

信頼される AI システムを支える基盤技術  
2020 年度採択研究代表者

2022 年度  
年次報告書

乾 健太郎

東北大学 大学院情報科学研究科  
教授

知識と推論に基づいて言語で説明できる AI システム

主たる共同研究者:

久木田 水生 (名古屋大学 大学院情報学研究科 准教授)

黒橋 禎夫 (京都大学 大学院情報学研究科 教授)

戸次 大介 (お茶の水女子大学 基幹研究院 准教授)

## 研究成果の概要

本研究では、自分の判断を言語で説明することができ、対話的な説明コミュニケーションを通して人の判断を支援する AI システムの設計論の構築を目指している。2021 年度は、以下の研究を実施した。

乾グループは、記号推論と深層ニューラルネット (DNN) の融合によって「自然言語に基づく記号推論」を実現する研究を進めている。2022 年度は、Cross-stitch 機構によって KG とテキストの情報・知識を高度に補う合う KG-Text 混合グラフ埋め込みを実現し、関係抽出モデルの遠距離学習課題において複数のデータセットで世界最高性能を達成した (Dai et al. EMNLP'22)。また、推論システムを訓練・評価するための KG-Text 混合型 QA データセットの構築に取り組み、(i) 因果関係 QA データセットに半構造化した説明を付与した因果関係説明タスクのデータセット (Brassard et al. LREC'22)、および (ii) 既存 KG と Wikipedia 記事を組み合わせる回答を導く能力を問う日本語説明付きマルチホップ QA データセットのパイロット版を構築した (北村ら NLP'23)。さらに、記号推論の例として算術推論課題を取り組み、構成性に関するニューラル言語モデルの汎化能力等について統制実験を行った (Kudo et al. EACL'2023, Aoki et al. EACL-Findings'23)。

戸次グループは、2022 年度は、研究実施項目【DNN と相互作用する高階論理推論】において、Neural DTS の学習アルゴリズムの提案、言語モデル RNN-CCG のパーザ部実装、Coq タクティクによる DTS の自動証明のテスト実装などを行った。また、研究実施項目【説明可能 AI の証明論的意味論】において、英語の比較構文の分析、数量詞の含意と推意データセットの構築と評価、日本語 CCG ツリーバンクの批判的検討と新たな構成法の提案、などの研究を行った。さらに【ファクトチェック支援システム】に関する研究項目について、仕様書から状態遷移図を自動生成する研究が行われた。以上の研究成果に加え、国際会議 LENLS19 の開催等によって、成果の普及活動や若手研究者の育成にも力を入れた。

黒橋グループは映画推薦ドメインにおける知識に基づく説明対話コーパスの構築を継続した。今年度は高品質な対話コーパスを収集するため、アノテーションの専門会社に依頼し、1,275 件の対話 (合計 22,280 発話) からなる高品質な知識に基づく説明対話コーパスを構築した。対話中に登場する各エンティティには対話者の知識と興味の状態をアノテーションしている。また、昨今の大規模言語モデルやそれに基づく対話システムの進展を受け、それらの知識や計算能力を電卓や知識ベースといった複数の外部ツールを活用することで補完・拡張する手法を提案した。化学の専門知識を要する数値推論タスク NumGLUE Task2 に提案法を適用し、圧倒的な世界最高性能を達成した。成果論文はトップ国際会議 ACL2023 に採択された。加えて、対話システムの進展に伴い、その振る舞いを広く詳細に分析するには言語解析が依然重要と考え、汎用言語モデルに基づく日本語統合解析器 KWJA を開発した。KWJA は汎用言語モデルを基盤とすることで、形態素解析や構文解析といった基本的な解析から述語項構造解析や談話関係解析といった高度な解析まで、各解析タスクに特化した既存の解析器と遜色ない精度で実行する。成果論文はトップ国際会議 ACL2023 に System Demonstration Paper として採択された。

久木田グループは、昨年度に引き続き信頼を醸成する人工知能と人間のコミュニケーションの在り方についての人間社会科学的観点から研究を遂行した。今年度はフェイクニュースに関連した

人々の心理と行動についての研究として、オンラインでのファクトチェック行動を調査し、分析した。この研究と並行して人工知能やロボットに対する人々の態度についてフィクション作品の分析を通じて研究を行い、それに基づいて人工知能が信頼されるためには人工知能をどのように発展・利用するのが良いかという規範的な側面についての研究を行った。またフェイクニュースに関連する、専門家と市民を招いたワークショップを開催した。

#### 【代表的な原著論文情報】

- 1) Daisuke Bekki, Ribeka Tanaka, and Yuta Takahashi (2022) Learning Knowledge with Neural DTS. In Proceedings of the 3rd Natural Logic Meets Machine Learning (NALOMA III), pp.17-25, Galway, Ireland, Association of Computational Linguistics.
- 2) Qin Dai\*, Benjamin Heinzerling\*, Kentaro Inui (\*equal contributions). Cross-stitching Text and Knowledge Graph Encoders for Distantly Supervised Relation Extraction. In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022), pp.6947–6958, Dec 2022.
- 3) Izumi Haruta, Koji Mineshima, and Daisuke Bekki (2022) Implementing Natural Language Inference for Comparatives. Journal of Language Modelling, Volume 10, No.1, pp.139-191.
- 4) Tatsuro Inaba, Hirokazu Kiyomaru, Fei Cheng, and Sadao Kurohashi. GPT-3 Can Leverage Multiple External Tools for Reasoning. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL), July 2023.
- 5) Keito Kudo, Yoichi Aoki, Tatsuki Kuribayashi, Ana Brassard, Masashi Yoshikawa, Keisuke Sakaguchi and Kentaro Inui. Do Deep Neural Networks Capture Compositionality in Arithmetic Reasoning? In Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics (EACL2023), pp.1343-1354, May 2023.
- 6) 久木田水生、「AI の ELSI と研究倫理」、『電子情報通信学会誌』、Vol. 106 No. 3, pp. 203-206、2023 年 3 月。
- 7) 久木田水生、「人工知能とリスク分析文化」、『Nextcom』、50 号、13-21、2022 年 6 月。
- 8) Nobuhiro Ueda, Kazumasa Omura, Takashi Kodama, Hirokazu Kiyomaru, Yugo Murawaki, Daisuke Kawahara and Sadao Kurohashi. KWJA: A Unified Japanese Analyzer Based on Foundation Models. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL): System Demonstrations, July 2023.