

Using Genetic Algorithm for Word Alignment Model

Nway Nway Han

nwaynwayhan@ucsm.edu.mm

University of Computer Studies, Mandalay, UCSM

Aye Thida

ayethida@ucsm.edu.mm

University of Computer Studies, Mandalay, UCSM

Word alignment is a key task for every innovative statistical machine translation (SMT) system. An alignment is the arrangement of two or more alignments between the parallel sentences. The problem of word alignment in SMT is to find the strong alignment in the corresponding sentence pairs. Moreover, the popular word alignment models need bilingual corpora to align the words in the parallel corpus. But for the Myanmar Language which is inflected and it is also a language scarce resource. For this reason, we developed a manually sentence aligned bilingual corpus which has three thousand sentence pairs and created a gold standard (GS) corpus to measure the alignment error rate of the system. This paper explores a new unsupervised word alignment model for these tasks based on the genetic algorithm (GA). Experimental results are show that our model reduces the alignment error rate by 7.55% AER on the baseline.