X-Similarity: Computing Semantic Similarity between Concepts from Different Ontologies

Euripides G.M. Petrakis, Giannis Varelas,
Angelos Hliaoutakis, Paraskevi Raftopoulou
Department of Electronic and Computer Engineering
Technical University of Crete (TUC)
Chania, Crete, GR-73100, Greece
petrakis@intelligence.tuc.gr, varelas@softnet.tuc.gr,
angelos@softnet.tuc.gr, paraskevi@intelligence.tuc.gr

Abstract

Semantic Similarity relates to computing the similarity between concepts (terms) which are not necessarily lexically similar. We investigate approaches to computing semantic similarity by mapping terms to an ontology and by examining their relationships in that ontology. More specifically, we investigate approaches 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). The most popular semantic similarity methods are implemented and evaluated using WordNet and MeSH. The focus of this work is also on cross ontology methods which are capable of computing the semantic similarity between terms stemming from different ontologies (WordNet and MeSH in this work). This is a far more difficult problem (than the single ontology one referred to above) which has not been investigated adequately in the literature. *X-Similarity*, a novel cross-ontology similarity method is also a contribution of this work. All methods examined in this work are integrated into a semantic similarity system which is accessible on the Web.

1. Introduction

This work aims at providing robust tools for standardizing the quality and delivery of information across communicating information sources. In this set-up, information is acquired from several disparate sources in several formats using different language terminologies. Interpreting the meaning of this information is left to the users. This task can be highly subjective and time consuming. It depends on the user's level of expertise to decide on the exact meaning of information acquired. This is a common problem in applications such as information retrieval on the Web or information communication on a grid. This is also an interoperability issue and refers both to the need for computers to communicate, but also to the need for the users to work together and understand each other (across all specialties and levels of expertise). Intelligent middleware is needed to help users and computers perform this task. Ontologies offer the means by which such middleware can be developed so that consistent information can be communicated between users or processed by computers.

For making knowledge commonly understandable, it is agreed that the concepts or terms characterizing different communicating sources (i.e., humans or computers) are represented by ontologies. It is common that the communicating sources use the same ontology. However, it is also possible that a different ontology is employed to represent the concepts of each knowledge source. Then, knowledge sharing (between humans or computers) is related to mapping or comparing concepts in the same or across different ontologies (specific to each source). To relate concepts or entities between different sources (the same as for answering user queries involving such concepts or entities), the concepts extracted from each source must be compared in terms of their semantic meaning.

Semantic similarity relates to computing the similarity between concepts which are not lexically similar. This is an important problem in Natural Language Processing (NLP) and Information Retrieval (IR) research and has received considerable attention in the literature. Several algorithmic approaches for computing semantic similarity have been proposed. Detection of similarity between concepts or entities is possible if they share common attributes or if they are linked with other semantically related entities in an ontology (e.g., [3,8]). To relate concepts in different ontologies, semantic similarity works by discovering linguistic relationships or affinities between ontological terms across different ontologies (e.g., [10]).

We present a critical evaluation of several semantic similarity approaches using two well known taxonomic hierarchies (or ontologies) namely WordNet[1] and MeSH[2]. WordNet is a controlled vocabulary

[1] http://wordnet.princeton.edu

[2] http://www.nlm.nih.gov/mesh

and thesaurus offering a taxonomic hierarchy of natural language terms developed at Princeton University. MeSH (Medical Subject Heading) is also a controlled vocabulary and thesaurus developed by the U.S. National Library of Medicine (NLM) offering a hierarchical categorization of medical terms. Similar results for MeSH have not been reported before in the literature.

We investigate approaches for computing semantic similarity by mapping terms or concepts to an ontology and by examining their relationships in that ontology (WordNet for natural language terms and MeSH for medical terms). Comparing concepts that belong to different ontologies is a far more difficult problem that has not been investigated adequately in the literature. *X-Similarity*, a novel cross-ontology semantic similarity method is also proposed and evaluated as part of this work. *X-Similarity* demonstrates very promising performance improvements over existing methods (e.g., [10]). All methods are implemented and integrated into a semantic similarity system which is accessible on the Web[3].

2. Related Work and Background

Issues related to semantic similarity algorithms along with issues related to computing semantic similarity on WordNet and MeSH are discussed below.

2.1. WordNet

WordNet is an on-line lexical reference system developed at Princeton University. WordNet attempts to model the lexical knowledge of a native speaker of English. WordNet can also be seen as ontology for natural language terms. WordNet v.2.0 contains around 100,000 terms, organized into taxonomic hierarchies. Nouns, verbs, adjectives and adverbs are grouped into synonym sets (*synsets*). The synsets are also organized into senses (i.e., corresponding to different meanings of the same term or concept). The synsets (or concepts) are related to other synsets higher or lower in the hierarchy defined by different types of relationships. The most common relationships are the *Hyponym/Hypernym* (i.e., Is-A relationship), and the *Meronym/Holonym* (i.e., Part-Of relationship). There are nine noun and several verb Is-A hierarchies (adjectives and adverbs are not organized into Is-A hierarchies). Figure 1 illustrates a fragment of the WordNet Is-A hierarchy.

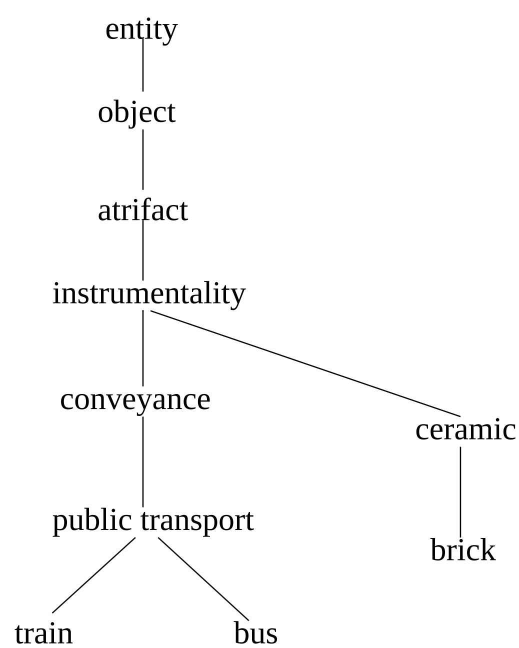

Figure 1. A fragment of the WordNet Is-A hierarchy.

2.2. MeSH

MeSH (Medical Subject Headings) is a taxonomic hierarchy (ontology) of medical and biological terms (or concepts) suggested by the U.S National Library of Medicine (NLM). MeSH terms are organized in Is-A taxonomies with more general terms (e.g., “chemicals and drugs”) higher in a taxonomy than more specific terms (e.g., “aspirin”). There are 15 taxonomies with more than 22,000 terms (in MeSH edition 2004). A term may appear in more than one taxonomies. Each MeSH term is described by several properties, the most

[3] http://www.intelligence.tuc.gr/similarity

important of them being the *MeSH Heading (MH)* (i.e., term name or identifier), *Scope Note* (i.e., a text description of the term) and *Entry Terms* (i.e., mostly synonym terms to the MH). Entry terms also include stemmed MH terms and are sometimes referred to as *quasi-synonyms* (they are not always exactly synonyms). In this work, entry terms are treated as synonyms. Each MeSH term is also characterized by its MeSH tree number (or code name) indicating the exact position of the term in a MeSH tree taxonomy (e.g., “D01,029” is the code name of term “Chemical and drugs”). Figure 2 illustrates a fragment of the MeSH Is-A hierarchy.

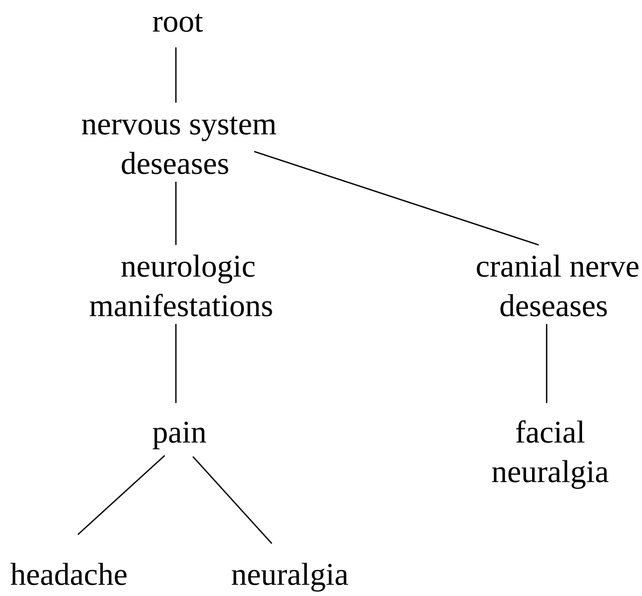

Figure 2. A fragment of the MeSH Is-A hierarchy.

2.3. Semantic Similarity

Several methods for determining semantic similarity between terms have been proposed in the literature and some of them have been tested on WordNet[4]. We present an evaluation for a more complete and up-to-date set of methods and we also investigate cross ontology methods. Similar results on MeSH have not been reported in the literature. Similarity measures apply only for nouns (and verbs in WordNet) and for Is-A relationships. Taxonomic properties for adverbs and adjectives do not exist. Semantic similarity methods are classified into four main categories:

Edge Counting Methods: Measure the similarity between two terms (concepts) as a function of the length of the path linking the terms and on the position of the terms in the taxonomy [7,9,13,3,2].

Information Content Methods: Measure the difference in information content between two terms as a function of their probability of occurrence in a corpus [6,8,4,1]. In this work, information content is computed according to [11]: WordNet is used as a statistical resource for computing the probabilities of occurrence of terms. More general concepts (higher in the hierarchy) with many hyponyms convey less information content than more specific terms (lower in the hierarchy) with less hyponyms. This approach is independent of the corpus and also guarantees that the information content of each term is less than the information content of its subsumed terms. This constraint is common to all methods of this category. Computing information content from a corpus does not always guarantee this requirement. The same method is also applied for computing the information content of MeSH terms.

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 [12].

Hybrid methods [10] combine the above ideas: Term similarity is computed by matching synonyms, term neighborhoods and term features. Term features are further distinguished into parts, functions and attributes and are matched similarly to [12].

---

[4] http://marimba.d.umn.edu/cgi-bin/similarity.cgi

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). Semantic similarity methods can also be distinguished between:

Single Ontology similarity methods that assume that the terms, which are compared, are from the same ontology (e.g., WordNet).

Cross Ontology similarity methods, which compare terms from different ontologies (e.g., WordNet and MeSH).

Edge counting and information content methods work by exploiting structure information (i.e., position of terms) and information content of terms in a hierarchy and are best suited for comparing terms from the same ontology. Because the structure and information content of different ontologies are not directly comparable, cross ontology similarity methods usually call for hybrid or feature based approaches. The focus of this work is on both single and cross ontology methods.

3. Cross Ontology Semantic Similarity

A recent contribution by Rodriguez [10] proposed a framework for comparing terms stemming from the same or from different ontologies. The similarity between terms *α* and *b* is computed as a weighted sum of similarities between synonym sets (synsets), features and terms neighborhoods:

$$Sim(a,b) = w \cdot S_{synsets}(a,b) + u \cdot S_{features}(a,b) + v \cdot S_{neighborhoods}(a,b), \tag{1}$$

with *w*, *u* and *v* denoting the relative importance of the three similarity components. Features are further specialized into “parts”, “attributes” and “functions”. For example, in WordNet $S_{features}$ is implemented as the matching of terms having the Part-Of relationship. Notice that no Part-Of relationships are defined in MeSH and this term is omitted when this method is applied on MeSH. Assuming that all terms in the neighborhoods of terms *α* and *b* as well as their features (i.e., their corresponding parts, attributes and functions) are also represented by synsets, each similarity component is computed by Tversky [12] as

$$S(a,b) = \frac{|A \cap B|}{|A \cap B| + \gamma(a,b)|A \setminus B| + (1-\gamma(a,b))|B \setminus A|}, \tag{2}$$

where *A*, *B* denote synsets of terms ***a***, *b* and *A\B* denotes the set of terms in A but not in B (the reverse for *B/A*). Parameter *γ(α,b)* is computed as a function of the depth of the terms *α* and *b* in their taxonomy:

$$\gamma(a,b) = \begin{cases} \dfrac{depth(a)}{depth(a)+depth(b)}, & depth(a) \le depth(b); \\ 1-\dfrac{depth(a)}{depth(a)+depth(b)}, & depth(a) > depth(b), \end{cases} \tag{3}$$

X-Similarity relies on matching between synsets and term description sets. The term description sets contain words extracted by parsing term definitions (“glosses” in WordNet or “scope notes” in MeSH). Two terms are similar if their synsets or description sets or, the synsets of the terms in their neighborhood (e.g., more specific and more general terms) are lexically similar. First, we propose replacing Equation 2 by plain set similarity

$$S(a,b) = \frac{|A \cap B|}{|A \cup B|}, \tag{4}$$

where *A* and *B* denote synsets or term description sets. Because not all terms in the neighborhood of a term are connected with the same relationship, we propose that set similarities are computed per relationship type (e.g., Is-A and Part-Of for WordNet and only Is-A for MeSH) as

$$S_{neighborhoods}(a,b) = \max \frac{|A_i \cap B_i|}{|A_i \cup B_i|}, \tag{5}$$

where *i* denotes relationship type. Equation 5 suggests computing the similarity between term neighborhoods, by matching same type relationships between synsets of more specific and of more general terms (i.e., for each term, the union of the synsets of all terms up to the root of each term hierarchy is taken) and by taking their maximum. The above ideas are combined into a single formula as follows

$$Sim(a,b) = \begin{cases} 1, & if\ S_{synsets}(a,b) > 0; \\ \max\{S_{neighborhoods}(a,b), S_{descriptions}(a,b)\}, & if\ S_{synsets}(a,b) = 0. \end{cases} \tag{6}$$

$S_{descriptions}$ denotes matching of term description sets. $S_{descriptions}$ and $S_{synsets}$ are computed according to Equation 4. Notice that, two terms with at least one common synonym term are 100% similar.

The following summarize the differences between *X-Similarity* and [10]:

- Parameter γ in [10] suggests taking into account the depth of the terms in the two ontologies. However, cross ontology matching should not depend on ontology structure information (i.e., it has to be independent of the depth of the two terms in their ontologies). *X-Similarity* does not use γ.
- Matching by Equation 2 unreasonably penalizes similarity for non common words in term definitions or synsets. However, word content of term descriptions may vary significantly from one ontology to another (from a few words in one ontology to a paragraph in another ontology). In *X-Similarity*, this is handled by Equation 4, which suggests matching based only on common words: The more common words the two definitions have, the more similar the terms are (and the reverse).
- According to [10] appropriate weights must be specified in Equation 1. The method did not show how to compute good weights (e.g., such weights can be computed by machine learning by decision trees). In the experiments with [10] below all weights are set to 1. *X-Similarity* does not use weights.

Notice that both, the method by Rodriguez [10] and *X-Similarity* can also be used for matching terms from the same ontology (WordNet or MeSH in this work).

Table 1 illustrates the term descriptions (in XML) of the terms *Hypothyroidism* and *Hyperthyroidism* taken from WordNet and MeSH respectively. Their similarity is computed by Equation 6 as $\max\{S_{neighborhoods}, S_{descriptions}\} = 0.387$. Notice that $S_{synsets} = 0$ (i.e., the two terms have no common synsets).

Table 1. XML descriptions of two terms from WordNet and MeSH.

WordNet term: **Hypothyroidism**	MeSH term: **Hyperthyroidism**
<term> hypothyroidism <definition> An underactive thyroid gland; a glandular disorder Resulting from insufficient production of thyroid hormones. </definition> <synset> Hypothyroidism </synset> <hypernyms> glandular disease, disorder, condition, state </hypernyms> <hyponyms> myxedema, cretinism </hyponyms> </term>	<term> hyperthyroidism <definition> Hypersecretion of Thyroid Hormones from Thyroid Gland. Elevated levels of thyroid hormones increase Basal Metabolic Rate. </definition> <synset> Hyperthyroidism </synset> <hypernyms> disease, thyroid, Endocrine System Diseases, diseases </hypernyms> <hyponyms> thyrotoxicosis, thyrotoxicoses </hyponyms> </term>

4. Evaluation of Semantic Similarity Methods

In the following, we present a comparative evaluation of the similarity methods referred to above. All data sets used in the experiments below are available on the Web[5].

4.1. Semantic Similarity on WordNet

In accordance with previous research [8], we evaluated the results obtained by applying the semantic similarity methods of Section 2.3 to the same pairs used in the experiment by Miller and Charles [6]: 38 undergraduate students were given 30 pairs of nouns and were asked to rate the similarity of each pair on a scale from 0 (not similar) through 4 (perfect synonymy). The average rating of each pair represents a good estimate of how similar the two words are.

We compared the computed similarity scores for the same terms as in Miller and Charles with the human relevance results reported there. The similarity values obtained by all competitive computational methods (all senses of the first term are compared with all senses of the second term) are correlated with the average scores obtained by the humans in [6]. The higher the correlation of a method the better the method is (i.e., the more it approaches the results of human judgments).

Table 2 shows the correlation obtained by each method. Jiang and Conrath [1] suggested removing one of the pairs from the evaluation. This increased the correlation of their method to 0.87. The method by Li et. al. [3] is among the best and it is also the fastest. These results lead to the following observations:

- Information Content methods perform very well and close to the upper bound suggested by [8].

[5] http://www.intelligence.tuc.gr/similarity

- Methods that consider the positions of the terms in the hierarchy (e.g., [3]) perform better than plain path length methods (e.g., [7]).
- 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 (e.g., [10]) are mainly targeted towards cross ontology similarity applications where edge counting and information content methods do not apply.
- *X-Similarity* performs at least as good as [12,10].

Table 2. Evaluation of Edge Counting, Information Content, Feature-based and Hybrid semantic similarity methods on WordNet.

Method	Method Type	Correlation
Rada [7]	Edge Counting	0.59
Wu [13]	Edge Counting	0.74
Li [3]	Edge Counting	0.82
Leacock [2]	Edge Counting	0.82
Richardson [9]	Edge Counting	0.63
Resnik [8]	Information Content	0.79
Lin [4]	Information Content	0.82
Lord [5]	Information Content	0.79
Jiang [1]	Information Content	0.83
Tversky [12]	Feature-Based	0.73
X-Similarity	Feature-Based	0.74
Rodriguez [10]	Hybrid	0.71

4.2. Semantic Similarity on MeSH

An evaluation of Semantic Similarity methods on MeSH has not been reported in the literature before. For the evaluation, we designed an experiment similar to that by Miller and Charles [6] for WordNet: we asked Dr. Qi at Dalhousie University to compile a set MeSH term pairs. Dr. Qi proposed a set of 49 pairs and for each pair she provided an estimate of similarity between 0 (not similar) and 4 (perfect similarity). We created a form-based interface[6] with all pairs on the Web and we asked medical experts to enter their evaluation (the interface is still accepting results by experts world-wide). So far, we accepted results from 12 experts**.**

The analysis of the results revealed that: (a) Some medical terms are more involved, or ambiguous leading to ambiguous evaluation by many users. For each pair, the standard deviation of their similarity (over all users) was computed. Pairs with standard deviation higher than a user defined threshold $t = 0.8$ were excluded from the evaluation. (b) Medical experts were not at the same high level of expertise and (in some cases) gave unreliable results. For each user, we computed the standard deviation of their evaluation (over all pairs). We excluded users who gave significantly different results from the majority of other users. Overall, 13 out of the 49 pairs and 4 out of the 12 users were excluded from the evaluation.

Following the same procedure as in the WordNet experiments, the similarity values obtained by each method (all senses of the first term are compared with all senses of the second term) are correlated with the average scores obtained by the humans. The correlation results are summarized in Table 3. These results lead to the following observations:

- Edge counting and information content methods perform about equally well. However, methods that consider the positions of the terms (lower or higher) in the hierarchy (e.g., [3]) perform better than plain path length methods (e.g., [7,9]).
- Hybrid and feature based methods (the same as our proposed method) exploiting properties of terms (e.g., scope notes, entry terms) perform at least as well as information content and edge counting methods (exploiting information relating to the structure and information content of the underlying taxonomy), implying that term annotations in MeSH are significant information by themselves and that it is possible to design even more effective methods by combining information from all the above sources (terms annotations, structure information and information content).
- *X-Similarity* performs at least as good as other good methods (e.g., [1,8,10]).

[6] http://www.intelligence.tuc.gr/similarity/mesh

Table 3. Evaluation of Edge Counting, Information Content, Feature-based and Hybrid semantic similarity methods on MeSH.

Method	Method Type	Correlation
Rada [7]	Edge Counting	0.50
Wu [13]	Edge Counting	0.67
Li [3]	Edge Counting	0.70
Leacock [2]	Edge Counting	0.74
Richardson [9]	Edge Counting	0.64
Resnik [8]	Information Content	0.71
Lin [4]	Information Content	0.72
Lord [5]	Information Content	0.70
Jiang [1]	Information Content	0.71
Tversky [12]	Feature-Based	0.67
X-Similarity	Feature-Based	0.71
Rodriguez [10]	Hybrid	0.71

4.3. Cross Ontology Semantic Similarity on WordNet and MeSH

An evaluation of cross ontology Semantic Similarity methods has not been reported elsewhere. In accordance with the previous evaluations, we designed an experiment similar to that by Miller and Charles [6] using the same 43 MeSH term pairs of Section 4.2. We observed that in 40 (out of the 43) pairs, at least one of the two terms is also a WordNet term. For the evaluation, one of the two terms in each pair is considered to be a WordNet term (ignoring the fact that it is also a MeSH term) and is compared with the second MeSH term (ignoring the fact that it can be also a WordNet term). The similarity values obtained by the two methods considered in Section 3 are correlated with the average scores obtained by the humans in the experiment of Section 4.2. Table 4 summarizes the correlation results.

Table 4. Evaluation of Cross Ontology similarity methods.

Method	Method Type	Correlation
X-Similarity	Feature-Based	0.70
Rodriguez [10]	Hybrid	0.55

Our proposed method achieves 15% better correlation than the state-of-the-art method [10]. This is mostly due to the use of parameter γ (Equation 3) and due to matching by Equation 2. *X-Similarity* handles both these cases by omitting γ and by not penalizing for non-common attributes.

5. Semantic Similarity System

All methods are implemented and integrated into a semantic similarity system which is accessible on the Web[7]. Each MeSH or WordNet term is represented by its tree hierarchy (as in Figure 1 and Figure 2) which is represented by an XML file (as in Table 1) stored in the XML repository. These XML files are created by the XML generator using the WordNet XML Web-Service[8]. The purpose of this structure is to facilitate access to terms (as well as to their hyponyms or hypernyms) stored in an XML repository. The terms are indexed by their name of identifier (otherwise accessing a term would require exhaustive searching through the entire WordNet or MeSH files). The information content of all terms is also computed in advance and stored separately in the Information Content (IC) database. The user is provided with several options at the user interface (e.g., ontology selection, method selection, sense selection). The user is also assisted by a suggestion tool for selecting terms from an ontology (e.g., show me terms which contain string "ball").

[7] http://www.intelligence.tuc.gr/similarity
[8] http://wnws.sourceforge.net

6. Conclusions

We experimented with several semantic similarity methods for computing the conceptual similarity between natural language terms using WordNet and MeSH. To our knowledge, similar experiments with MeSH haven't been reported elsewhere. The experimental results indicate that it is possible for these methods to approximate algorithmically the human notion of similarity reaching correlation (with human judgment of similarity) up to 83% for WordNet and up to 74% for MeSH. This work also presents *X-Similarity*, a novel semantic similarity method which has been shown to out-perform the state-of-the-art cross ontology matching method [10] by 15%. All methods are implemented and integrated into a semantic similarity system which is accessible on the Web.

Acknowledgements

We are grateful to Prof. Evangelos Milios of Dalhousie University for his constructive comments, to Dr. Qiufen Qi of Dalhousie University for preparing the MeSH terms for the evaluation. We are also grateful to Spyros Argyropoulos of TUC who implemented the Web interface of the Semantic Similarity system. This work was supported by project BIOPATTERN (FP6-NoE, Project No 508803) and by project TOWL (FP6-Strep, Project No 26896).

References

[1] Jiang J.J. and Conrath D.W. (1998). Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy. In: *Intern. Conf. on Research in Computational Linguistics*, Taiwan.

[2] Leacock C. and Chodorow M. (1998). Combining Local Context and WordNet Similarity for Word Sense Identification in WordNet. In: Christiane Fellbaum, editor, *An Electronic Lexical Database*, pp. 265-283. MIT Press.

[3] Li Y., Bandar Z.A., and McLean D. (2003). An Approach for Measuring Semantic Similarity between Words Using Multiple Information Sources. *IEEE Trans. on Knowledge and Data Engineering*, 15(4), 871-882.

[4] Lin D. (1993). Principle-Based Parsing Without Overgeneration. In: *Annual Meeting of the Association for Computational Linguistics (ACL'93)*, pp. 112-120, Columbus, Ohio.

[5] Lord P.W., Stevens R.D., Brass A., and Goble C.A. (2003). Investigating Semantic Similarity Measures across the Gene Ontology: the Relationship between Sequence and Annotation. *Bioinformatics*, 19(10), 1275-83.

[6] Miller G. and Charles W.G. (1991). Contextual Correlates of Semantic Similarity. *Language and Cognitive Processes*, 6, 1-28.

[7] Rada R., Mili H., Bicknell E., and Blettner M. (1989). Development and Application of a Metric on Semantic Nets. *IEEE Trans. on Systems, Man, and Cybernetics*, 19(1), 17-30.

[8] Resnik O. (1999). Semantic Similarity in a Taxonomy: An Information-Based Measure and its Application to Problems of Ambiguity and Natural Language. *Journal of Artificial Intelligence Research*, 11, 95-130.

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

[10] Rodriguez M.A. and Egenhofer M.J. (2003). Determining Semantic Similarity among Entity Classes from Different Ontologies. *IEEE Trans. on Knowledge and Data Engineering*, 15(2), 442-456.

[11] Seco N., Veale T., and Hayes J. (2004). An Intrinsic Information Content Metric for Semantic Similarity in WordNet. Tech. report, University College Dublin, Ireland.

[12] Tversky A. (1997). Features of Similarity. *Psycological Review*, 84(4), 327-352.

[13] Wu Z. and Palmer M. (1994). Verb Semantics and Lexical Selection. In: *Annual Meeting of the Associations for Computational Linguistics (ACL'94)*, pp. 133-138, Las Cruces, New Mexico.