

# Ontology Cleaning by Mereotopological Reasoning

José A. Alonso-Jiménez, Joaquín Borrego-Díaz, Antonia M. Chávez-González  
Departamento de Ciencias de la Computación e Inteligencia Artificial– Universidad of Sevilla  
Avda. Reina Mercedes s.n. 41012-Sevilla, Spain  
{jalonso, jborrego, tchavez}@us.es

## Abstract

*A mereotopological semantics to manage ontologies is presented. The aim is to provide a formal basis for ontology cleaning. It allows us to arrange, in a consistent manner, the concepts in early steps of the building of an ontology as well as to repair anomalies. The semantics supports cleaning cycle that combines several AI techniques as closed world assumption, default reasoning on taxonomies and knowledge acquisition.*

## 1. Introduction

It is generally accepted that for a satisfactory transition from the actual WWW to the Semantic Web it will be needed to deal with evolving ontologies. Ontologies provide common understanding in fields as Knowledge Management and electronic commerce, because they play a key role in the reasoning services for Knowledge Bases (KB) in the Semantic Web [8] [5].

The aim of this paper is to provide a formal semantics to support the repairing of provisional ontologies. It is based on a mereotopological representation of (incomplete) specifications of the concepts of the ontology. The method starts obtaining a spatial representation of the ontology, based on the mereotopological theory RCC [4], a sound theory for qualitative spatial reasoning, which is interpreted here as a theory on spatial representation of concepts. Two types of actions (topological or reticular arrangements) can be used on the 2D representation in order to clean the KB, repairing some anomalies. Since the actions work on a model of the interpretation, the resultant representation models the final KB. Thus, the method preserves consistency. Additional steps require the interaction with the beliefs of the user.

A motivation of the semantics is to satisfy the need of improvements of the current data cleaning systems [9]: a clear separation between the logical specification of data transformation and their physical transformation, an explanation

of the reasoning behind cleaning results and the possibility of interactive facilities to tune data cleaning programs.

Description Logic (DL) will be used to represent meta-data. In this way we focus here on issues not depending on the specific representation in the Semantic Web. Description Logic is a sound formalism to give a clear semantics to several tools for Knowledge Representation in the Semantic Web (see e.g. [12] [3]).

We aim to provide a formal semantics we can analyse with and repair some types of anomalies that arise from a *cleaning-cycle* applied to Knowledge Bases associated to complex ontologies [1]. In particular, anomalies that come up from lacks in the Knowledge Domain; that is, in the ontology (TBox, in terms of DL). With the tool presented here, the argument justifying an anomaly is analyzed, for improving the KB through its transformation. The study of an argument (an argument is a pair  $\langle \Pi, \varphi \rangle$  where  $\Pi$  is a subset of the KB and  $\Pi \vdash \varphi$ ) is easier than global analysis of a KB; this is due to its relative small size and it provides us with an interesting way to face up the management of inconsistent information in KB [7]. We will reason with spatial representations of arguments, in particular. It is clear that the technique is valid for full KB, although usefulness perhaps is lost because of the complexity of graphic representation (if the number of concepts is high).

The paper is structured as follows. Section 2 is devoted to show some particular features of the problem. The process where the spatial interpretation is used is explained in section 3. After presenting an interpretation of the qualitative reasoning formalism as a theory on relationships between concepts (section 4), each step of the process is shown in sections 5 to 8. Finally some remarks will be done.

### 1.1. Description Logics

Description Logics (<http://dl.kr.org>) deal with the representation of structured concepts. The concepts are described with a language with specific features, as conjunction, quantifiers on attributes of concepts, etc. Formally, DL is a subset of first order logic. Thus it inherits a formalized semantics, as opposite to some early formalisms.

\* Work partially supported by the project TIC-137 of the *Plan Andaluz de Investigación*.

Amenable calculus for several DL systems exist, whose computational complexity is known (see e.g. [6]).

In this logic, a KB,  $\Sigma$ , is a pair  $\langle \mathcal{T}, \mathcal{A} \rangle$ , where  $\mathcal{A}$  is a set of facts (the *extensional* component or Abox) and  $\mathcal{T}$  is a set of relations among concepts (the *intensional* component or Tbox). The following KB on the family ontology will be our running example

$$\mathcal{T} = \begin{cases} \text{Woman} \sqsubseteq \text{Person} \sqcap \text{Female} \\ \text{Man} \sqsubseteq \text{Person} \sqcap \neg \text{Woman} \\ \text{Father} \sqsubseteq \text{Man} \sqcap \exists \text{hasChild}.\text{Person} \\ \text{Father} \sqsubseteq \text{Parent} \end{cases}$$

$$\mathcal{A} = \begin{cases} \text{Father}(\text{John}) & \text{Man}(\text{John}) \\ \text{Female}(\text{RuPaul}) & \text{Man}(\text{RuPaul}) \\ \text{Woman}(\text{Ann}) & \exists \text{hasChild}.\text{Person}(\text{Ann}) \end{cases}$$

The set of elementary concepts of  $\Sigma$ ,  $\text{concepts}(\Sigma)$ , is the set of the concepts needed to build every formula of  $\Sigma$  via boolean combinations.

## 2. Anomalies in provisional ontologies

Knowledge Bases in DL may be affected by classical anomalies. It is known that there are anomalies which do not come from inconsistencies, and some of them can not be avoided in practice. We must bear in mind the possible dynamic nature of ontologies (they are neither complete nor definitive). In some respects – and in the first phases of their building – they must be considered as *provisional*.

Incompleteness of a KB must be understood in two ways: the logical incompleteness (with respect to a kind of questions), and the lack of concepts or roles (incompleteness with expressive nature). The matter of verification of KB in DL inherits the matter of logical inconsistency.

Another kind of anomalies due to expressive poverty exists. These anomalies arise -from the point of view of the user- because of lack of an exact profile of several concepts. When it happens, the user works on beliefs not even explicit in the KB. Such concepts, which exist in provisional ontologies, will be named *notions*. The existence of notions in an ontology implies that two concepts covered by the same notion can not be distinguished, namely, the ontology is coarse. It is necessary to locate and eventually to acquire this knowledge.

On the other hand, neglected development of the ontology (it may occur for example, when multiple programmers work on the same ontology project), leads to a validation problem which is different from classical problem of validation in Knowledge Based Systems: the ontology does not fit with the user beliefs about his/her framework, or it is both hard to use and to be understood by the others. Messy ontologies are definitely a risk in the management of large KB.

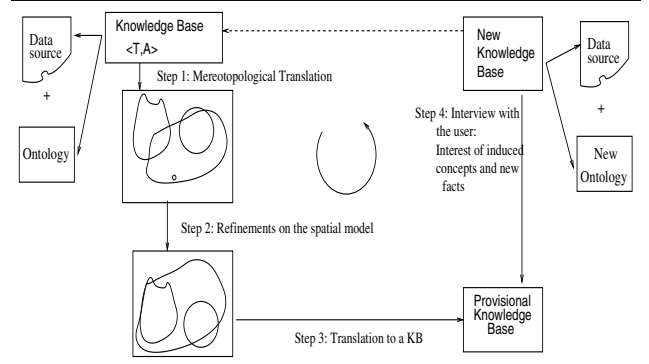


Figure 1. The ontology cleaning process

## 3. Sketch of the repairing process

The initial KB,  $\Sigma$ , is a pair TBox/Abox (usually, the hypothesis of an anomalous argument). Let us briefly sketch the four steps of the cleaning process (figure 1):

1. Firstly, a constraint satisfaction problem (CSP) on the spatial relational calculus RCC8 (or RCC5) is produced by a cognitively sound translation of the TBox to RCC formulas. The problem is solved obtaining a consistent scenario, represented in 2D. Facts of the Abox are added as points.
2. The user is requested to make *reticular* or *topological* arrangements on the graphical representation. Reticular arrangements represent refinements of relationships between concepts and topological arrangements imply substitution of a relationship by another one, disjoint with the former one but *cognitively near* of that. The user must lastly think that she/he has a fair RCC8(5) representation. This gives a table of spatial constraints on notions of the provisional ontology.
3. A translation from RCC formulas to DL formulas is applied. In the new KB, some relationships have changed and new concepts may have been induced.
4. Finally, the user interprets (or discards) the new objects that the translation may induce.

The result of the process is a new KB modified by the beliefs of the user. Indeed, the process may be a cycle, because it is possible that the ontology will change, new data will be induced, and they may lead to a new revision.

The foundational principle (argued by the practice) we assume is that an acceptable small set of concepts must have a clear spatial representation. In other case, this set is messy.

## 4. Mereotopological notions

The *Region Connection Calculus* (RCC) is a topological approach to qualitative spatial representation and reasoning on *spatial entities*, which are non-empty regular sets

$A_{DC} : DC(x, y) \leftrightarrow \neg C(x, y)$ (x is disconnect from y)
$A_P : P(x, y) \leftrightarrow \forall z[C(z, x) \rightarrow C(z, y)]$ (x is part of y)
$A_{PP} : PP(x, y) \leftrightarrow P(x, y) \wedge \neg P(y, x)$ (x is proper part of y)
$A_{EQ} : EQ(x, y) \leftrightarrow P(x, y) \wedge P(y, x)$ (x is identical with y)
$A_O : O(x, y) \leftrightarrow \exists z[P(z, x) \wedge P(z, y)]$ (x overlaps y)
$A_{DR} : DR(x, y) \leftrightarrow \neg O(x, y)$ (x is discrete from y)
$A_{PO} : PO(x, y) \leftrightarrow O(x, y) \wedge \neg P(x, y) \wedge \neg P(y, x)$ (x partially overlaps y)
$A_{EC} : EC(x, y) \leftrightarrow C(x, y) \wedge \neg O(x, y)$ (x is externally connected to y)
$A_{TPP} : TPP(x, y) \leftrightarrow PP(x, y) \wedge \exists z[EC(z, x) \wedge EC(z, y)]$ (x is a tangential prop. part of y)
$A_{NTPP} : NTPP(x, y) \leftrightarrow PP(x, y) \wedge \neg \exists z[EC(z, x) \wedge EC(z, y)]$ (x is a non-tang. prop. part of y)

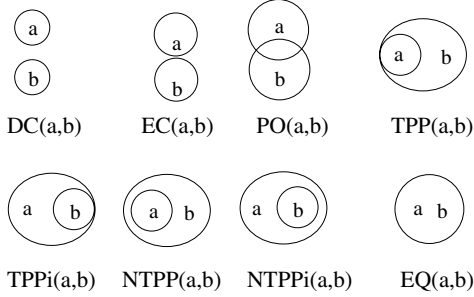


Figure 2. Axioms of RCC and RCC-8 relations

[4]. The basic relation between regions is the connection relation  $C(x, y)$ , which is interpreted as “the closures of  $x$  and  $y$  intersect”. The axioms are two basic axioms on  $C$ ,

$$A_1 := \forall x[C(x, x)], A_2 := \forall x, y[C(x, y) \rightarrow C(y, x)]$$

and several axioms on the main relationships (see fig. 2).

The eight jointly exhaustive and pairwise disjoint relations of figure 2 form the calculus RCC-8, that has been deeply studied in [16]. The RCC5 calculus is defined on the set  $\{DR, PP, PO, PPi, EQ\}$ . In [13] it is empirically analyzed which of RCC8 or RCC5 is more adequate for qualitative spatial reasoning, from the psychological point of view, and the authors have decided on RCC8. We will use *reticular* and *topological motions*. They are cognitively adequate motions which have a *continuous* nature. The first ones are refinements of relationships (downward motions in the lattice of the RCC relationships) and topological motions are motions of *least topological distance* (see fig. 3 and 3). The *reticular projection* on  $RCC_n$  ( $n = 5, 8$ ) is  $R \mapsto \hat{R} := \{R' \in RCC_n : R' \subseteq R\}$ .

#### 4.1. RCC as meta-ontology

For our purposes, RCC must be interpreted as a theory on properties of concepts; namely a *meta-ontology*: an ontology about the relationships among concepts of an ontology. The relation  $C(x, y)$  is intended as: *it seems that  $x, y$  have elements in common, but we can not affirm that they have a common subconcept*. An approach may be, given two concepts  $C_1, C_2$ ,  $C(C_1, C_2)$  holds iff  $\Sigma \models C_1 \sqcap C_2 \equiv \perp$ .

From this definition, every relation on RCC can be interpreted. The cognitive interpretation of the representation is completed by means of identifying the spatial border of a region with the elements which are capable of changes in their conceptual assignments.

However, the above interpretation is not usable in practice; we have to weaken that interpretation (sect. 7).

#### 5. Step 1: mereotopological interpretation

Each concept of  $\Sigma$  is interpreted as a region in the plane,  $I : concepts(\Sigma) \rightarrow \mathcal{P}(\mathbb{R}^2)$ , where  $A \mapsto A^I \subseteq \mathbb{R}^2$  (we will write  $A^I = A$ ). In order to carry out this interpretation, a translation of  $\Sigma$  to a CSP problem on RCC8 (or RCC5) is applied, translating the formulas of TBox to a set of RCC formulae as follows:

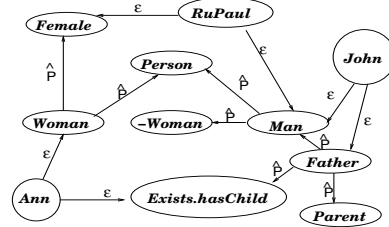
$$(C \sqsubseteq D)^* = \{P(C, D)\},$$

$$(C \sqsubseteq D_1 \sqcap D_2)^* = \{P(C, D_1), P(C, D_2)\},$$

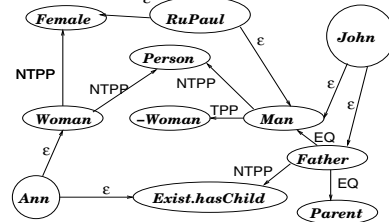
$$(C \sqsubseteq D_1 \sqcup D_2)^* = \{O(C, D_1), O(C, D_2)\}$$

(in the last case, it is possible that, for example,  $C \cap D_1 = \emptyset$ , but user will discard it later) and  $T^* = \bigcup_{F \in T} F^*$ .

Each fact is translated as  $A(a) \mapsto a \in A$ . For our example,



making reticular projection on RCC8, and solving the CSP,



The consistent scenario is spatially represented by a set of (not necessarily connected) regular regions of the plane (fig. 4). This representation always exists [15].

#### 6. Step 2: spatial arrangements

In this step the user is requested to make reticular and/or topological arrangements on the map (see fig. 3). By introducing new secondary regions, new concepts might be in-

troduced too. For our example, the new picture is figure 4. With respect to the interpretation of the facts in ABox, these facts must be intended from a *Closed World Assumption on spatial location* point of view: If the user accepted  $d \notin A^I$ , then we reason as if  $\neg A(d)$  was true.

### 7. Step 3: translating to a new KB

When the user believes that the current spatial *scenario* is a sound representation, it is translated to a new KB. It is necessary to make some remarks on the spatial scenario. Firstly, the spatial scenario may be inadequate with respect to the original KB (for example, when topological arrangements are used). On the other hand, the spatial relationships may be inadequate with respect to the *mental* ontology believed by the user. This anomaly is detected when the user refuses the translated KB, needing of a new graphical refinement. Actually, the translation must be applied to the table of relationships that the scenario represents.

Next we define the translation  $R \mapsto R^*$  of each relation on RCC to a set of DL formulae by recursion on the order of the RCC axiomatization (fig. 2). The relation  $a \in A$  is translated to  $A(a)$ . From now, “element” means “spatial interpretation of a constant symbol”.

1.  $C(A, B)^*$ : if regions  $A$  and  $B$  have an element in common in the representation, then  $C(A, B)^* = \emptyset$ . Otherwise a Skolem constant is introduced,  $c_{A,B}$ , and  $C(A, B)^* = \{A(c_{A,B}), B(c_{A,B})\}$ .
2.  $DC(A, B)^* = \{A \sqsubseteq \neg B\}$ .
3. It should be  $P(A, B)^* = \bigcup_{D \cap A \neq \emptyset} C(D, B)^*$ . But in practice, this induces many Skolem constants. It is chosen  $P(A, B)^* = \{A \sqsubseteq B\}$ .
4.  $PP(A, B)^*$ : if there exists a region  $D$  such that  $C(B, D) \wedge \neg C(A, D)$ , then  $PP(A, B)^* = P(A, B)^*$ . If such a region does not exist, a new concept name is introduced,  $N_{B \setminus A}$ , and  $PP(A, B)^* = P(A, B)^* \cup \{N_{B \setminus A} \sqsubseteq B, N_{B \setminus A} \sqsubseteq \neg A\}$ .
5.  $EQ(A, B)^* = \{A \equiv B\}$ .
6.  $O(A, B)^*$ : if there exists  $D$  such that  $P(D, A) \wedge P(D, B)$ , then  $O(A, B)^* = P(D, A)^* \cup P(D, B)^*$ . Otherwise we introduce a new concept name,  $N_{A \cap B}$ , and then  $O(A, B)^* = \{N_{A \cap B} \sqsubseteq A, N_{A \cap B} \sqsubseteq B\}$ . The spatial interpretation of the new concept is minimal, that is, with no elements.
7.  $DR(A, B)^*$ : It does not exist  $D$  such that  $P(D, A) \wedge P(D, B)$ , but it should exist an element in  $A \cap B$ . In that case, we reason by default: we introduce a concept of *abnormal* type,  $Ab_{A,B}$  and then  $DR(A, B)^* = \{Ab_{A,B}(d) : d \in A \cap B\}$ . Else,  $DR(A, B)^* = \emptyset$ .
8.  $PO(A, B)^*$  contains  $O(A, B)^*$  and reasoning with  $\neg P(A, B)$  and  $\neg P(B, A)$  in a similar way as we did at (4), we have  $PO(A, B)^* = O(A, B)^* \cup \{N_{B \setminus A} \sqsubseteq B, N_{B \setminus A} \sqsubseteq \neg A, N_{A \setminus B} \sqsubseteq A, N_{A \setminus B} \sqsubseteq \neg B\}$ .

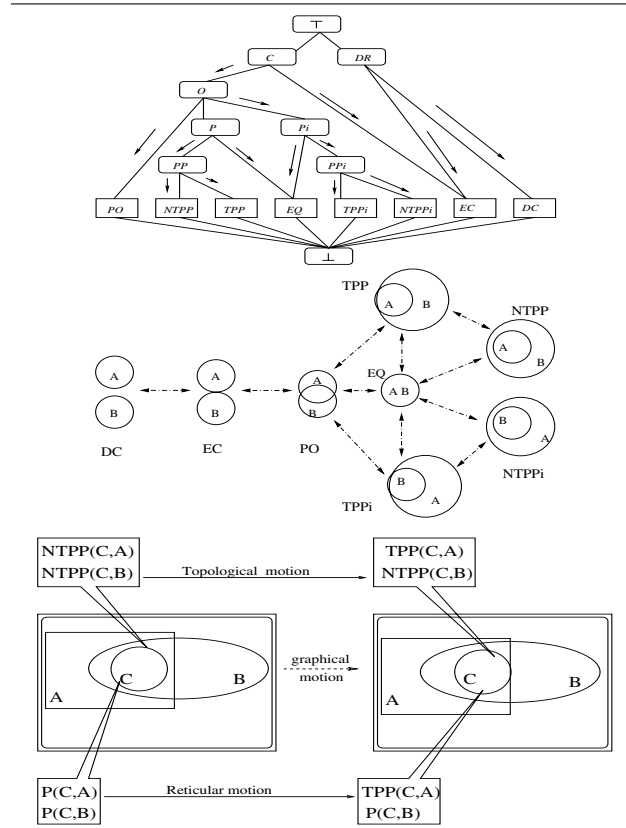


Figure 3. Reticular (up) and topological (down) motions in RCC8, and an example.

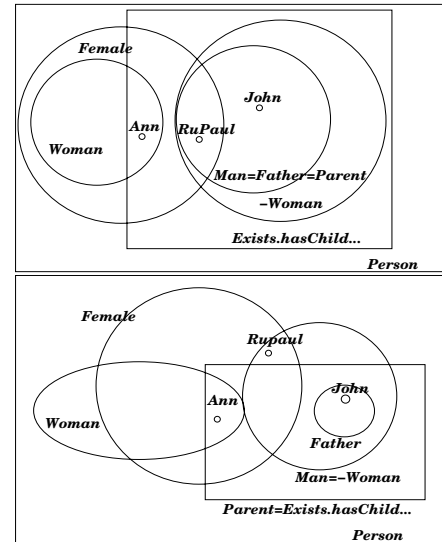


Figure 4. Representation of the solution and the result after making arrangements

	Female	Woman	Man	Father	Parent	Person	$\neg$ Woman	$\exists$ hasChild.P
Female	EQ	PO	PO	DC	PO	NTPP	PO	PO
Woman	PO	EQ	DC	DC	PO	NTPP	EC	PO
Man	PO	DC	EQ	NTPPi	PO	NTPP	EQ	PO
Father	DC	DC	NTPP	EQ	NTPP	NTPP	NTPP	NTPP
Parent	PO	PO	PO	NTPPi	EQ	NTPP	PO	EQ
Person (P)	NTPPi	NTPPi	NTPPi	NTPPi	NTPPi	EQ	NTPPi	NTPPi
$\neg$ Woman	PO	EC	EQ	NTPPi	PO	NTPP	EQ	PO
$\exists$ hasChild.P	PO	PO	PO	NTPPi	EQ	NTPP	PO	EQ

Figure 5. Table of spatial relationships among the concepts

9.  $EC(A, B)^*$  contains  $C(A, B)^*$  and the formulae we had obtained reasoning with  $\neg O(A, B)$  like in (7) are added. Thus, we have  $C(A, B)^*$  in first place. It can be  $\emptyset$ ; in this case, a concept is added,  $Ab_{A,B}$ , and  $EC(A, B)^* = \{Ab_{A,B}(d) : d \in A \cap B\}$ . In other case,  $C(A, B)^* \neq \emptyset$ , there exists a new element  $c_{A,B}$  and  $EC(A, B)^* = C(A, B)^* \cup \{Ab_{A,B}(c_{A,B})\}$ .

10.  $TPP(A, B)^*$ : If there exists  $C$  such that  $EC(C, A) \wedge EC(C, B)$ , then  $TPP(A, B)^* = PP(A, B)^* \cup EC(C, A)^* \cup EC(C, B)^*$ . Otherwise, it is necessary to add a new concept name,  $N_{A \supset B}$ , and then  $TPP(A, B)^* = PP(A, B)^* \cup EC(N_{A \supset B}, A)^* \cup EC(N_{A \supset B}, B)^*$ .

The intended spatial interpretation of  $N_{A \supset B}$  is *minimal*, that is, with no elements inside it.

11.  $NTPP(A, B)^*$  contains  $PP(A, B)^*$  and, reasoning as we do in (1), we can affirm that there is no region externally connected to both  $A$  and  $B$ . Thus, if there does not exist a region  $D$  such that  $EC(A, D)$ , then  $NTPP(A, B)^* = PP(A, B)^*$ . In other case  $NTPP(A, B)^* = PP(A, B)^* \cup O(D, B)^*$ .

With the translation it provides a model. Thus

**Theorem.** The KB obtained from a scenario by above interpretation is consistent.  $\blacksquare$

The notation of the new concept names does not refer to any intended feature. For example, it is not initially true that  $N_{A \cap B} \equiv A \cap B$ . In our running example, a table of relations is generated, see figure 5, and the KB obtained is

$$\mathcal{A}' = \begin{cases} \text{Female(Rupaul)}, \text{Man(Rupaul)} \\ \text{Father(John)}, \text{Man(John)} \\ \text{Parent(John)}, \text{Woman(Ann)} \\ \exists \text{hasChild.Person(Ann)}, \text{Female(Ann)} \\ \text{Woman}(C_{\text{Woman}, \neg \text{Woman}}) \\ \neg \text{Woman}(C_{\text{Woman}, \neg \text{Woman}}) \\ N_{\text{Man} \cap \text{Female}}(\text{Rupaul}) \\ Ab_{\text{Woman}, \neg \text{Woman}}(C_{\text{Woman}, \neg \text{Woman}}) \end{cases}$$

$$\mathcal{T}' = \begin{cases} \text{Father} \sqsubseteq \text{Person} \sqcap \text{Parent} \sqcap \text{Man} \\ \text{Father} \sqsubseteq \neg \text{Woman} \sqcap \neg \text{Female} \\ \text{Female} \sqsubseteq \text{Person} \\ \text{Man} \equiv \neg \text{Woman} \\ \text{Woman} \sqsubseteq \text{Person} \\ \text{Man} \sqsubseteq \text{Person} \\ \text{Parent} \sqsubseteq \text{Person} \\ \exists \text{hasChild.Person} \equiv \text{Parent} \\ N_{\text{Woman} \cap \text{Female}} \sqsubseteq \text{Woman} \sqcap \text{Female} \\ N_{\text{Woman} \setminus \text{Female}} \sqsubseteq \text{Woman} \sqcap \neg \text{Female} \\ N_{\text{Female} \setminus \text{Woman}} \sqsubseteq \text{Female} \sqcap \neg \text{Woman} \\ N_{\text{Parent} \cap \text{Woman}} \sqsubseteq \text{Parent} \sqcap \text{Woman} \\ N_{\text{Parent} \setminus \text{Woman}} \sqsubseteq \text{Parent} \sqcap \neg \text{Woman} \\ N_{\text{Woman} \setminus \text{Parent}} \sqsubseteq \text{Woman} \sqcap \neg \text{Parent} \\ N_{\text{Man} \cap \text{Female}} \sqsubseteq \text{Man} \sqcap \text{Female} \\ N_{\text{Man} \setminus \text{Female}} \sqsubseteq \text{Man} \sqcap \neg \text{Female} \\ N_{\text{Female} \setminus \text{Man}} \sqsubseteq \text{Female} \sqcap \neg \text{Man} \\ N_{\text{Parent} \cap \text{Female}} \sqsubseteq \text{Parent} \sqcap \text{Female} \\ N_{\text{Parent} \setminus \text{Female}} \sqsubseteq \text{Parent} \sqcap \neg \text{Female} \\ N_{\text{Female} \setminus \text{Parent}} \sqsubseteq \text{Female} \sqcap \neg \text{Parent} \\ N_{\text{Parent} \setminus \text{Man}} \sqsubseteq \text{Parent} \sqcap \neg \text{Man} \\ N_{\text{Man} \setminus \text{Parent}} \sqsubseteq \text{Man} \sqcap \neg \text{Parent} \end{cases}$$

## 8. Step 4: Evaluation by the user

One of the aims of this step is to give a name for new concepts (*if something does not have a name, it does not exist*): the user must make the effort of interpreting the result. Moreover, the user must decide which elements belong to the new concepts, if they are topologically close. For example, if  $Rupaul \in N_{\text{Female} \setminus \text{Woman}}$ . It is also possible that a new concept may be discarded. This might happen if the graphic representation made by the user has become inadequate for her/his beliefs.

In our case, the user decides to denote  $N_{\text{Female} \setminus \text{Woman}}$  as *Crossdresser*. And  $N_{\text{Parent} \cap \text{Woman}}$  is named *Mother*. Thus

$$\mathcal{T}'' = \begin{cases} \text{Father} \sqsubseteq \text{Parent} \sqcap \text{Man} \sqcap \neg \text{Female} \\ \text{Female} \sqsubseteq \text{Person}, \quad \text{Woman} \sqsubseteq \text{Person} \\ \text{Man} \sqsubseteq \text{Person}, \quad \text{Parent} \sqsubseteq \text{Person} \\ \text{Man} \equiv \neg \text{Woman} \\ \exists \text{hasChild.Person} \equiv \text{Parent} \\ \text{Crossdresser} \sqsubseteq \text{Female} \\ \text{Crossdresser} \sqsubseteq \neg \text{Woman} \\ \text{Mother} \sqsubseteq \text{Parent} \sqcap \text{Woman} \end{cases}$$

and the new Abox:

$$\mathcal{A}'' = \mathcal{A}' \cup \{\text{Mother}(\text{Ann}), \text{Crossdresser}(\text{Rupaul})\}$$

Since the user believes that  $EC(Woman, \neg Woman)$ , the translation produces  $Ab_{Woman, \neg Woman}$ , and  $Ab_{Woman, \neg Woman}(C_{Woman, \neg Woman})$  is included into the Abox, but no relation between  $Ab_{Woman, \neg Woman}$  and  $Woman$  or  $\neg Woman$  is added. It is preferable to be a notion. (It is not decided if it is a new concept or an special case to discard). It would be possible that  $N_{Man \setminus Father}$  appears. To assign  $UnmarriedMan$  to  $N_{Man \setminus Father}$  means that the user thinks that this concept is a notion still (the intended notion  $UnmarriedMan$  is more restricted). This could lead to subsequent revisions. All the actions described here preserve consistency.

## 9. Conclusions and future work

We showed how to use RCC in order to provide a formal semantics to repair ontologies by means of the mereotopological theory RCC. This work allows us to extend [2] (on arguments for spatial databases are analysed) to general ontologies. Note, also, that our initial aim is not a graphical tool to manage ontologies (as the hyperbolic inference engine [14]). We propose a spatial semantics to support ontological cleaning to solve the absence of a formal semantics on which to support graphical arrangements of spatial representations. However it is evident that the implementation of a cleaning tool based in this semantics will be a future work.

The cycle is not, strictly speaking, a tool for reasoning services based on entailment. Non monotonic reasoning is used in several steps of the cleaning process. This is more related with the mereological analysis of concepts, as [10].

There exist two research lines we are currently studying on. First, although in this paper we do not deal with a spatial representation of the roles of KB, this feature should be added. And second, the study of the relation between the original KB and the new one is to be done. In fact, it may consider the tool as a learning process (based on generalization and refinement of concepts) or, properly, an interactive knowledge acquisition mechanism. Thus it is interesting to design metrics reflecting such a process by means of adapting techniques from clausal learning [11].

## References

- [1] J.A. Alonso-Jiménez, J. Borrego-Díaz, A.M. Chávez-González, M. A. Gutiérrez-Naranjo and J. D. Navarro-Marín. A Methodology for the Computer-Aided Cleaning of Complex Knowledge Databases. Proc. of IECON 2002. IEEE Press. pp. 1806–1811 (2002).
- [2] J.A. Alonso-Jiménez, J. Borrego-Díaz, A. M. Chávez-González, M. A. Gutiérrez-Naranjo and J. D. Navarro-Marín, Towards a Practical Argumentative Reasoning with Qualitative Spatial Databases. Proc. of 16th Int. Conf. on

- Industrial and Engineering Applications of Artificial Intelligence and Expert Systems IEA/AIE 2003, Lecture Notes in AI n. 2718, pp. 789-798, Springer-Verlag (2003).
- [3] G. Antoniou and F. van Harmelen. Web Ontology Language: OWL. In S. Staab and R. Studer (eds.) Handbook on Ontologies in Information Systems, pp.67-92, Springer-Verlag (2003).
- [4] A. G. Cohn, B. Bennett, J. M. Gooday and N. M. Gotts. Representing and Reasoning with Qualitative Spatial Relations about Regions. In O. Stock (ed.) Temporal and Spatial Reasoning, pp. 97-134. Kluwer (1997).
- [5] Y. Ding, D. Fensel, M. Klein, B. Omelayenko and E. Schulten. The role of ontologies in e-Commerce. In S. Staab and R. Studer (eds.) Handbook on Ontologies in Information Systems, pp. 593-616, Springer-Verlag, (2003).
- [6] F. M. Donini, M. Lenzerini, D. Nardi, and A. Schaerf. Reasoning in description logics. In G. Brewka (ed.) Foundation of Knowledge Representation, pp. 191-236. CSLI-Publications (1996).
- [7] M. Elvang-Goransson and A. Hunter. Argumentative logics: Reasoning from classically inconsistent information. Data and Knowledge Engineering, 16, pp. 125-145 (1995).
- [8] D. Fensel: Ontologies: Silver Bullet for Knowledge Management and E-Commerce. Springer-Verlag (2000).
- [9] H. Galhardas, D. Florescu, D. Sasha, E. Simon, C.A. Saita. Declarative Data Cleaning: Language, Model and Algorithms. Proc. of 27th Very Large Databases Conference VLDB. Morgan Kauffman (2001).
- [10] N. Guarino and C. Welty. Towards a methodology for ontology-based model engineering. In Proc. of ECOOP-2000 Workshop on Model Engineering (2000).
- [11] M. A. Gutiérrez-Naranjo, J. A. Alonso-Jiménez and J. Borrego-Díaz. A quasi-metric for Machine Learning. Advances in Artificial Intelligence-IBERAMIA 2002. Lecture Notes in AI n. 2527, pp. 193–203 (2002).
- [12] I. Horrocks, D.L. McGuinness, C. A. Welty. Digital Libraries and Web-Based Information Systems. In F. Baader et al. (ed.) The Description Logic Handbook. pp. 436-459. Cambridge University Press (2003).
- [13] M. Knauff, R. Rauh and J. Renz. A cognitive assessment of topological spatial relationships: Results from an empirical evaluation. Proc. of the 3rd Int. Conf. on Spatial Information Theory (COSIT'97). Lecture Notes in Computer Science n. 1329, pp. 193–206 (1997).
- [14] J. Lamping, R. Rao, and P. Pirolli. A focus + context technique based on hyperbolic geometry for visualizing large hierarchies. in ACM Conference on Human Factors in Computing Systems CHI'95 (1995).
- [15] J. Renz. A Canonical Model of the Region Connection Calculus. Proc. of the 6th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'98) (1998).
- [16] J. Renz, B. Nebel. On the Complexity of Qualitative Spatial Reasoning: A Maximal Tractable Fragment of the Region Connection Calculus. Artificial Intelligence 108(1-2), pp. 69-123 (1999).