# Mereotopological Connection

Anthony G. Cohn School of Computing, University of Leeds, Leeds, UK

Achille C. Varzi Department of Philosophy, Columbia University, New York, USA

(Final version published in Journal of Philosophical Logic 32:4 (2003), 357-390)

**Abstract.** The paper outlines a model-theoretic framework for investigating and comparing a variety of mereotopological theories. In the first part we consider different ways of characterizing a mereotopology with respect to (i) the intended interpretation of the connection primitive, and (ii) the composition of the admissible domains of quantification (e.g., whether or not they include boundary elements). The second part extends this study by considering two further dimensions along which different patterns of topological connection can be classified—the strength of the connection and its multiplicity.

#### 1. Introduction

In recent years there has been an outgrowth of theories for representing space and time in a qualitative way based on a primitive notion of topological connection. Applications include natural language semantics, formal ontology, cognitive geography, and spatial and spatiotemporal reasoning in artificial intelligence (see [12, 37, 43, 60, 64] for overviews). Most of this work has been influenced directly by the seminal contributions of Clarke [16, 17], which in turn were based on the theory of extensive connection outlined by Whitehead in *Process and Reality* [67]. However, some theories have been developed on independent grounds and with different motivations (for instance, as an attempt to overcome the intrinsic limits of mereology in dealing with the notion of an integral whole, following in the footsteps of Husserl's *Logical Investigations* [45]), and it is interesting to see how topology has itself become a point of connection among a number of previously unrelated research areas.

Unfortunately, this variety of outlooks corresponds to a variety of theories that are not always in agreement on the basic terms. In some cases, the disagreement is a sign of genuine philosophical dissension; in other cases it simply reflects a difference in the intended application of the theories. In other cases still, the disagreement is due to a different understanding of the connection primitive itself. This is not surprising, since the ordinary set-theoretic account of the topological notion of connection rests on the distinction between open and closed entities (sets) and since Bolzano [6] this distinction has been regarded as problematic when the entities in question are interpreted on a spatiotemporal domain (see [66, 68] for discussion). Indeed, the difficulty in applying standard point-set topology to ordinary space is one of the guiding reasons behind the development of many connection-based theories. As mereology (the theory of parthood or overlap) was initially developed as an alternative to set theory in the constructional analysis of the world of spatiotemporal entities (at least in the pioneering works of Leśniewski [47] and Leonard and Goodman [48], the latter under the emblematic rubric of Calculus of Individuals), likewise the theory of connection is typically viewed as an alternative to set theory in the analysis of those spatial and spatiotemporal phenomena that exhibit topological structure. The resultant theories are sometimes called, quite aptly, mereotopologies. And the lack of a unified framework bear witness to the difficulty of the task, in spite of its practical significance.

Our aim in this paper is to go some way in the direction of such a unified framework. Independently of any foundational or practical concerns, in what follows we delineate a model-theoretic framework for investigating the logical space of mereotopological theories and comparing the main options in light of their intended models. In the first part (Sections 2–6) we introduce the basic notation and terminology and we consider different ways of characterizing a mereotopology with respect to the intended interpretation of the connection primitive(s). We also consider, albeit briefly, the way theories may differ with regard to the composition of their intended domains—e.g., whether or not they allow for any boundary elements. In the second part (Sections 7–9) we extend this study by considering two further dimensions along which different patterns of topological connection can be classified—the strength of the connection and its multiplicity.

#### 2. The Frame of Reference

Our approach is neutral with regard to whether mereotopology is meant to apply to space, time, or space-time. Most of what follows holds true of domains of arbitrary dimensions, of which space and time can be seen as special cases. However, in our illustrations we shall focus primarily on the spatial domain. This will hopefully be an aid to intuition, and will allow us to exploit rather closely a terminology that has recently become rather widespread precisely under the impact of recent work in the area of spatial reasoning.

Let us also emphasize that our focus will be on the logical spectrum of theories concerned with the topological structure of space, as opposed to things *located in* space. This makes our study independent of questions of location, which call for a different sort of theory (see [11, 12] for some work in this direction). Thus, the domain of quantification of these theories consists exclusively of spatial items such as points, lines, regions. Moreover, we are only interested in these theories insofar as their account of the connection relation is concerned, ignoring other important topological notions such as, for instance, compactness.

Within these limits, our purpose is taxonomic in the following sense. Given theory X, the intended interpretation of the basic relation of topological connection in X can be described in terms of ordinary point-set theoretic notions. That is, the *semantics* of X can be specified in settheoretic terms (though, of course, the *purpose* of X may be to provide an independent account of topology, dispensing with set-theoretic notions altogether)<sup>1</sup>. Our aim is to compare various theories on the basis of such a common set-theoretic semantics. Given a topological space T, we want to investigate how the basic predicates of different mereotopological theories get interpreted when the variables are taken to range over elements of T.

<sup>&</sup>lt;sup>1</sup> Thus, for instance, both Whitehead [67] and Clarke [16] present their systems as formalizing the relation 'x is connected with y' as a rendering of 'x and y share a common point', although points are not meant to be included in the domain of quantification of their theories. Similarly, Randell *et al.* [53] suggest the renderings 'there is no distance between x and y' as well as 'the closures of x and y share a common point'; the former is the favored interpretation, but the latter makes it possible to classify and compare their system with others.

Clearly, a lot depends on which elements of T are included in the domain of quantification. This is one major source of disagreement among different theories in the literature. For instance, in certain theories the intended interpretation of the connection predicate is explained (or can be explained) in terms of the full make up of T even if the theories themselves require that only some elements of T be included in their domain of quantification. In particular, we shall be interested in comparing theories with regard to whether or not they include unextended boundary elements (elements with empty interiors, such as points, lines, and surfaces). Theories that do include such elements in their domain of quantification will be called boundary-based theories; the others will be called boundary-free. It is clear that this opposition is reflected in the account of defined topological relations and predicates-for instance, it crucially affects the way in which the opposition between open and closed regions can be expressed in the language of the theory. Our purpose is to investigate this and related issues in abstract terms, and to illustrate in some detail the cases corresponding to actual theories that play a prominent role in the literature.

Another important factor is the kind of topological space one considers. In particular, one may draw a line between theories that take space to be dense (a normal space) and those that do not. Most accounts in the literature are of the first kind, but there are exceptions. In the following we shall remain neutral on this issue and work with arbitrary topological spaces, though eventually we shall also consider some questions that turn out to relate to the density assumption.

#### 3. Definition Schemas for Mereotopology

Topological spaces can be conveniently characterized in terms of closure operators. A *closure* operator on a set A is a function c associating with each subset x of A a subset c(x) satisfying the following four statements:

If *c* is a closure operator on *A*, the set of fixed points  $T(c) = \{x \mid A: c(x) = x\}$  is the *topology* on *A* associated with *c*. The pair (*A*, *T*(*c*)) is then called a

topological space, and the elements of T(c) are called the *closed* sets of the topological space. It is a fundamental fact of all topological spaces (due essentially to Kuratowski [46]) that the set T(c) of closed sets contains both A and and is closed under the formation of finite unions and arbitrary intersections. Moreover, one verifies that the closure of a set x is always the smallest closed set including x. Likewise, let us define the interior of x, written i(x), to be the greatest open set included in x, where a subset of A is called *open* if and only if (iff) it is the relative complement A-x of a closed subset x. Then one verifies that the set O(c) of all open sets contains both A and and is closed under the formation of finite intersections and arbitrary unions. For convenience, let us also introduce the notion of a boundary. Given any subset x of A, the boundary of x can be conveniently defined as the set b(x) = c(x) - i(x), and a set z is called a boundary (or boundary element) in A iff it is the boundary of a subset of A.

Now let T = (A, T(c)) be any topological space. We shall initially focus on the following three ways of characterizing a relation of connection between subsets of A (see Figure 1):

$$\begin{array}{rcl} C_{1}(x, y) & x & y \\ C_{2}(x, y) & x & c(y) & \text{or } c(x) & y \\ C_{3}(x, y) & c(x) & c(y) \end{array}$$

With ' $C_n$ ' understood as the connection predicate in the object language of a theory, the double arrow is to be construed as a semantic relation specifying the relevant intended interpretation. (We postpone the details to Section 4.) Thus, intuitively, on the first account the connection relation holds between any two sets that share an element; on the second it holds iff one of the sets shares an element with the closure of the other; and on the third it holds iff it is the closures of both sets that share an element.

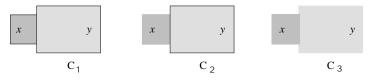


Figure 1: The three C relations (limit cases); a solid line indicates closure.

As we shall see, these notions correspond—or can be made to correspond, under suitable conditions—to the main variants found in the literature.<sup>2</sup> However, to get a proper picture of the alternatives offered by these options, two more parameters must be considered, corresponding to the ways in which the relation of parthood and the operation of fusion can be characterized in terms of connection:

$$P_n(x, y) =_{df} z(C_n(z, x) \quad C_n(z, y))$$
(1 n 3)  
$${}_n x =_{df} z y(C_n(y, z) \quad x( \quad C_n(y, x)))$$
(1 n 3)

Intuitively: x is part<sub>n</sub> of y iff whatever is connected<sub>n</sub> to x is also connected<sub>n</sub> to y, and the fusion<sub>n</sub> of all -ers is that thing (if it exists at all) that connects<sub>n</sub> precisely to those things that .<sup>3</sup> Most theories define these notions in terms of the same connection relation that is assumed as a topological primitive, in which case the above reduce to ordinary definitions in the object language of the theory. However, this need not be the case, and in fact we shall see that an important family of theories stem precisely from the intuition that parthood and connection cannot be defined in terms of each other. This effectively amounts to using two distinct primitives—two notions of connection (one of which is used in defining parthood), or a notion of connection and an independent notion of parthood. Accordingly, and more generally, we shall consider the entire space of mereotopological theories that result from the options determined by the above definitions when 1 n 3. That is to say, we shall work with an object language in which all three connection predicates are available as primitives, and we

Such relations do not correspond to any notion of connection found in the literature, so we shall not consider them here. We shall come back to them indirectly in Section 8.

<sup>3</sup> Here and below we assume the definite descriptor ' ' to be contextually defined in Russellian fashion:

 $(x) =_{df} x(y(y'_x y = x)).$ 

<sup>&</sup>lt;sup>2</sup> Other relations of connection could be defined using the interior operator instead of the closure operator—for instance:

 $C_4(x, y) \qquad x \quad i(y) \quad \text{or } i(x) \quad y \\ C_5(x, y) \qquad i(x) \quad i(y)$ 

where ' ' and ' ' stand for any well-formed expressions of the object language in which '*x*' occurs free and '  $\frac{y}{x}$ ' denotes the expression obtained from ' ' by properly substituting every free occurrence of '*x*' by an occurrence of '*y*'. It would be interesting, but too complex for our present purposes, to consider the possibility of treating ' ' as a genuine term-forming operator against the background of a free quantification theory. See [55] for a first step in this direction.

shall model theories in which some such predicates are defined in term of others by adding suitable axioms in place of the corresponding definitions. (The case of  $_n$  is particularly complicated, also in view of the fact that

 $_{n}x$  will not be uniquely defined unless  $C_{n}$  is assumed to be extensional, i.e., unless it is assumed that distinct sets must have different connections<sub>n</sub>; we shall come back to this in the next section.)

To do so in a systematic manner, let us call a triple = i,j,k (where 1 i,j,k 3) a type. The first coordinate of a type is meant to indicate a corresponding relation of connection; the second coordinate indicates a corresponding choice for the definition of parthood; and the third coordinate indicates which connection relation is used in the definition of the fusion operator. For instance, 2,1,3 is the type associated with  $C_2$  as a primitive for topological connection, P1 as a parthood predicate (defined in terms of a different connection primitive  $C_1$ ), and  $_3$  as a fusion operator (defined in terms of  $C_3$ ). If the three coordinates of a type are all equal, then the type is *uniform* and corresponds to the case in which the primitive for topological connection is the only primitive used in defining all other mereotopological notions. (The other primitives do not enter in the theory, or can be shown to be equivalent to open formulas involving only the chosen primitive.) However, a type need not be uniform, and this is meant to reflect the possibility of relative independence among the three notions of connection, parthood, and fusion.

Using the notion of a type, the following notation provides a convenient generalization of the notation introduced above:

 $C_{i,j,k} (x, y) =_{df} C_i(x, y)$   $P_{i,j,k} (x, y) =_{df} P_j(x, y)$   $i,j,k x =_{df} k^x$ 

We can then define a number of customary mereotopological relations by relativizing them to our types . To simplify notation, we shall assume bound variables to range exclusively over non-empty sets.

O(x, y)	$=_{df}$	$z(\mathbf{P}(z,$	$x)  \mathbf{P}(z, y))$		x -overlaps $y$
A ( <i>x</i> , <i>y</i> )	$=_{df}$	C(x, y)	$\neg O(x, y)$		x -abuts y
E ( <i>x</i> , <i>y</i> )	$=_{df}$	P(x, y)	P ( <i>y</i> , <i>x</i> )		x -equals y
$PP\ (x, y)$	$=_{df}$	P ( <i>x</i> , <i>y</i> )	$\neg P(y, x)$		x is a proper -part of $y$
TP $(x, y)$	$=_{df}$	P(x, y)	$z(\mathbf{A}(z,x))$	A $(z, y)$ )	<i>x</i> is a tangential -part of <i>y</i>

BP ( <i>x</i> , <i>y</i> ) PO ( <i>x</i> , <i>y</i> ) TO ( <i>x</i> , <i>y</i> ) IO ( <i>x</i> , <i>y</i> )	$=_{df} P(x, y) \neg TP(x, y) =_{df} z(P(z, x) TP(z, y)) =_{df} O(x, y) \neg P(x, y) \neg P(y, x) =_{df} z(TP(z, x) TP(z, y)) =_{df} z(IP(z, x) IP(z, y)) =_{df} O(x, y) \neg IO(x, y)$	x is an interior -part of y x is a boundary -part of y x properly -overlaps y x tangentially -overlaps y x internally -overlaps y x boundary -overlaps y
x-y k (x) i (x) e (x) c (x)	$=_{df} z x ( P (z, x)) =_{df} z (P (z, x) P (z, y)) =_{df} z (P (z, x) P (z, y)) =_{df} z (P (z, x) P (z, y)) =_{df} z \neg O (z, x) =_{df} z IP (z, x) =_{df} i (k (x)) =_{df} k (e (x)) =_{df} c (x) - i (x) =_{df} z O (z, z)$	<ul> <li>-product of ers</li> <li>-sum of x and y</li> <li>-product of x and y</li> <li>-difference of x and y</li> <li>-complement of x</li> <li>-interior of x</li> <li>-exterior of x</li> <li>-closure of x</li> <li>-boundary of x</li> <li>-universe</li> </ul>
Rg (x) Op (x) Cl (x) Re (x) Cn (x)	$=_{df} yBP(x, y) =_{df} yIP(y, x) =_{df} E(x, i(x)) =_{df} E(x, c(x)) =_{df} E(i(x), i(c(x))) =_{df} y z(E(x, y+z) C(y, z)) =_{df} P(x, y) Cn(x)$	<pre>x is a -boundary x is a -region x is -open x is -closed x is -regular x is -connected x is a -connected part of y</pre>

Depending on the structure of , the notions thus defined may receive different interpretations, hence the gloss on the right should not be taken too strictly. One intended interpretation of the binary relations relative to the Euclidean plane  $R^2$ —an interpretation that justifies the gloss—is illustrated in Figures 2 and 3. We shall call this the *standard interpretation*. However, the exact semantic import of these definitions may change radically from one framework to another, depending on the type and on the constraints in the model theory. Our aim is precisely to investigate this variety of interpretations. We shall do so by considering boundary-based accounts first, and then boundary-free accounts. This effectively corresponds to interpreting the quantifiers in the above definitions as ranging over two different sorts of subsets of the underlying topological space.

Boundary-based theories include in the domain of quantification all boundaries (sets of points that have no interior). Boundary-free theories, by contrast, do not include such elements in their domain of quantification.

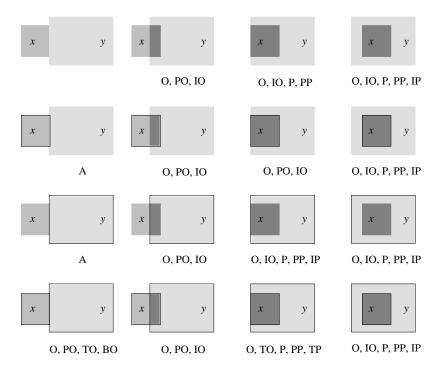


Figure 2: Standard interpretation of the mereotopological relations on the Euclidean plane  $R^2$ . (The labels under each diagram indicate relations that hold between *x* and *y*, in this order.)

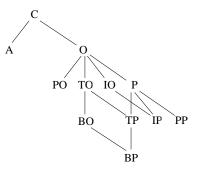


Figure 3: Partial ordering by implication of the mereotopological relations of Figure 2.

Moreover, in both cases further constraints may derive from specific conditions on the relevant topology. For instance, boundary-based theories typically assume that all boundary elements qualify as closed; boundaryfree theories typically assume all open sets to be *regular*, i.e., to coincide with the interior of their own closure (this rules out, for instance, sets obtained by removing an interior point from an open set).

Incidentally, one could also draw a line here between theories that include the empty set in their domain of quantification and theories that do not. Classical topology does, of course. However, mereotopologies typically do not follow this course: virtually every account in the literature, whether boundary-based or boundary-free, agrees on restricting the domain of quantification by leaving out the empty set. Intuitively this corresponds to the idea that nothing can occupy an empty region of space (along with the thought that an empty region of space can hardly be associated with clear identity criteria: where would it be located?). In the following, we shall focus on this course, which is why the above list of definitions is given on the understanding that variables only range over non-empty sets. We shall therefore confine ourselves to a comparison between boundarybased and boundary-free mereotopologies without the null element. An extension of our results to mereotopologies with the null element is rather trivial and we leave it to the reader.

One final remark is in order. It concerns the possibility of extending our set of defined mereotopological predicates and operators by relying on higher-order notions of connection. For instance, using parthood and fusion one can define overlap (the mereological counterpart of set intersection) and closure; but then one could use these notions to introduce a corresponding variety of connection predicates, which in turn can be used to define corresponding notions of parthood and fusion, and so on.

It is not difficult to include this possibility into our general picture. Let us simply amend our notion of type by including a fourth coordinate, indicating the order at which the predicate is defined. This can be done by recursive definition:

if 1 i,j,k 3, then i,j,k,0 is a type; if is a type and 1 i,j,k 3, then i,j,k, is a type; nothing else is a type.

The basic types give us the same as above:

$$C_{i,j,k,0} (x, y) =_{df} C_i(x, y) P_{i,j,k,0} (x, y) =_{df} P_j(x, y) _{i,j,k,0} =_{df} _kx$$

But the inductive types allow us to introduce higher-order connection relations:

$$C_{1,j,k,} (x, y) =_{df} O(x, y)$$
  

$$C_{2,j,k,} (x, y) =_{df} O(x, c(y)) O(c(x), y)$$
  

$$C_{3,j,k,} (x, y) =_{df} O(c(x), c(y))$$

Using these notions, the long list of definitions given above can be iterated, yielding further mereotopological predicates and operators. Some of these will of course collapse, but not necessarily all. The question of whether and when the distinctions between basic and higher-order notions can be dismissed is itself an interesting subject to be explored, but we shall not include it in our agenda here. (Actually one can have many more relations by allowing the entries in our list of definitions to involve relations of different types in the definienda—for instance, one could define hybrid forms of overlap:

$$O_{1,2}(x, y) =_{df} z(P_{1}(z, x) P_{2}(z, y))$$

However, this way of generating new relations is probably not interesting and will also be ignored in the following.)

# 4. General Facts

Before proceeding to a comparative analysis of mereotopological theories, we note here some general facts.

First of all, let us be more explicit about the model-theoretic apparatus. We assume a first-order language with identity  $L=\{C_1, C_2, C_3\}$  whose non-logical vocabulary consists of the three connection predicates. The semantics follows a standard first-order account. Only notice that we are interested in models that are based on some topological space T=(A, T(c)), i.e., models M=(U, f) whose domain U is a non-empty subset of (A)and whose interpretation function f treats each connection predicate 'C<sub>n</sub>' as indicated at the beginning of Section 3 (relative to T). Such models are called *canonical*. Given a model M = (U, f), the notion of an L-formula being satisfied by a model M relative to a finite sequence of elements  $a_1, ..., a_n$  U (notation  $M \models [a_1, ..., a]$ ) is defined as usual, and so are all the other semantic notions. In particular, M satisfies a formula *simpliciter* iff is satisfied by M relative to every finite sequence of elements, in which case we also say that is *true* in M. We shall also use the notation M to indicate the relation defined by in M:

$$^{M} =_{\mathrm{df}} \{ a_{1}, ..., a_{n} \quad U^{n} \colon M \models [a_{1}, ..., a_{n}] \},$$

where *n* is the number of distinct variables occurring free in . And, for convenience, the same notation will be extended to defined predicates, so that, for instance,  $P_n^M$  is a name for the relation  $P_n(x_1, x_2)^M$ , where  $x_1$  and  $x_2$  are the first two variables in the alphabetic order. ( $C_n^M$  is another name for  $f(C_n)$ .)

Now, it is easy to see that the following formulas are true in every canonical model for all types :

(C1) C (x, x)(C2) C (x, y) C (y, x).

In other words, each connection relation  $C^M$  is *reflexive* and *symmetric*. (When is a basic type, reflexivity follows from the assumption that the universe contain only non-empty sets, and symmetry follows from the commutativity of ; when is inductive, these properties follow by definition.) Also, the following are always logically true in view of the definition of P :

(P1) P 
$$(x, x)$$
  
(P2) P  $(x, y)$  P  $(y, z)$  P  $(x, z)$ .

Thus, each parthood relation  $P^M$  is *reflexive* and *transitive* in every canonical model M.

Another important property that is often associated with parthood is *antisymmetry*. There are two formulations of this property, depending on whether we use -equality (E) or plain equality (=). The first formulation

(P3) P (x, y) P (y, x) E (x, y).

is obviously true by definition. However, the second formulation,

$$(P3_{=})$$
 P  $(x, y)$  P  $(y, x)$   $x = y$ ,

is stronger and may fail in some models. For instance, if the universe includes only two sets with a non-empty intersection (but does not include the intersection itself), then (P3 \_) is false for each basic type  $\therefore$  Indeed, the requirement that P<sup>M</sup> be antisymmetric in the sense of (P3 \_) is logically equivalent to the requirement that P<sup>M</sup> be *extensional* in the following sense:

$$(P4_{=})$$
  $z(P(z, x) P(z, y)) x = y,$ 

which in turn is equivalent to the requirement that  $C^M$  be likewise extensional:

$$(C3_{=})$$
  $z(C(z,x) C(z,y)) x=y.$ 

(The first equivalence follows by (P1 ) and (P2 ), while the second follows directly from the definition of P .) These requirements are stronger than the corresponding versions for E :

(C3) 
$$z(C(z, x) = C(z, y)) = E(x, y).$$
  
(P4)  $z(P(z, x) = P(z, y)) = E(x, y).$ 

These latter are logically true. But whether a model satisfies (P4  $_{=}$ ) and (C3  $_{=}$ ) depends crucially on the relevant closure operator c and on which sets are included in the universe U. Figures 4 and 5 show that there are models satisfying or falsifying any combination of the first three basic instances of (C3  $_{=}$ ) (i.e., the three instances obtained by taking to be a basic type), thus showing the relative independence of the three sorts of extensionality. This represents a significant parameter in comparing competing theories. In the following, we shall call a model M -extensional iff it satisfies the extensionality axioms for C<sup>M</sup> and P<sup>M</sup> and, consequently, the antisymmetry axiom for P<sup>M</sup>.

Let us now see how the various notions of connection, parthood, and fusion are related. For the basic case where = i,j,k,0, the relationships between the three C 's is easily stated: they are ordered by increasing strength. In other words, every canonical model satisfies the following conditionals:

This follows immediately from (A1), as illustrated in Figure 1. More generally, for any pair of types  $_1 = i_1, j, k$ , and  $_2 = i_2, j, k$ , the following holds whenever  $i_1 \quad i_2$ :

 $(C4_{i_1i_2}) C_1(x, y) C_2(x, y).$ 

(This can be shown by induction on .)

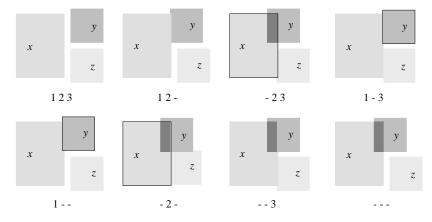


Figure 4. Relative independence of extensionality. Each diagram illustrates a model with universe  $U = \{x, y, z\}$ ; the inclusion/exclusion of 'i' in the label under a diagram indicates whether the corresponding model is or is not i,j,k,0 -extensional.

	i=1		i=2			i=3			
	x	у	z	x	у	z	x	у	z
123	x	у	z	x	у	z	x	у	z
12-	x	у	z	x	у	z	<i>x</i> , <i>y</i>	х, у	z
-23	<i>x</i> , <i>y</i>	х, у	z	<i>x</i> , <i>y</i> , <i>z</i>	х, у	<i>X</i> , <i>Z</i>	<i>x</i> , <i>y</i> , <i>z</i>	х, у	<i>x</i> , <i>z</i>
1-3	x	у	z	х, у	х, у	z	<i>x</i> , <i>y</i> , <i>z</i>	х, у	<i>x</i> , <i>z</i>
1	x	у	z	х, у	<i>x</i> , <i>y</i>	z	<i>x</i> , <i>y</i>	х, у	z
-2-	<i>x</i> , <i>y</i>	х, у	z	<i>x</i> , <i>y</i> , <i>z</i>	х, у	<i>x</i> , <i>z</i>	<i>x</i> , <i>y</i> , <i>z</i>	<i>x</i> , <i>y</i> , <i>z</i>	<i>x</i> , <i>y</i> , <i>z</i>
3	<i>x</i> , <i>y</i>	х, у	z	х, у	<i>x</i> , <i>y</i>	z	<i>x</i> , <i>y</i> , <i>z</i>	х, у	<i>x</i> , <i>z</i>
	<i>x</i> , <i>y</i>	х, у	z	<i>x</i> , <i>y</i>	<i>x</i> , <i>y</i>	z	<i>x</i> , <i>y</i>	х, у	z

Figure 5. Tabulation of the connection patterns depicted in Figure 4. Each row corresponds to a model. For each value of i, the column under x lists the items in the universe to which x is connected, and likewise for y and z. A model is i,j,k,0 extensional iff the corresponding three cells are pairwise distinct (shaded areas).

The three parthood predicates are not, in general, related in a similar fashion. In fact, no instance of the following *inclusion* schema is generally true when 1 = 2:

$$(P5_{i_1i_2}) P_1(x, y) P_2(x, y).$$

A glimpse at Figures 4 and 5 is sufficient to see that there are models satisfying or falsifying any combination of the first three parthood relations, thereby showing their relative independence. (Each pattern in the diagram illustrates a model that satisfies P (x,y) iff it is not –extensional.) However, all of these models are, in a sense, non-intended, and one might want to rule them out precisely by assuming one or more instances of (P5<sub>i,i2</sub>), along with some form of extensionality (which some authors regard as an essential feature of all connection relations [26]). For instance, every model in which P<sup>M</sup><sub>1</sub> is the relation of set inclusion satisfies extensionality as well as (P5<sub>32</sub>). This follows immediately from the fact that x = y always implies c(x) = c(y).

With the fusion operator the situation is more complex. Let us extend our notation to this case by setting:

$$(x)^{M} =_{df} (z = x)^{M},$$

where z is, say, the first variable foreign to  $\therefore$  If x is uniquely defined, this yields an *n*-ary operation, where *n* is the number of free variables in  $\therefore$ (In particular, 0-ary operations are just singletons in U.) But x need not be uniquely defined. For one thing, there is no guarantee that U be closed under fusions unless one requires so. In the most general case, such a requirement amounts to assuming the principle of unrestricted fusion:

(C4) 
$$x = z(z = x)$$
.

Let us say that a model satisfying (C4) for a given formula is *-fused* for .<sup>4</sup> Then the point is that unless M is *-fused* for , there is no guarantee that the *existence* condition for x be satisfied. Hence one may get  $(x)^{M} =$ , which is not an element of U. On the other hand, if M is not

<sup>&</sup>lt;sup>4</sup> One could distinguish many variants of this principle, depending on what additional requirements one may want to impose on besides satisfiability. For instance, one may want to rule out disconnected fusions (see [14] for a classical discussion of this view.) On this and related aspects of the idea behind (C4) we refer to [12].

-extensional, then there is no guarantee that the *uniqueness* condition be satisfied. Hence  $(x)^M$  may fail to be a function. As an example, take = 1,j,k,0 and let be the formula  $P_3(x, z)$ . Then  $(x)^M$  should map every u U to the unique set a U such that, for all b U, c(a) c(b) iff c(b) c(c) for some c u. But there may not be such a unique set a unless M is -extensional. For instance, if both the interior and the closure of a are included in U, then both would satisfy the required conditions.

In view of the above, it is not possible to establish general relationships between the operations defined by the various . In general, the determination of the necessary and sufficient conditions that a model must satisfy in order for it to be -fused for a formula is, as far as we can see, a difficult open problem.

# 5. Boundary-based Theories

We now proceed to examine in some detail the logical space of the theories that result from the options discussed above.

Let = i,j,k, ' be a type. A theory which formalizes topological connection by C , parthood by P , and fusion by we call a *-theory*. There are of course many distinct *-*theories, depending on how the basic predicates are axiomatized. Here we consider some indicative examples, confining ourselves to the case '=0 (zero-order theories). We begin in this section with boundary-based *-*theories, i.e., theories whose models include boundary elements; in the next section we shall move on to boundary-free theories.

Consider first the case where is uniform (i=j=k). In this case a theory could be formulated within a proper fragment of the language L, namely the fragment  $L_i=\{C_i\}$ , and with '=Oone can in principle distinguish three main options (1 i 3). As it turns out, however, we may note immediately that none of these options is viable.

(a) The option i=1 yields implausible topologies in which the boundary of a region is never connected to the region's interior (since the boundary and the interior never share any points).

(b) The option i=2 yields implausible mereologies in which every boundary is part of its own complement (since anything connected to the former is connected to the latter).

(c) The option i=3 yields implausible mereotopologies in which the interior of a region is always connected to its exterior (so that boundaries make no difference) and in which the closure of a region is always part of the region's interior.

There is also a sense in which these theories trivialize all mereotopological distinctions in the presence of boundaries. For (a)–(c) imply that if

is uniform, any canonical model that includes the boundaries of its elements satisfies the conditional:

$$C(x,y) = O(x,y).$$

(This is obvious for i=1. For i=2 or 3, it follows from (b) and (c), which imply that every object overlaps its complement.) Hence, in every such model the -abut predicate A defines the empty relation, and so do the predicates of tangential and boundary parthood (TP, BP) and tangential and boundary overlap (TO, BO). We take these results to show that if boundaries are admitted in the domain, uniformly typed theories are inadequate. In fact, this applies not only to uniform types, but to all types where i=j. (See [4, 38] for related material.)

Moving on to non-uniform types, we may note that some theories have been explicitly proposed in the literature, specifically for the case = 2,1,1,0. An early example is to be found in [9], though the topological primitive there is Op rather than C. (One gets a definitionally equivalent characterization of C via the definitions of Section 2. A similar warning applies to some other theories discussed below.) Other examples may be found in [49, 50, 57, 63, 64]. Since parthood P is not defined in terms of the connection primitive C, these theories need at least two distinct primitives (corresponding to the parameters 1 and 2 in the type); but since fusion is typically understood using the same primitive as parthood, a third primitive is not needed (whence the equality of the second and third coordinates in the type).

These theories typically represent an attempt to reconstruct ordinary topological intuitions on top of a mereological basis. In fact, it is immediate from the definition that in this case C corresponds to the notion of connection of ordinary point-set topology: two regions are connected if the closure of one intersects the other, or vice versa. Moreover, P is typically assumed to satisfy the relevant extensionality and inclusion principles.

Thus, a minimal theory of this kind is typically axiomatized using  $(C1_2)$ ,  $(C2_2)$ ,  $(P1_1)$ ,  $(P2_1)$ ,  $(P3_1)$ ,  $(P5_{12})$ . If we also add the fusion principle  $(C4_1)$ , the result is a mereotopology subsuming what is known as classical extensional mereology [12, 54], in which P defines a complete Boolean algebra with the null element deleted. And if we add the following:

(A1') P (x, c (x))

(A2') P (c (c (x)), c (x))

(A3') E (c (x) + c (y), c (x + y))

the result is what may be called a full mereotopology, in which c behaves like the standard Kuratowski closure operator. (A0 has no analogue due to the lack of a null element.)

All of these theories, of course, must account in some way for the intuitive difficulties that arise out of the notion of a boundary, and correspondingly of the distinction between open and closed entities. For instance, Smith [57] considers various ways of supplementing a full mereotopology with a rendering of the intuition that boundaries are ontologically dependent entities [13], i.e., can only exist as boundaries of some open entity (contrary to the ordinary set-theoretic conception). In our notation the simplest formulation of this intuition is given by the axiom:

(B1) BP (x, y) z(Op (z) BP (x, c (z))).

(A more general formulation can be given using a primitive predicate to express the dependence relation; see [32] for a study of the relevant axiomatics.) Other proposals exploit a distinction between "bona fide" boundaries, corresponding to natural discontinuities (the edge of an island), and the "fiat" boundaries induced by our cognitive and/or social practices (the borders of Wyoming). See [59] for a discussion of these options.

It is also noteworthy that all theories of this sort have type 2,1,1,0. We conjecture that this is indeed the only viable option. For instance, it is easy to see that any 1,2,k,0 -theory would immediately run into the troubles mentioned in (a)–(b) above.

# 6. Boundary-free Theories

Though the idea of a uniform type appears to founder in the case of boundary-based theories, it has been taken very seriously in the context of boundary-free theories, i.e., theories that leave out boundaries from the universe of discourse in the intended models. Theories of this sort are rooted in [25, 67] and have recently become popular under the impact of Clarke's formulation in [16, 17] (see also [38]). Clarke's own is a 1,1,1,0 - theory, and some later authors followed this account (e.g. [1, 2, 52]). However, one also finds examples of theories of type 2,2,2,0 (e.g. in [42, 51]) as well as of type 3,3,3,0 (especially in the work of Cohn *et al.* [20, 22,41,53], which has led to an extended body of results and applications in the area of spatial reasoning; see [33] for an independent example of a type 3,3,3,0 theory.) Indeed, to our knowledge all boundary-free theories in the literature are uniformly typed: this is remarkable but not surprising, since the main difficulties in reducing mereology to topology lies precisely in the presence of boundaries.

Now, by definition, a boundary-free -theory admits of no boundary elements. In axiomatic terms, this is typically accomplished by adding a further postulate to the effect that everything is a region (i.e., has interior parts):

(R)  $x \operatorname{Rg}(x)$ ,

which implies the emptiness of the relations BP and BO, hence of Bd. So b (x) is never defined in this case. However, let us emphasize that even in a boundary-free theory boundaries *may* be included, not in a model's domain, but in the topological space relative to which the model is defined. Moreover, it is worth noting that such theories typically afford some indirect way of modeling boundary talk, e.g., as talk about infinite series of extended regions. (Cf. [5, 17, 31].) In this sense, these theories do have room for boundary elements, albeit only as higher-order entities.

Note also that an axiom such as (R) gives us a way of studying the spectrum of boundary-free theories in terms of their boundary-tolerant counterparts. To this end, define the *-region relativization* of a formula , written Rg:

$$C (y,x)^{Rg} =_{df} C (y,x)$$

$$(\neg )^{Rg} =_{df} \neg ( {}^{Rg} )$$

$$( )^{Rg} =_{df} {}^{Rg} {}^{Rg}$$

$$( x )^{Rg} =_{df} x(Rg (x) {}^{Rg} )$$

$$( x )^{Rg} =_{df} x(Rg (x) {}^{Rg} )$$

By ordinary induction on , one immediately verifies that the following holds in every (canonical) model:

$$x \operatorname{Rg}(x)$$
 (  $\operatorname{Rg}$ ).

It follows that in general a formula is a theorem of a boundary-free - theory iff its relativization  $R_g$  is a theorem of the boundary-tolerant theory obtained by dropping (R).

More specifically, consider now the three main options mentioned in the previous section, where is a basic uniform type of the form i,i,i,0. Unlike their boundary-based counterparts, none of these options yields a collapse of the distinction between tangential and interior parthood (TP, IP) or between tangential and interior overlap (TO, IO). However, the three options diverge noticeably with regard to the distinction between open and closed regions (Op, Cl). The general picture is as follows.

(a) The case i=1 allows for the open/closed distinction, yielding theories in which the relation of abutting (A) is a prerogative of closed regions (open regions abut nothing). As a corollary, such theories determine nonstandard mereologies that violate the supplementation principle:

(S) z(P(z, x) = O(z, y)) = P(x, y)

(It is enough to take *y* open and *x* equal to the closure of *y*.) This is so even if the theory includes the extensionality axioms  $(P3_{1=})$ ,  $(C3_{1=})$ , or  $(P4_{1=})$ . For although extensionality guarantees that a closed region *x* is never part of its own interior *y*, this is due to a mereological difference (a boundary) that cannot be found in the domain of regions. This is a feature that some authors have found unpalatable: as Simons [54] put it, one can discriminate regions that differ by as little as a point, but one cannot discriminate the point. There are also some topological peculiarities that follow from the choice of  $C_1$  as a connected to its complement, hence that the universe is bound to be disconnected. This was noted in [1, 17], where the suggestion is made that self-connectedness should be redefined accordingly:

Cn' (x) =<sub>df</sub> y 
$$z(E(x, y+z)) = C(c(y), c(z))).$$

This, however, is just a way of saying that self-connectedness must be defined with reference to a different notion of connection (namely, the notion obtained by taking i=3.)

(b) The case i=2 also allows for the open/closed distinction, but yields theories in which the relation of abutting may only hold between two regions one of which is open and the other closed in the relevant contact area. This results in a rather standard topological apparatus, modulo the absence of boundary elements. However, also in this case the mereology is bound to violate (S). (Again, just take *y* open and *x* equal to the closure of *y*.)

(c) The case i=3 is the only one where the open/closed distinction dissolves: in this case every region turns out to be -equal to its interior as well as to its closure. This follows from (P3), i.e., equivalently, from (C3) or (P4). This means that -theories of this sort cannot be extensional—in fact, they yield highly non-standard mereologies. However, this is coherent with the fundamental idea of a boundary-free approach. For one of the main motivations for going boundary-free is precisely to avoid the many conundrums that seem to arise from the distinction between open and closed regions [41]. In addition, and for this very same reason, such theories can validate (S), thereby eschewing the problem mentioned in (a) and (b).

We are not aware of any non-uniformly typed boundary-free theories. However, one may imagine that such theories could also alleviate some of the unpalatable properties of the uniformly typed mereotopologies mentioned in (a) and (b). For example, a type of the form 1,3,k,0 would correspond to a mereotopology in which a type-1 notion of connection is combined with a type-3 parthood relation that satisfies the supplementation principle (S). Similarly with a type of the form 2,3,k,0. It would certainly be interesting to pursue abstract studies in this direction. We hope our framework may constitute a first step towards this possibility.

## 7. An Orthogonal Dimension of Variety

So far, we have only been concerned with the different ways of interpreting the connection relation *vis-à-vis* the options made available by the open/closed distinction. In each case, we have followed the familiar course of explaining connection in terms of boundary sharing, irrespective of the size (dimension) of the relevant boundary. This means that two regions qualify as connected even if they share a single boundary point, as illustrated in Figure 6. For many purposes, however, this may be considered too weak a form of connection. For example, a worm cannot pass from the interior of one apple to another, which touch just at a point, without becoming visible to the exterior-so from the worm's point of view we might as well say that the apples are not "sufficiently" connected. In the remainder of the paper we extend our study by taking a closer look at the options that are available in this respect: two regions may share a single boundary point or a boundary section of increasing extent. This gives us a second, orthogonal dimension along which varieties of topological connection can be classified—the strength of the connection. Indeed, one can find mereotopologies (e.g. [12, 39]) in which a predicate is defined to distinguish simple point connection from a stronger form of connection (where a worm could travel from one body to another without becoming visible to the exterior). There are also whole mereotopological theories that have been built taking the stronger notion of connection as primitive (see e.g. [7]). So the general question arises as to what varieties of strength there might be. Further dimensions may then be obtained by taking into consideration the topological structure of the regions allowed in the domain of quantification-for instance, the distinction between one-piece and multipiece regions.



Figure 6. A very weak kind of connection.

Examining the limiting cases, it would appear that there are four main degrees of strength whereby two regions x and y may be said to be connected, depending on whether their common boundary is (a) a single point (or, more generally, a boundary of dimension d-2 if x and y are of dimension d), (b) an extended boundary portion (of dimension d-1), (c) a maximal boundary, or (d) a complete, all-encompassing boundary. These four cases are illustrated in Figure 7. There are twelve patterns overall, since each case can be construed differently depending on which connection relation  $C_n$  is used to cash out the concept of a "common boundary".

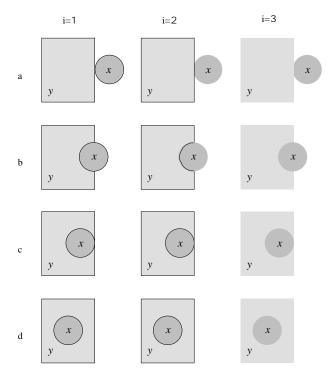


Figure 7. Connection relations of increasing strength (limit cases): four main patterns for each basic type of connection illustrated in Section 3.

Before proceeding to a more precise characterization, however, a few remarks are in order. First, here as elsewhere our illustrative examples are two-dimensional, but this is simply for ease of drawing; the relevant distinctions are applicable to higher dimensions as well (though obviously not to lower dimensions). For instance, in 3D space the patterns in the first row of Figure 7 correspond to the case where two bodies touch at a point, but we also want to include here the case where two cubes (for example) touch along an edge. The two cases are distinct and could be further ordered in terms of increasing strength, but from the present perspective we shall treat them as equally weak: in both cases a worm cannot travel from one region to the other without becoming visible to the exterior. Our definitions below will implement this intuition.

Second, the patterns in the two bottom rows of Figure 7 may suggest that the corresponding connection relations are asymmetric; it appears as

though one of the two regions *surrounds* the other. However it is not our intention to model this closely related concept. Clearly if y surrounds x then the two regions share a "maximal" boundary as in the figure—a boundary that is maximally connected relative to x. Yet this need not be the only way that this sort of connection can be achieved. For example, suppose x and y are defined by splitting of 2D Euclidean space into two half planes: then neither region surrounds the other, but they nonetheless share a boundary that is maximal in this sense. Other examples of maximal connection without surrounding will become apparent as soon as we consider multi-piece regions (see Section 9).

Third, it is worth pointing out that connections of type c need not have just a single point of "imperfection" where the exterior touches both boundaries. Consider, for instance, the configuration depicted in Figure 8*a*. By contrast, notice that the configuration in Figure 8*b* is a case of connection of type b rather than type c: the shared boundary is not maximal in the relevant sense, as it connects to other portions of the boundaries of both regions. And notice that two regions may be connected in more than one way, as in Figure 8*c*; in such cases we want to say that the weaker pattern of connection (here: of type a) holds along with the stronger ones (here: a connection of type b).

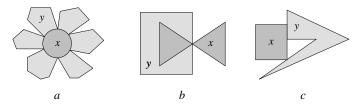


Figure 8. Some examples of imperfect connections.

Finally, it is important to stress that the informal characterizations given above—based on the concept of a boundary—involve several complications, including the emphasis on external connection rather than connection in general. We will by contrast give a set of definitions that do not depend on boundaries but rather rely on the intuition afforded by the example of the worm and the apple mentioned above. These definitions will try to capture the notion of a path connecting the interiors of two regions without touching their common exterior.

Rather than using the term "path", which already has a specific meaning in topology, let us speak of "conduits". Intuitively, a conduit between two regions x and y is a self-connected region z that overlaps the interiors of both x and y, and we may say that a conduit z is *direct* or indirect depending on whether or not the difference between z and x and the difference between z and y are also self-connected (i.e., intuitively, whether or not conduit z crosses the boundaries of x and y only once, which effectively forces z to "start" in x and "end" into the interior of y). Moreover, let us say that a direct conduit z is *ideal* iff, for every internally selfconnected region w which is part of z, the difference of z and w is not a direct conduit between x and y—where a region is internally self-connected iff its interior is self-connected. Thus, intuitively, a conduit may be "pinched" to a point, but an ideal direct conduit is a direct conduit that is minimal with respect to the number of "pinchings". Formally, where is any type, we can express these three notions of conduit as follows.<sup>5</sup>

We can now use these concepts to define four new kinds of connection relations, i.e., more precisely, to distinguish the four degrees of strength for each connection relation of type :

<sup>&</sup>lt;sup>5</sup> Here, as elsewhere, our definitions are given against the general set-theoretic background of Section 3. It is understood that some of these definitions may not be significant for some mereotopological theories. For example, a theory that does not assume the existence of an interior i (x) for every region x in the domain may not rely on the given definition of ideal -conduit. An alternative definition that does not rest on this assumption (but rather on the existence of a closure) can be obtained by replacing the clause 'Cn (i (w))' by the following conditional:

 $<sup>\</sup>begin{array}{rrrr} x & y & z((E & (w, x+y) & IP & (c & (z), x)) \\ & v(Cn & (v) & P & (z, v) & O & (y, v) & IP & (c & (v), c & (w)))). \end{array}$ 

It is easy to verify that these four kinds of relations are ordered in terms of increasing strength for each type . Also, in the domain of extended regions each relation is both reflexive and symmetric, exactly as the generic connection relations discussed in Section 3.

These definitions capture the intended notion of a worm's path. Since every direct conduit is self-connected, the notion of *minimal* connection  $C_{a}$  is just C :

$$C_{a}(x, y) = C(x, y)$$

(Compare Figure 9a.) Next, since a direct conduit cannot connect with the common complement, the *extended* notion of connection  $C_{b}$ , (x, y) forces the contact area to be wider (Figure 9b). For the maximal notion of connection,  $C_c(x, y)$ , the intuition is that rather than just requiring some direct conduit not to intersect the complement, this must be true for every direct conduit. Thus every "path" that starts in x, crosses x's boundary, and enters y must do so directly without intersecting the complement. The ideality restriction is necessary because otherwise there could be an additional component of the direct conduit which overlapped the complement (Figure 9c): one direct conduit is ideal, and illustrates a direct conduit that is not connected to the complement; the other direct conduit connects with the complement, but it is not ideal, since a component could be removed whilst it remained a direct conduit between x and y. Notice that this latter direct conduit with the outer component removed is still connected to the complement; the definition of  $C_{d}$  (x, y) ensures that the non-existence of such a conduit results in a case of *perfect* connection.

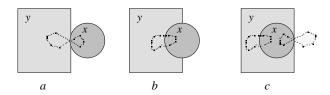


Figure 9. Characterizing the strength of a connection pattern in terms of conduits.

# 8. Discussion

It is worth taking a second look at Figure 7. The intended interpretation is that the differently shaded areas represent distinct, non-overlapping re-

gions (x and y). Thus, although y starts as a rectangle on the top row, it changes shape in each successive row, ending up with a hole inside it. Notice, however, that if we reinterpreted the figure so that y remained a rectangle, then in the bottom three rows the two regions would overlap mereologically. More precisely, in the limit case where the strength of the connection is due exclusively to external contact (without overlap), x would partially overlap, be a tangential part, and be an interior part of y, respectively. This fact deserves consideration. For it shows that there exists a natural homomorphism between the vertical variety of the connection relations of Figure 7 and the horizontal variety (known in the literature as a "conceptual neighborhood diagram" [21, 34] or a "continuity network" [21]) of the basic mereotopological relations of Figure 2.

It also bears emphasis that although one can imagine treating the new dimension of variety as determining a corresponding set of primitive connection relations, the relations introduced in the first part of the paper are analytically more primitive as the definitions of Section 7 rely on those of Section 3. Moreover, in certain circumstances one can move from one degree of strength to the next by progressive movement of one region, as the simple patterns of Figure 7 illustrate, whilst it is always necessary for a region to undergo topological change to move between the basic patterns of Figure 1—specifically, at least one region needs to grow or lose its boundary. This may explain why the relations most familiar in the literature are of the weakest sort: as we have already seen,  $C_{a, 1,1,1,0} = C_1$  is connection  $\hat{a} \, la \, \text{Clarke} \, [16, 17]; \, C_{a, 2,1,1,0} = C_2$  is the analogue of standard topological connection [57, 61]; and  $C_{a, 3,3,3,0} = C_3$  is "RCC" connection [20, 41, 53]. Occasionally, however, stronger relations have also been considered. For example:

(a)  $C_{b, 3,3,3,0}$  is the relation of connection in the sense of Borgo *et al.* [7],<sup>6</sup> which is also essentially Bennett's [3] "firm" connection. (Strictly speaking, this latter notion is defined as a kind of external connection rather connection more generally; see also [44]);

(b)  $C_{c, 2,1,1,0}$  is "niche" connection as described by Smith and Varzi [58], which is in turn is closely related to the notion of tangential surround defined in [57];

<sup>&</sup>lt;sup>6</sup> This relation is essentially equivalent to the relation  $C_5$  of footnote 2.

(c)  $C_{d, 2,1,1,0}$  is closely related to the notion of non-tangential surround defined by Smith [57] (except that this notion is asymmetric).

The significance of the distinctions we have been making can also be illustrated with reference to specific domain examples. We have already mentioned the case of a worm travelling from one body to another, and this is illustrative of a familiar pattern in 3D space. For another example, consider a chunk of Swiss cheese. A hole hidden in the interior—an internal cavity—is  $C_d$ -connected to the piece of cheese (more precisely, it is  $C_d$ , connected for some —but we shall omit the second index when it is of no relevance): a worm cannot leave the hole without going through the cheese. If the worm starts digging, eventually the hole will "open up"—then it will be merely  $C_b$ -connected to the cheese. And at the "magical moment" when the worm sees the light for the first time—when the worm breaks through the last layer of cheese—the hole is  $C_c$ -connected to the cheese, though only for an instant. (See [10, ch. 6].)

Further interesting examples arise in the geographic domain. For instance, consider the portion of the USA illustrated in Figure 10a. Utah and New Mexico are only C<sub>a</sub>-connected (as are Colorado and Arizona). Utah and Colorado are  $C_b$ -connected (as are many other states, e.g. Arizona and Nevada). Utah is C<sub>c</sub>-connected with the sum of all the states except for New Mexico and Utah itself. Finally, Utah is  $C_d$ -connected with the sum of all the other states. One might wonder what motivates taking the sum of a set of States and treating it as a uniform region, but one can easily imagine that there might be some property (e.g., social, economic, meteorological, political) which only some set of States might share. Another, essentially identical example exists in France, where four departments all meet at a single point—see Figure 10b. There are many other geographic examples of these relationships. For example, Vatican City and Italy are C<sub>d</sub>connected, since the former is a separate country surrounded by the latter. The same relationship holds between the state of San Marino and Italy as also between the former countries of East and West Germany (because West Berlin was completely surrounded by East German territory).

Finally, it is worth pointing out that the patterns of interaction displayed in Figure 7 become considerably more complex as soon as other shapes are contemplated. For example, consider replacing the circular shape x with a simple concave shape, as in Figure 11. If we visualize the

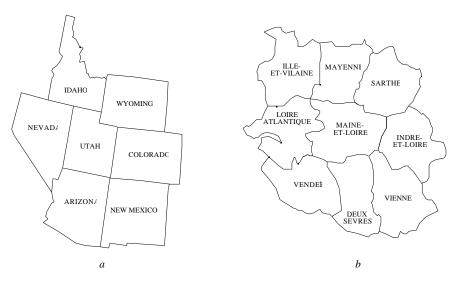


Figure 10. Geographic examples of connections of various strengths.



Figure 11. Connection relations involving a concave, boomerang-shaped region.

sequence of pictures as the "boomerang" sinks into the "water", a variety of topological configurations ensue, with single and multiple connections.

It is tempting to conjecture that the number of connections between two regions might also count as an indicator of the strength of their connection. However, this is not entirely straightforward. There is, for instance, no reason to think that a number of single points of connection are "as good as" a single extended connection. Moreover, we have already noted that two regions may be connected in more than one way (see again Figure 8c) and there is no obvious limit to the patterns of interaction between regions of different shapes and topological genus (e.g., regions with holes). Thus, this line of investigation is bound to be of considerable complexity. (See [36] for a thorough study of  $C_1$ -connections in boundarybased theories—i.e., where  $C_1$  amounts to mereological overlap.)

#### 9. Multi-piece Regions—A Third Dimension?

The intuitive, boundary-based characterization of the four kinds of connection relations introduced in section 7 was meant to apply for one-piece or multi-piece regions alike. However, the formal, conduit-based definitions of these relations may fail for multi-piece regions. Specifically, it is the definitions of the relations  $C_{c_1}$  and  $C_{d_1}$  that may fail, because these relations are defined by demanding that no ideal direct conduit exists satisfying a particular condition; adding a new, separate piece to a region will mean that many new ideal direct conduits exist (from the new piece to the other region) and some of these may fail the condition. For example, consider the configuration depicted in Figure 12; there is certainly an ideal direct conduit that connects *x* to the exterior of *y*.



Figure 12. A peculiar connection pattern involving a connected and a disconnected (multi-piece) region.

If we wish to handle multi-piece regions, then the remedy is fairly straightforward: we must add a further condition selecting a particular component from x and y for the condition to hold over. To this end, let us first define the notion of a maximal self-connected part, or constituent:

$$MCP(x, y) =_{df} CP(x, y) \qquad z(CP(z, y) P(x, z)) = E(x, z))$$

Then we can revise the definitions of  $C_{c,}$  (maximal connection) and  $C_{d,}$  (perfect connection) accordingly:

$$C_{c,}(x,y) =_{df} C(x,y) \quad x' \quad y'(MCP(x',x) \quad MCP(y',y)$$
  

$$\neg \quad z(IdCd(z,x',y') \quad O(z,k(x+y))))$$
  

$$C_{d,}(x,y) =_{df} C(x,y) \quad x' \quad y'(MCP(x',x) \quad MCP(y',y)$$
  

$$\neg \quad z(IdCd(z,x',y') \quad C(z,k(x+y))))$$

Considering the notion of multi-piece regions now leads to the idea that the degree of connection between the various constituents of a multipiece region is a third dimension of variation of the connection relation. Consider two multi-piece regions x and y. The weakest form of connection is that each has a single constituent, x' and y' respectively, which is connected to the other by one of the twelve connection relations. A stronger form of multi-piece connection is that *every* constituent of one region connects to some constituent of the other. And by quantifying appropriately, we can come up with a new variety of connection relations.

Let us add a third subscript, chosen from the initial portion of the Greek alphabet to indicate this variety. Where = i,j,k, ' is any type and

 $\{a, b, c, d\}$ , we can define the following four relations, indexed by , , , and :

$C_{, y} =_{df}$	x' y'(MCP(x',x))	MCP(y',y) <b>0</b>	$C_{(x,y)}$
$\mathbf{C}  (x,y) =_{\mathrm{df}}$	x'(MCP(x',x))	y'(MCP (y',y)	C(x,y)))
	<i>y</i> ′(MCP ( <i>y</i> ′, <i>y</i> )	x'(MCP(x',x))	C((x,y))))
$C_{\mu}(x,y) =_{df}$	x'(MCP(x',x))	<i>y'</i> (MCP ( <i>y'</i> , <i>y</i> )	C(x,y)))
	<i>y</i> ′(MCP ( <i>y</i> ′, <i>y</i> )	x'(MCP(x',x))	C(x,y))))
C (x,y) = <sub>df</sub>	<i>x' y'</i> (MCP ( <i>x'</i> , <i>x</i> )	MCP (y', y)	C(x,y)

These relations are illustrated in Figure 13 for the case i = 1. The first row illustrates four distinguishing cases of C<sub>,a,1</sub>, C<sub>,a,1</sub>, C<sub>,a,1</sub>, and C<sub>,a,1</sub><sup>-</sup> connection, in order from left to right. The next rows illustrate the corresponding patterns for = b, c, d, respectively. Notice that for and only one of the two regions can have multiple constituents. For C<sub>,d,1</sub> we show two alternative ways of achieving the relationship, but note that the left-hand configuration is not extendible further, while the right-hand configuration can be extended by adding further lighter-shaded constituents. Also notice that, in the figure, all individual connections are of the same kind; multiple connections of varying types are of course possible, though it would take us too long to analyze them here.

As with the relations of section 7, it is not difficult to think of concrete examples in which these relations are instantiated—for instance in the geographic domain. Consider a typical city in the United States in which all the avenues and streets are orthogonal and each street crosses each avenue and vice versa: in this case we can assert that  $C_{++}(x, y)$ , where x is the

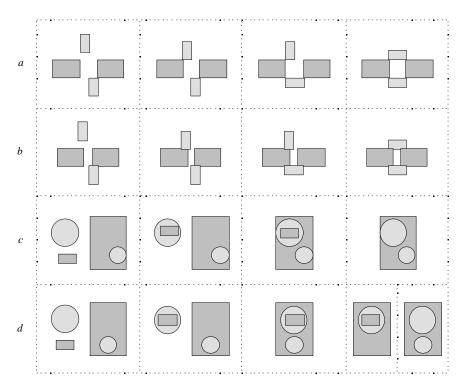


Figure 13. Varieties of multiple connection.

-fusion of the avenues and y the -fusion of the streets. In the case of a city like Manhattan only the weaker relationship  $C_{-,-}(x,y)$  holds, because only Broadway extends all the way to the southernmost street of the island. The -fusions of all American rivers and all American states stand in a  $C_{-,-}$  -relationship, since every river is connected to some state and vice versa. Finally, an example of  $C_{-,-}$  is afforded by the relationship between the -fusion of all airplanes and the -fusion of all airports. At any time, some airplane is connected to (at) an airport, though there is no guarantee that this holds for all airplanes.

Two final remarks are in order. First, we have been careful to define these relations so that they are symmetric: the first and last definitions have quantifiers of the same type, which naturally commute, whereas the other definitions contain a "vice versa" conjunct. On the other hand, the first two definitions yield reflexive relations but the third and fourth do not. This is obvious in view of the fact that a multi-piece region x cannot, by definition, have each of its constituents connected to the others. There appears to be no natural way to make these definitions reflexive. We include them nonetheless as examples of interesting connection relations that violate this familiar postulate on the concept of mereotopological connection. (In certain domains it would also be natural to consider connection relations which are not symmetric; for example one way streets, valves, and hyperlinks.)

Second, it is appropriate to point out that two further relations can be defined which relate in a natural way to  $C_{\mu}$  and  $C_{\mu}$ , where the quantifiers are different (hence non-commutative). This can be done by weakening the "vice versa" clause to a disjunction rather than a conjunction:

C ( <i>x</i> , <i>y</i> ) = <sub>df</sub>	x'(MCP(x',x))	y'(MCP (y',y)	C(x,y)))
	<i>y</i> ′(MCP ( <i>y</i> ′, <i>y</i> )	x'(MCP(x',x))	C((x,y))))
C ', (x,y) = <sub>df</sub>	x'(MCP(x',x))	<i>y</i> ′(MCP ( <i>y</i> ′, <i>y</i> )	C(x,y)))
	<i>y'</i> (MCP ( <i>y'</i> , <i>y</i> )	x'(MCP(x',x))	C(x,y))))

The resulting relations (see Figure 14) are weaker than C  $_{/}$  and C  $_{/}$ , respectively, but stronger than C  $_{/}$  and C  $_{/}$ . Thus, overall the relations are ordered in terms of increasing strength according to the pattern , ', ', ', ', '

#### **10. Final Comments**

The analysis of connection and connection-based theories presented here certainly does not exhaust all possibilities. For example, we have not explicitly investigated irregular regions or regions of higher-order topological genus (i.e., regions with holes) and this may bear explicit investigation. Equally, the analysis could be extended to connection relations between spatial entities of differing dimensions (cf. [19, 30, 35, 40]).

Another possibility is to extend the analysis to cover for instance the notion of "weak connection" defined in [1], the Brentanian notion of connection [8, 15, 56], or the notion of connection through "fiat" boundaries [59]. Yet another notion of connection is given by a pair of linked (interlocked) tori [27, 65].

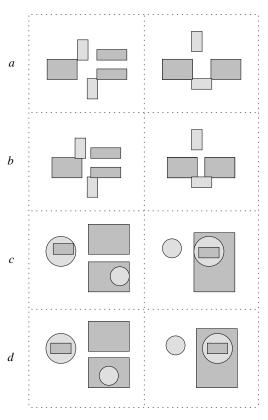


Figure 14. Further connection relations with multi-piece regions.

It is also worth pointing to the work of Egenhofer and Franzosa [28, 29], who present a calculus that allows, at the cost of arbitrary complexity, the possibility of classifying any topological distinct situation. (Compare [18] for a related proposal.)

The variety of mereotopological connection relations is very rich indeed. We hope to have gone some way in the direction of a unified framework that allows one to see the forest besides the many trees.

## Acknowledgements

This paper grew out of two conference presentations, [23] and [24]. We are grateful to the audiences of both meetings for helpful discussion. We are also grateful to Maureen Don-

nelly, Claudio Masolo, Philippe Muller, and Laure Vieu for comments on earlier drafts. The support of the EPSRC under grants GR/K65041 and GR/M56807 is gratefully ac-knowledged.

#### References

- Asher N., Vieu L., 1995, 'Toward a Geometry of Common Sense: A Semantics and a Complete Axiomatization of Mereotopology', in *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, San Mateo (CA): Morgan Kaufmann, pp. 846–852.
- [2] Aurnague M., Vieu L., 1993, 'A Three-Level Approach to the Semantics of Space', in C. Z. Wibbelt (ed.), *The Semantics of Prepositions: From Mental Processing to Natural Language Processing*, Berlin: Mouton de Gruyter, pp. 393–439.
- [3] Bennett, B., 1996, 'Carving Up Space: Steps Towards the Construction of an Absolutely Complete Theory of Spatial Regions', in J. J. Alfres, L. M. Pereira, and E. Orlowska (eds.), *Proceedings of the Fifth European Workshop on Logics in Artificial Intelligence (JELIA'96)*, Berlin: Springer, pp. 337–353
- [4] Biacino L., Gerla G., 1991, 'Connection Structures', Notre Dame Journal of Formal Logic 32, 242–247.
- [5] Biacino L., Gerla G., 1996, 'Connection Structures: Grzegorczyk's and Whitehead's Definitions of Point', *Notre Dame Journal of Formal Logic* 37, 431–439.
- [6] Bolzano B., 1851, Paradoxien des Unendlichen, Leipzig: Reclam (Eng. trans. by D. A. Steele, Paradoxes of the Infinite, London: Routledge & Kegan Paul, 1950).
- [7] Borgo S., Guarino N., Masolo C., 1996, 'A Pointless Theory of Space Based on Strong Connection and Congruence', in L. Aiello, J. Doyle, and S. Shapiro (eds.), *Principles of Knowledge Representation and Reasoning: Proceedings of the Fifth International Conference*, San Francisco (CA): Morgan Kaufmann, pp. 220–229.
- [8] Brentano F., 1976, Philosophische Untersuchungen zu Raum, Zeit und Kontinuum, hrsg. von S. Körner and R. M. Chisholm, Hamburg: Meiner (Eng. trans. by B. Smith, Philosophical Investigations on Space, Time and the Continuum, London: Croom Helm, 1988).
- [9] Cartwright R., 1975, 'Scattered Objects', in K. Lehrer (ed.), *Analysis and Meta-physics*, Dordrecht: Reidel, pp. 153–171.
- [10] Casati R., Varzi A. C., 1994, *Holes and Other Superficialities*, Cambridge, MA, and London: MIT Press/Bradford Books.
- [11] Casati R., Varzi A. C., 1996, 'The Structure of Spatial Localization', *Philosophi-cal Studies* 82, 205–239.
- [12] Casati, R., Varzi, A. C., 1999, Parts and Places. The Structures of Spatial Representation, Cambridge (MA) and London: MIT Press/Bradford Books.
- [13] Chisholm R. M., 1984, 'Boundaries as Dependent Particulars', Grazer philosophische Studien 10, 87–95.
- [14] Chisholm R. M., 1987, 'Scattered Objects', in J. J. Thomson (ed.), On Being and

Saying: Essays for Richard Cartwright, Cambridge (MA): MIT Press, pp. 167-173.

- [15] Chisholm R. M., 1992/3, 'Spatial Continuity and the Theory of Part and Whole. A Brentano Study', *Brentano Studien* 4, 11–23.
- [16] Clarke B. L., 1981, 'A Calculus of Individuals Based on "Connection", Notre Dame Journal of Formal Logic 22, 204–218.
- [17] Clarke B. L., 1985, 'Individuals and Points', Notre Dame Journal of Formal Logic 26, 61–75.
- [18] Clementini, E., Di Felice, P., 1998, 'Topological Invariants for Lines', IEEE Transactions on Knowledge and Data Engineering 10, 38–54
- [19] Clementini, E., Di Felice, P., Oosterom, P., 1993, 'A Small Set of Formal Topological Relationships Suitable for End User Interaction', in D. Abel and B. C. Ooi (eds.), Advances in Spatial Databases, Proceedings of the Third International Symposium (SSD'93), Berlin: Springer, pp. 277–295.
- [20] Cohn, A. G., Bennett, B., Gooday, J., Gotts. N., 1997, 'RCC: A Calculus for Region Based Qualitative Spatial Reasoning', *GeoInformatica* 1, 275–316.
- [21] Cohn A. G., Hazarika S. M., 2001, 'Qualitative Spatial Representation and Reasoning: An Overview', *Fundamenta Informaticae* 45, 1–29.
- [22] Cohn A. G., Randell D. A., Cui Z., 1995, 'A Taxonomy of Logically Defined Qualitative Spatial Regions', *International Journal of Human-Computer Studies* 43, 831–846.
- [23] Cohn A. G., Varzi A. C., 1998, 'Connection Relations in Mereotopology', in H. Prade (ed.), *Proceedings of the 13th European Conference on Artificial Intelli*gence, Chichester: Wiley, pp. 150–154.
- [24] Cohn A. G., Varzi A. C., 1999, 'Modes of Connection', in C. Freksa and D. M. Mark (eds.), Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science, Berlin: Springer, pp. 299–314.
- [25] De Laguna T., 1922, 'Point, Line, and Surface, as Sets of Solids', Journal of Philosophy 19, 449–461.
- [26] Düntsch I., Orlowska E., 2000, 'A Proof System for Contact Relation Algebras', *Journal of Philosophical Logic* 29: 241–262.
- [27] Egenhofer M. J., Clementini E., Di Felice P., 1994, 'Toplogical Relations Between Regions with Holes', *International Journal of Geographical Information Systems* 8, 129–144.
- [28] Egenhofer M. J., Franzosa R. D., 1991, 'Point-Set Topological Spatial Relations', International Journal of Geographical Information Systems 5, 161–174.
- [29] Egenhofer M. J., Franzosa R. D., 1995, 'On the Equivalence of Topological Relations', *International Journal of Geographical Information Systems* 9, 133–152.
- [30] Egenhofer, M., Herring, J., 1994, 'Categorizing Topological Spatial Relationships between Point, Line and Area Objects', in *The 9-Intersection: Formalism and Its Use for Natural Language Spatial Predicates*, Technical Report 94-1, National Center for Geographic Information and Analysis, Santa Barbara.

- [31] Eschenbach C., 1994, 'A Mereotopological Definition of "Point", in C. Eschenbach, C. Habel, and B. Smith (eds.), *Topological Foundations of Cognitive Sci*ence. Papers from the Workshop at the First International Summer Institute in Cognitive Science (Buffalo 1994), University of Hamburg, Reports of the Doctoral Program in Cognitive Science, No. 37, pp. 63–80.
- [32] Fine K., 1995, 'Part–Whole', in B. Smith and D. W. Smith (eds.), *The Cambridge Companion to Husserl*, New York: Cambridge University Press, pp. 463–485.
- [33] Fleck M. M., 1996, 'The Topology of Boundaries', Artificial Intelligence 80, 1-27.
- [34] Freksa C., 1992, 'Temporal Reasoning Based on Semi-Intervals', Artificial Intelligence 54, 199–227.
- [35] Galton A. P., 1996, 'Taking Dimension Seriously in Qualitative Spatial Reasoning', in W. Wahlster (ed.), Proceedings of the 12th European Conference on Artificial Intelligence (ECAI 96), Chichester: Wiley, pp. 501–505.
- [36] Galton A., 1998, 'Modes of Overlap', *Journal of Visual Languages and Comput*ing 9, 61–79.
- [37] Galton, A. P., 2000, Qualitative Spatial Change, Oxford: Oxford University Press.
- [38] Gerla G., 1995, 'Pointless Geometries', Chapter 18 of F. Buekenhout (ed.), Handbook of Incidence Geometry, Amsterdam: Elsevier, pp. 1015–1031.
- [39] Gotts N. M., 1994a, 'Defining a 'Doughnut' Made Difficult', in C. Eschenbach, C. Habel, and B. Smith (eds.), *Topological Foundations of Cognitive Science. Papers from the Workshop at the First International Summer Institute in Cognitive Science*, University of Hamburg, Reports of the Doctoral Program in Cognitive Science, No. 37, 105–129.
- [40] Gotts N. M., 1996, 'Formalising Commonsense Topology: The INCH Calculus', in H. Kautz and B. Selman (eds.), *Proceedings of the Fourth International Sympo*sium on Artificial Intelligence and Mathematics (AI/Math-96), Fort Lauderdale (FL), pp. 72–75.
- [41] Gotts N. M., Gooday J. M., Cohn A. G., 1996, 'A Connection Based Approach to Common-Sense Topological Description and Reasoning', *The Monist* 79, 51–75.
- [42] Grzegorczyk A., 1960, 'Axiomatizability of Geometry Without Points', *Synthese*, 12, 109–127.
- [43] Guarino, N. (ed.), 1998, Formal Ontology in Information Systems. Proceedings of FOIS-98, Amsterdam: IOS Press.
- [44] Hazarika S. M., Cohn A. G., 2001, 'Qualitative Spatio-Temporal Continuity', in D. R Montello (ed.), Spatial Information Theory: Foundations of Geographic Information Science, Berlin: Springer, pp. 92–107.
- [45] Husserl E., 1901, Logische Untersuchungen. Zweiter Band. Untersuchungen zur Phänomenologie und Theorie der Erkenntnis, Halle: Niemeyer (2nd ed. 1913; Eng. trans. by J. N. Findlay, Logical Investigations, London: Routledge & Kegan Paul, 1970).
- [46] Kuratowski, C., 1922, 'Sur l'opération A<sup>-</sup> de l'Analysis Situs', Fundamenta Mathematicae 3, 182–199.

- [47] Leśniewski S., 1916, Podstawy ogólnej teoryi mnogości. I, Moskow: Prace Polskiego Kola Naukowego w Moskwie (Eng. trans. by D. I. Barnett, 'Foundations of the General Theory of Sets. I', in S. Leśniewski, Collected Works, ed. S. J. Surma, J. Srzednicki, D. I. Barnett, and F. V. Rickey, Dordrecht: Nijhoff, 1992, Vol. 1, pp. 129–173).
- [48] Leonard H. S., Goodman N., 1940, 'The Calculus of Individuals and Its Uses', *Journal of Symbolic Logic* 5, 45–55.
- [49] Pianesi F., Varzi A. C., 1996a, 'Refining Temporal Reference in Event Structures', *Notre Dame Journal of Formal Logic* 37, 71–83.
- [50] Pianesi F., Varzi A. C., 1996b, 'Events, Topology, and Temporal Relations', *The Monist* 78, 89–116.
- [51] Pratt I., Schoop D., 1998, 'A Complete Axiom System for Polygonal Mereotopology of the Real Plane', *Journal of Philosophical Logic* 27, 621–658.
- [52] Randell D. A., Cohn A. G., 1989, 'Modelling Topological and Metrical Properties in Physical Processes', in R. J. Brachman, H. J. Levesque, and R. Reiter (eds.), *Principles of Knowledge Representation and Reasoning. Proceedings of the First International Conference*, Los Altos (CA): Morgan Kaufmann, pp. 357-368.
- [53] Randell D. A., Cui Z., Cohn A. G., 1992, 'A Spatial Logic Based on Regions and Connection', in B. Nebel, C. Rich and W. Swartout (eds.), *Principles of Knowledge Representation and Reasoning. Proceedings of the Third International Conference*, Los Altos (CA): Morgan Kaufmann, pp. 165–176.
- [54] Simons P. M., 1987, Parts. A Study in Ontology, Oxford: Clarendon Press.
- [55] Simons P. M., 1991, 'Free Part-Whole Theory', in K. Lambert (ed.), *Philosophical Applications of Free Logic*, Oxford: Oxford University Press, pp. 285–306.
- [56] Smith B., 1997, 'Boundaries', in L. Hahn (ed.), *The Philosophy of Roderick Chisholm*, LaSalle (IL): Open Court, pp. 534–561.
- [57] Smith, B., 1996, 'Mereotopology: A Theory of Parts and Boundaries', *Data and Knowledge Engineering* 20, 287–304.
- [58] Smith, B., Varzi, A. C., 1999, 'The Niche', Noûs 33, 214–238.
- [59] Smith, B., Varzi, A. C., 2000, 'Fiat and Bona Fide Boundaries', *Philosophy and Phenomenological Research* 60, 401–420.
- [60] Stock O. (ed.), 1997, Spatial and Temporal Reasoning, Dordrecht: Kluwer.
- [61] Tarski A., 1929, 'Les fondements de la géométrie des corps', Ksiega Pamiatkowa Pierwszkego Polskiego Zjazdu Matematycznego, suppl. to Annales de la Société Polonaise de Mathématique 7, 29–33.
- [62] Tarski A., 1935, 'Zur Grundlegung der Booleschen Algebra. I', Fundamenta Mathematicae, 24, 177–198.
- [63] Tiles J. E., 1981, Things That Happen, Aberdeen: Aberdeen University Press.
- [64] Varzi A. C., 1996a, 'Parts, Wholes, and Part-Whole Relations: The Prospects of Mereotopology,' *Data and Knowledge Engineering* 20, 259–286.
- [65] Varzi A. C., 1996b, 'Reasoning about Space: The Hole Story', *Logic and Logical Philosophy* 4, 3–39.

- [66] Varzi A. C., 1997, 'Boundaries, Continuity, and Contact', Noûs 31, 26-58.
- [67] Whitehead A. N., 1929, *Process and Reality. An Essay in Cosmology*, New York: Macmillan.
- [68] Zimmerman D. W., 1996, 'Indivisible Parts and Extended Objects: Some Philosophical Episodes from Topology's Prehistory', *The Monist* 79, 148–180.