On the Translation of Qualitative Spatial Reasoning Problems into Modal Logics

W. Nutt

German Research Center for

Artificial Intelligence (DFKI)

Stuhlsatzenhausweg 3

66123 Saarbrücken, Germany

E-mail `Werner.Nutt@dfki.de`

Abstract

Among the formalisms for qualitative spatial reasoning, the Region Connection Calculus and its variant, the constraint algebra RCC8, have received particular attention recently. A translation of RCC8 constraints into a multimodal logic has been proposed by Bennett, but in his work a thorough semantical foundation of RCC8 and of the translation into modal logic is missing.

In the present paper, we give for the first time a rigorous foundation for reasoning in RCC8. To represent qualitative relationships between regions, we introduce a language of *topological set constraints,* which generalizes RCC8. We formulate a semantics for our language that interprets regions as subsets of topological spaces. Using McKinsey and Tarski's topological interpretation of the modal propositional logic **S4**, we reduce reasoning about topological constraints to reasoning in that logic. We show that reasoning in the general language is PSPACE-complete. As a special case, we obtain also a reduction of constraint solving in RCC8.

1 Introduction

An approach to qualitative spatial reasoning that has received considerable attention is the so-called *Region Connection Calculus* (RCC), which has been introduced by Randell, Cui, and Cohn [9]. A specialization of RCC is the calculus RCC8. Similar to Allen's calculus for temporal reasoning [1], which is based on 13 elementary relations that can hold between time intervals, in RCC8, there are eight elementary topological relations that can hold between regions. The base relations are used to specify qualitative constraints that hold between regions. Then, the reasoning tasks are to determine whether a set of constraints is consistent or whether it entails another one.

Different semantics of RCC8 have been considered, which essentially differ in the way that "regions" are interpreted. Applications in geographical information systems, for instance, motivate to interpret a region as a connected subset of the plane whose boundary is a continuous curve or, more specifically, as a polygon [3]. For this semantics, however, it is not known whether reasoning is decidable.

As an alternative, Bennett has proposed to interpret regions as subsets of a topological space [2]. There is a close connection between propositional languages describing sets in topological spaces and modal propositional logics, which has already been pointed out by McKinsey and Tarski [7, 8]. Bennett gave a translation of RCC8 constraints to formulas in a multimodal logic. However, in his work a thorough semantical foundation of RCC8 and of the translation into modal logic is missing. In particular, it is unclear whether the translation preserves the satisfiability of constraints. Nonetheless, other researchers have taken the correctness of Bennett's translation for granted and based their own work upon it. For instance, Renz and Nebel applied it to identify maximal tractable fragments of RCC8 [10], and Haarslev et al. [4] used it to develop algorithms for spatioterminological reasoning.

In the present paper, we give for the first time a rigorous foundation for reasoning in RCC8. To represent topological relationships between regions, we introduce a language of *topological set constraints,* which generalizes RCC8. For example, we can express in the new language relations that involve more than two regions. We formulate a semantics for our language that interprets regions as subsets of topological spaces. Using McKinsey and Tarski's topological interpretation of the modal propositional logic **S4**, we reduce reasoning about topological constraints to reasoning in that logic. Algorithms for the latter are well-known (see e.g. [5]). Through our translation, they become applicable to spatial reasoning problems. We show that reasoning in the general language is PSPACE-complete. As a special case, we obtain also a reduction of constraint solving in RCC8. Interestingly, the work by Renz and Nebel on the complexity of RCC8 still applies to our translation, although it differs from the one given by Bennett.

Figure 1 illustrates each base relation of RCC8 by a pair X, Y of planar regions with continuous boundary curve such that the relation holds between

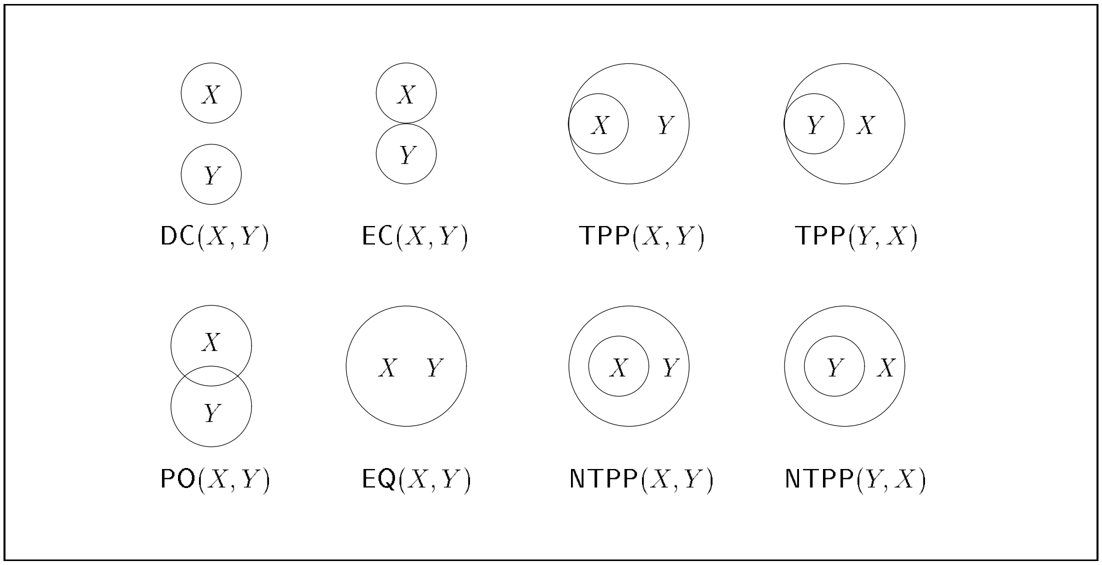

Figure 1: A graphical representation of the RCC8 relations

them. The "interior" of such a region is the region area minus the boundary curve. The regions are "disconnected," written $\mathsf{DC}(X, Y)$, if they are disjoint. They are "externally connected," written $\mathsf{EC}(X, Y)$, if their interiors are disjoint, but the regions are not. They "partially overlap," written $\mathsf{PO}(X, Y)$, if their interiors intersect, but none is a subset of the other. They are "equal," written $\mathsf{EQ}(X, Y)$ if they are identical. The region X is a "tangential proper part" of Y, written $\mathsf{TPP}(X, Y)$, if X is a subset of Y, but not of the interior of Y. Finally, X is a "non-tangential proper part" of Y, written $\mathsf{NTPP}(X, Y)$, if X is contained in the interior of Y. Since tangential and non-tangential proper part are asymmetric relations, their converses can also hold between X and Y.

In the rest of the paper, we remind the reader of the basic notions of point set topology so far as they are relevant for our subject (Section 2), define topological set constraints and show how to formally express with them the RCC8-relations (Section 3), recall the topological interpretation of modal logic (Section 4), and then reduce reasoning about set constraints to reasoning in **S4** (Sections 5 and 6). Finally, we translate set constraints into formulas of a multimodal logic with an additional **K**-operator such that satisfiability is preserved (Section 7).

2 Reminder on Point Set Topology

Point set topology—or simply topology—is a mathematical theory that deals with properties of space that are independent of size and shape. In topology, one can define concepts such as interior, exterior, and isolated points of regions, boundaries and connected components of regions, and connected regions. Point set topology is therefore a possible framework for qualitative spatial reasoning.

In this section, we review the basic concepts of topology, with particular emphasis on interior operators. For further information on the topic we refer to standard textbooks like [6].

2.1 Definitions

A topological space is a pair $\mathcal{T} = (U, \mathcal{O})$ where U is a nonempty set, called the *universe*, and $\mathcal{O}$ is a set of subsets of U, called the *topology* of $\mathcal{T}$, such that the following holds:

1. $\emptyset \in \mathcal{O}$ and $U \in \mathcal{O}$;
2. if $O_1 \in \mathcal{O}$ and $O_2 \in \mathcal{O}$, then $O_1 \cap O_2 \in \mathcal{O}$;
3. if O_i, $i \in I$, is a (possibly infinite) family of elements of $\mathcal{O}$, then

$$\bigcup_{i \in I} O_i \in \mathcal{O}.$$

The elements of $\mathcal{O}$ are called the *open sets* of $\mathcal{T}$. For example, in Euclidean space, a set O is open if for each point $p \in O$ there is a circle surrounding p that is contained in O.

Let $\mathcal{T} = (U, \mathcal{O})$ be a topological space, $S \subseteq U$ a subset of U, and $p \in U$ a point. Then p is an *interior point* of S if there is an open set $O \in \mathcal{O}$ such that $p \in O$ and $O \subseteq S$. We denote the *set of interior points* of S as $i(S)$. The set $i(S)$ is the largest open set contained in S.

Proposition 2.1 *Let $\mathcal{T} = (U, \mathcal{O})$ be a topological space and $i\colon 2^U \to 2^U$ be the mapping that associates to every subset of U its set of interior points. Then for all S, $T \subseteq U$ we have*

1. $i(U) = U$;
2. $i(S) \subseteq S$;
3. $i(S) \cap i(T) = i(S \cap T)$;
4. $i(i(S)) = i(S)$.

The point p is a *touching point* of S if every open set containing p has a nonempty intersection with S. Finally, p is a *boundary point* of S if p is a touching point of S and of its complement $U \setminus S$. We denote the set of touching points of S as $cl(S)$, and the set of boundary points of S as $bd(S)$. We call $i(S)$ the *interior* of S, $cl(S)$ the *closure* of S, and $bd(S)$ the *boundary* of S. We say that a set S is *closed* if it contains all its touching points. Closure and boundary, openness and closedness, can all be expressed in terms of interiors of sets.

Proposition 2.2 *Let $\mathcal{T} = (U, \mathcal{O})$ be a topological space and $S \subseteq U$. Then*

- *$cl(S) = U \setminus i(U \setminus S)$, and $bd(S) = cl(S) \cap cl(U \setminus S)$*
- *S is open iff $S = i(S)$, and S is closed iff $S = cl(S)$.*

2.2 Interior Operators

In the preceding subsection, we have defined a topological space in terms of its open sets. Now we show that it is also possible to define it through the mapping i.

Let U be a set and $i\colon 2^U \to 2^U$ an operator that maps subsets of U to subsets of U. We say that i is an *interior operator* if it satisfies the properties (1) to (4) of Proposition 2.1. Every topology determines an interior operator.

Proposition 2.3 *Let $\mathcal{T} = (U, \mathcal{O})$ be a topological space and i be the operator that maps every subset of U to its interior. Then*

1. *i is an interior operator;*
2. *a set $S \subseteq U$ is open if and only if it is a fixpoint of i, that is, $i(S) = S$.*

The second statement of the above proposition says that not only determines the topology of $\mathcal{T}$ an interior operator, but the topology can also be reconstructed from the interior operator. We will show that also the converse is true. An interior operator determines a topology, from which the operator can be reconstructed. For an interior operator i on a set U we define $\mathcal{O}_i$ as the set of fixpoints of i, that is,

$$\mathcal{O}_i := \{O \subseteq U \mid O = i(O)\}.$$

Proposition 2.4 *If i is an interior operator, then $\mathcal{O}_i$ is a topology.*

Proposition 2.5 *Let i be an interior operator and $i_{\mathcal{O}_i}$ be the interior operator corresponding to the topology $\mathcal{O}_i$. Then $i = i_{\mathcal{O}_i}$.*

This shows that point set topology can equivalently be based on open sets and on interior operators. In the sequel, we will therefore talk about topological spaces as pairs (U, i), where i is an interior operator on U.

3 Set Expressions and Set Constraints

We now introduce a formal language in which we can make statements about relationships between subsets of a topological space. We assume that there is a countably infinite set of *variables* (written X, Y, Z). *Set expressions* (written s, t) are built up according to the syntax rule

$$s, t \quad \longrightarrow \quad X \mid \top \mid \bot \mid s \sqcap t \mid s \sqcup t \mid \overline{s} \mid \mathtt{I}s.$$

A *topological interpretation* $\mathcal{I}$ is a triple

$$\mathcal{I} = (U, i, d),$$

where (U, i) is a topological space, and d is a function that maps every variable to a subset of U. The function d can be extended in a unique way to set expressions such that the following identities hold:

$$\begin{aligned} d(\bot) &= \emptyset \\ d(\top) &= U \\ d(s \sqcap t) &= d(s) \cup d(t) \\ d(s \sqcup t) &= d(s) \cap d(t) \\ d(\overline{s}) &= U \setminus d(s) \\ d(\mathtt{I}s) &= i(d(s)). \end{aligned}$$

If no misunderstanding can arise, we will refer to topological interpretations simply as interpretations.

Elementary set constraints have the form

$$s \doteq t \qquad \text{or} \qquad s \not\doteq t,$$

where s, t are set expressions. Constraints of the form $s \doteq t$ are *positive*, while those of the form $s \not\doteq t$ are *negative*. *Complex set constraints* (written C, D) are obtained from elementary ones using the propositional connectives conjunction, disjunction, and negation. Thus, if C, D are set constraints, so are $C \wedge D$, $C \vee D$, and $\neg C$.

Next we define when an interpretation $\mathcal{I} = (U, i, d)$ *satisfies* a set constraint C (written $\mathcal{I} \models C$). For elementary constraints we have

$$\begin{aligned} \mathcal{I} \models s \doteq t \quad &\text{iff} \quad d(s) = d(t) \\ \mathcal{I} \models s \not\doteq t \quad &\text{iff} \quad d(s) \neq d(t). \end{aligned}$$

For conjunctions, disjunctions, and negations of constraints, satisfaction is defined as one would expect.

An interpretation $\mathcal{I}$ is a *model* of C if $\mathcal{I}$ satisfies C. A constraint is *satisfiable* if it has a model. Two constraints are *equivalent* if they have the same models. If $\mathcal{C}$ is a set of constraints, then $\mathcal{I}$ is a model of $\mathcal{C}$ if $\mathcal{I}$ is a model of every constraint in $\mathcal{C}$. The set $\mathcal{C}$ *entails* a constraint D (written $\mathcal{C} \models D$) if every model of $\mathcal{C}$ is also a model of D. As usual, entailment can be reduced to (un)satisfiability, since $\mathcal{C} \models D$ if and only if $\mathcal{C} \cup \{\neg D\}$ is unsatisfiable.

With our topological set constraints, we can now give a semantics to the RCC8-relations. The constraints range over variables that represent regions. Which sets are regions? First, regions should be nonempty. Moreover, among subsets of the plane, most authors consider lines or sets with cracks as pathological and rule them out as regions. Formally, such sets can be excluded by requiring that regions are *regular,* that is, they coincide with the closure of their interior. That a set X is regular and nonempty can be expressed by the constraint

$$X \doteq \overline{\mathtt{I}\overline{\mathtt{I}X}} \quad \wedge \quad X \not\doteq \bot. \tag{1}$$

We define now each base relation as a shorthand for a certain set constraint:

$$\begin{array}{rcl}
\mathsf{DC}(X,Y) & := & X \sqcap Y \doteq \bot \\
\mathsf{EC}(X,Y) & := & X \sqcap Y \not\doteq \bot \ \wedge\ \mathtt{I}X \sqcap \mathtt{I}Y \doteq \bot \\
\mathsf{PO}(X,Y) & := & \mathtt{I}X \sqcap \mathtt{I}Y \not\doteq \bot \ \wedge\ X \sqcap \overline{Y} \not\doteq \bot \ \wedge\ \overline{X} \sqcap Y \not\doteq \bot \\
\mathsf{EQ}(X,Y) & := & X \doteq Y \\
\mathsf{TPP}(X,Y) & := & X \sqcap \overline{Y} \doteq \bot \ \wedge\ X \sqcap \overline{\mathtt{I}Y} \not\doteq \bot \\
\mathsf{NTPP}(X,Y) & := & X \sqcap \overline{\mathtt{I}Y} \doteq \bot.
\end{array}$$

It is easy to verify that the eight base relations are mutually exclusive and cover all possible cases. That is, any two subsets of a topological space, regardless of whether they are regions or not, satisfy one and only one base relation.

A *basic* RCC8 *constraint system* over a set of region variables is a conjunction that contains for each variable the regularity constraint (1) and for some pairs of variables a base relation. In a *general constraint system,* there can be disjunctions of base relations between variables. Satisfiability and entailment of RCC8 constraints are defined as for arbitrary set constraints. Methods to decide the satisfiability of arbitrary topological set constraints are therefore also applicable to RCC8.

4 The Topological Interpretation of Modal Logics

In propositional logic, formulas are constructed from variables (again denoted as X, Y) and the constants $\top$ and $\bot$ (standing for *true* and *false*) with the connectives $\wedge$, $\vee$, and $\neg$.[1] In *modal* propositional logic, there is in addition a unary operator, which is often written as $\Box$. We denote such formulas with the letters ϕ, ψ.

The semantics of such formulas is usually given in terms of Kripke-interpretations. Such an interpretation consists of a set, whose elements are called *worlds.* The worlds are connected by a binary relation, the *reachability* relation. To each world, a propositional interpretation is associated that assigns truth values to the variables. A set of worlds with a reachability relation is called a *frame.*

[1] We also use, as shorthands, the connectives $\rightarrow$ and $\leftrightarrow$.

One singles out variants of modal logics by admitting only interpretations whose reachability relation has certain properties. The most general logic is the logic **K**, where all possible interpretations are allowed. For the logic **S4**, only interpretations are admitted whose reachability relation is *reflexive* and *transitive*. For reasons that will become obvious later on we write the modal operator in **S4** as **I**. The following schemas are valid in **S4**, that is, they hold in every **S4**-interpretation and for every modal formula ϕ:

1. $\mathbf{I}\top \leftrightarrow \top$;
2. $\mathbf{I}\phi \rightarrow \phi$;
3. $\mathbf{I}(\phi \wedge \psi) \leftrightarrow \mathbf{I}\phi \wedge \mathbf{I}\psi$;
4. $\mathbf{I}\mathbf{I}\phi \leftrightarrow \mathbf{I}\phi$.

These formulas are also characteristic for **S4** in the sense that a frame with the property that all interpretations over it satisfy all formulas of the above kind is reflexive and transitive.

There is an obvious analogy between the above formulas and the equations in Proposition 2.1 that characterize interior operators on topological spaces. It has been shown that the analogy is not accidental. To see this, let π be the bijection that maps set expressions s to modal propositional formulas $\pi(s)$ by mapping variables, $\top$ and $\bot$ to themselves, and by recursively replacing $s \sqcap t$, $s \sqcup t$, $\overline{s}$, and $\mathtt{I}s$ with $\pi(s) \wedge \pi(t)$, $\pi(s) \vee \pi(t)$, $\neg\pi(s)$, and $\mathbf{I}\,\pi(s)$, respectively. The bijection π abstracts from the fact that set expressions and propositional modal formulas are notational variants of each other. We write the inverse of π as π^{-1}.

A set term s is a *topological tautology* if $d(s) = U$ for all topological interpretations (U, i, d), that is, if s always denotes the entire space. A modal formula ϕ is *topologically valid* if $\pi^{-1}(\phi)$ is a topological tautology. The fundamental result that links the topological interpretation of set terms to modal logics is due to McKinsey and Tarski [7].

Theorem 4.1 (McKinsey and Tarski) *A modal propositional formula is topologically valid if and only if it is* **S4***-valid.*

5 A Deduction Theorem for Set Constraints

By Theorem 4.1, the satisfiability of an elementary constraint $s \neq \bot$ and the **S4**-satisfiability of $\pi(s)$ are equivalent properties. To make use of the theorem also for complex constraints, we prove a deduction theorem which will allow us to reduce entailment of elementary constraints to validity, and thus disentailment to satisfiability of elementary set constraints, which is equivalent to satisfiability in **S4**.

Let $\mathcal{I} = (U, i, d)$ be a topological interpretation and $U' \subseteq U$ be a nonempty subset of U. Then the *restriction* of $\mathcal{I}$ to U' is the interpretation $\mathcal{I}' = (U', i', d')$ that is defined by

- $i'(S) := i(S)$ for all $S \subseteq U'$;[2]
- $d'(X) := d(X) \cap U'$ for all variables X.

We denote the unique extension of d' to arbitrary set expressions again with the letter d'.

Lemma 5.1 *Let $\mathcal{I} = (U, i, d)$ be a topological interpretation, $U' \subseteq U$ be a nonempty subset of U, and $\mathcal{I}' = (U', i', d')$ be the restriction of $\mathcal{I}$ to U'. If U' is open, then for all set expressions s*

$$d'(s) := d(s) \cap U'.$$

Proof. The proof is by induction over the structure of set expressions. By definition of d', the claim holds for variables. The induction step for expressions of the form $\top$, $\bot$, $s \sqcap t$, $s \sqcup t$ and $\overline{s}$ involves only the properties of the Boolean connectives and is therefore omitted. We show the claim for expressions of the form $\mathtt{I}s$:

$$\begin{aligned} d'(\mathtt{I}s) &= i'(d'(s)) && (2)\\ &= i'(d(s) \cap U') && (3)\\ &= i(d(s) \cap U') && (4)\\ &= i(d(s)) \cap i(U') && (5)\\ &= i(d(s)) \cap U' && (6)\\ &= d(\mathtt{I}s) \cap U', && (7) \end{aligned}$$

where (2) holds because of the definition of d' for expressions, (3) because of the induction hypothesis, (4) because of the definition of i', (5) because of Property 3 of interior operators (see Proposition 2.1), (6) because U' is open, that is, a fixpoint of i, and (7) because of the definition of d for expressions. Note that in (6) we indeed make use of the fact that U' is open. □

For set expressions, we write $s \sqsubseteq t$ as a shorthand for $\overline{s} \sqcup t \doteq \top$.

Lemma 5.2 (Deduction Lemma) *Let s, t be set expressions. Then*

$$s \doteq \top \models t \doteq \top \qquad \text{iff} \qquad \models \mathtt{I}s \sqsubseteq t.$$

[2]It is easy to check that i' is again an interior operator.

Proof. "$\Leftarrow$" Suppose that $\mathcal{I} \models \mathtt{I}s \sqsubseteq t$ for all topological interpretations $\mathcal{I}$. Let $\mathcal{I} = (U, i, d)$ be an interpretation such that $\mathcal{I} \models s \doteq \top$, that is, $d(s) = U$. This implies $d(\mathtt{I}s) = i(d(s)) = i(U) = U$, and therefore $d(\overline{\mathtt{I}s}) = \emptyset$. Since $\mathcal{I} \models \overline{\mathtt{I}s} \sqcup t \doteq \top$, we have $d(t) = U$. Hence, $\mathcal{I} \models t \doteq \top$.
"$\Rightarrow$" Suppose that $s \doteq \top \models t \doteq \top$. We want to show that $\mathcal{I} \models \overline{\mathtt{I}s} \sqcup t \doteq \top$ for every interpretation $\mathcal{I} = (U, i, d)$. We distinguish whether $d(\mathtt{I}s)$ is empty or not. If $d(\mathtt{I}s) = \emptyset$, then $d(\overline{\mathtt{I}s}) = U$, and the claim holds. Suppose therefore that $d(\mathtt{I}s) \neq \emptyset$. Then $U' := d(\mathtt{I}s) = i(d(s))$ is a nonempty open subset of U. Consider the restriction $\mathcal{I}' = (U', i', d')$ of $\mathcal{I}$ to U'. Then we have

$$
\begin{aligned}
d'(s) &= d(s) \cap U' && (8)\\
&= d(s) \cap i(d(s)) && (9)\\
&= i(d(s)) = d(\mathtt{I}s) = U', && (10)
\end{aligned}
$$

where (8) holds because of Lemma 5.1, (9) because of the definition of U', and (10) holds, since $i(d(s)) \subseteq d(s)$ because of Property 2 of interior operators (see Proposition 2.1).
Since $d'(s) = U'$, we have that $\mathcal{I}' \models s \doteq \top$. Hence, $\mathcal{I}' \models t \doteq \top$, and therefore, $U' = d'(t) = d(t) \cap U'$, where the second equation holds because of Lemma 5.1. This implies that $d(\mathtt{I}s) = U' \subseteq d(t)$, which means that $\mathcal{I} \models \mathtt{I}s \sqsubseteq t$. □

One may conjecture that a stronger version of the Deduction Lemma holds, stating that $s \doteq \top \models t \doteq \top$ if and only if $\models s \sqsubseteq t$. However, this statement is not correct. For instance, we have that $X \doteq \top \models \mathtt{I}X \doteq \top$, because of Property 1 of interior operators in Proposition 2.1. But $X \sqsubseteq \mathtt{I}X$ does not hold in all interpretations.

Theorem 5.3 (Deduction Theorem for Topological Set Constraints) *Let s, t be set expressions. Then*

$$s \doteq \top \models t \doteq \top \qquad \text{iff} \qquad \models \mathtt{I}s \sqsubseteq \mathtt{I}t.$$

Proof. The theorem is an immediate consequence of the Deduction Lemma because $\models \mathtt{I}s \sqsubseteq \mathtt{I}t$ holds if and only if $\models \mathtt{I}s \sqsubseteq t$ holds.
The "if" statement is true because $i(T) \subseteq T$ for every set T in a topological space. The "only if" statement is true, since for sets S, T in a topological space, $i(S) \subseteq T$ implies $i(S) \subseteq i(T)$ because $i(S)$ is open and $i(T)$ is the largest open subset of T. □

6 Checking the Satisfiability of Set Constraints

We define conjunctive set constraints and translate them equivalently into sets of **S4**-formulas. Then we show that this is sufficient to decide satisfiability of arbitrary set constraints.

A set constraint is in *normal form* if every elementary constraint occurring in it is either of the form $s \doteq \top$ or $s \not\doteq \top$. Obviously, every set constraint can be rewritten in polynomial time into an equivalent one in normal form, making use of the fact that $s \doteq t$ is equivalent to $(\overline{s} \sqcup t) \sqcap (s \sqcup \overline{t}) \doteq \top$. A set constraint in normal form is *conjunctive* if it is a conjunction of elementary constraints, that is, if it has the form

$$s_1 \doteq \top \wedge \ldots \wedge s_m \doteq \top \wedge t_1 \not\doteq \top \wedge \ldots \wedge t_n \not\doteq \top. \tag{11}$$

As has been noted in [2], the negative conjuncts in a conjunctive constraint are independent of each other. That is, to determine satisfiability of the constraint it is sufficient to consider conjunctions of the positive conjuncts and one negative conjunct at a time.

Lemma 6.1 (Convexity) *The conjunctive set constraint* (11) *is satisfiable if and only for each* $j \in 1..n$ *the constraint*

$$s_1 \doteq \top \wedge \ldots \wedge s_m \doteq \top \wedge t_j \not\doteq \top \tag{12}$$

is satisfiable.

Proof. We denote the constraint (11) as C and the constraints (12) as C_j. Clearly, if C is satisfiable, then each C_j is satisfiable. Suppose, conversely, that each C_j is satisfiable. Then there are interpretations $\mathcal{I}_j = (U_j, i_j, d_j)$ such that $d_j(t_i) = U_j$ for all $i \in 1..m$ and $d_j(t_j) \neq U_j$. Without loss of generality, we can assume that the universes U_j are pairwise disjoint. We define a new interpretation $\mathcal{I} = (U, i, d)$ through

$$\begin{aligned} U &:= U_1 \cup \cdots \cup U_n \\ i(S) &:= i_1(S \cap U_1) \cup \cdots \cup i_n(S \cap U_n) \ \text{ for every } S \subseteq U, \text{ and} \\ d(X) &:= d_1(X) \cup \cdots \cup d_n(X) \ \text{ for every variable } X. \end{aligned}$$

It is easy to check that i is in fact an interior operator on U. Then $\mathcal{I}$ is an interpretation such that $\mathcal{I} \models s_i \doteq \top$ for all $i \in 1..m$ and $\mathcal{I} \models t_j \not\doteq \top$ for all $j \in 1..n$. Thus, $\mathcal{I}$ satisfies C. □

Theorem 6.2 (Translation) *The conjunctive set constraint* (11) *is satisfiable if and only for each* $j \in 1..n$ *the modal formula*

$$\mathbf{I}\,\pi(s_1) \wedge \ldots \wedge \mathbf{I}\,\pi(s_m) \wedge \neg\mathbf{I}\,\pi(t_j) \tag{13}$$

is **S4***-satisfiable.*

Proof. Let $s = s_1 \sqcap \cdots \sqcap s_m$. We observe that the constraint $s_1 \doteq \top \wedge \ldots \wedge s_m \doteq \top$ is equivalent to $s \doteq \top$, and that the formula $\mathbf{I}\,\pi(s_1) \wedge \ldots \wedge \mathbf{I}\,\pi(s_m)$ is **S4**-equivalent to $\mathbf{I}\,\pi(s)$. Thus, by Lemma 6.1 it suffices to show that $s \doteq \top \wedge t_j \not\doteq \top$ is satisfiable if and only $\mathbf{I}\,\pi(s) \wedge \neg\mathbf{I}\,\pi(t_j)$ is **S4**-satisfiable.
Now, $s \doteq \top \wedge t_j \not\doteq \top$ is satisfiable if and only if $s \doteq \top \not\models t_j \doteq \top$. By the Deduction Theorem (Theorem 5.3), this is the case if and only if $\not\models \mathtt{I}s \sqsubseteq \mathtt{I}t_j$, that is, $\overline{\mathtt{I}s} \sqcup \mathtt{I}t_j$ is not a topological tautology. Then, by McKinsey and Tarski (Theorem 4.1), $\neg\mathbf{I}\,\pi(s) \vee \mathbf{I}\,\pi(t_j)$ is not **S4**-valid, that is, $\mathbf{I}\,\pi(s) \wedge \neg\mathbf{I}\,\pi(t_j)$ is **S4**-satisfiable. □

Satisfiability checking of arbitrary constraints in normal form can be reduced to satisfiability checking of conjunctive constraints. To see this, note that we can equivalently rewrite a constraint into its disjunctive normal form, that is, into a disjunction of conjunctive constraints. The latter is satisfiable if and only if one of its disjuncts is satisfiable.

Alternatively, we can associate to each constraint a propositional formula by viewing the elementary constraints as propositions. Then we guess a truth assignment that satisfies the propositional formula. Next, we check whether our truth assignment can arise from a topological interpretation. To do so, we form a conjunction whose conjuncts are those elementary constraints for which we have guessed "true" and inverses of those for which we have guessed "false." (The inverse of $s \doteq \top$ is $s \not\doteq \top$ and vice versa.) The truth assignment can arise from a topological interpretation if and only if this conjunctive constraint is satisfiable.

From this consideration, we infer the complexity of satisfiability for arbitrary set constraints. In addition, we obtain an an upper bound for RCC8 by an argument that is much simpler than the one in [10].

Proposition 6.3 (Complexity) *Satisfiability of arbitrary topological set constraints is* PSPACE-*complete, while satisfiability of* RCC8 *constraints systems is* NP-*complete.*

Proof. Satisfiability of arbitrary set constraints is at least as hard as satisfiability in **S4**, which is PSPACE-complete [5]. On the other hand, as shown above, the satisfiability of a set constraint can be reduced in nondeterministic polynomial time to a linear number of **S4**-satisfiability tests. This gives us the upper bound. As regards RCC8, we observe that translating RCC8 constraint systems results in formulas where the maximal nesting depth of the modal operator **I** is two. Satisfiability of such formulas can be decided in nondeterministic polynomial time (cf. the algorithm in [5]). The lower bound has been proved in [10]. □

7 A Multimodal Encoding

Bennett has proposed a translation of RCC8 constraint systems into a multimodal logic with an **S4**-operator and a "strong **S5**"-operator, however, without proving the correctness of the translation [2].

We show now, more generally, that arbitrary topological set constraints can be translated into multimodal formulas that have an **S4**-operator **I** and a **K**-operator $\Box$ such that equivalence is preserved. We define the translation π as an extension of the mapping π between set expressions and **S4**-formulas by:

$$\pi(s \doteq \top) \quad := \quad \Box\mathbf{I}\,\pi(s) \tag{14}$$

$$\pi(s \doteq \bot) \quad := \quad \Box\mathbf{I}\,\neg\pi(s) \tag{15}$$

$$\pi(s \not\doteq \top) \quad := \quad \Diamond\neg\mathbf{I}\,\pi(s) \tag{16}$$

$$\pi(s \not\doteq \bot) \quad := \quad \Diamond\neg\mathbf{I}\,\neg\pi(s), \tag{17}$$

and for Boolean combinations of constraints as one would expect ($\Diamond$ is the standard shorthand for $\neg\Box\neg$). Under this mapping, the RCC8-relations are translated into the following formulas:

$$\begin{array}{rcl}
\pi(\mathsf{DC}(X,Y)) & = & \Box\mathbf{I}\,\neg(X \wedge Y) \\
\pi(\mathsf{EC}(X,Y)) & = & \Box\mathbf{I}\,\neg(\mathbf{I}\,X \wedge \mathbf{I}\,Y) \;\wedge\; \Diamond\neg\mathbf{I}\,\neg(X \wedge Y) \\
\pi(\mathsf{PO}(X,Y)) & = & \Diamond\neg\mathbf{I}\,\neg(\mathbf{I}\,X \wedge \mathbf{I}\,Y) \;\wedge \\
 & & \Diamond\neg\mathbf{I}\,\neg(X \wedge \neg Y) \;\wedge\; \Diamond\neg\mathbf{I}\,\neg(\neg X \wedge Y) \\
\pi(\mathsf{EQ}(X,Y)) & = & \Box\mathbf{I}\,(\neg X \vee Y) \;\wedge\; \Box\mathbf{I}\,(X \vee \neg Y) \\
\pi(\mathsf{TPP}(X,Y)) & = & \Box\mathbf{I}\,(\neg X \vee Y) \;\wedge\; \Diamond\neg\mathbf{I}\,\neg(X \wedge \neg\mathbf{I}\,Y) \\
\pi(\mathsf{NTPP}(X,Y)) & = & \Box\mathbf{I}\,(\neg X \vee \mathbf{I}\,Y).
\end{array}$$

The intuition behind this encoding is that we need the modality $\Box$ to appropriately combine the positive and negative constraints.

The following elementary proposition (see e.g. [5]) shows that the condition for satisfiability of a conjunction of formulas with boxes and diamonds is analogous to the one for satisfiability of a conjunction of positive and negative elementary set constraints in Lemma 6.1.

Proposition 7.1 *Let $\phi_1, \ldots, \phi_m$, $\psi_1, \ldots, \psi_n$ be multimodal formulas. Then*

$$\Box\phi_1 \wedge \cdots \wedge \Box\phi_m \wedge \Diamond\psi_1 \wedge \cdots \wedge \Diamond\psi_n$$

is satisfiable if and only if for each $j \in 1..m$ the formula

$$\phi_1 \wedge \cdots \wedge \phi_m \wedge \psi_j$$

is satisfiable.

Interestingly, Proposition 7.1 would not be true if we used **I** instead of $\Box$. Consider for instance the **S4**-formulas

$$\begin{array}{rcl}
\phi & := & (\mathbf{I}\,X \vee \mathbf{I}\,\neg X) \\
\psi_1 & := & X \\
\psi_2 & := & \neg X.
\end{array}$$

Then $\phi \wedge \psi_1 = (\mathbf{I}\,X \vee \mathbf{I}\,\neg X) \wedge X$ and $\phi \wedge \psi_2 = (\mathbf{I}\,X \vee \mathbf{I}\,\neg X) \wedge \neg X$ are both **S4**-satisfiable, but $\phi \wedge \psi_1 \wedge \psi_2 = \mathbf{I}\,(\mathbf{I}\,X \vee \mathbf{I}\,\neg X) \wedge \neg \mathbf{I}\,\neg X \wedge \neg \mathbf{I}\,\neg(\neg X)$ is not. Thus, indeed a second modal operator is needed for the translation.

Together with the Translation Theorem 6.2, Proposition 7.1 yields our claim for conjunctive constraints, from which it can generalized to arbitrary constraints.

Theorem 7.2 *Let C be a topological set constraint. Then C is satisfiable if and only if $\pi(C)$ is satisfiable.*

If we want to translate only constraints in normal form, we can base the translation function on the Deduction Lemma 5.2 instead of Theorem 5.3, because we do not have to consider negated elementary constraints. Then, instead of the four cases (15) to (17), we need only consider two, and a translation function π', that keeps satisfiability invariant, can then be defined by

$$\begin{aligned} \pi'(s \doteq \top) &:= \Box \mathbf{I}\,\pi(s) \\ \pi'(s \not\doteq \top) &:= \Diamond \neg \pi(s). \end{aligned}$$

Translating the RCC8-constraints with π' results in slightly simpler modal formulas:

$$\begin{aligned} \pi'(\mathsf{DC}(X,Y)) &= \Box \mathbf{I}\,\neg(X \wedge Y) \\ \pi'(\mathsf{EC}(X,Y)) &= \Box \mathbf{I}\,\neg(\mathbf{I}\,X \wedge \mathbf{I}\,Y) \ \wedge \ \Diamond(X \wedge Y) \\ \pi'(\mathsf{PO}(X,Y)) &= \Diamond(\mathbf{I}\,X \wedge \mathbf{I}\,Y) \ \wedge \\ & \quad \Diamond(X \wedge \neg Y) \ \wedge \ \Diamond(\neg X \wedge Y) \\ \pi'(\mathsf{EQ}(X,Y)) &= \Box \mathbf{I}\,(\neg X \vee Y) \ \wedge \ \Box \mathbf{I}\,(X \vee \neg Y) \\ \pi'(\mathsf{TPP}(X,Y)) &= \Box \mathbf{I}\,(\neg X \vee Y) \ \wedge \ \Diamond(X \wedge \neg \mathbf{I}\,Y) \\ \pi'(\mathsf{NTPP}(X,Y)) &= \Box \mathbf{I}\,(\neg X \vee \mathbf{I}\,Y). \end{aligned}$$

8 Conclusion

We have defined syntax and semantics of topological set constraints, which generalize the RCC8 constraints, a prominent formalism for qualitative spatial reasoning. For instance, we can say now that a certain region consists of at least two nonempty disconnected components, which is not possible in RCC8. We have proved that reasoning about our constraints can be reduced to reasoning in a multimodal logic with an **S4** and additional **K**-operator. Thus, we have provided for the first time a rigorous theoretical foundation for the research into topological constraint languages like RCC8.

Acknowledgement

This research was supported by the Esprit Long Term Research Project 22469 "Foundations of Data Warehouse Quality" (DWQ).

I am grateful to H.-J. Bürckert, D. Hutter, B. Nebel, and J. Renz for discussions about the subject of the paper. In particular, Nebel and Renz helped me to understand the area and to eliminate errors in earlier drafts.

References

[1] J.F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843, 1983.

[2] B. Bennett. Modal logics for qualitative spatial reasoning. *Journal of the Interest Group in Pure and Applied Logic*, 4(1):23–45, 1996.

[3] M. Grigni, D. Papadias, and Chr.H. Papadimitriou. Topological inference. *Proc. IJCAI'95*, 901–907, August 1995.

[4] V. Haarslev, C. Lutz, and R. Möller. Foundations of spatioterminological reasoning with Description Logics. *Proc. KR'98*, 112–123, June 1998.

[5] J.Y. Halpern and Y. Moses. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence*, 54(3):319–380, 1992.

[6] J.L. Kelley. *General Topology.* Van Nostrand, Princeton (New Jersey, USA), 1960.

[7] J.C.C. McKinsey and A. Tarski. The algebra of topology. *Annals of Mathematics*, 45:141–191, 1944.

[8] J.C.C. McKinsey and A. Tarski. Some theorems about the sentential calculi of Lewis and Heyting. *J. Symbolic Logic*, 13(1):1–15, 1948.

[9] D.A. Randell, Zh. Cui, and A.G. Cohn. A spatial logic based on regions and connection. In *Proc. KR'92*, 165–176, October 1992.

[10] 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.