Antony Galton

**Abstract** Whereas mereology, in the strict sense, is concerned solely with the partwhole relation, mereotopology extends mereology by including also the notion of connection, enabling one to distinguish, for example, between internal and peripheral parts, and between contact and separation. Mereotopology has been developed particularly within the Qualitative Spatial Reasoning research community, where it has been applied to, amongst other areas, geographical information science and image analysis. Most research in mereotopology has assumed that the entities being studied may be subdivided without limit, but a number of researchers have investigated mereotopological structures based on discrete spaces in which entities are built up from atomic elements that are not themselves subdivisible. This chapter presents an introductory treatment of mereotopology and its discrete variant, and provides references for readers interested in pursuing this subject in further detail.

# **1** From Mereology to Mereotopology

Mereology, as the theory of parts and wholes, leads to a set of five jointly exhaustive and pairwise disjoint (JEPD) relations that may hold between any pair of entities X and Y that come under its purview, namely

X is a proper part of Y	PP(x,y)
X coincides with Y	EQ(x,y)
X partially overlaps Y	PO(x,y)
X contains Y as a proper part	PPI(x,y)
X is disjoint from Y	DR(x,y)

Antony Galton

University of Exeter, UK e-mail: apgalton@ex.ac.uk

For the logical development, we first stipulate that the primitive relation of parthood (P) is reflexive and transitive:<sup>1</sup>

$$\mathsf{P}(\mathsf{x},\mathsf{x})\,,\tag{A1}$$

$$\mathsf{P}(\mathsf{x},\mathsf{y})\wedge\mathsf{P}(\mathsf{y},\mathsf{z})\to\mathsf{P}(\mathsf{x},\mathsf{z})\;. \tag{A2}$$

We then define overlap as possession of a common part:

$$O(x,y) =_{def} \exists z (P(z,x) \land P(z,y))$$
(D1)

and go on to define the five relations listed above as follows:

$$\mathsf{PP}(\mathsf{x},\mathsf{y}) =_{\mathsf{def}} \mathsf{P}(\mathsf{x},\mathsf{y}) \land \neg \mathsf{P}(\mathsf{y},\mathsf{x}) \,, \tag{D2}$$

$$\mathsf{EQ}(\mathsf{x},\mathsf{y}) =_{\mathsf{def}} \mathsf{P}(\mathsf{x},\mathsf{y}) \land \mathsf{P}(\mathsf{y},\mathsf{x}) , \qquad (D3)$$

$$\mathsf{PO}(\mathsf{x},\mathsf{y}) =_{\mathsf{def}} \mathsf{O}(\mathsf{x},\mathsf{y}) \land \neg \mathsf{P}(\mathsf{x},\mathsf{y}) \land \neg \mathsf{P}(\mathsf{y},\mathsf{x}) \,, \tag{D4}$$

$$\mathsf{PPI}(\mathsf{x},\mathsf{y}) =_{\mathsf{def}} \mathsf{PP}(\mathsf{y},\mathsf{x}) \,, \tag{D5}$$

$$\mathsf{DR}(\mathsf{x},\mathsf{y}) =_{\mathrm{def}} \neg \mathsf{O}(\mathsf{x},\mathsf{y}) \,. \tag{D6}$$

This system of relations is known in the Qualitative Spatial Reasoning (QSR) community as RCC5, the five-element Region Connection Calculus.<sup>2</sup>

If the terms of the formal language are interpreted as referring to *spatial* entities, which we here call *regions*, it is generally felt that mereology alone does not provide sufficient expressive power to be useful for QSR. In addition to parthood and the relations derived from it, we need also to be able to distinguish between, on the one hand, *internal* and *peripheral* parts, and on the other, between *contact* and *separation*. To express these, a primitive relation C (for contact, or *connection*) is introduced, and stipulated to be reflexive and symmetric:

$$C(x,x)$$
, (A3)

$$C(x,y) \rightarrow C(y,x)$$
. (A4)

The minimal relationship between P and C is that anything connected to an entity is automatically connected to anything that entity is part of:

$$\mathsf{P}(x,y) \to \forall z(\mathsf{C}(z,x) \to \mathsf{C}(z,y)) \;. \tag{A5}$$

Using P and C as primitives, we can now define a number of important additional relations as follows:

<sup>&</sup>lt;sup>1</sup> We adopt the usual convention in presenting first-order theories that free variables in formulae presented as axioms or theorems are understood to be universally quantified, so that, e.g.,  $P(x,y) \land P(y,z) \rightarrow P(x,z)$  is to be read as if it were written  $\forall x \forall y \forall z (P(x,y) \land P(y,z) \rightarrow P(x,z))$ .

 $<sup>^2</sup>$  The Region Connection Calculus was introduced, though not under that name, in [16]. The version there presented is RCC8, described below; explicit recognition of RCC5 in QSR came later. Strictly speaking, this set of mereological relations should only be called a connection calculus if they are defined in terms of connection rather than, as here, in terms of parthood.

• X is *disconnected* from Y:

$$\mathsf{DC}(\mathsf{x},\mathsf{y}) =_{\mathsf{def}} \neg \mathsf{C}(\mathsf{x},\mathsf{y}) . \tag{D7}$$

• X is *externally connected* to Y:

$$\mathsf{EC}(\mathsf{x},\mathsf{y}) =_{\mathsf{def}} \mathsf{C}(\mathsf{x},\mathsf{y}) \land \neg \mathsf{O}(\mathsf{x},\mathsf{y}) \,. \tag{D8}$$

• X is a *tangential part* of Y:

$$\mathsf{TP}(\mathsf{x},\mathsf{y}) =_{\mathrm{def}} \mathsf{P}(\mathsf{x},\mathsf{y}) \land \exists \mathsf{z}(\mathsf{C}(\mathsf{x},\mathsf{z}) \land \neg \mathsf{O}(\mathsf{z},\mathsf{y})) \ . \tag{D9}$$

(i.e., X is a part of Y that is connected to something disjoint from Y).

• X is a *non-tangential part* of Y:

$$\mathsf{NTP}(\mathsf{x},\mathsf{y}) =_{\mathrm{def}} \mathsf{P}(\mathsf{x},\mathsf{y}) \land \forall \mathsf{z}(\mathsf{C}(\mathsf{x},\mathsf{z}) \to \mathsf{O}(\mathsf{z},\mathsf{y})) \;. \tag{D10}$$

(i.e., X is a part of Y that is only connected to things that overlap Y).

Note that any part of a region must be either a tangential part or a non-tangential part of it, but not both. In particular, a region is a non-tangential part of itself if and only if it is not connected to any region disjoint from it and is therefore a union of one or more connected components of the whole space.

The system of eight JEPD relations known as RCC8 comprises DC, EC, PO, EQ, TPP (defined as the conjunction of PP and TP), NTPP (the conjunction of PP and NTP), and the inverses of TPP and NTPP.

The logical language here is denoted  $\mathscr{L}_{P,C}$ , and comprises all first-order formulae in which the non-logical language is restricted to the two binary predicates P and C — all formulae containing the other RCC8 relations being reducible to formulae containing just P and C, via the definitions given above. Systems of this kind, which combine the mereological notion of parthood with the topological notion of connection, are called *mereotopologies*.

Mereotopologies are normally interpreted as referring to regions which can be indefinitely subdivided. This is expressed by positing the formula

$$\exists y \mathsf{PP}(y, x)$$
 (N1)

as an axiom. The domain of such an interpretation is usually taken to be some collection of non-empty subsets of  $\mathbb{R}^n$  for some positive integer *n* (typically either 2 or 3). In order to ensure infinite subdivisibility, only infinite subsets should be considered as possible domain elements, but this still leaves open many different possible such collections, for example

- all infinite subsets of  $\mathbb{R}^n$
- all non-empty open subsets of  $\mathbb{R}^n$
- all non-empty regular open<sup>3</sup> subsets of  $\mathbb{R}^n$

<sup>&</sup>lt;sup>3</sup> A regular open set is a set that is equal to the interior of its closure; a regular closed set is equal to the closure of its interior.

- all non-empty regular closed subsets of  $\mathbb{R}^n$  [6]
- all open polygonal (polyhedral, etc) subsets of  $\mathbb{R}^n$  [13, 14, 15]

In these interpretations, it is usual to interpret the predicates P and C as standing for the following relations:

- *X* is part of *Y* if and only if  $X \subseteq Y$ .
- X is connected to Y if and only if  $\overline{X} \cap \overline{Y} \neq \emptyset$ , i.e., the topological closures of X and Y have at least one point in common.

The interpretation of parthood in terms of the subset relation explains why the domain has to be restricted to non-empty sets: if the empty set were allowed, then any pair of regions would overlap, since they would have the empty set as a common part.

It should be noted that under any of these interpretations, parthood is necessarily *antisymmetric*, satisfying the formula

$$\mathsf{P}(\mathsf{x},\mathsf{y}) \land \mathsf{P}(\mathsf{y},\mathsf{x}) \to \mathsf{x} = \mathsf{y} , \qquad (A6)$$

and thus *extensional*, meaning that two distinct entities cannot have exactly the same parts:

$$\forall z(P(z,x) \leftrightarrow P(z,y)) \rightarrow x = y.$$
 (T1)

From now on we shall assume that P denotes an antisymmetric relation; a consequence of this is that EQ(x, y) becomes equivalent to x = y, meaning that the symbol EQ can be dropped.

Connection, on the other hand, need not be extensional: that is, it does not necessarily follow that two entities are identical if they are connected to exactly the same things. In the first two models above, connection is not extensional; for example, if the domain of discourse consists of all infinite subsets of  $\mathbb{R}^n$ , an open set and its closure are connected to exactly the same sets, yet they are not identical. In the last three models listed above, connection *is* extensional, that is, they satisfy

$$\forall z(C(x,z) \leftrightarrow C(y,z)) \rightarrow x = y.$$
(N2)

In such models, if X is connected to everything Y is connected to, then X is part of Y, which means that the converse of (A5) holds. In this case, parthood can be characterised exactly in terms of connection, as follows:

$$\mathsf{P}(\mathsf{x},\mathsf{y}) \leftrightarrow \forall \mathsf{z}(\mathsf{C}(\mathsf{x},\mathsf{z}) \to \mathsf{C}(\mathsf{y},\mathsf{z})) \ . \tag{N3}$$

It is easy to see that (N2) and (N3) together imply (A6) and hence (T1). If we have (N3), we can use it to *define* P, leaving just the one primitive predicate C. In this case the logical language can be reduced to  $\mathscr{L}_C$ .

It should be noted that although individual terms in  $\mathscr{L}_{P,C}$  refer to regions, by interpreting them as denoting subsets of  $\mathbb{R}^n$  we are implicitly postulating a universe of points, even though these cannot be referred to in the language. To avoid this unsatisfactory situation, Stell [24] showed that the terms of RCC could be interpreted

4

as referring to elements of structures called Boolean Connection Algebras (BCAs), which do not presuppose that regions are collections of points. This notion was generalised by Li and Ying [12] to Generalized Boolean Connection Algebras, which as well as subsuming Stell's BCAs can provide models for the discrete versions of RCC which we turn to next.

# 2 Discrete Mereotopology and Adjacency Spaces

If we wish to interpret the mereotopological predicates over *discrete* domains, in which entities are *not* indefinitely subdivisible, it is no longer possible to define parthood in terms of connection. In a discrete domain, every entity decomposes into *atoms*, which have no proper parts. We define

$$Atom(x) =_{def} \neg \exists y \mathsf{PP}(y, x) \tag{D11}$$

Then it can be seen that (N1) is equivalent to  $\neg \exists x Atom(x)$ .

If A is an atom which is connected to its complement  $A^c$ , then anything connected to A must either be A itself or overlap  $A^c$ , and hence in either case must be connected to  $A^c$ ; but on the other hand A is obviously *not* part of its complement,<sup>4</sup> so we have

$$\forall z(C(a,z) \rightarrow C(b,z)) \land \neg P(a,b)$$

(where a and b denote A and  $A^c$  respectively), contradicting (N3).

For *discrete* mereotopology, then, both P and C are needed as primitive predicates in the logical language. How should they be interpreted? The subsets of  $\mathbb{R}^n$  listed above are no longer appropriate, and an obvious substitute here would be to use subsets of  $\mathbb{Z}^n$ , i.e., sets of points with integer coordinates. How should C be interpreted in this case? Since overlap is a form of connection,<sup>5</sup> we need only concern ourselves with the interpretation of non-overlapping connection, i.e., EC.

An example is illustrated in Fig. 1, where the atomic regions are shown as unit squares, which can be mapped in the obvious way to elements of  $\mathbb{Z}^2$ . The external connection between the two differently shaded regions depends on the fact that the squares labelled 'a' and 'b', one from each region, are *adjacent* to each other, and it is this notion of adjacency which forms the basis for a general way of interpreting the connection predicate in discrete mereotopologies. We do not confine ourselves to subsets of  $\mathbb{Z}^n$  but rather to regions defined over a more general class which we define as follows:

<sup>&</sup>lt;sup>4</sup> At least, this is obvious so long as "part" and "complement" are understood in the usual sense; however, Roy and Stell [21] showed that by replacing the ordinary set-theoretical complement operation by a weaker operation, the dual pseudo-complement, defined over a class of structures called dual p-algebras, one obtains a model of discrete space in which (N3) holds.

 $<sup>^5</sup>$  It follows from (A5) that  $\forall x \forall y (O(x,y) \rightarrow C(x,y)).$ 

#### Antony Galton



Fig. 1 External connection in discrete space

**Definition:** An *adjacency space* is a non-empty set *U* of entities called *cells* together with a reflexive, symmetric relation  $\sim \subseteq U \times U$ , called *adjacency*.

An adjacency space can be regarded as a graph; but this does not mean that the theory of adjacency spaces is identical to graph theory. An important difference emerges when we consider substructures. A graph is specified by a set V of vertices and a set E of edges, where each edge joins an element of V to an element of V. A subgraph is specified by a subset  $V' \subseteq V$  of the vertices and a subset  $E' \subseteq E$  of the edges, with the proviso that each edge in E' joins an element of V' to an element of V'. There is no requirement that an edge in E which happens to join an element of V' to an element of V' nust be in E'. Thus there can be many different subgraphs of (V, E) all of which have the vertex-set V'. In adjacency spaces, the adjacency or otherwise of two cells is fixed by the *space* and is automatically inherited by the substructures. Thus a substructure of an adjacency space can be specified by giving its cells alone, without reference to adjacency.

These substructures, which may be thought of as aggregates of cells, are called *regions*, and it is these entities that discrete mereotopology is primarily concerned with, not the cells themselves. A cell might, indeed, be considered to be the aggregate which is composed of precisely that cell and nothing else; but for theoretical purposes it is convenient to specify a region in terms of the (non-empty) *set* of cells which make it up, and in that case a one-cell region is conceptually distinct from its only cell, the former being, in fact, the singleton set of the latter. Therefore an interpretation *I* of the logical language  $\mathscr{L}_{\mathsf{P},\mathsf{C}}$  over an adjacency space  $(U, \sim)$  is specified as follows:

- Each individual term t of the logical language denotes a non-empty subset  $t^I \subseteq U$ .
- A formula  $P(t_1, t_2)$  is interpreted to mean that  $t_1^I \subseteq t_2^I$ .
- A formula  $C(t_1, t_2)$  is interpreted to mean that there are cells  $x \in t_1^I$  and  $y \in t_2^I$  such that  $x \sim y$ .

Thus two regions are regarded as connected so long as some cell in one is adjacent to some cell in the other.<sup>6</sup>

<sup>&</sup>lt;sup>6</sup> These ideas were presented, without explicit use of the term "adjacency space", in [4]. The term "adjacency space" was used in [5].

The theory of Discrete Mereotopology (DM), as we shall understand it in this paper, comprises all and only those formulae of the language  $\mathscr{L}_{P,C}$  which are satisfied by every adjacency space under the scheme of interpretation just specified. It is easy to see that DM includes all the formulae thus far introduced with labels beginning with either A or T: the formulae (A1), (A2), (A6), and (T1) characterising parthood, (A3) and (A4) characterising connection, and (A5) relating parthood and connection. <sup>7 8</sup>

DM does *not* include the converse of (A5), meaning that the definitional reduction of P to C given by (N3) is not available here. The other important non-theorem is (N1), which expresses the infinite subdivisibility of regions that is characteristic of non-discrete (dense or continuous) models; instead, DM includes the formula

$$\forall \mathsf{x} \exists \mathsf{y} (\mathsf{P}(\mathsf{y},\mathsf{x}) \land \mathsf{Atom}(\mathsf{y})) , \tag{A7}$$

which says that every region has an atomic region as part. The predicate Atom is clearly satisfied by just the singleton subsets of the universe U; and *every* subset of U has at least one singleton subset; under the interpretation, these two sets count as a region and an atomic part of that region.

It will be convenient to introduce a predicate AP to say that one region is an atomic part of another; this is straightforwardly defined as follows:

$$AP(x,y) =_{def} Atom(x) \land P(x,y), \qquad (D12)$$

enabling us to rewrite (A7) as  $\forall x \exists y AP(y,x)$ .

The mereotopological relations of atoms are much simpler than those of general regions. In particular, if A overlaps B, where A is an atom, then the common part of A and B cannot be a proper part of A and must therefore be A itself. We thus have

$$Atom(x) \to (O(x, y) \to P(x, y)). \tag{T2}$$

An important mereological principle, which forms part of General Extensional Mereology, is the *Strong Supplementation Principle* [22, p.29], which states that any region that is *not* part of a given region must have a part that does not overlap that region:

$$\neg \mathsf{P}(\mathsf{y},\mathsf{x}) \to \exists \mathsf{z}(\mathsf{P}(\mathsf{z},\mathsf{y}) \land \neg \mathsf{O}(\mathsf{z},\mathsf{x})) \tag{A8}$$

<sup>&</sup>lt;sup>7</sup> Of course, as written, not all of these are  $\mathscr{L}_{P,C}$  formulae; they become so when the predicates other than P and C are expanded in accordance with their definitions, given by the formulae whose labels begin with D.

<sup>&</sup>lt;sup>8</sup> Formulae whose labels beginning with T logically follow from those with labels beginning with A; thus the latter can be regarded as *axioms* and the former as *theorems*. However, the distinction is somewhat arbitrary (since there are in principle many different ways of assigning "A" and "T" labels) and only comes into its own when we wish to consider to what extent reasoning about adjacency spaces can be accomplished purely by means of symbolic manipulation of  $\mathscr{L}_{P,C}$  formulae, without reference to any interpretation. In that case it becomes of interest whether or not there is a finitely-specifiable set of  $\mathscr{L}_{P,C}$  formulae whose logical consequences comprise all and only the true formulae of discrete mereotopology — in short, whether this theory can be completely axiomatised.

In combination with (A7), this leads to a powerful extensionality principle for DM, namely

$$\forall z (AP(z, x) \leftrightarrow AP(z, y))) \rightarrow x = y. \tag{T3}$$

Whereas, in order to show that two regions are the same, (T1) requires us check that they agree in *all* their parts, with (T3) it suffices to check that they agree in just their atomic parts.

To see how this follows from (A8), suppose we have

$$\forall z (AP(z, x) \leftrightarrow AP(z, y)). \tag{(*)}$$

We must show that x = y. Suppose not; then by (A6), either  $\neg P(x,y)$  or  $\neg P(y,x)$ . Without loss of generality we may assume the former. By (A8), this means there is a region u such that

$$\mathsf{P}(\mathsf{u},\mathsf{x}) \land \neg \mathsf{O}(\mathsf{u},\mathsf{y}) \,. \tag{**}$$

By (A7), there is a region v such that AP(v, u), and by transitivity therefore AP(v, x). By (\*) this means that AP(v, y). We now have  $P(v, u) \wedge P(v, y)$ , so O(u, y), which contradicts (\*\*).

It is clear that the Strong Supplementation Principle is satisfied when regions are interpreted as subsets of an adjacency space, so both (A8) and (T3) belong to DM. The upshot of this is that we can define a region uniquely by characterising its atomic parts. We will use this in the following fashion: if a predicate  $\phi$  is defined by a rule of the form

$$\phi(\mathsf{x}) =_{\operatorname{def}} \forall \mathsf{z}(\operatorname{\mathsf{Atom}}(\mathsf{z}) \to (\mathsf{P}(\mathsf{z},\mathsf{x}) \leftrightarrow \psi(\mathsf{z},\mathsf{x})))$$

then it follows that any two regions with the property  $\phi$  are identical.

The formulae (A1), (A2), (A6), and (A8) together constitute the axiomatic basis for the system designated **EM** (for Extensional Mereology) in [26]. The addition of (A7) yields Atomic Extensional Mereology **AEM**.

The study of discrete mereotopology can be pursued on two levels, which we may loosely characterise as "set-theoretical" and "logical". At the set-theoretical level, the class of adjacency spaces are treated as mathematical objects in their own right, independently of any particular logical language chosen for describing them. At the logical level, on the other hand, one focusses on the particular first-order language  $\mathscr{L}_{P,C}$ , which is the common language in which to express mereotopological theses, regardless of whether discrete or continuous spaces are intended. The set-theoretical level provides a metalanguage within which one can specify interpretations of the logical level. While much of what is said at one level can be transposed easily to the other, it is important to maintain a clear conceptual distinction between them. This is supported here by a typographical distinction: formulae at the logical level are always printed in a sanserif font.

Moving back and forth between the levels we can investigate what formulae of  $\mathscr{L}_{P,C}$  are satisfied in adjacency spaces (these formulae constituting the theory of DM), and conversely which properties of adjacency spaces can be expressed in the language. We can then ask whether the theory of adjacency spaces, insofar as it can

8

be expressed in  $\mathscr{L}_{P,C}$ , is axiomatisable, i.e., whether there is a finitely specifiable set of  $\mathscr{L}_{\mathsf{P},\mathsf{C}}$  formulae from which all and only the true theorems of DM follow as logical consequences.

In the remainder of this paper we present a few of the most important features of DM, briefly discuss its relation to some other approaches, and describe an area in which it is being applied.

#### **3 Examples of Adjacency Spaces**

The adjacency space in Fig. 1 is underdetermined, in that we did not specify how the relation  $\sim$  was to be defined. It was assumed that the reader would naturally understand that the cells labelled 'a' and 'b' were to count as adjacent. In fact there are (at least) two different, and equally natural, ways of understanding adjacency in  $\mathbb{Z}^2$ . Under *orthogonal adjacency*, only cells which share an edge count as adjacent; thus each cell is adjacent to four cells other than itself. We denote this relation  $\sim_4$ . Orthodiagonal adjacency is where cells count as adjacent so long as they share at least one boundary point - either along an edge or at a corner; each cell is adjacent to eight others, and hence we denote this relation  $\sim_8$ . These two adjacency relations are defined as follows:

- $(x,y) \sim_4 (x',y')$  iff  $|x-x'| + |y-y'| \le 1$ .  $(x,y) \sim_8 (x',y')$  iff  $|x-x'| \le 1$  and  $|y-y'| \le 1$ .

Both of these spaces are homogeneous in the sense that all cells "look the same"; more formally, for any  $x, y \in U$  there is a bijective adjacency-preserving function from *U* to *U* which maps *x* onto *y*.

The adjacency spaces  $(\mathbb{Z}^2, \sim_4)$  and  $(\mathbb{Z}^2, \sim_8)$  have played a prominent part in work on discrete spaces, mainly because they are the most natural spaces within which to model *digital pictures*, as seen, for example, on a computer screen in which the display is produced by assigning colour values to each element in a rectangular array of pixels (see, e.g., [20, 8]).

Other homogeneous adjacency spaces correspond to tessellations of triangles or hexagons. In the triangular case, there are two possible adjacency relations:  $X \sim_3 Y$ if triangles X and Y share an edge, and  $X \sim_{12} Y$  if they share at least one boundary point. With the hexagonal lattice there is only one natural adjacency relation,  $\sim_6$ , which holds between hexagons that share an edge. All these cases are illustrated in Fig. 2.

Homogeneous adjacency spaces do not need to be infinite. Familiar examples of finite spaces are provided by the five platonic solids. The faces of a dodecahedron, for example, can be thought of as a 12-element adjacency space, where adjacency is interpreted as edge-sharing between the pentagonal faces.

Beyond these examples, adjacency spaces do not have to be homogeneous. Nonhomogeneous tessellations include the triangulated irregular networks (TIN) used in Geographical Information Science. An example is shown in Fig. 3(a). As with the



Fig. 2 Adjacency relations in regular planar tessellations

homogeneous tessellation of squares, two kinds of adjacency can be defined on a TIN: either adjacency along edges only, or adjacency at edges and vertices. The advantage of a hexagonal tessellation is that this ambiguity does not arise, and for this reason we will use such tessellations for our illustrative examples in what follows — though for practical convenience, the hexagonal tessellation will be represented in the form of an isomorphic "staggered squares" grid, as shown in Fig. 3(b).



**Fig. 3** (a) A triangular irregular network (TIN). (b) A grid of staggered squares, isomorphic to the regular hexagonal tessellation.

# 4 Mereotopological Relations on Adjacency Spaces

Given the interpretation of the relations P and C over an adjacency space, the interpretations of all the relations defined in terms of P and C, such as the RCC8 relations, become fixed. Following the standard convention in model theory, we write  $(U, \sim) \models R[X, Y]$  to mean that the relation denoted by a predicate *R* (defined in  $\mathscr{L}_{P,C}$ ) holds between the elements  $X, Y \in U$ . Thus for example we have, for regions X, Y:

- $(U, \sim) \models \mathsf{DC}[X, Y]$  iff there are no cells  $x \in X$  and  $y \in Y$  such that  $x \sim y$ .
- $(U, \sim) \models \mathsf{EC}[X, Y]$  iff  $X \cap Y = \emptyset$  and there are cells  $x \in X$  and  $y \in Y$  such that  $x \sim y$ .
- $(U, \sim) \models \mathsf{TP}[X, Y]$  iff  $X \subseteq Y$  and there are cells  $x \in X$  and  $y \notin Y$  such that  $x \sim y$ .
- $(U, \sim) \models \mathsf{NTP}[X, Y]$  iff for all cells  $x \in X$ , if  $x \sim y$  then  $y \in Y$ .
- $(U, \sim) \models \mathsf{EQ}[X, Y]$  iff X = Y.

Examples of all the RCC8 relations are illustrated in Fig. 4, using regions defined on the "staggered squares" grid.



Fig. 4 RCC8 relations in an adjacency space

It should be noted that *every* region is a non-tangential part of U since the consequent of the defining condition,  $y \in Y$ , is always true when Y = U.

The set of subsets of *U* forms a Boolean algebra under the usual set-theoretic operations of union, intersection, and complement, with *U* itself acting as the top element (generally notated 1 or  $\top$ ) and  $\emptyset$  as the bottom element (notated 0 or  $\bot$ ). Since  $\emptyset$  is not a region, the regions just fall short of being a Boolean algebra: they form a *quasi-Boolean algebra*. In mereotopology, therefore, the Boolean operations are appropriately restricted so that neither their range nor their domain contains the empty set, as we show below.

The universe U can be characterised in  $\mathscr{L}_{P,C}$  as the region which every other region is part of. This can be expressed by the predicate U, defined by

$$U(x) =_{def} \forall y P(y, x) \tag{D13}$$

That such a region exists is stated by the formula

$$\exists \mathsf{x}\mathsf{U}(\mathsf{x}),\tag{A9}$$

which is generally accepted as an axiom in RCC, whether in a discrete or continuous setting. By antisymmetry of P, it straightforwardly follows from this that there can be at most one universal region; in terms of adjacency structures we have

$$(U, \sim) \models \mathsf{U}[X]$$
 if and only if  $X = U$ .

We can therefore introduce a constant symbol U to denote the unique universal region, defined contextually as follows:

$$\phi(\mathsf{U}) =_{\mathrm{def}} \forall \mathsf{x}(\mathsf{U}(\mathsf{x}) \to \phi(\mathsf{x})) , \qquad (D14)$$

where  $\phi$  stands for any open formula with one free variable.

Since not all pairs of regions have a Boolean product (intersection), we cannot represent it by a function symbol in  $\mathscr{L}_{P,C}$ ; instead we define a relational predicate Prod, the intended meaning of Prod(x, y, z) being that z is the intersection of x and y, defined as follows:

$$\mathsf{Prod}(\mathsf{x},\mathsf{y},\mathsf{z}) =_{\mathsf{def}} \forall \mathsf{v}(\mathsf{P}(\mathsf{v},\mathsf{z}) \leftrightarrow \mathsf{P}(\mathsf{v},\mathsf{x}) \land \mathsf{P}(\mathsf{v},\mathsf{y})) \ . \tag{D15}$$

Similarly, the Boolean sum (union) is defined by

$$Sum(x, y, z) =_{def} \forall v (O(v, z) \leftrightarrow O(v, x) \lor O(v, y)), \qquad (D16)$$

and the Boolean difference by

$$\mathsf{Diff}(\mathsf{x},\mathsf{y},\mathsf{z}) =_{\mathsf{def}} \forall \mathsf{v}(\mathsf{P}(\mathsf{v},\mathsf{z}) \leftrightarrow \mathsf{P}(\mathsf{v},\mathsf{x}) \land \neg \mathsf{O}(\mathsf{v},\mathsf{y})) \ . \tag{D17}$$

The existence of regions playing the role of z in these formulae is stated by the following formulae, which also specify the conditions on x and y for such a z to exist:

$$O(x,y) \rightarrow \exists z Prod(x,y,z)$$
, (A10)

$$\exists z Sum(x, y, z), \qquad (A11)$$

$$\neg \mathsf{P}(\mathsf{x},\mathsf{y}) \to \exists \mathsf{zDiff}(\mathsf{x},\mathsf{y},\mathsf{z}) . \tag{A12}$$

As with (D13), it follows from (D15), (D16), and (D17) that products, sums, and differences, where they exist, are unique. Note that (A12), in conjunction with (A1), logically implies (A8), meaning that the latter could be relegated to the status of a theorem (with a 'T' label) rather than an axiom; however, we shall retain the designation (A8) to avoid confusion.

The complement of a non-universal region can be defined as its difference from U, i.e.,

$$Compl(x,y) =_{def} Diff(U,x,y)$$
.

Since we always have P(v, U), this may be expanded as

$$\mathsf{Compl}(\mathsf{x},\mathsf{y}) =_{\mathsf{def}} \forall \mathsf{v}(\mathsf{P}(\mathsf{v},\mathsf{y}) \leftrightarrow \neg \mathsf{O}(\mathsf{v},\mathsf{x})) . \tag{D18}$$

It is easy to show that, for non-empty sets  $X, Y \subseteq U$ ,  $(U, \sim) \models \mathsf{Compl}[X, Y]$  if and only if  $X = Y^c$ , thus ensuring that the  $\mathscr{L}_{\mathsf{P},\mathsf{C}}$ -definable predicate Compl captures the set-theoretic relation of complementation insofar as it applies to regions in adjacency spaces. It does not *immediately* follow from this, of course, that Compl behaves like complementation in arbitrary models of DM, but that this is so is shown by the following theorem:

$$\mathsf{Compl}(\mathsf{x},\mathsf{y}) \leftrightarrow \neg \mathsf{O}(\mathsf{x},\mathsf{y}) \land \mathsf{Sum}(\mathsf{x},\mathsf{y},\mathsf{U}) , \tag{T4}$$

which says that one region is the complement of another if and only if they are disjoint regions whose sum is the universe. A corollary of this is that complementarity is mutual:

$$Compl(x,y) \rightarrow Compl(y,x)$$
 (T5)

The proofs of these theorems are given in Appendix 1.

With the addition of (A9), (A10), (A11), and (A12) to **EM** we obtain the system of Closed Extensional Mereology designated **CEM** by Varzi [26]. Li and Ying [12] show that every model of **CEM** is isomorphic to a complete quasi-Boolean algebra, and therefore that **ACEM**, the atomic variant of **CEM** with the additional axiom (A7) is isomorphic to an atomic complete quasi-Boolean algebra,

In standard mereotopology, where there are no atomic regions, and parthood is defined in terms of connection, the definition of these Boolean operations in  $\mathscr{L}_{P,C}$  has proved somewhat problematic, it being difficult to demonstrate that the required interrelationships hold when connection is taken into account. In particular, (D18) does not suffice to captured the desired behaviour and needs to be supplemented by an additional axiom, represented here by the formula

$$\mathsf{Compl}(\mathsf{x},\mathsf{y}) \to \forall \mathsf{z}(\mathsf{C}(\mathsf{z},\mathsf{y}) \leftrightarrow \neg\mathsf{NTP}(\mathsf{z},\mathsf{x})) \tag{T6}$$

In DM, however, it can be proved that (T6) follows from (D18) and the existing axioms, as demonstrated in Appendix 2.

# **5** Quasi-Topological Operators

The relation NTP picks out those subregions of a given region which are disconnected from the complement of the region: the neighbours of each cell in the subregion are all in the region itself. The union of all the non-tangential parts of a region thus consists of *all* the cells in the region whose neighbours are also all in the region:

$$\left| \left\{ X \mid (U, \sim) \models \mathsf{NTP}[X, R] \right\} = \left\{ x \in U \mid \forall y (x \sim y \to y \in R) \right\} \right|$$

This set is called the *(discrete) interior* of region *R* and is denoted  $int_D(R)$ . So long as it is non-empty, it is of course a region itself.

While *int<sub>D</sub>*, considered as an operator on sets of cells, is a total function, when considered as an operator on regions it is only a partial function, since if *int<sub>D</sub>*(*R*) is empty, *R* does not have an interior region. In the language  $\mathscr{L}_{P,C}$ , therefore, the notion of interior is expressed by means of a relational predicate Int(x,y), meaning that y is *an* interior of x, defined by:

$$\mathsf{Int}(\mathsf{x},\mathsf{y}) =_{\mathsf{def}} \forall \mathsf{z}(\mathsf{P}(\mathsf{z},\mathsf{y}) \leftrightarrow \mathsf{NTP}(\mathsf{z},\mathsf{x})) \tag{D19}$$

Thus a region is an interior of R if and only if its parts are all and only the nontangential parts of R. As discussed earlier, it suffices, in fact, to consider just atomic parts, so we have

$$\mathsf{Int}(\mathsf{x},\mathsf{y}) \leftrightarrow \forall \mathsf{z}(\mathsf{Atom}(\mathsf{z}) \to (\mathsf{P}(\mathsf{z},\mathsf{y}) \leftrightarrow \mathsf{NTP}(\mathsf{z},\mathsf{x}))). \tag{D19'}$$

It now follows that a region cannot have two distinct interiors, since if  $Int(x,y_1) \land Int(x,y_2)$  then from (D19) we have  $\forall z(P(z,y_1) \leftrightarrow P(z,y_2))$  so in particular  $P(y_2,y_1)$  and  $P(y_1,y_2)$ , whence  $y_1 = y_2$  by antisymmetry of P (A6). The appropriateness of D19 follows from the easily demonstrated fact that, if  $Y \neq \emptyset$ ,  $(U, \sim) \models Int[X, Y]$  if and only if  $Y = int_D(X)$ . Thus Int does capture in  $\mathscr{L}_{P,C}$  the relationship between a region and its discrete interior, so long as the latter is non-empty.

Any connected component of U, and any union of such connected components (including U itself), is its own interior, since it is a non-tangential part of itself. If a region has no non-tangential parts, and hence no interior, we describe it as "thin". In Fig. 5(a), the left-hand region has its interior shaded a darker grey; the right-hand region is thin, since all of its cells have at least one neighbour outside the region. Note that U. and any of its connected components, cannot be thin since it is its own interior; in particular, therefore, a single cell is a thin region so long as it is connected to at least one other cell, but if it is a connected component of U (and thus an isolated cell, disconnected from the rest of the space), it is not thin.



**Fig. 5** Discrete interiors and closures. In (a), a two-part region (*all shading*) and its discrete interior (*dark grey*); in (b), a region (*mid-grey*) and its discrete closure (*all shading*)

We refer to the *discrete* interior in order to distinguish  $int_D$  from the topological interior operator *int*, which does not apply to adjacency spaces since they are not defined as topologies. The two operators share a number of common features, notably:

- $int_{(D)}(U) = U$ ,
- $\forall X(int_{(D)}(X) \subseteq X)$ ,
- $\forall X \forall Y (X \subseteq Y \rightarrow int_{(D)}(X) \subseteq int_{(D)}(Y))$ .

The most important *difference* between the discrete and topological interior operators is that whereas the latter is idempotent, i.e.,

• 
$$\forall X(int(int(X)) = int(X))$$
,

this is not, in general true for the former, since, in an adjacency space, the only regions for which  $int_D(X) = X$  are unions of connected components of the space. In view of both the similarities and the differences, we call  $int_D$  a quasi-topological operator.

In topology, the *closure* of a set is defined as the complement of the interior of its complement:  $cl(X) = (int(X^c))^c$ . The analogous operation in adjacency spaces gives us another quasi-topological operator, the *discrete closure*,

$$cl_D(R) = (int_D(R^c))^c = \{x \mid \exists y(x \sim y \land y \in R)\} = \bigcap \{X \mid (U, \sim) \models \mathsf{NTP}[X, R]\}.$$

Thus the discrete closure of a region R consists of all those cells which are adjacent to an element of R; it is the intersection of all regions which R is a non-tangential part of.

As with the interiors, the discrete closure shares some properties with the topological closure, namely

• 
$$cl_{(D)}(\emptyset) = \emptyset$$

• 
$$\forall X(X \subseteq cl_{(D)}(X))$$

•  $\forall X (X \subseteq C(D)(X))$ , •  $\forall X \forall Y (X \subseteq Y \to cl_{(D)}(X) \subseteq cl_{(D)}(Y))$ ,

but not:

•  $\forall X(cl(cl(X)) = cl(X))$ .

As with interiors, the only regions in adjacency space for which  $cl_D(X) = X$  are the unions of connected components of the space.

Analogously to discrete interior, we can define the discrete closure relation in  $\mathscr{L}_{\mathsf{P},\mathsf{C}}$  by

$$\mathsf{Cl}(\mathsf{x},\mathsf{y}) =_{\mathsf{def}} \forall \mathsf{z}(\mathsf{P}(\mathsf{y},\mathsf{z}) \leftrightarrow \mathsf{NTP}(\mathsf{x},\mathsf{z})) , \qquad (\mathsf{D20})$$

which says that *y* is the closure of *x* if and only if *x* is a non-tangential part of any region *y* is part of.

Our decision to develop mereology with a universal element but no null element leads to a certain asymmetry. In set-theoretical interpretations the asymmetry shows up as the fact that the universal set is recognised as determining a region but the empty set is not. A consequence of this is that, unlike the discrete interior, discrete closure is a total function on regions: not every region has a discrete interior, but every region has a discrete closure. This means that we can also define the closure *function* cl by

$$\phi(\mathsf{cl}(\mathsf{x})) =_{\mathrm{def}} \forall \mathsf{y}(\mathsf{Cl}(\mathsf{x},\mathsf{y}) \to \phi(\mathsf{y})), \tag{D21}$$

where  $\phi$  stands for any open formula with one free variable. Fig. 5(b) shows a region (dark grey cells) and its discrete closure (the region plus the lighter grey cells).

Within  $\mathscr{L}_{P,C}$ , in order to characterise the relationship between discrete closure and interior, we need to use the predicate Compl already defined in (D18). The

relationship between discrete interior and discrete closure is then expressed in  $\mathscr{L}_{P,C}$  by the formula

$$\mathsf{Compl}(\mathsf{x},\mathsf{y}) \land \mathsf{Int}(\mathsf{y},\mathsf{z}) \land \mathsf{Compl}(\mathsf{z},\mathsf{w}) \to \mathsf{Cl}(\mathsf{x},\mathsf{w}), \tag{T7}$$

a proof of which is given in Appendix 3.

From now on, we shall use the words 'closure' and 'interior' on their own to mean the *discrete* closure and interior; if we need to refer to the topological operators, we shall do so explicitly. A useful way of characterising the closure and interior in an adjacency space is to use the idea of the *neighbourhood* of a cell, defined as follows:

$$N(x) = \{ y \in U \mid x \sim y \}$$

Thus neighbourhoods are in fact the closures of atoms, and the closure of any region  $R \subseteq U$  is the union of the neighbourhoods of its constituent cells:

$$cl_D(R) = \bigcup \{ N(x) \mid x \in R \}.$$

The interior comprises those cells whose neighbourhoods are parts of the region:

$$int_D(R) = \{x \in U \mid N(x) \subseteq R\}$$

Useful operations result from combining closure and interior, in either order. If a set X includes some thin spikelike parts, these will disappear when the interior is taken, and will not be restored if closure is then applied. Thus  $cl_D(int_D(X))$  is essentially like X but with any thin parts removed (Fig. 6, left). On the other hand, if the region has any thin holes or fissures, these will be filled in by the closure operation and not be opened out again when interior is applied. Thus  $int_D(cl_D(X))$ is like X but with any thin holes or fissures filled in (Fig. 6, right). Regions which lack spikes or fissures, i.e., for which  $X = cl_D(int_D(X)) = int_D(cl_D(X))$ , may be called *regular*.

Referring back to the earlier discussion of the distinction between adjacency spaces and graphs, it should be noted that Stell [25] has reformulated these quasitopological operators in terms of two kinds of complementation definable on graphs. Recall that a subgraph is specified by giving both its vertices and its edges. The *negation*  $\neg G$  of a subgraph *G* consists of all the vertices of *U* that are not in *G*, and all the edges of *U* joining vertices in  $\neg G$ . The *supplement*  $\sim G$  consists of all the edges of *U* that are not in *G*, and all the vertices of *V* that are incident with an edge in  $\neg G$ . Then the subgraphs  $\neg \sim G$  and  $\sim \neg G$  correspond to  $int_D(G)$  and  $cl_D(G)$  respectively. It is worth noting that Stell defines a *region* in a graph to be subgraph *G* such that  $\neg \neg G = G$ ; thus a region, in this sense, includes all the edges of *U* that join vertices of *G* to a vertices of *G*; it can therefore can be specified just by giving its vertices, and thus corresponds to a region in adjacency space.



**Fig. 6** Left: A region (*mid grey*) with its interior (*dark grey*) and the closure of the interior (*heavy outline*). Right: The same region, with its closure (*all shading*) and the interior of the closure (*heavy outline*)

# 6 Measures of Size and Distance

From now on we assume that the universe is connected, meaning that every region (other than the universe itself) is connected to its complement:

$$\forall X \subset U((U, \sim) \models \mathsf{C}[X, X^c]) .$$

This means that the universe consists of a single connected component, itself, and is therefore the only region which is its own closure and interior (the empty set also has this property, but it is not a region).<sup>9</sup>

We have already characterised a *thin* region as one with empty interior. More generally we can define the *thickness* of a region as the number of successive interior operations required to reduce the region to nothing. For a region R of thickness n we have the sequence

$$R, int_D(R), int_D(int_D(R)), \dots, int_D^n(R) = \emptyset$$
.

Thus for  $X \subseteq U$  we define

$$Thickness(X) =_{def} \begin{cases} 0 & (\text{if } X = \emptyset) \\ n+1 & (\text{if } X \neq \emptyset \text{ and } Thickness(int_D(X)) = n) \\ \infty & (\text{Otherwise}) \end{cases}$$

The thin regions are then those with thickness 1. The universe, U, since it is its own interior, can never be reduced to nothing in this way, so its thickness is infinite

<sup>&</sup>lt;sup>9</sup> In standard treatments of mereotopology this is posited as an axiom:  $Compl(x,y) \rightarrow C(x,y)$ .

— even if it is finite in the sense of containing only finitely many cells.<sup>10</sup> Regions other than the universe can only have infinite thickness if the universe is infinite. Thickness provides a measure of how 'substantial' a region is.

A region of thickness *n* is the union of a sequence of n - 1 shells surrounding a central *core* of thickness 1. If *Thickness*(*R*) = *n*, then the shells of region *R* are

$$R \setminus int_D(R), int_D(R) \setminus int_D(int_D(R)), \dots, int_D^{n-2}(R) \setminus int_D^{n-1}(R), int_D^{n-1}(R)$$

where the final term in the sequence is the core, whose interior is empty.

We can perform an analogous construction, starting from a region *R* and repeatedly forming the closure, thus building up a sequence of regions:

$$R, cl_D(R), cl_D(cl_D(R)), cl_D(cl_D(cl_D(R))), \ldots$$

If the universe is infinite, this process may go on for ever, depending on the starting point *R*, otherwise, given that the universe is connected, we will reach a point where  $cl_D^n(R) = U$ . Give the complementarity of closure and interior, this occurs when *Thickness*( $R^c$ ) = *n*. We can think of the sequence of closures being built up by the successive addition of outer shells,

$$cl_D(R) \setminus R, cl_D(cl_D(R)) \setminus cl_D(R), \dots, cl_D^n(R) \setminus cl_D^{n-1}(R), \dots$$

The notion of *distance* is usually defined in terms of shortest paths; but as we shall see, in adjacency spaces it can also be defined in terms of closures.

A *path* of length *n* from cell *x* to cell *y* is a sequence  $x_0, x_1, ..., x_n$  such that  $x_0 = x$ ,  $x_n = y$ , and for i = 1, ..., n,  $x_{i-1} \sim x_i$ . We can prove that there is a path of length *n* from *x* to *y* if and only if  $y \in cl_D^n(\{x\})$ . We use induction on *n*:

*Base case* (n = 0). A path of length 0 from *x* to *y* consists of a single point  $x_0$  which must be equal to both *x* and *y*. Clearly this exists if and only if x = y. Since  $cl_D^0(\{x\}) = \{x\}$  this means that  $y \in cl_D^0(\{x\})$  as required.

Induction step (from n - 1 to n). Assume the result holds for n - 1. If  $x_0, x_1, \ldots, x_n$  is a path of length n from x to y, then  $x_0, x_2, \ldots, x_{n-1}$  is a path of length n - 1 from x to  $x_{n-1}$ . By hypothesis such a path exists if and only if  $x_{n-1} \in cl_D^{n-1}(\{x\})$ . Since  $x_{n-1} \sim x_n = y$ , this means that  $y \in cl_D(cl_D^{n-1}(\{x\})) = cl_D^n(\{x\})$ , as required.

This gives us a natural measure of the distance between two cells:

$$d(x,y) = \min_{n \in \mathbb{N}} (y \in cl_D^n(\{x\})) .$$

From the above, this is the length of a shortest path from *x* to *y*, and it is easy to see that it is a true metric, i.e.,

• d(x, y) = 0 if and only if x = y,

<sup>&</sup>lt;sup>10</sup> An alternative definition of thickness, which would allow the thickness of the universe to be finite, would be to restrict the definition given in the text to non-*U* regions, and define  $Thickness(U) =_{def} \max_{X \subset U} Thickness(X) + 1$ , it being understood that this expression evaluates to  $\infty$  if there is no upper bound to the thickness of non-universal regions.

- d(x,y) = d(y,x)
- $d(x,z) \le d(x,y) + d(y,z)$  (triangle inequality).

Note that there will not, in general, be a unique shortest path between two cells. A familiar example is the adjacency space  $(\mathbb{Z}^2, \sim_4)$ . The points (m, n) and (m+1, n+1) are linked by two minimal paths, of length 2, one going via (m, n+1) and the other via (m+1, n); in general, between the points (m, n) and (m+h, n+k) there are  ${}^{h+k}C_h$  minimal paths, of length h+k. This is the "Manhattan" or "city-block" distance.

An obvious generalisation of distance to regions gives us the *proximal distance* between two regions, defined as the smallest number of closure operations that can be applied to X to produce a region that overlaps Y; as before, this is equivalent to the more familiar definition as the shortest distance between a cell in one region and a cell in the other:

$$pd(X,Y) = \min_{n \in \mathbb{N}} (cl_D^n(X) \cap Y \neq \emptyset) = \min_{\substack{x \in X \\ y \in Y}} d(x,y).$$

Unfortunately proximal distance is not a true metric since (1) the proximal distance between distinct but overlapping regions is zero, and (2) the triangle inequality does not hold, i.e., we can have regions X, Y, Z such that pd(X, Z) > pd(X, Y) + pd(Y, Z).

A more satisfactory measure of distance for regions is the *Hausdorff distance*, defined as the greatest distance between any point in one of the regions and the nearest point in the other:

$$hd(X,Y) = \max\left(\max_{x \in X} \min_{y \in Y} d(x,y), \max_{y \in Y} \min_{x \in X} d(x,y)\right)$$

If  $\max_{x \in X} \min_{y \in Y} d(x, y) = n$ , then for any  $x \in X, y \in Y$  we have  $y \in cl_D^n(\{x\}) \subseteq cl_D^n(X)$ , so  $Y \subseteq cl_D^n(X)$ , and likewise with X and Y reversed. Thus an equivalent formulation in terms of closures is

$$hd(X,Y) = \min \left\{ n \in \mathbb{N} \mid X \subseteq cl_D^n(Y) \land Y \subseteq cl_D^n(X) \right\}.$$

The Hausdorff distance between two regions in adjacency space is thus the smallest n such that each region is within the nth closure of the other. Unlike proximal distance, Hausdorff distance is a true metric.

It should be noticed that while the Hausdorff distance between a region and its closure is always 1, i.e.,  $hd(X, cl_D(X)) = 1$ , this is not necessarily the case for a region and its interior. In fact

$$hd(X, int_D(X)) = 1$$
 if and only if  $cl_D(int_D(X)) = X$ .

Such an X is called a *regular closure set* in [23].

### 7 Relation to Mathematical Morphology

Mathematical Morphology (MM) comprises a set of mathematical tools for manipulating images. Readers familiar with MM will recognise a clear similarity between our discrete interior and closure operations and the erosion and dilation operators of that theory. Here we make this relationship explicit. While the theory of MM may be developed both for continuous and discrete images, for our purposes we will consider only the discrete case: in this case we are working with  $\mathbb{Z}^2$ . An *image* is any subset of this set.

In Mathematical Morphology, there is no pre-defined adjacency relation. Instead, erosion and dilation may be performed with respect to an arbitrary *structuring element* which in effect determines which points are to count as adjacent. A structuring element is itself an image, typically small. Given an image *X* and a structuring element *B*, we define

• The *dilation* of *X* by *B* is the image

$$X \oplus B = \{x + b \mid x \in X, b \in B\}.$$

• The *erosion* of *X* by *B* is the image

$$X \ominus B = \{ y \in \mathbb{Z}^2 \mid \forall b \in B(y + b \in X) \}.$$

Addition here is coordinate-wise, i.e., treating points as vectors. The dilation of a region by *B* expands *B* by replacing each of its points by a copy of *B*; the location of the copy depends on where *B* itself is with respect to the origin. It is usual to assume that the origin is one of the points of *B*, otherwise dilation will result in a displacement of the image as well as an expansion. If the structuring element is taken to be a  $3\times3$  square centered on the origin, then the dilation of any image will be exactly its closure with respect to the adjacency relation  $\sim_8$ . If instead we take a cross-shaped structuring element consisting of the origin and the four points orthogonally adjacent to it (like the leftmost image in Fig. 2), dilation then gives closure with respect to the relation  $\sim_4$ .

Erosion removes the outer part of the image, retaining a point only if a copy of the structuring element anchored on that point would lie entirely within the image. So long as the structuring element has central symmetry (i.e., B = -B), erosion and dilation are related exactly as interior and closure. More generally we have

$$X \oplus B = (X^c \ominus (-B))^c ,$$

where  $-B = \{-x \mid x \in B\}$ .

MM makes much use of operations called *opening* and *closing*, defined as follows:

Opening:  $X \circ B = (X \ominus B) \oplus B$ , Closing:  $X \bullet B = (X \oplus B) \ominus B$ .

These correspond to the DM operations  $cl_D(int_D(X))$  and  $int_D(cl_D(X))$  illustrated in Fig. 6.

Mathematical Morphology may appear to be in some ways more general, and in other ways less so, than Discrete Mereotopology, but both these appearances are misleading.

The sense in which MM may appear to be more general than DM is that it allows arbitrary structuring elements; but this generality could be recovered in DM by defining a different adjacency relation on  $\mathbb{Z}^2$  for each possible structuring element. For example, if the structuring element consists of just the points  $\{(0,0), (1,0), (0,1)\}$ , then the corresponding adjacency relation would relate any point (x,y) to (x,y), (x+1,y), (x,y+1) and nothing else.

On the other hand, as usually presented in terms of structuring elements, MM would appear to presuppose spaces which are homogeneous in the sense that they allow a copy of the same structuring element to be located at each point in the space. As we have seen, however, DM is equally happy in non-homogeneous spaces where such arbitrary translation of structuring elements does not make sense; the discrete closure and interior operations of DM work just as well in this setting as with homogeneous spaces, but the most familiar forms of MM gain no purchase in this context.

This is not, however, the full story, since a number of researchers have investigated forms of MM which allow variable structuring elements—see for example [19, 27]. And it is certainly true more generally that the study of MM is a much more mature research area than that of DM, and as a result has been developed to a considerably greater degree of mathematical sophistication and generality — see for example [1].

#### 8 Relation to Digital Topology

Adjacency spaces are examples of a general class of mathematical structures called *closure spaces* [2]. A closure space is a pair (U, cl), where U is any set, and cl is a function mapping each set  $X \subseteq U$  to a set  $cl(X) \subseteq U$  such that

1.  $cl(\emptyset) = \emptyset$ , 2.  $X \subseteq cl(X)$ , 3.  $cl(X \cup Y) = cl(X) \cup cl(Y)$ .

This notion generalises topological closure, which in addition to satisfying conditions 1-3 also satisfies the idempotency rule

4. cl(cl(X)) = cl(X),

which as we have noted is not, in general, satisfied by sets in an adjacency space.

Although the discrete closure operation in an adjacency space is not a topological closure (since it is not idempotent), one can define topological spaces associated with any adjacency space. A trivial way of doing this is to specify the *discrete*  topology on U, that is, the topology under which cl(X) = X for every  $X \subseteq U$ . Much more interesting is to extend U by including in the universe not just the original elements of U (conceived of as atomic regions) but also boundary elements where these atomic regions adjoin one another, and boundaries of those, etc, depending on the dimensionality one wishes to confer on the space. This is illustrated in Fig. 7, where, on the left is shown part of an adjacency space in the form of a regular hexagonal lattice, and on the right is shown a space which includes, in addition to the hexagons of the original lattice, a set of line segments representing the boundaries of the hexagons and a set of points representing the ends of the line segments. This space can be made into a topology by specifying that the closure of any set consists of the elements of that set together with all their bounding elements. Thus the closure of one hexagonal tile consists of the hexagon together with its six bounding lines and its six bounding points, and the closure of a line element is the line together with its two bounding points. It is easy to see that this closure operation satisfies the conditions (1)-(4) above, and therefore defines a topological space. Such topological spaces, if finite, are called *cellular complexes*, and these are investigated in the context of applications to image analysis in [11].

The topological spaces obtained in this way from rectangular grids of the form  $\mathbb{Z}^n$  are called *Khalimsky spaces* [7, 10]. See also [9] for a discussion of the relationship between these topological approaches and graph-based approaches such as DM.



Fig. 7 An adjacency space in the form of a hexagonal grid (*left*), and the topological space obtained by the addition of bounding lines and points (*right*)

## 9 An Application

Discrete Mereotopology has been applied to the analysis of histological images [17, 18]. Real-world images invariably have imperfections which means that when standard segmentation algorithms are applied to them in order to extract information about the entities pictured, the resulting structures are not always in conformity with theoretical models — e.g., one might find cell nuclei overlapping the boundaries of their cytoplasm, or distinct tissue types wrongly labelled. DM can be used to identify maximally parsimonious ways of repairing such ill-formed images, using "conceptual neighbourhood diagrams" [3] to identify sequences of operations

that can transform an existing inappropriate structure to one that is in conformity with expectation. This method works at the level of individual image pixels, and can therefore harness the power of mathematical morphology alongside DM; but by considering a coarser segmentation of the image one can exploit the ability of DM to allow reasoning about arbitrary adjacency spaces.

To illustrate, Fig. 8(a) shows a Haemotoxylin and Eosin stained section of an odontogenic keratocyst lining. Image-processing techniques are used to extract theoretical cell boundaries from this image, defining "virtual cells" or "v-cells" in the epithelial compartment separating the background free space at the top of the image from the connective tissue at the bottom. The segmentation into v-cells is shown in Fig. 8(b). Individual v-cells, as well as the whole block of connective tissue and the background space, can be regarded as atomic regions of an adjacency space. The vcells adjacent to the connective tissue form what is called the basal layer. If V is the region in the image consisting of the v-cells, and C is the region corresponding to the connective tissue, then the basal layer B can be identified as  $V \cap cl_D(C)$ . By taking successive closures of the basal layer we can segment the epithelium into layers as shown in Fig. 8(c). This operation allows one to derive a more meaningful measure of tissue thickness, for example, than crude measures involving pixel counts or Euclidean distance. Such measures can provide important diagnostic criteria for histopathology. By taking a single target cell within the segmented image, one can similarly use the closure operation to generate nested rings of v-cells, as shown in Fig. 8(d), which can again provide useful information, at the cellular level, on local tissue architecture.



(a) Stained epithelium section



(b) Segmentation into "virtual cells" (v-cells)



(c) Layering of epithelial v-cells



(d) Nested shells of v-cells around a target cell

**Fig. 8** Application of discrete closure operation to a histological image (images courtesy of Prof. Gabriel Landini and Dr D. Randell)

#### Acknowledgments

I would like to thank Professor Gabriel Landini and Dr David Randell of the School of Dentistry, University of Birmingham, for supplying the images used in Fig. 8 and for providing critical commentary on a draft of this chapter. Dr John Stell from the School of Computing, University of Leeds, also provided useful feedback on an earlier draft of this paper.

#### References

- Isabelle Bloch, Henk Heijmans, and Christian Ronse. Mathematical morphology. In Marco Aiello, Ian Pratt-Hartmann, and Johan van Benthem, editors, *Handbook of Spatial Logics*, pages 857–944. Springer, 2007.
- Eduard Čech. *Topological Spaces*. John Wiley, Chichester, 1966. Revised edition by Zdeněk Frolík and Miroslav Katětov.
- Christian Freksa. Temporal reasoning based on semi-intervals. Artificial Intelligence, 54:199– 227, 1992.
- A. P. Galton. The mereotopology of discrete space. In C. Freksa and D. M. Mark, editors, Spatial Information Theory: Cognitive and Computational Foundations of Geographic Science, pages 251–66, New York, 1999. Springer-Verlag.
- A. P. Galton. Continuous motion in discrete space. In A. G. Cohn, F. Giunchiglia, and B. Selman, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Seventh International Conference (KR2000)*, pages 26–37, San Francisco, CA, 2000. Morgan Kaufmann.
- N. M. Gotts. An axiomatic approach to topology for spatial information systems. Technical Report 96.25, University of Leeds School of Computer Studies, 1996.
- Efim Khalimsky, Ralph Kopperman, and Paul R. Meyer. Computer graphics and connected topologies on finite ordered sets. *Topology and its Applications*, 36:1–17, 1990.
- T. Y. Kong and A. Rosenfeld. Digital topology: Introduction and survey. Computer Vision, Graphics and Image Processing, 48:357–93, 1989.
- T. Y. Kong and A. Rosenfeld. Digital topology: a comparison of the graph-based and topological approaches. In G. M. Reed, A. W. Roscoe, and R. F. Wachter, editors, *Topology and Category Theory in Computer Science*, pages 273–89. Clarendon Press, Oxford, 1991.
- T. Yung Kong, Ralph Kopperman, and Paul R. Meyer. A topological approach to digital topology. *The American Mathematical Monthly*, 98(10):901–17, December 1991.
- 11. V. A. Kovalevsky. Finite topology as applied to image analysis. *Computer Vision, Graphics, and Image Processing*, 46:141–61, 1989.
- Sanjiang Li and Mingsheng Ying. Generalized region connection calculus. Artificial Intelligence, 160:1–34, 2004.
- Ian Pratt and Oliver Lemon. Ontologies for plane, polygonal mereotopology. Notre Dame Journal of Formal Logic, 38(2):225–45, 1997.
- 14. Ian Pratt and Dominik Schoop. A complete axiom system for polygonal mereotopology of the real plane. Technical Report UMCS-97-2-2, University of Manchester, March 1997.
- Ian Pratt-Hartmann and Dominik Schoop. Elementary polyhedral mereotopology. *Journal of Philosophical Logic*, 31:469–498, 2002.
- 16. D. A. Randell, Z. Cui, and A. G. Cohn. A spatial logic based on regions and connection. In B. Nebel, C. Rich, and W. Swartout, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Third International Conference (KR'92)*, pages 165–76, San Mateo, Calif., 1992. Morgan Kaufmann.

#### 24

- D. A. Randell and G. Landini. Discrete mereotopology in automated histological image analysis. In *Proc. Second ImageJ user and developer Conference, Luxembourg, 2008*, pages 151– 156, 2008.
- D. A. Randell, G. Landini, and A. Galton. Discrete mereotopology for spatial reasoning in automated histological image analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013. To appear, March 2013.
- J. B. T. M. Roerdink and H. J. A. M. Heijmans. Mathematical morphology for structures without translation symmetry. *Signal Processing*, 15(3):271–277, 1988.
- Azriel Rosenfeld. Digital topology. American Mathematical Monthly, 86(8):621–30, October 1979.
- A. J. Roy and J. G. Stell. A qualitative account of discrete space. In M. J. Egenhofer and D. M. Mark, editors, *Proceedings of GIScience 2002*, pages 276–290. Springer, LNCS 2478, 2002.
- 22. Peter Simons. Parts: a Study in Ontology. Clarendon Press, Oxford, 1987.
- Michael B. Smyth and Julian Webster. Discrete spatial models. In Marco Aiello, Ian Pratt-Hartmann, and Johan van Benthem, editors, *Handbook of Spatial Logics*, pages 713–798. Springer, 2007.
- J. G. Stell. Boolean connection algebras: A new approach to the region-connection calculus. *Artificial Intelligence*, 122:111–136, 2000.
- J. G. Stell. The representation of discrete multi-resolution spatial knowledge. In A. G. Cohn, F. Giunchiglia, and B. Selman, editors, *Principles of Knowledge Representation and Rea*soning: Proceedings of the Seventh International Conference (KR2000), pages 38–47, San Francisco, CA, 2000. Morgan Kaufmann.
- Achille Varzi. Parts, wholes, and part-whole relations: The prospects of mereotopology. *Data and Knowledge Engineering*, 20:259–286, 1996.
- Jacques G. Verly and Richard L. Delanoy. Adaptive mathematical morphology for range imagery. *IEEE Transactions on Image Processing*, 2(2):272–275, 1993.

#### **Appendix 1: Proof of (T4) and (T5)**

The first theorem to be proved is

$$\forall x \forall y (Compl(x, y) \leftrightarrow \neg O(x, y) \land Sum(x, y, U))$$
(T4)

First, suppose Compl(x, y), so since P(y, y) we have  $\neg O(y, x)$  by (D18) and therefore

$$\neg \mathsf{O}(\mathsf{x},\mathsf{y}). \tag{1}$$

Let v be any region; if  $\neg O(v,x)$ , then P(v,y) (since Comp(x,y)), so O(v,y). Hence  $O(v,x) \lor O(v,y)$ . Since v is arbitrary, we always have O(v,U), and hence we have  $O(v,U) \leftrightarrow O(v,x) \lor O(v,y)$ , i.e.,

$$Sum(x, y, U)$$
(2)

Conversely, suppose we have  $\neg O(x, y) \land Sum(x, y, U)$ .

Let P(u, y), so that whenever P(w, u) we have also P(w, y). Therefore from  $\neg O(x, y)$ , i.e.  $\neg \exists w(P(w, y) \land P(w, x))$ , we infer  $\neg \exists w(P(w, u) \land P(w, x))$ , i.e.,  $\neg O(u, x)$ . Hence we have shown  $\forall u(P(u, y) \rightarrow \neg O(u, x))$ . Now suppose  $\neg O(u,x)$ . We must show P(u,y). Suppose not; then from  $\neg P(u,y)$ , by (A8), there is a region w such that  $P(w,u) \land \neg O(w,y)$ . From  $\neg O(w,y)$ , since Sum(x,y,U) we have O(w,x) (since O(w,U) in any case). From P(w,u) and O(w,x) it is an easy deduction that O(u,x), contradicting our assumption. Hence we have shown  $\forall u(\neg O(u,x) \rightarrow P(u,y))$ .

We now have 
$$\forall u(P(u, y) \leftrightarrow \neg O(u, x))$$
, i.e.,  $Compl(x, y)$ .

The proof of (T5) is now straightforward:

 $\mathsf{Compl}(x,y) \Leftrightarrow \neg \mathsf{O}(x,y) \land \mathsf{Sum}(x,y,\mathsf{U}) \Leftrightarrow \neg \mathsf{O}(y,x) \land \mathsf{Sum}(y,x,\mathsf{U}) \Leftrightarrow \mathsf{Compl}(y,x).$ 

### **Appendix 2: Proof of (T6)**

The theorem to be proved is

$$\forall x \forall y (Compl(x, y) \to \forall z (C(z, y) \leftrightarrow \neg NTP(z, x)))$$
(T6)

Assuming

$$Compl(a,b),$$
 (1)

suppose, first, that C(c,b). From (1), since P(b,b), we have  $\neg O(b,a)$ , hence we have  $C(c,b) \land \neg O(b,a)$  and therefore  $\neg NTP(c,a)$  by (D10). Hence we have  $\forall z(C(z,b) \rightarrow \neg NTP(z,a))$ .

Next, suppose we have  $\neg NTP(c, a)$ . Then from (D10), either we have

$$\neg P(c,a)$$
 (2a)

or there is a region d such that

$$C(c,d) \wedge \neg O(d,a). \tag{2b}$$

In the former case, from (1) and (T5) we have Compl(b, a), so (2a) implies O(c, b), which implies C(c, b). In the latter case (2b), from (1) and  $\neg O(d, a)$  we have P(d, b). Then from C(c, d) and P(d, b) we have C(c, b) by (A5). Thus in either case we have C(c, b) and we have proved  $\forall z (\neg NTP(z, a) \rightarrow C(z, b))$ .

Combining the results and generalising give us (T6).

# **Appendix 3: Proof of (T7): Relationship of discrete interior and closure**

The theorem to be proved is

$$Compl(x,y) \land Int(y,z) \land Compl(z,w) \to Cl(x,w)$$
(T7)

Using the definitions (D19, D20, D18), this means that from

$$\forall x (P(x,b) \leftrightarrow \neg O(x,a)) \tag{1}$$

$$\forall x (P(x,c) \leftrightarrow NTP(x,b))$$
(2)

$$\forall x (P(x,d) \leftrightarrow \neg O(x,c))) \tag{3}$$

we must derive

$$\forall x (P(d, x) \leftrightarrow NTP(a, x)) \tag{4}$$

From (T5) we can rewrite (1) and (3) as

$$\forall x (P(x,a) \leftrightarrow \neg O(x,b)) \tag{5}$$

$$\forall x (P(x,c) \leftrightarrow \neg O(x,d)) \tag{6}$$

so from (2) and (6) we have

$$\forall x(\neg O(x,d) \leftrightarrow \mathsf{NTP}(x,b)) \tag{7}$$

Then from (1) and (7) (since NTP(x, b) implies P(x, b)) we have

$$\forall x (\neg O(x,d) \rightarrow \neg O(x,a)) \tag{8}$$

Suppose  $\neg P(a,d)$ . Then by (A8) there must be a region e such that  $P(e,a) \land \neg O(e,d)$ . By (8) this would imply  $P(e,a) \land \neg O(e,a)$ , a contradiction. Therefore we have P(a,d), and therefore

$$\forall x (P(d, x) \to P(a, x)) \tag{9}$$

Suppose P(d,g). and let f be any region connected to a, i.e., C(a,f). Suppose  $\neg O(f,d)$ . Then by (6) we have P(f,c) and so by (2), NTP(f,b). Therefore any region connected to f must overlap b. Since we have C(a,f) this means that O(a,b). But from (1) we know that  $\neg O(a,b)$  and we have a contradiction. Therefore O(f,d), and therefore, since P(d,g), we have O(f,g).

Thus we have

$$\forall x (\mathsf{P}(\mathsf{d}, x) \to \forall y (\mathsf{C}(\mathsf{a}, y) \to \mathsf{O}(y, x))) \tag{10}$$

Combining (9) and (10) (and using (D10)) we get

$$\forall x (P(d, x) \to NTP(a, x))$$
(11)

which is one half of (4).

For the converse, let  $\mathsf{NTP}(\mathsf{a},\mathsf{e})$ ; we must show that  $\mathsf{P}(\mathsf{d},\mathsf{e})$ . Assume on the contrary that

$$\neg \mathsf{P}(\mathsf{d},\mathsf{e})$$
 (12)

By (A8), there is a region f such that P(f,d) and  $\neg O(f,e)$ .

From P(f,d) we have, by (3),  $\neg O(f,c)$ , and therefore

$$\neg \mathsf{P}(\mathsf{f},\mathsf{c}) \tag{13}$$

From  $\neg O(f, e)$ , i.e.,  $\neg O(f, e)$ , we have, since NTP(a, e),

$$\neg \mathsf{C}(\mathsf{a},\mathsf{f}) \tag{14}$$

Also from  $\neg O(f,e)$ , given NTP(a,e), and therefore P(a,e), we have  $\neg O(f,a)$ , and therefore, by (1), P(f,b).

We will show that in fact NTP(f, b).

To this end we must show that anything connected to f overlaps b. Suppose  $\neg O(z,b)$ , i.e.,  $\neg O(z,b)$ . By (5) this implies P(z,a), and therefore, from (14),  $\neg C(z,f)$  (here we are using a theorem  $C(x,y) \land P(x,z) \rightarrow C(y,z)$  — we've already asserted this, as (5)). Hence if C(z,f) it follows that O(z,b). Thus we have NTP(f,b).

By (2) this gives P(f,c), contradicting (13). Hence assumption (12) is false, and we conclude, as required, that P(d,e).

We have now shown that  $\forall x(NTP(a,x) \rightarrow P(d,x))$ , which, in combination with (11), gives us (4).

28