

審査の結果の要旨

氏 名 グェン ルー トゥイ ガン

本論文は、共参照解析を特に医療・生物学文献上のタンパク質名に対して行うために、その評価の基盤となる評価用大規模データ基盤の構築ならびに評価方法の確立を行い、さらにそれに基づいて得られた知見に基づき、新たにタンパク質名に対し特化した新規手法を開発し、それによって医療・生物学文献上のタンパク質名の共参照解析の精度を大幅に向上させることに成功したものである。

本論文は七章からなり、第一章は本論文における共参照解析を医療・生物学文献に特化して研究を行う必要性あるいは動機付けを明らかにしている。第二章は、本論文で用いられる各種用語の解説、共参照解析に関連した関連研究の概略を紹介している。第三章では、対象分野によって共参照解析が異なる性質を持つことを、各種文献集合に対して解析を行うことで明らかにしている。第四章では、そのような共参照解析を医療文献に特化して評価する際の評価方法の提案、ならびに実際に評価する際に重要な基盤となる評価用大規模データ基盤の構築に関して議論されている。この構築された大規模データ基盤ならびに評価手法は、実際にコンベンション型国際ワークショップにおいて参加アルゴリズムの評価に用いられた。また、このデータ基盤は、さらにその後も研究コミュニティの重要な共有資産として用いられており、すでに多数の引用があるなど実際に研究コミュニティへの大きな貢献となっている。第六章では、そのコンベンション型国際ワークショップに参加した多数のチームの共参照解析アルゴリズムの比較を、第五章で構築された大規模データ基盤を用いて行い、各チームのアルゴリズムの特徴を議論するとともにそれぞれの手法の評価し、医療・生物学文献上の共参照解析において何が重要であるかについて議論を行っている。第六章では、第五章で得られた知見に基づき、医療・生物学文献上に特化した共参照解析を大幅に高精度化するため、語句のセマンティック情報を共参照解析の精度向上に活用した新たな手法を提案し、それによって従来の他手法と比較して大幅な精度向上ができたことを実験を通して明らかにしている。第七章では、本論文で提案した手法と有効性を総括し、さらに今後考察すべき課題についての展望が示されている。

なお、本論文の第三章はJin-Dong Kim氏と、第四、五、六章は、Jing-Dong Kim氏、三輪誠氏、松崎拓也氏、辻井潤一氏との共同研究であるが、論文提出者が主体となって立案、分析、検証を行ったもので、論文提出者の寄与が十分であると判断する。

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。