

平成19年度 福井大学工学研究科予算重点配分に対する報告書

研究課題	構文知識による語の機能を付加した英文検索手法 —中上級者のための科学技術英文作成支援システムの検討—
種別	萌芽的研究
研究代表者の所属・職・氏名	知能システム工学専攻・教授・小倉久和
研究分担者の所属・職・氏名	知能システム工学専攻・研究補佐員・諏訪いずみ 知能システム工学科・卒業研究生・中村奈津希
研究成果の概要	<p>英語で論文を作成する機会が増加しているが、中上級者でも適切な単語や言い回しを選ぶのに苦勞する場合が多い。たとえば、Roget の類語辞典は適切な動詞を選ぶのには欠かせないし、動詞と副詞の適切な組み合わせは研究社の新編英和活用大辞典などでチェックする必要がある。Webster などの英語辞書でも用例を確認する必要がある。近年の電子辞書はこれらの辞書の検索を統合的に行うことができ、論文作成では必須のアイテムである。しかし、これらの辞書はどちらかといえば文系寄りであり、理工系の学術論文でしばしば使われる表現でも採録されていないことがある。実際には、英文論文を読んだときに気付いた表現や常套句などをメモしておき、それを参照しながら論文を作成することになる。</p> <p>現在、全文データベースとしてのコーパスが入手しやすくなり、また、インターネットには多数の理工系学術英語論文が公開されている。それらをデータベースとして豊富な文例集として活用することができる。本プロジェクトでは、専門分野の学術論文を主としたコーパスあるいはインターネットの論文を利用し、辞書による単語検索と語の文法的構文機能を利用した用例検索手法を検討・実装することを試みた。シソーラスはフリーの WordNet を利用して類義語、関連語を検索する機能を検討し、文例の構文解析は MontyLingua を利用した。</p> <p>このシステムには、論文作成中にある内容を表現するのに用いた動詞や名詞などの単語について違和感がある場合、類語や関連語を検索し、さらにもとの単語や類語・関連語による文例の検索をする機能を実装した。ところで、いくつもの語に対して単純に文例を検索すると意図しない文例が多数検索されてしまう。本システムでは、文例検索において、多義単語についてはその品詞を特定でき、他の語との組合せについて、動詞では適切な副詞や前置詞など、名詞では適切な動詞や形容詞、前置詞などとの組合せも指定して検索結果を表示する。たとえば、“innovation＋動詞”や“analysis＋形容詞”のように主題となる名詞を主語にとる動詞や名詞を修飾する形容詞、“keep＋副詞”のように動詞を修飾する副詞にはどのようなものがあるか、といった例文検索が可能となる。これは、文字列の一致によって検索した文例を構文解析することにより、このような語の品詞や他の語との組合せの特徴を用いて適切な文例のみを抜き出している。</p> <p>以上により、利用者の意図する用例の絞り込みを多様な角度から行うことができ、中上級者向けの英文検索機能を実現できる可能性を示すことができた。</p>

整理 番号	
----------	--

(事務で記載)

(/)

	<p>本年度は、形態素解析・構文解析ツールを使用して、単語の品詞を用いた検索機能を実現したが、実際に使うためにはユーザインタフェースを向上させる必要がある。検索された例文中の語を選択すると同義語・類義語を検索するリンク機能、例文の構文木を表示する構文解析機能、といった、使い勝手のよいインタフェースが欠かせない。</p> <p>今後の研究課題であるが、日本語シソーラスと英語シソーラスの対応づけをすることが、支援機能を向上させるためには欠かせない。実際に語を検索するためには、日本語単語と英語単語の意味関係のある程度把握している必要がある。しかし、初級者はもちろんであるが、中上級者であっても、専門分野から少し離れた近隣の分野では、その分野の用語について日本語と英語との意味関係を把握するのは困難である。このようなことは、共同研究をしていて論文を執筆するときにはしばしば経験する。日本語の意味関係に基づくシソーラスと英語の意味関係に基づくシソーラスは、当然ながら 1 対 1 に対応するものではない。どのようにして対応させるか検討することは、今後の重要な課題の 1 つである。</p>																											
配分額及び経費の支出額内訳	<table border="0"><tr><td>配分研究費</td><td style="text-align: right;">計</td><td></td></tr><tr><td>支出項目</td><td></td><td></td></tr><tr><td>学会論文発表・参加費・旅費</td><td style="text-align: right;">約</td><td>円</td></tr><tr><td>科学技術日英・英日コーパス辞典</td><td style="text-align: right;">約</td><td>円</td></tr><tr><td>EDR 専門用語辞書</td><td style="text-align: right;">約</td><td>円</td></tr><tr><td>EDR 英語共起辞書</td><td style="text-align: right;">約</td><td>円</td></tr><tr><td>ノートブックパソコン</td><td style="text-align: right;">約</td><td>円</td></tr><tr><td>データバックアップメディア</td><td style="text-align: right;">約</td><td>円</td></tr><tr><td style="text-align: right;">計</td><td></td><td>円 (不足分は、研究基盤経費から補填した)</td></tr></table>	配分研究費	計		支出項目			学会論文発表・参加費・旅費	約	円	科学技術日英・英日コーパス辞典	約	円	EDR 専門用語辞書	約	円	EDR 英語共起辞書	約	円	ノートブックパソコン	約	円	データバックアップメディア	約	円	計		円 (不足分は、研究基盤経費から補填した)
配分研究費	計																											
支出項目																												
学会論文発表・参加費・旅費	約	円																										
科学技術日英・英日コーパス辞典	約	円																										
EDR 専門用語辞書	約	円																										
EDR 英語共起辞書	約	円																										
ノートブックパソコン	約	円																										
データバックアップメディア	約	円																										
計		円 (不足分は、研究基盤経費から補填した)																										
	<p>関連業績</p> <ol style="list-style-type: none">1. Analysis of command frequency and command sequence grammar in IDS Zhou J., Shirai H., Kuroiwa J., Odaka T., Ogura H. Proceedings of the 2008 IEEE Conference on Soft Computing in Industrial Applications (2008.6)2. 検索者の意図を反映した英文検索システムの検討と実装 中村奈津希, 諏訪いづみ, 白井治彦, 黒岩丈介, 小高知宏, 小倉久和 平成 20 年度電気関係学会北陸支部連合大会 (2008.9)3. 可視化したシソーラスを用いた段落一貫性評価指標の解析 板倉由知, 白井治彦, 黒岩丈介, 小高知宏, 小倉久和 平成 20 年度電気関係学会北陸支部連合大会 (2008.9)4. 単語のセマンティック距離を用いた英語作文の数量化評価 周健, 板倉由知, 小高知宏, 小倉久和 平成 20 年度電気関係学会北陸支部連合大会 (2008.9) <p>受賞</p> <ol style="list-style-type: none">5. ふくいソフトウェアコンペティション2008 なつきんぐの英単語発ケンジャー (中村奈津希) 福井県ソフトウェア奨励賞 (2008. 11)																											

※ 記入欄の長さを適宜変更してもよいが、最大3頁を超えないこと。できるだけ簡潔に記入のこと。記述のフォントの大きさは10.5ポイント以上。

整理番号	
------	--

(事務で記載)