

## 科学研究費助成事業 研究成果報告書

令和 4 年 6 月 10 日現在

機関番号：11301

研究種目：研究活動スタート支援

研究期間：2020～2021

課題番号：20K23314

研究課題名（和文）深層学習とシンボリックな推論を統合する基礎技術の開発と文章読解システムへの応用

研究課題名（英文）Development of a Method to Combine Deep Learning and Symbolic Reasoning and its Application to Machine Reading Comprehension

研究代表者

吉川 将司 (Yoshikawa, Masashi)

東北大学・タフ・サイバーフィジカルAI研究センター・助教

研究者番号：80883470

交付決定額（研究期間全体）：（直接経費） 1,900,000円

研究成果の概要（和文）：我々は、深層学習による自然言語処理を発展させる鍵は、その技術と記号推論を組み合わせることにあると考える。これにより、データ効率が良く、テキストの変異に頑健であり、推論過程の透明性を持ったシステムの開発できる。しかし、深層学習の鍵は誤差逆伝搬法であるため、離散的な記号関数は本来的に相性が悪い。本課題では、Gumbel-Softmaxトリックを応用し、このようなシステムの勾配による学習を可能にする。この手法を用い、深層学習ベースの自然言語処理における未解決問題である数量推論に取り組む。我々は深層言語モデルに電卓を組み込むことでこの問題に対処する。

研究成果の学術的意義や社会的意義

深層学習言語モデルの高い推論能力が喧伝される一方で、少し入力进行操作するだけで思わぬ挙動を示すなど脆弱性も分野内でよく知られている。記号推論の仕組みを外部知識として組み込むことで、深層モデルのブラックボックス性を軽減でき、推論過程の透明性のみならず、期待される推論過程を教え込むことでより頑健なシステムを構築することが可能になる。数量は言語中に頻出である一方で自然言語処理では見過ごされがちである。金融文章の解析など応用にも直結する課題であるが、本研究の取り組みにより数量処理に大きな前進をもたらすことが期待される。

研究成果の概要（英文）：We believe the key to advance further the state of deep learning (DL)-based natural language processing (NLP) is combining its technologies with those of symbolic reasoning. By the combination, it will be possible to develop an AI system that is (1) more data-efficient, and (2) more robust to variations of texts (e.g., domains), with (3) the more visible inference process essential to symbols. However, as a major key to the success of DL is end-to-end training using backpropagation, it is challenging to incorporate discrete symbolic function within a neural network. In this project, we propose a simple extension of the Gumbel-Softmax trick to overcome the problem and enable gradient-based learning of such a neuro-symbolic system.

Using the method, we tackle numerical reasoning, which is an unsolved problem in DL-based NLP. We approach this problem by incorporating an arithmetic calculator layer within a DL-based reasoning model.

研究分野：自然言語処理

キーワード：自然言語処理 深層学習 文章読解 機械学習 記号推論 数量推論

## 1. 研究開始当初の背景

深層学習の到来により自然言語処理システムがより複雑なタスクに応用可能となり、性能評価ベンチマーク上では人間に接近するような高い推論能力が報告されている。一方で、それらシステムが入力テキストから推論、解答を行う際に、人間の推論の仕方とは大きく異なった、非本質的な根拠(訓練セット中の事例のみに有効なパターン)に依拠した動作を示すことが報告されており、人間であれば無視できる操作を入力テキストに加えるだけで思わぬ誤答をしてしまうなどの問題(ショートカット学習; Gururangan et al., 2018)が報告され、推論をより頑健にするモデルの構造や学習法が模索されている。

推論問題の中でも言語モデルに最も難しいとされているものとして数量推論がある。例えば、数量推論を要する文章読解は「太郎がリンゴを3つ、花子がリンゴを5つ持っています。より多くのりんごを持っているのは誰ですか?」のようなものを問う課題であるが、文、文章の構造の解析、理解などの基礎的な能力と同時に数量に関する推論能力を要求する。つまり、無限の要素からなる数空間の構造の把握に、「 $3 + 2 = 5$ 」のような数値間の関係知識(四則演算など)を活用できるようにモデルを学習する必要があるが、言語の一般的な処理能力とともにこのような知識を獲得することは、有限個の訓練事例を用いた標準的な教師あり学習でアプローチすることは困難に思われる。実際に近年のWebスケールの大量のデータで事前訓練した巨大言語モデルを持ってしても打開できておらず(Brown et al., 2020)、数量推論用に特殊に訓練されたモデルでも未知の大きな数を処理することができない(Wallace et al., 2019, Kim et al., 2021)。また、言語モデルにこのような数量推論課題を解かせるように訓練したときに、上述のショートカット学習が起こっていることも指摘されており(Al-Negheimish et al., 2021)、多層に複雑なこの課題に対し、新たな技術的ブレークスルーが必要であるように思われる。

## 2. 研究の目的

我々は、上述の深層学習ベースの自然言語処理モデルの推論能力の問題について、データから帰納的に推論の方法を獲得する深層学習と、人間による対象の問題の構造に関する演繹的知識を実装した古典的な記号処理の技術を組み合わせることによって打開できると考える。後者の典型例としては数量推論(電卓)、論理推論(自動定理証明)、グラフ探索、データベースへのクエリなどが挙げられる。演繹的な知識を言語モデルに組み込むことにより、なんでもありの学習ではなく、推論の仕方を望むように制限できる上、推論に必要な知識をモデルの外部に括り出し学習不要とすることができ、より強力な言語処理システムを構築できると考える。また、深層モデルの内部で離散的な記号を扱うことは、深層学習の応用で大きな課題の1つであるモデルの解釈可能性向上にも大きな役割を果たす。

本研究課題では、これら記号推論モジュールを深層学習モデル内の1層として取り込み、誤差逆伝搬法でend-to-endにモデル全体を訓練することで、深層ネットワークが記号推論モジュールを活用するような学習を仕向けることを目指す。深層ネットワークは微分可能な関数の合成であり、誤差逆伝搬法(合成関数の微分則)でネットワーク全体を最適化していくことが学習の鍵であるため、離散的な処理を行う記号推論を層として含めて学習することは従来に不可能となり、これが技術的に解決すべき大きな課題となる。

基礎技術の開発は一般性のあるものを目指す、具体的な応用課題として数量推論に取り組み、深層言語モデルに電卓(文字列として数を入出力する四則演算関数)を層として与える。電卓をモデルの中間に含め、入力テキストをもとに電卓でどのような計算をしたいか判断させ、また電卓の計算結果を活用しながら入力の問題を解くように仕向けた。この仕組みにより、数量推論問題には電卓を使った解法を選択するように促し、また数に関する知識の大部分を深層モデルのパラメータ内部で表現する必要がなくなることから、学習時に未知であった巨大な数などに関する推論も可能になると期待する。

## 3. 研究の方法

記号処理の仕組みを深層学習モデルの中間層として取り込み、全体を誤差逆伝搬法で学習するためには、記号処理関数に対して何らかの方法で微分則を定義する必要がある。この点は大きな課題であり、既存のEstimate and Replace法(Jacovi et al., 2019)をもとにした手法など様々に検討したが、Gumbel-Softmax trick(Jang et al., 2016)を拡張することによって、学習可能性に関する理論的な保証つきの手法を構築することができた。

既存の数量推論タスクDROP(Dua et al., 2019)をもとにして、小規模高精度な評価用データをグループ研究者の人手で、また中規模な評価用データをクラウドソーシングで構築して実験した。これによって、記号処理関数をただ微分可能にし、深層学習モデル内部に組み込めたとしても、学習の設定によっては必ずしもモデルが記号処理(電卓)を活用した解法を学習から獲得するわけではないことがわかった。いかに電卓の活用をモデルに促すことができるか、より詳細な調査を行うため、問題の形式やデータセット全体の分布を制御可能である人工データを用いた実験に移行し、その上で電卓を活用した学習を行うための条件や、学習方法の調査を行った。

#### 4. 研究成果

人工データを用いた実験の結果、構築した電卓つき文章読解モデルは一部のタイプの問題に対し、期待される形で電卓を活用し、既存手法よりも高度な数量推論能力を発揮することが観察されたが、別の種類の問題に対しては、パラメータに対する正則化、事前学習などの微妙な調整を行なって初めて電卓を活用した解法を獲得することがわかった。いずれの問題に対しても、電卓つきモデルは学習時に未知であった巨大な数が入力文章中に含まれていたとしても高精度に解答可能となったが、これは既存のフルに深層学習ベースのモデルには不可能であったことである。この成果は、国内学会で発表し (NLP2022)、国際学会論文を執筆中である。

人工データを用いて言語モデルの能力を評価するという方法には様々な応用の可能性がある。メインの研究課題の解決は困難を極め、二転三転したが、このアイデアの応用として深層学習モデル一般に関する調査の研究に展開できた。上述の電卓つき文章読解モデルモデルの拡張を念頭に、一般の深層言語モデルが文処理を行う際に文の統語構造に従った処理を行っているかに関する調査や、また同モデルの多段言語推論の能力に関わる限界の調査や、それを改善する手法に関する研究を行っている。これらいずれの成果についても、国内学会で発表し (青木ら NLP2022, 松本ら NLP2022, 工藤ら JSAI2022)、国際学会投稿に向けて準備中である。

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 0件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 Ouchi Hiroki, Suzuki Jun, Kobayashi Sosuke, Yokoi Sho, Kuribayashi Tatsuki, Yoshikawa Masashi, Inui Kentaro	4. 巻 9
2. 論文標題 Instance-Based Neural Dependency Parsing	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 1493 ~ 1507
掲載論文のDOI（デジタルオブジェクト識別子） 10.1162/tacl_a_00439	査読の有無 無
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 -

〔学会発表〕 計6件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 吉川将司, 乾健太郎
2. 発表標題 離散記号処理に対する近似的な微分構造の考察と数量推論を要する文章読解問題への応用
3. 学会等名 言語処理学会第27回年次大会
4. 発表年 2021年

1. 発表者名 吉川将司, Benjamin Heinzerling, 乾健太郎
2. 発表標題 End-to-End 学習可能な記号処理層の検討と数量推論への応用における課題の分析
3. 学会等名 言語処理学会第28回年次大会
4. 発表年 2022年

1. 発表者名 青木洋一, 工藤慧音, Ana Brassard, 栗林樹生, 吉川将司, 乾健太郎
2. 発表標題 多段の数量推論タスクに対する適応的なモデルの振る舞いの検証
3. 学会等名 言語処理学会第28回年次大会
4. 発表年 2022年

1. 発表者名 松本悠太, 吉川将司, Benjamin Heinzerling, 乾健太郎
2. 発表標題 四則演算を用いた Transformer の再帰的構造把握能力の調査
3. 学会等名 言語処理学会第28回年次大会
4. 発表年 2022年

1. 発表者名 工藤慧音, 青木洋一, 吉川将司, 乾健太郎
2. 発表標題 数量推論モデルに対する人工データを活用した表層的な手がかりに頑健な評価法の検討
3. 学会等名 NLP若手の会第16回シンポジウム
4. 発表年 2021年

1. 発表者名 工藤慧音, 青木洋一, Ana Brassard, 栗林樹生, 吉川将司, 乾健太郎
2. 発表標題 深層学習モデルの数量推論能力の評価とメタ学習活用の試み
3. 学会等名 2022年度人工知能学会全国大会(第36回)
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
---------	---------