

研究報告書

「非テキストデータと接続可能なテキスト解析・推論技術の開発」

研究タイプ: 通常型

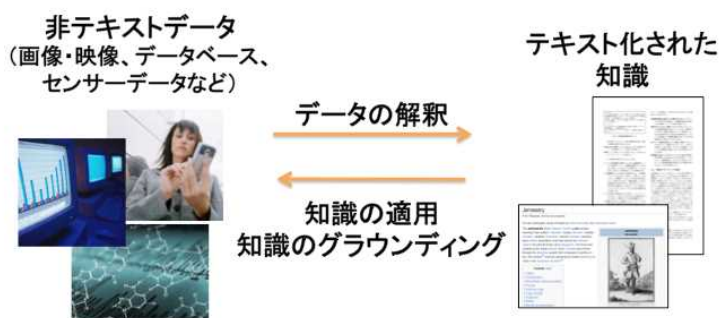
研究期間: 平成 25 年 10 月～平成 29 年 3 月

研究者: 宮尾 祐介

1. 研究のねらい

本研究は、自然言語テキストで書かれた情報と、非言語データ(画像・映像、データベース、時系列数値データなど)とを統合的に理解して推論を行うフレームワークを想定し、そのために必要となる自然言語の意味表現と、テキスト・非テキストデータの意味解析・推論技術を構築することを目指す。ビッグデータの効果的な利活用のためには、単に個々のデータを解析・検索・可視化するだけでなく、現在までに得られている知識を適用・応用して新たなデータや知識を創出する必要がある。後者のプロセスは、現在のところ研究者やデータサイエンティストなどの人間に全面的に頼らざるをえず、ビッグデータを利用した研究開発のボトルネックの一つと考えられる。様々

なデータから得られる知識は自然言語を用いて表出・伝達・蓄積されるため、テキスト化された知識と非テキストデータに埋もれた情報を統合的に利用するためには、これらをつなぐ枠組みが必要である。



そこで、自然言語処理で用いられる自然言語の意味表現を出発点として、この意味表現を介してテキストデータや非テキストデータを接続することを可能にする枠組みについて研究を行う。これまでの研究で、テキスト間の意味的關係や推論關係(含意、矛盾、パラフレーズなど)を計算するための意味表現として、依存構造に基づく構成的意味表現と、それを利用した推論手法について研究を進めてきた。これを非テキストデータに適用することができれば、自然言語テキストと同様の推論を、テキスト・非テキストデータをあわせて行うことができる。

このような技術を実現するためには、自然言語の意味表現と解析技術、および非テキストデータに対して意味表現を求める技術の 2 つが必要である。前者は、テキスト間含意關係認識というタスクとして、自然言語処理において研究が進められている。ただし、非テキストデータと接続するために適した意味表現の設計や、認識精度の向上が課題として残っている。一方、後者の問題については、今のところ研究分野として確立されていない。例えば、「アジアの多くの都市では、モバイルネットワークの利用は夜間に集中している」というテキストがあった時、これを実世界データ、すなわち各都市におけるモバイル通信の利用統計と自動的に対応付けることは自明ではない。逆に、各都市におけるモバイル通信の利用統計があった時、そこからこのような言明を自動生成することも既存技術では不可能である。ここでの問題は、「アジ

ア)「都市」「モバイルネットワーク」「利用」「夜間」といった各単語、および単語間の意味的關係をどのように利用統計データと対応づけるか、という点にある。本研究では、以上の 2 つの問題について、自然言語テキスト、画像、データベース、時系列データを対象に、基盤技術の研究を行う。

2. 研究成果

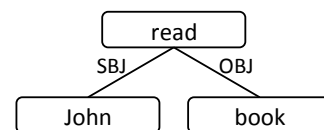
(1) 概要

本研究では、以下の 4 項目について研究を進めた。研究項目 A. 自然言語テキスト間の含意関係認識の研究では、依存構造に基づく意味表現を利用して、頑健かつ高精度なテキスト間含意関係認識を行う手法を提案した(論文 5)。また、英語や日本語において、テキスト間含意関係認識、意味構造解析、およびこれらを利用する応用タスクの研究を推進するために、共有リソースの構築およびタスク提案を行った(論文 3, 4)。研究項目 B. 記号化された非テキストデータに対する解析技術の研究では、自然言語の質問文を大規模データベース (Freebase や DBPedia などのリンクドデータ) に対するクエリに自動変換することで、質問応答を行う技術について研究を行った(論文 2)。これは、自然言語の構文・意味構造をクエリという形式言語に変換することで、自然言語とデータベースをつなぐ技術である。また、日本語においてこのような研究を行った例がないため、既存の英語のデータセットを翻訳することで、日本語の評価用データセットを構築した。研究項目 C. 記号化されていないデータに対する解析技術の研究では、画像に対して依存構造に基づく意味表現を計算する手法を提案し、映像検索タスクにおいて有効性を示した。また、この意味解析技術を株式市場データに適用し、同様の意味解析が可能であることを示した。さらに、映像データに対する意味解析の研究を進めるために、自然言語による説明文を付与した映像データの開発を行った。研究項目 D. これらの技術を応用したプロトタイプシステムの開発では、上記の技術の応用例として、画像に対して自然言語により検索を行うシステムを実装し、本研究で開発した技術の有効性を示した。

(2) 詳細

研究項目 A. 「自然言語テキスト間の含意関係認識の研究」

本研究項目では、依存構造に基づく構成的意味表現を利用して頑健かつ高精度なテキスト間含意関係認識を行う手法について研究を行った。依存構造に基づく構成的意味表現とは、右図(“John reads books”の意味表現)に示すよう



に、単語とその間の構文的・意味的關係(依存関係)の木あるいはグラフで文の意味を表現するものである。述語論理などに基づく意味表現と比べて、直感的にわかりやすく、かつ解析がしやすいという長所がある。一方で、他の形式論理と比べて記述力が弱いという短所があるが、自然言語の文の意味や推論関係を表すには、ほとんどの場合この表現で十分であることを示した。さらに、この意味表現は構文構造との対応関係が明確であるため、構文構造を用いたパラフレーズ認識手法などを組み合わせることができる。英語におけるテキスト間含意関係認識の標準データセットを用いた実験では、人手で構築されたシソーラスである WordNet や、統計的手法で類義フレーズを認識することができる word embedding を用いることにより、高精度を達成することを示した(論文 5)。本手法は、日本語のテキスト間含意関係認識や法

律文書におけるテキスト間含意関係認識にも適用され、有効性を示した。また、本手法をさらに拡張し、一階述語論理では記述することが難しい一般化量子子に関して高精度かつ高効率な推論が可能であることを示した。

また、テキスト間含意関係認識は英語での研究がほとんどであったため、日本語における研究を推進するため、日本語の評価用データセットを構築し、評価型タスクを運営した(論文4)。本タスクには、国内外から多くのチームが参加し、日本語のテキスト間含意関係認識の研究の標準データとなっている。また、欧州を中心とした評価型ワークショップ CLEF や、自然言語の意味解析に関する評価型ワークショップ SemEval において、意味解析やその応用に関する評価用データを開発し、評価タスクを運営した(論文3)。

研究項目 B. 「記号化された非テキストデータに対する解析技術の研究」

本研究項目では、関係データベースやリンクトオープンデータなど記号化されているデータベースに対し、自然言語テキストとの対応関係を計算する手法について研究を行った。特に、大規模リンクトデータ(Freebase や DBPedia など)を知識源として用いて、自然言語の質問文で質問応答を行うタスクにフォーカスして研究を進めた。本タスクで代表的な評価型タスクである Question Answering over Linked Data (QALD) で提供されているデータから抜粋した例を下に示す。

質問文: Which river does the Brooklyn Bridge cross?

答え: http://dbpedia.org/resource/East_River

QALD では、質問文は自然言語文として与えられ、答えとして Uniform Resource Identifier (URI) を返すことが求められる。このタスクは、与えられたリンクトデータ(この例の場合は DBPedia)に答えが存在することを前提としている。すると、データベースに対する検索クエリを作ることができれば、答えを得ることができる。上記の例については、以下のようなクエリを作ればよい。

```
SELECT DISTINCT ?uri WHERE {  
  res:Brooklyn Bridge dbo:crosses ?uri  
}
```

詳細は割愛するが、これは下図のような部分グラフを検索することに相当する。



したがって、本タスクは、自然言語文を上図のようなグラフに変換することが目的となる。この例から明らかなように、本タスクは、自然言語の単語やフレーズ(Brooklyn Bridge や does ... cross など)を、データベースのノードやリンクに変換し、それらを適切に組み合わせる、という問題に帰着される。

本研究では、このタスクに対し、木構造を変換するフレームワークである Tree Transducer を応用する手法について研究を行った(論文2)。Tree Transducer は、入力の木構造を有限の変換規則を用いて出力の木構造に変換する手法であり、これまで機械翻訳などに応用されている。そこで、自然言語の構造を表す木構造を入力とし、データベースクエリを表す木構造を出力する手法を提案した。機械翻訳と比べると、本タスクでは利用できる学習データが圧倒的に小さい。そこで、単語やフレーズをノードやリンクに変換する部分には教師なし学習を、変換ルールを学習する部分には教師あり学習を用い、双方の組み合わせを最適化すること

で変換規則の学習・適用を行う手法を提案した。これにより、少ない学習データからでも高精度な変換規則が得られることを示した。

また、上記のタスクはこれまで主に英語で研究が行われてきたため、既存の標準データセットの英語データを翻訳することで、日本語データセットを構築した。このデータセットは、日本語の質問文をデータベースクエリに変換する研究などで利用されている。

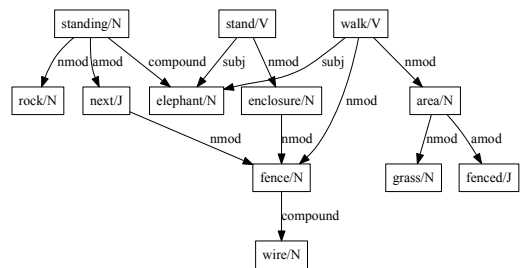
また、上述のタスクとは逆に、自然言語テキストからデータベースを自動構築する研究について、共同研究を行った。情報抽出技術を利用することで、大規模テキストから高精度で関係知識を獲得することができ、質問応答において有効性を示した。

研究項目 C. 「記号化されていないデータに対する解析技術の研究」

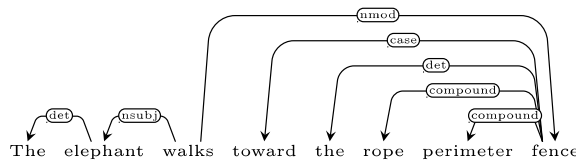
ビッグデータの多くは、画像・映像データ、センサーデータなどの時系列データなど、記号化されていないデータである。本研究項目では、主に画像データに着目し、上記研究項目で用いている依存構造に基づく意味表現を計算する手法について研究を進めた。



本研究では、右上図のような画像データを入力とし、その画像の意味内容を示す右下図のようなグラフ構造を出力する手法を提案した（論文 1）。このグラフ構造は、上述の研究で用いた依存構造を元としており、ノードが単語、エッジが単語間の関係を表している。このような形式で画像の意味表現が得られれば、上記のような意味解析・推論技術を組み合わせることができ、様々な応用が期待できる。



多様な画像に対して高精度でこのような出力を認識するモデルを学習するためには、大規模な学習データが必要である。しかし、このようなグラフ構造を、大量の画像に対して人手で構築するのは不可能である。そこで提案手法では、画像と説明文がペアとなった既存の大規模リソースを利用し、画像と意味表現のペアのデータを構築した。まず、説明文に対して自然言語の構文解析器を適用することで、構文・意味構造が得られる。たとえば、上記の画像に対する説明文 “The elephant walks towards the rope perimeter fence.” という文に対しては、下図のような構造が得られる。



一つの画像に対し、複数の説明文が与えられているため、すべての説明文に対して構文解析を行い、その結果を前処理・マージすることで、上記のような意味表現を得ることができる。

このような学習データが得られれば、入力画像に対してグラフ構造を予測するモデルを学習することができる。提案手法では、画像処理において複数の物体や概念を高精度で認識することができる深層学習手法を応用し、入力画像に対して各エッジを予測するモデルを提案し

た。これにより、人間が記述した説明文から得られる意味表現の一致率と同程度の精度で、意味表現に自動認識が実現できることが示された。また、この技術を応用することで、映像検索タスクにおいて有意な精度向上が見られることを示した。

また、本研究項目では、時系列データとして株式市場データや、映像データに対して同様の技術の研究開発を進めるため、研究用データの構築を行った。これらのデータを用いた予備実験では、画像に対する意味解析と同様に、意味表現の自動認識が可能であることを示している。

研究項目 D. 「応用プロトタイプシステムの開発」

これまでに述べた基盤技術の応用可能性を示すため、画像データに対して自然言語で検索を行うプロトタイプシステムを実装した。本システムでは、大量の画像データに対してあらかじめ依存構造を計算し、検索エンジンでインデックスを作成しておく。この検索エンジンは、自然言語テキストの検索で広く用いられているものである。つまり、画像データに対し、あたかもテキストデータかのように検索を行うことができる。本システムは意味表現を用いた検索を実現しているため、同じ単語でも異なる意味関係を区別するような検索や、否定などの論理関係を用いた検索が可能であるなど、今後の様々な応用が期待される。

3. 今後の展開

本さがけ研究では、自然言語と非テキストデータを意味表現を介してつなぐ技術について研究を行い、特に自然言語間の意味的推論、自然言語とデータベース、および自然言語と画像を対象とした研究において、大きな成果を挙げてきた。これまでの研究成果では、自然言語の意味表現と、データベース・画像それぞれをつなぐ技術の可能性について示すことができた。しかし、これらの技術を実社会の応用につなげるには、精度の大幅な向上が必要である。このためには、画像やデータベースをそのまま用いるだけでなく、自然言語テキストやセンサーなど、外部知識を有効活用することが必要であると考えられる。また、本研究成果のさらなる可能性として、意味表現を介して複数の非テキストデータをつなぐ技術も考えられる。たとえば、データベースに蓄積された情報と、画像・映像データに埋もれた情報をつなぐことができれば、これまででない応用技術が生まれる可能性がある。

具体的な応用アプリケーションとしては、画像やデータベースを利用した自然言語検索、質問応答システム、対話システムなどが考えられる。たとえば、観光案内の対話システムを考えると、あらかじめ想定している質問だけでなく、ユーザから与えられる画像・映像データと気象や交通データの情報を統合して、観光情報を提示したり新たな情報を蓄積するといった応用が考えられる。現時点では、このような知的システムの実現はまだ遠いと言えるが、本研究で開発した技術が重要な基盤となると期待される。

4. 評価

(1) 自己評価

本研究課題は、自然言語処理だけでなく、データベース検索や画像処理など、複数分野にまたがった知見や技術が必要とされる、きわめてチャレンジングなものであった。本研究を開始する前は、自然言語処理の研究経験しかなく、他分野に関する知識がほとんどない状況であり、研究の推進は困難であったと言える。しかし、本研究がさがけ研究者をはじめとして多

くの研究者と接点をもつ機会となり、現在ではさまざまな分野の研究者やエンジニアと共同研究を行い、研究論文を発表するまでに至っている。たとえば、画像に対して意味構造を認識する技術については、自然言語処理と画像処理の最先端技術を組み合わせる必要があり、当初の計画以上の成果が得られたと評価できる。

研究費は、主にデータ作成、研究補助員の雇用、および実験のためのクラウドサービスに用いた。データ作成は、新たな研究領域を開拓する本研究では必要不可欠であり、研究費が有効活用できたと考えている。研究補助員は、主にデータ作成とソフトウェア開発の補助にあたり、研究開発の効率的な推進に大きく寄与した。また、提案手法の評価実験のためには大きな計算リソースが必要であるが、クラウドサービスを活用することにより、管理コストや導入コストを大幅に削減することができた。

今後の展開で述べたように、本研究の成果はまだ基礎研究の段階であるが、自然言語とその他多様なデータを統合的に利用する場面は社会の様々な活動で現れるため、将来的な波及効果は大きい。画像・映像データの意味理解は、上で述べた観光案内をはじめとして様々な質問応答システムや対話システムに応用できるほか、視覚情報を利用する知的ロボットにも必須の技術である。また、株式データなどの時系列・数値データの意味理解は、企業における戦略判断、政策の妥当性や効果の検証など、自然言語を主とした知的活動において様々なデータをエビデンスとして必要とする場面に応用できる可能性がある。

(2) 研究総括評価(本研究課題について、研究期間中に実施された、年2回の領域会議での評価フィードバックを踏まえつつ、以下の通り、事後評価を行った)。

有史以来、人類が生み出した知識は様々な形態で流通・蓄積されてきた。代表的な形態である書籍や論文は、本文、図、表などで構成されており、電子化されたものだけでも大量に存在する。これらをビッグデータとみなし、本文、図、表などの表現形式の違いを超えて、コンピュータが意味まで読み取って処理を行うことができれば、大きな社会的価値を生み出すであろう。

本研究は、自然言語のテキスト、画像、データベースの表現形式の違いを超えて、統合的に利用するための研究を行うものである。自然言語研究で一般的な意味構造を中心に、テキスト間の含意関係の認識、テキストによるデータベースの問い合わせ、画像からの意味構造の抽出に関する研究を行っている。そして、これらの要素技術を統合して自然言語で問い合わせ可能な画像データベースのプロトタイプも構築している。研究の過程では評価に必要なデータの整備も行っている。これらに関する論文を、トップカンファレンスなどで多数発表しており、優れた学術成果をあげている。

今後、適切な応用分野を設定して、実用化を目指すことを期待したい。また、自然言語処理、画像理解、機械学習などの広範な技術分野に関連する研究テーマであることから、これらの分野の研究者をリードして研究に取り組むことも期待したい。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

1. Sang Phan, Yusuke Miyao, Duy-Dinh Le, Shin'ichi Satoh. Video Event Detection by Exploiting Word Dependencies from Image Captions. Proceedings of COLING 2016. 2016. pp. 3318-3327.
2. Pascual Martínez-Gómez and Yusuke Miyao. Rule Extraction for Tree-to-Tree Transducers by Cost Minimization. Proceedings of EMNLP 2016. 2016. pp. 12-22.
3. Stephan Oepen, Marco Kuhlmann, Yusuke Miyao, Daniel Zeman, Silvie Cinková, Dan Flickinger, Jan Hajič, Angelina Ivanova, and Zdeňka Urešová. Towards Comparability of Linguistic Graph Banks for Semantic Parsing. Proceedings of LREC 2016. 2016. pp. 3991-3995.
4. Suguru Matsuyoshi, Yusuke Miyao, Tomohide Shibata, Chuan-Jie Lin, Cheng-Wei Shih, Yotaro Watanabe and Teruko Mitamura. Overview of the NTCIR-11 Recognizing Inference in TExt and Validation (RITE-VAL) Task. Proceedings of NTCIR-11. 2014. pp. 223-232.
5. Ran Tian, Yusuke Miyao, Takuya Matsuzaki. Logical Inference on Dependency-based Compositional Semantics. Proceedings ACL 2014. 2014. pp. 79-89.

(2)特許出願

なし

(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)

受賞

日本学術振興会 日本学術振興会賞 自然言語の構文解析・意味解析の研究とその応用
2015年2月

情報処理学会 長尾真記念特別賞 自然言語の深い構文・意味解析の研究とその応用 2014
年6月