Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

A Cluster-Based Approach for Semantic Similarity in the Biomedical Domain

Hisham Al-Mubaid and Hoa A. Nguyen
University of Houston-Clear Lake, Houston, TX 77058, USA
{hisham, nguyenh3308}@uhcl.edu

Abstract—**We propose a new cluster-based semantic similarity/distance measure for the biomedical domain within the framework of *UMLS*. The proposed measure is based mainly on the cross-modified path length feature between the concept nodes, and two new features: (1) the common specificity of two concept nodes, and (2) the local granularity of the clusters. We also applied, for comparison purpose, five existing general English ontology-based similarity measures into the biomedical domain within *UMLS*. The proposed measure was evaluated relative to human experts' ratings, and compared with the existing techniques using two ontologies (*MeSH and SNOMED-CT*) in *UMLS*. The experimental results confirmed the efficiency of the proposed method, and showed that our similarity measure gives the best overall results of correlation with human ratings. We show, further, that using *MeSH* ontology produces better semantic correlations with human experts' scores than *SNOMED-CT* in all of the tested measures.**

I. Introduction

The natural language processing (NLP) literature includes a number of approaches to measure the semantic similarity, or more generally, *relatedness* between two terms or concepts [1-4,7-9,11-15,17]. Measures of semantic similarity and relatedness can be used in applications such as sense disambiguation, information extraction and retrieval, classification and ranking, detection of redundancy, and detection and correction of malapropisms. In this paper, we investigate the semantic similarity measures that are based on ontology structure and propose a new cluster-based semantic similarity/distance measure. We apply these measures into the biomedical domain. We evaluated the proposed measure along with other existing ontology-based measures using the *MeSH* and *SNOMED-CT* thesauruses/ontologies, within *UMLS* framework [5, 18]. The experimental results with two datasets of biomedical terms and two ontologies confirmed the efficiency of the proposed measure. The proposed measure (*Sem*) achieved the highest results of correlation with human scores in the two datasets using both ontologies.

UMLS: The Unified Medical Language System (*UMLS*) project started at the National Library of Medicine (NLM) in 1986 [5,18], with one of the objectives is to help interpret and understand medical meaning across systems [5]. It consists of three main knowledge sources: Metathesaurus (*MeSH*, *SNOMED-CT* thesauruses, etc.), Semantic Network, and SPECIALIST Lexicon & Lexical Tools [13].

MeSH: *MeSH*, stands for Medical Subject Headings, [5, 19], is one of the source vocabularies used in *UMLS. MeSH* includes about 15 high-level categories, and each category is divided into subcategories and assigned a letter: A for Anatomy, B for Organisms and C for Diseases, and so on.

SNOMED-CT: *SNOMED-CT*, stands for Systemized Nomenclature of Medicine Clinical Term [5, 6, 18], was included in *UMLS* in May 2004 (2004AA) [5]. SNOMED-CT is a comprehensive clinical terminology [6], and the current version contains more than 360,000 concepts, 975,000 synonyms and 1,450,000 relationships organized into 18 hierarchies.

In this paper, we use term "*concept node*" to refer to a concept class represented as a node in the ontology and that contains a set of synonymous concepts. The similarity between two concepts that belong to the same node reaches maximum, and the similarity of two concepts is the similarity of the two concept classes (nodes) containing them.

II. Semantic Similarity

In this paper, we focus only on those semantic similarity measures that use ontology as primary information source.

A. Ontology-Structure-Based Measures

Most of the semantic similarity measures that are based on the structure of ontology are actually based on path length (shortest path length) between two concept nodes, and/or depths of concept nodes in the *"IS A"* hierarchy tree. Some of the WordNet-based measures are: Path length [13], Wu & Palmer [17], Leacock & Chodorow [7], and Li et al. [8]. Choi & Kim [2] proposed a semantic similarity measure in their work for solving the problem of topic distillation and applied it on *yahoo* category tree.

B. Information-Based Measures

The information-based measures use information content (IC) of concept nodes derived from the ontology hierarchy structure and corpus statistics. In WordNet-based measures, some information-based measures are: Resnik [14], Jiang & Conrath [4] and Lin [9].

1-4244-0033-3/06/$20.00 ©2006 IEEE.

III. Method For Semantic Similarity

The primitive approach to find the similarity/distance between two concept nodes is to find the shortest path length between them on the hierarchy tree (Rada et al. [13]). Let us consider, for example, a fragment of ontology showing two clusters as in Figure 1. The first cluster, *A*, contains concepts a_i; and the second cluster, *B*, contains concepts b_i. Depth of cluster A is 4 and depth of cluster B is 3 (*by node counting)*. The path length between a_2 and a_6 is 3 using node counting. The path length between a_1 and b_1 is also 3. Thus, the similarity in these two cases is the same according to Path length measures. However, intuitively speaking, the similarity between a_1 and b_1 is less than the similarity between a_2 and a_6 as the latter two concepts (a_2 & a_6) lie at a lower level in the hierarchy and share more information. Thus, path length-based measures such as Path length [13] and Leacock & Chodorow [7] give the same similarity value for these two pairs while the measure of Wu & Palmer [17] uses the depth feature only. Therefore, we should take the *specificity* of concepts into account by using the *depths* of the concept nodes. The least common subsumer (LCS) of two concept nodes determines the common specificity of two concept nodes (*e.g.* $LCS(a_2,a_6)=a_1$ & $LCS(a_1,b_1)=r$), therefore we use LCS for computing *common specificity* of two concept nodes. Furthermore, local density such as link strength/weight also affects the similarity. One way of measuring the local density is using information content of concepts based on corpus statistic [4,8,9,15]. And since there is no standard corpus in biomedical domain, we use only ontology-based features as properties of semantic similarity. We also want to further examine the *local specificity* of a concept node by considering the cluster containing that concept node. The following example explains the effect of cluster on *local concept specificity*. Let us consider, for example, a fragment of ontology showing two clusters as in Figure 1. We define the specificity of a concept c in cluster C as follow:

$$\mathrm{spec}(c) = \frac{\mathrm{depth}(c)}{\mathrm{depthC}} \tag{1}$$

where depthC is the depth of cluster C, and spec(c) $\in[0,1]$. We notice that spec(c) = 1 when the concept c is a leaf node in the cluster C. Then, in Figure 1, the specificity of a_3 and b_3, is calculated as follow:

$\mathrm{spec}(a_3)=3/4= 0.75$

$\mathrm{spec}(b_3)=3/3= 1.00$

Thus, the specificity of b_3 (1.00) is more than specificity of a_3 (0.75), even though their depths are equal. Thus, b_3 has more specificity within its cluster than a_3 as it lies further down towards the bottom in its cluster. Therefore, we should take into account the local granularity of clusters as a feature that most existing measures that use ontology structure as primary information source do not take it into consideration.

IV. The Proposed Semantic Measure

Before discussing the details of the proposed approach, we present our rules and assumptions to be satisfied in the proposed measure.

A. Rules and Assumptions

We want to combine all the semantic features discussed above in one measure in an effective and logical way. We summarize our intuitive rules and assumptions in the following:

Rule R1: The semantic similarity scale system shows (reflects) the degree of similarity of pairs of concepts comparably in one cluster or in cross-cluster. This rule ensures that the mapping of cluster 1 to cluster 2 does not deteriorate the similarity scale of any cluster.

Rule R2: The semantic similarity must obey local cluster's similarity rule as follow:

- **Rule R2.1:** The shorter the distance between two concept nodes in the hierarchy tree, the more they are similar.
- **Rule R2.2**: Lower level pairs of concept nodes are semantically closer (*more similar*) than higher level pairs.
- **Rule R2.3:** The maximum similarity is reached when the two concept nodes are the same node in the hierarchy tree.

Before presenting the details of the proposed measure, we present and explain our assumptions about the semantic law function:

Assumption A1: Logarithm functions are the universal law of semantic distance.

Exponential-decay functions are universal law of stimulus generalization for psychological sciences [16]. We use logarithm (inverse of exponentiation) for semantic distance. We argue that non-linear combination approach is the optimum approach for combining semantic features. Rule R2.3 shows that when the two concept nodes are the same node (i.e. they are identical or synonymous), the semantic similarity must reach highest similarity regardless of other features, and so, we should use non-linear approach to combine the features. Therefore, we need another assumption.

Assumption A2: Non-linear function is the universal combination law of semantic similarity features.

B. New Feature: Common Specificity Feature

Besides the path length feature, we use in our measure the depth of concept nodes effectively to improve performance. The least common subsumer (LCS) node of two concepts C_1 and C_2 determines the *common specificity* of C_1 and C_2 in the cluster. So we measure the specificity of two concepts by finding the depth of their LCS node and then scaling this depth by the depth *D* of the cluster as follow:

$$\mathrm{CSpec}(C_1,C_2) = D - \mathrm{depth}(\mathrm{LCS}(C_1,C_2)) \tag{2}$$

where D is the depth of the cluster. Thus the $CSpec(C_1,C_2)$ feature determines the "*common specificity*" of two concepts in the cluster. The smaller the common specificity value of two concept nodes, the more they share information, and thus the more they are similar.

C. Single Cluster Similarity

In single cluster, the local granularity of the cluster is not considered as there is only one single cluster. We have two features to combine: Path length and the Common specificity given by Eq.(2). When the two concept nodes are the same node then path length will be 1 (using node counting), and so the semantic distance value must reach the minimum regardless of CSpec feature by rule R2.3 (recall the semantic *distance* is the inverse of semantic *similarity*). Therefore, we use product of semantic distance features for combination of features. By applying Rules R1, R2 and the two assumptions, the proposed measure for a single cluster is:

$$\mathrm{Sem}(C_1,C_2) = \log\left((\mathrm{Path}-1)^{\alpha} \times (\mathrm{CSpec})^{\beta} + k\right) \tag{3}$$

where $\alpha>0$ and $\beta>0$ are contribution factors of two features; k is a constant; LCS is the least common subsumer of two concepts; and Path is the path length of the shortest path between the two concept nodes. To insure the distance is positive and the combination is non-linear, k must be greater or equal to one ($k\geq1$). We use k=1 in our experiments. When two concept nodes have path length of 1 (Path=1) using node counting (*i.e.,* they are in the same node in the ontology), they have a semantic distance (Sem) equals to zero (i.e. maximum similarity) regardless of common specificity feature.

D. Cross-Cluster Semantic Similarity

In cross-cluster semantic similarity, to measure the similarity between two concepts (C_1 & C_2), there are four cases depending on the positions of the two concept nodes within the clusters of the ontology. Let us assume that the cluster that has the largest depth is the main cluster (call it the *primary* cluster) on which the semantic features from all other clusters will be scaled to this cluster's scale-level. Let us, further, call all other remaining cluster *secondary* clusters. Then, we have four cases as follows:

Case 1: *Similarity within the Primary Cluster:*

If the two concept nodes occur in the primary cluster then we treat this case as similarity within single cluster Eq.(3) discussed in section IV(C).

***Case 2*:** *Cross-Cluster Similarity:*

In this case, one of the two concept nodes belong to the primary cluster while the other is in a secondary cluster, and the LCS of two concept nodes is the *global root* node, which belongs to the two clusters. This technique does not affect the scale of the *CSpec* feature of the primary cluster. The common specificity is then given as:

$$\mathrm{CSpec}(C_1,C_2) = \mathrm{CSpec}_{primary} = D_{primary} - 1 \tag{4}$$

where $D_{primary}$ is the depth of the primary cluster. The root is the LCS of the two concept nodes in this case. The path between the two concept nodes passes through two clusters having different granularity degrees. The portion of the path length that belongs to the secondary cluster is in scale of granularity different from that of the primary cluster, and thus, we need to convert it (level it) into primary cluster scale-level as follows.

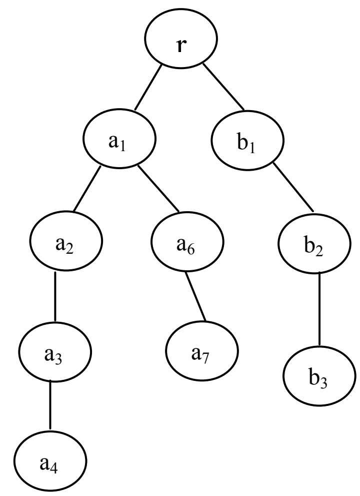

Figure 1. **A fragment of two clusters in ontology**

<u>The Cross-Cluster Path length Feature</u>**:** The path length between two concept nodes (c_1 & c_2) is computed by adding up the two shortest path lengths from the two nodes to their LCS node (*their LCS is the root*). For example, in Figure 1, for the two concept nodes (a_3, b_3), the LCS is the root *r*. So, we measure the path length between a_3 and b_3 as:

$$\mathrm{Path}(C_1,C_2) = d_1 + d_2 - 1 \tag{5}$$

such that: $d_1 = d(a_3, \mathrm{root})$ and $d_2 = d(b_3, \mathrm{root})$,
where $d(a_3, \mathrm{root})$ is the path length from the root *r* to node a_3 ; and similarly $d(b_3, \mathrm{root})$ is the path length from *r* to b_3. Notice that the root node is counted twice, so we subtract one in Eq.(5). We notice here that the densities or granularities of the two clusters are in different scales. Then, we want to scale the portion of the path length in the secondary cluster into the primary cluster's scale-level. The cluster containing a_3 has higher depth, and then it's the *primary* cluster, and the cluster containing b_3 is the *secondary*. The granularity rate of the *primary* cluster over the *secondary* cluster for the common specificity feature is:

$$\mathrm{CSpecRate} = \frac{D_1 - 1}{D_2 - 1} \tag{6}$$

where (D_1-1) and (D_2-1) are maximum common specificity values of the primary and secondary clusters respectively. The granularity rate, *PathRate*, of path length feature for the *primary* cluster over the *secondary* cluster is given by:

$$\mathrm{PathRate} = \frac{2D_1 - 1}{2D_2 - 1} \tag{7}$$

where $(2D_1\text{-}1)$ and $(2D_2\text{-}1)$ are maximum path values of any two nodes in the primary and secondary clusters respectively. Following Rule R1, we convert d_2 in Eq.(5) to the primary cluster as follows:

$$d'_2 = \text{PathRate} \times d_2 \quad (8)$$

This new path length d'_2 reflects path length of the second concept to the LCS relative to primary cluster's path length feature scale. Applying Eq.(8), we obtain path length between 2 concept nodes in primary cluster scale as follow:

$$\text{Path}(C_1, C_2) = d_1 + \text{PathRate} \times d_2 - 1 \quad (9)$$

$$\text{Path}(C_1, C_2) = d_1 + \frac{2D_1 - 1}{2D_2 - 1} \times d_2 - 1 \quad (10)$$

Finally, the semantic distance between two concept nodes is given as follow:

$$\text{CSpec}(C_1, C_2) = D_{primary} - 1 \quad (11)$$

$$\text{Sem}(C_1, C_2) = \log\left((\text{Path} - 1)^{\alpha} \times (\text{CSpec})^{\beta} + k\right) \quad (12)$$

Case 3: *Similarity within a Single Secondary Cluster*

The third case is when the two concept nodes are in a single secondary cluster. Then the semantic features, in this case, must be converted to primary cluster's scales for the two features, Path and CSpec, as follow:

$$\text{Path}(C_1, C_2) = \text{Path}(C_1, C_2)_{secondary} \times \text{PathRate} \quad (13)$$

$$\text{CSpec}(C_1, C_2) = \text{CSpec}(C_1, C_2)_{secondary} \times \text{CSpecRate} \quad (14)$$

$$\text{Sem}(C_1, C_2) = \log\left((\text{Path} - 1)^{\alpha} \times (\text{CSpec})^{\beta} + k\right) \quad (15)$$

where $\text{Path}(C_1, C_2)_{secondary}$ and $\text{CSpec}(C_1, C_2)_{secondary}$ are the Path and CSpec between C_1 & C_2 in the *secondary* cluster.

Case 4: *Similarity within Multiple Secondary Clusters*

The fourth case occurs when the two concept nodes are in two different secondary clusters. In this case, one of the secondary clusters acts *temporarily* as a *primary* cluster to calculate the semantic features (path and CSpec) using cross-cluster approach as in *Case 2* above. Then, the *semantic distance* is computed using *Case-3* to scale the features (again) to the scale-level of the primary cluster.

V. Experiments and Results

A. Datasets

There are no standard human rating datasets for semantic similarity in biomedical domain. To evaluate our methods, however, we used a (*published*) dataset of 30 concept pairs from Pedersen et al. [12], (we call it *Dataset 1*) which was annotated by 3 physicians and 9 medical index experts. Each pair was annotated on a 4-point scale: *practically synonymous, related, marginally related, and unrelated*. For space limitation, Table 1 contains only the first 10 pairs of this dataset. The average correlation between physicians is 0.68, and between experts is 0.78. Because the experts are more than the physicians, and the *agreement* between experts (0.78) is higher than the correlation between physicians (0.68), we can assume that the experts' rating scores are more reliable than the physicians' rating scores, and so we use experts' scores in our experiments.

Table 1. Dataset 1: first 10 medical term pairs with physicians' and experts' scores

Concept 1	Concept 2	Phys.	Expert
Renal failure	Kidney failure	4.0000	4.0000
Heart	Myocardium	3.3333	3.0000
Stroke	Infarct	3.0000	2.7778
Abortion	Miscarriage	3.0000	3.3333
Delusion	Schizophrenia	3.0000	2.2222
Congestive heart failure	Pulmonary edema	3.0000	1.4444
Metastasis	Adenocarcinoma	2.6667	1.7778
Calcification	Stenosis	2.6667	2.0000
Diarrhea	Stomach cramps	2.3333	1.3333
Mitral stenosis	Atrial fibrillation	2.3333	1.3333

Table 2. Dataset 2: first 10 medical term pairs with human scores

Concept 1	Concept 2	Human
Anemia	Appendicitis	0.031
Meningitis	Tricuspid Atresia	0.031
Sinusitis	Mental Retardation	0.031
Dementia	Atopic Dermatitis	0.062
Acquired Immunodeficiency Syndrome	Congenital Heart Defects	0.062
Bacterial Pneumonia	Malaria	0.156
Osteoporosis	Patent Ductus Arteriosus	0.156
Amino Acid Sequence	Anti Bacterial Agents	0.156
Otitis Media	Infantile Colic	0.156
Hyperlipidemia	Hyperkalemia	0.156

The second dataset we used (we call it *Dataset 2*) contains 36 biomedical (*MeSH*) term pairs [3]. The human scores in this dataset are the average evaluated scores of reliable doctors**.** Table 2 contains the first 10 pairs of this dataset. We used the *UMLSKS* browser [18] for *SNOMED-CT*, and *MeSH* Browser [19] for *MeSH* to get information on the terms in the two datasets.

B. Experiments and Results

In the experiments, we assumed that the two features contribute equally to semantic similarity (*i.e.,* $\alpha = \beta = 1$) and conducted comparisons with four other structure-based semantic similarity measures. All the measures use node counting for path length and for depth of concept nodes except Li et al measure which uses edge/link approach [8].

Table 4. Absolute values of correlations with human judgments for all measures on 29 pairs of *Dataset 1*

Measure	Dataset 1	
	MeSH (rank)	**SNOMED-CT (rank)**
Path Length	0.744 (5)	0.254 (5)
Wu & Palmer	0.794 (4)	0.296 (4)
Leacock& Chodorow	0.857 (2)	0.431 (2)
Choi & Kim	0.725 (6)	0.152 (6)
Li et al.	0.852 (3)	0.371 (3)
Sem *(proposed)*	0.863 (1)	0.665 (1)
Average	***0.806***	***0.362***

Table 5. Absolute values of correlation with human judgments for all measures on 34 pairs of *Dataset 2*

Measure	Dataset 2	
	MeSH (rank)	**SNOMED-CT (rank)**
Path Length	0.765 (5)	0.586 (5)
Wu & Palmer	0.811 (4)	0.686 (3)
Leacock& Chodorow	0.820 (3)	0.677 (4)
Choi & Kim	0.673 (6)	0.440 (6)
Li et al.	0.830 (1)	0.694 (2)
Sem *(proposed)*	0.825 (2)	0.735 (1)
Average		

Out of the 30 pairs of *Dataset 1*, we were able to find only 25 pairs in *MeSH*, and 29 pairs in *SNOMED-CT*. For the four pairs that were not found in *MeSH* and found in *SNOMED-CT*, we calculated average similarity of the most related concepts to each one of them, so we have 29 pairs in *MeSH* and *SNOMED-CT* in total. We also found 34 pairs out of the 36 pairs of *Dataset 2* in *SNOMED-CT*, so we use only these 34 pairs in the experiments. Table 4 and Table 5 shows the correlation results with human rating scores experimented on *MeSH* and *SNOMED-CT* using *Dataset 1* and *Dataset 2*.

C. Discussion

Tables 4 and 5 show that the proposed measure achieves the best correlations with human ratings and ranks #1 in almost all experiments with two ontologies. These results confirm the efficiency of our proposed measure and the '*goodness*' of the new features. If we take the average correlation of all experiments, for each measure, we find that *Sem* achieves 23.9% higher that the average of the other measures. Moreover, we notice that by testing with *Dataset 1* using *SNOMED-CT* (Table 4) most structure-based measures perform very low except *Sem* because the specificity (granularity) of *SNOMED-CT* is much more than that of *MeSH*. So our measure, by using specificity feature, significantly outperforms the other five measures. We observe that with *Dataset 1*, all measures get the same ranks in both *MeSH* and *SNOMED-CT* (Table 4). Another observation is that both Path length and Choi & Kim maintain their ranks in both ontologies and in the two datasets. The average correlations of all measures in Table 4 and Table 5 show that *MeSH* gives higher correlations with human scores than *SNOMED-CT*, that is, all measures perform better in *MeSH* than in *SNOMED-CT*.

VI. Conclusion

This paper presents a new ontology-based semantic similarity measure as a cluster-based approach. The proposed measure was evaluated in the biomedical domain using two datasets of biomedical term pairs and two different ontologies (*MeSH & SNOMED-CT*) within the *UMLS*. The main contribution of this paper is the new measure with new features (common specificity and local granularity) that are combined non-linearly in the semantic similarity measure. The experimental results have proven the efficiency of the proposed measure and the new features relative to human judgments and compared with five other semantic measures. Compared with other measures, the proposed measure produced the best overall correlation result with human judgments in both datasets and in two ontologies. The experimental results demonstrated, further, that *MeSH* ontology produces better semantic similarity correlations with human ratings than *SNOMED-CT* over most of the tested measures.

Reference

[1] J.E. Caviedes and J.J. Cimino, "Towards the development of a conceptual distance metric for the *UMLS,*" *Journal of Biomedical Informatics,* vol. 37, no. 2, pp. 77-85, 2004

[2] I. Choi and M. Kim 2003**, "** Topic distillation using hierarchy concept tree," *Proceedings of the 26th annual int'l ACM SIGIR conference on Research and development in information retrieval,* 2003.

[3] A. Hliaoutakis, "Semantic Similarity Measures in MeSH Ontology and their application to Information Retrieval on Medline," Master's thesis, Technical University of Crete, Greek, 2005.

[4] J.J. Jiang and D.W. Conrath, "Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy," *Proc. ROCLING X*, 1997.

[5] R. Kleinsorge, C. Tilley, and J. Willis," Unified Medical Language System (UMLS) Basics"**.** Available: http://www.nlm.nih.gov/ research/umls/pdf/UMLS_Basics.pdf

[6] H. Kuntz and M. V. Berkum, "SNOMED CT® A standard Terminology for Healthcare". Available: http://www.sst.dk/upload/informatik_og_sundhedsdata/sundhedsinformatik/terminologi/kuntz_vanberkum_snomedct_30mar05.pdf

[7] C. Leacock, and M. Chodorow, "Combining local context and WordNet similarity for word sense identification," *In Fellbaum, C., ed., WordNet: An electronic lexical database,* MIT press. 1998.

[8] Y. Li, Z.A. Bandar, and D. McLean. "An approach for measuring semantic similarity between words using multiple information sources". IEEE Transactions on knowledge and data engineering, Vol. 15, No. 4, 2003.

[9] D. Lin, "An Information-Theoretic Definition of Similarity," *Proc. Int'l Conf. Machine Learning*, July 1998.

[10] G.A. Miller, "WordNet: A Lexical Database for English," *Comm. ACM*, vol. 38, no. 11, pp. 39-41, 1995.

[11] H. A. Nguyen. and H. Al-Mubaid, "New Ontology-Based Semantic Similarity Measure for the Biomedical Domain," *Proceedings of IEEE GrC06*, 2006.

[12] T. Pedersen, S. Pakhomov and S. Patwardhan. "Measures of Semantic Similarity and Relatedness in the Medical Domain," University of Minnesota Digital Technology Center Research Report *DTC 2005/12.*

[13] R. Rada, H. Mili, E. Bichnell, and M. Blettner, "Development and Application of a Metric on Semantic Nets," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 17-30, 1989.

[14] P. Resnik., "Using information content to evaluate semantic similarity," *In Proc. of the 14th International Joint Conference on Artificial Intelligence*, 448–453. Montreal, Canada, 1995.

[15] R. Richardson, A.F. Smeaton, and J. Murphy, "Using WordNet as a Knowledge Base for Measuring Semantic Similarity," Working paper CA-1294, School of Computer Applications, Dublin City Univ., Dublin, 1994.

[16] R.N. Shepard, "Towards a Universal Law of Generalisation for Psychological Science," *Science,* vol. 237, pp.1317-1323, 1987.

[17] Z. Wu and M. Palmer, "Verb semantics and lexical selection," *In 32nd Annual Meeting of the Association for Computational Linguistics*, pp. 133–138,1994.

[18] UMLS Browser-UMLS Knowledge Source Server. Available: http://umlsks.nlm.nih.gov

[19] MeSH Browser: www.nlm.nih.gov/mesh/2005/MBrowser.htm