

# **Feature Reduction in Transition-based Dependency Graph Construction**

László Kovács, László Csépanyi-Fürjes  
University of Miskolc

Dependency grammar is an important tool in semantical analysis of Hungarian text sources. There are two main approaches in the literature for construction of dependency graphs. One approach is the family of transition-based methods, while the other model uses direct graph construction operations starting from an initial dependency graph. In the transition-based approach, the engine detects the special features of the source text and determines the next construction steps using a machine learning method. Traditionally, the feature set is constructed manually, and some of the features may be irrelevant or redundant.

In this paper, we present here two proposed methods for the implementation of the feature reduction. The first method uses entropy-based algorithm (information gain) and the other uses the correlation-based approach. For the test evaluation, we use the sentence bank of UD (Universal Dependency) homepage. We use corpuses from two different languages: English and Hungarian.

Our experiences show that the accuracy depends significantly from the applied method whether it supports only projective grammars or also non-projective structures are involved. Regarding the correlation-based approach, we could get initially a very good progress, but later it was very time-consuming to cover all special features of the language.