A Study of Critical Approaches in WSD for Telugu Language Nouns: Current State of the Art

J.Sreedhar , Dr.S.Viswanadha Raju, Dr.A.Vinaya Babu

Abstract— *Word Sense Disambiguation (WSD) is the process of differentiating among senses of words. WSD plays a vital role to reduce the ambiguity about the words in the telugu language. Natural Language Processing (NLP) is a system which explores various methodologies to forecast the ambiguity between human languages. In the field of computational linguistics, some of the results have already been obtained even though, a number of important research problems have not been solved yet. In this article assessment of the Current State of the Art about "Critical Approaches in Word Sense Disambiguation for Telugu nouns" was discussed and further it contains short descriptive taxonomy of the NLP and WSD.*

Index Terms— NLP, WSD, POS, Ambiguity, Disambiguation and IR.

1 INTRODUCTION

In daily life every human being communicates with the language. So language is the vehicle for human communication. These languages are called Natural Languages. Processing of these languages computationally is called Natural Language Processing (NLP). For computing, Artificial Intelligence is the major area for processing of natural languages. Languages broadly classified into 2-types. They are Scripted Languages and Non Scripted Languages.

Scripted Languages can be represented with literature are known as scripted languages. These languages can able to visualize the information in a formatted text. Scripted languages are majorly divided into two types depending upon the popularity of the usage. They are English and Non English. English language is globally adapted language in human race. Many systems are designed in English due to global adaptation.

Non Scripted Languages cannot be represented with literature are known as nonscripted languages. These languages cannot be able to visualize the information in a formatted text. So we cannot able to write and read in a scripted manner. Just we can speak and listen to these languages. In rural areas maximum people are using non scripted languages. These languages don't have grammar rules and other regulations.

Ambiguity is the common phenomena in all the natural languages. Sometimes while speaking people are unable to understand the context. This will occur due to word ambiguity. Word ambiguity is not a major problem for human beings since through conversation they can resolve it. When this problem is switched to or turned to machine processing , it creates lot of difficulty to convert context into structured data.

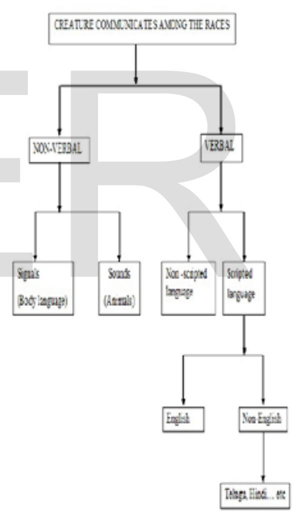

Figure1: Classification of languages

Language is the communication media for creatures among the races. Language can able to exchange the information among the races and it is the evolution criteria for the technology. Development of the languages can able to exchange the thoughts, views, suggestions in an understandable way. The development in the human races has been observed when the communication among the people started to increase from the rock age. Communication is broadly of two types such as Verbal and Nonverbal.

Nonverbal communication is adopted by the animal races. Nonverbal communication is the process of exchange of information with signs and sounds. The signs that looking seriously can able to understand the angriness. There are no words or written script for the non-verbal communication.

- *J.Sreedhar, Research Scholor, Department of CSE, JNTUK ,Kakinada, India-533003. Mobile No: 08790423564, (e-mail: sreedharyd@gamil.com)*
- *Dr. S.Viswanadha Raju, Professor & HOD – CSE,JNTUHCEJ,Hyderaba*
- *Dr.A.Vinaya Babu,Professor of CSE,JNTUH,Hyderabad.*

ISSN 2229-5518 s

This type of communication is observed in daily life.

Verbal communication is associated with alphabets, words, sentences etc. The perfectness of the language depends upon the grammar rules associated with it. Verbal languages can be classified as scripted and nonscripted languages.

The organization of the paper is as follows: Section 2 describes Taxonomy of Word Sense Disambiguation in Natural Language Processing; Section 3 explains Current State of the Art and Section 4 deals with Conclusion followed by References.

2 TAXONOMY OF THE WORD SENSE DISAMBIGUATION IN NLP

2.1 WSD

Word sense disambiguation (WSD) is the ability to computationally determine which sense of a word is activated by its use in a particular context. Classification tasks are studied in the area of NLP (for an introduction see Manning and Schutze [1] in 1999 and Jurafsky and Martin [2] in 2000), such as parts-of-speech tagging (i.e., the assignment of parts of speech to target words in context), named entity resolution (the classification of target textual items into predefined categories), text categorization (i.e., the assignment of predefined labels to target texts) etc.

We can distinguish two variants of the generic WSD task: Lexical sample where a system is required to disambiguate a restricted set of target words usually occurring one per sentence. Supervised systems are typically employed in this setting, as they can be trained using a number of hand-labeled instances (training set) and then applied to classify a set of unlabeled examples (test set).

All-words WSD, where systems are expected to disambiguate all open-class words in a text (i.e., nouns, verbs, adjectives, and adverbs). This task requires wide-coverage systems. Consequently, purely supervised systems can potentially suffer from the problem of data sparseness, as it is unlikely that a training set of adequate size is available which covers the full lexicon of the language of interest. On the other hand, other approaches, such as knowledge-lean systems, rely on full-coverage knowledge resources, whose availability must be assured. your manuscript electronically for review.

2.2 NLP

Natural Language Processing (NLP) is a technique which is computerized approach to analyze text that is based on theories as well as technologies and both. NLP is a health research and development area in Artificial Intelligence (AI). It can be defined in various forms depending upon the scholar in the history that means which is not having a unique definition.

Definition: Natural Language Processing (NLP) is theoretically motivated methods and techniques which are selected for the accomplishment of particular type of language. It is used in analyzing and representing a human communication at one or more level of linguistic analysis. The purpose is to achieve human like languages processing, for a range of tasks or applications.

hen Natural language processing goal is to accomplish human like language processing. The word processing is very calculated that should not be replaced with "understanding". Formerly natural language processing (NLP) is referred to Natural Language Understanding (NLU). A NLU system would be able to paraphrase an input text, translates the text into another language, answer questions about the contents of the text and draw inferences from the text. Information retrieval systems (IR) works are based on NLP. This system is used to provide more accurate results to the users. The goal of the NLP system is to display true meaning of the intent of the user query.

3 CURRENT STATE OF THE ART

Here in Current State of the Art shows the up to date research approaches and their solutions in a standard manner. The following are the some of the issues discussed in the earlier stages.

Walker [3,47] proposed an algorithm which is considering a thesaurus; each word is assigned to one or more subject categories in the thesaurus. There are several subjects assigned with a word then it is assumed that they correspond to different senses of the word. Black applied walker's approach to choose five different words and achieved accuracies of 50% .

Wilk [4,47] suggested that dictionary glosses are too short to result reliable disambiguation. Later he developed a context vector approach that expand the glosses with related words which allows for matching to be based on one or more words in the year 1990 by using the Longman's dictionary of contemporary English (LDOCE). Walker's approach has controlled definition vocabulary of appx 2200 words which increase the likelihood of finding overlap among word sense.

Lesk [5] developed various ideas for future research and in fact several issues he raised to continue the research even today. Lesk algorithm be used to disambiguate all the words in a sentence at once, or should it proceed sequentially, from one word to the next. If it did proceed sequentially, should the previously assigned senses influence the outcome of the algorithm for following words.

Quillian[6] in the mid 1960 said that the way to use the content of a machine readable dictionary to make inferences about word meaning and proposed the semantic network representation of dictionary contents. Here the node represents each meaning of the word, for defining the concept in the dictionary; this node is used to connect words. Content words in the definitions are in turn connected to the words that are used to define them to create a large web of words.

Cowie[7,47] said that the Lesk algorithm is capable of disambiguation all the words in the sentence simultaneously. Computation complexity of such an undertaking is enormous and makes it difficult in practice. The simulated annealing method is used to search the senses in sentence of all words. An exhaustive search has done to find a solution that globally op-

ISSN 2229-5518 s
timized the assignment.

Kozima and Furugori [8] constructed a network from LDOCE glosses that consist of nodes representing the controlled vocabulary and links to show the co-occurrence of these words in glosses. They define a measure based on spreading activation that results in a numeric similarity score between two concepts.

Pedersen, Banerjee and Patwardhan[10] suggested that Semantic relatedness to perform word sense disambiguation is measured by an algorithm. It finds its root in the original Lesk algorithm which disambiguates a polysemous word. It picks that sense of the target word whose definition has the most words in common with the definitions of other words in a given window of content. Lesk's intuition was that related word senses will be defined using similar words. The overlap in their definitions will indicate their Relatedness, a algorithm that performs disambiguation using any measure, that return a relatedness or similarity score for pairs of word senses.

Nitwa and Nitta [11] developed that Context vectors derived from co-occurrence statistic of large corpora and vectors derived from the path length in a network that represent their co–occurrence in dictionary definitions. They construct a Quillian style network, words that occur together in definitions are linked and those words are linked to the words that are used in definitions and so forth. They evaluate Wilks context vector method of disambiguation and find that dictionary context is more suitable source of co-occurrence information than other corpora.

Sussna[12,44,48,49] proposed a disambiguation algorithm assigns a sense to each noun in a window of context by minimizing a semantic distance function among their possible senses. While this is quite similar to our approach of disambiguation. His disambiguation algorithm is based on a measure of relatedness among nouns that he introduces. This measure requires that weights be set on edges in the WordNet noun hierarchy, based on the type of relation the edge represents. His measure accounts for is- a relations, as well as has-part, is-a-part-of, and antonyms.

Agirre and Rigau [13] introduced a similarity measure based on conceptual density and apply it to the disambiguation of nouns. It is based on the is-a hierarchy in WordNet, and only applies to nouns. This measure is similar to the disambiguation technique proposed by Wilks, in that it measures the similarity between a target noun sense and the nouns in the surrounding context.

Rivest [14] in 1987 is proposed a decision list algorithm. It describes an ordered set of rules for categorizing test instances (in the case of WSD, for assigning the appropriate sense to a target word). It can be seen as a list of weighted "if-then-else" rules.

Kelly and Stone[15] in 1975 proposed decision tree algorithm. It explores a predictive model used to represent classification rules with a tree structure that recursively partitions the training data set. Each internal node of a decision tree represents a test on a feature value, and each branch represents an outcome of the test. A prediction is made when a terminal node (i.e., a leaf) is reached.

Naïve Bayes [25] proposed a probabilistic classifier algorithm based on the application of Bayes' theorem. McCulloch and Pitts [16] in 1943 proposed a neural network which is an interconnected group of artificial neurons that uses a computational model for processing data based on a connectionist approach. Pairs of input feature, desired response are input to the learning program. The aim is to use the input features to partition the training contexts into nonoverlapping sets corresponding to the desired responses.

Cottrell[17] in 1989 employed neural networks to represent words as nodes: the words activate the concepts to which they are semantically related and vice versa. The activation of a node causes the activation of nodes to which it is connected by excitory links and the deactivation of those to which it is connected by inhibitory links (i.e., competing senses of the same word).

Veronis and Ide [18] in 1990 built a neural network from the dictionary definitions of the Collins English Dictionary. They connect words to their senses and each sense to words occurring in their textual definition.

Tsatsaronis et al. [19] in 2007 successfully extended their approach to include all related senses linked by semantic relations in the reference resource that is WordNet.

Towell and Voorhees [20] in 1998 found that neural networks perform better without the use of hidden layers of nodes and used perceptrons for linking local and topical input features directly to output units (which represent senses).

Boser et al. [21] in 1992 is based on the idea of learning a linear hyperplane from the training set that separates positive examples from negative examples. The hyperplane is located in that point of the hyperspace which maximizes the distance to the closest positive and negative examples (called support vectors). In other words, support vector machines (SVMs) tend at the same time to minimize the empirical classification error and maximize the geometric margin between positive and negative examples.

SVM has been applied to a number of problems in NLP, including text categorization [Joachims [22] in 1998], chunking [Kudo and Matsumoto [23] in 2001], parsing [Collins [24] in 2004], and WSD Escudero et al. [25] in 2000, Murata et al. [26] in 2001, Keok and Ng [27] in 2002].

Klein and Florian et al. [28,42] in 2002 studied the combination of supervised WSD methods, achieving state-of-the-art results on the Senseval-2 lexical sample task. Brody and Navigli et al.[29,45,50] in 2006 reported a study on ensembles of unsupervised WSD methods. When employed on a standard test set, such as that of the Senseval-3 all-words WSD task, ensemble methods overcome state-of-the-art performance among unsupervised systems (up to +4% accuracy). Single classifiers can be combined with different strategies: here they introduce majority voting, probability mixture, rank-based combination, and AdaBoost.

Kilgarriff and Grefenstette [30] in 2003, viewing the Web as corpus, which is an interesting idea which has been and is currently exploited to build annotated data sets, with the aim to relieve the problem of data sparseness in training sets. They can annotate such a large corpus with the aid of monosemous relatives by way of a bootstrapping algorithm similar to Yarowsky's [31] in 1995, starting from a few number of seeds. As a result, they can use the automatically annotated data to train WSD classifiers.

Banerjee and Pedersen [32] suggested that the main advantage of the original Lesk algorithm. Network relations provided in WordNet, rather than simply consider the glosses of the surrounding words in the sentence. The concept network of WordNet is exploited to allow for glosses if word senses. There are related to the words in the context to be compared. The glosses of surrounding words in the text are expanded to include glosses of those words, which are related through relations in WordNet. They also suggest a set of n one word matches weighted less heavily than scoring schema such that a match of n consecutive words in two glosses.

Agirre et al. [33,50] in 2001 studied the performance of topic signatures in disambiguating a small number of words and found out that they do not seem to provide a relevant contribution to disambiguation. In contrast, in a recent study on large-scale knowledge resources, Cuadros and Rigau [34] in 2006 showed that automatically acquired knowledge resources perform better than hand labeled resources when adopted for disambiguation in the Senseval-3 lexical sample task.

Gale et al. [35,43] in 1992b suggested an unsupervised methods have the potential to overcome the knowledge acquisition bottleneck which is, the lack of large-scale resources manually annotated with word senses. These approaches to WSD are based on the idea that the same sense of a word will have similar neighboring words. They are able to induce word senses from input text by clustering word occurrences, and then classifying new occurrences into the induced clusters.

Schutze [36] in 1992 described a set of unsupervised approaches which are based on the notion of context clustering. Each occurrence of a target word in a corpus is represented as a context vector. The vectors are then clustered into groups, each identifying a sense of the target word. A historical approach of this kind is based on the idea of word space.

Widdows and Dorow [37] in 2002 defines the construction of a cooccurrence graph which is based on grammatical relations between words in the contex. Van Dongen [38] in 2000 suggested the Markov clustering algorithm is applied to determine the word senses, which is based on an expansion and an inflation step, aiming, respectively, at inspecting new more distant neighbors and supporting more popular nodes.

Veronis [39] in 2004 was proposed an adhoc approach called HyperLex. Here a cooccurrence graph is built such that nodes are words occurring in the paragraphs of a text corpus in which a target word occurs and an edge between a pair of words is added to the graph if they cooccur in the same paragraph. Each edge is assigned a weight according to the relative cooccurrence frequency of the two words connected by the edge.

Brin and Page[40,41,46] in 1998 explored an alternative graph-based algorithm for inducing word senses is PageRank. PageRank is a well-known algorithm developed for computing the ranking of web pages and is the main ingredient of the Google search engine. It has been employed in several research areas for determining the importance of entities whose relations can be represented in terms of a graph.

4 Conclusion

In this paper we presented the current state of the art about word sense disambiguation and also we understood the need and necessity of WSD. We analyzed, measured many approaches and found the right path towards to extract the nouns in WSD. So we are very particular about our future work in this direction

Acknowledgment

We are very thankful to all the esteemed authors in a reference list, to make this research article in a better shape and in a right direction.

References

[1] Manning, C. and Schutze, H. Foundations of Statistical Natural Language Processing. MIT Press, Cambridge, MA, 1999.

[2] Jurafsky, D. and Martin, J. Speech and Language Processing. Prentice Hall, Upper Saddle River, NJ, 2000.

[3] Basak Mutlum, Word Sense Disambiguation http://www.denizyuret.com/students/bmutlum/index.htm

[4] Y. Wilks, D. Fass, C. Guo, J. McDonald, T. Plate, B. Slator, Providing machine tractable dictionary tools, Machine Translation 5, 99–154, 1990.

[5] Manish Sinha, Mahesh Kumar, Prabhakar Pande, Lakshmi Kashyap and Pushpak Bhattacharyya, , Hindi Word Sense Disambiguation, International Symposium on Machine Translation, Natural Language Processing and Translation Support Systems, Delhi, India, November, 2004 http://www.cse.iitb.ac.in/~pb/papers/HindiWSD.pdf

[6] M. Quillian, Semantic memory, in: M. Minsky (Ed.), Semantic Information Processing, the MIT Press, Cambridge, MA, pp. 227–270, 1968.

[7] J. Cowie, J. Guthrie, L. Guthrie, Lexical disambiguation using simulated annealing, in: Proceedings of the 14th International Conference on Computational Linguistics, Nantes, France, pp. 359–365,1992.

[8] H. Kozima, T. Furugori, Similarity between words computed by spreading activation on an english dictionary, in: Proceedings of the

6th Conference of the European Chapter of the Association for Computational Linguistics, Utrecht, pp. 232–239,1993.
[9] J. Veronis, N. Ide, Word sense disambiguation with very large neural networks extracted from machine readable dictionaries, in: Proceedings of the 13th International Conference on Computational Linguistics, Helsinki, pp. 389–394,1990.
[10] S. Banerjee, T. Pedersen, Extended gloss overlaps as a measure of semantic relatedness, in: Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence, Acapulco, pp. 805–810, 2003.
[11] Y. Niwa, Y. Nitta, Co-occurrence vectors from corpora versus distance vectors from dictionaries, in: Proceedings of the Fifteenth International Conference on Computational Linguistics, Kyoto, Japan, pp. 304–309, 1994.
[12] M. Sussna, Word sense disambiguation for free-text indexing using a massive semantic network, in: Proceedings of the Second International Conference on Information and Knowledge Management, pp. 67–74, 1993.
[13] E. Agirre, G. Rigau, Word sense disambiguation using conceptual density, in: Proceedings of the 16th International Conference on Computational Linguistics, Copenhagen, pp. 16–22, 1996.
[14] Rivest, R. L. Learning decision lists. Mach. Learn. 2, 3, 229–246, 1987.
[15] Kelly, E. and Stone, P. Computer Recognition of English Word Senses. Vol. 3 of North Holland Linguistics Series. Elsevier, Amsterdam, The Netherlands, 1975.
[16] Mcculloch, W. and Pitts, W. A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. 5, 115–133, 1943.
[17] Cottrell, G. W. A Connectionist Approach to Word Sense Disambiguation. Pitman, London, U.K, 1989.
[18] Veronis, J. and Ide, N. Word sense disambiguation with very large neural networks extracted from machine readable dictionaries. In Proceedings of the 13th International Conference on Computational Linguistics (Coling, Helsinki, Finland). 389–394, 1990.
[19] Tsatsaronis, G., Vazirgiannis, M., and Androutsoppulos, I. Word sense disambiguation with spreading activation networks generated from thesauri. In Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI, Hyderabad, India). 1725–1730, 2007.
[20] Towell, G. and Voorhees, E. Disambiguating highly ambiguous words. Computat. Ling. 24, 1, 125–145, 1998.
[21] Boser, B. E., Guyon, I. M., and Vapnik, V. N. A training algorithm for optimal margin classifiers. In Proceedings of the 5th Annual Workshop on Computational Learning Theory (Pittsburgh, PA). 144–152, 1992.
[22] Joachims, T. Text categorization with support vector machines: Learning with many relevant features. In Proceedings of the 10th European Conference on Machine Learning (ECML, Heidelberg, Germany). 137–142, 1998.
[23] Kudo, T. and Matsumoto, Y. Chunking with support vector machines. In Proceedings of NAACL (Pittsburgh, PA). 137–142, 2001.
[24] Collins,M. Parameter estimation for statistical parsing models: Theory and practice of distributionfree methods. In New Developments in Parsing Technology, H. Bunt, J. Carroll, andG. Satta, Eds. Kluwer, Dordrecht, The Netherlands, 19–55, 2004.
[25] Escudero, G., Marquez, L., and Rigau, G.Naive Bayes and exemplar-based approaches to word sense disambiguation revisited. In Proceedings of the 14th European Conference on Artificial Intelligence (ECAI, Berlin, Germany). 421–425, 2000b.
[26] Murata, M.,Utiyama, M.,Uchimoto, K.,MA, Q., and Isahara,H. Japanese word sense disambiguation using the simple Bayes and support vector machine methods. In Proceedings of the 2nd International Workshop on Evaluating Word Sense Disambiguation Systems (Senseval-2, Toulouse, France). 135–138, 2001.
[27] Keok, L. Y.and NG, H. T. An empirical evaluation of knowledge sources and learning algorithms for word sense disambiguation. In Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP, Philadelphia, PA). 41–48, 2002.
[28] Klein, D., Toutanova, K., Ilhan, T. H., Kamvar, S. D., and Manning, C. D. Combining heterogeneous classifiers for word-sense disambiguation. In Proceedings of the ACL workshop on Word Sense Disambiguation: Recent Successes and Future Directions (Philadelphia, PA). 74–80, 2002.
[29] Brody, S.,Navigli, R., and Lapata,M. Ensemble methods for unsupervised WSD. In Proceedings of the 44th Annual Meeting of the Association for Computational Linguistics joint with the 21st International Conference on Computational Linguistics (COLING-ACL, Sydney, Australia). 97–104, 2006.
[30] Kilgarriff. A. and Grefenstette,G. Introduction to the special issue on the Web as corpus. Computat. Ling. 29, 3, 333–347, 2003.
[31] Yarowsky, D. Unsupervised word sense disambiguation rivaling supervised methods. In Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics (Cambridge, MA). 189–196, 1995.
[32] S. Banerjee, T. Pedersen, Extended gloss overlaps as a measure of semantic relatedness, in: Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence, Acapulco, pp. 805–810, 2003 .
[33] Agirre, E. and Martinez, D. Learning class-to-class selectional preferences. In Proceedings of the 5th Conference on Computational Natural Language Learning (CoNLL, Toulouse, France). 15–22, 2001.
[34] Cuadros, M. and Rigau, G. Quality assessment of large scale knowledge resources. In Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP, Sydney, Australia).534–541, 2006.
[35] Gale, W. A., Church, K., and Yarowsky, D. A method for disambiguating word senses in a corpus. Comput. Human. 26, 415–439, 1992b.
[36] Schutze, H. Dimensions of meaning. In Supercomputing ’92: Proceedings of the 1992 ACM/IEEE Conference on Supercomputing. IEEE Computer Society Press, Los Alamitos, CA. 787–796, 1992.
[37] Widdows, D. and Dorow, B. A graph model for unsupervised lexical acquisition. In Proceedings of the 19th International Conference on Computational Linguistics (COLING, Taipei, Taiwan). 1–7, 2002.
[38] Van Dongen, S. Graph Clustering by Flow Simulation, Ph.D. dissertation. University of Utrecht, Utrecht, The Netherlands, 2000.
[39] Veronis, J. Hyperlex: Lexical cartography for information retrieval. Comput. Speech Lang. 18, 3, 223–252, 2004.
[40] Brin, S. and Page, M. Anatomy of a large-scale hypertextual Web search engine. In Proceedings of the 7th Conference on World Wide Web (Brisbane, Australia). 107–117, 1998.
[41] E. Agirre and A. Soroa "Personalizing pagerank for word sense disambiguation", Proceedings of the 12th conference of the European chapter of the Association for Computational Linguistics (EACL-2009), pp.33 -41 2009
[42] Ping ChenBowes, C. ; Wei Ding ; Choly, M. “Word Sense Disambiguation with Automatically Acquired Knowledge” , 46 – 55,ISSN :1541-1672,August 2012.
[43] R. Navigli and M. Lapata "An Experimental Study of Graph Connectivity for Unsupervised Word Sense Disambiguation", IEEE Trans.

Pattern Anal. Mach. Intell., vol. 32, no. 4, pp.678 -692 2010.
[44] Lei Guo, Xiaodong Wang, Jun Fang, 'Ontology Clarification by Using Semantic Disambiguation', 978 -1-4244-1651-6/08, IEEE 2008.
[45] S. P. Ponzetto, R. Navigli, "Knowledge-rich Word Sense Disambiguation rivaling supervised systems," Proc. of the 48th Annual Meeting of the Association for Computational Linguistics (ACL), pp. 1522-1531, 2010.
[46] Roberto Navigli and Paola Velardi, "Structural semantic interconnections: A knowledge-based approach to word sense disambiguation", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 27, No. 7, pp. 1075-1086, July 2005.
[47] Boshra F. ZoponAl_Bayaty and Dr.Shashank Joshi"Word Sense Disambiguation (WSD) and Information Retrieval (IR): Literature Review" ijarcsse,Volume 4, Issue 2,ISSN: 2277 128X, February 2014.
[48] Lucia Specia, Sujay Kumar Jauhar, RadaMihalcea, SemEval 2012 Task 1: English Lexical Simplification, in Proceedings of the SemEval-2012 Workshop on Semantic Evaluation Exercises, Montreal, Canada, June 2012.
[49] Kulkarni, M.Comput. Eng. Dept., V.J.T.I., Mumbai, India,Sane, S. "An ontology clarification tool for word sense disambiguation"2011 3rd International Conference, IEEE - Electronics Computer Technology (ICECT),292 – 296, E-ISBN :978-1-4244-8679-3, 8-10 April 2011.
[50] RadaMihalcea, "Knowledge-Based Methods for WSD", e-ISBN 978-1-4020-4809-2, Springer 2007.