The Groningen Meaning Bank

Johan Bos

University of Groningen, The Netherlands

What would be a good method to provide a large collection of semantically annotated texts with formal, deep semantics rather than shallow? In this talk I will argue that (i) a bootstrapping approach comprising state-of-the-art NLP tools for semantic parsing, in combination with (ii) a wiki-like interface for collaborative annotation of experts, and (iii) a game with a purpose for crowdsourcing, are the starting ingredients for fulfilling this enterprise. The result, known as the Groningen Meaning Bank, is a semantic resource that anyone can edit and that integrates various semantic phenomena, including predicate-argument structure, scope, tense, thematic roles, animacy, pronouns, and rhetorical relations. A single semantic formalism, Discourse Representation Theory, embraces all these phenonema by taking meaning representations of texts rather than sentences as the units of annotation.