

Abstract

Elena Irimia and Alexandru Ceausu. Dependency-based translation equivalents for factored machine translation

One of the major concerns of the machine translation practitioners is to create good translation models: correctly extracted translation equivalents and a reduced size of the translation table are the most important evaluation criteria. This paper presents a method for extracting translation examples using the dependency linkage of both the source and target sentence. To decompose the source/target sentence into fragments, we identified two types of dependency link-structures - super-links and chains - and use these structures in different combinations to set the translation example borders. The option for the dependency-linked n-grams approach is based on the assumption that a decomposition of the sentence in coherent segments, with complete syntactical structure and which could also account for extra-phrasal syntactic dependency (e.g.: subject-verb relation), would guarantee “better” translation examples and would make a better use of the storage space. The performance of the dependency-based approach is measured with the BLEU-NIST score and in comparison with a baseline system.