Semantic Nearest Neighbor Search in OWL Ontologies

Nicola Fanizzi, Claudia d'Amato, and Floriana Esposito

Dipartimento di Informatica, Università degli studi di Bari
Campus Universitario, Via Orabona 4, 70125 Bari, Italy
{fanizzi|claudia.damato|esposito}@di.uniba.it

Abstract. A *nearest neighbor* search procedure is presented, for retrieving resources in knowledge bases expressed in OWL. The procedure exploits a semi-distance for annotated resources, that is based on a number of dimensions corresponding to a committee of features represented by OWL concept descriptions. The procedure can retrieve resources belonging to query concepts expressed in OWL, by analogy with other training instances, on the grounds of the classification of the nearest ones w.r.t. the dissimilarity measure. Besides, it may also be able to suggest new assertions that are not logically entailed by the knowledge base due to open world semantics. In the experimentation, where we compare the performance of the procedure to running a reasoner, we show that it can be quite accurate and augment the scope of its applicability, improving w.r.t. previous prototypes that adopted other semantic measures.

1 Introduction

In the perspective of resource retrieval, purely logical approaches pursued so far, in the context of the Semantic Web, may fall short in terms of noise-tolerance and efficiency. Hence, *analogical* methods applied to multi-relational domains appear particularly well suited, since they are known to be more efficient and noise-tolerant, which is very important in contexts where knowledge is intended to be acquired from distributed sources. To this purpose, a relational distance-based framework for retrieving resources contained in semantic knowledge bases has been devised to infer inductively consistent class-membership assertions that may be not logically derivable due to the open-world semantics. The main idea is that similar individuals, by analogy, should likely belong to the extension of similar concepts.

Specifically, we present a retrieval procedure that constitutes a multi-relational extension [5] of the well-known *Nearest Neighbor* approach (henceforth, *NN*) [10]. These algorithms may be quite efficient because they require checking query-membership for a limited set of training instances on such concepts and making a decision on the classification of new instances.

From a technical viewpoint, extending the NN setting to work on multi-relational representations, such as concept languages like OWL, required suitable metrics whose definition could not be straightforward. In particular, a theoretical problem has been posed by the *Open World Assumption* (OWA) that is generally made in the target context, differently from typical databases settings where the *Closed World Assumption* (CWA) is the standard. Indeed the NN algorithms are devised for simple classifications where classes are assumed to be pairwise disjoint which is quite unlikely in the Semantic Web context.

As pointed out in [3], most of the existing measures focus on the similarity of atomic concepts within hierarchies or simple ontologies. Moreover they have been conceived for assessing *concept* similarity. Conversely, for our purposes, a notion of similarity between *individuals* is required. Recently, dissimilarity measures for specific description logics concept descriptions have been proposed [3, 4]. Although they turned out to be quite effective for the inductive tasks of interest, they are still partly based on structural criteria (a notion of normal form) which determine their main weakness: they are hardly scalable to deal with standard languages, such as OWL-DL, commonly used for ontologies and knowledge bases.

A new semantic pseudo-metric [7] is exploited in order to overcome these limitations. Following the distance-induction method proposed in [9], the proposed measures are based on a committee of features (concepts) onto which individuals are projected for being compared. As such, these measures are not absolute, yet they depend on the knowledge base they are applied to. However, the measures are suitable for a wide range of languages, since they merely depend on the discernibility of the input individuals w.r.t. a fixed set of concepts. The choice of optimal committees may be performed in advance through randomized search algorithms [7].

Such measures have been integrated in the NN procedure presented in [4]. Essentially the classification of a resource w.r.t. a query concept is performed by selecting the closest resources in the knowledge base and then determining the membership through a weighted voting procedure based on the neighbor similarity.

The resulting system allowed for an experimentation of the method on performing instance retrieval with real ontologies drawn from public repositories comparing its predictions to the assertions that were logically derived by a standard reasoner. These experiments show that the novel measure considerably increases the effectiveness of the method with respect to past experiments where the same procedure was integrated with other dissimilarity measures [4].

The paper is organized as follows. The basics of the instance-based approach applied to the standard representations are recalled in Sect. 2. The next Sect. 3 presents the semantic similarity measures adopted in the retrieval procedure. Sect. 4 reports the outcomes of experiments performed with the implementation of the procedure. Possible developments are finally examined in Sect. 5.

2 Resource Retrieval as Nearest Neighbor Search

2.1 Representation and Inference

In the following sections, we assume that concept descriptions are defined in terms of a generic sublanguage based on OWL-DL that may be mapped to *Description Logics* with the standard model-theoretic semantics (see the handbook [1] for a thorough reference).

A *knowledge base* $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$ contains a *TBox* $\mathcal{T}$ and an *ABox* $\mathcal{A}$. $\mathcal{T}$ is a set of axioms that define concepts. $\mathcal{A}$ contains factual assertions concerning the resources, also known as individuals. Moreover, the *unique names assumption* may be made on the ABox individuals, that are represented by their URIs. The set of the individuals occurring in $\mathcal{A}$ will be denoted with $\mathsf{Ind}(\mathcal{A})$.

As regards the inference services, like all other instance-based methods, our procedure may require performing *instance-checking* [1], which roughly amounts to determining whether an individual, say a, belongs to a concept extension, i.e. whether $C(a)$ holds for a certain concept C. Note that because of the OWA, a reasoner may be unable to give a positive or negative answer to a class-membership query. This service is provided proof-theoretically by a reasoner.

2.2 The Method

Query answering boils down to determining whether a resource belongs to a (query) concept extension. Here, an alternative inductive method is proposed for retrieving the resources that likely belong to a query concept. Such a method may also be able to provide an answer even when it may not be inferred by deduction, Moreover, it may also provide a measure of the likelihood of its answer.

In *similarity search* [10] the basic idea is to find the most similar object(s) to a query one (i.e. the one that is to be classified) with respect to a similarity (or dissimilarity) measure. We review the basics of the k-NN method applied to the Semantic Web context [4] context.

The objective is to induce an approximation for a discrete-valued target hypothesis function $h : \mathit{IS} \mapsto V$ from a space of instances *IS* to a set of values $V = \{v_1, \ldots, v_s\}$ standing for the classes (concepts) that have to be predicted. Note that normally $|\mathit{IS}| \ll |\mathsf{Ind}(\mathcal{A})|$ i.e. only a limited number of training instances is needed especially if they are prototypical for a region of the search space. Let x_q be the query instance whose class-membership is to be determined. Using a dissimilarity measure, the set of the k nearest (pre-classified) training instances w.r.t. x_q is selected: $NN(x_q) = \{x_i \mid i = 1, \ldots, k\}$.

In its simplest setting, the k-NN algorithm approximates h for classifying x_q on the grounds of the value that h is known to assume for the training instances in $NN(x_q)$, i.e. the k closest instances to x_q in terms of a dissimilarity measure. Precisely, the value is decided by means of a weighted majority voting procedure: it is simply the most *voted* value by the instances in $NN(x_q)$ weighted by the similarity of the neighbor individual.

The estimate of the hypothesis function for the query individual is:

$$\hat{h}(x_q) := \operatorname*{argmax}_{v \in V} \sum_{i=1}^{k} w_i \delta(v, h(x_i)) \qquad (1)$$

where δ returns 1 in case of matching arguments and 0 otherwise, and, given a dissimilarity measure d, the weights are determined by $w_i = 1/d(x_i, x_q)$.

Note that the estimate function $\hat{h}$ is defined extensionally: the basic k-NN method does not return an intensional classification model (a function or a concept definition), it merely gives an answer for the instances to be classified. It should be also observed that this setting assigns a value to the query instance which stands for one in a set of pairwise disjoint concepts (corresponding to the value set V). In a multi-relational setting this assumption cannot be made in general. An individual may be an instance of more than one concept.

The problem is also related to the CWA usually made in the knowledge discovery context. To deal with the OWA, the absence of information on whether a training instance x belongs to the extension of the query concept Q should not be interpreted negatively, as in the standard settings which adopt the CWA. Rather, it should count as neutral (uncertain) information. Thus, assuming the alternate viewpoint, the multi-class problem is transformed into a ternary one. Hence another value set has to be adopted, namely $V = \{+1, -1, 0\}$, where the three values denote, respectively, membership, non-membership, and uncertainty, respectively.

The task can be cast as follows: given a query concept Q, determine the membership of an instance x_q through the NN procedure (see Eq. 1) where $V = \{-1, 0, +1\}$ and the hypothesis function values for the training instances are determined as follows:

$$h_Q(x) = \begin{cases} +1 & \mathcal{K} \models Q(x) \\ -1 & \mathcal{K} \models \neg Q(x) \\ \ \ 0 & \textit{otherwise} \end{cases}$$

i.e. the value of h_Q for the training instances is determined by the entailment[1] the corresponding assertion from the knowledge base.

Note that, being based on a majority vote of the individuals in the neighborhood, this procedure is less error-prone in case of noise in the data (e.g. incorrect assertions) w.r.t. a purely logic deductive procedure, therefore it may be able to give a correct classification even in case of (partially) inconsistent knowledge bases.

It should be noted that the inductive inference made by the procedure shown above is not guaranteed to be deductively valid. Indeed, inductive inference naturally yields a certain degree of uncertainty. In order to measure the likelihood of the decision made by the procedure (individual x_q belongs to the query concept denoted by value v maximizing the argmax argument in Eq. 1), given the nearest training individuals in $NN(x_q, k) = \{x_1, \dots, x_k\}$, the quantity that determined the decision should be normalized by dividing it by the sum of such arguments over the (three) possible values:

$$l(\textit{class}(x_q) = v | NN(x_q, k)) = \frac{\sum_{i=1}^{k} w_i \cdot \delta(v, h_Q(x_i))}{\sum_{v' \in V} \sum_{i=1}^{k} w_i \cdot \delta(v', h_Q(x_i))} \tag{2}$$

Hence the likelihood of the assertion $Q(x_q)$ corresponds to the case when $v = +1$.

3 A Semantic Pseudo-Metric for Individuals

As mentioned in the first section, various attempts to define semantic similarity (or dissimilarity) measures for concept languages have been made, yet they have still a limited applicability to simple languages [3] or they are not completely semantic depending also on the structure of the descriptions [4]. Moreover, for our purposes, we need a function for measuring the similarity of individuals rather than concepts. It can be observed that

[1] We use $\models$ to denote entailment, as computed through a reasoner.

individuals do not have a syntactic structure that can be compared. This has led to lifting them to the concept description level before comparing them (recurring to the notion of the *most specific concept* of an individual w.r.t. the ABox [1], yet this makes the measure language-dependent. Besides, it would add a further approximations as the most specific concepts can be defined only for simple DLs.

For the NN procedure, we intend to exploit a new measure that totally depends on semantic aspects of the individuals in the knowledge base.

3.1 The Family of Measures

The new dissimilarity measures are based on the idea of comparing the semantics of the input individuals along a number of dimensions represented by a committee of concept descriptions. Indeed, on a semantic level, similar individuals should behave similarly with respect to the same concepts. Following the ideas borrowed from [9], totally semantic distance measures for individuals can be defined in the context of a knowledge base.

More formally, the rationale is to compare individuals on the grounds of their semantics w.r.t. a collection of concept descriptions, say $\mathsf{F} = \{F_1, F_2, \ldots, F_m\}$, which stands as a group of discriminating *features* expressed in the OWL-DL sublanguage taken into account.

In its simple formulation, a family of distance functions for individuals inspired to Minkowski's norms L_p can be defined as follows [7]:

Definition 3.1 (family of measures). *Let $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$ be a knowledge base. Given a set of concept descriptions $\mathsf{F} = \{F_1, F_2, \ldots, F_m\}$, a family of dissimilarity functions $d_p^{\mathsf{F}} : \mathsf{Ind}(\mathcal{A}) \times \mathsf{Ind}(\mathcal{A}) \mapsto [0, 1]$ is defined as follows:*

$$\forall a, b \in \mathsf{Ind}(\mathcal{A}) \quad d_p^{\mathsf{F}}(a, b) := \frac{1}{|\mathsf{F}|} \left[\sum_{i=1}^{|\mathsf{F}|} \mid \pi_i(a) - \pi_i(b) \mid^p \right]^{1/p}$$

where $p > 0$ and $\forall i \in \{1, \ldots, m\}$ the projection function π_i *is defined by:*

$$\forall a \in \mathsf{Ind}(\mathcal{A}) \quad \pi_i(a) = \begin{cases} 1 & F_i(a) \in \mathcal{A} \qquad (\mathcal{K} \models F_i(a)) \\ 0 & \neg F_i(a) \in \mathcal{A} \qquad (\mathcal{K} \models \neg F_i(a)) \\ 1/2 & otherwise \end{cases}$$

The superscript F will be omitted when the set of features is fixed.

The alternative definition for the projections, requires the entailment of an assertion (instance-checking) rather than the simple ABox look-up; this can make the measure more accurate yet more complex to compute unless a KBMS is employed maintaining such information at least for the concepts in F.

3.2 Discussion

It is easy to prove [7] that these functions have the standard properties for pseudo metrics (i.e. semi-distances [10]):

Proposition 3.1 (pseudo-metric). *For a given a feature set* F *and* $p > 0$, d_p *is a pseudo-metric.*

It cannot be proved that $d_p^{\mathsf{F}}(a,b) = 0$ iff $a = b$. This is the case of *indiscernible* individuals with respect to the given set of features F. To fulfill this property several methods have been proposed involving the consideration of equivalent classes of individuals or the adoption of a supplementary meta-feature F_0 determining the equality of the two individuals.

Compared to other proposed dissimilarity measures [3, 4], the presented functions do not depend on the constructors of a specific language, rather they require only (retrieval or) instance-checking for computing the projections through class-membership queries to the knowledge base.

The complexity of measuring he dissimilarity of two individuals depends on the complexity of such inferences (see [1], Ch. 3). Note also that the projections that determine the measure can be computed (or derived from statistics maintained on the knowledge base) before the actual distance application, thus determining a speed-up in the computation of the measure. This is very important for algorithms that massively use this distance, such as all instance-based methods.

The measures strongly depend on F. Here, we make the assumption that the feature-set F represents a sufficient number of (possibly redundant) features that are able to discriminate really different individuals. The choice of the concepts to be included – *feature selection* – is beyond the scope of this work (see [7] for a randomized optimization procedure aimed at finding optimal committees). Experimentally, we could obtain good results by using the very set of both primitive and defined concepts found in the knowledge base.

Of course these approximate measures become more and more precise as the knowledge base is populated with an increasing number of individuals.

4 Experimentation

4.1 Experimental Setting

In order to test the NN procedure integrated with the pseudo-metric proposed in the previous section, we have applied it to retrieval problems on random queries.

To this purpose, a number of OWL ontologies was selected, namely: FINITE STATE MACHINES (FSM), SURFACE-WATER-MODEL (SWM), part of SCIENCE and NEW TESTAMENT NAMES (NTN) from the Protégé library[2], the Semantic Web Service Discovery dataset[3] (SWSD) and the FINANCIAL ontology[4]. Tab. 1 summarizes the details of these knowledge bases.

For each ontology, 30 queries were randomly generated by composition of primitive or defined concepts. The performance was evaluated comparing the decisions made by the NN procedure to those returned by a standard reasoner[5] as a baseline.

[2] `http://protege.stanford.edu/plugins/owl/owl-library`

[3] `https://www.uni-koblenz.de/FB4/Institutes/IFI/AGStaab/Projects/xmedia/dl-tree.htm`

[4] `http://www.cs.put.poznan.pl/alawrynowicz/financial.owl`

[5] We employed PELLET v. 1.5. See `http://pellet.owldl.com`

Table 1. Data concerning the ontologies employed in the experiments.

knowledge base	DL language	#concepts	#object prop.	#data prop.	#individuals
FSM	$\mathcal{SOF}(D)$	20	10	7	37
SWM	$\mathcal{ALCOF}(D)$	19	9	1	115
SCIENCE	$\mathcal{ALCIF}(D)$	74	70	40	331
NTN	$\mathcal{SHIF}(D)$	47	27	8	676
SWSD	$\mathcal{ALCH}$	258	25	0	732
FINANCIAL	$\mathcal{ALCIF}$	60	17	0	1000

The parameter k was set to $\log |\mathsf{Ind}(\mathcal{A})|$ depending on the number of individuals in the ontology. Yet we found experimentally that much smaller values could be chosen, resulting in the same classification. We employed the simpler version of the distance (d_1^{F}) using all the concepts in the knowledge base for determining the set F.

4.2 Results

Standard IR measures. Initially the standard IR measures precision, recall, and F_1-measure were employed to evaluate the system performance. The outcomes are reported in Fig.2. For each knowledge base, we report the average values obtained over the 30 queries as well as their standard deviation and minimum-maximum ranges of values.

It is possible to note that precision and recall are generally quite good except in the experiment with the SWSD ontology where precision was significantly lower. Namely, SWSD turned out to be more difficult (also in terms of recall) for two reasons: a very limited number of individuals per concept was available and the number of concepts is larger than in other knowledge bases. For the other ontologies scores are quite high, as testified also by the F-measure values. The results in terms of recall are also more stable than those for recall as proved by the limited variance observed, whereas some queries turned out to be quite difficult w.r.t. the correctness of the answer.

The reasons for precision being less than recall are probably related to the OWA. Indeed, in a many cases it was observed that the NN procedure deemed some individuals as relevant for the query issued while the DL reasoner was not able to assess this relevance and this was computed as a mistake while it may likely turn out to be a correct inference when judged by a human agent.

Because of the problem issued by the OWA, in some cases it could not be (deductively) ascertained whether a resource was relevant or not for a given query. Thus explicitating both the rate of inductively classified individuals and the real nature of the mistakes would be needed. This leads to consider different indices.

Alternative measures. In previous works [4], we had employed the following indices for the evaluation:

- *match rate*: rate of individuals whose classification matched the reasoner decision;
- *omission error rate*: rate of individuals for which inductive method could not determine whether they were relevant to the query (or not) while they were actually relevant according to the reasoner;

Table 2. Experimental results in terms of standard IR measures: average ± standard deviation and [min.;max.] intervals.

	precision	recall	F-measure
FSM	89.22 ± 15.88 [28.60;100.00]	91.63 ± 12.41 [50.00;100.00]	90.26 ± 14.46 [36.39;100.00]
SWM	73.35 ± 11.66 [52.90;93.80]	89.56 ± 9.35 [73.30;97.50]	80.52 ± 10.55 [62.04;93.80]
SCIENCE	94.55 ± 6.03 [86.70;99.70]	97.12 ± 2.78 [93.50;99.70]	95.79 ± 4.45 [89.97;99.70]
NTN	78.73 ± 9.98 [34.60;95.60]	92.28 ± 4.58 [85.30;99.70]	84.63 ± 7.84 [49.23;97.61]
SWSD	55.30 ± 11.01 [31.90;74.10]	70.59 ± 10.37 [56.80;86.20]	61.51 ± 9.68 [41.03;79.69]
FINANCIAL	89.57 ± 19.48 [22.40;99.70]	97.80 ± 5.06 [84.70;100.00]	92.43 ± 15.47 [35.75;99.85]

- *commission error rate*: rate of individuals inductively found to be relevant to the query concept, while the reasoner assigned them to its negation (and vice-versa);
- *induction rate*: rate of individuals whose relevance (or irrelevance) relevant w.r.t. the query concept could be determined by the inductive method, while this classification could not be derived logically by the reasoner.

Tab. 3 reports the outcomes in terms of these new indices. Preliminarily, it is important to note that, in each experiment, the commission error was low or absent. This means that the search procedure is quite accurate: it did not make critical mistakes i.e. cases when an individual is deemed as an instance of a concept while it really is an instance of a disjoint one. Furthermore, the rate of omission errors was quite low, yet it is more frequent for the considered ontologies especially when few disjointness axioms were specified. A noteworthy difference was observed for the case of the FINANCIAL ontology for which we found the lowest match rate and the highest variability in the observed results over the various query concepts.

Comparing these outcomes to those reported in other works on the same task [4], where the highest average match rate observed was about 80%, we find a significant increase of the performance due to the accuracy of the new measure. Also the elapsed time (not reported here) was much less with the new measure: once the values of the projection functions are pre-computed, the efficiency of the classification, which depends on the similarity computation gains a lot of speed-up.

The usage of all concepts for the feature committee F made the measure quite accurate, which is the reason why the procedure resulted quite conservative as regards inducing new assertions. In many cases, it matched rather faithfully the reasoner decisions (the top k nearest neighbors had null distance w.r.t. the query instance). Namely, we found that a choice for lower values for k could have been made, for in many cases the decision on the correct classification was easy to make even on account of fewer (the closest) neighbor instances. This yielded also that the likelihood of the inference made (see Eq. 2) turned out quite high.

Table 3. Results with alternative indices: average ± standard deviation and [min.;max.] intervals.

	match r.	commission e.r.	omission e.r.	induction r.
FSM	94.51 ± 6.63	5.49 ± 6.63	0.00 ± 0.00	0.00 ± 0.00
	[73.00;100.00]	[0.00;27.00]	[0.00;0.00]	[0.00;0.00]
SWM	85.38 ± 5.69	0.00 ± 0.00	2.68 ± 0.92	11.95 ± 5.37
	[75.70;98.30]	[0.00;0.00]	[0.90;4.30]	[0.90;20.90]
SCIENCE	97.31 ± 1.97	0.00 ± 0.00	0.98 ± 0.61	1.71 ± 1.41
	[94.60;99.40]	[0.00;0.00]	[0.30;1.80]	[0.30;3.60]
NTN	88.06 ± 6.95	0.00 ± 0.00	2.12 ± 0.77	9.83 ± 7.12
	[74.60;95.40]	[0.00;0.00]	[0.30;3.40]	[4.30;24.30]
SWSD	85.40 ± 4.96	0.00 ± 0.00	4.76 ± 1.86	9.84 ± 3.97
	[74.50;92.20]	[0.00;0.00]	[2.70;8.70]	[4.00;19.00]
FINANCIAL	93.34 ± 11.55	6.30 ± 11.55	0.01 ± 0.03	0.35 ± 0.06
	[54.80;99.70]	[0.00;44.70]	[0.00;0.10]	[0.30;0.50]

Cases of induction are particularly interesting because they suggest new assertions which cannot be logically derived by a deductive reasoner and they might be used to *complete* a knowledge base [2], e.g. after being validated by an ontology engineer. Eq. 2 should be employed to assess the likelihood of the candidate assertions and hence decide on their inclusion in the ABox.

5 Conclusions and Outlook

This paper explored the application of a distance-based procedure for semantic search to knowledge bases represented in OWL. To this purpose, a novel family of language-independent semantic pseudo-metrics was exploited. Specifically, these measures were integrated in a nearest neighbor search procedure which can be employed for solving approximate retrieval problems.

This turns out to be more effective w.r.t. purely logical methods, especially in the presence of incomplete (or noisy) information in the knowledge bases. Experiments made on various ontologies showed that the method is quite effective and also robust since it seldom made commission errors during the various runs. As expected for an instance-based learning method, the overall performance depends on the number (and distribution) of the available training instances.

As regards the dissimilarity measures, we argue that more efficiency may be reached when statistics (on class-membership) are maintained by the KBMS [8]. Besides, so far, the subsumption relationships among concepts in the feature committee are not explicitly exploited, which might likely make similarity measurements more accurate.

Further developments can be also foreseen as concerns the choice of good feature committees. Namely, since measures are very dependant on this choice, some immediate lines of investigations arise: studying how to maintain limited-sized concepts committees, yet saving those sets which altogether are endowed of a real discriminating power.

Randomized optimization procedures can be used to learn maximally discriminating sets of features, by allowing the composition of concepts through the specific constructors made available by the representation language of choice [7]. This can be accomplished especially well when large sets of individuals are available for the ontologies. Namely, part of the entire data can be drawn in order to learn optimal feature sets, in advance with respect to the next stage.

As mentioned, the measures can be adopted in other instance-based machine learning methods which can be applied to several further tasks. For instance, the measures have been exploited in a hierarchical conceptual clustering algorithm where clusters would be formed grouping individual resources on the grounds of their similarity [6].

References

[1] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P. Patel-Schneider, editors. *The Description Logic Handbook*. Cambridge University Press, 2003.

[2] F. Baader, B. Ganter, B. Sertkaya, and U. Sattler. Completing description logic knowledge bases using formal concept analysis. In M. Veloso, editor, *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, pages 230–235, Hyderabad, India, 2007.

[3] A. Borgida, T.J. Walsh, and H. Hirsh. Towards measuring similarity in description logics. In I. Horrocks, U. Sattler, and F. Wolter, editors, *Working Notes of the International Description Logics Workshop*, volume 147 of *CEUR Workshop Proceedings*, Edinburgh, UK, 2005.

[4] C. d'Amato, N. Fanizzi, and F. Esposito. Reasoning by analogy in description logics through instance-based learning. In G. Tummarello, P. Bouquet, and O. Signore, editors, *Proceedings of Semantic Web Applications and Perspectives, 3rd Italian Semantic Web Workshop, SWAP2006*, volume 201 of *CEUR Workshop Proceedings*, Pisa, Italy, 2006.

[5] W. Emde and D. Wettschereck. Relational instance-based learning. In L. Saitta, editor, *Proceedings of the 13th International Conference on Machine Learning, ICML96*, pages 122–130. Morgan Kaufmann, 1996.

[6] N. Fanizzi, C. d'Amato, and F. Esposito. A hierarchical clustering procedure for semantically annotated resources. In R. Basili and M.T. Pazienza, editors, *Proceedings of the 10th Congress of the Italian Association for Artificial Intelligence, AI*IA2007*, volume 4733 of *LNAI*, pages 266–277. Springer, 2007.

[7] N. Fanizzi, C. d'Amato, and F. Esposito. Induction of optimal semi-distances for individuals based on feature sets. In D. Calvanese, E. Franconi, V. Haarslev, D. Lembo, B. Motik, A.-Y. Turhan, and S. Tessaris, editors, *Working Notes of the 20th International Description Logics Workshop, DL2007*, volume 250 of *CEUR Workshop Proceedings*, Bressanone, Italy, 2007.

[8] I. R. Horrocks, L. Li, D. Turi, and S. K. Bechhofer. The Instance Store: DL reasoning with large numbers of individuals. In V. Haarslev and R. Möller, editors, *Proceedings of the 2004 Description Logic Workshop, DL 2004*, volume 104 of *CEUR Workshop Proceedings*, pages 31–40. CEUR, 2004.

[9] M. Sebag. Distance induction in first order logic. In S. Džeroski and N. Lavrač, editors, *Proceedings of the 7th International Workshop on Inductive Logic Programming, ILP97*, volume 1297 of *LNAI*, pages 264–272. Springer, 1997.

[10] P. Zezula, G. Amati, V. Dohnal, and M. Batko. *Similarity Search – The Metric Space Approach*. Advances in database Systems. Springer, 2007.