Research Journal of Applied Sciences, Engineering and Technology 4(12): 1815-1821, 2012
ISSN: 2040-7467

Submitted: February 28, 2012 Accepted: March 16, 2012 Published: June 15, 2012

Studying of Semantic Similarity Methods in Ontology

[1]Vahideh Reshadat and [2]Mohammad-Reza Feizi-Derakhshi
[1]Department of Computer, Shabestar Branch, Islamic Azad University, Shabestar, Iran
[2]Department of Computer, University of Tabriz, Tabriz, Iran

Abstract: Humans are able to easily judge if a pair of concepts are related in some way. Understanding of how humans are able to perform this task is not easy. Semantic similarity denotes computing the similarity between concepts, having the same meaning or related information, which are not necessarily lexically similar. Semantic similarity between concepts plays an important role in Semantic Web, knowledge sharing, Web mining, semantic sense understanding and text summarization. This also is an important problem in Natural Language Processing and Information Retrieval Researches. These techniques are becoming important components of most of the Information Retrieval (IR), Information Extraction (IE) and other intelligent knowledge based systems. Therefore it has received considerable attention in the literature. Ontology has a good hierarchical structure of concepts. In the ontology, semantic information can be realized through the semantic relationship of concepts. Ontology-based semantic similarity techniques can estimate the semantic similarity between two hierarchically expressed concepts in a given ontology or taxonomy. Semantic similarity is usually computed by mapping concepts to ontology and by examining their relationships in it. The most popular semantic similarity methods are implemented and evaluated using WordNet and MeSH. Several algorithmic approaches for computing semantic similarity have been proposed. This paper discusses the various approaches used for identifying semantically similar concepts in ontology.

Key words: Conceptual similarity, ontology, semantic similarity, similarity method, taxonomy, wordnet

INTRODUCTION

With the rapid development of network information, it is increasingly dependent on information retrieval technology to find required information. The purpose of information retrieval is to assist users in finding information they are looking for. Information retrieval is currently being applied in a variety of application domains from database systems to web information search engines. The goal of information retrieval process is to retrieve information relevant to a given request and also retrieve all the relevant information eliminating the nonrelevant information. However, the traditional method based on keyword cannot inform the semantic information of resources itself. The need of semantic understanding the web information becomes urgent. But there are lots of jobs should be done before the web information is understood by both human and machine. Such as page annotation, ontology construction, knowledge reasoning, semantic searching and natural language processing. The semantic similarity of concepts is the base of these and also other jobs such as knowledge sharing, Web mining and semantic sense understanding.

Humans are able to judge the relatedness of words (concepts) relatively easily and are often in general agreement as to how related two words are. For example, few would disagree that “pencil” is more related to “paper” than it is to “boat”. Semantic similarity relates to computing the similarity between concepts which are not necessarily lexically similar. It is possible to develop methods capable of detecting similarities between conceptually similar documents even when they don’t contain lexically similar terms. The lack of common terms in two documents does not necessarily mean that the documents are not related. Computing text similarity by classical information retrieval models (e.g., Vector Space, Probabilistic, Boolean) is based on lexical term matching. However, two terms can be semantically similar (e.g., can be synonyms or have similar meaning) although they are lexically different. Therefore, classical retrieval methods will fail to associate documents with semantically similar but lexically different terms (Butanisky and Hirst, 1999). The ontological theory and technology originate from the field of knowledge engineering and artificial intelligence can deal with the problem of natural language understanding and reasoning mechanism based on semantics. We present an evaluation of several semantic similarity approaches using two well known taxonomic hierarchies (or ontologies) namely WordNet (http:// worelnet.princeton.edu) and MeSH (http://www. nlm.nith.gov/mesh). WordNet is a controlled vocabulary and thesaurus offering a taxonomic hierarchy of natural language terms developed at Princeton University. MESH (Medical Subject Heading) is also a controlled vocabulary

Corresponding Author: Vahideh Reshadat, Department of Computer, Shabestar Branch, Islamic Azad University, Shabestar, Iran

and thesaurus developed by the U.S. National Library of Medicine (NLM) offering a hierarchical categorization of medical terms.

This research discusses the survey of different similarity measuring methods used to compare and find very similar concepts of ontology. A comparison of various similarity computation methods is discussed and the paper is concluded with a summary of the discussed issues.

METHODOLOGY

Semantic similarity methods: The purpose of this survey is to review the similarity computation methods. In this section, we discuss about various similarity methods. The similarity methods are:

- Edge-based approaches
- Depth-based approaches
- Information Content approaches
- Feature-based approaches

These approaches are reviewed in the rest of this paper.

Edge-based methods: The measure of edge-based methods turns the concepts, properties and instances into an ontology. Nodes in ontologies denote concepts of ontology and edges connecting the nodes denote different relations between concepts. A simple way to measure semantic similarity in a taxonomy is to evaluate the distance between the nodes corresponding to the items being compared. The shorter distance results in high similarity. Path length-based methods measure the similarity between two concepts c_1, c_2 as a function of the path linking the terms in the taxonomy and of the position of the terms in the taxonomy (Rada *et al*., 1989).

Shortest path length methods: Shortest path measure is a simple and powerful measure in hierarchical semantic nets. The first measure has to do with how close in the taxonomy the two concepts are. Distance can be successfully used to assess the conceptual distance between sets of concepts when used on a semantic net of hierarchical relations. In (Rada *et al*., 1989; Rada and Bicknell, 1989) shortest path length approach is followed assuming that the number of edges between terms in a taxonomy is a measure of conceptual distance between concepts. He used a matrix formed by all the words on the shortest path connecting the two concepts to measure the semantic similarity.

$$\text{distRada}(c_i, c_j) = \text{Minimal number of edges in a path from } c_i \text{ to } c_j \quad (1)$$

This method yields good results. Since the paths are restricted to IS-A relation, the path lengths corresponds to conceptual distance. Moreover, the experiment has been conducted for specific domain ensuring the hierarchical homogeneity. The drawback with this approach is that, it is compatible only with commonality and difference properties and not with identity property. According to commonality property the similarity between two concepts x and y is related to what they have in common and also in the difference property, the similarity between two concepts x and y is related to their differences. Identity property expresses that the similarity between two identical concepts is the highest possible value for similarity. This work is the base of edge-counting method. The connect-power of each edge is considered equal. That isn't so consistent with human intuition.

Hirst and St-Onge (1998) measure considers all the relations defined in WordNet. All links in WordNet are classified as Upward (e.g., part-of), Downward (e.g., subclass) or Horizontal (e.g., opposite-meaning). Further, they also describe three types of relations between words- extra-strong, strong and medium-strong. Any two words are related by one of these types of relations if they conform to certain rules.

The idea behind this measure of semantic relatedness is that two lexicalized concepts are semantically close if their Word Net synsets are connected by a path that is not too long and that does not change direction too often (Hirst and St-Onge, 1998). Although this measure gives a different perspective of similarity between two concepts, it seems to poorly perform (Butanisky and Hirst, 1999) mainly because it lies in its tendency to wander than in the use of concept relationships (Hliaoutakis, 2005).

Weighted shortest path length methods: There is another simple edge-counting approach. In this method, weights are assigned to edges. In brief, weighted shortest path measure is a generalization of the shortest path length. Extending the shortest path length measure, the use of weighted links is proposed to compute the similarity between two concepts. The weight of a link may be affected by:

- The density of the taxonomy at that point
- The depth in the hierarchy
- The strength of connotation between parent and child nodes

Then, computing the distance between two concepts is translated into summing up the weights of the traversed links instead of counting them (Wang and Hirst, 2011). Given a node/concept c in Word Net, depth refers to the number of nodes between c and the root of Word Net, (i.e., the root has depth zero, its hyponyms depth one and so on). There are more variations in the definition of density, but it is usually defined as the number of edges leaving c (i.e., its number of child nodes) or leaving its parent node(s) (i.e., its number of sibling nodes).

Irregular density often results in unexpected conceptual distance measures. Most concepts in the middle to high sections of the hierarchical network, being spatially close to each other, would therefore be deemed to be conceptually similar to each other. In order to account for the underlying architecture of the semantic network, Lee *et al*. (1993) proposed a new method called Knowledge-Based Extended Boolean Model (kb-ebm) which is a document ranking method to calculate the conceptual distance or closeness between a Boolean query and a document by suggesting an effective weighting scheme for queries and documents. kb-ebm provides high quality document rankings by using term dependence information from IS-A hierarchies. Lee argued that the semantic distance model should handle weighted indexing schema and variable edge weights. To determine weights the structural characteristics of the semantic network are typically considered, such as local density network, depth of a node in a hierarchical, type of link and the strength of an edge link.

Yang and Powers (2005) presents a new model to measure semantic similarity in the taxonomy of Word Net. They proposed a new model to measure semantic similarity in the taxonomy of Word Net, based on a variation of edge-counting. They constructed Concept Graph base on Word Net and weight each connect type of edges and measured the similarity by taking account of both IS-A and PART-OF relationships in Word Net.

Depth-based methods: Sussna (1993) introduced a depth-relative scaling approach, based on the observation that siblings deep in a tree are more closely related than siblings higher in the tree. He defined two edges representing inverse relations for each edge in a taxonomy. The weight attached to each relation r is a value in the range [min_r; max_r]. The point in the range for a relation r from concept c_1 to c_2 depends on the number nr of edges of the same type, leaving c_1, which is denoted as the type specific fanout factor:

$$W(c_1 \rightarrow r\ c_2) = max_r\text{-}\{max_r\text{-}min_r/n_r(c_1)\} \quad (2)$$

The two inverse weights are averaged and scaled by depth d of the edge in the overall taxonomy. The distance between adjacent nodes c_1 and c_2 are computed as:

$$Dist_{sussna}\ (c_1, c_2) = w\ (c_{1\rightarrow r}\ c_2)+(c_{1\rightarrow r'}\ c_2)/2d \quad (3)$$

where r is the relation that holds between c_1 and c_2 and r' is its inverse. The semantic distance between two arbitrary concepts c_1 and c_2 is computed as the sum of distances between the pairs of adjacent concepts along the shortest path connecting c_1 and c_2. In the overall computation of the distance between two adjacent nodes in the network, the two inverse weights for an edge connecting the two nodes are averaged and divided by the depth of the edge within the overall taxonomy.

Wu and Palmer (1994) proposed a measure of semantic similarity in their paper on the semantic representation of verbs in computer systems and its impact on lexical selection problems in machine translation. Wu and Palmer defined their conceptual similarity between a pair of concepts c_1 and c_2 based on position of concepts $c_1$and c_2 in the taxonomy relatively to the position of the most specific common concept c. As there may be multiple parents for each concept, two concepts can share parents by multiple paths. The most specific common concept c is the common parent related with the minimum number of IS-A links with concepts c_1 and c_2.

Leacock and Chodorow (1998) proposed an approach for measuring semantic similarity as the shortest path using IS-A hierarchies for nouns in Word Net. Since only noun hierarchies are considered, this measure is restricted to finding relatedness between noun concepts. The different noun hierarchies are combined into a single hierarchy by introducing a topmost node, subsuming all the topmost nodes in all the noun hierarchies. This ensures the existence of a path between all synsets in the taxonomy. So the fact that this measure takes into account the depth of the taxonomy in which the synsets are found means that the behavior of the measure is profoundly affected by the presence or absence of a unique root node. The proposed measure determines the semantic similarity between two synsets (concepts) by finding the shortest path and scaling by the depth of the taxonomy.

Another similarity measure is proposed by Li *et al*. (2003) which was intuitively and empirically derived, combines the shortest path length between two concepts and the depth in the taxonomy of the most specific common concept in a non-linear function. This method explores the determination of semantic similarity by a number of information sources, which consist of structural semantic information from a lexical taxonomy and information content from a corpus. To investigate how information sources could be used effectively, a variety of strategies for using various possible information sources are implemented. This measure combines information sources nonlinearly.

Alvarez and Lim (2007) presented a novel algorithm for scoring the Semantic Similarity (SSA) between words. Given two input words w_1 and w_2, SSA exploits their corresponding concepts, relationships and descriptive glosses available in WordNet in order to build a rooted weighted graph Gsim. The output score is calculated by exploring the concepts present in Gsim and selecting the minimal distance between any two concepts c_1 and c_2 of w_1 and w_2 respectively. In fact Li and Alvarez both consider that there are three main factors that should be taken into account in the edge-counting by Word Net:

- The shortest path length of C_1 and C_2
- The overlap between gloss of C_1 and C_2
- The depth of the Nearest Common Ancestor (NCA)

Information content based methods: These methods measure the difference in information content of the two terms as a function of their probability of occurrence in a corpus (Hliaoutakis *et al*., 2008). In this method rather than counting edges in the shortest path, the maximum information content of the least upper bound between two concepts are selected.

According to the idea of information content based measure, the more IC (Information Content) two concepts share, the more similar the concepts are. Statistical information from large corpora is used to estimate the information content of concepts. The idea of information content was introduced by Resnik (1995) In brief, information content of a concept measures the specificity or the generality of that concept, i.e. how specific to a topic the concept is. For example, a concept like sprinkler is a highly topical concept and would have high information content. On the other hand, a more general concept such as artifact would have much lower information content.

To find information content first the frequency of occurrence of every concept in a large corpus of text is computed. Every occurrence of a concept in the corpus adds to the frequency of the concept and to the frequency of every concept subsuming the concept encountered. We note that by this process the root node includes the frequency count of every concept in the taxonomy. It is incremented for every occurrence of every concept in the taxonomy.

If there is a common ancestor concept c of two concepts and p(c) stands for use frequency of c, the IC of c is as formula (4):

$$IC(c) = -\log p(c) \tag{4}$$

In a Concept Tree (CT) of ontology, leaf nodes have the most IC and it decrease with the depth increase. In another word, the more abstract a concept is, the less IC it has. The frequency of hypernyms is the sum of frequency of all its hyponyms and its own. So if a global root exists, its IC will be zero.

Resnik defines the semantic relatedness of two concepts as the amount of information they share in common. He goes on to elaborate that the quantity of information common to two concepts is equal to the information content of their lowest common subsumer (lcs) the lowest node in the hierarchy that subsumes both concepts:

$$Sim_{Resnik}(c_1,c_2) = IC(lcs(c_1,c_2)) \tag{5}$$

where IC determines the information content of a concept and lcs (c_1, c_2) finds the lowest common subsuming concept of concept c_1 and c_2.

The Resnik measure depends completely upon the information content of the lowest common subsumer of the two concepts whose relatedness we wish to measure. It takes no account of the concepts themselves. This leads to somewhat “coarser” relatedness values. For example, the concept pair car and bicycle will have the same measure of semantic relatedness as the pair car and all terrain bicycle because both pairs of concepts have the same lowest common subsumer. This measure has a lower bound of 0 and no upper bound.

Lin (1993) measure uses both the amount of information needed to state the commonality of two terms and the information needed to fully describe these terms. However, the Resnik measure depends solely on the information content of the shared parents and there are only as many discrete scores as there are ontology terms. By using the information content of both the compared terms and the shared parent the number of discrete scores is quadratic in the number of terms appearing in the ontology (Lord *et al*., 2003), thus augmenting the probability to have different scores for different pairs of terms. Consequently, using this measure to compare the terms of an ontology can have a better ranking of similarity than the Resnik measure.

A measure introduced by Jiang and Conrath (1997) addresses the limitations of the Resnik measure. Jiang combines the edge-counting based and IC-based measure and also takes edge weights into account. An edge weight is determined by several factors, such as depth of nodes, total edge count linking the nodes (local density), average local density of the hierarchy and edge type. Jiang and Conrath (1997) approach uses the notion of information content in the form of the conditional probability of encountering an instance of a child-synset given an instance of a parent-synset. Thus the information content of the two nodes, as well as that of their most specific subsumer, plays a part. Notice that this formula measures semantic distance, the inverse of similarity:

$$simjcn(c_1,c_2) = IC(c_1)+IC(c_2)-(2.IC(lcs(c_1,c_2))) \tag{6}$$

Lord *et al*. (2003) investigated the use of ontological annotation to measure the similarities in knowledge content or semantic similarity between entries in a data resource. They found that the semantic similarity calculated from annotations correlates well with the sequence similarity. After that many new approaches have been proposed specifically for measuring the semantic similarity on Gene Ontology.

Wang *et al*. (2006) considers both graph-based measure and IC-based measure compute the similarity by the hierarchy information of ontology only, but ignore a lot of important information, such as concepts' property. So he proposes the HIC-AIC algorithm. It takes the hierarchy and attribute of the ontology into account and combines the two measures. The HIC stands for the IC of hierarchy and the IC of NCA is used to measure it. AIC stands for the IC of inherited attribute. It is calculated by the IC of inherited properties.

Feature-based methods: Feature based methods measure the similarity between two terms as a function of their properties (e.g., their definitions or "glosses" in WordNet or "scope notes" in MeSH) or based on their relationships to other similar terms in the taxonomy. Common features tend to increase the similarity and (conversely) non-common features tend to diminish the similarity of two concepts (Tversky, 1977; Hliaoutakis *et al*., 2008).

However, the features of a term contain valuable information concerning knowledge about the term. This measure considers also the features of terms in order to compute similarity between different concept, while it ignores the position of the terms in the taxonomy and the information content of the term.

Tversky (1997) proposed a measure which based on the description sets of the terms. It is supposed that each term is described by a set of words indicating its properties or features. Then, the more common characteristics two terms have and the less non-common characteristics they have, the more similar the terms are.

Cross ontology: Semantic similarity measures can also be distinguished between Single Ontology and Cross ontology. Single Ontology similarity methods, which assume that the terms which are compared are from the same ontology (e.g., MeSH). Cross ontology measures that are capable of comparing terms from different ontologies. Because the structure and information content between different ontologies cannot be compared directly, cross ontology approaches usually call for hybrid or feature based measures. For example, two terms are similar if they have similar spelling or definition or they are related with other terms which are similar. Notice also that cross ontology methods may also be used to measure semantic similarity between terms from the same ontology (Hliaoutakis, 2005).

Rodriguez and Egenhofer (2003) proposed a framework for comparing terms stemming from the same or from different ontologies. The similarity function determines similar entity classes by using matching methods over synonym sets, semantic neighborhoods and distinguishing features that are further classified into parts, functions and attributes (e.g. considering the term college, a function is to educate, its parts may be roof and floor and other attributes can be architectural properties) (Hliaoutakis, 2005). The similarity between terms is computed as a weighted sum of similarities between synonym sets (synsets), features and terms neighborhoods.

Euripides *et al*. (2006) authors suggested X-Similarity, as a novel cross-ontology similarity method to computing the semantic similarity between natural language terms (using WordNet as the underlying reference ontology) and between medical terms (using the MeSH ontology of medical and biomedical terms). Since the most popular semantic similarity methods are implemented and evaluated using WordNet and MeSH. The focus of this method is on cross ontology methods which are capable of computing the semantic similarity between terms stemming from different ontologies (WordNet and MeSH in this method).

COMPARISON

Several semantic similarity measures for computing the conceptual were examined. Edge counting and information content methods work by exploiting structure information (i.e., position of terms). Evaluation of Edge Counting, Information Content, Feature based and Hybrid semantic similarity methods on WordNet shows that Information Content methods perform very well and close to the upper bound suggested by Resnik. Methods that consider the positions of the terms in the hierarchy (Li *et al*., 2003) perform better than plain path length methods. Methods exploiting the properties (i.e., structure and information content) of the underlying hierarchy perform better than Hybrid and Feature based methods, which do not fully exploit this information. However, Hybrid and Feature based methods are mainly targeted towards cross ontology similarity applications where edge counting and information content methods do not apply.

Comparison with human judgments is the ideal way to evaluate a measure of similarity or semantic relatedness. There are lots of discusses about the results of comparison of the measures to human similarity judgments. The first human similarity judgment was done by Rubinstein and Goodenough (1965) using two groups totaling 51 subjects to perform synonymy judgments on 65 pairs of nouns and this in turn been the basis of the comparison of similarity measures. Miller and Charles (1991) repeated Rubinstein and Goodenough (1965) original experiment, they used a subset of 30 noun pairs from the original list of 65 pairs, where ten pairs were from the high level of synonymy, ten from the middle

Table 1: Comparison between similarity measures (Euripides *et al*., 2006; Yang and Powers, 2005)

Measure	Increase with commonality	Decrease with difference	Symmetric property	Position in hierarchy	Normalized in [0, 1]
Rada	yes	yes	yes	no	yes
Hirst	no	yes	no	yes	no
Wu	yes	yes	yes	yes	yes
Leacock	no	yes	yes	yes	no
Li	yes	yes	yes	yes	yes
Jiang	yes	yes	yes	yes	no
Resnik	yes	no	yes	yes	no
Lord	yes	no	yes	yes	yes
Richardson	yes	yes	yes	yes	yes
Rodriguez	yes	yes	no	no	yes
Tversky	yes	yes	no	no	yes
Lin	yes	yes	yes	yes	yes

level and ten from the low level. The correlation of these experiments was 0.97. The correlations for the (Resnik, 1999; Jiang and Cornath, 1997; Hirst and St-Onge, 1998; leacock and Chodorow, 1998) are available in (Butanisky and Hirst, 1999; Budanitsky, 1999).

The key properties of the similarity measures presented are summarized in Table 1.

CONCLUSION

Several semantic similarity methods for computing the conceptual similarity between natural language concepts in an ontology and between ontologies were discussed. The experimental results in the earlier researches indicate that it is possible for these methods to provide better correlation values with human judgment of similarity.

Distances between concepts in a hierarchy are not evenly. distributed and network density, node depth, type of link and link strength are the determining factors of the distances, An important observation and a desirable property of most semantic similarity methods is that they assign higher similarity to terms which are close together (in terms of path length) and lower in the hierarchy (more specific terms), than to terms which are equally close together but higher in the hierarchy (more general terms).

REFERENCES

Alvarez, M.A. and S. Lim, 2007. A Graph Modeling of Semantic Similarity between Words. Proceeding International Conference on Semantic Computing, pp: 355-362.

Butanisky, A. and G. Hirst, 1999. Semantic Distance in Word Net: an Experimental, Application-oriented Evaluation of Five Measures. Workshop on Word Net and Other Lexical Resources, 2nd meeting of the North American Chapter of the Association for Computational Linguistics, Pittsburgh, PA.

Budanitsky, A., 1999. Lexical Semantic Relatedness and its Application in Natural Language Processing. Technical Report, University of Toronto, Canada.

Euripides, G., M. Petrakis, G. Varelas, A. Hliaoutakis and P. Raftopoulou, 2006. X-Similarity: Computing Semantic Similarity between Concepts from Different Ontologies. JDIM, 4(4): 233-237.

Hirst, G. and D. St-Onge, 1998. Lexical Chains as Representations of Context for the Detection and Correction of Malapropisms. In Proceedings of Fellbaum, pp: 305-332.

Hliaoutakis, A., 2005. Semantic Similarity Measures in MESH Ontology and their application to Information Retrieval on Medline. M.S. Thesis, Technical University of Crete, Greek.

Hliaoutakis, A., G. Varelas, E. Voutsakis, G. Euripides, M. Petrakis and E. Milios, 2008. Information Retrieval by Semantic Similarity Dept. of Electronic and Computer Engineering Technical University of Crete (TUC) Chania, Crete, Greece.

Jiang, J. and W. Conrath, 1997. Semantic Similarity Based On Corpus Statistics and Lexical Taxonomy. In: Proceedings of International Conference Research on Computational Linguistics.

Leacock, C. and M. Chodorow, 1998. Combining local context and Word Net Similarity for Word Sense Identification. In: Christiane, F., (Ed.), Word Net: An Electronic Lexical Database. MIT Press, Cambridge, MA.

Lee, J., M. Kim and Y. Lee, 1993. Information retrieval based on conceptual distance in is-a hierarchies. J. Doc., 49(2): 188-207.

Li, Y., Z.A. Bandar and A.D. McLean, 2003. An approach for measuring semantic similarity between words using multiple information sources. IEEE T. Knowl. Data En., 15: 871-882.

Lin, D., 1993. Principle-Based Parsing without over Generation. In: Annual Meeting of the Association for Computational Linguistics (ACL'93), Columbus, Ohio, pp: 112-120.

Lord, P.W., R.D. Stevens, A. Brass and C. Goble, 2003. Investigating semantic similarity measures across the Gene Ontology: The relationship between sequence and annotation. Bioinformatics, 19: 1275-1283.

Miller, G.A. and W.G. Charles, 1991. Contextual correlates of semantic similarity. Lang. Cognitive Proc., 6(1): 1-28.

Rada, R. and E. Bicknell, 1989. Ranking documents with a thesaurus. JASIS, 40(5): 304-310.

Rada, R., H. Mili, E. Bicknell and M. Blettner, 1989. Development and application of a metric on semantic nets. IEEE T. Syst. Man Cyb., 19(1): 17-30.

Resnik, P., 1995. Using information content to evaluate semantic similarity in taxonomy. In: Proceedings of the 14th International Joint Conference on Artificial Intelligence, Montreal.

Resnik, P., 1999. Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language. J. Artif. Intell. Res., 11: 95-130.

Rodriguez M.A. and M.J. Egenhofer, 2003. Determining Semantic Similarity among Entity Classes from Different Ontologies. IEEE T. Knowl. Data En., 15(2): 442-456.

Rubinstein, H. and J.B. Goodenough, 1965. Contextual correlates of synonymy. Commun. ACM, 8(10): 627-633.

Sussna, M., 1993. Word Sense Disambiguation for Free-text Indexing Using a Massive Semantic Network. Proceedings of CIKM.

Tversky, A., 1977. Features of similarity. Psychol. Rev., 84(4): 327-352

Wang, G., Y. Wang and M. Guo, 2006. An Ontology-Based Method for Similarity Caculation of Concepts in the Semantic Web. In: Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, pp: 1538-1542.

Wang, T. and G. Hirst, 2011. Refining the Notions of Depth and Density in Word Net-based Semantic Similarity Measures. Department of Computer Science University of Toronto, Canada.

Wu, Z. and M. Palmer, 1994. Verb Semantics and Lexical Selection. 32nd Annual Meeting of the Association for Computational Linguistics, pp: 133-138.

Yang, D. and D.M.W. Powers, 2005. Measuring Semantic Similarity in the Taxonomy of Word Net. In: Proc. Twenty-Eighth Australasian Computer Science Conference (ACSC2005), pp: 315-322.