

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)	自然言語処理学 (渡辺 太郎 (教授))		
学籍番号	2111203	提出日	令和 5年 1月 19日
学生氏名	早川 明男		
論文題目	Towards Personalized Text Simplification		
要旨			
<p>Text Simplification aims to rewrite texts for easier understanding while keeping their main ideas. To simplify texts, humans execute rewriting operations such as deleting unnecessary information, replacing words or phrases with simpler ones, and splitting sentences. Just as some individuals may prefer sentences with fewer words, while others may prefer sentences with simpler vocabulary, the appropriateness of simplification depends on readers. Therefore, the ideal automatic text simplification systems should be personalized, though current automatic text simplification studies focus on simplicity for native language speakers.</p> <p>Towards personalized text simplification, which can simplify text in the most appropriate way for each individual, there are challenges in terms of obtaining personal information and generating diverse sentences. In this thesis, we take two approaches as a step towards realizing personalized text simplification.</p> <p>First, we create JADES, a text simplification dataset in Japanese targeted at people with a specific attribute, namely, intermediate-level non-native speakers. Targeted simplification is conducted by an expert according to public vocabulary and grammar profiles, and JADES comprises 3,907 complex-simple sentence pairs.</p> <p>Second, we build an operation-controllable text simplification model, which can control output sentences by selecting operations, including deletion and splitting. Through operation identification using word alignment techniques, we demonstrate that our model has controllability in operation with relatively small parallel data.</p>			