

Source-Language Features and Maximum Correlation Training for Machine Translation Evaluation

Ding Liu and Daniel Gildea
Department of Computer Science
University of Rochester
Rochester, NY 14627

December 16, 2007

We propose three new features for MT evaluation: source-sentence constrained n-gram precision, source-sentence reordering metrics, and discriminative unigram precision, as well as a method of learning linear feature weights to directly maximize correlation with human judgments. By aligning both the hypothesis and the reference with the source-language sentence, we achieve better correlation with human judgments than previously proposed metrics. We further improve performance by combining individual evaluation metrics using maximum correlation training, which is shown to be better than the classification-based framework.