Data-driven Approaches to Parsing and Semantic Composition

10. and 11.12.2019 – Tübingen University, Germany

Syntactic parsing assigns structure to phrases. Composition models, on the other hand, assign meaning vectors to phrases. Parsers are known to benefit from distributional word representations while composition models can again profit from (morpho-)syntactic information. Given this interplay of parsing and composition, we would like to investigate in this workshop whether parsers and models of distributional semantics can mutually inform each other above the word level.

Phrase composition and its relation to parsing raise a number of important questions:

- How can distributional representations of phrases be created?
- What kind of semantic and/or syntactic information do these representations encode?
- What are possible underlying units beyond single words which can be represented via composition?
- What kind of parsing formalisms benefit composition models for phrases and what kind of phrase representations benefit parsing?
- Which criteria do corpora have to fulfill to be most suited as a data resource for distributional methods?
- Do we need semantic composition for (syntactic) parsing?

In this workshop, we want to shed light on these questions from the perspective of distributional semantics, parsing and corpus linguistics. Among the topics this workshop will cover are: semantic composition modeling, parsing, and the intersection of the two tasks; distributional semantics; automatically annotated large-scale corpora. A series of invited talks from speakers of different specialisations will be concluded with a final panel discussion.

Organizers: Corina Dima, Daniël de Kok, Patricia Fischer,
             Erhard Hinrichs, Tobias Pütz