

Visual Ontology Cleaning: Cognitive Principles and Applicability

Joaquín Borrego-Díaz and Antonia M. Chávez-González

Departamento de Ciencias de la Computación e Inteligencia Artificial.
E.T.S. Ingeniería Informática-Universidad de Sevilla.
Avda. Reina Mercedes s.n. 41012-Sevilla, Spain
{jborrego, tchavez}@us.es

Abstract. In this paper we connect two research areas, the Qualitative Spatial Reasoning and visual reasoning on ontologies. We discuss the logical limitations of the mereotopological approach to the visual ontology cleaning, from the point of view of its formal support. The analysis is based on three different spatial interpretations which are based in turn on three different spatial interpretations of the concepts of an ontology.

1 Introduction

It is commonly accepted that for achieving a satisfactory sharing of knowledge in the envisioned Semantic Web (SW), it will be necessary to build ontologies. They play a key role in the reasoning services for Knowledge Bases (KB) in the SW [17]. Practical management of ontologies, such as extension, refinement and versioning technologies will be essential tasks.

From the point of view of the Knowledge Representation (KR) paradigm, ontology revision comes from the fact that the discourse domain may not be faithfully represented by an ontology (a well known working principle in KR). In many cases, end-users need to interact and transform the ontology. Even if the ontology designer thinks that the ontology is final, the end-user may think the opposite, or simply that the ontology is incorrect. In fact, it should be feasible to achieve the agreement designer-user. This agreement is essential for the assimilation of SW technologies into non-academic community portals, for example.

Unfortunately, several reasons obstruct the agreement. The first one is that end users do not know the logical formalisms behind ontology web languages. Therefore, the user can not know hidden principles on which ontologies are built. It does not help to increase the understanding of technologies involved in SW tools. Anyway, this fact might not be important if he uses amenable technologies for representing/repairing the anomalies founded in its own ontology project. Visual encodings are very interesting for such purposes.

End-user preferences on visual representation are well known in other related fields such as Formal Concept Analysis or Data Engineering. The spatial

* Supported by project TIN2004-03884 of Spanish Ministry of Education and Science, cofinanced by FEDER funds.

metaphor is a powerful tool in human information processing. The user will feel encouraged to repair the anomaly, although there exist some obstacles: on the one hand, visual reparation may not be corresponded by a logical reparation of the ontology source. This occurs if there is no a formal semantics for supporting the change; on the other hand, repairs can be logically complex. Domain experts often underestimate the amount of time required to produce an ontology, and consequently they build an ontology based on a large scope. The resulting conceptual ontologies are consequently a mix of both domain and task ontology concepts which are hard to manage [22].

Paraphrasing [23], visual cleaning of ontologies is important for future end-users of ontology debugging systems due mainly to three reasons:

1. It allows the user to summarize ontology contents.
2. User's information is often fuzzily defined. Visualization can be used to help the user to get a nice representation.
3. Finally, visualization can therefore help the user to interact with the information space.

There is not a generally accepted representation mechanism that translates every possible changes in the visual representation into the specification of the ontology. In fact, this is an interesting problem in the design of visual reasoning tools. Current end-user tools are mostly based on facilitating the understanding of the ontology (see e.g. [14], [21]) facilitating very limited graphical changes to the user. In order to augment such features, we need formally sound mappings between (visual) representations and Knowledge Bases (KBs) (expressed, for example, in Description Logics). Note that such mappings have to translate logical notions for supporting the logical impact of arrangements on the spatial representation (for example, when new concepts are inserted). These issues are critical and we need to solve them in order to integrate solutions in systems for visual representation of information [14]. This goal is far away of being achieved for classical Information Visualization (IV) tools. IV is the use of computer-supported, interactive and visual representations of abstract data to amplify cognition [9]. The goal of Visual Ontology Cleaning (VOC) should be to reason spatially for visually debugging and repairing of ontologies. Therefore, it should have additional features, different from classical user analysis, querying and navigation/browsing.

A second limitation concerns to the scalability of debugging problem. It is hard to manage visual representations of large ontologies, although a broad spectrum of tools has been designed [14]. It is sometimes sufficient to locate which small portion of the ontology supports the anomaly. This task can be facilitated by Automated Reasoning Systems (ARS). ARS are useful both for debugging ontologies [1] [2][27] (although several foundational problems exist [4]), and for the computer-assisted evolution of robust ontologies [6] [7]. If an ARS find a proof of an anomalous result, it can recover an *argument* from the proof. An argument is a pair formed by a consistent portion of the KB and the entailed result (an *ontological argument*). In this case, it seems natural to consider the repairing of the self argument to cut such inference [3]. The argument is often a very

small portion of the ontology, and therefore it is easily represented. Despite its modest size, it provides more useful information about the anomaly than the full ontology. Furthermore, the reparation is model-based. That is, the user keep in mind the intended model that ontology represents, and this model induces the changes. Therefore, argument repairing is a relatively easy task. Thus, for the purposes of this paper, we might consider the ontological argument as the anomalous ontology.

In this paper we investigate some KR issues behind the sound mereotopological representation of the conceptualization induced by small ontologies (namely the mentioned ontological arguments) [3]. The intended aim of such representation is the understanding and repairing of ontologies. Specifically, those considered as anomalous (although consistent) due to errors in the concept structure. In this paper the spatial representation and algorithmic repairing are described; we do not describe here the (future) implementation. The advantage of this approach is that visual reparation stage hides formal semantics that supports the change, facilitating in this way its use by non experts.

The structure of the paper is as follows. The next section introduces a well-known mereotopological approach to qualitative spatial reasoning. Section 4 is devoted to study whether the Mereotopological reasoning can be interpreted as an abstract metaontology. It is proved, using logical types (sect. 3) that such interpretation does not support the work with instances inaccurately classified. In sections 5,6 we suggest new approaches based on the representation of frontiers and vague regions, respectively. The paper concludes with some closing remarks about the presented framework.

2 Mereotopology and First Cognitive Principle

The thesis that supports this paper is the following principle

Main Cognitive Principle (MCP): If we aim to use spatial reasoning for cleaning ontologies, we have to provide a theory on spatial entities for translating the impact of spatial arrangements into revisions of the ontology source.

In order to satisfy the MCP, a theory on Qualitative Spatial Reasoning (QSR) has to be selected. In this way, the following sub-principle is chosen:

First Cognitive Principle (CP1): The concepts of a conceptualization associate to a clear ontology can be topologically represented by means of regular non-empty regions.

That is, there is a model of the ontology whose universe is the bidimensional or tridimensional space, and that model interprets concept symbols as regions. It is evident that the represented knowledge will depend of topological relations. The starting-up of CP1 needs of a robust theory to reason with spatial regions.

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

Fig. 1. Axioms of RCC

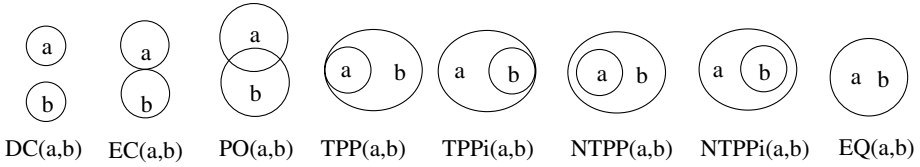


Fig. 2. The relations of RCC8

Additionally, the theory must facilitate the knowledge interchange between the ontology and spatial models.

The selected theory is the well known *Region Connection Calculus* (RCC) [13]. RCC is a mereotopological approach to QSR; it describes topological features of spatial relationships. It has been used in several subfields of AI, for example, in GIS and spatial databases [24] [16].

In RCC, the *spatial entities* are non-empty regular sets. The ground relation is the *connection*, $C(x, y)$, with intended meaning: “*the topological closures of x and y intersect*”. The basic axioms of RCC are

$$\forall x[C(x, x)] \quad \forall x, y[C(x, y) \rightarrow C(y, x)]$$

and a set of definitions on the main spatial relations (fig. 1), jointly with another set of auxiliary axioms (see [13]).

The set of binary relations formed by the eight jointly exhaustive and pairwise disjoint (JEPD) relations given in figure 2 is denoted by RCC8. If this set is thought as a calculus for Constraint Satisfaction Problems (CSP), every set of basic relations is considered. This calculus has been deeply studied by J. Renz and B. Nebel in [26]. Other interesting calculus is RCC5, based on the set $\{DR, PO, PP, PPi, EQ\}$. Roughly speaking, the main difference between RCC5 and RCC8 is that the latter one allows one to represent knowledge that depends on topological frontiers, while the former one does not allow. The cognitive impact of this distinction on the spatial representation of a concept has to be discussed (as we will do, in fact). Nevertheless, it has been empirically constated that RCC8 is more adequate than RCC5 as a tool for representing topological relations discriminated by humans [19].

3 Background: 1-Types in Ontologies

For analyzing the first spatial interpretation showed in this paper types are required. The use of types is a classic tool in Model Theory (see e.g. [11]). We succinctly describe here their main features. Given a consistent ontology Σ , a 1-*type* is a (maximal) set of formulas $\{\varphi_k(x)\}_k$ finitely satisfiable. That is, such that for any $F \subseteq \mathbb{N}$ (F finite), the theory

$$\Sigma + \{\exists x \bigwedge_{k \in F} \varphi_k(x)\}$$

is consistent. The ontology Σ *realizes* the type if that theory is consistent when $F = \mathbb{N}$. Equivalently, Σ realizes the type if the theory $\Sigma + \{\varphi_k(a) : k \in \mathbb{N}\}$ is consistent, being a a new constant symbol.

For a correct definition of types it has to consider First Order Logic (FOL) formulas. Types contain formulas that can not be translated into DL. Nevertheless, general FOL formulas are not needed in the proof of Theorem 1 below. Only the constructors *negation* and *intersection* are needed.

The set $S(\Sigma) = \{p : p \text{ is a type of } \Sigma\}$ is the *space of types* for Σ . For a concept C , let $S_\Sigma(C) := \{p \in S(\Sigma) : C \in p\}$ be the set of types containing C . It will be denoted by $[C]$ if Σ has been previously fixed. It is easy to see that

$$[C] \cap [D] = [C \wedge D] \text{ and } [\neg C] = S(\Sigma) \setminus [C]$$

Given a model $I \models \Sigma$ and a an element of its universe, the *type of a in I* is $[a] = tp_I(a) = \{\varphi(x) : I \models \varphi(a)\}$. Fixed I , it will be denoted $[a]$.

4 RCC as a Meta-ontology (I): Strong Interpretation

The use of RCC to visually represent the concepts turns RCC8 into an ontology on conceptual relations. The idea can be translated in different ways.

The straightforward approach consists in interpreting the concepts as regions in some model of the theory. Thus, in the *strong interpretation*, the intended meaning of $C(x, y)$ is: *there exist a common element in the concepts x, y in some model I of the ontology*.

Definition 1. (Strong Interpretation of RCC as a metaontology) *Two concepts C_1, C_2 of an ontology Σ are Σ -connected (denoted by $C_\Sigma(C_1, C_2)$) if*

$$\Sigma \not\models C_1 \sqcap C_2 \equiv \perp$$

The remaining RCC relations can be interpreted by means of its corresponding definition (depicted in fig. 1). Note that the strong interpretation works on abstract spatial encodings of Σ . That is, it does not work on a concrete spatial interpretation of concepts. The following result states a logical limitation of the strong interpretation of RCC as meta-ontology.

Theorem 1. *The strong interpretation does not discriminate RCC8 as ontological relations between concepts. Concretely, it has the following characterizations:*

1. $C_\Sigma(\mathbf{C}_1, \mathbf{C}_2) \iff S_\Sigma(\mathbf{C}_1) \cap S_\Sigma(\mathbf{C}_2) \neq \emptyset$
2. $DC_\Sigma(\mathbf{C}_1, \mathbf{C}_2) \iff S_\Sigma(\mathbf{C}_1) \cap S_\Sigma(\mathbf{C}_2) = \emptyset$
3. $P_\Sigma(\mathbf{C}_1, \mathbf{C}_2) \iff S_\Sigma(\mathbf{C}_1) \subseteq S_\Sigma(\mathbf{C}_2)$
4. $PP_\Sigma(\mathbf{C}_1, \mathbf{C}_2) \iff S_\Sigma(\mathbf{C}_1) \subsetneq S_\Sigma(\mathbf{C}_2)$
5. $EQ_\Sigma(\mathbf{C}_1, \mathbf{C}_2) \iff \Sigma \models \mathbf{C}_1 \equiv \mathbf{C}_2.$
6. $O_\Sigma(\mathbf{C}_1, \mathbf{C}_2) \iff C_\Sigma(\mathbf{C}_1, \mathbf{C}_2)$
7. $PO_\Sigma(\mathbf{C}_1, \mathbf{C}_2) \iff \begin{cases} S_\Sigma(\mathbf{C}_1) \cap S_\Sigma(\mathbf{C}_2) \neq \emptyset \wedge S_\Sigma(\mathbf{C}_1) \not\subseteq S_\Sigma(\mathbf{C}_2) \wedge \\ \wedge S_\Sigma(\mathbf{C}_2) \not\subseteq S_\Sigma(\mathbf{C}_1) \end{cases}$
8. $DR_\Sigma(\mathbf{C}_1, \mathbf{C}_2) \iff DC_\Sigma(\mathbf{C}_1, \mathbf{C}_2)$
9. *If $\mathbf{C}_1, \mathbf{C}_2$ and $R \in \{EC, TPP, NTPP, TPPi, NTPPi\}$, then $\neg R_\Sigma(\mathbf{C}_1, \mathbf{C}_2)$.*

Proof. Let us only show three cases (the complete proof appears in [10]).

- (1) By definition, $C_\Sigma(\mathbf{C}_1, \mathbf{C}_2)$ means $\Sigma \not\models \mathbf{C}_1 \sqcap \mathbf{C}_2 \equiv \perp$. Let I be a model with $\mathbf{C}_1^I \cap \mathbf{C}_2^I \neq \emptyset$. By interpreting a new constant a as an element of the intersection, it certifies that $\Sigma \cup \{(\mathbf{C}_1 \sqcap \mathbf{C}_2)(a)\}$ is consistent. In fact,

$$[a] \in [\mathbf{C}_1 \sqcap \mathbf{C}_2]$$

therefore, $[\mathbf{C}_1 \sqcap \mathbf{C}_2] \neq \emptyset$. Hence it holds $S_\Sigma(\mathbf{C}_1) \cap S_\Sigma(\mathbf{C}_2) \neq \emptyset$. The converse is trivial: if $p \in S_\Sigma(\mathbf{C}_1) \cap S_\Sigma(\mathbf{C}_2)$, any element realizing this type in a model I is an element of $\mathbf{C}_1^I \cap \mathbf{C}_2^I$. Thus, $C_\Sigma(\mathbf{C}_1, \mathbf{C}_2)$.

- (3) Assume proved (1) and (2). By definition of P , $P_\Sigma(\mathbf{C}_1, \mathbf{C}_2)$ means

$$\forall \mathbf{C}_3 (S_\Sigma(\mathbf{C}_3) \cap S_\Sigma(\mathbf{C}_1) \neq \emptyset \rightarrow S_\Sigma(\mathbf{C}_3) \cap S_\Sigma(\mathbf{C}_2) \neq \emptyset) \quad (\dagger)$$

We proceed to show that this condition is equivalent to $S_\Sigma(\mathbf{C}_1) \subseteq S_\Sigma(\mathbf{C}_2)$ (thus, (3) will be proved):

It suffices to show that the condition is necessary.

Let us consider $p \in S_\Sigma(\mathbf{C}_1)$ (that is, $\mathbf{C}_1 \in p$). Suppose, contrary to our claim, that $p \notin S_\Sigma(\mathbf{C}_2)$. Hence $\mathbf{C}_2 \notin p$. Since p is maximal, $\neg \mathbf{C}_2 \in p$. Therefore, $\mathbf{C}_1 \in p$ and $\neg \mathbf{C}_2 \in p$, thus

$$p \in S_\Sigma(\mathbf{C}_1) \cap S_\Sigma(\mathbf{C}_1 \sqcap \neg \mathbf{C}_2)$$

Then it follows that $S_\Sigma(\mathbf{C}_1) \cap S_\Sigma(\mathbf{C}_1 \sqcap \neg \mathbf{C}_2) \neq \emptyset$. Consequently, by (\dagger)

$$S_\Sigma(\mathbf{C}_2) \cap S_\Sigma(\mathbf{C}_1 \sqcap \neg \mathbf{C}_2) \neq \emptyset$$

From this, we obtain $p' \in [\mathbf{C}_1 \sqcap \neg \mathbf{C}_2]$ and $p' \in [\mathbf{C}_2]$. This is impossible because $\{\mathbf{C}_1 \sqcap \neg \mathbf{C}_2, \mathbf{C}_2\}$ can not be contained in a type (it is inconsistent).

- (5) Suppose proved from (1) to (4). By definition, $EQ_\Sigma(\mathbf{C}_1, \mathbf{C}_2)$ is

$$P_\Sigma(\mathbf{C}_1, \mathbf{C}_2) \wedge P_\Sigma(\mathbf{C}_2, \mathbf{C}_1)$$

By (3), this is equivalent to $S_\Sigma(\mathbf{C}_1) \subseteq S_\Sigma(\mathbf{C}_2)$ and $S_\Sigma(\mathbf{C}_2) \subseteq S_\Sigma(\mathbf{C}_1)$. Therefore $S_\Sigma(\mathbf{C}_1) = S_\Sigma(\mathbf{C}_2)$.

Finally, the last condition is equivalent to $\Sigma \models \mathbf{C}_1 \equiv \mathbf{C}_2$. It is trivial to see that it is sufficient. Let us see that it is necessary:

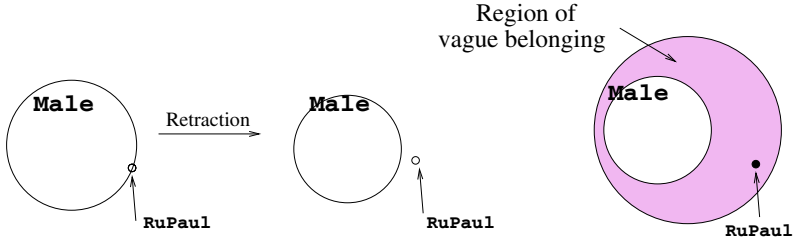


Fig. 3. Two kinds of representation of a problematic instance in the frontier (left, right) and solving by retraction the first one (middle) (right)

Let $I \models \Sigma$ and a some realization of C_1 in I , $I \models C_1(a)$. Thus,

$$[a] = \{\phi : I \models \phi(a)\} \in S_\Sigma(C_1)$$

So $[a] \in S_\Sigma(C_2)$. Therefore, $I \models \forall x(C_1(x) \rightarrow C_2(x))$. In DLs, $I \models C_1 \sqsubseteq C_2$. Similarly, $I \models C_2 \sqsubseteq C_1$ holds. Thus, $\Sigma \models C_1 \equiv C_2$.

It could be pointed out that the above interpretation might be not usable in practice. The connection between concepts is based on entailment. Therefore, it can have high algorithmic complexity when we deal with expressive description logics (the consistency of concepts and, hence, satisfiability, is EXPTIME-hard in \mathcal{ALC} , cf. [8]). Strong connection reduces to subsumption (hence to a satisfiability problem):

$$C_\Sigma(C_1, C_2) \text{ if and only if } \Sigma \not\models C_1 \sqsubseteq \neg C_2$$

Theorem 1 says that, thinking RCC as a metaontology, this theory can not represent *frontier-sensitive* knowledge. Nevertheless, sometimes it seems cognitively natural to consider the frontier of a concept as formed by elements in which the user has not confidence on its correct classification in the concept. This can occur when the user works with a *notion*, a rough idea of the concept that attempts to specify. The notion becomes in a *concept* when its behavior is constrained by new axioms that relate it with the former concepts. By allocating the problematic instances near of the frontier it will be easy to discard them by a simple topological movement (a retraction, see fig. 3).

5 RCC as Meta-ontology (II): Weak Interpretation

As it has been commented, strong interpretation can be not advisable due to its computational complexity for ontological arguments of larger size. Other limitation is that our cognitive capability is reduced to RCC5. In this section a new interpretation is introduced. This is based on the following idea: if it were possible to represent RCC8-relations in visual encodings, the topological frontiers of the regions could be endowed of cognitive features. The following principle is useful for both the *weak interpretation* defined bellow and the *vague interpretation* defined in the next section (see figure 3).

Second Cognitive Principle (CP2): The frontier of a spatial interpretation of a concept C represents the individuals with doubtful membership to C .

The ontology RCC has been used as formal support for CP2 in ontology cleaning [3]. The idea is based on the translation of relations between concepts into spatial relationships among spatial representations of these concepts. It is, at the same time, based on studies about the relation between logical consistency of Constraint Satisfaction Problems in RCC and spatial consistency [25]. The main step of the cited cleaning cycle consists in a translation of logical information on conceptual relations of the ontology into a CSP on spatial relations. Solving this CSP, a spatial encoding of conceptual relations is obtained. This approach is useful to repair arguments suffering anomalies due to the conceptualization. It works with *spatial interpretations*.

Definition 2. A *spatial interpretation* I of Σ is a interpretation in the language of Σ , such that $I : \text{concepts}(\Sigma) \cup \text{indiv}(\Sigma) \rightarrow \Omega$, where Ω is a T_3 connected topological space such that $I \models \Sigma$ and for each $C \in \text{concepts}(\Sigma)$, $I(C)$ is an open regular set in Ω and for each $a \in \text{indiv}(\Sigma)$, $I(a)$ is a point.

A spatial model of Σ is a spatial interpretation which is a model of Σ .

The following theorem guarantees that the weak interpretation is useful for analysing knowledge bases in DL, since it preserves the consistency.

Theorem 2. [10] *The CSP associate to Σ is spatially consistent if Σ is consistent.*

Moreover, it is possible to obtain a spatial model on the plane formed by polygonal regions [25].

Once formalized the notion of spatial model, spatial arrangements can be formally justified and classified [10]. In figure 4 we present a simple example of visual repairing (of the ontology of Figure 7, left) following a cleaning cycle presented in [3] (see fig. 6). Note that, although one can think that the ontology source is correct, it is assumed that the end-user thinks it is anomalous (he thinks that it is an ontological argument that deficiently classifies to *Rupaul*). The cycle is based on the following stages:

1. First, it builds a constraint satisfaction problem (CSP) on the spatial relational calculus RCC8 (or RCC5). 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 into DL is applied. In the resulting ontology, some relationships have changed and new concepts may have been induced.

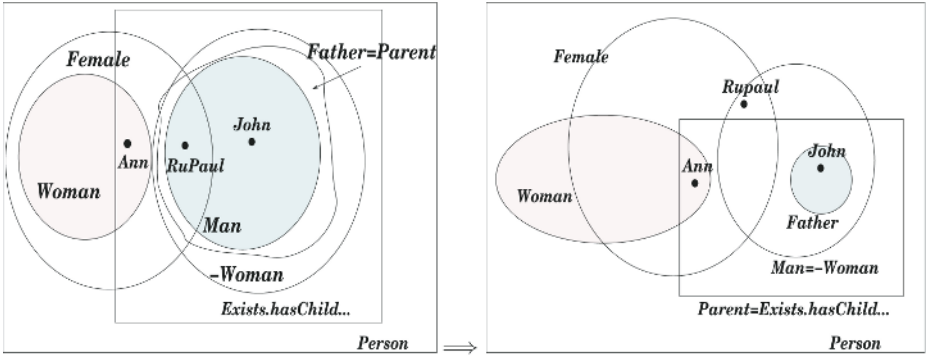


Fig. 4. Spatial representation of an anomalous argument (left) and the solution proposed (right)

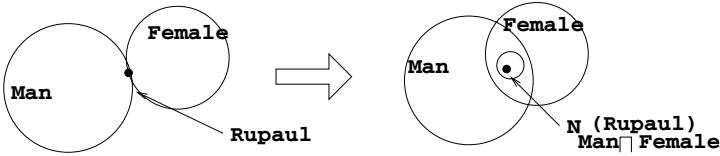


Fig. 5. New Notion induced. The user has to interpret/discard it, because it transforms $DR_{\Sigma}(Man, Female)$ into $PO_{\Sigma}(Man, Female)$

4. Finally, the user interprets (or discards) the new objects (individuals or concepts) that the translation may induce (see Fig. 5).

The result of the process is a new ontology modified according to the preferences of the user. The impact of this arrangement on the ontology is shown in fig. 7 (right). Furthermore, in complex or huge ontologies, it is convenient to use, for cleaning tasks, other relaxed spatial interpretations. Since the ontology we are representing is regarded as defective, it is possible to make a spatial characterization less detailed than the offered one by the CSP [3].

There is a natural relationship between strong and weak interpretations.

Theorem 3. [10] *The following conditions are equivalent:*

1. $C_{\Sigma}(C_1, C_2)$
2. *There is a spatial interpretation I of Σ such that $I \models C(C_1, C_2)$.*

6 RCC as Meta-ontology (III): Vague Interpretation

Strong and weak interpretation work with precise regions. In both cases, the interpretation of C is a subset. The *vague interpretation* deals with the spatial interpretation of the concepts by *vague* regions. A vague region can be represented by means of two regular regions although there are other options such as egg-yolk [12], topological spaces with pulsation [10], rough sets, etc.

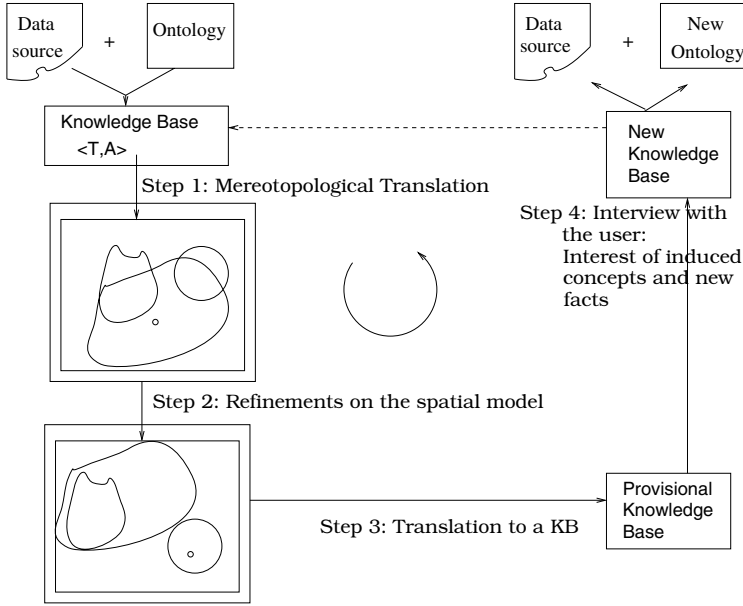


Fig. 6. Visual Ontology Cleaning Cycle [3]

In order to work with vague regions, a robust extension of RCC ([6, 7, 10]) is needed. The extension needs the re-interpretation of the ontology. In figure 8 we present one of the seven possible robust extensions of the ontology RCC given in [6]. The interpretation is based on *pulsation*. A pulsation in a topological space $\Omega = (\mathcal{X}, \mathcal{T})$ is a map that associates to each regular set X a set $\sigma(X)$ such that its closure contains the closure of X ; $\overline{X} \subseteq \sigma(X)$. In Fig. 8, the topological interpretation of the new relation $I(a, b)$ is $PP(a, b) \wedge EQ(\sigma(a), \sigma(b))$ (see Fig. 8, right). The reasoning of vague regions is based on the following principle:

Third Cognitive Principle (CP3): Given a spatial interpretation I , the region $\sigma(I(C)) \setminus I(C)$ represents the set of individuals with doubtful membership to C

In order to apply this principle for visual encoding, it considers the concept and its *approximate definition* in the ontology.

From now on, it is assumed that Σ is an unfoldable DL-ontology, that is, the left-hand sides of the axioms (defined concepts) are atomic and the right hand sides contain no direct or undirect references to defined concepts.

Definition 3. Let Σ be a DL ontology. The approximate definition according to Σ is a map σ that associates to any $C \in \text{concepts}(\Sigma)$ a DL-formula as follows:

$$\sigma(C) = \begin{cases} C, & \text{if } C \text{ is a defined concept} \\ \bigcap \{D : C \sqsubseteq D \in \Sigma\}, & \text{if } C \text{ is a primitive concept} \\ \top, & \text{if } C \text{ is an atomic concept} \end{cases}$$

$$\begin{array}{l}
\mathcal{T} = \left\{ \begin{array}{l} \text{Woman} \sqsubseteq \text{Person} \sqcap \text{Female} \\ \text{Man} \sqsubseteq \text{Person} \sqcap \neg \text{Woman} \\ \text{Father} \sqsubseteq \text{Man} \sqcap \exists \text{hasChild.Person} \\ \text{Father} \sqsubseteq \text{Parent} \end{array} \right. \\
\mathcal{A} = \left\{ \begin{array}{l} \text{Father}(\text{John}) \\ \text{Man}(\text{John}) \\ \text{Female}(\text{RuPaul}) \\ \text{Man}(\text{RuPaul}) \\ \text{Woman}(\text{Ann}) \\ \exists \text{hasChild.Person}(\text{Ann}) \end{array} \right. \\
\mathcal{T}' = \left\{ \begin{array}{l} \text{Father} \sqsubseteq \text{Parent} \sqcap \text{Man} \sqcap \neg \text{Female} \\ \text{Female} \sqsubseteq \text{Person} \\ \text{Woman} \sqsubseteq \text{Person} \\ \text{Man} \sqsubseteq \text{Person} \\ \text{Parent} \sqsubseteq \text{Person} \\ \text{Man} \equiv \neg \text{Woman} \\ \exists \text{hasChild.Person} \equiv \text{Parent} \\ \text{Crossdresser} \sqsubseteq \text{Female} \\ \text{Crossdresser} \sqsubseteq \neg \text{Woman} \\ \text{Mother} \sqsubseteq \text{Parent} \\ \text{Ab}_{\text{Woman}, \neg \text{Woman}} \sqsubseteq \text{Woman} \sqcap \neg \text{Woman} \end{array} \right. \\
\mathcal{A}' = \left\{ \begin{array}{l} \text{Female}(\text{Rupaul}), \text{Man}(\text{Rupaul}) \\ \text{Father}(\text{John}), \text{Man}(\text{John}) \\ \text{Parent}(\text{John}), \text{Woman}(\text{Ann}) \\ \exists \text{hasChild.Person}(\text{Ann}), \text{Female}(\text{Ann}) \\ \text{Woman}(\text{C}_{\text{Woman}, \neg \text{Woman}}) \\ \neg \text{Woman}(\text{C}_{\text{Woman}, \neg \text{Woman}}) \\ \text{N}_{\text{Man} \sqcap \text{Female}}(\text{Rupaul}) \\ \text{Ab}_{\text{Woman}, \neg \text{Woman}}(\text{C}_{\text{Woman}, \neg \text{Woman}}) \\ \text{Mother}(\text{Ann}) \\ \text{Crossdresser}(\text{RuPaul}) \end{array} \right.
\end{array}$$

Fig. 7. Ontology before (left) and after the spatial repairing (right)

Two concepts $C_1, C_2 \in \text{concepts}(\Sigma)$, will be said Σ -**connected** under σ , which will be denoted by $C_{\Sigma}^{\sigma}(C_1, C_2)$, if $C_{\Sigma}(\sigma(C_1), \sigma(C_2))$. The formula $\sigma(C_1)$ will be named the associate **notion** to C_1 in Σ .

The notion is defined for any concept. Nevertheless, in practice, this definition is not used intensively for atomic concepts (in the analysis of anomalies). It is that because the undefinition of the notions of an atomic concept can be deliberated: they are primitive concepts of the ontology (abstract concepts in many cases). Thus, it is not advisable to force to the user to refine them. The spatial idea of Σ -connection under σ is obviously that of the topological connection of the pulsation of sets. Now, $\sigma(C)$ represents a DL formula associate to a concept C . Notice that the new connection is related with the previous one:

$$C_{\Sigma}^{\sigma}(C_1, C_2) \iff \Sigma \not\models \sigma(C_1) \sqcap \sigma(C_2) \equiv \perp$$

Therefore, given two concepts C_1, C_2 in the ontology Σ and $R \in \text{RCC8}$,

$$R_{\Sigma}^{\sigma}(C_1, C_2) \iff R_{\Sigma}(\sigma(C_1), \sigma(C_2))$$

However, there is no cognitive reason to consider frontiers in vague regions, because the undefinition is represented by $\sigma(I(C)) \setminus I(C)$. Thus, RCC5 is more adequate in this case [12]. Starting with C_{Σ}^{σ} , it is possible both to classify *all* the relative positions between concepts/notions and, in due course, to repair them by using spatial reasoning preserving consistency [10].

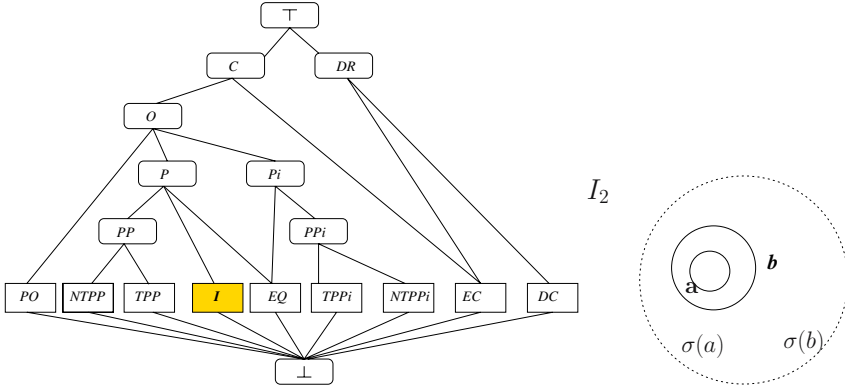


Fig. 8. A *robust* ontological extension of RCC by insertion of an uncertain relation (left) and its spatial interpretation (right)

6.1 Application: Advicement of Visual Repairs

The following principle is actually a working principle for refining specifications. Its soundness has to be accepted by the user in that application.

Fourth Cognitive Principle (CP4): If the RCC5-relation between spatial concepts does not agree with that of their corresponding notions, it may be necessary to adjust the spatial representation in order to ensure the agreement

To illustrate how CP4 can be applied in a simple case, let us consider

$$\Sigma = \begin{cases} \text{Omnivorous} \sqsubseteq \text{Carnivorous} \sqcap \text{Herbivorous} \\ \text{Carnivorous} \sqsubseteq \text{Animal}, \quad \text{Herbivorous} \sqsubseteq \text{Animal} \\ \text{Omnivorous}(\text{Bear}) \end{cases}$$

We deduce that $PO_{\Sigma}(\text{Carnivorous}, \text{Herbivorous}) \wedge EQ_{\Sigma}^{\sigma}(\text{Carnivorous}, \text{Herbivorous})$.

A simple way to solve this conflict, under the consideration of the user, relies on adding two constants b_1, b_2 to accurate the partial overlapping of notions (in the notion of a concept but not in the other one). In such a way, $PO_{\Sigma}(\mathbf{C}_1, \mathbf{C}_2)$ and $PO_{\Sigma'}^{\sigma}(\mathbf{C}_1, \mathbf{C}_2)$ holds. After this step, the new knowledge base is:

$$\Sigma' = \begin{cases} \text{Omnivorous} \sqsubseteq \text{Carnivorous} \sqcap \text{Herbivorous} \\ \text{Carnivorous} \sqsubseteq \text{Animal} \sqcap \neg\{b_2\}, \quad \text{Herbivorous} \sqsubseteq \text{Animal} \sqcap \neg\{b_1\} \\ \text{Herbivorous}(b_2), \text{Carnivorous}(b_1), \text{Omnivorous}(\text{Bear}) \end{cases}$$

Therefore, $PO_{\Sigma'}^{\sigma}(\text{Carnivorous}, \text{Herbivorous}) \wedge PO_{\Sigma}^{\sigma}(\text{Carnivorous}, \text{Herbivorous})$.

Graphically, the spatial adjustment consists in inserting two skolem constants in the spatial encodings of the notions (see Fig. 9, left). The interpretation of new constants b_1, b_2 (actually they are Skolem constants) is requested to the user. There are other more complex cases that imply more complex ontological revisions (figure 9 shows two examples based on individual insertion).

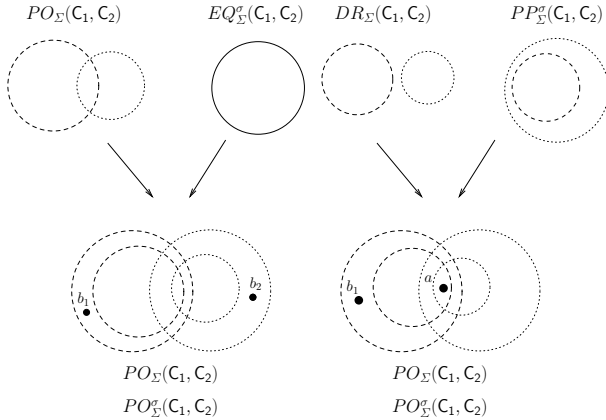


Fig. 9. Two examples of spatial arrangements based on CP4

6.2 Auxiliary Principles

Once CP4 is applied, an optional principle is useful for nonmonotonic reasoning on the visual encoding:

Fifth Cognitive Principle (CP5): Once CP4 is satisfied and the visual repairing is done, analyse whether is sound the following transformation on the final ontology: $C \sqsubseteq \sigma(C) \in \Sigma \mapsto C \equiv \sigma(C)$

CP5 is a completion-based principle that would allow to find a *definitional ontology*. That is, it can allow to transform the ontology source into an ontology which satisfies most of definitional principles given in [5]. However, the logical categoricity is weakened by some sort of spatial categoricity.

7 Closing Remarks

There exist a great number of methods for visual representation of ontologies, supporting a variety of tasks such as data analysis and queries ([14], [15]). However these works are mostly focused on visual representation and they lack both inference mechanisms and formal semantics. Unlike such visual encodings, RCC8 (RCC5) representations outlined here operate beyond just primarily mapping the ontology information/conceptualization structure. RCC8-based spatial encoding provides formal semantics where spatial arrangements mean ontology revision. Moreover, the encoding establishes a correspondence between the conceptualization implicit in the Ontology and a realm well known to the user. We described several spatial encodings based on different mereotopological interpretations of the ontologies. The spatial encoding are a sort of concept map [18] (because it identifies the interrelationships among concepts) enhanced with sound reasoning on the representation. Furthermore, we exploit logical features of RCC to analyse the impact of revision on the self ontology.

It is worth pointing out that it would be possible for that anomaly to come from other reasons different from the conceptualization. Future will be focused on mereotopological encodings of roles, in order to assist the user in the repair of anomalous arguments caused by roles. This phase is essential prior to the full implementation of a VOC-system. Finally, note that the visual encoding can be unmanageable for medium size ontologies. Our reparation method is argumentative (that is, it does not use representations of the whole ontology). However, if we want a whole representation, it would be interesting to adapt the spatial semantics to work with other visual encodings as the hyperbolic plane [20].

References

1. J. A. Alonso-Jiménez, J. Borrego-Díaz, A. M. Chávez-González and J.D. Navarro-Marín, A Methodology for the Computer-Aided Cleaning of Complex Knowledge Databases. *Proc. 28th Conf. of IEEE Industrial Electronics Soc. IECON 2002*, pp. 1806-1812, 2003.
2. J. 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, *16th Int. Conf. on Industrial & Eng. Appl. of Artificial Intelligence and Expert Systems IEA/AIE 2003*, LNAI 2718, Springer, 2003, pp. 789-798.
3. J.A. Alonso-Jiménez, J. Borrego-Díaz, A. M. Chávez-González, Ontology Cleaning by Mereotopological Reasoning. *DEXA Workshop on Web Semantics WEBS 2004*, pp. 132-137 (2004).
4. J. A. Alonso-Jiménez, J. Borrego-Díaz, A. M. Chávez-González and F. J. Martín-Mateos, Foundational Challenges in Automated Data and Ontology Cleaning in the Semantic Web, *IEEE Intelligent Systems*, 21(1):42-52 (2006).
5. B. Bennett, The Role of Definitions in Construction and Analysis of Formal Ontologies, *Logical Formalization of Commonsense Reasoning (2003 AAAI Spring Symposium)*, 27-35, AAAI Press, 2003.
6. J. Borrego-Díaz and A. M. Chávez-González, Extension of Ontologies Assisted by Automated Reasoning Systems, *10th Int. Conf. on Computer Aided Systems Theory, EUROCAST 2005*, LNCS 3643, pp. 247-253, Springer, 2005.
7. J. Borrego-Díaz and A. M. Chávez-González, Controlling Ontology Extension by Uncertain Concepts Through Cognitive Entropy, *Proc. of ISWC'05 Workshop on Uncertainty Reasoning for the Semantic Web*, pp. 56-66 (2005).
8. Calvanese, K., de Giacomo, G., Lenzerini, M., Nardi, D. Reasoning in Expressive Description Logics. In Alan Robinson y Andrei Voronkov (eds.) *Handbook of Automated Reasoning*, pp. 1581-1634 Elsevier Science Pub. (2001).
9. S. Card, J. Mckinlay and B. Shneiderman (eds.) *Readings in Information Visualization: Using Vision to Think*, Morgan Kaufman, 1999.
10. A. M. Chávez-González, *Automated Mereotopological Reasoning for Ontology Debugging* (spanish), Ph.D. Thesis, University of Seville, 2005.
11. C.C. Chang, H.J. Keisler, *Model Theory*, North Holland, 1977.
12. A.G. Cohn and N.M. Gotts, The 'Egg-Yolk' Representation of Regions with Indeterminate Boundaries in P. Burrough and A. M. Frank (eds), *Proc. GISDATA Specialist Meeting on Geographical Objects with Undetermined Boundaries*, GIS-DATA Series, vol. 3, Taylor and Francis, pp. 171-187 (1996).

13. A. G. Cohn, B. Bennett, J. M. Gooday and N. M. Gotts. Representing and Reasoning with Qualitative Spatial Relations about Regions. chapter 4 in O. Stock (ed.), *Spatial and Temporal Reasoning*, Kluwer, Dordrecht, 1997.
14. C. Fluit, M. Sabou, F. Harmelen. Ontology-based Information Visualization, in V. Geroimenko, C. Chen (eds.), *Visualizing the Semantic Web*. Springer (2003).
15. C. Fluit, M. Sabou, F. Harmelen. Supporting User Tasks Through Visualization of Light-weight Ontologies, in S. Staab and R. Studer (eds.), *Handbook on Ontologies in Information Systems*, Springer-Verlag (2003).
16. M. Grohe and L. Segoufin, On first-order topological queries, *ACM Trans. Comput. Log.* 3(3) (2002) pp. 336-358.
17. 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).
18. B. Gaines, M. Shaw, Concept Maps as Hypermedia Components, *Int. Journal of Human-Computer Studies* 43(3):323-361 (1995).
19. M. Knauff, R. Rauh and J. Renz, A Cognitive Assessment of Topological Spatial Relations: Results from an Empirical Investigation, *Proc. 3rd Int. Conf. on Spatial Inf. Theory (COSIT'97)*, LNCS 1329, 193-206, Springer-Verlag, Berlin (1997).
20. J. Lamping, R. Rao, and P. Pirolli. A focus + context technique based on hyperbolic geometry for visualizing large hierarchies, *ACM Conference on Human Factors in Computing Systems (CHI'95)* (1995).
21. Y. Mao, Z. Wu, H. Chen, X. Zheng, An Interactive Visual Model for Web Ontologies, *Proc. Knowl. Eng. Systems KES 2005*, LNAI 3862, pp. 866-872 (2005).
22. H. Mizen, C. Dolbear and G. Hart, Ontology Ontogeny: Understanding How an Ontology is Created and Developed, *Proc. First International Conference on Geospatial Semantics (GeoS 2005)*, LNCS 3799, pp. 15,29, 2005.
23. F. Murtagh, T. Taskaya, P. Contreras, J. Mothe and K. Englmeier, Interactive Visual Interfaces: A Survey, *Artificial Intelligence Review* 19:263-283, 2003.
24. C. Papadimitriou, D. Suciuc and V. Vianu, Topological Queries in Spatial Databases, *J. Computer System Sci.* 58 (1):29-53 (1999).
25. J. Renz. A Canonical Model of the Region Connection Calculus. *Proc. 6th Int. Conf. on Principles of Knowl. Rep. and Reasoning (KR'98)* (1998).
26. 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): 69-123 (1999).
27. D. Tsarkov, A. Riazanov, S. Bechhofer, and I. Horrocks, Using Vampire to Reason with OWL. *2004 Int. Semantic Web Conference (ISWC 2004)*, LNCS 3298 Springer, 2004, pp. 471-485.