Frame Semantics for Language Technology

The potential impact of FrameNet and other frame-semantic resources in general for language technology (LT) is evident. However, the measurable effect of adding frame-semantic information to LT applications like question-answering and information extraction has been consistently low. The Recognising Textual Entailment (RTE) Challenge, established some five years ago, has turned out to be an extremely valuable meta-task, a test-bed for evaluating the degree to which different linguistic resources and computational methods enhance the inferential capacities of NLP systems. The task is to decide for pairs of natural-language sentences called "text" and "hypothesis", respectively, whether the latter can be plausibly inferred from the former. The only RTE systems making substantial use of FrameNet information were developed by the Saarbruecken SALSA project (Burchardt et al. 2006) and submitted to the RTE 2 and RTE 3 shared tasks. The evaluation coincided with the observations mentioned above: Addition of frame-semantic information led only to minor improvements over system versions using the shallow feature of word-overlap only.

To investigate the reasons of the discrepancy between frame-semantic potential and poor practical outcome, we conducted a series of studies, which started out from the manual annotation of RTE data. The most central annotation result is the FATE corpus (Burchardt and Pennachiotti 2008). It contains the 800 text-hypothesis pairs of the RTE2 test set annotated with frame structures taken from the FrameNet database. With the FATE corpus as a gold standard, we conducted experiments to independently evaluate possible sources of the poor effect of frame-semantic information. In particular, we investigated the coverage of the FrameNet resource on RTE data, which turned out to be surprisingly good, the performance of the shallow semantic parser, which is noisy, as expected, and the appropriateness of knowledge modeling decisions: We looked into the discriminative power of different features extracted from the gold-standard frame-annotation for the RTE classification task, with results that were more ambivalent and complex than in the former cases.

This paper reports the SALSA work on frame-based textual entailment recognition, discusses the results of the empirical studies and the conclusions to be drawn in order to overcome the "implementation gap". In particular, it suggests models for the future collaboration between frame-semantic lexicography and computational linguistics to fully leverage the indubitable value of frame semantic information for language technology.