

In my bachelor thesis I decided to focus on disambiguation of Czech morphology. This task is important in particular in the area of natural language translation, where it takes part in preprocessing the text intended for translation in order to eliminate ambiguity in part of speech and other morphological categories. This ambiguity would cause problems in subsequent phases of translation or unacceptable growth of translation's time demands. I chose statistical approach to this problem, which is in comparison with other possible methods faster, more universal and able to select word category in all cases. I founded my application KDTagger, which I created within the framework of this bachelor thesis, on the theory of Hidden Markov Models. My aim was to create such a program, which would be universal in operating system and the way of use. KDTagger allows the experts to adjust every important linguistic parameter while preserving comfort use for beginners. My work also includes extensive testings of the program KDTagger, which I performed on the Czech newspaper texts from Prague Dependency Treebank version 2.0. The program can be however applied on arbitrary natural language without not even the smallest change.