Abstract

Part of speech (POS) cataloguing is the process of allocating the part of speech tag or other philological class sign to each and every word in a sentence. In many Natural Language Processing presentations such as word intellect disambiguation, information recovery, information handling, analyzing, interrogating, and machine interpretation, POS tagging is reflected as the one of the basic obligatory tool. Categorizing the uncertainties in language philological items is the puzzling objective in the procedure of emerging an effectual and correct POS Tagger. Works survey displays that, for Indian lingoes, POS taggers were established only in Hindi, Punjabi, Bengali and Dravidian languages. Some POS taggers were also established generic to the Hindi, Telugu and Bengali tongues. All scheduled POS taggers were grounded on diverse Tag-set, established by diverse organization and individuals. This paper speaks the various developments in POS-taggers and POS-tag-set for Indian language, which is very essential computational verbal tool needed for many natural language processing (NLP) presentation [15].
- Jyoti Singh, Nisheeth Joshi, Iti Mathure, "Development of Marathi Part of Speech Tagger Using Statistical Approach";
- Hasan Fahim Muhammad, Zaman Naushad Uz and Mumit Khan, (2007) "Comparison of Unigram, Bigram, HMM and Brill’s POS Tagging Approaches for some South Asian Languages"; In proceeding of Center for Research on Bangla Language Processing.
Index Terms

Computer Science
Artificial Intelligence

Keywords
Tag-set  Ambiguity  Trigram  HMM  NPL  Tokenized  Indian Languages.