Abstract

This paper discusses the UNL Enconversion of Tamil sentences. The rich morphology of Tamil enables the Enconversion process to be based on morpho-semantic features of the words and their preceding and succeeding context. The use of case relation indicating morphological suffixes, POS tag and word level semantics allows the rule based Enconversion to be
independent of the syntactic structure of the sentence. These UNL graphs are used to build a conceptual level index.

Reference

- M. Mahdi Shafiei and Evangelos E. Milios A Statistical Model for Topic Segmentation and Clustering, Canadian Conference on AI 2008: 283-295
- Qi Sun, Runxin Li, Dingsheng Luo and Xihong Wu, Text Segmentation with LDA-Based Fisher Kernel, Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics, June 15-20, 2008, Columbus, Ohio, USA.
- Sathiyamurthy, K., T.V. Geetha, Association of domain concepts with Educational objectives for E-Learning, ACM Compute 2010, Bangalore, India.
- Wei Li, Y., Andrew McCallum Pachinko Allocation: DAG-Structured Mixture Models of Topic Correlations, Proceedings of the Twenty-Third International Conference (ICML 2006),

Index Terms

Computer Science
Machine Learning

Key words

Latent Dirichlet Allocation (LDA)
Ontology
Pedagogy