

信頼される AI の基盤技術
2021 年度採択研究代表者

2022 年度
年次報告書

谷中 瞳

東京大学 大学院情報理工学系研究科
講師

意思決定を支援する言語と非言語の論理関係認識

研究成果の概要

第二年次では、研究項目 1「言語・非言語データの深い意味解析システムの構築」と研究項目 2「推論過程が解釈可能な、言語と非言語の論理関係認識システムの構築」の成果として、(i)表とテキスト間の数量関係を高速に推論できる意味解析・推論システムの構築、(ii)テキスト間の時間関係を頑健に推論できる意味解析・推論システムの構築、(iii) 画像生成モデルを用いて言い換えフレーズの知識を論理推論システムに補完する手法の 3 点に取り組み、自然言語処理のトップ国際会議 ACL2022、計算言語学の国際会議 LENLS2022 で発表を行なった。また、研究項目 3「推論の信頼性を保証する評価システムの構築」の成果として、(iv)日本語の構成的な推論データセットの構築に取り組み、自然言語処理のトップジャーナル誌 TAACL2022 に採択された。

(i)では、Wikipedia に含まれる半構造化表データを非言語データの対象として、表とテキストの意味を論理表現に変換し、表中の情報に対して文の真偽をモデル検査で判定する言語・非言語データの意味解析・推論システムを提案し、既存の深層学習の手法と比較して正確かつ高速に数量関係を捉えられることを確認した。(ii)では、日本語のテキストの時間情報を論理式に自動変換し、テキスト間の時間関係を頑健に推論できる意味解析・推論システムを提案し、既存の日本語の時間推論ベンチマークで最高精度を達成した。(iii)では、テキスト間の推論に不足する言い換えフレーズの知識を補完する手法として、テキストから画像を生成する深層学習のモデルを用いて必要な知識を特定し補完する手法を提案し、既存の論理推論システムに対する有用性を示した。(iv)では、語彙的な意味だけでなく能動態・受動態などの統語的な意味、否定・量化などの論理的な意味との構成性を考慮した日本語の含意関係認識と類似度計算のベンチマーク JSICK を構築し、現在の深層学習に基づく推論システムが日本語の柔軟な語順を捉えられていないことを明らかにした。

【代表的な原著論文情報】

- 1) Tomoya Kurosawa, Hitomi Yanaka, Logical Inference for Counting on Semi-structured Tables, Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop (ACL), pp.84-96, 2022.
- 2) Tomoki Sugimoto, Hitomi Yanaka, Compositional Semantics and Inference System for Temporal Order based on Japanese CCG, Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop (ACL), pp.104-114, 2022.
- 3) Akiyoshi Tomihari and Hitomi Yanaka, Logical Inference System with Text-to-Image Generation for Phrase Abduction, Proceedings of the 19th International Workshop on Logic and Engineering of Natural Language Semantics (LENLS), 2022.
- 4) Hitomi Yanaka and Koji Mineshima, Compositional Evaluation on Japanese Textual Entailment and Similarity, Transactions of the Association for Computational Linguistics (TAACL), Volume 10, pp.1266-1284, 2022.