

Semantic Relatedness Measure Using Object Properties in an Ontology

Laurent Mazuel and Nicolas Sabouret

Laboratoire Informatique de Paris 6 - LIP6, 104 av du Président Kennedy,
75016, Paris, France

{laurent.mazuel,nicolas.sabouret}@lip6.fr

Abstract. This paper presents a new semantic relatedness measure on ontologies which considers especially the object properties between the concepts. Our approach relies on two hypotheses. Firstly, using only concept hierarchy and object properties, only a few paths can be considered as “semantically corrects” and these paths obey to a given set of rules. Secondly, following a given edge in a path has a cost (represented as a weight), which depends on its type (*is-a*, *part-of*, etc.), its context in the ontology and its position in this path. We propose an evaluation of our measure on the lexical base WordNet using *part-of* relation with two different benchmarks. We show that, in this context, our measure outperforms the classical semantic measures.

1 Introduction

Whereas semantic similarity focuses on common points in the concepts definitions, semantic relatedness permits to take into account functional relations between concepts. Thus, as stated by Resnik [1], semantic similarity “represents a special case of semantic relatedness”, that only uses the subsumption relation of the knowledge representation. For example, “car” and “gasoline” have a low similarity degree but a high relatedness degree [1]. However, automatic computation of a relatedness degree is considered to be much more difficult than computing a similarity measure. For this reason, much work on semantic measures has focused on computing similarity degrees using well-known hierarchy of concepts (*e.g.* MeSH [2], WordNet [3]). Recent work on relatedness measure tries to adopt a different point of view with the use of glosses [4,5], since a gloss contains a functional description of the concept (by means of other concepts), or terms frequencies in Internet using search engine [6,7]. However, no work, since work of Hirst & St-Onge in 1998 [8], has focused on the issue of semantic relatedness using heterogeneous relations in a graph-based knowledge representation, as semantic network or ontology limited to hierarchy and object properties. In this paper, semantic relatedness between two concepts in a semantic net can be materialized by a path, starting from one concept, following different kinds of relations (subsumption (*is-a*), meronymy (*part-of*) or any other domain specific relation) to the other concept. However, the Hirst & St-Onge measure assumes that all edges have the same weight (*i.e.* the information-content of all edges is uniform in the ontology), which has been demonstrated to be an incorrect hypothesis [1].

Computing semantic relatedness in graph-based knowledge base rather than using glosses is however a challenging and important issue. Indeed, the great majority of existing Human-Machine interaction systems with Natural Language (questions/answering systems, dialogue systems...) makes use of ontologies for the semantic interpretation of requests [9,10,11]. Moreover, in web services composition, it is often interesting to compute the semantic similarity between two ontologies before searching for an alignment (to avoid useless attempts if the ontologies don't represent the same data) [12,13]. However in most work, this ontology is reduced to a hierarchy of concepts and the semantic interpretation relies only on similarity measure, although the literature underlines the need for semantic relatedness measures [14,9,15].

Some recent work has tried to define semantic relatedness using complex ontology languages, as OWL-Lite [16]. For instance, in [15], the authors take into account negation, intersection or disjunction of classes, to compute the relatedness between web services. However, these approaches are currently not evaluated nor implemented and remain at a theoretical point of view. Hence, define a semantic relatedness measure based on ontology limited to concept hierarchy and object properties which can be evaluated can be seen as a first step towards semantic relatedness in complex and real ontologies.

In this paper, we present and evaluate a new semantic relatedness measure on an ontology voluntarily limited to concept hierarchy and object properties with the following attributes:

- It considers a set of patterns to filter the paths which are not “semantically correct”. Actually, when using more relations than only the subsumption, there exists a lot of paths between two concepts and only a sub-set of these paths are “semantically correct” (*i.e.* their structure correspond to a semantic meaning).
- It uses the information-theoretic definition of semantic similarity to weight the hierarchical edges in the graph,
- It computes the weight of non-hierarchical edges.
- It combines these three points in a unique measure.

The next section presents related work on semantic similarity and semantic relatedness. Moreover, it outlines the problem of finding a semantically correct path in a graph-based knowledge representation. Section 3 presents our approach for computing semantic relatedness using subsumption links and heterogeneous links. Section 4 presents the evaluation and discusses the results.

2 Related Work

The first part of this section presents work on similarity measures. The second part deals with semantic relatedness measures and discusses the problem of finding a semantically correct path in a given ontology.

2.1 Semantic Similarity Measures

This kind of measures is also called *taxonomic measures* or *attributional measures* [17]. An intuitive way to compute quickly semantic similarity between two nodes of a hierarchy is to count the number of edges in the shortest path between these two nodes.

The idea captured with this hypothesis is that the semantic distance of two concepts is correlated with the length of the shortest path to join these concepts. This measure was firstly defined by Rada [2]:

$$dist_{rad}(c_1, c_2) = len(c_1, c_2)$$

where len is the length of the shortest path (in terms of the number of edges) between c_1 and c_2 in the hierarchy. However, this measure relies upon the assumption that each edge carries the same amount of information, which is not true in most ontology [1]. Thus, many other formulae extend this measure by computing weights on edges using additional information, like the depth of each concept in the hierarchy [18,19] or the *closest common parent* of the two nodes [20,19]. All these methods are called *edge-based measures*.

On the contrary, *node-based measures* associate a weight to each node. This weight represents the *information content* (IC) of the concept¹. The more specialised a concept is, the more the weight will be. The first node-based similarity measure, proposed by Resnik in [1], is defined by the information content of the *closest common parent* (*ccp*) of the two concept c_1 and c_2 :

$$sim_{res}(c_1, c_2) = IC(ccp(c_1, c_2))$$

where $IC(c)$ is the information content of the concept c . Many other propositions have been made after Resnik to combine the IC of the two target nodes and their ccp (e.g. [22,23]).

An interesting attempt to mix node-based and edge-based methods is the Jiang & Conrath measure [23]. They define their distance measure as “derived from the edge-based notion” (by analogies with the Rada measure [2]) “by adding the information content as a decision factor”. In this measure, each edge is linked to a weight and the semantic distance is computed by adding all the edge weight along the shortest path. The weight $LS(x, y)$ (“ LS for link strength”) of an edge $\{x, y\}$ between the node x and the node y is computing regarding to their information content:

$$LS(x, y) = |IC(x) - IC(y)|$$

Then, the Jiang & Conrath measure is defined as:

$$\begin{aligned} dist_{JC}(c_1, c_2) &= \sum_{\{x,y\} \in sp(c_1, c_2)} LS(x, y) \\ &= IC(c_1) + IC(c_2) - 2 \times IC(ccp(c_1, c_2)) \end{aligned}$$

where $sp(c_1, c_2)$ is the shortest taxonomic path between c_1 and c_2 .

¹ We don't focus in this article on the way to compute the information content of a concept. The most classical ways are the Resnik's approach with a corpus [1] and the Seco's approach with taxonomy analysis [21].

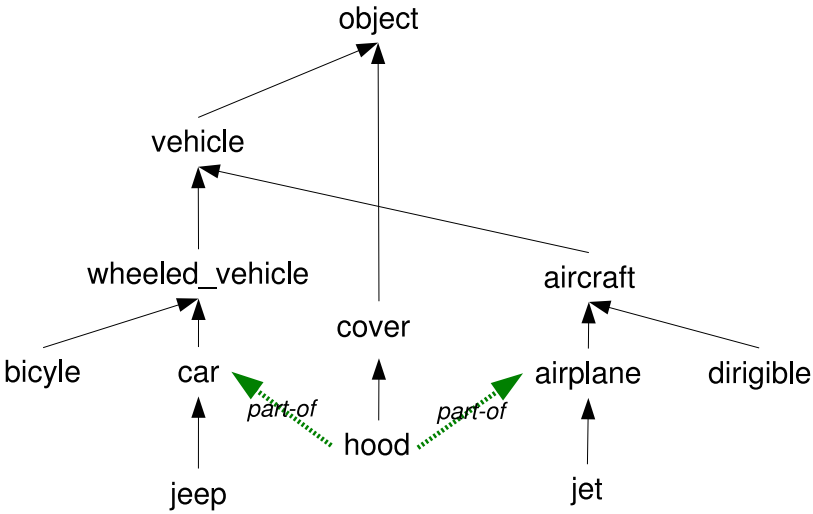


Fig. 1. Example of paths with relations. The black arrows represent *is-a* links.

2.2 Semantic Relatedness Measures and Notion of Semantically Correct Path

This kind of measures is also called *relational measure* [17]. For instance, these measures can capture the semantic link between couples like *gasoline-car* or *bee-honey* [1]. To do this, they have to consider several kinds of relations (and not only the taxonomic relation) as *partOf*, *madeWith*, etc. However, using different kinds of relations will give rise to the problem of multiple existing paths. Indeed, if the shortest path is unique in a hierarchy, many possible paths exist and are conceivable in a graph-based knowledge representation although most of them are not *correct semantically* [24,8]. For this reason, any relational measure must provide (implicitly or explicitly) a set of constraints to ensure that a path is semantically correct.

For instance, in Aleksovski’s work [24], a path is considered to be semantically correct if and only if no hierarchical links appear after a non-hierarchical one. Hence, the path² $\{cover, includes, hood, part-of, airplane, is-a, aircraft\}$ in figure 1 is incorrect for Aleksovski since $(airplane, is-a, aircraft)$ follows the non-hierarchical relation $(hood, part-of, airplane)$.

In [8], Hirst & St-Onge associate a direction in *Upward*, *Downward* and *Horizontal* for each relation type and give three rules to define a semantically correct path in terms of the three directions. It is interesting to note that the rule R1 of Hirst & St-Onge (i.e. “No other direction may precede an upward link”) subsumes the conclusion of Aleksovski work (*is-a* being an Upward link). Finally, Hirst & St-Onge enumerate only 8 patterns of semantically-correct paths which match their three rules: $\{U, UD, UH, UHD, D, DH, HD, H\}$. The difficulty is to determine the direction of each relation. For instance, considering the WordNet [25] relations, the authors define in their paper

² In this article, we will represent a path in a graph using a set-notation as $\{x_1, e_{1,2}, x_2, e_{2,3}, \dots, x_n\}$ where x_i are nodes and $e_{i,i+1}$ are links type.

the hypernymy and meronymy relations as Upward link, the hyponymy and holonymy relations as Downward links and the synonymy and antonymy relations as Horizontal links. As a consequence, for Hirst & St-Onge, the path from *jeep* to *airplane* in figure 1 is semantically correct (pattern *UD*) whereas the path from *bicycle* to *hood* is not (*Upward* link after a *Downward* link), even if these two paths have both the same length.

Once a semantically-correct path has been found, one must evaluate its cost so as to determine whether the concepts are semantically close or far (taking into account that several candidate paths can exist). In Hirst & St-Onge's work, the semantic relatedness value is computed by the following formula:

$$rel_{hso}(c_1, c_2) = C - len(c_1, c_2) - k \times turns(c_1, c_2)$$

where C and k are constant defined empirically ($C = 8$ and $k = 1$ in [8]), len the length of the shortest path between c_1 and c_2 *considering all relations and the patterns* and $turns$ the number of changes of direction in this path. The main drawback of this measure is that it considers, as Rada measure does with similarity, that each edges of each types represents the same information content. However, apart Hirst & St-Onge measure, few work has been made on semantic relatedness measures using heterogeneous relations.

In the next section, we define our semantic relatedness measure, which mixes the information theoretic approach of non-uniformity of the hierarchical edge values, the Hirst & St-Onge patterns to find semantically correct paths and a new proposition to compute a non-hierarchical link weight.

3 Our Semantic Relatedness Measure

Preliminary note. The measure we present in this section is analogue to a distance (*i.e.* a non-relatedness measure): the lower is the score, the higher are the concepts related. However, we will show in section 3.3 a linear transformation from this distance measure to a bounded relatedness measure.

Our work is based on the assumption (used in information-theoretic measure for similarity) that two different hierarchical edges do not carry the same information content, and extends this assumption to non-hierarchical links. We first present the computation of a weight for a single-relation path. We then explain how it can be combined for mixed-relation paths. Finally, we present our measure on the set of semantically correct weighted paths.

3.1 Single-Relation Path

We call *single-relation path* a path whose edges are all of the same type X . For instance, in figure 1, the path $\{jeep, is-a, car, is-a, wheeled_vehicle, is-a, vehicle\}$ is a single-relation path of type *is-a*. To compute the weight W of a single-relation path, we separate hierarchical relations (X is the *is-a* or the *includes* relation) and non-hierarchical relations.

Let us consider a path $path_X(x, y)$ between two concepts x and y in the ontology, following *only* the relation X :

- If X is a hierarchical relation, we chose to consider the Jiang & Conrath definition for the weight of an edge (see section 2.1):

$$W(path_{X \in \{is-a, includes\}}(x, y)) = |IC(x) - IC(y)|$$

Note that even if *is-a* and *includes* are symmetric relation, the definition of single relation-path does not allow them to be in the same path. For instance, the path $\{jeep, is-a, car, is-a, wheeled_vehicle, includes, bicycle\}$ is not a single-relation path but the mix of the two single relation-path $\{jeep, is-a, car, is-a, wheeled_vehicle\}$ and $\{wheeled_vehicle, includes, bicycle\}$. This definition is consistent with the similarity measure of Jiang & Conrath defined in section 2.1.

- If X is not a hierarchical relation, we cannot use the information content of nodes, because this value is computed regarding to the hierarchy structure ([21,1]). We then suggest a new proposition for computing this weight. This weight formula is based upon two parts:
 1. A static weight TC_X , which corresponds to the “strength” of a given relation type. This strength will represent the maximum information content that this kind of link can carried. For a given link of type X , if we have $TC_X < 1$, then this type of link is considered being informative and the cost of this edge must be low. For $TC_X = 1$, the cost will be equals the cost of a hierarchical link and for $TC_X > 1$ the edge will be costly. This allows us to study if different kinds of relations carry different types of information. For instance, in most system the meronymy will have a $TC_{hasPart}$ inferior to 1 and the antonymy relation will, in the contrary, have a $TC_{antonymy}$ superior to 1.
 2. A formula to compute the impact of the length of the path for the cost of the path. We want that the cost of a path with a single relation X must respect the following properties: 1) it increases with the length of the path, 2) it is bounded by TC_X which represents the worst possible value for an X -relation path (*i.e.* the value of an infinite-length path that uses only X relations) and 3) it follows a polynomial function to be compared with the *log* progression of IC value for hierarchical links.

Actually, information-theoretic measures [1,3] have outlined the adequacy of the *log* function to compute the weight of a node (and, by extension, an edge). However, to respect our three requirements and since the *log* function is not bounded, we use the $n/n+1$ function to simulate a logarithmic bounded function (figure 2).

As a consequence, the weight of $path_X(x, y)$ when X is not a hierarchical relation, is defined by:

$$W(path_X(x, y)) = TC_X \times \left(\frac{|path_X(c_1, c_2)|}{|path_X(c_1, c_2)| + 1} \right)$$

3.2 Mixed-Relation Path

Let us consider a path $path(x, y)$ between two concepts x and y in the ontology. It can be factorized as an ordered set of n single-relation sub-paths such as:

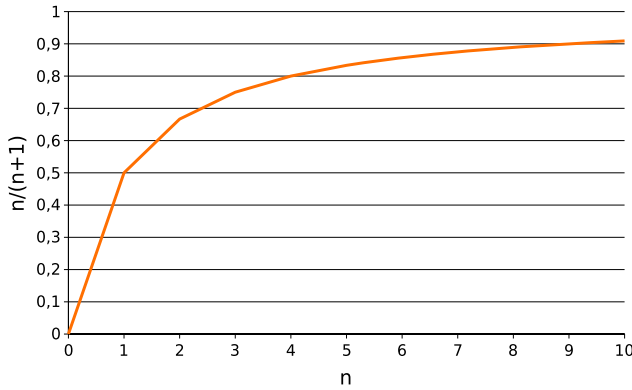


Fig. 2. The progression of the function $n/n+1$, which is bounded and close to the \log function

$$path(x,y) = path_{X_1}(x,z_1) \oplus path_{X_2}(z_1,z_2) \oplus \dots \oplus path_{X_n}(z_{n-1},y)$$

Therefore, there exist several possible factorizations of a given path $path(x,y)$. We define the minimal factorization $T_{min}(path(x,y))$ as the factorization which minimizes the value n . As a consequence, in the minimal factorization $T_{min}(path(x,y))$ of n sub-paths, we have the property $\forall i \in [1, n - 1], X_i \neq X_{i+1}$.³ This property can be used to build the $T_{min}(path(x,y))$ factorization.

For instance, in the figure 1, $T_{min}(path(jeep,object)) =$

$$\left\{ \begin{array}{l} \{jeep, is-a, car\}, \\ \{car, has-part, hood\}, \\ \{hood, is-a, cover, is-a, object\} \end{array} \right\}$$

Note that $\{jeep, is-a, car\}$ and $\{hood, is-a, cover, is-a, object\}$ are two disjoint sub-paths in $T_{min}(path(x,y))$ even if they have the same link type, because they don't have any common nodes.

We finally suggest to define the weight of the path as the sum of the weights of all sub-paths of the minimal factorization. Hence, the weight of the mixed path $path(x,y)$ is then defined as the sum of all sub-paths of T_{min} :

$$W(path(x,y)) = \sum_{p \in T_{min}(path(x,y))} W(p)$$

For instance, let consider the preceding path $path(jeep,object)$ with⁴ $IC(jeep) = 1.0$, $IC(car) = 0.68$, $IC(hood) = 1.0$, $IC(cover) = 0.55$, $IC(object) = 0.08$ and the relation factor $TC_{has-part} = 0.4$. The weights of the three single-relation paths are:

³ Proof by contradiction: Assume that in $T_{min}(path(x,y))$ with n sub-paths, $\exists i \in [1, n - 1], X_i = X_{i+1}$. This would mean that $path_{X_i}(z_i, z_{i+1}) \oplus path_{X_{i+1}}(z_{i+1}, z_{i+2})$ is a single-relation sub-path. Thus, there exists a factorization of $path(x,y)$ with $n - 1$ single-relation sub-paths, which is smaller than the minimal factorization $T_{min}(path(x,y))$. This is a contradiction. Therefore, in $T_{min}(path(x,y))$, we have the property $\forall i \in [1, n - 1], X_i \neq X_{i+1}$.

⁴ The five IC values are real values computes with Seco algorithm [21] and WordNet 3.0.

- $W(\{jeep, is-a, car\}) = |IC(jeep) - IC(car)| = 0.32$
- $W(\{car, has-part, hood\}) = 0.4 \times \frac{1}{1+1} = 0.2$
- $W(\{hood, is-a, cover, is-a, object\}) = |IC(hood) - IC(object)| = 0.92$

And finally, we sum these weights and we obtain that the weight of this path from *jeep* to *object* is $W(path(jeep, object)) = 0.32 + 0.2 + 0.92 = 1.44$.

3.3 Final Measure

To compute the semantic distance between two concepts, we consider only the semantically correct paths between these two concepts and we will select the best one.

We chose to use the Hirst & St-Onge rules (see section 2.2) to filter the semantically correct paths. Let us consider two concepts c_1 and c_2 . We note $\pi(c_1, c_2)$ the set of acyclic paths between c_1 and c_2 and $HSEO : \pi(c_1, c_2) \rightarrow \mathbb{B}$ the function such that $HSEO(p)$ is *true* if and only if p is a valid path *w.r.t.* Hirst & St-Onge patterns. Our semantic distance between c_1 and c_2 is then defined as follows:

$$dist(c_1, c_2) = \min_{\{p \in \pi(c_1, c_2) | HSEO(p) = true\}} W(p)$$

Let consider again the example $path(jeep, object)$. The path using the *part-of* relation gives a weight of 1.44. The hierarchical path between *jeep* and *object* is (figure 1) $\{jeep, is-a, car, is-a, wheeled_vehicle, is-a, vehicle, is-a, object\}$ and its weight is $|IC(jeep) - IC(object)| = 0.92$. Thus, in this example, the relational path does not give more information than the hierarchical path and the final semantic score obtained is: $dist(jeep, object) = 0.92$. It is an expected result, since *object* is a very general concept and does not contain relations with other concepts. On the contrary, if we consider the distance from *jeep* to *hood* the relational path $\{jeep, is-a, car, has-part, hood\}$ correspond to a weight of 0.52 whereas the hierarchical path (using the *ccp object*) correspond to a weight of 1.50. In this last example, the relation path gives the final result: $dist(jeep, hood) = 0.52$.

Note that, *by construction*, the hierarchical path between c_1 and c_2 (corresponding to the semantic similarity measure) is a semantically correct path and has value $dist_{JC}(c_1, c_2)$.⁵ Thus, since $dist(c_1, c_2)$ is the minimum value considering all semantically correct paths, $dist(c_1, c_2) \leq dist_{JC}(c_1, c_2)$. As a consequence, if IC is bounded between $[0, IC_{max}]$ (which is the case with the Seco formula [21], with $IC_{max} = 1.0$), the weight of a path can be bounded in $[0, 2 \times IC_{max}]$ ($2 \times IC_{max}$ being the maximal value of the Jiang & Conrath distance). Thus, our distance measure can, if necessary, be linearly transformed into a relatedness measure [1,23]:

$$rel(c_1, c_2) = 2 \times IC_{max} - dist(c_1, c_2)$$

4 Evaluation

The purpose of our evaluation is to show the relevancy of our hypothesis on weight and path-validity for semantic relatedness measure. In this section, we present the evaluation on two different benchmarks using WordNet relations and several values for TC_X .

⁵ Our measure can be seen as a relatedness generalisation of the Jiang & Conrath similarity measure.

4.1 Implementation

The Miller & Charles test [26] is well-known for semantic measure evaluation (e.g. [3,23,22,1]). This test is composed of 30 couples of words. For each couple of words, a significant number of persons had associated a value between 0 and 4 of “synonymy judgment”. This gives us a vector of 30 semantic rates which can be used as a benchmark for semantic computation. Testing a semantic measure simply consists of computing the correlation factor (usually the Pearson-product moment correlation factor) between this Miller & Charles vector and the vector generated by the computer using the semantic measure.

However, since this test was defined for “synonymy judgment”, its accuracy to test a semantic relatedness measure is not evident. Most of the chosen couples do not have any functional relationship and the subjects were explicitly asked to evaluate the *synonymy* between the words. Only a few couple (such as “journey-car” or “furnace-stove”) appears to have a possible non-hierarchical relation. It is interesting to note that in known similarity evaluations (as [3]), these couples are the ones which mostly differ from the human reference. However, since “semantic similarity represents a special case of semantic relatedness”, a semantic relatedness measure should at least work on the Miller & Charles test, even if it does not allow to validate the relational aspects of the measure. This is the reason why we first evaluate our measure on the Miller & Charles benchmark. The human reference for the Miller & Charles test was taken in [3].

Table 1. Pearson product-moment correlation factor for Rada, Resnik, Lin, Jiang & Conrath, Hirst & St-Onge and our approach ($TC_X = 0.4$)

Measures	Correlation	
	Miller & Charles	WordSimilarity-353
Rada	0.638	0.249
Resnik	0.804	0.375
Lin	0.836	0.377
Jiang & Conrath	0.880	0.362
Hirst & St-Onge	0.847	0.380
Our measure, $TC_X = 0.4$	0.902	0.400

Then, we needed another benchmark with to test the relational part of our measure. We have found the WordSimilarity-353⁶ test [27], that was essentially constructed with couples of words which are relationally connected (e.g. “computer-keyboard”, “telephone-communication”, etc.). Classical similarity measures logically tend to fail this test (i.e. their correlation factors are very low). We can expect that our measure will outperform similarity measure, but we will see that the choice of the ontology can limit this impact.

To compute our semantic score, we will consider as an ontology the *noun* sub-part of the lexical base WordNet 3.0 [25], since this knowledge base is easily accessible and

⁶ <http://www.cs.technion.ac.il/~gabr/resources/data/wordsim353/wordsim353.html>

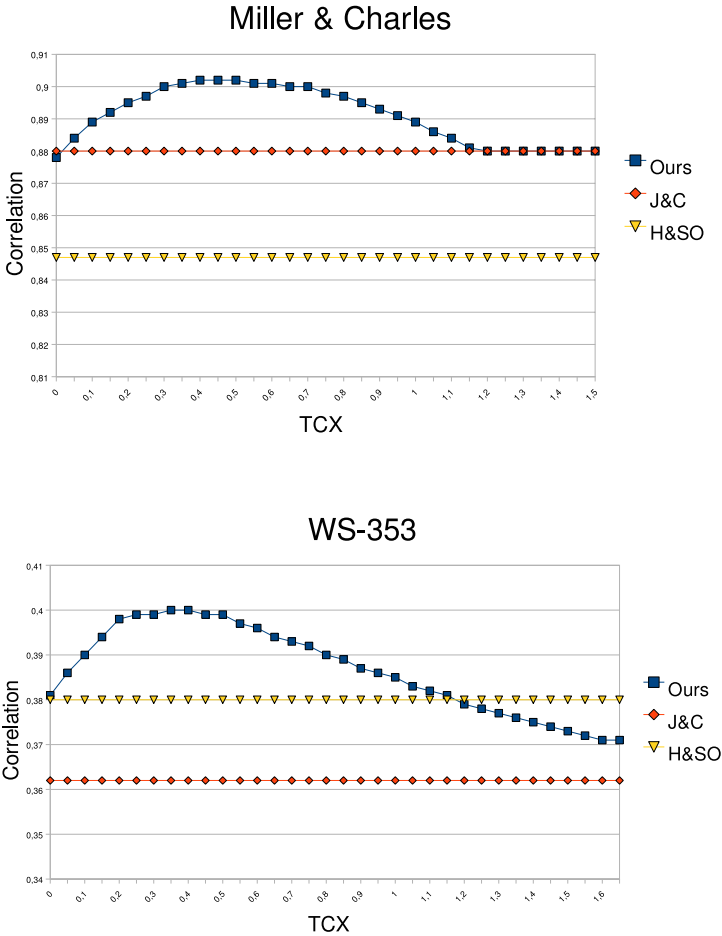


Fig. 3. Correlation factor evolution regarding to the TC_X factor in Miller & Charles and WS-353 test. The maximal value is obtained for $TC_X = 0.4$ in both test.

contains a large number of concepts. For the WS-353 test, we have removed 9 couples of words that did not exist in the *noun* part of WordNet (e.g. the couple “fighting-defeating”). Because of WordNet relation definition, we will consider only the *part-of* relations for non-hierarchical relations⁷. Moreover, we chose to consider only one fixed maximal weight TC_X for all the 6 *part-of* relations of WordNet.

Since we cannot anticipate the correct value for this $TC_{part-of}$, we evaluated all values from 0 (free non-hierarchical links) to 1.5 (very costly non-hierarchical links) by steps of 0.05.

⁷ In WordNet, it is separated in three meronym relations (*meronym_member*, *meronym_part*, *meronym_substance*) and three holonym relations (*holonym_member*, *holonym_part*, *holonym_substance*).

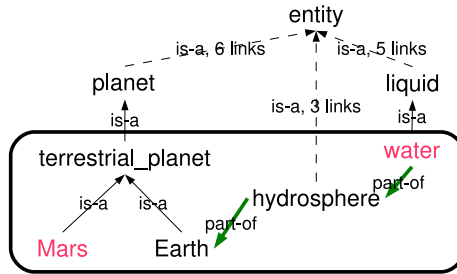


Fig. 4. Example of semantically bad path for Hirst & St-Onge patterns. From “water” to “Mars”, the shortest path in a graph-based view is not semantically correct since the last link (*terrestrial_planet, includes, Mars*) is not authorised.

The information content IC of a given node is computed using the Seco formula [21], which does not require a corpus to give a good representation of IC and is bounded by $[0, 1]$. We consider for correlation comparison 4 classical similarity measure (Rada [2], Resnik [1], Lin [22] and Jiang & Conrath [23]) and one relatedness measure (Hirst & St-Onge, see section 2.2).⁸

The results are given in the table 1. The evolution of the correlation factor regarding to the value of TC_X is given in the figure 3.

4.2 Discussions

In both tests, our measure outperforms the correlation of the others measures. Moreover, to our knowledge, it is the first time that a semantic measure based on WordNet reaches a correlation of 0.4 for the WS-353 test (see Strube work [5] for the last known evaluation).

Since, the Miller & Charles test is based on similarity, the major part of couples use a hierarchy-only path and, thus, our result corresponds to the Jiang & Conrath result. However, some results are different with the use of non-hierarchical link. For instance, if we consider the “furnace-stove” couple, the common closest concept is “artifact”, which makes a weak relation. Using relation, our measure was able to find the path “furnace *has-part* grate *part-of* stove” which uses the functional common property between the two concepts. This path is also identified by the Hirst & St-Onge measure, but the Hirst & St-Onge suffers from the uniformity hypothesis for the edge weight. Our measure takes advantages of the two paradigms: the semantically correct path for semantic relatedness and the information-theoretic approach to refine result.

Also note that no relevant link for the “journey-car” couple was found, neither by our measure nor by the Hirst & St-Onge one. This comes from the fact that the “common-sense relation” between these two concepts is not expressed in terms of meronym/holonym relations as defined in WordNet. This underlines the lack of relation types in WordNet, which is a limitation to computes semantic relatedness.

⁸ To be able to compare correlation factors, we have recalculated them for all these measures with WordNet 3.0.

As expected, relatedness measures (Hirst & St-Onge and our) do better than similarity measure on the WS-353 test, since this test correspond to couples of words which are connected relationally. Relatedness measures can find relational functions between words (*e.g.* “keyboard-computer”, etc.) whereas similarity measures only consider the hierarchy. In addition, the Hirst & St-Onge patterns permits to invalidate some path which are very shorter than the hierarchical, but incorrect in a semantic point of view (figure 4). However, some couples were not connected in WordNet and, thus, by our measure (*e.g.* “telephone-communication”, etc.). This can be explained, again, by the lack of relations types in WordNet. Moreover, as stated by Strube [5], the WS-353 test contains many couples which are connected by common-sense link and that cannot be connected in WordNet (*e.g.* “popcorn-movie”, etc.). Then, it explains why it is very difficult to obtain a good correlation using WordNet for this test. For this reason, we believe that it will be very difficult to go beyond the 0.35-0.4 limit on the WS-353 test using only WordNet as an ontology.

5 Conclusion and Future Work

In this paper, we have presented a new semantic measure to evaluate the semantic relatedness between two concepts. This measure makes use of the Hirst & St-Onge [8] patterns for semantically correct paths and the information-theoretic paradigm introduced by Resnik [1]. We have implemented our hypotheses on WordNet with the Miller & Charles [26] and the WS-353 test [27] and have shown that our measure outperforms known measures.

Note that by construction, our relatedness measure is always higher than the Jiang & Conrath similarity measure. Thus, we should always “fail” on couples with a J&C result higher than the human reference. However, one can hardly conceive an ontology that contains a relation between two concepts that are not associated from the human point of view. This would mean that this knowledge base would not be consistent with the domain. Moreover, if these links yet exists, we can study for a new weight allocation to invalidate it.

In addition, the evaluation underlines the lack of non-hierarchical relations in WordNet, as first mentioned in [8]. For instance, in WordNet, there is no relational path between concepts like “journey-car” or “telephone-communication”. This allows us to conclude that to use the capabilities of a semantic relatedness measure on ontologies, we need a *real* domain ontology. For this reason, our next aim is to test the impact of our measure on the performance rate of our semantic heterogeneity management system for multi-agent system [28]. Firstly, this measure will be used to enhance the alignment of the two agent’s ontologies, and secondly will be used for semantic interpretation of requests exchanged. This kind of result will allow us to conclude for the scalability of our approach in different applications.

Another open issue is the allocation of the weight for each link type. In our evaluation, we used a single TC_X and the best results are obtained in both tests for $TC_X = 0.4$. However, there is no proof that this value will always lead to the best result, nor than having one single TC_X for all edges is appropriate. One idea could be to query a domain expert to fix an initial set of value the weights TC_X . The expert should know what

kind of relation is important for semantic proximity. Since we consider the application of our semantic heterogeneity system in semantic interpretation of natural language commands (considering that an user is a special agent [29]), we think that the user feedback can be used as a background knowledge for a reinforcement learning algorithm [30,31] on the weight evolution. When the user confirms the system's interpretation of the command (*i.e.* the selected path in the ontology), the TC_X factor on the concerned edges will decrease. On the contrary, if the user denies the command, the TC_X factor will increase. This algorithm, currently in study, should be in charge to learn the optimal weight (or to accurate the expert weight) of the TC_X function, regarding to a given ontology.

Our final objective is to propose and to evaluate a measure of semantic similarity for more complex language of knowledge representation. For instance, Hau & al [15] has proposed a similarity measure based upon the Lin measure template [22] which considers OWL restriction, cardinality, intersection, etc. Alas, this measure was neither evaluated nor implemented and remains a theoretical proposition. We believe that, based on such work, it is however possible to extend our measure to model specific relations between concepts, such as intersection or disjunctive classes.

References

1. Resnik, P.: Using information content to evaluate semantic similarity in a taxonomy. In: 14th International Joint Conference on Artificial Intelligence (IJCAI 2005), pp. 448–453 (1995)
2. Rada, R., Mili, H., Bicknell, E., Blettner, M.: Development and Application of a Metric on Semantic Nets. *IEEE Transactions on Systems, Man, and Cybernetics* 19, 17–30 (1989)
3. Budanitsky, A., Hirst, G.: Evaluating wordnet-based measures of semantic distance. *Computational Linguistics* 32, 13–47 (2006)
4. Patwardhan, S., Pedersen, T.: Using WordNet Based Context Vectors to Estimate the Semantic Relatedness of Concepts. In: Proc. of the EACL 2006 Workshop Making Sense of Sense - Bringing Computational Linguistics and Psycholinguistics Together (2006)
5. Strube, M., Ponzetto, S.: WikiRelate! Computing semantic relatedness using Wikipedia. In: Proc. of AACL, vol. 6, pp. 1419–1424 (2006)
6. Cilibrasi, R., Vitanyi, P.: Automatic Extraction of Meaning from the Web. In: Proc. IEEE International Symposium on Information Theory, pp. 2309–2313 (2006)
7. Iosif, E., Potamianos, A.: Unsupervised Semantic Similarity Computation using Web Search Engines. In: International Conference on Web Intelligence (WI), pp. 381–387. IEEE, Los Alamitos (2007)
8. Hirst, G., St-Onge, D.: Lexical chains as representation of context for the detection and correction malapropisms. In: Fellbaum, C. (ed.) *WordNet: An Electronic Lexical Database*, pp. 305–332. MIT Press, Cambridge (1998)
9. Corby, O., Dieng-Kuntz, R., Faron-Zucker, C.: Querying the Semantic Web with the CORESE search engine. In: Press, I. (ed.) Proc. of the ECAI 2004, Valencia, pp. 705–709 (2004)
10. Milward, D., Beveridge, M.: Ontology-based dialogue systems. In: Proc. 3rd Workshop on Knowledge and reasoning in practical dialogue systems (IJCAI 2003), pp. 9–18 (August 2003)
11. Dzikovska, M.O., Allen, J.F., Swift, M.D.: Integrating linguistic and domain knowledge for spoken dialogue systems in multiple domains. In: Proc. of IJCAI 2003 Workshop on Knowledge and Reasoning in Practical Dialogue Systems (2003)

12. Euzenat, J., Shvaiko, P.: *Ontology matching*. Springer, Heidelberg (DE) (2007)
13. Maedche, A., Staab, S.: Measuring similarity between ontologies. In: Gómez-Pérez, A., Benjamins, V.R. (eds.) *EKAW 2002*. LNCS (LNAI), vol. 2473, pp. 251–263. Springer, Heidelberg (2002)
14. Eliasson, K.: Case-Based Techniques Used for Dialogue Understanding and Planning in a Human-Robot Dialogue System. In: *Proc. of IJCAI 2007*, pp. 1600–1605 (2007)
15. Hau, J., Lee, W., Darlington, J.: A Semantic Similarity Measure for Semantic Web Services. In: *Proc. Workshop on Web Service Semantics (2005)*
16. Smith, M.K., Welty, C., McGuinness, D.L.: Owl web ontology language guide (February 2004), <http://www.w3.org/TR/owl-guide/>
17. Turney, P.: Similarity of Semantic Relations. *Computational Linguistics* 32(3), 379–416 (2006)
18. Sussna, M.: Word Sense Disambiguation for Free-text Indexing Using a Massive Semantic Network. In: Bhargava, B.K., Finin, T.W., Yesha, Y. (eds.) *Proc. of the 2nd International Conference on Information and Knowledge Management (CIKM 1993)*, Washington, DC, USA, November, pp. 67–74. ACM, New York (1993)
19. Wu, Z., Palmer, M.: Verb semantics and lexical selection. In: *32nd. Annual Meeting of the Association for Computational Linguistics*, New Mexico State University, Las Cruces, New Mexico, pp. 133–138 (1994)
20. Zhong, J., Zhu, H., Li, J., Yu, Y.: Conceptual graph matching for semantic search. In: *ICCS 2002. Proceedings of the 10th International Conference on Conceptual Structures*, London, UK, pp. 92–196. Springer, Heidelberg (2002)
21. Seco, N., Veale, T., Hayes, J.: An Intrinsic Information Content Metric for Semantic Similarity in WordNet. In: *Proc. ECAI 2004, the 16th European Conference on Artificial Intelligence*, pp. 1089–1090 (2004)
22. Lin, D.: An information-theoretic definition of similarity. In: *Proc. 15th International Conf. on Machine Learning*, pp. 296–304. Morgan Kaufmann, San Francisco, CA (1998)
23. Jiang, J., Conrath, D.: Semantic similarity based on corpus statistics and lexical taxonomy. In: *Proc. on International Conference on Research in Computational Linguistics*, Taiwan, pp. 19–33 (1997)
24. Aleksovski, Z., ten Kate, W., van Harmelen, F.: Exploiting the structure of background knowledge used in ontology matching. In: *Proc. Workshop on Ontology Matching in ISWC 2006, CEUR Workshop Proceedings (2006)*
25. Fellbaum, C. (ed.): *WordNet, An Electronic Lexical Database*. MIT Press, Cambridge (1998)
26. Miller, G., Charles, W.: Contextual correlates of semantic similarity. *Language and Cognitive Processes* 6(1), 1–28 (1991)
27. Finkelstein, L., Gabrilovich, E., Matias, Y., Rivlin, E., Solan, Z., Wolfman, G., Ruppin, E.: Placing search in context: the concept revisited. In: *WWW 2001. Proceedings of the 10th international conference on World Wide Web*, pp. 406–414. ACM Press, New York (2001)
28. Ferber, J.: *Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA (1999)
29. Mazuel, L., Sabouret, N.: Generic command interpretation algorithms for conversational agents. *Web Intelligence and Agent Systems* 6(2) (April 2008)
30. Sutton, R.: Learning to predict by the methods of temporal differences. *Machine Learning* 3(1), 9–44 (1988)
31. Watkins, C., Dayan, P.: Q-learning. *Machine Learning* 8(3), 279–292 (1992)