Abstract

In this paper, we have proposed a hierarchically organized semantic lexicon in Bangla and also a graph based edge-weighting approach to measure semantic similarity between two Bangla words. We have also developed a graphical user interface to represent the lexical organization. Our proposed lexical structure contains only relations based on semantic association. We have included the frequency of each word over five Bangla corpuses in our lexical structure and also associated more details to words such as, whether the words are mythological or not, whether it can be used as verb or not, in order to use the word as a verb which word should be appended to it etc. As we have earlier discussed, this lexicon can be used in various applications like categorization, semantic web, and natural language processing applications like, document clustering, word sense disambiguation, machine translation, information
Design and Development of a Bangla Semantic Lexicon and Semantic Similarity Measure

retrieval, text comprehension and question-answering systems.

References


Index Terms

Computer Science
Text Processing
Keywords
Bangla SynNet  Semantic Similarity  Category  Concept  Sub-concept  Cluster