

平成21年度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 特別研究員奨励費 4. 研究期間 平成20年度～平成21年度
5. 課題番号 2 0 9 5 4 5
6. 研究課題名 大規模なコーパスを用いた機械学習による名詞句の項構造解析

7. 研究代表者

研究者番号	研究代表者名	所属部局名	職名
	フリガナ コマチ マモル 小町 守	情報科学研究科	特別研究員 (DC2)

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
	フリガナ		
	フリガナ		
	フリガナ		
	フリガナ		
	フリガナ		

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。

本年度は、本研究のベースとなった事態性名詞の述語項構造解析の研究に関して、構文パターンと既存の項構造解析モデルを組み合わせた手法をまとめた論文が出版された。しかしながら、年次計画1年目の後半に予定していた大規模なコーパスからリソースを自動構築する研究が理論面においても実践面においても発展を見せ、研究計画を変更したために、本研究の2年度目に予定していた名詞句の自動項構造解析は実施しなかった。以下では研究計画を変更したあとの研究について述べる。

本年度の研究上の貢献は2つある。1つ目は、1年度目においてグラフ理論に基づく自然言語処理における半教師あり手法（ブートストラッピング）について行った研究をまとめ、発表したことである。グラフ理論的観点からブートストラッピングを解析した研究はこれまでになく、ブートストラッピングに内在する問題点を理論的に明らかにした一方、リンク解析で用いられている手法を自然言語処理にも適用し、語義曖昧性解消のタスクで高い性能が得られることを示した。さまざまなヒューリスティックを用いるブートストラッピングと比較して、使用時に決めなければならない変数の数が少ないため、扱いが容易であることも提案手法の利点の一つである。

2つ目は、ヤフー研究所との共同研究により、検索クリックスルーログをコーパスとした意味カテゴリ辞書の自動構築の研究を行い、検索クリックスルーログが意味カテゴリの獲得に有用な情報抽出源であることを示した。また、日本最大規模のウェブデータを対象として、並列分散環境でも実行できるようなスケーラブルなアルゴリズムを提案し、有効性を示した。

また、本年度は年次計画で予定していた統計的機械翻訳への応用に替え、3ヶ月アメリカ合衆国 Apple Inc. に渡航し、統計的自然言語処理（統計的かな漢字変換）に関する研究を行った。

10. キーワード

- (1) 自然言語処理 (2) 機械学習 (3) コーパス
- (4) 大規模データ (5) 半教師あり学習 (6) ブートストラッピング法
- (7) 国際研究者交流 (8) アメリカ
- (裏面に続く)

11. 研究発表（平成21年度の研究成果）

〔雑誌論文〕 計（ 3 ）件 うち査読付論文 計（ 3 ）件

著者名		論文標題			
小町守・飯田龍・乾健太郎・松本裕治		名詞句の語彙統語パターンを用いた事態性名詞の項構造解析			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
自然言語処理	有	17・1	2010	141-159	

著者名		論文標題			
小町守・工藤拓・新保仁・松本裕治		Espresso型ブートストラッピング法における意味ドリフトのグラフ理論に基づく分析			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
人工知能学会論文誌	有	25・2	2010	233-242	

著者名		論文標題			
小町守・牧本慎平・内海慶・颯々野学		ラブラシアンラベル伝播による検索クリックスルーからの意味カテゴリ獲得			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
人工知能学会論文誌	有	25・1	2010	196-205	

〔学会発表〕 計（ 1 ）件 うち招待講演 計（ 0 ）件

発表者名		発表標題		
Mamoru Komachi, Shimpei Makimoto, Kei Uchiumi, Manabu Sassano		Learning Semantic Categories from Clickthrough Logs		
学会等名		発表年月日	発表場所	
Joint conference of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing		2009年 8月4日	シンガポール (国:シンガポール)	

〔図書〕 計（ 1 ）件

著者名		出版社		
言語処理学会		共立出版		
書名		発行年	総ページ数	
言語処理学事典		2009	913	

12. 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

13. 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。

--