A Comparative Assessment of Ontology Weighting Methods in
Semantic Similarity Search

Antonio De Nicola1, Anna Formica2, Michele Missikoff2, Elaheh Pourabbas2 and Francesco Taglino2

1Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA),
Casaccia Research Centre, Via Anguillarese 301, I-00123, Rome, Italy
2Istituto di Analisi dei Sistemi ed Informatica (IASI) “Antonio Ruberti”, National Research Council,
Via dei Taurini 19, I-00185, Rome, Italy

Keywords: Weighted Reference Ontology, Semantic Similarity, Information Content, Probabilistic Approach.

Abstract: Semantic search is the new frontier for the search engines of the last generation. Advanced semantic search
methods are exploring the use of weighted ontologies, i.e., domain ontologies where concepts are associated
with weights, inversely related to their selective power. In this paper, we present and assess four different
ontology weighting methods, organized according to two groups: intensional methods, based on the sole ontology
structure, and extensional methods, where also the content of the search space is considered. The
comparative assessment is carried out by embedding the different methods within the semantic search engine
SemSim, based on weighted ontologies, and then by running four retrieval tests over a search space we have
previously proposed in the literature. In order to reach a broad audience of readers, the key concepts of this
paper have been presented by using a simple taxonomy, and the already experimented dataset.

1 INTRODUCTION

Search engines represent today the killer application of the Web and can be found in every and all possible Web applications. For instance, if you need to find a place on Google Maps, or you are looking for a friend on Facebook, or you want to discover the last song of your preferred singer on YouTube or Spotify, you always go through a search facility. Since the first appearance of general purpose search engines on the Web, such as Yahoo! and AltaVista in the Nineties, followed a few years later by Google and, almost a decade afterwards, by Bing (just to name the popular ones), their technology has been constantly evolving. Such an evolution brought continuous enhancements of search strategies, algorithms, and, last but not least, indexes, directories, vocabularies, and other supporting metadata. Among metadata, semantic annotation has emerged as an important enrichment of digital resources, necessary to support the evolution of search engines towards semantic similarity search. A semantic annotation consists of a set of concepts, taken from an ontology, that characterize a resource. In (Formica et al., 2008), (Formica et al., 2013), (Formica et al., 2016), the authors addressed the semantic annotation and retrieval in accordance to a probabilistic approach, based on a Vector Space Model proposed in the context of text mining and retrieval, where text documents are represented by feature vectors. In our case, we deal with any kind of digital resources (not only text documents), and the features that characterize a resource correspond to concepts in a reference ontology. Therefore we refer to such a vector of features as an Ontology Feature Vector (OFV). The adoption of ontologies is the base of semantic search, representing a marked evolution from the traditional keyword based retrieval methods. In an ontology based search engine, the matchmaking process can take place between a user request vector and the annotation vectors associated with the digital resources in the search space. A significant enhancement of semantic search consists in the use of probabilistic similarity reasoning methods. Within these approaches, concept similarity is computed considering the contextual knowledge represented by the ontology, with its (topo)logical structure (essentially, the ISA hierarchy). This approach requires each concept in the ontology be associated with a weight related to the level of specificity of the concept in the resource space. The introduction of concept weights yields a new breed of weighted ontologies, see for instance (Abioui et al., 2018), (Sánchez et al., 2011).
majority of them share the idea that the weight of a concept corresponds to the probability that selecting at random a resource, it is characterized by a set of features including one representing such a concept, or one of its descendants in the ontology. Then, the higher the weight of a concept the lower its specificity. For instance, the concept student has a smaller weight than person since the former is more specific than the latter. Therefore, in formulating a query, the lower the weights of the concepts, the higher their selective power, and a more focused answer set is returned.

The performance of a semantic search engine depends on the semantic matchmaking method and the approach used to weigh the reference ontology. In this paper, we focus on the analysis of four different approaches for weighting the concepts of an ontology, and we carry out an experiment in order to assess the analyzed ontology weighting methods.

The presented methods are divided according to two groups (Sánchez et al., 2011): (i) extensional methods (also known as distributional methods), where the concept weights are derived by taking into account both the topology of the ISA hierarchy and the content of the resource space, also referred to as dataset, (ii) intensional methods (also known as intrinsic methods), where the concept weights are derived on the basis of the sole topology of the ISA hierarchy.

In this paper, we selected the semantic similarity method SemSim (Formica et al., 2013) in order to evaluate the assessment of the four methods. In the mentioned paper, the authors illustrate that SemSim outperforms the most representative similarity methods proposed in the literature, i.e., Dice, Cosine, Jaccard, and Weighted Sum. The SemSim method requires: i) a dataset consisting of a set of resources annotated according to a given ontology, and ii) a method for associating weights with the concepts of the ontology. Then, SemSim has been conceived to compute the semantic similarity between a given user request and any annotated resource in the dataset. With respect to this work, in the mentioned paper we considered only two weighting methods, i.e., the frequency and the probabilistic approaches. In this paper, they correspond to the Annotation Frequency Method and the Top Down Topology Method, respectively. Note that, in order to be coherent with the results given in (Formica et al., 2013), in this paper we keep the same experimental setting, in particular, the reference ontology and the dataset presented in the mentioned work.

The next section gives a brief overview about ontology weighting. Section 3 provides the basic notions concerning weighted ontologies and ontology based feature vectors and proposes a probabilistic model for weighted ontologies. Section 4 describes in detail the four methods. Section 5 illustrates the assessment of the methods and, finally, Section 6 concludes.

2 RELATED WORK

According to the extensional methods, also referred to as distributional (Sánchez et al., 2011), the information content of a concept is in general estimated from the frequency distribution of terms in text corpora. Hence, this type is based on the extensional semantics of the concept itself as its probability can be derived on the basis of the number of occurrences of the concept in the text corpora. This approach was used in (Jiang and Conrath, 1997), (Resnik, 1995), and (Lin, 1998) to assess semantic similarity between concepts. Other proposals include the inverse document frequency (IDF) method, and the method based on the combination of term frequency (TF) and the IDF (Manning et al., 2008). In our work, we derived the concept frequency method and the annotation frequency method, respectively, from those used in (Resnik, 1995) and the IDF.

According to the intensional methods, also referred to as intrinsic (Sánchez et al., 2011), information content is computed starting from the conceptual relations existing between concepts and, in particular, from the taxonomic structure of concepts. With this regard, one of the most relevant methods is presented in (Seco et al., 2004). This is based on the number of concepts’ hyponyms and the maximum number of concepts in the taxonomy. In (Meng et al., 2012), the authors present a method derived from (Seco et al., 2004) but they also consider the degree of generality of concepts and, hence, their depth in the taxonomy. In (Sánchez et al., 2011), the authors claim that the taxonomical leaves are enough to describe and differentiate two concepts because ad-hoc abstractions (e.g., abstract entities) rarely appear in a universe of discourse, but have an impact on the size of the hyponym tree. In (Hayuhardhika et al., 2013), the authors propose to use the density factor to estimate concept weights on the basis of the sum of inward and outward connections with other concepts against the total number of connections in the ontology. Finally, just to mention one more example, (Abiou et al., 2018) takes into account both the taxonomic structure and other semantic relationships to compute weights of concepts.

In this work, first of all we focus on a tree-shaped taxonomy organized as an ISA hierarchy and, within
the above mentioned classification, we investigate two extensional and two intensional methods. In particular, with regard to the extensional methods, we address semantic annotations of resources rather than text corpora.

3 A WEIGHTED ONTOLOGY AS A PROBABILISTIC MODEL

In line with (Formica et al., 2013), (Formica et al., 2016), an ontology Ont is a taxonomy defined by the pair:

\[ Ont = < C, \text{ISA} > \]

where \( C = \{K_i\} \) is a set of concepts and ISA is the set of pairs of concepts in \( C \) that are in subsumption (\( \text{subs} \)) relation:

\[ \text{ISA} = \{(K_i, K_j) \in C \times C | \text{subs}(K_i, K_j)\} \]

where \( \text{subs}(K_i, K_j) \) means that \( K_i \) is a child of \( K_j \) in the taxonomy. In this work, we assume that the hierarchy is a tree. A Weighted Reference Ontology (WRO) is then defined as follows:

\[ WRO = < Ont, w > \]

where \( w \), the concept weighting function, is a probability distribution defined on \( C \), such that given \( K \in C \), \( w(K) \) is a decimal number in the interval [0...1].

The WRO is then used to annotate each resource in the Universe of Digital Resources (UDR) by means of an OFV. An OFV is a vector that gathers a set of concepts of the ontology \( Ont \), aimed at capturing the semantic content of the corresponding resource. The same also holds for a user request, and is represented as follows:

\[ ofv = (K_1, \ldots, K_n) \]

where \( C_i = C_i \in C, i = 1, \ldots, n \)

A normalized OFV is an OFV where if a concept appears, none of its ancestors appears. Note that, when an OFV is used to represent a user request, it is referred to as semantic Request Vector (RV) whereas, if it is used to represent a resource, it is referred to as semantic Annotation Vector (AV). They are denoted, respectively, as follows:

\[ rv = (R_1, \ldots, R_n), \ AV = (A_1, \ldots, A_m) \]

where \( \{R_1, \ldots, R_n\} \cup \{A_1, \ldots, A_m\} \subseteq C \). We assume that also AVs and RVs are normalized OFVs.

In the following, consider an ontology \( Ont = < C, \text{ISA} > \) and a dataset defined as a set of annotated resources, where different resources can also have the same annotations. For each \( K_i \in C \), let \( X_{K_i} \) be a boolean variable, where \( 1 \leq i \leq q \) and \( q = |C| \). According to the semantics of the ISA relationship, we assume that the set of variables associated with the concepts of the ontology are dependent. Each annotation \( AV = (A_1, \ldots, A_m) \) in the dataset can also be represented as:

\[ [X_{A_1} = 1, \ldots, X_{A_m} = 1] \]

Analogously, any OFV can also be represented according to the above notation.

Table 1: Simple dataset.

<table>
<thead>
<tr>
<th>Resource</th>
<th>Annotation Vector</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( AV_1 = (A, B) )</td>
</tr>
<tr>
<td></td>
<td>( AV_2 = (C) )</td>
</tr>
<tr>
<td></td>
<td>( AV_3 = (B) )</td>
</tr>
<tr>
<td></td>
<td>( AV_4 = (C, D) )</td>
</tr>
</tbody>
</table>

In order to better illustrate this point, let us consider the very simple taxonomy shown in Figure 1. According to this taxonomy, we have the following boolean variables: \( X_T, X_A, X_B, X_C, X_D \), corresponding to the concepts \( T, A, B, C, D \), respectively. For example,
the variables $X_C$ and $X_A$ are dependent because $C$ is a child of $A$. Therefore $X_C = 1$ implies $X_A = 1$, according to the semantics of the ISA hierarchy. Furthermore, with regard to the dataset, we assume the UDR is composed by the four resources $r_1$, $r_2$, $r_3$, and $r_4$, annotated as shown in Table 1. According to the notation given in (1), for instance $av_1 = (A, B)$ can also be represented as $[X_A = 1, X_B = 1]$. In the AF method, given a concept $K$, its relative frequency is the number of occurrences of $K^+$ divided by the number of occurrences of all concepts in the set of all annotation vectors (AVs). In formal terms, we have:

$$p(K) = \frac{n(K^+)}{N} \quad (4)$$

where $n(K^+)$ is the total number of occurrences of the concepts in $K^+$ ($K$ and its descendants in the taxonomy, as defined previously), and $N$ is the number of occurrences of all the concepts in the AVs.

Therefore, the bag of all possible outcomes $S$ is formed by all the occurrences of the concepts in the AVs defined in the dataset, and an event $<X_K = 1>_S$ corresponds to the occurrences of the concept $K$ and its descendants in $S$.

Let us consider the running example, defined according to Figure 1 and Table 1. In this case, the set $S$ is defined as follows:

$$S = \{[X_A = 1], [X_B = 1], [X_C = 1], [X_D = 1], [X_C = 1], [X_D = 1]\}.$$  

For instance, consider the event $<X_A = 1>_S$. We have:

$$<X_A = 1>_S = \{[X_A = 1], [X_C = 1], [X_C = 1], [X_D = 1]\}.$$  

As a result, according to Eq. (2), we have:

$$p(A) = p(<X_A = 1>_S) = 4/6 = 2/3.$$  

Similarly, in the other cases:

$$p(T) = p(<X_T = 1>_S) = 1/3 \quad p(B) = p(<X_B = 1>_S) = 1/3 \quad p(C) = p(<X_C = 1>_S) = 1/3 \quad p(D) = p(<X_D = 1>_S) = 1/6.$$  

Annotation Frequency Method (AF). The AF method is also referred to as frequency in (Formica et al., 2013). In the AF method, given a concept $K$, its relative frequency is the number of annotation vectors
containing \( K \), or a descendant of it, divided by the total number of annotation vectors. Therefore we have:

\[
p(K) = \frac{|A^+_K|}{|A^V|} \tag{5}
\]

where \( A^V \) is the set of all the annotation vectors in the dataset, and \( A^+_K \) is the subset of \( A^V \) containing the concept \( K \) or a descendant of it.

The bag of all possible outcomes \( S \) is represented by the bag of the outcomes corresponding to the \( AV \)s in the UDR, and an event \( < X_K = 1 >_S \) corresponds to the occurrences of the \( AVs \) containing a concept in \( K^+ \).

Consider the running example:

\[
S = \{ [X_A = 1, X_B = 1], [X_C = 1], [X_B = 1], [X_C = 1, X_D = 1] \}.
\]

For instance, in the case of the event \( < X_A = 1 >_S \) we have:

\[
< X_A = 1 >_S = \{ [X_A = 1, X_B = 1], [X_C = 1] \}.
\]

and:

\[
p(A) = p(< X_A = 1 >_S) = 3/4.
\]

Similarly, in the other cases, we have:

\[
p(T) = p(< X_T = 1 >) = 1,
\]

\[
p(B) = p(< X_B = 1 >) = 1/2,
\]

\[
p(C) = p(< X_C = 1 >) = 1/2,
\]

\[
p(D) = p(< X_D = 1 >) = 1/4.
\]

### 4.2 Intensional Methods

With respect to the previous methods, the intensional, or topology-based, methods illustrated in this section follow an axiomatic approach, and therefore do not require a dataset and a set of possible outcomes \( S \).

**Top-Down Topology-based Method (TD).** The TD method has been introduced in (Formica et al., 2008), and successively extensively experimented in (Formica et al., 2013) (where it has been referred to as probabilistic). Here, we briefly recall it for reader’s convenience. In order to compute the probabilities of concepts in the reference ontology, this method adopts a uniform probabilistic distribution along the ISA hierarchy following a top-down approach. In particular, the root of the hierarchy has the probability equal to 1, and the probability of a concept \( K \) of the ontology is computed as follows:

\[
p(K) = \frac{p(parent(K))}{|children(parent(K))|} \tag{6}
\]

In our running example, according to this approach, the probabilities of the concepts in Figure 1 are defined as follows:

\[
p(T) = 1, \quad p(A) = 1/2, \quad p(B) = 1/2
\]

\[
p(C) = 1/4, \quad p(D) = 1/4.
\]

**Intrinsic Information Content Method (IIC).** The IIC method is based on an axiomatic approach, which has been conceived in order to compute the information content of concepts (Seco et al., 2004). The authors define the information content of a concept in a taxonomy as a function of its descendants. In particular, they claim that the more descendants a concept has the less information it expresses. Therefore, concepts that are leaves are the most specific in the taxonomy, and their information is maximal.

Formally, they define the intrinsic information content (iic) of a concept \( K \) as follows:

\[
iic(K) = 1 - \frac{\log(|desc(K)| + 1)}{\log(|C|)} \tag{7}
\]

where the \( desc(K) \) is the set of the descendants of the concept \( K \), and \( C \) is the set of the concepts in \( Ont \). Note that the denominator assures that the iic values are in \([0, \ldots, 1]\). The above formulation guarantees that the information content decreases monotonically. Moreover, the root node of the taxonomy yields an information content value equal to 0.

For instance, consider the taxonomy shown in Figure 1. The information contents of the concepts are:

\[
iic(T) = 0, \quad iic(A) \equiv 1 - \frac{\log(2+1)}{\log(5)} = 0.32
\]

\[
iic(B) = 1, \quad iic(C) = 1, \quad iic(D) = 1.
\]

### 5 ASSESSMENT OF METHODS

In this section, in order to carry out an assessment of the four methods illustrated in the previous section, we first recall the SemSim method.

#### 5.1 SemSim

The SemSim method has been conceived to search for the resources in the resource space that best match the RV, by contrasting it with the various AV, associated with the searchable digital resources (Formica et al., 2013). This is achieved by applying the semsim function, which has been defined to compute the semantic similarity between OFV. In SemSim, the probabilities of concepts are used to derive the information content (iic) of the concepts that, according to (Lin, 1998), represents the basis for computing the concept similarity. In particular, according to the information theory, the iic of a concept \( K \) is defined as:

\[
iic(K) = -\log(w(K))
\]
Table 2: Annotation Vectors (dataset).

\[ \begin{align*}
\text{av}_1 &= (\text{InternationalHotel}, \text{FrenchMeal}, \text{Cinema}, \text{Flight}) \\
\text{av}_2 &= (\text{Pension}, \text{VegetarianMeal}, \text{ArtGallery}, \text{ShoppingCenter}) \\
\text{av}_3 &= (\text{CountryResort}, \text{MediterraneanMeal}, \text{Bus}) \\
\text{av}_4 &= (\text{CozyAccommodation}, \text{VegetarianMeal}, \text{Museum}, \text{Train}) \\
\text{av}_5 &= (\text{InternationalHotel}, \text{ThaiMeal}, \text{IndianMeal}, \text{Concert}, \text{Bus}) \\
\text{av}_6 &= (\text{SeasideCottage}, \text{LightMeal}, \text{ArcheologicalSite}, \text{Flight}, \text{ShoppingCenter}) \\
\text{av}_7 &= (\text{RegularAccommodation}, \text{RegularMeal}, \text{Salon}, \text{Flight}) \\
\text{av}_8 &= (\text{InternationalHotel}, \text{VegetarianMeal}, \text{Ship}) \\
\text{av}_9 &= (\text{FarmHouse}, \text{MediterraneanMeal}, \text{CarRental}) \\
\text{av}_{10} &= (\text{RegularAccommodation}, \text{EthnicMeal}, \text{Museum}) \\
\text{av}_{11} &= (\text{RegularAccommodation}, \text{LightMeal}, \text{Cinema}, \text{Bazaar}) \\
\text{av}_{12} &= (\text{SeasideCottage}, \text{VegetarianMeal}, \text{Shopping}) \\
\text{av}_{13} &= (\text{Campsite}, \text{IndianMeal}, \text{Museum}, \text{RockConcert}) \\
\text{av}_{14} &= (\text{RegularAccommodation}, \text{RegularMeal}, \text{Museum}, \text{Bazaar}) \\
\text{av}_{15} &= (\text{InternationalHotel}, \text{PictureGallery}, \text{Flight}) \\
\text{av}_{16} &= (\text{Pension}, \text{LightMeal}, \text{ArcheologicalSite}, \text{CarRental}, \text{Flight}) \\
\text{av}_{17} &= (\text{AlternativeAccommodation}, \text{LightMeal}, \text{RockConcert}, \text{Bus}) \\
\text{av}_{18} &= (\text{CozyAccommodation}, \text{VegetarianMeal}, \text{Exhibition}, \text{ArcheologicalSite}, \text{Train}) \\
\text{av}_{19} &= (\text{CountryResort}, \text{VegetarianMeal}, \text{Concert}, \text{Bus}) \\
\text{av}_{20} &= (\text{Campsite}, \text{MediterraneanMeal}, \text{ArcheologicalSite}, \text{Attraction}, \text{CarRental}) \\
\text{av}_{21} &= (\text{AlternativeAccommodation}, \text{LightMeal}, \text{Concert}, \text{Bus}) \\
\text{av}_{22} &= (\text{FarmHouse}, \text{LightMeal}, \text{RockConcert}, \text{Train})
\end{align*} \]

The \( \text{semsim} \) function is based on the notion of similarity between concepts (features), referred to as \( \text{consim} \). Given two concepts \( K_i, K_j \), it is defined as follows:

\[
\text{consim}(K_i, K_j) = \frac{2 \times \text{IC}(\text{lab}(K_i, K_j))}{\text{IC}(K_i) + \text{IC}(K_j)}
\]

where the \( \text{lab} \) represents the least abstract concept of the ontology that subsumes both \( K_i \) and \( K_j \). Given an instance of \( RV \) and an instance of \( AV \), say \( rv \) and \( av \) respectively, the \( \text{semsim} \) function computes the \( \text{consim} \) for each pair of concepts belonging to the set formed by the Cartesian product of the \( rv \) and \( av \).

However, we focus on the pairs that exhibit high affinity. In particular, we adopt the exclusive match philosophy, where the elements of each pair of concepts do not participate in any other pair. The method aims to identify the set of pairs of concepts of the \( rv \) and \( av \) that maximizes the sum of the \( \text{consim} \) similarity values. In particular, given \( rv = \{R_1, \ldots, R_n\} \) and \( av = \{A_1, \ldots, A_m\} \) as defined in Section 3, let \( S \) be the Cartesian Product of \( rv \) and \( av \), i.e., \( S = rv \times av \), then, \( P(rv, av) \) is defined as follows:

\[
P(rv, av) = \{ P \in S \mid \forall (R_i, A_j), (R_k, A_k) \in P, R_i \neq R_k, A_j \neq A_k, |P| = \min\{n, m\} \}.
\]

Therefore, \( \text{semsim}(rv, av) \) is given by:

\[
\text{semsim}(rv, av) = \frac{\max_{P \in P(rv, av)} \{ |R_i, A_j| \in P \} \sum \text{consim}(R_i, A_j) \}}{\max\{n, m\}}.
\]
Table 5: Results of $\text{SemSim}$ about $rv_2$.  

<table>
<thead>
<tr>
<th>AV</th>
<th>Extensional</th>
<th>Intensional</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HJ</td>
<td>CF</td>
</tr>
<tr>
<td>$av_1$</td>
<td>0.72</td>
<td>0.59</td>
</tr>
<tr>
<td>$av_2$</td>
<td>0.21</td>
<td>0.41</td>
</tr>
<tr>
<td>$av_3$</td>
<td>0.16</td>
<td>0.24</td>
</tr>
<tr>
<td>$av_4$</td>
<td>0.10</td>
<td>0.34</td>
</tr>
<tr>
<td>$av_5$</td>
<td>0.10</td>
<td>0.39</td>
</tr>
<tr>
<td>$av_6$</td>
<td>0.20</td>
<td>0.36</td>
</tr>
<tr>
<td>$av_7$</td>
<td>0.71</td>
<td>0.67</td>
</tr>
<tr>
<td>$av_8$</td>
<td>0.10</td>
<td>0.36</td>
</tr>
<tr>
<td>$av_9$</td>
<td>0.10</td>
<td>0.23</td>
</tr>
<tr>
<td>$av_{10}$</td>
<td>0.40</td>
<td>0.30</td>
</tr>
<tr>
<td>$av_{11}$</td>
<td>0.10</td>
<td>0.29</td>
</tr>
<tr>
<td>$av_{12}$</td>
<td>0.30</td>
<td>0.10</td>
</tr>
<tr>
<td>$av_{13}$</td>
<td>0.10</td>
<td>0.19</td>
</tr>
<tr>
<td>$av_{14}$</td>
<td>0.44</td>
<td>0.30</td>
</tr>
<tr>
<td>$av_{15}$</td>
<td>0.86</td>
<td>0.69</td>
</tr>
<tr>
<td>$av_{16}$</td>
<td>0.25</td>
<td>0.40</td>
</tr>
<tr>
<td>$av_{17}$</td>
<td>0.10</td>
<td>0.35</td>
</tr>
<tr>
<td>$av_{18}$</td>
<td>0.10</td>
<td>0.28</td>
</tr>
<tr>
<td>$av_{19}$</td>
<td>0.10</td>
<td>0.34</td>
</tr>
<tr>
<td>$av_{20}$</td>
<td>0.10</td>
<td>0.28</td>
</tr>
<tr>
<td>$av_{21}$</td>
<td>0.10</td>
<td>0.36</td>
</tr>
<tr>
<td>$av_{22}$</td>
<td>0.10</td>
<td>0.32</td>
</tr>
<tr>
<td><strong>Corr</strong></td>
<td>1.00</td>
<td>0.81</td>
</tr>
</tbody>
</table>

5.2 Validation

In order to analyze the four methods illustrated in the previous sections, we refer to the experiment presented in (Formica et al., 2013). In that experiment, the taxonomy shown in Figure 2 has been considered, and four request vectors, namely $rv_i$, $i = 1, ..., 4$, which are recalled in Table 3. In the same experiment, 22 annotated resources have been defined, which are represented by their annotation vectors $av_1, av_2, ..., av_{22}$ as recalled in Table 2. In our approach they represent the dataset. In the experiment, the $\text{SemSim}$ values were computed against the 22 annotation vectors, and the correlation index ($\text{Corr}$) against human judgment (HJ) scores was calculated. The HJ scores were computed by asking to a group of 21 people to evaluate the similarity among each request vector and the annotation vectors defined in Table 2. In the same work, the authors demonstrated that the Annotation Frequency Method (AF) (referred to as frequency in the mentioned paper) outperforms some of the most representative similarity methods defined in the literature (i.e., Dice, Jaccard, Cosine, and Weighted Sum).

In our work, for each request vector, we apply $\text{SemSim}$ by using the four weighting methods illustrated above. In Tables 4, 5, 6, 7 the results about $rv_1$, $rv_2$, $rv_3$, $rv_4$ are shown. In particular, we observe that the AF method still achieves a higher correlation with HJ with respect to all the other considered methods, i.e.,

Table 6: Results of $\text{SemSim}$ about $rv_3$.  

<table>
<thead>
<tr>
<th>AV</th>
<th>Extensional</th>
<th>Intensional</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HJ</td>
<td>CF</td>
</tr>
<tr>
<td>$av_1$</td>
<td>0.10</td>
<td>0.30</td>
</tr>
<tr>
<td>$av_2$</td>
<td>0.62</td>
<td>0.73</td>
</tr>
<tr>
<td>$av_3$</td>
<td>0.29</td>
<td>0.34</td>
</tr>
<tr>
<td>$av_4$</td>
<td>0.10</td>
<td>0.36</td>
</tr>
<tr>
<td>$av_5$</td>
<td>0.10</td>
<td>0.34</td>
</tr>
<tr>
<td>$av_6$</td>
<td>0.31</td>
<td>0.49</td>
</tr>
<tr>
<td>$av_7$</td>
<td>0.10</td>
<td>0.38</td>
</tr>
<tr>
<td>$av_8$</td>
<td>0.10</td>
<td>0.30</td>
</tr>
<tr>
<td>$av_9$</td>
<td>0.12</td>
<td>0.34</td>
</tr>
<tr>
<td>$av_{10}$</td>
<td>0.18</td>
<td>0.39</td>
</tr>
<tr>
<td>$av_{11}$</td>
<td>0.78</td>
<td>0.79</td>
</tr>
<tr>
<td>$av_{12}$</td>
<td>0.38</td>
<td>0.45</td>
</tr>
<tr>
<td>$av_{13}$</td>
<td>0.10</td>
<td>0.35</td>
</tr>
<tr>
<td>$av_{14}$</td>
<td>0.42</td>
<td>0.58</td>
</tr>
<tr>
<td>$av_{15}$</td>
<td>0.10</td>
<td>0.24</td>
</tr>
<tr>
<td>$av_{16}$</td>
<td>0.31</td>
<td>0.42</td>
</tr>
<tr>
<td>$av_{17}$</td>
<td>0.10</td>
<td>0.44</td>
</tr>
<tr>
<td>$av_{18}$</td>
<td>0.18</td>
<td>0.35</td>
</tr>
<tr>
<td>$av_{19}$</td>
<td>0.10</td>
<td>0.41</td>
</tr>
<tr>
<td>$av_{20}$</td>
<td>0.22</td>
<td>0.38</td>
</tr>
<tr>
<td>$av_{21}$</td>
<td>0.10</td>
<td>0.45</td>
</tr>
<tr>
<td>$av_{22}$</td>
<td>0.10</td>
<td>0.42</td>
</tr>
<tr>
<td><strong>Corr</strong></td>
<td>1.00</td>
<td>0.85</td>
</tr>
</tbody>
</table>

6 CONCLUSION

In this paper, we presented a comparative assessment of the performances of four different methods for ontology weighting. The results of this work reveal that, in general, the extensional methods outperform
Table 7: Results of SemSim about rv4.

<table>
<thead>
<tr>
<th>RV</th>
<th>Extensional</th>
<th>Intensional</th>
</tr>
</thead>
<tbody>
<tr>
<td>HJ</td>
<td>0.10</td>
<td>0.36</td>
</tr>
<tr>
<td>CF</td>
<td>0.31</td>
<td>0.44</td>
</tr>
<tr>
<td>AF</td>
<td>0.38</td>
<td>0.38</td>
</tr>
<tr>
<td>TD</td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td>IIC</td>
<td>0.10</td>
<td>0.33</td>
</tr>
</tbody>
</table>

A Comparative Assessment of Ontology Weighting Methods in Semantic Similarity Search

Table 8: Summary of correlations.

<table>
<thead>
<tr>
<th>RV</th>
<th>Extensional</th>
<th>Intensional</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF</td>
<td>0.92</td>
<td>0.90</td>
</tr>
<tr>
<td>AF</td>
<td>0.81</td>
<td>0.83</td>
</tr>
<tr>
<td>TD</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>IIC</td>
<td>0.87</td>
<td>0.88</td>
</tr>
</tbody>
</table>

the intensional ones. Furthermore, among the extensional methods, the AF method exhibits the best correlation with human judgment. However, there are cases where the extensional methods may require more elaboration, e.g., when the resource space is highly dynamic, and then it is more appropriate to rely on intensional methods.

REFERENCES


