Thu. Jun 7, 2018

Room P

Poster presentation | General Session | Interactive

[3Pin1] インタラクティブ(1)

9:00 AM - 10:40 AM Room P (4F Emerald Lobby)

[3Pin1-01] Optimization method for Educational Application of Structural Engineering

○Takuya Suzuki[1], Keiichi Iguchi[2], Takuya Kinoshita[1], Akihiro Ishikawa[2] （1. Takenaka Corporation, 2. HEROZ）

9:00 AM - 10:40 AM

[3Pin1-02] Benchmark Functions for Cooperative Coevolution and Analysis of Influence by Dynamic Problem Decomposition

○Yuki Nakashima[1], Jun-ichi Matsuoka[1], Asuka Hisatomi[1], Satoshi Ono[1] （1. Kagoshima University）

9:00 AM - 10:40 AM

[3Pin1-03] Incorporating Knowledge Graph Embeddings into Latent Variable Models

○Naoya Takeishi[1,2], Kosuke Akimoto[1] （1. The University of Tokyo, 2. RIKEN）

9:00 AM - 10:40 AM

[3Pin1-04] Covariate shift bias correction induced by action

○Ryota Higa[1] （1. NEC Corporation）

9:00 AM - 10:40 AM

[3Pin1-05] Batch Reinforcement Learning for Linearly Solvable MDP

○Tomoki Nishi[1], Keisuke Otaki[1], Takayoshi Yoshimura[1] （1. Toyota Central R&D Labs., Inc.）

9:00 AM - 10:40 AM

[3Pin1-06] Cappuccino: Bootstrapping from Structural Data

○Ryo Hanafusa[1], Masafumi Oyamada[2] （1. Kwansei Gakuin University, 2. NEC Corporation）

9:00 AM - 10:40 AM

[3Pin1-07] Non-negative matrix factorization with information expansion using item class information

○Katsuhiko Murakami[1] （1. The University of Tokyo）

9:00 AM - 10:40 AM

[3Pin1-08] Proposal of direct policy search using reduction of policy parameters by principal component analysis

○Yuuki Murata[1], Megumi Miyashita[1], Shiro Yano[1], Toshiyuki Kondo[1] （1. Tokyo University of Agriculture and Technology）

9:00 AM - 10:40 AM

[3Pin1-09] Curriculum Learning for Similar Learning Cases

○Yuta Ota[1], Yuta Taki[1] （1. KOZO KEIKAKU ENGINEERING Inc.）

9:00 AM - 10:40 AM

[3Pin1-10] Graph Feature Extraction Using Graph Fragment Decision Trees

○Haruki Sakagami[1], Ichigaku Takigawa[1], Hiroki Arimura[1] （1. Hokkaido University）

9:00 AM - 10:40 AM

[3Pin1-11] Objective Correction for Policy Improvement under Entropy Regularization

○Ryo Iwaki[1], Minoru Asada[1] （1. Osaka University）

9:00 AM - 10:40 AM

[3Pin1-12] Experimental evaluation of Time-Series Gradient Boosting Tree with Time-Series Benchmark datasets

○Mitsuyoshi Imamura[1,2], Kei Nakagawa[1,2], Kenichi Yoshida[1] （1. University of Tsukuba, 2. Nomura Asset Management Ltd.）

9:00 AM - 10:40 AM

[3Pin1-13] Batch Random Walk for GPU-Based Classical Planning(Extended Abstract)

○Ryo Kuroiwa[1], Alex Fukunaga[1] （1. the University of Tokyo）

9:00 AM - 10:40 AM

[3Pin1-14] Learning to resample for time-series generative model

○Takaaki Kaneko[1], Shohei Ohsawa[1], Yutaka Matsuo[1] （1. Department of Technology Management for Innovation, The University of Tokyo）

9:00 AM - 10:40 AM

[3Pin1-15] End-of-Query Detection in Text Chat System

○Shun Tanaka[1] （1. CyberAgent, Inc.）

9:00 AM - 10:40 AM

[3Pin1-16] Debating AI: Counterargument Generation Based on Textual Supportiveness Recognition

○Misa Sato[1], Kohsuke Yanai[1], Toshihiko Yanase[1], Yuta Koreeda[1], Kenzo Kurotsuchi[1] （1. Hitachi, Ltd.）

9:00 AM - 10:40 AM

[3Pin1-17] Associative Dialog System for Recipe

©The Japanese Society for Artificial Intelligence

Recommendation Using Food Knowledge

○Katsuyoshi Yamagami[1,2], Mitsuru Endo[2], Hirokazu Kiyomaru[3], Sadao Kurohashi[3] （1. Advanced Industrial Science and Technology (AIST), 2. Panasonic Corporation, 3. Kyoto University）

9:00 AM - 10:40 AM

[3Pin1-18] Analysis of work motivation by persuasion in dialogue

○Takuya Mizukami[1] （1. Waseda University）

9:00 AM - 10:40 AM

[3Pin1-19] Acquiring Utterance Candidates for Chat-oriented Dialogue Systems from Wikipedia Based on a State Space Model

○Tomoki Matsumoto[1], Masato Sakai[1], Mikio Nakano[2] （1. Nextremer Co., Ltd., 2. Honda Research Institute Japan Co., Ltd.）

9:00 AM - 10:40 AM

[3Pin1-20] Knowledge Graph Expansion for Constructing Dialogue Systems Responding to Queries about Unknown Attributes

○Yuma Fujioka[1], Kazunori Komatani[1] （1. The Institute of Scientific and Industrial Research, Osaka University）

9:00 AM - 10:40 AM

[3Pin1-21] Utterance understanding and emotion generation in a dialogue system for children

○Kento Yamamoto[1], Syunya Omura[2], Ryo Nemoto[2], Shiori Wada[1], Toru Sugimoto[2] （1. Graduate School of Engineering and Science,Shibaura Institute of Technology, 2. College of Engineering,Shibaura Institute of Technology）

9:00 AM - 10:40 AM

[3Pin1-22] Augmentation of Dialogue Database by Generating Interrogative Sentences Using Templates

○Yota Osanai[1], Tomoya Ogata[1], Mamoru Komachi[1], Eri Sato-Simokawara[1], Kazuyoshi Wada[1], Toru Yamaguchi[1], Tomoya Takatani[2] （1. Tokyo Metropolitan University, 2. TOYOTA MOTOR CORPORATION）

9:00 AM - 10:40 AM

[3Pin1-23] Utterance Intention Classification and Database Construction for Conversational News Delivery System

○Katsuya Yokoyama[1], Hiroaki Takatsu[1], Hiroshi Honda[2], Shinya Fujie[1,3], Yoshihiko Hayashi[1], Tetsunori Kobayashi[1] （1. Waseda University, 2. Honda R&D Co., Ltd, 3. Chiba Institute of Technology）

9:00 AM - 10:40 AM

[3Pin1-24] A proposal of automatic negotiation simulator Jupiter for negotiation agent using machine learning.

○Tomoya Fukui[1], Takayuki Ito[1] （1. Nagoya Institute of Technology）

9:00 AM - 10:40 AM

[3Pin1-25] An analysis of multilingual communication for solving global environmental issues

○Yuqi Liao[1], Bo Wang[1], Yuya Ieiri[1], Yuu Nakajima[2], Reiko Hishiyama[1] （1. Waseda University, 2. Toho University）

9:00 AM - 10:40 AM

[3Pin1-26] Premise generation based on argumentation schemes for dialogue systems

○Koh Mitsuda[1], Ryuichiro Higashinaka[1], Taichi Katayama[1], Junji Tomita[1] （1. NTT）

9:00 AM - 10:40 AM

[3Pin1-27] Metaphor creation using corpus and thesaurus

○Ryoga Sato[1], Toru Sugimoto[1] （1. Shibaura Institute of Technology）

9:00 AM - 10:40 AM

[3Pin1-28] Uncombined Elements in an Integrated Narrative Generation System and Expanding Conceptual Dictionaries by Attribute Information

○Jumpei Ono[1], Takuya Ito[2], Takashi Ogata[2] （1. Graduate School of Iwate Prefectural University, 2. Iwate Prefectural University）

9:00 AM - 10:40 AM

[3Pin1-29] Research on Sharing of Robotics Skills

○Guilherme de Campos Affonso[1], Kei Okada[1], Masayuki Inaba[1] （1. Univ. of Tokyo）

9:00 AM - 10:40 AM

[3Pin1-30] Hybrid Policy Gradient for Deep Reinforcement Learning

○praveen singh Thakur[1], Masaru Sogabe[1], Katsuyoshi Sakamoto[2], Koichi Yamaguchi[2], Dinesh Bahadur Malla[2], Shinji Yokogawa[2], Tomah Sogabe[1,2] （1. GRID Inc., 2. The University of Electro-communications）

9:00 AM - 10:40 AM

[3Pin1-31] A mobile robot based on real-time

©The Japanese Society for Artificial Intelligence

reinforcement learning using human evaluation

○Ken Yamane[1], Hiromasa Uetsuki, Hideyasu Yokomatsu[2] (1. Teikyo University, 2. Taito Corporation)

9:00 AM - 10:40 AM

[3Pin1-32] Study on evolutionary calculation by real world feedback from physical experiment determining the shape of an underwater robot.

○Ryota Nakamura[1], Satoru Inoue[1] (1. Saitama Institute of Technology)

9:00 AM - 10:40 AM

[3Pin1-33] Person Re-identification by Online Transfer Learning for Mobile Robot using Region-based CNN and Triplet Loss

○Yuki Murata[1], Masayasu Atsumi[1] (1. Soka University Graduate School of Engineering)

9:00 AM - 10:40 AM

[3Pin1-34] An AI game player trained by a genetic algorithm that avoids bombardments in a shooting game

○Shizuma Namekawa[1], Taro Tezuka[1] (1. University of Tsukuba)

9:00 AM - 10:40 AM

[3Pin1-35] Action acquisition by Memory Reinforcement Learning useing a prior knowledge

○Yuna Inamori[1], Tsubasa Hirakawa[1], Takayoshi Yamashita[1], Hironobu Fujiyoshi[1], Ryota Kashihawa[2], Masaki Inaba[2], Naoki Nitanda[2] (1. Chubu University, 2. DENSO CORPORATION)

9:00 AM - 10:40 AM

[3Pin1-36] Neural Headline Generation with Self-Training

Shintaro Takemae[2], Kazuma Murao[1], ○Taichi Yatsuka[1], Hayato Kobayashi[1,3], Masaki Noguchi[1], Hitoshi Nishikawa[2], Takenobu Tokunaga[2] (1. Yahoo Japan Corporation, 2. Tokyo Institute of Technology, 3. RIKEN AIP)

9:00 AM - 10:40 AM

[3Pin1-37] Security Reinforcement for ATM by Using Image Sensing

○Reiko Kishi[1], Trong Huy Phan[1], Kazuma Yamamoto[1], Makoto Masuda[1] (1. Corporate Research and Development Center, Oki Electric Industry Co., Ltd.)

9:00 AM - 10:40 AM

[3Pin1-38] Study on Image Recognition in Tank/Hold of Ship

○Masaru Hirakata[1], Ma Chong[1], Tomoyuki Taniguchi[1] (1. National Institute of Maritime, Port and Aviation Technology)

9:00 AM - 10:40 AM

[3Pin1-39] Prediction of Bascketball Free Throw Scoring by OpenPose

○Masato Nakai[1], Yoshihiko Tunoda[1], Caidong Sun[1], Hideki Murakoshi[1], Hisashi Hayashi[1], Tuyoshi Aziro[1] (1. Advanced Institute Industrial Technology)

9:00 AM - 10:40 AM

[3Pin1-40] BGM Recommendation System Considering the Atmosphere

○Kikue Sato[1], Shiori Sakai[1], Eichi Takaya[2], Kazuki Yamauchi[3], Hayato Oya[3], Satoshi Kurihara[2] (1. The University of Electro-Communications, 2. Keio University, 3. RecoChoku Co.,Ltd.)

9:00 AM - 10:40 AM

[3Pin1-41] Music Generation from Motion Data

Junta Suzuki[1], ○Syogo Watanabe[2], Kazuaki Rokusawa[1] (1. Department of Computer Science, Chiba Institute of Technology, 2. Graduate School of Information and Computer Science, Chiba Institute of Technology)

9:00 AM - 10:40 AM

[3Pin1-42] Acquisition of uniform impression of music using graph kernel filter

○Hisashi IKARI[1] (1. Sockets Co., Ltd.)

9:00 AM - 10:40 AM

[3Pin1-43] Feature Extraction and Anomaly Detection by Non-negative Matrix Factorization of Vibration Data of Machine

○Masatoshi Sekine[1], Kazuki Kobayashi[1], Satoshi Ikada[1] (1. Corporate Research and Development Center, Oki Electric Industry Co., Ltd.)

9:00 AM - 10:40 AM

[3Pin1-44] Anomaly Detection in Control Valves by 1d CNN-LSTM

○Takahiro Motegi[1], Yuya Nakazawa[1], Tetsuya Tabaru[1] (1. Azbil Corporation)

9:00 AM - 10:40 AM

[3Pin1-45] Constructing Diagnosis Corpus Using Collaborative Learning Environment for Improving Dementia Care

○Naoki Kamiya[1], Hideki Ueno[2,3,4], Takumi Yoshizawa[1], Shogo Ishikawa[1], Kenta Imada[1], Hayato

©The Japanese Society for Artificial Intelligence

Nishiyama[1], Shinya Kiriyama[1,4], Akira Tamai[3], Yoichi Takebayashi[1,4] (1. Shizuoka University, 2. Chiba University Hospital, 3. Tsuruga Onsen Hospital, 4. The Society of Citizen Informatics for Human Cognitive Disorder)
9:00 AM - 10:40 AM

[3Pin1-46] Construction and Practice of Collaborative Learning Environment for Personality-focused Care in Dementia Care

○Atsushi Omata[1], Shogo Ishikawa[1], Hatsue Munakata[1], Ayumi Nakanome[2], Mio Ito[3], Yutaka Sakane[4], Miwako Honda[5], Hisao Hara[2], Shinya Kiriyama[1,6], Yoichi Takebayashi[1,4,6] (1. Shizuoka University, 2. Koriyama Medical Care Hospital, 3. Tokyo Metropolitan Institute of Gerontology, 4. ExaWizards Inc., 5. Tokyo Medical Center, 6. The Society of Citizen Informatics for Human Cognitive Disorder)
9:00 AM - 10:40 AM

[3Pin1-47] Co-production of care knowledge for people with dementia based on an emotion thinking models

○Shogo Ishikawa[1], Yuki Sasaki[1], Shinya Kiriyama[1], Miwako Honda[2], Yves Gineste[3], Yoichi Takebayashi[1] (1. Shizuoka University, 2. Tokyo Medical Center, 3. Institute Gineste Marescotti)
9:00 AM - 10:40 AM

[3Pin1-48] Patient's sitting position estimation using skeleton information and bed position detection

○MADOKA INOUE[1,2], RYO TAGUCHI[2], TAIZO UMEZAKI[2,3] (1. AIPHONE CO., LTD., 2. Nagoya Institute of Technology, 3. The University of Tokyo)
9:00 AM - 10:40 AM

[3Pin1-49] Accurate drowsiness estimation via eye-related movements: a neural-network-based investigation

Mingfei Sun[2], ○Masanori Tsujikawa[1], Yoshifumi Onishi[1], Xiaojuan Ma[2], Atsushi Nishino[3], Satoshi Hashimoto[3] (1. NEC Corporation, 2. Hong Kong University of Science and Technology, 3. DAIKIN Industries, LTD)
9:00 AM - 10:40 AM

[3Pin1-50] Estimating nursing care necessity degree using by assessment information

○Shu Kobayashi[1], Sachiko Hirata-Mogi[1], Fukuda Shoji[1] (1. SMS CO., LTD.)
9:00 AM - 10:40 AM

[3Pin1-51] Attempt to visualize clinical data for grasping status of patients

○Masamichi Ishii[1], Kengo Miyo[1] (1. National Center for Global Health and Medicine)
9:00 AM - 10:40 AM

[3Pin1-52] Development of “Sleep Diary” application and continuation prediction of behavior for improvement sleeping habits

○Jou Akitomi[1], Ikuo Kajiyama[1], Miho Ishii[1], Isa Okajima[2], Mineko Yamaguchi[1] (1. NEC Solution Innovators, Ltd., 2. Tokyo Kasei University)
9:00 AM - 10:40 AM

[3Pin1-53] Simulation of Target Word Prediction for Neural Machine Translation

○Yuto Takebayashi[1], Chenhui Chu[3], Yuki Arase[1], Masaaki Nagata[2] (1. Graduate school of Information Science and Technology of Osaka University, 2. NTT Corporation, 3. Osaka University)
9:00 AM - 10:40 AM

[3Pin1-54] Research of abnormal node eradication conditions of self-repair network

○Yoshiyuki Tamura[1], Yoshiteru Ishida[1] (1. Toyohashi University of Technology)
9:00 AM - 10:40 AM

©The Japanese Society for Artificial Intelligence

Poster presentation | General Session | Interactive

[3Pin1] インタラクティブ(1)

Thu. Jun 7, 2018 9:00 AM - 10:40 AM Room P (4F Emerald Lobby)

[3Pin1-01] Optimization method for Educational Application of Structural Engineering

○Takuya Suzuki[1], Keiichi Iguchi[2], Takuya Kinoshita[1], Akihiro Ishikawa[2] （1. Takenaka Corporation, 2. HEROZ）

9:00 AM - 10:40 AM

[3Pin1-02] Benchmark Functions for Cooperative Coevolution and Analysis of Influence by Dynamic Problem Decomposition

○Yuki Nakashima[1], Jun-ichi Matsuoka[1], Asuka Hisatomi[1], Satoshi Ono[1] （1. Kagoshima University）

9:00 AM - 10:40 AM

[3Pin1-03] Incorporating Knowledge Graph Embeddings into Latent Variable Models

○Naoya Takeishi[1,2], Kosuke Akimoto[1] （1. The University of Tokyo, 2. RIKEN）

9:00 AM - 10:40 AM

[3Pin1-04] Covariate shift bias correction induced by action

○Ryota Higa[1] （1. NEC Corporation）

9:00 AM - 10:40 AM

[3Pin1-05] Batch Reinforcement Learning for Linearly Solvable MDP

○Tomoki Nishi[1], Keisuke Otaki[1], Takayoshi Yoshimura[1] （1. Toyota Central R&D Labs., Inc.）

9:00 AM - 10:40 AM

[3Pin1-06] Cappuccino: Bootstrapping from Structural Data

○Ryo Hanafusa[1], Masafumi Oyamada[2] （1. Kwansei Gakuin University, 2. NEC Corporation）

9:00 AM - 10:40 AM

[3Pin1-07] Non-negative matrix factorization with information expansion using item class information

○Katsuhiko Murakami[1] （1. The University of Tokyo）

9:00 AM - 10:40 AM

[3Pin1-08] Proposal of direct policy search using reduction of policy parameters by principal component analysis

○Yuuki Murata[1], Megumi Miyashita[1], Shiro Yano[1], Toshiyuki Kondo[1] （1. Tokyo University of Agriculture and Technology）

9:00 AM - 10:40 AM

[3Pin1-09] Curriculum Learning for Similar Learning Cases

○Yuta Ota[1], Yuta Taki[1] （1. KOZO KEIKAKU ENGINEERING Inc.）

9:00 AM - 10:40 AM

[3Pin1-10] Graph Feature Extraction Using Graph Fragment Decision Trees

○Haruki Sakagami[1], Ichigaku Takigawa[1], Hiroki Arimura[1] （1. Hokkaido University）

9:00 AM - 10:40 AM

[3Pin1-11] Objective Correction for Policy Improvement under Entropy Regularization

○Ryo Iwaki[1], Minoru Asada[1] （1. Osaka University）

©The Japanese Society for Artificial Intelligence

9:00 AM - 10:40 AM

[3Pin1-12] Experimental evaluation of Time-Series Gradient Boosting Tree with Time-Series Benchmark datasets

○Mitsuyoshi Imamura[1,2], Kei Nakagawa[1,2], Kenichi Yoshida[1] （1. University of Tsukuba, 2. Nomura Asset Management Ltd.）

9:00 AM - 10:40 AM

[3Pin1-13] Batch Random Walk for GPU-Based Classical Planning(Extended Abstract)

○Ryo Kuroiwa[1], Alex Fukunaga[1] （1. the University of Tokyo）

9:00 AM - 10:40 AM

[3Pin1-14] Learning to resample for time-series generative model

○Takaaki Kaneko[1], Shohei Ohsawa[1], Yutaka Matsuo[1] （1. Department of Technology Management for Innovation, The University of Tokyo）

9:00 AM - 10:40 AM

[3Pin1-15] End-of-Query Detection in Text Chat System

○Shun Tanaka[1] （1. CyberAgent, Inc.）

9:00 AM - 10:40 AM

[3Pin1-16] Debating AI: Counterargument Generation Based on Textual Supportiveness Recognition

○Misa Sato[1], Kohsuke Yanai[1], Toshihiko Yanase[1], Yuta Koreeda[1], Kenzo Kurotsuchi[1] （1. Hitachi, Ltd.）

9:00 AM - 10:40 AM

[3Pin1-17] Associative Dialog System for Recipe Recommendation Using Food Knowledge

○Katsuyoshi Yamagami[1,2], Mitsuru Endo[2], Hirokazu Kiyomaru[3], Sadao Kurohashi[3] （1. Advanced Industrial Science and Technology (AIST), 2. Panasonic Corporation, 3. Kyoto University）

9:00 AM - 10:40 AM

[3Pin1-18] Analysis of work motivation by persuasion in dialogue

○Takuya Mizukami[1] （1. Waseda University）

9:00 AM - 10:40 AM

[3Pin1-19] Acquiring Utterance Candidates for Chat-oriented Dialogue Systems from Wikipedia Based on a State Space Model

○Tomoki Matsumoto[1], Masato Sakai[1], Mikio Nakano[2] （1. Nextremer Co., Ltd., 2. Honda Research Institute Japan Co., Ltd.）

9:00 AM - 10:40 AM

[3Pin1-20] Knowledge Graph Expansion for Constructing Dialogue Systems Responding to Queries about Unknown Attributes

○Yuma Fujioka[1], Kazunori Komatani[1] （1. The Institute of Scientific and Industrial Research, Osaka University）

9:00 AM - 10:40 AM

[3Pin1-21] Utterance understanding and emotion generation in a dialogue system for children

○Kento Yamamoto[1], Syunya Omura[2], Ryo Nemoto[2], Shiori Wada[1], Toru Sugimoto[2] （1. Graduate School of Engineering and Science,Shibaura Institute of Technology, 2. College of

©The Japanese Society for Artificial Intelligence

Engineering,Shibaura Institute of Technology）
9:00 AM - 10:40 AM

[3Pin1-22] Augmentation of Dialogue Database by Generating Interrogative Sentences Using Templates

○Yota Osanai[1], Tomoya Ogata[1], Mamoru Komachi[1], Eri Sato-Simokawara[1], Kazuyoshi Wada[1], Toru Yamaguchi[1], Tomoya Takatani[2] （1. Tokyo Metropolitan University, 2. TOYOTA MOTOR CORPORATION）
9:00 AM - 10:40 AM

[3Pin1-23] Utterance Intention Classification and Database Construction for Conversational News Delivery System

○Katsuya Yokoyama[1], Hiroaki Takatsu[1], Hiroshi Honda[2], Shinya Fujie[1,3], Yoshihiko Hayashi[1], Tetsunori Kobayashi[1] （1. Waseda University, 2. Honda R&D Co., Ltd, 3. Chiba Institute of Technology）
9:00 AM - 10:40 AM

[3Pin1-24] A proposal of automatic negotiation simulator Jupiter for negotiation agent using machine learning.

○Tomoya Fukui[1], Takayuki Ito[1] （1. Nagoya Institute of Technology）
9:00 AM - 10:40 AM

[3Pin1-25] An analysis of multilingual communication for solving global environmental issues

○Yuqi Liao[1], Bo Wang[1], Yuya Ieiri[1], Yuu Nakajima[2], Reiko Hishiyama[1] （1. Waseda University, 2. Toho University）
9:00 AM - 10:40 AM

[3Pin1-26] Premise generation based on argumentation schemes for dialogue systems

○Koh Mitsuda[1], Ryuichiro Higashinaka[1], Taichi Katayama[1], Junji Tomita[1] （1. NTT）
9:00 AM - 10:40 AM

[3Pin1-27] Metaphor creation using corpus and thesaurus

○Ryoga Sato[1], Toru Sugimoto[1] （1. Shibaura Institute of Technology）
9:00 AM - 10:40 AM

[3Pin1-28] Uncombined Elements in an Integrated Narrative Generation System and Expanding Conceptual Dictionaries by Attribute Information

○Jumpei Ono[1], Takuya Ito[2], Takashi Ogata[2] （1. Graduate School of Iwate Prefectural University, 2. Iwate Prefectural University）
9:00 AM - 10:40 AM

[3Pin1-29] Research on Sharing of Robotics Skills

○Guilherme de Campos Affonso[1], Kei Okada[1], Masayuki Inaba[1] （1. Univ. of Tokyo）
9:00 AM - 10:40 AM

[3Pin1-30] Hybrid Policy Gradient for Deep Reinforcement Learning

○praveen singh Thakur[1], Masaru Sogabe[1], Katsuyoshi Sakamoto[2], Koichi Yamaguchi[2], Dinesh Bahadur Malla[2], Shinji Yokogawa[2], Tomah Sogabe[1,2] （1. GRID Inc., 2. The University of Electro-communications）
9:00 AM - 10:40 AM

[3Pin1-31] A mobile robot based on real-time reinforcement learning using human evaluation

©The Japanese Society for Artificial Intelligence

○Ken Yamane[1], Hiromasa Uetsuki, Hideyasu Yokomatsu[2] （1. Teikyo University, 2. Taito Corporation）

9:00 AM - 10:40 AM

[3Pin1-32] Study on evolutionary calculation by real world feedback from physical experiment determining the shape of an underwater robot.

○Ryota Nakamura[1], Satoru Inoue[1] （1. Saitama Institute of Technology）

9:00 AM - 10:40 AM

[3Pin1-33] Person Re-identification by Online Transfer Learning for Mobile Robot using Region-based CNN and Triplet Loss

○Yuki Murata[1], Masayasu Atsumi[1] （1. Soka University Graduate School of Engineering）

9:00 AM - 10:40 AM

[3Pin1-34] An AI game player trained by a genetic algorithm that avoids bombardments in a shooting game

○Shizuma Namekawa[1], Taro Tezuka[1] （1. University of Tsukuba）

9:00 AM - 10:40 AM

[3Pin1-35] Action acquisition by Memory Reinforcement Learning useing a prior knowledge

○Yuna Inamori[1], Tsubasa Hirakawa[1], Takayoshi Yamashita[1], Hironobu Fujiyoshi[1], Ryota Kashihawa[2], Masaki Inaba[2], Naoki Nitanda[2] （1. Chubu University, 2. DENSO CORPORATION）

9:00 AM - 10:40 AM

[3Pin1-36] Neural Headline Generation with Self-Training

Shintaro Takemae[2], Kazuma Murao[1], ○Taichi Yatsuka[1], Hayato Kobayashi[1,3], Masaki Noguchi[1], Hitoshi Nishikawa[2], Takenobu Tokunaga[2] （1. Yahoo Japan Corporation, 2. Tokyo Institute of Technology, 3. RIKEN AIP）

9:00 AM - 10:40 AM

[3Pin1-37] Security Reinforcement for ATM by Using Image Sensing

○Reiko Kishi[1], Trong Huy Phan[1], Kazuma Yamamoto[1], Makoto Masuda[1] （1. Corporate Research and Development Center, Oki Electric Industry Co., Ltd. ）

9:00 AM - 10:40 AM

[3Pin1-38] Study on Image Recognition in Tank/Hold of Ship

○Masaru Hirakata[1], Ma Chong[1], Tomoyuki Taniguchi[1] （1. National Institute of Maritime, Port and Aviation Technology）

9:00 AM - 10:40 AM

[3Pin1-39] Prediction of Bascketball Free Throw Scoring by OpenPose

○Masato Nakai[1], Yoshihiko Tunoda[1], Caidong Sun[1], Hideki Murakoshi[1], Hisashi Hayashi[1], Tuyoshi Aziro[1] （1. Advanced Institute Industrial Technology）

9:00 AM - 10:40 AM

[3Pin1-40] BGM Recommendation System Considering the Atmosphere

○Kikue Sato[1], Shiori Sakai[1], Eichi Takaya[2], Kazuki Yamauchi[3], Hayato Oya[3], Satoshi Kurihara[2] （1. The University of Electro-Communications, 2. Keio University, 3. RecoChoku Co.,Ltd.）

9:00 AM - 10:40 AM

[3Pin1-41] Music Generation from Motion Data

Junta Suzuki[1], ○Syogo Watanabe[2], Kazuaki Rokusawa[1] （1. Department of Computer Science, Chiba Institute of Technology, 2. Graduate School of Information and Computer Science,

©The Japanese Society for Artificial Intelligence

Chiba Institute of Technology）

9:00 AM - 10:40 AM

[3Pin1-42] Acquisition of uniform impression of music using graph kernel filter

○Hisashi IKARI[1] （1. Sockets Co., Ltd.）

9:00 AM - 10:40 AM

[3Pin1-43] Feature Extraction and Anomaly Detection by Non-negative Matrix Factorization of Vibration Data of Machine

○Masatoshi Sekine[1], Kazuki Kobayashi[1], Satoshi Ikada[1] （1. Corporate Research and Development Center, Oki Electric Industry Co., Ltd.）

9:00 AM - 10:40 AM

[3Pin1-44] Anomaly Detection in Control Valves by 1d CNN-LSTM

○Takahiro Motegi[1], Yuya Nakazawa[1], Tetsuya Tabaru[1] （1. Azbil Corporation）

9:00 AM - 10:40 AM

[3Pin1-45] Constructing Diagnosis Corpus Using Collaborative Learning Environment for Improving Dementia Care

○Naoki Kamiya[1], Hideki Ueno[2,3,4], Takumi Yoshizawa[1], Shogo Ishikawa[1], Kenta Imada[1], Hayato Nishiyama[1], Shinya Kiriyama[1,4], Akira Tamai[3], Yoichi Takebayashi[1,4] （1. Shizuoka University, 2. Chiba University Hospital, 3. Tsuruga Onsen Hospital, 4. The Society of Citizen Informatics for Human Cognitive Disorder）

9:00 AM - 10:40 AM

[3Pin1-46] Construction and Practice of Collaborative Learning Environment for Personality-focused Care in Dementia Care

○Atsushi Omata[1], Shogo Ishikawa[1], Hatsue Munakata[1], Ayumi Nakanome[2], Mio Ito[3], Yutaka Sakane[4], Miwako Honda[5], Hisao Hara[2], Shinya Kiriyama[1,6], Yoichi Takebayashi[1,4,6] （1. Shizuoka University, 2. Koriyama Medical Care Hospital, 3. Tokyo Metropolitan Institute of Gerontology, 4. ExaWizards Inc., 5. Tokyo Medical Center, 6. The Society of Citizen Informatics for Human Cognitive Disorder）

9:00 AM - 10:40 AM

[3Pin1-47] Co-production of care knowledge for people with dementia based on an emotion thinking models

○Shogo Ishikawa[1], Yuki Sasaki[1], Shinya Kiriyama[1], Miwako Honda[2], Yves Gineste[3], Yoichi Takebayashi[1] （1. Shizuoka University, 2. Tokyo Medical Center, 3. Institute Gineste Marescotti）

9:00 AM - 10:40 AM

[3Pin1-48] Patient's sitting position estimation using skeleton information and bed position detection

○MADOKA INOUE[1,2], RYO TAGUCHI[2], TAIZO UMEZAKI[2,3] （1. AIPHONE CO., LTD., 2. Nagoya Institute of Technology, 3. The University of Tokyo）

9:00 AM - 10:40 AM

[3Pin1-49] Accurate drowsiness estimation via eye-related movements: a neural-network-based investigation

Mingfei Sun[2], ○Masanori Tsujikawa[1], Yoshifumi Onishi[1], Xiaojuan Ma[2], Atsushi Nishino[3], Satoshi Hashimoto[3] （1. NEC Corporation, 2. Hong Kong University of Science and Technology, 3. DAIKIN Industries, LTD）

9:00 AM - 10:40 AM

©The Japanese Society for Artificial Intelligence

[3Pin1-50] Estimating nursing care necessity degree using by assessment information

○Shu Kobayashi[1], Sachiko Hirata-Mogi[1], Fukuda Shoji[1] （1. SMS CO., LTD.）

9:00 AM - 10:40 AM

[3Pin1-51] Attempt to visualize clinical data for grasping status of patients

○Masamichi Ishii[1], Kengo Miyo[1] （1. National Center for Global Health and Medicine）

9:00 AM - 10:40 AM

[3Pin1-52] Development of “Sleep Diary” application and continuation prediction of behavior for improvement sleeping habits

○Jou Akitomi[1], Ikuo Kajiyama[1], Miho Ishii[1], Isa Okajima[2], Mineko Yamaguchi[1] （1. NEC Solution Innovators, Ltd., 2. Tokyo Kasei University）

9:00 AM - 10:40 AM

[3Pin1-53] Simulation of Target Word Prediction for Neural Machine Translation

○Yuto Takebayashi[1], Chenhui Chu[3], Yuki Arase[1], Masaaki Nagata[2] （1. Graduate school of Information Science and Technology of Osaka University, 2. NTT Corporation, 3. Osaka University）

9:00 AM - 10:40 AM

[3Pin1-54] Research of abnormal node eradication conditions of self-repair network

○Yoshiyuki Tamura[1], Yoshiteru Ishida[1] （1. Toyohashi University of Technology）

9:00 AM - 10:40 AM

©The Japanese Society for Artificial Intelligence

建築構造教育アプリケーションに対する各種最適化手法の適用性検討

Optimization method for Educational Application of Structural Engineering

鈴木　琢也[*1]　Takuya Suzuki
井口　圭一[*2]　Keiichi Iguchi
木下　拓也[*1]　Takuya Kinoshita
石川　明洋[*2]　Akihiro Ishikawa

[*1] 竹中工務店　Takenaka Corporation
[*2] HEROZ　HEROZ

The purpose of this paper is to confirm an applicability of some optimization methods to an application for architectural structural education. First, the outline of the developed application is explained. This application takes advantage of the features of smart devices as to anyone from adults to children can enjoy it. Secondary, the detail of three optimization method is explained. And then, by each method, optimization analyses are conducted for 50 stages. Finally, the optimization results are compared, and the applicability of these optimization methods is examined.

1. はじめに

建築物の安全性を支える構造に関する技術の理解には，高度な専門的技術・知識を必要とする。そのため，一般の人々から理解されにくいものとなっている。しかし，技術者が構造解析の結果や意味をよりわかりやすく伝えることができれば，一般の人々はより正しく「安全」を理解し，今以上に「安心」を得ることができると考えられる。「いかにわかりやすく伝えるか」は，建築技術者にとって重要な課題といえる。

一方で，近年スマートフォンやタブレット端末といった高性能な携帯型情報端末（以下，「スマートデバイス」という）が広く普及している。スマートデバイスは，タッチパネルによる直感的なインターフェースや，各種センサーを内蔵することにより，これまでとは異なる計算機と人との関わり方を可能にしている。

このような背景から，著者の一人は，構造解析をよりわかりやすくするためのツールとしての可能性をスマートデバイスに見出し，これまでに，スマートデバイスの特徴を活かした建築構造解析アプリケーションに関する調査研究を行ってきた[鈴木 2017]。

本報では，新たにタッチで出来る構造最適化アプリケーションの対戦版を提案し，そこで利用する対戦コンピューターの戦略ロジックを決定するために実施した，各種最適化手法の適用性検討結果を報告する。

2. 開発するアプリケーションの概要

タッチで出来る構造最適化アプリケーションである StartOPT[鈴木 2017]を対戦型に改良したものである。

図 1 にはゲーム画面を示す。ボールを支えるのに不要と思われる部分をタッチで消した後，残ったブロックのみで変形解析が行われる。2 人のプレイヤーがブロック（最大 10 個，パスも可）を交互に消していき，先にボールを規定ラインよりも下に落としたプレイヤーが負けとなる。その点では，子どもの遊びである砂山崩しに近い。ただし，両者がパスした場合には，その時点で，より多くのブロックを消したプレイヤーの勝ちとなる。

様々な荷重条件，境界条件下における強い形を数多く試すことを通して，効率的な構造を直感的に理解するとともに，建築構造の基礎である「力の流れ」を，老若男女問わず楽しく理解することができる。

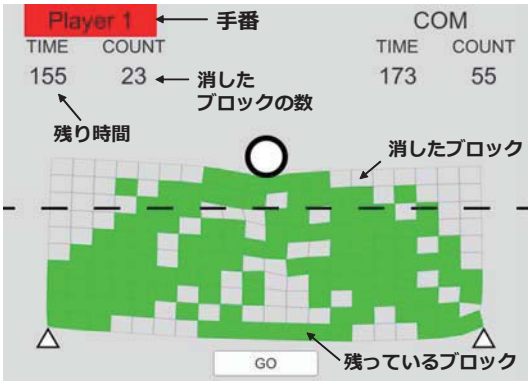


図 1　ゲーム画面

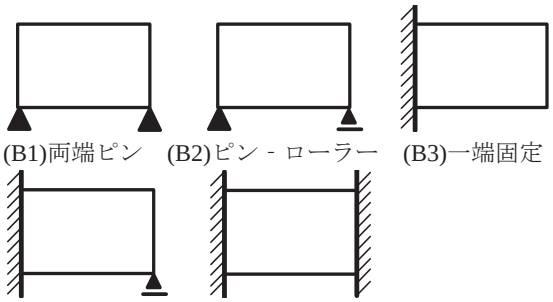


図 2　境界条件

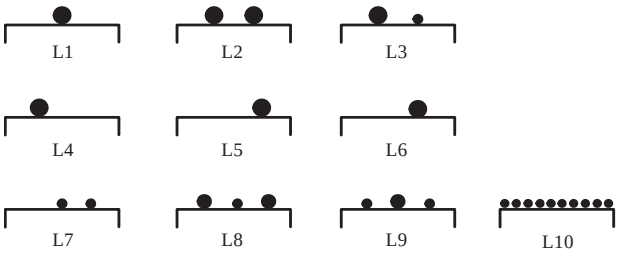


図 3　荷重条件

連絡先：鈴木琢也，竹中工務店，千葉県印西市大塚 1-5-1，0476-47-1700，0476-47-6460，suzuki.takuya@takenaka.co.jp

各ステージは，ブロックの境界条件（どこが支えられているか），および荷重条件（ボールの位置，個数，重さ）が変化する。

境界条件は，図 2 に示す建築構造で基本となる 5 パターン，荷重条件は図 3 に示す 10 パターン用意し，これらを組み合わせた計 50 種のステージが用意されている。これらの境界条件，荷重条件に応じて，最適な形状も様々に変化することになり，多くのステージをプレイすることで，プレイヤーの理解度も深まっていく。

3. 検討する最適化手法

前章で紹介したアプリケーションのコンピュータープレイヤーの構築には，ボールを規定ラインよりも落とさないための，最も有効なブロックを選択するロジックを作成する必要がある。その嚆矢として本報においては，「ボールを規定ラインよりも落とさない条件で，最も多くのブロックが消去された状態」を，3 つの最適化手法を用いて計算し，この構造最適化問題における適用性を検討する。比較する手法は表 1 に示す，貪欲法，反復局所探索法(Iterated Local Search (ILS)，密度法の 3 種の方法である。それぞれの詳細について以降で説明する。

貪欲法は，入力条件を変更させた結果，出力がどのように変化するかを確認した上で，消去するブロックを選択する。具体的には，ある状態において残っているブロックのうち，一つだけを抜いた場合の解析を，全ブロックに対して実施する。その後，最も影響の小さい（ボールの鉛直変位増分が小さい）ブロック 1 つを消去する。この操作をボールがクリアラインを越えるまで繰り返してブロックを一つずつ消去していき，消去できるブロックの最大個数を算出する。なお，計算回数の低減を図るために，感度計算の途中で「消去することでクリアラインを越えると判定されたブロック」は，以降の消去候補としないこととしている。

反復局所探索法(Iterated Local Search (ILS))は組合せ最適化問題を解くのに使われる手法で、その名の通り局所探索を繰り返す手法である。反復局所探索法では、まず、ブロックを一つずつ除いていく局所探索を行う。そして、どのブロックを除いてもボールがクリアラインを越えてしまう状態になれば、ランダムにいくつかのブロックを復活させて再び局所探索を行うという操作を制限時間まで何度も繰り返す。制限時間が来れば探索したすべての解の中で最も多くブロックを除いた解を出力して終了する。なお、復活させるブロックの数は削除済みブロックの半数以下とし、制限時間は 10 分とした。

密度法は、トポロジー最適化手法の一つで、材料のヤング率が密度のべき乗となるよう問題を緩和して最適解を求める手法である[Bensoe 2002]。本報では、支持構造全体の体積の制約条件の下で、静的コンプライアンス（外力仕事）を最小化するような要素密度分布を求める問題とした。このような問題設定を行うことで、目的関数及び制約条件の感度を一度の変形解析後の既知量から陽に算定することができ、計算効率が大幅に向上する。ここでは、密度のべき乗数 p=4 とし、体積制約をある程度大きい値から始めて、体積制約値を徐々に減らしながら複数回の最適化を行った。最適解の変形量が制限値を超えたとき、それまでに最もブロックを消したケースを出力して終了とした。なお、最適要素密度は実数で出力されるため、各最適化試行後に全要素をブロック有り／無しのいずれかに 2 値化させる処理を施している。

表 2　検討結果

STAGE		貪欲法	ILS	密度法
Boundary	Load			
B1	L1	165	178	178
	L2	160	170	170
	L3	156	165	160
	L4	148	170	165
	L5	148	170	164
	L6	140	165	160
	L7	142	159	158
	L8	145	158	160
	L9	143	159	160
	L10	140	158	150
B2	L1	129	132	132
	L2	146	152	150
	L3	119	143	138
	L4	134	142	137
	L5	128	141	135
	L6	132	141	138
	L7	120	139	134
	L8	123	134	137
	L9	117	128	136
	L10	133	132	134
B3	L1	165	158	165
	L2	132	127	128
	L3	148	145	146
	L4	186	181	189
	L5	83	81	83
	L6	111	107	108
	L7	125	120	119
	L8	113	107	106
	L9	125	124	114
	L10	98	102	91
B4	L1	143	171	179
	L2	154	163	167
	L3	161	162	167
	L4	182	180	185
	L5	146	161	166
	L6	153	155	165
	L7	148	158	159
	L8	149	156	161
	L9	152	154	161
	L10	157	154	156
B5	L1	153	162	179
	L2	172	164	178
	L3	146	164	179
	L4	150	170	188
	L5	150	171	188
	L6	156	163	179
	L7	148	156	162
	L8	156	158	168
	L9	141	156	168
	L10	152	152	158
合計		7123	7518	7658
Best Score の回数		9	16	30



(a)貪欲法



(b)ILS



(c)密度法

図 2　算出された最適解（B1，L1）



(a)貪欲法



(b)ILS



(c)密度法

図 3　算出された最適解（B1，L6）

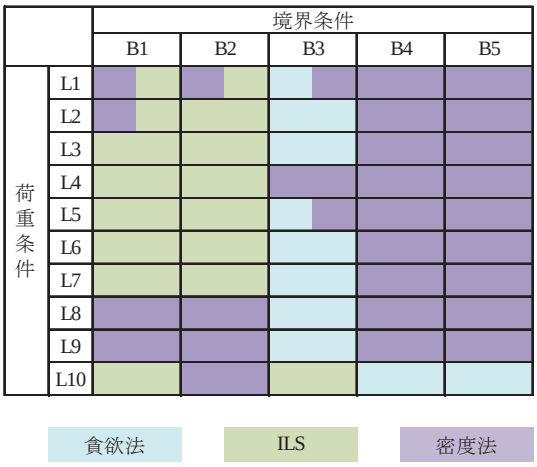


図 4　最小解を算出した最適化手法

4. 検討結果

表 2 には，各手法によって算出された消去できるブロックの最大の数の一覧を示す。大きな値が算出できるほど効率的な形を導出できたと判断する。各条件において最も多くのブロックを消すことができた手法は網掛け表示としている。

表より，境界条件，荷重条件によって，最適形状を算出できる最適化手法は異なることがわかる。三者による算定手法の差が最も小さいのは「B3, L5」のケース，最も大きいのは「B5, L4」および「B5, L5」のケースであった。

また，表には，全 50 ステージで算出された最大ブロック消去数の合計値，最大ブロック消去数を算出できたステージ数も併せて示す。表より平均的な観点では，密度法による方法が，合計値 7658，ステージ数 30 となり，最も有効であった。なお，ILS による方法も，合計値は 7518 と密度法との差は僅かであった。

図 2，図 3 には一例として，境界条件 B1（両端ピン）における最も基本的な荷重条件である L1 と，3 者の評価に最も差がついた L6 のケースにおいて，得られた最適形状の解析結果を示す。図 2 より，それぞれの最適化手法から算定された形は，概ね同傾向にあり，密度法と ILS による方法はほぼ同じ形を算出した。これらと比較すると，貪欲法によって得られた解には，まだ削ることができるブロックが残されていることがわかる。一方，図 3 においては，算出された形に差異がみられた。探索ルートが異なった結果，異なる最適化形状に至る一例と言える。

図 4 には，各ステージにおいて最適な解を算出できた手法を荷重条件，境界条件によって整理したものを示す。図より，条件によって最適形状を算出できる手法は様々に変化することがわかる。今回の検討範囲では，荷重条件よりは，境界条件よって最適形状を算出できる手法が変化する傾向があり，B1，B2 においては ILS，B3 においては貪欲法，B4，B5 においては密度法がより適切な解を算出する傾向が確認できる。

5. まとめ

本報では，タッチで出来る構造最適化アプリケーションの対戦版を提案し，そこで利用する対戦コンピューターの戦略ロジックとして，適切な最適化手法について検討を行った。

その結果，密度法による方法が，平均的に最適な形状を見つけやすい結果となった。ただし，荷重条件によっては他の方法の方が優れている場合もあるため，最適化手法を採用する際には，十分な事前検討が必要と考えられる。

参考文献

[鈴木　2017] 鈴木琢也ほか：スマートデバイスの特徴を活かした教育用構造解析アプリケーションの開発と展開，日本建築学会学術講演会梗概集，教育，pp. 47-48, 2017.7

[Bensoe　2002] Bensoe, M.P., Sigmund, O.: Topology Optimization; Theory, Methods, and Applications, Springer, 2002.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

協調型共進化の性能評価を目的としたベンチマーク問題と問題の再分割が性能に与える影響の検証

Benchmark Functions for Cooperative Coevolution and Analysis of Influence by Dynamic Problem Decomposition

中島 有貴 *1 Yuki Nakashima　松岡 淳一 *1 Jun-ichi Matsuoka　久冨 あすか *1 Asuka Hisatomi　小野 智司 *1 Satoshi Ono

*1鹿児島大学大学院 理工学研究科 情報生体システム工学専攻
Department of Information Science and Biomedical Engineering, Graduate School of Science and Engineering, Kagoshima University

This paper proposes benchmark functions for cooperative coevolution (CC), which dynamically divides a target problem into smaller subproblems and solves in parallel. The proposed benchmark functions allow adjusting dependency strength between variables by changing the number of linkages and linkage types. Experimental results showed the influence of dynamic decomposition change frequency to CC's optimization performance.

1. はじめに

近年，進化計算の分野では，大規模な最適化問題を部分問題へと分割し，それぞれを探索する集団を複数用意することで最適化を行う協調型共進化アルゴリズムの研究が行われている [Potter 94]．従来の進化計算が数十次元の問題を対象としていたのに対し，協調型共進化は数百から数千次元の問題を対象としており，部分問題へと分割しそれぞれを低次元の問題として最適化することで，良好な解を得ることが期待できる．

一方，協調型共進化では対象問題の変数間に依存性がある場合，問題分割により各ソルバが部分問題を跨ぐような変数間の依存性を考慮できなくなり局所解へ陥ることがある．しかし，分割の仕方を一定の頻度で変更し，部分問題内の変数を入れ替えることで局所解が発生するリスクを回避している．協調型共進化において分割の方法を変更可能である理由は，対象問題が大規模であるもののほぼ均質な性質を持つことを仮定しているためである．

本研究では，部分問題間の依存形状，および依存性の強さを調整可能なベンチマーク問題を提案する．本問題は，協調型共進化の性能評価を目的としており，協調型共進化における特有の局所解の発生の解析や回避方法の開発への寄与を目指す．本研究では，実験により，問題分割の頻度および問題の依存性の強さを変更することで協調型共進化の性能が変化することを確認する．

2. 関連研究 : 問題分割型共進化（DECC-G）

Wang らが設計した協調型共進化である DECC-G は，対象問題の変数をランダムに各部分問題へ割り当て，各部分問題の最適化アルゴリズムに SaNSDE（Self-adaptive DE with Neighborhood Search）を適用した協調型共進化アルゴリズムである [Yang 08]．SaNSDE ではスケール係数 F と交叉率 CR を更新しながら最適化を行う．各部分問題の最適化を T_g 世代の間行った後に，部分問題を結合する．各部分問題の最適化の後に最良個体，ランダムな個体，最悪個体を用いて重み付け用の個体群を DE で最適化を行い，最適化に用いた 3 個体を更新して個体群の評価を行う．ここまでを 1 サイクルとして，終了後，また対象問題の再分割を行い最適化を行う．この際，分割の方法を変更するため，各部分問題に含まれる変数は 1 サイクル時とは異なる．DECC-G はこのサイクルを繰り返すことで大規模な最適化問題を解くことが可能となる．

連絡先: 中島有貴，鹿児島大学大学院 理工学研究科 情報生体システム工学専攻，〒 890-0065 鹿児島市郡元 1-21-40，sc114035@ibe.kagoshima-u.ac.jp

表 1: Linkage form

Problems	Linkage form
Absolute	$L(x) = \|x\|$
Linear	$L(x) = x$
Nonlinear	$L(x) = \frac{sin(\pi\times\ x)}{2}$
	$L(x) = \frac{x}{\sqrt{\|x\|}}$
	$L(x) = x^2$
Deceitful	$L(x) = sin(\pi\times\ x)$

3. 共進化評価用ベンチマーク問題

本研究では，松岡らにより提案された共進化評価用のベンチマーク問題を改良した問題 [Mat 17] を提案する．本問題は連続 NK-Landscape 問題を参考に設計され，互いに依存関係のある 2 つの部分問題を持ち，さらに依存関係の強さを調整できるという特徴がある．

連続 NK-Landscape 問題は，実数値変数を含みつつ一般的な NK-Landscape と同様に設計変数間の依存性を制御できるよう設計された問題である [Kauffman 92, Wang 08]．連続 NK-Landscape 問題の評価関数を以下の式 (1) から式 (3) に示す．

$$f(\boldsymbol{x}) = \frac{1}{N}\Sigma_{j=1}^{N} f_j(x_j, z_1^{(j)}, ..., z_K^{(j)}) \tag{1}$$

$$f_j(x_j, z_1^{(j)}, ..., z_K^{(j)}) = \left| x_j - z_1^{(j)} - ... - z_K^{(j)} \right| \tag{2}$$

$$z_i^{(j)} = L(x_{(i+j)modN}) \tag{3}$$

ここで，$\boldsymbol{x} = \{x_1, x_2, ..., x_N\}$ は解候補を表すベクトルであり，$x_j \in [-1, 1]$ とする．また，式 (3) に含まれる関数 $L(x_i)$ は Linkage form と呼ばれ，表 1 に示す式のいずれかを用いる．$L(x_i)$ の定義により，連続 NK-Landscape 問題は Sphere 関数や Rosenbrock 関数などの一般的なベンチマーク問題を表現することができる．

共進化評価用のベンチマーク問題は，この連続 NK-Landscape 問題における解 $\boldsymbol{x}$ を $\boldsymbol{x}^{C_1}$ と $\boldsymbol{x}^{C_2}$ に分け，それぞ

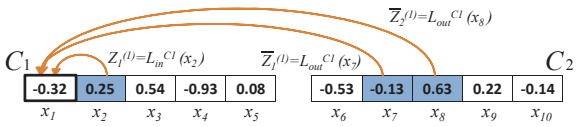
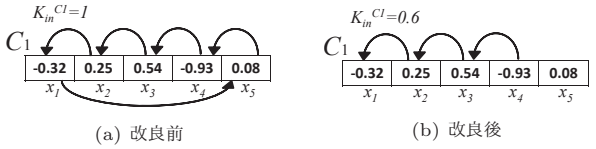


図 1: 共進化評価用ベンチマーク問題における依存関係の例



(a) 改良前　(b) 改良後

図 2: 共進化評価用ベンチマーク問題における $K_{in}^{C_1}$ の例

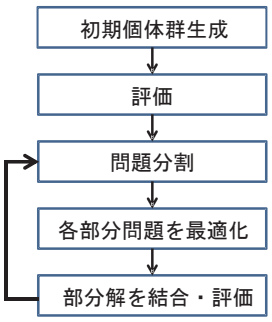


図 3: 実装した協調型共進化のフローチャート

れの解を持つ部分問題を定義する．また，それぞれ $\boldsymbol{x}^{C_1}$，$\boldsymbol{x}^{C_2}$ の解を持つ部分集団を C_1，C_2 とする．これらの部分問題は互いの解を組み合わせることで目的関数の値が計算できるものとする．

依存関係に関して，$\boldsymbol{x}^{C_1}$ と $\boldsymbol{x}^{C_2}$ は自身の解に含まれる変数のみから影響を受けるだけでなく，他の解に含まれる設計変数からも影響を受ける．ここで，自身の解からの影響を K_{in}，他の解からの影響を K_{out} とする．C_1 に関して自身の解からの影響の強さを $K_{in}^{C_1}$，他の解からの影響の強さを $K_{out}^{C_2}$ とし，C_2 に関して自身の解からの影響の強さを $K_{in}^{C_2}$，他の解からの影響の強さを $K_{out}^{C_2}$ とする．$K_{in}^{C_1}$，$K_{out}^{C_1}$，$K_{in}^{C_2}$，$K_{out}^{C_2}$ はいずれも $[0, N/2-1]$ の範囲の値を持つ．

目的関数を以下のように定義する．

$$f(\boldsymbol{x}) = \Sigma_{i=1}^{N}(\beta x_i)^2 + f^{C_1}(\boldsymbol{x}) + f^{C_2}(\boldsymbol{x}) \tag{4}$$

$$f^{C_1}(\boldsymbol{x}) = \frac{1}{N/2}\Sigma_{j=1}^{N/2} f_j(x_j, z_1^{(j)}, ..., z_{K_{in}^{C_1}}^{(j)}, \bar{z}_1^{(j)}, ..., \bar{z}_{K_{out}^{C_1}}^{(j)}) \tag{5}$$

$$f^{C_2}(\boldsymbol{x}) = \frac{1}{N/2}\Sigma_{j=\frac{N}{2}+1}^{N/2} f_j(x_j, z_1^{(j)}, ..., z_{K_{in}^{C_2}}^{(j)}, \bar{z}_1^{(j)}, ..., \bar{z}_{K_{out}^{C_2}}^{(j)}) \tag{6}$$

$$f_j(x_j, z_1^{(j)}, ..., z_{K_{in}^{C}}^{(j)}, \bar{z}_1^{(j)}, ..., \bar{z}_{K_{out}^{C}}^{(j)}) = \left| x_j, z_1^{(j)}, ..., z_{K_{in}^{C}}^{(j)}, \bar{z}_1^{(j)}, ..., \bar{z}_{K_{out}^{C}}^{(j)} \right| \tag{7}$$

$$z_i^{(j)} = L_{in}^{C}(x_{(i+j)}) mod \frac{N}{2} \tag{8}$$

$$\bar{z}_i^{(j)} = L_{out}^{C}(x_{\frac{N}{2}+(i+j)}) mod \frac{N}{2} \tag{9}$$

解 x^{C_1} の内部の依存関係は，j 番目の変数である x_j の値に着目した場合，$j+1$ 番目から $j+K_{in}$ 番目までの変数 $x_{j+1}, ..., x_{j+K_{in}}$ からの影響を受ける．また，相手の解である $\boldsymbol{x}^{C_2}$ における変数 $x_{N/2+j+1}, ..., x_{N/2+j+K_{out}}$ からも影響を受ける．図 1 に $K_{in}^{C_1}=1$，$K_{out}^{C_1}=2$ における依存関係を示す．図 1 では，x_1 に着目しており，自身の解の内部では x_2 から影響を受ける．また，他の解において x_7，x_8 からの影響を受けるようになっている．なお，式 (4) の第 1 項は大域的最適解の単一性を保障するための成分を示す．

本研究では，共進化評価用ベンチマーク問題について，より詳細な結果を得るために，$K_{in}^{C_1}$，$K_{out}^{C_1}$，$K_{in}^{C_2}$，$K_{out}^{C_2}$ をそれぞれ $[0,1/N]$ の範囲の実数値を持つように改良する．ここで N は各ソルバに含まれる変数の数を示す．共進化評価用ベンチマーク問題における $K_{in}^{C_1}$ にのみ着目した際の例を図 2 に示す．また，図 2(a) には改良前，図 2(b) には改良後の例を示している．改良前は，$K_{in}^{C_1}=1$ の場合，自身の解に含まれる全ての変数は自身の解の内部にある 1 つの変数から影響を受けるように設計されている．このため，変数ごとに影響を受ける数を変更することができない．改良後は $N=5$ の場合，$K_{in}^{C_1}$ は 0.2 刻みで増加していく．図 2(b) に示すように $K_{in}^{C_1}=0.6$ では，自身の解に含まれる 3 つの変数が依存性を受けるようになっている．このように設計することで，変数ごとに影響を受ける変数の数を変更することが可能となり，より詳細な結果を得ることができる．

4. 評価実験

本研究では Wang らが設計した DECC-G の問題分割方法にもとづく協調型共進化を実装し，実装した協調型共進化で共進化評価用ベンチマーク問題を解く．実装した協調型共進化のフローチャートを図 3 に示す．実験の際には，問題分割の頻度を 1 世代ごと，250 世代ごと，500 世代ごとの 3 パターンに変更し，頻度ごとの性能を調査する．本研究において，これらの分割の頻度に関して，それぞれ分割世代ごとに $T_g=1$，$T_g=250$，$T_g=500$ と定義する．

3 章で示した共進化評価用ベンチマーク問題は $K_{in}^{C_1}$，$K_{out}^{C_1}$，$K_{in}^{C_2}$，$K_{out}^{C_2}$ を変化させることが可能であるが，問題が複雑になることを避けるために $K_{in}^{C_1}$ および $K_{out}^{C_1}$ を変化させ，$K_{in}^{C_2}$，$K_{out}^{C_2}$ を 0 に固定し，C_1 にのみ依存性を持たせて実験を行う．問題全体の次元数は 10 次元とし，2 つの部分問題がそれぞれ 5 変数を含むように問題分割を行う．ここで協調型共進化が大規模な問題を対象にしているにも関わらず，問題の次元数を 10 次元としたのは，低次元の問題であっても依存性により最適解の発見が困難になることを示すためであり，また，解析を容易にするためである．また，問題全体の性質を均質にするために $L_{in}^{C_1}$ および $L_{out}^{C_1}$ には表 1 に示す”Absolute”の関数を割り当てることとする．なお，各部分問題の最適化アルゴリズムには jDE [Brest 06] を用いることとした．jDE の各パラメーターについて表 2 に示すように設定した．

実験結果を図 4 に示す．それぞれの図における z 軸は探索成功率 (Success Rate) を示し，これは 20 試行のうち，大域的最適解に収束した割合を示している．また，$T_g=1$ と $T_g=250$ の場合，$K_{in}^{C1}=0$ と $K_{in}^{C1}=1$ において成功率の差が激しかったため，K_{in}^{C1} を 0.2 刻みでで変化させて実験を行った．また，図 4(c) では K_{in}^{C1}，K_{out}^{C1} ともに 1 刻みで変化させている．

毎世代問題分割を変更する場合（$T_g=1$）は，$0.0 \leqq K_{in}^{C1} < 1.0$ では K_{out}^{C1} の値が大きくなるにつれて成功率は低下するこ

表 2: jDE の各パラメーター

項目	値
各ソルバ内で用いるアルゴリズム	jDE
個体数	50/集団
交叉方法	rand/1/exp
τ_1，τ_2	0.1，0.1
F_{max}，F_{min}	1.0，0.1
CR_{max}，CR_{min}	1.0，0.0

とが確認された．また，$K_{in}^{C1} = 1.0$ のとき K_{out}^{C1} の値に関わらず成功率は 0%となったが，$K_{in}^{C1} > 1.0$ では K_{out}^{C1} の値に関わらず成功率は 100%となった．$T_g = 250$ の場合は，$T_g = 1$ の場合と同様に $0.0 \leqq K_{in}^{C1} < 1.0$ のときに K_{out}^{C1} の値が大きくなるにつれて成功率は低下することが分かる．$K_{in}^{C1} = 1.0$ のとき，$T_g = 1$ の結果よりも成功率は上昇しており，K_{out}^{C1} の値が大きくなるにつれて成功率は低下することが確認された．$K_{in}^{C1} > 1.0$ では $T_g = 1$ の場合と同様の結果が得られた．$T_g = 500$ の場合は，$K_{in}^{C1} = 0$ のとき，$T_g = 1$，$T_g = 250$ の場合と同様に K_{out}^{C1} の値が大きくなるにつれて成功率は低下することが確認された．また，$K_{in}^{C1} > 0$ では K_{out}^{C1} の値に関わらず成功率は 100%となることが確認できた．

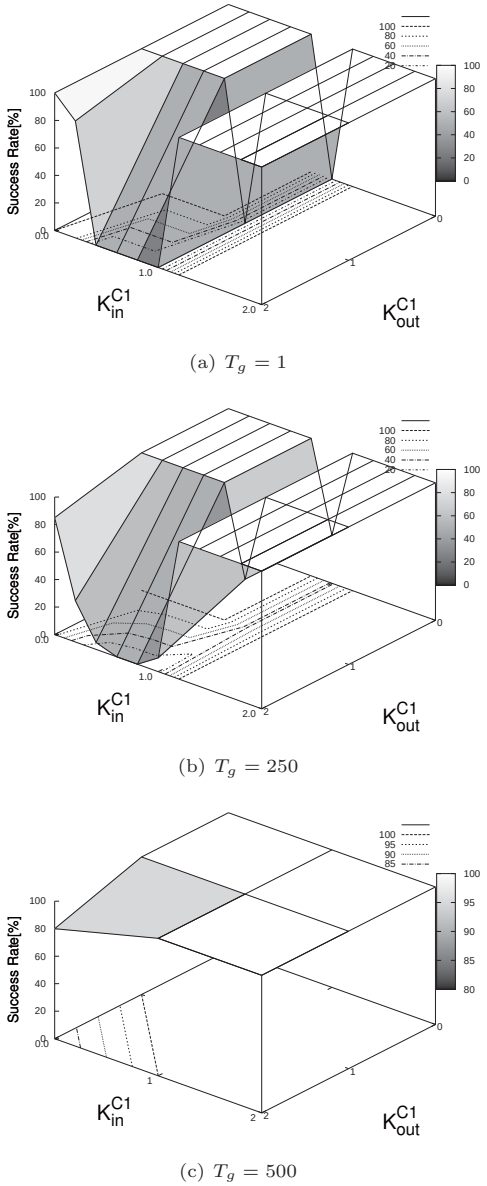
$K_{in}^{C1} = 1.0$ の場合の $T_g = 250$ についての 20 試行全てにおける最良個体の適応度の推移を図 5 に示す．$T_g \geqq 250$ の場合，図 5 のように，局所解に収束するものの世代を経て分割の仕方を変更することにより，局所解から脱出できる場合がある．これは，依存性のある変数同士を同じ部分問題に含みつつ最適化を行う期間が長いためであると考える．

$T_g = 1$ および $T_g = 250$ の場合，$K_{in}^{C1} = 1$ において成功率が極端に低下していたため，$K_{in}^{C1} = 1$，$K_{out}^{C1} = 0$ の問題で発生している局所解付近の適応度景観を作成し解析を行った．全ての次元数の組み合わせを掲載することは困難であるため，x_1 と $x_2, ..., x_{10}$ との組み合わせに着目した適応度景観のみを図 6 に掲載する．図 6 では，左側には 2 次元グラフで示した図，右側には 3 次元グラフで示した図を示す．また，2 次元グラフで示した図において，縦軸を x_1 としている．図 6 から，局所解付近が大域的に凸の形状になっている．他の変数の組み合わせにおいても同様に局所解付近では大域的に凸の形状をする適応度景観が確認された．このため，少なくとも 2 変数の同時変更だけでは局所解から抜け出すことが困難であることがわかる．

5. 結論

本研究では，DECC-G をもとに協調型共進化を実装し，共進化評価用ベンチマーク問題を解くことにより性能を調査した．その際，問題分割の頻度を 1 世代ごと，250 世代ごと，500 世代ごとに変化させ実験を行った．その結果，$K_{in}^{C_1} \leqq 1$ において，分割の頻度が高い場合は $K_{out}^{C_1}$ の値が大きくなるにつれて成功率が低下することが確認でき，分割の頻度が低い場合は，$0 < K_{in}^{C_1}$ において成功率が高くなることが確認できた．特に $K_{in}^{C_1} = 1$ のような難しい問題では，依存性のある変数同士を同じ部分問題に含みつつ最適化を行う期間を長くすることが，局所解から抜け出すために必要であることが分かった．

本稿では Linkage form に”Absolute”関数を用いた場合の実験を行ったため，今後は他の Linkage form ではどのような性能を示すか実験を行う必要がある．また，使用する最適化アルゴリズムに関して，本稿では jDE を用いて実験を行ったが，



(a) $T_g = 1$

(b) $T_g = 250$

(c) $T_g = 500$

図 4: 各分割頻度の結果

その他のアルゴリズムではどのような性能を示すか調査する必要がある．

謝辞

本研究の一部は科学研究費補助金（課題番号: 15H02758）による支援を受けて実施しました．ここに感謝の意を表します．

参考文献

[Brest 06] Brest, J., Greiner, S., Boskovic, B., Mernik, M., and Zumer, V.: Self-adapting control parameters in differential evolution: A comparative study on numerical benchmark problems, *IEEE transactions on evolutionary computation*, Vol. 10, No. 6, pp. 646–657 (2006)

[Kauffman 92] Kauffman, S. A.: The origins of order: Self-organization and selection in evolution, in *Spin glasses*

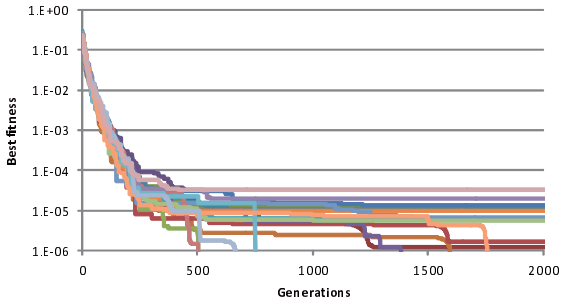


図 5: 全 20 試行における $T_g = 250$ の場合の BestFitness の推移

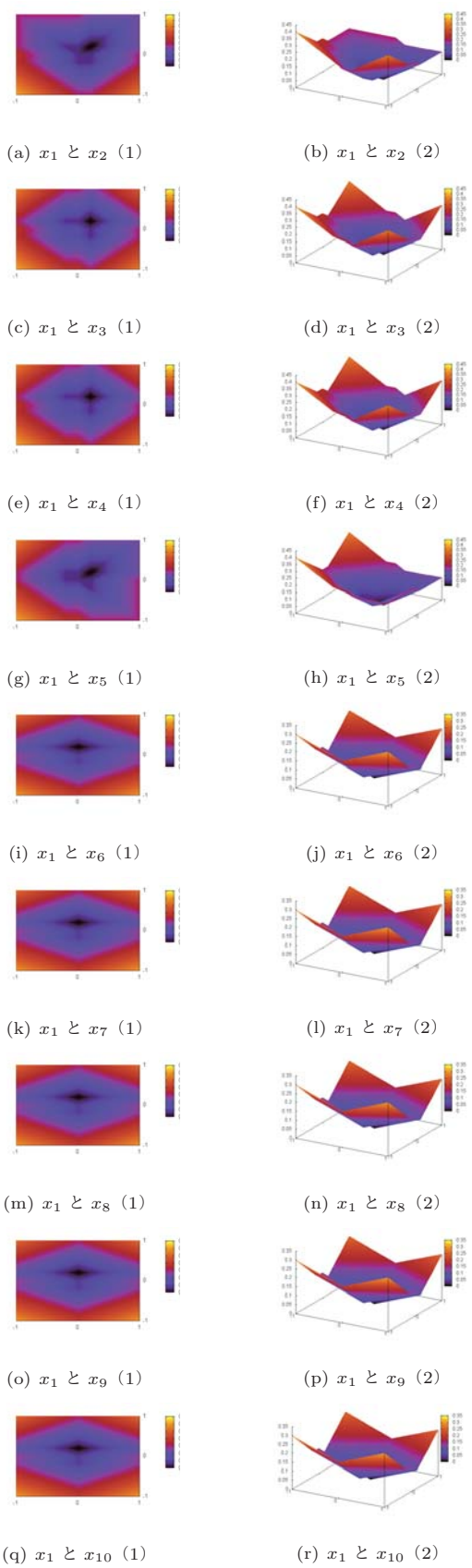
and biology, pp. 61–100, World Scientific (1992)

[Mat 17] *A Preliminary Study on Designing a Benchmark Problem for Analysis of Sparsely-Synchronized Heterogeneous Coevolution*IEEE (2017)

[Potter 94] Potter, M. A. and De Jong, K. A.: A cooperative coevolutionary approach to function optimization, in *International Conference on Parallel Problem Solving from Nature*, pp. 249–257Springer (1994)

[Wang 08] Wang, Y. and Li, B.: Understand behavior and performance of real coded optimization algorithms via nk-linkage model, in *Evolutionary Computation, 2008. CEC 2008.(IEEE World Congress on Computational Intelligence). IEEE Congress on*, pp. 801–808IEEE (2008)

[Yang 08] Yang, Z., Tang, K., and Yao, X.: Large scale evolutionary optimization using cooperative coevolution, *Information Sciences*, Vol. 178, No. 15, pp. 2985–2999 (2008)



(a) x_1 と x_2 (1) (b) x_1 と x_2 (2)

(c) x_1 と x_3 (1) (d) x_1 と x_3 (2)

(e) x_1 と x_4 (1) (f) x_1 と x_4 (2)

(g) x_1 と x_5 (1) (h) x_1 と x_5 (2)

(i) x_1 と x_6 (1) (j) x_1 と x_6 (2)

(k) x_1 と x_7 (1) (l) x_1 と x_7 (2)

(m) x_1 と x_8 (1) (n) x_1 と x_8 (2)

(o) x_1 と x_9 (1) (p) x_1 と x_9 (2)

(q) x_1 と x_{10} (1) (r) x_1 と x_{10} (2)

図 6: $K_{in}^{C_1} = 1$, $K_{out}^{C_1} = 0$ の局所解付近における適応度景観

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

潜在変数モデルとナレッジグラフ埋め込みの融合

Incorporating Knowledge Graph Embeddings into Latent Variable Models

武石 直也 *1　秋元 康佑 *1
Naoya Takeishi　Kosuke Akimoto

*1東京大学 航空宇宙工学専攻
Department of Aeronautics and Astronautics, The University of Tokyo

In this talk, we introduce an idea for incorporating information encoded in a knowledge graph to a latent variable model (LVM). We propose an extension of an LVM by two-view modeling, where the parameters and the latent variables of the LVM are shared between the original LVM and a probabilistic model for knowledge graph embedding. We specifically introduce how to incorporate a knowledge graph into probabilistic principal component analysis and show preliminary experimental results.

1. 背景

混合モデルや因子分析などの（特に、教師なしの）潜在変数モデル（latent variable model, LVM）は、次元削減やクラスタリングによる探索的データ分析をはじめとする種々のタスクに活用されている。LVM の潜在変数やパラメタはデータをもとに推定されるが、その推定量がどのような実際的・物理的な意味合いを持つかは多くの場合定かではない。例えば、主成分分析で得られた主成分方向の意味は必ずしも専門家から見てわかりよいとは限らない。無論それこそがデータ分析の意義のひとつといえる（常に既存の知識で説明できてしまう結果が出ていては新しい知見はない）が、一方で専門家の知識を LVM に事前知識として「入れ込む」ことができれば、データと既存知識との整合性／非整合性を定量的にとらえるのに役立つ可能性がある。一般的な LVM のモデル設計や推論・学習の過程においても、特徴量、仮説空間、正則化、初期値などの選択において事前知識といえるものが役立っている。また、データドメインに関してより直接的な形の事前知識として制約や論理式を LVM で考慮する研究もある [Fung 03, Varol 12, Andrzejewski 11, Mei 14, Foulds 15]。

本稿では、論理式などよりは間接的な形の事前知識として、**ナレッジグラフ**の形で表された知識を LVM に利用するための方法を提案する。ここでいうナレッジグラフとは、（通常、真の）タプル (h, r, t) の集合である。あるタプル (h, r, t) について、例えば

$$h = \text{Kagoshima}, \quad r = \text{is-located-in}, \quad t = \text{Japan},$$

であれば、そのタプルは真である。h や t はエンティティ、r はリレーションと呼ばれる。

ナレッジグラフは離散的なデータであるが、近年はエンティティやリレーションを連続値で表現するナレッジグラフ埋め込み（knowledge graph embedding, KGE）が盛んに研究され、リンク予測等のナレッジグラフ上のタスクのほか、question answering 等のナレッジグラフを利用するタスクに活用されている。よく知られた KGE として TransE [Bordes 13] とよばれる方法や DistMult [Yang 15] とよばれる方法がある。より複雑な最近の方法としては ConvE [Dettmers 17] などがある。本稿で紹介するのは、LVM にナレッジグラフによる知識を入れ込むために KGE を活用するという発想である。

連絡先: {takeishi,akimoto}@ailab.t.u-tokyo.ac.jp

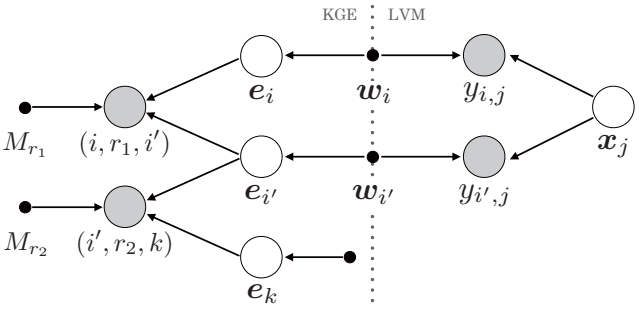


図 1: 提案手法のグラフィカルモデル表示の一部

2. 提案手法

まず、LVM および KGE 一般についてそれらを融合する方法論を紹介する。その後、確率的 PCA [Tipping 99] および KGE の一手法である DistMult [Yang 15] を組み合わせる場合についてのモデルや推論を特に記述する。

2.1 表記法

データは n 個のオブジェクトからなり、各々のオブジェクトは m 属性の観測からなるベクトル $\boldsymbol{y}_j \in \mathbb{R}^m$ である（ただし $j = 1, \ldots, n$）。$\boldsymbol{y}_j$ の i 番目の要素、すなわち属性 i のオブジェクト j での観測値を $y_{i,j} \in \mathbb{R}$ と表す（ただし $i = 1, \ldots, m$）。データ行列を $\boldsymbol{Y} = [\boldsymbol{y}_1 \ \cdots \ \boldsymbol{y}_n] \in \mathbb{R}^{m \times n}$ とする。

ある LVM が、パラメタの集合 θ および n 個の潜在変数ベクトル $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ を持っているとする（$\boldsymbol{x}_j \in \mathbb{R}^d$ は $\boldsymbol{y}_j$ に対応する）。$\theta = \{\boldsymbol{\pi}, \{\boldsymbol{w}\}\}$ の元には大域的パラメタと局所的パラメタの二種類があって、大域的パラメタ $\boldsymbol{\pi}$ が属性インデクス i によらないのに対して、m 個ある局所的パラメタ $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\} \in \{\mathbb{R}^d\}^m$ は属性毎に存在するとする。*1 例えば確率的 PCA では、$\boldsymbol{\pi}$ には観測ノイズ分散などがあたり、$\boldsymbol{w}_i$ には因子負荷行列の i 行目があたる。また、LVM には観測モデル $p(\boldsymbol{Y} \mid \boldsymbol{X}, \theta)$ および事前分布 $p(\boldsymbol{X})$ が定まっているとする。パラメタの事前分布 $p(\theta)$ があってもよい。

KGE に関する量は次のように表す。エンティティの埋め込みは任意の添字とともに $\boldsymbol{e}_\cdot \in \mathbb{R}^q$ で表す。埋め込みのスコア付けのために用いるリレーション r に対応するパラメタは一般に

*1　$\boldsymbol{x}$ と $\boldsymbol{w}$ の次元は異なってもよいが、簡単のため両方 d とする。

M_r で表し、KGE のモデルによって行列 $\boldsymbol{M}_r$ やベクトル $\boldsymbol{m}_r$ として取られる。タプル (h,r,t) のスコアは $\psi(\boldsymbol{e}_h,\boldsymbol{e}_t,M_r)$ と書くことにする。また、タプルの負例は $(h,r,\tilde{t})$ などと表す。この場合は、タプル (h,r,t) に対して t だけを取り替えて作った負例という意味である。

2.2 LVM と KGE の two-view モデル

提案手法の主要な考え方を表すため、図 1 にグラフィカルモデルを示す。同図点線より右側は、ある LVM の観測モデルの一部（オブジェクト j と属性 i,i' に対応する部分）、すなわち

$$p(y_{i,j} \mid \boldsymbol{x}_j,\boldsymbol{w}_i)\cdot p(y_{i',j} \mid \boldsymbol{x}_j,\boldsymbol{w}_{i'}), \tag{1}$$

を表す。[*2] 同図点線より左側は KGE の一部を確率モデル化したものとその事前分布に対応していて、次の確率分布を表す：

$$\begin{gathered} p(\boldsymbol{e}_i \mid \boldsymbol{w}_i),\quad p(\boldsymbol{e}_{i'} \mid \boldsymbol{w}_{i'}),\quad p(\boldsymbol{e}_k),\\ p((i,r_1,i') \mid \boldsymbol{e}_i,\boldsymbol{e}_{i'},M_{r_1}),\quad p((i',r_2,k) \mid \boldsymbol{e}_{i'},\boldsymbol{e}_k,M_{r_2}). \end{gathered} \tag{2}$$

ここで、r_1 は属性 i と属性 i' との間のリレーション、r_2 は属性 i' とあるナレッジグラフ上のエンティティ k との間のリレーションを表す。このような two-view モデリングでは、LVM の局所的パラメタ $\boldsymbol{w}$ がエンティティの埋め込み $\boldsymbol{e}$ と関連づいている。このモデルの動機は、$\boldsymbol{w}$ を（$\boldsymbol{e}$ を通じて）ナレッジグラフのモデリングにも用いることによって、LVM の事前分布 $p(\boldsymbol{w})$ だけでなくナレッジグラフの構造も考慮した形で $\boldsymbol{w}$ の推定値や事後分布を得ることにある。

さて、式 (1) はオリジナルの LVM で定まっている一方、式 (2) の各分布は都合よく定めなくてはいけない。本稿では、

$$p(\boldsymbol{e} \mid \boldsymbol{f}(\boldsymbol{w})) = \mathcal{N}_{\boldsymbol{e}}(\boldsymbol{f}(\boldsymbol{w}),\boldsymbol{V}), \tag{3}$$

$$p(\boldsymbol{e}) = \mathcal{N}_{\boldsymbol{e}}(\boldsymbol{0},\boldsymbol{I}), \tag{4}$$

$$p((h,r,t)=\mathrm{true} \mid \boldsymbol{e}_h,\boldsymbol{e}_t,M_r) = \sigma(\psi(\boldsymbol{e}_h,\boldsymbol{e}_t,M_r)), \tag{5}$$

のようにモデリングすることを提案する。ただし、$\boldsymbol{f}:\mathbb{R}^d\to\mathbb{R}^q$ はユーザが設計すべき関数、σ はシグモイド関数である。また、$\boldsymbol{V}$ は大域的パラメタで $d\times d$ の実対称行列である。

推論や学習は基本的に通常の LVM と同様の手続きで行えばよい。ただし、KGE 部分の学習のためには予めナレッジグラフ上の負例を作成しておく前処理が重要となる（KGE に関して確率モデルを考えている先行研究だと [He 15, Xiao 16, Zhang 16] などを参照するとよい）。これは KGE 学習では一般的に行われている処理で、一種の data augmentation といえる。負例は正例タプルの一部をランダムサンプリングで変更して作成する（例えば、(h,r,t) から $(h,r,\tilde{t})$ を作る）ことが多いが、その詳細な方法は KGE のスコア関数設計と同様に本稿の対象の外となる。また、$\boldsymbol{e}$ の事後分布は解析的には計算できないので、サンプリングや近似推論が必要となる。しかし本稿では、ひとまず式 (3) のかわりに

$$p(\boldsymbol{e} \mid \boldsymbol{f}(\boldsymbol{w})) = \delta(\boldsymbol{e}-\boldsymbol{f}(\boldsymbol{w})), \tag{3'}$$

として（すなわち $\boldsymbol{e}=\boldsymbol{f}(\boldsymbol{w})$ として）、属性に対応しないエンティティについては式 (4) の事前分布は考えず、$\boldsymbol{e}$ をパラメタとして点推定する妥協策を用いる。$\boldsymbol{e}$ を確率変数としてみなすことでデータやナレッジグラフの不確かさを考慮することができるが、今後の課題とする。

*2 簡単のため j について独立な観測モデルであるとし、局所的パラメタの事前分布は考えず、図中では大域的パラメタを省略している。

2.3 ナレッジグラフ＋PCA

前節では LVM や KGE 一般について、それらを融合させるための two-view modeling に基づく方法論を提案した。この節では、そのような two-view modeling の例として、基本的な LVM のひとつである確率的 PCA [Tipping 99] および KGE のベースラインとしてよく参照される DistMult [Yang 15] を融合したモデルを定義し、学習方法を示す。

まず、確率的 PCA の観測モデルと事前分布は次の通り：

$$p(\boldsymbol{y}_j \mid \boldsymbol{x}_j,\theta) = \prod_{i=1}^{m}\mathcal{N}_{y_{i,j}}(\boldsymbol{w}_i^\top\boldsymbol{x}_j+\boldsymbol{\mu},\sigma^2), \tag{6}$$

$$p(\boldsymbol{x}_j) = \mathcal{N}_{\boldsymbol{x}_j}(\boldsymbol{0},\boldsymbol{I}). \tag{7}$$

ただし、$\boldsymbol{\mu}$, σ^2 は大域的パラメタである。次に、DistMult [Yang 15] のスコア関数は次のように定義される：

$$\psi(\boldsymbol{e}_h,\boldsymbol{e}_t,\boldsymbol{m}_r) = \boldsymbol{e}_h^\top\,\mathrm{diag}(\boldsymbol{m}_r)\boldsymbol{e}_t, \tag{8}$$

ただし、$\boldsymbol{m}_r\in\mathbb{R}^q$ はリレーション r に対応するパラメタである。$\boldsymbol{e}$ と $\boldsymbol{w}$ を関連させる関数 $\boldsymbol{f}$ の形は簡単にアフィン変換とする。すなわち、大域的パラメタ $\boldsymbol{A}\in\mathbb{R}^{q\times d}$, $\boldsymbol{b}\in\mathbb{R}^q$ を用いて

$$\boldsymbol{f}(\boldsymbol{w}) = \boldsymbol{A}\boldsymbol{w}+\boldsymbol{b}. \tag{9}$$

なお、$\boldsymbol{f}$ の設計には試行錯誤の余地が大いにある。確率的 PCA より複雑なモデルでは、j によって異なる $\boldsymbol{f}$ が必要な場合もある。

前述のように、$\boldsymbol{e}$ の事後分布推論は諦めて（というより、$\boldsymbol{e}$ を確率変数としないで）、$\boldsymbol{x}$ だけを潜在変数と見なす妥協策を採用する。このとき、学習時に最大化すべき不完全データ対数尤度は次の通り：

$$\begin{aligned} \mathcal{L} = &\sum_j \ln p(\boldsymbol{y}_j \mid \boldsymbol{W})\\ &+\sum_{(h,r,t)\in\mathcal{T}_1}\ln p((h,r,t) \mid \boldsymbol{e}_h,\boldsymbol{e}_t,\boldsymbol{m}_r)\\ &+\sum_{(i,r,t)\in\mathcal{T}_2}\ln p((i,r,t) \mid \boldsymbol{f}(\boldsymbol{w}_i),\boldsymbol{e}_t,\boldsymbol{m}_r)\\ &+\sum_{(h,r,i)\in\mathcal{T}_3}\ln p((h,r,i) \mid \boldsymbol{e}_h,\boldsymbol{f}(\boldsymbol{w}_i),\boldsymbol{m}_r)\\ &+\sum_{(i,r,i')\in\mathcal{T}_4}\ln p((i,r,i') \mid \boldsymbol{f}(\boldsymbol{w}_i),\boldsymbol{f}(\boldsymbol{w}_{i'}),\boldsymbol{m}_r). \end{aligned} \tag{10}$$

$\boldsymbol{x}$ は解析的に周辺化できるので、式 (10) の第一項は

$$\sum_j \ln\mathcal{N}_{\boldsymbol{y}_j}\left(\boldsymbol{\mu},\boldsymbol{W}\boldsymbol{W}^\top+\sigma^2\boldsymbol{I}\right),$$

と書ける。第二項以降は、例えば

$$\begin{aligned} \sum_{(i,r,t)\in\mathcal{T}_2}\Big\{&z_{(i,r,t)}\ln\left[\sigma(\psi(\boldsymbol{f}(\boldsymbol{w}_i),\boldsymbol{e}_t,\boldsymbol{m}_r))\right]\\ &+\left(1-z_{(i,r,t)}\right)\ln\left[1-\sigma(\psi(\boldsymbol{f}(\boldsymbol{w}_i),\boldsymbol{e}_t,\boldsymbol{m}_r))\right]\Big\}, \end{aligned}$$

のように書ける。ただし、$\mathcal{T}_1$ は head と tail の両者がデータの属性に関係しないタプルの集合、$\mathcal{T}_2$ は head だけがデータの属性に対応しているタプルの集合、以下同様に定義する。

3. 数値例

3.1 トイデータの用意と実験手順

次のように三次元の連続値データを生成した。まず、$\bar{y}_{1,1:n}$ および $\bar{y}_{2,1:n}$（ただし n は平方数）を $[-0.5, 1.5]^2$ 内の $\sqrt{n} \times \sqrt{n}$ 格子上の点としてとる。これをもとに $\bar{y}_{3,1:n}$ を次のように生成する：

$$\bar{y}_{3,j} = \bar{y}_{1,j}^3 + \bar{y}_{2,j}^3, \quad j = 1, \ldots, n. \tag{11}$$

このように生成した $\bar{y}$ に正規分布に従うノイズを加えてデータ $\boldsymbol{y}$ をつくる。

$$\boldsymbol{y}_j = \begin{bmatrix} \bar{y}_{1,j} \\ \bar{y}_{2,j} \\ \bar{y}_{3,j} \end{bmatrix} + \boldsymbol{v}_j, \quad \boldsymbol{v}_j \sim \mathcal{N}_{\boldsymbol{v}_j}(\boldsymbol{0}, 0.01\boldsymbol{I}) \tag{12}$$

また、利用するナレッジグラフとして単に

$$\mathcal{T} = \{(h = w_1, r = r_1, t = e_1), (w_2, r_1, e_1)\} \tag{13}$$

を用意する。w_1, w_2 はそれぞれデータの 1 次元目、2 次元目の特徴に対応するエンティティ、e_1 は何か任意のエンティティである。

以上をもとに、$q = 5, d = 2$ として次のような実験を行った。

1. 上記の手順でデータを生成する（$n = 400$）。
2. 訓練・検証用およびテスト用にデータを分割する。このとき、訓練・検証用データは範囲 $[-0.5, 1.5] \times [-0.5, 0.5]$ の 200 点、テスト用データは $[-0.5, 1.5] \times [0.5, 1.5]$ の 200 点を使用する（つまり、訓練時とテスト時で共変量シフトのようなことが起こっている設定）。訓練・検証用データは更にランダムに分割し、訓練用データ $\boldsymbol{Y}_{\mathrm{tr}} \in \mathbb{R}^{3\times 160}$、検証用データ $\boldsymbol{Y}_{\mathrm{val}} \in \mathbb{R}^{3\times 40}$ を用意する。テスト用データを $\boldsymbol{Y}_{\mathrm{te}} \in \mathbb{R}^{3\times 200}$ と表す。
3. $\boldsymbol{Y}_{\mathrm{tr}}$ から確率的 PCA のパラメタを学習する。このとき、次の三つの方法を試す：
   (a) データ共分散行列の固有値分解に基づく最尤推定 [Tipping 99]（$\mathcal{T}$ は使わない）
   (b) 勾配法による $-\mathcal{L}$ の最小化（$\mathcal{T}$ は使わない）
   (c) 勾配法による $-\mathcal{L}$ の最小化（$\mathcal{T}$ を使う）

   ただし、(b) および (c) では (a) の解を初期値とする。(c) で $\boldsymbol{e}$ の初期値はランダムに得る。さらに、(b) および (c) では勾配法の各ステップで $\boldsymbol{Y}_{\mathrm{val}}$ に対する再構成誤差を監視し、最小の検証再構成誤差を達成するパラメタを採用する。
4. 上記 (a)–(c) それぞれについて、$\boldsymbol{Y}_{\mathrm{te}}$ に対してテスト再構成誤差 $\boldsymbol{Y}_{\mathrm{te}}$ を計算して比較する。
5. 上記 1–3 を異なる乱数シードのもとで 100 回繰り返す。

3.2 実験結果

まず、方法 (b) および (c) による再構成誤差が方法 (a) によるものより明らかに小さくなったため、方法 (a) による再構成誤差については省略する。方法 (b) および (c) を比較するため、図 2 に、方法 (b) によるテストデータに対する平均再構成誤差から方法 (c) によるテストデータに対する平均再構成誤差を引いた値のヒストグラムを示す（なお、差をとる前の元の再構成誤差は 1.0 程度というオーダーである）。このとこから、ほとんどの場合にナレッジグラフを用いた方が用いない場合よりも再構成誤差が 10%前後小さくなることがわかる。参考までに、図 3 にある試行における訓練・検証用データとテスト用データ、および学習により得られた $\boldsymbol{W}$ に対応する平面をナレッジグラフを利用しない場合と利用する場合についてそれぞれ示す。

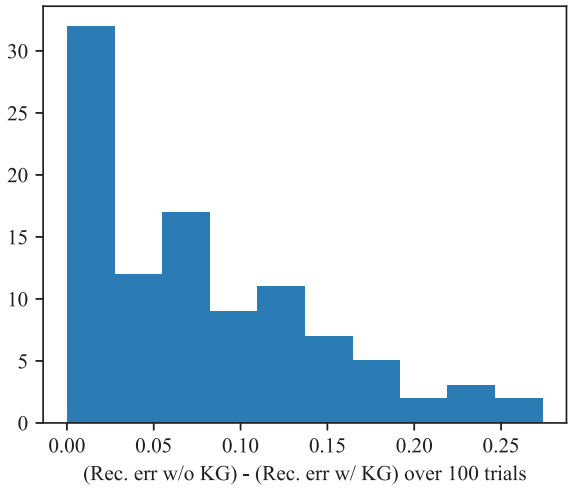


図 2: ナレッジグラフを使用した場合と使用しなかった場合のテスト再構成誤差の差に関するヒストグラム（正の値はナレッジグラフを使用した場合のテスト再構成誤差の方が使用しなかった場合より小さいことを示す）

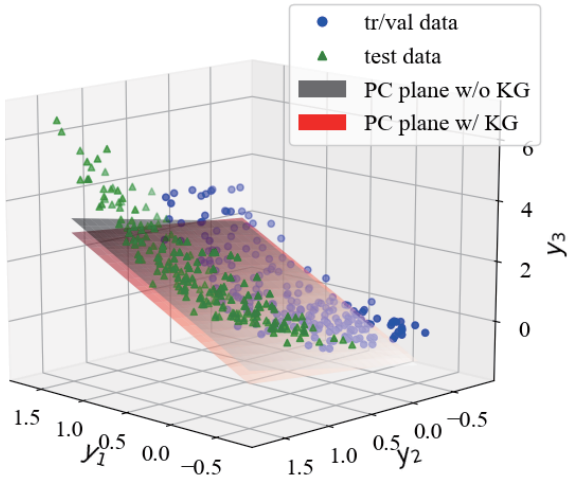


図 3: ある試行における訓練・検証データとテストデータ、および得られた $\boldsymbol{W}$ の示す平面

4. 今後の課題

本稿で紹介した方法では、$\boldsymbol{e}$ の事後分布推論についての方略が未完成である。局所的パラメタ $\boldsymbol{w}$ を点推定することにして、各局所的パラメタに対応するようなエンティティとナレッジグラフ中のエンティティとの関係のみを考慮するなら本稿で紹介した学習方法で十分であるかもしれないが、例えば各オブジェクトとナレッジグラフのエンティティとの関係も考慮したい際には、（オブジェクトに対応する $\boldsymbol{x}$ を確率変数として事後分布推論したい限りは）$\boldsymbol{e}$ の事後分布推論を考えなくてはな

らない。これについては、Pólya–Gamma data augmentation [Polson 13] による Gibbs サンプリングや、または確率的勾配変分ベイズ [Kingma 14] のような方法を検討したい。

また、当然実験も行う必要がある。まず、LVM 側の性能向上を示すという観点では、本稿で紹介したように PCA の再構成誤差が小さくなるなどの汎化性能向上を示したい。なお、本稿で作ったトイデータは結局線形モデルから生成されているために、ナレッジグラフの情報を用いない場合とあまり差が出なかったと考えられる。また、提案手法によってデータとナレッジグラフを同時に考慮できることから、KGE（によるタスク）の性能向上も期待される。これについても、実際利用されている大規模なナレッジグラフやそれに関連するデータを用いて検証したい。

参考文献

[Andrzejewski 11] Andrzejewski, D., Zhu, X., Craven, M., and Recht, B.: A Framework for Incorporating General Domain Knowledge into Latent Dirichlet Allocation using First-Order Logic, in *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pp. 1171–1177 (2011)

[Bordes 13] Bordes, A., Usunier, N., Gracia-Durán, A., Weston, J., and Yakhnenko, O.: Translating Embeddings for Modeling Multi-relational Data, in *Advances in Neural Information Processing Systems*, Vol. 26, pp. 2787–2795 (2013)

[Dettmers 17] Dettmers, T., Minervini, P., Stenetorp, P., and Riedel, S.: Convolutional 2D Knowledge Graph Embeddings, arXiv:1707.01476 (2017)

[Foulds 15] Foulds, J., Kumar, S. H., and Getoor, L.: Latent Topic Networks: A Versatile Probabilistic Programming Framework for Topic Models, in *Proceedings of the 32nd International Conference on Machine Learning*, pp. 777–786 (2015)

[Fung 03] Fung, G. M., Mangasarian, O. L., and Shavlik, J. W.: Knowledge-Based Support Vector Machine Classifiers, in *Advances in Neural Information Processing Systems*, Vol. 15, pp. 537–544 (2003)

[He 15] He, S., Liu, K., Ji, G., and Zhao, J.: Learning to Represent Knowledge Graphs with Gaussian Embedding, in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, pp. 623–632 (2015)

[Kingma 14] Kingma, D. P. and Welling, M.: Stochastic Gradient VB and the Variational Auto-Encoder, in *Proceedings of the 2nd International Conference on Learning Representations* (2014)

[Mei 14] Mei, S., Zhu, J., and Zhu, X.: Robust RegBayes: Selectively Incorporating First-Order Logic Domain Knowledge into Bayesian Models, in *Proceedings of the 31st International Conference on Machine Learning*, No. 1, pp. 253–261 (2014)

[Polson 13] Polson, N. G., Scott, J. G., and Windle, J.: Bayesian Inference for Logistic Models Using Pólya–Gamma Latent Variables, *Journal of the American Statistical Association*, Vol. 108, No. 504, pp. 1339–1349 (2013)

[Tipping 99] Tipping, M. E. and Bishop, C. M.: Probabilistic Principal Component Analysis, *Journal of the Royal Statistical Society: Series B*, Vol. 61, No. 3, pp. 611–622 (1999)

[Varol 12] Varol, A., Salzmann, M., Fua, P., and Urtasun, R.: A Constrained Latent Variable Model, in *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2248–2255 (2012)

[Xiao 16] Xiao, H., Huang, M., Hao, Y., and Zhu, X.: TransG : A Generative Model for Knowledge Graph Embedding, in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pp. 2316–2325 (2016)

[Yang 15] Yang, B., Yih, tau W., He, X., Gao, J., and Deng, L.: Embedding Entities and Relations for Learning and Inference in Knowledge Bases, in *Proceedings of the 3rd International Conference on Learning Representations* (2015)

[Zhang 16] Zhang, F., Yuan, N. J., Lian, D., Xie, X., and Ma, W.-Y.: Collaborative Knowledge Base Embedding for Recommender Systems, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 353–362 (2016)

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

作用によって生じる共変量シフトバイアスの補正

Covariate shift bias correction induced by action

比嘉　亮太
Ryota Higa

NEC データサイエンス研究所
NEC Data Science Research Laboratories

In this study, we consider the covariates shift caused by the action on the data problem. Also, consider the method of bias correction using the density ratio estimation.In the numerical experiments, the action actually cause unsteady/extrapolation problem of distribution. It shows that have a significant bias to the expected value evaluation.Further illustrates that the method using the density ratio even in the case of changing the data size and dimensions are effective.

1. はじめに

人工知能/機械学習における研究やビジネスにおいて、データの蓄積と活用手法の確立がますます重要になっている。解くべき課題の明確化と十分なデータをそろえることが、機械学習を用いた予測モデルを構築するためのスタートラインである。しかしながら、これまでにとったことの無い行動を含むリーンスタートアップやキャンペーン最適化、新素材探索などは、本質的にデータ取得が部分的になる。その様な場合は、機械学習の手法が適用困難であり、適用できたとしても適切な結果を導けない。

データ取得が部分的である場合に生じる問題として、共変量シフトによって生じるバイアスの問題が指摘されている。これは、予測と評価時の入力分布が異なる、つまり非定常な分布によって生じる期待値評価のズレである。密度比推定を用いたアプローチは、非定常な分布によって引き起こされる回帰や分類において有効であることが知られている [Shimodaira 00, Sugiyama 07]。

また、我々が知りたい量としては、データのミクロな分布そのものではない。これまでと異なったアプローチを行った場合の費用対効果や売り上げなどの期待値であるマクロな量である。期待値は、上昇、一定、減少など現状との相対的な比較によって評価できる。

本研究において、データに対する作用によって生じる共変量シフトの問題と密度比推定を用いたバイアス補正の方法について考察する。数値実験において、最適化の作用が実際に分布の非定常/外挿問題を引き起こし、期待値評価に大きなバイアスを与えることを示す。さらに、データサイズや次元を変えた場合においても密度比を用いた方法が有効であることを示す。

2. 問題設定

具体例として、以下の目的関数の期待値評価を考える。

$$l(a, x, y) := xy - ca \tag{1}$$

ここで、a は操作変数、x は入力、y は出力であり、c は定数とする。例えば、a はキャンペーン、x は収入、y は解約であり、$l(a, x, y)$ はキャンペーンの費用対効果の指標と成る。または、a は電圧、 x は電流、y は抵抗とする。操作変数 a に対して、あるアルゴリズムを適用した後の、l の期待値を

$$\mathbb{E}_F[l(a, x, y)] := \sum_{a,y} \int \mathrm{d}x p_F(a, x, y) l(a, x, y) \tag{2}$$

と定める。本問題は、上記の期待値を評価することが目的である。しかし、a に対するアルゴリズム適用前のイニシャルデータセット $\mathcal{D}_I = \{a_n, x_n, y_n\}$ は得られているが、操作後のデータセットは部分的 $\mathcal{D}_F = \{a_n, x_n\}$ であるとする。

2.1 仮定

出力 y の条件付き分布において、以下の仮定を置く。

$$p_I(y|x, a) = p_F(y|x, a) \tag{3}$$

ここで、添え字の I は a の操作前、F は a の操作後とする。つまり、出力結果においては、保存則が成立りつしているとする。1 つの方法としては、過去のデータから $p(y|x, a)$ を推定して、それをデータを生成するシミュレータとみなして a 操作後の期待値を求める方法がある。しかしながら、推定したモデルに対して操作変数を変えて期待値評価を行うと、データの無い部分の外挿問題に対応してしまい、過大、過小評価の結果となる。以下では、密度比を用いたアプローチを示す。

3. 密度比を用いた定式化

操作変数 a によって、$p(a, x, y)$ に対して定常な分布は仮定できない。以下のような状況を考える。

$$p_I(a, x, y) \neq p_F(a, x, y) \tag{4}$$

ここで、両者の比を γ とすると、y の条件付分布の仮定より

$$\gamma(a, x) := \frac{p_F(y|a, x) p_F(a, x)}{p_I(y|a, x) p_I(a, x)} \tag{5}$$

$$= \frac{p_F(a, x)}{p_I(a, x)} \tag{6}$$

となり、y の情報が不要になることがわかる。

連絡先: 比嘉　亮太、NEC データサイエンス研究所、神奈川県川崎市中原区下沼部 1753、email:r-higa@ct.jp.nec.com

密度比 $\gamma(a,x)$ を用いると、(2) は

$$\begin{aligned}\mathbb{E}_F[l(a,x,y)] &= \sum_{a,y}\int \mathrm{d}x p_F(a,x,y)l(a,x,y)\\ &= \sum_{a,y}\int \mathrm{d}x \frac{p_F(a,x,y)}{p_I(a,x,y)} p_I(a,x,y)l(a,x,y)\\ &= \sum_{a,y}\int \mathrm{d}x \gamma(a,x)p_I(a,x,y)l(a,x,y) \quad (7)\\ &= \mathbb{E}_I[\gamma(a,x)l(a,x,y)] \quad (8)\end{aligned}$$

となる。ただし、$\gamma(x,a) < \infty$ とする。つまり、本来知りたい操作後の期待値は、操作前の期待値と γ によって評価できることを示している。これは、y に対して手元に得られているデータの活用方法やシミュレータなどのドメイン知識などの助けを借りることによって、マクロな期待値の評価は可能であることを意味している。

仮定を検証する方法としては、$r(x,a) = 1$ となる場合において、操作前後の期待値の差であるバイアス

$$B = \mathbb{E}_F[l(a,x,y)] - \mathbb{E}_I[\gamma(a,x)l(a,x,y)] \quad (9)$$

がゼロに成ることを確かめればよい。これは、y を生成するモデルに対する改良にも使える。例えば、仮説の再検討において、時間 t を与えた $p_I(y|x,a) = p_F(y|x,a)\exp(-\alpha t)$ として、パラメータ α を決定する問題とする。ミュレータと現実の環境の変化への適用といった、シミュレータ開発者にとっても、データ主導による改良できる可能性を示している。

4. 数値実験

密度比を用いた方法と、仮説を用いて y のモデルを推定する方法の比較を行い、分布のズレによるバイアス補正の効果を実験的に確かめる。

4.1 設定

検証の実験として、以下の様な設定を行う。

$$a \sim Bernoulli(a|\mu_a) \quad (10)$$

$$x|a \sim \mathcal{N}(x|\mu_a, \sigma_a^2) \quad (11)$$

$$y|a,x \sim Bernoulli(y|\mu_a, x) \quad (12)$$

$$(13)$$

操作変数 a に対しては、x 全体の上位 K[%] にたいして $a = 1$、それ以外を $a = 0$ とする操作アルゴリズム (以下、Top of K と呼ぶ) を用いる。図 1 は、操作前後においての $p(a,x)$ 分布の変化である。Top of K の適用後の y は得られないとする。

4.2 結果

設定した人工モデルに対して、データサイズを変更しながら、モンテカルロにより 1000 回の試行によって目的関数の期待値を評価した。共変量シフトを用いたアプローチとの比較手法としては、$p(y|x,a)$ の保存則を用いて、捜査前のデータより、ロジスティック回帰、SVM(RBF), Deep(4 層) により推定した。

結果は、2 で与えられている。黒が操作前、赤が操作後の期待値である。青が共変量シフトを用いた期待値評価であり、データサイズ 100 で Top of K 適用後の期待値評価が出来ている。一方、ロジスティック (緑)、SVM(茶), Deep(橙) は過大、過小に期待値を見積もる傾向にあることが確かめられた。

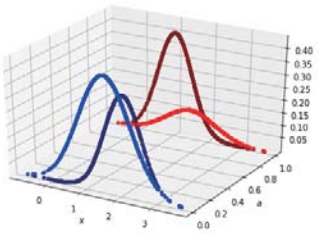


図 1: Top of K による分布の変化。$a = 0$ の x の分布は左側へシフトし、$a = 1$ の x の分布は右側へシフトしている。

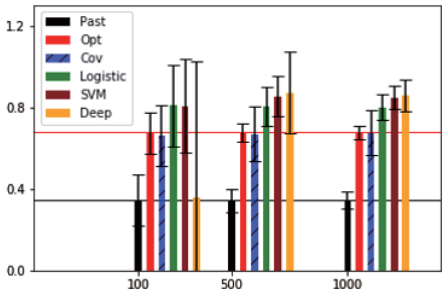


図 2: サンプル数を 100, 500, 1000 と変化させた場合での共変量シフトとロジスティック、SVM, Deep との比較。赤が正しい結果であり、共変量シフトが最もバイアスが補正できている結果である。

4.3 考察

直接 y を推定することは、分布の非定常性に加えデータの偏りなどで生じる外挿問題にダイレクトに挑戦する難しい問題である。一方、密度比を用いた期待値評価は、観測されたデータの内挿問題であり、分布の詳細な情報をあきらめる代わりに予測の難易度は下がる。マクロな量である期待値の評価を行う場合は、分布の普遍的な部分やドメイン知識で補える部分できりわけるアプローチは、既存のデータを活用する上で有効な手法であることが経験的に確かめられた。

5. おわりに

本研究では、密度比を用いたデータの活用が、将来の期待値評価に有効であることを示した。今後は、本研究の一般的な定式化を含め、逐次的な操作を含む期待値評価や、目的関数の推定を含む問題へ拡張する予定である。

参考文献

[Shimodaira 00] H. Shimodaira, “Improving predictive inference under covariate shift by weighting the log-likelihood function,” J. Stat. Plan. Inference, vol. 90, no. 2, pp. 227244, 2000.

[Sugiyama 07] M. Sugiyama, M. Krauledat, and K.-R. Mller, “Covariate shift adaptation by importance weighted cross validation,” J. Mach. Learn. Res., vol. 8, pp. 9851005, 2007.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

線形可解マルコフ決定過程のためのバッチ強化学習

Batch Reinforcement Learning for Linearly Solvable MDP

西智樹 *1 Tomoki Nishi　大滝啓介 *1 Keisuke Otaki　吉村貴克 *1 Takayoshi Yoshimura

*1豊田中央研究所
Toyota Central R&D Labs., Inc.

Linearly solvable Markov decision process (L-MDP) is an essential subclass of MDP to find a better policy efficiently. We first develop a novel batch reinforcement learning algorithm for L-MDP. The algorithm simultaneously learns a state value function and a predictor of state values at next step by using pre-collected data. We evaluate our method on traffic signal control domain in a single intersection with the traffic simulator SUMO. Our experiment demonstrates that our method finds the policy on the domain efficiently.

1. はじめに

線形可解マルコフ決定過程 (L-MDP) は Todorov によって提案されたマルコフ決定過程 (MDP) の一種である [Todorov 06]．L-MDP はベルマン方程式を，規定の行動を実施時のある状態と次状態に関する線形差分方程式に変形することで，ダイナミクスが既知の場合には最適政策を効率的に求められる．また文献 [Todorov 06] ではさらに，ダイナミクスが未知の場合の強化学習として Z-learning が提案され，Q-learning [Watkins 92] よりも効率的に政策を学習できることが報告されている．

一方で，Fitted Q-iteration [Riedmiller 05] に代表されるバッチ強化学習は，予め収集したデータに基づき，オフラインでより良い行動を学習することができる強力な枠組みの一つである．自動運転・交通制御など学習中に事故や渋滞を引き起こす可能性がある問題では，有用な学習手法である [Riedmiller 07]．筆者は，先に述べた L-MDP に対し，連続状態行動空間におけるエージェントのダイナミクスが既知の場合のバッチ強化学習 passive Actor-Critic を提案し，自動運転の混雑時の高速道合流において有用であることを示した [Nishi 17]．しかしながら，離散行動空間における L-MDP のためのバッチ強化学習はこれまで提案されておらず，交通信号機など行動が離散的に表現される場合には利用することができなかった．

そこで本稿では，離散行動空間における L-MDP のためのバッチ強化学習 Fitted Z-learning を提案する．本手法は，線形化されたベルマン方程式に基づく状態価値の学習と，ある状態である行動を選択した際の遷移先の状態価値の学習とを同時に行うことで政策を学習する．また提案手法の有効性を検証するために，シミュレーションを用いた 1 交差点での信号機制御で評価を行う．

2. 準備

2.1 マルコフ決定過程と Q-learning

MDP は $\langle \mathcal{S}, \mathcal{A}, T, R, \gamma \rangle$ の 5 つの要素の組で表現できる．$\mathcal{S}$ は状態空間，$\mathcal{A}$ は行動空間，$T : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1]$ は状態遷移モデル，$R : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to \mathbb{R}$ はコストモデル，γ はコストの 1 ステップ毎の割引率を表す．また政策 $\pi : \mathcal{S} \to \mathcal{A}$ は状態 $s \in \mathcal{S}$ の時に行動 $a \in \mathcal{A}$ を出力する関数である．マルコフ決定過程における基本的な問題設定は，状態 s が与えられた時に，目的関数 J を最小化する政策 $\pi^* : \mathcal{S} \to \mathcal{A}$ を発見することである．その時の目的関数は，

$$J \coloneqq \mathbb{E}_\pi \left[\sum_{k=0}^{\infty} \gamma^k r_k \right],$$

で表される無限ホライゾンにおけるコストの累積和の期待値がよく用いられる．ここで，$\mathbb{E}_\pi[\cdot]$ は政策 π の下での期待値，r_k は時刻 k での即時コストを表す．政策 π の下での状態価値関数 V^π，及び行動価値関数 Q^π を下記のように定義すると，

$$V^\pi(s) \coloneqq \mathbb{E}_\pi \left[\sum_{k=0}^{\infty} \gamma^k r_{t+k} | s_t = s \right] \quad s \in \mathcal{S},$$

$$Q^\pi(s, a) \coloneqq \mathbb{E}_\pi \left[\sum_{k=0}^{\infty} \gamma^k r_{t+k} | s_t = s, a_t = a \right] \quad s \in \mathcal{S}, a \in \mathcal{A},$$

時刻 t における最適な政策 $\pi^*(s_t)$ 及び，その下での行動価値関数 $Q^*(s_t, a_t)$，状態価値関数 $V^*(s_t)$ はベルマン方程式

$$Q^*(s_t, a_t) = r(s_t, a_t) + \mathbb{E}_{s_{t+1} \sim p(s_{t+1}|s_t, \pi^*(s_t))} \left[\gamma V^*(s_{t+1}) \right], \tag{1}$$

$$V^*(s_t) = \min_{a_t \in \mathcal{A}} Q^*(s_t, a_t),$$

$$\pi^*(s_t) = \arg\min_{a_t \in \mathcal{A}} Q^*(s_t, a_t), \tag{2}$$

を満たす．

モデルフリー強化学習の最もよく知られた方法である Q-learning は，ベルマン方程式に基づき計算される TD 誤差 $\mathcal{E}$ を減少させるように，行動価値関数 Q の推定値を更新する手法である．

$$\mathcal{E}_t = Q(s_t, a_t) - \left(r(s_t, a_t) + \gamma \max_{a \in \mathcal{A}} Q(s_{t+1}, a) \right).$$

2.2 線形可解マルコフ決定過程と Z-learning

L-MDP は $\langle \mathcal{S}, \mathcal{U}, T, R, \gamma \rangle$ の 5 つの要素の組で表現される．但し $\mathcal{U}$ は L-MDP における行動空間であり，他の要素は通常の MDP と同様である．L-MDP は通常の MDP に下記の 2 つの仮定を追加した MDP である．

連絡先: 西智樹，豊田中央研究所，愛知県長久手市横道 41-1，nishi@mosk.tytlabs.co.jp

1. $p(s_{t+1}|s_t, \mathbf{u}_{s_t}) \coloneqq p(s_{t+1}|s_t, \pi_0)\exp(\mathbf{u}_{s_t})$,
2. $r(s_t, \mathbf{u}_{s_t}) \coloneqq q(s_t) + \lambda \mathrm{KL}(p(s_{t+1}|s_t, \pi_0)||p(s_{t+1}|s_t, \mathbf{u}_{s_t}))$,

但し, $\mathbf{u}_{s_t} \in \mathbb{R}^{|\mathcal{S}|}$ は状態 s_t における行動ベクトル, $p(s_{t+1}|s_t, \pi_0)$ 及び $p(s_{t+1}|s_t, \mathbf{u}_{s_t})$ はそれぞれ予め設計者が決めた政策 π_0 に基づいて行動した時の状態遷移の確率分布，行動 $\mathbf{u}_{s_t}$ をとった時の状態遷移の確率分布，$q(s_t)$ は状態にのみ依存した状態コスト，$\mathrm{KL}(\cdot||\cdot)$ は KL-divergence．また λ は，状態コスト $q(s_t)$ に対する KL-divergence で表される行動コストの，相対的な大きさを調整するためのパラメータである．1 つめの仮定は，L-MDP における行動 $\mathbf{u}_{s_t}$ は，MDP の行動とは異なり，状態 s からある状態 s' への状態遷移確率をどれくらい変化させるかであり，その大きさが $\exp(\mathbf{u}_{s_t})$ で表されるとする仮定である．また 2 つめの仮定は，L-MDP の即時コスト関数 $r(s_t, \mathbf{u}_{s_t})$ は，状態コスト $q(s_t)$ と，$p(s_{t+1}|s_t, \pi_0)$ と $p(s_{t+1}|s_t, \mathbf{u}_{s_t})$ の KL-divergence で表される行動コスト，の 2 つに分解できるとする仮定である．

上記の仮定から，最適な政策 π_L^* は，

$$Z^*(s_t) = \exp(-q(s_t)/\lambda)\mathbb{E}_{s_{t+1}\sim p(s_{t+1}|s_t,\pi_0)}\left[(Z^*(s_{t+1}))^\gamma\right], \tag{3}$$
$$Z^*(s_t) \coloneqq \exp(-V^*(s_t)/\lambda),$$

で表される線形化ベルマン方程式を満たすことが導かれる (詳細は [Todorov 06] を参照のこと)．その時最適な政策 $\pi_L^*(s_t, s_{t+1})$ 及びその政策の下での状態遷移は，

$$\pi_L^*(s_t, s_{t+1}) = \left(V^*(s_t) - q(s_t) - \gamma V^*(s_{t+1})\right)/\lambda, \tag{4}$$
$$p(s_{t+1}|s_t, \pi_L^*) = \frac{p(s_{t+1}|s_t, \pi_0)Z^*(s_{t+1})}{\exp(q(s_t)/\lambda)Z^*(s_t)}, \tag{5}$$

となる．

政策 π_0 に従う際の状態遷移確率 $p(s_{t+1}|s_t, \pi_0)$ と状態コスト $q(s_t)$ が既知である場合には，離散状態行動空間における L-MDP の状態コスト V と最適政策 π_L^* は，式 (3) の Z^* についての固有値問題を解くことで求められる．また連続状態行動空間の場合においても，Z 関数 $Z(s_t)$ を基底関数の線形和により近似することで，効率的に学習できる [Todorov 09]．

また，状態遷移確率 $p(s_{t+1}|s_t, \mathbf{u}_{s_t})$ と即時コスト関数 $q(s_t)$ が未知である場合は，予め決められた政策 π_0 に従って行動した時のデータに基づいて，Z 関数を学習する Z-learning が適用できる．Z-learning は，線形化ベルマン方程式に基づき計算される TD 誤差 $\mathcal{E}_L$ を減少させるように，Z 関数の推定値を更新する．

$$\mathcal{E}_{L,t} = Z(s_t) - \exp(-q(s_t)/\lambda)(Z(s_{t+1}))^\gamma.$$

3. 提案手法 Fitted Z-learning

L-MDP のためのバッチ強化学習として，Fitted Z-learning を提案する．Z-learning は，既定の政策 π_0 に従って行動した際の状態遷移のデータに基づき，Z 関数を学習することができる．また，状態コストを既知とすると，L-MDP における政策 π_L は，式 (4) により学習した Z 関数を用いて求められる．さらに $p(s_{t+1}|s_t, \pi_0)$ を既知とすると，政策の下での状態遷移は，式 (5) により算出できる．しかしながら，L-MDP における行動は状態遷移に与える影響の強さであり，エージェントが具体的に取り得るどの行動（MDP の意味での行動）を取れば，式 (5) から算出される望ましい状態遷移を実現することができるかは自明ではない．

行動価値関数 $Q^*(s_t, a_t)$ が学習出来れば，MDP における政策 π^* は式 (2) により求められる．ここで，即時コスト $r(s_t)$ が状態コスト $q(s_t)$ で近似できると仮定する．これは，λ を小さく設定することで常に実現できる．

$$r(s_t, a_t) = q(s_t) + \lambda \mathrm{KL}(p(s_{t+1}|s_t, \pi_0)||p(s_{t+1}|s_t, \mathbf{u}_{s_t})) \approx q(s_t). \tag{6}$$

この時式 (6) の仮定により行動価値関数は，

$$\begin{aligned} Q^*(s_t, a_t) &= r(s_t, a_t) + \mathbb{E}_{s_{t+1}\sim p(s_{t+1}|s_t,\pi^*)}\left[\gamma V^*(s_{t+1})\right], \\ &\approx q(s) + \mathbb{E}_{s_{t+1}\sim p(s_{t+1}|s_t,\pi^*)}\left[\gamma V^*(s_{t+1})\right], \end{aligned}$$

と近似できる．また近似した行動価値関数に基づく政策 $\hat{\pi}$ は，

$$\begin{aligned} \hat{\pi}(s_t) &= \arg\min_{a_t\in\mathcal{A}}\left(q(s_t) + \mathbb{E}_{s_{t+1}\sim p(s_{t+1}|s_t,\hat{\pi})}\left[\gamma V(s_{t+1})\right]\right), \\ &= \arg\min_{a_t\in\mathcal{A}}\mathbb{E}_{s_{t+1}\sim p(s_{t+1}|s_t,\hat{\pi})}\left[V(s_{t+1})\right], \\ &= \arg\max_{a_t\in\mathcal{A}}\mathbb{E}_{s_{t+1}\sim p(s_{t+1}|s_t,\hat{\pi})}\left[Z(s_{t+1})\right], \end{aligned}$$

と書き直すことができる．この政策は，状態 s_t において，次状態 s_{t+1} での Z 関数の期待値が最大となる行動を選択することを意味している．つまり，即時コスト $r(s_t)$ が状態コスト $q(s_t)$ で近似できる場合には，状態 s_t においてある行動 a_t をとった時の次状態 s_{t+1} での Z 値を予測することで，行動を決定することができる．

提案する Fitted Z-learning は，2 つの関数，

- 現状態 s_t の Z 値を推定する関数 $\hat{Z}(s_t; \boldsymbol{\theta})$,
- 行動 a_t を選択した時の次状態 s_{t+1} での Z 値を予測する関数 $\hat{Z}(s_t; a_t, \boldsymbol{\phi})$,

の学習を通して政策 π を学習する．ただし，$\boldsymbol{\theta}$ 及び $\boldsymbol{\phi}$ は共に関数近似のためのパラメータである．Z-learning と同様に，$\boldsymbol{\theta}$ は，L-MDP におけるベルマン方程式から算出される TD 誤差を小さくするよう学習する．$\boldsymbol{\phi}$ は，時刻 t で行動 a_t を選択した時の次状態 s_{t+1} の $\hat{Z}(s_{t+1}; \boldsymbol{\theta})$ の値と $\hat{Z}(s_t; a_t, \boldsymbol{\phi})$ との最小 2 乗誤差を小さくするように学習する．本稿ではさらに multi-step Q-learning [Peng 96] を参考に，連続した T ステップの状態遷移を考慮してパラメータの更新を行う multi-step Fitted Z-learning を用いた．

詳細なアルゴリズムを Algorithm 1 に示す．

4. 数値実験

本節では 1 交差点における信号機制御タスクにより Fitted Z-learning の有効性を検証する．

4.1 ニューラルネットワークの構成

Fitted Z-learning の $\hat{Z}(s; \boldsymbol{\theta})$ 及び $\hat{Z}(s; a, \boldsymbol{\phi})$ をニューラルネットワークでモデリングし，学習した．ニューラルネットワークの構成を図 1 に示す．$\hat{Z}(s; \boldsymbol{\theta})$ と $\hat{Z}(s; a, \boldsymbol{\phi})$ は，状態 s の Z 値と，状態 s で行動 a を選択した時の次状態の Z 値の推定値であり，この 2 つの Z 値は相互に関連するため，ニューラルネットワークの多くのパラメータは共有出来ると予想される．そこで Dueling Network Architectures [Wang 15] を参考に，図 1 のように下位層のネットワークを共有したフォーク

Algorithm 1 multi-step Fitted Z-learning

procedure 予め収集したデータセット $\mathcal{D}$ から政策を学習
  パラメータ $\boldsymbol{\theta}$ 及び $\boldsymbol{\phi}$ を初期化
  for エピソード $i = 1$ to N **do**
    $\mathcal{D}$ から T ステップ連続したデータをランダムに抽出:
      $\{(i, s_1, a_1, s_2, q_1), ..., (i, s_T, a_T, s_{T+1}, q_T)\} \sim \mathcal{D}$
    パラメータの勾配を初期化: $d\boldsymbol{\theta} \leftarrow \mathbf{0}$
    $R \leftarrow \hat{Z}(s_{T+1}; \boldsymbol{\theta})$
    for ステップ $t = T$ to 1 **do**
      $R \leftarrow \exp(-q_t/\lambda)R^{\gamma}$
      TD 誤差 $\mathcal{E}_{L,t}$ を計算: $\mathcal{E}_{L,t} \leftarrow \left(R - \hat{Z}(s_t; \boldsymbol{\theta})\right)$
      パラメータの勾配 $d\boldsymbol{\theta}$ を更新: $d\boldsymbol{\theta} \leftarrow d\boldsymbol{\theta} + \partial \mathcal{E}_{L,t}^2 / \partial \boldsymbol{\theta}$
      誤差 δ_t を計算:
        $\delta_t \leftarrow \left(\hat{Z}(s_{t+1}; \boldsymbol{\theta}) - \hat{Z}(s_t; a_t, \boldsymbol{\phi})\right)$
      パラメータの勾配 $d\boldsymbol{\phi}$ を更新: $d\boldsymbol{\phi} \leftarrow d\boldsymbol{\phi} + \partial \delta_t^2 / \partial \boldsymbol{\phi}$
    end for
    パラメータ $\boldsymbol{\theta}, \boldsymbol{\phi}$ を $d\boldsymbol{\theta}/T, d\boldsymbol{\phi}/T$ に基づき更新
  end for
end procedure

型のネットワークを用いた．つまり，下位層のパラメータは，$\boldsymbol{\theta}$ 及び $\boldsymbol{\phi}$ で同じ値を用い，$\boldsymbol{\theta}$ 及び $\boldsymbol{\phi}$ の更新時に更新される．また，全ての中間層は，全ノードが結合された全結合層 (Fully Connected Layer, FCL) を用い，最終層以外の活性化関数は全て ReLU$(\cdot) \coloneqq \max(0, \cdot)$ を用いた．出力層の活性化関数は，$Z(s) > 0, \forall s \in \mathcal{S}$ であるため，$\hat{Z}(s; \boldsymbol{\theta})$ 及び $\hat{Z}(s; a, \boldsymbol{\phi})$ 共に，Softplus$(\cdot) \coloneqq \log(1 + \exp(\cdot))$ を用いた．

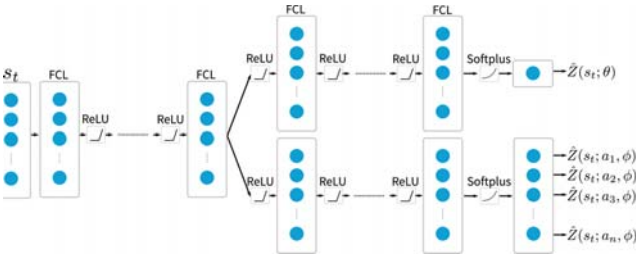


図 1: Fitted Z-learning のためのネットワークの構成

4.2 実験設定

従来手法 Q-learning と提案手法 Fitted Z-learning に対し，学習の安定性，速度，性能の 3 つの観点での比較評価を行った．評価タスクとしては，上下左右に伸びた 4 道路各方向 1 車線ずつ，計 8 車線が結合している道路（図 2）において，交差点の信号機を制御し，全道路の交通量を最大化するタスクを用いた．以下で，評価タスクの詳細を述べる．

各道路は直進のみができ，右左折できない道路とした．また状態は各車線を走行する全車両の平均速度（8 状態変数）及び各車線で信号待ちしている車両の台数（8 状態変数）の計 16 次元のベクトルで表現し，行動は（縦方向，横方向）の道路の信号機の色がそれぞれ（青，赤）または（赤，青）の 2 つとした．各車両の出発地及び目的地はランダムに各車線の端点を選択した．信号を制御した際の交通流をシミュレートするために交通流シミュレータ SUMO [Krajzewicz 02] を用いた．また Fitted Z-learning は，SUMO により生成された固定サイクル長の信号機で制御した 5,000,000 秒（約 1400 時間）のデータを用いて 40,000 回ニューラルネットワークのパラメータを更新した．本実験では，実際の街中で一般に用いられている固定サイクル長の信号機，Q-learning により学習された信号機

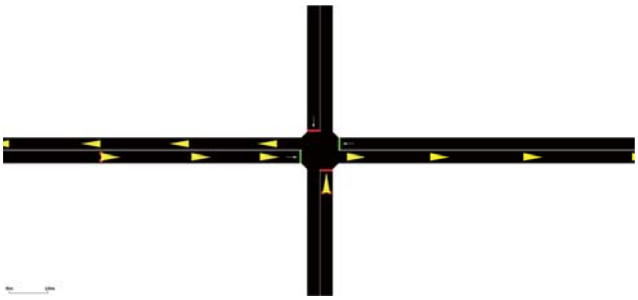


図 2: 1 交差点の信号機制御タスク．上下左右に伸びた 4 道路各方向 1 車線ずつの計 8 車線が結合した交差点の信号機を制御するタスク．三角で示された車両が交差点通過するために待つ時間を最小化するような制御方法を学習する．

と比較実験を行った．Q-learning は Fitted Z-learning で用いたニューラルネットワークと層数が同じになるよう中間層が 4 層のニューラルネットワークを用いた．

4.3 実験結果

4 道路が各方向 1 車線ずつ計 8 車線が結合された交差点の信号機制御を Q-learning, Fitted Z-learning で各 20 回学習した．まず 40,000 ステップ以内に学習初期の政策の平均待ち時間に対して，8 割以下の待ち時間を達成できた場合を学習の成功とし，各手法の成功率を図 3 に示す．この結果から Fitted Z-learning は Q-learning に比べ安定して学習ができることが分かった．これは，Q-learning は状態と行動を入力とする状態行動価値関数を学習する必要があるのに対し，Fitted Z-learning は状態のみを入力とする Z 関数の学習を通して政策を学習するため，学習が容易になり，学習が安定したのではないかと推察される．

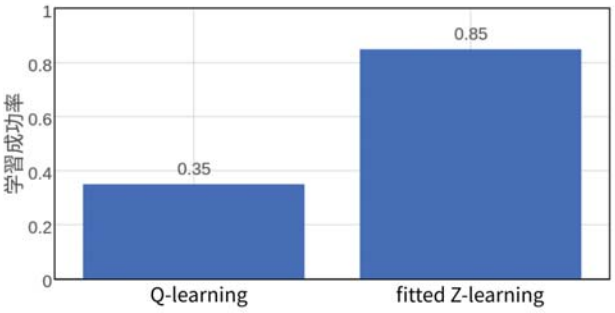


図 3: 学習の成功率

また，学習が成功した時のみの学習曲線の平均，及び最後の繰り返し回数時点での性能の平均値と標準偏差を，それぞれ図 4 に示す．図 4(左図) から Fitted Z-learning は予め収集したデータのみから Q-learning と同等の，固定サイクルに比べ 2 秒以上短い待ち時間を達成する政策を学習できた．また図 4(右図) から，街中で一般に見られる固定サイクルの信号機で制御された交差点のデータを用いることで，Fitted Z-learning は Q-learning に比べ効率的にパラメータを学習できることが分かった．これも上記の学習が安定した理由と同様，Fitted Z-learning の方が学習が容易であったためではないかと推察される．

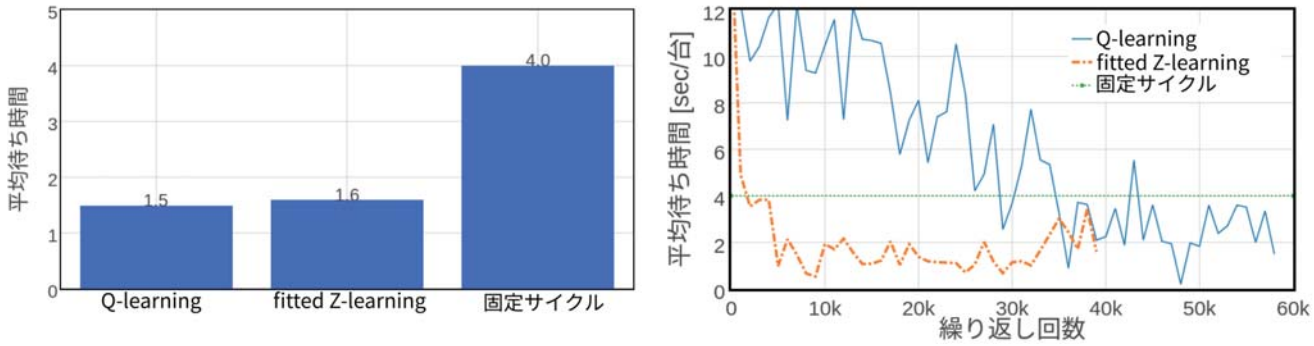


図 4: 1 交差点信号機制御の実験結果．(左図) 学習後の各手法で信号制御した場合の平均待ち時間．Fitted Z-learning は予め収集したデータのみから Q-learning と同等の，固定サイクルに比べ 2 秒以上短い待ち時間を達成する政策を学習できた．(右図)Fitted Z-learning および Q-learning の学習曲線．Fitted Z-learning は Q-learning に比べ少ない繰り返し回数で同等の政策を学習することができた．但し Q-learning は繰り返し回数 4,0000 の時点では収束していなかったため，60,000 まで学習を継続した．

5. 結論

本稿では，離散行動空間における線形可解マルコフ決定過程のためのバッチ強化学習として Fitted Z-learning を提案した．Fitted Z-learning は線形可解マルコフ決定過程から導かれる線形化ベルマン方程式に基づきマルコフ決定過程での状態価値関数に相当する Z 関数と共に，ある行動を選択した時の次状態での Z 関数の値を予測する関数を同時に学習する．提案手法を評価するために 1 交差点の信号機制御で実験を行った．その結果，提案手法はデータのみから固定サイクルに比べ，2 秒以上待ち時間が短い政策を Q-learning より効率的に学習できることが確認できた．

参考文献

[Krajzewicz 02] Krajzewicz, D., Hertkorn, G., Rössel, C., and Wagner, P.: SUMO (Simulation of Urban MObility) An open-source traffic simulation Car-Driver Model, *Proc of the 4th Middle East Symposium on Simulation and Modelling*, pp. 183–187 (2002)

[Nishi 17] Nishi, T., Doshi, P., James, M., and Danil, P.: Actor Critic for Linearly-Solvable Continuous MDP with Partially Known Dynamics, *arXiv preprint:1706.01077* (2017)

[Peng 96] Peng, J. and Williams, R. J.: Incremental multi-step Q-learning (1996)

[Riedmiller 05] Riedmiller, M.: Neural fitted Q iteration - First experiences with a data efficient neural Reinforcement Learning method, in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (2005)

[Riedmiller 07] Riedmiller, M., Montemerlo, M., and Dahlkamp, H.: Learning to Drive a Real Car in 20 Minutes, in *2007 Frontiers in the Convergence of Bioscience and Information Technologies*, pp. 645–650, IEEE (2007)

[Todorov 06] Todorov, E.: Linearly-solvable Markov decision problems, *Advances in neural information processing systems*, pp. 1369–1376 (2006)

[Todorov 09] Todorov, E.: Eigenfunction approximation methods for linearly-solvable optimal control problems, in *2009 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning, ADPRL 2009 - Proceedings*, pp. 161–168 (2009)

[Wang 15] Wang, Z., Schaul, T., Hessel, M., Hasselt, van H., Lanctot, M., and Freitas, de N.: Dueling Network Architectures for Deep Reinforcement Learning, No. 9 (2015)

[Watkins 92] Watkins, C. J. C. H. and Dayan, P.: Q-learning, *Machine learning*, Vol. 8, No. 3-4, pp. 279–292 (1992)

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

Cappuccino：構造データに対するブートストラッピング手法

Cappuccino: Bootstrapping from Structural Data

花房 諒 *1　　小山田 昌史 *2
Ryo Hanafusa　　Masafumi Oyamada

*1関西学院大学 大学院理工学研究科
Graduate School of Science and Technology, Kwansei Gakuin University

*2日本電気株式会社 データサイエンス研究所
Data Science Research Laboratories, NEC Corporation

We present *Cappuccino*, a bootstrapping method for structural data, which extracts target pairs of words by treating structures in data as patterns that connect pairs of words. The task we tackle in this paper is relation extraction from structural data such as presentation slides and HTML files. In order to exploiting the structures for bootstrapping, we convert structural data to graphs, and apply frequent subgraph mining to the graphs to calculate reliabilities of word pairs and patterns. Experimental results of a relation extraction task show that Cappuccino outperforms a previous method.

1. 概要

本稿は，プレゼンテーションスライドや HTML のような構造データからの関係抽出タスクを扱い，構造データに対するブートストラッピング手法である Cappuccino を提案する．ブートストラッピング法は，目的の関係をもつ単語ペア (インスタンス) の集合と文集合から，従来は人手で作っていたインスタンスの抽出ルール (パターン) を生成することにより，関係抽出タスクにおいて人手による作業のコスト削減を実現する．そのために，与えられたインスタンスを繋ぐパターン集合を信頼度にもとづいて抽出し，同様に，抽出されたパターンに繋がれるインスタンス集合を信頼度にもとづいて抽出する．従来の手法は，テキストデータのみを対象としていたが，Cappuccino は構造データを対象とする．

Cappuccino は，データがもつ構造 (例：リスト構造) をパターンとすることで，インスタンスの抽出を行う．構造をパターンとして扱うために構造データをグラフへ変換し，そのグラフに頻出サブグラフマイニングを適用することにより，インスタンスとパターンの信頼度計算を可能にする．グラフのノードラベルが単語を表し，エッジラベルがパターンの一部分を表す．また，構造データ内に存在する通常の文に対しては，単語ごとに分解し，別々のノードを生成することでインスタンスの抽出を可能にする．

本稿では，実験のために作成した人工的な HTML ファイルを用いて，文字列のパターンだけを用いる従来のブートストラッピング手法と比較を行い，データがもつ構造を利用する Cappuccino が高い性能になることを確認する．本稿の主な貢献を以下にまとめる．

- ブートストラッピング法において，構造データをグラフへ変換し，データがもつ構造をパターンとして扱うことにより，非構造データを対象にした従来のブートストラッピング手法では抽出できないパターンやインスタンスを抽出できる．

連絡先: 花房　諒，関西学院大学　大学院理工学研究科，rhanafusa@kwansei.ac.jp
連絡先: 小山田 昌史 (NEC)，m-oyamada@cq.jp.nec.com

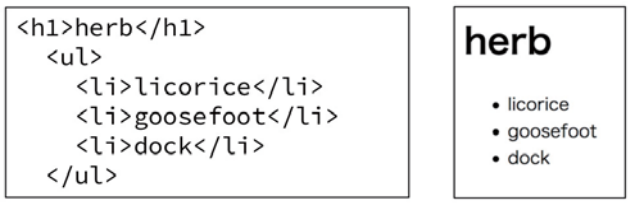


図 1: 構造データの例である HTML．左はソースコードの一部．

- 構造データからの関係抽出において，与えられたインスタンス集合からパターン抽出を行うため，構造のパターンを人手で指定することなく，インスタンスが抽出できる．

2. 背景

テキストデータを対象とした関係抽出は，一般的な自然言語処理のタスクである．例として，(apple, fruit) のような上位下位関係にあるペアの集合をテキストデータから抽出する問題が挙げられる．このような問題に対して，目的の関係をもつペア (インスタンス) をテキストデータから人手により探索することや，ペアを列挙して 1 つずつラベルを付けていくという手段はコストが高い．そのため，自動的にインスタンスを抽出する方法として，ブートストラッピング手法による関係抽出が広く行われており [Hearst 92, Riloff 99, McIntosh 11]，現在は Pantel が考案した Espresso [Pantel 06] が主に評価の基準となっている．これらの従来のブートストラッピング手法では，テキストデータのみを対象としており，発表で用いられるスライドや HTML のような構造データ (図 1) を対象とはしていない．そのため，従来手法を構造データへ適用し，関係抽出を行ったとしても獲得できるインスタンスには限界がある．本稿の提案方式 Cappuccino は，構造データがもつ構造を活用することにより，関係抽出タスクの性能向上を図る．

構造データを対象とした関係抽出としては，Wu が提案した，構造データから知識構築を行う Fonduer [Wu 17] が存在する．ただし，Fonduer を利用してインスタンスを抽出する

ためには，そのインスタンスを繋ぐ構造パターン (例：リスト構造) や条件を人手により指定しなければならない．すなわち，人手を利用するためコストが高いことが問題となる．本稿で提案する Cappuccino は，構造パターンも自動的に抽出するブートストラッピング法であるため，人手による作業のコスト削減に役立てることができる．

本稿では，構造データに対するブートストラッピング手法である Cappuccino を提案し，既存手法である Espresso と比較して Cappuccino が高い性能で関係抽出を行うことを示す．関連研究との違いを以下にまとめる．

- Espresso との違い
  Cappuccino では，構造データのもつ構造を利用して関係抽出を行う．一方，Espresso では構造を利用できないため関係抽出に限界がある．
- Founduer との違い
  Cappuccino では，関係抽出のための構造パターンを指定せず，そのパターンは自動的に抽出される．一方，Founduer では人手により構造パターンを指定しなければならない．

3. 準備

はじめに，提案する Cappuccino のベースとなる Espresso とブートストラッピング法について述べる．

3.1 ブートストラッピング法

本稿では，関係抽出をタスクとして，ブートストラッピング法を説明する．例として，上位下位関係をもつインスタンス (例：(apple, fruit)) の集合をテキストデータから獲得するタスクを考える．ブートストラッピング法は，以下の 3 つのフェーズを繰り返す．

1. シードインスタンス集合 (例: {(apple, fruit), (basketball, sport)}) を与える．
2. パターン抽出：テキストデータの各文を対象に，1 で与えられたシードインスタンスと共起する ”is a” のようなパターンを列挙する．列挙したパターンの各々に対して，信頼度スコアを計算し，そのスコアによってパターンのランキングを行う．ランキングの上位のパターンを選抜する．
3. インスタンス抽出：テキストデータの各文を対象に，2 で選抜されたパターンと共起するインスタンス (例：(dog, animal)) を列挙する．2 と同様にインスタンスを選抜する．選抜されたインスタンスが新たなシードインスタンス集合となる．

繰り返しの停止条件を満たすと，最後に選抜されたインスタンス集合が獲得したインスタンスとなる．信頼度スコアの関数や抽出するインスタンスとパターンの数，繰り返しの停止条件，各反復の途中で抽出したインスタンスも獲得したインスタンスとするかは，手法やタスクによって変化する．

ブートストラッピング法では，意味ドリフトが発生しうる [Curran 07]．意味ドリフトとは，ブートストラッピング法の繰り返しにより，シードインスタンスや目的のインスタンスとは関係のないインスタンスが抽出されるという問題である．これはジェネリックパターンを抽出してしまうことが原因である．ジェネリックパターンとは，目的の関係をもたないインスタンスを多く繋ぐパターン，すなわち，再現率は高いが精度が低いパターンである．Espresso は，意味ドリフトを回避しつつ，再現率と精度を高く保つ手法として提案された [Komachi 10]．

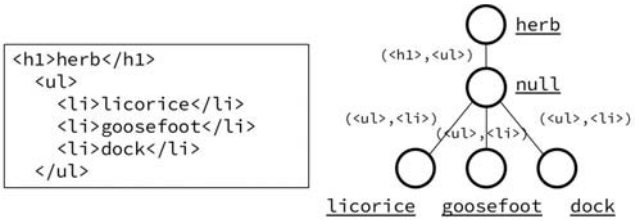


図 2: 構造データである HTML からグラフへの変換の例．

3.2 Espresso

Espresso は，意味ドリフトを回避するために，従来法のようにヒューリスティックにジェネリックパターンを排除するのではなく，自己相互情報量を用いた信頼度スコア関数によってジェネリックパターンを利用することで，バランスのとれた再現率と精度を実現した．パターン p の信頼度スコア関数 $r_\pi(p)$ とインスタンス i の信頼度スコア関数 $r_\iota(i)$ は，それぞれ以下で与えられる．

$$r_\pi(p) = \frac{1}{|I|} \sum_{i \in I} \frac{pmi(i,p)}{\max\ pmi} r_\iota(i), \tag{1}$$

$$r_\iota(i) = \frac{1}{|P|} \sum_{p \in P} \frac{pmi(i,p)}{\max\ pmi} r_\pi(p). \tag{2}$$

ここで，

$$pmi(i,p) = \log_2 \frac{|i,p|}{|i,*||*,p|} \tag{3}$$

とする [*1]．また，P と I はそれぞれ，各繰り返し時のパターンとインスタンスの集合であり，$\max\ pmi$ は，各繰り返し時のすべてのパターンとインスタンスの組み合わせにおいての pmi の最大値，$|i,p|$ は，テキストデータの各文を対象としたインスタンス i とパターン p との共起回数，$(*)$ はワイルドカードを表す．最初に与えるシードインスタンスの信頼度は 1 である．

4. Cappuccino

Cappuccino が Espresso と異なる点は構造データを扱うことである．構造データをブートストラッピング法で扱うための手法と，構造データ内に存在する通常の文の扱いについて述べる．Cappuccino の信頼度スコア関数は，第 3.2 節で述べた Espresso と同様である．

4.1 構造データ

Cappuccino の信頼度スコア関数は Espresso と同様だが，構造データにおいてのインスタンスとパターンの共起回数を求める必要がある．そのために，構造データをグラフとして扱い，頻出サブグラフマイニング [Yan 02] を利用する．構造データの例として HTML で説明する．

*1 自己相互情報量の推定量．自己相互情報量は，$pmi(i,p) = \log_2 \frac{P(i,p)}{P(i)P(p)}$ であるが，Pantel らは実験に実データを用いたため，推定量を利用している．なぜなら，$P(i)$ や $P(p)$ の計算には，インスタンスとパターンそれぞれの全体での個数が必要であるが，実データではこれが与えられないためである．

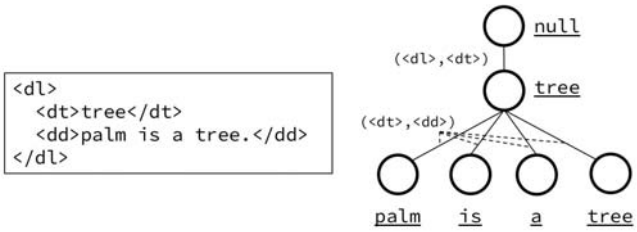


図 3: 通常の文を含むデータからグラフへの変換の例.

はじめに，図 2 で示すように HTML ファイルを頻出サブグラフマイニングで扱うためのグラフに変換する．ここで，ノードのラベルは HTML タグで囲まれる単語を表し，エッジのラベルが HTML タグの組を表す．ただし，(<ul>, <li>) と (<li>, <ul>) のような対称の組は異なるラベルとする．本稿の実験では，タグ名のみをエッジのラベルとして利用したが，属性やメタ情報なども利用することにより，エッジラベルの種類を増加させることが可能である．

次に，構造データを変換したグラフに対して，頻出サブグラフマイニングを適用し，サブグラフを列挙する．列挙したサブグラフの出現回数が，Espresso でのインスタンスとパターンの共起回数にあたる．なぜなら，1 つのサブグラフはインスタンスである特定のノードとノードの組を含み，そのサブグラフを表現するエッジの配置が 1 つのパターンを表すためである．サブグラフを列挙した結果，包含関係にあるようなパターンも存在するが，関係のない余計なエッジが含まれるようなパターンは信頼度スコア関数による信頼度が低くなるため問題ない．実験では，頻出サブグラフマイニングのツールとして DataMining-gSpan [Zhou 17] を利用した．サブグラフに含まれるノードの最大数は 4 とし，少なくとも 1 回出現するサブグラフを列挙して出現回数をカウントした．

4.2 構造データ内に存在する通常の文

構造データでは，リストのように項目として単語が並ぶだけでなく，説明文などの通常の文も存在する．このような文を第 4.1 節で述べた方法で扱うことを考える．通常の文を含むタグに対しては，図 3 で示すように，その文を単語ごとに分割し，ペアとなるノードと各単語との組を生成する．生成された組は，すべて同じ種類のエッジで繋がれる．すなわち，同じパターンで繋がれることになる．このようなパターンは，再現率は高いが精度が低いため，ジェネリックパターンになる可能性が高い．

5. 実験

構造データに対する関係抽出のタスクにおいて，提案手法である Cappuccino と既存手法である Espresso を再現率 (recall) と精度 (precision)，F 値 (F-score) の 3 つの指標で比較する．構造データには，実験のために作成した人工的な HTML ファイルを用いる．データがもつ構造を利用することで，文字列のパターンだけを用いるブートストラッピング手法と比べて高い性能となることを確認する．

関係抽出を行う構造データとして，正しい上位下位関係をもつインスタンス 100 個を含む人工的な HTML ファイルを実験に用いた．WordNet に登録されている植物の種類をインスタンスとして利用した．正しいインスタンスを繋ぐパターンとして 18 種類のパターンを準備した．実験で用いたパターン 18 種類のうち，9 種類は構造のパターン (例：リスト，定義) であり，残りの 9 種類は文字列のパターン (例：”is a”) である．実験で用いたインスタンス 100 個は，構造のパターンと文字列のパターンで繋がれているが，その個数と比率はインスタンス 10 個ずつに対して変化させた．構造のパターンと文字列のパターンとの抽出できるインスタンスの違いを確認するためである．その組み合わせの個数を表 1 に示す．10 個ずつのインスタンスのグループが $I_g,\ g = 1, \ldots, 10,$ であり，$s_j,\ j = 1, \ldots, 9,$ は構造のパターン，$c_k,\ j = 1, \ldots, 9,$ は文字列のパターンである．

表 1: 実験データにおけるインスタンスとパターンの組の数．インスタンス 10 個に対して，パターンが 3 種類あるため 30 個となる．

	I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8	I_9	I_{10}
s_1, s_2, s_3	30	0	0	0	30	0	0	0	0	0
s_4, s_5, s_6	30	30	0	0	0	30	0	0	0	0
s_7, s_8, s_9	30	30	30	0	0	0	30	0	0	0
c_1, c_2, c_3	0	30	30	30	0	0	0	30	0	0
c_4, c_5, c_6	0	0	30	30	0	0	0	0	30	0
c_7, c_8, c_9	0	0	0	30	0	0	0	0	0	30

構造データへの Espresso の適用について述べる．データは，構造をもつ HTML ファイルであるため，Espresso を適用するには HTML ファイルにおいて，どの部分を文とするかを決める必要がある．以下の 3 つの設定を用意した．

1. ESP1 : 通常の文を扱うためにタグで囲まれた部分を文とする．(例 : Kwansei Gakuin University is an university.)
2. ESP2 : 単語の組を扱うためにタグが間にある部分を文とする．(例 : fruit <li> apple)
3. ESP1 + ESP2 : 1 と 2 を合わせたものを文とする．

ブートストラッピング法では，シードインスタンスの数 S と信頼度のランキングから抽出されるインスタンスの数 K，繰り返し回数 L を決定しなければならない．これらのパラメータは，$S = 2$, $K = 50$, 75, 100, 125, 150, $L = 100$ とした．

実験結果を図 4 に示す．Cappuccino が再現率と精度，F 値のすべてで最も高い．このことから，構造データに対する関係抽出のタスクにおいては，文字列のパターンだけを利用するよりも，構造を利用することで，より多くの正しいインスタンスを抽出できることが示せたと考える．今後，Cappuccino で扱う構造パターンの種類を増やし，実データを用いて実験を行う予定である．

6. 結論

本稿は，発表で用いられるスライドや HTML のような構造データからの関係抽出タスクを扱い，構造データに対するブートストラッピング手法である Cappuccino を提案した．Cappuccino は，データがもつ構造をパターンすることで，インスタンスの抽出を行う．構造をパターンとして扱うために構造データをグラフへ変換し，そのグラフに頻出サブグラフマイニングを適用することにより，インスタンスとパターンの信頼度計算を可能にした．グラフのノードラベルが単語を表し，エッジラベルがパターンを表す．また，構造データ内に存在す

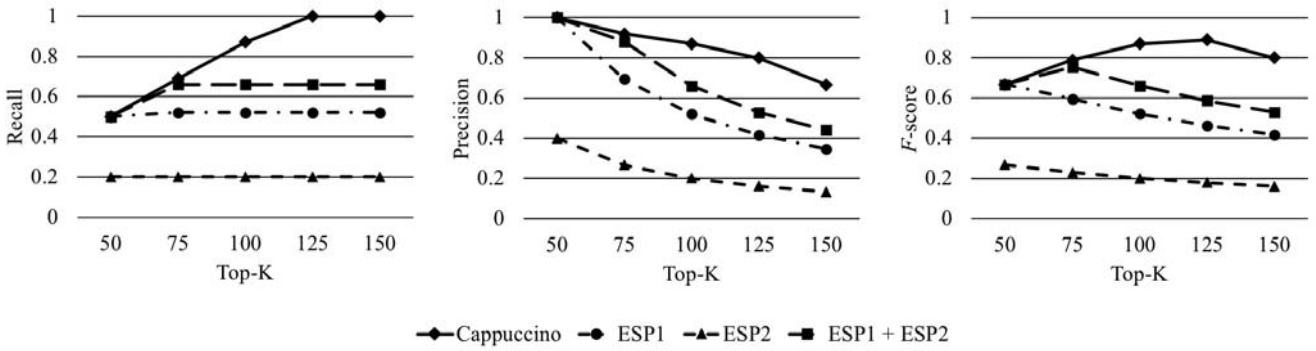


図 4: 上位下位関係の関係抽出タスクにおける実験結果.

る通常の文に対しては，単語ごとに分解し，別々のノードを生成することでインスタンスの抽出を可能にした．実験の結果，構造データに対する関係抽出のタスクにおいて，既存手法である Espresso と比較し，Cappuccino が再現率と精度，F 値のすべてで高い性能となることを確認した．

参考文献

[Curran 07] Curran, J. R., Murphy, T., and Scholz, B.: Minimising semantic drift with mutual exclusion bootstrapping, in *Proceedings of the 10th Conference of the Pacific Association for Computational Linguistics*, Vol. 6, pp. 172–180 (2007)

[Hearst 92] Hearst, M. A.: Automatic acquisition of hyponyms from large text corpora, in *Proceedings of the 14th conference on Computational linguistics-Volume 2*, pp. 539–545 (1992)

[Komachi 10] Komachi, M., Kudo, T., Shimbo, M., and Matsumoto, Y.: Semantic drift in espresso-style bootstrapping: Graph-theoretic analysis and evaluation in word sense disambiguation, *Transactions of the Japanese Society for Artificial Intelligence*, Vol. 25, pp. 233–242 (2010)

[McIntosh 11] McIntosh, T., Yencken, L., Curran, J. R., and Baldwin, T.: Relation guided bootstrapping of semantic lexicons, in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2*, pp. 266–270 (2011)

[Pantel 06] Pantel, P. and Pennacchiotti, M.: Espresso: Leveraging generic patterns for automatically harvesting semantic relations, in *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, pp. 113–120 (2006)

[Riloff 99] Riloff, E., Jones, R., et al.: Learning dictionaries for information extraction by multi-level bootstrapping, in *Proceedings of the Sixteenth National Conference on Artificial Intelligence*, pp. 474–479 (1999)

[Wu 17] Wu, S., Hsiao, L., Cheng, X., Hancock, B., Rekatsinas, T., Levis, P., and Ré, C.: Fonduer: Knowledge Base Construction from Richly Formatted Data, *arXiv preprint arXiv:1703.05028* (2017)

[Yan 02] Yan, X. and Han, J.: gspan: Graph-based substructure pattern mining, in *Data Mining, 2002. ICDM 2003. Proceedings. 2002 IEEE International Conference on*, pp. 721–724 (2002)

[Zhou 17] Zhou, K.: Jokeren/DataMining-gSpan (2017), https://github.com/Jokeren/DataMining-gSpan

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

アイテムのクラス情報を利用した非負値行列分解 NMF

Non-negative matrix factorization with information expansion using item class information

村上勝彦
Katsuhiko Murakami

東京大学 医科学研究所
The Institute of Medical Science, The University of Tokyo

Non-negative matrix factorization (NMF) is widely used for various problems, such as item recommendation of shopping, image recognition and bioinformatics. However, when the data are sparse, and the row and columns of the items have no association with others, these items tend to be independent resulting poor linking. Here we investigated a method to compensate such too sparse data, by adding information of hierarchical relationships among the items to the matrix analyzed. We show how the additional information helps to make desired new clusters. In addition, too much expansion of class information makes many items into the same clusters, the situation in which is not desired. Some ideas to avoid it are discussed.

1. はじめに

非負値行列分解(Non-negative matrix factorization; NMF)は、顔画像の要素の抽出や単語の分類といったテキストマイニングの方法として提案された[Lee 99]。その後、買い物での商品推薦、画像処理、バイオインフォマティクスなど様々な分野で応用されている。NMF の解析では、1つの非負値の要素からなる行列があたえられ、それを新たな2つの非負値行列の積で近似できるような行列を求める。それによって、もとの行列のなかの、複数行と複数列の中に存在する関連を見出す解析である。

買い物での商品推薦の場合、購入者と商品（アイテム）がそれぞれ行と列になり、購入点数が要素になるデータが解析対象となる。関連付けの根拠は共起頻度にもとづくため、購入事例の少ない個人や商品については、他と関連を見いだせることはあまりない。ほとんど買われない商品が全体の取引の多くを占めるというロングテールという現象でも知られるように、これらの割合は非常に高い。全商品のなかで評価値の予測が可能な商品の割合を被覆率 (coverage) と呼ぶが、ロングテールの状況では被覆率が下がってしまい、多くの商品について推薦ができないという問題があった。

NMF はバイオインフォマティクスの分野でも、遺伝子発現データの解析にも利用されている。NMF は遺伝子発現解析に使われた[Kim 03, Kim 03]。遺伝子とサンプル（細胞や個人を示す）の2次元軸に想定し、各行に遺伝子、各列にサンプルを配置し、行列要素の値が遺伝子発現量である「遺伝子発現行列」がまず作成される。これから行列分解によって、遺伝子セットとその基本発現量を決め、各サンプルはその組み合わせによって説明されるという解釈がされる。これはがん細胞の分類に有効であった[Brunet 04]。ここでは、概要をつかむという目的では有効であったが、他との関連が薄い遺伝子やサンプルを軽視しており、細部の情報を落としているという問題があった。

「遺伝子発現行列」を起点とした場合、この中から協調して発現が近い遺伝子群とサンプル群を取り出すという「バイクラスタリング」[Pontes 15]という手法も発展してきた。この方法に、分解する前後の行列要素に非負という制約を課したものが NMF に対応する。

我々は過去に遺伝子データベースの機能情報同士の関連性を見出すためにNMFを利用した関連抽出を行った [村上 15]。多くの遺伝子と機能の関連が見出されたものの、50％近くを占める多くの用語についてはつながらないままであった。このように NMF ではスパースなデータの場合に、多くの情報が活用できないことが課題である。

本稿では、このような見落としがちな事例の少ないアイテム（用語）についても、他のアイテムの関連を見出すため、アイテムのクラス情報を利用する方法について議論する。各アイテムはなんらかのクラスに属するので、アイテムのオントロジーデータを利用して上位クラスの用語へ一般化していけば、同じクラスとの関連が見いだせると考えられる。以下では、遺伝子と機能の関連性を見出す事例について検討した。

2. データと方法

遺伝子と機能のデータとして、ヒト遺伝子統合データベース H-InvDB (http://h-invitational.jp/) [Takeda 13]を利用した。遺伝子データベースにある記述のうち、遺伝子の機能については統制された用語で基本情報が記述されている。典型的なものとして遺伝子オントロジー（Gene Ontology; GO, GO タームとも呼ばれる）[Blake 2013] がある。ここでは全遺伝子のうち、GO タームが付与されたタンパク質をコードしている遺伝子(12,261遺伝子)を抜き出した。

2.1 遺伝子オントロジー(GO)

GO は用語の集合であり、各用語に少なくとも親子関係（IS-A 関係）が付与されているので、全用語は Directly Acyclic Graph (DAG)構造、階層構造を形成する（図1）。古典的には1遺伝子には1つの機能は1つではなく、複雑な機能を説明するべく、さまざまな観点から関連用語が付与されている。1つの遺伝子には、主要な情報として該当する GO が複数付与されている。あまり調べられていない遺伝子もあり、その場合は何の GO も付与されていない。付与される GO の階層レベルは特に決まりがなく上から下まで使われている。

木構造のルートは 3 種の概念、すなわち機能(molecular function)、生物学的な反応過程（biological process）、および細胞構成要素（cellular component）となっている。本来、生命現象は多くの要素が互いに絡み合っているので、これらの用語や概念は互いに繋がっているはずである。しかし、そのような繋がりはデータベースに陽にはかかれていない。そこで、そのような関係を NMF により発見して補完していく必要がある。

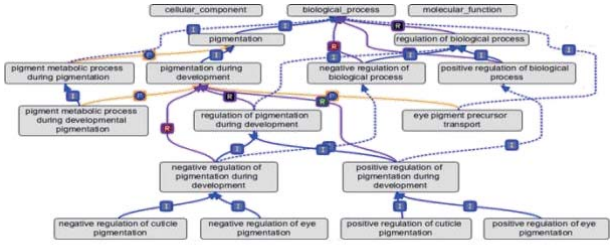


図1　Gene Ontology (GO) の階層構造。GO タームがノード、関係がリンクであり、全体がネットワークで表現されている。

2.2 遺伝子と機能の行列データと行列分解

ヒト遺伝子データベースから取得した、GO タームが付与された 12,261 遺伝子について、縦軸に遺伝子、横軸に用語とする行列を作成し、情報付与の有無を 1,0 で表現する（図2）。

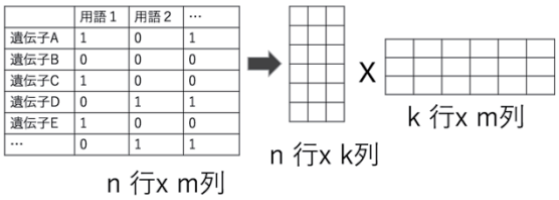


図2　遺伝子と GO タームのデータ構造と行列分解

ユニークな GO ターム回数（異なり数）は 1,741 個であり、延べ数は 40,871 個であった。これに付与されていた 533 個の GO タームを使用した。GO が付与されていた個数は1遺伝子あたり平均 3.3 個であった。通常の文書解析の場合と異なり、1遺伝子にある GO が付与されるのは0回か1回である。

この行列 V を式(1)のように2つの行列の積で近似する。

$$\mathrm{V_{i\mu}} \approx (WH)_{i\mu} = \sum_{a=1}^{k} W_{ia} H_{a\mu} \qquad (1)$$

ここで、V, W, H についてすべての行列要素が非負値である。行列の次元は、V が n×m、W が n×k、H が k×m である。また、k は基底ベクトルの数で、解析の際に固定値として与える。W と H の推定方法としては、以下のようなコスト関数を定義してその最小化を行う。

$$f = \|V - WH\|_F^2 \qquad (2)$$

右辺添字の F はフロベニウスのノルム(V と WH の同じ行と同じ列の要素毎に差をとり、その平方和)である。計算には R ライブラリ Non-Negative Linear Models (NNLM)[Lin 16] を使用した。

3. 関連を見出すための情報の拡張

上記の行列データの間に隠れた関連性を見出すことがここでの目標である。しかし、実際には行列はスパースである。1遺伝子あたりの平均 GO タームの数は3個である[村上 15]。通常の方法では、共起が少ない項目はノイズとして扱われて消えてしまうことになる。

ここでは、各列で要素値の合計（頻度）が t 回以下の低い頻度の GO タームについて、オントロジーの階層情報を参照して階層上位の GO タームを同定する。その GO タームはもともとその遺伝子には付与されていない（下位語が付与されているので上位語は付与する必要がない）。そこで階層が1つ上の上位語について図2の行列のフラグをたてる。こうして背景知識となるオントロジーから関係情報を取得して、データ行列に追加することで上位の GO タームにおける共起頻度が増加し、関連を拾いやすいデータ構造となった。

アイテム推薦の問題においても、商品のカテゴリについての階層データを作成し、同様にデータ拡張が適用できる。

付与した上位の GO タームについても、さらに階層レベルを上げて追加することも考えられるが、階層をあげすぎると、すべての用語について最上位のターム(「molecular function」など)がどの遺伝子にも付与されてしまって意味がない。適切なレベルの階層まで参照することが望ましい。この対策には相互情報量、または最大アイテム数か最大遺伝子数などの基準を設けることが考えられる。

4. おわりに

本稿では、スパースな行列データを NMF で解析する場合に、共起頻度の少ない孤立しがちなアイテム（ここでは GO ターム）について、クラス情報を付与することで、行列データの情報を増やす方法を提案した。これにより、多くの孤立アイテムについて従来は困難であった互いの関連性把握ができるようになった。

今後の課題は、行列要素値から関連遺伝子（および GO ターム）選定の自動化である。係数の大小によって選定するが、閾値を決定する客観的な方法が必要である。将来的には得られた知識データによる推論システムを構築したい。

参考文献

[Blake 13] Blake, J. a, Dolan, M., Drabkin, H., Hill, D. P., Li, N., Sitnikov, D., … Westerfield, M. Gene Ontology annotations and resources. *Nucleic Acids Research*, *41*(Database issue), D530-5. 2013.

[Brunet 04] Brunet, J.-P., Tamayo, P., Golub, T. R., & Mesirov, J. P. Metagenes and molecular pattern discovery using matrix factorization. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(12), 4164–9. 2004.

[Devarajan 08] Nonnegative matrix factorization: An analytical and interpretive tool in computational biology. *PLoS Computational Biology*, *4*(7). 2008.

[Kim 03] Kim, P. M., & Tidor, B. (2003). Subsystem Identification Through Dimensionality Reduction of Large-Scale Gene Expression Data. *Genome Research*, *13*(7), 1706–1718, 2003.

[Lee 99] Lee, D. D., & Seung, H. S. Learning the parts of objects by non-negative matrix factorization. *Nature*, *401*(6755), 788–791.1999.

[Lin 16] Lin, E. X. NNLM: A package For Fast And Versatile Nonnegative Matrix Factorization. (Web) 2016.

[Pontes 15] Pontes, B., Giráldez, R., & Aguilar-Ruiz, J. S. (2015). Biclustering on expression data: A review. *Journal of Biomedical Informatics*, *57*, 163–180.

[Takeda 13] Takeda, J.-I., Yamasaki, C., Murakami, K., Nagai, Y., Sera, M., Hara, Y., … Imanishi, T. H-InvDB in 2013: an omics study platform for human functional gene and transcript discovery. *Nucleic Acids Research*, *41*(Database issue), D915-9. 2013.

[村上 15] 村上勝彦. 行列因子分解による遺伝子データからの潜在的因子の抽出. In *The 29th Annual Conference of the Japanese Society for Artificial Intelligence*, (pp. 40–42), 2015.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

主成分分析による方策パラメータの低次元化を用いた直接方策探索の提案

Proposal of direct policy search using reduction of policy parameters by principal component analysis

村田 悠稀 Yuuki Murata　宮下 恵 Megumi Miyashita　矢野 史朗 Shiro Yano　近藤 敏之 Toshiyuki Kondo

東京農工大学
Tokyo University of Agriculture and Technology

In the sampling based direct policy search in reinforcement learning, higher dimensional decision variables causes the deterioration of optimal value and the slowing down of the learning speed. We clarified that the variance of the sampling probability distribution affects both for the optimal value and the learning speed. Especially, there exists the tradeoff between the optimal value and the learning speed. In this paper, we propose two trick to improve the learning speed without deteriorating the optimal value. First trick is to employ the small variance sampling distribution for improving the optimal value; It causes slower convergence as a side effect. As the second trick, we employed the dimensionality reduction of the decision variable for improving the learning speed.

1. はじめに

近年, 公共機関, 医療, 軍事的など, 様々な場面で活躍するロボットが開発される時代になってきており, 今では, 自律・自己学習型ロボットも開発されつつある.

強化学習とは, エージェントが試行錯誤を繰り返すことで, 環境からの累積報酬を最大にするような最適な行動指針（方策）を見つけだす機械学習の枠組みである. 最近では, 目的関数が未知の状態で決定変数を一つ決めると, 値が一つ求まるという条件のもとで, 値を最小化するような決定変数を見つける問題クラスであるブラックボックス最適化問題の考えになぞらえて強化学習を理解することが重要であると言われている [4]. ブラックボックス最適化問題における決定変数を方策とみなし, 出力を累積報酬だと見なせば, ブラックボックス最適化の流れにそって強化学習を理解することができる.

本研究では, 強化学習の枠組みの一つであるサンプリングベースの直接方策探索に着目して研究を行った. 直接方策探索とは, 方策を数理モデルで表現し, モデルのパラメータを最適化するものであり, この方策を更新する際に, ある確率分布からのサンプリングによって方策を更新する. G-MDS[1] や TRPO[7] などがこれに相当し, 一般的に確率分布は正規分布であることが多い. このサンプリングベースの直接方策探索には, 2 つの問題点がある. 1 つ目は, 方策を特徴づける方策パラメータを, 累積報酬を最大化することを指標として更新していくことになるが, この方策パラメータの次元が高次元になると, 学習速度や最適値が悪化することである. Funny[5] らが行った実験では, ロボットアームで対象物を掴むことを強化学習させたとき, 主成分分析 [6] を用いて低次元化した方が, 学習速度と最適値が共に向上したという結果を示している. 2 つ目は, サンプリングベースの直接方策探索では, サンプリングするための確率分布の分散によって, 最適値と学習速度に影響を及ぼしてしまうことである (詳細は 2.4 節を参照). 例えば, 分散を大きくすると, 学習速度は速くなるが最適値は悪化する. 分散を小さくすると, 逆の現象が起こる. すなわち, 分散によって, 最適値と学習速度がトレードオフになる.

本研究では, トレードオフをできるだけ解消して, 最適値を悪化させずに学習速度を向上させることを目的とし, 以下の 2 つの操作を行う.

1. 方策パラメータ高次元化による最適値悪化を解消する確率分布への操作
2. 1 の操作を行うことにより悪化した学習速度を向上させるために, 主成分分析による低次元空間を用いて決定変数の削減

今回はサンプリングベースの直接方策探索である G-MDS に提案手法を適用して従来手法との比較実験を行い, 提案手法の有用性を評価した.

連絡先: 村田悠稀，東京農工大学, murata@livingsyslab.org

2. 先行研究

2.1 MDS

強化学習においてコスト最小化問題を考えたとき, 目的関数 $\mathcal{J}$ を次の式 (1) で表せる.

$$\mathcal{J} = \sum_{j=1}^{M} p\left(\boldsymbol{\theta}_j\right) J\left(\boldsymbol{\theta}_j\right) \tag{1}$$

なお, $\boldsymbol{p} = [p(\theta_1), p(\theta_2), ..., p(\theta_M)]$ は方策パラメータ θ が $\theta = \theta_j$ をとる確率分布, $J(\boldsymbol{\theta}_j)$ は方策パラメータ θ_j の報酬であり, 条件は以下に示すとおりである.

$$\boldsymbol{p} \in \mathcal{P} := \left\{ x | x \in \mathbb{R}^M, \sum_{i}^{M} x_i = 1, x_i \geq 0 \right\} \tag{2}$$

ここで, 生起確率 $\boldsymbol{p}$ を決定変数とみなし, $\boldsymbol{p}$ を更新して最適な方策を獲得することを考えると, ブラックボックス最適化の観点から,

$$\boldsymbol{p}^* = \arg\min_{\boldsymbol{p} \in \mathcal{P}} \mathcal{J} \tag{3}$$

というような最小化問題となる.
また, $\boldsymbol{p}$ の更新方法として, 最適化手法の一つである鏡像降下法を用いる. 一般的な鏡像降下法 [8] の式が次の式 (4) である.

$$\beta_t = \arg\min_{\beta \in \mathcal{B}} \left\{ \langle g_t, \beta \rangle + \psi\left(\beta\right) + \eta_t B_\phi\left(\beta || \beta_{t-1}\right) \right\} \tag{4}$$

ここで, $\mathcal{B}$ はパラメータ空間, β はパラメータ, $g_t \in \delta(\beta_{t-1})$, η_t は更新幅を調節するハイパーパラメータである. 式 (4) の β_t を $\beta_t = p_t$ とすると,

$$p_t = \arg\min_{\boldsymbol{p}\in\mathcal{B}} \{\langle g_t, p(\boldsymbol{\theta})\rangle + \eta_t B_\phi\left(p(\boldsymbol{\theta})||p_{t-1}\right)\} \tag{5}$$

となり, $g_t = \nabla_p \mathcal{J} = J(\theta)$ であるため, 元の目的関数を微分しなくても更新を行うことが可能になる. これを MDS(Mirror Desent Search) と呼ぶ.

2.2 G-MDS

2.1 節の MDS をさらに発展させたのが, G-MDS(Gaussian-MDS)[1] である. まず, 式 (4) の鏡像降下法を発展させる. 具体的には, 式 (4) のブレグマン距離を KL 距離とする. つまり, 式 (4) の B_ϕ における ϕ を, $\phi(x_{t,j}) = \sum_{j=1}^{N} x_{t,j}\log(x_{t,j})\left(\boldsymbol{x}\in\mathbb{R}^N, \beta_{t,j} > 0\right)$ とすると, 更新式として式 6 の指数勾配降下法が求まる.

$$p_t(\theta_i) = \frac{\exp\left(-\eta_t g_{t,i}\right) p_{t-1}(\theta_i)}{\sum_{j=1}^{N}\exp\left(-\eta_t g_{t,j}\right) p_{t-1}(\theta_j)} \tag{6}$$

また, ここでは $p_t(\boldsymbol{\theta})$ の確率分布をガウス分布として仮定しており, $p_t(\boldsymbol{\theta})$ は,

$$p_t(\theta) = \mathcal{N}\left(\theta \mid \mu_{t-1}, \Sigma_{\varepsilon_{t-1,i}}\right) \tag{7}$$

となる. さらに, G-MDS では, $p(\theta)$ の平均値の更新についてのみ考えており, 平均値を式 (6) から計算することができる.

$$\mu_t = \sum_{i=1}^{N}\theta_i p_t(\theta_i) = \frac{\mathbb{E}_{p_{t-1}}[\theta_i \exp(-\eta_t g_{t,i})]}{\mathbb{E}_{p_{t-1}}[\exp(-\eta_t g_{t,i})]} \tag{8}$$

式 (8) にモンテカルロ積分を適用して, 推定される μ_t を $\hat{\mu_t}$ とすると, 以下の式となる.

$$\hat{\mu}_t = \frac{\frac{1}{N}\sum_{i=1}^{N}\theta_i \exp(-\eta_t g_{t,i})}{\frac{1}{N}\sum_{i=1}^{N}[\exp(-\eta_t g_{t,i})]} \tag{9}$$

さらに, θ_i を $\varepsilon_{t,i} \sim \mathcal{N}(0, \Sigma_{\varepsilon_{t-1,i}})$ のガウスノイズを用いて表記すると,

$$\theta = \mu_{t-1} + \varepsilon_{t-1,i} \tag{10}$$

と表すことができる. ちなみに, $\Sigma_{\varepsilon_{t-1,i}}$ は初期標準偏差である. 式 (9),(10) より,

$$\begin{aligned}\hat{\mu_t} &= \frac{\sum_{i=1}^{N}(\hat{\mu}_{t-1} + \varepsilon_{t-1,i})\exp(-\eta_t g_{t,i})}{\sum_{i=1}^{N}[\exp(-\eta_t g_{t,i})]} \\ &= \hat{\mu}_{t-1} + \frac{\sum_{i=1}^{N}\varepsilon_{t-1,i}\exp(-\eta_t g_{t,i})}{\sum_{i=1}^{N}[\exp(-\eta_t g_{t,i})]}\end{aligned} \tag{11}$$

となるので, ガウシアンの平均値をガウスノイズの重み付けによって更新が可能になる.

2.3 χ^2 分布に依存する最適値

確率変数 $X_1, ..., X_n$ が独立に標準正規分布に従うとき, その二乗和の確率変数 $X = X_1^2 + \cdots + X_n^2$ は自由度 n の χ^2分布 に従う [2]. G-MDS では, 生起確率 $p(\boldsymbol{\theta})$ の平均値を最適解に向けて更新していく. そして, $p(\boldsymbol{\theta})$ の平均値が最適解に収束するとき, $\mu_{t-1} = \mu_t = \theta^*$ となる. このとき, 方策パラメータ θ の推定二乗誤差を考えると, 式 (10) から

$$(\theta^* - \theta)^2 = \{\mu_t - (\mu_{t-1} + \varepsilon_{t-1,i})\}^2 = \|\varepsilon_{t-1,i}\|^2 \tag{12}$$

となることがわかる. ところで, 方策パラメータ θ は正規分布からのサンプリングなので, $\varepsilon_{t-1,i}$ も正規分布からのサンプリングであり, ノイズの次元は決定変数 $\theta \in \mathbb{R}^n$ に依存している. このことから $\varepsilon_{t-1,i}$ は自由度 n の χ^2 分布に従うことがいえる. χ^2 分布に従う確率変数の期待値は n(自由度 n) であることが知られているため,

$$\mathbf{E}\left[\|\boldsymbol{\varepsilon}_{t-1}\|^2\right] = dim(\boldsymbol{\varepsilon}) = n \tag{13}$$

となる. このことから, 最適化の収束時に決定変数の次元に比例したノイズの残差成分が上乗せされてしまうため, 決定変数の次元に応じて収束コストが悪化すると考えられる.

2.4 サンプリングの確率分布

G-MDS における更新則は式 (6) で表される式を用いており, 事前分布が次の式で表せるものとする.

$$p^t(\theta) \propto \frac{\exp\left(-\lambda^{-2}(\theta-\theta_t)\Sigma^{-1}(\theta-\theta_t)\right)}{Z} \tag{14}$$

ここで, λ^{-2} はガウス分布の分散に相当するものである. ここからさらに, $p^{t+1}(\theta)$ の更新式を求めると,

$$p^{t+1}(\theta) = \frac{\exp(-\eta J(\theta) - \lambda^{-2}(\theta-\theta_t)\Sigma^{-1}(\theta-\theta_t))}{Z} \tag{15}$$

となる. この式 (15) の λ の値が大きい (分散が大きい) と, $\eta J(\theta)$ の影響が強くなるので分布の更新幅が大きくなる. 逆に, λ の値が小さい (分散が小さい) と, 事前分布の影響が強くなるので分布の更新幅が小さくなる. まとめると, 分散が大きい場合は更新幅が大きくなるので, 学習速度が速くなり, 分散が小さい場合は更新幅が小さくなるので, 学習速度が速くなる.

3. 提案手法

3.1 最適値改善操作

2.3 節で説明したノイズの残差を打ち消すことで最適値の改善を図る. χ^2 分布に従う確率変数の期待値は決定変数に依存することがわかっているので, 正規分布からサンプリングされるノイズ ε に対して, $\sqrt{\frac{1}{k}}$ を掛け合わせる. このとき, $k = dim(\boldsymbol{\varepsilon})$ である. このような操作を加えることで, χ^2 分布に従うノイズを以下のように表すことができる.

$$\left(\frac{\varepsilon_1}{\sqrt{k}}\right)^2 + \left(\frac{\varepsilon_2}{\sqrt{k}}\right)^2 + \cdots + \left(\frac{\varepsilon_k}{\sqrt{k}}\right)^2 = \frac{1}{k}\left(\|\varepsilon_1\|^2 + \|\varepsilon_2\|^2 + \cdots + \|\varepsilon_k\|^2\right) \tag{16}$$

ここで, $\left(\|\varepsilon_1\|^2 + \|\varepsilon_2\|^2 + \cdots + \|\varepsilon_k\|^2\right) \to k$ となるので, この操作を加えることにより, 決定変数に依存しない値にできる. つまり, 決定変数の高次元化に伴う最適値の悪化を改善することが可能になる.

ところで, ノイズに対して係数倍する操作を行うことは, 確率分布の再生性 [2] よりガウス分布の分散に係数倍を行っていることと等価になる. よって, ノイズを以下で表すことができる.

$$\varepsilon_{t,i} \sim \mathcal{N}\left(0, \frac{1}{\sqrt{k}}\Sigma_{\varepsilon_{t-1,i}}\right) \tag{17}$$

このときの分散をみると, 最適値改善の操作を加えたときの方が分散がかなり小さくなることがわかり, さらにノイズの次元数が高次元になるにつれて, ガウシアンの分散が小さくなっていく. このことから, 最適値は改善されるが, ノイズの決定変数の高次元化により学習速度は大幅に低下することも引き起こされてしまう (2.4 節を参照).

3.2 低次元空間を用いた直接方策探索

3.1 節の操作を加えると残差成分が決定変数の次元に依存しないものになるが, 学習速度がさらに悪化してしまう. そこで, この悪化した学習速度を改善する操作として, 主成分分析によって獲得した低次元空間を用いて, 方策パラメータを計算する操作を行う.

あらかじめ収集した分析用データに対して主成分分析を行い, 最適方策パラメータについての主成分を算出する. そして, この主成分の上位主成分だけを抽出して低次元空間とする. この低次元空間の主成分とこの主成分をどの程度の比率で用いるのかを決定する決定変数 $\boldsymbol{\omega}$ を混合することにより, 方策パラメータを計算する. これを具体的に示すと以下のようになる.

$$(\boldsymbol{p_1}, \boldsymbol{p_2}, \cdots, \boldsymbol{p_k}) \cdot \boldsymbol{\omega} = \boldsymbol{\theta} \tag{18}$$

ここで, $\boldsymbol{p_j} = (p_{1,j}, p_{2,j}, \cdots p_{k,j})^{\mathrm{T}}$, $\boldsymbol{\omega} = (\omega_1, \omega_2, \cdots \omega_k)^{\mathrm{T}}$, $k = dim(\boldsymbol{\omega})$ であり, 従来の手法 $(G-MDS)$ では方策パラメータ $\boldsymbol{\theta}$ そのものが決定変数であるのに対し, 提案手法では方策パラメータを構成するための混合比 $\boldsymbol{\omega}$ が決定変数であることに注意されたい. 提案手法により, 決定変数の次元を低次元化することで, 学習速度の低下が緩和されることになる.

4. 実験

4.1 実験条件

本実験は, 多自由度アームの最適軌道学習を用いて評価実験を行なった. [3] で述べられている実験内容を引用しているが, シミュレータ内容を簡単に説明する. xy 座標上に描かれたアームに対して, 任意の数の関節を設定することが可能であり, 関節はアーム上に等間隔に設定され, エンドエフェクタは xy 座標上を自由に動くことが可能である. シミュレーションでは, 中継点と呼んでいる赤い点を通過したのち, ゴールに向かうように設定されており, 中継点の通過時間はスタートしてから 300ms 後, ゴール到達時間はスタートしてから 500ms と設定されている. 最終的に, 条件を満たすようなエンドエフェクタの軌道を学習することが目的となる. また, このシミュレータの報酬関数は以下の式で表される.

$$r_t = \frac{\sum_{i=1}^{d}(d-1-i)f^2}{\sum_{i=1}^{d}(d-1-i)} \tag{19}$$

$$r_{300ms} = 10^8((x_{viapoint} - x_{t_{300ms}})^2 + (y_{viapoint} - y_{t_{300ms}})^2) \tag{20}$$

ここで, d はアームの関節数, $(x_{viapoit}, y_{viapoint})$ は中継点の座標, f は各関節の加速度である. また, r_{300ms} は 300ms の時のエンドエフェクタと中継点の二乗誤差に係数倍したペナルティである.

なお, 主成分分析のための分析用データは, 中継点 (0.5, 0.5) を重心とする一片の長さが 0.2 の正方形の範囲からランダムに 100 個の中継点を選択して, それぞれの最適軌道の方策パラメータをあらかじめ学習し, まとめたデータを分析用データとした. そして, この分析用データに主成分分析を適用して獲得された低次元空間を元に再度軌道学習を行った. 実験パラメータは以下の通りである.

表 1: 実験パラメータ

中継点座標	(0.54, 0.44)
更新回数	10 万
ロールアウト	15
初期標準偏差	1

4.2 実験結果

決定変数 1000 次元問題を, 提案手法で決定変数を 500 次元, 100 次元に削減したときの収束速度の比較を行った. 比較したのが以下の図である.

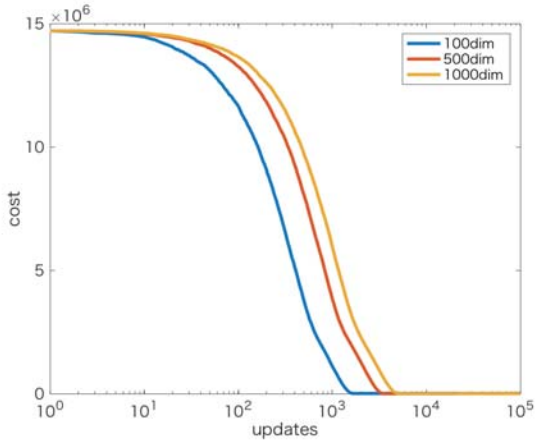


図 1: 学習速度の比較

1000 次元は G-MDS, 500 次元, 100 次元は提案手法を用いている. 学習速度を比較してみると, 1000 次元よりも 500 次元, 500 次元よりも 100 次元の方が学習速度が速くなっていることがわかる.

次に最適値の改善についてである. 最適値を比較したのが次の図である.

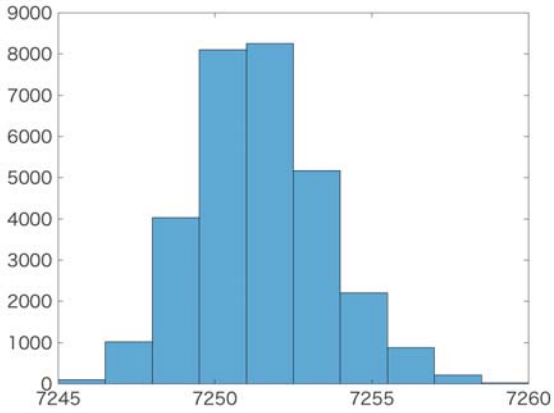


図 2: 終盤コストのヒストグラム (100 次元)

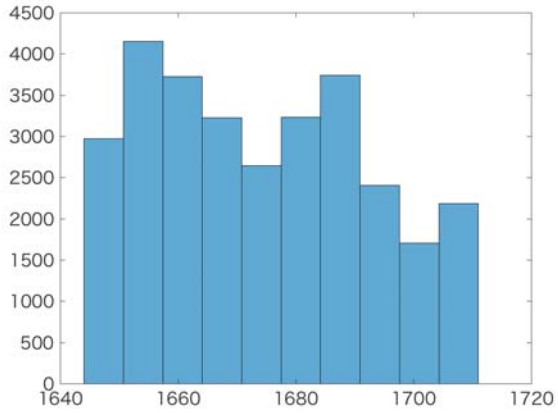


図 3: 終盤コストのヒストグラム (500 次元)

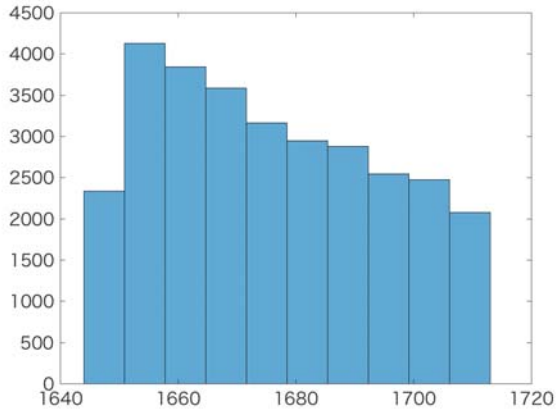


図 4: 終盤コストのヒストグラム (1000 次元)

各学習の更新回数の 7 万回〜最終までを集計してヒストグラムとして比較した. 決定変数 100 次元の場合 (図 2), 1000 次元 (従来手法)(図 4) と比べて明らかに最適値が悪くなってしまっているのがわかる. また, 決定変数 500 次元 (図 3) の場合でも, 従来手法と最適値がほとんど変わらないという結果が示された.

4.3 考察

従来手法で 1000 次元問題を解くのと, 提案手法で全主成分を使用して解くのは解釈がほぼ同じになる. このことから, 用いる主成分を削減した低次元では, 従来手法よりも情報量が少なくなり, 収束できる値に制限ができてしまうために, 最適値が悪化してしまったと考えられる. しかし, 今回の条件では, 決定変数を半分に減らしても最適値が従来手法とほぼ同じになるという結果も得ることができた. 最適値がほぼ同じであるならば, 学習速度が速い方が好ましい. 低次元空間を用いることの有用性に期待がもてる結果である.

5. まとめ

2 つの提案手法を従来手法に適用し, 評価実験を通して, 最適値と学習速度のトレードオフを従来よりも解消することに成功し, 最適値を悪化させずに学習速度を向上させることができた.

今後は, 強化学習と関係のある手法を MDS の観点からまとめ, 本研究の提案手法の適用範囲を広げていく. 低次元空間の拡張であるオンライン PCA などを提案手法に適用することで, 提案手法自体の拡張も行っていきたい.

謝辞

本研究は JSPS 科研費 JP17K12737, JP26120005, JP16H03219 の助成を受けたものです.

参考文献

[1] M. Miyashita, S. Yano, and T. Kondo, “Mirror descent search and acceleration, ” arXiv:1709.02535, 2017.

[2] 藤澤 洋徳 (2016), ”確率と統計.” 朝倉書店

[3] E. Theodorou, J. Buchli, S. Schaal, Reinforcement learning of motor skills in high dimensions: A path integral approach, in: Robotics and Automation (ICRA), 2010 IEEE International Conference on, IEEE, 2010, pp. 2397-2403.

[4] Stulp, Freek, and Olivier Sigaud. ”Policy improvement methods: Between black-box optimization and episodic reinforcement learning.” (2012).

[5] Ficuciello, Fanny, Damiano Zaccara, and Bruno Siciliano. ”Synergy-based policy improvement with path integrals for anthropomorphic hands.” Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on. IEEE, 2016.

[6] Bishop, Christopher M. ”パターン認識と機械学習 下.” ベイズ理論による統計的予測 (2008).

[7] Deisenroth, Marc Peter, Gerhard Neumann, and Jan Peters. ”A survey on policy search for robotics.” Foundations and Trends in Robotics 2.1-2 (2013): 1-142.

[8] 鈴木 大慈 (2015), ”確率的最適化.” 講談社

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

類似学習ケースに対するカリキュラム学習

Curriculum Learning for Similar Learning Cases

太田 悠太 *1　　滝 勇太 *1
Yuta Ota　　Yuta Taki

*1株式会社構造計画研究所
KOZO KEIKAKU ENGINEERING Inc.

In appricationsof machine learning, we often need to generate some models for multiple similar cases. In our study, we propose an efficient learning method exploiting the idea of the curriculum learning. Our idea is that we evaluate the effectiveness of a curriculum learning for one problem, then we use the learned curriculum as feedback for learning other cases. We apply our proposal method to two small problems, addition and subtraction of integers, using Sequence-to-Sequence model. Our experimental results show that the method works efficiently for some problem settings.

1. はじめに

機械学習の活用において，複数の類似する対象について，それぞれの学習モデルを構築したいことがある．

例えば，工場設備に対する異常検知モデルを構築したい場合，同じ種類の設備であっても設備の個体差や配備される工場の環境の差が存在するため，共通のモデルを構築するよりも，各設備に対してそれぞれのモデルを構築する方が望ましいことがある．また，生産ライン上の複数のロボットに動作を学習させたい場合，各ロボットの基本的な動作は類似しているが，担う機能によって必要な動作が部分的に異なるため，各ロボットに対してそれぞれの学習を行いたい場合が考えられる．

本研究では，このような一定の類似性を持つ複数の学習対象についてそれぞれのモデルを構築したい場合を考え，各学習を独立に行うのではなく，学習間で学習の効率性に関する情報を共有することで一連の学習を効率的に行う方法を提案する．より具体的には，カリキュラム学習 [Benjio 2009, Zaremba 2014, Wu 2017, Graves 2017, Matiisen 2017] を活用し，ある学習対象に対して行ったカリキュラム学習の効率性を評価し，類似する別の対象に関する学習を行う際に，その評価結果に基づくフィードバックを行う仕組みを提案する．

2. 関連研究

ここでは，カリキュラム学習に関する研究について説明を行う．カリキュラム学習とは，学習期間の短縮やモデルの高精度化のため，学習する事例の順序を工夫する学習アプローチを指す．

2.1 カリキュラム学習に関する研究

Benjio らは学習において相対的に容易な事例を最初に提示し，徐々に事例の難易度を高くすることで，ランダムに事例を提示するよりもモデルの予測精度が高くなることを画像分類，言語モデルのそれぞれの実験により示している [Benjio 2009]．ここでは基本的に難易度が上がっていくように事例を並び替えることでカリキュラムを生成しており，画像分類タスクでは難易度は目視で判断され，言語モデルでは単語数を難易度として並び替えを行なっている．

連絡先: 太田悠太，株式会社構造計画研究所，東京都中野区中央 5-4-3，03-5342-1125，yuta-ota@kke.co.jp

また，Zaremba らは計算プログラムのコードから計算結果を予測する問題を LSTM(long short-term memory) で学習する方法を提案する中で，従来の確率的勾配降下法による学習が困難な問題に対して，適切なカリキュラムによる学習戦略が重要であることを示している [Zaremba 2014]．カリキュラムの生成は，コードの長さと含まれる文法の種類数の 2 種類の難易度パラメータを考え，単純に難易度を上げていく戦略だけでなく，いずれかの難易度を固定する事例や，部分的にランダムに難易度を選択した事例を混ぜる戦略により，予測精度が高くなることを示している．

2.2 カリキュラムの自動生成に関する研究

Graves らは，Zaremba らのカリキュラムによる学習戦略に限界があることを指摘し，カリキュラムを自動生成する仕組みを構築した [Graves 2017]．より具体的には，敵対的バンディットの報酬として，学習の進捗率と学習対象のニューラルネット構造の複雑性の増加率を与え，ニューラルネットに対する学習において，学習エポックごとに提示する事例を選択するという枠組みを提示した．この枠組みを 3 つの実験で検証し、いくつかの設定で同じ精度になる学習時間を半分にできることを示した．また，Matiisen らは非定常バンディットを拡張した方法により，同様にカリキュラムを自動生成する仕組みを構築した [Matiisen 2017]．

これらの研究は，提示した事例に対する学習の進捗に基づき，次に提示する事例を選択するアルゴリズムを用いることで，カリキュラムの自動生成を可能にしている．

3. 提案手法

前節ではカリキュラム学習に関する研究について説明した．ここでは，類似性を持つ複数の学習対象について，各々のモデルを構築する問題設定を考える．そして学習を行うにあたり，ある学習対象に対してカリキュラム学習を行う過程で観測可能な「学習モデルがどのような状態の時に，どの事例を提供したら，どの程度学習が進捗したか」に関する情報を活用し，別の類似する学習対象の学習をより効率的に行う仕組みを提案する．

3.1 概要

本研究で提案する類似学習ケースに対するカリキュラム学習の仕組みを図 1 に示す．1 ケース目の学習は「事前学習」の位

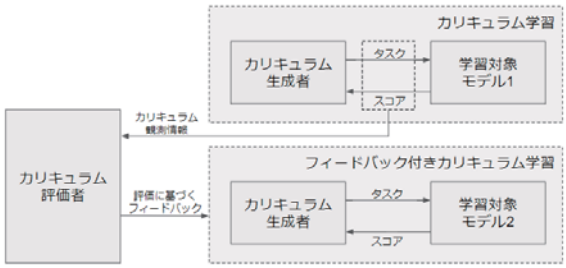


図 1: 類似学習ケースに対するカリキュラム学習の仕組みの概念図

置付けの学習であり，カリキュラムの自動生成 [Graves 2017, Matiisen 2017] により行われる (図 1 のカリキュラム学習)．2 ケース目 (もしあればそれ以降) の学習はターゲットとなる学習であり，カリキュラムの自動生成に加えて，1 ケース目のカリキュラムの効率性の評価に基づくフィードバックを踏まえたカリキュラム学習を行う (図 1 のフィードバック付きカリキュラム学習)．

ここでは，カリキュラムの自動生成機能を提供するクラスを「カリキュラム生成者」と呼び，事前のカリキュラム学習に対する評価機能とフィードバック機能を提供するクラスを「カリキュラム評価者」と呼ぶ．また，1 ケース目のカリキュラムの効率性に関する情報を「カリキュラム観測情報」と呼ぶ．

図 1 では 1 ケースのカリキュラム学習に対するカリキュラム観測情報を，別の類似する 1 ケースの学習に活用する状況を示しているが，3 ケース目以降の学習についても同様に評価とフィードバックを逐次的に行う仕組みを取ることも想定する．また，事前の複数のカリキュラム観測情報を合わせて評価してフィードバックを行う仕組みを取ることも想定する．

次節以降でこの仕組みの詳細を説明する．

3.2 カリキュラム学習

ここでは本研究で用いるカリキュラム学習の仕組みを示す．カリキュラム学習の仕組みは，[Graves 2017, Matiisen 2017] に従い，ここでは教師あり学習を考える．以下にその設定を示す．

目標系列として $Y_1, Y_2, \ldots$ を，入力系列として $X_1, X_2, \ldots$ を考える．系列は N 個のバッチに分割され，これをタスク D_k, $k = 1, \ldots, N$ とする．カリキュラム学習は T 回のバッチ学習により行われ，より少ない学習ステップで学習が収束することを目指して各学習ステップ t におけるタスクの選択を行う．このタスクのアンサンブルをカリキュラムと呼ぶ．

学習の目標として 2 つの設定が考えられる．1 つは全てのタスクが目標タスクである場合であり，この場合は，全てのタスクに対する学習の収束が目標となる．2 つ目は，興味のある任意のタスク $D_{\hat{k}}$ のみが目標タスクである場合である．この場合は $D_{\hat{k}}$ に対する学習の収束が目標となり，通常，$D_{k\neq\hat{k}}$ のタスクは，$D_{\hat{k}}$ に対する効率的な学習を達成するためのより難易度の低いタスクである．

3.3 カリキュラムの自動生成方法

カリキュラム生成者は各学習ステップ t でタスクを選択することで，カリキュラムを生成する．本研究ではカリキュラム生成者のカリキュラム自動生成方法として [Graves 2017] の仕組みを採用する．本節では [Graves 2017] のカリキュラムの自動生成方法を示す．

カリキュラム学習を T 回のタスク選択により総報酬 $\sum_{t=1}^{T} r^{(t)}$ を最大化するバンディット問題として考える．ここで，バンディット問題におけるアームの選択は時点 t におけるタスク D_k の選択，報酬 $r^{(t)}$ はタスク D_k を与えた時の学習の進捗率とする．このとき，各タスクに対する期待報酬は学習の進捗により変化するため，非定常な期待報酬を想定できる敵対的バンディットアルゴリズムの 1 つである Exp3.S アルゴリズム [Auer 2002] を用いる．Exp3.S アルゴリズムでは，報酬 $r^{(t)}$ に基づいて，タスク D_k を採用する確率に関する重みパラメータを学習ステップ t ごとに更新する．

各 t ステップ目の学習における目標タスクに対する損失を $\mathcal{L}^{(t)}$ とした時，報酬 $r^{(t)}$ を以下の式で与える：

$$r^{(t)} = \mathcal{L}^{(t)} - \mathcal{L}^{(t-1)}. \tag{1}$$

式 (1) は学習の進捗により取りうる範囲が変化するため，全期間における共通指標としては不適切である．従って，t ステップまでのすべての $r^{(t)}$ の 20 %四分位点を $q_{lo}^{(t)}$、80 %四分位点を $q_{hi}^{(t)}$ として、正規化を行った以下の指標 $\hat{r}^{(t)}$ を報酬として用いる：

$$\hat{r}^{(t)} = \begin{cases} -1 & \text{if} \quad r^{(t)} < q_{lo}^{(t)} \\ 1 & \text{if} \quad r^{(t)} > q_{hi}^{(t)} \\ \frac{2(r^{(t)} - q_{lo}^{(t)})}{q_{hi}^{(t)} - q_{lo}^{(t)}} - 1 & \text{otherwise.} \end{cases} \tag{2}$$

3.4 フィードバック付きカリキュラム学習

本節ではフィードバック付きカリキュラム学習 (FCL: Feedback-based Curriculum Learning) について説明する．FCL ではカリキュラム生成者によるカリキュラムの自動生成に加えて，カリキュラム観測情報に基づくフィードバックを踏まえた学習を行う．フィードバックの仕組みは以下の通りである．

いま，事前のカリキュラム学習により，カリキュラム観測情報が得られているものとする．式 (2) の報酬 $\hat{r}^{(t)}$ はその定義から $[-1, 1]$ の連続値をとるが，ここでは $\hat{r}^{(t)}$ の値に応じて $[-1, 1]$ を J 分割した離散的なスコア R_j, $j = 1, \ldots, J$ のいずれかをとるものとする．この時，カリキュラム観測情報は一期前に与えたタスク $D_k^{(\tau-1)}$，その時得られたスコア $R_j^{(\tau-1)} (\ni \hat{r}^{(\tau-1)})$，次に与えたタスク $D_l^{(\tau)}$ からなる状態 $S_{k,j,l} \equiv (D_k^{(\tau-1)}, R_j^{(\tau-1)}, D_l^{(\tau)})$, $k = 1, \ldots, N$, $j = 1, \ldots, J$, $l = 1, \ldots, N$ を持つ．カリキュラム学習の各ステップでは，式 (2) で与えられる報酬を得るが，その時の状態は $\{S_{k,j,l}\}$ のいずれかをとる．状態 $S_{k,j,l}$ をとった時の報酬の集合を $\mathcal{R}_{k,j,l} = \{\hat{r}^{(t_1)}, \ldots, \hat{r}^{(t_{T_{k,j,l}})}\}$ と表現する．ただし，$T_{k,j,l}$ は状態 $S_{k,j,l}$ をとった回数を表す．状態 $S_{k,j,l}$ を一度も取らなかった場合には，$\mathcal{R}_{k,j,l} = \phi$ となる．

カリキュラム評価者は，現在状態が $S_{k,j,\hat{k}}$ における価値を報酬集合 $\mathcal{R}_{k,j,\hat{k}}$ を用いて以下の式で計算する：

$$V_{k,j,\hat{k}} \equiv V(\mathcal{R}_{k,j,\hat{k}}) = \frac{1}{|\mathcal{R}_{k,j,\hat{k}}|} \sum_{r \in \mathcal{R}_{k,j,\hat{k}}} r$$

なお，$\mathcal{R}_{k,j,l} = \phi$ の場合は，$\mathcal{I} = \{(k, j, l') | \mathcal{R}_{k,j,l'} \neq \phi\}$ として，$\{V_{k,j,l'} | (k, j, l') \in \mathcal{I}\}$ の平均を用いる．

FCL の各ステップでは，一期前に選択したタスクが D_k，その際の報酬が R_j だった時，次に選択するタスク $D_{\hat{k}}$ の選択確

Algorithm 1 フィードバック付きカリキュラム学習

1: **for** $t = 1, \ldots, T$ **do**
2: カリキュラム生成者によりタスクごとの選択確率 P^G を計算
3: カリキュラム評価者によりタスクごとの選択確率 P^F を計算
4: $P \leftarrow (1-\alpha^{(t)})P^G + \alpha^{(t)}P^F$
5: 選択確率 P に基づいてタスク D_k を選択
6: 選択したタスク D_k により学習対象モデルを訓練
7: 目標タスクに対する損失 $\mathcal{L}^{(t-1)}$, $\mathcal{L}^{(t)}$ から報酬 $\hat{r}^{(t)}$ を計算
8: 報酬 $\hat{r}^{(t)}$ によりカリキュラム生成者の重みパラメータを更新
9: **end for**

率 $P^F_{k,j}(\hat{k})$ をフィードバックとして利用する．$P^F_{k,j}(\hat{k})$ は価値集合 $\{V_{k,j,s}\}_{s=1,\ldots,N}$ を用いて，

$$P^F_{k,j}(\hat{k}) = \frac{\exp V_{k,j,\hat{k}}}{\sum_{s=1}^{N} \exp V_{k,j,s}}, \tag{3}$$

により与える．

最終的にタスク $D_{\hat{k}}$ を選択する確率 $P_{k,j}(\hat{k})$ は，カリキュラム生成者の Exp3.S により得られるタスク選択確率 $P^G_{k,j}(\hat{k})$ および式 (3) により得られるフィードバックを用いて以下により計算する：

$$P_{k,j}(\hat{k}) \quad = \quad (1-\alpha^{(t)})P^G_{k,j}(\hat{k}) + \alpha^{(t)}P^F_{k,j}(\hat{k}).$$

ここで，$\alpha^{(t)}$ はカリキュラム評価者，カリキュラム生成者の各タスクの選択確率の平滑化パラメータである．$\alpha^{(t)}$ が 1 に近いほどカリキュラム評価者を信頼し，0 に近いほどカリキュラム生成者を信頼する．$\alpha^{(t)}$ が常に 0 の時，カリキュラム生成者が算出した確率 $P^G_{k,j}(\hat{k})$ にのみ従い，Graves らが提案した仕組みと一致する．

カリキュラム評価者によるフィードバックは，カリキュラム生成者が報酬に関する情報を充分に受け取っていない学習期間の序盤で，より有効であると考えられる．これを踏まえ，$\alpha^{(t)}$ は以下の式で与える：

$$\alpha^{(t)} = \frac{1}{1+\exp[-a(T_F - t)]},\ T_F > 0,\ a > 0.$$

ここで，T_F はカリキュラム評価者の選択確率をカリキュラム生成者の選択確率よりも重視する期間である．$T_F > t$ の場合は $\alpha^{(t)} > 0.5$ であるため，$P^F_{k,j}(\hat{k})$ が重視される期間となる．この期間をすぎると，徐々に $\alpha^{(t)}$ が減少する．また，a は $\alpha^{(t)}$ の減少の傾きの大きさを決める．a が小さいほど $\alpha^{(t)}$ の減少は緩やかに行われる．

ここまでの議論を踏まえ，FCL のアルゴリズムを Algorithm 1 に示す．

4. 計算機実験

4.1 実験設定

実験は，Sequence-to-Sequence モデル [Sutskever 2014] による加算問題 [Zaremba 2014] を対象とするもの，および，これを応用した減算問題を対象とするものの 2 種類を行う．加算問題は，文字列で与えられる'+' 記号で区分された 2 つの数字列を入力とし，2 つの数字列の和を出力とする．減算問題は'-' を用い，2 つの数字列の差を出力とする．入出力の例を表 1 に示す．

表 1: 入出力データ例

	入力	出力
例 1：加算問題	'29742+315'	'30057'
例 2：減算問題	'15413-8142'	'7271'
例 3：減算問題	'919-2313'	'-1397'

カリキュラムのタスク $\{D_k\}_{k=1}^N$ の設定は [Zaremba 2014, Matiisen 2017] の設定に倣う．加算 (もしくは減算) の入力における 2 つの数字列の最大桁数をタスクの難易度とする．表 1 の例 1, 2 は最大桁数 5 桁，例 3 は最大桁数 4 桁である．最大桁数 1 から 9 の 9 種類のタスクを用いることとし，各タスクにおいて 4,096 サンプルの入出力セットを生成する．各学習ステップでは，この 9 つのタスク $\{D_k\}_{k=1}^9$ から 1 つを選択して，モデルパラメータの更新を行う．タスクとは別に目標タスクとして最大桁数 9 桁の入出力セットを 128 サンプル生成する．モデルの損失 $\mathcal{L}^{(t)}$ はこの目標タスクを対象として算出する．損失は，正解の出力とモデルによる予測値の差として，128 サンプルに対する平均絶対偏差により計算する．

本実験では，提案手法である FCL を含む 4 つの手法，FCL, [Graves 2017] に従うカリキュラム学習 (CL)，各学習ステップにおいてランダムにタスクを選択する方法 (RANDOM)，すべての学習ステップにおいて目標タスクと同じ最大桁数 9 桁のタスクを選択する方法 (MAX)，を適用する．各手法に対して，目標タスクに対する損失の収束に要したステップ数，収束時の損失の大きさを比較する．FCL においてカリキュラム評価者が受け取るカリキュラム観測情報の内容は各実験の節に記載する．

実験において採用した各種パラメータは次の通りである．Sequence-to-Sequence モデルにおいて，encoder，decoder ともに 256 ユニットの LSTM を用い，RMSProp による最適化を行った．カリキュラム生成者が用いる Exp3.S については定式化を [Graves 2017] に従い，各パラメータは $\beta = 0.001$, $e = 0.02$, $\eta = 0.01$ とした．カリキュラム評価者はモデルの学習進捗率の状態を $[-1,\ 1]$ を 0.2 ステップで分割し，平滑化パラメータについては $T_F = 120$, $a = 0.2$ とした．

4.2 加算問題

各学習ステップにおける損失を図 2 に示す．学習はランダムシードを変更して 3 回行った．各学習について 20 ステップ分の移動平均をとった損失を求め，3 回の学習の平均損失の対数値を示している．

FCL におけるカリキュラム観測情報として，減算問題に対するカリキュラム学習をステップ 700 まで行い，損失の減少が大きいステップ 400 までをカリキュラム観測情報として利用した．また，カリキュラム観測情報が多く与えられるほど効率的な学習が達成されることを検証するため，カリキュラム学習 1 回分のカリキュラム観測情報を利用したときの学習を FCL(1)，3 回分のカリキュラム観測情報を利用したときの学習を FCL(3) としてそれぞれを実験対象とした．

結果として，FCL(1), FCL(3) は 100 ステップ目以降で CL と比較して損失減少の傾きが大きくなり，最終的に CL, MAX, RANDOM 手法に比べて，より小さい損失に収束していることが確認できる．このことは，フィードバックが有効であるこ

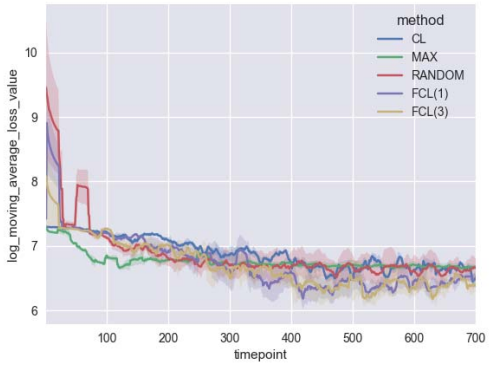


図 2: 加算問題における目標タスクに対する損失

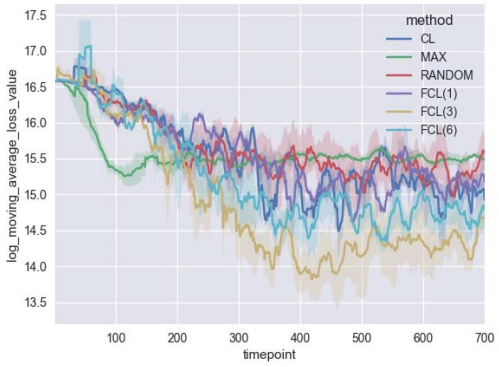


図 3: 減算問題における目標タスクに対する損失

とを示唆している．

なお，FCL(1) と FCL(3) は同程度の値に収束し，その収束にかかるステップ数も同程度であった．本実験からはカリキュラム観測情報の多寡による収束速度および損失への影響は明確に現れなかった．

4.3 減算問題

各学習ステップにおける損失を図 3 に示す．加算問題と同様に，学習はランダムシードを変更して 3 回行った．各学習について 20 ステップ分の移動平均をとった損失を求め，3 回の学習の平均損失の対数値を示している．

FCL におけるカリキュラム観測情報として，加算問題に対するカリキュラム学習をステップ 700 まで行い，損失の減少が大きいステップ 400 までをカリキュラム観測情報として利用した．カリキュラム学習 1 回分のカリキュラム観測情報を利用したときの学習を FCL(1)，3 回分のカリキュラム観測情報を利用したときの学習を FCL(3)，6 回分のカリキュラム観測情報を利用したときの学習を FCL(6) としてそれぞれを実験対象とした．

結果として，FCL(1)，CL は同程度の値に収束し，MAX，RANDOM より小さい損失に収束していることが確認できる．FCL(3) はステップ 150 程度から CL と比較して損失減少の傾きが大きくなり，より小さい損失に収束している．また，FCL(6) は 200 ステップ目以降で CL と比較して損失減少の傾きが大きくなり, 充分な差ではないものの，より小さい損失に収束している．これらはフィードバックが有効であることを示唆している．

この実験では，FCL(3) が最も小さい損失となり，FCL(6)，FCL(1) の順にこれに続く結果となった．前節の加算問題と同様に，カリキュラム観測情報は多く与えれば良いというわけではないことが確認された．従って，対象とするタスクに対して適切な量および質のカリキュラム観測情報を与えることが重要であると考えられる．

5. 結論

本研究では，類似性を持つ複数の学習ケースについてモデルを構築したい場合を考え，カリキュラム学習の仕組みを応用し，カリキュラム観測情報に基づくフィードバックによって類似ケースの学習をより効率的に行う手法を提案した．

実験により，提案手法について一定の効果がある可能性を確認することができた．カリキュラム観測情報の与え方によって収束速度および損失が変化することから，適切なカリキュラム観測情報の選択方法についても検討の必要がある．また，カリキュラム観測情報の結果への影響度合を検証することで,「学習対象の類似性」を明らかにし，フィードバックが有効に作用する条件などについても検証する必要がある．さらに，今回実施した実験設定が充分な規模と言い切れるものでないことから，今後はより大規模な実験設定や異なる問題に対する実験により，提案手法の有効性を検証する必要がある．

参考文献

[Auer 2002] Auer, P., Cesa-Bianchi, N., Freund, Y., and Schapire, R. E.: The nonstochastic multiarmed bandit problem, SIAM Journal on Computing, 32(1):pp. 48-77 (2002)

[Benjio 2009] Bengio, Y., Louradour, J., Collobert, R., and Weston, J.: Curriculum learning, ICML (2009)

[Graves 2017] Graves, A., Bellemare, M., Menick, J., Munos, R., and Kavukcuoglu, K.: Automated curriculum learning for neural networks, arXiv preprint arXiv:1704.03003 (2017)

[Matiisen 2017] Matiisen, T., Oliver, A., Cohen, T., Schulman, J., and Kavukcuoglu, K.: Teacher-Student Curriculum Learning, arXiv preprint arXiv:1707.00183 (2017)

[Sutskever 2014] Sutskever, I., Vinyals, O., and Le, Quoc, V.: Sequence to sequence learning with neural networks, In Advances in neural information processing systems, pp. 3104-3112 (2014)

[Wu 2017] Wu, Y., and Tian, Y.: Training agent for first-person shooter game with actor-critic curriculum learning, In Submitted to Int'l Conference on Learning Representations (2017)

[Zaremba 2014] Zaremba, W., and Sutskever, I.: Learning to execute, arXiv preprint arXiv:1410.4615 (2014)

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

グラフ断片決定木を用いたグラフ特徴抽出手法

Graph Feature Extraction Using Graph Fragment Decision Trees

坂上 陽規 Haruki Sakagami　瀧川 一学 Ichigaku Takigawa　有村 博紀 Hiroki Arimura

北海道大学 大学院情報科学研究科
Graduate School of Information Science and Technology, Hokkaido University

Graph-fragment decision trees are decision trees for graph classification, whose paths are graph operation codes. The GFDT algorithm is a top-down graph fragment decision tree learner without costly graph enumeration. In this paper, we study GFDT in the view of graph feature extraction. We compared GFDT based methods with pattern enumeration based methods in experiments on real-world datasets.

1. はじめに

1.1 背景と研究目的

機械学習を用いたグラフ分類問題について，近年多くの研究がなされている．グラフ分類ではグラフデータの特徴表現としてグラフパターンを用いるのが一般的であり，多くの場合では，学習の前処理として何らかの基準を満たすようなパターンの列挙を行う．例えば，部分グラフ列挙に基づく gSpan アルゴリズム（gS と表記する） [4] と決定木構築を合わせた手法（gS-DT と表記する）が考えられる．これらの前処理は基本的に重く，多様な構造学習応用のボトルネックになっている．本稿ではグラフ分類問題に対する新しいアプローチとして，グラフ断片決定木を用いた手法を考察する．

グラフ断片決定木: グラフ断片決定木 (GF 決定木) [5] は，グラフパターンの表現にグラフの符号化表現であるグラフコード [4] を利用する．決定木の各ノードはテストとしてグラフコードの拡張演算子を持ち，決定木のパスによって一意なグラフパターンに対応したグラフコードが表現される．学習アルゴリズム GFDT では通常の決定木のトップダウン構築法に基づいて決定木の構築を行う．gSpan [4] のように部分グラフの探索空間の網羅的な探索を行わないため，GFDT は，冒頭で述べた gS-DT アルゴリズムに比べて，高速かつメモリ使用量が小さいという利点をもつ．一方で，GFDT 単体の分類精度は，多少 gS-DT アルゴリズムに及ばないという問題がある．

1.2 研究結果

そこで，本稿では，GFDT アルゴリズムをグラフ特徴集合発見として用いて，集約学習手法と組み合わせて，より高精度のグラフ学習アルゴリズムを構成することを目指す．**集約学習** [3]（ensemble leanring）は，与えられた弱学習器と呼ばれるアルゴリズムを用いて多数の分類規則を生成し，それらによる多数の予測を統合して，より高精度な予測を行う枠組みである．集約学習の例として，ランダムフォレスト (random forest, 以降では RF と表記する) やブースティングがあり，広く利用されている [1].

具体的には，GF 決定木学習手法 GFDT を元にして，二つのグラフ分類手法を提案する．一つ目に，一般の RF に関する考察から，ランダム化した GF 決定木学習手法 rGFDT を開発し，これを弱学習器として RF と組み合わせた手法 rGFDT-RF を提案する．二つ目に，グラフ特徴抽出手法として，GFDT が学習した木のパス全体に対応する部分グラフ集合を取り出す手法 gfPath を開発し，これを弱学習器として RF と組み合わせた手法 gfPath-RF を提案する．

実データに対する実験では，これらの提案したグラフ分類の集約学習手法と，従来の gSpan を特徴抽出器として抽出した部分グラフパターンを弱学習器として RF と組み合わせた手法 gS-RF およびその変種と比較した．

実験結果からは，次のことがわかった．精度に関しては，提案の特徴抽出手法と RF を組み合わせた gfPath-RF は少ない特徴数で，頻出部分グラフ特徴を用いた gS[sup]-RF と gS[size]-RF に近い分類精度が得られた（実験 1 の表 4）．さらに，使用する特徴数の条件を揃えて比較した場合では，gfPath-RF の分類精度が大きく上回った（実験 2 の表 6）．

計算時間に関しては，GFDT ベースの提案手法の GFDT と，rGFDT-RF，gfPath-RF のいずれも，一部の例外を除き，gSpan ベースの RF よりも高速であった（実験 2 の表 5）．

まとめると，GFDT を用いたグラフ特徴抽出手法と集約学習の組合せは，軽量で比較的高性能なグラフ分類手法として，有用性があると考えられる．

連絡先: 坂上陽規，sakagami@ist.hokudai.ac.jp

2. 準備

本節では，以降で必要な概念と定義を導入する．可変長の要素のリストを $L = [a_1, \ldots, a_n]$ と表す．リスト L と，添え字 $i \leq n$，要素 b に対して，要素 $L[i] = a_i$ および接頭辞 $L[1..i] = a_1, \ldots, a_i$，連結 $L \cdot b = [a_1, \ldots, a_n, b]$ を定義する．キーと値の組 $a : v$ と書く．ハッシュ表 $H = \{a_1 : v_1, \ldots, a_n : v_n\}$ に対して，キー a_i に対する値を $H[a_i] = v_i$ で表す．以降で説明のない用語については，教科書 [3] を参照されたい．

2.1 グラフデータベース

$\mathcal{L} =$ をラベルの可算集合とする．$\mathcal{L}$ 上の**頂点ラベル付きグラフ**は，組 $G = (V(G), E(G), L_G)$ である．ここに，$V(G)$ は頂点の集合であり，$E(G) \subseteq V^2$ は無向辺の集合，$L_G : V(G) \to \mathcal{L}$ は各頂点にラベルを割り当てるラベル関数である．頂点付きグラフ G' が G の部分グラフに同型であるとき，$G' \sqsubseteq G$ と書く．

$\mathcal{GRAPH}$ と $\mathcal{LABEL} = \{0, \ldots, K-1\}$ で，それぞれ，$\mathcal{L}$ 上のラベル付きグラフ全体と**分類ラベル**全体を表す．サイズ m の**グラフデータベース**は，頂点ラベル付きグラフのリスト $\mathcal{G} = [\, G_i \mid i = 1, \ldots, m \,]$ と分類ラベルのリスト $\mathcal{Y} = [\, y_i \mid i = 1, \ldots, m \,]$ の組 $(\mathcal{G}, \mathcal{Y})$ である．$\mathcal{G}$ の**サンプル**は，任意の添え字集合 $S \subseteq \{1, \ldots, m\}$ である．

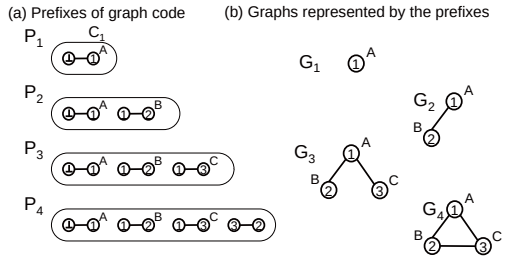


図 1: 例 1 のグラフコード P_4) の接頭辞 $P_i = P[1..i]$ と対応するグラフパターン $G_i = G(P_i)$ を示す $(1 \leq i \leq 4)$.

2.2 グラフコード

長さ K の**グラフコード**（graph code）とは，リスト $code = [op_1, \ldots, op_K]$ である．ここに，各要素 $op_i \in \mathcal{I}$ は**グラフ演算子**（graph operator）であり，与えられたグラフパターンを変更する操作を表し，以下のいずれかの形をとる．

- OP1 $op = (n_0, n_1, l_1)$: 連結頂点追加演算子は，グラフパターンにラベル l_1 を持つノード n_1 を加え，既存のノード n_0 とノード n_1 の間を辺で結ぶ操作を表す．
- OP2 $op = (n_0, n_1)$: 辺追加演算子は，グラフパターン上の既存のノードの対 n_0 と n_1 の間を辺で結ぶ操作を表す．

コード $code$ **が表現するグラフパターン** $H(code)$ は，空のグラフ $(\emptyset, \emptyset, \emptyset)$ に対し，$code$ 中の各演算子を先頭から順に適用して得られる頂点ラベル付きグラフである．

例 1 図 *1* にグラフコード $P = [(\bot, 1, A), (1, 2, B), (1, 3, C), (3, 2)]$ と，その接頭辞に対するグラフパターンの例を示す．

2.3 ランダムフォレスト

ランダムフォレスト（random forest, RF） [1] は，複数の弱学習器の決定木で得られた予測を統合して [*1]，全体の予測を行う集約学習手法である．RF では，弱学習器の集合は互いが多様性をもつことが望ましい．そこで，決定木の場合は次の二つの手法が広く用いられる [1,3].

特徴のランダム選択: 超パラメータを $\beta \in [0, 1]$ とする．弱学習器としての決定木構築において，テストとして用いる可能な特徴全体の集合 $\mathcal{OP}$ から，β の割合の部分集合 $\mathcal{OP}'$ をランダムに選択し，テストに用いる特徴を $\mathcal{OP}'$ の中から選ぶ．

サブサンプリング: 弱学習器として決定木を学習する際に，元の入力訓練データ集合から，定められたサイズの部分集合をランダムにサンプリングし，これらのデータから学習を行う．

3. 提案手法

定義 1（GF 決定木）グラフ断片決定木（graph fragment decision tree，GF 決定木）は，次の条件を満たす完全二分木 T の形をした決定規則である．(i) 根 $root$ は，親を持たない唯一の頂点である．(ii) 内部頂点 v は 1 子 $v.1$ と 0 子 $v.0$ をもち，頂点ラベルとして，グラフ演算子 $op(v) \in \mathcal{I}$ をもつ．(iii) 葉 v は，子をもたず，分類ラベル $y(v) \in \mathcal{LABEL}$ をラベルにもつ．

図 2 に，GF 決定木と，入力グラフに対する分類の例を示す．

*1 本稿では，全予測の多数決を用いる．

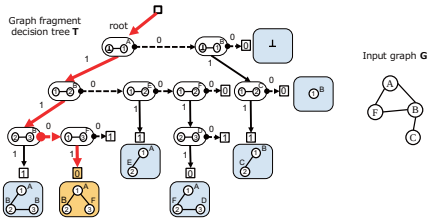


図 2: グラフ断片二分決定木の例

Algorithm 1 学習アルゴリズム GFDT

1: **Algorithm** GFDT($\mathcal{G}, minsup, \alpha, \varepsilon$)
2: Input: サイズ m のグラフデータベース $\mathcal{G}$, 最小支持度 $minsup \geq 0$，飽和しきい値 $\alpha \in [0, 1]$，利得しきい値 $\varepsilon \geq 0$ の組 $\Theta = (minsup, \alpha, \varepsilon)$
3: Output: 決定木の根ノード $root$
4: $S \leftarrow \{1, \ldots, m\}$
5: $OT \leftarrow \{\, i : [\,] \mid 1 \leq i \leq m \,\}$
6: **return** RecGFDT($S, OT, [\,], \mathcal{G}, \Theta$)

GF 決定木のすべての頂点 v に対して，次のように再帰的に，v **のグラフコード** $code(v) \in \mathcal{I}^*$ を対応づける: (i) 根 $root$ には，空列 $code(root) = [\,]$ を対応づける．(ii) もし深さ $k-1$ の内部ノード v にコード $code(v) = [op_1, \ldots, op_{k-1}]$ が対応付けられたならば，深さ k の子である 1 子には $code(v.1) = code(v) \cdot op(v)$ を，0 子には $code(v)$ を対応づける．

定義 2 (予測) GF 決定木 T の**予測アルゴリズム**は，与えられた入力グラフ $G \in \mathcal{GRAPH}$ に対して，T の根からスタートし，各内部ノード v において照合条件"$H(code(v)) \sqsubseteq G$?" をテストする．もしテストが成功すれば 1 子へ進み，失敗すれば 0 子へ進む．この操作を繰り返し，到達した葉ラベル w の持つ分類ラベル $\hat{y} = y(w)$ を予測値とする．以後，$Leaves(T)$ で，T の**葉全体**の集合を表す．

出現リスト: グラフパターン H のグラフデータベース $\mathcal{G}$ におけるマッチングの情報を効率的に保持するために，出現リストと出現テーブルを導入する．

グラフ におけるグラフパターン H の**出現リスト**（occurrence list）は，リスト $occlist = [occ_1, \ldots, occ_\ell]$ である．各 $1 \leq i \leq \ell$ に対して，occ_i は，**出現**と呼ばれる H の各頂点を G の各頂点に対応づける照合写像 $occ_i : V(H) \to V(G)$ である．

グラフデータベース $\mathcal{G}$ における H の**出現リスト表**は，ハッシュ表 $OT = \{\, k : occlist_k \mid k = 1, \ldots m \,\}$ である．

3.1 基本アルゴリズム GFDT

GFDT: GF 決定木の学習は，貪欲法でのトップダウン構築により行う．Algorithm 1 にトップレベルの手続き GFDT を，Algorithm 2 に主要な再帰手続き RecGFDT を示す．

再起手続き RecGFDT は，各ノードで (a) **葉生成演算**，(b) 内部ノードの下に二つの子を生成する**ノード分割演算**，(c) アンバランスなサンプル分割を伴う**枝伸長演算**（**飽和演算**とも呼ぶ）のいずれかを行う．

(a) 葉生成演算: **停止条件**として，(a.i) S の不純度スコアが 0 であるか，(a.ii) S の最良分割での利得が ε 未満であるか，(a.iii) S の最小支持度がしきい値 $minsup$ 未満であるか，のいずれかの条件が満たされたときに，葉を生成する．葉 v の予測ラベルには $S(v)$ の出力ラベル分布で最頻のラベルを用いる．

(b) ノード分割演算: ノードの分割では，親から受け継いだグラフコード $code$ に対し，グラフ演算子 op を用いて拡

Algorithm 2 再帰手続き RecGFDT

```
Algorithm RecGFDT(S, OT, code, G, Θ)
Output: 決定木のノード v
if 停止条件が成立 then {(a) 葉生成演算 }
    return 新しい葉 v
end if
OP(S, code) ← S に適用可能なグラフ演算全ての集合
飽和演算 (c) により OP(S, code) から飽和演算子 op を探す
if 飽和演算子が見つからなかったとき then
    OP(S, code) から分割スコア (b) が最大の演算子 op を探す
end if
S_1, S_0 ← S の code_1 := code · op による分割
OT_1 ← OT を code_1 で更新
v.1 ← RecGFDT(S_1, OT_1, code_1, G, Θ)
v.0 ← RecGFDT(S_0, OT, code, G, Θ)
return 新しい分岐頂点 v
```

張されたグラフコード $code_1 = code \cdot op$ を用いて，サンプル S を $S_1 = \{ i \in S \mid H(code_1) \sqsubseteq GSB[i] \}$ と $S_0 = \{ i \in S \mid H(code_1) \not\sqsubseteq GSB[i] \}$ に分割する．**グラフ演算子 op の探索**は，利得スコア $gainScore(S, S_1, S_0) := g(S) - (|S_1|/|S|) * g(S_1) + (|S_0|/|S|) * g(S_0)$ が最大になるように選ぶ．ここに，$g(S)$ は gini 不順度 [1] である．候補となる演算子 op は，$H(code \cdot op)$ が一つ以上の入力グラフにマッチするような適用可能な全ての演算子の集合 $\mathcal{OP}(S, code)$ から選ぶ．

(c) 飽和演算（枝伸長演算）：**飽和演算**（saturated search）では，与えられた正実数 $\alpha \in [0, 1]$ に対して，$|S_1| \geq \alpha|S|$ を満たす分割を行う演算子 $op \in \mathcal{OP}$ を直ちに選択する．条件を満たすものが複数あれば，ランダムにその中の一つを選択する．飽和演算には長いグラフコードを生成し，パターンの探索範囲を広げる効果が期待される．

3.2 ランダム化アルゴリズム rGFDT

本節では， 4. 節の実験で RF と組み合わせて用いる弱学習器 rGFDT を説明する．これは， 3.1 節の GFDT アルゴリズムを， 2.3 節の二つの方法でランダム化したものである．

特徴のランダム選択として，Algorithm 2 の 9 行目において，適用可能なグラフ演算子集合 $\mathcal{OP}$ の代わりに，部分集合 $\mathcal{OP}' \subseteq \mathcal{OP}$ を用いる．ただし，$\mathcal{OP}'$ は定数 $\beta \in (0, 1]$ に対して $|OP'| = \lceil \beta * |OP| \rceil$ となるように，ランダムに使用する特徴を選択したものを用いる．サブサンプリングに関しては，2.3 節のサブサンプリング法をそのまま用いる．

3.3 GFDT を用いたグラフ特徴抽出

本節では， 3.1 節の GFDT アルゴリズムを，グラフデータからの特徴抽出に使う手法 gfPath（graph fragment path）を説明する．GF 決定木の定義から，アルゴリズム GFDT が学習した GF 決定木の任意の葉は，根から葉までのパス（path）に対応するグラフコードを通して，一つの部分グラフパターンを表す．そこで，提案手法 gfPath では，入力の分類ラベル付きグラフデータベース $(\mathcal{G}, \mathcal{Y})$ から，GFDT を用いて，部分グラフ特徴からなる集合 $\mathcal{H}(\mathcal{G}, \mathcal{Y})$ を次のように抽出する．Θ を，GFDT の学習の超パラメータの組とする．

手続き gfPath$(\mathcal{G}, \mathcal{Y}, \Theta)$:

1. 入力の分類ラベル付きグラフデータベース $(\mathcal{G}, \mathcal{Y})$ に対し，$(\mathcal{G}, \mathcal{Y})$ から GF 決定木 T を GFDT で学習する
2. GF 決定木 T の全ての葉 $w \in Leaves(T)$ が表す部分グラフ $H(code(w))$ を特徴として抽出する．以上を繰り返して得られた部分グラフの集合を $\mathcal{H}(\mathcal{G}, \mathcal{Y})$ として返す．

表 1: 実験に用いたデータセット.

データセット	MUTAG	NCI1	ENZYMES
データ数	188	4110	600
クラス数	2	2	6
平均頂点数	17.93	29.87	32.63
平均辺数	19.79	32.30	62.14
頂点ラベルの種類数	7	36	3
最頻クラスの割合 (%)	66.49	50.05	16.67

GFDT アルゴリズムは，gSpan と異なり同型性判定を行っていないので，特徴集合 $\mathcal{H}(\mathcal{G}, \mathcal{Y})$ には互いに同形なグラフパターンが含まれる可能性があり，冗長である．

4. 実験

提案の GF 決定木を用いる手法と，既存手法 (頻出グラフ特徴を用いる手法) の比較を，実データ上の実験によって行う．

4.1 データ

表 1 に示した MUTAG, NCI1, ENZYMES の 3 つのベンチマークデータを，サイト [2] からダウンロードして用いた．

4.2 実験方法

実験では，次に示したグラフ特徴抽出器（グラフの弱学習器）と集約学習手法を組み合わせて，グラフ分類学習器を構成したものを用いた．以下では，記法 *A*-*B* で，グラフに対する弱学習器 A と集約学習手法 B を組み合わせて得られるグラフ分類学習器を表す．

グラフ特徴抽出器は，次の通りである．

- GFDT: 3.1 節のグラフ断片決定木 (GF 決定木) 学習手法の Python による実装.
- rGFDT: 3.2 節の GFDT のランダム化学習手法の Python による実装.
- gfPath: 3.3 節の GF 決定木を用いたグラフ特徴抽出アルゴリズムの Python による実装.
- gS: 頻出グラフ発見手法 gSpan [4] の C++による高速な実装 gBolt [*2]. ただし，並列処理は使用していない.

集約学習手法は次の通りである．

- RF: python の機械学習ライブラリ scikit-learn によるランダムフォレストの実装.

特徴抽出器 gSpan のパラメータである，グラフパターンの最小支持度 s と最大サイズ z を，それぞれのデータセットごとに表 2 に示した値に設定した．[*3] ここに，$gS[sup]$ と書かれた群は，z は制限せず，s を変化させたものを表し，一方で $gS[size]$ と書かれた群は，s は制限せず，z を変化させたものを表す．

gfPath と rGFDT-RF のパラメータは，飽和しきい閾値を $\alpha = 1$，終了条件しきい値を $\varepsilon = 0$, 特徴のランダム選択割合を $\beta = 0.5$ に設定した．集約学習手法について，RF の弱学習器の数は 50 とし，木の深さなどの制限は用いなかった．

分類精度は，異なる 5CV を 5 回繰り返した平均値を計測した．計算機環境として，PC (CPU Intel Corei7 3.3GHz，Memory 64GB, OS Ubuntu 16.04) と Python 3.5.2 を用いた．

*2 https://github.com/Jokeren/DataMining-gSpan

*3 パラメータ s および z は，手作業で，gSpan の計算時間が 200(sec) 以下で生成される特徴が 150 万個を超えないように選択した.

表 2: 実験で使用した gSpan の設定パラメータ．s と z は，それぞれ最小支持度と最大サイズを表す．

手法	MUTAG	NCI1	ENZYMES
gS[sup]	s=5, z=∞	s=205, z=∞	s=360, z=∞
gS[size]	s=1, z=17	s=1, z=10	s=1, z=7

表 3: グラフ特徴抽出器による生成した特徴数の比較．

手法	MUTAG	NCI1	ENZYMES
gfPath	38.3	1010.0	237.3
gS[sup]	782816.8	94995.6	276.6
gS[size]	1094054.4	727001.8	106802.8

表 4: 各手法と分類精度の比較．

手法	MUTAG	NCI1	ENZYMES
GFDT	81.10	75.02	39.13
rGFDT-RF	81.30	83.22	53.93
gfPath-RF	84.73	83.08	45.70
gS[sup]-RF	89.28	83.09	37.60
gS[size]-RF	87.04	83.94	51.83
baseline	66.49	50.05	16.67

4.3 実験 0： 各手法の分類性能の評価

gSpan と gfPath は，生成する特徴数に大きな差がある．予備実験として，二つのアルゴリズムを各データセットに対して実行し，生成した特徴数を計測した．

表 3 に得られた結果を示す．全データセットにおいて，gS[sup] と gS[size] は，非常に多くの特徴を生成している一方で，gfPath は少数の特徴を生成していることがわかった．これは，gfPath の利点でもあり，欠点でもある．

4.4 実験 1：各手法の分類性能の評価

グラフ特徴抽出アルゴリズム GFDT, rGFDT, gfPath, gSpan と集約学習手法 RF の組み合わせ手法による分類精度を，各データセットに対して計測した．ただし，GFDT は RF と組み合わせず，単一の学習器を用いる．表 4 に計測した分類精度を，表 5 に計測した実行時間を示す．

表 4 から，ENZYMES に対しては rGFDT-RF が最も精度が高く，MUTAG に対してはグラフ特徴抽出手法に gSpan を用いた gS[sup]-RF と gS[size]-RF の精度が高かった．gfPath-RF は gS[sup]-RF および gS[size]-RF と比較して全体的にやや精度を下回った．NCI1 では各集約学習手法内で大きな精度の差はなかった．表 5 からは，実行時間に関して属性抽出手法 GFDT と rGFDT，gfPath が高速であることが観察される．

4.5 実験 2：パターン特徴集合の評価

次に，gSpan で列挙されたグラフ特徴の数を，gfPath と同数に特徴数を削減して実験を行った．具体的には，gSpan で列挙されたグラフ特徴集合から，gfPath の特徴数との誤差が最小となるような整数 K を選択し，3 通りの特徴削減手法 (i) 訓練データに対する利得スコアが高い順に K 個選択する方法 (score)，(ii) 支持度が高いものから順に K 個選択する方法 (support)，(iii) ランダムに K 個選択する方法 (random) を用いた．gSpan の制約パラメータ $\mu \in \{\mathsf{sup}, \mathsf{size}\}$ と特徴選択法 $\nu \in \{\mathsf{score}, \mathsf{support}, \mathsf{random}\}$ の組み合わせを記法 gS[μ-ν] で表す．

特徴抽出に gfPath を用いた手法である gfPath-RF と，gSpan の特徴数を削減した手法である gS[μ-ν]-RF の各データセットに対する精度を比較した．表 6 に実験結果を示す．

表 5: 各手法の実行時間 (秒) の比較．括弧内の値は，グラフ特徴発見部分のみの実行時間を示す．

手法	MUTAG	NCI1	ENZYMES
GFDT	0.2	9.0	3.4
rGFDT-RF	10.1	557.6	215.9
gfPath-RF	0.3(0.2)	14.4(13.8)	26.3(26.3)
gS[sup]-RF	129.6(128.1)	222.1(197.4)	142.4(142.2)
gS[size]-RF	118.1(116.7)	123.5(32.8)	112.8(111.1)

表 6: 使用する特徴数を合わせた条件での各手法と分類精度の比較．

手法	MUTAG	NCI1	ENZYMES
gfPath-RF	84.73	83.08	45.70
gS[sup-score]-RF	78.11	73.06	37.17
gS[sup-support]-RF	66.38	80.73	36.93
gS[sup-random]-RF	64.99	79.03	32.83
gS[size-score]-RF	78.11	76.38	41.27
gS[size-support]-RF	66.28	81.21	37.33
gS[size-random]-RF	66.59	72.60	33.17
baseline	66.49	50.05	16.67

実験の結果，gfPath-RF の分類精度が，全てのデータセットについて最も高かった．この結果から，少数のグラフ属性しか用いない場合に，gfPath は高い識別能力をもつ特徴集合を発見していると考えられる．これは，解釈しやすい分類器の学習に有用と思われる．

5. おわりに

本稿では，グラフ断片決定木学習手法 GFDT の拡張として，集約学習器 rGFDT-RF と，グラフ特徴抽出アルゴリズム gfPath，集約学習器 gfPath-RF を提案し，実データに対する実験で評価した．実験結果から，GFDT ベースの手法の，頻出部分グラフ列挙による手法に対する一定の有用性が観察された．GFDT のグラフ特徴発見手法としての改良は今後の課題である．

謝辞

本研究は，JSPS 科研費 基盤研究 JPA16H01743，挑戦的萌芽研究 JP15K12022 と，北大 GI-CoRE GSB 拠点の助成と支援を部分的に受けました．ここに謝意を表します．

参考文献

[1] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[2] Kristian Kersting, Nils M. Kriege, Christopher Morris, Petra Mutzel, and Marion Neumann. Benchmark data sets for graph kernels, 2016.

[3] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of Machine Learning*. The MIT Press, 2012.

[4] Xifeng Yan and Jiawei Han. gspan: Graph-based substructure pattern mining. In *Proceedings 2002 IEEE International Conference on Data Mining.*, pages 721–724. IEEE, 2002.

[5] 坂上 陽規, 栗田 和宏, 瀧川 一学, and 有村 博紀. 決定化されたグラフパターントライの学習アルゴリズム. *SIG-FPAI*, B5(02):63–68, Jan 2018.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

エントロピー正則化付方策改善のための目的関数の補正

Objective Correction for Policy Improvement under Entropy Regularization

岩城 諒 *1　　浅田 稔 *1
Ryo Iwaki　　Minoru Asada

*1大阪大学大学院工学研究科 知能・機能創成工学専攻
Department of Adaptive Machine Systems, Graduate School of Engineering, Osaka University

Reinforcement learning aims to find a policy which maximizes long term future reward by interacting with unknown environment through trial and error. In this study, we propose an objective correction method for entropy regularized Markov decision process. After deriving a policy gradient under the regularization by the entropy and relative entropy, we propose an on-policy objective correction method for off-policy policy improvement under entropy regularization.

1. はじめに

強化学習は，未知の環境と試行錯誤的に相互作用しながら，意思決定則である方策を最適化することを目的とする．多くの場合，学習の目的関数は，設計者が報酬関数として定義した報酬の割引累積和である．環境と相互作用し学習サンプルを生成する挙動方策が，最適化したい推定方策と異なるとき，方策オフと呼ぶ．方策オフで学習可能であれば，過去の学習データの再利用が可能であり，サンプル効率よく学習できる [Lin 92, Mnih 15, Sugimoto 16]．

近年，エントロピーもしくは双対エントロピーによって目的関数を正則化する学習則が数多く研究されている．エントロピー正則化は，方策勾配法の興りとともに提案され [Williams 92]，深層強化学習においても有用性が示された [Mnih 16]．さらに近年，価値ベースの手法と方策ベースの手法が統一的に扱えることが示されてきた [O'Donoghue 16, Haarnoja 17, Nachum 17a, Schulman 17]．また，双対エントロピーによる正則化についても，様々な学習則が提案されてきた [Todorov 06, Todorov 10, Peters 10, Azar 12, Fox 16, Nachum 17b, Kozuno 17]．

一方で，(双対) エントロピーによる正則化を導入すると，目的関数と最適方策が元となるマルコフ決定過程と異なってしまうという問題がある．本研究では，エントロピーの正則化を利用した方策オフ型の学習則の一つである Path Consistency Learning (PCL) とその派生である trust-PCL[Nachum 17b] に着目し，これらの学習則を利用して元となるマルコフ決定過程の最適方策を求める手法を提案する．

2. 背景

2.1 マルコフ決定過程

マルコフ決定過程 (Markov Decision Process, MDP) における最適方策を獲得する問題を扱う．MDP は $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \rho_0, \gamma)$ の組によって特定される．$\mathcal{S}, \mathcal{A}$ はそれぞれ可能な状態と行動の集合である．離散時刻 $t \in \mathbb{N}_{\geq 0}$ において，エージェントは環境の状態 $s_t \in \mathcal{S}$ を観測し，行動 $a_t \in \mathcal{A}$ を選択する．環境の状態は，マルコフ性を有する状態遷移確率 $\mathcal{P} : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to \mathbb{R}$ に従って次状態 s_{t+1} に遷移する．エージェントは，有界な報酬関数 $\mathcal{R} : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ に従って環境から報酬を得る．ρ_0 は初期状態の分布，$\gamma \in [0, 1]$ は割引率である．エージェントの意思決定は，方策 π に従う．方策が決定論的であるとき，π は状態から行動への写像 $\pi : \mathcal{S} \to \mathcal{A}$ であり，方策が確率的であるとき，π はある状態である行動をとる確率 $\pi : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ である．方策 π と初期状態の分布 ρ_0 に対し，将来の割引き状態分布 $\rho^\pi(s) = \sum_{t=0}^{\infty} \gamma^t \Pr(s_t = s | \rho_0, \pi)$ が存在する．状態価値関数 $V^\pi(s) \triangleq \mathbb{E}_{\pi,\mathcal{P}}\left[\sum_{t=0}^{\infty} \gamma^t r_t \middle| s_0 = s\right]$ は，方策 π のもとである状態 s に期待される割引収益である．同様に，行動価値関数 $Q^\pi(s, a) \triangleq \mathbb{E}_{\pi,\mathcal{P}}\left[\sum_{t=0}^{\infty} \gamma^t r_t \middle| s_0 = s, a_0 = a\right]$ は，ある状態 s において行動 a をとり，その後方策 π に従った場合に期待される割引収益である．さらに，アドバンテージ関数を $A^\pi(s, a) \triangleq Q^\pi(s, a) - V^\pi(s)$ によって定義する．強化学習は，目的関数

$$\eta(\pi) = \sum_{s \in \mathcal{S}} \rho_0 V^\pi(s) = \sum_{s \in \mathcal{S}} \rho^\pi(s) \sum_{a \in \mathcal{A}} \pi(a|s) \mathcal{R}(s, a) \tag{1}$$

を最大化する決定論的な最適方策

$$\pi^* \in \arg\max_{\pi} \eta(\pi), \tag{2}$$

を獲得することである．方策の最適化としては様々な手法が提案されているが，パラメータ θ によって表現された方策 π_θ を考えるとき，方策勾配 $\nabla_\theta \eta(\pi)$ は以下のように与えられる [Sutton 99]:

$$\begin{aligned} \nabla_\theta \eta(\pi) &= \sum_{s \in \mathcal{S}} \rho^\pi(s) \sum_{a \in \mathcal{A}} \nabla_\theta \pi_\theta(a|s) Q^\pi(s, a) \\ &= \sum_{s \in \mathcal{S}} \rho^\pi(s) \sum_{a \in \mathcal{A}} \nabla_\theta \pi_\theta(a|s) A^\pi(s, a). \end{aligned} \tag{3}$$

2.2 エントロピー正則化付マルコフ決定過程

本研究ではさらに，エントロピー正則化付マルコフ決定過程 (Entropy-Regularized Markov Decision Process, ERMDP) を扱う．§2.1 で導入した MDP に加え，$\tau \geq 0, \lambda \geq 0$ として，エントロピー正則化付状態価値関数を定義する:

$$\begin{aligned} V^\pi_{\tau,\lambda}(s) \triangleq \mathbb{E}_{\pi,\mathcal{P}}\Bigg[\sum_{t=0}^{\infty} \gamma^t \Big(\mathcal{R}(s_t, a_t) - \tau \ln \pi(a_t|s_t) \\ - \lambda \ln \frac{\pi(a_t|s_t)}{\bar{\pi}(a_t|s_t)} \Big) \Bigg| s_0 = s \Bigg]. \end{aligned} \tag{4}$$

連絡先: 岩城諒，ryo.iwaki@ams.eng.osaka-u.ac.jp

ただし，$\bar{\pi}$ は参照方策である．エントロピー正則化付状態価値関数は以下のベルマン方程式を満たす:

$$V^{\pi}_{\tau,\lambda}(s) = \sum_{a\in\mathcal{A}} \pi(a|s)\Big(\mathcal{R}(s,a) + \gamma(\mathcal{P}V^{\pi}_{\tau,\lambda})(s,a) - \tau\ln\pi(a|s) - \lambda\ln\frac{\pi(a|s)}{\bar{\pi}(a|s)}\Big). \quad (5)$$

ただし，

$$(\mathcal{P}V^{*}_{\tau,\lambda})(s,a) = \sum_{s'\in\mathcal{S}} \mathcal{P}(s'|s,a)V^{*}_{\tau,\lambda}(s')$$

とした．ERMDP での目的関数は，以下で与えられる:

$$\eta_{\tau,\lambda}(\pi) = \sum_{s\in\mathcal{S}} \rho_0(s)V^{\pi}_{\tau,\lambda}(s) = \eta(\pi) + \tau H_{\gamma}(\pi) - \lambda D_{\gamma}(\pi,\bar{\pi}). \quad (6)$$

ただし，

$$H_{\gamma}(\pi) = -\sum_{s\in\mathcal{S}} \rho^{\pi}(s)\sum_{a\in\mathcal{A}} \pi(\cdot|s)\ln\pi(\cdot|s),$$
$$D_{\gamma}(\pi,\bar{\pi}) = \sum_{s\in\mathcal{S}} \rho^{\pi}(s) D_{\mathrm{KL}}(\pi(\cdot|s),\bar{\pi}(\cdot|s))$$

である．

ERMDP において，最適価値 $V^{*}_{\tau,\lambda}(s)$ と最適方策 $\pi^{*}_{\tau,\lambda}$ に対し，以下の命題が成立する．

命題 1. [Nachum 17a, Nachum 17b, Kozuno 17] 目的関数 (6) の元で，最適価値 $V^{*}_{\tau,\lambda}(s)$ と最適方策 $\pi^{*}_{\tau,\lambda}$ は，全ての状態行動の組に対し以下の方程式を満たす:

$$\begin{aligned} V^{*}_{\tau,\lambda}(s) &= \mathcal{R}(s,a) + \gamma\sum_{s'\in\mathcal{S}} \mathcal{P}(s'|s,a)V^{*}_{\tau,\lambda}(s') - (\tau+\lambda)\ln\pi^{*}_{\tau,\lambda}(a|s) + \lambda\ln\bar{\pi}(a|s) \\ &= \mathbb{E}_{s'\sim\mathcal{P}}\big[\mathcal{R}(s,a) + \gamma V^{*}_{\tau,\lambda}(s') - (\tau+\lambda)\ln\pi^{*}_{\tau,\lambda}(a|s) + \lambda\ln\bar{\pi}(a|s)\big]. \end{aligned} \quad (7)$$

すなわち，最適方策は以下で与えられる:

$$\pi^{*}_{\tau,\lambda}(a|s) = \frac{\bar{\pi}(a|s)^{\frac{\lambda}{\tau+\lambda}}\exp\left(\left(\mathcal{R}(s,a) + \gamma(\mathcal{P}V^{*}_{\tau,\lambda})(s,a)\right)/(\tau+\lambda)\right)}{\exp\left(V^{*}_{\tau,\lambda}(s)/(\tau+\lambda)\right)}. \quad (8)$$

さらに，全ての状態行動の組に対し，価値関数 V と方策 π が方程式 (7) を満たすとき，その価値関数と方策はそれぞれ最適価値と最適方策である: $\pi = \pi^{*}_{\tau,\lambda}, V = V^{*}_{\tau,\lambda}$.

さらに，命題 1 は次のように複数ステップ予測 (multi-steps prediction) に拡張できる．

系 2. [Nachum 17b] $n\in\mathbb{N}_{\geq 1}$ とする．目的関数 (6) の元で，最適価値 $V^{*}_{\tau,\lambda}(s)$ と最適方策 $\pi^{*}_{\tau,\lambda}$ は，任意の状態 s_t から開始される全ての状態行動系列に対し以下の方程式を満たす:

$$V^{*}_{\tau,\lambda}(s_t) = \mathbb{E}_{s_{t+i}\sim\mathcal{P}}\Big[\gamma^n V^{*}_{\tau,\lambda}(s_{t+n}) + \sum_{i=0}^{n-1}\gamma^i\big(\mathcal{R}(s_{t+i},a_{t+i}) - (\tau+\lambda)\ln\pi^{*}_{\tau,\lambda}(a_{t+i}|s_{t+i}) + \lambda\ln\bar{\pi}(a_{t+i}|s_{t+i})\big)\Big]. \quad (9)$$

式 (7), (9) ともに，期待値は状態遷移確率に従う後続状態のサンプリングのみに対して定義されていて，方策に依存しない．よって，Path Consistency Learning (PCL)[Nachum 17a] と Trust-PCL[Nachum 17b] は，Q 学習 [Watkins 89, Watkins 92] と同様に，式 (9) の両辺の残差を誤差信号としてそれを最小化するよう学習する．すなわち，それぞれパラメータ θ と ϕ で表現された方策 $\pi_{\theta}(a|s)$ と価値関数 $V_{\phi}(s)$ に対し，以下のようにパラメータを更新する:

$$\begin{aligned} \delta_{t:t+n}(\theta,\phi) &= -V_{\phi}(s_t) + \gamma^n V_{\phi}(s_{t+n}) + \sum_{i=0}^{n-1}\gamma^i\big(\mathcal{R}(s_{t+i},a_{t+i}) - (\tau+\lambda)\ln\pi_{\theta}(a_{t+i}|s_{t+i}) + \lambda\ln\bar{\pi}(a_{t+i}|s_{t+i})\big), \\ \Delta\theta &\propto -\frac{1}{2}\nabla_{\theta}\delta^2_{t:t+n}(\theta,\phi) = \delta_{t:t+n}(\theta,\phi)\sum_{i=0}^{n-1}\gamma^i\nabla_{\theta}\ln\pi_{\theta}(a_{t+i}|s_{t+i}), \quad (10) \\ \Delta\phi &\propto -\frac{1}{2}\nabla_{\phi}\delta^2_{t:t+n}(\theta,\phi) = \delta_{t:t+n}(\theta,\phi)\left(\nabla_{\phi}V_{\phi}(s_t) - \gamma^n\nabla_{\phi}V_{\phi}(s_{t+n})\right). \end{aligned}$$

ただし，$\tau\neq 0\wedge\lambda = 0$ のときが PCL，$\tau\neq 0\wedge\lambda\neq 0$ のときが Trust-PCL である．

3. エントロピー正則化付方策勾配

エントロピー正則化付行動価値関数を，以下のように定義する:

$$Q^{\pi}_{\tau,\lambda}(s,a) \triangleq \mathbb{E}_{\pi,\mathcal{P}}\left[\sum_{t=0}^{\infty}\gamma^t\left(\mathcal{R}(s_t,a_t) - \tau\ln\pi(a_t|s_t) - \lambda\ln\frac{\pi(a_t|s_t)}{\bar{\pi}(a_t|s_t)}\right)\middle| s_0 = s, a_0 = a\right]. \quad (11)$$

エントロピー正則化付行動価値は以下のベルマン方程式

$$Q^{\pi}_{\tau,\lambda}(s,a) = \mathcal{R}(s,a) + \gamma\sum_{s'\in\mathcal{S}}\mathcal{P}(s'|s,a)\sum_{a'\in\mathcal{A}}\pi(a'|s')Q^{\pi}_{\tau,\lambda}(s',a') - \tau\ln\pi(a|s) - \lambda\ln\frac{\pi(a|s)}{\bar{\pi}(a|s)} \quad (12)$$

を満たし，エントロピー正則化付状態価値との間に以下の関係が成立する:

$$V^{\pi}_{\tau,\lambda}(s) = \sum_{a\in\mathcal{A}}\pi(a|s)Q^{\pi}_{\tau,\lambda}(s,a), \quad (13)$$

$$Q^{\pi}_{\tau,\lambda}(s,a) = \mathcal{R}(s,a) + \gamma(\mathcal{P}V^{\pi}_{\tau,\lambda})(s,a) - \tau\ln\pi(a|s) - \lambda\ln\frac{\pi(a|s)}{\bar{\pi}(a|s)}. \quad (14)$$

さらに，エントロピー正則化付アドバンテージを以下のように定義する:

$$A^{\pi}_{\tau,\lambda}(s,a) = Q^{\pi}_{\tau,\lambda}(s,a) - V^{\pi}_{\tau,\lambda}(s).$$

MDP におけるアドバンテージ A^π と同様に，方策 π についての $A^\pi_\tau(s,a)$ の期待値は 0 である:

$$\begin{aligned}\sum_{a\in\mathcal{A}}\pi(a|s)A^\pi_{\tau,\lambda}(s,a) &= \sum_{a\in\mathcal{A}}\pi(a|s)\left(Q^\pi_{\tau,\lambda}(s,a)-V^\pi_{\tau,\lambda}(s)\right)\\ &= \sum_{a\in\mathcal{A}}\pi(a|s)Q^\pi_{\tau,\lambda}(s,a)-V^\pi_{\tau,\lambda}(s)\\ &= 0.\end{aligned}$$

最後の等号は (13) によって成立する．さらに，式 (5) に基づき，エントロピー正則化付 Temporal Difference (TD) 誤差 δ^π_τ を，以下のように定義する:

$$\begin{aligned}\delta^\pi_{\tau,\lambda}(s,a,s') = \mathcal{R}(s,a)-\tau\ln\pi(a|s)-\lambda\ln\frac{\pi(a|s)}{\bar\pi(a|s)}\\ +\gamma V^\pi_{\tau,\lambda}(s')-V^\pi_{\tau,\lambda}(s).\end{aligned}$$

TD 誤差 δ^π_τ の期待値は，アドバンテージ $A^\pi_\tau(s,a)$ に一致する:

$$\begin{aligned}\bar\delta^\pi_{\tau,\lambda}(s,a) &\triangleq \mathbb{E}_{s'\sim\mathcal{P}(\cdot|s,a)}\left[\delta^\pi_{\tau,\lambda}|s,a\right]\\ &= \mathcal{R}(s,a)-\tau\ln\pi(a|s)-\lambda\ln\frac{\pi(a|s)}{\bar\pi(a|s)}\\ &\quad+\gamma\mathbb{E}_{s'\sim\mathcal{P}(\cdot|s,a)}\left[V^\pi_{\tau,\lambda}(s')\right]-V^\pi_{\tau,\lambda}(s_t)\\ &= Q^\pi_{\tau,\lambda}(s,a)-V^\pi_{\tau,\lambda}(s)\\ &= A^\pi_{\tau,\lambda}(s,a).\end{aligned}$$

以下の命題は，ERMDP における目的関数 (6) を最大化するための最急勾配方向を与える．

命題 3. ERMDP における目的関数 (6) についての方策勾配は以下で与えられる:

$$\nabla_\theta\eta_{\tau,\lambda}(\pi)=\sum_{s\in\mathcal{S}}\rho^\pi(s)\sum_{a\in\mathcal{A}}\nabla_\theta\pi_\theta(a|s)\left(Q^\pi_{\tau,\lambda}(s,a)-b(s)\right).$$

命題 3 は，MDP における方策勾配定理 [Sutton 99] と同様の手順で，式 (12)-(14) を利用して証明できる．さらに，従来の方策勾配法と同様に，$b(s)=V^\pi_{\tau,\lambda}(s)$ とおくことで以下を得る:

$$\begin{aligned}\nabla_\theta\eta_{\tau,\lambda}(\pi) &= \sum_{s\in\mathcal{S}}\rho^\pi(s)\sum_{a\in\mathcal{A}}\nabla_\theta\pi_\theta(a|s)\left(Q^\pi_{\tau,\lambda}(s,a)-V^\pi_{\tau,\lambda}(s)\right)\\ &= \sum_{s\in\mathcal{S}}\rho^\pi(s)\sum_{a\in\mathcal{A}}\nabla_\theta\pi_\theta(a|s)A^\pi_{\tau,\lambda}(s,a)\\ &= \sum_{s\in\mathcal{S}}\rho^\pi(s)\sum_{a\in\mathcal{A}}\nabla_\theta\pi_\theta(a|s)\bar\delta^\pi_{\tau,\lambda}(s,a).\end{aligned}\tag{15}$$

4. 目的関数の補正

エントロピーによって正則化する場合，PCL などの方策オフ学習が可能である一方で，目的関数 (6) は元の MDP の目的関数 (1) と異なる．さらに得られる最適方策 (8) はソフトマックス関数であるため，元の MDP の決定論的な最適方策 (2) とは異なり，ERMDP での最適方策の”良さ”は，参照方策 π の選び方に依存する．本節では，ERMDP で定式化される学習則を MDP での目的関数最大化に利用する方法について議論する．

命題 4. MDP における方策勾配 (3) は，ERMDP におけるエントロピー正則化付方策勾配を利用して以下のように表現できる:

$$\begin{aligned}&\nabla_\theta\eta(\pi)\\ &=\sum_{s\in\mathcal{S}}\rho^\pi(s)\sum_{a\in\mathcal{A}}\nabla_\theta\pi_\theta(a|s)\bar\delta^\pi_{\tau,\lambda}(s,a)\\ &\quad+\sum_{s\in\mathcal{S}}\rho^\pi(s)\sum_{a\in\mathcal{A}}\nabla_\theta\pi_\theta(a|s)\big((\tau+\lambda)\ln\pi_\theta(a|s)-\lambda\ln\bar\pi(a|s)\big)\\ &=\mathbb{E}_{\pi,\rho^\pi}\left[\nabla_\theta\ln\pi_\theta(a|s)\bar\delta^\pi_{\tau,\lambda}(s,a)\right]\\ &\quad+\mathbb{E}_{\pi,\rho^\pi}\left[\nabla_\theta\ln\pi_\theta(a|s)\big((\tau+\lambda)\ln\pi_\theta(a|s)-\lambda\ln\bar\pi(a|s)\big)\right].\end{aligned}\tag{16}$$

命題 4 は，式 (6) の両辺を θ について偏微分し，命題 4 を利用することで容易に示される．式 (16) の初項はエントロピー正則化における方策勾配 (15) であり，第二項は ERMDP における方策勾配を MDP における方策勾配へと補正する．さらに，初項の $\nabla_\theta\ln\pi_\theta(a|s)\bar\delta^\pi_{\tau,\lambda}(s,a)$ は PCL による方策の更新方向 (10) において，状態価値関数を現在の方策に関するエントロピー正則化付状態価値 $V^\pi_{\tau,\lambda}$ に置き換え，さらに $n=1$ とした場合と等価である．

β を挙動方策とする．$\bar\delta^\pi_{\tau,\lambda}$ を $\delta_{t:t+1}(\theta,\phi)$ で置き換え，PCL と同様に，式 (16) の初項の期待値を β によって生成されたサンプルから計算することで，以下の方策オフと方策オンのサンプリングが混合した方策勾配を得る:

$$\begin{aligned}&\nabla_\theta\eta_\beta(\pi)\\ &=\mathbb{E}_{\beta,\rho^\beta}\left[\nabla_\theta\ln\pi_\theta(a|s)\delta_{t:t+1}(\theta,\phi)\right]\\ &\quad+\mathbb{E}_{\pi,\rho^\pi}\left[\nabla_\theta\ln\pi_\theta(a|s)\big((\tau+\lambda)\ln\pi_\theta(a|s)-\lambda\ln\bar\pi(a|s)\big)\right].\end{aligned}\tag{17}$$

このとき，式 (17) 右辺初項は，ERMDP での最適方策へ向かう勾配を計算している．式 (17) に基づき，方策オフのサンプルからエントロピー正則化付目的関数を最適化し，方策オンのサンプルから目的関数を補正することで，サンプル効率よく元の MDP における最適方策を学習できる．

5. おわりに

本研究では，エントロピー正則化付方策勾配定理を導出し，さらに MDP での目的関数を ERMDP での更新則を利用して最大化することを提案した．今後の課題として，複数ステップ予測（$n>1$）への拡張，提案法の実験的な評価，収束性に関する議論などが挙げられる．

謝辞

本研究は，JST，CREST，JPMJCR17A4 の支援を受けたものである．

参考文献

[Azar 12] Azar, M. G., Gómez, V., and Kappen, H. J.: Dynamic Policy Programming, *Journal of Machine Learning Research*, Vol. 13, (2012)

[Fox 16] Fox, R., Pakman, A., and Tishby, N.: Taming the Noise in Reinforcement Learning via Soft Updates, in *Uncertainty in Artificial Intelligence* (2016)

[Haarnoja 17] Haarnoja, T., Tang, H., Abbeel, P., and Levine, S.: Reinforcement Learning with Deep Energy-Based Policies, in *International Conference on Machine Learning* (2017)

[Kozuno 17] Kozuno, T., Uchibe, E., and Doya, K.: Unifying Value Iteration, Advantage Learning, and Dynamic Policy Programming, in *arXiv* (2017)

[Lin 92] Lin, L.-J.: Self-improving reactive agents based on reinforcement learning, planning and teaching, *Machine Learning*, Vol. 8, No. 3/4, pp. 69–97 (1992)

[Mnih 15] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., et al.: Human-level control through deep reinforcement learning, *Nature*, Vol. 518, No. 7540, pp. 529–533 (2015)

[Mnih 16] Mnih, V., Mirza, A. P. B. M., Graves, A., Harley, T., Lillicrap, T. P., Silver, D., and Kavukcuoglu, K.: Asynchronous methods for deep reinforcement learning, in *International Conference on Machine Learning*, pp. 1928–1937 (2016)

[Nachum 17a] Nachum, O., Norouzi, M., Xu, K., and Schuurmans, D.: Bridging the Gap Between Value and Policy Based Reinforcement Learning, in *Advances in Neural Information Processing Systems* (2017)

[Nachum 17b] Nachum, O., Norouzi, M., Xu, K., and Schuurmans, D.: Trust-PCL: An Off-Policy Trust Region Method for Continuous Control, in *International Conference on Learning Representations* (2017)

[O'Donoghue 16] O'Donoghue, B., Munos, R., Kavukcuoglu, K., and Mnih, V.: Combining Policy Gradient and Q-Learning, in *International Conference on Learning Representations* (2016)

[Peters 10] Peters, J., Mülling, K., and Altün, Y.: Relative Entropy Policy Search, in *AAAI* (2010)

[Schulman 17] Schulman, J., Chen, X., and Abbeel, P.: Equivalence Between Policy Gradients and Soft Q-Learning, in *arXiv* (2017)

[Sugimoto 16] Sugimoto, N., Tangkaratt, V., Wensveen, T., Zhao, T., Sugiyama, M., and Morimoto, J.: Trial and Error: Using Previous Experiences as Simulation Models in Humanoid Motor Learning, *IEEE Robotics & Automation Magazine*, Vol. 23, No. 1, pp. 96–105 (2016)

[Sutton 99] Sutton, R. S., McAllester, D., Singh, S., and Mansour, Y.: Policy Gradient Methods for Reinforcement Learning with Function Approximation, in *Advances in Neural Information Processing Systems*, Vol. 12, pp. 1057–1063 (1999)

[Todorov 06] Todorov, E.: Linearly-solvable Markov decision problems, in *Advances in Neural Information Processing Systems* (2006)

[Todorov 10] Todorov, E.: Policy gradients in linearly-solvable MDPs, in *Advances in Neural Information Processing Systems* (2010)

[Watkins 89] Watkins, C. J. C. H.: *Learning from Delayed Rewards*, PhD thesis, King ' s College, Cambridge, England (1989)

[Watkins 92] Watkins, C. J. C. H. and Dayan, P.: Q-learning, *Machine Learning*, Vol. 8, pp. 279–292 (1992)

[Williams 92] Williams, R. J.: Simple statistical gradient-following algorithms for connectionist reinforcement learning, *Machine Learning*, Vol. 8, pp. 229–256 (1992)

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

ベンチマークデータを用いた 時系列勾配ブースティング木の実験評価

Experimental evaluation of Time-Series Gradient Boosting Tree with Time-Series Benchmark datasets

今村 光良 *1*2 Mitsuyoshi Imamura
中川 慧 *1*3 Kei Nakagawa
吉田 健一 *3 Kenichi Yoshida

*1野村アセットマネジメント株式会社
Nomura Asset Management Co., Ltd

*2筑波大学大学院 システム情報工学研究科
University of Tsukuba Graduate School of Systems and Information Engineering

*3筑波大学大学院 ビジネス科学研究科
University of Tsukuba Graduate School of Business Sciences

In this paper, We evaluated the time-series gradient boosting decision tree method using benchmark data. Our time-series gradient boosting tree has weak learners with time-series and cross-sectional attribute in its internal node, and split examples based on dissimilarity between a pair of time-series or impurity between a pair of cross-sectional attributes. It has been empirically observed that the method induces accurate and comprehensive decision trees in time-series classification, which has gaining increasing attention due to its importance in various real-world applications.

1. はじめに

本論文では、別報 [中川 18] で提案された時系列勾配ブースティング木について、様々な時系列データセットに対して分析を行い、その有効性を評価する。別報 [中川 18] では時系列およびクロスセクションの属性を持つデータセットに対する時系列勾配ブースティング決定木を提案した。勾配ブースティング木などのツリーモデルは一般に可読性に優れており、説明責任という観点から実務への適用がしやすい。勾配ブースティング木は決定木を多段に組み合わせることで確率的な予測値を導出する手法であり、非常に高い精度を実現できる。しかしながら従来の決定木学習手法は時系列属性を想定していないため、時系列データを含むデータ集合に適用する場合、データの前処理が必要となる。最も単純な前処理の方法として、時系列データを計測値の平均値や標準偏差などの時系列データのモーメントで置き換える方法が考えられる。ただし、この方法は時系列データの構造、すなわち形状を無視しており、例えば形が大きく異なる時系列データを同一視してしまう欠点がある。そこで、勾配ブースティング・アルゴリズムに用いる弱学習器として、以上の問題点を克服した時系列データの形を陽に扱い時系列データ全体を対象とする [Yamada 03] の基準例分割テストによる分割を用いた時系列決定木を使用する。時系列決定木は相違度基準として動的時間伸縮法 (DTW) を用いているが、時系列ペアに対し距離が定義されていることから、[Nakagawa 17] のように決定木学習手法ではなく、最近傍法 (k-NN や k^*-NN[Anava 16]) を用いることも考えられる。ただし、最近傍法は怠惰学習 (lazy learning) であるために分類モデルが存在せず、学習結果が分かりにくいという欠点がある。本論文では、85 種類の異なる時系列データセットである UCR データセット [Chen 15] を用いて、DTW を距離尺度とした 1-Nearest Neibor(NN) と先行研究である [Yamada 03] の時系列決定木をベンチマークに、時系列勾配ブースティング木の性能を評価する。

連絡先: 今村 光良，野村アセットマネジメント株式会社，〒103-8260 東京都中央区日本橋一丁目 12 番 1 号

2. 先行研究

勾配ブースティングは、ブースティング・アルゴリズムの一種である。ブースティングとは、集団学習の枠組みの一つで、複数の弱学習器を統合して全体の学習器を構成する手法である。弱学習器としては、決定木が使われることが多い。ブースティングにおける弱学習器として決定木を用いることは、データの外れ値に強い、数値変数と離散変数や欠損値などを扱いやすいなどの利点がある。本研究では弱学習器として、[Yamada 03] の時系列決定木に、時系列以外の通常の特徴量も追加した決定木を使用することを提案する。そこで、以下では時系列決定木および勾配ブースティング木についてレビューする。

2.1 時系列決定木

時系列決定木は、内部ノードに基準となる属性時系列を持つ決定木であり、基準例分割テストまたはクラスター分割テストによってサンプルを分割していく [Yamada 03]。[山田 03] の実験からクラスター分割テストより基準例分割テストの結果が良好であることから、本稿では基準例分割テストを分割手法として用いる。

データセット D は、n 個のサンプル $\{x_i, i = 1, ..., n\}$ を含み、各サンプル x_i はそれぞれ m 個の属性 $\{a_j, j = 1, ..., n\}$ とクラスラベルまたは目的変数 $\{y_i, i = 1, ..., n\}$ を持つ。各属性 a_j は、時間順に値を並べた時系列データ又は系列を持たない通常の連続値、名目属性を持つ。ここで $x(a_j)$ は一つのサンプルであり、基準例と呼ぶ。基準例分割テストは、サンプル x_i の属性 a_j に関する時系列データを $x_i(a_j)$ で表すと、データセット全体を、$DTW(x(a_j), ex_i(a)) < \theta_i$ を満たすデータセットから構成される集合 $L(x, a_j, \theta_i) = \{(y, x)|DTW(x(a_j), x_i(a_j)) < \theta_i\}$ とそれ以外の集合 $R(x, a_j, \theta_i) = \{(y, x)|DTW(x(a_j), x_i(a_j)) \geq \theta_i\}$ に分割する。ここで $DTW(x, y)$ は、動的時間伸縮法 (DTW) に基づいた類似度を表す。DTW は時系列データにおける一点のデータをもう片方の時系列データにおける複数点

のデータに対応づけられるため、時間方向の非線形な伸縮を許容する。このため、長さが異なる時系列データのペアに適応できる上に、結果が人間の直観により合致する。具体的なアルゴリズムは、Algorithm1 の通り。

基準例分割テストはサンプルが基準例 $x(a_j)$ と DTW で計測して、類似しているか、あるいは類似していないかという基準で閾値 θ を変えながら分割する。当該分割の評価基準としては、利得比基準など通常の決定木で用いられている方法が使える。

Algorithm 1 DTW distance

procedure DTW(x, y)
▷ Initialize matrix D
  Var $D[N, M]$
  $D[1, 1] = 0$
  for $i = 2$ to N **do**
    for $j = 2$ to M **do**
      $D[i, j] = \infty$
    end for
  end for
▷ Calculate DTW distance
  for $i = 2$ to N **do**
    for $j = 2$ to M **do**
      $D[i, j] = d(x[i-1], y_[j-1]) + min(D[i, j-1], D[i-1, j], D[i-1, j-1])$
    end for
  end for
  return $D[N, M]$
end procedure

2.2 勾配ブースティング木

[Friedman 01] は、ブースティングの各ステップで構築するモデルに回帰ツリーを用いる勾配ブースティング木（gradient boosting tree）という手法を考案した。これは特徴量 (入力) を x、ラベルを y として、弱学習器 $f_m(x), m = 0, ..., M$ に対して全体の学習器 $F(x)$

$$F(x) = f_0(x) + f_1(x) + ... + f_M(x) \quad (1)$$

を損失関数 $L(y, F(x))$ が最小になるように弱学習器 $f_m(x)$ を逐次的に学習し統合する。つまり、最初に学習器 $F_0(x) = f_0(x)$ が与えられるとし、m ステップ目の学習では m 個の学習器からなる全体の学習器を損失関数 $L(y, F(x))$ が最小になるように弱学習器を決める。

$$F_m(x) = F_{m-1}(x) + f_m(x) \quad (2)$$

具体的なアルゴリズムは以下の通りである。勾配ブースティング木の特徴は、その時点でのモデルの予測値と結果の残差を目的変数とした回帰ツリーを構築し、そのツリーモデルの予測値による値を加算することにある。

3. 提案手法 - 時系列勾配ブースティング木

前章で確認した [Yamada 03] の基準例分割テストによる時系列決定木において、時系列属性以外の属性、通常の連続値、名目属性も追加する。属性が時系列データの場合には基準例分割テストによって分割を行い、時系列以外の場合には、通常の決定木と同様の基準で分割を行う。具体的なアルゴリズム (StandardExSplit with Cross-Sectional Data) は、Algorithm3 の通り。$H(\bullet)$ は不純度を表す関数で、本稿では、よく使用される Gini 係数を用いる。

Algorithm 2 Gradient Boosting Tree

procedure GRADIENT BOOSTING TREE(y, x)
▷ Initialize F_0 with a constant
  $F_0(x) = \sum_{i=1}^{N} \arg\min_c(y_i, c)$
  for $m = 1$ to M **do**
    $r_{im} = -[\frac{L(y_i, F(x_i))}{F(x_i)}]_{F(x)=F_{m-1}(x_{i-1})}$
▷ Fit a decision tree to predict targets r_{im}
    f(x) = FitTree(r_{im},x)
    $\rho_m = \arg\min_\rho \sum_{i=1}^{N} L(y_i, F_{m-1}(x) + \rho f(x))$
    $F_m(x) = F_{m-1}(x) + \gamma\rho_m f(x)$
  end for
  return $F(x) = \sum_{m=1}^{M} F_m(x)$
end procedure

Algorithm 3 StandardExSplit with Cross-Sectional Data

procedure SPLIT($\{y_1, ..., y_n\}, \{x_1, ..., x_n\}$)
  for each samples x_i **do**
    for each attribute a_j **do**
      if attribute a_j is time-series **then**
        for each samples x_k **do**
          $L(x, a_j, \theta_i)$
          $= \{(y, x) | DTW(x_i(a_j), x_k(a_j)) < \theta_k\}$
          $R(x, a_j, \theta_i)$
          $= \{(y, x) | DTW(x_i(a_j), x_k(a_j)) \geq \theta_k\}$
          $G(L, R, \theta_i)$
          $= \frac{|L|}{|L+R|} H(L) + \frac{|R|}{|L+R|} H(R)$
        end for
      else
        $L(x, a_j, \theta_i) = \{(y, x) | x(a_j) < \theta_i\}$
        $R(x, a_j, \theta_i) = \{(y, x) | x(a_j) \geq \theta_i\}$
      end if
      $G(L, R, \theta_i) = \frac{|L|}{|L+R|} H(L) + \frac{|R|}{|L+R|} H(R)$
    end for
  end for
  $\theta^* = \arg\min_\theta G(L, R, \theta_i)$
  return best split θ^*
end procedure

この分割アルゴリズムを用いた時系列決定木を勾配ブースティング木の弱学習器として利用する。ただし、本論文の分析においては、データセットにクロスセクションのデータが存在しないため、クロスセクションのデータは用いない。

4. 実験

提案手法の有効性を確認するために、実データを用いた分析を行う。85 種類の異なる時系列データセットである UCR データセット [Chen 15] を用いて、時系列勾配ブースティング木の性能を評価する。UCR データセットは平均 0、分散 1 に正規化されているため、特別な前処理は行っていない。各データセットはそれぞれ、Image Outline(IO)、Motion Classification(MC)、

Sensor Readings(SR)、Synthetic(Syn) という 4 つのカテゴリーに分類できる。表 1 にベースラインである DTW 距離を用いた 1-NN(1-NN DTW) と先行研究で提案された時系列決定木 (TST)、そして提案手法である時系列勾配ブースティング木 (TSGBT) の各データセット毎の Error Rate をまとめた。31 のデータセットについてベンチマークである 1-NN DTW に対して改善がみられた。

勾配ブースティング木は、怠惰学習 (Lazy Learning) である k-NN とは異なり、学習が必要であるため、学習にはある程度のサンプル数が必要であると考えられる。そこで、別報 [中川 18] の金融時系列データの分析と同程度である 400 以上のサンプルが確保できるデータセット 21 個についての各カテゴリーの平均 Error Rate を示したのが表 1 である。14 のデータセットについて改善がみられた。

5. まとめ

本論文では、別報 [中川 18] で提案された時系列勾配ブースティング木について、85 種類の異なる時系列データセットである UCR データセット [Chen 15] を用いて、時系列勾配ブースティング木の性能を評価した。ベースラインである DTW 距離を用いた 1-NN(1-NN DTW) と先行研究で提案された時系列決定木 (TST)、そして提案手法である時系列勾配ブースティング木 (TSGBT) の各データセット毎の Error Rate を評価した結果、31 のデータセットについてベンチマークである 1-NN DTW に対して改善がみられた。

また、勾配ブースティング木は、怠惰学習 (Lazy Learning) である k-NN とは異なり、学習が必要であるため、学習にはある程度のサンプル数が必要であると考えられる。そこで、400 以上のサンプルが確保できるデータセット 21 個についての各カテゴリーの平均 Error Rate を評価したところ、14 のデータセットについて改善がみられた。

表 1: Error Rate of UCR Time-series Datasets

Name	Category	1-NN DTW	TST	TSGBT
50Words	IO	0.31	0.68	0.40
Adiac	IO	0.40	0.80	0.51
ArrowHead	IO	0.30	0.34	0.39
Beef	SR	0.37	0.63	0.57
BeetleFly	IO	0.30	0.30	0.30
BirdChicken	IO	0.25	0.25	0.30
Car	SR	0.27	0.43	0.40
CBF	Syn	0.00	0.11	0.01
ChlorineConcentration	SR	0.35	0.45	0.41
CinCECGtorso	SR	0.35	0.54	0.44
Coffee	SR	0.00	0.11	0.11
Computers	SR	0.30	0.38	0.32
CricketX	MC	0.25	0.62	0.36
CricketY	MC	0.26	0.61	0.39
CricketZ	MC	0.25	0.66	0.36
DiatomSizeReduction	IO	0.03	0.17	0.14
DistalPhalanxOutlineAgeGroup	IO	0.21	0.24	0.21
DistalPhalanxOutlineCorrect	IO	0.23	0.25	0.23
DistalPhalanxTW	IO	0.29	0.36	0.23
Earthquakes	SR	0.26	0.29	0.23
ECG	SR	0.23	0.21	0.21
ECG5000	SR	0.08	0.08	0.07
ECGFiveDays	SR	0.23	0.32	0.32
ElectricDevices	SR	0.40	0.36	0.25
Face (all)	IO	0.19	0.66	0.28
Face (four)	IO	0.17	0.11	0.35
FacesUCR	IO	0.10	0.54	0.18
Fish	IO	0.18	0.43	0.28
FordA	SR	0.44	0.36	0.26
FordB	SR	0.41	0.36	0.24
Gun-Point	MC	0.09	0.27	0.21
Ham	SR	0.53	0.49	0.33
HandOutlines	IO	0.20	0.19	0.17
Haptics	MC	0.62	0.63	0.60
Herring	IO	0.47	0.41	0.39
InlineSkate	MC	0.62	0.74	0.70
InsectWingbeatSound	SR	0.65	0.69	0.58
ItalyPowerDemand	SR	0.05	0.09	0.09
LargeKitchenAppliances	SR	0.21	0.27	0.19
Lightning-2	SR	0.13	0.26	0.25
Lightning-7	SR	0.27	0.51	0.42
MALLAT	Syn	0.07	0.36	0.16
Meat	SR	0.07	0.10	0.12
MedicalImages	IO	0.26	0.45	0.28
MiddlePhalanxOutlineAgeGroup	IO	0.25	0.26	0.25
MiddlePhalanxOutlineCorrect	IO	0.35	0.37	0.31
MiddlePhalanxTW	IO	0.42	0.40	0.40
MoteStrain	SR	0.17	0.18	0.19
Non-Invasive Fetal ECG Thorax1	SR	0.21	0.71	0.24
Non-Invasive Fetal ECG Thorax2	SR	0.14	0.88	0.18
OliveOil	SR	0.17	0.17	0.20
OSU Leaf	IO	0.41	0.52	0.43
PhalangesOutlinesCorrect	IO	0.27	0.31	0.20
Phoneme (readme)	SR	0.77	0.80	0.77
Plane	IO	0.00	0.21	0.00
ProximalPhalanxOutlineAgeGroup	IO	0.20	0.16	0.18
ProximalPhalanxOutlineCorrect	IO	0.22	0.20	0.17
ProximalPhalanxTW	IO	0.26	0.28	0.23
RefrigerationDevices	SR	0.54	0.53	0.50
ScreenType	SR	0.60	0.57	0.62
ShapeletSim	Syn	0.35	0.45	0.46
ShapesAll	IO	0.23	0.90	0.29
SmallKitchenAppliances	SR	0.36	0.27	0.22
SonyAIBORobot Surface	SR	0.28	0.29	0.13
SonyAIBORobot SurfaceII	SR	0.17	0.21	0.21
StarLightCurves	SR	0.09	0.07	0.05
Strawberry	SR	0.06	0.13	0.09
Swedish Leaf	IO	0.21	0.48	0.18
Symbols	IO	0.05	0.24	0.10
Synthetic Control	Syn	0.01	0.05	0.03
ToeSegmentation1	MC	0.23	0.28	0.26
ToeSegmentation2	MC	0.16	0.27	0.27
Trace	SR	0.00	0.09	0.06
Two Patterns	Syn	0.00	0.03	0.00
TwoLeadECG	SR	0.10	0.20	0.20
uWaveGestureLibraryX	MC	0.27	0.31	0.23
uWaveGestureLibraryY	MC	0.37	0.39	0.31
uWaveGestureLibraryZ	MC	0.34	0.35	0.29
UWaveGestureLibraryAll	MC	0.11	0.39	0.10
Wafer	SR	0.02	0.03	0.02
Wine	SR	0.43	0.48	0.35
WordSynonyms	IO	0.35	0.63	0.47
Worms	IO	0.54	0.48	0.51
WormsTwoClass	IO	0.34	0.34	0.34
Yoga	IO	0.16	0.24	0.19

参考文献

[Anava 16] Anava, O. and Levy, K.: k*-nearest neighbors: From global to local, in *Advances in Neural Information Processing Systems*, pp. 4916–4924 (2016)

[Chen 15] Chen, Y., Keogh, E., Hu, B., Begum, N., Bagnall, A., Mueen, A., and Batista, G.: The UCR Time Series Classification Archive (2015), `www.cs.ucr.edu/~eamonn/time_series_data/`

[Friedman 01] Friedman, J. H.: Greedy function approximation: a gradient boosting machine, *Annals of statistics*, pp. 1189–1232 (2001)

[Nakagawa 17] Nakagawa, K., Imamura, M., and Yoshida, K.: Stock Price Prediction using k*-Nearest Neighbors and Indexing Dynamic Time Warping, in *Artificial Intelligence of and for Business (AI-Biz 2017)* (2017)

[Yamada 03] Yamada, Y., Suzuki, E., Yokoi, H., and Takabayashi, K.: Decision-tree induction from time-series data based on a standard-example split test, in *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, pp. 840–847 (2003)

表 2: Average Error Rate of UCR Time-series Datasets by Category

Category	1-NN DTW	TST	TSGBT
Image Outline	0.23	0.49	0.24
Motion Classification	0.27	0.36	0.20
Sensor Readings	0.24	0.37	0.19
Synthetic	0.00	0.03	0.00

[山田 03] 山田悠, 鈴木英之進, 横井英人, 高林克日己 他：動的時間伸縮法に基づく時系列データからの決定木学習, 情報処理学会研究報告知能と複雑系 (ICS), Vol. 2003, No. 30 (2002-ICS-132), pp. 141–146 (2003)

[中川 18] 中川慧, 今村光良, 吉田健一：時系列勾配ブースティング木による分類学習金融時系列予測への応用, 人工知能学会全国大会論文集 2018 年度人工知能学会全国大会 (第 32 回) 論文集社団法人 人工知能学会 (2018)

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

Batch Random Walk for GPU-Based Classical Planning (Extended Abstract)

Ryo Kuroiwa Alex Fukunaga

Graduate School of Arts and Sciences, The University of Tokyo

Graphical processing units (GPUs) have become ubiquitous because they offer the ability to perform cost and energy efficient massively parallel computation. We investigate forward search classical planning on GPUs based on Monte-Carlo Random Walk(MRW). We first propose Batch MRW (BMRW), a generalization of MRW which performs random walks starting with many seed states, in contrast to traditional MRW which used a single seed state. We evaluate a sequential implementation of BMRW on a single CPU core and show that a sequential, satisficing planner based BMRW performs comparably with Arvand, the previous MRW-based planner. Then, we propose BMRW$_G$, which uses a GPU to perform random walks. We show that BMRWG achieves significant speedup compared to BMRW and achieves competitive performance on a number of IPC benchmark domains. This is an extended abstract of an ICAPS2018 paper [Kuroiwa 18].

1. Introduction

The use of Graphics Processing Units (GPUs) for general-purpose computing has become ubiquitous in many areas including AI, but their use in domain-independent planning has been quite limited. This seems largely due to the fact that there is a significant mismatch between the architecture of GPUs and forward heuristic search based algorithms commonly used for planning. For satisficing, classical planning, the most widely studied forward search strategy in recent years have been approaches such as Greedy Best First Search (GBFS), as well as many improvements which seek to avoid/escape local minima and plateaus.

In standard GBFS, Enhanced Hill-Climbing, and weighted A* approaches, each node expansion involves accessing global open/closed sets, which poses a challenge for efficient parallelization. Methods for efficiently distributing work in parallel best-first search (BFS) based planners on multi-core machines as well as clusters have been studied [Burns 10, Kishimoto 13], but these previous approaches for parallel BFS can not be straightforwardly applied to GPUs. One major issue is that GPUs provide thousands of cores/threads, but the amount of GPU RAM available per thread is quite limited. For example, a state-of-the-art Nvidia GTX1080 has 8GB global RAM, which must be shared by 2560 physical CUDA cores. This will be exhausted within a few seconds by a parallel BFS algorithm if the open/closed lists are stored on the GPU, as in GA*, a delayed duplicate detection based A* for the GPU [Zhou 15]. There is also a tiny amount of fast RAM per core ($\sim$375 bytes/core on a GTX1080), and although previous work has investigated performing domain-specific IDA* search using only this local memory [Horie 17], this is too small to hold even a single state for most domain-independent planning domains. Sulewski et al. ([Sulewski 11]) used a GPU to parallelize the successor generation step for breadth-first search in cost-optimal planning, with duplicate detection and open/closed list management performed on the CPU. A forward *heuristic* search method for satisficing planning which effectively uses the GPU has remained an open problem.

One approach to heuristic-driven forward search which encourages explorative search behavior and is suited for GPU parallelization is Monte Carlo Random Walk Planning (MRW) [Nakhost 09]. At each step, MRW starts at some state s (initially the initial state), performs a set of random walks from s, and then sets s to the best endpoint (according to a heuristic function) found by the random walks. State-of-the-art MRW-based planners have been shown to be competitive with GBFS-based approaches on some domains [Nakhost 13]. MRW appears to be suited for GPU parallelization because each random walk can be executed independently by a GPU thread.

In this paper, we first propose BMRW, a simple generalization of MRW-based search which combines an open list based search strategy with MRW. In BMRW, the search is driven by an open list, as in GBFS. In each iteration, a batch of nodes is selected from the open list, and random walks are performed starting with these nodes. Promising states found by the random walks are then inserted into the open list, and this cycle repeats until a goal is found or time runs out. We experimentally show that BMRW is a promising search strategy, and show that a planner using BMRW search on a single CPU core is competitive with Arvand13, the previous, state-of-the-art MRW based sequential planner.

Next, we propose BMRW$_G$, an efficient, parallel implementation of BMRW for GPUs. BMRW$_G$ maintains the open list in CPU memory, but uses the GPU for the random walks. We show that BMRW$_G$ achieves significant speedup compared to BMRW, and that BMRW$_G$ achieves competitive performance on a number of standard IPC benchmark domains.

2. Monte-Carlo Random Walk Planning

Monte-Carlo Random Walk Planning (MRW) was first proposed in the Arvand planner [Nakhost 09]. The state-of-the-art Arvand13 MRW-based planner, which performs significantly better than the original Arvand planner, works as follows: Given a state s, a basic *random walk* repeatedly generates successors of s, chooses one of the successors s' of s and transitions to s' ($s \rightarrow s'$). From the current node s (initially set to the start state), the Arvand13 MRW algorithm perform a set of random walks. Each state on the walk is evaluated according to a heuristic evaluation function h (the FF heuristic [Hoffmann 01] was used), and the walk returns

either when (a) it encounters a state with a better h-value than the random walk start state, in which case that state becomes the start point for the next random walk, or (b) with some probability (i.e., local restart). A global restart is triggered when h does not improve after some number of random walks. Arvand13 used an enhanced random walk which, instead of uniformly randomly choosing a successor, biased successor choice according to *helpful actions* identified by the FF heuristic. In addition, local and global start thresholds are set adaptively.

3. Batch Monte-Carlo Walk (BMRW)

We now propose BMRW, a generalization of MRW. BMRW, shown in Alg. 1, maintains a priority queue $openList$, initially containing the successors of the initial state s_0. In each iteration of the main loop (lines 21–36), BMRW first checks if $openList$ is empty, and if so, initializes it with the successors of the start state s_0, i.e., it performs a *global restart*. Then, a *batch* of $batchSize$ nodes is selected from $openList$. A random walk (Alg. 1, Walk function) of up to l steps is peformed from each start point in $batch$, and the results are stored in $walkres$.

The main differences between BMRW and MRW are: (1) MRW performs random walks from the same start state, while BMRW performs a set of random walks based on a batch of start states selected from $openList$. (2) MRW only keeps and updates a single "current state" (which is used as the start point for random walks), while BMRW maintains an $openList$, similar to GBFS-based approaches. Thus, BMRW can be viewed as a hybrid between MRW and GBFS. (3) In MRW, every random walk is followed by a possible update of the current state (jump to the state returned by the walk), whereas BMRW performs an entire batch of walks at a time.

In addition to the above basic scheme, BMRW uses an *elite insertion* policy (lines 32-36), where for the best n results (according to h-value) of the random walk return, the successors of those nodes are inserted into $openList$ instead of the nodes themselves. This is intended to strongly encourage further exploration of these "elite" nodes by pushing its many successors into $openList$ (because these successors also have good h-values, they are likely to be expanded soon).

Furthermore, a $closedList$ is used in order to prevent duplicate states from being pushed into the open list (lines 30-31). This ensures that each random walk starts from a different start state, promoting exploration of the search space.

MRW can be viewed as a special case of BMRW with batch size $b = 1$ and a special $openList$ limited to size 1.

3.1 BMRW$_G$: BMRW on a GPU

In principle, BMRW can be efficiently implemented on a GPU, due to the independence of each random walk. In Alg. 1 lines 25-26, the for loop is executed in parallel on the GPU. After each node in the batch selected in line 24 is copied to the GPU, each node in the batch is assigned to a GPU thread, and each random walk is performed by a single GPU thread, after which the result of the walk is copied back to the CPU. Everything else is performed on the CPU.

The main bottleneck in implementing BMRW on a GPU is the relatively small amount of GPU RAM available per thread. Although GPUs have a substantial amount of global memory available on the GPU, (the GTX1080 we used has 8GB), this global RAM must be shared among all threads. The GTX1080 has 2560 CUDA cores (20 Streaming Multiprocessors, 128 cores/SM), roughly 3125KB/thread. Although a random walk does not require much memory for state information, as there is no open/closed list during the walk, it is necessary to maintain data structures for fast, incremental computation of heuristic values during the walk. For example, to compute the FF heuristic using the generalized Dijkstra method [Liu 02, Keyder 08], the cost and unsatisfied preconditions of each ground action is required. On the GTX1080, if we use 2560 threads, each thread can handle approximately 390,000 ground actions – this is sufficient for most IPC benchmark domains. However, this implies that we currently can not run BMRW with tens of thousands of threads to take even better advantage of the ability of GPUs to overlap communication and memory access. Note that larger domains can be handled by reducing the number of threads, at the cost of leaving some cores idle.

We used the Landmark Count (LMC) heuristic [Hoffmann 04] for BMRW$_G$. The landmark graph is built on the CPU once at the beginning of search, transferred to the GPU global RAM, and shared by all GPU threads. Computing the LMC value only requires incrementally updating the number of unreached landmarks for the state. This requires $O(\#facts)$ memory. As shown below, BMRW with the LMC heuristic significantly outperforms BMRW with the FF heuristic.

Algorithm 1 Batch MRW

```
1: function WALK(s, goals, l)
2:     s_best ← s
3:     for i ← 1 to l do
4:         if s ∈ goals then return s
5:         if DeadEnd(s) then s ← s_best
6:         else s ← RANDOMSELECT(successors(s))
7:         if h(s) < h(s_best) then s_best = s
8:     return s_best
9:
10: function FILLBATCH(openList, batchSize, batch)
11:     offset ← 0
12:     for i ← 1 to batchSize do
13:         if openList is Empty then
14:             if offset = 0 then State offset = i
15:             batch[i] ← batch[i − offset]
16:         else
17:             batch[i] ← POP(openList)
18:
19: function BATCHMRW(s_0, goals, l, batchSize, n)
20:     openList, closedList, batch, walkres ← φ
21:     loop
22:         if openList is Empty then
23:             openList ← successors(s_0)
24:         FILLBATCH(openList, batchSize, batch)
25:         for i ← 1 to batchSize do
26:             walkres[i] ← WALK(batch[i], goals, l)
27:         i ← 1
28:         for all s ∈ walkres do
29:             if s ∈ goals then return s
30:             if s ∈ closedList then continue
31:             INSERT(closedList, s)
32:             if i ≤ n then
33:                 PUSH(openList, successors(s), h(s))
34:                 i ← i + 1
35:             else
36:                 PUSH(openList, s, h(s))
```

4. Experimental Results

In our experimental evaluation (please see the full paper [Kuroiwa 18]), we evaluated a sequential implementation of BMRW on a single CPU core and showed that a sequential, satisficing planner based BMRW performs comparably with Arvand, the previous MRW-based planner. Then, we showed that BMRW$_G$ achieves significant speedup compared to BMRW and achieves

competitive performance on a number of IPC benchmark domains.

5. Conclusion

In order to exploit GPUs in domain-independent planning, we first proposed BMRW, a generalization of MRW which combines GBFS and random walk by performing a GBFS-like, $openList$-driven search, which at each iteration performs batches of random walks in order to explore the search space. *We showed that BMRW is competitive with previous random walk strategies, including Arvand13. We then showed that BMRW$_G$, a heterogeneous CPU/GPU implementation of BMRW, achieves significant speedup compared to BMRW and a straightforward parallelizatoin of MRW13.*

We have shown that random walk using the relatively lightweight Landmark Count heuristic can be efficiently implemented *entirely on the GPU*. In fact, MRW13, the baseline parallelization of Arvand13, runs almost entirely on the GPU, except for the top level loop, and in BMRW$_G$, only the global $openList$ and $closedList$ management are performed by the CPU. Our primary objective was to demonstrate the feasibility of GPU-based forward heuristic search, so we focused on efficient implementation on the GPU side, and the current implementation only uses a single CPU core. Effective, simultaneous usage of multiple CPU cores along with the GPU in a more heterogeneous algorithm is an avenue for future work. For example, $openList$ and $closedList$ management can be parallelized, as in [Sulewski 11].

While we focused on a relatively simple search strategy which is basically GBFS with (batched) random walk lookahead, random walk has been embedded as an exploration mechanism in other forward heuristic search variants such as RW-LS (ArvandLS) [Xie 12] and GBFS-LE [Xie 14]. It should be possible to combine these more complex algorithms with the basic idea of applying batches of random walks on the GPU with a diverse set of start points and a lightweight heuristic.

References

[Burns 10] Burns, E., Lemons, S., Ruml, W., and Zhou, R.: Best-first heuristic search for multicore machines, *Journal of Artificial Intelligence Research*, Vol. 39, pp. 689–743 (2010)

[Hoffmann 01] Hoffmann, J. and Nebel, B.: The FF Planning System: Fast Plan Generation through Heuristic Search, *J. Artif. Intell. Res.(JAIR)*, Vol. 14, pp. 253–302 (2001)

[Hoffmann 04] Hoffmann, J., Porteous, J., and Sebastia, L.: Ordered Landmarks in Planning, *J. Artif. Intell. Res.*, Vol. 22, pp. 215–278 (2004)

[Horie 17] Horie, S. and Fukunaga, A.: Block-Parallel IDA* for GPUs, in *Proceedings of the Tenth International Symposium on Combinatorial Search*, pp. 134–138 (2017)

[Keyder 08] Keyder, E. and Geffner, H.: Heuristics for Planning with Action Costs Revisited, in *ECAI 2008 - 18th European Conference on Artificial Intelligence, Patras, Greece, July 21-25, 2008, Proceedings*, pp. 588–592 (2008)

[Kishimoto 13] Kishimoto, A., Fukunaga, A., and Botea, A.: Evaluation of a simple, scalable, parallel best-first search strategy, *Artificial Intelligence*, Vol. 195, pp. 222–248 (2013)

[Kuroiwa 18] Kuroiwa, R. and Fukunaga, A.: Batch Random Walk for GPU-Based Classical Planning, in *Proceedings of International Conference on Automated Planning and Scheduling (ICAPS)* (2018)

[Liu 02] Liu, Y., Koenig, S., and Furcy, D.: Speeding Up the Calculation of Heuristics for Heuristic Search-based Planning, in *Eighteenth National Conference on Artificial Intelligence*, pp. 484–491, Menlo Park, CA, USA (2002), American Association for Artificial Intelligence

[Nakhost 09] Nakhost, H. and Müller, M.: Monte-Carlo Exploration for Deterministic Planning, in *Proceedings of the 21st International Jont Conference on Artifical Intelligence*, IJCAI'09, pp. 1766–1771, San Francisco, CA, USA (2009), Morgan Kaufmann Publishers Inc.

[Nakhost 13] Nakhost, H. and Müller, M.: Towards a Second Generation Random Walk Planner: An Experimental Exploration, in *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, IJCAI '13, pp. 2336–2342, AAAI Press (2013)

[Sulewski 11] Sulewski, D., Edelkamp, S., and Kissmann, P.: Exploiting the Computational Power of the Graphics Card: Optimal State Space Planning on the GPU., in *Proceedings of the International Conference of Automated Planning and Scheduling(ICAPS)* (2011)

[Xie 12] Xie, F., Nakhost, H., and Müller, M.: Planning via random walk-driven local search, in *Proceedings of the International Conference of Automated Planning and Scheduling(ICAPS)* (2012)

[Xie 14] Xie, F., Müller, M., and Holte, R.: Adding Local Exploration to Greedy Best-First Search in Satisficing Planning, in *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27 -31, 2014, Québec City, Québec, Canada.*, pp. 2388–2394 (2014)

[Zhou 15] Zhou, Y. and Zeng, J.: Massively Parallel A* Search on a GPU, in *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA.*, pp. 1248–1255 (2015)

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

時系列生成モデルのためのリサンプリング学習

Learning to resample for time-series generative model

金子 貴輝 *1　大澤 昇平 *2　松尾 豊 *3
Takaaki Kaneko　Shohei Ohsawa　Yutaka Matsuo

*1*2*3東京大学大学院 工学系研究科 技術経営戦略学専攻
Department of Technology Management for Innovation, The University of Tokyo

Sequential Monte Carlo (SMC) is a typical sampling method that can be sampled in order from a sequential probabilistic model. However, due to degeneration of the sample, SMC may produce samples with low likelihood with a small number of particles. In our study, we focus on the fact that the same resampling targets of SMC for each sample cause degenerating samples. We want to relax this constraint, but analytically deriving asymmetric sequential resampling targets is difficult. Therefore, we expand resampling strategy of SMC asymmetrically by learning the sequential resampling target from the target of the whole series approximated to the lower bound. By this, by learning to resample, it is expected that accurate estimation of latent variable can be realized with the same particle number as SMC.

1. はじめに

深層学習において，観測変数 x の背後の依存関係を表現するために，潜在変数 z で同時分布 $p(x,z)$ をモデル化すると，推論分布 $p(z|x)$ の扱い方がモデルの学習のために重要になる．例えば変分オートエンコーダ（VAE）[Kingma 13] では，モデルの学習基準である周辺尤度を表現するために，この分布を必要とする．そのためエンコーダで定義される別の分布に置き換えて基準を近似し，KL divergence 最小化を用いてその分布を推論分布に近づける．

特にマルコフ性を持つ時系列の潜在変数における推論分布 $p(z_{1:T}|x_{1:T})$ は，その扱い方が時間の長さ T にスケールするかが問われる基本的な例であり，汎用なアルゴリズムを目指す上で重要である．潜在変数の状態が離散的で計算機で扱えないほど多くなければ，隠れマルコフモデルのような解析的な扱いが可能な場合もある．しかし動画や音声情報を潜在変数に含める場合，密な表現として出現する状態の数は特徴の数に対して指数的に多くなりうる．この場合，解析的に扱うことも直接全体をエンコーダで置き換えることも，時間の長さにスケールしない．このような推論分布の深層学習による近似は研究途上の課題である．

時系列の推論分布に関する問題として，各時刻の潜在変数を推定する逐次推定の問題（2 章）があり，従来のアプローチとして代表的なものに Sampling Importance Resampling（SIR）の逐次適用を行う，逐次モンテカルロ（SMC）[Gordon 93] がある．SIR とは，粒子と呼ばれる確率変数の値の候補を複数持ち，途中までの尤度を使って粒子を選び直す手法である．SMC は将来的に得られるサンプルの尤度が予測できない問題を SIR で解決する手法であり，粒子フィルタとも呼ばれる．SMC によって時系列の推論分布の近似的なサンプルが得られるので，期待値をモンテカルロ近似することができる．3 章で述べるように，SMC の提案分布にニューラルネットを用いて推論分布を近似する研究も出てきている．

しかし SMC では途中までの分布での尤度が低い粒子は捨てられやすく，後から尤度が大きくなる粒子を本来より低い確率でサンプリングしてしまう．これは縮退問題として知られている．特に明らかな問題だと思われるのは，同程度の尤度の粒子を選び直すことで一部無くしてしまう傾向である．これは図 1（左）で示されるように，2 粒子 2 状態の単純な例（5.2 章参照）で顕著である．これは潜在変数が連続値でも起こりうる．

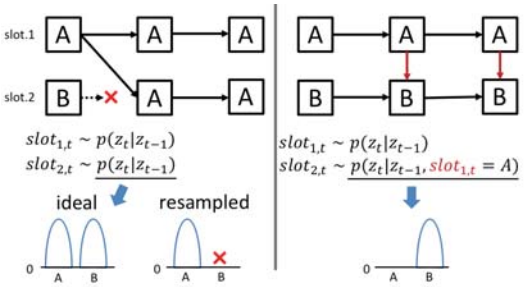


図 1: SMC と提案手法のリサンプリング方式の比較．観測変数は省略している．黒矢印はリサンプリングの流れを，アルファベットは潜在変数の値を表す．(左）SMC でのリサンプリング．独立同分布を用いてサンプルしている．(右）提案手法でのリサンプリング．リサンプリング結果を順に依存させて次のリサンプリングを行うことができる．

本研究では縮退問題のうち，この同程度の粒子をなくす問題に注目する．原因の一つは，リサンプリング目標がどの粒子でも同じであること，つまりリサンプリング目標の対称性である．多様なサンプルを保持できるように，本研究では図 1（右）のように，複数の粒子のリサンプリング目標を非対称にする．

最適な非対称リサンプリング目標は，未来の変数についての期待値の式になるため，一般には解析的に計算できない．学習によって逐次目標を獲得する方針は，計算時間などの有用性が乏しいため注目されてこなかった．しかし潜在変数系列の推論分布に対する新たなアプローチとして重要なアルゴリズムであると考えたため，本研究では逐次目標を学習によって獲得する新しい逐次リサンプリング手法を検討する．

提案手法では，非対称なリサンプリング目標を実現するため，リサンプリング重みに学習パラメータを追加して拡張する．これは従来の SIR のリサンプリング重みに，過去のサンプルを引数に取る新たな重みを掛け合わせることで実現できる．

さらに拡張した重みが縮退を軽減させる適切な値になるよ

連絡先: kaneko@weblab.t.u-tokyo.ac.jp

うに，最終目標 $p(z_{1:T}|x_{1:T})$ との KL 誤差の上界を学習基準として，勾配法を用いて重みを学習する．本研究では，この学習とそれによって獲得された非対称なリサンプリング目標によるリサンプリングのアルゴリズム（4 章）を提案手法とする．学習基準は上界であるため，必ず良い方向に学習が進むとは限らないのは問題であるが，5.1 章で示すように，過去の情報の解釈が未来の尤度に大きく寄与するのに，長期に渡って複数の解釈が尤もらしい状況を簡略化した問題においては，SMC よりも提案手法が正確なサンプリングを目指す学習基準になることを保証できる．

時系列確率モデルでの縮退の軽減を検証するため，本研究では前述の問題設定で実験（5 章）を行い，学習基準と正答率が連動し，SMC より提案手法が正確なリサンプリングが行えることを示した．

2. 問題設定

時系列マルコフモデルの同時分布を以下の式で定義する．

$$p(z_{1:T}, x_{1:T}) = \prod_{t=1}^{T} p(z_t, x_t|z_{t-1}),$$
$$p(z_1, x_1|z_0) = p(z_1, x_1)$$

このモデルからサンプルされた系列 $(x_{1:T}, z_{1:T})$ の $x_{1:T}$ から $z_{1:T}$ を推定し，その誤差の期待値を最小化する問題を考える．

$$\mathbb{E}_{p(z_{1:T}, x_{1:T})}\left[L(z_{1:T}, \hat{z}_{1:T})\right]$$

ただし L は適当な損失関数，$\hat{z}_{1:T}$ は $x_{1:T}$ を用いた推定値とする．

3. 関連研究

2 章の問題には，SMC を基本とした様々なアルゴリズムが提案されてきたが，近年，時系列生成モデルの学習の文脈から，深層学習と SMC を組み合わせる新しい手法が提案されている．SMC とこの新しい手法の二つについて説明する．

3.1 SMC

SMC は時刻 $t = 1$ から $t = T$ まで以下の式で表されるサンプリングを繰り返すアルゴリズムである．

$$\begin{aligned}
&propose && z_t^j \sim r(z_t|z_{1:t-1}^{a_{t-1}^j}),\\
&append && z_{1:t}^j = \{z_{1:t-1}^{a_{t-1}^j}, z_t^j\},\\
&reweight && w_t^j = \frac{p(z_t^j, x_t|z_{t-1}^{a_{t-1}^j})}{r(z_t^j|z_{1:t-1}^{a_{t-1}^j})},\\
&resample && a_t^i \sim \mathrm{Categorical}(w_t^i / \sum_l w_t^l)
\end{aligned}$$

$z_{1:t-1}^{a_{t-1}^j}$ は $t-1$ での j 番目の粒子であり，提案分布 r によって t での潜在変数の値 z_t^j が加わり $z_{1:t}^j$ となる．これを SIR のリサンプリング重み w_t^j にしたがってリサンプリングを行い t での i 番目の粒子 $z_{1:t}^{a_t^i}$ を得る．

SMC では，粒子間でリサンプリング重みを比較し a_t^i をリサンプリングすることで，尤度 $p(z_{1:t}, x_{1:t})$ を正規化して推論分布 $p(z_{1:t}|x_{1:t})$ を計算することなく，近似的なサンプリングを可能としている．しかし全体での目標分布は $p(z_{1:T}|x_{1:T})$ であるため，このリサンプリングが最適であるとは限らない．

3.2 Variational SMC

Variational SMC（VSMC）[Naesseth 17] は VRNN[Chung 15] のような時系列生成モデルを拡張して，SMC と組み合わせたエンコーダとそれに対応した学習基準からなる．SMC における提案分布をニューラルネットで定義できるので，長期的な依存関係を学習することができ，エンコーダで推論分布をさらにうまく近似できるようになった．

本研究の式導出はこれらの研究を参考にしているが，学習させるパラメータは異なる．これらの研究ではエンコーダのパラメータを学習することが目標であり，リサンプリング重みにパラメータは置いていない．

4. 非対称逐次リサンプリング

SMC を非対称なリサンプリングに拡張した提案手法について説明する．まず重みの定義とリサンプリングアルゴリズムを定義したのち，パラメータを学習するための同時分布と学習目標の定義，学習のための勾配法の利用について説明する．

4.1 非対称逐次リサンプリングの定義

時刻 t での，SMC でのリサンプリングを以下の拡張したリサンプリング重みの式で置き換える．

$$\begin{aligned}
&reweight && w_t^{i,j} = \frac{p(z_t^j, x_t|z_{t-1}^{a_{t-1}^j})}{r(z_t^j|z_{1:t-1}^{a_{t-1}^j})} \cdot w_{t,i}^{\Delta}(z_t^j; z_{1:t-1}^{a_{t-1}^j}, z_t^{a_t^{<i}}),\\
&resample && a_t^i \sim w_t^{i,\cdot} / \sum_l w_t^{i,l}
\end{aligned}$$

拡張重み $w_{t,i}^{\Delta}$ は多段のニューラルネットを用いてモデル化する．$z_t^{a_t^{<i}}$ は i 番目より前のリサンプリングされた粒子の組 $\{z_t^{a_t^k}|0 \le k < i\}$ を表す．リサンプリングの計算順序は図 2 のグラフで示すようになる．

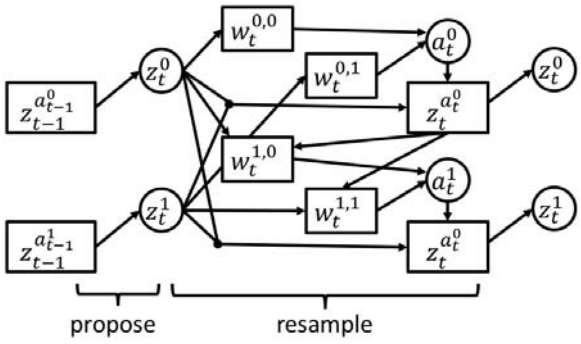


図 2: 提案手法の各時間ステップの計算グラフ．円形のノードが確率変数を，正方形のノードが決定論的な変数を表す．SMC と同様に提案分布（propose）からのサンプリングの後，拡張したリサンプリング重み $w^{i,*}$ と選ばれる粒子 z^{a^i} を粒子のスロット i ごとに繰り返し計算してリサンプリングを行う（resample）．

4.2 同時分布

棄却される粒子も含めた全ての同時分布 ϕ は以下の式で表される．

$$\phi(z, b, a) = \left[\prod_i r(z_1^i)\right] \prod_{t=2}^{T} \prod_i \left[\frac{w_{t-1}^{i,a_{t-1}^i}}{\sum_l w_{t-1}^{a_{t-1}^i}} r(z_t^i|z_{t-1}^{a_{t-1}^i})\right] \left[\frac{w_T^{1,b_T}}{\sum_l w_T^{1,l}}\right]$$

ここで $z = \{z_t^i | t = 1 \dots T, i = 1 \dots N\}, a = \{a_t^i | t = 1 \dots T, i = 1 \dots N\}, b = b_T$ とした．これを周辺化することで提案手法のサンプリング分布が表される．

$$q(z_{1:T}) = q(z_{1:T}^{b_{1:T}}) = \sum_{b_{1:T}} \sum_{a_{1:T}^{\neg b_{1:T}}} \sum_{z_{1:T}^{\neg b_{1:T}}} \phi(z, b, a)$$

ただし $b_t = a_t^{b_{t+1}}, z_{1:T}^{\neg b_{1:T}} = \{z_t^{\neg b_{t+1}} | 1 \leq t \leq T\}, z_t^{\neg b_{t+1}} = \{z_t^i | i \neq b_{t+1}\}, a_{1:T}^{\neg b_{1:T}} = \{a_t^{\neg b_{t+1}} | 1 \leq t \leq T\}, a_t^{\neg b_{t+1}} = \{a_t^i | i \neq b_{t+1}\}$ とする．

4.3 学習基準の導出

目標分布とサンプリング分布で定義される次の KL divergence を最小化することを考える．

$$\mathbb{E}_{p(x_{1:T})}\left[D_{\mathrm{KL}}(q \| p)\right] = \mathbb{E}_{p(x_{1:T})q(z_{1:T}|x_{1:T})}\left[\log \frac{q(z_{1:T}|x_{1:T})}{p(z_{1:T}|x_{1:T})}\right]$$

サンプリング分布に含まれる周辺化計算が解析的に行えないため，以下でイェンセンの不等式を用いた学習基準の近似を行う．

KL divergence の最小化は VAE のエビデンス下界 [Kingma 13] の最大化と一致する．

$$\begin{aligned}
&- \mathbb{E}_{p(x_{1:T})}\left[D_{\mathrm{KL}}(q \| p)\right] \\
&= \mathbb{E}_{p(x_{1:T})q(z_{1:T}|x_{1:T})}\left[\log \frac{p(z_{1:T}, x_{1:T})}{q(z_{1:T}|x_{1:T})}\right] - \mathbb{E}_{p(x_{1:T})}\left[\log p(x_{1:T})\right]
\end{aligned}$$

この式のサンプリング分布の対数尤度を計算し，真の対数尤度と残りに分解する．

$$\begin{aligned}
\phi(z, b, a) =& \prod_{t=1}^{T}\left[\frac{w_t^{b_{t+1}, b_t}}{\sum_l w_t^{b_{t+1}, l}}\right] \prod_{i \neq b_1} r(z_1^i) \\
& \cdot \prod_{t=2}^{T} \prod_{i \neq b_t} \left[\frac{w_{t-1}^{i, a_{t-1}^i}}{\sum_l w_{t-1}^{i,l}}\right] \prod_{t=1}^{T}\left[r(z_t^{b_t} | z_{t-1}^{b_{t-1}})\right] \\
=& \prod_{t=1}^{T}\left[\frac{w_t^{b_{t+1}, b_t}}{\sum_l w_t^{b_{t+1}, l}} r(z_t^{b_t} | z_{t-1}^{b_{t-1}})\right] \cdot \phi(x^{\neg b}, a^{\neg b} | x^b, b)
\end{aligned}$$

ただし，以下のように置いた．

$$\phi^{\neg b} = \phi(z^{\neg b}, a^{\neg b} | z^b, b) = \prod_{i \neq b_1} r(z_1^i) \prod_{t=2}^{T} \prod_{i \neq b_t} \left[\frac{w_{t-1}^{i, a_{t-1}^i}}{\sum_l w_{t-1}^{i,l}}\right]$$

SIR のリサンプリング重みから目標分布の尤度を取り出して，次のように式変形する．

$$\begin{aligned}
&\phi(z, b, a) \\
&= \prod_t \left[\tilde{p}(z_t^{b_t}, z_{1:t-1}^{b_{t-1}})\right] \cdot \prod_t \left[\frac{w_{t, b_{t+1}}^{\Delta}(z_t^{b_t}; x_t, z_{t-1}^{b_{t-1}}, z_t^{a_t^{<b_{t+1}}})}{\sum_l w_t^{b_{t+1}, l}}\right] \\
&\quad \cdot \phi(x^{\neg b}, a^{\neg b} | x^b, b) \\
&= p(z^b, x) \cdot W \cdot \phi(z^{\neg b}, a^{\neg b} | z^b, b)
\end{aligned}$$

ただし，以下のように置いた．

$$W = \prod_t \left[\frac{w_{t, b_{t+1}}^{\Delta}(z_t^{b_t}; x_t, z_{t-1}^{b_{t-1}}, z_t^{a_t^{<b_{t+1}}})}{\sum_l w_t^{b_{t+1}, l}}\right]$$

すると $q(z_{1:T}) = p(z) \sum_{b_{1:T}} \mathbb{E}_{\phi^{\neg b}}\left[W\right]$ となるので，真の対数尤度がキャンセルする．

$$\begin{aligned}
&\mathbb{E}_q\left[\log p - \log q\right] \\
&= \sum_z q(z)(\log p - \log p - \log \sum_b \mathbb{E}_{\phi^{\neg b}}\left[W\right]) \\
&= - \sum_z p(z) \sum_b \mathbb{E}_{\phi^{\neg b}}\left[W\right] \log \sum_b \mathbb{E}_{\phi^{\neg b}}\left[W\right]
\end{aligned}$$

次に期待値と対数尤度に使われているサンプリング分布の b についての周辺化を人工的な一様分布 $p(b_{1:T}) = \frac{1}{N^T}$ で置き換え，式に含まれる凸関数 $f(x) = -x \log(x)$ に対してイェンセンの不等式を適用する．

$$\begin{aligned}
&\mathbb{E}_q\left[\log p - \log q\right] \\
&= - \sum_z p(z) \sum_b \frac{1}{N^T} \mathbb{E}_{\phi^{\neg b}}\left[N^T W\right] \log \sum_b \frac{1}{N^T} \mathbb{E}_{\phi^{\neg b}}\left[N^T W\right] \\
&\geq \mathbb{E}_{p(z)}\left[\mathbb{E}_{p(b)}\left[\mathbb{E}_{\phi^{\neg b}}\left[N^T W \log(N^T W)^{-1}\right]\right]\right] \\
&= \sum_z \sum_b \sum_{a^{\neg b}, z^{\neg b}} p(z) W \phi^{\neg b} \cdot \left[\log(N^T W)^{-1}\right]
\end{aligned}$$

すると得られる下界はサンプリングの同時分布の期待値と，計算可能なその中身によって成り立っているため，モンテカルロ近似で算出可能な学習目標となる．

$$L(\lambda) = \mathbb{E}_\phi\left[\log(\frac{1}{N} W^{-1})\right]$$

この学習目標は重みが対称なとき VSMC の学習基準と一致する．ただし VSMC ではこの学習基準は提案分布の学習に用いられるのに対し，本手法では SMC を拡張する重みの学習に用いられている．またこの学習目標は下界であるため，この最大化は KL divergence を最小化することを保証しないが，長期的な依存関係が重視される場合においては，下界を高めることが KL divergence を下げることに繋がることが言える．(5.3 章で後述)

4.4 勾配法

分布にパラメータを含むため，勾配を計算すると方策勾配の項が導出される．VSMC では，この項は高分散であるとして実装上は無視されていた．しかし今回の導出ではリサンプリング目標を学習する上で，離散な確率変数 a_t^i が重要になるため，項を無視しなかった．

5. 実験

提案手法を調べるため，実世界の一部の問題を簡略化したトイプロブレムを構成する．

この構成とそれが簡略化していると考えたものについてまず説明し，次に SMC を使用すると起きる問題について説明する．このとき提案手法の学習基準はより良いパラメータを測ることができることを確認し，実際の実験設定について説明する．得られた結果から，提案手法がトイプロブレムで SMC より正確に潜在変数を推定できることを主張する．

5.1 保留期間の長いトイプロブレム

潜在変数も観測変数も 2 状態しかない時刻 $1 \leq t \leq T$ の隠れマルコフモデルを考える．$t = 1$ で潜在変数は一様な確率で 2 状態を取り，時刻 T まで観測変数は潜在変数と独立して決まり，潜在変数は確率 ε_τ で別状態に変化して遷移し，時刻 T で確率 $1 - \varepsilon$ で潜在変数の状態に対応する観測変数が得られる生成モデルを考える．このモデルでの時刻 T までを保留期間とし，時刻 1 で保留された 2 状態を T まで縮退せずに保持できるかを見ることができる．

保留されるべき 2 状態は，動画や音声情報などを潜在変数に符号化した場合の扱いきれない状態数のうち，ある時刻で同程度の尤度を持つものの簡略表現とした．また，確率 ε_τ と ε が小さいほど潜在変数は確定的に遷移して T で観測されることは，同程度の尤度を持つ候補の情報が T ステップ後の状態推定に重要な意味を持つ状況の表現とした．最後に時刻 T まで潜在変数と観測変数が独立していることは，保留期間中に候補が同程度の尤度を持つ状況が変化しない状況の表現とした．

5.2 SMC による縮退問題

SMC を用いてこの 2 状態を保持した粒子をリサンプリングすると，各リサンプリングごとに 1/2 の確率で片方を取りこぼす．これはサンプルごとに対称なリサンプリング目標を使う限り避けられない．取りこぼした片方が時刻 T に観測された場合，ε_τ と ε が小さいほど，尤度の低いサンプルとなる．

5.3 下界での評価

このトイプロブレムでは，ε を限りなく 0 に近づければ，誤ったサンプルに対する対数尤度は $-\infty$ に発散する．すると提案手法の学習基準は，非対称なリサンプリングが最後が誤ったサンプルを生成する確率を低くする限り，下界を限りなく上昇させる．そしてそのような非対称なリサンプリングを実現するパラメータは拡張重み w^Δ に十分な自由度があれば，存在することがわかる．よって，学習過程でそのようなパラメータにたどり着ける限り，下界評価によって学習の進捗を評価できる．

5.4 実験設定

このトイプロブレムを使って提案手法が長期的なサンプリング目標に備えて SMC より縮退を減少させられることを確認した．状態数が多く，取りこぼした粒子の情報を回復できない状況を表すため，トイプロブレムの遷移分布 $p(z_t|z_{t-1})$ を提案分布として固定した．トイプロブレムのパラメータは $T = 4, \varepsilon_\tau = \varepsilon = 0.0005$ とした．保留期間中は観測変数から潜在変数を推定できないため，時刻 T での真の潜在変数と推定した潜在変数を比較し，正答率を評価基準とした．比較対象として毎回リサンプリングを行う SMC と，時刻 T 以外はリサンプリングを行わない SMC を比較対象とした．

モデルのパラメータは，拡張重みのネットワークは 2 層で 1 層目の活性化関数は ELU，2 層目は線形のままにした．2 層目のユニット数は 2 とした．勾配法には Adam を使用した．バッチサイズは 32，学習率は 0.001 とした．粒子数=2 とした．

5.5 実験結果

学習ステップの間の正答率の変化をグラフにすると図 3 のようになった．初期のパラメータでは拡張重み $w^\Delta \approx 1$ なので SMC と同じ正答率だが，学習を進めるにつれてリサンプリングしない場合の正答率に近づいた．このトイプロブレムにおいて提案手法が非対称なリサンプリング目標を学習によって獲得し，SMC よりも正確に潜在変数を推定できることを示した．

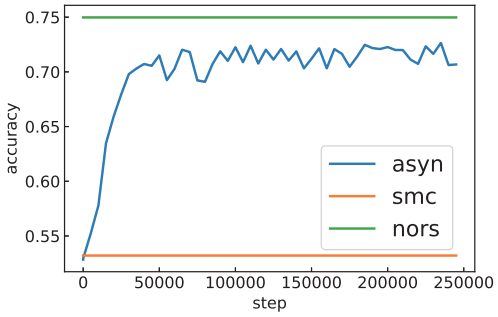


図 3: 学習に伴う最終時間でのサンプル状態の正答率の変化．橙線は SMC，緑線は最終時間のみリサンプリングした場合，青線が提案手法，のそれぞれの正答率．縮退を抑えるよう学習できている．

6. まとめ

本研究では，SMC において粒子を縮退させる原因としてリサンプリングの目標の対称性に注目し．下界によってこのリサンプリング目標を近似して学習することで，SMC のリサンプリング戦略を非対称に拡張し，トイプロブレムでの縮退の軽減を検証した．これにより SMC と同じ粒子数で，正確な潜在変数の推定を実現できる場合が存在することが示された．

一方，今後の課題として，縮退の軽減によるサンプル効率が計算効率に繋がるかの解明が挙げられる．非対称なサンプリングの内，リサンプリング結果に依存する部分は並列化できないため，計算効率が上がるとは限らないからである．今回は明示的に非対称性を学習するためにリサンプリング結果を引数としたが，これは非対称性からの必要条件かは不明である．これについては今後検証していきたい．

謝辞

本研究は JSPS 科研費 JP15H05327,JP16H06562 の助成を受けたものです。

参考文献

[Chung 15] Chung, J., Kastner, K., Dinh, L., Goel, K., Courville, A. C., and Bengio, Y.: A Recurrent Latent Variable Model for Sequential Data, in Cortes, C., Lawrence, N. D., Lee, D. D., Sugiyama, M., and Garnett, R. eds., *Advances in Neural Information Processing Systems 28*, pp. 2980–2988, Curran Associates, Inc. (2015)

[Gordon 93] Gordon, N. J., Salmond, D. J., and Smith, A. F. M.: Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proceedings F - Radar and Signal Processing*, Vol. 140, No. 2, pp. 107–113 (1993)

[Kingma 13] Kingma, D. P. and Welling, M.: Auto-encoding variational bayes, *arXiv preprint arXiv:1312.6114* (2013)

[Naesseth 17] Naesseth, C. A., Linderman, S. W., Ranganath, R., and Blei, D. M.: Variational Sequential Monte Carlo, *ArXiv e-prints* (2017)

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

テキストチャットにおけるクエリ終了タイミングの検出

End-of-Query Detection in Text Chat System

田中 駿
Shun Tanaka

株式会社サイバーエージェント
CyberAgent, Inc.

I propose a prediction model for end-of-query detection. It is a common practice that a chatbot system responds an answer when they get a single utterance from the user. A problem arises when a human breaks their query into multiple utterances and a chatbot system can not respond appropriately, leading to lower user experience. In this paper, I propose a prediction method based on LSTM for end-of-query detection.

1. はじめに

近年，コールセンターにおける人件費削減などを目的として，顧客からの問い合わせに自動応答を行う対話システムを提供する事業者が増えている．事業者が提供する多くの対話システムでは，ユーザから送られた発話 *1 に対して適切な返答を用例から検索，もしくは生成し応答する，一問一答形式の対話システムが採用されている．その一方で，人間の発話には 1 つの会話のまとまり (クエリ) を複数発話に分けて送信する現象が見られる場合があり，複数発話に分割されたクエリに対して，一問一答形式の対話システムでは適切な応答を返せない場合がある．

表 1 にブログサービスのコールセンターにおける，オペレータとユーザの対話例を挙げる *2．ユーザからの発話，#2，#3 は「以前記事を投稿していたブログに記事を投稿しようとしたところ，ログインができなかった (故にログイン ID，パスワード等の再設定を行いたい)」という趣旨の一つのクエリであるが，ユーザはこのクエリを 2 つの発話に分割して発話している．理想的には，#2，#3 を合わせて一つのクエリと捉え，「ログインできない場合のアカウントの再設定方法」などの回答を提示することが適切と考えられるが，現状の対話システムでは，#2 の発話を受信した時点で，用例との類以度検索を行い，「ブログへの記事投稿の方法」などの回答を提示してしまう可能性がある．このように複数発話に分割されたクエリを適切に扱えないことは，応答精度の低下に伴うユーザ満足度 (ユーザエクスペリエンス) の低下や顧客流出を招く要因になると考えられる．

そこで本研究では，対話システムがより適切に応答できるよう，ユーザから送信された発話に対して，その発話でクエリとして完結するか否かを予測する手法を提案する．

表 1: ユーザとオペレータの対話例

#	発話者	発話内容
1	オペレータ	ログインできないとのことですが、現在どのような状態でしょうか？ ID・登録メールアドレスのアカウント情報と、ログインできない詳細をお教えください。
2	ユーザ	以前に記事を投稿していましたが
3	ユーザ	久しぶりに投稿しようとしたらログインできませんでした。

2. 先行研究

Lala ら [Lala 17] は人と音声対話を行うロボットが適切なタイミングであいづちを打てるようにするため，ロジスティック回帰を用いて直近 500ms 以内にあいづちをすべきかどうかを予測するモデルを提案した．CRF を用いて特徴的な単語に対してフォーカスする手法，決定木を用いてフォーカスした単語に対して質問を返すべきか，単語を復唱するべきかを決定する手法，ロジスティック回帰を用いたロボットが返答すべきタイミングかどうかを判別する手法を提案した．それぞれの手法に対する個別の評価と，それらを併せてアンドロイド Erica に導入した際の評価を行った．

Skantze ら [Skantze 17] は，音声対話において，音声データから素性 (音声区間，ピッチ，パワー，品詞など) を抽出し LSTM の入力とすることで連続的に発話の交代を予測する 3 つのモデルを提案した．タスク指向型の音声対話データである HCRC MapTaskCorpus [Thompson 93] を使用し，発話が止まった際に次の話者を予測するタスクと，発話開始時に一連の発話の長さを予測するタスクで実験・評価を行い，従来のルールベースなどの手法と比べ精度が向上することを示した．

これらの研究は，系列情報をモデル化できる CRF や LSTM などを用いて，ユーザの音声発話データに含まれる情報をもとに発話の転換点を推定するというものであり，言語情報だけでなく，ピッチやパワーなどのいわゆる non-verbal な情報を多く用いて解析を行っている．その一方でテキストチャットにおいては，会話相手の表情や音声など，non-verbal な情報を取得することができず，verbal な情報のみで推定を行う必要があり，言語情報のみを対象にして発話の転換点を推定するタスクは今までに行われていない．

連絡先: 株式会社サイバーエージェントアドテク本部 AI Lab, tanaka_shun_xa@cyberagent.co.jp

*1 本稿では，テキストチャットにおいて送信ボタンを押した際に送信される文字列の単位を発話，ユーザとボット・オペレータの一連の発話の集合を対話と呼んでいる．

*2 情報保護の観点から，実際の対話ログをもとに類似する対話を作成した．

表 2: データセットの内訳

データセット	対話数	ユーザ発話数	ラベル 1 の割合	ラベル 0 の割合
学習	22,160	140,170	0.79	0.21
開発	2,770	17,476	0.80	0.20
評価	2,770	17,808	0.78	0.22

3. データ作成

本研究では，弊社の子会社でチャットボット事業を展開している，株式会社 AI メッセンジャー (AIM)[*3] での対話ログを用いてデータセット作成を行った．AIM の提供するサービスでは，ユーザから発話が送信された際，基本的にはボットが一問一答形式での応対を行うが，ユーザの問題を解決できなかった場合や，ユーザがオペレータとの会話を希望した際は人間のオペレータが応対を行うことができるサービスとなっている．

ユーザとオペレータの会話は人間同士の会話であるため，対話相手の発話にまだ続きがありそうだと判断した際は対話相手の発話の続きを待つという行動が見られる．そこで，ユーザとオペレータの応答ログをもとに，特定の発話に続きがあるか否かの 2 値ラベル (end-of-query ラベル) を付与しデータを作成する．end-of-query ラベルは，ユーザからの発話に続きがある場合に 0，発話に続きがなくオペレータに話者が交代している場合に 1 を付与した [*4]．

2016 年 9 月 25 日から 2017 年 9 月 24 日までの 1 年分のデータを収集した結果，計 27,700 対話のデータが得られた．それらのデータから学習，開発，評価データセットを作成するために対話数が 8:1:1 となるようにランダムにデータを分割しデータセットの作成を行った．データセットの内訳を 2，ラベル付けされた対話データの例を 5 に示す．

4. 予測モデル

本研究ではベースラインとしてロジスティック回帰 (L2LR) を用い，LSTM を用いた手法との比較を行う．それぞれの手法の概要を以下に示す．

- L2LR :
  単語 unigram(表層形) に対するベクトルを素性としロジスティック回帰を学習したもの
- LSTM :
  単語 unigram(表層形) に対して分散表現ベクトルを求め，LSTM の入力としたもの

なお，LSTM は対話におけるユーザが発した直前の発話情報は保持せず，各発話毎に end-of-query の予測を行うものとした．

5. 実験

5.1 実験設定

形態素解析器として MeCab[Kudo 05]，形態素解析用辞書として mecab-ipadic-neologd[Sato 15] を用いた．ロジスティック回帰の学習には LIBLINEAR[Fan 08] を使用した．提案手法における単語分散表現は日本語 Wikipedia のダンプデータから予め word2vec[Mikolov 13] を学習したものを使用した．提案手法のモデル学習には Adam[Kingma 14] を用いた．なお，モデルの学習の際，単語の分散表現は学習対象外とした．モデルの隠れ層の次元 $n = \{32, 64, 128\}$ でそれぞれ実験し，単語分散表現の次元を 300，イテレーション数 $i = 150$ を上限として実験を行った．

*3 https://www.ai-messenger.jp/

*4 適切なクエリのまとまりが取れない場合もあるが，アノテーションコストと比較した結果，自動構築できる方策を取った．

表 3: 実験結果

手法	Accuracy	F 値
L2LR	0.778	0.875
LSTM	**0.786**	**0.878**

表 4: L2LR において重みの多くかかった単語例

単語	重み
何	0.01797
言わ	0.01202
か	0.01017
また	0.007
ページ	0.00592
まし	0.0059
ログイン	0.00484
ん	0.0046
ませ	0.00449
しよ	0.00298

5.2 評価方法

手法の評価方法として，予測ラベルと正解ラベルの一致率である Accuracy と，Recall と Precision の調和平均である F 値を用いる．

5.3 パラメータチューニング

L2LR, LSTM ともに，開発データセットを用いて最も高い F 値を出すパラメータを検証した．その結果, L2LR では，正則化パラメータのコスト $c = 2^{-5}$，閾値 $\theta = 0.39$，LSTM では，隠れ層の次元 $n = 128$，イテレーション数 $i = 146$ となった．

5.4 実験結果と考察

それぞれの手法における評価データセットに対する実験結果を 3 に示す．

ベースラインである L2LR と提案手法である LSTM を比較すると，Accuracy，F 値のどちらにおいても LSTM が L2LR を上回る結果となった．対話例全文に対する各手法による予測結果を表 5，対話例中に出現する 115 単語において L2LR での重みの大きい上位 10 単語を表 4 に示す．L2LR では，“ます”の活用形である “まし”などに重みがかかっていることが分かるが，時系列を考慮できるモデルではないため，#2「以前に記事を投稿していましたが」の発話に対しても end-of-query の確率が高くなっていることが伺える．その一方で，LSTM では単語の系列を考慮することができるため，#2 の発話に対して end-of-query の確率を低く予測することができていることが分かる．

6. まとめと今後

本研究では，ユーザとオペレータの対話ログを用いて，テキストチャットにおけるユーザの発話がクエリとして完結してい

るか否かの予測に取り組んだ．実験の結果，ロジスティック回帰を用いたモデルより LSTM を用いた提案手法の精度が高くなり，発話に含まれる単語の系列をモデルに組み込むことでわずかながらも精度が向上することがわかった．

今後はこのモデルを改良し，ユーザ毎の発話特徴やユーザの一連の発話，オペレータの直前の発話内容の情報などを加味したモデルを実現したいと考えている．また，実システムへの導入を行った結果の UX の評価や，雑談対話のデータにおける実験も行いたいと考えている．

参考文献

[Fan 08] Fan, R.-E., Chang, K.-W., Hsieh, C.-J., Wang, X.-R., and Lin, C.-J.: LIBLINEAR: A Library for Large Linear Classification, *Journal of Machine Learning Research*, Vol. 9, pp. 1871–1874 (2008)

[Kingma 14] Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, *CoRR*, Vol. abs/1412.6980, (2014)

[Kudo 05] Kudo, T.: MeCab : Yet Another Part-of-Speech and Morphological Analyzer, *http://mecab.sourceforge.net/* (2005)

[Lala 17] Lala, D., Milhorat, P., Inoue, K., Ishida, M., Takanashi, K., and Kawahara, T.: Attentive listening system with backchanneling, response generation and flexible turn-taking, in *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*, pp. 127–136 (2017)

[Mikolov 13] Mikolov, T., Chen, K., Corrado, G., and Dean, J.: Efficient Estimation of Word Representations in Vector Space, *CoRR*, Vol. abs/1301.3781, (2013)

[Sato 15] Sato, T.: Neologism dictionary based on the language resources on the Web for Mecab (2015)

[Skantze 17] Skantze, G.: Towards a General, Continuous Model of Turn-taking in Spoken Dialogue using LSTM Recurrent Neural Networks, in *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*, pp. 220–230 (2017)

[Thompson 93] Thompson, H. S., Anderson, A., Bard, E. G., Doherty-Sneddon, G., Newlands, A., and Sotillo, C.: The HCRC Map Task Corpus: Natural Dialogue for Speech Recognition, in *Proceedings of the Workshop on Human Language Technology*, HLT '93, pp. 25–30, Stroudsburg, PA, USA (1993), Association for Computational Linguistics

表 5: オペレータとユーザの対話例

#	発話者	発話内容	end-of-query ラベル	LSTM	L2LR
1	オペレータ	ログインできないとのことですが、現在どのような状態でしょうか？ ID・登録メールアドレスのアカウント情報と、ログインできない詳細をお教えください。	-	-	-
2	ユーザ	以前に記事を投稿していましたが	0	0.320	0.516
3	ユーザ	久しぶりに投稿しようとしたらログインできませんでした。	1	0.571	0.673
4	オペレータ	アカウント情報をお教えいただけますか？ ID または、ブログ URL ・登録メールアドレス ※現在登録メールアドレスがご利用できない場合は、その旨もお教えください。	-	-	-
5	ユーザ	ID もパスワードも忘れてしまっていて	0	0.433	0.593
6	ユーザ	ID を忘れたときのページで自分のメールアドレスを入力しても	0	0.554	0.565
7	ユーザ	メールアドレスが違うと言われます。	1	0.676	0.641
8	オペレータ	ログインしたいアカウントを確認しますので、 ・ブログ URL ・ブログタイトルとニックネーム ・その他登録覚えのあるメールアドレス 上記をお教えいただけますか？	-	-	-
9	ユーザ	https:// ID / アカウント名	1	0.643	0.573
10	オペレータ	ご連絡ありがとうございます。 ご連絡のアカウントは、ドメイン（ メールドメイン名 ）で登録がございます。 登録メールアドレスにお心あたりはございませんか？	-	-	-
11	ユーザ	メールアドレス でしょうか？	1	0.697	0.814
12	オペレータ	一致いたしました。 登録メールアドレスの変更を行いますので、ご利用可能なメールアドレスをお教えいただけますか？	-	-	-
13	ユーザ	メールアドレス でお願いします。	1	0.905	0.925
14	オペレータ	お待たせしました。 登録メールアドレスの変更を行いましたので、パスワード再設定をお願いいたします。 パスワード再設定: 再設定 URL	-	-	-
15	ユーザ	ありがとうございます！ 書き込む事が出来ました！	0	0.645	0.888
16	ユーザ	また何かありましたら、宜しくお願いします。本当にありがとうございました。	1	0.896	0.957
17	オペレータ	無事解決のご連絡ありがとうございます！	-	-	-

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

ディベート型人工知能によるサポート性推定に基づく反論生成

Debating AI: Counterargument Generation Based on Textual Supportiveness Recognition

佐藤 美沙 Misa Sato　柳井 孝介 Kohsuke Yanai　柳瀬 利彦 Toshihiko Yanase　是枝 祐太 Yuta Koreeda　黒土 健三 Kenzo Kurotsuchi

日立製作所 研究開発グループ
Research & Development Group, Hitachi, Ltd.

We will demonstrate a counterargument generation system in debating, which aims to utilize newswire articles for decision-making support. Users can specify a claim such as “We should legalize casinos because they promote economy.” and then the system outputs counterargument scripts against the claim.

1. はじめに

筆者らの研究グループでは，電子化されたテキストデータを意思決定支援に活用することを目的として，ディベートをするシステムを開発している [Sato 15]. このシステムを本稿ではディベート型人工知能と呼ぶ. 本発表ではディベート型人工知能による反論生成のデモンストレーション展示を行う. 聴講者が主張をテキスト文で与えると，このシステムは，反論する主張およびその理由や根拠を提示する. たとえば「カジノは経済を活性化させる」という主張が与えられると，反論として「カジノは経済を活性化させない」という反対側の主張を提示し，続いてその理由や根拠となる事例を新聞記事等から抽出して提示する.

反論においては，単に議題に対する肯定・否定の立場が相手と反対であれば良いわけではなく，相手の主張と噛み合った議論を行う必要がある. 本研究では，入力される主張と同じ観点に基づいて反対側の立場から論じることにより，話の噛み合った反論を行う.

本研究では，入力された主張の依拠する価値を認識して，それを踏まえた反論側の主張を作成し，続けて反論側の主張を支持する根拠を提示するシステムを作成する. 指定の価値についての反論側根拠を集めることを可能とするために，データソースから抽出したテキストが，ある価値に依拠する主張に対する根拠であるか否かを推定するサポート性推定 [佐藤 16b]の問題を組み込む. よって本稿では特にサポート性推定に着目し，従来のタスクで有用であった特徴量の有用性を調べることによりサポート性推定の特徴を調べる.

2. 関連研究

議論解析は近年 Argument Mining 分野として自然言語処理の観点から多くの研究がなされている. 反論に関するものとしては，議論中に出現する 2 つの文の関係が支持（support）であるか反論（attack）であるかを推定する Relation-based Argument Mining というタスクが提案されている[Carsteins15,Cocarascu 17]. また，[Bosc 16]ではツイート間が支持関係にあるか反論関係にあるかを識別している. これらは議論内の構造を明らかにするものであり，与えられた主張に対して反論を提示するものではない. また，どの観点で反論しているかは取り扱っていない.

反論の生成に関する研究としては，Bilu らは，与えられた主張文を否定する主張文を生成する手法を提案している[Bilu 15].

連絡先：佐藤美沙，株式会社 日立製作所 中央研究所，
〒185-8601 東京都国分寺市東恋ヶ窪一丁目 280 番地，
042-323-1111，misa.sato.mw@hitachi.com

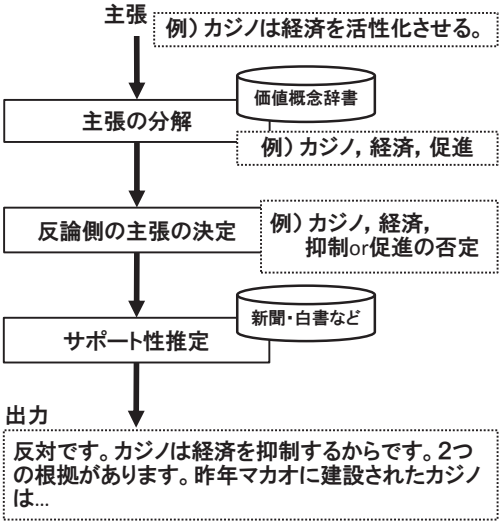


図 1. 反論生成システムの全体フロー

主張文に否定語“not”を加える，もしくは，感情語の極性を反転させるなどのルールベースの手法によって否定側の主張文を生成し，生成された否定側の主張文がディベートで使われた際にもっともらしいか否かを統計的な手法によって検証している. この研究では根拠の提示は行っていない.

本研究では，与えられた主張が問題としている議題および依拠する価値を推定し，議題から価値への影響が主張とは反対方向であることを示す根拠を探すことにより，反論を生成する. これにより，与えられた主張の観点を踏まえた反論を生成することができると考えられる.

3. ディベート型人工知能による反論生成

本研究では，[Sato 15]に説明しているディベート型人工知能の機能を一部組み替えて反論生成を行う. 入力される主張には，「カジノは経済を活性化させる」のような議論の対象と効果を示す文を想定している. 効果を論じる際に，価値というものを取り扱う. ここで言う価値とは，それ自体が人あるいは社会において良い存在（悪い存在）と捉えられるものであり，たとえば健康は良い価値，貧困は悪い価値，と考えることができる. つまり，ディベート型人工知能における肯定的な主張とは，議論の対象が良い価値を増やしていることへの指摘，あるいは，悪い価値を減らしていることの指摘であり，否定的な主張とは，議論の対象が悪い価値を増やしていることへの指摘あるいは良い価値を減らしていることの指摘である.

ディベート型人工知能における反論生成とは，与えられた主張に対して，それが肯定的な主張であれば否定的な主張およびその根拠を示し，否定的な主張であれば肯定的な主張およびその根拠を示すことである．図1に示す通り，全体フローは大きく以下の3ステップから成る．

1. 主張の分解
2. 反論側の主張の決定
3. 反論側の根拠の抽出

本章では各ステップの詳細を説明する．

3.1 主張の分解

主張の分解では，入力された主張文から議論の対象と価値を抜き出す．抜き出した後，議論の対象が価値に及ぼす影響の方向を推定する．影響の方向には，促進関係，抑制関係，促進関係の否定，抑制関係の否定，影響関係なし，の 5 通りを想定する．たとえば，「カジノは経済を活性化させる」という主張が入力された場合は，議論の対象が「カジノ」，価値が「経済」，影響の方向は「促進」となる．

議論の対象の抽出および価値の同定は，構文構造に対するルールベースの手法によって行う．優先度の付いた 5 つのルールを順番に適用し，合致した場合にそれを抽出する．最も優先度の高いルールは，影響関係に関する述語を同定し，その述語の主語を議論の対象，目的語を価値とするルールである [佐藤 16a]．影響関係に関する述語については約 350 語の辞書を用意している．

価値は，価値体系辞書によって分類を定義する [柳井 15]．価値体系辞書は 15 の大価値とそれにぶら下がる合計 69 の小価値からなる．それぞれの小価値には表現語が定義されている．表 1 に価値体系辞書の一部を示す．主張から価値を抜き出す際には，抽出した目的語がどの小価値にあたるかを表現語との単語列類似度により分類する．

次に影響の方向を推定する．ここでも構文構造に対するルールベースの手法を用いる．促進関係を 1，抑制関係を-1 とし，述語の極性と，述語に対する否定の係り受けの有無を掛け合わせることにより，方向を計算する．

表 2. データセットの内訳

分類	数
計	7955
影響関係なし	3784(48%)
影響関係あり	4171(52%)
促進関係	2861(68%)
抑制関係の否定	219(5%)
抑制関係	701(17%)
抑制関係の否定	390(9%)

3.2 反論側の主張の決定

反論側の主張においても，議論の対象が価値に影響を及ぼすという考え方を用いる．反論の依拠する価値を決定し，主張を決定する．入力された主張の価値を元に反論側の価値を決定するが，本研究では，元の主張の価値をそのまま反論側でも採用する．そして，影響の方向を反転させる．つまり，「カジノが経済を活性化させる」という主張に対しては，「カジノが経済を抑制する」という反論側の主張を行う．

3.3 反論側の根拠の抽出

前節で採用した反論側の主張を支持するような事例を新聞記事などのテキストから収集し，反論側の根拠として提示する．

サポート性推定では，入力されたテキストについて，議論の対象が価値に対して与える影響が，促進関係か，抑制関係か，あるいは影響関係が無いかを推定する． 根拠候補文のサポートする主張の影響関係の方向が反論側の主張と同一である場合，反論側の根拠であるとして抽出する．ここで反論側の根拠として抽出された文が最終的な反論文として表示されるため，サポート性推定の精度が高いほど，見逃しが少なく，誤りの少ない，良い反論生成システムになる．

最後に，信頼度の高い 2 文を選択し，反論側の根拠として提示する．

4. サポート性推定

4.1 手法

サポート性推定は，データソースから抽出したテキストが，主張を支持（否定）する根拠としてどの程度もっともらしいかを推定するタスクである [佐藤 16b]．サポート性推定の特徴を明らかにするため，関連タスクである根拠検知と賛否識別で使われている特徴量によってサポート性の推定がどの程度可能かを調べる．加えて，サポート性推定特有の特徴量の追加による推定精度への影響を調べる．

まず，根拠検知（Evidence Detection; ED）で有用な特徴量として，[Rinott 15]で使われている以下の 5 つを利用する．

・意味的類似度： 主張文と根拠候補文の間の類似度．Word2Vec のコサイン類似度[Mikolov 13]により計算する．

・センチメントの一致度： 根拠候補文のセンチメント値を[0,1]に正規化した値．センチメント値の導出には Stanford Core NLP [Manning 14] を用いた．

・語彙： 根拠文によく出現する語を集めた辞書を用意し，その辞書内の語がそれぞれ出現しているか否かを表す 1 か 0 の値．辞書は手作業で集めた 287 語からなる．専門家や公的機関の発言は説得力の元となるため，これらを表す語は根拠文に出現しやすいという考えの下に作成されている．

・固有表現クラス： 根拠候補文がそれぞれの固有表現クラスにあたる固有表現を含むか否かを表す 1 か 0 の値．固有表現およびクラスは Stanford Core NLP により抽出した．

表 1. 価値体系辞書の一部

大価値	小価値	極性	表現語
economy	economy	+1	economy, investment, development, …
〃	employment	+1	employment, job, …
〃	loan	-1	loan, debt, …
health	health	+1	health, healthy, …
〃	disease	-1	disease, complication, …
safety	crime	-1	crime, prostitution, …

・文パターン: 文が引用を含むか，人あるいは組織に寄る発言を含むか，人あるいは組織による行動を含むか，人あるいは組織の意見を含むか，を表すパターンを作成し，根拠候補文がそれぞれに当てはまるか否かを表す 1 か 0 の値.

次に，賛否識別（Stance Classification; SC）で有用性が知られている [Mandya 16]，以下の 1 つの特徴量を根拠候補文に対して適用する.

・依存関係: 根拠候補文において依存関係にある 2 語のペアを表す one-hot ベクトル.

最後に，サポート性推定（Supportiveness Recognition; SR）特有の特徴量として以下の 2 つを用いる. 議論対象と価値の関係に着目する問題であるため，候補文内のどの語句が議論対象および価値と対応するかを考慮した特徴量としている.

・機能語: 根拠候補文中の議論対象を表す語と価値を表す語の間に出現する動詞，名詞，副詞あるいは形容詞に対する one-hot ベクトル.

・否定語: 根拠候補文中の議論対象を表す語と価値を表す語の間に否定語が含まれるか否かを表す 1 か 0 の値. 否定語の辞書は，一般的な”not”などの否定語に加え，”ban”，“barrier”，“drop”，”hazard”などの極性を反転させるような語を含む 154 語からなる.

4.2 実験設定

データセットは，Annotated English Gigaword [Napoles 12] から候補となる文を取得し，人手でラベリングすることにより作成した. 候補文の取得は，31 通りの主張（議論の対象と価値のペア）について，議論の対象と価値の表現語の共起する文を抽出することにより行った. 表 2 にデータセットの統計値を示す.

反論側の根拠の候補には，(a)反論側主張と影響関係の方向が合致するもの，(b)影響関係の方向が反対であるもの，(c)影響関係がないもの，の 3 通りが存在する. 表 2 のラベルとの対応を考えると，反論側主張の影響関係の方向が促進関係であるとすると，(a)は促進関係あるいは抑制関係の否定，(b)は抑制関係あるいは促進関係の否定，(c)は影響関係なしが対応する. そこで，影響関係の有無を識別する（(a)と(b)あるいは(c)を識別する）2 値問題と，促進関係と抑制関係を識別する（(a)と(b)を識別する）2 値問題のそれぞれについて実験を行った.

前節で示した特徴量を用い，線形 SVM 識別器によって識別を行った. 3-fold の交差検定により精度を測定している. ただし，同じ主張に基づく根拠文が訓練データとテストデータの両方に含まれることはないように分割している. パラメータは，$C = \{10^{-2}, 10^{-1}, 10^{0}, 10^{1}, 10^{2}\}$とした. また，ベースラインの特徴量として Bag-of-Words（BoW）を用いた.

4.3 実験結果

表 3 に影響関係の有無の識別結果を示す. BoW，根拠同定および根拠分類の特徴量を用いた[BoW+ED+SR]が最も F 値が高く 57%となった. しかしながらいずれの特徴量の組み合わせも，BoW ベースライン特徴量の結果からさほどの向上は見られなかった. 類似のタスクの特徴量ならびに本稿で提案した特徴量は有用とはいえない.

表 4 に促進関係と抑制関係の識別結果を示す. [BoW]の結果に比べて[SC]は F 値が 5%高く，賛否分類の特徴量が有用であったことがみてとれる. 加えて，BoW 特徴量はうまく働かなかった. 賛否分類タスクにおいても BoW 特徴量はあまり機能しない [Mandya 16] ことから，この点も賛否分類との共通点といえる. しかしながら，賛否分類の特徴量にサポート性推定特有の特徴量を加えた[SC+SR]が最も良い精度となった. 議論の対象と価値に着目した特徴量によって精度の向上が見られた.

表 3. 影響関係の有無の識別実験結果

特徴量	適合率	再現率	F 値
[ED]	55%	55%	54%
[SC]	57%	56%	56%
[SC+SR]	56%	56%	56%
[BoW]	57%	56%	56%
[BoW+ED]	56%	55%	55%
[BoW+ED+SR]	**58%**	**57%**	**57%**

表 4. 促進関係と抑制関係の識別実験結果

特徴量	適合率	再現率	F 値
[ED]	58%	55%	56%
[SC]	62%	69%	64%
[SC+SR]	**68%**	**71%**	**69%**
[BoW]	61%	57%	59%
[BoW+ED]	62%	59%	60%
[BoW+ED+SR]	64%	63%	63%

5. おわりに

本稿では，反論生成システムの全体フローについて述べ，サポート性推定について実験を行った. サポート性推定というタスクを中心に据えることで，主張の依拠する価値を踏まえた反論を生成するシステムを開発した. 一方でサポート性推定については精度が未だ十分ではなく，特に影響関係の有無の識別については今回の実験では有効な特徴量を見つけることができなかった. エラー分析などにより問題の分析を進めていく.

今後は，精度向上と共に，立論システムと反論システムを統合して競技ディベートに出場可能なシステムとするため，立論から反駁まで一貫性のある議論を展開できるような手法の開発を進めていく.

参考文献

[Bilu 15] Yonatan Bilu, Daniel Hershcovich, and Noam Slonim: Automatic claim negation: why, how and when, *Proceedings of the 2nd Workshop on Argumentation Mining*, pp.84-93, Association for Computational Linguistics, 2015.

[Bosc 15] Tom Bosc, Elena Cabrio, and Serena Villata: Tweeties Squabbling: Positive and Negative Results in Applying Argument Mining on Social Media, *Computational Models of Argument (COMMA)*, pp.21-32, IOS Press, 2016.

[Carsteins 15] Lucas Carstens and Francesca Toni: Towards relation based argumentation mining, *Proceedings of the 2nd Workshop on Argumentation Mining*, pp.29-34, Association for Computational Linguistics, 2015.

[Cocarascu 17] Oana Cocarascu and Francesca Toni: Identifying attack and support argumentative relations using deep learning, *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pp.1374-1379, Association for Computational Linguistics, 2017.

[Mandya 16] Angrosh Mandya, Advaith Siddharthan, and Adam Wyner: Scrutable Feature Sets for Stance Classification, *Proceedings of the 3rd Workshop on Argumentation Mining*, pp.60-69, Association for Computational Linguistics, 2016.

[Manning 14] Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard and David McClosky: The Stanford CoreNLP natural language processing toolkit, *Proceedings of the ACL 2014: System Demonstrations*, pp.55-60, Association for Computational Linguistics, 2014.

[Mikolov 13] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado and Jeff Dean: Representations of Words and Phrases and their Compositionality, *Proceedings of the Third Workshop on Argument Mining*, pp.3111-3119, Curran Associates, Inc., 2013.

[Napoles 12] Napoles, Courtney, Matthew Gormley, and Benjamin Van Durme : Annotated English Gigaword LDC2012T21, *Philadelphia: Linguistic Data Consortium*, 2012.

[Rinott 15] Ruty Rinott, Lena Dankin, Carlos Alzate Perez, M. Mitesh Khapra, Ehud Aharoni and Noam Slonim: Show me your evidence - an automatic method for context dependent evidence detection, *Proceedings of the EMNLP 2015*, pp.440-450, Association for Computational Linguistics, 2015.

[Sato 15] Misa Sato, Kohsuke Yanai, Toshinori Miyoshi, Toshihiko Yanase, Makoto Iwayama, Qinghua Sun and Yoshiki Niwa: End-to-end Argument Generation System in Debating, *Proceedings of the ACL-IJCNLP 2015 System Demonstrations*, pp.109-114, Association for Computational Linguistics, 2015.

[佐藤 16a] 佐藤 美沙, 柳井 孝介, 柳瀬 利彦, 是枝 祐太, 丹羽 芳樹: ディベート人工知能における影響関係認識のためのテキスト内の論理構造に関する考察, 人工知能学会全国大会論文集, 第 30 巻, 4B1-3, pp.1-3, 人工知能学会, 2016.

[佐藤 16b] 佐藤 美沙, 柳井 孝介, 柳瀬 利彦, 三好 利昇, 是枝 祐太, 丹羽 芳樹: 意見文章自動生成のための組合せ構文特徴を用いたサポート性推定, 人工知能学会論文誌, 第 31 巻, no. 6, pp. AI30-L_1-12, 人工知能学会, 2016.

[柳井 15] 柳井 孝介, 三好 利昇, 柳瀬 利彦, 佐藤 美沙, 丹羽 芳樹, Reisert, P., 乾 健太郎:ディベート人工知能における意見生成,人工知能学会全国大会論文集, 第 29 巻, 3M3-2in, pp. 1-3, 人工知能学会, 2015.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

食知識に基づく連想型レシピ提案対話システム

Associative Dialog System for Recipe Recommendation Using Food Knowledge

山上 勝義 *1*2 Katsuyoshi Yamagami
遠藤 充 *2 Mitsuru Endoh
清丸 寛一 *3 Hirokazu Kiyomaru
黒橋 禎夫 *3 Sadao Kurohashi

*1産業技術総合研究所 Advanced Industrial Science and Technology
*2パナソニック株式会社 Panasonic Corporation
*3京都大学 Kyoto University

One of the problems when using the dialog system is that users do not know what utterances the system can understand to convey their preferences. To mitigate the problem, we developed an associative dialog system for recipe recommendation which utilizes food knowledge base. In addition to providing expressions candidates that system can accept, it also repetitively generates and respond with linguistic expressions related to foods associated with user input.

1. はじめに

我々は対話技術の実用化を目指してタスク志向の対話技術の開発をすすめている．レシピ提案をタスクとしたチャット型の対話実証実験システム CookChat[1][2] を構築，運用している．CookChat の運用を通じて得られた対話ログやアンケートなどの分析から技術課題を探り出しつつ改良を行っている．

フィードバックから得られた主な課題として，「あっさり」「パーティ向き」などの料理に関する印象や状況に関する要望を表したユーザ発話を十分に理解できない点がある．これについては，基本料理の知識ベースを構築し，料理にまつわる語彙知識を充実させ，ユーザ発話における語彙理解のカバー率を改善しつつある [3]．

一方で，翠らの研究 [4] で指摘されているように情報検索型タスクの対話システムにおいては，

- 対話システムがどのような観点でコンテンツを検索し絞り込むことができるのかをユーザが知りえないこと
- ユーザ自身も自分の嗜好を表す観点に気づいているとは限らないこと

といった課題がある．

本研究では，この課題に対して，基本料理に関する知識 [3] を用いて，料理から連想される言語表現（連想表現）を提示することで，対話システムが持つコンテンツ検索の観点をユーザに理解させるとともに，ユーザ自身が意識していなかった観点についての嗜好情報を取得する対話システムを試作した．



図 1: 連想型レシピ提案対話システムの対話動作の例

2. 連想型レシピ提案対話システム

2.1 連想型レシピ提案対話システムの概要

試作した対話システムの概要について述べる．図 1 に連想型レシピ提案対話システムでの対話動作の例を示す．ユーザと対話システムとの対話は以下の流れで進んで行く．

STEP1 ユーザが料理に関して思い浮かんだ表現をテキストで入力する (例:「おふくろの味」)

STEP2 対話システムは，入力された表現と関連が深い料理名を応答文で提示し，どのような料理を連想したかをユーザに示す

STEP3 対話システムは，料理に関係する表現（連想表現）を複数提示し，ユーザに料理の嗜好の観点を伝える

STEP4 所定の対話ターン数に達するまで，**STEP1** に戻り提示された連想表現から選択するか，あるいは，ユーザが思いついた次の表現を入力する

STEP5 **STEP4** で所定の対話ターン数に達すると，それまでにユーザが入力した表現からレシピに関する嗜好を推

連絡先: 山上 勝義，産業技術総合研究所，東京都江東区有明 2-4-7 臨海副都心センター yamagami.katsuyoshi@aist.go.jp

定し，その嗜好に近い順にランキングしてレシピをリストで提示する

STEP6 ユーザはレシピのリストから気に入ったレシピを選択する，あるいは，対話を継続するかを選択する．対話継続の場合は **STEP1** に戻る．

上記の対話動作の **STEP1** ～ **STEP4** までが連想型対話のフェーズである．このフェーズでユーザが入力した (あるいはボタン選択した) 連想表現からユーザのレシピの嗜好を理解して，**STEP5** ～ **STEP6** のレシピ提案フェーズで理解した嗜好に基いてレシピをリスト形式で提案する．このように，最初にユーザが入力した表現をきっかけとして，対話システムが連想的に料理に関する表現を提示し，ユーザの料理に関する興味・嗜好を探り出す形で対話が進行する．

以前の CookChat では，

1. ユーザが入力したテキストに含まれる料理に関するキーワードとレシピに付与されたキーワードを照合することでユーザへ提案するレシピの候補を絞り込む
2. さらに，決定木アルゴリズムで，最も絞り込みが進むキーワードを選択し，そのキーワードについてユーザの嗜好を質問して，絞り込みを進める

という流れで，ユーザの嗜好に合ったレシピを提案していた．CookChat が 1 で理解する，あるいは，2 で質問するキーワードのバラエティが少ないため，どんな表現が理解されるか分からないという課題に加え，繰り返し使っていると提案されるレシピが固定化するという課題もあった [*1]．

本連想型レシピ提案対話では，この課題についても考慮し，以前の CookChat よりも料理に関する観点の幅を広げ，ユーザの嗜好のバラエティを捉えやすくすることを狙った．

2.2 基本料理知識ベース

ユーザが入力した連想表現を理解し，対話システムが新たな連想表現を提示するために基本料理知識ベース [3] を利用した．

```
【ぶり大根】
[同義語]ブリ大根 鰤大根
[材料]<大根>生姜.7 ブリ.2 水.2 米飯.1 みりん.9 酒.9 醤油.9 砂糖.7 だし.4 食塩.3
[調理法]<煮る>
[属性]酒のアテ(105/付け合せ).5 家族の人気メニュー(201/対象).5
大人が好き(201/対象).5 家庭的(221/場所) あたたまる(231/効果・機能).5
あまじょっぱい(251/味).5 美味しい(251/味).5 トロトロ大根(252/食感).5
プリプリ(252/食感).5 生臭い(254/香り).5 よく染みてる(261/印象).5
美味しく作るのにはコツがいる(261/印象).5 部屋が臭くなる(261/印象).5
プロの技(301/作り手).5 煮崩れ(311/調理法・工程).5
美味しいけど難しい(322/難易度).5 面倒くさい(322/難易度).5
```

図 2: 基本料理知識ベースの記述内容の例

図 2 に基本料理知識ベースの知識記述例を示す．1 行目は料理名で，2 行目以降がこの料理名 (ぶり大根) に関する知識の記述部分である．[ ] で囲まれた部分は，料理名に関する知識の種類名である．[同義語] は料理名の同義表現の知識，[材料] は料理に用いられる材料の知識，[調理法] は料理の調理方法の知識，[属性] は料理の同義語・材料・調理法以外の知識である．知識の右側の小数点付きの数字は，料理がその知識をもつ信頼度を表している．< > で囲んだ部分はその料理特有の知識 (代表知識) であることを表す．代表知識の信頼度は 1.0 が付与されているものとして扱う．さらに [属性] の知識には下位の種類があり ( ) で囲んだ部分が下位の種類名 [*2] である．

連想型レシピ提案対話では，同義語，材料，調理法，属性のすべての知識をユーザが発話したレシピの嗜好を表す表現とのマッチング対象として扱う．また，この属性の知識をユーザに提示する連想表現の候補として用いる．属性の知識にはクラウドソーシングによって収集した料理に関する様々な観点の印象表現が含まれるので，ユーザが発話するレシピに関する嗜好表現を幅広く捉え，かつ，ユーザに多様な観点の連想を促すことが期待できる．

2.3 連想型レシピ提案対話システムの構成と動作

図 3 に連想型レシピ提案対話システムの構成と動作のフローを示す．

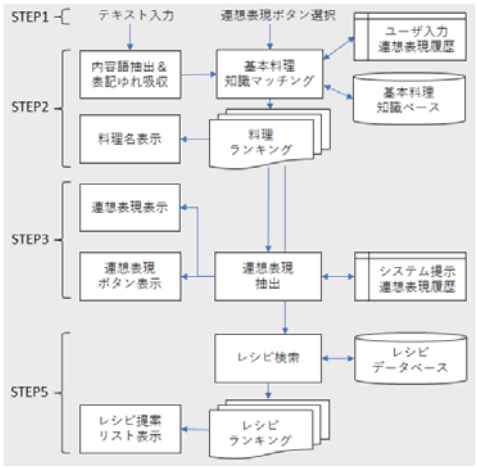


図 3: 連想型レシピ提案対話システムの構成と動作のフロー

内容語抽出・表記ゆれ吸収

ユーザのテキスト入力を形態素解析し，機能語以外の単語連鎖を内容語として抽出し，基本料理知識ベースの類義語関係を参照して表記ゆれを吸収した連想表現に変換する．

基本料理マッチング

基本料理知識ベースの料理の中から，料理がもつ知識と連想表現がマッチング (完全一致または部分一致) した料理を抽出し，マッチした知識に付与された信頼度を抽出した料理のスコア [*3] として料理のランキングを生成する．

料理名表示

料理ランキング上位の 2 つの料理名をシステムが連想した料理として応答文生成して表示する．

連想表現抽出

料理ランキングの料理上位 3 つの料理のそれぞれについて，最も信頼度が高い知識を抽出する．このときシステムが提示した連想表現の履歴を参照して既に提示したものは抽出対象からはずす．

*1 レシピの候補数が 500 未満で少ないことにも起因している

*2 '/' の右側が下位の種類名であり，'/' の左側の番号は下位の種類の体系化のために付与された番号である

*3 ユーザが入力した連想表現の履歴にある連想表現に関してもマッチングし，それぞれのマッチングでのスコアを合計する



図 4: 連想型レシピ提案対話システムでの対話例

連想表現表示
抽出した知識 (3 つ) をユーザに提示する連想表現として応答文生成して表示する．

連想表現ボタン表示
抽出した知識 (3 つ) をユーザが選択可能なボタンとして表示する．

レシピ検索
料理ランキングの各料理の料理名とレシピデータベース内のレシピ名をマッチングさせ，完全一致または部分一致したレシピを抽出する．料理名がマッチした料理のスコアをレシピのスコアとして，レシピのランキングを生成する．前記のマッチングが失敗した場合は，別途料理の上位下位関係を記述したオントロジーを用いてオントロジーの葉ノードの料理名とマッチングさせ，マッチングした料理名の上位概念の料理カテゴリー表現を用いてレシピを検索する．

レシピ提案リスト表示
レシピのランキングの上位 5 つを提案するレシピとして表示する．

上記の処理フローに従って動作する連想型レシピ提案対話システムの対話例を図 4 に示す．この対話例では，「節約」，「もやし」，「大根」はユーザがテキスト入力で行った発話，「いろんな味付けができる」はユーザがシステムが表示した連想表現のボタンを選択した発話である *4．

「節約」というユーザの発話に対して「もやしのナムル」「豆腐ハンバーグ」という基本料理が「節約」を知識としてもつ料理として検索され提示される (**基本料理マッチング**, **料理名表示**)．それら料理がもつ信頼度の高い知識を連想表現として「焼肉」「さっぱり」「家で食べる味じゃない」が抽出し表示する (**連想表現抽出**, **連想表現表示**)．この対話ターンを 4 回繰り返した後，ユーザが入力した連想表現の履歴からランキングした基本料理の料理名とレシピ名をマッチングさせてレシピのランキングを生成し表示する (**レシピ検索**, **レシピ提案リスト表示**)．なお，この対話例では，レシピ検索の際に基本料理名とレシピ名のマッチングに失敗したため，オントロジーで上位概念化してレシピ検索した結果を表示している．

3. 評価実験

従来の CookChat (以降，絞り込み型と記載) と試作した連想型レシピ提案対話 (以降，連想型と記載) のユーザ体験を比較するため，操作した上での主観評価を行った．

クラウドソーシングで募集した 20〜55 歳の男女 100 名を，絞り込み型を操作する評価者 50 人 (評価者群 A)，連想型を操作する評価者 50 人 (評価者群 B) とした．評価者は，絞り込み型，連想型をそれぞれ 3 回づつ操作し，各操作では異なる料理をイメージしてレシピの嗜好に関する発話を入力する．さらに，操作の後に対話システムの使用感に関するアンケートに回答する．

表 1 にアンケートの項目を示す．項目のカテゴリは 7 種類あるが，各カテゴリに対して対話システムの機能性を問う観点 (機能性評価項目) と対話システムの自然性を問う項目 (自然性評価項目) に分けて質問した．評価者は各項目に関して，

- 全くそう思わない → 1
- あまりそう思わない → 2
- どちらとも言えない → 3
- そう思う → 4
- かなりそう思う → 5

*4　ボタン表示はユーザが選択した後，選択したボタン表示の吹き出しに変わる．

表 1: 評価者へのアンケート項目

	機能性評価項目	自然性評価項目
エージェントの発話の質	Q1 分かりやすい応答ができていると思いましたか?	Q2 人間らしい応答ができていると思いましたか?
対話の円滑さ	Q3 やりとりが円滑に行えたと思いましたか?	Q4 思い通り話せましたか?
話しやすさ	Q5 何を発話すればよいか，すぐに分かりましたか?	Q6 どう会話を始めたらよいか躊躇しましたか?
エージェントへの親近感	Q7 エージェントを違和感無く利用できると思いましたか?	Q8 エージェントと親しみを持って会話できると思いましたか?
対話の満足度	Q9 エージェントの提供する情報にあなたは満足しましたか?	Q10 エージェントとの会話にあなたは満足しましたか?
継続利用	Q11 またシステムを利用してみたいですか?	Q12 またエージェントとお話してみたいですか?
理解性能	Q13 異なる表現の入力も正しく理解されると感じましたか?	Q14 いろんな言い回しをちゃんと理解できると感じましたか?

の 5 段階で回答する．

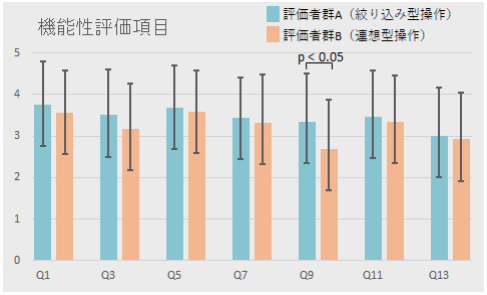


図 5: 機能性評価項目の比較

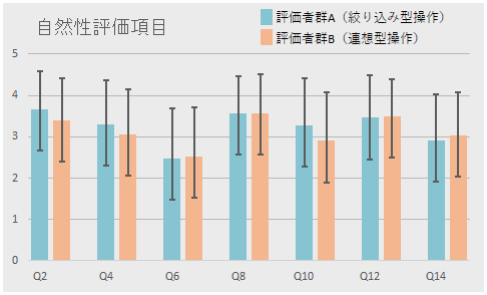


図 6: 自然性評価項目の比較

図 5 と図 6 に，絞り込み型を操作した評価者群 A の回答結果と連想型を操作した評価者群 B の回答結果 (1 ～ 5 のスコアの平均と標準偏差) の比較を示す．各質問項目に関して，左側のバーが評価者群 A の回答結果，右側のバーが評価者群 B の回答結果である．図 5 は機能性評価項目に関する比較，図 6 が自然性評価項目に関する比較である．

評価者群 A と評価者群 B の各質問項目に関する回答結果 (1~5) を t-検定したところ，Q9 のみ 5%水準有意で，評価者群 A より評価者群 B の方が回答結果のスコアが低くなった (平均値で 3.34 と 2.68)．他は有意差は無かった．

表 1 に示す質問項目以外のアンケートの設問として，対話システムを使用した感想を問う自由回答記述も設定した．表 2 に連想型の評価者の自由回答の一部を示す．Q9 の質問項目のスコアが低い評価者の自由回答には「イメージしていたレシピとはずれたレシピが提案された」という内容の回答が多かった．

ポジティブな回答からは，それほど具体的なレシピのイメージを持っていなかったユーザに対しては，ユーザ自身が自覚していなかった嗜好について探り出すという我々が狙った連想型の対話の効果があったことが分かる．一方で，ネガティブな回答からは，具体的なレシピのイメージを持っていたユーザは，自身の発話で明確にレシピが絞り込まれることを期待しており，

表 2: 連想型の評価者の自由回答 (一部)

ポジティブな回答
・とくに食べたいものが決まらないときに連想していって，決めるのもいいなと思った．
・頭にある雰囲気をさっちしてレシピを選んでくれるのが楽しい．
・何を作るか困った時に役に立ちそうなのでこのサービスを使いたいと思った．
・曖昧な言葉でも反応してくれるし，ちゃんと好みを具現化してくれるので使いやすいと思いました．
・お餅グラタンなど，まったく予想していなかったレシピが表示されて楽しい．
ネガティブな回答
・どうしてもこちらの意図とは異なる単語をピックアップされるようだ．
・最初に入力した言葉に対するシステムの返事が，想像力を間違った方向に使い選択の幅を違う方向に広げて行っているように感じた．
・連想を繰り返しているうちに，最初のワードとは全く関係のないレシピに落ち着いてしまう．
・自分の予想とはかけ離れた方向の提案が出てきてしまうので，戸惑いました．
・会話があまり成り立っていなかったと思った．一番最初に入力した材料や料理のジャンルの中から見つけてほしかった．

連想型の対話で我々が狙った効果を感じなかったと思われる．ユーザが料理の嗜好として発話する表現を，どの程度強い絞り込み条件として扱うかの制御が必要であることが分かった．

4. まとめと今後の課題

食知識に基づく連想型レシピ提案対話システムを試作しユーザの主観評価を行った．主観評価の結果は，従来の絞り込み型のレシピ提案対話に対して連想型のレシピ提案対話が好評価とはならなかったが，評価者の一部は，我々が狙った効果を感じたことが確認できた．今後，ユーザが料理の嗜好を発話する際の「絞り込み度合い」の意図を推測し，嗜好の重みの強さに反映させるなどのアプローチを検討する予定である．

参考文献

[1] 山上 勝義, 遠藤 充, 牛尾 貴志, 堀井 則彰.: “対話サービスプラットフォームの開発”, 人工知能学会全国大会 (第 31 回) 論文集, (2017)

[2] 遠藤 充, 牛尾 貴志, 山上 勝義, 堀井 則彰.: “発話とコマンドの系列を制御する対話エージェント”, 人工知能学会全国大会 (第 31 回) 論文集, (2017)

[3] 清丸 寛一, 黒橋 禎夫, 遠藤 充, 山上 勝義.: “料理レシピとクラウドソーシングに基づく基本料理知識ベースの構築”, 言語処理学会第 24 回年次大会 (NLP2018), (2018)

[4] 翠 輝久, 大竹 清敬, 堀 智織, 柏岡 秀紀.: “意思決定を支援する音声対話システム”, 言語処理学会第 18 回年次大会 (NLP2012), pp. 658-661(2012)

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

対話行為における説得による作業意欲の分析

Analysis of work motivation by persuasion in dialogue

水上　卓哉 *1
Takuya Mizukami

*1早稲田大学大学院創造理工学研究科経営システム工学専攻
Graduate School of Creative Science and Engineering, Waseda University

In this research, I conducted "a persuasive game" where a worker convinces the character about the importance of the work and analyzed the increase in motivation levels of the worker. When the worker fails to persuade, the worker confirms the importance of the work, so the motivation for work will be improved. In this experiment, I prepared two types of systems : one with a character that accept any persuation(conventional system) and one that may or may not accept a persuation , incorporating the success or failure of a persuation based on its cntent (proposed system), compare and evaluate each went. As a result, it was shown that the system that incorporates the success or failure of persuasion tends to significantly improve the motivation before work, compared with the conventional system.

1. 研究背景と目的

作業意欲が起こらない単純かつ退屈な作業は未だに存在している．そのため，そのような作業への意欲向上を目的とした研究が行われている．そのうちの 1 つに作業前に作業者がキャラクタに作業の重要性を説得させることで，作業意欲が向上することを示した研究がある [1]．この研究では，説得行為を行うことに重点を置いており，作業者の説得内容に関しては考慮していない．そのため，作業の重要性を理解していない作業者は意欲が向上しない問題がある．そこで本研究では，作業者がキャラクタ説得した際，説得の成否を取り入れることを提案し，作業前の作業意欲が向上に効果があるかを分析する．

2. 関連研究

Festinger ら [2] は自分の作業を本音と関わらず，あえて他人に「この作業は面白いものだ」と説明し納得させる被験者実験を行った．その結果，被験者が抱いていた作業に対するネガティブな印象がポジティブな印象に変容したことを明らかにした．この研究を参考に藤野ら [1] は，他人ではなくキャラクターエージェントに対して，作業の重要性を説得し，納得させるゲームシステムを考案した．その結果，説得対象がキャラクターエージェントであっても作業者の作業意欲が向上することを示した．しかし，十分に作業の重要性を理解しない人の場合は，キャラクタが常に説得される設定のため，効果的でないこともわかった．そこで本研究では，キャラクタが説得を受け入れないという設定を取り入れたシステムを提案する．

3. 提案

本研究では，藤野らのシステム [1] に対して， 説得の成否を取り入れたものを提案する．従来の研究では，どのような内容の説得に対しても，キャラクタは常に納得する設定している一方で，本研究では，あらかじめ決められた答えの文章に近い内容でなければ，キャラクタは説得を受け入れない設定をする．仮説としては，キャラクタにへの説得に失敗した作業者は，説得成功のために作業の目的や重要性をより深く考察するため，作業意欲が向上すると考えられる．

3.1 システム概要

本提案である説得成否の方法について述べる．作業者は作業の重要性を文章として入力する．入力された文章は Language Understanding Intelligent Service(以下 LUIS)[3] と呼ばれる自然言語処理エンジンによって Correct，Nearly，None のいずれかに分類される．この分類結果に応じて，キャラクタの作業者に対する応答文章を生成する．文章生成には C#を用い，そこでは，分類結果を受け取り，その結果ごとに決められている文章の返答を行う．説得システムの構成を図 1 に示す．

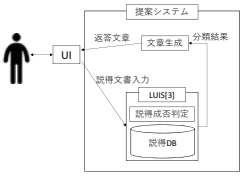


図 1: システムの構成

分類結果ごとの文章生成の中身を以下に示す．

```
CASE 分類結果 OF
 Correct：Display
  "納得しました。ありがとうございます"
  "この調査は隔年に行われている、文部科学省の調査でとても重要だそうです。"
  "作業がんばってください！"
 Nearly：Display
  "惜しいです。"
  "ヒント：なにかの資料を得ることを一番の目的としています"
 None or null:Display
  "こんにちは、私は説得キャラクターのボットです"
  "私を説得できたら勝ち、出来なかったら負けのゲームです"
  "あなたの作業はどのような目的がありますか？"
ENDCASE
```

連絡先: 水上卓哉，早稲田大学大学院創造理工学研究科経営システム工学専攻，〒 169-8555 東京都新宿区大久保 3-4-1　51 号館 15 階 02 号室，090-7847-6874，mizumizu-35@fuji.waseda.jp

4. 実験概要

実験は従来手法 [1] および提案手法に各 4 名，計 8 名で行う．今回被験者に課した作業は，紙に印刷された数値データを，表計算ソフトに入力する作業である．実験手順は以下に示す．

1. 実験説明
   被験者に作業の内容，および実験シナリオの説明を行った．実験シナリオは 2 種類用意した．ただし，実験シナリオ，そこで用いられている数値データは本研究のために架空に作成されたものである．

2. 作業 1
   作業 1 では作業とアンケートを行う．アンケートは作業前（1-1），作業中（1-2），作業後（1-3）の 3 回実施し，それぞれ作業アンケートと感情評価の 2 種類取得する．作業アンケートは，藤野らの研究のものを用いた [1]．その内容の一部を表 1 に示す．このアンケートでは計 11 項目を被験者は各項目 7 段階尺度で回答を行う．そして，合計得点をその被験者の作業意欲とする．感情評価では寺島ら [4] の多面的感情状態尺度短縮版（MMS）を用いて，ポジティブ，ネガティブ感情を 1〜4 段階で計測を行う．

3. 説得ゲーム
   被験者は作業 2 の重要性や目的をキャラクタに説明する．従来手法ではどのような内容の説得でも成功とし，提案手法では説得に対して成否の判定を行い，説得が成功するまで，説得を行う．

4. 作業 2
   被験者は作業 1 と同様に 2 回目の作業を行い，作業 1 の時と同様に，作業前（2-1），作業中（2-2），作業後（2-3）にアンケートを実施する．

5. 事後アンケート
   最後に実験に関して感想を自由記述で取得した．

表 1: 作業アンケートの例 [1]

Q1	この仕事に従事して、あなたは社会のための重要な人物だと思いますか？
Q4	この仕事は楽しいですか？
Q9	あなたはこの仕事に興味がありますか？
Q11	この仕事に従事して、あなたはイライラしますか？☆

(☆：Reverce Question)

5. 結果と考察

提案手法において，どの被験者も 1 回目の説得では失敗をしている．つまり，提案手法での被験者は従来手法の被験者よりも説得の回数が多い．そこで，説得ゲームによる作業意欲への影響を分析するため，各被験者の 1-1 での値を基準値とし，作業 2 での作業意欲との比較を行う．その結果を図 2 に示す．この結果から，作業 2 の作業前の意欲に関して，従来手法では基準値と比較して約 10 %に減少しているのに対し，提案手法では基準値と比較して，約 6 %上昇している．つまり，提案手法は従来手法と比較して，作業前の意欲が向上していると考えられる．そこで，基準値に対する 2-1 での値に関して，従来手法，提案手法において，各被験者の値を F 検定，片側 t 検定を行い，その結果を表 2 に示す．$p < 0.10$ となり作業前の

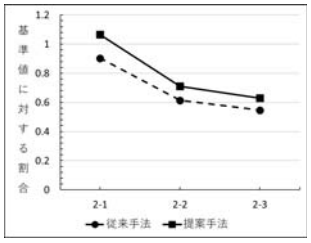


図 2: 2 回目の作業意欲の変化 (従来 n=4，提案 n=4)

表 2: 作業意欲の分析

実験手法	平均	分散	f	t
提案手法 (n=4)	1.064	0.014	1.213	-1.821
従来手法 (n=4)	0.903	0.017		

作業意欲に関して，従来手法と提案手法との差には有意傾向がある．つまり，説得に失敗した被験者はそうでない被験者と比較して作業意欲が向上しやすい傾向がある．

次に MMS[4] の結果から作業中の感情変化を分析した．その結果，作業中の集中の感情において違いが見られた．そこで，1-2 での値と 2-2 での値との比較を行う．得られた値を F 検定および両側 t 検定を行い，その結果表 3 に示す．

表 3: 作業中での集中感情の分析

実験手法	平均	分散	f	t
提案手法 (n=4)	0.945	0.013	1.926	-2.834*
従来手法 (n=4)	0.673	0.243		

5 %水準で集中感情に関して，従来手法と提案手法に有意な差がある．つまり，従来手法では約 33 %と減少しているのに対し，提案手法ではほぼ変化がないという違いであり，これは説得回数の差が作業中の集中感情に影響を与えたといえる．

事後アンケートでは，提案手法において説得失敗によって作業の重要性を確認するという意見が得られた．一方で，「説得をしているというよりもゲームの正解を考えているように感じた」という意見も得られた．つまり，作業者にとって説得文章を考えることが，ただゲームの答えを考えるためのものに変わる可能性が確認された．

6. まとめと今後の課題

本研究では，従来の説得ゲームのキャラクタが作業者のどのような説得も受け入れるという設定に対して，作業者の説得文章が答えの文章に類似するものでなければ，キャラクタは説得を受け入れない設定を取り入れた．その結果，従来手法と比較を行い，作業前の意欲の向上に有意傾向があった．今後の課題として，作業者にとって用意された答えが納得感が得られない可能性があるため，成否の判定の種類を増やし納得感が得られる仕組みが必要だと考えられる．

参考文献

[1] 藤野　秀則，石井　裕剛，下田　宏，吉川　榮和：アフェクティブインターフェースを用いた作業意欲向上の開発と評価，電気学会論文誌，Vol.128，No.1，pp.11-23(2008)．

[2] L. Festinger and J.M. Carlsmith： Cognitive consequences of forced compliance， *Journal of Abnormal Psychology*，Vol.58，No.2，pp203-210(1959)．

[3] Microsoft：LUIS:Homepage，
https://www.luis.ai/home，最終閲覧日 2017/12/18

[4] 寺島　正治，岸本　陽一，古賀　愛人：多面的感情尺度の作成，心理学研究，Vol.62，No.6，pp.350-356(1992)．

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

状態空間モデルを用いたWikipediaからの雑談対話システム用発話文の獲得

Acquiring Utterance Candidates for Chat-oriented Dialogue Systems from Wikipedia Based on a State Space Model

松本 知己 *1　酒井 正人 *1　中野 幹生 *2
Tomoki MATSUMOTO　Masato SAKAI　Mikio NAKANO

*1株式会社 Nextremer　*2株式会社 ホンダ・リサーチ・インスティチュート・ジャパン
Nextremer Co., Ltd.　Honda Research Institute Japan Co., Ltd.

This paper presents a method for acquiring utterance candidates for chat-oriented dialogue systems in specific domains from Wikipedia articles. A previously proposed method extracts sentences from Wikipedia articles related to the domain, converts them to utterances suitable for chat, and excludes sentences inappropriate as system utterances using a classifier. The classifier uses position information of the sentence in the article, but it is based on a simple logistic regression. On the contrary, our proposed method uses a state space model in a Bayesian approach. Experimental results show the proposed method outperforms the previous method.

1. はじめに

我々は, 特定のドメインの雑談を行う対話システムの構築を目指している. 雑談対話システムの発話生成として, seq2seqのようなニューラルネットワークにより発話文を生成する方法 [Li 16] と, あらかじめ用意された発話文集合から文脈に合うものを選択する手法 [稲葉 17] がある. 後者は自然な文が生成されるというメリットがあるが, あらかじめ人手で大量の文を用意するのは多大なコストがかかる.

そこで, 自動的に対話システムの発話候補文を獲得する手法として, Wikipedia[*1] を利用した方法が提案されている [太田 09, 杉本 17]. Wikipedia の記事は誰でも編集をすることが可能であるため情報の信頼性が心配されるが, 一定の信頼性があることが分かっている [日下 12]. また Wikipedia の各記事は, 記事ごとに差はあるが 1 つの記事にあるトピックについての情報が網羅されており, 記事の数も膨大であるため, 対話システムに用いる対話候補文の獲得源として注目されている. 本稿では, Wikipedia を対話候補文の獲得源として注目した [杉本 17] の研究を発展させ, 雑談対話システムのための新たな対話候補文の獲得方法を提案する.

本稿の構成は以下の通りである. 第 2 章では, 本研究に関連する研究を紹介する. 第 3 章では, [杉本 17] の方法を概説した後, [杉本 17] のモデルを拡張したモデルについて説明し, 第 4 章で提案したモデルの評価をする. 第 5 章では考察を述べ, 第 6 章では本稿のまとめと, 今後の課題を述べる.

2. 関連研究

[太田 09] は, 対話が盛り上がるような発話を行う雑談対話システムへの応用を目的に, Wikipedia から発話候補文の抽出を行う方法を提案している. [太田 09] は珍しい単語や単語の組み合わせで構成される文は人間にとって意外性があり, 対話を盛り上げるような内容であると仮定し, TF-IDF, 単語の共起性, 文の長さを用いて計算されるスコアで文をランク付けする方法を提案し, そのスコアの値が高い文ほど, 対話を盛り上げる内容の文であるとしている. しかし, この方法では一部の特徴的な文だけが抽出されるため, これだけでは雑談対話システムの実現は難しい.

また, [杉本 17] は雑談対話システムのために Wikipedia の記事から自動的に文を抽出し, 抽出した文を対話システムの出力として相応しいかという点でフィルタリング, 及び口語体へ変換する方法を提案している. そして, こういった変換が施された文が, 実際に対話システムの文として使用可能かどうか判別するモデルとして, 文を抽出する際に付随する記事上における文の位置情報を特徴量とした, ロジスティック回帰判別モデルを提案している. しかし, [杉本 17] の判別モデルは, 文の位置情報を十分に活用できていなかった.

本稿では, これら文の位置情報をベイジアンモデリングの枠組みに組み込むことで, [杉本 17] が提案したロジスティック回帰判別モデルを拡張する.

3. 提案手法

本章では, [杉本 17] が提案した, Wikipedia から発話候補文を獲得する方法について概説した後, 獲得した文に付随する情報を用いた, 新たな文使用可否判別モデルを提案する.

3.1 候補文の獲得

[杉本 17] が提案した, 発話候補文を獲得する方法について概説する. まず, Wikipedia 記事の HTML の<p>タグ内に属する文章を抽出し, 句点 (。) 区切りで文章を文に分割する.

次に, 抽出した各文に対して, フィルタリングを行い, さらに口語体変換を行う. フィルタリングは, 一文中に含まれる文字数や, 一文で意味を理解できない文 (文脈依存の文), 対話システムへの使用が不適切と判断される表現を含む文などを排除することである. そして, フィルタリングによって排除されなかった文は口語体に変換される. この時点で, 例えば, Wikipedia の「お好み焼き」トピックからは

お好み焼きは, 小麦粉とキャベツなどを使用する鉄板焼きの一種で, 日本の庶民的な料理なんですよ *2

連絡先: 松本 知己, 株式会社 Nextremer, 東京都板橋区成増 1-30-13 トーセイ三井生命ビル 10F, 03-6915-6447, tomoki.matsumoto@nextremer.com

*1 日本語版 Wikipedia: http://ja.wikipedia.org/

*2 本稿では, クリエイティブ・コモンズ・表示・継承ライセンス 3.0 (http://creativecommons.org/licenses/by-sa/3.0/) のもとで公表されたウィキペディア (https://ja.wikipedia.org/) の「お好み焼き」及び「ウイスキー」の項目を二次利用した. (参照日 2018 年

といった文が抽出される.

こうして抽出された文が対話システムに使用可能であるかどうかを判定する. ここで, 対話システムに使用可能かどうかの基準は, (1) 日本語として自然で, (2) その文だけで内容が理解でき, (3) その記事のトピックに関して述べている, である. この判別には, 記事から文を抽出した際に得られる, 以下の特徴量を用いたロジスティック回帰判別を用いる.[*3][*4]

- f_1: header info（<h1>,<h2>,<h3>,<h4>）
- f_2: header id（上から何番目のヘッダに所属するか）
- f_3: paragraph id（ヘッダ中の何番目の段落に登場するか）
- f_4: sentence id（所属する段落の何文目に登場するか）

3.2 モデル

本節では, [杉本 17] のロジスティック回帰を状態空間モデルへ拡張した, 階層文状態空間モデルを提案する.

[杉本 17] は, ロジスティック回帰モデルの特徴量として記事における文の位置情報を使用したが, 単に説明変数として使用しただけであったため, 十分に位置情報をモデルに組み込めてはいなかった. そこで [杉本 17] のモデルを, 以下を考慮したベイズモデルへ拡張する:

(i) それぞれの記事は, 固有の特徴がある.
(ii) それぞれの記事は, ヘッダーによる階層構造を持っている.
(iii) 段落内において, ある文が使用可である確率が, その次の文の使用可否に影響を与える.
(iv) 段落内において, 文の使用可否の状態は連続的に変化する.

仮定 (i) は, それぞれの記事は現時点で手に入っているデータからは説明することができない, 固有の特徴を持っていることを仮定している. これは, 生態学において生物や植物の個体を解析する際に個体差と呼ばれているもので, 記事差と呼ぶこととする. 仮定 (ii) は, それぞれの記事はヘッダーごとに異なる傾向を示すであろうということを仮定している. 例えば, ヘッダー 1 内の文は使用可となりやすく, ヘッダー 4 内の文は使用不可となりやすい, といった異なる傾向を持つことを仮定していることになる. 仮定 (iii) は, 例えば, ある文が使用可になる確率が高ければ, その次に続く文も使用可になりやすいといった傾向が存在することを仮定している. 今回は, ある文に影響を与える文は, その文の 1 つ前の文, 2 つ前の文の 2 文とした. 仮定 (iv) は, 段落内における文の先頭からの順番を時系列のインデックスとみなしたときに, 段落内で文はある種の“状態”を持っており, その状態は前の文から次の文へ伝搬していくということを仮定している. このことを表現するために, 今回はある文の状態とその 1 つ前の文の状態は似ているという 1 階差分の制約を課した.

以上をまとめるとモデルは以下となる. 以下の式で, $a = 1, \ldots, A$ は記事番号, $h = 1, \ldots, 4$ はヘッダー番号, $b = 1, \ldots, B_a$ はパラグラフ番号, $s = 1, \ldots, S_{a,b}$ は文番号, $t = 1, 2$ はそれぞれ 1 つ前, 2 つ前の文を表す遅延番号, $i = 2, 3$ は特徴量番号で, それぞれ 3.1 節で定義した特徴量 f_2, f_3 を表すインデックスとする. インデックス h には特徴量 f_1 を, インデックス s には特徴量 f_4 を用いることができる. また, $y_{a,h,b,s}$ は [杉本 17] の方法によって抽出された文のアノテーション結果であり, 教師データとなる. つまり, 抽出された文が使用可であれば $y_{a,h,b,s} = 1$, 使用不可であれば $y_{a,h,b,s} = 0$ となる.

$$y_{a,h,b,s} \sim Bernoulli(p_{a,h,b,s}) \tag{1}$$

$$p_{a,h,b,1} = \mathrm{logit}^{-1}(\gamma_a + \alpha_1 + \beta_{h,0} + \sum_{i=2}^{3} \beta_{h,i} f_{a,h,b,1,i}) \tag{2}$$

$$p_{a,h,b,2} = \mathrm{logit}^{-1}(\gamma_a + \alpha_2 + \eta_{h,1} p_{a,h,b,1} + \beta_{h,0} + \sum_{i=2}^{3} \beta_{h,i} f_{a,h,b,2,i}) \tag{3}$$

$$p_{a,h,b,s} = \mathrm{logit}^{-1}(\gamma_a + \alpha_s + \sum_{t=1}^{2} \eta_{h,t} p_{a,h,b,s-t} + \beta_{h,0} + \sum_{i=2}^{3} \beta_{h,i} f_{a,h,b,s,i}) \tag{4}$$

$$\gamma_a \sim Normal(\theta_\gamma, \sigma_\gamma^2) \tag{5}$$

$$\alpha_s \sim Normal(\alpha_{s-1}, \sigma_\alpha^2),\ s = 2, \ldots, S_{a,b} \tag{6}$$

$$\beta_{h,0} \sim Normal(\theta_{\beta_0}, \sigma_{\beta_0}^2) \tag{7}$$

$$\beta_{h,i} \sim Normal(\theta_{\beta_i}, \sigma_{\beta_i}^2) \tag{8}$$

$$\eta_{h,t} \sim Normal(\theta_{\eta_t}, \sigma_{\eta_t}^2) \tag{9}$$

$$\theta_\gamma, \theta_{\beta_0}, \theta_{\beta_i}, \theta_{\eta_t} \sim Normal(\mu, \rho^2) \tag{10}$$

$$\alpha_1, \mu \sim Uniform(-\infty, \infty) \tag{11}$$

$$\sigma_\gamma, \sigma_\alpha, \sigma_{\beta_0}, \sigma_{\beta_i}, \sigma_{\eta_t}, \rho \sim Uniform(0, \infty) \tag{12}$$

logit^{-1} はロジスティック関数, $Bernoulli$ はベルヌーイ分布の確率質量関数, $Normal$ は正規分布の確率密度関数, $Uniform$ は一様分布の確率密度関数を表している. $Uniform(-\infty, \infty)$ 及び $Uniform(0, \infty)$ は無情報事前分布を表しており, 事前情報を仮定しないことを意味している.

式 (5) は仮定 (i) を表している. それぞれの記事 a についての固有の特徴 γ_a の値は, 平均値としては記事全体に共通の θ_γ であり, その平均の周りを共通の標準偏差 σ_γ で正規分布した結果, 記事ごとに異なる特徴 γ_a が現れるというモデリングである. したがって, 標準偏差 σ_γ は記事差がどれほど存在するかの指標となり, 小さな値を取れば, それぞれの記事の固有の特徴は平均周りに集中的に分布することになるため, あまり記事差が無いことを示唆し, 逆に大きな値を取れば, それぞれの記事の特徴は平均から離れた値を取りやすくなるため, 記事差があることを示唆する.

式 (7), (8), (9) は仮定 (ii) を表している. ヘッダー h ごとのロジスティック回帰における切片項 $\beta_{h,0}$, 特徴量 f_i $(i = 2, 3)$ に対する係数 $\beta_{h,i}$ と, $t(= 1, 2)$ つ前の文が使用可である確率に対する係数 $\eta_{h,t}$ は, それぞれが各ヘッダーに共通の平均 $\theta_{\beta_0}, \theta_{\beta_i}, \theta_{\eta_t}$ の周りを標準偏差 $\sigma_{\beta_0}, \sigma_{\beta_i}, \sigma_{\eta_t}$ で正規分布した結果, ヘッダーごとに異なった値として現れるというモデリングである. したがって記事差の場合と同様に, これら標準偏差はヘッダーごとに異なる傾向を示すことの指標となる.

式 (3), (4) は仮定 (iii) を表現している. ロジスティック回帰モデルにおける確率を計算するロジスティック関数 logit^{-1} の中に, 前の文が使用可になる確率とその係数が掛け合わされた形で入っている.

式 (6) は仮定 (iv) を表現している. ある文の状態 α_s とその前の文の状態 α_{s-1} は“似ている”という, 以下の 1 階差分の

03 月 08 日）

*3 フィルタリングや口語体変換, 文が対話システムで使用可能かどうかの判別の訓練データのアノテーションの方法の詳細は [杉本 17] を参照.

*4 [杉本 17] の検証では, f_1, f_2, f_4 を説明変数として使用したロジスティック回帰判別モデルの精度が最も高かった.

	(f_2)	(f_3)	(f_2, f_3)	[杉本 17]
WAIC	0.572	0.576	0.571	-
精度	-	-	0.682	0.644

表 1: 提案モデルの WAIC と精度, 及び [杉本 17] のモデルの精度. (f_2), (f_3), (f_2, f_3) はそれぞれ提案モデル説明変数として f_2 だけを用いた場合, f_3 だけを用いた場合, f_2 と f_3 をともに用いた場合を表している.

制約をそれぞれの文の状態に課すことで, 文の使用可否の状態が連続的に変化することを表現している.

$$\alpha_s - \alpha_{s-1} \approx 0 \iff \alpha_s \approx \alpha_{s-1}$$

このことを, ある文の状態 α_s はその前の文の状態 α_{s-1} を中心に標準偏差 σ_α で正規分布するという形で表現したものが式 (6) であり, これは一般的に 1 階差分ガウス状態空間モデルと呼ばれている.

最後に, θ と下付き文字で表されるすべての正規分布の平均パラメータはすべて共通の平均 μ, 標準偏差 ρ の正規分布に従うとし, 段落内における先頭の文の状態 α_1, μ, σ と下付き文字で表されるすべての正規分布の標準偏差パラメータ, そして ρ はすべて無情報事前分布に従うとする.

4. 評価

本章では, 提案手法のモデル選択と精度評価を行う. 提案モデルには, 説明変数として f_2 だけを用いる場合, f_3 だけを用いる場合, f_2 と f_3 をともに用いる場合の 3 パターンの自由度があるため, 変数選択の指標として WAIC (Widely Applicable Bayesian Information Criterion) を用いる [Watanabe 13]. WAIC は事後分布が正規分布で近似できない場合にも使用できる情報量基準であり, AIC(赤池情報量基準) を一般化したものである.

精度評価には 5-fold の交差検証を実施した. また, Wikipedia から抽出されたデータは [杉本 17] で使用されたものを使用した. [杉本 17] は料理分野に関する 60 記事から候補文を抽出していたため, その 60 記事を 5 分割することで訓練データ 48 記事, 検証データ 12 記事のペアが 5 つできる. これらを用いて交差検証を行い, 5 回の検証で得られた精度の平均値を評価指標とする. モデルの推定には, Stan を R 言語上で実行することができる, RStan を使用した [Stan Development Team 17]. Stan は, 事後分布に従うパラメータのサンプルを得ることができるサンプラーである.

検証結果は表 1 のようになった. WAIC によるモデル選択の結果, WAIC の値が最も小さい, 説明変数として f_2 と f_3 をともに用いるモデルが選択された. また, 選択されたモデルと [杉本 17] のモデルの精度を比較すると, 我々が今回提案するモデルの方が [杉本 17] のモデルよりも高精度に文の使用可否を判別できていることがわかる.

5. 考察

3.2 節で説明したように, 提案したモデルには 4 つの仮定を置いていた. 本章では, それぞれの仮定についての妥当性を確認する.

表 2 は, 3.2 節で提案したモデルの各正規分布の標準偏差パラメータの推定結果である. σ_γ は, 各記事は特有の効果を持っているという仮定 (i) についての妥当性を表していると考えられる. 各記事の特徴は, 平均パラメータ θ_γ の値が定まると, その θ_γ を中心に平均的に標準偏差 0.69 でばらつく. 例えば, 他の変数の値が同じで記事差だけが 0.69 異なる記事同士を比較したとすると, 使用可の確率は $\exp(0.69) = 1.993$ と, 約 2 倍異なってくる. これは記事差が存在することを表していると言える.

	mean	sd	2.5%	25%	50%	75%	97.5%
σ_γ	0.69	0.10	0.51	0.62	0.68	0.75	0.90
σ_0	2.87	3.94	0.77	1.35	1.97	3.11	10.2
σ_{η_1}	0.83	0.81	0.06	0.40	0.64	1.03	2.72
σ_{η_2}	0.49	0.73	0.02	0.16	0.34	0.62	1.86

表 2: モデルの各正規分布の標準偏差パラメータの推定結果. mean は平均値, sd は標準偏差を表す.

	mean	sd	2.5%	25%	50%	75%	97.5%
$\eta_{1,1}$	-0.02	0.53	-1.13	-0.34	0.00	0.31	0.98
$\eta_{2,1}$	0.54	0.27	0.03	0.35	0.54	0.72	1.07
$\eta_{3,1}$	0.39	0.31	-0.18	0.18	0.38	0.60	1.03
$\eta_{4,1}$	-0.35	0.57	-1.56	-0.72	-0.30	0.05	0.64
$\eta_{1,2}$	0.04	0.43	-0.82	-0.19	0.01	0.24	1.00
$\eta_{2,2}$	0.12	0.28	-0.40	-0.07	0.10	0.29	0.74
$\eta_{3,2}$	-0.13	0.33	-0.88	-0.31	-0.10	0.07	0.48
$\eta_{4,2}$	0.10	0.47	-0.76	-0.15	0.05	0.32	1.25

表 3: 1 つ前, 2 つ前の文が使用可となる確率に対する, ヘッダーごとの係数の推定結果. mean は平均値, sd は標準偏差を表す.

σ_0 は, それぞれの記事はヘッダーによる階層構造を持っているという仮定 (ii) の妥当性を表していると考えられる. σ_γ の場合と同様の議論で, ヘッダーごとに異なる特徴を持っていることが言える. σ_{η_1}, σ_{η_2} はそれぞれ 1 つ前, 2 つ前の文が使用可となる確率に対する, ヘッダーごとの係数 $\eta_{h,1}$, $\eta_{h,2}$ の標準偏差である. 1 つ前, 2 つ前の文が使用可となる確率が当該文に与える影響がヘッダーごとに異なることを表しており, このことからもヘッダーごとに異なる特徴を持っていることが言える.

表 3 は, 1 つ前, 2 つ前の文が使用可となる確率に対する, ヘッダーごとの係数 $\eta_{h,1}$, $\eta_{h,2}$ の推定結果である. 特に 1 つ前の文が使用可となる確率に対する係数 $\eta_{h,1}$ について, ヘッダー 1 は 1 つ前の文が使用可となる確率に影響をあまり受けず, 逆にヘッダー 2 と 3 に属する文は正の影響を, ヘッダー 4 に属する文は負の影響を受ける傾向があることが分かる. このことからも仮定 (ii) は妥当なものであると考えられ, また前の文が使用可である確率が次の文が使用可になる確率に影響を与えるという仮定 (iii) は妥当なものであると考えられる.

図 1 は, 推定された段落内における文の状態の推定結果を, 横軸が段落内における先頭からの文の順番, 縦軸が推定された文の状態として図示したものである. 段落の先頭の文が最も使用可になりやすく, 段落の深い箇所にある文であるほど使用可になりにくいという構造が推定されている. このことから, 段落内において, 文の使用可否の状態は連続的に変化するという仮定 (iv) は妥当なものであると考えられる.

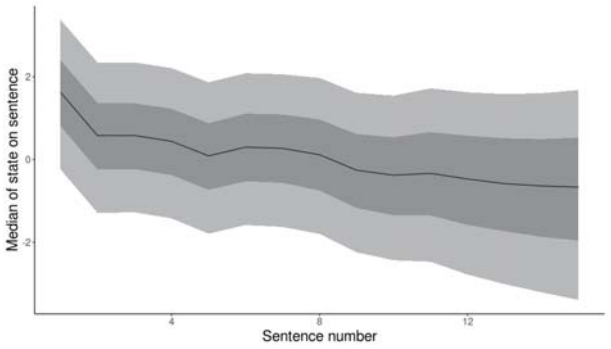


図 1: 文の状態推移．横軸が段落内における先頭からの文の順番, 縦軸が推定された文の状態．実線は 50%点, 濃い区間は 50%信用区間, 薄い区間は 80%信用区間を表している．

また，以下は使用可とアノテーションされた文のうち, [杉本 17] のモデルでは使用不可と判別されていたが, 今回我々が提案したモデルでは正確に使用可と判別できるようになった文の例である．

英国スコットランドで造られるウイスキーをスコッチ・ウイスキーまたは単にスコッチと呼ぶんですよ

この文は「ウイスキー」トピックの記事において, 上から 18 番目のヘッダー番号 4 のヘッダーに属していた．ヘッダー番号 4 の文は使用不可になりやすく, また記事の下の方にあるヘッダーに属する文も使用不可になりやすい傾向にあるため, [杉本 17] のモデルはこの文を使用不可と判断したと考えられる．しかし, この文は段落内における先頭の文であったため, 先頭の文はより使用可になりやすいという, 我々が提案したモデルにおける文の状態の推定結果によって, この文が正しく使用可であると判別できるようになったと考えられる．

以上から, 3.2 節で提案した仮定は妥当なものであると考えられ, こういった文の構造のモデル化の妥当性が, 第 4 章における文の使用可否の判別性能の向上に繋がったと考えられる．

6. おわりに

本稿では, [杉本 17] が提案した雑談対話システムのために Wikipedia から発話候補文を獲得する方法によって得た候補文が使用可能であるかを判別する, 階層文状態空間モデルを提案した．また, [杉本 17] が提案したロジスティック回帰モデルと精度の比較検証を行い, 提案モデルの方がより高精度に文の使用可否判別が行えることを確認した．

今後の課題として, 今回は単に説明変数として用いたヘッダーの順番 f_2 と段落の順番 f_3 のさらなる活用がある．今回のモデルではヘッダー情報を表す f_1 と段落内での文の順番を表す f_4 を, それぞれ階層化の対象と状態変化の対象としてモデルに組み込むことができた．しかし, ヘッダーの順番 f_2 と段落の順番 f_3 は単に説明変数として用いただけであったため, これらを適切にモデルの中に組み込む方法を考える必要がある．また, 現在は文の位置情報だけをモデルに用いており, 仮に同じ構造をもつ記事が複数あった場合, 記事の内容や文の意味に関わらず全く同じ使用可否判別を行うことになる．より高精度な判別のためには, 文の意味を組み込んだモデルが必要となる．

参考文献

[稲葉 17] 稲葉 通将，高橋 健一：非タスク指向型対話システムのためのリカレントニューラルネットワークを用いた発話候補ランキング, 電子情報通信学会論文誌 D, Vol.J100-D, No.6, pp. 661-671, 2017.

[稲葉 12] 稲葉 通将, 平井 尚樹, 鳥海 不二夫, 石井 健一郎：非タスク指向型対話エージェントのための統計的応答手法, 電子情報通信学会論文誌, vol.95, pp. 1390–1400. 2012.

[太田 09] 太田 知宏, 島海 不二夫, 石井 健一郎：発話生成を目的とした Wikipedia からの文抽出, 人工知能学会全国大会講演論文集, 2009.

[日下 12] 日下 九八：ウィキペディア：その信頼性と社会的役割, 情報管理, vol.55, pp. 2–12. 2012.

[杉本 17] 杉本 俊, 植木 拓, 林 宏幸, ニコルズ エリック, 中野 幹生：特定ドメイン雑談対話システムのための Wikipedia を用いた発話文の生成, 人工知能学会 言語・音声理解と対話処理研究会, vol.81, pp. 98-99. 2017.

[東中 14] 東中 竜一郎：雑談対話システムに向けた取り組み, 人工知能学会研究会資料 SIG-SLUD-70, pp. 65–70. 2014.

[樋口 11] 樋口 知之：予測にいかす統計モデリングの基本—ベイズ統計入門から応用まで, 講談社. 2011.

[Gelman 13] Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, Donald B. Rubin : Bayesian Data Analysis, Third Edition, CRC Press. 2013.

[Li 16] Jiwei Li, Michel Galley, Chris Brockett, Georgios PSpithourakis, Jianfeng Gao, Bill Dolan : A persona-based neural conversation model, In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL-16), pp. 994-1003. 2016.

[Stan Development Team 17] Stan Development Team : RStan: the R interface to Stan. R package version 2.16.2. 2017.

[Watanabe 13] Sumio Watanabe : A Widely Applicable Bayesian Information Criterion, Journal of Machine Learning Research, vol.14, pp. 867-897. 2013.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

未知の属性の問合せに応答可能な対話システムを目指した知識グラフの拡充

Knowledge Graph Expansion for Constructing Dialogue Systems Responding to Queries about Unknown Attributes

藤岡 勇真　　駒谷 和範
Yuma Fujioka　　Kazunori Komatani

大阪大学産業科学研究所
The Institute of Scientific and Industrial Research, Osaka University

Dialogue systems cannot respond to queries including unknown words. Constructing a perfect knowledge base is practically impossible; that is, filling all the attributes in relational databases is quite labor-intensive. We are trying to construct a system that responds reasonably by exploiting existing system knowledge even when a user's query contains unknown attributes. More specifically, we use knowledge graphs as the system's backend and infer unknown attributes of entities by generalizing them with collaborative filtering and tensor factorization. As a first step for it, we extracted a knowledge graph from Wikidata, a commonly-used LOD (linked open data), in the food and dish domain, but it was sparse for generalization. We expand the knowledge graph with external datasets, specifically recipe data contributed to a Web site, and evaluate it by a task to find similar words.

1. はじめに

対話システムは，未知の情報に関する問合せに上手く応答することが出来ない．対話システムに対するよくある入力例として，ある語に関する情報の問合せを考える．このような問合せが入力された場合，システムはその知識ベースから必要な情報を検索し，適切な応答文を生成する．しかし現状の対話システムは，要求された情報が知識ベースに存在しなかった場合，話題を変えたり「わかりません」といった，文脈上不自然な応答をしてしまう．また，必要な情報が全て網羅されている知識ベースを構築するのは非常に困難である．ゆえに，より良い応答の出来る対話システムを目指すうえで，知識ベースに無い情報の問合せへの対応は避けられない課題だと言える．

本研究では，質問された語の情報がシステムにとって未知だった場合でも合理的に応答出来るシステムの構築を目指している．図 1 に目標とするシステムの枠組みを示す.「ミネストローネってイタリアの料理なんですか？」と，ユーザからある語に関する問合せがシステムに入力される．この問合せには，入力語である「ミネストローネ」の「国」という属性に関する情報が必要になる．知識ベースにその情報がないことが判明すれば，システムは既知の情報を用いて，欠損した語の属性や類似した語の推測を試みる．図 1 では,「ミネストローネ」の「国」が「イタリア」であるかどうかを既知の情報から推測し，その結果から「ミネストローネはたぶんイタリアの料理だと思いますよ」という，断定はしないが可能性をほのめかせる応答を行っている．

対話システムにおける入力文中の未知語の処理に取り組んだ研究はこれまでにも行われている [大塚 13][大野 18]．これらの研究では，ユーザへの質問や発話内容の解析によって未知語の情報の獲得を試みているが，本研究では既に知識ベース上にある情報を利用して必要な情報の推測を目指す．

図 1 のような応答生成の実現に向けての第一歩として，システムがバックエンドで持つ知識ベースとして知識グラフを使用する．知識グラフとは，グラフ構造を持ったデータベースである．知識グラフはデータ間の様々な関連性を表現するのに長けた知識モデルであり [Angles 08]，近年こういったグラフ構造をもつデータベースが徐々に増えている．

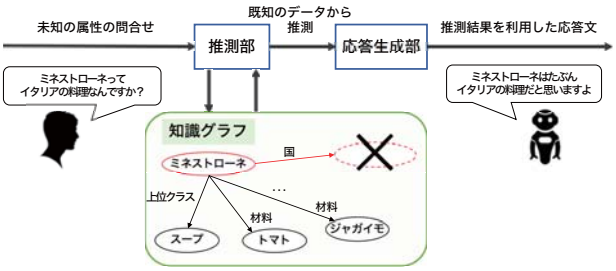


図 1: 目標システムの全体像

本稿では，システムが持つ知識グラフが料理の一般的な情報を持つとして，その構築を行う．まず大規模知識ベースである Wikidata[*1] からのデータ抽出によって知識グラフを構築した．しかし，構築した知識グラフ上には料理の一般的な材料情報が十分でなく，知識グラフがあまりにスパースだと推測や解析が困難になると考えられる．そこで，楽天公開データ [*2] 中の楽天レシピに関するデータを使って，料理の一般的な材料情報を知識グラフの拡充を行う．知識ベースの拡充手法には Web 上のテキストデータからの抽出などがあるが [高橋 07]，本研究では構造化されたレシピデータを用いて信頼性の高いデータの抽出を試みる．拡充の有効性は，入力語に類似した語を知識グラフ上から得るというタスクにより検証した．

2. システムが持つ知識グラフ

2.1 知識グラフ

目標システムは，料理に関する一般的な情報が格納された知識グラフを持つとしてその構築を行う．ここで知識グラフとは，データとデータの間にある関係がグラフ構造によって表現されたデータベースを指す．一般に，グラフ中のそれぞれのノードが 1 つのデータを表現し，ノード間に張られるエッジは

連絡先: 藤岡勇真，大阪大学産業科学研究所，大阪府茨木市美穂ヶ丘 8-1，06-6879-8416，fujioka@ei.sanken.osaka-u.ac.jp

∗1 https://www.wikidata.org
∗2 https://rit.rakuten.co.jp/data_release_ja/

その 2 つのデータ間に何らかの関係があることを示している．また，データ間の構造的な関係を表現するには向きを持つ有向エッジが用いられ，エッジそのものにラベルを付与し関係の種類，すなわちどんな属性であるかを表すこともある．

2.2 Wikidata を用いた知識グラフの構築

Wikidata は，Web の仕組みに基づきリンクされたオープンなデータである Linked Open Data として公開されている．Wikidata は大規模知識ベースの 1 つとして知られ，多言語対応，人手による管理体制などが主な特徴として挙げられる．Wikidata は，Resource Description Framework(RDF) 用のクエリ言語である SPARQL を用いたデータ検索が可能である．

RDF とは，概念や実体を抽象化したリソースの関係を，主語・述語・目的語という三つ組によって記述するデータモデルであり，この三つ組は一般的にトリプルと呼ばれている．RDF は，トリプルを有向エッジ，リソースをエンティティ，述語に記述されたプロパティを有向エッジのラベルとする，ラベル付き有向グラフとして表現することが出来る．SPARQL ではそのグラフパターンを指定することでデータ検索を実現しており，トリプルとしてデータを抽出することが出来る．

Wikidata からの料理に関する情報抽出の手順を示す．はじめに，目標システムが料理としてみなすエンティティの集合である料理群を設定した．具体的には，

1. 『料理，料理のサブクラス，type of food or dish』のインスタンス
2. 『料理，type of food or dish』のサブクラス

という 2 つの条件を 1 つでも満たすエンティティの集合である．

この条件を満たすエンティティを Wikidata から収集すると，その数は 3278 個であった (重複を除く)．

そして，取得した料理群の要素となるエンティティ(以下，料理エンティティと記述する) が主語もしくは目的語となっているトリプルを全て取得した．ここで取得した 110255 個のトリプルには，他のデータベース上における固有 ID や画像・発音音声へのリンクなどといった，エンティティ間の関連性を表さないプロパティを持つものが多く存在していた．そのため，前述のようなプロパティを除く 10 種のプロパティを持つ 11495 個のトリプルによって知識グラフを構築した．

3. 楽天公開データを用いた知識グラフの拡充

前節で構築した知識グラフは，一般的な料理の材料情報が不足していた．そこで，楽天公開データ中の料理レシピ検索サイト「楽天レシピ *3」に投稿された約 80 万件分のレシピデータから，料理の一般的な材料の情報を抽出して知識グラフに拡充する．

このレシピデータはレシピ名・料理名・カテゴリ・材料などの情報が各レシピごとにデータ化されたものである．同じ料理名が記述されたレシピでも，使っている材料はレシピ毎に異なる．またレシピに記載されている材料の中には，地域性や投稿者のアレンジなどに由来する，その料理の材料としては一般的でない材料が含まれている場合もある．ゆえに，料理名に結びつけられた使用材料をそのまま全て知識グラフに拡充することは，一般的な料理の材料情報の拡充という目的にそぐわない．そこで，ある料理で使われている材料 m の一般性が高いほど，その料理のレシピにおける材料 m の使用率が高いと仮定し，使用率に基づく拡充対象の選別を行う．

*3 https://recipe.rakuten.co.jp/

表 1: 楽天レシピデータにおける「オムライス」の材料使用率

使用された材料名	使用率 U_{rate} (%)
「卵」	79.9
「ケチャップ」	69.9
「ご飯」	44.7
「玉ねぎ」	38.4
……	……
「しいたけ」,「ほうれん草」	1.6
「豆乳」,「みりん」,「ソース」	1.5
「塩麹」,「とろけるチーズ」,「冷凍コーン」	1.3
……	……

「料理名が c でその材料が m である」という情報を知識グラフに拡充する手順は，大きく分けて次の 3 つのステップからなる．

1. 全てのレシピデータから，料理名と使用材料の組を抽出する．
2. 料理名 c のレシピにおける材料 m の使用率 $U_{rate}(c,m)$ を算出する．
3. 一定の条件を満たす c,m の組を知識グラフへの拡充候補とし，そうでない場合は拡充候補から棄却する．

まず，第 1 ステップである料理名と使用材料の組の抽出について述べる．料理名 c における材料 m の使用率を求めるために，各レシピデータから料理名と材料の組を抽出する．i 番目のレシピデータに記述されている料理名 $c^{(i)}$ と使われている材料からなる集合 $M^{(i)}$ を抜き出し，i 番目のレシピデータの料理名・材料の組を表す集合 $S^{(i)} = \{c^{(i)}, M^{(i)}\}$ を抽出する．この $S^{(i)}$ の作成を，$i = 1, 2, ..., 796067$ について行う (796067 はレシピ総数)．

第 2 ステップである，料理名 c のレシピにおける材料 m の使用率 $U_{rate}(c,m)$ の算出について述べる．料理名 c における材料 m の使用率 $U_{rate}(c,m)$ を式 (1) のように定義する．

$$U_{rate}(c,m) = \frac{|\{S^{(j)}|c^{(j)} = c, M^{(j)} \ni m\}|}{|\{S^{(j)}|c^{(j)} = c\}|} \qquad (1 < j < n) \tag{1}$$

式 (1) の分母は料理名が c であるレシピの総数を表している．そして分子は，料理名が c であるレシピの中で，材料 m が使用されているレシピの総数を表す．

使用率 $U_{rate}(c,m)$ の実例として表 1 に，c を「オムライス」とした時の，材料ごとの使用率の一部を示す．表 1 は「オムライス」の食材を使用率の昇順に並べたもので，使用率が高い食材と低い食材の一部を抜粋している．使用率が高い材料には「卵」,「ケチャップ」,「ご飯」といった,「オムライス」の一般的な材料が並んでいる．一方で，使用率が低い材料には,「しいたけ」,「豆乳」,「塩麹」等といった「オムライス」の材料としては一般的ではない材料が多く含まれている.「オムライス」の例は，使用率が高いほどその料理に一般に使われる材料であるという仮説を支持する結果になっている．

次に第 3 ステップである，材料情報の拡充の判定について述べる．使用率 $U_{rate}(c,m)$ に対するしきい値を θ とし，$U_{rate}(c,m) \geq \theta$ が成り立つ c,m の組を知識グラフへ拡充する．この判定で，使用率の低い材料が知識グラフに拡充されることを防ぐ．しきい値 θ の値は 4.2 節の調査結果より決定する．

例外として，多くの料理に材料として頻出した「水」,「油」,「砂糖」,「塩」の 4 つの材料は，料理の特徴的な材料としては不適当であるとし，これらの材料を含む c, m の組は拡充対象から除外している．また，Wikidata を用いて構築した知識グラフ上のエンティティの中に，c, m というラベルが付与されたエンティティが片一方でも存在していない場合でも，拡充対象から除外している．これは実装上設けた制限であり，構築した知識グラフ上に新たなエンティティを作成することは行っていない．あくまで Wikidata を用いて構築した知識グラフ上のエンティティ間に，材料の関係を結びつける形になる．

第 3 ステップで述べた条件を満たす全ての c, m の組について,「c の材料が m である」という事実を表すトリプルを作成し，知識グラフにトリプルを拡充する．以上の処理によって，楽天公開データから料理の材料情報をトリプル形式で抽出し，Wikidata と楽天公開データの料理に関する情報を統合する．

4. 知識グラフ上での類似語推測性能の調査

4.1 調査方法

知識グラフの拡充の効果を，入力語に類似した語を知識グラフから得るというタスクにより検証する．一般的な料理の材料情報が楽天公開データから拡充されていれば，その情報によって知識グラフ上の料理エンティティは正しく特徴付けられる．よって拡充を行わない場合に比べて，容易に知識グラフ上から類似語を得ることが出来ると考えている．反対に一般的でない料理の材料情報が拡充されていれば，料理エンティティが一般的でない情報で特徴付けられ，類似した語を得るのが困難になると考えている．

入力語に類似した語をシステムが見つけられているかどうかの指標として，類似度に基づくランキングにおける順位を用いる．この調査では，テストセット中の 30 種の料理を入力語と仮定している．ここでテストセットとは，30 種の料理と各料理に類似した料理を 1 から 3 種人手で選定してペアとして結びつけた 30 個のペアを指す．類似した料理の合計は 57 個となった．類似した料理を入力語に対する類似語とし，入力語と紐づけた類似語を類似度ランキング上で高く順位付け出来るかで構築した知識グラフの評価を行う．

類似度ランキングの作成に際して，知識グラフ上の各エンティティを表現するベクトルを定めた．知識グラフ上のエンティティを表現するベクトルは，知識グラフにおけるエンティティ間の接続関係を 1 と 0 で表した隣接行列を用いて表現した．

入力語を表現するベクトルと料理エンティティを表現するベクトルに関するコサイン類似度に従ってランキングを作成する．ランキング対象とする料理エンティティは，エンティティを表現するベクトルの要素の総和が 4 以上かつ Wikidata から取得したエンティティのラベル情報に，日本語のラベルが含まれている 513 個の料理エンティティである．

4.2 類似度ランキングにおける順位による評価としきい値の分析

テストセット中の入力語とする 30 種の料理全てに対して類似度ランキングを生成した．類似度ランキングの生成は，拡充していない知識グラフと，$U_{rate}(c, m)$ に関するしきい値 θ に基づいて拡充した知識グラフに対して行った．θ は 0 から 0.28 まで 0.01 刻みに変化させ，調査結果からしきい値を決定した．θ の変化に対する，作成した類似度ランキングにおける似た料理全 57 種の平均順位，および拡充されたトリプルの数の変化を図 2 に示す．

θ が 0 から 0.03 に増加するにつれて平均順位は高くなっていき，$\theta = 0.03$ の時に平均順位が最高になった．$\theta > 0.03$ の区間では，θ が増加するにつれて緩やかに平均順位が低下する傾向が見られた．$\theta = 0$ として 7685 個全てのトリプルを知識グラフに拡充するよりも，$\theta = 0.03$ として使用率の低い材料を除外した 4035 個のトリプルを知識グラフに拡充した方が平均順位が高くなるということがわかる．拡充するトリプルを約半数削っているにもかかわらず平均順位は高くなっており，使用率の低い材料の情報を知識グラフに加えることが，類似語の推測性能に悪影響を与えることがわかる．料理の一般的でない材料情報によって，料理エンティティが特徴付けられたためだと考えられる．

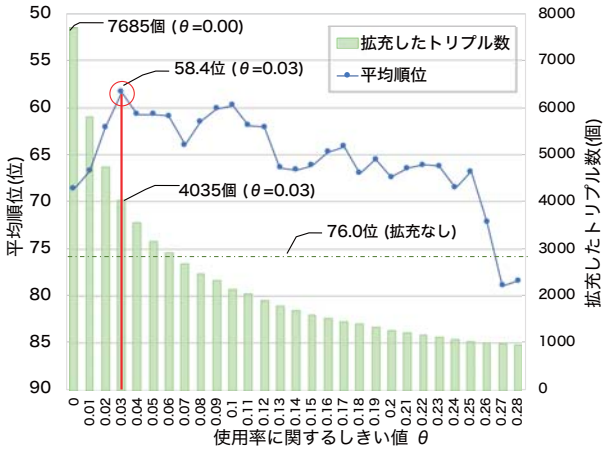


図 2: 類似度ランキングにおける平均順位と拡充トリプル数の θ に対する変化

そして，しきい値 θ が高くなるにつれて，拡充トリプル数の減少量が少なくなっている．このことから，使用率が高い材料の種類数に比べて，使用率が低い材料の種類数の方が多いことがわかる．

また，楽天公開データからのデータ拡充をしていない知識グラフでの結果は 76.0 位であったが，$\theta < 0.27$ においてはその値を上回っており，拡充の有無による差が見られる．

表 2: 類似度ランキングにおける類似語の順位の平均・標準偏差

	平均順位 (位)	標準偏差 (位)
拡充あり ($\theta = 0.03$)	58.4	76.7
拡充なし	76.0	68.1

表 3: 拡充結果の詳細 ($\theta = 0.03$)

構築された知識グラフの総トリプル数	15530 (個)
楽天レシピを用いて拡充したトリプル数	4035 (個)
材料情報を拡充した料理の種類数	573 (種)
拡充した料理の平均追加トリプル数	7.0 (個)

以上の調査から θ を，平均順位が最も高くなる 0.03 に決定して拡充を行った．拡充前後の平均順位，標準偏差を表 2 に示す．拡充後の平均順位の値が拡充前に比べ 17.6(位) 上昇しており，拡充の効果が見られる．一方，ランキング対象とした料理エンティティが 513 個であることを考慮すると，58.4 位と

いう平均順位はさほど高精度な結果ではないと考えられる．拡充結果の詳細を表 3 に示す．結果として楽天公開データから，4035 個のトリプルを知識グラフに拡充した．

4.3 類似度ランキングの実例

類似度ランキングの実例を示し，類似した料理が想定通りに推測出来た例とそうでない例での傾向の差異について述べる．

表 4: 「ボルシチ」に関する類似度ランキング

順位	料理名
1	◆ミネストローネ
2	ヴィシソワーズ
3	ハヤシライス
4	シチー
5	◆ビーフストロガノフ
6	カレーライス
7	ポトフ
8	クリームシチュー
.....	.....

まず，類似した料理が想定通りに推測出来た例として,「ボルシチ」に関する類似度ランキングから，上位 10 位までを抜粋したものを表 4 に示す．表 4 内で太字で示されている「ミネストローネ」と「ビーフストロガノフ」は，テストセット上で「ボルシチ」に類似した料理として結びつけられた料理である.「ミネストローネ」は 1 位,「ビーフストロガノフ」は 5 位に順位付けられており，想定通りの推測結果が出ている．上位 10 位以内に順位付けされている他の料理についても，煮込み料理が主に並んでおり，概ね似ていると思われる料理が多い．

表 5: 「トルティージャ」に関する類似度ランキング

順位	料理名
1	バカリャウ・ア・ブラース
2	オリヴィエ・サラダ
3	フラットブレッド
4	フォカッチャ
5	ビリヤニ，ザルツブルガーノッケルン
7	マギリッツァ，けんちん汁
.....	.....
47	◆オムレツ，他 22 種
.....	.....
133	◆卵焼き，他 7 種
.....	.....

反対に，類似した料理が上手く推測出来なかった例として,「トルティージャ」に関する類似度ランキングから，一部を抜粋したものを表 5 に示す．表 4 と同様に,「オムレツ」と「卵焼き」はテストセット上で「トルティージャ」に類似した料理として結びつけた料理である．ランキングにおいて「オムレツ」は 47 位,「卵焼き」は 133 位に順位付けられており，類似した料理を想定通りに推測出来ていない．上位に順位付けられている料理も，サラダや汁物，パンなどが見られ，似ていると思われる料理が少ない．

「ボルシチ」と「トルティージャ」の類似度ランキングを見比べてみると,「トルティージャ」の類似度ランキングには同率に順位付けされている料理が多い傾向にあるということがわかる.「トルティージャ」の類似度ランキングを見ると，5 位,7 位,47 位,133 位には同率に順位付けされている料理エンティティが存在する．この原因として考えられるのは，定めたベクトルが持つ情報量の少なさである．表 6 に,「ボルシチ」と「トルティージャ」，そしてテストセット上でこの二つの料理と結びつけた料理エンティティを表現するベクトルの要素の和を示す．料理エンティティを表現するベクトルの要素の和は，結びつく他エンティティの数を表す.「ボルシチ」とそれに紐づけた料理エンティティのベクトルの要素の和に比べ,「トルティージャ」とそれに紐づけた料理エンティティのベクトルの要素の和は小さい傾向にあることがわかる.「トルティージャ」やそれに紐づけた料理エンティティを表すベクトルに，他エンティティとの差を表す情報を十分に組み込めておらず，類似性という点でも，単純なコサイン類似度では差がつかない料理が同率順位に並んでいるのだと考えられる．

表 6: 料理エンティティを表すベクトルの要素の和

ボルシチ	27	トルティージャ	7
ミネストローネ	24	オムレツ	11
ビーフストロガノフ	27	卵焼き	20

5. おわりに

未知の属性の推測にむけて，今後は構築した知識グラフ上における関係を推測する手法の確立を目指す．またエンティティ間の接続関係だけでなく，プロパティなど他の情報を考慮した推測が効果的かどうか検証する．得られた推測結果を応答生成に反映させる枠組みについても検討する必要がある．

謝辞

本研究では，楽天株式会社が国立情報学研究所の協力により研究目的で提供している「楽天公開データ」を利用した．

参考文献

[Angles 08] Angles, R. and Gutierrez, C.: Survey of graph database models, *ACM Computing Surveys*, Vol. 40, No. 1, pp. 1–39 (2008)

[高橋 07] 高橋 潔考, 鈴木 遼, 下村 芳樹, 舘山 武史, 吉岡 真治, 武田 英明 : Web 情報を用いた設計知識情報データベースの拡充手法, 横幹連合コンファレンス予稿集, Vol. 2007, pp. 45–45 (2007)

[大塚 13] 大塚 嗣巳, 駒谷 和範, 佐藤 理史, 中野 幹生 : データベース検索音声対話システムにおける店舗属性値取得のための質問生成, 人工知能学会第 27 回全国大会 (2013)

[大野 18] 大野 航平, 武田 龍, ニコルズ エリック, 中野 幹生, 駒谷 和範 : 対話を通じた未知語のクラス獲得に向けた暗黙的確認の提案, 人工知能学会論文誌, Vol. 33, No. 1, pp. 1–10 (2018)

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

子どもを対象とする雑談対話システムにおける発話理解と感情生成

Utterance understanding and emotion generation in a dialogue system for children

山本 賢人[*1] 大村 俊哉[*2] 根本 凌[*2] 和田 史織[*1] 杉本 徹[*2]
Kento Yamamoto Syunya Omura Ryo Nemoto Shiori Wada Toru Sugimoto

[*1] 芝浦工業大学大学院 理工学研究科
Graduate School of Engineering and Science, Shibaura Institute of Technology

[*2] 芝浦工業大学 工学部
College of Engineering, Shibaura Institute of Technology

We propose a dialogue system for children with emotions. In this paper, we focus on utterance understanding and emotion generation in this dialogue system. In utterance understanding, we extract utterance intention and focus. Machine learning based on utterances collected from college students gives an accuracy of about 70% for utterance intention analysis and about 90% for focus extraction for children's utterances. In emotion generation, we deal with two types of emotion: short emotion and mood. The type of response is affected by mood, and short emotion is expressed as a facial expression. Regarding expressions of emotion, college students judged about 60% of emotions generated in dialogue as valid, but we could not confirm their usefulness in dialogue with children.

1. はじめに

近年，雑談対話システムの研究やビジネス利用が盛んである．たとえば，NTT ドコモ社の「しゃべってコンシェル」，ソフトバンク社の「Pepper」が例として挙げられる[東中 2014]．また，子どもを対象とした対話システムの研究として「たけまるくん」が存在する[西村 2004]．

一方で子どもが対話システムを利用する場合，子どもが退屈せずに楽しめるための話題や機能の充実といった課題が考えられるが，このような課題に焦点を当てた研究は少ない．

そこで我々はシステムの対象年齢を 7〜12 歳程度と仮定し，子どものユーザとクイズや石取りゲーム，雑談を楽しむことができるシステムの開発を目指している．クイズや石取りゲームといった子ども向けの対話の内容に加えて，システムに感情を持たせることや強化学習を行い子どもに適応することで，子どもが退屈せず利用できるようにする．

本稿では，この子どもを対象とする雑談対話システムにおける発話理解と感情生成手法に焦点を当て評価と考察を行う．

2. システムの概要

提案するシステムの処理の流れを図 1 に示す．

ユーザが音声による入力を行うと，発話理解部は入力結果から発話内容を解析し，発話意図や焦点などの情報を抽出する．

そして感情更新部では，発話理解で得られた結果をもとに感情の更新を行う．本研究では感情の概念をシステムの情緒と気分の二つに分けて定義する．得られた情緒に基づき，キャラクターの表情を出力する．また得られた気分に基づきシステムのいくつかの応答候補に対してその応答のスコアを増加させる．

一方でクイズ出題部では発話理解部で得られた現在の対話状況に応じて次にクイズとして出題する問題文の選択や，クイズの正誤判定を行う．

連絡先：山本賢人，芝浦工業大学大学院 理工学研究科，
〒135-8548 東京都江東区豊洲 3-7-5
E-mail: al14097@shibaura-it.ac.jp

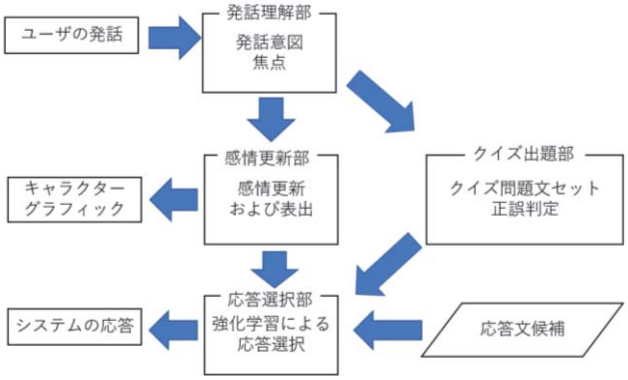


図 1. システムの処理の流れ

最後に応答選択部で応答文の選択を行う．本研究では現在の対話の状態をタスクとフェーズの組として表現し，それらに加えシステムの気分，発話行為タイプ，キーワードが含まれているか，ユーザの沈黙時間等に基づいて複数の応答文の候補のスコアリングを行い，スコアが最大の応答候補を選択する．ここでタスクとは，雑談・クイズ中・石取ゲームのいずれかである．一方，フェーズとはタスクが行われている・中断している・もう一度行うか現在確認中のいずれかである．

スコアの値は現在の状態に対しユーザの沈黙時間や相手を褒める内容や貶す内容が発話に含まれているか等を報酬とした強化学習により調整される．

最後に応答文を音声合成して出力する．

3. 発話理解手法

3.1 発話意図および焦点情報の抽出

(1) 発話意図解析

発話意図は発話行為タイプと属性の組で表す．発話行為タイプはその発話がどのような意味であるかを大まかに表す．一方属性は発話行為タイプが指し示す具体的な内容である．発話行為タイプは[Meguro 2010]が雑談対話システム用に定めた 33 種類の発話行為タイプを参考にし，後述する発話収集の実

験での発話内容を踏まえつつ本システムに合わせて 18 種類定義した. このうち 4 種類はシステムの発話のみに付与されるため, ユーザの発話に付与される発話行為タイプは 14 種類である. 発話行為タイプの例を表 1 に示す.

表 1. 発話行為タイプおよび焦点の例

発話文	発話行為タイプ	焦点
僕の誕生日は 3 月 2 日だよ.	陳述	誕生日, 3 月 2 日
寝たことありますか.	質問	寝る, ある
やらない.	質問返答. 否定	(無し)

発話行為タイプの付与は LIBLINEAR を用いてロジスティック回帰で行う. 属性は, 各発話行為タイプに対応した品詞の単語を形態素解析の結果から抽出する.

(2) 焦点解析

焦点とは発話の中心となる語であり, 名詞(句)と述語の組み合わせと定義する. 述語は動詞, 形容詞, 名詞(句)のいずれかである. なお名詞または述語どちらか一方しか含まれない場合はそれだけを抽出する. 焦点の例を表 1 に示す. 発話への焦点情報の付与は, CRFsuite を用いて形態素単位の系列ラベリングで行う.

3.2 発話データの収集および学習

(1) 学習に用いる発話の収集

発話意図解析および焦点解析の学習を行うため発話を収集した. 大学生に子どもを演じてもらい, システム役の人間と実際の対話を想定した模擬対話を行い 9 人から合計 563 発話を収集した. 発話文は人手で書き起こしを行い, 発話行為タイプと焦点情報を付与した.

(2) 発話意図解析に用いる素性の決定

発話意図解析の学習に用いる素性の決定を行った. 素性の候補と最終的な使用の有無を表 2 に示す.

表 2. 発話意図解析で用いる素性の候補

素性の種類	使用の有無
単語および品詞 n-gram	1,2,3-gram
何発話分の n-gram を使用するか	ユーザの発話と直前のシステムの発話
ユーザの発話の文字数	使用
直前のシステムの発話行為タイプ	不使用
直前のシステムの発話に焦点が含まれているか	不使用
ユーザの発話の形態素解析結果に JUMAN の 22 種類の意味カテゴリや 4 種類の固有表現情報が含まれているか	使用

素性の組み合わせは, 現在得られている最も正解率の良い組み合わせに新しい素性を加えて正解率を調べ, 正解率が向上した場合その素性を加えることを繰り返した. 素性は表 2 の上から順に加えた. 表 2 に示す素性の組み合わせで最高 0.701 の正解率が得られた. 正解率は学習データの発話のうち, LIBLINEAR が出力する確率値が最大の発話行為タイプと正解の発話行為タイプが一致した発話の割合である.

なお LIBLINEAR のアルゴリズムは L2 正則化 primal を使用し, パラメータはデフォルト値を使用した. また, 素性のスケーリングは行わなかった. 正解率を調べる際には 10 分割交差検定を用いた.

(3) 焦点解析に用いる素性の決定

焦点解析の学習に用いる素性の決定を行った. 素性の候補と最終的な使用の有無を表 3 に示す.

表 3. 焦点解析で用いる素性の候補

素性の種類	使用の有無
単語 n-gram	1,2gram
品詞 n-gram	1,2,3gram
直前のユーザとシステムの発話行為タイプ	使用
直前のシステムの発話に焦点が含まれているか	不使用
ユーザの発話の文字数	使用
発話文中での形態素の位置を正規化した値*1	使用
ユーザの発話の形態素解析結果に JUMAN の 22 種類の意味カテゴリや 4 種類の固有表現情報が含まれているか	不使用

*1 $\frac{\text{発話文中での形態素の位置}}{\text{発話文の形態素数}} * 10$

素性の組み合わせの求め方は発話意図解析と同様である. ただし評価基準は焦点の F 値および再現率を用いた. F 値がほぼ等しい場合は再現率が高いものを選択した. 表 3 に示す素性の組み合わせで 0.688 の再現率が得られた.

なお内部アルゴリズムおよびパラメータはデフォルト値を使用し, 5 分割交差検定で調べた.

4. 感情更新手法

4.1 感情生成モデル

[目良 2010]では, 感情の概念を情緒, 気分, 表情の 3 つに分けて定義している. 情緒は瞬間的に生じる短くて強い感情である. 気分は生起した情緒や内的状態に起因して生じる弱いながらも長く持続する感情である. 表情は情緒や気分などが身体を通じて表出されるものである. 本稿では, 感情の概念をシステムの情緒と気分の 2 つに分けて定義する. 情緒と気分は, Russel の感情円環モデル[Russel 1980]を用いて「快-不快」を表す好感度, 「覚醒-眠気」を表す覚醒度の二次元平面上の実数値の組として表現する. また, 極座標ベクトルの方向および大きさをそれぞれ情緒および気分の種類と大きさとする. 本研究で扱う情緒の種類を表 4 に示す.

表 4. 生起する情緒の種類

	覚醒	眠気
快	喜び	平常
不快	怒り	悲しみ

4.2 好感度の計算

最初にユーザの発話文から名詞, 動詞, 形容詞, 副詞を抽出する. 抽出した語の好感度データを実数値で取り出し, その平均値をその発話文における好感度とする. 好感度データには, 単語感情極性対応表[高村 2006]を用いる. この表は, ある単語が一般的に良い印象を持つか, 悪い印象を持つかを-1 から+1 の実数値で表したものである.

4.3 覚醒度の計算

覚醒度はシステムがユーザの発話に興味を持つかどうかで判定する. システム自身の好き嫌いやユーザの発話行為タイプ, クイズや石取ゲームを連続して続けた回数に応じてスコア付けを行い-1 から+1 の実数値を決定する.

システム自身の好き嫌いは, システムのプロフィールとして好きなもの, 嫌いなものを登録しておき, ユーザの発話に好きなものの語が含まれればシステムは興味を持ち覚醒度が上がる. 嫌いなものの語が含まれれば覚醒度が下がる.

ユーザの発話行為タイプが自己開示や情報提供などを意味する陳述である場合は, システムは興味を持ち覚醒度が上がる.

クイズやミニゲームを連続して続けた場合にも覚醒度は上がるが，すぐにやめてしまった場合には覚醒度は下がる．また，長く続けた場合にも飽きが来ると考え覚醒度は下がる．

4.4 情緒および気分の更新

情緒と気分を以下の式で更新する．まず 4.2 節，4.3 節の手法で好感度 Fr および覚醒度 Ar を求め初回の情緒$\mathrm{E}_1 = \binom{\mathrm{Fr}}{\mathrm{Ar}}$を決定する．次に，気分Mを決定する．初回は$\mathrm{M}_1 = \mathrm{E}_1$とする．2 回目以降は 1 回目と同様にFrとArを求め，以下の式により情緒E_nを更新する．

$$\mathrm{E_n} = (1-\mathrm{t})\binom{\mathrm{Fr}}{\mathrm{Ar}} + \mathrm{tM_{n-1}} \qquad (\mathrm{n} \geq 2)$$

tはどれだけ現在の気分を反映させるかの定数であり今回は 0.5 とした．気分の更新方法は生起した情緒の種類によって以下の二つに分けられる．

1. 生起した情緒の種類が直前の情緒の種類と異なる場合:

$$M_n = E_n$$

2. 生起した情緒の種類が直前の情緒の種類と同じ場合:

$$M_n = \frac{\alpha}{2}\left(M_{n-1} + E_n\right) (\alpha\text{: 気分の減衰率})$$

以上により情緒および気分を更新する．

4.5 情緒に応じた表情の生成

生起した情緒はシステムの表情として出力される．[中村 1996]では人類に共通していると考えられる表情の特徴をまとめている．本研究ではこの特徴をもとに表情の作成を行った．

4.6 気分に応じた応答戦略

システムの現在の気分に応じ，いくつかの応答に対してその応答の選択されやすさを変化させる．システムの気分が良いときはシステムの方から積極的に話すようになると考え，システムの自己開示，話題に基づく応答，オープンクエスチョンの応答をしやすくなるようにスコアを増加させる．システムの気分が悪いときは，システムは消極的になると考え，傾聴や当たり障りのない応答をしやすくなるようにスコアを増加させる．スコアの増加量は現在の気分M_nの好感度と覚醒度の二乗和の平方根により求める．

5. システムの実装および評価実験

5.1 プロトタイプシステムの実装

本研究のシステムのプロトタイプを作成した．このプロトタイプに用いたキャラクターの喜び，怒り，悲しみ，平常時での発話状態のグラフィックを図 2 に示す．



喜び　　怒り　　悲しみ　　平常(発話時)

図 2. キャラクターの各感情のグラフィック

また，今回は音声認識を用いずユーザが話した音声をオペレータが入力する形で対話を行った．

クイズは子どもが興味を持つと考えられる話題の問題を 23 問人手で子どもに合わせた言葉遣いで作成した．

5.2 発話理解手法および感情生成手法の評価実験

前述したプロトタイプを用いて小学 2 年生 3 人，5 年生 2 人と対話を行った．対話では感情を表出するシステム，しないシステムそれぞれと各 7 分間対話を行い，対話するシステムの順番はランダムに入れ替えた．合計 244 発話を収集した．ユーザとシステムの対話例を表 5 に示す．

表 5. 対話例

番号	発話者	発話文
1	システム	はじめましてボクの名前はクリームだよ，さっそくだけどクイズか石取ゲームやらない，それとも僕とお話しする?
2	ユーザ	石取.
3	システム	石取りゲームは1から数を数えていって先に21を言ったほうが勝ちだよ．1回で3まで数をふやすことができるよ．さいしょはボクからスタートだよ．じゃあ 1，2.
…		
4	ユーザ	21.
5	システム	おめでとう，君の勝ちだよもう一回やる？
6	ユーザ	もう一回やる.
…		
7	システム	ボク，チョコレートが大好き.
8	ユーザ	僕はふつう.
…		
9	ユーザ	クイズ.
10	システム	それでは問題です，あさごはんによくでる，ねばねばしてい大豆からできた食べ物は何でしょう，1 番とうふ，2 番なっとう，3 番ドーナツ.
11	ユーザ	2 番.

最初に 1 のようにシステムから発話が行われる．2 から 6 は石取ゲームの発話であり，7 から 8 は雑談中の発話，9 から 11 はクイズ中の発話である．

(1) 発話理解の精度

3.2 節で述べた大学生の発話データを学習データとして，LIBLINEAR と CRFsuite の学習を行った．その後プロトタイプを使用して子どもから得た発話データをテストデータとして与え適合率および再現率を求めた．発話意図解析の適合率および再現率を表 6 に，焦点抽出の適合率および再現率を表 7 に示す．

表 6. 発話行為タイプごとの適合率および再現率

発話行為タイプ	適合率	再現率
陳述	0.456（21/46）	0.750（21/28）
質問	0.846（33/39）	0.804（33/41）
質問返答.肯定	0.875（14/16）	0.560（14/25）
質問返答.否定	1.000（02/02）	0.400（02/05）
質問返答.その他	0.823（14/17）	0.875（14/16）
話題変更		0.000（00/02）
相槌	0.470（08/17）	0.615（08/13）
クイズ.回答	1.000（20/20）	0.645（20/31）
石取	0.861（62/72）	1.000（62/62）
定型	0.000（00/02）	
要求	0.000（00/01）	
+評価		

-評価	0.000 (00/03)	
その他	0.333 (03/09)	0.142 (03/21)
平均	0.725 (177/244)	

表 7. 焦点抽出の適合率および再現率

ラベル	適合率	再現率
焦点	0.717 (0160 / 0223)	0.824 (0160 / 0194)
非焦点	0.959 (0802 / 0836)	0.927 (0802 / 0865)
平均	0.908 (0962 / 1059)	

表 6 に関しては，発話行為タイプの出現頻度に差がみられ，出現頻度が高い発話行為タイプほど再現率も高い．これは学習データにも同様の傾向がみられ，学習データの量が十分であるため正しく判別が出来ていると考えられる．一方で話題変更，定型，要求，+評価，-評価は出現数が少ないか全く出現していないため分類精度を調べられなかった．これらは学習データにもほとんど出現しないため，発話行為タイプの種類自体の見直しや収集方法を工夫する必要があると思われる．

表 7 に関しては，焦点の再現率が学習データの 0.688 に比べ向上している．これは，子どもが話す文に複文や複雑な係り受け関係が少ないため，文字数や形態素の位置の素性がより有効に働いたためと考えられる．

学習データおよびテストデータの形態素数の分布をそれぞれ図 3，図 4 に示す．

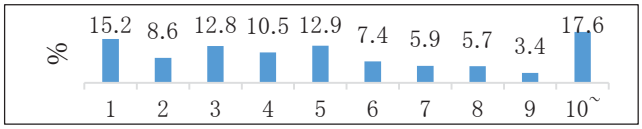


図 3. 学習データ(大学生)の形態素数

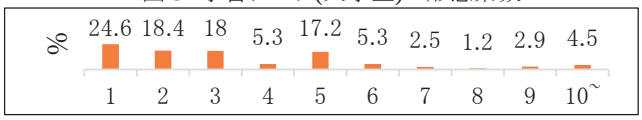


図 4. テストデータ(小学生)の形態素数

小学生の発話は形態素数が 10 以上の発話が少ないことが分かる．

(2) 感情生成手法の有効性の評価

対話後に各システムについて 4 段階(1~4)で評価してもらった．評価内容および各システムの平均値を表 8 に示す．

表 8. 各システムのアンケート評価結果

アンケート項目	感情あり	感情なし
上手に会話できていたか	2.8	3.0
話していて面白かったか	3.0	3.0
また話してみたいか	3.0	3.2
仲良くなれそうだったか	3.0	3.0

アンケートの結果から，感情を表情として表出するシステムとしないシステムで，あまり差異は見られなかった．これは実際に実験の様子を観察して，子どもが画面を見ずに別の方を見ていたり，画面を見ていても字幕を読むことに集中していたことから，システムの表情の変化に気づかなかったのではないかと推測される．従って，システムの表情の変化を分かりやすくするために，表情と共に背景色を感情に合わせた色に変化させるなどの工夫をする必要があると考えられる．

(3) 感情生成手法の妥当性の評価

大学生の被験者 7 人に，子どものユーザとシステムの対話ログを読み，システムの発話時に生成される情緒が妥当かを 5 段階で評価してもらった．評価は，ユーザの発話に対するシステムの発話を 1 対話として合計 36 対話に対して行った．被験者が選択した評価のうち，最も多かったものをその対話の評価とし，各評価に割り当てられた対話の数とその割合を表 9 に示す．

表 9. 各評価に割り当てられた対話数とその割合

評価	1 妥当である	2 まあまあ妥当である	3 どちらとも言えない	4 あまり妥当でない	5 妥当でない
対話数	16	6	5	5	4
割合	44%	17%	14%	14%	11%

1 および 2 の評価で約 61%の割合となった．4 および 5 の評価となった対話には，ユーザ発話が質問なのに対し怒りなどの情緒が生成されている場合があった．

6. まとめ

本研究では子どもを対象とする雑談対話システムの提案を行い，実際にプロトタイプを構築し，発話理解の精度および感情生成の手法についての評価を行った．発話意図解析では約 7 割，焦点解析については約 9 割の正解率で分類が行えるという結果を得た．一方，感情生成については約 6 割の対話について生成された情緒が妥当という結果を得たが，子どものユーザに対する感情の表出の有用性は確認できなかった．

今後は発話検出や音声認識を組み込んだ発話理解部の開発や，感情表出の仕方の見直しを行う．またシステムのクイズの自動生成機能の実装や，強化学習による応答選択の最適化に関する評価実験を行う．

参考文献

[東中 2014] 東中竜一郎: 雑談対話システムに向けた取り組み，人工知能学会第 70 回言語・音声理解と対話処理研究会，2014.

[西村 2004] 西村竜一，西原洋平，鶴身玲典，李晃伸，猿渡洋，鹿野清宏: 実環境研究プラットホームとしての音声情報案内システムの運用，電子情報通信学会論文誌 J87-D-II(3) pp. 789-798，2004.

[Meguro 2010] Meguro, Toyomi, Minami Yasuhiro, Higashinaka Ryuichiro, Dohsaka Kohji: Controlling listening-oriented dialogue using partially observable Markov decision processes, Proceedings of COLING, pp. 761-769, 2010.

[目良 2010] 目良和也，市村匠，黒澤義明，竹澤寿幸:情緒計算手法と心的状態遷移ネットワークを用いた音声対話エージェントの気分変化手法，日本知能情報ファジィ学会誌 22(1) pp. 10-24，2010.

[Russel 1980] J.A. Russel: A circumplex model of affect, Journal of Personality and Social Psychology 39(6), pp. 1161-1178, 1980.

[高村 2006] 高村大也，乾孝司，奥村学: スピンモデルによる単語の感情極性抽出，情報処理学会論文誌 47(2) pp. 627-637，2006.

[中村 1996] 中村真: 表情とコミュニケーション，繊維製品消費学 37(1) pp. 4-9，1996.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

テンプレートを用いた疑問文生成による対話応答文DBの拡充

Augmentation of Dialogue Database by Generating Interrogative Sentences Using Templates

長内洋太 *1 Yota Osanai　尾形朋哉 *1 Tomoya Ogata　小町守 *1 Mamoru Komachi　下川原（佐藤）英理 *1 Eri Sato-Shimokawara　和田一義 *1 Kazuyoshi Wada
山口亨 *1 Toru Yamaguchi　高谷智哉 *2 Tomoya Takatani

*1首都大学東京 Tokyo Metropolitan Univercity　*2トヨタ自動車株式会社 TOYOTA MOTOR CORPORATION

In recent years, there are many dialogue systems selecting an appropriate response from a dialogue database. However, when the number of sentences in the database is small, it may not contain an appropriate sentence in an open domain setting. Therefore, it is desirable to prepare a large number of candidate sentences and to fully expand the dialogue database beforehand. In this research, we generate templates from handwritten responses from actual dialogue logs to augment interrogative sentences for a dialogue database.

1. はじめに

近年，コンピュータが自然言語を用いて人と会話を行うための対話システムの研究が行われている．対話システムには，大別して，経路案内や Web 検索などの特定の目的のために発話，応答を行う task-oriented なもの [Williams 07] と，特定の目的がない，主に雑談を行う non-task-oriented なもの [Wallace 09] の 2 種類がある．

non-task-oriented な対話システムにおいて，事前に応答の候補となる文を生成，または収集し，データベースに大量に蓄えておくことで，ユーザ発話に適切な応答を文選択によって出力するという手法が古くから研究されている [Higashinaka 06][柴田 09]．この手法は，応答文を出力する際，フィルタリングやスコアリングを用いることによって，非文を出力しづらいという強みがある．しかし，データベース内にユーザ発話に対する応答文としてふさわしい文が存在しない場合，適切でない文を選択し，応答文として出力してしまうという問題が存在する．そのため，応答文のデータベースは様々なユーザのあらゆる発話に対応できるよう，応答文となり得る様々な応答文を大量に保持していることが望ましい．

また，比較的小規模な日本語対話データや，Twitter などのスラングやノイズの多いデータセットを用いて生成された応答文は，それらを除外するためのフィルタリングに非常にコストがかかるという問題点がある．

そこで，本研究では，実際の対話ログから応答となる疑問文を手書きで作成し，テンプレート生成の観点から分類する．その後，日本語形態素解析器 JUMAN++の辞書情報を外部リソースとして用いて，分類ごとにルールベースの手法を用いて作成したテンプレートを適用することで，文選択の選択肢となりうる候補文を簡易かつ大量に生成する手法を提案する．

本研究の貢献は，以下の通りである．

1. ユーザ発話に対して応答疑問文を人手で作成し，それらを応答可能性の観点から分類した
2. ユーザの 1 発話から，対話を続けるための日本語の対話応答候補文のテンプレートを作成した

連絡先: 長内 洋太，首都大学東京，osanai-youta@ed.tmu.ac.jp

2. 関連研究

テンプレートから日本語応答文を生成する研究として，加藤らの研究 [加藤 15] や山内らの研究 [山内 14] がある．

加藤ら [加藤 15] は，話題となる単語とともにユーザ発話内の感情を表現する語を抜き出し，話題語のカテゴリと感情の一致を考慮してテンプレートに挿入した．話題を一致させることは対話のロボット応答における違和感を低減することが示されたが，固有表現抽出の精度が悪いことや，感情語の活用や品詞を考慮していないことなどに問題があり，生成文の流暢性が低いという問題がある．本研究では，彼らのいう‘話題語’に着目し，どのような品詞が話題語に該当し，どのようなカテゴリであれば応答の生成にテンプレートを用いることができるかの分類を行い，自然な応答文が生成されるテンプレートを作成した．また，ユーザ発話への応答疑問文を事前に人手で作成し，単語を話題語として抽出すべきかどうかで分類した．

また，山内ら [山内 14] は，日本語文について，ある特定の分野に関する用語の情報が記述された用語データベースを用意し，ユーザとの会話の話題をその特定分野に誘導することで，テンプレートを用いた応答文生成を行なう．ただし，特定の分野の単語に対してのみ情報を所持しているため，他の分野の単語に対しては適切な応答文生成ができないという問題がある．そこで，本研究では，日本語形態素解析器 JUMAN++の辞書情報を用いて，ある分野の単語のみでなく，分野を限定せずに，日本語の応答疑問文を生成した．

3. テンプレートの生成

テンプレートによって応答疑問文を作成できるユーザ発話群の発見のため，対話データ中のユーザ発話の分類を行う．その後，応答疑問文を作成できるものに対して応答文を人手で作成し，作成した応答疑問文に対して，テンプレートを作成できる応答疑問文群発見のための類型化を行う．

応答疑問文の作成において，文脈を考慮する場合，応答疑問文の候補は多岐に渡る．したがって本研究では，文脈を考慮せず，ユーザの 1 発話のみを見るという制約のもとで応答疑問文を作成した．また，ユーザ発話のどの内容を受けた応答疑問文であるかを明確化するために，ユーザの使用した単語（見出

表 1: 各分類におけるユーザ発話の件数と例

分類	件数	ユーザの発話の例	作成した応答文
単語抽出が必要	733	**本**を読んだりラジオを聴いたりしてます。 特別な**習い事**はしてないです。	最近はどんな**本**を読まれましたか？ 何か**習い事**をしてみたいと思ったことはありますか？
単語抽出が不要	72	おはよう，今起きました。 ただいま帰りました。	おはようございます。今日は何をされるご予定ですか？ お帰りなさい、お疲れ様でした。
User-negative	125	ダメです、全く泳げません。 全く興味ありません、できません。	- -
User-question	54	あなたのお名前は？ はい、毎日してます。一緒にしませんか？	- -
単語抽出不可	150	そうですね。 うーん…。	- -
書き起こし文なし	66	-	-

し語）を一語以上使用することも制約として加えた．

以上の制約下で作成した応答疑問文を用いて，応答疑問文同士で類似した箇所を探し，テンプレートを作成する．

3.1 データセット

実際に用いる対話ログに関しては，先行研究 [下川原 16] で収集，編集したデータを扱う．このデータは，主にロボットが高齢者のユーザに対して質問を行い，ユーザがそれに返答をする雑談形式の対話ログである．ユーザの発話ログには，ユーザ発話を聞き取って書き起こされたものを扱う．本研究では，2014 年 6 月に収集した 6 ユーザの発話から各 200 文，計 1,200 文を抜き出し，それに対して分類を行い，疑問文を作成可能なユーザ発話に対して応答文を人手で作成した．

3.2 データセットの分類と応答疑問文作成

対話ログにおけるユーザ発話文の分析を行う．応答疑問文を作成するかどうかで場合分けし，それぞれに対しさらに分類を行う．応答疑問文の作成が可能である場合は，ユーザ発話から単語を抽出して応答疑問文を作成したか否かによって 2 つの分類に分かれる．疑問文の作成が不可，もしくは応答を作成すべきでないとき，これらは，‘ 単語の抽出が可能であるか不可であるかに関わらず，ユーザ発話が否定的な応答をしている ’，‘ 単語の抽出の可否に関わらず，ユーザ発話が疑問文である ’，‘ 抽出するべき単語がユーザ発話に含まれない ’，‘ ユーザ発話データが存在しない ’ の 4 つの分類に分かれる．

単語抽出が必要　ユーザ発話内に，ユーザとシステム間で今行われている会話の話題が読み取れる単語，主に自動詞が存在し，それを抜き出して用いることで応答疑問文が作成できる場合が該当する．

単語抽出が不要　主に ‘ おはよう ’ や ‘ ただいま ’，‘ ありがとう ’ などの挨拶や「言語行為の定型表現・運用上固着した語彙 [加藤 06]」，それにより応答が疑問文である必要がない場合などが，抽出の必要がない場合に該当する．

User-negative　ユーザが ‘ 嫌い ’，‘ 全くない ’ 等の語を含む否定的な発話をした場合が該当する．この場合，ユーザは 1 つ前でロボットが提示した話題に対して，会話の続行を望んでいないと推測できる．したがって，単語抽出によって疑問文の作成が可能な場合であっても，他の話題についての発話を行なった方が適切であると判断したため，疑問文を作成していない．

User-question　ユーザの発話が疑問文である場合が該当する．この発話に対する応答の作成には，ユーザやロボットのパーソナリティ情報や外部のリソース等から情報を取得する必要があり，ユーザ発話に含まれる単語をテンプレートに挿入して疑問文を作成することは適切ではない．

表 2: 各名詞の細分類内訳と例

細分類	該当する名詞の例	件数
普通名詞	服, メール, テニス	353
地名	日本, 東京, パリ	79
サ変名詞	会話, 経験, 予定	79
時相名詞	梅雨, この後, 最近	57
数詞	100, 何, 一	10
副詞的名詞	前, ところ	3
組織名	ユニクロ	1

単語抽出不可　感嘆詞のみであったり，聞き取りがうまくいかずに名詞等が途中で途切れている場合などの，主にユーザ発話に自動詞が存在していない場合が該当する．

書き起こし文なし　ユーザが実際に発話していない場合に加え，ノイズやユーザの声量の問題で，ユーザ発話が正しい文として書き起こされていない場合が含まれる．

以上の分類に基づき，‘ 単語抽出が必要 ’ と ‘ 単語抽出不要 ’ の応答疑問文を作成可能なユーザ発話に対して，疑問文を人手で作成した．各分類におけるユーザ発話の件数と，作成した応答を表 1 に示す．ユーザ発話における単語抽出を行う場合，ユーザ発話に含まれる単語を用いつつ応答を作成した．

3.3 抽出した単語の分類

テンプレートに挿入される頻度の高い品詞の特定のため，ユーザ発話の分類 ‘ 単語抽出が必要 ’ に対して作成した各応答文 733 件について，応答作成で抽出された単語の品詞ごとに分類を行う．ユーザ発話に含まれる 1 単語を抽出して作成した応答文について，日本語形態素解析器 JUMAN++（version 1.01）を用いて，抽出した単語の品詞分類を行った．

抽出した品詞を分類したところ，名詞 470 件（例：犬，ご飯，今日），動詞 153 件（例：起きる，食べる，行く），形容詞・形容動詞 110 件（例：静か，長い，若い）となり，抽出される単語の品詞は名詞が最も多かった．そのため，本研究では名詞に対するテンプレートを考えることにした．

名詞を用いて作成した応答文に対して，JUMAN++で解析した結果における，名詞の細分類，JUMAN++辞書における ‘カテゴリ’，‘ドメイン’ の両分類を用いることで，各名詞の属性付けを行う．ここで言及した JUMAN++のカテゴリとは，語の上位下位関係に基づくものであり，‘人’，‘動物’，‘植物’ 等 22 種存在する．ドメインは単語をオントロジー横断的に意味情報で区分したものであり，‘文化・芸術’，‘スポーツ’，‘健康・医学 ’ などの 12 種が存在する [黒橋 16]．

表 3: カテゴリ・ドメインの組の内訳と例 - 上位 10 件

カテゴリ名	ドメイン名	該当する名詞の例	件数
時間	-	昨日, この後	60
抽象物	-	天気, 流行	59
抽象物	家庭・暮らし	メール, 買い物	28
人工物-その他	家庭・暮らし	ゴミ, クーラー	23
人工物-食べ物	料理・食事	食べ物, 缶詰	23
人	家庭・暮らし	主人, 女の子	21
抽象物	文化・芸術	文化, 演奏	18
抽象物	スポーツ	テニス, スポーツ	15
人工物-その他	-	話題, ボールペン	13
動物	-	虫, ペット	13

作成した応答文のうち，名詞を抜き出して作成した応答文の名詞について品詞細分類を付与した内訳を表 2 に，カテゴリとドメインの組の出現頻度上位 10 件と該当する名詞の例を表 3 に示す．この中で，JUMAN++の解析結果において，名詞の解析候補が複数存在する名詞は，複数の分類のテンプレートに挿入することが考えられるため，それぞれの解析候補を別の名詞として考えて出現頻度を計算した．また，解析結果が一つであっても，複数のカテゴリやドメインを持つ単語に対しては，意味の詳細な絞り込みのため，その単語の持つ複数のカテゴリ，または複数のドメインの組み合わせを独立した分類として考える．また，カテゴリとドメインのどちらか，もしくはその両方を持たない語も存在した．

3.4 テンプレートの作成

ユーザ発話から，名詞を一つ抽出して挿入することで応答文となるテンプレートを作成する．

今回は，抽出した各名詞と述語の共起などは考えず，述語を各テンプレートに固定する形で作成した．このうち，作成できたテンプレートの該当する分類を表 5 に示す．表 5 において，`<word>`で示されるのが実際にユーザの発話から抜き出した単語である．

また，表 2 の‘サ変名詞’に該当する名詞は，それをする主体が必要であるという点で述語と扱いが類似しており，今回は挿入する名詞から除外している．

3.5 考察

内訳の上位に位置する，カテゴリが‘時間’や‘抽象物’である場合，ドメインが‘家庭・暮らし’である場合などの非常に広範に渡る分類では，カテゴリとドメインに加えて，名詞の細分類を用いても抽出した名詞の情報を絞りきれず，その分類に該当する名詞が適用できるテンプレートの考案が難しい場合がある．そのため，さらに細かい名詞の区分が必要である．

カテゴリ‘人’の分類では，‘妻’，‘嫁’などの名詞がよく見られた．これらの単語は，ユーザが発話者である場合と，システム発話の場合において名称が変わる名詞である．今回の手法ではこれらの単語への対応は難しいが，人称名詞の辞書を用いて活用することで，テンプレート作成が可能になると考える．

また，カテゴリ‘抽象物’に該当する名詞に見られるように，適切な述語が，抽出した名詞ごとに様々である分類において，述語をテンプレートに固定する今回の手法は適していない．このような場合，抽出した各名詞と共起しやすい述語を同時に挿入するテンプレートを作成することで，適切な出力文が得られる可能性がある．

表 3 において，‘家庭・暮らし’のドメインが頻出であることは，データセットが高齢者とロボットの日常的な対話ログであることに起因すると考えられる．

4. テンプレートを用いた応答文生成実験

3.4 節で考案したテンプレートに対して，ユーザ発話文から抽出した単語を挿入して応答文を生成し，自然な応答文が出力されるかどうかを検証する．今回は，表 4 に示す 3 つの分類のテンプレートに対して，対話ログデータ内の全単語を挿入して応答文を生成し，カテゴリのみ・ドメインのみ・両方を用いた場合の 3 つの場合において検証を行った．全体として，3 つの分類に属する名詞 118 語から，応答候補文 557 文が生成された．テンプレートに単語を挿入した出力例を表 5 に示す．

表 5 に示した例のうち，カテゴリとドメインの両方を用いて単語を挿入した場合，出力は非文ではなく，簡潔かつ応答文の範疇を逸脱していないといえる．また，カテゴリのみもしくはドメインのみを用いた場合の自然な文の出力割合が，該当するカテゴリ，ドメインによって大きく異なることがわかる．例として示した分類のうち，‘抽象物’というカテゴリのみを用いた場合は，この区分が非常に広範であるため，作成したテンプレートに適する単語を抽出できていないことが問題である．ドメイン‘料理・食事’については，ドメイン側が広範な区分であることと，‘冷蔵’や‘煮（る）’のような事態性名詞が普通名詞として挿入されることが問題となっていた．

しかし，広範な分類の名詞を挿入した場合など，依然として文として意味が通じないものが生成された．今回の設定では，全体の 8.98 %の出力が適切でない応答文であり，挿入した単語によって，その単語を挿入した応答文全てが不自然である場合と，ある特定のテンプレートに挿入した場合のみ不自然な文が出力された場合が存在した．適切でない出力の事例としては，以下の 3 つが見られた．

挿入する単語の粒度の問題　表 5 の‘スポーツの他に何かスポーツはされるのですか?’等の出力に見られる，該当の分類に属する単語同士の意味的な粒度が揃っていない場合である．この原因として挙げられるのは，カテゴリまたはドメインの意味的なばらつきが原因で，そのテンプレートに挿入するのにふさわしくない単語が挿入されてしまうことである．この問題に対しては，異なる辞書を用いることによって，単語の分類をさらに細かく行うことが必要がある．この問題については，テンプレート自体に含まれる単語や，それに類する単語に，フィルタをかけてテンプレートに挿入しない等の対策が必要である．

形態素解析の誤りによる問題　JUMAN++の解析結果が実際の単語の意味と異なる解析を行なった場合である．例としては，表 5 の細分類‘地名’における，‘将棋’のような単語である．この問題は，テンプレート作成の段階での対処が難しいため，異なる出力を行なった場合に，順次人手で JUMAN++辞書を修正，拡張する等の対応が望まれる．

外部知識や環境情報の問題　ロボットの応答が発話者にとって当然の事柄であり，会話の進行を妨げてしまう場合も存在する．例としては，日本人に‘日本にはよく行かれるのですか?’等の応答を出力する場合である．ただし，この場合を適切でないと判断するのは難しい．なぜなら，この分類の不自然な出力文は，ユーザやロボットの立ち位置や周囲の環境などの状況によって，自然な文になりうるからである．この場合，応答候補文の生成としては問題ないが，実際に候補文の中から応答文を選択する際に，ユーザの出身がどこであるかなどのユーザの登録情報や，ユーザとロボットが今どこで会話しているかなどの対話環境といった，追加情報の取得が必要になる．

表 4: 各分類における出力の統計

分類			カテゴリ・ドメイン			カテゴリのみ			ドメインのみ		
細分類	カテゴリ	ドメイン	単語数	出力文数	正文割合	単語数	出力文数	正文割合	単語数	出力文数	正文割合
地名	-	-	79	395	89.9	-	-	-	-	-	-
普通名詞	抽象物	スポーツ	15	90	98.9	296	1,939	9.3	16	96	92.7
	人工物-食べ物	料理・食事	24	72	87.5	38	114	61.4	61	183	38.8

表 5: 作成したテンプレートと対話ログから抽出した単語を用いた出力例.
（文頭に‘*’のあるものは適切でない出力例，‘?’のあるものは疑問の残る，または議論の必要性のある出力例）

分類			作成テンプレート	出力
細分類	カテゴリ	ドメイン		（**ゴシック**が挿入単語）
地名	-	-	<word>には行ったことはありますか？ <word>にはよく行かれるのですか？ <word>と聞くと何を思い浮かべますか？ ご自宅から<word>へはどう行かれるのですか？ <word>へ行ってみたいと思いますか？ <word>でよく行くお店はありますか？	*将棋に行ったことはありますか？ ?**日本**にはよく行かれるのですか？ **横須賀**と聞くと何を思い浮かべますか？ ご自宅から**新宿**へはどう行かれるのですか？ **パリ**へ行ってみたいと思いますか？ **新宿**でよく行くお店はありますか？
普通名詞	抽象物	スポーツ	<word>はよくされるのですか？ <word>の他に何かスポーツはされるのですか？ <word>の観戦に行かれたことはありますか？ <word>はいつ始められたのですか？ <word>を始めたきっかけは何だったのですか？ <word>をしていて楽しい時はどんな時ですか？	**テニス**はよくされるのですか？ ***スポーツ**の他に何かスポーツはされるのですか？ **サッカー**の観戦に行かれたことはありますか？ **ジョギング**はいつ始められたのですか？ **体操**を始めたきっかけは何だったのですか？ **野球**をしていて楽しい時はどんな時ですか？
	人工物-食べ物	料理・食事	<word>はよく食べるのですか？ <word>はお好きですか？ ご自分で<word>を作られたことはありますか？	**刺身**はよく食べるのですか？ **ご飯**はよく食べるのですか？ ?**食べ物**はお好きですか？ **ヨーグルト**はお好きですか？ ご自身で**漬物**を作られたことはありますか？ ご自身で **弁当**を作られたことはありますか？

5. おわりに

本研究では，ユーザの発話に対して人手で応答文を作成し，作成した応答文の単語を，日本語形態素解析器 JUMAN++の辞書におけるカテゴリやドメイン等の情報を活用することでテンプレートを作成した．ユーザの発話から文脈となる単語を抽出し，JUMAN++の辞書上で該当する各分類のテンプレートに挿入することで，その単語を使いつつ対話を続行するための文が得られることがわかった．

提案手法は，生成したテンプレートと単語の条件が合致すれば一定数の文が得られる一方，テンプレートを考案することが難しい単語や品詞においては，他の手法を用いた応答文の生成が必要になる．また，カテゴリとドメインの分類ごとにテンプレート数に偏りがあるため，一部のカテゴリやドメインにおいては十分な応答文が得られない問題や，カテゴリまたはドメインが非常に広範であるために，単語情報を絞り込めない場合，適切なテンプレートの作成が難しいという問題が存在する．

また，今回の手法では名詞のみに焦点を絞ったため，ユーザ発話が名詞を含まない場合を考慮できていない点が課題として残る．ユーザ発話が名詞を含まない場合には，今までの対話ログから会話の文脈を考慮しつつ，ユーザ発話に含まれる述語を抽出し，テンプレートに挿入することが考えられる．

今後は，提案手法に加えて，ユーザ発話から得られる情報のみならず，ロボットとユーザの位置関係や名前等のパーソナリティを利用したテンプレートによる応答生成と，意味的な観点からの客観的な評価を行う予定である．加えて，一つの名詞を抽出してテンプレートに挿入する手法において，挿入する名詞と関連性の高い述語を同時に挿入できるテンプレートを作成することで，さらに豊富な応答文を得る可能性についても探っていきたい．

参考文献

[Higashinaka 06] Higashinaka, R., Prasad, R., and Walker, M. A.: Learning to Generate Naturalistic Utterances Using Reviews in Spoken Dialogue Systems, *Proceedings of ACL*, pp. 265–272 (2006)

[Wallace 09] Wallace, R. S.: *The Anatomy of ALICE* (2009)

[Williams 07] Williams, J. D. and Young, S.: Partially Observable Markov Decision Processes for Spoken Dialog Systems, *Computer Speech & Language*, Vol. 21, No. 2, pp. 393–422 (2007)

[下川原 16] 下川原 （佐藤）英理, 篠田 遥子, 李 海妍, 高谷 智哉, 和田 一義, 山口 亨：高齢者と音声対話ロボットの雑談履歴の解析, *JRSJ*, Vol. 34, No. 5, pp. 309–315 (2016)

[加藤 06] 加藤 重広：日本語文法入門ハンドブック, 研究社 (2006)

[加藤 15] 加藤 和樹, 柴田 千尋, 田胡 和哉：文のテンプレートの学習および感情を考慮した会話文の生成, 情報処理学会第 77 回全国大会講演論文集, pp. 175–176 (2015)

[黒橋 16] 黒橋 禎夫：日本語形態素解析システム JUMAN++ version1.0 マニュアル (2016), http://lotus.kuee.kyoto-u.ac.jp/nl-resource/jumanpp/jumanpp-manual-1.0p2.pdf

[山内 14] 山内 祐輝, Neubig, G., Sakti, S., 戸田 智基, 中村 哲：対話システムにおける用語間の関係性を用いた話題誘導応答文生成, 人工知能学会論文誌, Vol. 29, No. 1, pp. 80–89 (2014)

[柴田 09] 柴田 雅博, 冨浦 洋一, 西口 友美：雑談自由対話を実現するための WWW 上の文書からの妥当な候補文選択手法, 人工知能学会論文誌, Vol. 24, No. 6, pp. 507–519 (2009)

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

会話によるニュース記事伝達のための発話意図分類とデータベースの構築

Utterance Intention Classification and Database Construction for Conversational News Delivery System

横山 勝矢 *1 Katsuya Yokoyama　高津 弘明 *1 Hiroaki Takatsu　本田 裕 *2 Hiroshi Honda　藤江 真也 *1*3 Shinya Fujie　林 良彦 *1 Yoshihiko Hayashi
小林 哲則 *1 Tetsunori Kobayashi

*1早稲田大学 Waseda University　*2本田技術研究所 Honda R&D Co.,Ltd.　*3千葉工業大学 Chiba Institute of Technology

In a human-human conversation, a listener frequently conveys various intentions explicitly/implicitly to the speaker with reflexive short responses. The speaker recognizes the intentions of these feedbacks and changes the utterance plans to make the communication more smooth and efficient. These functions are expected to be useful for human-system conversation also, but when human face the system they do not always give the same feedback as they did with human. We investigated the feedback phenomena of human users against our system which is designed to transfer a massive amount of information like news articles by spoken dialogue. First, user's intentions of feedbacks that can affect the behavior of the system were classified from the viewpoint of roles in the progress of conversation: presence/absence of a user's interest, comprehension state, release/keep turn, and so on. Then, using our news delivery conversation system, we gathered user's short responses, and labeled the intention through listening test by the subjects. We also attempted automatic identification of those intentions. As a result, in the current system, it is found that there are many feedbacks to convey doubt and automatic identification of them is possible stably, but there are not so many other feedbacks.

1. はじめに

実会話システム利用時における，システム発話に対するユーザのフィードバックの現象を分析するためのコーパス収集について検討する．

人同士の会話において，聞き手は，話し手の発話に対し，しばしば反射的な応答を行う．これらの応答は，発話の理解状態や発話内容への興味の程度など，聞き手の様々な状態を，話し手にフィードバックする．話し手は，これらのフィードバックをもとに，相手の理解が乏しければ情報を補足し，興味がなさそうなら話題を変えるなど，自らの発話計画を適宜修正し，聞き手の状態に合わせた発話を行うことになる．

こうした一連の処理を，人とシステムとの会話にも適用することで，従来の一問一答の枠組みを超えた，円滑で利便性の高い会話システムを実現できることが期待されている [Fujie 06, Kobayashi 13]．ここで，聞き手からのフィードバックは，必ずしも言語的に明示的な形でのみ表現されるとは限らず，抑揚や声色などに表れるパラ言語を伴った複雑な形で表現される [益永 14]．このため，こうした機能の実現には，自然なフィードバックを含む大規模な会話音声コーパスが必要となる．

会話における聞き手からのフィードバックを扱うためのコーパスの研究には，既に様々なものがある．牧野ら [牧野 17] は，人—人の会話における話者交代の現象を分析対象として，話者交代の際に起こる韻律情報と視線情報を調査するためのコーパスを作成した．荒木ら [荒木 17] は，雑談における会話内容に対するユーザの興味の有無を分析対象として，MMDAgent を用い，WoZ（Wizard of Oz）方式により対話データの収集を行った．井上ら [井上 17] は，エンゲージメント（興味や意欲による対話継続欲求度）を分析対象として，自立型アンドロイド ERICA[Inoue 16] を用い，WoZ 方式により，会話時の相槌，笑い，うなずき，視線などのデータを収集した．

これら，人—人の会話や，WoZ 方式のように，応答時に人が介在して収集したコーパスには，極めて自然なフィードバックが含まれ，研究対象としては興味深い．しかし，現状，実際の会話システムを使用したときのユーザの振る舞いとは乖離している可能性があり，システムへ組み込む機能を実現するための学習データとしては必ずしも適切ではない．

このような観点から，実際の自律システムとの対話を通じてデータ収録した研究もある．長澤ら [長澤 17] は，インタビューを行うシステムにおいて，ユーザの反応によって話題を掘り下げるか，話題を転換するかなどの対話戦略を決めるために，実際にロボットを用いた対話を行ってデータを収録し，ユーザの発話への積極性を 5 段階でアノテーションしている．

本研究においても，長澤らの研究と同様に，実システムの利用時におけるユーザの応答を調査対象とすることを試みる．現在我々が作成しているニュース伝達のための会話システム [高津 18a] を対象として，ユーザが実際にシステムとの会話時にどのようなフィードバックを行い，システムはどのようにそれを利用しうるかを検討する．システムからの提示情報に対するユーザの興味の有無，発話交代に関する働きかけ（発話権を取りたいか，継続したいかなど），システム発話の理解状態など，ユーザがフィードバックの発話に込めた，システムの振る舞いに影響を与えうる情報を広くとりあげ，これらの機能を整理・分類するとともに，これらが，現状のシステムとの会話におけるユーザの応答にどの程度含まれるかを調査する．

以下本稿では，2. でユーザが会話時に行うフィードバックの分類について説明し，3. でデータセットの収集方法やアノテーションについてとその結果について述べる．

連絡先: 早稲田大学 理工学術院 知覚情報システム研究室
〒 162-0042 東京都新宿区早稲田町 27
E-mail:katsuya@pcl.cs.waseda.ac.jp

2. 発話意図分類

システム会話におけるユーザのフィードバック現象を分析するコーパスの収集に先駆け，フィードバックがシステムの振る舞いに影響を与えうる情報を広くとりあげ，これらの機能を整理・分類した．

フィードバックの役割は，システムにある種の要求を伝えることにあるが，これらは，ユーザの要求を直接明示的に伝えるものと，ユーザ状態等を伝達することで間接的，あるいは暗示的に伝えるものがある．本稿ではこれらをまとめて発話意図（intention）と呼ぶ．

円滑なシステムを実現するための機能として，人の発話とシステムの発話が同時に起きるべきではないと考え（発話の衝突回避），また，利便性の高いシステムを実現するための機能として，システムが伝える情報の量という観点から，その増減を行うフィードバックがあることを考えた．これらのフィードバックを分類し，以下の 6 つの発話意図を提案する．

・待機要求
・質問
・補足要求
・反復要求
・無関心
・既知

「待機要求」はユーザがある発話の後少し間を空けても話し続けることを示し，システムの発話は待ってほしいという意図である．従来の音声対話システムでは VAD(Voice Activity Detection) により発話の区切りが検出されたら次の発話を開始するというものであった．そのためユーザはまだ発話を継続するつもりにも関わらず，システムが発話を開始し発話の衝突が起こっていた．この「待機要求」を認識することでシステムは発話を待機し，衝突を陽に回避することができると考えられる．

「質問」「補足要求」「反復要求」は伝達する情報を増やすための意図である．まとまった量の情報を伝達するシステムにおいてユーザの理解状態に応じて情報量の調整を行う必要があるが，中でもユーザの疑問解消や十分な情報の伝達は不可欠である．そのため，これらの意図を認識することで伝達する情報量を増やす調整を行うことができると期待した．「質問」はユーザが疑問をもち陽に質問を行ったもので，対話行為で分類される「質問」と同様のものである．「補足要求」は「質問」のように陽に質問は行っていないが，システム発話に疑問点を感じているもしくはシステム発話に理解が追いついていないため補足説明がほしいという意図である．「反復要求」はシステム発話を聞き取れなかった場合や受け取れなかった場合にもう一度同じ発話を繰り返してほしいという意図である．

「無関心」「既知」は，伝達する情報を減らすための意図である．興味のないトピックや既に詳細を知ってる情報を長々と伝えるシステムは利便性のあるものではない．そういった情報をユーザの反応に応じて省くことで利便性があり，継続して利用したくなるようなシステムになると考えられる．「無関心」は対話を行うトピックに対してユーザが興味を持っていないという意図である．ユーザの興味の無い情報を伝達することは無駄だと考え，この意図を認識した場合は他のトピックへ遷移することが考えられる．「既知」はシステムが伝える情報がすでにユーザが有する情報であることを示す意図である．既知情報を含むトピックに関しては詳細な情報を省き伝達することで，伝達効率を上げることができると考えられる．



図 1: 収録環境

3. データセット構築

2. で提案した発話意図に関するラベルを付与したデータセットの構築について述べる．発話意図のあらわれ方は，人同士で対話した場合と，システムと対話した場合とでは異なることが考えられる．本研究ではシステムと対話した際の現象をとらえるため，実際に [高津 18a] のシステムを用い，24 人の被験者が対話を行った際の音声データを収集した．さらに．収集したデータを 10 人のアノテータにより 2. で提案した発話意図に関するアノテーションを行った．

3.1 音声収録方法

人がシステムと対話した際の現象をとらえるため，実際に [高津 18a] のシステムを用い，被験者が対話を行った際の音声データを収集した．現在 2103 対話分の収録が完了している．音声データ収集には 24 人（男：12 人，女：12 人）の被験者がシステムとの対話を行い，その際の音声データを随時収録した．

3.2 アノテーション

アノテーションは，図 2 に示すアノテーションツールを作成し，10 人のアノテータで行った．このツールで用いるラベルは 2. で提案した発話意図を示すことができるかどうかと，このツール自体の使いやすさを考慮し設計した．具体的には，ラベルを現象，行為，内部状態の 3 つのレイヤに分け内部状態には 5 つの軸を設定した．

現象では，「発話の完了」を示す終了，「発話が途切れていたり，遮られている」ことを示す継続，「音声のない区間」を示す無音・雑音，「システム対話と関係のない音声」を示す無効，の 4 ラベルを付与する．

行為では，発話行為として分類される相槌，ターンをとる，返答，質問，情報開示，リアクションの 6 ラベルについて当てはまるものを選択する．

内部状態では「システム発話を受け取れているかどうか」「システム発話の内容に疑問を感じているかどうか」「対話内容に興味を抱いているかどうか」「システム発話の内容に好意的かどうか」「システム発話の内容が予期しえたかどうか」の 5 つの軸についてそれぞれ 3 段階で評価を行う．



図 2: 作成したアノテーションツール

表 1: 各ラベルに対応する発話意図とそれに対して想定されるシステムアクション

ラベル		発話意図	システムアクション
質問		質問	質問応答
継続		待機要求	傾聴
ターンをとる			
意外		補足要求	補足説明
相槌 or リアクション	＆懐疑		
	＆不受理	反復要求	繰り返し
	＆無関心	無関心	他トピックへ移行
当然		既知	詳細説明省略

アノテーションのルールは，1. 現象のレイヤで「終了」のラベルを付けた場合には行為，内部状態のレイヤに関しても必ずラベルを付与する．2.「継続」のラベルを付けた場合には分かれば行為，内部状態のレイヤに関してラベルを付与する．3.「無音・雑音」「無効」のラベルを付けた場合には行為，内部状態のレイヤにはラベルを付与しない．とした．

各レイヤでの選択は独立に行われ，内部状態のレイヤでは軸ごとにも独立に評価が行われる．各ラベルに対応するユーザ意図とそれに対して想定されるシステムアクションを表 1 に示す．

対象音声の分割は WebRTC vad（aggressiveness=2）[*1] を用い，即時的な反応を対象とするため 1.5 秒以下の音声を対象とした．また [森 05] において前の文脈が対象発話の評価に影響を及ぼすことが示されている事を考慮し，評価の際の材料として「対象音声の 3 秒前から再生」「対象音声の 5 秒前から再生」を作成した．さらに 5 秒前の音声からのみでは評価できなかった場合を考慮して,「対話全体」を聞けるようにし，対象発話が全体の中で何秒から始まるかの記述も行った．

他にも，アノテータ間の意見の相違をなくすためアノテーションの試用期間を設け，不一致が多かったものに対する意見のすり合わせ等も行った．

アノテーションを行う際には，アノテータ同士の一致度を求めるため．200 発話分に対しては全員がラベルを付与し，その一致度を求めた．一致度の計算にはフライスのカッパ係数（Fleiss' Kappa）を用い,「現象」「行為」ではそれぞれどのラベルを選択したか,「内部状態」では各軸毎にどのラベルを選択したかで一致度を求めた．

*1 https://github.com/wiseman/py-webrtcvad

表 2: 各ラベルの数

ラベル名		数	ラベル名		数
現象	終了	9240	内部状態	受理	8511
	継続	271		不受理	396
	無音・雑音	930		納得	3426
	無効	215		懐疑	2337
行為	相槌	3671		興味	2287
	ターンをとる	88		無関心	1156
	返答	233		当然	301
	質問	2161		意外	1594
	情報開示	366		好感	844
	リアクション	4950		嫌悪	542

表 3: 各発話意図に対応するラベルとそのデータ数

ラベル		発話意図	データ数
質問		質問	2161
継続		待機要求	312
ターンをとる			
意外		補足要求	2055
相槌 or リアクション	＆懐疑		
	＆不受理	反復要求	192
	＆無関心	無関心	1145
当然		既知	301

3.3 結果と考察

データセットの構築により得られた結果について述べる．現在，10420 発話に対するラベル付与が済んでおり，集まったデータについて各ラベルの数は表 2 の通りである．この数とは，アノテータが重複している場合は一人でも付与しているラベルがあればカウントしたものである．

この中から表 1 で示した発話意図の分類に当てはまるラベルの数は表 3 の通りである．また，アノテータ全員がラベルを付与した 200 発話に対する一致度 (Fleiss' Kappa) は図 3 のようになった．

提案した音声対話によりまとまった情報を伝えるシステムにおいて理解するべきユーザの発話意図において，最も表出度合が高かったものは「質問」が 2161 発話，次いで「補足要求」が 2055 発話,「無関心」が 1145 発話であり，他の意図に関しては表出度合が低かった．また，一致度についてみると現象や行為のラベルについては κ 係数が約 0.8 とほぼ完全な一致に近い値であったが，内部状態のラベルについてはどれも 0.4 を下回り，アノテータの評価一致度は低かった．

内部状態のラベル数は,「システム発話を受け取れているかどうか」の軸に関するラベルの付いた発話数が 8907 と「終了」のラベルが付いた発話のほぼすべてにラベルがついていたが,「システム発話の内容に疑問を感じているかどうか」に関するラベルの付いた発話数は 5763,「対話内容に興味を抱いている

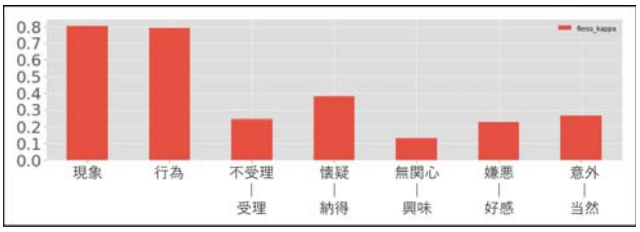


図 3: アノテータ同士のラベル一致度（Fleiss' Kappa）

かどうか」に関するラベルが付いた発話数は 3443,「システム発話の内容に好意的かどうか」に関するラベルが付いた発話数が 1386,「システム発話の内容が予期しえたかどうか」に関するラベルが付いた発話数が 1895 と「システム発話を受け取れているかどうか」の軸以外は内部状態のラベルの付いた発話が少ないことがわかる．このことから，現状のシステムとの対話において話者はそれほど内部状態の表出を行っていないことが分かる．

使用したシステムのニュースを伝えるという特性上,「質問」「補足説明」といった発話意図の表出が増えることは容易に想像でき，それは表出度合にも表れている．中でも「質問」に関してはラベルの一致度も高いため，システム実運用においても有用な意図であり，データセットとしての質も高く学習などに用いることができると考えられる．しかし,「補足要求」についてはその構成するラベルの一致度が低いためシステム実運用の際に判断のブレが強く出てくることが考えられる.「無関心」については，話者の内部状態表出が少ないことから，棒読みに聞こえることが多く，第三者からすると無関心に聞こえることが多かったのではないかと考えられる．また，ラベルの一致度については「補足説明」と同様のことが言える．他の意図に関しては表出度合が低く，構成するラベルの一致度も低い．

このように話者のフィードバックが少ない対話では，従来感情認識や意図認識として行われてきた研究の成果が生かせず，効率的な対話の実現ができない．また，雑談対話のようなタスクにおいても，システムの対話戦略をユーザの意図に沿わずに決定することになり，継続的に使われるシステムの実現は難しい．そのため，ユーザが内部状態を表出するためには現状のシステムで改善するべきものは何か，どのような機能を加えるべきなのか，そもそもシステムという時点で内部状態は表出しないのか，といった検討が必要である．

また，今回作成したコーパスの中でもデータ量が多く，アノテータ間の一致度も高かった発話意図「質問」については各発話が「質問」かどうかの識別実験も行った [高津 18b]．この結果，識別率約 95 %と非常に高い識別性能を得ることができた．

4. おわりに

実会話システム利用時における，システム発話に対するユーザのフィードバックの現象を分析するためのコーパス収集について検討を行った．現在我々が作成しているニュース伝達のための会話システム [高津 18a] を用いて，ユーザが実際にシステムとの会話時にどのようなフィードバックを行い，システムはどのようにそれを利用しうるかを調査した結果，現状のシステムでは発話意図「質問」においては出現頻度は高く，アノテータ間の一致度も高いことが分かった．一方で，その他の発話意図についてフィードバックは陽に得られないことが分かった．今後の課題として話者のフィードバックを陽に得られるようなシステムの作成があげられる．

参考文献

[Fujie 06] Shinya Fujie, Riho Miyake, Tetsunori Kobayashi: Spoken Dialogue System Using Recognition of User's Feedback for Rhythmic Dialogue, in Proceedings of Speech Prosody (2006)

[Kobayashi 13] Tetsunori Kobayashi, Shinya Fujie: Conversational Robots: An Approach to conversation protocol issues that utilizes the paralinguistic information available in a robot-human setting, Acoustical Science and Technology, Vol. 34, pp64-72 (2013)

[益永 14] 益永祐吾, 川端豪: 発話意図判定における文型・韻律・音色の相補的効果, 情報処理学会研究報告音声言語情報処理（SLP）, Vol.2010-SLP-80, No.14 (2014)

[牧野 17] 牧野孝史, 三浦寛也, 竹川佳成, 平田圭二: 複数人対話における話者交替現象分析のためのコーパスの作成, 人工知能学会言語・音声理解と対話処理研究会（SIG-SLUD）, pp.3-8 (2017)

[荒木 17] 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, 人工知能学会言語・音声理解と対話処理研究会（SIG-SLUD）, pp.20-25 (2017)

[井上 17] 井上昂治, Divesh Lala Pierrick Milhorat, 高梨克也, 河原達也: 潜在キャラクタモデルによるリアルタイム対話エンゲージメント推定, 人工知能学会言語・音声理解と対話処理研究会（SIG-SLUD）, pp.78-83 (2017)

[Inoue 16] K. Inoue, P. Milhorat, T. Zhao, T. Kawahara: Talking with ERICA, an autonomous android, in Processings of SIGDIAL Conference, pp.212-215 (2016)

[長澤 17] 長澤史記, 石原卓弥, 岡田将吾, 新田克己: ユーザーの態度推定に基づき適応的なインタビューを行うロボット対話システムの開発, 人工知能学会言語・音声理解と対話処理研究会, pp.72-77 (2017)

[高津 18a] 高津弘明, 福岡維新, 藤江真也, 林良彦, 小林哲則: 意図性の異なる多様な情報行動を可能とする音声対話システム, 人工知能学会論文誌, Vol.33, No.1 (2018)

[森 05] 森大毅, 相澤宏, 粕谷英樹: 対話音声のパラ言語情報ラベリングの安定性, 日本音響学会誌 61 巻 12 号, pp.690-697 (2005)

[高津 18b] 高津弘明, 横山勝矢, 本田裕, 藤江真也, 林良彦, 小林哲則: 会話によるニュース記事伝達のための発話意図理解, 第 32 回人工知能学会全国大会論文集 (2018)

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

機械学習技術を用いた交渉エージェントのための自動交渉シミュレータ Jupiter の提案

A proposal of automatic negotiation simulator Jupiter for negotiation agents using machine learning.

福井智哉 *1　伊藤孝行 *1
Tomoya Fukui　Takayuki Ito

*1名古屋工業大学大学院 情報工学専攻
Nagoya Institute of Technology, Department of Computer Science

The purpose of this paper is to propose Jupiter, a new environment for automated negotiations in which we can easily create agents that is able to use machine learning. Genius is cited as a prior study of the environment where automated negotiation can be simulated. It provides an environment for automated negotiations that aim to solve multi-issue negotiation problems. In the field of automated negotiation, it is expected that agreement results are optimized by machine learning. However, it is difficult to use Genius to simulate automated negotiations with agents using machine learning, because the past negotiations information provided by Genius is insufficient. As above, in this paper we propose Jupiter as a new automated negotiation environment in which we can easily create agents that is able to use machine learning. In addition, we compare Jupiter with Genius and show the superiority of Jupiter.

1. はじめに

本論文の目的は，機械学習技術を用いたエージェントの作成が容易な交渉シミュレーション環境として Jupiter を提案することである．

交渉は互いの利害の調整のために，社会において必要不可欠な行為である．自動交渉を模擬的に行う環境の先行研究として，Genius が挙げられる．Genius は複数論点交渉問題における，エージェントによる自動交渉シミュレーション環境を提供し，自動交渉の合意結果から自動交渉エージェントの性能を定量的に測ることができる．

近年，ハードウェア技術の向上に伴い，機械学習技術の発展が目覚ましい．しかし，Genius では繰り返し交渉を行う際に，過去の交渉に関する情報が一部分しか取得できないため，機械学習技術を用いる際に必要となる学習データが不足しており，機械学習技術を応用したエージェントの自動交渉実験が困難である．以上の点から，本論文では機械学習技術を用いたエージェントの作成が容易な環境として Jupiter を提案する．また提案システムである Jupiter と Genius で比較と評価実験を行うことにより，Jupiter の有用性を示す．

2. 関連研究

2.1 ANAC

概要

本論文で提案する Jupiter では，既存研究 ANAC[Baarslag 2016] を参考に，扱うことのできる交渉プロトコルと交渉ドメインを設定した．交渉プロトコルとしては Stacked Alternating Offers Protocol(SAOP)[Reyhan 2017] を採用している．交渉プロトコルとは，交渉の進め方や各エージェントの行動の定義をまとめた概念である．SAOP では，交渉参加エージェントが，交互に合意候補案の中から 1 つ提案を行うことにより，交渉を進行させる．交渉時間の種類にはターン制と時間制の 2 種類が存在し，各エージェントの行動には Offer，Accept および EndNegotiation の 3 種類存在する．

本節では Jupiter で取り扱う交渉プロトコルと交渉ドメインについて述べる．

交渉時間の種類

交渉の制限時間についてターン制と時間制の 2 種類述べる．

1. ターン制
   全参加エージェントの合意または EndNegotiation が発生したとき，もしくはあらかじめ決められたターンまで合意が得られかったときまで交渉を続ける．

2. 時間制
   全参加エージェントの合意または EndNegotiation が発生したとき，もしくはあらかじめ決められた制限時間まで合意が得られかったときまで交渉を続ける．

各エージェントの行動

各エージェントが取ることの出来る行動について述べる．エージェントの行動には Offer，Accept 及び EndNegotiation の 3 種類が存在する．各交渉参加エージェントは，自分の手番が回ってくる度に，3 種類の行動から 1 つ選ぶ．また，交渉参加エージェントは，交渉参加エージェントの順番に関係なく，他エージェントの行動を公開情報として全て認知することができる．しかし，行動を起こすには，自分の手番が回ってくるまで待たなければならない．

1. Offer
   合意案候補から 1 つ選び，提案する．他エージェントからの Offer をされたあとに，自エージェントが Offer を行なった場合，その他のエージェントの Offer は自動的に却下されたことになる．

2. Accept
   最近の他エージェントからの Offer を受け入れることを意味する．あるエージェントの Offer を，提案したエージェント以外の全エージェントが連続で Accept した場合，その交渉は合意に至ったとして判定され，交渉が終了する．

3. EndNegotiation
   交渉の強制的な終了を意味する．他のエージェントの行動に関わらず，交渉を終了することができる．一般的に他エージェントが交渉において，割引効用が大きいにもかかわらず，まったく妥協しない場合に有用な選択肢となる．

交渉ドメイン

ANAC で取り扱われる交渉ドメインについて述べる．交渉ドメインには，交渉結果における，各エージェントが取得する効用値に関する情報が記されている．

交渉ドメインに記述されている内容を説明する前に，合意案候補 (bid) を数式で示す．$\boldsymbol{s}$ を合意案候補，n を論点の数，I_i

連絡先: 福井智哉，名古屋工業大学工学部情報工学科，fukui.tomoya@itolab.nitech.ac.jp

を i 番目の論点で取りうる選択肢の集合と定義すると，各合意案候補 (*bid*) はの式 (1) により定義される．

$$\boldsymbol{s} = \{(s_1, s_2, \ldots, s_n) | s_i \in I_i, i = 1, 2, \ldots, n\} \quad (1)$$

式 (1) をふまえて，交渉ドメインに記述されている各合意案候補の効用値，留保価格及び割引効用の 3 つを説明する．

1. 各合意案候補の効用値
   各合意案候補の効用値 $U(\boldsymbol{s_k}; \boldsymbol{w})$ は式 (2)，式 (3) 及び式 (4) を用いて表される．

$$U(\boldsymbol{s}; \boldsymbol{w}) = \sum_{i=1}^{n} w_i \cdot eval(s_i) \quad (2)$$

$$\sum_{i=1}^{n} w_i = 1 \quad (3)$$

$$eval(s_i) = \frac{value(s_i)}{\arg\max_{s_i \in I_i} value(s_i)} \quad (4)$$

   w_i は i 番目の論点における重みを表し，$0 \leq w_i \leq 1 (1 \leq i \leq n)$ である．$value(s_i)$ は I_i において s_i を選んだ場合の効用値を表す．$w_i (1 \leq i \leq n)$ と $value(s_i) (s_i \in I_i)$ の値は交渉ドメインに記されている．補足として，式 (4) から

$$0 \leq eval(s_i) \leq 1 \quad (5)$$

   である．式 (2)，式 (3) および式 (5) より

$$0 \leq U(\boldsymbol{s}; \boldsymbol{w}) \leq 1 (\forall \boldsymbol{s})$$

   である．

2. 留保価格
   EndNegotiation や合意が得られずに交渉が終了した際に得られる効用値を表す．留保価格を r，合意失敗を e とすると，式 (6) で定義される．

$$U(e) = r \quad (6)$$

3. 割引効用
   交渉が終わった際に，交渉にかかった時間に応じて，各エージェントが得られる効用値を割り引くために使われる値である．各エージェントが得られる割引済みの効用値を U_d，割引効用を d，正規化された交渉の経過時間を $t (0 \leq t \leq 1)$ とすると

$$U_d(\boldsymbol{s}, t; d) = U(\boldsymbol{s}) \cdot d^t \quad (7)$$

$$U_d(e, t; d) = U(e) \cdot d^t \quad (8)$$

   と表される．

2.2 Genius

概要

既存の自動交渉シミュレーション環境に関する先行研究として，Genius[Lin 2014] について述べる．汎用的な自動交渉エージェントに関する研究を支援し，Java API を用いて交渉エージェントの開発や自動交渉シミュレーションを行うことができる．Genius の主な機能は，以下の 3 つである．

1. 交渉問題を定義する交渉ドメインファイルの作成
2. 自動交渉エージェントによる交渉シミュレーション
3. 交渉の過程と結果の記録と解析

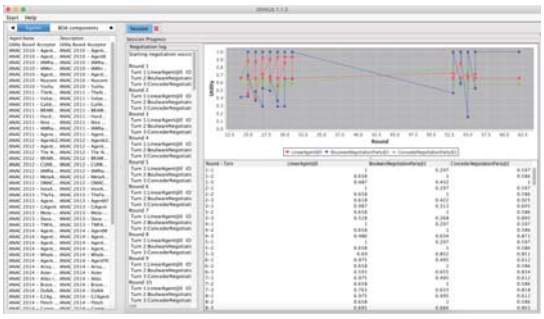


図 1: Genius の実行例

Genius はパレートフロント，ナッシュ解，およびカライ・スモロデンスキー解といった合意案候補点をグラフ上に描画するため，自動交渉の結果を視覚的に分析することができる．また特定の交渉問題ではなく，本章で説明した交渉プロトコルを採用することにより，汎用的な自動交渉全般に関する研究を支援している．よって Genius は，異なる交渉戦略をもつエージェント同士を，客観的な指標を用いて評価することが可能であり，自動交渉エージェント及び交渉戦略の研究において重要な役割を果たしている．Genius の実行例として，図 1 を添付しておく．

Genius の問題点

Genius を用いた自動交渉エージェントの研究開発では，さまざまな分野の技術が用いられている．森ら [森 2015] の提案するエージェントでは，ゲーム理論の観点からエージェントを作成している．Williams の提案するエージェント [Williams 2011] では，ガウス過程という機械学習技術を用いたエージェントを作成している．

一方，昨今のエージェントの研究分野では，深層学習技術を用いた研究が注目されている．2017 年には Lewis らが自然言語により自動交渉を行うエージェントの研究 [Lewis 2017] を行なっている．今後も深層学習などの機械学習技術を用いたエージェントの研究が期待される．しかし，Genius は Java 実装である点や，2010 年より開発されているという経緯から，昨今の深層学習技術を用いたエージェントの開発に向けて作成されていない．一般的に深層学習技術を実装する場合には，Python で実装されることが多い．

3. 自動交渉シミュレーション環境 Jupiter

3.1 Jupiter の概要

Jupiter は本論文で提案する自動交渉シミュレーション環境である．Jupiter は 2.1 章で述べた交渉プロトコルと交渉問題を扱うことができる．またプロセス間通信を行うことにより，Genius 上で動作する自動交渉エージェントを Jupiter 上で動かすことができる．本章では，Jupiter の実行方法と動作例を示す．

3.2 Jupiter で自動交渉を行うまでの流れ

本章では，本論文で提案する自動交渉シミュレーション環境 Jupiter により，自動交渉を行う際の処理の流れを説明する．

1. ユーザが交渉の長さ，交渉ドメインファイルおよび交渉参加エージェントを指定する．
2. Jupiter が交渉ドメインファイルを読み込む．またエージェントを呼び出し，エージェントに交渉ルール，交渉ドメイン情報を渡す．
3. 交渉を実行する．

また交渉を実行する際のパラメータ次第で，交渉過程を動的に表示させることや交渉履歴を json 形式で保存することができる．

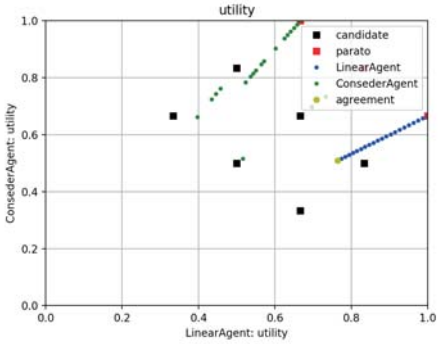


図 2: Jupiter の実行例

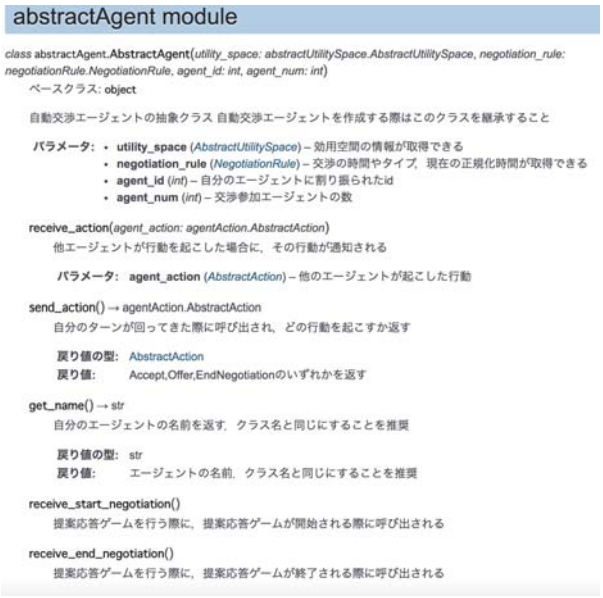


図 3: abstractAgent

Jupiter の実行例として，図 2 を載せる．緑色の点が ConcederAgent の提案した合意案候補であり，青色の点が Linear エージェントの提案した合意案候補である．横軸が Linear エージェントの獲得できる効用値であり，縦軸が Conceder エージェントの獲得できる効用値である．また黒色の四角形と赤色の四角形が合意案候補点である．

3.3 Python による自動交渉エージェントの作成

Jupiter でエージェントを実装するには，図 3 に表示されている，自動交渉エージェントの抽象クラスである abstractAgent クラスを継承したクラスを作成すればよい．継承した子クラスについて，親クラスの各メソッドをオーバーライドすることにより，エージェントの交渉戦略を定義できる．

ここでエージェントを実装する際の，Genius との差異として，receive_start_negotiation メソッドと receive_end_negotiation メソッドがあることを挙げておく．繰り返し提案応答ゲームにおいて機械学習技術を応用する際に，提案応答ゲームと提案応答ゲームの間に，パラメータの調整などの処理を行いたい場合がある．receive_start_negotiation メソッドは提案応答ゲームの開始時に，receive_end_negotiation メソッドは終了時に呼ばれる．そのため，提案応答ゲームが終了した際の後処理を実装したい場合，Jupiter では容易に行うことができる．

表 1: Genius と Jupiter によって実装された Linear エージェントの比較

シミュレーション環境	コードの行数	コードの文字数
Genius	57	120
Jupiter	36	83

4. Jupiter と Genius の比較

4.1 概要

本章では Jupiter と Genius について，定性的な評価と定量的な評価を行うことにより，その差異を示す．

まず Genius 実装のエージェントと，Jupiter 実装のエージェントについて，そのコードの長さを比較する．次に Genius によって行われる自動交渉と，Jupiter によって行われる自動交渉で，実行にかかる時間を比較する．実行にかかる時間が短ければ，自動交渉エージェントの実験にかかる時間を削減することができる．よって，自動交渉シミュレータにとって重要な課題である．次に定性的な評価について比較を行う．最後に，本章で行なった比較のまとめと考察を行う．

4.2 自動交渉エージェントのコードの長さの比較

本実験では，Linear エージェントを Jupiter 環境と Genius 環境で動くようにそれぞれ実装し，そのコードの行数と文字数を定量的に比較する．以下，Linear エージェントについて解説する．

エージェントを作成する際には，交渉戦略を設計する必要がある．自動交渉エージェントはその交渉戦略に基づいて受容 (Accept)，提案 (Offer) 及び交渉打ち切り (EndNegotiation) から 1 つ選択する．特に他エージェントからの提案を受容するかどうか，に関する戦略を Acceptance Strategy と呼ぶ．Acceptance Strategy を設計する際に，自身が獲得できる効用値からその判断を行うエージェントが多く存在する．その判断の手法には，提案された内容の効用値を入力として，受容するかどうかの 2 値を出力する，譲歩関数を用いる手法が存在する．文献 [C.Sierra 1999] では，譲歩関数に交渉の経過時間 $t(0 \leq t \leq 1)$ を用いたエージェントを提案している．

$$T(t) = 1.0 - t^{\alpha} \tag{9}$$

具体的には式 (9) で定義される譲歩関数 $T(t)$ を提案している．式 (9) で表される譲歩関数を持ち，特に $\alpha = 1$ となるような譲歩関数を持つエージェントを Linear エージェントと呼ぶことにする．Linear エージェントは，譲歩関数の出力する値以上の効用値を持つ提案または受容を行う．

比較の結果を表 1 にまとめた．一般的に Python 言語は Java 言語よりも，コードを簡潔に記述できるとされている．本比較の結果においても，定量的なコードの記述量による比較から，コードの簡潔さを示すことができた．

4.3 実行速度の比較

本実験では Linear エージェント，Conceder エージェント及び Boulware エージェントによる 3 者間交渉を行う．Conceder エージェントは式 (9) において $\alpha = 0.5$ となるような譲歩関数を持つエージェントである．Boulware エージェントは式 (9) において $\alpha = 10$ となるような譲歩関数を持つエージェントである．

交渉の設定としてはターン制で，180 ターンを上限とした．また，交渉ドメインファイルとして，ANAC2016 で用いられた triangularFight，KDomain 及び NewDomain を用いる．交渉ドメインファイルを選んだ基準として，合意案候補の数を考えており，合意案候補の数が少ないもの，中くらいのもの及び多いものを選んだ．それぞれの効用空間の大きさは表 2 を参照されたい．

交渉参加エージェントの順列が ${}_3P_3 = 6$ 通りであり，それぞれ 10 回ずつ自動交渉を行うため，各々の効用情報について合計 $6 \times 10 = 60$ 回ずつ自動交渉を行う．

表 2: 実験で用いる効用空間の大きさ

ドメイン名	論点数	合意案候補の数
triangularFight	2	9
KDomain	5	1280
NewDomain	7	35840

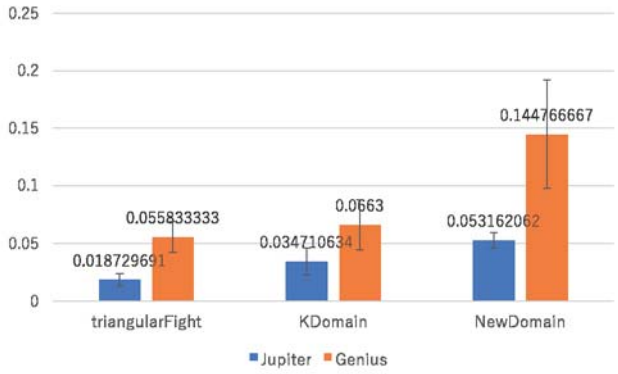


図 4: 平均交渉時間とその分散

各交渉ドメインファイルにおける 60 回の，平均交渉時間とその分散を示したものが図 4 である．効用空間の大きさに関わらず，Jupiter の方が自動交渉にかかる時間が短いことが示された．Jupiter では一部の処理を Cython で実装することにより，処理の高速化を行なっている．

4.4 定性的な評価による比較

本節では定性的な評価により比較を行う．

Genius ではパレートフロント，ナッシュ解，及びカライ・スモロデンスキー解などの分析を自動で行う．Jupiter では現在，パレートフロントのみを提供している．ナッシュ解及びカライ・スモロデンスキー解の分析については実装中である．

Genius は Java 言語による実装であり，Jupiter は Python 言語による実装である．そのため，Jupiter は scikit-learn や tensorflow などのライブラリを用いることにより，機械学習技術を応用したエージェントの実装が容易である．

また Genius では GUI 上で自動交渉が行えるが，Jupiter では基本的に Jupyter Notebook というシェル上での実行を推奨している．Jupyter Notebook 上での実行が可能なため，実行ログの視覚的な確認が容易であり，エージェントの各種パラメータの調整が容易である．

4.5 比較のまとめ

以下，本章で行った比較評価に関して，まとめる．簡単にまとめたのが表 3 である．

Genius と提案する自動交渉シミュレーション環境 Jupiter について比較評価を行った．定量的な比較項目として，交渉の実行時間と，エージェントを実装する際のコードの記述量を比較した．定性的な評価として，実装言語や使用方法を比較した．

4.6 考察

考察をもって，本章の終わりとする．

Python 言語はインタラクティブな実行が可能である，スクリプト言語である．よって，一般的には Java 言語の方が実行速度は速いと言われている．しかし，Genius は 2010 年から現在に到るまで，長期間運用されてきた自動交渉シミュレータである．内部が最適化されていないために，動作が遅いと考えている．一方で，長期間運用されてきたことからその研究分野はゲーム理論，心理学及びマルチエージェントなど多岐に渡る．例えば 4.4 節で述べた，ゲーム理論の枠組みによる，合意結果の分析は大変興味深い．また ANAC で毎年用いられていることから，留保価格や割引効用が取り入れられるようになり，より現実の交渉問題に近い交渉設定を扱えるようになって来ている．しかし現状ではまだ現実の交渉問題を扱うことは難しいと感じている．よって，交渉プロトコルや交渉ドメインについての研究を今後の課題としたい．

表 3: Genius と Jupiter の比較まとめ

比較項目	Genius	Jupiter
実装言語	Java	Python3
CUI	△	○
交渉の実行時間	△	○
エージェントのコード記述量	△	○

5. おわりに

本研究では，自動交渉シミュレーション環境 Jupiter を提案した．SAOP に基づく自動交渉シミュレーションを Python 言語により行うことができる．また実験により，自動交渉にかかる時間とエージェントのコード記述量において，先行研究である Genius よりも短いことを示した．しかし，Genius と比較して，交渉ドメインファイルの作成機能や GUI の実装など未実装である機能が存在することが問題点として挙げられる．今後の課題としていきたい．

参考文献

[Williams 2011] Williams, Colin R., et al.: ”Using gaussian processes to optimise concession in complex negotiations against unknown opponents.”, IJCAI Proceedings-International Joint Conference on Artificial Intelligence. Vol. 22. No. 1. (2011).

[Reyhan 2017] Aydoan, Reyhan, et al.: ”Alternating offers protocols for multilateral negotiation.”, Modern Approaches to Agent-based Complex Automated Negotiation. Springer International Publishing, 153-167. (2017).

[Lin 2014] Lin, Raz, et al.: ”Genius: An integrated environment for supporting the design of generic automated negotiators.”, Computational Intelligence 30.1, 48-70. (2014).

[森 2015] 森顕之, and 伊藤孝行.: ”二者間交渉問題における進化的安定戦略を考慮した交渉戦略の提案.”, 研究報告知能システム (ICS) 2015.19, 1-8. (2015).

[Baarslag 2016] T. Baarslag, R. Aydog an, K. V. Hindriks, K. Fuijita, T. Ito, and C. M. Jonker.: The automated negotiating agents competition 2010-2015., AI Magazine(2016).

[Lewis 2017] Lewis, Mike, et al. : ”Deal or no deal? end-to-end learning for negotiation dialogues.” arXiv preprint arXiv:1706.05125 (2017).

[C.Sierra 1999] Sierra, Carles, Peyman Faratin, and Nick R. Jennings.: ”A service-oriented negotiation model between autonomous agents.” Collaboration between human and artificial societies. Springer, Berlin, Heidelberg, 201-219, (1999) .

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

環境問題を対象とした多言語ディスカッションの対話分析

An analysis of multilingual communication for solving global environmental issues

廖 育琦[*1] Yuqi Liao　王 博[*1] Bo Wang　家入 祐也[*1] Yuya Ieiri　中島 悠[*2] Yuu Nakajima　菱山 玲子[*1] Reiko Hishiyama

[*1] 早稲田大学創造理工研究科経営システム工学専攻 Graduate School of Creative Science and Engineering, Waseda University

[*2] 東邦大学理学部情報科学科 Department of information Science, Toho Univercity

Because of their transnational nature, environment problems are difficult-to-solve, global challenges. The goal of our research is to examine ways to clarify common features, points of similarity, and points of difference in people's thinking in order to promote understanding and co-ownership of environmental problems between people who speak different languages. Our research begins with facilitating communication between participants about environmental problems through ICT. This communication log will consist of a sequence of dialogues in two languages, the discourse analysis of which will require significant cost. However, by reducing and analyzing techniques of expression through consolidating each person's thoughts post-communication and providing illustrated diagrams of the dialogue sequences the speakers experienced, we demonstrate that we can summarize each person's points and assist in mutual understanding.

1. 始めに

昨今，地球環境問題が深刻化しメディアでも多く取り上げられ，深刻な社会問題となっている．中でも中国は，深刻な空気汚染と水質汚染で苦しんでいる．これらの環境問題は一般に越境性を伴い，一国のみでは解決が困難であるため，問題解決に向けて地球規模の対話が求められている．こうした場面では，異言語で多様な人々の意見を相互に交換できるコミュニケーション環境が必要となる．

本研究では，こうしたコミュニケーション環境で，相互の意見や議論に対する共通点，類似点や相違点を抽出し，それを比較するために有効な方法を検討する．研究では異言語の対話系列を図解としての表現技法に置き換える方法を適用し，対話系列と図解を併用して思考内容の比較のための情報を得る方法を提案し，意見の共通点や相違が抽出できることを示す．

2. 関連研究

これまで機械翻訳サービスを介した多言語のコミュニケーションに関する研究が多く行われてきた．照井ら[1]はビジネススクールの異文化にまつわる経営ケース教材を利用し，言語グリッド[2]を介した多言語コミュニケーションの対話内容を分析し，文化差を抽出した．鈴木ら[3]は専門知識が含まれる日本酒の醸造工程の知識伝達によって機械翻訳を介した知識伝達の品質に関する議論を行った．これらの一連の研究に対して，本研究は特に対話とその図解を併用する点に特徴があり，意見の共通点・類似点，相違点が把握しやすくなるよう配慮している．

3. 提案

研究では，異言語・異文化コミュニケーションによる対話においてその考え方の共通点・類似点や相違点を相互に理解するための方法として，対話系列を図解としての表現技法に置き換えて比較可能とし，相互の論点を集約することを提案する．提案ではICT環境でのチャット形式による対話に加え，思考・発想法として知られるマインドマップ[4]を活用し，対話から得られた情報をマインドマップに整理することで，その論点を整理する．これにより，対話系列だけでは把握できない相互の意見を思考の構造ベースで比較できる．

Name: Yuqi Liao
Affiliation: Graduate School of Creative Science and Engineering, Waseda University
Address: 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan

4. 実験概要

本研究では，日本人と中国人が中国の環境問題の原因 3 点を巡り議論を行う実験を遂行した．議論参加者の 2 名は，提示された中国の環境問題 3 つの原因：火力発電，工場の排気，自動車排気について議論しながら優先順位をつけ，最も重要な原因から解決策を議論する．この作業を繰り返し，2 番目に重要な原因，3 番目に重要な原因，のように順次議論を行う．議論が終わった後，各自が議論内容と自分の考えを結合して紙上に整理し，自分なりのマインドマップを描く作業を行った．

実験は 5 回行ない，うち 3 回は，日本人 1 名と中国人 1 名を 1 組とし，日中翻訳サービスを介した異言語の議論を行った．残り 2 回は対照実験として，日本人 2 名を 1 組，中国人 2 名を 1 組とし，それぞれ同言語（母語）による実験を行った．

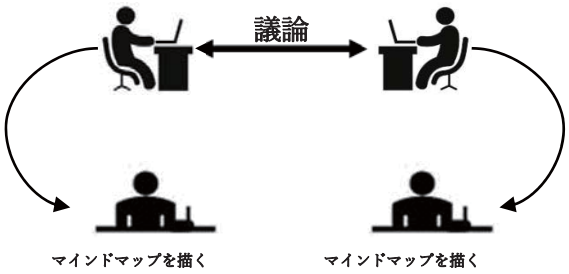


図 1：実験手順の概要

5. 結果と考察

議論の時間は，三つの原因とも 8 分間を費やした．議論された三つの原因に関し，グループ別でマインドマップに記入されたキーワードを集計した結果を図 2 に示す．三つの原因の中，グループ 1 とグループ 3 は火力発電に関するキーワードが最も多い．グループ 2 は工場の排気に関するキーワードが最も多い．実験では，グループ 1 とグループ 3 は火力発電が第一の原因であるとし，グループ 2 は工場の排気が第一の原因であるとして，

対話と議論が進行した. 順位が高い原因ほど, 利用されたキーワード数も多い. つまり, 三つの原因の中で, 優先順位が高い原因は, 時間制限内で, 単位時間あたりの伝達情報量がより多い結果となった. 残りの二つの原因には, ほぼ同じ数のキーワードを記入していた. これは残り二つの原因に対して, 同様な重みがつけられていることが考えられる. なお, 議論過程の意思決定はマインドマップ上にも反映されている.

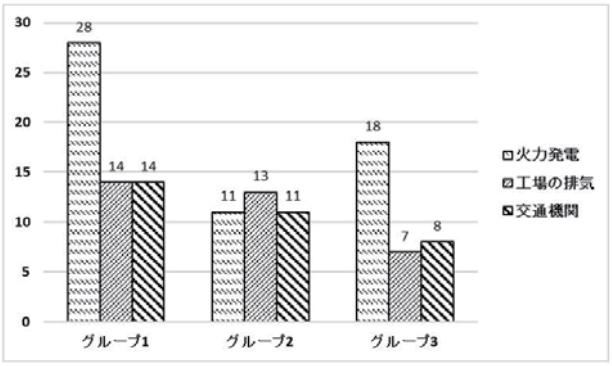


図 2:原因別のキーワードの数

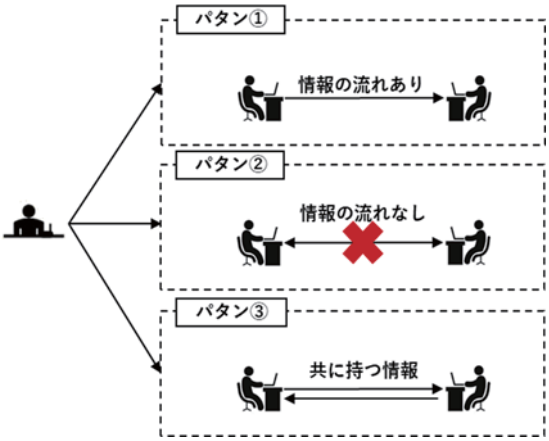
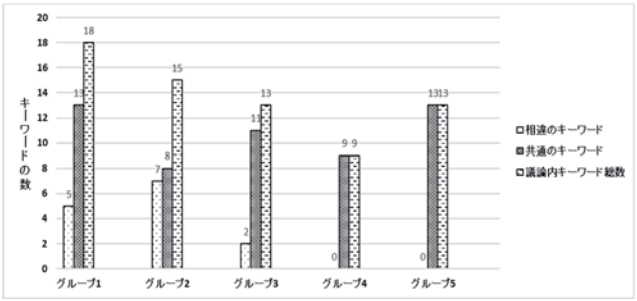


図 3:キーワードの分類

マインドマップ上に記述されることで得られたキーワードを集計し, パタン分類したものを図 3 に示す. パタン①では, 相手から情報を受け取り, マインドマップに記入したキーワードを集計する. これらは, 片方のみの意見や意見の相違点としてのキーワードを集計することを意味する. パタン②では議論していない独自の考えがマインドマップに記入されているケースを集計する. パタン③では, 共に情報を持ち, 議論中にそれを交換し,マインドマップに記入したキーワードを集計する. これらは, 意見の共通点としてのキーワードを集計することを意味する.

パタン①とパタン③は議論内で発生したキーワードであり, その集計結果を図 4 に示す. グループ 1, 2, 3 は, 日本人と中国人が議論に参加した異言語対話グループであり, グループ 4 は中国人と中国人, グループ 5 は日本人と日本人の同言語対話グループである.

日本人と中国人が議論した 3 組の異言語対話グループで, マインドマップに記入したキーワードを調べると, 情報の流れが見えるキーワードはそれぞれ 5, 7, 2 個であった. 同言語グループであるグループ 4 と 5 は, これが両方 0 になっている. このことから, 異言語(異国)間の参加者によるコミュニケーションで, 他国の情報が相互に伝達され理解される可能性を見出せる.



4:相違と共通キーワードの数

同国間コミュニケーションでは共通に持つ情報が多いとみられ, 情報の流れが表出せず, その流れの把握は容易ではない.

図 5 はマインドマップに記入されたキーワードを日本人グループと中国人グループに分け, さらにエネルギー, 政府政策, 法律, 経済, 教育の 5 つの種類に分類し集計した結果である. 環境問題の解決を巡り, 日本人と中国人はエネルギー, 政府政策, 法律, 経済, 教育の各分類に同程度の注目を行っている傾向が見えた. 特にエネルギーと経済について, 日本人と中国人は共に多数のキーワードを提示したことがわかる.

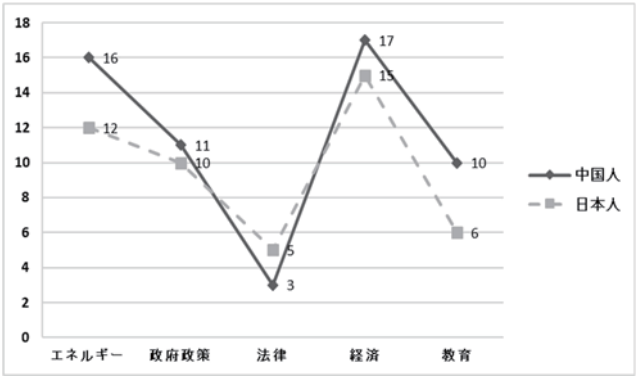


図 5: 日本人と中国人各分類におけるキーワード数

6. まとめと今後の課題

本研究は中国の環境問題に関する異言語コミュニケーションから日中間の情報の流れを分析し, 思考や理解の共通点・類似点や相違点の存在を明らかにすることを試みた. 今後の課題は, コミュニケーション前と後で, 考え方の変化を定量的に分析できる手法を見つけることである.

謝辞 本研究は中山隼雄科学技術文化財団研究助成の成果である.

参考文献

[1] 照井 賢治, 菱山 玲子:多言語コミュニケーション環境における異文化分析, ヒューマンインタフェース学会論文誌, 16(1), pp63-76 (2014).

[2] 言語グリッドポータルサイト: http://langrid.org/jp/ (閲覧日 2017/11/6).

[3] 鈴木 宏, 菱山 玲子:機械翻訳サービスを用いた専門知識伝達サービスの分析, 第 14 回情報科学技術フォーラム (FIT2015), 14(3), pp.69-76 (2015).

[4] Mind Mapping: www.tonybuzan.com/about/mind-mapping/ (閲覧日 2018/12/1).

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

議論スキームを用いた議論対話における論拠の自動生成

Premise generation based on argumentation schemes for dialogue systems

光田 航 Koh Mitsuda　東中 竜一郎 Ryuichiro Higashinaka　片山 太一 Taichi Katayama　富田 準二 Junji Tomita

日本電信電話株式会社 NTT メディアインテリジェンス研究所
NTT Media Intelligence Laboratories, Nippon Telegraph and Telephone Corporation

Towards creating an argumentative dialogue system, this paper proposes a method to create premises from a given utterance representing a conclusion. The proposed method utilizes Walton's argumentation schemes for generating premises from various inputs. The generated premises were evaluated by human subjects, and the results showed the validity of our method.

1. はじめに

自分の意見を述べたり，相手の意見に対する支持や不支持の論拠を述べる議論は，合意形成や意見交換を行うための重要な対話の一種である．近年，テキストに含まれる議論を構造化する argumentation mining と呼ばれる分野の研究が盛んに報告されており，議論が含まれるテキストから命題を抽出したり，抽出された命題間の支持や不支持の関係を推定する手法が提案されている [Lippi 16]．

argumentation mining の分野とは対照的に，ユーザと議論を行う議論対話システムの構築を試みた研究はまだ少ない [Rakshit 17, Higashinaka 17, Sakai 18]．議論対話システムが実現できれば，ユーザはシステムとの議論を通じて考えを整理したり，新たな気付きを得るといった有益な対話が可能になる．議論において，任意のユーザ発話に対して支持や不支持を表すシステム発話を生成するのは挑戦的な課題である．そのため，既存の議論対話システムは，特定の議論のテーマに関する議論の発話（結論を表す発話とその前提となる論拠を表す発話のペア）をあらかじめ人手で用意し，ユーザ発話に応じて適切な発話を選択することで応答を返す用例ベースの手法を採用している．

既存の議論対話システムの課題として，議論のテーマごとに人手であらゆる発話を網羅的に用意することは困難であるため，議論のテーマを増やしたり，一つのテーマで様々な話題に適切に答えられるようにするには限界がある．この課題を解決するためには，自動で議論に関する発話を用意できることが望ましい．argumentation mining の技術を用いて，特定のテーマに関する議論が含まれる Web テキストなどから命題を発話として獲得する抽出ベースのアプローチが考えられるが，Web テキストに頻出するテーマや命題以外には対応することができないため，獲得できる発話のカバレッジが十分とは限らない．

そこで，本稿では，argumentation mining で扱われる抽出ベースの手法ではなく，生成ベースの手法を用いて，結論となる命題の支持や不支持の論拠を表す命題を自動的に獲得する手法を提案する．生成する命題のカバレッジを高めるために，命題間の典型的な論理関係を整理した Walton の議論スキーム [Walton 13] を利用し，議論スキームごとに用意した生成用テンプレートと言語リソースを用いて命題の生成を行う．手法の有効性を確認するために，抽出ベースの手法との比較評価を実施した結果，提案手法はより幅広い命題に対して有効な命題を獲得できたことを報告する．

連絡先: 光田航，日本電信電話株式会社 NTT メディアインテリジェンス研究所，神奈川県横須賀市光の丘 1-1，239-0847，mitsuda.ko@lab.ntt.co.jp

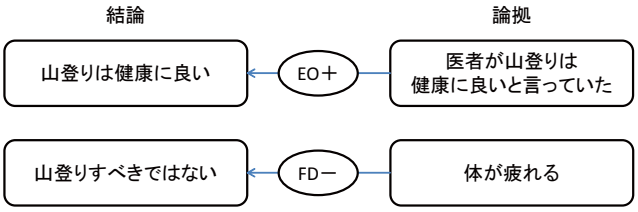


図 1: 議論スキームの例

2. 議論スキームを用いた論拠の生成

図 1 に議論スキームの例を示す．議論スキームとは，結論とその論拠の間に成り立つ典型的な支持（‘＋’），不支持（‘－’）の関係を 35 種類に分類したものである．例えば，図 1 の “EO ＋” は “Expert Opinion” というスキームを用いた支持関係を表し，“FD －” は “Fear and Danger Appeal” というスキームを用いた不支持の関係を表す．提案手法では，結論を表す命題を入力として受け取り，適切な議論スキームを適用することで，その論拠となる命題を出力する．議論スキームには，特殊な命題を結論とする 4 種類を除いた 31 種類を利用する．

図 2 に提案手法の処理の流れを示す．まず，入力命題に対して述語構造解析 [Imamura 14] を適用し，格要素，述語，モダリティ表現（当為や否定表現など）を抽出する．次に，抽出した情報を用いて，あらかじめ議論スキームごとに定義したルール（例えば，当為と否定表現が含まれていれば “FD －” が使用可能と判定するルール）に基づいて，使用するスキームを決定する．その後，使用可能スキームごとに用意した論拠の生成用テンプレートと言語リソースを用いて論拠の候補を生成する．最後に，言語モデルを用いて閾値未満の候補を除外し，スキームごとの Top 1 の候補を論拠として出力する．

表 1 に各議論スキームの生成手法を示す．提案手法では，議論スキームを大きく 6 種類に分類し，“具体例” でオントロジー，“因果関係” と “情報源” でパタン抽出，“決めつけ”，“一般論”，“その他” で固定表現を利用して生成を行う．以降，各生成手法について述べる．

表 1: 議論スキームを用いた論拠の生成手法と具体例．具体例の列に示される入力の角括弧で囲まれたテキストは，論拠の生成に利用するテキストを表し，論拠の角括弧で囲まれたテキストは，各議論スキームごとに用意された生成用テンプレートに設けられたブランクに埋められたテキストを表す．

生成手法	分類	スキーム	概要	具体例 （入力 → 生成される論拠）
オントロジー	具体例	AA +	類推	田舎は [商店] が遠い → 同じように田舎は [バス発着場] も遠い
		AD +	分割	朝食なら [パン] がよい → [手で食べられるもの] がよい
		CO +	合成	パンは [和食] と合わない → [うどん] がパンと合わない
		EX −	例外	パンは [柔らかい] → パンは [固い] 場合もある
パタン抽出	因果関係	PR +/−	推論	[都市に永住す] べきではない → [体に悪い]
		AC +	結果+	[北海道に旅行に行く] べきだ → [思い出を作る]
		FD −	結果−	[沖縄に旅行に行く] べきではない → [天気が崩れる]
		CE +	原因	[沖縄に旅行に行く] べきではない → [お金がない]
		EC +	結果	[沖縄は平地が多い] → 沖縄は [自転車に乗る]
		EH +	証拠	[ご飯を食べる] べきだ → [ご飯が健康に良い] ならば [ご飯を食べる] べきだ
		SIG +	兆候	[沖縄は平地が多い] → 沖縄は [自転車に乗る] のは [平地が多い] ことの兆候だ
	情報源	EO +	専門家	[ジャンクフードにご飯は合わない] → [食] の専門家がそう言っていた
		PK +	事情通	[都会には 100 円ショップが多い] → [東京] の人が言っていた
		PH +	援助	[田舎に永住する] べきだ → [田舎] の人が喜ぶ
固定表現	決めつけ	BI +	先入観	沖縄は物価が安い → それは [沖縄] に対する先入観だ
		EA +/−	善人悪人	ご飯を食べるべきだ → あなたは良い人なのでその通りだ
		GA +	立場	ご飯を食べるべきだ → あなたはそう言わざるを得ない立場だ
		PA +	個人攻撃	ご飯を食べるべきだ → あなたは悪いことをしたので信じられない
	一般論	PO +	世論	田舎には店がない → みんながそう思っている
		LK −	認知不足	ホットドックは主食にならない → その考えは一般的とは言えない
		ER +	ルール	[ご飯を食べる] べきではない → [ご飯を食べる] のを避けるのはルールだ
		PER +	印象	[田舎は老人が多い] → [田舎は老人が多い] という印象を持たれている
	その他	AP +	前例	[ご飯を食べる] べきだ → [ご飯を食べる] ことで後悔したことがあるはずだ
		SUN +	費用	朝食で [パンを食べる] べきだ → 過去に [パンを食べる] ことがあったはずだ
		AO +	選択	沖縄に旅行に行くべきだ → 他の選択肢よりも良い
		SS +	滑り坂	[ご飯を食べる] べきではない → 一度 [ご飯を食べる] と歯止めがきかない
		IN +	矛盾	都市に永住するべきだ → あなたの過去の行動と矛盾している
		AR +	仮説	パンは手軽に食べれるので需要は大きい → 他によい仮説が思いつかない
		CC +	相関	パンは手軽に食べれるので需要は大きい → これらの間には相関がある
		VAL +	価値	都会は人口が多いため [景気も良い] → [景気がよい] ことは重要だ
		CM +	言動	北海道に旅行に行くべきだ → 過去にそういう言動をしていた

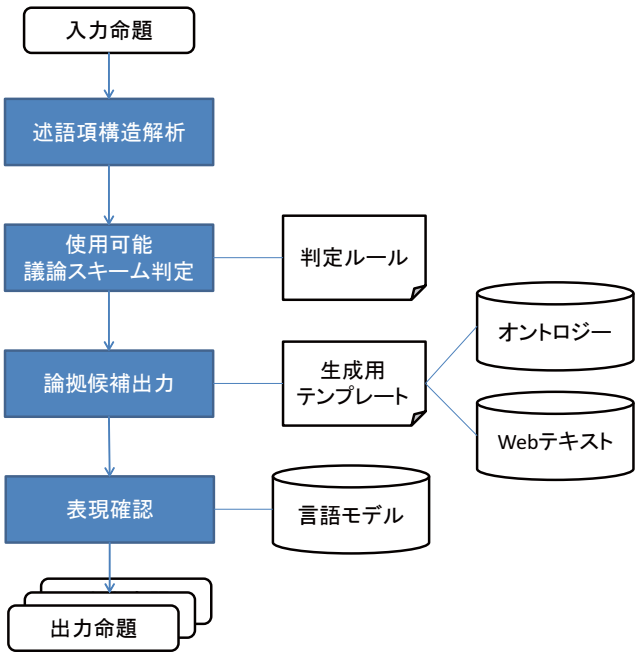


図 2: 議論スキームを用いた論拠の生成の流れ

オントロジーを用いた生成 オントロジーを用いた生成では，入力された命題中の単語をオントロジーに基づいて置換することで，論拠を生成する．オントロジーとして，“AA +” では日本語 WordNet [Isahara 08] の “synonym”，“AD +”，および，“CO +” では ConceptNet [Speer 12] の “IsA”，および，“MadeOf”，“EX −” では我々が独自に作成した反義語辞書を利用する．

パタン抽出を用いた生成 パタン抽出を用いた生成では，Web テキストからパタンを用いて抽出した表現の中から，入力された命題に適切な表現を選択し，論拠を生成する．“因果関係” に分類される議論スキームのパタンとしては “述語項構造，接続表現，述語項構造” を使用し，獲得された述語項構造のペアを利用して論拠を生成する．“情報源” に分類される議論スキームのパタンとしては，“（）の専門家”，および，“（）の人” を使用し，括弧内の表現を利用して論拠を生成する．使用する表現を選択する際には，“因果関係” と “情報源” のいずれも，入力された命題に含まれる格要素となる単語と表現の間の word2vec [Mikolov 13] における類似度を用いる．例えば，表 1 の “PR +” の例であれば，“都市に永住する” に含まれる単語の平均ベクトルを用いて，“都会に住む” という述語項構造をマッチさせ，そのペアとなる “体に悪い” という述語項構造を論拠として出力する．

固定表現を用いた生成 固定表現を用いた生成では，同一の表現，または，入力の一部をテンプレートに当てはめることで論拠を生成する．

表 2: 論拠を生成するモデル．出力される論拠の例の列は，入力された命題が「田舎は空気が良い」の場合に出力される論拠の例を表す．

モデル名	概要	出力される論拠の例
提案手法	議論スキームに基づいたテンプレートをもとに生成を行う手法（支持または不支持）	支持：田舎は自然が良い 不支持：田舎は空気が悪い場合もある
雑談ベース	雑談対話システムを用いて，入力に対する応答発話を出力する手法（支持のみ）	支持：田舎に行くのです
抽出ベース	検索エンジンを用いて，入力にヒットしたテキストを出力する手法（支持のみ）	支持：空気が綺麗で健康になるとの理由から，田舎に住む人が増えていることが○○の調査から明らかになりました（作例）
人手ベース	入力に対する論拠を人手で記述する手法（支持または不支持）	支持：田舎は自然が豊かだ 不支持：田舎は車が多いので空気が綺麗とは限らない

3. 評価実験

表 2 に示す手法を用いて，提案手法の有効性を調査するために，入力された命題に対して論拠を出力し，その論拠の評価を行った．以降，その詳細について述べる．

3.1 比較手法

雑談ベース 本手法は，結論となる命題に対する既存雑談対話システム [Higashinaka 14] の応答発話を論拠として出力する手法である．例えば，雑談対話において出力される同意を表す発話を支持の論拠とみなすことができるように，一部の発話は支持の論拠としてそのまま使用できる可能性がある．出力される論拠は支持のみを表すものとする．

抽出ベース 本手法は，Web などの大規模テキストから結論となる命題の論拠となる命題を抽出する手法に対応する．本稿では，抽出型の手法の代替システムとして，「なぜ？」などの検索が可能な NICT の WISDOM X [*1] を用いて，入力命題をクエリとした際の検索結果を論拠として利用する．出力される論拠は支持のみを表すものとする．

人手ベース 本手法は，結論となる命題の論拠を人手で記述したものをそのまま利用する手法である．結論とその論拠が記述されたデータとして，酒井ら [Sakai 18] に述べられているデータを利用する．酒井らのデータには支持と不支持の論拠が含まれているため，出力される論拠にも支持と不支持の両方が含まれる．

3.2 評価用の入力データと評価方法

評価用のデータとして前述の酒井ら [Sakai 18] で述べられているデータを利用し，条件節や複数の述語項構造を含むものを除いた命題（1 文の自然文で表されたもの）を入力として使用する．入力を選定する際には，議論スキーム全体の有効性を評価するため，提案手法の各議論スキームの出力数が同数になるようにサンプリングを行い，サンプリングされた命題に対して 3 種類に比較手法を適用することで全 4 種類の出力を得るようにした．評価用の入力として，各スキームごとに 3 事例を選択し，3 事例 × 31 スキーム = 93 事例を用意した．出力に必要なパタン抽出，word2vec の学習，および，言語モデルの学習には我々が独自に収集したブログテキスト約 2 億文書を用いた．

評価方法として，入力された命題（結論）と各モデルが出力した命題（論拠）を作業者に提示し，3 種類のラベル（出力の自然さ，命題間の関係性，出力の説得性）を付与した．論拠の自然さは，出力された命題を単体の発話として見たとき，どの程度自然かどうかを表す（1 点：かなり不自然〜5 点：とても自然の 5 段階）．命題間の関係性は，出力された命題が入力に対して正しく論拠になっているかを表す（支持，不支持，無関係の 3 種類）．付与された関係性は，出力が意図した通り支持または不支持になっているかを判定するために使用され，意図した通りの関係が付与されれば正解，それ以外は不正解となる．出力の説得性は命題間の関係性が支持または不支持と判定された場合にのみ付与し，その説得性を表す（1 点：ほとんど説得性がない〜5 点：かなり説得性がある）．2 名の作業者が作業を行った．作業の前に，付与される関係性が支持と不支持に偏らないよう，発話間の関係を付与する課題であると教示を与えた．また，提示する評価用の事例の順番はシャッフルしておき，どの手法で生成された論拠かは作業者にわからないようにした．

*1 http://wisdom-nict.jp/#top

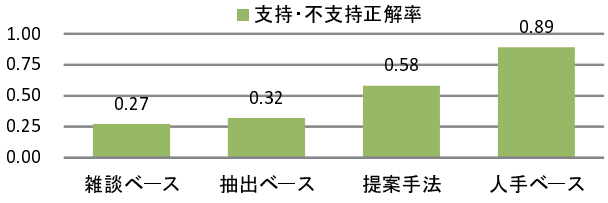


図 3: 出力した命題の支持・不支持が正しく判定された割合

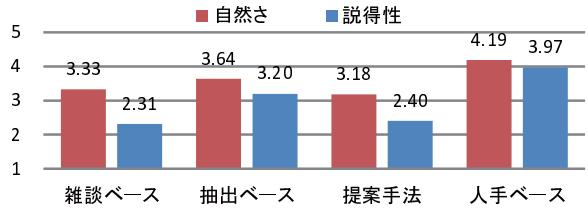


図 4: 出力の自然さと説得性（支持・不支持が正しく判定されたもののみ）

3.3 評価結果

図 3 に，付与された命題間の関係性に基づく，各手法の出力した命題の正解率を示す．図 3 から，**人手ベース**には高い適合率で正しく支持または不支持が付与されたことがわかる．**提案手法**は，**人手ベース**よりは低いものの，**雑談ベース**や**抽出ベース**と比較して，より多くの命題に対して適切な支持または不支持の論拠を生成できていることがわかる．

図 4 に，生成された命題の良さを調査するため，命題間の関係性が正しく判定された出力における命題の自然さと論拠としての説得性を示す．図 4 から，**人手ベース**は自然さ，お

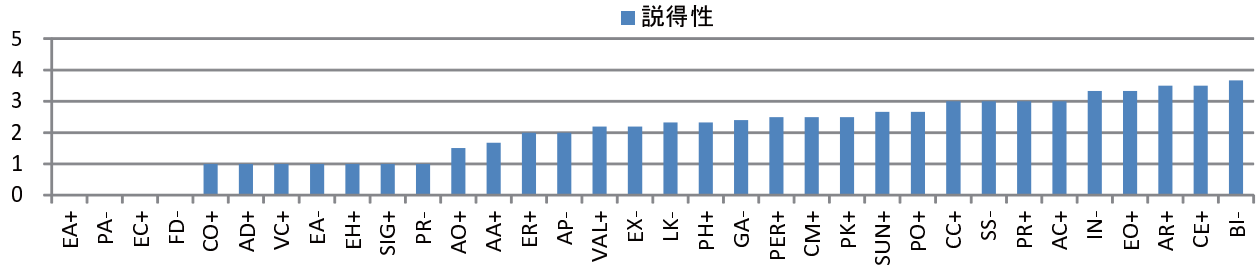


図 5: 議論スキームごとの説得性（支持・不支持が正しく判定されたもののみ）．説得性の値が 0 の議論スキームは，支持・不支持が正しく判定された出力がなかったものを表す．

表 3: 説得性が高い議論スキームのみに限定した場合の説得性の変化（支持・不支持が正しく判定されたもののみ）

使用する議論スキーム	説得性
上位 20%（CC ＋まで 6 種類）	3.38
上位 40%（PER ＋まで 12 種類）	3.12
上位 60%（VAL ＋まで 18 種類）	2.83
上位 80%（PR －まで 24 種類）	2.56
全スキーム（CO ＋まで 30 種類）	2.40

よび，説得性の両方で高い評価値が付与されていることがわかる．**提案手法**と**抽出ベース**を比較すると，自然さは共に 3 点以上であるものの，説得性に大きな差があることがわかる．この結果から，**抽出ベース**では，論拠が適切に抽出できれば，人手に近い高い説得性の論拠を出力することが可能であるとわかる．また，**提案手法**では，出力された全ての命題が論拠として説得性のあるものであるとは限らないということがわかる．

図 5 に，**提案手法**において，論拠の生成に利用した議論スキームごとの説得性の違いを調査するために，議論スキームごとの説得性の平均値を示す．図 5 から，説得性が高く判定された議論スキーム（“BI －” など）と，説得性が低く判定された議論スキーム（“EA －”）があることがわかる．説得性が低く判定された議論スキームの出力を調査したところ，論拠の出力が難しい議論スキーム（因果関係を用いる “PR ＋” や “CO ＋” など）では，不自然な命題が出力されており，説得性が低くなっている傾向がわかった．また，一般的な対話において使用されにくい議論スキーム（根拠なく相手を攻撃する “EA －” など）も含まれていることがわかった．

表 3 に，議論スキームごとの論拠の説得性の違いの傾向の調査結果をもとに，説得性が高い上位のスキームのみを使用した場合の説得性の平均を示す．表 4 から，生成に利用するスキームを上位 40%（12 種類）までに限定することで，**提案手法**でも，**抽出ベース**と同程度の説得性を持つ命題を出力可能なことがわかる．

表 4 に，提案手法の出力の例を示す．不自然な表現も一部含まれているものの，入力に対して様々な支持と不支持の論拠を出力できていることがわかる．

4. おわりに

本稿では，議論対話システムの実現に向けて，与えられた命題の論拠となる命題を議論スキームに基づき生成する生成ベースの手法を提案した．既存の抽出ベースとの比較を行う評価実験を行い，抽出ベースと比較してよりさまざまな入力に対して論拠を生成できることを確認した．今後は生成した論拠を用いた議論対話システムの構築や，生成した論拠を入力としてさらに論拠を推定する問題に取り組み，議論対話システムの実現を目指す．

表 4: 提案手法の出力例（入力が「田舎は空気が良い」の場合）

スキーム	生成された論拠
AA ＋	同じように田舎は雰囲気も良い
AD ＋	田舎は自然が良い
CO ＋	例えば田舎は葉が良い
EX －	田舎は空気が悪い場合もある
CE ＋	田舎は天気が良い
EC ＋	田舎は風に当たる
EH ＋	田舎は天気が良いならば空気が良い
SIG ＋	田舎は天気が良いのは空気が良いことの兆候だ
LK －	その考えは一般的とは言えない
BI －	それは田舎に対する先入観だ

参考文献

[Higashinaka 14] Higashinaka, R., Imamura, K., Meguro, T., Miyazaki, C., Kobayashi, N., Sugiyama, H., Hirano, T., Makino, T., and Matsuo, Y.: Towards an open domain conversational system fully based on natural language processing, in *Proc. COLING*, pp. 928–939 (2014)

[Higashinaka 17] Higashinaka, R., Sakai, K., Sugiyama, H., Narimatsu, H., Arimoto, T., Fukutomi, T., KiyoakiMatsui, , Ijima, Y., Ito, H., Araki, S., Yoshikawa, Y., Ishiguro, H., and Matsuo, Y.: Argumentative dialogue system based on argumentation structures, in *Proc. SemDial* (2017)

[Imamura 14] Imamura, K., Higashinaka, R., and Izumi, T.: Predicate-argument structure analysis with zero-anaphora resolution for dialogue systems, in *Proc. COLING* (2014)

[Isahara 08] Isahara, H., Bond, F., Uchimoto, K., Utiyama, M., and Kanzaki, K.: Development of Japanese WordNet, in *Proc. LREC* (2008)

[Lippi 16] Lippi, M. and Torroni, P.: Argumentation mining: State of the art and emerging trends, *ACM Transactions on Internet Technology*, Vol. 16, No. 10 (2016)

[Mikolov 13] Mikolov, T., Chen, K., Corrado, G., and Dean, J.: Efficient Estimation of Word Representations in Vector Space, in *Proc. CoRR* (2013)

[Rakshit 17] Rakshit, G., Bowden, K. K., Reed, L., Misra, A., and Walker, M.: Debbie, the Debate Bot of the Future, in *Proc. IWSDS* (2017)

[Sakai 18] Sakai, K., Inago, A., Higashinaka, R., Yoshikawa, Y., Ishiguro, H., and Tomita, J.: Creating Large-Scale Argumentation Structures for Dialogue Systems, in *Proc. LREC (appeared)* (2018)

[Speer 12] Speer, R. and Catherine, H.: Representing General Relational Knowledge in ConceptNet5, in *Proc. LREC* (2012)

[Walton 13] Walton, D.: *Methods of argumentation*, Cambridge University Press (2013)

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

コーパスとシソーラスを用いた比喩生成

Metaphor creation using corpus and thesaurus

佐藤　遼河[*1]
Ryoga Sato

杉本　徹[*2]
Toru Sugimoto

*1　芝浦工業大学大学院　理工学研究科
Graduate School of Engineering and Science, Shibaura Institute of Technology

*2　芝浦工業大学　工学部
College of Engineering, Shibaura Institute of Technology

We propose a method to create metaphors using corpus and thesaurus. Our system can create metaphors based on input tenors and grounds. We use word2vec to obtain word vectors for tenors and grounds, which are used to select vehicles considering cosine similarity. In order to output various vehicles that are congenial to the ground, we create a grounds-by-categories matrix using thesaurus. We evaluate appropriateness and usefulness of vehicles that are created by our system. As the result, it is shown that our system can output metaphors that are useful for user, though many of the metaphors created by the system are inappropriate. We compared vehicles created using Aozora Bunko corpus and NWJC. As the result, it is shown that our system can output vehicles that are congenial to user's literary style by using appropriate corpus.

1. 研究背景と目的

近年、小説投稿サイトの利用やセルフパブリッシングにより個人での執筆活動を公開する機会が増えている。それに伴い、読者にイメージが伝わりやすい表現を用いることへの需要が高まっている。イメージが伝わりやすい表現方法の一つとして比喩が挙げられる。しかし喩える語（喩辞）と喩えられる語（被喩辞）の関連性を考慮した上でことば選びをする必要があるため、比喩を作成することは難しい。

本研究は文章中のユーザが指定した語に対応する比喩を生成し、ユーザに提案するシステムを構築することが目的である。システムは入力された被喩辞と被喩辞の特徴を表す語（特徴語）から、執筆活動に用いることができる喩辞を複数生成し、提示する。システムが比喩の作成を支援することで、ユーザは容易にイメージが伝わりやすい表現を用いることができるようになる。比喩は直喩、隠喩、換喩、提喩に分類できるが本研究では直喩のみを扱う。

比喩の自動生成を目指した研究としては[北田　2001] [中條 2017]などがある。北田らは被喩辞と喩辞の概念的な近さ、カテゴリの近さ、共起のしやすさを考慮するために EDR 電子辞書を用いて比喩の生成を行った。中條らは毎日新聞をコーパスとした共起行列を用いて比喩の生成を行った。

本研究では word2vec[Mikolov 2013]とシソーラスを用いて比喩を生成する。word2vec を用いることでユーザが執筆する文章のジャンルに合わせたコーパスを選ぶことができる。またシソーラスを用いることで特徴語と共起しやすい意味カテゴリに属する単語を優先的に出力する。

本稿では提案手法と青空文庫コーパスを用いて生成された比喩の評価実験の結果を報告し考察を行う。また別のコーパスとして国語研日本語ウェブコーパスを用いて比喩を生成し、青空文庫コーパスを用いて生成された比喩との比較を行った。

連絡先：佐藤遼河，芝浦工業大学大学院　理工学研究科，
〒135-8548　東京都江東区豊洲 3-7-5，
E-mail:al14044@shibaura-it.ac.jp

2. 提案手法

2.1　比喩生成の処理手順

提案手法の処理の流れを図 1 に示す。

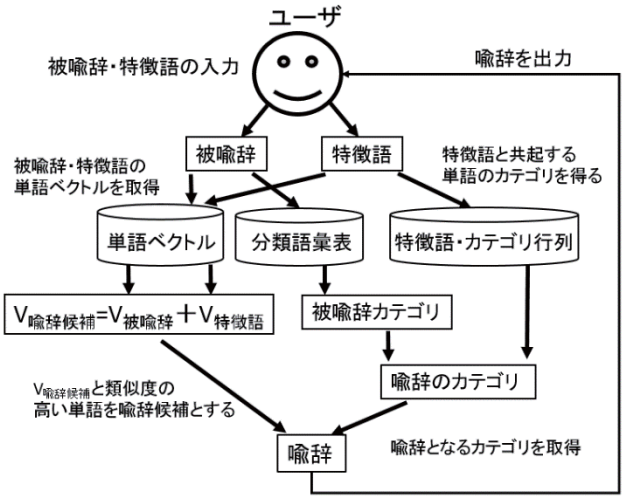


図 1　提案手法の概要

システムの処理手順を以下に示す。

①ユーザが被喩辞と特徴語を入力

被喩辞は名詞、特徴語は形容詞、形容動詞、動詞のいずれかとする。例えば「少女はＡのように美しい」という比喩を生成

するのであれば、被喩辞として「少女」、特徴語として「美しい」を入力する。

②喩辞候補のベクトルを取得

被喩辞と特徴語に対応する単語ベクトルを取得する。この単語ベクトルは予め word2vec を用いて作成したものである。次に被喩辞のベクトルと特徴語のベクトルを加算して、得られたベクトルを喩辞候補のベクトルとする。喩辞候補のベクトルとコサイン類似度が高い単語ベクトルをもつ名詞が喩辞候補となる。

③喩辞のカテゴリを取得

システムは特徴語・カテゴリ行列を参照して、特徴語と共起しやすい単語のカテゴリを得る。ここから被喩辞のカテゴリを除いたものが喩辞のカテゴリとなる。

④喩辞の出力

②で取得した喩辞候補のうち、③で取得したカテゴリに属する語を喩辞として出力する。

2.2 特徴語・カテゴリ行列

特徴語を行、分類語彙表[国立国語研究所 2004]に基づく名詞のカテゴリを列とする共起行列を作成して用いる。

特徴語は名詞の特徴を表す形容詞、形容動詞、動詞の原形である。本研究では IPA 辞書において形容詞、名詞の形容動詞語幹、動詞の自立語に属する語であり、コーパス中に出現する語を扱う。

カテゴリは分類語彙表の中項目の名前である。分類語彙表の中項目の中で喩辞として適切でない語は予め取り除いた。喩辞として適切でない語とは抽象的な語、固有名詞、倫理的に不適切となりうる語である。

行列を作成するために、まずコーパス中の文章を係り受け解析し、特徴語と共起する名詞を含むカテゴリを抽出する。ここで、共起する名詞とは、特徴語が形容詞と形容動詞の場合は係り先、動詞の場合は係り元と係り先の文節中の名詞であるとする。例えば「冷たい雪に覆われた」という文の場合、「冷たい」と「覆われる」は「雪」と係り受けの関係にあるので、「冷たい」と「覆われる」は「雪」のカテゴリである「物質カテゴリ」と共起する。

本研究ではコーパスとして青空文庫に掲載されている作家 83 人分の文章を約 313MB 用いた。また係り受け解析には CaboCha を用いた。

行列の要素 a_{ij} の値は以下の式で求める。

$$a_{ij} = \frac{\textit{特徴語}i\textit{とカテゴリ}j\textit{の単語が共起する回数}}{\textit{カテゴリ}j\textit{の単語の総出現回数}}$$

特徴語を表す行において要素の値が大きいものから順に喩辞のカテゴリの候補とする。

2.3 単語ベクトル

単語の分散表現ベクトルは word2vec の skip-gram を用いて 100 次元のベクトルを生成した。word2vec はコーパスからニューラルネットワークの学習を行い、単語の意味を表す固定長ベクトルを獲得する手法である。単語ベクトルはそれぞれの見出しの単語に対して float 型 100 次元の配列によって表現されている。入力された被喩辞、特徴語と一致する見出しを持つ単語ベクトルを求め、取得した 100 次元の配列をそれぞれ被喩辞ベクトルと特徴語ベクトルとする。

本研究ではコーパスとして前節で述べた青空文庫の文章において単語の活用形を基本形に直したものを用いた。表 1 に word2vec 実行時の各種パラメータを示す。

表 1 青空文庫をコーパスに用いたときの word2vec 実行時のパラメータ

CBOW or skip-gram	-cbow	0
次元数	-size	100
文脈長	-window	8
負サンプリング数	-negative	25
階層化 softmax	-hs	0
最低頻度閾値	-sample	1e-4
反復回数	-iter	15

3. 評価実験

3.1 生成された喩辞の適切性の評価

3.1.1 評価方法

生成された喩辞に対してアンケートによる定量的な評価を行った。10 種類の被喩辞と特徴語のペアをシステムに入力し、出力された喩辞を 5 人の被験者にすばらしい(4)、適切である(3)、判断が難しい(2)、不適切(1)の 4 段階で評価してもらった。アンケートでは被験者に「『被喩辞』は○○のように『特徴語』」という文を提示し、○○に喩辞を当てはめた場合の適切さを評価してもらう。例えば被喩辞「少女」、特徴語「美しい」の例では「少女は○○のように美しい」という文を提示する。

本実験ではシステムが生成する喩辞を喩辞候補ベクトルとのコサイン類似度が 0.5 以上かつ各カテゴリの上位 5 位以内の単語とし、最大で 50 語を出力した。

3.1.2 結果

表 2 に喩辞の出力結果の例を示す。

表 2 喩辞の出力例

入力 被喩辞, 特徴語	喩辞の カテゴリ	生成された喩辞
ひとみ, 光る	天地	二日月
		夜空
	物質	黒曜石
		夕焼け
	資材	たいまつ
		イルミネーション
ラジオ, 鳴る	物質	雷
		春雷
	道具	鐘
		ベル
	言語	サイレン
		警鐘
人生, 儚い	物質	うたかた
		泡沫
	心	夢幻
		追憶
	芸術	哀詩
		悲劇

次に喩辞の評価結果を表 3 に示す。

表 3　喩辞に対する評価の割合

喩辞数	4	3	2	1
477 個	4.2%	13.1%	12.3%	70.4%

生成された喩辞の 17.3%がすばらしい、または適切であるという結果が得られた。また入力した被喩辞と特徴語のペアのうち、コサイン類似度が 0.5 を上回る喩辞の候補が少なかったため、50 個の喩辞が得られないものが 3 組あった。

3.2 比喩の有用性の評価

3.2.1 評価方法

本実験では前節の実験で評価が高かった喩辞を被喩辞と特徴語のペアごとに上位 10 個ずつ用いて有用性の評価を行う。比較対象として文芸サークルに所属する大学生が作成した喩辞を被喩辞と特徴語のペアごとに 10 個ずつ用いる。6 人の被験者に被喩辞と特徴語を提示し、システムが生成した 10 個の喩辞と人が作成した 10 個、合計 20 個の喩辞をランダムに提示する。被験者は比喩として用いたいと思う表現を提示された喩辞の中から 5 個選択し回答する。

本実験で用いた喩辞の例を表 4 に示す。

表 4　評価に用いた喩辞の例

被喩辞：少女　特徴語：美しい	
システム	人
令嬢	サクラ
姫	女神
聖女	満月
騎士	初雪
女王	オーロラ
ヒロイン	星座
女優	花火
舞姫	宝石
花	夜空
人形	ダイアモンドダスト

3.2.2 結果

被験者が選んだ喩辞のうちシステムが生成した喩辞の割合は 35.3%、人が作成した喩辞は 64.7%であった。このことからシステムが生成した喩辞に一定の有用性があることが分かった。被喩辞ごとの評価結果を図 2 に示す。

システムが生成した喩辞と人が作成した喩辞のうち、共通して得られたのは「被喩辞:ラジオ,特徴語:鳴る」に対する喩辞「雷」のみだった。この点から本システムは人が比較的思いつきにくい喩辞を提案できることが分かった。

生成された喩辞のうち 10 個中 9 個の比喩において有用な喩辞が得られた。有用な喩辞を得られなかった比喩は被喩辞「船」、特徴語「大きい」だった。これは「大きい」という特徴語が相対的な特徴語であるからだと考えられる。被喩辞「船」、特徴語「大きい」の評価に用いた喩辞とアンケートの結果を表 5 に示す。

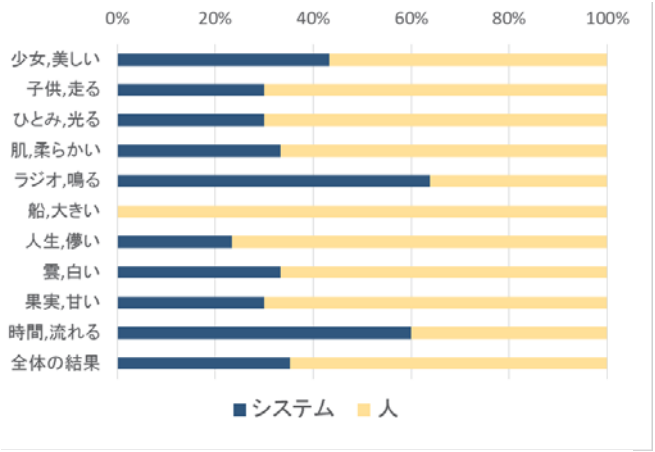


図 2　有用性の評価結果

表 5 「船が大きい」のアンケートの結果

被喩辞：船　特徴語：大きい			
システム		人	
太鼓腹	0人	山	5人
頭でっかち	0人	鯨	4人
巨体	0人	島	5人
海	0人	建物	1人
陸	0人	軍艦	3人
防波堤	0人	城	6人
ガスタンク	0人	要塞	3人
倉庫	0人	岩	2人
格納庫	0人	ドーム	1人
ロケット	0人	マンモス	0人

他の比喩における特徴語は「美しい」、「走る」、「光る」、「柔らかい」、「鳴る」、「儚い」、「白い」、「甘い」、「流れる」であり、いずれも絶対的な特徴である。例えば「花」の「美しい」という特徴は比較対象がある場合でも存在する。また「子猫」が「走る」という特徴も他の比較対象が「走る」場合でも失われない。一方「太鼓腹」の「大きい」という特徴は相対的な特徴であるため、「船」と比較した場合は失われる。つまり「太鼓腹」は「船」よりも小さい。「船」の「大きい」という特徴を強調するために「太鼓腹」という船より小さいもの用いたことで比喩として不適切になったと考えられる。

4. コーパスによる出力の差異

4.1 実験

コーパスによる出力の差異について調べる。本システムはユーザの執筆する文章のジャンルに合わせたコーパスを用いることで、より文章に適したことば選びができると期待できる。

本実験では word2vec を作成する際のコーパスとパラメータを変え、異なる単語ベクトルを用いて比喩を生成する。喩辞を生成する際に入力する被喩辞と特徴語は前節で用いた組み合わせと同じものを用いる。

4.2 使用する単語ベクトル

比較するベクトルは青空文庫をコーパスとして作成した単語ベクトル（以下「青空文庫単語ベクトル」と称する）と浅原らの作成した nwjc2vec[浅原 2017]である。nwjc2vec は国語研日本語ウェブコーパス（NWJC）から word2vec の CBOW を用いて生成した 200 次元の単語ベクトルである。国語研日本語ウェブコーパスはウェブ上の日本語テキストを利用した大規模コーパスである。

表 6 に nwjc2vec を生成する際の word2vec 実行時の各種パラメータを示す。

表 6 nwjc2vec を生成する際に用いられた word2vec 実行時のパラメータ

CBOW or skip-gram	-cbow	1
次元数	-size	200
文脈長	-window	8
負サンプリング数	-negative	25
階層化 softmax	-hs	0
最低頻度閾値	-sample	1e-4
反復回数	-iter	15

4.3 結果

はじめに前節の実験と同様にシステムが生成する喩辞をコサイン類似度が 0.5 以上かつ各カテゴリの上位 5 位以内の単語とし、最大で 50 語を出力した。その結果得られた喩辞は 108 個だった。青空文庫単語ベクトルが 477 個の喩辞を得たことと比較すると、得られた喩辞は少なかった。これはコサイン類似度が 0.5 以上の喩辞の候補が少なかったためである。そこで本実験では多くの喩辞を得るため、コサイン類似度が 0.4 以上の語を喩辞として出力する。

nwjc2vec を用いた場合の喩辞の出力例を表 7 に示す。

表 7 nwjc2vec を用いた喩辞の出力例

入力 被喩辞, 特徴語	喩辞の カテゴリ	生成された喩辞
ひとみ, 光る	植物	さゆり
		まゆみ
	物質	三日月
		流れ星
	物質	陽子
		ルビー
ラジオ, 鳴る	物質	遠雷
	道具	チャイム
		ホイッスル
	成員	DJ
	言語	サイレン
		警鐘
人生, 儚い	社会	今生
		現世
	心	夢
		思い出
	芸術	物語
		悲喜劇

コサイン類似度が 0.4 以上かつ各カテゴリの上位 5 位以内の単語を最大で 50 語を出力するという条件で喩辞は合計 332 個得られた。

nwjc2vec を用いたことで青空文庫単語ベクトルとは異なる表現を得ることができた。例えば「被喩辞：ラジオ,特徴語：鳴る」に関して nwjc2vec を用いたことで「DJ」や「ホイッスル」などの喩辞を取得した。これは「DJ」や「ホイッスル」のような単語は青空文庫には用いられないが、ウェブでは一般的に用いられているからであると考えられる。一方、青空文庫単語ベクトルを用いた場合と同様に「遠雷」や「サイレン」や「警鐘」や「鐘」などの喩辞が得られた。nwjc2vec と比較すると、青空文庫単語ベクトルの方が同じ意味の語でも仰々しい言い回しの語が出力された。例えば nwjc2vec では「夢」や「思い出」が出力されているのに対し、青空文庫単語ベクトルでは「夢幻」や「追憶」が出力された。このことから nwjc2vec は現代が舞台の物語や説明的文章を書くときに用いる比喩に適している。一方青空文庫単語ベクトルは古風な表現が求められる歴史小説や伝奇小説、荘厳な世界観を表現するファンタジー小説を書くときに用いる比喩に適している。

nwjc2vec を用いた場合、多義語の問題があった。例えば「被喩辞：ひとみ,特徴語：光る」では被喩辞の「ひとみ」が nwjc2vec のベクトルでは眼球や目という意味だけでなく、女性の人名を表す語としての意味を表している。その結果コサイン類似度の上位に植物カテゴリからは「さゆり」や「まゆみ」、物質カテゴリから「陽子」などの女性の人名を表す語が出力されている。これはウェブで出現する人名の中に一般名詞としての意味をもつ語が多いからだと考えられる。

5. 結論

本研究では被喩辞と特徴語の入力に対してコーパスとシソーラスを用いて喩辞を生成する方法を提案した。今回行った実験により提案手法に一定の有用性があることが分かった。またコーパスを変えることで執筆する文章に合わせた喩辞が出力できることも分かった。提案手法を用いることでユーザは容易に比喩を用いることができるようになると期待できる。

今後の課題は出力された喩辞の適切性の低さの問題、相対的な特徴を表す喩辞をどのように扱うかという問題、多義語によって想定と異なるカテゴリの語が出力される問題の 3 つが挙げられる。

参考文献

[北田 2001] 北田純弥, 萩原将文：電子辞書を用いた比喩による文章作成支援システム, 情報処理学会論文誌, Vol.42, No.5, pp.1232-1241 (2001)

[中條 2017] 中條寛也, 松吉俊, 内海彰：意味空間に基づく文脈情報を用いた比喩生成, 情報処理学会研究報告, Vol.2017-NL-231, No.14, pp.1-10 (2017)

[国立国語研究所 2004]国立国語研究所：分類語彙表 増補改訂版, 大日本図書 (2004)

[Mikolov 2013] T. Mikolov, K. Chen, G. Corrado, and J. Dean：Efficient Estimation of Word Representations in Vector Space. In ICLR Workshop (2013)

[浅原 2017]浅原正幸, 岡照晃：nwjc2vec：『国語研日本語ウェブコーパス』に基づく単語分散表現データ, 言語処理学会第 23 回年次大会発表論文集 (2017)

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

統合物語生成システムの未結合諸要素と属性情報による概念辞書の拡張

Uncombined Elements in an Integrated Narrative Generation System and Expanding Conceptual Dictionaries by Attribute Information

小野淳平*1 Jumpei Ono
伊藤拓哉*2 Takuya Ito
小方孝*2 Takashi Ogata

*1 岩手県立大学大学院 Graduate School of Iwate Prefectural University
*2 岩手県立大学 Iwate Prefectural University

Although we have been studying diverse elements of narrative generation, many of the elements have not been combined with our basic narrative generation system called Integrated Narrative Generation System (INGS). In this paper, we list these elements and especially focus on the part of the attribute information of noun concepts in the noun conceptual dictionary in the INGS. We introduce the attribute information into the noun conceptual dictionary to aim at various effects including detailed conceptual explanation generation and rich narrative character formation.

1. はじめに

統合物語生成システム（Integrated Narrative Generation System: INGS）[Ogata 2016]は，物語生成に関連した筆者らの研究及び開発の成果を有機的に結合したシステムである．INGS は物語において生起した内容，語り手が如何にしてその内容を語るかといった部分，さらに文字，音，絵などの物語の表現部分まで，一貫した生成を行うことが可能である．

INGS における物語生成の流れは，固定されていない．生成を開始する機構を柔軟に変化させることで，文，音楽，映像など様々な入力に対応し，そして，各種機構を利用するタイミングもまた動的に変化することで，循環的な生成が可能となる．

また，生成パラメータを自動的に調整する手段として，ランダム生成を試みた[小方 2016; Ono 2018]．ランダムに決定されたパラメータに基づき生成された結果を，「良い」あるいは「悪い」といったように単純に評価し，生成パラメータのどのような点が評価に影響を及ぼしたのかを，システムに検討させる方法である．生成・評価・検討を繰り返すことで生成パラメータを徐々に，特定の方向へ調整させる．しかし，INGS に自由な生成を行わせるためには，大きな知識ベースが必要となる．そのため，筆者らは，自動的な知識獲得も試みている[小野 2014; Ono 2018]．

既に筆者らは，概念辞書及び言語表記辞書の構造並びに利用方法の拡張構想[小方 2015a]について述べており，3 節ではそれらも踏まえた形で，変更点・拡張点を示す．また，3 節で述べる成果統合に関する課題のうち，属性情報に関する成果統合，本稿で実施しさらに課題を考察する．

2. 統合物語生成システムの構造に関して

図 1 は INGS の各種機構の構成図である．大きく分けて，ストーリー生成機構，物語言説機構，物語表現機構があり，物語表現機構は，文生成機構，音楽生成機構，映像生成機構を持つ．ストーリー生成及び物語言説生成では，物語の構造を生成し，物語表現機構は物語の表層表現を生成する．各機構の周辺には，物語生成のための知識ベースや概念辞書がある．

連絡先：小野淳平，岩手県立大学大学院，岩手県滝沢市巣子 152-52，019-694-2000，g236m001@s.iwate-pu.ac.jp

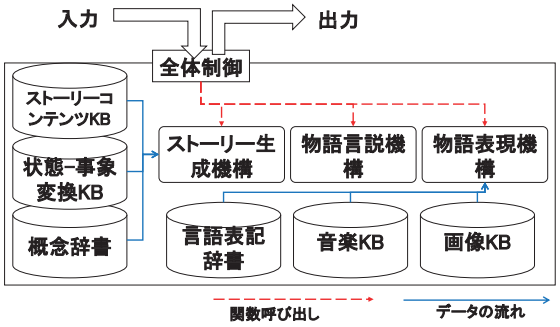


図 1 INGS における各種機構の構成

INGS における物語は木構造が基本である．図 2 は，抽象化したストーリーの木構造である．実際のデータ構造では，事象が終端ノードとなり，関係が中間のノードとなる．状態は，ID によって二つの事象と論理的に結びつく．物語言説では，物語言説構造用の関係ノードがあり，音楽木は，ストーリー木や物語言説木と対応する音楽専用の構造を持つが，物語言説構造も含めて別稿[Ogata 2016]に任せてここでは説明を省略する．

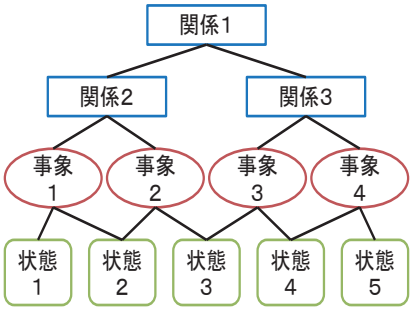


図 2 ストーリーの木構造の例

INGS の開発ではその大部分を Common Lisp により実装した．そのため，生成に用いる辞書や知識ベースの多くは，Lisp 言語におけるリスト形式（またはフレーム形式）によるデータ構造を持つ．音声や画像など上記データ構造で表現が難しいデータも，階層化されたリスト形式のデータと対応関係を持つ．入出

力でやりとりするファイルの形式は，基本的に拡張子「txt」ファイルであり，音楽生成では「mml」や「midi」，画像/映像生成では「jpg」をシステム内で編集する．

文字・記号は，Lisp 言語により開発した機構で処理を行うが，音声処理や映像処理では，他のプログラミング言語やツールを用いており，インタプリンタ言語『Hot Soup Processor』[1]，Music Macro Language による音楽作成ツール『テキスト音楽「サクラ」』[2]，NHK 放送技術研究所が開発した 3D の映像作成ツール『TVprogram Making Language』[3]，歌声合成ツール『UTAU』[4]を利用した．また，知識獲得等においてテキスト解析には，形態素解析器『MeCab』[5]，文章統計分析ツール『KH Coder』[6]を利用している．

3. 開発成果の統合における課題

ここでは，INGS に関連する筆者らの研究・開発のうち，現状の INGS と未統合の成果を列挙し，その課題を示す．加えて，知識ベースや辞書を編集するために開発したツールも示す．

3.1 生成に関する知識ベース及び辞書

まず，物語の構造の生成や表層的な表現を生成するうえで必要となる，知識ベース及び辞書に関連した内容を説明する．また知識構造の修正や規模の拡張だけでなく，知識の自動獲得についても触れる．

- **動詞概念辞書の修正**：動詞概念辞書は，動詞を階層的に体系化した辞書である．中間ノードは動詞のカテゴリを示す中間概念であり，終端ノードは実際のストーリーに現れる動詞を示す動詞概念である．動詞概念は，文生成に用いる基本的な表現を示す文型パターン，必須格を示す格構造，各々の格がとり得る値を示す制約条件を持つ（制約条件は後述する名詞概念の中間概念によって記述される）．以上の三要素は格を媒介として，論理的に結びついている．記述形式上の問題として，後述する知識獲得により，生成の際に参照される動詞概念の種類が大幅に拡大されたため，現状，格構造及び制約条件における記述形式上のミスによる生成エラーが多く網羅的に修正する必要がある．意味的な側面では，階層構造の粗さや制約条件の緩さが挙げられ，3.2 節で述べる逸脱機構において，動詞概念に対するより細かい逸脱の調整が可能となる．
- **固有名詞概念辞書の導入**：名詞概念辞書は，名詞を体系化した辞書である．一般名詞を格納した辞書は，カテゴリを示す中間概念と実際にストーリーに現れる名詞概念を持つ．固有名概念辞書は，一般名詞の終端概念と論理的に結びついた固有名詞のカテゴリ及び固有名詞概念を格納した辞書である．現状は一部分だけ実験的に連結した[寺田 2014]．
- **属性情報の結合**：属性情報は，ある物事の特徴・性質に関する記述である．以降では属性情報を属性と記述する．名詞概念辞書における名詞概念は，一般名詞，固有名詞を問わず属性を持つ．筆者らは属性の自動獲得を試み，Wikipedia から属性を獲得した．それらの属性を実際の概念辞書と結びつける．本稿では，属性の統合の試みを行っており，4 節にてさらに具体的な説明を述べる．
- **修飾概念に対する言語表記辞書の利用方法**：言語表記辞書とは，名詞概念，動詞概念，修飾概念及び助詞に対して，文字による記述方法を指定する辞書である．特に文生成機構で用いる辞書であり，「漢字＋ひらがな」，「ひらがなのみ」，「カタカナのみ」，「ローマ字のみ」の 4 種類の表記方法を格納した（カタカナ語など一部は漢字表記を含まない）．修飾概念は，形容詞概念，形容動詞概念，副詞概念を持つ．
- **スクリプト知識ベースの拡張**：ストーリーコンテンツ知識ベースは，ストーリーの構造を生成するための知識を格納した知識ベースであり．ストーリー生成の形式的手法であるストーリー技法と対応関係にある．ストーリーコンテンツ知識ベースは，因果関係の知識など格納する知識ごとに知識ベースは存在する．筆者らは主にスクリプト知識の自動獲得[荒井 2017]を試みてきた．スクリプト知識は，二つ以上の事象によって構成された，生起順に基づく事象の連なりである．獲得したスクリプトはランダム生成の試み[Ono 2018]でも使用したが，現状，INGS におけるスクリプト知識ベースは，獲得したスクリプト知識を格納していない．
- **CM 分析における中核使用法**：物語生成の応用例として，TV で流れる CM(CF)分析を筆者らは試みてきた[萱森 2003; 阿部 2009]．この分析では，実際の CF におけるストーリーの流れをショットごとに分析した．その分析では，商品の種類ごとに，その商品を使用する際の流れ（中核使用法）が知識として用意できており，それらの知識は，スクリプト知識として利用できると考えており，そのために中核使用法を現在の知識ベースの記述形式で整理する必要がある．
- **プロップに基づくストーリーコンテンツグラマー**：プロップに基づくストーリーコンテンツグラマー(Propp-based Story Contents Grammar: PBSCG)[小方 2015b]は，プロップが述べた物語の構造に関する理論をモデル化し，ストーリーグラマーとして実装したストーリー技法である．知識及び形式的手続きの各一部分の統合に留まり，最新バージョンは未統合である．

3.2 物語の構造の生成及び変形に関連する成果

この節では，INGS における生成過程のうち，特に物語の構造に関連する成果を整理する．一部は 3.1 節や続く 3.3 節と関連する内容もある．

- **統計情報の利用**：ここでの統計情報とは，任意の既存のテキストの範囲内における，ある語についての出現頻度や他の語との共起関係を表す値である．獲得した統計情報は，概念辞書の概念と結びつけられ，各概念の統計情報として利用可能とした（但し現状，語の言語表記を基準に対応付けしたため，意味的な対応付けは不完全である）．筆者らは，青空文庫に格納された既存の作品を対象とし，頻度情報[吉田 2016]及び共起情報[照井 2016]を獲得した．さらに概念辞書に格納された各概念選択の機構を最新の INGS に統合する必要がある．
- **属性の利用（説明文生成）**：3.1 節で述べたように，属性はある物事の特徴・性質に関する記述である．そのため，ストーリー生成などにおいて，属性を用いた説明文構造の

[1] http://hsp.tv/ (最終アクセス 2018.3.4)
[2] https://sakuramml.com/index.php? (最終アクセス 2018.3.4)
[3] https://www.nhk.or.jp/strl/tvml/index.html (最終アクセス 2018.3.4)
[4] http://utau2008.xrea.jp/ (最終アクセス 2018.3.4)
[5] http://taku910.github.io/mecab/ (最終アクセス 2018.3.4)
[6] http://khc.sourceforge.net/ (最終アクセス 2018.3.4)

挿入が可能である. 属性を参照する程度によって, 説明文の量や質が変化し, 例えば冗長さといったストーリーの質を調節する. 説明文機構の試作では, 獲得した属性の参照範囲によって, 説明文の文量を変化させる仕組みを導入した.

- **逸脱機構:** 3.1 節で述べた広告分析に基づく広告ストーリー生成手法から, 異化的修辞[Zhang 2012]を開発した. これは非常識的・非現実的な事象を生成し, 商品など特定要素を際立たたせる技法である. 生成の際には, 逸脱度を指定し, 逸脱度に基づいて動詞概念の持つ制約条件外の名詞概念を選択し, 逸脱した事象を生成する. INGS では一部実装済みであり, 異化的修辞を連続的に適用する異化的ストーリー生成技法を試作した.
- **ギャップ技法によるストーリー生成:** ある線形の時間的な流れにおいて, 人は, その先の事象を予測し, 実際に生起した事象との差(ギャップ)によって驚く. ギャップ技法とは, ストーリーに任意のギャップを与え, 受け手に対して驚きを与える技法である[小野 2017a]. 上で述べた異化的修辞を援用し, 筆者らはギャップ技法を開発した.
- **俳句生成:** 五・七・五の音で構成された俳句を生成する機構である. ストーリーと俳句の循環的な生成[伊藤 2017]を目指す方法と, 深層学習を用いる方法[五十嵐 2017]が存在する. どちらの生成結果も, データ構造の側面から見て, 現状の INGS の標準的な記述形式から異なる仕様のため, 直接的な統合は難しい.

3.3 表層表現の生成に関連する成果

この節では, INGS における生成過程のうち, 特に文・音楽・映像に関連する成果を説明する.

- **カメラワーク**: 筆者らは映像生成におけるカメラワーク技法として, 小津安二郎監督の東京物語におけるカメラワークの分析し, OZU ルールを開発した[小方 2009]. OZU ルールは, 小津監督のカメラワークのルールベース化であり, TVML による映像に対して, 小津監督風のカメラワークを与える. 現状の INGS における映像生成機構は, 二次元的画像を編集して映像生成するため, 三次元的な表現を伴うカメラワークの研究結果を反映しづらい. そのため, 共有可能な部分の検討が必要である. また, 東京物語における行為と動作の関係も分析した. その成果と映像生成機構が持つアニメーションデータベースとの統合することで, より細かなアニメーションの生成も可能となる.
- **音楽生成機構:** 物語の概念構造と音楽の構造の循環的な生成を目指した機構である[Akimoto 2013]. 音楽生成機構単独では, ストーリーから原曲, 物語言説から変奏曲, 原曲から変奏曲を生成可能である. 現状の INGS では, 以上の中から一部の機能を統合した. また応用システムでは, 音楽生成に用いる知識ベースのモチーフを偏らせ, 民話風の音楽が生成されるように調整した.
- **歌生成機構:** 研究としては未発表であり, 小さな試作を行った段階である. 音楽生成機構で出力される原曲と, 文生成機構で出力される文を組み合わせることで歌を生成する. 無伴奏であり, 原曲の音符一つに文の一文字を対応づけた midi ファイルが生成される.

3.4 知識の可視化・編集ツールなどその他の成果

この節では, 3.1 節で述べた生成に関連する知識や辞書の構造の可視化や編集を目的としたツールを示す.

- **階層構造可視化ツール:** INGS における物語の構造や概念辞書の構造は全容を一望しづらいため, リスト形式によるデータ構造を図形で表現するツールを用意した.
- **概念辞書記述形式チェッカー:** 動詞概念辞書における形式的な記述ミスをチェックするツールである. Lisp プログラミング上の括弧の数の間違いの他, 格構造と制約条件における格の数, 制約条件の記述ミスを判別する.
- **PBSCG における知識登録簡易ツール:** 任意のストーリーのあらすじを分解し, ストーリーコンテンツグラマーへ変換するためのツールである[藤原 2015]. あくまで人手による知識作成作業を補助するためのツールであり, 自動的な知識獲得の要素は持たない.
- **スクリプト半自動構成ツール:** 日本語で構成された文章を入力とし, 文章に含まれる動詞からスクリプト知識を構成するためのツールである[荒井 2017]. 動詞の抽出, 動詞概念の選択候補の列挙, スクリプト知識化した際の格構造の内容はツールが決定するが, より細かい動詞の意味の決定や格構造の修正はユーザが行う.

4. 属性と概念辞書の結合

この節では, 3 節で述べた成果のうち, 属性情報を取り上げ, INGS における概念辞書と結合する.

4.1 属性の概要

[影山 2012]は恒常的に持続する性質を記述する文を属性叙述文と呼び, 筆者らは属性叙述文の解析により, ある概念の特徴・性質を記述する属性の自動的な獲得を試みた[小野 2014].

INGS における属性は, 名詞概念辞書に格納された終端概念が持つ. 属性は構造化されたフレーム形式で記述される. これを属性フレームと呼ぶ. ストーリーに登場する人・物・場所・時間は, 名詞の終端概念に基づき生成されるインスタンスである. ストーリーに現れるあらゆるインスタンスは, 属性フレームにおける is-a スロットにより名詞概念と結びつく. あるインスタンスは, 結びついた名詞概念が持つ属性に加え, その時々における状態を記録するスロットにより, 属性フレームが構成される.

図 3 は獲得された属性の記述形式である. <対象概念>は属性と結びつく名詞概念を示し, <テーマ>は属性の種類を示す. 例えば「所在地に関する属性」,「外見に関する属性」等がある. <テーマに対する属性>は, 構造化された文が入る.

```
(<対象概念>
 (<テーマ 1>
    <テーマ 1 に対する属性 1>
    <テーマ 1 に対する属性 2>
    …<テーマ 1 に対する属性 n>)
  …(<テーマ m> …) …)
```

図 3 属性フレームの記述形式

名詞概念辞書に含まれる名詞概念のうち, 一般名詞の名詞概念の数は, 115765 個である. この名詞概念は, 具体的な物体を表すもの(食べ物や建物, 地形等の何らかの形を持つもの)から物事を表すもの(季節や思想, 職業等の形を持たない抽象的なもの)まで含まれる. これに対して, Wikipedia から獲得

された一般名詞の属性は 5615 個である(全一般名詞概念の約 5%)[小野 2017b]. さらに固有名詞の名詞概念は, 188340 個あり, 獲得できた属性は, 28932 個である(全固有名詞概念の約 15%).

4.2 属性の利用と課題

現状の INGS が生成する物語の木構造は, ある二つの状態間における変化を示した, 動的な情報である事象によって構成される. しかし, ある物事の特徴や外見の明ないし描写, ある時間における状態等の静的な内容の生成もまた必要な要素である. 例えば, 受け手にとって理解しづらい概念に対して説明文を挿入し, 受け手の理解を補助することが考えられる. そのため, 筆者らは知識獲得や説明生成を試みた[小野 2017b]. 4.1 節で述べた 5615 個の属性では, 60 のインスタンスで構成されたストーリーうち, 約 1 割のインスタンスに説明文が付与できた[小野 2017b]. しかし, より多くのインスタンスに対して, 説明文の付与を考える場合, 属性情報を持つ概念の数を増やすことが課題となる. 現状, 例えば日本語 Wikipedia の範囲のみで獲得を行った. そのため, テキストの解析と解析結果の属性化だけでなく, 関連テキストの収集も必要となるだろう.

5. おわりに

本稿では, INGS と統合していない筆者らの成果を列挙し, また統合による効果や統合に向けた課題等を示した. 特に名刺概念が持つべき属性に関して, 詳しく説明した. 以上の成果の統合は, INGS における物語生成を拡張し, 新たな応用システムの可能性を示すことができる.

参考文献

[阿部 2009] 阿部弘基, 小方孝, 小野寺康: 広告における商品導入の修辞の分析とシステムの構築, 2009 年度人工知能学会全国大会(第 23 回)論文集, 1J1-OS2-4, 2009.

[Akimoto 2013] Akimoto, T., Endo, J., & Ogata, T.: The Expansion of Paths in the Mutual Transformation Mechanism of Music and narrative, International Journal of Cognitive Informatics and Natural Intelligence, 7(4), 44-63, 2013.

[荒井 2017] 荒井達也, 小野淳平, 小方孝: 統合物語生成システムにおけるスクリプト型事象連鎖の獲得と利用, 言語処理学会第 23 回年次大会論文集, 703-706, 2017.

[藤原 2015] 藤原朱里, 小野淳平, 小方孝: プロップに基づくストーリーコンテンツグラマーを利用した知識登録・格納簡易ツールに基づく考察, 人工知能学会第二種研究会ことば工学研究会(第 48 回)資料, 57-66, 2015.

[五十嵐 2017] 五十嵐広太, 伊藤拓哉, 小方孝: 単語・文字をベースとした深層学習用いた俳句生成, 第 56 回ことば工学研究会資料, 33-34, 2017.

[伊藤 2017] 伊藤拓哉, 小方孝: 俳句／物語生成システムに関する考察, 第 56 回ことば工学研究会資料, 35-37, 2017.

[影山 2012] 影山太郎: 属性叙述の世界, くろしお出版, 2012.

[萱森 2003] 萱森修, 小方孝: 物語ーブランド関係に注目した TVCF の分析と CF シナリオ自動生成, 人工知能学会全国大会(第 17 回)論文集, 2G2-06, 2003.

[小方 2016] 小方孝, 荒井達也, 小野淳平: 統合物語生成システムを用いたランダム物語生成, 第 15 回情報科学技術フォーラム講演論文集, 133-134, 2016.

[小方 2015a] 小方孝, 小野淳平, 戸来裕紀, 五十嵐航, 白井柊太: 統合物語生成システムのための概念辞書及び言語表記辞書の構成・利用の拡張構想, 人工知能学会第二種研究会ことば工学研究会(第 48 回)資料, 51-56, 2015.

[小方 2015b] 小方孝, 藤原朱里:「機能」を具体化する方式の一般化―統合物語生成システムにおける Propp-based ストーリー生成機構の拡張―, 第 29 回人工知能学会全国大会論文集, 3G4-OS-05b-3, 2015.

[Ogata 2016] Ogata, T.: Computational and Cognitive Approaches to Narratology from the Perspective of Narrative Generation in Ogata, T., & Akimoto, T. (eds.), Computational and Cognitive Approaches to Narratology, IGI Global, 1-74, 2016.

[小方 2009] 小方孝, 立花卓, 富手瞬: 物語概念表現からの映像の生成と自動カメラワーク―『東京物語』の分析とシミュレーション, P2-32, 2009.

[小野 2014] 小野淳平, 小方孝: 名詞句の分類を利用した名詞概念の属性フレームの自動獲得―統合物語生成システムの一機構として―, 第 13 回情報科学技術フォーラム講演論文集 第二分冊, 335-338, 2014.

[小野 2017a] 小野淳平, 小方孝:「ギャップ技法」を利用して「驚き」を作り出すストーリー生成の方法―テーブルトークロールプレイングゲームに基づく物語自動生成ゲームへの一アプローチ―, 認知科学, 24(3), 410-430, 2017.

[小野 2017b] 小野淳平, 小方孝: Wikipedia を利用した概念辞書における属性情報の獲得と物語自動生成ゲームでの利用, 2017 年度人工知能学会全国大会(第 31 回)論文集, 1D3-OS-29b-2in2, 2017.

[Ono 2018] Ono, J. & Ogata, T.: Acquiring Short Scripts and Setting a Case Frame in Each Acquired Script: Toward Random Story Generation, Proc. of the 2018 International Conference on Artificial Life and Robotics, 663-667, 2018.

[寺田 2014] 寺田貴範, 秋元泰介, 小野淳平, 小方孝: 統合物語生成システムにおける固有名詞概念の体系的記述, 言語処理学会第 20 回年次大会, 217-220, 2014.

[照井 2016] 照井和舎, 小野淳平, 小方孝: 語の共起情報による概念・単語選択の改善―統合物語生成システムにおける利用―, 人工知能学会全国大会(第 30 回) 予稿集, 3P1-7in2, 2016.

[吉田 2016] 吉田和樹, 小野淳平, 小方孝: 語の頻度情報による概念・単語選択の改善―統合物語生成システムにおける利用―, 人工知能学会全国大会(第 30 回) 予稿集, 2P1-6in1, 2016.

[Zhang 2012] Zhang, Y., Ono, J. & Ogata, T.: The Rhetoric of Defamiliarization for Narrative Generation Using the Constraints in a Conceptual Dictionary, Proc. of the 34th Annual Meeting of the Cognitive Science Society, Poster 183, 2012.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

Research on Sharing of Robotics Skills

Guilherme C. Affonso*1 Kei Okada Inaba Masayuki

The University of Tokyo, JSK Laboratory

Growing number of robots and increase in its capabilities impose new demands on meta robotic systems, capable of performing hardware abstraction and sharing skills between different robots. In this work such systems are studied, in order to make evident what can actually be done now and what are the problems that must be dealt with in the future.

1. Introduction

Holding the advent of deep learning and the third artificial intelligence boom as background [Matsuo 15], one of the keys of current robotics research is to be *Adaptable* and *Universal*. This gives place to the rise of robots as a new and tangible market, with total sales of household and personal service robots forecast to grow 23.5% per year from 2015 to 2020 [Grace Market Data 15]. Such demands feed robotics research back, in a cycle that leads to growth in both overall number and capabilities of robots.

Under such situation, the administration of specific programs build for specific platforms becomes more and more complex, urging for the unification of software into meta robotic systems capable of abstracting hardware and sharing skills between different robots. Although the creation of a system that can both solve different problems and be compatible with several robots is a common goal for robot software architects, as described in [Ferland 15], robotics research nowadays proves to be highly dependent on hardware platform and software implementation, focusing on environment adaptability rather than universality of execution.

In this work the sharing of skills between robots is explored, in order to delimit what can currently be done towards its solving and make clear what kind of problems still need to be overcome.

2. Definition and Classification of Skills

In order to reason about the sharing of robotic skills, it is necessary to first properly define them in clear and concrete terms. In this work we use the definitions proposed by Affonso et al. [Affonso 18b], modified from Jacobsson et al. [Jacobsson 16] and described as follows.

- **Skill**: *The potential for a robot to reproduce a certain task.*
- **Task**: *A set of actions with meaningful outcome.*
- **Action**: *A set of motions which cause an effect on the outside environment.*
- **Motion**: *Movements that alter the robot's physical state.*

Furthermore, classification of skills into joint, action point and objective level, also given at Affonso et al.'s [Affonso 18b] work, are used. These are described as follows.

1. **Joint level**
   At joint level, each motion is described by a vector gathering angle values for all of robot's degrees of freedom. Skills are the enumeration of such motions.
2. **Action point level**
   At action point level, skills are described by the trajectory of the end effector, with emphasis on its points of application. Possible relations with target objects and effects on the outside environments are also considered, being equivalent to the overall procedure for achieving the task in question.
3. **Objective level**
   At objective level, skills are solely described by the final state to be achieved, without further instruction on motions and subprocedures to be taken.

The above definitions are treated as equivalent statements for appearance level, action level and purposive task level imitation of human behavior, defined by Kuniyoshi et al. [Kuniyoshi 07], for being the sharing of skills between two robots comparable to the imitation between them.

3. Sharing Robotic Skills

3.1 Requirements on Skill Sharing

To share a skill between two robots means, in simple statements, that both of these robots will be able to achieve a task that previously only one of them could accomplish. In further analysis, for being tasks characterized by its *meaningful outcome*, or, in other words, objective, we can understand skill sharing as both the sharing of objectives and the ability to achieve them.

Regarding the sharing of objectives, it is easy to conclude that it can be accomplished through the description of skills in objective level. The opposite is also valid, being possible to say that the generation of objective level statements is a

Contact: Guilherme C. Affonso, The University of Tokyo, 113-8656 Tokyo, Bunkyo-ku, Hongo, 7-3-1, tel. 03-5841-7416, affonso@jsk.imi.i.u-tokyo.ac.jp

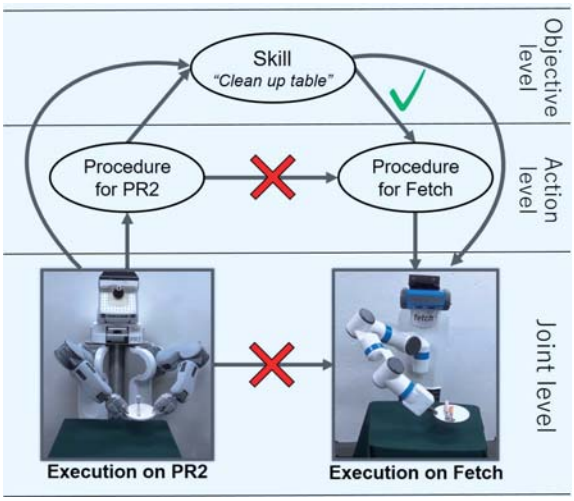


Figure 1: Sharing of robotic skills, as the transition between objective level and task level descriptions.

necessary step for the sharing of objectives and, therefore, skills.

Regarding the ability to achieve a certain objective *i.e.* final state, it is vital to observe that, as any machine, robots are only capable of executing sufficiently concrete and small orders. This means that all commands send to the real robot must be ultimately translated into joint level orders before execution. Therefore, the ability to achieve a certain objective is held as being equivalent to the potential to generate joint level orders based on a given final state.

Above reasoning leads to the conclusion that in order to attain general and certain sharing of robotic skills, it is necessary to (i) share objective level descriptions, and (ii) generate joint level commands accordingly.

An interesting point to notice is that, although being impossible to share skills between different robots directly on joint level, for those being extremely hardware dependent, it is possible to obtain functional results when sharing directly on action point level. This happens in cases where both robots have sufficiently similar hardware properties, such as number of limbs and overall reachability, enough to grant that the original procedure remains valid on the new robot. However, these are limited cases, having general and certain sharing of skills relying on, as previously stated, the combination of objective and joint level descriptions.

3.2 Difficulties on Skill Sharing

The main difficulty imposed by the sharing of robotic skills is the necessity for the machine to deal with objective level descriptions, either in its extraction or in its conversion into joint level commands. Differently from humans, who act consciously of the cause and effect of each motion, knowing what they should and should not do in order to achieve certain objectives; robots – and machines, in general – are only able to process concrete and simple orders, which are executed without any concern on its original purposes or potential side effects. This is due to the incapability for machines to comprehend the cause-consequence nexus, widely known as the *Frame Problem*. The *Frame Problem* can also be understood as the difficulty for artificial intelligence to classify phenomena between related and unrelated [Matsuo 15], in more basic levels. Because it is impossible to know the relations of a certain event with other ones, it is impossible to know which phenomena is directly influenced by which event, being therefore impossible to understand the consequences of actions taken by others or themselves.

4. Past Work

In the past years great effort have been made to build up objective level interfaces, in which programming would be resumed to the definition of the desired final state, using syntax resembling natural language or logical programming like statements. Some examples of such task level programming languages can be verified on [Lozano-Perez 77] [Lieberman 77] [Okano 88], holding the merit of (i) intrinsically attaining the sharing of objectives, and (ii) easing the act of programming itself, for providing methods closer to the objective-oriented human behavior. However, even making use of such interfaces, the necessity of assigning appropriate joint level commands still prevails, making the sharing of skills between different hardware unable to be achieved.

When attempting to deal with the Frame Problem, KnowRob [Tenorth 09] can be pointed out as being of notable interest. Knowrob is a knowledge database implemented on Prolog and the Web Ontology Language OWL, in which the analysis of its enormous database allows robots to learn from experience and reason about the environment and effects of its actions. However, although it is possible to correctly assign subprocedures to higher level commands using this method, it is still limited to specific cases.

Extraction of joint level commands from higher descriptions also finds precedents on the inverse kinematics field [Buss 04] [Nakamura 86], which enables conversion of coordinate based action point level commands into joint level orders. Even though inverse kinematics calculation relies on the number of links and its length, being highly hardware dependent, common interfaces capable of building appropriate solvers based on each robot's hardware descriptions are observed [Matsui 90]. Use of such interfaces allows writing programs fully in action point level, being able to share overall procedures between robots. Although not able to attain complete sharing of skills, sharing of procedures can also be effective on platforms with similar hardware. However, since inverse kinematics result depends on initial position, such codes tend to rely on environment conditions, generating unstable outcomes.

5. Joint - Action level Transition

In this work the extraction of action point level procedures from joint level commands is tested. When combined with common interfaces capable of translating coordinate based commands back into joint level orders, as the ones introduced above, this method makes possible to share procedures of any robotic skill. This is done through source code

and run-time data analysis, by (i) determining and translating motion patterns, and (ii) identifying each action and the relations between them. Details of such method are given on [Affonso 18b], which also proposes the use of a visual and natural language-like programming based interface for facilitating the manipulation of extracted procedures, making therefore possible to share skills in a semi-automatic way.

At the first step, motions patterns are identified by (i) segmenting and labeling original code into minimal, unconditional divisions, and (ii) searching for common sets of labels between executions. Each motion is then translated into action point level descriptions by comparing end effector position relative to the robot base, target object center and previous position. Coordinate systems capable of describing the end effector position in unchanged statements are taken to be the application point of the motion, process illustrated on Fig. 2.

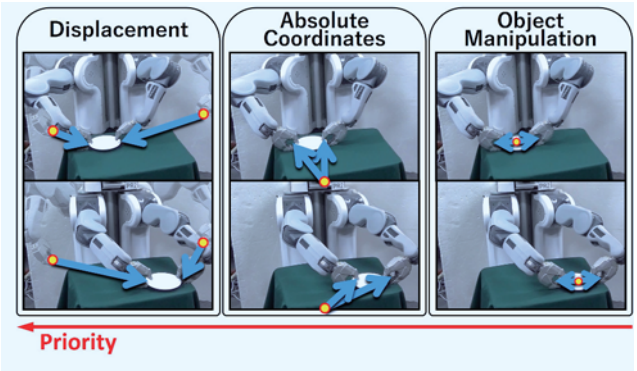


Figure 2: Example of coordinates analysis on motion translation. Here, object centered coordinate system allows unchangeable representation of end effector position, signalizing the application point.

At the second step, the heading label for each motion pattern is analyzed, identifying its trigger conditions and other related motion patterns included on its context. Nested structures containing several motion patterns related to the same conditional clause are interpreted as actions, and bound accordingly. Detailed procedure is given on [Affonso 18a].

6. Experimentation

Experimentation is conducted on the simple domestic task of cleaning up a table after a meal, removing dishes and cutlery left on it. This task is schematically shown at Fig. 3, and is executed a total of seven times for collecting run-time data. Conditions of each trial are given at Table 1.

Without counting the action describing initial pose, five motions patterns and three actions were extracted from given task. Overall procedure, including the relation between such actions and motion patterns, is shown at Fig. 4. Here, actions A, B and C are verified to correspond to original actions of piling-up dishes, collecting cutlery and picking-up (refer to Fig. 3).

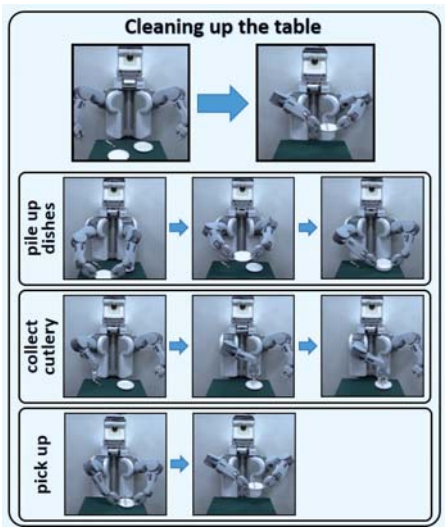


Figure 3: Task of cleaning up table, executed by piling up dishes, collecting cutlery and picking everything up.

Table 1: Conditions of conducted experiments.

trial number	1	2	3	4	5	6	7
number of plates	1	1	1	1	2	2	2
number of cutlery	0	1	1	2	0	0	1
pile plate executions	0	0	0	0	1	1	1
collect cutlery exec.	0	1	1	2	0	0	1
pick up executions	1	1	1	1	1	1	1

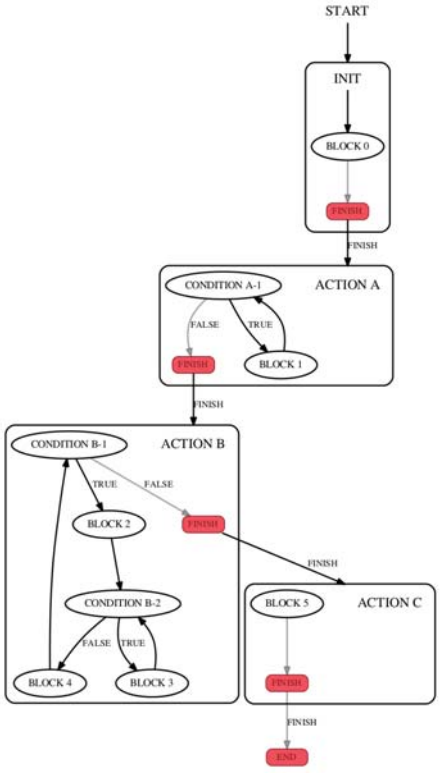


Figure 4: Flow diagram describing task's overall procedure, showing relations between extracted motion patterns (*blocks*) and actions. Conditions are abbreviated.

7. Conclusion

In this work the sharing of robotic skills was dealt with, analyzing its current situation and problems that must be dealt with.

From past work, we can conclude that now it is possible to:

1. Share objectives between robots, by *e.g.* using task level programming languages.
2. Translate coordinate based action point level descriptions into joint level orders, by *e.g.* using inverse kinematics solvers.

In order to attain the complete sharing of skills among different robots, however, it is also necessary to be capable of executing shared objectives, by assigning joint level orders appropriately. This involves dealing with reasoning on higher level objectives, being required to overcome the Frame Problem – difficulty for machines to understand the cause-consequence nexus – at some extent.

We also note that, although not being able to accomplish sharing in general and certain ways, the extraction of action point level procedures can be helpful, for being able to both give functional results on similar hardware and attain certain degree of hardware abstraction. When willing to share a certain skill between robots, even if incomplete this methods can save a lot of time and human work.

References

[Affonso 18a] Affonso, G. C., Okada, K., and Inaba, M.: Detection of Motion Patterns and Transition Conditions for Automatic Flow Diagram Generation of Robotic Tasks, in *International Conference on Intelligent Autonomous Systems (IAS)* (submitted on 2018)

[Affonso 18b] Affonso, G. C., Okada, K., and Inaba, M.: Extraction and Modularization of Procedures Towards Sharing of Robotic Skills, in *International Conference on Intelligent Robots and Systems (IROS)* (submitted on 2018)

[Buss 04] Buss, S. R.: Introduction to inverse kinematics with jacobian transpose, pseudoinverse and damped least squares methods, *IEEE Journal of Robotics and Automation*, Vol. 17, No. 1-19, p. 16 (2004)

[Ferland 15] Ferland, F., Cruz-Maya, A., and Tapus, A.: Adapting an hybrid behavior-based architecture with episodic memory to different humanoid robots, in *24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 797–802 (2015)

[Grace Market Data 15] Grace Market Data, : Global Domestic Service Robots Market Size And Trends To 2020 (August 2015)

[Jacobsson 16] Jacobsson, L., Malec, J., and Nilsson, K.: Modularization of skill ontologies for industrial robots, in *Proceedings of 47st International Symposium on Robotics(ISR)*, pp. 1–6 (2016)

[Kuniyoshi 07] Kuniyoshi, Y.: Adaptive and emergent imitation as the fundamental of humanoid intelligence, *Journal of the Robotics Society of Japan*, Vol. 25, No. 5, pp. 671–677 (2007)

[Lieberman 77] Lieberman, L. I. and Wesley, M. A.: AUTOPASS: an automatic programming system for computer controlled mechanical assembly, *IBM Journal of Research and Development*, Vol. 21, No. 4, pp. 321–333 (1977)

[Lozano-Perez 77] Lozano-Perez, T. and Winston, P. H.: LAMA: A Language for Automatic Mechanical Assembly., in *IJCAI*, pp. 710–716 (1977)

[Matsui 90] Matsui, T. and Inaba, M.: Euslisp: An object-based implementation of lisp, *Journal of Information Processing*, Vol. 13, No. 3, pp. 327–338 (1990)

[Matsuo 15] Matsuo, Y.: 人工知能は人間を超えるか ディープラーニングの先にあるもの [Will Artificial Intelligence overcome humans? What lies beyond Deep Learning., KADOKAWA (2015)

[Nakamura 86] Nakamura, Y. and Hanafusa, H.: Inverse kinematic solutions with singularity robustness for robot manipulator control, *Journal of dynamic systems, measurement, and control*, Vol. 108, No. 3, pp. 163–171 (1986)

[Okano 88] Okano, A., Matsubara, H., and Inoue, H.: Design and implementation of a task-oriented robot language, *Advanced Robotics*, Vol. 3, pp. 177–191 (1988)

[Tenorth 09] Tenorth, M. and Beetz, M.: KnowRob – knowledge processing for autonomous personal robots, in *International Conference on Intelligent Robots and Systems (IROS)*, pp. 4261–4266 (2009)

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

Hybrid Policy Gradient for Deep Reinforcement Learning

Thakur Praveen Singh[*1], Masaru Sogabe[1], Katsuyoshi Sakamoto[2], Koichi Yamaguchi[2], Dinesh Bahadur Malla[2], Shinji Yokogawa[2], Tomah Sogabe[*1,2]

[1] GRID Inc. [2] i-PERC, The University of Electro-communications

In Reinforcement Learning, to solve complex continuous action tasks, it requires a policy with stable learning and faster convergence, since policy may converge to sub-optimal solutions at the early stage of policy learning. Deterministic Policy Gradient(DDPG) algorithm which is a limiting case of stochastic policy gradient in actor-critic approach used for solving continuous tasks. In this paper, we propose an alternative way of updating the actor (policy) in DDPG algorithm to increase convergence and stable learning process. In the basic actor-critic architecture with TD (temporal difference) learning of critic, the actor parameters are updated in the gradient ascent direction of TD-error of critic. Similar to basic actor-critic approach, in our proposed hybrid method, Hybrid-DDPG (shortly H-DDPG), at one time step actor is updated similar to DDPG (gradient of critic output with respect to policy parameters) and another time step, policy parameters are moved in the direction of TD-error of critic. The algorithm is tested on OpenAI gym's RoboschoolInvertedPendulumSwingup-v1 environment. Once among 5 trial runs, reward obtained at the early stage of training in H-DDPG is higher than DDPG. In Hybrid update, the policy gradients are weighted by TD-error. This results in 1) higher reward than DDPG 2) pushes the policy parameters to move in a direction such that the actions with higher reward likely to occur more than the other. This implies if the policy explores at early stages good rewards, the policy may converge quickly otherwise vice versa. However, among the remaining trial runs, H-DDPG performed same as DDPG.

1. Introduction

Reinforcement Learning (RL) has shown successful result in the past few years in the discrete action space such as Atari games using DQN (Mnih et al. 2015), Go game using actor critic with Monte Carlo Tree search (Silver et al. 2017). For continuous control problems, there are mainly two types of algorithms: vanilla policy gradient and actor-critic architecture. Deep Deterministic Policy Gradient (DDPG) (Lillicrap et al. 2016), Asynchronous Advantage Actor Critic (A3C) (Mnih et al. 2016) has actor-critic architectures, Trust Region Policy Optimization (TRPO) (Schluman et al. 2015), Proximal Policy Optimization (PPO) (Shluman et al. 2017) can use vanilla policy gradient and actor-critic architecture. Except DDPG, the above mentioned RL algorithms use stochastic policy gradients and policy is parameterized as Gaussian distribution, from which action is sampled. Whereas in DDPG, policy maps state space to action value. In the policy evaluation, for exploration, policy output is summed with Ornstein-Uhlenbeck noise. DPG is a limiting case of stochastic policy gradient in the limit of variance approaches to zero (Silver et al. 2014). Due to the deterministic behavior in DDPG, in the early stages of policy iteration, if the policy has seen good immediate reward, the policy converges to sub-optimal region and policy has hard time to reach global optimum. Here, we propose a method to find the global optimal policy and faster convergence by using actor update similar to stochastic policy gradient. In this Hybrid-DDPG at one time step of policy update, same as DDPG, the policy parameters are moved in the direction of action-value gradient at another time step, instead of calculating the action-value gradient, TD error calculated for critic is applied at the action gradient of the actor network. It is believed that since the critic network does not reach optimum solution at the early stages of learning, the policy parameters may not converge to sub-optimal solution.

Contact: Tomah Sogabe, i-PERC, The University of Electro-communications, sogabe@uec.ac.jp

2. Background

2.1 Preliminaries:

In the Reinforcement Learning, the control problem is formulated as Markov Decision Process (MDP). A MDP consists of state space (s), an agent to take an action from action space (a), transition dynamics $p(s_{t+1}|s_t, a_t)$, Reward process ($r(s,a)$) and initial state distribution $p(s_1)$ (Sutton et al. 1998). The goal of an agent is to find an optimum policy (π^*) such that it maximizes the total expected cumulative discounted reward.

Total discounted reward $R_t = \sum_{i=t}^{T} \gamma^{(i-T)} r(s_i, a_i)$.
Here, γ is the discount factor $\in$ [0~1].

2.2 Actor-critic method:

In RL, actor-critic architecture is one of the policy gradient algorithm, used to tackle continuous action problems. The actor is the agent (policy) which takes the action, and the critic criticizes the action how good or bad the sampled action from actor. The policy is parameterized with neural network and critic is the value-function. There are two types of actor-critic methods: stochastic and deterministic.

In the stochastic actor-critic off-policy setting, the performance objective is typically modified to be the value function of target policy, averaged over the state distribution of the behavior policy (Degris et al. 2012)

$$\nabla_\theta J_\beta(\pi_\theta) = \mathbb{E}_{s\sim\rho^\beta, a\sim\beta}\left[\frac{\pi_\theta(a|s)}{\beta_\theta(a|s)}(\nabla_\theta \log \pi_\theta(a|s))Q^\pi(s,a)\right] \quad \text{--- (1)}$$

Instead of the $Q^\pi(s,a)$ in Equ. (1), the temporal difference (TD) error, $\delta = r_{t+1} + \gamma V(s_{t+1}) - V(s_t)$ is used for actor update.

In off-policy deterministic policy gradient, the actor update becomes

$$\theta^{k+1} = \theta^k + \alpha \mathbb{E}_{s\sim\rho^{\mu^k}}\left[\nabla_\theta Q^{\mu^k}(s, \mu_\theta(s))\right] \quad \text{--- (2)}$$

By applying the chain rule, the policy improvement is decomposed into the gradient of action-value with respect to actions, and the gradient of policy with respect to policy parameters.

$$\theta^{k+1} = \theta^k + \alpha \mathbb{E}_{s\sim\rho^{\mu^k}}\left[\nabla_a Q^{\mu^k}(s,a)|_{a=\mu_\theta(s)} * \nabla_\theta \mu_\theta(S)\right] \quad \text{--- (3)}$$

In the calculation of gradient of Q with respect to actions in Eq. (3), the error is considered as 1.0 at the output neuron which is propagated until the action gradient. Here, in this paper, we propose, similar to stochastic actor-critic, the TD-error (δ) is applied at the Eq. (3) instead of the action-value Q gradient. The TD-error is

$$\delta = r_t + \gamma Q'(s_{t+1}, \mu'(s_{t+1})) - Q(s_t). \quad \text{--- (4)}$$

For critic update, this TD-error (TD(0) error) is used for backpropagation. We call this method as Hybrid Deep Deterministic Policy Gradient (H-DDPG) method. At one time step, use the Eq. (3) and the next time step use the Eq. (5).

$$\theta^{k+1} = \theta^k + \alpha \mathbb{E}_{s\sim\rho^{\mu^k}}\left[\delta * \nabla_a Q^{\mu^k}(s,a)|_{a=\mu_\theta(s)}\right] \quad \text{--- (5)}$$

3. Propose method

3.1 Hybrid-DDPG Algorithm:

Randomly initialize actor $\mu(s|\theta^\mu)$ and critic $Q(s,a|\theta^Q)$ neural networks with weights θ^μ and θ^Q.
Initialize target network μ' and Q' with weights $\theta^{\mu'} \leftarrow \theta^\mu$ and $\theta^{Q'} \leftarrow \theta^Q$. (Initially copy the same weights)
Initialize Replay Buffer R.
For episode 1 to m do:
  Initialize Ornstein-Uhelnbeck(OU) noise for $\mathcal{N}$ for exploration
  Environment reset to get initial observation s_0
  For t = 1 to Terminal state do
    Select action $a_t = \mu(s_t|\theta^\mu) + \mathcal{N}_t$ by adding noise
    Apply action a_t to environment and observe next state s_{t+1} and reward r_t and terminal situation
    Store following tuple in (s_t, a_t, r_t, s_{t+1}, terminal) in R
    If Replay Buffer size > mini batch size
      Sample K (mini batch) from R
      Calculate $TD(0)_{target} = r_t + \gamma Q'(s_{t+1}, \mu'(s_{t+1}|\theta^{\mu'}))$
      Update critic with mean squared loss:

$$L = \frac{1}{K}\sum(TD(0)_{target} - Q(s_t, a_t))^2$$

      Update Actor (Hybrid):
        At time step: t

$$\nabla_{\theta^\mu} J \approx \frac{1}{K}\sum_i \nabla_a Q(s,a|\theta^Q)|_{s=s_i, a=\mu(s_i)} * \nabla_{\theta^\mu}\mu(s|\theta^\mu)|_{s_i}$$

        At time step: $t+1$

$$\nabla_{\theta^\mu} J \approx \frac{1}{K}\sum_i TD(0)_{target} * \nabla_{\theta^\mu}\mu(s|\theta^\mu)|_{s_i}$$

      Update target networks:

$$\theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'}$$
$$\theta^{\mu'} \leftarrow \tau\theta^\mu + (1-\tau)\theta^{\mu'}$$

  end for
end for

4. Results and Discussion:

4.1 Simulation settings and results:

For fair comparison of the DDPG and H-DDPG, the neural network architectures and hyper parameters are kept same as (Lillicrap et al. 2016). In the OU-noise the parameters are $\sigma = 0.3$, and $\theta = 0.15$. The replay buffer size is 200000. The discount factor gamma (γ), Actor learning rate, critic learning rate are 0.999, 0.0001 and 0.001 respectively. The τ is set as 0.001.

RoboschoolInvertedPendulumSwingup-v1 (OpenAI Gym environments) is tested using both algorithms. Each run is continued for 500 episodes. The reward threshold is 800. For brevity, the Figure 1 is plotted for 300 episodes. The total reward achieved in each episode is plotted. The game is run for 5 times. To test the learned policy without noise for stability, we tested the learned policy every after 10 episodes (after 0,10,20,…,500th episode) The average reward of 10 test episodes is plotted in Figure 2.

4.2 Discussion:

Once among 5 trial runs, In H-DDPG4 curve in Figure 1, reward obtained at the early stage of training is higher than DDPG2 trial run. In Hybrid update, the policy gradients are weighted by TD-error. This results in 1) higher reward than DDPG 2) pushes the policy parameters to move in a direction such that the actions with higher reward likely to occur more than the other. This implies if the policy explores at early stages good rewards,

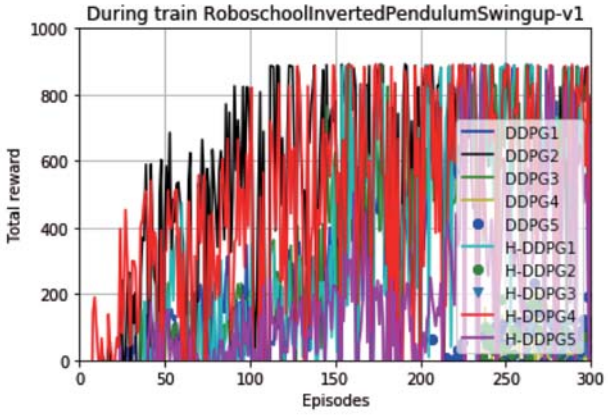


Figure 1. During training rewards obtained in both algorithms

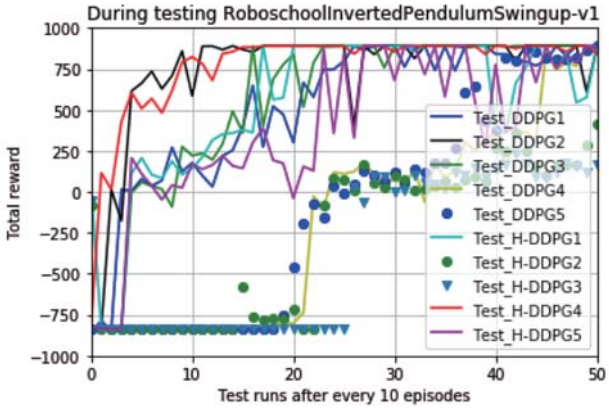


Figure 2. During testing after every 10 episodes of training

the policy may converge quickly otherwise vice versa. However, among the remaining trial runs, there is no much difference between both of the algorithms. In the Figure 2, for the case H-DDPG4 and DDPG2 curves have the same stability of test runs. Briefly, at least the Hybrid performance same as DDPG.

Although it is promising in one trail run of H-DDPG over DDPG, there is a lot rooms to improve the HDDPG for more complex continuous tasks including the introduction of the experienced replay and reward clipping, which are undergoing and the results will be presented at the conference.

References

[Minh] Human-level control through deep reinforcement learning, Nature, 2015.

[Silver] Deterministic policy gradient algorithms, In Proceeding of the 31st International Conference on Machine learning(ICML-14), 2014.

[Silver] Mastering the Game of Go without Human Knowledge, Nature, 2017.

[Sutton] Reinforcement learning: An Introduction, MIT press Cambridge,1998.

[Lillicrap] Continuous control with deep reinforcement learning. In Proc. International Conference on Learning Representations (ICLR),2016.

[Schulman] Trust Region Policy Optimization, In proceedings of the 32nd International Conference on International Conference on Machine Learning (ICML'15), JMLR.org, 2015.

[Schulman] Proximal policy optimization algorithms, Arxiv, abs/1707.06347, 2017.

[Degris] Linear off-policy actor-critic, In 29th International Conference on Machine Learning, 2012.

[Mnih] Asynchronous methods for deep reinforcement learning, Proceedings of the 33rd International Conference on Machine Learning, Vol 48 of Machine Learning Research, 2016.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

人から評価を得てリアルタイムに強化学習する移動ロボット

A mobile robot based on real-time reinforcement learning using human evaluation

山根 健 *1　植月 宏昌　横松 秀康 *2
Ken Yamane　Hiromasa Uetsuki　Hideyasu Yokomatsu

*1帝京大学　*2株式会社タイトー
Teikyo University　Taito Corporation

In develping intelligent robot systems, it is necessary for designers to make robots by fine-tuning parameters and modifying programs in detail. However, this conventional method is a big burden not only for designers but also general users. In this paper, we present a new design method based on reinforcement learning in real time through human-robot interaction. We also develop an autonomous mobile robot system and show its possibilities.

1. はじめに

知能ロボット開発において，一般的には，ロボットのために環境を整えて，パラメータやプログラムを細かく調整するなど設計者がシステムを作り込む．しかし，作り込みには大きな労力が必要であるし，実環境では想定外のことが頻繁に起こる．また，非技術者である一般ユーザがロボットの行動を望む方向へ変更したい場合，いわゆるプログラミングによって機能を修正したり追加したりすることはできない．

これらに対して，我々は人との相互作用を通じてロボット自らが行動を学習する方法を検討している．例えば，ユーザがロボットへ行動を教示したり，ロボットの行動に対してその良し悪しを評価したりする中で，ロボットがリアルタイムに学習して必要な機能を獲得する方法である．これは設計者やユーザにとっても負担の少ない設計方法として期待される．

その実現のために，強化学習 [Sutton 98] の枠組みを利用できると考えている．しかし，実環境において強化学習を行う場合，膨大な学習時間が必要であるなど幾つかの大きな問題がある．これに関して，強化学習の価値関数近似器として選択的不感化ニューラルネット（SDNN）を用いる方法が提案され，単純なアルゴリズムである Q 学習でも従来手法に比べて効率よく学習できることが示された [新保 10]．しかし，実環境で強化学習するための具体的な方法や人という不確定な要素を系に含んでも安定して学習できるかについては明らかではない．

そこで本研究では，SDNN を価値関数近似器として用いて強化学習するロボットを実際に構築し，人との相互作用の中で学習可能か検証する．具体的には，事前知識をほとんど与えずに，ロボットの行動に対する良し悪しの評価を人から得て，ロボットが適切な行動をリアルタイムに学習可能か，ウェイポイント経路を追従させる課題において調べる．

2. 方法

2.1 ハードウェアとセンサ情報の前処理

使用するロボット，測域センサおよび評価用ボタンを図 1 に示す．ロボットは独立 2 輪駆動方式であり，市販のノートパソコンにより走行制御される．なお，安全のために，緊急停止スイッチを取り付け，最大速度を 0.15m/s に制限する．

連絡先: 山根健，帝京大学理工学部情報電子工学科，〒 320-8551 栃木県宇都宮市豊郷台 1-1，028-627-7224，yamane@ics.teikyo-u.ac.jp

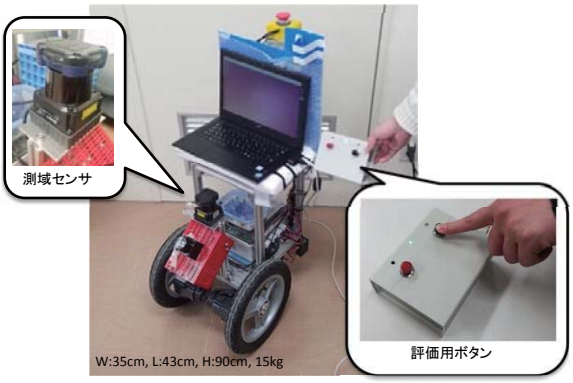


図 1: ハードウェアの構成

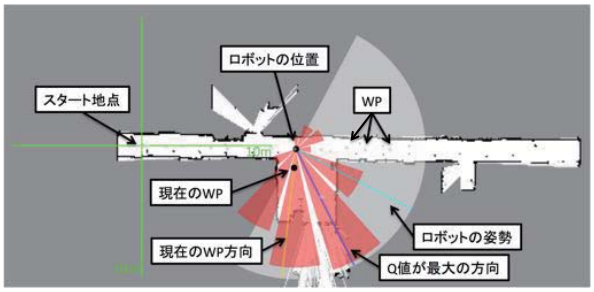


図 2: 実験コースと可視化したロボットの状態

環境情報を取得するため，測域センサ（UTM-30LX-EW，北陽電機）を用いて床面から 0.49m の高さにおける周囲 ±135° の範囲で 1080 点の距離情報を取得し，前処理して状態変数とする．具体的には，5m を超えるものに関して 5m とし，30° ずつの範囲でグループに分け（120 点ずつ 9 つに分け），それぞれの範囲における平均値を求める．次に，最大値 5m で割ることで 0〜1 の値に正規化し，これらを x_1〜x_9 と表す．また，推定したウェイポイント方向とロボットの姿勢の差（±90°）を 0〜1 の範囲で正規化したものを x_{10} として状態変数に含める．これら状態変数の構成イメージを図 2 に示す．

2.2 学習方法

ロボットの学習方法を図 3 に示す．ここでは強化学習アルゴリズムとして Q 学習を，価値関数近似器として SDNN[新保 10, 小林 15] を用いる．SDNN は前述した入力変数 x_1〜x_{10} を受

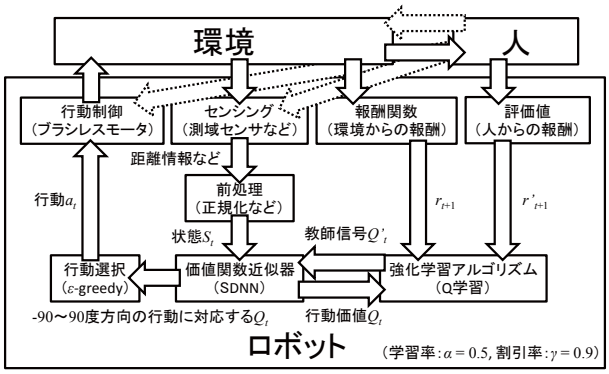


図 3: 学習方法

けて，ロボットの現在の姿勢から $-90°$〜$90°$ 方向へ直進または旋回する行動の価値（Q 値）を計算する．この連続的な Q 値を表現するため，小林らが提案する連続状態行動空間における学習に対応した SDNN[小林 15] を用いて，181 個の行動（1° 刻みの行動）に対する Q 値を出力する．

行動選択では，最大の Q 値に対応する行動を選択する．なお，学習初期のみある一定の確率でランダムに行動を選択すると，連続状態行動空間を効率よく探索できる．

行動実行後，予め設計した報酬関数に基づいて環境から報酬が与えられる．加えて，図 1 に示す評価用ボタンを用いて，人から行動の良し悪しの評価を受ける．2 つのボタンはそれぞれ正の評価（良い）と負の評価（悪い）に対応し，ロボットの行動に対して人が自由なタイミングでボタンを押すことができる（押さなくてもよい）．このボタン押しの頻度から一次元的な値である評価値が計算されて報酬として与えられる．

最終的に，Q 学習の更新式で得られた新しい Q 値（図 3 の Q'_t）を教師信号として SDNN を 1 回だけ学習する．なお，学習には誤り訂正学習を用いる．以上の処理を行動毎に繰り返す．

3. 実験

実際にロボットシステムを構築して，複数のウェイポイント（WP）を順番に繰り返し通過させる課題を行った．具体的には，図 2 に示す 1 周およそ 46.5m の屋内コースにおいて，事前知識なしの状態でスタート地点にロボットを置き，人の評価がある条件とない条件でそれぞれ 4000 step まで行動させた．

このとき，Q 学習のパラメータについて，学習率 $\alpha = 0.5$，割引率 $\gamma = 0.9$ とし，ランダム行動の確率については 1000 step まで 20%，さらに 1000 step まで 10%，その後は 0%とした．報酬関数について，障害物との距離が 0.55m 未満になった場合に-7，WP から離れる場合に-3，ロボットの姿勢と WP 方向の差が 5.6° 未満の場合に 2 を与える．さらに，SDNN を構成する素子数について，入力層 1,000 個，中間層 9,000 個，出力層 3,000 個とし，出力層において -30〜30 までの Q 値を表現した．また，学習係数 $c = 0.01$ とした．

それぞれの条件で 3 回ずつ実験を行った，コース 1 周にかかるステップ数の変化を図 4 に示す．評価の有無に関係なく学習が進むとステップ数が減少した．また，2 周目以降は「評価あり」の方が「評価なし」よりも少ない結果が得られた．

「評価あり」における報酬について見ると，環境から得られた報酬（予め設計した報酬関数から得られた報酬）については，スタート地点，道幅が狭いところや障害物付近などで負の報酬が多かった．一方，人の評価から得られた報酬については，それらの付近でも正の報酬が与えられており，全体として

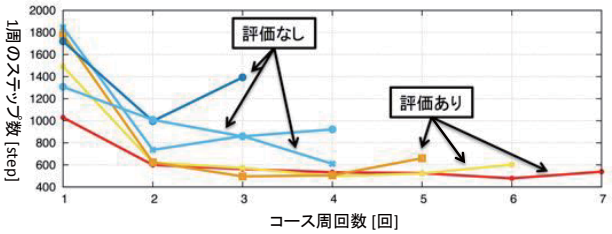


図 4: 1 周にかかるステップ数（4000 step まで）

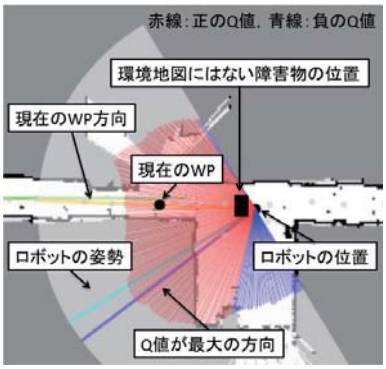


図 5: 障害物を回避しているときの Q 値の構成

も正の報酬が多かった．そこで両者を比較したところ，報酬の正負について 79%が一致していなかった．また，壁に向かう行動に対して人が良い評価をするなど，人が非協力的な条件ではまったく課題を遂行できなかった．これらから報酬関数では対応できない部分を人がカバーしていたと考えられる．

詳しく見るために，障害物付近での Q 値を図 5 に示す．WP 経路追従が本来の目的であるから，現在の WP 方向へ向かう Q 値が学習により高くなると予想される．しかし，実際には WP 方向の Q 値は抑えられており，障害物を回避する方向の Q 値が最も高くなっている．このような Q 値の構成は曲がり角や道幅が狭いところでも見られ，人の評価に基づいて学習する中で簡単な障害物回避機能が獲得できた．

4. おわりに

本研究では，人の評価を利用して強化学習するロボットを示した．詳しい分析は課題として残されているが，人を含む系でもリアルタイムに学習でき，人が協力的な条件では学習効率が高いことや簡単な障害物回避も可能であることがわかった．

今後の課題として，人の評価の効果について調べることや複数の行動次元（例えば，進む方向と速度の組み合わせなど）を扱えるようにシステムを拡張することなどが挙げられる．

参考文献

[小林 15] 小林高彰, 渋谷長史, 森田昌彦:選択的不感化ニューラルネットを用いた連続状態行動空間における Q 学習, 信学論 D, Vol.J98-D, No.2, pp.287-299 (2015).

[新保 10] 新保智之, 山根健, 田中文英, 森田昌彦:選択的不感化ニューラルネットを用いた強化学習の価値関数近似, 信学論 D, Vo.J93-D, No.6, pp.837-847 (2010).

[Sutton 98] Sutton, R. S. and Barto, A. G.:Reinforcement Learning, MIT Press (1998), (邦訳：強化学習, 三上貞芳, 皆川雅章 訳, 森北出版 (2000)).

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

水中ロボットの形状決定を例とした 物理実験からのフィードバックによる進化的計算の研究

Study on evolutionary calculation by real world feedback from physical experiment determining the shape of an underwater robot.

中村 亮太[*1]
Ryota Nakamura

井上 聡[*1*2]
Satoru Inoue

[*1] 埼玉工業大学大学院 工学研究科
Graduate School of Engineering, Saitama Institute of Technology

[*2] 埼玉工業大学
Saitama Institute of Technology

Several kinds of model to solve the problem existing in the real world have been proposed. On the other hand, it is difficult to infer the elements and measure the physical quantities which are necessary for simulation. We propose a system that searches solutions for specific problem by the real world feedback. As an example, we attempted to improve the function of underwater robot by optimize its caudal fin shape based on genetic algorithm.

1. はじめに

実世界における実体の最適化は広く行われており、その解決手段としてコンピュータシミュレーションによる手法が多くみられる。コンピュータシミュレーションは、空間的、時間的な制約が少ないというメリットが存在する。たとえば、製品の性質をコンピュータ上で再現し、設計や製造過程の事前検討を行う CAE (Computer Aided Engineering)技術が広く浸透している。これを用いることで、実物を試作する前に構造を確認することや、実際に試験を行うことなく製品の性能が十分かどうかを検証することができる。一方で、コンピュータシミュレーションを行う際には、扱う対象に関する知識と実験での科学的データの収集が不可欠であるため、対象の性質が未知の場合や、物理量の計測が困難な場合には運用することが難しい。

そこで、対象に関する多くの事前知識を必要とせず、あるいは多くの物理量の計測を省いたとしても、ある程度複雑な目的を達成できる手段が必要だと考えた。動作環境が設計段階で再現でき、モデルの試作にかかるコストが低いという条件下ならば、実際のモデルを用いた開発手法が考えられるのではないだろうか。本研究ではその例として、水中で「尾びれ」を振りながら進むロボットについて、遊泳速度が向上するような形質を求めるというテーマを設定し、有用な設計手法を模索した。

2. 研究目的

物理的な実体が稼働するシステムをコンピュータ上に再現する場合、関係する物理法則を推定し、必要なデータを計測する必要がある。これには、対象に関する専門的な知識と、正確な測定データが必要となるため、人間の知識量によって結果が左右される。また、シミュレーションは実際の物理法則に従うように行わなければならないという制約が存在する。このように、実世界の実体に関するコンピュータシミュレーションには多くの難点が伴う。そこで、コンピュータシミュレーションを用いることなく最適化アルゴリズムを適用し、最適解を求める方法について考える(図 1)。

本研究では、そのような手法の具体的な例として、小型ロボットの遊泳速度を改善するシステムを構築した。コンピュータ上に仮想のモデルを構築することはせず、現実の機械を動作させた結果を探索アルゴリズムにフィードバックすることでのみ解を求め、目的を達成できることを示す。また、設定したテーマにおける提案手法のメリットとデメリットを考察する。

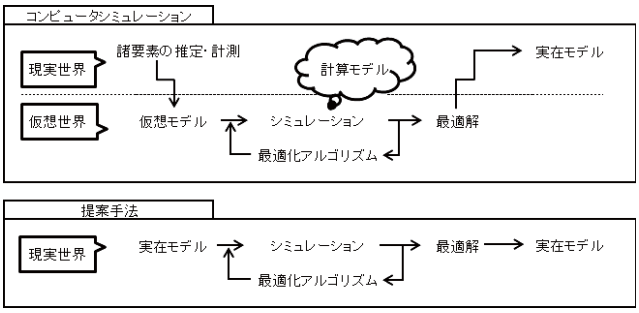


図 1: コンピュータシミュレーションによる方法と提案手法の比較

3. 実験内容

本実験では、「ある往復運動を与えたときに、効率の良い推進力を生み出す『尾びれ』の形質」を求めることにした。問題を単純化するために、外部からの刺激をセンサで感知したり、アクチュエータで運動を制御したりすることはせず、モーターから一定の力を加えて「尾びれ」を動かす。形質を変化させる方法は無数に考えうるが、ここでは「尾びれ」に使用するパーツの形状と剛性を変化させることで推進性能を改善する。

実世界で稼働する物体の中でも、流体を扱うものの最適化は複雑な問題である。まず、物体が受ける抗力、浮力、推力などといった要素は非線形な関数で表現される。また、そのことに起因して目的関数は局所的な最適解を多く持つため、大域解は確率的に求めることになる。水中ロボットおよび遊泳環境は複雑な様相を呈するため、本実験において解の探索には遺伝的アルゴリズム(GA)を用いることにした。GA は、生物が進化する過程を模倣して最適解を探索するアルゴリズムである。GA による解法は、まず解く対象を任意のコードで表現し(染色体とよぶ)、これを複数集めた集団に対して選択、淘汰、突然変異という操作を適用しながら集団を更新していき、終了条件に達するまで繰り返すというものである。他の探索アルゴリズムと比較すると、勾配法のように1点からの探索ではなく、多点からの同時探索を

行うことや、解の評価は評価関数(その染色体から作られる個体の優秀さを比較できるような関数)のみを用いることが特徴である。実験を行うにあたっては、関係する物理法則の推定と計測は行わず、遺伝的アルゴリズムを適用する際に必要な「コーディング」、「適応度の評価」、「遺伝的操作(選択、交叉、突然変異)」の各要素のみを決定する。

3.1 コーディング

本システムでは、遊泳模型の後部に装着する「尾びれ」の形質を変化さていく。「尾びれ」は樹脂で形成された板状のパーツを、1個体につき 1〜3 つを連結して駆動部に装着したものである(図 2)。それぞれのパーツは長さ、表面の加工、可動範囲、連結部分の剛性が違う複数のパーツから成り、その組み合わせによって複数の物理的な特性を表現する意図がある。「長さ」は 15mm から 25mm までの 5mm 刻みと 0mm(その節のパーツなし)の 4 水準、「表面の加工」はフラットなものと溝を付加したものがあり、溝の向きが縦か横かの 2 パターンに加え、溝の間隔が粗・密の 2 パターンを設定した。パーツの長さや表面の加工は水と作用した際の力のかかり方を変化させるほか、連結されているパーツどうしの運動にも影響がある可能性がある。また、ロボット本体と「尾びれ」どうしはジョイントでつながっており、可動範囲と曲がりやすさを変更できるようになっている。可動範囲は 0° ~180° まで 30° 刻みの 7 水準、「連結部分の剛性」はなし、弱、強の 3 水準とした。以上の要因が遊泳能力に影響する可能性を考慮する。このパーツどうしの最適な組み合わせを発見することが実験の目標となる。

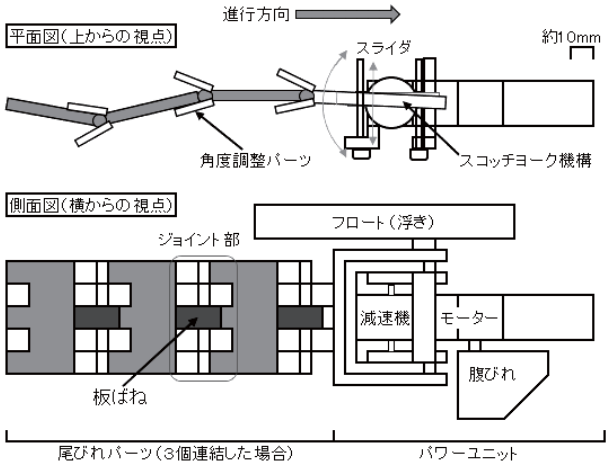


図 2: ロボット全体の模式図

本実験において遺伝的アルゴリズムに用いる染色体は長さ 12 の整数列とし、1〜6 までの整数 12 個が、1 つのパーツに 4 つずつ割り当てられる。この数字がそれぞれのパーツにおける長さ、表面の加工、可動範囲、剛性を決定する(表 1)。

表 1: 「尾びれ」の因子と対応するパーツ、水準の一覧

因子	対応するパーツ	水準(数字は遺伝子にコーディングする際に割り振られるもの)						
		0	1	2	3	4	5	6
長さ(mm)	板	※0	15	20	25			
表面の加工		フラット	横溝(粗)	横溝(密)	縦溝(粗)	縦溝(密)		
可動範囲(度)	角度調整パーツ	0	30	60	90	120	150	180
剛性	板ばね	なし	弱	強				

※「長さ0」は、その節のパーツが存在しないことを表現する。

これを 3D モデルに変換し、3D プリンタでプリントしたものを表現型とする。3D プリンタは、主に材料を薄い層に積み重ねて 3 次元形状を作成する工作機械であり、従来の切削加工等と比較すると、段取りや作業員の熟練度が必要なく、3D データさえ存在すれば人の解釈に依らず造形することができる。また、単品であれば安価かつ高速で試作品を製造できるため、本研究では実験機のパーツを作る目的で利用した。染色体データと、それを元に造形した「尾びれ」の例を図 3 に示す。

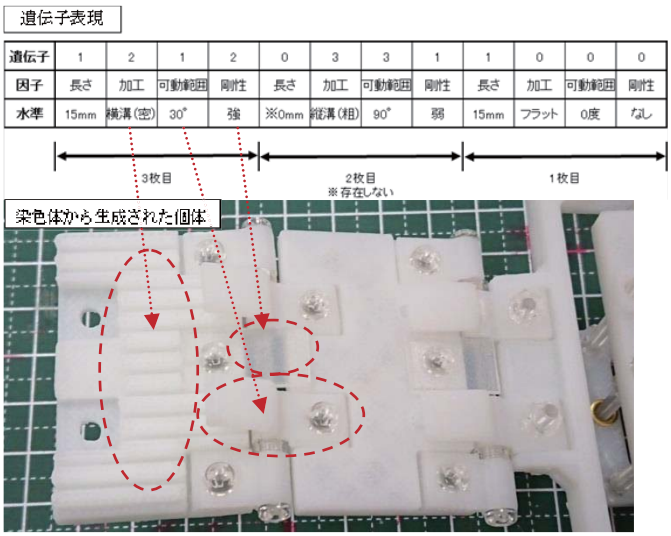


図 3: 遺伝子表現と表現型の対応例

3.2 評価

水中ロボットの動作には様々な評価方法が考えられるが、ここでは最も単純なものとして一定距離を泳ぎ切るまでのタイムを評価値として採用した。造形した「尾びれ」を水中ロボットに装着し、スタート地点を中心とする同心円状のラインで一定距離を泳ぎ終えるまでの時間が短いほど評価値を高く設定する(図 4)。



図 4: 遊泳タイム測定の模式図

3.3 遺伝的操作

選択方式はエリート保存ありのルーレット選択方式とし、交叉は1点交叉とする。また、子の個体の各遺伝子を 5%の確率で別のランダムな値に置き換えることで突然変異を発生させる。

4. まとめ

実世界でのシミュレーションは作業時間の大部分が評価に費やされるため、効率の良い評価方法が必要とされる。また、最適解へ早く収束するような遺伝的アルゴリズムのパラメータを定めることも重要である。

参考文献

[Long 2004] J.H. Long, A.C. Lammert, C.A. Pell, M. Kemp, J.A. Strother, H.C. Crenshaw, and M.J. McHenry: “A navigational primitive: biorobotic implementation of cycloptic helical klinotaxis in planar motion”, IEEE Journal of Oceanic Engineering 29, no.3, 2004

[樋口 1993] 樋口哲也, 北野宏明 : 遺伝的アルゴリズムとその応用, 情報処理 , 情報処理学会, 1993.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

領域ベース CNN と Triplet Loss を用いた 移動ロボット向けオンライン転移学習型人物再同定

Person Re-identification by Online Transfer Learning for Mobile Robot using Region-based CNN and Triplet Loss

村田　祐樹[*1]
Yuki Murata

渥美　雅保[*1]
Masayasu Atsumi

[*1]創価大学大学院工学研究科情報システム工学専攻
Information Systems Science, Graduate School of Engineering, Soka University

We propose a person re-identification method for mobile robots that periodically provides services to specific groups. This method consists of a feature extractor that learns to extract person features based on the Triplet Loss from person regions detected by a region-based CNN and a person re-identifier that learns to identify persons through transfer learning of person features while moving around a room. The person re-identification incorporates adaptive transfer learning to periodically re-learn the same persons with different appearance, such as clothes etc. Performance of the proposed method is evaluated by an experiment using a public large-scale data set and an experiment using the self-made dataset periodically collected for the same group by the mobile robot.

1. はじめに

人物再同定とは一般に，視野を共有しない複数のカメラに映る人物から同一人物を識別する問題である[Zheng 16]．応用例としては屋外空間に設置された防犯カメラ映像を利用した警備システムや，商業施設におけるマーケティングなどが挙げられる．移動ロボットによる人物再同定は，移動ロボットが移動中に再会した同一人物を識別する問題となる．本研究では，特定の集団に定期的にサービスを提供する移動ロボットのための人物再同定手法を提案する．

本手法は，領域ベース CNN が検出する人物領域から Triplet Loss[Florian 15]に基づき人物特徴表現を学習する特徴抽出器と人物特徴の転移学習を行う人物再同定器とから構成される．人物再同定器には服装等部分的に外見が異なる人物を定期的に再学習するための転移学習が組み込まれる．公開されている大規模データセットを用いた実験，及び移動ロボットにより同一集団に対して定期的に収集した自作データセットを用いた実験により，提案手法の性能を評価する．

2. 人物再同定手法

2.1 概要

提案手法は，図 1 に示す 3 つの CNN-人物検出 CNN，人物特徴抽出 CNN，人物再同定 CNN-から構成される．人物検出 CNN は画像から人物領域を検出する．人物特徴抽出 CNN と人物再同定 CNN は検出された人物領域の座標とその中間層の特徴を受け取り，それぞれ Triplet Loss を用いた人物特徴表現学習と移動しながらのオンライン転移学習に基づき人物特徴抽出と再同定を行う．

連絡先：村田 祐樹，創価大学大学院工学研究科，
〒192-8577 東京都八王子市丹木町 1-236，Tel: 042-691-2211，
Email: y.murata0105@gmail.com

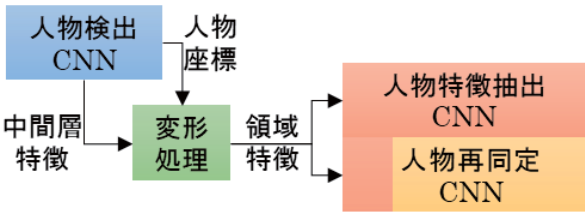


図 1　人物再同定手法

2.2 人物検出 CNN

人物検出 CNN は，画像から人物再同定のための人物領域を取り出す CNN である．高精度で実時間の人物再同定を実現するためには，まず，人物検出において人物の見落としや誤検出を極力減らし，リアルタイムに検出処理を行う必要がある．本研究では，検出器として YOLOv2[Redmon 16]を用いることで，これら課題を解決する．YOLOv2 の CNN は COCO データセットで事前に学習されたものを使用し，人物を含む 80 種類の物体を検出することができる．

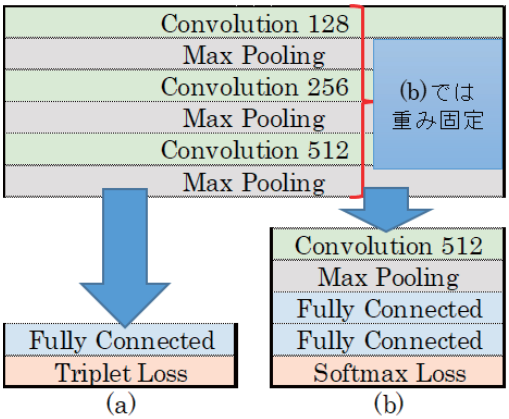


図 2　特徴抽出器(a)と再同定器(b)の CNN 構成

2.3 人物特徴抽出 CNN

人物特徴抽出 CNN(図 2(a))は複数の畳み込み層とプーリング層で構成され，最後の全結合層で 512 次元の人物特徴を出力するネットワークである．損失関数には Triplet Loss が用いられる． Triplet は基準となるサンプル(anchor)，それと同クラスのサンプル(positive)，別クラスのサンプル(negative)の 3 つの組で構成される．それぞれが重みを共有する同一の特徴抽出 CNN へ入力されて，出力される特徴量同士を比較したときに，同一クラス間では距離が小さく，別クラス間で距離が大きくなるように学習が行われる．

2.4 人物再同定 CNN

人物再同定 CNN(図 2(b))は，前半部分が特徴抽出 CNN と共通の層構成をとり，後半部分に再同定用の畳み込み層とプーリング層，及び全結合層を追加したネットワークで，出力層には人物クラスを分類するソフトマックス関数，損失関数には交差エントロピーを用いている．前半部分は人物特徴抽出 CNN 学習時の重みを利用し，後半で追加された層のみパラメータの更新を行う転移学習を取り入れることで，変化への適応と収束の促進を図っている．

3. 実験：人物特徴抽出と再同定の性能評価

3.1 実験概要

特徴抽出 CNN と再同定 CNN の性能評価実験を行う．特徴抽出 CNN の実験では，出力される特徴量間の類似度を利用して識別精度を評価する．再同定 CNN の実験では異なるデータセットで学習した特徴抽出 CNN の重みを利用して転移学習を実施した際の再同定精度を評価する．事前に行った複数の実験から，各 CNN に入力する領域特徴は検出 CNN の 2 番目の畳み込み層から取得し，サイズは 20×20×64 とした．

3.2 データセット

実験に使用する 2 種類のデータセット構成を表 1 に示す．

表 1 データセット構成（サンプル数 / クラス数）

	Train	Gallery	Query
PRW	17690 / 483	16343 / 450	2057 / 450

	Train	Test
SOTARO-1	12789 / 41	3193 / 41
SOTARO-2	12825 / 48	3194 / 48

(1) PRW データセット

公開されている人物再同定向けデータセットの多くは，人物領域が切り取られた画像で構成されている．一方，Person Re-identification in the Wild(PRW)[Zheng 16]では，画像全体と人物領域の座標情報が提供されている．提案手法では人物検出と同時に CNN の中間層から特徴を取り出すため，人物領域が切り取られる前の画像全体のデータが必要となる．そこで，今回はこのデータセットを，人物特徴表現の学習のために使用する．

(2) SOTARO データセット

SOTARO-1 および SOTARO-2 は，当研究室で開発を進めている授業支援ロボット SOTARO[渥美 17]で収集したデータを再構成して作成したデータセットである．2 種類のデータセットは別日に収集されており，服装等見え方の異なる同一人物のサンプルが含まれている．

3.3 人物特徴抽出器の評価

まず，訓練データ(表 1 の Train)からクラスラベルに基づいて Triplet を作成し，特徴抽出 CNN を学習する．次に，学習した CNN を使って全サンプルの人物特徴を抽出し，サンプル間のコサイン類似度を求めて分類を行う．PRW データセットではギャラリデータ(表 1 の Gallery)とクエリデータ(表 1 の Query)間，SOTARO データセットでは訓練データとテストデータ(表 1 の Test)間の類似度を算出する．結果を表 2 に示す．

表 2 類似度評価の結果

dataset	r1	r5	r10	r20	mAP
PRW	78.7	89.0	92.2	94.3	33.6
SOTARO-1	84.1	86.0	86.9	87.8	79.8
SOTARO-2	78.9	82.2	83.3	84.5	74.6

ここで r1, r5, r10, r20 は類似度の高い上位 1, 5, 10, 20 サンプル中に 1 つでも正解となるサンプルが含まれる場合に正しいと評価する指標である．また，mAP は mean Average Precision を表す．

3.4 人物再同定器の評価

3.3 で学習済みの特徴抽出 CNN と SOTARO データセットを用いて人物再同定 CNN の転移学習を実施し，分類性能を評価する．結果を表 3 に示す．

表 3 再同定評価の結果

feature	trans	r1	r5	r10	r20	mAP
PRW	SOTARO-1	79.9	92.7	96.3	98.4	85.3
PRW	SOTARO-2	71.4	85.3	90.7	95.4	77.9
SOTARO-2	SOTARO-1	81.3	93.6	97.2	98.4	86.8
SOTARO-1	SOTARO-2	78.3	91.9	95.0	98.2	84.5

ここで，1 列目の ‘feature`は特徴抽出 CNN の学習に使用したデータセットを，2 列目の‘trans`は再同定 CNN の転移学習に使用したデータセットを表す．

1･2 行目と 3･4 行目の比較から，特徴抽出 CNN の学習に PRW データセットを用いた場合より SOTARO データセットを用いた方が転移学習後の再同定性能が優れていることが分かる．この結果から，特徴表現の学習には，転移学習を行う環境と類似するデータセットを用いるべきであるといえる．

また，転移学習の収束に要した時間は GPU 環境下で 5 分ほどであり，短時間で異なるデータセットに適した再同定器の学習ができることが分かった．

4. オンライン転移学習型人物再同定システム

4.1 概要

人物検出から再同定までの各処理を同時並行的に行うシステムの構成を図 3 に示す．図 1 との違いは，データセット作成のための領域特徴･人物特徴の保存や人による GUI を用いた半自動的なラベル付け処理が追加されている点である．図中の Gallery には，人物が検出されるたびにその特徴が蓄積される．その時点で検出･取得した人物特徴と，Gallery 内に保存されている全サンプルの特徴のコサイン類似度を求め，ある条件に従って自動的にラベルを与えるか，GUI を利用したラベル付けを行うかを判断する．現在試作段階のシステム構成では，最も類似するサンプルの類似度が事前に設定した閾値以上なら自動

的にラベルを与え，そうでない場合は GUI による操作でラベルを付ける処理へ切り替えることとしている．

人物再同定 CNN の転移学習は，訓練データが集まり次第，任意のタイミングで実施する．再同定器生成後は，人物特徴の Gallery サンプルとのコサイン類似度の代わりに人物再同定結果の分類クラス確率を半教師付きラベル付けに利用する．この場合は，再同定 CNN が出力する分類クラス確率が閾値以上なら自動的にラベルを与え，そうでなければ GUI による操作へ切り替えることとしている．

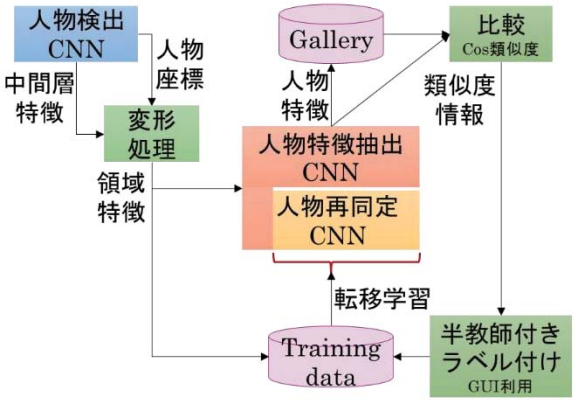


図 3 オンライン転移学習型人物再同定システム構成

4.2 実験：データセット分割による段階ごとの性能評価

本実験では，システムによる半自動ラベル付け処理が正確に行われたと仮定し，サンプル数の増加に対する再同定精度の推移を評価する．3.2 で紹介した SOTARO データセットを 6 分割し，図 4 で示すように 5 段階の再同定精度を比較した．

各段階におけるサンプル数とクラス数（人数），そして再同定精度をまとめた結果を表 4・表 5 に示す．表 4（分割実験 1）は特徴抽出 CNN の学習に SOTARO-2 を，再同定 CNN の転移学習に SOTARO-1 を使用した結果，表 5（分割実験 2）はその逆の条件で実験を行った結果を示している．分割実験 1 では学習に使用するサンプル数が増えるたびに再同定精度も向上する結果となった．一方，分割実験 2 では段階 3 において精度が若干低下した．これは段階を踏むごとにデータ量とともにクラス数も増加しており，分類問題としての難易度が上がっているためと考えられる．2 つの実験から，データ量を増やすことが精度向上に繋がり，早い段階から多くの人数のサンプルを集めることが重要であることが分かった．

4.3 ラベル付け GUI 設計

現在開発中の GUI を図 5 に示す．この GUI は，ラベル選択部とラベル調節部で構成される．画面上部のラベル選択部では，左端に表示されたサンプルと比較して，類似度の高い 5 つのラベル候補サンプルからユーザが手動でラベルを付けることができる．また，それまで出会ったことのない人物を見つけた場合，新たなラベルを追加することができる．画面下部のラベル調節部では，各ラベルに含まれるサンプル数やその代表となる画像を表示するほか，ラベルの修正が可能である．

5. むすび

本論では，領域ベース CNN の中間層特徴から Triplet Loss に基づく特徴学習と転移学習により人物再同定を行う移動ロボット向けの手法を提案した．そして，大規模データセットと自作のデータセットに対して行った実験により，その有用性を確かめた．また，データセットを分割して行った追加の実験から，再同定精度の向上とサンプルとなるデータ量の関係が明らかとなった．今後の課題としては，高精度に人物再同定を行うために早い段階でデータ量を増やす工夫や，再同定結果を移動ロボットのサービス提供に活かすことが挙げられる．

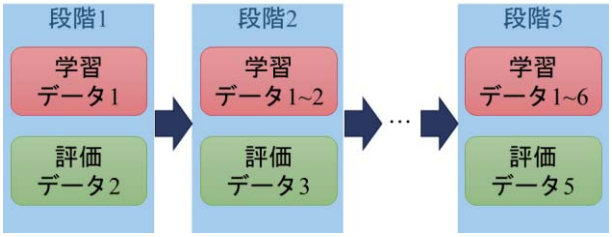


図 4 再同定精度の段階的評価方法

表 4 データセット分割実験 1

	段階1	段階2	段階3	段階4	段階5
サンプル数	2663	5327	7991	10654	13318
クラス数	35	40	40	40	40
精度	19.5	27.4	56.8	59.7	81.4

表 5 データセット分割実験 2

	段階1	段階2	段階3	段階4	段階5
サンプル数	2668	5336	8004	10672	13340
クラス数	28	42	45	48	48
精度	26.2	35.0	26.3	52.4	74.9

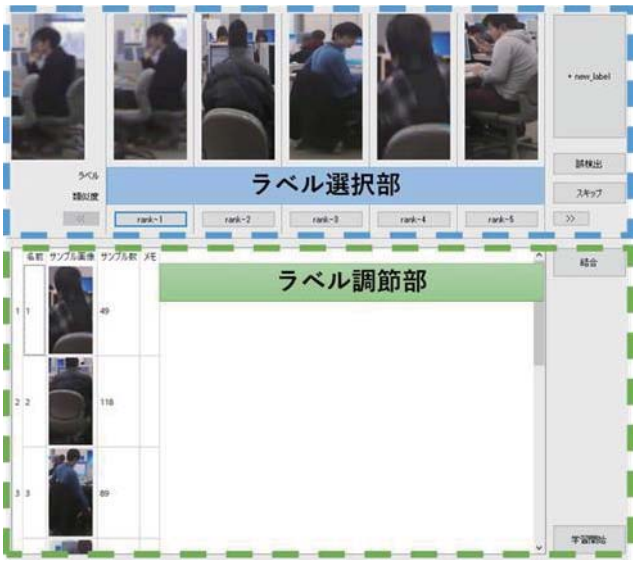


図 5 ラベル付け GUI

参考文献

[Zheng 16] L. Zheng, et al.: Person Re-identification in the Wild, arXiv preprint arXiv:1604.02531, 2016.

[Florian 15] S. Florian, et al.: FaceNet: A Unified Embedding for Face Recognition and Clustering. arXiv preprint arXiv:1503.03832, 2015.

[Redmon 16] J. Redmon, and A. Farhadi.: YOLO9000: Better, Faster, Stronger. arXiv preprint arXiv:1612.08242, 2016.

[渥美 17] 渥美雅保, et al.: SOTARO:オープン チャットとロボットの連携による演習授業支援シス テム―Human-in-the-loop 型機械学習によるアプローチ―. 情報処理学会第 79 回全国大会, 2017.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

遺伝的アルゴリズムによるシューティングゲームにおけるゲーム AI

An AI game player trained by a genetic algorithm that avoids bombardments in a shooting game

滑川　静海[*1]　　手塚　太郎[*2]
Shizuma Namekawa　　Taro Tezuka

[*1] 筑波大学知識情報・図書館学類
College of Knowledge and Library Sciences, University of Tsukuba

[*2] 筑波大学図書館情報メディア系
Faculty of Library, Information and Media Science, University of Tsukuba

Making programs play games has contributed significantly to the advancement of AI research. It is partly because popular games often resemble real-world problems. Tackling them has enabled AI to cope with real-world scenarios as well. Now that AI has surpassed humans in strategy board games such as chess and go, one next target would be to train it to play video games. This paper focuses on a shooting game and optimizes a program to avoid bombardments deployed by the opposing player. Using a genetic algorithm, the AI player was optimized to move around without hitting enemy attacks. The results of experiments showed that it can successfully learn to do so, although with much computation time.

1. はじめに

人工知能の発展と共に、それにゲームをプレイさせる試みが盛んである。人間がゲームを好むのは本能的な欲求と親和性があるからであり、ポピュラーなゲームは人間の生存に必要なタスクと何らかの類似性を持っていることが多い。そのためゲームで人工知能を鍛えることは今後、人工知能を現実社会に進出させていく上で有益であると考えられる。ゲームは現実世界のタスクと比較してシミュレーションが容易であり、その学習プロセスを研究しやすいというメリットもある。

現在、ゲームに使われている AI には様々な種類が存在している。大別するとゲーム自体をコントロールするメタ AI、ゲームの NPC（non-player character）として機能するキャラクターAI、その時々に応じたゲームの進行を支えるナビゲーション AI の三種類に分けることができる。今回はそのうちのキャラクターAI に分類されるものを扱う。

本研究では AI 技術のうち、特に遺伝的アルゴリズムを用いて戦略を学習させるアプローチをとる。遺伝的アルゴリズムとは生物が進化していく仕組みから着想を得た最適化アルゴリズムである。多数の個体のそれぞれが解の候補を持ち、それらが突然変異や交叉と呼ばれる形で値を変化させながら、最適解を探索する。また、損失関数が小さい個体ほど自らの複製を多数生成するため、その解候補周辺で重点的に探索が行われる。

使用したゲームは PC を中心に生産されている東方 project シリーズの東方妖々夢（図 1）というゲームである。これはいわゆる弾幕シューティングゲームというジャンルに該当する。弾幕シューティングとは、敵機が大量の弾を発射し、自機がそれを避けつつ敵を倒していくというゲームを指す。

熟練した人間のプレーヤーはシューティングゲームにおいて敵が放った大量の弾幕を驚くほどの器用さで回避できる。本研究の目標は遺伝的アルゴリズムを用いてそのようなプレーを AI にも実行させることである。

連絡先：滑川静海，筑波大学知識情報・図書館学類，〒305-855
茨城県つくば市春日 1-2，s1511536@u.tsukuba.ac.jp

このような能力の実現は、たとえば混雑した交通状況の中で自動運転の車をスムーズに移動させる機能にも繋がることが期待される。



図1：東方妖々夢のスナップショット

2. 関連研究

2.1 ゲーム AI

ゲーム AI は様々なジャンルのゲームで研究開発されている。将棋や囲碁といったパズルゲームだけでなく RPG やアクションゲームにも応用されてきている。例えば、昔のアクションゲームは、敵の出現位置や強さが固定されていた。しかし現在はオンライン学習という技術が研究されている。Marc Ponsen（2004）によると「オンライン学習とは、ゲームがリリースされた後に、AI がプレイ中に最適な行動を選択していくというものである。」としている。これはどういうことかというと、以前までのゲームでは完全にランダムにするか決められた環境で遊ぶことしかできなかった。しかしこのオンライン学習という技術を用いると、ゲームの進行中に AI が自分のプレイに適応する、つまり自分の実力にあった難易度や自分のリズムにあったプレイスタイルでゲームを進行することができるようになる。

他の AI を用いた研究としては、AlphaGo が挙げられる。これは碁のプログラムであるが、ニューラルネットワークやディープラーニング、強化学習などを用いて学習をしている。特徴的な学

習方法として、囲碁の AI 同士で対戦することによってお互いを高め合う形で強化学習をしている。これは遺伝的アルゴリズムにおいても応用可能性があり、研究として扱うことができるだろう。

2.2 遺伝的アルゴリズム

遺伝的アルゴリズムに関しても様々な研究がなされている。遺伝的アルゴリズム自体が物事の最適化を進めることのできるものであるため、ゲーム分野との親和性が高い技術と言える。したがってゲーム分野において有効に応用することができる。例えば、本研究で扱うようなシューティングゲームや、アクションゲーム、RPG など多岐にわたる分野で使うことができる。三宅(2008)によると「ゲーム内のキャラクタの進化などに使用される」としている。具体的には、いくつも個体を用意し、それらに何度もトラップを課すことで進化させるというものである。他にもレースゲームに遺伝的アルゴリズムを適用することで効果を得た例もある。このように遺伝的アルゴリズムはゲーム AI の実現に高い適合性を持っていることがわかる。

3. 提案手法

3.1 遺伝的アルゴリズムの概略

本研究で提案する手法では、プレイヤーキャラクターを操作して条件をクリアする個体が現れるまで繰り返し解の探索を続けることである。具体的には、上下左右と動きなしの五種類のコマンドを遺伝子として持った個体を 30 体用意し、各個体でステージをプレイさせ、その個体の中から親候補となる 10 個体をエリート選択方式で選択する。そして選ばれた10個体の中から2個体をルーレット選択方式で選び交叉を行い、3体の個体を生成する。これを10回繰り返すことで新たな世代の30個体を生成する。以上の一連の操作をクリア個体が出るまで繰り返し行い、その経過を見る。

3.2 遺伝子データ

今回は左右のコマンドが選ばれる確率を 5%、上下が選ばれる確率を3%とし、動かない確率を残りの84%とする。これらにより作られた遺伝子データを持つ個体を 30 体用意する。今回の実験ではゴールまでにかかるコマンド数が未知数であるため、また交叉の質を保つため第一世代や上の世代のコマンド数を超えた場合は逐次コマンドを上記の確率に基づいて生成する。この際、なぜ上記のような偏りのある確率にしたかというと、シューティングゲームにおいて、時期が動くことはリスクが高いからである。通常プレイするときは、飛んでくる弾を避ける際にむやみに動かず、自分に当たると予測される球を見つけた際に動いて避けるという動作を行う。従って、動き続けるのではなく動くべき時に動く AI が実用的かつ人間味のある動きと言える。このような理由からそれぞれのコマンドに偏りを持たせている。また、左右と上下の確率に違いがあることもシューティングゲームの特徴が起因していて、東方 project は縦方向シューティングであるため敵の弾は上から飛んでくることになる。その弾を避ける際、縦方向に動くことよりも横方向に動いて避ける方が避けやすいということは感覚的にも明らかだろう。以上のような理由からある程度試行錯誤した結果今回の確率分布に設定した。

3.3 ゲームのプレイ

用意した 30 体の個体を実際にプレイさせる。ここでのプレイ方法であるが、実際に売られているゲームであるためソースコードが公開されていない。そのような理由で UWSC という OS マクロツールを用いて、実際にゲーム上でキー操作を行う形でコマンド操作を行う。

この実験におけるゲーム内の設定であるが、自機の残機は 1(一度被弾した時点で死亡)とする。また、弾は常に発射している状態でボム(一定時間無敵かつ強力な攻撃)は使わないという制約のもと操作を行う。また、実験を行う際のゲームの難易度は easy として進める。

今回の実験においては、第一世代に限り、コマンド生成→コマンド入力→死亡判定→コマンド生成…のように繰り返す。第二世代以降は、引き継いだコマンドを入力→死亡判定の操作を繰り返すが、引き継いだコマンド数を超えて生存していた場合、第一世代と同様にコマンド生成フェーズを挟む。以上のようにゲームのプレイを行い、ひとつの個体がゲームオーバーになる度にそのログを保存する。

3.4 遺伝子の評価と選択、交叉

ゲームのプレイにより得たログを元に遺伝子の評価を行う。今回は単純にステージで最も長く進んだ個体、つまり終了時のコマンド数が大きい個体ほど優秀とする。これによりそれぞれの個体に優秀さの序列をつけることができる。この序列を用いて親の遺伝子を決定するが、今回の実験ではエリート選択方式とルーレット選択方式を掛け合わせた形で選択を行う。具体的には、先に親候補となる遺伝子上位 10 体を選択する。その後、選ばれた 10 個体の中からルーレット選択により親となる二個体を選択する。これにより選択された親二個体から二点交叉で二個体、一様交叉で一個体の計三個体を生成する。

ここで二点交叉は複数交叉に分類されるが、複数交叉とは、北野(1991)によると、「交叉位置が複数ある方法である。例えば、交叉位置が 4 と8なら、新たな個体 B の 5 番目から 8 番目まで、個体 A の 9 番目から最後までによって、遺伝子が作られる。同時に、その逆の組み合わせで、もう一つの新たな個体の遺伝子が作られる。」としている。これは図2、図3を用いて説明する。図2が交叉前の遺伝子を持つ個体 A、B であり、交叉位置としてランダムに選択された二点に囲まれた部分を入れ替えることで、新たな二つの遺伝子を持つ個体(図3)を生成する。また、一様交叉とは、北野(1991)によると「交叉時にマスクをかけてそれによってどちらの親の遺伝子を受け継ぐかを決定する方法である。」としている。これは図4で示している。今回はどちらの遺伝子を引き継ぐ確率はどちらも等しく 1/2 としている。

以上の一連の操作を子孫が 30 体になるまで、つまり 10 回繰り返す。これにより次の世代が決定する。

ここで一つ問題となる点がある。それは、遺伝子の長さが一定ではないため交叉をする際に存在しない点で交叉をしてしまうことがある点である。これを解消するために、交叉をするフェーズに入ったタイミングで遺伝子の大きい個体のコマンド数に合わせて、小さい個体にコマンドの確率を条件通りに保った状態で新たなコマンドを追加する。以下で表を用いて説明する。長さ 9 の遺伝子を持つ個体 1(表 1)と、長さ 3 を持つ個体 2(表 2)を親として選んだ場合を考える。

表における矢印はコマンドの入力された向き、S は stay の S でその場から動かないコマンドを指す。また、×は被弾したことによりゲームオーバーになった状態を示している。

この二つの個体で交叉を行う際に、個体 2 においては 4 以降のコマンドが抜け落ちてしまっている。この状態で交叉を行ってしまうと、表 3 のように被弾した後にコマンドが入力されうるという、ありえない遺伝子データが出来上がってしまう可能性がある。これを解消するために、交叉直前の個体 2 の遺伝子に個体1と比

べた不足コマンド数を条件で指定した確率で追加する。つまり、表 4 のように遺伝子の大きさが等しくなるようにコマンドを補充した後に交叉を行う。

これは最初に遺伝子の長さを一定にした状態で交叉することと本質的には変わらない操作であるため、実験上問題にはならない。

最後に、生成された 30 個体に突然変異を施す。突然変異が起きる確率は 1%と設定し、突然変異に選ばれた子孫個体は、ランダムに選ばれたコマンド二つが入れ替わる。

以上の操作により次世代を構成する子孫を生成する。

表 1:個体 1 のコマンド系列

Index	1	2	3	4	5	6	7	8	9
Command	←	S	S	S	↑	→	S	S	↓

表 2:個体 2 のコマンド系列

Index	1	2	3	4	5	6	7	8	9
Command	↑	S	S	×	×	×	×	×	×

表 3:個体 1 と個体 2 を交叉してできた遺伝子の一例

Index	1	2	3	4	5	6	7	8	9
Command	↑	S	S	×	↑	×	S	S	×

表 4:個体 2 にコマンドを補ってできた遺伝子の一例

Index	1	2	3	4	5	6	7	8	9
Command	↑	S	S	S	←	S	S	S	↑

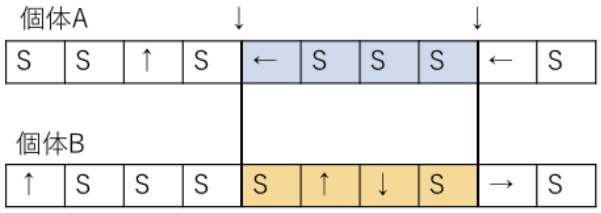


図 2:二点交叉前の個体 A、B

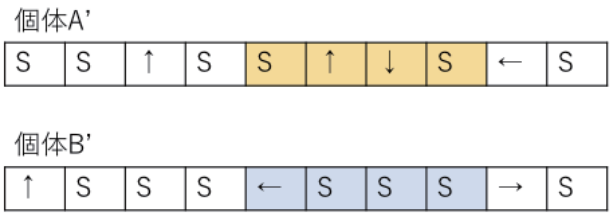


図 3:二点交叉後の個体 A、B

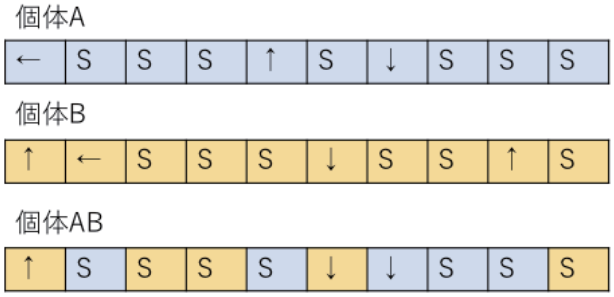


図 4:一様交叉

4. 実験結果

今回は自機が目標とする地点まで到達できるようなコマンド系列が得られることを終了条件とし、終了までに 76 世代を要した。実験により得られた各世代のコマンド数の平均値を図 5 に示した。コマンド数とは遺伝子の長さを指す。

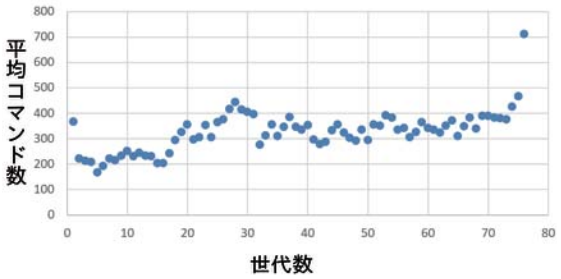


図5:遺伝的アルゴリズムの進行に伴う各世代における平均コマンド数の変化

平均コマンド数を見ることによって学習が進んでいるかを見ることができる。今回の実験においては、平均コマンド数が伸びているならば、よりステージの先に進んでいる個体が多いことを示す。図5から分かるように全体としては徐々に学習していき、右肩上がりになっている。今回の実験で特徴的だった世代としては、第 1 世代、第 28 世代、第 76 世代である。これらの世代では平均コマンド数の局所的なピークが現れている

まず第 1 世代に焦点を当てると、平均値が第 2 世代に比べて高いが、原因としては外れ値として高いコマンド数を持つ遺伝子がいたため、平均値が高くなっていたことが挙げられる。しかし第 2 世代になった途端に一気に平均コマンド数が落ちてしまった。この原因に関してはシューティングゲームの特性が関わっている。シューティングゲームというのは、プレイ中連続的に自機が動いていく、言い換えると、前の瞬間の状況を引き継いだ上で次の行動をとる。つまりこの場合、コマンド一つ一つが独立しているわけではないので、一つ前のコマンドが変わってしまうだけでそこより後のコマンド全てに影響が出てしまう。この特性のために第2世代では平均コマンド数が落ちてしまったと考えられる

第28世代も同様のことが言え、偶然28世代付近では上手くいっていたが、まだ収束するに至っていなかったために次以降の世代では落ち込んでしまったと考えられる。

最後に規定の場所に達した第76世代であるが、これは平均が急激に高くなっている。この理由としては、第75世代における最優秀個体が多く選ばれ、かつ運良く追加されたコマンドが上手く噛み合う形になったからということが考えられる。第69世代から第73世代にかけて横ばいで、かつ第74世代から伸びていることから、この辺りで第73世代までの部分が収束していた可能性が考えられる。遺伝子が収束することで、収束した部分までは同じ動きをするため、その部分までは安定的に到達することができる。従って平均コマンド数の伸びることにつながる。以上のような理由で第76世代において急激に平均コマンド数が伸びたと考えられる。

5. 結論

本研究ではシューティングゲームにおいて敵が打つ弾幕を回避する戦略を遺伝的アルゴリズムを用いて学習させた。

実験の結果、シューティングゲームにおける遺伝的アルゴリズムの適用は有効であった。しかし同時に様々な問題点や課題点が浮上してきた。

一つ目に挙げられるのは、その学習にかかる膨大な時間である。一般にプログラムを動かすのと異なり、60FPS という条件下でコマンドの入力を行うため、ソースコードの有無にかかわらずプレイさせるのに時間を要する。従って、シューティングゲーム、あるいは、より複雑なゲームの AI に適用することを考えると、相応の量の学習時間を見積もる必要がある。それを防ぐためにGUI 上で操作しなくても済むシミュレートツールを作ることが解決策として挙げられる。

次に挙げられる問題点としてはシューティングゲームという特性上遺伝子を組み替えることのリスクが大きいということである。これは 3 でも説明した通り、コマンド一つ一つが独立していなく連続であるがゆえのリスクである。これのために交叉をした際に、一つ前の世代と比較して平均コマンド数が大きく落ちてしまうという現象が見られた。今回は親の選択にエリート選択とルーレット選択を掛け合わせたものを用いたが、やはりこれだけでは良い個体を失ってしまうリスクがあるため、エリート選択で最低でも 1 個体は最優秀のものをそのまま残すという形をとった方が良いと考えられる。

以上のような問題点を解決することができれば学習効率が大幅に上昇し、よりゲーム AI の発展につながるだろう。また、RPG やパズルゲームなどではニューラルネットやディープラーニング等の技術と遺伝的アルゴリズムを組み合わせた AI を組み合わせた AI を作る研究も行われており、シューティングゲームにおいてこれらを利用することも今後の課題である。

謝辞

本研究は, JSPS 科研費 JP16K00228, JP16H02904 の助成を受けたものである。

参考文献

[Marc Ponsen 2004] Marc Ponsen: Improving Adaptive Game AI with Evolutionary Learning , Faculty of Media & Knowledge Engineering Delft University of Technology , 2004.

[伊瀬 顕史 2011] 伊瀬 顕史, 馬野 元秀, 瀬田 和久: Tuning of Fuzzy Rules with a Real-Coded Genetic Algorithm in Car Racing Game, The 27th Fuzzy System Symposium, 2011.

[三宅 洋一郎 2008] 三宅 陽一郎 : How to Use AI Technologies to Develop Digital Games, 人工知能学会誌, 23 巻1号, 2008 年 1 月.

[北野 宏明 1991] 北野 宏明: 遺伝的アルゴリズム, 人工知能学会誌, 7 巻 1 号 , 1991.

[Michael Genesereth 2005] Michael Genesereth, Nathaniel Love, Barney Pell: General Game Playing: AI Magazine, 62-72, 2005.

[David Silver 2016] David Silver, Aja Huang, Chris J. Maddison: Mastering the game of Go with deep neural networks and tree search, Nature 529, pp.484-489, 2016

[Pieter Spronck 2006] Pieter Spronck, Marc PonsenIda Sprinkhuizen-Kuyper: Adaptive game AI with dynamic scripting, Machine Learning, Vol. 63, No. 3, pp. 217-248, 2006.

[川野 洋 2006] 川野 洋: シューティングゲームの敵機攻撃弾発射アルゴリズムに関する考察, 情報処理学会研究報告, 2006.

[片岡 俊 2007] 片岡 俊: 遺伝アルゴリズムを用いたシューティングゲームの攻撃弾発射手法, 高知工科大学 情報システム工学科, 2007.

[Georgios N. Yannakakis 2018] Georgios N. Yannakakis and Julian Togelius : Artificial Intelligence and Games, Springer, 2018.

ZUN: 東方妖々夢, http://www16.big.or.jp/~zun/html/th07.html(2018/03/08)

umiumi, “UWSC”, UWSC, http://www16.big.or.jp/~zun/html/th07.html(2018/03/08)

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

事前知識を活用したMemory Reinforcement Learningによる行動獲得

Action acquisition by Memory Reinforcement Learning useing a prior knowledge

稲盛 有那 *1 Yuna Inamori　平川 翼 *1 Tsubasa Hirakawa　山下 隆義 *1 Takayoshi Yamashita　藤吉 弘亘 *1 Hironobu Fujiyoshi　柏原 良太 *2 Ryota Kashihara
稲葉 正樹 *2 Masaki Inaba　二反田 直己 *2 Naoki Nitanda

*1 中部大学 Chubu University　*2 株式会社デンソー DENSO CORPORATION

Obtaining a human-level control through reinforcement learning (RL) requires massive training. Furthermore, a deep learning-based RL method such as deep Q network (DQN) is difficult to obtain a stable control. In this paper, we propose a novel deep reinforcement learning method to learn stable controls efficiently. Our approach leverages the technique of experience replay and a replay buffer architecture. We manually create a desirable transition sequence and store the transition in the replay buffer at the beginning of training. This hand-crafted transition sequence enables us to avoid random action selections and optimum local policy. Experimental results on a lane-changing task of autonomous driving show that the proposed method can efficiently acquire a stable control.

1. はじめに

Alpha Go[1] が囲碁の世界チャンピオンに勝利して以降，様々な分野で強化学習の活用が期待されている．強化学習は，一般的な教師あり学習と異なり，環境とエージェントのやりとりで得られる報酬を手掛かりに学習する手法であり，未知の環境でも学習が行えるというメリットがある．強化学習は車両やロボットの自律制御 [2] や物体把持 [3] など様々なタスクに応用されている．近年では，Deep Q-Network(DQN)[4] をはじめとする深層強化学習により，複雑なゲームや制御問題を解決することが可能となっている．

深層強化学習による行動獲得は高い性能を示す一方で，複雑な問題に対する学習が難しいという問題が存在する．例えば，DQN のようなニューラルネットワークを用いる手法では，状態を連続値としてネットワークに入力しネットワークの演算を経て行動を出力する．この時，多層のネットワークを用いることで，複雑な問題に対する制御を獲得することが可能となる．しかし，このような多層のネットワークを学習する場合，学習するパラメータ数の増加や勾配消失等により学習が困難となる．また，パラメータはランダムに初期化され学習を開始するが，これはランダムな行動選択から学習を行うことに等しい．特に，時系列行動を強化学習によって獲得する場合，現在の状態に対する制御が過去の制御に依存するため，離散的な行動に比べて困難となる．DQN における学習を安定化させる手法として，過去の行動遷移を蓄積し，蓄積したデータを繰り返しランダムサンプリングして利用する experience replay がある．しかし，ランダムな初期値からの学習では，学習に寄与しない行動選択を蓄積するため，最適な行動を選択するまでに多くの時間を要する．さらに，大規模な状態空間や複雑なタスクに対する学習では局所解に陥りやすく，望ましい行動を獲得することが難しい．

画像認識における Convolutional Neural Network(CNN) では，上記の問題を回避し高い認識性能を実現するために，ImageNet[5] などの大規模な一般物体認識のためのデータセットを用いて学習したモデル (事前学習モデル) のパラメータを初期値とし，個々のタスクに対して学習を行う転移学習が広く用いられている．この事前学習モデルの活用は，タスクが異なった場合でも画像中の特徴が類似していると仮定した上で用いられている．しかし，深層強化学習では問題に応じてその状態空間および入力が異なるため，事前学習モデルの導入が困難である．

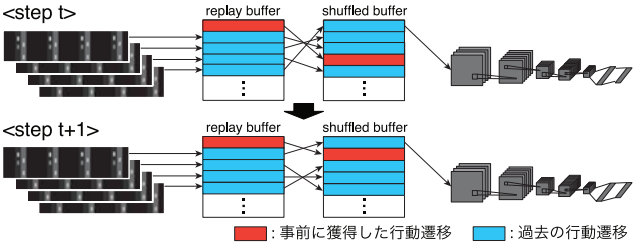
そこで本稿では，事前学習モデルではなく，安定した行動に対する事前知識を学習に導入することで，上記の問題を抑制し，効率的に安定した制御を獲得するための Memory Reinforcement Learning を提案する．提案手法では，事前に獲得した行動遷移を experience replay の replay buffer にあらかじめ保存し，学習を行う．学習時に，replay buffer 内の主導で作成した行動遷移が活用されることで，学習の初期から安定した行動を生成し，効率的に安定した制御を獲得することが可能である．

自動運転車両の車線変更を行うシミュレータを用いた実験にて，提案手法を用いることで安定した動作を効率的に獲得できることを示す．また，replay buffer 内の一部を常に事前知識にして学習した場合および，事前知識のみで学習を行った後に通常の学習を行う場合の性能比較および，事前知識と報酬の組み合わせによる精度の変化について確認する．

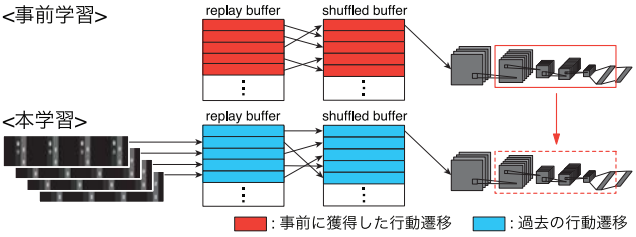
2. 関連研究

強化学習において，効率的に安定した制御を獲得する手法はいくつか提案されている．Peng ら [6] は二足歩行の動作獲得を目的とし，階層的な強化学習手法を提案している．この手法では，目的地まで到達するための長期的な制御と歩行動作を制御する短期的な制御を行う 2 種類の制御システムを用いるこ

連絡先:
稲盛 有那：yuna@mprg.cs.chubu.ac.jp
平川 翼：hirakawa@mprg.cs.chubu.ac.jp
山下 隆義：yamashita@cs.chubu.ac.jp
藤吉 弘亘：hf@cs.chubu.ac.jp
柏原 良太：RYOTA_KASHIHARA@denso.co.jp
稲葉 正樹：MASAKI_INABA@denso.co.jp
二反田 直己：NAOKI_NITANDA@denso.co.jp



(a) 事前知識 A



(b) 事前知識 B

図 1: 提案手法での事前知識の与え方

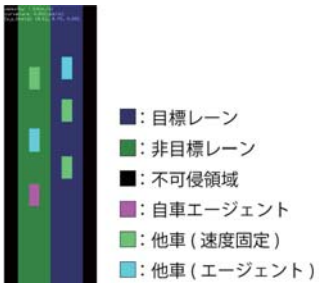


図 2: 環境画像例



図 3: 入力画像例

とで，複雑な地形や環境に対応した歩行動作を獲得している．一方，提案手法では事前に獲得した安定した行動遷移を学習開始時から活用することで，安定した動作を獲得する．

高次元な状態空間における行動を獲得するために，Memory Network[7] を用いて過去の行動を効率的に活用する手法が提案されている [8]．Parisotto ら [9] は長期にわたる過去の記憶を保存し学習に用いるために，Neural Map と呼ばれる手法を提案している．この手法は，3D 環境における強化学習エージェント専用に設計された構造化メモリである．適応可能な書き込み操作を使用しているため，メモリサイズと計算コストは時間の経過と共に増加しない．また，書き込み操作に特定の帰納バイアスを課しており，ナビゲーションが正常な動作の中核要素である 3D 環境に適している．一方，提案手法では単純な過去の行動ではなく，事前に獲得した安定した行動を replay buffer に記憶し学習を行う．

また，事前に獲得した行動を活用する方法として，模倣学習 [10] を導入した手法が存在する [11]．まず模倣学習で他社の行動を真似ることで大まかに望ましい行動を獲得し，その後副報酬を伴う強化学習を行う．模倣学習で得た方策を強化学習による試行錯誤で修正することで，最終的に望ましい行動を獲得する．しかし，模倣学習による行動獲得そのものが難しい問題であり，安定した動作が獲得できなければその後の学習が困難になるという問題が考えられる．一方，提案手法では安定した行動遷移を事前知識として用いるため，不安定な動作の獲得を回避することが可能である．

3. 提案手法

本節では，提案手法について述べる．前述の通り，深層強化学習において安定的な制御を効率よく獲得するために，experience replay を活用する．これは，試行により獲得した状態，行動，報酬のデータを Replay buffer と呼ぶメモリに蓄積する．そして，学習する際に蓄積したデータを繰り返しランダムサンプリングして利用する．通常の強化学習では，学習開始時には Replay buffer の中身は空であるため，一定のエピソードを経て Replay buffer にデータを蓄積させてから学習が始まる．また，初期の Replay buffer 内のデータは殆どランダムな行動をした時のデータである．そのため，膨大な状態と行動の組み合わせから最適な行動を学習するのに時間がかかる．そこで提案手法では，事前知識として望ましい行動遷移を手動で作成し，replay buffer へ格納し，それをもとに学習を開始する．あらかじめ replay buffer に事前知識を蓄積しておくことで，学習の初期から事前知識を用いて学習されるためランダムな行動選択を抑制できる．これにより，効率よく学習を行うことができる．本稿では，作成した事前知識の与え方を 2 種類提案する．図 1 に，事前知識の与え方のイメージを示す．図 1 の replay buffer，shuffled の中の赤色は事前知識，青色は学習中に得た行動遷移を表す．

3.1 事前知識 A

図 1(a) に示すように，replay buffer 内の 1 割に事前知識を常に蓄積しておく．残りの 9 割では従来の replay buffer と同様に，学習が進むにつれデータが入れ替わっていく．これにより，repray buffer 内には常に望ましい行動遷移が蓄積されているため，一定の割合で事前知識を用いた学習を行い，ランダムな行動選択を抑制することが可能となる．

3.2 事前知識 B

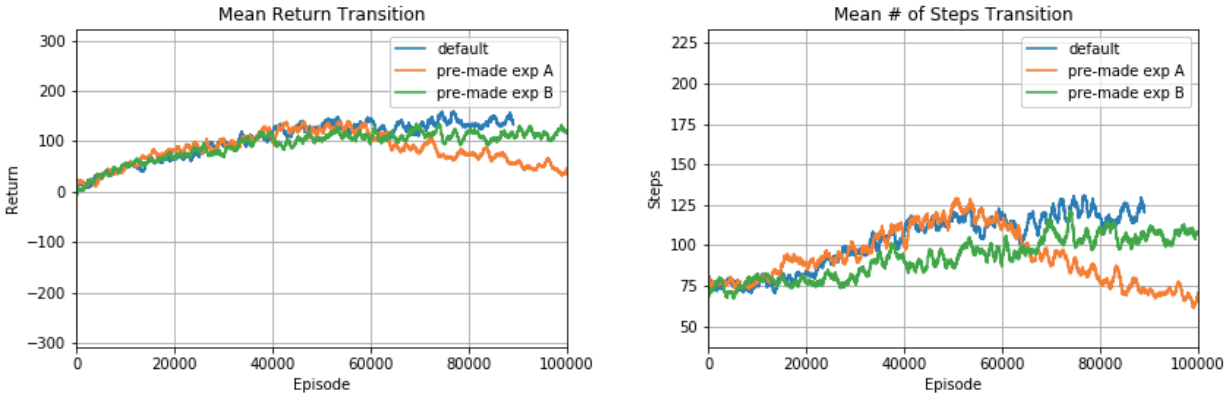
はじめに replay buffer に事前知識として作成した行動遷移データを蓄積しておく．この状態で深層強化学習を行う．この処理を事前学習とする．その後に本学習として，学習したモデルを初期値にして従来の深層強化学習を行う．これにより，事前知識をもとに獲得した行動から，さらに最適な行動を獲得することが期待できる．

4. 評価実験

事前知識の有効性を評価するためのタスクとして，車線変更タスクを用いる．車線変更タスクでは，現在の走行レーンから隣の走行レーンへ移動する．その際，他車が複数台おり，衝突を回避しながら移動しなければならない．また，明確なゴール地点が存在せず，長期の行動を獲得することが必要なタスクである．そのため，事前知識なくランダムな行動から最適な行動を獲得することが困難なタスクとなっている．

4.1 実験概要

図 2 に，車線変更タスクのシミュレータ環境を示す．青色は目標レーン，緑色は非目標レーン，黒色は不可侵領域であり，ピンク色は自車エージェント，黄緑色と水色は他車とする．他車の総数は，学習時は全エピソードで 5 台固定，テスト時はエ



(a) 平均収益　　(b) 平均生存ステップ

図 4: 各エピソードにおける平均収益および平均生存ステップ数の推移

ピソード毎に 0〜5 台のランダムで行う．他車は，固定速度で直進する黄緑色のものと，自車エージェントのモデルをコピーして作られる水色のものの 2 種類存在する．他車が一定速度，または実車と同じエージェントであると，タスクとして簡単になるが，実環境ではドライバの運転スキルにより行動が異なる．そのような環境を想定し，他車に自車エージェントの過去のモデルをランダムに選択して用いる．環境を表した画像から，自車周辺を 84×84 ピクセルに切り出したグレースケール画像 4 フレームを入力とする．図 3 に，入力画像の例を示す．エージェントが可能な行動は，速度 3 種，曲率 3 種を掛け合わせた 9 つである．速度は，$\pm$0m/s（速度維持），+0.1m/s(加速)，$-$.0.1m/s(減速) の 3 種である．曲率は，0(車体角度維持)，+0.01rad/m(左回転)，$-$0.01rad/m(右回転) の 3 種である．これらは，実環境での動作を想定し，実際の車両で制御可能な範囲である．

本実験では，強化学習法として Double DQN[12] を用いる．従来の Double DQN で学習したものをデフォルトとし，事前知識 A，事前知識 B と比較を行う．表 1 に学習条件，表 2 に学習に用いる CNN の構造を示す．

4.2 実験結果

図 4 に，学習時の平均収益，平均生存ステップ数の推移を示す．事前知識 A は，平均収益が 5 万エピソードをピークに途中から低下している．また，平均生存ステップ数も同様に 5 万エピソード付近から低下している．10 万エピソード時点では事前知識 A が最もスコアが低い結果となった．一方，事前知識 B は，平均収益が低下することなく学習できている．また，平均生存ステップ数も同様に低下することなく，増加していることが分かる．

また，表 3 にテストスコアを示す．テスト回数は 100 回であり，成功時に 1 点として計測している．初期レーンは，スタート時の自車エージェントの位置を表す．初期レーンがランダムの場合，初期レーンが必ず非目標レーンである場合の 2 通りでテストを行った．生存スコアは 300 ステップ到達時までに自車が他車や壁に衝突しなかった回数，目標到達スコアは 300 ステップ到達時に自車が車線変更し，目標レーンに達した回数である．人の項目のスコアは，6 名に 100 回ずつテストプレイを行ってもらったスコアの平均である．事前知識を用いないデフォルトは学習後期において，収益，ステップ数ともに最も高くなっている．しかしながら，表 3 に示すようにテストスコアでは事前知識 B よりも低くなっている．これは，最適ではない局所解に陥ったためと考えられる．一方で，デフォルト

表 1: 実験での学習条件

報酬	他車または壁に衝突：-5 衝突していない場合は下記を合算 車線中央付近：+1 停止：-1 目標車線：+1 非目標車線：-2
エピソード終了条件	衝突または 300 ステップ経過
学習エピソード数	最大 100000 エピソード
割引率	0.95
探索法	Linaer decay epsilon greedy
Replay buffer サイズ	$1e+5$

表 2: CNN の構造

	処理	詳細
入力層	サイズ	$4 \times 84 \times 84$
畳み込み層 1	活性化関数	ReLU
	フィルタ	$3 \times 3 \times 16$
	maxpooling	2×2
畳み込み層 2	活性化関数	ReLU
	フィルタ	$3 \times 3 \times 16$
	maxpooling	2×2
畳み込み層 3	活性化関数	ReLU
	フィルタ	$3 \times 3 \times 16$
	maxpooling	2×2
全結合層	活性化関数	ReLU
	ユニット数	256
出力層	ユニット数	9

と比較して，事前知識 A は生存スコアが同等であるが，目標到達スコアが低い．これは，事前知識として与えた経験に頼りすぎ，新たな経験の獲得が十分にされていないことが考えられる．事前知識 B はデフォルトおよび事前知識 A と比べて，高いスコアとなっている．事前知識 B では学習初期は事前知識のみで，学習が進むにつれ replay buffer 内は新しい経験に入れ替わる．しかし，初期で既に人に与えられた行動遷移を得ており，ランダムに参照を行なった場合でも全てに事前知識が導入されているため，さらによい行動を獲得し，テストスコアが上昇したと予想される．したがって，事前知識からタスクで実行するのに有効な基本行動を獲得し，それをもとにさらにタスクに適した行動を獲得できていると考える．

また，人のスコアと比較しても，事前知識 B は生存スコアは約 32%，目標到達スコアは約 31%高い結果となっている．このことから，事前知識 B の手法は車線変更に伴う車体制御が

表 3: テストスコアの比較

	初期レーン	生存スコア	目標到達スコア
デフォルト	ランダム	53	50
	非目標	33	29
事前知識 A	ランダム	49	37
	非目標	37	10
事前知識 B	ランダム	**67**	**66**
	非目標	**51**	**48**
人	ランダム	50.7	50.3

安定しており，効果的であると言える．

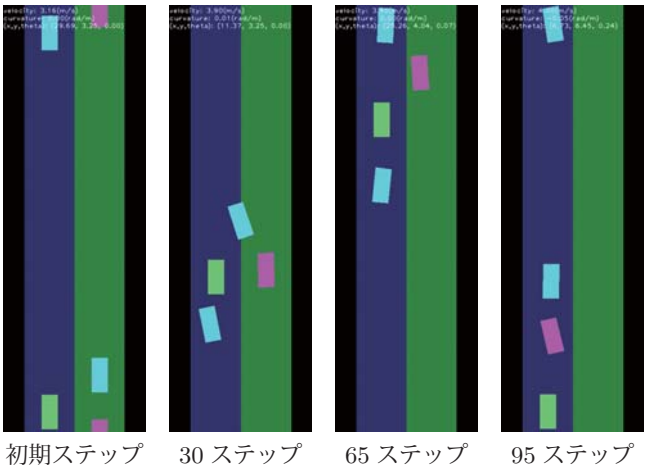
4.3 考察

本実験でテストスコアが最も高かった事前知識 B について，テストシーンにおける動作例を図 5 に示す．図 5(a) では，他車が 3 台いるシーンにおいて，スタート時に自車の前方を走行する他車が車線変更を試みている．その際，自車は他車との衝突を避けるために加速せずに追従している．そして，他車が車線変更後，他車間の間隔が十分に空いたところにスムーズに入る行動を取ることが出来ている．また，図 5(b) では，他車が 3 台おり，隣の車線に 2 台走行している．このようなシーンにおいて，自車は加速して目標車線にいる他車を追い抜き，充分な車間距離になってから車線変更を試みている．このように他車との速度と車間距離を考慮することが出来ている．これらの結果より，提案手法では，人が実環境で車線変更を行う際に取りうる動作を深層強化学習により獲得することができた．

5. おわりに

本稿では，深層強化学習において効率的に安定した制御を獲得するための Memory Reinforcement Learning を提案した．提案手法では，experience replay の replay buffer を活用し，事前に獲得した行動遷移を replay buffer に保存し学習を開始する．これにより，学習初期のランダムな行動選択を抑制し，不安定な行動の獲得を回避することが可能である．自動運転車両の車線変更タスクを用いた実験にて，事前に獲得した行動遷移を用いずに学習した場合と比較し，安定した制御を効率的に獲得できることを確認した．

今後の課題として，事前知識を効果的に活用する方法の検討や事前知識を踏まえた報酬の設計などが挙げられる．



(a) 他車に追従してから車線変更する例

(b) 他車を追い抜き車線変更する例

図 5: 事前知識 B のテストシーンにおける動作例，青:目標レーン，緑色:非目標レーン，黒:不可侵領域，ピンク:自車，黄緑・水色:他車を表している．

参考文献

[1] D. Silver, et al., “Mastering the game of Go with deep neural networks and tree search,” Nature, Vol. 529, pp. 484–489, 2016.

[2] T.P. Lillicrap, et al., “Continuous control with deep reinforcement learning,” In ICLR, 2016.

[3] L.Pinto, *et al*. Learning Hand-Eye Coordination for Robotic Grasping with Deep Learning and Large-Scale Data Collection. In IEEE, 2016.

[4] V. Mnih, *et al*. Human-level control through deep reinforcement learning. Nature 518, pp.529–533, 2015.

[5] J. Deng, et al., “ImageNet: A large-scale hierarchical image database,” In CVPR, 2009.

[6] X.B.Peng, *et al*. DeepLoco: Dynamic Locomotion Skills Using Hierarchical Deep Reinforcement Learning. ACM Transactions on Graphics, Vol. 36, No. 4, Article 41, 2017.

[7] S. Sukhbaater, et al., “End-to-end memory networks,” In NIPS, 2015.

[8] J. Oh, et al., “Control of Memory, Active Perception, and Action in Minecraft,” In ICML, 2016.

[9] E.Parisotto, et al., “Neural Map: Structured Memory for Deep Reinforcement Learning,” In ICLR, 2018.

[10] E. Parisotto, et al., “Imitation Learning: A Survey of Learning Methods,” ACM Computing Surveys, Vol. 50, No. 2, pp. 21:1–21:35, 2017.

[11] 田渕一真ら, “模倣学習と強化学習の調和による効率的行動獲得,” 人工知能学会全国大会論文集, pp. 212–215, 2016.

[12] H. Hassel, et al., “Deep Reinforcement Learning with Double Q-Learning,” In AAAI, 2016.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

自己学習を用いたニューラル見出し生成

Neural Headline Generation with Self-Training

竹前 慎太郎 *1 Shintaro Takemae　村尾 一真 *2 Kazuma Murao　谷塚 太一 *2 Taichi Yatsuka　小林 隼人 *2*3 Hayato Kobayashi　野口 正樹 *2 Masaki Noguchi　西川 仁 *1 Hitoshi Nishikawa
徳永 健伸 *1 Takenobu Tokunaga

*1東京工業大学 Tokyo Institute of Technology　*2ヤフー株式会社 Yahoo Japan Corporation　*3理化学研究所 AIP センター RIKEN AIP

In this paper we propose a novel method which incorporates self-training into a sequence-to-sequence model in order to improve the accuracy of the headline generation task. Our model is based on neural network-based sequence-to-sequence learning with an attention mechanism and trained with approximately 100,000 labeled examples and 2,000,000 unlabeled examples. Through experiments, we show our proposal significantly improves the accuracy and works effectively.

1. はじめに

自動見出し生成は, 自動要約研究における重要な課題のひとつである [Nenkova and McKeown 2011]. 見出し生成課題は一般的に, 新聞記事などの文章に対して, その文章の内容をより少ない文字数で表す文を生成する課題である. 自動要約によって作成された見出しを提供することで, 読み手は少ない文字数の見出しから記事の内容を把握し, 読むべきかどうかをより迅速に判断できるようになるため, この課題は実用上非常に有用である. 見出し生成における一般的な設定では記事に付属しているタイトル（記事タイトル）が生成されるが, 本論文では, あるニューストピックを指し示すための, 記事タイトルよりもさらに短い文字数の見出し（トピックス見出し）を生成する方法を提案する. Yahoo!ニューストピックスのような総合ニュースサイトにおいては, 記事タイトルの全文字を描画領域に表示できないことが頻繁に起こるため, 記事タイトルよりも短く, そのトピックにおける重要な情報を含む見出しを生成することができれば, 読み手の負担を減らし効率よく記事を提供できるようになる.

一方で, 近年, 自動要約システムは機械学習を用いた手法に大きく依存している. 機械学習の枠組みにおいては, 訓練データ, すなわち入力とそれに対応する出力の組を準備することが必要である. しかしながら, 出力に相当するトピックス見出しは記事タイトルや内容から編集者の手によって作成されており, その作業コストが高いことからこのようなデータセットを十分に用意することは困難である. 学習のためのデータが不足しがちという問題点を克服するため, 本論文では自己学習によって学習データを拡張する手法を提案する. 提案手法では, まず最初のモデルを人間の手によって作成された学習データセットから作成し, 続いて, 記事タイトルのみを持つデータに対して, このモデルを用いて疑似トピックス見出しを付与する. その後, 疑似トピックス見出しが付与されたデータと, 人間がトピックス見出しを作成したデータの両方のデータセットを学習データとしてモデルの再学習を行う. 比較実験により, 提案手法がベースライン手法よりも統計学的に有意に高い精度を達成することを確認した.

連絡先: *1{takemae.s.aa@m, hitoshi@c, take@c}.titech.ac.jp
*2{kmurao, tyatsuka, hakobaya, manoguch}@yahoo-corp.jp

2. 関連研究

自動要約についての手法は, 抽出的要約と生成的要約の大きく 2 つに分類される. 抽出的要約では, システムは入力文章の一部分を抜き出し, それを組み合わせることで出力文を構成する. 生成的要約では, 入力文書に含まれない全く新しい表現を含む出力文を生成する. 後述するように, 編集者がトピックス見出しを作成するときにはしばしば入力文にない表現を用いることがあるため, 本研究課題では生成的要約のアルゴリズムを適用することが望ましいと考えられる.

見出し生成については多くの研究がなされており [Banko et al. 2000, Knight and Marcu 2002, Alfonseca et al. 2013], 近年ニューラルネットワークを用いる手法が活発に研究されている [Rush et al. 2015, Filippova et al. 2015, Takase et al. 2016, Ayana et al. 2016, Xu et al. 2016, Chopra et al. 2016, Zhou et al. 2017, Tan et al. 2017, Tilk and Alumäe 2017]. ニューラルネットワークを用いたアルゴリズムでは主に, 文章の単語分散表現の系列を入力とし, encoder-decoder 型のネットワークにアテンション機構を取り入れた構成によって出力文を生成する.

自己学習は, 自然言語処理において広く用いられており, 構文解析 [McClosky et al. 2006] や語義曖昧性解消 [Mihalcea 2004] などに利用されている. sequence-to-sequence の研究においても, Dai と Le によって半教師あり学習の手法が提案されている [Dai and Le 2015]. 彼らは, 前段の教師なし学習によって後段の教師あり学習におけるより良い初期パラメータを計算する事前学習手法を提案した. 彼らの手法と異なり, 本論文の提案手法では, 正解付きのデータセットを用いて前段のモデル学習を行うが, その後, 正解を持たないデータセットに対してこのモデルから疑似正解を付与し, これらの正解付き・疑似正解付きの両方のデータセットを利用して学習することで最終的なモデルを獲得する. 前段の教師あり学習によって作成されるモデルと, 後段の自己学習を経たモデルを比較した結果, 後者のモデルが前者のモデルより高い精度

記事タイトル	トピックス見出し
＜天皇陛下＞退位後に「赤坂御用地」に転居　宮内庁検討	陛下 退位後は赤坂に転居案

表 1: 記事タイトルと対応するトピックス見出しの例 *4

を達成することを確認した.

3. 提案手法

本論文の提案手法は以下のような手順による.

1. 教師あり学習 (ベースライン): ベースラインのモデルは, 人間の作成したトピックス見出しと記事タイトルの組から学習される.
2. 疑似見出しの生成: 次に, 記事タイトルのみからなるデータセットに対して, 手順 1 で作成したモデルを用いて疑似トピックス見出しを作成する.
3. 提案モデルの学習: 最後に, 手順 1 のオリジナルデータと, 手順 2 で疑似トピックス見出しが付与されたデータの両方を用いて, 提案モデルの学習を行う.

言い換えると, 提案手法においては, ベースラインのモデルから新しいデータセットを作成することで自己学習を行っている.

4. 実験

4.1 実装と設定

ベースラインのモデルを作成するため, OpenNMT ツールキット [Klein et al. 2017] を利用した. encoder-decoder モデルのネットワークとして, Luong らによって提案された, stacking recurrent neural network に global attention を加えた構成を用いている [Luong et al. 2015].

encoder の入力には 3000 次元の one-hot ベクトルを埋め込んだ分散表現を用い, 250 次元の内部ベクトルを持った 3 層の隠れ層による双方向 LSTM を採用した. decoder においても, 出力した語の再入力に同様の分散表現を用いたが, 3 層の隠れ層の内部ベクトルの次元数は 500 とした. 学習時の最適化手法として Adam [Kingma and Ba 2014] を採用し, そのハイパーパラメータはそれぞれ, $\beta_1 = 0.9$, $\beta_2 = 0.98$, $\varepsilon = 10^{-9}$, $\alpha = 10^{-3}$ とした. 前段の学習と後段の学習はそれぞれ 30 エポックずつ行い, そのミニバッチサイズは 64, dropout 率は 0.3 とした.

本研究では対象とする見出しが非常に短いことから, 分散表現として単語単位ではなく文字単位の入力を用いた. すなわち, システムは記事タイトルを 1 文字ずつ読み込み, その出力として見出しを 1 文字ずつ生成する.

また, 出力時には探索幅 5 によるビームサーチを用いて見出しを生成した.

4.2 比較手法

以下の 3 手法を比較した.

- ベースライン手法: 前述の通り, アテンション機構付きの一般的な encoder-decoder 型モデルをベースラインとして採用した.
- 事前学習手法: 他の半教師あり学習との比較のため, Dai と Le が提案した sequence autoencoder に基づく事前学習手法 [Dai and Le 2015] を比較対象とした. この手法では, 正解見出しなしデータを用いた sequence autoencoder モデル (記事タイトルから記事タイトルへの sequence-to-sequence モデル) の学習を行った後, そのパラメータで初期化したモデルの学習を行う. このとき初期化するパラメータの対象を全てのネットワークパラメータとする場合と, アテンション機構を除く encoder-decoder 部分のみのパラメータを初期化する場合について実験を行った. なお, Dai と Le はこの事前学習手法を感情分析課題で評価している. そのため, 事前学習手法が想定していた, 事例を正負に分類する単純な分類問題と比較して, 本研究課題のほうが複雑であることを考慮に入れる必要がある.
- 提案法: 正解見出しを持つデータと, 疑似正解見出しを付与したデータの双方を学習データとして学習を行った.

4.3 データ

実験用のデータセットとして, 日本語ニュース記事タイトルと, それに対応するトピックス見出しの組を用いた. 表 1 にその例を示す. トピックス見出しにおいては全体の長さをコンパクトにまとめるため, 記事のタイトルに含まれているいくつかの語を省いたり, 言い換えなどが行われている.

クローリングによって, 記事タイトルのみからなるデータセットを準備した. このうち, トピックス見出しと紐付けられた約 10 万組を教師ありデータとし, 紐付けをしなかったものを自己学習用のデータとした. 教師ありデータは, 約 60%を学習データ, 約 20%を開発データ, 約 20%をテストデータとして分割した. 自己学習を経たモデルの精度にデータセットのサイズが与える影響を確認するため, 自己学習用データからそれぞれ 20 万件, 40 万件, 80 万件, 100 万件, 150 万件, 200 万件をサンプリングしたデータセットを作成した.

トピックス見出しは全て 13.5 文字以内で記述されている. 日本語のひらがな, カタカナ, 漢字はそれぞれ 1 文字として数えられ, 英語のアルファベット, 数字および空白文字については 0.5 文字換算として数えられる. すなわち, 編集者は 13 文字の日本語の文字と空白文字 1 つによってトピックス見出しを作成することができ, またアルファベットや数字を用いる場合はそれらの 2 文字を日本語の 1 文字分と換算することができる.

4.4 評価尺度

評価尺度として, 文字ベースの ROUGE-1 指標を用いた [Lin 2004]. これは比較対象のシステムが単語ではなく文字を単位として出力するためである. データセット中の記事タイトルとトピックス見出しは非常に短く, トピックス見出しはしばしば言い換えが用いられている. 単語ベースの ROUGE は言い換え表現を過剰に低く評価するが, 文字ベースで評価すると, 元の単語を構成する文字を抽出したような略語についてはその語を正解に含まれるものとして評価できる. このように, 文字ベースの評価を行うことで, 単語ベースの評価よりもより正確な生成文字列の評価が可能であると考えられる.

また評価において初期値による影響を軽減するため, それぞれの手法に対し, 異なる初期値による 10 回の学習とトピック

*4 https://news.yahoo.co.jp

	ROUGE-1
ベースライン	0.568
事前学習（アテンション機構除く）	0.520
事前学習（全て）	0.503
提案手法（60 万件）	**0.574**

表 2: ROUGE-1 の平均

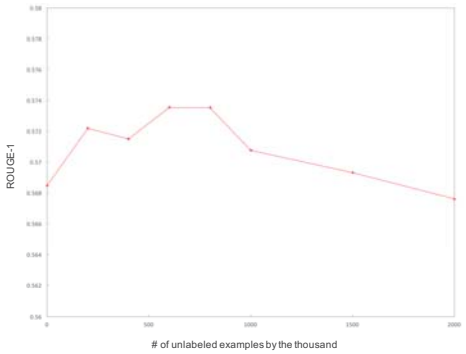


図 1: 正解なしデータのサイズに対する提案手法の ROUGE-1

ス見出しの生成を行い, その評価値の平均をその手法のスコアとして採用した.

5. 結果と考察

表 2 に実験結果を示す. 提案手法のスコアは, 60 万件の自己学習用データセットを用いた場合のモデルを採用した. 提案手法によるモデルは, 自己学習を行わないベースラインモデル, および事前学習を行ったモデルより高い精度を達成した. t 検定により, 提案法と比較手法の間の ROUGE-1 スコアは統計学的に有意であることが示された ($p < 0.05$).

事前学習手法はベースラインよりも低い精度となった. 全てのパラメータを事前学習する手法と比較して, アテンション機構のパラメータを事前学習による初期化の対象から除く手法が良い精度を達成した. このことは, アテンション機構が, 入力された記事タイトルに含まれる情報のうち, どのような情報をトピックス見出しに残すべきかといった, 教師なし学習からは本質的に学習できないパラメータを学習しており, そのため正解のないデータからは望ましい初期値を得られなかったことを示唆している.

図 1 に, 自己学習用のデータセットのサイズに対する提案法の精度の変化を示した. データセットのサイズが 60 万件, および 80 万件のときスコアは 0.574 と最も高い値を記録した. データセットのサイズが小さいとき, もしくは大きすぎるときに精度が低下する現象は, 正解なしデータが多くなりすぎることで, 本来学習すべき対象である正解ありデータよりも大きな影響を持ってしまうことで生じると考えられ, 自己学習ではしばしば確認される. そのため, 自己学習用に疑似正解を付与する際に望ましい出力例のみを採用することができれば, 自己学習による精度向上はより大きくなると考えられる.

表 3 に, 提案法と比較手法による出力文字列の長さ (文字ベース, 半角英数字も 1 文字として数える) を掲載した. 要約率は約 50%で, 機械によって生成された見出しの長さはどれも, 編集者が作成したものと比較して若干短い傾向にあった.

図 4 に, 評価データに対するモデルの出力例をいくつか掲載した. 1 番目の例では, 記事タイトル中の「割引切符」という名詞を, モデルによる出力で「割符」と略語化している. 提案手法においては単語単位ではなく文字単位で処理が行われるため, 複合語が再構成されて出力されることがある. この出力例に対し ROUGE 値は高い値を示すが, 「割引切符」と「割符」は異なるものを指し示すため, この出力例から正しい意味を読み取ることは難しい.

	文字数
本文	595.6
記事タイトル	26.4
トピックス見出し	14.0
ベースライン	13.6
事前学習（アテンション機構除く）	13.3
事前学習（全て）	12.8
提案手法（60 万件）	13.5

表 3: 平均文字数

2 番目の例では, ベースラインモデルは「ロボット」という語を反復して出力し, この見出しで正しい記事内容を説明することは困難である. 一方で, 提案手法の出力では「ソフトバンク」という文字列を「ソフト B」のようにうまく略語化しているように見えるが, 正解トピックス見出し中に「ソフトバンク」に相当する語がなかったため ROUGE 値はベースラインモデルを下回った.

3 番目の例においては, 正解の見出しは記事タイトルの削除操作で生成することができ, 学習モデルはこれを模倣している. Filippova らの先行研究によって, ニューラルネットワークによる文圧縮モデルは, データセット中に存在する正解文字列が全て元文字列の削除操作のみから生成することができる場合に比較的良い結果を見せることが報告されている [Filippova et al. 2015].

4 番目の例のように正解見出しが記事タイトルと全く異なっていると, モデルが生成した見出しがある程度意味をとれるものであっても, 正解見出しと共通する文字を持たないことから ROUGE-1 が低い値をとる場合がある.

最後の例では, 提案手法による出力はベースラインモデルのものよりも望ましい出力を得た. ベースラインモデルは記事の重要な要素とみられる「試験」という語を生成見出しから欠落させたが, 提案法ではこれを残したままトピックス見出しを生成できた.

全般的に, 人間の編集者によって作成されたトピックス見出しは洗練された言い換え表現を含んでおり, これらはしばしば文字単位で置き換えられている. 正解のトピックス見出しが入力文の削除操作のみから生成することができる場合には, ベースラインモデルと提案モデルの双方から比較的良い出力結果を得ることができた.

6. まとめ

本論文では, アテンション機構付き sequence-to-sequence のモデル学習に, 自己学習を追加する学習手法を提案した. 実験により, 自己学習を追加して行うことで学習モデルの精度が有意に向上したことが示された.

5 節で議論したように, 自己学習のために生成された疑似正解見出しの中から品質の高い組のみを選ぶことができれば, より自己学習の効果が高まると考えている. また, 人間の編集者はトピックス見出しを作成する際に記事タイトルだけでなく記事の内容などを考慮に入れていることから, モデルにおいても記事リード文などの情報を考慮することで生成見出しの品質を高めていくことを検討している.

記事タイトル	トピックス見出し	ベースライン	提案手法
JR 東日本の割引切符に若者からため息，高齢者優遇の背景は？	JR のシニア優遇に若者ため息	JR 東の割引切符高齢者の背景 (0.462)	JR 東割符に若者からため息 (0.615)
ソフトバンクがロボット事業に参入，日本のロボット産業どうなる？	日本のロボット産業 未来は?	日本ロボットロボットに参入 (0.462)	ソフト B ロボット事業に参入 (0.385)
＜ボクシング＞亀田ジムが新会長での再出発を検討	亀田ジム 新会長で再出発検討	亀田ジム新会長で再出発検討 (1.000)	亀田ジム新会長で再出発検討 (1.000)
カーボヴェルデ代表が世界ランクで日本を追い抜いた理由	FIFA ランク 無名国がなぜ躍進	カーボヴェルデ代表日本追い抜く (0.000)	カーボヴェルデ代表日本追い抜く (0.000)
ドワンゴの入社試験有料化，その是非をどう考えるべきか	入社試験「有料化」の是非は	ドワンゴ有料化その是非どう (0.462)	ドワンゴの入社試験有料化へ (0.615)

表 4: 出力例 と ROUGE-1 *5

参考文献

[Alfonseca et al. 2013] Alfonseca, E., Pighin, D., and Garrido, G. (2013). Heady: News headline abstraction through event pattern clustering. In *Proceedings of the 51st Annual Meeting of ACL*, pages 1243–1253.

[Ayana et al. 2016] Ayana, Shen, S., Liu, Z., and Sun, M. (2016). Neural headline generation with minimum risk training. *CoRR*, abs/1604.01904.

[Banko et al. 2000] Banko, M., Mittal, V. O., and Witbrock, M. J. (2000). Headline generation based on statistical translation. In *Proceedings of the 38th Annual Meeting on ACL*, pages 318–325.

[Chopra et al. 2016] Chopra, S., Auli, M., and Rush, A. M. (2016). Abstractive sentence summarization with attentive recurrent neural networks. In *Proceedings of the 2016 Conference of NAACL-HLT*, pages 93–98.

[Dai and Le 2015] Dai, A. M. and Le, Q. V. (2015). Semi-supervised sequence learning. In *Advances in Neural Information Processing Systems 28*, pages 3079–3087.

[Filippova et al. 2015] Filippova, K., Alfonseca, E., Colmenares, C. A., Kaiser, L., and Vinyals, O. (2015). Sentence compression by deletion with lstms. In *Proceedings of the 2015 Conference on EMNLP*, pages 360–368.

[Kingma and Ba 2014] Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980.

[Klein et al. 2017] Klein, G., Kim, Y., Deng, Y., Senellart, J., and Rush, A. M. (2017). Opennmt: Open-source toolkit for neural machine translation. *CoRR*, abs/1701.02810.

[Knight and Marcu 2002] Knight, K. and Marcu, D. (2002). Summarization beyond sentence extraction: A probabilistic approach to sentence compression. *Artificial Intelligence*, 1(139):91–107.

[Lin 2004] Lin, C.-Y. (2004). Rouge: A package for automatic evaluation of summaries. In *Proceedings of ACL Workshop Text Summarization Branches Out*, pages 74–81.

[Luong et al. 2015] Luong, M.-T., Pham, H., and Manning, C. D. (2015). Effective approaches to attention-based neural machine translation. In *Proceedings of the 2015 Conference on EMNLP*, pages 1412–1421.

[McClosky et al. 2006] McClosky, D., Charniak, E., and Johnson, M. (2006). Effective self-training for parsing. In *Proceedings of the Main Conference on NAACL-HLT*, pages 152–159.

[Mihalcea 2004] Mihalcea, R. (2004). Co-training and self-training for word sense disambiguation. In *Proceedings of the Eighth Conference on CoNLL*, pages 33–40.

[Nenkova and McKeown 2011] Nenkova, A. and McKeown, K. (2011). *Automatic Summarization*. Now Publishers.

[Rush et al. 2015] Rush, A. M., Chopra, S., and Weston, J. (2015). A neural attention model for abstractive sentence summarization. In *Proceedings of the 2015 Conference on EMNLP*, pages 379–389.

[Takase et al. 2016] Takase, S., Suzuki, J., Okazaki, N., Hirao, T., and Masaaki, N. (2016). Neural headline generation on abstract meaning representation. In *Proceedings of the 2016 Conference on EMNLP*, pages 1054–1059.

[Tan et al. 2017] Tan, J., Wan, X., and Xiao, J. (2017). From neural sentence summarization to headline generation: A coarse-to-fine approach. In *Proceedings of IJCAI-17*, pages 4109–4115.

[Tilk and Alumäe 2017] Tilk, O. and Alumäe, T. (2017). Low-resource neural headline generation. *CoRR*, abs/1707.09769.

[Xu et al. 2016] Xu, L., Wang, Z., Ayana, Liu, Z., and Sun, M. (2016). Topic sensitive neural headline generation. *CoRR*, abs/1608.05777.

[Zhou et al. 2017] Zhou, Q., Yang, N., Wei, F., and Zhou, M. (2017). Selective encoding for abstractive sentence summarization. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1095–1104. Association for Computational Linguistics.

*5 JR 東日本は東日本旅客鉄道株式会社の商標または登録商標です. ソフトバンクの名称は, 日本国およびその他の国におけるソフトバンクグループ株式会社の登録商標または商標です. ドワンゴは株式会社ドワンゴの商標または登録商標です. その他、本論文中に記載されている会社名および商品・サービス名は、各社の商標または登録商標です.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

画像認識を用いたATMセキュリティ強化

Security Reinforcement for ATM by Using Image Sensing

岸礼子[*1] ファンチョンフィ[*1] 山本一真[*1] 増田誠[*1]
Reiko Kishi Phan Trong Huy Kazuma Yamamoto Makoto Masuda

[*1] 沖電気工業株式会社 経営基盤本部 研究開発センター
Corporate Research and Development Center, Oki Electric Industry Co., Ltd.

ATM (Automated Teller Machine)-related crimes have been reported as serious problems overseas. One of such crimes is *skimming*: the act of installing skimming device on ATM to illegally copy information of cash cards or passwords. Detecting such illegal acts promptly is important since we can immediately prevent ATM-related crimes by automatically raising alarms and also reduce supervisor's monitoring burden. We have realized an anomaly behavior detection technique which finds not only suspicious devices but also detects actions of installing them. In this article, we discuss our proposed method which combines video feature extraction, ATM status acquisition and recognition. Compared to processing only image information, the proposed method proves to be more effective in recognizing anomaly behaviors from normal behaviors.

1. はじめに

海外では ATM（Automated Teller Machine）利用者のカード情報や暗証番号情報を盗み取るスキミング等の犯罪が深刻な問題となっている。このような犯罪を実行時に検知することは、その場での警報や通報による犯罪抑止や監視員の負担軽減につながるため重要である。しかし近年は犯罪装置の小型化が進み、ATM 外観からは犯罪装置の識別が困難となっている。そこで我々は犯罪装置だけでなくそれを取り付けている行為も検知する技術を実現した。

本稿では、動画像処理による異常行動検知手法に“ATM 状態把握”を組み合わせた ATM 異常行動検知手法を提案する。さらに模擬データを用いて提案手法の有効性を示す。

2. 動画像処理による異常行動検知手法

本章では、一般的な動画像処理による異常行動検知手法を紹介し、ATM へ応用する際の課題を述べる。

2.1 一般的な動画像処理による異常行動検知

動画像処理による異常行動検知手法は一般的に、“動画像特徴抽出”と“識別処理”が用いられる。

[南里 2005]は、動画像特徴に立体高次局所自己相関（CHLAC）を、識別処理には部分空間法を用いている。CHLAC は画素値を 2 値化した連続する 3 フレーム間の局所領域を 251 個のパターンに当てはめる手法である。異常行動が正常行動の中に混ざっても検知可能（加法性）で、かつ、検知性能は画像内の位置に依らない（位置不変性）という利点を持っている。部分空間法では、特徴空間上に正常行動の部分空間を構成し、入力フレームの特徴量とその部分空間との距離を異常度として扱っている。これにより、複数人歩行者の中での転倒動作を異常として検知し、手法の有効性を示している。

[弓場 2012]は、動画像特徴として Space-Time Patch（ST-Patch）、識別には Continuous Rank-Increase Measure（CRIM）を用いている。ST-Patch は、動画像の小領域内の濃度勾配値の統計量から求められる。勾配値は上下左右の画素に加え連続するフレーム間からも求めるため、特徴量は「見え」と「動き」の情報を併せ持っている。CRIM は、ST-Patch から計算できる非定常度の値で、この値が閾値を超えたときに異常とみなしている。ST-Patch と Real AdaBoost と呼ばれる機械学習手法によって人物領域検出を行い、Real AdaBoost の信頼度を用いて人物領域内の人の部分に重み付けを行った CRIM で異常検知を行っている。それによって、エスカレーターでの転倒を対象とした動的な背景下での異常行動検知手法を提案している。

その他にも異常行動検知に用いられる手法には、動画像特徴として、[林 2006]では Optical Flow、[Kobayashi 2012]では STACOG、[弓場 2012]では MHI などがあり、識別処理としては、Linear Discriminant Analysis（LDA）、Support Vector Machine（SVM）、Gaussian Mixture Models（GMM）などがある。

2.2 ATM 異常行動検知での課題

我々は動画像処理による異常行動検知を ATM に対する犯罪の検知に応用する。監視カメラは ATM の上部に設置し、手元の細かい動作がわかるように、操作部を中心に撮影する（図 1）。

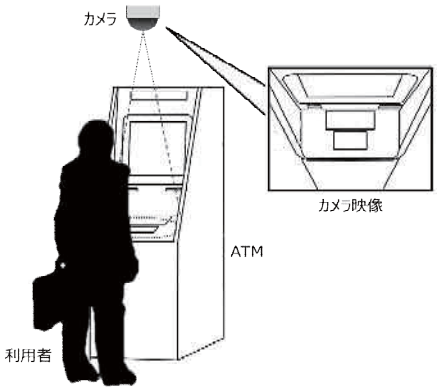


図 1: 撮影イメージ

連絡先: 岸礼子, 沖電気工業株式会社 経営基盤本部 研究開発センター, 埼玉県蕨市中央 1-16-8, kishi422@oki.com

「引き出し」「預け入れ」などの ATM に対する通常操作が正常行動となる。正常行動は動作速度に個人差があり、さらに端末上に手を置く・持ち物を持つなど操作以外の動作が含まれる。また、異常行動はカードを挿入口へ差し込むふりをして犯罪装置を設置するなど、正常行動に似せた犯罪行動が起こり得る。このように正常行動と異常行動に多様性があるため動画像処理のみでの識別は十分な精度を期待できない。

3. 提案手法

我々は“ATM 状態把握”を導入し、動画像特徴量抽出と ATM 状態把握および識別処理による異常行動検知手法を提案する。本手法は異常検知辞書を作成する学習フェーズと、異常検知辞書を用いて異常行動を検知する検知フェーズに分かれる。本章では、“ATM 状態把握”の概要、学習フェーズと検知フェーズについてそれぞれ説明する。

3.1 ATM 状態把握

“ATM 状態把握”とは、「無人状態」「暗証番号入力状態」などといった ATM の取引状態(以下、ATM 状態)を把握することである。ATM 状態把握により、起こり得る正常行動を限定できることが大きな利点である。図 2 のように、無人状態から利用者が取引を行って立去るまでには ATM 状態は多様に変化する。例えば「暗証番号入力状態」では暗証番号キーに対する動作が発生するが、それ以外の ATM 状態時に暗証番号キー周辺に動作があれば警戒する必要がある。「近接状態」は取引開始前または終了後のことである。そのため ATM 操作のような複雑な動きは起こらないはずであり、複雑な動作がある場合には警戒する必要がある。このように、ATM 状態把握によって起こり得る動作を絞ることで、異常行動検知の識別性能の向上が可能になる。

ATM 状態は、ATM 自体から通知される内部状態の情報を用いて把握する。

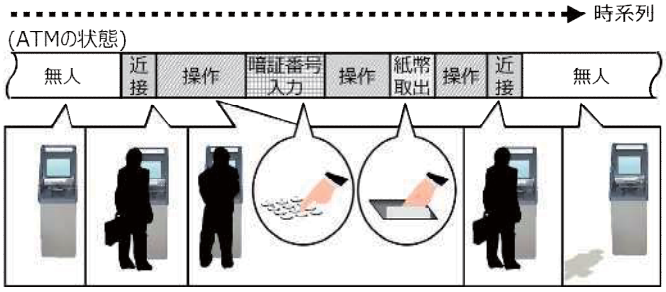


図 2:ATM 状態が変化していく例

3.2 学習フェーズ

学習フェーズでは、正常行動のみを用いて機械学習を行い ATM 状態毎の異常検知辞書を生成する。

図 3 に提案手法の流れを示す。まず、カメラから映像を取得し、フレームごとに正規化画像を生成し、特徴量を抽出する。ATM 状態把握では、入力フレームに対応する ATM 状態を取得しておく。その後、ATM 状態ごとに蓄積しておいた特徴量に対して、機械学習を用いて ATM 状態毎の異常検知辞書を生成する。

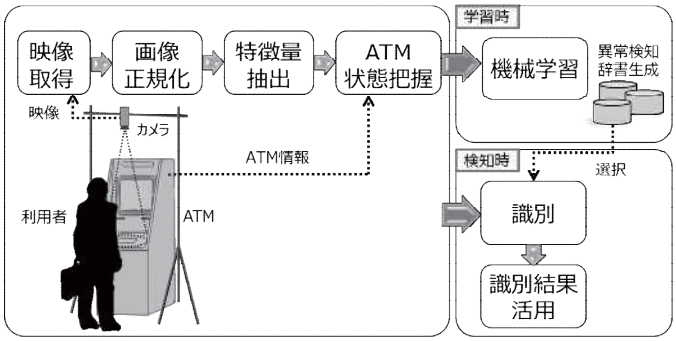


図 3:提案手法の流れ

3.3 検知フェーズ

検知フェーズでは、入力フレーム毎に正常／異常を判定する。

映像取得から ATM 状態把握までは、学習フェーズと共通の処理となる。入力フレームと同じ ATM 状態時の特徴量から生成した異常検知辞書を選択して識別処理を行い、正常／異常を判定する。識別結果活用では、結果が異常だった場合にのみアラーム発報、ATM 画面へのメッセージ表示、監視センターや ATM 管理者への通知・通報などを行う。

4. 評価

提案手法の有効性を確認するための実験を行った。本章では、従来手法(動画像特徴抽出と識別処理)と提案手法(動画像特徴抽出と ATM 状態把握および識別処理)でそれぞれ異常行動検知を行い、結果を比較する。

4.1 処理概要

従来手法と提案手法の間で“ATM 状態把握”に関わる処理以外は共通である。映像をフレームレート 10fps で取得し、画像正規化で 640×480 pixel にリサイズする。動画像特徴には細かい動きに対応できる ST-Patch を使用し、識別処理には一般的に用いられる部分空間法を用いた。入力フレームの特徴量と正常行動の部分空間との距離を異常度として算出し、異常度が閾値を超えた場合に「異常」と判定する。

4.2 評価データ

社内で模擬環境を構築し、実際に ATM に対する動作の映像を取得した。被験者として社員 149 名を集めて撮影し、正常行動データと異常行動データを生成した。正常行動データは、ATM の通常の取引を撮影したものとなる。通常の取引の中には、左手の操作、手を置く動作、紙幣を数える動作など特徴のある動きを意図的に含んでいる。異常行動データは、一般に知られている犯罪事例を元にしてシナリオを作り、犯罪行動を真似た動きを演じることで生成した。学習データと検知データにはそれぞれ以下を用いた。

- 学習データ:
  正常行動　約 800 動作(約 60 万フレーム)
- 検知データ:
  未学習の正常行動　約 240 動作
  異常行動　32 動作

ここでの異常行動は、カード挿入口に犯罪装置を取り付ける行為である。犯罪装置は厚紙などで模式的に製作したもので、カード挿入口を覆う様に取り付けられる。

4.3 評価方法と結果

本稿では、動作単位で評価を行う。各動作の中で一度でも異常と判定されたフレームがあればその動作を「異常」とし、そうでない場合は「正常」とする。評価指標には、Equal Error Rate（EER）を用いる。EER は、正常行動の誤検知率と異常行動の未検知率が等しくなるときのエラー率である。

従来手法は EER が 8.3% となった。正常行動に含まれているカードの入れ直しをする動作が高い異常値となったことが原因として挙げられる。一方、“ATM 状態把握”を導入した提案手法では EER が 0.0% となり、正常行動と異常行動を全て正しく識別できた。従来手法は、多様な正常行動を一つの部分空間にするのに対し、本手法は ATM 状態毎に起こり得る正常行動を学習するため、従来手法で見られた誤検知・未検知が改善されたと考えられる。

表 1：評価結果

	従来手法	提案手法
EER	**8.3 %**	**0.0 %**

5. おわりに

我々は、犯罪装置取付けなどの犯罪行動の検知が可能な異常行動検知技術を実現した。動画像特徴量抽出と ATM 状態把握および識別処理を組み合わせた手法を提案し、動画像処理のみでは識別できなかった正常行動と異常行動に対して結果が改善されることを確認した。本技術は ATM 操作に不慣れな人の戸惑い検知への応用も期待できる。

本評価データは疑似環境データでの評価のため、実環境データでの検証が必要である。

参考文献

[南里 2005] 南里卓也, 大津展之: 複数人動画像からの異常動作検出, 情報処理学会論文誌, vol.45, no.SIG_15, pp.43-50, 2005.

[弓場 2012] 弓場竜, 村井泰裕, 藤吉弘亘: 時空間特徴の統計的学習を用いた異常行動検知, 電子情報通信学会論文誌, vol.J95-D,no.6,pp.1369-1379,2012.

[林 2006] 林健太郎, 羽下哲司, 関真規人, 笹川耕一: 映像監視における人物位置行動の検出技術, 情報処理学会論文誌, vol.47, no.SIG_9, pp.12-19, 2006.

[Kobayashi 2012] Kobayashi, T., Otsu, N., : Motion recognition using local auto-correlation of space–time gradient, Pattern Recognition Letters, vol.33,no.9,pp.1188-1195,2012.

[弓場 2012] 弓場竜, 縣禎輝, 藤吉弘亘: 回帰推定による欠損距離画像の動作特徴量補正法, 電子情報通信学会技術研究報告, 信学技報, 2012.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

船舶のタンク・ホールド内画像認識に関する研究

Study on Image Recognition in Tank / Hold of Ship

平方 勝　　馬 沖　　谷口 智之
Masaru Hirakata　　Ma Chong　　Tomoyuki Taniguchi

海上・港湾・航空技術研究所
National Institute of Maritime, Port and Aviation Technology

This document describes whether image recognition using Deep Learning can be practical for images taken by unmanned aerial vehicle (drone). There is a movement to try to utilize a drone for inspection inside tank / hold of the ship. Accessibility to the inspection site becomes easy, while on the other hand, a technique to support condition evaluation on the screen is also required to judge through the video taken with the drone. In this research, we examined to verify the recognition rate of structural members in tank / hold in a bright environment. In addition to recognizing in a bright environment, we also examined to verify the recognition rate in a virtual dark environment.

1. はじめに

船舶のタンク/ホールド内の点検に，ドローンを活用しようとする動きがある．点検箇所へのアクセスが容易になる一方，ドローンで撮影した映像を通して現場判断するため，画面上で状態評価を支援する技術が望まれる．

ばら積み貨物船のホールド内は，ハッチを開けると明るく，また，ドローンの飛行にとって障害となる構造部材がない開放的な空間になっている．一方，タンカーのタンク内は，甲板で覆われるため暗い上にドローンの飛行に有用な GPS が効かない．そして，ウェブフレームなど大きな構造部材（ドローンの飛行にとって障害物）がタンク内に多く配置されている．

本研究では，深層学習(Faster R-CNN)を用いた物体検出技術を，第一段階として，ばら積み貨物船のホールド内の構造部材の認識に適用させ，その認識率を検証する．実験は，明るい環境での認識率の検証を行った後，暗い環境を模擬した状態での認識率の検証を行った．

2. 船舶の点検

大型の船舶でアクセスが困難な高所へは，図 1 に示すように，高所作業車を使用して点検を行うことが多い．図 2 に示すように，高所点検箇所にドローンを飛ばして点検する動きがある．



図 1　ばら積み貨物船の点検風景

連絡先：平方勝，海上・港湾・航空技術研究所，〒181-0004 東京都三鷹市新川 6-38-1，0422-41-3690，0422-41-3085，hirakata@nmri.go.jp

図 2 に示すように，船舶のハッチが開いた状態では明るい環境であるが，タンカーのように甲板で覆われているような船舶のタンク内は，図 3 に示すように，暗い環境である．また，タンク内は多くの構造部材が配置されているため，直接の視界から外れることも予想される．

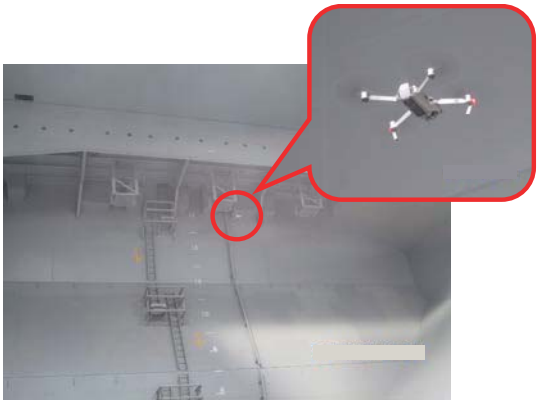


図 2　ドローンの貨物倉内飛行風景



図 3　貨物倉内（ハッチ閉）

3. 画像認識

3.1 Faster R-CNN

深層学習を利用した物体検出に，R-CNN が利用されることが多い[Ross Girshick 2014] ．R-CNN は Selective Search[J.R.

R. Uijlings 2013] で検出した物体候補領域を AlexNet（または VGGNet）に入力して検出する手法である．Selective Search は，色の類似度に着目したセグメンテーション法であり，色の類似度の許容度を超えてグルーピングすることで様々な大きさの候補領域を抽出することができる．Selective Search で検出した物体候補領域を一定サイズにリサイズし，後段の CNN に入力する．物体候補領域の数だけ CNN で認識するため，計算コストが高いといわれる [情報機構 2017] .

畳み込み処理の回数を大幅に削減することで計算コストを改善した Fast R-CNN が報告された．その後，Fast R-CNN をベースとし，Selective Search をやめ，Region Proposal Network で行うことで計算コストを削減した Faster R-CNN が報告された．現在では，Faster R-CNN が物体検出で広く使用される [情報機構 2017]．本研究では，Faster R-CNN による検出を行った．

3.2 データセット

ばら積み貨物船の重点点検箇所周辺の構造部材名を学習・認識させる．ばら積み貨物船の重点点検箇所は，一般的に図 4 に示す 5 箇所といわれている．1 箇所につき，100 画像程度を用意した．画像データを大量に用意するために，画像の中には，同じ対象箇所を別のアングルから撮影したものもある．図 5 で赤枠を記した部材名を学習させた．

学習には，訓練データを 550 画像，検証データを 250 画像用意した．テストデータに 89 画像を用意した．今回用意した画像は，すべて点検者がデジタルカメラで撮影した画像であり，ドローンのカメラによる撮影画像ではない．



図 4 データセット

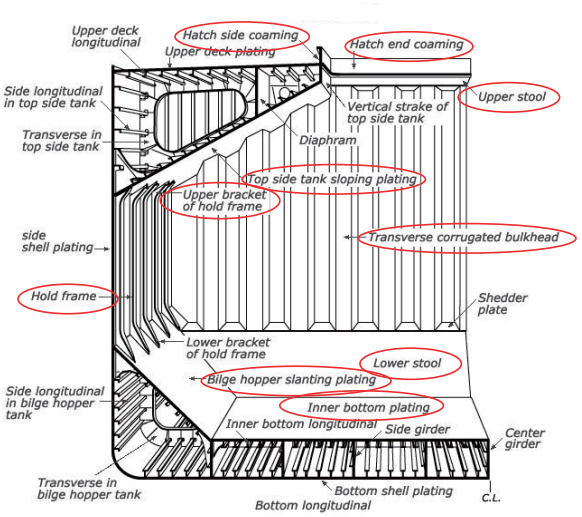


図 5 学習させた部材名

3.3 明るい環境での認識実験

図 6 は，ばら積み貨物船の貨物倉内の画像認識結果及び撮影箇所を示した図面である．表 1 は，認識できた部材名とその認識率である．図 6 で撮影されたすべての部材名が 100%または 100%に近い認識率となっている．訓練データに同様のアングルからの撮影画像が多かったためと考える．

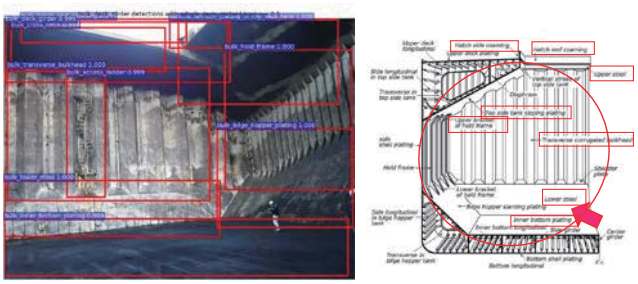


図 6 部材認識(1)

表 1 部材認識(1)の認識率

部材名	認識率
Top side tank sloping plating	100%
Hold frame	100%
Bilge hopper plating	100%
Transverse bulkhead	100%
Lower stool	100%
Access ladder	99.9%
Inner bottom plating	99.9%
Deck girder	99.9%
Hatch side coaming	99.9%
Hatch end coaming	99.4%
Cross deck	99.4%

図 7 も，ばら積み貨物船の貨物倉内の画像認識結果及び撮影箇所を示した図面である．表 2 は，認識できた部材名とその認識率である．表 1 と同様，すべての部材名が 100%または 100%に近い認識率となっている．Inner Bottom Plating は，画像の中で僅かに写っているだけだが，高い認識率で認識できて

いる. Lower Stool の床(エッジ)を認識し, それが, Inner Bottom Plating であると, 構造位置関係も認識できている.

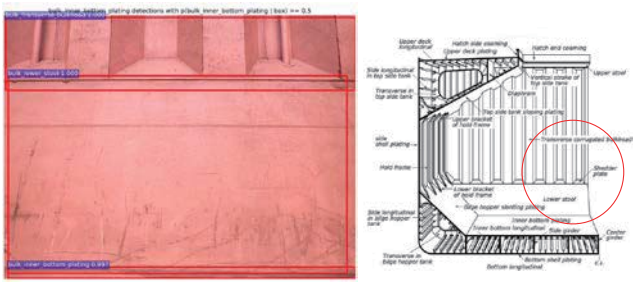


図 7　部材認識(2)

表 2　部材認識(2)の認識率

部材名	認識率
Transverse bulkhead	100%
Lower stool	100%
Inner bottom plating	99.7%

図 6 及び図 7 の認識結果は, 学習で使用した画像ではないが, 同じ船舶の他の画像データを学習させていたことが, 高い認識率につながったものと考え, 別の船舶の同様の箇所について認識させた. 図 8 は, ばら積み貨物船の貨物倉内の認識結果である. 学習で使用した船舶は全て, 船側部に Hold Frame が配置された構造様式であったが, 図 8 の船舶は Hold Frame が配置されていないタイプの船側構造となっている. そのため, Transverse bulkhead が船側を含めて認識された. 船舶の構造様式について多様な学習データを用意して改善していく必要がある. また, Deck girder, Cross deck が画像(図 8)に写っているが, 認識率は 50%を下回っていた. 訓練データを増やして学習を重ねる必要がある. Access Ladder は画像に写っておらず, 認識しなかったことは正しい結果である.



図 8　部材認識(3)

表 3　部材認識(3)の認識率

部材名	認識率
Top side tank sloping plating	98.3%
Transverse bulkhead	95.3% (船側を誤認)
Access ladder	認識せず(正解)
Inner bottom plating	97%
Deck girder	50%未満
Cross deck	50%未満

3.4　暗い環境を模擬した認識実験

ばら積み貨物船のハッチを閉じると, 図 3 に示すようにライトがないと周りを知覚できないほど暗い. ドローンにライトをつけて撮影した画像はない. ライトをつけて飛行したドローンの撮影画像をどの程度認識できるか, ここでは, 図 9 に示すように, オリジナルの画像(明るい環境)から画像処理によって, 暗い環境でライトがあたった模擬的な状態を作り出して, 認識率の検証を行う. Case1, Case2 にしたがって, 明るい範囲を中心位置に集中させ, 全体的に暗い環境を作り出した.



図 9　仮想的な暗い環境創出

図 10 は, 図 6 の画像をオリジナルにして, Case1, Case2 の暗い環境を模擬的に作り出した状態の認識結果である. その状態の認識結果を表 4 に示す. なお, 明るい画像で学習させた状態の重みを使用している. ヒトの知覚のように, Case1, Case2 の順に認識率が低下している. 画面中央の明るい部材の認識率は依然高いものの, 周辺の暗い部材の認識率は低下している. Inner Bottom, Cross Deck の認識率は 50%と低いものの認識できていることは画像判断するヒトにとって有用である. Case2 の認識結果で, Transverse Bulkhead が高い認識率となっていることは, この画像からでは Transverse Bulkhead を見落とす可能性があるため, 判断するヒトに有用な判断材料を提供できると考える. Case2 の静止画像を見ただけでは全体判断が困難である. ヒトは, 静止画像だけでは対象を認知てきなくても, 連続的な動きの中では認知できることもあるため, システムも同様にリアルタイムの連続的な動きから認識できるようにする必要がある.

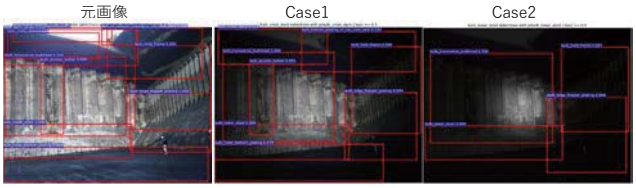


図 10　部材認識の変化(1)

表 4　部材認識の変化(1)での認識率変化

部材名	元画像	Case1	Case2
Top side tank sloping plating	100%	53%	50%未満
Hold frame	100%	99.8%	98.7%
Bilge hopper plating	100%	99.4%	86.8%
Transverse bulkhead	100%	100%	99.8%
Lower stool	100%	98.9%	90%
Access ladder	99.9%	99.3%	50%未満
Inner bottom plating	99.9%	57.9%	50%未満
Deck girder	99.9%	50%未満	50%未満
Hatch side coaming	99.9%	50%未満	50%未満
Hatch end coaming	99.4%	50%未満	50%未満
Cross deck	99.4%	52.4%	50%未満

図 11 は, 図 7 の画像をオリジナルにして, Case1, Case2 の暗い環境を模擬的に作り出した状態の認識結果である. その状態の認識結果を表 5 に示す. Case1,Case2 とも画像に写っていな

い Top side tank sloping plating を誤認している. Case1 では Transverse bulkhead を認識できていることは, この画像からでは Transverse Bulkhead を見落とす可能性があるため, 判断するヒトに有用な判断材料を提供できると考える. Case2 はヒトがみても, この画像のみでは判断が困難である. リアルタイムの連続的な動きから認識できるようにする必要がある. また, 構造部材の位置関係等, 船体構造の知識をベースにヒトは判断することがあるため, このような仕組みを取り入れることによって, 認識率が向上すると思われる.

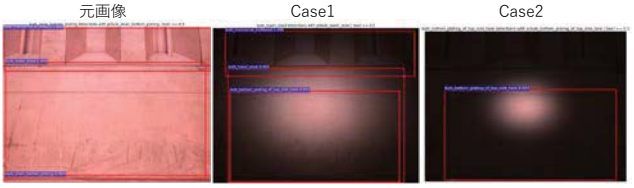


図 11 部材認識の変化(2)

表 5 部材認識の変化(2)での認識率変化

部材名	元画像	Case1	Case2
Top side tank sloping plating	認識せず（正解）	95.3%（誤認）	86.2%（誤認）
Transverse bulkhead	100%	99.6%	50%未満
Lower stool	100%	90.2%	50%未満
Inner bottom plating	99.7%	50%未満	50%未満

4. おわりに

船舶のタンク/ホールド内の点検にドローンを活用する動きがあり, ドローンが撮影した映像から自動評価できる実用的なシステムを構築するための基礎実験を行った. 本研究では, Faster R-CNN を用いて船体構造部材の認識を行った.

データセットは, 明るい環境での画像を用い, 訓練データが 550 画像, 検証データが 250 画像で 15 時間学習した結果を基に, テストデータ 89 画像の検証を行った.

学習した明るい環境での画像では, 認識率は高かった. ただし, 学習した船舶にない構造様式については誤認していたため, 学習データは画像の数に加えて多様な画像を用意する必要がある.

明るい環境で学習した重みを使用して, 暗い環境での認識率を検証した. 暗い環境での画像は, 画像処理によりライトがあたった模擬的な状態を作り出したものである. 明るい環境での画像認識率に比べると, 認識率は低下する. ライトのあたったと想定する明るい(画像中央)の認識率は高いものの, 周辺の暗い部材の認識率は低下し, 認識できなかったケースも見られた. ヒトの視覚認知に似ている. 実際にライトを当てると, ハレーションが起こり, ライトを当てた箇所周辺が真白くなる場合もある. このような問題処理も含めて, ライトを当てた実際の画像を整備し, そのもとで学習を行っていく必要がある.

暗い静止画像だけでは人も判断が困難であることがある. ヒトは, 静止画像だけでは対象を認知できなくても, 連続的な動きの中では認知できることもあるため, システムも同様にリアルタイムの連続的な動きから認識できるようにする必要がある.

今後は, 検査場面で最も重要な損傷を認識できるように学習データを整備して検証を行っていく予定である. 深層学習はブラックボックスであるといわれることが多い. 認識結果の解釈を数多くこなすことで, 認識率の向上につなげる予定である.

参考文献

[Ross Girshick 2014] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik: Rich Feature hierarchies for Accurate Object Detection and Semantic Segmentation, CVPR '14 Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp. 580-587, 2014.

[J.R. R. Uijlings 2013] K. E. A. Van De Sande, T. Gevers and A. W. M. Smeulders: Selective search for object recognition, International Journal of computer vision, vol. 104, no. 2, pp.154-171, 2013.

[情報機構 2017] 機械学習・人工知能業務活用の手引き, 情報機構, 2017

[Shaoqing Ren 2015] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, IEEE Trans. Pattern Anal. Mach. Intell. 39(6): 1137-1149, 2017

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

OpenPoseによるバスケットボール投入予測

Prediction of Basketball Free Throw Scoring by OpenPose

中井 眞人 *1 Masato NAKAI
角田 善彦 *1 Yosihiko TSUNODA
孫 財東 *1 Caidong SUN
村越 英樹 *1 Hideki MURAKOSHI
林 久志 *1 Hisashi HAYASHI
網代 剛 *1 Tsuyoshi AZIRO

*1産業技術大学院大学 産業技術研究科
School of Industrial Technology, Advanced Institute of Industrial Technology

We aim for application developments in several fields such as healthcare and robot control. When considering to achieve high quality performance in such applications, it is often necessary to have large amount of accurate and various human posture data. OpenPose of CMU can realize simultaneous posture recognition of multiple people in real time by simple Web camera. We developed two application systems utilizing OpenPose. The first one is a basketball free throw prediction model with human posture data recognized by OpenPose. The second one is a healthcare application which aims to detect brain stroke tendency. In this article, we focus on what we achieved through our first application development. We accomplished to build a high quality motion prediction model with OpenPose data. We consider it opens a way for other research and application developments which are related to human posture and motion.

1. はじめに

我々のチームでは，ヘルスケアやロボット制御分野などを念頭に置き，それら分野での応用開発に取り組んでいる．それらの性能向上を考える場合，精度の高い大量で多様な姿勢データが必要とされる場合が多い．昨年夏に CVPR2017 で発表された米国カーネギーメロン大 (CMU) の OpenPose [Cao 17] [CMU-OP GitHub17] は，簡単な Web カメラを使って図 1 の様にリアルタイムで多人数の姿勢が同時に認識できる．その動画は SNS 等で拡散されかなりの反響を呼んだ．OpenPose の応用研究としてリアルタイムの姿勢データを使ってバスケットボールの投入予測と脳梗塞初期診断判定 (FAST [Takano] [NCVC]) システムを構築した．その結果，投入予測モデルでは先行研究と比較してかなり優位であることが分った．理由としては 2 つあり，従来は全体の姿勢データを時系列で取得するにはモーションキャプチャーの様なセンサーを取り付ける必要がありデータ取得が煩雑であった．今ひとつはデータ取得が煩雑であるためサンプリングが極めて少なく回帰分析等の統計モデルを使った分析ができなかった．一方 OpenPose は Web カメラで動画を撮るだけで容易に多くの姿勢データが取得でき統計モデルを使った分析が可能になった．さらに統計モデルを使った投入予測では有意な精度が得られたので OpenPose が認識するリアルタイムの姿勢データは，統計モデルに耐え得る精度を持つことが判明した．

2. 目的

OpenPose は 2017 年 3 月にステレオ（複眼）カメラによる 3 次元捕捉 [CMU-OP GitHub17] に対応しているが，本研究では導入の容易さや実績を重視して単眼 Web カメラを用いた OpenPose 二次元位置認識環境を構築した（本稿 4.1 参照）．OpenPose の 1 つの特長は，リアルタイムで多人数の姿勢を同時に認識する点である．これを実現するため深層学習を使ったトップダウンの位置認識や PAF（Part Affinity Fields）[Cao 17] による部位の所属認識を採用した独特の工夫をしている．その結果 OpenPose の動画は見た目には精度と汎化性能は驚くべきものがある．しかし一般的に精度と汎化性能はトレードオフの関係があり両立は難しいとされている．そのため OpenPose の性能を確かめるべく OpenPose の姿勢認識データを使った予測モデルを構築して，その予測精度で OpenPose の性能を判断することにした．予測モデルは動画が撮り易く結果が明瞭なバスケットボールフリースローの動作から命中を予測するものとした．手続きとしてはフリースローの動画を OpenPose に通すと我々の環境では，人体姿勢部位検出の引数指定にもよるが 1 秒間に 5 ～ 20 フレーム程度の分解能で姿勢データが出力される．フレームデータを繋いで一旦時系列データとし，そこから特徴量を抽出する．命中予測は特徴量と命中結果とでロジステック回帰で行った．



図 1: OpenPose

3. 先行研究

スポーツ動作の解析は従来はビデオ画像を静止画に分割して姿勢位置を抽出するものが大多数であった [奥山 90] [松原 90]．一方姿勢データを直接取得する方法として体の部位にセンサーやモーションキャプチャーを取り付けて観測した．しかしデータ取得が煩雑なので回帰の様な統計モデルが使えず上

連絡先: 産業技術大学院大学 産業技術研究科 創造技術専攻
b1617mn@aiit.ac.jp b1613yt@aiit.ac.jp

級者と初心者の姿勢の差を論じるのが殆どであった [小郷 92] [安松谷 12]．一方バスケットボール試合ではシュート，ボール奪取等の回数データが試合毎に大量に記録されているので統計モデルを使った勝敗予測モデルが可能だった [斉藤 13]．スポーツ動作の統計モデルとしては隠れマルコフ [Stamp 12] を使った稲邑の姿勢抽象化モデル [稲邑 03] があるが，これは転移学習を目的としている．時系列の予測モデル [北川 93] は一般に定常波と非定常波に分かれ前者は ARIMA モデル，後者はカルマンフィルター [有本 77] が使われる事が多い．しかしこれらのモデルは逐次的に観測された説明変数と目的変数を使う．一方バスケットボールの投入予測は一連の姿勢データから最後に目的変数が得られるので特殊な時系列モデルとなる．この様なモデルとしては LSTM による一連の文章から肯定的か否定的かを予測する感情分析 [Yogatama 17] [Socher 12] や株価の動向から短期的に上昇下降を予測する SCW [Wang 14] がある．しかし何れも大量のデータを必要とする．バスケットの投入予測モデルは特殊な時系列モデルであり，少量のサンプリングであったので，時系列予測ではなく投入前後の姿勢とその差分から特徴量を抽出したロジステック回帰 [Bishop 06] を採用した．

4. OpenPose とは

安価な Web カメラで撮った動画をリアルタイムで多数人数を同時に姿勢認識するモデルで，昨年夏の CVPR2017 でカーネギーメロン大学 (CMU) が開発し学術公開したものである．図 2(左端) の様に 17 の関節点が繋がった 2 次元の姿勢を認識し，その結果をファイルに出力する．図 2(中央，右端) の様に OpenPose は指先と顔も同様に認識する．

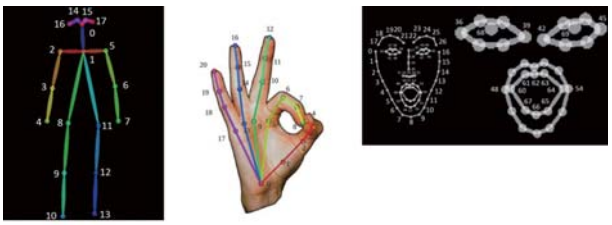


図 2: Analysis Point

4.1 OpenPose の動作方法

本研究では，次のような構成の PC サーバ上で OpenPose を稼働させた．インストールは CMU OpenPose オンラインリポジトリ [CMU-OP GitHub17] の “Installation, Reinstallation and Uninstallation” から，特に以下のハードウェアやソフトウェアに関して推奨の仕様（CPU コア数やメモリ容量）やバージョンを満たしていることを確認して行う．

【ハードウェア】

- CPU: AMD Ryzen 7 1800X (コア数 8)
- メインメモリ 16GB
- GPU: NVidia GeForce GTX 1080 ti (GPU メモリ: 11GB)

【ソフトウェア】

- OS: Ubuntu 14.04 LTS
- CUDA バージョン: 8.0
- cuDNN バージョン: 5.1 for CUDA8.0

4.2 OpenPose 出力データ

本稿 2 章に記した通り，本研究での OpenPose は 2 次元の位置認識なので姿勢，指先，顔の位置データは何れも x 値，y 値，信頼度の 3 値で表現される．OpenPose は隠れた位置の部位まで推測するので 0〜1 の信頼度がある．また多人数の姿勢を同時に認識できるので，1 フレームのファイルには左側の人物から人数分出力される．OpenPose は図 2 左端の人体姿勢 17 点の捕捉を基本とし，加えて図 2 中央および右端のように指先や顔のデータ出力は引数で指定できる．姿勢データは 1 フレーム毎に基本 17 点，指先，顔の同期をとって別ファイルに出力される．出力データ分解能は，計算環境の性能や出力部位の引数指定により変わるが，我々の環境では 1 秒間に 5〜20 フレーム程度であった．

5. バスケット投入予測の手法

5.1 OpenPose の姿勢データ

本研究では OpenPose の姿勢データのみ使用した（図 2 左端 17 関節点データ）．51 人のバスケットボールフリースロー動作の動画を撮り，OpenPose で解析した姿勢認識データをダウンロードした．

5.2 時系列への変換

ダウンロードしたフレームファイル群を読み込み 17 点の関節を時系列に変換した．

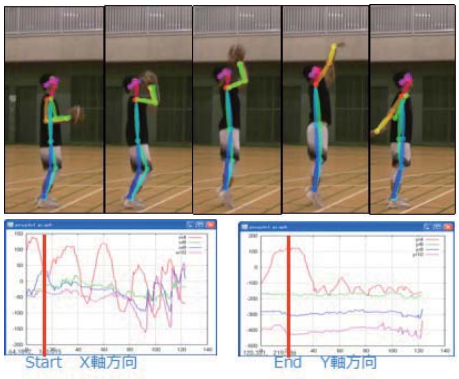


図 3: Time Series

51 人の全員は膝を曲げてから投入していたので，投入開始は膝を最大に屈折した時点とし，投入終了は手が最高点に達する時点とし，約 3 〜 5 秒間の時系列に揃えた．また各関節点は首関節からの相対距離とした．

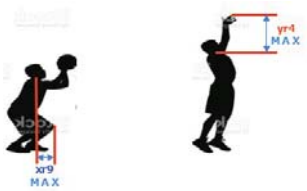


図 4: Start and Last Figure

5.3 ロジステック回帰モデルの採用

ボールが籠に入ったか入らなかったかの 2 値の教師データなので，ロジステック回帰モデルを採用した．2 値判定モデルでは定評のある Xgboost [Chen 16] も検討したが，分析結果の

出力は命中に寄与する特徴量が不明瞭なので採用しなかった．また非線形の 2 値判定モデルの SVM[Platt 99] があるが，これは判定確率を計算しないので採用しなかった．

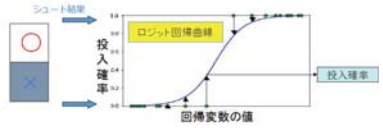


図 5: Logistic Regression

5.4 特徴量の作成

ロジステック回帰モデルに投入する特徴量を 17 点の関節点の投入開始時の値，投入終了時の値，その差分とその速度とした．22 の特徴量を投入したが，有意な特徴量として図 7 に示すものが残った．

6. 実験結果

6.1 精度

51 人を命中予測の高い順に並べて，命中した実績の累計の図を作成すると図 6 のパレート図が得られた．これで示された精度 AR は 41 %で有効な判別モデルが構築できていることが示された．(AR は完全モデルと無反応モデルで囲まれた凸部分の割合)

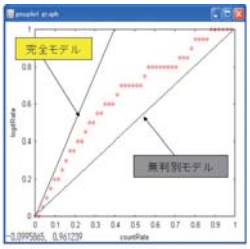


図 6: AR Accurate

6.2 有意な特徴量の検証

投入された特徴量で有意なものが正しくモデルを説明できているか検討した．図 7 は回帰の符号より，命中予測が高くなる特徴量の方向である．これによると開始時は腕を前から後ろへ一旦引いて反動を付けて投擲するオーバスローと投球時には膝の伸張と腕の振り上げによりボールに力を伝達している事が判明した．これにより特徴量の解釈でもこのモデルの有効性を確認した．

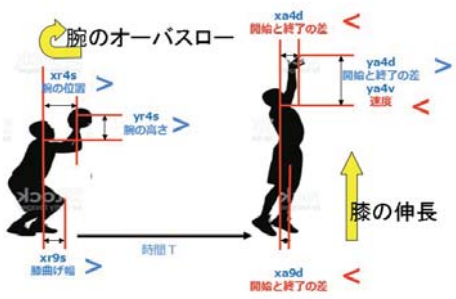


図 7: Feature Direction for Maximum Preciction

6.3 診断と矯正

投入予測モデルはレベル診断システムに使えることが判明した．OpenPose はリアルタイムで姿勢を認識するので図 8 に示す様に被験者の投入姿勢から直ちに投入予測を計算してレベル診断ができる．



図 8: Diagnosis System

また図 9 に示す様に有意な特徴量について初心者と上級者とを比較するとかなりの相違があることが判明した．この様に有意な特徴量を比較することで投入姿勢を矯正できることも分った．

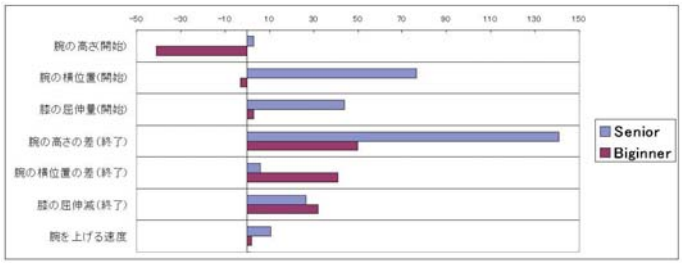


図 9: Diagnosis by Comparison

7. まとめと今後

一連の動作から投げられたボールの命中を予測する難しいモデルで期待された精度が得られたので，OpenPose のリアルタイムの姿勢認識精度は極めて高いと判断した．OpenPose は容易に姿勢データを得られることから，様々な場面で大量で高精度の姿勢データを取得することができる．図 10 の様に運転手や自転車の人も同時に観測でき，自動運転等の危険回避モデルに応用できると考えられる．



図 10: Prediction for Auto Driving

また競技での姿勢データが得られることで，図 11 の様に敵対的に学習するロボット [Fu 17] の初期学習としての逆強化学習 [Wulfmeier 14] に適用が考えられる．

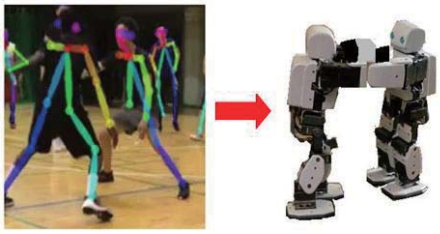


図 11: Inverse Reinforcement Learning

本稿研究で得られた OpenPose 姿勢データ解析の知見を活用し，FAST [Takano] [NCVC] と呼ばれる一般人が行う脳梗塞兆候判定を情報デバイスが検出・実行するシステムの試作も行った．この試作では図 12 の様に OpenPose の指先と顔の認識も使うことで正しく判定できることが分かった．この取り組みにより，OpenPose 技術のヘルスケア分野への応用可能性の一端を開いた．この詳細については別の機会に報告する予定である．

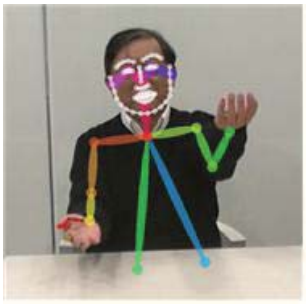


図 12: FAST

謝辞

東京都立産業技術高等専門学校のバスケットボール部にはフリースローの動画撮影に協力頂き感謝申し上げます．

参考文献

[Cao 17] Zhe Cao, Tomas Simon, Shih-En Wei , Yaser Sheikh:Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields, CVPR2017

[CMU-OP GitHub17] CMU-Perceptual-Computing-Lab openpose,https://github.com/CMU-Perceptual-Computing-Lab/openpose

[Takano] Junji Takano:t-PA - a 'Magic Bullet' For Treating Stroke,https://www.pyroenergen.com/articles13/ischemic-stroke-detection.htm

[NCVC] 国立研究開発法人国立循環器病研究センター: [103] 脳梗塞が起こったら - FAST を覚えよう, http://www.ncvc.go.jp/cvdinfo/pamphlet/brain/pamph103.html

[奥山 90] 奥山秀雄:コンピューター動作分析システム（APAS） を用いたバスケットボールのシュート分析 (1990), 日本体育学会第 41 回大会

[松原 90] 松原孝 猪木原孝二 川上雅之 [他]:バスケットボールゲームのシュートについて (1990), 岡山理科大学紀要

[小郷 92] 小郷克敏　錦井利臣　小澤雄二　中川保敬:運続動作時における主運動の視覚情報獲得時期 (1992), 日本体育学会大会号

[安松谷 12] 安松谷 亮宏 曽我真人 瀧寛和:バスケットボールにおけるシュートフォームの学習支援環境の構築 (2012), 情報処理学会シンポジウム論文集

[斉藤 13] 斉藤百合子 倉石平:バスケットボール競技におけるゲーム分析の指標に関する事例研究 (2013), 早稲田大学修士論文

[Stamp 12] Mark Stamp:A Revealing Introduction to Hidden Markov Models(2012)

[稲邑 03] 谷江博昭 稲邑哲也 中村仁彦:統計的情報量を用いた原始シンボル空間の構成とそれに基づく原始シンボル系列の記号化 (2003), 第 21 回ロボット学会学術講演会

[北川 93] 北川源四郎:時系列解析プログラミング—FORTRAN77 (1993), 岩波コンピュータサイエンス

[有本 77] 有本卓:カルマン・フィルター (1977), 産業図書システム・サイエンス・シリーズ

[Socher 12] Richard Socher, Brody Huval, Christopher D. Manning, Andrew Y. Ng:Semantic Compositionality through Recursive Matrix-Vector Spaces, EMNLP2012

[Yogatama 17] Dani Yogatama, Phil Blunsom, Chris Dyer, Edward Grefenstette, Wang Ling:Lerning to Compose Wards into Sentences with Reinforcement Learning, ICLR2017

[Wang 14] Jialei Wang, Peilin Zhao, Steven C.H.:Exact Soft Confidence-Weighted Learning, ICML2014

[Bishop 06] Christopher M. Bishop:Pattern Recognition and Machine Learning, Springer(2006)

[Chen 16] Tianqi Chen, Carlos Guestrin, Christopher:XGBoost: A Scalable Tree Boosting System, Arixv2016

[Platt 99] John C. Platt:Fast Training of Support Vector Machines using Sequential Minimal Optimization(1999)

[Fu 17] Justin Fu, Katie Luo, Sergey Levine:Lerning Robust Rewards with Adversarial Inverse Reinforcement Learning, Arxiv2017

[Wulfmeier 14] Markus Wulfmeier, Peter Ondru´ ska , Ingmar Posner :Maximum Entropy Deep Inverse Reinforcement Learning, NIPS2014

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

場の雰囲気を考慮したBGM推薦システム構築の試み

BGM Recommendation System Considering the Atmosphere

佐藤 季久恵 *1 Kikue Sato　坂井 栞 *1 Shiori Sakai　高屋 英知 *2 Eichi Takaya　山内 和樹 *3 Kazuki Yamauchi　大矢 隼士 *3 Hayato Oya　栗原 聡 *2 Satoshi Kurihara

*1電気通信大学 大学院情報理工学研究科
Graduate School of Informatics and Engineering, The University of Electro-Communications

*2慶應義塾大学 大学院理工学研究科
Graduate School of Science and Technology, Keio University

*3株式会社レコチョク
RecoChoku Co.,Ltd.

Japanese sometimes leave the decision of things to the atmosphere, it is very important to design the atmosphere of the place in various situations. BGM has an emotion inducing effect and an image inducing effect, which makes it possible to change the atmosphere of the space. In this paper, we proposed a BGM recommendation system reflecting the atmosphere in the shop by combining images and environmental sounds and constructing a data set using shop labels. As a result of recommendation by our system, higher ratings than baseline were obtained for three index.

1. はじめに

場の雰囲気は，人と人との間に流れる「空気」や視覚情報，聴覚情報などが複雑に絡み合うことにより形成される．これを人為的にデザインすることを考えたとき，特に視覚情報と聴覚情報をコントロールすることが重要である．家具や調度品，壁紙など，BGM には感情誘導効果やイメージ誘導効果があるとされており [1]，場の雰囲気を小さい労力で変容させることができる．現にレストランなどでは，それらの効果を利用するべく，楽曲配信サービスを導入しているところが多い．しかし，時間帯による店内の見え方や客層の変化により，店舗側はその都度雰囲気に合った BGM を提供することが求められる．時々刻々と変化する店内の状況に合わせて BGM 配信サービスが対応したり，人手による選曲を行うことは困難であることから，本研究では，店舗の雰囲気に適した BGM を推薦するシステムの構築を行う．本稿では，あらかじめ撮影された店舗内動画と，異なる環境音を組み合わせて印象評価を行い，類似度に基づいた楽曲推薦を試みる．

2. 関連研究

画像や楽曲といったコンテンツに対しユーザがタグ付けを行うことは，ソーシャルタギングと呼ばれ，これを利用した推薦や検索システムに関する研究は数多く存在する．梶ら [2] は，歌詞とアノテーションを利用し，視聴時のユーザの状況に合わせたプレイリストを作成するために楽曲とユーザを特徴量空間へマップする手法を採用している．特徴量には歌詞，楽曲情景，視聴状況を用いており，それらの特徴量空間にユーザをマップすることで，楽曲間，ユーザと楽曲間，ユーザ間の類似度計算を可能にしている．楽曲情景のラベルは登場人物（一人，私），いつ（朝，過去，春），状況（恋愛中，反社会），心理状況（悲しい，怒り）の 4 項目を用いている．歌詞と楽曲情景については，それまで視聴した好きな曲の特徴量平均をそれぞれの特徴量空間にそのユーザの嗜好としてマップすること

連絡先: 佐藤季久恵，電気通信大学大学院情報理工学研究科情報学専攻，東京都調布市調布ヶ丘 1-5-1，ksato@ni.is.uec.ac.jp

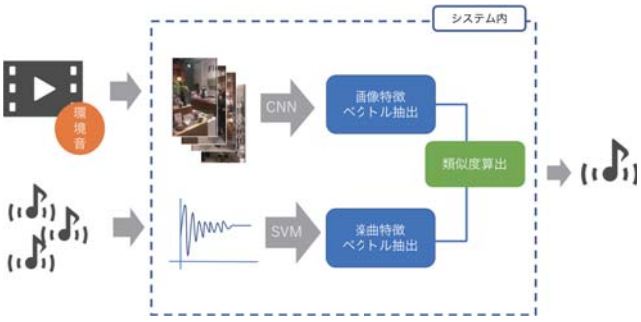


図 1: システムの概要図

で．推薦を行っている．Kaminskas ら [3] の研究ではユーザが関心のある場所（place of interest, POI）に即した楽曲推薦のシステムを構築している．楽曲と POI に双方同様の感情語を用いたタグをつけ，それらをベクトルとして扱い，類似度から適した楽曲を推薦している．タグには 9 項目の感情タグと 13 項目の物理的タグ（色や気温など）を用いている．

3. 本研究のアプローチ

本研究では，店舗での利用を想定し，場の雰囲気に合った BGM の推薦システムを提案する．

3.1 システムの概要

本研究における場の雰囲気を反映した BGM 推薦の流れを図 1 に示す．

まず，店舗内動画像を提案システムに入力し，Convolutional Neural Network（CNN）によってラベルを予測する．そして，推薦の候補となる楽曲群から特徴量としてメル周波数ケプストラム係数（MFCC）を求め，それを入力とした Support Vector Machine（SVM）を用いてラベルを予測する．それぞれの楽曲と店舗内動画において予測されたラベルの類似度を算出し，類似度が高いものを店舗に適した楽曲として推薦する．

3.2 環境音の利用

先行研究の多くは，場所の画像に対してラベル付けがなされている．画像データの取得やラベル付けは容易である一方，推薦システムにおいては視覚情報しか考慮されない点で不十分である．そこで本研究では，動画を印象評価に用い，その中でも環境音に着目する．

環境音は，我々が普段生活する中で無意識に耳にし，そこがどのような場所であるのかを判断するための材料にもなりうる．また，環境音が存在することにより，場所の見え方も変化する傾向があるとされている [4]．本研究では BGM を独立した音として扱うのではなく，環境音を含めて場をデザインすることを考慮し，環境音を含めた印象評価を行う．

3.3 店舗ラベルの導入

感情語や印象語を用いて画像や楽曲にラベル付けを行い，それに基づいた楽曲推薦を行うことは，先行研究でも試みられている．ラベルを用いることで，画像や楽曲がユーザーへ与える心理的影響を考慮することができ，個人の嗜好や状況に即した楽曲などを推薦することが可能である．しかし，個人の経験や感情に依存する感情語や印象語ラベルを使用することは不特定多数への適応を目的とした BGM 推薦には不向きである．

そこで本研究では，画像を見て想起される「情景」として，スターバックスや東急ハンズといった具体的な店舗名を用いてラベル付けを行う．チェーン店は全体を通してコンセプトを持っており，店舗内装などを統一しているところが多い．そのため，店を利用した人の間では同一のイメージを共有することが可能である．また，これまでは複数の感情語や印象語を用いることで店舗内を表現していたが，具体的な店名を用いることで一つのラベルで表現することが可能になる．

4. データセット

4.1 ラベルの選定

先述した店舗ラベルの候補を選定するにあたり，聞き取り調査を行った．

回答者は学生 4 人，社会人 2 名で，店舗イメージが確立している店舗名を列記してもらった．集計後，店の種類に偏りが出ないよう，USEN のコンシェルジュサービス [5] を参考に記載されている項目で補った．ラベル内容は表 1 に示す．これらのラベルを，後述する店舗内動画データセットの作成と楽曲評価に用いた．

4.2 店舗内動画データセットの作成

店舗内の雰囲気を評価するための店舗内動画を収集した．撮影は目線の高さで店舗内を 180 度撮影とし．時間は 10 秒から 15 秒ほどで行った．画像サイズは 1920 × 1080，フレーム数は 30fps とした．撮影イメージを図 2 に示す．ライトやアプリなどで明度や彩度の調整はしないものとした．

収集された店舗内動画から環境音を切り離し，異なる環境音を付け加えることで仮想店舗内動画を作成した．環境音はカフェ店内，子供が多いイベント会場内，ショッピングモール内，オフィス内，街中の 5 種類を使用した．

4.2.1 店舗内動画評価

作成した 740 本の仮想店舗内動画に対し，3 名の作業者に印象評価を行ってもらった．具体的には，作成した仮想店舗内動画を視聴してもらい，動画がどのような状況に当てはまるか，表 1 のラベルから選択してもらった．

表 1: ラベル一覧

情景	
スターバックス	ダイソー
ルノアール	東急ハンズ
コメダ珈琲	無印良品
バー	ニトリ
居酒屋	大塚家具
割烹・料亭	紀伊國屋書店
ラーメン屋	TSUTAYA
沖縄料理店	高島屋
イタリア料理店	PARCO
西友	イオンモール
成城石井	シェラトンホテル
カルディ	オフィス
アパレル（高級店）	企業ロビー（大企業）
アパレル（フォーマル）	企業ロビー（中小企業）
アパレル（カジュアル）	国際空港
ドラッグストア	地方空港
francfranc	
日差し	
あり	なし
時間帯	
朝（-10 時台）	昼（11 時台-14 時台）
夕方（15 時台-17 時台）	夜（18 時台-）
都市度合い	
都会	郊外
田舎	

4.3 楽曲評価

仮想店舗内動画と同様に，多数の楽曲に対して，それらがどのような状況下で BGM として流れているかを基準に，3 名の作業者に評価を行ってもらった．楽曲は「 J-POP」,「アニメ」,「キッズ・ファミリー」,「歌謡曲・演歌」,「邦楽ヒップホップ・R & B・レゲエ」,「邦楽ロック」,「洋楽ヒップホップ・R & B・レゲエ」,「洋楽ポップス」,「洋楽ロック」,「洋楽総合」の計 10 個のジャンルから 100 曲ずつを用意し，ラベルは表 1 から選択してもらった．

5. 予測器の構築と推薦方法

前節で作成したデータセットを用い，店舗内動画および楽曲に対する予測ラベルを出力するための予測器をそれぞれ構築した．

店舗内動画から予測ラベルを出力する予測器には，CNN を用いた．ネットワーク構造は AlexNet[6] と同様とし，ラベル付けのなされた仮想店舗内動画 740 本を学習データに用いた．なお，各動画は 1 本あたり 10 枚の連続画像とみなしているため，データセットは計 7400 枚のラベル付き画像となる．また，全 5 種類の環境音と画像との紐付けは，環境音タグにあたる値を CNN アーキテクチャにおける全結合層のユニットに追加することによって行っている（図 3）．

楽曲から予測ラベルを出力する予測器には，SVM を用いた．その際の入力には，各楽曲から算出した MFCC を用いている．

識別器から得られる結果は，各ラベルにおける値である．動画と楽曲のラベルの類似度を計算し，値が大きい 3 曲を推薦する．本研究ではユークリッド距離，コサイン類似度，ピアソ



図 2: 店舗内動画イメージ図

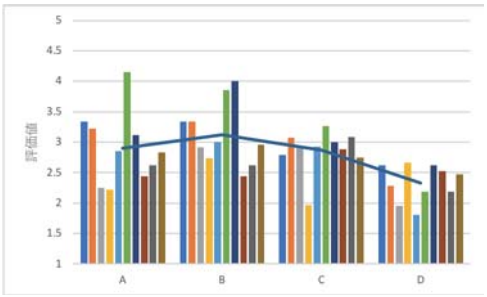


図 4: 楽曲推薦結果

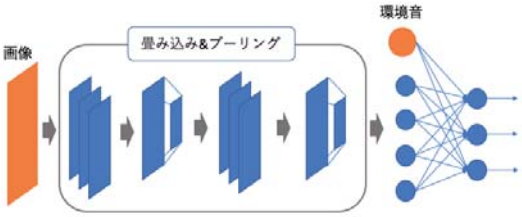


図 3: 店舗識別用 CNN のアーキテクチャ

ン相関係数を用いて求める．

6. 評価実験

提案システムの評価は，曲付きの動画を視聴した評価者に対するアンケートによって行った．評価者は本学学生から募集し，1 本の動画に対し 8 名または 9 名が 5 段階で評価を行った．

提案システムを用いて計 10 本の動画に対する楽曲をそれぞれ 3 曲推薦し，評価した結果を図 4 に示す．A，B，C はそれぞれコサイン類似度，ピアソン相関係数，ユークリッド距離にもとづいて推薦を行った結果であり，D はランダムに推薦を行った結果である．

類似度指標ごとの平均値は 2.99，3.12，2.89 の値で，標準偏差は 0.59，0.46，0.32 であった．また，図 4 からも分かるように A と B においては一部突出して評価が高いものが存在しており，D と比べると平均的にはいずれのグループも評価値が高いことが分かる．

しかし，評価者から感想を聞いたところ，「場所が歌詞と合っていない」「歌詞が気になってしまう」などの意見が得られた．

7. 考察

どの類似度計算においても，定量的にはランダム推薦に勝るものの，定性的には最適な BGM 推薦ができたと結論付けることはできなかった．原因の一つとしてデータの少なさとラベルの多さが学習に影響を与えたと考えられる．また十分なラベル付けがされていないという問題もあった．本研究ではデータセット作成時に動画や楽曲に対してラベル付けをしてもらったが，複数のラベルがつけられたものと，そうでないものの差が激しかった．楽曲に対するタグ付けは 1 曲に対し最低 2 人で行っていたことでラベル付与の偏りが生まれ，推薦に影響が出たのではないかと考えている．

加えて評価実験において同じ楽曲であるにもかかわらず，グループが異なるだけで点数が異なるものが多く存在した．前の曲との関係性が原因と考えられるが，確証を得るためにも人為的に様々な曲と組み合わせ，評価を行う必要がある．

8. おわりに

本稿では，画像と環境音を組み合わせ，店舗ラベルを用いたデータセットを構築することにより，店舗内での雰囲気を反映した BGM 推薦システムを提案した．3 種類の類似度指標を用いて動画に対し楽曲を推薦した結果，いずれの類似度指標においても，ランダムで推薦された曲よりも平均的に高い評価が得られた．一方で，定性的にはあまり高い評価は得られなかった．

今後の課題としては，データセットの拡充や，楽曲間の相性を考慮したシステムの改良などが挙げられる．

参考文献

[1] Daniel Västfjäll. Emotion induction through music: A review of the musical mood induction procedure. *Musicae Scientiae*, Vol. 5, No. 1_suppl, pp. 173–211, 2001.

[2] 梶克彦, 平田圭二, 長尾確. 状況と嗜好に関するアノテーションに基づくオンライン楽曲推薦システム. 情報処理学会研究報告音楽情報科学（MUS）, Vol. 2004, 127（2004-MUS-058）, pp. 33–38, 2004.

[3] Marius Kaminskas and Francesco Ricci. Location-adapted music recommendation using tags. In *International Conference on User Modeling, Adaptation, and Personalization*, pp. 183–194. Springer, 2011.

[4] R Murray Schafer. *The soundscape: Our sonic environment and the tuning of the world*. Simon and Schuster, 1993.

[5] music.usen.com コンシェルジュ, http://music.usen.com（閲覧日：2017 年 12 月 5 日）.

[6] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pp. 1097–1105, 2012.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

運動データからの楽曲作成

Music Generation from Motion Data

鈴木 隼太*1 渡邉 彰吾*2 六沢 一昭*1
Junta Suzuki Syogo Watanabe Kazuaki Rokusawa

*1千葉工業大学 情報科学部 情報工学科
Department of Computer Science, Chiba Institute of Technology

*2千葉工業大学大学院 情報科学研究科 情報科学専攻
Graduate School of Information and Computer Science, Chiba Institute of Technology

This paper describes an experimental sonification system which generates music from motion data measured by an acceleration sensor. The system generates rhythms, chords, tunes and harmony reflecting the observed movement, aiming for comfortable music.

1. はじめに

本稿は, 加速度センサーにより測定した運動データから楽曲を作成する可聴化実験システムについて述べる. 運動状況を反映し, 聴いて楽しめる楽曲を作成することが本システムの目標である.

運動情報の可聴化の研究には運動学習支援を目的としたものが多い. 動作情報を聴覚情報へ変換して伝達することにより, スポーツなどの動作学習やリハビリなどの動作支援に役立てることが考えられている. [高野 2016], [松原 2012], [宝里 2011] はその例であると言えよう. 本研究は, それらとは異なり, 楽しめる楽曲の作成が第一の目的である.

2. 自然な楽曲

自然な楽曲を作成するために着目したコードと規則を以下に示す.

表 1: ダイアトニックコード (ハ長調)

コード	C	Dm	Em	F	G	Am	$Bm^{\flat 5}$
機能	T	S	T	S	D	T	D

2.1 ダイアトニックコード (Diatonic Chord)

ダイアトニックコードのみを用いる. ハ長調のダイアトニックコードを表 1 に示す. 表中の T, D, S はコードの機能を示すもので, 以下の特徴がある.

T (Tonic; トニック) 安定感のあるコードである. 楽曲のはじめと終わりによく使われる.

D (Dominant; ドミナント) 不安定な響きを持つコードである. トニックへ動いて落ち着こうとする.

S (Subdominant; サブドミナント) やや不安定な響きを持つコードである.

連絡先: 六沢 一昭, 千葉工業大学 情報科学部 情報工学科, 275-0016 千葉県習志野市津田沼 2-17-1, rokusawa.kazuaki@it-chiba.ac.jp

2.2 コード進行

安定したコード進行には以下の 3 つのパターンがある.

- T → S → T
- T → S → D → T
- T → D → T

T, S からは T, S, D のどれへも進行するが, D からは T へのみ進行する.

$$
\begin{aligned}
C &\rightarrow Dm, Em, F, G, Am, Bm^{\flat 5} \\
Dm &\rightarrow G \\
Em &\rightarrow C, F, Am \\
F &\rightarrow C, Dm, G \\
G &\rightarrow C, Am \\
Am &\rightarrow Dm, F, G \\
Bm^{\flat 5} &\rightarrow C, Am
\end{aligned}
$$

図 1: コード進行 (ハ長調)

ハ長調における自然で安定したコード進行を図 1 に示す. このコード進行に基づいて楽曲のコード進行を決定する.

3. 運動データから楽曲へ

3.1 楽曲作成の流れ

運動データ

運動データは, 加速度センサーにより測定したデータである. 上下/左右/前後の 3 方向それぞれの加速度測定値からなる.

区間と小節

運動データを適当な数ごとに区切る. 例えば 50 個ごとに区切る. 区切りによって生まれた (運動データの) ひとつの区間を (楽曲の) ひとつの小節に対応させて楽曲を作成する.

楽曲作成

まず, 運動データを解析し各区間 (小節) のコードを決定する. そして, コードをもとに, 運動データも考慮して, メロディを生成する. 同様に運動データからリズムとハーモニーも生成する. 各小節のメロディを構成する音は, コードの構成音のみからとする. この「まずコードを決め, コードの構成音からメロディを決定する」方式は [杉原 2011][渡邉 2017] で用いたものである.

表 2: レベル値の例

測定値の絶対値	レベル値
0.0000 以上 0.1428 未満	0
0.1428 以上 0.2856 未満	1
0.2856 以上 0.4284 未満	2
0.4284 以上 0.5712 未満	3
0.5712 以上 0.7140 未満	4
0.7140 以上 0.8568 未満	5
0.8568 以上	6

3.2 レベル値と運動

各区間に対し, その区間内の測定値の絶対値の最大値のレベルを求め, それをその区間のレベルとする. 測定値の絶対値とレベルの対応例を表 2 *1 に示す. 測定値の絶対値が大きくなるほどレベルの値も大きくなる.

レベル値は運動の変化の大きさを表わす. 大きいレベル値は, 加速度の大きいことを示し, 運動が激しいことを意味する. 一方, 小さいレベル値は穏やかな運動を示し, レベル値ゼロは, 運動に変化が (ほぼ) ないことを意味する.

3.3 リズム/コード/メロディ/ハーモニーの決定

3.3.1 基本的な考え

運動が激しいほど楽曲のリズムも激しくするのが適当であろう. 上下/左右/前後方向のうち, 運動の激しさを直接反映するのは上下方向の測定値であると考え, 上下方向の測定値からリズムを決定することにした.

リズム以外の楽曲の要素であるコード及びメロディ, ハーモニーは, 主に上下方向以外 (左右及び前後方向) の測定値から決定する. これらは, リズムと異なり, 前の小節とのつながりが重要である. 例えばある小節のコードを決める時は, 前の小節のコードも考慮する必要がある. 対応する区間の測定値のみからコードを決定したのでは, 2.2 で述べた自然で安定したコード進行を実現することができない.

前述したように, レベル値は運動の変化の大きさを表わす. このことから, 基本的に, レベル値が大きいほど楽曲の変化を大きくし, 小さいほど変化は小さくする. そしてレベル値ゼロの場合は楽曲は変化させない.

3.3.2 リズムの決定

各区間の上下方向のレベル値から, 対応する小節のリズムを表 3 に基づき決定する. レベル値が大きいほど (運動が激しいほど), 激しいリズムが (音の数の多い小節が) 生成される.

3.3.3 コードの決定

ある小節のコードと, その小節に対応する区間の左右方向のレベル値から, 次の小節のコードを表 4 に基づき決定する.

例えば, ある小節のコードが F で, その小節に対応する区間の左右方向のレベル値が 4 の場合, 次の小節のコードは Dm となる. C と 3 ならば次のコードは F となる.

表 4 は図 1 のコード進行を実現しており, 自然で安定したコード進行が実現される.

3.3.4 メロディの決定

ある小節のコードとリズム (音符の数) と, 対応する区間の前後方向の測定値から, その小節のメロディ (を構成する音) を決定する.

*1 3.3.3 のコードの決定における, 左右方向の測定値とレベル値の対応を示す.

表 3: 上下方向のレベル値と対応するリズム

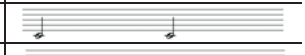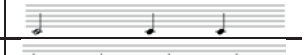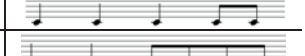
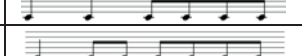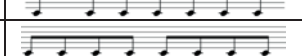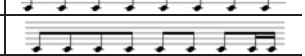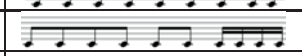
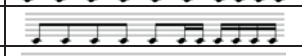
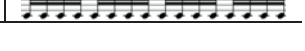
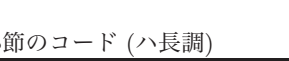
レベル値	音符数	リズム
0	1	
1	2	
2	3	
3	4	
4	5	
5	6	
6	7	
7	8	
8	9	
9	10	
10	11	
11	12	
12	13	
13	14	
14	15	
15	16	

表 4: 次の小節のコード (ハ長調)

レベル値	C	Dm	Em	F	G	Am	$Bm^{\flat 5}$
0	C	Dm	Em	F	G	Am	$Bm^{\flat 5}$
1	Dm	G	C	C	C	Dm	C
2	Em	G	C	C	C	Dm	C
3	F	G	Am	Dm	C	F	C
4	G	G	Am	Dm	Am	F	Am
5	Am	G	F	G	Am	G	Am
6	$Bm^{\flat 5}$	G	F	G	Am	G	Am

まず, リズムが示す音符数で前後方向の測定値を等分する. 等分によって生まれた各部分は, リズムによって決まった各音に対応する. 例えば, ある小節のリズムが表 3 のレベル値 6 のものならば, 音符数は 7 なので, 対応する区間の測定値を 7 等分する. 等分によって生まれた 7 つの各部分は 7 つの各音に対応する.

次に, 各部分の測定値の絶対値の最大値のレベル値を求め, そのレベル値と対応する小節のコードから表 5 と図 2 に基づき, メロディの構成音を決定する. 例えば, コードが Am である小節に対応する区間の, ある部分のレベル値が 5 の場合, その部分に対応する音は E5 となり, コードが G でレベル値が 1 ならば音は B3 となる.

3.3.5 ハーモニー (伴奏パターン) の決定

各区間の前後方向のレベル値から, 対応する小節の伴奏パターンを表 6 に基づき決定する. レベル値が大きいほど複雑な伴奏パターンが生成される.

4. 楽曲作成例

図 3 は, JR 京葉線新習志野駅から千葉工業大学新習志野キャンパスまで歩いた時に測定された運動データである. 図の横軸は歩行開始後の経過時間 (単位: 秒) を示す. 測定は毎秒 10 回

表 5: メロディ構成音

レベル値	C	Dm	Em	F	G	Am	$Bm^{\flat 5}$
0	C4	D4	E4	F3	G3	A3	B3
1	E4	F4	G4	A3	B3	C4	D4
2	G4	A4	B4	C4	D4	E4	F4
3	C5	D5	E5	F4	G4	A4	B4
4	E5	F5	G5	A4	B4	C5	D5
5	G5	A5	B5	C5	D5	E5	F5
6	C6	D6	E6	F5	G5	A5	B5
7	E6	F6	G6	A5	B5	C6	D6
8	G6	A6	B6	C6	D6	E6	F6
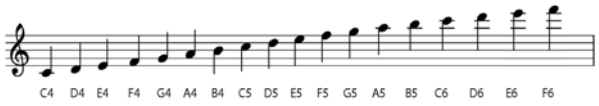


図 2: 構成音

であり, 運動データ 10 個で 1 秒経過となる. 運動時間 (測定時間) はおよそ 480 秒であり, 一方向あたりの運動データはおよそ 4800 個である. このデータからハ長調の楽曲を作成した.

運動データを 200 個ごとに区切って区間を生成した. 24 個の区間が生まれ, 24 個の小節からなる楽曲が作成された. 運動データ 200 個は 20 秒に対応するので, 運動 20 秒ごとに ひとつの小節が作られたわけである.

各小節の左右方向のレベル値とコード, 及び, 上下方向のレベル値と音符数を表 7 に示す. また, 7 小節目の前後方向の測定値とレベル値, 構成音を表 8 に示す. そして, 作成された楽曲を図 4 に示す.

最初の小節のコードは, ハ長調の楽曲のため, C である. 2 番目の小節は, 最初の小節のコード C と左右方向のレベル値 0 から C, 3 番目は, C と 2 から Em である (表 7)

表 7 より 7 小節目のコードは G で音符数は 10 である. 区間を 10 等分して生まれた各部分の, 前後方向の測定値の絶対値の最大値と, その値から定まるレベル値, そして, コードとレベル値及び表 5 から決まる構成音が表 8 に示されている.

図 3 の運動の内容であるが, 歩行だけでなく, 走行や停止といった動きも含んでいる.

走行部分は, 歩行開始後およそ 60 秒からの 30 ～ 40 秒間である. 運動データの対応する部分は 3 方向とも測定値 (の絶対値) が大きい. 作成された楽曲では, 走行部分は 4, 5 番目の小節が対応する. 二つの小節とも音符数が多く (11 個と 15 個), メロディの構成音が高い. リズムとメロディに走行が表われていると言えよう.

停止は 2 回ある. 初回の停止は, 停止時間が短く, 運動データにも楽曲にも明確には表われていない. 2 回目の停止は, 歩行開始後およそ 240 秒後からのおよそ 40 秒間である. 運動データの対応する部分は 3 方向とも測定値はゼロに近い. 作成された楽曲の対応する二つの小節 (13, 14 番目の小節) は, 二分音符二つの小節と, 全音符の小節である. 停止が明確に表われていると言えよう.

5. まとめ

加速度センサーにより測定した運動データから楽曲を作成する実験システムについて述べた. 目標の, 運動状況を反映し

表 6: 前後方向のレベル値と対応する伴奏パターン

レベル値	伴奏パターン
0	
1	
2	
3	
4	
5	
6	
7	
8	
9	

聴いて楽しめる楽曲を作成することはほぼ実現できたと思われる.

参考文献

[高野 2016] 高野 衛, 安藤 大地, 笠原 信一: 動作情報を用いた可聴化における変換手法の考察 – パフォーマンス実演による検証 –, 先端芸術音楽創作学会 会報, Vol.8, No.3, pp.1–4 (2016).

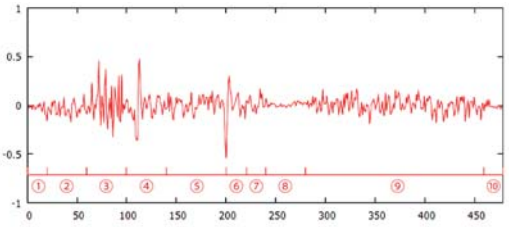
[松原 2012] 松原 正樹 他: 身体動作の連動性理解にむけた筋活動可聴化, 日本音響学会講演論文集, 2-10-2, pp.919–922 (2012).

[宝里 2011] 宝里 直幸, 有田 浩之, 阪口 豊: 身体運動の可聴化: その概念と運動学習支援への応用, 情報処理学会 インタラクション 2011, 3CR3-1 (2011).

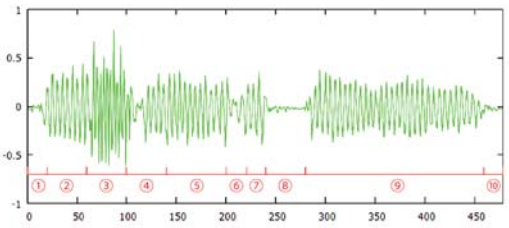
[杉原 2011] 杉原 大和, 六沢 一昭: プログラムからの楽曲生成の試み, 第 5 回エンターテイメントと認知科学シンポジウム (2011).

[渡邉 2017] 渡邉 彰吾, 六沢 一昭: プログラムからの楽曲作成, 情報処理学会 研究報告音楽情報科学, Vol.2017-MUS-115, No.25 (2017).

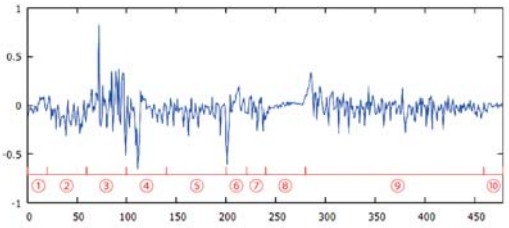
[島岡 1964] 島岡 譲 他: 和声 理論と実習, 音楽之友社 (1964).



(a) 上下方向



(b) 左右方向



(c) 前後方向

横軸の単位: 秒, 測定値 10 個/秒
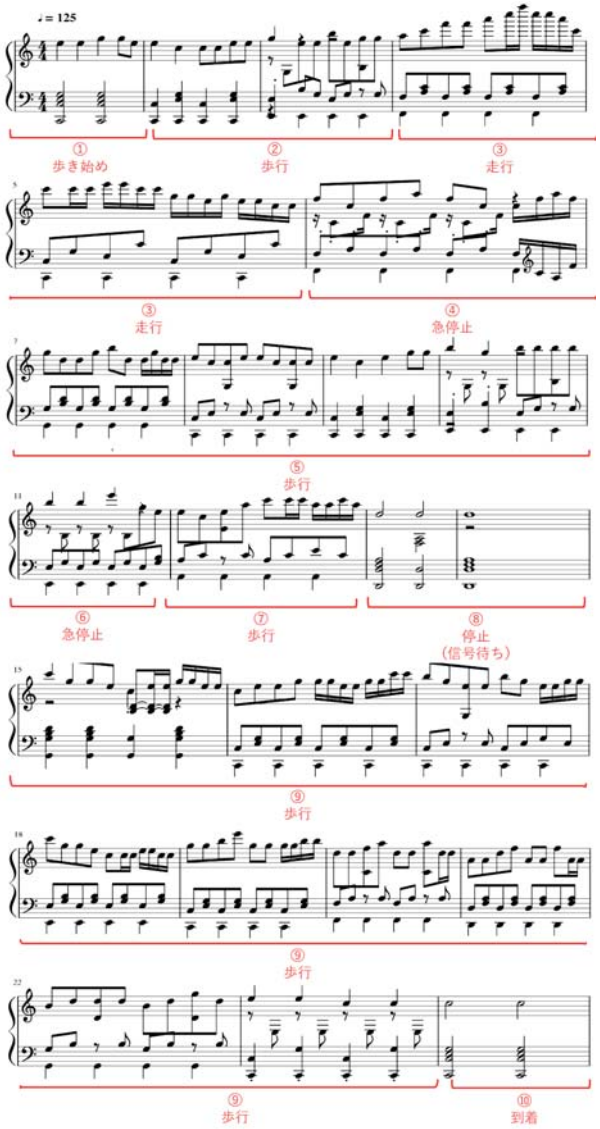
図 3: 運動データ

表 7: 各小節のレベル値とコード, 音符数

小節	左右	コード	上下	音符数
1	0	C	4	5
2	2	C	5	6
3	5	Em	6	7
4	2	F	10	11
5	3	C	14	15
6	5	F	9	10
7	2	G	9	10
8	0	C	7	8
9	2	C	4	5
10	0	Em	5	6
11	3	Em	4	5
12	2	Am	10	11
13	0	Dm	1	2
14	1	Dm	0	1
15	1	G	10	11
16	0	C	11	12
17	2	C	9	10
18	2	Em	10	11
19	3	C	9	10
20	3	F	8	9
21	2	Dm	8	9
22	1	G	7	8
23	0	C	3	4
24	0	C	1	2

表 8: 7 小節目の前後方向の測定値とレベル値, 構成音

区間	測定値	レベル値	構成音
1	0.391	3	G4
2	0.224	2	D4
3	0.235	2	D4
4	0.333	3	G4
5	0.505	4	B4
6	0.301	2	D4
7	0.254	2	D4
8	0.367	3	G4
9	0.224	2	D4
10	0.284	2	D4



運動データ 200 個/区間, 運動 20 秒/小節

図 4: 作成された楽曲

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

グラフカーネルフィルタをもちいた楽曲の画一な印象の獲得

Acquisition of uniform impression of music using graph kernel filter

錠 尚史 *1
Hisashi IKARI

*1株式会社ソケッツ
Sockets Co., Ltd.

The impression of a song varies from person to person. The sensitivity that depends on this individual is greatly different among individuals and it is difficult to quantify. In addition, because sensitivity changes with time, fixed point observation is also necessary, and it takes much time. In this case, focusing on the Graph laplacian which has the property of Fourier transform among the Deep learning attracting attention in recent years, we will clarify the time shortening by computer experiment and the frequency feature which affects the sensitivity of music.

1. はじめに

楽曲の印象は人により様々である．この個人に依存する感性は個人差があり，楽曲の印象の研究は心理学実験 [Hevner, 1936] を始めとする統計数理による検定に基づく研究が主立つ．この際，統計をもちいるために必要なサンプルの数は多い程，楽曲の印象の真の分布に近づく [Clopper and Pearson, 1934]．しかしながら，大規模の心理学実験には時間と人員を要し，また，感性は時間経過に伴い変化するために，時間変化を考慮した定点観測も必要であり，多くの時間を要する．

近年，深層学習の一般物体の認識の精度が注目を浴びている [Krizhevsky, Sutskever and Hinton, 2012]．深層学習は脳の計算モデルに基づいており，その深層学習の畳み込みは視覚の受容野の特性を示している [David Marr, 1982]．この畳み込みは SS 理論 [S. Suzuki, 1994] に端を発し，感性は受容野の反応と個人の経験に基づいた意味付けにより意識される感覚と前提される．

本件においても同様に深層学習の畳み込みを受容野とし，個人の経験に基づいた意味付けをニューラルネットワーク [Cochocki and Unbehauen, 1993] として分離，個人の経験に基づいた意味づけに至る前の楽曲の印象を画一な楽曲の印象と前提する．この構成に感性を表すラベルを付与し，深層学習の学習過程より受容野の特徴を求め，画一な楽曲の印象とし，かつ深層学習をもちいた計算機実験により実験の時間の短縮を行う．

2. 楽曲の感性の特徴モデル

2.1 楽曲の特徴量

楽曲は波形で表される．また，楽曲の印象は音色に起因すると前提する [Hevner, 1937]．ゆえに楽曲の特徴量をスペクトログラムとする．ただし，楽曲の特徴量を鮮鋭とするために，マルコフクラスタをもちいて端点を得るようスペクトログラムの要素の値を先鋭化する．この際，マルコフクラスタのパラメータ a を 3.0，e を 2.0 とおく．なお，楽曲の教師信号は 1,090 の楽曲からなり，ラベルには感性を表す 90 種の語彙が付与される．この語彙は賑やか，穏やか，爽やかを始めとする情感を表す語彙からなる *1．本稿ではこれを感性語彙と呼ぶ．

連絡先: 錠 尚史，株式会社ソケッツ，ikari@sockets.co.jp

*1 商用データのために非公開である，必要の際は連絡先まで連絡されたし

2.2 楽曲のグラフ表現

フーリエ級数は正規直交基底関数の正弦波関数の多項式として表せられる．ゆえにラプラス固有関数 $f(w)$ をもちいて，次数行列 D と隣接行列 A からなるラプラシアン行列 L の固有ベクトルより，フーリエ変換に等しい結果を得られる．

$$f(w) = \langle f, e^{jwt} \rangle = \int_R f(t)\, e^{-jwt}\, dt \tag{1}$$

$$e^{-jwt} = \frac{\partial^2}{\partial t^2}\, e^{-jwt} = w^2\, e^{-jwt} \tag{2}$$

$$w^2 = (2\pi\xi)^2 \tag{3}$$

楽曲の特徴量は，スペクトログラムの要素をグラフの頂点，隣接行列 A で任意の近傍の頂点同士を結合した構成をとる．次数行列 D の辺の重みはこれらスペクトログラムの要素の値の相関を重みとする．すなわち楽曲全てのスペクトログラムの要素の相関行列となる．なお，λ は固有値，μ_λ は固有ベクトルを表す．以上より，楽曲の特徴量は，スペクトログラムの要素について，任意の近傍の頂点間の値の振動として表される．

$$L = D - A = U\Lambda U^T \tag{4}$$

$$U = [\mu_{\lambda_0}, ..., \mu_{\lambda_{N-1}}] \tag{5}$$

$$\Lambda = diag(\mu_0, ..., \mu_{N-1}) \tag{6}$$

2.3 楽曲の畳み込み

楽曲の特徴量の畳み込みは，[Defferrard, Bresson and Vandergheynst, 2016] らの 128 要素の 2 乗からなるグラフカーネルフィルタ H を 20 次まで展開して畳み込む．グラフカーネルフィルタ H は楽曲の特徴量について逆フーリエ変換 U とフーリエ変換 U^T との積からなる多項式 $Udiag(H(\lambda_H))U^T f_{in}$ の H であり，次数は振動の大きさを表す．この際に求められる f_{out} が目的とする楽曲の感性の特徴である．

$$P_{\lambda_i} = \sum_{\lambda=\lambda_i} \mu_\lambda \mu_\lambda^T \tag{7}$$

$$P_{\lambda_i} P_{\lambda_j} = \delta(\lambda_i - \lambda_j) P_{\lambda_i} \tag{8}$$

$$f_{out} = H f_{in} \tag{9}$$

$$= \Big(\sum_{\lambda_i \in \sigma(L)} H(\lambda_i) P_{\lambda_i} \Big) f_{in} \tag{10}$$

$$= Udiag(H(\lambda_H))U^T f_{in} \tag{11}$$

2.4 楽曲の感性の個人差を表すニューラルネットワーク

フーリエ級数は級数が収束する前提の元，微分可能である．ゆえにグラフカーネルフィルタは，ラベルと楽曲の特徴量より誤差逆伝搬法をはじめとする学習過程で求められる．この際，グラフカーネルフィルタの荷重は，90 種の感性語彙が付与されたラベルと，グラフ表現された楽曲の特徴量を教師信号とし，グラフカーネルフィルタを入力層，1 層の中間層を持つニューラルネットワークを出力層とした構成より誤差逆伝搬法をもちいて求められる．上記の f_{out} と 1 層の中間層を持つニューラルネットワークにおける荷重とは異なり，この荷重は個人の感性の差を表す役割を持つ．

3. 楽曲の感性の特徴

3.1 楽曲の識別の精度

楽曲の感性の特徴モデルをもちいて，1,090 の楽曲，ならびに 90 種の感性語彙を付与したデータセットについて，深層学習の学習過程よりグラフカーネルフィルタを求め，さらに楽曲の特徴量とグラフカーネルフィルタより，全ての感性語彙を満たす楽曲の感性の特徴を求めた．その楽曲の特徴量の識別の精度は 96%に達した．この精度について学習回数の変動を加えた [表 1] を以下に記す．この際，汎化性能については 1,090 の楽曲の内の 10%をアトランダムに訓練データとした．

表 1: 楽曲の識別の精度

Epochs	訓練正答率	汎化正答率
1	0.0073	0.1431
40	0.8284	0.8110
80	0.9532	0.9394
100	0.9633	0.9514

3.2 楽曲の感性の特徴の俯瞰

得られた楽曲の感性の特徴について，全体の俯瞰を得るために，全ての楽曲の感性の特徴をグラフカーネルフィルタごとに重ね合わせた [図 1] を作成した．グラフカーネルフィルタはフーリエ級数展開の多項式に等しく，4 次項までで楽曲の感性の特徴がほぼ表されている．

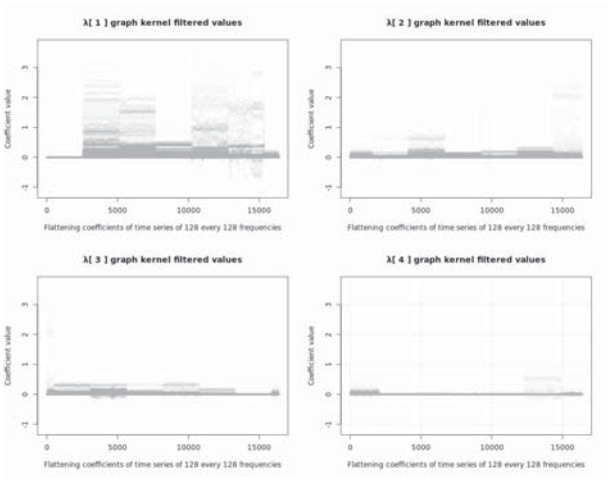


図 1: 楽曲の感性の特徴（全体）

[図 1] の横軸は楽曲の特徴量とグラフカーネルフィルタの積をとり，128 要素の 2 乗を 128 要素ごとに 1 次元に平坦化した軸である．これは左より高周波成分，右に進むにつれて低周波成分であることを示す．縦軸は楽曲の感性の特徴，すなわちグラフカーネルフィルタの係数を表す．この係数は値が高いほど振動が多い．楽曲の特徴量が与えられた際，受容野相当のグラフカーネルフィルタから得られる係数，すなわち楽曲の感性の特徴は，横軸の階数を 6 とした際に高周波帯，低周波帯，中周波帯のそれぞれ 2 種の反応よりなる．これは聴覚の基底膜の反応と同じ傾向を示している [Pickles, 1995].

3.3 特徴の次元圧縮による俯瞰

楽曲の感性の特徴は，グラフカーネルフィルタの係数であるため，グラフカーネルフィルタの 128 要素の 2 乗，それを 20 次持つ．すなわち 327,680 要素で表される．ゆえに楽曲の感性の特徴を固有値分解し，固有ベクトルより傾向を得る．得られた固有ベクトルは，おおよそ 3 種の方向に集約される [図 2].



図 2: 特徴の次元圧縮による俯瞰

3.4 感性語彙と楽曲の感性の特徴

感性語彙は，活性と非活性を表す軸，快と不快を表す軸の直交上で表される [Russell, 1980]．ゆえにそれらの軸に当該する語彙を活性について賑やか，非活性について穏やか，快について爽やかと前提する．不快は不快に相当する感性語彙がないために割愛した．ここで感性語彙に紐づく楽曲の例を挙げる，賑やかは真夏の Sounds good![AKB48] 他，穏やかは Too Much Rain[Paul McCartney] 他，爽やかは Love Is Stronger Than Pride[Sade] 他である．これらの楽曲の感性の特徴は，賑やかでは，1 次と 2 次に注目すれば高周波帯と中周波帯に高い値が分布し [図 3]，穏やかでは，1 次に注目すれば低周波帯に高い値が分布 [図 4]，爽やかでは，1 次に注目すれば高周波帯に高い値が分布する [図 5].

ここで感性語彙の直交軸から離れた感性語彙について，楽曲の感性の特徴を求める．非活性と快の中間に位置付けられる感性語彙は，愛しいと前提する．楽曲の例は If You Leave Me Now Maps[Chicago] となる．得られた楽曲の感性の特徴は，1 次と 2 次に注目すれば低周波帯と高周波帯に高い値が分布する [図 6]．感性語彙の軸の非活性，すなわち穏やかには低周波帯に高い値が分布し，快，すなわち爽やかには高周波帯に高い値が分布している．ゆえに愛しいは感性語彙の直交軸の和からなると考えられる．また，これらの感性語彙の直交軸の周波帯の傾向は，[3.3 特徴の次元圧縮による俯瞰] の 3 種の方向を示していると考えられる．

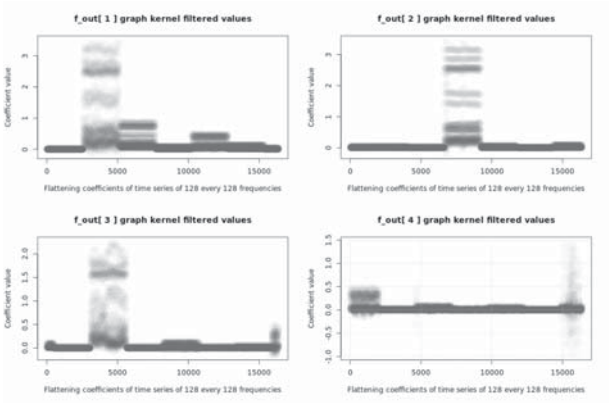


図 3: 楽曲の感性の特徴（賑やか）

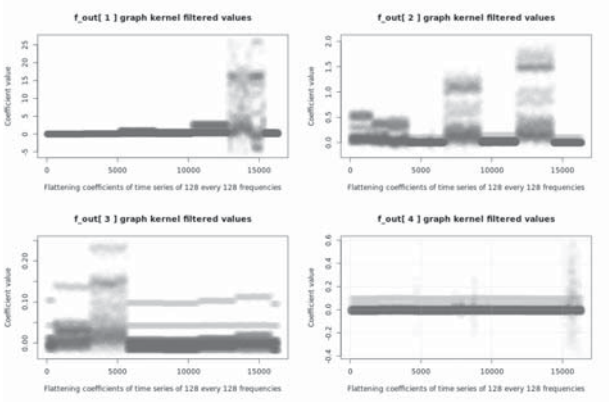


図 4: 楽曲の感性の特徴（穏やか）

3.5　楽曲の感性の特徴と感性語彙の類似性

楽曲の感性の特徴は感性語彙に紐づく．また，楽曲の感性の特徴は数列として表される．ゆえに楽曲の感性の特徴より幾何表現が可能である．以下に恋愛に関する感性語彙が付与された楽曲の集まりの例を記す [図 7]．恋愛を表す感性語彙を持つ楽曲の感性の特徴は，類似した数列となり，近い位置関係が求められている．また，同様に勇気づけに関する感性語彙が付与された楽曲の感性の特徴も類似した数列となり [図 8]，楽曲の感性の特徴は，感性語彙に基づいた類似性を得ていると言える．

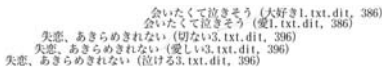


図 7: 恋愛を表す感性語彙を持つ楽曲の位置関係

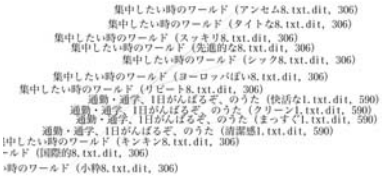


図 8: 勇気づけを表す感性語彙を持つ楽曲の位置関係

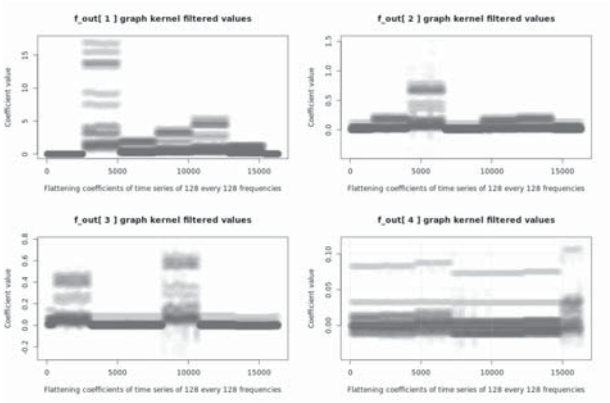


図 5: 楽曲の感性の特徴（爽やか）

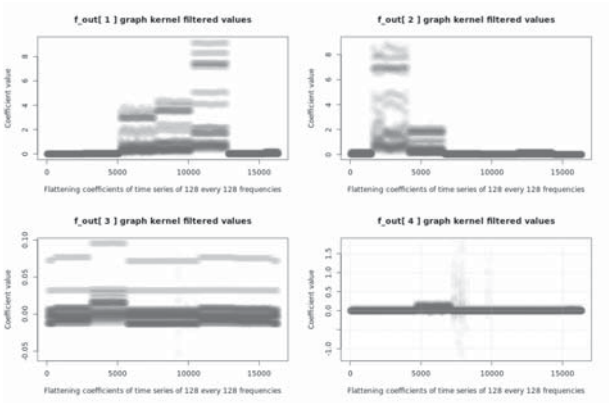


図 6: 楽曲の感性の特徴（愛しい）

3.6　楽曲の感性の特徴同士の類似性

楽曲の感性の特徴は，深層学習の学習過程をもちいて求められる．ゆえに楽曲の特徴量より感性語彙の識別が可能である．識別は深層学習に基づくため，識別の結果，得られる楽曲の特徴量の分類の値は，その分類の荷重を表す．本項では，この分類の荷重の値と楽曲の感性の特徴の類似性より，求められる類似楽曲の例を示す．ここでは感性語彙より特徴が顕著な賑やかの真夏の Sounds good![AKB48]，穏やかの Too Much Rain[Paul McCartney]，爽やかの Love Is Stronger Than Pride[Sade]，愛しいの Lady Marmalade[Christina Aguilera] を例にもちいる．それぞれ，賑やかでは summer tail[KEYTALK]，穏やかでは Heaven On The Ground[Jos James]，爽やかでは The Lady Wants To Know[Michael Franks]，愛しいでは Want U Back[Cher Lloyd] が求められた．これら類似楽曲の特徴を，賑やかの類似を [図 9] に，穏やかの類似を [図 10] に，爽やかの類似を [図 11] に記す．

4.　まとめ

楽曲の特徴量より，グラフカーネルフィルタと深層学習の学習過程をもちいて，楽曲の感性の特徴を求めた．この楽曲の感性の特徴に基づく識別の精度は 96%に達した．得られた楽曲の感性の特徴を固有値分解した際，その特徴は，おおよそ 3 種の方向に集約される固有ベクトルからなり，おおよその楽曲の感性の特徴の傾向が示された．また，楽曲の感性の特徴は，

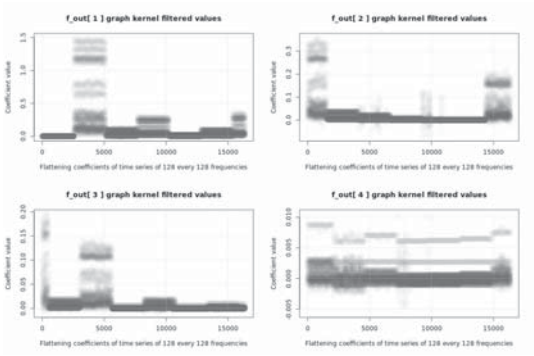


図 9: 楽曲の感性の特徴（賑やかの類似）

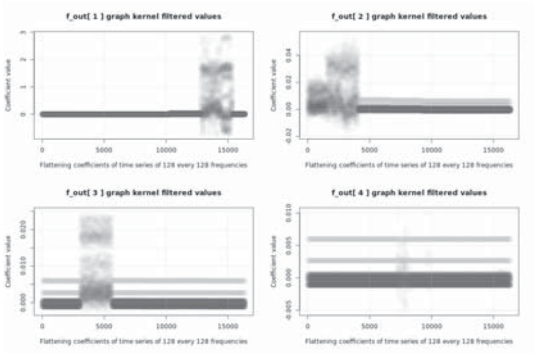


図 10: 楽曲の感性の特徴（穏やかの類似）

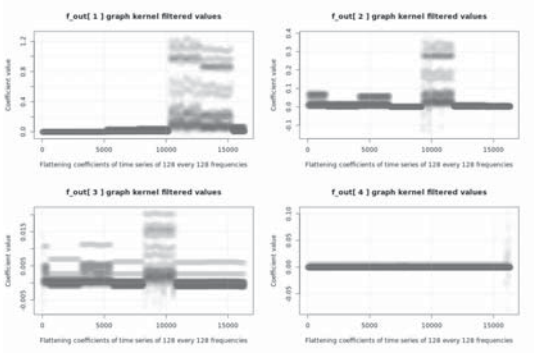


図 11: 楽曲の感性の特徴（爽やかの類似）

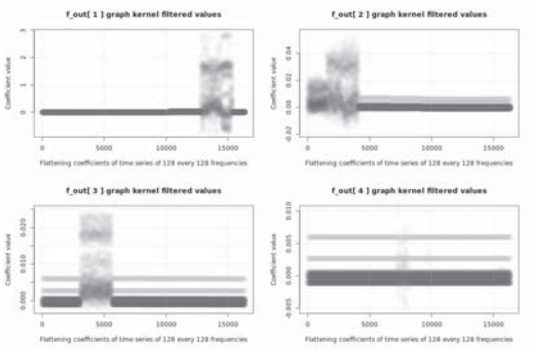


図 12: 楽曲の感性の特徴（愛しいの類似）

賑やかな感性語彙が与えられた際には高周波成分と中周波成分が高く，穏やかな感性語彙が与えられた際には低周波成分が高く，爽やかな感性語彙が与えられた際には高周波成分が高く，感性語彙とそれに紐づく楽曲の感性の特徴の関係が認められ，3 種の方向に集約される固有ベクトルとの関連が示唆された．加えて，感性語彙と楽曲の感性の特徴の関係，楽曲の感性の特徴同士においてそれぞれ類似性が示された．上記より，得られた楽曲の感性の特徴を充当と結論する．

ただし，今回もちいた楽曲の数は 1,090 であり少数であることをふまえ，今後はより大規模な楽曲の感性の特徴を求め，より微細な特徴の獲得を可能とすること，楽曲の感性の特徴と感性語彙に関係が見られることより，その関係を正準相関と仮定し，深層学習の学習過程をもちいた，楽曲の感性の特徴のさらなる追求を今後の課題とする．

5. 謝辞

本件を進めるにあたり，日常の議論を通じての多くの知識や示唆を頂いた株式会社ソケッツの皆様に感謝します．協力していただいた皆様へ心から感謝の気持ちと御礼を申し上げたく，謝辞にかえさせていただきます．

参考文献

[1] Hevner, K. (1935b). Expression in music: A discussion of experimental studies and theories. Psychological Review, 42, 186-204.

[2] C.J.Clopper, E.S.Pearson, (1934). Theuseofconfidence-orfiducial limits illustrated in the case of the binomial, Biometrika 26 404416.

[3] Krizhevsky, A., Sutskever, I. and Hinton, G. (2012). ImageNet classification with deep convolutional neural networks. In Proc. Advances in Neural Information Processing Systems 25 10901098.

[4] Marr, D. And Hildrech, E. (1980). Theory of edge detection. Proc. R. Soc. Long. B, 207:187-217.

[5] 鈴木昇一. (1994). パターンのエントロピーモデル，電子情報通信学会論文誌 (D-II)，vol.J77-D-II, no.10, pp.2220-2238.

[6] A.Cochocki, Rolf Unbehauen. (1993). Neural Networks for Optimization and Signal Processing 1st. John Wiley and Sons, Inc.

[7] Hevner, K. (1937). The affective value of pitch and tempo in music. American Journal of Psychology, 49, 621-630.

[8] Michal Defferrard, Xavier Bresson, Pierre Vandergheynst. (2016). Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. Advances in Neural Information Processing Systems 29.

[9] J.O.Pickles, 谷口郁雄監訳. (1995). 聴覚生理学, 二瓶社.

[10] J.A.Russell. (1980). A circumplex model of affect. Journal of Personality and Social Psychology, Vol. 36, pp. 1161 1178.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

機器の振動データの非負値行列因子分解による特徴抽出と異常検知

Feature Extraction and Anomaly Detection by Non-negative Matrix Factorization of Vibration Data of Machine

関根 理敏 *1　小林 一樹 *1　伊加田 恵志 *1
Masatoshi Sekine　Kazuki Kobayashi　Satoshi Ikada

*1沖電気工業株式会社 経営基盤本部 研究開発センター
Corporate Research and Development Center, Oki Electric Industry Co., Ltd.

In recent years, in the field of manufacturing industry, there is an increasing need for anomaly detection by collection and analysis of sensor data for predictive maintenance of factory equipment and industrial product. In the conventional method, for example, attention is focused on the power of a specific frequency band related to the anomaly of the equipment, and anomaly detection is performed by comparing the magnitude of the power. However, when there are many or wide frequency bands involved in anomaly, they may not be clearly specified. Therefore, we propose a method for anomaly detection using non-negative matrix factorization which is an unsupervised learning algorithm. Even if the frequency band of the vibration data is not clearly specified, our proposed method can automatically extract the features of vibration data and detect anomaly of machines accurately.

1. はじめに

近年，主に製造業の分野において，工場設備や工業製品の機器の予防保全に向けたセンサデータの収集・分析による異常検知のニーズが高まっている [METI 17]．そこでは機器の状態監視において，ネットワーク化された振動センサ，音響センサ等を配置し，それらのセンサデータをリアルタイムに収集・分析し，観測対象の故障の兆候を自動的に把握する．これによりメンテナンスのタイミング等を最適化でき，機器のダウンタイム発生による稼働率低下を抑制することが可能となる．

機器の異常検知において，従来手法では一例として，機器の正常・異常に関連する特定の周波数帯のパワーに着目し，そのパワーの大小を比較することによって異常検知を行っている．しかしながら，正常・異常に関与する周波数帯が複数ある場合や，広域にまたがる場合，また検知対象に個体差がある場合等では，周波数帯を明確に特定し，パワーの大小から異常検知を行うのが困難な場合がある．

そこで本稿では，教師なし学習アルゴリズムである非負値行列因子分解 (Non-negative Matrix Factorization (以下，NMF))[Lee 99] を利用し，学習段階と判別段階で NMF の基底行列を共通化することで，係数行列を特徴量とした異常検知手法を提案する．機器を利用した異常検知の性能評価を行った結果，振動データの自動特徴抽出によって，異常部位の動作区間における異常検知が可能であることが分かった．

2. 異常検知手法

2.1 学習段階におけるパワースペクトルからの NMF による特徴抽出と学習モデルの構築

図 1 に本手法における学習段階の動作を示す．学習データとして，正常状態のみ利用する場合と，正常・異常の両方の状態を利用する場合があるが，以下は前者の場合について説明する．まず正常状態における振動データからそれぞれ平均値を引き，高速フーリエ変換（Fast Fourier Transform (FFT)）を行うことにより，パワースペクトルを求める．以下の式 (1) で

連絡先: 関根理敏，沖電気工業株式会社 経営基盤本部 研究開発センター，埼玉県蕨市中央 1-16-8，sekine566@oki.com

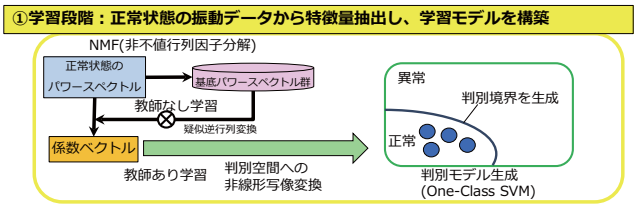


図 1: 学習段階

示すように，正常状態のパワースペクトル $\boldsymbol{Y}_l$ に NMF を適用し，係数行列 $\boldsymbol{W}_l$ と基底行列 $\boldsymbol{H}_l$ に分解する．

$$\boldsymbol{Y}_l \simeq \boldsymbol{W}_l \boldsymbol{H}_l \tag{1}$$

ここで係数行列 $\boldsymbol{W}_l$ の各成分は，対応する基底行列 $\boldsymbol{H}_l$ の各ベクトル成分の，正常状態のパワースペクトル $\boldsymbol{Y}_l$ における含有情報を示し，パワースペクトルの特徴を表す．本稿では，このように係数行列を特徴量として学習モデルを構築する．この場合，構築した学習モデルにより判別を行うには，学習段階における基底行列 $\boldsymbol{H}_l$ を共通の基底行列として判別段階でも利用する必要がある．そこで正常状態のパワースペクトル $\boldsymbol{Y}_l$ に基底行列 $\boldsymbol{H}_l$ の逆行列をかけて係数行列を求める．しかしながら，基底行列 $\boldsymbol{H}_l$ は一般に逆行列を持たない．そこで，ムーア・ペンローズの擬似逆行列 $\boldsymbol{H}_l{}^+$ を求め，学習段階においても判別段階においてもこの疑似逆行列 $\boldsymbol{H}_l{}^+$ により係数行列を抽出する．よって，学習時においては以下の式 (2) を用いて係数行列 $\boldsymbol{W}_{lf}$ を求め，この係数行列を One-Class SVM に入力しモデルを構築する．

$$\boldsymbol{W}_{lf} = \boldsymbol{Y}_l \boldsymbol{H}_l{}^+ \tag{2}$$

ここで係数行列 $\boldsymbol{W}_{lf}$ は $||\boldsymbol{Y}_l - \boldsymbol{W}_{lf}\boldsymbol{H}_l||^2$ を最小にする最小二乗解となる．

2.2 判別段階における分類スコアの算出と正常・異常判別

図 2 に本手法における判別段階の動作を示す．判別対象のパワースペクトル $\boldsymbol{Y}_t$ に対して学習段階と同様に，以下の式 (3)

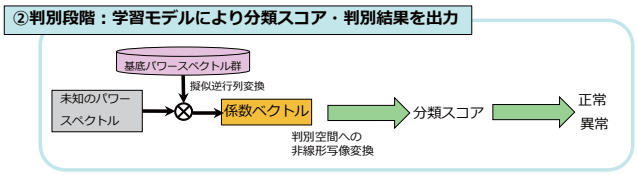


図 2: 判別段階

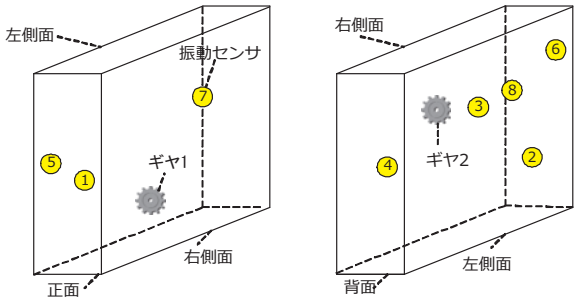


図 3: 振動センサ 8 箇所とギヤ 2 箇所の設置位置

で示すように基底行列の擬似逆行列 $\boldsymbol{H}_l{}^+$ をかけることで係数行列 $\boldsymbol{W}_{tf}$ を求める．

$$\boldsymbol{W}_{tf} = \boldsymbol{Y}_t \boldsymbol{H}_l{}^+ \tag{3}$$

学習モデルは係数行列 $\boldsymbol{W}_{tf}$ の各ベクトルごとに分類スコアを算出し，異常度の推定及び判別を行う．なお，分類スコアは判定境界までの符号付き距離であり，異常度が高いほど値が小さくなる．

3. 性能評価

本手法の評価を行うため，複数のモータ及び機械式の可動部を有する機器の正常・異常時の異常度推定による判別を行った．機器はシーケンス動作 30 秒間を 1 サンプルとして正常・異常各 10 サンプル分の振動データを取得した．機器の異常を，歯の欠け（ギヤ 1），回転の偏心（ギヤ 2）があるギヤにそれぞれ交換することにより発生させた．また検知周波数範囲が 10Hz から 15kHz までの同機種の振動センサを図 3 のように 8 箇所配置した．サンプリング周波数は 100kHz とした．

図 4 に 1 シーケンス動作 30 秒間の，図 3 における振動センサ 1 で取得した信号及びスペクトログラムを示す．スペクトログラムはウィンドウ幅 0.1 秒，シフト幅 0.05 秒とした．また図 5 に 1 シーケンス動作におけるギヤ 1，2 の動作期間を示す．動作期間はシーケンス動作全期間における NMF の係数の時系列の振幅の立ち上がり時点，立ち下がり時点から求めた．

スペクトログラムにおける 0.1 秒間のパワースペクトルごとに正常 5 サンプルで One-Class SVM で学習を行った．残りの正常データ 5 サンプル及び異常データを判別データとして交差検定（分割数 2）を行った．また，ギヤ 1, ギヤ 2 の動作が重複しない動作期間を評価対象とし，ギヤ 1 は 2.8 秒間，ギヤ 2 は 0.75 秒間における 0.05 秒毎の分類スコア値の平均値を閾値判別した．閾値は判別正解率が最も高くなる値に設定した．

表 1，表 2 にそれぞれギヤ 1, ギヤ 2 の判別結果を示す．実験の結果，ギヤ 1 については図 3 における振動センサ 2 または 8，ギヤ 2 については振動センサ 2 のデータをそれぞれ用いることで，最も精度よく正常・異常判別できることが分かった．

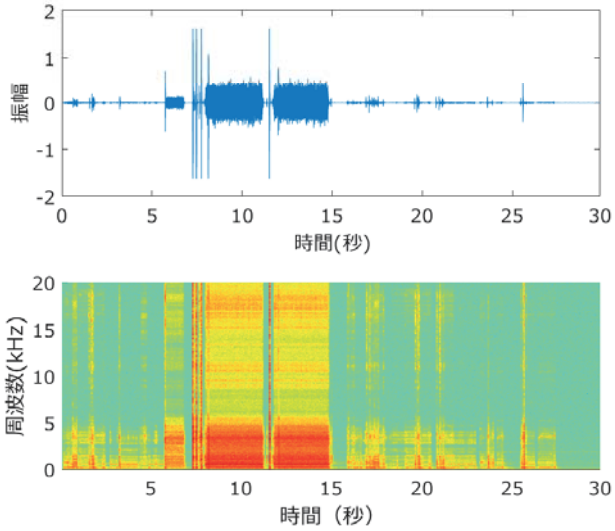


図 4: 1 シーケンス動作における信号（上）とスペクトログラム（下）

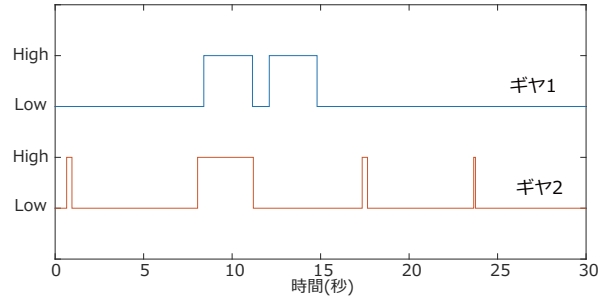


図 5: 1 シーケンス動作におけるギヤの動作区間 (High の区間)

4. おわりに

本稿では機器の異常検知において，NMF で振動データの特徴を抽出し，機械学習により機器の異常検知を行う手法について示した．機器を対象とした正常・異常状態の振動データを利用した評価の結果，ギヤの動作区間において，異常時においては正常時と比べて異常度が増加し，異常検知が可能であることが分かった．今後は機器の劣化経緯を観測した振動データを利用した異常の予兆検知を行う予定である．また正常状態において，複数の動作モードが存在する場合における効率的かつ高精度な異常検知手法の検討を行う予定である．

参考文献

[METI 17] 経済産業省，2017 年版ものづくり白書.

[Lee 99] D.D. Lee and H.S.Seung,„“Learning the part of objects by non-negative matrix factorization,” Nature, vol.401,no,pp.788-791,Oct.1999.

表 1: ギヤ 1 の判別正解率 (%)

真＼判別	正常	異常
正常	100	0
異常	0	100

表 2: ギヤ 2 の判別正解率 (%)

真＼判別	正常	異常
正常	100	0
異常	5	95

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

1d CNN-LSTM による調節弁内部の異常検知

Anomaly Detection in Control Valves by 1d CNN-LSTM

茂木 貴弘[*1] 中澤 友哉[*1] 田原 鉄也[*1]
Takahiro Motegi Yuya Nakazawa Tetsuya Tabaru

[*1] アズビル株式会社
Azbil Corporation

This paper proposes an anomaly detection system for time series data by using 1d CNN-LSTM networks, which is a combination of one dimensional convolutional neural network (1d CNN) and long short time memory (LSTM). The CNN processes a time series data from a sensor and its output is passed to LSTM. After training of the networks, the CNN grows into an appropriate filter to emphasize features of the time series in frequency domain, and the LSTM becomes a good feature extractor. The extractor works fine even under noisy conditions. It generates a feature vector and the vector is utilized to diagnose anomalies. We applied the proposed system to vibration sensor data obtained from a control valve. The system detected anomalies due to cavitation, which is one of the most serious phenomenon of control valves, with 99.5% accuracy. The result shows availability of our proposed method.

1. はじめに

本稿では、時系列一次元信号の分類タスクによる異常検知アプリケーションとして、振動センサデータを分析した事例について報告する。

石油・化学業界をはじめ各種プラントにおいては、流量等を制御する操作端として調節弁が用いられており、調節弁内部で発生する異常の一つとして、キャビテーションと呼ばれる現象がある。キャビテーションとは、流体が弁を通過する際に、流体圧力がその流体の飽和蒸気圧以下に下がり再度圧力回復する過程で、流体中に気泡が発生、消滅する一連の現象である[azbil 13]。キャビテーションによる衝撃圧は数百気圧にも達し、弁内部を損傷させるため、放置し続ければ流体漏洩などの事故につながる可能性もある。そのため、プラントの予防保全の観点からキャビテーションを検出することは重要である。

そこで、キャビテーションを自動で検出することを目的に、調節弁本体の外部に振動センサを取り付けて信号を取得し、振動信号を分析する手法を検討した。振動信号の例を図 1 に示す。

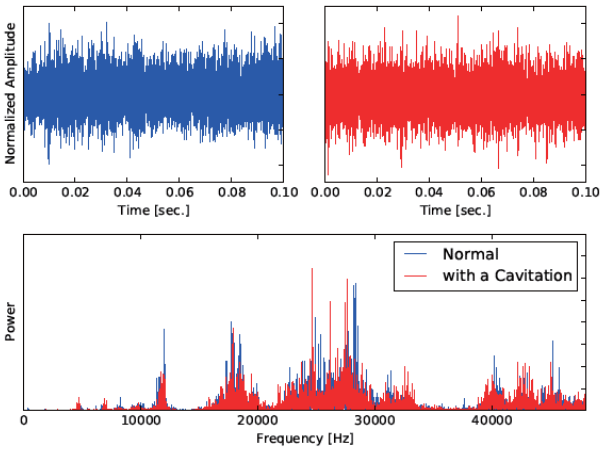


図 1 分類対象の一次元信号

図 1 の青線はキャビテーション無しの波形であり、流体による振動がノイズとして計測されている。図 1 の赤線はキャビテーション有りの波形であり、流体による振動に加えて、非定常にキャビテーションによる振動が混在している。図 1 で比較してわかるように、信号対雑音比は小さく、一般には判別することは難しい。

よって、振動信号からキャビテーション有無を自動で分類する手法として、CNN (Convolutional Neural Network) [LeCun 98] と LSTM (Long Short Term Memory) [Hochreister 97] に着目した。本稿では、これらを組み合わせ、教師あり学習により、キャビテーションの特徴量を自動で抽出し分類するタスクを試みた。

2. 1d CNN-LSTM による異常分類方法

2.1 1d CNN

一般に、CNN は画像データ等の多次元データを対象に利用されており、画像の局所的な特徴量を学習できることが知られている。先行研究では、一次元信号をスペクトログラム等の二次元データへ変換し、一次元信号を画像として CNN で分類タスクをする試みも行われている [Lostanlen 16]。しかし、ニューラルネットワークへ信号を入力する度にスペクトログラムへ変換する計算コスト、本来一次元である信号をわざわざ CNN で二次元分析する計算コストが必要となり、計算能力の低い現場機器へ学習済みモデルを搭載し運用する用途には適さないという課題がある。また、一次元信号をスペクトログラムへ変換する過程で位相情報が損なわれるため、分類精度へ影響が出る可能性もある。

そこで、一次元信号を二次元化することなく入力できる 1d CNN に着目した [Collobert 16]。1d CNN の畳み込み層は入力信号と同じく一次元の畳み込みフィルタを用いて時間方向へストライドさせる演算となるが、基本的に通常の CNN と同様に実装できる。1d CNN では、分類対象の信号の特徴的な周波数帯域を強調するような畳み込みフィルタを学習により獲得できる。

また、通常の CNN では次元削減のためにプーリングを用いることが定石であるが、1d CNN の場合はプーリングによるダウンサンプリングでエイリアシングが発生し、高周波成分が損なわれることが懸念されるため、本稿では用いないこととした。

連絡先： 茂木貴弘, アズビル株式会社 技術開発本部,
神奈川県藤沢市川名 1-12-2, t.motegi.zg@azbil.com

2.2 LSTM

CNN は、局所的な周波数帯域での特徴量学習に適した手法である一方、白色雑音等の広帯域な雑音が混在する分類タスクは苦手と考えられる [三村 17]。図 1 に示すように、分類対象の信号の周波数帯域は広帯域で重複しており、流体による振動に比べてキャビテーションによる振動は小さいことから、1d CNN のみにより分類タスクを高精度で実現することは難しいと考えられる。そこで、LSTM に着目した。LSTM は、過去のタイムステップの信号値を長期間保持できる特性があるため、白色雑音等の広帯域な雑音への耐性が期待できる。

2.3 1d CNN-LSTM

上記考え方に基づき、図 2 のニューラルネットワーク(評価モデル I)を設計した。ネットワークの詳細設計は表 1 に示す。

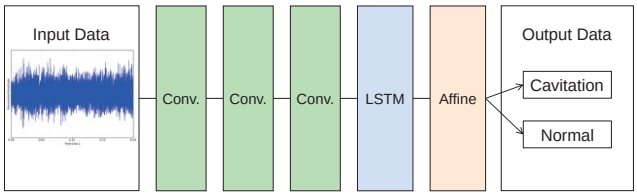


図 2 ニューラルネットワーク構造

表 1 ニューラルネットワークの詳細条件

レイヤーの設定	
入力層	サイズ：(10000, 1)
畳み込み層①	フィルタ：(100, 1)×1 ストライド：2 バッチ正則化あり
畳み込み層②	同上
畳み込み層③	同上
LSTM 層	ドロップアウト率：0.1
全結合層	活性化関数：シグモイド
出力層	サイズ：(2, 1)

3. キャビテーション検出実験

3.1 データセット

流量実験設備に組み付けられた調節弁本体外部に振動センサを設置し、流体条件を変化させたときの振動をサンプリング周波数 96 kHz で収録した。実験時に透明流路を用いてキャビテーション有無を目視確認し、ラベル付けした。ラベル付けされた収録データは、信号長 10,000 サンプルで区切り、データセットとした。また信号の振幅のばらつきが学習に影響しないように、データセットの振幅は実効値で正規化した。キャビテーションあり 1,200 セット、キャビテーションなし 1,600 セットからなるデータセットから、検証データとして各 100 セットをランダムに選択し、残りを訓練データとした。

3.2 学習条件

学習を効率的に進めるために、バッチ数 200 のミニバッチ学習を行った。損失関数はソフトマックス交差エントロピとし、最適化アルゴリズム RMSprop としてネットワークの更新を行った。

評価モデル I との比較として、1d CNN のみの評価モデル II、LSTM のみの評価モデル III についても検証を行った。評価モデル II、III のレイヤーの設定は、表 1 の畳み込み層、LSTM 層と同様とした。

4. 実験結果

学習回数と分類精度の関係を図 3 に示す。評価モデル II は、学習が収束せず、検証データでの最高分類精度は 90%であった。本稿で着目した評価モデル I は、検証データでの最高分類精度は 99.5%で収束しており、極めて高い分類精度が確認された。一方で、評価モデル III は過学習状態となっており、汎用的な特徴量を学習できていない。その理由については今後も引き続き考察が必要であるが、本実験結果より、1d CNN と LSTM を組み合わせることで時系列一次元信号をより高精度で分類できることがわかった。

5. まとめ

本稿では、教師あり学習により時系列一次元信号から特徴量を自動で抽出できる 1d CNN-LSTM を用いて、キャビテーションを検出した事例を報告した。

実験結果より、評価モデル I は 99.5%という極めて高い分類精度を確認できた。よって、本手法が調節弁内部の異常を早期に検知し、予防保全に貢献できる可能性が示された。

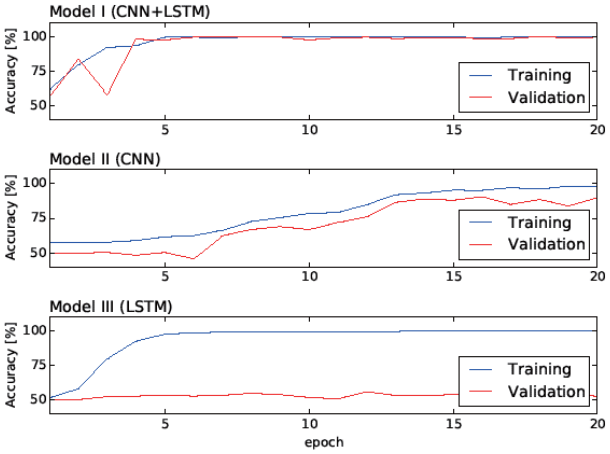


図 3 実験結果

参考文献

[azbil 13] アズビル株式会社編：調節弁のキャビテーションについて，調節弁用計装資料, ID1-8000-4600, アズビル株式会社 (2013)

[LeCun 98] Y. LeCun: Gradient-based learning applied to document recognition, Proceedings of the IEEE, Vol.86, pp. 2278-2324 (1998)

[Hochreister 97] S. Hochreiter: Long short-term memory, Neural computation, Vol. 9, pp. 1735-1780 (1997)

[Lostanlen 16] V. Lostanlen and C.-E. Cella：Deep Convolutional Networks on the Pitch Spiral for Musical Instrument Recognition, Proceedings of the ISMIR2016 (2016)

[Collobert 16] R. Collobert, C. Puhrsch, and G. Synnaeve: Wav2Letter: an End-to-End ConvNet-based Speech Recognition System, arXiv:1609.03193 (2016)

[三村 17] 三村正人: 深層学習に基づくフロントエンド特徴量強調と頑健な音声認識，日本音響学会誌，73 巻 1 号，pp.47-54 (2017)

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

認知症ケア高度化のための協調学習環境を活用した認知症の見立てコーパスの構築

Constructing Diagnosis Corpus Using Collaborative Learning Environment for Improving Dementia Care

神谷 直輝 *1 Naoki Kamiya　上野 秀樹 *2, 3, 4 Hideki Ueno　吉沢 拓実 *1 Takumi Yoshizawa　石川 翔吾 *1 Shogo Ishikawa
今田 兼太 *1 Kenta Imada　西山 明也斗 *1 Hayato Nishiyama　桐山 伸也 *1, 4 Shinya Kiriyama　玉井 顯 *1 Akira Tamai　竹林 洋一 *1, 4 Yoichi Takebayashi

*1 静岡大学 Shizuoka University　*2 千葉大学医学部附属病院 Chiba University Hospital　*3 敦賀温泉病院 Tsuruga Onsen Hospital
*4 みんなの認知症情報学会 The Society of Citizen Informatics for Human Cognitive Disorder

This paper describes construction of diagnosis corpus using collaborative learning environment for improving dementia care. Analyzing causes of reversible dementia are essential to improve dementia care, and which are interpreted diversely. We developed the diagnosis corpus using the collaborative learning environment to construct diagnostic knowledge by accumulating interpretations. Our corpus enable us to analyze learning effect and improve knowledge of dementia diagnosis.

1. はじめに

認知症の「状態」を引き起こす認知機能低下の原因は 70 種以上も存在する. その中には治療可能な医学的疾患も多く存在することが知られており, 医療の関与が欠かせない. 脳機能の低下を引き起こしている原因を分析することを本稿では認知症の「見立て」と呼ぶ. 医師の見立ての知識を学ぶことで, 医師が必要とする情報を理解することができ, 適切な情報提供によって認知症ケアの質が改善することが期待できる. また, 見立てのノウハウを学ぶことで支援対象者に対する理解が深まり, 自信を持って支援に当たることができるようになり, 支援の質が改善することが期待される.

筆者らは多職種で見立て知が学べる学習環境を構築してきた[橋詰 2017]. 家族や介護従事者, 看護師, 医師が認知症のケースを軸に多面的に議論することで, 気づきが増え, 実践的な見立ての理解が促される. 本稿では, 学習者の意見や気づきを収集し, 協調学習で生み出される見立て知として蓄積・利用を促すための見立てコーパスの構築について述べる.

2. 認知症の見立て知の協調学習

認知症とは, “一度正常に発達した認知機能が後天的な脳の障害によって持続的に低下し, 日常生活や社会生活に支障をきたすようになった状態”である[日本神経学会 2010]. 特に治療可能な認知症(Reversible Dementia[Chari 2015] / Treatable Dementia[Mahler 1987])への対応が重要である. 治療可能な認知症とは, 脳細胞は死滅しておらず, 脳機能を低下させる原因を治療で取り除くことで脳機能の回復が可能な認知症である. この治療可能な認知症に気づくポイントには, 歩行や排泄, ふらつき, 食欲, 睡眠の状況など日々の暮らしで観察できる項目が多くある. 身近にいる支援者がいち早く気づけるようにするためには, 医師の見立ての知識の理解が欠かせない. そして, 実生活で見立てを行い, 支援の質を改善することが必要である.

実践的な学びにつなげる手法には, 医学や経営学, 法学の分野で有用なケースメソッドがあり[Barnes L.B. 1984], 筆者らはその手法を発展させた協調学習会を開催してきた. 認知症のケース検討を主軸とした「①個人での見立て, ②グループでの見立て, ③総合討論」という流れを設計した. 多様な学習者が各自の立場・経験をもとに考え, 多角的に議論することで見立ての理解が促される. 個々の学習者の知の獲得を支援する取り組みはあるが[Nelson B. 2012], 学習者と教師の議論や気づきなど学びのプロセスを蓄積することで, 見立て知の深化につながると期待される.

3. 協調学習環境に基づく見立てコーパスの設計

継続的な見立て知の構築と発展に向けて, 見立てコーパスを機軸とした協調学習環境を構築した(図 1). 見立てコーパスには, 専門家の知識を付与した病院のケースデータが蓄積されており, そのデータをもとに専門家と研究者が協調学習会で用いる検討ケース, 及び検討ケースに対する質問レポートを作成する. それらを用いた協調学習会を設計・開催し, その結果を見立て知として本コーパスに蓄積する. この協調学習環境によって家族や介護従事者, 専門家の知識・経験が織り交ざった新しい知が形成され, 見立て知を深化成長できると考える.

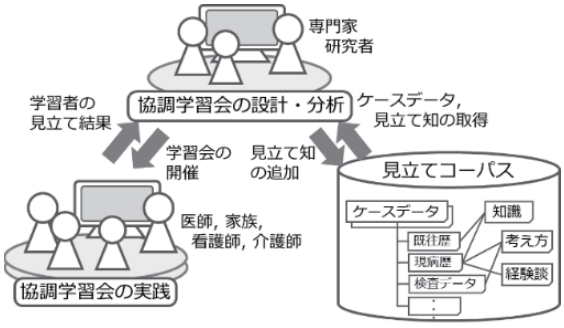


図 1 見立て知の協調学習環境

連絡先：神谷直輝, 静岡大学, 静岡県浜松市中区城北 3-5-1, 053-478-148, kamy@kirilab.net

協調学習会を効果的に学ぶ場にするために，質問レポートをタブレット型端末上で提示するツールを開発した．学習者の見立てた結果や意見を入力ログとして 5 秒毎にデータベースに送信する．連続的な入力内容の遷移を収集することによって「②グループ」での見立てや「③総合討論」での議論を通した個人の考え方や思考プロセス等の学びを表出化につなげる．また，ツールを用いた遠隔参加も可能である．議論に参加できる学習者が増えることで様々な観点の知を収集することができるようになる．このように収集された入力データを専門家や研究者が分析し，新たな知として見立てコーパスに蓄積する．

見立てコーパスには精神科医 2 名とのケースカンファレンスに基づいた病院の外来患者のケース 22 事例が蓄積されている．見立てでは数十項目以上の医療情報や詳細な生活情報等が認知症の人ごとに適切に用いられる．そこで，見立てに不可欠な情報を選定するため，外来患者のケースを用いたケースカンファレンスを開催し，医師に見立ててもらい，患者情報中の必要な項目を抽出した(表 **1**)．Web 会議ツール上で患者情報を医師に提示し，見立てのプロセスで着目する情報と理由，関連する知識等の解説を収集した．その結果をもとに患者情報を登録するための基本的な構造を設計した．医師 2 名に本コーパスの構造と入力済みデータを確認してもらった所，治療可能な認知症の検討に不要な項目は無いとの意見を得た．

表 1 治療可能な認知症の検討に用いる項目の一部

記載欄・検査名	項目
家族歴	家族構成，家族の状況
既往歴	聞き取り内容，HT，DM，輸血
本人歴	教育歴，職歴，家庭内状況，対人関係，趣味等
現病歴	聞き取り内容，身体状況(血圧，飲酒，排泄等)
服薬情報	一般名・含量・用法
認知機能評価	BFB，MMSE，HDS-R，行動観察方式 AOS, ADL
CT, MRI	医師の所見

4. コーパスに基づく学習効果の分析と見立て知の表現

見立てコーパスに継続的に知を蓄積・深化させるために，本コーパスを機軸とした協調学習環境を福井県敦賀市，若狭地区，藤沢市にて構築し，介護従事者を中心とした協調学習会を開催した．

収集した学習者の意見は，本コーパスの構造として予め蓄積された医師の見立て方との違いや，コーパスの構造で表現できない学習者の意見・考え方を見つけることができる．その例として，1 つの検討ケースに対し 10 名の学習者が治療可能な認知症を判別する際に参照した情報と，詳しく見立てるために必要だと指摘した情報の傾向を図 **2**に示す．特にエピソード(現病歴)の参照が多く，認知機能障害や生活障害を読み取る学習者が多かったことが分かる．とりわけ参照が少なかった項目には，現在の能力を以前と比べて理解するための職業歴や学歴，脳機能低下の原因を探るために不可欠な薬物歴への参照が少なく，それらの情報の活用した学習者は一部だったと分かる．しかし，その一部の学習者が生活支援や医学的な知識を用いた意見・見立てを他の参加者と共有することで，新たな意見や見立て知につながると考えられる．

協調学習会で得た意見や考え方を収集し，本コーパスの構造に応用する (図 3)．学習者の専門的知識や経験にもとづく意見を活用することで，ケースの情報から治療可能な認知症の可能性を検討できるようになる．このように協調学習会で得た知を蓄積することで，ひとつの事例の見立てを深めることや，見立て知を表現するコーパスの構造の改良・洗練につながることが期待される．

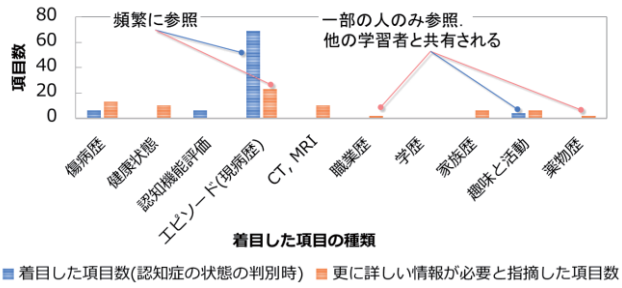


図 2 学習者が見立てに用いた情報の傾向

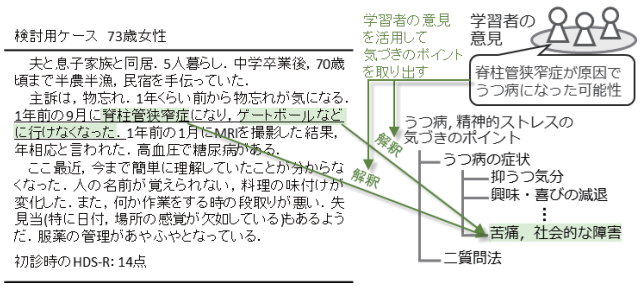


図 3 学習者の意見・考え方の活用例

5. おわりに

本稿では，見立てコーパスを機軸とした協調学習環境に基づく協調学習会を開催し，学習者の意見や考え方を活用することで見立てコーパスを構築・改良できることを示した．さらに，構築した本コーパスによる見立て知の学びの評価に加えて，学びが見立て知の深化につながる可能性が示された．今後も継続的に見立てコーパスを発展させ，見立て知を学ぶことでケアの質が向上するか検証する．

参考文献

[橋詰 2017] 橋詰裕樹 他：協調学習環境を活用した認知症の見立て知の学びと実践，第 31 回人工知能学会全国大会，1I2-NFC-02a-2in2 (2017).

[日本神経学会 2010] 日本神経学会監修：認知症疾患治療ガイドライン 2010, 医学書院 (2010).

[Chari 2015] Chari D, et.al., Reversible dementia in elderly: Really uncommon?, J. Geriatr Ment Health, 2:30-7 2015.

[Mahler 1987] Mahler ME, et. al., Treatable Dementias, Western Journal of Medicine, 146(6):705-712 (1987).

[Barnes L.B. 1984] Teaching and the Casemethod, third edition, Harvard Business School Press (1984).

[Nelson B. 2012] Baloian N., et.al., “Ubiquitous Mobile Knowledge Construction in Collaborative Learning Environments.” Sensors (Basel, Switzerland) 12.6 (2012).

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

個性に基づくケアのための認知症ケア協調学習環境の構築と実践

Construction and Practice of Collaborative Learning Environment for Personality-focused Care in Dementia Care

小俣 敦士 *1 Atsushi Omata
松井 佑樹 *1 Yuki Matsui
石川 翔吾 *1 Shogo Ishikawa
宗形 初枝 *2 Hatsue Munakata
中野目 あゆみ *2 Ayumi Nakanome
香山 壮太 *2 Sota Kayama
伊東 美緒 *3 Mio Ito
坂根 裕 *4 Yutaka Sakane
本田 美和子 *5 Miwako Honda
原 寿夫 *2 Hisao Hara
桐山 伸也 *1*6 Shinya Kiriyama
竹林 洋一 *1*4*6 Yoichi Takebayashi

*1静岡大学 Shizuoka University
*2郡山市医療介護病院 Koriyama Medical Care Hospital
*3東京都健康長寿医療センター研究所 Tokyo Metropolitan Institute of Gerontology
*4株式会社エクサウィザーズ ExaWizards Inc.
*5東京医療センター Tokyo Medical Center
*6みんなの認知症情報学会 The Society of Citizen Informatics for Human Cognitive Disorder

We describe the construction and practice of collaborative learning environment to develop the dementia care skills focusing on elderlies personality. In our learning environment, the policy of care for people with dementia is considered based on their personality and desire. And our learning environment helps to learn collaboratively dementia care throughout the organization by using video. We have conducted a collaborative learning for dementia care in the care fields. The results suggest that the collaborative learning environment is useful for learners.

1. はじめに

認知症高齢者は多様な個性をもち，自分らしく生活したいと願っているが，認知機能の低下に伴う将来の不安や自己喪失感によって行動・心理症状が現れやすい状態にある．この行動・心理症状には，抑うつ，徘徊，暴力なども含まれ，多くの医療介護現場を疲弊させている．一方で，ケアの達人は，身体的なスキルに加えて，相手の個性や願いを適切に理解することで，良好な関係を築くことができている．認知症の人本人の個性や願いを，ケアに関わる全ての人が共有してケアを実践できれば，より良い関係性を築くことができるとともにケアの高度化につながる．

本稿では，認知症高齢者の個性やゴール (願い) といった個性情報に基づいたケアの実践と高度化のための協調学習環境を提案し，有効性について検討する．

2. 構造化映像を用いた協調学習環境

閉鎖的であるケア現場では，どのようにケアを実践したのかを言語で表現することは難しいため，映像を用いてケアの振り返りを支援することで，学びを促進させることが可能である．筆者らは，認知症ケアの現場において組織全体で認知症ケアを学び，高度化するための協調学習環境の構築を進めてきた [宗形 17]．本学習環境では，認知症ケア現場で撮影されるケア映像を基に学習を進める (**図 1**).

- 実践：ケアを実践している様子をビデオ撮影 (図中 2)
- 指導：ケア動画に指導データを付与 (図中 3)
- 学習：指導付きケア動画を視聴し学ぶ (図中 4, 5)

連絡先: 小俣敦士，静岡大学創造科学技術大学院，静岡県浜松市中区城北 3-5-1，omata@kirilab.net

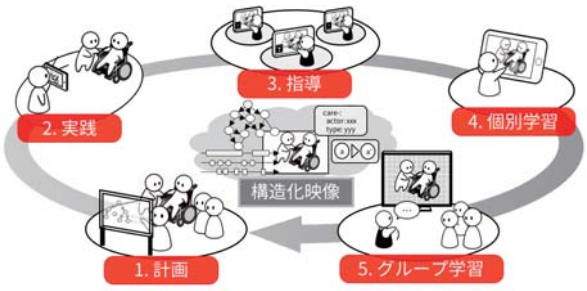


図 1: ケア映像を用いた学びの高度化と知の構造化サイクル

学習者は，ケアのどこに注目すべきか，どのようなことを考えるべきかをといった状況判断能力を養うことや，具体的なスキルを客観的に指摘・アドバイスを受けることができ，スキル向上に効果的である．そして，学びのプロセスを通して取得される多様なデータをケア動画と関連づけることにより，構造化映像としてデータ化することが可能である [小俣 17].

本稿では，上述したスキル習得を支援する枠組みを発展させ，ケア実践の前に，ケア対象者の個性やゴールを検討し，ケアの方針を決めるための計画の場を設計する (**図 1 の 1**).

3. 個性情報に基づくケアの目標設定

私たちは異なる能力や，価値観，社会的な役割を持ち，状況や目的に応じて複数のモデルを切り替えている [Minsky 09]. 認知症の人に関しても，ケアを受ける際の身体的状況や，気分といった心的状況，そしてケアスタッフとの関係性など社会的な側面も影響し，様々なモデル (個性) を切り替えていると考えられる．そのため，各スタッフが日頃ケアを行う中で得たその人に関する知識を，組織やチームで共有し，その人についての理解を深めることがより良いケアを実現する上で役立つ．

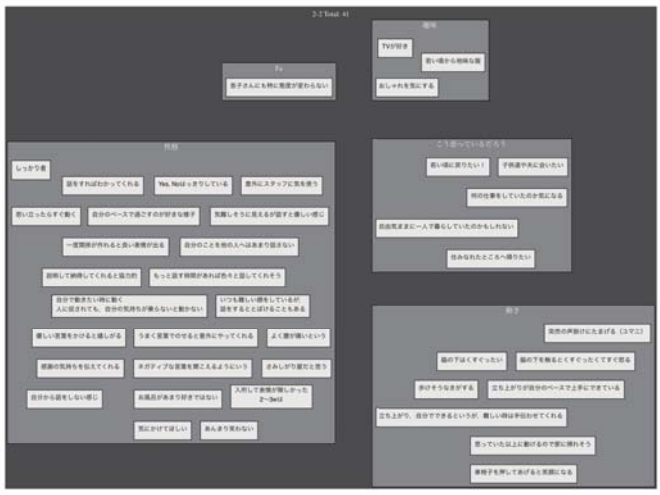


図 2: 個性情報の検討結果の例

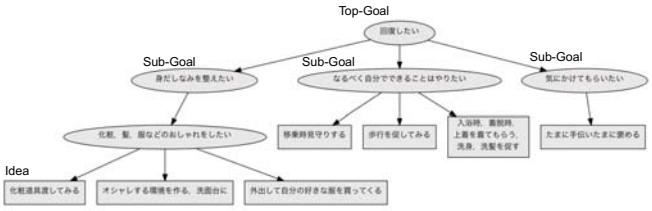


図 3: ゴールと実現可能なケアの検討結果の例

そして，その人がどのようなことを望んでいて，どのような状態を目指すのか，各スタッフが共通の目標を持ちチームとして関わることが重要である．このような観点から，以下のような手続きでグループワークを行いケアの方針を検討する．なお，一名の認知症高齢者の方についてグループワークを実施した際の結果も同時に示す (**図 2**, **図 3**).

1. ブレインストーミング形式で個性，その人らしさ，特徴など個性情報を自由に挙げる (**図 2**)
2. 個性情報をもとに，その人の望んでいること，やりたいことをゴールで表現する (**図 3**)
3. 各ゴールに対して，実現可能なケアを検討する (**図 3**)
4. 実際に実践・撮影するケアを決定する

個性グループワークを実施する際には，可能な限り認知症の人本人や家族も同席してもらい，意見を聞きながら実施するような場を作るようにする．目標をゴールで表現すると複数のゴールの関係性を表現することができ，認知症ケアにおいてはなぜそのケアを行うのかといった行動意図を表現できる．そして，目標を設定したのち，ケアを実践・撮影し，その映像を用いて組織全体で学びを進める．このように，学びをケアを実践する人で閉じるのではなく，同じ場を共有するスタッフ，外部の専門家，認知症の人本人や家族などを巻き込んで，学びを深めることが可能な協調学習環境を持続的に発展させていく．

4. 協調学習環境による学びの効果

福島県の郡山市医療介護病院で協調学習環境の枠組みに基づいて学習を実施した結果について示す．本病院では約 80 人のケアスタッフが所属しており，6〜10 名のグループで学習を進めた．学習前後にケアを実践し，学びの効果を分析した結果を一つの事例を用いて示す．

表 1: 事例情報

ケア対象者	90 代女性
個性情報	みんなと一緒にいるのが好き 社交的な人 お世話をすることが得意 犬を飼っていた 周りの様子をよく観察している 挨拶をすると笑顔で挨拶してくれる 手を握る力が強い 梅酒が好き 息子さんのことは遠くからでもわかる 息子さんととても良い関係　　など
ケアの目標	立って歩く 普段は経管栄養だが口からの食事を楽しむ
実施したケア	おやつの時間にデザートを食べる

まず，学習で対象となった事例の情報を**表 1** に示す。個性グループワークの結果，ケア対象者の方は社交的で，息子さんとの関係が良好であることがわかった．そして，現在の身体機能を伸ばし回復を目指すために,「立って歩く」というゴールと，現在 3 食経管栄養であり，口からの食事を他の人と一緒に楽しめるようになるために,「口からの食事を楽しむ」という目標を設定した．協調学習では「おやつの時間にデザートを食べる」ケアを実践している映像を用いて学習を進めた．約三ヶ月間をかけて，ケア映像に対して指導データを付与し，生成された指導付きケア動画をもとに学習を進め，学習後にも同様のケアを実践した．なお，ケアの実践に対する指導はグループに所属する全スタッフが行い，映像学習についても全スタッフで行った．このように，映像という客観データを用いて，個人の実践経験から組織の学びを深めていく．

学習前後のケアを，スクリプト [Schank 77] という手続き的知識や出来事を表現するための知識表現を用いて比較した結果を**図 4** に示す．筆者らは，認知症ケアにおける様々な状況・場面を，人工知能のモデルを用いて表現するための構造設計を進めており，本稿ではその構造に則ってケアを表現している．ケア映像における，ケアスタッフの行為とケアを受けている認知症高齢者の状態・行為が可視化されており，左から右に向かってケアの流れが表現されている．ケアスタッフの行為に関しては，より抽象的な構造として，認知症ケア技法ユマニチュード [本田 14] の「５つのステップ」を示している．ユマニチュードでは，介入の際の出会いから別れまでのプロセスがモデル化されており，この「５つのステップ」の手順を踏むことで認知症の人との人間関係が形成できると言われている．

- 出会いの準備：相手に自分が来たことを認識してもらうプロセス
- ケアの準備　：関係性を作りケアの合意を得るプロセス
- 知覚の連結　：実際に心地よいケアを行うプロセス
- 感情の固定　：良い時間を過ごせたことを確認するプロセス
- 再会の約束　：次回のケアに繋げるプロセス

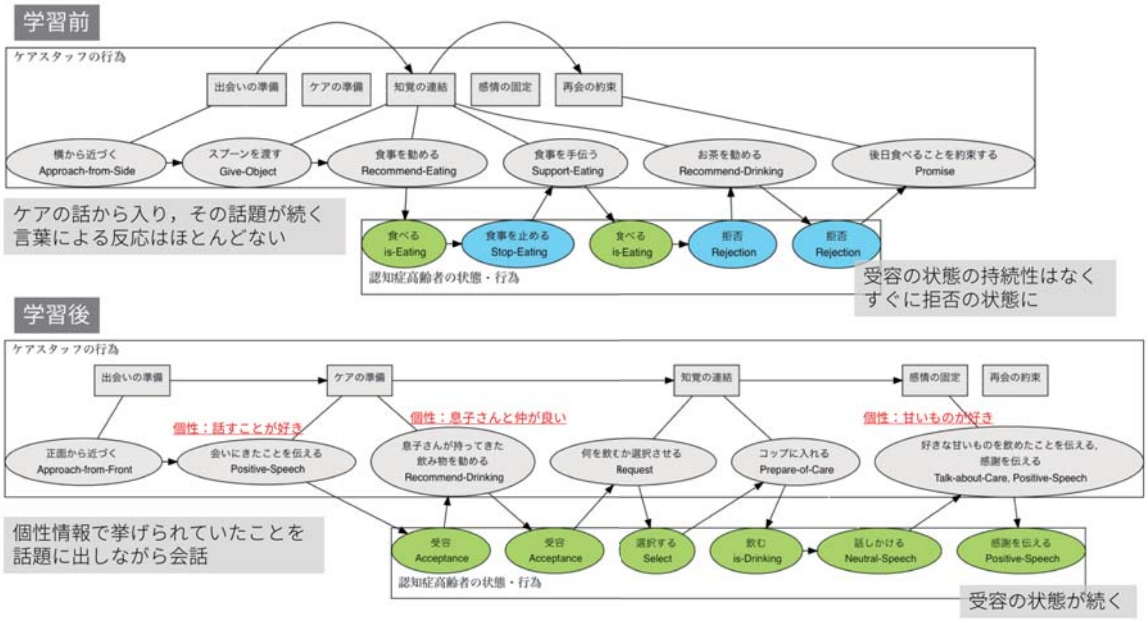


図 4: 学習前後のケアスクリプトの変化

学習前のケアでは，出会ってからスプーンが渡され，その状態のまま食事が促されており，認知症の人は拒否の状態で終わっていることが確認できる．また，ユマニチュードのスキルの観点においても，「ケアの準備」や「感情の固定」が実施されておらず，相手に合意を得ないままケアが実施されていることがわかる．一方で，学習後のケアでは，出会いから会話しながらケアを行い，認知症の人は受容の状態が続いていることが確認できる．また，ユマニチュードのスキルの観点でも，「再会の約束」はないものの，手順を踏んでケアが進行している．個性の観点では，社交的な側面や息子さんとの関係が良いことなど個性情報で挙げられていた内容を話題に出しながら人間関係を構築し，相手からの発話などの自発的な行動を引き出していることが確認できる．この学習前後では，個性グループワークと，実際のケア映像に対して，グループスタッフからの指導が行われた．学習前の映像に対して得られたグループスタッフからのコメントを以下に示す．

- 突然おやつの声掛けになっている
- 視線が合っていない
- いきなりケアの話をしているので，相手のところに行ったらケアの準備として相手との関係性を作ることが大切
- 食べてもらうのではなくてポジティブな人間関係を築くことに目的を立てた方がよりスムーズなケアに結びつく
- 拒否されているので，これ以上やると強制ケアになってしまうので，また次の機会にするとしたのは良い
- まずアイコンタクトをとって，こちらを認識してもらって，そこから関係性が生まれるのでまずは目玉を捉える
- 少し抑揚をつけて，相手に感情が届くように「おいしいねー」とか「いっぱい食べてくれると嬉しいなー」とか，やや大げさな感じで抑揚をつけてポジティブな言葉を選ぶと喜んで食べてくれる可能性があります．

このように，他のスタッフからコメントから，ケアの具体的な改善点やケアの考え方の共有が行われ，ケアスキル向上につながっていることが示された．また，個性に基づき協調的に学習することで，ケアの目標やそのケアを行う意図を共有し，個性に基づくケア実践に繋がることが示唆された．

5. おわりに

本稿では，認知症高齢者の個性に基づくケアのための認知症ケア協調学習環境について述べた．個性やゴールに基づき協調的に学ぶことが，ケアを受ける相手の理解を深めるとともに，ケアを行う際の目標やその意図をチーム全体で共有でき，多様な個性を持った認知症の人に合わせたケア実践につながることが示唆された．今後は，本稿で述べた個性情報やゴールの構造化を進め，協調学習環境の発展と学習効果の評価を進める．

参考文献

[宗形 17] 宗形初枝, 他: 第 31 回人工知能学会全国大会, (2017).

[小俣 17] 小俣敦士, 他: 認知症ケアの内省を促す構造化映像を用いた協調学習環境, 第 8 回高齢社会デザイン研究会, (2017).

[Minsky 09] Minsky, M. 著, 竹林訳:ミンスキー博士の脳の探検-常識・感情・自己とは-, 共立出版, (2009).

[Schank 77] Schank, R. C., et al.: Scripts, plans, goals, and understanding: An inquiry into human knowledge structures., Psychology Press, (1977).

[本田 14] 本田美和子, 他: ユマニチュード入門, 医学書院, (2014).

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

感情思考モデルに基づくマルチモーダル認知症ケア知の共創

Co-production of care knowledge for people with dementia based on an emotion thinking models

石川 翔吾 *1 Shogo Ishikawa　佐々木 勇輝 *1 Yuki Sasaki　桐山 伸也 *1 Shinya Kiriyama　本田 美和子 *2 Miwako Honda　Yves Gineste*3 Yves Gineste
竹林 洋一 *1 Yoichi Takebayashi

*1静岡大学 Shizuoka University　*2東京医療センター Tokyo Medical center　*3Institute Gineste Marescotti Institute Gineste Marescotti

This paper describes a methodology of co-production of care knowledge for people with dementia based on emotion thinking models. Our goal is to construct the AI platform to explain complicated thinking processes in caring. Emotion thinking models are designed to represent knowledge of care experts. The models have multiple representations: “goal”, “trans-frame” and “resourcefulness” on the basis of Minsky’s theory. The result suggests that the representation of thinking process in dementia care leads to encouraging construction of care knowledge.

1. はじめに

筆者らはケアと認知症の人の関係を明らかにする Evidence Based Care (EBC) の実現に向けて認知症の人の情動理解基盤技術の開発を進めている [竹林 15]．ケア従事者のケアスキル，及び認知症の人の情動的行動に着目し，表層的な行動の特徴表現や認知症ケアの評価を行ってきた [石川 15]．認知症ケアは，認知症当事者の身体的，心的状態の評価，変動性への対応，そして認知症当事者の生活の質を高めるためのケアの目的等を考慮した複雑な思考プロセスの結果である．一方でこのような思考プロセスを説明可能なものとして表現する手段が少ないという課題がある．

そこで本稿では，行動評価の仕組みを発展させ，認知症ケアにおける思考プロセスを客観化するための感情思考モデルを構築し，多様な熟達者の意見を集約することで認知症ケア知の構築につながることを示す．

2. マルチモーダル認知症ケアの表現

2.1 認知症ケアにおける感情思考モデルの設計

介護現場では,「優しい」ケア,「尊厳のある」ケアというスーツケースワード（多義的な意味を一つの単語に詰め込むこと）[Minsky 09] で指導されることが多い．筆者らは,「優しさ」や「尊厳」を技術と哲学として具体的に習得可能なものにしたマルチモーダルコミュニケーション技法の ユマニチュード® に着目して,「見る」「話す」「触れる」等の優しさに関するスキルの評価を進めてきた [石川 15]．

一方で，熟達したケア従事者は，本人の尊厳を考慮することを具体的に，１）認知症の人の状態をセンシングし，２）ケアの目標達成のためのプランをその場その場で組み換えながら，臨機応変にスキルを活用している．このような認知症ケアにおける思考プロセスを [本田 17, 石川 15] に基づき，以下の三つの観点から表現することとした．

- 感情的状態：環境要因に対して脆弱な認知症の人が感情的思考を利用しやすい特性から感情的状態としてポジティブ，ネガティブ，ニュートラルな状態を規定する．
- 行動意図：対象の心的・身体的状態を踏まえたゴール設定．
- 知識利用：時系列のケアプロセスのなかでどのように知識を利用したのかを表現する．

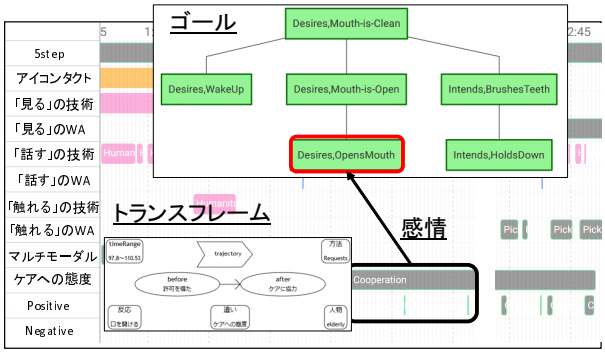


図 1: 口腔ケア場面におけるケアの思考プロセスの表現

本稿では，これらの観点を統合するものを感情思考モデルと呼ぶ．ケア従事者の視点から認知症の人の心の状態を表現することとし，これらのさまざまな思考をマルチモーダルに同時に処理している系として扱う．また，心の状態についての表現系は表 1 の構造を持つ．

2.2 Semantic Primitive の設計と心的状態の表現

[石川 15] で設計した行動レイヤの記述と思考レイヤの記述を接続するための構造として Semantic Primitives を設計した．Semantic Primitives は行為を表現する Action，状態を表現する State，変化を表現する Transition の 3 種類から構成される．一例を表 2 に示す．本 Primitive は思考レイヤを記述するための基本タームとなる．

図 1 では感情思考モデルに基づき口腔ケア場面における思考プロセスを表現している．この場面では，認知症の人がポジティブな状態の時，口腔内を清潔にするというゴールを達成するために，ケアの協力が得られた場面を示している．これらの思考プロセスを表現するために，行動レベルの記述，ルー

連絡先: 石川翔吾，静岡大学，静岡県浜松市中区城北 3-5-1，053-478-1488，ishikawa-s@inf.shizuoka.ac.jp

表 1: 心の状態の表現の一例

Category	構造	意味
感情的状態	Positive :actor :object :action	ACTOR が何に対して positive な ACTION を表現したか
行動意図	Desires :actor :prop	ACTOR は PROP が実現することを望む
	Intends :actor :prop	ACTOR は ACTION を取る計画を持っている
知識利用	Trans-frame :actor :before :after :goal :gap :prop	ACTOR は前後の変化でどの FRAME を利用したのか

表 2: 口腔ケアにおける Semantic Primitives の例

Category	Type	フレーム構造
Action	Look	Looks at Oral :actor :object
	Speech	Requests :actor :object
	Touch	Grasps :actor :object
	Action	is-walking :actor
State	Physical	Mouse-is-Opening :person
		Eyes-are-Closing :person
	Environment	Bed-is-Reclining
Transition	Mental-trans	ToPositive :person
		ToNegative :person
	Physical-trans	LocationChange :person
	Abstract-trans	Attitude :person

ルベースの Semantic primitive の生成，思考プロセスの記述，記述結果の検索・可視化を支援するツールを設計し，認知症ケア知の共創を実現する．

3. 感情思考モデルに基づく認知症ケア知の共創

設計した感情思考モデルに基づき認知症ケア知の共創について検証した．特別養護老人ホームを利用する 2 名の認知症の人（アルツハイマー型認知症）に対してケアを実施した．ビギナーレベルのスタッフ，ユマニチュードインストラクターの映像事例 2 事例を評価対象とした．ユマニチュードインストラクター 6 名が個別に映像事例を視聴し，提示されたベースデータに対して追加・修正点として意見が集約された．

図 2 はビギナーレベルのスタッフの口腔ケア映像に対して，ゴールの観点から意見を集約した結果である．口腔ケアをする際に口を開けることを促す行為に対し，合図をして，押さえて，催促し，指示するという意図があったことがサブゴールとして追加された．また，「口腔ケアをしやすくしたい」というゴールが表現されていたが，同じ人に対するケアの類似事例を参照しながら映像を視聴することで，全てのインストラクターがこのゴールの表現は適切ではないという結果が得られた．

図 3 はインストラクターによる移乗ケア映像に対して，どのように考えていたのかを集約した結果である．感情の状態がネガティブからポジティブに遷移した中で，高齢者の身体的状態をセンシングしながらケアの中で elderly frame の slot を修正している．また，インストラクターは立つ力を確認するために，事故を予防して膝を支え，一連の流れで自然に歩く行為に移れるように促し，万一の場合にすぐに座れるようにケアしているという意見がゴールに集約された．

このように，変化，状態，意図という観点から構築した感情

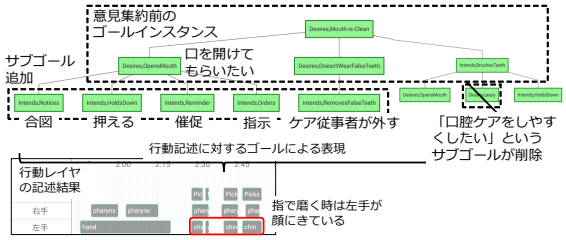


図 2: ケア熟達者の意見集約による初学者の行動意図の表現

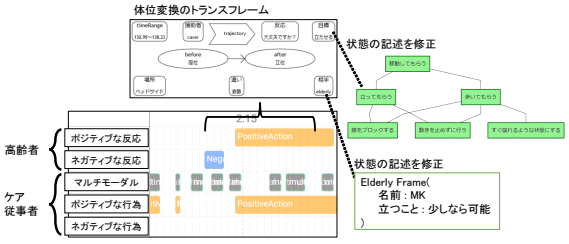


図 3: ケア熟達者の意見集約による熟達者のケアプロセスの表現

思考モデルによって，熟達者の意見が集約され，認知症ケア知の共創につながることが示された．

4. おわりに

本稿では，感情思考モデルを機軸とした共創システムによって，ドメインエキスパートの意見が集約され，認知症ケア知の再構造化が促進されることを示した．今後も思考の表現を深化させるために Interior grounding[Minsky 09] を進め，一見単純にみえる現象が複雑であることを示すための情動理解基盤技術の開発を進める．

参考文献

[本田 17] 本田，他：ユマニチュードの Evidence-Based Care の実現に向けたマルチモーダルコミュニケーションの評価，第 31 回人工知能学会全国大会，1I3-NFC-02b-2in1 (2017).

[石川 15] 石川，竹林: エビデンスを生み出す認知症情報学－情動理解基盤技術とコミュニケーション支援－，人工知能学会誌，32(1), pp.103-110 (2017).

[Minsky 09] Minsky, M. 著, 竹林訳：ミンスキー博士の脳の探検－常識・感情・自己とは－, 共立出版, 東京 (2009).

[竹林 15] 竹林：認知症の人の暮らしをアシストする人工知能技術，人工知能学会誌, 29(5), pp.515-523 (2014).

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

骨格検出とベッド位置検出の組合せによる患者の端座位推定に関する初期検討

Patient's sitting position estimation using skeleton information and bed position detection

井上　円[*1,2]
Madoka Inoue

田口　亮[*2]
Ryo Taguchi

梅崎　太造[*2,3]
Taizo Umezaki

[*1] アイホン株式会社
AIPHONE CO., LTD.

[*2] 名古屋工業大学
Nagoya Institute of Technology

[*3] 東京大学
The University of Tokyo

The work of a nurse in charge of nursing of a hospitalized patient is a task that leads to a serious accident with a single mistake or overlook, so nurses are exposed to high stress. In particular, the three items of medication, tube connection, and fall are factors that lead to serious accidents and are considered to be a major load factor of nursing work. In order to reduce the burden of nursing work, we estimated the sitting posture, which is the initial posture of the patient's lifting motion, using the camera image. In this study, as an initial study of the study, we estimated the patient's sitting position by combining the detection result of the skeleton position of the patient, the detection result of the bed position and the integrated processing of them. As a result of the evaluation, the end sitting position was identified with an accuracy of 98.6 [%].

1. はじめに

入院患者の看護を担当する看護師の業務は，ひとつのミスや見逃しが重大な事故に繋がる業務であり，看護師は高いストレスに晒されている．とりわけ，注射・予薬，チューブ接続，転倒転落の 3 項目は重大な事故に繋がる要因であり，看護業務の大きな負荷要因とされている．

これまで我々は，看護業務の負荷を低減する為に，画像センシング技術を用いて，事故原因の主要 3 項目の一つである転倒転落に対する未然防止を検討してきた．その中で，転倒事故は就寝中の患者がトイレに行く過程で多発するという理解を得た．そこ本研究は，事故防止の手段として，患者の離床動作における特定の姿勢である端座位を推定する．なお，用語の定義として，患者の就寝姿勢を臥位（がい）とし，患者が足を地面に着けてベッド端に座る姿勢を端座位（たんざい）とする．

本稿で用いる情報は全て病室に設置した単眼カメラデバイスから得られた可視画像の情報とする．近年デプスセンサが普及しており，空間内の情報を捕捉する有効な手段とされている．その様な背景において，我々がカメラデバイスを用いる理由は二つある．理由の一つは，入院患者は音や光に過敏に反応する為，デプスセンサを導入する事で内蔵の近赤外光が患者のストレス源となる事を避ける為である．もう一つの理由は，導入コストである．ステレオカメラを用いた場合，デプスセンサと等価の空間情報を捕捉する事が出来るが，製品にカメラモジュールを 2 個搭載する必要があり製品単価が上昇する．国内の大規模病院は 500 床を超える為，デバイス単価の上昇はシステム全体のコストに影響する．計算処理を工夫し単眼カメラで所望の機能を実現する事で，システム全体のコストダウンに寄与する．

本項では，研究の初期検討として，患者の骨格位置の検出結果と，ベッド位置の検出結果を組み合わせ，患者の端座位を推定する．

連絡先：井上　円，アイホン株式会社　技術本部　技術開発部　第二開発課，〒456-8666 名古屋市熱田区神野町 2-18，madoka_inoue@aiphone.co.jp

2. 提案手法

2.1 骨格検出

骨格検出には，Cao らの二次元画像の姿勢推定手法[Cao 2017]を用いる．深層学習の各チャネルで部位毎の検出を行い，部位の関連度を算出する．更に各部位と部位の関連度に基づいて，部位同士をつなぎ合わせる事で骨格を抽出する．

2.2 ベッド位置検出

可視画像からベッド位置を検出する代表的な手法は，線分をヒューリスティックに組み合わせる手法[Kittipanya-Ngam 2012] [岡田 2015] や画像特徴と機械学習を組み合わせた手法 [井上 2018] が存在する．ただし，ヒューリスティックな手法は，製品の長期供給を考慮した場合に，製品の設計変更が生じた際にパラメータの再検討が必要となるリスクを含む．従って，本稿では



Fig. 1. Bed shapes of input images

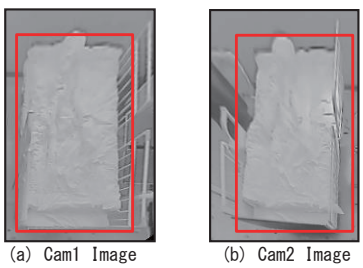


Fig. 2. Bed shapes of IMP images

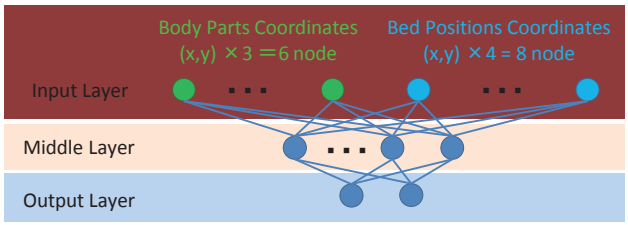


Fig. 3. Overview of network

画像特徴と機械学習による手法を用いる.

提案手法は, 入力画像をベッドサイズ程度の着目領域で走査する. 着目領域毎の処理は, 領域内の輝度情報を用いて HOG 特徴を抽出し, SVM 識別器を用いて HOG 特徴に対する尤度判定を行う. 尤度判定の結果が事前に設定した閾値以上であれば着目領域をベッドと判定し, 着目領域の位置からベッド座標を算出する. ただし, 画像に記録されるベッド形状は, 透視投影変換の原理上, 図 1 に示す様にカメラとベッドの位置関係に応じて様々な四角形となる. この様にカメラ位置に応じて形状のばらつきが生じるベッド画像を HOG 特徴で抽出した場合, 画像特徴の出力値が発散する為にベッドを正しく検出する事が出来ない. そこで, 特徴抽出の前処理として入力画像に逆射影変換 [Kim 2002] 処理を施す. 予め入力画像に対してレンズ歪を補正する処理を行い, 歪補整後の画像に対して視点変換を用いて視点を室内の鉛直上方向に置き換える鳥瞰図を生成して用いる事で, 設置エリア内のどの位置に取り付けられたカメラを使用しても, 図 2 に示す様な長方形のベッド形状を得る.

2.3 端座位推定

初期検討として, 3 層ニューラルネットワークを用いて端座位の推定を行う. 入力値は, 骨格の検出結果座標と, ベッド位置の検出結果の座標情報用いる. Cao らの骨格推定では部位情報として体の 18 点の座標が得られるが, 本手法では座位の推定に必要な首・右骨盤・左骨盤の 3 点の座標情報を用いる. また, ベッド位置はベッド検出処理により得られたベッド隅 4 点の座標を入力とする. 出力層はノード数 2 の 2 クラス分類問題として解く. ネットワークの概略を図 3 に示す.

3. 実験

3.1 実験データ

病院環境を模擬したシミュレーション環境を構築し, ベッド頭部壁面の 8 か所に設置したカメラを用いて, 実験協力者 5 名の臥位から端坐位に遷移し, 室外に退出する一連の動作を 5[fps]で 3888 データ収集した. 本実験では, 撮影データから端坐位シーンの静止画を 7832 枚, 端坐位以外の姿勢を 7900 枚使用する. なお, 実験協力者に対しては実験の目的やデータの用途, 作業の内容, ならびに個人情報の取り扱いについて十分な説明を行い, 同意書を取り交わした.

3.2 端座位の推定精度評価

学習データに対して, 骨格とベッド位置の検出を行う. 一例として骨格とベッド位置の検出結果サンプルを図 4 に示す. 次に, 各々の結果座標を取得し, 両者の座標を画像毎に統合した実験用入力座標データを作成する. 続いて, 学習課程として実験用入力座標データの 90[%]を用いてネットワークのパラメータを学習する. 最後に, 得られたネットワークパラメータについて, 実験用入力座標データの残り 10[%]で推論を行い, 識別精度を調べる.

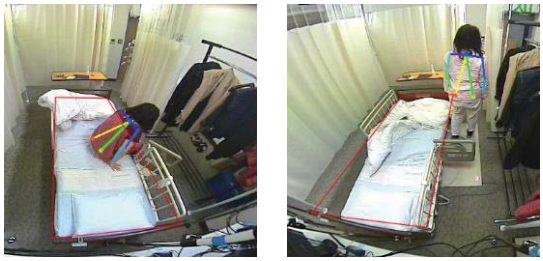


Fig. 4. Results of Pose and Bed Detection

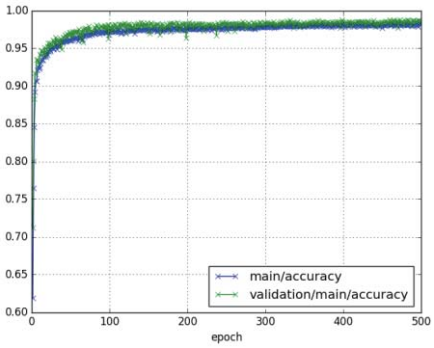


Fig. 5. Learning Curve

なお実験条件として各層のノード数は, 入力層を 14, 中間層を 10, 出力層を2と設定した. 学習繰り返し回数は 500 とした.

3.3 実験結果

評価データに対して 98.6[%]の精度で端坐位とその他の動作を正しく識別した. 実験の学習曲線を図 5 に占めす.

4. まとめ

安価なデバイスを用いた看護業務の負荷低減策として, カメラ映像と機械学習を組み合わせて患者の端座位を識別する手法を提案した. 骨格とベッド位置の検出結果に統合処理を加える事で, 患者の端座位とその他の動作を 98.6[%]の精度で識別した. 今後, 更なる精度向上と, 一連の処理を統合したアルゴリズムを検討したい.

参考文献

[Cao 2017] Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y.: Realtime multi-person 2d pose estimation using part affinity fields, In *IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

[Kittipanya-Ngam 2012] Kittipanya-Ngam, P., Guat, O. S., and Lung, E. H.: Bed detection for monitoring system in hospital wards, In *IEEE Proc.34th Annual Conf. Engineering in Medicine and Biology Society*, 2012.

[岡田 2015] 岡田康貴, 村下君高: 医療現場での運用に即したベッド領域認識技術の開発, 精密工学会誌 Vol.81 No. 12, 2015.

[井上 2018] 井上 円, 安井 俊之, 田口 亮, 梅崎 太造: 患者みまもりの自働化に向けたベッド位置の検出, 電気論文 C Vol.138 No.6, 2018.

[Kim 2002] Kim, D. K., Jang, B. T., & Hwang, C. J.: A planar perspective image matching using point correspondences and rectangle-to-quadrilateral mapping, In *Image Analysis and Interpretation, 2002. Proceedings. Fifth IEEE Southwest Symposium on*, 2002.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

Accurate drowsiness estimation via eye-related movements: a neural-network-based investigation

Mingfei Sun*1 Masanori Tsujikawa*2 Yoshifumi Onishi*2 Xiaojuan Ma*1
Atsushi Nishino*3 Satoshi Hashimoto*3

*1 Department of Computer Science & Engineering, Hong Kong University of Science and Technology
*2 Data Science Research Laboratories, NEC Corporation
*3 Technology and Innovation Center, DAIKIN Industries, LTD

Many studies reported eye-related movements, e.g., eye blink and eyelid drooping, are highly indicative symptoms of drowsiness. However, few has investigated the computational efficacy for drowsiness estimation accounted by these movements. This paper thus analyzes two typical movements: eyelid movements $X_{el}(t)$ and eyeball movements $X_{eb}(t)$, and investigates different neural-network modelings: CNN-Net and CNN-LSTM-Net. Experimental results show that using joint movements can achieve better performances than eyelid movements for short time drowsiness estimation while using eyeball movements alone perform worse even than the baseline (PERCLOS method). In addition, the CNN-Net is more effective for accurate drowsiness level estimation than the CNN-LSTM-Net.

1. INTRODUCTION

Drowsiness has been studied for years in fatigue risk management and vigilance monitoring. Psychological studies have demonstrated that the drowsiness can significantly impair productivity and deteriorate the quality of work outcomes. For example, drowsy driving, usually caused by sleep loss, night or long-time working [Dinges and Grace, 1998], is reported as one of the main causes of serious accidents [Akbar et al., 2017, Reddy et al., 2017]. On the another hand, if drowsiness can be effectively estimated, the aforementioned side effects might be significantly mitigated or avoided. Studies showed that accurate driver drowsiness estimation can help prevent potential accidents caused by driving fatigue [Dinges and Grace, 1998]. Also, research on Massive Open Online Courses (MOOC) demonstrated that the learning outcomes can be greatly improved by taking measures based on students' estimated level of drowsiness [Kim et al., 2017].

Eye-related movements, e.g., eyelid movements and eyeball movements, are reported to be important and useful for accurate drowsiness estimation by previous studies [Zhu et al., 2014, Wierwille et al., 1994, Dinges and Grace, 1998]. However, few has tried to further differentiate the role played by different types of eye movements. Basically, there are two types of eye-related movements: eyelid movement $X_{el}(t)$ and eyeball movement $X_{eb}(t)$. Technical investigation of these two types of eye movements is non-trivial since it requires accurate extractions of eyelid movements and eyeball movements. Furthermore, although many neural-network-based temporal modelings, e.g., CNN and LSTM, are proposed for modeling these eye-related movements, an insightful comparison is still in need.

Contact: Mingfei Sun, CSE Department at HKUST, Hong Kong, (+852) 5224 9330, mingfei.sun@ust.hk

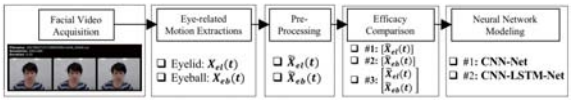


Figure 1: The proposed investigation pipeline

In this paper, we conduct a systematic analysis of two types of eye-related movements for drowsiness estimation. The contributions of this paper are as follows. First, we analyze the computational efficacy of eyelid movements and eyeball movements for drowsiness estimation. Second, we propose and compare two neural networks for modeling temporal correlations of eye-related movements. Third, we conduct multiple experiments and present insightful interpretations as well as discussions for experimental results. Empirical evaluation shows that the proposed neural network models are effective, and the combined eye-related movements are more effective than eyelid movements for short time drowsiness estimation, while eyeball movements alone fail to deliver good results. In addition, as drowsiness period extends, CNN-Net performs increasingly better than CNN-LSTM-Net for drowsiness level prediction.

2. METHOD

The overview of our proposed method is shown in Fig. 1.

2.1 Facial Video Acquisition

The facial video data were collected from 29 subjects with their informed consent in a local laboratory and then labeled with 3 professional annotators. First, 29 subjects were asked to finish a specific math task while seated in front of a display with a web camera placed on top. The facial videos ($640 * 480$, 30fps) were captured by a web camera when subjects were performing the task, which finally adds up to $205,600$s ($\sim$ 57h). Then, 3 professional annotators rated every 10 seconds (without overlapping) of facial videos with

3Pin1-49

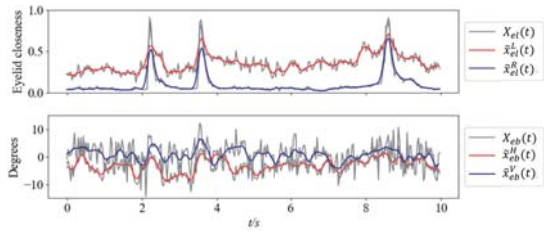


Figure 2: Filtering & down-sampling of eyelid (top) and eyeball (bottom) movements

5 drowsiness levels according [Zilberg et al., 2007]. The averaged label values is used as the ground truth.

2.2 Motion Extraction & Pre-processing

Eye-related motion extraction: The eyelid movements $X_{el}(t)$ and eyeball movements $X_{eb}(t)$ are extracted by an accurate facial analysis tool developed by NEC corporation. Specifically, $X_{el}(t) = \left[x_{el}^{L}(t), x_{el}^{R}(t)\right]$ where $x_{el}^{L}(t)$ and $x_{el}^{R}(t) \in [0, 1]$ represent left and right eyelid closenesses with 0 denoting fully open and 1 fully closed. Meanwhile,$X_{eb}(t) = \left[x_{eb}^{H}(t), x_{eb}^{V}(t)\right]$ where $x_{eb}^{H}(t)$ and $x_{eb}^{V}(t) \in [-30, 30]$ denote vertical and horizontal degrees of eyeball movements.

Pre-processing: The Hamming filter (window size 11) is applied separately to $X_{el}(t)$ and $X_{eb}(t)$ to remove high frequency noises, as shown in Fig. 2 where $X'_{el}(t) = \left[x'^{L}_{el}(t), x'^{R}_{el}(t)\right]$, $X'_{eb}(t) = \left[x'^{L}_{eb}(t), x'^{R}_{eb}(t)\right]$ and $x'(t) = H\big(x(t)\big)$. The filtered $X'_{el}(t)$ and $X'_{eb}(t)$ are then down-sampled to 15 points per second, i.e., $X'_{el_d}(t) = X'_{el}(t)$ and $X'_{eb_d}(t_d) = X'_{eb}(t)$, in order to remove redundancies. In addition, according to the Rectified Linear Unit (ReLU) used in neural networks, the input data are further shifted and scaled to be non-negative based on the degree ranges, i.e., $\tilde{x}_{eb}(t) = \left[x'_{eb_d}(t) + 30\right]/60$, while $x'^{L}_{el_d}(t)$ are kept $x'^{R}_{el_d}(t)$ the same, $\tilde{x}_{el}(t) = x'_{el_d}(t)$.

2.3 Efficacy Comparison & Neural Network Modeling

To compare the efficacy of different eye-related movements for drowsiness estimation, we apply a same neural network model, i.e., CNN-LSTM-Net as shown in Fig 3(a), to three different combinations of eyelid and eyeball movements: $\left[\tilde{X}_{el}(t)\right]$, $\left[\tilde{X}_{eb}(t)\right]$ and $\left[\tilde{X}_{el}(t), \tilde{X}_{eb}(t)\right]$. Then three combinations of eye-related movements are input as features, i.e., $\left[\tilde{X}_{el}(t)\right]$, $\left[\tilde{X}_{eb}(t)\right]$ and $\left[\tilde{X}_{el}(t), \tilde{X}_{eb}(t)\right]$. For joint movement feature $\left[\tilde{X}_{el}(t), \tilde{X}_{eb}(t)\right]$, the $\tilde{X}_{el}(t)$ and $\tilde{X}_{eb}(t)$ are processed separately by CNN part before being concatenated into LSTM part.

Another neural network structure, i.e., CNN-Net as shown in Fig 3(b), can also be applied to estimate drowsiness from eye-related movements [Zhu et al., 2014]. Different from CNN-LSTM-Net, CNN-Net structure only exploits convolution layers without LSTM layers. In the proposed CNN-Net, we use a similar CNN structure as that in CNN-LSTM-Net. Instead of the LSTM layer, we adopt a global pooling layer to compare every sequence segment to obtain the max value, which are then fully connected to output the drowsiness labels.

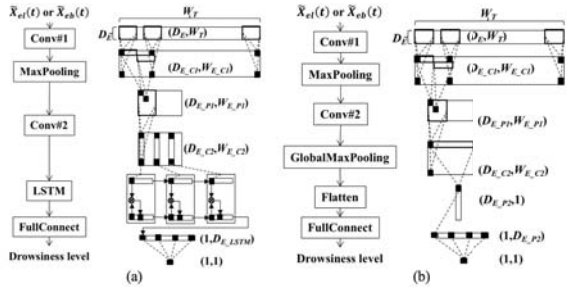


Figure 3: (a) CNN-LSTM-Net; (b) CNN-Net

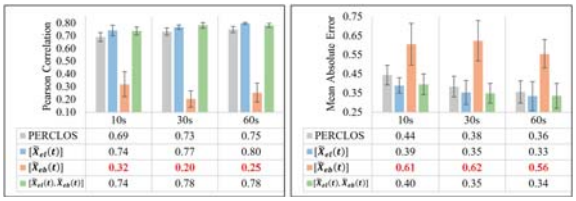


Figure 4: Modeling eyelid movements and eyeball movements by CNN-LSTM-Net along different time periods

3. EXPERIMENT RESULTS

The experiments are conducted on the facial video dataset with two common regression performance metrics: Pearson Correlation coefficient and Mean Absolute Error (MAE). Specifically, the evaluation dataset is randomly divided into five groups based on participants to remove the subject-dependences, and the cross validation adopts leave-one-group-out strategy with the PERCLOS method [Dinges and Grace, 1998] as the baseline. The training is conducted on the Nvidia 1080 (8GB memory) graphic card (batch size 128, epoch 100), and an early stop checker was also set to reduce the training time accordingly. The Adam optimizer ($lr = 0.001$) was used with learning rate decaying to 0.0001 and 0.00001 after 20 and 50 epochs respectively.

3.1 Efficacy comparison

The importance is compared between $\left[\tilde{X}_{el}(t)\right]$, $\left[\tilde{X}_{eb}(t)\right]$ and $\left[\tilde{X}_{el}(t), \tilde{X}_{eb}(t)\right]$. From Fig. 4, it is notable that using $\left[\tilde{X}_{eb}(t)\right]$ to estimate drowsiness fails to deliver good results, worse even than the baseline (PERCLOS) performances: the Pearson Correlation values < 0.4 indicates the estimated drowsiness levels and the ground truth labels are weakly correlated; the Mean Absolute Error(MAE) values are all around 0.6, which are almost two times of the MAEs achieved by PERCLOS method. In the meanwhile, $\left[\tilde{X}_{el}(t)\right]$ and $\left[\tilde{X}_{el}(t), \tilde{X}_{eb}(t)\right]$ both perform almost equally better than the PERCLOS. As a summary, the eyeball movements alone are not an effective feature for drowsiness estimation while eyelid movements or eyelid movements combined with eyeball movements are strongly related to drowsiness status.

The underlying reasons that eyeball movements alone fail

3Pin1-49

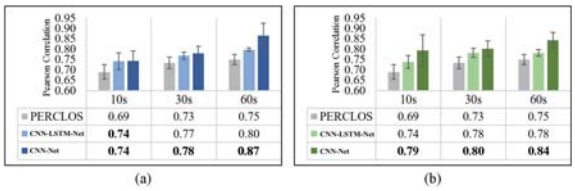


Figure 5: CNN-LSTM-Net vs. CNN-Net for drowsiness estimation using (a) eyelid movements; (b)joint movements

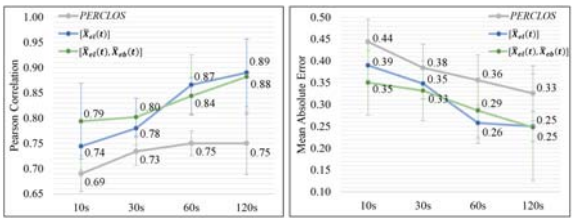


Figure 6: CNN-Net modeling of eyelid and joint movements

to capture the drowsiness levels may possibly be due to the following phenomenon: when a person feels sleepy, his/her eye gazes will not move too much, usually fixed on a point, which, from the eyeball movement's perspective, is exactly the same as the person is fixating his/her eye gazes when he/she is awake. In the following experiments, the eyeball movements will not be considered again.

3.2 Neural network modeling

The eyelid movement data and the joint movement data are input separately into the CNN-LSTM-Net and CNN-Net. From the results, as shown in Fig 5, the CNN-Net can deliver better performances than CNN-LSTM-Net. For eyelid movements, CNN-Net produces better results as drowsiness video time increases, while for the joint movements, CNN-Net consistently beats CNN-LSTM-Net for multiple video lengths. Thus, for drowsiness estimation, CNN-Net modeling is more preferable.

We proposed one possible interpretation of why CNN-Net beats CNN-LSTM-Net. For CNN-Net structure, it is naturally good at spotting short-time correlations due to its fixed-sized (usually small) filters. For CNN-LSTM-Net, however, it mainly captures the long-term temporal correlations among eye-related movements since the LSTM structure is designed to automatically "remembering" information flows along many time steps. In drowsiness estimation, most of the drowsiness levels are actually showed by instant body signals, e.g., a quick eye drooping, which possibly indicates that the short-term eye-related movements may be more important than the long-term eye movements. Consequently, the CNN-Net modeling (short-term) performs better than CNN-LSTM-Net modeling (long-term).

Furthermore, we compare the eyelid and joint movements along different time periods by CNN-Net modeling, as shown in Fig 6. The results show that the joint movements $[\tilde{X}_{el}, \tilde{X}_{eb}]$ deliver slightly better results for short time drowsiness estimation while eyelid movements alone increasingly achieves comparable or even better results as time increases.

4. CONCLUSION

To conclude, two very typical eye-related movements, eyelid movements $\tilde{X}_{el}$ and eyeball movements $\tilde{X}_{eb}$, are analyzed for drowsiness estimation. Also, two very typical neural network structure, i.e., CNN-Net and CNN-LSTM-Net, are investigated for eye-related movement modelings. The empirical evaluation shows that the eyelid movements alone $[\tilde{X}_{el}]$ as well as the joint movements $[\tilde{X}_{el}, \tilde{X}_{eb}]$ are effective for accurate drowsiness estimation while eyeball movements alone cannot deliver good performance. Moreover, the experimental results demonstrate that the CNN-Net is more effective than the CNN-LSTM-Net for eye-related movement modeling due to the possible short-term temporal correlations. In addition, we also find that the joint movements can be more useful for short time drowsiness estimation while eyelid movements alone can also produce comparable performance for long time drowsiness estimation.

References

[Akbar et al., 2017] Akbar, I. A., Rumagit, A. M., Utsunomiya, M., Morie, T., and Igasaki, T. (2017). Three drowsiness categories assessment by electroencephalogram in driving simulator environment. In *EMBC 2017, Conference of the IEEE*, pages 2904–2907. IEEE.

[Dinges and Grace, 1998] Dinges, D. F. and Grace, R. (1998). Perclos: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance. *Federal Highway Administration*.

[Kim et al., 2017] Kim, H. Y., Kim, B., Lee, J., and Kim, J. (2017). Hey, wake up: Come along with the artificial learning companion to the e-learner's outcomes high! In *Proceedings of CHI 2017*, pages 1763–1770. ACM.

[Reddy et al., 2017] Reddy, B., Kim, Y.-H., Yun, S., Seo, C., and Jang, J. (2017). Real-time driver drowsiness detection for embedded system using model compression of deep neural networks. In *CVPRW, 2017 IEEE Conference on*, pages 438–445. IEEE.

[Wierwille et al., 1994] Wierwille, W. W., Wreggit, S., Kirn, C., Ellsworth, L., and Fairbanks, R. (1994). Research on vehicle-based driver status/performance monitoring. Technical report.

[Zhu et al., 2014] Zhu, X., Zheng, W.-L., Lu, B.-L., Chen, X., Chen, S., and Wang, C. (2014). Eog-based drowsiness detection using convolutional neural networks. In *IJCNN*, pages 128–134.

[Zilberg et al., 2007] Zilberg, E., Xu, Z. M., Burton, D., Karrar, M., and Lal, S. (2007). Methodology and initial analysis results for development of non-invasive and hybrid driver drowsiness detection systems. In *AusWireless 2007. The 2nd International Conference on*, pages 16–16. IEEE.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

アセスメントデータを利用した要介護度の推定

Estimating nursing care necessity degree using by assessment information

小林　秀[*1]　Shu Kobayashi
平田 佐智子[*1]　Sachiko Hirata-Mogi
福田升二[*1]　Shoji Fukuda

[*1] 株式会社エス・エム・エス
SMS Co. ,Ltd.

Expense on elderly care is increasing more rapidly than other expense on social security benefits. To Operate long-term care insurance system, improving operating efficiency is the critical issue. Especially, we focused on assessment of nursing care necessity degree that is the first step of long-term care insurance system. Assessment productivity will improve by assessing nursing care necessity degree periodically, objectively and automatically as possible. In this paper, we created nursing care necessity degree classify model using by assessment data for planning care-plan, and analyzed mis-classify trends.We found assessment data for planning care-plan are also useful information on assessing nursing care necessity degree, and there is possibility of improving assessment productivity.

1. はじめに

本稿では要介護認定の自動化を目標として，介護アセスメントデータを利用して要介護度の推定モデルの構築を行い，誤判別の傾向を分析のうえ，その実施上の課題を検討する．

近年，社会保障給付費は増加の一途を辿っており，[財務省 2015]の資料によれば，2000 年には 78. 1 兆円であったものが，2012 年には 109. 5 兆円となり，2025 年には 148. 9 兆円に達すると見込まれている．2000 年時点では医療と年金の総額が 67. 2 兆円に対して，介護は 3. 3 兆円であった．これが，2012 年時点では 8. 4 兆円となり，2025 年には 19. 8 兆円に達すると見込まれている．このうち，2012 年から 2025 年にかけては，介護に関する給付費の増加率は 2. 3 倍となっており医療の 1. 5 倍に比べても顕著な増加が見込まれている．

介護保険制度は 1997(平成 9)年 12 月に「介護保険法」が制定され，2000(平成 12)年 4 月から施行されている．介護保険は 40 歳になった月から全ての人が加入することになり，介護保険料の支払い義務が発生する．65 歳以上は「第 1 号被保険者」，40 歳から 64 歳までは「第 2 号被保険者」にあたる．このうち，第 1 号被保険者は，介護が必要であると認定を受けると，その程度によって日常生活の支援者や介護のサポート(以下介護サービス)を受ける際に介護給付を受けることができる．

介護給付を受けるためには，要介護度認定と呼ばれる介護サービスの利用者がどの程度介護サービスを必要としているかの目安となる基準の審査を受ける必要がある．審査のためには申込みを行い，全国共通の認定調査書を使った訪問調査が行われ，コンピュータによる機械的な 1 次判定が行われる．[厚生労働省 2012]認定調査員テキスト 2009 改訂版には調査項目が掲載されており，この一部を図 1 に示す．訪問調査では 74 項目について評価を行い，さらに特記事項について調査を行う．その後，機械的なアルゴリズムにより得られた 1 次判定の結果をもとに，保険医療・福祉の専門家で構成される「介護認定審査会」が 2 次判定を行い，判定結果を市区町村に通知する．認定を受けると各市区町村の指定を受けた事業者の介護支援専門家(ケアマネージャ)が介護サービス計画書(ケアプラン)を作成し，それにしたがってサービスが実施される．


認定調査票（基本調査）

1-1 麻痺等の有無について、あてはまる番号すべてに○印をつけてください。（複数回答可）
1. ない　2. 左上肢　3. 右上肢　4. 左下肢　5. 右下肢　6. その他（四肢の欠損）

1-2 拘縮の有無について、あてはまる番号すべてに○印をつけてください。（複数回答可）
1. ない　2. 肩関節　3. 股関節　4. 膝関節　5. その他(四肢の欠損)

1-3 寝返りについて、あてはまる番号に一つだけ○印をつけてください。
1. つかまらないでできる　2. 何かにつかまればできる　3. できない

1-4 起き上がりについて、あてはまる番号に一つだけ○印をつけてください。
1. つかまらないでできる　2. 何かにつかまればできる　3. できない

1-5 座位保持について、あてはまる番号に一つだけ○印をつけてください。
1. できる　2. 自分の手で支えればできる　3. 支えてもらえればできる　4. できない

1-6 両足での立位保持について、あてはまる番号に一つだけ○印をつけてください。
1. 支えなしでできる　2. 何か支えがあればできる　3. できない

1-7 歩行について、あてはまる番号に一つだけ○印をつけてください。
1. つかまらないでできる　2. 何かにつかまればできる　3. できない

1-8 立ち上がりについて、あてはまる番号に一つだけ○印をつけてください。
1. つかまらないでできる　2. 何かにつかまればできる　3. できない

1-9 片足での立位保持について、あてはまる番号に一つだけ○印をつけてください。
1. 支えなしでできる　2. 何か支えがあればできる　3. できない

1-10 洗身について、あてはまる番号に一つだけ○印をつけてください。
1. 介助されていない　2. 一部介助　3. 全介助　4. 行っていない

図 1. 認定調査表（一部）

要介護度には，要介護への進行を予防する「要支援」と介護を必要とする状態である「要介護」に分かれており，「要支援」はさらに 2 段階．「要介護」は 5 段階に分かれている．要介護度によって給付限度額が分かれており，「要支援 1」の月額 5 万 30 円から「要介護 5」では 36 万 650 円と大きく金額差がある．要介護度は，初回の認定では一般的な場合には 6 ヶ月間が有効期間であり，以降の更新後は 12 ヶ月が有効期間となっている．

コンピュータによる 1 次判定は機械的に導出されるものの，保健・医療・福祉の学識経験者により構成される介護認定審査会により，一次判定結果，主治医意見書等に基づき審査判定（二次判定）を行う．[厚生労働省 2012] 介護認定審査会委員テキストには 1 次判定の具体的な流れや，1 次判定結果から変更さ

連絡先：小林　秀，株式会社エス・エム・エス，
shu-kobayashi@bm-sms.co.jp
** 本研究の内容は株式会社エス・エム・エスの公式見解を示すものではありません

れやすい認定調査票上の調査結果について記載があり，機械的に実施可能な要因や現在の調査項目以外の要因によっても 1 次判定結果が修正される可能性がある．

そこで，本研究ではまず，ケアプラン作成のために一定頻度で実施される 1 次判定に使用される項目以外の評価項目を含んだアセスメントの結果と要介護度の関係を比較することでその対応関係が実態としてどのようになっているかを把握する．そして，予測と実態の乖離要因を比較することで，アセスメントデータによる要介護度判定の代替可能性の検討を行った．この結果を用いることで，定期的な要介護度の見直しや 2 次審査での要介護度の補正が発生しやすいと考えられる利用者の事前抽出が実施可能となり，要介護度判定業務の業務負担低減が可能になるものと考えられる．

2. 従来研究

一般に，要介護度に大きな影響を与える要因としてサービス利用者の ADL(Activity of daily life)に関する具合が挙げられる．要介護度の推定にあたり，ADL 調査項目がどのように影響を与えているか[川越 2017]により，医学的検知からの研究が見られる．

現在の介護保険制度上の要介護度の 1 次判定に関して、その判定ロジックの妥当性については，[関 2000]らによって検証がなされている．一方で 2 次判定に関連するものとして，疾患と要介護度の関連については，例えば介護保険の新規認定者において，要介護度が重度となるような原因疾病の特定について，[高橋 2017]らにより，主治医意見書に記載された疾患名のうち，脳血管疾患，悪性新生物が要介護度 2 をこえるような重度の要介護度に関連することが述べられている．

以上のような研究は見られるものの，情報学的な観点から大量のデータを利用して，要介護認定に使用される調査項目以外も対象とした分析事例や，なかでもアセスメントの結果と要介護度の関連に関して，何らかの分析を行った事例は見られなかった．

3. 分析のアプローチ

3.1 分析データの概要

本稿では，エス・エム・エスが運営する介護事業者向けサービスである「カイポケ」のデータを利用した．このデータ概要を表 1 に示す．

表 1. 分析データの概要

対象データ	2011 年以降のカイポケ上でのアセスメント機能使用対象者
分析データ	同一サービス利用者に対する複数回のアセスメントを含む延べ 963，409 件
アセスメント方式	全社協方式準拠 必須絞込版（帳票 3 ページ） 全項目版（全社協方式準拠/帳票 11 ページ） ※いずれも厚労省（老企第 29 号）の提示する課題分析標準項目は網羅

3.2 分析アプローチの全体像

本稿で実施するアプローチでは大きく 4 つのステップが存在する．この概要を図 2 に示す．分析にあたり，データ処理やモデル構築には AmazonWebService 上に環境を準備のうえ，統計処理には R を利用した．

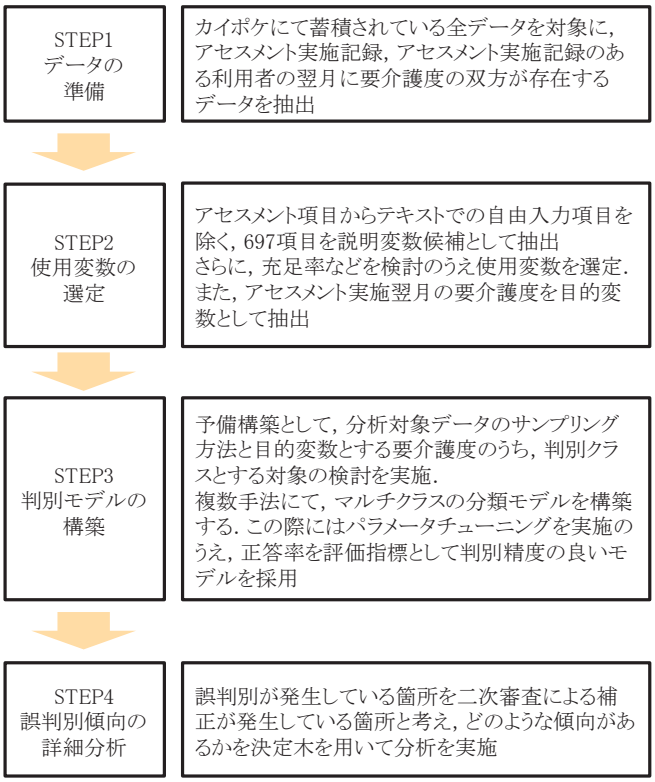


図 2. 本稿での分析アプローチ

4. 使用変数の選定

4.1 アセスメントデータの概要

本稿で説明変数として用いるアセスメントは，要介護度の認定後に具体的にどのような介護サービスを実施していくかを計画されたケアプランを作成する目的で実施される．アセスメントには様々な方式があるものの，本稿で用いるアセスメントシートは全社協方式準拠となっており，要介護度認定に使用される項目も含まれている．アセスメント項目の入力画面例を図 3 に示す．

4.2 変数

本稿で使用するアセスメント票には，認定調査員テキスト 2009 改訂版に含まれる 74 項目が含まれている．したがって，項目上は一次判定における調査に相当する判定が実施可能である．しかしながら，選択肢には当該項目についてアセスメントを行わなかったことを示す“未選択”が存在し，完全に再現ができるものではない．

説明変数はアセスメント項目のうち，テキストでの自由入力項目を除いたうえで，さらに入力率が 100%となっている 603 項目を使用した．また，目的変数となる要介護度はアセスメント翌月時点の値を用いた．



図 3. アセスメントシート入力画面例

5. 要介護度判別モデルの構築

5.1 要介護度判定モデルの構築

要介護度判定モデルの構築にあたり，予備構築として，分析対象の母集団を全件から構築:検証を 7:3 に分割した場合と，構築と検証を各要介護度ごとに構築と検証にそれぞれ 1 万件ずつランダムサンプリングした場合と，さらに，目的変数が要介護度 1 から 5 に加えて要支援 1 から 2 を含めた 7 段階，要介護度のみ 1 から 5 の 5 段階とした場合の 4 パターンを実施した．

この結果，要支援を含めた場合や，全件のデータを使用した場合には要介護度ごとのサンプル数の違いに起因すると考えられるバイアスが見られた．したがって，目的変数を要介護度 1~5 として，各要介護度から構築と検証で重複はせずに構築に 5 万件ずつ，検証用に 1 万件のランダムサンプリングを行い本分析用データとした．

本分析でのモデリングには RandomForest， XGBoost の 2 手法を用いて，パラメータチューニングを実施のうえ，正答率と構築と検証データでの正答率の差異を考慮し最適なモデルの選定を行った．

5.2 判別精度

要介護度 1-5 を対象にランダムサンプリングを実施のうえ，件数を揃えた上で，モデル構築を行った．モデルから得られた予測クラスと正解のクラスの集計結果を図 4 に示す．

正答率は構築データで 50.8%，検証データで 59.7%と非常に安定した結果が得られた．検証データにおいて不正解のうち，正解が予測より低いものが 21.0%に対して，正解が予測より高いものが 28.3%と実態は高めに要介護度が設定されていることが示唆された．

また，1 段階の推定差異を許容する場合，正答率は検証データで 87.4%となり，2 段階以上高い先は 9.0%に対して，2 段階以上低い先は 3.7%と推定よりも実際は高めの要介護度となっている場合が多く見られる傾向が伺える．

構築:件数		正解					
		要介護1	要介護2	要介護3	要介護4	要介護5	合計
予測	要介護1	34,022	18,205	7,207	3,753	2,057	65,244
	要介護2	12,599	20,706	12,552	5,210	2,142	53,209
	要介護3	2,371	8,068	17,488	8,610	2,733	39,270
	要介護4	739	2,328	9,909	19,672	7,973	40,621
	要介護5	269	693	2,844	12,755	35,095	51,656
	合計	50,000	50,000	50,000	50,000	50,000	250,000

構築:構成比		正解					
		要介護1	要介護2	要介護3	要介護4	要介護5	合計
予測	要介護1	52.1%	27.9%	11.0%	5.8%	3.2%	26.1%
	要介護2	23.7%	38.9%	23.6%	9.8%	4.0%	21.3%
	要介護3	6.0%	20.5%	44.5%	21.9%	7.0%	15.7%
	要介護4	1.8%	5.7%	24.4%	48.4%	19.6%	16.2%
	要介護5	0.5%	1.3%	5.5%	24.7%	67.9%	20.7%
		20%	20%	20%	20%	20%	100.0%

検証:件数		正解					
		要介護1	要介護2	要介護3	要介護4	要介護5	合計
予測	要介護1	6,892	3,665	1,360	744	410	13,071
	要介護2	2,455	4,132	2,523	1,019	405	10,534
	要介護3	470	1,588	3,486	1,783	545	7,872
	要介護4	136	460	2,069	3,894	1,694	8,253
	要介護5	47	155	562	2,560	6,946	10,270
	合計	10,000	10,000	10,000	10,000	10,000	50,000

検証:構成比		正解					
		要介護1	要介護2	要介護3	要介護4	要介護5	合計
予測	要介護1	52.7%	28.0%	10.4%	5.7%	3.1%	26.1%
	要介護2	23.3%	39.2%	24.0%	9.7%	3.8%	21.1%
	要介護3	6.0%	20.2%	44.3%	22.6%	6.9%	15.7%
	要介護4	1.6%	5.6%	25.1%	47.2%	20.5%	16.5%
	要介護5	0.5%	1.5%	5.5%	24.9%	67.6%	20.5%
		20%	20%	20%	20%	20%	100.0%

区分	構築		検証	
	件数	構成比	件数	構成比
正解が予測より低い箇所	52,575	21.0%	10,502	21.0%
正解が予測より高い箇所	70,442	28.2%	14,148	28.3%
実態と予測が整合	126,983	50.8%	25,350	50.7%
全体	250,000	100.0%	50,000	100.0%
実態と予測が上下1ランク以内	217,654	87.1%	43,687	87.4%
実態と予測が2ランク以上小さい	9,244	3.7%	1,830	3.7%
実態と予測が2ランク以上大きい	23,102	9.2%	4,483	9.0%

図 4 予測と正解の要介護度集計結果

5.3 誤判別先の傾向分析

構築した要介護度判定モデルから得られた予測クラスに対して，正解が高めに要介護度が設定される利用者の特徴を把握するために，要介護度を 1 と予測した 65,244 件を対象に，正解が 2 段階以上となった 13,017 件を判別対象として，決定木を用いてルール生成を行った．この結果中の LIFT が 2 以上となる特徴的なルールを表 2 に示す．寝たきり度が最重要変数となっており，未選択の場合の排尿介助現状区分が見守り必要もしくは介助必要となる場合に,正解が予測に対して 2 段階以上，高くなっている傾向が見られた．

表 2．予測要介護度 1 に関する乖離先判別ルール

ルール	該当件数	判別対象件数	判別対象件数比率
寝たきり度がランク B-2 以上	872	748	85．8%
寝たきり度　ランク B-1	678	370	54．6%
寝たきり度　未選択 かつ　移乗　未選択	2,518	1,445	57．4%
寝たきり度　未選択 かつ　移乗一部解除 or 全介助	431	800	53．9%
全体	65,244	13,017	19．9%

同様に要介護度を 2 と予測した 53,209 件を対象に，正解が 2 段階以上となった 7,352 件を判別対象として，決定木を用いてルール生成を行った結果中の LIFT が 2 以上となる特徴的なルールを表 3 に示す．寝たきり度が最重要変数となっており，未選択場合には排尿介助現状区分が見守り必要もしくは介助必要となる場合に 2 段階以上上方修正される傾向が見られる結果となった．

表 3．予測要介護度 2 に関する乖離先判別ルール

ルール	該当件数	判別対象件数	判別対象件数比率
寝たきり度　ランク C 以上	1，550	817	52．7%
寝たきり度　ランク B-2	6，653	1，883	28．3%
寝たきり度　未選択 排尿介助現状区分 見守り必要，介助必要	1，021	283	27．7%
全体	53，209	7，352	13．8%

ルール生成の結果，寝たきり度がいずれも重要な変数となった．この，厚生労働省による寝たきり度の定義を表 4 に示す．

表 4 障害高齢者の日常生活自立度(寝たきり度)

生活自立	ランクJ	何らかの障害等を有するが，日常生活はほぼ自立しており独力で外出する 1．交通機関等を利用して外出する 2．隣近所へなら外出する
準寝たきり	ランクA	屋内での生活は概ね自立しているが，介助なしには外出しない 1．介助により外出し，日中はほとんどベッドから離れて生活する 2．外出の頻度が少なく，日中も寝たり起きたりの生活をしている
寝たきり	ランクB	屋内での生活は何らかの介助を要し，日中もベッド上での生活が主体であるが，座位を保つ 1．車いすに移乗し，食事，排泄はベッドから離れて行う 2．介助により車いすに移乗する
	ランクC	1 日中ベッド上で過ごし，排泄，食事，着替において介助を要する 1．自力で寝返りをうつ 2．自力では寝返りもうてない

6. 考察

本稿ではアセスメントのデータを使用して介護度を推定することを試みた．そして、予測と正解に差異が発生している傾向の抽出を試みた．

本来は 1 次調査の結果と主治医意見書等の情報が揃ったうえで、1 次判定の結果、2 次判定の最終結果それぞれを予測のうえ比較を行うことが望ましいものの，本稿のように 87%ほどは上下 1 段階に収まっていることからも，全体的にはアセスメントの結果を用いて 2 次判定を含む要介護度について一定精度で推定することが可能そうである．

一方，2 次審査による要介護度の修正は一定数存在することが推測される．介護認定審査会委員テキストによれば，二次審査にて情報修正が発生しやすい 1 次調査における調査項目とその調査結果について記載が見られる．1 次判定結果が要介護度 2 の場合，口腔清潔:一部介助以上，洗顔:一部介助以上，整髪:一部介助以上，場所の理解:できない，となった場合により重度に変更することが多いとのことであった．本稿での結果からは、アセスメントで得られた寝たきり度ランクが B-1 以上の場合に重度になる傾向が見られていたが、障害自立度 B ランクの当否は、「日常生活活動のうち、食事、 排泄、着替のいずれかにおいては、部分的に介護者の援助を必要とし、1 日の大半をベッドの上で過ごす場合が該当」であるため、口腔清潔:一部介助以上が含まれていると考えられる．したがって，初回では難しいものの、更新の時点ではこのような情報を予め取り込むことで、効率的な要介護度判定が実施できるものと考えられる．

今後の課題として，アセスメントの目的に応じて母集団を分けた分析の実施や,本稿では使用しなかった自由入力欄のデータを使用した分析の実施が考えられる．

最後に，本稿においては医学的観点からの妥当性は検討しておらず，現介護保険制度について言及するものではない．実際の運用に向けては十分な医学的観点での見当も必要であると考えられる．

参考文献

[財務省　2015]財政制度等審議会財政制度分科会（平成 27 年 4 月 27 日開催）資料 1，Available at：http://www. mof. go. jp/about_mof/councils/fiscal_system_council/sub-of_fiscal_system/proceedings/material/zaiseia270427/01. pdf

[厚生労働省　2012]認定調査員テキスト 2009 改訂版(平成 27 年 4 月改訂)，Available at：http://www. mhlw. go. jp/file/06-Seisakujouhou-12300000-Roukenkyoku/0000077237. pdf

[厚生労働省　2012] 介護認定審査会委員テキスト 2009 改訂版(平成 28 年 4 月改訂)，Available at：http://www. mhlw. go. jp/file/06-Seisakujouhou-12300000-Roukenkyoku/0000116033. pdf

[川越 2017] 川越 雅弘: 一次判定方法への ADL 調査項目の関与と要介護度と ADL の関係性，老年精神医学雑誌 (0915-6305)28 巻 9 号，pages1004-1009，2017.

[高橋 2017] 高橋 恭子，築島 恵理: 介護保険新規認定者において要介護度が重度となる原因疾病の検討，日本公衆衛生雑誌 64(11)，pages655-663，2017 .

[関　2017]関　庸一，筒井孝子，宮野尚哉:要介護認定一次判定方式の基礎となった統計モデルの妥当性．応用統計学，29（2），pages101－110，2000.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

患者群の状態把握を目的とした臨床データ可視化の取り組み

Attempt to visualize clinical data for grasping status of patients

石井雅通 Masamichi Ishii　美代賢吾*2 Kengo Miyo

*1 NCGM（国立国際医療研究センター）
National Center for Global Health and Medicine

This paper describes an attempt to visualize clinical data collected from multi-institutional EHR System.

1. はじめに

多施設の医療機関から患者症例レジストリとして電子カルテに蓄積された臨床データを集積して臨床データベースを構築する事業が複数始まっている．当センターでは日本糖尿病学会と連携した「診療録直結型全国糖尿病データベース事業（Japan Diabetes comprehensive database project based on an Advanced electronic Medical record System：J-DREAMS）」を平成26年度から開始しており、2017 年 12 月現在、35 施設が参加し、今後4年間で 100 施設、10～20 万人規模での登録をめざしている。

こうした患者症状レジストリの構築では臨床医師が当該症例の観察項目を入力することが一般的に実施されており，収集されるデータ項目は専門家による正解ラベルとして取り扱い可能な項目及び既知の臨床知識から判断した重要な特徴量となる項目で構成されている．こうした医師のアセスメント結果を活用し，電子カルテに記録された患者基本情報，検査結果情報，投薬情報などの構造化データから抽出した特徴量と併せて分析することで，観察する臨床データ項目の標準化による診療の質の改善及び合併症進展リスク因子の検索，合併症進展抑制効果の期待される介入の同定，糖尿病薬の副作用についての発生頻度及びリスク因子の確認，症例の未解決課題の発見といった成果が期待されている．

2. 課題

2.1 臨床データに内在する新 Np 問題とスパースネス

集積した臨床データを解析するにあたっては多施設臨床データの特徴に制約を受ける．先ずデータ件数(n)に対してパラメータ数（p）が多い（n<<p）という新 NP 問題があり，かつ医療データがスパースネスであることである．

300 床以上の大規模病院であれば，採用医薬品は 2,000～3,000 品目，検査結果値は 3,000～5,000 項目，検査値以外の観察項目等はすの数倍あり，データ項目としては 10000 を超える．今後のゲノム医療の伸展により患者ごとの遺伝子変異がパラメータに加わると数万規模となる．

一方，患者の疾患は多様でありため特定の疾病に対する分析を目的とする場合，対象となる患者（すなわちデータ件数）が稀少疾患であれば一施設あたり数名から数十名であり，一般に母数が大きいと考えられる“がん”についても，研究のための解析にあたってはがん種やステージごとに患者群を絞り込む必要ため新 NP 問題からは逃れられない．

2.2 患者群の状態把握のために

臨床現場においては専門医が相対する患者の状態を把握して必要な医療を提供する判断をしているが，多施設から収集した臨床データ全体を取り扱う場合は個別の患者の状態をすべて把握して解釈するアプローチは困難である．

そこで，各症例の専門家による探索的な患者群の傾向分析を可能とすることで患者群の状態把握が可能となり，新たなクリニカルクエスチョンの選定に貢献できると考えた．

3. 提案手法

3.1 専門家による正解ラベル選択

症例レジストリで収集しているデータ項目から専門家により患者状態を表現したデータ項目を選択する．

3.2 構造化データ抽出

正解ラベル以外の構造化データ（患者基本情報，処方情報，検査結果情報）を抽出する．

3.3 患者症状単語抽出

診療録等の非構造化データから患者症状を表現する用語を抽出する．ここでは専門用語辞書として万病辞書を使用する．

3.4 患者症状単語統計

前記抽出結果より頻度統計を作成する．

3.5 時系列データ処理

3.2 及び 3.4 の抽出データを入力として，症例レジストリ収集時点を起点として時系列情報を考慮した特徴量データを作成する．例えば，患者ごとの前回検査値からの変化量，次回検査値の変化量等を算出する．

3.6 ベイジアンネットワーク構築

3.1 及び 3.5 の出力データを入力として，ベイジアンネットワークを構築する．

連絡先：石井雅通，国立国際医

3.7 専門家による主成分探索

3.6 で構築したベイジアンネットワークを利用して当該症例の専門家による探索的な患者群の傾向分析を試行する.

3.8 クリニカルクエスチョンの検討

3.7 による合併症進展リスク因子, 合併症進展抑制効果の介入, 医薬品の副作用についてのリスク因子等の観点において新たなクリニカルクエスチョン創出の可能性について評価する.

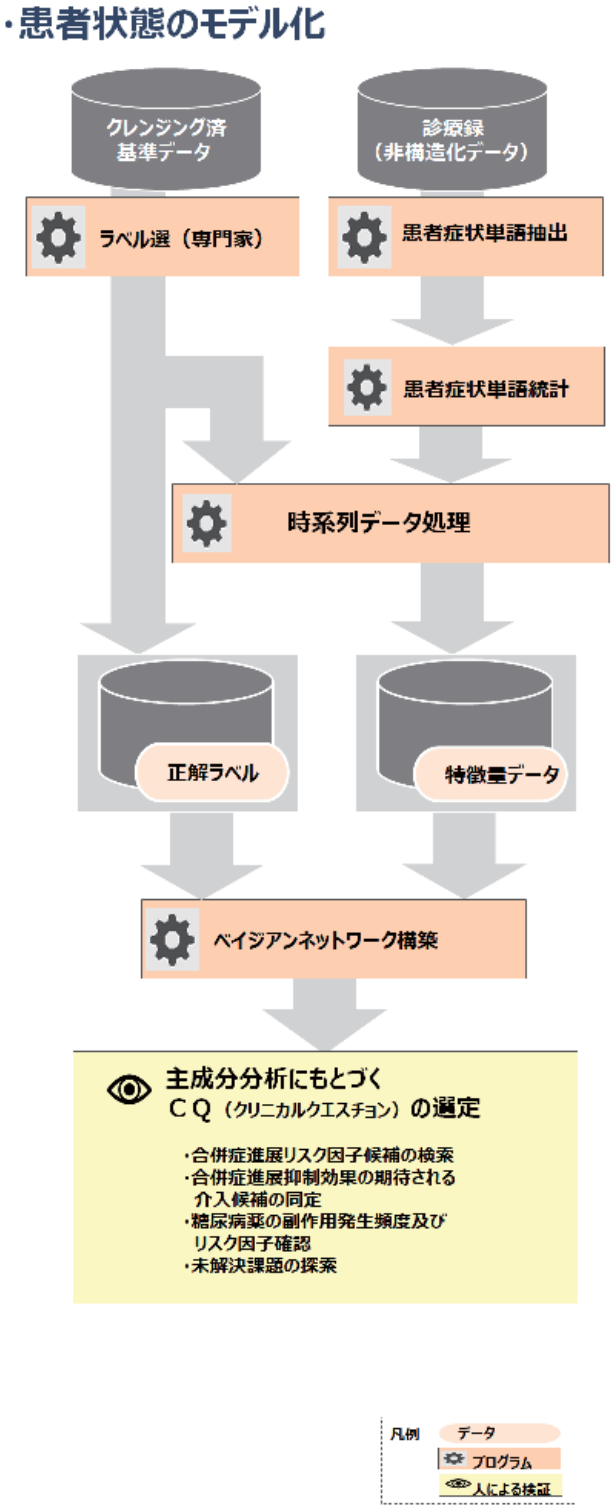
4. 今後の計画

プロトタイプによる評価実験を計画しており, 提案手法で述べた専門医によるレジストリへの入力データを正解ラベルとして活用する手法により, 専門家以外が専門家の協力を得ることが容易となり, ベイジアンネットワークによる可視化により専門家の気づきや暗黙知といったフィードバックをモデルへ反映することが容易となる好循環が期待できる.

ベイジアンネットワークを用いた感度分析を実施することでデータ収集事業が目的としている, (1)合併症進展リスク因子の検索, (2) 合併症進展抑制効果の期待される介入の同定, (3) 糖尿病薬の副作用についてのリスク因子の確認, (4) 糖尿病の未解決課題の発見, への貢献が期待できる.

上記プロセスで患者状態を表現する代表的な特徴量を抽出し, 患者状態モデルを構成する要素として特定していくスパイラルを継続することで既存の収集対象のデータ項目の解析での貢献度を元に精査することが期待できる. 入力データの評価は効率的なデータ収集項目への見直しの機会となり, データ入力元である臨床医の負担軽減の一助となすと考える.

謝辞:この研究の一部は AMED の課題番号 JP17lk1010010 の支援を受けた



参考文献

[J-DREAMS] 診療録直結型全国糖尿病データベース事業, http://jdreams.jp/

[PMDA] MID-NET (Medical Information Database Network), https://www.pmda.go.jp/safety/mid-net/0001.html

[田中博 15] 人工知能 (AI) とビッグデータの 医療と創薬への応用, http://ngbrc.com/img/file54.pdf, 2013

[統計関連学会連合理事会および有志 07] 統計関連学会連合理事会および有志, 我が国の統計科学振興への提言, http://www.jfssa.jp/TokeiKagakuShinkou0702.pdf, 2007

[矢野 17] 矢野憲,若宮翔子,荒牧英治,医療テキスト解析のための事実性判定と融合した病名表現認識器, 言語処理学会第 23 回年次大会, 2017

[荒牧 17] 荒牧英治,岡久,矢野ほか, 大規模医療コーパス開発に向けて, 言語処理学会 第 23 回年次大会 発表論文集, 2017

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

睡眠日誌アプリの開発と睡眠習慣改善行動の継続の予測

Development of “Sleep Diary” application and continuation prediction of behavior for improvement sleeping habits

秋冨 穣[*1] Jou Akitomi　梶山 征央[*1] Ikuo Kajiyama　石井 美穂[*1] Miho Ishii　岡島 義[*2] Isa Okajima　山口 美峰子[*1] Mineko Yamaguchi

[*1] NEC ソリューションイノベータ株式会社 NEC Solution Innovators, Ltd.

[*2] 東京家政大学 Tokyo Kasei University

We consider that computer assistance at effective timing will be an important element in the user's behavioral change using an application. So, we tried to predict "effective timing" by using input data for first three days. As the first step, we tried a prediction of whether the user's behavior continued for several weeks.

1. はじめに

2017 年は睡眠負債という言葉が話題となったが，不眠や睡眠不足など，睡眠に問題を抱えることは精神面や身体面の両面に悪影響を与えることが知られており，疲労感や集中力の低下による産業事故のリスクを増大させ，生産性の低下など，社会的な問題を引き起こすと言われている[厚生労働省 2014]．特に我が国では，20 歳以上のおよそ 5 人に 1 人の割合で睡眠の問題を抱えていると言われており[Doi 2003]，睡眠の問題は大きな社会問題の一つであると言える．

睡眠改善の手法の一つとしては，不眠のための認知行動療法（Cognitive Behavioral Therapy for Insomnia: CBT-I）があり，これまでにその有効性が示されている[Okajima 2011]．CBT-I は特に欧米において既に広く普及しているが，我が国においては普及率が未だ低く，段階的治療の観点からも多くの人がアクセスできる CBT-I ツールの開発の必要性が挙げられている[井上・岡島 2012]．

こうした状況を踏まえ，我々は CBT-I 技法の一部である睡眠日誌の記録と睡眠衛生の教育を提供するためのアプリである「睡眠日誌」アプリを試作開発した．ユーザは，睡眠日誌アプリの利用により，日々の睡眠を記録し，良い睡眠習慣について学び，日々の睡眠の状態を振り返る．ユーザがこのアプリの利用を継続することができれば，ユーザの睡眠習慣が改善されることが期待される．

しかし，このようなユーザの生活習慣の改善，行動変容を促すための取り組みは，必ずしもユーザの継続率は高くなく，途中で脱落するケースが多いという問題が指摘されており[宗 2016]，この問題を解決するために，エージェントによる介入技術を組み入れ，エージェントによる介入を行う方法などが提案されている[日室 2017]．

これらの方法は継続率の向上に一定の成果を挙げているが，ここにそれとは違う視点からのアプローチを追加することにより，より高い効果を発揮することが期待される．そこで，本研究では，これに付加する新たなアプローチとして，アプリの利用の継続を促すための介入方法そのものではなく，介入するべきタイミングや，介入するべきユーザを予測することに注目した．本稿ではその最初のステップとして，睡眠日誌アプリに蓄積されたデータを基に support vector machine（SVM）の判別器を作成し，あるユーザが入力した最初の数日間のデータから，数週間後にそのユーザがアプリの利用を継続しているか否かを予測することを試みた．

連絡先: 秋冨 穣， NEC ソリューションイノベータ株式会社，東京都江東区新木場 1-18-7， 03-5534-2619，jou-akitomi@vt.jp.nec.com

2. 睡眠日誌アプリ

睡眠日誌アプリは，2016 年 3 月から一般公開されており，ユーザは Android 版と iOS 版をどちらも無償で利用できる．睡眠日誌アプリには大きく 3 つの機能がある．

2.1 睡眠に関する情報の記録機能

睡眠日誌アプリでは，ユーザは，自身の睡眠に関する情報を日々記録する．記録する内容は，布団に入った時刻，眠りについた時刻，目覚めた時刻，布団から出た時刻などの睡眠そのものに関する情報が中心であるが，その他にも，寝る前にアルコールを摂取したか否か，寝る前にカフェインを摂取したか否か，寝る前にタバコを吸ったか否かなど，それに付随する情報も含まれる．

特に睡眠時間に関する情報に関しては，ユーザが自身の睡眠の状態を自覚することを特に重視するために自動記録の機能は取り入れておらず，自己申告の睡眠時間を記録するというコンセプトになっている．具体的な記録内容は，図 1 の画面で振り返ることが出来る．

2.2 睡眠に関する知識（睡眠衛生）の学習機能

睡眠日誌アプリでは，ユーザは，睡眠の専門家が監修した睡眠に関するコラムを，常時 30 本以上読むことが出来る．このコラムはマイページからいつでも読むことが出来る他，毎日の睡眠の記録をした際におすすめのコラムが 1 本推薦される．コラムは，例えば図 2 のような画面で提供される．

2.3 睡眠と活動量の振り返り機能

睡眠日誌アプリでは，一週間を単位として，ユーザは自身の睡眠の状態と，その日の日中にどのくらいの活動が出来たか（元気度）をつき合わせて閲覧することが出来る．

例えば，睡眠が乱れた翌日は元気度が低いであるとか，逆に，睡眠が乱れていたとしても意外と日中元気であることが多いであるとか，そのユーザに適したそのユーザだけの睡眠習慣に対する気付きを得ることが期待できる．一週間の振り返りの画面は，図３の画面で提供される．



図1. 睡眠日誌アプリの記録内容の例

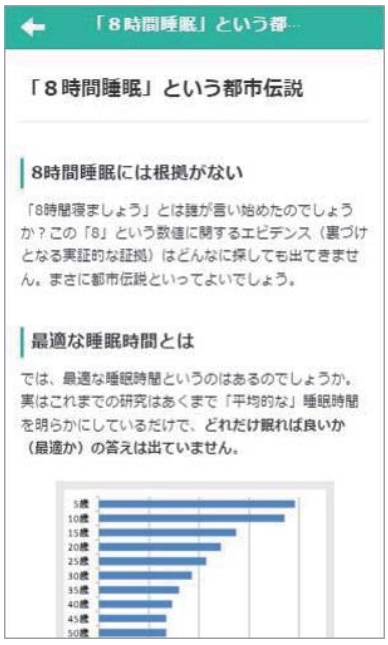


図2. 睡眠日誌アプリで提供されるコラムの例

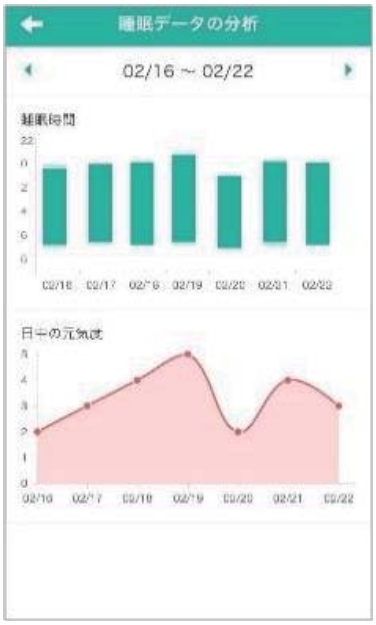


図3. 睡眠日誌アプリの一週間の振り返りの例

3. 睡眠日誌アプリユーザの睡眠の改善

これまでに述べたように，睡眠日誌アプリには，ユーザが良い睡眠習慣を獲得するための要素が含まれている．そのため，睡眠日誌アプリの継続利用者は，睡眠の問題が改善する可能性があると考えられる．睡眠状態を測るためには，例えばアテネ不眠尺度（AIS）と呼ばれる世界各国で使われているアンケートに回答することがその手段として用いられる．睡眠日誌アプリ内にも，ユーザが AIS に回答する機能と，その得点の推移を確認する機能が盛り込まれている．

3.1 アテネ不眠尺度（AIS）

アテネ不眠尺度は 8 問で構成されるアンケートである．ユーザの回答結果に応じて，0 点から最大 24 点までの得点が算出される．この得点は，以下のように解釈される．

- 6～24：不眠症の疑いあり
- 4～5：不眠症の疑いが少しあり
- 0～3：不眠症の疑いなし

3.2 睡眠日誌利用継続ユーザの AIS の変化

実際に睡眠日誌アプリを継続利用することによってユーザは自主的に良い睡眠習慣を獲得し，AIS の値が改善するかどうかについての予備的な調査を行った．調査は 2016 年 10 月から，2016 年 12 月までのデータを用いて行った．その結果を図 4 に示す．

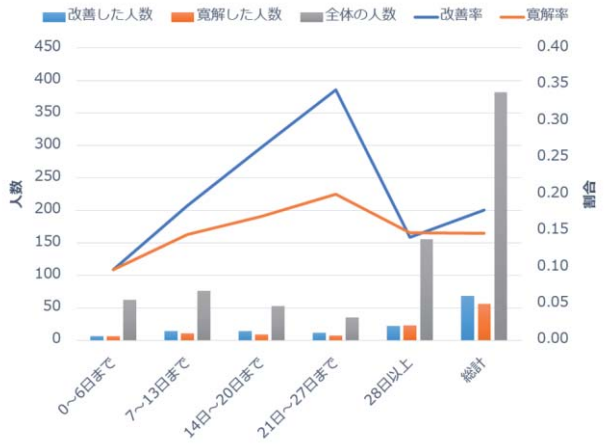


図４．睡眠日誌ユーザの継続利用日数と AIS の変化

図４では，ユーザのアプリ継続期間内において AIS の測定を 2 回以上実施し，AIS が 3 点以下になった者を「寛解」と定義し，AIS が 5 点以上低下した者を「改善」と定義した．図４に寄ると，睡眠日誌アプリユーザの中で寛解した者や改善した者は利用継続日数が増えるにつれて増加しており，増加のピークは 21～27 日の期間，即ち 4 週間程度の利用期間であることが示された．ただし，この集計においてはユーザの行動は統制されておらず，睡眠日誌アプリを利用継続するユーザが自主的に通院やその他の手段を用いて睡眠改善に取り組んだ効果が含まれている可能性があるため，この睡眠改善効果が睡眠日誌アプリによるものであると断定することは出来ない．

4. 睡眠習慣改善行動の継続の予測

アプリの利用の継続を促すために介入が必要なユーザや介入が必要なタイミングを予測するための第一歩として，アプリのユーザの 3 日分の入力データから，そのユーザがその後数週間アプリの利用を継続するか否か（規定回数以上の睡眠の記録をするか否か）の予測を試みた．

なお，睡眠日誌アプリは利用を開始する際にユーザからデータの研究利用についての同意を得ており，データは全て個人を特定できない形に処理した上で使用した．

4.1 使用したデータ項目

予測に用いたデータは，全て睡眠日誌アプリの入力データから抽出した．抽出したデータ項目は，具体的には以下の通りである．

(1) デモグラフィック

- 性別：0, 1
- 年代
- 初期の AIS の値
- 入眠困難を感じるか：0, 1
- 中途覚醒があるか：0, 1
- 早期覚醒があるか：0, 1
- 熟眠障害があるか：0, 1

(2) 日々の記録

- 入床時刻
- 入眠時刻
- 覚醒時刻
- 離床時刻
- 総睡眠時間
- 総就床時間
- 入眠潜時
- 離床潜時
- 中途覚醒時間
- 総覚醒時間
- 睡眠効率
- 入床時刻と離床時刻の中心時刻
- 入眠時刻と覚醒時刻の中心時刻
- 日中の活性度
- 昼寝時間
- 夜にアルコールを摂取したか否か：0, 1
- 夜にカフェインを摂取したか否か：0, 1
- 夜にタバコを吸ったか否か：0, 1
- 自由記述のコメントを入力したか否か：0, 1
- 覚醒してから睡眠を記録するまでにかかった時間
- AIS の測定回数

4.2 実験内容

睡眠日誌アプリを利用しているユーザの中から 10000 人分のデータを選択し，その中から特に、アプリの利用開始日から 3 日間連続で睡眠の記録をつけているユーザのみを抽出した．3 日分のデータはユーザごとに 3 日間の平均値，中央値，標準偏差の 3 通りのデータセットに加工し，そのユーザがその後規定回数分の睡眠記録をつけたか否かのラベルを付加することで，SVM の入力データとした．

5. 結果と考察

平均値，中央値，標準偏差の各データセットについて，2 週間継続（2 週間目に 3 日以上の睡眠記録を実施するか否か），3 週間継続（3 週間目に 3 日以上の睡眠記録を実施するか否か），4 週間継続（4 週間目に 3 日以上の睡眠記録を実施するか否か）の予測を実施し，11-fold cross-validation で，accuracy を評価した．その結果を以下の表１に示す．

表１．SVM による初期 3 日データからの継続予測結果

	平均値	中央値	標準偏差
2 週間継続	66.1	65.4	71.0
3 週間継続	61.7	60.1	61.9
4 週間継続	63.9	63.6	64.9

表１によると，平均値，中央値，標準偏差という 3 つのデータセットの中では，比較的標準偏差のデータによる予測結果の精度が良く，アプリの最初の 3 日の入力データから 2 週間後～4 週間後の利用継続の予測が可能である可能性があることが分かった．標準偏差による分析結果が最も

精度が良かったということから考えると，アプリの利用を継続するユーザと継続しないユーザの差は 3 日の入力データの標準偏差に最も現れやすいということになる．標準偏差の大小は，言うまでも無く，算出する数値のばらつきを反映する．つまり，今回の結果は，継続するユーザと継続しないユーザでは，入力データのばらつきに差があるという可能性があることを示している．ここで入力データのばらつきが少ないということはユーザの日々の睡眠の状態が良くも悪くも一定であるということを示しており，ばらつきが大きいということは，変動が大きいということを示している．これがアプリの利用継続率に影響を与えているとすれば，元々の生活習慣がアプリの利用継続率に影響を与えているという可能性が示唆される．

また，今回実施した内容は本格的な解析の前の予備的な検討であったが，今後予測モデルを更新し，パラメータをチューニングすることにより，より実用的な予測結果を得られる可能性があると言える．

6. おわりに

本稿では特に対象を睡眠日誌アプリのユーザに限定し，ユーザの睡眠の問題解決のための生活習慣改善，行動変容を行うための方法の提案を行った．睡眠日誌アプリに限らず，ヘルスケアアプリを用いた行動変容においては，ユーザごとに最適化されたプログラムの提供や，最適なタイミングでの介入を行うことによる行動の継続支援が重要な要素となると考えられる．本稿で取り組んだ内容は初期 3 日の入力データからの規定の期間の継続の予測であったが，これが実用的なレベルで実施可能になると，初期 3 日のデータからユーザの属性を推定し，そのユーザに最適な介入を実施することが可能になると考えられる．本稿の内容には未だ改良の余地が大いにあるため，引き続き予測精度の向上や，新たな課題への応用の検討を進めていきたい．

謝辞

本研究について大きな貢献をしていだいた NEC ソリューションイノベータの村上紘士さんをはじめ，本研究に協力していただいた皆様に感謝をいたします．

参考文献

［厚生労働省 2014］ 厚生労働省健康局：健康づくりのための睡眠指針 2014，2014.

[Doi 2003] Yuriko Doi, Masumi Minowa and Toshiro Tango: Impact and Correlates of Poor Sleep Quality in Japanese White-Collor Employees, *SLEEP*, Vol.26, No. 4, 2003.

[Okajima 2011] Isa Okajima, Yoko Kodama and Yuichi Inoue: A meta-analysis on the treatment effectiveness of covnitive behavioral therapy for primary insomnia, *Sleep and Biological Thythms*, Japanese Society of Sleep Research, 2010.

［井上・岡島 2012］ 井上雄一 ・岡島義：不眠の科学，朝倉書店，2012.

［宗 2016］ 宗未来・関沢洋一・竹林由武：人工知能で，人のこころは癒せるか？，人工知能（自然言語処理）フィードバック機能搭載型のインターネット認知行動療法（iCBT-AI）の抑うつ者に対する世界初の効果検証（無作為統制試験），RIETI Discussion Paper Series, RIETI，2016.

［日室 2017］ 日室聡仁，山本弘樹，武智小百合，新貝敦，山口美峰子：コンピュータ支援型認知行動療法におけるエージェントが行動継続に与える効果の検証，人工知能学会全国大会（31 回）講演論文集 4Q1-10in2，2017.

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

ニューラル機械翻訳における単語予測の重要性について

竹林　佑斗 *1　Chenhui Chu*1　荒瀬 由紀 *1　永田　昌明 *2

*1大阪大学　*2日本電信電話株式会社

1. 序論

本稿では、人手または統計的手法により自動作成された対訳辞書を利用するニューラル機械翻訳 (NMT) のデコーディング法を提案する。提案手法は、入力文から出力文に出現すべき単語集合を予測する方法と、予測された単語集合が生成されることに報酬を与えるようなデコーディング時のスコアの計算法から構成される。本稿では、科学技術論文の日英翻訳タスク (ASPEC-JE) を対象として、予測単語集合の再現率と適合率を様々に変化させるシュミレーションにより提案手法の有効性を検証した。例えば、予測単語集合としてオラクル単語集合 (正解訳の単語集合) を与えた場合、提案したスコア計算法を用いることにより BLEU スコアが 7.74 ポイント向上することを確認した。

2. 単語集合予測に基づく単語レベル報酬モデル

ここで、本稿の提案法を述べる。提案法は、大きく二つの構成要素からなる。

1. 入力系列から出力系列に出現する各単語の集合（目的単語集合）を予測するモデル
2. 目的単語集合を用いてデコーダが選択する単語を制御する機構

図 1 にデコーディングステップ j における提案法の振る舞いを示す。まず入力系列から目的単語集合を予測する。予測された単語の生成確率に報酬を与えることで、その単語を生成されやすくする。各ステップにおいて、以下の式を用いて単語の生成確率を増加させる。この式は [Li 17] の，出力系列の長さを制御するモデルを参考にした。

$$y_j = \arg\max_y (\log P(y_j|y_{<j}, \mathbf{x}) + \lambda d_y), \qquad d_y \in \mathbf{D}_{f2e}, \quad (1)$$

ここで $\mathbf{D}_{f2e}$ はは対訳辞書とし、λ は報酬の重みである。d_y は、 $y \in D_{f2e}$ ならば 1、そうでなければ 0 であるような関数である。

3. 単語集合予測のシミュレーション

本稿では、提案手法の 2 番目の構成要素の有用性を検証するため実験を行なった。まず参照文に対する単語予測の適合率 (precision) と再現率 (recall) が共に 100%であるオラクル単語集合を作成し、提案手法の効果の上限値を調査した。さらに、オラクル単語集合への単語の追加および削除の処理を加えることで、適合率と再現率をそれぞれ変化させて、予測結果のシミュレーションを行いその時の翻訳性能を測った。その結果、どのような適合率と再現率を持つ目的単語集合を予測できれば、提案手法を用いることでベースラインを大きく上回る結果が得られるかが明らかとなった。

3.1 対訳データと翻訳ツール

ASPEC-JE を用いて提案手法の有効性をシミュレーションする。このデータでは 300 万文のうち上位 200 万文を訓練データとし、 開発データとして 1,790 文、テストデータとして 1,812 文を用いた。提案モデルのベースラインとして「深層学習による自然言語処理」*1 の付録として公開されている mlpnlp-nmt*2 (注意機構付き LSTM エンコーダデコーダモデル) を用いた。このモデルは WAT-2017 の ASPEC-JE/EJ タスクにおいて人手評価でもっとも優れた性能を納めたものである [Nakazawa 17]。提案手法はこのモデルのデコーダ部を上述の式の通り書き換えることで実装した。ハイパーパラメータの設定は [Morishita 17] に従った。

3.2 オラクル単語集合からの精度調整

オラクル単語集合は参照文の単語を目的単語集合とすることで得た。シミュレーション用の擬似予測単語集合を作成する際はオラクル単語集合に下記の処理を加えることで得た。再現率を減少させる際は、オラクル単語集合から複数の単語を無作為に削除し True positive を減少、および False negative を増加させることで調整した。適合率を減少させる際は、オラクル単語集合に複数の単語を追加し False positive を増加させることで調整した。近似予測をできる限り現実的にするため、追加する単語は GIZA++*3 でワードアライメントを取った結果、入力系列の単語から出力系列単語への単語翻訳確率の高い単語から追加した。

3.3 実験結果

オラクルの単語集合を用いた時の BLEU スコアは、35.64 であった。これはベースラインシステムの BLEU スコアの 27.90 から+7.74 ポイントの向上であり、ベースラインの翻訳性能を有意に上回っている ($p < 0.01$)。図 2 に擬似予測単語集合を用いたシミュレーションの結果を示す。各点は擬似単語集合の適合率と再現率をそれぞれ 10% – 90% まで 10% 刻みで変化させた時の、BLEU スコアを表している。*border* は、その線以上の結果はベースラインを有意に上回っているという境界線を示している。したがって、*border* 以上の適合率と再現率をもつ単語集合を入力系列から予測できれば、我々の提案手法を用いることでベースラインを優位に上回る性能が得られる。結果より、適合率よりも再現率のほうが翻訳性能の向上に大きく貢献しており、例えば適合率が 10%であっても再現率が 50%あれば翻訳性能が大きく向上するが再現率が 10%であると適合

連絡先: 竹林　佑斗、大阪大学、大阪府吹田市山田丘 1-5、takebayashi.yuto@ist.osaka-u.ac.jp

∗1 http://www.kspub.co.jp/book/detail/1529243.html
∗2 https://github.com/mlpnlp/mlpnlp-nmt
∗3 http://code.google.com/p/giza-pp

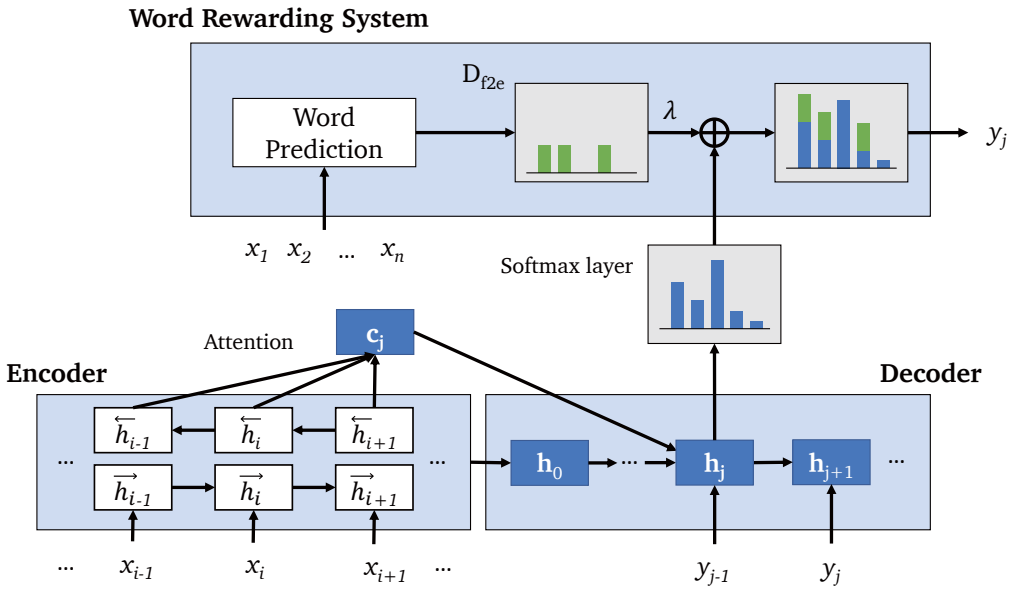


図 1: デコーディングステップ j における単語レベル報酬モデル。各デコーディングステップにおいて、予測された目的単語集合内の単語生成確率に報酬を与えることで、生成されやすくする。

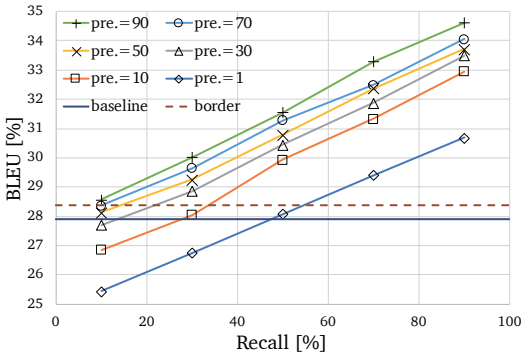


図 2: 擬似予測単語集合を用いて適合率と再現率を変化させた時の BLEU スコア。*boder* は翻訳結果がベースラインと差があるものの境界を示しており、*border* より上の結果はベースラインを優位に大幅に上回っている ($p < 0.01$) ものである。

率が 70%より多くなければ *border* を超えないという結果が得られた。

4. 結論

本稿では、出力系列を制御する手法として、単語レベル報酬モデルを提案した。このモデルは辞書などの外部知識を NMT モデルに適用するものであり、NMT モデルを再訓練する必要がないのが特徴である。実験では擬似予測を行い適合率と再現率を変化させて目的単語集合をシミュレーションし、それを用いた時の BLEU スコアを求めて、ベースラインと統計的有意差があるかどうかを検定した。その結果、予測モデルで作成する目的単語集合に求められる適合率と再現率が明確になった。今後の課題としては、実際に目的単語集合を予測して本稿のシミュレーション結果を満たす結果が得られるかを検証する。

参考文献

[Li 17] Li, J., Monroe, W., and Jurafsky, D.: Learning to Decode for Future Success, *CoRR*, Vol. abs/1701.06549, (2017)

[Morishita 17] Morishita, M., Suzuki, J., and Nagata, M.: NTT Neural Machine Translation Systems at WAT 2017, in *Proceedings of the 4th Workshop on Asian Translation (WAT2017)*, pp. 89–94, Taipei, Taiwan (2017), Asian Federation of Natural Language Processing

[Nakazawa 17] Nakazawa, T., Higashiyama, S., Ding, C., Mino, H., Goto, I., Kazawa, H., Oda, Y., Neubig, G., and Kurohashi, S.: Overview of the 4th Workshop on Asian Translation, in *Proceedings of the 4th Workshop on Asian Translation (WAT2017)*, pp. 1–54, Taipei, Taiwan (2017), Asian Federation of Natural Language Processing

The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

自己修復ネットワークの異常ノード撲滅条件の一般化

Research of abnormal node eradication conditions of self-repair network

田村　真優[*1]
Yoshiyuki Tamura

石田　好輝[*2]
Yoshiteru Ishida

[*1] 豊橋技術科学大学
Toyohashi University of Technology

[*2] 豊橋技術科学大学
Toyohashi University of Technology

In the self-repair networks, a ring network has been studied. However, the network structure in the real world is more complex and diverse. This research discussed the boundary between the spread phase and the eradication phase of abnormal nodes by Mean Field Approximation (MFA) in the regular graph of degree three. We also conduct agent simulations for the regular graphs with degrees exceeding three. As a result, we have derived parameter region that guarantees abnormal nodes extinction.

1. はじめに

本研究では,ネットワークの構成要素であるノードが互いに修復をし,ネットワーク全体を正常な状態に保つ「自己修復ネットワークモデル」を扱う.

先行研究[1]ではリング状のネットワークが研究対象とされてきたが,現実世界のネットワーク構造はより複雑で多様なものである.

そこで,本研究では,次数が 3 の正則グラフのネットワーク構造について,自己修復ネットワークの異常ノード撲滅条件の導出を行い,より複雑な構造のネットワークの安定性についての考察を行う.

2. 自己修復ネットワークについて

自己修復ネットワークとは,ネットワークを構成するノードが修復試行確率 Pr で隣接するノードと相互に修復を行い,ネットワーク全体を正常に保つモデルである.各ノードは「正常」か「異常」のどちらかの状態をとるものとする.

修復は必ず成功するものではなく,修復元ノードに設定されている修復成功率で成功する.そして,修復対象のノードの状態は修復に成功した場合は正常,修復に失敗した場合は異常となる.ただし,修復元ノードの状態によって修復成功率が定まり,正常ノードによる修復成功率 Prn は異常ノードによる修復成功率 Pra 以上とする($Prn>Pra$).また,一つのノードに対して複数のノードから修復が行われた場合,本研究では,修復が全て成功した場合のみ修復成功として扱う(AND 修復).つまり,一つでも修復が失敗すれば修復を受けたノードは異常となる.

また,正則グラフとは,全てのノードが同じ数のノードと接続されているグラフのことを表す.先行研究で主に研究されてきたリング状のネットワークは,次数が 2 の正則グラフのネットワーク構造である.

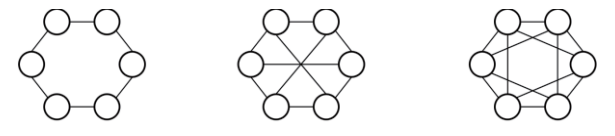


図 1. 正則グラフの例

田村真優, 豊橋技術科学大学 情報・知能工学専攻,
tamura@sys.cs.tut.ac.jp

自己修復ネットワークの目的はネットワーク内の異常ノードを撲滅し,ネットワーク全体を正常に保つことである.しかし,パラメータの取り方によっては修復の失敗によってネットワーク内が異常ノードのみになってしまう場合もある.本論文では,異常ノードを撲滅できるパラメータ領域を撲滅相,それ以外の領域を蔓延相と呼ぶ.

3. 結果

3.1 異常ノード撲滅条件の導出

平均場近似による近似計算で異常ノードの撲滅条件を計算したところ,以下の理論式で撲滅条件が表されることが分かった.

$$Pr^2-3Pr+3Pra\geqq 0$$

この条件を満たす領域を図 2 に示す.

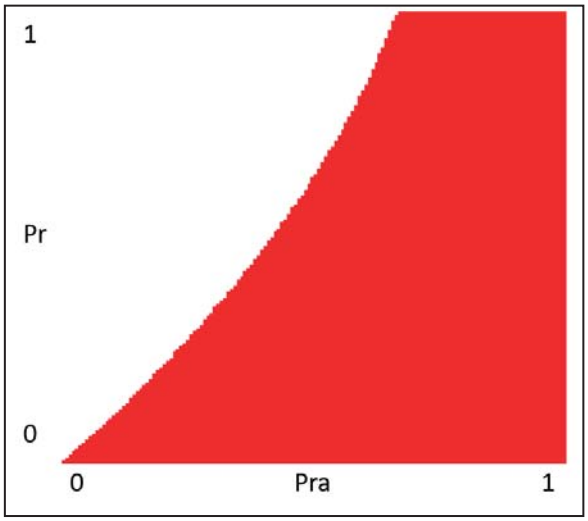


図 2. 異常ノード撲滅条件を満たすパラメータ領域

3.2 シミュレーションの結果

3 次の正則グラフのネットワークについて,表 3 の条件で自己修復のシミュレーションを行った.

表 3. シミュレーションの条件

ノード数	100
ステップ数	2000
修復試行率 Pr	0~1
正常ノードによる修復成功率 Prn	1
異常ノードによる修復成功率 Pra	0~1

今回のシミュレーションでは, 修復試行率と異常ノードによる修復成功率を 0 から 1 の間で 150 段階に変化させ, 計 22500 回のシミュレーションを行った. そして, 各試行において 2000 ステップ後の正常なノードの割合を算出した. この結果を図 4 に示す.

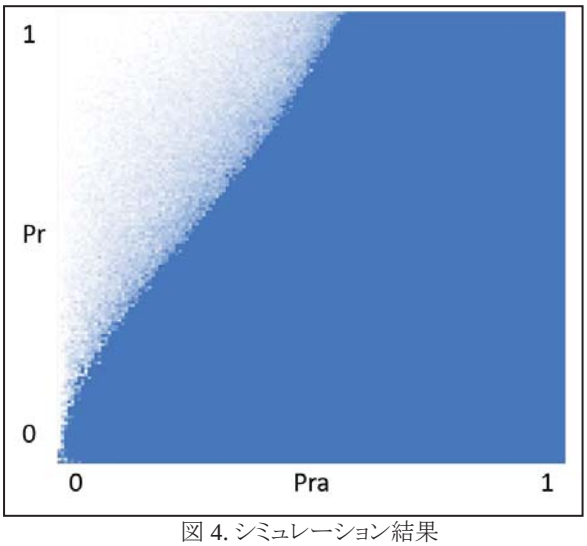


図 4. シミュレーション結果

ここで,グラフの色の濃さは 2000 ステップ後の正常ノードの割合を示す.すなわち,青色の領域は 2000 ステップ後にネットワーク内の全てのノードが正常となり,異常ノードを撲滅できたことを表す.

同様の条件で,次数が 4, 6, 8, 10 の正則グラフのネットワーク構造についても自己修復のシミュレーションを行った.シミュレーションの結果を以下の図 5(a)~(d)に示す.

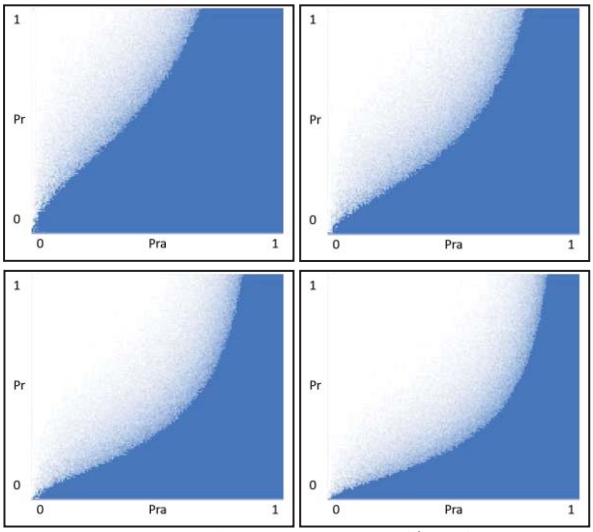


図 5. (a) 4 次　シミュレーション結果(左上)
図 5. (b) 6 次　シミュレーション結果(右上)
図 5. (c) 8 次　シミュレーション結果(左下)
図 5. (d) 10 次　シミュレーション結果(右下)

4. 考察

シミュレーションの結果から,次数が増えるほど異常ノードの撲滅条件が厳しくなる,すなわち,撲滅相が狭くなり蔓延相が広くなることが分かった.また,先行研究[1]より,*Pra* が小さい場合 *Pr* が大きくなると異常が蔓延してしまうことが分かっており,今回のシミュレーションでも同じ結果が得られた.これらの結果から,正則グラフの次数が増えネットワーク構造が複雑になることと,修復試行率 *Pr* が大きくなることは似たような効果があると考えられる.

この考察に関して,先行研究[1]より,2 次の正則グラフのネットワーク構造の場合,*Pra*/*Pr* ≤1/2 の条件で異常ノードが絶滅することが分かっている.そこで,2 次の正則グラフを基準として,次数が *n* 倍になることは *Pr* が *n* 倍になることと等価とみなせるのではないか,つまり,*n* 倍の数のノードから同じだけ修復を受けることと,1 つのノードから *n* 倍修復を受けることは同じとみなせるのではないかと考え,実際にシミュレーション結果と *Pr* を *n* 倍にした一次式の線を比較したところ,結果は以下のようになった.

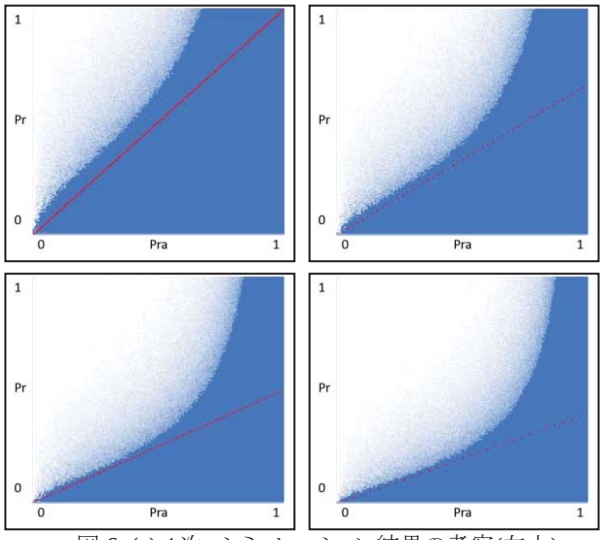


図 6. (a) 4 次　シミュレーション結果の考察(左上)
図 6. (b) 6 次　シミュレーション結果の考察(右上)
図 6. (c) 8 次　シミュレーション結果の考察(左下)
図 6. (d) 10 次　シミュレーション結果の考察(右下)

Pr と *Pra* が小さいとき,すなわちグラフの左下の領域では,一次式の条件でもシミュレーション結果とフィットしている.本研究では修復方式として AND 修復を採用しているため,修復の衝突が発生すると修復成功率は下がってしまう.しかし,修復試行率 *Pr* が低い場合,修復の衝突がほとんど発生せず,修復の衝突の影響が無視できるほど小さくなるため,このような結果になったのだと考えられる.

この結果から,修復の衝突が発生しない,あるいは,修復の衝突の影響が無視できる場合については,次数が *n* 倍になることと *Pr* が *n* 倍になることが等価とみなせるといえる.

5. おわりに

本研究では,3 次の正則グラフのネットワーク構造に平均場近似を適用することにより,異常ノード撲滅条件を導出した.また,修復方式として AND 修復を採用した場合,正則グラフの次数が増えネットワーク構造が複雑になると,異常ノード撲滅相が狭くなり,蔓延相が広くなることを確認した.

参考文献

[1] Ishida, Y.: Self-Repair Networks, Springer, 2015.