

# The topo-approach to spatial representation and reasoning

Marco Aiello  
Dipartimento di Informatica e Telecomunicazioni  
Via Sommarive 14 - 38050 Trento  
Università di Trento  
<http://www.dit.unitn.it/~aiellom>  
[aiellom@dit.unitn.it](mailto:aiellom@dit.unitn.it)

## ABSTRACT

Commonsense knowledge about the surrounding physical world and quantitative theories of space, such as metric geometry, can be viewed as two extremes on how human beings relate to space. Qualitative spatial representation and reasoning places itself in between these two approaches. Qualitative spatial reasoning is a set of high-level theories which abstract from the quantitative details and attempt to mimic the human commonsense knowledge about space as much as possible. Successful approaches to spatial reasoning may impact many application areas of AI, most notably, robotics, computer vision in its broader sense, and natural language processing.

In this paper, we briefly overview a modal approach to spatial representation and reasoning, called *topo-approach*, presented in the PhD thesis “Spatial Reasoning: Theory and Practice,” winner of the AI\*IA dissertation award for 2003.

## 1 Reasoning about space

Spatial structures and spatial reasoning are essential to perception and cognition. Much day-to-day practical information is about what happens at certain spatial locations. Moreover, spatial representation is a powerful source of geometric intuitions that underlie general cognitive tasks. How can we represent spatially located entities and reason about them? To take a concrete domestic example: when we are setting a table and place a spoon, what are the basic spatial properties of this new item in relation to others, and to the rest of the space? Not only, there are further basic aspects to perception: we have the ability to compare different visual scenes, and recognize objects across them, given enough ‘similarity’. More concretely: which table settings are ‘the same’? This is another task for which logic provides tools.

Constraining space within the bounds of a logical theory and using related formal reasoning tools must be

performed with particular care. One cannot expect the move from space to formal theories of space to be complete. Natural spatial phenomena will be left out of logical theories of space, while non-natural spatial phenomena could try to sneak in (cf. the account of Helly’s theorem implications on diagrammatic reasoning in [29]). Paraphrasing Ansel Adams’ concern of space bound in a photograph [1], one could say that space in nature is one thing; space confined and restricted in the bounds of a formal representation and reasoning system is quite another thing. Connectivity, parthood, and coherence should be correctly handled and expressed by the formalism, not aiming at a complete representation of space, but focusing on expressing the most perspicuous spatial phenomena.

The preliminary and fundamental step in devising a spatial reasoning framework lays, thus, in the identification of which spatial behaviors the theory should capture and, possibly, in the identification of which practical uses will be made of the framework. A key factor is in appropriately balancing the right amount of expressive power, completeness with respect to a specific class of spatial phenomena and computational complexity. The blend of expressivity and tractability we are aiming at points us in the direction of modal logics as a privileged candidate for the formalization task (see for instance [18]). We assume the reader has some basic knowledge of modal logics and its best-known Kripke semantics (also named possible world semantics). Strangely enough, even assuming knowledge of Kripke semantics, we are going to make little use of it, and rather resort to topological semantics, introduced about 30 years earlier by Tarski [42]. Modern modal logics of space need old modal logic semantics.

The attention to spatial reasoning stems, in the case of the presented research effort, from the interest in applications in the domains of image processing and computer vision, therefore, the sub-title of the thesis *Theory and Practice*. But this is only one of the many motivations for which spatial logics have been considered in the past. These range from the early philosophical efforts

[48, 31] to recent work motivated by such diverse concerns as spatial representation and vision in AI [40, 36], semantics of spatial prepositions in linguistics [27, 49], perceptual languages [22], or diagrammatic reasoning [26, 25]. The resulting logics are diverse, too. Theories differ in their primitive objects: points, lines, polygons, regions (contrast [42] against [43]). Likewise, theories differ in their primitive spatial relations: such as inclusion, overlap, touching, ‘space’ versus ‘place’, and on how these should be interpreted: [35, 17, 13]. There are mereological theories of parts and wholes, topological ones (stressing limit points and connection) and mereotopological ones (based on parthood and external connection). Systematic accounts of the genesis of spatial vocabulary date back to Helmholtz’ work on invariants of movement, but no generally agreed primitive relations have emerged on the logic side. Moreover, axioms differ across theories: [21] vs. [34] vs. [20]. Even our modal approach has its predecessors: of which we mention [38, 41, 17, 46, 15, 30].

The remainder of the paper is organized as follows. In Section 2, we review the topo-approach to spatial reasoning. In Section 3, we look at various modal languages of increasing expressive power. Section 4 deals with two AI applications of modal spatial logics. Conclusions are summarized in Section 5.

## 2 The topo-approach

In the topo-approach regions of a topological space are denoted by modal formulas. Consider, for instance, the real plane with its standard topology and draw a spoon on it, as shown in Figure 1. One may want to identify the handle of the spoon, or the small shallow bowl, or the point where these two components are joined. This can be achieved by ‘naming’ the spoon with a proposition letter, say  $p$ , and then by means of boolean and modal operators identifying the other relevant regions of the spoon  $p$ . Lets see how this is feasible using a modal logic.

In the 30s, Tarski provided a topological interpretation and various completeness theorems [32, 37] making the modal logic S4 the basic logic of topology. In the topological interpretation of a modal logic, each propositional variable represents a region of the topological space, and so does every formula. Boolean operators such as negation  $\neg$ , or  $\vee$ , and  $\wedge$  are interpreted as complement, union and intersection, respectively. The modal operators diamond and box, become the topological closure and interior operators. More precisely, the modal logic S4, called  $\mathcal{L}$  onward, consists of:

- a set of proposition letters  $P$ ,
- two constant symbols  $\top, \perp$ ,
- Boolean operators  $\neg, \wedge, \vee, \rightarrow$ , and
- two unary modal operators  $\Box, \Diamond$ .

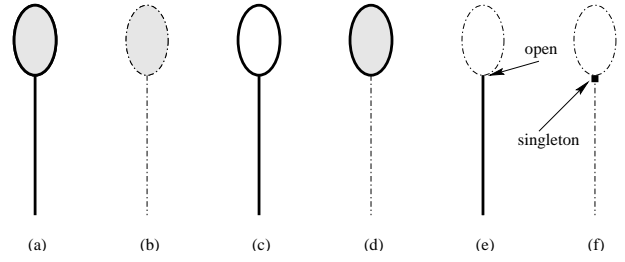


Figure 1: A formula of the language  $\mathcal{L}$  identifies a region in a topological space. (a) a spoon,  $p$ . (b) the containing part of the spoon,  $\Box p$ . (c) the boundary of the spoon,  $\Diamond p \wedge \Diamond \neg p$ . (d) the container part of the spoon with its boundary,  $\Diamond \Box p$ . (e) the handle of the spoon,  $p \wedge \neg \Diamond \Box p$ . In this case the handle does not contain the junction point handle-container. (f) the joint point handle-container of the spoon,  $\Diamond \Box p \wedge \Diamond (p \wedge \neg \Diamond \Box p)$ : a singleton in the topological space.

Formula	Interpretation
$\top$	the universe
$\perp$	the empty region
$\neg \varphi$	the complement of a region
$\varphi \wedge \psi$	intersections of the regions $\varphi$ and $\psi$
$\varphi \vee \psi$	union of the regions $\varphi$ and $\psi$
$\Box \varphi$	interior of the region $\varphi$
$\Diamond \varphi$	closure of the region $\varphi$

Figure 2: Formulas of  $\mathcal{L}$  and their intended meaning.

Formulas are built by means of the following recursive rules:

- $p$  such that  $p \in P$  is a well formed formula,
- $\top, \perp$  are well formed formulas,
- $\neg \varphi, \varphi \vee \psi, \varphi \wedge \psi$  are well formed formulas if  $\varphi$  and  $\psi$  are well formed formulas,
- $\Box \varphi$  and  $\Diamond \varphi$  are well formed formulas if  $\varphi$  is well formed formula.

In Figure 2, the intended meaning of some basic formulas is summarized. The intuitions about the language are reflected in its semantics. Topological models  $M = \langle X, O, \nu \rangle$  are topological spaces  $(X, O)$  plus a valuation function  $\nu : P \rightarrow \mathcal{P}(X)$ .

**Definition 1 (topological semantics of  $\mathcal{L}$ )** Truth of modal formulas is defined inductively at points  $x$  in topological models  $M$ :

$M, x \models \perp$	never
$M, x \models \top$	always
$M, x \models p$	iff $x \in \nu(p)$ (with $p \in P$ )
$M, x \models \neg\varphi$	iff not $M, x \models \varphi$
$M, x \models \varphi \wedge \psi$	iff $M, x \models \varphi$ and $M, x \models \psi$
$M, x \models \varphi \vee \psi$	iff $M, x \models \varphi$ or $M, x \models \psi$
$M, x \models \varphi \rightarrow \psi$	iff if $M, x \models \varphi$ , then $M, x \models \psi$
$M, x \models \Box\varphi$	iff $\exists o \in O :$ $x \in o \wedge \forall y \in o : M, y \models \varphi$
$M, x \models \Diamond\varphi$	iff $\forall o \in O :$ if $x \in o$ , then $\exists y \in o : M, y \models \varphi$

As usual we can economize by defining  $\varphi \vee \psi$  as  $\neg\varphi \rightarrow \psi$ , and  $\Diamond\varphi$  as  $\neg\Box\neg\varphi$ .

One of Tarski's early results was this. Universal validity of formulas over topological models has the modal logic  $\mathcal{L}$  as a sound and complete proof system. The axiomatization is:

$$\Box\top \quad (\text{N})$$

$$(\Box\varphi \wedge \Box\psi) \leftrightarrow \Box(\varphi \wedge \psi) \quad (\text{R})$$

$$\Box\varphi \rightarrow \varphi \quad (\text{T})$$

$$\Box\varphi \rightarrow \Box\Box\varphi \quad (4)$$

Modus Ponens and Monotonicity are the only rules of inference

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

In addition, consider the following derived theorem of  $\mathcal{L}$ :

$$\Box A \vee \Box B \leftrightarrow \Box(\Box A \vee \Box B) \quad (\text{or})$$

Axiom (N) says that the whole space is open. (R) is the finite intersection condition on a topological space. Next, (or) says that open sets are closed under finite unions. (Closure under *arbitrary* unions requires an infinitary extension of the modal language.) Finally, axiom (T) says every set contains its interior, and (4) expresses inflationarity of the interior operator. Further principles of  $\mathcal{L}$  may define special notions in topology. For instance, the derived rule

$$\text{if } \Box(\varphi \leftrightarrow \Diamond\Box\varphi), \text{ then } \Box(\Box\neg\varphi \leftrightarrow \Box\Diamond\Box\neg\varphi)$$

says that if a set is closed regular, so is its 'open complement'.

There is a natural question on whether we are 'correctly' interpreting modal languages with respect to topological spaces. The first answer is due to McKinsey and Tarski.

**Theorem 1**  $\mathcal{L}$  is the complete logic of any metric separable dense-in-itself space.

This, though, is not the only completeness result available for the topological version of S4. In fact, in [11] we prove completeness using modern modal logic techniques such as the construction of canonical models.

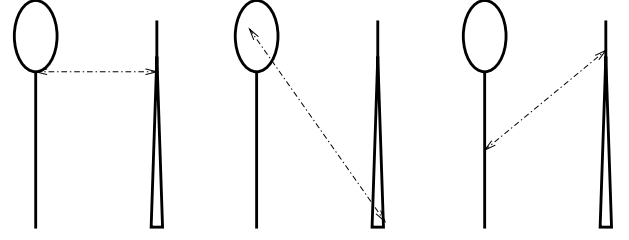


Figure 3: A spoon is bisimilar to a 'chop-stick'.

**Theorem 2** For any set of formulas  $\Gamma$ ,

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

In [11] we also derive a number of results relating S4 to specific topological spaces such as finite well-connected topological spaces and the Cantor space. We also introduce the logic of serial sets for the real line, that is, the logic of regions which are finite unions of convex sets of the real line. The properties of finite unions of convex sets have proved to be useful on the practical side [3, 2].

## 2.1 Topological bisimulation

Once we have a language for expressing properties of visual scenes, we can also formulate *differences* between such scenes. This brings us to the notion of 'sameness' for spatial configurations associated with  $\mathcal{L}$ , and hence to techniques of comparison. The following is the topological version of a well-known notion from modal logic and computer science [44, 33].

**Definition 2 (topological bisimulation)** Consider the language  $\mathcal{L}$  and two topological models  $\langle X, O, \nu \rangle$ ,  $\langle X', O', \nu' \rangle$ . A *topological bisimulation* is a non-empty relation  $\rightleftharpoons \subseteq X \times X'$  such that if  $x \rightleftharpoons x'$  then:

- (i)  $x \in \nu(p) \Leftrightarrow x' \in \nu'(p)$  (for any proposition letter  $p$ )
- (ii) (*forth condition*):  $x \in o \in O \Rightarrow \exists o' \in O' : x' \in o'$  and  $\forall y' \in o' : \exists y \in o : y \rightleftharpoons y'$
- (iii) (*back condition*):  $x' \in o' \in O' \Rightarrow \exists o \in O : x \in o$  and  $\forall y \in o : \exists y' \in o' : y \rightleftharpoons y'$

We call a bisimulation *total* if it is defined for all elements of  $X$  and of  $X'$ . We overload the symbol  $\rightleftharpoons$  extending it to models with points:  $\langle X, O, \nu \rangle, x \rightleftharpoons \langle X', O', \nu' \rangle, x'$  requires also that  $x \rightleftharpoons x'$ . If only the atomic clause (i) and the forth condition (ii) hold, we say that the second model *simulates* the first one.

To motivate this definition, one can look at the 'topological dynamics' of the back and forth clauses, seeing how they make  $x, x'$  lie in the same 'modal setting'. Further motivations come from a match with modal formulas, and basic topological notions.

**Example 1 (spoon and chop-stick)** *Is a spoon the same as a chop-stick?* The answer depends of course on how we define this cutlery. Suppose we let the spoon be a closed ellipse plus a touching straight line and the chop-stick a straight line touching a closed triangle (cf. Figure 3). Let us regard both as the interpretation of some fixed proposition letter  $p$  in their respective models. Then we do have a topo-bisimulation by matching up (a) the two ‘junction points’, (b) all points in the two handles, and likewise for (c) the interiors, (d) the remaining boundary points, and (e) all exterior points in both models.

Many more examples and cutlery related pictures of topologically bisimilar and not spaces can be found in [10].

Crucially, modal spatial properties are *invariant for topo-bisimulations*:

**Theorem 3** *Let  $M = \langle X, O, \nu \rangle$ ,  $M' = \langle X', O', \nu' \rangle$  be models with bisimilar points  $x \in X$ ,  $x' \in X'$ . For all modal formulas  $\varphi$ ,  $M, x \models \varphi$  iff  $M', x' \models \varphi$ .*

To clinch the fit, we need a converse. In general this fails, and matters become delicate (see [18]). The converse does hold when we use an *infinitary* modal language—but also for our finite language over special classes of models. Here is a nice illustration: finite modally equivalent pointed models are bisimilar.

**Theorem 4** *Let  $M = \langle X, O, \nu \rangle$ ,  $M' = \langle X', O', \nu' \rangle$  be two finite models,  $x \in X$ , and  $x' \in X'$  two points in them such that for every  $\varphi$ ,  $M, x \models \varphi$  iff  $M', x' \models \varphi$ . Then there exists a bisimulation between  $M$  and  $M'$  connecting  $x$  and  $x'$ .*

## 2.2 Topo-bisimilar reductions

In many contexts, bisimulations and simulations are used to find minimal models. This is useful, for instance, to find minimal representations for labeled transition systems having certain desired properties modally expressible. Topo-bisimulation can be used for finding a minimal representation for a determined spatial configuration. For example, consider a spoon with two handles, as depicted in Figure 6.a. The spoon has 7 ‘salient’ points; these satisfy the formulas reported in Figure 4. It is easy to find a  $\mathcal{L}$  Kripke model satisfying the 7 formulas above, for instance, the one in Figure 6.a. By a bisimulation one ‘reduces’ it to a minimal similar one. The topo-bisimilar reduction is presented in the table on the right of Figure 6.

From the reduced model one can ‘reconstruct’ the pictorial example, that is, a spoon with only one handle, Figure 6.b. Checking the topo-bisimilarity of Figure 6.a and Figure 6.b is an easy task to perform. We do not spell out the general method used here for transforming topological models into Kripke ones (and back); but it should be fairly clear from the example.

Point	Formula
1	$\Box p$
2	$\Diamond p \wedge \Diamond \neg p$
3	$\Box \neg p$
4	$p \wedge \neg \Diamond \Box p \wedge \Diamond \Box \neg p$
5	$\Diamond \Box p \wedge \Diamond (p \wedge \neg \Diamond \Box p)$
6	$p \wedge \neg \Diamond \Box p \wedge \Diamond \Box \neg p$
7	$\Diamond \Box p \wedge \Diamond (p \wedge \neg \Diamond \Box p)$

Figure 4: Formulas true at points of the model in Figure 5.

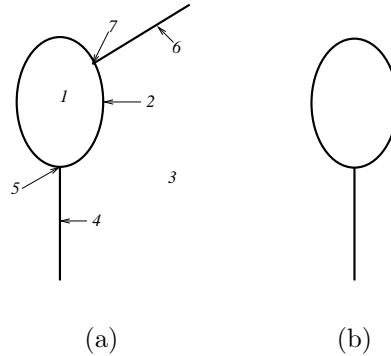


Figure 5: The reduction of a topological model to a minimal topo-bisimilar one.

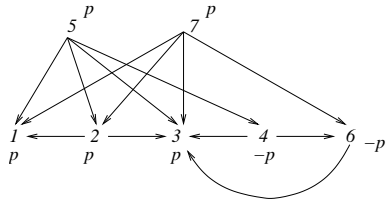
The claim is not that one should move back and forth from topological and Kripke semantics to find minimal models. Our goal is to show that topo-bisimulations enable the reduction of spatial models in the same way that bisimulations enable the reduction of Kripke models. A general algorithm for deciding topo-bisimulation is still missing, but for a specific class of models it does exist [6].

## 2.3 Games that compare visual scenes

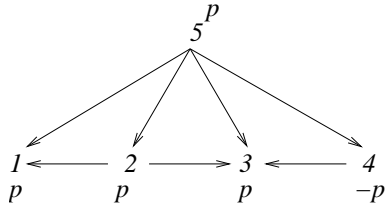
Topo-bisimulation is a global notion of comparison. But in practice, we are interested in *fine-structure*: what are the ‘simplest differences’ that can be detected between two visual scenes? For this purpose, we introduce *topo-games* between first-order models [24]. Similarity and difference between visual scenes will then have to do with *strategies* for players comparing them.

**Definition 3 (topological game)** Consider two models  $\langle X, O, \nu \rangle$ , and  $\langle X', O', \nu' \rangle$ , a natural number  $n$  and two points  $x_1 \in X$ ,  $x'_1 \in X'$ . A *topological game* of length  $n$ , with starting points  $x_1, x'_1$ —notation  $TG(X, X', n, x_1, x'_1)$ —consists of  $n$  rounds between two players: Spoiler and Duplicator. Each round proceeds as follows:

- (i) Spoiler chooses a model  $X_s$  and an open  $o_s$  containing the current point  $x_s$  of that model



(a)



(b)

(a)	(b)
1	1
2	2
3	3
4	4
5	5
6	4
7	5

Figure 6: The reduction of the spoons of Figure 5 via a bisimulation on the corresponding Kripke models. In the table, the bisimulation relation.

- (ii) Duplicator chooses an open  $o_d$  in the other model  $X_d$  containing the current point  $x_d$  of that model
- (iii) Spoiler picks a point  $\bar{x}_d$  in Duplicator's open  $o_d$  in the  $X_d$  model
- (iv) Duplicator finally picks a point  $\bar{x}_s$  in Spoiler's open  $o_s$  in  $X_s$

The points  $\bar{x}_s$  and  $\bar{x}_d$  become the new current points of the  $X_s$  and  $X_d$  models, respectively. After  $n$  rounds, two sequences have been built:

$$\{x_1, o_1, x_2, o_2, \dots, o_{n-1}, x_n\} \{x'_1, o'_1, x'_2, o'_2, \dots, o'_{n-1}, x'_n\}$$

with  $x_i \in o_i$ , and  $o_i \in O$  (analogously for the second sequence). After  $n$  rounds, if  $x_i$  and  $x'_i$  (with  $i \in [1, n]$ ) satisfy the same atoms, Duplicator *wins*. (Note that Spoiler already wins 'en route', if Duplicator fails to maintain the atomic match.) A *winning strategy* ('w.s.' for short) for Duplicator is a function from any sequence of moves by Spoiler to appropriate responses which always ends in a win for Duplicator. The same notion applies to Spoiler. An *infinite topological game* is one without a finite limit to the number of rounds. In this case, Duplicator wins if the matched points continue to satisfy the same atoms.

**Example 2 (playing on spoons)** Consider the three configurations in Figure 7. (a) The leftmost game starts with a point on the boundary of the spoon versus an interior point of the other spoon. Spoiler can win this game in one round by simply choosing an open set on the

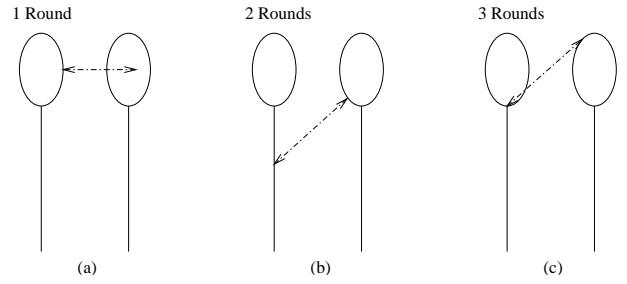


Figure 7: Games on two spoons with two different starting points. On top, the number of rounds needed by Spoiler to win.

right spoon completely contained in its interior. Duplicator's open response must always contain a point not in the spoon, which Spoiler can then pick, giving Duplicator no possible response. (b) In the central game, a point on the handle is compared with a boundary point of the spoon's container. Spoiler can again win the game, but needs two rounds this time. Here is a winning strategy. First, Spoiler chooses an open on the left spoon containing the starting point but without interior points. Any open chosen by Duplicator on the other spoon must contain an interior point. Spoiler then picks such an interior point. Duplicator's response to that can only be a boundary point of the other model (on the handle) or a point outside of the spoon. In the latter case, she loses at once – in the former, she loses in one round, by reduction to the previous game. (c) Finally, on the left the junction between handle and container is compared with a boundary point of the container. In this game, Spoiler will choose an open on the right model, avoiding points on the handle of the spoon. Duplicator is forced to choose an open on the left containing points on the handle. Spoiler then picks such a handle point. Duplicator replies either with an interior point, or with a boundary point of the right spoon. Thus we are back with game (b), and Spoiler can win in the remaining two rounds.

The fine-structure provided by games measures differences in terms of the minimum number of rounds needed by Spoiler to win. These same differences may also be formulated in terms of our modal language. To see this, we need the notion of *modal rank*, being the maximum number of nested modal operators in a formula. For instance, the modal ranks of the formulas in Figure 1:  $p$ ,  $\Box p$ ,  $p \wedge \neg \Box p$ ,  $\Diamond \Box p$ ,  $p \wedge \neg \Diamond \Box p$ ,  $\Diamond \Box p \wedge \Diamond (p \wedge \neg \Diamond \Box p)$ , are 0, 1, 1, 2, 2, and 3, respectively. Here is the relation between games,  $\mathcal{L}$  and topological models.

**Theorem 5 (adequacy)** topological game!adequacy Duplicator has a w.s. in  $TG(X, X', n, x, x')$  iff  $x$  and  $x'$  satisfy the same formulas of modal rank up to  $n$ .

This is the usual version of adequacy: slanted toward similarity. But in our pictorial examples, we rather

looked at Spoiler. One can also set up the proof of Theorem 5 so as to obtain an *effective correspondence* between (a) winning strategies for Spoiler, (b) modal ‘difference formulas’ for the initial points. Here is an illustration.

**Example 3 (matching strategies with formulas)**

Look again at Figure 7. The strategies described for Spoiler are immediately linked to modal formulas that distinguish the two models. Suppose the spoons are denoted by the proposition letter  $p$  and hence the background by  $\neg p$ . In the game on the left,  $\Box p$  is true of the starting point of the right spoon, and its negation  $\Diamond \neg p$  is true of the starting point of the other spoon. The modal depth of these formulas is one and therefore Spoiler can win in one round. In the central case, a distinguishing formula is  $\neg \Diamond \Box p$ , which holds for the starting point on the left spoon, but not for that on the right. The modal depth is 2, which is the number of rounds that Spoiler needed to win the game. Finally, a formula of modal depth 3 that is only true of the point on the left spoon of the leftmost game is:  $\Diamond(p \wedge \neg \Diamond \Box p)$ . The negation of this formula is true on the other starting point, thus justifying Spoiler’s winning strategy in 3 turns.

There is still more fine-structure to these games. E.g., visual scenes may have several modal differences, and hence more than one winning strategy for Spoiler. Also, recall that topo-games can be played *infinitely*. Then the winning strategies for Duplicator (if any) are precisely the various *topo-bisimulations* between the two models.

**2.4 Topo-distance**

Topo-bisimulations are an equivalence relation, so one may very well use them to define identity of patterns. Via simulations, one can also consider issues of a pattern being a sub-pattern of another one. Then, topo-games are a refining notion of topo-bisimulations. Therefore, one may use topo-games to define a measure of difference among spatial patterns. Think of it this way. The less it takes Spoiler to win a game, the more different must the spatial patterns be, the more *unsimilar*. On the opposite, the longer Duplicator can resist, the more *similar* are the spatial patterns. In the limit, if Duplicator can resist forever, i.e., in the infinite round game, the two patterns are topologically bisimilar. Now comes the technical problem. Topo-games are defined as a way of comparing two given topological models, exactly in the spirit of the original definition of first-order model comparison games à la Ehrenfeucht-Fraïssé, but we need a similarity measure on the whole class of models; we need a measure that behaves uniformly across all models for  $\mathcal{L}$ .

The first intuition on turning model comparison games into a similarity measure may be misleading in a pessimistic direction. To get to a similarity measure, we

need to define a distance in terms of topo-games. Distances require considering more than just two models at a time. Consider, for example, three models and the three model comparison games that can be played. The formulas, the points and open sets picked in the three games may be completely unrelated one game from each other, therefore, one may be discouraged and conjecture that model comparison games are not related across different models of the same class.

Even though the remark on the unrelatedness of the strategies for different games is true. It turns out that there is still an interrelation between model comparison games over two given models and the whole class. Most importantly, the relation can be defined to satisfy the three properties defining a distance measure. Here is how.

**Definition 4 (isosceles topo-distance)** Consider the space of all topological models  $T$ . *Spoiler’s shortest possible win* is the function  $spw : T \times T \rightarrow \mathbb{N} \cup \{\infty\}$ , defined as:

$$spw(X_1, X_2) = \begin{cases} n & \text{if Spoiler has a winning strategy} \\ & \text{in } TG(X_1, X_2, n), \\ & \text{but not in } TG(X_1, X_2, n - 1) \\ \infty & \text{if Spoiler does not have a} \\ & \text{winning strategy in} \\ & TG(X_1, X_2, \infty) \end{cases}$$

The *isosceles topo-model distance* (*topo-distance*, for short) between  $X_1$  and  $X_2$  is the function  $tmd : T \times T \rightarrow [0, 1]$  defined as:

$$tmd(X_1, X_2) = \frac{1}{spw(X_1, X_2)}$$

The distance was named ‘isosceles’ since it satisfies the triangular property in a peculiar manner. Given three models, two of the distances among them (two sides of the triangle) are always the same and the remaining distance (the other side of the triangle) is smaller or equal. In [3, 2], it is shown that indeed *tmd* behaves as a distance measure.

**Theorem 6 (isosceles topo-model distance)** *tmd* is a distance measure on the space of all topological models.

**3 Extensions**

So far we have presented a modal language of basic topology and we have equipped ourselves with a number of tools: (1) a topological semantics to interpret formulas as regions; (2) topological bisimulations to assess the equivalence of patterns, which comes with the notion of simulations and of reductions; (3) topological games to compare any two given patterns; and, finally,

(4) topological distance to measure globally the similarity of any pattern expressible with the given language. Though, the language  $\mathcal{L}$  used to introduce the topo-approach is not very powerful. One can merely express simple topological properties at a given point. Thus, one may wonder how general the topo-approach to space is. In [9, 6], we identify a number of modal languages for which the same approach is viable. Some of these are already well-known in the field of spatial reasoning, such as the extension of S4 with a universal operator [17] others are new. Here we simply provide an overview following two roads: that of extending the logical power of  $\mathcal{L}$  and that of extending the geometrical power.

### 3.1 Extensions of the logical power

An extremely useful technique in modal logics to gain expressive power without leaving the guarded area of decidable languages is to add a modal operator. For instance, if one needs to express notions connected to equality of states in Kripke semantics, one may add a difference operator  $D\varphi$  which reads “there is a state different from the current one that satisfies  $\varphi$ .” This is exactly what the sort of extensions we are after. We consider important topological relations not captured by  $\mathcal{L}$  alone which can be safely expressed by ‘adding’ appropriate new modal operators [23].

The first limitation to overcome is  $\mathcal{L}$ ’s locality. The formulas are evaluated at points and provide local information, e.g., the point  $x$  is in the open set given by the intersection of the interior of  $\varphi$  and  $\psi$  ( $M, x \models \Box\varphi \wedge \Box\psi$ ). By this information we know a lot about the point  $x$ , but very little about the set denoted by  $\Box\varphi \wedge \Box\psi$ , we merely know that there is one point satisfying it, the point  $x$ . Introducing an universal (or global) modality is the solutions to this problem. For instance, with  $\mathcal{L}+$ (the universal modality) one is able to express whether a topological space is connected or not, which is clearly a global property of the space and not a local one of some points of the space.

The truth definition of  $\mathcal{L}$  is extended with the following:

$$\begin{aligned} M, x \models E\varphi & \quad \text{iff} \quad \exists y \in X, M, y \models \varphi \\ M, x \models U\varphi & \quad \text{iff} \quad \forall y \in X : M, y \models \varphi \end{aligned}$$

The definition reads, for  $E\varphi$ , “there exists a point in the model satisfying  $\varphi$ ,” and dually for  $U\varphi$ , “all the points in the model satisfy  $\varphi$ .”

Some basic facts are that  $U$  and  $E$  modalities follow the axiomatization of S5 [17] and that it is possible to identify normal forms [6]. But, in the context of this article, the more interesting feature is that notions of topo-bisimulation, of topo-games and topo-distance exist. The adequate version of topo-bisimulation follows the definition of that for  $\mathcal{L}$  with the addition that the relation must be defined for all points of the two models. As for the games, the difference lies in the fact that these are no longer defined for a given starting point and that two types of rounds are now possible. One as in

the games for  $\mathcal{L}$  plus the following one:

$$\text{global} \left\{ \begin{array}{l} (i) \quad \text{Spoiler chooses a model } X_s \text{ and picks} \\ \quad \text{a point } \bar{x}_s \text{ anywhere in } X_s \\ (ii) \quad \text{Duplicator chooses a point } \bar{x}_d \text{ anywhere} \\ \quad \text{in the other model } X_d \end{array} \right.$$

It is up to Spoiler to decide which type of round he wants to engage Duplicator in.

The universal extension is not the only possibility to enhance the logical power. An alternative is that of using hybrid modal references, cf. [12]:

$$\begin{aligned} M, x \models @_A\varphi & \quad \text{iff} \quad \forall y \in \nu(A) M, y \models \varphi \\ M, x \models @_a\varphi & \quad \text{iff} \quad \exists y \in \nu(A) M, y \models \varphi \end{aligned}$$

The last increase of logical power we refer to is that of using an operator analogue to the well-known temporal logics operator Until [28]. If one abstracts from the temporal behavior and interprets the modality in spaces with dimensionality greater than one, one gets an operator expressing something to be valid *up to a certain boundary* region, a sort of fence surrounding the current region. Here is a natural notion of spatial ‘Until’ in topological models:

$$\begin{aligned} M, x \models \varphi\mathcal{U}\psi & \quad \text{iff} \quad \exists A : O(A) \wedge x \in A \wedge \forall y \in A. \varphi(y) \wedge \\ & \quad \forall z (z \text{ is on the boundary of } A \wedge \psi(z)) \end{aligned}$$

A graphical representation of the Until operator is presented in Figure 8. Its expressiveness is strictly richer than that of the basic modal language of space.

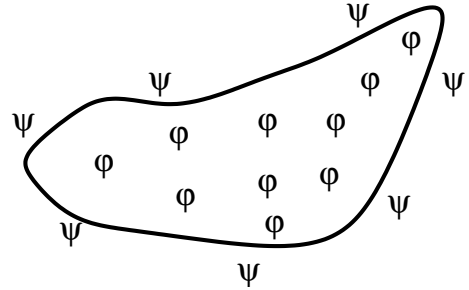


Figure 8: The region involved in  $\varphi\mathcal{U}\psi$ .

### 3.2 Geometrical extensions

Rather than simply using more expressive logical operators, an alternative way of extending the expressive power of a modal language of space is that of enriching the spatial structure of the models. A first elementary example is the property of a point’s being in the convex closure of a set of points. That is, there exists a segment containing the points whose end-points are in the set. Capturing convexity modally involves a standard similarity type, that of frames of points with a ternary relation of betweenness:  $M, x \models C\varphi$  iff

$$\exists y, z : M, y \models \varphi \wedge M, z \models \varphi \wedge x \text{ lies in between } y \text{ and } z$$



This definition is slightly different from the usual notion of convex closure. It is a one-step convexity operator whose countable iteration yields the standard convex closure. One-step convexity exhibits a modal pattern for an existential binary modality:

$$\exists yz : \beta(yxz) \wedge \varphi(y) \wedge \varphi(z)$$

We are now moving in the domain of affine spaces. These have a strong modal flavor, as shown by [15, 14, 47], where two roads are taken. One merges points and lines into one sort of pairs (point, line) equipped with two incidence relations. The other has two sorts for points and lines, and a matching modal operator. But there are more expressive classical approaches to affine structure. In [43], Tarski gave a full first-order axiomatization of elementary geometry in terms of a ternary betweenness predicate  $\beta$  and quaternary equidistance  $\delta$ . More precisely, betweenness  $\beta(xyz)$  means that the point  $y$  lies in between  $x$  and  $z$ , allowing  $y$  to be one of these end-points. Line structure is immediately available by defining *collinearity* in terms of betweenness:

$$xyz \text{ are collinear iff } \beta(xyz) \vee \beta(yzx) \vee \beta(zxy)$$

Ternary betweenness models a binary betweenness modality  $\langle B \rangle$ :  $M, x \models \langle B \rangle(\varphi, \psi)$  iff

$$\exists y, z : \beta(yxz) \wedge M, y \models \varphi \wedge M, z \models \psi$$

Note that this is a more standard modal notion than the earlier topological modality: we are working on frames, and there are no two-step quantifiers hidden in the semantics. But what about our topo-approach, can we find an adequate notion of bisimulation for affine languages? The answer is again positive.

**Definition 5 (affine bisimulation)** Given two affine models  $\langle X, O, \beta, \nu \rangle$ , and  $\langle X', O', \beta', \nu' \rangle$ , an *affine bisimulation* is a non-empty relation  $\rightleftharpoons \subseteq X \times X'$  such that, if  $x \rightleftharpoons x'$ :

- (i)  $x$  and  $x'$  satisfy the same proposition letters,
- (ii) (*forth condition*):  $\beta(yxz) \Rightarrow \exists y'z' : \beta'(y'x'z')$  and  $y \rightleftharpoons y'$  and  $z \rightleftharpoons z'$
- (iii) (*back condition*):  $\beta'(y'x'z') \Rightarrow \exists yz : \beta(yxz)$  and  $y \rightleftharpoons y'$  and  $z \rightleftharpoons z'$

where  $x, y, z \in X$  and  $x', y', z' \in X'$ .

We may be after even more structure than just affine point and line patterns. Tarski's equidistance also captures *metric* information. There are various primitives for this. Tarski used quaternary equidistance—while ternary equidistance would do just as well ( $x, y$  and  $z$  lie at equal distances). An alternative choice, which we stress in [9], is that of considering a ternary relation of relative nearness (originally introduced in [45]):  $N(x, y, z)$  iff

$$y \text{ is closer to } x \text{ than } z \text{ is, i.e., } d(x, y) < d(x, z) \\ \text{where } d(x, y) \text{ is any distance function.}$$

This is meant very generally. The function  $d$  can be a geometrical metric, or some more cognitive notion of visual closeness (van Benthem's original interest [45]; cf. also Gärdenfors 'Conceptual Spaces'), or some utility function with metric behavior. Randell et al. [36] develop the theory of comparative nearness for the purpose of robot navigation, related to the region calculus RCC [35]. Note that relative nearness defines equidistance by  $Eqd(x, y, z) : \neg N(x, y, z) \wedge \neg N(x, z, y)$ , furthermore, Tarski's quaternary equidistance is expressible in terms of  $N$  as well. The relation naturally maps to a binary modal operator defined as:  $M, x \models \langle N \rangle \phi, \psi$  iff

$$\exists y, z : M, y \models \psi \wedge M, z \models \varphi \wedge N(x, y, z)$$

With this operator one can amuse himself with various well-known geometrical constructions [6], such as, the Mascheroni construction (where one can build elementary geometry by using the compass alone), Voronoi diagrams (where regions of influences of a given set of points are separated), and Delaunay triangulation (where the space is partitioned into triangles).

Finally, one may consider vector spaces as the models for modal languages. For brevity, we omit the treatment of this here and refer to [6] for a speculation on the relation between mathematical morphology [39, 19] and linear and arrow logics.

## 4 Applications

The topo-distance—build on the games played by two antagonist which challenge each other in finding differences among spatial patterns—is a similarity measure between spatial configurations. There are many possible uses for similarity measures in computer vision. A robot may compare what it sees with a database of known patterns. A natural language system may give various semantics to the phrase 'part-of', 'same', 'just like' when the object of qualification are spatial. Another possibility is that of using similarity measures for retrieving images from a database.

We have approached this last task with the prototype IRIS (Image RetrIeval based on Spatial relationships) [3, 2]. The similarity measure of IRIS is based on the topo-distance for the language  $\mathcal{L}$  extended with the universal modality. This allows expressing region properties such as the 'man inside the car' or the 'ball touching the foot' in a pictorial manner. The prototype has shown the applicability of the topo-approach. On the negative side, the distance assigns similarity values which are not close to human intuition. In addition, the system shows brittleness as the quality of the segmentation of the images decreases.

Another application of spatial reasoning we have investigated is that of logical structure detection in the context of document image analysis. The task is that of reconstructing the intended meaning of document images analyzing the layout of the document. The approach uses

a topological language of the plane with weak geometrical expressive power. Extensive experimentation has been performed showing positive results [7, 8, 4]. The results are even more encouraging if one considers the heterogeneity of the document collections used for experimentation (different journals, magazines, one-page ads, etc.).

## 5 Conclusions

The research underlying [6] is an attempt to bring together two research areas: the standard mathematical approach to space (topology, geometry, and linear algebra) with a computational analysis of visual processing tasks. To build such a bridge, we proposed a modal logic approach, which connects up with both: one, more tractable levels of spatial structure inside mathematical theories; and two, more expressive power in computational tasks. The results in [6] show the connection meaningful by providing a number of tools which are both useful for ‘deconstructing mathematics’ and for the analysis and redesign of computational tasks. In particular, the topo-approach we propose is a framework for topological reasoning with a modal language of visual patterns, emphasizing bisimulation and comparison games as a means of calibrating similarity of visual scenes. Moreover, a pleasing side-effect is a new take on elementary topology. Laying the basis for a more ambitious program of ‘modal geometry’, exploring new fine-structure of tractable fragments of geometry; just as modal logic itself does for first-order logic.

Our analysis of space and of applications of spatial theories is only a small step which generates more questions than answers. We identified many new open problems along the way in the thesis. Thus, our work may also serve as a pilot study for a broader modal geometry developed with a view to potential applications.

## 6 Acknowledgments

The results presented in this article and in the dissertation should be considered mine as much as of my co-authors which I thank. In particular, I am grateful to my thesis supervisor Johan van Benthem. Part of the research described here has been funded by the Institute for Logic, Language and Computation [5], the Italian and Dutch national research councils (CNR, NWO). Conference and lab visits have been supported by AI\*IA, CSLI (Stanford) and the Jelia2000 organizing committee.

## REFERENCES

- [1] A. Adams. *The Negative*. Brown Little and Company, 1981. Page 129.
- [2] M. Aiello. Computing Spatial Similarity by Games. In F. Esposito, editor, *Conference of the Italian Association for Artificial Intelligence (AI\*IA-01)*, number 2175 in Lecture Notes in Artificial Intelligence, pages 99–110. Springer, 2001.
- [3] M. Aiello. A Spatial Similarity Measure based on Games: Theory and Practice. *Journal of the IGPL*, 10(1):1–22, 2002.
- [4] M. Aiello. Document Image Analysis via Model Checking. *AI\*IA Notizie*, XV(1):45–48, 2002.
- [5] M. Aiello. Lo Isntitute for Logic, Language and Computation dell Università di Amsterdam. *AI\*IA Notizie*, XV(3):31–32, 2002. In Italian.
- [6] M. Aiello. *Spatial Reasoning: Theory and Practice*. PhD thesis, ILLC, University of Amsterdam, 2002. DS-2002-02.
- [7] M. Aiello, C. Monz, L. Todoran, and M. Worring. Document Understanding for a Broad Class of Documents. *International Journal of Document Analysis and Recognition IJDAR*, 5(1):1–16, 2002.
- [8] M. Aiello and A. Smeulders. Thick 2D Relations for Document Understanding. *Information Sciences*, 2004. To appear.
- [9] M. Aiello and J. van Benthem. A Modal Walk through Space. *Journal of Applied Non-Classical Logics*, 12(3–4):319–363, 2002.
- [10] M. Aiello and J. van Benthem. *Logical Patterns in Space*, pages 5–25. In Barker-Plummer et al. [16], 2002.
- [11] M. Aiello, J. van Benthem, and G. Bezhanishvili. Reasoning about Space: the Modal Way. *Journal of Logic and Computation*, 2004. To appear.
- [12] C. Areces. *Logic Engineering: The Case of Description and Hybrid Logics*. PhD thesis, University of Amsterdam, 2000.
- [13] N. Asher and L. Vieu. Toward a Geometry of Common Sense: a semantics and a complete axiomatization of mereotopology. In *IJCAI95*, pages 846–852. International Joint Conference on Artificial Intelligence, 1995.
- [14] Ph. Balbiani. The modal multilogic of geometry. *Journal of Applied Non-Classical Logics*, 8:259–281, 1998.
- [15] 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.
- [16] D. Barker-Plummer, D. Beaver, J. van Benthem, and P. Scotto di Luzio, editors. *Words, Proofs, and Diagrams*. CSLI, Stanford, 2002.

- [17] B. Bennett. Modal Logics for Qualitative Spatial Reasoning. *Bulletin of the IGPL*, 3:1 – 22, 1995.
- [18] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, 2001.
- [19] I. Bloch. Using mathematical morphology operators as modal operators for spatial reasoning. In *ECAI 2000, Workshop on Spatio-Temporal Reasoning*, pages 73–79, 2000.
- [20] R. Casati and A. Varzi. *Parts and Places*. MIT Press, 1999.
- [21] B. Clarke. A calculus of individuals based on ‘connection’. *Notre Dame Journal of Formal Logic*, 23(3):204 – 218, 1981.
- [22] M. Dastani, R. Scha, and B. Indurkha. An algebraic method for solving proportional analogy problems involving sequential patterns. In *Mind II: Computational Models of Creative Cognition*, pages 1–15, 1997.
- [23] M. de Rijke. *Extended Modal Logic*. PhD thesis, ILLC, University of Amsterdam, 1993.
- [24] K. Doets. *Basic Model Theory*. CSLI Publications, Stanford, 1996.
- [25] C. Gurr. Theories of Visual and Diagrammatic Reasoning: Foundational Issues. In *Formalizing Reasoning with Diagrammatic and Visual Representations*, Orlando, Florida, October 1998. AAAI Fall Symposium.
- [26] E. Hammer. *Logic and Visual Information*. CSLI and FoLLi, Stanford, 1995.
- [27] A. Herskovits. Language, spatial cognition, and vision. In O. Stock, editor, *Spatial and Temporal Reasoning*, pages 155–202. Kluwer, 1997.
- [28] J. Kamp. *Tense logic and the theory of linear order*. PhD thesis, University of California, Los Angeles, 1968.
- [29] O. Lemon. *Comparing the Efficacy of Visual Languages*, pages 47–69. In Barker-Plummer et al. [16], 2002.
- [30] O. Lemon and I. Pratt. Complete Logics for QSR: A Guide to Plane Mereotopology. *Journal of Visual Languages and Computing*, 9(1):5–21, 1998.
- [31] S. Lesniewski. The foundation of mathematics. *Topoi*, 2:7–52, 1983. Translation of the paper of 1931.
- [32] J. McKinsey and A. Tarski. The Algebra of Topology. *Annals of Mathematics*, 45:141–191, 1944.
- [33] D. Park. Concurrency and Automata on Infinite Sequences. In *Proceedings of the 5th GI Conference*, pages 167–183, Berlin, 1981. Springer Verlag.
- [34] I. Pratt and D. Schoop. A complete axiom system for polygonal mereotopology. *Journal of Philosophical Logic*, 27:621–658, 1998.
- [35] D. Randell, Z. Cui, and A. Cohn. A Spatial Logic Based on Regions and Connection. In *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning (KR’92)*, pages 165–176. San Mateo, 1992.
- [36] D. Randell, M. Witkowski, and M. Shanahan. From Images to Bodies: Modelling and Exploiting Occlusion and Motion Parallax. In *Proc. of Int. Joint Conference on Artificial Intelligence (IJCAI-01)*, 2001.
- [37] H. Rasiowa and R. Sikorski. *The Mathematics of Metamathematics*. Panstwowe Wydawnictwo naukowe, 1963.
- [38] K. Segerberg. “Somewhere else” and “Some other time”. In *Wright and Wrong*, pages 61–64, 1976.
- [39] J. Serra. *Image Analysis and Mathematical Morphology*. Academic Press, 1982.
- [40] M. Shanahan. Default Reasoning about Spatial Occupancy. *Artificial Intelligence*, 74(1):147–163, 1995.
- [41] V. Shehtman. Modal Logics of Domains on the Real Plane. *Studia Logica*, 42:63–80, 1983.
- [42] A. Tarski. Der Aussagenkalkül und die Topologie. *Fund. Math.*, 31:103–134, 1938.
- [43] A. Tarski. What is Elementary Geometry? In L. Henkin and P. Suppes and A. Tarski, editor, *The Axiomatic Method, with Special Reference to Geometry and Physics*, pages 16–29. North-Holland, 1959.
- [44] J. van Benthem. *Modal Correspondence Theory*. PhD thesis, University of Amsterdam, 1976.
- [45] J. van Benthem. *The Logic of Time*. Reidel, Dordrecht, 1983.
- [46] Y. Venema. *Many-Dimensional Modal Logic*. PhD thesis, ILLC, University of Amsterdam, 1992.
- [47] Y. Venema. Points, Lines and Diamonds: a Two-Sorted Modal Logic for Projective Planes. *Journal of Logic and Computation*, 9(5):601–621, 1999.
- [48] A. Whitehead. *Process and Reality*. The MacMillan Company, New York, 1929.
- [49] Y. Winter and J. Zwarts. A Semantic Characterization of Locative PPs. In *SALT7*, 1997.