

平成23年度科学研究費助成事業（科学研究費補助金）実績報告書（研究実績報告書）

1. 機関番号

1	4	6	0	3
---	---	---	---	---

 2. 研究機関名 奈良先端科学技術大学院大学
3. 研究種目名 基盤研究(A) 4. 研究期間 平成23年度～平成25年度
5. 課題番号

2	3	2	4	0	0	2	0
---	---	---	---	---	---	---	---
6. 研究課題名 大域情報を利用した同時処理による自然言語解析手法の研究

7. 研究代表者

研究者番号	研究代表者名	所属部局名	職名
10211575	マツモト ユウジ 松本 裕治	情報科学研究科	教授

8. 研究分担者(所属研究機関名については、研究代表者の所属研究機関と異なる場合のみ記入すること。)

研究者番号	研究分担者名	所属研究機関名・部局名	職名
90311589	シンボ マサシ 新保 仁	情報科学研究科	准教授
80379528	アサハラ マサユキ 浅原 正幸	情報科学研究科	助教
60581329	コマチ マモル 小町 守	情報科学研究科	助教

9. 研究実績の概要

下欄には、当該年度に実施した研究の成果について、その具体的内容、意義、重要性等を、交付申請書に記載した「研究の目的」、「研究実施計画」に照らし、600字～800字で、できるだけ分かりやすく記述すること。また、国立情報学研究所でデータベース化するため、図、グラフ等は記載しないこと。

従来、自然言語処理分野で行われていた様々なレベルの解析を独立に行なうのではなく、同時に行なう手法、さらに、局所的な情報だけに頼るのではなく、大域的な情報を利用した解析手法を実現することを目的として研究を行った。形態素解析およびその上位の固有表現認識や統語解析との同時最適化を実現するため、本年度は、機械学習に基づく形態素解析と浅い統語解析のパラメータを同時学習する手法を提案した。係り受け解析と並列構造解析の同時処理については、並列構造を知ることが係り受け解析の誤り訂正にどの程度貢献できるかを実験によって確認し、効果的なアノテーション作業への道筋を明らかにした。述語項構造解析については、大規模コーパスから得られる述語の項に関する情報の利用について研究を行った。また、複数の項の値の同定を同時に行なう手法について研究を行い、成果を得た。単語の使用文脈の関係に基づく意味的類似度の計算については、文脈ベクトルに基づく類似度として類似度行列上のカーネルを用いているが、ベクトルが高次元の場合に生じる問題としてハブとなる点の存在が問題になることが明らかになった。その性質の解明について研究を行った。

これらの研究の基本データとなるタグ付きコーパスを構築するため、日本語係り受けと述語項構造のアノテーション作業を行った。また、タグ付きコーパスを格納するコーパス管理ツールに次のような機能拡張を行った。一つは、係り受け解析と並列構造解析を重ねて表示するインタフェースの構築、もう一つは、このツールから係り受け解析システムを呼び出して文あるいは文の一部を再解析させる機能である。

10. キーワード

- | | | | |
|------------|-----------|----------|-------------|
| (1) 自然言語処理 | (2) 形態素解析 | (3) 統語解析 | (4) 述語項構造解析 |
| (5) 機械学習 | (6) 言語資源 | (7) 同時処理 | (8) 大域情報 |

11. 現在までの達成度

下欄には、交付申請書に記載した「研究の目的」の達成度について、以下の区分により自己点検による評価を行い、その理由を簡潔に記述すること。また、国立情報学研究所でデータベース化するため、図、グラフ等は記載しないこと。
<区分>①当初の計画以上に進展している。②おおむね順調に進展している。③やや遅れている。④遅れている。

(区分)	②おおむね順調に進展している。
(理由)	本年度予定していた研究の多くについて、計画通りの進展があった、一部の研究、例えば類似度計算、については、外部発表を行なうまでの成果を得ていないが、解決すべき問題を明らかにするという進捗があった。一方、述語項構造解析については、述語間の関係や複数の項の同時推定、項同定とラベル同定の同時処理など、当初計画以上の研究を行い、成果を挙げることができた。

12. 今後の研究の推進方策

本研究課題の今後の推進方策について簡潔に記述すること。研究計画の変更あるいは研究を遂行する上での問題点があれば、その対応策なども記述すること。また、国立情報学研究所でデータベース化するため、図、グラフ等は記載しないこと。

今年度の研究では取り扱わなかった浅い統語解析と文全体の解析、全域的な情報と局所的な情報の相補的あるいは同時利用、統語解析と述語項構造解析等の意味解析との同時解析へ研究を進展させて行く。本研究や関連研究で利用可能なタグ付きコーパスの構築と、種々の解析の同時処理を利用した効率良いタグ付きコーパスの構築法を明らかにして行く。
--

13. 研究発表（平成23年度の研究成果）

※ 「13. 研究発表」欄及び「14. 研究成果による産業財産権の出願・取得状況」欄において記入欄が不足する場合には、適宜記入欄を挿入し、それによりページ数が増加した場合は、左端を糊付けすること。

〔雑誌論文〕 計（ 1 ）件 うち査読付論文 計（ 1 ）件

著 者 名	論 文 標 題				
Katsumasa Yoshikawa, Sebastian Riedel, Tsutomu Hirao, Masayuki Asahara, and Yuji Matsumoto	Coreference Based Event-Argument Relation Extraction on Biomedical Text				
雑 誌 名	査読の有無	巻	発 行 年		最初と最後の頁
Journal of Biomedical Semantics	有	2	2	0 1 1	—
掲載論文のDOI（デジタルオブジェクト識別子）					
10.1186/2041-1480-2-S5-S6					

著 者 名	論 文 標 題				
雑 誌 名	査読の有無	巻	発 行 年		最初と最後の頁
掲載論文のDOI（デジタルオブジェクト識別子）					

著 者 名	論 文 標 題				
雑 誌 名	査読の有無	巻	発 行 年		最初と最後の頁
掲載論文のDOI（デジタルオブジェクト識別子）					

〔学会発表〕計（ 5 ）件 うち招待講演 計（ 0 ）件

発表者名	発表標 題		
Ai Azuma, Yuji Matsumoto	Multilayer sequence labeling		
学 会 等 名	発表年月日	発 表 場 所	
Conference on Empirical Methods in Natural Language Processing	2011年7月28日	Edinburgh, U.K.	

発表者名	発表標 題		
Yuta Hayashibe, Mamoru Komachi and Yuji Matsumoto	Japanese Predicate Argument Structure Analysis Exploiting Argument Position and Type		
学 会 等 名	発表年月日	発 表 場 所	
Proceedings of the 5th International Joint Conference on Natural Language Processing	2011年11月9日	Chiang Mai, Thailand	

発表者名	発表標 題		
Katsumasa Yoshikawa, Masayuki Asahara and Yuji Matsumoto	Jointly Extracting Japanese Predicate-Argument Relation with Markov Logic		
学 会 等 名	発表年月日	発 表 場 所	
Proceedings of the 5th International Joint Conference on Natural Language Processing	2011年11月9日	Chiang Mai, Thailand	

発表者名	発表標 題		
Yanyan Luo, Masayuki Asahara and Yuji Matsumoto	Dual Decomposition for Predicate-Argument Structure Analysis		
学 会 等 名	発表年月日	発 表 場 所	
Proceedings of the 7th International Conference on Natural Language Processing and Knowledge Engineering	2011年11月28日	University of Tokushima Japan	

発表者名	発表標 題		
林部祐太, 小町守, 松本裕治	文脈情報と格構造の類似度を用いた日本語文間述語項構造解析		
学 会 等 名	発表年月日	発 表 場 所	
情報処理学会研究報告 第201回自然言語処理研究会	2011年5月17日	東京大学	

〔図 書〕 計（ 1 ）件

著 者 名	出 版 社		
松本裕治（編集）	朝倉書店		
書 名		発行年	総ページ数
言語と情報科学		2011	216

14. 研究成果による産業財産権の出願・取得状況

〔出 願〕 計（ 0 ）件

産業財産権の名称	発明者	権利者	産業財産権の種類、番号	出願年月日	国内・外国の別

〔取 得〕 計（ 0 ）件

産業財産権の名称	発明者	権利者	産業財産権の種類、番号	取得年月日	国内・外国の別
				出願年月日	

15. 備考

※ 研究者又は所属研究機関が作成した研究内容又は研究成果に関するwebページがある場合は、URLを記載すること。

--