Exploiting WordNet glosses to disambiguate nouns through verbs

Donato Barbagallo, Leonardo Bruni, and Chiara Francalanci
Department of Electronics and Information
Politecnico di Milano
Milano, Italy
{barbagallo,bruni,francala}@elet.polimi.it

***Abstract*—This paper presents an unsupervised graph-based algorithm for word sense disambiguation based on WordNet glosses. The algorithm exploits the contributions of verbs in identifying the correct senses of nouns. Due to the complexity of WordNet's semantic network, we have defined disambiguation as a similarity optimization problem and solved it through a genetic algorithm. Testing compares the performance of our algorithm with that of a traditional method based on Wu-Palmer similarity measure. Our approach shows an overall precision of about 68% and a statistically significant average increase of precision of about 3% with respect to the traditional algorithm.**

Keywords - word sense disambiguation; WordNet; genetic algorithm.

I. Introduction

Word Sense Disambiguation (WSD) is the ability of identifying the meaning of words in a given sentence. It represents a fundamental research problem in Natural Language Processing with many practical applications, such as search engines, information retrieval, and sentiment analysis.

WSD has made a considerable progress in the last few years and can now obtain good results through supervised algorithms. Though results can be very precise, the literature recognizes the high costs and strong feasibility limits of these techniques, due to their need for context-dependent annotated corpora [6]. On the other hand, unsupervised techniques can also be applied to WSD. In this field, the term unsupervised is usually referred to techniques that are not necessarily knowledge-free, since some kind of knowledge base, i.e. dictionaries or computational lexicons, is needed [2]. These knowledge bases usually provide a context-independent sense inventory and relations among senses, which can be exploited to perform WSD. Though the literature tends to recognize that supervised methods usually outperform unsupervised ones, from a cost-benefit analysis point of view it can be still more convenient to invest and develop unsupervised methods. Indeed, in some applications of WSD, such as information retrieval, perfect word sense information would be of limited utility [5].

The literature has often combined unsupervised methods based on semantic networks such as WordNet [12] with the so-called similarity measures. These measures assume that two words are similar when they appear in a similar context, e.g., in the same sentence of paragraph, and contexts are similar when they contain similar words [28]. According to these assumptions, word senses whose definitions have the highest score of similarity are assumed to be the correct ones [27][28].

In this paper, we present a new unsupervised method to disambiguate nouns based on WordNet and on the concept of similarity. The innovative aspect of this method is that it is able to exploit WordNet glosses and verbs and create the link between nouns and verbs sub-graphs, outperforming traditional approaches based only on nouns. We compare the performance of our algorithm with a classical disambiguation approach based on nouns and the Wu-Palmer similarity measure.

There are many techniques in literature to solve the similarity optimization problem. This work uses a genetic algorithm to solve the similarity problem. Although genetic algorithms are global search heuristics and their results are not guaranteed to be optimal solutions, they are also known to outperform other optimization techniques when the problem space is large, as in the case of a semantic network made of hundreds of thousands of words. Moreover, genetic algorithms perform better when the solution space is discontinuous and include multiple local optima [30].

The remainder of this paper is structured as follows. Section II explains the background knowledge behind our algorithm. Section III presents the algorithms, Section IV describes the experiments and discusses the results. Section V presents related research and highlights the innovative aspects of our work. Finally, Section VI shows some conclusion and presents future research directions.

II. Methodology and Technology Resources

In this section, the methodology and the technology resources are briefly described.

A. Genetic Algorithm

Evolutionary computation techniques [11] make use of Darwin's evolutionary principles and translate them into heuristic algorithms that can be used to search for optimal solutions to a problem. In a search algorithm, the objective is to find the best possible solution in a fixed amount of time. When the search space grows in size, an exhaustive search becomes quickly unfeasible. The key aspect distinguishing an evolutionary search algorithm from more traditional heuristic algorithms is that it is *population-based*. Through the adaptation of successive generations of a large number of individuals, an evolutionary algorithm performs an efficient search.

Genetic algorithms are a particular class of evolutionary algorithms. In a genetic algorithm, a potential solution is represented by a *chromosome*, usually encoded as an array of

 ISBN: 978-1-61208-104-5

bits or characters. A single bit or a set of bits coding part of the solution is called *gene*. In turn, an *allele* is one of the possible instances of the gene.

The first population is typically randomly generated. Then a measure of goodness (necessarily domain dependent) is computed for each chromosome. Guided by this quantitative information, together with a set of genetic operators like crossover and mutation, genetic algorithms move from one population of chromosomes to a new population. Typically, the evolution terminates when either a fixed number of generations has been created or the fitness value of a chromosome reaches a target threshold.

Genetic algorithms have been used for many applications like optimization, classification, prediction, economy, ecology and automatic programming.

B. WordNet

WordNet [12] is a freely available lexical database for the English language that organizes nouns, verbs, adjectives and adverbs into hierarchies of *synonym sets* or *synsets*. Each synset groups words with a unique meaning and it has a gloss that describes the concept that it represents. For example, the synset composed by the words {apartment, flat} represents the concept defined by the gloss "a suite of rooms usually on one floor of an apartment house". Many glosses are extended with the addition of some examples of usages of the concept that they describe.

WordNet is organized as a network of concepts linked by semantic relations, like *hypernym*, *hyponym*, *meronym*, *holonym*, and *antonym*. However, these relations do not cross part of speech boundaries. Thus, semantic relations are tied to a particular part of speech, creating different and separate sub-graphs for nouns, verbs, adjectives and adverbs. In our experiments we use WordNet 3.0 and we focus on the hypernym hierarchy that represents the most complete set of relations.

III. THE WSD SYSTEM

We follow two strategies to perform the disambiguation of nouns. The first, called *base algorithm*, relies only on the information carried by the nouns in a sentence. The second, called *enhanced algorithm*, aims at improving the results exploiting also the information that can be extracted from verbs through disambiguated glosses.

A. Base algorithm

The basic idea of our WSD system, also exploited in [10], relies on the assumption that terms that appear in the same sentence tend to be semantically similar. The genetic algorithm is used to find a set of senses that maximizes the similarity between the terms to be disambiguated. Because both the number of possible senses for each word and the cardinality of the set of words to disambiguate can be large the search space become huge. Thus genetic algorithms are a suitable solution for this kind of problem.

Similarity is a widely used concept. According to Budanitsky and Hirst [1], it is possible to make a distinction between semantic similarity and relatedness: semantic similarity is a special kind of relatedness between two words and denotes the degree of semantic association between them. Measures of relatedness can be made across part of speech boundaries and are not tied to the is-a relation. However, for the purpose of this paper, we refer to both kinds of measure with the term similarity. Many similarity measures have been proposed, such as information content [4], Lin [7], Jiang-Conrath [3], Banarjee-Pedersen [9], Wu-Palmer [8]. The whole set of cited measures has been compared through some preliminary experiments showing that Wu-Palmer can be considered the measure providing the best results.

Wu-Palmer defines the similarity of two concepts by measuring how closely they are related in the hierarchy, i.e., the similarity measure between a pair of concepts *c1* and *c2* is:

$$sim_{WP}(c1, c2) = (2 * N3)/(N1 + N2 + 2 * N3) \qquad (1)$$

where N1 is the number of nodes on the path from c1 to c3 (the least common superconcept of c1 and c2), N2 is the number of nodes on the path from c2 to c3, and N3 is the number of nodes on the path from c3 to the root of the hierarchy.

Table I shows some examples of measures between pairs of nouns computed by (1), where with city#1 we mean the first sense of “city”, with animal#1 we mean the first sense of “animal” and so on. The first sense of “turkey” is “large gallinaceous bird with fan-shaped tail; widely domesticated for food” and its second sense is “a Eurasian republic in Asia Minor and the Balkans”. As we can expect, the concept of “city” intended in its first sense, i.e., “a large and densely populated urban area”, is more similar with the second sense of “turkey” than with the first sense. Analogously, the concept of “animal” intended in its first sense, i.e., “a living organism characterized by voluntary movement”, is more similar with the first sense of “turkey” than with the second sense.

TABLE I. SIMILARITY BY WU-PALMER’S MEASURE

	turkey#1	turkey#2
city#1	0.20	0.75
animal#1	0.67	0.29

In our system the similarity measure presented above represents the core of the fitness function of the genetic algorithm. Each solution is represented by a chromosome that is encoded as a sequence of positive integer numbers. Each gene of a chromosome is a possible sense of a term. The fitness value for each chromosome is computed as follows:

$$F = \sum_i \sum_{j=i+1} sim(s(w_i), s(w_j)) \qquad (2)$$

where w_i and w_j are two terms, $s(w_i)$ and $s(w_j)$ are the candidate senses of w_i and w_j.

The similarity measure used in (2) is slightly different from the measure in (1). In order to perform better on general documents, the original value is weighted by the frequency of the words' sense, because in general context, words tend to assume their more frequent meaning. In WordNet, word senses are ordered by their frequency of use, i.e. the most frequent senses are indicated with lower ordinal numbers. So we define the new similarity measure as:

$$sim\left(s(w_i), s(w_j)\right) = \left(\frac{1}{n_i} + \frac{1}{n_j}\right) * sim_{WP}\left(s(w_i), s(w_j)\right) \quad (3)$$

where n_i and n_j denote the ordinal number of $s(w_i)$ and $s(w_j)$ as reported by WordNet.

B. Enhanced algorithm

We have observed that the information carried by the nouns may not be enough. For example, if we want to disambiguate a sentence like "I ate a tasty turkey for Christmas" using the Base Algorithm, the set of nouns used for the disambiguation is composed by the words {turkey, Christmas}. Just considering the couple it is not clear whether the noun *turkey* has to be interpreted as fowl or Eurasian republic. On the other hand, if the verb *eat* is added to the set, it becomes clear that the former is the right meaning, because the verb carries contextual information.

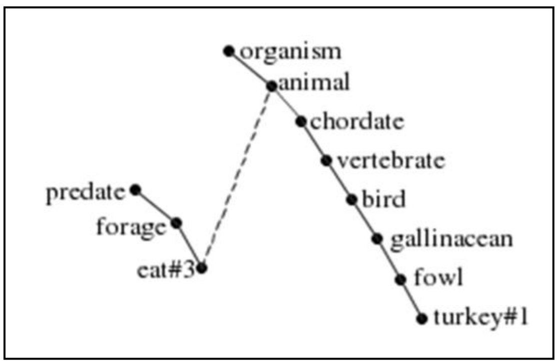

Figure 1. Example of a new verb-noun relation in WordNet

Similarity measures can be applied only to pairs of words of the same part of speech. To deal with this limit we make use of an additional resource [13] where word forms from the definitions (called *glosses*) in WordNet's synsets are manually linked to the context-appropriate sense in WordNet.

In order to take into account this new kind of information we extended the size of the chromosome defined in the previous algorithm. For each sense of each verb found in the sentence of which we want to disambiguate the terms we extract a noun from its annotated gloss and then a new gene is added to the chromosome whose possible value is only the sense with which it was tagged in the gloss. We provide a simple example using the set {{turkey, Christmas},{eat}}. The base form of the chromosome will have two genes since there are two nouns in the set. The verb *eat* in WordNet has six senses, so we add six "monosemic" genes to the chromosome, each one representing one noun extracted from the gloss of each sense of the verb. The new chromosome is shown in Table II. Figure 1 graphically represents the example and shows how easily it is possible to create a relation (dashed line) between a noun and a verb, while unbroken lines represent hypernym hierarchies for the verb eat (synset #3) and the noun turkey (synset #1).

TABLE II. EXAMPLE OF CHROMOSOME

Gene no.	Noun	Alleles
1	turkey	1-5
2	christmas	1-2
3	solid_food	1
4	meal	1
5	animal	1
6	way	1
7	resource	1
8	action	4

When one of the new genes is involved in the computation of the similarity value, the frequency with which is weighted is the frequency of the verb that has generated the gene.

IV. EVALUATION AND DISCUSSION

Experiments are performed using the JGAP [29] library to implement the genetic algorithm. Due to the intrinsic heuristic nature of genetic algorithms, we performed several tests with different settings of the parameters of the genetic operators. These tests have highlighted that result deltas are irrelevant. In this section, we present the results obtained by applying the default configuration of genetic operators as provided by the genetic algorithm implementation in JGAP. Similar tests executed by varying the population size have highlighted also the fact that precision does not increase significantly when the population size overgrows ten chromosomes. Given that WordNet word senses are ordered by frequency of use, the first ten senses are sufficient to cover the common usage of words.

We have performed a sentence by sentence analysis for two main motivations: *(i)* in our case, the disambiguation process is part of a larger project on sentiment analysis (cf. [31]) considering short sentences, such as tweets or blog posts, and *(ii)* we found out that the number of words to disambiguate and precision are uncorrelated, as shown later in this section.

Algorithm performance is measured in terms of precision and recall. Following [32], precision is defined as the number of correct disambiguated senses divided by the total number of answers reported; recall is defined as the number of correct disambiguated senses divided by the total number of senses. Since our methods can assign a sense for every word, precision equals recall.

The results of the experiments are evaluated on SemCor Corpus, the sense-tagged version of the Brown Corpus, by automatically comparing the sense-tags in SemCor with those computed by our algorithms. We carried out experiments over 19 randomly selected SemCor files (br-a02, c01, e04, e27, f10, f22, f43, g18, g19, g28, h18, j04, j12, j20, j57, j70, k04, l18, r05).

Because of the heuristic nature of genetic algorithms we run each test ten times in order to have an empirical assessment of the variability of the results.

 ISBN: 978-1-61208-104-5

Table III shows the comparison statistics between our base and enhanced algorithms. It is worth noting how enhanced algorithm outperforms base algorithm in precision. Indeed Base Algorithm shows an overall average precision of 64.39, while enhanced algorithm obtains an average of 67.78. A t-test on the average results obtained by the two algorithms file by file through the ten simulations reveals how these differences are statistically significant ($t = -6.719$, $p < 0.001$). Moreover, enhanced algorithm should also be preferred to base algorithm because of its lower standard deviation, which guarantees more coherent results through the simulations. Results also show how there is a high variance in the results among the 19 files. The maximum difference between the two approaches has been found in file br-g18 where the precision obtained by enhanced algorithm was 7.74% above the precision obtained by base algorithm. In all runs enhanced algorithm outperforms base algorithm. Finally, the maximum precision values, obtained with file br-e27, show how enhanced algorithm is able to exceed the threshold of 80% (81.10%), while base algorithm is less precise with a 77.41%.

TABLE III. COMPARISON STATISTICS BETWEEN BASE AND ENHANCED ALGORITHMS

Algorithm	precision (mean)	precision (standard deviation)	maximum precision value
Base	64.39	6.1	77.41
Enhanced	67.78	5.8	81.10

It is interesting to note how there does not exist a relation between the number of nouns in a single sentence that has been analyzed and the corresponding precision results. Unexpectedly, base algorithm does not show correlation between the two variables ($r = 0.120$, $p < 0.001$), thus implying that context window size is not correlated to precision. An even more significant result is obtained with enhanced algorithm ($r = 0.074$, $p < 0.001$), supporting our initial experimental decision of running analyses sentence by sentence rather than paragraph by paragraph.

Despite the promising results, we noted that similarity alone is not sufficient. Indeed, usually there can be different possible set of senses of the nouns that are fairly plausible, in a given sentence.

TABLE IV. SIMILARITY DIFFERENCES BETWEEN SEMCOR SENSES AND BASE ALGORITHM'S BEST SOLUTION

	SemCor	Base Algorithm
Target	5	3
Chart	1	2
Additives	1	1
Similarity score	0.24	1.11

Table IV shows this drawback using the sentence "The target chart quickly and briefly tells you which additives do what." extracted from the file named br-e27 of the SemCor Corpus. By computing the overall similarity measure on both sets of senses, we obtain a value of 0.24 in the SemCor sense-tagged set and a value of 1.11 in the set computed by our base algorithm. The meaning of the senses in the latter set clearly indicates that words are strongly related: the fifth sense of "target" is "the goal intended to be attained" and the first sense of "chart" is "a visual display of information", while the third sense of "target" is "the location of the target that is to be hit" and the second sense of "chart" is "a map designed to assist navigation by air or sea". We are currently studying if and how this observation can be exploited in order to improve the precision of the disambiguation process.

Checking some results by hand we have also noted that in SemCor some words have been sense-tagged with a meaning that tough it is not totally wrong, it is at least ambiguous. This fact can be explained by an example. The file br-l18 contains the sentence "I asked her why she couldn't do it tomorrow, but it seems the muse is working good tonight and she's afraid to let it go" where the word "muse" has been tagged with the meaning of "in ancient Greek mythology any of 9 daughters of Zeus and Mnemosyne; protector of an art or science". Although this meaning is not completely wrong, it is definitely more correct the meaning of "the source of an artist's inspiration".

More generally, there are some cases where the sense assigned in SemCor is right, but, nevertheless not necessarily unambiguous. This observation raises the question whether the fine granularity of WordNet is appropriate for the word sense disambiguation task as discussed in [19].

V. RELATED WORK

The idea of using similarity among synsets in WordNet is not original. Much literature has tried to exploit WordNet semantic relations for WSD. In particular Zhang et al. [10] have implemented a genetic algorithm for noun disambiguation based on the Wu-Palmer measure of similarity and on SemCor word frequency. Their results are considerable as they obtained an overall 71.98% precision on the general SemCor testbase. Yarowsky [14] presented an unsupervised learning algorithm whose performance (overall 96% accuracy) is comparable to that of supervised algorithms. Yarowsky's algorithm applies two constraints to the properties of human language to discriminate among senses, i.e., one sense per collocation and one sense per discourse. Recently, Social Network Analysis has gained interest in WSD through the use of its classical graph connectivity metrics. Navigli and Lapata [2] used local centrality and global graph measures, showing that the former outperforms the latter and is comparable to the current state of the art. Unsupervised graph-based methods have been exploited also by Mihalcea [15]. In this work, synstet similarity is defined, similarly to Lesk [27], as a function of the number of common tokens in the definitions of word senses. This algorithm obtains an overall precision of 54.2%, being able to disambiguate nouns, and also verbs, adjectives, and adverbs. Recently, Navigli and Velardi [17] introduced the Structural Semantic Interconnections (SSI) algorithm that detects relevant semantic patterns of word senses through the use of a context-free grammar, obtaining a precision of 86% for nouns and almost 70% for verbs.

Several works have also attempted to use other resources in addition to WordNet [12]. In particular, they have focused

ISBN: 978-1-61208-104-5

on ontologies such as OntoNotes [18] and SUMO [16]. More specifically, these works are built on top of SVM-based supervised algorithms. Zhong et al. [19][20], based on OntoNotes, perform domain adaptation experiments trained using the knowledge sources of local collocations, part-of-speech, and surrounding words. The results of these papers highlight the importance of having an appropriate level of sense granularity. On the other hand, authors in [21] performed semantic disambiguation for Spanish. They used semantic classes instead of senses, based on the SUMO ontology. This approach allows collecting a larger number of examples for each class while polysemy is reduced, improving the accuracy of semantic disambiguation. In turn, works in [25][26] have proposed ways to exploit additional knowledge given by domain information. Specifically, Magnini et al. [26] proposed an extended version of WordNet called WordNet Domains obtaining an average 70% precision, while the work in [25] proposes a preliminary algorithm including domain information.

The need of an augmented version of WordNet has been formalized in [22] and [23]. The inclusion of logics and the exploitation of glosses to connect verbs and nouns have been explicitly called for. Naskar and Bandyopadhyay [24] have implemented a variation of the Lesk algorithm using eXtended WordNet [23] and its glosses to disambiguate nouns, verbs, and adjectives obtaining an 85% overall precision.

The main difference between our algorithm and other algorithms is that we are now able to deal with one of the main drawbacks of WordNet when using similarity measures, i.e., with the fact that the organization of words in hierarchies does not cross part of speech boundaries. Indeed, by extracting disambiguated nouns from the disambiguated glosses of verbs we create a new relation in WordNet that links each sense of each verb to one or more nouns, making it possible to process verbs through related nouns. This new kind of relation gives suggestions, in an automatic manner, about which nouns are used with which verbs in natural language.

VI. CONCLUSION AND FUTURE WORK

In this paper, we presented a new algorithm that uses WordNet disambiguated glosses to create a relation between nouns and verbs in WordNet network. Our results suggest that the information provided by the new relation can be significantly helpful in the context of WSD. We have tested our algorithm on 19 randomly chosen SemCor files and we have found that it is able to outperform an algorithm based only on nouns and on Wu-Palmer similarity measure. Our algorithm has been able to reach an 81.10% precision on file br-e27 and an average of 67.78%.

The analysis of the results of our experiments also highlighted three main drawbacks: *(i)* though manually tagged, SemCor disambiguated words can present very ambiguous synsets that could even be considered wrong; *(ii)* as previous literature has pointed to, WordNet is a too fine-grained resource for WSD; *(iii)* the well-established methodology based on similarity often leads to wrong solutions since the right synsets are not necessarily the most similar. The first two mentioned issues strictly depend on the used tools, indeed, as shown in [19] precision would benefit from having a more coarse-grained resource such as OntoNotes [18]. Regarding the third issue, we are working in two directions: *(i)* on the development of a tool that allows the addition of contextual information to WordNet creating new types of relations, e.g., adjective-noun, further improving the presented Enhanced Algorithm; *(ii)* integrating the tool with domain-specific ontologies that could be used when dealing with documents in a specific context.

As a further development, we are considering to exploit domain knowledge. We have also run a preliminary version of new algorithms that include WordNet Domains in the WSD process, but they need to be refined since results are not promising, probably due to the nature of WordNet Domains which is too coarse-grained. Another important evolution of our algorithm is to focus on the disambiguation of other parts of speech, especially verbs, that could significantly help improve the overall sentence disambiguation, and adjectives, which could be useful for our application context, i.e., sentiment analysis.

REFERENCES

[1] A. Budanitsky, and G. Hirst, "Semantic distance in WordNet: An experimental, application-oriented evaluation of five measures," In Workshop on WordNet and Other Lexical Resources, Second meeting of the North American Chapter of the Association for Computational Linguistics. Pittsburgh, June 2001.

[2] R. Navigli, and M. Lapata, "An Experimental Study of Graph Connectivity for Unsupervised Word Sense Disambiguation," *IEEE Trans. Pattern Anal. Mach. Intell.* 32, 4, pp. 678-692, Apr. 2010.

[3] J. Jiang, and D. Conrath, "Semantic similarity based on corpus statistics and lexical taxonomy," Proc. International Conference on Research in Computational Linguistics. Taiwan, 1997.

[4] P. Resnik, "Using information content to evaluate semantic similarity in a taxonomy," Proc. 14th International Joint Conference on Artificial Intelligence. Montreal, 1995.

[5] P. Resnik, and D. Yarowsky, "Distinguishing systems and distinguishing senses: new evaluation methods for Word Sense Disambiguation," *Nat. Lang. Eng.* 5, 2, pp. 113-133, June 1999.

[6] T.H. Ng, "Getting Serious about Word Sense Disambiguation," Proc. ACL SIGLEX Workshop Tagging Text with Lexical Semantics: Why, What, and How?, pp. 1-7, 1997

[7] D. Lin, "Using syntactic dependency as a local context to resolve word sense ambiguity," Proc. 35th Annual Meeting of the Association for Computational Linguistics, pp. 64-71. Madrid, July 1997.

[8] Z. Wu, and M. Palmer, "Verb semantics and lexical selection," Proc. 35th Annual Meeting of the Association for Computational Linguistics, pp. 133-138. Las Cruces, NM, July 1997.

[9] S. Banerjee, and T. Pedersen, "Extended gloss overlaps as a measure of semantic relatedness," Proc. International Joint Conference on Artificial Intelligence, volume 18, pp. 805-810, 2003.

[10] C. Zhang, Y. Zhou, and T. Martin, "Genetic word sense disambiguation algorithm," Proc. Second International Symposium on Intelligent Information Technology Application (IITA 08), 2008.

[11] S. N. Sivanandam, and S. N. Deepa, "Introduction to Genetic Algorithms," Springer, 2008.

[12] C. Fellbaum, "WordNet: An electronic lexical database," MIT Press, Cambridge, MA.

[13] Princeton WordNet Gloss Corpus. http://wordnet.princeton.edu/glosstag.shtml, last access: 07-19-2010.

 ISBN: 978-1-61208-104-5

[14] D. Yarowsky, “Unsupervised word sense disambiguation rivaling supervised methods,” In *Proceedings of the 33rd Annual Meeting on Association For Computational Linguistics*. Annual Meeting of the ACL. Association for Computational Linguistics, Morristown, NJ, pp. 189-196, June 1995.

[15] R. Mihalcea, “Unsupervised large-vocabulary word sense disambiguation with graph-based algorithms for sequence data labeling,” In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*. Human Language Technology Conference. Association for Computational Linguistics, Morristown, NJ, pp. 411-418, 2005.

[16] I. Niles and A. Pease, “Towards a standard upper ontology,” In *Proceedings of the international Conference on Formal ontology in information Systems - Volume 2001*. FOIS '01. ACM, New York, NY, pp. 2-9, 2001.

[17] R. Navigli, and P. Velardi, “Structural Semantic Interconnections: A Knowledge-Based Approach to Word Sense Disambiguation,” *IEEE Trans. Pattern Anal. Mach. Intell.* 27, 7, pp. 1075-1086, July 2005.

[18] S. S. Pradhan, E. Hovy, M. Marcus, M. Palmer, L. Ramshaw, L., and R. Weischedel, “OntoNotes: A Unified Relational Semantic Representation,” In *Proceedings of the international Conference on Semantic Computing*. ICSC. IEEE Computer Society, Washington, DC, pp. 517-526, 2007.

[19] Z. Zhong, H. T. Ng, and Y. S. Chan, “Word sense disambiguation using OntoNotes: an empirical study,” In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Annual Meeting of the ACL. Association for Computational Linguistics, Morristown, NJ, pp. 1002-1010, 2008.

[20] Z. Zhong, and H. T. Ng, “Word sense disambiguation for all words without hard labor,” In *Proceedings of the 21st international Jont Conference on Artifical intelligence*. H. Kitano, Ed. International Joint Conference On Artificial Intelligence. Morgan Kaufmann Publishers, San Francisco, CA, pp. 1616-1621, 2009.

[21] R. Izquierdo-Bevia, L. Moreno-Monteagudo, B. Navarro, A. Suarez, “Spanish All-Words Semantic Class Disambiguation Using Cast3LB Corpus,” Lecture Notes in Computer Science numb 4293, Springer-Verlag, pp. 879-888, 2006.

[22] P. Clark, C. Fellbaum, J. R. Hobbs, P. Harrison, W. R. Murray, and J. Thompson, “Augmenting WordNet for deep understanding of text,” In *Proceedings of the 2008 Conference on Semantics in Text Processing*, ACL Workshops. Association for Computational Linguistics, Morristown, NJ, pp. 45-57, 2008.

[23] D. Moldovan, and V. Rus, “Explaining answers with extended wordnet,” In Proc. of ACL’01, 2001.

[24] S. K. Naskar, and S. Bandyopadhyay, “Word Sense Disambiguation Using Extended WordNet,” In *Proceedings of the international Conference on Computing: theory and Applications*, ICCTA. IEEE Computer Society, Washington, DC, pp. 446-450, 2007.

[25] S. G. Kolte, S. G. Bhirud, “Word Sense Disambiguation Using WordNet Domains,” In *Proceedings of the 2008 First international Conference on Emerging Trends in Engineering and Technology* ICETET. IEEE Computer Society, Washington, DC, pp. 1187-1191, 2008.

[26] B. Magnini, C. Strapparava, G. Pezzulo, and A. Gliozzo, “The role of domain information in Word Sense Disambiguation,” *Nat. Lang. Eng.* 8, 4, pp. 359-373, 2002.

[27] M. Lesk, “Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone,” In *Proceedings of the 5th Annual international Conference on Systems Documentation*. V. DeBuys, Ed. SIGDOC '86. ACM, New York, NY, pp. 24-26, 1986.

[28] Y. Karov, and S. Edelman, “Similarity-based word sense disambiguation,” *Comput. Linguist.* 24, 1, pp. 41-59, Mar. 1998.

[29] K, Meffert, et al., “JGAP – Java Genetic Algorithms and Genetic Programming Package”. http://jgap.sf.net.

[30] R. K. Belew, and M. D. Vose, “Foundations of genetic algorithms 4,” Morgan Kaufmann, 1997.

[31] D. Barbagallo, C. Cappiello, C. Francalanci, and M. Matera, “A Reputation-based DSS: the INTEREST Approach,” In *Proceedings of ENTER 2010: International Conference On Information Technology and Travel&Tourism*, February 2010.

[32] E. Agirre, and G. Rigau, "Word sense disambiguation using conceptual density," Proc. of COLING, 1996.