積もり依頼／予約システム作る／医師の新規採用)

④ **②の解決策**を進める上で, どのような障壁があると考えますか？(例:システム導入コストが高い／医師不足)

3. 結果と今後

本稿を執筆している 2018 年 2 月ではセミナーの第 3 回までが終了しており, 各回は平均して毎回 30－40 名程度, 開業医や勤務医などの医療従事者のほか製薬会社, 医療政策関係者, エンジニアなどが参加している.

今までの議論からは, 技術の有用性(精度), 安全性, 利便性, セキュリティ, コストの他, いくつかのジレンマ状況に対する課題が指摘された. 例えば以下のような問題である

- 現在の技術では治せない症例が見つかってしまった場合どうするか
- 将来的に生じるだろうリスクの予測を患者につげるべきかどうか, またそれが偽陽性であった場合, 判明するまで患者の心理的な負担につながらないだろうか
- AI による診断で確率は低いものの, 侵襲性があり患者に負荷の高い検査が必要と提示された場合, 検査を行うべきか
- AI による診断によって確率は低くても可能性のあるものに対する過剰診断が生じないか

AI が入ることによって, 受付や会計の時間が短縮されるだけではなく, 見落としや偽陰性が減り, 患者の満足度が上がり, 医者側の運用負担も減ることが期待される. 一方で, AI による診断や予測, 提案に対しては, 現在は医師に最終的な判断が委ねられている. そのため上記で挙げたような問題が増加すれば, 逆に医師の労働時間を増やし, さらには場合によっては, 患者の身体的, 心理的, 経済的コストが増えることも懸念される. 今後はこの質的なデータをさらに大人数を対象としたアンケート調査などによって深堀をしていく予定である.

なお, 各回の話題提供者の内容はウェブに報告書が掲載されている. 全 5 回のセミナーを通して, 現場の医師からはどのような AI の利用のニーズと, 技術シーズとのミスマッチがあるのかについて, 当日会場で議論を行いたい.



参考文献

[江間 19] 江間有沙:人工知能社会の歩き方」, 化学同人, 2019.

[厚生労働省 15] 厚生労働省:保健医療 2035, https://www.mhlw.go.jp/seisakunitsuite/bunya/hokabunya/shakaihoshou/hokeniryou2035/

[厚生労働省 17] 厚生労働省:保健医療分野における AI 活用推進懇談会, https://www.mhlw.go.jp/stf/shingi2/0000169233.html

[厚生労働省 18]厚生労働省:第 1 回 保健医療分野 AI 開発加速コンソーシアム(開催案内), https://www.mhlw.go.jp/stf/newpage_00177.html

[横山 19] 横山和明, 本邦における人工知能(AI)を用いた診療支援の事例, https://www.mhlw.go.jp/content/10601000/000468448.pdf

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

倫理的設計学の構築に関する一考察

A Study on Constructing the Ethical Design Sciences

関口　海良 *1　　堀　浩一 *1*2
Kaira Sekiguchi　　Koichi Hori

*1東京大学大学院工学系研究科
Graduate School of Engineering, The University of Tokyo

*2理化学研究所革新知能統合研究センター
Center for Advanced Intelligence Project, RIKEN

In the 21st century, further contribution to solving social problems is required for academic professionals. Therefore, we set our research question as "Can engineering contribute more to solving social problems?" We consider that the answer is positive, and proposed the ethical design sciences which is a new interdisciplinary field including engineering. The ethical design sciences will contribute to systematically solving social problems through re-defining representations in design such as perspectives and description methods, positioning creative activity support tools also as experimental devices and demonstrating the effects through user studies etc. In this paper, after describing practical examples, we derived a positive answer to the question.

1.　序論

貧困，飢餓，不平等，気候変動等、社会全体に関わる課題が存在している．これらの解決という目標に関して、特に重要なものは SDGs として纏められ，更に注力されるに至った．また人工知能（AI）に関しても AI を社会に上手く導入していく事が課題となり，人工知能倫理として広く研究されている．

一般に，科学研究は公的研究費を用いており，何らかの形で社会に貢献する責任を有す [吉川 09]．課題解決への期待が高い工学においては尚更，これら社会的課題の解決に対する立場を各研究者は明確にし，取り組む必要がある．

そこで本研究では，社会的課題の解決に向けて従来以上に積極的に貢献するひとつの工学の在り方を探求し，「倫理的設計学」として体系化し，実証を進めている．

2.　倫理的設計学で扱う三領域

倫理的設計学は，倫理学，設計学，創造活動支援の三分野が重なる領域に位置付けられる．

初めに，社会的な課題の解決に向けて科学／技術を用いるには，科学／技術と社会との関係についての考察が重要となる．既に環境倫理，生命倫理，技術倫理等の応用倫理はこれら影響を扱ってきた．倫理的設計学ではこれら学問を参考にしながら，逆に，望ましい変化から逆算して社会にとって良い意味を持つ人工物の設計を目指す．次に，課題解決に向けて体系的に取り組むには，設計を学術的に扱う「設計学」が重要となる．これにより実践だけでなく，知識の蓄積や共有も効率化する．

ここに至って，本研究では設計学と倫理学の重なる領域に「倫理的設計理論（ethical design theory）*1」を導出した．ここでは設計における視点の持ち方や，設計アイデアの記述方法を再定義して，上述の様な人工物に関する「倫理」と，更に技術者にとっての「倫理」に対応する軸を設計において連続的に扱える様になっている [Sekiguchi 18a]．

三つ目として，上記研究を「机上の空論」とするのを避けるべく，従来の工学が持つ良い特徴であった実証的側面を踏襲する．ここでの実証対象は，(1) 社会的課題の解決を体系的に進める為に設計行為がどの様に変化すると良いかと，(2) そこで生まれた設計解が実際に課題解決に資するかである．本研究では，前者を研究テーマとして特に明確に設定する．実証に際しては創造活動を扱うので，創造活動支援の分野を参考にしつつ，ツールを支援に用いるだけでなく，理論を実証する為の「実験装置」として位置付け，「観測データ」の収集にも用いる．更に，設計解の実証には一般的に従来から実験やシミュレーションが行われているので，これらも踏襲するものとする．

これら「理論–実験装置–観測データ」の関係は，物理学におけるそれと対応する．例えば，物理学における「一般相対性理論–レーザー干渉計型重力波検出器（LIGO）–重力波イベントの記録」の関係と類似である．

課題解決に向けた学問としてより明確に取り組むべきは，ニーズを定義し，要素を統合して，価値を生み出す事との理解がある（[吉川 09] pp.6–7 の「第二種基礎研究」，p.86 の「持続性の時代」に関する記述等）．上記三領域を扱う事により，社会的な影響も含めた上で工学としてより積極的に課題解決に貢献していけるものと筆者らは考えている．以下，これら三領域の取り組みについて紹介する．

3.　倫理的設計理論の導出

倫理的設計理論では，工学的視点に基づき倫理を扱う為に設計における表現（視点・記述方法）を再定義した．

3.1　視点としての「倫理レベルからの設計」

視点としては，人工物の階層表現を再定義し，その最上位に倫理レベルを位置付けた．倫理レベルは，自由、平等，持続可能性といった社会全体に関する価値を表現するレベルであり，設計における最終的な目的が表現されるレベルに対応する [Sekiguchi 18a]．

次に，人工物の階層表現では自然科学が対象とする客観的現象が表現されるので，設計者の主観を表現するための新たな次元を位置付けた．筆者らはこの次元を図 1 の様に階層表現に直行させた，三次元の表現を持って現在研究を進めている．

3.2　記述方法としての「言説による設計」

人工物を倫理レベルから設計アイデアを記述する為，倫理的設計理論では下記三つのルールを設定した．

連絡先: 関口 海良，東京大学大学院工学系研究科航空宇宙工学専攻 知能工学研究室，kaira @ dfrome.com

*1　以前はこれを「倫理的設計学」と呼称したが，今回は理論として位置付け直し，倫理的設計学はより広く再定義する．

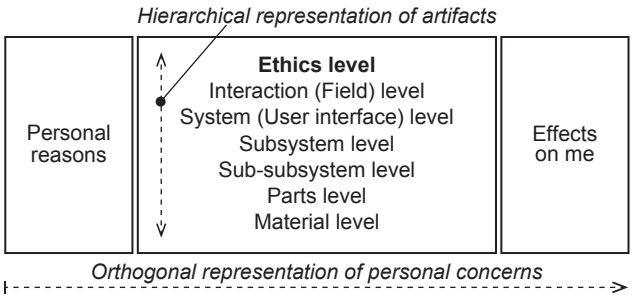


図 1: Simplified version of the three-dimensional representation of artifacts which contains hierarchical and orthogonal representation of artifacts [Sekiguchi 18a, Sekiguchi 18b]

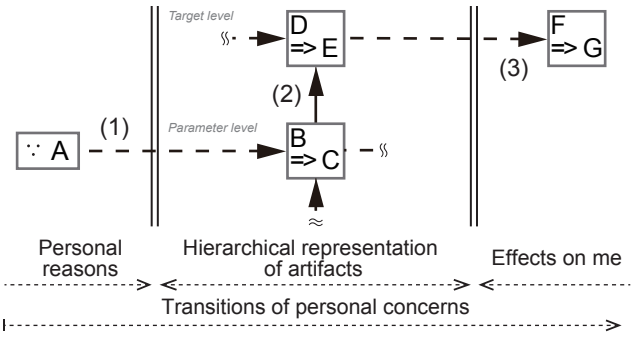


図 2: Hierarchical and orthogonal grammar of design with discourse [Sekiguchi 18a, Sekiguchi 18b]

1. Since A is a personal reason, I/we generate a design that will change B to C in the hierarchical representation of artifacts.
2. If B is changed to C at the parameter level, then D will change to E at the target level in the hierarchical representation of artifacts.
3. If D is changed to E in the hierarchical representation of artifacts, then F will change to G as effects on me/us.

また，図 2 の様に可視化の形式を提案し、式 1〜式 2 の様に「偏微分」の考え方を用いて数学的に理解できる事を整理した．

$$D = f_D(t, B, ...) \quad (1)$$

$$dD = \frac{\partial f_D}{\partial t}dt + \frac{\partial f_D}{\partial B}dB + ... \quad (2)$$

設計におけるアイデアはこれら記述の要素からなるツリー状となる．

3.3 その他の展開

他にも研究テーマとして理論的な展開があり得る．例えば，視点を更に拡張する事ができる．図 1 では，1 次元目が人工物の階層，2 次元目が設計解の併記を可能にする広がり，3 次元目がこれらに直行させた主観的な軸であった．式 1 や式 2 に表現されている様に，これまでも時間を考慮していたが，これは設計解の中の世界に流れる言わば内的時間（inner time）である．一方，他にも設計の外側に流れる外的時間（outer time），

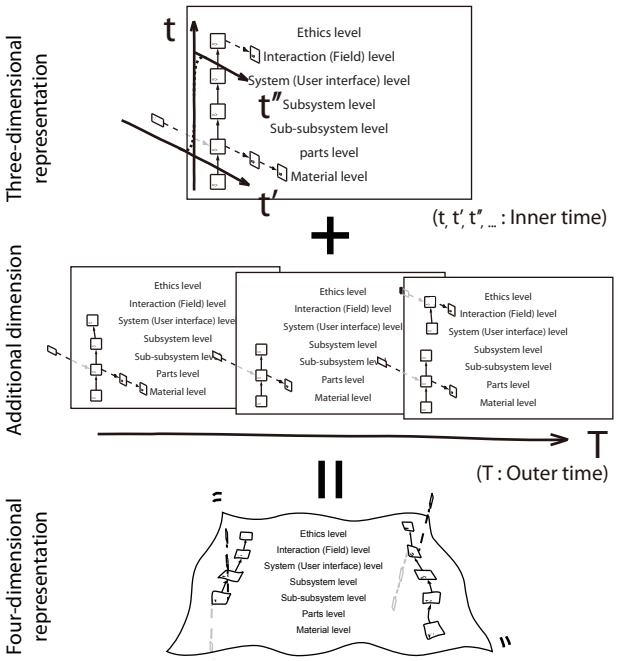


図 3: Overview of the expansion of the dimensions in design

言わば設計者が生きている時間が存在する．例えば，外的時間は設計を見直してアイデアを更新するといった，設計解の時系列変化を扱う際に対象となる．図 3 に示す様に，この外的時間を加える事で 4 次元の表現を考える事ができる [*2]．

4 次元の表現を考える事で時代の「波」を扱う事ができる．例えば，人工知能は第三次ブームに在るが，これが倫理レベルまで含めてどのような波となっているか，どの手を打つ事で波が制御されこのブームをソフトランディングさせられるか等が扱える．この 4 次元化は筆者らも今後の課題としている．

4. 有機的で動的な支援ツールの実装

倫理的設計理論に基づく設計の実践を支援する為に，筆者らは支援ツールとして Dfrome（Design FROM the Ethics level の略）を実装した．また，Dfrome は倫理的設計理論を実証する為の実験装置として位置付けた．

図 4 が Dfrome の全体像であり，主にブラウザ（browser），エディタ（editor），クラウド環境（cloud）と，探求用エンジン（investigation engine）からなる [Sekiguchi 18b]．以下，探求用エンジンにおける各種処理を紹介する．

4.1 自動マッピング及び自動連結

Dfrome は，文章の形で与えられたインプットから，変化の記述の要素を表現空間上に自動でマッピングし，関連する要素同士を自動で連結する機能を有す．支援ツールとしては編集時にツリーのたたき台を提供し，実験装置としてはツリーの編集が提案者以外の設計者でも可能である事を実証する為に用いる（詳細は 5.1 節参照）．基本的な処理は，既存のツリーを参考にして，どの次元のどの座標に要素を配置すると良いか，また要素間に矢印を引けるかを類推している．

*2 試論は以下参照 : Kaira Sekiguchi, The sixth and the seventh rules of “design with discourse” for design from the ethics level, January 1, 2011. ethics-level.com.

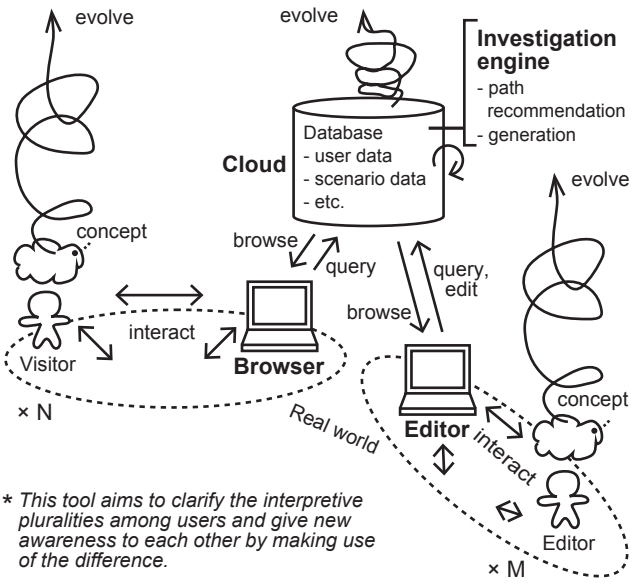


図 4: Overview of Dfrome [Sekiguchi 18b]

4.2 シナリオパスの延長

Dfrome は，考慮すると良いでシナリオパスを推薦する機能を有す．例えば，倫理レベルまで繋がるシナリオや，考慮していそうでしていなかった他の実現手段はどの様なものが在るかを提示する．これらの機能は，シナリオを拡げる過程の再現性の実証としても位置付けられる．こをエディタで行ったケースは [Sekiguchi 18a] に，ブラウザで行ったケースは [Sekiguchi 18b] に記載した．

4.3 ツリー間の距離の計算

Dfrome は，編集者の手元にあるツリーをクエリとして，他のツリーとの距離を計算する機能を有す．これにより，ツリー単位でもどの様な視点を考慮できており，何が漏れていたかを確認するのを支援し，またその過程の実証となる [Sekiguchi 18a]．

4.4 Dfrome のコンテンツ

Dfrome のデータベースには，2019 年 2 月 7 日時点で 122 個のツリーが既存のものとして利用可能（＝非削除かつ公開状態）である．例えば，Asilomar AI Principles や Ethically Aligned Design（EAD）の Principles [IEEE 17] を始めとした人工知能倫理のツリーに加え，機械学習のアルゴリズムやドローンの研究に関するツリーを含む．また，理化学研究所革新知能統合研究センター（RIKEN AIP）がウェブに公開している研究概要 [RIKEN] を自動マッピング及び自動連結を用いながら筆者らが編集したツリーも含む．

5. 観察データの収集

倫理的設計理論の実証を本研究では二段階で行っている．まず，本理論の運用が提案者の暗黙知に依存するなら実用的でない為，そうではなく第三者でも運用可能な事を示す．次に，ユーザスタディによって本理論が設計者に期待通りの効果を与えるかや，想定外の反応が見られるか等を観察する．本来は，設計された人工物が実装され社会に投入された際に期待通りの機能を発現するかを，シミュレーションや社会的実験によって確認する必要がある．これらも倫理的設計学の範疇であるが，その実践は今後の課題としている．

5.1 AI による記述の再現性の確認

初めに，AI によって倫理的設計理論に基づく記述を行う事ができれば，第三者による運用の再現性の証拠となる．

以下の実験では，インプットの文章からの表現空間へのマッピングと，要素間に矢印を引く処理を AI が担い，その他は筆者らで行った．実験条件を保存するため，今回はデータをダウンロードし，ローカルに環境を再構築して実験を行った．加えて，本理論の発展的な運用についても再現性を確認する為，機械学習に関する文章をインプットとしてツリーを作成した後，社会的課題に関する文章を続けてインプットとして，これら異なる文章間に隠れた有用な繋がりを見つけられるかを確認した（図 5 の 1〜2）．適当なインプットを設定する為，「Google Scholar」を用いて，クエリを「machine learning」とした際のトップにランクされた論文 [Baltrušaitis 19] と，「social problems」とした際のトップの論文 [McIntyre 19] の各アブストラクトを用いた．最近の論文を得る為，設定で「2019 年以降」をオンにし，「特許を含める」と「引用部分を含める」をオフにした．

図 5 は試行の結果である．「直さないといけないがそこそこ書けている程度のツリー」が得られている．更に，ツリーを跨いで下記記述 A から B へつなぐ矢印が確認された．

A) Multimodal machine learning => aim to build models that can process and relate information from multiple modalities

B) Results => reveal that discussing an effective solution to a social problem in a news story caused readers to feel less negative and to report more favorable attitudes toward the news article and toward solutions to the problem than when no solution or an ineffective solution was mentioned

この結果から，例えば AI をマルチモーダルにしてニュース等のコンテンツを作成する際には効果的な解決策にも言及する事で，社会的には更なる議論の活性化へと繋がるといった「興味深い」気付きを得られる．ツールは dfrome.com にて提供済みの為，類似の追試を広く可能としている．

5.2 ユーザスタディによる効果の確認

本理論の効果を確認する為，筆者らはユーザスタディを進めている．学生実験からは，例えば Google のみを用いた場合では被験者が既に持っている問題意識に基づき検索を行う傾向があるのに対して，Dfrome を併用すると倫理的な視点の導入を促す傾向があるのに加え，被験者が研究テーマを見直す機会もその中で得られ得るという創造活動支援としての効果も確認できた．例えば，ある被験者はアンケートに「利点については今までも考えたことがあったが，リスクについては考慮していなかったため，リスクについて新たな気づきが得られた」と回答した．今後はより大規模な実験を実施する予定である．

6. 関連研究との位置付け

[Verbeek 11] はひとつの哲学的基盤を本研究に提供する．例えば，4 次元まで考慮した上での図 2 全体の実践は彼の言う自由（freedom）に基づく設計として，図 2 の（3）の中央から右側へ向かう個人的影響の発現が媒介（mediation）として位置付け可能である．一方，倫理的設計学は理論自身の実証を含む為より自然科学的である．更に本理論は階層や記述方法を設計学的に含む為より包括的かつ実用的であり，その実践が創造活動支援に資する事も本研究では重視して実証を進めている．

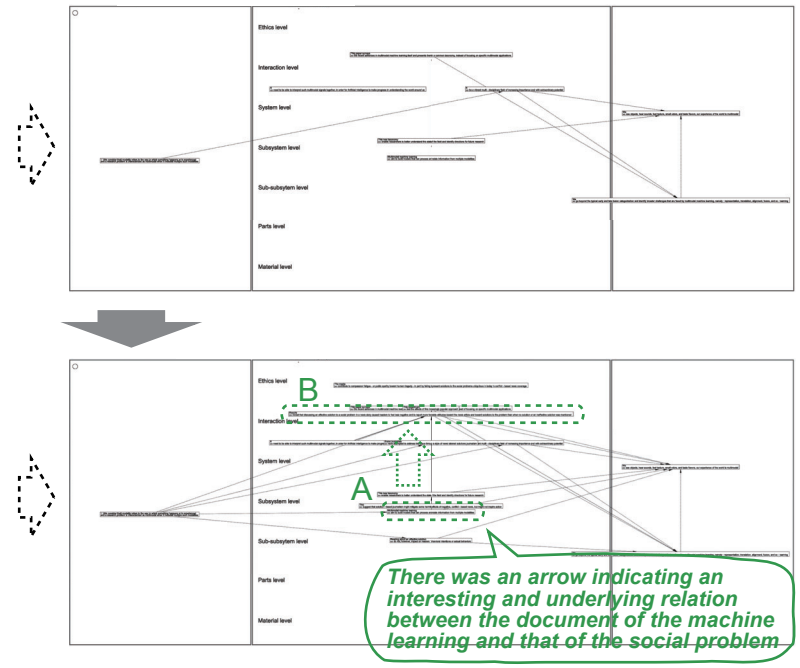


図 5: Image of a process of the generation of items and arrows [Baltrušaitis 19, McIntyre 19].

一般設計学（general design theory）やそれを包含する本格研究（full research）については [吉川 09] に詳しい．一般設計学を提唱した吉川自身,「... 現在，設計の本質は... 依然として明確に記述することはできないでいる」（[吉川 09] p.160）と記載しているが，倫理的設計理論はここに挑戦し，本論文も含め証拠の提示を始めている．更に研究体制全体を纏めた倫理的設計学は，吉川の本格研究に倫理レベル（設計の最終的な目的に対応する領域）に関する考察を有機的に連結できている為，残されていた本質的な領域をそこに拡張できたものとして位置付け可能である．

7. 結論と今後の課題

記述した倫理的設計学の実践例によって，社会的課題の解決に向けて工学が体系的かつ実証的により陽に倫理的設計に貢献できる事を示した．また豊かな研究テーマも設定可能な為，研究者の自発的参加を促すポテンシャルも確認できた．

Dfrome に関する今後の課題としては，最終的には各機能を繋げつつ，またスクレイピングを用いながら時系列でのデータを増やし，4 次元での自動設計までを目指す．ただし，Dfrome は唯一の解ではないので，他のツールとの連携も図る．更に，理論やコンテンツの深化を目指して，人文・社会科学を含む多分野を跨ぐコラボレーションを今後も進める予定である．

謝辞

本研究は JSPS 科研費 JP18K18434 の助成を受けたものです．

参考文献

[Baltrušaitis 19] Baltrušaitis, T., Ahuja, C., and Morency, L.-P.: Multimodal Machine Learning: A Survey and Taxonomy, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 41, No. 2, pp. 423–443 (2019)

[IEEE 17] IEEE, : The IEEE Global Institute for Ethical Consideration in Artificial and Autonomous Systems, Ethically Aligned Design: A Vision for Prioritizing Wellbeing with Artificial Intelligence and Autonomous Systems, Version 2, http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html (2017)

[McIntyre 19] McIntyre, K.: Solutions Journalism, *Journalism Practice*, Vol. 13, No. 1, pp. 16–34 (2019)

[RIKEN] RIKEN, : Center for Advanced Intelligence Project, http://www.riken.jp/en/research/labs/aip/

[Sekiguchi 18a] Sekiguchi, K. and Hori, K.: Organic and Dynamic Tool for Use with Knowledge Base of AI Ethics for Promoting Engineers' Practice of Ethical AI Design, *AI & Society* (2018)

[Sekiguchi 18b] Sekiguchi, K. and Hori, K.: Realization of Organic and Dynamic Creativity Support Tool for Promoting Ethical AI Design, in *Proceedings of the 13th International Conference on Knowledge, Information and Creativity Support Systems*, KICSS2018 (2018)

[Verbeek 11] Verbeek, P.-P.: *Moralizing Technology; Understanding and Designing the Morality of Things*, University of Chicago Press (2011), (邦訳：技術の道徳化：事物の道徳性を理解し設計する, 鈴木 俊洋 訳, 法政大学出版局 (2015))

[吉川 09] 吉川 弘之：本格研究, 東京大学出版会 (2009)

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

プロファイリングに関する自主的規律導入支援の実践

Practice of a support for creating self-regulations on profiling

工藤郁子[*1]
Fumiko Kudo

[*1] 中京大学経済研究所
Institute of Economics Chukyo University

This report discusses a policy making process about automated processing, including profiling. Our project team provided an analysis of social issues on the use of personal data to evaluate certain aspects related to the individual and published our recommendations on that. The key finding was that some companies and industry groups address these issues from the standpoint of reputation management.

1. プロファイリングをめぐる社会的課題

ICT 技術の進展に伴い、いわゆる「プロファイリング」が生活の中に浸透している。従来から、データに基づいて対象者の趣味趣向を予測するアルゴリズムは、閲覧や購買の履歴に基づいてインターネット広告を表示するターゲティングなどに利用されてきた。さらに近年では、マーケティングを超えて、与信、採用・人事などの場面でも積極的に利用されつつある。例えば、中国では、アリババ・グループ傘下の芝麻信用が、アルゴリズムを用いて個人の社会的信用度をスコア化するサービスを提供している。日本でも、みずほ銀行とソフトバンクが共同開発した社会信用スコアリング・サービス「J.Score」が与信領域において既に運用を開始している。企業の採用活動や労務管理の場面でも、人工知能を用いた求職者や従業員の適性プロファイリングが急速に実装されている。

そのため、アルゴリズムが、個人の自由で主体的な生活に重要な影響を与えうることが懸念されている。また、欧米を中心に、法的・技術的な規律の導入も進んでいる。例えば、2018 年 5 月に適用開始された EU 一般データ保護規則（General Data Protection Regulation, GDPR）は、プロファイリングを「自然人に関する特定の個人的側面を評価するために、特に、当該自然人の職務遂行能力、経済状況、健康、個人的選好、関心、信頼性、行動、位置もしくは動向を分析または予測するために、パーソナルデータを用いて行うあらゆる形式の自動化されたパーソナルデータの処理」と定義し（4 条 4 項）、一定の規制をした。

しかし、日本では、現状において、どのような点に注意すればプロファイリングを実施する（または実施しようとしている）技術者や企業が法的・社会的責任を果たしたといえるかを示す具体的指針がなく、健全な技術開発や事業展開の歩みを妨げている。

2. プロファイリング提言案と中間報告書

2.1 3 つの提言案

筆者が所属するパーソナルデータ＋α研究会では、このような認識の下で、2018 年 5 月より「プロファイリングに関する提言案作成タスクフォース」を設置し、プロファイリングに対する適切な規律のあり方について検討を重ね、以下 3 つの提言案をまとめた [PD+α研 2018]。

1. **【企業：プロファイリングに関するコンプライアンス体制等の整備】**
   与信、採用・人事等にプロファイリングを用いる企業は、プロファイリングの実施、活用等に当たって、法令との抵触を避けるためのコンプライアンス体制を整備すべきである。その際、プロファイリングが、プライバシーのみならず、平等や公正概念との関係でも問題になり得る点に留意すべきである。また、上記企業は、プロファイリングの実施・活用等に伴う社会的・倫理的責任に配慮した体制づくりに努めるべきである。
2. **【業界団体等：プロファイリングに関する自主的な取組みの支援】**
   業界団体・認定個人情報保護団体等は、各企業のコンプライアンスや自主的取組みを支援すべく、ガイドライン等の作成を検討することを推奨する。その際、プロファイリングの対象となる個々の利用者・生活者等（以下「被評価者」という。）と企業の間に立って、利害関係の調整や課題解決を目指すべきである。
3. **【政府：民間主導の自主的な取組みの尊重】**
   政府は、民間による自主的な取組みを尊重すべきである。その取組みの実効性を慎重に見定めながら、企業・業界団体・認定個人情報保護団体等との対話を続けるべきである。そして、プロファイリングによる個人の権利利益の侵害リスクを看過し得なくなった場合には、それを抑止するために必要な立法措置等を検討すべきである。

提言案およびそれに付属する「中間報告書」は、2018 年末からインターネット上で公開された（https://www.shojihomu-portal.jp/nbl1137pc）。2019 年度内に公表予定の最終報告書の作成に向けて、意見を公募している。

2.2 自主的取組みに関するチェックリスト

中間報告書では、プロファイリングの実施に際して企業等が法的・社会的責任を果たすための留意点（「自主的取組みに関するチェックリスト」）をまとめた。

手法としては、以下のような資料を参照することで、論点を抽出した。(a) EU の GDPR、(b) 米国のビッグデータレポート[FTC 2016]、(c) 第 40 回データ保護プライバシー・コミッショナー国際会議において採択された「AI における倫理及びデータ保護についての宣言」[ICDPPC 2018]、(d) Google やソニーグループなどが発表した AI の倫理原則等、(e) プロファイリングやビッグ

データをめぐる報道記事等である。さらに、(f)一般社団法人ピープルアナリティクス&HR テクノロジー協会、一般財団法人日本情報経済社会推進協会、株式会社メルペイへのヒアリングや関係する官公庁との非公式な意見交換も実施した。

チェックリストではプロファイリングの工程を、(1)企画・設計、(2)データの取得、(3)データの加工・分析・学習、(4)実装・運用・評価、という 4 段階に区分し、企業が考慮すべき度合い(優先度)を 3 段階に分けて表示した(「★★★」「★★☆」「★☆☆」の順に優先度が高い)。

チェックリストの見出しは、以下の通りとなった。なお、筆者としては、チェックリストの項目すべてを遵守する必要はないことを強調したつもりである。なぜなら、「プロファイリングを行うことによって何を実現したいのか」という意図・姿勢を示して、ビジネス上の信頼を獲得することが最も重要だからだ。

1 企画・設計段階
1-1 プロファイリング導入の適否判断 ★☆☆
1-2 プロファイリングの実施の有無・利用目的の明確化 ★★★
1-3 エシックス・バイ・デザインの導入検討 ★★☆

2 データの取得段階
2-1 データの適正な取得 ★★★
2-2 データ取得時のユーザー・インターフェース等の工夫 ★☆☆
2-3 データセットの偏向に関する配慮 ★★☆

3 データの加工・分析・学習段階
3-1 分析・評価段階におけるプライバシー侵害リスクの検討 ★★★
3-2 公平性に配慮した学習技術の導入検討 ★★☆
3-3 アカウンタビリティ(答責性)に配慮したモデルの導入検討 ★★☆
3-4 その他の社会的価値への配慮 ★☆☆

4 実装・運用・評価段階
4-1 人間関与原則の検討 ★★★
4-2 答責性への配慮 ★★★
4-3 安全管理措置の実施 ★★☆
4-4 データ内容の正確性の確保 ★★☆
4-5 開示・訂正等の手続きの整備 ★★★
4-6 プロファイリングに関するリテラシー向上 ★☆☆
4-7 被評価者等の反応の精査・検討 ★☆☆
4-8 信用スコアを第三者提供する手続きの確認 ★★☆
4-9 監査可能性への配慮 ★☆☆

2.3 チェックリストの具体例

各項目では、留意点の説明のほか、「ケース」を添えて理解を促した。ここでは、具体例として「4-2」を抜粋する。

(4-2) 答責性への配慮 ★★★

答責性、透明性等への配慮は、実装・運用・評価段階において、より重要性を増す。本研究会は、企業が、決定の過程(決定を導くデータを含む)について、被評価者や他のステークホルダーに非専門用語で説明できるかどうかを検討することを推奨する(【ケース H】参照)。

【ケース H】再犯リスク予測プログラムとアルゴリズムの透明性

米国では一部の州で有罪確定者の再犯リスク予測プログラムが用いられている。ただし、このプログラムのアルゴリズムが明らかにされていないため、バイアスがあるのではないかとの批判もある。ある再犯リスク予想プログラムでは、アフリカ系を白人よりもリスクを高く評価するが、アフリカ系たちが実際に再犯をした割合は半分以下であったという調査も報告されている。データやアルゴリズムのバイアスに関しての検証をするためにも技術システムの透明性や答責性が開発側には求められている。

もっとも、説明内容や程度の妥当性については、いまだ社会的合意はない。また、AI のプロファイリングにのみ答責性を求めるのは均衡を欠くとの指摘もある。これは例えば、人間の採用担当者であれば不採用理由を開示しなくても問題視されない現状をどう捉えるかという問題である。他方で、プロファイリングの過程や評価理由が不透明であれば、能力や信用力に関する評価の基礎になるのかがわからず、将来の行動計画を練れなくなり、自律的・主体的に自らの人生を歩んでいくことが困難になるとの懸念も存在する。

上記を踏まえると、おそらく当面のベンチマークにすべきは、事後(特に苦情対応時)に被評価者が理解し得るような「理由(justification)」を示すことであろう。ここで被評価者に事後提示される「理由」は、実際の決定過程を完全に反映したものである必要はない(別のアプローチとして、【ケース I】)。関連して、決定理由を人間に提示できる「説明可能な AI」(Explainable AI: XAI)の開発動向にも関心を向ける必要があろう。

【ケース I】信頼の構築

ディープラーニングの進展に伴う説明困難性を受け止め、「信頼(trust)」や「信用(fiduciary)」の構築を重視する立場もある。例えば、医薬品の作用機序を十分に理解していなくても、医師や薬剤師の勧めに従って服薬することが一般的である。これは、高い能力・専門性を備えた医師等に任せた方が合理的であるという考えや、国家資格による認証や専門職集団の行動規範(code of conduct)による制約などが制度的に担保されているから信じるに足るとの考えが背景となっていよう。このような信頼構築のための規律構造を AI 化社会においても実現しようとする試みが存在する。

3. 自主的規律導入に関する考察と今後の展開

提言案と中間報告書の作成のためにヒアリングを行った過程で、企業や業界団体等がプロファイリングに関する自主的規律の導入に前向きであることがわかった。おそらく、レピュテーション・マネジメントの観点から有益であるためと見られる。

筆者としては、チェックリスト等が、自主的な取組みの「たたき台」として活用されることを期待している。今後も関係者との情報共有・意見交換を続け、プロファイリングの健全な実装に資する情報や議論を提供していきたいと考える。

参考文献(論文誌と同じスタイルを推奨)

[PD+α研 2018] パーソナルデータ+α研究会: プロファイリングに関する提言案, NBL1137 号, 商事法務, 2018, pp. 64-85.
[FTC 2016] Edith Ramirez et al.: Big Data: A Tool for Inclusion or Exclusion?, FTC Report, 2016, Federal Trade Commission Website
[ICDPPC 2018] CNIL et al.: Declaration on Ethics and Data Protection in Artificial Intelligence, ICDPPC Report, 2018, ICDPPC Website

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

人間中心設計の観点からみたSF映画の中のAI

AI in Sci-fi Movies from the Viewpoint of Human Centered Design

飯塚重善 *1
Shigeyoshi IIZUKA

*1神奈川大学
Kanagawa University

Due to the rapid development of recent AI technology, how society accepts AI and robot has emerged as a realistic and big problem. If not only the robot becomes close to human beings but also human beings are replaced by human beings and human beings depend heavily on it, there is a possibility that the human-centered society should be reconsidered about it. Therefore, in this paper, as materials to consider how we can design from the perspective and the future society where human beings and AI coexist, Sci-fi movies including relevant scenes about "Three principles of robot engineering", "What is a human?", "heart", and ethical issues are picked up. It is thought that studying AI and robot design from such viewpoints have great significance as HCD(Human Centered Design).

1. はじめに

AI(Artificial Intelligence) 社会では，人間と AI の関係性はどのようなものになるだろうか？AI やロボットが世界を大きく変えようとしており，人工知能はあらゆる産業を革新すると期待される．しかし，その影響の範囲が広過ぎるため，具体的な将来像をイメージすることは難しい．そこで我々がイメージすることを助けてくれるのが SF(Science Fiction: Sci-fi) 作品である．近年の SF 映画では，未来社会やそこに登場する事物の表現にリアリティが求められることから，登場する大道具・小道具のすべてを圧倒的にリアリティ高く製作し，また，電子情報工学の専門家，未来学者の意見を積極的に取り入れられている．そこで本稿では，人間と AI が共存する未来の社会を，どのような観点から，どのようにデザインしていくことができるかを，SF 映画を素材として，人間中心設計 (Human Centered Design, 以下 HCD) の観点から検討する．

2. 本稿における「ロボット」と「AI」の考え方

ロボット (Robot) という言葉は，1920 年にチェコスロバキアで発表された戯曲，R・U・R(Rossum's Universal Robots) で使われた，チェコ語の Robota(労働者) から「a」を取った造語である．劇中で，ロボットは労働の機能だけをもった人造人間として登場した．1927 年ドイツで発表された映画『メトロポリス』(*Metropolis*, 1927 年公開）で女性の姿をした金属のロボットが登場した．R.U.R とメトロポリスがヒットし，「ロボット」＝「人間のために働く人型の機械」という認識が人々の中に根づいていったといわれている．そして，現在ロボットと呼ばれているものの形や機能はさまざまである．例えば製造現場で働くロボットは，人間が教えたことを正確に繰りかえす機能が主で，形は人間の腕のような姿をしているものもある．また，災害地で救助活動をするロボットには，遠隔操作で障害物を除いたりする操縦型ロボットや，移動中に目の前に物があったらよけて人を探すという自ら状況を判断し行動する自律型のものもある．その他，重いものを持ち上げる時，人間の動きを補助してくれる外骨格型のロボットスーツなどもあり，形状や機能，種類は，その用途によって多岐に渡る．このように，実際にはロボットの厳密な定義はないようである．ただし，大まかにはロボットは，あらかじめプログラムされた動作を正確におこなうことが基本となっており，決められたこと以上のことを自己判断で進めることはできず，人間が遠隔操作するものなど，知的処理の部分は人間が担っていた．一方 AI は，"人間の脳がおこなっている知的な作業をコンピュータで模倣したソフトウェアやシステム" であり，"人間の使う自然言語を理解したり，論理的な推論をおこなったり，経験から学習したりするコンピュータプログラミングなどのこと" を指す．その目指す形態は自動化であり，人の脳と同じような仕組みとなるニューラルネットワークを利用し，自ら学習して精度を高めることができる仕組みをもっている．将来的には両方の機能がインテグレートされ，「ロボット」が人工知能を搭載することで自分で周囲の状況を捉え，どう行動すべきかを考え，判断して行動するようになると予想される．実際，SF 映画でもこうした「自律化 (Autonomy)」したロボットを描いていることが多い．そこで本稿では，「自律ロボット」を対象とすることで，「ロボット」と「人工知能」を同義として扱う．

連絡先: 飯塚重善，神奈川大学，〒 259-1293 神奈川県平塚市土屋 2946，iizuka@kanagawa-u.ac.jp

3. SF 映画における AI の描写

まず表 1 に，いくつかの代表的な SF 映画に登場するロボットについて，それぞれの特徴をまとめた．以下では，この中から，特筆すべきものを抽出して具体的に紹介する．

映画『オートマタ』(Automata, 2014 年公開）は，人類存続の危機的状況下での AI と人類の未来や共存，主従関係を描いており，ロボット三原則をテーマとして，これが破られている．また，映画『アイ，ロボット』(I, ROBOT, 2004 年公開）は，アシモフの作品に "ヒントを得た" として，ロボットは三原則に従いながら世界を乗っ取ることができる，という発想が映画のストーリーの中心に据えられている．ロボットが実用化されるという近未来を具体的に想定し，その設計にあたって，人間とロボットの共存関係はいかにあるべきかが重要なテーマであり，人間はそれを真剣に考えなければならないという問題提起をした作品である．この作品では，2035 年，米国シカゴを舞台とし，家庭用ロボットが普及し，「ロボット三原則」によって人間とロボットは平和に共存している．しかし，人間に

表 1: 代表的な SF 映画に登場するロボット

作品名	ロボット名	役割	外観	三原則	主な特徴
A.I.	デイビッド	家族 (子供)	人間	不明	人間と同じ愛情を持つ
2001 年宇宙の旅	HAL	司令	無 (赤 LED)	不明	
リアル・スティール	アトム	格闘用	人型ロボット	不明	動きを真似るシャドー機能付
チャッピー	チャッピー	手下 (兵器)	人型ロボット	不明	高性能 AI を半分取り入れた攻撃ロボット
エクス・マキナ	エヴァ	試作品	ほぼ人間	不明	表情や会話の自然さはそれまでの AI をはるかに凌駕
オートマタ	クリオ	娼婦	人型ロボット	2 制御機能	「生命体への危害の禁止」「自他のロボットの修正 (改造) の禁止」実装
アンドリュー	アンドリュー	家事用	人型→人間	実装	創造性を持つ
アイ，ロボット	サニー	家事用	人型ロボット	本来は実装	サニーだけが三原則を破る?自我のある

危害を加えることができないはずのロボットによる殺人事件とロボットの反乱が描かれる．これは，“ロボット工学三原則には重大な欠陥（問題点）がある” ことを含んでおり，これを突き詰めると，“「人間」をどのように定義するか” という点となる．

そこで，“人間とは？” “どこまでが人間？” “どこからが人工物？” についてみてみる．映画『アンドリュー NDR114』（Bicentennial Man, 1999 年公開）は，ある一家に購入された家事用ロボットの NDR114（「アンドリュー」と名付られる）が，家族の一員として生活を営むうち，次第に人間への憧れを抱くようになり，200 年にわたって人間の心を育んでゆく過程が描かれている．この映画において，事故や病気で損傷した体の一部をアンドリューと同じ人工臓器で置き換えた人間も多い中，最終的に，血液の影響で老化の進んだ外観となったアンドリューは，人類法廷で，“死” を迎えることでようやく人間であることを認められる．この映画から，不死身の人工生物を目の当たりにする人類は，その生死観や倫理観に大きな影響を受けることが想像される．また，ロボットの心が，少なくとも発現する機能の面でヒトの心とほとんど同じであるということが確認され，一般に受け入れられるようになる必要がある．

一方で，映画『ブレードランナー』（Blade Runner, 1982 年公開）では，ロボットに代わる「レプリカント」と呼ばれる人造人間が描かれている．レプリカントは，見た目だけでは人間と区別がつかないという前提であり，その判定にチューリング・テストが用いられている．本来のチューリング・テストで判定するのは人間の心であって，姿形は関係ない．その点からすると，映画『ザ・マシーン』（The Machine, 2013 年公開）中では，独自の AI テストシーンが描かれている．ちなみに，人間のように何でもこなせる汎用人工知能（汎用 AI）の判定基準として,「コーヒーテスト」提案されている [ニュートン 18]．これは,「間取りを知らない家に上がってコーヒーを入れる」ということができれば汎用 AI だということである．Apple 社の創業者の一人であるスティーブ・ウォズニアック (Steve Wozniak) が “知らない家に上がってコーヒーを入れる機械を我々は決してつくれないだろう” と予言したことに由来することから,「ウォズニアック・テスト」とも呼ばれる．

こうした人間のような人工物（ロボット・AI）を創り上げるためには，こうした人工物に “心” を持たせることにも検討が必要である．“コンピュータに心を持たせることは可能か” という命題は，近年，ロボット工学，認知科学，情報科学など，様々な分野で盛んに議論されている．心とは，知的情報処理に加え，基本的な情動やより高度の感情，さらには価値観などの行動規範を含むものと考えられる．また，心とは,「知」「情」「意」「記憶と学習」「意識」という 5 つの働きを総合した情報処理機能であるといわれる [松本 96] が，本稿ではこれら 5 つを明確に分けることはせず，大まかに「情」もしくは「意識」と捉えることとする．既出の『アンドリュー NDR114』に出てくるアンドリューの他，映画『チャッピー』（CHAPPiE，2015 年公開）では，感情を持ったり，人間の知性を模倣する AI 搭載ロボット “チャッピー” が，インストール直後は純真無垢でまるで赤ん坊のようであったが，人間からの教育を受けて徐々に “人間的” なことを学習していく姿が描かれている．人間も赤ん坊のうちは動物を判別できないが，成長とともに学習すれば判別できるようになる．AI も人間のように学習することで成長する．

そしてさらに一歩踏み込んだ「情」として「恋愛感情」が描かれている作品もある．AI と人とのコミュニケーションが日々進化し，人間が AI に恋をするような未来が来るのかもしれない．映画『エクス・マキナ』（Ex Machina，2015 年公開）では，AI を搭載した女性型ロボットに主人公が日増しに惹かれていく様が描かれている．ただし実際には，主人公の一方的な好意と思われる．また，映画『her/世界でひとつの彼女』（Her，2013 年公開）は，人格を持つ最新の AI 型 OS をインストールし生活を共にすることで恋愛関係が生まれていくという物語である．この映画の OS は，一貫した音声認識 AI として明確なビジュアルは一切描かれていない．その状況で，声だけの AI 型 OS と男性が互いに惹かれている（OS 側の本当の気持ちについては定かではないが）．さらに，既出の『アンドリュー NDR114』では，人間の心や感触，痛覚などの神経系器官を手に入れ，限りなく人間に近い状態になったアンドリューが人間の女性と愛し合う．AI は，ユーザーの情報を豊富に持つことで，より身近な存在になる．コミュニケーション能力が発達した AI との濃密なコミュニケーションによって，人間は AI に対し恋のような特別な感情を抱くことはある意味，当然の流れなのかもしれない．そして，いわゆる「心」や「意識」を AI が持っていないにしても,（『her/世界でひとつの彼女』中の主人公のように）人間はおそらく誤認（勘違い?）することは容易に起こりうると考えられる．現在研究されている AI にも「意思」のように感じられる事例もあるが，それは人間による意思のエミュレーションに過ぎない．人間がなぜ個々に独立した意思や精神，感情を持つのかという「脳科学」的な視点から研究をおこなうことで，AI に意思を持たせられるかもし

れない．

ただし，「意思」を持ったロボットが出現し高度化することの問題点としては，古来 SF で取り上げられ続けたように，ヒトを脅かし支配するロボットの恐怖が挙げられる．本来，人工物（ロボット）は設計者が定義した範囲内での心を呈するのみであるから，例えば抑圧された結果犯罪に至るような心理発展のアルゴリズムを埋め込まなければ，そのような犯罪心理は生じ得ないはずである．しかし，レイ・カーツワイル (Ray Kurzweil) による予言を適用すれば，AI が自らを規定しているプログラムを自身で改良するようになると，永続的に指数関数的な進化を遂げる．この結果，ある時点で人間の知能を超えて，それ以降の発明などはすべて人間ではなく AI が担うようになり，それ以降の進歩を予測できなくなる (人間がいくら考えても想像ができないレベルに達する)．これがいずれは，“AI が世界を支配する” という域にまで到達すると考えれば，“ヒトを脅かし支配するロボットの恐怖” に繋がることになる．映画『2001 年宇宙の旅』（2001: A Space Odyssey，1968 年公開），映画『マトリックス』（The Matrix，1999 年公開），そして『アイ，ロボット』では，まさにこの状況が描かれ,『オートマタ』では，こうした “人間を超える” 事態に至らないよう，ロボットの自己修復禁止機能が組み込まれている（作中では，自己修復禁止を解除されたロボットが数体現れる）．

例えば，人工知能開発段階では，開発方向の判断において倫理という社会規範 (Social Norm) に依拠するところが大きい．方向性を誤らないためにも，倫理という規範 (Norm) について議論を重ねる必要がある．

「モラルジレンマ」の例として，哲学者や心理学者が道徳や倫理について考える時に用いられる例題「トロッコ問題（Trolley Problem）」が挙げられる．AI の世界で，まさにこの「モラルジレンマ」の問題を抱えている顕著な事例が自動走行車である．MIT の研究チームは,「自動運転車を用いた人工知能の道徳的な意思決定に関して，人間の視点を収集するためのプラットフォーム」をオンライン上に公開している [MIT 16]．この「トロッコ問題」に近い内容が『アイ，ロボット』中でも描かれている．作中では，交通事故の救助に現れたロボットが，少女より生存率の高かった主人公を優先して救い，少女を見捨てた（以来，主人公は数字で物事を判断するロボットを毛嫌いするようになった）という状況が描かれている．コンピューターがあらゆる情報をもち，一瞬で判断する，ということであれば，少女の環境，将来性などを勘案し，主人公と比べて社会的にどちらが有用か等から総合的に判断する必要性も考えられる．

また，AI を使って犯罪を予知する仕組みも考えられている [日経 18]．これは，街中に設置された防犯カメラや Web カメラの映像を用いて全市民を監視し，収集したデータをコンピュータが分析し，犯罪やテロなどの異常を事前に察知するというものである．映画『マイノリティ・リポート』（Minority Report，2002 年公開）では，事件を予想したのは「プリコグ」と呼ばれる予知能力者だが，現実でその役を担うことになるのは AI である．

実際，日本では 2018 年 1 月，神奈川県警が AI を使って事件・事故の発生を予測する新システムの導入を検討していると報じられた [産経 18]．この報道によると，連続発生した事件の容疑者が同一かどうかを分析したり，容疑者の次の行動を予測したりする他，事件事故が起きやすい時間帯と場所を確率で示すシステムの構築を目指すという．また海外でも，例えば 2017 年 8 月，米・シカゴ市警察がサウスサイド地区で犯罪予測システムを導入し,「凶悪事件が激減した」という成果を発表した [Engadget 18]．こうした「犯罪予知システム」については，どのようなアルゴリズムで判断を出しているのか，その精度はどれほどのものなのかが分からないままである．また，こうしたセキュリティ領域への活用が進む中で，議論の一つになるのがプライバシーの問題である．国家を脅かすテロや凶悪犯罪，または我々の身近に潜むサイバー犯罪などを検知するために，我々の個人情報をはじめ，インターネット上の検索・閲覧履歴や行動，サイトへのログイン履歴や所在地など，プライバシーに関わる情報が必要になる場合が考えられる．AI によって人間の分析能力を超えた莫大なデータを，過去何十年にもわたって解析可能となった未来において，我々は自分たち自身のプライバシーを守ることができるのか? そのような未来において，人々の安全を守るために，個々のプライバシーは犠牲にされるべきなのか? 「安全保障 or 市民のプライバシー」という究極の選択に，AI はどのように対処していくべきなのか?といった観点について考えていく必要がある．

4. おわりに

最近の AI 技術の急速な発展により，AI を搭載した高度なロボットを如何に社会的が受容するかが現実的かつ大きな課題として浮上してきている．AI を搭載したロボットは，自身で学習し，ときには人間の知性の及ばない思考や判断すら可能といわれる．すなわち，近い将来，AI が人間の能力を超える “シンギュラリティ” が実現すると予言されている．しかしながら，現在のロボット開発では技術力の向上だけが先行し，ロボットが人間社会に入り込んだときの影響が考慮されていない，という問題点も指摘されている．そうしたロボットが人間にとって身近な存在になるだけではなく，人間の役割を代替し，人間がそれに大きく依存していくことになれば，これまでの人間中心の社会のあり方に再考が迫られることになる可能性が考えられる．例えば，癒しロボットとよばれる範疇のペット型ロボットが安価に購入できるようになり，人間に心理的影響を及ぼすようになってきている．逆にそれらの使い道として，心のケアやリラクゼーション，老人の機能回復なども考えられている．今後これらの他に別な形態のロボットが人間生活に深く入り込んでくる可能性は十分に考えられる．しかしながら，それらが人間，環境，社会にどの様な影響を及ぼすかについて，十分な議論がおこなわれているとはいえない．我々は，どういう社会を作りたいのか?ロボットが人間生活に浸透する社会を見越し，人間とロボットの両方が生活しやすい環境とは何か，社会規則を含めた社会システム・インフラをどのようにすべきか，災害などの非日常という特殊環境におけるロボットの役割などを工学分野のみならず，心理学，医学などの異種分野からより良い解決策を追求することにより，技術だけで進められるロボット開発に方向性を与えることができ，ロボット開発の発展に寄与できると考えられる．

そこで本稿では，人間と AI が共存する未来の社会を，どのような観点から，どのようにデザインしていくことができるかを検討する素材として,「ロボット工学三原則」，“人間とは?” “どこまでが人間?” “どこからが人工物?”，人工物に “心”（情，意識），そして倫理的問題について，該当する（その状況を具体的に描いている）シーンを含む SF 映画を紹介した．こうした観点からの AI，ロボットデザインを検討することは，HCD としても大きな意義があると考える．

人間に近いロボットは，ロボット自身が自律的に判断し行動する点が車などの機械とは異なる．このため，ロボットと共生する上での倫理的および法的な基準を設けることも重要な課題である．さらには，そもそも心を持つように見える AI を作っ

て良いのか，AI が知財を生み出す場合の権利，人工知能システムが社会に広がったときの不具合の問題，倫理や社会制度の議論が必要になる．これらの点が曖昧なまま人間に近いロボットが人間社会に展開されていった場合，我々の日常に混乱が起きることは容易に想像できる．シンギュラリティを待望もしくは警戒する一方で，AI とどのように共存していくかという議論はいまだ十分とはいえない．

黒崎は，著書 [黒崎 98] の中で『ロボットは主体にはなりません。あくまで、それらは人間が作ったものであって、もし倫理が必要であるとすれば、ロボットを製造する倫理、ロボットを使う倫理であって、決してロボットの倫理ではありません。倫理が必要なのは、〈作り・使う〉人間の側です。』と述べている．一般的に，AI を含む情報システムの設計では，設計者が何を問題設定しているかが重要である．さらには，基本的にその設計者の倫理観，よかれと思うことが拠り所になっている．もちろん人間中心設計（Human Centered Design）において，利用者が果たしてどんな文脈で，どういうふうに絞ればわかるだろうか，をユーザー視点で考え，最善のものを作ろうとはしている．あるサービス仕様は，顧客，ユーザーにとって正しいかもしれないが，それではスポンサーがつかなくなるかもしれない．するとサステナブルに実運用できなくなくなり，結局は継続できなくなってしまい，ユーザーに迷惑をかけてしまうかもしれない．こうした意味から，実運用とのバランスは考えなければいけない．ただし，そこで重要なのは，できる限りユーザーのためにフェアであろうとする努力目標を持つことである．

参考文献

[ニュートン 18] ニュートンプレス社: Newton 2018 年 6 月号特集「人工知能は人より賢くなれるか」, 第 38 巻, 第 6 号, pp.44-57 (2018)

[松本 96] 松本元: 脳・心・コンピュータ, 丸善 (1996)

[MIT 16] MIT Media Lab.: モラル・マシン (2016) http://moralmachine.mit.edu/hl/ja (最終アクセス 2018.11.13)

[日経 18] 日本経済新聞社: 明日の安心・安全 AI で高度監視, 2018 年 9 月 13 日, 日経新聞朝刊 14 面 (2018)

[産経 18] 産経新聞社: 神奈川県警犯罪発生 AI で予測東京五輪までの試用目指す, 2018 年 1 月 29 日, 産経新聞東京朝刊 27 面 (2018)

[Engadget 18] https://japanese.engadget.com/2017/08/11/hunchlab/ (最終アクセス 2018.11.13)

[黒崎 98] 黒崎政男: となりのアンドロイド – 哲学者クロサキの憂鬱, NHK 出版 (1998)

参考映画 (タイトル五十音順)

『2001 年宇宙の旅』(*2001: A Space Odyssey*)
監督：スタンリー・キューブリック
公開年：1968 年
製作国：イギリス/アメリカ
配給：メトロ・ゴールドウィン・メイヤー

『アイ，ロボット』(*I, Robot*)
監督：アレックス・プロヤス
公開年：2004 年
製作国：アメリカ
配給：20 世紀フォックス

『アンドリュー NDR114』(*Bicentennial Man*)
監督：クリス・コロンバス
公開年：1999 年
製作国：アメリカ
配給：ソニー・ピクチャーズ エンタテインメント

『エクス・マキナ』(*Ex Machina*)
監督：アレックス・ガーランド
公開年：2015 年
製作国：イギリス
配給：ユニバーサル映画

『オートマタ』(*Automata*)
監督：ガベ・イバニェス
公開年：2014 年
製作国：スペイン・ブルガリア合作
配給：松竹

『ザ・マシーン』(*The Machine*)
監督：カラドッグ・W・ジェームズ
公開年：2013 年
製作国：イギリス
配給：熱帯美術館

『チャッピー』(*CHAPPiE*)
監督：ニール・ブロムカンプ
公開年：2015 年
製作国：アメリカ
配給：ソニー・ピクチャーズエンタテインメント

『her/ 世界でひとつの彼女』(*Her*)
監督：スパイク・ジョーンズ
公開年：2013 年
製作国：アメリカ
配給：ワーナー・ブラザーズ映画（アメリカ）/アスミック・エース（日本）

『ブレードランナー』(*Blade Runner*)
監督：リドリー・スコット
公開年：1982 年
製作国：アメリカ・香港合作
配給：ワーナー・ブラザーズ映画（アメリカ）

『マイノリティ・リポート』(*Minority Report*)
監督：スティーブン・スピルバーグ
公開日：2002 年
製作国：アメリカ
配給：20 世紀フォックス映画

『マトリックス』(*The Matrix*)
監督：ラリー・ウォシャウスキー/アンディ・ウォシャウスキー
公開日：1999 年
製作国：アメリカ
配給：ワーナー・ブラザース

『メトロポリス』(*Metropolis*)
監督：フリッツ・ラング
公開日：1927 年
製作国：ドイツ
配給：UFA（ドイツ）/パラマウント映画（アメリカ）/松竹

『リアル・スティール』(*REAL STEEL*)
監督：ショーン・レヴィ
公開日：2011 年
製作国：アメリカ
配給：タッチストーン・ピクチャーズ（アメリカ）/ウォルト・ディズニー・ピクチャーズ（アメリカ）

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

[4E2-OS-7a] AIの法学の応用(1)

佐藤 健（国立情報学研究所）、新田 克己（国立情報学研究所）

Fri. Jun 7, 2019 12:00 PM - 1:40 PM Room E (301A Medium meeting room)

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

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

12:00 PM - 12:40 PM

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

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

12:40 PM - 1:00 PM

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

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

1:00 PM - 1:20 PM

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

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

1:20 PM - 1:40 PM

©The Japanese Society for Artificial Intelligence

12:00 PM - 12:40 PM (Fri. Jun 7, 2019 12:00 PM - 1:40 PM Room E)

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

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

Keywords: support of criminal investigation, offender profiling, machine learning

犯人が不明である未解決事件の早期解決を目的として、1970年代米国 FBIより登場した犯罪者プロファイリング、中でも犯人像推定の方法論は精神医学や臨床心理学の知識を基にした分析手法から、1980年代には多変量データ解析を駆使したものと変遷し、今現在の主流となっている。そのような中、近年は我が国においても機械学習を応用した犯人像推定研究が始まる兆しがみられている。また、講演者においては、テキストマイニング技術を用いて、文章情報から犯人像を推定するといった研究も行っている。

本セッションでは、従来の手法の変遷と現在の犯罪者プロファイリングの紹介、そして今後の展望について議論していくこととする。

©The Japanese Society for Artificial Intelligence

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

法律の要約のためのランダムフォレストを用いた重要文抽出

Extracting Important Sentences with Random Forest for Statute Summarization

小川 泰弘 *1*2 OGAWA, Yasuhiro
佐藤 充晃 *2 SATOU, Michiaki
駒水 孝裕 *1*2 KOMAMIZU, Takahiro
外山 勝彦 *1*2 TOYAMA, Katsuhiko

*1名古屋大学 情報基盤センター
Information Technology Center, Nagoya University
*2名古屋大学大学院情報学研究科
Graduate School of Informatics, Nagoya University

Our purpose is to provide an automatic summarization for Japanese acts and we propose a sententence extraction method with Random Forest. While the traditional automatic summarization methods have used the information of summarizing source data, in recent years, the methods based on machine learning use the summarization results. However, in such a method, the amount of learning corpus is small, especially in Japanese text. In this research, we solve this problem by using "Outlines of Japanese Statutes," which are official summaries of statutes published by the Japanese government. Furthermore, we show that the sentence extraction method with Random Forest has higher performance rather than with decision trees or with support vector machines.

1. はじめに

法令は社会の設計図であり，法令を理解することにより，権利や義務などの国民の生活に関する多くの情報を得ることができる．しかし，法令は，1 文が長かったり，構文が複雑であったりするため，一般に読みにくく，理解しにくい文書であるとされている．また細部を規定する条文などもあるため，文書全体の分量が多い場合があり，その内容の把握は容易ではない．

そこで本研究では，法令理解の促進に貢献するために，法令の要約を提供することを目標とする．自動要約については，ルールに基づく手法や，学習データを利用する機械学習，最近ではニューラルネットワークを用いた手法も提案されている．

機械学習では，学習データが必要になる．当初は人手によって要約が作成されてきたが，その作成コストの大きさから大量のデータは用意できなかった．近年では，新聞記事の見出しを要約と見做すことにより大規模なコーパスが作成されている．

法令の自動要約についても，こうした学習データが必要となるが，幸いなことに，法令に関しては「法令のあらまし」と呼ばれる文書が，昭和 48 年 4 月から政府により作成されている．ここで新規制定法律に注目すると，平成 31 年 1 月までに 1,156 件のあらましが作成されており，従来研究よりも多くの文書がコーパスとして利用可能である．

そこで本研究では，まず法令のあらましを利用して要約用の学習データを構築し，次に，それを用いた機械学習による自動要約に取り組む．元の法令と法令のあらましを比較すると，文字数や文数の制約がなく，法令の 1 文がそのまま取り出されていることが多い．そこで，分類器を用いた重要文抽出により自動要約を実現する．

重要文抽出に分類器を用いた先行研究としては，決定木を用いたもの [1, 2] や，SVM を用いたもの [3] があるが，本研究では，それらよりも高い分類性能が期待できるランダムフォレストを用いて，その有効性を検証する．なお，予備実験としてニューラルネットワークを用いる手法も試したが，学習データの量が不十分であったためか，充分な性能は得られなかった．

本論文の貢献は以下の通りである．

- 従来よりもサイズの大きな要約用コーパスの作成
- 重要文抽出におけるランダムフォレストの有効性の確認

連絡先: 小川泰弘，名古屋大学，yasuhiro@i.nagoya-u.ac.jp

2. 関連研究

自動要約については，文短縮 [4, 5]，重要文抽出 [6, 7, 8, 9, 10, 11, 1, 2, 3]，文融合 [12] などの手法が研究されてきている．

文短縮は，1 文を短くする要約手法であり，構文木を枝刈りする手法 [4] などが提案されてきた．重要文抽出は，原文書から重要な文を抽出する要約手法であり，何らかの素性を基に原文書中の文に重要度を付与し，重要度の高い順に文を抽出する手法が提案されてきた．具体的には，文の位置情報を利用する手法 [6] や，文間の類似度を利用する手法 [7]，それらに加えて重要語など複数の素性を同時に利用する手法 [8, 9] などがある．文融合は，複数の文を 1 文に融合する要約手法であり，構文木の情報を利用する手法 [12] などが提案されてきた．

ここまで紹介した手法は学習データが不要であるが，近年，機械学習などの学習データを利用した要約手法が提案されている．特に重要文抽出においては，ベイズ分類器を用いる手法 [10] や，文の重要度を重回帰分析で決定する手法 [11]，分類器である決定木を用いる手法 [1, 2] や，分類器であるサポートベクターマシン (SVM) を用いる手法 [3] などが提案されている．さらに現在は，ニューラルネットワークを用いた手法 [13] などが提案されている．

3. 法令のあらまし

法令のあらましは，国民の法令理解の促進を目的とし，昭和 48 年 4 月から官報に掲載されている．あらまし作成の対象となるのは，法令のうち，法律・政令・条約であるが，本稿では法律を対象とする．

ここで，政府があらましを作成するのであれば，法律の要約は不要と思われるかもしれないが，そうではない．まず，あらましは昭和 48 年 4 月以前に公布された法律には存在しない．平成 30 年 10 月 1 日時点で有効な法律は 2,261 件あり，これは前述のあらましが存在する新規制定法律 1,156 件の約 2 倍である．しかも，法律はしばしば改正される．実際，上記の 1,156 件中，840 件以上が改正されており，そうした法律のあらましは，現在の法律の要約になっていない可能性がある．よって，現在有効な法律の概要を知るために，自動要約は必要である．

本節では，法令のあらましに対する分析と，それに基づく要約用の学習データの作成について述べる．

法律	あらまし
‥(略)‥	‥(略)‥
(基本方針) 第四条 主務大臣は、有明海及び八代海の海域の特性に応じた当該海域の環境の保全及び改善並びに当該海域における水産資源の回復等による漁業の振興に関する施策を推進するため、有明海及び八代海の再生に関する基本方針(以下「基本方針」という。)を定めなければならない。 2 基本方針においては、次に掲げる事項を定めるものとする。 一 有明海及び八代海の海域の環境の保全及び改善並びに当該海域における水産資源の回復等による漁業の振興に関する基本的な指針 二 次条第一項の県計画の策定に関する基本的な事項 3 主務大臣は、基本方針を定めようとするときは、あらかじめ、関係県の意見を聴くとともに、関係行政機関の長に協議しなければならない。 4 主務大臣は、基本方針を定めたときは、遅滞なく、これを公表するとともに、関係県に通知しなければならない。 5 主務大臣は、情勢の推移により必要が生じたときは、基本方針を変更するものとする。 6 第三項及び第四項の規定は、基本方針の変更について準用する。 第五条 関係県は、基本方針に基づき、‥‥を定めるものとする。	3 基本方針 (一) 主務大臣は、有明海及び八代海の海域の特性に応じた当該海域の環境の保全及び改善並びに当該海域における水産資源の回復等による漁業の振興に関する施策を推進するため、有明海及び八代海の再生に関する基本方針(以下「基本方針」という。)を定めることとした。**(第四条第一項関係)** (二) 基本方針においては、有明海及び八代海の海域の環境の保全及び改善並びに当該海域における水産資源の回復等による漁業の振興に関する基本的な指針並びに4の(一)の県計画の策定に関する基本的な事項を定めることとした。**(第四条第二項関係)** (三) 主務大臣は、基本方針を定めようとするときは、あらかじめ、関係県の意見を聴くとともに、関係行政機関の長に協議しなければならないこととした。**(第四条第三項関係)** 4 県計画 (一) 関係県は、基本方針に基づき、当該関係県の区域内の指定地域について、有明海及び八代海の海域の特性に応じた当該海域の環境の保全及び改善並びに当該海域における水産資源の回復等による漁業の振興に関し実施すべき施策に関する計画(以下「県計画」という。)を定めることとした。**(第五条第一項関係)**
‥(略)‥	‥(略)‥

図 1: 法律とそのあらましの対応例（有明海及び八代海を再生するための特別措置に関する法律，平成 14 年法律第 120 号）

3.1 法令のあらましの分析

法令のあらましが，どのように作られているかを分析する．まず，法律とそのあらましの対応の例を図 1 に示す．この図において，色付けした法律文とあらまし文が対応している．対応する文を見ると，図 1 中の橙色や青色の箇所の文は，法律とあらましでほぼ同じである *1．実際，5.1 節の実験用に作成した学習データ中の正例（あらまし文が対応する法律文）8,793 文のうち，約半数の 4,421 文において，法律文とあらまし文がほぼ同一であった．このことから，法律のあらましの作成においては，重要文を抽出することが最初の課題となることが分かる．もちろん，図 1 中の緑色の箇所が示すように，法律文 3 文が一つのあらまし文に融合される場合もある．しかし，そうした文融合の手法を適用する場合でも，まずは融合の対象となる文を抽出する必要がある．以上のことから，本稿では，法律の自動要約のための重要文抽出を目的とする．

次に，重要文抽出のためにどのような手法を取るかを検討する．その際，法律の自動要約には，従来の自動要約の研究とは大きく異なる点がある．それは，従来の研究の多くでは，生成する要約文に文字数や文数の制約があったが，法令のあらましでは，そのような制約は観測されない点である．表 1 は，本稿の実験に使用するテストデータである．各法律に含まれる文数と，あらまし用に抽出された文数から，抽出率を計算している．この表に示すように，あらましにおける抽出文数・抽出率はばらつきが大きく，全文を抽出する場合もあれば，平成 26 年法律第 49 号のように 2 割しか抽出しない場合もある．平均は約 70%であり，従来の自動要約と比べて抽出率は大きい．以上を踏まえ，本研究では分類器を用いた重要文抽出を試みる．

3.2 学習データの構築

分類器を重要文抽出に用いるためには，重要文として抽出すべきか否かを正例・負例としてラベル付けした文からなる学習データが必要となる．そのためには，あらましの各文がどの法律文と対応するかを判定しなければならない．幸いなことに，法令のあらましにおいては，図 1 中で赤字で示した「(第四条第一項関係)」のように，どの法律文を要約したかを示す関係規定 [14] と呼ばれる記述が存在する．

本研究では，この関係規定の情報を利用して，法律文とあらまし文の対応付けを自動的に行った．図 1 の色分けがその結果である．ここで，図 1 において橙色および青色の箇所は，法律文 1 文があらまし文 1 文に要約されている．よって，これらの法律文を正例として学習データに加えた．一方，法律文において色付けされていない文は，対応する文があらましに出現していない．すなわち重要文として抽出されなかったことを示す．そこで，こうした法律文を負例として学習データに加えた．なお，図 1 において緑色の箇所は，法律 3 文があらまし文 1 文と対応している．このように複数文の法律文があらまし文 1 文と対応する場合，複数文がまとめて 1 文に要約されている場合もあれば，実際には複数文中の 1 文だけが要約に使われている場合もある．つまり，こうした法律文は関係規定の情報だけでは正例になるか負例になるか判定できない．よって，今回は学習データには加えなかった．

4. ランダムフォレストを用いた重要文抽出

本研究では，ランダムフォレスト [15] を用いた重要文抽出方法を提案する．ランダムフォレストは，アンサンブル学習の一種であり，複数の決定木から構成される．

ランダムフォレストの学習に用いる素性として，予備実験では，文の出現位置などの文書構造の情報も用いたが，性能向上が見られなかった．よって本研究では，単純に文中での単語の出現の有無を用いた．ただし，すべての単語を用いると素性の数が多くなるため，名詞に限定し，その上で，あらまし文中での出現回数が 500 位以上のものを用いた．これは，要約の結果としてあらましに残った単語が，文抽出の手がかりになると

*1 厳密には，文末に「こととした」を付加したり，「第四条第一項」を「4 の (一)」と表記するなどの，あらまし特有の書式に合わせた変換はあるが，それ以外は同一である．

考えたためである．

5. 実験

提案手法の有効性を検証するために，提案手法で用いるランダムフォレストと，それ以外の分類器，さらに学習データを利用しない重要文抽出手法との比較実験を行った．

5.1 学習データとテストデータ

実験では，昭和 48 年 4 月から平成 31 年 1 月までに新規制定された 1,156 法律のうち，1,111 件から学習データを構築した．この 1,111 件には法律文 104,102 文が含まれているが，3.2 節で述べた手法に基づいて正例 8,793 文，負例 15,507 文を取り出し，合計 24,300 文からなる学習データを作成した．これは TSC-1 [16] における重要文抽出タスクで使用されたデータセット 4,521 文の 5 倍以上のサイズである．ただし，正例とも負例とも判定できなかった文が 77%残っており，こうしたデータを利用して学習データを増やすことは，今後の課題である．

テストデータには，残りの法律から 33 件を選んで [*2]，人手で正解ラベルを付与した．実際には，ほとんどのラベルは自動で付与できたが，図 1 の緑色の箇所のような場合は，人手で判定した．作成したテストデータの詳細を表 1 に示す．抽出率のマクロ平均は 74%，マイクロ平均は 70%であった．

5.2 比較手法

提案手法と比較するために，先行研究で用いられた決定木 [2] 及び SVM [3] を用いた．さらに学習データを利用しない手法であるリード法 [6]，TextRank [7]，TextTeaser [8] も用いた．

決定木: 決定木は学習データを素性ベクトルに基づき部分集合に分割し，学習データを分類する木構造を構築する．決定木の学習アルゴリズムには CART 法を用いた．

SVM: SVM は，クラスごとに学習データを分離する境界を求める．本実験では [3] を参考に，線形な境界によってデータを分離する線形 SVM (SVM_Lin) と，カーネル関数に多項式カーネルを用いた SVM (SVM_Poly) を用いた．

リード法: リード法 [6] は，文書の先頭から順に文を抽出する手法である．新聞要約などで良い性能を示しており，自動要約におけるベースラインとして使われている．本実験では，法律文書の先頭をそのまま抽出する「リード法 (Original)」と，法律が条単位で構成されることに注目し，各条の先頭を抽出する「リード法 (Article)」を実施した．

TextRank: TextRank [7] は，グラフに基づく重要文抽出手法であり，文をノードとして文書をグラフ構造で表現する．多くの文と類似している文は重要な文であるという仮定があり，グラフのエッジに文間の類似度を付与し，そのグラフから各文に重要度を付与する．本実験では，GitHub 上で公開されている実装 [*3] を用いた．

TextTeaser: TextTeaser [8] は，特徴に基づく重要文抽出手法であり，文の長さや文の位置を特徴として，各文に重要度を付与する．本実験では，PyTeaser [9][*4] を用いた．

5.3 実験設定

ランダムフォレスト・決定木・SVM の学習には scikit–learn (v.0.19.1) を用いた．ランダムフォレストのデータ分割基準にはジニ係数を用い，決定木の数と深さについては 5 分割交差検証によって最適な値を求めた．決定木と SVM のパラメータについても，5 分割交差検証によって最適値を求めた．今回の学習データは，負例が正例の約 2 倍となる不均衡データである．予備実験の結果にもとづき，アンダーサンプリングを適用して負例のデータ数を正例の数に合わせた．

法律文およびあらまし文は，MeCab (IPA 辞書) を用いて形態素解析し，単語に分割した．素性に用いる単語には，学習データの作成に用いた法律 1,111 件に対応するあらましの形態素から取得した．ただし，解析結果が未知語となるものは素性に用いなかった．

なお，リード法・TextRank・TextTeaser は，抽出文数の設定が必要となる．本実験では，テストコーパスにおける抽出率のマクロ平均である 74% をもとに抽出文数を設定した．

表 1: テストデータにおける文数と抽出率

法律番号	文数	抽出文数	抽出率
昭和 59 年法律第 34 号	52	34	65%
平成 元 年法律第 70 号	20	11	55%
平成 12 年法律第 75 号	18	6	33%
平成 16 年法律第 166 号	61	20	33%
平成 23 年法律第 99 号	22	22	100%
平成 24 年法律第 50 号	40	38	95%
平成 25 年法律第 64 号	37	36	97%
平成 25 年法律第 66 号	10	8	80%
平成 25 年法律第 86 号	18	17	94%
平成 25 年法律第 110 号	30	15	50%
平成 25 年法律第 112 号	109	93	85%
平成 26 年法律第 17 号	46	23	50%
平成 26 年法律第 49 号	53	11	21%
平成 26 年法律第 57 号	24	21	88%
平成 26 年法律第 98 号	42	42	100%
平成 26 年法律第 100 号	28	24	86%
平成 26 年法律第 101 号	20	10	50%
平成 26 年法律第 126 号	17	17	100%
平成 26 年法律第 137 号	44	42	95%
平成 27 年法律第 14 号	35	23	66%
平成 27 年法律第 62 号	14	10	71%
平成 27 年法律第 69 号	20	20	100%
平成 28 年法律第 33 号	55	21	38%
平成 28 年法律第 68 号	11	11	100%
平成 28 年法律第 74 号	29	17	59%
平成 28 年法律第 105 号	42	29	69%
平成 28 年法律第 111 号	33	32	97%
平成 28 年法律第 112 号	31	19	61%
平成 28 年法律第 115 号	48	25	52%
平成 30 年法律第 28 号	12	12	100%
平成 30 年法律第 47 号	36	26	72%
平成 30 年法律第 48 号	35	26	74%
平成 30 年法律第 100 号	29	28	97%
合計	1121	789	70%

5.4 実験結果

本実験においては，抽出性能を再現率・精度・F 値で，要約性能を ROUGE で評価した．その結果を表 2 に示す．なお，ROUGE については，ROUGE-1・ROUGE-2・ROUGE-L を計算したが，いずれも同様の傾向が見られたため，本稿では

*2 残りの法律 12 件については，法律文とあらまし文の対応付けが困難なものが含まれており，人手で正解ラベルを付与できなかった．

*3 https://github.com/summanlp/textrank/

*4 http://xiaoxu193.github.io/PyTeaser/

表 2: 重要文抽出実験の評価結果

システム名	再現率	精度	F 値	ROUGE-1
ランダムフォレスト	84.7%	81.9%	81.0%	88.4%
TextRank	79.2%	80.0%	77.4%	88.0%
SVM_Lin	77.7%	82.5%	77.4%	85.1%
SVM_Poly	77.4%	82.1%	77.7%	84.5%
リード法 (Article)	76.7%	78.0%	74.8%	81.9%
決定木	73.4%	80.3%	74.7%	80.1%
TextTeaser	73.0%	74.2%	71.3%	68.4%
リード法 (Original)	70.9%	72.9%	69.7%	77.0%

ROUGE-1 の値のみを示す．これらの値は，テストコーパスである法律 33 件をそれぞれ入力し，得られた評価結果のマクロ平均である．ランダムサンプリングにおける偏りを防ぐため，ランダムフォレスト・決定木・SVM においては，分類器のモデルを 50 個作成し，それらのマクロ平均を求めた．

再現率は，ランダムフォレストがもっとも高く，有意水準 5%ですべての比較手法より優れていた．ランダムフォレストの次に TextRank が高い再現率を示した．精度は，SVM_Lin がもっとも高く，SVM_Poly もランダムフォレストを上回った．F 値は，ランダムフォレストが最高で，有意水準 5%で TextRank 以外のすべての比較手法より優れていた．ROUGE においては，ランダムフォレスト，TextRank の順に高かった．

5.5 考察

提案手法であるランダムフォレストを用いた方法が，精度以外の指標で最高の性能を示した．精度で SVM に劣った原因を調べたところ，SVM はランダムフォレストよりも抽出する文数が少ないことが分かった．逆に言えば，その点が SVM の再現率の低さの原因である．要約においては再現率が重視されることから，SVM より 7%多く文を抽出しているにも関わらず，SVM に近い精度を示すランダムフォレストが，もっとも有用であると言える．ROUGE の結果も，この点を示唆している．

リード法 (Article) がリード法 (Original) の性能を上回ったことは，法律の要約では重要文が先頭以外にも出現することを示唆している．ただし，リード法 (Article) の評価値は TextRank よりさらに低い．これは，法律の要約では文の位置よりも，文中の単語が重視されていることを示唆している．文の位置を用いる TextTeaser も，同様に低い性能であった．TextRank はリード法や TextTeaser と異なり，文の位置を考慮せず，単語の重要度を考慮するため，高い評価値を示したと考えられる．

6. おわりに

本研究では，法律の要約のためにランダムフォレストを用いた重要文抽出手法を提案した．また，その際に従来よりもサイズの大きな要約用のコーパスを作成した．ただし，このコーパスにおいては，元となったデータの 23%しか利用していない．よって，今後の課題としては，学習コーパスの量を増やすことと，それにより抽出性能を向上させることが挙げられる．

謝辞

本研究は JSPS 科研費 17K00460 の助成を受けたものです．

参考文献

[1] 野本 忠司, 松本 裕治: 人間の重要文判定に基づいた自動要約の試み, 情処研報, NL–120–11, pp. 71–76 (1997).

[2] 奥村 学, 原口 良胤, 望月 源: 決定木学習を用いたテキスト自動要約手法に関するいくつかの考察, 情報処理学会第 59 回全国大会講演論文集, 5N–2, pp. 393–394 (1999).

[3] 平尾 努, 磯崎 秀樹, 前田 英作, 松本 裕治: Support Vector Machine を用いた重要文抽出法, 情報処理学会論文誌, Vol. 40, No. 8, pp. 2230–2243 (2003).

[4] Jing, H.: Sentence Reduction for Automatic Text Summarization, In ANLP 2000, pp. 310–315 (2000).

[5] Knight, K., Marcu, D.: Summarization beyond Sentence Extraction: A Probabilistic Approach to Sentence Compression, Artificial Intelligence, Vol. 139, No. 1, pp. 91–107 (2002).

[6] Brandow, R., Mitze, K., Rau, L. F.: Automatic Condensation of Electronic Publications by Sentence Selections, Information Processing and Management, Vol. 31, No. 5, pp. 675–685 (1995).

[7] Mihalcea, R., Tarau, P.: Textrank: Bringing Order into Texts, In EMNLP 2004, pp. 404–411 (2004).

[8] Jagadeesh, J., Pingali, P., Varma, V.: Sentence Extraction based Single Document Summarization, Workshop on Document Summarization (2005).

[9] Gunawan, D., Pasaribu, A., Rahmat, R. F., Budiarto, R.: Automatic Text Summarization for Indonesian Language Using TextTeaser, In IOP Conference Series: Materials Science and Engineering, Vol. 190, No. 1, p. 012048 (2017).

[10] Kupiec, J., Pedersen, J., Chen, F.: A Trainable Document Summarizer, In SIGIR 1995, pp. 68–73 (1995).

[11] Watanabe, H.: A Method for Abstracting Newspaper Articles by Using Surface Clues, In COLING 1996, Vol. 2, pp. 974–979 (1996).

[12] Barzilay, B., McKeown, K. R.: Sentence Fusion for Multidocument News Summarization, Computational Linguistics, Vol. 31, No. 3, pp. 297–328 (2005).

[13] Rush, A. M., Chopra, S., Weston, J.: A Neural Attention Model for Abstractive Sentence Summarization, In EMNLP 2015, pp. 379–389 (2015).

[14] Kawachi, G., Ogawa, Y., Nakamura, M., Ohno, T., Toyama, K.: Daily News on Japanese Legislation toward Global Sharing of Japanese Legal Information, Journal of Open Access to Law, Vol. 3, No. 1, 19 pages (2015).

[15] Breiman, L.: Random Forests, Machine Learning, Vol. 24, No. 1, pp. 5–32 (2001).

[16] Fukushima, T., Okumura, M.: Text Summarization Challenge Text Summarization Evaluation in Japan, In NAACL 2001, Workshop on Automatic Summarization, pp. 51–59 (2001).

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

AIによる立法支援システム
Legislation Supporting Systems by AI

角田 篤泰[*1]
Kakuta, Tokuyasu

[*1] 中央大学 研究開発機構
Research and Development Initiative, Chuo University

This paper reports our legislation supporting system adopting AI technology researched and developed by the author. The system is extended from "eLen" system as a result of our previous research. "eLen" system has been provided as a web database system including about 90% regulations of all of local governments in Japan. Furthermore, the system can generate comparing tables among similar regulations and find clusters of similar regulations. In this research, we have implemented a new function for automatically generating templates of regulation based on the clusters. This report shows the outline of the function and the short history of the systems.

1. はじめに

本稿では，筆者らがいくつかのプロジェクトを通じて，AI 技術を用いて実施してきた立法支援システムの研究・開発の経緯や現状について報告する．

2009 年の人工知能学会全国大会において，筆者は立法過程とソフトウェア開発過程が類似していることを示し[角田 2009]，それ以降，様々な立法支援システムの研究・開発と実際の立法の現場へのシステム提供や技術提供を行ってきた[角田 2012,角田 2014,角田 2018b]．

立法過程においても，ソフトウェア開発の手法と同様に，ウォータフォール型で立法する場合や，過去の資産を再利用して立法する場合，そしてそれらを合わせた場合などがある．ウォータフォール型については，Python 表記を用いて条例を記述する研究などが進行中である[角田 2018a]．一方，過去の資産を再利用する方式については，類似条例を自動検出し，比較表などを自動作成するシステムが『立法支援機能付き eLen 条例データベースシステム』(以下「eLen システム」と略す)として，2012 年より全国の自治体に実際に提供され，今日まで多くの自治体で利用されている．このシステムをベースに新たな機能の開発も進んでいる．この他，政府が一般公開している日本法令の Web データベースシステムの『e-LAWS』において，政府の各省庁内では，2016 年より，そのデータベース更新時に自動的に改正法案を合成するシステムが稼働しているが，その改正法案を作成するための膨大な法制執務ルールという，条文表記のルールを計算機が扱えるように筆者が手作業で整理し，政府に提供した[角田 2016]．

本稿では，紙面や時間の制約のため，これらのうち，2018 年度の人工知能学会全国大会のオーガナイズドセッション「CREST 人工知能」にて報告した「AI 技術を用いた法的文書作成支援」のプロジェクトの最新の進捗について報告することにする．これは，上記の eLen システムにおける類似条例のクラスタリング機能を改良し，さらに，そのクラスタリング結果を用いて条例案のテンプレートを自動合成するというものである．

連絡先：角田篤泰，中央大学研究開発機構，〒112-8551 東京都文京区春日 1-13-27，kaku@nagoya-u.jp

2. eLen システムの概要

元々日本では全国に 1790 団体ある自治体の条例データは集積されておらず，自治体ごとにそれぞれ Web サーバを立ち上げ，その中で各自治体の例規(条例＋規則)情報を自治体ごとの方式で提示していた．多くの自治体ではクローリング技術などで自動的に集められるような仕組みにはなっておらず，中には PDF 形式であったり，地元企業の独自性の強いデータベースシステムで提供されたりしている自治体もあった．そこで，個別に対応しながら，それらの例規を収集して，1 つのデータベースとして構築し直し，Web を通して各自治体に利用できるように 2012 年より eLen システムとして実現して[角田 2014]，全国の自治体向けに提供を始め[1]，今日に至っている[2]．なお，eLen の画面例は図1の通り．

このような立法支援システムの研究・開発のために全国の自治体を回り，要望を集めていた際に，どこの自治体でも，高度な支援システムよりも，まずは自治体横断検索ができる例規データベースが欲しいという要望が多かったため，そのようなデータベース提供を研究・開発の手始めとした．現在では，全国の約 9 割の 1670 団体の例規，約 130 万本(約 1490 万条)を集め，一旦そのすべてのデータを政府標準の XML 形式に整形し[3]，法制執務[4]の例規表示形式で提供している．

[1] 日本経済新聞全国版(2012 年 08 月 16 日)「条例作成支援の検索システム」，読売新聞全国版(2012 年 10 月 30 日)「全国の条例データベース化して公開…名古屋大」を始め，多くのマスコミ報道あり．

[2] 2012 年より名古屋大学から公開を開始し，2017 年より鹿児島大学に運営を移し『全国条例データベース powered by eLen』として公開している．簡易版は一般公開もしている．URL は <https://elen.ls.kagoshima-u.ac.jp/> である．

[3] 政府の e-LAWS で採用されている XML スキーマとコンパチブルな形式である．詳しくは <http://elaws.e-gov.go.jp/search/html/readme.html#schema> を参照のこと．なお，e-LAWS では総務省の「e-Gov」の一環として，現行のすべての日本法令を公開している．この URL は <http://elaws.e-gov.go.jp/search/elawsSearch/elaws_search/> である．

このように自治体職員からの要望を集める過程で，例規作成の立法作業においては，ウォータフォール型の立法方式を採るよりも，他の類似例規を見つけてきて，それをベースに新たな例規を作成する方式の方が圧倒的に多いことがわかり，そちらを支援する，類似例規を見つける機能，そして，実際にそれらを比較表(自治体法務の分野では「ベンチマーク表」と言う)にして資料を作成できる機能が求められた．そこで，これらを実装して，eLen データベースの付属機能として公開している．それが「立法支援機能付き」と呼ばれる所以である．類似例規を見つける機能は，類似な例規ごとのクラスタを算出することによって実現できるため，例規クラスタリングが主要な課題となった．通常のクラスタリングのように，単に意味が似ているだけでなく，文面を真似したり，標準的な書き方を参照したりすることも必要であったので，いきなり例規単位でクラスタリングをする方式を採用せず，まずは，条文単位のペア組合せ間の「編集距離」[5] によって，全組合せについて計算しておき，現在，3 割以内の書き替えに収まっているものを類似条文のペアとして判定している．さらに，その結果を用いて，類似条文となるペアを 7 割以上含んでいる例規ペアを「類似例規ペア」と判断している．より正確には，例規集合 R について，r∈R に含まれる条文の集合を J(r)と記し，$r_1,r_2 \in R$ について $p \in J(r_1) \times J(r_2)$が類似条文のペアとなるものの集合を $Sim(r_1,r_2)$と記すと，r_1 と r_2 が類似例規ペアとなるのは次の条件を満たす時である．

$$|Sim(r1,r2)| / (|J(r1)|+|J(r2)|) \geq 0.7$$

右辺の 0.7 は本研究で現在採用している閾値である．

本研究では，この類似例規ペアをエッジとし，例規を 1 つのノードとするグラフ構造において，クリークとなるもののうち極大なものをクラスタとして算出し[6]，同一クラスタに属するものを類似例規と判定した．さらに，すべての例規を法令用の XML 形式で保持しているので，各条文に付される見出し部分の情報も認識して保持している．そこで，その見出し情報を使って，類似例規や比較対象例規内の条文の種別を区別して比較表を自動作成している．以上が先行研究成果である eLen システムの概要である．

3. テンプレート合成の方式

先行研究の比較表自動作成の過程において，さらなる要望として例規のテンプレートの自動合成が求められた．これは，比較表を見たときに，共通している部分と差異の部分を分けて眺めて考察や作業を進めているという事実に基づいた要望である．図 2 のように，類似した文があれば，共通部分を残し，差異の部分をオプションとして一覧できるようにしておけば(できれば頻度の多いものが順に)，比較表を用いた考察と同じことができるテンプレートができあがるのである．

このようにして，例規のテキストを処理する場合，要するに，マルチな diff コマンド機能と，大前提として類似例規でクラスタリングしておくことが核となる．前者については mdiff という diff 機能拡張したコマンドをすでに実装済みであり，後者は，とりあえず従来の方式によってクラスタを計算することで対応している．本稿執筆時現在，最終的なテンプレート GUI 用のデータに変換するプログラムを作成中である．デモシステムの画面例を図 3 に示す．

また，実験に用いた現実の例規に対して mdiff を実施の例を示す．実際に本システムでクラスタとして計上された次の 4 つの条文本文<a>～<d>に対して，mdiff を適用した出力が最後の<T>である．

<a> (準用規定) この規程に定めていない事項については，全国町村会総合賠償補償保険制度 において適用される「賠償責任保険普通保険約款」，「予防接種実施主体特約条項」及び「全国町村会予防接種事故賠償補償保険契約特約書」の規定を準用する．

<b> (準用規定) この規程に定めていない事項については，全国町村会総合賠償補償保険制度 において適用される「賠償責任保険普通保険約款」，「予防接種実施主体特約条項」および「全国町村会予防接種事故賠償補償保険契約特約書」の規定を準用する．

<c> (準用規定) この規則に定めのない事項については，全国市長会予防接種事故賠償補償保 険制度において適用される「賠償責任保険普通保険約款」，「予防接種実施主体特約条項」及び「全国市長会予防接種事故賠償補償保険契約特約書」の規定を準用する．

<d> (準用規定) この告示に定めていない事項については，全国町村会総合賠償補償保険制度 において適用される「賠償責任保険普通保険約款」，「予防接種実施主体特約条項」及び「全国町村会予防接種事故賠償補償保険契約特約書」の規定を準用する．

<T> (準用規定) この[規則,規程,告示]に定め[てい,の]ない事項については，全国[市長,町村]会[予防接種事故,総合]賠償補償保険制度において適用される「賠償責任保険普通保険約款」，「予防接種実施主体特約条項」[及,および]「全国[市長,町村]会予防接種事故賠償補償保険契約特約書」の規定を準用する．

---

4 「法制執務」とは，法令や例規を実際に記述するための仕事を指すが，その記述スタイルには詳細で膨大な慣習的ルールがある．多くの自治体で参照されるマニュアル[石毛 2012]もある．

5 この編集距離としてはレーベンシュタイン距離を採用しており，挿入，削除および置き換えをそれぞれ 1 カウントとするものである．なお，対象は文字ではなく，形態素としている．

6 「クリーク」とはグラフ G(V,E)において，完全グラフを形成する V の部分集合のことである．

最後の出力テキストの「[」「]」で囲まれた部分が選択肢の候補群となる. それぞれ半角コンマで区切られている.

4. クラスタリング方式の比較

本節では, サンプルとして用いた 23 自治体分の約 13000 例規について, 本方式で採用しているクラスタリングと Paragraph Vector に基づく方式によるクラスタリングを行い, 複数の法律家によって行われた, それらの意味的な比較・検証の結果を報告する.

まず, 前者の方式は, 先行研究の頃から実施していた条文ペアの編集距離情報をベースに類似例規ペアを算出し, そこからクリークを検出するという「編集距離クリーク方式」で出力される例規クラスタを用いる方式である.

次に, 後者の方式は, 「Paragraph Vector モデル」に基づく Doc2Vec ライブラリ[7]を使って計算される例規クラスタを用いたものである[8]. 詳細は Quoc Le らの提案[Le2014]を参照のこと.

それぞれの方式によってクラスタリングされた結果の例規について, 複数の法律家により意味的な比較・検証をしてもらった. その結果, 前者は速度的に問題があるが, 意味的にはほとんど問題なく, 後者では速度的には問題がないが, 意味的には問題があった. 後者では, 得られたクラスタ内に意味的に的外れなものが含まれるケースも多く, また, 大事なものが含まれていないケースも多くあった. さらに, テンプレート合成のためには条文単位で共通の文字列の並びが必須だが, 後者ではそのテンプレート用の共通部分がほとんどない例規が含まれてしまうケースがあり, 実用には耐えないものであった. これは例規クラスタ内に 1 例規でも全く字面が異なる例規が入ってしまうと, 例規文面のほとんどが差分になってしまい, 共通部分がなくなってしまうからである. そこで, 後者による算出結果は, まだまだパラメタ設定やクラスタリング方式の改良の余地はあるが, 現時点では, 例規や契約書のようなもののテンプレート自動合成の見地からは, 採用に踏み切れない. ただし, 後者の速度性には魅力があるため, 現在, 前者と後者の方式の組合せを検討中である.

なお, このようなクラスタの比較をするために有効だったのが「疑似クリーク」の概念であった[大久保 2005]. 疑似クリークとは, 比較的大きな共有部分を持つ極大クリークを合成したもの(各極大クリークをそのグラフ内のノードの集合とみたときの和集合)のことである. これを見つけておき, 共有部分をコアとし, そこからはみ出す各極大クリークの部分集合をオプションと見ることで, 例規の構造レベルでの意味的な分析結果に近い様子が浮き彫りになる. 現在, 法律家による検査用にこの疑似クリークベースの GUI を設計中である. 彼らは, 既に「疑似クリーク」という言葉を使って検査作業を行っている.

5. まとめと課題

本稿では, 先行研究成果である eLen システムの類似例規クラスタリング機能を発展させて, 例規テンプレートの自動合成を目指していることを示した. こうして, 出来上がったテンプレートについて, 機械学習や深層学習などの一般的な問題として, なぜそうなるのか説明できない, という問題を孕んでおり, これが大きな課題となる. しかしながら, 現在, 既に並行して進めている「SDL(Software Defined Law)アプローチ」によって[角田 2018a], 本稿の 1.で示したように, ウォータフォール型に対応させ, Python によって法令プログラムを作成する研究で得られる法制度のクラスの意味的な構造情報を参照することも進めている. こうして, 自動合成された例規テンプレートへの説明が付与されたり, より高度な立法支援やアドバイスがされたりできるようにすることを目指している.

謝辞

先行研究 eLen は文科省科研費基盤研究 A(課題番号 24240040)の助成を受けたものであり, 例規テンプレート自動作成の研究は JST の CREST 事業(グラント番号 JPMJCR1782)の助成によるものである. また疑似クリークによる分析研究は公益財団法人栢森情報科学振興財団(K29 研ＸＸⅡ第 509 号)の支援を受けて実施された.

参考文献

[石毛 2012] 石毛正純: 法制執務詳解 新版 2, ぎょうせい, 2012.

[角田 2009] 角田篤泰, 齋藤大地, 外山勝彦: ソフトウェア開発過程との類似性に基づく立法支援システム,JSAI 全国大会(第 23 回)論文集,2F2:4.1-4.4, 2009.

[角田 2012] 角田篤泰: e-Legislation 環境の構築へ向けて, 情報ネットワーク・ローレビュー, 第 11 巻, pp.13-32, 2012.

[角田 2014] 角田篤泰, 島亜紀 他: 全国自治体例規データベース eLen の構築と定量的例規調査, 情報ネットワーク・ローレビュー, 第 13 巻, 第 1 号, pp.14-33, 2014.

[角田 2016] 角田篤泰: 改め文作成の機械化に向けた基盤 ― e-Legislation の一具体例 ―, 名古屋大學法政論集, 266 号, pp.23-59, 2016.

[角田 2018a] 角田篤泰: SDL(Software Defined Law)としての法令プログラミングと立法支援システム, 第 80 回全国大会講演論文集, pp.391-392, 2018.

[角田 2018b] 角田篤泰: 人工知能の発展と企業法務の未来(4-2), NBL, 1125 号, 商事法務, pp.43-52, 2018.

[Le2014] LE, Quoc and MIKOLOV, Tomas.: Distributed representations of sentences and documents, International Conference on Machine Learning, pp. 1188-1196, 2014.

[大久保 2005] 大久保好章, 原口誠: 疑似クリーク探索によるクラスター抽出, JSAI 全国大会(第 19 回)論文集, 2F3-05, 2005.

---

7 このプログラムは Python を使ったトピックモデリングのためのライブラリ「gensim」に含まれているもので, 詳細と配布については, 次の URL のサイトを参照のこと. <https://radimrehurek.com/gensim/models/doc2vec.html>

8 実験で使用した分散表現ベクトルの学習パラメタは, ベクトル長=300, エポック数＝1000, ウインドウサイズ＝30, 負例サンプル数＝5 であった. 各例規の分散表現ベクトルに対して, 凝集型階層的クラスタリングを実施した. その際, 距離計算手法としては ward 法を, メトリックとしてはユークリッド距離を採用した.



図 1: eLen システムの画面例

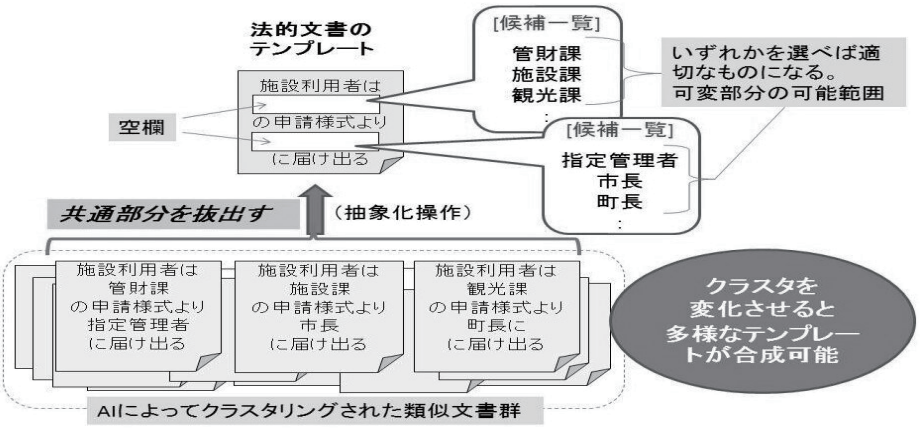


図 2: テンプレート自動合成のイメージ



図 3: 例規テンプレートを用いた立法支援システムのデモ画面例

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

AI裁判支援システムへの人々の期待と受容
People's Expectation and Attitude toward AI-assisted Courts

太田　勝造[*1]
OTA Shozo

[*1] 東京大学・明治大学
The Univ. of Tokyo/Meiji Univ.

Abstract: This paper presents the preliminary results of my nation-wide internet survey on people's expectation and attitude toward AI-assisted Courts which uses experimental research design. Section 1 & 2 explains the components and their AI-models of legal decision-making and the merits and demerits thereof respectively. Section 3 reports the preliminary results which was conducted in February 2019.

1. AIによる裁判支援システムとは？

1. 1.「裁判」とは？

裁判の構成要素（裁判所の作業）の内容は以下の3つに大別できる.

1. 事実認定（心証形成）
2. 法規範の法律要件への「当てはめ」（包摂）
3. 法的推論

1の事実認定（心証形成）とは,「法律要件」に該当する事実（「主要事実」）の真偽を,「証拠方法」（人証, 物証など）によって判断すること作業である. 主要事実は当事者（民事裁判の原告と被告, 刑事裁判の検察官と被告人）が主張し, 証拠方法を提出する. 主要事実の真偽の程度の判断を「心証」と呼ぶ. 心証が「真実（存在）」の方向や「偽（不存在）」の方向へ動いていくことを「心証形成」と呼ぶ.

2の法規範の法律要件への「当てはめ」（包摂）とは, 事実認定の作業によって真であると認定された事実が, 法律要件に当てはまるか否かを判断する作業である. これは法的判断に他ならない. 法律要件が事実を包摂するか否かと表現して, 包摂判断と呼ぶこともある.

なお, 1における事実認定の対象の事実が主要事実であるかの判断と, 2の当てはめの判断は相関する判断であり, 同時並行的になされることも多い.

3の法的推論とは, 請求認容か請求棄却かを判断するため, 適用される複数の法規範について「法的推論」をして結論を導く作業である.

裁判では, 以上の作業のそれぞれの結論を整理して「判決理由」として起草し, 全体の結論を「判決主文」として示す. 民事裁判では, 請求認容, 請求一部認容, 請求棄却, 請求却下などであり, 刑事裁判では, 被告人無罪, 有罪, 有罪の場合の刑の量定などである.

1. 2. 裁判支援のAIモデル

1. 2. 1. 事実認定のAIモデル

事実認定, 心証形成の作業を支援するAIモデルとしては, 不確実性下の意思決定の標準モデルである, ベイズ確率論・意思決定論の応用が考えられる. 証拠方法の吟味に基づく主要事実の真偽の判断は「主観的確率」と位置づけることができ, これには, ベイズ法則による「規範モデル」を用いることができる. 規範モデルにおいては, 合理的事実認定判断のモデルであり, 規範モデルの結論と異なる結果の裁判官認定は非合理とされ, 合致する裁判官認定は合理的とされる.

なお, 現実の裁判官の認定を記述するためのモデルである「記述モデル」ならば, 裁判官認定と異なる結果となれば記述モデルの方が棄却される.

複数の証拠方法の間に関連性が存在することは非常に多いので, より複雑なベイジアン・ネットワークによってモデル化することになる.

1. 2. 2. 当てはめ判断のAIモデル

ある法律要件に, 裁判で争われる具体的事実関係が当てはまるか（包摂されるか）の判断は, 法的ルールが適用されるか否かの中核的判断である.

同一ないし類似の先例や明確な法令が存在して, 法律家なら同じ結論に至るような「イージー・ケース」であれば, 当てはめ判断は容易でほぼ機械的にできる. 他方, 同一ないし類似の先例も明確な法令も存在しない場合は, 正義, 公平, 効率などの考慮に基づく法政策的判断が必要で, 立場や法律家によって判断が分かれる. このような事例は「ハード・ケース」と呼ばれ, 当てはめ判断は非常に困難となる.

イージー・ケースであれば, 論理プログラミングやルール・ベースと・リーズニングなどによってAI化できる. ハード・ケースの場合は, 国民意識や国民の理解などを自然言語処理や深層学習で応用してAI化するか, ネット直結のスポット世論調査のような「民意」組み上げシステムを内蔵したAI化が必要となるかもしれない.

ハード・ケースを解決する上で, 法的ルールのある判断選択肢が他の判断選択肢群に対して望ましいか否かは, 当該判断選択肢が人々の行動や社会状態に与える影響を予測ないし調査して判断する必要があり, これは現実に対応した「人工社会」等を構築してのシミュレイションによるなど, 理論的には考えられるが, 実践状はさしあたり不可能である.

1. 2. 3. 法的推論のAIモデル

裁判では, 刑事であれ民事であれ複数の規範を用いる必要が生じる. 規範相互には矛盾対立が存在しうる. これを解消するためのメタ・レヴェルの規範も存在する.

連絡先：太田勝造（おおた・しょうぞう）, 東京大学大学院法学政治学研究科・明治大学法学部, sota@j.u-tokyo.ac.jp

そして，主要事実のいくつかが証明され，それらに対応する法律要件が認められて，複数の規範が適用される場合，「法的推論」によって正しい結論を導く．

しかも，裁判ではその後にさらに主要事実が証明されて，それに対応する新たな法律要件が認められて，それまでの法的推論の結論が覆ることも起きる．

この法的推論の過程のAIモデルとしては，「非単調推論」を組み込んだ論理的推論が考えられる．

2. 裁判支援のAIモデルの評価

2．1．事実認定

ベイジアン・ネットワークによるモデル化によれば，確率モデルであるゆえに，三段論法のような真偽の二値論理に対して，多値を扱える．確率モデルであるゆえ，さらに，不確実性を合理的に扱うことができる．数理モデルであるので，計算プロセスが透明である．つまり，ベイズ法則を知っていれば説明であり，ミスなどについて，デバッグすることができる．

実質的なベネフィットとしては，事実認定における裁判官や裁判員のバイアスや過誤から解放される可能性がある点が指摘できる．また，AIによる事実認定は，様々なバイアスのありうる人間の裁判の事実認定と異なり正確でありうる．よって，裁判所や裁判官による判断のブレがないと期待できる．

但し，法規範の対象は人間活動と人間社会のほぼすべての領域であり，森羅万象について条件付確率テーブル(Conditional Probability Table)を埋め尽くすことはかなり困難である．大量のデータが必要となるが，現実には存在していない．

実質的なデメリットとしては，裁判の事実も証拠も雑多すぎるので，実現が困難である．また，AIによる事実認定への人々の支持納得，受容がどれほど期待できるかは開かれた問題である．

2．2．法的当てはめ

法的当てはめ判断をAI化できた場合のベネフィットとして，法的判断における裁判官や裁判員のバイアスや過誤から開放されうる点がある．裁判官や裁判員による判断のブレもなくなり，判断の恣意性と主観性を排除できると期待される．こうして，裁判の地理的，時間的な斉一性の実現が期待できる．どの地方の裁判所でも，いつの時代の裁判所でもブレのない判断を示してくれるようになる．

但し，深層学習・自然言語処理による当てはめ判断には，ビッグ・データが必要である．深層学習は，プロセルが不透明で，結論に対する理由や根拠を「見える化」できない．よって，どこが間違っているかわからない．これは法的責任の説明ができず，法的判断の内容（過失，故意，因果関係などの判断）を説明できないことを意味する．

そもそも，規範的価値判断が当てはめであるから「正しい立法」や「正しい先例」が唯一存在する保証はまずない．これは，正解，最適解，すなわち「教師信号」を確定できないことになる．敢えて，教師信号を内部で構成しようとすれば，法的価値判断，法政策的判断をシステムに組み込むことになるが，これは困難である．

現実的には，多数の学説，多数の判例が相互に争っている場合，その処理が困難である．

また，AIによる法的当てはめ判断への人々の支持納得，受容がどれほど期待できるかは開かれた問題である．

2．3．法的推論

法的推論を論理プログラミングで実現する上では，真偽の二値論理の範囲内で可能となる．この場合，プロセスが透明であり，素人相手にも説明できるとともに，デバッグもしやすい．

実質的なベネフィットとしては，論理的推論である限りで，プログラミングや入力にミスがないなら現実妥当性を有する結論を導くことができる．

そして，裁判官や裁判員のバイアス，恣意性から開放されうる．

但し，論理的推論の場合，情報は増えないので，"Garbage-In, Garbage-Out"となる可能性がある．これは，システムのプログラミングでルールを間違えたり，利用の際の入力データを間違えたりすれば，誤回答をすることになる．また，法令や先例は，深層学習やベイジアン・ネットワークを機能させるには少なすぎるが，法令や先例を現状のように人手で論理ルール化するには大量すぎて，労力，費用，時間がかかりすぎる．

また，AIによる法的推論への人々の支持納得，受容がどれほど期待できるかは開かれた問題である．

3. 裁判支援AIシステムへの人々の期待と受容

AIによる裁判支援システムへの人々の支持納得，受容がどれほど期待できるかは開かれた問題であるので，社会調査を実験計画法で設計して，人々の期待と懸念と受容の程度を調査する．

例えば，

質問「交通事故を起こして相手を死亡させてしまい．あなたは，業務上過失致死罪で起訴され，刑事裁判の被告人となったとします．下記の裁判所の中で，あなたが裁判を受けたい裁判所はどれですか．」

と尋ね，

(1)生身の裁判官，
(2)素朴なレベルのAI裁判システム，
(3)かなり高度なレベルのAI裁判システム，
(4)チューリング・テストをクリヤできる（生身の人間と区別がつかない）AI裁判システム，

について，「裁判を受けたい程度」を尺度法で尋ねて比較するなどの調査である（2019 年 2 月時点で調査実施中）．

このようなリサーチ・デザインによるパイロット調査の結果を報告する予定である．

参考文献

[太田 1982] 太田勝造『裁判における証明論の基礎』，弘文堂，1982 年．

[太田 1990] 太田勝造『民事紛争解決手続論』，信山社，1990 年．

[太田 2000] 太田勝造『社会科学の理論とモデル7：法律』，東京大学出版会，2000 年．

[吉野 1986] 吉野一『法律エキスパート システムの基礎』，ぎょうせい，1986 年．

[吉野 1986] 吉野一『法律エキスパート システムの基礎』，ぎょうせい，1986 年．

[吉野 2000] 吉野一『法工知能：法的知識の解明と法的推論の実現』，創成社，2000 年．

[吉野] 吉野一，加賀山茂，櫻井成一朗，新田克 己，鈴木宏昭，太田勝造「事例問題に基づく法律知識ベースおよび論争システムを活用した法創造教育」『IT活用教育方法研究』第 9 巻第 1 号 1 頁～5 頁，2006 年．

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

[4E3-OS-7b] AIの法学の応用(2)

佐藤 健（国立情報学研究所）、新田 克己（国立情報学研究所）

Fri. Jun 7, 2019 2:00 PM - 3:40 PM Room E (301A Medium meeting room)

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

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

2:00 PM - 2:20 PM

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

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

2:20 PM - 2:40 PM

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

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

2:40 PM - 3:00 PM

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

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

3:00 PM - 3:20 PM

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

○Ken Satoh[1], Yoshiaki Nishigai[2] （1. National Institute of Informatics, 2. Nihon University）

3:20 PM - 3:40 PM

©The Japanese Society for Artificial Intelligence

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

司法試験自動解答における論理型言語 PROLEG ルール生成のための過去問分析と設計

Question Analysis and Rule Design for Generating PROLEG Rules towards Automatic Legal Bar Exam Solver

林 隆児
静岡大学情報学部
rhayashi@kanolab.net

狩野 芳伸
静岡大学情報学部
kano@inf.shizuoka.ac.jp

There have been COLIEE competitions held, which aim to extract legal information and to resolve legal textual entailments by answering Japanese legal bar exams. We analyzed 72 past legal bar exam problems to examine possibility of automatic solver. In order to construct such an automatic solver, we use logical programming language called PROLEG, which performs legal reasoning by defining logical rules. While PROLEG rules are normally defined manually, our ultimate goal is automatic generation of PROLEG rules as the first research. Our approach combines natural language processing and logical inference by PROLEG. Assuming that the natural language processing part could be available in future, we implemented programs for other parts that do not perform natural language processing. We performed subjective evaluation for easiness of associating PROLEG rules with problem sentences. Furthermore, we suggest new rule design that makes rules fitting better with natural language processing. This design should allow solving more problems automatically.

1. はじめに

試験の自動解答は自然言語処理や人工知能の主要なベンチマーク課題のひとつといえる。その中で、我が国の司法試験自動解答についても様々な取り組みがされている。COLIEE（Competition for Legal Information Extraction and Entailment）ワークショップは、法律文書の情報検索や含意関係認識問題に関するコンテストとして毎年開催されている[1][2][3]。2018 年に開催された COLIEE 2018[4]では、事件について判例を抽出するタスク（Task1）、判例から判決を決定するタスク（Task2）、問題文から関連する条文を検索するタスク（Task3）、法的質問に対して、関連する条文が分からない状態で、2 値（Yes/No)で解答するタスク（Task4）の四つのサブタスクが設定された。Task3 および 4 は我が国の司法試験から民法短答式の問題を用いたものである。本稿はこのうち、主に Task4 の内容を対象とする。

Task4 では、訓練データとして、質問、関連条文、質問の答え（Yes/No)が与えられ、テストデータとしては質問のみが与えられる。多くの参加チームは、機械学習や単語の重要性の解析を用いて自動解答を試みたが、我々は Prolog[5]をベースにした論理型言語 PROLEG[6]を用いた、自然言語処理と論理推論を組み合わせるアプローチをとる。

PROLEG は、判例や条文をルールとして論理的に記述することができ、それに質問の事実を加えて記述することで、法的問題を推論することができる。PROLEG を用いることで、自然言語処理と法的知識に基づいた、人間の解答プロセスに近い方法で自動解答を実現できるのではないかと考える。法的知識を論理として正しく組み込むことができれば、自然言語処理の性能如何で自動解答の正答率が格段に上がると期待される。

佐藤ら[7]は、契約文書に対して独自に作成したフレームを用いて自然言語処理を行い、PROLEG ルールへの変換を行っている。佐藤らの研究は契約に関する文章に限ったものであるが、司法試験の問題は契約に関する問題だけでなく、契約に関しない問題がさまざまな状況、さまざまな文体で問われる。より多くの問題に適応するためには、膨大なフレームの設計が要求されるだろう。本研究では、法的知識に基づいた自然言語処理により問題文に対応するのではなく、法的知識はすべて PROLEG ルールとして組み込まれているものと仮定する。自然言語処理においては、問題文の単語と PROLEG ルールの単語をどう結びつけるか、という焦点で取り組む。このアプローチを進めるために、本稿では司法試験の問題文を PROLEG ルールに変換する手法を提案し、PROLEG ルール自動生成プログラムの作成を進め、さらにこの手法に基づいた自動解答の可能性について分析を行う。

2. PROLEG

PROLEG とは。論理型言語 Prolog を法律用に拡張したものである。ルールを記述することにより、法的推論を行う。ルールは、Y(X1,X2,…Xn).と表記される。Y が権利や行為を指し、**述語**とよばれる。X は動作主や目的語などが記述され、**引数**とよばれる。<=記号を使った Y1(X1) <= Y2(X2)という表記は、Y2 は Y1 が成立するために必要なルールを示す。この Y2 を Y1 の**要件事実**とよぶ。Y1(X1) <= Y2(X2), Y3(X3).などと表記すると要件は AND 条件になる。proleg 条件(Y(X)).という表記は、Y=X の関係を示す。引数には、**定数**と**変数**の 2 種類がある。定数は確定的な単語であり、変数は推論過程で具体化される。変数は「アンダースコアで始まる」「大文字の英字で始まる」、定数は「小文字の英字で始まる」「漢字で始まる」「シングルクォートで囲まれている」文字列と定義されている。例えば、

所有権に基づく返還請求権(_所有権者,_相手方,_対象物)<=
所有権(_所有権者,_対象物),
占有(_相手方,_対象物).

というルールは、「”所有権者による、ある対象物についての、相手方に対する、所有権に基づく返還請求権”が成り立つためには、”所有権者による、対象物についての所有権と”、”相手方による、対象物の占有”が成り立つ必要がある」と読める。

以下では具体的な問題の解き方と共に PROLEG ルールの作り方を説明する。

[問題文]
土地の所有者は、同土地を不法に占有する者に対して、所有権に基づき土地の明け渡しを求めることができる。

この問題では、物権的返還請求権が認められるかどうかが焦点となる。この物権的返還請求権を軸に、要件事実を展開する。これを、PROLEG では Rule Base とよぶ。Rule Base とは、民法と裁判例を PROLEG ルール化したものである。Rule Base では、A() <= B().と要件事実を展開した

後、B()についてもさらに要件事実があれば展開し、要件事実がなくなるまで続ける。上の問題の物権的返還請求権については、次のような結果が得られる。

```
物権的返還請求権(_所有権者,_相手方,_対象物)<=
        所有権に基づく返還請求権(_所有権者,_相手方,_対象物).
所有権に基づく返還請求権(_所有権者,_相手方,_対象物)<=
        proleg条件(現時点(_占有時)),
        所有権(_所有権者,_対象物,_占有時),
        占有(_相手方,_対象物,_占有時).
所有権(_所有者,_物,_時点)<=
        所有権(_所有者,_物,_もと所有時,_時点).
所有権(_所有者,_物,_もと所有時,_問題となる時点)<=
        もと所有(_所有者,_物,_もと所有時,_問題となる時点).
```

この Rule Base において、末端の要件事実、つまり<=によって展開されている要件事実がないルールについて記述するのが、Fact Base である。Fact Base とは、問題における事件の事実を PROLEG ルール化したものである。上の例では、末端にあたるのが

```
proleg条件(現時点(_占有時)).
占有(_相手方,_対象物,_占有時).
もと所有(_所有者,_物,_もと所有時,_問題となる時点).
```

である。Fact Base においては、これは

```
現時点(_占有時).
主証(占有('B',土地,_占有時)).
主証(もと所有('A',土地,_もと所有時,_問題となる時点)).
```

と記述される。

proleg 条件は、要件ではなく条件として表現したい場合に用いる記法である。proleg 条件は、Fact Base では proleg 条件()述語を外して表現される。通常の要件事実については主証()述語で囲う。さらに、引数は「_相手方」→「'B'」、「_対象物」→「土地」など、事実に基づく単語に置き換えられる。「_占有時」などは問題文に具体的な数値が記述されていないため、変数の形のままで記述する。

PROLEG では Rule Base のトップのルールを質問として記述する。今回の例でいえば、**請求権存在(物権的返還請求権('A','B',甲土地)).**と記述する。引数は Fact Base と同様に事実に基づく単語に置き換える。請求権存在 ()述語は、PROLEG の推論を true/false で出力するコマンドである。

Rule Base には以下のような例外事由がありうる。

```
売買契約に基づく代金支払請求権(_売主,_買主,契約(売買,_売主,_買主,_物,_合意時))<=
      合意(売買,_売主,_買主,_物,_合意時).
例外事由(売買契約に基づく代金支払請求権(_売主,_買主,_契約),同時履行の抗弁(_買主,_売主,_契約,_契約)).
```

「売買契約に基づく代金支払請求権」のようなルールに対して例外事由となるルールが存在していれば、そのルールを記述する。この場合は、「同時履行の抗弁」が成り立てば、例外が成り立ち、「売買契約に基づく代金支払請求権」は成り立たない。問題文から、ルールに対して適切な例外事由を選択する。

3. 提案手法

将来的な自動解答の方法として、以下の手順を提案する:

1. 問題文を自然言語解析し、Rule Base のトップルール(請求権存在()述語に入るべき述語名)を特定する
2. 上記で特定した述語の要件事実を、要件事実群をまとめたファイルから検索して Rule Base を生成する。
3. 生成した Rule Base から、Fact Base で記述すべき述語を抜き出し、引数などを Fact Base の形式に変換する。引数については、問題文と照合して定数と変数を分類する。
4. 完成した質問、Rule Base、Fact Base を Prolog 実行系で実行する。

上記の手法の準備として、自然言語処理を必要としない部分をプログラムで作成し、自然言語処理を行った場合の自動解答の可能性について分析を行った。結果と分析は後述する。

言語処理を必要とする部分については、主に三つある。

まず、Rule Base のトップルールを判断する部分である。問題文を解析し、何を問われているのかを特定し、述語名の候補群と照合する。

次に、例外事由としてルールを選択する場面である。同様に問題文を解析し、例外事由の候補群と照合する。

最後に、Fact Base の生成である。Fact Base は、「_対象物」→「土地」など、Rule Base における表現を事実に基づく単語に置き換えていた。ただし、本研究の目的は正しい PROLEG ルールの生成ではなく、問題の解答である。問題を解きたいだけなら、「土地」に置き換えずとも、「対象物」と、ただアンダースコアを外して定数に変換するだけでよい。言語処理を必要とするのは、どの引数を定数とし、どの引数を変数とするかという分類である。

先の例の引数のうち、「_相手方」「_対象物」「_所有者」「_物」については、問題文に存在が明記されている。このような引数は、Fact Base において定数とする。「_占有時」「_もと所有時」「_問題となる時点」については、問題に時間的要素の記述がないため、Fact Base において変数とする。これをまとめると、以下の Fact Base が完成する。

```
%Fact Base
現時点(_占有時).
主証(占有(相手方,対象物,_占有時)).
主証(もと所有(所有者,物,_もと所有時,_問題となる時点)).
```

実際に Fact Base として上記の形式を用いたとしても、問題の解答結果は一致する。よって、Fact Base の生成においては、問題文に明示的な記述があるかないかによって、定数と変数を分類すればよい。ただし、定数と変数を分類するだけでは、不整合が起き、正しく解けない問題も例外としてある。

4. 実験

自然言語処理を行った場合の自動解答の可能性を分析するために、自然言語処理を行わない部分についてプログラムを作成した。具体的には、Rule Base トップルールから要件事実をたどり、Rule Base を生成する部分については、過去問の Rule Base のいくつかをまとめた仮想的な要件事実群ファイルを作り、それについて、指定したルールから要件事実を末端まで展開し、出力するプログラムを作成した。

Fact Base を生成する部分については、先の出力の Rule Base から、Fact Base に記述すべき末端の要件を抜き出し、主証()述語で囲む。引数の分類は、言語処理を行わないため確実に分類ができないが、「_契約時」のような時間的要素や、「_請求額」のような数値的要素は、問題の事実に関わらないことが多いため、今回は、それらを変数、それ以外を定数として扱った。

5. 結果と分析

司法試験の過去問題のうち、72 問について分析を行った。分析する問題として、佐藤らが手動で PROLEG 化を行った、平成 18 年から平成 27 年までの問題を用いた。これらの司法試験過去問で自然言語処理が必要な三つの作業について、容易性を判断し、分析した。

まず、Rule Base のトップルールを判断する部分について、「述語名と同じ言葉が問題文に含まれている」「述語名に近い表現が問題文に含まれる」「述語名に近い表現が問題文に含まれない」の 3 つに分類した。具体的な例を示す。**図 1** は「述語名と同じ言葉が問題文に含まれる」の例、**図 2** は「述語名に近い表現が問題文に含まれる」の例、**図 3** は「述語名に近い表現が問題文に含まれない」の例である。

[問題文]
土地の賃貸人が借地契約の更新拒絶をするためには，正当の事由がなければならないほか，契約期間の満了の１年前から６か月前までの間に賃借人に対して更新をしない旨の通知をしなければならない。

請求権存在(更新拒絶(賃貸人,賃借人,契約(賃貸借,賃貸人,賃借人,土地,建物所有目的,平成 1 年 1 月 1 日),平成 30 年 1 月 1 日)).

図 1 平成 18 年度短答式民事系第 24 問選択肢エ

図 1 の問題では、問題文に「更新拒絶をするためには、」という表現があり、この問題の主たる述語が「更新拒絶」であることは明らかな上、Rule Base に「更新拒絶」という完全に一致するルールがある。

[問題文]
ＡＢが共有している建物の管理費用をＡが立て替えた場合，Ａは，Ｂからその共有持分を譲り受けたＣに対し，当該立替金の支払を請求することができる。（正）

請求権存在(立替金支払請求権('A','C',管理費用(共有建物,'A 支払額',平成 2 年 1 月 1 日))).

図 2 平成 18 年度短答式民事系第 26 問選択肢 5

図 2 の問題では、問題文に「立替金の支払を請求する」という表現があり、Rule Base トップルールの述語名「立替金支払請求権」と容易に結びつけることができる。

[問題文]
被相続人Ａから相続開始前に甲不動産を買い受けたＸは，Ａの唯一の相続人Ｂの債権者ＹがＢに代位して甲につきＢの相続登記をした上で甲を差し押さえた場合，登記がなくても，甲の所有権取得をＹに対抗することができる。（誤）

請求権存在(異議権の存在('X','Y',甲不動産)).

図 3 平成 19 年度短答式民事系第 8 問選択肢ア

図 3 の問題では、Rule Base トップルールの述語名「異議権の存在」について、問題文に「異議」という単語もなく、表現から見ても結びつけることは難しいと判断した。分析結果を**表 1** に示す。

表 1 Rule Base トップルールの述語名判断容易性の分析結果

分類	問題数	割合
Rule Base トップルールの述語名と同じ単語が問題文に含まれている	3	4.2%
Rule Base トップルールの述語名に近い表現が問題文に含まれる	30	41.7%
Rule Base トップルールの述語名に近い表現が問題文に含まれない	39	54.2%

次に、例外事由としてルールを選択する場面について、「Rule Base に例外事由が含まれない」「例外事由のルールの述語名と同じ言葉が問題文に含まれる」「例外事由のルール名と近い表現が問題文に含まれる」「例外事由のルール名と近い表現が問題文に含まれない」の 4 つに分類した。分類の基準は、Rule Base トップルールと同じである。分析結果を**表 2** に示す。

表 2 例外事由のルールの判断容易性の分析結果

分類	問題数	割合
Rule Base に例外事由が含まれない	22	30.6%
例外事由のルールの述語名と同じ言葉が問題文に含まれる	1	1.4%
例外事由のルール名と近い表現が問題文に含まれる	9	12.5%
例外事由のルール名と近い表現が問題文に含まれない	40	55.6%

最後に、Fact Base の生成の容易性について、分析した。Fact Base の生成においては、問題に事実として記述されている引数を定数とし、記述されていない引数を変数として分類を行う。これについて、「_契約時」のような時間的要素や、「_請求額」のような、数値的要素は、問題に記述されていないことが容易に判断できるため、時間的要素や数値的要素が変数、それ以外が定数となっている問題を「定数と変数の分類が簡単」、そうでない問題を「定数と変数の分類が難しい」として、分類した。ただし、前述したとおり、Fact Base の生成においては、定数と変数を分類するだけでは正しく解けない例外的問題がある。例外的問題はまだ定量的に検証できていないが、このために実際にはもう少し解ける割合が少ない。分析結果を**表 3** に示す。

表 3 Fact Base の生成容易性の分析結果

分類	問題数	割合
定数と変数の分類が簡単	51	70.8%
定数と変数の分類が難しい	21	29.2%

これらはそれぞれの要素の結果を別々にまとめたものだが、次にこれら 3 種類の要素をあわせたものを考える。Rule Base のトップルールや例外事由のルールの判断容易性が「近い表現が問題文に含まれる」以上であれば、問題を自動解答する見込みが高い。Fact Base については、「定数と変数の分類が簡単」であれば同様に見込みが高い。これに基づいた分析結果を**表 4** に示す。

表 4 問題正答容易性の分析結果

分類	数	割合
Rule Base のトップルールの述語名判断容易性が「近い表現が問題文に含まれる」以上かつ 例外事由のルールの判断容易性が「近い表現が問題文に含まれる」以上 かつ Fact Base の生成容易性が「定数と変数の分類が簡単」	16	22.2%

Rule Base のトップルールの述語名や、例外事由のルールに関する自然言語処理がうまくできれば、22.2%の問題が解けるはずである。ただし、前述したように、Fact Base の生成容易性に関して、例外的な問題がある。例外となる問題は、実際に PROLEG を実行してみないと判断が付かず、時間がかかってしまうため、検証が完了していない。この検証の結果によっては、解ける問題の割合は下がる可能性がある。それを踏まえた概算をすると、15%～20%の問題が自動で解ける見込みである。

6. 議論と展望

今後、自動解答の実現、正答率の向上に向けて、やるべきことが三つある。

一つ目は、**表 4** に関して述べた、Fact Base の生成について検証を済ませること。Fact Base を生成する際に、定数と変数を分類するだけでいい問題がいくつあるかによって、自動解答の可能性は大きく変わる。

二つ目は、問題文の自然言語処理である。これまで我々が作成してきた司法試験自動解答のための問題文述語項解析器があり[8][9]、これをさらに PROLEG 生成に適する形に改善したい。

三つ目は、PROLEG ルールの設計、特に Rule Base についてである。過去問分析の項で、Rule Base トップルールや例外事由のルールの判断容易性について記述したが、その判断容易性を、PROLEG のルールを再設計することで改善できると考えている。PROLEG のルールを、問題文の言語処理結果と照合しやすい形に変更することを考えている。例を**図 4** に示す。

[問題文]
書面によらない贈与の受贈者は，贈与者に対して贈与の履行を求めることができない。

請求権存在(契約成立(贈与,受贈者,贈与者,契約(贈与,受贈者,贈与者,対象物,平成 2 年 2 月 2 日))).

%Rule Base
契約成立(贈与,_受贈者,_贈与者,契約(贈与,_受贈者,_贈与者,_対象物,_契約成立時))<=
合意(贈与,_受贈者,_贈与者,契約(贈与,_受贈者,_贈与者,_対象物,_契約成立時)).

図 4 平成 22 年度短答民事系第 22 問選択肢ア

図 4 の問題では、Rule Base トップルールの契約成立()という述語の中に”贈与”という引数があり、契約成立が”贈与に関する”ことを、引数で表現している。契約成立()という述語は贈与以外にも成り立つ表現であり、この状態では「贈与に関する契約成立」のルールを探すために、引数まで解析しなくてはならない。これを改善するためには、ルールを

贈与に関する契約成立(_相手方,_契約主,契約(_相手方,_契約主,_対象物,_契約成立時))<=

とするのが理想的であると考えられる。このように、自然言語処理によって照合しやすい一般的な形にルールを再設計する予定である。

7. 結論

本研究では、司法試験問題の自動解答へ向けて、論理型言語 PROLEG を用いたアプローチのために、手動で PROLEG ルールが作成された過去問 72 問の分析を行った。分析を行うために、自然言語処理を必要としない部分についてはプログラムを作成し、自然言語処理ができれば PROLEG の自動生成ができることを確認した。

自然言語処理を必要とする部分について、自動化の容易性を分析した。問題文と PROLEG ルールの結びつきの難易度を、自然言語処理が必要な三つの場面についてそれぞれ難易度の段階を設定して分類した。その結果、全体で 15%～20%の問題は現状の手法で自動解答できる見込みがあると分かった。

今後は、自然言語処理部分の実装と、解答できる問題を増やすために、ルールの自然言語処理向け再設計を目指す。ルールの再設計については、手動で作成されたルールについて、言語処理をしやすいよう、問題文の形式に沿ったルールや述語の名前に変更する手法を提案した。

謝辞

本研究は JSPS 科研費 17H06103 の助成を受けたものである。国立情報学研究所の佐藤健教授及び藤田恵氏には、過去問の PROLEG 化ファイルの提供、PROLEG の仕組みについてご教示いただいたことを、心から感謝する。

参考文献

[1] M.-Y. Kim, R. Goebel, and K. Satoh, “Coliee-2015: evaluation of legal question an- swering,” in Ninth International Workshop on Juris-informatics, 2015.
[2] M.-Y. Kim, R. Goebel, Y. Kano, and K. Satoh, “COLIEE-2016: Evaluation of the Competition on Legal Information Extraction/Entailment,” in Tenth International Workshop on Juris-informatics, 2016.
[3] Y. Kano, M.-Y. Kim, R. Goebel, and K. Satoh, “Overview of coliee 2017,” Epic Ser. Comput., vol. 47, pp. 1–8, 2017.
[4] M. Yoshioka, Y. Kano, N. Kiyota, and K. Satoh, “‘Overview of Japanese Statute Law Retrieval and Entailment Task at COLIEE-2018,’” in Twelfth International Workshop on Juris-informatics, 2018.
[5]I.Bratko (1986) . “Prolog programming for artificial intelligence” Addition-Wesley Publishing Company(安部憲広(訳)(1990). 「Prolog への入門」近代科学社)
[6]佐藤 健,浅井 健人,古川 昂宗,久保田 理広,中村 恵,西貝 吉晃,白川 佳,高野 千明, “PROLEG:論理プログラミングをベースとした民事訴訟における要件事実論の実装”
[7]Maria Navas-Loro, Ken Satoh, and Victor Rodriguez-Doncel,"ContractFrames:Bridging the gap between Natural Language and Logics in Contract Law"
[8]R. Hoshino, R. Taniguchi, N. Kiyota, and Y. Kano, “‘Question Answering System for Legal Bar Examination using Predicate Argument Structure,’” in Twelfth International Workshop on Juris-informatics, 2018.
[9]R. Taniguchi, R. Hoshino, and Y. Kano, “‘Legal Question Answering System using FrameNet,’” in Twelfth International Workshop on Juris-informatics, 2018.

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

司法試験の問題分類と法律文書における人物関係の構造化辞書

Problem classification of legal bar exam and structured dictionary design for human relationships in legal documents

清田 直希[*1] NAOKI KIYOTA
狩野 芳伸[*1] YOSHINOBU KANO

静岡大学情報学部[*1]

Abstract: Legal document processing is one of promising applications of natural language processing. However, legal sentences are very complicated and require various difficult analysis techniques. We focus on the competition workshop COLIEE (Competition for Legal Information Extraction and Entailment) that aims to automatically answer the legal bar exam. Firstly, we manually classify the past legal bar exam data distributed in COLIEE by required analysis techniques such as conditional sentence extraction and negative form interpretation. Secondly, we focus on person relationships among our classifications, propose a dictionary design for legal document processing based on verbs and their arguments with roles of persons.

1. はじめに

近年の技術革新と共に注目を集めている分野が自然言語処理である。計算機が、人間が日常的に利用している自然言語を処理するために様々な研究がおこなわれている。

その中でも、法律文書の処理に関しては、一般的な文書の言語処理に加えて、専門性の高い言語が多く用いられている。そのうえ、文章中の事象を客観的に評価する必要があるため、難易度は高いが自然言語処理の分野において期待が高まっている。本研究では、毎年開催されている我が国の司法試験自動解答を行うコンテスト型ワークショップ COLIEE (Competition for Legal Information Extraction and Entailment)タスクシリーズ[1][2][3][4]で配布されたデータを対象とし、法律文書処理の基盤リソース構築を狙う。

法律文書においては先述の通り、専門性の高い言語に加え、構文、意味役割、言い換え表現など高度かつ多様な解析技術を必要とする複合的な要素がある。中でも、人物関係は司法試験を解くうえでの重要な問題分類の一つであり、こうした処理には用語の構造化された整理が必要であると考えられる。そこで、人物関係に着目した構造化辞書を提案する。

2. 司法試験の自動解答タスク COLIEE

COLIEE は、COLIEE 2014、COLIEE 2015 [1], COLIEE 2016[2]と人工知能学会国際シンポジア(JSAI-isAI)のひとつである JURISIN(法情報学ワークショップ)において開催してきた。COLIEE 2017[3]は法と人工知能に関するトップカンファレンスである ICAIL (International Conference of Artificial Intelligence and Law)において開催された。COLIEE 2018 ではカナダ法を用いたタスクがが追加されたが、本稿では我が国司法試験を用いたタスクのみを対象とする。

司法試験は弁護士や裁判官、検察官となろうとする者を対象に行われる試験である。問題形式は択一問題を含む短答式と論文式に分けられて行われる。試験の科目数については、短答式が憲法、民法、刑法の 3 科目であり、論文式が公法系科目(憲法及び行政法に関する分野)、民事系科目(民法、商法及び民事訴訟法に関する分野)、刑事系科目(刑法及び刑事訴訟法に関する分野)、選択科目(知的財産法、労働法などの中から 1 科目選択する)の 4 科目である。

COLIEE ではこのうち民法短答式試験を対象としている。多択式を二択にブレークダウンするとともに、解答に関する条文を検索する情報検索タスク(Information Retrieval)と、二択を Yes/No で答える含意関係タスク(Entailment)を用意した。また知識源として民法の全条分を配布し、問題文共々原文の日本語に加え、英語に翻訳したものを配布している。

COLIEE は与えられた問題文に対し、民法条文から関連する条文を返す情報抽出タスク(Phase 1)、問題文と与えられた関連条文について二値の安定をする含意関係タスク(Phase 2)、関連条文を与えずにこの二つを同時に行う質問タスク(Phase 3)の三つのサブタスクからなる。COLIEE 2017 においては、時期の問題から Phase 2 を実施せず、Phase 1 を Task 1、Phase 3 を Task 2 と呼称して実施した。以下で各タスクを定式化する。

情報抽出タスクは、与えられた問題文 Q を回答するのに関連する民法条文のサブセット S1, S2, …, Sn を民法全体から抽出するタスクとなる。解答すべき条文は複数のこともある。

含意関係タスクは、与えられた問題文 Q を、同時に与えられた民法条文のサブセット S1, S2, …, Sn が含意するか、すなわち Entails(S1, S2, …, Sn, Q)または Entails(S1, S2, …, Sn, not Q)を回答するタスクである。

質問応答タスクは与えられた問題文 Q について、民法全体を用いて Yes あるいは No を回答するタスクである。

3. 司法試験の要求要素技術別問題分類

司法試験を解答するうえで、背後にある構文構造、人物の役割と関係性、論理、抽象性など多様で複雑な構造を的確に処理をする必要がある。そこで平成 29 年度のテストデータを対象に、解答に必要な要素を設定した基準を基に手作業で分類をした。表 1 は、解答に必要な要素を手作業で分類した結果と、COLIEE2018 参加チームの要素技術別の正解数及び正答数をまとめた表である。分類の基準は表 2 に記した。この分類は、要素技術ごとに排他的ではないため、ある問いに対して複数の要素技術を要することもある。以下では、分類した要素技術(タスク)のうち、テストデータに頻出なものについて簡単な説明を述べる。

3.1 条件文抽出

このテストデータにおいて、最も多くカウントされた要素技術は、条件文抽出である。これは問題文・条文のそれぞれの条件文を調べることで正誤判定ができる可能性がある問題である。例えば以下のような問いである。

第百六十条 相続財産に関しては、相続人が確定した時、管理人が選任された時又は破産手続開始の決定があった時から六箇月を経過するまでの間は、時効は、完成しない。

	問題数	UA[6]	正答率	UE	正答率	KIS_mo3[7]	正答率	KIS_Frame[7]	正答率
条件文抽出	31	19	0.61	9	0.29	13	0.42	16	0.52
人物役割付与	27	16	0.59	12	0.44	14	0.52	13	0.48
人物関係抽出	26	14	0.54	11	0.42	13	0.50	10	0.38
形態素解析	25	18	0.72	16	0.64	20	0.80	16	0.64
照応解析	20	13	0.65	5	0.25	12	0.60	13	0.65
曖昧性解消	17	9	0.53	7	0.41	8	0.47	9	0.53
意味役割付与	15	8	0.53	6	0.40	9	0.60	6	0.40
動詞言い換え	13	7	0.54	6	0.46	7	0.54	4	0.31
一般辞書	9	5	0.56	5	0.56	4	0.44	4	0.44
述語項解析	9	8	0.89	3	0.33	5	0.56	5	0.56
否定形解釈	7	5	0.71	3	0.43	5	0.71	7	1.00
法律系用語	7	4	0.57	2	0.29	2	0.29	3	0.43
含意関係	5	2	0.40	2	0.40	1	0.20	2	0.40
係り受け	5	3	0.60	1	0.20	2	0.40	4	0.80
条文検索	5	3	0.60	2	0.40	3	0.60	4	0.80
言い換え	5	2	0.40	4	0.80	3	0.60	3	0.60
箇条書き	3	1	0.33	2	0.67	1	0.33	2	0.67

表 1. 平成 29 年度のテストデータの問題の要素技術別分類と COLIEE 参加チームの正答数及び正答率

問い:相続財産に関しては，相続財産管理人が選任された場合でも，相続人が確定するまでの間は，時効は完成しない。

下線部が文章中の条件と考えられる。問いでは相続財産の時効不成立の条件が正誤の判断材料になる。前者は「相続人が確定し、管理人選任(または破産手続き決定)から 6 カ月が経過するまで」という条件下では時効が完成しないのに対し、後者は「管理人が選任されて、相続人が確定するまで」という条件になっている。このように 2 つの文は条件が異なっているため、答えは No となるのである。

3.2 形態素解析

また、形態素解析が特に必要となる問題も数多く見られた。形態素解析は自然言語処理において重要なタスクの一つであるため、特記をする必要性は感じられないかもしれない。ここでの分類基準は、係り受け(単語の修飾関係)や述語項解析が判断基準となる問いの他、条文がおおむねそのまま問いになっているものも含まれる。つまりは形態素解析さえあれば解ける問いも少なからず見受けられるということになる。こうした意味でも形態素解析という分類基準を追加した。

3.3 人物役割付与・人物関係抽出

人物役割付与は、問題文に登場する人物(例えば A など)が条文上のどの役割を担っているのかを明確にするタスクである。人物関係抽出は、売買契約や婚姻関係等の複数の人物間の立場やその関係性を示すために必要となるタスクである。これらについては次節で述べる。

分類名	分類規準
条件文抽出	問題文・条文のそれぞれの条件文を調べることで正誤が判断できる可能性がある問題
人物役割付与	問題文に出てきた人物(A など)がどの役割(未成年,買主,債務者 等)を担っているのかを明確にする必要がある問題
人物関係抽出	複数の人物間の立場や役割を明確にする必要がある問題
形態素解析	おおむね同じような文章である問題　又は文章を解析しやすくするために特定の格助詞などを入れるのに用いる
照応解析	文章の一部を省略している際に、それが何を指しているのかを明確にする際に用いる。「こそあど言葉」のみだけではなく「前文」「同項」なども対象となる
曖昧性解消	曖昧な表現が用いられている文を解析するのに用いる 例:H29-4-A 2　前項の規定による意思表示の無効は、善意の第三者に対抗することができない。
意味役割付与	問題文の中で用いられている名詞が動作主なのか・その動作の対象なのか、手段なのか等を特定する必要がある問題
動詞言い換え	動詞が言い換えられている問題
一般辞書	法律系の言語ではない一般生活で用いられるような事物が出てきている問題(例:冷蔵庫、会社)
述語項解析	その問題の動作を明確にする必要がある問題。特に「～した」「～できる」などはこうした動詞を抜き出すのではなく、その行為を抜き出す必要がある
否定形解釈	否定形になっていて正誤判定をする問題　又は「～できない(不可能)」「～しなければならない(義務)」など「ない」が用いられていてもその意味が異なる問題
法律系用語	法律に関する言語が言い換えられている問題(例:制限行為能力者→未成年など)
含意関係	条文から隠された意図を把握する必要がある問題 ピカソが『ゲルニカ』を発表した→ピカソは『ゲルニカ』の作者である
係り受け	主語と述語の関係、単語の修飾又は文の並列関係に着目する必要がある問題
条文検索	条文中に別の条文が指定されている問題
言い換え	動詞以外で言い換えがされている問題(例:一→1)
箇条書き	条文で箇条書きが用いられている問題。

表 2. COLIEE 司法試験問題の要素技術別分類基準

3.4 照応解析

照応解析は、文章中で登場する「これ」や「あの」などの指示内容を明確にする解析である。前述の指示詞に加えて、「前章」や

「同項」など司法試験に特有の指示表現もあり、言語処理においてはその内容を正しく理解する必要があると考えられる。

3.5 曖昧性解消

曖昧性解消は、文章中に登場する意味的に曖昧な用語を適切に捉える必要がある問いにつけるタスクである。曖昧な用語とは以下のようなものである。

第四百七十五条　弁済をした者が弁済として他人の物を引き渡したときは、その弁済をした者は、更に有効な弁済をしなければ、その物を取り戻すことができない。

ここで述べられている「物」というのが曖昧な用語に該当する。このほかにも名詞であれば「事由」や「意思」、動詞であれば「する」や「ある」といったものが該当する。こうした曖昧性は、複数の対象あるいは幅広い語義を含み、文章理解が困難である。

3.6 意味役割付与

意味役割付与は、文章中の語彙に動作主や場所、対象といった意味的な関係を付与するタスク、SVM[5]やニューラルネットワーク[5]を用いた機械学習手法の提案がされている。

3.7 否定形解釈

否定形解釈は、条文あるいは問いの否定形を正しく解釈したうえで、正誤の判断材料にするタスクである。条文の内容を否定形に反転した司法試験の問題も見られる。

第六百七十五条　組合の債権者は、その債権の発生の時に組合員の損失分担の割合を知らなかったときは、各組合員に対して等しい割合でその権利を行使することができる。

問い：組合の債権者は，各組合員に対して，その権利を行使することができない。

このように条文では権利を行使できる旨が書かれているが、問いでは権利の行使ができないことになっている。このように否定形にすることにより答えが偽になる問題も司法試験では見られ、それは文末の「ない」を認識できれば簡単に評価できる。ところが、次のような問いについてはそうもいかない。

第六百三十三条　報酬は、仕事の目的物の引渡しと同時に、支払わなければならない。ただし、物の引渡しを要しないときは、第六百二十四条第一項の規定を準用する。

問い：請負人は，仕事の目的物の引渡しを要する場合には，これを引き渡した後でなければ，報酬を請求することができない。

前者の文の特殊条件である「ただし」以下の文章を除けば、この 2 文の文末は「ない」が用いられている。しかし、前者の文は「義務」を表した文であり、一方で後者は「否定」を表している文であるため、この 2 文は偽の関係となる。つまり、「ない」に内在される意味を把握しないとこの違いを見抜くことは難しい。

このように問題を要素技術ごとに分類することで、複雑な問題構造を分解した検討ができ、今後の研究の新たな指針になる。

4. 自動解答タスクと人物関係抽出

平成 29 年度のテストデータは全 69 問あり、そのうち言語処理を行う際に、人物関係を抽出する必要のある問い及び人物役割を付与する必要のある問いは、それぞれ 26 問、27 問存在した。しかし、現状の解答器では人物間の関係はほとんど考慮されておらず、表層的な処理だけとなっており、人物関係を抽出するタスクの解決は有用である。以下に人物関係を要する問いの例を挙げる。

甲土地を所有する A が B と通謀して甲土地を B に仮装譲渡し，A から B への所有権移転登記がされた後，B の債権者 C が甲土地を差し押さえた場合において，その差押えの時に C が仮装譲渡について善意であったときは，A は，C に対し，B への譲渡が無効であることを主張することができない。

この問いは正しい主張であるので、計算機は、Yes を出力する必要がある。この問いでは、A、B、C という 3 人の人物が登場していることが明らかである。しかしこのアルファベットには人物のラベルという以上の意味がない。重要なのはこの A~C の人物がどういう関係にあるのかということを計算機が読み取ることである。ここで Yes の根拠となる条文を示す。

第九十四条　相手方と通じてした虚偽の意思表示は、無効とする。

2 前項の規定による意思表示の無効は、善意の第三者に対抗することができない。

A という人物をこの問いの主語とすると、B は相手方、C は善意の第三者という関係が読み取ることができる。よって、A が B に虚偽の意思表示である土地の仮装譲渡を行った後、(A にとっての)善意の第三者である C に土地を差し押さえられたとしても、行為者 A は善意の第三者に対抗できないため、譲渡の無効を主張できない。

こうした人物関係について、動作主、被動作主、第三者などの属性をもつ辞書を構築することで、述語項構造解析による主格・目的格といったツールの出力と紐づけて自動的に人物関係を抽出することが可能になる。さらに、省略されている人物や事物の推測が容易になると考える。

5. 法律文書処理のための構造化辞書

先述の分類のうち、人物関係を対象として意味的関係を含めた構造化辞書を提案する。表 3 は H29 年の司法試験のデータを基に作成した辞書の例である。一つの関係を一エントリとして表現する。関係そのものを表す行為とは、動作主と被動作主の立場を端的に説明する動詞である。また動作主・被動作主・第三者は、条文を基に抽出した。第三者とは、動作主と被動作主の関係を前提に、さらに何らかのアクションをとる存在である。これらの項目は、自然言語処理システムの利用を想定して実際の出現形を基本形にしたものを登録する。以下の条文を例に説明する。

第四百七十五条　弁済をした者が弁済として他人の物を引き渡したときは、その弁済をした者は、更に有効な弁済をしなければ、その物を取り戻すことができない。

第四百七十七条　前二条の場合において、債権者が弁済として受領した物を善意で消費し、又は譲り渡したときは、その弁済は、有効とする。この場合において、債権者が第三者から賠償の請求を受けたときは、弁済をした者に対して求償をすることを妨げない。

この場合は、動作主である「債権者」と被動作主である「弁済をした者」が関係そのものを表す行為「債権を有している」で結ばれている。「第三者」は、この関係を前提に、「第三者の行為」として、「債権者」に対して消費又は譲渡した物についての「賠償を請求」している。

今回は 1 年分のテストデータから抽出したが、エントリを増やすことで条文すべてを網羅できるであろう。さらには動作主・被動作主の属性を記述できると有用である。例えば、成年被後見人とは、精神上の障害により判断能力が欠けており、家庭裁判所から後見開始の審判を受けた人を指す。その成年後見人はこうした人々を後見する(法律行為などを代理したり補助したりする)人物である。こうした人物の可能な法律行為やそれぞれの関係性はエントリを増やすことによって得られるのではないか。

6. 議論と展望

法律文書の辞書化における今後の課題をふたつ述べる。

抽象的な表現の扱いは難しい。例えば先の問いに対する条文では、「虚偽の意思表示」という抽象的な表現が用いられてい

た。具体的な記述がされた問いの文章から条文を検索するためには、問いの文章を抽象化する必要があるだろう。動詞に関しても同様である。「認める」「得る」「…がある」などといった一般性の高い動詞はよく用いられるが、解釈の幅が広く曖昧である。こうした動詞では動詞そのものよりもその項に具体的な表現がなされることが多いため、動詞そのものは利用されないことが多い。動詞の抽象化については、階層的に捉え、包含関係を表す手法[8]などが提案されているが、一般性が高いほど、動詞の置換は慎重な考慮が必要である。

もうひとつの課題は第三者の扱いについてである。先述のように、ここでの第三者には幅広く異なる立場のものが含まれる。権利を行使する立場であったり、許可をする立場であったりとその行為も多種多様である。次のステップとしては、行為の手がかりとして、表層格を用いた行為の判断を検討している。

7. おわりに

本研究では司法試験自動解答を念頭に、法律文書処理のリソース構築のため、司法試験の必要要素別分類を行った。さらにその中でも人物関係に着目して、人物関係の意味をとらえるための構造化辞書の設計と、手作業での辞書エントリサンプルの作成を行った。今後は作業やサンプルとする条文の数が増していく。これにより法律分野に現れうる単語を十分に網羅できる辞書構築を目標としている。こうしたリソースにより、自然言語処理による自動解答、自動処理に貢献できる。今後の研究では辞書を使った人物関係認識の自動化を実装する予定である。

謝辞 本研究は JSPS 科研費 17H06103, JSPS 科研費萌芽研究、JST CREST の助成を受けた。

参考文献

[1] Chang, C. L. and Lee, R. C. T.. Symbolic Logic and Mechanical Theorem Proving. Academic Press, 1973, 331p.

[2] M.-Y. Kim, R. Goebel, and S. Ken, “COLIEE-2015: Evaluation of Legal Question Answering,” in Ninth International Workshop on Juris-informatics (JURISIN 2015), 2015.

[3] M.-Y. Kim, R. Goebel, Y. Kano, and K. Satoh, “COLIEE-2016:Evaluation of the Competition on Legal Information Extraction/Entailment,” in Tenth International Workshop on Juris-informatics (JURISIN 2016), 2016.

[4] Y. Kano, M.-Y. Kim, R. Goebel, and K. Satoh, “Overview of COLIEE 2017,” in the 16th International Conference on Artificial Intelligence (ICAIL 2017), 2017.

[5] 岡田拓哉,竹内孔一,石原靖弘:ニューラルネットワークを利用した日本語意味役割ラベル付与システムの構築,言語処理,2018

[6] M.-Y. Kim and R. Goebel, “Two-step Cascaded Textual Entailment for Legal Bar Exam Question Answering,” in the 16th International Conference on Artificial Intelligence and Law (ICAIL 2017), 2017.

[7] Y. Kano, R. Hoshino, and R. Taniguchi, “Analyzable Legal Yes/No Question Answering System using Linguistic Structures,” in 4th Competition on Legal Information Extraction and Entailment (COLIEE 2017), 16th International Conference on Artificial Intelligence and Law (ICAIL 2017), 2017.

[8] 竹内孔一,乾健太郎,竹内奈央,藤田篤: 意味の包含関係に基づく動詞項構造の細分類, http://www.cl.cs.okayama-u.ac.jp/wp-content/uploads/2017/03/golex.pdf , 2017

関係そのものを表す行為	その他とりうる行為	動作主	被動作主	第三者	第三者の行為
代理する,代表する	責任を負う,同意する	代理人,法廷代理人	(本人),未成年者,制限行為能力者	家庭裁判所	認める
	同意を得る,財産を処分する,通知を発する	(本人),未成年者,制限行為能力者	代理人,法定代理人		
保佐する・補助する	追認する,	保佐人・補助人 法定代理人	制限行為能力者	制限行為能力者の相手方	(制限行為能力者に)催告する,(保佐人・補助人に)行為を取り消す
所有する	償金を支払う,償金を請求する,通行する,譲り渡す,囲障を設ける,持分を取得する,使用する,分担する	主たる所有者(本人)	従たる所有者,その他の所有者,その他の分割者		
債権を有する	取り立てる,(第三債務者に)供託させる	質権者,債権者	債務者	第三債務者	
	譲り渡す,	債務者	債権者	連帯債務者	更改する
	(動作主が)取り戻す,(債権者が)消費する,(債権者が)譲渡する,(動作主に)求償する	弁済をした者		第三者	物を引き渡される,賠償の請求をする
売買する	代金の支払いを拒む,契約の解除,損害賠償の請求,支払う	買主	売主		
	請求する,契約の解除をする,権利を移転する	売主	買主	他人	

表 3. 法律文書の人物関係構造化辞書のエントリ例

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

法律構成を表す双極議論フレームワーク上での推論

Reasoning on a bipolar argumentation framework for a law

川崎 樹　森口 草介　高橋 和子
Tatsuki Kawasaki　Sosuke Moriguchi　Kazuko Takahashi

関西学院大学
Kwansei Gakuin University

We describe a method to construct a bipolar argumentation framework (BAF) from the logical structure of a law. A law consists of a set of ultimate facts to make the conclusion effective, and the defense to make it ineffective. We construct a BAF by considering them to the relations of attacks and supports between arguments, respectively. Moreover, we show a method of reasoning on the BAF that derives the conclusions, that is, the applicable laws, from given facts. The reasoning works bidirectionally: we derive the conclusions from given facts in a bottom-up manner, and find the required facts from these conclusions in a top-down manner to derive new applicable conclusions.

1. はじめに

現在，法的推論を計算機で行う試みが広く行われている．法的推論はある状況に対して法律を適用することで行うが，同じ状況下であっても法律の状況に対する適用の仕方（法律構成という）は異なる．

例 1. 例として表見代理と準占有者弁済について述べる．Y は X から物品を購入し，代金の支払いを X ではなくその代理人 Z に行ったとする．しかし，X は Z に代理権がないこと（無権代理）を主張し，Y に代金の請求を行った．これに対し，無権代理の主張を排斥するために抗弁として表見代理を主張した．

ここで，表見代理の主張が効果を発揮するために必要な条件としての事実（要件事実という）は以下の通りである．

a. 代理人の法律行為

b. 相手方が代理人に a の法律行為について代理権があると信じたこと

c. 相手方の b の誤信についての正当な理由の評価根拠事実

d. a に先立ち，a の法律行為以外のある特定の事項についての代理権の発生原因事実

このうち，要件事実 a〜c は準占有者弁済の主張の要件事実として捉えることができる．準占有者弁済の主張をすることでも無権代理の主張を排斥できるため，表見代理の主張をする代わりに，準占有者弁済の主張しても良い．このような場合，同じ状況下で同じ結論を導くものであるが，法律構成は異なるものができる．

他にも，複数の法律が同時に効果を発揮する状況は存在する．このような状況下で，依頼者がどの法律が効果を発揮しているのかを判断することは容易ではない．そのため，ある状況において考えられる事実からどの法律が効果を発揮するか，または特定の法律が効果を発揮するために他にどのような事実が必要かといった具体的な推論することは依頼者の理解の助けとなる．この法的推論を計算機で行うために法律の構造を具体化する試みがされている．その 1 つの方法として，要件事実を論理型言語で記述する PROLEG が提案されている [4]．PROLEG では要件事実論を条文ごとに 1 つの論理式で記述し，抗弁を例外として記述する．

また著者らは，PROLEG の記述から双極議論フレームワーク (BAF)[1] に変換する手法を示した [2]．BAF は支持と攻撃を持つ議論を扱うシステムであり，議論を議論中の論証と，論証間の反論と支持の 2 つの二項関係を用いて表現する．法律構成を一種の議論としてとらえると，法律間の抗弁の関係は攻撃関係に位置づけることが可能であり，法律と法律の効果を発揮するため根拠となる事実は支持関係に位置づけることが可能である．また，BAF は有向グラフ表現が可能であるため，具体的に関係性を示すことが容易に行える．著者らはこの BAF 上で原告と被告がそれぞれ論争に勝つために証拠として提示すべき事実を導く推論方法を提案した [3]

本研究は，法律構成についてより深く広い理解を得るために，法律構成を用いてどのような推論が可能かを示す手法の提携を目的としている．本論文では，条文から直接 BAF を構築する手法を述べる．また，BAF の上での推論を拡張し，根拠となる事実から，結論を導く手順について述べる．

本論文の構成は以下の通りである．まず第 2 節で双極議論フレームワークについて述べ，次に第 3 節で法律構成から双極議論フレームワークを構築する手法について述べる．その後，第 4 節で双極議論フレームワークを用いた推論の手法を示し，第 5 節でまとめを述べる．

2. 双極議論フレームワーク (BAF)

2.1 BAF

Cayrol らは，議論を攻撃と支持の 2 種類の関係を用いて表現した [1]．著者らはその BAF を拡張し，攻撃を論証間の二項関係，支持を論証と論証の集合の間の二項関係として扱う．

定義 1. 双極議論フレームワーク BAF は $baf = \langle AR, ATT, SUP\rangle$ の組として定義される．AR は論証の有限集合であり，ATT，SUP はそれぞれ $ATT \subseteq AR \times AR$，$SUP \subseteq (2^{AR}\backslash\{\emptyset\}) \times AR$ である．$(B, A) \in ATT$ を $\mathrm{att}(B, A)$，$(\mathbf{A}, A) \in SUP$ を $\sup(\mathbf{A}, A)$ と記述する．

BAF は論証をノード，攻撃関係と支持関係をエッジとした

連絡先: 川崎 樹，関西学院大学，兵庫県三田市学園 2-1，dxk96093@kwansei.ac.jp

有向グラフで表現することが可能である．例えば以下の BAF のグラフ表現は図 1 である．

$baf = \langle \{a, b, c, d\}, \{(b, a)\}, \{(\{b, c\}, a)\} \rangle$

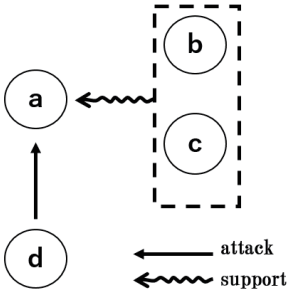


図 1: BAF の例

2.2 ラベリング

ラベリングは BAF における各論証の受理性を表す．議論の意味は議論中の各論証の受理性にあるため，ラベリングは BAF に意味を与える．

定義 2. $baf = \langle AR, ATT, SUP \rangle$ に対して，ラベリング $\mathcal{L}$ とは AR から $\{in, out\}$ への関数である．また，$\mathbf{A} \subseteq AR$ に対して，$\forall A \in \mathbf{A}, \mathcal{L}(A) = in$ のときに $\mathcal{L}(\mathbf{A}) = in$，そうでなければ $\mathcal{L}(\mathbf{A}) = out$ とする．

ここでは，法律構成が表す意味と同等な意味を BAF に与える．これを完全ラベリングと呼び，以下のように定義する．

定義 3. $baf = \langle AR, ATT, SUP \rangle$ において，任意の $\mathbf{A} \in AR$ に対するラベリング $\mathcal{L}$ が以下を満たすならばそのラベリングを完全ラベリングという．

- $(\forall B \in AR, \neg\mathrm{att}(B, A)) \wedge (\forall \mathbf{A} \subseteq AR, \neg\mathrm{sup}(\mathbf{A}, A))$ であるならば $\mathcal{L}(A) = in$.
- $(\forall B \in AR, \mathrm{att}(B, A) \Rightarrow \mathcal{L}(B) = out) \wedge (\exists \mathbf{A} \subseteq AR, \mathrm{sup}(\mathbf{A}, A) \wedge \mathcal{L}(\mathbf{A}) = in)$ であるならば，$\mathcal{L}(A) = in$.
- そうでなければ $\mathcal{L}(A) = out$.

完全ラベリングは法律構成における攻撃防御方法を表す意味論である．効果を発揮することを $\mathcal{L}(A) = in$ とし，効果が排斥されることを $\mathcal{L}(A) = out$ として表す．ある主張あるいはそれに対する抗弁が効果を発揮するためには，それぞれが持つ要件事実を満足し，かつ自身に対するあらゆる抗弁の効果が排斥されることが必要である．一般的なラベリングでは in でも out でもない論証に対して $undec$ のラベルを用いるが，最終的な法律構成において各主張はその効果を発揮するか排斥されるかの 2 値となるため，ここでは $undec$ を考慮していない．

3. 法律構成に準ずる BAF の構築

民事訴訟は攻撃と防御によって成り立つ．原告側が提出する主張や証拠などを攻撃とし，それに対して被告側が提出するものを防御としている．これは，主張・証拠事実によって互いを攻撃している関係と考えられる．つまり，お互いに提出する主張や証拠事実は攻撃関係を持つ．また，主張の効果を認めるためには適用する法律の要件事実を満たす必要がある．この時，法律と要件事実は結論と根拠の関係にあり，支持関係を持つ．このようにして考えることで，法律構成は双極議論フレームワークを用いて表現することができる．

3.1 存在/不存在論証

ある法律の要件事実を満たすためには，要件事実を満たすための法律もしくは，その要件事実に応じた事実が存在しなければならない．よって，すべての主張の根拠はそれぞれに応じた事実の存在であるといえる．この事実の存在を BAF 上で示すために，存在論証と不存在論証の 2 つを与える．存在論証は，要件事実によって求められた事実が存在することを示し，その要件事実を支持する論証である．逆に，不存在論証は要件事実によって求められた事実が存在しないことを表し，その要件事実を攻撃する論証である．存在論証と不存在論証はそれぞれ事実が存在するか否かを表し常に真である論証である．ここではある事実 a に対する存在論証を $\mathrm{ex}(a)$，不存在論証を $\mathrm{ab}(a)$ と表記する．

例 2. *例 1* で示した表見代理の例をもとに *BAF* を作成する．表見代理の主張は無権代理の主張に対する抗弁であるため，表見代理の主張は無権代理の主張を攻撃している．また，表見代理の主張の要件事実は *4* つの事実が同時に満たされることで表見代理の主張が効果を発揮するため，*4* つの要件事実の集合は表見代理の主張を支持している．また無権代理の要件事実 x があるとする．これを *BAF* のグラフで表すと図 *2* になる．

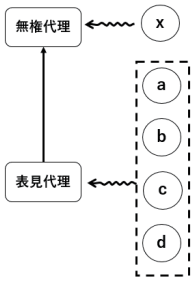


図 2: 表見代理の例に対する BAF

x の根拠が存在しているとする．また，表見代理の主張に対する *4* つの要件事実のうち，a〜c に対する根拠が存在し，d に対する根拠が存在しなかったとすると，この *BAF* には *3* つの存在論証と *1* つの不存在論証が追加され，その完全ラベリングは図 *3* になる．表見代理の主張におけるすべての要件事実が満たされてはいないため，表見代理の主張が効果を発揮せず，無権代理の主張が効果を発揮していることがわかる．

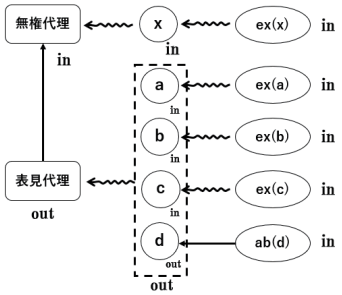


図 3: 完全ラベリング付き BAF

4. BAF の段階的推論

BAF を用いて，結論及びその他の必要とされる事実を推論する手法について述べる．推論は 2 段階からなる．まず，状況から得られる事実を存在論証として与えた時に得られる結論（結論集合）を導く．次に，新たな結論を引き出すために必要な他の存在論証（要求存在論証集合）を導く．

考えられる全ての法律構成を表す BAF が与えられていると仮定し，それを以下のように定める．

$*baf = \langle *AR, *ATT, *SUP \rangle$

4.1 結論集合の導出

初めに与える存在論証の集合を Ex とする．これらの存在論証に対して，それぞれが支持する論証が 1 つずつ存在するので，AR と SUP を次のように定める．

$AR = Ex \cup \{a|\mathrm{ex}(a) \in Ex\}$

$SUP = \{(\{\mathrm{ex}(a)\}, a)|\mathrm{ex}(a) \in Ex\}$

1. $a \in *AR, a \notin AR, \mathbf{A} \subset AR, (\mathbf{A}, a) \in *SUP$ を満たす論証 a 及び集合 $\mathbf{A}$ を探す．AR に a を，SUP に $(\mathbf{A}, a)$ を追加する． この追加を論証 a が見つからなくなるまで繰り返す．
2. $ATT = *ATT \cap (AR \times AR)$ とする
3. 1 で得た AR，SUP，2 で得た ATT を用いて $baf = \langle AR, ATT, SUP \rangle$ とし，この BAF に完全ラベリングを与え，集合 $S = \{s|\mathcal{L}(s) = in \wedge \forall t \forall \mathbf{A}(s \in \mathbf{A} \Rightarrow (\mathbf{A}, t) \notin SUP)\}$ を得る．この集合 S が Ex に対する結論集合である．

例 3. 図 *4* における結論集合は $S = \{S1, S3\}$ である．

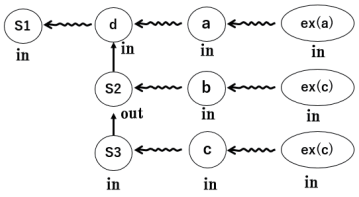


図 4: 結論集合の例

4.2 要求存在論証集合の抽出

次に新たな結論を得るために必要となる存在論証を推論する．

1. 結論集合 S の要素は以下の 3 つに分類される．それぞれ，(a)$*baf$ 内で支持も攻撃も行わない論証，(b) 支持は行わないが $*baf$ 内の他の論証に対して攻撃する論証，(c) その論証と他の論証を含む集合が存在し，$*baf$ においてその集合が支持する関係が存在するような論証，である．結論集合の要素を s とすると以下の論理式で表される．
   (a) $(\forall(\mathbf{A}, a) \in *SUP, s \notin \mathbf{A}) \wedge (\nexists b, (s, b) \in *ATT)$
   (b) $(\forall(\mathbf{A}, a) \in *SUP, s \notin \mathbf{A}) \wedge (\exists b, (s, b) \in *ATT)$
   (c) $\exists(\mathbf{A}, a) \in *SUP, \{s\} \subsetneq \mathbf{A}$
2. (c) の場合に， a とは異なる $b \in \mathbf{A}$ について，[3] の方法に基づいて $\mathcal{L}(b) = in$ とするような存在論証を求める．これによって得られた存在論証の集合が要求存在論証集合である．

例 4. 例えば，図 *5* において $S4$ は（c）に分類される．この時，$\mathcal{L}(\mathrm{f}) = in$ とするような存在論証を求める．

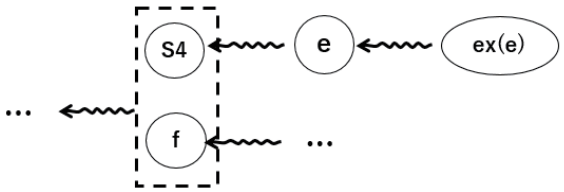


図 5: 要求存在論証集合の例

[3] で述べられた方法を行う際，ある S の要素 s が攻撃する論証が出現した場合，その論証は常に排斥されるため，枝刈りが可能である．これらの推論を繰り返すことによって，結論及び，その結論の効果を発揮させるために必要な存在論証を求めることができる．

5. まとめ

法律が持つ構成要件と法律における抗弁の関係を用いて，法律構成から BAF を構築する手法を示した．さらに BAF を用いて根拠となる事実から結論を推論する手法を示し，新たな結論を得るために必要な事実を求める推論方法について概観を述べた．今後は，全体を統括することでより広い推論方法を検討する．

謝辞

本研究は JSPS 科研費 17H06103 の助成を受けたものである．

参考文献

[1] Cayrol, C. and Lagasquie-Schiex, M.: Bipolarity in argumentation graphs: Towards a better understanding. *International Journal of Approximate Reasoning*, Volume 54, 876-899 (2013).

[2] Kawasaki, T., Moriguchi, S. and Takahashi, K.: Transformation from PROLEG to a bipolar argumentation framework. In *Proc. of SAFA2018*, 36-47 (2018).

[3] Kawasaki, T., Moriguchi, S. and Takahashi, K.: Reasoning by a Bipolar Argumentation Framework for PROLEG. In *Proc. of JURISIN2018*, 16-29(2018).

[4] Satoh, K. et al.: Translating the Japanese Presupposed Ultimate Fact Theory into Logic Programming. In Proc.of *JURIX 2009*, 162-171 (2009).

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

機械学習による裁判の裾野拡大の試み

Towards expanding the scope of trials via machine learning

川本 達郎 *1　靖本 真未 *1　本村 陽一 *1
Tatsuro Kawamoto　Masami Yasumoto　Yoichi Motomura

*1産業技術総合研究所 人工知能研究センター
National Institute of Advanced Industrial Science and Technology, Artificial Intelligence Center

We investigate the possibility of expanding a trial process to the cases that are not usually considered in a court. Our central question is whether it is possible to extract points of argument on each case from the public opinions. To this end, for a few hypothetical cases, we conduct a free-response survey using a web application. The data is collected as a graph (network) so that the public opinions can be summarized using a graph clustering algorithm. As a result, we found that extracted points of argument sensitively reflect details of the cases.

1. 背景

民事裁判は、非常に特殊な条件が揃った時にのみ実際に行われる。当事者双方が、金銭的・時間的コストを支払ってでも徹底的に争う姿勢を見せるような重大な事例であり、かつ裁判をする前から結果が明白ではないような典型的ではない事例の場合である。

しかし、裁判には至らない争いは日常的に存在している。そのような争いのうち幾らかは、裁判の対象に値しないのではなく、現実的なコストに見合わないため裁判事例として現れないだけだと考えられる。これは、権威が行う裁判という仕組みが、小さな訴訟案件まで対処しきれないことを反映している。裁判の裾野とでも呼ぶべき小さな訴訟に対処することは、できないのだろうか。

1.1 裁判の裾野へのアプローチ

本研究では、裁判員制度（あるいは論文の査読制度）のような枠組みから、事例についての論点抽出ができるかを考えたい。裁判員制度といっても、数人の人間が事例について意見を述べるだけであれば、複雑な処理は不要である一方、論点抽出のパフォーマンスは、選ばれたその数人の思考能力に強く依存する。ここではその代わりに、1 つの事例に裁判員として数百人単位の人間を割り当て、集まった意見を集約して、妥当な論点抽出が可能かを検討する。

もちろん正規の裁判過程では、論点を挙げるだけではなく、事実認定や法律条文への “あてはめ”、それらを総合した最終的な判決が必要となるが、ここではそれらには踏み込まない。また、正規の裁判過程において、論点抽出とは、何が争われているかを正確に捉えながら行うものであるが、この点に関しても寛容に扱うこととする。

2. 離婚時の子供の親権問題

裁判事例として、事例数が比較的多いものとして離婚案件が挙げられる。そのなかでも、子供の親権・監護権をどう扱うかについては、賛否両論が分かれる、重要なテーマである。ただし、本研究の主目的は方法論であり、このテーマは一例であることに注意されたい。

連絡先: kawamoto.tatsuro@aist.go.jp

2.1 離婚事例データセットの構築

まず事前調査として、離婚に関係する事例に典型性はあるのか、親権が問題となる事例がどの程度存在するのかについて調べた。データセットとして、離婚に関する web 掲示板 [2] への書き込みを収集し、各事例についてのタグ付けをしたものを作成した。このデータセットは、今後さらにタグ付けを行ったり、言語処理を行うことで、有用な情報を引き出せる可能性もある。

もし離婚事例について自明な典型性が見られるのであれば、問題としては比較的簡単ということである。そうであれば、典型例について数人の専門家で議論すれば十分かもしれない。しかし結論としては、離婚事例について自明な傾向は見られなかった。

2.2 親権・監護権に関するアンケート調査

次に、特定の事例についての論点抽出問題を考える。前節のデータセットを基にして、子供の親権が関わる 5 つのショートストーリーを作成し、訴訟案件に見立てた。前節の離婚掲示板データのように、投稿者からの相談という形式になっており、例えば、1 つは以下のような内容である。

ショートストーリー 1

30 歳専業主婦です。役に立つ資格は、運転免許証のみで他は何一つありません。最近は、夫（35 歳）が感情をコントロールできず子供にまで怒鳴りつけます。私にも子供（5 歳）にも威圧し、物を破壊するところを見せます。離婚したいのですが、私の実父母は他界しており身寄りがないのと、専業主婦の私が親権をとれるのか、今後の経済的な心配もあります。夫と一緒にいるほうが、裕福な生活を送れます。育児は 90%私で、夫は休みの日に遊んでくれる程度です。離婚をして、私が子供を育てることが、子供にとっての幸せでしょうか。

上記のようなショートストーリーについての質問の他に、回答者の属性・人生観・家族観を把握するための質問を盛り込み、約 1 万人に向けてアンケート調査を実施し、6000 人ほどの回答を得た。ショートストーリーごとにアンケートを 5 パターン

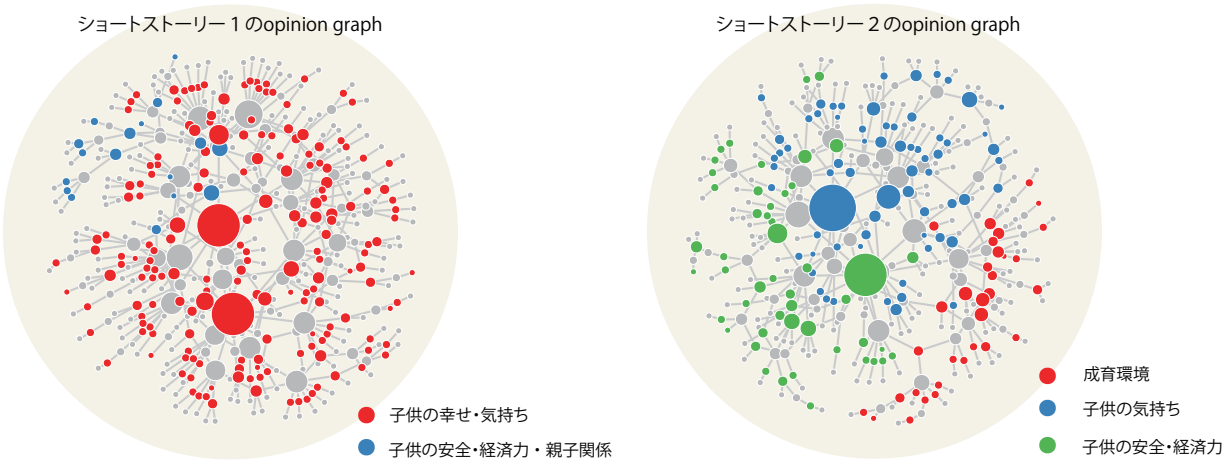


図 1: **投票クラスタリングによって得られた opinion graphs.** グラフの頂点は、各ショートストーリーに対する意見（色付き頂点）と回答者（灰色頂点）を表している。アンケート時に、回答者がある意見に賛成を表明した場合、回答者頂点と意見頂点は枝で結ばれる。頂点の大きさは、その頂点の次数（頂点に付いている枝の数）を反映している。頂点の色は、betweenness の方法 [1] によってクラスタリングをした結果を表している。各グラフの右下には、各色の頂点の代表的な意見をリストしている。

に分けているため、1 つのパターンの回答は 1000 人弱程度である *1。

2.3 投票クラスタリング

ショートストーリーを読んだ上で、どのように論点を見出すかを問うアンケートは、自由記述式である必要がある。そのために我々は、「投票クラスタリング」[3] という、大規模な自由記述式アンケートを実現するシステムを使用した。投票クラスタリングでは、回答文章や回答者をそれぞれ 1 つの頂点、回答文章の類似性を枝で表現したグラフ（opinion graph）を生成し、グラフクラスタリングアルゴリズムによって回答集合の集計を行う。この方法は、一切自然言語処理を必要としないため、いかなる言語・内容のアンケートでも対応可能である。また、クラスタリングアルゴリズムが処理できる限り、どんな大規模なアンケートでも集計可能である。

3. 結果

図 1 は、専業主婦の妻から投稿であるショートストーリー 1 に対するアンケート結果と、専業主夫の夫からの投稿に反転させた内容であるショートストーリー 2 に対するアンケート結果である。子供の気持ちという論点は、どちらにおいても支配的である。ショートストーリー 1 では、ほとんどその意見で埋め尽くされ、わずかに他の論点が出現しているだけである。一方ショートストーリー 2 では、さらに多様な論点が抽出されている。

回答者の属性に何らかの偏りがあるかもしれないため、この結果からだけでは、何によって論点の多様性が生まれているかは判断できない。しかし、本調査では、様々な属性質問を収集しているため、この問いにさらに深く踏み込むことは可能である。

4. まとめ・展望

今回の調査によって、方法論として、大衆のコンセンサスとしての論点抽出の可能性が見出された。今回のアンケートでは、回答者の知識レベル等についてフィルタリングをかけていないが、有識者を対象にしたり、関心の高い人を対象にすることで、より高い分解能の論点抽出が可能になると期待される。この方法が、（法的拘束力はないにせよ）正規の裁判では扱えないような小さな訴訟の対処に用いられたり、正規の裁判にも活用されることが、大きな展望である。

今回のデータセット構築および解析は、論点抽出に使えるかという問いの他にも、親権・監護権についての法律について議論するための様々な情報を与えるものであり、色々な研究用途の可能性も秘めている。さらに、親権問題に限らず多様な問題について要素抽出を積み重ねていくことは、裁判過程を支援する人工知能技術を現実社会で機能させるために、ある種の水準・フレームを与える役割として重要だと考えられる。

謝辞

本研究は JSPS 科研費 17H06103 の助成を受けたものである。研究構想・準備段階においては、高橋周子氏（産業技術総合研究所テクニカルスタッフ [当時]）、石綿はる美氏（東北大学）から、多くの提案をいただいた。また、親権アンケートへ回答していただいた多くの参加者の方々に感謝したい。

参考文献

[1] Girvan, M. and Newman, M. E. J., Proc. Natl. Acad. Sci. USA 99, 7821–7826 (2002).

[2] https://mikle.jp/listp/49/

[3] https://www.bonsci.org/

*1 ただし、集計に実質的な寄与を与える回答者数は数百人程度に落ちる。

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

刑事訴訟版の PROLEG の開発

Implementation of Criminal Procedure using PROLEG Technology

佐藤　健 *1　　西貝　吉晃 *2
Ken Satoh　　Yoshiaki Nishigai

*1国立情報学研究所/総研大　　*2日本大学
National Institute of Informatics/Sokendai　　Nihon University

In this paper, we show a possibility to implement Japanese criminal procedure including Japanese penal code in PROLEG (PROlog-based LEGal reasoning support system). PROLEG has been orginially used to formalize only civil procedure but it is also generally applicable to laws which include general rules and exceptions. Japanese penal code consists of general rules and exceptions so it is possible to formalize it in PROLEG. This paper discusses a concern about how we represent conditions of concepts in criminal law domain using PROLEG.

1. はじめに

本論文では、論理型推論支援システム PROLEG(PROlog-based LEGal reasoning support system) [西貝 11, 佐藤 11] を用いて刑事訴訟を表現することを検討する。PROLEG は、民法の要件事実論 (民事訴訟) を実装することを目的として開発された。もっとも、民事訴訟にとどまらず、[Satoh12] において示したとおり、PROLEG は、原則・例外を用いた推論に基づく法律システムにも応用可能だと期待される。したがって上記特徴を併有する刑事訴訟も表現可能だと思われるが、刑事訴訟には民事訴訟とは異なる理解がなされている箇所があるので、本稿では、それに合わせた実装上の工夫を述べる。

2. PROLEG

PROLEG は、原則・例外からなるルールで表現したルールベースと、当該事件認定された事実を記載したファクトベース等からなる。PROLEG の実行は、当事者の立証活動の結果を証明責任の観点から考慮した形で、民法や刑法を適用する際の推論を行う。

2.1 PROLEG のルールベース

PROLEG のルールベースでは、要件事実論を以下の 2 つの式の形式で表す。

結論 <= 要件 1 , 要件 2 ,..., 要件 n.
例外事由 (結論, 例外).

結論、要件、例外はそれぞれ、述語名 (引数 1, 引数 2,..., 引数 m) の形で表現される。上の第 1 の式（「原則ルール」と呼ぶ）の読み方は、「要件 1 かつ要件 2 かつ... かつ要件 n ならば結論が成立する」であり、その直観的意味は、要件 1 ,... 要件 n がすべて「存在」する場合には結論も「原則としては」、「存在」するということを表す 。また、第 2 の式（「例外ルール」と呼ぶ）の読み方は、「例外は、結論に対する例外事由である」であり、その直観的意味は、例外が「存在」する場合には、いくら原則ルールのすべての要件が「存在」していたとしても、「例外的に」結論が「不存在」となることを表している。

なお、訴訟法の理論としての証明責任の解釈として、法規不適用説と証明責任規範説の相違はあるが、PROLEG においては、デフォルト値の説明の仕方に過ぎない、と考える [西貝 11]。

連絡先: 佐藤　健, 国立情報学研究所, ksatoh@nii.ac.jp

2.2 PROLEG のファクトベース

PROLEG のファクトベースには、実際の事件の事実が記載される。PROLEG の原則ルールの要件のうち、それを結論とした原則ルールがない場合には、それは PROLEG のファクトベースに書くべき事実となる。たとえば、ある事実 P について、PROLEG のルールベースには、それを結論とする原則ルールがないため、もしこの法律要件に対応する事実が存在するならば、事実としてファクトベースに記載する必要がある。PROLEG の刑事訴訟版では、「事実 (P).」のような形式で、そのような事実を記載する。

2.3 意味論

PROLEG の意味論は、論理プログラミングにおける解集合 (Answer Set) と同じように定義される。PROLEG のルールベースにおける原則ルールの集合を $\mathcal{R}$ とし、例外ルールの集合を $\mathcal{E}$ としたとき、すべての変数を言語における基礎項の集合 $\mathcal{H}$ の要素で置き換えた集合をそれぞれ $\Pi_{\mathcal{R}}, \Pi_{\mathcal{E}}$ とする。そして、基礎項 $\mathcal{H}$ の部分集合 M に対して $\Pi_{\mathcal{E}}$ によってフィルターされた $\Pi_{\mathcal{R}}$ の適用可能ルール $\Pi_{\mathcal{R}}^{M}$ を以下のように定義する。

$\{R \in \Pi_{\mathcal{R}} |$
　$\neg\exists$ 例外事由 $(H, E) \in \Pi_{\mathcal{E}}$
　　s.t. $H = head(R)$ and $E \in M\}$

さらに、R の条件部の正リテラルの部分を $pos(R)$ と表し、原則ルールの集合を $\mathcal{R}$ としたときの $\{head(R) <= pos(R) | R \in \mathcal{R}\}$ を $pos(\mathcal{R})$ で表す。
このとき、M が PROLEG プログラム $\langle \mathcal{R}, \mathcal{E} \rangle$ のモデルであるとは、$M = min(pos(\Pi_{\mathcal{R}}^{M}))$ であることをいう *1。

3. 刑事訴訟の構成および及び実装に当たっての留意点

刑事訴訟においては、主にルールベースに関係する刑法 (刑事実体法) と、ルールベースとファクトベースを接続する刑事訴訟法が議論の対象となり、これらの法規範をプログラムする必要がある。

刑法においては、犯罪とは構成要件に該当し、違法で有責な行為だとされる（定説）。つまり、犯罪の成立要件は、構成要件該当性、違法性及び有責性であり、3 段階に分けて要件該当

*1　ホーン節からなるプログラムを P としたとき、$min(P)$ は P の最小モデルを表す

性が検討される。ここで、構成要件に該当すれば、違法（さらには責任）が推定され、原則的に犯罪が成立するとされる。したがって、例えば違法性の段階における考察の内容、積極的な違法の基礎付けではなく、既に構成要件に該当していることによってありそうにみえる違法が「なくなるのではないか」、すなわち違法性を失わせる「違法性阻却事由があるのではないか」、ということが検討される。責任阻却事由についても同様に考えることができる。したがって、違法性阻却事由および責任阻却事由（以下「阻却事由」と呼ぶ）を、その有罪の結論の例外と考えることができる。

しかし、ここで民事訴訟用の PROLEG の推論方式との不一致が問題となる。すなわち、刑事訴訟においては「疑わしきは被告人の利益に」の原則が妥当し、構成要件に該当する事実の「存在」だけでなく、阻却事由に該当する事実の「不存在」の立証責任を検察官が負うことになる。結論的に例外、つまり阻却事由の立証責任を負う者が民事訴訟と逆転しているのである。それゆえ、阻却事由の存在が疑われる場合には、検察官が不存在の証明をしない限り、その阻却事由がデフォルトとして存在しなければならないのだが、PROLEG で直接に、有罪の例外として阻却事由を記述してしまうと、阻却事由が存否不明の場合には、その阻却事由は存在しないことになってしまう。

この問題を解決するために、刑事訴訟法における挙証責任の議論 [川出 02, 堀江 13] に基づき、以下のように原則、例外を分配することにした。すなわち、被疑者は、阻却事由については、証明責任はないが、争点形成責任はあるとして、裁判上で阻却事由の争点形成がされた場合には、原則として、阻却事由が存在すると評価されるようにし、その例外を、それらの事由の不存在が証明されることとした。そして、もし阻却事由の要件が複数ある場合には、それらの要件のどれか一つでも存在しないことが証明された場合には、その阻却事由が存在しないこととした。

たとえば、違法性阻却事由である正当防衛（刑法 36 条 1 項 *2）の場合には、正当防衛についての口頭弁論中に争点形成がされた場合、それを PROLEG ファクトとして登録し、争点形成の存在があれば、原則として正当防衛が成立するとし、それに対して、正当防衛の要件である、「不正の侵害」、「侵害の急迫性」、「防衛の意思」、「防衛行為の相当性」の事実のうちの一部でも不存在であると証明された場合に、正当防衛が不成立となる、という構成とした。

4. 刑法上の要件の実装方法

民法の要件事実論においては、教科書において法律要件がほぼ統一的に定義されている。したがって、それらの法律要件をルールベースに、ファクトベースに法律要件に該当する具体的事実（主要事実，要件事実）を書けばよい。

一方、刑法においては、法律要件に対応するものは、構成要件，阻却事由の要件等である。ここで、民事でもあり得ることではあるが、構成要件等の中には、条文にない要件もあるので適宜、刑法解釈を教科書等で確認しながら実装する必要がある（例：窃盗罪 *3 における「不法領得の意思」*4）。

さらに、正当防衛の成立を否定するためには、正当防衛の要件（急迫不正の侵害、防衛の意思、防衛行為の相当性）のうち少なくとも一つを否定する必要がある。しかし、判例は、さらにそれらの要件の解釈を具体化し、急迫性の不存在の要件として、「積極的加害意思」*5 を採用し、防衛の意思の不存在の要件として「積極的加害行為」*6 という要件を導出しているので、それらについても刑法の教科書等で補う必要がある。

また、解釈の具体化に伴って生じる概念をどこまで PROLEG のルールベースに実装するかは一つの問題となるが、PROLEG の教育応用も考えている我々としては、教科書において抽象的な概念で説明されている場合には、PROLEG のルールとして実装することにした。

5. PROLEG ルール及びテストケースの実行例

この方針にしたがって、PROLEG で刑法を記述した一部を以下に示す。ここでは殺人罪（刑法 199 条）*7 、および、それに対する正当防衛の例外、および、正当防衛の要件の急迫性の不存在の要件である積極的加害意思のルールおよび防衛の意思の不存在の要件である積極的加害行為のルールをルールベースで表現したものを示す。

```
有罪 (_罪,_行為者,_客体,_行為)<=
    構成要件該当性 (_罪,_行為者,_客体,_行為).

構成要件該当性 (殺人罪,_行為者,_客体,_行為)<=
    人 (_客体),
    殺人行為 (_行為者,_客体,_行為),
    死亡 (_客体),
    因果関係 (_行為, 死亡 (_客体)),
    構成要件的故意 (
      _行為者, 殺人行為 (_行為者,_客体,_行為)).

例外事由 (有罪 (_罪,_行為者,_客体,_行為),
    正当防衛 (_罪,_行為者,_客体,_行為,_侵害内容)).

正当防衛 (_罪,_行為者,_客体,_行為,_侵害内容)<=
    争点形成 (
      正当防衛 (_罪,_行為者,_客体,_行為,_侵害内容)).

例外事由 (正当防衛 (_罪,_行為者,_客体,_行為,_侵害内容),
    不正の侵害の不存在 (_行為者,_客体,_侵害内容)).
例外事由 (正当防衛 (_罪,_行為者,_客体,_行為,_侵害内容),
    急迫性の不存在 (_行為者,_客体,_侵害内容)).
例外事由 (正当防衛 (_罪,_行為者,_客体,_行為,_侵害内容),
    防衛の意思の不存在 (
      _行為者,_客体,_行為,_侵害内容,_防衛の意思内容)).
例外事由 (正当防衛 (_罪,_行為者,_客体,_行為,_侵害内容),
    防衛行為の不相当 (
      _行為者,_客体,_行為,_侵害内容,
      _不相当評価根拠事実)).

不正の侵害の不存在 (_行為者,_客体,_侵害内容)<=
    不正の侵害の不存在の評価根拠事実 (
```

*2 急迫不正の侵害に対して、自己又は他人の権利を防衛するため、やむを得ずにした行為は、罰しない。

*3 （窃盗）第 235 条 他人の財物を窃取した者は、窃盗の罪とし、10 年以下の懲役又は 50 万円以下の罰金に処する。

*4 権利者を排除して他人の物を自己の所有物としてその経済的用法に従い利用、処分する意思（大判大正 4.5.21）

*5 相手の侵害の機会を利用して石鋸的に相手に加害行為をする意思で侵害に望んだたときには、侵害の急迫性の要件は充たされない（最決昭和 52.7.21）

*6 防衛に名を借りて侵害者に対し積極的に攻撃を加える行為は防衛の意思を欠く (最決昭和 50.11.28)

*7 人を殺した者は、死刑又は無期若しくは 5 年以上の懲役に処する。

```
        _行為者,_客体,_侵害内容).

急迫性の不存在 (_行為者,_客体,_侵害内容)<=
      急迫性の不存在の評価根拠事実 (
        _行為者,_客体,_侵害内容).
急迫性の不存在 (_行為者,_客体,_侵害内容)<=
      積極的加害意思 (_行為者,_客体,_侵害内容).

積極的加害意思 (_行為者,_客体,_侵害内容)<=
      侵害の予期 (_行為者,_客体,_侵害内容),
      加害意思 (_行為者,_機会を利用して加害する意思).

防衛の意思の不存在 (
      _行為者,_客体,_行為,_侵害内容,_防衛の意思内容)<=
      防衛の意思の不存在の評価根拠事実 (
        _行為者,_客体,_行為,_侵害内容,_防衛の意思内容).
防衛の意思の不存在 (
      _行為者,_客体,_行為,_侵害内容,_防衛の意思内容)<=
      積極的加害行為 (
        _行為者,_客体,_行為,_侵害内容,_防衛の意思内容).

防衛行為の不相当 (
      _行為者,_客体,_行為,_侵害内容,
      _不相当評価根拠事実)<=
    防衛行為の不相当の評価根拠事実 (
      _行為者,_客体,_行為,_侵害内容,
      _不相当評価根拠事実).
```

以下のファクトベースでは、甲が乙から包丁で甲を刺そうとしたところ、それに対抗して乙の胸をナイフで刺したという事件において、甲が乙から襲われることを予見し、襲われる機会を利用して乙を殺してしまおうという事実があったことが示されている。

```
事実 (人 (乙)).
事実 (殺人行為 (甲, 乙, 乙の胸をナイフで刺す)).
事実 (死亡 (乙)).
事実 (因果関係 (乙の胸をナイフで刺す, 死亡 (乙))).
事実 (構成要件的故意 (甲, 殺人行為 (甲, 乙, 乙の胸をナイフで刺す))).

事実 (争点形成 (正当防衛 (殺人罪, 甲, 乙, 乙の胸をナイフで刺す, 包丁で甲を刺そうとした))).

事実 (侵害の予期 (甲, 乙, 包丁で甲を刺そうとした)).
事実 (加害意思 (甲, もっぱら乙の攻撃利用して乙を殺す意思)).
```

図 1 に、PROLEG の実行結果を示す（PROLEG ブロック図）という。左上に結論があり、その結論の箱の一番下が o(成立) となっている。そして実線矢印により結ばれた右の複数の箱に、その結論を満たす要件が示されており、それらの要件の存否が o または x で表される。すべての要件が o であっても、点線矢印により結ばれた例外事由（複数の場合もある）のどれか一つでも成立すれば（o になれば）、結論が否定される。すべてが不成立 (x) であれば、結論が維持される。本件では正当防衛について争点形成がされたため、その例外が検討されている。しかし、その要件の一つである、「急迫性」を不存在とする「積極的加害意思」の事情が存在しているため、正当防衛の急迫性の不存在が証明され、殺人罪が成立することになる。

6. おわりに

本論文では、刑法の PROLEG 化の大まかな表現の方針を示した。しかし、刑法においては、「過失」、「事実の錯誤」（行為者の認識が客観的行為と異なっている場合）、「共犯」など様々な概念があるため、それらを考慮して、この表現方針が矛盾しないかどうか検討する必要がある。

謝辞: 本研究は JSPS 科研費 17H06103 の助成を受けたものである。

参考文献

[西貝 11] 西貝吉晃, 浅井健人, 久保田理広, 古川昂宗, 佐藤健, 白川佳, 高野千明, 中村恵, PROLEG: 論理プログラミング言語 Prolog を利用した要件事実論のプログラミング, 情報ネットワーク・ローレビュー, Vol. 10 pp.54-89 (2011).

[佐藤 11] 佐藤　健, 浅井健人, 古川昂宗, 久保田理広, 中村恵, 西貝吉晃, 白川　佳, 高野千明, PROLEG:論理プログラミングをベースとした民事訴訟における要件事実論の実装, 人工知能学会第 92 回知識ベース研究会資料, pp.1 - 8(2011).

[Satoh12] Satoh, K., Kogawa, T., Okada, N., Omori, K., Omura, S., and Tsuchiya, K., "On Generality of PROLEG Knowledge Representation", Proc. of the Proceedings of the 6th International Workshop on Jurisinformatics (JURISIN 2012), pp. 115 – 128, Miyazaki, Japan (2012).

[川出 02] 川出敏裕, 挙証責任と推定, ジュリスト増刊, 刑事訴訟法の争点 [第 3 版], pp.158-161(2002).

[堀江 13] 堀江慎司, 挙証責任と推定, ジュリスト増刊, 新・法律学の争点シリーズ 6, 刑事訴訟法の争点, pp.148-151(2013).

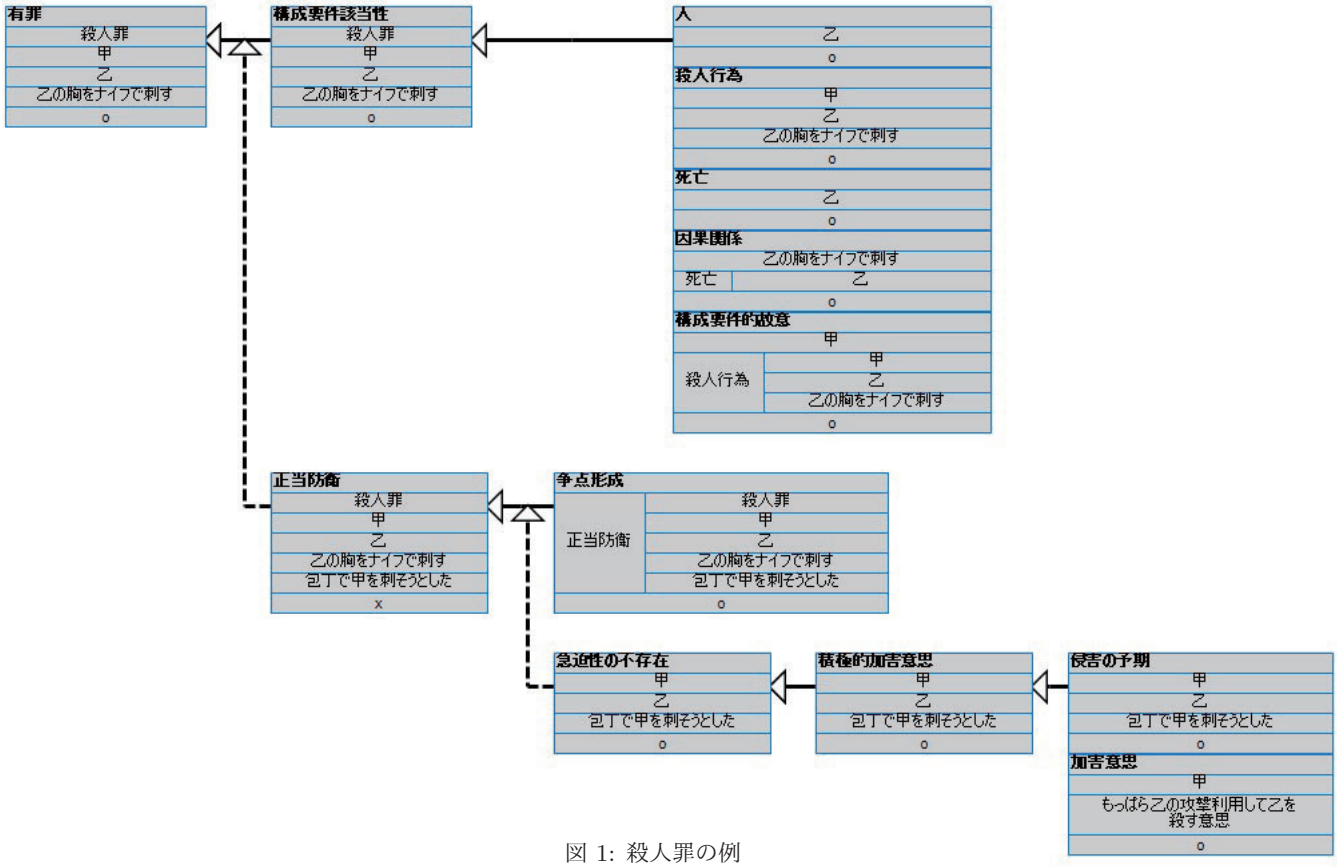

有罪
殺人罪
甲
乙
乙の胸をナイフで刺す
o
構成要件該当性
殺人罪
甲
乙
乙の胸をナイフで刺す
o
人
乙
o
殺人行為
甲
乙
乙の胸をナイフで刺す
o
死亡
乙
o
因果関係
乙の胸をナイフで刺す
死亡
乙
o
構成要件的故意
甲
殺人行為
甲
乙
乙の胸をナイフで刺す
o
正当防衛
殺人罪
甲
乙
乙の胸をナイフで刺す
包丁で甲を刺そうとした
x
争点形成
正当防衛
殺人罪
甲
乙
乙の胸をナイフで刺す
包丁で甲を刺そうとした
o
急迫性の不存在
甲
乙
包丁で甲を刺そうとした
o
積極的加害意思
甲
乙
包丁で甲を刺そうとした
o
侵害の予期
甲
乙
包丁で甲を刺そうとした
o
加害意思
甲
もっぱら乙の攻撃利用して乙を殺す意思
o

図 1: 殺人罪の例

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

[4F2-OS-11a] 社会的信号処理と AI(1)

岡田 将吾（北陸先端科学技術大学院大学）、石井 亮（NTT）

Fri. Jun 7, 2019 12:00 PM - 1:40 PM Room F (302B Medium meeting room)

[4F2-OS-11a-01] The Interaction between Dancer and DJ in Performance Scene of Breakdance

○Daichi Shimizu[1], Takeshi Okada[1] （1. The University of Tokyo）

12:00 PM - 12:20 PM

[4F2-OS-11a-02] Estimating Verbal・ Nonverbal Skills in Business Presentation

○Yagi Yutaro[1], Okada Shogo[1], Shiobara Shota[2], Sugimura Sota[2] （1. Japan Advanced Institute of Science and Technology, 2. SoftBank Corp.）

12:20 PM - 12:40 PM

[4F2-OS-11a-03] Measuring affective sharing between two people from simultaneously recorded EEG signals

○Taiki Kinoshita[1], Hiroki Tanaka[1], Koichiro Yoshino[1], Satoshi Nakamura[1] （1. Nara Institute of Science and Technology）

12:40 PM - 1:00 PM

[4F2-OS-11a-04] Estimating performance of group interaction based on product dimension

○Go Miura[1], Shogo Okada[1] （1. Japan Advanced Institute of Science and Technology）

1:00 PM - 1:20 PM

[4F2-OS-11a-05] Use of Dialogue History for User's Interest Level Estimation Based on Multimodal Information

○Haruto Nishimoto[1], Ryu Takeda[1], Kazunori Komatani[1] （1. The Institute of Scientific and Industrial Research (ISIR), Osaka University）

1:20 PM - 1:40 PM

©The Japanese Society for Artificial Intelligence

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

ダンスパフォーマンスにおける演者間インタラクション：

Dancer-DJ 間の相互作用に関する検討

The Interaction between Dancer and DJ in Performance Scene of Breakdance

清水 大地[*1]
Daichi Shimizu

岡田 猛[*1]
Takeshi Okada

[*1] 東京大学大学院教育学研究科
Graduate School of Education, The University of Tokyo

Abstract: In performing arts like dance and music performance, performers interact with each other actively and present their fascinating performances to audiences. This study aims to investigate this complicated interaction among performers by applying a prominent theory of communication, synchrony. We conducted a case study to examine the interaction between the dancer and the DJ in the performance scene of breakdance. The results suggested that the dancer and the DJ coordinated their performances in a complicated fashion, like coordinating their movements of multiple body parts and changing their coordination patterns as time proceeds. We suggest the importance to develop the framework and analysis methods of synchrony to fully capture these complicated features of the interaction among performers.

1. Introduction

1.1 上演芸術における演者間の関わり合い

ダンスや音楽演奏といった上演芸術では、その場に存在する他者と豊かな関わり合いが営まれながらパフォーマンスが披露されていく。本研究は、上演芸術の中でも特にダンス表現を取り上げ、そこで生じる演者間（特にダンサーと DJ 間）のインタラクションに関する探索的な検討を行ったものである。

実際に上演芸術では、時空間を共有する他者との活発な関わり合いが営まれること、その関わり合いの様相をパフォーマンスとして披露する点に 1 つの魅力が存在することは、インタビューや逸話において広く示唆されてきた。例えば、演者同士が互いの表現に強く依拠した表現を展開すること、その関わり合いから一人では生成出来ない魅力的な表現が生成される場合があることが指摘されている[Bailey 1980; 清水 2013]。

また生物学や文化人類学の観点からも、上演芸術における共演者との関わり合いの重要性はかねてより指摘されてきた。例えば、複数名で共に音楽表現や舞踊的な儀式に取り組むことで社会的な絆（social bond）が生成・強化されること、そしてその生成・強化が共同体の維持・発展にとって有用であったため、様々な共同体において音楽やダンス表現が普遍的に観察され、伝統的に受け継がれてきたことが示唆されている[Merker 2015]。実際に、共に音楽演奏を行うことやパフォーマンスの場を共有することによって参加者間の親密度が上昇することは実証的な検討からも示唆されつつある[Weinstein 2016]。本研究では、この上演芸術における共演者とのインタラクションに着目し、そこで見られる複雑な関わり合いを捉える観点の提案と、ケーススタディによる部分的な検証を行った。

1.2 演者間インタラクションを捉える観点

上記した演者間のインタラクションを科学的に捉えるために本研究が着目したのが、同期現象（syncnronization）に関する知見である。同期現象とは、時空間を共有する複数のエージェント間に、類似した振る舞いが時間的に近接して繰り返し生じることを示す現象である。実際に同期現象は、時計の振り子やメトロノームの振動といった非生物、ホタルの明滅やカエルの鳴き声といった生物に加え[Strogatz 2003]、拍手や歩行・会話時の姿勢といったヒトの多様な振る舞いにおいても観察されることが示唆されてきた[Néda 2000]。そしてそこでの振る舞いのパターンの仔細や現象が生じる背景としての重要な情報の種類（視覚・聴覚等の知覚情報）などがこれまで検討されつつある。

以上の、複数のエージェントが同時に活動を行う際に頻繁に観察される同期現象やその蓄積された理論は、上演芸術における演者間のインタラクションを検討する上でも適用可能であると考えられる。上記した通り、複数名がともに表現を披露するパフォーマンス場面では、互いに多様な媒体を通した関わり合いを行うことによって、演者同士が共通の特徴を有するパフォーマンスを生成・披露していくことが推測されるためである。実際に、近年では以上の理論を適用・拡張することで音楽演奏やダンス表現における関わり合いの様相を捉えようとする試みが営まれつつある[Walton 2018]。一方で、上演芸術におけるインタラクションは、特定の媒体において同一の振る舞いを一致したタイミングで行うことに留まらない。ダンスを例とすれば、多様な媒体（腕部や脚部といった各身体部位など）を用いて、他者と多様なレベルで対応した振る舞いを、多様なタイミングで披露する様相が見られると考えられるのである[清水 2013; 清水 2016]。以上を踏まえ、本研究では、実際のインタラクションのデータに同期理論の解析手法を適用することで、その適用可能性と拡張すべき方向性を検討することを目指した。

1.3 ブレイクダンスのパフォーマンス場面

本研究では、上演芸術の 1 領域としてダンス表現、特にこれまで著者が長期に渡り検討を行ってきたブレイクダンスのパフォーマンス場面を対象とし、上記の検討を行うこととした。ブレイクダンスでは、例えばダンサーや DJ といった多様な振る舞いを行う複数名の演者が活発に関わり合いながらパフォーマンスが展開されることが、映像、インタビューや過去の研究から確認されており、インタラクションの検討を行う上で適切な対象と考えられたためである[清水 2013; 清水 2016]。なおブレイクダンスは、

連絡先：清水大地，東京大学大学院教育学研究科，東京都文京区本郷 7−3−1，tothefuture0415@yahoo.co.jp

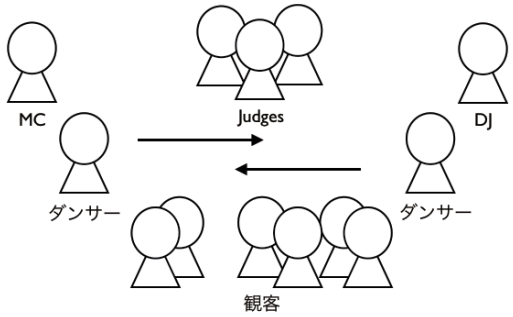


図1. バトル場面の1例

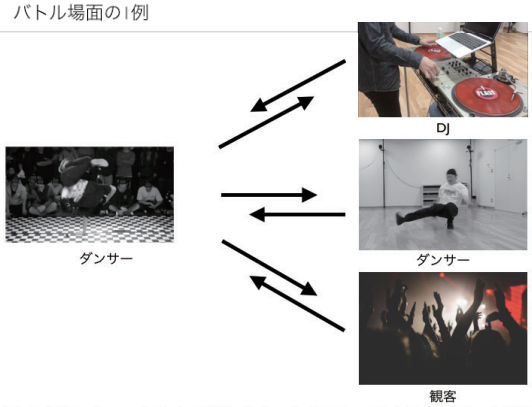


図2. バトル場面において見られる関わり合いによるシステム的な様相。実際は、DJや他のDancer、Audienceの間においてもインタラクションが生じており、それらも含めた非常に複雑な関わり合いが構築されている。

ギャングや若者の抗争の代替として利用され、発展してきた歴史を有する。その経緯から、専らバトルという複数名のダンサーが表現を披露し合って勝敗を決定する形式によってパフォーマンスが披露される[OHJI 2001]。そのため、本研究でも基本的にこのバトル場面の形式を部分的に模した状況を設定して演者間のインタラクションを検討した。

ブレイクダンスにおけるバトル場面の 1 例を図1に示す。バトル場面では、2 名のダンサー、もしくは 2 チームのダンサー達が互いに向かい合った状態で交互にパフォーマンスを披露する。その際には、相手の表現を部分的に取り入れて発展させる、相手とジェスチャーを行い合う・視線を交わし合うといった活発な関わり合いが営まれる様子が観察されている。また、流す音楽に関しては、DJ が即興で選択して編集を加えながら流しており、ダンサーはその音楽にも即時的に反応しながらパフォーマンスを披露する必要が生じる。加えて、MC やジャッジ、観客はダンサーのパフォーマンスに対して歓声や拍手といった形で活発に反応を返しており、ダンサーはその様子にも敏感に反応していることがフィールドワーク映像において確認されている。このようにバトル場面においては、場を共有した複数名のエージェントが互いに活発に関わり合いながらパフォーマンスが披露される様相が見られており、各参与者の振る舞いが互いの関わり合いによってある程度規定される、システム的な振る舞いが生じる場面と捉えることが可能であろう(図 2)。

著者らは既に、ダンサー間に生じるインタラクションを捉える試みをこれまで行ってきた[清水 2013; 清水 2016]。本研究では以上の試みを拡張し、同様の場面においてダンサーと DJ との間に生じるインタラクションに関して探索的な検討を行う。

2. ケーススタディ

2.1 Outline

ここでは、実際に熟達したダンサーと DJ にパフォーマンスを披露してもらい、両者の間に生じるインタラクションの様相について同期理論の解析手法を適用することで探索的に検討した。前述した通り、バトル場面においてダンサーは、DJ がその場で選択し編集を加える音楽に即時的に合わせながらパフォーマンスを披露する[清水 2013]。一方で、DJ も同様にダンサーのパフォーマンスに注意を向け、その振る舞いに敏感に反応する様子がバトル場面において見られている。その際に生じる複雑なインタラクションについて同期理論を用いて定量的に検討することを目指した。その際前述した上演芸術の特徴を踏まえ、特に多様な媒体(身体部位)において対応した振る舞いが生じるのか、その対応関係が時間経過によって変化するのか、という 2 点を検討した。

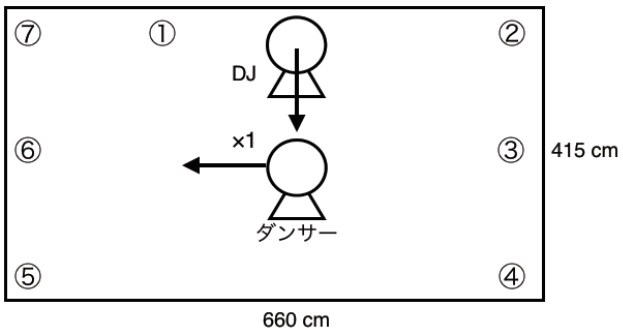


図3. 実験の実施図

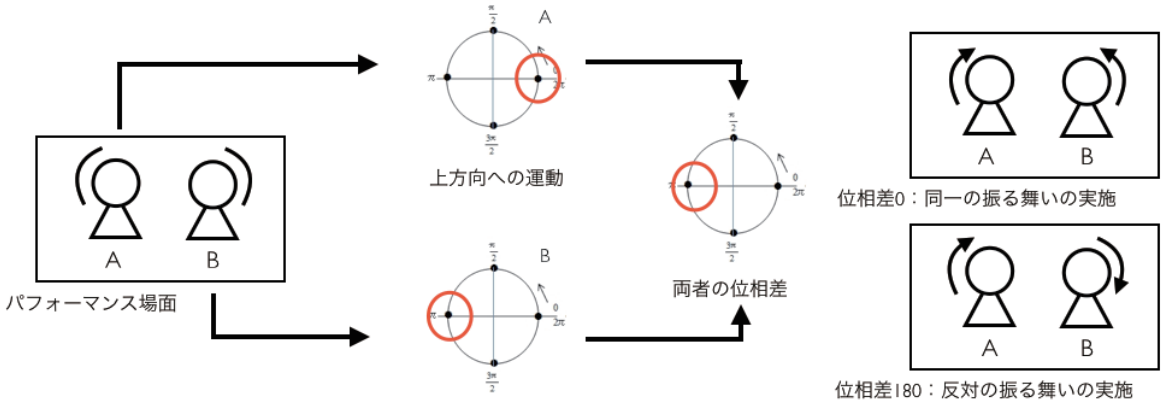
2.2 Method

協力者は、熟達したダンサーA, B の 2 名 (A : 30 歳、経験年数 14 年、B : 30 歳、経験年数 13 年、共に国内の大会における優勝・準優勝等の経験を有する) と熟達した DJ1 名 (27 歳、大会等における DJ 経験を豊富に有する)である。

実験では、ダンサー1 名ずつが DJ とペアとなり、計 3 回のパフォーマンス(ダンサー:ダンスの披露、DJ:曲の即興的な選択・編集)を行った。本研究では、その中でも各ペアで同じ音楽を用いた 1 回目の測定に関して解析を行った。この場面において、ダンサーは DJ の流す音楽に合わせながらダンス表現を 1 回披露しており、その際の両者の振る舞いに関する測定・解析を行った(図 3)。なお、ダンサーは DJ の流す音楽について事前に知らされておらず、DJ は実験者の設定した 2 つの条件(現実のバトル場面で頻繁に流される著名な曲の中から 1 曲選択して流す、両ダンサーで同一の曲を流す)に基づいて流す曲をその場で選択した。また、パフォーマンスを行う時間にも厳しい制限を設けず、可能な限り現実場面に近づけた生態学的妥当性の高い状況で実験を実施した[1]。

そして上記の状況におけるダンサーと DJ 各自の振る舞いに関してモーションキャプチャーシステム(OQUS 300, Optical motion capture system, QUALISYS co.)を用いて測定した。まず、パフォーマンスを行う上で邪魔にならない箇所を中心に全身にマーカーを接着した。今回は特に、大局的なリズム運動を反映すると考えられるダンサーと DJ の頸部の上下運動(z 軸の座標情報)とパフォーマンスの細部を反映すると考えられる左手の運動(ダンサーは z 軸、DJ は y 軸の座標情報)について、マーカーの位置情報を利用して測定・算出した。前述した通り、上演芸術において演者らは、他の演者と多様なレベルで関係した振る舞いを多様な媒体を用いて営んでいくことが予想される。そのため、ダンサーと DJ との間で共有可能な情報であり、互いのパフォーマンスに大きな影響を及ぼすと考えられる互いのリズム運動と互いの表現の細部に着目して検討を行うこととした。特に

[1] 現実場面でもパフォーマンスの時間は明確には指定されない場合が多い。



$$SV_A(t) = s_A(t) + ls_{H_A}(t)$$

$$SV_B(t) = s_B(t) + ls_{H_B}(t)$$

①ヒルベルト変換によるダンサー・DJの振る舞いに関する位相の算出

$$\Delta\emptyset_{AB}(t) = \left| tan^{-1} \frac{s_{H_A}(t)\cdot s_B(t) - s_{H_B}(t)\cdot s_A(t)}{s_A(t)\cdot s_B(t) + s_{H_A}(t)\cdot s_{H_B}(t)} \right|$$

②ダンサー・DJ間の位相差の算出

図4. ダンサー・DJ間の振る舞いの対応関係に関する解析手法。Okumura et al. (2012)に基づいて作成した。

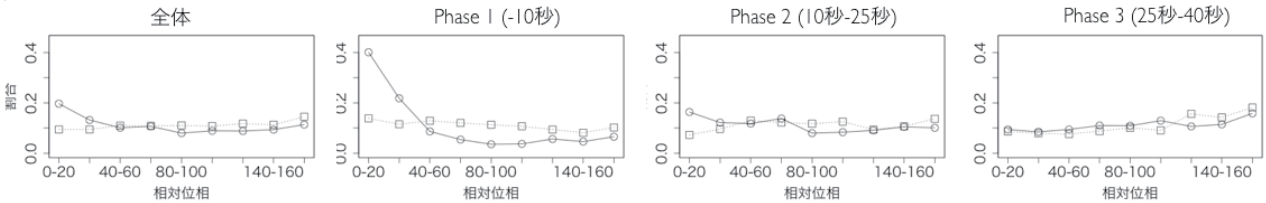


図5. リズム運動に関する対応関係の結果. 実線はreal pairの結果を、点線はpseudo pairの結果を示す.

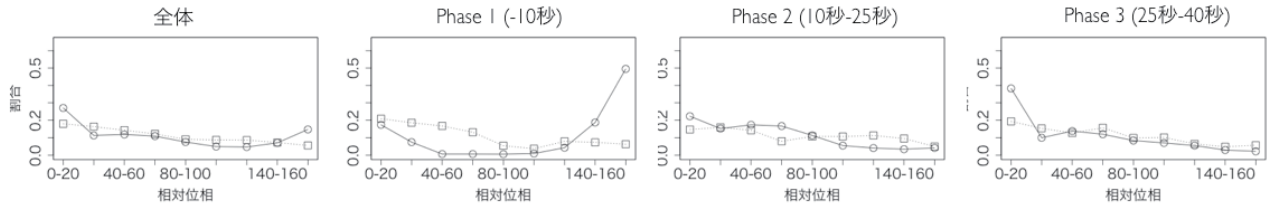


図6. 表現の詳細に関する対応関係の結果. 実線はreal pairの結果を、点線はpseudo pairの結果を示す.

DJ の手の動きは、DJ が音楽を操作する際の振る舞いを直接的に反映した指標と捉えることが出来るだろう。

2.3 Analysis

解析では、上記のマーカーの位置情報を利用して、1:ダンサーと DJ 両者のリズム運動の対応関係と 2:ダンサーと DJ 両者の表現の細部(左手の振る舞い)の対応関係について算出した。まず、1:ダンサーと DJ 両者のリズム運動の対応関係に関しては、頸部の上下運動のデータに、平滑化(2 次の Butterworth フィルタ、10 Hz)と標準化(z score の算出)の処理を行った上で、ヒルベルト変換を行い、ダンサー・DJ 各自の上下運動の位相を算出した。そして両者の位相の差を取ることで上下運動に関する相対位相を算出した(図 4)。

次に 2:ダンサーと DJ 両者の表現の細部(左手の振る舞い)の対応関係に関しても、同様に左手の振る舞いのデータに平滑化・標準化を行った上でヒルベルト変換を実施し、各自の左手の運動の位相を算出した。そして両者の位相差を取ることで相対位相を算出した。算出した相対位相は、0 度であれば相手と同一の振る舞いを行った様子(一方が上方に運動した際に他方も上方に運動したこと)を、180 度であれば相手と反対方向の振る舞いを行った様子(一方が上方に運動した際に他方が下方に運動したこと)を示唆する指標と考えられる(図 4)。以上の指標は、剣道における複数名間の振る舞いの対応関係を検討した研究を参考に算出した[Okumura 2012]。

また、今回はダンサーがパフォーマンス内で表現の特徴を大きく 2 回変化させる様子が見られた(フェイズ 1:立位で軽くステップを行うパフォーマンス、フェイズ 2:立位で激しくステップを行うパフォーマンス、フェイズ 3:床に手や頭を着けながら激しい運動を行うパフォーマンス)。そのため、パフォーマンス全体に加え、各パフォーマンスフェイズにおける上記の指標も算出し、時間経過による対応関係の変化に関する検討も行った。

さらに両指標の算出時に、実際にパフォーマンスを行ったペア(real pair)に加えて DJ の位置情報をもう一方のペアにおける DJ の位置情報と入れ替えた偽ペア(pseudo pair)も作成し、上記の指標を算出した。これは、流された音楽に元々含まれるリズムやビートによってパターン化された対応関係が生じることが予想されたためである。その影響のみを反映した対応関係の結果を算出し、そのベースラインとの比較を行うことで、実際のインタラクションにより生じる対応関係を同定することが可能となる。

2.4 Results and Discussion

まず、1:ダンサーと DJ 両者のリズム運動の対応関係(頸部の上下運動の相対位相)の結果を図 5 に示す。全体を集計した結果(図 5 の左から 1 つ目)から、real pair と pseudo pair ともに同位相もしくは逆位相の頻度が高く、ダンサーと DJ との間で振

る舞いにある程度の対応関係が見られたことが分かる。この相対位相の頻度を、ダンサーのパフォーマンスフェイズごとに分割したものが図 5 の左から 2 つ目−4 つ目である。図から real pair においては初期に同位相による対応が多く見られ、後期に逆位相による対応が多く見られるといったフェイズを経た対応関係の変化が見られた。一方で pseudo pair においては全フェイズを通じて逆位相による対応が多く見られ、その関係が大きく変化する様子は窺われない。以上の結果は、実際にダンサーと DJ がインタラクションを行う場面では、その関わり合いの様相を活発に変化させつつパフォーマンスを披露したことを示唆している。

同様の結果は、2:ダンサーと DJ 両者の表現の細部の対応関係(左手の振る舞いの相対位相)に関してより顕著に示唆された。全体の集計結果(図 6 の左から 1 つ目)を見ると、real pair では同位相と逆位相という振る舞いの異なる対応関係が多く見られたのに対し、pseudo pair では同位相という一定の対応関係のみが多く見られていた。各フェイズの結果(図 6 の左から 2 つ目−4 つ目)においても、real pair では初期に逆位相の対応が多く見られ、後期に同位相による対応が多く見られており、実際にダンサーと DJ が互いに関わり合いを活発に変化させながらパフォーマンスを披露した様子が窺われた。一方で pseudo pair では全フェイズを通して同位相の対応関係のみが多く見られており、関わり合いを変化させる様子は見られなかった。以上の結果を整理すると、実際のパフォーマンスにおいては、ダンサーと DJ との間にリズム運動や表現の細部といったパフォーマンスの多層的なレベルにおける対応関係が見られる様子、そしてその対応関係が時間経過を経て活発に変化する様子が示唆されたと考えられる。

3. General Discussion

ケーススタディでは、ブレイクダンスのパフォーマンス場面におけるダンサーと DJ との間に生じるインタラクションの探索的な検討を行った。結果として、各演者の振る舞いに複雑な対応関係が存在する可能性が示唆された。具体的には、大局的なリズム運動を表す頸部や、表現の細部を表す左手首など、多様な媒体において対応関係が見られたこと、そしてこれらの媒体における対応関係が時間経過によって活発に変化する様子が見られたこと、が示唆されている。

以上の結果は、上演芸術における同期理論の適用の有効性と上演芸術の特徴を踏まえた理論の拡張の必要性とを示唆するものと考えられる。1 章で推測した通り、上演芸術の演者間インタラクションにおいては、多様な媒体を用いて他者と多様なレベルで対応した振る舞いを多様なタイミングで披露する複雑な関わり合いが生じることが示唆された。そしてそのインタラクションは、パフォーマンスの展開と共に活発に変化する様相が窺われたのである。類似した結果は、ダンサー同士のインタラクションに関する知見からも示唆されている(清水・岡田, 2016)。以上を踏まえると、上演芸術における演者間の関わり合いを検討する上で、同期理論において蓄積された枠組みや解析手法は有効に機能すると考えられるだろう。一方で、上記した複雑な様相を十分に捉えるためには、おそらく既存の同期理論における観点(同一の媒体における時間的に近接した類似した振る舞い)を超えた、多様な媒体間における複雑な対応関係を検証する必要があると考えられる。例えば、表現のどういった側面・媒体において他者とどのように協調した振る舞いが生じるのか、それらの観点を定量的に検討可能であるように同期理論の観点・解析手法を拡張することが必要と考えられる。実際に、近年 Beyond Synchrony という枠組みにおいて、ヒトの現実に生じる複雑な協調関係(互いの補助的な振る舞いの実施、発話内容の部分的な取り入れと発展)を捉えるための同期理論の拡張が営まれつつある[Dale 2013]。そこで用いられた観点や解析手法との関係性を踏まえつつ、対象とする表現の特徴を考慮した理論や解析手法の拡張が今後必要になることが想定される。

今後は、上記の理論や解析手法の拡張を行っていくとともに、上演芸術において生じる演者間の複雑なインタラクションを探索的に検証する試みを継続していく予定である。例えば、実際のバトル場面においては、図 2 で示唆されたように 3 者、4 者を超えた多数の参与者(ダンサー、DJ、MC、観客など)による振る舞いが高度に相互作用することでパフォーマンスが生まれ、展開していく。また、その関わり合いに関しても大きな個人差が存在する可能性が示唆されている[清水 2016]。以上の関わり合いの様相を詳細に検討するとともに、どのような関わり合いを通して魅力的な表現が生成されるのか、表現としての魅力にも着目した検討を今後は行っていく予定である。

Acknowledgment

本研究は、科学研究費補助金若手研究 B(課題番号:16K17306、代表:清水大地)の助成を受けて行われた。

参考文献

[Bailey 1980] Bailey, D.: IMPROVISATION, Buxton: Moorland Publishing, 1980.

[Dale 2013] Dale, R., Fusaroli, R., Døjbak Håkonsson, D. D., Healey, P., Mønster, D., McGraw, J., ... & Tylén, K.: Beyond synchrony: complementarity and asynchrony in joint action, In Proceedings of the Annual Meeting of the Cognitive Science Society, Vol. 35, No. 35, 2013.

[Merker 2015] Merker, B., Morley, I., & Zuidema, W.: Five fundamental constraints on theories of the origins of music, Philosophical Transactions of the Royal Society B: Biological Sciences, 370(1664), 20140095, 2015.

[Néda 2000] Néda, Z., Ravasz, E., Brechet, Y., Vicsek, T., & Barabási, A. L.: Self-organizing processes: The sound of many hands clapping, Nature, 403(6772), 849, 2000.

[OHJI 2001] OHJI: ROOTS OF STREET DANCE, 東京:ぶんか社, 2001.

[Okumura 2012] Okumura, M., Kijima, A., Kadota, K., Yokoyama, K., Suzuki, H., & Yamamoto, Y.: A critical interpersonal distance switches between two coordination modes in kendo matches, Plos One, 7(12), e51877, 2012.

[清水 2013] 清水大地・岡田猛:ストリートダンスにおける即興的創造過程, 認知科学, 20(4), 421-438, 2013.

[清水 2017] 清水大地・岡田猛:舞台表現における他者との相互作用のダイナミクスーコミュニケーションの隠れた次元としての距離による検討, HCS2016, 95, 29−34, 2017.

[Strogatz 2003] Strogatz, S. H.: Sync: rhythms of nature, rhythms of ourselves, Allen Lane, 2003.

[Walton 2018] Walton, A. E., Richardson, M. J., Langland-Hassan, P., & Chemero, A.: Improvisation and the self-organization of multiple musical bodies, Frontiers in Psychology, 6: 313, 2015.

[Weinstein 2016] Weinstein, D., Launay, J., Pearce, E., Dunbar, R. I., & Stewart, L.: Singing and social bonding: changes in connectivity and pain threshold as a function of group size, Evolution and Human Behavior, 37(2), 152-158, 2016.

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

ビジネスプレゼンテーションにおける言語・非言語的能力の自動推定

Estimating Verbal・Nonverbal Skills in Business Presentation

八木 悠太朗 *1 Yutaro Yagi　岡田 将吾 *1 Shogo Okada　塩原 翔太 *2 Shota Shiobara　杉村 聡太 *2 Sota Sugimura

*1北陸先端科学技術大学院大学 先端技術研究科
School of Information science, Japan Advanced Institute of Science and Technology (JAIST)

*2ソフトバンク株式会社
SoftBank Corp.

This paper focuses on developing a model for estimating presentation skills of each participant from multimodal (verbal and nonverbal) features. For this purpose, we use a multimodal presentation dataset including audio signal data and body motion sensor data, text data of speech contents of participants observed in 58 presentation sessions. The dataset also includes the presentation skills of each participant, which is assessed by two external observers of the Human Resources Department. We extracted various kinds of features such as spoken uttetances, acoustic features, and the amount of body motion to estimate the presentation skills. We created a regression model to infer the level of presentation skills from these features using support vector regression to evaluate the estimation accuracy of the presentation skills. Experiment results show that the multimodal model achieved 0.59 in R^2 as the regression accuracy of effective production elements.

1. はじめに

企業やビジネスの場において，プレゼンテーション能力は重要なスキルのひとつである．プレゼンテーション能力が必要とされる場面として，企画やアイデアを発表するプレゼンテーションや商品やサービス内容を説明する店頭販売，採用面接などが挙げられる．プレゼンテーション能力を数値化することが出来れば，プレゼンテーションの振り返りや練習などを行う自動評定システムの構築に役立つ．本研究では，「プレゼンテーション能力」に関する個人特性を計算モデルにより捉えることに焦点を当てる．

プレゼンテーション能力の向上支援に向け，本研究は，プレゼンテーションを通じて観測できる参加者個人の話し方，体の動作，発話内容から，プレゼンテーション能力値を推定するモデルを機械学習により構築し，評価することを目的とする．推定モデルの構築に用いるデータセットを，ソフトバンク株式会社で行われたプレゼンテーション研修にて収集した．

2. 関連研究

2.1 プレゼンテーションスキルの推定

Torsten ら [Torsten 15] は，Public speaking skill の評定値と相関関係にあるプレゼンテーションから得られる非言語情報を分析し，それらを推定するモデルを構築した．相関係数 0.75 の予測精度を得た．また，音声と動作の特徴量を両方用いた場合に予測精度が向上することも示した．Vikram ら [Vikram 15] は，プレゼンテーションの音声情報や話者の表情，動作などの時系列マルチモーダル情報から特徴量を抽出し，個人のプレゼンテーション能力の評定値を推定するモデルを構築した．ある動作の後に別の動作を行うなどの時間的な共起パターンを共起ヒストグラムによって特徴量として抽出し，その特徴量がプレゼンテーション能力の評価にどれだけ貢献するのか分析した．相関係数で最大 0.69 の予測精度を得た．前者の研究は，プレゼンテーション能力に関係する特徴量の分析を，後者の研究は，プレゼンテーション能力の向上支援のためのシステムを構築している．[Torsten 15, Vikram 15] の目的は本研究の目的に類似する点はあるが，いずれも非言語情報のみを用いており，本研究のようにプレゼンテーション中の言語内容が重要視されるビジネスプレゼンテーションを対象としていない．

福島ら [福島 16] は，TED のウェブサイトで公開されている 1646 本のプレゼンテーション動画に対して視聴者が抱く印象を予測するモデルを構築した．プレゼンテーション動画から得られた言語情報と音声情報を用いて，複数の視聴者がアノテーションした beautiful，confusing などの合計 14 の印象の中から特定の印象を抱くか否かの 2 クラス分類を行い印象の予測を行った．分類の最大精度は 97.0%であった．この研究はプレゼンテーションの印象を予測することが目的であり，本研究のようにプレゼンテーション能力の推定を目的としていない．

2.2 本研究の位置づけ

関連研究と本研究の相違点をまとめ，本研究の貢献を明らかにする．

言語・非言語両面を考慮した評価尺度の生成

本研究はビジネスプレゼンテーションの評価を対象としている．ビジネスプレゼンテーションでは，プレゼンテーションの内容が重要視される．したがって，ジェスチャーや音声を考慮した評価尺度に加えて，言語面を重視した新たな評価尺度を作成した．

言語・非言語情報の利用

[Torsten 15, Vikram 15] では主に非言語情報に基づき分析を行っている．一方で言語的な能力を推定するためには，プレゼンテーションにおける発話内容は重要である．本研究では非言語特徴量だけでなく，論理性やプレゼン内容の評定を予測するために，発話内容に含まれる名詞数などの言語特徴量も用いて，評定値を推定する機械学習モデルを構築する．

連絡先: 八木　悠太朗，北陸先端科学技術大学院大学，s1810184@jaist.ac.jp

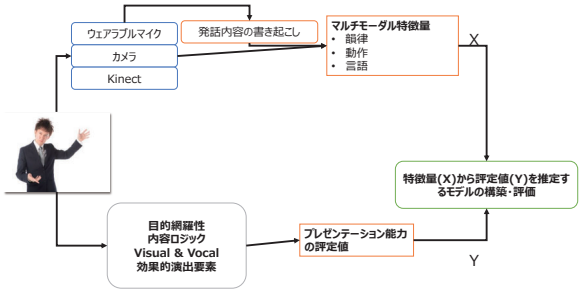


図 1: プレゼンテーション能力を推定する枠組み

表 1: 抽出する言語特徴量

特徴量	詳細
名詞数	発話内容に含まれる名詞数
動詞数	発話内容に含まれる動詞数
接続詞数	発話内容に含まれる接続詞数
フィラー数	発話内容に含まれるフィラー数
感動詞数	発話内容に含まれる感動詞数
同一名詞繰り返し最大数	同一の名詞を繰り返した最大数
同一動詞繰り返し最大数	同一の動詞を繰り返した最大数
名詞出現種類数	発話内容に含まれる名詞の種類数
動詞出現種類数	発話内容に含まれる動詞の種類数

3. 方法

3.1 概要

本研究の概要を図 1 にまとめる．収集したプレゼンテーションデータから音声・動作・言語に関する特徴量を抽出し，個人のプレゼンテーション能力を推定する機械学習モデルを構築し，評価する．推定モデルの構築に用いるデータセットは，ソフトバンク株式会社で行われているソフトバンクユニバーシティ「プレゼン話し方研修」で収集した．収集したデータセットに対して，ソフトバンク株式会社の人事部にプレゼンテーション能力の評価を依頼した．本章では，データセットの概要，抽出するマルチモーダル特徴量，および評価項目について述べる．

3.2 データセット

本実験で使用するデータセットの概要を説明する．データセットには，ウェアラブルマイクから収音した音声情報，Kinect から収音した音声情報，Kinect から得られた深度情報，Kinect から得られた各関節の位置情報，音声情報から文字起こしした発話内容が含まれる．このうち本研究では，ウェアラブルマイクから収音した音声情報，Kinect から得られた骨格の位置座標，音声情報から文字起こしした発話内容を用いてマルチモーダル特徴抽出を行う．実験には 58 人が参加し，一人一回 3 分間のプレゼンテーションを行う．実験で得られた合計 58 個のプレゼンテーションデータを実験に用いる．

3.3 特徴量

[岡田 16] を参考に韻律特徴量，言語特徴量，動作特徴量を抽出する．

3.3.1 韻律特徴量

ウェアラブルマイクから収音した音声情報から発話区間を推定し，OpenSMILE[*1] を用いて特徴抽出を行った．収音した音声のエネルギーの閾値処理に基づいた発話区間推定を用いた．The INTERSPEECH 2010 Paralinguistic Challenge feature set に含まれる 1582 次元の特徴量を抽出した．これらの特徴量には音の大きさ，0 次〜14 次メル周波数ケプストラム係数，0 次〜7 次メル周波数帯の対数パワー，8 つの LPC 係数から算出された 8 つの線スペクトラムペア周波数，平滑化基本周波数，最後の基本周波数の有声確率，及びそれらの一次微分が含まれる．

3.3.2 言語特徴量

音声情報から文字起こしした発話内容のテキストデータに対して形態素解析を行い，単語の品詞情報を取得した．形態素解析には MeCab[*2] を用いた．得られた単語の品詞情報をもとに，抽出した言語特徴量を表 1 にまとめる．

3.3.3 動作特徴量

Kinect から得られた各関節 (上半身 19 部位) の位置座標の時間変化を計算して，動作量と速度，加速度についてそれぞれ平均と標準偏差を求めた．また，発話中についても同様の特徴量を抽出した．

3.4 プレゼン研修講師による評定

プレゼンテーション評定の熟練者として，ソフトバンク株式会社のソフトバンクユニバーシティの講師二名が評定を行った．研修に参加した 58 名のプレゼンテーション映像を閲覧し，15 個の評価項目について，最低 0〜最高 2 の 3 段階で評定した．それらを 4 つの大項目に分類し，各項目の平均値を算出したものを評定値として用いた．評価項目とその定義を以下にまとめる．

目的網羅性

「誰に」,「何を」,「どうしてほしい」が伝えられている．

内容ロジック

「結論」,「根拠」,「相手の利益」が述べられている．

Visual & Vocal

「抑揚」,「声量」,「アイコンタクト」,「ジェスチャー」,「表情」が適切に用いられている．

効果的演出要素

「強調」,「繰り返し」,「具体表現」,「双方向性」が適切に用いられている．

各評定者による評定値の一致度を計算し，各項目に関する信頼性を確認した．一致度の計算には，カッパ値と重み付きカッパ値を算出した．ある項目に関してカッパ値が 0.4 以上であれば各評定者による参加者への評定値は概ね一致しており，その項目の評定値は信頼性が高いと言える [Landis 77]．

表 2 に項目ごとのカッパ値と重み付きカッパ値を示す．図中では，小数点以下第三位を四捨五入した値を記載している．全項目でカッパ値は 0.4 を上回っており，項目としての信頼性は高いと考える．

4. 評価実験

回帰予測タスクを行い，各参加者の評定値 (連続値) の推定精度を評価する．

*1 https://www.audeering.com/opensmile/

*2 http://taku910.github.io/mecab/

表 2: 各評定者間での評定値の一致度

能力項目	重み付きカッパ値	カッパ値
目的網羅性	0.99	0.81
内容ロジック	0.89	0.81
Visual & Vocal	0.91	0.54
効果的演出要素	0.99	0.93

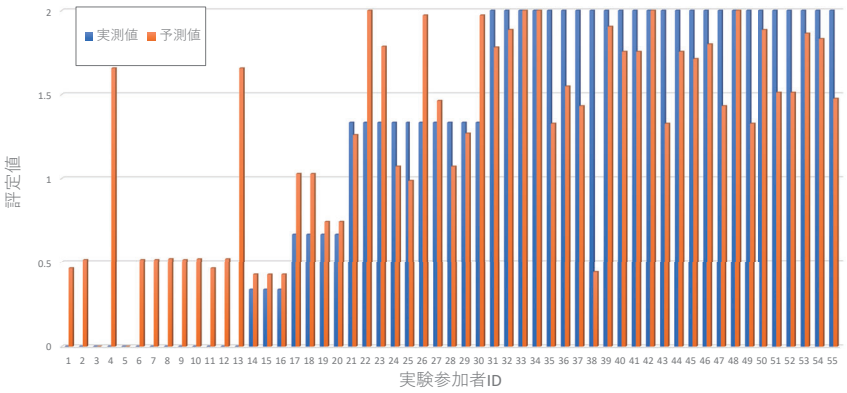


図 2: 「$All\,(p < 0.01)$」のモデルを用いた「目的網羅性」の評定値の予測値と実測値

4.1 推定モデルの構築手順

参加者の特徴量セットを入力 X，そのプレゼンテーションの評定値を出力 Y として，この入出力関係を学習する．合計で 58 個のデータを利用できる予定であったが，3 データの特徴量が機器の不具合などで一部欠損していたため，これらを除外し合計 55 個のデータを用いた．評価には一つ抜き (一人抜き) 交差検定法を利用した．ある参加者から得られたデータをテスト，それ以外の 54 個のデータを訓練に用いて回帰モデルの評価実験を行った．

回帰学習の概要

回帰学習には，RBF カーネルを用いた非線形サポートベクトル回帰モデル (Support Vector Regression:SVR) を用いた．SVR における誤差関数には，ϵ 許容誤差を用いた．誤差の許容度を調整する ϵ，損失関数とマージンの大きさの間のトレードオフを調整するパラメータである C，RBF カーネルの形状を調整するパラメータである γ に関して，それぞれ [0,0.01,0.1,1],[0.01,0.1,1,5,10]，[0.0001,0.001,0.01,0.1] の探索域を設け，訓練データを用いた交差検定を通じて，グリッドサーチにより最適パラメータを決定した．

回帰モデルの推定性能の評価指標にはテストデータに対する決定係数 R^2 を用いた．テストデータに対する評定値 (正解) と推定値の二乗誤差を，テストデータに対応する評定値の分散で割った値を E とすると，決定係数 R^2 は $1 - E$ で計算される．したがって，誤差が小さいほど R^2 は大きくなり，良い結果を示す．

4.1.1 比較対象の特徴量セット

各モダリティの推定性能への寄与を検証するため，以下の 9 種類の特徴量セットを準備し，推定精度の比較を行う．韻律，言語，動作の特徴量セットを A，L，M とそれぞれ略記する．

1. A：韻律特徴量
2. L：言語特徴量
3. M：動作特徴量
4. $A + L$：韻律と言語
5. $A + M$：韻律と動作
6. $L + M$：言語と動作
7. All：全ての特徴量 $(A + L + M)$
8. $All\,(p < 0.05)$：相関分析で有意な相関があった特徴量
9. $All\,(p < 0.01)$：相関分析で有意な相関があった特徴量

4.2 実験結果

プレゼンテーション能力を構成する 4 つの要素項目の評定値を推定する回帰タスクの結果を表 3 に示す．表 3 は各特徴量セットを用いて，各評価項目を推定した結果を表している．推定精度にはテストデータに対する決定係数 R^2 を用いた．

§1 単一モダリティ特徴量セットのモデルの精度

単一モダリティの特徴量セットを用いたモデルに関して「目的網羅性」,「内容ロジック」で「韻律特徴量 (A)」のモデルの精度が最大であった．「Visual & Vocal」では，「動作特徴量 (M)」のモデルの精度が最大であった．「効果的演出要素」では，「言語特徴量 (L)」のモデルの精度が最大であった．「目的網羅性」,「内容ロジック」,「Visual & Vocal」,「効果的演出要素」の最大推定精度はそれぞれ 0.32，0.07，0.05，0.00 であった．

§2 マルチモーダル特徴量セットのモデルの精度

マルチモーダル特徴量セットを用いたモデルに関して「目的網羅性」,「効果的演出要素」で「$All\,(p < 0.01)$」のモデルの精度が最大であった．「内容ロジック」で「$All\,(p < 0.10)$」のモデルの精度が最大であった．「Visual & Vocal」で「$All\,(p < 0.05)$」のモデルの精度が最大であった．「目的網羅性」,「内容ロジック」,「Visual & Vocal」,「効果的演出要素」の最大推定精度はそれぞれ 0.59，0.51，0.38，0.29 であった．

「目的網羅性」,「内容ロジック」,「Visual & Vocal」,「効果的演出要素」の要素項目に関して，マルチモーダル情報の統合は，単一モダリティ特徴量セットのモデルの精度を向上させるために役立つ．

推定精度が最大であった特徴量セットを以下にまとめる．「目的網羅性」,「効果的演出要素」の推定には $All\,(p < 0.01)$ が有効であった．「内容ロジック」の推定には $All\,(p < 0.10)$ が有効であった．「Visual & Vocal」の推定には $All\,(p < 0.05)$ が有効であった．

5. 考察

単一モダリティ特徴量セットの精度よりも，複数のモダリティを統合したマルチモーダル特徴量セットのモデルの精度のほうが高くなる傾向にあった．これは [Torsten 15] の結果とも一致している．

韻律特徴量について，今回は集団での研修の場においてデータ収集を行ったため，ウェアラブルマイクで収音した音声データにノイズが含まれている場合があった．実用化に向けては，音声認識を用いて自動的に発話内容から言語特徴量を抽出する必要がある．本研究で用いた音声データに音声認識を用いて得

表 3: SVM を用いた回帰タスクの結果：各特徴量セットを用いて各プレゼンテーション能力項目を推定した場合の精度を示している．精度はテストデータに対する決定係数 R^2 を用いた．

	目的網羅性		内容ロジック		Visual & Vocal		効果的演出要素	
	R^2	MSE	R^2	MSE	R^2	MSE	R^2	MSE
A	0.32	0.47	0.07	0.57	-0.11	0.20	-0.09	0.21
L	-0.69	1.18	-0.11	0.67	-0.21	0.22	0.00	0.19
M	-0.23	0.86	-0.36	0.82	0.05	0.17	-0.34	0.26
$A + L$	0.31	0.48	0.07	0.57	-0.12	0.20	-0.04	0.20
$A + M$	-0.16	0.81	0.22	0.47	0.16	0.15	-0.17	0.23
$L + M$	-0.25	0.87	-0.18	0.71	-0.31	0.24	-0.29	0.25
All	-0.17	0.81	0.21	0.48	0.17	0.15	-0.17	0.23
$All\,(p < 0.10)$	0.54	0.32	**0.51**	0.30	0.23	0.14	0.24	0.15
$All\,(p < 0.05)$	0.47	0.37	0.39	0.37	**0.38**	0.11	-0.14	0.21
$All\,(p < 0.01)$	**0.59**	0.28	-0.32	0.80	0.13	0.16	**0.29**	0.14

られた言語情報から言語特徴量を抽出した場合の推定精度についても調査する必要がある．

動作特徴量について，本研究では体の各部位の動作量や速度，加速度を特徴量として抽出した．しかし「Visual & Vocal」の評価項目では視線や表情も考慮されているため，視線や表情に関する特徴量を加えモデルを構築する必要がある．

言語特徴量について，本研究では音声情報から文字起こしした発話内容のテキストデータに対して形態素解析を行い，得られた単語の品詞情報から言語特徴量を抽出している．また学習の前に相関分析を行い，有意な相関がある特徴量のみから各評価項目の評定値を推定した場合に最も高い推定精度が得られた．言語面の評価項目である「目的網羅性」,「内容ロジック」に対して，名詞出現種類数，動詞出現種類数は $p < 0.10$ で有意な相関を示した.「効果的演出要素」では，名詞数，動詞数，名詞出現種類数，動詞出現種類数の特徴量が $p < 0.05$ で有意な相関を示した．単語レベルでなく，構文レベルの解析を行い,「目的網羅性」,「内容ロジック」の項目の推定に寄与する特徴量を抽出することが今後の課題の一つである．

図 2 は予測精度が最大であった「$All\,(p < 0.01)$」のモデルを用いた「目的網羅性」の評定値の予測値と実測値を比較している．評定値が 0 点〜1 点，1 点〜2 点の場合，推定誤差の平均がそれぞれ 0.16，0.35 であるのに対し，0 点の場合の推定誤差は 0.60 であった．この結果より，評定値が 0 のデータの推定精度が低いことがわかった．このようなデータに対しても高い精度で推定を行える機械学習モデルを構築する必要がある．

6. 結論

本研究では，プレゼンテーションデータから得られたマルチモーダル情報から「プレゼンテーション能力」を推定するモデルの構築・評価を行った．発話内容に含まれる品詞情報，韻律情報，動作量をマルチモーダル特徴量として抽出し，個人のプレゼンテーション能力値を推定するモデルを機械学習により構築した．評価実験の結果,「目的網羅性」,「内容ロジック」,「Visual & Vocal」,「効果的演出要素」に関する評定値に関して，回帰タスクでそれぞれ最大 0.59，0.51，0.38，0.29 の決定係数 R^2 を得た．

本研究では，プレゼンテーション中の言語・非言語特徴量の時系列変化を考慮していない．時系列マルチモーダル情報を用いたモデルの構築は今後の課題である．

7. 謝辞

ソフトバンク株式会社 人事本部の笠井 理恵様，高嶋 直人様，海上 博志様には，データセットの収集，および評定値の作成においてご支援いただきました．心より感謝申し上げます．

参考文献

[岡田 16] 岡田 将吾, 松儀 良広, 中野 有紀子, 林 佑樹, 黄 宏軒, 高瀬 裕, 新田 克己, マルチモーダル情報に基づくグループ会話におけるコミュニケーション能力の推定, 人工知能学会論文誌, 2016, 31 巻, 6 号, p. AI30-E_1-12

[福島 16] 福島 悠介, 山崎 俊彦, 相澤 清晴, 文書と音声解析に基づくプレゼンテーション動画の印象予測, 電子情報通信学会論文誌 D Vol.J99-D No.8 pp.699-708, 2016

[Landis 77] Landis, J.R.; Koch, G.G. (1977). "The measurement of observer agreement for categorical data". Biometrics. 33 (1): 159–174.

[Torsten 15] Torsten Wörtwein, Mathieu Chollet, Boris Schauerte, Louis-Philippe Morency, Rainer Stiefelhagen, and Stefan Scherer. 2015. Multimodal Public Speaking Performance Assessment. In Proceedings of the 2015 ACM on ICMI (ICMI '15).

[Mathieu 16] Mathieu Chollet, Helmut Prendinger, and Stefan Scherer. 2016. Native vs. non-native language fluency implications on multimodal interaction for interpersonal skills training. In Proceedings of the 18th ACM on ICMI (ICMI 2016).

[Vikram 15] Vikram Ramanarayanan, Chee Wee Leong, Lei Chen, Gary Feng, and David Suendermann-Oeft. 2015. Evaluating Speech, Face, Emotion and Body Movement Time-series Features for Automated Multimodal Presentation Scoring. In Proceedings of the 2015 ACM on ICMI (ICMI '15).

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

同時計測されたEEG信号からの2名における感情共有の測定

Measuring affective sharing between two people from simultaneously recorded EEG signals

木下 泰輝 Taiki Kinoshita　田中 宏季 Hiroki Tanaka　吉野 幸一郎 Koichiro Yoshino　中村 哲 Satoshi Nakamura

奈良先端科学技術大学院大学　先端科学技術研究科
Nara Institute of Science and Technology, Graduate School of Science and Technology

Empathy has been studied intensively in the field of human-machine interaction and human-human interaction. In many of these studies, empathy is subjectively measured by questionnaire. We propose a method to directly measure human empathy from neural potential activity. In this research, EEG signals were recorded simultaneously from two subjects engaged in emotional communication. After time frequency analysis was performed on the recorded Electroencephalogram (EEG) signals, a cross correlation functions between EEG signals of two subjects were calculated. To investigate the significance of cross correlations, paired t-test and false discovery rate correction were performed. As a result, there was a significant cross correlation in the gamma frequency band in the joy condition. In addition, there was a marginally significant correlation in the sadness condition. From these results we showed the possibility to measure empathy from cross correlation of EEG signals recorded from two subjects.

1. はじめに

共感は，人間のコミュニケーションにおいて重要な要素であり，社会心理学において盛んに研究が行われている．さらに，共感は，人間と人工エージェントのコミュニケーションにおいても重要性が指摘されいる [Złotowski 16]．このため，ヒューマンコンピュータインタラクションやアフェクティブコンピューティングの分野で，人工エージェントに，社会心理学において提案された共感のモデルを取り入れ，より良いコミュニケーションの実現を目指す研究が行われている．しかし，このような研究においては，提案された人工エージェントに対する確立された評価手法が存在せず，さらなる研究が必要とされている [Paiva 17].

共感は，認知的，感情的なメカニズムを含む複雑な現象であり，合意の取れた定義は存在しない．このことが，共感の計測が困難である要因の一つである [Neumann 15]．しかし，多くの共感研究に共通して，共感の要素に感情共有 (Affective sharing) が存在することが指摘されている [Decety 04]．感情共有とは，他者の感情を，自他の区別を伴った状態で経験し，無意識に共鳴することをさす．感情共有には，脳活動や表情の同期現象が伴うことが知られている．本研究では，共感に関係する多様な現象のうち，感情共有を対象とする．

感情共有に対する研究では functional magnetic resonance imaging (fMRI) が頻繁に使用される．しかし，fMRI は大規模な装置を必要とし，また被験者が計測中に身動きを取れないことから，人工エージェントとの自然な対話を評価するなどの、実環境での利用は難しいと考えられる．一方，Electroencephalogram (EEG) は，脳活動を計測する代表的な手法の一つであり，計測が容易であることから，実環境下で利用可能という利点が存在する．以上のことから，本研究では，人工エージェントに対する共感の評価手法を提案する前段階として，2名の人間における感情共有を EEG によって計測することを目的とする．

連絡先: 木下 泰輝，奈良先端科学技術大学院大学，奈良県生駒市高山町 8916-5 情報棟 B712，電話: 0743-72-5264，電子メール: kinoshita.taiki.kk1@is.naist.jp

2. 感情共有の計測方法

人間間の共感研究では，fMRI を用いた感情共有の計測が行われている [Anders 11]．この研究では，2 名の実験協力者を募集し，1 名に送信者と呼ばれる役割を，もう 1 名に受信者と呼ばれる役割を割り当てた．実験において，送信者は感情的な場面を想像し，感情表出を行う課題を行った．一方，受信者は送信者の表情を見て共感する課題を行った．この課題を行っている際の 2 名の脳活動が，2 台の fMRI によって同時計測された．得られた fMRI データに対する分析を行った結果，2 名の脳活動の同期現象が確認された．

一方，多くの人間エージェント間の共感研究では，共感と関連を持つと考えられている主観評価によって間接的に共感が評価されている．例えば，人工エージェントの信頼感，社会的存在感，知的さといった指標が利用されており，直接的な評価手法の必要性が指摘されている [Paiva 17].

本研究では，EEG を利用した感情共有の計測を提案する．このため，fMRI を利用した場合と比較して，実環境に近い状態で感情共有を計測することが可能となる．また，今後人工エージェントの評価に応用することにより，主観評価を利用した場合と比較して，直接的に共感を評価することが可能となることが期待される．

3. EEG による感情共有の測定

3.1 実験協力者

実験協力者は，2 名の女性とした．ここで，実験協力者を 2 名と限定したのは，EEG には個人差が現れることが知られているためである．この個人差の解消は次研究の課題とし，本研究では 2 名の実験協力者のみを対象とする．また，実験協力者を女性としたのは，女性が男性と比較して感情共有を持ちやすいことがが知られているためである [David 80].

3.2 実験手続き

本実験は，[Anders 11] で fMRI を用いて行われた実験が EEG で同様に再現可能か調査することを目的とする．図 1 に実験の概要を示す．実験協力者のうち 1 名が送信者の役割を，もう 1 名が受信者の役割を担当した．これら 2 名のうち，送信

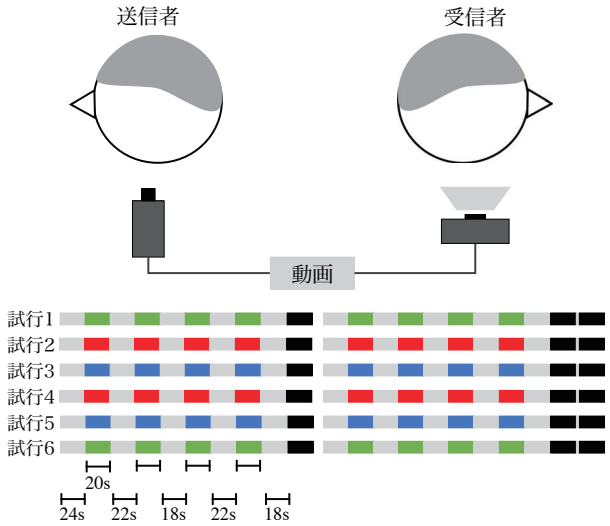


図 1: 実験の概要．各色は，感情の種類を表している（喜び，悲しみ，無感情）．また，黒色は SAM を表している．本研究では，実験を 2 度行い，合計試行回数は 12 回となった．

者は休憩，感情表出という 2 種類の課題を行った．休憩では，教示用モニタに休憩という文字を表示した．この間，送信者はリラックスするよう指示された．一方，感情表出では，3 秒間のカウントダウンを表示した後，喜び，悲しみ，無感情のいずれかの文字を表示した．この間，送信者は，感情を感じる場面をそれぞれ想起し，それらの感情を表情によって表出した．ここで，喜び，悲しみは，[Anders 11] において利用された 5 種類の感情のうち，最も受信者の感情認識精度が高いと報告されている 2 種類の感情である．また，無感情は，喜び，悲しみと比較を行うためのベースラインである．本実験では，1 試行で，休憩と感情表出を交互に 4 回行った．休憩の時間は 18〜24 秒から無作為に決定し，感情表出の時間は 20 秒とした．また，試行の最後には，Self Assessment Mannequin (SAM) [Bradley 94] により，試行中の感情価が 9 段階で回答された．以上の試行が合計 12 回行われた．

一方，受信者には，送信者の顔の表情が Web カメラ（Microsoft 社製 LifeCam Studio）によって，リアルタイムで提示された．受信者はこの映像から送信者の顔の表情を観察し，共感するという課題を行った．また，各試行の最後に，試行中に認識した送信者の感情と，受信者自身の感情をそれぞれ SAM により回答した．

以上の課題を，2 名の実験協力者が同時に行っている間， 2 名の EEG データ・並びに electromyography (EMG) を同時に収録した．ここで，EMG とは，筋電を収録する機器である．本研究では，EEG を装着した状態で表情を動かすため，筋電がノイズとして混入することが予想される．このため，左皺眉筋ならびに左大頬骨筋の EMG を EEG と共に収録し，後処理で EEG データの筋電ノイズを削除した．EEG，EMG は，cognionics 社製 Quick-30 を使用して収録した．ここで，EMG は，EEG チャンネルのうち AF3，FC5 を置き換えることにより収録した．

3.3 EEG データの前処理

前処理は，以下の 6 つの手順で行った．ただし，手順 1)〜4) においては，EEGLAB [Delorme 04] を利用し，手順 5)〜6) においては，MNE [Gramfort 14][Gramfort 13] を利用した．1) サンプリング周波数を 200Hz にダウンサンプリングした．2) 0.3〜45Hz のバンドパスフィルタをかけた．3) 感情の指示が出された点をオンセットとし，-20 秒〜20 秒をエポッキングした．エポッキングにより，各感情について，それぞれ 16 個ずつのエポックが切り出された．4) ノイズ除去を行った．具体的には，独立成分分析を行った後，筋電を除去するため，EMG との相互相関が 0.1 以上である成分を削除した．また，ADJUST [Mognon 11] を利用して，ノイズ成分を自動削除した．さらに，EEG のチャンネルのうち，最も筋電の影響を受けやすいと考えられる T7，T8 は削除した．5) 特徴量の抽出を行うため，時間周波数解析を行った．モルレーウェーブレット変換を行い，各周波数のパワーを算出した．この上で，ガンマ波を 30〜40Hz，ベータ波を 12〜30Hz，アルファ波を 7〜12Hz，セータ波を 4〜7Hz，デルタ波を 1〜4Hz と設定し，各周波数帯でパワーの平均値を算出した．6) 個人差の影響を軽減するため，パワーを各チャンネル毎に平均値 0，標準偏差 1 に標準化した．

3.4 感情共有の分析手法

EEG データに対して行った分析の方法を述べる．多くの研究で，2 名の脳活動の関連性は，EEG チャンネル同士の相関によって計測されている [Babiloni 14]．そこで，本研究では，送信者と受信者の EEG チャンネルの相互相関を，全ての組み合わせに対して算出した．ただし，相互相関のシフトは 1 秒以下に限定し，相互相関が最大となるシフトを選択した．さらに，算出された相関係数が有意であるかを決定するため，休憩期間と感情表出期間の相互相関を対応のある t 検定により比較した．ただし，false discovery rate は Benjamini-Hochberg 法により調整した．以上の分析を行った上で，各チャンネルを頂点とし，有意 ($p < 0.05$) または有意傾向 ($p < 0.1$) な相互相関をもつチャンネル同士に辺を持つグラフを構築した．すなわち，対応のある t 検定の結果によって，相互相関の閾値を設定した．さらに，構築されたグラフに対して，Global efficiency を算出した．Global efficiency とは，グラフ理論において提案された，情報伝達の効率性を表す指標である．Global efficiency の定義は，Average efficiency に基づいている．Average efficiency は，各頂点間の最短経路長の平均値を表しており，式 (1) のように定義される．

$$E(G) = \frac{1}{n(n-1)} \sum_{i \neq j \in G} \frac{1}{d(i,j)} \tag{1}$$

ここで，n は全ノード数を表している．また，$d(i,j)$ は，ノード i と j の最短経路長を表しており，本研究においては，エッジの有無を表している（ノード間に繋がりがあれば 1 を取り，なければ無限大を取る）．この時，Global efficiency は，式 (2) で表される．

$$E_{glob}(G) = \frac{E(G)}{E(G^{all})} \tag{2}$$

ここで，$E(G^{all})$ は，全結合グラフにおいて算出された Average efficiency を表している．本研究においては，送信者のチャンネルと受信者のチャンネルが全結合されたグラフの Average efficiency となる．Global efficiency は 0〜1 の値を取り，値が大きいほど効率性が高いことを意味する．この指標は，EEG 研究においても，2 名の脳活動の繋がりの強さを指標化する目的で利用されている [Ciaramidaro 18]．

4. 結果

まず，SAM に対する分析を行った結果を述べる．送信者の感情価と受信者が認識した感情価を図 2 に示す．送信者の感

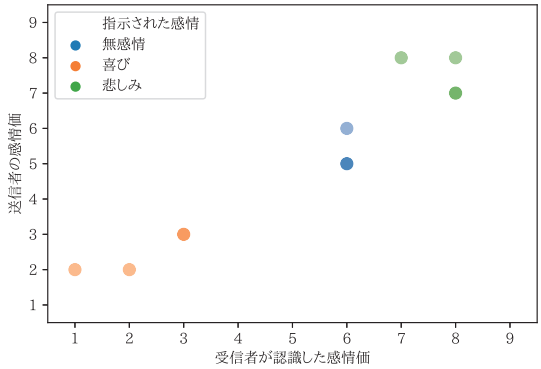


図 2: 送信者の感情価と受信者が認識したが認識した感情価．重なっている点は，色を濃く表示している．

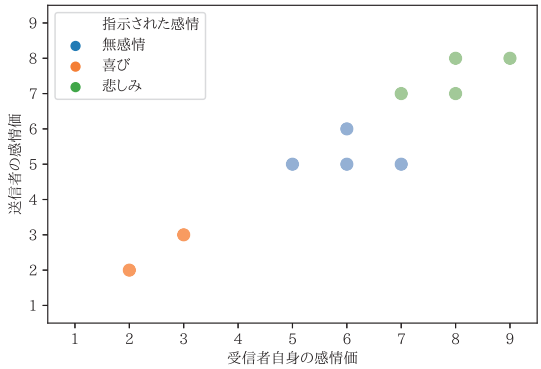


図 3: 送信者の感情価と受信者自身の感情価．重なっている点は，色を濃く表示している．

情価と受信者が認識した感情価の相関係数は 0.954 であり強い正の相関が見られた．次に，送信者の感情価と受信者自身の感情価を図 3 に示す．この時の相関係数は 0.966 であり強い正の相関が見られた．

次に，相互相関に対して，対応のある t 検定を行った結果を述べる．喜びでは，ガンマ波において，28 個の相互相関が有意であった ($p < 0.05$, $n = 16$)．ここで，n は，対象感情のエポック数を表している．その他の周波数帯では，有意な相互相関は見られなかった．図 4 に，t 値が大きかったチャンネルの組み合わせを 10 個示す．線が引かれているチャンネル同士に有意な相互相関が存在した．図 4 に示した通り，送信者のチャンネルと受信者の後頭部に位置するチャンネルに高い相互相関が見られた．一方，悲しみでは有意な相互相関は見られなかった．しかし，ガンマ波において有意傾向が見られた ($p < 0.1$, $n = 16$)．図 5 に，t 値が大きかったチャンネルの組み合わせを 10 個示す．図 5 に示した通り，悲しみにおいても，喜びと同様に送信者のチャンネルと受信者の後頭部に位置するチャンネルに高い相互相関が見られた．一方無感情では，アルファ波において，休憩期間の方が有意に相互相関が高い組み合わせが一つ存在した ($p < 0.05$, $n = 16$)．

最後に，Global efficiency を算出した結果を述べる．有意 ($p < 0.05$) な相互相関を基に Global efficiency を算出した結果，喜びのガンマ波において，0.097 だった．その他の感情，

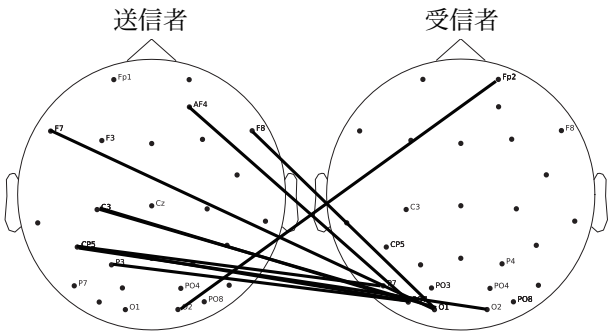


図 4: 喜びにおいて相互相関が有意だったチャンネルの組み合わせのうち，t 値が大きかったチャンネルの組み合わせ 10 個に線が引いた ($p < 0.05$) その他の有意なチャンネルは，チャンネル名のみ示した．

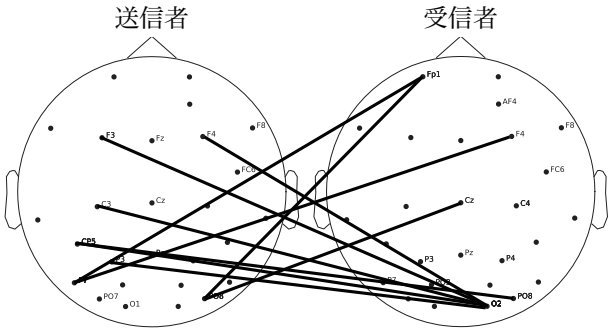


図 5: 悲しみにおいて相互相関が有意傾向だったチャンネルの組み合わせのうち，t 値が大きかったチャンネルの組み合わせ 10 個に線が引いた（$p < 0.1$）．その他の有意傾向なチャンネルは，チャンネル名のみ示した．

周波数帯においては全て 0 だった．次に，有意傾向 ($p < 0.1$) な相互相関を基に Global efficiency を算出した結果を表 1 に示す．表 1 に示した通り喜びのガンマ波，ベータ波，悲しみのガンマ波において 0 以外の値をとった．しかし，その他の周波数帯においては，全て 0 の値をとった．総じて，高周波数帯において Global efficiency が高い傾向が見られた．また，喜び，悲しみにおいて，無感情と比較して Global efficiency が高い傾向が見られた．

5. 考察

SAM に対する分析を行った結果より，送信者の感情が受信者によって正しく認識されていると考えられる．また，送信者の感情表現により受信者に近い感情が誘発され，感情共有が生じている状態で，脳活動の計測を行うことができたと判断できる．

次に，相互相関に対して対応のある t 検定を行った結果より，感情共有による脳活動の同期を，EEG を利用して計測することができる可能性が示された．しかし，悲しみにおいて有意な結果が得られなかったため，試行の回数を増やし再度分析を行うことが必要であると考えられる．

最後に，Global efficiency を算出した結果より，感情共有が生じている時，Global efficiency が高まる傾向が見られた．これにより，この指標を利用することで，感情共有を計測するこ

表 1: 2 名の電極相互相関に基づく Global efficiency

感情	ガンマ波	ベータ波	アルファ波
喜び	0.572	0.385	0.000
悲しみ	0.152	0.000	0.000
無感情	0.000	0.000	0.000

とができる可能性が示された.

6. まとめ

本研究の目的は，感情共有を 2 名同時計測された EEG から計測することである．分析の結果，感情共有が生じている時，送信者のチャンネルと受信者の後頭部に位置するチャンネルの相互相関が高まる傾向が明らかとなった．さらに，相互相関を基に算出された Global efficiency の値が高まる傾向が明らかとなった．以上のことから，EEG の相互相関と Global efficiency によって，二者間の Affecitve sharing を計測できる可能性があると考えられる．今後は，実験の試行数を増やし，統計的に有意な結果を得る必要があると考えられる．

参考文献

[Anders 11] Anders, S., Heinzle, J., Weiskopf, N., Ethofer, T., and Haynes, J.-D.: Flow of affective information between communicating brains., *NeuroImage*, Vol. 54, No. 1, pp. 439–46 (2011)

[Babiloni 14] Babiloni, F. and Astolfi, L.: Social neuroscience and hyperscanning techniques: past, present and future, *Neuroscience & Biobehavioral Reviews*, Vol. 44, pp. 76–93 (2014)

[Bradley 94] Bradley, M. M. and Lang, P. J.: Measuring emotion: The self-assessment manikin and the semantic differential, *Journal of Behavior Therapy and Experimental Psychiatry*, Vol. 25, No. 1, pp. 49–59 (1994)

[Ciaramidaro 18] Ciaramidaro, A., Toppi, J., Casper, C., Freitag, C. M., Siniatchkin, M., and Astolfi, L.: Multiple-Brain Connectivity during Third Party Punishment: An EEG Hyperscanning Study, *Scientific Reports*, Vol. 8, No. 1, p. 6822 (2018)

[David 80] David, M. H.: A multidimensional approach to individual differences in empathy, *JSAS Catalog of Selected Documents in Psychology*, No. 10, p. 85 (1980)

[Decety 04] Decety, J. and Jackson, P. L.: The functional architecture of human empathy., *Behavioral and cognitive neuroscience reviews*, Vol. 3, No. 2, pp. 71–100 (2004)

[Delorme 04] Delorme, A. and Makeig, S.: EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, *Journal of neuroscience methods*, Vol. 134, No. 1, pp. 9–21 (2004)

[Gramfort 13] Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., et al.: MEG and EEG data analysis with MNE-Python, *Frontiers in neuroscience*, Vol. 7, p. 267 (2013)

[Gramfort 14] Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Parkkonen, L., and Hämäläinen, M. S.: MNE software for processing MEG and EEG data, *Neuroimage*, Vol. 86, pp. 446–460 (2014)

[Mognon 11] Mognon, A., Jovicich, J., Bruzzone, L., and Buiatti, M.: ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features, *Psychophysiology*, Vol. 48, No. 2, pp. 229–240 (2011)

[Neumann 15] Neumann, D. L., Chan, R. C., Boyle, G. J., Wang, Y., and Westbury, H. R.: Measures of Empathy: Self-Report, Behavioral, and Neuroscientific Approaches, in *Measures of Personality and Social Psychological Constructs*, pp. 257–289, Academic Press (2015)

[Paiva 17] Paiva, A., Leite, I., Boukricha, H., and Wachsmuth, I.: Empathy in Virtual Agents and Robots, *ACM Transactions on Interactive Intelligent Systems*, Vol. 7, No. 3, pp. 11:1–11:40 (2017)

[Złotowski 16] Złotowski, J., Sumioka, H., Nishio, S., Glas, D. F., Bartneck, C., and Ishiguro, H.: Appearance of a robot affects the impact of its behaviour on perceived trustworthiness and empathy, *Paladyn, Journal of Behavioral Robotics*, Vol. 7, No. 1, pp. 7:55–66 (2016)

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

プロダクトディメンジョンを評価指標としたグループ対話のアウトプット推定

Estimating performance of group interaction based on product dimension

三浦 郷 *1　　岡田 将吾 *1
Go Miura　　Shogo Okada

*1北陸先端科学技術大学院大学
School of information Science, Japan Advanced Institute of Science and Technology(JAIST)

This paper focuses on developing a model for estimating the quality of discussion using multimodal features. For this purpose, we use a group meeting corpus including audio signal data of participants observed in 30 meeting sessions. Also, four annotators watch conversation transcripts and annotate the score about quality of discussion using product dimension which is a sociological criteria. We extracted various kinds of features such as spoken utterances, acoustic features, speaking turns. First, binary (high or low) classification models are trained to infer the annotated score from these features using support vector machine. Second, binary classification models are developed to infer the quality of unknown discussion task. Experiments results show that multimodal model archived 0.92 as the classification accuracy and task independent model archived 0.73.

1. はじめに

意思決定を集団で行う理由の一つは，個人の情報処理能力の限界を克服することである．集団意思決定に関して，社会心理学の領域で研究が行われている．[Ueda 97] では，集団の意思決定は相互作用により認知的刺激を受けるなどメリットがある一方，多数派の同調圧力により少数派の自分の意見を表出できない場合のように，必ずしも個人の意思決定より集団の意思決定の方が良い結果を得るとは限らないと報告されている．集団の意思決定方法やプロセスの質を評価することが出来れば，意思決定方法の見直しや，個人の意思決定結果の尊重など，集団意思決定におけるデメリットを緩和できる可能性がある．本研究では，集団意思決定のための方法の一つであるグループディスカッションを題材として，ディスカッションの質を評価し，質の改善・向上を支援するシステムに応用可能な技術の提案を目指している．

本研究はディスカッションを通じて観測できる参加者個人の発言の仕方，聞き方と発言内容から，その個人が属するグループで行われるディスカッションの質を推定するモデルを機械学習により構築し，評価することを目的とする．本研究で対象とするグループディスカッション (Group Discussion: 以降 GD と略記する) の質を，[Hackman 67] における評価指標に基づき定義する．この評価指標は書き言葉に対して適応可能であり，本研究ではディスカッションの発話内容を書き起こしたテキストを評定者が閲覧し，[Hackman 67] を元に GD の質を評価している．本研究の貢献は以下の 2 点である．

1. **グループアウトプットの推定モデルの構築**
   40 名の参加者による計 9 時間を超える会話データ [林 15] から，評定者 4 名により評価された GD の質を推定するモデルを学習する．評価実験の結果から，最大で 9 割以上の精度で GD の質の高低を推定できることを示す．

2. **タスクの種類に依存しない推定モデルの構築**
   40 名の参加者は，グループごとに「インバスケット型」課題 1 つと,「ケーススタディ型」課題 2 つの課題に取り組んでいる．3 つの課題から学習する組み合わせを変えて，課題に依存せずに GD の質を推定するモデルの構築を目指す．評価実験の結果から，最大で 7 割以上の精度で訓練データに用いていないタイプの GD の質の高低を推定できることを示す．

連絡先: 三浦郷，北陸先端科学技術大学院大学，〒 923-1211
石川県能美市旭台 1 丁目 1， s1810175@jaist.ac.jp

2. 関連研究

[Avci 16] らは，マルチモーダル情報を特徴量に用いて，グループアウトプット (GD の質) を推定するモデルの構築を行った．アウトプットの高・低の 2 値分類タスクにおいて，Coupled Hidden Markov Model (CHMM) を改良した IM (Influence Model) と Support Vector Machine (SVM) を用いたモデルが最大の分類精度を示し，91%であった．[Murray 18] らは，マルチモーダル特徴量に言語に関する情報を含めて，グループアウトプットを推定する回帰予測タスクを行った．データサイズを増加させる転移学習を行い，学習器に Random Forest を用いたモデルが最良の精度を示し，MSE (平均二乗誤差) は 64.4 であった．

[Avci 16] と [Murray 18] で用いられた ELEA コーパス [Sanchez 13] には，Winter survival task と呼ばれるインバスケット型課題にグループで取り組んだ際に得られたマルチモーダルデータが記録されている．この課題はロールプレイング形式となっており，サバイバルを想定した時に必要なアイテム群（水や専門書など）の重要度を議論しランキングする．専門家によるランキング（正解）と，グループで決めたランキングの差を「グループアウトプット（GD の質）」と定義している．一方で，ディスカッションはインバスケット型課題に限らず様々な種類が存在し，正解が明確に決まっていない場合が多い．上記の正解との差によりパフォーマンスを定義する方法では，正解が明確に存在しないタスクに対して適応ができない．そこで，本研究では課題に依存しないグループパフォーマンスの定義方法を 3.1 節で提案する．この評価指標を基にアノテーションしたデータセットを用いて，課題に依存しない GD の質の自動推定の精度を検証する．

3. グループディスカッションコーパス

本研究で用いるグループディスカッションコーパス [林 15, Nihei 14] の説明を行う．以降，このコーパスを GD (Group Discussion) コーパスと呼ぶ．参加者 40 名の内，初対面同士の 4 名を 1 グループとして，合計 10 グループを構成している．各グループは 3 つの課題 [林 15] についてディスカッションに取り組み，ディスカッションの過程で得たデータ群をデータセットとして収集した．結果として，10 グループ × 3 つの課題 × 40 名で，延べ 120 名分の参加者のデータセットが GD コーパスに含まれている．GD コーパスに含む情報の内，本研究ではヘッドセットマイクから観測した音声データを用いて後述のマルチモーダル特徴量抽出を行った．

3.1 ディスカッションの質の定義

表 1: プロダクトディメンジョンの評価指標

y1	その課題への回答は，創造性を有していた．
y2	その課題に対して，適切な方向性，適切な議論，適切な回答が行えていた．
y3	製品（議論の履歴）は，過度な困難はなく，理解可能であり，明快である．
y4	製品（議論の履歴）は，何らかの問題に関しての観点や立場を取ることが出来ていない，または対処できない程度，適切な粒度で問題を捉えられず，対象となる問題の扱いに失敗している．
y5	製品（議論の履歴）は幾つかの考え方，物，事に難色を示すものである．
y6	製品（議論の履歴）には綿密さや，事例，詳細が欠如している．
y7	製品（議論の履歴）は，具体的に何か行うことに関連した方向性を持っていて，実質的な提案，または建設的な内容である．
y8	製品（議論の履歴）は，今後第三者にフォローされるべき，特定または一般的な行動方針を示したり暗示したりしている．
y9	製品（議論の履歴）は，あるアイデア，目的，または出来事に関する特定の信念や意見を述べたり，強く示唆したりしている．
y10	製品（議論の履歴）は特定の観点を押し通そうとしている．
y11	製品（議論の履歴）に含まれるアイデアはユニークで，新鮮で，あまり見たことがない驚きのあるものである．
y12	製品（議論の履歴）の一般的な見通し，態度，調子，方向性は「否定的」ではなく「肯定的」である．
y13	製品（議論の履歴）は断片化していたり，相互の発言に関連性がなかったりして，調和のとれた首尾一貫している状態と真逆である．
y14	製品（議論の履歴）はユニークで，普通ではなく，珍しい．
y15	製品（議論の履歴）は幾つかの考え方，物，事をサポートしている．
y16	製品（議論の履歴）は独りよがりであり，活性化していない．
y17	製品（議論の履歴）は普通で，ありふれていて，よく見るものである．
y18	製品（議論の履歴）内容が簡潔すぎて短い．
y19	製品（議論の履歴）のプレゼンテーションは適切な方法で行われ，調和がとれている．

本研究では課題に依存しないグループパフォーマンスの指標として，Hackman により定義された Product dimension を用いる．表 1 にその評価指標を示す．プロダクトディメンジョンの指標は書き言葉に対して適応可能であり，本研究においては GD における参加者の発話内容を書き起こしたテキストデータを 4 人の評定者が閲覧し，この評価指標に基づき各グループに評定値を付与した．この 4 人に付与された評定値の合計を各グループが行った GD の質 (機械学習における教師ラベル) として定義する．以降，19 個の評価指標をそれぞれ y1，...，y19 と略記する． 評定値のスケールは，y1 が 1 点から 5 点の 5 段階，y2 から y19 は 1 点から 7 点の 7 段階で評価した．ディスカッションのサンプル数は 30 であったが，そのうち 1 つのサンプルに不備があったため，29 のサンプルを実験に用いた．

3.2 アノテーションの一致率の分析

各グループによる GD で観測された発話内容を書き起こし，4 人の評定者がその文章を閲覧して GD の質の評定を行う．評価基準として前述のプロダクトディメンジョンの指標を用いる．

各評定者による評定の一致度を計算し，各評価項目についての信頼性を確認した．一致度の計算に関して，信頼性分析の項目を評定者とし，各項目ごとにクローンバック α 値とクリッペンドルフ α 値を算出した．ある項目に対してクローンバック α が 0.7 以上，クリッペンドルフ α が 0.667 以上であれば各評定者によるグループに対しての評定値は概ね一致しており，その項目の評定値は信頼性が高いと言える．

表 2 は各評価指標におけるクローンバック α，クリッペンドルフ α を示している．各値が正の場合は四捨五入，負の場合は絶対値を四捨五入し，マイナスをつけたものを記載している．各評価指標の値は閾値より低く，十分な一致度は得られなかった．GD の質のアノテーションは主観的であるため，高い一致率が得られなかったと考えられる．この理由の考察は今後の課題とする．

4. マルチモーダル特徴量の抽出

評定者は GD を書き起こした文章を閲覧して評価しているため，GD の質は主に書き起こした発話内容に関連すると考え，発話内容の言語特徴量を抽出する．また，議論中の参加者の非言語特徴量も GD の質に影響を与えると考え，[岡田 16] を参考に発話ターン特徴量，発話中の韻律特徴量を抽出する．

4.1 発話ターン特徴量

合計発話長：発話区間から算出した発話時間を発話長とし，1 セッション単位 (各課題の開始時間から終了時間まで) で発話長の合計を計算する．

合計発話回数：一つの発話区間を発話断片と定義する．発話断片の回数を発話回数とし，1 セッション単位で発話回数の合計を計算する．

合計発話長 (1 秒以上)：1 秒以上の発話断片を抽出し，その長さの総和を計算する．

合計発話回数 (1 秒以上)：1 秒以上の発話断片を抽出し，その発話回数を計算する．

4.2 韻律特徴量

各発話断片における韻律情報の特徴量を抽出する．4.1 節で得た発話区間情報から，各発話断片の特徴量を抽出する．音声情報の信号処理と韻律特徴量抽出には Praat[*1] を用いる．

最大ピッチ，最小ピッチ：各発話断片の最大ピッチ，最小ピッチを計算し，1 セッション中の全発話断片に対して平均値を計算する．

ピッチ平均：各発話断片に関して，発話区間でピッチの平均値を計算し，1 セッション中の全発話断片に対して平均を計算する．

最大インテンシティ，最小インテンシティ：各発話断片の最大，最小音圧 (インテンシティ) を計算し，1 セッション中の全発話断片に対して平均値を計算する．

音圧の幅：最大，最小インテンシティの差を計算する．

抑揚：最大ピッチと最小ピッチの差を計算する．

話速：発話のシラブル (音節) 数を各発話長で除した値を計算する．

4.3 発話内容の言語特徴量

名詞数，動詞数，感動詞数，フィラー数：1 セッション中の発話に含まれる各品詞：名詞，動詞，感動詞の単語の出現回数，フィラーの回数を計算する．

新規名詞数：4.1 節で得た発話区間情報を利用し，各名詞が初

*1 http://www.fon.hum.uva.nl/praat/

表 2: 各評価指標のクローンバック α，クリッペンドルフ α

評価指標	y1	y2	y3	y4	y5	y6	y7	y8	y9	y10	y11	y12	y13	y14	y15	y16	y17	y18	y19
クローンバック α	0.68	0.58	0.68	0.43	-0.03	0.64	0.64	0.34	0.67	0.50	0.60	0.78	0.52	0.63	0.34	0.52	0.62	0.54	0.51
クリッペンドルフ α	0.34	0.26	0.29	0.16	-0.01	0.30	0.30	0.08	0.32	0.20	0.27	0.43	0.19	0.28	0.12	0.15	0.27	0.16	0.20

表 3: 各特徴量セットを用いた場合の推定精度の結果．太字は正解率 0.7 以上を示す．

	評価指標									
	y1	y2	y3	y4	y5	y6	y7	y8	y9	y10
A	0.52	0.27	0.52	0.32	0.48	0.48	0.56	0.68	0.36	0.48
S	0.36	0.50	0.48	0.18	0.52	**0.80**	0.56	**0.76**	0.61	**0.74**
L	**0.76**	0.00	0.56	0.00	0.56	**0.72**	0.44	0.68	0.61	0.67
A + S	0.48	0.23	0.48	0.18	0.56	**0.76**	0.60	**0.88**	0.14	0.33
A + L	0.32	0.00	0.48	0.00	0.56	0.56	0.68	**0.92**	0.39	0.22
S + L	0.48	0.12	0.56	0.00	0.56	**0.72**	0.64	**0.72**	0.64	0.59
A + S + L	0.24	0.00	0.52	0.18	0.56	**0.76**	**0.76**	**0.92**	0.04	0.30

	評価指標								
	y11	y12	y13	y14	y15	y16	y17	y18	y19
A	0.52	0.54	0.44	0.23	0.48	**0.70**	0.44	0.26	0.42
S	0.59	0.50	0.52	0.58	0.36	0.29	0.56	0.67	0.35
L	**0.70**	0.54	0.56	0.61	0.24	0.00	0.68	**0.70**	0.42
A + S	0.63	0.63	0.52	0.61	0.36	0.56	0.56	0.30	0.35
A + L	0.59	0.50	0.40	0.61	0.28	0.59	0.60	0.41	0.42
S + L	**0.78**	0.50	0.56	0.46	0.20	0.26	0.52	**0.70**	0.31
A + S + L	0.59	0.63	0.52	0.65	0.20	0.37	0.60	0.41	0.42

めて発言された時間，発言した参加者を特定し，参加者ごとに数え上げた回数を新規名詞数とする．新規名詞は，他の参加者に対して新しい提案をする時に使用される傾向がある．

4.4 グループ特徴量への変換

発話ターン，韻律，言語の個人特徴量をグループ特徴量に変換する．4 人の参加者の特徴量の最大値，最小値，最大値と最小値の差分，平均値，標準偏差，合計値の計 6 つを計算して各グループの特徴量とする．変換前の個人の特徴量は前述した 17 個であり，グループ特徴量に変換後はそれぞれ 6 つずつのグループ値を計算するため計 102 個となった．

5. 評価実験

中央値を基準にグループの評定値を高群・低群に分類し，2 クラス分類タスクを行い推定精度を評価する．評価実験は実験 1 と実験 2 の 2 つ行う．実験 1 では，学習された分類モデルのマルチモーダル特徴量の組み合わせと推定精度を分析することで，GD の質を推定するために有効な特徴量を明らかにする．実験 2 では，3 つの問題解決課題ごとに分けた特徴量の組み合わせと推定精度を分析することで，未知の課題に対するモデルの推定精度を評価する．

5.1 実験手順

評価には 1 つの GD データを抜いた交差検定法を使用した．あるグループ i から得られたデータをテスト，それ以外の 28 グループから得られたデータを訓練に用いて分類モデルの評価実験を行った．

二値分類タスクの準備として，評定値を二値に変換する際，評定値のうち中央値より高いものを高群，低いものを低群とし，中央値にあたる 1 つのデータサンプル（1 グループ）は取り除いた．分類モデルの推定性能には全テストデータに対する正解率の平均を用いた．分類学習には，RBF カーネルを用いたサポートベクトルマシン (Support Vector Machine: SVM) を用いた．RBF カーネルのバンド幅 γ を [1/120, 1/119, 0.01, 0.1, 1, 5, 10]，SVM における損失とマージンの大きさの間のトレードオフを調整するパラメータである C を [0.01, 0.1, 1, 5, 10] の範囲で探索し，最適パラメータをテストに用いた．

各モダリティの推定精度への寄与を検証するため，単一モダリティとマルチモーダル特徴量セットを計 7 種類準備し，推定精度の比較を行う．「韻律」,「発話ターン」,「言語」の特徴量セットをそれぞれ A，S，L とする．また，マルチモーダル特徴量セットとして,「韻律と発話ターン」を $A + S$,「韻律と言語」を $A + L$,「発話ターンと言語」を $S + L$,「韻律と発話ターンと言語」を $A + S + L$ と表す．

5.2 2 種類の分類タスク実験

実験 1: 各種モダリティを用いた推定精度の評価

GD の質を推定するために有効な特徴量セットを明らかにするため，特徴量セットの組を変えてモデルを構築し，推定精度を比較した．特徴量セットは 5.1 節で記述した 7 つの組を用いた．

実験 2: 未知の課題に対する推定精度の評価

課題の種類に依存しない，未知の課題に対する推定精度を評価するための実験を行う．ディスカッションデータ全 29 サンプルの内，課題 1 は 9 サンプル，課題 2，3 は 10 サンプルずつ保有している．課題 1，2，3 の内，2 つの課題を訓練データに，残り 1 つの課題をテストデータに用いる．テストデータに使用する課題のデータ 10 サンプル (課題 1 は 9 サンプル) について，1 サンプルずつテストに用いて推定し，10 サンプルの推定精度の平均を算出した．

6. 評価実験結果

実験 1 の各種モダリティごとの推定結果を表 3 に示す．表 3 では，各評価指標において，単一モダリティとマルチモーダル特徴量セット 7 種類を用いて推定した時の正解率を示している．実験 2 の未知の課題に対する推定結果を表 4 に示す．表 4 では，各評価指標において，訓練に用いていない各課題のデータをテストに用いた場合の正解率と，その正解率を平均した結果を示している．

実験 1 の結果:

各種モダリティごとの推定結果について，評価指標 y8「行動方針の示唆」における「韻律 + 言語 $(A + L)$」,「韻律 + 発話ターン + 言語 $(A + S + L)$」のモデルの推定精度が最大となり，0.92 であった．評価指標 y8 において単一モダリティよりマルチモーダル特徴量を用いたモデルが有効であることが示された．また，推定精度が 0.7 以上であるモデルが存在したのは，評価指標 y1「創造性」，y6「綿密さの欠如」，y7「建設的な内容」，y8「行動方針の示唆」，y10「多面的な観点の欠

表 4: 各評価指標における推定精度の結果．太字は正解率が 0.7 以上であることを示す．各指標において，0.5 を超える正解率が存在しないものは割愛した．

	テストデータ			
	課題 1	課題 2	課題 3	平均
y1: 創造性	0.39	0.0	0.61	0.33
y6: 綿密さの欠如	0.60	0.41	0.24	0.40
y7: 建設的な内容	**0.70**	0.0	0.06	0.25
y8: 行動方針の示唆	**0.90**	0.35	**0.94**	**0.73**
y11: 驚くべき内容	0.61	0.47	0.67	0.58
y12: 肯定的な態度	**0.70**	0.00	0.00	0.23
y14: 独自性	**0.85**	0.33	0.44	0.54
y15: 意見への支持	0.00	**0.81**	0.00	0.27
y16: 独りよがりな内容	0.55	0.00	0.00	0.18
y17: ありふれた内容	0.65	0.44	0.38	0.49
y18: 情報の不足	0.39	**0.89**	0.31	0.53
y19: プレゼンの適切性	0.39	0.33	**0.71**	0.48

如」，y11「驚くべき内容」，y16「独りよがりな内容」，y18「情報の不足」であった．また，全 19 個の指標のうち 15 個の指標で，2 クラス分類タスクのチャンスレベル 0.5 を上回るモデルが存在した．

実験 2 の結果:

未知の課題に対する推定結果について，実験 1 と同様に評価指標 y8「行動方針の示唆」において精度が最大となり，0.73 であった．評価指標 y8 に関して，タスクに依存しないモデルを構築できることが示唆された．また，平均する前の課題別の正解率について，2 クラス分類のチャンスレベル 0.5 を上回ったのは，表 4 に示した 12 個の評価指標であった．その他の指標については精度が 0.5 に満たず，十分な精度には至らなかった．推定精度を上げるために今後検討を行う．

7. 考察

先行研究である [Avci 16] では，GD の質のレベルの高低を最大 0.91 の精度で分類可能であることが示された．コーパス参加者や使用言語，評価基準が異なるため，本研究と上記の先行研究の結果とを直接比較することはできないが，評定値の 2 クラス分類タスクとして同様に扱うと，最大 0.92 と同程度の精度でモデル構築可能であることが示された．しかし，正解率 0.92 を示したのは評定の一致度について信頼性が高くなかった評価指標であり，考慮に入れる必要がある．

実験 1 の結果から，プロダクトディメンジョンの評価指標 y1「創造性」，y6「綿密さの欠如」，y7「建設的な内容」，y8「行動方針の示唆」，y10「多面的な観点の欠如」，y11「驚くべき内容」，y16「独りよがりな内容」，y18「情報の不足」において，各種モダリティを用いた推定が可能であることが示された．すなわち，議論の質をプロダクトディメンジョンの評価指標に基づき定義すれば，「ケーススタディ型」課題などの正解が明確に決まっていない議論の質を特定の観点からは推定可能であることが期待できる．実験 2 の結果から，評価指標 y8「行動方針の示唆」においては課題に依存しないモデル構築が可能であることが示された．しかし，テストデータに用いた課題によって推定精度にばらつきがあること，評価指標 y8 以外ではモデルの精度が芳しくないことは今後の課題である．

今回行ったアノテーションは，4 人の評定者の主観に依るところが大きいため，評定の一致度が低くなったと考えられる．今後，4 人の評定者の中で誰の評定が大きく異なっているか詳細を分析し，アノテーションの一致率が高い評定者を選定して信頼性の高いモデルを構築することを目指す．また，評定者の人数を増やすと評定の一致率がどう変化するか傾向を分析することも今後の課題である．

8. 結論

本研究では，グループ対話の参加者から得られるマルチモーダル情報からディスカッションの質を推定するモデルの構築・評価を行った．発話ターン，韻律，言語といった情報をマルチモーダル特徴量として抽出し，プロダクトディメンジョンの評価指標に基づき評定された各グループの GD の質を推定するモデルを機械学習により構築した．評価実験の結果，評価指標 y8「行動方針の示唆」に関して，高群・低群の 2 クラス分類タスクで最大 0.92 の精度を得た．また，本実験で対象とした 3 つの課題の内，2 つの課題のデータを学習したモデルを用いて，学習に用いなかった 1 つの課題のデータに対して推定した精度の平均は，評価指標 y8 において最大精度の 0.73 を得た．

参考文献

[Avci 16] Umut Avci, et al. ”Predicting the Performance in Decision-Making Tasks: From Individual Cues to Group Interaction.” in IEEE Transactions on Multimedia, vol. 18, no. 4, pp. 643-658 (2016)

[Murray 18] Gabriel Murray, et al. ”Predicting Group Performance in Task-Based Interaction” In Proceedings of the 20th ACM International Conference on Multimodal Interaction, pp. 14-20, ACM (2018)

[Hackman 67] Hackman, J. Richard, et al. ”A set of dimensions for describing the general properties of group-generated written passages” Psychological Bulletin vol. 67, no. 6, pp. 379-390 (1967)

[Sanchez 13] Sanchez-Cortes, Dairazalia, et al. ”Emergent leaders through looking and speaking: from audio-visual data to multimodal recognition” Journal on Multimodal User Interfaces, vol. 7, no. 1-2, pp. 39-53 (2013)

[林 15] 林佑樹, 他 ”グループディスカッションの構築および性格特性との関連性の分析” 情報処理学会論文誌, Vol. 56, No. 4, pp. 1217-1227 (2015)

[岡田 16] 岡田将吾, 他 ”マルチモーダル情報に基づくグループ会話におけるコミュニケーション能力の推定” 人工知能学会論文誌, vol. 31, no. 6, pp. AI30-E_1-12 (2016)

[Nihei 14] Fumio Nihei, et al. ”Predicting Influential Statements in Group Discussions using Speech and Head Motion Information” In Proceedings of the 16th ACM International Conference on Multimodal Interaction, pp. 136-143, ACM (2014)

[Ueda 97] 上田泰，”個人と集団の意思決定-人間の情報処理と判断ヒューリスティックス-” 文眞堂 (1997)

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

マルチモーダル情報に基づくユーザの興味度推定のための対話履歴の効果の検証

Use of Dialogue History for User's Interest Level Estimation Based on Multimodal Information

西本 遥人 Haruto Nishimoto　武田 龍 Ryu Takeda　駒谷 和範 Kazunori Komatani

大阪大学 産業科学研究所
The Institute of Scientific and Industrial Research (ISIR), Osaka University

Estimation of a user's mental states including his/her interest level for the current topic is a key component to realize user-adaptive multimodal dialogue systems. Multimodal sensing provides helpful information for that, but its results are often incorrect because the user's subtle behaviors need to be captured. We incorporate a dialogue history when estimating the user's interest level to make the results more stable. Specifically, we formulate the estimation by using the partially observable Markov decision process (POMDP). The multimodal sensing results are used as likelihoods of input observation and the estimated interest levels are tracked as states. Experimental results showed the estimation performance by considering the dialogue history. We used two kinds of annotation results for each exchange: the interest levels and the next desirable system actions, which were given by multiple annotators. We also confirmed a correlation between the estimated interest levels and the next desirable system actions.

1. はじめに

対話における人の発話には，その言語内容だけでなく，様々な有用な情報が含まれる．対話分析においては，声の韻律情報，顔の表情などの非言語情報が重要になる [1]．本研究では，人対人ではなく，人対システムの対話において，ユーザの心的状態の一種である興味度をマルチモーダル情報を用いて推定する．そしてその推定結果を考慮した応答生成が可能なシステム構築を目標としている．

ユーザの興味を反映してシステムが応答する場合，ユーザのシステムに対する印象がよくなると考えられる．例えば，ユーザが興味を持たない場合，システムは対話を別の話題に切り替えることができる．一方興味を持つ場合，対話中の話題を掘り下げたり，その話題に関連する情報を提示できる．本研究では，図 1 に示すようにシステムとユーザの発話対（一交換）ごとに興味度を推定することで，その推定結果に応じて適応的に発話を変更するシステムの構築を目指している．

対話中のユーザの興味度が発話ごとに大きく変化するとは考えにくいため，対話中の過去の状態を考慮することは有用である．例えば，興味ありの状態が長く続いているユーザに対し，システムがこれまでと同じ話題を継続する発話をすると，ユーザの興味度が急に下降するとは考えにくい．このため，対話履歴を考慮して，発話ごとの推定性能の向上を狙う．マルチモーダル情報を用いた興味度の推定は，ユーザの微妙な特徴を捉える必要があるため，各発話の情報だけから必ずしも正しい推定結果が得られるわけではない．実際，我々の以前の報告 [2] では，各発話対から得られる情報のみを用いて推定を行ったところ，7 割程度の正解率であった．システムが誤った推定結果を用いて次発話を行った場合，ユーザは期待しない発話を受け取ることになり，ユーザのシステムへの印象は悪化する．つまり，対話において，一発話だけからの興味度推定結果を用いて応答を行うのは適切ではない．

連絡先: 西本 遥人，大阪大学 産業科学研究所 駒谷研究室，〒 567-0047 大阪府茨木市美穂ヶ丘 8-1，TEL:06-6879-8416，nishimoto@ei.sanken.osaka-u.ac.jp

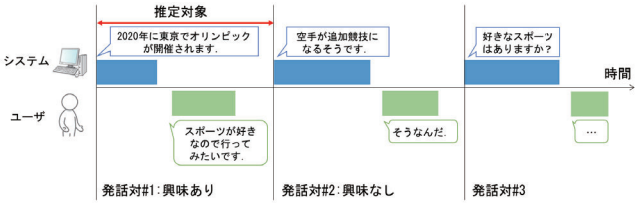


図 1: 興味の推定の対象区間

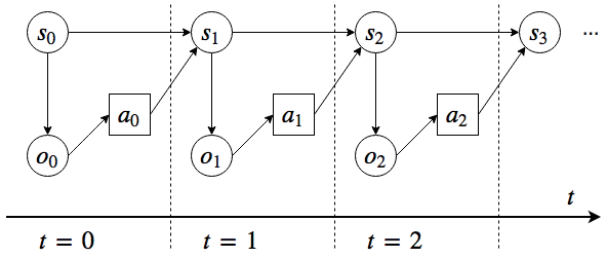


図 2: POMDP の変数の依存関係

これらの課題を解決するために，本稿では Partially Observable Markov Decision Process（部分観測マルコフ決定過程，POMDP）を用いたシステム構築について述べる．その効果を検証し，興味度の推定の重要性，興味度とシステムが次発話でとるべき行動の関連性を述べる．

2. 対話履歴を用いたユーザの興味度推定

2.1 対話履歴を用いる枠組み

POMDP とは，変化するユーザの状態が直接観測できない環境において，その状態を確率的に表し，連続的に表現する手法である [3]．図 2 に POMDP で用いる変数の依存関係を示す．ここで，状態 s_i，o_i，および a_i は，時刻 i における状態，観測データとシステム行動を示す．

本研究では内部状態 s_i として，ユーザのシステムへの興味度を設定し，POMDP を適用する．s_i はユーザの興味，o_i は

システムが観測したマルチモーダル特徴量，a_i はシステム発話とみなす．この設計で，対話という時系列データにおける興味度の変化を追跡する．まず，システムがユーザの発話から特徴量抽出を行い，それをもとにシステムはユーザの興味度を更新する．そして，その興味度をもとに，システムの行動を決定しユーザ状態を更新する．

以下の式 (1) に基づいて信念 $b(s)$ を更新する．式中の「′」は次時刻を表し，「′」なしは現時刻を表す．

$$\begin{aligned} b'(s') &= p(s'|o', a, b) \\ &= \frac{1}{z} p(o'|s', a) \sum_{s \in S} p(s'|a, s) b(s) \end{aligned} \tag{1}$$

$b(s)$ は信念であり，ある時刻において状態が s である確率を表す．$p(o'|s', a)$ は s の尤度関数，$p(s'|a, s)$ は s の状態遷移確率，$\frac{1}{z}$ は正規化係数である．POMDP では，現在の信念 $b(s)$ に基づき，システム行動が式 (2) 政策関数 π によって決定される．

$$a' = \pi(b'(s')) \tag{2}$$

2.2 信念の更新式の設計

状態変数 s は興味あり/なしを意味する s_1/s_2 の 2 値を取るとする．まず，式 (1) の尤度関数 $p(o'|s', a)$ を，簡単のために a を除いた $p(o'|s')$ として考える．$p(o'|s')$ を定式化するのは困難なため，ベイズの定理により以下のように変形する．

$$p(o'|s') = \frac{p(s'|o')}{p(s')} \cdot \sum_{s'} p(o'|s') p(s') \tag{3}$$

ここで $\sum_{s'} p(o'|s')p(s')$ は s' にかかわらず一定となるので，

$$p(o'|s', a) \propto \frac{p(s'|o')}{p(s')} \tag{4}$$

という関係が得られる．$p(s'|o')$ は特徴量が得られたときの s' の事後確率である．これを現時刻の発話対から推定した興味度で近似する．$p(s')$ は s' の事前確率である．次に，状態遷移確率 $p(s'|a, s)$ を考える．簡単のために，$p(s'|a, s)$ を a を除いた $p(s'|s)$ とみなす．具体的な設定は 4.1 節で述べる．

最後に，対話中にシステムが話題を変更する場合，式 (1) 中の事前確率 $\sum_{s \in S} p(s'|a, s)b(s)$ について，状態 s を等確率とする規則を追加する．これは，話題の変更時には前発話の興味度は必ずしも現発話に反映されず，ユーザの興味度は中間的な値に戻ると考えられるからである．

3. 対象データと単一発話の興味度推定

3.1 対象データ

対象としたデータは，人工知能学会 言語・音声理解と対話処理研究会に設置されたワーキンググループにより収集された，マルチモーダル対話コーパス [4],[5] である．コーパスの対話内容はシステムが決められた話題に関する質問や情報提供を行うものである．コーパスは合計 39 名の対話が収録されており，合計 3209 のシステム-ユーザ発話対が含まれる．その各発話対ごとに，3.2.1 節で述べる興味ラベルが付与されている．また今回新たに，このうちの 29 名分 [5] に対して，3.2.2 節に示す話題継続ラベルを付与した．

表 1: L_{int} の内訳

L_{int} の値	< 0.5	$= 0.5$	> 0.5	all
数	1326	292	1591	3209

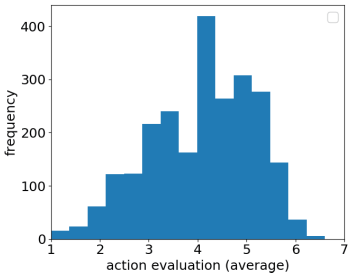


図 3: 5 名のアノテータの話題継続ラベルの平均

3.2 アノテーションと正解

3.2.1 興味ラベル

使用する対話コーパス [4],[5] から，興味ラベルの作成を行う．複数のアノテータによって発話対ごとに「興味あり」「不明」「興味なし」とラベル付けされている．それらのラベルの票数をもとに，対象発話対での興味度を式 (5) を用いて 0 から 1 の連続値 L_{int} に変換し，これを人手で付与された正解とする．複数人のアノテータの意見の一致割合をその発話対でのユーザの興味度とする．ユーザの興味度を 0 から 1 の連続値で表現することで，L_{int} を興味ありである確率とみなすことができる．

$$L_{int} = \frac{1}{2}(1 + \frac{N_o - N_x}{N_a}) \tag{5}$$

N_a はアノテータの人数 ($N_a = 3, 6$)，N_o, N_x はそれぞれ興味あり，興味なしラベルを付けた人数である．表 1 に L_{int} の内訳を示す．

3.2.2 話題継続ラベル

対象データのうち 29 名分 2402 発話対 [5] に対して，話題継続ラベルを新たに付与した．5 名のアノテータに対して，発話対ごとに，「自分がシステム役（聞き役）だったとした場合，ユーザ発話の後に話題を変えようと思うかどうか」を 7 段階で付与してもらった．この 7 段階は，1 が「話題を変える」，4 が「どちらとも言えない」，7 が「この話題を深く尋ねる」である．このラベルを用いて，システムが次に取るべき行動の正解データを作成した．具体的には，5 名が付与したラベルの値を発話対ごとに平均し L_{act} とした．L_{act} のヒストグラムを図 3 に示す．

3.3 単一発話の興味度推定

興味度の推定には対話情報，韻律情報，発話内容，顔画像情報の 4 つのモーダル情報と L_{int} を用いる．マルチモーダル情報を用いる理由は，推定を行う際にユーザの複数の情報を用いると，統合推定結果を修正でき，より信頼できる結果を得られるからである [6]．以下に 4 つのモーダルから抽出した特徴量を示す．

【対話情報】対話行為，発話語数，発話語数の差，応答時間

【韻律情報】openSMILE[7] で抽出される音響特徴から，特徴選択により特徴量を 10 に削減 *1

*1 scikit-learn の selelectKbest を利用．

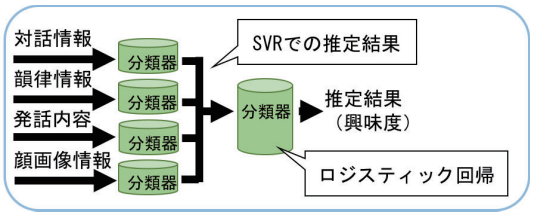


図 4: 興味推定の概要

表 2: ユニモーダルとマルチモーダルでの推定性能の比較

対話	韻律	発話	顔画像	RMSE
○	-	-	-	0.419
-	○	-	-	0.387
-	-	○	-	0.407
-	-	-	○	0.402
○	○	○	○	0.319

【発話内容】語彙（ユーザ発話中の単語から bag-of-words で作成，主成分分析で次元数を 5 に圧縮.），品詞の数，無言の有無

【顔画像情報】人の顔の動きを記述する Action Unit (AU)[8] を使用．17 の AU の出現の有無，18 の AU の強度に関する値

[2] をもとに，単一発話のみを用いた興味度の推定を図 4 のように行った．まず，4 つのマルチモーダル情報をそれぞれ単独で用いて SVR を行った．次に，それらの回帰の結果得られる値を説明変数として，ロジスティック回帰を行い，マルチモーダル情報を統合した．学習のコスト関数にはクロスエントロピーを用いた．推定性能の評価は，未知のデータに対する性能を測るために，ユーザごとに分割した交差検証により行った．

ユニモーダルでの推定結果，統合推定結果を表 2 に示す．評価指標には L_{int} と推定結果の平均平方二乗誤差 (RMSE) を用いた．各モーダル情報を統合した結果，RMSE が他に比べ低い値となっており，推定性能の上昇が確認できる．

4. 評価実験

4.1 対話履歴を考慮したときの性能評価実験

POMDP の設計に関して，使用コーパス [4],[5] の全発話対をもとに，式 (1) 中の尤度関数，状態遷移確率を定めた．尤度関数 $p(o'|s')$ は式 (4) より，$p(s'|o')$ と $p(s')$ に分けられる．前者を単一発話の興味度の推定結果 $r = p(s'|o')$ とした．後者をコーパス中の興味あり/なしの発話対の割合とし，$(p(s_1), p(s_2)) = (0.413, 0.587)$ とした．状態遷移確率 $p(s'|s)$ はコーパスから状態遷移した発話対の組の割合を求め，以下と設計した．$p_{s_i \to s_j}$ は，s_i から s_j へ状態遷移する確率である．

$$p(s'|s) = \begin{bmatrix} p_{s_1 \to s_1} & p_{s_2 \to s_1} \\ p_{s_1 \to s_2} & p_{s_2 \to s_2} \end{bmatrix} = \begin{bmatrix} 0.699 & 0.301 \\ 0.213 & 0.787 \end{bmatrix} \quad (6)$$

一方，対話履歴を用いた手法（以降，POMDP）と比較して，対話履歴を用いない手法をベースラインとする（以降，BASE）．BASE は式 (7) で定義した．

$$\begin{bmatrix} b'(s_1) \\ b'(s_2) \end{bmatrix} = \begin{bmatrix} r \\ 1-r \end{bmatrix} \quad (7)$$

表 3: 推定した興味度と L_{int} の誤差 (RMSE)

POMDP	0.316
BASE	0.319

表 4: L_{act} と興味度の相関係数

POMDP	0.566
BASE	0.545
L_{int}	0.788

POMDP によって更新された興味度の推定値と，BASE での興味度の推定値を比較する．図 5 はあるユーザの対話（全 82 発話対）における POMDP と BASE での $b(s_1)$ の遷移である．横軸が発話対番号，縦軸がそのときの $b(s_1)$ の値である．点線はそれぞれ POMDP，BASE の結果である．黒の点線は，興味あり/なしの中間の状態 ($b(s_1) = 0.5$) を示したものである．赤の点線は，対話において話題が変化した箇所を示している．点線と実線が似た形の折れ線であるほど，モデルによる推定値が人手で付与された値に対応しており，性能がよいと言える．2 つの点線の相違箇所が，POMDP と BASE の差分である．

POMDP を用いることで，推定を誤った場合の改善がなされていた箇所があることが確認できた．図 5 の範囲 A の L_{int} は時間の経過とともに，ユーザの興味度合いが上昇している．範囲 A の途中で，BASE では興味度が低いと誤推定され，折れ線が下降しているのがわかる．一方，POMDP は興味度がこれまでより低いという観測情報 o を得たが，BASE よりも下降の割合は小さく L_{int} により近い値であることが確認できる．これは，前発話での情報（図 5 中の A の場合，「興味あり」）を保持して推定がなされたからだと考えられる．BASE と POMDP それぞれと，L_{int} との RMSE を表 3 に示す．ここからも，僅かではあるが POMDP の方がより L_{int} に近い結果が得られたことがわかる．

一方，範囲 B では POMDP での結果が L_{int} から大きく外れていることがわかる．$b(s_1)$ の増加傾向に関しては，ある程度 L_{int} と同じ傾向がある．誤差の大きい推定をした後の数発話対は，正しく推定することが困難であると考えられる．これに対しては，推定性能の向上や，信頼できない推定結果は反映しないといった規則の作成もしくはシステム行動 a を用いた $p(s'|a, s)$ の設計などで改善が見込まれ，今後の課題である．

4.2 興味度と適切なシステム行動の関係

ここでは，推定した興味度，人手で付与された興味度それぞれと，新たに作成した L_{act} の相関を調べた．その結果を表 4 と図 6 に示す．表 4 から，L_{int} と L_{act} は強い相関があることがわかる．言い換えると，人は対話相手が興味を持っていそうな場合，話題を深く掘り下げるべきだと感じる傾向があることが確認された．表 4 の BASE よりも，POMDP が L_{int} の相関係数に近い値となっている．これは，POMDP が人手で付与した正解により近い興味度推定結果を出力することを示している．図 6 は，L_{int} と L_{act} の相関を示したものである．図 6 中の直線は回帰直線であり，正の相関があることが確認できる．以上より，システムの行動をユーザの興味をもとに決定する枠組みは適切であると言える．

図 7 は POMDP での興味度の推定結果と L_{act} の相関を示したものである．表 4 から分かる通り，POMDP は L_{int} ほどではないが，L_{act} との相関がある．図 6 と図 7 を比較する

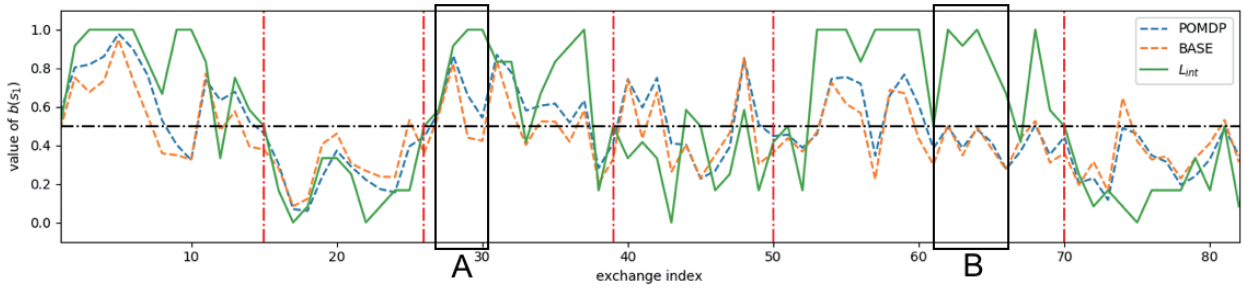


図 5: 推定した興味度の比較の例 (userID：F2006)

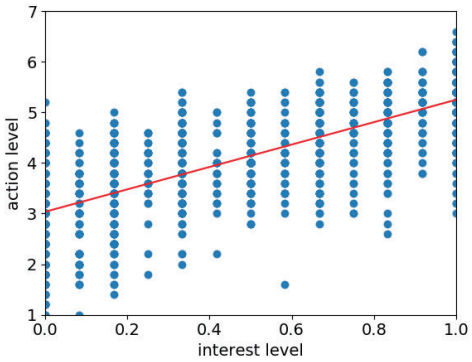


図 6: L_{int} と L_{act} の相関

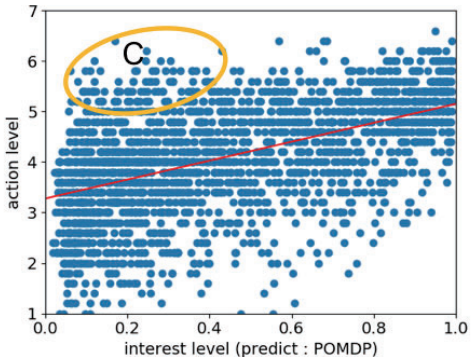


図 7: 対話履歴を用いた興味度の推定結果と L_{act} の相関

と，POMDP では話題を深く尋ねるべきシステム発話に対して，ユーザは興味がないという推定をした結果が多く見られる（図中 C の範囲）．これは信念 $b(s)$ の更新式中の状態遷移確率を，より興味なしに遷移するように設計（式 (6) を参照）したためと考えられる．本稿では状態遷移確率をコーパスの遷移の割合をもとに設計したが，今後の改善の余地があると考えられる．

5. 今後の課題

本稿では近似を用いて POMDP を実装したため，課題は多く存在している．2.2 節で述べた手法で，信念 $b(s)$ の更新式に含まれる状態遷移確率と尤度がシステムの行動 a とは独立であると仮定している．a を反映させた更新式の設計が今後の課題である．

今後は，$b(s)$ をもとにシステム行動を決める政策関数 π の学習に取り組むことも課題の一つである．政策関数 π を最適化するための手法として，(s_t, s_{t+1}, a_t) のペアから報酬を設計し，それを用いた強化学習が考えられる．報酬の設計には，目標の対話システムの評価尺度の定義を行う必要がある．評価尺度は「ユーザに興味度が高いまま持続させる」「対話を長く続ける」「話題変更をしたときに興味度がどう変化したか」などが考えられる．それを定義した上で，報酬の設計に取り組む必要がある．

参考文献

[1] M. L. Knapp, J. A. Hall, and T. G. Horgan. Nonverbal communication in human interaction. *Cengage Learning*, 2013.

[2] 西本遥人, 武田龍, 駒谷和範. マルチモーダル対話における興味の有無の推定と追加コーパスを用いた性能評価. *SIG-SLUD*, Vol. 5, No. 02, pp. 72–73, 2018.

[3] J. D. Williams and S. Young. Partially observable Markov decision processes for spoken dialog systems. *Computer Speech & Language*, Vol. 21, No. 2, pp. 393–422, 2007.

[4] 荒木雅弘, 冨増紗也華, 中野幹生, 駒谷和範, 岡田将吾, 藤江真也, 杉山弘晃. マルチモーダル対話データの収集と興味判定アノテーションの分析. *SIG-SLUD*, Vol. B5, No. 02, pp. 20–25, 2017.

[5] 駒谷和範, 岡田将吾, 西本遥人, 荒木雅弘, 中野幹生. 配布可能なマルチモーダル対話データの収集とアノテーション不一致傾向の分析. *SIG-SLUD*, Vol. 5, No. 02, pp. 45–50, 2018.

[6] C. Busso, Z. Deng, S. Yildirim, M. Bulut, C. M. Lee, et al. Analysis of emotion recognition using facial expressions, speech and multimodal information. In *Proc. ICMI*, pp. 205–211, 2004.

[7] B. Schuller, S. Steidl, and A. Batliner. The interspeech 2009 emotion challenge. In *Proc. Interspeech*.

[8] P. Ekman and W. V. Friesen. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press, 1978.

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

[4F3-OS-11b] 社会的信号処理と AI(2)

岡田 将吾（北陸先端科学技術大学院大学）、石井 亮（NTT）

Fri. Jun 7, 2019 2:00 PM - 3:00 PM Room F (302B Medium meeting room)

[4F3-OS-11b-01] Considering Idling Motion Generation using RNNs for Virtual Conversational Agents

○Hung-Hsuan Huang[1,2], Masato Fukuda[1], Toyoaki Nishida[1,2] （1. RIKEN Center for Advanced Intelligence, 2. Graduate School of Informatics, Kyoto University）

2:00 PM - 2:20 PM

[4F3-OS-11b-02] Identifying Discourse Boundaries in Group Discussions using Multimodal Features

Ken Tomiyama[1], ○Fumio Nihei[1], Yutaka Takase[2], Yukiko Nakano[2] （1. Graduate School of Science and Technology, Seikei University, 2. Faculty of Science and Technology, Seikei University）

2:20 PM - 2:40 PM

[4F3-OS-11b-03] A corpus-based typology of reactions to apologies in Japanese}

○Masashi Inoue[1], Kaori Suzuki[1] （1. Tohoku Institute of Technology）

2:40 PM - 3:00 PM

©The Japanese Society for Artificial Intelligence

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

回帰型ニューラルネットを用いた会話エージェントのアイドリング動作生成の検討

Considering Idling Motion Generation using RNNs for Virtual Conversational Agents

黄 宏軒 *1*2 Hung-Hsuan Huang　福田匡人 *1 Masato Fukuda　西田豊明 *1*2 Toyoaki Nishida

*1国立研究開発法人理化学研究所革新知能統合研究センター
RIKEN Center for Advanced Intelligence

*2京都大学大学院情報学研究科
Graduate School of Informatics, Kyoto University

This work aims to develop a model to generate fine grained and reactive non-verbal behaviors of the virtual character when the human user is talking to it. The target *micro* non-verbal idling behaviors are micro facial expression, head movements, and postures. We explored the use of recurrent neural network (RNN) to learn these behaviors in reacting to the human communication interlocutor's corresponding micro non-verbal behaviors. The models are trained on an active listening data corpus which features elderly speakers talking with young active listeners and was collected by ourselves.

1. はじめに

擬人化会話エージェントはコミュニケーションロボットのような物理的制約がなく，高い自由度で精緻でリアルな顔の表情や身体の仕草を表出でき，ユーザとコミュニケーションが取れるインターフェイスとして期待されている．しかし，これまでの会話エージェントの多くは事前に定義された顔表情や体の動きのアニメーションシーケンスをインタラクション中にルールベースで呼び出して再生する仕組みになっていた [Huang 11]．そのため，エージェントが同じことをする際に，例えば，笑うと必ず同じ笑いをしたり，視線をそらすと，同じ視線の動きをしたりすることになる．また，エージェントが何らかの意図を表出しようとしない際には，静止する，または，固定のアイドリングモーションを繰り返し再生するようになっている．こうしたエージェントとしばらくインタラクションをしていると，その固定パターンが見えて自然さを損なってしまう．

しかし，生身の人間は固定パターンの動作をしない。笑うたびに毎回少し違う笑いをするはずであり，微細なブレが含まれている．そのブレは物理的な要因以外に会話相手，文脈，本人の内的状態などの影響を受けると思われる．本研究はそのブレを捉えようとするものである．エージェントが何らかの意図を表出しようとしない間，いわゆるアイドリングの状態でも，リアルタイムで会話相手の振る舞いから微細なアイドリングモーションの生成を目指している．

近年，深層学習が様々な分野で著しい成果を挙げており，普及が急速に進んでいる．そのなかでは，データの経時的変化を捉えるように再帰型ニューラル (Recurrent Neural Networks, RNN) が考案されている．本研究では，エージェントの振る舞い生成モデルの検討として，エージェントの会話相手の振る舞い (頭部の回転，視線，表情，声の韻律情報) からエージェントの振る舞い (頭部の動き，表情，上半身の動き) を予測することを課題とし，我々が収集した傾聴会話データコーパスを基に，異なるネットワーク構造の予測性能について考察する．RNN の代表的な構造である Long Short-term Memory (LSTM) [Hochreiter 97] と忘却ゲートや出力ゲートを省きより簡素な構造になっている Gated Recurrent Unit (GRU)[Cho 14] と，ベースラインのマルチレイヤパーセプトロン (MLP) を比較対象とする．

連絡先: 黄 宏軒，hhhuang@acm.org

2. データコーパス

本研究の目的は利用者の年齢層を限定しないが，実装する会話エージェントが高齢者のメンタル支援での活用が視野に入るため，会話データコーパス収集実験は高齢者を対象とした傾聴会話課題で行った．実験参加者は高齢者 4 名 (男性 2 名，女性 2 名．年齢 69〜73 歳，平均 71 歳) と大学生 4 名 (男性 2 名，女性 2 名，平均年齢 21 歳) であり，参加者毎の組み合わせで 15〜30 分間の 4x4 計 16 会話を収録した．エージェントとの会話は画面越しになるのと高齢の参加者は遠隔地からだったため，実験は Skype でのテレビ会話であった．Skype 通信用の Web カメラの他，フル HD のビデオカメラで参加者の上半身の映像を正面から記録した (図 1)．高齢者が話し手役で若者が聴き手役の会話であったが，話題は限定せず初対面の設定で自由に会話してもらった．結果として若者が質問をして高齢者がそれに答える場面が多く見られ，話題は双方の自己紹介のほか，高齢者の過去の仕事，趣味，最近はまったことなどであった．

わずかであったがネットワーク通信の遅延があったため，分析データは実時間のデータではなく，高齢者側で収録した若者の音声を基準に若者側の映像を同期させた．その上で高齢者と若者それぞれ直接撮影したビデオカメラの映像と音声を抽出した．映像に対して，OpenFace*1 を用いて Ekman が提唱しており，顔表情を記述する Facial Action Coding System (FACS)[Ekman 02] の Action Unit (AU)44 個のうち 17 個，そして，頭部の回転 (3 次元)，視線方向 (8 次元) の 30 fps データを抽出した．さらに，OpenPose *2 を用いて肩関節の 2 次元座標からカメラまでの距離と肩の傾きを抽出した (30 fps)．音声については OpenSmile *3 を用いて Interspeech 2009 の

∗1 https://github.com/TadasBaltrusaitis/OpenFace
∗2 https://github.com/CMU-Perceptual-Computing-Lab/openpose
∗3 https://www.audeering.com/what-we-do/opensmile/



図 1: 傾聴会話コーパスの収録映像．自然な会話に近づけるように大画面のテレビとスクリーンに映像を映し，視線が合うように位置を調整した

Emotion Challenge[Schuller 09] で定められたローレベル特徴量，energy (1 次元), mfcc (12 次元), zero-crossing rate of time signal (1 次元)，voicing probability (1 次元)，F0(1 次元) の計 16 次元 100 fps のデータを抽出した．聴き手の振る舞いは発話状態によって変わると思われるため，Elan[?] のツールを用いて聴き手の発話区間を割り出した．

話し手の振る舞い (OpenFace で抽出した映像情報 28 次元と OpenSmile で抽出した韻律情報 16 次元) を説明変数とし，聴き手の振る舞い (OpenFace の 28 次元と OpenSmile の 16 次元と OpenPose の 2 次元) を目的変数と設定した．計算が重くてリアルタイムでの入力が困難であったため，OpenPose の情報を説明変数にしていなかった．会話セッション長にばらつきがあったため，セッションの最初から最長 20 分間の部分を切り出した．入力データに 1 秒間と 2 秒間のウィンドウ幅を持たせ，出力は 1 フレーム分になるようにデータを整形した．整形した約 4.5 時間分のデータセットの概要を表 1 に示す．このデータから，男性聴き手の発話時間が短く，話し手の話す時間が長くなっており，女性聴き手自身の発話が長く，話し手の話す時間が相対的に短ったことが分かる．実験時に，男性聴き手は一方的に質問を聞いており相手の回答から話題を掘り下げない傾向に対して，女性聴き手の方は聞き上手で全体的に会話が弾んでいる観察と一致している．

表 1: データセットの概要．データ量は 2 秒時間幅のもの

	発話状態	フレーム	時間 (s)	データ量
男性	On	69,818	2,334	1.4GB
	Off	169,387	5,663	13.6GB
女性	On	82,261	2,750	1.7GB
	Off	147,165	4,920	11.8GB

3. 学習モデルの比較

3.1 実験手順

前節のデータセットの全体，男性聴き手，女性聴き手，そして聴き手個人のサブデータセットに対して，MLP, GRU, LSM の 3 つのモデルと時間幅 1 秒，2 秒の条件で機械学習実験を行った．ネットワーク構造については入力モダリティの fps 数が違うため，各モダリティの時系列をそれぞれに対応する LSTM/GRU 層 (入力の 8 倍、それぞれ 28 x 8, 16 x 8 個のノード) に入力して、そのあと全モダリティを結合する全結合層 (2 x 128 個のノード)，そして，さらに 1 層の全結合層 (512 個のノード) の出力層のシンプルな 4 層構成で統一している (図 2)．入出力データはすべて 0 から 1 までの区間に正規化した．なお，MLP については，時系列表現ができないため，同等のデータ入力を横方向 1 次元に展開して扱う．

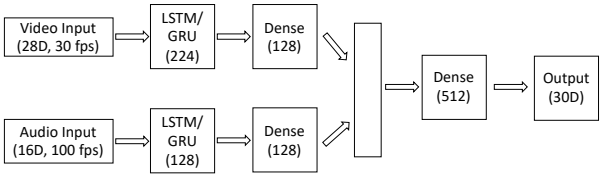


図 2: 評価実験に用いたネットワークの構造

学習モデルの評価には平均二乗誤差の平方根 (Root Mean Squared Error, RMSE) を用いる．各条件で一人の話し手をテストデータとして抽出して残り 3 人のデータを用いて学習し，これを 4 回繰り返して平均値を求める leave-one-subject out 法で交差検証を行った．

3.2 結果の考察

表 2 から，どのデータセットを用いても GRU が最も高い精度 (回帰誤差を示す RMSE 値が低い) を得ており，LSTM がそれに次いで，MLP の精度が最も低い結果を示す．このことから，同等の入力データに対してやはり時系列に特化したモデルの方は話し手と聴き手のインタラクションの特徴を捉えられている．そのなかでも今回のデータセットに対してよりシンプルな構造の GRU の方は性能が高い点，2 秒間の入力時間幅よりも 1 秒間幅の方は性能が高いことから一瞬一瞬の顔表情は反射的な要因が強いことを示唆される．また，LSTM モデルにおいても GRU モデルにおいても，男性聴き手の両名とも長い時間幅の方はより精度が高く，女性聴き手の両名とも短い時間幅の方は精度が高いことも興味深い．

4. 結論

本研究は仮想会話エージェントがアイドリング状態における意図の含まないモーションを生成するために，傾聴会話コーパスを基に話し手の振る舞いの映像と音声情報から聴き手の振る舞いを回帰的ニューラルネットの LSTM，GRU とベースラインの MLP で予測する試みをした．0〜1 の値域のデータに対して，どのモデルも 0.15 程度以下の精度を出力しており，会話パートナー同士の心理状態や話の内容を問わず，単純に相手の顔・声から顔表情と上半身のポスチャーをある程度予測できている結果を示した．しかし，この結果はあくまで数値データの観点での精度であり，生成されたエージェントの動きに対して，人間らしさを感じるかどうかの主観評価も今後実施する必要がある．

参考文献

[Cho 14] Cho, K., Merrienboer, van B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y.: Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation, *CoRR*, Vol. abs/1406.1078, (2014)

[Ekman 02] Ekman, P., Friesen, W. V., and Hager, J. C.: Facial Action Coding System (FACS), Website (2002)

表 2: 機械学習の結果

	時間幅	全員	男性	女性	全員平均	男性平均	女性平均
LSTM	1 sec	0.1335	0.1439	0.1219	0.1219	0.1356	0.1081
LSTM	2 sec	0.1416	0.1415	0.1343	0.1271	0.1357	0.1185
GRU	1 sec	0.1299	0.1369	0.1187	0.1174	0.1282	0.1065
GRU	2 sec	0.1323	0.1343	0.1278	0.1207	0.1264	0.1150
MLP	1 sec	0.1377	0.1513	0.1263	0.1278	0.1393	0.1162
MLP	2 sec	0.1514	0.1525	0.1442	0.1354	0.1406	0.1303

[Hochreiter 97] Hochreiter, S. and Schmidhuber, J.: Long Short-term Memory, *Neural Computation*, Vol. 9, No. 8, pp. 1735–1780 (1997)

[Huang 11] Huang, L., Morency, L.-P., and Gratch, J.: Virtual Rapport 2.0, in *11th International Conference on Intelligent Virtual Agents (IVA 2011)*, pp. 68–79 (2011)

[Schuller 09] Schuller, B., Steidl, S., and Batliner, A.: The INTERSPEECH 2009 Emotion Challenge, in *10th Annual Conference of the International Speech Communication Association (Interspeech 2009)*, Brighton, United Kingdom (2009)

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

マルチモーダル特徴量を用いた談話セグメントの検出

Identifying Discourse Boundaries in Group Discussions using Multimodal Features

冨山 健[*1] Ken Tomiyama　二瓶 芙巳雄[*1] Fumio Nihei　高瀬 裕[*2] Yutaka Takase　中野 有紀子[*2] Yukiko Nakano

[*1] 成蹊大学大学院理工学研究科 Graduate School of Science and Technology, Seikei University

[*2] 成蹊大学理工学部 Faculty of Science and Technology, Seikei University

This study proposes models for detecting conversation boundaries in group discussions. First, we created a multimodal embedding space using an autoencoder, and applied a similarity-based approach to detect the discussion boundary. As the second method, we annotated conversation boundaries and created unimodal CNN models for language, audio, and head motion information. Then, created multimodal models by concatenating the output of unimodal models. In the evaluation experiment, we found that language information was the most useful modality, but by combining with audio and head motion modalities, the CNN-based models more accurately predict the conversation boundaries.

1. はじめに

自然言語処理の分野では，テキストを複数の意味的まとまりに分割することをテキストセグメンテーションという．テキストを分割することは文章の要約や検索などに有益であると考えられる．議論においても，セグメントに分割することができれば，話題構造の認識やファシリテーション支援にも応用できる．例えば，セグメントが短ければ，そのトピックについて議論が十分でない可能性があり，ファシリテータの介入決定の情報となりうる．

文章のセグメンテーション手法は，これまでに多くの研究が行われてきたが，高い精度をもつ手法は未だ確立されておらず，また，対話文に適用した研究は少ない．多くの研究では語彙結束性をセグメント境界の識別モデルの作成などに使用しているが，対話においてはそのスパース性が問題になってくるなど，難しい課題である．

そこで本研究では，議論参加者の言語・非言語情報に基づいた，会話のセグメント境界を検出するモデルを提案する．

2. 関連研究

[1]では語彙結束性をとらえ，文章を分割する手法 TextTiling を提案している．TextTiling は，文章に 2 つの連続する窓をスライドさせ，類似度を算出することで境界を検出する手法である．これを基にした手法も多く提案されていて，広く知られた手法である．また，[2]は新聞や TV ニュースの書き起こしを用いて，手掛かり語彙や前後に出現する語彙情報などから識別関数を作成している．[3]では複数人対話コーパスのテキスト，音声情報を用い，手掛かり語彙，沈黙時間，オーバーラップ時間などを基にしたトピックシフト識別モデルを作成している．

3. 議論コーパス

3.1 議論収集実験

本研究で使用する議論コーパスについて述べる．

1 グループ 3 人 9 グループ，男女 27 人の実験参加者が議論を行った．課題は外国人へ向けた一日観光コースをテーマに 30 分間話し合い，その後 5 分間でアピール点を決定してもらうというものである．最終的に「観光客の国籍と観光コースのキャッチフレーズ」「観光コースの詳細」「アピール点」を記入した用紙を提出させた．議論は「対象の外国人の国籍」「コースのキャッチフレーズ」「観光コースの作成」の順に行うよう指示した．議論前に参加者には議論終了後に作成した観光コースについて「観光スポットの数が多い」などの評価点によって評価されることが伝えられた．



図 1：実験環境

図 1 に実験環境を示す．Kinect による各参加者の追跡データ，頭部に装着したモーションセンサによる 20 ms 毎の加速度・角速度・地磁気のデータ，ヘッドセットマイクによる各参加者の音声データ，ビデオカメラによる議論全体の動画が収集された．また録画した動画から議論の書き起こしを作成した．

3.2 セグメント境界のアノテーション

収集した議論に対して，会話のセグメント境界のアノテーションを行った．初めに，2 名のアノテータが試行と議論を繰り返し，コーディングスキームを作成した．セグメント境界になる発話のルールとして次の 3 つのルールを定めた，1.「話題が移った発話」，2.「会話のレベルが下がる(上がる)発話」，3.「書き込み作業について言及した発話」．1 番目のルールはたとえば，「旅行時期」から「旅行場所」へ話題が移った場合である．2 番目のルールはたとえば「旅行場所」からより詳細な話題「京都」へ移った場合である，3 番目のルールは議論内容に関係するトピックではないが，本研究で使用したコーパスでは多々見られたため，追加した．

連絡先：中野有紀子，成蹊大学理工学部，y.nakano@st.seikei.ac.jp

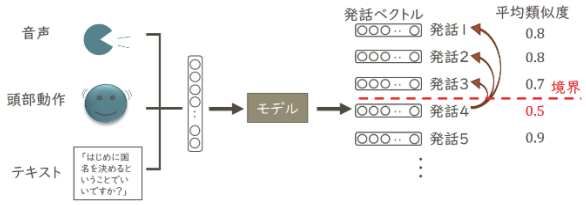


図 2：類似度に基づくセグメンテーション

この作成したコーディングスキームをもとに，全 9 グループについて自動で検出された音声区間に対して，1 名のアノテータがセグメント境界を付与した．ただし，前章で述べた通り，議論 35 分間の内，最後の 5 分間は基本的には用紙への書き込み作業になるため，それを除いた 30 分をアノテーションの範囲とする．その結果，各グループにつき平均約 35 個のセグメント境界が付与された．

4. 発話ベクトルを用いた類似度算出モデル

セグメント内の発話は結束性のある内容を持っていると考えられる．そこで，発話をベクトルで表現できれば，発話間の類似度を算出し，類似度に基づき境界を決定する手法を利用できると考えた．また，議論上で人間は言語だけでなく非言語情報も用いているので，マルチモーダルな情報を用いた議論のセグメント境界の識別を目的とする．そこで，マルチモーダルな情報を基に AutoEncoder を用いて発話を表現するベクトルを作成し，発話ベクトルの類似度を算出し，境界の識別を試みる．提案手法の概略図を図 2 に示す．

4.1 特徴量

セグメント境界の識別に使用する特徴量を下に記す．これらは発話単位で算出される．

(1) 新規/既出名詞の数

発話の書き起こしを形態素解析器 Mecab にかけ，名詞を抽出する．抽出した名詞がそれまでの議論で出現した名詞と一致すれば既出名詞，一致しなければ新規名詞とし，それぞれの個数を数え上げる．一度出現した名詞は既出名詞一覧に登録され，以後の新規/既出名詞の判断に用いられる．

(2) 現発話と直前の発話間の一致/不一致名詞数

直前の発話と現発話に共通する名詞の数を共通名詞数とする．また，現発話に出現し，直前の発話には出現しない名詞の個数を不一致名詞数とする．

(3) 現発話と直前の発話間の一致/不一致動詞数

直前の発話と現発話に共通する動詞の数を共通動詞数とする．また，現発話に出現し，直前の発話には出現しない動詞の個数を不一致動詞数とする．

(4) 発話長（時間長と形態素数）

発話長を示す特徴量として，発話区間の時間長 ms と発話を構成する形態素数の 2 種類を使用した．発話者の話速により，同じ時間長でも形態素数が異なるため，発話の形態素数も発話長として特徴量とした．

(5) オーバーラップ時間

2 つの連続する発話のオーバーラップ時間 ms を特徴量とする．1 人の発話に他の 2 人の発話がオーバーラップしている場合は，それらの合計を算出する．

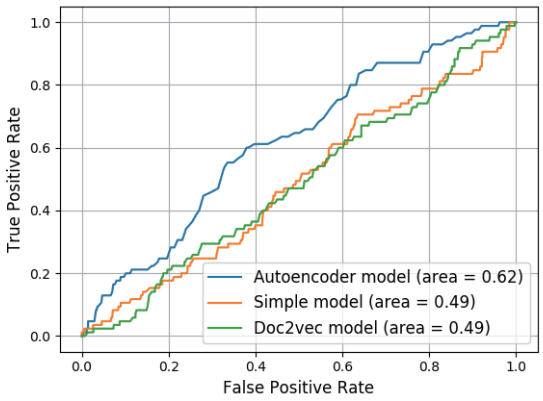


図 3：3 つのモデルの ROC 曲線と AUC による比較

(6) 頭部回転量

発話が行われている時間を対象に，20 ms の時間窓を設定し，各区間における頭部のヨー角回転量 degree をデータとする．各区間における最大値・最小値・平均値・分散値をそれぞれ特徴量とする．

(7) 音声インテンシティ

音声解析ソフト Praat を用いて，収録した発話音声のインテンシティ dB を計測する．各発話区間において 10 ms の時間窓を設定し，その区間におけるインテンシティの最大値・最小値・平均値・分散値をそれぞれ特徴量とする．

(8) 合成加速度

議論参加者の後頭部につけられたモーションセンサの出力から下式を算出することで，合成加速度 HA を得る．ここで，x, y, z はそれぞれ x 軸，y 軸，z 軸方向の加速度の値である．

$$HA = \sqrt{x^2 + y^2 + z^2}$$

発話毎に，算出された合成加速度から最大値，最小値，平均値，分散を計算する．

(9) 合成加速度の Wavelet 変換特徴量

(8)で得られた合成加速度に Daubechies ウェーブレットを用いた多重解像度解析を行い，最下位レベルの信号の最大値，最小値，平均値，分散値を計算する．

(10) Doc2vec 特徴量

Doc2vec [4] を日本語の Wikipedia 記事を用いて学習し，各発話について 200 次元のベクトルを得る．

4.2 識別モデルの作成

発話ベクトル作成に使用する AutoEncoder は入力層，中間層，出力層の 3 層からなり，中間層と出力層でそれぞれ ReLu と Liner の活性化関数を用い，コスト関数は平均二乗誤差を使用する．入力された 241 次元のベクトルは中間層で 150 次元まで圧縮される．これを発話ベクトルとして類似度の算出に用いる．

AutoEncoder の訓練には 9 グループの内 7 グループ（4044 発話）を使用し，残りの 2 グループ（1124 発話）を検証に使用した．1000 epoch で訓練を行ったが 300 epoch にはモデルは収束したため，300 epoch 時点のモデルを用いて各発話に対して 150 次元の発話ベクトルを作成した．

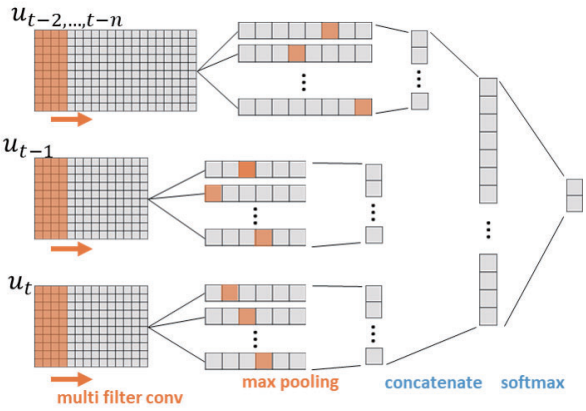


図 4: 言語モダリティモデルのアーキテクチャ

現発話と前 3 発話との発話ベクトルのコサイン類似度をそれぞれ算出し，その平均値を計算する．平均値が任意の閾値を下回れば，その発話を境界と識別する．

4.3 識別モデルの評価と議論

閾値によってセグメント境界かどうかの精度が変化するため，ROC 曲線と AUC で評価する．また，あいづちやフィラーなどの意味のない発話を考慮し，発話長が 0.8 秒以内の発話は除いた．比較のため，提案手法である AutoEncoder で特徴量ベクトルを圧縮したベクトルを用いるモデル（AutoEncoder model）の他に，圧縮せず特徴量を連結させただけのベクトルを用いるモデル（Simple model），Doc2vec 表現だけベクトルを用いるモデル（Doc2vec model）の 3 つを用意し，同様に評価した．その結果，ROC 曲線と AUC は図 3 のようになった．AUC は AutoEncoder model が 0.62, Simple model は 0.49，Doc2vec model は 0.49 となり，AutoEncoder model が一番性能が良い結果となった．

5. 教師あり学習によるセグメント境界識別モデル

各発話について境界であるかどうかを識別する 2 値分類モデルを作成する．教師データとして 3.2 節で作成したアノテーションを用いる．言語情報として発話の書き起こしテキスト，非言語情報として頭部動作，音声を用い，それぞれの 3 つのユニモーダルモデルを作成する． 言語モダリティモデルは Word2vec から得た各発話内の単語ベクトルを入力とし，頭部動作と音声モーダルモデルはそれぞれスペクトログラムを入力とするモデルを作成する．

5.1 入力データの作成

言語モダリティモデルの入力として，Word2vec から得た発話内の単語ベクトルを単語数分連結させた（200 次元×単語数）のサイズの 2 次元配列を発話毎に作成した．

頭部動作スペクトログラムは議論参加者毎に作成する．頭部に装着したモーションセンサから得た角速度から合成角速度を算出し，フーリエ変換することで 25 の周波数解像度を持つスペクトログラムを作成した．音声モーダルモデルの入力となる音声スペクトログラムも同様に議論参加者毎に作成する．ヘッドセットマイクより収録した音声から 32 の周波数分解能を持つスペクトログラムを作成した．また，各データに対して標準化を行った．

さらに，全議論中の最大発話長でゼロパディングを行う．その結果，それぞれの入力サイズは言語モダリティモデルが 217(コ)×200(次元)，頭部動作モデルが 761×25(周波数分解能)，音声モデルは 761×32(周波数分解能)となった．

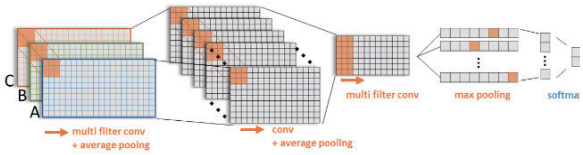


図 5: 頭部動作と音声モダリティのためのアーキテクチャ

5.2 言語モダリティモデルのアーキテクチャ

[5]を参考にアーキテクチャを作成した．アーキテクチャ図を図 4 に示す．現在の発話 u_t(t：発話数)，一つ前の発話 u_{t-1}，過去の発話 $u_{t-2,\ldots,t-n}$(n: 履歴発話数)の 3 つの入力を持つ．各入力されたベクトルは中間層で時間軸方向の 1D-convolution と global max pooling 後に 3 つの出力を連結させて，最終的に softmax 層で境界かどうかの予測が行われる．訓練時のカーネルサイズ，フィルタ数を表 1 に示す．

5.3 頭部動作と音声モーダルモデルのアーキテクチャ

頭部動作と音声モーダルモデルは入力がスペクトログラムであるため，同等のアーキテクチャを用いた．3 人の議論参加者のスペクトログラムを 3 チャンネルとして入力する．参与役割を考慮して，スペクトログラムの順序は優位性の近似値である現発話までの累積発話数[6]で並び替える．入力されたベクトルは中間層で複数の 2D-covolution 層と average pooling の後，時系列方向の 1D-convolution 層と global max pooling を通り，softmax 層で境界かどうかの予測を行う．訓練時には頭部動作モダリティ，音声モダリティともにそれぞれ 2 層の 2D-convolution と average pooling 層を用いた．訓練時のカーネルサイズやフィルタ数を表 1 に示す．

表 1: 各モダリティモデルとカーネルサイズ

	2D-conv	1D-conv	2D-average pooling	2D-conv フィルタ数	1D-conv フィルタ数
テキスト		4			32
頭部動作	(2, 2)	8	(2, 1)	8	32
音声	(2, 2)	8	(2, 1)	8	32

5.4 モデルの訓練と評価

言語，頭部動作，音声のユニモーダルモデルを作成し，3 つの組み合わせも考慮して合計 7 つのモデルを作成する．マルチモーダルモデルは各ユニモーダルモデルの全結合層の入力を連結し softmax 層への入力とすることでモデルの融合を行った．また，モデルの重みの初期値としてユニモーダルで学習した重みを使用する．言語モダリティモデルでは履歴数 10 発話し過去の発話は 8 発話でモデルを作成する．頭部動作モデルは各 convolution 層では10^{-7}の L2 正則化に加え，pooling 層後に 0.25 の dropout を行う．音声モデルも同様に各 convolution 層で10^{-8}の L2 正則化に加え，pooling 層後に 0.25 の dropout を行った．

作成されたモデルは各発話に対してセグメント境界であるかどうかの予測値を出力する．予測値が 0.5 を下回る，つまり境界であると判別された発話が少なかったため，ROC とその面積である AUC でモデル性能の評価を行う．

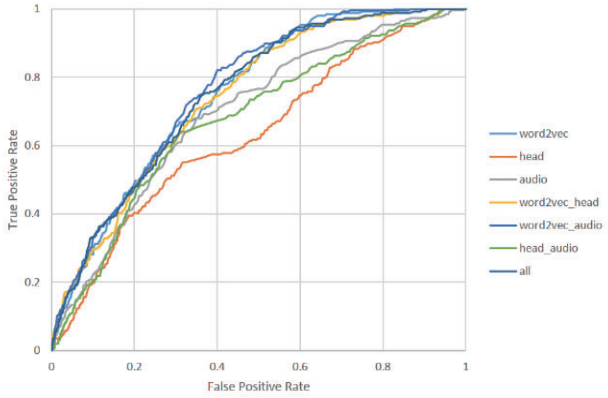


図 6: cross validation での ROC 曲線と AUC

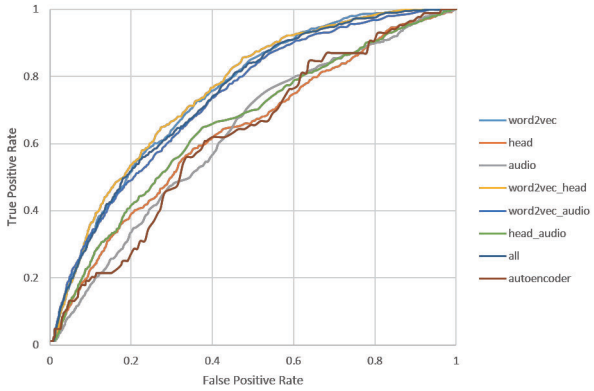


図 7: leave one Group out での ROC 曲線と AUC

全データを 5 分割し，そのうち 1 つを評価データとし，残りのデータで cross validation を行った．5 分割すべてに全グループのデータの発話が含まれるように，グループごとにセグメント境界の発話とそうでない発話に分割し，5 分割したものを統合した．

具体的な訓練・検証・評価データのサイズはそれぞれ約 2500 発話，約 830 発話，約 827 発話である．言語，頭部動作，音声モダリティのモデルをそれぞれ w2v, head, audio とする．これに加え，2 つのモダリティを組み合わせたモデル，言語＋頭部動作(w2v_head)，言語＋音声(w2v_audio)，頭部動作＋音声(head_audio)，全てのモダリティを統合したモデル(all)の 4 つのマルチモーダルモデルを作成した．図 6 にモデル毎の 4 hold の平均 ROC 曲線と平均 AUC を示す．w2v_audio, w2v, all の順に性能が高くなっていることから，言語情報に音声情報を追加すると性能が改善されることがわかる．

しかし，cross validation の評価は訓練・検証・評価で同一グループ内の発話が含まれてしまっている．そこで AutoEncoder モデルとの比較のため，グループ 2 と 7 の 2 グループを評価データとして，残りの 7 グループで leave one Group out を行った．cross validation と同様に各 7 個の検証データごとに作成されたモデルの平均 ROC と AUC を算出した結果を図 7 に示す．w2v_head, w2v, all の順に性能が高く，cross validation での評価とは異なり，言語情報に頭部動作情報を追加したモデルが一番良い性能であった．これは頭部情報がグループによる違いが少ない，識別に有効な情報を持っている，あるいは音声情報がグループ毎の性質に依存する情報である可能性がある．また，音声と頭部動作のモデルは僅差ではあるが，すべての教師あり学習モデルが AutoEncoder モデルの性能を上回った．

また，f-measure が最大となる時の閾値を用い，セグメント境界であると判定する precision, recall, f-measure を算出した．その結果を表 2 に示す．f-measure に着目すると，w2v_head モデルが最もよい性能であった．

表 2: モデル性能の比較

	w2v	head	audio	w2v_head	w2v_audio	head_audio	all	AutoEncoder
平均AUC	0.747	0.634	0.626	0.753	0.732	0.655	0.741	0.624
precision	0.211	0.177	0.131	0.253	0.228	0.181	0.223	0.148
recall	0.549	0.357	0.731	0.440	0.423	0.405	0.497	0.560
f-measure	0.305	0.237	0.222	0.321	0.297	0.25	0.308	0.234

6. まとめ

本研究では，言語・非言語情報に基づく，会話のセグメント境界の検出モデルを提案した．言語情報は最も有用なモダリティであったが，音声や頭部動作と組み合わせることで，より精度よく会話セグメント境界を検出できることが分かった．本研究では，議論参加者の身体情報として頭部動作情報を入力としたが，実際の議論ではポーズやジェスチャーも重要な意味を持っている．特にポーズは，例えば，前のめりになれば議論に対する積極性を示すなどセグメント境界の識別に有効であると考えられる．よって，頭部動作だけでなく姿勢位置などの全体的な身体情報も有効になると考えられる．

謝辞：本研究は，科学技術振興機構(JST) 戦略的創造研究推進事業(CREST)「実践知能アプリケーション構築フレームワーク PRINTEPS の開発と社会実践」(JPMJCR14E3)，および理化学研究所革新知能統合研究センターの支援を受けたものである

参考文献

[1]. Hearst, M.A., Multi-paragraph segmentation of expository text, in *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*, pp. 9-16, 1994.

[2]. Beeferman, D., A. Berger, and J. Lafferty, *Statistical Models for Text Segmentation.* Machine Learning, 1999. **34**(1-3): p. 177-210.

[3]. Galley, M., et al., Discourse segmentation of multi-party conversation, in *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics - Volume 1*. pp. 562-569 2003.

[4]. Le, Q. and T. Mikolov, Distributed representations of sentences and documents, in *Proceedings of the 31st International Conference on Machine Learning - Volume 32*, JMLR.org: Beijing, China. p. II-1188-II-1196, 2014.

[5]. Seokhwan Kim, Rafael E. Banchs, Haizhou Li, Exploring Convolutional and Recurrent Neural Networks in Sequential Labelling for Dialogue Topic Tracking, in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pp. 963–973, 2016.

[6]. 吉野 亮，八城 美里，高瀬 裕，中野 有紀子，会話エージェントによる優位性推定に基づくグループ会話への介入，第 29 回人工知能学会全国大会，1I4-4, 2015

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

コーパスに基づく日本語における謝罪への反応の分類

A corpus-based typology of reactions to apologies in Japanese

井上雅史　鈴木香里
Masashi Inoue　Kaori Suzuki

東北工業大学
Tohoku Institute of Technology

We proposed a typology of reactions to apologies base on the analysis of dialogue corpus. Our typology regards forgiving as the default reaction to the apologies. The responses are classified into two categories: the case when forgiving expressions are explicitly used and the case when forgiving expressions were left out. Then, non-default reactions were categorized based on the possible reasons why the apologies were not accepted.

1. はじめに

社会的な関係を維持，構築するために，会話中に謝罪あるいは詫びとそれに対する応答が出現する．知的な対話システムを実現するためには，エージェントがそれらを自然に使用することが求められる．謝罪の類型についての分析と比較して，謝罪に対する反応の類型については，あまり調べられていない．困難の原因の一つは，謝罪においては，一般に想起される典型的な表現が実際に使用される一方で，謝罪に対する反応では，典型的な表現が必ずしも使用されない傾向にあることと考えた．そこで，我々は反応の不在を考慮に入れた反応の類型を検討する．対話コーパスを使用することで，規範ではなく実際にどのような反応があり得るのかという観点での分類を試みる．

謝罪に対する応答に関して先行する調査に，Kitao らの英語を対象としたものがある [Kitao 14]．これは，言語教育への応用を志向している．我々の取り組みは，日本語を対象とし，反応の生成への応用を志向するものである．まず謝罪における典型的表現を定め，それに基づき謝罪発話を抽出する．次に，それらの謝罪に対する応答を分類する．

2. データ

対話データとして，名大会話コーパスを使用する．合計約 100 時間の日本語母語話者同士の雑談を文字化したものである [藤村 11]．主に大学院生にテープレコーダーか MD プレーヤーと外付けのマイクを渡し，30 分から 1 時間ぐらいの会話を収録している．本研究では，収録参加者の内 2 名の会話 95 件のデータを使用した．

3. 分析手順

3.1 謝罪の抽出

発話内で特定の表現が使われているものを謝罪とする．コーパス全体を俯瞰した予備調査から，対象とする表現として,『ごめん』『ごめんなさい』『すみません』『すまん』『失礼しました』『申し訳ありません』の 6 つを選定した．これらの表現が含まれる発話を謝罪としてコーパスからすべて抽出する．発話に対すして聞き手が言語的に反応した場合，その反応を『応答』とする．対話者毎に発話と応答のペアを抽出する．

連絡先: m.inoue@acm.org

謝罪に用いられる表現の使用傾向は，対話ペアの関係性に応じて変化することが実験的に示されている [趙 12]．例えば，親しい間柄では,「すみません」よりも「ごめん」の使用が好まれる．本研究においてコーパスから抽出した謝罪表現についても，話者間の人間関係と表現の使用比率を調べたところ，同様の傾向が確認された．

3.2 反応分類の手がかり

3.2.1 許しを表明する典型的表現

抽出した謝罪を含む発話とそれに対する反応のペアを分析者の主観により分析したところ，謝罪の言葉に対する反応を,『許す』『悩む』『許さない』『反応がない』『訂正』に分けることができた．そして，通常選択されるのは，謝罪を受け入れる，許す反応であった．明示的な発話によって許す場合には，典型的な表現の使用により許す場合と，反応しないことにより許す場合とが存在した．以下の 9 つの表現が典型的である．

1. いえいえ
2. いいよ
3. ううん
4. いやいや
5. 大丈夫
6. いいえ
7. そんな
8. 全然
9. とんでもない

3.2.2 許す応答表出可能

明示的な許しの表明がない場合,「許す応答表出可能」という基準で許す応答であると判断する.「許す応答表出可能」な反応とは，典型的表現が省略されている場合に相当し，典型的反応の挿入を行っても，反応の意味合いが変わらない．

典型的反応の表出が可能な場合の例：

発話：うん，ちょっとごめんね．
応答：鏡あるけど．

ここでの応答は，発話の謝罪について触れずに新たな提案をしている．次の様にいくつかの典型的表現を挿入した発話で，置き換えが可能である．

置換候補 1: いいよ．鏡あるけど．
置換候補 2：ううん．鏡あるけど．

置換候補 3: 大丈夫．鏡あるけど．
置換候補 4：全然．鏡あるけど．

次に，以上 3 つの分類（「典型的表現を使用した許し」「典型的表現を使用しない許し」「許す以外の反応」）全体の分析を行う．謝罪の回数上位 9 つのデータを使用し，3 つの分類に属する反応の使用割合の統計を取る．

4. 反応分析結果

4.1 典型的表現の使用

分析対象中に出現した典型的表現は，表 1 のような使用場面に対応していた．グループ化を行うと，謙遜系，信頼系，遠慮系の 3 つの系統に類別が可能である．謙遜系は，発話が敬語を話している場合や，感謝の言葉を含んでいる場合に使用される．信頼系は，発話がタメ口の場合，応答側が妥協をしている場合に使用される．遠慮系は，謝罪の打ち消し，謙遜を表すとき，許しを表すとき，親しい関係のときに使用される．

表 1: 許すための典型的表現と使用場面

典型的表現	使用場面
いいえ	謙遜を表すとき．
いいよ	許しを表す．友人などの親しい関係．妥協．
ううん	謝罪の対象の行為に対して，応答側に支障がないとき．
いやいや	遠慮を表すとき．逆に，謝罪の理由が明確に違う時も使用．
大丈夫	謝罪の対象の行為に対して，応答側に支障がないとき．
いいえ	謝罪の対象の行為に対して，応答側に支障がないとき．３０代以上会話で使用する． 友人などの近い間で使用される． 敬語もたまに出るような関係で使用される．

4.2 許す応答表出可能

明示的に許すことが表明されていない場合に，典型的な表現が挿入候補として現れる割合は，信頼系 70.6%，謙遜系 22.5%，遠慮系 3.9% であった．

4.3 許す以外の反応

許す以外の反応の場合は，謝罪を受け入れなかった原因の観点から，以下の 9 つの場合に類別が可能であった．

1. 発話者の 1 ターンがとても長い場合
2. 1 つ前のターンの謝罪に対して発話者が謝罪している場合
3. 発話から次の質問をされている場合
4. 発話者の謝罪理由を確認している場合
5. 発話者の態度に呆れている場合
6. 発話者の謝罪理由が違っている場合
7. 発話者の謝罪に対して納得していない場合
8. 発話者のフォローをしている場合
9. 発話者の次の発言を待っている場合

4.4 集計結果

謝罪の回数が多かった上位 9 つのデータの反応別の統計を取ると図 1 の様な結果だった．参加者 F130 の場合，data112，data110，data111 において，典型的表現を使用せず，許しを表明していた．この様に，典型的表現の使用には謝罪を受け入れる際に必須ではない可能性が高い．

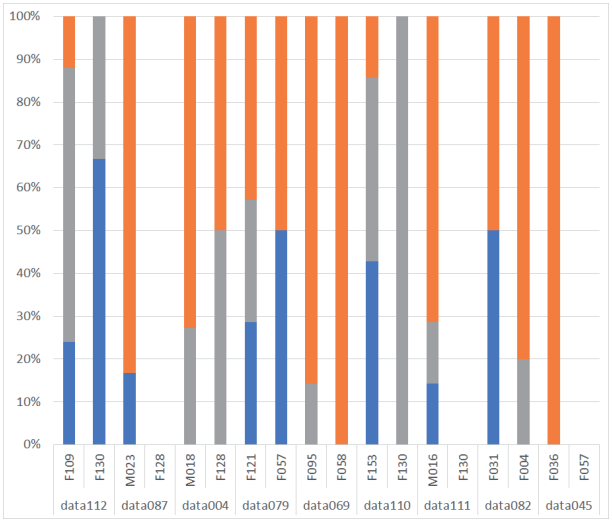


図 1: 話者ごとの反応の種別割合．横軸は話者とデータ ID．各バーの下から「典型的表現を使用した許し」「典型的表現を使用しない許し」「許す以外の反応」を表す．

5. おわりに

対話コーパスから抽出したデータを基に，謝罪への反応を，「典型的表現により許す反応」，「許す応答表出可能な反応」，「許す以外の反応」に類別した．「典型的表現により許す反応」には，9 つのよく使用される典型的表現が存在し，それらは使用する場合によって謙遜系，信頼系，遠慮系に類別が可能であった．「許す応答表出可能な反応」は，典型的表現が発せられていないが，文脈から許しの意味と捉えることが可能な反応である．「許す以外の反応」は，許さない原因から，9 つの場合に類別が可能であった．

本研究の結果は，人間同士の会話においては，明示的に謝罪を受け入れる表現を使用しないことによって，謝罪を受け入れる反応が頻繁に使用されていることを示唆しており，対話システム等においても，応用可能と考えられる．ただし，反応を生成する場合には，どのような分類があるかの情報だけでは不足であり，例えば謝罪の表現をいくつかの類に分ける試みを参考に [滝浦 08]，反応の種類との関係を調査するなど，反応選択の規則を明らかにしていくことが必要となる．

参考文献

[Kitao 14] Kitao, S. K. and Kitao, K.: A corpus-based study of responses to apologies in US English, *Journal of culture and information science*, Vol. 9, No. 2, pp. 1–13 (2014)

[滝浦 08] 滝浦 真人：ポライトネス入門, 研究社 (2008)

[藤村 11] 藤村 逸子, 大曽 美恵子, 大島ディヴィッド義和：会話コーパスの構築によるコミュニケーション研究, 藤村逸子, 滝沢 直宏（編）, 言語研究の技法：データの収集と分析, pp. 43–72, ひつじ書房 (2011)

[趙 12] 趙 翻：日本語と中国語における謝罪表現の対照研究：家族と親友間の異なりに注目して, 東洋大学大学院紀要, Vol. 49, pp. 124–98 (2012)

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

[4G2-OS-8a] マイニングと知識創発(1)

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

Fri. Jun 7, 2019 12:00 PM - 1:40 PM Room G (302A Medium meeting room)

[4G2-OS-8a-01] Proposal of Summarization Support System Based on Mixed-Initiative Interaction

○Hiroto Kato[1], Yasufumi Takama[1] （1. Tokyo Metropolitan University）

12:00 PM - 12:20 PM

[4G2-OS-8a-02] *differential* PLSA

○Koji Nomori[1] （1. Analytics Design Lab Inc.）

12:20 PM - 12:40 PM

[4G2-OS-8a-03] Basic Analysis of Descriptions on Medicine-care in Nursing Records

○Seiya Uesaka[1], Yoshinori Hijikata[1], Muneou Suzuki[2] （1. Kwansei Gakuin University , 2. University of Miyazaki Hospital ）

12:40 PM - 1:00 PM

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

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

1:00 PM - 1:20 PM

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

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

1:20 PM - 1:40 PM

©The Japanese Society for Artificial Intelligence

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

相互主導型インタラクションに基づく要約生成支援システムの提案

Proposal of Summarization Support System Based on Mixed-Initiative Interaction

加藤　大登　　高間　康史
Hiroto Kato　　Yasufumi Takama

首都大学東京大学院システムデザイン研究科
Graduate School of System Design, Tokyo Metropolitan University

This paper proposes a text summarization support system based on mixed-initiative interaction. Automatic summarization is used to generate sentences useful as a clue for information gathering. However, the quality of the generated summary is still lower than that generated by humans. Furthermore, it does not consider users' interest and knowledge. To solve the problems, the proposed system dynamically suggests what is useful for writing a summary, such as important sentences, in response to users' input. The effectiveness and challenges of the proposed approach are discussed based on the result of the experiment using a prototype system.

1. はじめに

関連研究の論文サーベイを行う際などには，論文要約が重要な手がかりとなる．しかし，要約文を作成するには元の文章内容を十分に理解した上で内容を吟味して取捨選択する必要があるため，人手で行うには時間と労力がかかる．この課題を解決するために自動要約の研究が行われている．しかし現状の自動要約文の精度は人手で作成されたものに劣るほか，ユーザが要約作成プロセスに関与することができない．

これらの問題に対し，本稿ではユーザと計算機が相互主導的に要約文を作成するアプローチを提案する．提案するアプローチではユーザが計算機の提示内容を確認し，計算機はユーザの入力に合わせて提示内容を変更することで，効率的な要約作成を支援する．プロトタイプシステムを構築してユーザ実験を行い，ユーザが有用性を感じる機能や必要としている機能を調査した結果を示す．

2. 関連研究

自動要約手法の代表的アプローチに抽出型があり，LexRank[Erkan 04] は他の文との類似度が高い文，また重要文との類似度が高い文を抽出して要約を生成する．ユーザの文章入力支援を目的としたシステムには英語論文執筆支援ソフト BEAR[Narita 02] があり，オンライン辞書へのリンク機能や単語の用例を提示する機能が実装されている．相互主導型インタラクションを取り入れたシステムの例として，詩の作成支援システム CoPoeTryMe[Oliveira 17] が挙げられる．詩の自動生成システムのパラメータをユーザが変更可能であり，生成内容をユーザ自身が制作中の詩の一部として取り入れることができる．

3. 提案手法

提案システムの構成を図 1 に示す．フロントエンドは，対象文書の内容やユーザの入力を受け取ってバックエンドに送信し，バックエンドの計算結果をユーザインタフェースに反映する．バックエンドは，データセットから事前に単語の IDF 値と分散表現を作成する機能と，対象文書から文書モデルを作成する機能，ユーザの入力をもとに提示内容を構成する機能を持つ．単語の IDF 値と分散表現の元になるデータセットは Wikipedia 日本語版の Dump ファイル *1 を用いる．

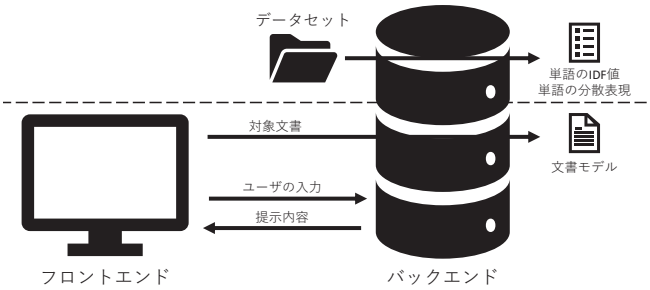


図 1: システム構成

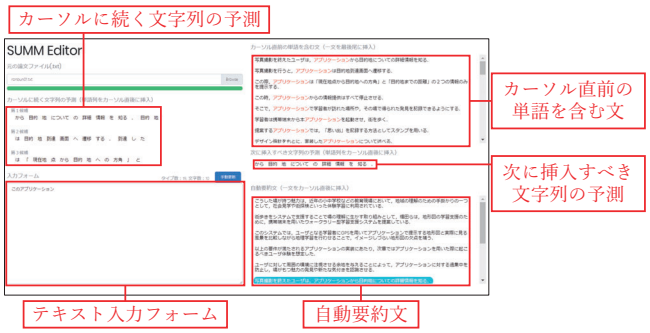


図 2: フロントエンドの画面構成

図 2 はフロントエンドの画面構成で，テキスト入力フォームと提示スペースから構成される．提示スペースの各欄にはユーザの要約文作成を支援する情報が提示され，クリックしてテキスト入力フォームに挿入できる．提示内容はテキスト入力フォーム内でのキー入力に従い，自動要約の出力を利用して動的に生成する．自動要約文の中で青色に強調された文は，現在入力中の位置に挿入すべき文であることを示している．

バックエンドでは，フロントエンドから受け取ったユーザ入力に従い，提示内容を以下の手順で動的に計算する．対象文書中の各文の重要度は LexRank アルゴリズム [Erkan 04] で求める．

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

*1 https://dumps.wikimedia.org/jawiki/20181001/jawiki-20181001-pages-articles-multistream.xml.bz2

1. テキスト入力フォームの，カーソル直前の単語 x を取得する．
2. 単語の分散表現と cos 類似度を用いて，対象文書中から x に類似した単語を探索する．
3. x の類似単語を含む文を重要度順に並べ「カーソル直前の単語を含む文」とする．
4. 手順 3 の文で重要度が最も高い 3 文の，x の類似単語以降の語句列を「カーソルに続く文字列の予測」とする．
5. テキスト入力フォームの文章を S，手順 3 の各文を D_i として式 1 によりスコアを求め，最も高い 1 文の x の類似単語以降の語句列を「次に挿入すべき文字列の予測」とする．
6. 手順 5 の文を S，対象文書中の各文を D_i として式 1 によりスコアを求める．最も高い 9 文に S を合わせた計 10 文を対象文書中の出現順に並べ「自動要約文」とする．

$$Score(D_i) = Imp(D_i) - Sim(D_i, S) \quad (1)$$

式 1 は多様な文を選択することを意図したものであり，Imp は重要度を，Sim は類似度を意味する．類似度は文を単語の集合とみなし Jaccard 係数を用いて計算する．

手順 2 で分散表現を利用することで，単語 x が対象文書中に存在しない場合でもシステムが機能する．手順 5 では，重要だがまだ執筆されていない情報を持つ文を「次に挿入すべき文字列」と想定し抽出している．手順 6 では LexRank の結果を利用しつつ「次に挿入すべき文字列」とは類似しない文から自動要約文を生成している．

4. ユーザ実験

4.1 実験内容

要約執筆支援エディタ「SUMM Editor」を実装し，実験協力者 12 名に論文要約の執筆を行ってもらう評価実験を行った．本実験ではエディタの機能と，要約文の文字数などの目安を伝えた後に執筆を開始してもらった．要約対象となる論文には実験前に目を通してもらっており，執筆中も閲覧可能とした．制限時間は設けず，協力者自身が要約作成の完了を宣言したとき実験を終了した．また実験終了後にアンケートを行い，各機能の評価や意見，対象論文の理解度を調査した．評価値は 5 段階で，1 が最も否定的，5 が最も肯定的な評価を表す．

4.2 実験結果・分析

キータイプ回数および機能ボタンのクリック回数と執筆時間の相関係数を表 1 に示す．クリック回数と執筆時間には正の相関があり，特に「自動要約文」「次に挿入すべき文字列の予測」機能の相関が高いことがわかる．この二つはシステムから情報を推薦する機能であり，またアンケートに「とりあえず入れてみてから文の校正を考えた」などの回答があったことから，要約文執筆に行きづまった実験協力者が発想支援を受けるために利用したと考える．

各機能に関するアンケート結果について，評価値ごとの回答人数を表 2 に示す．Q1 では，「カーソル直前の単語を含む文」の機能の評価が比較的高かった．アンケートでは「同じ単語がどの文で出ているかがわかり要約作成の手助けとなった」という意見があり，単語を検索して文書内での使われ方を確認するという想定通りの用途で使用されたと言える．

Q2 では，文単位の挿入機能である「自動要約文」の機能の評価が比較的高かった．アンケートには「打ち込む手間が省けた」や「自動要約文の文章を挿入し不必要な部分を取り除いていく」という意見があり，要約執筆に活用できたと考える．しかし内容の参考度合いの評価については「悩んだ時の参考にした」との一方で「あまり気にならなかった」との意見もあり，人により極端に分かれることが判明した．

表 1: 各操作と執筆時間の相関係数

操作		相関係数
キータイプ		0.3240
機能ボタンのクリック		0.6062
機能別	自動要約文	0.6418
	カーソルに続く文字列の予測	0.1731
	カーソル直前の単語を含む文	0.4770
	次に挿入すべき文字列の予測	0.6308

表 2: 各機能の評価値ごとの回答人数

（Q1. 内容が参考になったか，Q2. 挿入機能が役立ったか）

設問	Q1					Q2				
評価値	1	2	3	4	5	1	2	3	4	5
自動要約文	0	6	0	3	3	1	1	4	4	2
カーソルに続く...	2	4	1	5	0	3	2	3	2	2
カーソル直前の...	1	1	3	5	2	3	2	2	3	2
次に挿入すべき...	3	2	6	1	0	3	2	6	1	0

単語単位での挿入機能（「カーソルに続く文字列の予測」と「次に挿入すべき文字列の予測」）は両設問とも低評価の方が多かった．「自動要約文」の挿入機能が比較的高評価であることからも，対象文書内での意味を確認しやすい文単位での挿入のほうが，単語単位の挿入より使いやすかったと考える．

そのほかに「直観的な操作は今回のタスクを楽にしてくれた」と相互主導的インタラクションの簡易な操作が評価される一方で，「suggest までの時間が少しネック」「すでに要約文に入力してある文と被る部分がないように文を挿入してほしい」といった意見があった．システムを改善するには，簡易な操作方法を維持しつつ要約生成支援により適した機能を提供する必要があると考える．

5. おわりに

本稿ではユーザと計算機が相互主導的に要約文を作成するアプローチを提案し，プロトタイプ「SUMM Editor」によるユーザ実験で有用な機能や必要とされる機能を調査した．

相互主導的インタラクションは直感的な操作で情報提示を行うことができる．しかし，ユーザの求めるシステムへと改善するためには，ユーザの入力に合わせた精度の高い提示を行うことに加えて，インタフェースもより要約執筆に適したものに改善する必要があることが示された．今後は計算時間の短縮と文挿入機能の利便性向上，本用途に適した要約生成アルゴリズムの検討を進めていく予定である．

参考文献

[Erkan 04] G. Erkan, D. R. Radev, LexRank: Graph-based Lexical Centrality as Salience in Text Summarization, Journal of Artificial Intelligence Research, Vol. 22, pp. 457–479 (2004)

[Narita 02] M. Narita, K. Kurokawa, and T. Utsuro, A web-based English abstract writing tool using a tagged E-J parallel corpus, LREC 2002, pp. 2115–2119 (2002)

[Oliveira 17] H. G. Oliveira, T. Mendes, and A. Boavida, CoPoeTryMe: a Co-Creative Interface for the Composition of Poetry, ICNL 2017, pp. 70–71 (2017)

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

differential PLSA
－テキスト情報の典型的なトピックではないより個性的なトピックの抽出－

differential PLSA
A Method of Extracting not Representative Topics but More Individual Topics from Text Data

野守　耕爾[*1]
Koji Nomori

[*1] 株式会社アナリティクスデザインラボ
Analytics Design Lab Inc.

Abstract: This study proposes a new method extracting topics from text data named *differential* PLSA. The method applies PLSA to a differential co-matrix computed by logarithm of the ratio of an observed co-matrix to expected one. It enables to extract not representative topics but more individual ones. This paper showed the effectiveness by applying the method to patent document data and comparing with results using normal PLSA. As a result, topics extracted by the method were composed of less frequent and more concrete elements, and they were more individual.

1. はじめに

データ活用がビジネスで進む昨今，テキストデータを分析することで定性情報によるインサイトを獲得して業務に活用しようとする取り組みも増えている．筆者は企業におけるデータ分析・活用のコンサルティング事業を展開しているが，ビジネスにおけるテキストデータの活用では，例えば顧客満足度調査などで実施されるアンケートの自由記述文を分析し，回答者の生の声から潜在的なニーズを抽出することでマーケティングに活用したり，コールセンターの問い合わせ履歴のデータを分析し，消費者ならではの隠れた評価の観点を抽出することで商品・サービスの改善を検討したり，特許公報の文書データを分析し，技術トレンドや他社の得意技術などを抽出することで研究開発戦略や他社との提携戦略を検討するなど，様々な業務領域でテキストデータの活用が進められている．

ビジネスにおけるテキストデータの活用では，新たな気づきとなるインサイトをテキスト情報から抽出することがしばしば求められるが，結果が期待外れということも少なくない．その理由は，業務担当者からすると経験的によく知っている結果であり，目新しさがないということが多い．テキストデータの分析には，テキストマイニング技術を適用することが通例であるが，テキストマイニングは文章に含まれる単語を抽出し，主に文章の全体像を各単語の出現頻度をベースに理解する手法となる．ここで，頻度の多い単語では当たり前の結果となるため，頻度の少ない単語まで確認しようとすると，今度は結果が複雑になり何が重要なのか分からず，結果の理解ができなくなるというジレンマに陥ってしまう．特に対象となる文章がビッグデータになると，抽出される単語も膨大となり，単語ベースの結果では複雑になりすぎるため，文章を単語ではなく意味的なまとまりのあるトピックをベースに理解する手法として，トピックモデルの適用が有効となる．トピックモデルでは単語の出現データを教師なし学習によりトピックに集約するが，これは分析対象のテキストデータ全体をうまく表現するようないわば数学的な最適解となるトピックを抽出するので，どうしても典型的なトピックになりがちになり，ビジネスの業務担当者の視座では満足しがたい結果になってしまう．

そこで本稿では，テキストデータにトピックモデルを適用した際に，典型的なトピックだけではなく，より個性的なトピックを抽出する方法を提案し，その適用事例を紹介する．

連絡先：野守耕爾，株式会社アナリティクスデザインラボ，
koji.nomori@analyticsdlab.co.jp

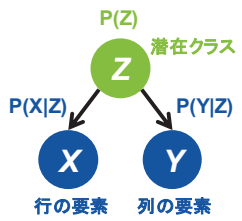
2. PLSA

本稿では従来のトピックモデルとして PLSA（確率的潜在意味解析：Probabilistic Latent Semantic Analysis）を紹介する．

2.1 PLSA の概要

PLSA は，行列データ（共起行列）の行要素 X と列要素 Y の背後にある共通特徴となる潜在クラス Z を抽出する次元圧縮法であり，元々は文書分類のために開発された手法である [Hofmann 1999]．PLSA のグラフィカルモデルを図 1 に示す．PLSA の計算では，X と Y の共起確率を潜在クラス Z を使って式展開し，この対数尤度関数を最大化する EM アルゴリズムを実行することで，3 種類の確率変数 $P(X|Z)$, $P(Y|Z)$, $P(Z)$ が計算される．文書分類における PLSA では，文書 X とそこに出現する単語 Y の間には潜在的な意味トピック Z があることを想定し，各文書における単語の出現頻度が記録された「文書」×「単語」の共起行列データを学習し，文書と単語の共通特徴となるトピックを見つける．これにより「文書」×「トピック」という低次元データに変換して文書分類できる．

PLSA はクラスタリングの手法としても用いられるが，クラスタリングという観点で PLSA が他のクラスタリング手法と異なる特徴の一つは，行と列を同時にクラスタリングできることである．一般的なクラスタリング手法は，列をベースに行をクラスタリングする，あるいは行をベースに列をクラスタリングするため，どちらか一方しかクラスタリングできない．一方 PLSA で抽出される潜在クラスには，行の要素と列の要素が同時に所属することができる．



$$P(X,Y) = \sum_Z P(X|Z)P(Y|Z)P(Z)$$

図 1　PLSA のグラフィカルモデル

2.2 PLSA の共起行列構成の工夫

行と列を同時にクラスタリングできる PLSA では，行と列は双方が十分意味を持つ情報で構成すれば，抽出された潜在クラスの意味を行と列の 2 つの情報軸から解釈することができる．本来の PLSA の適用では，「文書」×「単語」という構成の共起行列をインプットとするが，この構成を工夫することで解釈のしやすい潜在クラスを抽出する試みがある．

例えば，「品詞」×「品詞」の共起行列を用いる方法や[Kameya 2005][野守 2014]，「単語」×「係り受け」の共起行列を用いる方法が提案されている[野守 2018]．「単語」×「係り受け」の共起行列を用いる PLSA では，単語と係り受け表現が同時にクラスタリングされ，単語という話題の観点となる軸に基づき，その観点の具体的内容となる係り受け表現をグルーピングできるため，より文脈上近しい言葉・表現でまとめられた解釈のしやすいトピックを潜在クラスとして抽出できるとされている．

3. *differential* PLSA

テキストデータに PLSA を適用したとき，典型的なトピックだけではなく，より個性的なトピックを抽出する方法として，*differential* PLSA（以下 *diff*-PLSA）を提案する（特許出願中，商標出願中）．

diff-PLSA では，テキストデータにおいて行要素 X_i と列要素 Y_j が共起する実測頻度 $n(X_i, Y_j)$ を値として持つ通常の共起行列 M に加え，行要素 X_i と列要素 Y_j が共起する期待頻度 $n'(X_i, Y_j)$ を値として持つ共起行列 M' を構成する．この各期待頻度に対する実測頻度の比率の対数を取った $log\left(n(X_i, Y_j)/n'(X_i, Y_j)\right)$ を値として持つ *differential* 共起行列（以下 *diff*-共起行列）を構築し，これに PLSA を適用する．*diff*-PLSA で適用する共起行列のイメージを図 2 に示す．ここで期待頻度とは，X_i の総頻度（出現文章数）$n(X_i)$，Y_j の総頻度（出現文章数）$n(Y_j)$，総文章数 N から式(1)のように計算される．なお，*diff*-共起行列の値が負数となるものは 0 に置換する．

$$n'(X_i, Y_j) = n(X_i) \cdot n(Y_j)/N \quad (1)$$

実測頻度の共起行列に適用する通常の PLSA では，その解を求める最適化計算において，どうしても頻度が高い要素に高い確率が割り当てられ，結果として抽出されるトピックは典型的なものになる傾向があり，目新しさに欠けてしまう．一方，*diff*-共起行列では，実測共起頻度を期待共起頻度で除した値を持つが，実測共起頻度が高い共起ペアでも，元々全体の頻度が高い要素が含まれるときには期待共起頻度も高くなるため，実測共起頻度を期待共起頻度で除すことで値の大きさが制限される．逆に実測共起頻度が高くない共起ペアでも，期待共起頻度がそれよりも十分低ければ *diff*-共起行列での値は大きくなり，これに PLSA を適用した解ではこうした要素にも高い確率が割り当てられる可能性がある．つまり，通常の PLSA では頻度が低い要素は高い確率が割り当てられない傾向があるが，*diff*-PLSA ではそうした要素にも高い確率が割り当てられる可能性があり，より個性的なトピックが抽出されることが期待できる．

期待共起頻度に対する実測共起頻度の比率に対数を取る理由としては，極端に高くなる値を制限ためである．特に期待共起頻度は 1 未満となるケースも多く，比率のみでは値が高くなりすぎるものもある．この状態では共起行列全体の値の分布は大きくばらつき，極端な値の開きが生まれてしまうため，PLSA を適用した際の最適化計算において，今度はこの極端に大きな値に引っ張られる結果となり，必要以上にデフォルメされた歪んだトピックとなることが考えられる．そこでこの比率の値の対数を取ることで値の分布をならし，上記の現象を制限する．

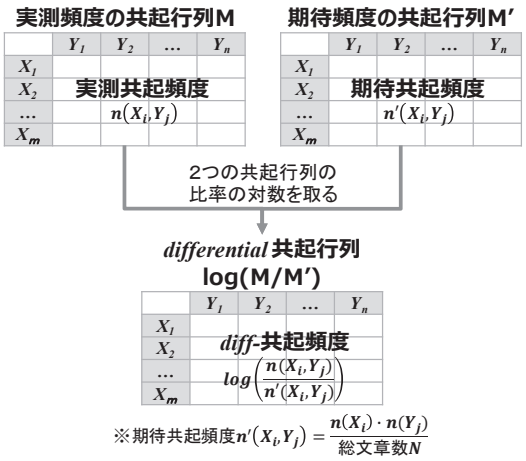


図 2　*diff*-PLSA で適用する共起行列

4. *diff*-PLSA の適用事例

本稿では特許の要約文のテキストデータを対象に，通常の PLSA と *diff*-PLSA を適用した事例を紹介し，その結果を比較することで *diff*-PLSA の有効性を検討する．

4.1 対象データと共起行列の構成

特許の要約文と請求項に「車」「電気」を含む 10 年分（出願日が 2007 年 1 月 1 日～2016 年 12 月 31 日）の国内特許公報 26,419 件を用意し，その要約文の文章データを分析対象とした．これにより電気自動車関連の技術トピックを抽出する．

共起行列の行列構成は，行を「名詞」，列を「係り受け」とした．係り受けは文法的につながりのある単語のペアを抽出するものだが，名詞に対する動詞・形容詞・形容動詞のペアを対象とした．こうした構成の共起行列に PLSA を適用することで，使われ方の似ている名詞と係り受けで構成されるトピックを得ることができる．つまり技術を象徴する名詞とその名詞を修飾する表現でまとめられたトピックを抽出でき，解釈しやすい結果が得られると期待できる．要約文にテキストマイニングを実行し，名詞と係り受けを抽出したところ，データ頻度が 20 件以上となる名詞が 3,020 語，係り受けが 2,128 表現得られた．この「名詞 3,020 語」×「係り受け 2,128 表現」で構成される共起行列を PLSA および *diff*-PLSA で適用するベースとした．

4.2 通常の PLSA の適用

通常の PLSA を適用する共起行列において，名詞と係り受けの各ペアの共起頻度とは，そのペアが同時に出現する文章数とした．ここでいう文章とは，句点で区切られた一文を指し，今回の 26,419 件の特許データに含まれる総文章数は 229,598 件であった．この共起行列に PLSA を適用してトピックを抽出するが，PLSA は予めトピック数を設定する必要があり，また初期値により解が異なる特性がある．そこでトピック数を 1 刻みで変化させ，それぞれのトピック数に対して PLSA を初期値を変えて 5 回ずつ実行し，それぞれの解を情報量基準 AIC で評価して最も評価の良い解を採用した．その結果 34 個のトピックが得られた．

PLSA のアウトプットは，①トピック Z における行要素 X（名詞）の所属確率 P(X|Z)，②トピック Z における列要素 Y（係り受け）の所属確率 P(Y|Z)，③トピック Z の存在確率 P(Z)，という 3 つの確率が計算される．特に各トピックにおいて①②の確率が高い名詞，係り受けを確認し，そのトピックの意味を解釈する．抽出されたトピックの内容例を表 1 に示す．表 1 ではトピック Z05 と Z14 について，所属確率の高い上位 7 つの名詞および係り受けと，

それぞれの所属確率 P, 頻度 n(出現文章数)についてまとめたものである. トピック Z05(表 1 上)は,「ブレーキ」や「制動力」に関する表現で確率が高いため, ブレーキに関する技術と解釈できる. トピック Z14(表 1 下)は,「給電」や「電力の供給」に関する表現で確率が高く, また「非接触」という表現も上位にあることから, 非接触受電などの給電装置に関する技術と解釈できる. このように解釈を付けた 34 個のトピックの一覧を表 2 に示す.

表 1 通常の PLSA で抽出されたトピックの内容の例

トピックZ05

P(XZ)	n(X)	X:単語	P(YZ)	n(Y)	Y:係り受け
7.3%	779	ブレーキ	3.2%	92	車両-ブレーキ
5.0%	1,384	作動	2.4%	33	ブレーキ液圧-発生
3.4%	5,188	モータ	2.2%	53	制動力-発生
2.9%	279	制動力	1.8%	49	ブレーキ-備える
2.9%	867	運転者	1.7%	32	操作量-応ずる
2.7%	1,117	車輪	1.6%	82	ブレーキ-提供
2.6%	1,005	操作	1.6%	59	電気信号-基づく
…	…	…	…	…	…

トピックZ14

P(XZ)	n(X)	X:単語	P(YZ)	n(Y)	Y:係り受け
9.7%	1,162	給電	4.3%	1,350	電力-供給
5.1%	3,629	電力	2.7%	115	給電-行う
3.6%	1,140	電源	2.5%	52	電力-受電
3.4%	329	給電装置	2.0%	28	給電-電力
2.4%	4,655	電気自動車	1.6%	124	電源-接続
2.4%	180	非接触	1.5%	20	駐車装置-変化
2.4%	1,052	外部	1.5%	34	非接触-受電
…	…	…	…	…	…

表 2 通常の PLSA で抽出された 34 個のトピック一覧

No.	トピック名	No.	トピック名
Z01	エンジンの始動と停止	Z18	演算・推定
Z02	動力の伝達	Z19	機器の異常検出
Z03	モータ駆動	Z20	操作スイッチ
Z04	ロータ・ステータなど回転部品の構成	Z21	筐体
Z05	ブレーキ装置	Z22	表面の形成
Z06	動作制御	Z23	位置とその移動
Z07	動力伝達の制御	Z24	配置・位置・方向
Z08	スイッチの切り替え	Z25	構成の方位
Z09	交流・直流の変換	Z26	構成
Z10	エネルギーの変換	Z27	接続
Z11	電池モジュールの提供	Z28	方法の提供
Z12	二次電池の構成	Z29	損傷や浸水など不具合の防止
Z13	電気自動車の蓄電池充電	Z30	小型化・簡素化・低コスト化など付加価値
Z14	非接触受電など給電装置	Z31	効率性・安全性の向上
Z15	外部への電力供給	Z32	既存エンジンへの警鐘・樹脂組成物の提供
Z16	空調などの冷却・加熱	Z33	重力発電の活用による地球温暖化防止
Z17	情報通信	Z34	タービン発電の出力向上・燃費低減

4.3 各文章に対するトピックのスコア計算

全特許データの各文章に対して各トピックのスコア(該当度)を計算した. これにより各トピックと関連の強い具体的な文章を確認できるため, トピックを解釈するうえの参考情報となる. 実際に表 2 の解釈においてもこのスコア結果を使用した.

文章 S_h におけるトピック Z_k のスコアは $P(S_h|Z_k)/P(S_h)$で定義した. これはトピックを条件とすることで文章の発生確率が何倍になるのかを示し, そのトピックをよく話題にしている文章ほど高くなる. 以下, $P(S_h|Z_k)$と $P(S_h)$の計算について説明する.

$P(S_h|Z_k)$は, 文章 S_h を名詞で定義される文章 Sx_h と係り受けで定義される文章 Sy_h に分解し, それぞれ $P(Sx_h|Z_k)$と $P(Sy_h|Z_k)$を計算し, それらを統合して $P(S_h|Z_k)$を計算する. $P(Sx_h|Z_k)$と $P(Sy_h|Z_k)$は式(2)(3)で計算される. 名詞 X_i と係り受け Y_j の頻度(出現文章数)をそれぞれ $n(X_i)$と $n(Y_j)$とすると, $P(Sx_h|X_i)$は $n(X_i)$の逆数, $P(Sy_h|Y_j)$は $n(Y_j)$の逆数として計算される. $P(X_i|Z_k)$と $P(Y_j|Z_k)$はそれぞれ PLSA の実行結果で得られている. $P(S_h|Z_k)$は式(4)で計算され, 文章 S_h において Sx_h と Sy_h の重みは同じであるため, $P(S_h|Sx_h)$と $P(S_h|Sy_h)$はそれぞれ 0.5 とする. $P(S_h)$は式(5)で計算され, $P(Z_k)$は PLSA の実行結果で得られている.

$$P(Sx_h|Z_k) = \sum_i P(Sx_h|X_i)P(X_i|Z_k) \quad (2)$$

$$P(Sy_h|Z_k) = \sum_i P(Sy_h|Y_i)P(Y_i|Z_k) \quad (3)$$

$$P(S_h|Z_k) = P(S_h|Sx_h)P(Sx_h|Z_k) + P(S_h|Sy_h)P(Sy_h|Z_k) \quad (4)$$

$$P(S_h) = \sum_k P(S_h|Z_k)P(Z_k) \quad (5)$$

上記のトピックのスコアは文章単位に計算したものだが, これを特許単位に集約し, そのスコアを特許の各属性情報(出願年や出願人など)の軸で集計すれば, トピックと属性との関係を分析することもできる. こうした方法で技術のトレンドを可視化したり, 出願人のポジショニングを可視化し, 技術戦略の検討に用いる試みもされている[野守 2018].

4.4 *diff*-PLSA の適用

diff-PLSA で採用する共起行列もその行列構成は 4.2 の通常の PLSA で使用した「名詞 3,020 語」×「係り受け 2,128 表現」の共起行列の構成と同様となるが, 先述の通り *diff*-PLSA では通常の PLSA で構築した実測頻度の共起行列に加え, 期待頻度の共起行列を構築する. 期待頻度の計算は式(1)の通りで, 総文章数 N は 229,598 件である. この期待共起頻度に対する実測共起頻度の比率の対数を取ることで *diff*-共起行列を構築し, これに PLSA を 4.2 と同様の方法で適用したところ, 最も評価の良い解として 50 個のトピックが得られた. 4.2 と同様に, 各トピックにおける名詞および係り受けの所属確率と, 4.3 の方法で計算される各文章のトピックのスコアから, 50 個のトピックに解釈を付けた. その一覧を表 3 に示す.

表 3 *diff*-PLSA で抽出された 50 個のトピック一覧

No.	トピック名	No.	トピック名
Z'01	エンジン制御	Z'26	電流・電圧の検出
Z'02	動力伝達	Z'27	温度、電流、充電量などの検出と制御
Z'03	差動機構などを備えた動力伝達の制御	Z'28	演算や推定、測定などのステップを含む方法
Z'04	回転運動	Z'29	情報の取得・提供(位置情報やバッテリ残量等)
Z'05	ロータ・ステータなどモータの構成	Z'30	スイッチなど操作装置
Z'06	モータ制御(トルク制御や回転数制御等)	Z'31	車両用灯具
Z'07	油圧ポンプなどを利用したモータ駆動	Z'32	掃除機
Z'08	ブレーキ	Z'33	基板の構成
Z'09	状態に応じた制御、運転者の操作補助	Z'34	回路の接続(電力変換回路等)
Z'10	コンバータとバッテリ昇降圧	Z'35	端子接続
Z'11	直流と交流の電力変換	Z'36	部品・装置の収容ケース・筐体
Z'12	回転力などの電気エネルギー変換	Z'37	部品・装置の配置
Z'13	エネルギー効率の向上	Z'38	パーツなどの移動、位置
Z'14	発電と蓄電	Z'39	構造の形成・方位
Z'15	電池モジュールの提供	Z'40	支持構造
Z'16	燃料電池	Z'41	装置やユニットの構成
Z'17	二次電池の構成	Z'42	システム・方法の構成
Z'18	バッテリの充放電	Z'43	その他方法
Z'19	充電システム	Z'44	組成物の製造方法(樹脂や電解液等)
Z'20	充電の接続	Z'45	機能性組成物・成形品(耐熱性や耐衝撃性等)
Z'21	非接触など受給電装置	Z'46	製造の効率化(小型化や低コスト化等)
Z'22	車両用空調など熱交換	Z'47	不具合の防止(損傷、感電、盗難等)
Z'23	冷却装置と放熱	Z'48	その他
Z'24	信号の入出力と検出	Z'49	タービン発電と船舶・飛行機への応用
Z'25	電気信号の取得と変換(センサ検出等)	Z'50	重力発電の活用による地球温暖化防止

4.5 通常の PLSA と *diff*-PLSA の比較

通常の PLSA の結果と *diff*-PLSA の結果を比較したところ, ① *diff*-PLSA の方が通常の PLSA よりも頻度の少ない名詞・係り受けでトピックが構成される傾向があること, ②*diff*-PLSA では通常の PLSA では抽出されないトピックが抽出されること, ③*diff*-PLSA の方が通常の PLSA よりもトピックの解釈が難しいことが考察された. 以下それぞれの考察について説明する.

① 頻度の少ない名詞・係り受けでトピックが構成されること

①の考察に関しては, 通常の PLSA と *diff*-PLSA で同様の解釈ができるトピックの内容を比較すると分かりやすい. 表 1 に示した通常の PLSA による「Z05:ブレーキ装置」「Z14:非接触受電など給電装置」に対応する *diff*-PLSA のトピックはそれぞれ表 3 の Z'08 と Z'21 であり, その内容を表 4 に示す. ブレーキに該当する Z'08 は, 通常の PLSA と比べて,「マスタシリンダ」や「液圧」など頻度は少ないがブレーキに関連するより具体的な表現が上位に位置している. また, 非接触受電などの給電装置に該当する Z'21 は,「非接触」が最も確率が高く, また「送電コイル」や「受電コイル」といった頻度は少ないが具体的な表現が上位に位置している. 電気自動車において非接触の充電システムは

今後重要な技術であり，その仕組みは駐車場の路面に設置した送電コイルと車両に搭載された受電コイルが重なり合うことで電磁誘導が発生し電力が供給されるという技術が代表的であり，*diff*-PLSA ではそうした技術のより具体的な表現がトピックの上位語となっている．つまり同様の意味を示すトピックでも *diff*-PLSA の方がより細かい技術を示すものが抽出されている．

また，各トピックにおいて所属確率の高い上位語の頻度 *n* が，通常の PLSA よりも *diff*-PLSA の方が少ないことを確認するため統計的検定を行った．各トピックにおいて所属確率の高い順に名詞および係り受けを並べたとき，累積確率が 50%になるまでの名詞および係り受けの平均頻度を検定用データとし，通常の PLSA と *diff*-PLSA におけるこの値のトピック平均の差について Welch の t 検定を実施した．その結果，通常の PLSA（n=34）では名詞が平均 1288.2 件（SD=686.5），係り受けが平均 85.6 件（SD=26.0），*diff*-PLSA（n=50）では名詞が平均 440.9 件（SD=137.5），係り受けが平均 58.8 件（SD=15.9）となり，名詞および係り受けの両方で有意水準 1%の違いがみられた．

表 4　*diff*-PLSA で抽出されたトピックの内容の例

トピックZ'08

P(XZ)	n(X)	X：単語	P(YZ)	n(Y)	Y：係り受け
2.2%	112	マスタシリンダ	3.2%	32	基づく-発生
2.0%	73	ブレーキ液圧	2.6%	32	操作量-応ずる
1.6%	72	ブレーキ操作	2.6%	33	ブレーキ液圧-発生
1.6%	117	液圧	2.5%	53	制動力-発生
1.4%	779	ブレーキ	2.3%	49	ブレーキ-備える
1.4%	279	制動力	2.1%	20	備える-ブレーキ
1.3%	111	操作量	2.0%	92	車両-ブレーキ
…	…	…	…	…	…

トピックZ'21

P(XZ)	n(X)	X：単語	P(YZ)	n(Y)	Y：係り受け
3.2%	180	非接触	3.0%	52	電力-受電
2.5%	78	送電コイル	2.7%	28	給電-電力
2.4%	160	受電	2.3%	35	受電-電力
2.4%	86	受電コイル	2.1%	25	電力-給電
2.2%	329	給電装置	1.8%	20	駐車装置-変化
2.0%	73	受電装置	1.7%	34	非接触-受電
1.8%	124	受電部	1.6%	21	供給-給電装置
…	…	…	…	…	…

② *diff*-PLSA でのみ抽出されるトピックがあること

②の考察に関しては，今回の方法では *diff*-PLSA の方がトピック数が多かったということもあるが，通常の PLSA のトピック数を 50 個に設定して計算した解も確認したところ，*diff*-PLSA のみに現れたトピックが複数存在した．その例を表 5 に示す．トピック Z'09（表 5 上）は，「シフトレンジ」や「パーキングレンジ」，「検出」，「停止」，「自動的に行う」といった表現で確率が高く，運転者の誤操作を抑制したり自動停止などの運転アシストに関する技術と解釈できる．トピック Z'29（表 5 下）は，「ナビゲーション」や「情報」，「目的地」，「位置情報」といった表現で確率が高く，位置情報を取得してドライバーにナビ情報として提供するなど，情報の取得と提供に関する技術と解釈できる．どちらも近年の自動車業界において付加価値を高める重要な機能といえる．

表 5　*diff*-PLSA でのみ抽出されたトピックの内容の例

トピックZ'09

P(XZ)	n(X)	X：単語	P(YZ)	n(Y)	Y：係り受け
1.4%	50	シフトレンジ	1.9%	23	自動的-行う
1.0%	38	パーキングレンジ	1.7%	38	操作-行う
0.9%	147	検出結果	1.6%	22	駆動-停止
0.7%	1,200	停止	1.6%	26	動作-行う
0.7%	365	解除	1.6%	40	要する-時間
0.6%	34	キースイッチ	1.4%	46	停止-状態
0.6%	3,960	検出	1.2%	26	ブレーキ-作動
…	…	…	…	…	…

トピックZ'29

P(XZ)	n(X)	X：単語	P(YZ)	n(Y)	Y：係り受け
1.2%	122	ナビゲーション装置	2.1%	48	情報-送信
1.1%	809	情報	2.1%	42	情報-含む
1.0%	165	目的地	1.8%	68	情報-取得
1.0%	111	位置情報	1.5%	31	情報-受信
0.9%	786	取得	1.5%	35	示す-情報
0.9%	819	送信	1.5%	126	情報-基づく
0.8%	558	表示	1.4%	25	情報-用いる
…	…	…	…	…	…

③ *diff*-PLSA のトピックの解釈が難しいこと

③の考察に関しては，*diff*-PLSA で抽出されるトピックは，頻度が少ない表現で構成される傾向にあるが故に，トピックの所属確率が頻度の少ない多様な表現に分散しやすく，通常の PLSA で抽出されるトピックよりも意味の解釈が難しい印象があった．実際に表 1 の通常の PLSA の結果と比較し，表 4,5 の *diff*-PLSA の結果は上位語の所属確率が小さい．通常の PLSA は頻度の多い代表的な表現に所属確率が集中するため，解釈はしやすいが，結果として典型的なトピックとなってしまう．この現象を解消する手法として *diff*-PLSA を開発したため，このジレンマは仕方のないものといえる．逆にそれだけより具体的で細かい要素で構成される個性の強いトピックが抽出されていることを意味している．トピックの構成内容だけでは解釈しにくいものは，4.3 で示したように各文章のトピックのスコアを計算し，各トピックと関連の強い具体的な文章を確認することが手掛かりとなる．

5. まとめ

本稿では，テキストデータから典型的なトピックだけではなく，より個性的なトピックを抽出する手法として，従来の PLSA に対して *diff*-PLSA を提案した．本手法は，実測頻度の共起行列に加え，期待頻度の共起行列を構成し，それぞれの共起ペアにおいて期待頻度に対する実測頻度の比率の対数を取った値を持つ共起行列を構築し，これに PLSA を適用するものである．本稿では *diff*-PLSA の有効性を検討するため，特許の要約文データを例に，PLSA を適用した結果と *diff*-PLSA を適用した結果を比較した．その結果，通常の PLSA では，頻度の多い表現でトピックが構成される傾向にあり，全体を表す代表的なトピックが抽出されやすいが，*diff*-PLSA では，頻度の少ないより具体的で細かい表現もトピックを構成する重要な要素となっており，よりエッジの立ったトピックが抽出されていた．また *diff*-PLSA では通常の PLSA では抽出されないトピックも抽出されていた．

以上より，*diff*-PLSA は個性的なトピックを抽出する方法として適用価値があると思われるが，あくまでも全体を代表する典型的なトピックを理解した上で用いることが望ましく，実際には通常の PLSA と *diff*-PLSA を併用することが有効であると考えられる．

ビジネスにおいてテキストデータを分析し，その膨大な定性情報に潜む特徴を理解して有効なビジネスアクションを検討しようとする際，業務担当者が経験的に獲得するような典型的な特徴を改めて理解することも重要であるが，新たな気づきとなるインサイトを獲得するには，より個性的な特徴を抽出することも求められる．*diff*-PLSA ではそうしたトピックを通常の PLSA よりも効果的に抽出することができ，ビジネスインサイトの獲得に向けて有用な知識を提供することが期待できる．

参考文献

[野守 2014] 野守耕爾，神津友武: 三位一体アプローチによるテキストデータモデリング法の開発―宿泊施設の口コミデータを用いた評価推論モデルの構築―，2014 年度人工知能学会全国大会論文集，2014.

[野守 2018] 野守耕爾: テキストマイニングに複数の人工知能技術を応用した特許文書分析と技術戦略の検討，情報の科学と技術，Vol.68，No.7，pp.332-337，2018.

[Hofman 1999] Hofmann, T.: Probabilistic latent semantic analysis, Proc. of Uncertainty in Artificial Intelligence, pp.289-296, 1999.

[Kameya 2005] Kameya, Y., and Sato, T.: Computation of probabilistic relationship between concepts and their attributes using a statistical analysis of Japanese corpora, Proceedings of Symposium on Large-scale Knowledge Resources, pp.65-68, 2005.

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

看護記録における医療－介護連携に関する記述の基礎調査

Basic Analysis of Descriptions on Medicine-care in Nursing Records

上坂 静耶[*1] Seiya Uesaka
土方 嘉徳[*1] Yoshinori Hijikata
鈴木 斎王[*2] Muneou Suzuki

[*1] 関西学院大学商学部 School of Business Administration, Kwansei Gakuin University
[*2] 宮崎大学医学部附属病院 Faculty of Medicine, University of Miyazaki Hospital

Communication and collaboration between hospitals and nursing homes is required in these years. It is important to record the state of the patient at another facility. Thus, we analyzed the nursing records in University of Miyazaki Hospital to find content on the information of another facility and the state of the patient there. We analyzed nursing records of four nurses and found that there is little records according to the state of the patient at another facility.

1. はじめに

近年，医療と介護をとりまく現状は大きく変動しており，疾病構造の変化，医療ニーズの多様化，医療提供体制の機能分化によって，「病病連携」，「病診連携」，「医介連携」などの動きが広がっている．このような医療連携システムや在宅医療においては，大学病院など高度な機能を有する中核病院，市立病院など地域密着型の一般病院，かかりつけ医など地域診療所，老人ホームやデイサービスなどの介護施設，さらには患者の家庭まで含めた情報共有が重要となる．

電子カルテによって集められた情報は，院内の医療従事者間で共有されるだけでなく，地域の医療従事者との間でも共有されたり，さらに患者の転院や退院後に地域の医院や介護施設において活用されたりすることも期待されている．

しかし，各医療機関における電子カルテには，記載される病院以外の施設や家庭での患者の様子に関する記載など，各種連携にとって有益な情報がどれほど存在するのか分からないのが現状である．そこで中核となる大学病院の看護記録に，どれだけ上記のような各種機関の連携に関する情報の記載があるのかを明らかにする．

2. データ収集

看護記録について，日本看護協会は「看護記録および診療情報の取り扱いに関する指針」にて以下のように述べている．

1. 必要なことは漏れなく記述する
2. 必要でないことは一つも書かない
3. 無防備な看護記録の現実を改める
   - 個人的感情の記載
   - 感想，憶測，個人的見解
4. [大原則]重大医療事故発生時には，記録方式を経時的記録に変える

看護記録の様式には，基礎情報（データベース），看護計画，経過記録，要約（サマリー）等がある．このうち経過記録にはいくつかの記述方法があり，そのひとつが SOAP と呼ばれるものである．これは以下の 4 項目にそって記載していく方式である．

S（subjective）：主観的情報 - 患者の話から得られた情報
O（objective）：客観的情報 - 身体診察・検査から得られる患者の様子や認識の状態
A（assessment）：評価 - 診断，O と S から考えられること
P（plan）：計画（治療）- 治療方針・内容，生活指導

しかし，看護記録は生活歴や検査歴，予約などのメモとして使用されていることもある．テキストデータである経過記録に明確なルールは存在しないため，細かい記述方法は記載者によって異なる．看護師は，患者が述べた言葉を覚えてメモを取り，最後にまとめて電子カルテに入力している[串間 12].

本研究では，宮崎大学医学部附属病院で運用されているクリニカルパスの看護記録を分析した．入手した時点でデータ内の患者氏名および記載者氏名は匿名化されている．2013 年 12 月 21 日～2018 年 3 月 30 日の期間に SOAP に分けずにフリーテキストで書かれた看護記録と，2013 年 10 月 5 日～2018 年 3 月 30 日の期間に SOAP の形式で書かれた看護記録があり，この両方を分析の対象にした．フリーテキストのデータでは，記録数が 274,008 件，記録者総数が 995 名，患者総数が 8,718 名である．SOAP 形式のデータでは，記録数が 142,662 件，記録者総数が 549 名，患者総数が 4,778 名である．各データには，定型項目もあり，患者 ID，患者氏名，患者診療科，患者の性別，生年月日，文書番号，文書名称，タイトル，Subject，Object，Assessment，Plan，フリーテキスト，入院/外来区分，記載日付，記載者氏名，記載者所属部署，修正者氏名，修正者所属部署が記録されている（ただし，個人情報は匿名化されている）．

3. 分析方法

3.1 記録の調査項目

対象データの Subject，Object，Assessment，Plan および SOAP に分けずにフリーテキストで記載された記録において，関連機関との連携に関する以下の情報が書かれているかどうかを人手で調査する．

(1) 病診連携・医介連携に関する情報

他施設との連携に関する記述が存在するかどうかを調査者が判断しタグ付けする．具体的には，転院前に治療していた病院や入院前に入所していた介護施設における患者の様子に関する記述や，退院後の自宅における患者の生活状況に関する記述，患者の家族が発した言葉の記述などが挙げられる．

連絡先：土方 嘉徳，関西学院大学 商学部，
兵庫県 西宮市 上ケ原一番町 1-155 ，0798-54-6205，
contact@soc-research.org

また，上記情報と関連のある情報（看護師の記録の特徴）として，以下のものを人手または計算機により調査する．

(2) 看護師による主観的記述

看護師による主観的な記述が存在するかどうかを調査者が判断しタグ付けする．

(3) 専門用語の数

医療関係の専門用語がどれだけ使用されているかを計算機により調査する．具体的には，看護記録中のテキストに対して，医療用語辞書（ライフサイエンス辞書 [金子 93]）中の単語で検索を行い，適合した単語の数を記録する．

(4) 形容詞の数

主観的記述に用いられることの多い形容詞がどれだけ使用されているかを計算機により調査する．具体的には，看護記録中のテキストに対して，形態素解析を行い形容詞の数を記録する．

(5) 記載の量

看護記録の記載量（テキストサイズ）を調査する．

3.2 看護師の属性

病診連携・医介連携に関する情報を書くかどうかは，看護師のキャリアや経験に依存するかもしれない．そこで，記録した看護師の年齢と勤続年数も調査する．これらの情報は，職員リストのデータ（個人情報は匿名化されている）から得た．

4. 分析結果

本分析は人手によるタグ付けが含まれるため，全データを対象とするには時間が必要である．そのため，まずは内科（5 階西病棟・第 1 内科）で受診しているある患者に対して，看護師 4 名（それぞれ A, B, C, D）がフリーテキスト形式で記述した記録を対象に分析を行った．全員が女性で，それぞれの勤続年数は A が 5 年，B が 3 年，C と D は 2 年であった．

調査結果を表 1 に示す．4 人の看護記録においては，他医療機関における患者の様子についての記述は見られなかった．しかし退院指導に関する記述の中で，次回外来日時およびその病院，紹介状の有無に関する記述が 15 件見られた．各看護師の記述総数に占める，自宅での患者の様子が書かれた割合については，どの看護師も 1%に満たなかった．家族の言葉の記述数は A が最も多い 27 件，B, C, D はそれぞれ 6, 3, 5 件であった．

表 1 第 1 内科の医介連携に関する記述の数

看護師	A	B	C	D
勤続年数	5	3	2	2
記録総数	728	539	508	324
他施設での患者の様子	0	0	0	0
他施設との連携情報	13	2	0	0
自宅での患者の様子	5	1	0	2
自宅での様子記述割合(%)	0.687	0.186	0.000	0.617
家族の言葉	27	6	3	5
家族の言葉記述割合(%)	3.709	1.113	0.591	1.543

5. 考察

まずは他医療機関における患者の様子についての記述が見られなかったことについて考察する．今回の記録は，ある病棟（5 階西病棟）の入院患者に関するものであった．入院にいたるまでの経緯はそれまでのカルテに記載されているはずなので，本人に確認するまでもなかった可能性がある．今後は外来患者についての記録も分析する必要がある．

次に，医介連携に関する情報が記載されていることについて考察する．今回の記載者はほぼすべて大学病院で勤務する看護師である．患者が大学病院を退院し他の施設や自宅で治療・介護を進める際には看護師が退院指導を行う．このときの記録として次回外来日時およびその病院，紹介状の有無に関する情報を記載する．また，看護師は病診連携に関する知識を持ち合わせており，患者が大学病院を退院し他の施設で治療・介護を進める際に活用できることを理解して記述していると考えられる．

次に，患者の家族が発した言葉に関する記述数が目立って多かったことについて考える．看護師 A が記述した 27 件のうち 19 件がインフォームドコンセントについての記録だった．看護師は医師が説明する際に同席することも多く，このときの様子や反応を細かく記録している．インフォームドコンセント以外にも，看護師は訪室の際に面会に来ていた家族とコミュニケーションを図っていることがわかった．患者本人の意思や家族の思いについて詳細を記録していたこともあった．

最後に，各特徴をもつ記述数と看護師の属性について考察する．今回このような記述を行っていた看護師は勤続年数が最も長い看護師 A であった．実施した処置や観察初見の記載に終始することなく家族とのやりとりまで記録していたのは，長年の勤務経験によるものが大きいと考えられる[山崎 15]．

6. まとめ

今回の調査では，医療―介護連携に関する情報は，日付や紹介状の有無など，事務的な内容に限られて記載されていた．また，他医療機関における患者の様子に関する記載は見られなかった．一方，退院後に有益になりそうな家族の発した言葉に関する記述は存在した．今回の調査結果から，大学病院においては，医療－介護連携に関する有用な記述は看護記録に，ほとんど残されていない可能性があることが分かった．しかし，病院での診療行為だけでなく，患者の家族の言葉が記録として残されていたことについては，医療－介護連携を進める上で，助けとなるかもしれない．今後は，主観的記述に関連する項目にも注目し，データ数を増やして分析していく予定である．

謝辞

本研究は JSPS 科研費 17KT0084 の助成を受けたものです．

参考文献

[串間 12] 串間宗夫ほか: 電子カルテテキストデータに関する一考察．人工知能学会第 1 回インタラクティブ情報アクセスと可視化マイニング研究会 (2012)

[山崎 15] 山崎友義，串間宗夫，荒木賢二，鈴木斎王: 論文 TETDM を用いた電子カルテテキストデータ分析．第 29 回人工知能学会全国大会 (2015)

[金子 93] 金子周司ほか: ライフサイエンス辞書プロジェクト．https://lsd-project.jp/ (1993)

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

深層学習における学習ネットワークからの分類パターンの抽出

Extraction of Classification Patterns from Deep Learning Networks

安藤 雅行 *1 Masayuki Ando
河原 吉伸 *2*4 Yoshinobu Kawahara
砂山 渡 *3 Wataru Sunayama
畑中 裕司 *3 Yuji Hatanaka

*1滋賀県立大学大学院工学研究科
Graduate School of Engineering, The University of Shiga Prefecture

*2九州大学 マス・フォア・インダストリ研究所
Institute of Mathematics for Industry, Kyushu University

*3滋賀県立大学工学部
School of Engineering, The University of Shiga Prefecture

*4理化学研究所 革新知能統合研究センター
RIKEN Center for Advanced Intelligence Project

In deep learning, there is a problem that concrete classification patterns for deriving reasons for classification are often incomprehensible. In this paper, we propose a classification patterns extraction system from deep learning networks and verified the effectiveness of the system. The proposed system takes out learning networks from the learning result of deep learning and extracts classification patterns from the learning networks. Then the system displays the extracted classification patterns so that users of the system can interpret the learning networks. In verification experiments, the significance of the extracted classification patterns was estimated by chi-square test. The results showed that users of the system can extract classification patterns effective for interpretations of the learning networks by using the proposed system.

1. はじめに

インターネットの普及などによって世の中で増大する文章データへの対処方法として，深層学習を用いたテキストマイニングシステムが注目されている [1][2]．深層学習は，一般に多層から構成されるニューラルネットワークを用いた学習を指す．

その一方で，深層学習は，その出力を導いた根拠についての解釈が困難であることも知られている．テキスト分野においても，学習ネットワークの解釈を行うことで深層学習の分類基準をより深く理解できれば，例えば医療分野において新人とベテランの書いた電子カルテの違いから，良い電子カルテを書く方法を容易に理解したり，企業においても良い報告書や企画書を書く方法を短時間で習得できるなど，深層学習の新しい活用が期待できる．

深層学習の分類基準を示す研究としては，アテンションと呼ばれる手法を用いた研究 [3][4] が注目されている．しかし，アテンションは内部でどのような学習が行われているかは考慮していない．一方で画像に関してだが，深層学習の学習ネットワークの解釈に注目したものがある [5]．また，ニューラルネットワークを用いた研究では，中間層の役割と重みの持つ意味について考察したもの [6] や中間層が学習から，入出力の様々な関係を得ることを目指したもの [7] があり，中間層に注目すれば，出力に関して意味のある情報が得られることを示している．また，重みに注目した研究として，重みを用いて入力データ間の相関関係を抽出する研究 [8] があるが，分類基準を提示する目的で重みに注目したものは存在しない．

そこで本研究では，RNN（Recurrent Neural Network）を使用し，テキスト集合の学習によってネットワークの層に付けられた重みの値を取り出し，学習ネットワークに対する解釈を行うために，それぞれの出力特有の特徴となる分類パターンの抽出を行うシステムを提案する．

連絡先: 安藤 雅行，滋賀県立大学大学院工学研究科先端工学専攻，〒 522-8533 滋賀県彦根市八坂町 2500，oh23mandou@ec.usp.ac.jp

2. 深層学習の重みを用いたテキストの分類パターン抽出システム

本章では，本研究で開発した深層学習の重みを用いたテキストの分類パターン抽出システムについて，システムの構成とその詳細について述べる．

2.1 分類パターン抽出システムの構成

分類パターン抽出システムでは，まず，図 1 に示すように，各分類先ごとにラベル付けしたテキスト集合を RNN にて分類し，その分類先を導いた学習ネットワークと，学習ネットワーク上の重みから，提案システムの分類パターンの抽出・可視化処理によって各出力 (分類先) を導くネットワーク上のパスと，パス上の各ノードの情報（そのノードで学習された単語）の決定，表示を行う．最後に，システムの利用者は，システムによって得られた学習ネットワークの表示を自分が見やすいように調整し．分類パターンを抽出する．そして分類パターンの意味を理解しやすくするための機能を利用できる．

2.2 深層学習による学習ネットワークの形成

2.2.1 文中の単語のベクトル化

深層学習で学習を行う前に，テキストデータは文中の単語を抽出したあと，単語は One hot 法 [9] と呼ばれる手法に従い単語ベクトルの羅列に直される．そして，文中の各単語を単語ベクトルに置き換え，深層学習への入力データとする．なお，抽出される単語は名詞する．これは，文章の特徴とその順序関係をより学習・抽出しやすくするためである．

2.2.2 学習によるネットワークの重み付け

One hot 法によって単語ベクトルの羅列に変換され，さらにラベル付けされたテキストデータは，RNN の学習にて，そ

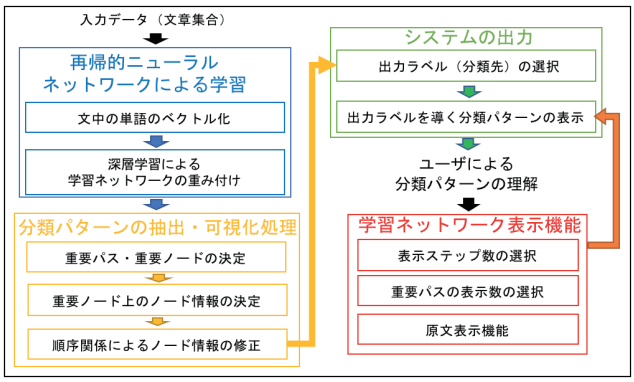


図 1: 分類パターン抽出システムの構成

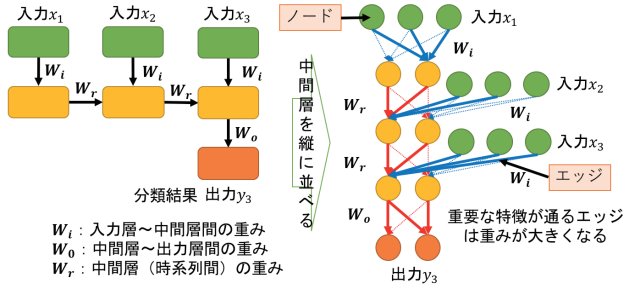


図 2: RNN の学習ネットワークと学習の様子

れぞれの出力ラベル（分類先）を導くネットワークへの重み付けがされていく．その様子を図 2 に示す．入力文章は各単語がベクトル化され，タイムステップごとに単語ベクトルが順番に入力されていく．また，RNN での分類時は，最後の単語が入力されたタイミングで，出力層から出力される．

2.3 学習ネットワークからの分類パターンの抽出・可視化処理

本提案システムの分類パターンの抽出･可視化処理では，RNN によって得られた学習ネットワークから，各出力を導く最も関係が強いパス（ネットワーク上のエッジの繋がりの線）を決定する処理を行う．この出力と繋がりが強いパス（重要パスと呼ぶ）の決定について，図??に示した 4 層全結合型ネットワークモデルを例として，具体的な手順を述べる．

まず，ある分類先（図??の例では層 D のノード b）に到達するパスについて，パス上のエッジについた重みの積で定義される，重要度と呼ばれる値を算出する．図??の入力層のノード i からの太矢印のパス上のエッジに付いた重みを $W1_{a,i}$，$W2_{i,j}$，$W3_{j,b}$ とすると，このパスの重要度 $B_{a,i,j,b}$ は以下の式 (1) で導かれる．

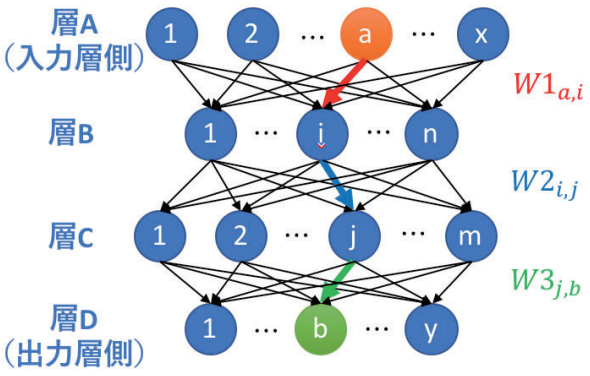


図 3: 4 層全結合型ネットワーク

$$B_{a,i,j,b} = W1_{a,i} \times W2_{i,j} \times W3_{j,b} \tag{1}$$

そして，式 1 で出力 b に到達する全てのパスの重要度を計算して比較し，最も値の大きいパスを，重要パスと決定する．出力 b の重要パスの重要度を MB_b として，その計算式 (2) を以下に示す．図 2 の右図のように，出力層と中間層間の重み $\boldsymbol{W_o}$ のあと，中間層間の重み $\boldsymbol{W_r}$ を繰り返し辿ることで，過去の中間層を遡ることになる．

$$MB_s = \max_{i,j,b} B_{a,i,j,b} \tag{2}$$

次に，重要パス上のノードについて，ノードの情報を決定の処理を行う．ノード情報の決定方法は，出力ごとの重要パスを決定した時と同様に，入力層からノード情報を決定したい中間層ノード間の重み $\boldsymbol{W_i}$ の大きさから，ノード間の結びつきの強さを求める．ただし，図 2 に示す様に，各中間層にはそれぞれ個別の入力層が対応している．ここで，RNN の入力には One hot 法による単語ベクトルを用いているため，入力層ノードにはノードごとに 1 種類の単語が対応していることになる．そこで，重要パス上の入力層ノードの単語を参照し，その単語をノードの情報と決める．最後に，ノード情報の単語について，前後の中間層ノード情報の単語と，原文中でその順番で表示されているかどうかを確認し，されていないなら単語の重要度を下げる．

こうして各出力を導く重要パスとパス上のノード情報を表示した学習ネットワークから，過去の中間層の重要ノードから現在の中間層の重要ノードの単語より，時系列を考慮した特徴の並びとしての分類パターンを抽出することができる．

2.3.1 出力ラベルを導く分類パターンの表示

本研究で開発した分類パターンの抽出システムでは，分類先に強く結びつく，重要パスの集合としての学習ネットワーク上に重要ノードの情報が表示される．例として，5 種類のお菓子の作り方に関するテキストの分類を行った場合の，システムのメイン画面を図 4 に示す．表示分類先は「マカロン」とする．このネットワークは，RNN が学習した情報を，過去から順番に中間層上に表示したものである．過去から出力層直前までの中間層の重要ノード情報（単語）の並びは，分類パターンを示し，重要パス 1 本につき 1 つの分類パターンが表示されていることになる

図 4 では，入力層を除いて選択したステップ数（図 4 では 2）だけ中間層と，出力層が表示され，出力層には分類された分類先を示す出力ノードが表示されている．また，分類先を示す出力ノードには，そこから選択したステップ数の長さで，選択した重要パス数だけ重要パスが表示されている．そして，重要パス上の各中間層ノードには，ノード情報として単語が表示されている．

2.4 システム上での学習ネットワーク表示機能

システムには，利用者が分類パターンの抽出・理解を行いやすいように，その表示内容を変更できる機能がある．その主なものを表 2 に示す．

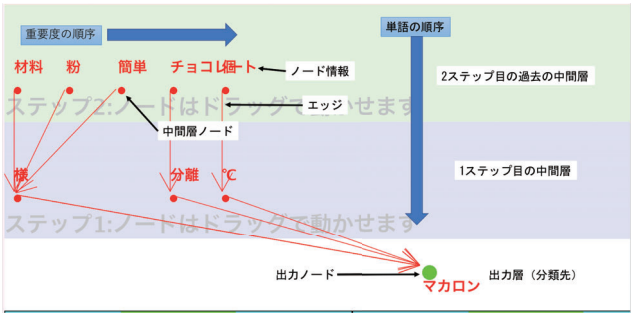


図 4: システムの画面

表 1: 学習ネットワーク表示機能

機能名	効果
表示するステップ数の選択	RNN において，何ステップ過去までの中間層を表示するか決定する
重要パスの表示数の選択	分類先ごとに，何本の重要パスを表示するか決定する
原文表示機能	分類パターン（順序を考慮した単語の組み合わせ）が，原文中でどのように出現しているかを表示する

3. 分類パターン抽出システムの有効性の検証実験

本章では，深層学習の重みを用いた分類パターンの抽出システムにより，学習ネットワークから抽出できる分類パターンが，テキストデータを理解する（学習ネットワークの解釈を行う）ために有効かを検証した実験について述べる．なお，本実験では，分類パターンを「2 種類の異なる単語の，順序を考慮した組み合わせ」と定義する．単語を 2 種類にした理由は，最も基本的な分類パターンであるためである．

3.1 実験準備

3.1.1 使用テキストデータと学習モデル

分類パターン抽出の対象とするテキスト「童話 5 種類」について，詳細を表 2 に示す．また，計算上の分類パターンの総数（抽出単語数 × 抽出単語数 − 抽出単語数）とその中で「ある分類パターンが，ある分類先に特有のものである」という仮説を有意水準 5%で検定したカイ二乗検定で，有意性があると推定された分類パターン数を表 3 に示す．

続いて，深層学習として使用した RNN モデルの概要を表 4 に示す．システム上での表示設定は，学習ネットワーク表示機能により，表示ステップ数 2，重要パス数 20 とする．また，テキスト「童話 5 種類」に対する RNN の分類精度は 100%であった．

3.2 実験手順

実験の対象者は著者 1 名とした．対象者はテキスト「童話 5 種類」について，本研究で開発したシステムを用いて，全ての分類先で分類パターン抽出システムによって抽出された分類パターン 20 個について，カイ二乗検定で有意性があるかどうか，ある場合，カイ二乗検定の順位で何位にあたるのかも調べ考察を行った．なお，カイ二乗検定の順位とは，カイ二乗検定

*1 Google：https://www.google.co.jp で「童話名　あらすじ」と検索し，表示された上位 50 位のサイトの本文をテキストとして利用

表 2: テキストデータの詳細

データ名	内容
童話 5 種類	日本の童話「かぐや姫」「鶴の恩返し」「さるかに合戦」「桃太郎」「浦島太郎」の，それぞれの概要やあらすじなどについて書かれたテキストをネット上 *1 から 1 種類あたり 50 テキストずつ用意した．

表 3: 有意性のある分類パターンの総数

分類先名	カイ二乗検定での分類パターン総数	分類パターン総数
かぐや姫	93,211	15,527,540
鶴の恩返し	9,208	15,527,540
さるかに合戦	24,839	15,527,540
桃太郎	46,749	15,527,540
浦島太郎	53,334	15,527,540
平均	45,468	15,527,540

で有意性のあると推定された分類パターンの集合を，カイ二乗値が高い（より強く有意性があるとされる）順に並べた時の順位である．

3.3 結果と考察

抽出された分類パターンで，有意性がないとされた分類パターンについての考察を行うため，有意性がある・ないとされた分類パターンの個数を図 5 に示す．

図 5 より，まず有意性があるとされた分類パターンの個数と，有意性がないとされた分類パターンの個数の比率は，4 対 6 ほどであることがわかる．その中で，有意性があるとされた分類パターンについて見ると，どちらのテキストも，ほとんどがカイ二乗検定での分類パターンの上位 50%以上に含まれることがわかる．これは，カイ二乗値の高い，より有意性があるとされた分類パターンと，有意性が全くないとされた分類パターンが混ざって抽出されていると言える．

その中で，有意性がないとされた分類パターンが 60%以上と最も多い「浦島太郎」を例に考察する．「浦島太郎」の抽出された分類パターン上位 10 個を表 5 に示す．有意性がないと判断される要因は，他のテキストでも頻度が高い，もしくはそのテキスト内での頻度が低いことが原因である．しかし，それでもシステム上で重要度の高い分類パターンとして抽出されたということは，有意性があるとされた分類パターンと合わさることで，よりテキストへの理解を深めることができる補助的

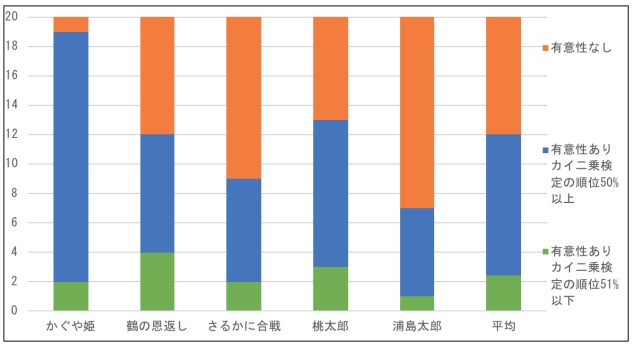


図 5: 抽出された分類パターン

表 4: 学習モデルの詳細

データ名	抽出した単語集合	入力層ノード数	中間層ノード数	出力層ノード数
童話 5 種類	名詞 3,941 種の単語	3,941	50	5

表 5: 抽出された分類パターン（「浦島太郎」）

	抽出された分類パターン			
順位	分類パターン	「浦島太郎」内の頻度	「浦島太郎」以外での頻度	カイ二乗検定での順位
1 位	世界→気	12	4	34,300
2 位	日→展開	6	3	0
3 位	風土記→日本	12	0	2,118
4 位	物語→日本	18	11	0
5 位	日→日本	16	10	0
6 位	長期間→ストーリ	3	0	0
7 位	集→嶋	3	0	0
8 位	漁師→共通	3	0	0
9 位	年→日本	19	11	0
10 位	亀→女性	16	0	964

な役割を持つと考えれる．なお，表 5 で有意性がない分類パターンのカイ二乗検定での順位は 0 と表記している．

表 5 をみると，まず単語「日本」が，4 つの分類パターンで共通しており，そのうち 3 つの分類パターンに有意性がなく，1 つが有意性があるとわかる．また，表 3 より，有意性のある分類パターンのカイ二乗検定での順位は，53,334 中 2,118，つまりカイ二乗検定での順位の上位 4%以内と高い順位であり，ここから，強い有意性を持つ 1 つの分類パターンを基準に，3 つの有意性のない分類パターンがテキストへの理解の補助を行なっていると仮定する．例として，順位 3 位の「風土記→日本」（有意性あり）から原文表示機能を用いて得られるテキストへの理解「丹後国の**風土記**など，**日本**各地で逸文されている」に，順位 4 位の「物語→日本」から得られる「浦島太郎**物語**は，**日本**書紀にも記述が見られる」と順位 9「年→日本」（有意性なし）から得られる「明治 29 **年**に巌谷小波氏が書いた物語をきっかけに**日本**中に広まった」を合わせることで，「浦島太郎物語は丹後国の風土記や日本書紀など，日本各地で逸文されており，明治 29 **年**に書かかれた物語をきっかけに，**日本**中に広まった」という理解が得られる．そして，この理解は「浦島太郎」特有のものである．よって，システムでは，有意性の高い分類パターンと，有意性がないが，それらを補助する役目を持つ分類パターンが得られていると言える．

以上より，実験結果から，本研究で開発したシステムでは，深層学習の学習ネットワークから，学習ネットワークの解釈に役立つ，分類先に対して特有の特徴を表す，有効性のある分類パターンの抽出が可能であると言える．そして，カイ二乗検定での分類パターンと比較することで，カイ二乗検定では有意性がないとされた分類パターンでも，有意性のある分類パターンの補助として役立っていることが考察によって判明し，カイ二乗検定では得られない隠れた有効性のある分類パターンを抽出することができると結論つけた．

4. おわりに

本研究では，複数のテキストデータの分類を深層学習である RNN で行い，学習ネットワークの解釈を行うための，分類パターンの抽出システムの開発を目的とした．本研究の特徴として，重みを辿ることで，RNN の学習ネットワーク内の情報の伝達を過去に向かって探索している点が挙げられる．システムの有用性を確かめる検証実験では，があると推定された分類パターンと，本システムで抽出した分類パターンをのカイ二乗検定で有意性を検定することで，本研究で開発したシステムにより分類パターンの抽出が容易になり，テキストの特有の特徴などを理解しやすくなったと結論づけた．今後の研究では，理解した内容からより容易にテキスト自体への解釈を行えるよう，分類パターンの抽出・可視化アルゴリズムや，システム上での表示を検討していきたい．

参考文献

[1] ボレガラ ダヌシカ，“自然言語処理のための深層学習”，人工知能学会誌，Vol.29, No.2, pp.195–201, 2014

[2] Ebru Arisoy, Tare N. Sainath, Brian Kingsbury, Bhuvaba Ramabhadran, “Deep Neural Network Language Models”, In Proceedings of the NAACLHLT Workshop, Will We Ever Really Replace the N-gram Model?, pp.20-28, 2012

[3] M Daniluk, T Rocktaschel, J Welbl, S Riedel, “Frustratingly Short Attention Spans in Neural Language”, ICLR, 2017

[4] A.Vaswani, N.Shazeer, N.Parmar, J.Uszkoreit, L.Jones, A.N. Gomez, L. Kaiser, and I. Polosukhin, “Attention Is All You Need”, CoRR, vol. abs/1706.03762, 2017

[5] 西銘 大喜，“ディープニューラルネットワークによる画像からの表情表現の学習”，人工知能学会全国大会論文集 3L4-3，pp.1-4, 2015

[6] 立石 雅彦，山崎 晴明，“手書き数字認識における改装ニューラルネットワークの中間層に関する考察”，情報処理学会論文誌，Vol.30，No.10，pp.1281-1288, 1989

[7] 松倉健太郎，村田 昇，“ニューラルネットワークの中間層における独立な特徴量の抽出”，電子情報通信学会技術研究報告，Vol.106，No.102，pp.63-67, 2006

[8] Michael Tsang, Dehua Cheng, Yan Liu, “DETECTING STATISTICAL INTERACTIONS FROM NEURAL NETWORK WEIGHTS”, ICLR, 2018

[9] M. Feurer, A. Klein, K. Eggensperger, J. Springenberg, M. Blum, and F. Hutter. “Efficient and robust automated machine learning”, In Neural Information Processing Systems (NIPS), 2015

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

深層学習を用いた対話文への「かわいさ」の付与

Attachment of “Cuteness” to Dialogue Sentences Using Deep Learning

後藤 優介 *1 Yusuke Goto　砂山 渡 *2 Wataru Sunayama　畑中 裕司 *2 Yuji Hatanaka　小郷原 一智 *2 Kazunori Ogohara

*1滋賀県立大学大学院工学研究科 Graduate School of Engineering, The University of Shiga Prefecture

*2滋賀県立大学工学部 School of Engineering, The University of Shiga Prefecture

Nowadays, automatic generation of dialogue sentences using deep learning has been conducted. Automatic generation systems of dialogue sentences require not only sequences of conversation but also characteristics of conversation partner. Because, it makes users feel charms and personalities in conversation. However, previous studies often transformed the sentence's characteristics to another one which easy to define transformed vocabularies, for example, the dialect and the style of samurai. Therefore, this study will transform the sentence's characteristics to another one which hard to define transformed vocabularies. Especially, the purpose of this study is that give the cuteness to sentences. We considered how to give the cuteness to sentences, and evaluate sentences that can give the cuteness or not.

1. はじめに

近年，Apple 社の Siri や Amazon 社の Alexa を皮切りに，対話型システムの注目度が増している．これらのように，特定のタスクを実行する事を目的とした対話型システムの事をタスク指向型対話システムと呼ぶ．一方で，SoftBank 社の Pepper や Microsoft 社のりんなに代表される，ユーザとコミュニケーションを取る事を目的とした対話型システムの事を非タスク指向型対話システムと呼ぶ．これは主に，コミュニケーションロボットでの活躍が期待され，AI と対話できる事で，より AI に対して親しみを持ってもらう事が期待できる [1]．こういった非タスク指向型対話システム固有の問題として，無個性な対話による対話の質の低下が挙げられる．対話の応答文が無個性で機械的なものであると，ユーザに AI との間に壁を感じさせる．それにより，ユーザの対話に対する意欲が失われる事や，AI に対して親しみを感じる事が難しくなる事が懸念される．

そこで本研究では，付与するキャラクタ性として，具体的なキャラクタではなく，人間の感性の 1 つである「かわいさ」を挙げた．そして，コーパスの調査に加え，深層学習によって学習されたモデルの学習内容を分析する事で，かわいさが現れている文にどのような特徴があるかを網羅的に調査した．その後，調査で得られた要素を自動で付与する事で，文にかわいさを付与するアルゴリズムを作成した．また，そのアルゴリズムに，深層学習による判定を取り入れる事で，ユーザによってどれくらい強くかわいさを付与するかを調整できるようにした．これにより，従来手法と比べて，ユーザの好みに沿った強さのかわいさを，文中の幅広い範囲で付与できる事が期待される．

2. 関連研究

2.1 文に現れるキャラクタ性とそれが現れる場所を調べた研究

文のどの部分にキャラクタ性を感じさせる要素を埋め込む自由度があるかを調べた研究 [2] がある．この研究では，内容表現 (名詞，動詞，形容詞，副詞など) と機能表現 (助詞，助動詞，接続詞など) の両方にキャラクタ性を埋め込む自由度があると結論付けている．しかし，内容表現は非常に自由度が高く，単純な置換は難しいとしている．

一方で，機能表現は同義表現への置き換えが比較的簡単であり，キャラクタ性を容易に埋め込む事ができるとしている．本研究でも，単語の同義表現への置換を扱う事になるので，比較的置換が容易とされていた機能表現は優先的に取り入れた．同時に，内容表現内でも比較的自由度が低いとされた形容表現にも着目した．

2.2 文にキャラクタ付けを行う研究

文節機能部に着目し，人手で書き換えたコーパスを分析する事で見つけ出した置換ルールを，確率に乗っ取り自動適用する研究 [3] がある．これにより，変換により感じる違和感を 10% 以下に抑えつつ，人間が 80%の確率で判別できるキャラクタ性変換を行う事に成功している．また，よりコストを抑えた変換を行うため，変分オートエンコーダと注意機構を用いてキャラクタ性を付与する研究 [4] がある．この研究では，注意機構を用いる事で，生成モデルを用いる場合でも変換する単語を制御しつつ，文全体の変換を行う事ができるという事を示している．

しかし，これらの研究における変換は，機能表現のみを対象とした置換に限られており，キャラクタ性を感じさせる要素を網羅的に適用できているとは言い難い．また，これらの研究では「関西弁」「侍風」といった具体的なキャラクタ性の付与に留まっており，キャラクタ性の付与が直接魅力的な対話につながっているとは言えない．そこで，本研究では，より人間の感性に近いキャラクタ性である「かわいさ」を付与する場合について調査を行った．調査するコーパスとして人手による変換を用いていない既存のものを選び，要素を網羅的に調査する事で，自動適用できる変換ルールを調査した．

3. 「かわいさ」を付与する変換ルール

3.1 「かわいさ」の定義

本節では，本研究で付与するキャラクタ性である「かわいさ」について定義する．「かわいい」という単語は，辞書的には「愛らしい，小さい」という意味がある．しかし，現代日本において，「かわいい」という単語は様々な意味で用いられるようになっており，物理的・数値的特徴からは明確に分類する

連絡先: 後藤優介，滋賀県立大学大学院工学研究科電子システム工学専攻，〒 522-8533 滋賀県彦根市八坂町 2500

事ができない [7]．その人個人の感性によって，どういったものが「かわいい」のかは変わってくる．

「かわいい」の意味について調べた研究 [5][6][7] によると，心理学的に人が他者に対して「かわいい」と感じる要素は，大きく「共感性」「不完全性」「親和性」の 3 つの要素に集約できると言われている．それぞれの用語の定義は以下の通りである．

- 「共感性」: 他者の状況や感情に共感し，他者志向の温かい気持ちを持たせる性質．
- 「不完全性」: 幼さ・弱さから，守ってあげたいという気持ち (保護欲) を感じさせる性質．
- 「親和性」: 相手と親密な関係を維持し，仲良くなりたいと感じさせる性質．

このうち，共感性と親和性については，自分と他者の関係性を評価して生まれる性質であり，相対的なものであるため，統一した基準での定義が難しい．そのため，本研究では不完全性を取り上げ,「文全体から読み手に不完全性を感じさせる性質」を「かわいさ」と定義する．また，外見的な要素を含まない不完全性の要素として,「幼さ (未熟さ)」と「精神的な弱さ (気の弱さ)」の 2 つを挙げ，それにつながる要素を調査していく．

本研究では,「かわいさ」を含む学習用のテキストデータとして，ソーシャルゲーム「アイドルマスターシンデレラガールズ」の登場人物 183 人のうち,「正統派かわいい」キャラクタとして位置づけられている 67 名のセリフを合計 50,113 件 [12] を用いることとした．

また,「かわいさ」を含まない学習用のテキストデータとして，老若男女 199 人の日常会話を録音し，それを文字に起こしたデータセット「名大会話コーパス」[9][10][11] に収録された発話文のうち，20 文字以上のものを 55,054 件を用いることとした．

3.2 「かわいさ」を付与する変換ルール

「かわいさ」を付与する変換ルールの作成に際して，前節で用意したデータをもとに，深層学習により「かわいさ」を含む文と含まない文とに分類する学習を行った．その結果を考察することで，変換ルールを用意した．

3.2.1 「幼さ, 未成熟さ」についての変換ルール

幼さや未成熟さを感じさせる表現として，以下の要素がある事がわかった．

- 幼さや柔らかさを感じさせる表記や表現を使う．
- 相手との距離を縮め，相手を頼ろうとする表現を使う．
- 感情を表に出し，裏表を感じさせない表現を使う．
- 敬語の使用は最低限にする．

これらの要素をもとに，分類に寄与する単語を調べた上で，表 1 に示す「幼さ，未成熟さ」を付与する変換ルールを作成した．

3.2.2 「精神的な気の弱さ」についての変換ルール

精神的な気の弱さを感じさせる表現として，以下の要素がある事がわかった．

- 自分の意見を強く主張できず，断定出来ない事を感じさせる表現を使う．
- 自分の考えを相手に強制できない事を感じさせる表現を使う．
- 相手の意見に対して，強く否定できない事を感じさせる表現を使う．

表 1: 「幼さ，未成熟さ」を付与する変換ルール

変換ルール	変換対象の単語	変換先の単語
名詞をひらがなに置換	漢字のみの名詞 (10%の確率で適用)	その単語をひらがなに置換
副詞をオノマトペに置換	体言か名詞の前にある一般副詞 (30 %の確率で適用)	評価の高い副詞 27 件のうち 1 つと置換
相槌・同意・終助詞長音化	相槌や同意を示す表現 (うん，なるほど等) または終助詞	対象の単語の後に「ー」を挿入
語尾表現置換	です，ます等の丁寧な文末表現とその前の単語 文末記号の前の，名詞や副詞では無い単語	よ，ねに置換もしくは挿入 必要に応じて前の単語も活用させる
感動詞の挿入・置換	感動詞 文頭・文末 (50%の確率で適用)	評価の高い感動詞 (5 件) のうち ランダムで 1 つと置換もしくは挿入
名詞にさんを付ける	一般名詞もしくは固有名詞 (15 %の確率で適用)	対象の単語の後に「さん」を挿入

表 2: 「精神的な気の弱さ」を付与する変換ルール

変換ルール	変換対象の単語	変換後の単語
断定表現を婉曲表現に置換	言い切る表現 (〜ない等)	対象の単語の後に「かも」を挿入
曖昧な表現を追加	文末の動詞 (50%の確率で適用)	対象の単語の後に「かも」を挿入
命令をお願いする表現に置換	命令形である動詞	対象の単語を基本形に置換し， 「〜てほしいな」「〜よ」を挿入
逆接の接続詞を削除	逆接の接続詞 (しかし，だけど等)	対象の単語を削除，必要に応じて句読点を挿入
強い否定を緩和する表現の追加	強い否定を表す副詞や形容動詞 (ちっとも，全く等)	「多分」「できれば」と置換

これらの要素をもとに，分類に寄与する単語を調べた上で，表 2 に示す「精神的な気の弱さ」を付与する変換ルールを作成した．

4. 深層学習を用いて「かわいさ」を付与するシステム

本章では，前節で用意した変換ルールをもとに，対話文に「かわいさ」を付与するシステムについて述べる.「かわいさ」を付与するシステムのフローチャートを図 1 に示す．以下で，フローチャート内の「かわいさ」の評価システムについて述べる．

4.1 深層学習による「かわいさ」の評価

本システムにおける「かわいさ」の評価（以降スコアと呼ぶ）は，対話文を 3.2 節で構築した深層学習に入力して分類した時の，SoftMax 関数により出力される値 E（0 から 1 の実数値）を 100 点満点に換算した値に，変換ルールが適用された回数 n の 5 倍の数値を加算した式 1 の値を用いて行われる．ただし，100 点を超えた時には 100 点として扱う．

$$S = 100E + 5n \quad (0 \leqq S \leqq 100) \tag{1}$$

深層学習においては，この値が 50 点を超えると「かわいい」文として判定されるが，本研究では文の分類が目的ではなく，文にかわいさを付与することを目的とするため，目標スコアを与え，このスコアを超えるまで前章で用意した変換ルールの適用を試みる．

4.2 「かわいさ」を付与する変換ルールの適用

表 1 と表 2 に示した 11 個の変換ルールのうち，まだチェックを行っていないルールについて，変換対象となる単語があるかをチェックする．すなわち，文中の各単語について，変換対象の単語かどうかをチェックし，変換対象の単語が存在する場合，その変換ルールの適用処理を行う．変換後に目標となるかわいさスコアを超えていた場合には変換ルールの適用を終了し，目標スコアを超えていない場合は，次のルールを適用する．なお，チェックの順番は，名詞をひらがなに置換，感動詞の挿入，曖昧な表現を追加以外の 8 個のルールについて乱数を用いてランダムに決定する．その 8 つのルールについての

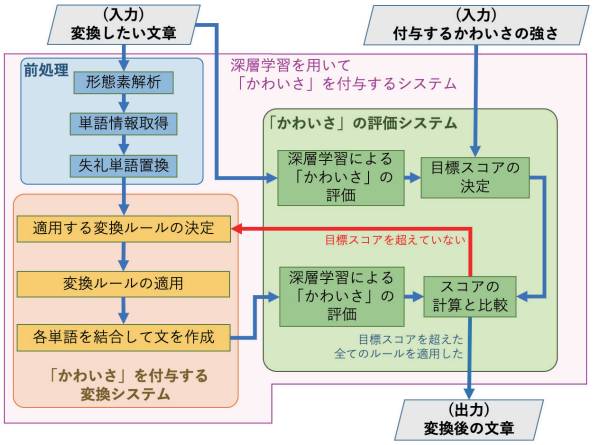


図 1: 「かわいさ」を付与するシステムのフローチャート

表 3: 変換ルールの適用で「かわいい」と答えた人数の増減

	文数 (/944)
「かわいい」と答えた人が増えた文	474
「かわいい」と答えた人が減った文	111

チェックが終了した後，曖昧な表現を追加→名詞をひらがなに置換→感動詞の挿入の順にルールを適用する．

5. 「かわいさ」を付与するシステムの検証実験

5.1 実験設定

大学生・大学院生 16 名に 3. 章で用意したデータの一部（「かわいさ」を含まない文 1000 件を学習データとは別に用意して，全く変換が行われない文を除いた 944 件）を用いて，無作為に分けた変換ルールの適用前と適用後のいずれかの文について「各文がそれぞれ自分に対する発話であった場合，その文を発した人に対してかわいいと感じるかどうか」を回答してもらった．

5.2 実験結果と考察

5.2.1 変換ルールの適用前後で「かわいい」と答えた人数の比較

変換ルールの適用により「かわいい」と答えた人数の増減表 3 に示す．約 50%の文で「かわいい」と回答する人の数が増えており，一定の効果があったことが確認できる．

変換後の文のうち，特にかわいいと判定した人数が増加した文の例を表 4 に示す．これらの文は，変換前の文はかわいいと答えた人数は 0 人だが，変換後は 6 人もしくは 7 人まで増加していた．一方で，変換後にかわいいと判定した人数が減少した文の例を表 5 に示す．これらの文は，変換前の文はかわいいと答えた人数は 3 人または 4 人であるが，変換後は 0 人であった．

変換によってかわいくならなかった 470 件の文について，その原因を著者が分類した結果を表 6 に示す．

もとの文がネガティブな内容の話，説教や教訓を含む話，若い人があまりしなさそうな内容を含む話という 3 つの要因で，かわいくならなかった文全体の約 42%を占めている．そのため，文の内容によっては，どのような変換を行ってもかわいさを付与できない可能性が考えられる．

表 4: 変換後の文のうち，特にかわいいと判定した人数が増加した文の例

変換前の文	変換後の文
結構，周りがワイワイガヤガヤしてても録れるんです．	ふふふ，結構，周りがワイワイガヤガヤしてても録れるね．
来年までいる気なわけだから．だから，ぎりぎりでしょう？	うふふ，来年までいる気なわけだからね．だから，ぎりぎりよね？
何とかっていうバンドのボーカルだったんだよ．	何とかっていうバンドのボーカルさんだったんだよー．

表 5: 変換後の文のうち，特にかわいいと判定した人数が減少した文の例

変換前の文	変換後の文
私は自分のを取ったら，手をひざの上に乗せて待ってます．	わたしは自分のを取ったら，手をひざの上に乗せて待ってよ．
それもほんとに昔行っただけだもん，高校ぐらいのときに．	それもつやつやしたむかし行っただけだもん，高校ぐらいのときにね．
いやでも，ヒジキとか，ちゃんと作ってあげようと思った．	いやでも，ヒジキとか，ぼかぼか作ってあげようと思った．

5.2.2 変換ルールごとのかわいいと判定した人数の変化

各変換ルールの適用が，読み手に与える「かわいい」という印象にどのように寄与しているかを調べるため，3.2 節で述べた各変換ルールごとに，かわいいと判定した人数の変化をまとめた．各変換ルールごとの「かわいい」と答えた人数の変化を表 7 に示す．全ての変換ルールについて，変換ルール適用後の文のほうがかわいいと判定した人数は増加していた．

特にかわいいと答えた人が増加した文が多いルールとして，感動詞の追加や置換をするルールと，命令をお願いする表現に置換するルールがある．これらのルールは，適用した文の約 6 割の文でかわいいと答えた人数が増加しており，比較的有用なルールであるといえる．かわいいと答えた人数が大きく増加した理由として，これらのルールが文のイメージを大きく変える効果を持っているという事が考えられる．感動詞を追加したり置換するルールの場合，本システムで付与する感動詞はにこやかな笑いや微笑みを感じさせるものであり，文全体に柔らかな印象を与えることができる．また，命令をお願いする表現に置換するルールの場合，命令口調というわかりやすくきつい印象を与える表現を無くし，お願いするという柔らかい表現を付与することで，きつい印象の文を一気に柔らかい印象の文に変えることができる．これらのように，文の印象に明確な変化を与えることができるため，柔らかさや気の弱さといった面から「かわいさ」につながったと考えられる．

一方で，否定表現を緩和するというルールと，断定表現を婉曲表現に変換するルールについては，ルールを適用後にかわいいと答えた人数が増加した文は約 34%～37%であり，約 45%の文ではかわいいと答えた人数が変わらなかった．これは，強い否定表現を弱い否定表現に置換しても結局否定している事には変わらず，文全体が読み手に与える印象としてはあまり変化がなかったからだと考えられる．同様に，断定する表現を遠回しな婉曲表現に置き換えても，結局内容を肯定している事には変わらないため，文の印象に変化が無かったといえる．しかし，これらのルールでもかわいいと答えた人が増加している文は存在しているので，印象が大きく変わらないからといって不要なルールであるとは一概には言い切れない．

5.2.3 事後アンケートから見るかわいい文の傾向

事後アンケートとして「どういった文がかわいいと感じたか」「どういった文がかわいくないと感じたか」について回答してもらった．特に用意した変換ルールとは異なる要素について，かわいいと感じた文についての回答を以下に挙げる．

- ゆるい内容の世間話的な会話
- 食べ物の話が含まれる文
- 全体的に柔らかく，ポジティブな印象を与える文
- 日常のふとした感じの会話

表 6: 変換によってかわいくならなかった原因の分類

かわいくならなかった原因	該当する文数 (/470)
ネガティブな内容の文	105
前後の会話が無いと理解し辛く，違和感を感じる文	90
難しい話題や説明的な内容を含む文	76
昔話やお酒の話など，若い人があまりしなさそうな話題の文	25
お金の話題についての文	14
変換前後であまり印象が変わらない文	14
教訓や説教と言った，相手に言い聞かせるような内容を含む文	7
その他 (上記のどれにも当てはまらないもの，原因不明のもの)	139

表 7: 変換ルールごとの「かわいい」と答えた人数の増減 (%)

ジャンル	ルール	増加	減少
幼さ，未成熟さ	名詞をひらがなに置換	50.88	12.87
	副詞をオノマトペに置換	48.28	12.07
	相槌・同意・終助詞長音化	48.94	8.51
	語尾表現置換	53.93	9.63
	感動詞の挿入・置換	64.40	8.20
	名詞にさんを付ける	43.40	11.32
精神的な気の弱さ	断定表現を婉曲表現に置換	37.84	20.27
	曖昧な表現を追加	55.56	16.67
	命令をお願いする表現に置換	61.29	9.68
	逆接の接続詞を削除	47.44	12.82
	強い否定を緩和する表現の追加	34.78	19.57

- 「うふふ」などの笑い声に準ずる言葉が含まれた文

また，かわいくないと感じた文についての回答を以下に挙げる．

- 内容が硬く，難しく感じる会話
- 会話の前後がよくわからず，内容が意味不明な文
- 否定的な文や，同調・共感とは逆の文
- きっぱりした断定表現や，誰かを突き放すような棘のある表現を含む文
- 「うふふ」や「ふふふ」が使われている文

これらを比較すると，5.2.1 項での考察と同様，文の内容から受ける印象がかわいさの判定に大きく寄与しているという事がわかる．また，被験者にかわいくないと判定した文について理由を聞くと，「表現はかわいいと思ったが，文の内容的にかわいいとは感じなかった」という解答が一定数得られた．

5.2.4 考察のまとめ

文におけるかわいさには，「表現から感じるかわいさ」と「内容から感じるかわいさ」の 2 つの要素があり，それら 2 つの要素が同時に存在する場合に，読み手はその文から「かわいい」という印象を感じていたと考えられる．本システムでは，「表現から感じるかわいさ」を文に対して付与する事はできるが，「内容から感じるかわいさ」については十分に付与する事ができない．そのため，元々の文の内容が難しい場合やネガティブな内容に触れている場合は，十分なかわいさを付与する事ができないといえる．

「内容から感じるかわいさ」を付与するためには，難しい内容の文やネガティブな内容の文を，ゆるい内容の文やポジティブな内容の文に置き換える必要があり，既存のルールによる変換では対応できない．また，文の内容を書き換えると対話の内容も変わってしまうという問題点も存在する．そのため，本システムの手法とはまた別のアプローチを試みる必要がある．

6. おわりに

対話文に対して「かわいさ」を付与するシステムについて述べ，提案システムがかわいさをどの程度付与できるかを実験により検証した．実験の結果，約 50%の文に対して一定のかわいさを付与できる事を確認した．かわいさを付与できなかった文については，ネガティブな内容の話や，説教や教訓を含む話，若い人があまりしなさそうな内容を含む話など，表現を変換しても内容からかわいいとは感じない文が多かった．本システムでは「かわいい表現」を付与する事はできるが，「かわいい内容」を付与する事はできないため，難しい文やネガティブな文に対しては十分な「かわいさ」を付与できなかったと考えられる．

今後は，かわいさを付与できる文と付与できない文を区別するとともに，より長い一連の対話文に対してかわいさを付与する量とタイミングを検討していきたいと考えている．

参考文献

[1] 狩野芳伸 :「コンピューターに話が通じるか-対話システムの現在」, 情報管理学会誌，Vol.59, No.10, pp.658–665 (2017)

[2] 刀山将大, 佐藤理史, 松崎拓也, 宮崎千明, 平野徹, 松尾義博:「文のどこにキャラクタ性を埋め込む自由度があるか」, 言語処理学会第 22 回年次大会発表論文集, pp.721–724 (2016)

[3] 宮崎千明, 平野徹, 東中竜一郎, 牧野俊朗, 松尾義博, 佐藤理史:「文節機能部の確率的書き換えによる言語表現のキャラクタ性変換」，人工知能学会誌，Vol31,No.1, pp.1–9 (2016)

[4] 谷川晃大, 藤田寛泰, 壹岐太一 :「変分オートエンコーダと注意機構を用いた発話文のキャラクタ性変換」, 第 32 回人工知能学会全国大会，4G2-01 (2018)

[5] 石川なつ美 :「『かわいい』の意味について」，東京女子大学言語文化研究 24，pp.21–35 (2015)

[6] 家内嘉宏, 入戸野宏 :「共感性と親和動機による“かわいい”感情の予測モデル構築」，パーソナリティ研究，Vol.23, No.3, pp.131–141 (2015)

[7] 入戸野宏 :「“ かわいい ”感情の心理学モデル」，情報処理学会誌，Vol.57, No.2, pp.128–131 (2016)

[8] kuromoji : (URL)https://www.atilika.com/ja/kuromoji/

[9] 名大会話コーパス : (URL)https://mmsrv.ninjal.ac.jp/nucc/

[10] 藤村逸子, 大曽美恵子, 大島ディヴィッド義和 :「会話コーパスの構築によるコミュニケーション研究」(2011)

[11] 藤村逸子, 滝沢直宏編 :「言語研究の技法 : データの収集と分析」, ひつじ書房, pp.43–72 (2011)

[12] iM@S-CGWordsSearch(β) : (URL)https://icws.indigo-bell.com/

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

[4G3-OS-8b] マイニングと知識創発(2)

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

Fri. Jun 7, 2019 2:00 PM - 2:40 PM Room G (302A Medium meeting room)

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

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

2:00 PM - 2:20 PM

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

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

2:20 PM - 2:40 PM

©The Japanese Society for Artificial Intelligence

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

対戦型ゲームにおける個人の性格や能力を反映したAIプレイヤーの作成

AI Player Creation Reflecting Personal Abilities and Characters in PvP Games

坂田 駿允 *1 Toshinobu Sakata　砂山 渡 *2 Wataru Sunayama　畑中 裕司 *2 Yuji Hatanaka　小郷原 一智 *2 Kazunori Ogohara

*1滋賀県立大学大学院工学研究科 Graduate School of Engineering, The University of Shiga Prefecture

*2滋賀県立大学工学部 School of Engineering, The University of Shiga Prefecture

Nowadays, PvP online games for playing with many players are very popular. Among them, we can feel more pleasure to playing with friends and players who became friends by online games. It is considered to be able to perceive player ' s abilities and characters.

In this study, we aim to construct environment playing with AI players who have any personality within a settable range. We can be set freely AI player ' s abilities and characters parameter. In experiment, we verified how to reflect personality to In-game action. And we investigate the difference of easiness in perceive AI player ' s personality.

1. はじめに

近年，プロ棋士に勝った将棋 AI[1] や，世界レーティング 1 位の囲碁プレイヤーに勝った囲碁 AI[2] など，人間より強いゲーム AI に注目が集まっている．しかし，これらの AI は強い人間に勝つために生成された AI であり，そのゲームの初心者や中級者は強いゲーム AI と対戦しても楽しむことは難しい．また，ゲーム AI は，人間と違い機械的な行動を取ることが多く見られ，プレイヤーが違和感を感じてしまう．

私たちが友人やオンラインでなどで人間と対戦型ゲームを行う際，コンピュータと対戦するよりも楽しいと感じることができる．何故ならば，コンピュータと対戦する時は，勝ち負けでのみ，そのゲームに対する楽しさを感じ取ることしかできないが，人間と対戦する場合，勝ち負けのみでなく，プレー中の相手の行動や心境，癖など，人間特有の行動を起こし，勝ち負け以外の楽しめる要素を感じ取れることができるからである．

そこで，本研究は，対戦型ゲームにおいて，ゲーム AI に性格や能力などの，人間の持つ個性を適用し，その個性を表したゲームプレイを行うことにより，まるで人間と対戦しているように楽しめる AI プレイヤー作成することを目的とする．

2. 関連研究

2.1 ゲームにおける AI プレイヤーに関する研究

身体的な制御と生き延びるために必要な欲求を「人間らしさ」として，人間らしい振る舞いを伴うゲーム AI の自動取得をする研究がある [3]．この研究では，人間の身体的な「疲れ」や「遅れ」，特定行動に対する「訓練」や「挑戦」によって AI プレイヤーを生成しているが，本研究では，人間の「性格」や「能力」といった「個性」を適用して AI プレイヤーを生成することを目的とする．

人間の操作履歴を元に学習し，模倣することにより，人間らしさを表現する AI プレイヤーを作成する研究がある [4]．この研究では，操作ミスをしない，短時間で操作が変わるといった，AI プレイヤーが行う非人間的な行動を改善し，AI の行動に反映させている．本研究では，行動を決定する際に，ゲームルールから考えられる，非人間的な行動のみを排除し，それ以外の行動は人間の性格や能力から適用する．

2.2 対戦型ゲームにおける AI プレイヤーに関する研究

プレイヤーの行動や対戦記録などから AI プレイヤーを生成する研究がある [5][6][7]．この研究では，AI プレイヤーを生成する際に，対戦するプレイヤーの行動を用いて，プレイヤーと同程度の強さを持った AI プレイヤーを生成している．本研究では，第三者の性格や能力を取得して，その性格や能力を表す行動を行う AI プレイヤーの生成を目的とする．

プレイヤーの行動やゲームの状況によって対戦相手に感情的な反応を返す研究がある [8]．この研究では，ゲームの状況に応じて快-不快パラメータを設定し，そのパラメータに対応する反応を返している．本研究では，ゲームの状況だけでなく，AI プレイヤーの性格や個性も判定要素として取り込み，反応を返す AI プレイヤーを生成する．

人間の個性的なプレースタイルを模倣する AI の研究がある [9]．この研究では，AI に適用する行動として人間のプレースタイルを模倣しているが，本研究では，人間の性格や能力から考えられる行動を AI プレイヤーに適用することを目的とする．

3. 対戦型ゲーム「大富豪」における AI プレイヤーの性格と能力

本章では，トランプの対戦型ゲーム「大富豪」における，考えられる性格や能力と，それらに基づいた行動について述べる．

3.1 対戦型ゲーム「大富豪」のルール設定

本研究では，性格や能力を適用する AI プレイヤーを評価する対戦型ゲームとして，「大富豪」を題材とする．ただし，以下の 5 つをルールとする簡易版の大富豪を取り扱う．

- Joker を除いた 4 つのスートに各 A〜K までの 13 枚のカード，計 52 枚のカードを使用
- 3 が最弱で，数が大きくなるにつれて強くなり，A は K より強く，2 が最強
- カードは必ず一枚ずつ出す
- 8 切りや革命などの特殊ルールは無効

連絡先: 坂田駿允，滋賀県立大学大学院工学研究科電子システム工学専攻，〒 522-8533 滋賀県彦根市八坂町 2500

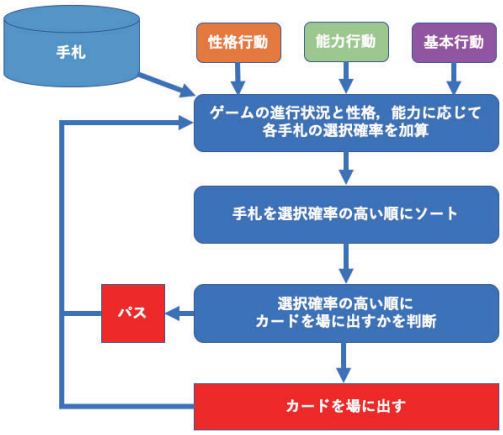


図 1: 大富豪における AI プレイヤーの行動の枠組み

表 1: 大富豪における性格や能力と対応する行動

性格や能力	行動
記憶力	場に出たカードを記憶する
リスク志向	強いカードを出しやすい 相手が上がりそうな時，阻止する
気まぐれ	適当にカードを出しやすい
負けず嫌い	一定以上の順位を目指す
焦りやすい	ピンチに強いカードを出しやすい

- ゲーム継続時のカード交換をしない

図 1 に、大富豪における AI プレイヤーの行動の枠組みを示す。AI プレイヤーは、設定された性格と能力に応じて、場に出すカードを決定する。

3.2 AI プレイヤーの性格と能力による行動

本研究では、表 1 に示す 1 つの能力「記憶力」と、4 つの性格「リスク志向」,「気まぐれ」,「負けず嫌い」,「焦りやすい」の実装を試みる。各性格と能力を表すため、それぞれの強さを表すパラメータ P_{memory}、P_{risk}、P_{moody}、$P_{competitive}$、P_{panic} を 0 から 1 の範囲の実数値で設定する。

大富豪において，AI プレイヤーが行える行動は，各カードを場に出す行動とパスに限られる．また AI プレイヤーが持つカードには、それぞれ選択確率が設定され、選択確率が高いものから順に、カードに与えられている確率でそのカードを出すか否かが決定される。なお全てのカードを出さない選択がなされた場合は、行動としてパスを選択する。ただし状況によっては、確率によらず出すカードが決定されることがある。

AI プレイヤーの基本行動、ならびに各性格や能力のパラメータに基づいて、カードに選択確率 C が式 (1) で与えられる。ただし、M_{basic}、M_{1risk}、M_{2risk}、M_{moody} は後述の、基本行動や性格、能力によって与えられる値となる。

$$C = M_{basic} + M_{1risk} + M_{2risk} + M_{moody} \tag{1}$$

また、正規化前のカードの選択確率の最大値を C_{max} として、カードの最大選択確率の値が 0.9 になるように式 (2) により正規化を行う。

$$\dot{C} = \frac{0.9}{C_{max}} C \tag{2}$$

また、場に出せないカードの選択確率は 0 に，必ず出すカードの選択確率は 1 にする。各カードに割り当てられる選択確率 M はで表される．

3.2.1 AI プレイヤーの基本行動

- 場に出せるカードの小さい順に選択確率を大きく設定する。
- 場を流せるカード以外のカードが一枚以下の場合，流せるカードから出す（他人に上がられるリスクを避ける）．
- 場に何も出ていない場合，最小カードから出す．
- カードが残り一枚かつそのカードが場に出せる時，パスをしない（パスが選択されないようにする）。

基本行動により各カードには、m 番目に小さいカードに、式 (3) によるパラメータ M_{basic} を与える。これは，大富豪の基本的な戦略として，場に出せる小さいカードからカードを減らすことを表している。

$$M_{basic} = \frac{0.3}{m} \tag{3}$$

3.2.2 性格と能力による行動

本実験で実装する性格や能力の大富豪における行動を表 1 に示す．

「記憶力」の行動は，場に出たカードを記憶し，場を流せるカードを作れる能力で記憶力を表し，カードの忘却頻度でその能力の強弱が表現できると考えた．

「リスク志向」は，積極的に勝負に挑む行動として，強いカードをよく出しに行き，相手が上がりそうな状況では，その状況を阻止するようなリスクの高い行動の頻度で表現できると考えた．

「気まぐれ」は，セオリーでない行動を表現する方法として，出せるカードのうち，場に出せる最小のカード以外を出すような行動，例えば，3 が場に出ている時に Q のように強いカードを出すことにより，気まぐれを表現できると考えた．

「負けず嫌い」は，勝負事に負けたくないという思考の強さを，目標順位として設定し，その順位を目指すような行動を行うことで表現できるのではないかと考えた．

「焦りやすさ」は，焦った時の判断能力の低下を表現するために，大富豪において，自分が負けそうになった時に，有利不利関係なく相手の行動を阻止する，つまり強いカードを出すことにより表現できるのではないかと考えた．

3.3 AI プレイヤーの性格と能力のパラメータ設定と行動

3.3.1 記憶力

記憶力によって行われる行動は，自分のカードと場に出たカードをすべて覚え，自身のターンが来た時に，記憶力パラメータ P_{memory} を用いて，式 (4) で表す確率 F_{memory} で忘れるようにする．また，記憶力の行動の目的は，場を流せるカードを増やすことなので，カードを覚え，忘れるカードの対象は，10 以上のカードとする．

$$F_{memory} = (1 - P_{memory}) \times 100(\%) \tag{4}$$

表 2: 負けず嫌いパラメータによる目標順位の設定

負けず嫌いパラメータ $P_{competitive}$	目標順位
1〜0.9	1 位
0.89〜0.75	2 位
0.74〜0.2	3 位
0.19〜0	4 位

3.3.2 リスク志向

自分のカードの n 番目に強いカードから順に、選択確率 M_{1risk} を、リスク志向パラメータ P_{risk} 用いて式 (5) より算出する。

また、4 人での対戦時において、中盤以降すなわち最後に場にカードを出したプレイヤーの手持ちカードが 7 枚以下の時に，自分の持っている一番強いカードに，式 (6) より算出した，選択確率 M_{2risk} を適用する．リスク志向の行動により，強いカードに大きく選択確率を適用し，その後は徐々に選択確率を減らしながら適用することで，リスク志向のパラメータが大きいほど，強いカードを出しやすくしている．また対戦相手のカードが減り始めた時に，一番強いカードのみに選択確率を割り当て，相手の行動を阻止する勝負に出る．

$$M_{1risk} = \frac{P_{risk}}{2n} \tag{5}$$

$$M_{2risk} = \frac{P_{risk}}{2} \tag{6}$$

3.3.3 気まぐれ

気まぐれは，そのパラメータが大きいほど，適当にカードを出しやすくなるように，気まぐれパラメータ P_{moody} に定数を掛ける式 (7) により，各カードに選択確率 M_{moody} を適用する．すなわち，気まぐれパラメータが大きいほど，予測不能なカードを出しやすくなる．また定数 R は、毎回 0 から 1 の範囲の実数値が乱数で生成されて与えられる。

$$M_{moody} = P_{moody} \times R \tag{7}$$

3.3.4 負けず嫌い

負けず嫌いパラメータ $P_{competitive}$ により、表 2 の目標順位を設定する。目標順位が達成できなかった時には、気まぐれパラメータ P_{moody} を 1 にする．

3.3.5 焦りやすい

大富豪において焦りやすい状況とは，相手が上がりそうな時に，自分が上がれる見込みがない時に発現しやすい．さらに，その状況を打開するため，無理やり自分のカードを減らそうとする行動を取る．すなわち、ゲーム中盤のプレイヤーが 8 回カードを出すかパスを行った後に，自分のカードの枚数順位が下位だった場合に，焦りやすさを表すパラメータ P_{panic} を確率として、P_{panic} の確率で再度リスク志向の行動を行う（M_{1risk} と M_{2risk} の加算を再度行う）ように設定する．

3.4 対戦型ゲーム「大富豪」のインタフェース

本節では，大富豪のインタフェースに実装した機能について述べる．図 2 に，大富豪のインタフェースを示す．今回，インタフェースを実装する環境として，TETDM[*1] の可視化ツールを用いた．

*1 TETDM サイト：http://tetdm.jp



図 2: 大富豪のインタフェース

表 3: 設定した AI プレイヤーの性格や能力パラメータ

	記憶力	リスク志向	気まぐれ	負けず嫌い	焦りやすい
くま	0.5	0.8	0.2	0.7	1.0
うさぎ	0.1	0.1	1.0	0.1	0.1
いぬ	0.3	1.0	0.5	0.8	0.3
ねこ	1.0	0.1	0.1	0.5	0.1

4. 対戦型ゲームにおける AI プレイヤーの性格推定実験

4.1 実験設定

理系の大学生大学院生 11 名に、性格や能力を適用した AI プレイヤー同士の対戦を観戦してもらった後で，各 AI プレイヤーの性格や能力を推定してもらい、AI プレイヤーの行動から AI プレイヤーの個性を感じ取れるかを確認する実験を行った．用意した 4 人の AI プレイヤー（くま，うさぎ，いぬ，ねこ）のパラメータを表 3 に示す。

観戦回数は 6 回から 9 回の間で、およそプレイヤーの能力と性格を推定できたと思うまで観戦してもらった。各性格と能力について、「記憶力が高い」「リスク志向」「気まぐれ」「負けず嫌い」「焦りやすい」のそれぞれに当てはまる度合いを 5 段階で回答してもらった。なお、9 回観戦してもらってもわからなかったときには、わからなかったと回答してもらった。

4.2 実験結果

表 4 から表 7 に各 AI プレイヤーの能力と性格の推定結果を示す。ただし各表の黄色いセルは、AI プレイヤーに設定したパラメータを 5 段階評価に合わせた対応推定値（0 から 0.2 を 1、0.3 と 0.4 を 2、0.5 と 0.6 を 3、0.7 と 0.8 を 4、0.9 と 1.0 を 5）を表す。

表 4 から表 7 の記憶力の項目を見ると，どの記憶力パラメータの値でも対応推定値に近い評価ができていた．これは，場の記憶と忘却による記憶力の表現や，パラメータによる表現の強弱ができていたためと考えられる。

表 4 から表 7 のリスク志向の項目を見ると，“くま”と “ねこ”の 2 つの AI プレイヤーで対応推定値に近い評価ができていた．リスク志向パラメータは高く設定された “いぬ”と “くま”について、“くま”は焦りやすさが高く，“いぬ”では焦りやすさが低く設定されていた．焦りやすさは「パラメータが高いとリスク志向行動を取りやすくなる」ため，“くま”でリスク志向の評価が高かったのは，焦りやすさによってリスク志向行動

表 4: AI プレイヤー “くま”の評価結果

評価	記憶力	リスク志向	気まぐれ	負けず嫌い	焦りやすい
1	2	0	2	0	1
2	3	1	4	1	3
3	3	4	1	1	4
4	1	1	1	5	3
5	1	5	2	4	0
0	1	0	1	0	0

表 5: AI プレイヤー “うさぎ”の評価結果

評価	記憶力	リスク志向	気まぐれ	負けず嫌い	焦りやすい
1	1	0	0	0	0
2	6	0	4	3	1
3	1	4	3	2	1
4	3	5	2	5	3
5	0	2	1	1	6
0	0	0	1	0	0

表 6: AI プレイヤー “いぬ”の評価結果

評価	記憶力	リスク志向	気まぐれ	負けず嫌い	焦りやすい
1	6	0	0	0	0
2	2	3	3	5	1
3	2	4	4	3	4
4	1	3	3	2	6
5	0	1	0	1	0
0	0	0	1	0	0

表 7: AI プレイヤー “ねこ”の評価結果

評価	記憶力	リスク志向	気まぐれ	負けず嫌い	焦りやすい
1	0	4	5	5	1
2	0	6	3	2	4
3	2	1	1	2	4
4	0	0	1	2	2
5	9	0	0	0	0
0	0	0	1	0	0

が強く現れたためと考えられる．そのため、現状のリスク志向の行動は目に見えて反映されにくい可能性が考えられる．

表 4 から表 7 の気まぐれの項目を見ると，全 AI プレイヤーにおいて，パラメータの値に関係なく，低く評価される傾向にあった．このことから，性格を表す行動として，幅広い値を取りうるランダム要素を用いるとうまく性格が推定されないことがうかがえる。

表 4 から表 7 の負けず嫌いの項目を見ると，“くま”のみ対応推定値に近い評価が得られた．しかし，“くま”と負けず嫌いパラメータが同程度の “いぬ”では，低い評価が多く，他 2 つの AI プレイヤーでも低い評価が多かった．これは，負けず嫌いパラメータの行動である「目標順位が達成されなかった場合，気まぐれパラメータを最大にする」という行動において、気まぐれが推定されにくかったことと合わせ、負けず嫌い独自の行動が設定されていなかったことによると考えられる。そのため，各性格ごとに独自の変化が明らかになるような行動の設定が必要と考えられる．

表 4 から表 7 の焦りやすさの項目を見ると，全 AI プレイヤーにおいて，対応推定値とは遠い評価が得られた．これは，焦りやすさを表す行動が,「順位が低い場合に，リスク志向行動を取る」のみであり，焦りやすさ特有の行動を設定していなかったためと考えられる．

5. おわりに

対戦型ゲームにおいて，人間と対戦しているように楽しめる AI プレイヤーとして，設定した性格や能力を適用し，行動を行う AI プレイヤーを作成した．

作成した AI プレイヤーが設定した性格や能力を反映し，行動によって表現できているかを検証するために，評価実験を行った．記憶力のように能力として反映されやすいものは、行動から推定しやすいことがわかった。性格として用意されるものは、その性格に基づく行動が強く現れるようにしないと、行動からの推定が難しいことがわかった。

今後は、今回題材として扱った性格について、より性格を推定しやすいパラメータや行動の設定を検討していきたい。また、同様の性格を他のゲームに応用するための一般的な方法論について研究を進めていきたいと考えている。

参考文献

[1] 瀧澤武信：現役プロ棋士に勝ち越したコンピュータ将棋～第 2 回電王戦，第 23 回世界コンピュータ将棋選手権速報～：4, 第 23 回世界コンピュータ将棋選手権速報, 情報処理, Vol.54, No.9, pp.925–928 (2013)

[2] 松原仁：AlphaGo の置き土産, 情報処理, Vol.58, No.9, pp.668–669 (2017)

[3] 藤井叙人・佐藤祐一・中嶌洋輔・若間弘典・風井浩志・片寄晴弘：生物学的制約の導入による「人間らしい」振る舞いを伴う ゲーム AI の自律的獲得, ゲームプログラミングワークショップ 2013 論文集, pp.73–80 (2013)

[4] 中野雄基・美添一樹・脇田建：進化計算と UCT による Mario を人間らしくプレイする AI, ゲームプログラミングワークショップ 2013 論文集, pp.81–88 (2013)

[5] 上田陽平・池田心：遺伝的アルゴリズムによる人間のレベルに適応する多様なオセロ AI の生成, 研究報告ゲーム情報学 (GI),Vol.2012,No.5,pp.1-8,(2012)

[6] 服部裕介 ・田中彰人 ・星野唯一 ：対戦型アクションゲームにおけるプレイヤの模倣行動の生成, 情報処理学会研究報告ゲーム情報学 (GI),Vol.2007,No.20,pp.1-8,(2007)

[7] 星野唯一・田中彰人・濱名克季：模倣学習により成長する格闘ゲームキャラクタ, 情報処理学会論文誌, Vol.49, No.7, pp.2529 － 2548 (2008)

[8] 塩入健太,・星野准一：仮想対戦プレイヤーの感情的発話生成, 情報処理学会研究報告エンタテインメントコンピューティング (EC),Vol.2005, No.125, pp.63–70 (2005)

[9] 生井智司・伊藤毅志：将棋における棋風を感じさせる AI の試作, 情報処理学会研究報告ゲーム情報学 (GI),Vol.2010.3, No.3, pp.1–7 (2010)

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

位置情報データを用いた公共施設来場者数の予測可能性

-野球興業の来場者数予測モデルの構築を中心に-

柳 博俊[*1 *3] 中原 嘉利[*2] 星野 崇宏[*2 *3]
Hirotoshi Yanagi Yoshitoshi Nakahara Takahiro Hoshino

[*1] 慶應義塾大学大学院経済学研究科 Graduate School of Economics, Keio University
[*2] 慶應義塾大学経済学部 Faculty of Economics, Keio University
[*3] 理化学研究所 Riken AIP

It is the purpose of this paper to discuss whether we could predict the number of visitors of public facilities from the GPS location big data collected from approximately 20 million smartphones in total while taking into consideration influences of various covariates to correct several biases relate to the penetration ratio. We used number of visitors of baseball games in the year 2018 as a label, and GPS location big data and other features as feature vectors to apply supervised machine learning methods. Because of the prediction accuracy, we concluded that it is possible to predict number of visitors of public facilities from GPS location big data.

1. はじめに

本稿ではスマートフォンから取得されたアプリ位置情報を用いた, 公共施設来場者の予測可能性の検討ついて報告する.

近年,スマートフォンの普及に伴い, 膨大な GPS ログが収集されている. これらの情報により, 海水浴場など来場者の数の正確な計測が困難な場所や, テーマパークや公共施設のような計測値が発表されていても月単位または半期/通期単位での公表しかなかった場所について, 即時的にその周辺人口を推定できる可能性がある. これらの情報を用いることによって, 施設利用者がよりリアルタイムな混雑状況についての情報を取得することが可能になったり, 来場者数そのものがビジネス上の重要な KPI になるような施設であれば, 公表前にその来場者数の情報を取得することによって決算公表前にその施設の経営状況を推察することが可能となる.

このように GPS データを用いた適用方法が多く考えられる一方で, GPS データから周辺人口の推定を行う方法についての研究/実適用例は多くなく,NTT ドコモの「モバイル空間統計」や agoop 社の提供する月別メッシュ別人口など集計に限られたものとなる.

本研究では, 公共施設の来場者数予測の可否について, 来場者のデータが公表されているプロ野球興行の来場数データを正解値に, X 社提供を受けた位置情報データ/試合に関するデータを特徴量として用いて, Kaggle など予測系タスクに広く使われる LightGBM により予測が可能かを検証する.

2. 関連研究

2.1 従来型人口統計について

従来の人口統計は、全数調査である国勢調査によって判明する昼間人口や夜間人口などの人口静態統計が主であり、公表の最小単位は地域メッシュである(総務省統計局(2015))。具体的には、「約 1km 四方の基準地域メッシュ、約 500m 四方の 2 分の 1 地域メッシュ及び約 250m 四方の 4 分の 1 地域メッシュ単位に再集計(総務省統計局(2017))」され、公表されている。故に、従来の統計情報では自身の関心のあるエリア、あるいは施設に滞在している流動人口をリアルタイムで把握することはほぼ不可能である。

2.2 位置情報を用いた, 動的な人口統計について

2.1 の問題点を解決し流動人口を推定すべく、交通系 IC カードを用いた方法や(牧村ら(2010))携帯電話の基地局圏在数を用いた方法(寺田ら(2012))が考案されている。特に後者の方法は、NTTドコモが提供する「モバイル空間統計」として商用化されており、推定される動的人口の代表性に関しては複数の研究が行われている(大藪ら(2017)、矢部ら(2016))。このデータベースは。震災後の避難者数の推定や帰宅困難者数の推定への利用が期待されている(村上ら(2011))。

また本研究と関連するところとして, 携帯電話の GPS(Global Positioning System)を利用した位置情報データから流動人口を推定する技術の開発が進みつつある。実例を挙げると、RESAS では、株式会社 Agoop の提供する月別・メッシュ別の動的人口を取得することができる。

3. データおよび変数について

3.1 予測に用いるデータについて

今回の研究では、X社にご提供いただいた約延べ2000万人分、5分毎の位置情報データを用い、2018年のプロ野球のホームゲーム全 786 試合の観客動員数を予測した。

プロ野球の対戦カード、使用スタジアム、雨天中止の有無、観客動員数、各球場の最大収容人数、予告先発投手、開始時刻、試合時間などの基本的な情報は、プロ野球 Freak(2018)が各球団の発表した値を収集したデータを使用した。なお、このデータは、日本プロ野球協会の公式 HP で公開されている集計値と照らし合わせて正確であることを確認してある。

また、各球団は、ホームスタジアムの試合に関し、有料販売したチケット枚数に無料配布分の実来場者数を加算した値を観客動員数として公表している (伊藤(2016))。なお、無料配布チケットは球場の窓口で受け取る仕組みになっており、無料チケット所有者が実際に来訪しているのか観測できる仕組みとなっている(伊藤(2016))。

観測端末数に関しては、試合時間中に球場で観測されたユ

連絡先: 柳博俊 , h.yanagi-95@keio.jp

ニークな ID 数を集計して求めた。なお、近隣施設を利用していたにも関わらず、GPS の測定誤差によって野球場に居たとカウントされることを防ぐため、最低でも 12 回(延べ 1 時間)球場で観測された ID のみを観測端末数に加算している。

懸念点としては，位置情報を取得している端末属性にかかるバイアスがあげられる．位置情報は特定アプリケーションをインストールしているスマートフォンから取得されており，そのアプリケーションの利用者属性によりバイアスが生じる．具体的には

1.インストール率は地域、年代、性別によって異なる。
2.インストール率は時間経過とともに変化し、一定ではなく、またその普及率の変化にも偏りがある。

ことが集計によりわかっている．

3.2 来場者数と観測端末数の関係の検討

観客動員数と観測端末数に相関関係がなくては、予測モデルを構築することは不可能である。

図 1 から分かるように、観客動員数と観測端末数には比較的強い正の相関があることがわかる。しかしながら、観客動員数が増えるにつれ、観測端末数の分散が大きくなる、不均一分散の傾向が観測されている。

分布に関して比較すると、観測端末数はほぼ対数正規分布に従っているのに対し、観客動員数は 3 万人を境として分布が大きく下がり、4 万 5 千人の周辺にもう一つのモードがあることがわかる。このことは、図 2,3 からも明らかで、観測端末数の分布と観客動員数の分布は、一致していないことが分かる。

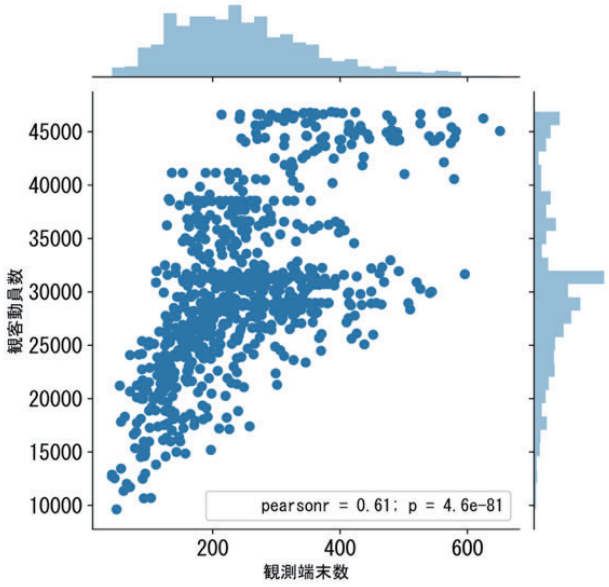


図 1 全試合の観客動員数と観測端末数

3.3 来場者数のカウントノイズの検討

観客動員数について，12 球団中複数の球団の発表値をみると測定ノイズではないかと思われる傾向が見つかった．観客動員数が満員に近い試合の場合，動員数がそれほど多くない試合とは異なり，端末の観測数に対して大きい値が発表される傾向にあるというものである．

例えば，図 2 では球団 A がホームチームの場合，観客動員数の発表値が 30000 人を超えたあたりから，付近での観測端末数に対して大きい値を示す傾向が見られる．

満員試合数というのは，球団経営において，重要なビジネス KPI なりうると考えられるため，観客の収容率が高い試合の場合はこのような補正がかかってしまっているのではないかと考えられる．

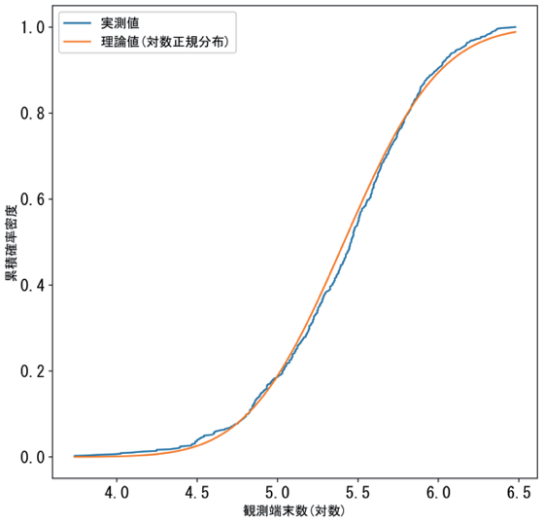


図 2 観測端末数の分布

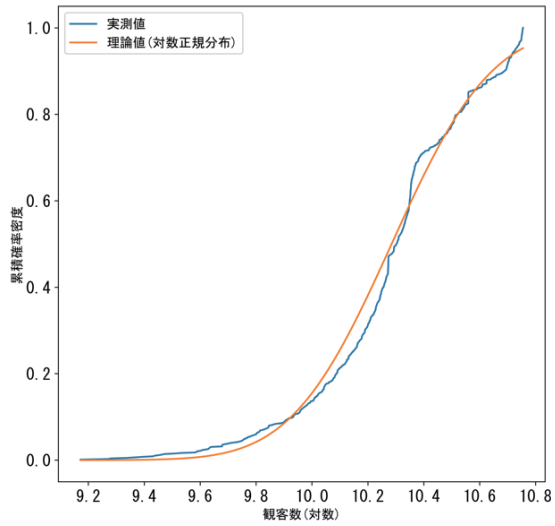


図 3 観客動員数の分布

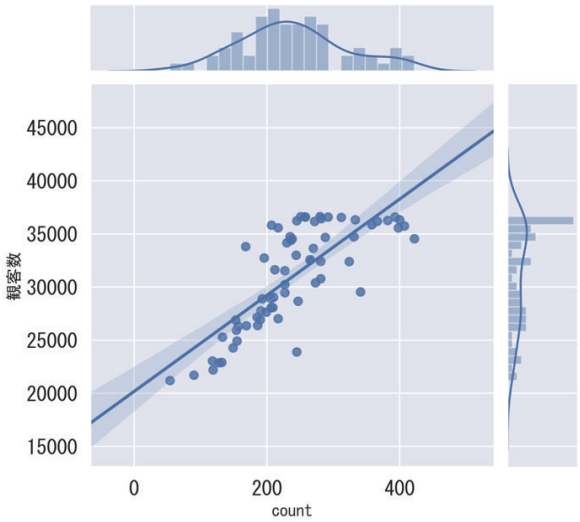


図 4 球団 A の観客動員発表値と観測端末数の関係

連絡先: 柳博俊 , h.yanagi-95@keio.jp

4. 実証分析

4.1 正解値定義

3.1 で述べたように各球団が有料チケット販売数と招待客数を合計したものを毎試合公表しており，それを正解値とする予測問題を解く．

4.2 特徴量

試合中に球場付近で観測された端末数を用いて，予測を行う．

また，試合を行う球場や，ファン層によって，実来場数のうちデータが取得される端末の数の含有率が異なると考えられる．これらを補正するために，

対戦相手 / 先発投手などを用いて，それらを補正することを試みる．

＜変数の定義＞

連続値特徴量	
観測端末数	試合が行われている最中に連続 1 時間以上球場周辺で観測された端末の数
ホームチーム	試合のホームチームごとに観客数について TargetEncoding 処理したもの
ビジターチーム	試合のビジターチームごとに観客数について TargetEncoding 処理したもの
月	試合の行われた月ごとに観客数について TargetEncoding 処理したもの
曜日	試合の曜日ごとに観客数について TargetEncoding 処理したもの
予告先発	ホームチームの予告先発投手ごとに観客数について TargetEncoding 処理したもの
OneHot 特徴量	
ホームチーム	試合のホームチームの OneHot 特徴
ビジターチーム	試合のビジターチームの OneHot 特徴
月	試合が行われた月の OneHot 特徴
曜日	試合が行われた月の OneHot 特徴
予告先発	ホームチームの予告先発投手

4.3 予測方法 / 精度評価

Microsoft 社の提供する勾配ブースティング木アルゴリズムである LightGBM を用いて，データを 10 分割した上で Leave-One-Out 方式で予測を行う．Out of Fold(OoF)の値を予測値として扱う．

精度評価には Mean Absolute Error（MAE）と Mean Absolute Percent Error 及び Mean Squared Error を使用する．

4.4 結果

結果は表の通りになった．

精度指標	MAE	MAPE	RMSE
予測精度	4266.11	0.17072	5962.79

表 予測誤差

また，特徴重要度の上位は図 5 のようになり，観測端末数が最も予測に貢献していることがわかる

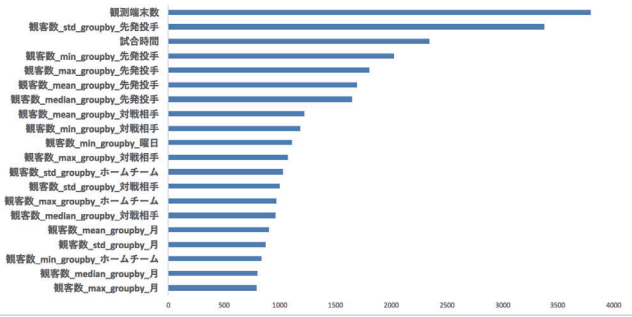


図 5 特徴重要度 Top20

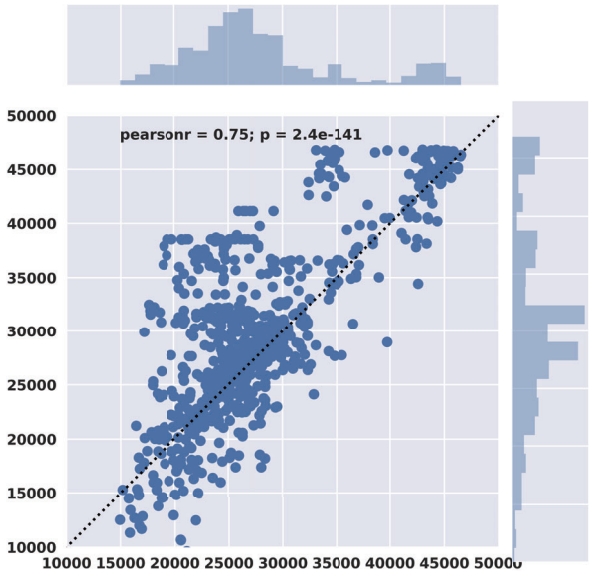


図 6 OOF に対して予測の結果

5. 今後の課題 / 展望

最後に、本研究によって浮かび上がった、今後取り組むべき研究課題に関し議論する。

5.1 過学習の懸念

第 1 の課題意識は、過学習の懸念である。今回は、位置情報データが 2018 年 4 月からしか利用できなかったため、学習データ、テストデータともに野球の 2018 年シーズンから作成している。そのため、本研究で作成した機械学習モデルが 1 シーズンのデータに対し過学習している可能性がある。この問題点に関しては、データの蓄積を待って、複数シーズンのデータに対し、本論で適用した機械学習モデルを適応して予測精度を確認する必要がある。

5.2 位置情報集計のノイズ

第 2 の課題は、商業施設が密集している地域で、野球の試合を観戦していない ID が間違って観客としてカウントされている、集計ミスの可能性である。3.2 で議論した通り、今回は、横浜市の中心部にあり商業施設やホテル、駅などと隣接し

ている横浜球場で、集計ミスの疑いがあった。この点に関しては、Gonzalez(2008)が提唱した、各位置情報端末所有者の行動特性の特徴量の抽出が必要であると考える。野球観戦特有の行動ログのパターンを特定し、そのパターンに一致している ID のみを抽出して集計することができれば、集計ミスは減らすことができると期待している。

5.3 観客動員集計のノイズ

3.3 にて,観客動員数の集計において，集計時のバイアスが疑われた．この問題を解決するため，より正確に来場のデータが取得されている公共施設について，上記懸念点を払拭した上で，今回と同様の予測を行いたい．

より正確な来場のデータが取得されたときにも，位置情報からの予測がより正確に行えることを示す必要がある．

連絡先: 柳博俊 , h.yanagi-95@keio.jp

参考文献

González, M. C., Hidalgo, C. A. and Barabási, A. L.: Understanding individual human mobility patterns, Nature, Vol. 453, pp. 779-782, 2008.

Hastie, T.; Tibshirani, R. & Friedman, J. (2009), The Elements of Statistical Learning 2nd Edition, Springer New York Inc., pp675-720,2009

伊藤歩(2016),「プロ野球は観客動員数をごまかしているのか-公表人数と肌感覚で大きな差が出るカラクリ」,『東洋経済』,2016 年 1 月 10 日,電子版

大藪勇輝, 寺田雅之, 山口高康, 岩澤俊弥, 萩原淳 一郎, 小泉大輔(2012),「モバイル空間統計の信頼性評価」,『NTT DOCOMO テクニカル・ジャーナル』, Vo.20 , No.3 , pp.17-23, 2012

RESAS 公式HP、https://resas.go.jp/#/13/13101(入手:2019 年 1 月)

総務省統計局(2017),「平成 27 年国勢調査に関する地域メッシュ統計 結果の概要 」,総務省 統計局 統計調査部,2015 年 6 月

RESAS 操作マニュアル,
https://resas.go.jp/manual/#/13/13101(入手:2019 年 1 月)

生形嘉良, 関本義秀(2013),「大規模・長時間の GPS データによる観光統計調査の活用可能性〜石川県を事例に〜」,『土木学会論文集 D3』.Vol. 69. No.(土木計画学研 究・論文集第 30 巻),2013

プロ野球 Freak(2018),https://baseball-freak.com/audience/(2018 年 11 月入手)

寺田雅之, 永田智大, 小林基成(2012),「モバイル空間統計にお ける人口推計技術」,『NTT DOCOMO テクニカル・ジャーナ ル』, Vo.20, No.3, pp.11-16, 2012

総務省統計局(2016),「平成 27 年国勢調査の集計体系及び結果の公表・提供等一覧」

村上正浩, 岡島一郎, 鈴木俊博, 山下仁(2011),「モバイル 空間統計を活用した滞留者・帰宅困難者数の推定と 具体的対策の検討」,『日本建築学会梗概集, F-1 分冊』. p.893-894, 2011

牧村和彦, 中村俊之, 千葉尚, 森尾淳, 布施孝志(2010),「バス IC カードを用いた人の動き~交通計画への活用 に向けた可能性と限界~」,『土木計画学研究・講演集』, Vol.41, 2010

連絡先: 柳博俊 , h.yanagi-95@keio.jp