

研究報告書

「大規模テキストからの知識獲得と深層学習・意味解析による知識活用」

研究期間：2019年4月～2021年3月
研究者番号：50222
研究者：栗田 修平

1. 研究のねらい

自然言語による文章を解析する手法は、深層学習の採用によりその精度が大きく向上した。しかし、自然言語の意味を解析し、現実世界のなかで実際に人間の言語指示等に従わせるシステムを作成していくためには、単純に深層学習を限られた学習コーパスに利用するだけでなく、学習コーパスからは得られないような多様な知識を、モデルが様々なリソースから学習し処理できるようになることが必要であるだろう。これまでの自然言語処理では、専ら人手で作成されたラベル付きコーパスに頼ってモデル学習が行われていた。このようなデータセットは、一つ一つの自然言語文や文章に対して非常に詳細な情報を付与している一方で、作成には膨大なコストがかかり、利用可能なデータサイズも限られている。より高度な構文・意味解析のためには、本質的により大規模なデータから情報や知識を取る必要があると考えられる。ACT-I 期間中には日本語の省略解析に対し大規模なラベルなしコーパスを学習に用いる手法で省略解析に対する世界最高性能を達成した。加速フェーズでは、これらの自然言語文解析手法を利用し、大規模なコーパスや外界知識源から既存のコーパスには存在しない知識をモデルに学習させることを目指す。また、これら大規模コーパスや知識ベース等から学習される知識に加えて、自然言語処理をなるべく現実世界に近い環境で学習・評価を行う手法を模索する。ロボット等への将来的な応用を意識し、シミュレーション環境中でのエージェントに対して、自然言語文で指示を与え、深層強化学習等により学習させることを研究提案し、自然言語を用いた解析モデルにはどれだけ知的なことが可能であるか調査する。この目的のため以下の課題に具体的に取り組む。

- [1] 大規模コーパスからの知識獲得による意味解析の精緻化
- [2] Wikipedia 等からの知識を利用した新しい事前学習手法
- [3] 強化学習と自然言語文によるシミュレーション環境下での報酬付け

これらのうち、[1]と[2]は互いに深く関連し合うテーマであり、事前学習モデルを活用して双方の課題に取り組むとともに、[3]の課題では、得られた自然言語の意味理解手法を、現実世界に近い条件での課題に適用することを目的に、自然言語文の解析モデルとシミュレーション環境下での視覚や動作のモデルとを統合したニューラルネットワークモデルの作成を行う。

2. 研究成果

(1) 概要

テーマ[1] 大規模コーパスからの知識獲得による意味解析の精緻化では ACT-I 期間中の省略解析等の研究成果に引き続き、自然言語文章の高精度な解析システムを作成することを目指した。既存の限られたアノテーション付きコーパスには出現しない文の解析精

度を高めるために、Wikipedia 等の大規模なコーパスを利用したモデル学習を行う。これにより、ACT-I 期間に引き続いた研究テーマである英語の意味依存構造解析についてコペンハーゲン大学 Anders Søgaard 先生と共同で 56th Annual Meeting of the Association for Computational Linguistics (ACL2019)にて”Multi-Task Semantic Dependency Parsing with Policy Gradient for Learning Easy-First Strategies”(研究成果リスト[1])の発表を行った。加えて、英語の意味役割ラベル付与に対し事前学習済み手法および強化学習の応用を検討し、自然言語処理の国際会議に投稿中である。

テーマ[2] Wikipedia 等からの知識を利用した新しい事前学習手法では、Wikipedia 等の知識を処理して学習に用いることで、あらかじめ基本的な知識が備わったモデルを作成し、別のタスクに応用することを検討した。この研究は筑波大学・理化学研究所の中山功太君と共同で、複数のモデルを Wikipedia から作られた少量のアノテーション付きコーパスでタスク別学習させ、得られた複数のモデルを教師として用いてより大規模なコーパスを処理することで、精度と大規模なコーパスからの知識獲得の両方を行う手法を提案した。この手法については、国内学会での発表を行い、国際会議への投稿準備を進めている。加速フェーズ期間の後半からテーマ[3] 強化学習と自然言語文によるシミュレーション環境下での報酬付けに取り組んだ。このテーマでは人間に与えられた指示に従って、実際の家屋環境を模した 3D シミュレータの中を移動する視覚と言語によるナビゲーション課題について研究を進め、学習済みの条件付き言語モデルを利用して視覚や動作から指示文章を逆に予測する手法で、人間の指示に従ったナビゲーションを可能にするとともに、個別の動作選択と言語指示との対応付けを明らかにする可視化手法を提案した。この研究はニューヨーク大学 Kyunghyun Cho 先生と共同で行われた。研究成果は 37th ICML Workshop on Learning in Artificial Open Worlds (ICML2020 LAOW)にて部分的な発表を行った後に、”Generative Language-Grounded Policy in Vision-and-Language Navigation with Bayes’ Rule,”として 9th International Conference on Learning Representations (ICLR2021) (研究成果リスト[2])に採択され、発表を予定している。

これらの成果に加えて、ACT-I の小林亮太先生と進めていた共同研究である”Reconstructing neuronal circuitry from parallel spike trains”について雑誌 Nature Communications に採択された。(研究成果リスト[2])

(2) 詳細

研究テーマ[1] 大規模コーパスからの知識獲得による意味解析の精緻化

本研究テーマでは、ACT-I 期間に引き続いた研究テーマである英語の意味依存構造解析についてコペンハーゲン大学 Anders Søgaard 先生と共同で 56th Annual Meeting of the Association for Computational Linguistics (ACL2019)にて”Multi-Task Semantic Dependency Parsing with Policy Gradient for Learning Easy-First Strategies”の発表を行った。加えて、英語における意味役割付与研究について、大規模な事前学習済みモデルを深層強化学習と組み合わせて解析するシステムについて研究を進めた。本システムについては、現在、自然言語処理の国際会議に投稿中である。

研究テーマ[2] Wikipedia 等からの知識を利用した新しい事前学習手法

本研究テーマでは、大規模な事前学習モデルを作成及び応用する箇所については研究テーマ[1]と同様に扱うとともに、理化学研究所関根先生及び筑波大学・理化学研究所の中山功太君と共同で、固有表現抽出タスクに対し、複数のモデルを Wikipedia から作られた少量のアノテーション付きコーパスでタスク別学習させ、得られた複数のモデルを教師として使用することで、より大規模なコーパスを処理し、大規模な疑似教師コーパスを作成する手法を提案した。これによりモデルの解析精度の向上と大規模なコーパスからの知識獲得の両方を同時に進めることが可能となった。この手法については、国内学会での発表を行い、現在、国際会議への投稿準備を進めている。

研究テーマ[3] 強化学習と自然言語文によるシミュレーション環境下での報酬付け

本課題分野は、言語の意味を理解するモデルをより現実に近い環境で評価するものであり、将来的には、視覚と言語の統合モデルや、ロボティクスなどへの応用が見込まれる、国内外で近年とても大きな注目を集めている分野である。本研究テーマでは、実際の家屋環境を模した 3D シミュレータの中を利用して、仮想のエージェントを言語によって指示された場所まで移動させる視覚と言語によるナビゲーション課題について研究を進めた。事前学習された条件付き言語モデルを動作選択に使用し、言語情報と視覚や動作情報を直接結びつける新しい手法を提案した。このように画像からの条件付き言語モデルをナビゲーションに直接使用する手法は、国内外に類を見ない新しいアプローチである(図1)。また、本手法を取るメリットとして、与えられた言語入力と、モデルの動作選択との対応関係を可視化することができることが示された(図2)。

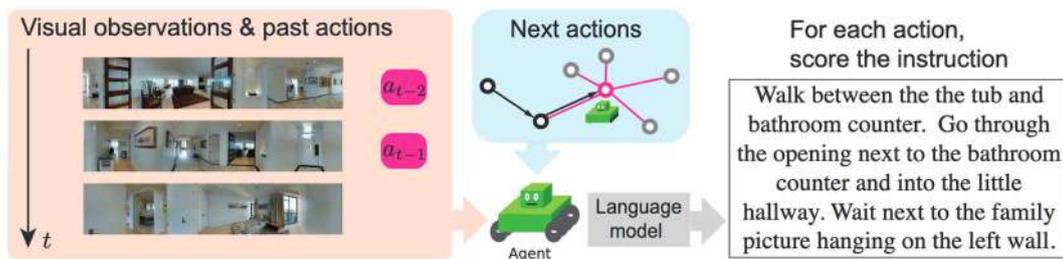


図1:事前学習された条件付き言語モデルをモデル学習に使用する手法。既存手法では、視覚情報と言語の対応関係が必ずしも明確ではなかったが、提案手法では、明示的に入力画像を自然言語出力へと対応付けるように学習し、自然言語の予測スコアを動作選択に使用する。

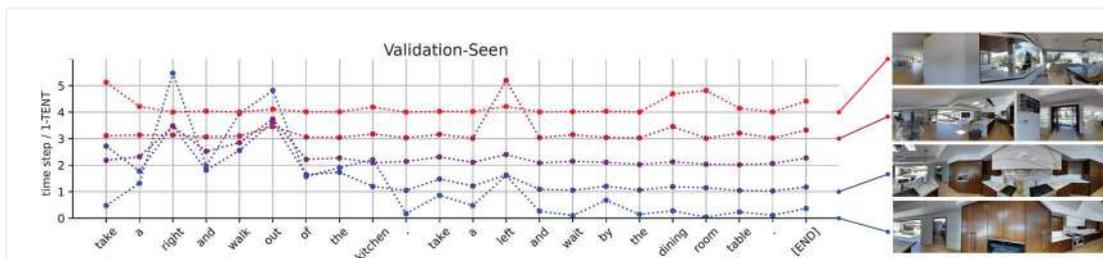


図2:トークンごとの予測エントロピーの可視化。各時刻(縦軸)にて、次の動作を決める上で大きな予測スコアを持つ指示文章中のトークン(単語)を可視化している。グラフの振幅が大きいトークンほど、その時刻の動作選択に大きな影響を持っていることを示している。

本研究成果は 37th ICML Workshop on Learning in Artificial Open Worlds (ICML2020 LAOW)にて部分的な発表を行った後に、” Generative Language-Grounded Policy in Vision-and-Language Navigation with Bayes’ Rule,”として 9th International Conference on Learning Representations (ICLR2021)に採択され、発表を予定している。

3. 今後の展開

ACT-I 加速フェーズ研究期間では、大規模なコーパスから知識を獲得し、自然言語の理解に役立てるシステムを研究すると同時に、自然言語を理解するための手法を現実世界での応用を意識した課題に応用することを目指して研究を進めた。特に研究期間の後半には、自然言語を理解するシステムについて、現実世界に近い仮想環境の中で外部から与えられた指示に従わせる手法について研究を行った。この分野は、将来的に視覚と言語の統合モデルや、ロボティクスなどへの応用が見込まれ、国内外で、近年、トップ国際会議でデータセット整備やモデルの提案が活発に行われ、ロボティクスへの応用も試行されるなど、とても大きな注目を集めている。今後は、自然言語を理解するシステムの長所を生かして、人間の持つ日常的な常識を理解して行動するシステムや、より長期的な時系列での自然言語の意味での計画立案手法、および画像認識技術と組み合わせること、ロボティクスへの応用を視野に入れたシステムの作成に向けて取り組んでいく予定である。

4. 自己評価

ACT-I での研究目的について、特に研究テーマ [3] 強化学習と自然言語文によるシミュレーション環境下での報酬付けでは、画像による条件付き言語モデルを動作選択に使用し、言語情報と視覚や動作情報を直接結びつけるという、国内外に類を見ない新しい手法を提案する等、当初の予定を上回る大きな進展があったと認識している。また、研究テーマ[1]および[2]にても、ACT-I 研究時に得られた成果や加速フェーズ研究時に得られた成果の国際会議での発表およびその準備を進めている。研究実施体制については、特に期間の後半にて、新型コロナウイルスの国際的な蔓延により、当初の予定よりも海外ニューヨーク大学訪問予定を切り上げて帰国する必要が生じるなどの影響があったが、概ね、計画通りに進展したものと認識している。今後は研究テーマ[1]および[2]で得られた成果を研究テーマ[3]のように現実世界に近い条件下で評価するとともに、画像認識やロボティクス技術とのコラボレーションなど、より社会的な実装に近い領域で、ACT-I で得られた自然言語処理技術を応用していくことが見込まれる。その上で、

ACT-I 期間での研究は、自然言語の意味理解に関する基礎研究と、画像やロボティクへの応用研究への橋渡しとなる非常に独創的な位置づけの研究課題となったと考えている。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

1. Shuhei Kurita and Anders Søgaard, Multi-Task Semantic Dependency Parsing with Policy Gradient for Learning Easy-First Strategies, Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL2019), 2019, pages 2420–2430, Long paper.
2. Ryota Kobayashi, Shuhei Kurita, Anno Kurth, Katsunori Kitano, Kenji Mizuseki, Markus Diesmann, Barry J. Richmond & Shigeru Shinomoto, Reconstructing neuronal circuitry from parallel spike trains, Nature Communications, 2019, Volume 10, Article number: 4468.
3. Shuhei Kurita and Kyunghyun Cho, Generative Language-Grounded Policy in Vision-and-Language Navigation with Bayes' Rule, Ninth International Conference on Learning Representations (ICLR2021), May 2021. To appear.

(2) 特許出願

研究期間累積件数:0 件

(3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

1. 中山功太, 栗田修平, 小林暁雄, 関根聡. Pre-Distillation Ensemble:リソース構築タスクのためのアンサンブル手法. 言語処理学会第 26 回年次大会, オンライン, (2020.3.16–19).
2. Shuhei Kurita and Kyunghyun Cho, Toward Understanding Language-Grounded Agents in Vision-and-Language Navigation, 37th International Conference on Machine Learning Workshop on Learning in Artificial Open Worlds (ICML2020 LAOW), Online (July 2020).
3. 栗田修平, Kyunghyun Cho. 視覚と言語によるナビゲーション課題への言語に対応付けられた生成的な方策. 言語処理学会第 27 回年次大会, (2021.3.15–19) .
4. 中山功太, 栗田修平, 馬場雪乃, 関根聡. 能動的サンプリングを用いたリソース構築共有タスクにおける予測対象データ削減. 言語処理学会第 27 回年次大会, (2021.3.15–19).
5. 栗田修平, 河原大輔, 黒橋禎夫. ニューラルネットワークを利用した中国語の統合的な構文解析. 2019, Vol.26, No.1, pp. 231–258. (言語処理学会論文賞受賞および招待講演. 2020 年 3 月.)