Abstract

Function tagging is one of the essential steps in Myanmar to English machine translation system. In this paper we propose a set of function tags for Myanmar and address the question of assigning function tags to Myanmar words. A small functional annotated tagged corpus manually serves as the training data because the large scale Myanmar Corpus is unavailable at
present. Part of the challenge of statistical function tagging for Myanmar language comes from
the fact that Myanmar has free phrase order and a complex morphological system. In the task
of function tagging, we use the output of morphological analyzer which tags the function of
Myanmar sentences with correct segmentation, POS (part-of-speech) tagging [1] and chunking
information [2]. We use Naïve Bayesian statistics to disambiguate the possible function tags of
each word in the sentence. Function tagging can be exploited by NLP applications such as
syntactic and semantic analysis, information retrieval and machine translation. Experiments
show that our analysis achieves a good result with simple sentences and complex sentences.

Reference

- Myint, P. H. Assigning automatically Part-of-Speech tags to build tagged corpus for
- Myint, P. H., Chunk Tagged Corpus Creation for Myanmar Language. In Proceedings of
  the ninth International Conference on Computer Applications, Yangon, Myanmar, 2011.
  Brown University, August.
- Charniak, E. 1997. Statistical parsing with a context-free grammar and word
  statistics. In Proceedings of the Fourteenth National Conference on Artificial Intelligence,
  pages 598-603, Menlo Park.
- Versteegen, L. 1999. The Simple Bayesian Classifier as a Classification Algorithm.
- Blaheta, D., and Johnson, M. 2000. Assigning function tags to parsed text. In
  Proceedings of the 1st Annual Meeting of the North American Chapter of the Association
  Advisor-Eugene Charniak.
  Asia Conference on Language, Information and Computation (PACLIC 2009), Hong Kong,
  2009.
- Lintean, M., and Rus, V. Naive Bayes And Decision Trees For Function Tagging. In
  Proceedings of the International Conference of the Florida Artificial Intelligence Research
- Soe, S. P. 2010. Aspects of Myanmar Language , Myanmar Department, University of
  Foreign Language.
- Myanmar Thudda, vol. 1 to 5 in Bur-Myan, Text-book Committee, Basic Edu., Min. of
  Department, University of Educaion.

Index Terms

Computer Science
Natural Language

Processing
**Key words**

Function tagging     function tagset     Myanmar  
language  
functional annotated tagged corpus