

Prague Treebanking for Everyone: A two-day tutorial

November 28-29, 2006

Krystal Hotel, Jose Martiho 407/2, Prague, Czech Republic

<http://ufal.mff.cuni.cz/vmc>

The tutorial will introduce the Prague Dependency Treebank project, which aims at a complex manual annotation of a substantial amount of naturally occurring sentences in continuous Czech texts. The Prague Dependency Treebank has three levels of annotation: morphological, analytical (describing surface syntax in a dependency fashion) and tectogrammatical, which combines syntax and sentence semantics into a language meaning representation, keeping the dependency structure as the core of the annotation structure but adding basic coreferential links, topic/focus annotation, and a detailed semantic labeling of every sentence unit. More Prague treebanks (the Czech Academic Corpus, the Prague Czech-English Dependency Treebank, and the Prague Arabic Dependency Treebank) will be introduced as well. In addition to the data, all the treebank and data processing tools will be discussed in details.

This tutorial is intended for students, researchers, and practitioners in natural language processing who want to see how many of the broadly annotated data and the annotation and data processing tools have been built in the Prague treebanking projects. The fact that the annotations and tools can be used in a general way could be a strong motivation for all attendees.

Organizer

Institute of Formal and Applied Linguistics
Charles University
Malostranské nám. 25
118 00 Prague, Czech Republic
tel.: +420-221 914 223, fax: +420-221 914 304

