

審査の結果の要旨

論文提出者氏名：鷺尾光樹

この課程博士学位請求論文の審査は、(主査)加藤恒昭、(副査)川崎義史、山口和紀(以上、東京大学大学院総合文化研究科)、林良彦(早稲田大学)、関根聡(理化学研究所)の5名によって行われた。公開審査は令和2年1月20日(月)17時から18時30分まで、18号館コラボレーションルーム2において行なわれた。論文題目は「単語間意味関係知識の獲得と応用」である。以下、審査結果の要旨を報告する。

鷺尾光樹氏の博士論文は、上位下位関係(包摂関係)や対義関係など、単語やその意味の間に成り立っている関係である単語間意味関係の工学的な扱いを論じている。本論文では、単語間意味関係知識をコーパス(文書集合)から自動的に獲得する手法を提案するとともに、そのように獲得された単語間意味関係知識を自然言語処理技術に応用した事例を提示している。提案されている手法は従来手法に比べて獲得できる知識の網羅性が高く、近年多くの自然言語処理技術で活用されているニューラルネットワーク技術との親和性が高い形式を有している。これらの点を活用することで、辞書知識を扱うふたつの自然言語処理技術応用、語義記述文の自動生成と語義記述文からの単語の表現獲得、において高い性能が出せることを示している。

本論文は、序論と結論を含めて、6つの章から構成されている。序論である第1章では、従来の人手構築による単語間意味関係知識の問題点が語られ、より網羅性が高く、ニューラルネットワーク技術との親和性が高い形式を持つ適切な単語間意味関係知識の必要性が述べられる。そして、自然言語処理技術において、そのような知識を活用する範囲を広げる必要があるという、本論文の動機が述べられている。

第2章では、単語間意味関係知識についてのこれまで研究が概観される。単語間意味関係の定義に始まり、人手構築された知識の問題点が述べられた後、単語間意味関係知識の自動獲得についての先行研究が説明される。獲得された知識の質を評価する手法の紹介がなされた後、それらの知識が自然言語処理技術の中でどのように活用されているかの現状が示されている。

第3章で、単語間意味関係知識をコーパスから自動獲得する手法が提案される。提案手法では、依存構造のパス等で表現される、ふたつの単語が文中でどのように共起したかという関係パターンを用いて意味関係知識が獲得される。従来の手法では、関係パターンを用いて獲得できる単語間意味関係知識はコーパス上で実際に共起した単語ペアに関するものに限定されていたが、本提案ではニューラルネットワークの汎化能力を用いることで、単語の表現が得られている任意の単語ペアについて、その間の意味関係知識を獲得することができる。単語間意味関係類似度の評価ベンチマークと、意味関係識別タスクのデータセットで評価を行い、提案手法によって得られた単語間意味関係知識の有効性が確認されている。

第4章では、単語間意味関係知識の新たな応用として、これを辞書知識処理に適用する手法を提案している。辞書知識処理には、辞書構築のための語義記述文の自動生成と、辞書に含まれる語義記述文からの単語の表現獲得がある。辞書によって定義される見出し語とその意味を定義する語義記述文との間の対応関係をモデリングする処理である。語義記述文は見出し語と意味関係を持つ単語が用いられているので、その関係を単語間意味関係知識を用いてモデルに組み込むことで、より適切な辞書知識処理が行えることが示されている。

第5章では、本博士論文中で行われた提案をあらためて考察し、獲得された単語間意味関係知識の性質が議論されている。あわせて、残された課題として、ハイパーパラメータの調整を含む手法の網羅的な検証の必要性、どのような意味関係知識が獲得されているかのより詳細な分析の必要性が挙げられている。

最終章の第6章は、全体のまとめにあてられている。

本論文の評価として、単語のペアと関係パターンとの共起をニューラルネットを用いて汎化するという着眼がシンプルでありながら、獲得される単語間意味関係知識の適切性と網羅性を高めることにおいて非常に強力なものであって、極めて秀逸であると、認識された。そして、この着眼に基づいて具体化を進め、適切な実験を行って優れた性能を確認していった研究の進め方とともに、行った提案の限界について考察し堅実に見定めていることも、好意的に高く評価された。加えて、獲得された単語間意味関係知識の応用として、語義記述文の自動生成という困難な生成問題に取り組み、最先端の結果を得ていることも卓越した点としてあげられた。

一方で、全体に対する客観的数値評価に力が入れられるあまり、細部の誤り分析、個々の事例に対する考察が不十分ではないかとの問題点も指摘された。また、タスクにおける性能評価にとどまらず、得られた知識がどのような性質を持っているのか、望ましい性質を得るためにはどうすればよいかという観点から、更に実験を進めるべきであったとのコメントもあった。これらはすべてもっともな指摘であり、論文中でも残された課題として挙げられているものであった。

しかし、こうした欠点は全体の価値を揺るがすほどのものではなく、鷲尾氏の論文は、これまでになく高い質を有する単語間意味関係知識の新しい獲得方法を提案した力作であり、その提案はさまざまな自然言語処理技術応用へ利用されていく可能性を秘めている。全体として学術的な価値が高く、この分野における優れた研究成果として十分な評価に値するものであるとの審査員全員の合意を得た。したがって、本審査委員会は博士（学術）の学位を授与するにふさわしいものと認定する。