

審査の結果の要旨

氏名 谷中 瞳

本研究は、自然演繹に基づく論理推論を用いた文間関連性の評価の研究で、5章より構成されている。

第1章は序論であり、本研究の背景、課題および貢献、さらに本論文の構成が示されている。文と文とがどのような意味的關係にあるかという文間の関連性を計算する技術は、情報検索や文書分類、質問応答などの自然言語処理の基盤を築く重要な技術である。この技術に関する2つの自然言語処理のタスクとして、2文間の意味的類似度を予測する文間類似度学習と、一方の文が他方の文の意味を含むか否かを予測する含意関係認識がある。このタスクに対するアプローチとして、文の意味をベクトルで表し、類似度や含意関係を予測する機械学習と、文の意味を論理式で表し、直接文間の含意関係を計算する論理推論がある。そこで両アプローチを組み合わせた手法として、論理推論の実行結果を特徴量として用いてモデルを学習し、類似度や含意関係を高精度で予測する手法が提案されている。しかし、推論の実行結果は推論から得られる情報のほんの一部にすぎず、前提から結論を導くまでの推論の実行過程も、文間の関連性の特徴を表しているはずである。そこで本研究の目的は、推論の実行過程が文間の関連性を表す特徴量として有用であることを明らかにし、文間類似度や含意関係を高精度で計算する手法を構築することである。本研究の貢献は、1. 高精度な文間類似度学習モデルの実現、2. 高精度な含意関係認識システムの実現、3. フレーズ間知識を考慮した高精度な論理推論の実現の3点である。

第2章では、本研究全体で扱う、文の意味表現の導出方法と論理推論による文間の意味的関係の計算方法について述べられている。本研究では文の高度な意味計算が実現可能な、高階論理式による文の意味表現を扱う。文の意味を論理式に変換するためにはまず、組合せ範疇文法(CCG)に基づく構文解析によって、文を構文木に変換する。次に、構文木に基づいてラムダ計算による意味合成を行い、高階論理式を導出する。また、本研究では推論に適したNeo-Davidsonian Semanticsに基づく意味表現を採用する。論理式間の意味的関係は自然演繹に基づく論理推論に基づいて計算する。自然演繹は効率的な自動推論を実現する定理証明支援器が存在するため、推論の実行過程がモニタリング可能である。

第3章では、文間の意味的関係の推論の実行過程を用いて、文間類似度や含意関係を高精度で計算する方法が提案されている。提案手法では、文を論理式に変換した上で、

2文間の双方向の含意関係について自然演繹による推論を試みる。またこのとき、必要に応じて、文間の意味的關係を正しく推論するために必要な語彙知識を公理として追加し、推論を試みる。ここまでの推論の実行過程から、推論に必要とした語彙知識、推論規則の適用回数といった様々な特徴量を導出する。これらの特徴量と、表層情報から得られた特徴量を組み合わせてランダムフォレストモデルで学習し、類似度と含意関係を予測する。複数のデータセットを用いて提案手法の評価を行った。その結果、文間類似度学習と含意関係認識のいずれのタスクにおいても高精度を達成した。また、含意関係認識のタスクでは、SICK データセットを用いた評価で最高精度を達成した。

第4章ではフレーズ間知識の公理生成が述べられている。第3章のエラー分析の結果、論理推論を用いて文間の意味的關係を高精度で計算するためには、推論に不足するフレーズ間知識を補完する方法を確立することが重要であることがわかった。そこで、自然演繹に基づいて文間の含意關係の推論を試みる際に、実行過程から不足するフレーズ間知識を特定し、補完する手法を提案する。提案手法を用いてフレーズ間知識の公理を生成し、含意関係認識のタスクで評価を行った結果、純粋な論理推論のアプローチとして最高精度を達成した。また、単語や構文構造に基づく既存のフレーズアライメント手法では特定できなかった、非連続なフレーズや反義のフレーズが特定できた。

第5章では結論と今後の課題がまとめられている。

以上を要するに、本研究では文の意味を高階論理式に変換し、推論の実行過程を特徴量として用いて学習することで、文間類似度と含意関係を高精度で予測できるようにした。加えて、推論の実行過程は、推論に不足するフレーズ間知識の特定にも有用であることを示した。よって本論文は博士（工学）の学位請求論文として合格と認められる。