Approximate Measures of Semantic Dissimilarity under Uncertainty

Nicola Fanizzi, Claudia d'Amato, 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. We propose semantic distance measures based on the criterion of approximate discernibility and on evidence combination. In the presence of incomplete knowledge, the distance measures the degree of belief in the discernibility of two individuals by combining estimates of basic probability masses related to a set of discriminating features. We also suggest ways to extend this distance for comparing individuals to concepts and concepts to other concepts.

1 Introduction

In the context of reasoning in the Semantic Web, a growing interest is being committed to alternative inductive procedures extending the scope of the methods that can be applied to concept representations. Among them, many are based on a notion of similarity such as *case-based reasoning* [4], *retrieval* [3], *conceptual clustering* [7] or *ontology matching* [6]. However this notion is not easily captured in a definition, especially in the presence of uncertainty.

As pointed out in the seminal paper [2] concerning similarity in Description Logics (DL), most of the existing measures focus on the similarity of atomic concepts within simple hierarchies. Besides, alternative approaches are based on related notions of *feature* similarity or *information content*. All these approaches have been specifically aimed at assessing concept similarity.

In the perspective of crafting inductive methods for the aforementioned tasks, the need for a definition of a semantic similarity measure for *individuals* arises, that is a problem that so far received less attention in the literature. Some dissimilarity measures for individuals in specific DL representations have recently been proposed which turned out to be practically effective for the targeted inductive tasks [3], however they are still partly based on structural criteria which determine also their main weakness: they can hardly scale to complex languages.

We devised a new family of dissimilarity measures for semantically annotated resources, which can overcome the aforementioned limitations [8]. Our measures are mainly based on Minkowski's measures for Euclidean spaces [18] induced by means of a method developed in the context of relational *machine learning* [14].

We extend the idea a notion of *discernibility* borrowed from *rough sets* theory [13] which aims at the formal definition of vague sets (concepts) by means of their approximations. In this paper, we propose (semi-)distance measures based on semantic discernibility and on evidence combination [16, 5, 15].

Namely, the measures are based on the degree of discernibility of the input individuals with respect to a committee of features, which are represented by concept descriptions expressed in the concept language of choice. One of the advantages of these measures is that they do not rely on a particular language for semantic annotations. However, these new measures are not to be regarded as absolute, since they depend both on the choice (and cardinality) of the features committee and on the knowledge base they are applied to. These measures can easily be computed based on statistics on individuals that are likely to be maintained by knowledge base management systems designed for storing instances (e.g. [10]), which can determine a potential speed-up in the measure computation during knowledge-intensive tasks.

Furthermore, we also propose a way to extend the presented measures to the case of assessing concept similarity by means of the notion of *medoid* [11], i.e., in a categorical context, the most centrally located individual in a concept extension w.r.t. a given metric.

The remainder of the paper is organized as follows. In the next section, we recall the basics of approximate semantic distance measures for individuals in a DL knowledge base. Hence, we extend the measures with a more principled treatment of uncertainty based on evidence combination. Conclusions discuss the applicability of these measures in further works,

2 Semantic Distance Measures

Since our method is not intended for a particular representation, in the following we assume that resources, concepts and their relationships may be defined in terms of a generic representation that may be mapped to some DL language with the standard model-theoretic semantics (see the handbook [1] for a thorough reference).

In this context, 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 concept definitions. $\mathcal{A}$ contains assertions concerning the world state. The set of the individuals (resources) occurring in $\mathcal{A}$ will be denoted with $\mathsf{Ind}(\mathcal{A})$. Each individual can be assumed to be identified by its own URI. Sometimes, it could be useful to make the *unique names assumption* on such individuals.

As regards the inference services, our procedure requires performing *instance-checking* and the related service of *retrieval*, which will be used for the approximations.

2.1 A Simple Semantic Metric for Individuals

Aiming at distance-based tasks, such as clustering or similarity search, we have developed a new measure with a definition that totally depends on semantic

aspects of the individuals in the knowledge base [8], following ideas borrowed from metric learning in clausal spaces [14].

Indeed, for our purposes, we needed functions to measure the (dis)similarity of individuals. However individuals do not have a syntactic (or algebraic) structure that can be compared. Then the underlying idea is that. on a semantic level, similar individuals should behave similarly with respect to the same concepts. A way for assessing the similarity of individuals in a knowledge base can be based on the comparison of their semantics along a number of dimensions represented by a set of concept descriptions (henceforth referred to as the *committee*). Particularly, the rationale of the measure is to compare them on the grounds of their behavior w.r.t. a given set of concept descriptions, say $\mathsf{F} = \{F_1, F_2, \ldots, F_k\}$, which stands as a group of discriminating *features* expressed in the language taken into account.

We begin with defining the behavior of an individual w.r.t. a certain concept in terms of projecting it in this dimension:

Definition 2.1 (projection function). *Given a concept $F_i \in \mathsf{F}$, the related* projection function

$$\pi_i : \mathsf{Ind}(\mathcal{A}) \mapsto \{0, 1/2, 1\}$$

is defined:
$\forall a \in \mathsf{Ind}(\mathcal{A})$

$$\pi_i(a) := \begin{cases} 1 & \mathcal{K} \models F_i(a) \\ 0 & \mathcal{K} \models \neg F_i(a) \\ 1/2 & \textit{otherwise} \end{cases}$$

The case of $\pi_i(a) = 1/2$ corresponds to the case when a reasoner cannot give the truth value for a certain membership query. This is due to the *Open World Assumption* normally made in this context. Hence, as in the classic probabilistic models uncertainty is coped with by considering a uniform distribution over the possible cases.

Now the discernibility functions related to the committee concepts which compare the two input individuals w.r.t. these concepts through their projections:

Definition 2.2 (discernibility function). *Given a feature concept $F_i \in \mathsf{F}$, the related* discernibility function

$$\delta_i : \mathsf{Ind}(\mathcal{A}) \times \mathsf{Ind}(\mathcal{A}) \mapsto [0, 1]$$

is defined as follows:
$\forall (a, b) \in \mathsf{Ind}(\mathcal{A}) \times \mathsf{Ind}(\mathcal{A})$

$$\delta_i(a, b) = |\pi_i(a) - \pi_i(b)|$$

Finally, a whole family of distance functions for individuals inspired to Minkowski's distances L_p [18] can be defined as follows [8]:

Definition 2.3 (dissimilarity measures). *Let $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$ be a knowledge base. Given a set of concept descriptions $\mathsf{F} = \{F_1, F_2, \ldots, F_k\}$, a family of dissimilarity measures $\{d_p^{\mathsf{F}}\}_{p \in \mathbb{N}}$, contains functions*

$$d_p^{\mathsf{F}} : \mathsf{Ind}(\mathcal{A}) \times \mathsf{Ind}(\mathcal{A}) \mapsto [0, 1]$$

defined
$\forall (a, b) \in \mathsf{Ind}(\mathcal{A}) \times \mathsf{Ind}(\mathcal{A})$*:*

$$d_p^{\mathsf{F}}(a, b) := \frac{L_p(\pi_i(a), \pi_i(b))}{|\mathsf{F}|} = \frac{1}{k} \sqrt[p]{\sum_{i=1}^{k} \delta_i(a, b)^p}$$

Note that k depends on F and the effect of the factor $1/k$ is just to normalize the norms w.r.t. the number of features that are involved. It is worthwhile to recall that these measures are not absolute, then they should be also be considered w.r.t. the committee of choice, hence comparisons across different committees may not be meaningful. Larger committees are likely to decrease the measures because of the normalizing factor yet these values is affected also by the degree of redundancy of the features employed.

2.2 Example

Let us consider a knowledge base in a DL language made up of a TBox:
$T = \{$ $\mathsf{Female} \equiv \neg\mathsf{Male}$,
$\mathsf{Parent} \equiv \forall\mathsf{child}.\mathsf{Being} \sqcap \exists\mathsf{child}.\mathsf{Being}$,
$\mathsf{Father} \equiv \mathsf{Male} \sqcap \mathsf{Parent}$,
$\mathsf{FatherWithoutSons} \equiv \mathsf{Father} \sqcap \forall\mathsf{child}.\mathsf{Female}$
$\}$
and of an ABox:
$A = \{$ $\mathsf{Being}(\mathsf{ZEUS}), \mathsf{Being}(\mathsf{APOLLO}), \mathsf{Being}(\mathsf{HERCULES}), \mathsf{Being}(\mathsf{HERA})$,
$\mathsf{Male}(\mathsf{ZEUS}), \mathsf{Male}(\mathsf{APOLLO}), \mathsf{Male}(\mathsf{HERCULES})$,
$\mathsf{Parent}(\mathsf{ZEUS}), \mathsf{Parent}(\mathsf{APOLLO}), \neg\mathsf{Father}(\mathsf{HERA})$,
$\mathsf{God}(\mathsf{ZEUS}), \mathsf{God}(\mathsf{APOLLO}), \mathsf{God}(\mathsf{HERA}), \neg\mathsf{God}(\mathsf{HERCULES})$,
$\mathsf{hasChild}(\mathsf{ZEUS}, \mathsf{APOLLO}), \mathsf{hasChild}(\mathsf{HERA}, \mathsf{APOLLO})$,
$\mathsf{hasChild}(\mathsf{ZEUS}, \mathsf{HERCULES})$,
$\}$
Suppose $\mathsf{F} = \{F_1, F_2, F_3, F_4\} = \{\mathsf{Male}, \mathsf{God}, \mathsf{Parent}, \mathsf{FatherWithoutSons}\}$. Let us compute the distances (with $p = 1$):
$d_1^{\mathsf{F}}(\mathsf{ZEUS}, \mathsf{HERA}) = (|1-0| + |1-1| + |1-1| + |0-0|)\,/4 = 1/4$
$d_1^{\mathsf{F}}(\mathsf{HERA}, \mathsf{APOLLO}) = (|0-1| + |1-1| + |1-1| + |0-1/2|)\,/4 = 3/8$
$d_1^{\mathsf{F}}(\mathsf{APOLLO}, \mathsf{HERCULES}) = (|1-1| + |1-0| + |1-1/2| + |1/2-1/2|)\,/4 = 3/8$
$d_1^{\mathsf{F}}(\mathsf{HERCULES}, \mathsf{ZEUS}) = (|1-1| + |0-1| + |1/2-1| + |1/2-0|)\,/4 = 1/2$
$d_1^{\mathsf{F}}(\mathsf{HERA}, \mathsf{HERCULES}) = (|0-1| + |1-0| + |1-1/2| + |0-1/2|)\,/4 = 3/4$
$d_1^{\mathsf{F}}(\mathsf{APOLLO}, \mathsf{ZEUS}) = (|1-1| + |1-1| + |1-1| + |1/2-0|)\,/4 = 1/8$

2.3 Discussion

It is easy to prove that these dissimilarity functions have the standard properties for semi-distances [8]:

Proposition 2.1 (semi-distance). *For a fixed feature set* F *and* $p > 0$*, given any three instances* $a, b, c \in \mathsf{Ind}(\mathcal{A})$*. it holds that:*

1. $d_p^{\mathsf{F}}(a, b) \geq 0$ *and* $d_p^{\mathsf{F}}(a, b) = 0$ *if* $a = b$
2. $d_p^{\mathsf{F}}(a, b) = d_p^{\mathsf{F}}(b, a)$
3. $d_p^{\mathsf{F}}(a, c) \leq d_p^{\mathsf{F}}(a, b) + d_p^{\mathsf{F}}(b, c)$

This measure is not a distance since it does not hold that $a = b$ if $d_p^{\mathsf{F}}(a, b) = 0$. This is the case of *indiscernible* individuals with respect to the given committee of features F. However, if the *unique names assumption* were made then one may define a supplementary dimension for the committee (a sort of meta-feature F_0) based on equality, such that:
$\forall (a, b) \in \mathsf{Ind}(\mathcal{A}) \times \mathsf{Ind}(\mathcal{A})$

$$\delta_0(a, b) := \begin{cases} 0 & a = b \\ 1 & a \neq b \end{cases}$$

and then

$$d_p^{\mathsf{F}}(a, b) := \frac{1}{k+1} \sqrt[p]{\sum_{i=0}^{k} \delta_i(a, b)^p}$$

The resulting measures are distance measures.

Compared to other proposed dissimilarity measures [2, 3], 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.

So far we made the assumption that F may represent 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* – may be crucial. Therefore, we have devised specific optimization algorithms founded in *randomized search* which are able to find optimal choices of discriminating concept committees [8, 7].

The fitness function to be optimized is based on the *discernibility factor* [13] of the committee. Given the whole set of individuals $\mathsf{Ind}(\mathcal{A})$ (or just a hold-out sample to be used to induce an optimal measure) $HS \subseteq \mathsf{Ind}(\mathcal{A})$ the fitness

function to be maximized is:

$$\text{DISCERNIBILITY}(\mathsf{F}, HS) := \sum_{(a,b)\in HS^2} \sum_{i=1}^{k} \delta_i(a,b)$$

However, the results obtained so far with knowledge bases drawn from ontology libraries [7, 9] show that (a selection) of the (primitive and defined) concepts is often sufficient to induce satisfactory dissimilarity measures.

3 Dissimilarity Measures Based on Uncertainty

The measure defined in the previous section deals with uncertainty in a uniform way: in particular, the degree of discernibility of two individuals is null when they have the same behavior w.r.t. the same feature, even in the presence of total uncertainty of class-membership for both. When uncertainty regards only one projection, then they are considered partially (possibly) similar.

We would like to make this uncertainty more explicit[1]. One way to deal with uncertainty would be considering intervals rather than numbers in [0,1] as a measure of dissimilarity. This is similar to the case of imprecise probabilities [17].

In order to extend the measure, we propose an epistemic definition based on rules for combining evidence [5, 15]. The ultimate aim is to assess the distance between two individuals as a combination of the evidence that they differ based on some selected features (as in the previous section).

The distance measure that is to be defined is again based on the degree of belief of discernibility of individuals w.r.t. such features. To this purpose the probability masses of the basic events (class-membership) have to be assessed. However, in this case we will not treat uncertainty in the classic probabilistic way (uniform probability). Rather, we intend to take into account uncertainty in the computation.

The new dissimilarity measure will be derived as a combination of the degree of belief in the discernibility of the individuals w.r.t. each single feature. Before introducing the combination rule (that will have the measure as a specialization), the basic probability assignments have to be considered, especially for the cases when instance-checking is not able to provide a certain answer.

As in previous works [3], we may estimate the concept extensions recurring to their retrieval [1], i.e. the individuals of the ABox that can be proved to belong to a concept. Thus, in case of uncertainty, the basic probabilities masses for each feature concept, can be approximated[2] in the following way:

[1] We are referring to a notion of *epistemic* (rather than *aleatory*) probability [15], which seems more suitable for our purposes. See Shafer's introductory chapter in [16] on this distinction.

[2] In case of a certain answer received from the reasoner, the probability mass amounts to 0 or 1.

$\forall i \in \{1, \ldots, k\}$

$$m_i(\mathcal{K} \models F_i(a)) \approx |\mathsf{retrieval}(F_i, \mathcal{K})|/|\mathsf{Ind}(\mathcal{A})|$$
$$m_i(\mathcal{K} \models \neg F_i(a)) \approx |\mathsf{retrieval}(\neg F_i, \mathcal{K})|/|\mathsf{Ind}(\mathcal{A})|$$
$$m_i(\mathcal{K} \models F_i(a) \vee \mathcal{K} \models \neg F_i(a)) \approx 1 - m_i(\mathcal{K} \models F_i(a)) - m_i(\mathcal{K} \models \neg F_i(a))$$

where the $\mathsf{retrieval}(\cdot, \cdot)$ operator returns the individuals which can be proven to be members of the argument concept in the context of the current knowledge base [1]. The rationale is that the larger the (estimated) extension the more likely is for individuals to belong to the concept. These approximated probability masses become more precise as more information is acquired. Alternatively, these masses could come with the ontology as a supplementary for of prior knowledge.

As in the previous section, we define a discernibility function related to a fixed concept which measures the amount of evidence that two input individuals may be separated by that concept:

Definition 3.1 (discernibility function). *Given a feature concept $F_i \in \mathsf{F}$, the related* discernibility function

$$\delta_i : \mathsf{Ind}(\mathcal{A}) \times \mathsf{Ind}(\mathcal{A}) \mapsto [0, 1]$$

is defined as follows:
$\forall (a, b) \in \mathsf{Ind}(\mathcal{A}) \times \mathsf{Ind}(\mathcal{A})$

$$\delta_i(a, b) := \begin{cases} m_i(\mathcal{K} \models \neg F_i(b)) & \textit{if } \mathcal{K} \models F_i(a) \\ m_i(\mathcal{K} \models F_i(b)) & \textit{else if } \mathcal{K} \models \neg F_i(a) \\ \delta_i(b, a) & \textit{else if } \mathcal{K} \models F_i(b) \vee \mathcal{K} \models \neg F_i(b) \\ 2 \cdot m_i(\mathcal{K} \models F_i(a)) \cdot m_i(\mathcal{K} \models \neg F_i(b)) & \textit{otherwise} \end{cases}$$

The extreme values $\{0, 1\}$ are returned when the answers from the instance-checking service are certain for both individuals. If the first individual is an instance of the i-th feature (resp., its complement) then the discernibility depends on the belief of class-membership to the complement concept of the other individual. Otherwise, if there is uncertainty for the former individual but not for the latter, the function changes its perspective, swapping the roles of the two individuals. Finally, in case there were uncertainty for both individuals, the discernibility is computed as the chance that they may belong one to the feature concept and one to its complement,

The combined degree of belief in the case of discernible individuals, assessed using the *mixing combination rule* [12, 15], can give a measure of the semantic distance between them.

Definition 3.2 (weighted average measure). *Given an ABox $\mathcal{A}$, a dissimilarity measure for the individuals in $\mathcal{A}$*

$$d^{\mathsf{F}}_{avg} : \mathsf{Ind}(\mathcal{A}) \times \mathsf{Ind}(\mathcal{A}) \mapsto [0, 1]$$

is defined as follows:
$\forall (a,b) \in \mathsf{Ind}(\mathcal{A}) \times \mathsf{Ind}(\mathcal{A})$

$$d^{\mathsf{F}}_{avg}(a,b) := \sum_{i=1}^{k} w_i \delta_i(a,b)$$

The choices for the weights are various. The most straightforward one is, of course, considering uniform weights: $w_i = 1/k$. Another one is

$$w_i = \frac{u_i}{u}$$

where

$$u_i = \frac{1}{|\mathsf{Ind}(A) \setminus \mathsf{retrieval}(F_k, \mathcal{K})|} \text{ and } u = \sum_{i=1}^{k} u_i$$

It is easy to see that this can be considered as a generalization of the measure defined in the previous section (for $p = 1$).

3.1 Discussion

It can be proved that function has the standard properties for semi-distances:

Proposition 3.1 (semi-distance). *For a fixed choice of weights $\{w_i\}_{i=1}^{k}$, function d^{F}_{avg} is a semi-distance.*

The underlying idea for the measure is to combine the belief of the dissimilarity of the two input individuals provided by several sources, that are related to the feature concepts. In the original framework for evidence composition the various sources are supposed to be independent, which is generally unlikely to hold. Yet, from a practical viewpoint, overlooking this property for the sake of simplicity may still lead to effective methods, as the Naïve Bayes approach in Machine Learning demonstrates.

It could also be criticized that the subsumption hierarchy has not been explicitly involved. However, this may be actually yielded as a side-effect of the possible partial redundancy of the various concepts, which has an impact on their extensions and thus on the related projection function. A tradeoff is to be made between the number of features employed and the computational effort required for computing the related projection functions.

The discriminating power of each feature concept can be weighted in terms of information and entropy measures. Namely, the degree of information yielded by each of these features can be estimated as follows:

$$H_i(a,b) = -\sum_{A \subseteq \Theta} m_i(A) \log(m_i(A))$$

where 2^{Θ}, w.r.t. the *frame of discernment*[3] [16, 15] $\Theta = \{D, \overline{D}\}$. then, the sum

$$\sum_{(a,b)\in HS} H_i(a,b)$$

provides a measure of the utility of the discernibility function related to each feature which can be used in randomized optimization algorithms.

3.2 Extensions

Following the rationale of the average link criterion used in agglomerative clustering [11], the measures can be extended to the case of concepts, by recurring to the notion of medoids.

The *medoid* of a group of individuals is the individual that has the highest similarity w.r.t. the others. Formally. given a group $G = \{a_1, a_2, \ldots, a_n\}$, the medoid is defined:

$$\text{medoid}(G) = \operatorname*{argmin}_{a\in G} \sum_{j=1}^{n} d(a, a_j)$$

Now, given two concepts C_1, C_2, we can consider the two corresponding groups of individuals obtained by retrieval $R_i = \{a \in \mathsf{Ind}(\mathcal{A}) \mid \mathcal{K} \models C_i(a)\}$, and their resp. medoids $m_i = \text{medoid}(R_i)$ for $i = 1, 2$ w.r.t. a given measure d_p^{F} (for some $p > 0$ and committee F). Then the function for concepts can be defined as follows:

$$d_p^{\mathsf{F}}(C_1, C_2) := d_p^{\mathsf{F}}(m_1, m_2)$$

Similarly, if the distance of an individual a to a concept C has to be assessed, one could consider the nearest (resp. farthest) individual in the concept extension or its medoid. Let $m = \text{medoid}(\mathsf{retrieval}(C))$ w.r.t. a given measure d_p^{F}. Then the measure for this case can be defined as follows:

$$d_p^{\mathsf{F}}(a, C) := d_p^{\mathsf{F}}(a, m)$$

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

4 Concluding Remarks

We have proposed the definition of dissimilarity measures over spaces of individuals in a knowledge base. The measures are not language-dependent, differently from other previous proposals [3], yet they are parameterized on a committee of concepts. Optimal committees can be found via randomized search methods [8].

[3] Here D stands for the case of discernible individuals w.r.t. F_i, $\overline{D}$ for the opposite case, and some probability mass may be assigned also to the uncertain case represented by $\{D, \overline{D}\}$.

Besides, we have extended the measures to cope with cases of uncertainty by means of a simple evidence combination method.

One of the advantages of the measures is that their computation can be very efficient in cases when statistics (on class-membership) are maintained by the KBMS [10]. As previously mentioned, the subsumption relationships among concepts in the committee is not explicitly exploited in the measure for making the relative distances more accurate. The extension to the case of concept distance may also be improved. Hence, scalability should be guaranteed as far as a good committee has been found and does not change also because of the locality properties observed for instances in several domains (e.g. social or biological networks).

A refinement of the committee may become necessary only when a degradation of the discernibility factor is detected due to the availability of somewhat new individuals. This may involve further tasks such as *novelty* or *concept drift* detection.

4.1 Applications

The measures have been integrated in an instance-based learning system implementing a nearest-neighbor learning algorithm: an experimentation on performing semantic-based retrieval proved the effectiveness of the new measures, compared to the outcomes obtained adopting other measures [3]. It is worthwhile to mention that results where not particularly affected by feature selection: often using the very concepts defined in the knowledge base provides good committees which are able to discern among the different individuals [9].

We are also exploiting the implementation of these measures for performing conceptual clustering [11], where (a hierarchy of) clusters is created by grouping instances on the grounds of their similarity, possibly triggering the induction of new emerging concepts [7].

4.2 Extensions

The measure may have a wide range of application of distance-based methods to knowledge bases. For example, logic approaches to ontology matching [6] may be backed up by the usage of our measures, especially when concepts to be matched across different terminologies are known to share a common set of individuals. Ontology matching could be a phase in a larger process aimed at data integration. Moreover metrics could also support a process of *(semi-)automated classification* of new data also as a first step towards ontology evolution.

Another problem that could be tackled by means of dissimilarity measures could be the *ranking* of the answers provided by a *matchmaking* algorithm based on the similarity between the concept representing the query and the retrieved individuals.

Acknowledgments

The authors would like to thank the anonymous reviewers for their suggestions.

References

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

[2] 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.

[3] 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.

[4] M. d'Aquin, J. Lieber, and A. Napoli. Decentralized case-based reasoning for the Semantic Web. In Y. Gil, E. Motta, V. Benjamins, and M. A. Musen, editors, *Proceedings of the 4th International Semantic Web Conference, ISWC2005*, number 3279 in LNCS, pages 142–155. Springer, 2005.

[5] D. Dubois and H. Prade. On the combination of evidence in various mathematical frameworks. In J. Flamm and T. Luisi, editors, *Reliability Data Collection and Analysis*, pages 213–241. 1992.

[6] J. Euzenat and P. Shvaiko. *Ontology matching*. Springer, 2007.

[7] N. Fanizzi, C. d'Amato, and F. Esposito. Evolutionary conceptual clustering of semantically annotated resources. In *Proceedings of the 1st International Conference on Semantic Computing, IEEE-ICSC2007*, pages 783–790, Irvine, CA, 2007. IEEE Computer Society Press.

[8] N. Fanizzi, C. d'Amato, and F. Esposito. Induction of optimal semi-distances for individuals based on feature sets. In *Working Notes of the International Description Logics Workshop, DL2007*, volume 250 of *CEUR Workshop Proceedings*, Bressanone, Italy, 2007.

[9] N. Fanizzi, C. d'Amato, and F. Esposito. Instance-based query answering with semantic knowledge bases. 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 254–265, Rome, Italy, 2007. Springer.

[10] 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.

[11] A.K. Jain, M.N. Murty, and P.J. Flynn. Data clustering: A review. *ACM Computing Surveys*, 31(3):264–323, 1999.

[12] C. K. Murphy. Combining belief functions when evidence conflicts. *Decision Support Systems*, 29(1-9), 2000.

[13] Z. Pawlak. *Rough Sets: Theoretical Aspects of Reasoning About Data*. Kluwer Academic Publishers, 1991.

[14] 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.

[15] K. Sentz and S. Ferson. Combination of evidence in Dempster-Shafer theory. Technical Report SAND2002-0835, SANDIATech, April 2002.
[16] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
[17] P. Walley. *Statistical reasoning with imprecise probabilities*. Chapman and Hall, London, 1991.
[18] P. Zezula, G. Amato, V. Dohnal, and M. Batko. *Similarity Search – The Metric Space Approach*. Advances in Database Systems. Springer, 2007.