

⑫ On A Grammar Formalism, Knowledge Bases and Tools for Natural Language Processing in Logic Programming

H.Sano*, F.Fukumoto(ICOT,日本)

発表要旨

自然言語処理技術は、文法の枠組み、文法規則の記述形式や辞書データベース、そして解析手法といった要素技術の均衡の許で発展するものであるし、その成果も総合的な積み重ねによって得られる。論理プログラミングのパラダイムは、正しい調和を支援する枠組である。本稿は、分析対象を日本語とする自然言語処理のための幾つかの要素技術についてICOTの成果を報告する。

まず、我々が構築してきた日本語文法の枠組について解説する。この文法枠組は、日本語に関する最新の研究成果を取り入れたもので機械処理に適合するよう組立てられている。文法枠組の検証のために、2つの具体的な実現のアプローチを採った。LUGは句構造ベースの文法規則の記述形式で、単一化文法の枠内で形式化されている。さらに一つの方法は、RDGと呼ばれ伝統的な依存文法に基礎を置いている。本稿で説明する双方の文法規則の記述形式は、いずれも論理プログラミングの枠内において実現し、具体的に動作を確認した。

語の持つ統語上の性質や意味的な特徴は、語彙知識としてまとめられた。我々が構築した幾つかの言語知識ベースを記載の情報を説明しながら紹介する。最後に、自然言語処理の成果を実際的なものとする、ことばの分析や実験のためのツールについて概説する。

質疑応答

質問：文法規則の扱う範囲はどうなっていますか。

回答：LUGでは、基本的な言語現象に限られています。話ことばは対象にしていません。RDGでは、新聞の社説を扱うことができます。

質問：800という文法規則は一般に比べて大変多いようですが、どうしてですか。

回答：形態素システムを文法が持っているからです。

質問：800のうちどのくらいが実質的に、いわゆる構文の解析に使われていますか。

回答：30～40規則です。