Beyond MaltParser -- Advances in Transition-Based Dependency Parsing

The transition-based approach to dependency parsing has become popular thanks to its simplicity and efficiency. Systems like MaltParser achieve linear-time parsing with projective dependency trees using locally trained classifiers to predict the next parsing action and greedy best-first search to retrieve the optimal parse tree, assuming that the input sentence has been morphologically disambiguated using a part-of-speech tagger. In this talk, I survey recent developments in transition-based dependency parsing that address some of the limitations of the basic transition-based approach. First, I show how globally trained classifiers and beam search can be used to mitigate error propagation and enable richer feature representations. Secondly, I discuss different methods for extending the coverage to non-projective trees, which are required for linguistic adequacy in many languages. Finally, I present a model for joint tagging and parsing that leads to improvements in both tagging and parsing accuracy as compared to the standard pipeline approach.

Bio: Joakim Nivre is Professor of Computational Linguistics at Uppsala University and currently visiting scientist at Google, New York. He holds a Ph.D. in General Linguistics from the University of Gothenburg and a Ph.D. in Computer Science from Vaxjo University. Joakim's research focuses on data-driven methods for natural language processing, in particular for syntactic and semantic analysis. He is one of the main developers of the transition-based approach to syntactic dependency parsing, described in his 2006 book Inductive Dependency Parsing and implemented in the MaltParser system. Joakim's current research interests include the analysis of mildly non-projective dependency structures, the integration of morphological and syntactic processing for richly inflected languages, and methods for cross-framework parser evaluation. He has produced over 150 scientific publications, including 3 books, and has given nearly 70 invited talks at conferences and institutions around the world. He is the current secretary of the European Chapter of the Association for Computational Linguistics.

* LTI colloquium: http://www.cs.cmu.edu/~roni/11700/