ABSTRACT

Statistical based machine translation is a natural language processing technique based on statistics by using learning model from a group of sentences (corpus) as training data. Two factors which has major influence in training process are word alignment and reordering model.

This final project conducted some tests to know and understand evectiveness of some word alignment method and reordering model on statistical based translator if used to translate sentences from English to Indonesian. The tests was done by train corpus using certain word alignment and reordering model, and use the trained systems to translate a group of untrained test sentences.

The tests results shows that grow-diag and final-and method gives the best average score. While for reordering model, though the results doesn't point on certain model as the best, it was concluded that using reordering model would gives a significant increase in system's performance, especially for English-Indonesian pair.

Keyword: statistical based machine translation, word alignment, reordering model