

Coreference from the Cross-lingual Perspective

Michal Novák

The subject of this thesis is to study properties of coreference using cross-lingual approaches. The work is motivated by the research on coreference-related linguistic typology. Another motivation is to explore whether differences in the ways how languages express coreference can be exploited to build better models for coreference resolution. We design two cross-lingual methods: the bilingually informed coreference resolution and the coreference projection. The results of our experiments with the methods carried out on Czech-English data suggest that with respect to coreference English is more informative for Czech than vice versa. Furthermore, the bilingually informed resolution applied on parallel texts has managed to outperform the monolingual resolver on both languages. In the experiments, we employ the monolingual coreference resolver and an improved method for alignment of coreferential expressions, both of which we also designed within the thesis.