京都大学人文科学研究所共同研究実績・活動報告書

（3 年計画の 3 年目）

1. 研究課題

東アジア古典文献コーパスの実証研究

Empirical Research on Digital Analysis of Classical Chinese Texts

2. 研究代表者氏名

安岡孝一

Koichi Yasuoka

3. 研究期間

2016 年 04 月 - 2019 年 03 月（3 年度目） 期間は 2020 年 3 月までC班として 1 年延長

4. 研究目的

2010 年以来、我々が構築を続けてきた漢文コーパスは、MeCab を用いた形態素解析手法を、漢文処理に適用するものである。この漢文コーパスでは 4 階層の品詞体系を採用しており、その第 2 層は「名詞」「代名詞」「数詞」「動詞」「前置詞」「副詞」「助動詞」「助詞」「感嘆詞」の 9 種類の品詞で構成される。すなわち我々は、従来の漢文文法等で見られた「形容詞」を廃止しているのだが、これが動詞類全体にどのような影響を及ぼしているのかは、必ずしも十分に検討できていない。 本共同研究では、漢文コーパスにおける動詞類の実証研究をおこなう。すなわち、実際のコーパスにおいて「動詞」「前置詞」「副詞」「助動詞」の 4 つのふるまいを研究し、さらに下層の意味素性と小素性についても、現在の品詞体系の妥当性を検証する。

5. 本年度の研究実施状況

平成 30 年度は、Universal Dependencies を用いて漢文を記述する手法に対し、Edwin George Pulleyblank の『Outline of Classical Chinese Grammar』の各例文を記述することで、その有効性の検証をおこなった。この結果、この手法の有効性に、かなりの確信を持ったことから、『孟子』『論語』『大学』『中庸』の全文を Universal Dependencies で記述すべく、調査と作業をおこなっている。また、この手法によって構築した古典中国語の依存文法解析エンジンの能力を検証すべく、大学入試センター試験『国語』の問題のうち、漢文の本文部分に対して、どの程度の自動解析がおこなえるかを検証中である。

6. 研究成果の概要

古典中国語(漢文)における形態素解析手法を発展させて、さらに文法解析へと展開すべく、数々の手法を検討した。具体的には、Chomsky 流の文法解析手法およびその亜種は古典中国語への適用が難しく、Ｍｅ ｌ ь ч ｕ ｋ 流の依存文法(Dependency Grammar)による解析手法が、古典中国語においては非常に有用であることが明らかとなった。この知見にもとづき、現代的な依存文法記法である Universal Dependencies を用いて、『孟子』『論語』『大學』『中庸』の文法記述をおこなって、デジタル・コーパスの形

で WWW で公表した。また、『孟子』『論語』『大學』『中庸』等を機械学習した依存文法解析エンジンを制作した上で、大学入試センター試験『国語』の問題のうち、漢文の本文部分に対して、どの程度の自動解析がおこなえるかを検証中である。

7. 本年度の研究実施内容

2018-04-20 2018 年度活動方針
2018-05-12 MeCab-Kanbun to CoNLL-U
2018-06-08 Universal Dependencies による『Outline of Classical Chinese Grammar』の例文検討
2018-06-22 Universal Dependencies による『Outline of Classical Chinese Grammar』の例文検討
2018-07-06 Universal Dependencies による『Outline of Classical Chinese Grammar』の例文検討
2018-07-20 Universal Dependencies による『Outline of Classical Chinese Grammar』『孟子』の例文検討
2018-08-18 人文科学とコンピュータ第 118 回研究発表会『古典中国語 UD コーパスの IPFS を用いた表現の試み』
2018-09-07 M. Bernhard Karlgren「LE PROTO-CHINOIS, LANGUE FLEXIONELLE」再検討
2018-09-21 「古典中国語(漢文)の依存文法解析と直接構成素解析」ゲラチェック
2018-10-12 「漢文の依存文法解析と返り点の関係について」ゲラチェック
2018-10-26 UDPipe Visualizer
2018-11-16 Universal Dependencies による『大学』の例文検討
2018-12-07 Universal Dependencies による『孟子』の例文検討
2018-12-21 Universal Dependencies による『論語』の例文検討
2019-01-11 Universal Dependencies による『孟子』の例文検討
2019-01-25 センター漢文を MeCab+UDPipe で読む

8. 共同研究会に関連した公表実績

守岡知彦「古典中国語 UD コーパスの IPFS を用いた表現の試み」(情報処理学会研究報告, Vol.2018-CH-118 (2018 年 8 月), No.6, pp.1-7) 安岡孝一「Universal Dependencies にもとづく古典中国語(漢文)の依存文法解析」(センター研究年報 2018 (2018 年 10 月)) 安岡孝一「古典中国語(漢文)の依存文法解析と直接構成素解析」(漢字文献情報処理研究, 第 18 号(2018 年 10 月), pp.56-62) 安岡孝一「漢文の依存文法解析と返り点の関係について」(日本漢字学会第 1 回研究大会予稿集 (2018 年 12 月), pp.33-48)

9. 研究班員

所内
Christian Wittern、池田巧、守岡知彦、白須裕之
学外
山崎直樹(関西大学)、二階堂善弘(関西大学)、師茂樹(花園大学)、鈴木慎吾(大阪大学)

10. 共同利用・共同研究の参加状況

区分	機関数	参加人数				延べ人数			
		総計	外国人	大学院生	若手研究者	総計	外国人	大学院生	若手研究者
所内	1	6 (1)	1 (0)	0	1 (1)	40 (11)	11 (0)	0	11 (11)
学内		0	0	0	0	0	0	0	0
国立大学	1	1 (0)	0	0	0	5 (0)	0	0	0
公立大学		0	0	0	0	0	0	0	0
私立大学	2	3 (0)	0	0	0	15 (0)	0	0	0
大学共同利用機関法人		0	0	0	0	0	0	0	0
独立行政法人等公的研究機関		0	0	0	0	0	0	0	0
民間機関		0	0	0	0	0	0	0	0
外国機関		0	0	0	0	0	0	0	0
その他		0	0	0	0	0	0	0	0
計	4	10 (1)	1 (0)	0	1 (1)	60 (11)	11 (0)	0	0

※( )内には、女性数を記載

11. 本年度 共同利用・共同研究を活用して発表された論文数
参加研究者がファーストオーサーであるものを対象

総論文数	6(6)
国際学術誌に掲載された論文数	0(0)

※( )内には、拠点外の研究者による成果(内数)を記載

12. 費目の 30%を超える大幅な変更があった場合の変更理由
なし

13. 次年度の研究実施計画
本来は平成 30 年度で終了予定だったが、『論語』および『礼記』などに関する Universal Dependencies の適用事例を、もう少し拡大して研究したいことから、1 年の延長をおこないたい。

14. 次年度の経費

なし

15. 研究成果公表計画および今後の展開等

最終報告書に記載