

論文内容の要旨

博士論文題目 Dependency-based Predicate-Argument Structure Analysis Using Structured Learning and Named Entity Information
(依存構造と固有表現情報を用いた構造学習による述語項構造解析)

氏名 渡邊 陽太郎

(論文内容の要旨)

述語項構造とは、状態や事態を表す述語と、その述語と関係を持つ項によって構成される構造のことをいう。述語と項の関係には、項が述語に対してどのような意味役割（動作主、対象など）を持っているかが与えられる。述語項構造解析は、高次の意味解析を必要とする自然言語処理の応用において重要な要素技術である。

本研究の目的は、近年自然言語処理分野において主流となっている統語構造の表現形式である依存構造に基づく頑健な述語項構造解析の実現である。本論文では、(1) 構造学習に基づく述語項構造解析のためのモデル設計、(2) 述語項構造解析のための固有表現情報の獲得と利用、の二点について論じる。

述語項構造には、構造の要素間に二種類の依存関係が存在する。一つは、述語の意味（語義）によって項が制約される、また項によって語義が限定されるという相互の依存関係、もう一つは複数の項が同じ意味役割を持たないこと、述語の必須の項は基本的に存在するといった項の間の依存関係である。既存研究のアプローチでは、これら双方の依存関係を同時に考慮することが困難である。そこで、双方の依存関係を同時に捉える構造学習法に基づくモデルを提案する。評価実験では、語義と意味役割の双方の分類性能が向上し、さらに素性選択法を適用せずに既存研究に匹敵する性能が得られた。さらに、提案モデルを用いて依存構造と述語項構造の双方を解析するシステムを構築した。評価実験の結果、素性選択法を用いていない既存システムの中では最良の結果が得られた。

述語項構造解析をおこなう上で問題となるのは、文中に現れる固有表現の存在がある。これらは無数に存在し、未知の固有表現が多く現れるため、これらを抽象化した形で扱う必要がある。そこで、文中の固有表現に対して固有表現情報を付与することで、抽象化をおこなう。本論文では、(1) Web 文書からの固有表現情報の自動獲得、(2) 固有表現情報の述語項構造解析モデルへの導入、の二点について論じる。まず、Web 上の百科事典である Wikipedia から、記事に現れるリスト構造の性質を利用した固有表現の獲得手法について提案する。次に、固有表現情報の導入については、既存研究で用いられている方法に加えて、複数の項に関連する大域的な固有表現情報の導入を提案する。評価実験では、既存の導入方法と比較して高い性能が得られた。

氏名	渡邊 陽太郎
----	--------

(論文審査結果の要旨)

平成21年12月25日に開催した公聴会の結果を参考に平成22年2月17日に本博士論文の審査を行った。以下のとおり、本博士論文は、提案者が独立した研究者として、研究活動を続けていくための十分な素養を備えていることを示すものと認める。

文章理解において、事象や状態を表わす動詞、形容詞などの述語とその主語や目的語など直接的に意味的關係を持つ項の同定は重要なタスクである。渡邊 陽太郎は、本博士論文において、文章の述語とその項を同定する問題において、局所的な情報と大局的な情報の両者を利用し、かつ、素性獲得等の手間のかからない自動学習手法を提案した。また、固有名詞等の固有表現の自動獲得に関する研究を行い、獲得された固有表現の情報を利用することによって、述語項構造解析のさらなる精度向上が可能であることを示した。

本論文の貢献は、次のようにまとめることができる。

1. 述語項構造解析の問題では、述語の語義と項の間の関係、あるいは、複数の項の間の関係、のように互いに影響を及ぼし合う関係にある要素があり、述語の語義と個々の項の同定を個別の問題として解くのでは、互いが及ぼし合っている制約を利用することができない。本論文では、局所的な情報に基づく学習と大局的な情報を用いた学習を融合することによって、全体的な最適化による述語項構造解析の方法を示した。
2. 述語項構造解析には統語的情報が重要な役割を演じるが、文の統語解析は必ずしも完全に行えるわけではない。本論文では、統語的依存関係と述語項構造を融合した解決法を示し、自動学習に基づく手法が高い精度を達成することを示した。
3. Wikipediaを情報源として、固有表現を獲得する手法を提案した。特に、Wikipediaの文書がもつ構造情報を利用することで固有表現獲得の精度が向上できることを示した。
4. 固有表現の情報を利用することで、上記の述語項構造解析に対してさらなる精度向上が可能であることを実験的に示した。

以上の通り、自然言語における述語項構造解析に対して様々な自動学習法を提案した本研究は、独創性が高く、しかも実用的であり、自然言語処理の分野において高い貢献があると評価する。

よって、本論文は、博士(工学)の学位論文として価値あるものと認める。