

10. キーワード

- | | | | |
|------------|------------|--------------|----------|
| (1) 音声翻訳 | (2) 機械翻訳 | (3) 自然言語処理 | (4) 機械学習 |
| (5) アルゴリズム | (6) 音声言語処理 | (7) ソフトウェア工学 | (8) |

11. 現在までの進捗状況

(区分)(2) おおむね順調に進展している。

(理由)

データ収集の面では、今年度実施した作業により、日本語から英語への一方向の設定に関して比較的大規模な同時通訳コーパスを収集することができた。このデータは今後の研究を遂行するための重要な材料となると考えられる。本年度の研究は、全体として翻訳アルゴリズムの理論面を重視したものとなった。代表としては言語処理学会にて発表を行った翻訳アルゴリズムがあり、これは音声翻訳に限らず一般の機械翻訳について精度を向上させるアルゴリズムである。また、音声という限られた範囲の入出力だけでなく、より幅広いデータを対象として音声翻訳の知見を応用することについても考慮しており、本年度はその一つとしてプログラミング言語を入力、または出力とした翻訳手法についても研究を行い、一定の知見を得た。

12. 今後の研究の推進方策

(今後の推進方策)

データの収集に関しては今後も継続して行う。特に現在は日本語から英語、英語から日本語の各設定についてのデータ以外は取り扱っておらず、また本年度に対象としたデータも一般講演であり、より一般的な同時通訳、及び機械翻訳のアルゴリズムを研究するにあたって、他の言語対や対象分野に関するデータも収集する必要があると考えている。また、ここ2年ほどで機械翻訳に深層学習を取り入れた手法の研究が盛んに行われ始めており、特に2014年、2015年に多くの新たな手法が提案された。同時通訳に関しても深層学習に関する手法を導入する余地はあると考えられるため、これらについての研究も追加で行う予定である。現在のところ、深層学習を用いた機械翻訳手法は研究の進展が浅く、基礎的なアルゴリズムが提案されるに留まっている。このため、本研究でもまずは基礎的な知見を得るための手法の提案や実験を中心に考察を進める必要があると考えている。

13. 研究発表(平成27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(8)件/うち招待講演 計(0)件/うち国際学会 計(4)件

発表者名		発表標題	
Yusuke Oda, Graham Neubig, Sakriani Sakti, Tomoki Toda, and Satoshi Nakamura		Ckylark: A More Robust PCFG-LA Parser	
学会等名	発表年月日	発表場所	
In Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies(国際学会)	2015年05月31日 ~ 2015年06月05日	Denver, USA	

発表者名		発表標題	
Yusuke Oda, Graham Neubig, Sakriani Sakti, Tomoki Toda, and Satoshi Nakamura		Syntax-based Simultaneous Translation through Prediction of Unseen Syntactic Constituents	
学会等名	発表年月日	発表場所	
In Proceedings of the Joint Conference of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing(国際学会)	2015年07月26日 ~ 2015年07月31日	Beijing, China	

発表者名	発表標題	
Yusuke Oda, Hiroyuki Fudaba, Graham Neubig, Hideaki Hata, Sakriani Sakti, Tomoki Toda, and Satoshi Nakamura	Learning to Generate Pseudo-code from Source Code using Statistical Machine Translation	
学会等名	発表年月日	発表場所
In Proceedings of 30th IEEE/ACM International Conference on Automated Software Engineering(国際学会)	2015年11月09日 ~ 2015年11月13日	Lincoln, USA

発表者名	発表標題	
Hiroyuki Fudaba, Yusuke Oda, Koichi Akabe, Graham Neubig, Hideaki Hata, Sakriani Sakti, Tomoki Toda, and Satoshi Nakamura	Pseudogen: A Tool to Automatically Generate Pseudo-code from Source Code	
学会等名	発表年月日	発表場所
In Proceedings of 30th IEEE/ACM International Conference on Automated Software Engineering(国際学会)	2015年11月09日 ~ 2015年11月13日	Lincoln, USA

発表者名	発表標題	
札幌 寛之, 小田 悠介, Graham Neubig, 畑 秀明, Sakriani Sakti, 戸田 智基, and 中村 哲	機械翻訳を用いた擬似コード生成による学習者支援	
学会等名	発表年月日	発表場所
教育システム情報学会 第40回全国大会	2015年09月01日 ~ 2015年09月03日	徳島大学, 徳島県徳島市

発表者名	発表標題	
札幌 寛之, 小田 悠介, Graham Neubig, 吉野 幸一郎, and 中村 哲	統計的機械翻訳による自然言語からソースコードへの変換の検討	
学会等名	発表年月日	発表場所
NLP若手の会 第10回シンポジウム	2015年09月03日 ~ 2015年09月05日	和倉温泉 ホテル海望, 石川県七尾市

発表者名	発表標題	
札幌 寛之, 小田 悠介, Graham Neubig, 吉野 幸一郎, and 中村 哲	統計的機械翻訳を用いた自然言語からのソースコード生成	
学会等名	発表年月日	発表場所
言語処理学会 第22回年次大会	2016年03月08日 ~ 2016年03月10日	東北大学, 宮城県仙台市

発表者名	発表標題	
小田 悠介, 工藤 拓, 中川 哲治, and 渡辺 太郎	複数の事前並べ替え候補を用いた句に基づく統計的機械翻訳	
学会等名	発表年月日	発表場所
言語処理学会 第22回年次大会	2016年03月08日 ~ 2016年03月10日	東北大学, 宮城県仙台市

〔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(1)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
構文評価装置、翻訳装置、構文評価方法及び構文評価プログラム	小田 悠介, ほか	同左	特許、学内手続中	2016年02月03日	外国

(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

Ckylark - An Latent Annotated PCFG Parser http://odaemon.com/?page=tools_ckylark
--