

平成18年度 情報工学コース卒業研究報告要旨

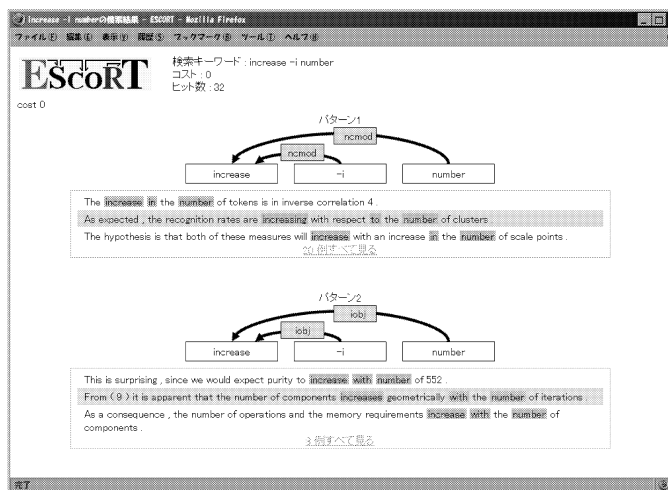
石川・松原 研究室	氏 名	江 川 誠 二
卒業研究題目	検索キーワード間の構文的関係を考慮する 英文用例検索システム	

大規模なコーパスを効果的に活用するためには、コーパス検索環境の利用が不可欠である。これまでに、依存構造を用いたコーパス検索手法が提案されている。この手法では、キーワード系列をクエリとして受け取り、あらかじめ依存構造が付与されたコーパスを参照しながら、キーワードをクエリ中の順で含むコーパスの各文において、キーワードが形成する依存構造パターンを同定し、文をパターン別に分類する。ユーザは、キーワードを入力するだけで依存構造を考慮した検索を行うことができる。

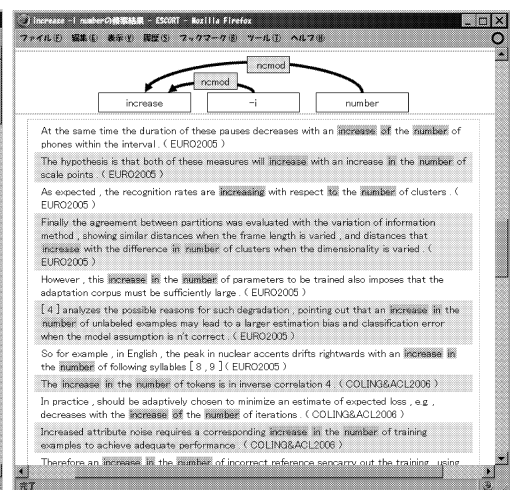
しかしこの手法では、各キーワード間の依存関係の有無のみを考慮し、その種類は利用しておらず、関係の種類に応じた分類はできない。

そこで本論文では、依存関係の種類を活用してコーパス中の文进行分类する方法を提案する。従来手法では、依存構造パターンを生成する際にキーワード間の依存関係の向きの情報のみを使用していた。これに対して、本手法では、向きに加えて、依存関係の種類の情報まで考慮してパターンの生成を行う。これにより、より細かく分類された検索結果をユーザに提示でき、ユーザの用例発見をスムーズにできる可能性がある。

本手法の有効性を確かめるため、本手法を用いて、英文用例検索システム ESCORT を実装した。システムの動作例により、本手法が従来手法に比べて、英文用例を有効に分類することを確かめた。今後の課題として、依存関係の種類と文の分類との関係を調べ、細かく分類されすぎている依存関係の種類をまとめることや、依存構造解析の誤った文を自動で検出し排除することなどが挙げられる。



英文用例検索システム ESCORT



パターンに属する用例の一覧表示

学会発表実績等

- 平成18年度電気関係学会東海支部連合大会 (2006.9)
“ラベル付き依存関係に基づく用例文検索”
- 言語処理学会第13回年次大会 (2007.3 発表予定)
“ラベル付き依存関係に基づく英文用例検索システム”