Abstract

Morphological analysis is an important component in many natural language processing Tasks. This project aims at giving a valuable well-defined morphological analysis of Maithili language. No computational work is found so far that is associated for this language. Published materials in this area are very few. Maithili is an eastern Indo-Aryan language spoken in the eastern and northern regions of Bihar state of India and the southeastern plains, known as tarai of Nepal. Maithili is a highly inflectional language so Finite state Transducer is most effective and efficient way for developing the morphological analyzer for Maithili language. In this project, we present Maithili morphological analyzer which gives us morphological analysis of Maithili.

The construction of morphological analyzer is divided into two phase. In first phase we building the lexicon file with the help of lexicon generator. In second phase we generate morphological processor. The approach used for developing the morphological processor is Morphological Analyzer. The morphological processor which does morphological analysis is known as a morphological analyzer. The Morphological analyzer need a dictionary of root words, file containing FST rules and the dictionary of suffixes. The inflectional rule are hand written. We have used the xfst tool for generating the finite states and our proposed morphological analyzer for the Morphological analysis of this language.

Keywords: The Maithili language, morphological analysis, corpus, finite state transducer