ISSN 2249-6343
International Journal of Computer Technology and Electronics Engineering (IJCTEE)
Volume 2, Issue 2

Word Sense Disambiguation Using Association Rules: A Survey

Samit Kumar, Neetu Sharma, Dr. S. Niranjan

***Abstract* --- Word sense disambiguation (WSD) is defined as the task of assigning the appropriate meaning (sense) to a given word in a text or discourse. The sense in which the word is used can be determined, most of the times, by the context in which the word occurs. Word sense ambiguity is a central problem for many established Human Language Technology applications (e.g., machine translation, information extraction, question answering, information retrieval, text classification, and text summarization). The context of an ambiguous word is regarded as a transaction record, the words in the context and the senses of the ambiguous word are regarded as items. If some items frequently occur together in some transactions (the context of the ambiguous word), then there must be some correlation between the items. The basic idea of the WSD algorithm based on mining association rules is: to discover the frequent item sets composed of the sense of the ambiguous word and its context by scanning its context database, which support degree is no less than the threshold of support degree; to produce the association rules X=>Y which confidence degree is no less than the threshold of the confidence degree from maximum frequent item sets; at last to determine the sense of the ambiguous word by choosing the sense which the most association rules deduced.**

***Index Terms*—Association Rules, Context, Machine Translation, Word Net, Word Sense Disambiguation.**

I. Introduction

Word sense disambiguation (WSD) is the ability to identify the meaning of words in context in a computational manner. WSD is considered an AI-complete problem, that is, a task whose solution is at least as hard as the most difficult problems in artificial intelligence for machine translation, information extraction, question answering, information retrieval, text classification [2]. It is the most open problem in Natural language processing. Lots of good work has been done, but still more to do to improve the accuracy of the system. However to date, no large scale, broad coverage accurate WSD system has been built. Human language is ambiguous, so that many words can be interpreted in multiple ways depending on the context in which they occur.

Manuscript received March 27, 2012.

***Samit Kumar**, Ph. D. Scholar, Department of Computer Science and Engineering, Mewar University, (email- samit1981@gmail.com Chittorgarh, Rajasthan, India.*

***Neetu Sharma**, Ph. D. Scholar, Department of Computer Science and Engineering, Mewar University, (email- neetush75@gnmail.com) Chittorgarh, Rajasthan, India.*

***S. Niranjan, Principle,** PDM College of Engineering, Bahadurgarh, (e-mail: niranjan.hig41@gmail.com)Haryana, India.*

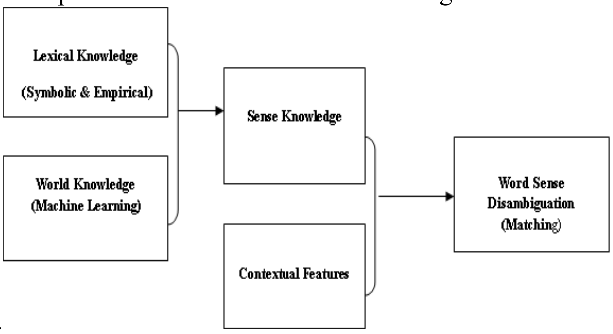
Some words have multiple meanings [1]. This is called Polysemy. E.g. bank can be a financial institute or a river shore. Sometime two completely different words are spelled the same. For example: Can can be used as model verb: You can do it, or as container: She brought a can of soda. This is called Homonymy. A word can have more than one sense, which can be determined by the context in which the word occurs. Word sense disambiguation is the task of assigning the appropriate meaning or sense to a given word in a context. Consider the word "bass", with two distinct senses:

(a) I can hear bass sounds.

(b) They like grilled bass.

The occurrences of the word bass in the two sentences clearly denote different meanings: low-frequency tones and a type of fish, respectively. To human it is obvious the first sentence is using the word "bass" in sense1 above, and in the second sentence it is used in sense2. Unfortunately, the identification of the specific meaning that a word assumes in context is only apparently simple. Its acknowledged difficulty does not originate from a single cause, but rather from a variety of factors. First, the task lends itself to different formalizations due to fundamental questions, like the approach to the representation of a word sense (ranging from the enumeration of a finite set of senses to rule-based generation of new senses), the granularity of sense inventories (from subtle distinctions to homonyms), the domain-oriented versus unrestricted nature of texts, the set of target words to disambiguate (one target word per sentence vs. "all-words" settings), etc. Second, WSD heavily relies on knowledge. In fact, the skeletal procedure of any WSD system can be summarized as follows: given a set of words (e.g., a sentence or a bag of words), a technique is applied which makes use of one or more sources of knowledge to associate the most appropriate senses with words in context. Knowledge sources can vary considerably from corpora (i.e., collections) of texts, either unlabeled or annotated with word senses, to more structured resources, such as machine readable dictionaries, semantic networks, etc. Without knowledge, it would be impossible for both humans and machines to identify the meaning, for example, of the above sentences and is called the knowledge acquisition bottleneck [Gale et al. 1992b]. The hardness of WSD is also attested by the lack of applications to real-world tasks. The exponential growth of the Internet community, together with the fast pace development of several areas of information technology (IT), has led to the production of a amount of unstructured data, such as document warehouses, Web pages, collections of scientific articles, blog corpora, etc.

As a result, there is an increasing urge to treat this mass of information by means of automatic methods. The word to be sense tagged always appears in a context. Context can be represented by a vector, called a context vector (word, features). Thus, we can disambiguate word sense by matching a sense knowledge vector and a context vector. The conceptual model for WSD is shown in figure I



.

Fig. I Conceptual Model for Word Sense Disambiguation

Lexical disambiguation in its broadest definition is nothing less than determining the meaning of every word in context, which appears to be a largely unconscious process in people. WSD is essentially a task of classification: word senses are the classes, the context provides the evidence, and each occurrence of a word is assigned to one or more of its possible classes based on the evidence.

II. BRIEF HISTORY

The task of WSD is a historical one in the field of Natural Language Processing (NLP). WSD was first formulated as a distinct computational task during the early days of machine translation in the late 1940s, making it one of the oldest problems in computational linguistics. Weaver (1949) introduced the problem in his now famous memorandum on machine translation:

If one examines the words in a book, one at a time through an opaque mask with a hole in it one word wide, then it is obviously impossible to determine, one at a time, the meaning of words. “Fast” may mean “rapid”; or it may mean “motionless”; and there is no way of telling which. But, if one lengthens the slit in the opaque mask, until one can see not only the central word in question but also say N words on either side, then, if N is large enough one can unambiguously decide the meaning.

At that time, researchers had already in mind essential ingredients of WSD, such as the context in which a target word occurs, statistical information about words and senses, knowledge resources, etc. Very soon it became clear that WSD was a very difficult problem, also given the limited means available for computation. The 1950s then saw much work in estimating the degree of ambiguity in texts and bilingual dictionaries, and applying simple statistical models. Zipf (1949) published his “Law of Meaning” that accounts for the skewed distribution of words by number of senses, that is, that more frequent words have more senses than less frequent words in a power-law relationship.

Kaplan (1950) determined that two words of context on either side of an ambiguous word was equivalent to a whole sentence of context in resolving power. Indeed, its acknowledged hardness [Bar-Hillel 1960] was one of the main obstacles to the development of MT in the 1960s. Madhu and Lytle (1965) calculated sense frequencies of words in different domains – observing early on that domain constrains sense – and then applied Bayes formula to choose the most probable sense given a context. During the 1970s the problem of WSD was attacked with AI approaches aiming at language understanding (e.g., Wilks [1975]). However, generalizing the results was difficult, mainly because of the lack of large amounts of machine-readable knowledge.

The 1980s were a turning point for WSD. Large-scale lexical resources and corpora became available so handcrafting could be replaced with knowledge extracted automatically from the resources (Wilks et al. 1990).The 1990s led to the massive employment of statistical methods and the establishment of periodic evaluation campaigns of WSD systems, up to the present days. Dictionary-based WSD had begun and the relationship of WSD to lexicography became explicit. For example, Guthrie et al. (1991) used the subject codes (e.g., Economics, Engineering, etc.) in the Longman Dictionary of Contemporary English (LDOCE) (Procter 1978) on top of Lesk’s method.

Although dictionary methods are useful for some cases of word sense ambiguity (such as homographs), they are not robust since dictionaries lack complete coverage of information on sense distinctions. The 1990s saw three major developments: WordNet became available, the statistical revolution in NLP swept through, and Senseval began.

WordNet (Miller 1990) pushed research forward because it was both computationally accessible and hierarchically organized into word senses called synsets. Today, English WordNet (together with WordNet for other languages) is the most-used general sense inventory in WSD research. Statistical and machine learning methods have been successfully applied to the sense classification problem. Today, methods that train on manually sense-tagged corpora (i.e., supervised learning methods) have become the mainstream approach to WSD, with the best results in all tasks of the Senseval competitions. Brown et al. (1991) were the first to use corpus-based WSD in statistical MT. Before Senseval, it was extremely difficult to compare and evaluate different systems because of disparities in test words, annotators, sense inventories, and corpora.

Senseval was first discussed in 1997 (Resnik and Yarowsky 1999; Kilgarriff and Palmer 2000) and now after hosting three evaluation exercises has grown into the primary forum for researchers to discuss and advance the field. Its main contribution was to establish a framework for WSD evaluation that includes standardized task descriptions and an evaluation methodology. It has also focused research, enabled scientific rigor, produced benchmarks, and generated substantial resources in many languages (e.g., sense-annotated corpora), thus enabling research in languages other than English.

At the Senseval-3 workshop (Mihalcea and Edmonds 2004) there was a general consensus (and a sense of unease) that the traditional explicit WSD task, so effective at driving research, had reached a plateau and was not likely to lead to fundamentally new research. This could indicate the need to look for new research directions in the field, some of which may already be emerging, for instance the use of parallel bilingual corpora. In the 4th workshop, SemEval-2007, the nature of the tasks evolved to include semantic analysis tasks outside of word sense disambiguation. SemEval-2010 encouraged tasks for different languages, cross-lingual tasks, and tasks that are relevant to particular NLP applications such as machine translation, information retrieval and information extraction.

Chunhui Zhang, Yiming Zhou and Trevor Martin [3] proposed a novel unsupervised genetic word sense disambiguation (GWSD) algorithm. They first uses WordNet to determine all possible senses for a set of words, then a genetic algorithm is used to maximize the overall semantic similarity on this set of words. GWSD is tested on two sets of domain terms and obtains good results. A weighted genetic word sense disambiguation (WGWSD) algorithm is then proposed to disambiguate words in a general corpus.

Myunggwon Hwang Chang Choi Byungsu Youn Pankoo Kim [4] proposed new WSD method. For WSD, they design a relation structure (RS) that measures relationship and distance between terms using WordNet . Based on the RS, they could grasp an exact sense of word. Through an experiment and evaluation, this method was as wonderful as a structural semantic interconnection (SSI) that was the best research for WSD.

S.K.Jayanthi and S. Prema[5] explored the development and subsequent evaluation of word sense disambiguation sing Brill tagger for several search engines which demonstrates increased precision from a sense based vector space retrieval model over traditional TF*IDF techniques.

Jiang Tao, Tan Ah-Hwee, and Wang Ke[6] proposed a systematic approach for discovering knowledge from free-form textual Web content. Specifically, they presented an automatic semantic relation extraction strategy to extract RDF metadata from textual Web content and an algorithm known as GP-Close for mining generalized patterns from RDF metadata. The experimental result proved that the GP-Close algorithm based on mining closed generalization closures can substantially reduce the pattern redundancy and perform much better than the original generalized association rule mining algorithm Cumulate in terms of time efficiency. The pattern analysis based on human validation shows that the proposed method is promising and useful.

Gaona, M.A.R. Gelbukh, A. Bandyopadhyay, S. [7] presented a measure for sense assignment useful for the simple Lesk algorithm. They use word co-occurrences of the gloss and the context, which is statistical information retrieved from the Web. In the SemCor data this method always gives an answer. On the Senseval 2 data, variant of the Lesk method outperformed some other Lesk-based methods.

Zhang Zheng, Zhu Shu [8] introduced and contrasted the main approaches of WSD prevailing in the world, and analyzes their advantages and disadvantages briefly. Then the author focused on the vector space model (VSM), and furthermore, puts forward a new method that uses the approach of multi-level sentence similarity (MLSS) computation in the VSM. The new method improves the accuracy of VSM method and overcomes the "bag of words" problem in VSM.

Xiao Guo, Dayou Li, Clapworthy, G. [10] proposed a conditional mutual information based selectional association to measure the selectional preference between two words in the same sentence. This selectional association is integrated conditional mutual information and a syntactic knowledge called link grammar. The selectional association is applied to indicative words selection for target words disambiguation. The experimental results show that this conditional mutual information based selectional association is able to select the appropriate word to indicate the appropriate meaning to the target word in different context.

Chatterjee, N.; Misra, R[11] presented a trainable model for word sense disambiguation (WSD) for resolving this ambiguity. The proposed model applies concepts of information theory to find the appropriate sense of a word when the context is known. Given a training text tagged with the correct senses of a particular word, their model learns to classify each occurrence of the target word with its correct sense in the unseen text.

S. G. Kolte and S. G. Bhirud[12] proposed the methodology for Word Sense Disambiguation based on domain information and WordNet hierarchy. Domain is a set of words in which there is a strong semantic relation among the words. The words in the sentence contribute to determine the domain of the sentence. The availability of WordNet domains makes the domain-oriented text analysis possible. The domain of the target word can be fixed based on the domains of the content words in the local context. This approach can be effectively used to disambiguate nouns. They presented the unsupervised approach to Word Sense.

M. Nameh, S.M. Fakhrahmad, M. Zolghadri Jahromi [14] presented a supervised learning method for WSD, which is based on Cosine Similarity. As the first step, they extract two sets of features; the set of words that have occurred frequently in the text and the set of words surrounding the ambiguous word. Then they presented the results of evaluating the proposed schemes and illustrate the effect of weighting strategies proposed.

Arindam Chatterjee, Salil Joshii, Pushpak Bhattacharyya, Diptesh Kanojia and Akhlesh Meena [15] shows that in almost all disambiguation algorithms, the sense distribution parameter *P(S/W)*, where *P* is the probability of the sense of a word *W* being *S*, plays the deciding role. The widely reported accuracy figure of around 60% for all-words-domain-independent WSD is contributed to mainly by *P(S/W)*, as one ablation test after another re-veals.Their experience of working with hu-man annotators who mark with WordNet sense ids, general and domain specific corpora brings to light the interesting fact that producing sense ids without looking at the context is a heavy cognitive load.

ISSN 2249-6343
International Journal of Computer Technology and Electronics Engineering (IJCTEE)
Volume 2, Issue 2

Sense annotators do form hypothesis in their minds about the possible sense of a word, but then look at the context for clues to accept or reject the hypothesis. Such clues are minimal, just one or two words, but are critical nonetheless

III. APPROACHES OF WSD

A. Knowledge Based Approaches

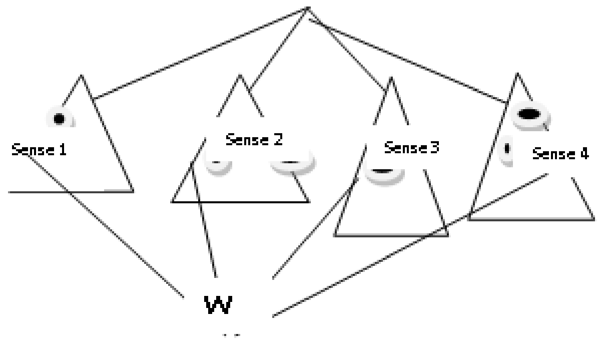
These approaches are mainly using external lexical resources such as dictionaries, thesaurus, WordNet etc. these are easy to implement because they require simple look up of a knowledge resources like a machine readable dictionary. Here no need of a corpus-tagged or untagged, since no training is involved. So many algorithms are suggested with these. Some of the algorithms are covering here.

B. Lesk Algorithm

This method is suggested by the scientist M.Lesk .According to him, a word is disambiguated by comparing the gloss of each of its senses to the glosses of every other word in the phrase. The sense whose gloss shares the largest number of words in common with the glosses of other words is selected as the correct sense. In 1990,wilki et.al. Proposed some modifications by calculating the frequency of co occurrence for the words in the definition text. From this, it is possible to derive several measures of the degree of relatedness among words. This metric is then used with the help of a vector method that relates each word and its context. By adding additional information fields such as Box codes, subject codes etc with dictionaries some of the authors attempted to improve the method in 1990's. For example, entry for bank in the Long man's Dictionary of Contemporary English (LDOCE) includes the subject code EC (Economics) for the financial senses of the bank. Given such subject codes, we can guess that expanded terms with the subject code EC will be related to this sense of the bank rather than any of the others.

C. Walker's algorithm

It is a thesaurus based approach. Here first finds the thesaurus category to which that sense belongs. Then calculate the score for each sense by using the context words. A context will add 1 to the score of the sense if the thesaurus category of the word matches that of the sense.
By using WordNet, it is possible to find the conceptual distance by analyzing the hyponyms. Once we find out the conceptual distance, conceptual density can be measured. If the conceptual distance is smaller, conceptual density will be higher. Let 'w' be the word to be disambiguated. w1, w2, w3 etc are the words in context. Following figure represents the concept of conceptual density. Each symbol represents the different senses of the word in context. Highest density will be obtained for the sub hierarchy containing more senses (Fig-II).



Word to be disambiguated: W
Context Words: W1, W2, W3, W4…….

Fig. II Different senses of word

D. Random Walk algorithm

In a sentence there may be more than one word which has different senses. In this approach, a vertex is created for each possible sense of each word in a text. By using definition base similarity, we can add weighted edges. A graph based ranking algorithm is then applied to find score of each vertex. Then the highest score vertex is selected as the correct sense (for each word).

E. Machine Learning Approaches

In machine learning approaches, systems are trained to perform the task of word sense disambiguation. In these approaches, what is learned is a classifier that can be used to assign as yet unseen examples to one of a fixed number of senses. These approaches vary as the nature of the training material, how much material is need, the degree of human intervention, the kind of linguistic knowledge used, and the output produced. But the system accuracy can definitely be improved by machine learning methods. These approaches can be mainly classified into two.

Supervised Learning

In such approaches, a learning system is presented with a training set consisting of feature encoded inputs along with their appropriate label, or category. The output of the system is a classifier system capable of assigning labels to new feature encoded inputs. Here a disambiguated corpus is available for training. There is a training set of exemplars where each occurrence of the ambiguous word 'w' is annotated with a semantic label. The task is to build a classifier which correctly classifies new cases based on their context of use. Two of the supervised algorithms applied to WSD in statistical language processing is:

Bayesian classification

This method is proposed by Gale et.al. It treats the context of occurrence as a bag of words without structure, but it integrates information from many words in the context window.

Information Theory

This one is proposed by Brownet et.al. This approach looks at only one informative feature in the context, which may be sensitive to text structure, but this feature is carefully selected from a large number of potential 'informants'.

Unsupervised Learning

In unsupervised learning we don't know the classification of the data in the training sample. It can often be viewed as a clustering task. Hyperlex and Lin's Approach are the main two algorithms used in these techniques several disambiguation systems have been developed for various languages like English, Tamil, Malayalam, Hindi, Chinese etc. Only some of the approaches to WSD were discussed here. So many approaches are implemented in both knowledge based and machine learning methods. Hybrid approaches by combining multiple knowledge sources and using tagged data are also one of the approaches to WSD.

IV. STATE-OF-THE-ART PERFORMANCE

We will briefly summarize the performance achieved by state-of-the-art WSD systems. First, homographs are often considered to be a solved problem. Accuracy above 95% is routinely achieved using very little input knowledge: for example, Yarowsky (1995) used a semi-supervised approach evaluated on 12 words (96.5%), and Stevenson and Wilks (2001) used part-of-speech data (and other knowledge sources) on all words using LDOCE (94.7%). Accurate WSD on general polysemy has been more difficult to achieve, but has improved over time. In 1997, Senseval-1 (Kilgarriff and Palmer 2000) found accuracy of 77% on the English lexical sample task, 5 just below the 80% level of human performance (estimated by inter-tagger agreement; however, human replicability was estimated at 95%; In 2001, scores at Senseval-2 (Edmonds and Cotton 2001) appeared to be lower, but the task was more difficult, as it was based on the finer grained senses of WordNet. The best accuracy on the English lexical sample task at Senseval-2 was 64% (to an inter-tagger agreement of 86%). Previous to Senseval-2, there was debate over whether a knowledge-based or machine learning approach was better, but Senseval-2 showed that supervised approaches had the best overall performance. However, the best unsupervised system on the English lexical sample task performed at 40%, well below the most frequent-sense baseline of 48%, but better than the random baseline of 16%.

By 2004, the top systems on the English lexical sample task at Senseval-3 (Mihalcea and Edmonds 2004) were performing at human levels according to inter-tagger agreement. The ten top systems, all supervised, made between 71.8% and 72.9% correct disambiguation compared to an inter-tagger agreement of 67%. The best unsupervised system overcame the most-frequent-sense baseline achieving 66% accuracy. The score on the all-words task was lower than for Senseval-2, probably because of a more difficult text. Senseval-3 also brought the complete domination of supervised approaches over pure knowledge-based approaches.

V. EVALUATION OF WORD SENSE DISAMBIGUATION SYSTEM

The metrics used to evaluate word sense disambiguation are precision and recall. Precision is defined as the proportion of correctly classified instances of those classified or it is the percentage of words that are tagged correctly, out of the words addressed by the system, while recall is the proportion of correctly classified instances of total instances or it is the percentage of words that are tagged correctly, out of all words in the test set.. Thus, the value of recall is always less than that of precision unless all instances are sense tagged.
Consider a test set of 100 words and suppose that 75 words are attempted by the system and out of these 50 words are correctly disambiguated. Then the precision and recall can be calculated as:
Precision=50/75
Recall=50/100
There are some gold standard test corpus are available like SEMCOR,SENSEVAL etc. Evaluation of the system actually includes the comparison with these standard test corpora.

VI. CONCLUSION

The main goal of this paper is to give an idea the people working in the field of natural language processing who want to learn about WSD. According to the extent to which major words in text are sense tagged, WSD tasks fall into two types: (1) tag all major words (nouns, verbs, adjectives and adverbs), and (2) tag some major words (usually nouns or verbs)
We can distinguish two variants of the generic WSD task:
—Lexical sample (or targeted WSD), 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.

ISSN 2249-6343
International Journal of Computer Technology and Electronics Engineering (IJCTEE)
Volume 2, Issue 2

On the other hand, other approaches, such as knowledge-lean systems, rely on full-coverage knowledge resources, whose availability must be assured. From the survey we can conclude that machine learning approaches are more capable for giving accurate senses. The reason is that knowledge based systems suffer from poor accuracies because of their complete dependence on dictionary defined senses, which don't provide enough clues. But the requirement of large corpus often renders learning algorithms unsuitable for resource poor languages, like Indian languages. We reviewed that lots of work has been done the area of word sense disambiguation, but still there are some area of word sense disambiguation which are to be explored to find most accurate sense of the word in relative context. We can apply some data mining tools to explore the new results. We propose a new method based on mining association rules, which can mine the association rules between the sense of the ambiguous word and its context, and the sense of the ambiguous word is determine by choosing the sense which is most association rule deduced. We can use the accuracy-coverage trade off to achieve high accuracy for WSD system. We can do this by disambiguating a word if the rule applicable to the word has high confidence.

REFERENCES

[1] Ramakrishnan and B. Prithviraj "Soft word sense disambiguation", 2004

[2] Andres Montoyo, Armando Suarez, German Rigau, Manuel Palomar "Combining Knowledge- and Corpus-based Word-Sense-Disambiguation Methods", Journal of Artificial Intelligence Research 23 (2005) 299-330.

[3] Chunhui Zhang, Yiming Zhou and Trevor Martin "Genetic Word Sense Disambiguation Algorithm" Second International Symposium on Intelligent Information Technology Application, IITA '08, Shanghai, 20-22 Dec. 2008, Vol 1, P 123 -127.

[4] Myunggwon Hwang Chang Choi Byungsu Youn Pankoo Kim "Word Sense Disambiguation Based on Relation Structure" International Conference on Advanced Language Processing and Web Information Technology, 2008, Dalian Liaoning, China. ALPIT '08, P 15 – 20.

[5] S.K.Jayanthi and S. Prema, Member, IACSIT "Word Sense Disambiguation in Web Content Mining Using Brill's Tagger Technique", International Journal of Computer and Electrical Engineering, Vol. 3, No. 3, June 2011.

[6] Jiang Tao, Tan Ah-Hwee, Senior Member, IEEE, and Wang Ke, "Mining Generalized Associations of Semantic Relations from Textual Web Content", IEEE Transactions on Knowledge and Data Engineering, VOL. 19, NO. 2, FEBRUARY 2007.

[7] Gaona M.A.R., Gelbukh A., Bandyopadhyay S., "Web-Based Variant of the Lesk Approach to Word Sense Disambiguation" Eighth Mexican International Conference on Artificial Intelligence, 2009. MICAI 2009, Guanajuato, P 103 - 107.

[8] Zhang Zheng, Zhu Shu "A New Approach to Word Sense Disambiguation in MT System" World Congress on Computer Science and Information Engineering, 2009, Los Angeles, CA, P 407 – 411.

[9] M.H.Margahny and A.A.Mitwaly "Fast Algorithm for Mining Association Rules" AIML 05 Conference, 19-21 December 2005, CICC, Cairo, Egypt.

[10] Xiao Guo, Dayou Li, Clapworthy, G." A Conditional mutual information based selectional Association and Word Sense Disambiguation" International Conference on Natural Language Processing and Knowledge Engineering, 2009. NLP-KE 2009, Dalian, P 1 – 7.

[11] Chatterjee N., Misra R. "Word-Sense Disambiguation using maximum entropy model" proceeding of International Conference on Methods and Models in Computer Science", 2009. ICM2CS 2009, Delhi, P1–4

[12] S. G. Kolte and S. G. Bhirud "WordNet: A Knowledge Source for Word Sense Disambiguation" International Journal of Recent Trends in Engineering, Vol 2, No. 4, November 2009.

[13] Bridget T. McInnes, Ted Pedersen, Ying Liu, Serguei Pakhomov, and Genevieve B. Melto "Using Second-order Vectors in a Knowledge-based Method for Acronym Disambiguation" Proceedings of the Fifteenth Conference on Computational Natural Language Learning (CoNLL 2011), June 23-24, 2011, pp. 145 - 153, Portland, Oregon.

[14] M. Nameh, S.M. Fakhrahmad, M. Zolghadri Jahromi," A New Approach to Word Sense Disambiguation Based on Context Similarity", Proceedings of the World Congress on Engineering 2011 Vol I, July 6 - 8, 2011

[15] Arindam Chatterjee, Salil Joshii, Pushpak Bhattacharyya, Diptesh Kanojia and Akhlesh Meena, "A Study of the Sense Annotation Process: Man v/s Machine", International Conference on Global Wordnets , Matsue, Japan,, Jan, 2012.

Samit Kumar, completed B.E. in Computer Science & Engineering in 2003 and M.Tech. in CSE (Object Oriented Database System) in 2007. Pursuing Ph.D. in Computer Science & Engineering. Have total 9 years of teaching experience.Presented 16 National Papers and 6 International Papers in IEEE sponsored conferences. Published one research paper in reputed International Journal.

Neetu Sharma, completed BE (CSE) in 1996. M.Tech (CSE) in 2010 from Panjab University. Presently doing Ph. D. in Computer Science and Engg.. Having total teaching experience of 15 years. Presented two papers in National and three papers in International conferences sponsored by IEEE**.**

S. Niranjan, did his M.Tech(Computer Engg.) from IIT Kharagpur in 1987. Completed Ph.D.(CSE) in 2004 and Ph.D.(I&CT) in 2007, Have total 26 years of teaching experience. Presently working as Principal in PDM College of Engg. Bahadurgarh(Haryana). Having 6 publications in journals and 30 other papers published in conferences.