

Title: Paragraph-Parallel based Neural Machine Translation Model with Hierarchical Attention

Abstract: Neural Machine Translation (NMT) has achieved great developments in recent years, but we still have to face two challenges: establishing a high-quality corpus and exploring optimal parameters of models for long text translation. In this paper, we first attempt to set up a paragraph-parallel corpus based on English and Chinese versions of the novels and then design a hierarchical attention model for it to handle these two challenges. Our encoder and decoder take segmented clauses as input to process the words, clauses, paragraphs at different levels, particularly with a two-layer transformer to capture the context from both source and target languages. The output of the model based on the original transformer is used as another level of abstraction, conditioning on its own previous hidden states. During this process, inter-clause and intra-clause contexts from source and target sides are introduced for translation prediction and both encoder and decoder can profit from contexts in complementary ways. Experimental results show that our hierarchical attention model significantly outperforms six competitive baselines, including ensembles.