

信頼される AI の基盤技術
2021 年度採択研究者

2021 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／理化学研究所 革新知能統合研究センター
講師／客員研究員

意思決定を支援する言語と非言語の論理関係認識

§ 1. 研究成果の概要

第一年次では、テキスト情報の意味表現の設計とパイロットシステムの構築に取り組み、国際ワークショップ 2 件 (LACompLing2021、LENLS2021) にて発表を行った。

一点目は、言語データを対象として、形式意味論のテンス・アスペクト解釈を考慮した意味表現の設計を行った。例えば、「私は解熱剤を投与する」という文では「投与する」というイベントの目的格である「解熱剤」が不可算名詞であるためイベントの到達点がないのに対して、「私は 3 回解熱剤を投与する」という文では 3 回解熱剤を投与した時がイベントの到達点となる。イベント間の時間関係を正確に推論するためには、前述のような文中のイベントの到達点の有無を特定し、意味表現に表す必要がある。この問題に対して、深層学習の分類モデルを用いて分布意味論に基づき文中のイベントの到達点の有無を特定し、記号論理を用いてイベント意味論に基づいた意味表現を導出する手法を提案し、計算言語学の国際ワークショップ LACompLing2021 で発表を行った¹⁾。

二点目は、時系列を考慮した高度な症例検索を実現するためには、症例テキスト中の時間関係認識を含む高度な意味解析が求められる。しかし、症例テキストには「解熱剤投与」「術後 3 週間以内」といった多様な意味関係から構成される複合語が含まれており、これらの複合語の意味をどのように解析するかが課題であった。そこで本研究では、症例テキスト中の複合語について正しい意味表示を導出可能にする手法として、複合語解析モデルを構築し、形式意味論の国際ワークショップ LENLS2021 で発表した²⁾。具体的にはまず、症例テキスト中の複合語の意味関係をアノテーションしたデータセットを構築した。次に、データセットを用いて深層学習に基づく複合語意味関係解析器を構築し、意味関係解析器と記号論理に基づく意味解析からなる複合語解析モデルを構築することで、症例テキスト中の複合語について構成的に意味表示を導出することができた。

【代表的な原著論文情報】

- 1) [Hitomi Yanaka](#), Towards Compositional Semantics and Inference System for Telicity, Proceedings of Logic and Algorithms in Computational Linguistics 2021 (LACompLing2021), online, 2021.
- 2) Mana Ishida, [Hitomi Yanaka](#), Daisuke Bekki, Compositional Semantics for Compound Words in Medical Case Retrieval, Proceedings of the 18th International Workshop on Logic and Engineering of Natural Language Semantics (LENLS18), 2021.
- 3) 小谷野華那, [谷中瞳](#), 峯島宏次, 福田浩司, 橋爪宏典, 戸次大介. 供述調書に現れる数量表現の推論テストセットの構築. 言語処理学会第 28 回年次大会, 2022.