

2023 年度年次報告書
信頼される AI の基盤技術
2021 年度採択研究代表者

谷中 瞳

東京大学 大学院情報理工学系研究科
准教授

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

第三年次では、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、シーン画像と、指差し情報と、「あの赤いものの隣に止まって」といった指差し情報に関連した言語指示に対して、指示が示す場所を特定するマルチモーダル指示理解タスクを提案した。さらに、そのタスクを解くためのシステムとして、論理と深層学習のアプローチを融合した Neural State Machine モデルを応用して、入力をグラフ表現に変換し、推論時にマルチモーダル情報間の相互作用の過程が解釈可能なマルチモーダル指示理解システムを構築した。本研究は株式会社本田技術研究所との共同研究の成果の一つであり、人工知能分野のジャーナル New Generation Computing に採択された。また、Vision-and-language モデルの埋め込み表現を用いてフレーズ間の含意関係や言い換え知識を推定し、文間の論理推論を行うシステムを構築した。本研究は情報科学分野のジャーナル IEEE Access に採択された。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、医用画像中の数量関係に関する文の真偽を判定する推論ベンチマーク MedVTE を構築し、現在の Vision-and-language モデルは医用画像中の数量関係の理解に課題があることを示した。本研究は自然言語処理のトップ国際会議 ACL2023 の医療系ワークショップ ClinicalNLP2023 に採択された。また、形式意味論と言語処理の知見を組み合わせることで日本語の大規模時間推論データセット JAMP を構築し、現在の大規模言語モデルが時間関係を十分に捉えられていないことを示した。本研究は ACL2023 Student Research Workshop に採択された。

【代表的な原著論文情報】

- 1) Aman Jain, Anirudh Reddy Kondapally, Kentaro Yamada, Hitomi Yanaka, Neuro-symbolic Reasoning for Multimodal Referring Expression Comprehension in HMI systems, <https://doi.org/10.1007/s00354-024-00243-8>, New Generation Computing, 2024.
- 2) Akiyoshi Tomihari and Hitomi Yanaka, Logic-based Inference with Phrase Abduction using Vision-and-Language Models, IEEE Access, pp.45645 - 45656, 2023.
- 3) Hitomi Yanaka, Yuta Nakamura, Yuki Chida, Tomoya Kurosawa, Medical Visual Textual Entailment for Numerical Understanding of Vision-and-Language Models, Proceedings of the 5th Clinical Natural Language Processing Workshop at ACL2023, pp.8-18, 2023.
- 4) Tomoki Sugimoto, Yasumasa Onoe, Hitomi Yanaka, JAMP: Controlled Japanese Temporal Inference Dataset for Evaluating Generalization Capacity of Language Models, Proceedings of the ACL2023 Student Research Workshop (SRW), pp. 57-68, 2023.
- 5) 谷中瞳, 峯島宏次, AI は言語の基盤を獲得するか: 推論の体系性の観点から, 認知科学, Vol.31 No.1, pp.27-45. 2024.