Abstract

The corpus based techniques in Machine Translation involves parallel corpora, but it is not applicable for the languages for which there are less or no parallel corpora available. In such case the Rule based machine Translation suits best. The main objective of our work is to build a translation system that translates English sentences to Tamil Sentences. Due to the less availability of parallel corpora for English to Tamil the system is implemented using a Hybrid Technique (the combination of both Rule Based Technique and Statistical Technique). The system is first implemented in a Rule Based approach which involves segmentation and tagging, Rule Based Reordering, Morphological Analyzing, and dictionary based translation to the Target language. Then the errors in the translated sentences are corrected by applying Statistical technique.

References

- Hui Yang, Anne de Roeck, Vincenzo Gervasi, Alistair Willis , and Bashar Nuseibeh, &quot;Analysing anaphoric ambiguity in natural language requirements&quot;, © Springer-Verlag London Limited, 2011.
Index Terms

Computer Science  Artificial Intelligence

Keywords
Segmentation And Tagging  Rule Based Reordering  Chunking  Morphological Analyzer