International Journal of Advanced Engineering: Artificial Intelligence, Robotics and Automation
ISSN XXXX-XXXX
Vol.:1 Issue: 1
Published online on: Dec 2024
PP. 26- 33

Word-Sense Disambiguation: an Aid to true translation between Natural Languages

Pratosh Bansal[1], Jayshri Bansal[2], Premanand Singh Chauhan[3], Rajesh Kumar Chakrawarti[4]
[1]Institute of Engineering &Technology, DAVV, Indore ,India
2Human Resource Development Centre,DAVV, Indore, India
3Vikrant University, Gwalior, India
[4]Sushila Devi Bansal College of Technology, Indore , India

***Abstract* —** Word-sense disambiguation (WSD) is a long-standing research area of Natural Language Processing (NLP) applications. This is the core of many NLP Systems and has become an important and very successful technique for the last 03 decades. By definition, WSD is a technic for seeking the right sense (meaning) of a word appearing in a sentence with a certain context. A single word may be *Polysemy* (i.e. multiple meanings) and even multiple words may represent *Synonymy* (i.e. similar meaning) by its nature. The *"Corpus/lexicon"* is a special kind of dictionary or knowledge base which is very helpful in Natural Language Translations. It is also called the heart of NLP. WordNet is the most important example of it. A WordNet is a lexical English database, available online and designed to use with the support of a control program. In this paper, a framework for WSD and its approaches has been provided through which the Natural Language Translation can be done. Further, the detailed knowledge of the artifacts of WSD and also the brief of various applications of WSD has been discussed.

Keywords — Natural Language Processing(NLP), Word-Sense Disambiguation(WSD), Machine Translation(MT), WordNet, Universal Networking Language(UNL), Ontology.

I. Introduction

India is a multilingual country and statistics prove that more than 20 languages are spoken by the various communities of it. These can include Hindi, Marathi, English, Telugu, Bengali etc. Hence Machine Translations and cross-lingual searches for these languages in the NLP systems a crucial task. In the field of computational linguistics (CL), this problem is widely recognized as Word-Sense Disambiguation (WSD), and computationally it can be defined as the problem of identifying the correct meaning (sense) of a particular word activated, during the use of that word in a specific context. These require large resources of high quality and coverage. Word-Sense Disambiguation approaches have shown high quality and efficient outputs in these scenarios along with the lexicons/corpora (e.g. WordNet) [1][2].

WSD can be considered as a task of classifications:

(i) Classes - *Word senses (meanings)*;

(ii) Evidence – Provided by the *Context*;

(iii) Occurrence - based on the evidence (called context), each word is allotted for more than one of its possible classes (called senses).

In simple words, we can say that the WSD is basically a relationship between "word", "context" and "sense" (or "name", "place" and "meaning").

A WordNet (or simply a lexicon/corpus) is an online lexical English database designed to be used with the support of a control program. The sets of synonyms have been prepared based on (i) nouns, (ii) verbs, (iii) adjectives and (iv) adverbs of English languages each representing a lexicalized concept which are linked through Semantic relations.

Machine Translation (MT) is a broad area that comes under NLP Systems. MT is a special application into the computers for translation of given texts from one Natural Language to another Natural Language. This can be broadly broken into two stages [3][4]:

(1) The source language – Understanding the sentence,

(2) The target language - Generating the sentence.

Corresponding Author's E-mail ID: *rajesh_kr_chakra@yahoo.com*

At both stages, WSD is required, because of a word available in the source language (SL) might produce two or more possible senses (meanings) for the asked target language (TL). E.g. the Hindi language word "□□□□" can be translated into English language either as "sleep" for its one sense in the context of daily routine or as "gold" for its other sense in the context of metal. Hence, to translate correctly any sentence, we always need to know what sense is essential in the given statement [5][6].

Universal Networking Language (UNL) is an important type of computer language which is responsible for enabling computer systems to process and disseminate knowledge across the barriers between the languages. By providing the Linguistic Infrastructure (LI) for information distribution, receiving, and understanding the concept of multilingualism, UNL enables the computer to inter-communicate for a source language with the target language. With the help the UNL System, multilingualism information might be accessed through natural languages.

Basically, The UNL uses the WSD approaches on both languages (i.e. the source and target) with the help of a lexicon and performs the efficient MT for that source language by providing the high quality and best covered output in the target language.

II. Related Works

A lot of researches have already been done for Machine Translations of Natural Languages with the help of Word Sense Disambiguation along with their applications. In continuation, researches are still in progress towards success and completion of important applications related to NLPs. In the next paragraph, brief literatures have been provided.

WordNet is a very long and large sized English language lexical database in which the (i) nouns, (ii) verbs, (iii) adjectives and (iv) adverbs of the English language are grouped into sets. These sets are called the cognitive synonyms (synsets) and each and every one is expresses a different concept. For download and use, WordNet is available freely and publicly. Hindi Wordnet is totally based on the idea of English WordNet. In other way, we can consider this as more than a conventional Hindi dictionary. It provides various links between the synsets (or synonym sets) which represents unique concept [7][8].

The authors and other researcher have provided a large numbers of Algorithms and Applications details for the WSD given by different peoples throughout the word. Few special articles on WordNet give knowledge for the construction of WordNet. Some researchers are helpful in building of various types WordNets for Indian languages as well as foreign languages [5][9]-12].

The authors have provided knowledge of consideration of the Hindi Wordnet as a fundamentally important task of NLP like, Hindi words disambiguation. Research also shows that, It have been first time attempted to use of WSD automatically for the Indian languages and plays a noteworthy role towards processing of Indian languages. A research is also available in which investigation for all the critical aspects related to various WSD approaches have been done. In this, success rates of various approaches and its usage in various applications related details have also been provided. *A* Tutorial on ACL is available in which the authors have tried to give the presentation for the WSD and its techniques [13]-[16].

A number of survey reports and researches are available for the Machine Translation systems developed in India as well as in abroad. Importance of the Poetry Translations has been discussed. Idea for developing the Idiom Translation system has also being provided [17-19].

III. WSD: Word-Sense Disambiguation

In the field of computational linguistics, Word-Sense Disambiguation is a central research problem. By consideration of the scope of Artificial Intelligence (AI), and at the beginning of the scientific interest in Machine Translation and Natural language processing, it was recognized. Basically, WSD is a technic for seeking the correct sense (meaning) to asked work appearing in a sentence with a specific context. WSD can also be considered as the task of classifications where Classes represents Word senses (meanings), Evidences(called context) are provided through the Context and Each occurrence of a word is allotted for more than one of its possible classes (called senses).

Occurrence - based on the evidences (called context), each word is allotted for more than one of its possible classes (called senses).

The central idea and heart of the problem for WSD is Word meaning. In the computational linguistics, the significance of WSD has been extensively accepted and acknowledged. The main idea behind the concept of using WSD is, for comparing the senses of a particular word with the context for the sentence along the contexts received through WordNet and then choosing the appropriate winner. This would be a right synset number generated by the WordNet displaying the correct sense of the word.

Some words may have two (02) or more possible sense (meaning). E.g., consider the given sentence-

Corresponding Author's E-mail ID: *rajesh_kr_chakra@yahoo.com*

पृथ्वी सोना चाहता है|

It can be translated in the following two ways as follows-

Prithvi wants to sleep.

Or

Prithvi wants gold.

Therefore, there is ambiguity for "□□□□" because it is being interpreted as "sleep" for its sense in context of daily routine or as "gold" for its sense in context of metal. When we look up in any dictionary for getting the meaning for a particular word, it can be seen that a word might have multiple meanings and of course some of them may be totally different from each other. Even some words may not be as easy to disambiguated because of they may have multiple senses which might be very-very close to each other. In some situations, disambiguation might not be possible altogether using only the given one sentence. On that situation, these scenarios may demand for look at the context at higher level (called the discourse) to disambiguate among the possible senses.

WSD is related to various computational areas of NLP and Artificial Intelligence (AI) which may encompasses: Machine Translation (MT), Speech Processing (SP), Information Retrieval (IR) and Hypertext Navigation is also associated with this, Cross Language Information Retrieval (CLIR), Knowledge Acquisition (KA), Question Answering (QA), Information Extraction (IE), Grammatical Analysis (GA), Content and Thematic Analysis (CTA), Statistical Machine Translation (SMT).

IV. Framework of WSD

A framework or Model using WSD for A Machine Translation from one Language to another language may include the five important artifacts under the task of NLP for the purpose of the efficient translation. The names of those artifacts are as follows:

(i) The Source Language (e.g. Hindi)
(ii) The Target Language (e.g. English)
(iii) Universal Networking Language (UNL)
(iv) A Corpus/Lexicon (e.g. WordNet)
(v) WSD Approaches

The framework or Model for the WSD may be look like Fig.1.

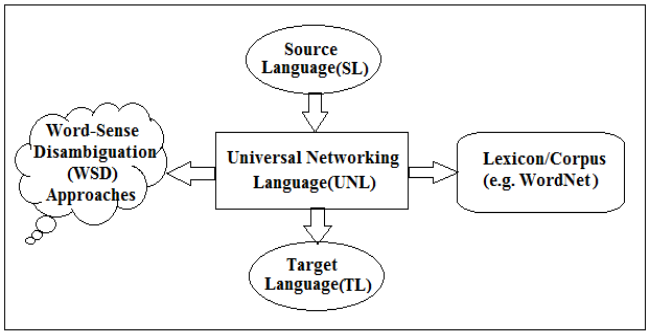

Fig. 1. WSD Framework.

Out of the above five, we are describing the first three artifacts at here and the remaining Two important will be described on later points. Here the Source language is that language for which we have to apply the translation e.g. we have taken Hindi as an example of source language. The target language is that in which we want the output e.g. we have taken English as example of target language.

Universal Networking Language (UNL) is an important type of computer language which is responsible for enabling the computer systems to process and disseminate the knowledge across the barriers between the languages. With providing the Linguistic Infrastructure (LI) for information distribution, receiving, and understanding concept of multilingualism, UNL enables the computer to inter-communicate for a source language with the target language. With the help the UNL System, multilingualism information might be accessed through natural languages. In our Example the UNL is English language again.

According to WSD Framework, the Source Language will be converted in to the Target Language by the UNL by using an appropriate approach of WSD and gaining the information from the Lexicon/Corpus (WordNet). How the WSD approaches works and how the information is organized on the WordNet, we will be discussing

Corresponding Author's E-mail ID: *rajesh_kr_chakra@yahoo.com*

on the proceeding points. At both stages, WSD is required, because of a word available in the source language (SL) might produce two or more possible sense (meaning) for the asked target language (TL).

V. WSD Approaches

Basically, WSD is a technic of finding the correct sense (meaning) for a particular word which may have multiple senses and can be used for a given sentence. WSD approaches classification might be done based on the use of main knowledge source for the sense differentiation for a particular word. Classification of WSD algorithms is also a very critical task. Although, various resources and different techniques are used by WSD algorithms, yet we have classified them according to the utilization of the resources and use of the processing techniques. Almost, these all approaches, works on the principle of assigning a *window* for content words (n) and around the word to be preserved for disambiguation on the corpus, and then by analyzing statistically those surrounding words (n). Broadly, there are four (04) conventional approaches to WSD [14][15]:

[1] *Knowledge and Dictionary-based methods for WSD*

In Knowledge–WSD, the knowledge is drawn from dictionaries and/or raw text. In other words, an external information source is used for the context of the words with information, without using any corpus evidence. The external sources include:

(i) Thesaurus
(ii) Lexical Knowledge bases (Lexicons)
(iii) Machine Readable Dictionaries (MRD's)

The main application areas are (a) Machine Translation, (b) Statistical Machine Translation, (c) Speech Processing, (d) Text Processing, and (e) Question Answering etc. are majorly based on Knowledge based approaches.

[2] *Supervised Learning Methods for WSD*

Supervised Learning methods are based on such kind of assumption at where enough evidence may be provided by the context its own for the words disambiguating (for this purposes the, world knowledge and their reasoning is reasoning are deemed unnecessary). Sense-annotated corpus is used to train from. With the use of these methods, a sample of training data is being created at where a target word may be manually annotated for defining the specific sense from a pre-determined set of available possibilities. In other words, methods works on labeled training set. The sense model is trained by the sense-tagged corpora, which produces the possibility of the contextual features to the word sense. A manually sense-tagged text included by the classifier along with machine learning techniques. These approaches are responsible for reducing WSD into classifier problems where the most appropriate sense is assigned to the target word, based on the particular context in which it occurs from pre-determined set of available possibilities. These methods use the following kinds of the classifiers:

(i) Naïve Bayesian,
(ii) Exemplar Based,
(iii) Lazy Boosting Algorithm,
(iv) Decision Lists.

The main application areas are (a) Information Retrieval, (b) Cross Language Information Retrieval, and (c) Question Answering etc. at where supervised learning methods may be applied.

[3] *Unsupervised Learning methods for WSD*

These methods discard any king of external source information completely and directly work on the raw-annotated corpora. These methods are also called as Word-Sense Discrimination methods. WSD researchers have found the greatest challenge during the use of these methods. These methods identify the patterns without the usages and benefits of any manually labeled and externally knowledge information in a large sample of data. Further, these patterns are further used for dividing the data by forming the clusters. In a cluster, each member is generally more common to the other members of own cluster rather than the clusters of others. As these approaches don't require any training corpus hence requires less computing power and time. These methods are very-very suitable for applications related to machine translation systems and information retrieval systems. It can also be believed that the unsupervised learning will be responsible for overcoming the knowledge acquisition issues as they don't dependent on manual effort. Unsupervised Learning methods are classified as follows:

(i) LSA - Latent Semantic Analysis,
(ii) Parallel Text,
(iii) Spreading Activation Networks,

Corresponding Author's E-mail ID: *rajesh_kr_chakra@yahoo.com*

The main application areas are (a) Information Retrieval, (c) Classification of Documents, (c) Synonymy Text, (d) Simulating and Word Sorting etc. at where unsupervised learning methods may be applied.

[4] *Semi/Minimally-Supervised Learning methods for WSD*

Semi/Minimally-Supervised methods use learning sense classifier, with minimal human supervision, from annotated data. A secondary external source of knowledge is used for support like small/thin annotated corpus which is basically a seed data responsible for bootstrapping process, or it can be a word-aligned bilingual corpus. Many WSD algorithms use these methods because availability of training data is very less in number, which allow labeled and unlabeled data both. Semi/Minimally-Supervised methods are classified into as follows:

(i) Bootstrapping Approach
(ii) Algorithm given by Yarowsky
(iii) Algorithm based on Self-Training
(iv) Algorithm based on Co-Training

The main application areas of Semi/Minimally-Supervised Learning methods are (a) Information Retrieval and (b) Cross Language Information Retrieval etc. at where these may be applied.

VI. WordNet As Corpus/Lexicon and its semantic Relations

Any System, which has been developed to process the natural languages, must include the enough information related to the words and their senses (meanings) definitely, because of the meaningful sentences are basically mixtures of meaningful words. Although, traditionally, this information can be collected through the dictionaries directly but entries available in the dictionaries have been made for the suitability of human and not for machines.

WordNet is an effective and systematic collection of mostly all found existing lexicographic information for traditional computing to the modern era computing. A WordNet is an online lexical database for English language specially designed to use with program under control. Psycholinguistic theory of human lexical memory inspired in this designing. Every word sense is specified with set of word-forms (also called simply synset or synonym). Basically, synsets have been created for content words of English language as (i) noun, (ii) verb, (iii) adjective & (iv) adverb. Each denotes a lexicalized concept. Synonym sets are linked through Semantic relations.

We define, a Lexical matrix, for a set of Words of Pair*(f, s)* for the description of a language, where *f* is a Word-form as a heading for the column and *s* is a Word-sense (meaning) as a heading for rows and it represents an element taken from available set of senses (meanings). An entry shown under the cell of the Lexical matrix specifies that the word-form in corresponding to the column can be used suitably expressing the meaning of respective row. The word-form is called as *polysemous* if there are more than two entries (i.e. two meanings) in the same column. Similarly, the two word-forms are called as *synonymous* (according to the context) if there are more than two entries for the same row. This inferences that same word forms might have different meanings and same meanings can be denoted by different word forms [9][11][12]. The Table I: depicts the Lexical Matrix for the word **"□□□□"**.

TABLE I: THE LEXICAL MATRIX FOR THE WORD "□□□□".

Word Senses (Meanings s) ↓	Word Forms(f) → f1	f2	f3	---	fn
S1	स्वर्ण	सोना	कंचन		
S2		सोना		शयन	
S3		सोना			सोनापाठा

Sn					

Basically, A WordNet is a network words which is having collection of the senses of the words. In this network, a node (containing word sense) represents a synset in reality which might be considered as the role object under the WordNet. Further, every word sense node (synset) within the WordNet is relatively linked to other synsets through the well-known lexical relations and the semantic relations. Semantic relations are made for linking of the synsets whether lexical relations are made for linking of the words.

Multiple and differentiable semantic relations might be constructed among the words along with their senses than which have already been specified into WordNet. The semantic relations in WordNet were chosen in such

Corresponding Author's E-mail ID: *rajesh_kr_chakra@yahoo.com*

way because fundamentally they apply English knowledge throughout and familiarity for that. Hence, users don't needed to have advanced training of linguistics to understand all concepts related to linkages. Broadly, The following semantic relations have been included in the WordNet[8][10]:

[1] *Synonymy vs. Antonymy (similar and opposite):* Synonymy is the main and fundamental relation of the WordNet. A set of synonyms (synsets) are used to represent senses (meaning) of a word in the WordNet. Symmetric relations among the word and forms are represented by Synonymy. On the otherside, the opposite meanings between the two words is expressed by, Antonymy.

[2] *Hyponymy vs. Hypernymy ('is a kind of' Relation):* The semantic relation used for capturing the sub-set hood between two synsets called Hyponymy. It is transitive as well as asymmetrical. On the other side, a semantic relation for capturing the super-set hood between two synsets called Hypernymy. It is the reverse of hyponymy.

[3] *Meronymy vs. Holonymy ('Part-whole' Relation):* These two are also important and special kinds of the semantic relation amongst the synsets. In case of *Meronymy*, if the two concepts like A and B are fundamentally related in a way that A is one of the constituent of B, then A is called as the meronym of B. On the opposite side B is called as the holonym of A as it is the case of Holonymy. The meronymy relation is transitive as well as asymmetrical also and Holonymy is the reverse of it. Holonymy is mainly used for constructing a part-of-hierarchy.

[4] *Troponymy (manner):* The elaboration for one verb along to another verb with a specific manner is known as Troponym. This represents an action for the manner example A is a Troponym of B if A is to B with some manner anyhow.

[5] *Entailment:* A relationship between two verbs is represented by the Entailment. If fact of a verb A is sensibly followed by the fact of verb B, It means the verb A entails B. It is one way relation hence sometimes also said that entailment is unilateral.

[6] *Gradation*: It is an important lexical relation. It is useful to represent the intermediary concepts between the two concepts available with opposite relations.

VII. Ontologies of WordNet

'Ontology' is an engineering artifact which has been developed by a specific vocabulary used for describing certain realities and a set of clear assumptions regarding the intended sense (meaning) of the vocabulary. Ontology describes a formal specifications to certain domain for understanding on sharing basis, formal and machine manipulate model of particular interest of domain. In computer science, ontology is a proper and formal arrangement of knowledge in terms of set of various concepts of prescribed for domain specific and through their relationships for those concepts. It can be used as to provide the reason for availabilities of the entities in specified domain and might also be used for describing the domain. In other words, ontology provides shared information, which can be further used to model specific domain like the object type, concepts/properties that exist and their relations. Hence the ontology helps to describe Grammatical, Morphological, Syntactic, Semantic and other specific features of a Word.

The ontology, of a WordNet system, is a hierarchical concepts arrangement, entities categorization and actions. A separate ontological hierarchy is available for each and every Part-of-speech (POS) as (i) noun, (ii) verb, (iii) adjective and (iv) adverb. Each synset (set containing the words having similar meaning) is mapped onto the handsome place under the ontology. A single synset might have multiple ancestors. Fig. 2 shows an ontology node for the Part-of-speech (POS)[7][8].

Fig. 2. Ontology specifying Parts-of-speech (POS).

Corresponding Author's E-mail ID: *rajesh_kr_chakra@yahoo.com*

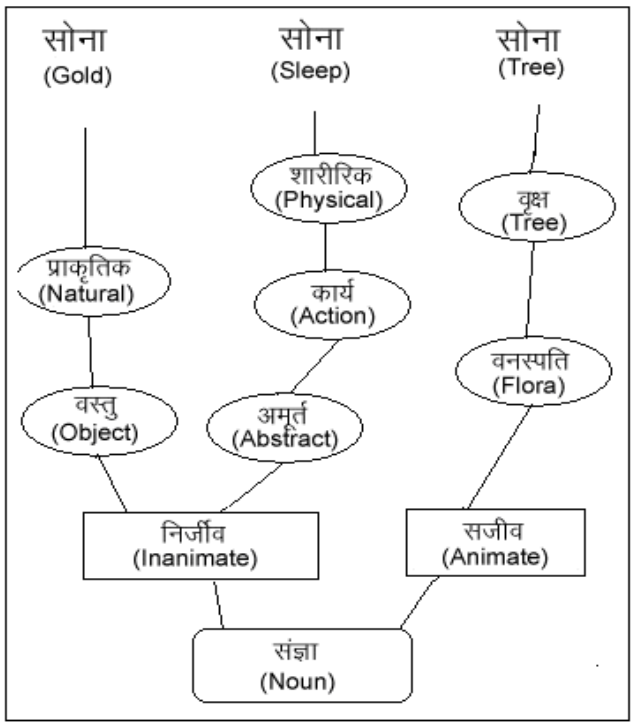

Fig. 3. Hierarchy of a word "□□□□" under the Ontology.

And the Fig. 3 shows the example of the individual detailed hierarchy for Parts-of-speech (POS) namely Noun word "□□□□"[11]. According to this diagram, the things become very clear i.e. the existence of the word "□□□□". This diagram comprises all the necessary belonging to the word "□□□□"[7][8].

According to the above figure, it is clear that if we want translate a sentence written in one language e.g. Hindi to another language e.g. English, the WordNet may play an important role in this conversion. Because the WordNet properly describes each and every construct of sentence by the effective mean. And with the help of the WSD approaches we can complete our conversion properly.

VIII. Result And Discussion

In this paper, the artifacts of the Natural Language Translation for one language to another language have been described. For this purpose, an example of Hindi to English translation has been taken in the consideration. The knowledge from the two important Corpuses; One is English i.e. WordNet developed at MIT press, Cambridge MA, USA and another is Hindi i.e. Hindi WordNet developed at IIT Bombay, India, have been taken. The concept of the Universal Networking Language (UNL) has been discussed in brief for defining the steps of the Translation i.e. how the translation will be taking place.

One of the most important things is that the idea of Word-Sense Disambiguation (WSD) approaches have been discussed in detail for the successful and the efficient translation; because without the use of any process/algorithm a translation is incomplete. With the use of all these approaches along with the all artifacts, the Word-Sense Disambiguation (WSD) may be the Best approaches for the successful, efficient, high-quality and true translation for a Natural Language Translation or Machine Translation (MT) for the Natural Language Processing (NLP) with the proper use of the Corpuses or Lexicons[16].

IX. Future Scope

In the present work, the detail knowledge of the Word-Sense Disambiguation (WSD), its utility and importance for the Natural Language Translation/ Machine Translation (MT) have been provided. But this is theoretical aspect only. Although the knowledge for the one kind of the translation i.e. Hind to English only has been discussed. But we are working for the implementation in an important domain called *poetry* translations from Hindi to English through the computer systems. For this purpose, the data sets have been prepared well and tools for implementation have been decided. In near future, the true translation for poetries might be provided and WSD approaches might be an effective input in this regard [16][17][19].

Corresponding Author's E-mail ID: *rajesh_kr_chakra@yahoo.com*

References

[1] R. Tandon, "Wordnet Sense Disambiguation using Hindi Wordnet", p. 1-2, February 18, 2009. [Online]. Available: Department of Computer Science & Engineering, IIT Kanpur,https://www.cse.iitk.ac.in/users/cs365/2009/proj/RashishTandon.pdf.

[2] R. Durgesh, "Machine Translation in India: A Brief Survey", National Centre for Software Technology (NCST), Mumbai, India.

[3] R. K. Chakrawarti, H. Mishra and P. Bansal, "Review of Machine Translation Techniques for Idea of Hindi to English Idiom Translation", *International Journal of Computational Intelligence Research,* vol.13, no. 5, May 2017.

[4] H. Mishra, R. K. Chakrawarti and P. Bansal, "A new approach for Hindi to English idiom translation", *International Journal on Computer Science and Engineering,* vol. 9, no.7, May 2017.

[5] E. Agirre and P. Edmonds, "Word Sense Disambiguation: Algorithms and Applications", Springer, pp. 1–28. Available: www.wsdbook.org/index.html

[6] A. Kaur, "Development of an Approach for Disambiguating Ambiguous Hindi postposition", International Journal of Computer Applications, 5(9), August 2010.

[7] "Wordnet", Princeton University. Princeton, NJ, USA. http://wordnet.princeton.edu/

[8] "Hindi Wordnet", Center for Indian Language Technology (CFILT) Solutions, IIT Bombay, Mumbai, India. http://www.cfilt.iitb.ac.in/wordnet/webhwn/

[9] G. A. Miller, B. Richard, F. Christiane, G. Derek and J. M. Katherine, "Introduction to WordNet: an on-line lexical database", *International Journal of Lexicography*, vol. 3 no. 4, pp. 235 – 244, 1990.

[10] G. A. Miller, "WordNet: a lexical database for English", *Communications of the ACM*, vol. 38 no. 11, pp. 39 – 41, November 1995.

[11] F. Christiane, (Ed.), "WordNet. An electronic lexical database", Cambridge, MA: MIT Press, 1998.

[12] P. Bhattacharyya, "IndoWordNet", *Lexical Resources Engineering Conference 2010, LREC 2010, Malta, May, 2010*.

[13] M. Sinha, M. K. Reddy, P. Pande, L. Kashyap and P. Bhattacharyya, "Hindi Word Sense Disambiguation", *International Symposium on Machine Translation, Natural Language Processing and Translation Support Systems, Delhi, India, November 2004.*

[14] S. K. Dwivedi, P. Rastogi, "Critical Analysis of WSD Algorithms", *International Conference on Advances in Computing, Communication and Control (ICACCC), Mumbai, Maharashtra, India, January 23-24, 2009.*

[15] T. Pedersen and R. Mihalcea, "Advances in Word Sense Disambiguation", Tutorial at ACL, June 25, 2005.

[16] R. K. Chakrawarti and P. Bansal. "Approaches for Improving Hindi to English Machine Translation System", Indian Journal of Science and Technology, vol. 10 no. 16, April 2017.
DOI: 10.17485/ijst/2017/v10i16/111895

[17] D. Genzel, J. Uszkoreit and F. Och, *"Poetic" Statistical Machine Translation: Rhyme and Meter, Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, MIT, Massachusetts, USA, 9-11 October 2010*. 2010 Association for Computational Linguistics, pp. 158–166.

[18] H. Mishra, R. K. Chakrawarti and P. Bansal, "Implementation of Hindi to English Idiom Translation System", In: Kamal R., Henshaw M., Nair P. (eds) *International Conference on Advanced Computing Networking and Informatics. Advances in Intelligent Systems and Computing, vol. 870,* Springer, Singapore, November 2018
https://doi.org/10.1007/978-981-13-2673-8_39

[19] R.K. Chakrawarti and P. Bansal, "Hindi-To-English Machine Translation System for Primary Education", *Doctoral Conference (DocCon-2016), Janardan Rai Nagar Rajasthan Vidyapeeth University, In collaboration with CSI and ACM Udaipur Chapter*, Udaipur, Rajasthan, March 2016.

Corresponding Author's E-mail ID: *rajesh_kr_chakra@yahoo.com*