Representing Uncertain Concepts in Rough Description Logics via Contextual Indiscernibility Relations

Nicola Fanizzi[1], Claudia d'Amato[1],
Floriana Esposito[1], and Thomas Lukasiewicz[2,3]

[1] LACAM – Dipartimento di Informatica, Università degli Studi di Bari
Via Orabona 4, 70125 Bari, Italy
{fanizzi|claudia.damato|esposito}@di.uniba.it
[2] Computing Laboratory, University of Oxford
Wolfson Building, Parks Road, Oxford OX1 3QD, UK
thomas.lukasiewciz@comlab.ox.ac.uk

Abstract. We investigate on modeling uncertain concepts via *rough description logics*, which are an extension of traditional description logics by a simple mechanism to handle approximate concept definitions through lower and upper approximations of concepts based on a rough-set semantics. This allows to apply rough description logics for modeling uncertain knowledge. Since these approximations are ultimately grounded on an indiscernibility relationship, the paper explores possible logical and numerical ways for defining such relationships based on the considered knowledge. In particular, the notion of context is introduced, allowing for the definition of specific equivalence relationships, to be used for approximations as well as for determining similarity measures, which may be exploited for introducing a notion of tolerance in the indiscernibility.

1 Introduction

Modeling uncertain concepts in description logics (DLs) [1] is generally done via numerical approaches, such as probabilistic or possibilistic ones [2]. A drawback of these approaches is that uncertainty is introduced in the model, which often has the consequence that the approach becomes conceptually and/or computationally more complex. An alternative (simpler) approach is based on the theory of *rough sets* [3], which gave rise to new representations and *ad hoc* reasoning procedures [4]. These languages are based on the idea of *indiscernibility*.

Among these recent developments, *rough description logics* (RDLs) [5] have introduced a complementary mechanism that allows for modeling uncertain knowledge by means of crisp approximations of concepts. RDLs extend classical DLs with two modal-like operators, the lower and the upper approximation. In the spirit of rough set theory, two concepts approximate an underspecified (uncertain) concept as particular

[3] Alternative address: Institut für Informationssysteme, Technische Universität Wien, Favoritenstr. 9-11, 1040 Wien, Austria; email: lukasiewicz@kr.tuwien.ac.at.

sub- and super-concepts, describing which elements are definitely and possibly, respectively, elements of the concept.

The approximations are based on capturing uncertainty as an indiscernibility relation R among individuals, and then formally defining the upper approximation of a concept as the set of individuals that are indiscernible from at least one that is known to belong to the concept:

$$\overline{C} := \{a \mid \exists b : R(a,b) \wedge b \in C\} .$$

Similarly, one can define the lower approximation as

$$\underline{C} := \{a \mid \forall b : R(a,b) \rightarrow b \in C\} .$$

Intuitively, the upper approximation of a concept C covers the elements of a domain with the typical properties of C, whereas the lower approximation contains the prototypical elements of C.

This may be described in terms of necessity and possibility. These approximations are to be defined in a crisp way. RDLs can be simulated with standard DLs without added expressiveness. This means that reasoning can be performed by translation to standard DLs using a standard DL reasoner.

The pressing issue of efficiency of the reasoning has to be solved. So far, reasoners are not optimized for reasoning with equivalence classes, which makes reasoning sometimes inefficient. To integrate equivalence relations into RDL ABoxes, other ways may be investigated. Inspired by recent works on semantic metrics [6] and kernels, we propose to exploit semantic similarity measures, which can be optimized in order to maximize their capacity of discerning really different individuals. This naturally induces ways for defining an equivalence relation based on indiscernibility criteria.

The rest of this paper is organized as follows. The basics of RDLs are presented in the next section. Then, in Section 3, contextual indiscernibility relations are introduced. In Section 4, a family of similarity measures based on such contexts is proposed along with a suggestion on their optimization. This also allows for the definition of tolerance degrees of indiscernibility. Conclusions and further applications of ontology mining methods are finally outlined in Section 5.

2 Rough Description Logics

In the following, we assume some familiarity with the basics of standard DL languages and their inference services (see [1] for further details).

As mentioned above, the basic idea behind RDLs is rather straightforward: one can approximate an uncertain concept C by giving upper and lower bounds. The upper approximation of C, denoted $\overline{C}$, is the set of all individuals that possibly belong to C. Orthogonally, the lower approximation of C, denoted $\underline{C}$, is the set of all individuals that definitely belong to C. Traditionally, this is modeled using primitive definitions, i.e., subsumption axioms. In pure DL modeling, the relation between C and its approximations $\underline{C}$ and $\overline{C}$ is $\underline{C} \sqsubseteq C \sqsubseteq \overline{C}$.

RDLs are not restricted to particular DLs, and can be defined for an arbitrary DL language $\mathcal{DL}$. Its RDL language $\mathcal{RDL}$ has the lower and upper approximation as additional unary concept constructors, that is, if C is a concept in $\mathcal{RDL}$, then also $\overline{C}$ and $\underline{C}$ are concepts in $\mathcal{RDL}$. The notions of *rough TBox* and *ABox*, as well as *rough knowledge base* canonically extend the usual notions.

Example 2.1 (Advertising Campaign). Suppose that we want to use some pieces of data collected from the Web to find a group of people to serve as addressees for the advertising campaign of a new product. Clearly, the collected pieces of data are in general highly incomplete and uncertain. The DL concept *Addressee* may now be approximated from below by all the definite addressees and from above by all the potential addressees. So, we can use a DL language to specify the TBox knowledge about the concept *Addressee*, and in the same time specify the TBox and ABox knowledge about which people are definite and potential addressees, i.e., belong to the two concepts $\underline{\textit{Addressee}}$ and $\overline{\textit{Addressee}}$, respectively.

A *rough interpretation* is a triple $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}}, R^{\mathcal{I}})$, where $\Delta^{\mathcal{I}}$ is a domain of objects, $\cdot^{\mathcal{I}}$ is an interpretation function, and $R^{\mathcal{I}}$ is an equivalence relation over $\Delta^{\mathcal{I}}$. The function $\cdot^{\mathcal{I}}$ maps RDL concepts to subsets and role names to relations over the domain $\Delta^{\mathcal{I}}$. Formally, $\mathcal{I}$ extends to the new constructs as follows:

- $\overline{C}^{\mathcal{I}} = \{a^{\mathcal{I}} \in \Delta^{\mathcal{I}} \mid \exists b^{\mathcal{I}} \in \Delta^{\mathcal{I}} : R^{\mathcal{I}}(a^{\mathcal{I}}, b^{\mathcal{I}}) \wedge b^{\mathcal{I}} \in C^{\mathcal{I}}\}$,
- $\underline{C}^{\mathcal{I}} = \{a^{\mathcal{I}} \in \Delta^{\mathcal{I}} \mid \forall b^{\mathcal{I}} \in \Delta^{\mathcal{I}} : R^{\mathcal{I}}(a^{\mathcal{I}}, b^{\mathcal{I}}) \rightarrow b^{\mathcal{I}} \in C^{\mathcal{I}}\}$.

Example 2.2 (Advertising Campaign cont'd). In order to define the definite and potential addressees for the advertising campaign of a new product, we may exploit a classification of the people into equivalence classes. For example, people with an income above 1 million dollars may be definite addressees for the advertising campaign of a new Porsche, while people with an income above 100 000 dollars may be potential addressees, and people with an income below 10 000 dollars may not be addressees of such an advertising campaign.

One of the advantages of this way of modeling uncertain concepts is that reasoning comes for free. Indeed, reasoning with approximations can be reduced to standard DL reasoning, by translating rough concepts into pure DL concepts with a special reflexive, transitive, and symmetric role.

A translation function for concepts $\cdot^t : \mathcal{RDL} \mapsto \mathcal{DL}$ is defined as follows (introducing the new role symbol R for the indiscernibility relation):

- $A^t = A$, for all atomic concepts A in $\mathcal{RDL}$,
- $(\overline{C})^t = \exists R.C$, and $(\underline{C})^t = \forall R.C$, for all other concepts C in $\mathcal{RDL}$.

The translation function is recursively applied on subconcepts for all other constructs. This definition can be extended to subsumption axioms and TBoxes.

For any DL language $\mathcal{DL}$ with universal and existential quantification, and symmetric, transitive, and reflexive roles, there is no increase in expressiveness, i.e., RDLs can be simulated in (almost) standard DLs: an $\mathcal{RDL}$ concept C is satisfiable in a rough interpretation relative to $\mathcal{T}^t$ iff the $\mathcal{DL}$ concept C^t is satisfiable relative to $\mathcal{T}^t$ [5].

Other reasoning services, such as subsumption, can be reduced to satisfiability (and finally to ABox consistency) in the presence of negation. As the translation is linear, the complexity of reasoning in an RDL is the same as of reasoning in its DL counterpart with quantifiers, symmetry, and transitivity.

Since RDLs do not specify the nature of the indiscernibility relation, except prescribing its encoding as a (special) new equivalent relation, we introduce possible ways for defining it. The first one makes the definition depend on a specific set of concepts determining the indiscernibility of the individuals relative to a specific context described by the concepts in the knowledge base. Then, we also define the relations in terms of a similarity measure (based on a context of features) which allows for relaxing the discernibility using a tolerance threshold.

3 Contextual Indiscernibility Relations

In this section, we first define the notion of a context via a collection of DL concepts. We then introduce indiscernibility relations based on such contexts. We finally define upper and lower approximations of DL concepts using these notions, and we provide some theoretical results about them.

It is well known that classification by analogy cannot be really general-purpose, since the number of features on which the analogy is made may be very large [7]. The key point is that indiscernibility is not absolute but, rather, an induced notion which depends on the specific contexts of interest. Instead of modeling indiscernibility through a single relation in the interpretation, one may consider diverse contexts each giving rise to a different equivalence relation which determines also different ways of approximating uncertain concepts.

We first recall the notion of projection function [8]:

Definition 3.1 (projection). Let $\mathcal{I}$ be a DL interpretation, and let F be a DL concept. The projection function $\pi_F^{\mathcal{I}} : \Delta^{\mathcal{I}} \mapsto \{0, \frac{1}{2}, 1\}$ is defined as follows:

$$\forall a \in \Delta^{\mathcal{I}} : \quad \pi_F^{\mathcal{I}}(a) = \begin{cases} 1 & \mathcal{I} \models F(a); \\ 0 & \mathcal{I} \models \neg F(a); \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

We define a *context* as a finite set of relevant features in the form of DL concepts, which may encode a kind of context information for the similarity to be measured [9].

Definition 3.2 (context). A *context* is a set of DL concepts $\mathsf{C} = \{F_1, \ldots, F_m\}$.

Example 3.1 (Advertising Campaign cont'd). One possible context C for the advertising campaign of a new product is given as follows:

$$\mathsf{C} = \{\mathit{SalaryAboveMillion}, \mathit{HouseOwner}, \mathit{Manager}\},$$

where *SalaryAboveMillion*, *HouseOwner*, and *Manager* are DL concepts.

Two individuals, say a and b, are indiscernible relative to the context C iff $\forall i \in \{1, \ldots, m\} : \pi_i(a) = \pi_i(b)$. This easily induces an equivalence relation:

Definition 3.3 (indiscernibility relation). Let $\mathsf{C} = \{F_1, \ldots, F_m\}$ be a context. The indiscernibility relation R_{C} induced by C is defined as follows:

$$R_{\mathsf{C}} = \big\{(a, b) \in \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}) \mid \forall i \in \{1, \ldots, m\} : \; \pi_i^{\mathcal{I}}(a) = \pi_i^{\mathcal{I}}(b)\big\}$$

Hence, one may define multiple such relations by considering different contexts.

Any indiscernibility relation splits $\Delta^{\mathcal{I}}$ in a partition of equivalence classes (also known as *elementary sets*) denoted $[a]_{\mathsf{C}}$, for a generic individual a. Each class naturally induces a concept, denoted C_a.

Example 3.2 (Advertising Campaign cont'd). Consider again the context C of Example 3.1. Observe that C defines an indiscernibility relation on the set of all people, which is given by the extensions of all atomic concepts constructed from C as its equivalence classes. For example, one such atomic concept is the conjunction of *SalaryAboveMillion*, *HouseOwner*, and *Manager*; another one is the conjunction of *SalaryAboveMillion*, *HouseOwner*, and $\neg$ *Manager*.

Thus, a C-*definable* concept has an extension that corresponds to the union of elementary sets. The other concepts may be approximated as usual (we give a slightly different definition of the approximations relative to those in Section 2).

Definition 3.4 (contextual approximations). Let $\mathsf{C} = \{F_1, \ldots, F_m\}$ be a context, let D be a generic DL concept, and let $\mathcal{I}$ be an interpretation. Then, the *contextual upper and lower approximations* of D relative to C, denoted $\overline{D}^{\mathsf{C}}$ and $\underline{D}_{\mathsf{C}}$, respectively, are defined as follows:

- $(\overline{D}^{\mathsf{C}})^{\mathcal{I}} = \{a \in \Delta^{\mathcal{I}} \mid \mathsf{C}_a \sqcap D \not\models \bot\}$,
- $(\underline{D}_{\mathsf{C}})^{\mathcal{I}} = \{a \in \Delta^{\mathcal{I}} \mid \mathsf{C}_a \sqsubseteq D\}$.

Fig. 1 depicts these approximations. The partition is determined by the feature concepts included in the context, each block standing for one of the C-definable concepts. The block inscribed in the concept polygon represent its lower approximation, while the red-hatched ones stand for its upper approximation.

These approximations can be encoded in a DL knowledge base through special indiscernibility relationships, as in [5], so to exploit standard reasoners for implementing inference services (with crisp answers). Alternatively new constructors for contextual rough approximation may be defined to be added to the standard ones in the specific DL language.

It is easy to see that a series properties hold for these operators:

Proposition 3.1 (properties). *Given a context $\mathsf{C} = \{F_1, \ldots, F_m\}$ and two concepts D and E, it holds that:*

1. $\underline{\bot}_{\mathsf{C}} = \overline{\bot}^{\mathsf{C}} = \bot$,
2. $\underline{\top}_{\mathsf{C}} = \overline{\top}^{\mathsf{C}} = \top$,
3. $\underline{D \sqcup E}_{\mathsf{C}} \sqsupseteq \underline{D}_{\mathsf{C}} \sqcup \underline{E}_{\mathsf{C}}$,

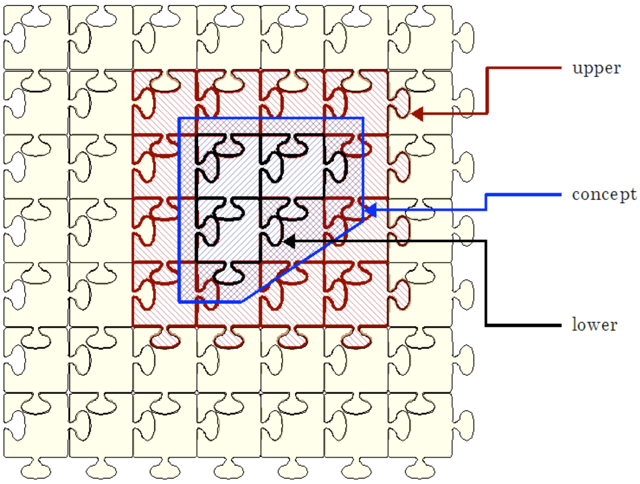

Fig. 1. Lower and upper approximations of rough concepts.

4. $\overline{D \sqcup E}^{\mathsf{C}} = \overline{D}^{\mathsf{C}} \sqcup \overline{E}^{\mathsf{C}}$,
5. $\underline{D \sqcap E}_{\mathsf{C}} = \underline{D}_{\mathsf{C}} \sqcap \underline{E}_{\mathsf{C}}$,
6. $\overline{D \sqcap E}^{\mathsf{C}} \sqsubseteq \overline{D}^{\mathsf{C}} \sqcap \overline{E}^{\mathsf{C}}$,
7. $\underline{\neg D}_{\mathsf{C}} = \neg \overline{D}^{\mathsf{C}}$,
8. $\overline{\neg D}^{\mathsf{C}} = \neg \underline{D}_{\mathsf{C}}$,
9. $\underline{\underline{D}_{\mathsf{C}}}_{\mathsf{C}} = \underline{D}_{\mathsf{C}}$,
10. $\overline{\overline{D}^{\mathsf{C}}}^{\mathsf{C}} = \overline{D}^{\mathsf{C}}$.

4 Numerical Extensions

We now first define rough membership functions. We then introduce contextual similarity measures, and we discuss the aspect of finding optimal contexts. We finally describe how indiscernibility relations can be defined on top of tolerance functions.

4.1 Rough Membership Functions

A rough concept description may include boundary individuals which cannot be ascribed to a concept with absolute certainty. As uncertainty is related to the membership to a set, one can define (rough) membership functions. This can be considered a numerical measure of the uncertainty:

Definition 4.1 (rough membership function). Let $\mathsf{C} = \{F_1, \ldots, F_m\}$ be a context. The C-rough membership function of an individual a to a concept D is defined by:

$$\mu_{\mathsf{C}}(a, D) = \frac{|(\mathsf{C}_a \sqcap D)^{\mathcal{I}}|}{|(\mathsf{C}_a)^{\mathcal{I}}|},$$

where $\mathcal{I}$ is the canonical interpretation [1].

Of course, this measure suffers from being related to the known individuals which conflicts with the open-world semantics of DL languages (unless an epistemic operator is adopted [10] or domain closure is assumed).

4.2 Contextual Similarity Measures

Since indiscernibility can be graded in terms of the similarity between individuals, we propose a new set of similarity functions, based on ideas that inspired a family of inductive distance measures [8, 6]:

Definition 4.2 (family of similarity functions). Let $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$ be a knowledge base. Given a context $\mathsf{C} = \{F_1, F_2, \ldots, F_m\}$, a family of similarity functions

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

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

$$s_p^{\mathsf{C}}(a, b) := \sqrt[p]{\sum_{i=1}^{m} \left| \frac{\sigma_i(a, b)}{m} \right|^p},$$

where $p > 0$ and the *basic similarity function* σ_i ($\forall i \in \{1, \ldots, m\}$) is defined by:

$$\forall a, b \in \mathsf{Ind}(\mathcal{A}) : \qquad \sigma_i(a, b) = 1 - |\pi_i(a) - \pi_i(b)|.$$

This corresponds to defining these functions model-theoretically as follows:

$$\sigma_i(a, b) = \begin{cases} 1 & (\mathcal{K} \models F_i(a) \wedge \mathcal{K} \models F_i(b)) \vee (\mathcal{K} \models \neg F_i(a) \wedge \mathcal{K} \models \neg F_i(b)); \\ 0 & (\mathcal{K} \models \neg F_i(a) \wedge \mathcal{K} \models F_i(b)) \vee (\mathcal{K} \models F_i(a) \wedge \mathcal{K} \models \neg F_i(b)); \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

Alternatively, in case of densely populated knowledge bases, this can be efficiently approximated, defining the functions as follows ($\forall a, b \in \mathsf{Ind}(\mathcal{A})$):

$$\sigma_i(a, b) = \begin{cases} 1 & (F_i(a) \in \mathcal{A} \wedge F_i(b) \in \mathcal{A}) \vee (\neg F_i(a) \in \mathcal{A} \wedge \neg F_i(b) \in \mathcal{A}); \\ 0 & (F_i(a) \in \mathcal{A} \wedge \neg F_i(b) \in \mathcal{A}) \vee (\neg F_i(a) \in \mathcal{A} \wedge F_i(b) \in \mathcal{A}); \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

The rationale for these functions is that similarity between individuals is determined relative to a given context [9]. Two individuals are maximally similar relative to a given concept F_i if they exhibit the same behavior, i.e., both are instances of the concept or of its negation. Conversely, the minimal similarity holds when they belong to opposite concepts. By the open-world semantics, sometimes a reasoner cannot assess the concept-membership, hence, since both possibilities are open, an intermediate value is assigned to reflect such uncertainty.

As mentioned, instance-checking is to be employed for assessing the value of the simple similarity functions. As this is known to be computationally expensive (also depending on the specific DL language), alternatively a simple look-up may be sufficient, as suggested by the first definition of the σ_i functions, especially for ontologies that are rich of explicit class-membership information (assertions).

The parameter p was borrowed from the form of the Minkowski's measures [11]. Once the context is fixed, the possible values for the similarity function are determined, hence p has an impact on the granularity of the measure.

Furthermore, the uniform choice of the weights assigned to the similarity related to the various features in the sum $(1/m^p)$ may be replaced by assigning different weights reflecting the importance of a certain feature in discerning the various instances. A good choice may be based on the amount of *entropy* related to each feature concept (then the weight vector has only to be normalized) [6].

4.3 Optimization of the Contexts

It is worthwhile to note that this is indeed a family of functions parameterized on the choice of features. Preliminary experiments regarding instance-based classification, demonstrated the effectiveness of the similarity function using the very set of both primitive and defined concepts found in the knowledge bases. But the choice of the concepts to be included in the context C is crucial and may be the object of a preliminary learning problem to be solved (*feature selection*).

As performed for inducing the pseudo-metric that inspired the definition of the similarity function [8], a preliminary phase may concern finding optimal contexts. This may be carried out by means of randomized optimization procedures.

Since the underlying idea in the definition of the functions is that similar individuals should exhibit the same behavior relative to the concepts in C, one may assume that the context C represents a sufficient number of (possibly redundant) features that are able to discriminate different individuals (in terms of a discernibility measure).

Namely, since the function is strictly dependent on the context C, two immediate heuristics arise:

- the *number* of concepts of the context,
- their discriminating power in terms of a *discernibility factor*, i.e., a measure of the amount of difference between individuals.

Finding optimal sets of discriminating features, should also profit by their composition, employing the specific constructors made available by the DL representation language of choice.

These objectives can be accomplished by means of randomized optimization techniques, especially when knowledge bases with large sets of individuals are available [8]. Namely, part of the entire data can be drawn in order to learn optimal feature sets, in advance with respect to the successive usage for all other purposes.

4.4 Approximation by Tolerance

In [4], a less strict type of approximation is introduced, based on the notion of *tolerance*. Exploiting the similarity functions that have been defined, it is easy to extend this kind of (contextual) approximation to the case of RDLs.

Let a *tolerance function* on a set U be any function $\tau : U \times U \mapsto [0, 1]$ such that for all $a, b \in U, \tau(a, b) = 1$ and $\tau(a, b) = \tau(b, a)$. Considering a tolerance function τ on U and a *tolerance threshold* $\theta \in [0, 1]$, a *neighborhood function* $\nu : U \mapsto 2^U$ is defined as follows:

$$\nu_\theta(a) = \{b \in U \mid \tau(a, b) \geq \theta\}.$$

For each element $a \in U$, the set $\nu_\theta(a)$ is also called the neighborhood of a.

Now, let us consider the domain $\Delta^{\mathcal{I}}$ of an interpretation $\mathcal{I}$ as a universal set, a similarity function s_p^{C} on $\Delta^{\mathcal{I}}$ (for some context C) as a tolerance function, and a threshold $\theta \in [0, 1]$. It is easy to derive an equivalence relationship on $\Delta^{\mathcal{I}}$, where the classes consist of individuals within a certain degree of similarity, indicated by the threshold: $[a]_{\mathsf{C}} = \nu_\theta(a)$. The notions of upper and lower approximation relative to the induced equivalence relationship descend straightforwardly.

Not that these approximations depend on the threshold. Thus, we have a numerical way to control the degree of indiscernibility that is needed to model uncertain concepts. This applies both in the standard RDL setting and in the new contextual one presented in the previous section.

5 Summary and Outlook

Inspired by previous works on dissimilarity measures in DLs, we have defined a notion of context, which allows to extend the indiscernibility relationship adopted by rough DLs, thus allowing for various kinds of approximations of uncertain concepts within the same knowledge base. It also saves the advantage of encoding the relation in the same DL language thus allowing for reasoning with uncertain concepts through standard tools obtaining crisp answers to queries.

Alternatively, these approximations can be implemented as new modal-like language operators. Some properties of the approximations deriving from rough sets theory have also been investigated.

A novel family of semantic similarity functions for individuals has also been defined based on their behavior relative to a number of features (concepts). The functions are language-independent being based on instance-checking (or ABox look-up). This allows for defining further kinds of graded approximations based on the notion of tolerance relative to a certain threshold.

Since data can be classified into indiscernible clusters, unsupervised learning methods for grouping individuals on the grounds of their similarity may be used for the definition of the equivalence relation [12, 8, 13]. Besides, it may also possible to learn rough DL concepts from the explicit definitions of the instances of particular concepts [14, 15, 16].

Acknowledgments. This work has been partially supported by the German Research Foundation (DFG) under the Heisenberg Programme.

References

[1] Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P., eds.: The Description Logic Handbook. Cambridge University Press (2003)

[2] Lukasiewicz, T., Straccia, U.: Managing uncertainty and vagueness in description logics for the Semantic Web. Journal of Web Semantics (2008) in press.

[3] Pawlak, Z.: Rough Sets: Theoretical Aspects of Reasoning About Data. Kluwer Academic Publishers (1991)

[4] Doherty, P., Grabowski, M., Lukaszewicz, W., Szalas, A.: Towards a framework for approximate ontologies. Fundamenta Informaticae **57** (2003) 147–165

[5] Schlobach, S., Klein, M.C.A., Peelen, L.: Description logics with approximate definitions - precise modeling of vague concepts. In: Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI 2007). (2007) 557–562

[6] d'Amato, C., Fanizzi, N., Esposito, F.: Query answering and ontology population: An inductive approach. In: Proceedings of the 5th European Semantic Web Conference (ESWC 2008). Volume 5021 of LNCS, Springer (2008) 288–302

[7] Mitchell, T.: Machine Learning. McGraw-Hill (1997)

[8] Fanizzi, N., d'Amato, C., Esposito, F.: Randomized metric induction and evolutionary conceptual clustering for semantic knowledge bases. In: Proceedings of the 16th International Conference on Knowledge Management (CIKM 2007), ACM Press (2007) 51–60

[9] Goldstone, R., Medin, D., Halberstadt, J.: Similarity in context. Memory and Cognition **25** (1997) 237–255

[10] Donini, F., Lenzerini, M., Nardi, D., Nutt, W.: An epistemic operator for description logics. Artificial Intelligence **100** (1998) 225–274

[11] Zezula, P., Amato, G., Dohnal, V., Batko, M.: Similarity Search – The Metric Space Approach. Advances in Database Systems. Springer (2007)

[12] Hirano, S., Tsumoto, S.: An indiscernibility-based clustering method. In: 2005 IEEE International Conference on Granular Computing, IEEE Computer Society (2005) 468–473

[13] Fanizzi, N., d'Amato, C., Esposito, F.: Conceptual clustering for concept drift and novelty detection. In: Proceedings of the 5th European Semantic Web Conference (ESWC 2008). Volume 5021 of LNCS, Springer (2008) 318–332

[14] Iannone, L., Palmisano, I., Fanizzi, N.: An algorithm based on counterfactuals for concept learning in the Semantic Web. Applied Intelligence **26** (2007) 139–159

[15] Lehmann, J., Hitzler, P.: A refinement operator based learning algorithm for the $\mathcal{ALC}$ description logic. In: Proceedings of the 17th International Conference on Inductive Logic Programming (ILP 2007). Volume 4894 of LNCS, Springer (2008) 147–160

[16] Fanizzi, N., d'Amato, C., Esposito, F.: DL-FOIL: Concept learning in description logics. In: Proceedings of the 18th International Conference on Inductive Logic Programming (ILP 2008). Volume 5194 of LNCS, Springer (2008) 107–121