

論文内容の要旨

博士論文題目 A Study on Operations used in Text Summarization
(テキスト要約における言語操作に関する研究)

氏名 竹内 和広

(論文内容の要旨)

自動要約に関する研究は長い歴史を持っているが、依然として、人間が作成する要約と機械が作成する要約の間には品質の面において大きく差がある。それは、自動要約システムの多くが、要約を重要な文を情報として抽出する操作の側面ばかりをとらえ、人間が要約を作成する際に行っている操作をテキスト処理の要素技術としてうまく整理されてこなかったことに原因があると考えられる。

本論文では、人間の作成した要約を分析し、人間が要約する際に行っている言語的処理を計算機で実現するためには何が必要であるかを調査・研究した。本論文の構成は大きくわけて3つの話題からなる。

第一に、人間が行っている要約を要約の元となった文章のテキスト構造との比較をもとに分析する予備実験を行った。テキスト構造とは文のつながりを観点にテキストを分析するための道具立てである。具体的には、新聞の報道記事のテキスト構造を手で分析し、人間の分析が一致する部分とゆれがある部分を考察した。次に、人間の分析に一致が見られる文間のつながりの良さに対してどのような言語的手がかりが貢献するかを機械学習の手法を用いて検討した。さらに、テキスト構造解析を行った記事に対しての要約を作成し、人間の分析に一致が見られる文のつながりの良さは、文の抽出よりもむしろ、要約文を生成するための操作のひとつ、文の結合操作と関連していることを指摘した。

第二に、予備実験で得られた知見にもとづき、人間が要約を作成する操作を詳細に分析した。分析対象は新聞の社説にひろげ、3人の要約作成者が作成した要約を対象としたテキスト分析を行った。分析には、要約文で使われている表現が要約元文章のどの部分から生成されたかを自動的に対応付ける手法を提案した。その結果、要約の大部分の表現は元文章中で統語的關係を持つ表現に基づいて生成されていることが分かった。要約と元文章で表現が異なる場合の言い換え操作についても、それらがどのような特質を持つかを分類し、さらに、予備実験で仮定した文結合操作が、社説の要約に対しても重要な働きを示すことを確認した。

第三として、要約作成の操作として頻繁に行われる、文を簡略にする操作を計算機上に実装する実験を行った。実装は、人間が作成した要約とその元文の対応付けを行い、その結果に基づいて生成されたものを教示データとして機械学習を行う手法を用いた。その結果、言語情報を複合的に用いることにより、人間が行う要約操作に関する知識を表現可能なことを示した。

結論として、本論文では、人間が作成した要約文生成に関わる操作を、テキスト分析で用いられる概念に基づき、その操作を分類し、特徴を分析した。さらに、そのような要約操作の一つを計算機上で実装し、その有効性を確認した。

| | |
|----|-------|
| 氏名 | 竹内 和広 |
|----|-------|

(論文審査結果の要旨)

平成14年1月24日に開催した公聴会の結果を参考に平成14年2月13日に本博士論文の審査を行った。以下のとおり、本博士論文は、提案者が独立した研究者として、研究活動を続けていくための十分な素養を備えていることを示すものと認める。

竹内 和広は、本博士論文において、文書要約において人間が行っていると考えられる種々の操作を分類検討し、その要因を探るとともに、特徴を明らかにして、要約研究に対する新たな視点を提供した。本論文では、以下のような3点から要約操作を分析し、一部計算機による実装を行ってその有効性を確認した。

1. 人間が行っている要約を要約の元となった文章のテキスト構造との比較をもとに分析する実験を行った。新聞の報道記事のテキスト構造を手で分析し、人間の分析に一致が見られる文間のつながりの良さに対してどのような言語的手がかりが貢献するかを機械学習の手法を用いて検討した。また、テキスト構造解析を行った記事に対しての要約を作成し、人間の分析に一致が見られる文のつながりの良さは、文の抽出よりもむしろ、要約文を生成するための操作のひとつ、文の結合操作と関連していることを指摘した。
2. 人間が要約を作成する操作を詳細に分析した。要約作成者が作成した要約を対象としたテキスト分析を行い、要約文で使われている表現が要約元文章のどの部分から生成されたかを自動的に対応付ける手法を提案した。その結果、要約の大部分の表現は元文章中で統語的關係を持つ表現に基づいて生成されていることが分かった。要約と元文章で表現が異なる場合の言い換え操作についても、それらがどのような特質を持つかを分類し、文結合操作が、社説の要約に対しても重要な働きを示すことを確認した。
3. 要約作成の操作としてよく使われる、文の簡略操作を計算機上に実装する実験を行った。実装は、人間が作成した要約とその元文の対応付けを行い、それに基づいて生成されたものを教示データとして機械学習を行う手法を用いた。その結果、言語情報を複合的に用いることにより、人間が行う要約操作に関する知識を表現可能なことを示した。

文書要約における人間の操作を詳細に分析して、自動化の可能性とその手法についての提案を行った本研究は、独創性が高く、今後の要約研究に大いに貢献すると評価する。

よって、本論文は博士(工学)の学位論文として価値あるものと認める。