Abstract

Polysemous Words can have more than one distinct meaning. Word sense disambiguation (WSD) is the ability to identify the exact meaning of such polysemous words in context in a
computational manner. WSD is considered as an AI-complete problem, that is, a task whose solution is at least as hard as the most difficult problem in Artificial Intelligence. In this paper, we propose an Integrated Kannada Word Sense Disambiguation system which includes a suite of high performance Natural Language Processing (NLP) modules implemented in Perl (Program Extraction and Reporting Language) to carry out word sense disambiguation task. The corpus builder module will construct the raw Kannada corpora using web. The proposed system uses randomly selected sentences from the corpora as a test bed for disambiguation. The electronic machine readable dictionary is built by Dictionary builder module using the corpora. The Target Word Sense Disambiguation module will disambiguate the potential ambiguous target words in a sentence. The polysemous verb in a sentence is disambiguated by Verb Sense Disambiguation module. The rule based disambiguator will disambiguate all ambiguous words with different lexical category. Experiments conducted and the results obtained have been described. The efficiency of the system proved to be reliable and extendable.

Reference


- Warmter, S. 1989. Integration of syntactic and semantic constraints for Structural noun
- Sampada. (online) http://sampada.net
- Kannada web blog. (online) http://kannadablogs.ning.com/
- Prajavani. (online) http://prajavani.net
- Parser. (online). http://ltrc.iiit.ac.in/analyzer/kannada/

Index Terms

Computer Science
Natural Language Processing
Key words

Kannada Word Sense Disambiguation

readable dictionary

Target Word

Verb Sense Disambiguation

Verbalizer

Kannada Corpus

Kannada machine