

論文審査の結果の要旨

氏名 シュタイナー ローランド

本論文は7章から成っている。第1章においては、本研究の目標である文法外入力 of 頑健な自然言語処理の必要性、既存の研究、本論文のアプローチが述べられている。

第2章においては、現在の多くの文法開発プロジェクトが採用している文法枠組である HPSG について解説されている。これは、素性構造を用いて単語、句の制約を記述した単一化ベースの枠組である。

単一化ベースのシステムの頑健性を増す方法としては基本的に二つある。一つは、文法により規定される制約を緩めることである。しかしながら、いくつかの理由によりそれは望ましくない。解析過程で得られる情報が減り、探索空間を大きく広げ、たとえ適格文が与えられたとしても非常に冗長な曖昧性を生み出してしまふ。もう一つの方法は、文中に出現する間違いに対し修正をしたり、許したりすることである。文法の受容範囲を超える解析を許す方法の一つは、“単一化強制”である。単一化強制を用いることにより、システムは全ての句、単語の組み合わせを計算し、文法的衝突（例えば、人称の不一致など）が生じる場合でも解析される。探索と修正を繰り返して解析する方法では、そのような文法的衝突も解析され正しい解析結果が得られる。しかしながら、そのような方法では、解析結果はあまりに一般的な解析結果（ありとあらゆる可能性を考慮した解析結果）になってしまい却って情報量が少なくなり、また、高効率なシステムを実装するのは難しい。また、一般的に言って柔軟性がない。

これらの理由から、本研究では、文法外入力を取り扱うためのユーザー定義可能なルールベースの機構が提案されている。ルールベースのアプローチは、そのルールを拡張することでカバレッジを拡張することができる一方、ルールに記述された場合だけが取り扱われるため、おかしい結果はほとんどあるいは全く生成されないという利点がある。パーズシステムが文法外であると判断した入力に基づいた(部分的な)結果を分類するために、本論文ではペナルティが使われる。ここで、ペナルティの大きさは文法侵害の大きさを示す。このアプローチによって、システムは、ありそうもない入力の解釈を考慮する前に、低いペナルティが与えられた最も可能性の高いパーズ解析を最初に追求することができる。

第3章では、まずメタ単一化の概念が導入されている。メタ単一化とは、ルールで統御された単一化プロセスで、特定の状況の下、素性構造上の特定の位置についての単一化過程を”修正し”、通常の単一化から逸脱するような単一化結果を生成することをユーザーに許す。

本論文では、メタ単一化を元に、単一化ベースのシステムにおける頑健性のための2つの主要な機構として、M-ルールと拡張IDスキーマが提案され、それらの詳細な分析が行われている。

第4章では、M-ルールについて詳説されている。M-ルールは、2ステップからなるアプローチで、ユーザー定義によって型階層を拡張することで、単一化が失敗するような特定の状況下においても単一化が成功するようにするものである。この型階層の拡張へ、型階層の任意の点を与えると一意に定まるもので、事前に拡張することで実行時のシステムに影響を与えない。本論文では、この拡張を生み出す正確なアルゴリズムとその特性が詳述されている。適度に拡張された型階層によって、ユーザー定義のM-ルールへ、以前であれば衝突を起こすような単一化において、衝突の原因となる素性構造の場所を見つけだして修復する、すなわち、正しい結果に置き換える。これらを通常の単一化に関連づけるために、本論文は実装の詳細と基本的枠組みの拡張についてのべるだけでなく、M-ルールとメタ単一化の理論的特性についても考察している。

第5章では、もう一つのアプローチである拡張IDスキーマについて詳述されている。拡張IDスキーマでは単一化操作を行う前に文法的衝突を発見し、直接正しい結果に変換する。HPSG文法枠組では単一化の対象となる二つの構文木構成要素は常にIDスキーマにより支配されるため、IDスキーマを拡張することが理想的である。その名前が示すように、拡張IDスキーマは文法的例外を直接検出し、取り扱いを行う。スキーマの変化はIDスキーマをより細かな部分に分割することにより定義される。この細かな部分を素性構造構築ブロック(FSBB)と呼ぶ。文法的例外を取り扱うためにユーザーは単一化を妨げるFSBBを見つけ取り除き、そして、意味のある出力を得るために他のFSBBを追加することになる。各々の例外を記述するルールは前もって定義された種類の現象に対してしか作用しないので、ユーザーはこのような現象を唯一に発見できるFSBBを追加できる。このように制約を追加することにより、探索空間を自在に制限することができる。

拡張IDスキーマの他の利点は、複数の例外を同時に扱えるように例外の記述を合成することができるということにある。そのような場合、全

ての例外定義から ID スキーマへの修正は同時に行われる。この方法の理論的特性は非単調性を避けられるということである。M-ルールと同様に、本論文では、拡張 ID スキーマの実装の詳細、拡張の可能性、および実験について報告されている。

第 6 章では、二つのアプローチの統合の可能性について議論されている。前者は主として意味論的な処理、後者は主として統語論的な処理を扱うため、統合は容易である。第 7 章において本論文の結論が与えられている。

なお、本論文は、辻井潤一氏との共同研究に基づいているが、論文提出者が主体となって分析及び検証を行なったもので、論文提出者の寄与が十分であると判断する。

従って、博士(理学)を授与できると認める。