

平成21年度 情報工学コース卒業研究報告要旨

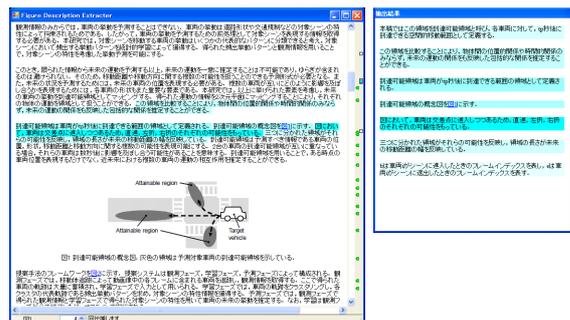
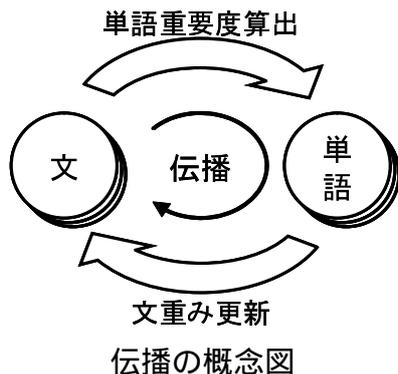
渡邊 研究室	氏 名	竹 島 亮
卒業研究題目	論文調査支援のための図表説明文の抽出	

近年インターネットの普及により、大量の論文を容易に入手することが可能となった。論文調査のために論文の入手にかかる時間は大幅に減ったが、人が論文を理解するのにかかる時間は未だ長く、改善することができれば論文調査をより効率的に行うことができる。自動要約など、文章を効率よく読む事を支援する研究が古くから行われてきたが、単純な要約では論文調査において必要とされる情報が含まれていない場合がある。本研究では、ユーザの意図を反映した論文調査支援を目的とする。

提案手法では、論文中に存在する図や表に着目し、図や表の理解を促進する文集合(図表説明文)を提示する。処理の手順や評価結果など、論文の重要な要素は図や表として表されていることが多い。なぜなら、内容を文章のみでなく図や表を用いて説明することで、大まかな内容を瞬時に読者に対して伝えることが容易になるためである。図や表は論文中で目立ち見つけることが容易であり、即座に大まかな意味を得ることができるため、知りたい情報の効率的な探索のために有用である。また、文章を読む上で先に全体像を把握しておくことが、文章の深い理解の助けになることが知られている。重要な内容を表す図や表を理解することは、論文の全体像の把握につながり、論文の深い理解を助けることができる。図や表を理解することは論文全体の理解を促進させるが、図や表は主に記号から構成されており、それだけで図や表が持つ意味をすべて理解するのは困難である。本研究では、論文中から図表説明文を抽出し、論文読解を支援する。

図表説明文の抽出は、論文中から目的に合う文を抽出するという点で、一種の要約と捉えることができる。要約のための重要文を抽出する手法に関する研究は多くなされてきたが、本研究の目的を達成するためには同じ論文でも画像によって抽出文を大きく変えなければならず適していない。文と単語は相互に依存しており、複数の類似した文に含まれる単語は類似した意味を持つことが知られている。提案手法では、この考え方を拡張し、文に与えた重みを単語を介し他の文に伝播させることで、図表説明文に適した文集合を抽出する。

提案した手法をもとにプロトタイプ・システムを構築した。プロトタイプ・システムを用いた評価実験の結果から、提案手法を用いることで論文中の重要な内容を早く見つけ出すことが可能となることを確認した。また、アンケートの結果から、本手法の論文読解支援に対する効果を確認した。



プロトタイプ・システム