

令和 6 年 6 月 19 日現在

機関番号：62615

研究種目：基盤研究(B)（一般）

研究期間：2021～2023

課題番号：21H03502

研究課題名（和文）複雑な作業を支援するための言語インタフェース技術

研究課題名（英文）Natural Language Interface Technology to Support Complex Tasks

研究代表者

相澤 彰子（Aizawa, Akiko）

国立情報学研究所・コンテンツ科学研究系・教授

研究者番号：90222447

交付決定額（研究期間全体）：（直接経費） 13,300,000円

研究成果の概要（和文）：本研究では、複雑な作業を言語でコンピュータに指示するための自然言語処理の研究に取り組む。ここでの「複雑な作業」とは、一問一答のQAだけでは解決できず、多段階の推論や対話的なやりとりが必要となる処理である。研究では、複数の文書からの情報を組み合わせて質問に答える多段階言語推論、自然言語文をSQLクエリや実行可能なプログラムに変換する意味解析、対話において複数の対話者が発話中で共通に参照する対象を同定する基盤化等の研究に取り組み、言語モデルの構築や評価のためのデータセットを構築し、モデルの訓練および分析評価を通して、有用性を示した。

研究成果の学術的意義や社会的意義

生成型の言語モデルの登場と急速な大規模化によって、言語モデルの推論能力は劇的に向上している。このような言語モデルは、さまざまな場面における人間の作業支援ツールとして大きな期待が寄せられる一方で、実際の環境において必要となる複雑な指示の理解やコマンド生成などの能力については、さらなる改良が求められている。また、複数の情報ソースに分散する情報をさがしあてて、質問への正確な回答を生成する処理は、ハルシネーション対策としても重要である。本研究では、複雑な推論タスクに関する複数のデータセットを構築して公開しており、今後の研究開発における言語モデルの訓練・評価に貢献するものである。

研究成果の概要（英文）：In this study, we addressed natural language processing for instructing computers to perform complex tasks in natural language. "Complex tasks" in this study are those processes that cannot be solved by solely using single-question answering, but require multi-step reasoning and conversational interactions. In the research period, we worked on multi-step machine reading comprehension, which combines information from multiple source documents to answer a single question; semantic parsing, which converts natural language sentences into SQL queries or executable programs; and mutual grounding, which identifies entities commonly referred to by multiple interlocutors in their utterances. We have constructed datasets for the construction and evaluation of language models, and demonstrated their usefulness through analytical evaluation of the models.

研究分野：研究成果の学術的意義や社会的意義

キーワード：言語モデル 多段階推論 対話システム 意味解析 基盤化

1. 研究開始当初の背景

(1) 近年では、モバイル端末やスマートスピーカーへの音声入力が一般的になり、情報家電や家庭用ロボットに対する会話での指示や、チャットボットによるユーザサポートも普及してきた。しかしながら、ユーザが言葉を使って指示できるのは、キーワード検索やスイッチ ON/OFF などの単純なコマンドに限られていて、これらを日常的な業務の場面に適用することはいまだ難しい。このような複雑な作業指示が困難である原因として、「ユーザ要求のあいまいさ」が考えられる。AI システムは、あいまいさを含むユーザの指示を、常識的な知識や推論などで補いながら理解し、適切に実行する必要がある。

(2) 上記の課題解決のため本研究では、ユーザの指示意図をシステムの実行コマンドに対応付けるための対話技術の研究に取り組む。対話を介して指示を与えることによって、一方的に人間が指示を与えるのではなく、実現可能なオプションを調べつつ、目的達成に向けて折り合いをつけて指示を明確化したり、人間がシステムの指示に従って作業したりすることが可能になる。

2. 研究の目的

(1) 本研究における「複雑な作業」とは、単純な質問応答だけでは解決できず、多段階の推論や対話的なやりとりが必要となる処理である。本研究では作業指示の問題を、「エンティティ参照の基盤化」と「作業手順の理解」の2つのレベルに整理して検討に取り組んだ。また、これらの基盤化を支えるための「共通知識の獲得」について研究を進めた。

3. 研究の方法

(1) 上記の3つの課題について、言語モデルの構築および評価のためのデータセットを構築するとともに、モデルの訓練および分析評価を通して有用性を示した。

(2) 「エンティティ参照の基盤化」について、対話において複数の対話者が発話中で共通に参照する対象を同定するための技術に焦点をあてて、時間変化を伴う環境でのタスク設計とコーパス構築を行い、これに基づき時間表現の言語解析およびベースラインとなる対話モデルの構築に取り組んだ (Udagawa et al., 2021)。

(3) 「作業手順の理解」について、発話内容をコンピュータで実行可能なコマンド列に変換するための技術に焦点を当てて、自然言語文を SQL クエリや実行可能なプログラムに変換する意味解析に関する研究に取り組んだ。特に、自然言語文中のエンティティとデータベースの構成要素 (項目名や値など) の対応付け手法の開発に取り組み、未知のデータベースにも適応可能な汎用的な解析手法の開発を進めた。そして、既存の SPIDER データセットを拡張して有効性を示した (Julavanich et al., 2022)。

(4) 「作業手順の理解」について、複数の文書からの情報を組み合わせて質問に答える多段階推論に注目して研究に取り組んだ。明示的な推論パスを付与したデータセットを構築して、推論パスを用いた補助タスクが言語モデルの性能、頑健性、汎化能力に及ぼす影響を調べるためのデータセットを構築して、代表的な多段階推論モデルの詳細な分析を行った (Ho et al., 2022)。また、言語モデルの運用において問題となるショートカットの回避および回答への説明出力の手法を提案した (Shinoda et al., 2021, Shinoda et al., 2022)。さらに、バイアスに寄与する重みの影響を取り除くための新しいバイアス除去手法を提案して有効性を示した (Meissner et al., 2022)。

(5) 「作業手順の理解」について、ブラウザ上に表示される自然言語タスク等に回答するため視覚・言語・アクションを統合した新たな言語モデルを提案して、ブラウザ上でのフォームへの入力やマウスクリック等を含むタスクを通して有効性を検証した (Iki et al., 2022)。

(6) 「共通知識の獲得」について、k 近傍言語モデルを用いて自然言語文中に出現する数字の埋め込み表現を表す手法を提案するとともに、数字の予測タスクのデータセットを構築して提案手法の有効性を示した (Sakamoto et al., 2021)。また、数字の埋め込み表現の応用として、文中の数字の誤り検出タスクを設計・構築した。

(7) 「共通知識の獲得」について、「多くの場合に成り立つ常識的な知識」と「特定の状況において与えられた情報から導かれる論理的な帰結」の2通りがあることに注目し、両者が矛盾する例

を含むデータセットを構築した。これにより、両者の推論方法の使い分けが高度な推論タスクにおいて重要であり、現状の言語モデルではまだ十分なレベルに到達していないことを示した (Kondo et al., 2023)。

4. 研究成果

(1) 生成型の言語モデルの登場と急速な大規模化によって、言語モデルの推論能力は劇的に向上している。このような言語モデルは、さまざまな場面における人間の作業支援ツールとして大きな期待が寄せられる一方で、実際の環境において必要となる複雑な指示の理解やコマンド生成などの能力については、さらなる改良が求められている。また、複数の情報ソースに分散する情報をさがしあてて、質問への正確な回答を生成する処理は、ハルシネーション対策としても重要である。

(2) 本研究では、特に複雑な指示文の理解を必要とする推論タスクに焦点をあてて複数のデータセットを構築して公開した。本研究で対象としたのは、言語モデルの性能が著しく向上する中でも、依然として人間には及ばない推論・常識タスクであり、本研究の成果は、今後の言語モデルの研究に貢献すると期待される。

5. 主な発表論文等

〔雑誌論文〕 計11件（うち査読付論文 10件 / うち国際共著 0件 / うちオープンアクセス 10件）

1. 著者名 Xanh Ho, Saku Sugawara, Akiko Aizawa	4. 巻 -
2. 論文標題 How Well Do Multi-hop Reading Comprehension Models Understand Date Information?	5. 発行年 2022年
3. 雑誌名 Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Short Papers)	6. 最初と最後の頁 470-479
掲載論文のDOI (デジタルオブジェクト識別子) 10.48550/arXiv.2210.05208	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Johannes Mario Meissner, Saku Sugawara, Akiko Aizawa	4. 巻 -
2. 論文標題 Debiasing Masks: A New Framework for Shortcut Mitigation in NLU	5. 発行年 2022年
3. 雑誌名 Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing	6. 最初と最後の頁 7607-7613
掲載論文のDOI (デジタルオブジェクト識別子) 10.18653/v1/2022.emnlp-main.517	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Thanakrit Julavanich and Akiko Aizawa	4. 巻 -
2. 論文標題 Measuring Text-to-SQL Semantic Parsing Model on the Question Generalizability	5. 発行年 2022年
3. 雑誌名 Proceedings of the 6th International Conference on Natural Language Processing and Information Retrieval (NLP4IR 2022)	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3582768.3582782	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Taichi Iki and Akiko Aizawa	4. 巻 -
2. 論文標題 Do BERTs Learn to Use Browser User Interface? Exploring Multi-Step Tasks with Unified Vision-and-Language BERTs	5. 発行年 2022年
3. 雑誌名 arXiv: 2203.07828	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.48550/arXiv.2203.07828	査読の有無 無
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Kazutoshi Shinoda, Saku Sugawara, and Akiko Aizawa	4. 巻 -
2. 論文標題 Penalizing Confident Predictions on Largely Perturbed Inputs Does Not Improve Out-of-Distribution Generalization in Question Answering	5. 発行年 2022年
3. 雑誌名 The AAAI-23 Workshop on Knowledge Augmented Methods for NLP (KnowledgeNLP)	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.48550/arXiv.2211.16093	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Kazutoshi Shinoda, Saku Sugawara, and Akiko Aizawa	4. 巻 -
2. 論文標題 Which Shortcut Solution Do Question Answering Models Prefer to Learn?	5. 発行年 2022年
3. 雑誌名 Proceedings of the 37th AAAI conference on Artificial Intelligence (AAAI)	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.48550/arXiv.2211.16220	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Taichi Iki, Akiko Aizawa	4. 巻 -
2. 論文標題 Effect of Visual Extensions on Natural Language Understanding in Vision-and-Language Models.	5. 発行年 2021年
3. 雑誌名 The 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP 2021)	6. 最初と最後の頁 2189-2196
掲載論文のDOI (デジタルオブジェクト識別子) 10.18653/v1/2021.emnlp-main.167	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Takuma Udagawa, Akiko Aizawa	4. 巻 9
2. 論文標題 Maintaining Common Ground in Dynamic Environments.	5. 発行年 2021年
3. 雑誌名 Transactions of the Association for Computational Linguistics	6. 最初と最後の頁 995-1011
掲載論文のDOI (デジタルオブジェクト識別子) 10.1162/tacl_a_00409	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Kazutoshi Shinoda, Saku Sugawara, Akiko Aizawa	4. 巻 -
2. 論文標題 Can Question Generation Debias Question Answering Models? A Case Study on Question-Context Lexical Overlap.	5. 発行年 2021年
3. 雑誌名 The 3rd Workshop on Machine Reading for Question Answering (MRQA), at the 2021 conference on Empirical Methods in Natural Language Processing (EMNLP)	6. 最初と最後の頁 63-72
掲載論文のDOI (デジタルオブジェクト識別子) 10.18653/v1/2021.mrqa-1.6	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Taku Sakamoto, Akiko Aizawa	4. 巻 -
2. 論文標題 Predicting Numerals in Natural Language Text Using a Language Model Considering the Quantitative Aspects of Numerals.	5. 発行年 2021年
3. 雑誌名 The Second Workshop on Knowledge Extraction and Integration for Deep Learning Architectures (DeeLIO), collocated with NAACL 2021	6. 最初と最後の頁 140-150
掲載論文のDOI (デジタルオブジェクト識別子) 10.18653/v1/2021.deelio-1.14	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Kazushi Kondo, Saku Sugawara, and Akiko Aizawa	4. 巻 -
2. 論文標題 Probing Physical Reasoning with Counter-Commonsense Context	5. 発行年 2023年
3. 雑誌名 The 61st Annual Meeting of the Association for Computational Linguistics (ACL),	6. 最初と最後の頁 603-612
掲載論文のDOI (デジタルオブジェクト識別子) 10.5715/jnlp.30.1261	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計3件 (うち招待講演 0件 / うち国際学会 0件)

1. 発表者名 阪本 拓功, 相澤 彰子
2. 発表標題 k近傍言語モデルを用いたテキスト中の数字の予測
3. 学会等名 人工知能学会 第36回全国大会 (JSAI2022)
4. 発表年 2022年

1. 発表者名 篠田一聡, 菅原朔, 相澤彰子
2. 発表標題 抽出型質問応答における相対位置バイアスの除去
3. 学会等名 言語処理学会第28回年次大会(NLP2022)
4. 発表年 2022年

1. 発表者名 鈴木淳平, 菅原朔, 相澤彰子
2. 発表標題 言語モデルに対するトークンのキャンセルアウト手法の比較
3. 学会等名 言語処理学会第28回年次大会(NLP2022)
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------