

平成18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 特別研究奨励費 4. 研究期間 平成18年度 ～ 平成18年度
5. 課題番号

1	8	・	5	0	7	2	2
---	---	---	---	---	---	---	---
6. 研究課題名 言語理論と機械学習の融合による照応解析

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	イダリュウ 飯田, 龍	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			

9. 研究実績の概要(国立情報学研究所でデータベース化するため、600字～800字で記入。図、グラフ等は記載しないこと。)

本研究では、機械翻訳や情報抽出のような自然言語処理の応用分野に必須な要素技術の一つである照応解析の問題を対象に、言語学分野で考察されてきた手がかりを効果的に学習ベースの解析手法に導入するかについて調査を行った。特に今年度の取り組みでは、ゼロ代名詞とその先行詞が同一の文内に出現する文内ゼロ照応解析の問題に焦点を当て、解析の手がかりとして有効にはたらくと考えられる文の構造情報を学習ベースの手法に効果的に導入する手法を提案した。この手法では、まず文の構造をゼロ照応の解析に役立つ粒度に抽象化し、この抽象化された構造を解析の際の素性の一つとして利用する。このような構造情報から分類に寄与する情報を学習するアルゴリズムとしてはブースティングに基づく学習アルゴリズムを利用し、実際に今回の手法が既存手法と比較して、劇的に解析精度の向上に貢献することを示した。この結果を国際学会である COLING-ACL 2007 で発表し、AFNLP Asian Best Paper Award を受賞した。

また、上述の問題を含めた述語項構造・共参照解析の問題を分析/評価/学習するには、それらの情報が付与されたタグ付与コーパスが必要となる。今年度はこの問題にも取り組み、具体的には日本語新聞記事を対象とした述語項構造・共参照タグ付与コーパスを作成した。この取り組みでは、厳密な規定のない照応・共参照の概念について既存のタグ付与コーパスを参考に、どのようにこれらの概念を扱うべきかを議論し、また実際にタグ付与を行う上で、各概念のどのような点が問題となるのかを経験的に明かにした。また、今回の作業の結果作成された大規模なタグ付与コーパスを一般に公開した(<http://cl.naist.jp/nldata/corpus/>)。

※ 成果の公表を見合わせる必要がある場合は、その理由及び差し控え期間等を記入した調書(A4判縦長横書1枚)を添付すること。

10. キーワード

- | | | |
|----------|-----------|------------|
| (1) 照応解析 | (2) ゼロ代名詞 | (3) 自然言語処理 |
| (4) | (5) | (6) |
| (7) | (8) | |
- (裏面に続く)

11. 研究発表(平成18年度の研究成果)

〔雑誌論文〕 計(4)件

著者名	論文標題		
Ryu Iida, Kentaro Inui, Yuji Matsumoto	Exploiting Syntactic Patterns as Clues in Zero-Anaphora Resolution		
雑誌名	巻・号	発行年	ページ
Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics		2006	pp.625-632

著者名	論文標題		
飯田龍, 小町守, 乾健太郎, 松本裕治	NAISTテキストコーパス: 述語項構造と共参照関係のアノテーション		
雑誌名	巻・号	発行年	ページ
情報処理学会自然言語処理研究会予稿集, NL-177-10	177・10	2007	pp.71-78

著者名	論文標題		
飯田龍, 小町守, 乾健太郎, 松本裕治	日本語書き言葉を対象とした述語項構造と共参照関係のアノテーション: NAISTテキストコーパス開発の経験から		
雑誌名	巻・号	発行年	ページ
言語処理学会第13回年次大会		2007	to appear

著者名	論文標題		
飯田龍, 乾健太郎, 松本裕治	ゼロ照応解析のための統語的パターンの学習		
雑誌名	巻・号	発行年	ページ
第5回情報科学技術フォーラム, FIT		2006	pp. 81-84

著者名	論文標題		
雑誌名	巻・号	発行年	ページ

著者名	論文標題		
雑誌名	巻・号	発行年	ページ

〔図書〕 計(0)件

著者名	出版社		
書名	発行年	総ページ数	

12. 研究成果による工業所有権の出願・取得状況

計(0)件

工業所有権の名称	発明者	権利者	工業所有権の種類、番号	出願年月日	取得年月日