Modal logics for metric spaces: open problems

Agi Kurucz[1], Frank Wolter[2], and Michael Zakharyaschev[1]

[1]Department of Computer Science, King's College London,
Strand, London WC2R 2LS, U.K.
[2]Department of Computer Science, University of Liverpool,
Liverpool L69 7ZF, U.K.
emails: {kuag,mz}@dcs.kcl.ac.uk, frank@csc.liv.ac.uk

Modal logics and their models have been used to speak about and represent *topological spaces* since the 1940s [22, 23, 16, 17]. Examples include Tarski's programme of algebraisation of topology ("*of creating an algebraic apparatus for the treatment of portions of point-set topology*," to be more precise) which involved modal logic **S4** [17], and the use of the extension of **S4** with the universal modality (and its fragments) for spatial representation and reasoning; see, e.g., [5, 18, 7, 8, 1, 9] and references therein.

Metric spaces are even more important mathematical structures that are fundamental for many areas of mathematics and computer science (recent examples include classification in bioinformatics, linguistics, botany, etc. using various similarity measures). A natural research programme is then to find out to which extent modal-like formalisms can be useful for speaking about metric spaces. Such a programme was launched in 2000 [21, 13, 14].

The aim of this note is to attract attention to the most important open problems and new directions of research in this exciting and promising area.

1 Distance spaces

Recall that a *metric space* is a pair (Δ, d), where Δ is a nonempty set (of points) and d is a function from $\Delta \times \Delta$ into the set $\mathbb{R}^{\geq 0}$ (of non-negative real numbers) satisfying the following axioms

identity of indiscernibles:	$d(x, y) = 0 \quad \text{iff} \quad x = y,$	(1)
triangle inequality:	$d(x, z) \leq d(x, y) + d(y, z),$	(2)
symmetry:	$d(x, y) = d(y, x)$	(3)

for all $x, y, z \in \Delta$. The value $d(x, y)$ is called the *distance* from the point x to the point y. Given a metric space (Δ, d), a point $x \in \Delta$ and a nonempty $Y \subseteq \Delta$, define the *distance $d(x, Y)$ from x to Y* by taking

$$d(x, Y) = \inf\{d(x, y) \mid y \in Y\},$$

and put $d(y, \emptyset) = \infty$. The distance $d(X, Y)$ between two nonempty sets X and Y is

$$d(X, Y) = \inf\{d(x, y) \mid x \in X,\ y \in Y\}.$$

Although acceptable in many cases, the defined concept of metric space is not universally applicable to all interesting measures of distance between points, especially those used in everyday life. Consider, for instance, the following two examples:

(i) If $d(x, y)$ is the flight-time from x to y then, as we know it too well, d is not necessarily symmetric, even approximately (just take a plane from London to Tokyo and back).

(ii) Often we do not measure distances by means of real numbers but rather using more fuzzy notions such as 'short,' 'medium' and 'long.' To represent these measures we can, of course, take functions d from $\Delta \times \Delta$ into the subset $\{1, 2, 3\}$ of $\mathbb{R}^{\geq 0}$ and define *short* $:= 1$, *medium* $:= 2$, and *long* $:= 3$. So we can still regard these distances as real numbers. However, for measures of this type the triangle inequality (2) does not make sense (short plus short can still be short, but it can also be medium or long).

Spaces (Δ, d) satisfying only axiom (1) will be called *distance spaces*.

Recall also that a *topological space* is a pair $(U, \mathbb{I})$ in which U is a nonempty set, the *universe* of the space, and $\mathbb{I}$ is the *interior operator* on U satisfying the *Kuratowski axioms*: for all $X, Y \subseteq U$,

$$\mathbb{I}(X \cap Y) \;=\; \mathbb{I}X \cap \mathbb{I}Y, \quad \mathbb{I}X \;\subseteq\; \mathbb{I}\mathbb{I}X, \quad \mathbb{I}X \;\subseteq\; X \quad \text{and} \quad \mathbb{I}U \;=\; U.$$

The operator dual to $\mathbb{I}$ is called the *closure operator* and denoted by $\mathbb{C}$: for every $X \subseteq U$, we have $\mathbb{C}X = U - \mathbb{I}(U - X)$. Thus, $\mathbb{I}X$ is the *interior* of a set X, while $\mathbb{C}X$ is its *closure*. X is called *open* if $X = \mathbb{I}X$ and *closed* if $X = \mathbb{C}X$.

Each metric space (Δ, d) gives rise to the *interior operator* $\mathbb{I}_d$ on Δ: for all $X \subseteq \Delta$,

$$\mathbb{I}_d X \;=\; \{x \in X \mid \exists \varepsilon > 0 \; \forall y \; (d(x, y) < \varepsilon \to y \in X)\}.$$

The pair $(\Delta, \mathbb{I}_d)$ is called the *topological space induced by* the metric space (Δ, d). The dual *closure operator* $\mathbb{C}_d$ in this space can be defined by the equality

$$\mathbb{C}_d X \;=\; \{x \in W \mid \forall \varepsilon > 0 \; \exists y \in X \; d(x, y) < \varepsilon\}.$$

Examples. We briefly remind the reader of a few standard examples of metric and topological spaces that will be used in what follows.

1 The *one-dimensional Euclidean space* is the set of real numbers $\mathbb{R}$ equipped with the following metric on it

$$d_1(x, y) \;=\; |x - y|.$$

Let $X \subseteq \mathbb{R}$. A point $x \in \mathbb{R}$ is said to be *interior* in X if there is some $\varepsilon > 0$ such that the whole open interval $(x - \varepsilon, x + \varepsilon)$ belongs to X. The interior $\mathbb{I}X$ of X is defined then as the set of all interior points in X. It is not hard to check that $(\mathbb{R}, \mathbb{I})$ is the topological space induced by the Euclidean metric d_1. Open sets in $(\mathbb{R}, \mathbb{I})$ are (possibly infinite) unions of open intervals (a, b), where $a \leq b$. The closure of (a, b), for $a < b$, is the closed interval $[a, b]$, with the end points a and b being its boundary.

2 In the same manner one can define *n-dimensional Euclidean spaces* based on the universes $\mathbb{R}^n$ with the metric

$$d_n(x, y) \;=\; \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

(in the definition of interior points x one should take n-dimensional ε-neighbourhoods of x).

3 Further well-known examples are *metric spaces on graphs*: the distance between two nodes of a graph is defined as the length of the shortest path between them. Special cases are the *tree metric spaces*.

4 A topological space is called an *Aleksandrov space* [2] if arbitrary (not only finite) intersections of open sets are open. Aleksandrov spaces are closely related to *quasi-ordered sets*, that is, pairs $\mathfrak{G} = (V, R)$, where V is a nonempty set and R a transitive and reflexive relation on V. Every such quasi-order $\mathfrak{G}$ induces the interior operator $\mathbb{I}_{\mathfrak{G}}$ on V: for $X \subseteq V$,

$$\mathbb{I}_{\mathfrak{G}} X \;=\; \{x \in X \mid \forall y \in V \;(xRy \to y \in X)\}.$$

In other words, the open sets of the topological space $\mathfrak{T}_{\mathfrak{G}} = (V, \mathbb{I}_{\mathfrak{G}})$ are the *upward closed* (or *R-closed*) subsets of V. It is well-known (see, e.g., [6]) that $\mathfrak{T}_{\mathfrak{G}}$ is an Aleksandrov space and, conversely, every Aleksandrov space is induced by a quasi-order.

2 Modal logics of distance spaces

The intended *distance models* we would like to talk about with the help of modal-like formalisms are structures of the form

$$\mathfrak{I} \;=\; \left(\mathfrak{D}, \ell_1^{\mathfrak{I}}, \ell_2^{\mathfrak{I}}, \dots, p_1^{\mathfrak{I}}, p_2^{\mathfrak{I}}, \dots\right) \tag{4}$$

where $\mathfrak{D} = (\Delta, d)$ is a distance space, the $\ell_i^{\mathfrak{I}}$ are some elements (or *locations*) of Δ and the $p_i^{\mathfrak{I}}$ are subsets of Δ. Distance models with the underlying distance space being a metric space will be called *metric models*.

We divide our languages designed for talking about distance models into two groups: those without *quantification over distances* and those that do use (explicitly or implicitly) such quantification.

2.1 Logics without quantification over distances

We introduce the following 'parameterised modalities' or 'bounded quantifiers:'

- $\exists^{=a}$ meaning 'somewhere at distance a,'
- $\exists^{<a}$ meaning 'somewhere at distance $< a$,'
- $\exists^{>a}$ meaning 'somewhere at distance $> a$,' and
- $\exists^{<b}_{>a}$ meaning 'somewhere at distance d with $a < d < b$,'

where a and b are some numbers from $\mathbb{R}^{\geq 0}$ (or rather $\mathbb{Q}^{\geq 0}$ to avoid the problem of representing real numbers and keep the language countable). Then one can also define duals like $\forall^{<b}_{>a}$ meaning 'everywhere within distance d for $a < d < b$,' etc. (As the expressive completeness result below shows, once we restrict ourselves to the 'modal' paradigm, our choice of operators is rather natural.)

More precisely, given a distance model $\mathfrak{I}$ of the form (4), we interpret our operators as

$$\begin{aligned}(\exists^{=a}\tau)^{\mathfrak{I}} &= \{x \in \Delta \mid \exists y\ (d(x,y) = a\ \wedge\ y \in \tau^{\mathfrak{I}})\},\\ (\exists^{<a}\tau)^{\mathfrak{I}} &= \{x \in \Delta \mid \exists y\ (d(x,y) < a\ \wedge\ y \in \tau^{\mathfrak{I}})\},\\ (\exists^{>a}\tau)^{\mathfrak{I}} &= \{x \in \Delta \mid \exists y\ (d(x,y) > a\ \wedge\ y \in \tau^{\mathfrak{I}})\},\\ (\exists^{<b}_{>a}\tau)^{\mathfrak{I}} &= \{x \in \Delta \mid \exists y\ (a < d(x,y) < b\ \wedge\ y \in \tau^{\mathfrak{I}})\},\end{aligned}$$

where $\tau^{\mathfrak{I}} \subseteq \Delta$.

The full 'modal' language of distance spaces. The full language $\mathcal{MS}$ of distance spaces with the operators $\exists^{=a}$, $\exists^{<a}$, $\exists^{>a}$, $\exists^{<b}_{>a}$ (and their duals $\forall^{=a}$, $\forall^{<a}$, etc.) interpreted as defined above was introduced and analysed in [14]. Formally, the expressions of this language are defined as follows:

$$\tau \quad ::= \quad p_i \mid \{\ell_i\} \mid \neg\tau \mid \tau_1 \sqcap \tau_2 \mid \exists^{=a}\tau \mid \exists^{<a}\tau \mid \exists^{>a}\tau \mid \exists^{<b}_{>a}\tau, \tag{5}$$

where $a, b \in \mathbb{Q}^{\geq 0}$ with $a < b$, and the ℓ_i are *location constants* (or *nominals*) interpreted by singleton sets. As expressions of the form τ are interpreted as subsets of distance spaces, we will call them (*spatial*) *terms*. Given some terms, we allow the language to say some simple things about them by means of *formulas* that are defined as follows:

$$\varphi \quad ::= \quad \tau_1 \sqsubseteq \tau_2 \mid d(\ell_1, \ell_2) = a \mid d(\ell_1, \ell_2) < a \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2$$

(in particular, we can express $\ell_i \in \tau$ and $\tau_1 = \tau_2$). Formulas are interpreted in distance models as *true* or *false* in the natural way. Various *logics* in (fragments of) this language can be obtained by restricting the class of distance spaces underlying our models, say, to the class of metric spaces.

Below we summarise what is known about the language $\mathcal{MS}$ interpreted over various classes of distance spaces. Perhaps the most important result is the following *expressive completeness theorem* [14] which describes precisely how the modal language $\mathcal{MS}$ is related to first-order logic over *metric models*:

> *Over the class metric models, the language $\mathcal{MS}$ is expressively complete for (or has the same expressive power as) the two-variable fragment of first-order logic with countably many unary predicates and binary predicates of the form $d(x, y) < a$ and $d(x, y) = a$ for $a \in \mathbb{Q}^{\geq 0}$.*

A rather transparent axiomatisation of formulas of $\mathcal{MS}$ that are valid in metric models has been given in [12] using some 'Gabbay-style' rules known from hybrid logic. To give the reader some idea of the axiomatisation, we observe first that $\exists^{\leq a}$ and $\exists^{>a}$ behave like normal modal 'diamonds,' $\forall^{\leq a}\tau \sqcap \forall^{>a}\tau$ is the *universal modality* $\boxtimes$, while $\exists^{>0}\tau$ is the *difference operator* (so in fact, nominals are expressible in $\mathcal{MS}$). Typical 'non-modal' axioms look as follows:

$$\begin{aligned}\tau &\sqsubseteq \forall^{\leq a}\exists^{\leq a}\tau,\\ \exists^{\leq a}\forall^{>b}\tau &\sqsubseteq \forall^{>a+b}\tau.\end{aligned}$$

It is also shown in [12] that $\mathcal{MS}$ does not have the Craig interpolation property over metric models.

The satisfiability[1] problem for $\mathcal{MS}$ over metric models was proved to be *undecidable* in [14]. It turns out that the 'doughnut' operators $\exists^{\leq a}_{>0}$ and the propositional constants $\top$ and $\bot$ are already enough for obtaining undecidability: this relatively small fragment can 'enforce' the $\mathbb{N} \times \mathbb{N}$ grid using the 'punctured' centers of circles, and so we can encode in it the undecidable $\mathbb{N} \times \mathbb{N}$ tiling problem.

An important class of distance models consists of those that are based on the *one-dimensional Euclidean space* $\mathbb{R}$. The situation here can be understood by embedding into quantitative temporal logics: the full language $\mathcal{MS}$ without the operator $\exists^{=a}$ (but with operators $\exists^{\leq a}_{\geq b}$ for $a \neq b$) turns out to be *decidable* over $\mathbb{R}$. It is EXPSPACE-complete under the binary coding and PSPACE-complete under the unary coding of parameters [11, 3]. Note, however, that the fragment of $\mathcal{MS}$ with the operators $\exists^{=a}$ only is *undecidable* over $\mathbb{R}$; see [3].

These undecidability results have motivated the study of 'well-behaved' fragments of $\mathcal{MS}$ over various classes of models. In particular, two such 'reasonable' fragments have been discovered.

The ($\exists^{\leq a}$, $\exists^{>a}$)-fragment. The terms of this fragment are formed as follows:

$$\tau \quad ::= \quad p_i \;\mid\; \{\ell_i\} \;\mid\; \neg\tau \;\mid\; \tau_1 \sqcap \tau_2 \;\mid\; \exists^{\leq a}\tau \;\mid\; \exists^{>a}\tau.$$

This fragment (together with several others without the doughnut operators) turns out to be *decidable* over various classes of distance models (over *metric models*, in particular), and even has the finite model property with respect to intended models (e.g., a term is satisfiable in a metric model iff it is satisfiable in a finite metric model) [14]. The computational complexity of these satisfiability problems was proved to be in non-deterministic exponential time in [14]. Using a different, carefully crafted 'Fisher–Ladner closure,' one can actually prove EXPTIME-completeness of these problems, provided that the numerical parameters are coded in *unary* [26].

Problem 1. *What is the complexity of the satisfiability problem for the* ($\exists^{\leq a}$, $\exists^{>a}$)*-fragment over metric models under the* <u>*binary*</u> *coding of parameters?*

Note that the ($\exists^{\leq a}$, $\exists^{>a}$)-fragment is *undecidable* over models based on Euclidean spaces $\mathbb{R}^n$, for $n \geq 2$ [14, 25].

Hilbert-style axiomatisations for the ($\exists^{\leq a}$, $\exists^{>a}$)-fragment over arbitrary distance models (and several subclasses) are provided in [13]. Observe that the universal modality and the difference operator (and so nominals) are still expressible in this fragment.

The ($\exists^{\leq a}$, $\exists^{<a}$)-fragment. The terms of this fragment are formed as follows:

$$\tau \quad ::= \quad p_i \;\mid\; \{\ell_i\} \;\mid\; \neg\tau \;\mid\; \tau_1 \sqcap \tau_2 \;\mid\; \exists^{\leq a}\tau \;\mid\; \exists^{<a}\tau$$

(observe that the universal modality and nominals are no longer expressible, and so the $\{\ell_i\}$ are not just syntactic sugar). In many contexts—e.g., if we represent a similarity measure between objects of a certain domain by means of a metric—we may not need operators of the form $\exists^{>a}$. The ($\exists^{\leq a}$, $\exists^{<a}$)-fragment extended with the *universal* and *existential modalities* $\boxminus$ and $\diamondminus$ was considered in [24]. The satisfiability problem for this language over *metric models* is EXPTIME-complete even if the numerical parameters are coded in *binary*, and enjoys the

[1]Throughout, it does not matter whether we consider term or formula satisfiability.

finite model property in the same sense as above. The crucial observation in the proof of this result is that the logic turns out to be complete with respect to *tree metric spaces*, a feature not shared by the richer languages considered above. An intriguing fact is that the fragments with only strict operators $\exists^{<a}$ and only non-strict ones $\exists^{\leq a}$ behave similarly, which perhaps reflects our everyday life disregard of the borders. Note that using both these operators we can say that the distance between two sets p and q is precisely a:

$$(p \sqcap \exists^{\leq a} q \neq \bot) \quad \wedge \quad (p \sqcap \exists^{<a} q = \bot)$$

The $(\exists^{\leq a},\ \exists^{<a})$-fragment is PSPACE-complete over models based on $\mathbb{R}$ under binary coding of parameters; see [15]. However, for $n \geq 2$, the $(\exists^{\leq a},\ \exists^{<a})$-fragment becomes *undecidable* over $\mathbb{R}^n$, even without nominals [14, 25]. Thus, no interesting decidable fragment of $\mathcal{MS}$ over $\mathbb{R}^2$ is known so far. Interesting candidates which might be decidable are given in the next open problem:

Problem 2. *Is the language with operators $\exists^{<a}$ only decidable over $\mathbb{R}^2$? What about the fragment with operators $\exists^{\leq a}$?*

Besides the full $\mathbb{R}^2$, natural and useful spaces to consider are bounded subspaces like $[0, 1] \times [0, 1]$. It is to be noted that the undecidability proofs mentioned above do not go through in this case.

Problem 3. *Investigate the satisfiability problem for fragments of $\mathcal{MS}$ over bounded subspaces of $\mathbb{R}^n$ (such as $[0, 1] \times [0, 1]$).*

Of course, the choice of the two fragments of $\mathcal{MS}$ discussed above is rather *ad hoc*. There are many open questions along these lines:

Problem 4. *Give a complete classification of the fragments of $\mathcal{MS}$ over various classes of distance models with respect to their satisfiability and axiomatisation problems. Given a class $\mathcal{C}$ of distance models, are there natural 'maximal' decidable fragments of $\mathcal{MS}$ over $\mathcal{C}$? If so, what is their computational complexity (under unary and binary coding of parameters)?*

We conjecture that a natural candidate for a 'maximal' decidable fragment of $\mathcal{MS}$ over various classes is the language with the operators $\exists^{\leq a}$, $\exists^{<a}$, $\exists^{>a}$, and $\exists^{\geq a}$.

2.2 Logics with quantification over distances

The language $\mathcal{MS}$ does not allow any *quantification over distances*. In particular, we can neither reason about the topology induced by a metric space nor compare distances without fixing their absolute values. A natural extension $\mathcal{QMS}$ of $\mathcal{MS}$ with quantification over distances can be obtained by allowing *individual variables* $x, y, z, \dots$ over $\mathbb{R}^{>0}$ or $\mathbb{Q}^{>0}$ in distance operators as well as quantification over these variables. Formally, the $\mathcal{QMS}$-terms are defined by adding to (5) terms of the form $\exists x\, \tau$ and by allowing *variables* $x, y, z, \dots$ along with concrete parameters in the distance operators, for example,

$$\exists^{=x}\tau, \quad \exists^{<x}\tau, \quad \exists^{>x}\tau, \quad \exists^{<x}_{>y}\tau, \quad \exists^{<x}_{>0}\tau.$$

To interpret $\mathcal{QMS}$-terms in distance models of the form (4), we also need *assignments* $\mathfrak{a}$ of *positive* real numbers $\mathfrak{a}(x) \in \mathbb{R}^{>0}$ to the individual variables x.[2] Then we have

$$\begin{array}{rcl}
(\exists^{=x}\tau)^{\mathfrak{I},\mathfrak{a}} & = & (\exists^{=\mathfrak{a}(x)}\tau)^{\mathfrak{I}} \\
(\exists^{<x}\tau)^{\mathfrak{I},\mathfrak{a}} & = & (\exists^{<\mathfrak{a}(x)}\tau)^{\mathfrak{I}} \\
(\exists^{>x}\tau)^{\mathfrak{I},\mathfrak{a}} & = & (\exists^{>\mathfrak{a}(x)}\tau)^{\mathfrak{I}} \\
(\exists^{<x}_{>y}\tau)^{\mathfrak{I},\mathfrak{a}} & = & (\exists^{<\mathfrak{a}(x)}_{>\mathfrak{a}(y)}\tau)^{\mathfrak{I}} \\
(\exists x\, \tau)^{\mathfrak{I},\mathfrak{a}} & = & \bigcup\{\tau^{\mathfrak{I},\mathfrak{b}} \mid \mathfrak{b}(y) = \mathfrak{a}(y), \text{ for } y \neq x\}\ .
\end{array}$$

Not much is known about the expressive power of this language. We conjecture, in particular, that the following problem can be solved in a positive way:

Problem 5. *Is the language $\mathcal{QMS}$ expressively complete for the two-sorted first-order logic where one sort is over $\mathbb{R}^{>0}$ and the other over the metric space underlying a given metric model, with only <u>two</u> variables of the second sort being allowed?*

However, we do have a number of interesting results for some fragments of $\mathcal{QMS}$.

'Modal' languages of metric and topological spaces. The terms of the fragment $\mathcal{MT}$ of $\mathcal{QMS}$ can be formed as follows:

$$\tau \quad ::= \quad p_i \mid \neg\tau \mid \tau_1 \sqcap \tau_2 \mid \exists^{<a}\tau \mid \exists^{\leq a}\tau \mid \exists x\forall^{<x}\tau \mid \forall x\exists^{<x}\tau \mid \forall x\forall^{<x}\tau \mid \exists x\exists^{<x}\tau,$$

where $a \in \mathbb{Q}^{\geq 0}$. (Observe the similarities between these 'directly closed' terms and expressions of Computational Tree Logic $\mathcal{CTL}$.) It is not hard to see that by adding similar 'non-strict' operators like $\forall x\exists^{\leq x}$ and $\exists x\forall^{\leq x}$ we do not increase the expressive power of the language. So in fact, what we obtain this way is an extension of the *nominal-free* ($\exists^{\leq a}$, $\exists^{<a}$)-fragment of $\mathcal{MS}$ above with the *topological interior* and *closure operators*

$$\mathbf{I}\tau \;=\; \exists x\forall^{<x}\tau, \qquad \mathbf{C}\tau = \forall x\exists^{<x}\tau, \tag{6}$$

and the *universal* and *existential modalities*

$$\boxed{\forall}\tau \;=\; \forall x\forall^{<x}\tau, \qquad \diamondsuit\!\!\!\exists\,\tau \;=\; \exists x\exists^{<x}\tau,$$

The intended meanings of these terms of course only 'work' in *metric models*:

$$(\exists x\,\forall^{<x}p)^{\mathfrak{I}} \;=\; \bigcup_{a\in\mathbb{R}^{>0}} (\forall^{<a}p)^{\mathfrak{I}}, \qquad (\forall x\,\exists^{<x}p)^{\mathfrak{I}} \;=\; \bigcap_{a\in\mathbb{R}^{>0}} (\exists^{<a}p)^{\mathfrak{I}},$$

$$(\boxed{\forall}\tau)^{\mathfrak{I}} = \begin{cases} \Delta, & \text{if } \tau^{\mathfrak{I}} = \Delta, \\ \emptyset, & \text{otherwise,} \end{cases} \qquad (\diamondsuit\!\!\!\exists\,\tau)^{\mathfrak{I}} = \begin{cases} \Delta, & \text{if } \tau^{\mathfrak{I}} \neq \emptyset, \\ \emptyset, & \text{otherwise.} \end{cases}$$

The language $\mathcal{MT}$ over metric models can be also regarded as an extension of the modal logic $\mathbf{S4}_u$ of topological spaces with the metric operators $\exists^{<a}$ and $\exists^{\leq a}$. Such a view was taken in [25] where this logic was first introduced and investigated.

[2]We quantify over *positive* real numbers rather than *non-negative* ones in order to obtain short and transparent definitions of standard topological operators; see (6). The expressivity of the language does not depend on this assumption.

Note first that this logic does not have the finite model property with respect to metric models because the topology induced by a finite metric space is trivial. For example, the term $p \sqcap \mathbf{C}\neg p$ is not satisfiable in any finite metric model, yet is satisfiable in every Euclidean space. Moreover, the logic in question is not compact in the sense that there is an infinite set Γ of terms such that, for every finite $\Gamma' \subseteq \Gamma$, there exists a model $\mathfrak{I}$ with $\bigcap_{\tau\in\Gamma'} \tau^{\mathfrak{I}} \neq \emptyset$, but there exists no model $\mathfrak{I}$ for which $\bigcap_{\tau\in\Gamma} \tau^{\mathfrak{I}} \neq \emptyset$. An example is given by the set of terms $\{\neg\mathbf{C}p\} \cup \{\exists^{<\frac{1}{n}} p \mid n \in \mathbb{N}^+\}$.

It turns out, however, that the intended metric models for this logic can be represented in the form of relational structures *à la* Kripke frames, which can be regarded as partial descriptions of scenarios that can be realised in metric models. This representation theorem—in fact, a generalisation of the McKinsey–Tarski [17] representation theorem for topological spaces—reduces reasoning with almost always infinite metric models to reasoning with finite relational models, which can be shown to be EXPTIME-complete even for the binary coding of the numerical parameters.

The formulas of $\mathcal{MT}$ that are valid in metric models can be axiomatised in a natural way (bearing in mind that both distance and topological operators are in fact normal modalities): we have the **S4**-axioms for **I** and **C**, standard axioms for $\forall^{<a}$ and $\exists^{<a}$ reflecting, in particular, the triangle inequality

$$\begin{aligned} \tau &\sqsubseteq \forall^{<a}\exists^{<a}\tau, \\ \exists^{\leq a}\exists^{\leq b}\tau &\sqsubseteq \exists^{\leq a+b}\tau, \\ \text{etc.} \end{aligned}$$

and only two axioms connecting metric and topology

$$\begin{aligned} \mathbf{C}\tau &\sqsubseteq \exists^{<a}\tau, \\ \exists^{<a}\mathbf{C}\tau &\sqsubseteq \exists^{<a}\tau, \end{aligned}$$

see [25] for more details.

Problem 6. *Investigate axiomatisation and satisfiability problems for $\mathcal{MT}$ over metric spaces whose induced topological spaces are* <u>*connected*</u>.

We conjecture that this logic can be axiomatised by adding the connectivity axiom

$$\diamondsuit\mathbf{I}p \sqcap \diamondsuit\mathbf{I}q \sqcap \boxdot(\mathbf{I}p \sqcup \mathbf{I}q) \quad \sqsubseteq \quad \diamondsuit(\mathbf{I}p \sqcap \mathbf{I}q)$$

to the axioms over arbitrary metric models, and that results on the satisfiability problem are similar to those for the arbitrary metric case.

Problem 7. *Investigate axiomatisation and satisfiability problems for $\mathcal{MT}$ over other interesting classes of metric and topological metric spaces.*

Problem 8. *What happens if we extend $\mathbf{S4}_u$ with the operators $\exists^{\leq a}$ and $\exists^{>a}$?*

Problem 9. *What happens if we extend $\mathcal{MT}$ with nominals?*

The satisfiability problem for $\mathcal{MT}$ over the Euclidean space $\mathbb{R}$ is *decidable*; see [11]. It becomes undecidable over models based on $\mathbb{R}^2$ (or over its various subspaces), as it contains the undecidable ($\exists^{\leq a}$, $\exists^{<a}$)-fragment of $\mathcal{MS}$, see above. However, the following questions are open:

Problem 10. *Is there a transparent axiomatisation of $\mathcal{MT}$ over $\mathbb{R}$? What is the computational complexity of the satisfiability problem?*

Problem 11. *Is the satisfiability problem for $\mathcal{MT}$ over $\mathbb{R}^2$ (or its subspaces) recursively enumerable? What happens if we omit the operators $\exists^{\leq a}$?*

The language of comparative similarity. We can be a bit more 'liberal' regarding the quantifier patterns of $\mathcal{MT}$ and (similarly to Computational Tree Logic $\mathcal{CTL}^+$) only require that the operators $\exists^{<x}$ and $\exists^{\leq x}$ cannot occur nested without an $\exists x$ in between. We then end up with the *similarity language* $\mathcal{SL}$ containing the '*closer operator*'

$$\tau_1 \leftleftarrows \tau_2 \;=\; \exists x\,(\exists^{\leq x}\tau_1 \sqcap \neg\exists^{\leq x}\tau_2).$$

Its semantical meaning in distance models of the form (4) is defined as follows:

$$(\tau_1 \leftleftarrows \tau_2)^{\mathfrak{I}} \;=\; \{x \in \Delta \mid d(x, \tau_1^{\mathfrak{I}}) < d(x, \tau_2^{\mathfrak{I}})\}\,. \tag{7}$$

In other words, $\tau_1 \leftleftarrows \tau_2$ is (interpreted by) the set containing those objects of Δ that are 'closer' (or 'more similar') to τ_1 than to τ_2.

This allows us to represent and reason about predicates like 'X is closer to Y than it is to Z' which are quite common in our everyday life ('the body was in the middle of the room, rather closer to the door than to the window').

The closer operator itself turns out to be quite powerful. Using it we can express (in metric spaces) the interior (and so the closure) operator by taking

$$\mathbf{I}\tau \;=\; \top \leftleftarrows \neg\tau.$$

Indeed, by the definition above, we have

$$(\mathbf{I}\tau)^{\mathfrak{I}} \;=\; \{x \in \Delta \mid d(x, \Delta - \tau^{\mathfrak{I}}) > 0\}.$$

We can also express the existential (and so the universal) modality:

$$\Diamond\tau \;=\; \tau \leftleftarrows \bot$$

because $d(x, \emptyset) = \infty$. Thus, the similarity language having the sole closer operator interpreted in metric models results in a logic that contains full $\mathbf{S4}_u$, and can again be regarded as a qualitative spatial formalism for reasoning about metric spaces with their induced topologies. We call it the *language of comparative similarity* and denote by $\mathcal{CSL}$.

One more interesting operator is

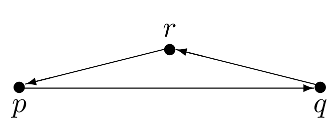
$$\tau_1 \leftrightarrows \tau_2 \;=\; \neg(\tau_1 \leftleftarrows \tau_2) \sqcap \neg(\tau_2 \leftleftarrows \tau_1)$$

which defines the set of points located at the same distance from τ_1 and τ_2.

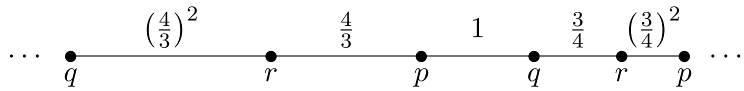
As a small illustrating example consider the formula

$$p \sqsubseteq (q \leftleftarrows r) \quad\wedge\quad q \sqsubseteq (r \leftleftarrows p) \quad\wedge\quad r \sqsubseteq (p \leftleftarrows q) \quad\wedge\quad p \neq \bot. \tag{8}$$

One can readily check that it is satisfiable in a three-point non-symmetrical 'graph model,' say, in the one depicted below where the distance from x to y is the length of the shortest directed path from x to y.



On the other hand, it can be satisfied in the following subspace of $\mathbb{R}$



The price we have to pay for the expressivity of the closer operator is that the satisfiability problem for $\mathcal{CSL}$ over natural classes of metric spaces becomes EXPTIME-hard (remember that $\mathbf{S4}_u$ is PSPACE-complete).

Special classes of distance spaces are the so-called *min-spaces* that satisfy the *min-condition*

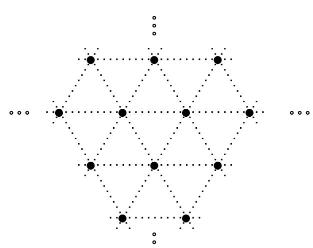
$$d(X,Y) \;=\; \inf\{d(x,y) \mid x \in X, y \in Y\} \;=\; \min\{d(x,y) \mid x \in X, y \in Y\},$$

for all sets X and Y. Over min-spaces, $\mathcal{CSL}$ is actually EXPTIME-complete (even if we extend it with distance operators $\exists^{<a}$ and $\exists^{\leq a}$, code the numerical parameters in binary, and allow nominals as well) [20]. Actually, the complexity remains the same no matter whether we assume symmetry (3) and the triangle inequality (2). Note that the term (8) is not satisfiable in any symmetric model satisfying the min-condition.

Rather unexpectedly, valid formulas of $\mathcal{CSL}$ are not even recursively enumerable when interpreted over finite subspaces (or arbitrary min-subspaces) of $\mathbb{R}$ or in $\mathbb{R}$ itself [19, 20]. This can be proved by a reduction of Hilbert's 10th problem on the unsolvability of Diophantine equations; see, e.g., [4] and references therein. The same holds for min-subspaces of $\mathbb{R}^n$, where $n \geq 2$. To give the reader some impression of what structures can be enforced on such subspaces of $\mathbb{R}^2$ by terms with the closer operator, consider the following formula

$$(p_0 \neq \bot) \quad \wedge \quad (p_1 \neq \bot) \quad \wedge \quad \bigwedge_{\substack{i,j<7 \\ j \neq i, i\oplus 1}} \big(p_i \sqsubseteq (p_j \leftrightarrows p_{i\oplus 1})\big),$$

where $\oplus$ is addition modulo 7. One can show that to satisfy this formula, a subspace of $\mathbb{R}^2$ must contain an infinite grid of the form



However, finding axiomatisations for $\mathcal{CSL}$ over other classes of models is open:

Problem 12. *Axiomatise the formulas of $\mathcal{CSL}$ that are valid in various classes of models.*

As concerns evaluating formulas of $\mathcal{CSL}$ in arbitrary (not necessarily min) metric spaces, we only know that satisfiability is decidable (and EXPTIME-hard).

Problem 13. *What is the computational complexity of the satisfiability problem for $\mathcal{CSL}$ over arbitrary metric models? What happens if we extend the language with nominals?*

Problem 14. *Gärdenfors* [10] *suggests that atomic terms of similarity languages should be interpreted by <u>convex</u> subsets of $\mathbb{R}^n$. Investigate the computational behaviour of $\mathcal{CSL}$ and its extensions under such or similar 'valuation restrictions.'*

Problem 15. *Characterise (un)decidable fragments of full $\mathcal{QMS}$ over various classes of models.*

Our similarity languages are 'crisp' in the sense that they operate with precise distances like 'the distance between two proteins is 3.1415926 . . . ' In practice such distances are only computed *approximately*. An interesting and important problem is the following:

Problem 16. *Develop logical formalisms capable of dealing with <u>non-crisp</u> distances, e.g., using vagueness/fuzziness/probability.*

Acknowledgments. The work on this paper was partially supported by U.K. EPSRC grants no. GR/S61973/01, GR/S61966/01, GR/S63175/01 and GR/S63182/01.

References

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

[2] P. S. Alexandroff. Diskrete Räume. *Matematicheskii Sbornik*, 2 (44):501–518, 1937.

[3] R. Alur, T. Feder, and T. Henzinger. The benefits of relaxing punctuality. *Journal of the ACM*, 43:116–146, 1996.

[4] J. Barwise, editor. *Handbook of Mathematical Logic*. North-Holland, Amsterdam, 1977.

[5] B. Bennett. Spatial reasoning with propositional logic. In *Proceedings of the 4th International Conference on Knowledge Representation and Reasoning*, pages 51–62. Morgan Kaufmann, 1994.

[6] N. Bourbaki. *General topology, Part 1.* Hermann, Paris and Addison-Wesley, 1966.

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

[8] M. Egenhofer and J. Herring. Categorizing topological relationships between regions, lines and point in geographic databases. Technical report, University of Maine, 1991.

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

[10] P. Gärdenfors. *Conceptual Spaces. The Geometry of Thought.* The MIT Press, 2000.

[11] Y. Hirshfeld and A. Rabinovich. Logics for real time: decidability and complexity. *Fundamenta Informaticae*, 62:1–28, 2004.

[12] O. Kutz. Notes on logics of metric spaces. *Studia Logica*, 2005. (In print).

[13] O. Kutz, H. Sturm, N.-Y. Suzuki, F. Wolter, and M. Zakharyaschev. Axiomatizing distance logics. *Journal of Applied Non-Classical Logic*, 12:425–440, 2002.

[14] O. Kutz, H. Sturm, N.-Y. Suzuki, F. Wolter, and M. Zakharyaschev. Logics of metric spaces. *ACM Transactions on Computational Logic*, 4:260–294, 2003.

[15] C. Lutz, D. Walther, and F. Wolter. Quantitative temporal logics: PSPACE and below. In *Proceedings of the Twelfth International Symposium on Temporal Representation and Reasoning*, Burlington, VT, USA, 2005. IEEE Computer Society Press.

[16] J.C.C. McKinsey. A solution of the decision problem for the Lewis systems **S2** and **S4**, with an application to topology. *Journal of Symbolic Logic*, 6:117–134, 1941.

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

[18] W. Nutt. On the translation of qualitative spatial reasoning problems into modal logics. In W. Burgard, T. Christaller, and A. Cremers, editors, *Advances in Artificial Intelligence. Proceedings of the 23rd Annual German Conference on Artificial Intelligence (KI'99)*, volume 1701 of *Lecture Notes in Computer Science*, pages 113–124. Springer, 1999.

[19] M. Sheremet, D. Tishkovsky, F. Wolter, and M. Zakharyaschev. 'Closer' representation and reasoning. In I. Horrocks, U. Sattler, and F. Wolter, editors, *International Workshop on Description Logics, (DL 2005)*, pages 25–36, 1999.

[20] M. Sheremet, D. Tishkovsky, F. Wolter, and M. Zakharyaschev. Comparative similarity, tree automata, and Diophantine equations. Submitted; available at `http://www.dcs.kcl.ac.uk/staff/mz/`, 2005.

[21] H. Sturm, N.-Y. Suzuki, F. Wolter, and M. Zakharyaschev. Semi-qualitative reasoning about distances: A preliminary report. In M. Ojeda-Aciego, I.P. de Guzmán, G. Brewka, and L. Moniz Pereira, editors, *JELIA*, volume 1919 of *Lecture Notes in Computer Science*, pages 37–56. Springer, 2000.

[22] A. Tarski. Der Aussagenkalkül und die Topologie. *Fundamenta Mathematicae*, 31:103–134, 1938.

[23] T. Tsao Chen. Algebraic postulates and a geometric interpretation of the Lewis calculus of strict implication. *Bulletin of the AMS*, 44:737–744, 1938.

[24] F. Wolter and M. Zakharyaschev. Reasoning about distances. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI 2003)*, pages 1275–1280. Morgan Kaufmann, 2003.

[25] F. Wolter and M. Zakharyaschev. A logic for metric and topology. *Journal of Symbolic Logic*, 70:795–828, 2005.

[26] F. Wolter and M. Zakharyaschev. On the computational complexity of metric logics. Manuscript, 2005.