論 文 の 内 容 の 要 旨

論文題目　Mechanisms for Combining Character and Word Level Representations in Natural Language Processing
（自然言語処理における文字レベルと単語レベルの表現を組み合わせるメカニズムの研究）

氏　名　バラズテノ　ホルヘアンドレス　ジャンミシェル

One of the first steps in any Natural Language Processing task is to generate vector representations for words. Classic approaches relied on humans manually engineering features for producing them, while later ones leveraged the Distributional Hypothesis and extracted meaning representations directly from word co-occurrence statistics. Recent neural-based approaches are capable of automatically learning highly generalizable distributed representations by exploiting the inductive biases encoded in neural network architectures. These architectures are also capable of combining the paradigmatic linguistic knowledge contained in huge corpora, with syntagmatic knowledge embodied in shorter utterances. Traditional word-based neural approaches for learning word representations are capable of modeling lexical and syntactic information, but they fail to acknowledge patterns occurring at the subword level. This limits their ability to properly model morphemic structure, thus producing low quality representations in highly inflectional languages, where certain words composed of frequent morphemes might appear rarely.

Recent work has shown that incorporating subword information into word-based models significantly improves word representation quality in a wide range of languages and domains. However, it has failed to provide significant insights on the specific methodologies for combining the subword and word level representations. In this thesis we first study the merits of modeling a specific instantiation of the subword domain, the character level, and then explore different mechanisms for combining it with word-level representations. We first show that modeling characters is not beneficial in the Natural Language Inference task embodied by the MutliNLI dataset

and argue this might be due to English being an analytical language, and thus not possessing a significant number of inflected words. We also hypothesize that the Natural Language Inference task in general and MultiNLI in particular might not benefit from increased lexical modeling power. Second, we show how a character-based pre-trained language model fine-tuned in the Implicit Emotion Recognition task performed almost as well as a model trained in at least an order of magnitude more data from the target domain. We argue this model was able to perform this well, despite having been pre-trained in a different domain and in less data, due to the high transferability of the patterns occurring at the subword level. Third we propose a Gating Mechanism for combining character and word-level representations. This mechanism is capable of automatically learning a measure of importance of both levels conditioned on the word level, and it discovered that to model rare words it had to increasingly rely on the character-level. Fourth, we argue that letting models learn character-derived word representations that are similar to their pre-trained word-level counterparts should provide increased modeling power, given the fact that they represent the same underlying meaning. Specifically, we force character-derived representations to resemble word-level representations by means of adversarial training and show that doing this significantly improves performance in various tasks and test sets in English.

Finally, since our findings pertain to the domain of learning word representations, they are roughly task-independent. The same gating architectures and adversarial methods used in combining character-derived and pre-trained word representations could be used in diverse applications such as SPAM detection, Part-of-Speech tagging, Named-Entity Recognition, Machine Translation, or Question Answering. In short, any system relying on word representations could potentially benefit from our findings.