

## 審査の結果の要旨

氏 名 狩野 芳伸

近年、自然言語処理分野の発展とともに、注釈付きコーパスや自然言語処理ツールなど利用可能な言語資源は飛躍的に増加してきた。自然言語処理の特徴として、同じタスクに対しさまざまなツールが公開されその性能を競っている一方で、適用するテキストのドメインが異なると同じツールでも大きく異なる振る舞いを示し、その変化が予測しがたいということがあげられる。また、サブタスクの組み合わせにより別のタスクを達成できることが多く、ツールの再利用の需要が非常に高いという特徴がある。すなわち、言語資源の組み合わせと、比較や評価といった実行結果の解析が欠かせないといえる。にもかかわらず、既存研究では、言語資源の開発はその組み合わせに不可欠な相互運用性を考慮せず行われることが多く、再利用やツールの比較、ドメインを変えての再評価なども困難であった。

本論文は、自然言語処理における相互運用性について、ユーザ作業の自動化という観点から設計方針を提案して、それを実装し、徹底した相互運用性を実現している。また、相互運用性に基づいて自動組み合わせ比較や視覚化といった汎用の解析機能を提案し、実タスクに適用してその有用性を示している。本論文は「**RICH ANALYSIS AND AUTOMATION FOR NATURAL LANGUAGE PROCESSING BASED ON INTEROPERABILITY**（自然言語処理のための相互運用性に基づく自動化と深い解析）」と題され、6章からなり英文で書かれている。

本論文の貢献は、大きく次の三点である。第一に、さまざまな側面から相互運用性を考察し、それを保証する設計方針に基づいて、世界最大規模の互換言語資源群を構築して提供している。第二に、言語資源の自動組み合わせ比較と自由に組み込み可能な評価機能、機械学習モデルの解析を含めた結果の視覚化といった解析機能を、相互運用性を損なうことなく設計し実装している。第三に、これらの機能と言語資源群を言語処理システムとして統合し、システムのインストールから言語資源の組み合わせ、比較評価、実行と結果解析までをプログラミングすることなく達成可能にし、高度なユーザビリティを実現している。

実験結果と実ユーザの体験からは、さまざまな分野での応用において、システムが実用的なタスクに適用可能な頑健さをもつと同時に、ユーザの自然言語処理タスク達成を質的にも量的にも飛躍的に効率化できることが明らかになった。自動組み合わせ機能を拡張したシステムアンサンブルの実験においては、BioNLP Shared Task 2009でのF1値評価におけるベストスコアを4ポイント上回る55.11を達成した。

以上のように、本論文は自然言語処理における相互運用性とそれに基づく解析手法について、自動化の観点から設計を提案し、それらを統合した実行環境を実装した上で、その有用性を実験により確認しており、コンピュータ科学上貢献するところが少なくない。

よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。