

# Combining Topological and Size Information for Spatial Reasoning

Alfonso Gerevini<sup>a</sup> Jochen Renz<sup>b</sup>

<sup>a</sup>*Dipartimento di Elettronica per l'Automazione, Università di Brescia,  
via Branze 38, I-25123 Brescia, Italy. Email: gerevini@ing.unibs.it*

<sup>b</sup>*Institut für Informationssysteme Technische Universität Wien,  
A-1040 Wien, Austria. Email: renz@dbai.tuwien.ac.at*

---

## Abstract

Information about the size of spatial regions is often easily accessible and, when combined with other types of spatial information, it can be practically very useful. In this paper we introduce four classes of qualitative and metric size constraints, and we study their integration with the Region Connection Calculus RCC-8, a well-known approach to qualitative spatial reasoning with topological relations.

We propose a new path-consistency algorithm for combining RCC-8 relations and qualitative size relations. The algorithm is complete for deciding satisfiability of an input set of topological constraints over one of the three maximal tractable subclasses of RCC-8 containing all the basic relations. Moreover, its time complexity is cubic and is the same as the complexity of the best known method for deciding satisfiability when only these topological relations are considered. We also provide results on finding a consistent scenario in cubic time for these combined classes.

Regarding metric size constraints, we first study their combination with RCC-8 and we show that deciding satisfiability for the combined sets of constraints is NP-hard, even when only the RCC-8 basic relations are used. Then we introduce RCC-7, a subalgebra of RCC-8 that can be used for applications where spatial regions cannot partially overlap. We show that reasoning with the seven RCC-7 basic relations and the universal relation is intractable, but that reasoning with the RCC-7 basic relations combined with metric size information is tractable. Finally, we give a polynomial algorithm for the latter case and a backtracking algorithm for the general case.

---

## 1 Introduction

Representing and reasoning about spatial information is important in several areas of AI and computer science, such as spatial information systems [7,22]

robot navigation [23,46], natural language processing [46], visual languages [21,33], qualitative simulation of physical processes [16,40,53] and common-sense reasoning [17].

Previous work in spatial reasoning has addressed various aspects of space, such as topology [4,9,10,12,20,54], direction [45], shape [24], size [52,65], distance or position [13]. However, most research on qualitative spatial reasoning has focussed on single aspects of space, while real world applications usually require to deal with more than just one spatial aspect. Representing and reasoning about, e.g., topological information only is often insufficient. Since different aspects are not independent from each other, an approach that treats each single aspect independently is not feasible. Research effort should therefore be directed towards integrating different aspects of space. In this paper we make a first step toward this direction: we introduce some types of qualitative and quantitative size information, and we study their integration with topological information.

Our work is based on RCC-8 [54], a well-known constraint language for topological spatial reasoning that is based on the Region Connection Calculus [5,6,14,19,35,36,55,57,59,64]. In this framework, regions are independent with respect to rotation, translation, and several other transformations of the underlying space which makes them very simple and natural. This has also been observed in cognitive evaluations [39,58]. The topological distinctions made by RCC-8 are essentially the same as those made by Egenhofer in his 9-intersection model [20]. There is a number of applications which involves these distinctions. For instance, they are used in state-of-the-art geographical information systems such as Oracle Spatial or Intergraph products, in spatial databases [50], for image retrieval [2], for visual languages [21,33], and in description logics for knowledge representation [35].

Information about the size of spatial regions is often easily accessible, it is commonly used in natural language descriptions of spatial configurations, and when combined with other types of spatial knowledge it can be very useful. In this paper we study the combination of topological information expressible by RCC-8 relations with qualitative and metric information about the size of spatial regions. Given a set of topological and size constraints on the spatial regions associated with the objects in the domain under consideration, a fundamental reasoning task consists of determining whether this information is consistent (i.e., the set of constraints is satisfiable). A reasoning algorithm for this task should take into account that neither qualitative nor quantitative size information is independent from topological information. As a very simple example, suppose we have three geographical regions A, B and C for which the only topological information available is that B is contained in A. In addition we know that A is smaller than C, and that C is smaller than B. The combined set of topological and relative size information is inconsistent, but we cannot

detect this by just independently processing the two kind of information, or by just expressing the size information as topological constraints.<sup>1</sup>

In general, topological information interacts with size information by imposing “containment constraints” stating that the size of a region  $x$  contained in another region  $y$  must be smaller than the size of  $y$ , and that the size of the union of all the regions contained in  $y$  must not exceed the size of  $y$ . Similarly, if we know that the size of a certain region is larger than the size of another one, we can infer that the first cannot be contained in or equal to the second.

Topological and size information could be independently treated by known reasoning algorithms, but, since topological and size information clearly can interact leading to inconsistencies when the separate sources of information are consistent, such reasoning would be inherently incomplete. As our results show, a correct and complete treatment of the information in the combined framework can be much harder than in the separate frameworks, depending on which topological relations are, and on whether spatial information is qualitative or quantitative.

RCC-8 is formed by a set of eight relations, called basic relations, and by all the possible unions of them. In general, deciding consistency (satisfiability) of a set of constraints in RCC-8 is NP-complete, but for three large subsets of RCC-8, called  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  and  $\mathcal{Q}_8$ , this problem can be solved in polynomial time [57,56]. These classes are the only maximal tractable subclasses of RCC-8 that contain all the basic relations. In the first part of the paper we consider a set of qualitative relations between region sizes forming a Point Algebra [42], which have been thoroughly studied in the context of temporal reasoning (e.g., [30,31,61–63]). We propose an algorithm, called BIPATH-CONSISTENCY, to combine RCC-8 and qualitative size constraints, and we study its properties when the input topological constraints are in either  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  or  $\mathcal{Q}_8$ . We prove that, even though our extended framework is more expressive than either  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  or  $\mathcal{Q}_8$  (and therefore it has a larger potential applicability), the problem of deciding consistency can be solved using our algorithm in cubic time (i.e., without additional worst-case cost). Moreover, this algorithm can be exploited to solve in cubic time the problem of computing a *consistent scenario* for a set constraints over either  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  or  $\mathcal{Q}_8$ , i.e., to compute a consistent refinement of each constraint in the set to one of its basic relations.

In the second part of the paper we study the combination of RCC-8 with metric information about the size of the spatial regions. We introduce three classes of metric spatial constraints. The first class is formed by metric constraints on the relative size of two regions (e.g., “the size of region  $x$  is at least 2.5 times the size of region  $y$ ”); the second class is formed by difference constraints

---

<sup>1</sup> Another more complex example illustrating interdependencies between topological and qualitative size constraints is given in Section 4.

between region sizes (e.g., “the size of  $y$  minus the size of  $x$  is less than or equal to 20 spatial units”); the third class is formed by constraints on the domain of the size of a region (e.g., “the size of region  $x$  is at least 25 spatial units and no more than 35”). Size difference constraints are analogous to the temporal distance constraints forming a *Simple Temporal Constraint Problem* or STP [18].

We show that, although these classes of constraints are independently tractable, deciding consistency becomes NP-hard when they are combined even with only the set of the basic RCC-8 relations. We then consider combinations of metric size relations with subsets of the RCC-8 basic relations, obtaining some tractable fragments. From a practical point of view, the most interesting of them is the set all the basic RCC-8 relations except the relation of partial overlapping. These seven basic relations can be disjunctively combined obtaining a subalgebra of RCC-8 that we call RCC-7.

RCC-7 is suited for applications dealing with spatial regions which cannot partially overlap. Obvious examples of this kind of regions are regions that correspond to solid-state physical objects, or geographic regions such as countries or administrative districts (this observation is made also in [34] and [11], where similar sets of topological relations are studied). For example, consider representing a hierarchy of geographical spatial information like topological constraints between regions of the same state, between a region and its state, between states of the same country, and between a state and its country. We have that no pair of regions, states, or countries can partially overlap with each other, while they can be disconnected, externally connected, or they can be a part of another spatial region (a certain state or country). A more detailed illustrative example will be given in Section 7.

We show that deciding consistency for a set of basic RCC-7 relations and metric size constraints is tractable, and we give a polynomial algorithm for solving this problem. We also investigate the complexity of deciding consistency for larger sets of RCC-7 relations. It turns out that the simple addition of the universal relation of RCC-7 to the basic relations of RCC-7 makes the consistency problem intractable. The universal relation is a practically interesting relation, since it is necessary for expressing the lack of topological information between regions.<sup>2</sup> Finally, we give a backtracking algorithm for dealing with the full class of RCC-7 relations and metric size information. This algorithm can also be used as an approximate algorithm for RCC-8 relations and metric size constraints.

---

<sup>2</sup> For instance, if we know that A is disconnected from B and that B is disconnected from C, but we do not know how A and C are topologically related (there is no constraint between them), then in the representation we should state that the relation between A and C is the universal relation.

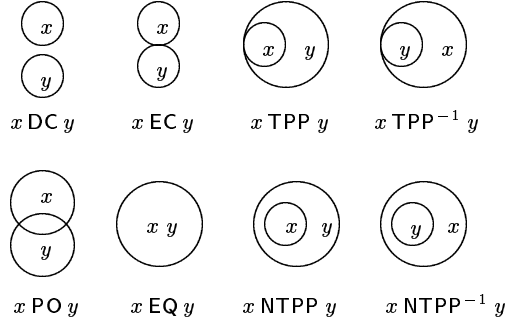


Fig. 1. Two-dimensional examples for the eight basic relations of RCC-8

The paper is organized as follows. Section 2 gives the necessary background on RCC-8, introduces RCC-7, and gives some basic results. Sections 3 and 4 deal with the combination of topological and qualitative size constraints; Sections 5 and 6 give our results regarding the combination of RCC-8 and RCC-7 with metric size constraints; Section 7 gives a worked example, and Section 8 our conclusions.

## 2 Topological Constraint Languages

In this section we present two relation algebras for topological constraint reasoning which are definable in the Region Connection Calculus [54], where spatial regions are non-empty regular subsets of a topological space, and can consist of more than one piece. First we briefly describe RCC-8, for which we summarize some known results and prove a technical lemma that will be used in our proofs. Then we introduce a new sub-algebra of RCC-8 which is important for applications where spatial regions cannot partially overlap. We call this algebra RCC-7.

### 2.1 The Region Connection Calculus RCC-8

RCC-8 [54] is a set of binary spatial relations formed by eight jointly exhaustive and pairwise disjoint relations, called *basic relations*, and by all possible unions of the basic relations (resulting in  $2^8$  different RCC-8 relations altogether). The basic relations are denoted by DC (DisConnected), EC (Externally Connected), PO (Partial Overlap), EQ (EQual), TPP (Tangential Proper Part), NTPP (Non-Tangential Proper Part), and their converses  $TPP^{-1}$  and  $NTPP^{-1}$  [54]; the set of basic relations is denoted  $\mathcal{B}$ . Figure 1 shows two-dimensional examples of these relations.

In the following, an RCC-8 relation will be written as a set of basic relations. E.g.,  $\{DC, PO, TPP\}$  denotes the union ( $\cup$ ) of the three basic relations DC, PO

and TPP. The *universal relation*, denoted by  $\{*\}$ , is the union of all the basic relations. Other operations on RCC-8 relations are intersection ( $\cap$ ), difference ( $\setminus$ ), converse ( $\smile$ ), and composition ( $\circ$ ). Intersection, difference and converse are defined in the standard way. Composition for the relations  $r_1$  and  $r_2$  is defined as follows:

$$r_1 \circ r_2 = \{\langle x, y \rangle \mid \exists z: \langle x, z \rangle \in r_1, \langle z, y \rangle \in r_2\}.$$

A table specifying composition of basic relations is given in [5]; composition of two non-basic relations  $R_1$  and  $R_2$  can be computed as the union of the composition of the basic relations in  $R_1$  and  $R_2$ .

Given a set  $\Theta$  of spatial constraints of the form  $xRy$ , where  $x, y$  are region variables and  $R$  is an RCC-8 relation, a fundamental reasoning problem is deciding the consistency (satisfiability) of  $\Theta$ . This problem is denoted by RSAT.  $\Theta$  is consistent if and only if there exists a *model* of  $\Theta$ , i.e., an assignment of spatial regions to the variables of  $\Theta$  such that all the constraints are satisfied. RSAT for RCC-8 constraints is NP-complete [57].

A set of constraints in RCC-8 involving  $n$  spatial variables can be processed using an  $O(n^3)$  time *path-consistency* algorithm, which makes the set *path-consistent* by eliminating all the impossible labels (basic relations) in every subset of constraints involving three variables [47,48]. Three subsets of RCC-8 for which RSAT can be solved in cubic time by enforcing path-consistency have been identified [57,56]. These subsets, which are denoted by  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  and  $\mathcal{Q}_8$ , are maximal with respect to tractability, i.e., if any RCC-8 relation is added to them, then RSAT becomes NP-complete. Moreover,  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  and  $\mathcal{Q}_8$  are the only maximal tractable fragments of RCC-8 containing all basic relations [56].  $\widehat{\mathcal{H}}_8$  contains 148 relations,  $\mathcal{Q}_8$  contains 160 relations, and  $\mathcal{C}_8$  contains 158 relations.  $\mathcal{NP}_8$  is a set of relations that together with the basic relations are intractable (RSAT for any relation of  $\mathcal{NP}_8$  combined with the set of the basic relations is NP-complete). It contains the following 76 relations which are not contained in any one of  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{Q}_8$ , or  $\mathcal{C}_8$  [56]:

$$\begin{aligned} \mathcal{NP}_8 = & \{R \mid \{\text{PO}\} \not\subseteq R \text{ and } (\{\text{NTPP}\} \subseteq R \text{ or } \{\text{TPP}\} \subseteq R) \\ & \text{and } (\{\text{NTPP}^{-1}\} \subseteq R \text{ or } \{\text{TPP}^{-1}\} \subseteq R)\} \\ & \cup \{\{\text{EC, NTPP, EQ}\}, \{\text{DC, EC, NTPP, EQ}\}, \\ & \{\text{EC, NTPP}^{-1}, \text{EQ}\}, \{\text{DC, EC, NTPP}^{-1}, \text{EQ}\}\}. \end{aligned}$$

The maximal tractable subsets contain the following relations [56]:

$$\widehat{\mathcal{H}}_8 = \{\text{RCC-8} \setminus \mathcal{NP}_8\} \setminus \{R \mid (\{\text{EQ, NTPP}\} \subseteq R \text{ and } \{\text{TPP}\} \not\subseteq R)\}$$

*Algorithm:* SCENARIO

*Input:* A path-consistent set  $\Theta$  of constraints  $\{x_i R x_j \mid i, j = 1 \dots n\}$  over either  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  or  $\mathcal{Q}_8$ .

*Output:* A consistent scenario  $\Theta_s$  of  $\Theta$ .

1.  $\mathcal{S} :=$  maximal tractable set containing the relations used in  $\Theta$  (either  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  or  $\mathcal{Q}_8$ );<sup>3</sup>
2.  $\Theta_s := \emptyset$ ;
3. for every  $xRy \in \Theta$  do
4.   if  $R \in \mathcal{B}$  then  $R' := R$
5.   else if  $\{\text{DC}\} \subseteq R$  then  $R' := \{\text{DC}\}$
6.   else if  $\{\text{EC}\} \subseteq R$  and  $\mathcal{S} \neq \mathcal{C}_8$  then  $R' := \{\text{EC}\}$
7.   else if  $\{\text{PO}\} \subseteq R$  then  $R' := \{\text{PO}\}$
8.   else if  $\{\text{NTPP}\} \subseteq R$  and  $\mathcal{S} = \mathcal{C}_8$  then  $R' := \{\text{NTPP}\}$
9.   else if  $\{\text{NTPP}^{-1}\} \subseteq R$  and  $\mathcal{S} = \mathcal{C}_8$  then  $R' := \{\text{NTPP}^{-1}\}$
10.   else if  $\{\text{TPP}\} \subseteq R$  then  $R' := \{\text{TPP}\}$
11.   else  $R' := \{\text{TPP}^{-1}\}$ ;
12.  $\Theta_s := \Theta_s \cup \{xR'y\}$ .

Fig. 2.  $O(n^2)$  time algorithm for computing a consistent scenario [56]

$$\begin{aligned} & \text{or } (\{\text{EQ}, \text{NTPP}^{-1}\} \subseteq R \text{ and } \{\text{TPP}^{-1}\} \not\subseteq R) \\ \mathcal{C}_8 &= \{\text{RCC-8} \setminus \mathcal{NP}_8\} \setminus \{R \mid \{\text{EC}\} \subseteq R \text{ and } \{\text{PO}\} \not\subseteq R \text{ and} \\ & \quad R \cap \{\text{TPP}, \text{NTPP}, \text{TPP}^{-1}, \text{NTPP}^{-1}, \text{EQ}\} \neq \emptyset\} \\ \mathcal{Q}_8 &= \{\text{RCC-8} \setminus \mathcal{NP}_8\} \setminus \{R \mid \{\text{EQ}\} \subseteq R \text{ and } \{\text{PO}\} \not\subseteq R \text{ and} \\ & \quad R \cap \{\text{TPP}, \text{NTPP}, \text{TPP}^{-1}, \text{NTPP}^{-1}\} \neq \emptyset\} \end{aligned}$$

We assume that a set of constraints  $\Theta$  contains one constraint for each pair of variables involved in  $\Theta$ , i.e., if no information is given about the relation holding between two variables  $x$  and  $y$ , then the universal constraint  $x\{*\}y$  is contained in  $\Theta$ . Another assumption that we make is that whenever a constraint  $xRy$  is in  $\Theta$ , also  $yR^\smile x$  is present.

We say that a set of constraints  $\Theta'$  is a *refinement* of  $\Theta$  if and only if the same variables are involved in both sets, and for every pair of variables  $x, y$ , if  $xR'y \in \Theta'$  and  $xRy \in \Theta$  then  $R' \subseteq R$ .  $\Theta'$  is a *consistent refinement* of  $\Theta$  if and only if  $\Theta'$  is a refinement of  $\Theta$  and both  $\Theta$  and  $\Theta'$  are consistent. A *consistent scenario*  $\Theta_s$  of a set of constraints  $\Theta$  is a consistent refinement of  $\Theta$  where the relation of every constraint in  $\Theta_s$  is a basic relation.

Some of the proofs of our results use particular consistent scenarios derived from a path-consistent set of topological constraints, and exploit some properties of these scenarios. Figure 2 gives an  $O(n^2)$  time algorithm for computing a consistent scenario of a path-consistent set of constraints over  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  and  $\mathcal{Q}_8$  developed in [56]. By applying SCENARIO to a path-consistent set of con-

straints over either  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  or  $\mathcal{Q}_8$  we obtain a particular consistent scenario  $\Theta_s$  of  $\Theta$ . Since exactly this scenario is used in the proof of the main theorem of Section 4, in the following lemma we prove the properties of  $\Theta_s$  that will be exploited.

**Lemma 1** *Let  $\Theta$  be a path-consistent set of constraints over either  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  or  $\mathcal{Q}_8$  and let  $\Theta_s$  be the output of  $\text{SCENARIO}(\Theta)$ . For each pair of variables  $x, y$  involved in  $\Theta$  the following holds:*

- (1) *The constraint  $x\{\text{EQ}\}y$  is contained in  $\Theta_s$  if and only if it is also contained in  $\Theta$ .*
- (2) *The constraint  $xR'y$  with  $R' \in \{\{\text{DC}\}, \{\text{EC}\}, \{\text{PO}\}\}$  is contained in  $\Theta_s$  if and only if  $xRy \in \Theta$  with  $R' \subseteq R$ .*
- (3) *In all the other cases,  $xRy \in \Theta$  is refined to one of  $x\{\text{TPP}\}y$ ,  $x\{\text{NTPP}\}y$ ,  $x\{\text{TPP}^{-1}\}y$ , or  $x\{\text{NTPP}^{-1}\}y$  in  $\Theta_s$ .*

**Proof.** As shown in [56], the output of  $\text{SCENARIO}(\Theta)$  is a consistent scenario  $\Theta_s$  if  $\Theta$  is consistent and contains only constraints over relations of either  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  or  $\mathcal{Q}_8$ . Thus, in order to prove the lemma, we only have to prove that the three specified conditions hold for  $\Theta_s$ . The first condition holds, since the constraint  $x\{\text{EQ}\}y$  can be added to  $\Theta_s$  only by Step 4 of  $\text{SCENARIO}(\Theta)$ . In this case  $x\{\text{EQ}\}y$  is also contained in  $\Theta$ . It follows from the definition of  $\mathcal{C}_8$  that  $\{\text{EC}\} \subset R \in \mathcal{C}_8$  only if  $R$  also contains  $\{\text{DC}\}$  or  $\{\text{PO}\}$ . Thus, whenever a constraint  $xRy \in \Theta$  contains one of  $\{\text{DC}\}$ ,  $\{\text{EC}\}$  or  $\{\text{PO}\}$ , it is refined by Steps 5 to 7 of  $\text{SCENARIO}(\Theta)$  to one of  $x\{\text{DC}\}y$ ,  $x\{\text{EC}\}y$  or  $x\{\text{PO}\}y$ . This proves the second condition. The only non-basic relations which are contained in one of  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  or  $\mathcal{Q}_8$  and which do not contain any of  $\{\text{DC}\}$ ,  $\{\text{EC}\}$ , or  $\{\text{PO}\}$  are  $\{\text{TPP}, \text{NTPP}\}$  and  $\{\text{TPP}^{-1}, \text{NTPP}^{-1}\}$  (which are contained in all three maximal tractable subsets),  $\{\text{TPP}, \text{EQ}\}$ ,  $\{\text{TPP}, \text{NTPP}, \text{EQ}\}$ ,  $\{\text{TPP}^{-1}, \text{EQ}\}$ , and  $\{\text{TPP}^{-1}, \text{NTPP}^{-1}, \text{EQ}\}$  (which are only contained in  $\widehat{\mathcal{H}}_8$  and  $\mathcal{C}_8$ ), and  $\{\text{NTPP}, \text{EQ}\}$  and  $\{\text{NTPP}^{-1}, \text{EQ}\}$  (which are only contained in  $\mathcal{C}_8$ ). Thus, the third condition is an immediate consequence of Steps 8 to 11 of  $\text{SCENARIO}$ , the previous conditions, and the fact that (under the specified preconditions)  $\Theta_s$  is a consistent scenario. ■

## 2.2 The Region Connection Calculus RCC-7

In some domains not every basic relation of RCC-8 is possible between spatial regions. Consider, for examples, domains in which we have solid-state physical objects. It is clear that spatial regions corresponding to these kind of objects cannot partially overlap, while an object can be part of another. For instance,

<sup>3</sup> If the set of relations used in  $\Theta$  is a subset of more than one of the maximal tractable subsets, then  $\mathcal{S}$  can be arbitrarily set to one of  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$ , or  $\mathcal{Q}_8$ .



a hard disk is usually a non-tangential proper part of a computer, while a floppy drive should better be tangential proper part of it. Another example which has been given by Grigni *et al.* [34] refers to geographic applications: “geographic regions and administrative subdivisions obviously can only meet or contain one another, but cannot overlap.” (Additional examples are given in Section 7.)

In order to deal with these kinds of applications where regions cannot partially overlap, we introduce a sub-algebra of RCC-8 which consists of all the RCC-8 relations except those containing PO. If regions cannot partially overlap each other, the seven remaining RCC-8 relations are jointly exhaustive and, thus, they form a set of basic relations from which we can derive a relation algebra. We denote this set as  $\mathcal{B}_7$  and the powerset of it, i.e., the full algebra consisting of  $2^7$  relations, as RCC-7. The RCC-7 universal relation is denoted by  $\{*_7\}$ . Composition of two RCC-7 basic relations can be derived from the composition table of RCC-8 [5] by eliminating the line and the column involving PO, by eliminating PO from each table entry, and by replacing  $*$  with  $*_7$ . Composition of two non-basic RCC-7 relations can be derived in the same way as the composition of two non-basic RCC-8 relations.

The complexity of reasoning over RCC-7 cannot be immediately derived from known complexity results for RCC-8, since the NP-hardness proofs of RCC-8 given in [57,56] involve the relation PO. Nevertheless, because  $\text{RCC-7} \subset \text{RCC-8}$ , it is clear that the intersections of RCC-7 with  $\widehat{\mathcal{H}}_8$ , with  $\mathcal{Q}_8$ , and with  $\mathcal{C}_8$  gives tractable subsets of RCC-7. However, the RCC-7 universal relation is not contained in these sets. Actually, it turns out that RSAT becomes NP-complete if we add  $*_7$  to the RCC-7 basic relations. We prove this by a reduction from the NOT-ALL-EQUAL-3SAT problem, the NP-complete problem of deciding whether a 3SAT formula is satisfiable in such a way that, for each clause at least one literal is assigned as *true*, and at least one literal is assigned as *false* [25].

**Theorem 2** *RSAT for  $\mathcal{S} = \mathcal{B}_7 \cup \{*_7\}$  is NP-complete.*

**Proof.** NP-hardness is proved by a polynomial reduction from NOT-ALL-EQUAL-3SAT that is similar to one used in [57]. Given an instance of this problem we derive a set of constraints over  $\mathcal{B}_7 \cup \{*_7\}$  that is satisfiable if and only if the instance can be positively decided. Every variable  $v$  of the NOT-ALL-EQUAL-3SAT instance is transformed into two RCC-8 constraints  $x_v\{R_t, R_f\}y_v$  and  $x_{-v}\{R_t, R_f\}y_{-v}$  corresponding to the positive and negative literal of  $v$ ; every literal occurrence  $l$  is transformed into an RCC-8 constraint  $x_l\{R_t, R_f\}y_l$ .  $R_t$  and  $R_f$  are mutually disjoint RCC-8 relations such that  $R_t$  holds if and only if the corresponding positive literal ( $x_v$ ) or negative literal ( $x_{-v}$ ) or literal occurrence ( $x_l$ ), respectively, is assigned as *true*, and  $R_f$  holds if and only if it is assigned as *false*. In order to guarantee that the following

equivalences are satisfied, we use RCC-8 “polarity constraints” which enforce this behavior:

- $x_v\{R_t\}y_v$  holds if and only if  $x_{-v}\{R_f\}y_{-v}$  holds,
- $x_v\{R_t\}y_v$  holds if and only if  $x_p\{R_t\}y_p$  holds for every positive literal occurrence  $p$  of  $v$ , and
- $x_v\{R_t\}y_v$  holds if and only if  $x_n\{R_f\}y_n$  holds for every negative literal occurrence  $n$  of  $v$ .

Additionally, we use RCC-8 “clause constraints” which enforce the clause requirement that, for every clause  $(i \vee j \vee k)$  of the NOT-ALL-EQUAL-3SAT instance,  $R_f$  must hold for at least one of the constraints  $x_i\{R_t, R_f\}y_i$ ,  $x_j\{R_t, R_f\}y_j$  and  $x_k\{R_t, R_f\}y_k$ , and  $R_t$  must hold for at least one of them.

Let  $\widehat{\mathcal{S}}$  be the closure of  $\mathcal{S}$  under composition, intersection, and inverse. By Theorem 5 in [57],  $\text{RSAT}(\widehat{\mathcal{T}})$  is NP-hard if and only if  $\text{RSAT}(\mathcal{T})$  is NP-hard, for any set of relations  $\mathcal{T}$  which contains the universal relation. NOT-ALL-EQUAL-3SAT can be reduced to  $\text{RSAT}(\widehat{\mathcal{S}})$  by using the following polarity and clause constraints, and having  $R_t = \{\text{TPP}\}$  and  $R_f = \{\text{TPP}^{-1}\}$ .

Polarity constraints (see Figure 3):

$$\begin{aligned} &x_v\{\text{EC}\}x_{-v}, y_v\{\text{TPP}, \text{TPP}^{-1}, \text{NTPP}, \text{NTPP}^{-1}\}y_{-v} \\ &x_v\{\text{DC}, \text{NTPP}\}y_{-v}, y_v\{\text{DC}, \text{NTPP}^{-1}\}x_{-v} \end{aligned}$$

Clause constraints (for every clause  $i \vee j \vee k$ ):

$$y_i\{\text{EQ}\}x_j, y_j\{\text{EQ}\}x_k, y_k\{\text{EQ}\}x_i.$$

All the relations used for  $R_t$ ,  $R_f$ , the polarity constraints, and for the clause constraints are contained in  $\widehat{\mathcal{S}}$ , and it is easy to see that the resulting set of topological constraints is satisfiable if and only the instance of NOT-ALL-EQUAL-3SAT under consideration can be positively decided. Thus,  $\text{RSAT}(\widehat{\mathcal{S}})$  is NP-hard and by Theorem 5 in [57],  $\text{RSAT}(\mathcal{S})$  is also NP-hard. Membership of  $\text{RSAT}(\mathcal{S})$  in NP is obvious given that  $\text{RSAT}(\mathcal{B}_7)$  is tractable. ■

This NP-completeness result is somewhat surprising since  $\text{RSAT}$  for the basic RCC-8 relations and  $\{*\}$  is clearly tractable. It suggests also that an algorithm for determining the consistency of a set  $\Theta$  of constraints over  $\mathcal{B}_7 \cup \{*_7\}$  can be obtained by enumerating the scenarios of  $\Theta$  using backtracking, and running a path-consistency algorithm for checking their consistency.

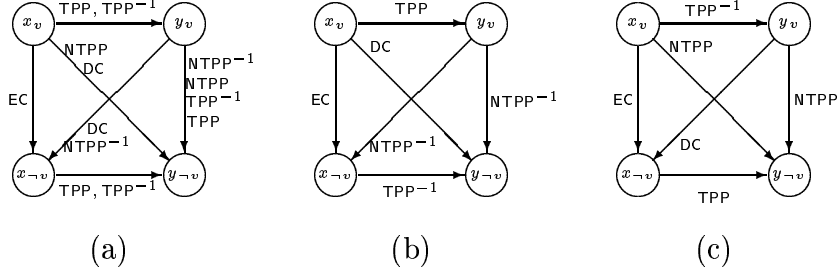


Fig. 3. The polarity constraints (a) for the transformation of NOT-ALL-EQUAL-3SAT assure that positive and negative literals of the same variable have opposite assignments: (b) and (c) are the only possible refinements of the corresponding relations to basic relations.

### 3 Combining Topological and Qualitative Size Relations

In this section we introduce  $\mathcal{QS}$ , a class of qualitative relations between region sizes, and we combine this class with RCC-8. We also give some technical results that will be used in the next section, where we present an algorithm for processing constraints in the combined framework. Since RCC-7 is a subset of RCC-8, the results we give in this and in the following sections are valid also for RCC-7. In case of notable differences, we will specify them.

We will assume that all the spatial regions are measurable sets in  $\mathbb{R}^n$  [3]. Note that this assumption does not compromise the computational properties of  $\widehat{\mathcal{H}}_8$ , because from [55] it follows that the regions of every consistent set of RCC-8 constraints can always be interpreted as measurable sets (e.g., as sets of spheres in  $\mathbb{R}^3$ ). We will also assume that the size of an n-dimensional region corresponds to its n-dimensional measure [3]. For example, the size of a sphere in  $\mathbb{R}^3$  corresponds to its volume. Moreover, we assume that space does not have an upper bound, i.e., for each region  $x$  there exists another region that contains  $x$ .

Given a set  $V$  of spatial region variables, a set of  $\mathcal{QS}$ -constraints over  $V$  is a set of constraints of the form  $size(x) S size(y)$ , where  $S \in \mathcal{QS}$ ,  $size(x)$  is the size of the region  $x$ ,  $size(y)$  is the size of the region  $y$ , and  $x, y \in V$ .

**Definition 3**  $\mathcal{QS}$  is the class formed by the following eight qualitative relations between the size of spatial regions:  $<$ ,  $>$ ,  $\leq$ ,  $\neq$ ,  $=$ ,  $\geq$ ,  $?$  and  $\emptyset$ , where  $?$  is the universal relation,  $\emptyset$  is the empty relation, and  $<$ ,  $>$ , and  $=$  are basic relations.

The relations of  $\mathcal{QS}$  are the same as those of the temporal Point Algebra [62,63,61,42] forming a relation algebra [60].

**Proposition 4** The relations of  $\mathcal{QS}$  form a Point Algebra.

$r$	$Sizerel(r)$	$r$	$Sizerel(r)$	$s$	$Toprel(s)$
TPP	$\models <$	DC	$\models ?$	$=$	$\models$ DC, EC, PO, EQ
NTPP	$\models <$	EC	$\models ?$	$>$	$\models$ DC, EC, PO, TPP <sup>-1</sup> , NTPP <sup>-1</sup>
TPP <sup>-1</sup>	$\models >$	PO	$\models ?$	$<$	$\models$ DC, EC, PO, TPP, NTPP
NTPP <sup>-1</sup>	$\models >$	EQ	$\models =$		

Table 1

Interdependencies of basic RCC-8 relations ( $r$ ) and basic QS relations ( $s$ )

The topological RCC-8 relations and the relative size relations are not independent from each other. Table 1 gives the size relations that are entailed by the basic RCC-8 relations, and the topological relations that are entailed by the basic size relations.  $Sizerel(R)$  indicates the strongest size relation entailed by the topological relation  $R$ , and  $Toprel(S)$  indicates the strongest topological relation entailed by the size relation  $S$ .<sup>4</sup> The interdependencies specified in Table 1 also hold for RCC-7, if PO is removed from all the entries. The dependencies from a non-basic relation  $R$  can be obtained by disjunctively combining the relations entailed by each basic relation in  $R$ . For example,  $\{TPP, EQ\}$  entails “ $\leq$ ”.

Since any topological relation – and any sub-relation thereof – entailed by the basic size relations is contained in  $\widehat{\mathcal{H}}_8$ , the following proposition is true.

**Proposition 5** *The relation  $R \in \text{RCC-8} \setminus \widehat{\mathcal{H}}_8$  of any constraint  $xRy$  can be consistently refined to a relation  $R' \in \widehat{\mathcal{H}}_8$ , if an appropriate size constraint between  $x$  and  $y$  is given. In particular, if definite size information is given, then  $R$  can always be consistently refined to a relation  $R' \in \widehat{\mathcal{H}}_8$ .*

For example, the RCC-8  $\setminus \widehat{\mathcal{H}}_8$  constraint  $x\{TPP, TPP^{-1}, NTPP, DC, EC\}y$  can be consistently refined to the  $\widehat{\mathcal{H}}_8$  constraint  $x\{TPP, NTPP, DC, EC\}y$ , if the size constraint  $size(x) \leq size(y)$  is given.

Before presenting an algorithm for processing RCC-8 constraints combined with qualitative size constraints, we need to give some further technical definitions and results that will be used in the next section to prove the formal properties of the algorithm.

**Definition 6 (Model for  $\Sigma$ )** *Given a set  $\Sigma$  of constraints in QS, we say that an assignment  $\sigma$  of spatial regions to the variables of  $\Sigma$  is a model of  $\Sigma$  if and only if  $\sigma$  satisfies all the constraints in  $\Sigma$ .*

**Definition 7 (Consistency for  $\Theta \cup \Sigma$ )** *Given a set  $\Theta$  of constraints in RCC-8 and a set  $\Sigma$  of constraints in QS,  $\Theta \cup \Sigma$  is consistent if there exists a model*

<sup>4</sup> A relation (or constraint)  $R$  is stronger than another relation (constraint)  $R'$  if and only if  $R \models R'$  and  $R' \not\models R$ .

of  $\Theta$  which is also a model of  $\Sigma$ .

The problem of deciding the consistency (or satisfiability) for a set of topological constraints  $\Theta$  and a set of size constraints  $\Sigma$  will be indicated with **RSAT** for  $\Theta \cup \Sigma$ . We say that a consistent scenario for a set  $\Theta$  of constraints is *size-consistent* relative to a set  $\Sigma$  of constraints if and only if there exists a model for the scenario that is also a model of  $\Sigma$ .

The next lemma states that non-forced equalities can be omitted from a path-consistent set of size constraints in  $\mathcal{QS}$  without losing consistency. The proof of this lemma is based on van Beek's algorithm for computing a consistent scenario for a set of constraints in the Point Algebra [61]. In this method the constraints are represented through a directed labeled graph, where the vertices represent the temporal variables and the edges represent the constraints between them. If the constraints are consistent, then the strongly connected components (SCCs) of the graph correspond to sets of variables that are equivalent (the strongest entailed constraint between each pair of variables in the same SCC is "="). The elements of each SCC are collapsed into a single vertex representing a class of variables that are equal to each other; each  $\leq$ -constraint relating variables belonging to different SCCs is consistently replaced by a  $<$ -constraint, and each  $\geq$ -constraint relating variables of different SCCs is consistently replaced by a  $>$ -constraint;  $\neq$ -constraints are ignored; finally, variables that are related by the universal relation are interpreted as different points (for more details on this algorithm see [61]).

**Lemma 8** *Let  $\Sigma$  be a path-consistent set of size constraints over  $\mathcal{QS}$  and  $\Sigma'$  the set of size constraints such that, for each constraint  $size(i) S size(j)$  in  $\Sigma$ ,*

- (1) *if  $S \in \{<, >\}$  then  $size(i) S size(j) \in \Sigma'$ ,*
- (2) *if  $S = "<="$  then  $size(i) < size(j) \in \Sigma'$ ,*
- (3) *if  $S = ">="$  then  $size(i) > size(j) \in \Sigma'$ ,*
- (4) *if  $S = "="$  then  $size(i) = size(j) \in \Sigma'$ ,*
- (5) *if  $S = "?"$  then  $size(i) \neq size(j) \in \Sigma'$ .*

*$\Sigma'$  is consistent and any model of  $\Sigma'$  is also a model of  $\Sigma$ .*

**Proof.** Any path-consistent set of PA-constraints is consistent [41,42] (if it does not contain the empty relation). Hence  $\Sigma$  is consistent. Moreover, as shown in [61],  $x = y$  belongs to a path-consistent set  $S$  of PA-constraints if and only if this is the strongest constraint between  $x$  and  $y$  entailed by  $S$ . It follows that, by applying van Beek's consistent scenario algorithm to  $\Sigma$ , we obtain a scenario that is also a consistent scenario for  $\Sigma'$ . Thus  $\Sigma'$  is consistent. Furthermore, it is clear that any model of  $\Sigma'$  is also a model of  $\Sigma$ , because the constraints of  $\Sigma'$  are stronger than the corresponding constraints in  $\Sigma$ . ■

Let  $\Theta$  be a set of constraints in RCC-8,  $\Sigma$  a set of constraints in  $\mathcal{QS}$ ,  $t_{ij}$  the relation between  $i$  and  $j$  in  $\Theta$ , and  $s_{ij}$  the relation between  $size(i)$  and  $size(j)$  in  $\Sigma$ . We say that:  $t_{ij}$  entails the negation of  $s_{ij}$  ( $t_{ij} \models \neg s_{ij}$ ) if and only if  $Sizerel(t_{ij}) \cap s_{ij} = \emptyset$ ;  $s_{ij}$  entails the negation of  $t_{ij}$  ( $s_{ij} \models \neg t_{ij}$ ) if and only if  $Toprel(s_{ij}) \cap t_{ij} = \emptyset$ .

**Proposition 9** *A consistent set  $\Theta$  of constraints in RCC-8 entails the negation of a  $\mathcal{QS}$  relation  $s_{ij}$  between  $size(i)$  and  $size(j)$  if and only if  $Sizerel(\hat{t}_{ij}) \cap s_{ij} = \emptyset$ , where  $\hat{t}_{ij}$  is the strongest entailed relation between  $i$  and  $j$  in  $\Theta$ .*

**Proof.** It follows from the fact that, for any  $i$  and  $j$ ,  $\Theta \models \neg s_{ij}$  if and only if  $\hat{t}_{ij} \models \neg s_{ij}$ , and from the definition of *Sizerel*. ■

**Proposition 10** *A consistent set  $\Sigma$  of constraints in  $\mathcal{QS}$  entails the negation of a RCC-8 relation  $t_{ij}$  between  $i$  and  $j$  if and only if  $Toprel(\hat{s}_{ij}) \cap t_{ij} = \emptyset$ , where  $\hat{s}_{ij}$  is the strongest entailed relation between  $i$  and  $j$  in  $\Sigma$ .*

**Proof.** It follows from the fact that, for any  $i$  and  $j$ ,  $\Sigma \models \neg t_{ij}$  if and only if  $\hat{s}_{ij} \models \neg t_{ij}$ , and from the definition of *Toprel*. ■

**Lemma 11** *Let  $\Theta$  be a consistent set of constraints in RCC-8,  $\Sigma$  a consistent set of  $\mathcal{QS}$ -constraints over the variables of  $\Theta$ ,  $t_{ij}$  the relation between  $i$  and  $j$  in  $\Theta$ , and  $s_{ij}$  the relation between  $size(i)$  and  $size(j)$  in  $\Sigma$ .*

- $t_{ij} \models \neg s_{ij}$  if and only if  $s_{ij} \models \neg t_{ij}$ ;
- $\Theta \models \neg \hat{s}_{ij}$  if and only if  $\Sigma \models \neg \hat{t}_{ij}$ .

**Proof.** It follows from Table 1 and Propositions 9–10. ■

In the next lemma  $t_{ij}$  indicates the *basic* relation between  $i$  and  $j$  in a consistent scenario  $\Theta_s$  for a set  $\Theta$  of topological relations.

**Lemma 12** *Let  $\Theta_s$  be a consistent scenario for a (consistent) set  $\Theta$  of topological constraints in RCC-8. It is possible to construct a model of  $\Theta$  that is also a model for the set  $\Sigma$  of size constraints obtained in the following way. For each  $i, j$  ( $i \neq j$ ):*

- (1) if  $Sizerel(t_{ij})$  is one of  $<$ ,  $>$ ,  $=$ , then  $size(i) < size(j)$ ,  $size(i) > size(j)$ , and  $size(i) = size(j)$ , respectively, is added to  $\Sigma$ ;
- (2) if  $Sizerel(t_{ij})$  is the universal relation, then one of  $size(i) < size(j)$  or  $size(i) > size(j)$  can be arbitrarily chosen to be added to  $\Sigma$  (provided that  $\Sigma$  remains consistent).

**Proof.** Let  $s_{ij}$  be the relation between the region sizes  $size(i)$  and  $size(j)$  in  $\Sigma$ . We show that it is possible to construct a model  $\theta$  of  $\Theta_s$  in which the

values (spatial regions) assigned to the variables satisfy  $\Sigma$ . Suppose that this were not true. Since  $\Theta_s$  and  $\Sigma$  are minimal sets, we would have that (a) there would exist  $h$  and  $k$  such that  $\Theta_s \models \neg s_{hk}$ , or (b) there would exist  $h'$  and  $k'$  such that  $\Sigma \models \neg t_{h'k'}$  (i.e., there is no model of  $\Theta_s$  satisfying  $s_{hk}$ , or there is no model of  $\Sigma$  consistent with  $t_{h'k'}$ .) Since by construction of  $\Theta_s$ , for any pair of variables in  $\Theta_s$  the strongest relation between  $i$  and  $j$  is  $t_{ij}$ , (a) can hold only if (a')  $t_{hk} \models \neg s_{hk}$  holds. For analogous reasons we have that (b) can hold only if (b')  $s_{h'k'} \models \neg t_{h'k'}$  holds. But both (a') and (b') cannot hold. In fact, since for any  $i, j$   $\text{Sizerel}(t_{ij}) \in \{<, >, =, ?\}$  (because  $t_{ij}$  is basic), by (1) it cannot be the case that  $t_{hk} \models \neg s_{hk}$ , and hence by Lemma 11 also  $s_{h'k'} \models \neg t_{h'k'}$  cannot hold. ■

#### 4 Reasoning about Topological and Qualitative Size Relations

A natural method for deciding the consistency of a set of RCC-8 constraints and a set of  $\mathcal{QS}$ -constraints, would be to first extend each set of constraints with the constraints entailed by the other set, and then independently check the consistency of the extended sets by using a path-consistency algorithm. However, as the example below shows, this method is not complete for  $\widehat{\mathcal{H}}_8$  constraints.

Another possibility, would be to compute the strongest entailed relations (minimal relations) between each pair of variables before propagating constraints from one set to the other. However, this method has the disadvantage that it is computationally expensive.<sup>5</sup>

Finally, a third method could be based on iteratively using path-consistency as a preprocessing technique and then propagating the information from one set to the other.<sup>6</sup> The following example shows that the information would need to be propagated more than once, and furthermore it is not clear whether in general this method would be complete for detecting inconsistency.

**Example.** Consider the set  $\Theta$  formed by the following  $\widehat{\mathcal{H}}_8$  constraints

$$x_0\{\text{TPP}, \text{EQ}\}x_2, x_1\{\text{TPP}, \text{EQ}, \text{PO}\}x_0, x_1\{\text{TPP}, \text{EQ}\}x_2, x_4\{\text{TPP}, \text{EQ}\}x_3,$$

<sup>5</sup> The best known algorithm for computing the minimal network of a set of constraints over either  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  or  $\mathcal{Q}_8$  requires  $O(n^5)$  time.

<sup>6</sup> A similar method is used by Ladkin and Kautz to combine qualitative and metric constraints in the context of temporal reasoning [38]. Note that imposing path-consistency is sufficient for consistency checking of a set of constraints over  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$ ,  $\mathcal{Q}_8$ , and  $\mathcal{QS}$ , but is incomplete for computing the minimal relations [61,57].

and the set  $\Sigma$  formed by the of following  $\mathcal{QS}$ -constraints

$$size(x_0) < size(x_2), size(x_3) \leq size(x_1), size(x_2) \leq size(x_4).$$

We have that  $\Theta$  and  $\Sigma$  are independently consistent, but their union is not consistent. Moreover, the following propagation scheme does not detect the inconsistency: (a) enforce path-consistency to  $\Sigma$  and  $\Theta$  independently; (b) extend  $\Sigma$  with the size constraints entailed by the constraints in  $\Theta$ ; (c) extend  $\Theta$  with the topological constraints entailed by the constraints in  $\Sigma$ ; (d) enforce path-consistency to  $\Theta$  and  $\Sigma$  again. In order to detect that  $\Theta \cup \Sigma$  is inconsistent, we need an additional propagation of constraints from the topological set to the size set.

Instead of directly analyzing the complexity and completeness of the propagation scheme illustrated in the previous example, we propose a new method for dealing with combined topological and qualitative size constraints. In particular, we propose an  $O(n^3)$  time and  $O(n^2)$  space algorithm, BIPATH-CONSISTENCY, for imposing path-consistency to a set of constraints in  $\text{RCC-8} \cup \mathcal{QS}$ . We prove that BIPATH-CONSISTENCY solves RSAT for any input set  $\Theta$  of topological constraints in either  $\widehat{\mathcal{H}}_8, \mathcal{C}_8$  or  $\mathcal{Q}_8$ , combined with any set of size constraints in  $\mathcal{QS}$  involving the variables of  $\Theta$ . Thus, despite this framework is more expressive than a purely topological one over the same set of relations (and therefore has a larger potential applicability), the problem of deciding consistency can be solved without additional worst-case cost.

BIPATH-CONSISTENCY is a modification of Vilain and Kautz’s path-consistency algorithm [62,63] as described by Bessi ere [8], which in turn is a slight modification of Allen’s algorithm [1]. The main novelty of our algorithm is that BIPATH-CONSISTENCY operates on a graph of *pairs* of constraints. The vertices of the graph are constraint variables, which in our context correspond to spatial regions. Each edge of the graph is labeled by a pair of relations formed by a topological relation in  $\text{RCC-8}$  and a size relation in  $\mathcal{QS}$ . The function  $\text{BIREVISION}(i, k, j)$  has the same role as the function  $\text{REVISE}$  used in path consistency algorithms for constraint networks (e.g., [47]). The main difference is that  $\text{BIREVISION}(i, k, j)$  considers pairs of (possibly interdependent) constraints, instead of single constraints.

A formal description of BIPATH-CONSISTENCY is given in Figure 4, where  $R_{ij}$  is a pair formed by a relation  $t_{ij}$  in  $\text{RCC-8}$  and a relation  $s_{ij}$  in  $\mathcal{QS}$ ;  $R_{ij} = \emptyset$  when  $t_{ij} = \emptyset$  or  $s_{ij} = \emptyset$ ;  $U_t$  indicates the universal relation in  $\text{RCC-8}$  and  $U_s$  the universal relation in  $\mathcal{QS}$ .

**Theorem 13** *Given a set  $\Theta$  of constraints in either  $\widehat{\mathcal{H}}_8, \mathcal{C}_8$  or  $\mathcal{Q}_8$ , and a set  $\Sigma$*

---

<sup>7</sup> As in the function  $\text{REVISE}$  given in [8], this step is used to avoid processing the triple  $i, j, k$  when it is known that  $R_{ij}$  would not be revised.



*Algorithm:* BIPATH-CONSISTENCY

*Input:* A set  $\Theta$  of RCC-8 constraints, and a set  $\Sigma$  of  $\mathcal{QS}$ -constraints over the variables  $x_1, x_2, \dots, x_n$  of  $\Theta$ .

*Output:* **fail**, if  $\Sigma \cup \Theta$  is not consistent; path-consistent sets equivalent to  $\Sigma$  and  $\Theta$ , otherwise.

1.  $Q \leftarrow \{(i, j) \mid i < j\}$ ; ( $i$  indicates the  $i$ -th variable of  $\Theta$ . Analogously for  $j$ )
2. *while*  $Q \neq \emptyset$  *do*
3. select and delete an arc  $(i, j)$  from  $Q$ ;
4. *for*  $k \neq i, k \neq j$  ( $k \in \{1..n\}$ ) *do*
5.     *if* BIREVISION( $i, j, k$ ) *then*
6.         *if*  $R_{ik} = \emptyset$  *then return fail*
7.         *else add*  $(i, k)$  *to*  $Q$ ;
8.     *if* BIREVISION( $k, i, j$ ) *then*
9.         *if*  $R_{kj} = \emptyset$  *then return fail*
10.         *else add*  $(k, j)$  *to*  $Q$ .

*Function:* BIREVISION( $i, k, j$ )

*Input:* three region variables  $i, k$  and  $j$

*Output:* true, if  $R_{ij}$  is revised; false otherwise.

*Side effects:*  $R_{ij}$  and  $R_{ji}$  revised using the operations  $\cap$  and  $\circ$  over the constraints involving  $i, k$ , and  $j$ .

1. *if* one of the following cases hold, *then return false*:<sup>7</sup>
  - (a)  $Toprel(s_{ik}) \cap t_{ik} = U_t$  and  $Sizerel(t_{ik}) \cap s_{ik} = U_s$ ,
  - (b)  $Toprel(s_{kj}) \cap t_{kj} = U_t$  and  $Sizerel(t_{kj}) \cap s_{kj} = U_s$
2.  $oldt := t_{ij}$ ;  $olds := s_{ij}$ ;
3.  $t_{ij} := (t_{ij} \cap Toprel(s_{ij})) \cap ((t_{ik} \cap Toprel(s_{ik})) \circ (t_{kj} \cap Toprel(s_{kj})))$ ;
4.  $s_{ij} := (s_{ij} \cap Sizerel(t_{ij})) \cap ((s_{ik} \cap Sizerel(t_{ik})) \circ (s_{kj} \cap Sizerel(t_{kj})))$ ;
5. *if*  $s_{ij} \neq olds$  *then*  $t_{ij} := (t_{ij} \cap Toprel(s_{ij}))$ ;
6. *if* ( $oldt = t_{ij}$ ) *and* ( $olds = s_{ij}$ ) *then return false*;
7.  $t_{ji} := Converse(t_{ij})$ ;  $s_{ji} := Converse(s_{ij})$ ;
8. *return true*.

Fig. 4. BIPATH-CONSISTENCY

*of constraints in  $\mathcal{QS}$  involving variables in  $\Theta$ , BIPATH-CONSISTENCY applied to  $\Sigma$  and  $\Theta$  solves RSAT for  $\Theta \cup \Sigma$ .*

**Proof.** It is clear that, if the algorithm returns **fail**, then  $\Sigma \cup \Theta$  is inconsistent. Otherwise (the algorithm does not return **fail**) both the output set of size constraints  $\Sigma_p$  and the output set  $\Theta_p$  of topological constraints are independently path-consistent. Hence, by Proposition 4 and the fact that a path-consistent set of constraints either in  $\widehat{\mathcal{H}}_8, \mathcal{C}_8, \mathcal{Q}_8$  or in a Point Algebra is consistent [57,41],  $\Sigma$  and  $\Theta$  are independently consistent.

Let  $\Theta_p$  be the path-consistent set of topological constraints given as output of BIPATH-CONSISTENCY applied to  $\Sigma$  and  $\Theta$ , and  $\Sigma_p$  the path-consistent set of the size constraints. We show that  $\Sigma_p \cup \Theta_p$  is consistent (and therefore that  $\Sigma \cup \Theta$  is consistent). In order to do that, we show that it is possible to construct a consistent scenario  $\Theta_s$  for  $\Theta_p$  in which the region variables can be interpreted as regions satisfying the constraints of  $\Sigma$ .

Let  $\Theta_s$  be a consistent scenario for  $\Theta_p$  in which, for any pair of variables  $i$  and  $j$ , the (basic) relation  $r_{ij}$  between  $i$  and  $j$  is

- EQ if  $i\{EQ\}j \in \Theta_p$ ,
- one of DC, EC, PO, if  $R \cap \{DC, EC, PO\} \neq \emptyset$ , where  $iRj \in \Theta_p$ ,
- one of TPP, NTPP,  $TPP^{-1}$ ,  $NTPP^{-1}$ , otherwise.

Lemma 1 guarantees the existence of  $\Theta_s$ . From  $\Theta_s$  we can derive an assignment to the variables of  $\Theta_s$  satisfying the constraints of  $\Sigma_p$  (and the topological constraints of  $\Theta_s$ ) in the following way. Let  $\Sigma'_p$  be the set of size constraints derived from  $\Sigma_p$  by applying the five transformation rules of Lemma 8, and let  $\sigma_p$  be a consistent scenario for  $\Sigma'_p$ . By Lemma 8  $\sigma_p$  is also a consistent scenario for  $\Sigma_p$  (and hence for  $\Sigma$ ).

For each pair of variables  $i$  and  $j$ , consider the size relation  $Sizerel(r_{ij})$  between  $i$  and  $j$ . By construction of  $\Theta_s$  and Steps 3–7 of BIREVISION (the subroutine used by BIPATH-CONSISTENCY to revise topological and size constraints), it is clear that if  $Sizerel(r_{ij})$  is one of “<”, “>”, “=”, then the relation between  $i$  and  $j$  in  $\Sigma'_p$  (and in  $\sigma_p$ ) is the same as  $Sizerel(r_{ij})$ . So, any assignment satisfying  $r_{ij}$  satisfies also the size relations between  $i$  and  $j$  in  $\Sigma'_p$  (and in  $\sigma_p$ ). Consider now the case in which  $Sizerel(r_{ij})$  is the indefinite relation (“?”). (Note that since  $r_{ij}$  is a basic relation it cannot be the case that  $Sizerel(r_{ij}) \in \{\leq, \geq, \neq\}$  – see Table 1.) We have that  $r_{ij}$  must be one of {DC}, {EC}, {PO}. Since  $\Sigma_p$  is consistent, by construction of  $\sigma_p$  and by Lemma 12 we can consistently assign regions to  $i$  and  $j$  satisfying  $r_{ij}$  and the size relations between  $i$  and  $j$  in  $\sigma_p$  (and hence in  $\Sigma'_p$ ). Consequently, by Lemma 8 from  $\Theta_s$  we can derive a consistent assignment satisfying the relations in  $\Sigma_p$  (and hence in  $\Sigma$ ). ■

**Theorem 14** *Given a set  $\Theta$  of constraints over RCC-8 and a set  $\Sigma$  of constraints in QS involving variables in  $\Theta$ , the time and space complexity of BIPATH-CONSISTENCY applied to  $\Sigma$  and  $\Theta$  are  $O(n^3)$  and  $O(n^2)$  respectively, where  $n$  is the number of variables involved in  $\Theta$  and  $\Sigma$ .*

**Proof.** Since any relation in QS can be refined at most three times, any relation in RCC-8 can be refined at most eight times, and there are  $O(n^2)$  relations, the total number of edges that can enter into Q is  $O(n^2)$ . For each arc in Q, BIPATH-CONSISTENCY runs BIREVISION  $2n$  times. BIREVISION has a constant time complexity. The quadratic space complexity is trivial. ■

**Theorem 15** *Given a set  $\Theta$  of constraints in either  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  or  $\mathcal{Q}_8$ , and a set*

$\Sigma$  of constraints in  $\mathcal{QS}$  involving variables in  $\Theta$ , RSAT for  $\Theta \cup \Sigma$  can be solved in  $O(n^3)$  time and  $O(n^2)$  space, where  $n$  is the number of variables involved in  $\Theta$  and  $\Sigma$ .

**Proof.** It follows from Theorems 13 and 14. ■

**Theorem 16** *Given a set  $\Theta$  of constraints in either  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  or  $\mathcal{Q}_8$ , and a set  $\Sigma$  of  $\mathcal{QS}$ -constraints involving variables in  $\Theta$ , a size-consistent consistent scenario  $\Theta_s$  for  $\Theta \cup \Sigma$  can be computed in  $O(n^3)$  time and  $O(n^2)$  space, where  $n$  is the number of variables involved in  $\Theta$  and  $\Sigma$ .*

**Proof.** From the proofs of Theorem 13 and Theorem 14, it follows that  $\Theta_s$  can be computed by first applying BIPATH-CONSISTENCY to  $\Theta$  and  $\Sigma$ , and then running the  $O(n^2)$  algorithm (SCENARIO) given in Figure 2 on the set of the topological constraints in the output of BIPATH-CONSISTENCY. ■

To conclude this section we observe that the efficiency of BIPATH-CONSISTENCY could be improved by a constant factor in the following way. For each pair of variables  $x_i$  and  $x_j$ , in the initial queue we have two triples  $(i, j, t)$  and  $(i, j, s)$ , which represent the topological and the size relations between  $x_i$  and  $x_j$  respectively. If the main routine takes the triple  $(i, j, t)$  from the queue, then, when we perform BIREVISION, step 4 is reduced to  $s_{ij} := (s_{ij} \cap \text{Sizerel}(t_{ij}))$ . Similarly, if the triple  $(i, j, s)$  is taken from the queue, then step 3 of BIREVISION is reduced to  $t_{ij} := (t_{ij} \cap \text{Toprel}(s_{ij}))$ . Moreover, step 1 of BIREVISION can be simplified according to whether  $(i, j, t)$  or  $(i, j, s)$  have been taken from the queue. Finally, we add  $(i, j, t)$  to the queue only when BIREVISION revises  $t_{ij}$  (similarly for  $(i, j, s)$ ). Since in the resulting algorithm BIREVISION performs fewer operations than in the original algorithm, this modification can speed up the process by a constant factor, if the input sets require many executions of BIREVISION.

## 5 Combining Topological and Metric Size Constraints

In this section we study the combination of topological information and metric size information, while in the next section we will focus on reasoning in this combined framework. As for the combination of topological information and qualitative size relations (Section 3), we consider some classes of metric size constraints that are known to be (independently) tractable. However, contrary to qualitative size constraints, handling metric size constraints in combination with topological constraints can be computationally very hard. This is mainly because the interdependencies between topological and metric size constraints are more intricate than those between topological and qualitative size constraints.

We will study the following practically useful classes of size constraints, assuming that for each region  $x$  in the input set of topological constraints there is an implicit size constraint stating that the size of  $x$  is larger than zero.

- *Metric relative size constraints* ( $\mathcal{MS}$ ). These are constraints of the form

$$size(x) R \alpha \cdot size(y),$$

where  $R$  is a relation in  $\{<, \leq, =, \neq, \geq, >\}$  and  $\alpha$  is a positive rational number. An example of this type of constraints is

$$size(x) \leq 2.5 size(y)$$

stating that the size of the spatial region  $x$  is smaller than 2.5 times the size of the spatial region  $y$ . Note that  $\mathcal{QS}$ -constraints are a special case of  $\mathcal{MS}$ -constraints.

- *Size difference constraints* ( $\mathcal{SD}$ ). These constraints are analogous to the temporal distance constraints forming a *Simple Temporal Constraint Problem* or STP [18,27]. They are of the form

$$size(y) - size(x) \in I,$$

where  $I$  is a (closed, open or semi-open) continuous interval of rational numbers, with a lower bound  $l$  and an upper bound  $u$ , such that  $l \leq u$ , and  $l \neq u$  if  $I$  is open or semi-open.<sup>8</sup> An example of this type of constraints is

$$size(y) - size(x) \in [1, 3.5],$$

stating that the difference between the size of  $y$  and the size of  $x$  is larger than or equal to 1, and smaller than or equal to 3.5.

- *Domain size constraints* ( $\mathcal{DS}$ ). These constraints are of the form

$$size(x) \in I,$$

where the lower bound  $l$  and an upper bound  $u$  of  $I$  satisfy  $l, u \geq 0$ , and  $I$  is open on the left if  $l = 0$ . An example of this type of constraints is

$$size(x) \in [1.2, 3],$$

stating that the size of  $x$  is a quantity between 1.2 and 3. As we will show,  $\mathcal{DS}$ -constraints can be formalized as a particular subclass of  $\mathcal{SD}$ -constraints. We distinguish  $\mathcal{DS}$  and  $\mathcal{SD}$  because the representation of a problem may need only the class of  $\mathcal{DS}$ -constraints. Thus, they could be computationally easier to treat than  $\mathcal{SD}$ -constraints. We will show that in general this is

<sup>8</sup> The lower and upper bounds can be  $-\infty$  and  $+\infty$ , respectively; in these cases the interval of the constraint is open on the left and on the right, respectively.

not the case. However, when  $\mathcal{DS}$ -constraints are restricted to *definite size constraints*, they can be processed more easily. A definite size constraint is a  $\mathcal{DS}$ -constraint where the lower and upper bounds of the interval are the same, e.g.,  $size(x) \in [3, 3]$ .<sup>9</sup>

These classes of metric size constraints can be used to express information handled by other size calculi that have been introduced in the literature (without a study of their interaction with topological information). Order of Magnitude reasoning [52] refines the qualitative size relations  $<$  and  $>$  of  $\mathcal{QS}$  by the introduction of relations like “slightly smaller” or “much smaller”. These relations are defined using a “tolerance parameter”  $\epsilon$  in the range  $]0, 0.4656[$  such that, e.g.,  $x$  is *slightly smaller than*  $y$  is defined as  $size(x)/size(y) \in ]1/(1 + \epsilon), 1[$ . This is equivalent to the  $\mathcal{MS}$ -constraints  $size(x) < size(y)$  and  $size(x) > 1/(1 + \epsilon) size(y)$ .

The  $\Delta$ -calculus [65] defines ternary relations of the form  $size(x)(R, \delta)size(y)$ , where  $R$  is either  $<$  or  $>$ , meaning  $size(x) = size(y) \pm \delta$ . An example of these constraints is “the size of  $x$  is larger than the size of  $y$  by the amount  $\delta$ ”. If  $\delta$  is a constant, this is equivalent to the  $\mathcal{SD}$ -constraint  $size(x) - size(y) \in [\delta, \delta]$ . If  $\delta$  is a multiple of  $size(x)$  or of  $size(y)$ , we can express this using  $\mathcal{MS}$ -constraints, e.g.,  $size(x)(<, size(y))size(y)$  is equivalent to  $size(x) = 2 size(y)$ . On the other hand, if  $\delta$  is a different variable, we cannot express this ternary relation using our classes of metric size constraints.

As for the combination of  $\mathcal{RCC-8}$  and  $\mathcal{QS}$ , in order to combine  $\mathcal{MS}$ ,  $\mathcal{DS}$  and  $\mathcal{SD}$  constraints with  $\mathcal{RCC-8}$ , we have to take into account size constraints that are entailed by the given topological information, as well as topological constraints that are entailed by the given size information. Regarding size constraints that are entailed by  $\mathcal{RCC-8}$  constraints, we have the same *Sizerel*-constraints of Table 1, which can be represented only in  $\mathcal{MS}$  and  $\mathcal{SD}$  (using only  $\mathcal{DS}$ -constraints we cannot express constraints on the relative size of two regions).

However, when we have metric size constraints, in addition to the size information of Table 1, it is necessary to extract from the topological information some non-binary constraints that we call *containment constraints*. These constraints state that the size of the union of all regions  $x_1, \dots, x_l$  ( $l > 1$ ) contained in a certain other region  $z$  must not exceed the size of  $z$ :

$$size(x_1 \cup \dots \cup x_l) \leq size(z),$$

---

<sup>9</sup> Definite size constraints can be used, for instance, to represent spatial regions corresponding to political or administrative districts, the exact sizes of which are usually known.

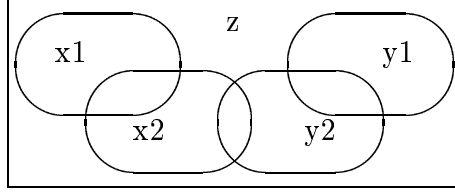


Fig. 5. Two-dimensional examples of topological configuration entailing three containment constraints:  $size(x1) + size(y2) \leq size(z)$ ,  $size(y1) + size(x2) \leq size(z)$  and  $size(x1) + size(y1) \leq size(z)$ .

where  $x_i \{TPP, NTPP\} z$  for  $i = 1, \dots, l$ . In order to make these constraints more accessible, we transform them into a form which uses only size information of named regions instead of size information of unions of named regions, i.e., containment constraints are expressed in the following form:

$$size(y_1) + \dots + size(y_k) \leq size(z),$$

where  $\{y_1, \dots, y_k\}$  is a subset of  $\{x_1, \dots, x_l\}$ , and  $y_i \{TPP, NTPP\} z$  for  $i = 1, \dots, k$ .

It is clear that this transformation heavily depends on the relations between the regions under considerations. If  $x_i \{TPP\} x_j$  or  $x_i \{NTPP\} x_j$  holds for some contained regions  $x_i$  and  $x_j$ , then  $size(x_i \cup x_j) = size(x_j)$ , and  $size(x_i)$  should not be involved in the containment constraint; if  $x_i \{EQ\} x_j$  holds, then only one of  $size(x_i)$  and  $size(x_j)$  should be involved. Note that this variable elimination will not lead us to a loss of information, because of the  $QS$ -constraints that are entailed by the topological relations. For example,  $y_1 \{TPP\} y_2$  and  $size(y_2) + size(w) < size(z)$  imply  $size(y_1) + size(w) < size(z)$ , because  $y_1 \{TPP\} y_2$  implies  $size(y_1) < size(y_2)$ . So, there is no loss of information if  $size(y_1) + size(w) < size(z)$  is not explicitly stated.

The case where two regions partially overlap is more difficult. If  $x_i \{PO\} x_j$  holds, then we have that  $max(size(x_i), size(x_j)) < size(x_i \cup x_j) < size(x_i) + size(x_j)$  holds. The lower bound is approached when the two regions overlap almost completely, while the upper bound is approached when the two regions overlap only at a very small part. Because of partially overlapping regions, it can be necessary to split a single containment constraint on the size of the union of named regions into multiple containment constraints involving only the size of named regions. For example, the following topological scenario involving five regions entails three containment constraints (see Figure 5):

$$\begin{aligned} & \{x1\{TPP\}z, x1\{PO\}x2, x1\{DC\}y1, x1\{DC\}y2, \\ & y1\{TPP\}z, y1\{PO\}y2, y1\{DC\}x1, y1\{DC\}x2, \\ & x2\{TPP\}z, x2\{PO\}x1, x2\{PO\}y2, x2\{DC\}y1, \\ & y2\{TPP\}z, y2\{PO\}y1, y2\{PO\}x2, y2\{DC\}x1\}. \end{aligned}$$

We now give more formal conditions under which a set of topological constraints entails a containment constraint. We restrict our analysis to topological constraints forming a consistent scenario.

A consistent scenario  $\Theta_s$  of RCC-8 constraints entails the  $\leq$ -containment constraint

$$size(x_1) + \dots + size(x_k) \leq size(z)$$

if we have that the regions  $x_1, \dots, x_k$  are contained in  $z$ , and they do not overlap each other, i.e., if the following conditions hold:

- (1)  $x_i Rz \in \Theta_s$  and  $R \in \{\{TPP\}, \{NTPP\}\}$ , for  $i = 1, \dots, k$ ;
- (2)  $x_i Rx_j \in \Theta_s$ ,  $R \in \{\{DC\}, \{EC\}\}$  and  $i \neq j$ , for  $i, j = 1, \dots, k$ .

When conditions (1) and (2) hold, it is clear that  $size(x_1 \cup \dots \cup x_k) = size(x_1) + \dots + size(x_k)$  holds.

Under certain conditions it is not possible that the sum of the sizes of the contained regions is equal to the size of the region containing them, and the corresponding  $\leq$ -containment constraint can be refined to a  $<$ -constraint. For RCC-7 scenarios there are exactly five conditions under which we can make this refinement. Of course, such conditions are sufficient also for RCC-8 scenarios, since RCC-7 scenarios are also RCC-8 scenarios. However, if PO is allowed (i.e., if we use RCC-8 scenarios), there are many additional conditions that imply this refinement. So far we have identified seven of them (given in [29]), but it is still an open problem whether these conditions are the only conditions to refine a  $\leq$ -containment constraint. In the following we specify the five conditions for RCC-7, and we prove that (for RCC-7) these are the only conditions under which we can refine a  $\leq$ -containment constraint. This result will be used in the next section to derive a polynomial algorithm for reasoning about basic relations in RCC-7 combined with metric size constraints, while, as we will show, reasoning with RCC-8 basic relation and metric size constraints is NP-hard.

A consistent scenario  $\Theta_s$  in RCC-7 entails the strict containment constraint

$$size(x_1) + \dots + size(x_k) < size(z)$$

if in addition to (1) and (2), we have that one of the following conditions holds:

- (a)  $x_i \{NTPP\} z \in \Theta_s$ , for  $i = 1, \dots, k$  (Figure 6.a illustrates this case for  $k = 3$ );
- (b) there exists a collection of  $m$  regions  $x_{p_1}$  to  $x_{p_m}$  ( $p_i \in \{1, \dots, k\}$ ,  $1 \leq i \leq m < k$ ) such that for each of the other  $k - m$  regions  $x_{p_j}$  ( $m < j \leq k$ )

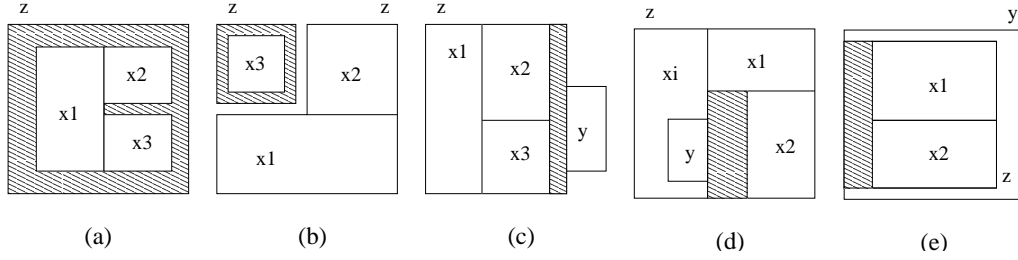


Fig. 6. Two-dimensional examples illustrating containment constraints. Note that all regions are rectangles. The part of  $z$  which cannot be covered by  $x_1, \dots, x_k$  is shaded.

- $x_{p_j}\{\text{DC}\}x_{p_i} \in \Theta_s$  and also  $x_{p_i}\{\text{NTPP}\}z$  (Figure 6.b illustrates this case for  $m = 1, k = 3$  and  $x_{p_1} = x_3$ );
- (c) there exists a spatial region  $y \in \Theta_s$  such that  $y\{\text{EC}\}z \in \Theta_s$ , and  $x_i\{\text{DC}\}y \in \Theta_s$  for  $i = 1, \dots, k$ ; (Figure 6.c illustrates this case for  $k = 3$ );
- (d) there exists a spatial region  $y \in \Theta_s$  such that  $y\{\text{TPP}\}x_i \in \Theta_s$  for some  $i \in \{1, \dots, k\}$ ,  $y\{\text{NTPP}\}z \in \Theta_s$ , and  $x_j\{\text{DC}\}y \in \Theta_s$  for  $j = 1, \dots, i - 1, i + 1, \dots, k$ ; (Figure 6.d illustrates this case for  $k = 3$ );
- (e) there exists a spatial region  $y \in \Theta_s$  such that  $y\{\text{TPP}^{-1}\}z \in \Theta_s$ , and  $x_i\{\text{NTPP}\}y \in \Theta_s$  for  $i = 1, \dots, k$ ; (Figure 6.e illustrates this case for  $k = 2$ );

Under conditions (1) and (2) it is clear that if all regions  $x_i$  are non-tangential proper part of  $z$ , then their union must be smaller than  $z$  (condition a). Suppose that two regions  $u$  and  $v$  are contained in a region  $w$ . Since  $w$  might consist of disconnected pieces, it does not follow that if  $u$  and  $v$  are disconnected, their sum is smaller than  $w$ . Only if, in addition,  $u$  or  $v$  is non-tangential proper part of  $w$ , then their sum must be smaller than  $w$ . This is generalized to multiple regions in condition (b). If there is a collection of the regions  $x_i$  which are all disconnected from the other regions  $x_j$  and which are all non-tangential proper part of the containing region  $z$ , then their sum must be smaller than  $z$ . Note that condition (a) is subsumed by condition (b) if  $m < k$  is relaxed to  $m \leq k$ .

Similarly, conditions (c)–(e) enforce that the regions contained in  $z$  cannot fill it in completely, because there is a part of the boundary of  $z$  that cannot be touched by any of these regions (conditions c and e), or because part of the boundary of a region contained in  $z$  can be touched neither by the boundary of  $z$ , nor by any of the other regions contained in  $z$  (condition d).

Given a containment constraint  $C$  involving a set  $S$  of contained regions, clearly  $C$  implies all the constraints involving a subset of  $S$  and the same containing region, because we assume that the size of each region is larger than zero. E.g.,  $\text{size}(x_1) + \text{size}(x_2) + \text{size}(x_3) \leq \text{size}(z)$  implies  $\text{size}(x_1) + \text{size}(x_2) < \text{size}(z)$ . Hence, it is sufficient to consider only the constraints involving a “maximal” set of contained regions. More formally, for each region  $z$  we will consider the maximal set of *covering regions*, which are defined as follows (for



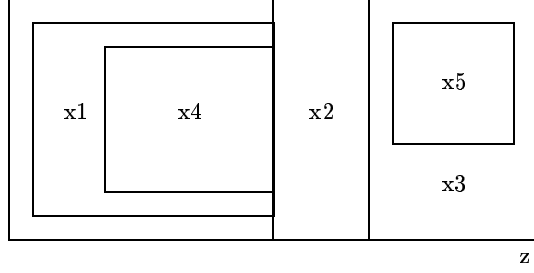


Fig. 7. Example of a set of covering regions for  $z$  ( $\{x_1, x_2, x_3\}$ ).  $x_5$  ( $x_4$ ) does not belong to this set because it is contained in  $x_3$  ( $x_1$ ) which is disconnected or externally connected with the other regions contained in  $z$ .

an example see Figure 7).

**Definition 17 (Set of Covering Regions)** Let  $\Theta_s$  be a consistent scenario and  $Q = \{x_1, \dots, x_k, z\}$  a set of variables in  $\Theta_s$  ( $k > 1$ ) satisfying conditions (1) and (2). We say that  $X = Q - \{z\}$  is a covering set of regions for  $z$  if and only if in  $\Theta_s$  there exists no region  $y$  which topologically contains a subset  $X' = \{x'_1, \dots, x'_h\}$  of the regions in  $X$ , and which is disconnected or externally connected with every  $x_j \in X - X'$  ( $1 \leq h < k$ ).

**Lemma 18** Let  $\Theta_s$  be a consistent scenario over RCC-7 and  $\{x_1, \dots, x_k\}$  a maximal set of covering regions for  $z$  in  $\Theta_s$  ( $k \geq 2$ ).  $\Theta_s$  entails

$$\text{size}(x_1) + \dots + \text{size}(x_k) \leq \text{size}(z)$$

if and only if (1) and (2) hold; while  $\Theta_s$  entails

$$\text{size}(x_1) + \dots + \text{size}(x_k) < \text{size}(z)$$

if and only if, in addition to (1) and (2), one of (a)–(e) holds.

**Proof.** The *if* direction of the claim is obvious, as illustrated by the five examples of Figure 6. Regarding the *only-if* direction, clearly (1) and (2) are necessary conditions, because we are assuming that the regions involved in a containment constraint cannot partially overlap, and because if (2) did not hold, then the sum of the sizes of the  $x_1, \dots, x_k$  may be larger than the size of  $z$ .<sup>10</sup> Moreover, in order to have a strict containment constraint (“<”), it is necessary that one of (a)–(e) holds. We prove this by induction over the number  $k$  of regions  $x_1, \dots, x_k$  contained in  $z$  and satisfying conditions (1) and (2).

*Induction base:*  $k = 2$ . This case can be proved by enumerating the six possible

<sup>10</sup> For example, if  $k = 2$  and  $x_1 \{EQ\} x_2$ , then the size of each single region  $x_1$  and  $x_2$  can be larger than half the size of  $z$ .

cases of  $x_1, x_2$  being either TPP or NTPP of  $z$ , and either DC or EC to each other. In all the other cases no containment constraint is entailed. From these constraints,  $size(x_1) + size(x_2) < size(x)$  is entailed only if either both  $x_1$  and  $x_2$  are non-tangential proper part of  $z$  (i.e., condition a), or if  $x_1$  and  $x_2$  are disconnected and  $x_1$  or  $x_2$  is non-tangential proper part of  $z$  (i.e., condition b). In addition to these cases,  $size(x_1) + size(x_2) < size(x)$  can be entailed when other constraints in  $\Theta_s$  enforce that a part of  $z$  cannot be covered by  $x_1$  and  $x_2$ . Since partial overlapping is not possible, and since  $\{x_1, x_2\}$  is a maximal set of covering regions for  $z$ , this can be enforced only by the presence of regions which are either part of  $x_1$  or  $x_2$ , or which are DC, EC,  $TPP^{-1}$  or  $NTPP^{-1}$  to  $z$ . If a region  $y$  is EC to  $z$  and DC to  $x_1$  and  $x_2$ , then the boundary of  $z$  where  $y$  connects to  $z$  cannot be covered by  $x_1$  or by  $x_2$  (condition c). If a region  $y$  is  $TPP^{-1}$  to  $z$  and  $NTPP^{-1}$  to  $x_1$  and  $x_2$ , then the boundary of  $z$  where  $y$  connects to  $z$  cannot be covered by  $x_1$  or by  $x_2$  (condition e). In all the other cases, the regions outside  $z$  cannot influence the containment constraints for  $z$ . If a region  $y$  is TPP of  $x_1$  ( $x_2$ ), DC to  $x_2$  ( $x_1$ ), and NTPP of  $z$ , then the boundary of  $x_1$  ( $x_2$ ) where  $y$  connects is inside  $z$  and cannot be connected to  $x_2$  ( $x_1$ ) (condition d). In no other case regions inside  $x_1$  or inside  $x_2$  can prevent  $x_1$  and  $x_2$  from completely covering  $z$  (because it is not possible to enforce that the regions contained in a region  $x$  completely cover  $x$ ).

*Induction assumption:*  $\Theta_s$  entails  $size(x_1) + \dots + size(x_k) < size(z)$  for  $k = n$  ( $n \geq 2$ ) only if, in addition to conditions (1) and (2), condition (a) or (b) holds.

*Induction step:*  $k = n + 1$  and  $\Theta_s$  entails  $size(x_1) + size(x_2) + \dots + size(x_k) < size(z)$ . If the sum of the  $n + 1$  regions must be smaller than the size of  $z$ , then a gap in  $z$  which can not be filled in by any of  $x_1, \dots, x_k$  must be forced. There are two cases to consider: (i)  $x_k \{NTPP\} z \in \Theta_s$  and (ii)  $x_k \{TPP\} z \in \Theta_s$  (one of them must hold by (1)).

Regarding (i), we have that, if  $x_k$  is disconnected from all the other  $n$  regions contained in  $z$ , then condition (b) must hold ( $x_k$  is a collection of regions satisfying the relations required by (b)). Otherwise, let  $x_i$  be a region that is externally connected with  $x_k$ , and  $x_{ik}$  the region corresponding to the union of  $x_k$  and  $x_i$ . By performing the following transformation to  $\Theta_s$  we can construct a set  $\Theta'_s$  of constraints such that  $\Theta_s$  entails  $size(x_1) + size(x_2) + \dots + size(x_k) < size(z)$  if and only if  $\Theta'_s$  entails

$$size(x_1) + \dots + size(x_{i-1}) + size(x_{i+1}) + \dots + size(x_{k-1}) + size(x_{ik}) < size(z).$$

- Replace  $x_k$  and  $x_i$  with  $x_{ik}$ ;
- remove each region  $y$  such that in  $\Theta_s$  either  $y \{NTPP\} x_i, y \{NTPP\} x_k, y \{EQ\} x_i, y \{EQ\} x_k, y \{TPP\} z$  and  $y \{TPP\} x_k$ , or  $y \{TPP\} z$  and  $y \{TPP\} x_i$ ;
- replace each pair of constraints  $x_k R_{kj} x_j$  and  $x_i R_{ij} x_j$  with  $x_{ik} \{DC\} x_j$ , if both

$R_{kj}$  and  $R_{ij}$  are equal to DC in  $\Theta_s$ , with  $x_{ik}\{\text{EC}\}x_j$ , if one of  $R_{kj}$  and  $R_{ij}$  is EC;

- replace  $x_iR_{iz}z$  and  $x_kR_{kz}z$  with  $x_{ik}\{\text{NTPP}\}z$  if both  $R_{iz}$  and  $R_{kz}$  are NTPP, with  $x_{ik}\{\text{TPP}\}z$ , if one of  $R_{iz}$  and  $R_{kz}$  is TPP;
- for each region  $y$  such that in  $\Theta_s$   $y\{\text{EC}\}z$  or  $z\{\text{TPP}\}y$  or  $z\{\text{NTPP}\}y$  holds, replace  $x_kR_{ky}y$  and  $x_iR_{iy}y$  with  $x_{ik}\{\text{EC}\}y$ , if  $R_{iy}$  or  $R_{ki}$  is EC, with  $x_{ik}\{\text{TPP}\}y$  if  $R_{iy}$  or  $R_{ky}$  is TPP, with  $x_{ik}\{\text{NTPP}\}y$  if both  $R_{iy}$  and  $R_{ky}$  are NTPP;
- for each region  $y$  such that  $y\{\text{NTPP}\}z \in \Theta_s$  and  $y\{\text{TPP}\}x_i \in \Theta_s$  (or  $y\{\text{TPP}\}x_k \in \Theta_s$ ), if  $y\{\text{EC}\}x_i \in \Theta_s$  ( $y\{\text{EC}\}x_k \in \Theta_s$ ) and  $y\{\text{DC}\}x_j \in \Theta_s$  for  $j = 1, \dots, k$ ,  $j \neq i$  and  $j \neq k$ , then replace  $y\{\text{TPP}\}x_i$  ( $y\{\text{TPP}\}x_k$ ) with  $y\{\text{NTPP}\}x_{ik}$ , otherwise replace it with  $y\{\text{TPP}\}x_{ik}$ .

The last step of the transformation is used to avoid that none of the conditions (a)–(e) holds in  $\Theta_s$  for  $\{x_1, \dots, x_k\}$ , while condition (d) holds in  $\Theta'_s$  for  $\{x_1, \dots, x_k, x_{ik}\} \setminus \{x_i, x_k\}$ .<sup>11</sup> It is easy to see that  $\Theta_s$  entails that the sum of the sizes of  $x_1, \dots, x_k$  is less than the size of  $z$  if and only if  $\Theta'_s$  entails the sum of the sizes of  $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{k-1}, x_{ik}$  is less than the size of  $z$ . Suppose that  $\Theta'_s$  entails this constraint. By the induction assumption we have that one of conditions (a)–(e) holds in  $\Theta'_s$ . If condition (a) holds, then (a) must hold also for the original  $n + 1$  regions of  $\Theta_s$  that are contained in  $z$ . If condition (b) holds, then there exists a collection of regions of  $\Theta'_s$  contained in  $z$  that are disconnected from all the other regions, and either this collection is formed by just  $x_{ik}$ , or it is formed by a set of regions that belongs to both  $\Theta'_s$  and  $\Theta_s$ . In both cases it is easy to see that also in  $\Theta_s$  there must be a collection of regions, each of which are non-tangential proper part of  $z$ , which are disconnected from all the other regions contained in  $z$ . Hence (b) must hold in  $\Theta_s$  as well. Similarly, by construction of  $\Theta'_s$ , if one of (c), (d) or (e) holds in  $\Theta'_s$ , then it must hold in  $\Theta_s$  as well.

Regarding case (ii), if  $x_k$  is externally connected with some region  $x_i$  contained in  $z$ , then we can prove that one of (a)–(e) must hold using the same argument used for case (i). If  $x_k$  is disconnected from  $x_1, \dots, x_{k-1}$ , then if two of these regions are externally connected, we can again prove that one of (a)–(e) must hold by considering their union, and using an argument similar to the argument used for case (i). Otherwise ( $x_1, \dots, x_k$  are all pairwise disconnected and  $x_k$  is TPP of  $z$ ), we have two cases to consider:

- there exists a region  $x_i$  such that  $x_i\{\text{NTPP}\}z \in \Theta_s$  and  $i \in \{1, \dots, k - 1\}$ ;
- there exists no such region.

If the first case holds, then (b) is satisfied. If the second case held and neither (c) nor (e) were satisfied, then, contrary to our assumption,  $\Theta_s$  would not entail  $size(x_1) + \dots + size(x_k) < size(z)$ , because we could assign to  $x_1, \dots, x_k$

<sup>11</sup> A longer version of the proof containing further details and an example of the transformation is given in a technical report [29].

regions which fill in  $z$  completely.<sup>12</sup> Hence also in this case the claim holds, because otherwise the induction hypothesis would be contradicted. ■

We now analyze the topological constraints that are entailed by the metric size information. These can be defined in terms of the entailed  $\mathcal{QS}$ -constraints. More precisely, let  $S$  be a  $\mathcal{QS}$ -relation,  $s$  is a basic relation in  $S$ , and  $C(x, y)$  a constraint between  $x$  and  $y$  in  $\mathcal{MS}$  or  $\mathcal{SD}$ . We have that the strongest topological constraint entailed by  $C(x, y)$ ,  $Toprel(C(x, y))$ , is defined as

$$Toprel(C(x, y)) = \{Toprel(s) \mid s \text{ is a } \mathcal{QS} \text{ basic relation and } C(x, y) \wedge xsy \text{ is consistent}\}.$$

For example, if  $C(x, y)$  is  $size(y) - size(x) \in [0, 3]$ , then  $Toprel(C(x, y)) = Toprel(<) \cup Toprel(=)$ .

$\mathcal{DS}$ -constraints can entail topological constraints only in combination with other  $\mathcal{DS}$ -constraints. For example,  $size(x) \in [1, 3]$  and  $size(y) \in [4, 5]$  entails  $size(x) \leq size(y)$ , and hence  $x\{\text{DC, EC, PO, TPP, NTPP}\}y$  (see Table 1). In our computational analysis of  $\mathcal{DS}$  we will exploit the fact that a set  $\Sigma$  of  $\mathcal{DS}$ -constraints can be translated in polynomial time into an equivalent set of  $\mathcal{SD}$ -constraints in the following way:

1. we introduce a new special region  $s$  which we assume having size zero, and which is disconnected with all the other regions involved in  $\Sigma$ ;<sup>13</sup>
2. we translate each  $\mathcal{DS}$ -constraint into a pair of  $\mathcal{SD}$ -constraints involving  $s$ , e.g.,  $size(x) \in [a, b]$  is translated into  $size(x) - size(s) \in [a, b]$ .

Clearly the translated set of constraints is consistent if and only if the original set of constraints is consistent.

To conclude this section, we define a model for a set of metric size constraints, and then give a formal definition of consistency for a set of topological constraints combined with a set of metric size constraints. Note that every model of a set of RCC-8-constraints satisfies all the implicit containment constraints.

**Definition 19 (Model for  $\mathcal{MS}$ ,  $\mathcal{SD}$ ,  $\mathcal{DS}$ )** *Given a set  $\Sigma$  of constraints in either  $\mathcal{MS}$ ,  $\mathcal{SD}$  or  $\mathcal{DS}$  we say that an assignment  $\sigma$  of spatial regions to the*

<sup>12</sup> We have that  $x_1, \dots, x_k$  are pairwise disconnected and that each of them is tangential proper part of  $z$ . Under these conditions clearly neither (a) nor (b) nor (d) hold. Thus, if neither (c) nor (e) hold, then we can interpret  $z$  as a region formed by  $k$  pieces, each of which is completely filled in by exactly one of  $x_1, \dots, x_k$ .

<sup>13</sup> Note that this is a particular region that can not be involved by the input constraints; according to RCC all the regions of these constraints must be non-empty regions. The use of this special region is inspired from a similar idea in the context of temporal reasoning [18].

variables of  $\Sigma$  is a model of  $\Sigma$  if and only if  $\sigma$  satisfies all the constraints in  $\Sigma$ .

**Definition 20 (Consistency for RCC-8 and  $\mathcal{MS}$ ,  $\mathcal{SD}$ ,  $\mathcal{DS}$ )** Given a set  $\Theta$  of constraints in RCC-8 and a set  $\Sigma$  of constraints in either  $\mathcal{MS}$ ,  $\mathcal{SD}$ , or  $\mathcal{DS}$   $\Theta \cup \Sigma$  is consistent if there exists a model of  $\Theta$  which is also a model of  $\Sigma$ .

## 6 Reasoning about Topological and Metric Size Constraints

In this section we analyze the computational complexity of RSAT for a set of topological constraints combined with metric size constraints. First we consider RCC-8, and we show that RSAT for topological scenarios combined with metric size constraints is intractable in general. Then we show that, when the input topological constraints form scenarios which do not involve some particular relations, RSAT can be solved in polynomial time.

Finally, we focus on the relations in RCC-7. As we have discussed in the introduction, RCC-7 is useful in domains where spatial regions cannot partially overlap, and in Section 7 we will give a worked example in one of such domains. Our results for RCC-7 include a polynomial algorithm for the set of basic relations, and a backtracking algorithm for the full calculus.

### 6.1 RCC-8 and Metric Size Constraints

Before giving our complexity results, we need to state a lemma which will be used in the proofs of the next theorems.

**Lemma 21** *Given any set  $\Theta$  of topological constraints in  $\{\{\text{DC}\}, \{\ast\}\}$ , by refining every universal relation to  $\{\text{PO}\}$  we obtain a consistent set of constraints.*

**Proof.** From the composition table of the basic relations of RCC-8 [5] it follows that any set of constraints over  $\{\{\text{DC}\}, \{\text{PO}\}\}$  with exactly one constraint for every pair of variables is path-consistent. Since path-consistency is sufficient for deciding consistency of sets of constraints over the RCC-8 basic relations, it follows that any set of constraints over  $\{\{\text{DC}\}, \{\text{PO}\}\}$  with exactly one constraint for every pair of variables is consistent. ■

In the following we indicate with co-RSAT the complementary problem of RSAT (i.e, determining whether a given set of constraints between spatial regions is *inconsistent*), and with  $\mathcal{B}$  the set of the eight basic relations of

RCC-8.

**Theorem 22** *co-RSAT for  $\mathcal{B} \cup \mathcal{MS}$  is NP-hard.*

**Proof.** We show that CLIQUE [25] is polynomially reducible to co-RSAT for  $\mathcal{B} \cup \mathcal{MS}$ . Let  $G = (V, E)$  be a graph, and  $k$  a positive integer such that  $k \leq |V|$ . From  $G$  we construct a set  $C$  of spatial constraints that is inconsistent if and only if  $G$  contains a clique of size  $k$  or more. The constraints of  $C$  are the following:

- (1) for each arc  $(x_1, x_2) \in E$ ,  $x_1\{\text{DC}\}x_2 \in C$ ;
- (2) for each  $x'_1, x'_2$  such that  $x'_1 \neq x'_2$  and  $(x'_1, x'_2) \notin E$ ,  $x'_1\{\text{PO}\}x'_2 \in C$ ;
- (3) for each  $x \in V$ ,  $x\{\text{NTPP}\}z \in C$ , where  $z$  is a new variable that is not present in  $V$ ;
- (4) for each  $x, y \in V$ ,  $(\text{size}(x) = \text{size}(y)) \in C$ ;
- (5)  $(\text{size}(z) = k \cdot \text{size}(x)) \in C$ , where  $x$  is an arbitrary element of  $V$ .

It is clear that the subset  $C_s$  of size constraints in  $C$  is independently consistent, while the subset  $C_t$  of topological constraints is independently consistent by Lemma 21, and the fact that for any set of regions there exists another region that contains all the regions of the set. So any model satisfying the constraints of (1), (2) can be extended to derive a model satisfying the constraints of (3) as well. Moreover, we have that

- (a) the strongest  $\mathcal{MS}$ -constraints entailed by  $C_t$  are  $\text{size}(x) < \text{size}(z)$ , for each  $x \in V$ , which are already entailed by the constraints in  $C_s$ . No  $\mathcal{MS}$ -constraint between variables in  $V$  is entailed by  $C_t$ ;
- (b) by extending  $C_t$  with the topological constraints entailed by  $C_s$ , we cannot refine any relation in  $C_t$ , which remains consistent. In fact, the strongest RCC-8 relations entailed by  $C_s$  are  $x\{\text{DC}, \text{PO}, \text{EC}, \text{EQ}\}y$ , for each  $x, y \in V$  ( $x \neq y$ ), and  $x\{\text{DC}, \text{EC}, \text{PO}, \text{TPP}, \text{NTPP}\}z$ , for each  $x \in V$ .

It follows that  $C$  is inconsistent if and only if  $C_t$  entails some containment constraint imposing that the sum of the size of  $k$  or more regions is smaller than the size of  $z$  (e.g.,  $\text{size}(x_i) + \text{size}(x_j) + \text{size}(x_k) < \text{size}(z)$ , for some  $x_i, x_j, x_k$  in  $V$  and  $k = 3$ ). Such constraint is implicit in  $C_t$  if and only if in  $C_t$  we have  $k$  or more variables that are constrained to be disconnected with each other. The *if* direction of this claim is obvious, while the *only if* direction follows from the following fact. If two regions  $x$  and  $y$  partially overlap, then we can interpret  $x$  and  $y$  in such a way that the size of their non-overlapping parts is an arbitrary small quantity  $\epsilon$ , i.e.,

$$\text{size}(x \cup y) = \max(\text{size}(x), \text{size}(y)) + \epsilon.$$

By construction of  $C_t$ , we have that  $C_t$  entails a containment constraint, imposing that the sum of the size of  $k$  or more regions is smaller than the size of  $z$ , if and only if  $G$  contains a clique of  $k$  or more vertices. ■

Note that in the proof of the previous theorem it was not necessary to explicitly derive containment constraints, and that we can use a polynomial reduction similar to the one we have used in this proof to prove that deciding whether there is an entailed containment constraint is NP-hard.<sup>14</sup> Hence the following claim is a Corollary of Theorem 22.

**Corollary 23** *Let  $\Theta_s$  a consistent scenario over RCC-8. Deciding whether there exists a set of regions  $\{x_1, \dots, x_k, z\}$  ( $k > 1$ ) such that  $\Theta_s$  entails*

$$size(x_1) + \dots + size(x_k) \leq size(z)$$

*is NP-hard.*

We now consider extending a set of topological constraints in  $\mathcal{B}$  with the class  $\mathcal{DS}$  of domain size constraints. The next theorem states that also for this class of spatial constraints co-RSAT is intractable.

**Theorem 24** *co-RSAT for  $\mathcal{B} \cup \mathcal{DS}$  is NP-hard.*

**Proof Sketch.** The proof is very similar to the proof of NP-hardness for  $\mathcal{BUMS}$  (Theorem 22). We show that CLIQUE is polynomially reducible to co-RSAT for  $\mathcal{B} \cup \mathcal{DS}$ . Let  $G = (V, E)$  be a graph, and  $k$  a positive integer such that  $k \leq |V|$ . From  $G$  we construct a set  $S$  of spatial constraints that is inconsistent if and only if  $G$  contains a clique of size  $k$  or more. The constraints of  $S$  are the same as the constraints of  $C$  in the proof of Theorem 22, except constraints (4) and (5) which are the following: for each  $x \in V$ ,  $(size(x) \in [1, 1]) \in S$ ;  $(size(z) \in [k, k]) \in S$ . □

We now examine metric information that can be expressed using the class of constraints  $\mathcal{SD}$ . Despite this class is tractable, its combination with the topological constraints over the simple class of the basic relations of RCC-8 leads again to intractability.<sup>15</sup>

**Theorem 25** *co-RSAT for  $\mathcal{B} \cup \mathcal{SD}$  is NP-hard.*

<sup>14</sup>The set  $C_t$  of topological constraints entails  $size(x_1) + \dots + size(x_k) \leq size(z)$  if and only if the graph  $G$  contains a clique of size  $k$ .

<sup>15</sup>Deciding the satisfiability of a set of constraints over  $\mathcal{SD}$  can be accomplished by using an all-pairs shortest-paths algorithm [18], or, more efficiently, using a one-source shortest-paths algorithm [26].

**Proof.** NP-hardness follows from Theorem 24, and the fact that, as we have previously described, any set of  $\mathcal{DS}$ -constraints can easily be translated into an equivalent set of  $\mathcal{SD}$ -constraints. ■

From the previous theorems it follows that RSAT for  $\mathcal{B} \cup \mathcal{MS}$ ,  $\mathcal{B} \cup \mathcal{DS}$ , and  $\mathcal{B} \cup \mathcal{SD}$ , is NP-hard.

**Theorem 26** *RSAT for  $\mathcal{B} \cup \mathcal{MS}$ ,  $\mathcal{B} \cup \mathcal{DS}$  and  $\mathcal{B} \cup \mathcal{SD}$  are NP-hard.*

**Proof.** It follows from Theorems 22, 24 and 25, and the fact that there exists a (trivial) polynomial time Turing reduction from any decision problem to its complement and vice-versa [25]. ■

The previous NP-hardness results impose a severe limit to the practical applicability of RCC-8 and metric size constraints. In fact, a correct and complete algorithm for consistency checking would require exponential cost even when only the basic relations of RCC-8 are used.

Since the source of this complexity are the (non-binary) containment constraints, we can derive a tractable class by excluding those topological relations that give rise to containment constraints. In particular, given a set  $\Theta$  of topological constraints over  $\mathcal{B} \setminus \{\{\text{DC}\}, \{\text{EC}\}\}$  or  $\mathcal{B} \setminus \{\{\text{TPP}\}, \{\text{NTPP}\}, \{\text{TPP}^{-1}\}, \{\text{NTPP}^{-1}\}\}$ , and a set  $\Sigma$  of size constraints, we can check the consistency of  $\Theta \cup \Sigma$  in polynomial time by (1) checking whether  $\Theta$  is consistent, (2) adding to  $\Sigma$  the strongest  $\mathcal{QS}$ -constraints that are entailed by  $\Theta$ , and (3) checking whether  $\Sigma$  is consistent. Completeness for this simple algorithm relies on the fact that once we have propagated the strongest entailed  $\mathcal{QS}$ -constraints (Step 2), there is no need to propagate metric size constraints to  $\Theta$  (we will give details of this in the proof of the main theorem of the next section). Thus we can state the following theorems.

**Theorem 27** *RSAT for  $\{\{\text{TPP}\}, \{\text{NTPP}\}, \{\text{TPP}^{-1}\}, \{\text{NTPP}^{-1}\}, \{\text{PO}\}, \{\text{EQ}\}\} \cup \mathcal{MS} \cup \mathcal{DS} \cup \mathcal{SD}$  is polynomial.*

**Theorem 28** *RSAT for  $\{\{\text{DC}\}, \{\text{EC}\}, \{\text{PO}\}, \{\text{EQ}\}\} \cup \mathcal{MS} \cup \mathcal{DS} \cup \mathcal{SD}$  is polynomial.*

From a topological point of view, the expressiveness of these tractable classes is quite limited, since they cannot express either that one region is contained in another region, or that two regions are disconnected or externally connected, which can be the case in many domains. However, the combined class of Theorem 27 could be useful for applications involving size constraints and topological containment reasoning, while the class of Theorem 28 could be useful for domains where the internal structure of the spatial regions is not



important.<sup>16</sup>

In the next section we will examine another class of topological relations, the basic relations of RCC-7, which can be combined with metric size constraints without losing tractability. The polynomial decision procedure for this class will be used to derive a general backtracking algorithm for the combination of RCC-7 and metric size constraints.

## 6.2 RCC-7 and Metric Size Constraints

It is interesting to observe that the NP-hardness proofs of Theorems 22, 24 and 25 do not work if PO is omitted from the topological class of relations, i.e., if we consider the set  $\mathcal{B}_7$  of RCC-7 basic relations. The following lemma will be used in the proof of the next theorem stating that RSAT for RCC-7 basic relations and metric size constraints can be solved in polynomial time.<sup>17</sup>

**Lemma 29** *Let  $\Theta$  be a consistent scenario of RCC-8 relations,  $\Delta$  a set of containment constraints entailed by  $\Theta$ , and  $\Sigma$  a set of size constraints in  $\mathcal{MS}$ ,  $\mathcal{DS}$ , or  $\mathcal{SD}$ . The consistency of  $\Sigma \cup \Delta$  can be determined in polynomial time.*

**Proof.** It follows from the fact that the constraints of  $\Sigma \cup \Delta$  are a special case of Lassez and McAloon's *Generalized Linear Constraints* for which satisfiability can be solved in polynomial time [44]. ■

**Theorem 30** *RSAT for  $\mathcal{B}_7 \cup \mathcal{MS}$  is polynomial.*

**Proof.** Let  $\Theta$  be the input set of topological constraints forming a scenario of RCC-7 relations, and  $\Sigma$  the input set of metric size constraints. We show that the algorithm of Figure 8 decides the consistency of RSAT for  $\Theta \cup \Sigma$  in polynomial time.

If  $\Theta$  is inconsistent then  $\Theta \cup \Sigma$  is clearly inconsistent (Step 1). Step 2 adds to  $\Sigma$  the qualitative size constraints that are entailed by  $\Theta$  using Table 1. Step 3 identifies, for each region  $x \in \Theta$ , the set  $IN_x$  of the regions that are contained in  $x$ . After Step 4,  $IN_x$  contains only those regions that are contained in  $x$ , and that are either disconnected or externally connected with each other. The constraints of  $\Theta$  involving these variables entail a containment constraint that is added to  $\Sigma$  by Step 5. By Lemma 18 and construction of the  $IN$ -sets (Steps 3 and 4), the constraints added by Step 5 to  $\Sigma$  are all the strongest (non-binary)

---

<sup>16</sup> For instance, consider a network of roads where two roads can be externally connected, disconnected, or they can partially overlap each other, but a road does not contain another road.

<sup>17</sup> Note that here we are mainly interested in proving tractability – more efficient specialized algorithms for solving these problems could be designed.

*Algorithm: B7\_METRICSIZE*

*Input:* A scenario  $\Theta$  of RCC-7 relations and a set  $\Sigma$  of metric size constraints;

*Output:* *True* if  $\Theta \cup \Sigma$  is consistent, *fail* otherwise.

1. Check whether  $\Theta$  is consistent; if  $\Theta$  is inconsistent, then return “fail”;
2. Add to  $\Sigma$  the strongest  $\mathcal{QS}$ -constraints that are entailed by  $\Theta$ ;
3. For each region  $x$  in  $\Theta$  compute the set  $IN_x$  of the regions in  $\Theta$  that are contained in  $x$ , i.e.,

$$IN_x = \{y \mid y\{\text{TPP}\}x \in \Theta \vee y\{\text{NTPP}\}x \in \Theta\};$$

4. For each  $i \in IN_x$ , if there exists  $j \in IN_x$  such that

$$iRj \in \Theta \wedge R \cap \{\text{EQ}, \text{TPP}, \text{NTPP}\} \neq \emptyset,$$

then remove  $i$  from  $IN_x$ ;

5. For each  $x \in \Theta$  such that  $IN_x = \{y_1, \dots, y_{|IN_x|}\}$  ( $|IN_x| > 1$ ), if one of the following two cases hold,
  - (a) for each  $y \in IN_x$ ,  $y\{\text{NTPP}\}x \in \Theta$ , or
  - (b) there exists a collection of regions  $p_1, \dots, p_m \in IN_x$  ( $1 \leq m < |IN_x|$ ) such that  $p_j\{\text{DC}\}p_i \in \Theta$  and  $p_j\{\text{NTPP}\}x \in \Theta$  for each pair  $p_j, p_i \in IN_x$  ( $1 \leq i \leq m < j \leq |IN_x|$ ), or
  - (c) There exists a region  $w$  such that  $w\{\text{EC}\}x \in \Theta_s$ , and for  $i = 1, \dots, |IN_x|$   $y_i\{\text{DC}\}w \in \Theta_s$ ;
  - (d) there exists a region  $w \in \Theta_s$  such that  $w\{\text{TPP}\}y_i \in \Theta_s$  for some  $i \in \{1, \dots, |IN_x|\}$ ,  $w\{\text{NTPP}\}x \in \Theta_s$ , and  $y_j\{\text{DC}\}w \in \Theta_s$  for  $j = 1, \dots, i-1, i+1, \dots, |IN_x|$ ;
  - (e) there exists a region  $w \in \Theta_s$  such that  $w\{\text{TPP}^{-1}\}x \in \Theta_s$ , and for  $i = 1, \dots, |IN_x|$   $y_i\{\text{NTPP}\}w \in \Theta_s$ ,
then add  $\text{size}(y_1) + \text{size}(y_2) + \dots + \text{size}(y_{|IN_x|}) < \text{size}(x)$  to  $\Sigma$ ,  
otherwise add  $\text{size}(y_1) + \text{size}(y_2) + \dots + \text{size}(y_{|IN_x|}) \leq \text{size}(x)$ ;
6. If  $\Sigma$  is consistent then return “true” otherwise return “fail”.

Fig. 8. Algorithm used in the proof of Theorem 30 for deciding the consistency of a scenario of RCC-7 relations and a set of metric size constraints.  $|IN_x|$  denotes the size of  $IN_x$ .

containment constraints that are entailed by  $\Theta$ . Steps 5a–5e correspond to checking conditions (a)–(e) of Section 5 that are illustrated in Figure 6. Clearly Steps 5a, 5c, 5d and 5e can be accomplished in polynomial time. Regarding Step 5b, since all the regions of  $IN_x$  are either disconnected or externally connected, it corresponds to computing the strongly connected components of the graph  $G = (V, E)$  with  $V = IN_x$ ,  $E = \{(y_i, y_j) \mid y_i\{\text{EC}\}y_j\}$ , and checking whether there is a strongly connected component where all the nodes (regions) are non-tangential proper part of  $x$ . This can be computed in time  $O(V + E)$

[15, p.488f].<sup>18</sup> Finally, Step 6 checks the consistency of  $\Sigma$  extended with the strongest size constraints entailed by  $\Theta$ .

Note that, once we have propagated from  $\Theta$  to  $\Sigma$  the strongest entailed  $\mathcal{QS}$ -constraints (Step 2), it is not necessary to propagate topological information from  $\Sigma$  to  $\Theta$ , because the relation of every constraint in  $\Theta$  is atomic. If we propagated topological relation from  $\Sigma$  to  $\Theta$ , then any topological relation  $r$  could only be refined to the empty relation. But if this were the case, then  $\Sigma$  would have to entail  $\neg r$ , which is impossible if  $\Sigma$  is consistent (because  $Sizerel(r)$  is one of “<”, “>”, “=” or “?”, and  $Sizerel(r) \in \Sigma$  by Step 2). It follows that if the extended set  $\Sigma'$  of size constraints computed by Steps 2 and 5 is consistent, then  $\Theta \cup \Sigma$  is consistent. In fact, since  $\Theta$  is consistent, if the set of the containment constraints and  $\mathcal{QS}$ -constraints entailed by  $\Theta$  are satisfiable together with the input size constraints, then there exists a model for  $\Theta$  that is also a model for  $\Sigma$ . This model can be constructed in the following way. Let  $\sigma$  be a model for  $\Sigma'$ ,  $\sigma_x$  the size of region  $x$  according to  $\sigma$ , and  $Q$  the set of  $\mathcal{QS}$ -constraints induced by  $\sigma$ , where a constraint  $xRy$  is induced by  $\sigma$  if and only if  $x$  and  $y$  are interpreted by  $\sigma$  as regions such that  $\sigma_x R \sigma_y$ . We can construct a model for  $\Theta$  where the size of each region  $x$  is equal to  $\sigma_x$ . Suppose that this were not the case, we would have that either some containment constraint cannot be satisfied, or there exists a pair of regions  $u, v$  such that  $uRv \in Q$  and  $\Theta \models \neg uRv$ . This cannot be the case, because  $\sigma$  satisfies all the size and containment constraints entailed by  $\Theta$ , which are propagated to  $\Sigma$  by Steps 2 and 5. It is easy to see that if the extended set  $\Sigma'$  of size constraints is inconsistent, then clearly also  $\Theta \cup \Sigma$  is inconsistent.

Regarding the complexity of the algorithm, it is sufficient to observe that (1) each of the Steps 1–5 can be accomplished in polynomial time with respect to the number  $n$  of the spatial regions in  $\Theta$ , (2) the number of size constraints added to  $\Sigma$  by Steps 2 and 5 is polynomial with respect to  $n$ , and (3) by Lemma 29 the consistency check of Step 6 can be accomplished in polynomial time. ■

The proof of the following theorem can be easily derived by observing that a set of  $\mathcal{SD}$ -constraints can be easily translated into an equivalent set of linear inequalities, and that, as we have previously described, a set of  $\mathcal{DS}$ -constraints can be translated into a particular set of  $\mathcal{SD}$ -constraints. It follows that the same algorithm given in the proof of Theorem 30 can be used for solving RSAT for  $\mathcal{B}_7 \cup \mathcal{MS} \cup \mathcal{DS} \cup \mathcal{SD}$ .

**Theorem 31** *RSAT is polynomial for  $\mathcal{B}_7 \cup \mathcal{MS} \cup \mathcal{DS} \cup \mathcal{SD}$ .*

<sup>18</sup>Note that if in  $G$  we have an edge  $(y_i, y_j)$ , then we also have the edge  $(y_j, y_i)$ . The case in which  $G$  has only one SCC is a special case of condition 5a; if  $G$  has more than one SCC, then all the regions involved in a SCC are disconnected from every other region, and condition 5b is satisfied.

*Algorithm:* RCC-7\_METRICSIZE

*Input:* A set  $\Theta$  of RCC-7 constraints and a set  $\Sigma$  of metric size constraints;

*Output:* *True* if  $\Theta \cup \Sigma$  is consistent, *fail* otherwise.

1. Enforce path-consistency to  $\Theta$ ; if the resulting set of topological constraints contains the empty relation, then return *fail*;
2. Check the consistency of  $\Sigma$ ; if  $\Sigma$  is inconsistent, then return *fail*;
3. Generate a consistent scenario  $\Theta_s$  of  $\Theta$  by using backtracking; if there is no consistent scenario, then return *fail*;
4. Add to  $\Sigma$  the strongest *QS*-constraints that are entailed by  $\Theta_s$  and check whether the resulting set  $\Sigma'$  is consistent; if  $\Sigma'$  is inconsistent, then Goto 3;
5. Propagate containment constraints from  $\Theta_s$  to  $\Sigma'$ , obtaining a set  $\Sigma''$  of constraints;
6. Check whether  $\Sigma''$  is consistent; if  $\Sigma''$  is consistent, then return *true*, otherwise Goto 3.

Fig. 9. High-level description of a backtracking algorithm for deciding the consistency of a set of topological and size constraints

*Remark.* The definite size constraints that we have introduced in the first part of Section 5 are computationally easier than the general *DS*-constraints. In particular, if a definite size constraint is specified for each region, then there is no need to propagate the size constraints that are entailed at the topological level. Each of these constraints can be independently tested using the known size of the regions. We can obtain an efficient algorithm for this combined class of constraints by revising *B7\_METRICSIZE* in the following way. In steps 2 and 5 we immediately check the entailed size constraints, instead of propagating them to  $\Sigma$ . If one of these constraints is not consistent with the corresponding definite size constraints, then return “no”, otherwise return “yes”. Step 6 is omitted.

In Section 2 we have proved that, while RSAT for  $\mathcal{B} \cup \{*\}$  is tractable, RSAT for  $\mathcal{B}_7 \cup \{*_7\}$  is NP-complete. Since  $\mathcal{B}_7 \cup \{*\}$  is clearly tractable, it is theoretically interesting to determine whether  $\mathcal{B}_7 \cup \{*\}$  plus metric size constraints is still tractable. We can easily prove that this is not the case by observing that the closure of  $\mathcal{B}_7 \cup \{*\}$  contains the relation  $\{\text{PO}\}$ . For instance, we have that  $x\{\text{NTPP}^{-1}\}y, y\{\text{EC}\}z, x\{\text{TPP}^{-1}\}w$  and  $w\{\text{NTPP}\}z$  imply  $x\{\text{PO}\}z$ . Thus, since  $\mathcal{B}_7 \cup \{*\}$  can be used to represent scenarios of RCC-8 relations, by Theorem 26 it follows that  $\mathcal{B}_7 \cup \{*\}$  plus metric size constraints is not tractable.

Since the combination of constraints over RCC-7 basic relations and metric size constraints is polynomial, we can use the algorithm of Figure 8 to derive a more general algorithm that decides the consistency of constraints over the full RCC-7 algebra combined with metric size constraints. This algorithm, that we call *RCC-7\_METRICSIZE*, is described in Figure 9. Given a set  $\Theta$  of constraints over RCC-7 and a set  $\Sigma$  of metric size constraints, *RCC-7\_METRICSIZE*

enumerates the consistent scenarios  $\Theta_s$  of  $\Theta$  by backtracking, and it decides the consistency of  $\Theta_s \cup \Sigma$  using our polynomial decision procedure. If  $\Theta_s \cup \Sigma$  is consistent for some scenario, then  $\Theta \cup \Sigma$  is consistent, otherwise it is inconsistent.

Enforcing path consistency (Step 1) requires cubic time. Steps 2, 3 and 5 can be accomplished by running known polynomial algorithms for solving a set of linear inequalities (e.g., [18,37,44]). Finally, the propagation of the containment constraints from  $\Theta$  to  $\Sigma'$  of Step 5 can be accomplished in polynomial time using Steps 3–5 of Algorithm `B7_METRICSIZE` (see Figure 8).

To conclude we observe that `RCC-7_METRICSIZE` can be used also as an approximate procedure for `RCC-8` and metric size constraints.<sup>19</sup> In this case we refine Step 3 to guarantee that, if a consistent scenario that does not involve `PO` exists, then this is derived before any scenario involving `PO`. This can be done by ordering the basic relations of each topological constraint in such a way that `PO` is the last relation to be selected. Also, the propagation of containment constraints of Step 5 is performed only for scenarios which do not involve `PO`. The resulting algorithm is not complete because it may fail to detect that the combined input sets are inconsistent. Clearly, if there exists a consistent scenario  $\Theta_s$  for the input set  $\Theta$  in which no pair of regions partially overlaps, and this scenario is consistent with the input set  $\Sigma$  of metric constraints, then the algorithm detects that  $\Theta_s \cup \Sigma$  is consistent. On the other hand, if there exists no such scenario, then the procedure may fail to detect that  $\Theta_s \cup \Sigma$  is unsatisfiable.

## 7 An Illustrative Example for `RCC-7` and Size Constraints

In this section we give a worked example illustrating the use of our results in a domain where regions cannot partially overlap.

We are considering locating the offices (or departments) of a certain company into an available space formed by several floors of a building (or of different buildings). Each office is formed by a cluster of several rooms, and some offices have a private rest room. Each office should be on one of the floors, since it would be inconvenient to have rooms (and collaborating people) of the same office on different floors (of possibly different buildings). Each floor has some

---

<sup>19</sup> A complete backtracking algorithm for the full `RCC-8` and metric size constraints would be computationally impractical, since in the worst case it would require a double exponential cost (there can be an exponential number of scenarios to generate, each of which could require exponential time just to identify and propagate the containment constraints).

“free area” that is not occupied by any office, and that is kept free for the main hallways and for a possible future enlargement of the company requiring larger offices or new offices. The free space on each floor should be least  $1/5$  of floor space. Some offices are concerned with related activities (e.g, the marketing and sales offices, or the personnel and payroll offices) should preferably be connected and hence stay on the same floor.

We have approximate information (minimum, maximum values) on the size of each office, and exact information on the size the overall available space for each floor, which we assume it is  $\frac{1}{n}$  of the overall available space, if we have  $n$  floors. Suppose the company under consideration has  $K$  offices ( $K_1 + K_2 + K_3 = K$ ): a group of  $K_1$  offices of size in a range  $S_1$  of possible sizes; a group of  $K_2$  offices of size in a range  $S_2$ ; a group of  $K_3$  offices of size in a range  $S_3$ . The exact values of  $K_1, K_2, K_3$  and  $S_1, S_2, S_3$  depend on the particular instance of the problem under consideration (for example, we could have an instance in which  $K_1 = K_2 = K_3 = 10, S_1 = [3, 5], S_2 = [4, 5], S_3 = [5, 6]$ ).<sup>20</sup>

Suppose we are considering an available space of overall size  $X$  distributed on  $N$  floors. (Again, the values of  $N$  and  $X$  depend on the particular instance of the problem under consideration.) In general, we are interested in determining whether all the constraints can be satisfied, if it is possible to place all the offices in the overall available space, and, if so, whether certain offices can be connected or stay on the same floor, on which floor they can be placed (especially if the available space on the floors is different), etc. These reasoning problems can be solved by running our algorithms for testing the satisfiability of the topological and size constraints of the problem, though the solutions that pass this test may still violate other practical constraints (see the discussion at the end of this section). A detailed illustration is given below, after the formalization of the domain using the constraints of our calculus.

### 7.1 Formalization of the domain

The general topological constraints of the domain are:

$Floor_i \{TPP\} Company$ , for  $i = 1..4$   
 $Office_i \{TPP, NTPP\} Company$ , for  $i = 1..K$   
 $Office_i \{DC, TPP, NTPP\} Floor_j$ , for  $i = 1..K, j = 1..4$   
 $FreeSpace_i \{TPP, NTPP\} floor_i$ , for  $i = 1..4$   
 $Floor_i \{DC\} Floor_j$ , for  $i, j = 1..4, i \neq j$

<sup>20</sup> Note that the domain could be generalized to the case in which the overall available space on each floor is different, we have arbitrarily many floors, and there are arbitrarily many offices with arbitrarily many different sizes. But for the sake of perspicuity we keep our illustrative example simple.

$Office_i \{DC, EC\} Office_j$ , for  $i, j = 1 \dots K, i \neq j$   
 $Rooms_i \{TPP, NTPP\} Office_i$ , for  $i = 1 \dots K$   
 $RestRoom_i \{TPP, NTPP\} Office_i$  (if  $Office_i$  has a rest room)  
 $RestRoom_i \{DC\} Rooms_j$ ,  $j = 1 \dots K$  (if  $Office_i$  has a rest room)

The size constraints are:<sup>21</sup>

$Size(Floor_1) = Size(Floor_2) = Size(Floor_3) = Size(Floor_4)$   
 $Size(Floor_1) = 1/4 Size(Company)$   
 $Size(FreeSpace_i) \geq 1/5 Size(floor_i)$ , for  $i = 1 \dots 4$   
 $Size(Office_i) \in S_1$ , for  $i = 1 \dots K_1$   
 $Size(Office_i) \in S_2$ , for  $i = K_1 + 1 \dots K_1 + K_2$   
 $Size(Office_i) \in S_3$ , for  $i = K_1 + K_2 + 1 \dots K$   
 $Size(Company) \in [X, X]$

## 7.2 Reasoning problem: consistency checking

*Problem:* Is it possible to place all the offices in the available space? I.e., is it possible to satisfy all the size and topological constraints of the problem instance?

This can be solved by running the backtracking algorithm `RCC-7_METRICSIZE`, which uses our polynomial algorithm for the basic relations (`B7_METRICSIZE`).

*Example for a problem instance*

If  $X = 60, K_1 = 1, K_2 = 2, K_3 = 3, S_1 = [4, 5], S_2 = [5, 6], S_3 = [9, 10]$ , then the answer to the problem is: NO.

Note that this is not trivial to determine, because the sum of the lower (and upper) bounds on the size of the  $K$  offices is less than the sum of the size of the floors, minus the relative  $1/5$  of space allocated to the free areas. This simple calculation is not sufficient to solve the problem because the offices cannot partially overlap floors. So, we cannot “optimize” space occupation by putting part of an office on a floor and another part on a different floor. This is not a restriction imposed by the model. It is rather a condition that is requested by the problem under consideration.

As an example of a positive instance of consistency checking, consider the previous instance with a different partition of the object groups:  $K_1 = 2, K_2 = 2, K_3 = 2$ . Clearly, the answer for this instance is YES, and the following is a

---

<sup>21</sup> Note that topology information is not sufficient to enforce that each office must be on one of the floors. The (topologically possible) case in which an office is inside the company and outside each the floors is excluded by the size information.

possible configuration for the offices: one of the two largest offices is on  $Floor_1$ , and the other one in  $Floor_2$ ; the two medium size offices are both on  $Floor_4$ ; the remaining offices are on  $Floor_3$ .

### 7.3 Reasoning problem: computing feasible scenarios

*Problem:* If the given set of constraints is consistent, compute a feasible (consistent) topological scenario satisfying the size constraints. Obviously the computed scenario gives a possible configuration for the position of all the objects in the office.

This problem can be solved by running `RCC-7_METRICSIZE`, with the simple addition that, when in Step 4 we check the consistency of  $\Sigma'$ , if this is consistent, then the algorithm returns  $\Theta_s$  as a feasible scenario. The problem can be easily generalized to the computation of *all* feasible scenarios.

### 7.4 Reasoning problem: querying relations and checking scenarios

*Problems:* Is a given scenario feasible? Is it possible to derive a scenario satisfying certain given relations? Which are the feasible topological relations between certain spatial regions?

The feasibility of a specific input scenario can be checked in polynomial time using `B7_METRICSIZE`. Queries can be answered by adding new constraints to the original set, and checking the consistency of the resulting set. This can be done by running the backtracking algorithm `RCC-7_METRICSIZE`.

*Instance 1:* If  $X = 60$ ,  $K_1 = 2$ ,  $K_2 = 2$ ,  $K_3 = 2$ ,  $S_1 = [4, 5]$ ,  $S_2 = [5, 6]$ ,  $S_3 = [9, 10]$ , is it possible that one of the medium size offices (e.g., the sales office  $Office_3$ ) and one of the largest size offices (e.g., the marketing office  $Office_5$ ) are externally connected?

*Additional constraints:*  $Office_3 \{EC\} Office_5$

*Answer:* NO.

Given the topological constraints, in all the consistent scenarios, if two offices are externally connected, then they must stay on the same floor. However, given the (inferred) maximum room that is available on each floor (i.e., at most 12.5) and the minimum room required by these two offices under consideration, there is no floor on which they can stay together (the extended set of constraints is inconsistent).



*Instance 2: If  $X = 60$ ,  $K_1 = 4$ ,  $K_2 = 4$ ,  $K_3 = 4$ ,  $S_1 = [2, 3]$ ,  $S_2 = [3, 4]$ ,  $S_3 = [7, 8]$ , which are the feasible topological relations between  $Office_1$  and  $Office_2$  (two of the smallest offices)?*

*Answer: {DC}*

The answer is obtained by checking the consistency of the constraints extended with either  $Office_1\{EC\}Office_2$  or  $Office_1\{DC\}Office_2$ . The addition of the first constraint leads to an inconsistent set. In fact, the problem instance is solvable only if each floor has exactly one office of each size.

The next example describes a case in which it is crucial to take into account the distinctions between externally connected and disconnected regions which are inside the same region, as well as the distinction between strict and non-strict containment constraints implied by the topological information.

*Instance 3: Suppose  $X = 60$ ,  $K_1 = 1$ ,  $K_2 = 1$ ,  $K_3 = 3$ ,  $S_1 = [4, 5]$ ,  $S_2 = [8, 9]$ ,  $S_3 = [10, 12]$ , and that  $Office_1$  consists of two rooms of size 2. Is it possible that the medium size office ( $Office_2$ ) has a rest room that is (externally) connected with the smallest office ( $Office_1$ )?*

*Additional constraints:*

$RestRoom_2 \{EC\} Office_1$   
 $Room_{1,i}\{TPP, NTPP\} Rooms_1$ , for  $i = 1, 2$   
 $Size(Room_{1,i}) \in [2, 2]$ , for  $i = 1, 2$

*Answer: NO.*

The rest room of  $Office_2$  cannot be externally connected with  $Office_1$ , because this office is formed by exactly two work rooms, and at least one of them would have to be connected with the external rest room, contrary to the preference (stated as a general constraint) that no rest room is externally connected with a work room. This inconsistency is detected by our techniques because the fact that  $RestRoom_2$  is externally connected with  $Office_1$ , but disconnected with every region ( $Rooms_{1,i}$ ) inside  $Office_1$ , forces the strict containment constraint

$$Size(Room_{1,1}) + Size(Room_{1,2}) < Size(Office_1),$$

which violates the implied fact that the two rooms of  $Office_1$  completely fills it in. This case is analogous to the case described in Figure 6(c), where we have  $RestRoom_2$  instead of  $y$ , and  $Room_{1,1-2}$  instead of  $x_1$ ,  $x_2$  and  $x_3$ .

## 7.5 Discussion

The office scenario is just a particular example of a whole class of problems which could be called *topological bin-packing problems*. In a standard bin-packing problem [25] we have a set of elements of known sizes, each of which should be assigned to one of several bins with given capacities in such a way that the sum of the sizes of all elements assigned to each bin does not exceed the capacity of the bin. In a topological bin-packing problem we can have exact or indefinite size information (e.g., minimum/maximum sizes), as well as qualitative and quantitative size constraints between elements of the domain. Another major difference is that the elements of the domain can be topologically characterized. This allows us to represent domains where, for example, certain elements form containment hierarchies, or have spatial boundaries that must/must not touch.

There are many other domains involving problems similar to the one in the previous office example. Consider for instance the organization of categories of merchandise in the (possibly non-adjacent) areas of a general store, or of book categories in the stack areas of a library; the allocation of cargos (sets of containers) to a fleet of ships in a transportation problem; storage of items in warehouses; etc. However, it should be noted that constraint satisfaction techniques such as the ones developed in this paper should not be expected to suffice in themselves for confirming satisfiability of all the requirements that arise in some practical application domain. In practice, there are always miscellaneous additional types of constraints besides those allowed for by a general constraint satisfaction tool. In particular, in any practical physical design domain involving topological and size constraints, there will always be additional constraints (whether on other spatial aspects, constructibility, aesthetics, conformity with bylaws, etc) which may render a possible design impractical. But the proposed techniques can efficiently make *infeasibility* judgments based on a subset of the available information, namely the topological and size information. In other words, they can serve as a method for generating potential solutions that can then be evaluated and refined taking into account additional information, or for pruning impossible solutions.

The algorithms developed in this paper significantly extend the subset of constraints that are considered for such purposes; they handle topological and size constraints simultaneously, and thus provide a more powerful method than any methods for topological constraints or size constraints alone. And as is evident from the examples just mentioned, there are many problem areas where such constraints comprise a subset of those that need to be satisfied.

## 8 Summary and Conclusions

Integrating different spatial aspects and studying their interdependencies is important for many applications of spatial reasoning, since most applications deal with more than just one spatial aspect. In this paper we have addressed the problem of integrating two basic types of spatial information, namely, topological information, expressed as RCC-8 constraints between spatial regions, and information about the size of the regions, expressed as qualitative or metric constraints.

Regarding qualitative size information, we have introduced a class of qualitative constraints on the relative size of spatial regions ( $QS$ ), and we have given a cubic time path-consistency algorithm, BIPATH-CONSISTENCY, for processing a set of constraints in  $RCC-8 \cup QS$ . When the input topological constraints are in either  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  or  $\mathcal{Q}_8$  (the only three maximal tractable subclasses of RCC-8 containing all the basic relations), this algorithm is correct and complete for deciding consistency. We have also proposed a cubic time algorithm for computing a consistent scenario for a set of topological constraints in  $QS$  and either  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  or  $\mathcal{Q}_8$ . BIPATH-CONSISTENCY is a general algorithm for handling two different sets of constraints which are not independent from each other, and hence can be applied to other types of constraints, e.g., Allen's interval constraints [1] and qualitative constraints on their duration. Of course, different classes of relations might need different completeness and complexity proofs.

In general, modeling a constraint between two variables which contains different types of information is not trivial. A possible alternative approach could be the definition of a new constraint language (or class of constraints), for which composition should be appropriately defined in order to apply a known path-consistency algorithm (yet the completeness of this approach is not immediately guaranteed). In this approach the interactions between different types of information are encoded in the relations, and this leads to a large number of different relations which can negatively affect the performance of the reasoning algorithms.<sup>22</sup> Our approach for combining constraints is different. It does not require to specify a new language, and it is more general because it can be applied to combine also other types of constraints (it requires only to specify the map of the interdependencies between the different types of constraints). Moreover, it does not require to increase the number or relations with respect to number of those in the original languages.

---

<sup>22</sup> Something similar to this has recently been done in the context of temporal reasoning by Pujari, Kumari and Sattar in order to model interval relations and duration relations [51]. Their new language contains 25 basic relations and  $2^{25}$  relations in total, while there are only 13 (Allen's) basic interval relations, and 3 basic duration relations.

Regarding metric size information, we have introduced three classes of constraints, and we have studied their combination with RCC-8 and with RCC-7, a subclass of RCC-8 which is important for applications involving spatial regions that cannot partially overlap (e.g., political and administrative regions, or solid-state physical objects). The interdependencies between topological and metric size constraints are more intricate than those between topological and qualitative size constraints. Contrary to the qualitative case, adding metric size information to  $\widehat{\mathcal{H}}_8$ ,  $\mathcal{C}_8$  or  $\mathcal{Q}_8$  leads to intractability. This is true even when the set of topological relations is restricted to only the set of the eight RCC-8 basic relations.

We have then shown that reasoning with the RCC-7 basic relations and metric size constraints is tractable, and we have given a polynomial time algorithm for deciding consistency in the combined class. Based on this algorithm, we have presented a general backtracking algorithm for constraints over the full RCC-7 algebra combined with metric size constraints.

An interesting open question is to what extent we can push the tractability limit further towards more expressive languages for qualitative size information combined with tractable subclasses of RCC-8. This could include a finer grained qualitative size calculus with relations such as “slightly smaller” or “much smaller”. Other open questions are whether RSAT for the RCC-8 basic relations and metric size constraints belongs to NP, and whether there are other interesting subclasses of RCC-7 that can be combined with metric size constraints without losing tractability. Finally, the next step to obtain a more expressive calculus which covers multiple aspects of space is to extend topological and size constraints with information about other spatial aspects, such as direction, distance, or shape. From a computational point of view, this is more promising when information is only qualitative, since reasoning with metric size constraints and topological constraints is already computationally hard.

## Acknowledgments

A preliminary version of some results contained in the first part of the paper appeared in the Proceedings of the Fourth International Conference on Principles and Practice of Constraint Programming, Springer-Verlag, 1998 [28]. The work of the first author has been partly supported by CNR project SCI\*SIA, and by a CNR short-term fellowship. The work of the second author has been supported by DFG as part of the project FAST-QUAL-SPACE (DFG special research effort on “Spatial Cognition”), and by a Marie Curie Fellowship of the EC programme “Improving Human Potential” (contract HPMF-CT-2000-00667). We thank Tony Cohn, Carola Eschenbach, Bernhard Nebel, Achille

Varzi, and the anonymous referees for their comments. A special thank to Len Schubert for a very helpful discussion.

## References

- [1] J.F. Allen. Maintaining knowledge about temporal intervals. *Communication of the ACM*, 26(1):832–843, 1983.
- [2] M. Aiello, C. Areces, and M. de Rijke. Spatial reasoning for image retrieval. In *Proceedings of the 1999 International Workshop on Description Logics (DL'99)*, 1999.
- [3] T. Apostol. *Mathematical Analysis*. Addison Wesley, 1974.
- [4] N. Asher and L. Vieu. Towards a geometry of common sense: A semantics and a complete axiomatization of mereotopology. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, pages 846–852, Montreal, Canada, August 1995.
- [5] B. Bennett. Spatial reasoning with propositional logic. In J. Doyle, E. Sandewall, and P. Torasso, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the 4th International Conference*, pages 51–62, Bonn, Germany, May 1994. Morgan Kaufmann.
- [6] B. Bennett. Modal logics for qualitative spatial reasoning. *Bulletin of the IGPL*, 4(1):23 – 45, 1996.
- [7] B. Bennett, A.G. Cohn, and A. Isli. Combining multiple representations in a spatial reasoning system. In *Proceedings of the 9th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'97)*, 1997.
- [8] C. Bessière. A simple way to improve path-consistency in Interval Algebra networks. In *Proc. of the 13th Nat. Conf. of the American Association for Artificial Intelligence (AAAI-96)*, pages 375–380, Portland, OR, 1996.
- [9] L. Biacino and G. Gerla. Connection structures. *Notre Dame Journal of Formal Logic*, 32(2):242–247, 1991.
- [10] S. Borgo, N. Guarino, and C. Masolo. A pointless theory of space based on strong connection and congruence. In L.C. Aiello, J. Doyle, and S.C. Shapiro, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the 5th International Conference*, pages 220–229, Cambridge, MA, November 1996. Morgan Kaufmann.
- [11] Z. Chen, M. Grigni, and C. Papadimitriou. Planar map graphs. In *30th Annual ACM Symp. on the Theory of Computing*, 1998.
- [12] B.L. Clarke. A calculus of individuals based on connection. *Notre Dame Journal of Formal Logic*, 22(3):204–218, 1981.

- [13] E. Clementini, P. di Felice, and D. Hernandez. Qualitative representation of positional information. *Artificial Intelligence*, 95(2):317–356, 1997.
- [14] A.G. Cohn, B. Bennett, J. Gooday, and N.M. Gotts. Qualitative spatial representation and reasoning with the region connection calculus. *GeoInformatica*, 1(3):275–316, 1997.
- [15] T. Cormen, C. Leiserson, and R. Rivest. *Introduction to Algorithms*, The MIT Press. 1990.
- [16] Z. Cui, A.G. Cohn, and D.A. Randell. Qualitative simulation based on a logical formalism of space and time. In *Proceedings of the 10th National Conference of the American Association for Artificial Intelligence*, pages 679 – 684, San Jose, CA, July 1992. MIT Press.
- [17] E. Davis. *Representation of Commonsense Knowledge*. Morgan Kaufmann Publishers, 1990.
- [18] R. Dechter, I. Meiri, and J. Pearl. Temporal Constraint Networks. *Artificial Intelligence*, 49:61–95, 1991.
- [19] I. Düntsch, H. Wang, and S. McCloskey. A relation algebraic approach to the region connection calculus. *Theoretical Computer Science*, 255(1-2):63–83, 2001.
- [20] M.J. Egenhofer. Reasoning about binary topological relations. In O. Günther and H.-J. Schek, editors, *Proceedings of the Second Symposium on Large Spatial Databases, SSD'91*, volume 525 of *Lecture Notes in Computer Science*, pages 143–160. Springer-Verlag, Berlin, Heidelberg, New York, 1991.
- [21] M.J. Egenhofer. Query processing in spatial-query-by-sketch. *Journal of Visual Languages and Computing*, 8(4):403–424, 1997.
- [22] M.J. Egenhofer and R.G. Golledge, editors. *Spatial and Temporal Reasoning in Geographic Information Systems*. Oxford University Press, Oxford, UK, 1998.
- [23] M.T. Escrig. *Qualitative Spatial Reasoning: theory and practice, Application to Robot Navigation*. IOS Press, 1998.
- [24] A. Galton and R.C. Meathrel. Qualitative outline theory. In *Proceedings of the 16th International Joint Conference on Artificial Intelligence*, pages 1061–1066, Stockholm, Sweden, August 1999.
- [25] M.R. Garey and D.S. Johnson. *Computers and Intractability—A Guide to the Theory of NP-Completeness*. Freeman, San Francisco, CA, 1979.
- [26] A. Gerevini, A. Perini, and F. Ricci. Incremental algorithms for managing temporal constraints. In *Proceedings of the 8th IEEE International Conference on Tools with Artificial Intelligence*, Toulouse, France, 1996. IEEE Computer Society Press.
- [27] A. Gerevini and M. Cristani. On finding solutions in temporal constraint networks. In *Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence (IJCAI-97)*, pages 1460–1465, Nagoya, Japan, 1997. Morgan Kaufmann.

- [28] A. Gerevini and J. Renz. Combining topological and qualitative size constraints for spatial reasoning. In *Proceedings of the Fourth International Conference on Principles and Practice of Constraint Programming (CP'98)*. Springer-Verlag, 1998.
- [29] A. Gerevini and J. Renz. Combining Topological and Size Information for Spatial Reasoning. Technical Report, DEA, Università di Brescia, February, 2002.
- [30] A. Gerevini and L. Schubert. On computing the minimal labels in time point algebra networks. *Computational Intelligence*, 11(3):443–448, 1995.
- [31] A. Gerevini and L. Schubert. Efficient algorithms for qualitative reasoning about time. *Artificial Intelligence*, 74:207–248, 1995.
- [32] M.C. Golumbic and R. Shamir. Complexity and algorithms for reasoning about time: A graph-theoretic approach. *Journal of the Association for Computing Machinery (ACM)*, 40(5):1128–1133, November 1993.
- [33] J.M. Gooday and A.G. Cohn. Using spatial logic to describe visual programming languages. *Artificial Intelligence Review*, 10:171–186, 1996.
- [34] M. Grigni, D. Papadias, and C. Papadimitriou. Topological inference. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, pages 901–906, Montreal, Canada, August 1995.
- [35] V. Haarslev, C. Lutz, and R. Möller. Foundations of spatioterminological reasoning with description logics. In A.G.Cohn, L. Schubert, and S.C. Shapiro, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the 6th International Conference*, pages 112–123, Trento, Italy, June 1998.
- [36] P. Jonsson and T. Drakengren. A Complete Classification of Tractability in RCC-5. *Journal of Artificial Intelligence Research (JAIR)*, 6:211–221, 1997.
- [37] N. Karmarkar. A new polynomial time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.
- [38] H.A. Kautz and P.B. Ladkin. Integrating metric and qualitative temporal reasoning. In *Proceedings of the Ninth National Conference on Artificial Intelligence (AAAI-91)*, pages 241–246. MIT Press, 12–19 July 1991.
- [39] M. Knauff, R. Rauh, and J. Renz. A cognitive assessment of topological spatial relations: Results from an empirical investigation. In *Proceedings of the 3rd International Conference on Spatial Information Theory (COSIT'97)*, volume 1329 of *Lecture Notes in Computer Science*, pages 193–206, 1997.
- [40] B. Kuipers. *Qualitative Reasoning : Modeling and Simulation With Incomplete Knowledge*. MIT Press, Cambridge, MA, 1994.
- [41] P.B. Ladkin and R. Maddux. On binary constraint networks. Technical Report KES.U.88.8, Kestrel Institute, Palo Alto, CA, 1988.

- [42] P.B. Ladkin and R. Maddux. On binary constraint problems. *Journal of the Association for Computing Machinery (ACM)*, 41(3):435–469, May 1994.
- [43] P.B. Ladkin and A. Reinefeld. Effective solution of qualitative interval constraint problems. *Artificial Intelligence*, 57(1):105–124, 1992.
- [44] J.L. Lassez and K. McAloon. A canonical form for generalized linear constraints. *Journal of Symbolic Computation*, 13:1–24, 1992.
- [45] G. Ligozat. Reasoning about cardinal directions. *Journal of Visual Languages and Computing*, 9(1):23–44, 1998.
- [46] W. Maaß, P. Wazinski, and G. Herzog. VITRA GUIDE: Multimodal route descriptions for computer assisted vehicle navigation. In *Proc. of the Sixth Int. Conf. on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems IEA/AIE-93*, pages 144–147, Edinburgh, Scotland, 1993.
- [47] A.K. Mackworth. Consistency in networks of relations. *Artificial Intelligence*, 8:99–118, 1977.
- [48] A.K. Mackworth and Eugene C. . Freuder. The complexity of some polynomial network consistency algorithms for constraint satisfaction problems. *Artificial Intelligence*, 25:65–73, 1985.
- [49] B. Nebel. Solving hard qualitative temporal reasoning problems: Evaluating the efficiency of using the ORD-horn class. *CONSTRAINTS*, 1(3):175–190, 1997.
- [50] C.H. Papadimitriou, D. Suciuc, and V. Vianu. Topological queries in spatial databases. *Journal of Computer and System Sciences*, 58(1):29–53, 1999.
- [51] A.K. Pujari, G.V. Kumari, and A. Sattar. *INDU: An Interval & Duration Network*. *Proceedings of the 13th Australian Joint Conference on Artificial Intelligence*. 2000.
- [52] O. Raiman. Order of magnitude reasoning. *Artificial Intelligence*, 51(1-3):11–38, 1991.
- [53] D.A. Randell and A.G. Cohn. Naive topology: modeling the force pump. In Peter Struss and Boi Faltings, editors, *Advances in Qualitative Physics*, pages 177–192. MIT Press, 1992.
- [54] A. Randell, Z. Cui, and A.G. . Cohn. A spatial logic based on regions and connection. In B. Nebel, W. Swartout, and C. Rich, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the 3rd International Conference*, pages 165–176, Cambridge, MA, October 1992. Morgan Kaufmann.
- [55] J. Renz. A canonical model of the Region Connection Calculus. In *Principles of Knowledge Representation and Reasoning: Proceedings of the 6th International Conference*, Trento, Italy, June 1998.
- [56] J. Renz. Maximal tractable fragments of the region connection calculus: A complete analysis. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI-99)*, 1999.



- [57] J. Renz and 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.
- [58] J. Renz, R. Rauh, and M. Knauff. Towards cognitive adequacy of topological spatial relations. In C. Freksa, W. Brauer, C. Habel, and K.F. Wender, editors, *Spatial Cognition II—Integrating abstract theories, empirical studies, formal models, and practical applications*, volume 1849 of *LNCS*, pages 184-197. Springer-Verlag, Berlin, Heidelberg, New York, 2000.
- [59] J.G. Stell. A new approach to the region-connection calculus. *Artificial Intelligence*, 122(1-2):111-136, 2000.
- [60] A. Tarski. On the calculus of relations. *Journal of Symbolic Logic*, 6:73-89, 1941.
- [61] P. van Beek. Reasoning about qualitative temporal information. *Artificial Intelligence*, 58(1-3):297-321, 1992.
- [62] M.B. Vilain, H.A. Kautz. Constraint Propagation Algorithms for Temporal Reasoning. In *Proceedings of the Fifth National Conference of the American Association for Artificial Intelligence (AAAI-86)*, pages 377-382, 1986.
- [63] M. Vilain, H.A. Kautz, and P. van Beek. Constraint propagation algorithms for temporal reasoning: a revised report. In D.S Weld and J. de Kleer, editors, *Readings in Qualitative Reasoning about Physical Systems*, pages 373-381. Morgan Kaufmann, San Mateo, CA, 1990.
- [64] F. Wolter and M. Zakharyashev. Spatio-temporal representation and reasoning based on RCC-8. In *Principles of Knowledge Representation and Reasoning Proceedings of the Seventh International Conference (KR'2000)*, pages 3-14, 2000.
- [65] K. Zimmermann. Measuring without measures: The delta-calculus. In *Proceedings of the 2nd International Conference on Spatial Information Theory (COSIT'95)*, volume 988 of *Lecture Notes in Computer Science*, pages 59-67, 1995.