Interpretable Neural Machine Translation from Translation to Post-Editing

氏　　名　出口 祥之

研 究 室 名　自然言語処理学研究室

主指導教員名（論文博士の場合は推薦教員名）　渡辺 太郎 教授

内 容 梗 概（1ページ目に収めること）

Neural machine translation (NMT) has achieved sufficient translation quality in the general domain, but not yet in the out-of-domain. Therefore, post-editing (PE), which manually corrects mistranslations, is still crucial, especially in fields where mistakes are not allowed, e.g., the medical domain. This dissertation tackles these problems from translation to post-editing using interpretable models. We firstly prevent the degradation of the translation quality in the out-of-domain. In previous work, kNN-MT adapted NMT models to various domains using the example-based approach; however, the example search is time-consuming and the decoding speed becomes two orders of magnitude slower than that of standard NMT. To improve the decoding speed of kNN-MT, we propose subset kNN-MT, which reduces the search space to the neighboring examples of the input sentence and employs the efficient computation using distance lookup table. Subset kNN-MT achieved a speed-up of up to 134.2 times and an improvement in BLEU score of up to 1.6 compared with kNN-MT in the De-En domain adaptation task.

We then aim to improve the efficiency of human PE. Previous automatic PE (APE) models attempt to correct the outputs of an MT model; however, many APE models are based on sequence generation, and thus their decisions are harder to interpret for human post-editors. We propose an edit-based PE model, which breaks the editing process into two steps, “error detection” and “error correction”. The detector model tags each MT output token whether it should be corrected and/or reordered while the corrector model generates corrected words for the spans identified as errors. Experiments on the WMT’20 En-De and En-Zh APE tasks showed that our model provides the editing process while maintaining PE performance comparable to that of a sequence-to-sequence model and outperforming previous edit-based model.