

ICS261: SSAD & Project

Team number	<i>39</i>
Project Title	<i>Word Sense Disambiguation in Crowd Sourcing Mode</i>
Document	SSAD Project Concept Document
Creation date	<i>4 September 2013</i>
Created By	<i>Abhinav Pandey, Ankit Kotak, Anubhav Srivastava, Jayanth Katuri, Nazrul Haque Athar</i>
Client	<i>LTRC</i>

Description

The project is about forming rules of translation between English and Hindi using word sense disambiguation(WSD) which governs the process of identifying the sense of a word in a sentence which usually has various meanings . Our aim is to collect data related to it from a variety of users in an interesting way .Given a word it has multiple senses and when translated to some other language it can be translated into multiple words out of which one is chosen based on context. We will be building a rich interface to collect this information from the user. Based on the context given by user , we further encourage the users to provide us with appropriate reasons for their choices . After this is accomplished , in the next part of the project we will be building an interface to implement the rules derived from the reasons provided by the users . This will help us in natural language processing and machine interpretation .

Profile of Users

The end user can be any person with basic understanding of English and Hindi languages . The user will be providing us with the meanings and translations of various words in English to Hindi and also the corresponding reasons for their choices . The user should have basic knowledge of computer regarding how to operate it , etc .

There will be administrators who will be managing the site and resolving any issues faced by the end-user .

Usage Model

The interface provided to the end-user will be highly interactive and appealing . At each step user will be provided with English sentences and later on will be asked for the meanings of particular words in those sentences in Hindi (Hints will be given for the same) . He will be also asked for the reasons for his/her choices regarding the context of those words . Based on the responses provided by the user , a database will be formed which will help in forming the general rules for translation .