

論文審査の結果の要旨

氏名 仲尾 由雄

本論文は5章からなる。第1章では研究の背景と目的が述べられている。ここでは、テキスト中の話題階層を検出する手法が取り上げられ、それをテキスト要約問題や複数テキストに含まれる関連話題の抽出問題に応用することの有用性について論じられている。なお、本研究で取り扱う話題階層とは、テキストの各部分が何について書かれたものかを階層的に表現したテキスト構造を意味する。第2章では、語彙的結束性に基づく話題階層検出手法が提案されている。この手法の特徴は、語彙の反復を手がかりに、ほぼ同じ大きさの話題に関する境界を検出できる点にあり、この性質により、テキストに含まれる様々な大きさの話題のまとまりを体系的に検出することが可能となる点にある。第2章ではさらに、この提案手法を3種類の性質の異なる文書に適用し、文書の論理構造と検出結果とを比較することにより、本手法で検出された話題階層が文書の論理構造とよく一致していることを確かめている。これにより、提案手法が、様々な粒度の話題を正しく検出できることの可能性を示している。第3章では話題階層を利用した、単一テキストの自動要約手法が提案されている。この手法は、長い文書の概要把握支援のために、話題を幅広く含む要約を作成することを想定しており、抽出話題数に応じたテキストの分割の仕方を、話題階層を利用して求めていることに特徴がある。人間による要約結果と比較することにより、話題階層の利用が単一テキストの自動要約に有効であることを示唆している。第4章では、複数の関連文書から関連箇所を抽出する手法が提案されている。本手法の特徴は、話題階層を利用して様々な大きさの話題区画を体系的に比較している点と、関連話題抽出用の閾値を話題階層に基づき自動設定している点にある。本章ではさらに提案手法を国会における代表質問と答弁との間の関連性を調べるという問題に適用し、その結果を、抽出区画対の対応関係の正しさ、抽出話題の網羅性・簡潔性、および、主要な話題のカバー率という3つの観点から評価することにより、複数文書からの関連話題抽出において、話題階層の利用が確かに有効であることを示している。第5章では本研究の結論、すなわち、話題階層は、適切な話題を幅広く抽出する上で有用であること、また、話題階層を利用して、様々な粒度の話題を体系的に比較することで、効率的に関連話題が抽出できることなどが述べられている。

本論文に述べられた研究成果は、情報科学その中でもとくにテキストからの話題検出や要約作成などに関する研究の進展に大きく寄与するものである。テキストの要約や複数テキストに含まれる関連話題の抽出などに関する情報技術への需要は、近年のインターネットの普及などによる電子文書や電子書籍の爆発的増大に伴って急速に高まっており、本研究の意義は学問的だけでなく社会的および産業的にも大変大きいと言える。一般にこの種の問題は、正解そのものに曖昧性があり、提案手法の有効性を正しく評価することはなかなか困難である。しかしながら、本研究では評価手法それ自体の有効性に関してさまざまな観点から考察を加え、評価結果にある種の客観性を与えることに成功している。また、語彙の反復だけを手がかりにした手法でどこまで様々な大きさの話題を体系的に検出できるかという点に問題を限定して研究を展開することにより、このような手法の有効性、応用性とその限界とを明確にすることに成功している。このような観点からも本研究の成果はテキストからの話題検出や要約作成などに関する研究の今後の進展に大いに貢献するものと期待される。

したがって、博士（理学）を授与できると認める。