Speaker: Josef van Genabith

Title: Parsing and Generation with Treebank-Based Wide-Coverage Probabilistic LFG Resources

Abstract:

Hand-crafting wide-coverage deep linguistic resources is time-consuming, expensive and hard to scale to unrestricted text. In this talk I present research on automatically acquiring wide-coverage, probabilistic Lexical-Functional Grammar resources from treebanks for English, Spanish, Chinese, German and French. I show how these resources can be used in a number of applications including parsing, generation, lexicon acquisition and MT evaluation. I compare the automatically acquired against hand-crafted resources (XLE and RASP). I relate our approach to research on inducing HPSG and CCG resources from treebanks and sketch some future work.

This is joint work with Aoife Cahill, Mick Burke, Ruth O'Donnovan, John Judge, Grzegorz Chrupala, Natalie Schluter, Yvette Graham, Conor Cafferkey, Ines Rehbein, Masanori Oya, Yuqing Guo, Yafa Al-Raheb, Jennifer Foster and Deirdre Hogan.