

## ⑮ A Parallel Cooperation Model for Natural Language Processing

S.Yamasaki\*, M.Turuta, I.Nagasawa, K.Sugiyama(富士通, 日本)

### 発表要旨

並列推論マシンを活用した自然言語解析のためのソフトウェア技術の開拓を目的として、形態素解析、構文解析、意味解析の並列協調を行なう方法を開発し評価した。

並列協調の原理として型推論を利用することにより、入力に応じて自動的に並列協調が実現される。(文字列からの解析、形態素列からの解析、文字形態素、解析途中の中間的構造が混在したものからの解析の実演により自動的並列協調を説明)

(型推論を利用した意味解析との並列協調により、主語と目的語の意味関係が動的に変化できることを実演)

(意味処理を並列協調させることの効果を、意味処理なしの解析と意味処理を協調させた解析を同一例文について実行し、意味処理を含めた方が解の個数が減るだけでなく、処理時間も減少することを実演)

(台数効果の実演として、同一例文を1台と16台で解析を実行し、解析時間の比率を見せた)

性能評価実験結果の説明。約14000語の辞書約600規則の文法、約700概念の意味体系規則を用いた実験の結果、32台のプロセッサを使用すると1台での処理に比べて約13倍の高速化が達成された。

### 質疑応答

質問：32台のプロセッサを使用したのに、なぜ32倍速くならないのか。

回答：この研究は並列協調技術にポイントがあるので、通信が本質的に必要である。このためにどうしても通信のオーバーヘッドがでてしまう。

質問：32台までは台数効果が線形にのびているようだが、もっと多くのプロセッサを使ってもそうなるのか。

回答：32台までのグラフから64台使うと約20倍の台数効果が期待されるが、これまでに使用した例文では64台では17倍から18倍程度しか台数効果がでていない。もっと長い例文なら20倍であるかもしれない。

質問：しかし、自然言語処理の対象として普通はそんなに長い例文はないのではないか。

回答：確かにそのとおりである。自然言語処理としてはむしろ並列協調させる要素を現在よりももっと増やす方が良いだろう。例えば現実世界からのフィードバックのような情報を協調させたい。

質問：デモに使った例文の解の個数は直観的に考えるよりも非常に多いように思えるが、いったいどうしてそんなに沢山の解がでてくるのか。

回答：ここで使用している文法は基本的にICOTで開発されたSFTBに準拠している。個々の解が出てくる理由を簡単に説明するのは非常に難しい。