Finite-state-based and Phrase-based Statistical Machine Translation

Josep M. Crego, Jos   B. Marino, Adria de Gispert

Technical University of Catalonia (UPC), Spain

This paper shows the common framework that underlies the translation systems based on phrases or driven by finite state transducers, and summarizes a first comparison between them. In both approaches the translation process is based on pairs of source and target strings of words (segments) related by word alignment. Their main difference comes from the statistical modeling of the translation context. The experimental study has been carried out on an English/Spanish version of the VERBMOBIL corpus. Under the constrain of a monotone composition of translated segments to generate the target sentence, the finite state based translation outperforms the phrase based counterpart.

Full Paper