Topological logics over Euclidean spaces

R. Kontchakov[1], I. Pratt-Hartmann[2] and M. Zakharyaschev[1]
[1]School of Computer Science and Information Systems, Birkbeck College, London, U.K.
[2]School of Computer Science, Manchester University, Manchester, U.K.

Abstract

In this paper we prove some results on the computational complexity of standard quantifier-free spatial logics with the connectedness predicate interpreted over the Euclidean spaces $\mathbb{R}$ and $\mathbb{R}^2$.

Topological logics with connectedness. A *topological logic* is a formal language whose variables range over subsets of topological spaces, and whose non-logical primitives denote fixed topological properties and operations involving these subsets. For example, let the function symbols $\cap$, $\cup$ and $\cdot^-$ denote the operations of intersection, union and topological closure, respectively; let the constant $\mathbf{0}$ denote the empty set; let the unary predicate c denote the property of connectedness; and let the binary predicate $\subseteq$ denote the subset relation. Then the formula

$$c(r_1) \ \wedge \ c(r_2) \ \wedge \neg(r_1 \cap r_2 \subseteq \mathbf{0}) \ \rightarrow \ c(r_1 \cup r_2) \tag{1}$$

states that the union of two intersecting connected sets r_1 and r_2 is connected; likewise, the formula

$$c(r_1) \ \wedge \ (r_1 \subseteq r_2) \ \wedge \ (r_2 \subseteq r_1{}^-) \ \rightarrow \ c(r_2) \tag{2}$$

states that, if r_1 is a connected set, and r_2 is sandwiched between r_1 and its closure, then r_2 is also connected. It is well known that these statements hold for any subsets r_1, r_2 of any topological space. As we might put it: formulas (1) and (2) are validities of the topological logic in question. Once the syntax of that logic has been made precise, it is natural to ask: what is the computational complexity of identifying such validities?

Formally, let F be a set of function symbols with fixed interpretations as operations on subsets of a topological space; and let P be a set of predicates, again with fixed interpretations as relations between subsets of a topological space. We denote by $\mathcal{L}(F, P)$ the set of quantifier-free first-order formulas over the signature (F, P). Using the obvious abbreviations, we may regard formulas (1) and (2) as belonging to the language $\mathcal{S}4_u c := \mathcal{L}(\{\cup, \cap, \cdot^-, \overline{\cdot}\}, \{c, =\})$, where the operator $\overline{\cdot}$ denotes complementation with respect to the containing space, and $=$ denotes the equality relation. An *interpretation* $\mathfrak{I}$ for this language consists of a topological space T and a map $r \mapsto r^{\mathfrak{I}}$ taking every variable to a subset of T. This map is extended to terms in the obvious way, and truth-values are assigned to atomic formulas according to the rules: $\mathfrak{I} \models \tau_1 = \tau_2$ iff $\tau_1^{\mathfrak{I}} = \tau_2^{\mathfrak{I}}$ and $\mathfrak{I} \models c(\tau)$ iff $\tau^{\mathfrak{I}}$ is connected. A formula φ is *satisfiable* if there exists an interpretation $\mathfrak{I}$ such that $\mathfrak{I} \models \varphi$, and *valid* if $\mathfrak{I} \models \varphi$, for all interpretations $\mathfrak{I}$. The properties of validity and satisfiability are thus dual in the usual sense.

If $\mathcal{L}$ is a topological language and $\mathcal{K}$ a class of interpretations, we write $\mathrm{Sat}(\mathcal{L}, \mathcal{K})$ to denote the problem of determining whether a given $\mathcal{L}$-formula is satisfied by some interpretation in $\mathcal{K}$. It is shown in [2] that $\mathrm{Sat}(\mathcal{S}4_u c, \textsc{All})$ is ExpTime-complete, where All denotes the class of all interpretations. Removing the connectedness predicates altogether, we obtain the language $\mathcal{S}4_u := \mathcal{L}(\{\cup, \cap, \cdot^-, \overline{\cdot}\}, \{=\})$—in essence the modal logic S4 with an additional universal modality, under the topological semantics of McKinsey and Tarski [4]. It is well-known that $\mathrm{Sat}(\mathcal{S}4_u, \textsc{All})$ is PSpace-complete.

By restricting the language in various ways, we obtain less expressive logics, having—in general—less complex satisfiability problems. A subset of a topological space is *regular closed* if it is the closure of an open set. The collection of regular closed sets of a topological space T forms a Boolean algebra under inclusion with top element $\mathbf{1} = T$ and bottom element $\mathbf{0} = \emptyset$. We denote this Boolean algebra by $\mathrm{RC}(T)$, and use the symbols $+$, $\cdot$ and $-$ to denote the obvious operations in $\mathrm{RC}(T)$; in addition, we write $\tau_1 \leq \tau_2$ in preference to $\tau_1 \subseteq \tau_2$. Now consider the language $\mathcal{B}c := \mathcal{L}(\{\cdot, +, -, \mathbf{0}, \mathbf{1}\}, \{c, =\})$. Let RegC be the class of interpretations over any domain of the form $\mathrm{RC}(T)$. Any such interpretation $\mathfrak{I}$ may then be extended to $\mathcal{B}c$-terms, and hence to $\mathcal{B}c$-formulas in the expected way. It turns out that, e.g., $c(r_1) \wedge c(r_2) \wedge (r_1 \cdot r_2 \neq \mathbf{0}) \rightarrow c(r_1 + r_2)$ is a validity of $\mathcal{B}c$ over RegC (cf. formula (1)). Since the property of being a regular closed set can be expressed in $\mathcal{S}4_u c$, we may regard $\mathcal{B}c$, interpreted over RegC, as a sub-language of $\mathcal{S}4_u c$, interpreted over All. Nevertheless, it is shown in [2] that $\mathrm{Sat}(\mathcal{B}c, \textsc{RegC})$ is still ExpTime-complete. For a language having intermediate expressive power between $\mathcal{B}c$ and $\mathcal{S}4_u c$, consider $\mathcal{C}c := \mathcal{L}(\{\cdot, +, -, \mathbf{0}, \mathbf{1}\}, \{C, c, =\})$, where the binary 'contact' predicate C is given the semantics: $\mathfrak{I} \models C(\tau_1, \tau_2)$ iff $\tau_1^{\mathfrak{I}} \cap \tau_2^{\mathfrak{I}} \neq \emptyset$. It follows from the above results that $\mathrm{Sat}(\mathcal{C}c, \textsc{RegC})$ is ExpTime-complete. On the other hand, by removing the predicate c from $\mathcal{C}c$, we obtain the language $\mathcal{C} = \mathcal{L}(\{\cdot, +, -, \mathbf{0}, \mathbf{1}\}, \{C, =\})$, whose satisfiability problem over RegC is shown in [7] to be NP-complete, and PSpace-complete if we restrict attention to connected spaces.

The language $\mathcal{C}$ is of particular interest, because it enables us to express the so-called 'RCC-8' relations—DC (disconnection), EC (external connection), PO (partial overlap), EQ (equality), TPP (tangential proper part) and NTPP (non-tangential proper part)—popularized in the seminal treatments of spatial logics by Egenhofer and Franzosa [1] and Randell *et al.* [5]. These relations are illustrated, for closed disc homeomorphs in $\mathbb{R}^2$, in Fig. 1. The RCC-8 predicates are of most

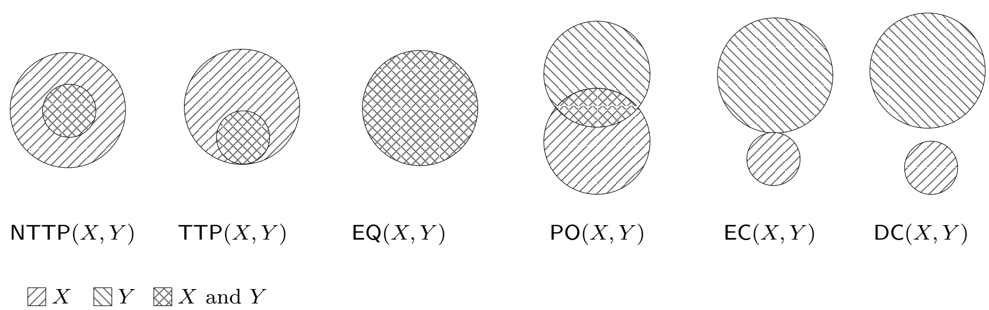

Figure 1: The RCC-8 relations illustrated for disc-homeomorphs in $\mathbb{R}^2$.

interest when interpreted over certain *geometrical* spaces: it is to these that we now turn.

Topological logics over $\mathbb{R}$ and $\mathbb{R}^2$. Consider the language $\mathcal{RCC}$-8c, whose only topological primitives are the RCC-8 predicates (illustrated above) together with the predicate c (connectedness); and consider the $\mathcal{RCC}$-8c-formula

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

Formula (3) 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 easy to see that this formula is satisfiable over $\mathrm{RC}(\mathbb{R}^2)$, but unsatisfiable over $\mathrm{RC}(\mathbb{R})$. Using simple facts about non-planar graphs, it is likewise easy to write an $\mathcal{RCC}$-8c-formula satisfiable over $\mathrm{RC}(\mathbb{R}^3)$, but not over $\mathrm{RC}(\mathbb{R}^2)$. Thus the satisfiability problems for $\mathcal{RCC}$-8c-formulas over these spaces are all different. More intriguingly, consider the $\mathcal{RCC}$-8c-formula

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

stating that r_1 is connected, and that any two of $r_1, \ldots, r_4$ touch at their boundaries without overlapping. Formula (4) *is* satisfiable over the regular closed subsets of $\mathbb{R}$, as shown in Fig. 2.

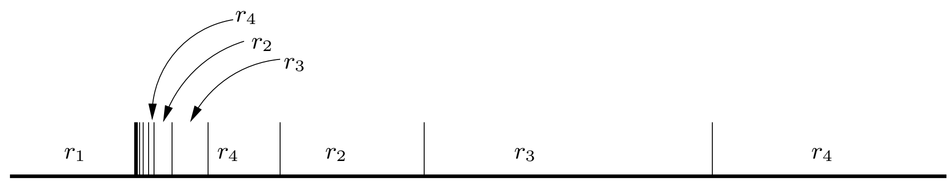

Figure 2: A configuration of regular closed regions in $\mathbb{R}$ satisfying the $\mathcal{RCC}$-8c-formula (4).

However, such an arrangement is only possible provided that at least two of the regions r_2, r_3 and r_4 have infinitely many components. If, for example, we chose to consider only interpretations over (say) the regular closed *semi-linear* (=*semi-algebraic*) subsets of $\mathbb{R}$, then this formula would count as unsatisfiable. Moral: simple topological logics featuring the connectedness predicate are sensitive to the underlying topological space, and indeed to the precise choice of subsets of that space over which its variables range.

Denote by $\mathrm{RC}(\mathbb{R}^n)$ the class of all interpretations over the regular closed subsets of $\mathbb{R}^n$, and by $\mathrm{RCP}(\mathbb{R}^n)$ the class of all interpretations over the regular closed, semi-linear subsets of $\mathbb{R}^n$. (The 'P' in RCP stands for 'polyhedron.') Although the example of formula (4) shows that $\mathrm{Sat}(\mathcal{RCC}\text{-}8c, \mathrm{RC}(\mathbb{R})) \neq \mathrm{Sat}(\mathcal{RCC}\text{-}8c, \mathrm{RCP}(\mathbb{R}))$, we nevertheless establish:

Theorem 1. *The problems* $Sat(\mathcal{B}c, \mathrm{RC}(\mathbb{R}))$ *and* $Sat(\mathcal{B}c, \mathrm{RCP}(\mathbb{R}))$ *are identical, and are* NP-*complete.*

Denote by $\mathcal{P}(\mathbb{R})$ the set of all interpretations over the power set of $\mathbb{R}$, and by $\mathcal{S}(\mathbb{R})$ the set of all interpretations over the set of semi-linear subsets of $\mathbb{R}$. It was shown in [3] that $\mathrm{Sat}(\mathcal{C}c, \mathrm{RC}(\mathbb{R}))$ and $\mathrm{Sat}(\mathcal{S}4_u c, \mathcal{P}(\mathbb{R}))$ are both PSPACE-complete. We present the following additional complexity results:

Theorem 2. *The problem* $Sat(\mathcal{RCC}\text{-}8c, \mathrm{RCP}(\mathbb{R}))$ *is* NP-*complete. The problems* $Sat(\mathcal{C}c, \mathrm{RCP}(\mathbb{R}))$ *and* $Sat(\mathcal{S}4_u c, \mathcal{S}(\mathbb{R}))$ *are both* PSPACE-*complete.*

Turning our attention from the real line to the Euclidean plane, we have:

$$\mathrm{Sat}(\mathcal{RCC}\text{-}8c, \mathrm{RC}(\mathbb{R}^2)) = \mathrm{Sat}(\mathcal{RCC}\text{-}8c, \mathrm{RCP}(\mathbb{R}^2)).$$

Thus, the easy separation result for the space $\mathbb{R}$ obtained using (4) no longer holds in $\mathbb{R}^2$. However, making the language just a little bit more expressive than $\mathcal{RCC}$-8 causes this equivalence to break down. Specifically, we have

$$\mathrm{Sat}(\mathcal{B}c, \mathrm{RC}(\mathbb{R}^2)) \neq \mathrm{Sat}(\mathcal{B}c, \mathrm{RCP}(\mathbb{R}^2)).$$

We remark that the simplest $\mathcal{B}c$-formula known to the authors that is satisfied by an interpretation in $\mathrm{RC}(\mathbb{R}^2)$, but not by one in $\mathrm{RCP}(\mathbb{R}^2)$, is much more complicated than formula (4), separating $\mathrm{Sat}(\mathcal{RCC}\text{-}8c, \mathrm{RC}(\mathbb{R}))$ from $\mathrm{Sat}(\mathcal{RCC}\text{-}8c, \mathrm{RCP}(\mathbb{R}))$. It is shown in [6] that $\mathrm{Sat}(\mathcal{RCC}\text{-}8, \mathcal{D}(\mathbb{R}^2))$ is NP-complete, where $\mathcal{D}(\mathbb{R}^2)$ is the domain of closed disc-homeomorphs in $\mathbb{R}^2$. A routine extension of this proof can be used to show

Theorem 3. *The problems* $Sat(\mathcal{RCC}\text{-}8c, \mathrm{RC}(\mathbb{R}^2))$ *and* $Sat(\mathcal{RCC}\text{-}8c, \mathrm{RCP}(\mathbb{R}^2))$ *coincide, and are* NP-*complete.*

The complexity—even the decidability—of $\mathcal{B}c$ and $\mathcal{C}c$ over $\mathrm{RC}(\mathbb{R}^2)$ and $\mathrm{RCP}(\mathbb{R}^2)$ remains open. However, it is possible to establish the following lower bounds:

Theorem 4. *The satisfiability problems for* $\mathcal{B}c$ *over* $\mathrm{RC}(\mathbb{R}^2)$ *and* $\mathrm{RCP}(\mathbb{R}^2)$ *are* PSPACE-*hard. The satisfiability problems for* $\mathcal{C}c$ *over* $\mathrm{RC}(\mathbb{R}^2)$ *and* $\mathrm{RCP}(\mathbb{R}^2)$ *are* EXPTIME-*hard.*

language	$\mathbb{R}$		$\mathbb{R}^2$	
	RC($\mathbb{R}$)	RCP($\mathbb{R}$)	RC($\mathbb{R}^2$)	RCP($\mathbb{R}^2$)
$\mathcal{RCC}$-8c	?	NP Thm. 2	NP Thm. 3	
$\mathcal{B}c$	NP Thm. 1		$\geq$PSPACE Thm. 4	$\geq$PSPACE Thm. 4
$\mathcal{C}c$	PSPACE [3]	PSPACE Thm. 2	$\geq$EXPTIME Thm. 4	$\geq$EXPTIME Thm. 4
	$\mathcal{P}(\mathbb{R})$	$\mathcal{S}(\mathbb{R})$	$\mathcal{P}(\mathbb{R}^2)$	$\mathcal{S}(\mathbb{R}^2)$
$\mathcal{S}4_u c$	PSPACE [3]	PSPACE Thm. 2	$\geq$EXPTIME Thm. 4	$\geq$EXPTIME Thm. 4

Table 1: Summary of results.

The known complexity results for the logics $\mathcal{RCC}$-8c, $\mathcal{B}c$, $\mathcal{C}c$ and $\mathcal{S}4_u c$ interpreted over $\mathbb{R}$ and $\mathbb{R}^2$ are summarized in Table 1.

We conjecture that $\mathcal{RCC}$-8c is still NP-complete over RC($\mathbb{R}$).

Acknowledgement. The results presented in this paper have been obtained in the framework of the UK EPSRC projects EP/E035248 and EP/E034942: 'Computational logic of Euclidean spaces.' Other directions of research on this project are (i) interpretations over $\mathbb{R}^3$ and higher-dimensional Euclidean spaces, (ii) topological logics with component counting (which usually increases complexity from EXPTIME to NEXPTIME), and first-order theories of topological signatures.

References

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

[2] R. Kontchakov, I. Pratt-Hartmann, F. Wolter, and M. Zakharyaschev. On the computational complexity of spatial logics with connectedness constraints. In I. Cervesato, H. Veith, and A. Voronkov, editors, *Proceedings of LPAR*, volume 5330 of *Lecture Notes in Computer Science*, pages 574–589. Springer, 2008.

[3] R. Kontchakov, I. Pratt-Hartmann, F. Wolter, and M. Zakharyaschev. Topology, connectedness, and modal logic. In C. Areces and R. Goldblatt, editors, *Advances in Modal Logic*, volume 7, pages 151–176. College Publications, London, 2008.

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

[5] 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.

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

[7] 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.