

先端科学技術研究科 修士論文要旨

所属研究室 (主指導教員)	知能コミュニケーション (中村 哲 (教授))		
学籍番号	2211018	提出日	令和 6年 1月 19日
学生氏名	石川 隆太		
論文題目	文内コンテキストを利用した分割統治ニューラル機械翻訳		
要旨			
<p>グローバル化に伴い、翻訳の需要が高まっており、ニューラル機械翻訳(NMT)が注目されている。NMTは質の高い訳文の生成が期待できる一方で、入力される文が長くなると訳抜けなどの問題が発生し、翻訳の品質が低下することがある。この課題に対し、統計的機械翻訳とニューラル機械翻訳それぞれにおいて、長文を短いセグメントに分割して翻訳し、並べ替えて繋げる分割統治的手法が提案されている。</p> <p>本研究では、既存の分割統治NMTに基づく、新しい分割統治手法を提案する。この手法では、(1) 構文解析によってSとラベル付けされた節を等位接続詞を基準として分割し、(2) 分割された各節を、その文中コンテキストを考慮できる形で、節翻訳用に調整されたモデルを用いて翻訳し、(3) 翻訳された節を別のsequence-to-sequenceモデルを使用して文翻訳を得るために集約する。ASPECを使用した英語から日本語への翻訳の実験結果から、特に41個以上の英単語を含む長い入力文において、事前訓練された多言語BARTモデルを使用したベースラインのTransformerベースのNMTと比較して、提案された手法がより良いBLEUをもたらすことを確認した。</p>			