

Abstract:

This thesis investigates the correspondence between two semantic formalisms, namely the tectogrammatical layer of the Prague Dependency Treebank 2.0 (PDT) and Robust Minimal Recursion Semantics (RMRS). It is a first attempt to relate the dependency based annotation scheme of PDT to a compositional semantics approach like RMRS.

An iterative mapping algorithm that converts PDT trees into RMRS structures is developed that associates RMRSs to each node in the dependency tree. Therefore, composition rules are formulated and the complex relation between dependency in PDT and semantic heads in RMRS is analyzed in detail. It turns out that structure and dependencies, morphological categories and some coreferences can be preserved in the target structures. Furthermore, valency and free modifications are distinguished using the valency dictionary of PDT as an additional resource.

The evaluation result of 81% recall shows that systematically correct underspecified target structures can be obtained by a rule-based mapping approach, which is an indicator that RMRS is capable of representing Czech data. This finding is novel as Czech, with its free word order and rich morphology, is typologically different from language that used RMRS thus far.