Logical Methods in Computer Science
Vol. 6 (3:7) 2010, pp. 1–43
www.lmcs-online.org

Submitted Jun. 17, 2009
Published Aug. 18, 2010

# SPATIAL LOGICS WITH CONNECTEDNESS PREDICATES *

ROMAN KONTCHAKOV [a], IAN PRATT-HARTMANN [b], FRANK WOLTER [c], AND MICHAEL ZAKHARYASCHEV [d]

[a] Department of Computer Science and Information Systems, Birkbeck College London
*e-mail address*: roman@dcs.bbk.ac.uk

[b] Department of Computer Science, Manchester University
*e-mail address*: ipratt@cs.man.ac.uk

[c] Department of Computer Science, University of Liverpool
*e-mail address*: frank@csc.liv.ac.uk

[d] Department of Computer Science and Information Systems, Birkbeck College London
*e-mail address*: michael@dcs.bbk.ac.uk

Abstract. We consider quantifier-free spatial logics, designed for qualitative spatial representation and reasoning in AI, and extend them with the means to represent topological connectedness of regions and restrict the number of their connected components. We investigate the computational complexity of these logics and show that the connectedness constraints can increase complexity from NP to PSpace, ExpTime and, if component counting is allowed, to NExpTime.

## 1. Introduction

The field of Artificial Intelligence known as *qualitative spatial reasoning* is concerned with the problem of representing and manipulating spatial information about everyday, middle-sized entities. In recent decades, much activity in this field has centred on *spatial logics*—formal languages whose variables range over geometrical objects (not necessarily points), and whose non-logical primitives represent geometrical relations and operations involving those objects. (For a recent survey, see [10].) The hope is that, by using a formalism couched entirely at the level of these geometrical objects, we can avoid the expressive—hence computationally expensive—logical machinery required to reconstruct them in terms of sets of points.

What might such qualitative spatial relations typically be? Probably the most intensively studied collection is the set of six topological relations illustrated—for the case of closed disc-homeomorphs in $\mathbb{R}^2$—in Fig. 1. These relations—$\mathsf{DC}$ (disconnection), $\mathsf{EC}$ (external connection), $\mathsf{PO}$ (partial overlap), $\mathsf{EQ}$ (equality), $\mathsf{TPP}$ (tangential proper part) and $\mathsf{NTPP}$ (non-tangential proper part)—were popularized in the seminal treatments of spatial

*1998 ACM Subject Classification:* F.4.1, I.2.4.
*Key words and phrases:* Spatial logic, topological space, connectedness, computational complexity.
* Some results of the paper were presented at LPAR 2008 and ECAI 2000.

LOGICAL METHODS IN COMPUTER SCIENCE DOI:10.2168/LMCS-6 (3:7) 2010

logics by Egenhofer and Franzosa [17] and Randell *et al.* [42]. Counting the converses of

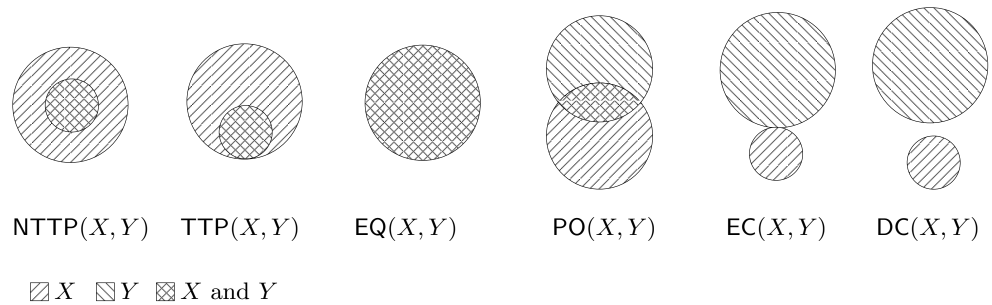



Figure 1: The $\mathcal{RCC}$-8-relations illustrated for disc-homeomorphs in $\mathbb{R}^2$

the two asymmetric relations $\mathsf{TPP}$ and $\mathsf{NTPP}$, the resulting eight relations are frequently referred to under the moniker $\mathcal{RCC}$-8 (for *region connection calculus*). To see how this collection of relations gives rise to a spatial logic, let $r_1$, $r_2$ and $r_3$ be disc-homeomorphs in $\mathbb{R}^2$, and suppose that $r_1$, $r_2$ stand in the relation $\mathsf{TPP}$, while $r_1$, $r_3$ stand in the relation $\mathsf{NTPP}$. A little experimenting with diagrams suffices to show that $r_2$, $r_3$ must stand in one of the three relations $\mathsf{PO}$, $\mathsf{TPP}$ or $\mathsf{NTPP}$. As we might put it, the $\mathcal{RCC}$-8-formula

$$\big(\mathsf{TPP}(r_1, r_2) \land \mathsf{NTPP}(r_1, r_3)\big) \to \big(\mathsf{PO}(r_2, r_3) \lor \mathsf{TPP}(r_2, r_3) \lor \mathsf{NTPP}(r_2, r_3)\big)$$

is valid over the spatial domain of disc-homeomorphs in the plane: all assignments of such regions to the variables $r_1$, $r_2$ and $r_3$ make it true. Similar experimentation shows that, by contrast, the formula

$$\mathsf{TPP}(r_1, r_2) \land \mathsf{NTPP}(r_1, r_3) \land \mathsf{EC}(r_2, r_3),$$

is *unsatisfiable*: no assignments of disc-homeomorphs to $r_1$, $r_2$ and $r_3$ make this formula true.

More generally, let $\mathcal{L}$ be a formal language featuring some collection of predicates and function symbols having (fixed) interpretations as geometrical relations and operations. The formulas of $\mathcal{L}$ may then be interpreted over any collection of subsets of some space $T$ for which the relevant geometrical notions make sense: we refer to the elements of such a domain of interpretation as *regions*. Let $\mathcal{K}$ be a class of domains of interpretation for $\mathcal{L}$. The notion of the *satisfaction* of an $\mathcal{L}$-formula by a tuple of regions, and, derivatively, the notions of *satisfiability* and *validity* of an $\mathcal{L}$-formula with respect to $\mathcal{K}$, can then be understood in the usual way. We call the pair $(\mathcal{L}, \mathcal{K})$ a *spatial logic*. If all the primitives of $\mathcal{L}$ are topological in character—as in the case of $\mathcal{RCC}$-8—we speak of a *topological logic*. For languages featuring negation, the notions of satisfiability and validity are dual in the usual sense. The primary question arising in connection with any spatial logic is: how do we recognize the satisfiable (dually, the valid) formulas? From an algorithmic point of view, we are particularly concerned with the decidability and complexity of these problems.

A second example will make this abstract characterization more concrete. Constraints featuring $\mathcal{RCC}$-8 predicates give us no means to *combine* regions into new ones; and it is natural to ask what happens when this facility is provided. Let $T$ be a topological space. A subset of $T$ is *regular closed* if it is the topological closure of an open set in $T$. The collection of regular closed sets forms a Boolean algebra with binary operations $+$, $\cdot$ and a unary operation $-$. Intuitively, we are to think of $r_1 + r_2$ as the *agglomeration* of $r_1$ and

$r_2$, $r_1 \cdot r_2$ as the *common part* of $r_1$ and $r_2$, and $-r$ as the *complement* of $r$. Further, the $\mathcal{RCC}$-8-relations illustrated above can be generalized, in a natural way, so that they apply to regular closed subsets of any topological space $T$. (Details are given below.) By augmenting the language $\mathcal{RCC}$-8 with the function symbols $+$, $\cdot$ and $-$, we obtain the more expressive formalism originally introduced in [57] under the name $\mathcal{BRCC}$-8 (*Boolean* $\mathcal{RCC}$-8), which we may interpret over any algebra of regular closed subsets of some topological space. Again, we can ask which of these formulas are satisfiable or valid over the class of domains in question. For instance, the formula

$$\mathsf{EC}(r_1 + r_2, r_3) \to \big(\mathsf{EC}(r_1, r_3) \vee \mathsf{EC}(r_2, r_3)\big),$$

asserting that, if the sum of two regions $r_1$ and $r_2$ stands in the relation of external connection to a region $r_3$, then one of the summands does as well, turns out to be valid; by contrast, the formula

$$\mathsf{EC}(r_1, r_2) \wedge \mathsf{EC}(r_1, -r_2),$$

asserting that $r_1$ is externally connected to both $r_2$ and its complement, is unsatisfiable.

Arguably, the topological primitive with the longest pedigree in the spatial logic literature is the relation now generally referred to as $C$ (for *contact*). Intuitively, two regions are to be thought of as being in contact just in case they either overlap or have touching boundaries. This relation was originally introduced by Whitehead ([56], pp. 294, ff.) under the name *extensive connection*, and formed the starting point for his region-based reconstruction of space. More recently, it has been studied within the framework of *Boolean contact algebras* [12, 16]. It turns out that, in the presence of the operations $+$, $\cdot$ and $-$, all of the $\mathcal{RCC}$-8-relations can be expressed in terms of the relations of equality and contact, and *vice versa*. Accordingly, and in order to unify these two lines of research, we shall denote the language $\mathcal{BRCC}$-8 by $\mathcal{C}$ in this paper.

One familiar topological property that has been notable by its absence from the spatial logic literature, however, is *connectedness* (or, as it is occasionally called, 'self-connectedness' [6]). This lacuna is particularly surprising given the recognized significance of this concept in qualitative spatial reasoning [10]. The availability of connectedness as a primitive relation greatly expands the expressive power of topological logics, and in particular increases their sensitivity to the underlying domain of quantification. For example, let the connectedness predicate $c$ be added to the language $\mathcal{RCC}$-8, yielding the language $\mathcal{RCC}$-8$c$; and consider the $\mathcal{RCC}$-8$c$-formula

$$\bigwedge_{1 \leq i \leq 3} c(r_i) \wedge \bigwedge_{1 \leq i < j \leq 3} \mathsf{EC}(r_i, r_j).$$

This formula states that regions $r_1$, $r_2$ and $r_3$ are connected, and that any two of them touch at their boundaries without overlapping. It is easily seen to be satisfiable over the domain of regular closed sets in $\mathbb{R}^2$; however, it is not satisfiable over the domain of regular closed sets in $\mathbb{R}$. For a non-empty, regular closed subset of $\mathbb{R}$ is connected if and only if it is a non-punctual, closed interval (possibly unbounded); and it is obvious that no three such intervals can touch in pairs without overlapping. More tellingly, consider the $\mathcal{RCC}$-8$c$-formula

$$c(r_1) \wedge \bigwedge_{1 \leq i < j \leq 4} \mathsf{EC}(r_i, r_j),$$

stating that $r_1$ is connected, and that any two of $r_1, \ldots, r_4$ touch at their boundaries without overlapping. This formula *is* satisfiable over the regular closed subsets of $\mathbb{R}$, as shown in Fig. 2. However, such an arrangement is only possible provided at least two of the regions

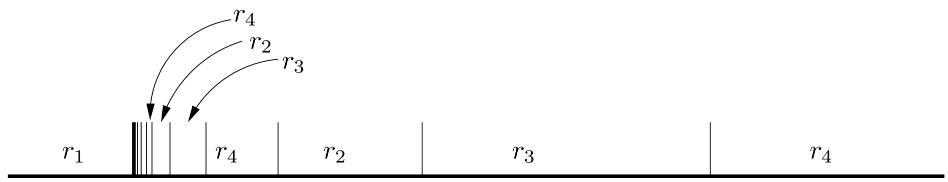



Figure 2: A configuration of regular closed regions in $\mathbb{R}$ satisfying the $\mathcal{RCC}$-8$c$-formula $c(r_1) \wedge \bigwedge_{1 \leq i < j \leq 4} \mathsf{EC}(r_i, r_j)$: the region $r_1$ is connected, while $r_2$, $r_3$ and $r_4$ have infinitely many components, with a common accumulation point on the boundary of the $r_1$-region.

$r_2$, $r_3$ and $r_4$ have infinitely many components. If—for whatever reason—our spatial ontology does not countenance regions with infinitely many components, the formula becomes unsatisfiable. Thus, simple logics featuring the connectedness predicate are sensitive to the underlying topological space, and indeed to the choice of subsets of that space we count as regions. By contrast, we shall see below that topological logics lacking the connectedness predicate—even ones much more expressive than $\mathcal{RCC}$-8—are remarkably insensitive to the spatial domains over which they are interpreted. More generally, the above examples give a hint of the interesting mathematical challenges which the property of connectedness presents us with in the context of almost any topological logic.

It is surprising that only sporadic attempts have been made to investigate the expressive power and computational complexity of topological logics able to talk about the connectedness of regions [8, 51, 57, 40]. The present paper rectifies this omission by introducing the unary predicates $c$ and $c^{\leq k}$ (for $k \geq 1$). We read $c(r)$ as 'region $r$ is connected' and $c^{\leq k}(r)$ as 'region $r$ has at most $k$ connected components.' Our aim is to provide a systematic study of the impact of these predicates on the computational complexity of the satisfiability problem for topological logics. We restrict attention in this paper to *quantifier-free* languages—i.e. those in which formulas are Boolean combinations of atomic formulas—in line with the constraint satisfaction approach of [47]—since first-order spatial logics are generally undecidable [27, 14, 11, 36].[1] For an overview of first-order topological logics, see [41].

Specifically, we consider three principal *base languages*, characterized by various collections of topological primitives, and investigate the effect of augmenting each of these base languages with the predicates $c$ and $c^{\leq k}$. The weakest of these base languages, denoted $\mathcal{B}$, features only the region-combining operators $+$, $\cdot$ and $-$, together with the equality predicate. Thus, $\mathcal{B}$ is essentially just the language of Boolean algebra equations: as such, this language can express no really characteristic topological properties; further, its satisfiability problem, when interpreted over the class of regular closed algebras of topological spaces, is easily seen to be NP-complete. If, however, we add the connectedness predicate $c$, we obtain the language $\mathcal{B}c$—a fully-fledged topological logic able to simulate (in a sense explained below) the contact relation $C$, and hence all the $\mathcal{RCC}$-8-relations. More ambitiously, we can

[1]One of the notable exceptions in this regard is Tarski's theory of *elementary geometry*, which can be regarded as a first-order spatial logic whose domain of interpretation is the set of points in the Euclidean plane. The precise computational complexity of this logic—essentially the first-order fragment of the system set out by Hilbert [29]—is still unknown, with the current lower bound being NExpTime [18] and the upper bound ExpSpace [4].

add to $\mathcal{B}$ all of the predicates $c^{\leq k}$ (for $k \geq 1$), to obtain the language $\mathcal{B}cc$. An indication of the resulting increase in expressiveness is that the satisfiability problem for the same class of interpretations jumps from NP to ExpTime (in the case of $\mathcal{B}c$) and NExpTime (in the case of $\mathcal{B}cc$).

Our next base language is $\mathcal{C}$ (alias $\mathcal{BRCC}$-8), which we encountered above. When interpreted over the class of all regular closed algebras of topological spaces, the satisfiability problem for this language is still NP-complete [57]. By extending $\mathcal{C}$ with the predicates $c$ and $c^{\leq k}$ (for $k \geq 1$), however, we obtain the more expressive languages $\mathcal{C}c$ and $\mathcal{C}cc$, whose satisfiability problems for the same class of interpretations again jump from NP to ExpTime and NExpTime, respectively.

Our final base language has its roots in the seminal paper by McKinsey and Tarski [37]. Following the modal logic tradition, we call it $\mathcal{S}4_u$, that is, Lewis' system $\mathcal{S}4$ extended with the universal modality. (For more information on the relationship between spatial and modal logic see [54, 21] and references therein.) The variables of this language may be taken to range over any collection of subsets of a topological space (not just regular closed sets), and its primitives include the operations of union, intersection, complement and topological interior and closure. Since the property of being regular closed is expressible in $\mathcal{S}4_u$, this language may be regarded as being more expressive than $\mathcal{C}$. When interpreted over the class of power sets of topological spaces, the satisfiability problem for $\mathcal{S}4_u$ is PSpace-complete. By extending $\mathcal{S}4_u$ with the predicates $c$ and $c^{\leq k}$ (for $k \geq 1$), however, we obtain the languages $\mathcal{S}4_uc$ and $\mathcal{S}4_ucc$, whose satisfiability problems, for the same class of interpretations, once again jump to ExpTime and NExpTime, respectively.

Thus, the addition of connectedness predicates to topological logics leads to greater expressive power and higher computational complexity. However, this increase in complexity is 'stable': over the most general classes of interpretations, the extensions of such different formalisms as $\mathcal{B}$ and $\mathcal{S}4_u$ with connectedness predicates are of the same complexity. Another interesting result is that, by restricting these languages to formulas with just one connectedness constraint of the form $c(r)$, we obtain logics that are still in PSpace, while two such constraints lead to ExpTime-hardness. In fact, if the connectedness predicate is applied only to regions that are known to be pairwise disjoint, then it does not matter how many times this predicate occurs in the formula: satisfiability is still in PSpace.

The rest of this paper is organized as follows. Section 2 presents the syntax and semantics of our base languages (together with some of their variants), and Section 3 extends these languages with connectedness predicates. Section 4 introduces the first main ingredient of our proofs—a representation theorem allowing us to work with Aleksandrov topological spaces rather than arbitrary ones. Such spaces can be represented by Kripke frames with quasi-ordered accessibility relations, and topological connectedness in these frames corresponds to graph-theoretic connectedness in the (non-directed) graphs induced by the accessibility relation. Based on this observation, we can prove the upper bounds in a more-or-less standard way using known techniques from modal and description logic; by contrast, the lower bounds are more involved and unexpected. Section 5 presents the proofs of these complexity results. Section 6 considers the computational behaviour of our topological logics when interpreted over various *Euclidean* spaces $\mathbb{R}^n$, and lists some open problems.

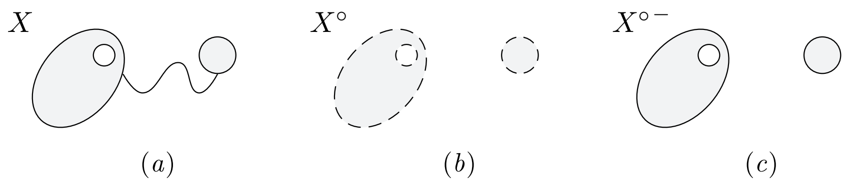



Figure 3: $(a)$ A closed set in the plane, $(b)$ its interior and $(c)$ the closure of its interior. Note that $X^\circ \subseteq X^{\circ -} \subseteq X$.

## 2. Background: topological logics without connectedness

A *topological space* is a pair $(T, \mathcal{O})$, where $T$ is a set and $\mathcal{O}$ a collection of subsets of $T$ containing $\emptyset$ and $T$, and closed under arbitrary unions and finite intersections. The elements of $\mathcal{O}$ are referred to as *open sets*; their complements are *closed* sets. If $\mathcal{O}$ is clear from context, we refer to the topological space $(T, \mathcal{O})$ simply as $T$. Given any $X \subseteq T$, the *interior* of $X$, denoted $X^\circ$, is the largest open set included in $X$, and the *closure* of $X$, denoted $X^-$, is the smallest closed set including $X$. These sets always exist. It is convenient, where the space $T$ is clear from context, to denote $T$ by $\mathbf{1}$, the empty set by $\mathbf{0}$, and, for any $X \subseteq T$, the complement $T \setminus X$ by $\overline{X}$. Evidently, $X^- = \overline{((\overline{X})^\circ)}$. If $X \subseteq T$, the *subspace topology* on $X$ is the collection of sets $\mathcal{O}_X = \{O \cap X \mid O \in \mathcal{O}\}$. It is readily checked that $(X, \mathcal{O}_X)$ is a topological space.

Let $T$ be a topological space. A subset of $T$ is called *regular closed* if it is the closure of an open set. We denote the set of regular closed subsets of $T$ by $\mathrm{RC}(T)$. It is a standard result (for example, [31], pp. 25–27) that, for any topological space $T$, the collection of sets $\mathrm{RC}(T)$ forms a Boolean algebra, with the top and bottom elements $\mathbf{1} = T$ and $\mathbf{0} = \emptyset$, respectively, Boolean operations given by

$$X + Y = X \cup Y, \qquad X \cdot Y = (X \cap Y)^{\circ -}, \qquad -X = (\overline{X})^-, \tag{2.1}$$

and Boolean order $\leq$ coinciding with the subset relation. In the context of the Euclidean plane $\mathbb{R}^2$, the regular closed sets are—roughly speaking—those closed sets with no 'filaments' or 'isolated points' (Fig. 3). When dealing with the Boolean algebra $\mathrm{RC}(T)$, for some topological space $T$, we generally write $X + Y$ in preference to $X \cup Y$ (though these are, formally, equivalent); similarly, we generally write $X \leq Y$ in preference to $X \subseteq Y$.

We establish a general framework for defining the topological languages studied in this paper. Fix a countably infinite set $\mathcal{R}$. We refer to the elements of $\mathcal{R}$ as *region variables* (or, more simply: *variables*) and denote them by $r$, $s$, etc. possibly with sub- or superscripts. Let $F$ be any set of function symbols (of fixed arities) and $P$ any set of predicate symbols (of fixed arities). In practice, the symbols in $F$ and $P$ may be assumed to have fixed topological interpretations, along the lines indicated in Section 1. For example, $F$ might contain function symbols denoting the operations $+$, $\cdot$ and $-$ on regular closed sets defined in (2.1); likewise, $P$ might contain predicates denoting the $\mathcal{RCC}$-8 relations. The $\mathcal{L}(F, P)$-*terms*, $\tau$, are given by the rule:

$$\tau \quad ::= \quad r \quad | \quad f^n(\tau_1, \ldots, \tau_n),$$

where $r$ is a variable in $\mathcal{R}$, $f^n$ a function symbol of arity $n$ in $F$, and the $\tau_i$ $\mathcal{L}(F, P)$-terms. The $\mathcal{L}(F, P)$-*formulas*, $\varphi$, are given by the rule:

$$\varphi \quad ::= \quad p^n(\tau_1, \ldots, \tau_n) \quad | \quad \varphi_1 \wedge \varphi_2 \quad | \quad \varphi_1 \vee \varphi_2 \quad | \quad \varphi_1 \rightarrow \varphi_2 \quad | \quad \neg\varphi,$$

where $p^n$ is a predicate symbol of arity $n$ in $P$, and $\varphi_1$, $\varphi_2$ are $\mathcal{L}(F,P)$-formulas. The *topological language* $\mathcal{L}(F,P)$ is the set of $\mathcal{L}(F,P)$-formulas. We shall write $\mathcal{L}$ in place of $\mathcal{L}(F,P)$ if $F$ and $P$ are understood. As usual, formulas of the form $p^n(\tau_1,\ldots,\tau_n)$ and $\neg p^n(\tau_1,\ldots,\tau_n)$ are called *literals*. Notice that, for the purposes of this paper, topological languages involve no quantifiers.

We now turn to the semantics of these languages. A *topological frame* is a pair of the form $(T,S)$, where $T$ is a topological space, and $S \subseteq 2^T$. We refer to the elements of $S$ as *regions*; there is no requirement for $S$ to be closed under any operations. A *topological model* on $(T,S)$ is a triple $\mathfrak{M} = (T,S,\cdot^{\mathfrak{M}})$, where $\cdot^{\mathfrak{M}}$ is a map from $\mathcal{R}$ to $S$, referred to as a *valuation*. Assuming the function symbols in $F$ and predicates in $P$ to have standard interpretations, any topological model $\mathfrak{M}$ determines the truth-value of an $\mathcal{L}$-formula in the obvious way. We write $\mathfrak{M} \models \varphi$ if the formula $\varphi$ is true in $\mathfrak{M}$.

Let $\mathcal{K}$ be a class of topological frames and $\varphi$ a formula of a topological language $\mathcal{L}$. We say that $\varphi$ is *satisfiable* over $\mathcal{K}$ if $\mathfrak{M} \models \varphi$ for some topological frame $(T,S)$ in $\mathcal{K}$ and some topological model $\mathfrak{M}$ on $(T,S)$; dually, $\varphi$ is *valid* over $\mathcal{K}$ if $\mathfrak{M} \models \varphi$ for every topological frame $(T,S)$ in $\mathcal{K}$ and every topological model $\mathfrak{M}$ on $(T,S)$. As usual, $\varphi$ is valid if and only if $\neg\varphi$ is not satisfiable. The *satisfiability problem* for $\mathcal{L}$-formulas over topological frames in $\mathcal{K}$ is the decision problem for the set

$$\mathit{Sat}(\mathcal{L},\mathcal{K}) \;=\; \bigl\{\varphi \in \mathcal{L} \mid \varphi \text{ is satisfiable over } \mathcal{K}\bigr\},$$

that is: given an $\mathcal{L}$-formula $\varphi$, decide whether it is satisfiable in a topological model based on a topological frame from $\mathcal{K}$. A *topological logic* is a pair $(\mathcal{L},\mathcal{K})$, where $\mathcal{L}$ is a topological language (whose primitives are taken to have fixed topological interpretations) and $\mathcal{K}$ a class of topological frames.

In the sequel, except where indicated to the contrary, we generally speak of *frames*, *models*, *logics* etc., taking the qualifier 'topological' to be implicit.

The primary motivation for introducing the notion of a frame $(T,S)$ is to provide a mechanism for confining attention to those subsets $S$ of the space $T$ which we regard as *bona fide* regions. For example, it is frequently observed (see, e.g., [22]) that no clear sense can be given to the question of whether a given physical object occupies a topologically closed, semi-closed or open region of space. Consequently, spatial logics in AI conventionally identify regions differing only with respect to boundary points. A convenient way to finesse the issue of boundary points in a topological space $T$ is to restrict attention to the regular closed sets $\mathrm{RC}(T)$. For, given any closed subset $X$ of $T$, there exists a unique $Y \in \mathrm{RC}(T)$ such that $X^\circ \subseteq Y \subseteq X$ (see Fig. 3). Moreover, these regular closed sets, as noted above, form a Boolean algebra with $+$, $\cdot$ and $-$ providing reasonable reconstructions of the intuitive operations of agglomeration, intersection, and complementation, respectively. In this paper, we shall be principally concerned with the classes of frames $\textsc{All}$ and $\textsc{RegC}$ given by

$$\textsc{All} = \{(T,2^T) \mid T \text{ a topological space}\},$$
$$\textsc{RegC} = \{(T,\mathrm{RC}(T)) \mid T \text{ a topological space}\}.$$

One word of caution: the Boolean algebras $\mathrm{RC}(\mathbb{R}^2)$ and $\mathrm{RC}(\mathbb{R}^3)$ include many sets which are not at all obviously suited to model regions occupied by physical objects. For this reason, we may decide to interpret our languages over topological frames $(T,S)$ where $S$ is a *sub-algebra* of $\mathrm{RC}(T)$, a restriction which turns out to have interesting mathematical consequences (see, e.g., [41]).

With these semantic preliminaries behind us, we now survey some of the most familiar topological logics occurring in the AI literature. This survey will occupy the remainder of this section. Remember: our aim in the sequel is to investigate the effect of increasing the expressive resources available to these logics by adding predicates expressing connectedness and related notions.

**The logic $\mathcal{RCC}$-8:** We begin with a formal account of the topological language $\mathcal{RCC}$-8, which we encountered in Section 1. As a preliminary, we define the following six binary relations on $\mathrm{RC}(T)$, where $T$ is any topological space:

$$\begin{array}{rll}
\mathsf{DC}(X,Y) & \text{iff} & X \cap Y = \emptyset, \\
\mathsf{EC}(X,Y) & \text{iff} & X \cap Y \neq \emptyset \text{ but } X^\circ \cap Y^\circ = \emptyset, \\
\mathsf{PO}(X,Y) & \text{iff} & X^\circ \cap Y^\circ,\ X^\circ \setminus Y \text{ and } Y^\circ \setminus X \text{ are all non-empty}, \\
\mathsf{EQ}(X,Y) & \text{iff} & X = Y, \\
\mathsf{TPP}(X,Y) & \text{iff} & X \subseteq Y \text{ but } X \not\subseteq Y^\circ \text{ and } Y \not\subseteq X, \\
\mathsf{NTPP}(X,Y) & \text{iff} & X \subseteq Y^\circ \text{ but } Y \not\subseteq X.
\end{array} \tag{2.2}$$

All of these relations except $\mathsf{TPP}$ and $\mathsf{NTPP}$ are symmetric. Counting the converses of $\mathsf{TPP}$ and $\mathsf{NTPP}$, we thus obtain eight binary relations altogether: these eight relations are easily seen to be jointly exhaustive and mutually exclusive over non-empty elements of $\mathrm{RC}(T)$. In Fig. 1, we illustrated them in the special case where the relata are closed disc-homeomorphs in the plane. We remark in passing that, when restricted to closed disc-homeomorphs in the plane, these relations are actually the atoms of a finite relation algebra [15, 35].

We now define the language $\mathcal{RCC}$-8 by

$$\mathcal{RCC}\text{-}8 \ = \ \mathcal{L}(\emptyset, \{\mathsf{DC}, \mathsf{EC}, \mathsf{PO}, \mathsf{EQ}, \mathsf{TPP}, \mathsf{NTPP}\}),$$

with the symbols $\mathsf{DC}$, $\mathsf{EC}$, $\mathsf{PO}$, $\mathsf{EQ}$, $\mathsf{TPP}$, $\mathsf{NTPP}$ taken to be binary predicates. There are no function symbols; so $\mathcal{RCC}$-8-terms are simply variables.

We always interpret this language over (some sub-class of) RegC—that is: variables are always taken to range over (certain) regular closed sets of (certain) topological spaces. The semantics for $\mathcal{RCC}$-8 may then be given by specifying the interpretations of the predicates in obvious way, thus:

$$\begin{array}{lll}
\mathfrak{M} \models \mathsf{DC}(r_1, r_2) & \text{iff} & \mathsf{DC}(r_1^{\mathfrak{M}}, r_2^{\mathfrak{M}}), \\
\mathfrak{M} \models \mathsf{EC}(r_1, r_2) & \text{iff} & \mathsf{EC}(r_1^{\mathfrak{M}}, r_2^{\mathfrak{M}}), \\
\textit{etc.} & &
\end{array}$$

Note the overloading of the symbols $\mathsf{DC}$, $\mathsf{EC}$, etc. here: on the left-hand sides of these equations, they are predicates of $\mathcal{RCC}$-8; on the right-hand side, they denote the relations on $\mathrm{RC}(T)$ defined in (2.2). Since these predicates will always be used with their standard meanings, no confusion need arise. In the literature, the language $\mathcal{RCC}$-8 is sometimes subject to the additional restriction that variables range only over *non-empty* regular closed subsets of the space in question. We do not impose this requirement, remarking, however, that non-emptiness is anyway expressible in $\mathcal{RCC}$-8 (on our interpretation) by adding conjuncts of the form $\neg\mathsf{DC}(r, r)$.

It is known that $\mathit{Sat}(\mathcal{RCC}\text{-}8, \textsc{RegC})$ is NP-complete [43]. (Proofs of all the complexity results mentioned in this section are discussed in Section 4.)

**The logic $\mathcal{BRCC}$-8:** Since, as we have observed, the regular closed subsets of a topological space form a Boolean algebra, it makes sense to augment $\mathcal{RCC}$-8 with function symbols denoting the obvious operations and constants of this Boolean algebra. The language $\mathcal{BRCC}$-8 (for *Boolean* $\mathcal{RCC}$-8, [57]) is defined by:

$$\mathcal{BRCC}\text{-}8 \;=\; \mathcal{L}(\{+,\cdot,-,\mathbf{0},\mathbf{1}\},\{\mathsf{DC},\mathsf{EC},\mathsf{PO},\mathsf{EQ},\mathsf{TPP},\mathsf{NTPP}\}),$$

where the function symbols $+$ and $\cdot$ are binary, $-$ is unary, and $\mathbf{0}$ and $\mathbf{1}$ are nullary (i.e. individual constants). Again, we confine attention to the class of frames RegC, with the function symbols interpreted as the obvious operations on regular closed sets. Formally:

$$\begin{array}{rclcrclcrcl}
(\tau_1+\tau_2)^{\mathfrak{M}} & = & \tau_1^{\mathfrak{M}}+\tau_2^{\mathfrak{M}}, & \qquad & (-\tau)^{\mathfrak{M}} & = & -(\tau^{\mathfrak{M}}), & \qquad & \mathbf{0}^{\mathfrak{M}} & = & \mathbf{0},\\
(\tau_1\cdot\tau_2)^{\mathfrak{M}} & = & \tau_1^{\mathfrak{M}}\cdot\tau_2^{\mathfrak{M}}, & & & & & & \mathbf{1}^{\mathfrak{M}} & = & \mathbf{1}.
\end{array}$$

Note the overloading of the symbols: on the left-hand side of these equations, they are function symbols of $\mathcal{BRCC}$-8; on the right-hand side, they denote the corresponding Boolean algebra operations in $\mathrm{RC}(T)$ as defined in (2.1). The predicates are interpreted in the same way as for $\mathcal{RCC}$-8.

Despite its increased expressive power, $\mathcal{BRCC}$-8 is in the same complexity class as $\mathcal{RCC}$-8—at least when interpreted over arbitrary topological spaces. That is: the problem $Sat(\mathcal{BRCC}\text{-}8,\textsc{RegC})$ is NP-complete [57]. (However, as we shall see below, this situation changes even under very mild restrictions on the class of frames.)

**The logic $\mathcal{C}$:** We can re-formulate $\mathcal{BRCC}$-8 more elegantly using the binary predicates $=$ (equality) and $C$ (*contact*). The language $\mathcal{C}$ is defined by:

$$\mathcal{C}=\mathcal{L}(\{+,\cdot,-,\mathbf{0},\mathbf{1}\},\{C,=\}).$$

As with $\mathcal{RCC}$-8 and $\mathcal{BRCC}$-8, we confine our attention to the class of frames RegC. The equality predicate $=$ denotes identity (as usual), and the contact predicate $C$ is interpreted as follows:

$$\mathfrak{M}\models C(\tau_1,\tau_2) \quad \text{iff} \quad \tau_1^{\mathfrak{M}}\cap\tau_2^{\mathfrak{M}}\neq\emptyset.$$

That is: two regions are taken to be in contact just in case they intersect. Notice that $C(\tau_1,\tau_2)$ is not equivalent to the condition $\tau_1\cdot\tau_2\neq\mathbf{0}$ (a shorthand for $\neg(\tau_1\cdot\tau_2=\mathbf{0})$), which states that the *interiors* of $\tau_1$ and $\tau_2$ intersect. In the context of any logic involving the function symbols $+$, $\cdot$, $-$, $\mathbf{0}$, $\mathbf{1}$ and the equality predicate, we standardly write $\tau_1\leq\tau_2$ as an abbreviation for $\tau_1\cdot(-\tau_2)=\mathbf{0}$; $\neg(\tau_1\leq\tau_2)$ is abbreviated by $\tau_1\not\leq\tau_2$.

Evidently, $C(\tau_1,\tau_2)$ is equivalent to $\neg\mathsf{DC}(\tau_1,\tau_2)$, and $=$ is just another symbol for $\mathsf{EQ}$; hence, $\mathcal{BRCC}$-8 is at least as expressive as $\mathcal{C}$. Conversely, it is easy to verify that the four remaining $\mathcal{RCC}$-8-relations can easily be equivalently expressed in $\mathcal{C}$, as follows:

$$\begin{array}{rcl}
\mathsf{EC}(\tau_1,\tau_2) & \leftrightarrow & (\tau_1\cdot\tau_2=\mathbf{0})\ \wedge\ C(\tau_1,\tau_2),\\
\mathsf{PO}(\tau_1,\tau_2) & \leftrightarrow & (\tau_1\cdot\tau_2\neq\mathbf{0})\ \wedge\ (\tau_1\not\leq\tau_2)\ \wedge\ (\tau_2\not\leq\tau_1),\\
\mathsf{TPP}(\tau_1,\tau_2) & \leftrightarrow & (\tau_1\leq\tau_2)\ \wedge\ C(\tau_1,-\tau_2)\ \wedge\ (\tau_2\not\leq\tau_1),\\
\mathsf{NTPP}(\tau_1,\tau_2) & \leftrightarrow & \neg C(\tau_1,-\tau_2)\ \wedge\ (\tau_2\not\leq\tau_1).
\end{array}$$

Hence, we may regard the languages $\mathcal{C}$ and $\mathcal{BRCC}$-8 as equivalent.

The predicate $C$ has an interesting history. Originally introduced by Whitehead [56] under the name 'extensive connection,' it provided the inspiration for many of the early approaches to topological logics in AI. To avoid confusion with the familiar topological property of *connectedness*, Whitehead's relation is now generally referred to as *contact*. Investigation of the contact-structure of the regular closed algebras of topological spaces gave rise to the study of so-called *Boolean connection algebras* (BCAs). The relationship between BCAs and the topological spaces that generate them is now well-understood [49, 16, 12, 13]. Our logic $\mathcal{C}$ (that is: $\mathcal{BRCC}$-8) is, in essence, the quantifier-free fragment of the first-order theory of BCAs. For an up-to-date account of this work, see [3].

**The logic $\mathcal{C}^m$:** As formulas in $\mathcal{C}$ are built from $\mathcal{C}$-terms using the binary predicates $\tau_1 = \tau_2$ and $C(\tau_1, \tau_2)$, they are not capable of expressing, for example, the predicate $\mathsf{EC}_3(\tau_1, \tau_2, \tau_3)$ stating that *three* regions $\tau_1, \tau_2, \tau_3$ are externally connected and have some common border. One way to extend the expressive power of $\mathcal{C}$ is to generalize the contact predicate and consider its extension $\mathcal{C}^m$ with arbitrary $k$-ary contact relations $C^k(\tau_1, \dots, \tau_k)$, for $k \geq 2$. The language $\mathcal{C}^m$ is defined by:

$$\mathcal{C}^m \;=\; \mathcal{L}(\{+, \cdot, -, \mathbf{0}, \mathbf{1}\}, \{C^k \mid k \geq 2\} \cup \{=\}).$$

Again, we confine attention to the class of frames RegC. The predicates $C^k$ are interpreted as follows:

$$\mathfrak{M} \models C^k(\tau_1, \dots, \tau_k) \quad \text{iff} \quad \tau_1^{\mathfrak{M}} \cap \dots \cap \tau_k^{\mathfrak{M}} \neq \emptyset.$$

The ternary predicate $\mathsf{EC}_3(\tau_1, \tau_2, \tau_3)$ above can now be expressed in a straightforward way:

$$\mathsf{EC}_3(\tau_1, \tau_2, \tau_3) \;=\; C^3(\tau_1, \tau_2, \tau_3) \wedge \mathsf{EC}(\tau_1, \tau_2) \wedge \mathsf{EC}(\tau_1, \tau_3) \wedge \mathsf{EC}(\tau_2, \tau_3),$$

which is not expressible in $\mathcal{C}$. Obviously, the predicates $C$ and $C^2$ have identical semantics; thus, $\mathcal{C}$ is a sub-language of $\mathcal{C}^m$. Again, the increased expressive power makes no difference to the complexity class: $Sat(\mathcal{C}^m, \text{RegC})$ is still NP-complete [21].

**The logic $\mathcal{B}$:** We mention at this point a sub-language of $\mathcal{C}$ so inexpressive that no distinctively topological facts can be expressed in it, but which will nevertheless prove significant in the sequel. The language $\mathcal{B}$, again interpreted over sub-classes of RegC, is defined by:

$$\mathcal{B} \;=\; \mathcal{L}(\{+, \cdot, -, \mathbf{0}, \mathbf{1}\}, \{=\}).$$

Thus, $\mathcal{B}$ is the language of the variety of Boolean algebras. In the present context, it can be seen as capturing the essential content of *mereology*—the logic of 'part-whole' relations. (For a discussion of the relationship between mereology and Boolean algebra, see [53, 28].) Trivially, $Sat(\mathcal{B}, \text{RegC})$ is NP-complete.

**The logic $\mathcal{S}4_u$:** Returning to matters topological, we come to the most expressive topological logic to have been considered in the literature. The language $\mathcal{S}4_u$ is defined by:

$$\mathcal{S}4_u \;=\; \mathcal{L}(\{\cup, \cap, \overline{\cdot}, \cdot^\circ, \cdot^-, \mathbf{0}, \mathbf{1}\}, \{=\}),$$

and we write $\tau_1 \subseteq \tau_2$ as an abbreviation for $\tau_1 \cap \overline{\tau}_2 = \mathbf{0}$. We interpret the terms of this language as follows:

$$\begin{array}{lll} (\tau_1 \cap \tau_2)^{\mathfrak{M}} = \tau_1^{\mathfrak{M}} \cap \tau_2^{\mathfrak{M}}, & (\overline{\tau})^{\mathfrak{M}} = \overline{\tau^{\mathfrak{M}}} = T \setminus \tau^{\mathfrak{M}}, & \mathbf{0}^{\mathfrak{M}} = \mathbf{0} = \emptyset, \\ (\tau_1 \cup \tau_2)^{\mathfrak{M}} = \tau_1^{\mathfrak{M}} \cup \tau_2^{\mathfrak{M}}, & (\tau^\circ)^{\mathfrak{M}} = (\tau^{\mathfrak{M}})^\circ, & \mathbf{1}^{\mathfrak{M}} = \mathbf{1} = T, \\ & (\tau^-)^{\mathfrak{M}} = (\tau^{\mathfrak{M}})^-, & \end{array} \qquad (2.3)$$

where $\mathfrak{M}$ is a model over some frame $(T, S)$. As before, we have deliberately equivocated between function symbols in our formal language and the operations they denote. Since these operations do not in general preserve the property of being regular closed, it is unnatural to confine attention to the frame class RegC. Accordingly, we standardly interpret $\mathcal{S}4_u$ over the class All of all frames.

The richer term-language of $\mathcal{S}4_u$ means that, even though $=$ is the only predicate, we are still able to formulate distinctively topological (not just mereological) statements. Consider, for example, the formula

$$\big(r_1^\circ \cap r_2^- \neq \mathbf{0}\big) \;\rightarrow\; \big(r_1^\circ \cap r_2 \neq \mathbf{0}\big).$$

This formula states that, if an open set $r_1^\circ$ intersects the closure of a set $r_2$, then it also intersects $r_2$. Thus, it is valid over the class of frames All.

The language $\mathcal{S}4_u$ may be regarded as the richest of all the languages considered here, in the following sense. Given a $\mathcal{C}^m$-term $\tau$, we define inductively the $\mathcal{S}4_u$-term $\tau^\dagger$ as follows:

$$\mathbf{1}^\dagger = \mathbf{1}, \qquad \mathbf{0}^\dagger = \mathbf{0}, \qquad r^\dagger \;=\; r^{\circ -} \quad (r \text{ a variable}),$$

$$(-\tau_1)^\dagger \;=\; \big(\overline{\tau_1^\dagger}\big)^-, \qquad (\tau_1 \cdot \tau_2)^\dagger \;=\; \big(\tau_1^\dagger \cap \tau_2^\dagger\big)^{\circ -}, \qquad (\tau_1 + \tau_2)^\dagger \;=\; \tau_1^\dagger \cup \tau_2^\dagger.$$

If $\varphi$ is a $\mathcal{C}^m$-formula, let $\varphi^\dagger$ be the $\mathcal{S}4_u$-formula obtained by replacing every occurrence of $\tau_1 = \tau_2$ in $\varphi$ with $\tau_1^\dagger = \tau_2^\dagger$ and every occurrence of $C^k(\tau_1, \ldots, \tau_k)$ in $\varphi$ with $\tau_1^\dagger \cap \cdots \cap \tau_k^\dagger \neq \mathbf{0}$. For any topological space $T$, the regular closed subsets of $T$ are exactly the sets of the form $X^{\circ -}$, where $X \subseteq T$. Hence, as the variable $r$ ranges over $2^T$, the $\mathcal{S}4_u$-term $r^\dagger = (r^\circ)^-$ ranges over exactly the regular closed subsets of $T$. Using this observation, it is readily checked that $\varphi$ is satisfiable over a frame $(T, \mathrm{RC}(T))$ if and only if $\varphi^\dagger$ is satisfiable over the frame $(T, 2^T)$. Thus, we may informally regard any logic $(\mathcal{L}, \text{RegC})$, where $\mathcal{L}$ is a fragment of $\mathcal{C}^m$, as contained within the logic $(\mathcal{S}4_u, \text{All})$.[2]

Furthermore, the logic $(\mathcal{S}4_u, \text{All})$ has essentially the same expressive power as the modal logic **S4** (under McKinsey and Tarski's [37] topological interpretation) extended with the universal and existential modalities $\forall$ and $\exists$ of [24]. More precisely, define $\mathbf{S4}_u$ to be the set of terms formed using the variables in $\mathcal{R}$ together with the function symbols

$$\cup, \quad \cap, \quad \overline{\cdot}, \quad \cdot^\circ, \quad \cdot^-, \quad \forall, \quad \exists, \quad \mathbf{0}, \quad \mathbf{1}.$$

Here, $\exists$ and $\forall$ are unary, with the remaining symbols having their usual arities. Given any interpretation $\mathfrak{M}$, we define $\tau^{\mathfrak{M}}$ for any $\mathbf{S4}_u$-term $\tau$ using (2.3) together with:

$$(\exists\tau)^{\mathfrak{M}} \;=\; \begin{cases} T & \text{if } \tau^{\mathfrak{M}} \neq \emptyset, \\ \emptyset & \text{if } \tau^{\mathfrak{M}} = \emptyset; \end{cases} \qquad (\forall\tau)^{\mathfrak{M}} \;=\; \begin{cases} T & \text{if } \tau^{\mathfrak{M}} = T, \\ \emptyset & \text{if } \tau^{\mathfrak{M}} \neq T. \end{cases}$$

Thus, $\exists$ is interpreted as the discriminator function, and $\forall$ as its dual. We say that an $\mathbf{S4}_u$-term $\tau$ is *valid* if $\tau^{\mathfrak{M}} = T$ for any model $\mathfrak{M}$ over any topological frame $(T, S)$. By replacing each equality $\tau_1 = \tau_2$ with the term $\forall\big((\tau_1 \cap \tau_2) \cup (\overline{\tau_1} \cap \overline{\tau_2})\big)$ and the Boolean connectives with the corresponding function symbols, we obtain a validity-preserving embedding of $\mathcal{S}4_u$ into $\mathbf{S4}_u$-terms. On the other hand, it is well known (see, e.g., [1, 30]) that every $\mathbf{S4}_u$-term can be equivalently transformed to a term without occurrences of $\forall$ and $\exists$ in the scope of $^\circ$, $^-$, $\forall$ and $\exists$. Any such term can easily be rewritten as an equivalent $\mathcal{S}4_u$-formula by replacing $\forall\tau$ and $\exists\tau$ with $\tau = \mathbf{1}$ and $\tau \neq \mathbf{0}$, respectively, and by replacing Boolean function symbols with the corresponding Boolean connectives. (Note, however, that this transformation in general results in an exponential increase in size.) As the validity and satisfiability problems for

[2]That $\mathcal{RCC}$-8 is a simple fragment of $\mathcal{S}4_u$ was first observed by Bennett [5]; see also [45, 39] (in fact, $\mathcal{RCC}$-8 and $\mathcal{BRCC}$-8 can be embedded into the modal logic $\mathcal{S}5$ [58]).

both $\mathbf{S4}$ and $\mathbf{S4}_u$ are known to be PSPACE-complete [34, 39, 2], it follows that the problem $Sat(\mathcal{S}4_u, \textsc{All})$ is PSPACE-complete as well.

In this section, we introduced the languages $\mathcal{RCC}$-8, $\mathcal{B}$, $\mathcal{BRCC}$-8 (=$\mathcal{C}$), $\mathcal{C}^m$ and $\mathcal{S}4_u$. The variables of these languages range over certain distinguished subsets of some topological space, and their non-logical symbols denote various fixed primitive topological relations and operations. The topological space in question and its collection of distinguished subsets together form a (topological) frame; and the pair of a language and a class of frames is a (topological) logic. In particular, we interpreted $\mathcal{RCC}$-8, $\mathcal{B}$, $\mathcal{BRCC}$-8 (=$\mathcal{C}$) and $\mathcal{C}^m$ over the frame-class RegC, and the language $\mathcal{S}4_u$ over the frame-class All. We explained how all of these logics can be seen as fragments of $(\mathcal{S}4_u, \textsc{All})$, which has essentially the expressive power of the modal logic $\mathbf{S4}_u$. We observed that the complexity of the satisfiability problems for all of these logics is known. However, none of the above languages can express the property of being a connected region. Our question is: what happens to the complexity of satisfiability when that facility is provided?

## 3. Topological logics with connectedness

A topological space $T$ is *connected* just in case it is not the union of two non-empty, disjoint, open sets. Note that this definition can be expressed by the following $\mathcal{S}4_u$-formula (see [51]):

$$\big(r^\circ \cup (\overline{r})^\circ = \mathbf{1}\big) \to \big((r = \mathbf{1}) \lor (r = \mathbf{0})\big).$$

A subset $X \subseteq T$ is *connected in* $T$ if the topological space $X$ (with the subspace topology) is connected. If $X \subseteq T$, a maximal connected subset of $X$ is called a (*connected*) *component* of $X$. Every set is the disjoint union of its components (of which there is always at least one); a set is connected just in case it has exactly one component. If $T$ and $T'$ are topological spaces, a function $f\colon T \to T'$ is *continuous* if the inverse image under $f$ of every open subset of $T'$ is open in $T$ (equivalently: if the inverse image of every closed subset of $T'$ is closed in $T$). The image of a connected set under a continuous function $f$ is always connected; in fact, if $X \subseteq T$ has $k \geq 1$ components then $f(X)$ has at most $k$ components.

The simplest way to introduce connectedness into topological logics is to restrict attention to frames over connected topological spaces. For example, consider the classes of frames given by

$$\textsc{Con} = \{(T, 2^T) \mid T \text{ a connected topological space}\},$$
$$\textsc{ConRegC} = \{(T, \mathrm{RC}(T)) \mid T \text{ a connected topological space}\}.$$

Thus, it makes sense to consider the problems $Sat(\mathcal{L}, \textsc{ConRegC})$ for $\mathcal{L}$ any of $\mathcal{RCC}$-8, $\mathcal{B}$, $\mathcal{C}$ or $\mathcal{C}^m$, as well as the problem $Sat(\mathcal{S}4_u, \textsc{Con})$. An alternative, and more flexible, approach, however, is to expand the languages in question. Let $c$ be a unary predicate. If $\mathcal{L}$ is one of the topological languages introduced in Section 2, denote by $\mathcal{L}c$ ($\mathcal{L}$ *with connectedness*) the result of augmenting the topological primitives of $\mathcal{L}$ by $c$. Formally, if $\mathcal{L} = \mathcal{L}(F, P)$,

$$\mathcal{L}c \;=\; \mathcal{L}(F, P \cup \{c\}).$$

The predicate $c$ is given the expected fixed interpretation as follows:

$$\mathfrak{M} \models c(\tau) \quad \text{iff} \quad \tau^{\mathfrak{M}} \text{ is connected.}$$

Thus, from the languages $\mathcal{RCC}$-8, $\mathcal{B}$, $\mathcal{C}$, $\mathcal{C}^m$ and $\mathcal{S}4_u$, we obtain $\mathcal{RCC}$-8$c$, $\mathcal{B}c$, $\mathcal{C}c$, $\mathcal{C}^mc$, $\mathcal{S}4_uc$.

Consider, for example, the language $\mathcal{B}c$, which includes the formula

$$\big(c(r_1) \ \wedge \ c(r_2) \ \wedge \ (r_1 \cdot r_2 \neq \mathbf{0})\big) \ \rightarrow \ c(r_1 + r_2). \tag{3.1}$$

It is a well-known fact that, if $T$ is a topological space, any two connected subsets of $T$ with non-empty intersection have a connected union; further, if $X$ is connected, and $X \subseteq Y \subseteq X^-$, then $Y$ is also connected. But if $r$ and $s$ are regular closed subsets of $T$, then $r^\circ \cup s^\circ \subseteq r + s \subseteq (r^\circ \cup s^\circ)^-$. It follows that (3.1) is valid over RegC. At the other end of the expressive spectrum, consider the language $\mathcal{S}4_u c$, which includes the formula

$$\big(c(r_1) \ \wedge \ (r_1 \subseteq r_2) \ \wedge \ (r_2 \subseteq r_1^-)\big) \ \rightarrow \ c(r_2). \tag{3.2}$$

Using one of the facts we have just alluded to, it follows that (3.2) is valid over All.

The predicate $c$ can be generalized in the following way. Let $c^{\leq k}$ be a unary predicate, where $k \geq 1$ is represented in binary. If $\mathcal{L}$ is one of the languages introduced in Section 2, we denote by $\mathcal{L}cc$ (*$\mathcal{L}$ with component counting*) the result of augmenting the topological primitives of $\mathcal{L}$ by all of the predicates $c^{\leq k}$ $(k \geq 1)$. Formally, if $\mathcal{L} = \mathcal{L}(F, P)$,

$$\mathcal{L}cc \ = \ \mathcal{L}(F, P \cup \{c^{\leq k} \mid k \geq 1\}).$$

The predicates $c^{\leq k}$ are given fixed interpretations as follows, where $\mathfrak{M}$ is a model over some frame $(T, S)$:

$$\mathfrak{M} \models c^{\leq k}(\tau) \quad \text{ iff } \quad \tau^{\mathfrak{M}} \text{ has at most } k \text{ components in } T.$$

Thus, from the languages $\mathcal{RCC}$-8, $\mathcal{B}$, $\mathcal{C}$, $\mathcal{C}^m$ and $\mathcal{S}4_u$, we obtain $\mathcal{RCC}$-8$cc$, $\mathcal{B}cc$, $\mathcal{C}cc$, $\mathcal{C}^m cc$, $\mathcal{S}4_u cc$. We write $\neg c^{\leq k}(\tau)$ as $c^{\geq k+1}(\tau)$ and abbreviate $c^{\leq 1}(\tau)$ by $c(\tau)$. Thus, we may regard $\mathcal{L}c$ as a sub-language of $\mathcal{L}cc$. To illustrate, consider the $\mathcal{B}cc$-formula

$$\big(c^{\leq k}(r_1) \ \wedge \ c^{\leq l}(r_2) \ \wedge \ (r_1 \cdot r_2 \neq \mathbf{0})\big) \ \rightarrow \ c^{\leq l+k-1}(r_1 + r_2). \tag{3.3}$$

Using the same argument as for (3.1), this formula is easily shown to be valid over RegC.

For rich topological languages, such as $\mathcal{S}4_u c$, the predicates $c^{\leq k}$ give us—in some sense—no expressive power that $c$ does not already give us. Let $\tau$ be an $\mathcal{S}4_u c$-term and $r_1, \ldots, r_k$ variables not occurring in $\tau$. Consider the $\mathcal{S}4_u c$-formulas

$$\big(\tau = \bigcup_{1 \leq i \leq k} r_i\big) \quad \wedge \quad \bigwedge_{1 \leq i \leq k} c(r_i), \tag{3.4}$$

$$\big(\tau = \bigcup_{1 \leq i \leq k+1} r_i\big) \quad \wedge \quad \bigwedge_{1 \leq i \leq k+1} (r_i \neq \mathbf{0}) \quad \wedge \quad \bigwedge_{1 \leq i < j \leq k+1} \big(\tau \cap r_i^- \cap r_j^- = \mathbf{0}\big) \tag{3.5}$$

together with some model $\mathfrak{M}$ over a topological frame $(T, S)$. Let us assume that $(T, S)$ has the property that, if $r \in S$, and $s$ is a component of $r$, then $s \in S$—a very natural requirement for topological frames. If (3.4) is true in $\mathfrak{M}$, then $\tau^{\mathfrak{M}}$ is seen to have at most $k$ components. Conversely, if $\tau^{\mathfrak{M}}$ has at most $k$ components, then, by modifying the regions assigned to $r_1, \ldots, r_k$ if necessary, we easily obtain a model $\mathfrak{M}'$ satisfying (3.4). It follows that, if $\varphi$ is an $\mathcal{S}4_u cc$-formula, then any instance of $c^{\leq k}(\tau)$ having positive polarity may be equisatisfiably replaced by (3.4) (with fresh variables $r_1, \ldots, r_k$). Similarly, any instance of $c^{\geq k}(\tau)$ having positive polarity may be equisatisfiably likewise replaced by (3.5). However, while the number of symbols in the predicate $c^{\leq k}$ is proportional to $\log k$, the number of symbols in (3.4) is proportional to $k$. That is: $\mathcal{S}4_u c$-formulas are in general exponentially longer than the $\mathcal{S}4_u cc$-formulas they replace. So, although the component-counting predicates $c^{\leq k}$ can usually be eliminated in this way, doing so may affect the complexity of the satisfiability problem.

The main contribution of this paper is to determine the computational complexity of the satisfiability problems for topological logics based on the languages $\mathcal{L}c$ and $\mathcal{L}cc$, where $\mathcal{L}$ is any of the languages introduced in Section 2. To date, there are only two known complexity results for such logics. On the one hand, according to [40], satisfiability of $\mathcal{S}4_u cc$-formulas over ALL is NEXPTIME-complete, which gives the NEXPTIME upper bound for all of the other logics considered in this section. On the other, according to [57], satisfiability of $\mathcal{C}$-formulas is PSPACE-complete over CONREGC.

## 4. Aleksandrov spaces.

The close connection between spatial logics and the modal logic $\mathbf{S4}_u$ mentioned above suggests that instead of topological semantics one may try to employ Kripke semantics (which gives rise to topological spaces with a very transparent structure) and the corresponding modal logic machinery.

Recall that *Kripke frames* for $\mathbf{S4}_u$ are pairs of the form $(W, R)$, where $W$ is a set, and $R$ is a reflexive and transitive relation on $W$; such frames are also called *quasi-orders*. Every quasi-order $(W, R)$ can be regarded as a topological space by declaring $X \subseteq W$ to be *open* if and only if $X$ is upward closed with respect to $R$, that is, if $x \in X$ and $xRy$ then $y \in X$. In other words, for every $X \subseteq W$,

$$X^\circ \;=\; \big\{x \in X \mid \forall y \in W \ (xRy \to y \in X)\big\}.$$

The most important property distinguishing topological spaces $T$ induced by quasi-orders is that arbitrary (not only finite) intersections of open sets in $T$ are open. Topological spaces with this property are called *Aleksandrov spaces*. It is also known (see, e.g., [7]) that every Aleksandrov space is induced by a quasi-order.

Another important feature of such topological spaces is that the topological notion of connectedness in $T$ coincides with the graph-theoretic notion of connectedness in the undirected graph induced by $(W, R)$. More precisely, one can easily check that a set $X \subseteq W$ is connected in $T$ if and only if, for any points $x, y \in X$, there is a path $x = x_1, \dots, x_n = y$ such that, for all $i$, $1 \le i < n$, we have $x_i \in X$ and either $x_i R x_{i+1}$ or $x_{i+1} R x_i$.

Henceforth, we shall identify an Aleksandrov space with the quasi-order generating it, alternating freely between topological and graph-theoretic perspectives. Denote by ALEK the class of *finite* Aleksandrov frames. A topological model based on an Aleksandrov space will be called an *Aleksandrov model*.

The next lemma, originating in [37] and [33], shows that, for many topological logics, it suffices to work with finite Aleksandrov spaces. It can be proved by the standard filtration argument (see, e.g., [9]).

**Lemma 4.1.** *For every finite set $\Theta$ of $\mathcal{S}4_u$-terms closed under subterms and every topological model $\mathfrak{M} = (T, S, \cdot^{\mathfrak{M}})$, there exist an Aleksandrov model $\mathfrak{A} = (T_A, 2^{T_A}, \cdot^{\mathfrak{A}})$ and a continuous function $f\colon T \to T_A$ such that $|T_A| \le 2^{O(|\Theta|)}$ and $\tau^{\mathfrak{A}} = f(\tau^{\mathfrak{M}})$, for every $\tau \in \Theta$.*

This lemma has a number of important consequences. First, it follows immediately that $Sat(\mathcal{S}4_u, \text{ALL}) = Sat(\mathcal{S}4_u, \text{ALEK})$. Using the translation $\cdot^\dagger$ of $\mathcal{B}$-terms and $\mathcal{C}^m$-formulas into $\mathcal{S}4_u$ defined in Section 2, we obtain $Sat(\mathcal{C}^m, \text{REGC}) = Sat(\mathcal{C}^m, \text{ALEK} \cap \text{REGC})$, etc. The PSPACE upper bound for $Sat(\mathcal{S}4_u, \text{ALL})$ follows from Lemma 4.1 and the fact (well-known in modal logic) that the model $\mathfrak{A}$ in Lemma 4.1 can be 'unravelled' into a forest of trees of

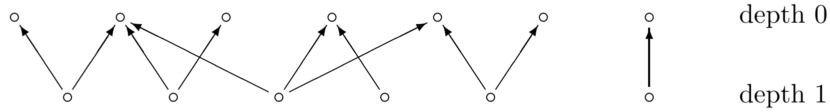



Figure 4: Quasi-saw.

clusters,[3] with the length of branches not exceeding the maximal size of terms in $\Theta$—the so-called *tree-model property* of $\mathcal{S}4_u$. Finally, the fact that $f$ is continuous guarantees that the number of components in $f(\tau^{\mathfrak{M}})$ does not exceed the number of components in $\tau^{\mathfrak{M}}$, which will be used in the case of logics with connectedness predicates.

For the topological logics with $\mathcal{B}$-terms only, say $\mathcal{C}^m cc$, even simpler Aleksandrov models are enough. Call a quasi-order $(W,R)$ a *quasi-saw* if $W = W_0 \cup W_1$, for some disjoint $W_0$ and $W_1$, and $R$ is the reflexive closure of a subset of $W_1 \times W_0$. In this case we also say that the points in $W_i$ are *of depth* $i$ in $(W,R)$; see Fig. 4. Aleksandrov models over quasi-saws will be called *quasi-saw models*.

**Lemma 4.2.** *For every finite Aleksandrov model* $\mathfrak{A} = (T_A, \mathrm{RC}(T_A), \cdot^{\mathfrak{A}})$*, with* $T_A$ *induced by a quasi-order* $(W, R_A)$*, there is a quasi-saw model* $\mathfrak{B} = (T_B, \mathrm{RC}(T_B), \cdot^{\mathfrak{B}})$ *such that* $T_B$ *is induced by* $(W, R_B)$ *with* $R_B \subseteq R_A$ *and, for every* $\mathcal{B}$*-term* $\tau$*,* (i) $\tau^{\mathfrak{B}} = \tau^{\mathfrak{A}}$*, and* (ii) $\tau$ *has the same number of components in* $\mathfrak{A}$ *and* $\mathfrak{B}$*.*

**Proof.**Let $W_0$ be the set of maximal points in $(W, R_A)$—the set of points from the final clusters in $(W, R_A)$, to be more precise—i.e.,

$$W_0 = \{v \in W \mid vR_Au \text{ implies } uR_Av, \text{ for all } u \in W\}.$$

In every final cluster $C \subseteq W_0$ with $|C| \geq 2$ we select some point and denote by $U$ the set of all such selected points. Then we set $V_0 = W_0 \setminus U$ and $V_1 = W \setminus V_0$, and define $R_B$ to be the reflexive closure of $R_A \cap (V_1 \times V_0)$. Clearly, $(W, R_B)$ is a quasi-saw, with $V_0$ and $V_1$ being the sets of points of depth 0 and 1, respectively. For every variable $r$, let $r^{\mathfrak{B}} = r^{\mathfrak{A}}$. As the extension of a $\mathcal{B}$-term $\tau$ in $\mathfrak{A}$ is regular closed and $\mathfrak{A}$ is finite, it is straightforward to show:

$$\text{if } y \in \tau^{\mathfrak{A}} \text{ then there exists } z \in V_0 \text{ such that } yR_Az \text{ and } z \in \tau^{\mathfrak{A}}. \tag{4.1}$$

We now prove (i) and (ii) by induction on the construction of $\tau$.

(i) The basis of induction follows from the definition.

*Case* $\tau = -\tau_1$. We have $x \in ((\overline{\tau_1})^-)^{\mathfrak{A}}$ iff

$$\begin{array}{ll} \text{[def.]} & \exists y \in W \ \big(xR_Ay \text{ and } y \notin \tau_1^{\mathfrak{A}}\big) \\ \text{iff [(4.1)]} & \exists y \in V_0 \ \big(xR_Ay \text{ and } y \notin \tau_1^{\mathfrak{A}}\big) \\ \text{iff [IH]} & \exists y \in V_0 \ \big(xR_By \text{ and } y \notin \tau_1^{\mathfrak{B}}\big) \\ \text{iff [def.]} & x \in ((\overline{\tau_1})^-)^{\mathfrak{B}}. \end{array}$$

*Case* $\tau = \tau_1 + \tau_2$. We have $(\tau_1 + \tau_2)^{\mathfrak{A}} = \tau_1^{\mathfrak{A}} \cup \tau_2^{\mathfrak{A}} = \tau_1^{\mathfrak{B}} \cup \tau_2^{\mathfrak{B}} = (\tau_1 + \tau_2)^{\mathfrak{B}}$, with the middle equation following by IH.

[3] A cluster in a quasi-order $(W,R)$ is any set of the form $\{x \in W \mid xRy \text{ and } yRx\}$, for some $y \in W$.

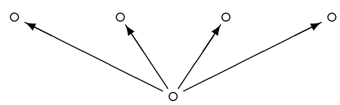



Figure 5: $k$-fork for $k = 4$.

(ii) As $R_B \subseteq R_A$, the number of connected components of $\tau^{\mathfrak{A}}$ in $\mathfrak{A}$ cannot be greater than the number of components of $\tau^{\mathfrak{B}}$ in $\mathfrak{B}$. Conversely, suppose that $X$ is a component of $\tau^{\mathfrak{A}}$ in $\mathfrak{A}$. As $\tau^{\mathfrak{A}}$ is regular closed, it is the closure under $R_A^-$ of the set of final clusters in $X$. It follows immediately from the definition of $R_B$ that all non-final points in $X$ have precisely the same $R_B$- and $R_A$-accessible final points. This means that $X$ is connected in $(W, R_B)$ as well. □

Recall now that every satisfiable $\mathcal{C}^m$-formula is satisfiable in a finite Aleksandrov model, and so, by Lemma 4.2, in a quasi-saw model (over RegC). The following lemma imposes restrictions on the branching factor and the number of points of depth 1 in such models. Let us call a *$k$-fork* any partial order of the form depicted in Fig. 5.

**Lemma 4.3.** *If a $\mathcal{C}^m$-formula $\varphi$ is satisfiable (over* RegC*) then it is satisfiable in a quasi-saw model over the disjoint union of $n$-many $k_i$-forks, $1 \leq i \leq n$, where $n \leq |\varphi|$ and each $k_i$ does not exceed the largest value $k$ such that the predicate $C^k$ occurs in $\varphi$.*

**Proof.** Without loss of generality, we may assume that all the literals in $\varphi$ involving equality are of the form $\tau = \mathbf{0}$ or $\tau \neq \mathbf{0}$. Suppose that $\mathfrak{M} \models \varphi$ and $\Lambda$ is the set of all literals of $\varphi$ that are true in $\mathfrak{M}$. By Lemmas 4.1 and 4.2, there is a quasi-saw model $\mathfrak{B} = (T_B, \mathrm{RC}(T_B), \cdot^{\mathfrak{B}})$ with $T_B$ induced by a partial order $(W_B, R_B)$ and $\mathfrak{B} \models \Lambda$ (and so $\mathfrak{B} \models \varphi$).

We construct a quasi-saw model $\mathfrak{A} = (T_A, \mathrm{RC}(T_A), \cdot^{\mathfrak{A}})$ induced by a disjoint union of forks $(W_A, R_A)$ and a map $f \colon T_A \to T_B$ as follows:

- for each literal $(\tau \neq \mathbf{0}) \in \Lambda$, we select a point $x \in \tau^{\mathfrak{B}}$ of depth 0, add a fresh 1-fork $(\{u, v\}, \{(u, v)\}^*)$ to $(W_A, R_A)$, where $R^*$ denotes the reflexive closure of $R$, and set $f(u) = x$, $f(v) = x$;
- for each literal $C^k(\tau_1, \dots, \tau_k) \in \Lambda$, either (i) there is a point $x$ of depth 0 in $\mathfrak{B}$ with $x \in \tau_1^{\mathfrak{B}} \cap \dots \cap \tau_k^{\mathfrak{B}}$ or (ii) there exists a point $y$ of depth 1 in $\mathfrak{B}$ such that $y \in \tau_1^{\mathfrak{B}} \cap \dots \cap \tau_k^{\mathfrak{B}}$, in which case there are (not necessarily distinct) points $x_1, \dots, x_k$ of depth 0 with $x_i \in \tau_i^{\mathfrak{B}}$; in the former case we add a fresh 1-fork $(\{u, v\}, \{(u, v)\}^*)$ to $(W_A, R_A)$ and set $f(u) = x$ and $f(v) = x$; in the latter case we add a fresh $k$-fork $(\{u, v_1, \dots, v_k\}, \{(u, v_i) \mid 1 \leq i \leq k)\}^*)$ to $(W_A, R_A)$ and set $f(u) = y$ and $f(v_i) = x_i$, for $1 \leq i \leq k$.

Define $\cdot^{\mathfrak{A}}$ by taking $v \in r^{\mathfrak{A}}$ iff $f(v) \in r^{\mathfrak{B}}$, for every $v$ of depth 0, and $u \in r^{\mathfrak{A}}$ iff there is $v \in r^{\mathfrak{A}}$ of depth 0 with $u R_A v$, for every $u$ of depth 1. By definition, $r^{\mathfrak{A}}$ is regular closed in $\mathfrak{A}$, and it is easily checked that $\mathfrak{A} \models \Lambda$. □

As an immediate consequence of Lemma 4.3 we obtain the following:

**Corollary 4.4.** *Sat*$(\mathcal{C}^m, \text{RegC})$*, Sat*$(\mathcal{C}, \text{RegC})$*, Sat*$(\mathcal{B}, \text{RegC})$ *and Sat*$(\mathcal{RCC}\text{-}8, \text{RegC})$ *are all* NP*-complete.*

The reader may wonder at this point whether lower complexities can be achieved if we consider various sub-logics of the logics mentioned in this corollary. The answer is that

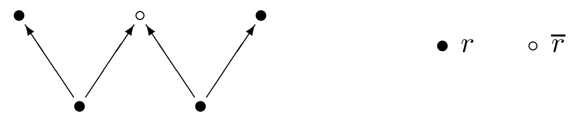



Figure 6: A quasi-saw model for $\neg c(r) \wedge c(\mathbf{1})$.

they can. For example, maximal tractable (polynomial) fragments of $\mathcal{RCC}$-8 were identified in [44, 46], and it was shown in [26] that the problem of determining the satisfiability of a conjunction of atomic formulas of $\mathcal{RCC}$-8 over the class RegC is NLogSpace-complete.

We close this section with some remarks about logics interpreted over connected spaces. The language $\mathcal{C}$ can distinguish between connected and disconnected spaces, because the formula

$$\neg C(r, -r) \;\wedge\; (r \neq \mathbf{0}) \;\wedge\; (r \neq \mathbf{1}) \tag{4.2}$$

is satisfiable, but only in models over disconnected spaces. By contrast, the languages $\mathcal{B}$ and $\mathcal{RCC}$-8 cannot distinguish between connected and disconnected spaces:

$$Sat(\mathcal{RCC}\text{-}8, \textsc{RegC}) = Sat(\mathcal{RCC}\text{-}8, \textsc{ConRegC}),$$
$$Sat(\mathcal{B}, \textsc{RegC}) = Sat(\mathcal{B}, \textsc{ConRegC}).$$

Indeed, suppose that an $\mathcal{RCC}$-8-formula $\varphi$ is satisfied in a quasi-saw model $\mathfrak{A}$ with the underlying quasi-order $(W_A, R_A)$ constructed in the proof of Lemma 4.3. To make this model connected, we can simply add to $W_A$ a new point $w$ of depth 0 and extend $R_A$ by the arrows $uR_Aw$, for all $u$ of depth 1 in $\mathfrak{A}$. It is easy to see that, under the same valuation as in $\mathfrak{A}$, $\varphi$ is satisfiable in the extended connected model. For the less expressive language $\mathcal{B}$, it suffices to add a new point of depth 1 to $\mathfrak{A}$ and connect it to all points of depth 0; details are left to the reader.

On the other hand, equipping even the weakest spatial logics such as $\mathcal{RCC}$-8 or $\mathcal{B}$ with the connectedness predicates (or even just interpreting them over connected spaces) invalidates various model-theoretic properties employed above—most notably the 'tree-model' property, heavily used in the proof of Lemma 4.3. Consider, for example, the $\mathcal{B}c$-formula $\neg c(r) \wedge c(\mathbf{1})$. Its smallest satisfying quasi-saw model is illustrated in Fig. 6. Note that this model cannot be transformed to a forest, because the underlying frame must stay connected. Indeed, as we shall see in the next section, to satisfy $\mathcal{B}c$-formulas, or $\mathcal{C}$-formulas in connected spaces, quasi-saw models with *exponentially* many points in the length of the formulas may be required. It is these phenomena that are responsible for the increased complexity of satisfiability which we shall encounter below.

## 5. Computational complexity

We are now in a position to prove tight complexity results for spatial logics in the range between $\mathcal{B}c$ and $\mathcal{S}4_u cc$. The NExpTime upper bound for all the logics considered in this paper was obtained in [40]:

**Theorem 5.1** ([40])**.** *$Sat(\mathcal{S}4_u cc, \textsc{All})$ is in* NExpTime.

The idea of the proof is based on the following observations. Let $\varphi$ be any $\mathcal{S}4_u cc$-formula. Evidently, $\varphi$ is satisfiable if and only if there exists a set $\Phi$ of $\mathcal{S}4_u cc$-literals, involving all the atoms occurring in $\varphi$, such that: (i) $\Phi$ is satisfiable; and (ii) $\Phi \cup \{\varphi\}$ is satisfiable in propositional logic (where we treat all the atoms as propositional variables). Since propositional satisfiability can be checked in NP, it suffices to restrict attention to $\mathcal{S}4_u cc$-formulas which are conjunctions of literals—i.e. those of the form:

$$\big(\rho = \mathbf{0}\big) \quad\wedge\quad \bigwedge_{i=1}^{m} \big(\tau_i \neq \mathbf{0}\big) \quad\wedge\quad \bigwedge_{i=1}^{n} c^{\leq k_i}(\sigma_i) \quad\wedge\quad \bigwedge_{i=1}^{p} c^{\geq k'_i}(\sigma'_i). \tag{5.1}$$

The conjuncts of the form $c^{\geq k'_i}(\sigma'_i)$ can be eliminated using the following lemma [40]:

**Lemma 5.2** ([40])**.** *Let $\varphi$ be any $\mathcal{S}4_u cc$-formula, and $\tau$ an $\mathcal{S}4_u$-term. Then, for every $n \geq 0$, there exists an $\mathcal{S}4_u c$-formula $\psi$, with $|\psi|$ bounded by a polynomial function of $n + |\tau|$, such that $\varphi \wedge \psi$ is satisfiable if and only if $\varphi \wedge c^{\geq 2^n}(\tau)$ is satisfiable.*

By repeated applications of Lemma 5.2, it is then a straightforward matter to transform (5.1), in polynomial time, into an equisatisfiable formula of the form

$$\psi \quad = \quad \big(\rho = \mathbf{0}\big) \quad\wedge\quad \bigwedge_{i=1}^{m} \big(\tau_i \neq \mathbf{0}\big) \quad\wedge\quad \bigwedge_{i=1}^{n} c^{\leq k_i}(\sigma_i). \tag{5.2}$$

Suppose that $\psi$ is true in some model $\mathfrak{M}$ over a space $T$, and let $\Theta$ be the set of terms occurring in $\psi$. Let $\mathfrak{A}$ and $f$ be as guaranteed by Lemma 4.1. Then $\tau^{\mathfrak{A}} = f(\tau^{\mathfrak{M}})$ for all $\tau \in \Theta$, and $f$ is continuous, whence $\mathfrak{A} \models \psi$. Thus, if $\psi$ is satisfiable, then it is satisfied over a topological space whose size is bounded by an exponential function of $|\psi|$, which gives the NExpTime upper bound of Theorem 5.1.

We turn next to the language $\mathcal{S}4_u c$. By similar reasoning to the above, we may without loss of generality confine attention to the problem of determining the satisfiability of formulas of the form

$$\psi \quad = \quad \big(\rho = \mathbf{0}\big) \quad\wedge\quad \bigwedge_{i=1}^{m} \big(\tau_i \neq \mathbf{0}\big) \quad\wedge\quad \bigwedge_{i=1}^{n} \big(c(\sigma_i) \wedge (\sigma_i \neq \mathbf{0})\big). \tag{5.3}$$

**Theorem 5.3.** *$Sat(\mathcal{S}4_u c, \textsc{All})$ is in* ExpTime.

**Proof.** The proof is by reduction to the satisfiability problem for propositional dynamic logic $\mathcal{PDL}$ with converse and nominals, which is known to be ExpTime-complete [23, Section 7.3]. Let $\psi$ be as in (5.3). Take two atomic programs $\alpha$ and $\beta$ and, for each $\sigma_i$, a nominal $\ell_i$. For a term $\tau$, denote by $\tau^{\ddagger}$ the $\mathcal{PDL}$-formula obtained by replacing in $\tau$, recursively, each sub-term $\vartheta^{\circ}$ with $[\alpha^*]\vartheta$. Thus the transitive and reflexive accessibility relation of the modal logic **S4** is simulated by $\alpha^*$, and the universal modality $\forall$ (see the end of Section 2) is simulated by $[\gamma]$, where $\gamma = (\beta \cup \beta^- \cup \alpha \cup \alpha^-)^*$. Consider now the formula

$$\psi' \quad = \quad [\gamma]\neg\rho^{\ddagger} \quad\wedge\quad \bigwedge_{i=1}^{m} \langle\gamma\rangle\tau_i^{\ddagger} \quad\wedge\quad \bigwedge_{i=1}^{n} \Big(\langle\gamma\rangle(\ell_i \wedge \sigma_i^{\ddagger}) \;\wedge\; [\gamma](\sigma_i^{\ddagger} \rightarrow \langle(\alpha \cup \alpha^-; \sigma_i^{\ddagger}?)^*\rangle\ell_i)\Big).$$

It is easy to see that $\psi'$ is satisfiable if and only if $\psi$ is satisfiable: the first conjunct of $\psi'$ states that $\rho$ is empty, the second that all $\tau_i$ are non-empty, the third states that each $\sigma_i$ holds at a point where $\ell_i$ holds and that from each $\sigma_i$-point there is a path (along $\alpha \cup \alpha^-$) to $\ell_i$ which lies entirely within $\sigma_i$. $\square$

If $\psi$ is parsimonious in its use of connectedness, we can do somewhat better. Denote by $\mathcal{S}4_uc^1$ the set of $\mathcal{S}4_uc$-formulas with *at most one* occurrence of an atom of the form $c(\tau)$.

**Theorem 5.4.** *$Sat(\mathcal{S}4_uc^1, \textsc{All})$ is* PSPACE*-complete.*

**Proof.** The lower bound follows from [34]. We sketch a nondeterministic PSPACE algorithm recognizing $Sat(\mathcal{S}4_uc^1, \textsc{All})$. We may assume without loss of generality that the input $\psi$ has the form (5.3), with $n \leq 1$. If $n = 0$, i.e. if $\psi$ does not contain a conjunct of the form $c(\sigma) \wedge (\sigma \neq \mathbf{0})$, then a standard satisfiability checking algorithm for $\mathbf{S4}_u$ is applied. Assume, then, that $n = 1$ in (5.3); and write $\sigma$ for $\sigma_1$. Set $\mathbf{B} = \{\overline{\rho}^\circ, \overline{\rho}\} \cup \{\tau, \overline{\tau} \mid \tau \in \mathit{term}(\psi)\}$, where $\mathit{term}(\psi)$ is the set of all sub-terms of $\psi$. A subset $\mathfrak{t}$ of $\mathbf{B}$ is called a *type for $\psi$* if $\overline{\rho}^\circ \in \mathfrak{t}$ and $\tau \in \mathfrak{t}$ iff $\overline{\tau} \notin \mathfrak{t}$, for all $\overline{\tau} \in \mathbf{B}$.

Now, guess a type $\mathfrak{t}_\sigma$ containing $\sigma$ and start $(m + 1)$ $\mathbf{S4}$-tableau procedures (see, e.g., [19, 25]) with inputs $\tau_1 \cap \overline{\rho}^\circ$, $\tau_2 \cap \overline{\rho}^\circ, \dots, \tau_m \cap \overline{\rho}^\circ$, and $\bigcap \mathfrak{t}_\sigma \cap \overline{\rho}^\circ$ in the usual way, expanding nodes branch-by-branch, and recovering the space once branches are checked. We may as well assume that the nodes of these tableaux are types. Suppose $\mathfrak{t}$ is a type occurring in one of them. If $\sigma \in \mathfrak{t}$, it suffices to check that $\mathfrak{t}$ can be connected by a $\sigma$-path of $\leq 2^{|\psi|}$ points to $\mathfrak{t}_\sigma$. This can be done in PSPACE by the following non-deterministic subroutine. We start with $\mathfrak{t}$ and count from 1 to $2^{|\psi|}$. At each step we guess a new type $\mathfrak{t}'$ with $\sigma, \overline{\rho}^\circ \in \mathfrak{t}'$, and check that

- either (i) $\tau \in \mathfrak{t}'$ for all $\tau^\circ \in \mathfrak{t}$, or (ii) $\tau \in \mathfrak{t}$, for all $\tau^\circ \in \mathfrak{t}'$ (in the former case, $\mathfrak{t}'$ is accessible from $\mathfrak{t}$, in the latter, the other way around);
- an $\mathbf{S4}$-tableau with root $\mathfrak{t}'$ can be constructed (which can be discarded after completion). Note that, although this tableau may contain types $\mathfrak{t}''$ with $\sigma \in \mathfrak{t}''$, these types can never threaten the connectedness of $\sigma$, since they are all accessible from the root $\mathfrak{t}'$ of the tableau, and so are connected to both $\mathfrak{t}$ and $\mathfrak{t}_\sigma$, by the transitivity of the accessibility relation.

If both checks are successful and $\mathfrak{t}' = \mathfrak{t}_\sigma$, the subroutine succeeds; if $\mathfrak{t}' \neq \mathfrak{t}_\sigma$ we set $\mathfrak{t} = \mathfrak{t}'$ and continue to the next step (provided that the step number $< 2^{|\psi|}$, otherwise the subroutine fails). Clearly, this subroutine succeeds if there is a $\sigma$-path connecting $\mathfrak{t}$ and $\mathfrak{t}_\sigma$ and fails in every computation otherwise; moreover, it requires only polynomial memory to store the tableau for $\mathfrak{t}'$ and the step number. □

To reduce notational clutter we denote, for any topological space $T$, the (singleton) frame-class $\{(T, 2^T)\}$ simply by $T$, and the (singleton) frame-class $\{(T, \mathrm{RC}(T))\}$ simply by $\mathrm{RC}(T)$. This notation is not entirely uniform, but it should be obvious what is meant.

As shown in [51] (see also Theorem 6.1 below), $Sat(\mathcal{S}4_u, \textsc{Con}) = Sat(\mathcal{S}4_u, \mathbb{R}^n)$ for any $n \geq 1$. Recalling now that the modal logic $\mathbf{S4}$ is PSPACE-hard, we immediately obtain the following:

**Corollary 5.5.** *$Sat(\mathcal{S}4_u, \textsc{Con})$ and $Sat(\mathcal{S}4_u, \mathbb{R}^n)$ are all* PSPACE*-complete for any $n \geq 1$.*

The proof of Theorem 5.4 can be generalized in various ways. For example, assume that $\psi = \psi_1 \wedge \psi_2$ is an $\mathcal{S}4_uc$-formula in which $\psi_2 = \bigwedge_{1 \leq i < j \leq n}(\sigma_i \cap \sigma_j = \mathbf{0})$, where $\{\sigma_1, \dots, \sigma_n\}$ is the collection of all $\sigma_i$ such that $c(\sigma_i)$ occurs in $\psi_1$. A straightforward extension of the algorithm in the proof of Theorem 5.4 shows that satisfiability of $\psi$ is still in PSPACE. Thus, if the connectedness predicate is applied only to regions that are known to be pairwise disjoint, then it does not matter how many times this predicate occurs in the formula: satisfiability is still in PSPACE.

Our next theorem gives matching lower bounds for Theorem 5.4 and Corollary 5.5.

**Theorem 5.6.** *$Sat(\mathcal{C}c^1, \textsc{RegC})$ is* PSPACE*-hard. In fact, the problems $Sat(\mathcal{C}, \textsc{ConRegC})$ and $Sat(\mathcal{C}, \mathrm{RC}(\mathbb{R}^n))$ for all $n \geq 1$ are* PSPACE*-hard.*

**Proof.** Let $L$ be a language in PSPACE. Then there is a polynomial-space-bounded *deterministic* Turing machine $M$ recognizing $L$. Without loss of generality, we may assume that, given some input $\vec{a} \in L$ on the tape, $M$ starts in the *initial state*, reaches the *accepting state* (with the resulting tape being empty and the head positioned over the first cell) and then moves to the *halting state*, from which no transition is possible. Moreover, throughout the computation the machine never goes to the left of the first cell and to the right of the $s$'th cell, where $s = p(|\vec{a}|)$ for some polynomial $p(\cdot)$.

Let $Q$ and $\Sigma$ be the set of states and the alphabet of $M$, respectively. The instructions of $M$ are of the form $(q, a) \to (q', a', d)$, $d \in \{+1, 0, -1\}$, with their standard meaning. A configuration of $M$ is a word $\mathfrak{c}$ of the form

$$a_1, \dots, a_{i-1}, (q, a_i), a_{i+1}, \dots, a_s, \tag{5.4}$$

where $a_1, \dots, a_s$ ($a_j \in \Sigma$) is the current contents of the tape, $q \in Q$ the current state, and $i$ the current position of the head. If a configuration $\mathfrak{c}'$ is obtained from a configuration $\mathfrak{c}$ by applying one instruction of $M$ then we write $\mathfrak{c} \to \mathfrak{c}'$.

It will be convenient for us to represent $M$ as the following set $\mathcal{T}$ of 4-tuples, where $t, b$ are two fresh auxiliary symbols (see Fig. 7):

- $(a, t, a, t)$ and $(a, b, a, b)$, for every $a \in \Sigma$,
- $(a', (q', b), (q', a'), b)$ and $(a', t, (q', a'), (q', t))$, for all $a' \in \Sigma$ and $q' \in Q$,
- $((q, a), t, (q', a'), b)$, for every instruction $(q, a) \to (q', a', 0)$ in $M$,
- $((q, a), t, a', (q', b))$, for every instruction $(q, a) \to (q', a', -1)$ in $M$,
- $((q, a), (q', t), a', b)$, for every instruction $(q, a) \to (q', a', +1)$ in $M$.

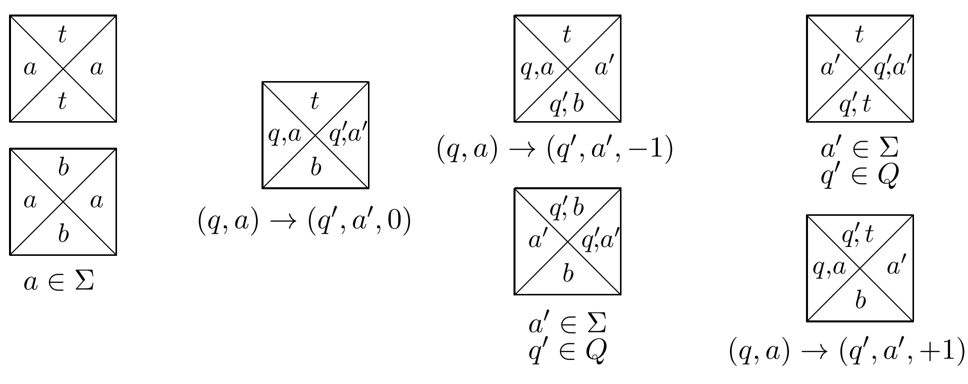



Figure 7: Tile types $\mathcal{T}$ for the Turing machine $M$.

We call these 4-tuples *tile types* and, for each $T \in \mathcal{T}$, denote its four components by $\mathit{left}(T)$, $\mathit{top}(T)$, $\mathit{right}(T)$ and $\mathit{bot}(T)$, respectively. Configurations of $M$ will be encoded on the left- and right-hand sides of the tile types in sequences $T_{k_1}, \dots, T_{k_s}$ such that

$$\mathit{top}(T_{k_s}) = t, \qquad \mathit{top}(T_{k_i}) = \mathit{bot}(T_{k_{i+1}}), \text{ for } 1 \leq i < s, \qquad \text{and} \qquad \mathit{bot}(T_{k_1}) = b. \tag{5.5}$$

By the definition of $\mathcal{T}$, every such sequence $T_{k_1}, \dots, T_{k_s}$ gives rise to two configurations $\mathfrak{c} = \mathit{left}(T_{k_1}), \dots, \mathit{left}(T_{k_s})$ and $\mathfrak{c}' = \mathit{right}(T_{k_1}), \dots, \mathit{right}(T_{k_s})$ of $M$ with $\mathfrak{c} \to \mathfrak{c}'$.

Now we describe the computations of $M$ in terms of $\mathcal{C}$-formulas. While constructing the formulas, we will assume that $\mathfrak{A}$ is a connected quasi-saw model induced by $(W, R)$ and $W_0$ is the set of points of depth 0 in $(W, R)$.

We need $s$ variables $T_k^1, \dots, T_k^s$, for each $T_k \in \mathcal{T}$, and three additional variables $B^0$, $B^1$ and $B^2$. Consider the following $\mathcal{C}$-formulas:

$$B^0 + B^1 + B^2 = \mathbf{1}, \tag{5.6}$$

$$B^0 \cdot B^1 = \mathbf{0}, \quad B^1 \cdot B^2 = \mathbf{0}, \quad B^2 \cdot B^0 = \mathbf{0}. \tag{5.7}$$

If the conjunction of (5.6)–(5.7) holds in $\mathfrak{A}$ then every point $x \in W_0$ is precisely in one of $(B^\ell)^{\mathfrak{A}}$, $0 \le \ell \le 2$. We will use the $B^\ell$ to introduce 'direction' in our quasi-saw model in the following sense. If $x_1, x_2 \in W_0$ and $zRx_i$, $i = 1, 2$, then there are only three possibilities:

- $x_1$ and $x_2$ are regarded as 'identical' whenever $x_1, x_2 \in (B^\ell)^{\mathfrak{A}}$, for $0 \le \ell \le 2$,
- $x_2$ is a 'successor' of $x_1$ whenever $x_1 \in (B^\ell)^{\mathfrak{A}}$ and $x_2 \in (B^{\ell\oplus 1})^{\mathfrak{A}}$, for $0 \le \ell \le 2$,
- $x_1$ is a 'successor' of $x_2$ whenever $x_1 \in (B^{\ell\oplus 1})^{\mathfrak{A}}$ and $x_2 \in (B^\ell)^{\mathfrak{A}}$, for $0 \le \ell \le 2$,

where $\oplus$ denotes addition modulo 3. (Here we remind the reader that $\tau_1 \cdot \tau_2 = \mathbf{0}$ holds in a quasi-saw model $\mathfrak{A}$ iff $\tau_1^{\mathfrak{A}}$ and $\tau_2^{\mathfrak{A}}$ contain no common points of depth 0. This means, in particular, that $(\tau_1 \cdot \tau_2)^{\mathfrak{A}} = \emptyset$ may hold even though $\tau_1^{\mathfrak{A}} \cap \tau_2^{\mathfrak{A}} \neq \emptyset$, i.e., $\mathfrak{A} \models C(\tau_1, \tau_2)$.)

Suppose also that the following formulas hold in $\mathfrak{A}$:

$$\textstyle\sum_{T_k \in \mathcal{T}} T_k^i = \mathbf{1}, \qquad 1 \le i \le s, \tag{5.8}$$

$$T_{k_1}^i \cdot T_{k_2}^i = \mathbf{0}, \qquad 1 \le i \le s, \quad T_{k_1}, T_{k_2} \in \mathcal{T}, \quad k_1 \neq k_2, \tag{5.9}$$

$$T_{k_1}^i \cdot T_{k_2}^{i+1} = \mathbf{0}, \qquad 1 \le i < s, \quad \mathit{top}(T_{k_1}) \neq \mathit{bot}(T_{k_2}), \quad T_{k_1}, T_{k_2} \in \mathcal{T}, \tag{5.10}$$

$$T_k^1 = \mathbf{0}, \qquad \mathit{bot}(T_k) \neq b, \quad T_k \in \mathcal{T}, \tag{5.11}$$

$$T_k^s = \mathbf{0}, \qquad \mathit{top}(T_k) \neq t, \quad T_k \in \mathcal{T}. \tag{5.12}$$

Then, by (5.8)–(5.9), for every $x \in W_0$ there is a unique sequence of tile types $T_{k_1}, \dots, T_{k_s}$ with $x \in (T_{k_i}^i)^{\mathfrak{A}}$, for $1 \le i \le s$. In this case we set $\mathit{left}^i(x) = \mathit{left}(T_{k_i})$, $\mathit{top}^i(x) = \mathit{top}(T_{k_i})$, etc. We also set $\mathit{left}(x) = \mathit{left}^1(x), \dots, \mathit{left}^s(x)$, $\mathit{right}(x) = \mathit{right}^1(x), \dots, \mathit{right}^s(x)$. Then, by (5.5) and (5.10)–(5.12), both $\mathit{left}(x)$ and $\mathit{right}(x)$ are configurations of $M$ and $\mathit{left}(x) \to \mathit{right}(x)$.

Consider now the following formulas, for $1 \le i \le s$, $0 \le \ell \le 2$, $T_{k_1}, T_{k_2} \in \mathcal{T}$:

$$\neg C(B^\ell \cdot T_{k_1}^i,\ B^{\ell\oplus 1} \cdot T_{k_2}^i), \qquad \mathit{right}(T_{k_1}) \neq \mathit{left}(T_{k_2}), \tag{5.13}$$

$$\neg C(B^\ell \cdot T_{k_1}^i,\ B^\ell \cdot T_{k_2}^i), \qquad k_1 \neq k_2. \tag{5.14}$$

Suppose that all of (5.6)–(5.14) are true in $\mathfrak{A}$. It easy to see that, if $x, y \in W_0$ and there exists $z \in W$ with $zRx$ and $zRy$, then:

$$\mathit{right}(x) = \mathit{left}(y) \text{ whenever } x \in (B^\ell)^{\mathfrak{A}} \text{ and } y \in (B^{\ell\oplus 1})^{\mathfrak{A}}, \text{ for } 0 \le \ell \le 2; \tag{5.15}$$

$$\mathit{left}(x) = \mathit{left}(y) \text{ and } \mathit{right}(x) = \mathit{right}(y) \text{ whenever } x, y \in (B^\ell)^{\mathfrak{A}}, \text{ for } 0 \le \ell \le 2. \tag{5.16}$$

Finally, we require the following formulas:

$$T_{i_1}^1 \cdot \ \cdots \ \cdot T_{i_s}^s \neq \mathbf{0}, \tag{5.17}$$

$$T_{j_1}^1 \cdot \ \cdots \ \cdot T_{j_s}^s \neq \mathbf{0}, \tag{5.18}$$

where $\mathit{left}(T_{i_1}), \dots, \mathit{left}(T_{i_s})$ is the initial configuration (with $\vec{a}$ written on the tape) and $\mathit{right}(T_{j_1}), \dots, \mathit{right}(T_{j_s})$ is the accepting configuration (with empty tape and the head scanning the first cell). Denote the conjunction of (5.6)–(5.18) by $\Psi(M, \vec{a})$. Clearly, the length of this $\mathcal{C}$-formula is polynomial in the size of $M$ and $\vec{a}$. We proceed to show: ($i$) if $\Psi(M, \vec{a})$ is satisfiable over ConRegC, then $M$ accepts $\vec{a}$; ($ii$) if $M$ accepts $\vec{a}$, then $\Psi(M, \vec{a})$ is satisfiable over $\mathrm{RC}(\mathbb{R}^n)$ for any $n \geq 1$. This proves the theorem.

Suppose $\Psi(M, \vec{a})$ is satisfiable over ConRegC. By Lemmas 4.1 and 4.2, it is satisfied in some model $\mathfrak{A}$ over a finite connected quasi-saw $(W, R)$. From (5.17) and (5.18), choose points $u, u'$ of depth 0 such that *left*$(u)$ is the initial configuration and *right*$(u')$ the accepting configuration. Since $(W, R)$ is connected, there exists a sequence $u_0, v_1, \dots, u_{m-1}, v_m, u_m$ with the points $u_i$ of depth 0 and the points $v_j$ of depth 1, such that: $u_0 = u$, $u_m = u'$, and, for all $i$ $(1 \leq i \leq m)$, $(v_i, u_{i-1}) \in R$ and $(v_i, u_i) \in R$. We may assume without loss of generality that the $u_i$ and $v_j$ are all distinct. It should be noted that *left*$(u_i)$ is not the accepting configuration for any $i \leq m$. For brevity, we write $\mathfrak{c}_i$ for *left*$(u_i)$ and $\mathfrak{c}_{m+1}$ for *right*$(u_m)$. We shall show that $\mathfrak{c}_0, \dots, \mathfrak{c}_m, \mathfrak{c}_{m+1}$ contains a sub-sequence that is an accepting run of $M$. From (5.15) and (5.16), and the fact that *left*$(x) \to$ *right*$(x)$ for any point $x$, we see that, for all $i$ $(0 \leq i \leq m)$, one of the following conditions holds: ($i$) $\mathfrak{c}_i = \mathfrak{c}_{i+1}$; ($ii$) $\mathfrak{c}_i \to \mathfrak{c}_{i+1}$; or ($iii$) $\mathfrak{c}_i \leftarrow \mathfrak{c}_{i+1}$. We shall presently establish the following claim:

**Claim 5.7.** *If $0 \leq j \leq m$ and $\mathfrak{c}_j \leftarrow \mathfrak{c}_{j+1}$, then there exists $k$ such that $j+1 < k \leq m$ and $\mathfrak{c}_k = \mathfrak{c}_j$.*

Taking this claim on trust for the moment, define the sub-sequence $\mathfrak{c}_{j_0}, \mathfrak{c}_{j_1}, \dots$ by setting $j_0$ to be the largest $j \leq m$ such that $\mathfrak{c}_j = \mathfrak{c}_0$, and, for $i \geq 0$, $j_{i+1}$ to be the largest $j \leq m$ such that $\mathfrak{c}_j = \mathfrak{c}_{j_i+1}$, until we eventually reach (say), $\mathfrak{c}_{j_K} = \mathfrak{c}_m$. It is then immediate from the claim that $\mathfrak{c}_{j_i} \to \mathfrak{c}_{j_i+1} = \mathfrak{c}_{j_{i+1}}$ for all $i$ $(0 \leq i < K)$, and we have the desired accepting run of $M$.

**Proof of Claim 5.7.** The claim is proved by (decreasing) induction on $j$. For $j = m$, the result is trivial, since $\mathfrak{c}_{m+1}$ has no successor configurations. Assume, then, that $0 \leq j < m$, and the claim holds for all larger values of $j$ up to $m$. Let $k$ be the largest number $(j+1 \leq k \leq m+1)$ such that $\mathfrak{c}_{j+1} = \mathfrak{c}_k$. Thus, $k \leq m$ (since $\mathfrak{c}_{j+1}$ is, by assumption, not the accepting configuration), and $\mathfrak{c}_k \to \mathfrak{c}_{k+1}$ (since otherwise, using the inductive hypothesis, we could find a larger value $k'$ with $\mathfrak{c}_{j+1} = \mathfrak{c}_{k'}$). Thus, $\mathfrak{c}_j \leftarrow \mathfrak{c}_{j+1} = \mathfrak{c}_k \to \mathfrak{c}_{k+1}$. But $M$ is deterministic, so $\mathfrak{c}_j = \mathfrak{c}_{k+1}$, completing the induction, and proving the claim. $\square$

The proof of the converse direction is straightforward: for an accepting computation $\mathfrak{c}_0 \to \dots \to \mathfrak{c}_{m+1}$ of $M$ on $\vec{a}$, we construct a model $\mathfrak{A}$ over $\mathrm{RC}(\mathbb{R})$. Define the closed intervals $I_0, \dots, I_m$ in $\mathbb{R}$ by

$$I_j = \begin{cases} (-\infty, 0], & \text{if } j = 0, \\ [i-1, i], & \text{if } 0 < j < m, \\ [m-1, +\infty), & \text{if } j = m. \end{cases}$$

Note that, given any valuation over $\mathrm{RC}(\mathbb{R})$ in which all the variables $T_k^i$ are interpreted as unions of the intervals $I_0, \dots, I_m$, we may meaningfully write statements of the form *left*$(I_j) = \mathfrak{c}$ and *right*$(I_j) = \mathfrak{c}$, for $0 \leq j \leq m$, and $\mathfrak{c}$ a configuration of $M$. Now define such a valuation $\cdot^{\mathfrak{A}}$ in which, for all $j$ $(0 \leq j \leq m)$, $I_j \subseteq (B^\ell)^{\mathfrak{A}}$ if and only if $j \equiv \ell \pmod 3$; and, for all $j$ $(0 \leq j < m)$, *left*$(I_j) = \mathfrak{c}_j$ and *right*$(I_j) = \mathfrak{c}_{j+1}$. It can be readily checked that the resulting model satisfies $\Psi(M, \vec{a})$. For $n > 1$, we may construct a model of $\Psi(M, \vec{a})$ over $\mathrm{RC}(\mathbb{R}^n)$ by cylindrification of $\mathfrak{A}$ in the obvious way. $\square$

**Corollary 5.8.** *Sat*($\mathcal{C}$, ConRegC), *Sat*($\mathcal{C}c^1$, RegC) *and Sat*($\mathcal{C}$, RC($\mathbb{R}^n$)) *for any* $n \geq 1$ *are all* PSpace-*complete.*

**Proof.** Follows from Theorems 5.4, 5.6 and Corollary 5.5. □

Having established a lower bound for $\mathcal{C}c^1$, we now proceed to do the same for the larger language $\mathcal{C}c$. Observe that when constructing a model for an $\mathcal{S}4_uc^1$-formula with one positive occurrence of $c(\tau)$, in the proof of Theorem 5.4, we could check the 'connectability' of two $\tau$-points by an (exponentially long) path using a PSpace-algorithm, because we did not need to keep in memory all the points on the path. However, if two statements $c(\tau_1)$ and $c(\tau_2)$ have to be satisfied, then, while connecting two $\tau_1$-points using a path, one has to check whether the $\tau_2$-points on that path can be connected by a path, which, in turn, can contain another $\tau_1$-point, and so on. The crucial idea in the proof below is to simulate infinite binary (*non-transitive*) trees using quasi-saws. Roughly, the construction is as follows. We start by representing the root $v_0$ of the tree as a point also denoted by $v_0$ (see Fig. 8), which is forced to be connected to an auxiliary point $w$ by means of some $c(\tau_0)$. On the connecting path from $v_0$ to $w$ we represent the two successors $v_1$ and $v_2$ of the root, which are forced to be connected in turn to $w$ by some other $c(\tau_1)$. On each of the two connecting paths, we again take two points representing the successors of $v_1$ and $v_2$, respectively. We treat these four points in the same way as $v_0$, reusing $c(\tau_0)$, and proceed *ad infinitum*, alternating between $\tau_0$ and $\tau_1$ when forcing the paths which generate the required successors. Of course, we also have to pass certain information from a node to its two successors. Such information can be propagated along connected regions. Note now that all points are connected to $w$. To distinguish between the information we have to pass from distinct nodes of even (respectively, odd) level to their successors, we have to use *two* connectedness formulas of the form $c(f_i + a)$, $i = 0, 1$, in such a way that the $f_i$ points form initial segments of the paths to $w$ and $a$ contains $w$. The $f_i$-segments are then used locally to pass information from a node to its successors without conflict. We now present the reduction in more detail.

**Theorem 5.9.** *Sat*($\mathcal{C}c$, RegC) *and Sat*($\mathcal{C}c$, ConRegC) *are* ExpTime-*hard.*

**Proof.** The proof is by reduction of the following problem. Denote by $\mathcal{D}_2^f$ the bimodal logic (with $\Box_1$ and $\Box_2$) determined by Kripke models based on the full infinite binary tree $\mathfrak{G} = (V, R_1, R_2)$ with *functional* accessibility relations $R_1$ and $R_2$. Consider the *global consequence relation* $\models_2^f$ defined as follows: $\chi \models_2^f \psi$ iff $\mathfrak{K} \models \chi$ implies $\mathfrak{K} \models \psi$, for every Kripke model $\mathfrak{K}$ based on $\mathfrak{G}$. This global consequence relation is ExpTime-hard, see, e.g., [52]. We construct a $\mathcal{C}c$-formula $\Phi(\chi, \psi)$, for any $\mathcal{D}_2^f$-formulas $\chi$, $\psi$, such that (*i*) $|\Phi(\chi, \psi)|$ is polynomial in $|\chi| + |\psi|$, (*ii*) if $\Phi(\chi, \psi)$ is satisfiable over RegC then $\chi \not\models_2^f \psi$, and (*iii*) if $\chi \not\models_2^f \psi$ then $\Phi(\chi, \psi)$ is satisfiable over ConRegC. While constructing $\Phi(\chi, \psi)$, we will assume that $\mathfrak{A}$ is a quasi-saw model induced by $(W, R)$ and $W_0$ is the set of points of depth 0 in $(W, R)$.

Let $sub(\chi, \psi)$ be the closure under single negation of the set of subformulas of $\chi$, $\psi$. For each $\varphi \in sub(\chi, \psi)$ we take a fresh variable $q_\varphi$, and for each $\Box_i\varphi \in sub(\chi, \psi)$, a pair of fresh variables $m_\varphi^{i,j}$, $j = 0, 1$. We also need fresh variables $a$ and $s_j^i$, for $j = 0, 1$ and $0 \leq i \leq 6$. Let $d = s_0^0 + s_1^0$. Intuitively, $d$ simulates the domain of the binary tree, where $s_0^0$ and $s_1^0$ stand for nodes with even and, respectively, odd distance from the root. Suppose that the

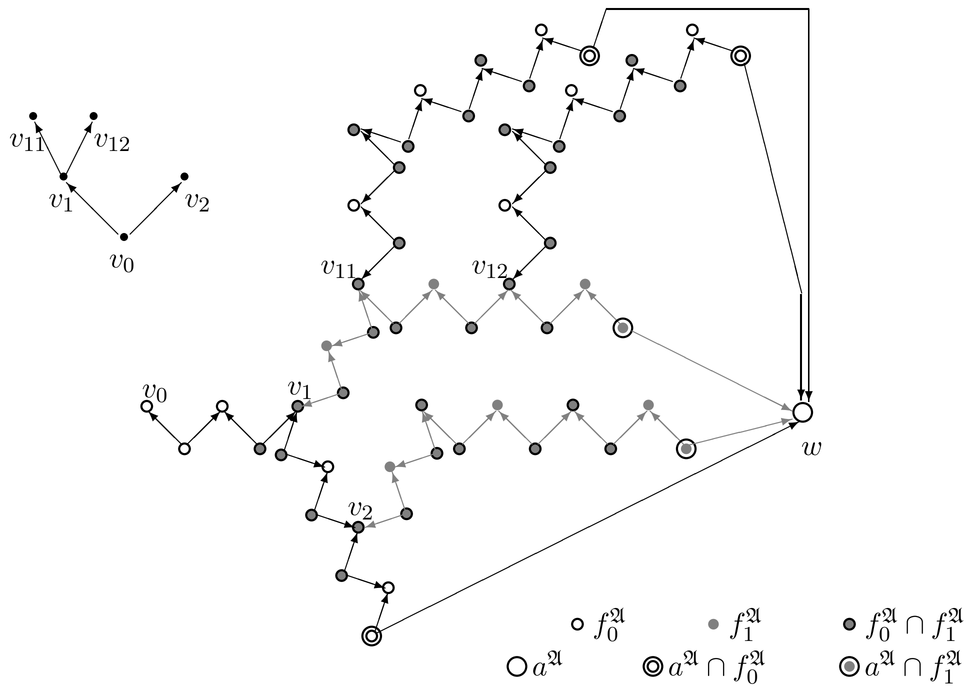



Figure 8: First 4 steps of encoding the full binary tree using 7-saws.

following $\mathcal{C}c$-formulas hold in $\mathfrak{A}$

$$\big(a = s_0^6\big) \ \wedge\ \big(a = s_1^6\big) \quad \wedge \quad \big(a \neq \mathbf{0}\big) \quad \wedge \quad c(f_0 + a) \quad \wedge \quad c(f_1 + a), \tag{5.19}$$

$$\bigwedge_{0 \leq k < k' \leq 6} \big(s_j^k \cdot s_j^{k'} = \mathbf{0}\big) \quad \wedge \quad \bigwedge_{\substack{0 \leq k < k' \leq 6 \\ |k - k'| > 1}} \neg C(s_j^k, s_j^{k'}), \tag{5.20}$$

where $f_j = s_j^0 + s_j^1 + s_j^2 + s_j^3 + s_j^4 + s_j^5$, for $j = 0, 1$. It follows that, for $j = 0, 1$, if there is a point $x_0 \in (s_j^0)^{\mathfrak{A}} \cap W_0$ then there is a (not necessarily unique) sequence of points $x_1, x_2, x_3, x_4, x_5$ from the same component of $f_j^{\mathfrak{A}}$ such that $x_i \in (s_j^i)^{\mathfrak{A}} \cap W_0$, $1 \leq i \leq 5$. Points $x_2$ and $x_4$ will be used to construct similar sequences for the two successors of the node represented by $x_0$: if (5.19)–(5.20) and

$$s_0^{2i} \ \leq\ s_1^0 \qquad \text{and} \qquad s_1^{2i} \ \leq\ s_0^0, \qquad \text{for } i = 1, 2, \tag{5.21}$$

hold in $\mathfrak{A}$ and $x_0 \in (s_j^0)^{\mathfrak{A}} \cap W_0$, then one can recover from $\mathfrak{A}$ the infinite binary tree with the root at $x_0$. The formula

$$\big(q_{\neg\psi} \cdot s_0^0 \neq \mathbf{0}\big) \quad \wedge \quad \big(d \ \leq\ q_\chi\big) \tag{5.22}$$

ensures then that there is $x_0 \in (s_0^0)^{\mathfrak{A}} \cap W_0$, the root of the tree, in which $\neg\psi$ holds, and $\chi$ holds everywhere in the tree, while the formulas

$$d \cdot q_{\neg\varphi} = d \cdot (-q_\varphi), \qquad\qquad d \cdot q_{\varphi_1 \wedge \varphi_2} = d \cdot (q_{\varphi_1} \cdot q_{\varphi_2}), \tag{5.23}$$

for all $\neg\varphi, \varphi_1 \wedge \varphi_2 \in sub(\chi,\psi)$, capture the meaning of the Boolean connectives from $sub(\chi,\psi)$ relativized to $d$. The formulas, for all $\Box_i\varphi \in sub(\chi,\psi)$ and $j = 0, 1$,

$$\neg C(f_j \cdot m_\varphi^{i,j},\ f_j \cdot (-m_\varphi^{i,j})), \tag{5.24}$$

$$s_j^0 \cdot q_{\Box_i\varphi} = s_j^0 \cdot m_\varphi^{i,j}, \tag{5.25}$$

$$s_j^{2i} \cdot m_\varphi^{i,j} = s_j^{2i} \cdot q_\varphi \tag{5.26}$$

are used to propagate information regarding $\Box_i\varphi$ along the components of $f_j$ using the markers $m_\varphi^{i,j}$. We define $\Phi(\chi,\psi)$ to be the conjunction of all the above formulas. Clearly, $|\Phi(\chi,\psi)|$ is polynomial in $|\chi| + |\psi|$ and contains only two occurrences of the connectedness predicate, $c$, in (5.19).

Suppose that $\Phi(\chi,\psi)$ is satisfied over RegC; we proceed to show that $\chi \not\models_2^f \psi$. By Lemmas 4.1 and 4.2, $\Phi(\chi,\psi)$ is satisfied in a finite quasi-saw model $\mathfrak{A}$ induced by some $(W,R)$. Denote by $W_0$ the set of points of depth 0 in $(W,R)$. Our aim is to construct, by induction, a Kripke model $\mathfrak{K}$ based on the full infinite binary tree $\mathfrak{G} = (V, R_1, R_2)$ such that $\mathfrak{K} \models \chi$ and $\mathfrak{K}, v_0 \not\models \psi$, for the root $v_0$ of $\mathfrak{G}$. The points of $V$ will be *copies* of some points in $d^{\mathfrak{A}} \cap W_0$. If $v \in V$ is a copy of $x \in W$, then we write $x = \varkappa(v)$.

Step 0. Take some $x_0 \in (q_{\neg\psi} \cdot s_0^0)^{\mathfrak{A}} \cap W_0$. It exists by (5.22). Take a fresh $v_0$ and let $\varkappa(v_0) = x_0$ and $\mathfrak{G}^0 = (\{v_0\}, \emptyset, \emptyset)$ and set $v_0 \models p$ for each propositional variable $p \in sub(\chi,\psi)$ such that $x_0 \in q_p^{\mathfrak{A}}$.

Step $n+1$. Suppose that we have already constructed $\mathfrak{G}^n = (V^n, R_1^n, R_2^n)$ and defined a valuation in $\mathfrak{G}^n$ for the variables in $sub(\chi,\psi)$. Let $v$ be a point of minimal co-depth in $\mathfrak{G}^n$ which does not have $R_1$- and $R_2$-successors yet, and let $\varkappa(v) = x \in (s_j^0)^{\mathfrak{A}} \cap W_0$, for some $j \in \{0,1\}$. We have $x \in f_j^{\mathfrak{A}}$, and so by (5.19), there is a finite path $\pi$ from $x$ to some $w \in (s_j^6)^{\mathfrak{A}} \cap W_0$ such that all the points on this path belong to $(f_j + a)^{\mathfrak{A}}$, and so to $\bigcup_{k=0}^6 (s_j^k)^{\mathfrak{A}}$. Fix such a path $\pi$. By the first conjunct of (5.20), $x \neq w$ and in view of the second conjunct of (5.20), of all $(s_j^k)^{\mathfrak{A}}$ on the path $\pi$ the set $(s_j^0)^{\mathfrak{A}}$ can only be in contact with (in fact, externally connected to) $(s_j^1)^{\mathfrak{A}}$. Take the last point $y_1 \in (s_j^1)^{\mathfrak{A}} \cap W_0$ on $\pi$. By a similar argument, the next point of depth 0 on $\pi$ can only be from $(s_j^2)^{\mathfrak{A}}$. Let $x_1$ be the last point in $(s_j^2)^{\mathfrak{A}} \cap W_0$ on $\pi$. In the same way we find the last point $y_2 \in (s_j^3)^{\mathfrak{A}}$ of depth 0 on $\pi$, and then the last point $x_2 \in (s_j^4)^{\mathfrak{A}} \cap W_0$ on $\pi$. Let $v_1, v_2$ be fresh copies of $x_1, x_2$, respectively, i.e., $\varkappa(v_i) = x_i$, $i = 1, 2$. Now we set $\mathfrak{G}^{n+1} = (V^n \cup \{v_1, v_2\}, R_1^n \cup \{(v, v_1)\}, R_2^n \cup \{(v, v_2)\})$, and $v_i \models p$ for each propositional variable $p \in sub(\chi,\psi)$ such that $x_i \in q_p^{\mathfrak{A}}$, for $i = 1, 2$. Note that, by (5.21), $x_1, x_2 \in (s_{j\oplus 1}^0)^{\mathfrak{A}}$, where $\oplus$ denotes addition modulo 2, and we can move on to Step $n+2$.

Step $\omega$. Finally, we set $V = \bigcup_{n<\omega} V^n$, $R_1 = \bigcup_{n<\omega} R_1^n$, $R_2 = \bigcup_{n<\omega} R_2^n$, $\mathfrak{G} = (V, R_1, R_2)$. Clearly, $\mathfrak{G}$ is the full binary tree with functional $R_1$ and $R_2$. The Kripke model $\mathfrak{K}$ is based on $\mathfrak{G}$ with the valuation defined by the inductive procedure above.

It remains to show by induction that, for every $\varphi \in sub(\chi,\psi)$ and every $v \in V$,

$$\varkappa(v) \in q_\varphi^{\mathfrak{A}} \qquad \text{iff} \qquad \mathfrak{K}, v \models \varphi.$$

The basis of induction follows from the definition, and the case of Boolean connectives from (5.23). Suppose now that $v$ is the point considered at Step $n+1$ above. Let $i = 1$ or 2 and $v_i \in V$ be such that $(v, v_i) \in R_i$ (it is defined uniquely as $R_i$ is functional). We have

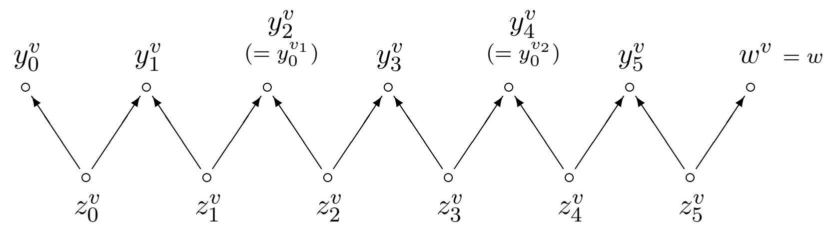



Figure 9: A 7-saw for $v$.

$\varkappa(v) \in (s_j^0)^{\mathfrak{A}}$, for either $j = 0$ or $j = 1$, and, by (5.21), $\varkappa(v_i) \in (s_{j\oplus 1}^0)^{\mathfrak{A}}$. Then $\varkappa(v) \in q^{\mathfrak{A}}_{\Box_i\varphi}$ iff, by (5.25), $\varkappa(v) \in (m_\varphi^{i,j})^{\mathfrak{A}}$ iff, by (5.24), $\varkappa(v_i) \in (m_\varphi^{i,j})^{\mathfrak{A}}$ iff, by (5.26), $\varkappa(v_i) \in q_\varphi^{\mathfrak{A}}$, iff, by IH, $\mathfrak{K}, v_i \models \varphi$ iff, by functionality of $R_i$, $\mathfrak{K}, v \models \Box_i\varphi$.

Suppose, finally, that $\chi \not\models_2^f \psi$; we proceed to show that $\Phi(\chi,\psi)$ is satisfied over ConRegC, thus completing the proof of the theorem. Let $\mathfrak{K}$ be a model for $\mathcal{D}_2^f$ based on the full infinite binary tree $\mathfrak{G} = (V, R_1, R_2)$ with root $v_0$ such that $\mathfrak{K} \models \chi$ and $\mathfrak{K}, v_0 \not\models \psi$. We construct a connected quasi-saw model $\mathfrak{A}$ satisfying $\Phi(\chi,\psi)$. The model $\mathfrak{A}$ will be induced by the quasi-saw $(W, R)$ constructed by induction from (infinitely many copies of) the 7-saws shown in Fig. 9. For each node $v$ of the infinite binary tree $\mathfrak{G}$, we take a fresh 7-saw $\mathfrak{S}^v = (S^v, R^v)$, where $S^v = \{y_i^v, z_i^v \mid 0 \leq i \leq 5\} \cup \{w^v\}$, $z_i^v R^v y_i^v$, for $0 \leq i \leq 5$, $z_i^v R^v y_{i+1}^v$, for $0 \leq i < 5$, and $z_5^v R^v w^v$, and identify the following points

$$y_2^v = y_0^{v_1}, \qquad y_4^v = y_0^{v_2}, \qquad w^{v_1} = w^{v_2} = w^v,$$

if $v_1$ and $v_2$ are the $R_1$- and $R_2$-successors of $v$. We present our construction in a step-by-step manner (see Fig. 8).

<u>*Step* 0.</u> We set $W^0 = \{y_0^{v_0}, w\}$, $R^0 = \emptyset$ and define a valuation $\cdot^{\mathfrak{A}_0}$ on $W^0$ by taking

- $(s_0^0)^{\mathfrak{A}_0} = \{y_0^{v_0}\}$, $\quad (s_1^0)^{\mathfrak{A}_0} = \emptyset$,
  $(s_0^i)^{\mathfrak{A}_0} = (s_1^i)^{\mathfrak{A}_0} = \emptyset$, for $0 < i \leq 5$, and $a^{\mathfrak{A}_0} = (s_0^6)^{\mathfrak{A}_0} = (s_1^6)^{\mathfrak{A}_0} = \{w\}$;
- $q_\varphi^{\mathfrak{A}_0} = \{y_0^{v_0} \mid \mathfrak{K}, v_0 \models \varphi\}$, for $\varphi \in sub(\chi,\psi)$;
- $(m_\varphi^{i,0})^{\mathfrak{A}_0} = \{y_0^{v_0}\}$, if $\mathfrak{K}, v_0 \models \Box_i\varphi$, and $(m_\varphi^{i,0})^{\mathfrak{A}_0} = \emptyset$, otherwise, and $(m_\varphi^{i,1})^{\mathfrak{A}_0} = \emptyset$, for $i = 1, 2$ and $\Box_i\varphi \in sub(\chi,\psi)$.

<u>*Step* $n+1$.</u> Suppose now that $v$ is a node of minimal co-depth in $\mathfrak{G}$ such that the constructed quasi-saw $(W^n, R^n)$ contains $y_0^v$ but does not contain a copy of $\mathfrak{S}^v$ for $v$ and let $v_1$ and $v_2$ be the $R_1$- and $R_2$-successors of $v$. Then we take a fresh 7-saw $\mathfrak{S}^v$ and 'hook' it to $(W^n, R^n)$ by identifying the first point of $\mathfrak{S}^v$ with $y_0^v$ and its last point with $w$ (see Fig. 8):

$$\begin{aligned} W^{n+1} &= W^n \cup \{y_1^v, y_0^{v_1}, y_3^v, y_0^{v_2}, y_5^v\} \cup \{z_i^v \mid 0 \leq i \leq 5\},\\ R^{n+1} &= R^n \cup \{(z_{2i-1}^v, y_0^{v_i}), (z_{2i}^v, y_0^{v_i}) \mid i = 1, 2\} \cup \{(z_0^v, y_0^v), (z_5^v, w)\}\\ &\quad \cup \{(z_i^v, y_i^v), (z_{i-1}^v, y_i^v) \mid i = 1, 3, 5\}, \end{aligned}$$

and define the valuation $\cdot^{\mathfrak{A}_{n+1}}$ by taking $p^{\mathfrak{A}_{n+1}} = p^{\mathfrak{A}_n} \cup p^{\mathfrak{A}_v}$, where $\cdot^{\mathfrak{A}_v}$ is defined for the new points as follows:

- $(s_j^0)^{\mathfrak{A}_v} = \{z_0^v\}$ and $(s_{j\oplus 1}^0)^{\mathfrak{A}_v} = \{z_{2i-1}^v, y_0^{v_i}, z_{2i}^v \mid i = 1, 2\}$;
- $(s_j^{2i})^{\mathfrak{A}_v} = \{z_{2i-1}^v, y_0^{v_i}, z_{2i}^v\}$ and $(s_{j\oplus 1}^{2i})^{\mathfrak{A}_v} = \{z_0^v \mid y_0^v \in (s_{j\oplus 1}^{2i})^{\mathfrak{A}_n}\}$, for $i = 1, 2$;
- $(s_j^i)^{\mathfrak{A}_v} = \{z_{i-1}^v, y_i^v, z_i^v\}$ and $(s_{j\oplus 1}^i)^{\mathfrak{A}_v} = \emptyset$, for $i = 1, 3, 5$;

- $a^{\mathfrak{A}_v} = (s_j^6)^{\mathfrak{A}_v} = (s_{j\oplus 1}^6)^{\mathfrak{A}_v} = \{z_5^v\}$,
- $q_\varphi^{\mathfrak{A}_v} = \{z_0^v \mid \mathfrak{K}, v \models \varphi\} \cup \{z_{2i-1}^v, y_0^{v_i}, z_{2i}^v \mid \mathfrak{K}, v_i \models \varphi, i = 1,2\}$, for $\varphi \in sub(\chi, \psi)$;
- $(m_\varphi^{i,j})^{\mathfrak{A}_v} = \bigcup_{i=0}^6 (s_j^i)^{\mathfrak{A}_v}$, if $\mathfrak{K}, v \models \Box_i\varphi$, and $(m_\varphi^{i,j})^{\mathfrak{A}_v} = \emptyset$, otherwise, and
$$(m_\varphi^{i,j\oplus 1})^{\mathfrak{A}_v} = \{z_0^v \mid y_0^v \in (m_\varphi^{i,j\oplus 1})^{\mathfrak{A}_n}\} \cup \{z_{2i-1}^v, y_0^{v_i}, z_{2i}^v \mid \mathfrak{K}, v_i \models \Box_i\varphi\},$$
for $\Box_i\varphi \in sub(\chi,\psi)$, $i = 1, 2$;

where $j = 0$ if $v$ is of *even* co-depth in $\mathfrak{G}$ and $j = 1$ otherwise, and $\oplus$ denotes addition modulo 2.

Step $\omega$. Finally, we set $W = \bigcup_{n<\omega} W^n$, $R = \bigcup_{n<\omega} R^n$ and $\cdot^{\mathfrak{A}} = \bigcup_{n<\omega} \cdot^{\mathfrak{A}_n}$.

Note that $y_1^v$ and $y_3^v$ are required to make the set of points in $f_j^{\mathfrak{A}}$ representing a node $v$ of $\mathfrak{G}$ disconnected from the subset of $f_j^{\mathfrak{A}}$ representing another node $v'$ of $\mathfrak{G}$ and thus satisfy (5.24); $y_5^v$ are required to satisfy the last conjunct of (5.20). We leave the remaining details to the reader. □

**Corollary 5.10.** *Sat*$(\mathcal{C}c, \text{RegC})$ *and Sat*$(\mathcal{C}c, \text{ConRegC})$ *are* ExpTime*-complete.*

**Proof.** Follows from Theorems 5.3 and 5.9. □

The argument of Theorem 5.9 goes through unproblematically when we restrict attention to the spaces $\mathbb{R}^n$ for $n \geq 3$, thus showing that *Sat*$(\mathcal{C}c, \text{RC}(\mathbb{R}^n))$ is ExpTime-hard for these values of $n$. If $n = 2$, however, this approach fails. The difficulty is that the construction of the model $\mathfrak{A}$ of $\Phi(\chi,\psi)$ from the Kripke structure $\mathfrak{K}$ will result in sets $(s_j^k)^{\mathfrak{A}}$ that have infinitely many components lying in a bounded region of the plane. These infinitely many components will give rise to accumulation points, meaning that we can no longer guarantee the truth of the formulas $\neg C(s_j^k, s_j^{k'})$ of (5.20) and (5.24). Fortunately, the proof can be rescued, either by means of the finite model property for $\models_2^f$, or, alternatively, via the following explicit construction based on polynomial-space alternating Turing machines. We remind the reader, in this connection, that APSpace = ExpTime (see e.g., [32]).

**Theorem 5.11.** *The problems Sat*$(\mathcal{C}c, \text{RC}(\mathbb{R}^n))$ *for all* $n \geq 2$ *are* ExpTime*-hard.*

**Proof.** Let $M$ be an alternating Turing machine that uses $\leq p(n)$ cells on each input of length $n$, for some polynomial $p(\cdot)$. Let $q_Y$ and $q_N$ be the accepting and rejecting states of $M$, respectively (they have no transitions from them). Without loss of generality we assume that every branch of the computation tree of $M$ is of *finite length* and its final state is either $q_Y$ or $q_N$ (indeed, for if it is not the case we can augment $M$ with a subroutine that at each step of $M$ increases the current step number written on an auxiliary tape and makes $M$ go into the rejecting state $q_N$ whenever the step number reaches the total number of configurations of $M$; as $M$ uses a polynomial working tape, we need only a polynomial number of cells on that auxiliary tape to represent the maximum number of configurations of $M$). All states of $M$ (except $q_Y$ and $q_N$) are divided into existential and universal states. We assume without loss of generality that there are exactly two transitions from each state, i.e., for $q \in Q \setminus \{q_Y, q_N\}$, $a \in \Sigma$ and $j = 1, 2$, we have $(q, a) \to_j (q', a', d)$, where $d \in \{-1, 0, +1\}$. We assume that $M$ never goes beyond special markers at the beginning and the end of the tape. We denote $p(n)$ by $s$, and employ propositional variables with the following intuitive meanings:

- $H_{q,i}$, for $1 \leq i \leq s$ and $q \in Q$, to say that the head is scanning the $i$th cell and the state is $q$,
- $S_{a,i}$, for $1 \leq i \leq s$ and $a \in \Sigma$, to say that the $i$th cell contains $a$,
- $A$ to say that the computation is accepting.

Let $\chi_M$ be the conjunction of the following formulas, for all $i$, $1 \leq i \leq s$,

$$\begin{array}{ll} H_{q,i} \wedge S_{a,i} \to \Box_j (H_{q',i+d} \wedge S_{a',i}), & \text{for } (q,a) \to_j (q',a',d),\ j=1,2,\ d \in \{-1,0,+1\},\\ H_{q,i} \wedge S_{a,k} \to \Box_j S_{a,k}, & \text{for each } k \neq i \text{ and } j = 1,2,\\ H_{q_Y,i} \to A, & \\ H_{q,i} \wedge (\Diamond_1 A \wedge \Diamond_2 A) \to A, & \text{for each universal state } q \in Q,\\ H_{q,i} \wedge (\Diamond_1 A \vee \Diamond_2 A) \to A, & \text{for each existential state } q \in Q. \end{array}$$

Let $\vec{a} = a^1, \dots, a^n$ be an input. For convenience, let $a^i$ be blank, for each $n < i \leq s$. Write

$$\psi_{\vec{a}} \quad = \quad (H_{q_0,1} \wedge S_{a^1,1} \wedge \cdots \wedge S_{a^s,s}) \to A,$$

where $q_0$ is the initial state of $M$. Clearly, $M$ accepts $\vec{a}$ iff $\chi_M \models_2^f \psi_{\vec{a}}$. Moreover, since we assume Turing machines to terminate each branch of a computation in $q_Y$ or $q_N$ after a finite number of steps, it follows that $M$ accepts $\vec{a}$ if and only if $A$ is satisfied at the root of every finite binary tree such that

(a) $\chi_M$ is satisfied in every point of the tree,
(b) every point satisfying $H_{q,i}$, for $q \in Q \setminus \{q_Y, q_N\}$ and $1 \leq i \leq s$, has a pair of successors (in particular: is not a leaf node),
(c) $H_{q_0,1} \wedge S_{a^1,1} \wedge \cdots \wedge S_{a^s,s}$ is satisfied at the root.

Denote by $\Phi'(\chi_M, \psi_{\vec{a}})$ the conjunction of formulas (5.19)–(5.20) and (5.22)–(5.26), constructed for $\chi_M$ and $\psi_{\vec{a}}$ in the proof of Theorem 5.9, together with the following replacement of (5.21), for $1 \leq k \leq s$ and $q' \in Q \setminus \{q_Y, q_N\}$:

$$q_{H_{q',k}} \cdot s_0^{2i} \ \leq\ s_1^0 \qquad \text{and} \qquad q_{H_{q',k}} \cdot s_1^{2i} \ \leq\ s_0^0, \qquad \text{for } i = 1,2. \tag{5.27}$$

If $\Phi'(\chi_M, \psi_{\vec{a}})$ is satisfied in a model over $\mathrm{RC}(\mathbb{R}^n)$, for $n \geq 2$, then it follows from the proof of Theorem 5.9 that $\chi_M \not\models_2^f \psi_{\vec{a}}$, and thus $M$ does not accept $\vec{a}$. Conversely, if $M$ does not accept $\vec{a}$, then there is a finite binary tree satisfying (a)–(c); we then construct a model of $\Phi'(\chi_M, \psi_{\vec{a}})$ over $\mathrm{RC}(\mathbb{R}^2)$ by representing nodes of the finite binary tree together with the 'sink' node $w$ (see Fig. 8) by rectangles of decreasing dimensions. □

Having established a lower bound for $\mathcal{C}c$, we now proceed to do the same for the larger language $\mathcal{C}cc$.

**Theorem 5.12.** *$Sat(\mathcal{C}cc, \mathrm{RegC})$ is* NExpTime*-hard. In fact, $Sat(\mathcal{C}cc, \mathrm{ConRegC})$ and $Sat(\mathcal{C}cc, \mathrm{RC}(\mathbb{R}^n))$ for all $n \geq 2$ are* NExpTime*-hard.*

**Proof.** The proof is by reduction of the NExpTime-complete $2^d \times 2^d$ *tiling problem* [55]: Given $d < \omega$, a finite set $\mathcal{T}$ of *tile types*—i.e., 4-tuples of colours $T = (\mathit{left}(T), \mathit{top}(T), \mathit{right}(T), \mathit{bot}(T))$—and a $T_0 \in \mathcal{T}$, decide whether $\mathcal{T}$ can tile the $2^d \times 2^d$ grid in such a way that $T_0$ is placed onto $(0,0)$. In other words, the problem is to decide whether there is a function $f$ from $\{(n,m) \mid n,m < 2^d\}$ to $\mathcal{T}$ such that $\mathit{top}(f(n,m)) = \mathit{bot}(f(n,m+1))$, for all $n, m+1 < 2^d$, $\mathit{right}(f(n,m)) = \mathit{left}(f(n+1,m))$, for all $n+1, m < 2^d$, and $f(0,0) = T_0$. We construct a $\mathcal{C}cc$-formula $\Theta(\mathcal{T}, T_0, d)$ such that $(i)$ $|\Theta(\mathcal{T}, T_0, d)|$ is polynomial in $|\mathcal{T}|$ and $d$, $(ii)$ if $\Theta(\mathcal{T}, T_0, d)$ is satisfiable over $\mathrm{RegC}$ then $\mathcal{T}$ tiles the $2^d \times 2^d$ grid, with $T_0$ being placed onto $(0,0)$, and $(iii)$ if $\mathcal{T}$ tiles the $2^d \times 2^d$ grid, with $T_0$ being placed onto $(0,0)$, then $\Theta(\mathcal{T}, T_0, d)$ is satisfiable over $\mathrm{RC}(\mathbb{R}^n)$, $n \geq 2$. While constructing the formula, we will assume that $\mathfrak{A}$ is a quasi-saw model induced by $(W, R)$ and $W_0$ is the set of points of depth 0 in $(W, R)$.

We partition all points of $W_0$ with the help of a pair of variable triples $H^0, H^1, H^2$ and $V^0, V^1, V^2$. Suppose that the formulas, for $0 \leq \ell \leq 2$,

$$H^0 + H^1 + H^2 = \mathbf{1}, \tag{5.28}$$

$$H^0 \cdot H^1 = \mathbf{0}, \qquad H^1 \cdot H^2 = \mathbf{0}, \qquad H^2 \cdot H^0 = \mathbf{0}, \tag{5.29}$$

and their $V$-counterparts hold in $\mathfrak{A}$. Then every point in $W_0$ is in exactly one of the $(H^\ell)^{\mathfrak{A}}$, $0 \leq \ell \leq 2$, and exactly one of the $(V^\ell)^{\mathfrak{A}}$, $0 \leq \ell \leq 2$ (these variables play the same role as the $B^\ell$ in the proof of Theorem 5.6).

To encode coordinates of the tiles in binary, we take a pair of variables $X_j$ and $Y_j$, for each $j$, $d \geq j \geq 1$. For $n < 2^d$, let $n_X$ be the $\mathcal{B}$-term $X'_d \cdot X'_{d-1} \cdot \ \cdots \ \cdot X'_1$, where $X'_j = X_j$ if the $j$th bit in the binary representation of $n$ is 1, and $X'_j = -X_j$ otherwise. For a point $u \in W_0$, we denote by $X(u)$ the binary $d$-bit number $n$, called the *$X$-value* of $u$, such that $u \in n_X^{\mathfrak{A}}$ (note that the $X$-value is defined uniquely for the points in $W_0$); the $j$th bit of $X(u)$ is denoted by $X_j(u)$. The term $m_Y$, the *$Y$-value* $Y(u)$ of $u$ and its $j$th bit $Y_j(u)$ are defined analogously. For a point $u \in W_0$ we write *coor*$(u)$ for $(X(u), Y(u))$. We will use the variables $X_i$ and $Y_j$ to generate the $2^d \times 2^d$ grid, which consists of pairs $(n_X, m_Y)$, for $n, m < 2^d$. Consider the following formulas, for $0 \leq \ell \leq 2$,

$$\neg C(X_k \cdot H^\ell,\ (-X_k) \cdot H^\ell), \qquad d \geq k \geq 1, \tag{5.30}$$

$$\neg C(X_j \cdot (-X_k) \cdot H^\ell,\ (-X_j) \cdot H^{\ell \oplus 1}), \qquad d \geq j > k \geq 1, \tag{5.31}$$

$$\neg C((-X_j) \cdot (-X_k) \cdot H^\ell,\ X_j \cdot H^{\ell \oplus 1}), \qquad d \geq j > k \geq 1, \tag{5.32}$$

$$\neg C((-X_k) \cdot X_{k-1} \cdot \ \cdots \ \cdot X_1 \cdot H^\ell,\ \ (-X_k) \cdot H^{\ell \oplus 1}), \qquad d \geq k > 1, \tag{5.33}$$

$$\neg C((-X_k) \cdot X_{k-1} \cdot \ \cdots \ \cdot X_1 \cdot H^\ell,\ \ X_i \cdot H^{\ell \oplus 1}), \qquad d \geq k > i \geq 1, \tag{5.34}$$

$$\neg C(X_d \cdot \ \cdots \ \cdot X_1 \cdot H^\ell,\ H^{\ell \oplus 1}), \tag{5.35}$$

where $\oplus$ denotes addition modulo 3. Suppose that $\mathfrak{A}$ satisfies (5.28)–(5.35). If $u, v \in W_0$ and $zRu$ and $zRv$, for some $z \in W$, then (cf. (5.15) and (5.16))

- $X(v) = X(u)$ whenever $u, v \in (H^\ell)^{\mathfrak{A}}$, for $0 \leq \ell \leq 2$,
- $X(v) = X(u) + 1 < 2^d$ whenever $u \in (H^\ell)^{\mathfrak{A}}$ and $v \in (H^{\ell \oplus 1})^{\mathfrak{A}}$, for $0 \leq \ell \leq 2$,
- $X(u) = X(v) + 1 < 2^d$ whenever $v \in (H^\ell)^{\mathfrak{A}}$ and $u \in (H^{\ell \oplus 1})^{\mathfrak{A}}$, for $0 \leq \ell \leq 2$.

Indeed, if $u, v \in (H^\ell)^{\mathfrak{A}}$ then, by (5.30), $X(u) = X(v)$. If $u \in (H^\ell)^{\mathfrak{A}}$ and $v \in (H^{\ell \oplus 1})^{\mathfrak{A}}$ then, by (5.35), $X(u) < 2^d - 1$. Let $k$ be the minimal number such that $u \in (-X_k)^{\mathfrak{A}}$, i.e., $u \in ((-X_k) \cdot X_{k-1} \cdot \ \cdots \ \cdot X_1 \cdot H^\ell)^{\mathfrak{A}}$. Then, by (5.33), $v \in X_k^{\mathfrak{A}}$ and, by (5.34), $v \in (-X_i)^{\mathfrak{A}}$, for all $i$, $k > i \geq 1$. By (5.31) and (5.32), $v \in (-X_j)^{\mathfrak{A}}$ iff $u \in (X_j)^{\mathfrak{A}}$, for all $d \geq j > k$. It follows that $X(v) = X(u) + 1$. The case of $v \in (H^\ell)^{\mathfrak{A}}$ and $u \in (H^{\ell \oplus 1})^{\mathfrak{A}}$ is similar.

Suppose now that (5.28)–(5.35) with their the $Y$-counterparts ($X_i$ and $H^\ell$ replaced by $Y_i$ and $V^\ell$, respectively) hold in $\mathfrak{A}$. It follows that for every $u, v \in W_0$ with $zRu$ and $zRv$, for some $z \in W$, we have

$$|X(u) - X(v)| \leq 1 \qquad \text{and} \qquad |Y(u) - Y(v)| \leq 1, \tag{5.36}$$


Figure 10: 8-neighbours vs. 4-neighbours.

which means that every point of depth 0 may be surrounded by all of its 8-neighbours (in the sense that they have a common predecessor of depth 1). We are, however, interested only in the 4-neighbours since only the 4-neighbours restrict the tile that can be assigned to the point. Formally, given a pair $(n, m)$ in the $2^d \times 2^d$ grid, denote by $\mathit{4\text{-}nb}(n, m)$ the set that consists of $(n, m)$ and its (at most four) neighbours in the grid, i.e., $(n-1, m)$, $(n+1, m)$, $(n, m-1)$, $(n, m+1)$; see Fig. 10. Let $G$ be a fresh variable, which will represent the points of the grid. Suppose now that in addition the formulas

$$\neg C(G \cdot X_1 \cdot Y_1,\ G \cdot (-X_1) \cdot (-Y_1)), \qquad \neg C(G \cdot (-X_1) \cdot Y_1,\ G \cdot X_1 \cdot (-Y_1)) \tag{5.37}$$

hold in $\mathfrak{A}$. Then, by (5.37), if $u, v \in G^{\mathfrak{A}} \cap W_0$ and $zRu$ and $zRv$, for some $z \in W$, then either $X_1(u) = X_1(v)$ or $Y_1(u) = Y_1(v)$, and thus, by (5.36), $\mathit{coor}(u) \in \mathit{4\text{-}nb}(\mathit{coor}(v))$ and $\mathit{coor}(v) \in \mathit{4\text{-}nb}(\mathit{coor}(u))$.

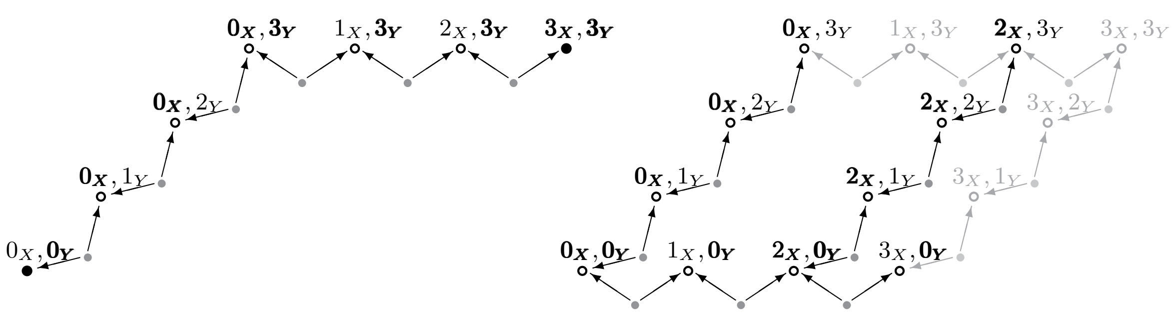
Suppose now that the following formulas are true in $\mathfrak{A}$ as well:

$$G \cdot 0_X \cdot 0_Y \neq \mathbf{0}, \qquad G \cdot (2^d - 1)_X \cdot (2^d - 1)_Y \neq \mathbf{0},$$
$$c(G \cdot (0_X + (2^d - 1)_Y)), \qquad c(G \cdot ((2^d - 1)_X + 0_Y)). \tag{5.38}$$

These constraints guarantee that in the connected set $(G \cdot (0_X + (2^d - 1)_Y))^{\mathfrak{A}}$ there are points $u_{(0,n)}$ and $u_{(n,2^d-1)}$, for $n < 2^d$, such that $\mathit{coor}(u_{(0,n)}) = (0, n)$ and $\mathit{coor}(u_{(n,2^d-1)}) = (n, 2^d - 1)$. Similarly for the connected set $(G \cdot (2^d - 1)_X + 0_Y)^{\mathfrak{A}}$. This gives us the border of the $2^d \times 2^d$ grid we are after (see Fig. 11 $(a)$). And the constraints

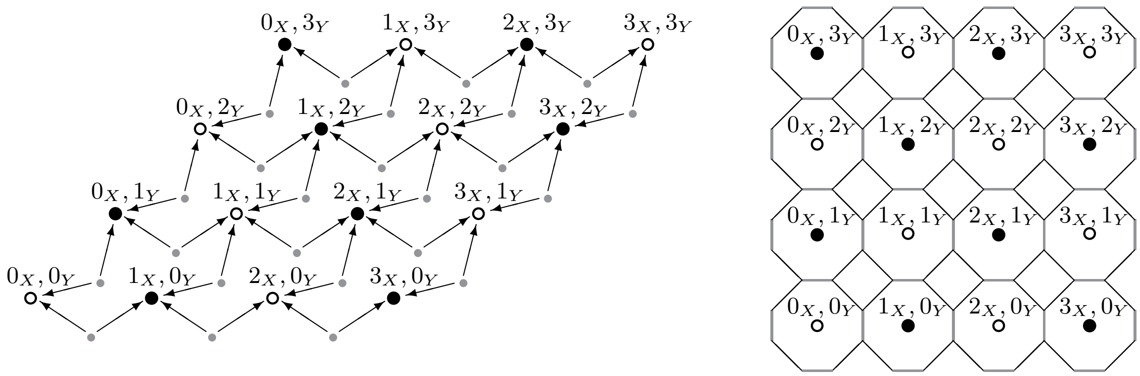
$$c(G\cdot((-X_1)+0_Y)), \qquad c(G\cdot(X_1+0_Y)), \qquad c(G\cdot(0_X+(-Y_1))), \qquad c(G\cdot(0_X+Y_1)) \tag{5.39}$$

ensure that we can find inner points of the grid (see Fig. 11 $(b)$).




$(a)$ $(b)$

Figure 11: Satisfying $(a)$ $c(G \cdot (0_X + (2^d - 1)_Y))$ and $(b)$ $c(G \cdot ((-X_1) + 0_Y))$, for $d = 2$, in a quasi-saw model.




$(a)$ $(b)$

Figure 12: Satisfying $\Theta(\mathcal{T}, T_0, d)$, for $d = 2$, in $(a)$ a quasi-saw model and $(b)$ $\mathbb{R}^2$.

It is to be noted, however, that in general $u \neq v$ even if $coor(u) = coor(v)$. In other words, the constructed points do not necessarily form a proper $2^d \times 2^d$ grid. Let

$$\mathbf{b} = \big(X_1 \cdot (-Y_1)\big) + \big((-X_1) \cdot Y_1\big) \qquad \text{and} \qquad \mathbf{w} = \big((-X_1) \cdot (-Y_1)\big) + \big(X_1 \cdot Y_1\big).$$

Points in $\mathbf{b}^{\mathfrak{A}}$ and $\mathbf{w}^{\mathfrak{A}}$ can be thought of as *black* and *white* squares of a chess board. Observe that if $u, v \in \mathbf{b}^{\mathfrak{A}} \cap W_0$ and $coor(u) \neq coor(v)$ then $u$ and $v$ cannot belong to the same component of $\mathbf{b}^{\mathfrak{A}}$. Thus, there are at least $2^{d-1}$ components in both $\mathbf{b}^{\mathfrak{A}}$ and $\mathbf{w}^{\mathfrak{A}}$. Our next pair of constraints

$$c^{\leq 2^{d-1}}(\mathbf{b}) \qquad \text{and} \qquad c^{\leq 2^{d-1}}(\mathbf{w}) \tag{5.40}$$

says that $\mathbf{b}^{\mathfrak{A}}$ and $\mathbf{w}^{\mathfrak{A}}$ have precisely $2^{d-1}$ components. In particular, if $u, v \in W_0$ belong to the same component of $\mathbf{b}^{\mathfrak{A}}$ then $coor(u) = coor(v)$. This gives a proper $2^d \times 2^d$ grid on which we encode the tiling conditions. The formulas

$$\textstyle\sum_{T_k \in \mathcal{T}} T_k = G, \tag{5.41}$$

$$T_{k_1} \cdot T_{k_2} = \mathbf{0}, \qquad \text{for } T_{k_1}, T_{k_2} \in \mathcal{T}, k_1 \neq k_2, \tag{5.42}$$

say that every point in $G^{\mathfrak{A}} \cap W_0$ is covered by precisely one tile and

$$\neg C(H^{\ell} \cdot V^{\ell'} \cdot T_{k_1},\ H^{\ell} \cdot V^{\ell'} \cdot T_{k_2}), \qquad \text{for } T_{k_1}, T_{k_2} \in \mathcal{T}, k_1 \neq k_2, \tag{5.43}$$

for $0 \leq \ell, \ell' \leq 2$, says that all points in the same component of $(H^{\ell} \cdot V^{\ell'})^{\mathfrak{A}}$ are covered by the same tile. That the colours of adjacent tiles match is ensured by

$$\neg C(H^{\ell} \cdot T_{k_1},\ H^{\ell \oplus 1} \cdot T_{k_2}), \qquad \text{for } T_{k_1}, T_{k_2} \in \mathcal{T} \text{ with } \mathit{right}(T_{k_1}) \neq \mathit{left}(T_{k_2}), \tag{5.44}$$

$$\neg C(V^{\ell} \cdot T_{k_1},\ V^{\ell \oplus 1} \cdot T_{k_2}), \qquad \text{for } T_{k_1}, T_{k_2} \in \mathcal{T} \text{ with } \mathit{top}(T_{k_1}) \neq \mathit{bot}(T_{k_2}), \tag{5.45}$$

for $0 \leq \ell \leq 2$. Finally, we have to say that $(0, 0)$ is covered with $T_0$:

$$0_X \cdot 0_Y \leq T_0. \tag{5.46}$$

One can check that the conjunction $\Theta(\mathcal{T}, T_0, d)$ of these $\mathcal{C}cc$-formulas is as required. For instance, a quasi-saw model $\mathfrak{A}$ satisfying $\Theta(\mathcal{T}, T_0, d)$ (if a tiling exists) is shown in Fig. 12 $(a)$: all points belong to $G^{\mathfrak{A}}$, solid white points to $\mathbf{w}^{\mathfrak{A}}$, solid black to $\mathbf{b}^{\mathfrak{A}}$ and grey to both of the sets. Note that each point of depth 0 has at most 4 predecessors (of depth 1) and the same number of neighbours (points of depth 0 that share a common point of depth 1).

The second statement of the theorem follows from the fact that the model constructed above can be embedded in $\mathrm{RC}(\mathbb{R}^n)$ for all $n \geq 2$: indeed, in $\mathbb{R}^2$, we may represent tiles as octagons (see Fig. 12($b$)), taking $G$ to be the union of all those tiles. Note that the tiles' corners are cut to ensure that every tile has at most 4 neighbours and thus $\mathbf{b}^{\mathfrak{A}}$ and $\mathbf{w}^{\mathfrak{A}}$ have precisely $2^{d-1}$ components. □

As a consequence of Theorems 5.1 and 5.12 we obtain:

**Corollary 5.13.** *Sat*$(\mathcal{C}cc, \textsc{RegC})$ *and Sat*$(\mathcal{C}cc, \textsc{ConRegC})$ *are* NExpTime*-complete.*

We now consider lower complexity bounds for the languages $\mathcal{B}c$ and $\mathcal{B}cc$, which will be obtained by reduction of satisfiability in $\mathcal{C}c$ and $\mathcal{C}cc$, respectively. The basic idea is that two connected closed sets are in contact if and only if their union is connected and they are non-empty; in other words, the formula

$$c(\tau_1) \wedge c(\tau_2) \rightarrow \big(C(\tau_1, \tau_2) \;\leftrightarrow\; c(\tau_1 + \tau_2) \wedge (\tau_1 \neq \mathbf{0}) \wedge (\tau_2 \neq \mathbf{0})\big)$$

is a $\mathcal{C}c$-validity. Thus, if a $\mathcal{C}cc$-formula contains a subformula $C(\tau_1, \tau_2)$, where $\tau_1$ and $\tau_2$ denote *connected* non-empty regions, then we may replace that subformula by $c(\tau_1 + \tau_2)$, thus eliminating an occurrence of $C$. The problem we face, of course, is that this 'reduction' of $C$ to $c$ cannot be directly applied to our formulas in which $\tau_1$ and $\tau_2$ are not necessarily connected and non-empty. The next three lemmas show how to overcome this problem.

We write $\varphi[\psi]$ to indicate that $\varphi$ contains a *positive* occurrence of $\psi$; then $\varphi[\psi/\chi]$ denotes the result of replacing all positive occurrences of $\psi$ in $\varphi$ by $\chi$. Recall that topological spaces are allowed to be empty. There is exactly one frame based on the empty topological space, and exactly one model over that frame. (It is trivial to check whether a formula $\varphi$ is satisfied in this model.) Accordingly, we define:

$$\epsilon_\varphi = \begin{cases} \mathbf{0} = \mathbf{1}, & \text{if } \varphi \text{ is satisfied over the empty space,} \\ \bot, & \text{otherwise.} \end{cases}$$

**Lemma 5.14.** *Let* $\varphi[C(\tau_1, \tau_2)]$ *be a* $\mathcal{C}cc$*-formula, and* $t, t_1, t_2$ *fresh variables. Then* $\varphi$ *is equisatisfiable* (*in* $\textsc{RegC}$ *and* $\textsc{ConRegC}$) *with the formula*

$$\varphi^* = \epsilon_\varphi \vee \Big(\varphi[C(\tau_1, \tau_2)/(t = \mathbf{0})] \;\wedge\; \Big((t = \mathbf{0}) \rightarrow c(t_1 + t_2) \wedge \bigwedge_{i=1,2} \big((t_i \neq \mathbf{0}) \wedge (t_i \leq \tau_i) \wedge c(t_i)\big)\Big)\Big).$$

**Proof.** If $\varphi$ is satisfiable over the empty space, there is nothing to prove; so assume otherwise. Evidently, $\models \bigwedge_{i=1,2}\big((t_i \neq \mathbf{0}) \wedge (t_i \leq \tau_i) \wedge c(t_i)\big) \wedge c(t_1 + t_2) \rightarrow C(\tau_1, \tau_2)$, whence $\models \varphi^* \rightarrow \varphi$. Conversely, let $\mathfrak{A} \models \varphi$, for an Aleksandrov model $\mathfrak{A} = (T, \mathrm{RC}(T), \cdot^{\mathfrak{A}})$; we then construct $\mathfrak{A}^* = (T, \mathrm{RC}(T), \cdot^{\mathfrak{A}^*})$, where $r^{\mathfrak{A}^*} = r^{\mathfrak{A}}$, for all variables occurring in $\varphi$. If $\mathfrak{A} \not\models C(\tau_1, \tau_2)$, then setting $t^{\mathfrak{A}^*} = T$ secures $\mathfrak{A}^* \models \varphi^*$. If, on the other hand, $\mathfrak{A} \models C(\tau_1, \tau_2)$, then there is $z \in \tau_1^{\mathfrak{A}} \cap \tau_2^{\mathfrak{A}}$. As the space is Aleksandrov, there is a *minimal* open subset $V$ of $\mathfrak{A}$ containing $z$. Set $t^{\mathfrak{A}^*} = \emptyset$ and $t_i^{\mathfrak{A}^*} = \tau_i^{\mathfrak{A}} \cap V^-$ ($1 \leq i \leq 2$). It is routine to check that both of the $t_i^{\mathfrak{A}^*}$ are regular closed and connected. Moreover, their sum is connected, since they share the point $z$. This secures $\mathfrak{A}^* \models \varphi^*$. □

We note that the above lemma does not work in any $\mathbb{R}^n$, as intersecting sets do not necessarily share an open subset.

Suppose $T$ is a topological space, and $S$ a regular closed subset of $T$. Then $S$ is itself a topological space (with the subspace topology), which has its own regular closed algebra. The following lemma is tedious to show in full generality, but simple for finite Aleksandrov spaces (which, by Lemma 4.1, is sufficiently general for our purposes). We state it here without proof.

**Lemma 5.15.** *If $S \in \mathrm{RC}(T)$, then $\mathrm{RC}(S) = \{S \cdot X \mid X \in \mathrm{RC}(T)\}$. Furthermore, denoting the Boolean operations in $\mathrm{RC}(S)$ by $+_S$, $\cdot_S$ and $-_S$, etc., we have, for any $X, Y \in \mathrm{RC}(S)$: (i) $X +_S Y = X + Y$, (ii) $X \cdot_S Y = X \cdot Y$, (iii) $-_S X = S \cdot (-X)$, (iv) $\mathbf{1}_S = S$ and $\mathbf{0}_S = \mathbf{0}$; (v) $X$ has the same number of components in $S$ and $T$.*

For a formula $\varphi$ and a variable $s$, define $\varphi_{|s}$ to be the result of replacing every maximal term $\tau$ occurring in $\varphi$ by the term $s \cdot \tau$. For any model $\mathfrak{M} = (T, \mathrm{RC}(T), \cdot^{\mathfrak{M}})$, define $\mathfrak{M}_{|s}$ to be the model over the topological space $s^{\mathfrak{M}}$ (with the subspace topology) obtained by setting $r^{\mathfrak{M}_{|s}} = (s \cdot r)^{\mathfrak{M}}$ for all variables $r$.

**Lemma 5.16.** *For any $\mathcal{C}cc$-formula $\varphi$, $\mathfrak{M} \models \varphi_{|s}$ iff $\mathfrak{M}_{|s} \models \varphi$.*

**Proof.** Using Lemma 5.15 $(i)$–$(iv)$, we can show by structural induction on terms that $(s \cdot \tau)^{\mathfrak{M}} = \tau^{\mathfrak{M}_{|s}}$, for any $\mathcal{B}$-term $\tau$. The result then follows by Lemma 5.15 $(v)$. ☐

**Lemma 5.17.** *Let $\varphi[\neg C(\tau_1, \tau_2)]$ be a $\mathcal{C}cc$-formula, and $s, t, t_1, t_2$ fresh variables. Then $\varphi$ is equisatisfiable in* RegC *with the formula*

$$\varphi^* = \epsilon_\varphi \vee \Big( \ (s \neq \mathbf{0}) \ \wedge \ \varphi[\neg C(\tau_1, \tau_2)/(t = \mathbf{0})]_{|s} \ \wedge \\ \big((t \cdot s = \mathbf{0}) \ \rightarrow \ \neg c(t_1 + t_2) \ \wedge \bigwedge_{i=1,2} \big(c(t_i) \wedge (\tau_i \cdot s \leq t_i)\big)\big)\Big),$$

*and $\varphi$ is equisatisfiable in* ConRegC *with $\varphi^* \wedge c(s)$.*

**Proof.** Again, we may assume that $\varphi$ is not satisfiable over the empty space. Evidently, $\bigwedge_{i=1,2} (c(t_i) \wedge (\tau_i \cdot s \leq t_i)) \wedge \neg c(t_1 + t_2) \rightarrow \neg C(\tau_1 \cdot s, \tau_2 \cdot s)$ is a $\mathcal{C}cc$-validity. So any model $\mathfrak{A}$ of $\varphi^*$ is a model of $\varphi_{|s}$, whence, by Lemma 5.16, $\mathfrak{A}_{|s} \models \varphi$. Note that $\mathfrak{A}_{|s}$ is connected whenever $\mathfrak{A} \models c(s)$. Conversely, suppose $\mathfrak{A} \models \varphi[\neg C(\tau_1, \tau_2)]$, for a quasi-saw model $\mathfrak{A}$ induced by $(W, R)$. Let $W_i$ $(i = 0, 1)$ be the set of points of depth $i$ in $(W, R)$. Without loss of generality, we may assume that every point in $W_0$ has an $R$-predecessor in $W_1$. If $\mathfrak{A} \models C(\tau_1, \tau_2)$, let $\mathfrak{A}^*$ be exactly like $\mathfrak{A}$ except that $s^{\mathfrak{A}^*}$ and $t^{\mathfrak{A}^*}$ are both the whole space. Then $\mathfrak{A}^* \models \varphi^*$. On the other hand, if $\mathfrak{A} \not\models C(\tau_1, \tau_2)$ then, for $i = 1, 2$, we add an extra point $u_i$ to $W$ to connect up the points in $\tau_i^{\mathfrak{A}}$ (if there are any). Formally, let $W^* = W \cup \{u_1, u_2\}$, where $u_1, u_2 \notin W$, and let $R^*$ be the reflexive closure of the union of $R$ and $\{(z, u_i) \mid z \in \tau_i^{\mathfrak{A}} \cap W_1\}$, for $i = 1, 2$. Note that $(W^*, R^*)$ is connected if $(W, R)$ is connected. Clearly, $W$ is a regular closed subset of the topological space $(W^*, R^*)$. Now define the model $\mathfrak{A}^*$ over $(W^*, R^*)$ by setting $s^{\mathfrak{A}^*} = W$, $t^{\mathfrak{A}^*} = \emptyset$, $t_i^{\mathfrak{A}^*} = \tau_i^{\mathfrak{A}} \cup \{u_i\}$ $(i = 1, 2)$, and $r^{\mathfrak{A}^*} = r^{\mathfrak{A}}$ for all other variables $r$. Thus, $\mathfrak{A} = \mathfrak{A}^*_{|s}$, whence, by Lemma 5.16, $\mathfrak{A}^* \models \varphi_{|s}$, and so $\mathfrak{A}^* \models \varphi[\neg C(\tau_1, \tau_2)/t \neq \mathbf{0}]_{|s}$. By construction, $\mathfrak{A}^* \models \bigwedge_{i=1,2} (c(t_i) \wedge (\tau_i \cdot s \leq t_i)) \wedge \neg c(t_1 + t_2)$. Thus, $\mathfrak{A}^* \models \varphi^*$. ☐

As a consequence of Lemma 5.14 and 5.17, we obtain

**Theorem 5.18.** *$Sat(\mathcal{B}c, \textsc{RegC})$ and $Sat(\mathcal{B}c, \textsc{ConRegC})$ are both* ExpTime*-complete. $Sat(\mathcal{B}cc, \textsc{RegC})$ and $Sat(\mathcal{B}cc, \textsc{ConRegC})$ are both* NExpTime*-complete.*

**Proof.** Given a $\mathcal{C}c$-formula, by repeated applications of Lemmas 5.14 and 5.17, we can compute an equisatisfiable $\mathcal{B}c$-formula in polynomial time. Similarly, given a $\mathcal{C}cc$-formula, we can compute an equisatisfiable $\mathcal{B}cc$-formula in polynomial time. It then follows from Theorem 5.9 that $Sat(\mathcal{B}c, \textsc{RegC})$ is ExpTime-hard, and from Theorem 5.12 that $Sat(\mathcal{B}cc, \textsc{RegC})$ is NExpTime-hard. Noting that Lemmas 5.14 and 5.17, and Theorems 5.9 and 5.12 all hold when we restrict attention to connected spaces, we obtain the remaining statements of the theorem. □

We mention here that although Lemma 5.17 does not hold for $\mathbb{R}^n$, a simple modification of the proof of Theorem 5.18 can be used to show the following:

**Theorem 5.19.** *$Sat(\mathcal{B}c, \mathrm{RC}(\mathbb{R}^n))$ is* ExpTime*-hard and $Sat(\mathcal{B}cc, \mathrm{RC}(\mathbb{R}^n))$ is* NExpTime*-hard, for any $n \geq 3$.*

**Proof.** Consider the formula $\Phi'(\chi_M, \psi_{\vec{a}})$ constructed in the proof of Theorem 5.11. It contains only negative occurrences of the predicate $C$. So, we can iteratively apply the transformation of Lemma 5.17 to obtain a $\mathcal{B}c$-formula $\Phi^*$. Denote by $s$ the product of all regions $s$ that have been used to relativize the formula. It follows from the proof of Theorem 5.11 that if $M$ does not accept $\vec{a}$ then $\Phi'(\chi_M, \psi_{\vec{a}})$ is satisfiable in a quasi-saw model, for which we construct a new quasi-saw model $\mathfrak{A}^*$ over $(W^*, R^*)$ as described in the proof of Lemma 5.17. It follows from the proof of Lemma 5.17 that $\mathfrak{A}^* \models \Phi^*$. We then embed the graph of the quasi-saw model $\mathfrak{A}^*$ into $\mathbb{R}^n$ by associating with each point of depth 0 an $n$-dimensional ball with attached cylinders, which touch only if there is a respective point of depth 1. As $\Phi^*$ is a $\mathcal{B}c$-formula and $\mathfrak{A}^* \models \Phi^*$, $\Phi^*$ is satisfied in the constructed model over $\mathrm{RC}(\mathbb{R}^n)$. Conversely, if $\Phi^*$ is satisfiable over $\mathrm{RC}(\mathbb{R}^n)$ then, by Lemmas 4.1 and 4.2, it is satisfied in a quasi-saw model $\mathfrak{A}$. As $\bigwedge_{i=1,2} (c(t_i) \wedge (\tau_i \cdot s \leq t_i)) \wedge \neg c(t_1 + t_2) \to \neg C(\tau_1 \cdot s, \tau_2 \cdot s)$ is a validity, $\mathfrak{A}$ is a model of $\Phi'(\chi_M, \psi_{\vec{a}})_{|s}$, whence, by Lemma 5.16, $\mathfrak{A}_{|s} \models \Phi'(\chi_M, \psi_{\vec{a}})$ and thus, $M$ does not accept $\vec{a}$.

For the second statement of the Theorem, we proceed in a similar fashion, using Lemma 5.17 to remove all (negative) occurrences of $C$ in the formula $\Theta(\mathcal{T}, T_0, d)$ constructed in the proof of Theorem 5.12. □

## 6. Topological logics over Euclidean spaces

So far, we have been mainly concerned with the computational complexity of various topological logics interpreted over the very general classes of frames All, Con, RegC and ConRegC. In this section, we discuss what happens when we restrict consideration to the *specific* topological spaces $\mathbb{R}$, $\mathbb{R}^2$ or $\mathbb{R}^3$.

For languages without connectedness predicates, there is little work to do. For the languages $\mathcal{B}$ and $\mathcal{RCC}$-8, we have, for all $n \geq 1$:

$$Sat(\mathcal{B}, \mathrm{RC}(\mathbb{R}^n)) = Sat(\mathcal{B}, \textsc{RegC}),$$
$$Sat(\mathcal{RCC}\text{-}8, \mathrm{RC}(\mathbb{R}^n)) = Sat(\mathcal{RCC}\text{-}8, \textsc{RegC}).$$

These equations can be established by embedding any model over RegC into one over the domain $\mathrm{RC}(\mathbb{R}^n)$ in such a way that the satisfaction of formulas in the relevant language is unaffected. A suitable embedding for $\mathcal{RCC}$-8 is described in [43]; the case of $\mathcal{B}$ is even more straightforward.

The corresponding equations fail for the languages $\mathcal{C}$, $\mathcal{C}^m$ and $\mathcal{S}4_u$. For example, we saw that the $\mathcal{C}$-formula (4.2) is satisfiable over frames in RegC, but only when the underlying space is disconnected. Thus, $\mathit{Sat}(\mathcal{C}, \mathrm{RC}(\mathbb{R}^n)) \neq \mathit{Sat}(\mathcal{C}, \textsc{RegC})$, whence $\mathit{Sat}(\mathcal{S}4_u, \mathbb{R}^n) \neq \mathit{Sat}(\mathcal{S}4_u, \textsc{All})$, for all $n \geq 1$. However, it turns out that (4.2) is, so to speak, the only fly in the ointment:

**Theorem 6.1** ([51]). *Let $T$ be any connected, dense-in-itself, separable metric space. Then $\mathit{Sat}(\mathcal{S}4_u, T) = \mathit{Sat}(\mathcal{S}4_u, \textsc{Con})$.*

Hence, $\mathit{Sat}(\mathcal{S}4_u, \mathcal{K}) = \mathit{Sat}(\mathcal{S}4_u, \textsc{Con})$ for any class of frames $\mathcal{K}$ included in Con and containing $\mathbb{R}^n$ for any $n$, and similarly, $\mathit{Sat}(\mathcal{C}, \mathcal{K}) = \mathit{Sat}(\mathcal{C}, \textsc{ConRegC})$ for any class of frames $\mathcal{K}$ included in ConRegC and containing $\mathrm{RC}(\mathbb{R}^n)$ for any $n$.

The predicate $c$, however, changes the above picture dramatically.

**Theorem 6.2.** *The problems $\mathit{Sat}(\mathcal{RCC}\text{-}8c, \mathrm{RC}(\mathbb{R}))$, $\mathit{Sat}(\mathcal{RCC}\text{-}8c, \mathrm{RC}(\mathbb{R}^2))$ and $\mathit{Sat}(\mathcal{RCC}\text{-}8c, \mathrm{RC}(\mathbb{R}^3))$ are all different.*

**Proof.** Let $r_i$ ($1 \leq i \leq 5$) and $r_{i,j}$ ($1 \leq i < j \leq 5$) be variables. As we observed in Section 1, the formula

$$\varphi_1 \;=\; \bigwedge_{1\leq i\leq 3} c(r_i) \wedge \bigwedge_{1\leq i<j\leq 3} \mathsf{EC}(r_i, r_j)$$

is not satisfiable over $\mathrm{RC}(\mathbb{R})$, since the non-empty, regular closed, connected subsets of $\mathbb{R}$ are exactly the closed, non-punctual intervals. On the other hand, $\varphi_1$ is visibly satisfiable over $\mathrm{RC}(\mathbb{R}^n)$ for all $n \geq 2$.

Now consider the formula

$$\varphi_2 \;=\; \bigwedge\Big\{\mathsf{DC}(r_{i,j}, r_{k,l}) \;\Big|\; 1 \leq i < j \leq 5, 1 \leq k < l \leq 5, \{i,j\} \cap \{k,l\} = \emptyset\Big\} \wedge \bigwedge\Big\{\mathsf{TPP}(r_i, r_{j,k}) \;\Big|\; 1 \leq i \leq 5, 1 \leq j < k \leq 5, i \in \{j,k\}\Big\} \wedge \bigwedge_{1\leq i<j\leq 5} c(r_{i,j}).$$

Think of the regions (assigned to) $r_1, \ldots, r_5$ as 'vertices' and the regions (assigned to) $r_{i,j}$ ($1 \leq i < j \leq 5$) as 'edges', with $r_{i,j}$ connecting $r_i$ and $r_j$. The literals of $\varphi_2$ having the form $\mathsf{TPP}(r_i, r_{j,k})$ ensure that 'vertices' (are non-empty and) are contained in any 'edges' which involve them. The literals of $\varphi_2$ having the form $\mathsf{DC}(r_{i,j}, r_{k,l})$ ensure that 'edges' lacking a common 'vertex' are not in contact. It is easy to construct a model for this formula over $\mathrm{RC}(\mathbb{R}^n)$ for all $n > 2$. To show that $\varphi_2$ is not satisfiable over $\mathrm{RC}(\mathbb{R}^2)$, suppose otherwise; we show, contrary to fact, that the graph $K_5$ has a plane embedding.

To avoid notational clutter, take $r_i$ ($1 \leq i \leq 5$) and $r_{i,j}$ ($1 \leq i < j \leq 5$) to denote regular closed sets of $\mathbb{R}^2$ satisfying $\varphi_2$. For any point $p$ in any of the $r_{i,j}$, let $\varepsilon'_p$ be the minimum Euclidean distance to any point $q$ in any $r_{k,l}$ such that $\{i,j\} \cap \{k,l\} \neq \emptyset$; let $\varepsilon_p = \max(1, \varepsilon'_p)$; and let $D_p$ be the closed disc centred on $p$ of radius $\varepsilon_p/3$. For all $i,j$ ($1 \leq i < j \leq 5$), let $r'_{i,j} = \bigcup\{D_p \mid p \in r_{i,j}\}$. The following are simple to verify: (i) $r'_{i,j}$ is regular closed; (ii) the open set $(r'_{i,j})^\circ$ is connected (hence path-connected) and contains both $r_i$ and $r_j$; (iii) for all $k$, $l$ ($1 \leq k < l \leq 5$) such that $\{i,j\} \cap \{k,l\} \neq \emptyset$, $r'_{i,j} \cap r'_{k,l} = \emptyset$.

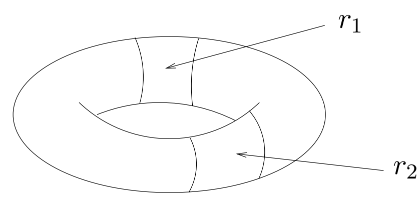



Figure 13: Two non-intersecting, connected, closed sets $r_1$ and $r_2$ on a torus: note that $\overline{r}_1$ and $\overline{r}_2$ are connected, but $\overline{r}_1 \cap \overline{r}_2$ is not.

For all $i$ ($1 \leq i \leq 5$), choose a point $v_i \in r_i$ and, for all $i$, $j$, ($1 \leq i < j \leq 5$), join $v_i$ and $v_j$ with an arc $\alpha_{i,j}$ lying in $r'_{i,j}$. It is straightforward to draw these arcs in such a way that, for each $i$, the arcs with $v_i$ as an endpoint meet only at $v_i$. However, if $\{i,j\} \cap \{k,l\} = \emptyset$, the arcs $\alpha_{i,j}$ and $\alpha_{k,l}$ do not intersect. □

Corresponding remarks apply to $\mathcal{B}$:

**Theorem 6.3.** *The problems* $Sat(\mathcal{B}c, \mathrm{RC}(\mathbb{R}))$, $Sat(\mathcal{B}c, \mathrm{RC}(\mathbb{R}^2))$ *and* $Sat(\mathcal{B}c, \mathrm{RC}(\mathbb{R}^3))$ *are all different.*

**Proof.** Almost identical to Theorem 6.2, noting that, for $r$, $s$ connected, $\mathsf{DC}(r,s)$ if and only if $r+s$ is not connected. □

It is not known whether $Sat(\mathcal{B}c, \mathrm{RC}(\mathbb{R}^3))$ and $Sat(\mathcal{B}c, \textsc{ConRegC})$ are different. However, $Sat(\mathcal{S}4_u, \mathbb{R}^n) \neq Sat(\mathcal{S}4_u, \textsc{Con})$ for all $n \geq 1$. To see this, we rely on the following fact:

**Theorem 6.4** ([38], p. 137)**.** *If* $r_1$ *and* $r_2$ *are non-intersecting closed sets in* $\mathbb{R}^n$*, and points* $x$ *and* $y$ *are connected in* $\overline{r}_1$ *and also in* $\overline{r}_2$*, then* $x$ *and* $y$ *are connected in* $\overline{r}_1 \cap \overline{r}_2$*.*

The formula

$$(r_1 \cap r_2 = \mathbf{0}) \ \wedge \ \bigwedge_{i=1,2} \big((r_i^- \subseteq r_i) \ \wedge \ c(\overline{r}_i)\big) \ \wedge \ \neg c(\overline{r}_1 \cap \overline{r}_2) \tag{6.1}$$

states that $r_1$ and $r_2$ are non-intersecting, closed, connected regions having connected complements, such that the intersection of their complements is not connected. This formula is not satisfiable over any $\mathbb{R}^n$, by Theorem 6.4. However, it is satisfiable over $T$ for many natural, connected topological spaces $T$. For example, let $T$ be a torus, and let $r_1$ and $r_2$ be interpreted as rings in $T$, arranged as in Fig. 13; it is then obvious that $r_1$ and $r_2$ satisfy (6.1).

Thus, for all of our base languages $\mathcal{L}$, the language $\mathcal{L}c$ (and therefore also $\mathcal{L}cc$) is more sensitive than $\mathcal{L}$ to the spatial domain over which it is interpreted. Since we know the complexity of the satisfiability problems for $\mathcal{L}c$ and $\mathcal{L}cc$ for very general classes of spatial domains, the question naturally arises as to the complexity of these problems for spatial domains based on low-dimensional Euclidean spaces.

For $n = 1$, we have a reasonably complete picture.

**Theorem 6.5.** $Sat(\mathcal{B}c, \mathrm{RC}(\mathbb{R}))$ *is in* NP.

**Proof.** *Future-past temporal logic formulas* (FP-formulas, for short) are constructed from propositional variables $p_i$, $i < \omega$, using the Boolean connectives $\wedge$, $\neg$, $\top$, and temporal operators $\Diamond_F$ ('some time in the future') and $\Diamond_P$ ('some time in the past'). A model $(\mathbb{R}, V)$ for FP consists of the real line $\mathbb{R}$ and a valuation $V$ mapping each propositional variable $p_i$ to a subset $V(p_i)$ of $\mathbb{R}$. The truth-relation $\models$ between pointed models $(\mathbb{R}, V, x)$ and FP-formulas $\phi$ is defined as follows (the clauses for the Boolean connectives are standard):

- $(\mathbb{R}, V, x) \models p_i$ iff $x \in V(p_i)$;
- $(\mathbb{R}, V, x) \models \Diamond_F \phi$ iff there is $y > x$ such that $(\mathbb{R}, V, y) \models \phi$, and symmetrically for $\Diamond_P \phi$.

An FP-formula $\phi$ is satisfiable if there exists $(\mathbb{R}, V, x)$ with $(\mathbb{R}, V, x) \models \phi$. Enumerating the region variables as $r_1, r_2, \ldots$, we define a translation $\cdot^*$ from $\mathcal{B}$-formulas to FP-formulas as follows:

$$\begin{aligned}
r_i^* &= p_i, & (-\tau)^* &= \neg\tau^*,\\
(\tau_1 \cdot \tau_2)^* &= \tau_1^* \wedge \tau_2^*, & (\tau_1 + \tau_2)^* &= \tau_1^* \vee \tau_2^*,\\
\mathbf{0}^* &= \bot, & \mathbf{1}^* &= \top,\\
(\tau_1 = \tau_2)^* &= \neg\Diamond_F\Diamond_P\neg(\tau_1^* \leftrightarrow \tau_2^*), & (c(\tau))^* &= \neg\Diamond_F\Diamond_P(\tau^* \wedge \Diamond_F(\neg\tau^* \wedge \Diamond_F\tau^*)),\\
(\neg\varphi)^* &= \neg\varphi^*, & (\varphi_1 \wedge \varphi_2)^* &= \varphi_1^* \wedge \varphi_2^*,\\
(\varphi_1 \vee \varphi_2)^* &= \varphi_1^* \vee \varphi_2^*,
\end{aligned}$$

where $\tau$, $\tau_1$, $\tau_2$ range over $\mathcal{B}$-terms, and $\varphi$, $\varphi_1$, $\varphi_2$ over $\mathcal{B}$-formulas. This translation can clearly be computed in polynomial time. It is routine to verify that a $\mathcal{B}c$-formula $\varphi$ is satisfiable over $\mathrm{RC}(\mathbb{R})$ if and only if $\varphi^*$ is satisfiable over the temporal flow $(\mathbb{R}, <)$. But the satisfiability problem for FP-formulas over $(\mathbb{R}, <)$ is known to be in NP (see, e.g., [20], Theorem 2.4). $\square$

The exact complexities of the problems $Sat(\mathcal{B}cc, \mathrm{RC}(\mathbb{R}))$, $Sat(\mathcal{RCC}\text{-}8c, \mathrm{RC}(\mathbb{R}))$ and $Sat(\mathcal{RCC}\text{-}8cc, \mathrm{RC}(\mathbb{R}))$ are not known, though an upper bound of PSpace is provided by our next theorem.

**Theorem 6.6.** *$Sat(\mathcal{S}4_u cc, \mathbb{R})$ is in* PSpace.

**Proof.** The proof is by reduction to the propositional temporal logic with Since and Until over the real line, for which satisfiability is known to be PSpace-complete [48]. Recall that *linear temporal logic formulas* (LTL-formulas, for short) are constructed from propositional variables $p_i$, $i < \omega$, using the Boolean connectives $\wedge$, $\neg$, $\top$, and binary temporal operators $\mathcal{S}$ ('since') and $\mathcal{U}$ ('until'). A model $(\mathbb{R}, V)$ for LTL consists of the real line $\mathbb{R}$ and a valuation $V$ mapping each propositional variable $p_i$ to a subset $V(p_i)$ of $\mathbb{R}$. The truth-relation $\models$ between pointed models $(\mathbb{R}, V, x)$ and LTL-formulas $\phi$ is defined as follows (the clauses for the Boolean connectives are standard):

- $(\mathbb{R}, V, x) \models p_i$ iff $x \in V(p_i)$;
- $(\mathbb{R}, V, x) \models \phi\mathcal{S}\psi$ iff there exists $y < x$ such that $(\mathbb{R}, V, y) \models \psi$ and for all $z$ with $y < z < x$, $(\mathbb{R}, V, z) \models \phi$;
- $(\mathbb{R}, V, x) \models \phi\mathcal{U}\psi$ iff there exists $y > x$ such that $(\mathbb{R}, V, y) \models \psi$ and for all $z$ with $x < z < y$, $(\mathbb{R}, V, z) \models \phi$.

An LTL-formula $\phi$ is satisfiable if there exists $(\mathbb{R}, V, x)$ with $(\mathbb{R}, V, x) \models \phi$.

Suppose now that we are given an $\mathcal{S}4_u cc$-formula $\varphi$. Without loss of generality, we may assume that no nested interior and closure operators occur in $\varphi$ (by introducing additional variables, we can always transform $\varphi$ into an equisatisfiable formula of length $2|\varphi|$ without nested topological operators). Let $K$ be the maximum number $k$ such that $c^{\leq k}(\tau)$ occurs in $\varphi$.

Our PSpace algorithm for checking the satisfiability of such a $\varphi$ proceeds as follows. To simplify the exposition, we introduce a new unary predicate $\infty(\cdot)$, with the following semantics (for models $\mathfrak{M}$ over $\mathbb{R}$):

$\mathfrak{M} \models \infty(\tau)$ iff some bounded interval includes infinitely many components of $\tau^{\mathfrak{M}}$.

(Note that $\infty$ is not part of the language $\mathcal{S}4_u cc$.) Now, we first guess a conjunction $\chi$ of literals of the forms $\tau_1 = \tau_2$, $\tau_1 \neq \tau_2$, $c^{=k}(\tau)$ and $c^{\geq K+1}(\tau)$ such that (i) for each subformula of $\varphi$ of the form $\tau_1 = \tau_2$, $\chi$ contains either $\tau_1 = \tau_2$ or $\tau_1 \neq \tau_2$, and (ii) for each subformula of the form $c^{\leq k}(\tau)$, it contains one of $c^{=0}(\tau), \ldots, c^{=K}(\tau), c^{\geq K+1}(\tau)$, where $c^{=k}(\tau)$ abbreviates $c^{\leq k}(\tau) \wedge c^{\geq k}(\tau)$. The literals in $\chi$ give in an obvious way the truth-values to the atoms in $\varphi$, and we check whether $\varphi$, as a propositional formula, is true under this valuation. If this is indeed the case, it only remains to verify whether $\chi$ is satisfiable over $\mathbb{R}$. This can be done by expressing $\chi$ as an equisatisfiable LTL-formula $\chi^*$.

To this end, we non-deterministically replace each literal of the form $c^{\geq K+1}(\tau)$ in $\chi$ with *either* $\infty(\tau)$ *or* the conjunction

$$(\tau = r_1 \cup r_2) \ \wedge \ (r_1^- \cap r_2^- \cap \tau = \mathbf{0}) \ \wedge \ c^{=K}(r_1) \ \wedge \ (r_2 \neq \mathbf{0}),$$

where $r_1$ and $r_2$ are fresh variables. It is not difficult to see that $\chi$ is satisfiable if and only if some conjunction $\hat{\chi}$ obtained in this way is also satisfiable; moreover, the length of $\hat{\chi}$ is linear in the length of $\chi$. Note that $\hat{\chi}$ is not necessarily an $\mathcal{S}4_u cc$-formula (because it may contain occurrences of $\infty$); however, all terms in $\hat{\chi}$ are $\mathcal{S}4_u$-terms.

Now, the meaning of $\mathcal{S}4_u$-terms over $\mathbb{R}$ can be expressed with the help of LTL-formulas: for an $\mathcal{S}4_u$-term $\tau$ we construct the LTL-formula $\tau^*$ according to the following definition:

$$\begin{array}{lll} r_i^* = p_i, & \mathbf{0}^* = \bot, & \mathbf{1}^* = \top, \\ \overline{\tau}^* = \neg\tau^*, & (\tau_1 \cap \tau_2)^* = \tau_1^* \wedge \tau_2^*, & (\tau_1 \cup \tau_2)^* = \tau_1^* \vee \tau_2^*, \\ (\tau^\circ)^* = (\tau^* \, \mathcal{S} \, \top) \wedge \tau^* \wedge (\tau^* \, \mathcal{U} \, \top) & & \end{array}$$

and $\tau^-$ is treated as an abbreviation for $\overline{\overline{\tau}^\circ}$. As $\tau$ does not have nested interior and closure operators, the length of $\tau^*$ is linear in the length of $\tau$. Let $\mathfrak{M}$ be a topological model over $\mathbb{R}$. Set $V_{\mathfrak{M}}(p_i) = p_i^{\mathfrak{M}}$. It can be shown by induction on the structure of an $\mathcal{S}4_u$-term $\tau$ that $(\mathbb{R}, V_{\mathfrak{M}}, x) \models \tau^*$ if and only if $x \in \tau^{\mathfrak{M}}$.

We now map each literal $\psi$ of $\hat{\chi}$ into an LTL-formula $\psi^*$ with (intuitively) the same meaning as $\psi$. First, we translate literals of the form $(\tau_1 = \tau_2)$ or $(\tau_1 \neq \tau_2)$ into LTL-formulas as follows:

$$\begin{array}{rcl} (\tau_1 = \tau_2)^* & = & \Box_F \Box_P (\tau_1^* \leftrightarrow \tau_2^*), \\ (\tau_1 \neq \tau_2)^* & = & \Diamond_F \Diamond_P \neg(\tau_1^* \leftrightarrow \tau_2^*), \end{array}$$

where $\Diamond_F \psi = \top \, \mathcal{U} \, \psi$, $\Box_F = \neg\Diamond_F\neg\psi$, $\Diamond_P \psi = \top \, \mathcal{S} \, \psi$, $\Box_P \psi = \neg\Diamond_P\neg\psi$.

Next, we translate literals of the form $\infty(\tau)$ into LTL-formulas as follows:

$$(\infty(\tau))^* = \Diamond_P \Diamond_F \left( (\neg(\tau^* \, \mathcal{U} \, \tau^*) \wedge \neg(\neg\tau^* \, \mathcal{U} \, \neg\tau^*)) \vee (\neg(\tau^* \, \mathcal{S} \, \tau^*) \wedge \neg(\neg\tau^* \, \mathcal{S} \, \neg\tau^*)) \right). \quad (6.2)$$

We claim that the formula on the right-hand side of (6.2) is true (at all points in $\mathbb{R}$) if and only if there exists a bounded interval of $\mathbb{R}$ which includes infinitely many components of $\tau$. For if $x \in \mathbb{R}$ is such that $(\mathbb{R}, V, x) \models \neg(\tau^* \mathcal{U} \tau^*) \wedge \neg(\neg\tau^* \mathcal{U} \neg\tau^*)$, under some valuation $V$, then every interval to the right of $x$ (i.e., whose left-hand endpoint is $x$) contains infinitely many components of $\tau$; and similarly for the second disjunct. Conversely, if there exists a bounded interval of $\mathbb{R}$ which includes infinitely many components of $\tau$, then these components must have some accumulation point $x$. But then either every interval to the right of $x$ contains infinitely many components of $\tau$, or every interval to the left of $x$ does (or both); in the former case, $x$ satisfies $\neg(\tau^* \,\mathcal{U}\, \tau^*) \wedge \neg(\neg\tau^* \,\mathcal{U}\, \neg\tau^*)$, in the latter case, the second disjunct in (6.2) is satisfied.

It is much harder to translate atoms of the form $c^{=k}(\tau)$. To do this, we will require LTL-formulas of the form:

$$\beta_\psi(\eta,\xi) \quad = \quad (\psi\wedge\eta)\,\mathcal{U}\,\big(((\psi\wedge\eta)\wedge((\neg\psi\wedge\xi)\,\mathcal{U}\,\top)) \;\vee\; (\neg\psi\wedge\xi)\big).$$

It is readily seen that if $\beta_\psi(\eta,\xi)$ is true at a point $x$ then there is $y > x$ such that either (i) $\psi\wedge\eta$ is true everywhere in $(x,y)$[4] and $\neg\psi\wedge\xi$ is true at $y$, or (ii) $\psi\wedge\eta$ is true everywhere in $(x,y]$ and $\neg\psi\wedge\xi$ is true in $(y,z]$, for some $z > y$. In other words, if $\beta_\psi(\eta,\xi)\wedge\psi$ is true at $x$ then $\eta$ is true at all points from $(x,\infty)$ that belong to the same connected component of $\psi$ as $x$, while $\xi$ is true immediately to the right of that connected component.

To count the connected components of (extensions of) terms $\tau$, we construct LTL-formulas $\theta^\tau_k$, where $k$ is a natural number not exceeding $K$. Take $\lfloor\log_2 K\rfloor+1$ fresh variables $v^\tau_n,\dots,v^\tau_1$ to represent a connected component number in binary. The formula $\theta^\tau_k$ contains the following conjuncts:

$$\begin{aligned}
&\tau^*\wedge v^\tau_i \;\to\; \big(\neg\tau^*\,\mathcal{U}\,\top \;\vee\; \beta_{\tau^*}(v^\tau_i,\top) \;\vee\; \Box_F(\tau^*\wedge v^\tau_i)\big), \qquad \text{for } n\geq i\geq 1,\\
&\tau^*\wedge \neg v^\tau_i \;\to\; \big(\neg\tau^*\,\mathcal{U}\,\top \;\vee\; \beta_{\tau^*}(\neg v^\tau_i,\top) \;\vee\; \Box_F(\tau^*\wedge \neg v^\tau_i)\big), \qquad \text{for } n\geq i\geq 1,\\
&\tau^*\wedge v^\tau_j\wedge\neg v^\tau_h \;\to\; \textsf{next-int}_{\tau^*}(v^\tau_j) \;\vee\; \Box_F\neg\tau^*, \qquad \text{for } n\geq j>h\geq 1,\\
&\tau^*\wedge \neg v^\tau_j\wedge\neg v^\tau_h \;\to\; \textsf{next-int}_{\tau^*}(\neg v^\tau_j) \;\vee\; \Box_F\neg\tau^*, \qquad \text{for } n\geq j>h\geq 1,\\
&\tau^*\wedge\neg v^\tau_h\wedge v^\tau_{h-1}\wedge\cdots\wedge v^\tau_1 \;\to\; \textsf{next-int}_{\tau^*}(v^\tau_h) \;\vee\; \Box_F\neg\tau^*, \qquad \text{for } n\geq h\geq 1,\\
&\tau^*\wedge\neg v^\tau_h\wedge v^\tau_{h-1}\wedge\cdots\wedge v^\tau_1 \;\to\; \textsf{next-int}_{\tau^*}(\neg v^\tau_i) \;\vee\; \Box_F\neg\tau^*, \qquad \text{for } n\geq h>i\geq 1,
\end{aligned}$$

where

$$\textsf{next-int}_{\tau^*}(\eta) \quad = \quad \beta_{\neg\tau^*}(\top,\eta) \;\vee\; \big(\tau^*\,\mathcal{U}\,\beta_{\neg\tau^*}(\top,\eta)\big).$$

It can be seen that if $\textsf{next-int}_{\tau^*}(\eta)\wedge\tau^*$ is true at $x$ then $\eta$ is true at some $y > x$ that belongs to the next connected component of $\tau$ to the right of $x$. The first two formulas ensure that inside the current connected component of $\tau$, the bits of the counter remain constant. The remaining conjuncts ensure proper counting; cf. (5.31)–(5.34). So, we set

$$\begin{aligned}
(c(\tau)^{=k})^* \quad = \quad \Diamond_F\Diamond_P\big(\mathbf{0}_\tau\wedge\tau^*\wedge((\tau^*\,\mathcal{S}\,\neg\tau^*)\vee\Box_P\neg\tau^*\vee\Box_P\tau^*) \;\;\wedge\;\; \theta^\tau_k\wedge\Box_F\theta^\tau_k\\
\wedge\;\; \Diamond_F(\mathbf{k}_\tau\wedge\tau^*\wedge((\tau^*\,\mathcal{U}\,\neg\tau^*)\vee\Box_F\neg\tau^*\vee\Box_F\tau^*)\big),
\end{aligned}$$

where $\mathbf{k}_\tau$ is the binary representation of $k$ using the counter variables $v^\tau_n,\dots,v^\tau_1$ (e.g., $\mathbf{0}_\tau = \neg v^\tau_n\wedge\cdots\wedge\neg v^\tau_1$). Clearly, the length of $(c(\tau)^{=k})^*$ is polynomial in the length of $\tau$ and $\log_2 K$.

Finally, we construct the LTL-formula $\chi^*$ by replacing each conjunct $\psi$ in $\hat\chi$ with $\psi^*$. Clearly, the length of $\chi^*$ is polynomial in the length of $\varphi$. We leave it to the reader to verify that $\chi^*$ is satisfiable if and only if $\hat\chi$ is satisfiable over $\mathbb{R}$. Since the satisfiability problem for LTL over $\mathbb{R}$ is in PSPACE [48], this completes the proof. □

[4]As usual, $(x,y] = \{z\in\mathbb{R}\mid x<z\leq y\}$ and $(x,y)=\{z\in\mathbb{R}\mid x<z<y\}$.

**Corollary 6.7.** *The problems $Sat(\mathcal{C}c, \mathrm{RC}(\mathbb{R}))$, $Sat(\mathcal{C}cc, \mathrm{RC}(\mathbb{R}))$, $Sat(\mathcal{S}4_u c, \mathbb{R})$ and $Sat(\mathcal{S}4_u cc, \mathbb{R})$ are all* PSpace*-complete; the problems $Sat(\mathcal{B}cc, \mathrm{RC}(\mathbb{R}))$, $Sat(\mathcal{RCC}\text{-}8c, \mathrm{RC}(\mathbb{R}))$ and $Sat(\mathcal{RCC}\text{-}8cc, \mathrm{RC}(\mathbb{R}))$ are all* (NP*-hard and) in* PSpace.

**Proof.** Theorems 5.6 and 6.6. □

This concludes our discussion of the complexity of satisfiability for topological logics interpreted over $\mathbb{R}$.

Over $\mathbb{R}^2$, topological logics become harder to analyze. The encodings used to obtain lower complexity bounds in Section 5 apply unproblematically to Euclidean spaces of dimension at least 2. In particular, Theorem 5.9 states that $Sat(\mathcal{C}c, \mathrm{RC}(\mathbb{R}^n))$ is ExpTime-hard, for all $n \geq 2$, whence $Sat(\mathcal{S}4_u c, \mathbb{R}^n)$ is ExpTime-hard, for all $n \geq 2$. Similarly, Theorem 5.12 states that $Sat(\mathcal{C}cc, \mathrm{RC}(\mathbb{R}^n))$ is NExpTime-hard, for all $n \geq 2$, whence $Sat(\mathcal{S}4_u cc, \mathbb{R}^n)$ is NExpTime-hard, for all $n \geq 2$.

Upper bounds for $Sat(\mathcal{C}c, \mathrm{RC}(\mathbb{R}^n))$, $Sat(\mathcal{S}4_u c, \mathbb{R}^n)$, $Sat(\mathcal{C}cc, \mathrm{RC}(\mathbb{R}^n))$ or $Sat(\mathcal{S}4_u cc, \mathbb{R}^n)$, where $n \geq 2$, are not known. However, for the smaller language $\mathcal{RCC}$-8, upper bounds are known from the literature in the case where the spatial domain is limited to certain well-behaved regions in $\mathbb{R}^2$. One such domain is the set $\mathbf{D}$ of closed disc-homeomorphs in $\mathbb{R}^2$, with which we began this paper. We mention the following remarkable fact, the proof of which is too involved to repeat here:

**Theorem 6.8** ([50]). *The problem $Sat(\mathcal{RCC}\text{-}8, (\mathbb{R}^2, \mathbf{D}))$ is* NP*-complete.*

Finally, we remark that, if $n \geq 3$, no upper complexity bound is currently known for the problem $Sat(\mathcal{B}c, \mathrm{RC}(\mathbb{R}^n))$, or, therefore, for any more expressive spatial logic.

## 7. Conclusion

In this paper, we have investigated the effect of augmenting various topological logics in the qualitative spatial reasoning literature with predicates able to express the property of connectedness. We considered three principal *base languages*: $\mathcal{B}$, the language of the variety of Boolean algebras; $\mathcal{C}$, the extension of the well-known language $\mathcal{RCC}$-8 with region-combining operations; and $\mathcal{S}4_u$, the extension of Lewis' system $\mathcal{S}4$ with a universal operator, under the topological interpretation of McKinsey and Tarski. And we considered two kinds of connectedness predicate: $c(r)$, for 'region $r$ is connected'; and $c^{\leq k}(r)$, for 'region $r$ has at most $k$ connected components.' For each base language $\mathcal{L}$, we defined the languages $\mathcal{L}c$ (by adding the predicate $c$) and $\mathcal{L}cc$ (by adding the predicates $c^{\leq k}(r)$ for $k \geq 1$); and we considered the complexity of the satisfiability problems for $\mathcal{L}$, $\mathcal{L}c$ and $\mathcal{L}cc$ over various natural (classes of) spatial domains, both very general—as in the case of RegC, ConRegC, All and Con—and also very specific—as in the case of $\mathrm{RC}(\mathbb{R}^n)$ and $\mathbb{R}^n$ for various $n$.

We showed that, whereas the base languages display a surprising indifference to the frames over which they are interpreted, the corresponding languages with connectedness predicates are highly sensitive in this regard. We also showed that the addition of connectedness predicates increases the complexity of satisfiability over general classes of frames—typically from NP or PSpace (for the base language $\mathcal{L}$) to ExpTime (for the corresponding language $\mathcal{L}c$) and NExpTime (for the language $\mathcal{L}cc$). We observed that this increase in complexity is 'stable': over the most general classes of frames, the extensions of such different formalisms as $\mathcal{B}$ and $\mathcal{S}4_u$ with connectedness predicates are of the same complexity. We further observed that by restricting these languages to formulas with just one connectedness

| | RegC | ConRegC | RC($\mathbb{R}^n$) $n \geq 3$ | RC($\mathbb{R}^2$) | RC($\mathbb{R}$) |
|---|---|---|---|---|---|
| $\mathcal{RCC}$-8 | NP Cor. 4.4 | | | | |
| $\mathcal{RCC}$-8$c$ | NP Cor. 4.4 | | | $\geq$NP | $\leq$PSpace,$\geq$NP |
| $\mathcal{RCC}$-8$cc$ | NP Cor. 4.4 | | | $\geq$NP | $\leq$PSpace,$\geq$NP |
| $\mathcal{B}$ | NP Cor. 4.4 | | | | |
| $\mathcal{B}c$ | ExpTime Thm. 5.18 | ExpTime Thm. 5.18 | $\geq$ExpTime Thm. 5.18 | $\geq$NP | NP Thm. 6.5 |
| $\mathcal{B}cc$ | NExpTime Thm. 5.18 | NExpTime Thm. 5.18 | $\geq$NExpTime Thm. 5.18 | $\geq$NP | $\leq$PSpace,$\geq$NP |
| $\mathcal{C}$ | NP Cor. 4.4 | PSpace Cor. 5.8 | | | |
| $\mathcal{C}c$ | ExpTime Cor. 5.10 | ExpTime Cor. 5.10 | $\geq$ExpTime Thm. 5.9 | $\geq$ExpTime Thm. 5.9 | PSpace |
| $\mathcal{C}cc$ | NExpTime Cor. 5.13 | NExpTime Cor. 5.13 | $\geq$NExpTime Thm. 5.12 | $\geq$NExpTime Thm. 5.12 | PSpace |
| $\mathcal{C}^m$ | NP Cor. 4.4 | PSpace | | | |
| $\mathcal{C}^m c$ | ExpTime | ExpTime | $\geq$ExpTime | $\geq$ExpTime | PSpace |
| $\mathcal{C}^m cc$ | NExpTime | NExpTime | $\geq$NExpTime | $\geq$NExpTime | PSpace |

| | All | Con | $\mathbb{R}^n$, $n \geq 3$ | $\mathbb{R}^2$ | $\mathbb{R}$ |
|---|---|---|---|---|---|
| $\mathcal{S}4_u$ | PSpace [34, 39, 2] | PSpace Cor. 5.5 | | | |
| $\mathcal{S}4_u c$ | ExpTime Thm. 5.3 | ExpTime Thm. 5.3 | $\geq$ExpTime | $\geq$ExpTime | PSpace |
| $\mathcal{S}4_u cc$ | NExpTime Thm. 5.1 | NExpTime Thm. 5.1 | $\geq$NExpTime | $\geq$NExpTime | PSpace Thm. 6.6 |

Table 1: Satisfiability complexity for the topological logics considered in this paper.

constraint of the form $c(r)$, we obtain logics that are still in PSpace, while two such constraints lead to ExpTime-hardness. Finally, we turned our attention to the complexity of the satisfiability problems for these languages when interpreted over Euclidean spaces, summarizing what is currently known and stating several open problems. The results obtained are summarized in Table 6.


## Acknowledgements

The work on this paper was partially supported by the U.K. EPSRC research grants EP/E034942/1 and EP/E035248/1.


## References


[1] M. Aiello and J. van Benthem. A modal walk through space. *J. of Applied Non-Classical Logics*, 12(3–4):319–364, 2002.

[2] C. Areces, P. Blackburn, and M. Marx. The computational complexity of hybrid temporal logics. *Logic J. of the IGPL*, 8:653–679, 2000.

[3] Ph. Balbiani, T. Tinchev, and D. Vakarelov. Modal logics for region-based theories of space. *Fundamenta Informaticae*, 81:29–82, 2007.

[4] M. Ben-Or, D. Kozen, and J. Reif. The complexity of elementary algebra and geometry. *J. of Computer and System Sciences*, 32:251–264, 1986.

[5] B. Bennett. Spatial reasoning with propositional logic. In *Proceedings of KR*, pages 51–62. Morgan Kaufmann, 1994.

[6] S. Borgo, N. Guarino, and C. Masolo. A pointless theory of space based on strong connection and congruence. In L. Aiello, J. Doyle, and S. Shapiro, editors, *Proceedings of KR*, pages 220–229. Morgan Kaufmann, 1996.

[7] N. Bourbaki. *General Topology, Part 1*. Hermann, Paris & Addison-Wesley, 1966.

[8] D. Cantone and V. Cutello. Decision algorithms for elementary topology I. Topological syllogistics with set and map constructs, connectedness and cardinailty composition. *Comm. on Pure and Appl. Mathematics*, XLVII:1197–1217, 1994.

[9] A. Chagrov and M. Zakharyaschev. *Modal Logic*, volume 35 of *Oxford Logic Guides*. Clarendon Press, Oxford, 1997.

[10] A. Cohn and J. Renz. Qualitative spatial representation and reasoning. In F. van Hermelen, V. Lifschitz, and B. Porter, editors, *Handbook of Knowledge Representation*, pages 551–596. Elsevier, 2008.

[11] E. Davis. The expressivity of quantifying over regions. *J. of Logic and Computation*, 16:891–916, 2006.

[12] G. Dimov and D. Vakarelov. Contact algebras and region-based theory of space: a proximity approach, I. *Fundamenta Informaticae*, 74(2–3):209–249, 2006.

[13] G. Dimov and D. Vakarelov. Contact algebras and region-based theory of space: a proximity approach, II. *Fundamenta Informaticae*, 74(2–3):251–282, 2006.

[14] C. Dornheim. Undecidability of plane polygonal mereotopology. In A. Cohn, L. Schubert, and S. Shapiro, editors, *Proceedings of KR*, pages 342–353. Morgan Kaufmann, 1998.

[15] I. Düntsch and E. Orlowska. A proof system for contact relation algebras. *J. of Philosophical Logic*, 29:241–262, 2000.

[16] I. Düntsch and M. Winter. A representation theorem for Boolean contact algebras. *Theoretical Computer Science*, 347:498–512, 2005.

[17] M. Egenhofer and R. Franzosa. Point-set topological spatial relations. *International J. of Geographical Information Systems*, 5:161–174, 1991.

[18] M. Fischer and M. Rabin. Super-exponential complexity of Presburger arithmetic. *Complexity of Computation, SIAM-AMS Proceedings*, 7:27–41, 1974.

[19] M. Fitting. *Proof Methods for Modal and Intuitionistic Logics*. D. Reidel, Dordrecht, 1983.

[20] D. Gabbay, A. Kurucz, F. Wolter, and M. Zakharyaschev. *Many-Dimensional Modal Logics: Theory and Applications*, volume 148 of *Studies in Logic*. Elsevier, 2003.

[21] D. Gabelaia, R. Kontchakov, A. Kurucz, F. Wolter, and M. Zakharyaschev. Combining spatial and temporal logics: expressiveness vs. complexity. *J. of Artificial Intelligence Research (JAIR)*, 23:167–243, 2005.

[22] A. Galton. *Qualitative Spatial Change*. Oxford University Press, Oxford, 2000.

[23] G. De Giacomo. *Decidability of Class-Based Knowledge Representation Formalisms*. PhD thesis, Università degli Studi di Roma 'La Sapienza', 1995.

[24] V. Goranko and S. Passy. Using the universal modality: gains and questions. *J. of Logic and Computation*, 2:5–30, 1992.

[25] R. Goré. Tableau methods for modal and temporal logics. In M. D'Agostino, D. M. Gabbay, R. Hähnle, and J. Posegga, editors, *Handbook of Tableau Methods*, pages 297–396. Kluwer Academic Publishers, 1999.

[26] A. Griffiths. *Computational Properties of Spatial Logics in the Real Plane*. PhD thesis, School of Computer Science, University of Manchester, 2008.

[27] A. Grzegorczyk. Undecidability of some topological theories. *Fundamenta Mathematicae*, 38:137–152, 1951.

[28] A. Grzegorczyk. The systems of Leśniewski in relation to contemporary logical research. *Studia Logica*, 3(1):77–95, 1955.

[29] D. Hilbert. *Grundlagen der Geometrie*. B.G. Teubner, Leipzig and Berlin, 1909.

[30] G. Hughes and M. Cresswell. *A New Introduction to Modal Logic*. Routledge, 1996.

[31] S. Koppelberg. *Handbook of Boolean Algebras*, volume 1. North-Holland, 1989.

[32] D. Kozen. *Theory of Computation*. Springer, 2006.

[33] S. Kripke. Semantical analysis of modal logic I: Normal modal propositional calculi. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 9:67–96, 1963.

[34] R. Ladner. The computational complexity of provability in systems of modal logic. *SIAM J. on Computing*, 6:467–480, 1977.

[35] S. Li and M. Ying. Extensionality of the RCC8 composition table. *Fundamenta Informaticae*, 55(3–4):363–385, 2003.

[36] C. Lutz and F. Wolter. Modal logics of topological relations. *Logical Methods in Computer Science*, 2, 2006.

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

[38] M.H.A. Newman. *Elements of the Topology of Plane Sets of Points*. Cambridge, 1964.

[39] W. Nutt. On the translation of qualitative spatial reasoning problems into modal logics. In W. Burgard, T. Christaller, and A. Cremers, editors, *Proceedings of KI*, volume 1701 of *LNCS*, pages 113–124. Springer, 1999.

[40] I. Pratt-Hartmann. A topological constraint language with component counting. *J. of Applied Non-Classical Logics*, 12:441–467, 2002.

[41] I. Pratt-Hartmann. First-order mereotopology. In M. Aiello, I. Pratt-Hartmann, and J. van Benthem, editors, *Handbook of Spatial Logics*, pages 13–97. Springer, 2007.

[42] D. Randell, Z. Cui, and A. Cohn. A spatial logic based on regions and connection. In B. Nebel, C. Rich, and W. Swartout, editors, *Proceedings of KR*, pages 165–176. Morgan Kaufmann, 1992.

[43] J. Renz. A canonical model of the region connection calculus. In A. Cohn, L. Schubert, and S. Shapiro, editors, *Proceedings of KR*, pages 330–341. Morgan Kaufmann, 1998.

[44] J. Renz. Maximal tractable fragments of the region connection calculus: a complete analysis. In *Proceedings of IJCAI*, pages 448–454. Morgan Kaufmann, 1999.

[45] J. Renz and B. Nebel. On the complexity of qualitative spatial reasoning: A maximal tractable fragment of the region connection calculus. In *Proceedings of IJCAI*, pages 522–527. Morgan Kaufmann, 1997.

[46] J. Renz and B. Nebel. On the complexity of qualitative spatial reasoning. *Artificial Intelligence*, 108:69–123, 1999.

[47] J. Renz and B. Nebel. Qualitative spatial reasoning using constraint calculi. In M. Aiello, I. Pratt-Hartmann, and J. van Benthem, editors, *Handbook of Spatial Logics*, pages 161–216. Springer, 2007.

[48] M. Reynolds. The complexity of the temporal logic over the reals. *Annals of Pure and Applied Logic*, 161:1063–1096, 2010.

[49] P. Roeper. Region-based topology. *J. of Philosophical Logic*, 26:251–309, 1997.

[50] M. Schaefer, E. Sedgwick, and D. Štefankovič. Recognizing string graphs in NP. *J. of Computer and System Sciences*, 67:365–380, 2003.

[51] V. Shehtman. "Everywhere" and "Here". *J. of Applied Non-Classical Logics*, 9(2–3), 1999.

[52] E. Spaan. *Complexity of Modal Logics*. PhD thesis, Universiteit van Amsterdam, 1993.

[53] A. Tarski. Zur Grundlegung der Booleschen Algebra. *Fundamenta Mathematicae*, 24:177–198, 1935.

[54] J. van Benthem and G. Bezhanishvili. Modal logics of space. In M. Aiello, I. Pratt-Hartmann, and J. van Benthem, editors, *Handbook of Spatial Logics*, pages 217–298. Springer, 2007.

[55] P. van Emde Boas. The convenience of tilings. In A. Sorbi, editor, *Complexity, Logic and Recursion Theory*, volume 187 of *Lecture Notes in Pure and Applied Mathematics*, pages 331–363. Marcel Dekker Inc., 1997.

[56] A. N. Whitehead. *Process and Reality*. New York: The MacMillan Company, 1929.

[57] F. Wolter and M. Zakharyaschev. Spatial reasoning in RCC-8 with Boolean region terms. In W. Horn, editor, *Proceedings of ECAI*, pages 244–248. IOS Press, 2000.

[58] F. Wolter and M. Zakharyaschev. Qualitative spatio-temporal representation and reasoning: a computational perspective. In G. Lakemeyer and B. Nebel, editors, *Exploring AI in the New Millenium*, pages 175–216. Morgan Kaufmann, 2002.