International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-3, January 2019

Design and Development of a Knowledge-Based Approach for Word Sense Disambiguation by using WordNet for Hindi

Pooja Sharma, Nisheeth Joshi

Abstract: ***Word Sense Disambiguation (WSD) aims at deciphering word meaning in terms of context with the help of computers. WSD is an accessible and challenging AI-complete problem. WSD analyses word tokens and determine the exact sense of the word being used according to the context. WSD is viewed asafundamental problem in Artificial Intelligence (AI) and Natural Language Processing (NLP). Our problem area involves finding the appropriate sense of a lexeme for available context and relationship between lexicons. This is done using natural language processing techniques which involve queries, NLP specific documents or output texts from Machine Translation (MT). MT automatically translates text from one native language into another. It can be performed on various natural languages like Urdu, Marathi, Punjabi, Bengali, English, and Hindi etc. The different application areas for word sense disambiguation involves SpeechProcessing, Information Retrieval (IR), lexicography, Text Processing and MTetc.With this article, we are exploring the knowledge-based technique for WSD for Hindi. This approach uses explicitly available lexical resources viz. lexicon and thesaurus. It involves incorporating word knowledge from external knowledge resources to removethe equivocalness of words. In this experiment, we tried to develop a WSD tool by considering a knowledge-based approach with WordNet of Hindi. The system uses knowledge-based LESK Algorithm for WSD for Hindi. Our proposed system gives the accuracy of about 71.4%.***

Keywords: Word Sense Disambiguation, LESK, WordNet

I. INTRODUCTION

Ambiguity finds its place amongst the critical characteristics of natural language. A word can have multiple meaning varying with the context in which it is used. WSD is characterised AI-complete issue which is related to NP-completeness. The complexity theory states that a problem with difficulty similar to solving major AI problems, is generally not because of a single issue, but from several factors. WSD is a process to determine the exact sense of the word for the context it is used in.

Example:

a.) गीता ने प्रतियोगिता मे भाग लिया |
Here, the 'भाग' meaning is a participant.

b.) माँ ने खेत के चार भाग किये|
Here, the 'भाग' meaning is division.

c.) पवन चोरी करके भाग गया |
Here, the 'भाग' meaning is running.

Manuscript published on 30 January 2019.
*****Correspondence Author(s)
Pooja Sharma, Department of Computer Science, Banasthali Vidyapith, Jaipur, Rajasthan, 304022, India.
Nisheeth Joshi, Department of Computer Science, Banasthali Vidyapith, Jaipur, Rajasthan, 304022, India.

WSD mainly relies on knowledge. The WSD system can be interpreted to work as follows: it takes as input an arrangement of words or sentence, then NLP techniques are applied which utilises at least one source of information to identify the best-suited senses of the phrase regarding the given context. Words possess a different meaning that varies with the changing context in which they are used and figure out which meaning of the word is intended in a particular context. This is one of the primary issues which are often experienced by any NLP framework. This is also identified as lexical-semantic ambiguity. WSD is an area of exploration in NLP, which is gaining popularity nowadays. This technique is amongst those NLP techniques, which can also be represented by job, execution, the source of knowledge, application, assumptions and computational complexity for WSD algorithms.

II. RELATED WORK

Chaplot&Salakhutdinov[4],utilised a WSD framework as the sentence or a little window of words around the real word as the setting for disambiguation. They use the formalism of the topic model to plan a WSD framework that scales directly with the number of words in the unique circumstance. Subsequently, this framework can use the entire report as the setting for a word to be disambiguated. The proposed strategy is a variation of Latent Dirichlet Allocation in which the theme extends for a record are supplanted by synset extends.

Montoyo et al.[11],worked for appointing the senses of the right words utilisingan e-lexicon as the word definitions sources. WSD is based on the following two methodologies:knowledge-based; corpus-based. The researchers have tried to join different sources of information, through blends of these two strategies.

Stevenson & Wilks, [3]to figure out the most helpful knowledge sources and whether their mix prompts enhanced outcomes. Display a sense tagger which utilises a few knowledge sources and framework attempts to disambiguate every single substance word in running content as opposed to constraining itself to treat a limited vocabulary of words.

Yadav & Husain [26],Hindi WordNet will be utilised for solving the ambiguity of words. Profound methodologies and shallow methodologies are two fundamental methodologies used as a part of preparing of language. Concentrate the words in the polysemous setting by an alternate way and measure their level of significance in deciding the importance of polysemy. In this procedure, consider the sentence collocation, as well as should additionally consider the punctuation and semantic to acquire more

*Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)*

learning in accordance with human psychological conduct models. Co-occurrence words, collocation words, and demonstratives have several degrees of constraint on deciding the polysemy sense. In this manner, they can be extricated from the corpus, lexicons, and learning source to build the bound together disambiguation information base, and afterwards, utilise them for disambiguation.

Bala[1], utilising Knowledge-based way to deal with WSD with the Hindi language. Dictionary and Thesauri are the outside lexical assets use by the approach. Appropriate senses will be assigned to a word in Hindi by using statistical strategy. Their work depicts the use of selectional restriction for the Hindi language which is again a knowledge-based approach.

Moro et al., [12] have reported on WSD and Entity Linking (EL). Both serve to the language's lexical vagueness. The contrast between these two techniques based on central is that in EL, and textual specification can be linked to the substance having a name, which might contain the correct sense; while with WSD there is an ideal connection amongst the word frame and an appropriate sense. An approach based on the graph was used for EL and WSD which chooses high coherence semantic interpretations.

Trivedi et al. [25], analyse the distinctive lexicon ways to deal with Word Sense Disambiguation, i.e. diagram based, ontology-based and information based. Knowledge-based methodologies are better yet require world information to be generally useful.

Pal et al. [15], Knowledge-based system have been proposed for Word Sense Disambiguation for any low asset language. As Knowledge-based, Bengali WordNet has been utilised as a part of this work. The conventional Lesk algorithm has been created as the standard information-based approach.

Naseer & Hussain [14],study no cited work will be there for resolving words ambiguity in the Urdu language. In this paper "Bayesian Classification" used for determining the specific kind of lexical ambiguityknown as WSD for URDU words.

Reddy et al. [19], works on techniques for adapting domains and utilising the involved documents released during the task were used to give ranking scores to the tokens and their corresponding senses. PersonalisedPageRank algorithm will be used to disambiguating the test data.It was applied to a graph constructed from the entire WordNet in which nodes were initialised with respective ranking scores of tokens and their senses. Their systems achieved comparable accuracy of 53.4 and 52.2 in the competition, which outperformed the most frequent sense baseline with accuracy 50.5.

Reddy et al. [20],used the ontological categories as per Hindi Wordnet as inventories of semantic type. They presented two unsupervised approaches with the name-Hierarchical Semantic Category Labeler (HSCL) and Flat Semantic Category Labeler (FSCL).The firsttechniquetakes in semantic categories in the form of a list, whereas the secondconsiders the hierarchy in a top-down model. Further,these methods use natural models based on probability, using which the category labelling became a simple table lookup with small computation and thus widening the scope of its application in real-time interactive systems. Parameswarappa& Narayana [17], proposed an Integrated Kannada WSD system which involved a suite of advanced executionNLP modules developed in Perl. The system used sentences,selected randomly through the corpora,like the set of an experiment for disambiguation. The target WSD will disambiguate the feasible ambiguous words targeted in a sentence. The Verb Sense Disambiguation module was used to disambiguate polysemous verb in a sentence. The rule-based disambiguator will disambiguate all ambiguous words with the different lexical category. They did experiments and described the obtained results. The systemefficiency was reliable by proof.

Kaur [7],chose to create an algorithm for treating ambiguous postpositions found in the Hindi language. She took this issue with the case study of existing Hindi-Punjabi Machine Translation System. Hence the disambiguation happened from the machine translation point of view. This was used for removing ambiguity from the corpus. Hindi postpositions were developed, and words were extracted from the corpus using the N-gram algorithm.

Pandey & Arora [16], Cross-Lingual methods have therefore been used for achieving WSD in languages other than English, where the vast amount of sensitive data available in English is used as a tool to make sense of texts in different languages for machines. They proposed a method outside the scope of parallel corpora. They generate comparable corpora using Wikipedia. This producedlarge data because the number of Wikipedia articles in Hindi is more than 100,000. Using extended Lesk to get the sense of ambiguous words from WordNet and transferring the sense to the Hindi text.Reddy and Inumella[18], model WSD as a Distributed Constraint Optimization Problem (DCOP), different knowledge sources information viewed as a constraint. DCOP algorithms can collectively enhance a lot of utility functions associated with these constraints. The researchers have elaborated on how utility functions for various knowledge sources can be designed. For evaluation, they havetaken a simple DCOP problem for all words WSD. The results are on par with state-of-art knowledge-based systems. Sankaran & Vijay-Shanker[21], analyses the influence of morphology in WSD for Tamil. The results do confirm that morphology affects the performance of a WSD algorithm to a considerable extent for Tamil. For evaluation purposes, we use the Decision List approach of Yarowsky (1995), a well-known WSD method. Mohapatra and Hembram[9], took a different approach to tackle WSD. They did an analysis on language understanding. This research can be termed- 'The law of conservation of thought'. Through this research, the structure of thought, its transformation and conservation along with the thought boundary detection procedure for disambiguation of the words in the text are analysed.

Baldwin et al. [2], have again considered the task of MRD-based WSD to extend the Lesk algorithm to study the impact of different tokenisation schemes on WSD performance and methods of definition extension. In their experiment on the Hinoki Sensebank and the Japanese Senseval-2 dictionary task, they explained that sense-sensitive definition extension over hypernyms, hyponyms, and synonyms, along with definition extension and tokenisation of word leads to most accurate WSD than both unsupervised and supervised baselines. While doing so, they explained the importance of ontology induction and widened the scope for the development of baseline unsupervised WSD methods.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Munusamy et al. [13], performed a study to define an object-net method for WSD. They proposed to model the basic senses that are meant to help the machine in undertaking the analysis and synthesis processes of meaning on their own. According to this methodology, the disambiguation process takes place based on context. From the original sentence, they detected its context, and then the actual meaning was found out using the correlation of elementary object meanings existing in the database of object-net. This is because even ambiguous words have only one meaning in a particular context, object or domain on which the sentence was written. According to the researchers this approach gave high precision of 96% of the verbs and 97% of nouns for data extraction from the database.

Moldovan and Novischi[10], have presented a pack of techniques and their results for the disambiguation of WordNet glosses based on semantics. Pointers were added from each word to its concept or synset by semantically disambiguating the words in the glosses. This enhanced the connectivity between the concepts of WordNet. They have shown how lexical chains and other applications can be built on WordNet. Heuristics were used to perform the semantic disambiguation of the WordNet glosses. The precision of this system was improved by comparing the disambiguation system and another WSD system. They disambiguated the entire WordNet 2.0 with the accuracy of 86%.

Tamilselvi and Srivastava [24], presented an approach for WSD based on cases using minimum features set. For making the disambiguation, they acceptedentirely two features: post-bigram and pre-bigram. In order to classify the instances for disambiguation, they adopted three steps: instance filtering POS based, instance or case identificationand case selection based on similarity measuring methods. They have used Euclidean, city-block and cosine methods as distant measuring functions. Input for this was set of the reduced form of cases. They usedartificial Neural Networks and K-nearest neighbouring algorithm. Between these two, KNN gave an accuracy of 81.75% from cosine cases using pre-bigram features.

Kolte and Bhirud[8],gave a new methodology for WSD. It was based on WordNet hierarchy and domain information. The domain of the sentence is determined by words in the sentence. The domain-based text analysis is possible due to the availability of WordNet domains. Depending upon the domains of the content words the domain of the target word can be set. This approach proved to be effective in disambiguating nouns. The researchers have presented the unsupervised approach to WSD using the WordNet domains. The model finds the target word domain and the sense related to this domain is understood by the correct sense.

Sharma et al. [22], classified WSD approaches on the basis of source of knowledge used in differentiating senses. They used methods that are based majorly on lexical knowledge bases, thesauri and dictionaries without utilising any corpus evidence. Such methods are known as dictionary-based and knowledge-based. They addressed WSD by using a combination of learning algorithms. This study has successfully compared machine learning algorithms for WSD.

III. METHODOLOGY

WSD is the work to identify the current interpretation of an ambiguous word in a context. Typically, one or more text is used for WSD. If the punctuation is neglected, text(T) can be seen as series of words (w_1, w_2,..., w_n), and thenwe candescribe WSD as a procedure of giving meaningful sense to some or all of the words, that means todefined a mapping(A) between words to their senses, such that $A(i) \subseteq Senses_D(w_i)$, *where* $Senses_D(w_i)$isthe set of senses for word $w_i$encoded in a dictionary(D) and $A(i)$is that subset of the senses of w_i which is most suitablein context T. Each word $w_i \in$Tcan assign more than one sense, only the most appropriate meaning is chosen, i.e.| A (i) | = 1.

WSD makes use of Dictionary or Thesauri to achieve better identification of all word senses. Knowledge resources can also vary depending on corpora of textual content, labelled or unlabeled with word senses or more arranged resources like dictionaries which are machine-readable, semantic systems etc. Knowledge of the words is the critical ingredient to distinguish the meaning whether you talk about humans or machines. Word sense disambiguation finds its applications in many areas such as information retrieval (IR), information extraction (IE), and speech recognition (SR). WSD is a critical element for word knowledge. WSD has seen different approaches. Most of the approaches are established on various statisticaltechniques. Few approaches require corpora which are tagged for the senses and othersemploy unsupervised learning. In thispaper, we acquireLeskapproach (Lesk 1986), which involveslooking for overlap between the wordsin given definitions with words fromthe text surrounding the word to bedisambiguated. The flow of WSD is shown in figure 1.

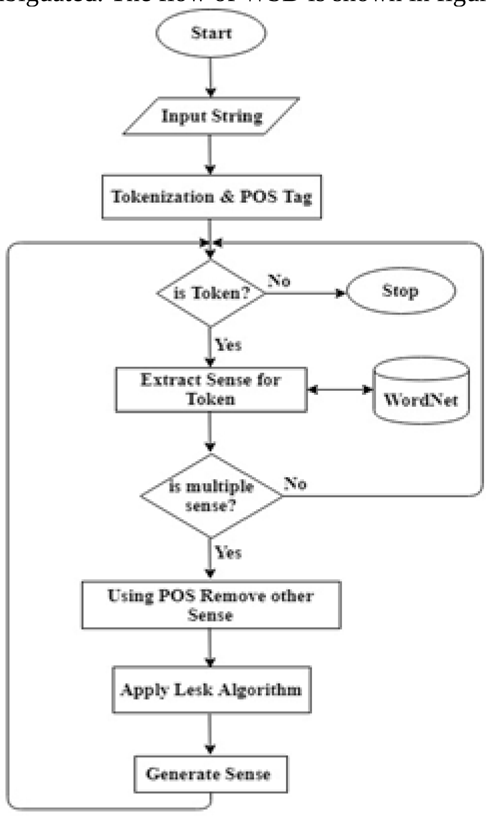

Figure 1: Flowchart of WSD

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

A. English WordNet

WordNet is defined as a dictionary wherewords are stored along withtheir meanings. However, itis differentfrom theconventional ones from numerous points of views. For example, words are arranged *semantically* in WordNet rather than alphabetically.
It is asystem of words or tokensjoinedby their lexical and semantic relations. It is a vastelectronicEnglish words database that can be nouns, adjective, verbsand adverb which are gathered into synonyms sets or synsets.Synsets are connected through the medium of lexical and conceptual-semantic relations.
It is a conjunction of dictionary and thesaurus which is made and preserved by Cognitive Science laboratory of Princeton University under the direction of Prof. George A. Millar (psychology).

B. Hindi WordNet

The Hindi WordNet is created under the guidance of Prof. Pushpak Bhattacharyya at the Centre for Indian Language Technology (CFILT), Dept. of Computer Science and Engineering, IIT Bombay [3].It is a system that unites a specific interpretive and semantic relationship between Hindi words. Itorganises the lexical data in terms of meaning.Words are added together according to their meaning equality in Hindi Wordnet.There can be interchangeable words in any reference. In the Hindi WordNet, there is a synonym set, or synset, for each wordrepresenting one lexical concept. Where a word has different meanings, in those cases it is done to remove the ambiguity.WordNet is a collection of synsets.

C. Synsets

Synset is a category of semantically equivalent data elements used for information retrieval. A set of more than one words and phrases ("collocations") is jointly referred to as "word forms" which can all divide the same meaning.

- A synonym set.
- Describe a particular meaning of a word.
- The smallest unit in wordnet.

Synsets are connectedto lexical and semantic relations. Every set ofword-meanings can be comprised ofwordforms set, also recognisedas synonym sets or synsets. It can be built by content words such as noun, adjective, verband adverb.

Example:"फल, परिणाम, अंजाम."

D. Lexical Matrix

It is an integralcomponent of the language system. It concerns the connection between word forms and their meanings. The lexical matrix is representedby Table 1. Word classes are taken as columnsheadings, and word-meanings are taken as rows heading. Rows only have synonymy while polysemy is represented in columns.

Table 1. The concept of Lexical Matrix

Word-Meaning	***Word-Forms*** Y1 Y2 Y3…………………Yn
X1	E1.1 E1.2
X2	E2.2
X3	E3.3
..	……
Xm	Em.n

Example, 'खग'of synset like{आकाशिायी, सायंग}gives the meanings'सायंग' (धातुकाबना हुआ पतला हतियार जो धनुष से चलाया जाता है)belongs to a synset, whose elements from a row in the matrix, and therow numbers will be a synset ID. 'खग'has several meanings, (पंख और चोचवालाद्विपदजिसकी उत्पत्ति अंडे से होती है)which comes in the column by the word.

E. Semantic relations

The lexical matrix isestablishedonto an importantcomponent of the human language organisation. Itaffirms the connection between form and meaning of the word. The Hindi WordNet is influencedthrough the English WordNet, and semantic relation utilises in lexical data. It has been used extensivelyin WordNet, and are primarily used to construct the lexicon, and the semantic relations are asfollow:

Table 2. Relations in Hindi WordNet

Relation	Meaning	Example
Hypernymy/ Hyponymy	Is-A (Kind-Of)	बेलपत्र Is A Kind- OF पत्ता
Entailment/ Troponymy	Manner-Of (for verbs)	खराटा लेना, नाकबजाना→सोना
Meronymy/ Holonymy	Has-A (Part-Whole)	जड़Is The Part-OF पेड़
Antonymy	holds between two words that express opposite meanings	मोटा→पतला
Causative	the pattern of making causation	खाना→ खिलाना

II. EXPERIMENTAL SETUP

A. LESK Algorithm

We have LESK Algorithm (given by Michael Lesk in 1986: using definition overlap for identifying senses of words in a given context) for words disambiguation.Machine-Readable Dictionary (MRD) is usually required in Overlap-Based Approaches.
LESK algorithm is illuminated as follows:
Suppose W is an ambiguous word and there is agroup of a context wordssets i.e. C inthe window surrounding it. We explore W in Hindi WordNet and then detect the senses of W. Hence, there is going to bemanysenses S for W. Let's find the most beneficial sense of W by using LESK algorithm. So, at last the most appropriate sense which has the maximum overlap will be the output of algorithm [23]. We can use simulated annealing to remove ambiguity if there is more than one word in the sentence.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Algorithm 1: Original LESK Algorithmimplemented byMichael Lesk in 1986

```
function LESK (word, sentence) returns best sense of word
best-sense ← 0
max-overlap ← 0
context ← set of words in sentence
for each sense in senses of word do
        content ← set of words in the gloss and examples of senses
        overlap ← COMPUTEOVERLAP (content, context)
        if overlap › max-overlap then
                max-overlap ← overlap
        best-sense ← sense
end return (best-sense)
```

The next step is to assign the sense count for each defined sense. Then weperform overlapping of the contextual meaning and the dictionary meaning.If they are found to be overlapping, we increment the instance count for each sense. The output describes the senses of the objective word is given. The sense with the maximalmarks is the right sense of the objective word.

Customized dictionary is used in the proposed system as an explanatory resource.It is trained to detect all the ambiguous words of Hindi and return their possible senses. If a word is not found in WordNet then it will store that word with senses. By performing overlapping

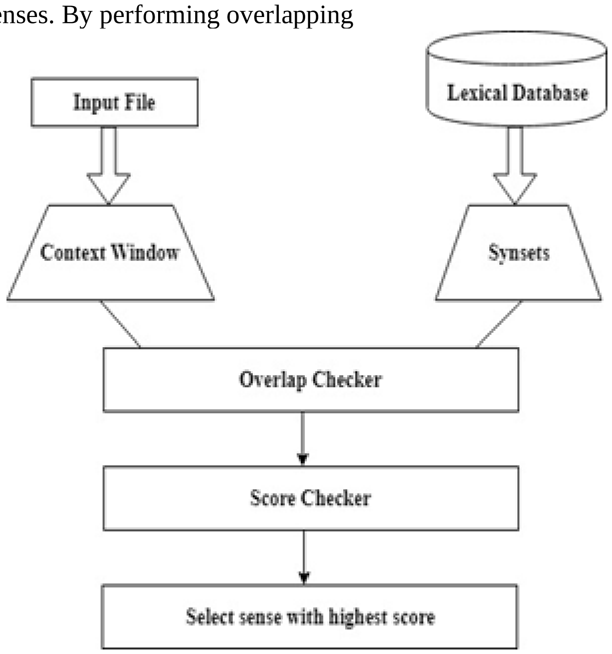

Figure 2. LESK Algorithm

between the given sense and context window for assigning a score to each sense.

B. Example:

By using the Hindi WordNet, we get the senses for the objective word 'सोना'. There are lots of senses for the word 'सोना'.

If we find the correct meaning as an example of this sentence:'राम के पास सोनेकी माला है।'

So, senses in Hindi WordNetfor the word 'सोना' are:

Query: 'सोने की माला'

1. सोना:As Noun

Sense 1: एक बहुमूल्य पीली धातु जिसके गहने आदि बनते हैं। (*आजकल सोने का भव आसमान छू रहा है।*)

Sense 2: सोने की क्रिया। (*सोने से शरीर को आराम मिलता है।*)

Sense 3: एक प्रकार का ऊँचा पेड़। (*सोनापाठा के बीज,छाल और फल दवा के रूप में काम आते हैं।*)

सोना: As Verb

Sense 1: लेटकर शरीर और मस्तिष्क को विश्राम देने वाली निद्रा की अवस्था में होना। (*थकावट के कारण आज वह जल्दी सो गया।*)

Sense 2: किसी कारण से रक्त संचार रुकने से हाथ या पैर के किसी भाग का सुन्न होना।(*सोते समय सीने के नीचे दब जाने के कारण मेरा हाथ सो गया है।*)

So, Hindi WordNet contains 5 different senses of word 'सोना'. If we link that 'सोना'word with 'माला'word then there will be several senses for the sentence 'सोने की माला'.

Senses for the word 'माला' in Hindi WordNet are:

2. माला: As Noun

Sense 1: मनका, फूल आदि को सूत आदि में गोलाकार पिरोकर बनाई हुई कोई वस्तु जो गले में पहनी जाती है (*उसके गले में मोतियों की माला सुशोभित हो रही थी।*)

Sense 2:बहुमंजिली इमारतों में ऊपर नीचे के विचार से बने मकान के स्तर (*मेरा घर सातवीं मंजिल पर है।*)

Sense 3:ऐसी परम्परा जिसमें एक ही प्रकार की वस्तुएँ, व्यक्ति या जीव एक दूसरे के बाद एक सीध में हों (*राशन की दुकान पर लोगों की पंक्ति लगी हुई थी।*)

So, Hindi WordNet contains 3 different senses of word 'माला'.

At last, the word'माला' has best sense 'गहना'according to LESK Algorithm. In the given sentence 'गहना' is an overlapping word, we get the outcome is'गहना'. In such process we are avoiding the stop words which are (का, के, की,है,हो, से, मे, नेetc.)

III. EVALUATION

In our paper, we have developed a tool for WSD by usingWordNet for Hindi. For this purpose, we considered a small corpus that contained 3000 ambiguous sentences.In this approach we evaluated word sense disambiguation using precision and recall. Precision (P)is correctly determining the percentage of correctly taggedwords out of the words addressed by our system i.e. 0.69 and recall (R) is defined as the number of correct answers given by the automatic system over the total number of answers i.e. 0.82. Until all the instances are tagged by our system, the value of the recall is always less accurate than precision.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Our system was able to produce 2142 correct senses out of the 3000 ambiguous words.

$$Accuracy = \frac{\#correct\ answer\ provided}{\#answer\ provided} \quad (1)$$

$$Accuracy = \frac{2142}{3000}\ X\ 100 = 71.4 \quad (2)$$

The overall accuracy of our system is 71.4.

IV. CONCLUSION

This is the paper in which we have centred mainly on WSD for Hindi. WSD has been a significant issue of NLP and AI,but finding the sense of words in respective contexts is the problem here. WSD requires a different kind of knowledge sources and techniques. We have utilised knowledge-based approach by implementing LESK algorithm for Hindi by using Hindi WordNet.

Our system gives 71.4% accuracy for finding a correct sense of an ambiguous word. We plan to add more knowledge sources like part-of-speech (POS) tagger etc. to improve the accuracy. It would be interesting to see how selectional restrictions can improve the accuracy of our current system.

REFERENCES

1. Bala, P. (2013). KnKnowledge-Basedpproach for Word Sense Disambiguation using Hindi Wordnet. The International Journal Of Engineering And Science (IJES), 2(4), 36-41.
2. Baldwin, T., Kim, S., Bond, F., Fujita, S., Martinez, D., & Tanaka, T. (2010). A reexamination of MRD-based word sense disambiguation. ACM Transactions on Asian Language Information Processing (TALIP), 9(1), 4.
3. Bhattacharyya, D. P. (2006). Hindi WordNet Data and Associated Software License Agreement. Indian Institute of Technology, Mumbai, CSE dept., TcTechnicaleport.
4. Chaplot, D. S., &Salakhutdinov, R. (2018). Knowledge-based word sense disambiguation using topic models. arXiv preprint arXiv:1801.01900.
5. Ekedahl, J., & Golub, K. (2004). Word sense disambiguation using Wordnet and the Lesk algorithm. Projektarbeten 2004, 17.
6. Gupta, M., Yadav, S., Sharma, S., & Surendra, Y. (2013). Word Sense Disambiguation Using Hindi WordNet and Lesk Approach. IPASJ International Journal of Computer Science (IIJCS), 1(6).
7. Kaur, A. (2010). Development of an approach for disambiguating ambiguous Hindi postposition. Department of Computer Science, Punjab University, India.
8. Kolte, S. G., &Bhirud, S. G. (2009). WordNet: a knowledge source for word sense disambiguation. International Journal of Recent Trends in Engineering, 2(4).
9. Mohapatra, R., &Hembram, L. (2010). Thought Boundary Detection in English Text through ththelawf Conservation of thought'for Word Sense Disambiguation. Language in India, 10(8).
10. Moldovan, D., &Novischi, A. (2004). Word sense disambiguation of WordNet glosses. Computer Speech & Language, 18(3), 301-317.
11. Montoyo, A., Suárez, A., Rigau, G., & Palomar, M. (2005). Combining knowledge-and corpus-based word-sense-disambiguation methods. Journal of Artificial Intelligence Research, 23, 299-330.
12. Moro, A., Raganato, A., &Navigli, R. (2014). Entity linking meets word sense disambiguation: a unifiedapproach. Transactions the Association for Computational Linguistics, 2, 231-244.
13. Munusamy, M., Arasu, G. T., Palanisamy, V., &Selvarajan, S. (2011). A Design and Development of Word Sense Disambiguation Algorithm for English Language Understanding for Database Access. Journal of Computer Science, 7(10), 1612.
14. Naseer, A., & Hussain, S. (2009). Supervised Word Sense Disambiguation for Urdu Using Bayesian Classification. Center for Research in Urdu Language Processing, Lahore, Pakistan.
15. Pal, A. R., Saha, D., & Pal, A. (2017). A KnKnowledge-basedethodology for Word Sense Disambiguation for Low Resource Language. Advances in Computational Sciences and Technology, 10(2), 267-283.
16. Pandey, P., & Arora, R. (2012). Cross-Lingual Word Sense Disambiguation using Wordnets and CoContext-basedapping.
17. Parameswarappa, S., & Narayana, V. N. (2011). Kannada word sense disambiguation for machine translation. International Journal of Computer Applications, 34(10).
18. Reddy, S., &Inumella, A. (2010, July). WSD as a distributed constraint opoptimisationroblem. In Proceedings of the ACL 2010 Student Research Workshop (pp. 13-18). Association for Computational Linguistics.
19. Reddy, S., Inumella, A., McCarthy, D., & Stevenson, M. (2010, July). IIITH: DoDomain-specificord sense disambiguation. In Proceedings of the 5th International Workshop on Semantic Evaluation (pp. 387-391). Association for Computational Linguistics.
20. Reddy, S., Inumella, A., Sangal, R., & Paul, S. (2009). All words unsupervised semantic category lalabellingorhiHindiIn Proceedings of the International Conference RANLP-2009(pp. 365-369).
21. Sankaran, B., & Vijay-Shanker, K. (2003). Influence of morphology in word sense disambiguation for TamilAU-KBC Research Centre.
22. Sharma, N., Kumar, S., & Niranjan, S. (2012). Using Machine Learning Algorithms for Word Sense Disambiguation: A Brief Survey. International Journal of Computer Technology and Electronics Engineering (IJCTEE) Volume, 2.
23. Sreedhar, J., Raju, S. V., Babu, A. V., Shaik, A., & Kumar, P. P. (2012). Word sense disambiguation: An empirical survey. International Journal of Soft Computing and Engineering (IJSCE) ISSN, 2231-2307.
24. Tamilselvi, P., & Srivatsa, S. K. (2011). Word Sense Disambiguation using CaCase-basedpproach with Minimal Features Set. Indian Journal of Computer Science and Engineering (IJCSE), 2(4), 628-633.
25. Trivedi, M., Sharma, S., &Deulkar, K. (2014). Approaches To Word Sense Disambiguation. International Journal of Engineering Research & Technology, 3(10), 645-647.
26. Yadav, P., & Husain, M. S. (2014). Study of Hindi Word Sense Disambiguation Based on Hindi Wordnet. IJRASET ISeSN.

AUTHORS PROFILE

Pooja Sharma received MCA degree from the Banasthali Vidyapith, Rajasthan, India in 2016 and currently doingM.Tech in computer science from Banasthali Vidyapith, Rajasthan, India. Her research activities mainly focused at the development of WSD algorithms and their techniques for Indian Languages and she also has interest in the field of NLP and AI.

Nisheeth Joshi Dr. Nisheeth Joshi is an Associate Professor in Department of Computer Science, Faculty Mathematics and Computing. He has done his PhD in the area of Natural Language Processing. Being involved in teaching and research for over 10 years, he has developed the art of explaining even the most complicated topics in a straight forward and easily understandable fashion. He also has vast experience of handling large scale research projects sponsored by various funding agencies. This has helped him in developing practical insights of complex AI systems. He has delivered invited talks in several conferences, workshops and training programs. He has authored several papers in international journals and conferences on various topics in AI and NLP.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)