A Review Of Literature On Word Sense Disambiguation

Ravi Mante[#1], Mahesh Kshirsagar[*2], Dr. Prashant Chatur[#3]

[#1]*Assistant professor,* [*2]*MTech (CSE),* [#3]*Head Of Department*
Department of computer science and engineering
Government college of engineering Amravati, Maharashtra

Abstract— **lexical ambiguity is a fundamental characteristic of language. Words can have more than one distinct meaning. Word sense disambiguation is defined as the problem of computationally determining which"sense"of a word is correct in given context. Word sense disambiguation is a task of classification where 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. There are various approaches for word sense disambiguation like knowledge based method, supervised method, unsupervised method and semi supervised method. There are various application of word sense disambiguation like machine translation, information extraction, information retrieval**

Keywords— **ambiguity, word sense, word sense disambiguation.**

I. INTRODUCTION

Word sense is a definition or meaning of a word. In any natural language each word may have more than one meaning that is each word can have many senses. Disambiguation means to remove all ambiguity. Word sense disambiguation is to find out which dictionary definition is correct in given context. For example suppose in GPS application if user asks that he wants to go the bank. So which location should GPS give? The noun bank has several meaning, it could be a financial institution to deposit a check or it could be a supply or a stock held in reserve for future use. In this what should be the correct sense it needs to understand how words relate with specific context. Humans can understand the context of word by simply looking at the surrounding words and comparing words in different ways. Same technique can be used by machines to determine the correct sense of the word in given context.

II. LITERATURE REVIEW

Word sense disambiguation 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 computer linguistic.

Weaver and Warren [2] acknowledged that context is crucial, and recognized that "statistical semantic studies should be undertaken, as a necessary primary step.

Zipf, George Kingsley [3] published his "Law of Meaning" that accounts for skewed distribution of words by number of senses, which is more frequent words have more senses than less frequent words in a power law relationship.

Kaplan, Abraham [4] determined that two words of context on either side of an ambiguous word was equivalent to a whole sentence of context in resolving power.

Madhu and Lytle [5] calculated sense frequencies of words in different domain observing early on that domain constrains and then applying Bayes formula to choose the most probable sense in given context.

Bar –Hillel [6] argued that "no existing or imaginable program will enable an electronic computer to determine that the pen is used in its 'enclosure' sense in following sentence: little john was looking for his toy box. Finally he found it .The box was in the pen. John was very happy.

Wilks [10] developed "preference semantics", one of the first systems to explicitly account for word sense disambiguation. The system used selection restrictions and a frame based lexical semantics to find a consistent set of word senses for the words in a sentence.

The 1980s were a turning point for word sense disambiguation .Large scale lexical resources and corpora became available so handcrafting could be replaced with knowledge extracted automatically from the resources. Lesks [8] used the overlap of word sense definitions in the Oxford Advanced Learners Dictionary Of Current English to resolve word senses. Given two target words in a sentence, the pair of senses whose definitions have the greatest lexical overlap are chosen.

In 1990 three major development happened Word Net Became available, statistical revolution in NLP swept through, and sensval began.. WorldNet pushed result forward because it was both computationally accessible and hierarchically organized into word senses called as synset.Before sensval it was extremely difficult to compare and evaluate different systems because of disparities in test words, annotators,sense inventories and corpora.Sensval was first discussed in 1997 and now hosting three evaluation exercises has grown into the primary forum for researchers to discuss and advance in the field.

III. APPROACHES OF WORD SENSE DISAMBIGUATION

Various approaches to WSD are often classified according to source of knowledge used in sense differentiation following are the techniques for word sense disambiguation

A. Knowledge Based Method[9]

It represents the distinct category in word sense disambiguation. These methods use lexical knowledge bases such as dictionary and thesauri and extract knowledge from word definitions and relation among word and senses [11]. Four main types of knowledge based method [12] are as follows

1) *The Lesk Algorithm* [8]: in this it computes the overlaps between words that are the number of words in common between the definitions of senses. For example consider the task of disambiguating the words pine and cone. The oxford Advanced Learners Dictionary defines four senses for pine and three senses for cone which are as follows
Pine:-
a) Seven kinds of evergreen tree with needle-shaped leaves
b) Pine
c) Waste away through sorrow or illness
d) Pine for something, pine to do something

Cone:-
a) solid body which narrows to a point
b) something of this shape, whether solid or hollows
c) fruit of certain evergreen trees(fir, pine)

Here the third definition of cone and first definition of pine have the largest overlap among all sense combination with three words *evergreen, tree and pine* in common.

2) *Measure of semantic similarity computed over semantic network [7,16]:* this include the methods for finding the semantic distance between concepts. Depending on size of context they are divided into two categories one is method applicable to a local context and second method applicable to global context. Semantic similarity is the most powerful constraint used in automatic word sense disambiguation. So by finding the smallest semantic distance we can find the appropriate senses of the word in given context.

3) *Heuristic method:* it consist of simple rules that can reliably assign sense to certain word categories which includes Most frequent sense, one sense per collocation and One sense per discourse

   a) *Most Frequent Sense:* in general it is found that one meaning of word occurs more often than other meaning. so by using the very simple disambiguation method with the help of word frequency data we can assign to each word its most frequent meaning[11].

   b) *One Sense Per Discourse:* according to Gale word tends to preserve its meaning across all its occurrences in a given discourse. If the meaning is identified in at least one such occurrence then it allows for the automatic disambiguation of all instances [13]

   c) *One Sense Per Collocation:* it states that a word tends to preserve its meaning when used in same collocation. Nearby words provide strong and consistent clues to the sense of a target word.

4) *Automatically or semi automatically acquired: -* [14]selectional preferences as a means of constraining the possible meaning of word.

B. Unsupervised corpus based method

One common sense of "unsupervised" literary means "not supervised" and which includes any method that does not use supervised learning. In this approach it bootstrap from small number of sense tagged training example and use that to build a simple model which tags a few more context. This method does not rely on external knowledge source such as machine readable dictionaries, concept hierarchies, or sense tagged text. They discriminate among word meaning based on information found in unannotated corpora and do not assign sense tags to words.

C. Supervised corpus based method:

It consist of training phase and testing phase .the training phase requires a sense annotated training corpus from which syntactic and semantic features are extracted to build a classifier .and in testing phase classifier picks the best sense of a word on the basis of surrounding words. Regarding automatic word sense disambiguation one of the most successful approaches in recent years is supervised learning from example in which classification models are induced from semantically annotated corpora. The supervised approach to word sense disambiguation uses semantically annotated corpora to train machine learning algorithm to decide which sense to choose in which context.

D. Semi supervised method:

An important issue of supervised method is the knowledge acquisition bottleneck. It is difficult to find the required minimum number of occurrences per each sense of word. To overcome the knowledge acquisition bottleneck faced by supervised method semi supervised method uses small annotated corpus.

IV. APPLICATION OF WORD SENSE DISAMBIGUATION

There are various application of word sense disambiguation some of which are as follows

1) *Machine translation:* Word Sense Disambiguation is required in machine translation for words that have different translations for different senses. For example whenever we translate any sentence from English to Hindi language it should give correct meaning. For example consider a sentence curiosity kills the cat. Here the meaning of this sentence if translated in Hindi gives different meaning because here the words have different senses in different context. If we take literal meaning of the word from one language then it translates it to another language then sometimes the translated sentence

does not give the same meaning as of original language.

*2) Information Extraction'*s is required for the accurate analysis of text in many applications. Word Sense Disambiguation is also used in text mining.

3) Information Retrieval: ambiguity must be resolved in some queries for example the noun "cricket" then what information should system retrieve because cricket is a insect as well as it is a sport so WSD plays very important role in information retrieval. Most of search engines do not use explicit semantics to prune out documents which are not related to user query. An accurate disambiguation of document base with possible disambiguation of the query words would allow it to eliminate document containing the same words used in different meaning and to retrieve document expressing the same meaning with different wordings [15]

a) Cross language information retrieval: in cross language information retrieval the user presents a query of the usual form but some of the document s may be written in a different language. For example user types query in English language and wants document in German language, here it needs to convert that English query in German language and then retrieve the document. Here the problem of ambiguity comes which is why word sense disambiguation is use.

b) Question Answering: Question Answering is the oldest natural language processing application. In this system specific questions are asked for example "when the computers are invented?" and it receives the concise answer rather than a set of relevant document. To retrieve concise pages we need to have correct sense of a particular word so word sense disambiguation plays a very important role.

V. CONCLUSIONS

In this paper we have studied what is word sense disambiguation and brief history about it and also various approaches for word sense disambiguation along with applications. Word sense disambiguation is very important in the field of natural language processing and also machine translation. Because every word is having different senses in different context. In order to get the correct meaning of the given sentence we should eliminate the ambiguity that is to choose the correct sense in given context hence word sense disambiguation plays very important role.

REFERENCES

[1] S. M. Metev and V. P. Veiko, *Laser Assisted Microtechnology*, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.

[2] Weaver ,Warren Translation.Mimographes,12 pp.Reprinted in William N.Locke &Donald A.Booth,eds.1955.Machine Translation of Languages,15-23.New York:John Wiley and sons.

[3] Zipf,George Kingslay "Human Behaviour and the principal of least effort: An introduction to human ecology".Cambridge,MA:Addison-Wesley.Reprinted by New York:Hafner,1972

[4] Kaplan, Abraham,"An experimental study of ambiguity and context. Mimographed,18pp,November 1950

[5] Madhu, Swaminathan and Dean W. Lytle."A figure of merit techniques for the resolution of non-grammatical ambiguity. Mechanical Translation,8(2):9-13,1965

[6] Bar-Hilel.The present status of automatic translation of languages.Advance in Computers,ed.by Franz Alt et al.91-163.New York:Academic press,1960

[7] Rada,Roy,Hafedh Mili,Ellen Bicknell & Maria Blettner"Development and application of a metric on semantic nets.IEEE Transactions on system,Man,and Cybernetics,19910:17-30,1989

[8] Lesk, Michael.Automated sense disambiguation using machine readable dictionaries: How to tell a pine cone from an ice cream cone. Proceedings of the 1986 ACM SIGDOC Conference, Toranto, Canada, 24-26.

[9] E. Agirre and P. Edmonds(eds), Word Sense Disambiguation: Algorithm and Applications,107-131.2007 Springer

[10] Wilks, Yorick, Preference semantics. Formal semantic of natural language,ed.by E.L.Keenam,III,329-348.Cambridge,UK:Cambridge University press,1975

[11] Gale, William,Ken Church & David Yarowsky"Estimating upper and lower bounds on the performance of word sense disambiguation programs. Proceedings of the annual meeting of the association for Computational Linguistics(ACL),Newark,U.S.A.,249-256,1992

[12] Gale, William,Ken Church & David Yarowsky"one sense per discourse. Proceedings of the DARPA Speech and Natural Language Workshop, New York,U.S.A,233-237,1992

[13] Yarowsky,David "One sense per collocation. Proceedings of the ARPA Human Language Technology Workshop,Plainsboro,U.S.A.,1993

[14] Brockmann,Carsten and Mirella Lapata "Evaluating and combining approaches to selectional preferences acquisition.Proceedings of the European Association for Computational Linguistics(EACL),Budapest,Hungary,27-34,2003

[15] ACM Computing Surveys, Vol. 41, No. 2, Article 10, Publication date: February 2009.

[16] Ankush Maind, A.V.Deorankar, Dr.P.N.Chatur "Web based measure of semantic similarity between words" IJCST vol 3, issue 4,oct-dec 2012