Fuzzy Semantic Distance Measures Between Ontological Concepts

Valerie Cross
Department of Computer Science and Systems Analysis,
Miami University
Oxford, OH, USA
crossv@muohio.edu

Abstract **- Recently, an emphasis has been placed on the use of ontologies for representing application domain knowledge. Determining a degree or measure of semantic similarity, semantic distance, or semantic relatedness between concepts from different systems or domains, is becoming an increasingly important task. This paper presents a brief overview of such measures between concepts within ontological representations and provides several examples of such measures found in the research literature. These measures are then examined within the framework of fuzzy set similarity measures. The use of a semantic similarity measure between elements that are part of a domain for which an ontological structure exists is explored in order to extend standard fuzzy set compatibility measures.**

I. INTRODUCTION

Determining a degree or measure of semantic similarity, or its inverse semantic distance, between concepts from different systems or domains, is becoming an increasingly important task due to growing access to heterogeneous and independent data repositories [1]. Much of the focus has been on information exchange such as models and information about models. An application domain of information interchange requires that each system have a representation of its knowledge. Recently, an emphasis has been placed on the use of ontologies for representing application domain knowledge. One of the most common and simply stated definitions is that an ontology is a specification of a shared conceptualisation [2]. An ontology specifies a shared vocabulary used to model a domain of interest. This vocabulary describes the type of objects and/or concepts that exist, their properties and relations. Standard relations such as is-a, part-of, and instance-of have predefined semantics.

Ontologies have been developed for many purposes [3]. In software systems, they provide reusability and information sharing. In information retrieval the search operation may use an ontology as metadata to help direct the information retrieval to more relevant sources. Researchers in areas such e-Commerce or geographical information systems are developing global standardized ontologies. But most agree that it is not feasible for each discipline or community to standardize. Information exchange technology should foster knowledge exchange by providing tools to enable semantic interoperability. Interoperability is established by discovering semantically appropriate mappings between different and independent ontologies.

The methods of representing an ontology are diverse and depend on the required level of detail and logic for the problem domain. In practice, a thesaurus, a simple concept hierarchy, a semantic net, a frame system, or a logical model may represent an ontology. For example, WorldNet, a terminological ontology, is a collection of categories organized by a partial order that is induced by inclusion [4]. A much more detailed ontology, Cyc [5] is an axiomatized ontology whose categories are distinguished by axioms and whose definitions are stated in logic. In addition to the diversity of the ontological representation, a variety of methods for measuring semantic similarity and semantic distance have been proposed. The main approaches are based on distance within an ontological structure [6,7], concept information content [8], concept feature matching [9], and a combination of these approaches [10, 11].

This paper presents a brief overview of semantic similarity and semantic distance measures and then provides specific examples from these approaches in order to discuss the issues when applying these measures with respect to matching ontological concepts and then overall ontology matching. Recently, researchers have begun to propose the use of fuzzy ontologies for information retrieval [12, 13] and for merging multiple ontologies in order to relate knowledge developed by different users [14]. Specific examples of semantic similarity and distance measures for use within fuzzy ontologies are proposed. Issues and problems with extending other semantic similarity and semantic distance measures for fuzzy ontologies are examined.

The presentation begins with Section 2 describing a common meta-model for representing ontologies and WordNet, a frequently used ontology for experiments with semantic similarity. Section 3 overviews several approaches for measuring semantic relatedness, similarity or distance between concepts within an ontology and attempts to categorize these approaches. Section 4 examines how fuzzy set similarity measures and aggregation can be used to extend the process of ontological concept matching. Section 5 presents conclusions and discusses future work.

II. ONTOLOGICAL REPRESENTATION

The term ontology, according to Webster's dictionary, means a "particular theory about the nature of being or the kinds of existents." Although first used in the area of philosophy, the term ontology has been used by researchers in a variety of areas such as artificial intelligence (AI), information retrieval (IR), database theory, linguistics, and eCommerce.

The methods of representing an ontology are diverse and depend on the required level of detail and logic needed by the problem domain. In practice, a thesaurus, a simple concept hierarchy, a semantic net, a frame system, or a logical model may represent an ontology. For example, WordNet , a terminological ontology, is a collection of categories organized by a partial order that is induced by inclusion [15]. A much more detailed ontology, Cyc [16] is an axiomatized ontology whose categories are distinguished by axioms and whose definitions are stated in logic.

Numerous languages for representing ontologies have been proposed. These languages differ not only in the expressiveness but also in the level of formality. Because the integration of and mapping between ontologies encoded in different languages is a difficult challenge [17], many researchers [18] investigating semantic mapping between ontologies, assume a common frame-based knowledge model designed to be compatible with OKBC [19]. This model serves as a generic knowledge representation compatible with many existing knowledge-representation systems.

The main components of an OKBC-compliant knowledge model are classes, slots (either for a relationship or an attribute in object-oriented terminology), facets and instances. A class is a collection of objects described by identical properties. Classes are organized into taxonomy or a specialization and generalization hierarchy, also referred to as a subclass-superclass hierarchy. The superclass represents a generalization of its subclasses, the subclass, a specialization of its superclass. Slots are associated with each class and are inherited by the subclasses. Slots (also known as properties) are named binary relations between a class and either another class or a primitive type (such as a string or a number). Facets constrain the values taken on by slots, for example, the minimum or maximum value of a slot. An actual member of a class is referred to as an instance of the class.

In this discussion, the detail level of the ontology that is assumed is a concept hierarchy. The other aspects of an OKBC compliant model such as slots or features internal to a concept are not the focus of this paper although they could be added in to enhance the performance of a semantic distance or similarity measure between concepts. The primary ontology that has been the basis for much of the research in semantic similarity is WordNet. Princeton University developed WordNet as a large-scale lexical database of English. Freely available, this electronic dictionary organizes nouns, verbs, adjectives and adverbs into synonym sets (synsets), each representing one underlying lexical concept.

WordNet is structured as a semantic network with sets of synonymous terms, or synsets, constituting its basic organization. The synsets differentiate word senses from each other and underlie lexical concepts. For example, the word "address" corresponds to two different lexical concepts, one related to making a speech and the other to designating a location. These lexical concepts are linked by a variety of relations such as hypernym (is-a), hyponym (subsumes), antonym, holonym, (has-a) and meronym (part-of) connections.

. The noun portion of WordNet is the most developed part of the network and within it the subsumption hierarchy (hypernymy/hyponym) makes up over 80% of the links. The top of the noun hierarchy contains 11 different abstract concepts that are the separate roots termed unique beginners, for example, entity ('something having concrete existence; living or nonliving') or event ('something that happens at a given place and time'). Many of the semantic distance measures were used only on only the noun part of WordNet and its subsumption hierarchy.

III. Semantic Similarity/Distance/Relatedness

Determining a measure of semantic relatedness between concepts within the same ontology or distinct ontologies is a need that pervades the exchange of information over the Semantic Web. Much research has focused on the use of ontologies for representing domain knowledge and for exchanging information between agents. Effective communication between agents using different ontologies, however, requires determining the semantic interoperability, i.e., the agreement between the two agents' ontologies [20]. In this paper the focus is not on overall matching of two different ontologies, but determining a degree of semantic relatedness, similarity or distance between the concepts.

Three different terms *similarity, distance* and *relatedness* are sometimes seen used interchangeably in the research literature. These terms however, are not identical. To distinguish between similarity and relatedness, the example in [21] illustrates that similarity is a special case of relatedness. The terms *car* and *gasoline* appear to be more closely related than the terms *car* and *bicycle,* even though *car* and *bicycle* are more similar. This one example shows one kind of relatedness based on a functional relationship such as "car uses gasoline." There are numerous other kinds of semantic relatedness based on the type of relationship between concepts such as subsumption (e.g., vehicle-car) and meronymy (e.g., car-wheel).

The term semantic distance presents even more difficulty when trying to determine its association with the other two. Much of the research literature supports the view that distance measures the opposite of similarity. Semantic distance, however, could be used with respect to distance between related concepts and distance between similar concepts. In this paper semantic distance is used for the opposite of both semantic similarity and semantic relatedness. The context should provide the basis for the correct interpretation.

A. *Network Distance Models*

Early research focused on using word ontologies to improve information retrieval. One of the most natural approaches [22] to determine semantic similarity in an ontology is to use its graphical representation and measure the distance between the nodes corresponding to the words or concepts being compared. The number of edges in the shortest path between the two concepts measures the distance between them. The shorter the distance the more similar the concepts are semantically.

One of the major and intuitively obvious arguments against using the edge count distance in measuring conceptual distance is the underlying assumption that edges or links between concepts represent uniform distances [23]. In most taxonomic ontologies, concepts that are higher in the hierarchy are more general than those lower in the hierarchy. An edge count of one between two general concepts naturally implies a larger distance than that between two more specific concepts. The example given in [23] illustrates this problem. The distance between *plant* and *animal* is 2 in WordNet since their common parent, is *living thing*. The distance between *zebra* and *horse* is also 2 since their common parent is *equine*. Intuitively, one would judge zebra and horse more closely related than plant and animal. Solely counting links between nodes is not sufficient.

To overcome the limitation of simple edge counting, the edges were weighted to reflect the difference in edge distances. Earlier approaches [22] [24] [25], hand weighted each edge. Since this approach is not practical for very large ontologies, others proposed methods of automatically weighting each link [23]. Based on the earlier criticism for simple edge counting, the automatic process uses several pieces of information about the edge in determining its weight: the depth, the density of edges at that depth, and the strength of connotation between parent and child nodes. The weights are reduced as one goes farther down the network since conceptual distance shrinks. The weight also is reduced in a dense part of the network since edges in a dense part are considered to represent smaller conceptual distances. Several measures that improve on the original edge count approach are presented.

One of the simplest adjustments made [26] is to scale the minimum edge count distance between c1 and c2 by the maximum depth D of a taxonomic hierarchy, i.e., they only use hyponymy, *is-a,* type links between concepts.

$$sim_{LC}(c1, c2) = \max[-\log(\min_{c1,c2}[len(c1, c2)]/2 * D)] \quad (1)$$

Another proposal for conceptual similarity between a pair of concepts c1 and c2 [27] scales the similarity based on the depth of the least common superconcept c3 between c1 and c2.

$$sim_{WP}(c1; c2) = 2 * N3 / (N1 + N2 + 2*N3) \quad (2)$$

where N1 is the length (in number of nodes) of the path from c1 to c3, N2 is the length of the path from c2 to c3, and N3 is the length of the path from c3 to the root of the hierarchy, i.e., the global depth in the hierarchy. Conceptual similarity can be converted to conceptual distance as

$$dist_{WP}(c1; c2) = 1 - sim_{WP}(c1; c2) = (N1 + N2)/(N1 + N2 + 2* N3) \quad (3)$$

In this equation, it is easy to see how an increase of depth decreases the distance between the two concepts.

Another proposal [28] incorporates all relation types in WordNet and thus is considered a measure of semantic relatedness. It also incorporates the direction of the link between the two nodes. The directions of links on the same path may vary (among horizontal (antonymy), upward (hyponymy and meronymy), and downward (hypernymy and holonymy)). Two concepts are semantically related if they are connected by a path that is not longer than an arbitrary fixed constant C and whose direction does not change too often (where d represents the number of changes in path direction and k is another constant.)

$$rel_{HS}(c1, c2) = C - \text{path length}(c1, c2) - k * d \quad (4)$$

A more sophisticated modification [29] also employs the different kinds of linking relationship within WordNet. Each edge maps to two inverse relations. Each type of relations *r* has a weight range between its own min_r and max_r. The actual value in that range for *r* depends on $n_r(X)$, the number of relations of type r leaving node X This value referred to as type specific fanout (TSF) factor incorporates the dilution of the strength of connotation between a source and target node as a function of the number of like relations that the source node has. This factor reflects that asymmetry might exist between the two nodes so that the strength of connotation in one direction differs from that in the other direction. The weight for the relation r between nodes X and Y is calculated as

$$w(XrY) = max_r - (max_r - min_r) / n_r(X), \quad (5)$$

and similarly for the inverse relation r', w(Yr'X). The two weights for an edge are averaged. The average is divided by the depth *d* of the edge within the overall network to produce the distance or weight between the concepts X and Y as

$$w(X,Y) = (w(XrY) + w(Yr'X)) / 2d. \quad (6)$$

The relative scaling by this depth is based on the intuition that siblings deep in a network are more closely related than only-siblings higher up. The semantic distance between two arbitrary nodes c1 and c2 is then computed as the sum of the distances between the pairs of adjacent nodes along the shortest path connecting c1 and c2.

B. Information Theoretic Models

The information-theoretic models of semantic similarity add to the information already present in the network by using a qualitatively different, knowledge source. The groundwork for much of this research [30,31] is founded on the insight that conceptual similarity between two concepts c1 and c2 may be judged by the degree to which they share information. The more information they share then the more similar they are. Different methods have been used to approximate that information content.

In [21], the information shared by two concepts c1 and c2 is approximated by the information content of the lowest superconcept c3 that subsumes them in the hierarchy. Another way of describing c3 is the most specific common ancestor of c1 and c2. Based on standard information theory, the information content of c3 is quantified as $-\log p(c3)$ where p(c3) is probability of encountering an instance of c3 in a specific corpus. The probability is based on using the corpus to perform a frequency count of all the words in the synset of concept c3 and in any synset of a descendent concept. The similarity between c1 and c2 is given as

$$sim_R(c1, c2) = -\log p(c3). \tag{7}$$

Note that this is intuitively satisfying since the higher the position of the c3 in the taxonomy, the more abstract c3 is, therefore, the lower the similarity between c1 and c2.

While most similarity measures increase with commonality and decrease with difference, sim_R only takes commonality into account. The semantic similarity between two concepts proposed in [32] takes both commonality and difference into account. It is argued that the similarity between two concepts is not about the concepts but about the instances of the concepts. As an example, if one says that a tree is similar to a bush, the set of all trees, is not being compared to the set of all bushes. Instead, a generic tree is being compared to a generic bush. The amount of information contained in $x1 \in c1$ and $x2 \in c2$ is

$$-\log p(c1) - \log p(c2)$$

where p(c1) is the probability that a randomly selected object x1 would belong to c1. If concept c3 is the most specific concept that subsumes both c1 and c2, then

$$sim_L(c1, c2) = 2 * \log p(c3) / (\log p(c1) + \log p(c2)) \tag{8}$$

This approach uses the shared information content in the subsuming concept and normalizes with sum of the unshared information content of both concepts.

Finally one method [10] began by trying to combine the network distance approach with information theoretic approach. They envisioned using the corpus statistics as a corrective factor to fix the problems with the weighted edge counting approaches and developed a general formula for the weight of a link between a child-concept c_c and its parent-concept c_p in a hierarchy. This formula incorporates the ideas of node depth, local density similar to the TSF and the link type. These ideas parallel the approach in [29].

They studied the roles of the density and depth components and concluded that they are not major factors in the overall edge weight and shifted their focus to the link-strength factor. The link-strength factor uses information content but in the form of conditional probability, i.e., the probability of encountering an instance of a child concept c1 given an instance of a parent concept c3. If the probabilities are assigned as in [21], then the distance between concepts c1 and c2 with concept c3 as the most specific concept that subsumes both is

$$dist_{JC}(c1, c2) = 2 * \log p(c3) - (\log p(c1) + \log p(c2)) \tag{9}$$

Note that this measure is but a different arithmetic combination of the same terms in Equation 8.

IV. Fuzzy Semantic Measures

In the previous section numerous measures have been proposed to determine the semantic similarity, relatedness or distance between concepts within an ontology. These measures have been classified into two primary approaches. The first uses distance with respect to path length between the two concepts within the link structure (the main link has been the IS-A taxonomic link). The second is based on shared information content of the most specific common ancestor of the two concepts. These two approaches can be related through the Wu and Palmer measure and the Lin measure. If the conditional probability of child concept k given its immediate parent concept p, $P(k \mid p)$ is the same for all child-parent pairs of concepts, then the two semantic similarity measures agree [10].

Although numerous researchers have emphasized that these measures are similarity measures based on distances within a taxonomy, these two measures may be viewed as a variation of Tversky's [9] parameterized ratio model of similarity:

$$S(X, Y) = f(X \cap Y) / [f(X \cap Y) + \alpha * f(X - Y) + \beta * f(Y - X)] \tag{10}$$

Let X represent the set of IS-A links from the root to concept c1 and Y represent the set of IS-A links from the root to concept c2. With $\alpha=\beta=1/2$, S becomes Dice's coefficient [dice] of similarity:

$$S_{dice}(X, Y) = 2*f(X \cap Y) / [f(X) + f(Y)] \tag{11}$$

With the function f simply the cardinality of the sets, the value of $f(X \cap Y)$ represents the cardinality of the intersection of the IS-A links on the path from the root to c1 and the IS-A links on the path from the root to c2, which is equivalent to N3, the path length from the root to c3, their most specific or lowest common ancestor. Here the assumption is that the weights on the links are 1. The value for f(X) is N1+N3 and for f(Y) is N2 + N3 so that equation 11 becomes equivalent to the Wu and Palmer similarity measure given in Equation 2. This semantic similarity measure may be generalized to a fuzzy semantic similarity measure if the weights for the IS-A links are replaced by membership degrees indicating the strength of the relationships between the parent and child concepts.

In the previous discussion a measure of fuzzy semantic similarity between two concepts from the same ontology is presented. From another viewpoint, the structure of the ontology could also be used to help define new fuzzy set similarity measures capitalizing on any of the fuzzy semantic similarity measures referred to as sim_{sem}. in the following examples.

If the elements of the two fuzzy sets being compared are located within an ontological structure then existing fuzzy set compatibility measures [33] can be extended to use sim_{sem} as a proximity relation between elements of the universe of discourse. For example, the compatibility between two fuzzy sets F_1 and F_2 over the universe of discourse U, can be measured using the standard sup-min set theoretic compatibility measure but enhanced with the sim_{sem}

$$S_{sup\text{-}min}(F_1, F_2) = \sup_{u \in U} \min(\mu_{F1}(u), \mu_{F2}(u)) \quad (12)$$

becomes

$$S_{sup\text{-}min}(F_1, F_2) = \sup_{(v,w) \in U \times U} \min(sim_{sem}(v, w), \mu_{F1}(v), \mu_{F2}(w)) \quad (13)$$

Other categories of fuzzy set similarity measures based on vector models can also be extended. For example, the Bhattacharyya distance for measuring divergence between two fuzzy sets is

$$S_{cos}(F_1, F_2) = \Sigma(\mu_{F1}(u)\, \mu_{F2}(u))/[(\Sigma\mu_{F1}(u)^2)^{1/2}\, (\Sigma\mu_{F2}(u)^2)^{1/2}] \quad (15)$$

and is extended as

$$S_{cos}(F_1, F_2) = (\Sigma\Sigma\, sim_{sem}(v, w)\, \mu_{F1}(v), \mu_{F2}(w)) / [(\Sigma\mu_{F1}(v)^2)^{1/2}\, (\Sigma\mu_{F2}(w)^2)^{1/2}] \quad (16)$$

There are numerous other fuzzy set compatibility measures that can be extended by incorporating a fuzzy semantic similarity between elements in a universe of discourse for which an ontological structure exists.

V CONCLUSIONS

This paper investigates several methods of measuring semantic agreement, i.e., semantic similarity, semantic distance and semantic relatedness, between concepts in an ontology. This task is becoming extremely critical in the process of exchanging information on the WWW. These measures are classified into two primary approaches. The first uses distance with respect to path length between the two concepts within the link structure (the main link has been the IS-A taxonomic link). The second is based on shared information content of the most specific common ancestor of the two concepts. These two approaches can be viewed as a variation of Tversky's parameterized ratio model of similarity. The various fuzzy semantic similarity measures sim_{sem} are considered from a slightly different perspective. Fuzzy set compatibility measures can incorporate them to better assess agreement between fuzzy sets whose elements are related through an ontological domain structure.

REFERENCES

[1] A.Sheth,"Changing Focus on Interoperability in Information Systems:From System Syntax, Structure to Semantics," *Interoperating Geographic Information Systems,* M.Goodchild, M.Egenhofer, R.Fegeas,and C.Kottman,eds.,pp.5-30,1999.

[2] Gruber, T. R. "A Translation Approach to portable ontology Specifications", *Knowledge Acquisition,* 5(2) 1993, pp. 199-220.

[3] Uschold, M.; Gringer, M. "ONTOLOGIES: Principles, Methods, and Applications" *Knowledge Engineering Review.* Vol. 11. N. 2. June. 1996.

[4]G.Miller,R.Beckwith,C.Fellbaum,D.Gross,andK.Miller, "Introduction to WordNet:An On-Line Lexical Database,"Int 'l J. Lexicography,vol.3,pp.235-244,1990.

[5] D. Lenat and R. Guha, Building Large Knowledge Based Systems: Representation and Inference in the CYC Project. Reading, Mass.:Addison-Wesley Publishing Company, 1990.

[6] J.Lee,M.Kim,and Y.Lee,"Information Retrieval Based on Conceptual Distance in IS-A Hierarchies,"J.Documentation,vol.49, pp.188-207,1993.

[7] N. Foo, B. Garner, A. Rao, and E. Tsui, "Semantic Distance in Conceptual Graphs", in "Current Directions in Conceptual Structure Research", (eds) P Eklund, T Nagle, J Nagle and L Gerhotz, Ellis Horwood, 1992,

[8] O.Resnik,"Semantic Similarity in a Taxonomy:An Information-Based Measure and Its Application to Problems of Ambiguity and Natural Language,"J.Artificial Intelligence Research,vol.11,pp.95-130,1999.

[9] A.Tversky,"Features of Similarity," Psychological Rev.,vol.84, pp.327-352,1977.

[10] Jay J. Jiang and David W. Conrath. 1997. Semantic similarity based on corpus statistics and lexical taxonomy. In *Proceedings of International Conference on Research in Computational Linguistics*,Taiwan.

[11] M. Andrea Rodríguez and Max J. Egenhofer, "Determining Semantic Similarity Among Entity Classes from Different Ontologies", *IEEE Transactions on Knowledge and Data Engineering* , 15 (2): 442-456, 2003.

[12] V. Cross and C. R. Voss, "Fuzzy Ontologies for Multilingual Document Exploitation," Proceedings of the 1999 Conference of NAFIPS, June, 1999, New York City, pp. 392-397.

[13] Dwi H. Widyantoro, John Yen: A Fuzzy Ontology-based Abstract Search Engine and Its User Studies. FUZZ-IEEE 2001: 1291-1294

[14] Parry, D. (2004). A Fuzzy Ontology for Medical Document Retrieval. In Proc. Australasian Workshop on Data Mining and Web Intelligence (DMWI2004), Dunedin, New Zealand. CRPIT, **32**. Purvis, M., Ed. ACS. 121-126.

[15] Miller, George A. (1995). WordNet: a lexical database for English. *Communications of the ACM* 38 (11), 39 - 41.

[16] D. Lenat and R. Guha, Building Large Knowledge Based Systems: Representation and Inference in the CYC Project. Reading, Mass.:Addison-Wesley Publishing Company, 1990.

[17] A. Sheth, "Changing Focus on Interoperability in InformationSystems: From System, Syntax, Structure to Semantics," Interoperating Geographic Information Systems, M. Goodchild, M. Egenhofer, R. Fegeas, and C. Kottman, eds., pp. 5-30, 1999.

[18] N. Noy and M. Musen. *PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment* In *Proceedings of the Seventeenth National Conference on Artificial Intelligence (AAAI-2000)*, pp. 450-455, 2000.

[19] Chaudhri, V.K., Farquhar, A., Fikes, R., Karp, P.D. and Rice, J.P. (1998). OKBC: A Programmatic Foundation for Knowledge Base Interoperability. In: *Proceedings of theFifteenth National Conference on Artificial Intelligence (AAAI-98)*, Madison, Wisconsin, AAAI Press.

[20] Cross, Valerie "Uncertainty in the Automation of Ontology Matching", Proceedings of ISUMA, University of Maryland, October, 2003.

[21] Resnik P., Using information content to evaluate semantic similarity in a taxonomy. In *,Proceedings of the the International Joint Conference on Artificial Intelligence (IJCAI) 448-453. hD. dissertation at the University of Pennsylvania*. appears as Technical Report 93-42, November.

[22] Rada, Roy, Hafedh Mili, Ellen Bicknell, and Maria Blettner (1989). Development and Application of a Metric on Semantic Nets. *IEEE Transactions on Systems, Man and Cybernetics* 19: 17-30.

[23] Richardson, R., A. Smeaton, and J. Murphy *Using WordNet as a Knowledge Base for Measuring Semantic Similarity between Words*. Technical Report, Working paper CA-1294, School of Computer Applications, Dublin City University: Dublin, Ireland. 1994

[24] Kim, Y and Kim, J "A model of knowledge based information retrieval with hierarchical concept graph" Journal of Documentation, 46, 1990, `113-116.

[25] Lee, J and Kim, M. "Information Retrieval based on conceptual distance in a is-a hierarchy." Journal of Documentation, 49, 1993, `188-207.

[26] Claudia Leacock and Martin Chodorow. 1998. Combining local context and WordNet similarity for word sense identi-fication. In Fellbaum 1998, pp. 265–283.

[27] [Wu and Palmer, 1994] Zhibiao Wu and Martha Palmer. Verb semantics and lexical selection. In Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics, Las Cruces, New Mexico , pages 133-138, June 1994.

[28] Graeme Hirst and David St-Onge." Lexical chains as representations of context for the detection and correction of malapropisms." In Christiane Fellbaum, editor, WordNet: An Electronic Lexical Database, chapter 13, pages 305-332. The MIT Press, Cambridge, MA, 1998.

[29] Sussna M. "Word Sense Disambiguation for Free-text Indexing Using a Massive Semantic Network," Proceedings of the second international conference on Information and knowledge management Washington, D.C., pp: 67 - 74 1993.

[30] Resnik P. Selection and Information : A Class based Approach to Lexical Relationships, *PhD. dissertation at the University of Pennsylvania*. appears as Technical Report 93-42, November.

[31] Resnik, Philip "Semantic Classes and Syntactic Ambiguity." In *Proceedings of the ARPA HumanLanguage Technology Workshop*, Princeton, 1993.

[32] Dekang Lin. 1998. An information-theoretic definition of sim-ilarity.In *Proceedings of the 15th International Conference on Machine Learning*, Madison, WI.

[33] V. Cross and T. Sudkamp, Similarity and Compatibility in Fuzzy Set Theory: Assessment and Applications, New York: Physica-Verlag, ISBN 3-7908-1458, 2002