Towards a Practical Argumentative Reasoning with Qualitative Spatial Databases*

José A. Alonso-Jiménez, Joaquín Borrego–Díaz,
Antonia M. Chávez-González, Miguel A. Gutiérrez-Naranjo and
Jorge D. Navarro-Marín

Dept. of Computer Science and Artificial Intelligence– University of Sevilla
{ jalonso, jborrego, tchavez magutier }@us.es

Abstract. Classical database management can be flawed if the Knowledge database is built within a complex Knowledge Domain. We must then deal with inconsistencies and, in general, with anomalies of several types. In this paper we study computational and cognitive problems in dealing qualitative spatial databases.

1 Introduction

Spatio-temporal representation and reasoning are topics that have attracted quite a lot of interest in AI. Since the spatial notions used by the humans are intrinsically qualitative, the reasoning about spatial entities, their properties and the relationship among them, are central aspects in several intelligent systems. But the problem is far to be solved in general. The spatial reasoning is more complex than the temporal one. The higher dimension of the things is not the unique problem. The topology is, in qualitative terms, hard to represent by formalisms with amenable calculus. The semantic of these representations offers incomplete support to our daily reasoning (the *poverty conjecture*: there is no purely qualitative, general purpose kinematics). Different ontologies have been proposed, but these are not of general purpose.

Among them, the theory called *Region Connection Calculus* (hereafter referred as RCC), developed by Randell, Cui and Cohn [3] have been extensively studied in AI [10], and in the field of Geographic Information Systems [1]. A common deficiency of the theories representing topological knowledge, is that either the full theory is computationally unacceptable or they fail to meet basic desiderata for these logics [8]. For constraint satisfaction problems there are algorithms to work with the relational sublanguage, and tractable subsets of the calculus RCC-8 (a relational sublogic of RCC) have been found [10]. The intractability of the full theory is mainly due to the complexity of its models (topological spaces with separation properties [6]). We propose a practical approach (using an automated theorem prover) to investigate the verification

* Work partially supported by the MCyT project TIC 2000-1368-C03-0 and the project TIC-137 of the *Plan Andaluz de Investigación*

problem of knowledge bases written in the language of RCC. In general, due to the complexity of the theories involved, the knowledge base may be inconsistent, although the environment was *well represented* by the relational part of the database. For a proper understanding of our framework, it is important to point out the following characteristics of the problem:

- The knowledge database has not ever completed (the user will write new facts in the future). Thus the difficulties begin with the future introduction of data.
- The *intensional theory* of the database is not clausal. Thus, is highly possible that Reiter's axiomatization of database theory [11] becomes inconsistent. Nevertheless, the self database represents a real spatial configuration.
- The knowledge base does not contain facts about all the relationships of the RCC language. It seems natural that only facts on primary relationships appear (we have selected for our work the relations `Connect`, `Overlaps` and `Part-of`, that one can consider as the *primary* relationships).

The above characteristics are important in order to classify the anomalies. The first one may produce inconsistencies that the user can repair, but the second one is a logical inconsistency and it is hard to solve. Thus, we have to reason with inconsistent knowledge. The last characteristic implies that the (logic-based) deduction of new knowledge must replace to solving methods for CSP.

Our problem is only an interesting example of the more general problem of cleaning incomplete databases in the Semantic Web: the *cleaning agent* must detect anomalies in knowledge bases written by the user (in structured text), and associated to a complex ontology. It is necessary to point out that it is not our aim to find inconsistencies in the domain knowledge. In [5] the authors it is shown an application of an automated theorem prover (the SNARK system) to provide a declarative semantics for languages for the Semantic Web, by translating first the forms from the semantic markup languages to first-order logic. The translation allows to apply the theorem prover to find inconsistencies. Our problem is not exactly that. We assume that the domain knowledge (the RCC theory and eventually the composition table for the relations of figure 3) is consistent, and that it is highly possible that RCC jointly with the database becomes inconsistent. However, in one of the experiments the theorem prover found an error in the composition table for the RCC-8 of [3]. Our problem has also another interesting aspect: the data inserted have not any spatial indexing.

2 The theory of RCC

The Region Connection Calculus is a topological approach to qualitative spatial representation and reasoning where the *spatial entities* are non-empty regular sets[1] (a good introduction to the theory is [3]). The primary relation between such regions is the connection relation $\mathtt{C}(\mathtt{x},\mathtt{y})$, which is interpreted

[1] A set x of a topological space is regular if it agrees with the interior of its closure.

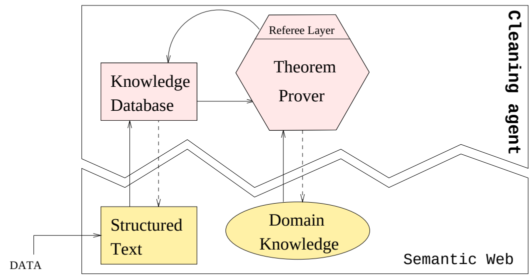


Fig. 1. Cleaning service process

as *"the closures of* $\mathtt{x}$ *and* $\mathtt{y}$ *intersect"*. The axioms of RCC are two basic axioms on $\mathtt{C}$, $\mathtt{A}_1 := \forall \mathtt{x}[\mathtt{C}(\mathtt{x},\mathtt{x})]$ and $\mathtt{A}_2 := \forall \mathtt{x},\mathtt{y}[\mathtt{C}(\mathtt{x},\mathtt{y}) \rightarrow \mathtt{C}(\mathtt{y},\mathtt{x})]$, plus several axioms/definitions on the main spatial relationships (see fig. 2).

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

Fig. 2. Axioms of RCC

The set of eight jointly exhaustive and pairwise disjoint relations shown in fig. 3 form the relational calculus RCC-8, that has been deeply studied by J. Renz and B. Nebel [10]. In that work the CSP problems associated to RCC-8 are classified in terms of (un)tractability. These problems are, in some cases, tractable, but the relational language can be too weak for some particular applications. The consistency/entailment problems in the full theory RCC have a complex behaviour. If we consider topological models, the problem is computationally unacceptable. The restriction of the problems to *nice* regions of $\mathbb{R}^2$ is also hard to compute [7].

The problem of a good representation of a model by a knowledge base arises. Concretely, we must consider three classes of models: the class of all models (according to the classical definition from first order logic), the class of the topological models, and the topological model $\mathbb{R}^n$ where the constants are interpreted as the regular sets under study (*the intended model*). Formally,

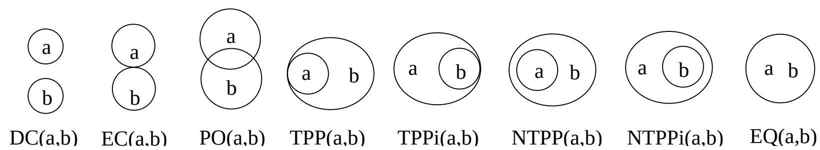


Fig. 3. The eight basic relations of RCC-8

Definition 1. *Let Ω be a topological space, and* $\mathtt{X}$ *be a finite set of constants. A structure Θ is called a* topological model *on Ω if it has the form*

$$\langle \mathcal{R}(\Omega)_{/\sim}, \mathtt{C}_\Theta, \{\mathtt{a}_\Theta \ : \mathtt{a} \in \mathtt{X}\}\rangle$$

where $\mathcal{R}(\Omega)$ is the class of regular sets, $\sim$ is the equivalence relation "the closures agree"[2], *$\mathtt{C}_\Theta$ is the intended interpretation of* $\mathtt{C}$ *and whenever $\mathtt{a} \in \mathtt{X}$, $\mathtt{a}_\Theta \in \mathcal{R}(\Omega)_{/\sim}$.*

Every structure is expanded to one in the full language of RCC, by the natural interpretation of the other relationships.

Theorem 1. *[6] If Ω is a nontrivial connected T_3-space, then the natural expansion of any topological structure on Ω to the full language is a model of* RCC.

3 Towards an automated argumentative reasoning

The logic-based argument theory is a formalism to reason with inconsistent knowledge [4]. An *argument* is a pair $\langle \Pi, \varphi\rangle$ where Π is consistent and $\Pi \vdash \varphi$. The argumentative structure of $\mathtt{K}$ is an hierarchy of arguments which offers a method to obtain useful knowledge from $\mathtt{K}$ with certain properties. Also, it provides a method to evaluate the robustness of an argument via *argument trees* [2]. However, this approach can not be directly applied on huge databases, since it needs, for example, to find all maximally consistent subsets of the database. The problem can be solved in practice by adapting the notion of *argument* to an automated theorem prover. For this work we choose OTTER [9], a resolution-based automated theorem prover, as inference system. Since it is not our aim to describe the methodology we need to work with the theorem prover, we simply assume that the system works in autonomous mode, a powerful feature of OTTER.

Definition 2.

1. *An* $\mathtt{O}$*-argument (an argument for OTTER) is a pair $\langle \Pi, \varphi\rangle$ such that OTTER obtains a refutation of $\Pi \cup \{\neg\varphi\}$ (that we write $\Pi \vdash^{\mathtt{O}} \varphi$).*
2. *If $\langle \Pi, \phi\rangle$ is an* $\mathtt{O}$*-argument, the length of $\langle \Pi, \phi\rangle$, denoted as $len(\langle \Pi, \phi\rangle)$, is the length of the refutation of $\Pi \cup \{\neg\varphi\}$ by OTTER.*

[2] The relation is necessary because of the extensionality of $\mathtt{P}$ given in the axiom $\mathtt{A_P}$. With this relation, the mereological relation $\mathtt{EQ}$ agree with the equality.

The argumentative structure can not be directly translated, because of the consistency notion. An associated automated model finder, MACE, may be considered for a complete description. By example, the argument class $\mathtt{A}\exists(\mathtt{K})$ may be adapted: $\mathtt{A}^{\mathtt{O}}\exists(\mathtt{K}) = \{\langle \Pi, \varphi \rangle \ : \ \Pi$ is consistent w.r.t. MACE and $\Pi \vdash^{\mathtt{O}} \varphi\}$.

4 Anomalies in complex knowldege bases

From now on, we will consider fix a topological model Θ, the spatial model whith we will work, and $\mathtt{K}$ a database representation of Θ (that is, $\Theta \models \mathtt{K}$). To simplify we assume that the model satisfies the unique names axiom. Three theories describe the model: the formalization of Reiter's database theory $\mathtt{T_{DB}}(\mathtt{K})$, the theory $\mathtt{RCC}(\mathtt{K})$, whose axioms are those of $\mathtt{K}$ plus RCC, and $\mathtt{RCC}(\mathtt{T_{DB}}(\mathtt{K}))$. The three theories has a common language, $\mathtt{L_K}$. The following is an intuitive ontology of the anomalies in RCC-databases suggested by the experiments:

A1: The contradictions of the base due to the bad implementation of the data (e.g. absence of some knowledge)
A2: The anomalies due to the inconsistency of the model: the theorem prover derives from the database the existence of regions which do not have a name (possibly because they have not been introduced by the user yet). This anomaly may also be due to the *Skolem's noise*, produced when we work with the domain closure axioms but the domain knowledge is not clausal.
A3: Disjuntive answers (a logical deficiency).
A4: Inconsistency in the Knowledge Domain.

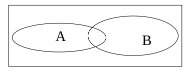
As we remarked, the anomalies come from several sources: the set may be inconsistent with the Domain Knowledge due to formal inconsistencies produced by wrong data, the database is not complete with respect to a basic predicate (the user will continue introducing data), etc. In fig. 4 the most common problem is shown. The system shows arguments with the Skolem function of the clausal form of $\mathtt{A_O}$ to questions as *"gives us a region which overlaps* $\mathtt{a}$*"*.

The Skolem functions come from four axioms of RCC, $\mathtt{A_P}, \mathtt{A_O}, \mathtt{A_{TPP}}$ and $\mathtt{A_{NTPP}}$, when they are clausified. It is possible to give a spatial interpretation of such functions. Even if THE spatial regions are semialgebraic sets, the Skolem functions can be semialgebraically defined [12]. In the practice, the spatial interpretation may be thought as a partial function. For example, the Skolem function for $\mathtt{A_O}$, $\mathtt{f_O}(\mathtt{x},\mathtt{y})$ gives the intersection region of x and y, if $\mathtt{O}(\mathtt{x},\mathtt{y})$. This idea allow us to eliminate useless results of type (A2) extending the theory RCC with a partial axiomatization of the intersection (see the first three axioms of figure 6).

5 Consistent databases and arguments

Definition 3. *Let Θ be a topological model. The* graph *of Θ, Θ_G, is the substructure of Θ whose elements are the interpretation of the constants.*

Definition 4. *Let* $\mathtt{K}$ *be a set of formulas.*




```
Database                                                    an OTTER's proof
------------------------------------------------------------------------
all x (x=A|x=B).                                             1 [] x=x.
A!=B.                                                        5 [] x!=A|y!=B|O(x,y).
all x y (x=A&y=B|x=B&y=A|x=A&y=A|x=B&y=B->O(x,y)).  14 [] P($f1(x,y),x)|
                                                                  -O(x,y).
all x y (x=A& y=B|x=B& y=A|x=A&y=A|x=B&y=B->C(x,y)).16 [] -P(x,A)|$Ans(x).
all x y ((exists z (P(z,x)&P(z,y)))<->O(x,y)).       27 [hyper,5,1,1] O(A,B).
all x (P(x,A)->$Ans(x)).                             65 [hyper,14,27]
                                                        P($f1(A,B),A).
                                                     66[binary,65.1,16.1]
                                                       $Ans($f1(A,B)).
```

Fig. 4. A simple anomaly and an O-argument

1. *The* world *of* K, W(K), *is the set of the interpretations in* Θ *of the constants in the language of* L_K.
2. *Consider an interpretation of the Skolem functions of the clausal form of* RCC. *The* cognitive neighborhood *of* K, $\Gamma(K)$, *is the least substructure of the expansion of* Θ *to the clausal language of* RCC, *containing* W(K).

It seems that the consistency of an argument depends only on its cognitive neighborhood. It is true for arguments with enough *credibility*.

Definition 5. *An* undercut *of* $\langle \Pi_1, \phi\rangle$ *is an argument* $\langle \Pi, \neg(\phi_1 \wedge \cdots \wedge \phi_n)\rangle$ *where* $\{\phi_1, \ldots, \phi_n\} \subseteq \Pi_1$. *The undercut is called* local *if* $\Gamma(\Pi) \subseteq \Gamma(\Pi_1)$

Definition 6. *An argument* $\langle \Pi, \alpha\rangle$ *is more conservative than* $\langle \Pi', \beta\rangle$ *if* $\Pi \subseteq \Pi'$ *and* $\beta \vdash^{\mathtt{O}} \alpha$.

Definition 7. *Let* T *be a set of formulas, and let* ϕ *be a formula of the clausal language of* T.

- *A clause has* Skolem's noise *if it has occurrences of Skolem function symbols.*
- *The* degree of credibility *of an argument* $\langle \Pi, \phi\rangle$ *is*

$$gr(\langle \Pi, \phi\rangle) = \frac{len(\langle \Pi, \phi\rangle) - |\{\eta \in Proof^{\mathtt{O}}(\Pi, \phi) \ : \ \eta \text{ has Skolem's noise}\}|}{len(\langle \Pi, \phi\rangle)}$$

The degree of credibility estimates the robustness of the argument according to the use of Skolem functions in the proof, functions which become *ghost* regions (the credibility degree of the argument shown in fig. 4 is 4/7).

Theorem 2. *If* $gr(\langle \Pi, \phi\rangle) = 1$ *then* $\langle \Pi, \phi\rangle \in \mathtt{A}\exists(\mathtt{RCC}(\mathtt{T}_{\mathtt{DB}}(\mathtt{K})))$ *and* $\Gamma(\Pi) \models \Pi + \phi$.

Corollary 1. *If $gr(\langle \Pi, \phi \rangle) = gr(\langle \Pi', \phi' \rangle) = 1$ and the first argument is an undercutting of the second one, then*

- $\Gamma(\Pi) \not\subseteq \Gamma(\Pi')$ *(there is not local undercutting with degree of credibility 1).*
- *If $\langle \Pi, \phi \rangle$ is a maximal conservative undercutting, then $\Gamma(\Pi') \subsetneq \Gamma(\Pi)$.*

The above corollary relates undercutting arguments and spatial configurations, and it may be useful to estimates the size of *argument trees* [2].

Definition 8. *Let* $\mathtt{K}$ *be a knowledge database for* Θ. *The base* $\mathtt{K}$

- *is* $\mathtt{C}$-complete *if whenever* $\mathtt{a}, \mathtt{b} \in \mathtt{L_K}$, *if* $\Theta \models \mathtt{C}(\mathtt{a}, \mathtt{b})$, *then* $\mathtt{C}(\mathtt{a}, \mathtt{b}) \in \mathtt{K}$.
- *is* extensional *for* $\mathtt{P}$ *if whenever* $\mathtt{a}, \mathtt{b} \in \mathtt{L_K}$

$$\mathtt{P}(\mathtt{a}, \mathtt{b}) \notin \mathtt{K} \quad \Longrightarrow \quad \exists \mathtt{c} \in \mathtt{L_K} \, [\mathtt{C}(\mathtt{c}, \mathtt{a}) \in \mathtt{K} \wedge \mathtt{C}(\mathtt{c}, \mathtt{b}) \notin \mathtt{K}]$$

- *is* refined *if whenever* $\mathtt{a}, \mathtt{b} \in \mathtt{L_K}$

$$\mathtt{O}(\mathtt{a}, \mathtt{b}) \in \mathtt{K} \quad \Longrightarrow \quad \exists \mathtt{c} \in \mathtt{L_K} \, [\{\mathtt{P}(\mathtt{c}, \mathtt{a}), \mathtt{P}(\mathtt{c}, \mathtt{b})\} \subseteq \mathtt{K}]$$

- recognize frontiers *if whenever* $\mathtt{a}, \mathtt{b} \in \mathtt{L_K}$ *such that* $\Theta \models \mathtt{P}(\mathtt{a}, \mathtt{b})$

$$\Theta \models \mathtt{TPP}(\mathtt{a}, \mathtt{b}) \Longleftrightarrow \exists \mathtt{c} \in \mathtt{L_K} \, [\{\mathtt{C}(\mathtt{c}, \mathtt{a}), \mathtt{C}(\mathtt{c}, \mathtt{b})\} \subseteq \mathtt{K} \wedge \{\mathtt{O}(\mathtt{c}, \mathtt{a}), \mathtt{O}(\mathtt{c}, \mathtt{b})\} \cap \mathtt{K} = \emptyset]$$

Theorem 3. *If* $\mathtt{K}$ *has the above four properties, then* $\Theta_{\mathtt{G}} \upharpoonright_{\mathtt{W(K)}} \models \mathtt{RCC}(\mathtt{T_{DB}}(\mathtt{K}))$.

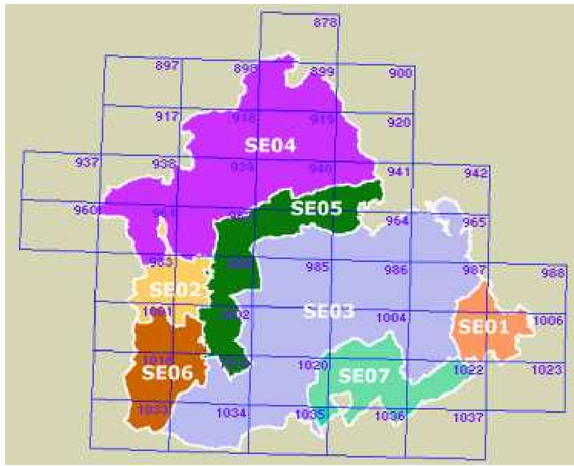
An useful parameter of the graph of Θ is the *compactness level.*

Definition 9. *The* compactness level *of Θ is the least $n > 0$ such that the intersection of any set of regions of Θ_G is equal to the intersection of n regions of the set.*

In general, a database is not refined. Notice that when a database $\mathtt{K}$ is refined, the Skolem function $\mathtt{f_O}$ is interpretable within $\mathtt{W(K)}$. Thus, we can add to the database a set of axioms with basic properties of such function, and $\mathtt{f_O}$ can be *syntactly defined* in K, and simplified by compactness level, if possible. If it is not refined, the partial definition is also useful (see fig. 7).

6 Experiments

We now report experiments with a spatial database on the relationships among three types of regions: counties, districts, and available maps on Andalucía, a Spanish autonomous region. The system works on a database built with the relationships of connection (`Connect`), nonempty-intersection (common subregion, `Overlaps`), and part-of (`Part-of`). Thus there exists hidden information, knowledge with respect to other topological relations among regions, not explicit in the database, that the theorem prover might derive (and, eventually, add to the database). The graph of Θ is formed by 260 regions, approximately, for which we have a database with 34000 facts (included the first-order formalization of



$$\begin{aligned}\mathtt{K} = \{&\langle \mathtt{Connect:\ SE04,\ SE05}\rangle, \langle \mathtt{Connect:\ SE04,\ Map-941}\rangle,\\ &\langle \mathtt{Overlaps:\ Map-920,\ SE04}\rangle, \langle \mathtt{Part-of:\ SE04,\ SEVILLA}\rangle \cdots\}\end{aligned}$$

Fig. 5. Partial view of the autonomous region and some facts from the database

databases, but the number can be reduced using some features of the theorem prover). This database has been made by hand, and possibly, it contains errors. The processed database has 40242 clauses (the processing takes 6.5 seconds). It has been used OTTER 3.2 on a computer with two Pentium III (800 Mhz) processors and 256 Mb RAM. The machine runs with Red Hat Linux operating system 7.0.

The database is C-complete but it is not refined. Thus, it is highly possible that the theorem prover detects anomalies of type (A2). It recognize accidentally the frontiers, and its compactness level is 2. The compactness level can be axiomatized and incorporated to the theory if we use the (partial) spatial interpretation of the Skolem function as *partial intersection* (see figure 6). This option allows us to obtain more acceptable results, reducing the anomalies (A2) and obtaining more arguments (see fig. 7)

$$\begin{array}{c}\mathtt{Int}(x,x) = x \qquad\qquad \mathtt{P}(x,y) \rightarrow \mathtt{Int}(x,y) = x\\ \mathtt{O}(x,y) \rightarrow \mathtt{Int}(x,y) = \mathtt{Int}(y,x)\\ \mathtt{O}(y,z) \wedge \mathtt{O}(x,\mathtt{Int}(y,z)) \rightarrow \mathtt{Int}(x,\mathtt{Int}(y,z)) = \mathtt{Int}(\mathtt{Int}(x,y),z)\\ \mathtt{O}(y,z) \wedge \mathtt{O}(x,\mathtt{Int}(y,z)) \rightarrow \left\{\begin{array}{l}\mathtt{Int}(x,\mathtt{Int}(y,z)) = \mathtt{Int}(y,z) \vee\\ \mathtt{Int}(x,\mathtt{Int}(y,z)) = \mathtt{Int}(x,y) \vee\\ \mathtt{Int}(x,\mathtt{Int}(y,z)) = \mathtt{Int}(x,z)\end{array}\right\}\end{array}$$

Fig. 6. An axiomatization of $\mathtt{f}_0$ (as Int) when the compactness level is 2

`P(x,Jaen) -> $Ans(x)`

Exp.	CPU time (sec.)	generated clauses	results	(A1)	(A2)	(A3)
(R1)	54.21	175	1	0	0	0
$(R1)^+$	55,20	180	1	0	0	0
(R2)	59	671	25	102	1	0
$(R2)^+$	60,26	677	25	0	2	0
(R3)	316	19,812	232	0	5	1
$(R3)^+$	320	31,855	287	0	5	1
(R4)	54.79	570	1	0	1	0
$(R4)^+$	55.6	575	1	0	1	0

Fig. 7. Experiment on the behaviour of OTTER on a question without and with ($^+$) axiomatization of the compactness level

`PP(x, Huelva) -> $Ans(x)`

Exp.	CPU time (sec.)	generated clauses	results	(A1)	(A2)	(A3)	(A4)
(R1)	2395.31	195,222	1	113	0	0	0
(R2)	2400	201,797	8	113	0	0	0
(R3)	2514.46	287,088	14	117	0	1	0
(R4)	54.15	286	0	1	0	0	0

Fig. 8. Statistics for a complex question

We selected the predicates `Part-of, Proper-part, Externally-connect` as targets of the experiments. Several results are in figures, 7, 8 and 9. (R1) shows the first correct answer to the question, (R2) shows the results 5 seconds later, (R3) shows the first useless result and (R4) shows statistics for the first error found.

It is not our aim to use the theorem prover as a simple database programming language. The idea is to ask to the system complex questions which are unsolvable by constraint satisfaction algorithms or simple SQL commands. The questions are driven to obtain knowledge on spatial relationships not explicit in the database (as `Proper-part` or boolean combination of complex spatial relations). Some of the questions require an excessive CPU time. Surprisingly, the time cost is justified: the theorem prover *thought* all the time on the database and it found many errors of the type (A1), errors which are not acceptable. The number of useless arguments of type (A2) obtained can be significatively reduced by the spatial interpretation of $\mathtt{f}_0$ function (see fig. 7).

As we remarked earlier, the theorem prover found an error in the composition table of RCC (type (A4)) working on a complex question (see fig. 9).

7 Conclusions and future work

In this paper we have focused on practical paraconsistent reasoning with qualitative spatial databases using logic-based argumentative reasoning. The problem

`EC(x, Sevilla) -> $Ans(x):`

CPU time (sec.)	generated clauses	results	(A1)	(A2)	(A3)	(A4)
3845	11,673,078	25	113	0	6	72

Fig. 9. Statistics of an experiment when the composition table of [3] produces errors

is an example of *cleaning* databases within complex domain knowledge, which is a promising field of applications in the Semantic Web. A spatial meaning of the arguments has been shown. The next challenge is to model the robustness of an argument estimating the *number of arguments for or against* a particular sequent by topological parameters on the graph of the model. The estimation will be useful when we work with very large spatial information.

References

1. B. Bennett. The application of Qualitative Spatial Reasoning to GIS. In R.J: Abrahart (ed.) Proceedings of the First International Conference on GeoComputation, 44-47, Leeds, (1996).
2. P. Besnard, A. Hunter: A logic-based theory of deductive arguments. Artificial Intelligence 128(1-2):203-235 (2001)
3. 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. Kluwer (1997).
4. M. Elvang-Goransson and A. Hunter. Argumentative logics: Reasoning from classically inconsistent information. Data and Knowledge Engineering 16:125-145 (1995).
5. R. Fikes, D. L. McGuinness, and R. Waldinger. A First-Order Logic Semantics for Semantic Web Markup Languages. Tech. Rep. n. KSL-01-01. Knowledge Systems Laboratory, Stanford University (2002).
6. N. Gotts. An Axiomatic Approach to Topology for Spatial Information Systems. Tech. Rep. n. 96.24. School of Computer Studies. University of Leeds (1996).
7. M. Grigni, D. Papadias, and C. Papadimitriou. Topological inference. In C. Mellish (ed.) Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI), (1):901-906. Morgan Kaufmann (1995).
8. O. Lemon and I. Pratt: Complete Logics for Qualitative Spatial Reasoning. Journal of Visual Languages and Computing 9:5-21 (1998).
9. W. McCune. OTTER's user manual. Argonne National Laboratory (1994). `http://www-unix.mcs.anl.gov/AR/otter/`
10. 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).
11. R. Reiter. Towards a Logical Reconstruction of Relational Database Theory. in M.L. Brodie, J.L. Mylopoulos and J.W. Schmidt (eds.) On Conceptual Modelling, pp. 191-233. Springer, N.Y. (1982).
12. L. P. D. van den Dries. Tame Topology and O-minimal Structures. Cambridge U. Press (1998).