Can neighboring relations help to anticipate upcoming discourse relations?

Fatemeh Torabi Asr and Vera Demberg
MMCI Cluster of Excellence, Saarland University

Discourse relations between text segments can be explicitly marked using discourse connectors (e.g., *because*), but by far not all discourse relations are marked explicitly. However, people are usually able to infer the intended relation. What kinds of cues do people rely on? Could they in principle use structural cues from the preceding discourse to *anticipate* upcoming discourse relations? This study tests whether information about neighboring discourse relations (e.g., the relation sense of neighboring discourse relation, whether a neighboring relation was marked with an explicit connector, whether an argument of the current relation embeds another relation or is embedded in some other relation) can help to automatically predict discourse relation senses. We specifically test in how far the incrementally available left context of a discourse relation (compared to no contextual information; and compared to the right context) could be a useful cue for identifying discourse relations.

We extract 15,371 implicit relations and their context from sections 1-23 of the Penn Discourse Treebank [2]. We distinguish between structural cues that occur in the left context of a relation (pre), cues in the right context (post) and cues that occur during the text span of the relation (mid). We compared the effect of pre, post and mid features in automatic classification of relation senses for the implicit discourse relations in 10% of the text after training on the remaining 90% (10-fold cross validation) using the Mallet Maximum Entropy classifier.

Using all structural discourse features, we observed a 5.28% improvement from 26.29% to 31.57%; sdev=0.01) in relation sense accuracy over the majority baseline (causal relation [1]). When using only pre or only post features for relations which had both pre and post features, relation sense identification accuracy was about 2% less (with pre features working slightly better than post features). Discourse relations which only had pre features were classified with an improvement of 3.12% over the baseline using the pre features, whereas improvement for post features on relations having only post features was smaller (1.29% over baseline). These results indicate that it's generally possible to use structural cues, and in particular left ones, from neighboring discourse to help infer discourse relations. Future work should investigate, in how far humans use this information.

1. Asr, and Demberg, (2012). Implicitness of discourse relations.

2. Prasad, Dinesh, Lee, Miltsakaki, Robaldo, Joshi, and Webber (2008). The Penn Discourse Treebank 2.0.