

信頼される AI の基盤技術
2020 年度採択研究代表者

2022 年度
年次報告書

菅原 朔

情報・システム研究機構 国立情報学研究所
助教

説明性の高い自然言語理解ベンチマークの構築

研究成果の概要

2022年度は、おもに以下の2つの研究成果が得られた。

【1. 読解問題において文章が問題の性質に与える影響の分析】

読解問題の収集では一般的にひとつのデータセットで少数のドメインの文章が利用されるが、作問に用いる文章が結果として執筆される質問にどのような影響を与えるのかは明らかにされていなかった。これが不明瞭なままだと、望ましい難易度・能力について質問を作成する上で、どのような文章を用いればよいか自明ではなく、効率のよいデータ収集の妨げになる可能性があった。そこで複数の異なるソースの文章を利用し、文章の読みやすさとドメインごとの違いが収集される質問の難易度・要求される能力の違いにどのような変化をもたらすか分析した。結果として、文章の難易度は質問の難易度とほとんど相関しないこと、異なる文章のドメインを利用することで多様な声質の質問を集めることが可能なことが示唆された。本論文はニューヨーク大学との共同研究であり、自然言語処理のトップ国際会議のひとつである ACL 2022 に採録された。

【2. 文脈に依存した常識推論を評価するデータセットの作成】

既存の常識推論のデータセットは一般的な状況下でも成り立つ常識的知識を問うことが多く、文脈に依存するような柔軟な推論を評価することが十分ではなかった。そこで既存のデータセットにおける短い物語文を利用し、その結末となる文を複数パターン用意した。それぞれの結末が答えになるような質問文を作成することで、質問という文脈に依存しながらもすべての選択肢が正答になりうるようなタスクを設計した。およそ 4,500 問の質問を作成したところ、テストセットにおいて教師なしの状況下ではその時点でもっとも高性能なシステムでも 6 割ほどの正解率しか達成できなかった(人間の正解率は 9 割程度)。本論文は自然言語処理のトップ国際会議のひとつである COLING 2022 に採録された。

【代表的な原著論文情報】

- 1) Saku Sugawara, Nikita Nangia, Alex Warstadt, Samuel R. Bowman, “What Makes Reading Comprehension Questions Difficult?”, *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL 2022)*, pp.6951-6971, 2022.
- 2) Mana Ashida, Saku Sugawara, “Possible Stories: Evaluating Situated Commonsense Reasoning under Multiple Possible Scenarios”, *Proceedings of the 29th International Conference on Computational Linguistics (COLING 2022)*, pp.3606-3630, 2022.