

SANSKRIT-ENGLISH TRANSLATOR WITH NLP TECHNIQUE

¹ PranaliBombale , ² AditiAuti ³ AmrutaBandewar , ⁴ Prof. Vina M. Lomte

Abstract:

The proposed system is based on the work towards developing a machine translation system for Sanskrit to English. In this project we are using direct approach in which the Sanskrit word is going to translated into the English word. This paper provides the overview of the Interlingua based Sanskrit English Machine translation. In the proposed system the given Sanskrit text is first converted to an intermediate notation called Interlingua (sanskrit). This document is then used for mapping to the target language, English, and generates the translated output. The proposed system works in the translation of simple and complex words in the given corpus.

IndexTerms - Component,formatting,style,styling,insert.

I. INTRODUCTION

Converting one natural language to other is a process of machine translation using application software. Machine-aided human translation is a sub-field of computational linguistics that finds the use of software to translate text or dialogue from one language to another. To process any translation, human or mechanical, the meaning of a manuscript in the original (source) language must be fully restored in the end language, i.e. the translation. It seems very straightforward on the surface but it is very much complex. Translation is not a mere word-for-word substitution. A translator must read and interpret and examine all of the elements in the text and know how each word may influence another. This requires wide knowledge and expertise in grammar, syntax etc. Mainly there are three types of machine translation techniques- direct approach, transfer based approach and interlingua based approach. Interlingual machine translation is one of the basic and traditional approaches to machine translation. In this approach, the source language, i.e. the text to be translated is transformed into an interlingua, i.e., an abstract language-independent representation. Sanskrit processes a rich grammar developed by Panini around 3000 years ago. In this project we are using direct approach in which the Sanskrit word is going to translated into the English word. This paper provides the overview of the Interlingua based Sanskrit English Machine translation

Keyword:

Sanskrit language, Translation, English language, grammar.

Related Work:

Akshar Bharathi et al. showed the details of Paninian framework [1], Parsing Free Word Order Languages in the Paninian Framework [2], and Karaka analysis [3]. The authors explains also explains the use of lexical functional grammar for mapping to grammatical relations[4]. The parsing of Sanskrit sentences using LFG is explained by Mrs. NamrataTapaswi et.al. [5]. Paul Kiparsky gives detailed description of different levels of Paninian framework with examples and rules of Ashtadhyayi and rule formation on different levels of Paninianframework[6]. Sudhir Kumar Mishra et.al. [7] gives a detailed study on the Karaka analysis system based on rules of Ashtadhyayi with examples. DeryleW. Lonsdale, Alexander M. Franz, and John R.

R. Leavitt presented the design and development of an interlingua for a large-scale MT project. They also discussed how the resulting Knowledge-based, Accurate Natural-Language Translation (KANT) interlingua handles complexity and development of different stages efficiently. It is developed in a balanced fashion with maximal coverage. They use, a recursive list-based structural representation of source sentences in this approach. An interlingua frame consists of a head concept, feature-value pairs, and semantic slots. It may contain nested interlingua frames. The source language expressions and semantic units from the domain were considered for the concept generation. The overall format is modeled using frame-based structures. The structure reflects deep semantic relationships between major constituents[8]. The Interlingua approach is based on the concept that MT must go beyond purely linguistic information, syntax and semantics, and should understand the content of texts. Sameh AlAnsary et.al.briefly reviews three of the most renowned interlingua-based machine translation projects, Distributed Language Translation (DLT), Universal Translator (UNITRAN) and KANT system. DLT, are search project developed in Utrecht, The Netherlands, is an interactive system developed to operate over computer networks. Translation is distributed between two independent terminals; one for the analysis and another for generation. UNITRAN is a translation system developed at Massachusetts Institute of Technology. The system operates bi-directionally between Spanish and English. KANT system has been developed at Carnegie- Melon University (CMU) in Pennsylvania, USA in 1989". KANT is the only Interlingua based MT system to be operational commercially. It has been used in translating English technical documents into French, Spanish and German.

Motivation:

Now a day's people want to learn different type of language from the online world. Also we all believe that most effective way to learn language is by own. For this purpose this system works. In Hindu hymns most of the records are written in Sanskrit. In which it includes Rigveda and shlokas. For understanding the meaning of Sanskrit words in natural language (ex English) we proposed this system

Mathematical Model

Mathematical model set theory $S = \{s, e, X, Y, \Phi\}$

s= Start of the program

1. Register/Login into the system
2. Provide Dataset (Sanskrit word & English Meaning).

e= End of the program

Find the Sanskrit word & English Meaning

$X = \text{Input of the program} = \{P, R\}$

$P = \text{Input data}$

$R = \text{Dataset}$

$Y = \text{Output of program} = \text{Sanskrit word related English Meaning}$

First, users provide Input data or File for specific output English Meaning.

Let R be the set of User Data

$D = \{D1, D2, D3... \dots\dots Dn\}$

Let A be the set of categories Dataset (English Meaning)

Therefore,

Overall Data is evaluated with the help of find the English meaning which basically represent helping people to interested learn Sanskrit Language.

Conclusion:

Many of the ancient and modern documents were written in Sanskrit, since we use an Machine translation system, it will translate Sanskrit to English language. The proposed system can be used for educational purposes, communication etc. The translation of Sanskrit to English is performed very well.

Acknowledgment

It gives us great pleasure in presenting the preliminary project report on 'SANSKRIT-MARATHITRANSLATOR WITH NLP TECHNIQUE'

I would like to take this opportunity to thank my internal guide for giving me all the help and guidance I needed I am really grateful to them for their kind support. Their valuable suggestions were very helpful. I am also grateful to HOD for her in dispensable support and suggestions.

Name of Students

¹ PranaliBombale , ² AditiAuti ³ Amruta Bandewar ,

REFERENCES:

- [1] SamehAlAnsary Department of Phonetics and Linguistics, Faculty of Arts, Alexandria University ElShatby, Alexandria, Egypt., “Interlingua-based Machine Translation Systems: UNL versus Other Interlinguas”,2014.
- [2] Jaideepsinh K. Raulji, “Sanskrit Machine Translation Systems: A Comparative Analysis”, International Journal of Computer Applications,2016.
- [3] P. Goyal, V. Arora and L. Behera, "Analysis of Sanskrit text: Parsing and semantic relations," in Sanskrit Computational Linguistics, Sanskrit Computational Linguistics, 200-218, 2009.
- [4] Sanskrit.jnu.ac.in/corpora/MSR-JNU-Sanskrit-Guidelines.htm, Annotation Guidelines for tagging Sanskrit using MSRI-JNUSanskrit tag set.
- [5] Marianne Starlander and Paula Estrella,” Looking for the best Evaluation Method for Interlinguabased Spoken Language Translation in the medicalDomain”, Proceedings of the 8th International NLPCS Workshop, Special theme: 20-21 August, 2011 pp 81–92.
- [6] Prof. Hans Uszkoreit “Survey of Machine Translation Evaluation”, EuroMatrix,Project funded by the European Community under the Sixth Framework Programme for Research and Technological Development.,2007, pp 49-53.
- [7] Ved Kumar Gupta, Prof. NamrataTapaswi, Dr. Suresh Jai,Knowledge representation of Grammatical constructs of Sanskrit language using Rule based Sanskrit to English MT, 2016 <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6524744>.
- [8] Dinesh kumar, Gurpreet Sing, POS Tagger for Morphology rich Indian languages, International Journal of Computer Applications (0975 – 8887) Volume 6– No.5, September 2010.