

## 平成28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 特別研究員奨励費 4. 研究期間 平成27年度～平成28年度

5. 課題番号 

1	5	J	1	0	8	3	9
---	---	---	---	---	---	---	---

6. 研究課題名 語のネットワーク解析手法を用いた文書の包括的一貫性の評価

## 7. 研究代表者

研究者番号	研究代表者名	所属部局名	職名
	コバヤシ ユウタ 小林 雄太	情報科学研究科	特別研究員(DC2)

## 8. 研究分担者

研究者番号	研究分担者名	所属研究機関名・部局名	職名

## 9. 研究実績の概要

【研究の目的】本研究は統計的機械学習による、一貫性等の文書評価・意味解析に関するモデル・指標の確立である。また、論文検索等の文書評価に利用可能なソフトウェアの開発に取り組む。当初のグラフ指標のみの手法では精度が不足するため、語や文書の分散表現を用いた手法へと方針の修正を行っており、取り組んだ課題は下記の2項目である。

(1)分散表現に関する研究：文書データは本文テキストのセクション構造や、引用ネットワーク等の構造的な情報を持つため、それらを考慮した分散表現の学習法について検討を行う。

(2)分散表現を用いた応用タスクに関する研究：論文検索の応用タスクにおいて分散表現の評価を行う。

(1),(2)の詳細を以下に記す。

(1)構造化された文書として自然言語処理分野の英語論文データを対象とし、語の分散表現とグラフの分散表現を組み合わせることで論文の分散表現学習を行った。まず、文字列情報を考慮した語の分散表現学習手法を用いて、本文テキスト中の各セクションの論文での構成要素(目的・手法・結論)の予測とベクトル化を行い、本文とその構成要素を考慮した論文ベクトルを学習した。次に、グラフの分散表現手法によるグラフ指標のモデル化により上記の論文ベクトルの更新を行った。具体的には、論文の引用グラフにラベル(引用の理由)を付与し、引用関係にある論文間で構成要素のベクトルの更新を行った。

(2)まず、論文データセットを用いた、「研究目的が似ている」といった論文の構成要素に基づいた類似論文検索タスクにおいて、精度の改善を確認した。また、論文解析の国際ワークショップにおいて、直感的に文書を理解可能な引用グラフ視覚化システムの発表を行った。さらに、列挙型共引用に着目し、論文の構成要素に基づく自動評価データセットを構築した。最後に、各セクションの分散表現を用いて内容の推移を定量化し、一貫性判定モデルを構築した。

## 10. キーワード

(1) 意味解析

(2) リンク解析

(3)

(4)

(5)

(6)

(7)

(8)

## 11. 現在までの進捗状況

(区分)

(理由)

28年度が最終年度であるため、記入しない。

## 12. 今後の研究の推進方策

(今後の推進方策)

28年度が最終年度であるため、記入しない。

## 13. 研究発表 (平成 28 年度の研究成果)

〔雑誌論文〕 計(0)件/うち査読付論文 計(0)件/うち国際共著論文 計(0)件/うちオープンアクセス 計(0)件

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	国際共著	
掲載論文のDOI (デジタルオブジェクト識別子)						
オープンアクセス						

〔学会発表〕 計(2)件/うち招待講演 計(0)件/うち国際学会 計(1)件

発表者名		発表標題	
小林雄太, 新保仁, 松本裕治		論文のテキストと引用グラフを考慮した分散表現学習	
学会等名	発表年月日	発表場所	
NLP若手の会 (YANS) 第11回シンポジウム	2016年08月28日 ~ 2016年08月30日	和歌山県白浜 ホテルシーモア	

発表者名		発表標題	
Yuta Kobayashi, Hiroki Teranishi, Masashi Shimbo, and Yuji Matsumoto		Learning scientific paper representations from text and citation graphs	
学会等名	発表年月日	発表場所	
International Workshop on Scientific Document Analysis (SCIDOCA 2016) (国際学会)	2016年11月15日 ~ 2016年11月16日	神奈川県日吉 慶應義塾大学日吉キャンパス	

〔図書〕 計(0)件

著 者 名		出 版 社	
書 名		発行年	総ページ数

## 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. 科研費を使用して開催した国際研究集会

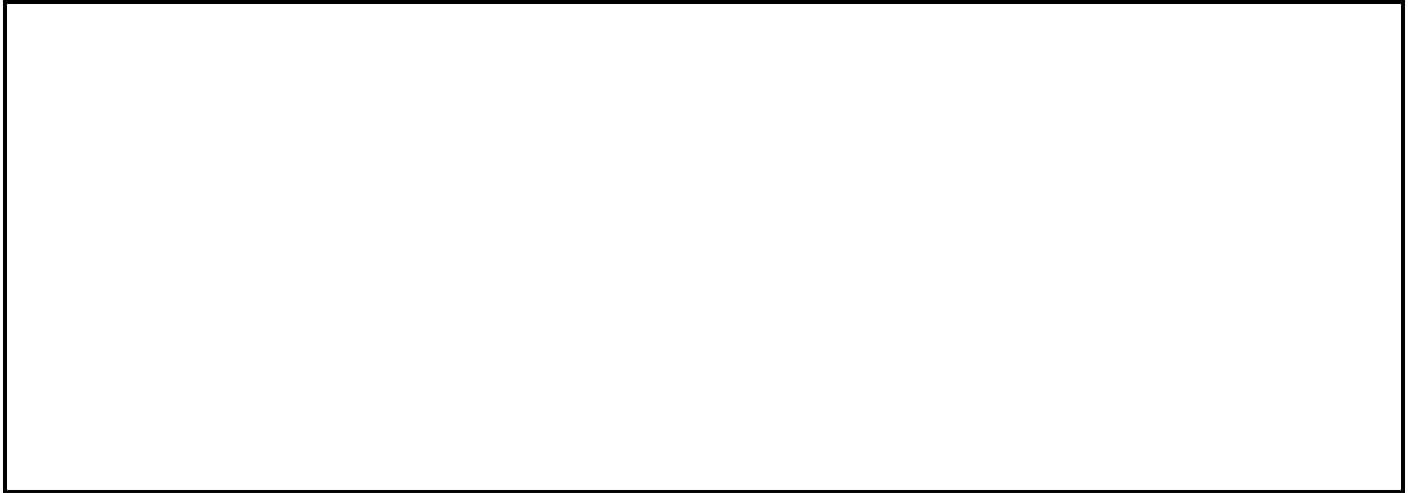
〔国際研究集会〕 計(0)件

国際研究集会名	開催年月日	開催場所

## 16. 本研究に関連して実施した国際共同研究の実施状況

(1) 国際共同研究： -

17. 備考

A large empty rectangular box with a black border, intended for handwritten notes or additional information.