
Reasoning About Space: The Modal Way

MARCO AIELLO, *Department of Information and Communication Technologies, University of Trento, Via Sommarive 14, 38050 Trento, Italy.*
E-mail: aiellom@ieee.org

JOHAN VAN BENTHEM, *Institute for Logic, Language and Computation, University of Amsterdam, Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands.*
E-mail: johan@science.uva.nl

GURAM BEZHANISHVILI, *Department of Mathematical Sciences, New Mexico State University, Las Cruces, NM 88003-0001, USA.*
E-mail: gbezhani@nmsu.edu

Abstract

We investigate the topological interpretation of modal logic in modern terms, using a new notion of bisimulation. We look at modal logics with interesting topological content, presenting, among others, a new proof of McKinsey and Tarski's theorem on completeness of $\mathbf{S4}$ with respect to the real line, and a completeness proof for the logic of finite unions of convex sets of reals. We conclude with a broader picture of extended modal languages of space, for which the main logical questions are still wide open.

Keywords: Spatial reasoning, modal logic, topological interpretation, topo-bisimulation, serial set.

1 Introduction and purpose

The topological interpretation is one of the oldest semantics for modal languages. Reading the modal box as the interior operator, one can show that the modal logic $\mathbf{S4}$ is complete with respect to arbitrary topological spaces. But there are classical results with much more mathematical content, such as McKinsey and Tarski's beautiful theorem that $\mathbf{S4}$ is also the complete logic of the reals, and indeed of any metric separable space without isolated points [20]. In the 1970s, Albert Dragalin started a programme of investigating modal logics of geometrical structures in Euclidean spaces, as reported in [26] containing Shehtman's results on logics of domains of the real plane (similar results were obtained independently by Goldblatt [17]). At the same time, Esakia started the study of modal diamonds interpreted as the topological derivative operator, resulting in an axiomatization of the base logic (later published in [15]), as well as that of scattered spaces [14]; in the late 1980s, Shehtman gave a similar axiomatization of Euclidean spaces [27] solving a problem posed in [20]. Also, much of the work on modal 'neighbourhood semantics' has a topological flavour, though the spatial aspect was never prominent. Overall, it seems fair to say that the topological interpretation has always remained something of a side-show in modal logic and intuitionistic logic, often tucked away in notes and appendices.

The 1990s marked extensive research in spatial logics from the AI point of view. An extension of Allen's approach to qualitative reasoning for time intervals [3] to spatial regions

[23] has been the seed for a number of modal studies of space in AI. The spatial version of the interval calculus, named the region connection calculus RCC, is a first-order theory with a topological predicate of connection. Since RCC is undecidable, Bennett [6] proposed a study of tractable fragments through an encoding into a multi-dimensional modal logic interpreted in topological spaces. Subsequently, Renz and Nebel have shown this encoding to be NP-complete [25]. Combinations of the modal encoding of RCC with various temporal theories have been proposed, from a computational viewpoint, in [32, 33]. Lemon and Pratt [19] provide an appealing analysis of spatial models and the design of complete logics for them.

The purpose of this paper is to establish a programme of the modal analysis of space – showing how one can get more generality, as well as some new questions. In particular, this paper contains (a) a modern analysis of the modal language as a topological formalism in terms of ‘topo-bisimulation’ (continuing [1]), (b) a new proof of McKinsey and Tarski’s Theorem (inspired by [21]), (c) an analysis of special topological logics on the reals, pointing towards a landscape of spatial logics above **S4**, and finally (d) an extension to richer modal languages of space, and their increased expressive power.

2 Modal language and topological semantics

2.1 Language and axioms

Let us first set the scene where we will operate. The basic language \mathcal{L} of propositional modal logic is composed of

- a countable set \mathcal{P} of propositional variables,
- Boolean connectives $\neg, \vee, \wedge, \rightarrow$,
- modal operators \Box, \Diamond .

The standard axiomatization of our central logic **S4** is

$$\begin{aligned} \Box(\varphi \rightarrow \psi) &\rightarrow (\Box\varphi \rightarrow \Box\psi) && \text{(K)} \\ \Box\varphi &\rightarrow \Box\Box\varphi && \text{(4)} \\ \Box\varphi &\rightarrow \varphi && \text{(T)} \end{aligned}$$

Modus Ponens and Necessitation are the only rules of inference:

$$\frac{\varphi \rightarrow \psi \quad \varphi}{\psi} \quad \text{MP} \qquad \frac{\varphi}{\Box\varphi} \quad \text{N}$$

For a closer fit to topological reasoning, however, it is better to work with an equivalent axiomatization of **S4**:

$$\begin{aligned} \Box\top &&& \text{(N)} \\ (\Box\varphi \wedge \Box\psi) &\leftrightarrow \Box(\varphi \wedge \psi) && \text{(R)} \\ \Box\varphi &\rightarrow \varphi && \text{(T)} \\ \Box\varphi &\rightarrow \Box\Box\varphi && \text{(4)} \end{aligned}$$

Modus Ponens and Monotonicity are the only rules of inference:

$$\frac{\varphi \rightarrow \psi \quad \varphi}{\psi} \quad \text{MP} \qquad \frac{\varphi \rightarrow \psi}{\Box\varphi \rightarrow \Box\psi} \quad \text{M}$$

As we shall see in a moment, these principles are valid in all topological spaces when we let formulas range over sets of points, with the Booleans as the obvious set operations, modal box as interior and dually, modal diamond as closure. On top of this base set, further modal axioms can be used to express more special topological properties. For example, an additional ‘axiom’ $\Box\varphi \leftrightarrow \varphi$ would say that each set is open, meaning that the spaces satisfying it have the discrete topology.

2.2 Topological completeness

The first semantic completeness proof for **S4** did not use the by now dominant Kripke models, which go back to Kanger, Hintikka, and Kripke in the 1950s. It was actually an earlier *spatial* completeness argument of [20], in terms of the following notions. Recall that a topological space is a pair $\langle W, \tau \rangle$, where W is a nonempty set and τ is a collection of subsets of W satisfying the following three conditions:

- $\emptyset, W \in \tau$;
- if $U, V \in \tau$, then $U \cap V \in \tau$;
- if $\{U_i\}_{i \in I} \in \tau$, then $\bigcup_{i \in I} U_i \in \tau$.

Let Int and $\bar{}$ denote the interior and closure operators of $\langle W, \tau \rangle$, respectively. It is well known (cf. [13]) that these satisfy the following clauses for all $X, Y \subseteq W$:

$$\begin{array}{ll} Int(W) = W & \bar{\emptyset} = \emptyset \\ Int(X \cap Y) = Int(X) \cap Int(Y) & \overline{X \cup Y} = \bar{X} \cup \bar{Y} \\ Int(X) \subseteq X & X \subseteq \bar{X} \\ Int(X) \subseteq Int(Int(X)) & \overline{\bar{X}} \subseteq \bar{X}. \end{array}$$

Moreover, there is a duality $Int(X) = W - \overline{W - X}$, and a topological space can also be defined in terms of an interior operator or a closure operator satisfying the above four clauses.

McKinsey and Tarski defined a valuation ν of formulas of \mathcal{L} into $\langle W, \tau \rangle$ by putting

- $\nu(p) \subseteq W$,
- $\nu(\neg\varphi) = W - \nu(\varphi)$,
- $\nu(\varphi \vee \psi) = \nu(\varphi) \cup \nu(\psi)$,
- $\nu(\varphi \wedge \psi) = \nu(\varphi) \cap \nu(\psi)$,
- $\nu(\varphi \rightarrow \psi) = (W - \nu(\varphi)) \cup \nu(\psi)$,
- $\nu(\Box\varphi) = Int(\nu(\varphi))$,
- $\nu(\Diamond\varphi) = \overline{\nu(\varphi)}$.

In definitions and arguments in this paper, we will often economize, and leave out the clauses for disjunction, implication and modal diamond, as these are automatic from the others. Now, call a triple $M = \langle W, \tau, \nu \rangle$ a *topological model*. A formula φ is said to be *true* in such a model M if $\nu(\varphi) = W$, and we say that φ is *topologically valid* if it is true in every topological model. Referring to the second axiomatization of **S4**, which highlights the interior operator, one easily sees its soundness:

If **S4** $\vdash \varphi$, then φ is topologically valid.

McKinsey and Tarski's pioneering achievement [20] was a proof of completeness (for a new proof, see Section 3 below):

If φ is topologically valid, then $\mathbf{S4} \vdash \varphi$.

Hence, for a modal logician, the topological semantics is adequate for $\mathbf{S4}$, or the other side of the same coin, for a topologist, $\mathbf{S4}$ axiomatizes the algebra of the interior operator completely.

2.3 The semantics amplified

In the rest of the paper we use a modern format for modal semantics. Given a topological model $M = \langle W, \tau, \nu \rangle$, we define when a formula φ is *true at a point* w by induction on the length of φ :

- $w \models p$ iff $w \in \nu(p)$;
- $w \models \neg\varphi$ iff not $w \models \varphi$;
- $w \models \varphi \wedge \psi$ iff $w \models \varphi$ and $w \models \psi$;
- $w \models \Box\varphi$ iff $(\exists U \in \tau)(w \in U \text{ and } (\forall v \in U)(v \models \varphi))$;

and hence, also

- $w \models \Diamond\varphi$ iff $(\forall U \in \tau)(w \in U \Rightarrow (\exists v \in U)(v \models \varphi))$.

This semantics for the modal language looks different from the usual one, where models have a binary accessibility relation between points, and $w \models \Box\varphi$ if φ is true in all relational successors of w . Nevertheless, there are strong analogies. All basic notions of Kripke models make sense for the topological semantics, too. Here are two typical examples that we shall need further on.

The topological semantics is *local* in that the truth value of a formula at a point w only depends on what happens inside the open neighbourhoods of that point. More precisely, consider any topological model M with a point w inside, which lies in some open set U . Now define the obvious *restriction* of M to a topological model $M|U$ by taking U for the new universe, letting the open sets be all the old open sets inside U , and putting $\nu'(P) = \nu(P) \cap U$. It is easy to show by induction on formulas that

$$w \models_M \varphi \text{ iff } w \models_{M|U} \varphi.$$

Thus, to determine truth values for modal formulas at a point w on the real line we only need to know how the model behaves in arbitrarily small open neighbourhoods around w . Or conversely, we can change the model at a distance from a point w , without affecting the original truth values.

Our second illustration concerns the proper *semantic invariance* for our modal language. The connection between M and $M|U$ is a special case of a more general model relation investigated at length in [1], including versions in terms of Ehrenfeucht–Fraïssé games.

DEFINITION 2.1 (Topo-bisimulation)

Suppose that two topological models $\langle W, \tau, \nu \rangle$ and $\langle W', \tau', \nu' \rangle$ are given. A *topo-bisimulation* is a nonempty relation $T \subseteq W \times W'$ such that if xTx' then

- (base): $x \in \nu(p)$ iff $x' \in \nu'(p)$ for any $p \in \mathcal{P}$;
- (forth condition): if $x \in U \in \tau$ then
 $(\exists U' \in \tau')(x' \in U' \ \& \ (\forall y' \in U')(\exists y \in U)(yTy'))$;
- (back condition): if $x' \in U' \in \tau'$ then
 $(\exists U \in \tau)(x \in U \ \& \ (\forall y \in U)(\exists y' \in U')(yTy'))$.

As an example, the identity relation on U is a topo-bisimulation between the above models M and $M|U$. This also shows that the preceding definition does not require *totality*: some points need not have links at all. But much rougher ‘contractions’ and ‘twists’ are also possible. For example, a continuous and open map between topological spaces gives rise to a topo-bisimulation. We recall that a map f between two topological spaces $\langle W_1, \tau_1 \rangle$ and $\langle W_2, \tau_2 \rangle$ is called *continuous* if $f^{-1}(V) \in \tau_1$ for every $V \in \tau_2$; f is called *open* if $f(U) \in \tau_2$ for every $U \in \tau_1$.

FACT 2.2

Let f be a continuous and open map between two topological spaces $\langle W_1, \tau_1 \rangle$ and $\langle W_2, \tau_2 \rangle$.

1. Suppose ν_2 is a valuation on $\langle W_2, \tau_2 \rangle$. Define a valuation ν_1 on $\langle W_1, \tau_1 \rangle$ by setting $\nu_1(p) = f^{-1}(\nu_2(p))$ for any $p \in \mathcal{P}$. Then f is a topo-bisimulation between the topological models $\langle W_1, \tau_1, \nu_1 \rangle$ and $\langle W_2, \tau_2, \nu_2 \rangle$.
2. Suppose ν_1 is a valuation on $\langle W_1, \tau_1 \rangle$ such that $w \in \nu_1(p)$ iff $v \in \nu_1(p)$ for any $p \in \mathcal{P}$ and $w, v \in W_1$ with $f(w) = f(v)$. Define a valuation ν_2 on $\langle W_2, \tau_2 \rangle$ by setting $\nu_2(p) = f(\nu_1(p))$ for any $p \in \mathcal{P}$. Then f is a topo-bisimulation between the topological models $\langle W_1, \tau_1, \nu_1 \rangle$ and $\langle W_2, \tau_2, \nu_2 \rangle$.

As follows from the above, topo-bisimulation is a coarse notion of similarity between topological spaces, much less fine-grained than homeomorphism or homotopy. But it is just right for the expressive power of the modal language.

FACT 2.3

If T is a topo-bisimulation between two models M, N such that wTv , then w, v satisfy the same modal formulas.

The statement of true versions of converse results is a much more delicate matter (cf. [8] and [11]). Here we give just one simple illustration.

FACT 2.4

If two worlds w and v satisfy the same modal formulas in two *finite* models M and N , then there exists a topo-bisimulation between these models which connects w with v .

2.4 Topological spaces and Kripke models

We briefly describe the well-known connection between topological and Kripke semantics for modal logic. The latter is a particular case of the former. To see this, we recall that an **S4**-frame (henceforth ‘frame’, for short) is a pair $\langle W, R \rangle$, where W is a nonempty set and R is a quasi-order (transitive and reflexive) on W . Call a set $X \subseteq W$ *upward closed* if $w \in X$ and wRv imply $v \in X$.

FACT 2.5

Every frame $\langle W, R \rangle$ gives rise to a topological space $\langle W, \tau_R \rangle$, where τ_R is the set of all upward closed subsets of $\langle W, R \rangle$.

It is easy to check that τ_R is a topology on W . Indeed, τ_R is a rather special topology on W ; for any family $\{X_i\}_{i \in I} \subseteq \tau_R$, we have $\bigcap_{i \in I} X_i \in \tau_R$. Such spaces are called *Alexandroff spaces*, in which every point has a least neighbourhood. In frames, the least neighbourhood of a point w is evidently $\{v \in W : wRv\}$, which is usually denoted by $R[w]$.

Conversely, every topological space $\langle W, \tau \rangle$ naturally induces a quasi-order R_τ defined by putting $wR_\tau v$ iff $w \in \overline{\{v\}}$. This is called the *specialization order* in the topological literature. Again it is easy to check that R_τ is transitive and reflexive, and that every open set of τ is R_τ -upward closed. Moreover, R_τ is anti-symmetric iff $\langle W, \tau \rangle$ satisfies the T_0 separation axiom (that is, any two different points are separated by an open set). Hence, R_τ is a partial order iff $\langle W, \tau \rangle$ is a T_0 -space.

Now it is easy to check that $R = R_{\tau_R}$ and $\tau \subseteq \tau_{R_\tau}$. Moreover, the following four conditions are equivalent: (i) $\tau = \tau_{R_\tau}$, (ii) every R_τ -upward closed set belongs to τ , (iii) every point of W has a least neighbourhood in $\langle W, \tau \rangle$, (iv) $\langle W, \tau \rangle$ is an Alexandroff space.

The upshot of all this is a one-to-one correspondence between quasi-ordered sets and Alexandroff spaces, and between partially ordered sets and Alexandroff T_0 -spaces. Since every finite topological space is an Alexandroff space, this immediately gives a one-to-one correspondence between finite quasi-ordered sets and finite topological spaces, and finite partially ordered sets and finite T_0 -spaces. It is straightforward to see that this also implies a one-to-one correspondence between order-preserving maps and continuous maps, as well as between p -morphisms and continuous and open maps.

3 General completeness

The preceding section shows that standard modal models are a particular case of a more general topological semantics. Hence, the known completeness of **S4** plus the topological soundness of its axioms immediately give us general topological completeness. Even so, we now give a direct model-theoretic proof of this result. It is closely related to the standard modal Henkin construction, and is much like completeness proofs for neighbourhood semantics (see [12]), but with some nice topological twists.

3.1 The main argument

Soundness is immediate, and so we move directly to completeness. Call a set Γ of formulas of \mathcal{L} (**S4**-)consistent if for no finite set $\{\varphi_1, \dots, \varphi_n\} \subseteq \Gamma$ we have that $\mathbf{S4} \vdash \neg(\varphi_1 \wedge \dots \wedge \varphi_n)$. A consistent set of formulas Γ is called *maximally consistent* if there is no consistent set of formulas properly containing Γ . It is well-known that Γ is maximally consistent iff, for any formula φ of \mathcal{L} , either $\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$, but not both. Now we define a topological space out of maximally consistent sets of formulas.

DEFINITION 3.1 (Canonical topological space)

The *canonical topological space* is the pair $S^\mathcal{L} = \langle W^\mathcal{L}, \tau^\mathcal{L} \rangle$, where:

- $W^\mathcal{L}$ is the set of all maximally consistent sets Γ_{max} ;
- $\tau^\mathcal{L}$ is the set generated by arbitrary unions of the following *basic sets* $B^\mathcal{L} = \{\widehat{\square\varphi} : \varphi \text{ is any formula}\}$, where $\widehat{\square\varphi} =_{\text{def}} \{\Gamma_{max} \in W^\mathcal{L} : \varphi \in \Gamma_{max}\}$. In other words, basic sets are the families of the form: $U_\varphi = \{\Gamma_{max} \in W^\mathcal{L} : \square\varphi \in \Gamma_{max}\}$.

Let us first check that $S^\mathcal{L}$ is indeed a topological space.

LEMMA 3.2

$B^{\mathcal{L}}$ forms a basis for the topology.

PROOF. We only need to show the following two properties:

- For any $U_\varphi, U_\psi \in B^{\mathcal{L}}$ and any $\Gamma_{max} \in U_\varphi \cap U_\psi$, there is $U_\chi \in B^{\mathcal{L}}$ such that $\Gamma_{max} \in U_\chi \subseteq U_\varphi \cap U_\psi$;
- For any $\Gamma_{max} \in W^{\mathcal{L}}$, there is $U_\varphi \in B^{\mathcal{L}}$ such that $\Gamma_{max} \in U_\varphi$.

Now (N) implies that $\Box\top \in \Gamma_{max}$ for any Γ_{max} . Hence, $W^{\mathcal{L}} = \widehat{\Box\top}$, and so the second item is satisfied. As for the first item, thanks to (R), one can easily check that $\Box(\widehat{\varphi \wedge \psi}) = \widehat{\Box\varphi} \cap \widehat{\Box\psi}$. Hence, $U_\varphi \cap U_\psi \in B^{\mathcal{L}}$, and so $B^{\mathcal{L}}$ is closed under finite intersections, whence, the first item is satisfied. ■

Next we define the canonical topological model.

DEFINITION 3.3 (Canonical topological model)

The *canonical topological model* is the pair $M^{\mathcal{L}} = \langle S^{\mathcal{L}}, \nu^{\mathcal{L}} \rangle$, where:

- $S^{\mathcal{L}}$ is the canonical topological space;
- $\nu^{\mathcal{L}}(p) = \{\Gamma_{max} \in X^{\mathcal{L}} : p \in \Gamma_{max}\}$.

The valuation $\nu^{\mathcal{L}}$ equates truth of a propositional variable *at* a maximally consistent set with its membership *in* that set. We now show this harmony between the two viewpoints lifts to all formulas.

LEMMA 3.4 (Truth lemma)

For all modal formulas φ ,

$$M^{\mathcal{L}}, w \models_{\mathcal{L}} \varphi \text{ iff } w \in \widehat{\varphi}.$$

PROOF. Induction on the complexity of φ . The base case follows from the definition above. The case of the Booleans follows from the following well-known identities for maximally consistent sets:

- $\widehat{\neg\varphi} = W^{\mathcal{L}} - \widehat{\varphi}$;
- $\widehat{\varphi \wedge \psi} = \widehat{\varphi} \cap \widehat{\psi}$.

The interesting case is that of the modal operator \Box . We do the two relevant implications separately, starting with the easy one.

⇐ ‘From membership to truth.’ Suppose $w \in \widehat{\Box\varphi}$. By definition, $\widehat{\Box\varphi}$ is a basic set, hence open. Moreover, thanks to axiom (T), $\widehat{\Box\varphi} \subseteq \widehat{\varphi}$. Therefore, there exists an open neighbourhood $U = \widehat{\Box\varphi}$ of w such that $v \in \widehat{\varphi}$, for any $v \in U$, and by the induction hypothesis, $M^{\mathcal{L}}, v \models_{\mathcal{L}} \varphi$. Thus $M^{\mathcal{L}}, w \models_{\mathcal{L}} \Box\varphi$.

⇒ ‘From truth to membership.’ Suppose $M^{\mathcal{L}}, w \models_{\mathcal{L}} \Box\varphi$. Then there exists a basic set $\widehat{\Box\psi} \in B^{\mathcal{L}}$ such that $w \in \widehat{\Box\psi}$ and $M^{\mathcal{L}}, v \models_{\mathcal{L}} \varphi$ for all $v \in \widehat{\Box\psi}$. By the induction hypothesis, $v \in \widehat{\varphi}$ for all $v \in \widehat{\Box\psi}$. Therefore, $\widehat{\Box\psi} \subseteq \widehat{\varphi}$. This implies that the logic **S4** can prove the implication $\Box\psi \rightarrow \varphi$. (If not, there would be some maximally consistent set containing both $\Box\psi$ and $\neg\varphi$.) But then we can prove the implication $\Box\Box\psi \rightarrow \Box\varphi$, and hence, using the **S4** transitivity axiom, $\Box\psi \rightarrow \Box\varphi$. It follows that $\widehat{\Box\psi} \subseteq \widehat{\Box\varphi}$, and so the world w belongs to $\widehat{\Box\varphi}$. ■

Now we can clinch the proof of our main result.

THEOREM 3.5 (Completeness)

For any set of formulas Γ ,

$$\text{if } \Gamma \models_{\mathcal{L}} \varphi \text{ then } \Gamma \vdash_{\mathbf{S4}} \varphi.$$

PROOF. Suppose that $\Gamma \not\vdash_{\mathbf{S4}} \varphi$. Then $\Gamma \cup \{\neg\varphi\}$ is consistent, and by the Lindenbaum Lemma it can be extended to a maximally consistent set Γ_{max} . By the Truth Lemma, $M^{\mathcal{L}}, \Gamma_{max} \models_{\mathcal{L}} \neg\varphi$, whence $M^{\mathcal{L}}, \Gamma_{max} \not\models_{\mathcal{L}} \varphi$, and we have constructed the required counter-model. \blacksquare

3.2 Topological comments

Let us now look at some topological aspects of this construction. In proving the box case of the Truth Lemma, we did not use the standard modal argument, which crucially invokes the *distribution axiom* of the minimal modal logic. Normally, one shows that, if a formula $\Box\phi$ does not belong to a maximally consistent set Γ , then there exists some maximally consistent successor set of Γ containing $\neg\phi$. This is not necessary in the topological version at this stage. We only needed the reflexivity and transitivity axioms, plus the Lindenbaum Lemma on maximally consistent extensions. The modal distribution axiom still plays a crucial role, but that was at the earlier stage of verifying that we had really defined a topology. This different perspective provides an additional proof-theoretic explanation why **S4** is the weakest axiom system complete for the topological semantics. Moreover, the divergence with the ‘standard’ argument explodes the prejudice that one single ‘well-known’ interpretation for a language must be the only natural one.

Comparing our construction given above with the standard modal Henkin model $\langle W^{\mathcal{L}}, R^{\mathcal{L}}, \nu^{\mathcal{L}} \rangle$ for **S4**, the basic sets of our topology $S^{\mathcal{L}}$ are $R^{\mathcal{L}}$ -upward closed. Hence, every open $S^{\mathcal{L}}$ is $R^{\mathcal{L}}$ -upward closed, and $S^{\mathcal{L}}$ is weaker than the topology $\tau_{R^{\mathcal{L}}}$ corresponding to $R^{\mathcal{L}}$. In particular, our canonical topological space is *not* an Alexandroff space.

Here are some further topological aspects of the above construction. First, it is worthwhile to compare Stone’s famous construction which uses the alternative basis $\{\widehat{\varphi} : \varphi \text{ any formula}\}$, yielding a space which we denote by $\langle W^{\mathcal{L}}, \tau^{\mathcal{S}} \rangle$. It is well known that $\langle W^{\mathcal{L}}, \tau^{\mathcal{S}} \rangle$ is homeomorphic to the Cantor space, and so, up to homeomorphism, $\langle W^{\mathcal{L}}, \tau^{\mathcal{S}} \rangle$ is compact, metric, 0-dimensional, and dense-in-itself. The basis of our topology, however, was the subfamily $\{\widehat{\Box\varphi} : \varphi \text{ any formula}\}$. Now every subtopology of one that is compact and dense-in-itself is also compact and dense-in-itself. Therefore, we get these same properties for our canonical topological space. But we can be more precise than this.

FACT 3.6

The canonical topology is actually the *intersection* of Kripke and Stone topologies.

In other words, $\tau^{\mathcal{L}} = \tau_{R^{\mathcal{L}}} \cap \tau^{\mathcal{S}}$. Indeed, since $\tau^{\mathcal{L}} \subseteq \tau_{R^{\mathcal{L}}}$ and $\tau^{\mathcal{L}} \subseteq \tau^{\mathcal{S}}$, obviously $\tau^{\mathcal{L}} \subseteq \tau_{R^{\mathcal{L}}} \cap \tau^{\mathcal{S}}$. Conversely, since every basic set $\widehat{\varphi}$ of Stone’s topology is $R^{\mathcal{L}}$ -upward closed iff $\widehat{\varphi} = \widehat{\Box\psi}$ for some ψ , we have $\tau_{R^{\mathcal{L}}} \cap \tau^{\mathcal{S}} \subseteq \tau^{\mathcal{L}}$, and so $\tau^{\mathcal{L}} = \tau_{R^{\mathcal{L}}} \cap \tau^{\mathcal{S}}$.

As an immediate consequence of the above observations we obtain:

COROLLARY 3.7

S4 is the logic of the class of all topological spaces which are compact and dense-in-itself.

Still, the canonical topological space $S^{\mathcal{L}}$ is neither 0-dimensional nor metric (it is not even a T_0 -space). So, $S^{\mathcal{L}}$ is not homeomorphic to the Cantor space. In Section 4 we show how to get completeness of **S4** with respect to the Cantor space by a different route.

3.3 Finite spaces suffice

We conclude Section 3 with an observation that is important for later arguments. The whole construction in the completeness proof would also work if we restricted attention to the *finite* language consisting of the initial formula and all its subformulas. All definitions go through for this restricted setting, and our arguments never needed to go beyond it. This means that we only get finitely many maximally consistent sets, and so nonprovable formulas can be refuted on *finite models*, whose size is effectively computable from the formula itself. (Note however that the obtained finite model is not necessarily dense-in-itself.)

COROLLARY 3.8

S4 has the *effective finite model property* with respect to the class of topological spaces.

Incidentally, this also shows that validity in **S4** is *decidable*, but we forego such complexity issues in this paper.

The resulting models have some interesting topological extras. Consider any finite frame $\mathcal{F} = \langle W, R \rangle$. We define some auxiliary notions. For any $w \in W$ let $C[w] = \{v \in W : wRv \ \& \ vRw\}$. Call a set C a *cluster* if it is of the form $C[w]$ for some w : the *cluster generated by w* . $C[w]$ is *simple* if $C[w] = \{w\}$, and *proper* otherwise. An element w of W is called *minimal* if vRw implies wRv for any $v \in W$. A cluster C is *minimal* if there exists a minimal $w \in W$ such that $C = C[w]$. Next, call \mathcal{F} *rooted* if there is $w \in W$ such that wRv for any $v \in W$: w is then a *root* of \mathcal{F} . This w need not be unique: any point from $C[w]$, the *initial cluster* of \mathcal{F} , will do.

Evidently, a finite Kripke frame \mathcal{F} is rooted iff it has only one minimal cluster. Topologically, this property is related to the notion of connectedness. We recall that a topological space $\langle W, \tau \rangle$ is *connected* if its universe cannot be written as a union of two disjoint open sets. $\langle W, \tau \rangle$ is *well-connected* if $W = U \cup V$ implies $W = U$ or $W = V$, for any $U, V \in \tau$. Obviously well-connectedness is a stronger notion than connectedness. It corresponds to $\langle W, R, \tau \rangle$ being rooted. For this observe that, dually, well-connectedness can be stated as follows:

For any two closed subsets C and D of $\langle W, \tau \rangle$, $C \cap D = \emptyset$ implies $C = \emptyset$ or $D = \emptyset$.

LEMMA 3.9

A finite Kripke frame is rooted iff the corresponding topological space is well-connected.

PROOF. Suppose $\langle W, R \rangle$ is a rooted Kripke frame with a root w , and $\langle W, \tau_R \rangle$ is the corresponding topological space. Let X_1 and X_2 be closed sets of $\langle W, \tau_R \rangle$ such that $X_1 \cap X_2 = \emptyset$. It is easy to see that a set $X \subseteq W$ is topologically closed iff it is *downward closed* in the ordering, that is $u \in X$ and vRu imply $v \in X$, for any $u, v \in W$. Now if both X_1 and X_2 are nonempty, then w belongs to both of them, which is a contradiction. Hence, one of them should be empty, and so $\langle W, \tau_R \rangle$ is well-connected.

Conversely, suppose $\langle W, R \rangle$ is not rooted. Then there are at least two different minimal clusters C_1 and C_2 in W . Since C_1 and C_2 are minimal clusters, they are downward closed, and hence closed in $\langle W, \tau_R \rangle$. Moreover, since they are different, $C_1 \cap C_2 = \emptyset$. Thus, $\langle W, \tau_R \rangle$ is not well-connected. ■

This allows us to improve on Corollary 3.8.

THEOREM 3.10

S4 is the logic of finite well-connected topological spaces.

PROOF. It suffices to observe the following. If a modal formula has a counter-example on a finite Kripke model, it fails at some point there. But then by standard modal locality it also fails in the submodel *generated* by that point and its relational successors, which is rooted and hence transforms into a well-connected topological space. ■

Again, there is a downside to such an upgraded completeness result. What it also means is that the basic modal language cannot *define* such a nice topological property as well-connectedness.

4 Completeness in real spaces

As early as 1944, McKinsey and Tarski [20] proved the following beautiful result, which is an expansion of a completeness theorem by Tarski [29] for intuitionistic propositional logic from 1938.

THEOREM 4.1 (McKinsey–Tarski)

S4 is the logic of any metric separable dense-in-itself space.

Most importantly, this theorem implies completeness of **S4** with respect to the real line \mathbb{R} , as well as with respect to the Cantor space \mathbf{C} . The original algebraic proof in [20] was very complex, the later more topological version in [24] is not much more accessible. Recently, Mints [21] replaced these by a much more perspicuous model-theoretic construction, extending earlier ideas of Beth and Kripke to get faster completeness of **S4** with respect to the Cantor space. We generalize its model-theoretic structure, using the topo-bisimulations of Section 2, and also provide a modification for completeness on the reals.

4.1 Cantorization

Our starting point is an arbitrary modal formula which is not provable in **S4**. We have already seen that such a nontheorem can be refuted on a finite rooted Kripke model. Now we show how to transform the latter into a topological counterexample on the Cantor space \mathbf{C} .

We think of the Cantor space as the countable topological product $\{0, 1\}^\omega$, where $\{0, 1\}$ is a two-element set with the discrete topology. Then the basic opens for the topology on \mathbf{C} are $U = \prod_{i \in \omega} U_i$, where all but finitely many U_i coincide with $\{0, 1\}$. To picture the Cantor space, we think of the full infinite binary tree \mathcal{T}_2 ; starting at the root, one associates 0 to every left-son of a node and 1 to every right-son (see Figure 1).

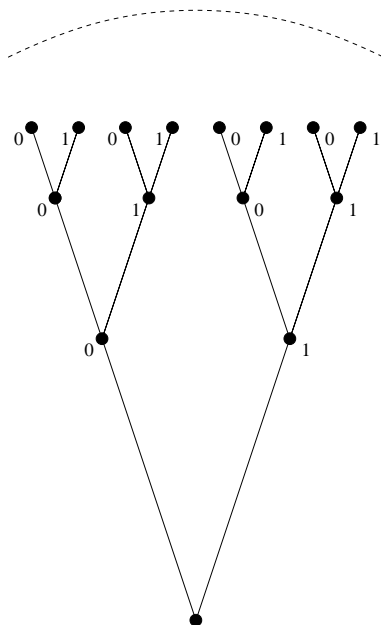
We call a countable sequence (t_1, \dots, t_n, \dots) of nodes of \mathcal{T}_2 an *infinite path* of \mathcal{T}_2 if t_1 is the root of \mathcal{T}_2 and each t_n is a node on the n th level of \mathcal{T}_2 . If the sequence is finite, then we call it a *finite path* of \mathcal{T}_2 . We think of points of the Cantor space as infinite paths of \mathcal{T}_2 . Therefore, if Σ denotes the set of all infinite paths of \mathcal{T}_2 , then Σ is in a one-to-one correspondence with \mathbf{C} . To transfer the topological structure of \mathbf{C} to Σ , for a finite path $X = (t_1, \dots, t_k)$ of \mathcal{T}_2 let

$$B_X = \{\sigma \in \Sigma : X \text{ an initial segment of } \sigma\}.$$

It is routine to observe that the sets B_X correspond to basic opens of \mathbf{C} . Thus, if we define topology τ_Σ on Σ by introducing

$$\mathcal{B} = \{B_X : X \text{ a finite path of } \mathcal{T}_2\}$$

as a basis, the following obvious fact holds true.

FIGURE 1. \mathcal{T}_2

FACT 4.2

$\langle \Sigma, \tau_\Sigma \rangle$ is homeomorphic to \mathbf{C} .

Now for any finite rooted Kripke model $M = \langle W, R, \nu \rangle$, we label \mathcal{T}_2 by the worlds of W recursively as follows.

DEFINITION 4.3

For any node $t \in \mathcal{T}_2$ the t -comb is the infinite leftward branch starting at t plus all one-step rightward successors of its nodes in \mathcal{T}_2 (see Figure 2).

For $w \in W$ let $w_1, \dots, w_{n(w)}$ denote some enumeration of $R[w]$.

- (1) We label the root of \mathcal{T}_2 by a root r of W .
- (2) We always repeat the same world as a label upwards to the left (see Figure 3).
- (3) If t is labelled by w , then we label the right nodes of the t -comb as follows: $w_1, w_2, w_3, \dots, w_{n(w)}$ and repeated modulo $n(w)$ (see Figure 4).

Suppose $\sigma \in \Sigma$. If σ is going infinitely to the left, that is σ is of the form $(X, 0, 0, 0, \dots)$, then by our labelling there exists $w \in W$ such that w labels every node of σ starting from some node on. In this case we say that w stabilizes σ . Otherwise, as node labels increase along the ordering of the model M , there exists a 'highest' cluster C of W such that every node of σ is labelled by an element of C starting from some node on. In this case we say that σ keeps cycling in C . Now we define a map $F : \Sigma \rightarrow W$ from branches to worlds by putting

$$F(\sigma) = \begin{cases} w & \text{if } w \text{ stabilizes } \sigma, \\ \rho(C) & \text{if } \sigma \text{ keeps cycling in } C \subseteq W, \text{ where } \rho(C) \text{ is some} \\ & \text{arbitrarily chosen representative of } C. \end{cases}$$

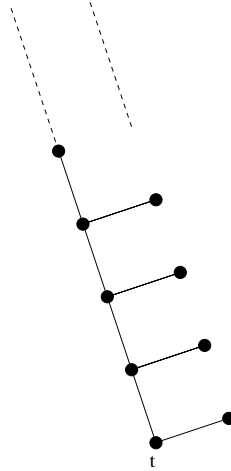


FIGURE 2. *t*-comb

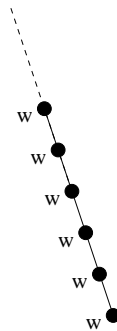


FIGURE 3. World repetition on the left

LEMMA 4.4

For any finite path X of \mathcal{T}_2 we have $F(B_X) = R[w]$, where w is the label occurring at the end of X .

PROOF. Since every branch $\sigma \in B_X$ carries label w , the definition of F implies that $F(B_X) \subseteq R[w]$. Conversely, if $v \in R[w]$, then by the above cyclic rule of labelling successors along a comb, there is a finite path Y extending X whose end is labelled by v . The infinite path $(Y, 0, 0, \dots) \in B_X$ is then mapped to v . ■

LEMMA 4.5

F is a continuous and open map from $\langle \Sigma, \tau_\Sigma \rangle$ onto $\langle W, \tau_R \rangle$.

PROOF. (a) F is onto. This follows immediately from the labelling of the t -comb, where t is the root of \mathcal{T}_2 .

(b) F preserves opens. By Lemma 4.4 basic opens B_X of Σ are mapped onto the basic opens $R[w]$ of the Alexandroff topology τ_R on W . It follows that F preserves arbitrary opens.

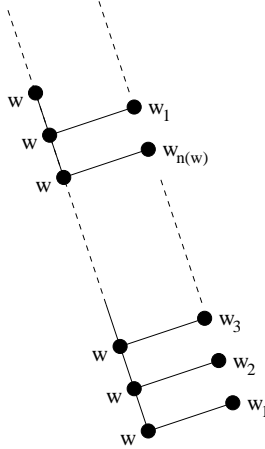


FIGURE 4. First t -comb labelling

(c) F reflects opens. It suffices to look at basic opens of the form $R[w]$. Now, for $w \in W$ let

$$U = \bigcup \{B_X : X \text{ a finite path of } \mathcal{T}_2 \text{ whose end is labelled by } v \in R[w]\},$$

Obviously $U \in \tau_\Sigma$. Hence, it is sufficient to show that $F^{-1}(R[w]) = U$. First, it follows from Lemma 4.4 that $F(U) = F(\bigcup B_X) = \bigcup F(B_X) \subseteq R[w]$. Therefore, $U \subseteq F^{-1}(R[w])$. Conversely, suppose $\sigma \in F^{-1}(R[w])$, i.e. $wRF(\sigma)$. Let X be a finite path of \mathcal{T}_2 such that X carries label w and the end of X is labelled by $F(\sigma)$. Then $\sigma \in B_X \subseteq U$, and so $F^{-1}(R[w]) \subseteq U$. ■

Now we define a valuation ν_Σ on Σ by putting $\nu_\Sigma(p) = F^{-1}(\nu(p))$.

COROLLARY 4.6

F is a total topo-bisimulation between $M_\Sigma = \langle \Sigma, \tau_\Sigma, \nu_\Sigma \rangle$ and $M = \langle W, R, \nu \rangle$.

PROOF. Apply Lemma 4.5 and Fact 2.2. ■

THEOREM 4.7

S4 is complete with respect to the Cantor space.

PROOF. Suppose **S4** $\not\models \varphi$. Then by Theorem 3.10 there is a finite rooted Kripke model M refuting φ . By Fact 4.2 and Corollary 4.6 there exists a valuation ν_C on the Cantor space C such that $\langle C, \nu_C \rangle$ is topo-bisimilar to M . Hence, φ is refuted on C . ■

4.2 Counterexamples on the reals

In the previous section we described how to transform counterexamples on a finite rooted Kripke model M into counterexamples on the Cantor space C . In this section we show how to transfer counterexamples from M to the open interval $(0, 1)$ of real numbers. As a result, we obtain a new proof of completeness of **S4** with respect to the real line \mathbb{R} . Our strategy is similar to that of Section 4.1. First we establish a one-to-one correspondence between $(0, 1)$

and a suitable set $\Lambda \subset \Sigma$ of infinite paths of the full infinite binary tree \mathcal{T}_2 . Then we transfer the topology from $(0, 1)$ to Λ . Finally, we label nodes of \mathcal{T}_2 by worlds of M in a way which allows us to define a valuation ν_Λ on Λ such that $\langle \Lambda, \tau_\Lambda, \nu_\Lambda \rangle$ is topo-bisimilar to M . This labelling is a bit more complicated than the previous one, but it shows the freedom one has on the tree within the ‘back and forth’ constraints of a modal simulation. It will readily follow that **S4** is complete with respect to $(0, 1)$, and hence with respect to \mathbb{R} .

We note that there exist three different types of infinite paths of \mathcal{T}_2 : (1) σ is going infinitely to the left if $\sigma = (X, 0, 0, 0, \dots)$; (2) σ is going infinitely to the right if $\sigma = (Y, 1, 1, 1, \dots)$; finally, (3) σ is infinitely zigzagging if it is neither going infinitely to the left nor infinitely to the right. Let Π denote the set of infinite paths of \mathcal{T}_2 that are either going infinitely to the left or are infinitely zigzagging. We also let $\Lambda = \Pi - \{(0, 0, 0, \dots)\}$. Then there is a one-to-one correspondence between Λ and $(0, 1)$. To see this recall the dyadic representation of a number from $(0, 1)$. Let $x \in (0, 1)$. To construct an infinite path $\sigma = (t_n)_{n \in \omega}$ of \mathcal{T}_2 representing x observe that either $x \in (0, \frac{1}{2})$ or $x \in [\frac{1}{2}, 1)$. In the former case put $t_1 = 0$ and in the latter case put $t_1 = 1$. Assume $x \in (0, \frac{1}{2})$. Then either $x \in (0, \frac{1}{4})$ or $x \in [\frac{1}{4}, \frac{1}{2})$. In the former case put $t_2 = 0$ and in the latter case put $t_2 = 1$. Continuing this process, we will get an infinite path $\sigma = (t_n)_{n \in \omega}$ of \mathcal{T}_2 representing x .

Note that there are two ways for the dyadic representation of a number $\frac{m}{2^n} \in (0, 1)$, where $m, n \in \omega$ and $0 < m < 2^n$; either as $(X, 1, 0, 0, 0, \dots)$ or as $(X, 0, 1, 1, 1, \dots)$. Therefore, if we throw away $(0, 0, 0, \dots)$ together with all infinite paths of \mathcal{T}_2 going infinitely to the right, we obtain a one-to-one correspondence between $(0, 1)$ and the remaining infinite paths of \mathcal{T}_2 . Hence, there exists a bijection, say f , between $(0, 1)$ and Λ . In order to transfer the topological structure of $(0, 1)$ to Λ , we need the following lemma.

LEMMA 4.8

The family $\mathcal{B} = \{(\frac{m}{2^n}, \frac{m+2}{2^n}) : m, n \in \omega; m + 2 \leq 2^n\}$ forms a basis for the topology on $(0, 1)$.

PROOF. Suppose $x \in (0, 1)$. For an arbitrarily small $\varepsilon > 0$, select $n \in \omega$ so that $\frac{1}{2^{n-1}} < \varepsilon$. Let $m \in \omega$ be the largest integer for which $\frac{m}{2^n} < x$. Then $\frac{m}{2^n} < x \leq \frac{m+1}{2^n} < \frac{m+2}{2^n}$. From $x \leq \frac{m+1}{2^n}$ it follows that $x - \frac{1}{2^n} \leq \frac{m}{2^n}$. Since $\frac{1}{2^n} < \varepsilon$, we have that $x - \varepsilon < \frac{m}{2^n}$. Also $\frac{m}{2^n} < x$ and $\frac{1}{2^{n-1}} < \varepsilon$ imply that $\frac{m+2}{2^n} < x + \frac{2}{2^n} < x + \varepsilon$. Therefore, there exist $m, n \in \omega$ such that $x - \varepsilon < \frac{m}{2^n} < x \leq \frac{m+1}{2^n} < \frac{m+2}{2^n} < x + \varepsilon$. Thus, $x \in (\frac{m}{2^n}, \frac{m+2}{2^n}) \subseteq (x - \varepsilon, x + \varepsilon)$, and so there exists a member of \mathcal{B} containing x and contained in $(x - \varepsilon, x + \varepsilon)$. This implies that \mathcal{B} forms a basis for the topology on $(0, 1)$. ■

For $(\frac{m}{2^n}, \frac{m+2}{2^n}) \in \mathcal{B}$ two cases are possible: (1) If m is even, then $m = 2k$ for some k , and so $(\frac{m}{2^n}, \frac{m+2}{2^n}) = (\frac{k}{2^{n-1}}, \frac{k+1}{2^{n-1}})$; and (2) if m is odd, then $m = 2k + 1$ for some k , and so $(\frac{m}{2^n}, \frac{m+2}{2^n}) = (\frac{2k+1}{2^n}, \frac{2k+3}{2^n})$. Therefore, we can think of \mathcal{B} as the union of two kinds of sets $\mathcal{B}_1 = \{(\frac{k}{2^n}, \frac{k+1}{2^n}) : k, n \in \omega; k + 1 \leq 2^n\}$ and $\mathcal{B}_2 = \{(\frac{2k+1}{2^n}, \frac{2k+3}{2^n}) : k, n \in \omega; 2k + 3 \leq 2^n\}$. It follows from the proof of Lemma 4.8 that if x does not have the form $\frac{a}{2^b}$, then there exist m, n such that $x \in (\frac{m}{2^n}, \frac{m+1}{2^n}) \subseteq (x - \varepsilon, x + \varepsilon)$. So the only reason we need elements of \mathcal{B}_2 is to find small neighbourhoods around the points of the form $\frac{a}{2^b}$.

Now we take the set of f -images of elements of \mathcal{B} as a basis for the topology τ_Λ on Λ . Then it should be clear that $\langle \Lambda, \tau_\Lambda \rangle$ is homeomorphic to $(0, 1)$. To understand better what the f -image of an element of \mathcal{B} is, for a finite path $X = (t_1, \dots, t_k)$ of \mathcal{T}_2 , let

$$C_X = \{\lambda \in \Lambda : X \text{ an initial segment of } \lambda\} = B_X \cap \Lambda$$

and

$$D_X = C_X - \{(X, 0, 0, 0, \dots)\}.$$

We observe that if $(X, 0, 0, 0, \dots)$ represents $\frac{k}{2^n}$, then $(X, 1, 1, 1, \dots)$ represents $\frac{k+1}{2^n}$, and so D_X represents $(\frac{k}{2^n}, \frac{k+1}{2^n})$. Therefore, elements of \mathcal{B}_1 are represented as D_X s, where X is a finite path of \mathcal{T}_2 . To represent elements of \mathcal{B}_2 we need the following definition.

DEFINITION 4.9

(1) We call $t, s \in \mathcal{T}_2$ *back-to-back nodes* if t is the end of the finite path $(X, 0, \underbrace{1, \dots, 1}_{n\text{-times}})$ and s is the end of the finite path $(X, 1, \underbrace{0, \dots, 0}_{n\text{-times}})$. In a picture, they sit on top of the following

W-formation in \mathcal{T}_2 (see Figure 5).

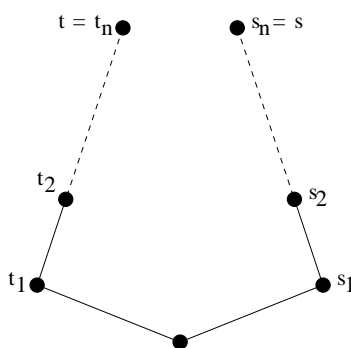


FIGURE 5. W-formation

(2) We call back-to-back nodes t, s *at least 3-deep* if $n \geq 3$.

Let $2k+3 \leq 2^n$. If $(X, 0, 0, 0, \dots)$ represents $\frac{2k+1}{2^n}$, where X ends in t , then $(Y, 1, 1, 1, \dots)$ represents $\frac{2k+3}{2^n}$, where Y ends in s and t, s are back-to-back nodes. Therefore, $D_X \cup C_Y$ represents $(\frac{2k+1}{2^n}, \frac{2k+3}{2^n})$. Thus, elements of \mathcal{B}_2 are represented as $D_X \cup C_Y$, where X and Y are finite paths of \mathcal{T}_2 ending in back-to-back nodes t, s . So if we define topology τ_Λ on Λ by taking the sets

$$\{D_X : X \text{ a finite path of } \mathcal{T}_2\}$$

and

$$\{D_X \cup C_Y : X \text{ and } Y \text{ finite paths of } \mathcal{T}_2 \text{ ending in back-to-back nodes } t, s\}$$

as a basis, the following obvious fact holds:

FACT 4.10

$\langle \Lambda, \tau_\Lambda \rangle$ is homeomorphic to $(0, 1)$.

Now for any finite rooted Kripke model $M = \langle W, R, \nu \rangle$, we label the nodes of \mathcal{T}_2 by worlds of W recursively as follows. For $w \in W$ let $w_1, \dots, w_{n(w)}$ denote some enumeration of $R[w]$.

- (1) We label the root of \mathcal{T}_2 by a root r of W .
- (2) We always repeat the same world as a label upwards on the left.
- (3) If t is labelled by w , then we label the right nodes of the t -comb as follows: $w_i, w, w_{i+1}, w, \dots, w_{n(w)}, w, w_1, w, \dots, w_{i-1}, w$ and repeated modulo $n(w)$ (see Figure 6).

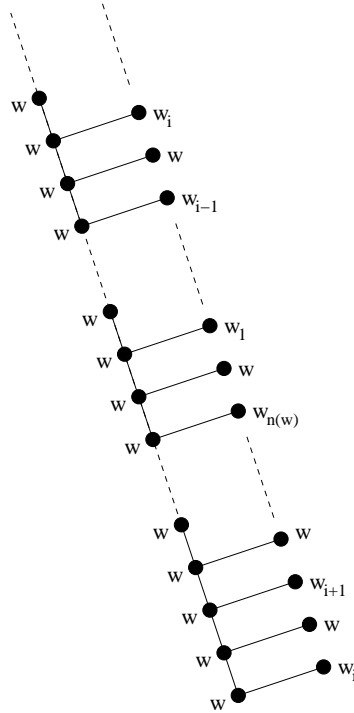


FIGURE 6. Second t -comb labelling

Clause (3) involves repetition of the initial label w in between successors of w . The point of this trick will become clear below. Also note that we left the exact starting point w_i for the successor labelling along the teeth of the t -comb unspecified. To make the choice of w_i , we look at the node s preceding t , if any. Given our inductive procedure, we only need to consider the case where t is a right-successor of s , as labels were merely repeated going leftward. Suppose that s is labelled by v . Then two cases are possible:

Case 1: w is a strong successor of v in M . Then we just put $w_i = w_1$.

Case 2: The two world labels are the same: $v = w$. Then, if s is the root of \mathcal{T}_2 , again put $w_i = w_1$. Otherwise, by our labelling procedure, the situation is as in Figure 7. That is, the node s is a left-successor of some node labelled with w , and the latter's right-successor is labelled by some successor $u \in R[w]$. In this case, put $w_i = u$.

Now we prove a crucial property of this construction.

LEMMA 4.11

If w labels t and v labels s , where t, s are at least 3-deep back-to-back nodes of \mathcal{T}_2 , then vRw .

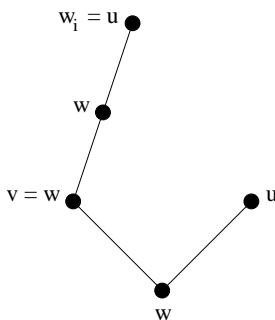


FIGURE 7. Case 2

PROOF. Let w_i label t_i and v_i label s_i as in Figure 5. Then $w = w_n$, $v = v_n$, and s_1 and t_2 are on the teeth of the same comb. So by our labelling procedure only the preceding two cases are possible.

Case 1: $w_1 = v_1$. Then vRw because $v = v_1$, $v_1 = w_1$ and w_1Rw since we always took R -successors going up any branch.

Case 2: v_1 is a strong successor of w_1 . Then by the repetition rule for labels, $w_2 = w_1$, and w_2 's first rightward successor, which in the worst case is w since the back-to-back nodes are at least 3-deep, is labelled v_1 by the above rule. So again $v = v_1Rw$. ■

The rest of the proof is similar to the proof for the Cantor space, but with some additional verifications. First, we define $G : \Lambda \rightarrow W$ by putting:

$$G(\lambda) = \begin{cases} w & \text{if } w \text{ stabilizes } \lambda, \\ \rho(C) & \text{if } \lambda \text{ keeps cycling in } C. \end{cases}$$

Thus, G is the restriction of the earlier F to Λ .

LEMMA 4.12

For any finite path X of \mathcal{T}_2 we have $G(D_X) = G(C_X) = R[w]$, where w is the label of the end t of X .

PROOF. The argument is like before. Since $D_X \subseteq C_X$, it is obvious that $G(D_X) \subseteq G(C_X)$. Also since every path $\lambda \in C_X$ carries label w , we have $G(C_X) \subseteq R[w]$. To see the third inclusion let $v \in R[w]$. By our labelling we find recurring repetitions of label v on the teeth of the t -comb. Let s be on the teeth of the t -comb labelled by v , and let Y be a finite path of \mathcal{T}_2 with the end s . Then X is an initial segment of Y , whence $(Y, 0, 0, 0, \dots) \in D_X$. By the definition of G we have $G(Y, 0, 0, 0, \dots) = v$. So $R[w] \subseteq G(D_X)$. ■

LEMMA 4.13

G is a continuous and open map from $\langle \Lambda, \tau_\Lambda \rangle$ onto $\langle W, \tau_R \rangle$.

PROOF. (a) G is onto. This follows immediately from the labelling of the t -comb, where t is the root of \mathcal{T}_2 .

(b) G preserves opens. It suffices to consider basic opens in the new topology, which are of the form D_X and $D_X \cup C_Y$, where the ends of X and Y are back-to-back nodes. By Lemma

4.12 the sets $G(D_X)$ and $G(D_X \cup D_Y)$ are mapped to unions of open sets of the form $R[w]$, which are open in the Alexandroff topology τ_R on W .

(c) G reflects opens. Again it suffices to look at basic opens in M . Let w be any world in W . We let

$$U_1 = \bigcup \{D_X : X \text{ a finite path of } \mathcal{T}_2 \text{ whose end is labelled by } u \in R[w]\}$$

and

$$U_2 = \bigcup \{D_X \cup C_Y : X, Y \text{ finite paths of } \mathcal{T}_2 \text{ whose ends } t, s \text{ are at least 3-deep back-to-back nodes with respective labels } v, u \text{ such that } u \in R[w]\}.$$

It is sufficient to show that $G^{-1}(R[w]) = U_1 \cup U_2$. First, consider the inclusion from right to left. Lemmas 4.11 and 4.12 imply that $G(U_1 \cup U_2) = G(U_1) \cup G(U_2) \subseteq R[w]$. In particular, for the case of U_2 we note that $G(D_X \cup C_Y) = G(D_X) \cup G(D_Y) = R[v] \cup R[u] = R[u] \subseteq R[w]$ because $wRuRv$. Therefore, $U_1 \cup U_2 \subseteq G^{-1}(R[w])$. For the inclusion from left to right, suppose that $\lambda \in G^{-1}(R[w])$, i.e. $wRG(\lambda)$. Now either λ is infinitely zigzagging or λ is going infinitely to the left.

Case 1: λ is infinitely zigzagging. Then it is in some open set D_X with X carrying label w and the end of X labelled by some $u \in R[w]$. Therefore, $\lambda \in U_1$.

Case 2: λ is of the form $(Y, 0, 0, 0, \dots)$. Then λ lies in an open set $D_X \cup C_Y$, where the ends of X and Y are back-to-back nodes. As the open neighbourhoods $D_X \cup C_Y$ can be chosen arbitrarily small, we can assume that the back-to-back end nodes of X, Y are at least 3-deep. Let v label the end of X and u label the end of Y . By Lemma 4.11 we have uRv . Also, since Y carries label w , wRu . Therefore, $\lambda \in D_X \cup C_Y \subseteq U_2$.

In either case we have $G^{-1}(R[w]) \subseteq U_1 \cup U_2$, thus the equality. ■

Finally, similar to the Cantor space argument, we define a valuation on Λ by copying the valuation ν of the model M using the inverse of G . By a reasoning exactly like before, we get:

COROLLARY 4.14

G is a total topo-bisimulation between $M_\Lambda = \langle \Lambda, \tau_\Lambda, \nu_\Lambda \rangle$ and $M = \langle W, R, \nu \rangle$.

This gives our main result.

THEOREM 4.15

1. **S4** is complete with respect to $(0, 1)$.
2. **S4** is complete with respect to the real line \mathbb{R} .

PROOF. (1) Suppose **S4** $\not\models \varphi$. Then by Theorem 3.10 there is a finite rooted Kripke model M refuting φ . By Corollary 4.14, M is topo-bisimilar to $M_\Lambda = \langle \Lambda, \tau_\Lambda, \nu_\Lambda \rangle$. Hence, M_Λ is refuting φ . Now since $\langle \Lambda, \tau_\Lambda \rangle$ is homeomorphic to $(0, 1)$, φ is refuted on $(0, 1)$.

(2) Suppose **S4** $\not\models \varphi$. Then by (1) there exists a valuation $\nu_{(0,1)}$ on $(0, 1)$ refuting φ . Now since $(0, 1)$ is homeomorphic to \mathbb{R} , φ is refuted on \mathbb{R} . ■

This provides an alternative proof of McKinsey and Tarski's original theorem. The method seems quite flexible, as we can vary tree labelling very widely, while keeping a topo-

bisimulation with the initial Kripke model. Another attraction is the immediate geometrical character of the procedure. Inductive tree labelling with recurring world labels is a bit like imposing fractal structure on a representation of the space.

Finally, we note that a minor further step establishes the following version of the completeness result.

COROLLARY 4.16

There exists a single valuation ν on \mathbb{R} falsifying all the nontheorems of **S4**.

PROOF. Enumerate all the nontheorems of **S4**. This can be done since the language of **S4** is countable. Let this enumeration be $\{\varphi_1, \varphi_2, \dots\}$. Since the interval $(n, n+1)$ is homeomorphic to \mathbb{R} , from Theorem 4.15 it follows that there exists a valuation ν_n on $(n, n+1)$ such that $\langle (n, n+1), \nu_n \rangle$ falsifies φ_n . (Note that we need not know anything about the shape of $\nu_n(\varphi_n)$.) For any $p \in \mathcal{P}$ let $\nu(p) = \bigcup_{n \in \omega} \nu_n(p)$ be the valuation of p on \mathbb{R} . Note that each $\langle (n, n+1), \nu_n \rangle$ is an open submodel of $\langle \mathbb{R}, \nu \rangle$, where the ‘identity embedding’ is a topo-bisimulation. Hence, the truth values of modal formulas do not change moving from each $\langle (n, n+1), \nu_n \rangle$ to $\langle \mathbb{R}, \nu \rangle$. Therefore, φ_n is still falsified on the whole \mathbb{R} for each n . Thus, we have constructed a single valuation ν on \mathbb{R} falsifying all the nontheorems of **S4**. ■

This corollary also shows that, though very different from the standard canonical Kripke model of **S4**, \mathbb{R} shares some of its universal properties.

4.3 Logical nonfiniteness on the reals

Recall that two formulas φ and ψ are said to be **S4-equivalent** if $\mathbf{S4} \vdash \varphi \leftrightarrow \psi$. It is well known that there exist infinitely many formulas in one variable which are not **S4-equivalent**. Consider the following list of formulas:

$$\begin{aligned} \varphi_0 &= p; \\ \varphi_n &= \varphi_{n-1} \wedge \diamond(\diamond\varphi_{n-1} \wedge \neg\varphi_{n-1}). \end{aligned}$$

We can easily construct a Kripke model in which all φ_n have different interpretations. Let $M = \langle \omega, R, \nu \rangle$, where ω denotes the set of all natural numbers, nRm iff $m \leq n$, and $n \in \nu(p)$ iff n is even. Then one can readily check that φ_n is true at all even points $\geq 2n$. Hence, every φ_n has a different interpretation on M . It implies that the φ_n are not **S4-equivalent**. Now we will give a topological flavour to this result by showing that interpreting a propositional variable as a certain subset of \mathbb{R} allows us to construct infinitely many **S4**-nonequivalent formulas of one variable. Corollary 4.16 already told us such a uniform choice must exist, but the proof does not construct $\nu(p)$ explicitly. The following argument does, and thereby also highlights the topological content of our modal completeness theorem in a more concrete fashion.

We use \square and \diamond instead of the standard notations *Int* and $\bar{}$ for the interior and closure operators of a topological space. This modal notation shows its basic use in topology because it allows us to write topological formulas in a much more perspicuous fashion.

To proceed further we need to recall the definition of Hausdorff’s residue of a given set. Suppose a topological space $\langle W, \tau \rangle$ and $X \subseteq W$ are given. $\varrho(X) = X \cap \diamond(\diamond X - X)$ is called the *Hausdorff residue* of X . Let $\varrho^0(X) = X$, $\varrho^1(X) = \varrho(X)$ and $\varrho^{n+1}(X) = \varrho\varrho^n(X)$.

X is said to be of *rank* n , written $r(X) = n$, if n is the least natural number such that $\varrho^n(X) = \emptyset$; X is said to be of *finite rank* if there exists a natural n such that X is of rank n ; X is said to be of *infinite rank* if it is not of finite rank.

$x \in X$ is said to be of rank n if $x \in \varrho^n(X)$, but $x \notin \varrho^{n+1}(X)$; $x \in X$ is said to be of finite rank if there exists a natural n such that x is of rank n ; x is said to be of infinite rank if it is not of finite rank.

Obviously X is of rank n iff the rank of every element of X is strictly less than n , and there is at least one element of X of rank $n - 1$; X is of finite rank iff there is a natural n such that the rank of every element of X is strictly less than n ; and X is of infinite rank iff there is no finite bound on the ranks of elements of X .

It is obvious that if we interpret p as a subset X of \mathbb{R} , then φ_n will be interpreted as $\varrho^n(X)$. So, in order to show that different φ_n are **S4**-nonequivalent, it is sufficient to show that there is $X \subset \mathbb{R}$ such that $X \supset \varrho(X) \supset \varrho^2(X) \supset \dots \supset \varrho^n(X) \supset \dots$.

PROPOSITION 4.17

There exists a subset X of \mathbb{R} such that $X \supset \varrho(X) \supset \varrho^2(X) \supset \dots \supset \varrho^n(X) \supset \dots$.

PROOF. We will construct X inductively. Fix a natural number k .

Step 1: Consider a sequence $\{x_{i_1}\}_{i_1=1}^\infty$ from $(k - 1, k)$ converging to $k - 1$, and put

$$X_1 = \{k - 1\} \cup \bigcup_{i_1=1}^\infty \{y_{i_2}^{i_1}\}_{i_2=1}^\infty,$$

where $\{y_{i_2}^{i_1}\}_{i_2=1}^\infty$ is a sequence from (x_{i_1+1}, x_{i_1}) converging to x_{i_1+1} . Note that

- $\diamond X_1 = X_1 \cup \{x_{i_1}\}_{i_1=1}^\infty$,
- $\diamond X_1 - X_1 = \{x_{i_1}\}_{i_1=1}^\infty$,
- $\diamond(\diamond X_1 - X_1) = \{k - 1\} \cup \{x_{i_1}\}_{i_1=1}^\infty$, and
- $\varrho(X_1) = \{k - 1\}$.

So, $k - 1$ is the only point of X_1 of rank 1, and $r(X_1) = 2$.

Step 2: Consider a sequence $\{x_{i_3}^{i_1, i_2}\}_{i_3=1}^\infty$ from $(y_{i_2+1}^{i_1}, y_{i_2}^{i_1})$ converging to $y_{i_2+1}^{i_1}$, and put

$$X_2 = \{k - 1\} \cup \bigcup_{i_1=1}^\infty \{y_{i_2}^{i_1}\}_{i_2=1}^\infty \cup \bigcup_{i_1=1}^\infty \bigcup_{i_2=1}^\infty \bigcup_{i_3=1}^\infty \{y_{i_4}^{i_1, i_2, i_3}\}_{i_4=1}^\infty,$$

where $\{y_{i_4}^{i_1, i_2, i_3}\}_{i_4=1}^\infty$ is a sequence from $(x_{i_3+1}^{i_1, i_2}, x_{i_3}^{i_1, i_2})$ converging to $x_{i_3+1}^{i_1, i_2}$. Note that $X_2 \supset X_1$, and

- $\diamond X_2 = X_2 \cup \{x_{i_1}\}_{i_1=1}^\infty \cup \bigcup_{i_1=1}^\infty \bigcup_{i_2=1}^\infty \{x_{i_3}^{i_1, i_2}\}_{i_3=1}^\infty$,
- $\diamond X_2 - X_2 = \{x_{i_1}\}_{i_1=1}^\infty \cup \bigcup_{i_1=1}^\infty \bigcup_{i_2=1}^\infty \{x_{i_3}^{i_1, i_2}\}_{i_3=1}^\infty$,
- $\diamond(\diamond X_2 - X_2) = \{k - 1\} \cup \bigcup_{i_1=1}^\infty \{y_{i_2}^{i_1}\}_{i_2=1}^\infty \cup \{x_{i_1}\}_{i_1=1}^\infty \cup \bigcup_{i_1=1}^\infty \bigcup_{i_2=1}^\infty \{x_{i_3}^{i_1, i_2}\}_{i_3=1}^\infty$,
- $\varrho(X_2) = \{k - 1\} \cup \bigcup_{i_1=1}^\infty \{y_{i_2}^{i_1}\}_{i_2=1}^\infty$, and
- $\varrho^2(X_2) = \{k - 1\}$.

So, the points of X_2 of rank 1 are $y_{i_2}^{i_1}$, for arbitrary i_1 and i_2 , $k - 1$ is the only point of X_2 of rank 2, and $r(X_2) = 3$.

Step n: Consider a sequence $\{x_{i_{2n-1}}^{i_1, \dots, i_{2n-2}}\}_{i_{2n-1}=1}^\infty$ from $(y_{i_{2n-2}+1}^{i_1, \dots, i_{2n-3}}, y_{i_{2n-2}}^{i_1, \dots, i_{2n-3}})$ converging to $y_{i_{2n-2}+1}^{i_1, \dots, i_{2n-3}}$, and put

$$X_n = \{k-1\} \cup \bigcup_{i_1=1}^\infty \{y_{i_2}^{i_1}\}_{i_2=1}^\infty \cup \dots \cup \bigcup_{i_1=1}^\infty \dots \bigcup_{i_{2n-1}=1}^\infty \{y_{i_{2n}}^{i_1, \dots, i_{2n-1}}\}_{i_{2n}=1}^\infty,$$

where $\{y_{i_{2n}}^{i_1, \dots, i_{2n-1}}\}_{i_{2n}=1}^\infty$ is a sequence from $(x_{i_{2n-1}+1}^{i_1, \dots, i_{2n-2}}, x_{i_{2n-1}}^{i_1, \dots, i_{2n-2}})$ converging to $x_{i_{2n-1}+1}^{i_1, \dots, i_{2n-2}}$. Also let

$$A = \{x_{i_1}\}_{i_1=1}^\infty \cup \dots \cup \bigcup_{i_1=1}^\infty \dots \bigcup_{i_{2n-2}=1}^\infty \{x_{i_{2n-1}}^{i_1, \dots, i_{2n-2}}\}_{i_{2n-1}=1}^\infty.$$

Then note that $X_n \supset X_{n-1} \supset \dots \supset X_2 \supset X_1$, and

$$\begin{aligned} \diamond X_n &= X_n \cup A, \\ \diamond X_n - X_n &= A, \\ \diamond(\diamond X_n - X_n) &= A \cup (X_n - [\bigcup_{i_1=1}^\infty \dots \bigcup_{i_{2n-1}=1}^\infty \{y_{i_{2n}}^{i_1, \dots, i_{2n-1}}\}_{i_{2n}=1}^\infty]), \\ \varrho(X_n) &= X_n - [\bigcup_{i_1=1}^\infty \dots \bigcup_{i_{2n-1}=1}^\infty \{y_{i_{2n}}^{i_1, \dots, i_{2n-1}}\}_{i_{2n}=1}^\infty], \\ \varrho^2(X_n) &= \rho(X_n) - [\bigcup_{i_1=1}^\infty \dots \bigcup_{i_{2n-3}=1}^\infty \{y_{i_{2n-2}}^{i_1, \dots, i_{2n-3}}\}_{i_{2n-2}=1}^\infty], \\ &\dots \\ \varrho^n(X_n) &= \{k-1\}. \end{aligned}$$

So, the points of X_n of rank 1 are

$$X_n - [\bigcup_{i_1=1}^\infty \dots \bigcup_{i_{2n-1}=1}^\infty \{y_{i_{2n}}^{i_1, \dots, i_{2n-1}}\}_{i_{2n}=1}^\infty],$$

the points of X_n of rank 2 are

$$\varrho^2(X_n) = \rho(X_n) - [\bigcup_{i_1=1}^\infty \dots \bigcup_{i_{2n-3}=1}^\infty \{y_{i_{2n-2}}^{i_1, \dots, i_{2n-3}}\}_{i_{2n-2}=1}^\infty],$$

and so on; finally, $k-1$ is the only point of X_n of rank n , and $r(X_n) = n+1$.

Now let X_1 be constructed in $(0, 1)$, X_2 in $(1, 2)$, X_n in $(n-1, n)$, and so on. We put

$$X = \bigcup_{n=1}^\infty X_n.$$

Then it is obvious that $n-1 \in \varrho^n(X)$ and $n-1 \notin \varrho^{n+1}(X)$, for any natural n . So, $X \supset \varrho(X) \supset \varrho^2(X) \supset \dots \supset \varrho^n(X) \supset \dots$, and X contains points of every finite rank. ■

REMARK 4.18

It is worth noting that the X constructed above does not contain elements of infinite rank. However, a little adjustment of the above construction will allow us to construct a subset of \mathbb{R} with an element of infinite rank. Actually, it is possible to construct a subset of \mathbb{R} containing elements of rank α , for any ordinal $\alpha < \aleph_1$.

Returning to our list of formulas, with p as the just constructed X , the interpretation of every φ_n in \mathbb{R} will be different. In the next section we show that if we restrict ourselves to only 'good' subsets of \mathbb{R} , then the situation will be drastically changed.

5 Axiomatizing special kinds of regions

As we saw in the previous section, by interpreting propositional variables as certain subsets of the real line \mathbb{R} , we can refute every nontheorem of **S4** on \mathbb{R} . Certainly not all subsets of \mathbb{R} are required for refuting the nontheorems of **S4**. In this section, we will analyse the complexity of the subsets of \mathbb{R} required for refuting the nontheorems of **S4**. Similarly to Section 4.5, we prefer to use \square and \diamond to denote the interior and closure operators of a topological space. For consistency we also use \wedge, \vee and \neg to denote the set-theoretical intersection, union, and complement.

5.1 Serial sets on the real line

To start with, consider subsets of \mathbb{R} with the simplest intuitive structure. Call $X \subseteq \mathbb{R}$ *convex* if all points lying in between any two points of X belong to X . In other words, X is convex if $x, y \in X$ and $x \leq y$ imply $[x, y] \subseteq X$. Every convex subset of \mathbb{R} has one of the following forms:

$$\emptyset, (x, y), [x, y], [x, y), (x, y], (-\infty, x), (-\infty, x], (x, +\infty), [x, +\infty), \mathbb{R}.$$

DEFINITION 5.1

Call a subset of \mathbb{R} *serial* if it is a finite union of convex subsets of \mathbb{R} . Denote the set of all serial subsets of \mathbb{R} by $\mathcal{S}(\mathbb{R})$. So,

$$\mathcal{S}(\mathbb{R}) = \{X \in \mathcal{P}(\mathbb{R}) : X \text{ is a serial subset of } \mathbb{R}\}.$$

Obviously the X constructed in Proposition 4.17 is not serial, and actually this was absolutely crucial in showing that X had points of any finite rank. Indeed, we have the following:

LEMMA 5.2

$r(X) = 0$ for any $X \in \mathcal{S}(\mathbb{R})$.

PROOF. First, $r(Y) = 0$ for any convex subset Y of \mathbb{R} . For, if Y is convex, then $\diamond Y \wedge \neg Y$ consists of at most two points, $\diamond(\diamond Y \wedge \neg Y) = \diamond Y \wedge \neg Y$, and $\varrho(Y) = Y \wedge \diamond(\diamond Y \wedge \neg Y) = Y \wedge (\diamond Y \wedge \neg Y) = \emptyset$. Hence, $r(Y) = 0$.

Now let X be a serial subset of \mathbb{R} . Then $X = \bigvee_{i=1}^n X_i$, where every X_i is a convex subset of \mathbb{R} , and actually we can assume that all X_i are disjoint. But then $\varrho(X) = \bigvee_{i=1}^n \varrho(X_i) = \emptyset$, and so $r(X) = 0$. ■

It follows that if we interpret p as a serial subset of \mathbb{R} , then all φ_n from the previous section define sets equal to each other.

Call a valuation ν of our language \mathcal{L} to subsets of \mathbb{R} *serial* if $\nu(p) \in \mathcal{S}(\mathbb{R})$ for any $p \in \mathcal{P}$. Since $\mathcal{S}(\mathbb{R})$ is closed with respect to \neg, \wedge and \square , we have that $\nu(\varphi) \in \mathcal{S}(\mathbb{R})$ for any serial valuation ν . Call a formula φ *S-true* if it is true in \mathbb{R} under a serial valuation. Call φ *S-valid* if φ is *S-true* for any serial valuation on \mathbb{R} . Let $L(\mathcal{S}) = \{\varphi : \varphi \text{ is } \mathcal{S}\text{-valid}\}$.

FACT 5.3

$L(\mathcal{S})$ is a normal modal logic over **S4**.

Obviously all φ_n from the previous section are $L(\mathcal{S})$ -equivalent. So, it is natural to expect that there are only finitely many formulas in one variable which are $L(\mathcal{S})$ -nonequivalent, and indeed that $L(\mathcal{S})$ is a much stronger logic than **S4**.

As a first step in this direction, we show that the Grzegorzcyk axiom

$$\mathbf{Grz} = \Box(\Box(p \rightarrow \Box p) \rightarrow p) \rightarrow p$$

belongs to $L(\mathcal{S})$.

LEMMA 5.4

Grz is \mathcal{S} -valid.

PROOF. **Grz** is \mathcal{S} -valid iff $X \subseteq \Diamond(X \wedge \neg\Diamond(\Diamond X \wedge \neg X))$ for any $X \in \mathcal{S}(\mathbb{R})$. Suppose $X \in \mathcal{S}(\mathbb{R})$. Since $\Diamond X \wedge \neg X$ is finite, $\Diamond(\Diamond X \wedge \neg X) = \Diamond X \wedge \neg X$. Hence $\Diamond(X \wedge \neg\Diamond(\Diamond X \wedge \neg X)) = \Diamond(X \wedge \neg(\Diamond X \wedge \neg X)) = \Diamond(X \wedge \neg(\Diamond X \wedge \neg X)) = \Diamond(X \wedge (\neg\Diamond X \vee X)) = \Diamond X$, which clearly contains X . So, $X \subseteq \Diamond(X \wedge \neg\Diamond(\Diamond X \wedge \neg X))$. \blacksquare

As a next step, we show that the axioms

$$\mathbf{BD}_2 = (\neg p \wedge \Diamond p) \rightarrow \Diamond \Box p, \text{ and}$$

$$\mathbf{BW}_2 = \neg(p \wedge q \wedge \Diamond(p \wedge \neg q)) \wedge \Diamond(\neg p \wedge q) \wedge \Diamond(\neg p \wedge \neg q),$$

bounding the depth and the width of a Kripke frame to 2, are \mathcal{S} -valid.

LEMMA 5.5

BD₂ and **BW**₂ are \mathcal{S} -valid.

PROOF. Note that **BD**₂ is \mathcal{S} -valid iff $\Diamond X \wedge \neg X \subseteq \Diamond \Box X$ for any $X \in \mathcal{S}(\mathbb{R})$, and that **BW**₂ is \mathcal{S} -valid iff $X \wedge Y \wedge \Diamond(X \wedge \neg Y) \wedge \Diamond(Y \wedge \neg X) \wedge \Diamond(\neg X \wedge \neg Y) = \emptyset$ for any $X, Y \in \mathcal{S}(\mathbb{R})$.

To show that $\Diamond X \wedge \neg X \subseteq \Diamond \Box X$ for any $X \in \mathcal{S}(\mathbb{R})$, suppose $x \in \Diamond X \wedge \neg X$. Then x is a limit point of X not belonging to X . Since X is serial, there is $y \in \mathbb{R}$ such that either $y < x$ and $(y, x) \subseteq X$, or $x < y$ and $(x, y) \subseteq X$. In both cases it is obvious that $x \in \Diamond \Box X$. So, $\Diamond X \wedge \neg X \subseteq \Diamond \Box X$.

To show that $X \wedge Y \wedge \Diamond(X \wedge \neg Y) \wedge \Diamond(Y \wedge \neg X) \wedge \Diamond(\neg X \wedge \neg Y) = \emptyset$ for any $X, Y \in \mathcal{S}(\mathbb{R})$, suppose $x \in X \wedge Y \wedge \Diamond(X \wedge \neg Y) \wedge \Diamond(Y \wedge \neg X)$. Then $x \notin \Box X$ and $x \notin \Box Y$. Hence, there exist $y, z \in \mathbb{R}$ such that $y < x < z$ and $(y, z) \cap (\neg X \wedge \neg Y) = \emptyset$, which means that $x \notin \Diamond(\neg X \wedge \neg Y)$. So, $X \wedge Y \wedge \Diamond(X \wedge \neg Y) \wedge \Diamond(Y \wedge \neg X) \wedge \Diamond(\neg X \wedge \neg Y) = \emptyset$. \blacksquare

The following is an immediate consequence of our observations.

COROLLARY 5.6

S4 \oplus **Grz** \oplus **BD**₂ \oplus **BW**₂ $\subseteq L(\mathcal{S})$.

In order to prove the converse, and hence complete our axiomatization of the logic of serial subsets of \mathbb{R} , observe that **S4** \oplus **Grz** \oplus **BD**₂ \oplus **BW**₂ is actually the complete modal logic of the following ‘2-fork’ Kripke frame $\langle W, R \rangle$, where $W = \{w_1, w_2, w_3\}$ and $w_1 R w_2, w_1 R w_3, w_2 R w_3, w_1 R w_2, w_1 R w_3$ (see Figure 8).

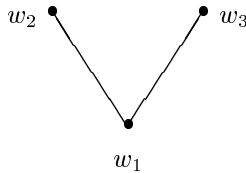


FIGURE 8. The 2-fork frame

Indeed, it is well known that **Grz** is valid on a Kripke frame iff it is a Noetherian partial order, that **BD**₂ is valid on a partially ordered Kripke frame iff its depth is bounded by 2, and that **BW**₂ is valid on a partially ordered Kripke frame of a depth ≤ 2 iff its width is bounded by 2. Now, denoting the logic of $\langle W, R \rangle$ by $L(\langle W, R \rangle)$, we have the following:

THEOREM 5.7

$$\mathbf{S4} \oplus \mathbf{Grz} \oplus \mathbf{BD}_2 \oplus \mathbf{BW}_2 = L(\langle W, R \rangle).$$

PROOF. Denote $\mathbf{S4} \oplus \mathbf{Grz} \oplus \mathbf{BD}_2 \oplus \mathbf{BW}_2$ by L . Obviously, $\langle W, R \rangle \models \mathbf{Grz}, \mathbf{BD}_2, \mathbf{BW}_2$. Hence, $L \subseteq L(\langle W, R \rangle)$. Conversely, since **Grz** is a theorem of L , every L -frame is a Noetherian partial order. Since **BD**₂ is a theorem of L , every L -frame is of depth ≤ 2 . Therefore, L has the f.m.p., and thus is complete with respect to finite rooted partially ordered Kripke frames of depth ≤ 2 . Since **BW**₂ is a theorem of L , it is obvious that the width of finite rooted L -frames is also ≤ 2 . But then it is routine to check that every such rooted frame is a p -morphic image of $\langle W, R \rangle$. Hence, $L(\langle W, R \rangle) \subseteq L$, and so $L = L(\langle W, R \rangle)$. ■

As a final move, we show that $\langle W, \tau_R \rangle$ is a continuous and open *serial* image of \mathbb{R} , meaning that there is a continuous and open map $f : \mathbb{R} \rightarrow W$ such that $f^{-1}(X) \in \mathcal{S}(\mathbb{R})$ for any subset X of W .

Recall that τ_R consists of the upward closed subsets of W , which obviously are \emptyset , $\{w_2\}$, $\{w_3\}$, $\{w_2, w_3\}$, and W . Fix any $x \in \mathbb{R}$ and define $f : \mathbb{R} \rightarrow W$ by putting

$$f(y) = \begin{cases} w_1 & \text{for } y = x, \\ w_2 & \text{for } y < x, \\ w_3 & \text{for } y > x. \end{cases}$$

Then it is routine to check that $f^{-1}(\emptyset) = \emptyset$, $f^{-1}(\{w_2\}) = (-\infty, x)$, $f^{-1}(\{w_3\}) = (x, +\infty)$, $f^{-1}(\{w_2, w_3\}) = (-\infty, x) \cup (x, +\infty)$, and $f^{-1}(W) = \mathbb{R}$. So, f is continuous. Moreover, for any open subset U of \mathbb{R} , if $x \in U$, then $f(U) = W$; and if $x \notin U$, then $f(U) \subseteq \{w_2, w_3\}$, which is always open. Hence, f is open. Furthermore, from the definition of f it follows that the f -inverse image of any subset of W is a serial subset of \mathbb{R} . So, $\langle W, \tau_R \rangle$ is a continuous and open serial image of \mathbb{R} .

As a trivial consequence of this observation, we obtain that for every valuation ν on $\langle W, R \rangle$ there is a serial valuation ν_S on \mathbb{R} such that $\langle W, R, \nu \rangle$ is topo-bisimilar to $\langle \mathbb{R}, \nu_S \rangle$. Hence, every nontheorem of $L(\langle W, R \rangle)$ is a nontheorem of $L(S)$, and we have the following:

COROLLARY 5.8

$$L(S) \subseteq L(\langle W, R \rangle).$$

Combining Corollaries 5.6 and 5.8 and Theorem 5.7, we obtain:

THEOREM 5.9

$$L(S) = L(\langle W, R \rangle) = \mathbf{S4} \oplus \mathbf{Grz} \oplus \mathbf{BD}_2 \oplus \mathbf{BW}_2.$$

5.2 Formulas in one variable over serial sets

This section provides some more concrete information about definability of serial sets. As $L(S)$ is the logic of the finite ‘2-fork’ frame, for every natural number $n \geq 0$, there are only finitely many $L(S)$ -nonequivalent formulas built from the variables p_1, \dots, p_n . In this section, we show that there are exactly 64 $L(S)$ -nonequivalent formulas in one variable, and describe them all.

THEOREM 5.10

Every formula in one variable is $L(\mathcal{S})$ -equivalent to a disjunction of the following six formulas:

- $\Box p$,
- $\Box \neg p$,
- $p \wedge \Box \Diamond \neg p$,
- $\neg p \wedge \Box \Diamond p$,
- $p \wedge \Diamond \Box \neg p \wedge \Diamond \Box p$, and
- $\neg p \wedge \Diamond \Box p \wedge \Diamond \Box \neg p$.

Hence, there are exactly 64 $L(\mathcal{S})$ -nonequivalent formulas in one variable.

PROOF. In line with our interest in tying up ‘modal’ and ‘topological’ ways of thinking, we will give two different proofs of this result. One proceeds by constructing the 1-universal Kripke model of $L(\mathcal{S})$, which is a standard technique in modal logic; the other is purely topological, using some obvious observations on serial subsets of \mathbb{R} .

First Proof. Since $L(\mathcal{S})$ is the logic of the ‘2-fork’ frame, we can easily construct the 1-universal Kripke model $\langle W_1, R_1, \nu_1 \rangle$ of $L(\mathcal{S})$, where $w_n \in \nu_1(p)$ iff n is even (see Figure 9). Now one can readily check that each point of W_1 corresponds to one of the six formulas in the condition of the theorem. Hence, every formula in one variable is $L(\mathcal{S})$ -equivalent to a disjunction of the above six formulas. Since there are exactly 2^6 different subsets of W_1 , we obtain that there are exactly 64 $L(\mathcal{S})$ -nonequivalent formulas in one variable.

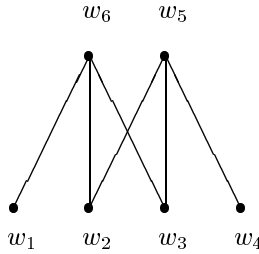


FIGURE 9. $\langle W_1, R_1, \nu_1 \rangle$

Second Proof. Observe that there exists a serial subset X of \mathbb{R} such that $\Box X \neq \Box \neg X \neq X \wedge \Box \Diamond \neg X \neq \neg X \wedge \Box \Diamond X \neq X \wedge \Diamond \Box \neg X \wedge \Diamond \Box X \neq \neg X \wedge \Diamond \Box X \wedge \Diamond \Box \neg X$. For example, let $x < y < z < u$, and take $X = [x, y) \cup (y, z) \cup \{u\}$. Then one can readily check that

- $\Box X = (x, y) \cup (y, z)$,
- $\Box \neg X = (-\infty, x) \cup (z, u) \cup (u, +\infty)$,
- $X \wedge \Box \Diamond \neg X = \{u\}$,
- $\neg X \wedge \Box \Diamond X = \{y\}$,
- $X \wedge \Diamond \Box \neg X \wedge \Diamond \Box X = \{x\}$, and
- $\neg X \wedge \Diamond \Box X \wedge \Diamond \Box \neg X = \{z\}$.

Hence, we can always interpret p as a serial subset of \mathbb{R} such that all the six formulas of the theorem correspond to different serial subsets of \mathbb{R} .

Now, let us prove that every subset of \mathbb{R} obtained by repeatedly applying \neg, \wedge, \square to a serial set X is equal to a finite (including the empty) union of the following serial subsets:

$$\begin{aligned} T_1 &= \square X, \\ T_2 &= \square \neg X, \\ T_3 &= X \wedge \square \diamond \neg X, \\ T_4 &= \neg X \wedge \square \diamond X, \\ T_5 &= X \wedge \diamond \square \neg X \wedge \diamond \square X, \text{ and} \\ T_6 &= \neg X \wedge \diamond \square X \wedge \diamond \square \neg X. \end{aligned}$$

For this first observe that $T_i \wedge T_j = \emptyset$ if $i \neq j$, and that $\bigvee_{i=1}^6 T_i = \mathbb{R}$. So, these six serial subsets of \mathbb{R} are mutually disjoint and jointly exhaustive. Next observe that $\neg T_i = T_j \vee T_k \vee T_l \vee T_m \vee T_n$, where $i, j, k, l, m, n \in \{1, 2, 3, 4, 5, 6\}$ are different from each other. Finally observe that $\square T_1 = T_1$, $\square T_2 = T_2$, and $\square T_3 = \square T_4 = \square T_5 = \square T_6 = \emptyset$.

Hence, every subset of \mathbb{R} obtained by repeatedly applying \neg, \wedge, \square to $\{T_1, \dots, T_6\}$ is a finite (including the empty) union of $\{T_1, \dots, T_6\}$.

Now suppose $Y \subseteq \mathbb{R}$ is obtained by repeatedly applying \neg, \wedge, \square to X . We prove by induction on the complexity of Y that Y is equal to a finite (including the empty) union of $\{T_1, \dots, T_6\}$.

Base case. Since $X = T_1 \vee T_3 \vee T_5$ (and $\neg X = T_2 \vee T_4 \vee T_6$), the base case (that is when $Y = X$) is obvious.

Complement. Suppose $Y = \neg Z$ and $Z = T_{i_1} \vee \dots \vee T_{i_k}$, where $i_1, \dots, i_k \in \{1, \dots, 6\}$. Then $Y = \neg(T_{i_1} \vee \dots \vee T_{i_k}) = \neg T_{i_1} \wedge \dots \wedge \neg T_{i_k}$. Since every $\neg T_{i_j}$ is equal to $\bigvee_{i_s \neq i_j} T_{i_s}$, using the distributivity law we obtain that $Y = \bigvee_{i_s, i_t \in \{1, \dots, 6\}} (T_{i_s} \wedge T_{i_t})$. Since for different i_s and i_t , $T_{i_s} \wedge T_{i_t} = \emptyset$, which is the empty union of T_{i_s} s, we finally obtain that Y is a finite union of $\{T_1, \dots, T_6\}$.

Intersection. Suppose $Y = Z_1 \wedge Z_2$, $Z_1 = T_{i_1} \vee \dots \vee T_{i_k}$ and $Z_2 = T_{j_1} \vee \dots \vee T_{j_m}$, where $i_1, \dots, i_k, j_1, \dots, j_m \in \{1, \dots, 6\}$. Similarly to the above case, using the distributivity law we obtain that Y is a finite union of $\{T_1, \dots, T_6\}$.

Interior. Suppose $Y = \square Z$ and $Z = T_{i_1} \vee \dots \vee T_{i_k}$, where $i_1, \dots, i_k \in \{1, \dots, 6\}$. Since T_{i_s} are mutually disjoint, $Y = \square T_{i_1} \vee \dots \vee \square T_{i_k}$. Now since $\{T_1, \dots, T_6\}$ is closed with respect to \square , we obtain that Y is a finite union of $\{T_1, \dots, T_6\}$.

Hence, every subset of \mathbb{R} obtained by repeatedly applying \neg, \wedge, \square to a serial set X is equal to a finite (including the empty) union of $\{T_1, \dots, T_6\}$. Since there are exactly 2^6 different subsets obtained as a union of $\{T_1, \dots, T_6\}$, we obtain that there are exactly 64 different subsets of \mathbb{R} obtained by repeatedly applying \neg, \wedge, \square to a serial set X . This directly implies that there are exactly 64 $L(S)$ -nonequivalent formulas in one variable. ■

The same technique can also be used to prove the normal form theorem over $L(S)$ for every formula with more than one propositional variable.

5.3 Countable unions of convex sets on the real line

Let us now be a bit more systematic. By Theorem 4.15, **S4** is the complete logic of \mathbb{R} , and hence sets of reals suffice as values $\nu(p)$ in refuting nontheorems. But how complex must these sets be? In first-order logic, for example, we know that completeness requires

atomic predicates over the integers which are at least Δ_2^0 . With only simpler predicates in the arithmetic hierarchy, the logic gets richer. In a topological space like \mathbb{R} , it seems reasonable to look at the Borel Hierarchy \mathfrak{B} . How high up do we have to go for our **S4**-counterexamples? One could analyse our construction in Section 4.3 to have an upper bound. But here we will state some more direct information.

Consider the set τ of all open subsets of \mathbb{R} . Let $\mathcal{B}(\tau)$ denote the Boolean closure of τ . Since $\mathcal{B}(\tau)$ contains all open subsets of \mathbb{R} , it is obvious that $\mathcal{B}(\tau)$ is closed with respect to \sqcap . Obviously $\mathcal{S}(\mathbb{R})$ is properly contained in $\mathcal{B}(\tau)$. It is natural to ask whether the elements of $\mathcal{B}(\tau)$ are enough for refuting all the nontheorems of **S4**. The answer is negative, the modal logic is still richer.

FACT 5.11

[10] The complete logic of $\mathcal{B}(\tau)$ is **Grz**.

Hence, we need to seek something bigger than $\mathcal{B}(\tau)$. Let $\mathcal{C}^\infty(\mathbb{R})$ denote the set of countable unions of convex subsets of \mathbb{R} . Since every open subset of \mathbb{R} is a countable union of open intervals, it is obvious that $\tau \subseteq \mathcal{C}^\infty(\mathbb{R})$. Let $\mathcal{B}(\mathcal{C}^\infty(\mathbb{R}))$ denote the Boolean closure of $\mathcal{C}^\infty(\mathbb{R})$. Since $\tau \subseteq \mathcal{C}^\infty(\mathbb{R})$, we also have $\mathcal{B}(\tau) \subseteq \mathcal{B}(\mathcal{C}^\infty(\mathbb{R}))$. It follows that $\mathcal{B}(\mathcal{C}^\infty(\mathbb{R}))$ is also closed with respect to \sqcap . Moreover, $\mathcal{B}(\tau)$ is properly contained in $\mathcal{B}(\mathcal{C}^\infty(\mathbb{R}))$, since the set \mathbf{Q} of rationals belongs to $\mathcal{B}(\mathcal{C}^\infty(\mathbb{R}))$ but does not belong to $\mathcal{B}(\tau)$.

THEOREM 5.12

[10] **S4** is the complete logic of $\mathcal{B}(\mathcal{C}^\infty(\mathbb{R}))$.

So, the Boolean combinations of countable unions of convex subsets of \mathbb{R} are all we need for refuting the nontheorems of **S4**. Since every countable union of convex subsets of \mathbb{R} belongs to the Borel hierarchy \mathfrak{B}_2 over the opens of \mathbb{R} , a very low level of the Borel hierarchy suffices for refuting nontheorems of **S4**. So, \mathfrak{B} itself is more than sufficient for refuting nontheorems of **S4**.

Summarizing, we constructed five Boolean algebras of subsets of \mathbb{R} forming a chain under inclusion: $\mathcal{S}(\mathbb{R}) \subset \mathcal{B}(\tau) \subset \mathcal{B}(\mathcal{C}^\infty(\mathbb{R})) \subset \mathfrak{B} \subset \mathcal{P}(\mathbb{R})$, where $\mathcal{S}(\mathbb{R})$ is the Boolean algebra of all serial subsets of \mathbb{R} , $\mathcal{B}(\tau)$ is the Boolean closure of the set of all open subsets of \mathbb{R} , $\mathcal{B}(\mathcal{C}^\infty(\mathbb{R}))$ is the Boolean closure of the set of all countable unions of convex subsets of \mathbb{R} , \mathfrak{B} is the Boolean algebra of all Borel subsets of \mathbb{R} over τ , and $\mathcal{P}(\mathbb{R})$ is the power set of \mathbb{R} . All of these Boolean algebras are closed with respect to \sqcap . The modal logic of the last three algebras is **S4**, that of the second one is **Grz**, and the modal logic of the first is the logic of the ‘2-fork’ Kripke frame.

5.4 Generalization to \mathbb{R}^n

We conclude this section by sketching the generalization of our results on the serial subsets of \mathbb{R} to sufficiently well-behaved ‘chequered subsets’ of a Euclidean space \mathbb{R}^n .

A set $X \subseteq \mathbb{R}^n$ is *convex* if all points lying in between any two points of X belong to X . It is said to be *serial* if X is a finite union of convex subsets of \mathbb{R}^n . Denote the set of all serial subsets of \mathbb{R}^n by $\mathcal{S}(\mathbb{R}^n)$. The main difference between \mathbb{R} and \mathbb{R}^n is that, unlike $\mathcal{S}(\mathbb{R})$, $\mathcal{S}(\mathbb{R}^n)$ is not closed with respect to the set-theoretical complement. For instance, a full circle is obviously a convex subset of \mathbb{R}^2 . However, its complement is not serial.

One natural way of overcoming this difficulty is to work with a smaller family of chequered subsets of \mathbb{R}^n , which also has a reasonable claim of being ‘the n -dimensional generalization of one-dimensional serial sets’.

A set $X \subseteq \mathbb{R}^n$ is said to be a *hyper-rectangular convex* if $X = X_1 \times \dots \times X_n$, where all the X_i are convex subsets of \mathbb{R} [7]. It is easy to see that every hyper-rectangular convex is a convex set in the usual sense, but not vice versa. For example, in the real plane a circle is obviously convex, but it is not a rectangular convex.

A set $X \subseteq \mathbb{R}^n$ is said to be *n-chequered* if it is a finite union of hyper-rectangular convex subsets of \mathbb{R}^n . Denote the set of all *n-chequered* subsets of \mathbb{R}^n by $\mathcal{CH}(\mathbb{R}^n)$. Obviously $\mathcal{CH}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$.

Note that unlike $\mathcal{S}(\mathbb{R}^n)$, $\mathcal{CH}(\mathbb{R}^n)$ does form a Boolean algebra. Moreover, $\Box X \in \mathcal{CH}(\mathbb{R}^n)$ for any $X \in \mathcal{CH}(\mathbb{R}^n)$. Hence, interpreting propositional variables as *n-chequered* subsets of \mathbb{R}^n , every formula of our language \mathcal{L} will also be interpreted as a *n-chequered* subset of \mathbb{R}^n .

This approach leads to a logic, which we sketch here. We call a valuation ν of \mathcal{L} to subsets of \mathbb{R}^n *n-chequered* if $\nu(p) \in \mathcal{CH}(\mathbb{R}^n)$ for any $p \in \mathcal{P}$. Since $\mathcal{CH}(\mathbb{R}^n)$ is closed with respect to \neg, \wedge and \Box , we have that $\nu(\varphi) \in \mathcal{CH}(\mathbb{R}^n)$ for any *n-chequered* interpretation ν . We call a formula φ *n-true* if it is true in \mathbb{R}^n under a *n-chequered* valuation. We call φ *n-valid* if φ is *n-true* for any *n-chequered* valuation on \mathbb{R}^n . Let $L_n = \{\varphi : \varphi \text{ is } n\text{-valid}\}$. Denote by $\langle W, R \rangle^n$ the *n*-times Cartesian product of the ‘2-fork’ frame $\langle W, R \rangle$ on itself. Let $L(\langle W, R \rangle^n)$ denote the logic of $\langle W, R \rangle^n$. A complete characterization of L_n is given by the following:

THEOREM 5.13
 $L_n = L(\langle W, R \rangle^n)$.

PROOF. See [9]. ■

In particular, the logic of chequered subsets of the real plane coincides with the logic $L(\langle W, R \rangle^2)$. An illustration of $\langle W, R \rangle^2$ is given in Figure 10. For interesting alternative approaches for the plane consult [22, 4].

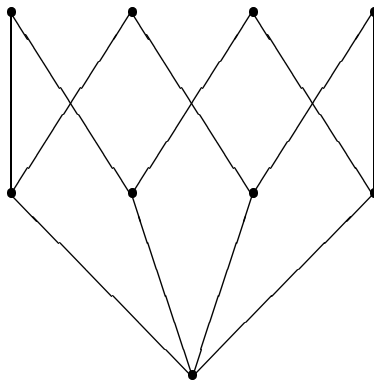


FIGURE 10. $\langle W, R \rangle^2$

6 A general picture

6.1 The deductive landscape

The logics that we have studied in this paper fit into a more general environment. Typical for modal logic is its lattice of deductive systems such as **K**, **S4**, **S5** or **GL**. These form a large family describing different classes of Kripke frames, with often very different motivations (cf., for instance, [18, 34]). Among the uncountably many modal logics, a small number is distinguished for one of two reasons. Logics like **S4** or **S5** were originally proposed as syntactic proof theories for notions of modality, and then turned out to be semantically complete with respect to natural frame classes, such as quasi-orders for **S4**. Other modal logics, however, were discovered as the complete theories of important frames, such as the natural numbers with their standard ordering. What about a similar landscape of modal logics for the topological interpretation?

Some well-known modal logics extending **S4** indeed correspond to natural classes of topological spaces. For example, it is easy to see that the ‘identity logic’ with the axiom $p \rightarrow \Box p$ axiomatizes the complete logic of all *discrete spaces*. And it also defines them semantically through the usual notion of *frame correspondence*, which can be lifted to the topological semantics in a straightforward manner. But already **S5** corresponds to a less standard condition, namely that every point x has an open neighbourhood all of whose points have x in all their open neighbourhoods. (Alternatively, this says that every open set is closed.) Also, even rich topological spaces do not seem to validate very spectacular modal logics, witness the fact that \mathbb{R} has just **S4** for its modal theory. We did find stronger logics with ‘general frames’ though, i.e. frames with a designated interior algebra of subsets, such as \mathbb{R} with the serial sets.

A related question is what becomes of the known general results on completeness and correspondence for modal logic in the topological setting. There appear to be some obstacles here. For example, the substitution method for Sahlqvist correspondence (cf. [11]) has only a limited range. It does work for axioms like the above $p \rightarrow \Box p$, where it automatically generates the corresponding first-order condition

$$(\forall x)(\exists U \in \tau)(x \in U \ \& \ (\forall y \in U)(y = x)),$$

i.e. discreteness. Likewise, it works for the **S5** axiom $p \rightarrow \Box \Diamond p$, where it produces the above-mentioned

$$(\forall x)(\exists U \in \tau)(x \in U \ \& \ (\forall y \in U)(\forall V \in \tau)(y \in V \rightarrow x \in V)).$$

The method also works for antecedents of the form $\Box p$, but things stop with antecedents like $\Diamond p$ or $\Box \Box p$. The reason is that, in the topological semantics, one modality \Box expresses a *two-quantifier* combination

$$\exists U \in \tau \text{ such that } \forall x \in U,$$

so that syntactic complexity builds up more rapidly than in standard modal logic, where each modality is one quantifier over relational successors of the current world. General correspondence or completeness results for topological modal logics therefore seem harder to obtain, and we may need different syntactic notions for them (see [16] for recent results in this direction).

6.2 The expressive landscape

In any case, the basic modal language seems too poor to express many properties of topological interest. A good example is *connectedness*, which cannot be modally defined. To see this, suppose there was a modal formula φ defining connected topological spaces in the sense of frame correspondence. Now consider the nonconnected discrete 2-element space with the universe $\{0, 1\}$. The formula φ must fail here under some valuation ν , say at point 0. Now consider the one-point model $\{0\}$ copying 0's valuation. The link between just the worlds 0 in the two models is a topo-bisimulation, as it is easy to see. But then, by modal invariance, φ would also fail in the connected one-point model: a contradiction.

Connectedness does have a definition in a modal language *extended* with a *universal modality* $U\varphi$ expressing that φ holds in all worlds of the model, and a dual existential modality $E\varphi$ expressing that φ holds in at least one world. This allows us to express connectedness in the following way [28, 1]:

$$U(\Diamond p \rightarrow \Box p) \rightarrow (Up \vee U\neg p).$$

This is one instance of a general trend in modal logic toward moderate expressive extensions of the base language. The $\{\Box, U\}$ language is a natural candidate, as it can formulate 'global facts' about topological spaces such as inclusion of one region into another. Many of our earlier techniques apply such as frame correspondence, bisimulation and related model constructions. (Cf. [1] for back-up to this section.) The general logic of this new language is known [6]. It is the system **S4+S5**, being **S4** for \Box , **S5** for U , plus the 'bridge axiom' $Up \rightarrow \Box p$. Moreover, according to [28], we have a natural extension of the McKinsey and Tarski theorem: the $\{\Box, U\}$ modal theory of \mathbb{R} , and indeed of any metric separable dense-in-itself space, is exactly **S4+S5** plus the given connectedness axiom. Thus, the concerns of this paper reproduce for richer modal languages, expressing more topological behaviour.

Indeed, modal languages can also have much stronger topological modalities, such as the following 'Until' operator generalizing two well-known notions from temporal logic:

x has an open neighbourhood all of whose interior points satisfy B while all its boundary points satisfy A.

And even further extensions are needed to deal with modal separation axioms, such as a space being Hausdorff, which requires even stronger 'modalities' definable in the monadic second-order language over topological spaces. One can then see the art of the field in choosing 'good fragments' out of this total language, admitting a good balance between expressive power and complexity.

Finally, the same modal methodology also extends to other similarity types. In particular, one can introduce geometrical structure. For example, the affine geometry of betweenness suggests a 'convexity modality' CA :

x lies in between two points that satisfy A.

This brings out differences between the spaces \mathbb{R}^n , as \mathbb{R} but no higher-dimensional \mathbb{R}^n satisfies the principle $CCA \leftrightarrow CA$. (A more extensive study of various modal languages for affine, projective, and metric geometries is made in [5, 31, 2].)

Thus, there is a lot of modal logic of space in between Tarski's work on the propositional language of topology [29, 20] and his work on the full first-order language of elementary geometry [30]. We see our work and much of the related literature we have cited as the beginnings of a systematic way of filling this gap.

Acknowledgements

We thank Yde Venema, Maarten Marx, and Nick Bezhanishvili of University of Amsterdam, Mai Gehrke and Patrick J. Morandi of New Mexico State University, and Mamuka Jibladze and Dimitri Patararaia of the Georgian Academy of Sciences for discussions. We also thank Ting Zhang of Stanford University for pointing out a flaw in the original proof of Lemma 4.13, as well as the referees for their comments which changed the original version of the paper.

References

- [1] M. Aiello and J. van Benthem. Logical patterns in space. In *Words, Proofs, and Diagrams*, D. Barker-Plummer, D. Beaver, J. van Benthem, and P. Scotto di Luzio, eds, pp. 5–25. CSLI, Stanford, 2002.
- [2] M. Aiello and J. van Benthem. A modal walk through space. *Journal of Applied Non-classical Logics*, **12**, 319–363, 2002.
- [3] J. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, **26**, 832–843, 1983.
- [4] P. Balbiani, J. Condotta, and L. Fariñas del Cerro. A model for reasoning about bidimensional temporal relations. In *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning (KR’98)*, A. G. Cohn, L. Schubert, and S. Shapiro, eds, pp. 124–130. Morgan Kaufmann, 1998.
- [5] Ph. Balbiani, L. Fariñas del Cerro, T. Tinchev, and D. Vakarelov. Modal logics for incidence geometries. *Journal of Logic and Computation*, **7**, 59–78, 1997.
- [6] B. Bennett. Modal logics for qualitative spatial reasoning. *Bulletin of the IGPL*, **3**, 1–22, 1995.
- [7] J. van Benthem. *The Logic of Time*, volume 156 of *Synthese Library*. Reidel, Dordrecht, 1983. [Revised and expanded, Kluwer, 1991].
- [8] J. van Benthem. *Exploring Logical Dynamics*, volume 156. CSLI Publications, Stanford — Cambridge University Press, 1996.
- [9] J. van Benthem, G. Bezhanishvili, and M. Gehrke. Euclidean hierarchy in modal logic. *Studia Logica*, 2003. To appear. Also available as Technical Report PP-2002-07, University of Amsterdam.
- [10] G. Bezhanishvili and M. Gehrke. A new proof of completeness of S4 with respect to the real line. Technical Report PP-2002-06, University of Amsterdam, 2002.
- [11] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, 2001.
- [12] B.F. Chellas. *Modal Logic: An Introduction*. Cambridge University Press, 1980.
- [13] R. Engelking. *General Topology*. Heldermann Verlag, 1989.
- [14] L. L. Esakia. Diagonal constructions, Löb’s formula and Cantor’s scattered spaces. In *Studies in Logic and Semantics*, pp. 128–143. Metsniereba, 1981. In Russian.
- [15] L. L. Esakia. Weak transitivity-restitution. In *Study in Logic*, Volume 8, pp. 244–254. Nauka, 2001. In Russian.
- [16] D. Gabelaia. *Modal Definability in Topology*. Master’s thesis, ILLC, University of Amsterdam, 2001.
- [17] R. Goldblatt. Diodorean modality in Minkowski space-time. *Studia Logica*, **39**, 219–236, 1980.
- [18] M. Kracht, M. de Rijke, H. Wansing, and M. Zakharyashev, eds. *Advances in Modal Logic*, volume 1. CSLI Publications, 1998.
- [19] O. Lemon and I. Pratt. On the incompleteness of modal logics of space: advancing complete modal logics of place. In *Advances in Modal Logics ’96*, M. Kracht, M. de Rijke, H. Wansing, and M. Zakharyashev, eds. CSLI and FoLLI, Stanford, 1996.
- [20] J. McKinsey and A. Tarski. The algebra of topology. *Annals of Mathematics*, **45**, 141–191, 1944.
- [21] G. Mints. A completeness proof for propositional S4 in Cantor Space. In *Logic at work: Essays dedicated to the memory of Helena Rasiowa*, E. Orłowska, ed. Physica-Verlag, Heidelberg, 1998.
- [22] I. Pratt and D. Schoop. Expressivity in polygonal, plane mereotopology. *Journal of Symbolic Logic*, **65**, 822–838, 2000.
- [23] D. Randell, Z. Cui, and A. G. Cohn. A spatial logic based on regions and connection. In *Proceedings of International Conference on Principles of Knowledge Representation and Reasoning (KR’92)*, pp. 165–176. San Mateo, 1992.
- [24] H. Rasiowa and R. Sikorski. *The Mathematics of Metamathematics*. Panstwowe Wydawnictwo Naukowe, 1963.

- [25] J. Renz and B. Nebel. On the complexity of qualitative spatial reasoning: a maximal tractable fragment of the region connection calculus. *Artificial Intelligence*, **108**, 69–123, 1999.
- [26] V. Shehtman. Modal logics of domains on the real plane. *Studia Logica*, **42**, 63–80, 1983.
- [27] V. Shehtman. Derived sets in Euclidean spaces and modal logic. Technical Report X-1990-05, University of Amsterdam, 1990.
- [28] V. Shehtman. ‘Everywhere’ and ‘Here’. *Journal of Applied Non-classical Logics*, **9**, 369–379, 1999.
- [29] A. Tarski. Der Aussagenkalkül und die Topologie. *Fund. Math.*, **31**, 103–134, 1938.
- [30] A. Tarski. What is elementary geometry? In *The Axiomatic Method, with Special Reference to Geometry and Physics*, L. Henkin and P. Suppes and A. Tarski, eds, pp. 16–29. North-Holland, 1959.
- [31] Y. Venema. Points, lines and diamonds: a two-sorted modal logic for projective planes. *Journal of Logic and Computation*, **9**, 601–621, 1999.
- [32] F. Wolter and M. Zakharyashev. Spatio-temporal representation and reasoning based on RCC-8. In *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning (KR2000)*, A. G. Cohn, F. Giunchiglia, and B. Selman, eds, pp. 3–14, 2000.
- [33] F. Wolter and M. Zakharyashev. Qualitative spatio-temporal representation and reasoning: a computational perspective. In *Exploring Artificial Intelligence in the New Millennium*. Morgan Kaufmann, 2002.
- [34] M. Zakharyashev, K. Segerberg, M. de Rijke, and H. Wansing, eds. *Advances in Modal Logic*, volume 2. CSLI Publications, 2000.

Received 25 October 2001