

Unsupervised dependency parsing is an alternative approach to identifying relations between words in a sentence. It does not require any annotated treebank, it is independent of language theory and universal across languages. However, its main disadvantage is its so far quite low parsing quality.

This thesis discusses some previous works and introduces a novel approach to unsupervised parsing. Our dependency model consists of four submodels: (i) edge model, which controls the distribution of governor-dependent pairs, (ii) fertility model, which controls the number of node's dependents, (iii) distance model, which controls the length of the dependency edges, and (iv) reducibility model. The reducibility model is based on a hypothesis that words that can be removed from a sentence without violating its grammaticality are leaves in the dependency tree.

Induction of the dependency structures is done using Gibbs sampling method. We introduce a sampling algorithm that keeps the dependency trees projective, which is a very valuable constraint.

In our experiments across 30 languages, we discuss the results of various settings of our models. Our method outperforms the previously reported results on a majority of the test languages.