

論文内容の要旨

博士論文題目 Probabilistic Logic Approach to Event Structure Analysis
(確率的論理による事象構造解析)

氏名 吉川 克正

(論文内容の要旨)

本論文では確率的論理を利用した事象構造の解析を行い、事象と項及びその関係の抽出について述べる。事象とは、文書中における物語を構成する“状態変化”を表す表現である。この事象表現が自然言語の文書において、重要な役割を果たすことはよく知られている。本研究では、このような事象表現に対する解析を、人間の持つ言語学的知識とコーパスからの統計情報の組み合わせにより行うことを目的とする。

本研究で扱う確率的論理は、マルコフ論理と呼ばれ、広く利用されている枠組みである。マルコフ論理は一階述語論理とマルコフネットワークを組み合わせた枠組みであるため、一階述語論理により人間の持つ言語知識を効率的に実装するとともに、マルコフネットワークにより、コーパスから様々な素性を学習して利用することができる。本研究ではこのマルコフ論理によって自然言語処理における三つのタスクを扱い、人間の言語知識とコーパスからの学習を効果的に組み合わせるモデルを提案する。

マルコフ論理による最大の利点は、複数の決定を同時に行える点である。現実のデータでは、データ間に様々な依存関係が考えられる。例えば、一つの文内に複数の述語とそれに対応する項が存在する場合、ただ一つの述語-項の組み合わせだけに着目しても正しい解を得られないことが多くある。なぜなら、同一文内に存在する述語-項関係には互いに依存関係があるからである。本研究の目的は、このような依存関係を人間の知識に基づく論理式により捉えることで、様々な自然言語処理のタスクにおいてより効果的なモデルを構築することにある。

本論文で扱うタスクはいずれも事象に関する関係抽出である。事象の中でも、特に動詞として働くものを中心に考え、三つの関係抽出タスクを行う。まず一つ目は日本語の新聞記事における述語-項関係抽出である。文内全ての単語を同時に考慮して文全体で最適な述語-項関係を捉えるモデルを構築している。このタスクでは定量的な評価だけでなく、定性的な評価も行う。二つ目は医学生物学文書における事象-項関係抽出である。このタスクにおける重要な点として、医学生物学文書という限定的な分野ではあるものの、共参照関係を事象-項関係抽出に利用することに成功したことが挙げられる。その結果、文境界を越えるような事象-項関係も扱うことができるようになる。三つ目は、時間順序関係推定と呼ばれ、事象表現、時間表現、文書作成日時に関してその時間的な順序を推定する問題である。時間的關係に対する大域的な論理制約を導入することで、文書全体における論理的整合性を向上させる試みである。

氏名	吉川 克正
----	-------

(論文審査結果の要旨)

平成 23 年 8 月 4 日に開催した公聴会の結果を参考に平成 23 年 9 月 7 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。

文章に記述された事象の間には時間関係、因果関係など様々な意味関係が存在し、その理解は文書理解における重要な課題となっている。吉川 克正は、本博士論文において、文章中に記述された事象間の関係を解析する種々の手法を提案し、また、提案手法の性能評価を行った。

1. 事象間の関係、および、事象を表現する用言の項となる名詞としての参照物の間に成り立つ種々の制約関係をマルコフ論理という確率的論理式を利用することで、大域的な制約を考慮した解析手法を提案した。これにより、人間による言語学的あるいは常識的な知見による知識を論理式のテンプレートとして記述し、その制約の強さをコーパスからの統計情報として学習することが可能になり、複雑な情報や知識が絡む問題への一つの解決策を示した。
2. 日本語新聞記事における述語項構造解析の問題に対し、文書中の名詞が述語の項になりやすいかどうかに関する大域的な制約をマルコフ論理によって記述し、これにより、述語項構造解析の性能向上を達成した。
3. 医学生物学文書における事象-項関係抽出タスクにおいて、名詞概念間の照応関係と事象-項関係を同時に考え、それらの間の制約をマルコフ論理によって記述することにより、同タスクの性能向上を達成した。
4. 文章中の事象の時間的順序関係の解析において、時間関係に関する大域的な制約を記述することにより、個々の事象間だけでなく、文章全体における事象間の時間的順序関係の整合性を考慮した解析を可能にし、時間関係解析の性能向上を達成した。

以上のように、マルコフ論理の枠組みを用いることによって、文書内の事象に関する様々な関係解析に関し、大域的な整合性を考慮することができる手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。