Paraconsistent Rough Description Logic

Henrique Viana⋆, João Alcântara⋆⋆ and Ana Teresa Martins⋆ ⋆ ⋆

Departamento de Computação, Universidade Federal do Ceará, P.O.Box 12166,
Fortaleza, CE, Brasil 60455-760

Abstract. In this paper, we introduce a paraconsistent extension of Rough Description Logics which allows the representation of incomplete and contradictory concepts, as well as the lower and upper approximations of these kinds of concepts. Furthermore, we use the notions of approximations, which can be applied successively, to make restrictions or relaxations in the concept queries with the objective of decreasing or increasing the number of results of the queries, respectively.

1 Introduction

In many applications of Artificial Intelligence not rarely query tasks result in empty answers. Similarly, it may happen that a query has too many answers as a result, and it is not always that all these answers are important. Some approaches, known as query refinement used to deal with uncertain knowledge, develop methods to settle these problems. One of these approaches is the Rough Set theory introduced by Z. Pawlak [13]. Rough Set theory may be applied for query refinement by resorting to query restriction or query relaxation. A query can be restricted in order to obtain only the necessary results, or it can be relaxed, aiming at increasing the number of its results. In this paper, we will focus on the definition of a rough set framework in Description Logics (DLs) fine-tuned to deal with query refinement and incomplete and contradictory information.

Rough Description Logics (RDLs) [5,9,10,14] have introduced a mechanism that allows modeling uncertain reasoning by means of approximations of concepts. RDLs extend classical DLs with two operations, the lower and the upper approximation. Both approximations are based on capturing uncertainty as an indiscernibility relation R, an equivalence relation (i.e., reflexive, symmetric and transitive). We can formally define 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:

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

Similarly, one can define the lower approximation of a concept as the set of individuals which for all ones that are indiscernible from, it is known that these ones belong to the concept:

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

Extending the ideas of rough set theory, in order to represent incomplete and contradictory information, the work presented in [16] introduced the notion of paraconsistent rough sets. Instead of employing approximations to classical sets, it was taken into consideration four-valued sets, which are intended to represent incomplete and contradictory information. Furthermore, a similarity relation (i.e., a sort of "indiscernibility" relation that is at least reflexive) is used instead of an equivalence relation, with the aim of modeling different levels of uncertainty.

In this work, we will adapt the notions of paraconsistent rough sets to DL, introducing a paraconsistent RDL, called $PR_{\mathcal{ALC}}$. This paraconsistent RDL is slightly different from the well-known paraconsistent DL presented in [11]. Furthermore, we introduce two similarity relations to deal with approximation of concepts and apply the notion of contextual approximation [5] in our approach as an alternative form of approximation of concepts. Finally, we present some reasoning tasks, related to query refinement, which can be applied with these introduced operations.

The paper is structured as follows. In Section 2, we present the notions of the four-valued calculus, sets and approximations defined in [16]. In Section 3, we introduce $PR_{\mathcal{ALC}}$, and we apply the contextual approximation to $PR_{\mathcal{ALC}}$ together with the similarity relations. In Section 4, we present some query tasks that can be used in $PR_{\mathcal{ALC}}$; in particular, the tight and loose approximations. Finally, in Section 5 we conclude the paper.

2 Paraconsistent Rough Sets

In order to represent incomplete and contradictory information, it was introduced a rough set framework taking into account four-valued sets, instead of elementary sets [16]. In these four-valued sets, an element may belong to a given set, or it may not belong to the set, or its membership in the set may be unknown due to incomplete information, or even inconsistent due to a contradictory evidence. Under this view, membership functions, set containment and set operations are also four-valued, where logical values are **t** (*true*), **f** (*false*), **i** (*inconsistent*) and **u** (*unknown*). Moreover, since the knowledge at hand is incomplete, instead of indiscernibility relations, the authors resort to similarity relations to group together elements that are close to each other. Consequently, the notions of upper and lower approximations extend the usual definitions of approximations in the rough set theory.

2.1 Four-valued Semantics

In [16], the language for the four truth values was adapted from Belnap's Logic [1], which is grounded on bilattices [6]. As it is known, in bilattices, two orderings of truth values are considered: *truth ordering* ($\leq_t$) and *knowledge ordering* ($\leq_k$). However, in [16], the construction of the language is slightly different from that employed by Belnap. The change is motivated by some results accounted in [12] as counterintuitive in Belnap's *truth ordering*. In order to give more details, let us regard the following example involving test of cars:

Example 1 Suppose that we have two cars a and b, belonging to the same man, where $Safe(a) = \mathbf{u}$ and $Safe(b) = \mathbf{i}$. We want to know if all of his cars are safe, i.e., $Safe(a) \wedge Safe(b)$, where $\wedge$ stands for the meet operation with respect to $\leq_t$ in Belnap's Logic. We have that $(\mathbf{u} \wedge \mathbf{i}) = \mathbf{f}$. However, in this case we do not know whether both cars are safe because we do not have any information about the safety of car a. In contrast to the answer obtained in Belnap's logic, $\mathbf{u}$ seems to be a more intuitive answer to this case. Similarly, if we want to check if the man has a safe car, i.e., $Safe(a) \vee Safe(b)$, Belnap's Logic results in the value $\mathbf{t}$. This differs from our intuition because we know that the safety of car b is unclear, since the results of the tests are contradictory, and we know nothing about the safety of car a. In such a case, $\mathbf{i}$ seems to be a more intuitive answer.

In [16], in order to overcome these unexpected results, they redefined $\leq_t$ as the smallest reflexive and transitive relation satisfying $\mathbf{f} \leq_t \mathbf{u} \leq_t \mathbf{i} \leq_t \mathbf{t}$. Consequently, the operations of meet ($\wedge$) and join ($\vee$) in the *truth ordering* are defined as the greatest lower bound and as the least upper bound in this new ordering, respectively, i.e., $(x \wedge y) = \mathrm{GLB}^t\{x, y\}$ and $(x \vee y) = \mathrm{LUB}^t\{x, y\}$. The truth table for the meet and join operations in $\leq_t$ as well as to the negation and implication can be seen in Table 1.

$\wedge$	**f**	**u**	**i**	**t**
f	**f**	**f**	**f**	**f**
u	**f**	**u**	**u**	**u**
i	**f**	**u**	**i**	**i**
t	**f**	**u**	**i**	**t**

$\vee$	**f**	**u**	**i**	**t**
f	**f**	**u**	**i**	**t**
u	**u**	**u**	**i**	**t**
i	**i**	**i**	**i**	**t**
t	**t**	**t**	**t**	**t**

$\hookrightarrow$	**f**	**u**	**i**	**t**
f	**t**	**t**	**t**	**t**
u	**u**	**u**	**i**	**t**
i	**i**	**i**	**i**	**t**
t	**f**	**u**	**i**	**t**

$\neg$	
f	**t**
u	**u**
i	**i**
t	**f**

Table 1. Truth tables for $\wedge, \vee, \hookrightarrow$ and $\neg$

The implication $\hookrightarrow$ naturally extends the usual two-valued implication, which can be defined as $(\neg x \vee y)$. Consequently, the implication has the following property: $(x \hookrightarrow y) \equiv (\neg y \hookrightarrow \neg x)$, but it does not satisfy the Modus Ponens if we assume that $\{\mathbf{t}, \mathbf{i}\}$ is the set of designated values. For instance, $(\mathbf{i} \hookrightarrow \mathbf{f}) \wedge \mathbf{i}$ does not result in $\mathbf{f}$. A more detailed explanation of the definition of $\hookrightarrow$ can be found in [16]. Finally, the semantics of $\forall$ and $\exists$ is given by

$$\forall x P(x) = \mathrm{GLB}^t_{x \in U}\{P(x)\} \text{ and } \exists x P(x) = \mathrm{LUB}^t_{x \in U}\{P(x)\},$$

where U is the universe set and $P(x)$ denotes that x has the property P, which is evaluated to one of the four truth values.

2.2 Four-valued Sets

Now, we present the notion of a four-valued set. Given the disjoint sets U and $\neg U$, where $\neg U = \{\neg x \mid x \in U\}$, a four-valued set A on U is defined as any subset of $U \cup \neg U$. Intuitively, $x \in A$ represents that there is an evidence that x is in A, and $(\neg x) \in A$ represents that there is an evidence that x is not in A. We assume that $\neg(\neg x)$ is equal to x. In this framework, set membership is four-valued and it extends the usual two-valued membership.

Set membership, denoted as $\bar{\in} : U \times 2^{U \cup \neg U} \to \{\mathbf{t}, \mathbf{f}, \mathbf{i}, \mathbf{u}\}$, is defined by

$$x \bar{\in} A = \begin{cases} \mathbf{t}, \text{ if } x \in A, (\neg x) \notin A, \\ \mathbf{f}, \text{ if } x \notin A, (\neg x) \in A, \\ \mathbf{i}, \text{ if } x \in A, (\neg x) \in A, \\ \mathbf{u}, \text{ if } x \notin A, (\neg x) \notin A. \end{cases}$$

The complement $\neg A$ of a four-valued set A is defined by $\neg A = \{\neg x \mid x \in A\}$, and the four-valued set inclusion is defined by $X \Subset Y = \forall x \in U(x \bar{\in} X \hookrightarrow x \bar{\in} Y)$. The four-valued intersection and union are defined as $x \bar{\in} (X \cap Y) = (x \bar{\in} X) \wedge (x \bar{\in} Y)$ and $x \bar{\in} (X \cup Y) = (x \bar{\in} X) \vee (x \bar{\in} Y)$, respectively.

A four-valued extension of rough sets is, then, defined in [16] by four-valued set approximations as follows. First, the equivalence relation is replaced by a similarity relation, which is also four-valued. A similarity relation extends the indiscernibility relation by gathering in the same class objects which are not necessarily indiscernible, but sufficiently close or similar. In other words, a similarity relation is constructed from the indiscernibility relation by relaxing the original conditions for indiscernibility. This relaxation can be performed in many ways, thus giving many possible definitions for similarity.

Definition 1 (Four-valued Similarity Relation) *[16] By a four-valued similarity relation σ we mean any four-valued binary relation on a universe set U satisfying at least the reflexivity condition, i.e., for any element x of the universe $(x, x) \bar{\in} \sigma = \mathbf{t}$. By the neighbourhood of an element $x \in U$ w.r.t. σ, we understand the four-valued set $\sigma(x)$ such that $y \bar{\in} \sigma(x) = (x, y) \bar{\in} \sigma$.*

The membership of an element x in the lower approximation of a four-valued set A is determined by verifying the set inclusion of its neighbourhood $\sigma(x)$ in A. The membership of an element x in the upper approximation is determined by computing the greatest membership value that an element of the universe may have in the intersection of $\sigma(x)$ and A.

Definition 2 (Lower/Upper Approximation) *[16] Let A be a four-valued set. Then, the lower and upper approximations of A w.r.t. the similarity relation σ, denoted by A_σ^+ and $A_\sigma^\oplus$, respectively, are defined by*

$$x \bar{\in} A_\sigma^+ = \sigma(x) \Subset A \text{ and } x \bar{\in} A_\sigma^\oplus = \exists y \in U[y \bar{\in} (\sigma(x) \cap A)].$$

3 Paraconsistent Rough Description Logic

Rough Description Logics (RDLs) [5,9,10,14] have introduced a complementary mechanism that allows modelling uncertain knowledge by means of approximations of concepts. RDLs extend classical DLs with two modal-like operations, the lower and the upper approximation. In the spirit of rough set theory, two concepts approximate an underspecified (uncertain) concept C as particular sub- and super-concepts, describing which elements are definitely and possibly, respectively, elements of C.

In this section, taking into consideration the approach presented in [16], we extend the RDLs formalisms to paraconsistent rough sets by introducing a four-valued DL general enough to encompass two kinds of similarity relations: those explicitly defined

and those defined in terms of a context. One distinctive aspect of our formalism is that it permits successive refinements of a query (see Section 4). In the sequel, we introduce the syntax, the semantics and reasoning tasks for this DL.

3.1 Paraconsistent Rough Description Logic $\mathcal{ALC}$

We introduce a paraconsistent rough extension of the description logic $\mathcal{ALC}$, called $PR_{\mathcal{ALC}}$.

Definition 3 (Alphabet) *The $PR_{\mathcal{ALC}}$ alphabet is composed of the symbols $\neg, \sqcap, \sqcup, \exists,$ $\forall, \top, \bot, \underline{\cdot}, \overline{\cdot}$ and three disjoint sets: the set of individuals N_I, the set of concept names N_C, the set of role names N_R and the set of similarity relations N_S.*

Definition 4 (Concepts) *Concepts in $PR_{\mathcal{ALC}}$ are defined by the syntax rules below, where C and D are concepts, A is an atomic concept, R is an atomic role and S is a similarity relation:*

$$C, D \longrightarrow A \mid \top \mid \bot \mid \neg C \mid C \sqcap D \mid C \sqcup D \mid \exists R.C \mid \forall R.C \mid \overline{C}^S \mid \underline{C}_S$$

Definition 5 (Semantics) *The semantics of $PR_{\mathcal{ALC}}$ individuals, atomic concepts, atomic roles and similarity relations is given by an interpretation $I = (\Delta^I, \cdot^I)$, where the domain Δ^I is a nonempty set of elements and $\cdot^I$ is a mapping function defined as follows: each individual $a \in N_I$ is mapped to $a^I \in \Delta^I$; each atomic concept name $A \in N_C$ is mapped to $A^I : \Delta^I \to \{\boldsymbol{t}, \boldsymbol{f}, \boldsymbol{i}, \boldsymbol{u}\}$; each atomic role name $R \in N_R$ is mapped to $R^I : \Delta^I \times \Delta^I \to \{\boldsymbol{t}, \boldsymbol{f}, \boldsymbol{i}, \boldsymbol{u}\}$; each similarity relation $S \in N_S$ is mapped to $S^I : \Delta^I \times \Delta^I \to \{\boldsymbol{t}, \boldsymbol{f}, \boldsymbol{i}, \boldsymbol{u}\}$ satisfying at least the reflexivity condition, i.e., $S^I(x, x) = \boldsymbol{t}$ for any $x \in \Delta^I$.*

Concepts can be interpreted inductively as follows where, for all $x \in \Delta^I$,

Syntax	Semantics
$\top$	$\top^I(x) = \boldsymbol{t}$
$\bot$	$\bot^I(x) = \boldsymbol{f}$
$\neg C$	$(\neg C)^I(x) = \neg(C^I(x))$
$C \sqcap D$	$(C \sqcap D)^I(x) = (C^I(x) \wedge D^I(x))$
$C \sqcup D$	$(C \sqcup D)^I(x) = (C^I(x) \vee D^I(x))$
$\exists R.C$	$(\exists R.C)^I(x) = \underset{y \in \Delta^I}{\mathbf{LUB}^t}(R^I(x, y) \wedge C^I(y))$
$\forall R.C$	$(\forall R.C)^I(x) = \underset{y \in \Delta^I}{\mathbf{GLB}^t}(R^I(x, y) \hookrightarrow C^I(y))$
$\overline{C}^S$	$(\overline{C}^S)^I(x) = \underset{y \in \Delta^I}{\mathbf{LUB}^t}(S^I(x, y) \wedge C^I(y))$
$\underline{C}_S$	$(\underline{C}_S)^I(x) = \underset{y \in \Delta^I}{\mathbf{GLB}^t}(S^I(x, y) \hookrightarrow C^I(y))$

The semantics of concepts is based on semantics of Fuzzy DLs [2], as well as the lower/upper approximations of a concept, which are related with fuzzy rough sets [4]. The main difference from our approach to others based on fuzzy sets is that we consider t-norms, t-conorms, implication functions, and negation functions as the operations of

conjunction ($\wedge$), disjunction ($\vee$), implication ($\hookrightarrow$) and negation ($\neg$), respectively, described for the four-valued calculus in [16] (see Table 1). Now, we define the notions of Terminological Box (TBox), Assertional Box (ABox) and ontology in $PR_{\mathcal{ALC}}$.

Definition 6 (TBox) *A TBox is a finite set of expressions of the form $C \sqsubseteq D$, the so-called general concept inclusion, where C and D are concepts in $PR_{\mathcal{ALC}}$.*

Definition 7 (ABox) *An ABox consists of a finite set of assertion axioms of the form $C(a)$ and $R(a,b)$, where C is a concept, R is a role, and a and b are individuals in $PR_{\mathcal{ALC}}$.*

When we want to refer to inclusion and assertion axioms indistinctly, we will just call them axioms. The semantics of inclusion and assertion axioms is given by the following table:

Syntax	Semantics
$C \sqsubseteq D$	$\mathbf{i} \leq_t \underset{x \in \Delta^I}{\text{GLB}^t}(C^I(x) \hookrightarrow D^I(x))$
$C(a)$	$\mathbf{i} \leq_t C^I(a^I)$
$R(a,b)$	$\mathbf{i} \leq_t R^I(a^I, b^I)$

Note that the semantics of concept inclusion $C \sqsubseteq D$ is derived from the four-valued calculus presented in [16], i.e., it is described w.r.t. implication $\hookrightarrow$, which is defined as $(a \hookrightarrow b) = \neg a \vee b$. We assume that $\{\mathbf{i}, \mathbf{t}\}$ is the set of designated values, therefore $C \sqsubseteq D$ holds iff the result of the implication is $\mathbf{i}$ or $\mathbf{t}$. For assertion axioms the idea is similar: $C(a)$ (resp. $R(a,b)$) holds iff $C(a)$ (resp. $R(a,b)$) is evaluated to one of the designated values.

Definition 8 (Ontology) *An ontology or knowledge base is a set composed by a TBox and an ABox.*

Definition 9 (Satisfiability) *The notion of satisfaction of a set of axioms ε by an interpretation $I = (\Delta^I, \cdot^I)$, denoted $I \models \varepsilon$, is defined as follows: $I \models \varepsilon$ iff I satisfies each element in ε. For an axiom α, if $I \models \alpha$ we say that I is a model of α. I satisfies an ontology O, denoted by $I \models O$, iff I is a model of each axiom of ontology O.*

Definition 10 (Logical Consequence) *An axiom α is a logical consequence of an ontology O, denoted by $O \models \alpha$, iff every model of O satisfies α.*

3.2 Contextual Approximation

In [5], it was introduced the notion of contextual indiscernibility relations in RDLs as a way of defining an equivalence relation based on the indiscernibility criteria. In particular, the notion of context is introduced, allowing the definition of specific equivalence relationships to be used for approximations. The main advantage of this approach is that the reasoning with equivalence classes becomes optimized, since the equivalence relations are discovered in the process of reasoning, differently from the traditional RDLs, where the equivalence relation must be explicitly defined.

In this subsection, we apply the notions of contextual approximation to $PR_{\mathcal{ALC}}$, extending the definitions of lower and upper approximations of a concept. We first present the notion of a context via a collection of $PR_{\mathcal{ALC}}$ concepts. We, then, introduce two specific similarity relations based on such contexts, and finally we define upper and lower approximations of concepts using these new similarities.

Definition 11 (Context) *A context is a finite nonempty set of DL concepts $\mathcal{C} = \{C_1, \ldots, C_n\}$.*

Two individuals a and b are indiscernible with respect to the context $\mathcal{C} = \{C_1, \ldots, C_n\}$ iff for all C_i, where $i \in \{1, \ldots, n\}, C_i^I(a) = C_i^I(b)$. This easily induces an equivalence relation:

Definition 12 (Indiscernibility Relation) *Let $\mathcal{C} = \{C_1, \ldots, C_n\}$ be a context. The indiscernibility relation $R_{\mathcal{C}}$ induced by $\mathcal{C}$ is defined as follows:*

$$R_{\mathcal{C}} = \{(a,b) \in \Delta^I \times \Delta^I \mid \text{ for all } C_i \text{ where } i \in \{1, \ldots, n\}, C_i^I(a) = C_i^I(b)\}.$$

Since we are dealing with incomplete information, a similarity relation should be more adequate to model relationships between individuals, because it allows to group together individuals that are close to each other, but not necessarily indiscernible. Now, we introduce a specific similarity relation, which is based on the work presented in [8]:

Definition 13 (Similarity Relation - unknown concepts) *Let $\mathcal{C} = \{C_1, \ldots, C_n\}$ be a context. The similarity relation $S_{\mathcal{C}}$ induced by $\mathcal{C}$ is defined as follows:*

$$S_{\mathcal{C}} = \{(a,b) \in \Delta^I \times \Delta^I \mid \text{ for all } C_i \text{ where } i \in \{1, \ldots, n\}, C_i^I(a) = C_i^I(b) \text{ or } C_i^I(a) = \boldsymbol{u} \text{ or } C_i^I(b) = \boldsymbol{u}\}.$$

The purpose of the similarity relation $S_{\mathcal{C}}$ is to approximate incomplete information by considering the truth value **u** as similar to **t**, **f** or **i** and vice versa. In $S_{\mathcal{C}}$ (the *do not care* relation), the relevant information is the only one which is asserted, i.e., to assure an individual has been evaluated to **u** in a concept is regarded as non-relevant. Therefore, an individual a is similar to b in a context $\mathcal{C}$ if for all concepts in $\mathcal{C}$, the interpretation of a and b are equal, or the interpretation of a or b is equal to **u**.

In order to approximate both contradictory and incomplete information, we introduce a second similarity relation, dubbed $P_{\mathcal{C}}$. In this new similarity relation, the truth values **t** and **f** are similar to **i**.

Definition 14 (Similarity Relation - unknown and inconsistent concepts) *Let $\mathcal{C} = \{C_1, \ldots, C_n\}$ be a context. The similarity relation $P_{\mathcal{C}}$ induced by $\mathcal{C}$ is defined as follows:*

$$P_{\mathcal{C}} = \{(a,b) \in \Delta^I \times \Delta^I \mid \text{ for all } C_i \text{ where } i \in \{1, \ldots, n\}, C_i^I(a) = C_i^I(b) \text{ or } C_i^I(a) = \boldsymbol{u} \text{ or } C_i^I(b) = \boldsymbol{u} \text{ or if } C_i^I(a) = \boldsymbol{t} \text{ then } C_i^I(b) = \boldsymbol{i} \text{ or if } C_i^I(a) = \boldsymbol{f} \text{ then } C_i^I(b) = \boldsymbol{i}\}.$$

In $P_\mathcal{C}$, whose definition was borrowed from [8], it is assumed that the information may be partially described because of our incomplete or contradictory knowledge. From this point of view, an individual a can be considered similar to b if the information obtained in a is also obtained in b. Hence, for a concept C, where $C^I(a) = \mathbf{t}$ and $C^I(b) = \mathbf{i}$, the individual a is similar to b since the truth value $\mathbf{t}$ is contained in $\mathbf{i}$. Note the converse is not true: b is not similar to a according to $P_\mathcal{C}$.

The contextual lower and upper approximation of a DL concept C w.r.t the context $\mathcal{C}$ is denoted by $\underline{C}_{\mathcal{R}_\mathcal{C}}$ and $\overline{C}^{\mathcal{R}_\mathcal{C}}$, respectively, where $\mathcal{R}$ is one of those similarity relations presented above. It is important to emphasize that, for approximations with indiscernibility and similarity relations, we have the following result:

Proposition 1 *[15] Given the concept C and the context $\mathcal{C}$, it holds that:*

$$\underline{C}_{P_\mathcal{C}} \sqsubseteq \underline{C}_{S_\mathcal{C}} \sqsubseteq \underline{C}_{R_\mathcal{C}} \sqsubseteq C \sqsubseteq \overline{C}^{R_\mathcal{C}} \sqsubseteq \overline{C}^{S_\mathcal{C}} \sqsubseteq \overline{C}^{P_\mathcal{C}}.$$

This holds because $P_\mathcal{C}$ is more general than $S_\mathcal{C}$, whilst, $S_\mathcal{C}$ is more general than $R_\mathcal{C}$.

4 Query Refinement

Rough set theory is an interesting candidate as a framework to be employed in query refinement. By definition, the upper approximation will add an individual to the result of the query as soon as it is related to one of the concepts already in the query, while the lower approximation will only retain an individual in the result if all related concepts are in the query. We can imagine a situation in which a query results in an empty answer; in this case, its upper approximation could be applied to possibly produce at least individuals related to the query. On the other hand, a query may result in many individuals, thus a lower approximation could be applied to possibly restrict the number of individuals in order to obtain those most related to the query.

However, the lower approximation may result in the empty query, being in this case too strict for query refinement. As for the upper approximation, it corresponds to a well known approach to query expansion. Nevertheless, it can be too flexible as a query expansion technique, resulting not only in an explosion of the query, but possibly even worse, in the addition of non-relevant individuals due to the ambiguous nature of some of the query concepts.

In this section, we combine the flexibility of the upper approximation with the strictness of the lower approximation by applying them alternately [3,4]. For instance, first we can expand the query by adding all the individuals that are known to be related to at least one of the query concepts. Next, we can reduce the expanded query by taking its lower approximation, thereby pruning away all previously added individuals suspected to be irrelevant for the query. The pruning strategy targets those individuals that are strongly related to the query concepts, but that do not belong to the expanded query. All these strategies of approximations are defined in $PR_{\mathcal{ALC}}$ in the sequel:

Definition 15 (Tight and Loose Upper/Lower Approximations) *Tight and loose upper/lower approximations are denoted by $\overline{\underline{C}}^S_{\ S}$, $\overline{\overline{C}}^{S^S}$, $\underline{C}_{S_S}$, and $\overline{\underline{C}}_S^S$, and are defined as*

Syntax	Semantics
$\underline{\overline{C}}^S_S$	$(\underline{\overline{C}}^S_S)^I(x) = \underset{y\in\Delta^I}{\mathrm{GLB}^t}(S^I(x,y) \hookrightarrow \underset{z\in\Delta^I}{\mathrm{LUB}^t}(S^I(y,z) \wedge C^I(z)))$
$\overline{\overline{C}^S}^S$	$(\overline{\overline{C}^S}^S)^I(x) = \underset{y\in\Delta^I}{\mathrm{LUB}^t}(S^I(x,y) \wedge \underset{z\in\Delta^I}{\mathrm{LUB}^t}(S^I(y,z) \wedge C^I(z)))$
$\underline{C_S}_S$	$(\underline{C_S}_S)^I(x) = \underset{y\in\Delta^I}{\mathrm{GLB}^t}(S^I(x,y) \hookrightarrow \underset{z\in\Delta^I}{\mathrm{GLB}^t}(S^I(y,z) \hookrightarrow C^I(z)))$
$\overline{\underline{C}_S}^S$	$(\overline{\underline{C}_S}^S)^I(x) = \underset{y\in\Delta^I}{\mathrm{LUB}^t}(S^I(x,y) \wedge \underset{z\in\Delta^I}{\mathrm{GLB}^t}(S^I(y,z) \hookrightarrow C^I(z)))$

Proposition 2 *[4] Given the concept C, a similarity relation S, and the indiscernibility relation $R_\mathcal{C}$, it holds that:*

- $\underline{C_S}_S \sqsubseteq \underline{C}_S \sqsubseteq C \sqsubseteq \overline{C}^S \sqsubseteq \overline{\overline{C}^S}^S$
- $\underline{C}_S \sqsubseteq \overline{\underline{C}_S}^S \sqsubseteq \overline{C}^S$ and $\underline{C}_S \sqsubseteq \overline{\underline{C}}^S_S \sqsubseteq \overline{C}^S$
- $\underline{C_{R_\mathcal{C}}}_{R_\mathcal{C}} \equiv \underline{C}_{R_\mathcal{C}} \sqsubseteq C \sqsubseteq \overline{C}^{R_\mathcal{C}} \equiv \overline{\overline{C}^{R_\mathcal{C}}}^{R_\mathcal{C}}$
- $\underline{C}_{R_\mathcal{C}} \equiv \overline{\underline{C}_{R_\mathcal{C}}}^{R_\mathcal{C}} \sqsubseteq \overline{C}^{R_\mathcal{C}}$ and $\underline{C}_{R_\mathcal{C}} \sqsubseteq \overline{\underline{C}}^{R_\mathcal{C}}_{R_\mathcal{C}} \equiv \overline{C}^{R_\mathcal{C}}$

Note that the application of tight and loose approximations w.r.t. an indiscernibility relation does not result in new answers to a query. Otherwise Proposition 2 suggests that if we resort to similarity relations, successive applications of approximations may result in different answers, making a similarity relation an interesting alternative to query refinement. For an application of $PR_{\mathcal{ALC}}$ to query refinement, let us consider an example considering similarity relations for incomplete and contradictory information.

Example 2 (Query Relaxation/Restriction) Let $x_1, x_2, x_3, x_4, x_5, x_6$ and x_7 be a set of individuals representing houses; *GoodLocation*, *Basement*, *Fireplace*, *Expensive*, *Cheap* and *Medium* be DL concepts; $\mathcal{C} = \{GoodLocation, Basement, Fireplace\}$ be a context and $I = (\Delta^I, \cdot^I)$ be an interpretation such that

$\Delta^I = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$, *GoodLocation*$^I = \{x_1, \neg x_2, x_3, \neg x_4, x_6, \neg x_7\}$,
Basement$^I = \{x_1, x_2, \neg x_2, \neg x_3, x_4, x_6, \neg x_6, x_7\}$,
Fireplace$^I = \{x_1, \neg x_2, \neg x_4, x_5, x_6, x_7\}$,
Medium$^I = \{\neg x_1, \neg x_2, x_3, x_4, x_5, \neg x_6, x_7\}$,
Expensive$^I = \{x_1, \neg x_2, \neg x_3, \neg x_4, \neg x_5, x_6, \neg x_7\}$,
Cheap$^I = \{\neg x_1, x_2, \neg x_3, \neg x_4, \neg x_5, \neg x_6, \neg x_7\}$.

Now, let us show our first example using query relaxation: suppose that we want to know what houses are expensive. Making a query with each house, we have that

$$I \models Expensive(x_1),\ I \not\models Expensive(x_2),\ I \not\models Expensive(x_3),\ I \not\models Expensive(x_4)\ \text{and}$$
$$I \not\models Expensive(x_5).$$

That is, x_1 is the only expensive house. But suppose that we want to know which houses are possibly expensive. Using query relaxation we will have that

$$I \models \overline{Expensive}^{S_\mathcal{C}}(x_1)\ \text{and}\ I \models \overline{Expensive}^{S_\mathcal{C}}(x_5).$$

Thus, x_1 and x_5 are possibly expensive. We can conclude that x_5 is similar to x_1, because x_1 is the only object which is expensive. If we use query relaxation again (loose upper approximation) we conclude that

$$I \models \overline{\overline{Expensive}^{S_c}}^{S_c}(x_1), I \models \overline{\overline{Expensive}^{S_c}}^{S_c}(x_3) \text{ and } I \models \overline{\overline{Expensive}^{S_c}}^{S_c}(x_5).$$

We have now that x_3 is loosely possibly an expensive object: because x_3 is similar to x_5, an object possibly expensive. Another example related to query refinement, but using query restriction: suppose that we want to know what houses are neither expensive nor cheap, i.e., medium value. Making a query with each house, we have that

$$I \not\models Medium(x_1), I \not\models Medium(x_2), I \models Medium(x_3), I \models Medium(x_4) \text{ and}$$
$$I \models Medium(x_5).$$

Using query restriction, we can conclude that

$$I \models \underline{Medium}_{S_c}(x_3), \text{ but } I \not\models \underline{Medium}_{S_c}(x_4) \text{ and } I \not\models \underline{Medium}_{S_c}(x_5).$$

That is, x_4 and x_5 do not have necessarily medium value. If we use query restriction again (tight lower approximation) we can conclude that

$$I \not\models \underline{\underline{Medium}_{S_c}}_{S_c}(x_3).$$

Thus x_3 surely does not necessarily have medium value (i.e., it is similar to an object that necessarily does not have medium value). Instead of using tight lower approximation, if we want to know what houses possibly have necessarily medium value, we may use loose lower approximation and conclude that

$$I \models \overline{\underline{Medium}_{S_c}}^{S_c}(x_3) \text{ and } I \models \overline{\underline{Medium}_{S_c}}^{S_c}(x_5).$$

With this example, we obtain that x_5 necessarily does not have medium value, but possibly necessarily it has medium value (because x_5 is similar to x_3, which has necessarily medium value).

Focusing on the similarity relation for inconsistent information, we have that $I \not\models Cheap(x_4)$ and $I \not\models \overline{Cheap}^{S_c}(x_4)$, but $I \models \overline{Cheap}^{P_c}(x_4)$. This means that $P_{\mathcal{C}}$ can be used to discover individuals related to contradictions within the context. Knowing $I \not\models \overline{Cheap}^{S_c}(x_4)$ and $I \models \overline{Cheap}^{P_c}(x_4)$, we infer that there is contradictory information in $\mathcal{C}$, and x_4 could be related to it. A similar intuition may be used for lower approximation to discover those individuals which certainly are not related to contradictions. For instance, as $I \models \underline{Medium}_{P_c}(x_7)$, the individual x_7 is certainly medium and not related to contradictions. On the other hand, as $I \models \underline{Medium}_{S_c}(x_3)$ and $I \not\models \underline{Medium}_{P_c}(x_3)$, although x_3 is certainly medium, it is related to contradictions. As we can see, in $PR_{\mathcal{ALC}}$, we can represent very elaborated query refinements.

5 Conclusion

In this paper, we introduced the paraconsistent rough description logic $PR_{\mathcal{ALC}}$, suitable to represent and approximate incomplete and contradictory concepts. Besides including in its language indiscernibility relations, our proposal permits to reason with more relaxed notions of similarity relations in order to characterise the upper/lower approximations of a concept/query. As consequence, many sophisticated kinds of query

refinements can be represented in $PR_{\mathcal{ALC}}$. Owing to the similarity relations, one of its distinctive aspects is that successive applications of approximations may result in successive refinements of a query.

As a future work, we will extend the $PR_{\mathcal{ALC}}$ to represent and approximate more sorts of incomplete knowledge, as in [8], where the incomplete information can be divided into several kinds of missing information. Consequently, new similarity relations can be introduced to model each kind of missing information. Another track to be explored is to introduce dominance relations [7], which are commonly employed to model preference between information and individuals.

References

1. N. D. Belnap. A useful four-valued logic. In J. Michael Dunn and G. Epstein, editors, *Modern Uses of Multiple-Valued Logic*, pages 8–37. D. Reidel, 1977.
2. F. Bobillo and U. Straccia. Supporting fuzzy rough sets in fuzzy description logics. In *ECSQARU*, pages 676–687, 2009.
3. M. De Cock and C. Cornelis. Fuzzy rough set based web query expansion. In *Proceedings of Rough Sets and Soft Computing in Intelligent Agent and Web Technology, International Workshop at WIIAT2005*, pages 9–16, 2005.
4. M. De Cock, C. Cornelis, and E. E. Kerre. Fuzzy rough sets: Beyond the obvious. In *Proceedings of FUZZ-IEEE2004*, volume 1, pages 103–108, 2004.
5. N. Fanizzi, C. d'Amato, F. Esposito, and T. Lukasiewicz. Representing uncertain concepts in rough description logics via contextual indiscernibility relations. In *URSW*, 2008.
6. M.L. Ginsberg. Multivalued logics: a uniform approach to reasoning in artificial intelligence. *Computational Intelligence*, 4:265–316, 1988.
7. S. Greco, B. Matarazzo, and R. Slowinski. Rough approximation of a preference relation by dominance relations. *European Journal of Operational Research*, 117(1):63–83, 1999.
8. J.W. Grzymala-Busse. Rough set strategies to data with missing attribute values. In *Foundations and Novel Approaches in Data Mining*, pages 197–212. 2006.
9. Y. Jiang, J. Wang, S. Tang, and B. Xiao. Reasoning with rough description logics: An approximate concepts approach. *Inf. Sci.*, 179(5):600–612, 2009.
10. C. Maria Keet. On the feasibility of description logic knowledge bases with rough concepts and vague instances. In *Description Logics*, 2010.
11. Y. Ma, P. Hitzler, and Z. Lin. Algorithms for paraconsistent reasoning with owl. In *The Semantic Web: Research and Applications. Proceedings of the 4th European Semantic Web Conference, ESWC2007*, pages 399–413. Springer, 2007.
12. J. Małuszyński, A. Szalas, and A. Vitória. A four-valued logic for rough set-like approximate reasoning. *T. Rough Sets*, 6:176–190, 2007.
13. Z. Pawlak. Rough sets. *International Journal of Information and Computer Science*, 11:341–356, 1982.
14. S. Schlobach, M. Klein, and Vrije Universteit Amsterdam. L.: Description logics with approximate definitions: Precise modeling of vague concepts. In *IJCAI 07*, 2007.
15. R. Słowiński and D. Vanderpooten. Similarity relation as a basis for rough approximations. *P.P. Wang, ed., Advances in Machine Intelligence and Soft-Computing. Vol.IV, Bookwrights, Raleigh, NC*, 1997.
16. A. Vitória, A. Szałas, and J. Małuszyński. Four-valued extension of rough sets. In *Proceedings of Third International Conference on Rough Sets and Knowledge Technology, Chengdu, China (17th–19th May 2008)*, volume 5009 of *LNCS*, pages 106–114, 2008.