

BF
311
T67
1994

KOGNITIONSWISSENSCHAFT

**Carola Eschenbach,
Christopher Habel &
Barry Smith**

Topological Foundations of Cognitive Science

**Papers from the Workshop at the *FISI-CS*
Buffalo, NY. July 9-10, 1994**

<Bericht Nr. 37, November 1994>

H

M

B

U

R

G



Contents

Topological Foundations of Cognitive Science Barry Smith	3
The Bounds of Axiomatisation Graham White	23
Rethinking Boundaries Wojciech Żelaniec	31
Sheaf Mereology and Space Cognition Jean Petitot	41
A Mereotopological Definition of 'Point' Carola Eschenbach	63
Discreteness, Finiteness, and the Structure of Topological Spaces Christopher Habel	81
Mass Reference and the Geometry of Solids Almerindo E. Ojeda	91
Defining a 'Doughnut' Made Difficult N.M. Gotts	105
A Theory of Spatial Regions with Indeterminate Boundaries A.G. Cohn and N.M. Gotts	131
Mereotopological Construction of Time from Events Fabio Pianesi and Achille C. Varzi	151
Computational Mereology: A Study of Part-of Relations for Multi-media Indexing Wlodek Zadrozny and Michelle Kim	173
Information about the authors	193

Topological Foundations of Cognitive Science

Introductory Talk at the FISI Workshop on 9–10 July 1994

Barry Smith

Department of Philosophy and Center for Cognitive Science

SUNY Buffalo

Buffalo NY 14260-1010

phismith@ubvms.cc.buffalo.edu

I shall attempt to set out the basic issues of this workshop by introducing the concepts at the heart of topology in an informal and intuitive fashion. Two well-known alternatives present themselves to this end; while these prove to be equivalent from the mathematical point of view, they point to distinct sorts of extensions and applications from the perspective of cognitive science.

1. The Concept of Transformation

On the one hand we can take as our starting point the notion of *transformation*. We note, familiarly, that we can transform a spatial body such as a sheet of rubber in various ways. We can invert it, stretch or compress it, move it, bend it, twist it or otherwise knead it out of shape. Certain properties of the body will in general be invariant under transformations such as these: transformations which are neutral as to shape, size, motion and orientation, transformations which are such that they cannot affect the possibility of our connecting two points on the surface or in the interior of the body by means of a continuous line. Let us provisionally use the term 'topological spatial properties' to refer to those spatial properties of bodies which are invariant under transformations such as these (broadly: transformations which do not affect the *integrity* of the body, path, or other sort of spatial structure with which we begin). Topological spatial properties will then in general fail to be invariant under more radical transformations, for example those which involve tearing or gluing or carving holes through a body, or which lead to a decomposition of the body into separate constituent parts.

The class of phenomena structured by topological spatial properties is wider than the class of phenomena to which, for example, Euclidean geometry, with its determinate Euclidean metric, can be applied. Everything (spatial) has topological spatial properties. (This holds, too, of mental images of spatially extended bodies: Cf. Brentano 1988.)

The property of being a (single, connected) body is a topological spatial property, as also are certain properties relating to the possession of holes (more specifically:

properties relating to the possession of tunnels and internal cavities). The property of being a *collection* of bodies and that of being an *undetached part* of a body, too, are topological spatial properties. Thus it is a topological spatial property of a pack of playing cards that it consists of this or that number of *separate* cards and it is a topological spatial property of my arm that it is *connected* to my body.

This concept of topological spatial properties can of course be generalized beyond the spatial case. Thus we can distinguish those properties of temporal structures which are invariant under transformations of (for example) stretching (slowing down, speeding up) and temporal translocation. Intervals of time, melodies, simple and complex events, actions and processes can be seen to possess topological properties. Thus the motion of a bouncing ball can be said to be topologically isomorphic to another, slower or faster, motion of, for example, a trout in a lake or a child on a pogo-stick.

2. The Concept of Boundary

On the other hand we can take as our starting point the intuitive notion of *boundary*. We begin, once again, with spatial examples. Imagine a spatial body, for example a solid and homogeneous metal sphere. We can then distinguish, with some intellectual effort, two disjoint parts of the sphere: on the one hand is its *exterior surface*; on the other hand is the difference between the sphere and this exterior surface (that which would result if, *per impossibile*, the latter could be subtracted from the former). Similarly in the temporal realm we can imagine an interval as being composed of its initial and its final points together with its *interior*: the result of subtracting, abstractly, the given points from the interval as a whole. The boundary of an entity is from the point of view of classical mathematical topology also the boundary of the complement of that entity. My outer surface is then also a boundary of that object which results when we imagine me as having been deleted from the rest of the universe. We can imagine, however, a variant topology which would recognize *asymmetric* boundaries, such as we find, for example, in the figure-ground structure as this is manifested in visual perception. As Rubin (1921) and many others have pointed out, the boundary of a figure is experienced as a part of the figure, and not simultaneously as boundary of the ground, which is experienced as running on behind the figure. Something similar applies also in the temporal sphere: the beginning and ending of a race, for example, are not in the same sense boundaries of any complement-entities as they are boundaries of the race itself.

Perhaps the best examples of asymmetrical boundaries are manifested in the conceptual sphere. We may imagine the instances of a concept arranged in a quasi-spatial way, as happens for example in familiar accounts of colour- or tone-space. Suppose that this is done in such a fashion that the most typical instances are located in the centre of the relevant region and the less typical instances located at distances from this centre in proportion to their degree of non-typicality. (Degree of non-typicality is thus employed to define a crude topological distance-function on the space of instances in question.) Boundary or fringe cases would then be those which are so untypical that even the slightest further deviation from the norm would imply that they are no longer instances of the given concept at all. The notion of *similarity* can be understood in this light as a

topological notion (Mostowski 1983): *a* is similar to *b* might mean, for example, that the colours of *a* and *b* lie close together in colour-space, so that they cannot be discriminated with the naked eye. The similarity relation is in general symmetric and reflexive, but it falls short of transitivity; thus it partitions the initial space not into tidily disjoint and exhaustive equivalence classes, but rather into overlapping *circles of similars* or *natural kinds* (and this falling short of the discreteness of partitions of the sort which are generated by equivalence relations is characteristic of non-trivial topological structures). The boundaries of these circles, constituted by the fringe-instances just referred to, may then be asymmetrical in the sense that the given fringe-instances need not be fringe-instances of any neighbouring natural kinds (or of any putative complement kinds): the transition from *red* to *yellow* may be continuous in this sense, but this is not the case in regard to the transition from, say, *dog* to *cat* or from *cyclo-octatetrene* to *cyclobutadiene*.

3. Closure

The two approaches briefly sketched above may be unified into a single system by means of the notion of *closure*, which we can think of as an operation of such a sort that, when applied to an entity *x* it results in a whole which comprehends both *x* and its (or the nearest circumcluding) boundaries. (We shall concentrate here primarily on definitions of closure in terms of boundaries, and shall mention only in passing the connections between closure and topological transformations.) We employ as basis of our definition of closure the notions of mereology. (Smith 1993, Varzi 1994) Some of the reasons why we shun the set-theoretical instruments employed in standard presentations of the foundations of topology will be set out below.

The operation of *closure* (*c*) is now defined in such a way as to satisfy the following axioms (where we shall employ ' \leq ' to signify the relation 'is a proper or improper part of', and where the range of variables can be assumed to be *regions* of one or other type).

(AC1) $x \leq c(x)$ (expansiveness)

(each object is a part of its closure)

(AC2) $c(c(x)) \leq c(x)$ (idempotence)

(the closure of the closure adds nothing to the closure of an object)

(AC3) $c(x \cup y) = c(x) \cup c(y)$ (additivity)

(the closure of the sum of two objects is equal to the sum of their closures)

These axioms were first set out by the Hungarian topologist Friedrich Riesz in 1906, and independently by the Pole Kazimierz Kuratowski in 1922. The axioms define a well-known kind of structure, that of a *closure algebra*, which is the algebraic equivalent of the simplest kind of topological space. The structure is algebraic in the sense that its definition does not involve the notion of point. Note that Kuratowski's original list includes in addition the following axiom, where 0 would be a null element (the empty set, or empty region) which would complete the underlying lattice from below:

$$(AC0) \quad c(0) = 0 \quad (\text{zero})$$

Note that various modifications and weakenings of the original Riesz-Kuratowski axioms are possible which, as has been shown by Ore, Hammer, Nöbeling, Netzer and others, still preserve the possibility of defining analogues of the standard topological notions of boundary, interior, etc. Thus we may drop the axiom of additivity, which, as Hammer puts it, might most properly 'be called the *sterility* axiom. ... it requires that two sets cannot produce anything (a limit point) by union that one of them alone cannot produce.' (1962, p. 65)

On the basis of the notion of closure we can define the standard notion of (symmetrical) *boundary*, $b(x)$, as follows:

$$(DB) \quad b(x) := c(x) \cap c(x') \quad (\text{boundary})$$

where x' is the mereological complement of x (that which remains when x is considered as having been removed from the universe as a whole) and ' \cap ' signifies mereological intersection. Note that it is a consequence of the definition of boundary here supplied that the boundary of an entity is in every case also the boundary of the complement of that entity.

It is possible to define an asymmetrical notion of "border" in standard topological terms, as the intersection of an object with the closure of its complement:

$$(DB^*) \quad b^*(x) = x \cap c(x')$$

As Zarycki showed (1927), the notion of border (like those of closure, interior, and boundary) can serve as a primitive in terms of which a system equivalent to the Kuratowski system can be defined.

Similarly we can define the notion of the interior, $i(x)$, of an object by:

$$(DI) \quad i(x) := x - b(x) \quad (\text{interior})$$

We may define a *closed object* as an object for which $x = c(x)$, and an *open object* as an object whose complement is closed. (We can then use these notions to relate the two approaches to topology distinguished above, in that topological transformations are transformations which map open objects onto open objects.) A closed object is, intuitively, an independent constituent. A closed object may however fall short of being unitary, i.e. it need not be *connected* in the sense that we can proceed from any one point in the object to any other and remain within the confines of the object itself. The notion of *connectedness*, too, is a topological notion, which we can define as follows:

$$(DCn) \quad Cn(x) := \forall yz(x = y \cup z \rightarrow \exists w(w \leq c(y) \cap c(z)))$$

A connected object is such that all ways of splitting the object into two parts yield parts

whose closures overlap. An alternative definition might be:

$$(DCn^*) \quad Cn^*(x) := \forall yz(x = y \cup z \rightarrow (\exists w(w \leq x \ \& \ w \leq c(y)) \vee \exists w(w \leq c(x) \ \& \ w \leq y)))$$

Examination reveals, however, that even two adjacent spheres which are momentarily in contact with each other will satisfy either condition of connectedness as thus defined. For certain purposes, therefore, it is useful to operate in terms of a notion of *strong connectedness* which rules out cases such as this. This latter notion may be defined as follows:

$$(DSCn) \quad Scn(x) := Cn^*(i(x))$$

4. Foundations of Cognitive Science: A Preliminary Survey

Topology is a branch of mathematics which deals with the large class of topological properties in the sense broadly indicated in the above, or with the topological spaces which the corresponding transformations define. Topology in this sense has been exploited by cognitive scientists above all in work on connectionist networks. In the papers that follow, however, the focus is not on connectionist mathematics and on the corresponding neural-network-like structures located on what Smolensky (1988) calls the 'sub-symbolic' level (which is to say on a level of cognitive architecture reputed to lie below that on which meanings are consciously articulated (see also Zeeman 1962)). Rather, these papers are concerned primarily with the application of topological ideas and methods to the 'symbolic' level of cognitive behaviour and to the corresponding objects. Thus we shall be concerned with topological features of the world as this is organized in meaningful human experience, and with topological features of this experience itself. Moreover, we shall be concerned with topology not (or not primarily) as a branch of mathematics. Rather we shall understand topology in a generalized fashion as a theory of a certain family of concepts centred around the notions of 'region', 'connectedness', 'boundary', 'surface', 'point', 'neighbourhood', 'nearness', etc. These concepts will be interpreted in a fashion that is initially inspired by standard mathematical treatments, the work here presented will however involve departures from standard mathematical topology in ways designed to meet the requirements of given subject-matters.

Work of the sort included in the present collection is not of course without precedent. In a range of cognitive science disciplines concepts and theories derived from topology have been utilized already in a variety of ways. Examples of existing work include:

- work on logical semantics in the field of modal and tense logics (for example under the rubric of 'interval semantics'): Tarski 1938, McKinsey and Tarski 1944, Freksa 1992, Cresswell 1978, 1979, 1985, van Benthem 1983, Allen 1981, 1984;
- analyses of natural-language expressions and of the categorizations of objects, events and situations effected thereby: Wildgen 1981, 1982, Heydrich 1982, Habel *et al.* 1988, 1993a, Habel 1989, 1990, Wunderlich and Herweg 1991, Kaufmann 1991, Aurnague and Vieu 1992; Bach 1986, Kamp 1979, Pianesi and Varzi 1994;

- work on verb-aspect and the mass-count distinction with a topological orientation: Mourelatos 1981, Galton 1984, Hoeksema 1985, Desclès 1989, Brandt 1989;
- topological work in cognitive linguistics by Lakoff, Talmy, Langacker, Jackendoff and others; see especially Talmy 1977ff., Brugman and Lakoff 1988, Lakoff 1989, Jackendoff 1991, and the related work of Petitot 1982ff., and Wildgen 1981ff. The importance of topology to the conceptual structuring effected by language is illustrated most easily in the case of prepositions. As Talmy notes, a preposition such as 'in' is magnitude neutral (in a thimble, in a volcano), shape neutral (in a well, in a trench), closure-neutral (in a bowl, in a ball); it is not however discontinuity neutral (in a bell-jar, in a bird cage);
- topological studies of the notion of semantic dependence in text-linguistics: Saloni and Trybulec 1974, Lipski 1974, 1975, Frawley 1980, Heydrich 1989;
- topological work in the theory of geographic information systems: Mark and Frank 1991, Freksa 1991, Herring 1991, Frank *et al.* 1992, Frank and Campari 1993;
- in the investigation of phenomenal spaces and in the spatial properties of mental representations: Freksa 1991, 1992, Varzi 1993, Smith 1993, Casati and Varzi 1994;
- in contributions to qualitative or 'naive' physics and to the field of common-sense reasoning: Randell and Cohn 1989, Thom 1990, Petitot and Smith 1990 and (in press), Cui *et al.* 1992, Randell *et al.* 1992, Smith 1992, Smith and Casati 1994; in studies of perceptual representation, motor learning, granularity, etc.: Hobbs 1985, Habel 1991, 1993, McCalla *et al.* 1992, Zimmermann and Freksa 1993, Zimmer 1994;
- in related naive-physical studies of phenomena such as cutting, friction, connection, grasping, assemblage, and so on: Forbus 1984, Hayes 1985, Hager 1985. Topological structures play a central role in studies of naive physics not least in virtue of the fact that even well-attested departures from true physics on the part of common sense (Bozzi 1958, 1959, McCloskey 1983) leave the topology and vectorial orientation of the underlying physical phenomena invariant: our common sense would thus seem to have a veridical grasp of the topology and broad general orientation of physical phenomena even where it illegitimately modifies the relevant shape and metric properties;
- in image-processing, for example in automatic analysis of X-ray images for medical and other purposes: Randell and Cohn 1989, Cui *et al.* 1992, Randell *et al.* 1992, Latecki 1992;
- philosophically oriented studies of topological categories: Adams 1973, 1984, 1986, Tiles 1981, Simons 1987, Smith 1987ff., Stroll 1988, Chisholm 1989, 1992/93, Bochman 1990, Eschenbach and Heydrich (in press).

In talking somewhat grandly of 'topological foundations for cognitive science', now, we are contending that the topological approach yields not simply a collection of insights and methods in selected fields, but an over-arching perspective for a range of different types of research across the breadth of cognitive science and a common language for the formulation of hypotheses drawn from a variety of seemingly disparate fields.

Initial evidence for the correctness of this view is provided not just by the scope of the inquiries referred to above, but also by the degree to which in different ways they overlap amongst themselves and support each other mutually.

One rationale behind the idea that the inventory of topological concepts can yield a single unified framework for cognitive science turns on the fact that, as we saw, as has often been pointed out (see e.g. Gibson 1986), boundaries are centres of salience not only in the spatial but also in the temporal world (the beginnings and endings of events, the boundaries of qualitative changes for example in the unfolding of speech events: cf. Petitot 1989). Moreover, topological properties are more widely applicable than are those properties (for example of a geometrical sort) with which metric notions are associated. (Topological phenomena may also be referred to as 'qualitative' phenomena.) Metric features have certainly proved highly useful for the purposes of natural science. Given the pervasiveness of qualitative elements in every cognitive dimension, however, and also the similar pervasiveness of notions like *continuity*, *integrity*, *boundary*, *prototypicality*, etc., we can conjecture that topology will be not merely sufficiently general to encompass the subject-matter of the different cognitive sciences, but also that it will have the tools to do justice to this subject-matter without imposing alien features thereon.

5. Topology vs. Set Theory

Some of the reasons for shunning a set-theoretic treatment of the fundamentals of topology for present purposes can now be deduced as follows. Imagine that we are seeking a theory of the boundary-continuum structure as this makes itself manifest in the realm of everyday human experience, then the standard set-theoretic account of the continuum, initiated by Cantor and Dedekind and contained in all standard textbooks of the theory of sets, will be inadequate for at least the following reasons:

1. The experienced continuum does not sustain the sorts of cardinal number constructions imposed by the Dedekindian approach. The experienced continuum is not isomorphic to any real-number structure; indeed standard mathematical oppositions, such as that between a dense and a continuous series, here find no application.
2. The set-theoretical construction of the continuum is predicated on the highly questionable thesis that out of unextended building blocks an extended whole can somehow be constructed. (Brentano 1988, Asenjo 1993) The experienced continuum, in contrast, is organized not in such a way that it would be built up out of particles or atoms, but rather in such a way that the wholes, including the medium of space, come before the parts which these wholes might contain and which might be distinguished on various levels within them.
3. The application of set theory to a subject-matter presupposes the isolation of some basic level of *Urelemente* in such a way as to make possible a simulation of all structures appearing on higher levels by means of sets of successively higher types. If, however, as holds in the case of investigations of the ontology of the experienced world, we are dealing with mesoscopic entities and with their mesoscopic constituents

(the latter the products of more or less arbitrary real or imagined divisions along a variety of distinct axes), then there are no *Urelemente* to serve as our starting-point. (Cf. Bochman 1990.) This idea is, incidentally, at the heart also of Gestalt-theoretical criticisms of psychological atomism, which in many respects parallel criticisms of set-theory-induced atomism of the sort presented here.

Of course, set theory is a mathematical theory of tremendous power, and none of the above precludes the possibility of reconstructing topological theories adequate for cognitive-science purposes also in set-theoretic terms, as is standardly done. The reservations stated above however suggest that the set-theoretic framework can yield at best a *model* of the experienced continuum and similar structures, not a theory of these structures themselves (for the latter are after all *not sets*). Our suggestion, then, is that mereotopology will yield more interesting research hypotheses, and in a more direct and straightforward fashion, than would be the case should we choose to work instead with set-theoretic instruments.

6. Topological Psychology

The idea of using topology as a foundation for cognitive science is not without precursors. Of these, the most important, and the most notorious, is the work on "topological and vector psychology" of the German Gestalt psychologist Kurt Lewin. It will suffice for our present purposes merely to illustrate some of the ways in which Lewin uses topological notions in his 1936.

We begin, following Lewin, with the opposition *thing* (intuitively: a closed connected unity) and *region* (intuitively: a space within which things are free to move). As Lewin points out, what is a thing from one psychological perspective may be a region from another: 'A hut in the mountain has the character of a thing as long as one is trying to reach it from a distance. As soon as one goes in, it serves as a region in which one can move about.' (1936, p. 116) We then define the notion of a *boundary zone* z between two regions m and n , as the region, foreign to m and n , which has to be crossed in passing from one to the other. The whole $m + n + z$ is then *connected* in the topological sense. (1936, p. 121)

We can further define the concept of a *barrier* defined as a boundary zone which offers resistance to passage of things between one region and another. Such resistance may then be asymmetric; thus it may be greater in one dimension than in the opposite direction. Barriers effect the degree of communication between one region and another, or in other words the *degree of influence* of the state of one region on that of another region. Hence the notion of degree of influence, too, need not be symmetric: the fact that a is in a certain degree of communication with b does not imply that b is in equally close communication with a . These notions applied psychological—no need for real influence—similar ideas worked out in greater sophistication later by Talmy in the linguistic sphere in his work on force dynamics (1988).

Two regions a and b are said to be parts of a *connected region* if a change of state of a results in a change of state of b . The notion of *connectedness*, too, is by what was said

earlier a matter of degree. In fact we can distinguish a hierarchy of degrees of interlinkage between regions, and here Lewin echoes discussions in the Gestalt-theoretical literature of the notions of 'strong' and 'weak' Gestalten (1936, pp. 173f.). A strong Gestalt may be defined as a complex with a high degree of connectedness between its parts; a weak Gestalt has a lesser but still some minimal connectedness between its parts, while a purely summative whole (an '*Und-Verbindung*' in Gestaltist terminology) has a zero degree of connectedness between its separate units. Interestingly, the notions central to Gestalt theory can be defined not merely on the basis of the notion of connectedness, but also in terms of structure-preserving transformations (Simons 1988).

We have deliberately introduced the above notions in a rather general way, abstaining from any specific applications to psychological matters. The notions are, as Lewin himself sometimes recognizes, *formal* in the sense that they can be applied indiscriminately to a wide range of different sorts of material regions. The general tendency of Lewin's writings, however, is to switch too unthinkingly to psychological applications, whereby his use of the concepts in question (concepts of 'barrier', 'path', '(psychological) locomotion', 'dynamic interdependence (as the main determinant of the topology of the person)', 'tension', 'resistance', 'inhibition', etc. seems often to remain on the level of metaphor. While some cognitive scientists may be content never to pass beyond this level in their investigations, Lewin's methodology rightly drew the attention of critics who pointed to a certain crucial shortfall in his use of mathematical notions in his writings. As was correctly pointed out, Lewin rarely makes the mathematical theory of such notions do any substantial work within the framework of his investigations. This criticism was put forward in an influential article by I. London (1944), an article which did much to thwart the further development of topological psychology (or of a topologically founded cognitive science) as Lewin had conceived it. Yet London's critique is in some respects ill-drawn: London neglects the degree to which certain aspects of Lewin's generalizations of standard topology have since shown themselves to be highly fruitful. These generalizations include:

1. the recognition that it is possible to construct topology on a non-atomistic, mereological basis which works in terms of *wholes (regions)* as well as parts—where London, a defender of the analytic method at the heart of physical science, holds that for scientific purposes 'all experience must be dealt with in bits' (p. 279);
2. the systematic employment of the notion of asymmetric boundary, a notion which turns out to be crucial in many cognitive spheres;
3. the employment of topological ideas and methods also in relation to finite domains of objects. London argues (pp. 288f.) that topology makes sense only in infinite domains; as Hammer, Latecki, and others have shown, however, it is possible to construct in rigorous fashion finitistic systems in which analogues of topological notions can be defined;
4. the recognition that topology can serve as a theoretical basis for a unification of diverse types of psychological facts (Cf. Back 1992, p. 52).

Much in London's critique has in addition been rendered nugatory by more recent

developments such as those referred to above, which have demonstrated that it is possible to go beyond the merely metaphorical employment of topological concepts in cognitive science and to exploit the logico-mathematical properties of these concepts for the theoretical purposes of cognitive science in genuinely fruitful way. Thus many of Lewin's ideas recall principles of "force dynamics" worked out in greater sophistication in the linguistic sphere by Talmy (1988).

7. Husserl's Mereotopology

Prior to Lewin, there is the example of the philosopher Husserl, whose *Logical Investigations* (1900/01) contain a formal theory of part, whole and dependence that is used by Husserl to provide a framework for the analysis of mind and language of just the sort that is presupposed in the idea of a topological foundation for cognitive science. The title of the third of Husserl's *Logical Investigations* is "On the Theory of Wholes and Parts" and it divides into two chapters: "The Difference between Independent and Dependent Objects" and "Thoughts Towards a Theory of the Pure Forms of Wholes and Parts". Unlike more familiar theories of wholes and parts, Husserl's theory concerns itself not merely with what we might think of as the vertical relations between parts and the wholes which comprehend them, but also with the horizontal relations of dependence and fusion between the different parts within a single whole. The theory of such part-part relations he then uses as a basis for an account of unity or integrity in a manner which anticipates later such accounts developed in explicitly topological terms. To put the matter simply and crudely: some of a whole exist merely side by side, they can be destroyed or removed from the whole without detriment to the residue. A whole all of whose parts manifest exclusively such side-by-sideness relations with each other is called a heap or aggregate or, more technically, a purely summative whole (an *Und-Verbindung*). In many wholes, however, and one might say in *all* wholes manifesting any kind of unity, certain parts stand to each other in relations of what Husserl called *necessary dependence* (which is sometimes, but not always, necessary *interdependence*). Such parts, for example the surface and interior of a sphere, the individual instances of hue, saturation and brightness involved in a given instance of colour, cannot, as a matter of necessity, exist, except in association with their complementary parts in a whole of the given type. There is a huge variety of such lateral dependence relations giving rise to correspondingly huge variety of different types of whole which more standard approaches of 'extensional mereology' (Simons 1987, Ch. 1) are simply unable to distinguish.

The connection between part and whole on the one hand and dependence on the other may be seen already in the fact that every whole can be regarded as being dependent on its own constituent parts. This thesis may amount to no more than the trivial claim that every object is such that it cannot exist unless all the objects which are, at different times, its parts, also exist at those times. Or it may consist in a non-trivial thesis to the effect that certain special sorts of objects are such as to contain 'integral parts' which must exist at all the times when these objects exist: their loss (for example the loss of brain or heart in a mammal) is sufficient to bring about the destruction of the whole. Or, finally, it may be transformed into the metaphysical thesis of mereological essentialism, i.e. into the

assertion that *every* spatio-temporal object is dependent in the non-trivial sense upon *all* its parts, so that the ship ceases to exist (becomes another thing) with the removal of the first splinter of wood. It is one not inconsiderable advantage of Husserl's theory that it allows a precise formulation of these and a range of related theses within a single framework that is rooted on ideas concerning part, whole and dependence which are consistent with our common intuitions. Both Stanislaw Lesniewski, the founder of mereology, and the linguist Roman Jakobson applied Husserl's ideas on parts, wholes and categories from the *Logical Investigations* in different areas of linguistics, in the early development of categorial grammar and of phonology, respectively. Thus Jakobson's account of distinctive features is as he himself admits an application of Husserl's idea of dependent moments from the third Investigation.

The underlying topological character of Husserl's thinking on dependence has been demonstrated by Fine (1995). The topological background of Husserl's work makes itself felt also in his treatment of the notion of *phenomenal fusion* (Casati 1991). After distinguishing dependent and independent contents (for example, in the visual field, between a colour- or brightness-content on the one hand, and a content corresponding to the image of a moving projectile, on the other), Husserl goes on to note that there is in the field of intuitive data an additional distinction, between *intuitively separated* contents, contents *set in relief from* or *separated off from* adjoining contents, on the one hand, and contents which are *fused* with adjoining contents, or which *flow over* into them without separation, on the other (Investigation III, §8, 449).

He points out that independent contents which are what they are no matter what goes on in their neighbourhood, need not have this quite different independence of separateness. The parts of an intuitive surface of a uniform or continuously shaded white are independent, but not separated. (*Loc. cit.*) Such content Husserl calls 'fused'; they form an 'undifferentiated whole' in the sense that the moments of the one pass 'continuously' [*stetig*] into corresponding moments of the other. (§9, 450)

That Husserl was at least implicitly aware of the topological aspect of his ideas, even if not under this name, is unsurprising given that he was after all a student of the mathematician Weierstrass in Berlin, and that it was Cantor, Husserl's friend and colleague in Halle during the period when the *Logical Investigations* were being written, who first defined the fundamental topological notions of open, closed, dense, perfect set, boundary of a set, accumulation point, etc., and Husserl consciously employed Cantor's topological ideas, not least in writings on the general theory of (extensive and intensive) magnitudes which make up one preliminary stage on the road to the third Investigation (see Husserl 1983, pp. 83f, 95, 413, etc. and compare §§22 and 70 of the *Prolegomena to the Logical Investigations*).

More generally, it is worth pointing out that the early development of topology on the part of Cantor and others was part of a wider project on the part of both mathematicians and philosophers in the 19th century to produce a *general theory of space*—to find ways of constructing fruitful generalisations of such notions as extension, dimension, separation, neighbourhood, boundary, distance, proximity, continuity, boundary, etc.—and Husserl, with Stumpf and other students of Brentano such as

Meinong, participated in this project. (See Husserl 1983, pp. 275-300, 402-410; Stumpf 1873; Meinong 1903, esp. §2: "Farbengeometrie Und Farbenpsychologie", and §5: "Der Farbenraum Und seine Dimensionen".) Significantly, the 1906 paper on "The Origins of the Concept of Space", in which Riesz first formulated the closure axioms at the heart of topology is in fact a contribution to formal phenomenology, a study of the structures of *spatial presentations*, in which the attempt is made to specify the additional topological properties which must be possessed by a mathematical continuum if it is adequately to characterise the continuity and order properties of our experience of space.

References

- Adams, Ernest W. 1973 "The Naive Conception of the Topology of the Surface of a Body", in P. Suppes (ed.), *Space, Time and Geometry*, Dordrecht/Boston: Reidel, 402-424.
- Adams, Ernest W. 1984 "On the Superficial", *Pacific Philosophical Quarterly*, 65, 386-407.
- Adams, Ernest W. 1986 "On the Dimensionality of Surfaces, Solids, and Spaces", *Erkenntnis*, 24, 137-201.
- Asenjo, F. G. 1993 "Continua without Sets", *Logic and Logical Philosophy*, 1, 95-128.
- Allen J. F. 1981 "An Interval-Based Representation of Temporal Knowledge", *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, Vancouver: IJCAI (Morgan Kaufmann), Volume 1, pp. 221-26.
- Allen J. F. 1984 "Towards a General Theory of Action and Time", *Artificial Intelligence*, 23, 123-54.
- Aurnague, M. and Vieu, L. 1992 "A Three-Level Approach to the Semantics of Space", in *The Semantics of Prepositions. From Mental Processing to Natural Language Processing*, Berlin: Mouton/de Gruyter.
- Bach E. 1986 "The Algebra of Events", *Linguistics and Philosophy*, 9, 5-16.
- Back, Kurt W. 1992 "This Business of Topology", *Journal of Social Issues*, 48 (special issue on: The Heritage of Kurt Lewin: Theory, Research, and Practice), 51-66.
- Bochman, Alexander 1990 "Mereology as a Theory of Part-Whole", *Logique et Analyse*, 129-130, 75-101.
- Bozzi, P. 1958 "Analisi fenomenologica del moto pendolare armonico", *Rivista di Psicologia*, 52, 281-302.
- Bozzi, P. 1959 "Le condizioni del movimento 'naturale' lungo i piani inclinati", *Rivista di Psicologia*, 53, 337-352.
- Brandt, Per Aage 1989 "Agonistique et analyse dynamique catastrophiste du modal et de l'aspectuel: Quelques remarques sur la linguistique cognitive de L. Talmy", *Semiotica*, 77, 151-162.
- Brentano, Franz 1988 *Philosophical Investigations on Space, Time and the Continuum*, translated by Barry Smith, London/New York/Sydney: Croom Helm.

- Brugman, Claudia and Lakoff, George 1988 "Cognitive Topology and Lexical Networks" in S. L. Small *et al.* (eds.), *Lexical Ambiguity Resolution: Perspectives from Psycholinguistics, Neuropsychology, and Artificial Intelligence*, San Mateo: Kaufmann,
- Casati, Roberto 1991 "Fusion", in H. Burkhardt and B. Smith (eds.), *Handbook of Metaphysics and Ontology*, Munich/Hamden/Vienna: Philosophia, 287–289.
- Casati, Roberto and Varzi, Achille 1993 "An Ontology for Superficial Entities I: Holes". In: N. Guarino & R. Poli (eds.), *Proceedings of the International Workshop on Formal Ontology in Conceptual Analysis and Knowledge Representation*, Padova: CNR-Ladseb, 127-148.
- Casati, Roberto and Varzi, Achille 1994 *Holes and Other Superficialities*, Cambridge, Mass. and London: MIT Press.
- Chisholm, Roderick M. 1989 "Boundaries", in Chisholm 1989a, 83–90.
- Chisholm, Roderick M. 1989a *On Metaphysics*, Minneapolis: University of Minnesota Press.
- Chisholm, Roderick M. 1992/93 "Spatial Continuity and the Theory of Part and Whole. A Brentano Study", *Brentano Studien*, 4, 11–23.
- Cresswell, M. J. 1978 "Adverbs of Space and Time", in F. Guenther, S. J. Schmidt, eds., *Formal Semantics and Pragmatics for Natural Languages*, Dordrecht: Reidel, pp. 171-99; reprinted in Cresswell 1985, pp. 41-66.
- Cresswell, M. J. 1979 "Interval Semantics for Some Event Expressions", in R. Bäuerle, U. Egli, A. n echow, eds., *Semantics from Different Points of View*, Berlin and Heidelberg: Springer-Verlag, pp. 90-116; reprint in Cresswell 1985, pp. 143-71.
- Cresswell, M. J. 1985 *Adverbial Modification: Interval Semantics and Its Rivals*, Dordrecht: Reidel.
- Cui, Z., Cohn, A. G. and Randell, D. A. 1992 "Qualitative Simulation Based on a Logical Formalism of Space and Time", in: *Proceedings of the AAAI-92*, Menlo Park, CA: AAAI Press, 679-684.
- Desclès, Jean-P. 1989 "State, Event, Process, and Topology", *General Linguistics*, 29, 159–200.
- Eschenbach, Carola and Wolfgang Heydrich (in press) "Classical mereology and restricted domains" in: N. Guarino & R. Poli (eds.), *Formal Ontology in Conceptual Analysis and Knowledge Representation*, Dordrecht: Kluwer.
- Fine, K. 1995 "Husserl's Theory of Part-whole", in B. Smith and D.W. Smith (eds.), *The Cambridge Companion to Husserl*, Cambridge: Cambridge University Press.
- Forbus, K. 1984 "Qualitative Process Theory", *Artificial Intelligence*, 24, 85–168.
- Frank, A. U., Campari, I. and Formentini, U. (eds.) 1992 *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space. Proceedings of the International Conference GIS - From Space to Territory: Theories and Methods of Spatio-Temporal Reasoning*, Pisa, Italy, September 1992.

- Frank, A. U. and I. Campari (eds.) 1993 *Spatial Information Theory. A Theoretical Basis for GIS. Proceedings of the European Conference COSIT'93, Marciana Marina, Elba, Italy, September 1993*, Berlin: Springer.
- Frawley, William 1980 "The Topology of Texts", *Papers in Linguistics: International Journal of Human Communication*, 13, 721-738.
- Freksa, C. 1991 "Conceptual Neighborhood and its Role in Temporal and Spatial Reasoning", in M. Singh & L. Trave-Massuyes (eds.) *Decision Support Systems and Qualitative Reasoning*. Amsterdam: North-Holland, 181-187.
- Freksa, C. 1992 "Temporal Reasoning Based on Semi-intervals", *Artificial Intelligence*, 54, 199-227.
- Freksa, Christian and Christopher Habel (eds.) 1990 *Repräsentation und Verarbeitung räumlichen Wissens*, Berlin: Springer.
- Freksa, C. and Röhrig, R. 1993 "Dimensions of Qualitative Spatial Reasoning", in N. Piera Carrete and M. G. Singh (eds.), *Qualitative Reasoning and Decision Technologies, Proc. QUARDET'93*, Barcelona: IMNE.
- Galton, Antony 1984 *The Logic of Aspect. An Axiomatic Approach*, Oxford: Clarendon Press.
- Gibson, J. J. 1986 *The Ecological Approach to Visual Perception*, Hillsdale, NJ: Lawrence Erlbaum.
- Habel, Christopher 1987 "Cognitive Linguistics: The Processing of Spatial Concepts", in *T.A. Informations - ATALA (Association pour le traitement automatique des langues)*, 28, 21-56.
- Habel, Christopher 1989 "zwischen-Bericht", in C. Habel, M. Herweg and K. Rehkämper (eds.), *Raumkonzepte in Verstehensprozessen*, Tübingen: Niemeyer, 37-69.
- Habel, Christopher 1990 "Propositional and Depictorial Representations of Spatial Knowledge: the Case of Path Concepts" in R. Studer (eds.), *Natural Language and Logic. Lecture Notes in Artificial Intelligence*, Berlin: Springer, 94-117.
- Habel, Christopher 1993 "Representing Space and Time: Discrete, Dense or Continuous?—Is that the Question?", Zentrum für Interdisziplinäre Forschung, Universität Bielefeld, Forschergruppe Mental Models in Discourse Processing, *Bericht 6/93*.
- Habel, Christopher and Simone Pribbenow 1988 "Gebietskonstituierende Prozesse", Lilog-Report Nr. 18, Stuttgart: IBM.
- Habel, Christopher, Herweg, Michael and Eschenbach, Carola 1988 "Vorüberlegungen zu einer Theorie der Gruppierungs- und Abgrenzungsprozesse beim Aufbau sprachlich angeregter mentaler Modelle", *Mitteilung Nr. FBI-HH-M-160/88*, Universität Hamburg, Fachbereich Informatik. (GAP-Arbeitspapier 1.)
- Habel, Christopher, Herweg, Michael and Rehkämper, Klaus (eds.) 1989 *Raumkonzepte in Verstehensprozessen*, Tübingen: Niemeyer.

- Habel, Christopher, Eschenbach, Carola, Fraurud, Kari, Herweg, Michael, Maienborn, Claudia, Rehkämper, Klaus 1993 *Intern strukturierte Objekte: Zur Verarbeitung von Pluralen und komplexen Situationen. Abschlußbericht des DFG-Projektes "Gruppierungs- und Abgrenzungsprozesse beim Aufbau sprachlich angeregter mentaler Modelle" (GAP)*, Universität Hamburg.
- Habel, Christopher, Simone Pribbenow and Geoffrey Simmons 1993 "Propositional and Depictorial Partonomies: A Hybrid Approach", Zentrum für Interdisziplinäre Forschung, Universität Bielefeld, Forschergruppe Mental Models in Discourse Processing, *Bericht 1/93*.
- Hager, Greg 1985 "Naive Physics of Materials: A Recon Mission", in *Commonsense Summer: Final Report*, Report No. CSLI-85-35, Center for the Study of Language and Information, Stanford University, Stanford, California.
- Hammer, P. C. 1960 "Kuratowski's Closure Theorem", *Nieuw Archief voor Wiskunde*, 8, 74-80.
- Hammer, P. C. 1961 "Extended Topology: Reduction of Limit Functions", *Nieuw Archief voor Wiskunde*, 9, 16-24.
- Hammer, P. C. 1961a "Extended Topology: The Wallace Functions of a Separation", *Nieuw Archief voor Wiskunde*, 9, 74-86.
- Hammer, P. C. 1962 "Extended Topology: Set-Valued Set-Functions", *Nieuw Archief voor Wiskunde*, 10, 55-77.
- Hammer, P. C. 1962a "Extended Topology: Additive and Sub-Additive Subfunctions", *Rendiconto del rcolo Matematico di Palermo*, Ser. 2, 11, 262-70.
- Hammer, P. C. 1964 "Extended Topology: Perfect Sets", *Portugaliae Mathematica*, 25, 27-34.
- Hammer, P. C. 1964a "Extended Topology: Structure of Isotonic Functions", *Journal für die reine und angewandte Mathematik*, 213, 174-86.
- Hammer, P. C. 1967 "Language, Approximation, and Extended Topologies", in I. Rauch and C. T. Scott (eds.), *Approaches in Linguistic Methodology*, Madison: University of Wisconsin Press, 33-40.
- Hammer, P. C. 1969 "Continuity", in P. C. Hammer (ed.), *Advances in Mathematical Systems Theory*, University Park and London: The Pennsylvania State University Press, 93-105.
- Hayes, Patrick J. 1985 "The Second Naive Physics Manifesto", in Hobbs, J. R. and Moore, R. C. eds. *Formal Theories of the Common-sense World*, Norwood: Ablex, 1-36.
- Herring, John R. 1991 "The Mathematical Modeling of Spatial and Non-Spatial Information in Geographic Information Systems", in Mark and Frank, eds., 313-350.
- Heydrich, Wolfgang 1982 *Gegenstand und Sachverhalt. Bausteine zu einer nominalistisch orientierten Semantik für Texte*, Hamburg: Buske.

- Heydrich, Wolfgang 1988 "Things in Space and Time", in J. S. Petöfi (ed.), *Text and Discourse Constitution. Empirical Aspects, Theoretical Approaches*, Berlin/New York: De Gruyter, 377-418.
- Heydrich, Wolfgang 1989 "Relevant Objects and Situations", in M. E. Conte, J. S. Petöfi and E. Soezer (eds.), *Text and Discourse Connectedness. Proceedings of the Conference "Text, Connexity, Coherence", Urbino, July, 1984*, Amsterdam: Benjamins, 413-25.
- Hobbs, J. R. 1985 "Granularity", *Proceedings of the 9th IJCAI*, 432-435.
- Hoeksema, J. 1985 *Categorical Morphology*, New York/London: Garland Publishing.
- Husserl, Edmund 1900/01 *Logische Untersuchungen*, 1st ed., Halle: Niemeyer, 2nd ed., 1913/21 (both now available in a comparative edition as *Husserliana XVIII-XIX*, The Hague: Nijhoff, 1975, 1984). Page references unless otherwise indicated are to Findlay's English translation of the 2nd ed. (*Logical Investigations*, London: Routledge, 1970), translation amended where necessary.
- Husserl, Edmund 1983 *Studien zur Arithmetik und Geometrie. Texte aus dem Nachlass, 1886-1901*, The Hague: Nijhoff (*Husserliana XXI*).
- Jackendoff, Ray 1991 "Parts and Boundaries", *Cognition*, 41, 9-45.
- Kamp, Hans 1979 "Events, Instants, and Temporal Reference", in R. Bäuerle, U. Egli and A. von Stechow (eds.), *Semantics from Different Points of View*, Berlin/Heidelberg: Springer-Verlag, pp. 376-417.
- Kaufmann, Ingrid 1991 "Semantik der wegbezogenen Präpositionen des Deutschen", *Arbeiten des Sonderforschungsbereichs 282: Theorie des Lexikons*, Nr. 8, Universität Düsseldorf.
- Kindt, Walther 1983 "Two Approaches to Vagueness: Theory of Interaction and Topology", in T. Ballmer and M. Pinkal, (eds.), *Approaching Vagueness*, Amsterdam: North-Holland, 361-392.
- Kuratowski, Kazimierz 1922 "Sur l'opération A⁻ d'analysis situs", *Fundamenta Mathematica*, 3, 182-99.
- Lakoff, George 1989 "Some Empirical Results about the Nature of Concepts", *Mind and Language*, 4, 103-129.
- Langacker, Ronald W. 1987 *Foundations of Cognitive Grammar*, 2 vols., Stanford: Stanford University Press.
- Langacker, Ronald W. 1993 "Clause Structure in Cognitive Grammar", *Studi Italiani di Linguistica Teorica e Applicata*, 22, 465-508.
- Latecki, Longin 1992 *Digitale und allgemeine Topologie in der bildhaften Wissensrepräsentation*, St. Augustin: infix.
- Lewin, Kurt 1936 *Principles of Topological Psychology*, New York and London: McGraw-Hill.
- Lipski, John M. 1974 "A Topology of Semantics Dependence", *Semiotica*, 12, 145-170.
- Lipski, John M. 1975 "Towards a Topology of Natural Languages", *Poetics*:

- International Review for the Theory of Literature*, 13, 5-17.
- London, Ivan D. "Psychologists' Misuse of the Auxiliary Concepts of Physics and Mathematics", *Psychological Review*, 51, 266-91.
- McCloskey, M. 1983 "Intuitive Physics", *Scientific American*, 248(4), 122-130.
- McKinsey, J. C. C. and Tarski, A. 1944 "The Algebra of Topology", *Annals of Mathematics*, 45, 141-91.
- Mark, David M. and Andrew U. Frank (eds.) 1991 *Cognitive and Linguistic Aspects of Geographic Space*, Dordrecht/Boston/London: Kluwer.
- Meinong, Alexius von 1903 "Bermerkungen über die Farbenkörper und das Mischungsgesetz", *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 33, 1-80, and in Meinong, *Gesamtausgabe*, vol. I, 497- 576.
- Mostowski, Marcin 1983 "Similarities and Topology", *Studies in Logic, Grammar and Rhetoric*, 3, 106-119.
- Mourelatos, A. P. D. 1981 "Events, Processes and States", in P. J. Tedeschi and A. Zaenen, eds., *Tense and Aspect (Syntax and Semantics, Vol. 14)*, New York: Academic Press, 191-211.
- Netzer, N. 1977 "Konvergenz in verallgemeinerten topologischen Räumen", *Mathematische Nachrichten*, 77, 245-62.
- Netzer, N. 1978 "Verallgemeinerte topologische Strukturen", *Jahrbuch Überblicke Mathematik*, 1978, 87-106.
- Nöbeling, G. 1954 *Grundlagen der analytischen Topologie*, Berlin: Springer.
- Ore, O. 1943 "Some Studies on Closure Relations", *Duke Mathematical Journal*, 10, 761-85.
- Ore, O. 1943a "Combinations of Closure Relations", *Annals of Mathematics*, 44, 514-33.
- Ore, O. 1946 "Mappings of Closure Relations", *Annals of Mathematics*, 47, 56-72.
- Perzanowski, Jerzy 1993 "Locative Ontology", *Logic and Logical Philosophy*, 1, 7-94
- Petitot, Jean 1982 "Sur la signification linguistique de la Théorie des Catastrophes", *Mathématique et Sciences Humaines*, 79, 37-74.
- Petitot, Jean 1989 "Morphodynamics and the Categorical Perception of Phonological Units", *Theoretical Linguistics*, 15, 1989, 25-71.
- Petitot, Jean 1989a "Hypothèse localiste, Modèles morphodynamiques et Théories cognitives", *Semiotica*, 77, 65-119.
- Petitot, Jean 1989b "On the Linguistic Import of Catastrophe Theory", *Semiotica*, 74, 179-209.
- Petitot, Jean 1990 "Le Physique, le Morphologique, le Symbolique. Remarques sur la Vision", *Revue de Synthèse*, 4, 139-183.
- Petitot, Jean 1991 "Morphodynamical Models and Topological Syntax", *Interdisciplinary Workshop on Compositionality in Cognition and Neural Networks*, Paris: CREA.

- Petitot, Jean 1991a "Syntaxe topologique et grammaire cognitive", *Langages*, 103, 97-128.
- Petitot, Jean 1992 "Natural Dynamical Models for Natural Cognitive Grammar", *International Journal of Communication*, 2, 81-104.
- Petitot, Jean 1992a *Physique du Sens*, Paris: Editions du CNRS.
- Petitot, Jean 1994 "Phenomenology of Perception, Qualitative Physics and Sheaf Mereology", in R. Casati, B. Smith and G. White (eds.), *Philosophy and the Cognitive Sciences*, Vienna: Hölder-Pichler-Tempsky, 387-408.
- Petitot, Jean and Smith, Barry 1990 "New Foundations for Qualitative Physics", in J. E. Tiles, G. T. McKee and C. G. Dean, eds., *Evolving Knowledge in Natural Science and Artificial Intelligence*, London: Pitman Publishing, 231-249.
- Petitot, Jean and Smith, Barry (in press) "Physics and the Phenomenal World", in R. Poli and P.M. Simons, eds., *Formal Ontology*, Dordrecht/Boston/Lancaster: Kluwer.
- Pianesi, F. and Varzi, A. C. 1994 "the Mereology of Event Structures", in P. Dekker and M. Stokhof, eds., *Proceedings of the 9th Amsterdam Colloquium*, Amsterdam: Institute for Language, Logic and Computation, 527-546. Extended version in this volume.
- Pogonowski, Jerzy 1979 "Topology and Semantics", *Lingua Posnaniensis*, 22, 75-85.
- Pribbenow, Simone 1993 *Räumliche Konzepte in Wissens- und Sprachverarbeitung - Hybride Verarbeitung von Lokalisierung*, Wiesbaden: Deutscher Universitätsverlag.
- Randell, David A. and Cohn, Anthony G. 1989 "Modelling Topological and Metrical Properties in Physical Processes", in R. J. Brachman, H. Levesque and R. Reiter (eds.), *Principles of Knowledge Representation and Reasoning*, Los Altos: Morgan Kaufmann.
- Randell, David A., Cui, Zhan and Cohn, Anthony G. 1992 "a Spatial Logic Based on Regions and Connection", in *Proceedings of the 3rd International Conference on Knowledge Representation and Reasoning*, Boston: Morgan Kaufmann.
- Rehkämper, Klaus 1991 *Sind mentale 'Bilder' bildhaft? - Eine Frage zwischen Philosophie und Wissenschaft*, Thesis, Hamburg.
- Riesz, F. 1906 "Die Genesis des Raumbegriffs", *Mathematische und naturwissenschaftliche Berichte aus Ungarn*, 24, 309-53.
- Rubin, E. 1921 *Visuell wahrgenommene Figuren. Studien in psychologischer Analyse*, Copenhagen: Gyldenalske.
- Riesz, F. 1909 "Stetigkeitsbegriff und abstrakte Mengenlehre", *Atti del IV Cong. Int. dei Mat.*, II, Rome, 18-24.
- Saloni, Zygmunt and Trybulec, Andrzej 1974 "Coherence of a Text and its Topology", *Semiotica*, 12
- Simons, Peter M. 1987 *Parts. A Study in Ontology*, Oxford: Clarendon Press.
- Simons, Peter M. 1988 "Gestalt and Functional Dependence", in Smith (ed.), 158-190.

- Smith, Barry 1987 "The Substance of Brentano's Ontology", *Topoi*, 6, 39-49.
- Smith, Barry 1992 "Characteristica Universalis", in K. Mulligan, ed., *Language, Truth and Ontology* (Philosophical Studies Series), Dordrecht/Boston/London: Kluwer, 50-81.
- Smith, Barry 1992a "Zum Wesen des Common Sense: Aristoteles und die naive Physik", *Zeitschrift für philosophische Forschung*, 46, 508-525.
- Smith, Barry 1993 "Ontology and the logistic analysis of reality", in N. Guarino and R. Poli (eds.), *Proceedings of the International Workshop on Formal Ontology in Conceptual Analysis and Knowledge Representation*, Institute for Systems Theory and Biomedical Engineering of the Italian National Research Council, Padova, Italy, 51-68. Revised version forthcoming in: G. Haefliger and P. M. Simons (eds.), *Analytic Phenomenology*, Dordrecht/Boston/London: Kluwer.
- Smith, Barry (in preparation) "Boundaries", in L. Hahn (ed.), *Roderick Chisholm* (The Library of Living Philosophers Series), Chicago and La Salle: Open Court.
- Smith, Barry (ed.) 1982 *Parts and Moments. Studies in Logic and Formal Ontology*, Munich: Philosophia.
- Smith, Barry (ed.) 1988 *Foundations of Gestalt Theory*, Munich and Vienna: Philosophia.
- Smith, Barry and Casati, Roberto 1994 "Naive physics", *Philosophical Psychology*, 7/2, 225-244.
- Smith, Barry and Mulligan, Kevin 1982 "Pieces of a Theory", in Smith, ed., 15-109.
- Smith, Barry and Mulligan, Kevin 1983 "Framework for Formal Ontology", *Topoi*, 3, 73-85.
- Smolensky, Paul 1988 "On the Proper Treatment of Connectionism", *Behavioral and Brain Sciences*, 11, 1-23.
- Stroll, Avrum 1988 *Surfaces*, Minneapolis: The University of Minnesota Press.
- Stumpf, C. 1873 *Über den psychologischen Ursprung der Raumvorstellung*, Leipzig: Hirzel.
- Talmy, Leonard 1977 "Rubber-Sheet Cognition in Language." In: *Papers from the 13th Regional Meeting, Chicago Linguistic Society*. W. Beach, et al. (eds.), University of Chicago.
- Talmy, Leonard 1978 "Figure and Ground in Complex Sentences", in: *Universals of Human Language: Syntax* (vol. 4). Joseph Greenberg (ed.), Stanford: Stanford University Press.
- Talmy, Leonard 1983 "How Language Structures Space", in H. Pick and L. Acredolo, eds., *Spatial Orientation: Theory, Research, and Application*, New York: Plenum Press, 225-82.
- Talmy, Leonard 1988 "Force Dynamics in Language and Cognition", *Cognitive Science*, 12, 49-100.
- Talmy, Leonard 1988a "The Relation of Grammar to Cognition", *Topics in Cognitive*

- Linguistics*, Brygida Rudzka-Ostyn (ed.), Amsterdam: John Benjamins, 165-205.
- Tarski, Alfred 1938 "Die Aussagenkalkül und die Topologie", *Fundamenta Mathematica*, 31, 103-34; Eng. trans., "Sentential Calculus and Topology", in *Logic, Semantics, Metamathematics*, Oxford: Clarendon Press, 1956, 421-454.
- Thom, René 1990 *Semio Physics: A Sketch*, Redwood City: Addison-Wesley.
- Tiles, J. E. 1981 *Things That Happen*, Aberdeen: Aberdeen University Press.
- van Benthem J. 1983 *The Logic of Time*, Dordrecht: Kluwer (2nd ed., 1991).
- Varzi, Achille C. 1993 "Spatial Reasoning in a Holey World", in P. Torasso (ed.), *Advances in Artificial Intelligence. Proceedings of the 3rd Congress of the Italian Association for Artificial Intelligence [Lecture Notes in Artificial Intelligence 728]*, Berlin et al.: Springer, 326-336.
- Varzi, Achille C. 1994 "On the boundary between mereology and topology", in R. Casati, B. Smith and G. White (eds.), *Philosophy and Cognitive Science. Proceedings of the 16th International Wittgenstein Symposium*, Vienna: Hölder-Pichler-Tempsky, 419-438.
- Wildgen, W. 1981 "Semantic Description in the Framework of Catastrophe Theory" in B. Rieger, ed., *Empirical Semantics. A Collection of New Approaches in the Field*, 2 vols., Bochum: Brockmeyer, 792-818.
- Wildgen, W. 1982 *Catastrophe Theoretic Semantics. An Elaboration and Application of René Thom's Theory*, Amsterdam: John Benjamins.
- Wildgen, Wolfgang 1994 *Process, Image, and Meaning: A Realistic Model of the Meaning of Sentences and Narrative Texts*, Amsterdam/Philadelphia: J. Benjamins.
- Wunderlich, Dieter and Michael Herweg 1991 "Lokale und Direktionale", in A. v. Stechow and D. Wunderlich, eds., *Semantik*, Berlin, New York: de Gruyter, 758-785.
- Zarycki, Miron 1927 "Quelques notions fondamentales de l'Analysis Situs aux point du vue de l'Algèbre de la Logique", *Fundamenta Mathematica*, 9, 3-15.
- Zeeman, E.C. 1962 "Topology of the Brain and Visual Perception", in M.K. Fort (ed.), *Topology of 3-manifolds*, New York: Prentice Hall, 57-64.
- Zimmer, A. C. and Koerndle, H. 1994 "A Gestalt Theoretic Account for the Coordination of Perception and Action in Motor Learning", *Philosophical Psychology*, 7, 24 9-265.
- Zimmermann, K. and Freksa, C. 1993 "Enhancing Spatial Reasoning by the Concept of Motion", in A. Sloman, D. Hogg, G. Humphreys, A. Ramsay, D. Partridge (eds.), *Prospects for Artificial Intelligence, Proceedings of AISB93, University of Birmingham*, Amsterdam: IOS-Press.

The Bounds of Axiomatisation

Graham White

International Academy for Philosophy

Obergass 75

FL-9494 Schaan

Liechtenstein

ggw10@phx.cam.ac.uk

This essay is an attempt at philosophical *Kulturkritik* – that is, at the critique of *philosophical* culture; the essay itself is only in a rather modest sense philosophical. The problem addressed is this. There is a certain genre of philosophical literature which consists in the construction or description of axiom systems. Here, it seems to be presumed, is a suitable role for philosophy: the indication of first principles, leaving the detailed work for other more specialised investigators. Is there anything to this?

The picture seems, at first sight, attractive. Axioms might be taken to be the basic elements out of which mathematical thought is constructed: the traditional role of philosophy is the investigation of just such basic elements: so, we could think, the philosophers might at last find a practical application for their work as they happily produce axiom systems to be used by the artificial intelligence community. Philosophers could also work critically, showing that certain areas of thought were based on the wrong axioms, and suggesting more appropriate foundations for them. However, there are several questionable elements in the above argument, namely:

Axioms as Basic Elements The idea that axioms are basic elements can be interpreted in several ways. This might be supposed to be a description of mathematical practice – i.e. one might suppose that mathematicians spend their time setting out axioms and drawing consequences from them. Or one might suppose that it is a description of mathematical ontology: that is, that axioms in some way refer to, or otherwise give one access to, items which are somehow ontologically basic. Or one might take a more normative stance, and say that axioms provide a foundation for mathematics, so that, regardless of actual mathematical practice, it would be important that mathematics could in principle be reconstructed axiomatically. All of these interpretations are to some extent questionable, and furthermore they rarely agree with one another. But more of this later.

The Traditional Role of Philosophy We are hardly on safer ground here. Conceptions of philosophy are (of course) historically variable. In particular, they have oscillated between the idea of philosophy as prior to other knowledge and the idea of philosophy as reflecting on and synthesising the results of the other sciences. Only the first would be consistent with the picture of axioms which we have sketched above. However, even if one attempts to enlist the aid of those philosophers who have conceived as philosophy as prior, one should be rather sobered by the following group of circumstances. Firstly, most of the philosophers who have done this have adhered to the ideal of demonstrative science set out in Aristotle's *Second Analytics*. Secondly, most of these philosophers have also thought of mathematics as a paradigm science. Thirdly, Aristotle's ideal of demonstrative science cannot possibly be an adequate description of mathematical practice, because it uses syllogisms and syllogisms are, after all, a decidable theory. Fourthly, hardly anyone seems to have noticed. Now one might argue that this does not matter, because Aristotle's theory of demonstrative science was perfectly good for what the philosophers wanted to do with it, that is, to use it for metatheoretical reflection. This may well be the case, but if so it is hardly an advertisement for philosophy as a useful art: in retrospect, it seems fortunate that scientific and mathematical practitioners have, by and large, ignored what philosophers have to say, and have thus been saved from the technical deficiencies of the philosophers' picture of science.

From the Consumers' Point of View The third element in the above argument is that axiom systems are useful: that is, that artificial intelligence specialists use them, so that after all there is some need for them. This might give us a pragmatic argument for the philosophical interest in axiomatisation, analogous to a famous criminal's reply to the question of why he robbed banks. ("That's where the money is.") Even though this may be true (in the sense that some philosophers have been paid money for doing this), it may distract attention from what is philosophically interesting. Neither may it be so interesting for the consumers. Axioms are, after all, only one of many possible formal descriptions of reality, or of mathematical practice, and others may be – and, I would argue, are – more interesting.

My position is as follows. I would argue in general that a group of techniques derived from category theory are more suitable for the purposes of both philosophy and of computer science than are axiomatic systems. From the point of view of computer science, one does not need much argument: one can simply point to the substantial body of work already achieved.¹ And from the point of view of philosophy, I would argue that these techniques give one an illuminating picture of mathematical practice, as well as a plausible model for computation, and that furthermore they are appropriate for the formalisation of a great deal of "naive physics" and other forms of common sense knowledge.²

So why axioms? That is, why do they have the prominence that they do? Probably the answer is simply conservatism; the idea of an axiomatic system was used for foundational purposes by Hilbert and others of that period, and a great deal of the

¹[2] is a good place to start.

²See [5].

philosophical discussion even now still refers to that group of ideas. The ideas used in such cross-disciplinary enterprises have an extraordinary permanence, because the basic translations – that *this* mathematical idea stands for *that* philosophical idea – are the way that both parties deal with the other subject. Philosophers, for example, may not be too familiar with such and such a mathematical concept, but if they know that, in philosophical terms, it “means” so and so, then they can deal with it. So these basic dictionaries between philosophy and mathematics are extremely difficult to change; it is even very difficult to try to examine them rationally.

1 Building Blocks of Mathematics

The idea of foundations of mathematics is almost entirely misleading. The late nineteenth century interest in foundations – the work of Kronecker and Frege and others – was motivated by a great deal of real disagreement among mathematicians about quite concrete matters: the validity or otherwise of mathematical proofs. This disagreement was not confined to philosophers’ favourite examples, such as the paradoxes of the infinite, but also extended to whole areas, in particular to the algebraic geometry of Severi and his school. Now the answer to a concrete crisis of this sort can also be very concrete: one can simply find a way in which mathematics can be done rigorously. The answer can even be very practical; one can learn how to do mathematics better, without the help of very much philosophy.

Similarly, Hilbert’s programme was a very concrete answer to a certain problem – the role of the infinite in mathematical thought – and the programme attempted to answer this by, firstly, providing a mathematical analysis of the concept of mathematical proof, and secondly by attempting to eliminate reference to the infinite. Now although all of this work – both the foundational work, Hilbert’s programme, and Gödel’s critique of it – is important, it should not be taken to say more than it does. It is standardly taken to be a description of mathematical practice, and it is not that; it is also standardly taken to be a description (or at least an attempted description) of ontologically basic mathematical entities, and it is not that.

To start with the question of basic entities. This is something that philosophers can hardly restrain themselves for looking for (it fits, after all, the philosophers’ favourite role of ontological customs officers, sniffing around the baggage of more normal people in search of ontological contraband, otherwise known as “queer entities”). But, even though the progressive reductions of the turn of the century foundations of mathematics certainly look like a decomposition of entities into more basic constituents, I think this appearance is deceptive. For one thing, the identities which these reductions establish can hardly be taken literally, as Benacerraf points out.³ Secondly, the direction in which reduction is taken to go often seems quite arbitrary: in mathematics, one can frequently reduce entities of type *A* to those of type *B*, but also vice versa. Such reductions may be convenient for mathematical purposes, but it is hard to argue that they show that one class of entities is ontologically more basic than the other. Here, is *A* more basic than *B*, or *vice versa*? In the

³[1]

case of physical reduction, one should remember, one has the idea of physical constituency, which indicate which direction reduction should go in, and force it to have an unambiguous direction. Concepts like constituency are lacking in mathematics. Many analogous concepts, which give a sense of direction to physical explanation, are also problematic in the mathematical case. Thus, in many physical cases, one has a classification of objects into sorts, and the idea of each sort is simpler than the idea of the members of that sort. (The idea of a mammal is simpler than the idea of a tapir.) However, this is not always the case in mathematics; categories are most helpfully regarded as the inhabitants of a certain two-category (the two-category of all categories), and the idea of a two-category is more complicated than that of a category. So the very idea of ontologically basic entities seems to be difficult to apply to mathematical ontology, without at least a good deal of analytical work on what it is to be a basic entity.

The critical investigation of axioms is also very questionable. What one would have to establish is that certain results are wrong because they are based on axioms, or foundational elements, of the wrong sort (set theory rather than mereology, for example). But this is, in the nature of things, hard to argue for, because most foundational setups for mathematics are intertranslatable, so it is a little difficult to see what harm ensues by using one rather than the other if the results are the same.

1.1 Mathematical Practice

I shall base a good deal of the positive argument of this paper on the concept of mathematical practice. This is important, for a variety of reasons. Firstly there is the *sine qua non* argument; whatever mathematical ontology might be, we have no other access to it than by way of mathematical practice, so that mathematical practice is implicitly involved in our philosophy of mathematics anyway. Now arguments like this are not altogether reliable – one can give a similar argument for the reduction of philosophy to epistemology, for example. However, when they fail, they generally fail because the domain they are applied to (epistemology, let us say) is more difficult to investigate, or more contentious, than is some other domain (let us say ontology). However, this is not the case for mathematical practice; its formal investigation is a flourishing branch of mathematics, namely category theory.

Secondly, the investigation of mathematical practice could well demystify concepts such as that of mathematical intuition, by showing how such concepts are linked to certain features of practice. (This would be similar to Husserl's introduction of eidetic intuition in connection with a practice, namely the practice of free imaginative variation.)

Thirdly, the consideration of mathematical practice might have a chance of bringing the philosophy of mathematics back into contact with mathematics and end the continuous recirculation among philosophers of the same tiny stock of mathematical examples; as McCarty puts it, "the presumptive foundations of mathematics turn out, at the hands of the philosophers, to issue in terms which would be utterly insufficient to explain in detail the manufacture of a paper clip."⁴ (Or, as a graduate

⁴[3, p. 326]

student at a prominent university said to me, "The philosophy of mathematics has got really interesting; you hardly have to know any mathematics at all to do it.")

Let us look at an example; the concept of dimension. This is a concept which is governed by one or two basic intuitions – that the dimension of a certain space is somehow a measure of the number of independent free variations that are possible within it, that the dimension of a subspace should be less than or equal to that of the containing space, that the dimension of a space should be an intrinsic property, and that the dimension of a space should be related to its behaviour under scaling. These intuitions are not laid down precisely, but seem, in mathematical practice, to be viewed as rough desiderata which a candidate notion of dimension ought to satisfy. This is significant, because there hardly seems to be any one concept of dimension which is clearly the right one.⁵ Thus, one common definition – that of Hausdorff dimension – is only definable for metric spaces, and is not as intrinsic as maybe it should be, since it is only preserved by maps of metric spaces which satisfy a Hölder condition. Furthermore, the counterexamples which one finds in this area do not seem to be merely pathological; many of them exhibit geometrical behaviour which is interesting in its own right (they are fractals, for example).

As an example, consider Baire space; it can be defined as the space of functions from the natural numbers to the natural numbers, and given a topology in which the basic opens are the subsets consisting of those functions which have given fixed values for the first k natural numbers, where k is arbitrary but fixed for each open set. Then it is easy to see that Baire space is isomorphic to the product of two copies of itself (one simply interleaves pairs of functions). Because the dimension of a product is the sum of the dimensions of its components (this is because dimension should measure the number of independent free variations), the only plausible dimensions to give to Baire space are zero and infinity. Both are problematic (infinity particularly so, since one can embed Baire space in the real line). But Baire space is not by any means simply pathological; it is used extensively in descriptive set theory and measure theory, and is an interesting object of study in its own right.

Furthermore, let us consider the family of concepts of dimension which are given by what are cohomology theories. Such theories are basically ways of systematically comparing the local and the global behaviour of topological spaces; the concept of a cohomology theory is very well defined (there are axioms called the Eilenberg-Steenrod axioms⁶ which such theories must satisfy). Each cohomology theory is appropriate for spaces of a certain sort, but here 'sort' is not defined axiomatically but by saying that such spaces are locally isomorphic to a given model: thus, manifolds are locally isomorphic to Euclidean space, CW complexes are isomorphic to polyhedra, schemes are locally isomorphic to the sets of solutions of algebraic equations. Each of these kinds of space has an associated cohomology theory – de Rham cohomology for manifolds, singular cohomology for CW complexes, and so on – which is defined by means of the appropriate comparisons between local and global

⁵See [4, Chapter 4].

⁶Notice that here is a case of axiomatisation which explicitly has nothing to do with basic entities of any sort, but rather with the standardisation of processes of reasoning which recur in many different concepts. Most occurrences of axiomatisation in mathematical practice seem to be of this sort rather than the foundational sort.

which are appropriate for that kind of space. One should notice that this procedure is essentially open ended; one simply has no grasp of the *totality* of theories satisfying the Eilenberg-Steenrod axioms. The list might, in principle, be extended at any time; the appropriate cohomology theory for schemes, namely étale cohomology, is an invention of the 60's. And each cohomology theory will carry with it an appropriate notion of dimension.

1.2 The Role of Axioms

This example shows that the role of axiomatics, and of exact definitions, in mathematics is perhaps rather other than it is commonly taken to be. According to the naive picture (mathematical objects are made out of basic mathematical entities, and the axioms of mathematics give us access to these), the axioms of a given theory ought to apply to the basic, simple objects out of which others are constructed, and furthermore the axioms and definitions ought to be epistemically the most entrenched parts of mathematical reasoning. Both of these fail to be the case here; the axioms we deal with (the Eilenberg-Steenrod axioms) are certainly not about basic entities, and the definitions and axioms are certainly not the most entrenched parts of this area of mathematics. What is most entrenched, and most resistant to change, are the rough desiderata which a theory of dimension ought to satisfy – that it should measure the number of independent possibilities of variation, and so on. It is these rough and open ended criteria which mathematicians appeal to when they are evaluating more formal mathematical work.

2 Conclusion

How does one evaluate this? Philosophically, there is clearly work to do; a critique of the role of basic entities in mathematical ontology, an investigation of the links between mathematical ontology and mathematical practice, and so on. These are clearly problems worth philosophical examination. However, from the point of view of *Kulturkritik*, other things are perhaps more important.

One is this: that the original hope was to find entities which would play a fundamental role in mathematical ontology, in mathematical practice, in our commonsense grasp of the world, and also be suitable for application in artificial intelligence. This is clearly an overdetermined problem; one would be very lucky to find a single group of entities which do all of this simultaneously. And, in fact, one can't; there are many senses of 'basic' at issue here, and this leads to tensions between the mathematical and the philosophical sides of the problem. These tensions are, of course, what makes the ontology of mathematics an interesting and difficult problem. But it is difficult not to see the original hope as somehow the result of a belief in magic, and similarly it is difficult not to see the critical investigation of axioms as somehow a belief in infection.

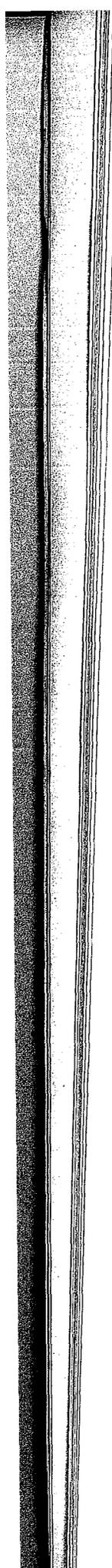
The second phenomenon is this. The problem that philosophers and artificial intelligence specialists are addressing with these efforts at axiomatisation is the problem of qualitative geometry: how can one investigate the geometrical properties of

entities in abstraction from their metric properties? It is a perfectly good problem (and, like many problems in computer science, it is a problem of attaining an appropriate level of abstraction). It is also a problem that has been, on the whole, solved; it was solved in the 60's by algebraic geometers and algebraic topologists. It is difficult and sophisticated work, but maybe necessarily so; qualitative mathematics is generally more difficult than quantitative mathematics, and it is hard to see that this should be any exception. Many of the themes which have surfaced in the philosophy/artificial intelligence community have been already discovered by the algebraic geometers (for example, the view of regions, rather than points, as primary). Furthermore, it is naturally presented in terms of category theory, and category theory has far better and simpler connections with theoretical computer science than does the theory of axiomatic systems. This is not magic, but perhaps a constellation of circumstances which works better than things normally do; it should certainly be investigated.

The third thing is this. There has been a great deal of polemical fuss made over the question of this or that axiom system. This is pointless: mathematicians don't, on the whole, use axiom systems, so the question of whether they are using the right one is a non-question (especially since normal mathematics seems to have been remarkably successful in dealing with the problems the philosophers are raising). But there are many genuine problems to be faced, which are both genuine philosophical problems and, at this juncture, important in a wider context. There are the claims of cognitive science to be evaluated, and there is also the problem of a growing anti-intellectualism, which seems to take the qualitative as a symbol for the non-theoretical. There is a great deal of work to be done here, work which will take philosophical acumen and also genuine metatheoretical insights. And it is important not to let a belief in magic distract one from more urgent tasks.

References

- [1] Paul Benacerraf. What numbers could not be. *Philosophical Review*, 74, 1965.
- [2] M. P. Fourman, P. T. Johnstone, and A.M. Pitts, editors. *Applications of Categories in Computer Science*. Number 177 in London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, 1992.
- [3] David Charles McCarty. Incompleteness in intuitionistic metamathematics. *Notre Dame Journal of Formal Logic*, 32:323-358, 1991.
- [4] Hans-Otto Peitgen, Hartmut Jürgens, and Dietmar Saupe. *Bausteine des Chaos: Fraktale*. Springer, Berlin, 1992.
- [5] S. K. Thomason. Free construction of time from events. *Journal of Philosophical Logic*, 18(1):43-67, February 1989.



Rethinking Boundaries*

Wojciech Żelaniec

State University of New York at Buffalo

Abstract

In this paper, I offer a re-examination of the idea that general topology could be a new basis, or an essential part of a new basis, for ontology conceived, in the traditional sense, as a “science of being.” I analyse a number of difficulties and I do *not* offer a way out, except, perhaps, that of giving up the name of “topology” for whatever *could* be a new basis for ontology.

I. Ontology and Topology

It is occasionally claimed—and in this volume you will find more than one example for it—that *topology* might not only be useful for knowledge representation research (in particular with regard to knowledge of spatial relations) but also serve as a new foundation for ontology, the science of being.

This article is a contribution to a critical appraisal of this claim. Such an appraisal is not possible without saying what one’s concept of ontology is. Let me, therefore, start out with this.

Ontology is an old philosophical discipline—though its name was coined only in the 17th century—and for many centuries it was, under the name of metaphysics, the central one.

It started, like almost everything in philosophy, with Plato and Aristotle, was continued by Plotinus, revived in the mediaeval Islamic and Jewish philosophy and the Christian Scholasticism, reaching its heights in the work of Thomas Aquinas and Duns Scotus, it was handed down to the modern times by original thinkers like Francisco Suárez, underwent a transformation in the work of the great Continental philosophers of the 17th century: Descartes, Malebranche, Spinoza and Leibniz, went through a period of decline in the 18th century, but resurrected in the early 19th century, first in the strange form of the late German Idealism (Schelling, Hegel), and in the work of solitary thinkers like Kierkegaard and Schopenhauer.

A major turn occurred when two parallel currents started to flow in the second half of the 19th century. (They have continued to flow—albeit in wide, and ever wider-growing, deltas—to this very day.) One of them—known as Neoscholasticism—

*Work supported by a grant from the Austrian *Fonds zur Förderung der Wissenschaftlichen Forschung*, grant number: P8661-HIS.

was a revival of the mediaeval Scholasticism which began as an attempt to re-read the mediaeval Scholastics in the search for what could make them relevant for the modern mind, but produced, and goes on producing, original work. The other current was a tradition inaugurated by Franz Brentano, a German-Austrian thinker of the late 19th and early 20th century, who was himself an offspring of the first current. He was, however, dissatisfied by what appeared to him, in Neoscholasticism, as too much traditionalism and too less interest in empirical data and in the discoveries of modern natural science. He developed, therefore, a family of highly original—though recognizably Aristotelian in their basic tenets—doctrines on various ontological, though not only ontological, topics. Brentano found numerous brilliant students and followers who, in their turn, had their own bright and enthusiastic students and followers, mainly in countries of Central Europe (which used to be a part of the then Austro-Hungarian Empire), and among whom Alexius Meinong, Stanisław Leśniewski, Edmund Husserl, and Roman Ingarden may be the best known.

The family of these Brentanians, grand-Brentanians and grand-grand-Brentanians, in the meantime partly transplanted to the United States, has been growing ever since, and expanding all over the world. This living (and lively) tradition has, in the meantime, become mature enough to afford to publish an authoritative work, with the title "Handbook of Metaphysics and Ontology,"¹ at which the reader will be well-advised to have a look, if she or he wants more technical information. But the British philosophy has not been intact from the ontological resurgence, either. Suffice it to mention the British Idealism, with T. H. Green and F.H. Bradley, a *sui generis* continuation of the German Idealism, and Moore and Russell's revolt against it, which did not fail to produce a good deal of ontological doctrines and arguments, within the framework of a philosophical system known as "Logical Atomism." In Russell's case, the influence of the above-mentioned Brentano student Meinong was significant, if only because Russell formed his views in opposition to Meinong's. In today's Anglo-American analytic philosophy, even some heirs of the tradition of the Vienna Circle, like David Lewis, have taken interest in ontological issues, and managed to produce valuable work in this area.

So much by way of a historical introduction. Let me now say a few words about the substance ontology. There are hardly any general theses to which ontologists of all these various traditions, schools and persuasions would subscribe. What they share, is rather a class of basis presuppositions, more often tacitly held than expressed. One of them is the presupposition that all existing things form a unity which is not like that of an agglomerate, or a heap, but which is governed by certain general principles. What these principles are, is, consequently, a major problem, or rather, a set of problems, shared by all ontologists. In fact, ontology, which is nothing like a mere catalogue of various existing things, can be defined as a "science of the general principles of being."²

The belief that there is a unity of everything that exists may seem, to the unin-

¹H. Burkhardt and B. Smith (eds.) *Handbook of Metaphysics and Ontology*, 2 vols., Munich/Hamden/Vienna: Philosophia, 1991.

²Or, in the famous formulation by Aristotle *Metaphysics* Γ, 1003a, 20–21, a science of being as being and the attributes belonging to it in virtue of its own nature.

tiated, utterly strange, and that for two reasons. Firstly, because most of the intellectual as well as practical work of humanity is devoted to questions pertaining to special domains of reality. We are interested in planting a tree, baking a loaf of bread, building a house, constructing a machine, splitting the atom, subduing a nation, cracking the jackpot, as well as in the various particular truths which may be helpful in doing all those various things. Seldom, if ever, are we interested in "everything that exists." Yet still, as ontologists believe, compartmentalized between various domains of practical and theoretical interest as reality is, it does form a unity, not only in virtue of common principles which govern it, but also in virtue of relations among those various domains. The second reason why the idea that "everything that exists" may not seem to be a respectable subject matter of a disciplined intellectual activity, is that this seemingly restrictive phrase "that exists" does not make sense. Are there things that do not exist? Which ones are they? Centaurs or golden mountains, perhaps? But ontologists are not easily daunted by queer entities like these; they think that if the existence of centaurs or golden mountains is a dubious case, then it must be cleared up, or at least, it must be explained why it is a dubious case.

Now, in the recent decades, ontologists of certain persuasions have become very much preoccupied with the latter sort of difficulty; their work has become, to a large extent, a search for, and an analysis of, "tricky examples." But nonetheless, the first problem, that of the general principles governing reality in its whole, has not been lost sight of, though few people today have the stamina to address that major problem of ontology directly. I should even venture to say that the popularity of topology and other mathematical applications in ontology is due mostly to the hope that advanced modern mathematics may be of decisive importance for ontology, in that it will help it to achieve firm results more quickly than it has been possible so far. For ontologists, despite their over two millennia of hard work, have produced only few firm results.

In what follows, I shall, therefore, concentrate mainly on the presupposition that there are some general principles of governing everything that exists.

Now, if there are such general principles, where are they to be sought? In other words, what *kind* of principles are they? In other words still, what is their nature? Ancient Pythagoreans thought they the principles at issue are *mathematical* in nature. Plato shared this view, to a certain extent, while Aristotle opposed it. The idea recurred, in various subtle versions later, prominently in the era of large meta-physical systems in the 17th century. It re-emerged with great vigour in early 20th century, when a mathematical theory of classes was developed. It began to look as if the theory of classes, later known, purged of its initial difficulties, as the theory of sets, could be a new basis for ontology, or indeed be itself a fundamental part of ontology.³ But today, after so many decades of only partly, at best, successful attempts actually to erect a solid edifice of ontology on the set-theoretical fundament, it is no longer certain if set theory alone would do. One of its most obvious flaws is that it is too abstract and unable to deal with qualitative richness and diversity

³Recall that Leśniewski called one of his three systems of what could have become an alternative theory of sets "ontology."

of reality. However, it is not only *qualitative* features of reality that set theory does not seem to be able to capture, at least in no natural way, but also some merely formal features of reality, e.g. the relations of dependence, or of boundedness, or of being inside or outside. It is such features that *topology* seems to be called upon to provide an theory of, a theory which, with suitable modifications, would become a fundament of ontology.

While I agree, however, that topology is conceptually richer than set theory—how could it not be, given that topological structures are usually defined on a family of *sets*—it is my contention that possible applications of topology to ontology are rather limited, and stretching them can be even dangerous in that it can introduce into ontology certain unexamined and possibly false assumptions.

II. Methodological problems

In what follows, I should concentrate on methodological problems troubling applications of topology to ontology, for there are quite of few of them, some of which can be summarily expressed in the question: how do we link topology to ontology? How do we find, identify, or impose topological structures in what constitutes the subject matter of ontology? And does the study of such structures afford us any new insights into the subject matter of ontology, i.e. insights which would not, or not easily, have been possible without the introduction of a topological structure? Is it not much rather, as with everything fashionable, the case that topological structures are imposed upon the subject matter of ontology in an artificial and arbitrary manner, so that the results of this work are simply things which we have always known, but in a new packaging?

We all, I think, agree that topology alone and as it stands, *is* not ontology, even if it can disguised as one. We could make the convention of, say, referring to sets as to “entities,” to spaces as to “universes,” to closures as to “closures” (in some mysteriously changed “ontological” sense) etc. This is not ontology, however, or perhaps it is just an ontology of those abstract mathematical objects with which topology deals. One could retort that this is a common problem of all disciplines to which mathematics is applied. This is true. But it is also true that not to every discipline mathematics is applied equally successfully. Mathematics gives more precision and rigour than we should otherwise have had. But it does this at a certain price: that of disregarding various aspects of that to which it is applied. In the case of ontology, we should ask ourselves the following questions: is the price not too high? Does the topological language and method as applied to ontology yield important ontological truths which we should not otherwise have learnt? Are these new truths—if any—so important as to justify putting up with the loss of other important truths which escape the topological methodology?

In any case, we need, at the very outset of the enterprise of introducing ontological methods to ontology, some link between the subject matter of ontology, as we know it from the hitherto existing ontological tradition, and topology. In other words, we need ontological models for topological structures.

This postulate is so obvious that it might seem superfluous to mention it. But how, exactly, are such desired models found?

1. Closure and dependence

We seem to have a number of ways at our disposal. We can note, for instance, that if we read " $cl(x)$ " (closure of x , a fundamental topological expression) in a particular ontologically interesting fashion, e.g. as " x plus everything on which x depends" or " x plus everything which depends on x ," $cl(x)$ might satisfy, so interpreted, the Riesz-Kuratowski axioms of the operation of closure. And these axioms define "a topology" (as mathematicians say), a basic topological structure with reference to which most other structures can be defined. This is very interesting, and a resulting ontology could look very neat, elegant, and economical. But there are several problems with it.

First of all, the relation of dependence is itself very difficult to understand. There is, strictly speaking, not *one* such relation, but several (as Ingarden has shown in his *Strife on the Existence of the World*): for instance, dependence with respect to coming into existence, dependence in the sense in which a part is dependent on its whole, and in the sense in which a universal is dependent on its instantiations, dependence with respect to staying in existence, finally, dependence of the determinate on the determinable, as e.g. the dependence of red on colour.⁴ And also (something skew to mathematical interest, because pure mathematics does not busy itself with particular objects): dependence as an individual, and dependence as a specimen of a natural kind, or as a token of a type. For instance, this performance of a piano sonata is dependent upon this particular pianist, and a number of other factors, though the type "performance of the piano sonata" is not so dependent, although some subtypes of this type might be dependent on this particular player and other circumstances on which the present item of the performance is dependent.

Note that we cannot say that on the abstract level of topology this is all the same, because this would be either begging the question (have we really checked that it is the same?) or it would involve substituting the well known topological closure for ontological dependence, while keeping the name of the latter. And this would amount to doing good old topology under an "ontological" guise (with new fashionable words, like "dependence" for "closure," for instance).

2. Inside, open sets, and what?

Another problem is that in topology, once we have defined a topological structure in terms of closure, we can define other operations, or other structures equivalent to the former. For instance, we can define the operation of taking the interior of x : $Int(x)$. But under the "ontological" interpretations like the afore-mentioned, interior in the topological sense does not have any natural ontological counterpart. Because, if the interior of x is defined either as the complement of the closure of the complement of x , or as the maximal open subset of x , the problem is that we do not seem to have, at first blush, natural ontological counterparts of either complement, or of open set. Saying that the interior of an entity is the sum of everything that is inside

⁴The reader will find a systematic classification of these types of dependence and their reciprocal relations in: Roman Ingarden, *Der Streit um die Existenz der Welt*, vol. I (*Existenzontologie*), Tübingen: Niemeyer, 1964, pp. 74–130. Interestingly, in one footnote (nr. 49, p. 115) Ingarden reproaches Husserl with disregarding these distinctions in the 3rd *Logical Investigations* and labelling *all* types of relations of dependence as formal.

the entity does not really help, for the problem is exactly *what* is inside an entity. What criteria shall we use in deciding whether this or that is inside of an entity? Such criteria will vary from one entity to another, and they must be found *before* the mathematical machinery is set in motion, so that one should wonder of what avail the latter should be, after all. Besides, is the inside of an entity really a *natural* ontological item? But if it is not, and the topological concept of an interior does not have any natural ontological counterparts, neither does boundary (standardly defined as the result of the subtraction of the interior of a set from its closure) have any natural ontological counterparts.

It is also difficult to see what "open sets," another important topological animal, might be, ontologically speaking (though we have a name for them handy, namely "open entities"). We can ask ourselves, as a matter of a simple thought-experiment, if Socrates is his own closure, that is to say, he is everything he depends upon (though he may not be, under some readings of "dependence"), what, then, could be an open part of Socrates? His brain, perhaps, or his ugliness? But this would mean that the Universe minus Socrates' brain or minus his ugliness is its own closure, which is a matter we must investigate into separately, as it were, without any help to be derived from our topological presuppositions. And it is not at all clear on what criteria we should decide whether the Universe minus Socrates' brain or minus his ugliness is its own closure: the very formulation of this problem sounds strange. So perhaps the ease with which "dependence" (conceived in a most general fashion) can be fitted into the Kuratowskian axiomatics of closure, does not point to any deep affinity of topology and ontology, but, rather, is merely accidental.

3. Boundaries

But somebody might say that the source of trouble is that we started introducing topology to ontology on too high a level of abstraction. We should have started with the observation that things do have boundaries, and we should have made the concept of boundary the main link between topology and ontology.

Maybe. But we should examine the observation, if indeed an observation it be, that things have boundaries. Is there not a multiplicity of senses in which we can speak of boundaries of things, analogous to the multiplicity of the relations of dependence in which they can stand to one another?

Now, an enthusiast of the topological approach in ontology might respond that there is one major, privileged, standard-setting, sense in which we can speak of boundaries of things, and that is: spatial, and temporal, boundaries of things.

This is certainly true. But then there are at least three objections that spring to one's mind almost immediately.

Firstly, if we are interested in spatial and temporal boundaries of things, this is a perfectly legitimate interest, and hardly a recent one, but why should we expect anything of ontology, and why should we think that this subject-matter would be relevant for ontology? There are many sciences, starting with applied mathematics, various sorts of engineering, and including psychology and cognitive science, which deal with spatial and temporal boundaries of things rather successfully. Even the old Eleatic paradoxes, which doubtlessly were a part of ontology at that time, have been

successfully solved by the theory of convergent infinite series, a mathematical theory invented in the 18th century. And topology can, and perhaps even does, contribute greatly to the progress of those sciences, the topological structures themselves being neatly defined in terms of the Euclidean metrics, for instance. Ontology, properly conceived, does not arrogate to itself anything of the domain of those respectable branches of science.

Rather, ontology examines, as Aristotle once put it, being as such and its essential attributes. But a cumbersome fact in this connection, and a *second* difficulty, is that for a good many categories of beings, being spatial or "in space" is not really an essential predicate. This is true particularly of various categories of things that are elements of our common-sense world, and of our everyday-experience. Things of vital interest to us, as animals, humans, and academics.

For instance, as animals, rational or otherwise, we are interested in food. But please note that neither bread nor wine are food solely in virtue of what they are as "chunks of stuff" occupying space. They are food also in virtue of how our digestive systems are constructed and equipped, and of how they—our digestive systems—react to bread and wine. Being food, like being a colour, is not an inherent property of a chunk of matter; it is, rather, a relational property straddling—not literally, because no spatial straddling is involved here—two chunks of matter, or rather species of such chunks. It is only because our digestive systems change very slowly over millennia and are therefore a stable factor, that we can incorrectly think that bread and wine are food simply as "chunks of stuff."

As humans, we are interested in communication with other humans. But communication involves various types of objects which do not have any spatial structure, either. For what is the spatial structure of signs? Signs are types of objects with meaning attached to them, and it is those objects, not the signs themselves, that have spatial, or more often, temporal structure.

As academics we are, I believe, primarily interested in tenured posts, but these, though they have a rather complex structure of their own, are not spatial objects, either, nor is their overall structure a spatial one.

Are states, and other political units, spatial objects (as is sometimes pretended)? They seem to be bound up with a certain spatial object, namely an area on the surface of the Earth, but they depend, for their existence upon many other entities that do not, at least at first look, appear to be spatial: structures of power, for instance.

We could multiply indefinitely examples of types of objects which are important to us, and without familiarity with which survival, let alone anything more demanding than mere survival, would be impossible, and which, nonetheless, are not constituted by any spatial properties.

What all such examples have in common, is, I believe, the fact that even though the objects in question do have some spatial component to them, it is not that component which makes them what they are, or not that spatial component alone. When we impose a topological structure on such objects, we usually anchor it in that spatial component, and we pretend that something important and fundamental can be learnt about those objects by studying that structure, and we thereby make out that spatial component to be itself important and fundamental. In this way,

reality begins to appear to us as basically mathematical, or—to be more specific—geometrical, as consisting of “chunks of stuff” (in the nowadays fashionable parlance) which are more or less neatly divided from each other. In this way, we have come back to where Pythagoreans were, though our ontology mathematicised is fascinated less by number and more by topological structures.

One of the sources of this error is the modern reverence for mathematics and the natural sciences, impregnated by mathematics, and fascination with relatively new branches of mathematics, such as topology. Another source is, I think, the old confusion caused by the fact that two different sorts of mental activity distinguished by Aristotle as “*aphairesis*” and “*khōrizein*” came to be translated into Latin by means of the same term “abstraction.” Very roughly speaking, the first activity abstracts from the physical bulk of things, and their qualitative properties bound up with it, and extracts their mathematical form, their shape, or, as we might say, their topological structure. The latter activity captures the “form” of a thing (in the specifically Aristotelian sense of the term “form”), its essence, or that which makes a given thing to what it is. These two different sorts of abstraction came to be distinguished in different ways from about the time of Avicenna at the latest, be it by means of different terms (Aquinas called the latter sort of abstraction “separation”), or by means of the concept of “degrees of abstraction.” Rudolphus Goclenius, one of the first, if not *the* first, philosopher to the term “ontology,” introduces ontology (science of being) as the discipline for which the third degree of abstraction is the proper method: ontology abstracts of matter “*tam secundum rem quam secundum rationem*” (with respect both to things themselves and to their concepts)—he tells us in his monumental *Lexicon Philosophicum*⁵—whereas mathematics employs abstraction of the second degree, where particularities due to—as we should say today—chemical properties of matter are disregarded, but conceptually, matter, as something that fills up space, is still taken into account.⁶ The identification of the mathematical form and the form in the sense of “that which makes the thing to what it is” has persistently recurred in major philosophical systems (cf. Cartesian *res extensa*), and has flown together in our modern usage of the word “form” in which these two senses are often present undistinguished. Ancient as this confusion is, there is no reason, I believe, to persist in it, let alone to aggravate it by extending mathematical techniques beyond the limits of their applicability.

And, as it happens, the distinction is not very difficult to observe. When we note that an object has some spatial component to it, we should ask ourselves if it is also constituted by that component, or if the existence of this component is a sufficient condition of the existence of the object in question. In most cases we shall see that it is not even a necessary condition of the existence of the object in question, though *some* spatial component (not necessarily the one which the object actually has, and not necessarily any specific one) might well be a necessary requirement.

⁵Frankfurt, 1613, p. 16.

⁶This is very much in line with the way in which Aristotle characterises the subject-matter of mathematics: shapes (*schemata*) of things, but taken in abstraction from the particular properties of the things they are shapes of. The mathematician concern himself with the curved, not with a snub nose (*Physics* I, 2, 193b30–194a11).

But could we not invent a deviant topology, tailor-made for non-spatial relations? Perhaps. But why should we still call it "topology"?

The *third* difficulty is that we should not uncritically assume that we know what the spatial and temporal boundaries of things are. Ontology is concerned with what is real, but consider how unequal, even *prima facie*, the degree of reality of boundaries of the following sorts of objects appears to be:

- planets, stars
- animals and humans
- continents, islands, peninsulas
- climatic zones
- states and other political units.

While it seems certain that entities of these categories are bounded, it is far from clear where, exactly, their boundaries are, and how they can be traced out. Sometimes, the boundary of an entity can be safely identified with one of its parts, e.g. with its skin, in the case of an animal. In other cases, the boundary is, not so much discovered in an entity, as ascribed to it, although non-arbitrarily; this is the case of continents. In still other cases, the boundary is ascribed completely arbitrarily. There appears, therefore, to be pretty little that an ontologist *qua* ontologist can say about boundaries in all generality, and little is won by adopting for this purpose a formal apparatus of mathematics: the real job has to be done on the level of various material ontologies: those of celestial bodies, of animals, of continents etc. That is to say: we can, of course, keep doing topology and pretend that we are doing ontology by inserting occasional references to "entities" or some other pieces of the ontological inventory, and if we find doing topology with such references more attractive than without them, all the better. But for it to be real ontology, we must supply evidence to the effect that references that we make are to objects that satisfy, and this not in a strained, far-fetched way, axioms of topology, deviant or not. In other words, we must show, or rather constantly be showing, that those objects do, *qua* objects of ontology, have some topological structure. And I predict that this will be the most difficult part of the job, far more difficult than the relatively simple topology that we shall be doing. We shall have to do a good deal of non-elementary material ontologies of different object-domains.

III. Prospects of the future

Doing material ontologies with a view for making topologies—deviant or not—applicable to them will a very unrewarding job, as I should also venture to predict, for the following reason. I showed earlier that defining a topology by recourse to the relation(s) of dependence is somewhat far-fetched because many natural topological concepts do not have equally natural counterparts in the domain of whatever is definable with recourse to those relations. Now, something analogous holds true for topologies definable with recourse to boundaries. Even if we knew very well what boundaries of various sorts of objects exactly are, we should still have the problem of finding natural applications of other topological concepts in the domain of boundaries and bounded objects themselves. The concept of boundary is not, as we should not forget, a central concept of topology. It can be, certainly, used as a

primitive in terms of which all the Kuratowski-Riesz axioms can be formulated, but the mathematician's point of view this would be rather unnatural, which explains why it is done so seldom, if ever. Topology as a branch of mathematics is not, in any of its various subbranches, interested primarily in objects, or in their boundaries (or their closures or interiors, for that matter). It is, rather, interested in whole domains of objects (called "topological spaces"), in special types of functions (called "homeomorphisms") and in the invariants of these functions. Now, in the domain of the subject-matter of ontology, it would be difficult to find anything like "ontologico-topological spaces" or "ontologically generalized homeomorphisms." And without them, the topology in ontology is bound to remain very elementary and inchoative.

But, as I said before, productive applications of topology to ontology will be possible no sooner than natural ontological models for at least *one*, be it even a non-central, topological concept have been found, for instance for the concept of boundary. Let it be stressed: no sooner than such models have really been found, rather than just assumed to be already there, simply because we have a word, like "boundary," which we can make ourselves guilty of a merely analogous use of (without its being clear what the basis of the analogy is).

And, to make a sort of compromise with those of us philosophers who are interested in the topological approach to ontology, we might propose that there should be something which could be called "pre-topological ontology," that is, a branch of ontology devoted to investigating objects of various sorts and types from the point of view of the applicability of different topological concepts to them. As I suggested above, there is a great deal of interesting truths that can be discovered in the process of investigating what it is for objects of different sorts to have boundaries, (which we can preliminarily construe as qualitative differentiations, or, in the mode of Kurt Lewin⁷, as "zones of resistance," or in some other fashion).

But even here, at this pre-topological stage, doubts might arise if it will be followed by some fully-fledged topological stage, since boundaries as we really find them in objects, display characteristics obviously deviant from those of boundaries in the topological sense. For instance, they are often not "thin": they are zones rather than skins. Yet, if it is the fashionableness of topology which would help us to reach a result like this, then we could say that topology has had a reason to be fashionable among philosophers.

⁷The Gestalt-psychologist defined barrier as a boundary zone which offers resistance to locomotion. He noted, however, that this resistance can "be different for different kinds of locomotion, for locomotion in different directions and at different points of the barrier. See his *Principles of Topological Psychology* New York/London: McGraw-Hill, 1936, p. 217. Very much in accordance with my above diagnosis, his "topology" remained very inchoate, at best, and he eventually gave up the name of "ontology," calling it "hodology" — science of paths, instead.

SHEAF MEREOLGY AND SPACE COGNITION

Jean PETITOT

CREA, 1 rue Descartes, 75005 Paris

EHESS, 54 Bd Raspail, 75006 Paris

petitot@poly.polytechnique.fr.

CONTENTS

Introduction.

- I. The morphodynamical schematization of predication.
- II. The indexical and modal status of spatial extension.
- III. Computational vision and receptive field jet calculus.
- IV. Husserl's *Erfahrung und Urteil*.
- V. Sheaves, Topoi and Logic.
 - 5.1. The concept of sheaf.
 - 5.2. Generalizations.
 - 5.3. The concept of topos.
 - 5.5. Sheaf mereology and boundaries.

Conclusion.

INTRODUCTION

Contemporary research concerning the cognitive links between perception, language, and action — see for instance Talmy's works — have revolutionized the dominating traditions of formal semantics. They have led to what Husserl already called in *Erfahrung und Urteil* a *genealogy of logic*, a perceptive genealogy of cognitive symbolic structures. Such a Gestalt-like perspective on symbolic structures raises a lot of new problems. It makes traditional simple — and even trivial — problems appear as complex, non trivial, ones.

It is one of these problems I want to address here, namely that of the most elementary and primitive forms of *predication in perceptive judgement*: “the snow is white”, “the sky is blue”, “the ball is red”, etc.

For classical elementary logic these atomic formulae “ S is p ” stay at the lowest level of complexity. Both their syntax and their semantic are trivial. But it is no longer the case if one takes into account the perceptive situations to which these statements refer. Indeed, these situations compel to introduce in the formalization of symbolic logical structures, *topological and morphological* structures of a completely different kind. If one

wants to do semantics that way, one is therefore committed to elaborate a *mathematically integrated* theory of topology and logic.

Let us see briefly why the problem of taking into account the perceptive roots of predication is a non trivial one. On the perceptive side we find a *geometrical descriptive eidetics* of the morphological (gestaltic) organization of perception. I use here the term “morphology” in the sense of Thom: a morphology is a set of qualitative discontinuities on a substrate space. But on the logical side we find a logical eidetics of judgement: a formal syntax and a formal semantics (e.g. a vericonditional theory of denotation). But between these two eidetics there exists a dramatic gap, what Thom (1988) called an “insuperable break” (*Esquisse de Sémiophysique*, p. 248). Our challenge here is to fill this gap using some sort of synthesis between logic and topology.

The gap is linked with deep traditional philosophical questions, for instance that of the relations existing between *analytic and synthetic* rules. Morphological eidetics depends on synthetic laws of perception. On the contrary, logical eidetics depends on analytic laws.

I. THE MORPHODYNAMICAL SCHEMATIZATION OF PREDICATION

Let us consider how René Thom has morphologically schematized a perceptual attributive judgement “*S is p*” such as “the sky is blue”. The truth conditions of such a statement depend on the way, as Husserl would say, the “intentions of signification” are *intuitively filled* (in the sense of an *Erfüllung*) at the *ante-predicative and prejudicative* level. Without an adequate description of this intuitive filling-in, one can’t elaborate a correct theory of predication.

Let W be the spatial extension of S . W is filled by a quality p belonging to a genus G (e.g. colours). The filling-in is then described by a map

$$f: W \rightarrow G, w \rightarrow f(w).$$

Husserl has given a deep analysis — but lacking of mathematical modeling — of this fact in his third Logical Investigation. I have shown that to model adequately *the unilateral dependence relation* linking the extension W with its dependent part — its *moment* — p , one has to interpret the map f as a *section* of the (trivial) *fibration*

$$\pi: E = W \times G \rightarrow W.$$

It is exactly the point of view mathematized by Thom.

Let me recall briefly what are fibrations and sections of fibrations.

Mathematically, a fibration is a differentiable manifold E endowed with a *canonical projection* (a differentiable map) $\pi: E \rightarrow M$ over another manifold M . M is called the *base* of the fibration, and E its *total space*. The inverse images $E_x = \pi^{-1}(x)$ of the points $x \in W$ by π are called the *fibers* of the fibration. They are the subspaces of E which are projected to points.

In general a fibration is required to be *locally trivial*, that is to satisfy the two following axioms:

(F₁) All the fibers E_x are diffeomorphic with a typical fiber F .

(F₂) $\forall x \in M, \exists U$ a neighborhood of x such that the inverse image $E_U = \pi^{-1}(U)$ of U is diffeomorphic with the direct product $U \times F$ endowed with the canonical projection $U \times F \rightarrow U, (x, q) \rightarrow x$. (See figures 1, 2).

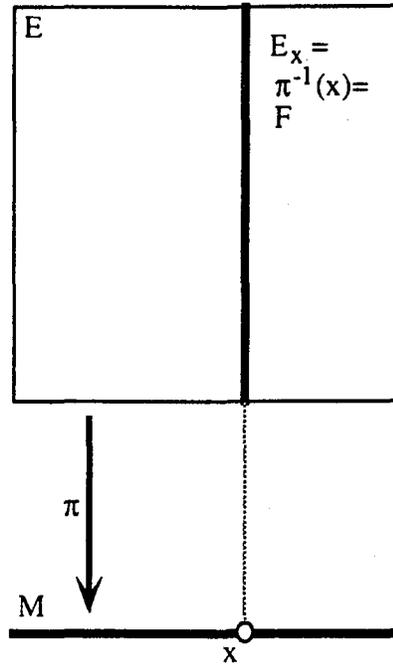


Figure 1

$$E_U = U \times F$$

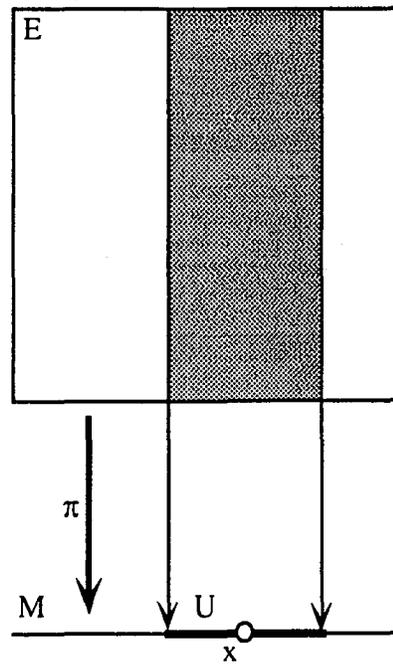


Figure 2

In our case, we have $M = W$ and $F = G$.

Let $\pi : E \rightarrow M$ be a fibration and let $U \subset M$ be an open subset of M . A *section* s of π over U is a *lift* of U to E which is compatible with π . More precisely, it is a map $s : U \rightarrow E$, $x \in U \rightarrow s(x) \in E_x$, i.e. such that $\pi \circ s = \text{Id}_U$. In general s is supposed to be continuous, differentiable, analytic. It can present discontinuities along a singular locus. (See figures 3, 4).

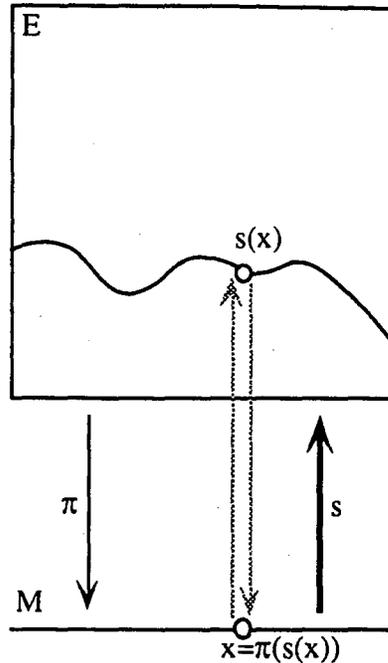


Figure 3

It is conventional to write $\Gamma(U)$ for the set of sections of π over U . If there exists a local trivialization of π over U (i.e. $E_U \rightarrow U \approx U \times F \rightarrow U$), it transforms every section $s : U \rightarrow E$ in a map $x \rightarrow s(x) = (x, f(x))$, that is in a map $f : U \rightarrow F$. Therefore, the concept of section generalizes the classical concept of map.

The main reason why it is of interest to conceive of the qualitative filling-in of spatial regions as sections of fibrations is that one can easily retrieve that way the dependence relations (i.e. the ontological hierarchy) between "essences". The fact that spatial extensions are ontologically prior to qualities such as colours is modeled by the fact that the extensions W are the *base spaces* of the fibrations, and can therefore exist in an *independent* way, while the qualitative genera G are the *fibers* of the fibrations, and can therefore only exist in a *dependent* way.

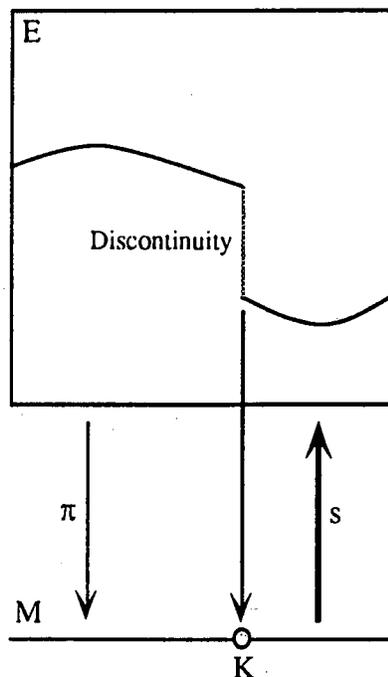


Figure 4

In general the genus G will be *categorized* and decomposed in sorts. Let p be such a category (e.g. “blue”) and let ∂p be its boundary in G . The perceptive statement “ S is p ” is morphologically interpreted by the following geometrical fact:

the image $f(W)$ of the section $f: W \rightarrow W \times G$ is encapsulated in the cylinder $W \times \partial p$.

These morphological descriptions can be generalized. If for instance there exists an order relation $<$ on G (e.g. p is lighter than p'), the statement $S(W) < S'(W')$ means that $\rho(f(w)) < \rho(f'(w'))$ where ρ is the second projection $\rho: W \times G \rightarrow G$ (that is $\rho(f(w))$ is the codomain of values of f).

The problem I want to address is very simple. In the framework of a semantic theory, what can be the link between two definitions of truth (for atomic formulae):

- (i) the tautological Tarskian one : “ S is p ” is true iff S is p ;
- (ii) the morphological Thomian one : “ S is p ” is true iff $f(w)$ is encapsulated in the cylinder $W \times \partial p$?

II. THE INDEXICAL AND MODAL STATUS OF SPATIAL EXTENSION

In the categorial syntactic structure of a statement such as “ S is p ”, the *specificity* of the spatial extension W of S is in some way *bracketed*, “implicitated”. Of course, one can focus on the extension W itself and consider its regions, its points, etc. But in doing so one changes the level of description. At the level at which S is processed as an individual entity sharing some properties — that is at the level at which the symbolic description “ S is p ” is relevant — the specificity of W is bracketed. This strange sort of

“*epoche*” concerns also the *relations* between qualities, as when we say “S is lighter than R”.

This problem has been well pointed out by Wittgenstein in his *Remarks on Colour* (see Mulligan 1992).

“A language-game: Report whether a certain body is lighter or darker than another. But now there’s a related one: State the relationship between the lightnesses of certain shades of colour. The form of the propositions in both language games is the same “X is lighter than Y”. But in the first it is an external relation and the proposition is temporal. In the second it is an internal relation and the proposition is timeless”.

The tautological aspect of Tarski's definition of truth is a symptom of this “*epoche*”. It is what remains when syntax and semantics uncouple themselves from their perceptive rooting.

As a matter of fact, one can say that *every* perceptive statement is in some sense analog to an *indexical one*. It refers to its perceptive situation in much the same way as an indexical statement refers to its pragmatic context. The intuitive *Erfüllung* of the intentions of signification operates in a *counterfactual* fashion and one needs therefore some sort of Kripke's semantics to take account of this fact.

If one wants to understand how classical logical structures can have a non classical *space-semantics*, one has to understand how classical variables can semantically denote *sections of fibrations* in such a way that extensions *can condition truth*.

III. COMPUTATIONAL VISION AND RECEPTIVE FIELD JET CALCULUS

Before tackling this point, I want to emphasize the *neuro-physiological* relevance of the geometrical concept of fibration.

The most influential specialists applying differential geometry to computational vision have hypothesized that cortical columns implement a *multi-scale jet calculus* and that it is essentially that way the brain does geometry. Let us consider for instance Jan Koenderink's perspective.

David Marr already showed at the end of the seventies that the retinal ganglion cells perform what is now called a *wavelet analysis* of the signal $I(x,y)$, that is a *multi-scale and spatially localized* analysis. They act as *filters*, that is they convolute the signal by the receptive profile of their receptive fields. Now, these receptive profiles are approximatively Laplacians of Gaussians ΔG . According to the fact that

$$\Delta G \approx I = \Delta(G \approx I)$$

one can conclude that such cells smooth locally the signal at a certain scale and compute its second partial derivatives.

Using such smoothed partial derivatives one can for instance detect qualitative salient *discontinuities*: it is Marr's celebrated *zero-crossing* criterium.

Now, zero-crossings are differential geometrical invariants which belongs to what is called the *2-jet* of the smoothed signal. Jets constitute an *intrinsic* version of the old idea of Taylor expansions. Let us take for instance the Taylor expansion of an \mathbb{R} valued differentiable map $I : \mathbb{R}^2 \rightarrow \mathbb{R}$ up to order 2 at a point (x_0, y_0) .

$$T_2 I(x, y) = I(x_0, y_0) + (x - x_0) \frac{\partial I}{\partial x}(x_0, y_0) + (y - y_0) \frac{\partial I}{\partial y}(x_0, y_0) + \frac{(x - x_0)^2}{2} \frac{\partial^2 I}{\partial x^2}(x_0, y_0) + (x - x_0)(y - y_0) \frac{\partial^2 I}{\partial x \partial y}(x_0, y_0) + \frac{(y - y_0)^2}{2} \frac{\partial^2 I}{\partial y^2}(x_0, y_0).$$

It is constructed via simple algebraic operations from the 8-uple

$$\{(x, y); I(x, y); (\partial_x I, \partial_y I); (\partial_x^2 I, \partial_{xy}^2 I, \partial_y^2 I)\}$$

Now we introduce a space $J^2 = \mathbb{R}^8$ with coordinates

$$\{(x, y); z; (\alpha, \beta); (\lambda, \mu, \gamma)\}$$

and we define the *2-jet* of I as the map $j^2 I : \mathbb{R}^2 \rightarrow J^2$ given by

$$j^2 I(x, y) = \{(x, y); I(x, y); (\partial_x I, \partial_y I); (\partial_x^2 I, \partial_{xy}^2 I, \partial_y^2 I)\}.$$

This map can be defined in a coordinate-free (geometrically intrinsic) manner. Actually J^2 is a *fibration* of base $\mathbb{R}^2_{(x,y)}$ and fiber $\mathbb{R}^6_{(z; (\alpha, \beta); (\lambda, \mu, \gamma))}$ and the 2-jet $j^2 I$ is a *section* of J^2 .

Jets are therefore spatial fields of differential data, and using them one can compute *geometrically relevant differential invariants and morphological features* (fold points, cusps, end points, curvature, etc.) using only *arrays of point-processors*.

If one introduces the hypothesis that the receptive fields of certain classes of cortical cells approximate partial derivatives of Gaussians of order > 2 (up to order 3 or 4) then one can give a *multi-scale* version of this jet calculus. Koenderink has strongly advocated the thesis that such a receptive field jet calculus is implemented in retinotopic arrays of cortical columns.

Let us recall briefly the *columnar structure* of the primary visual cortex (area 17 or area V1) extensively investigated since the pioneering works of Hubel, Wiesel and Mountcastle.

The basic functional module is a "cube" organized in a *retinotopic, columnar and layered* manner.

(i) The *retinotopic structure* preserves the topographic connexions of the retinian ganglion cells (the signal processing of which is transmitted and amplified by the lateral geniculate body). We get that way a *retinotopic fibration*, the base-space of which is constituted by *glued local receptive fields*.

(ii) The *columnar structure* (which is orthogonal to the cortex surface) is essentially constituted by *orientation columns*, columns of *ocular dominance*, and colour plugs. Orientation columns (the diameter of which is about 50μ) are set out orthogonally to the columns of ocular dominance. Their preferential orientation varies from 0° to 180° by 10° steps. They constitute *hypercolumns* the size of which is about 800μ -1mm.

(iii) The *layered structure* is composed of six layers. Its depth is about 1,8mm. The geniculate fibers project on layer IV. Layer VI is a feedback one. Layer V projects on the

colliculus. Layers II and III receive the axons of layer IV and project their efferent fibers in other cortical regions where different attributes (shape, colour, movement, stereopsis, etc.) are further processed in a *non* retinotopic way.

The orientation columns yield a beautiful example of a *neurally implemented fibration*. Actually, they implement the fibration $\pi: E \rightarrow W$ the base of which is the visual field and the fiber of which is the *projective line* $F = \mathbb{P}^1$ of plane directions. William Hoffman (1989) has explained how the cortex implements what is called the *contact structure* of E .

Let M be a n -manifold. A *contact element* $c = (x, H)$ of M is constituted by a point $x \in M$ and by a $(n-1)$ -hyperplane H of the tangent space $T_x M$ of M at x . (See figure 5).

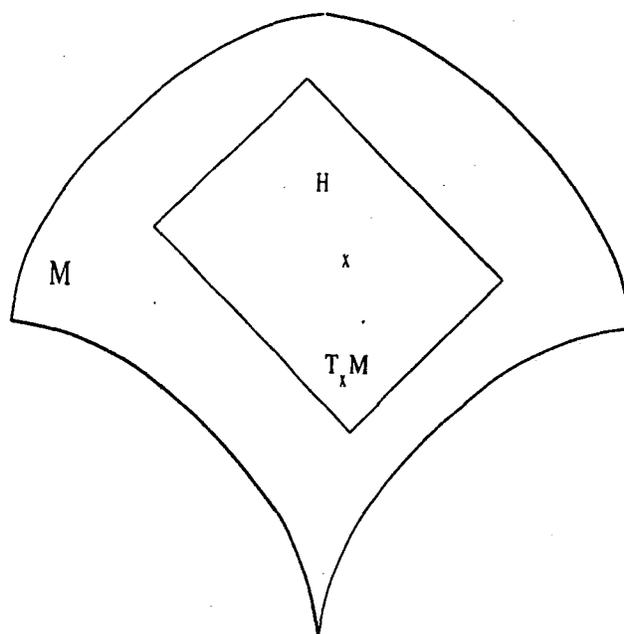


Figure 5

The *contact bundle* CM of M is the $(2n-1)$ -manifold of these contact elements. CM is a fibration — a *fiber bundle* — $\pi: CM \rightarrow M$, the fiber over x being the projective space $C_x M = \mathbb{P}(T_x M)$ of hyperplanes of the vector space $T_x M$. Here we have $n = 2$, the H s are the directions and therefore $C_x M \cong \mathbb{P}^1$. (See figure 6: M is represented as one-dimensional and in the second figure the directions are represented as points of \mathbb{P}^1).

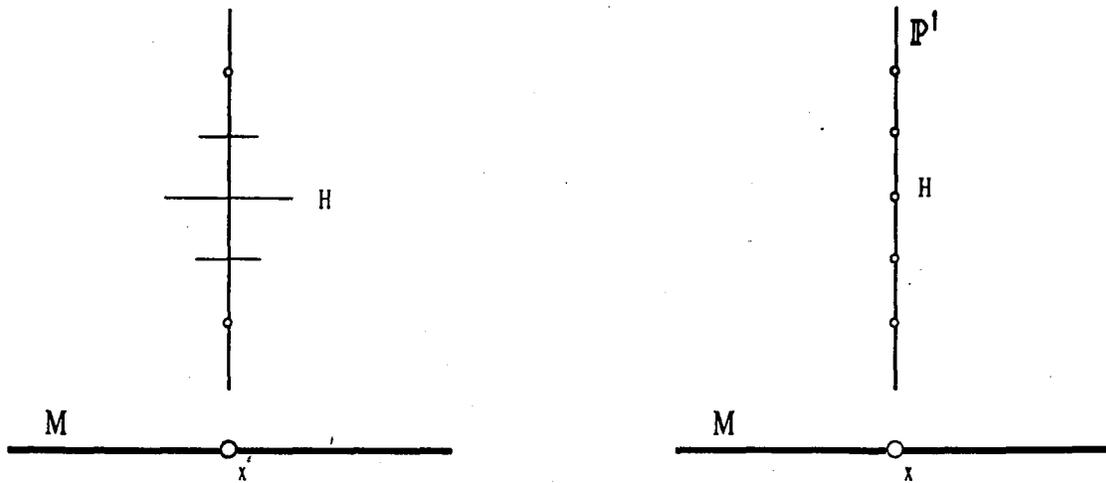


Figure 6

The contact bundle CM is isomorphic to the *projectivized* bundle of the *cotangent bundle* T^*M . Indeed a contact element (x, H) can be defined by a 1-form (a linear form) α_x on T_xM satisfying $\text{Ker}(\alpha_x) = H$. But $\lambda\alpha_x$ ($\lambda \neq 0$) defines the same H and therefore $CM \cong P(T^*M)$.

There exists on the $(2n-1)$ -contact bundle CM what is called a *canonical contact structure*. It reflects the *canonical symplectic structure* of T^*M . In general a *contact structure* on an N -manifold m is defined by a *non degenerate* field of $(N-1)$ -hyperplanes h_c which are called *contact hyperplanes* (non degeneracy forces N to be odd). Let $c = (x, H) \in m = CM$. A tangent vector

$$(\xi, \Theta) \in T_{c,m} = T_{(x,H)}CM$$

belongs to the contact $(2n-2)$ -hyperplane h_c iff its projection ξ on T_xM belongs to H (i.e. $c = (x, H)$ moves in such a way that the x -velocity ξ belongs to H). (See figure 7).

Let Γ be a sub-manifold of M and let $C_M\Gamma$ be the manifold of the $c \in CM$ tangent to Γ . $C_M\Gamma$ is always of dimension $n-1$ and is tangent to the field h_c . $C_M\Gamma$ is therefore an *(n-1)-integral manifold* of the canonical contact structure of CM . In the visual case $n = 2$. $m = CM$ is of $\text{dim} = 3$ (coordinates $x, y, \varphi = \text{direction of } H$). If $c = (x, H) \in CM$ and $(\xi, \Theta) \in T_{(x,H)}CM$, we have:

$$(\xi, \Theta) \in h_c \Leftrightarrow \xi \in H.$$

If Γ is a *curve* in M and if we consider the contact elements $c = (x, H)$ which are *tangent* to Γ (i.e. $T_x\Gamma = H$) we get a *curve* $C\Gamma$ (a $(n-1)$ -submanifold) in CM which is an *integral curve* of the canonical contact structure. Indeed, when x moves along Γ , (ξ, Θ) is such that $\xi \in T_x\Gamma$ and $T_x\Gamma = H$. Therefore $\xi \in H$ and therefore $(\xi, \Theta) \in h_c$. This shows that a boundary curve (a contour curve) is a section of a fibration and that it can therefore be processed by arrays of points processors.

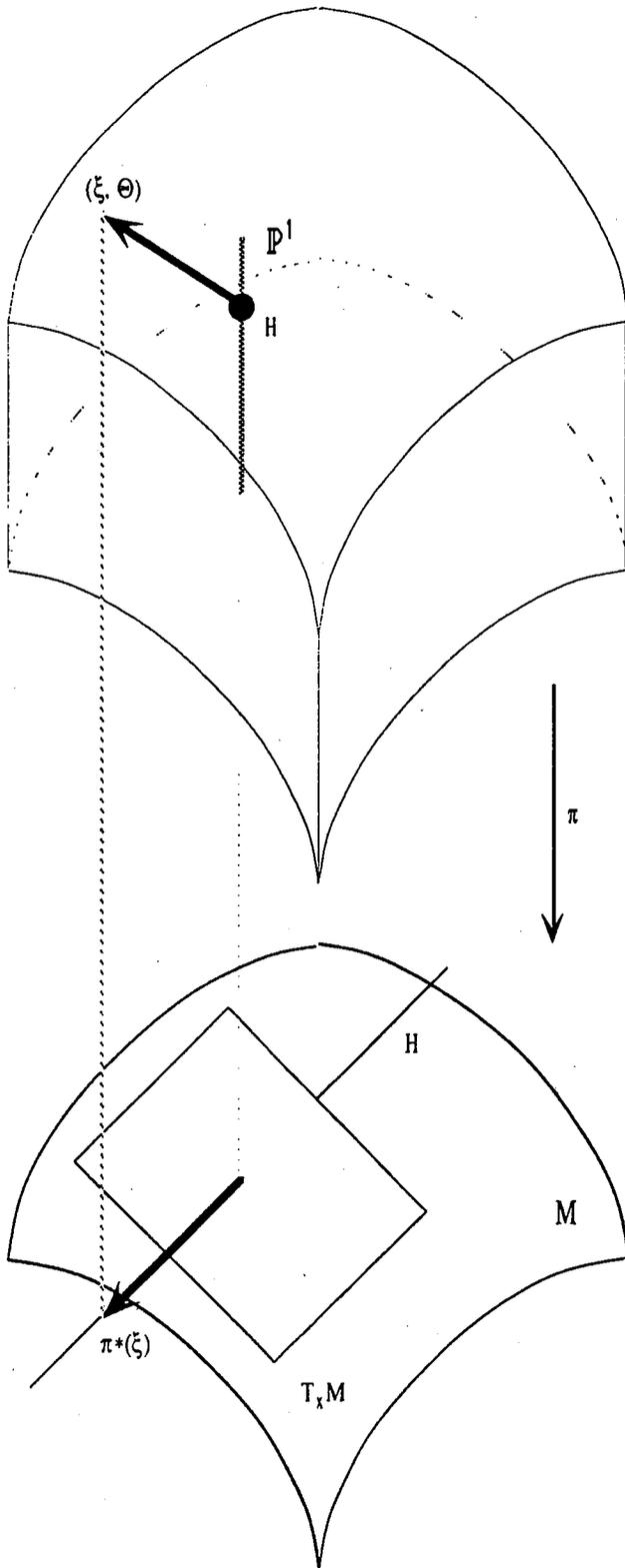


Figure 7

W. Hoffman explains very well how:

- (i) the concept of receptive field corresponds to the concept of local chart in differential geometry;
- (ii) the neurally implemented canonical contact structure explains the processing of visual contours as *integral curves*;
- (iii) the perceptual constancies are invariant structures describable by means of actions of symmetry Lie groups on the contact structure of the projective bundle of directions.

He concludes claiming that

“fibrations (...) are certainly present and operative in the posterior perceptual system if one takes account of the presence of “orientation” micro-response fields and the columnar arrangement of cortex”(p. 645).

For colours, the situation is a little bit more complex. There exist parvocellular layers in the lateral geniculate body constituted of ganglion cells sharing a *spectral antagonism* : (R+/G-), (G+/R-), (Y+/B-), (B+.Y-); there exist cortical plugs specialized in colour processing; there exist colour *blobs* in the area V1; and, according to Semir Zeki, the area V4 is specialized in colour processing. But it seems nevertheless plausible to hypothesize that colour is also processed through fibrations (with a dynamical antagonist spectral structure in the fibers).

This shed some light on the *morphological aspects* of perceptive statements.

Quid, now, of the logical ones ? How the morphological eidetics can be unified with a logic of judgement.

IV. HUSSERL'S *ERFAHRUNG UND URTEIL*

In *Erfahrung und Urteil*, *Untersuchungen zur Genealogie der Logik*, Husserl tries to clarify phenomenologically the *origin* of predication. He elaborates a theory of predicative statements such as “*S is p*”, in the framework of formal apophantics (a syntactic theory) and formal ontology (a semantic theory such as set theory). But his perspective is “genealogic”. According to him, classical logic *conceals* the fundamental problem of *evidence* in the logical concept of truth. By “evidence”, Husserl means here *perceptive presentation as immediate acquaintance* with the world. Predicative statements such as “*S is p*” are rooted in ante-predicative and pre-judicative experience and it is the nature of such a founding relation which constitutes for Husserl the main problem. His “basic theme” is “categorical judgement founded in perception” (p. 70). Now, according to him, formal logic neglects perceptual evidence.

“The formal character of logical Analytics consists in the fact that it does not consider the material quality of what is given prior to statements, and that it looks to substrates only in what concerns the categorial form they take in the statement” (p. 18).

He wants therefore to show that logic is originally based on the experience of the world, that the foundation of logic in perception has a universal scope and concerns its essence.

In *Erfahrung und Urteil* Husserl summarizes his analysis of perception previously done in the third *Logical Investigation, Ding und Raum* and *Ideen I*.

(i) The morphological analysis of the filling-in of spatial regions by qualities and the constitutive role of *qualitative discontinuities* in segmenting visual scenes (see the opposition between *Verschmelzung* and *Sonderung*).

(ii) The essentially *adumbrative* status of perception (*Abschattungslehre*). The object is a noematic pole unifying in a coherent way a continuous flux of adumbrations. Its noematic content expresses the *synthetic* relations of foundation between spatial extensions and their dependent moments such as contours or colours. It rule governs perceptive *anticipations* which, being filled-in by a "fluent variability" of adumbrations, are "indetermined and general", but which are nevertheless also *prescriptively ruled by typed relations*. For instance, there exists a "typical genericity" of the filling-in relations between extensions and qualities.

He analyzes then very precisely the origin of the "primitive logical categories" from the ante-predicative synthesis. The main problem is to understand the "categorical genesis" of the *substrate/determination categories* which transforms the *perceptive synthesis*, where a substrate *S* possesses a dependent moment *p*, in the *predication* "*S* is *p*". Husserl's main thesis is that this categorial transformation arises from a reflective thematization which enables consciousness to grasp the dependent moments and typifies the substrates in "subjects" and the dependent moments in "predicates". According to him:

"We find there the origin of the first "logical" categories" (p. 127).

Husserl points out therefore a fundamental conversion from the *synthetic* perceptive unity unifying an extension *S* with its dependent moments *p* to the *analytic* syntactic unity of the statement "*S* is *p*".

"In the most simple predicative judgement a *double information* is processed" (p. 247).

Indeed, underlying the syntactic "subject / predicate" information concerning the "functional forms" of the terms of the proposition, there exists another information concerning the "kernel forms" "substrate=independence" and "adjectivity=dependence". This underlying information is presupposed by the syntactic one. Predication is a process based on

"the covering of the kernel forms as syntactic material by the functional forms" (p. 248).

It is this logical typification of the synthetic dependance relations in syntactic categories I want now to model. As far as the dependence relations between extensions

and filling-in qualities can be schematized by sections of fibrations, we see that, if we want to model Husserl's "genealogical" analysis of logic, we need to understand *how variables which denote sections of fibrations can be syntactically typified in such a way that the resulting semantics could be that sort of Kripkean semantics we need to explain the "indexical" status of perceptive statements*. Moreover this must be done in the framework of formal apophantics and formal ontology.

Now, we will see that it is possible to achieve this goal using the tools of the *geometrical logic* yielded by *Topos theory*.

V. SHEAVES, TOPOI AND LOGIC

Some remarks to begin with.

1. Topos theory has been invented by Grothendieck for solving very sophisticated problems of algebraic geometry (he wanted to construct generalized cohomological theories). It can therefore seem strange to use it for solving elementary and non mathematical problems. But actually, when they are treated as *neuro-cognitive* ones, these "naïve" problems are not at all elementary and require sophisticated tools to be modeled adequately.

2. Category theory is certainly the best formal ontology we have now at disposal.

3. Some specialists of topos theory have already applied these tools to clarify some basic semantic problems. For instance Gonzalo Reyes (who works, with Eduardo Dubuc, Anders Koch, Ieke Moerdijk and Marta Bunge on the applications of topos theory to *Synthetic Differential Geometry*) has recently used these techniques to formalize the Kripkean theory of proper names as rigid designators.

4. In general, topos theory is used in formal logic as a tool for *typed λ -calculus* (see e.g. the works of John Mitchell, Philip Scott, Giuseppe Longo, etc.). As we will see, when variables are typed by objects in a topos, the categorical properties of the topos (to be a cartesian closed category, to possess a subobject classifier, etc.) leads to an intuitionistic "internal logic" which is a typed λ -calculus. As far as the Curry-Howard correspondance shows that (in intuitionistic logic) formulae can be treated as types, proofs as λ -terms, and the reduction of a proof by cut-elimination as the reduction of a λ -term to its normal form, topos theory has become fundamental for understanding the semantic of formal — and in particular of programming — languages.

In these applications, the geometrical origin of the sheaf and topos concepts is concealed. Here, we aim at a completely different sort of application. We want to use this geometrical origin to explain the symbolic ascent of logical forms from morphological ones.

As we already saw, in the "receptive field" perspective one needs to conceive of sections of fibrations as resulting from a glueing of local sections. This leads naturally to

the concept of *sheaf* which, even if it can seem a very abstract one, is actually very deeply rooted in spatial intuition.

5.1. The concept of sheaf.

For the concept of sheaf, the primitive topological concept is no longer that of a point but that of an open set (what Peter Johnstone calls “pointless topology”).

At an abstract level, a fibration is characterized by the sets of its sections $\Gamma(U)$ over the open sets $U \subset M$. If $s \in \Gamma(U)$ is a section over U and if $V \subset U$, we can consider the *restriction* $s|_V$ of s to V . The restriction is a map $\Gamma(U) \rightarrow \Gamma(V)$. It is clear that if $V=U$, then $s|_V=s$ and that if $W \subset V \subset U$ and $s \in \Gamma(U)$, then $(s|_V)|_W = s|_W$ (transitivity of the restriction). We get therefore what is called a *contravariant functor* $\Gamma : \mathfrak{o}^*(M) \rightarrow \mathbf{Sets}$ from the category $\mathfrak{o}(M)$ of the open sets of M in the category \mathbf{Sets} of sets. (The objects of $\mathfrak{o}(M)$ are the open sets of M , and its morphisms are the inclusions of open sets.)

Conversely, let Γ be such a functor — what is called a *presheaf* on M . To have a chance of being the functor of the sections of a fibration, Γ must clearly satisfy the two following axioms.

(S₁) Two sections which are locally equal must be globally equal. Let $u = (U_i)_{i \in I}$ be an open covering of M . Let $s, s' \in \Gamma(M)$. If $s|_{U_i} = s'|_{U_i} \forall i \in I$, then $s = s'$.

(S₂) Compatible local sections can be collated in a global one. Let $s_i \in \Gamma(U_i)$ be a family over $u = (U_i)_{i \in I}$. If the s_i are compatible, that is if $s_i|_{U_i \cap U_j} = s_j|_{U_i \cap U_j}$ when $U_i \cap U_j \neq \emptyset$, then they can be glued together : $\exists s \in \Gamma(M)$ such that $s|_{U_i} = s_i \forall i \in I$.

(S₁) and (S₂) can be expressed in a purely categorical manner. For instance (S₂) says that the arrow

$$e : s \rightarrow \{s|_{U_i}\}_{i \in I}$$

is the *equalizer*

$$\Gamma(U) \xrightarrow{e} \prod_i \Gamma(U_i) \begin{array}{c} \xrightarrow{p} \\ \xrightarrow{q} \end{array} \prod_{i,j} \Gamma(U_i \cap U_j)$$

of the two projections p, q corresponding to the inclusions $U_i \cap U_j \subset U_i$ and $U_i \cap U_j \subset U_j$.

In fact these axioms characterize a more general structure — and even more pervasive in contemporary mathematics — than the structure of fibration, namely the structure of *sheaf*. It can be shown that if the axioms (S₁) and (S₂) are satisfied, then one can represent the functor Γ by a general fibered structure $\pi : E \rightarrow M$ (called an “*étale*” space and which is not necessarily a locally trivial fibration) in such a way that $\Gamma(U)$ becomes the set of sections of π over U . In a nutshell, the fiber E_x — called in that case the *stalk* of the sheaf Γ at x — is the inductive limit (the colimit) :

$$E_x = \lim_{V \subset U \in \mathfrak{u}_x} \{(\Gamma(U), \Gamma(V \subset U))\}$$

(where \mathfrak{u}_x is the filter of the open neighborhoods of x). The stalk E_x is the set of *germs* s_x of sections at x . E is the sum of the E_x . If $s \in \Gamma(U)$, it can be interpreted as the map

$x \in U \rightarrow s(x) \in E_x$. The topology of E is then defined as the finest one making all these sections continuous.

5.2. Generalizations.

There are many generalizations of this basic situation. The best known is that of Grothendieck's topologies : open coverings are defined in terms of "sieves" and the concept of sheaf is generalized to this structure of "site".

Another generalization is that of *frames* and *locales*. One consider frames, that is lattices which share the properties of the lattices of open sets $\mathfrak{o}(X)$. They are complete distributive lattices with finite meets and general joins. Locales are the objects of the dual category (in the case of topological spaces the correspondance is

$$f : X \rightarrow Y \text{ continuous} \Rightarrow f^* : \mathfrak{o}(Y) \rightarrow \mathfrak{o}(X).$$

Points are then defined as morphisms $p : A \rightarrow 2 = \mathfrak{o}(1)$ (in the case of topological spaces, if $A = \mathfrak{o}(X)$ points correspond to true points $p : 1 \rightarrow X$). Let $Pt(A)$ be the set of points of A . A topology is defined on $Pt(A)$ by the subsets $\varphi(a) = \{p \in Pt(A) \mid p(a) = 1\}$. $\mathfrak{o}(Pt(A))$ is the best approximation of A by a "spatial" locale.

5.3. The concept of topos.

Now, the main point, is that the category of sheaves on a space M shares the categorical properties of a topos and that (intuitionistic) logical languages can therefore be *interpreted* in it. According to Lawvere's perspective, such a topos t can therefore be conceived of as a universe of discourse *the objects of which are variable entities depending on the elements U of $\mathfrak{o}(M)$* . $\mathfrak{o}(M)$ operates therefore as a set of "possible worlds". The "possible worlds" are here *spatial extensions*. This dependence on spatial extensions shares the two main properties we need:

(i) it is constitutive of the concept of logical truth and therefore of semantic relevance;

(ii) but it is not directly "visible" in the syntax of the internal logic of t and it is therefore syntactically irrelevant.

Actually, as long as the formal constructions in such a universe are constructive (in the intuitionistic sense) one can completely bracket and forget the spatial dependence at the syntactic level. We get therefore a very elegant explanation of the puzzling fact that spatial extension operates in perceptive statements in a rather "indexical" and "pragmatic" manner.

Ieke Moerdijk and Gonzalo Reyes (1991) have proposed an interesting philosophical commentary to Lawvere's main ideas.

"Topos theory has brought to light and given the means to exploit a complementarity (or duality) principle between *logic* and *structure*" (p. 10).

In the framework of classical model theory, a theory T is of the general form $T = S + I$ where S is the axiomatics of a type of structures and I a set-theoretical interpretation satisfying the "trivial" and "tautological" Tarskian semantics. But one can also interpret S

in a topos $\mathbf{t} = \mathbf{Sh}(c)$ by means of a *sheaf semantics*. One can then work in \mathbf{t} as one would work in a standard universe of sets provided one does it with a non classical (intuitionistic) logic reflecting the new interpretation.

Let us now recall briefly the structure of a topos. If A is a sheaf on M , $A(U)$ will denote the set $\Gamma_A(U)$ of sections of A over U . It is easy to show that the sheaves on a base space M constitute a category $\mathbf{Sh}(M)$.

5.3.1. Exponentials

$\mathbf{Sh}(M)$ is a *cartesian closed* category. This means that it has products and fibered products or pullbacks, a terminal object — classically denoted by 1 — and *exponentials* B^A . An exponential object is an object which “internalizes” in the objects of $\mathbf{Sh}(M)$ the morphisms $f: A \rightarrow B$. Such “internalization” of functorial structures are called *representable* functors. Technically, the functor $(\bullet)^A$ is the *right adjoint* of the functor $A \times (\bullet)$. This means that we have for every object C of $\mathbf{Sh}(M)$ a functorial isomorphism $\text{Hom}(C, B^A) \cong \text{Hom}(A \times C, B)$. E.g. for $C=1$, we get $\text{Hom}(1, B^A) \cong \text{Hom}(A, B)$. But an arrow $f: 1 \rightarrow B^A$ is like an “element” of B^A . In fact, if A is a sheaf, an arrow $s: 1 \rightarrow A$ is a *global section* of A , that is an element $s \in A(M)$.

If we take $C=B^A$ and Id_{B^A} , the right adjunction defines what is called a *counit* $\varepsilon: A \times B^A \rightarrow B$ such that for every $f: C \rightarrow B^A$ the associated $h: A \times C \rightarrow B$ is given by $h = \varepsilon \circ (1 \times f)$. The counit generalizes the map $(x, f) \rightarrow f(x)$ in set theory and is therefore called the *evaluation map*.

The sheaf B^A is defined using the evident restrictions $A|_U$ and $B|_U$ to open sets: $B^A(U) = \text{Hom}(A|_U, B|_U)$. It is called the “internal Hom” or the sheaf of germs of morphisms from A to B .

5.3.2. Subobject classifier

$\mathbf{Sh}(M)$ possesses also what is called a *subobject classifier* Ω , that is an object which “internalizes” the sets of subobjects, making the subobject functor representable. A subobject $m: S \rightarrow 0$. A is a monomorphism (an injective map in the case of the category of sets). This means that if f and g are two morphisms from an object R to S , then $m \circ f = m \circ g$ implies $f = g$. It is equivalent to say that the fibered product $S \times_A S$ defined by m is isomorphic with S . A subobject classifier is a monomorphism $\text{True}: 1 \rightarrow \Omega$ such that every subobject $m: S \rightarrow 0$. A can be retrieved from True by a pull-back:

$$\begin{array}{ccc} S & \rightarrow & 1 \\ \downarrow & & \downarrow \text{True} \\ A & \xrightarrow{\varphi_S} & \Omega \end{array}$$

We get therefore a functorial isomorphism $\text{Sub}(A) \cong \text{Hom}(A, \Omega)$. φ_S is called the *characteristic map* of the subobject S .

In the category **Sets** of sets, $\Omega = \{0, 1\}$ is the classical set of boolean truth-values. Here — and it is perhaps the main difference between a topos like $\mathbf{Sh}(M)$ and

the classical topos **Sets** —, $\Omega(U)$ depends essentially on the topological structure. It expresses the localization of truth in a sheaf topos.

By definition $\Omega(U) := \{W \subset U\}$. It is trivial to verify that Ω is a sheaf. The True map $\text{True} : 1 \rightarrow \Omega$ is defined by $\text{True}(U) : 1 \rightarrow U \in \Omega(U)$ that is by the *maximal* element of $\Omega(U)$: to be true over U is to be true “everywhere” over U . The global section $T \in \Omega(M)$ selected by True is nothing else than the whole base space M itself and $\text{True}(U)$ is its localization to U . If S is a subsheaf of the sheaf A , its characteristic map $\varphi_S : A \rightarrow \Omega$ is given by the maps $\varphi_S(U) : A(U) \rightarrow \Omega(U)$ which map $s \in A(U)$ to the largest $W \subset U$ s.t. $s|_W \in S$. It is easy to verify that the monic map $S \rightarrow A$ is effectively the pull-back of True by φ_S . As Mac Lane says: φ_S gives “the shortest path to truth” $W \subset U$ for $S \rightarrow A$.

5.3.3. Elements, properties and parts

In a topos, the morphisms $a : B \rightarrow A$ are called *generalized elements* of A , or elements *defined on* B (this denomination comes from algebraic geometric and, more precisely, from Grothendieck’s theory of schemes). Among the elements, the most important are those *defined on open sets* U , that is precisely the sections $s \in A(U)$. If $U \in \mathbf{o}(M)$, we can consider the Yoneda sheaf $y(U) = \{V \subset U\}$ and verify that $A(U) = \text{Hom}(y(U), A)$. The elements defined on the terminal object 1 are “global”. We will see that an arrow $\theta : A \rightarrow \Omega$ is a “predicate” for A , that is a “property” of its generalized elements. Among all predicates, there is the predicate $\text{True}_A : A \rightarrow 1 \rightarrow \Omega$. It is easy to verify that an element $a : B \rightarrow A$ factorizes through a subobject $S \rightarrow A$ iff $\text{char}_S(a) = \varphi_S \circ a = \text{True}_B$. φ_S is therefore the predicate of A which is true exactly for those elements of A which are in S . The *unicity* of φ_S expresses the *extensionality principle*.

Using the exponentials and the subobject classifier we can define the *parts* of an object A as another object $P(A) = \Omega^A$. We get the functorial isomorphisms:

$$\text{Sub}(A) = \text{Hom}(A, \Omega) = \text{Hom}(A \times 1, \Omega) = \text{Hom}(1, \Omega^A) = \text{Hom}(1, P(A)).$$

We have therefore $\Omega = P(1)$.

This shows that there are 3 equivalent descriptions of a subobject $m : S \rightarrow A$.

- (i) its “extension” S : we will see that it can be symbolized as in **Sets** by $\{a | \varphi_S(a)\}$;
- (ii) its characteristic map $\varphi_S : A \rightarrow \Omega$ which is a “predicate” of A ;
- (iii) the global section $s : 1 \rightarrow P(A)$ which is its “name”.

The evaluation map $\varepsilon_A : A \times \Omega^A = A \times P(A) \rightarrow \Omega$ is a “membership” predicate: if $a : B \rightarrow A$ is an element of A , and if $s : 1 \rightarrow P(A)$ is (the name of) a subobject of A , then $\varepsilon_A \circ (a \times s) = \text{True}_{B \times 1}$ iff $\varphi_S \circ a = \text{True}_B$, that is iff a is an element of S .

5.3.4. Towards logic

The existence of an intuitionistic “internal logic” in a topos depends essentially on the fact that, being the set of the open sets W of the topological space U , $\Omega(U)$ is (functorially) an *Heyting algebra*. Ω is therefore a sheaf of Heyting algebras (an Heyting

algebra object in $\mathbf{Sh}(M)$). The consequence is that the “external” set of subobjects $\text{Sub}(A)$ and the “internal” one $P(A)$ are also Heyting algebras, the canonical isomorphism $\text{Sub}(A) \cong \text{Hom}(1, P(A))$ being an isomorphism of Heyting algebras.

5.4. Topoi and logic

Now, the central fact is that a topos is exactly the categorical structure which is needed for doing logic. *But this logic is spatially localized.* For details see e.g. Saunders Mac Lane, Ieke Moerdijk 1992.

5.4.1. Types and localization

We can associate with each topos $\mathbf{Sh}(M)$ a formal language l_M called its Mitchell-Bénabou language, and a forcing semantics called its Kripke-Joyal semantics. The crucial point is that a sheaf X can be considered as a *type* for variables x which are interpreted as sections $s \in X(U)$ of X . *We get therefore at the same time a typification and a localization of the variables.* This achievement fits perfectly well with Husserl’s description and explains Wittgenstein’s remark. It provides them with a correct mathematical status.

(i) Sections are tokens denoted by variables belonging to types (species, essences). They are “concretely” particularized by the specification of their localization U and by their specific values. But as an element of type X , s particularizes an abstract unilateral relation of dependence, the relation “quality \rightarrow extension” which is constitutive of X .

(ii) The relations between particular sections $s \in X(U)$, $t \in Y(V)$ are *external*. The relations between X and Y are *internal*. Nevertheless the linguistic expressions which express them are formulas in the formal language l_M associated to $\mathbf{Sh}(M)$, and are *the same*.

5.4.2. Syntax

How are the terms and the formulas of l_M syntactically constructed? Here is a summary of their inductive construction.

A term σ of type X constructed using variables y, z of respective types Y, Z has a *source* $Y \times Z$ and is interpreted by a morphism $\sigma : Y \times Z \rightarrow X$ which expresses its structure.

(i) To each $X \in \mathbf{Sh}(M)$ considered as a type are associated variables x, x', \dots . They are interpreted by the identity map $1_X : X \rightarrow X$.

(ii) Terms $\sigma : U \rightarrow X$, $\tau : V \rightarrow Y$ of respective types X and Y yield a term $\langle \sigma, \tau \rangle$ interpreted by $\langle \sigma, \tau \rangle : W = U \times V \rightarrow X \times Y$ of type $X \times Y$.

(iii) Terms $\sigma : U \rightarrow X$, $\tau : V \rightarrow X$ of the same type X yield the term $\sigma = \tau$ of type Ω interpreted by

$$(\sigma = \tau) : W = U \times V \rightarrow X \times X \xrightarrow{\delta_X} \Omega$$

where δ_X is the characteristic function of the diagonal subobject $\Delta : X \rightarrow X \times X$.

(iv) A term $\sigma : U \rightarrow X$ of type X and a morphism $f : X \rightarrow Y$ yield by composition a term $f \circ \sigma$ of type Y .

(v) Terms $\theta : V \rightarrow Y^X$ and $\sigma : U \rightarrow X$ of respective types Y^X and X yield a term $\theta(\sigma)$ of type Y interpreted by

$$\theta(\sigma) : W = V \times U \rightarrow Y^X \times X \xrightarrow{e} Y$$

where e is the evaluation map.

(vi) In particular, terms $\sigma : U \rightarrow X$ and $\tau : V \rightarrow \Omega^X$ yield a term $\sigma \in \tau$ of type Ω interpreted by

$$\sigma \in \tau : W = V \times U \rightarrow X \times \Omega^X \xrightarrow{e} \Omega.$$

(vii) A variable x of type X and a term $\sigma : X \times U \rightarrow Z$ of type Z and of source $X \times U$ yield a term of type Z^X interpreted by $\lambda x \sigma : U \rightarrow Z^X$.

(viii) Ω is the type of the formulae of I_M . As it is an Heyting algebra, we get the logical operations of propositional calculus : $\phi \wedge \psi$, $\phi \vee \psi$, $\phi \Rightarrow \psi$, $\neg \phi$. It is easy to verify that if $\phi(x,y) : X \times Y \rightarrow \Omega$ is a formula, we can write the subobject of $X \times Y$ classified by its interpretation in the "set theoretic" manner : $\{(x,y) \in X \times Y \mid \phi(x,y)\}$.

(ix) One of the most remarkable facts of topoi theory is that it is possible to define *quantification* in a purely categorical manner. Let $f : A \rightarrow B$ be a morphism of $\mathbf{Sh}(M)$ and consider the "inverse image" functor $f^* : \text{Sub}(B) \rightarrow \text{Sub}(A)$ defined by composition with f . Its internal version is the morphism $P(f) : P(B) \rightarrow P(A)$. The fact is that $P(f)$ has *two adjoint functors* : a left adjoint one $\exists_f : P(A) \rightarrow P(B)$ and a right adjoint one $\forall_f : P(A) \rightarrow P(B)$. They generalize the two adjunctions in Sets. If $f : A \rightarrow B$, $S \subset A$ and $T \subset B$,

$$\exists_f(S) = \{b \in B \mid \exists a \in S (f(a) = b)\} = f(S),$$

$$\forall_f(S) = \{b \in B \mid \forall a \in A (f(a) = b \Rightarrow a \in S)\}$$

$$\{b \in B \mid f^{-1}(b) \subset S\},$$

$$\text{and } \begin{cases} f^*(T) \subset S \Leftrightarrow T \subset \forall_f(S) \\ S \subset f^*(T) \Leftrightarrow \exists_f(S) \subset T \end{cases}$$

We have:

$$S \in \exists_f(S)(U) \text{ iff } \{V \mid \exists t \in S(V) f(t) = s|_V\} \text{ covers } U,$$

$$S \in \forall_f(S)(U) \text{ iff } \forall V \subseteq U f^{-1}(s|_V) \subseteq S(V).$$

Now, let $\phi(x,y) : X \times Y \rightarrow \Omega$ be a formula of two variables in $\mathbf{Sh}(M)$. Let $p : X \rightarrow 1$ the canonical projection and $P(p) : P(1) \rightarrow P(X)$. We get the adjunctions:

$$\Omega^X = P(X) \begin{array}{c} \xrightarrow{\forall_p} \\ \xleftarrow{P(p)} \\ \xrightarrow{\exists_p} \end{array} P(1) = \Omega.$$

We consider $\lambda x \phi(x,y) : Y \rightarrow \Omega^X = P(X)$ and we get the formula of source Y :

$$\forall x \phi(x,y) = \forall_p \circ \lambda x \phi(x,y) : Y \rightarrow \Omega.$$

These categorical constructs show that the formal language l_M is, at the “linguistic” level, exactly of the same nature as the classical formal language of sets theory. The main difference is that we have introduced a subtle dialectics *between the type of the variables and their localization*.

5.4.3. Semantics

The Kripke-Joyal semantics of topoi is a *forcing* semantics generalizing Cohen’s one. A variable x of type X denotes a *section* $s \in X(U)$, that is a morphism $s : U \rightarrow X$ where U is now the sheaf defined by U . The semantic rules are rules $U \Vdash \varphi(s)$ (U forces $\varphi(s)$). Let $s : U \rightarrow X$ and let $\text{Im}(s) \in \text{Sub}(X)$ be the image of s . One defines

$$U \Vdash \varphi(s) := \text{Im}(s) \subseteq \{x \mid \varphi(x)\},$$

that is $U \Vdash \varphi(s)$ iff $U \xrightarrow{s} X \xrightarrow{\varphi} \Omega$ factorizes through $\{x \mid \varphi(x)\}$:

$$U \Rightarrow \{x \mid \varphi(x)\} \rightarrow 1 \xrightarrow{\text{True}} \Omega.$$

The semantic rules are:

- (i) $U \Vdash \varphi(s) \wedge \psi(s)$ iff $U \Vdash \varphi(s)$ and $U \Vdash \psi(s)$.
- (ii) $U \Vdash \varphi(s) \vee \psi(s)$ iff there exists an open covering $(U_i)_{i \in I}$ of U s.t. for every i , $U_i \Vdash \varphi(s|_{U_i})$ or $U_i \Vdash \psi(s|_{U_i})$ (intuitionistic rule for disjunction)
- (iii) $U \Vdash \varphi(s) \Rightarrow \psi(s)$ iff, for all $V \subseteq U$, $V \Vdash \varphi(s|_V)$ implies $V \Vdash \psi(s|_V)$.
- (iv) $U \Vdash \neg \varphi(s)$ iff there does not exist $V \subseteq U$, $V \neq \emptyset$ s.t. $V \Vdash \varphi(s|_V)$ (the negation is intuitionistic because Ω is a Heyting algebra and not a Boolean one).
- (v) $U \Vdash \exists y \varphi(s, y)$ (y being of type Y) iff there exist an open covering $(U_i)_{i \in I}$ of U and sections $\beta_i \in Y(U_i)$ s.t. for every $i \in I$ $U_i \Vdash \varphi(s|_{U_i}, \beta_i)$.
- (vi) $U \Vdash \forall y \varphi(s, y)$ iff for every $V \subseteq U$ and $\beta \in Y(V)$ we have $V \Vdash \varphi(s|_V, \beta)$.

The subtilities of this form of semantics result from the following problem. In a topos, one can interpret set-like expressions such as

$$\left\{ (x_i)_{i \in I} \in \prod_{i \in I} F_i \mid \varphi(x_i) \right\}$$

defined by $(a_i) \in \{(x_i) \mid \varphi(x_i)\}(U)$ iff $U \Vdash \varphi(a_i)$. But one has to be sure that such expressions define *subsheaves*. For this, one has to “sheafify” the sub-functors appearing in the set-like constructions.

5.5. Sheaf mereology and boundaries.

Two last points to conclude.

1. There is a natural *mereology* applying to sections of fibrations and sheaves. It satisfies all the axioms of mereo-topology (Smith 1993) except the one asserting the *unicity* of the universe of objects under consideration. Actually, the union of sections is a merging, a fusion, a glueing of compatible sections. The union of the sections s satisfying a predicate φ is therefore in general not a single section but only a set of *maximal* ones. (See Petitot 1994)

2. In perceptive statements, *boundaries* play a fundamental role. And one knows since Brentano that boundaries are somehow paradoxical entities. To tackle this point, we can use Lawvere's idea of *co-Heyting* algebras. In a Heyting algebra of open sets, the negation $\neg U$ of U is the *interior* of its complementary set, that is the largest open set V such that $U \cap V = \emptyset$. In a co-Heyting algebra of closed sets, the negation $\neg F$ of F is dually the smallest element H such that $F \cup H = 1$ (that is the closure of its complementary open set). One has

$$\neg(F \cap H) = \neg F \cup \neg H, \text{ but only } \neg(F \cup H) \subseteq \neg F \cap \neg G.$$

One defines then the *boundary* ∂F of F as the intersection $F \cap \neg F = \partial F$. ∂F is therefore defined by logical "contradiction". The boundary operator satisfies the Leibniz rule :

$$\partial(F \cap H) = (\partial F \cap H) \cup (F \cap \partial H).$$

Boundaries are characterized by $\partial B = B$ that is by $\neg B = 1$ or $\neg \neg B = 0$. In general, the double negation $\neg \neg F \subseteq F$ is the "regular core" of F (the closure of its interior). One has of course $F = (\neg \neg F) \cup \partial F$.

CONCLUSION

1. There exists a topological geometrical eidetics of the morphological structures of perception. This eidetics is neurologically relevant.
2. There exists an indexical status of spatial extensions in perceptive statements. This fact requires some sort of Kripkean modal semantics.
3. It is relevant to try to model Husserl's ante-predicative "genealogy" of logic.
4. The geometric logic yielded by topos theory can do the job. It has therefore a deep cognitive meaning.

BIBLIOGRAPHY

- ASPERTI, A., LONGO, G., 1991. *Categories, Types, and Structures*, Cambridge, MIT Press.
- HOFFMAN, W., 1989. "The visual cortex is a contact bundle", *Applied Mathematics and Computation*, 32.
- HUSSERL, E., 1954. *Erfahrung und Urteil, Untersuchungen zur Genealogie der Logik.*, Hamburg, Claassen&Goverts.
- MAC LANE, S., MOERDIJK, I., 1992. *Sheaves in Geometry and Logic*, New York, Springer.
- MOERDIJK, I., REYES, G., 1991. *Models for Smooth Infinitesimal Analysis*, Berlin, Springer.

- MULLIGAN, K., 1992. "Internal Relations", *Australian National University Metaphysics Conference*.
- PETITOT, J., 1990. "Le Physique, le Morphologique, le Symbolique. Remarques sur la Vision", *Revue de Synthèse*, 1-2, 139-183.
- PETITOT, J., 1992. *Physique du Sens*, Paris, Editions du CNRS.
- PETITOT, J., 1993. "Phénoménologie naturalisée et Morphodynamique", *Intellectica*, Paris.
- PETITOT, J., 1994. "Phenomenology of perception, Qualitative physics, and Sheaf mereology", *16th. International Wittgenstein Symposium, Philosophy and the Cognitive Sciences*, Kirchberg/Wechsel (forthcoming).
- PETITOT, J., SMITH, B., 1991. "New Foundations for Qualitative Physics", *Evolving Knowledge in Natural Science and Artificial Intelligence*, (J.E. Tiles, G.J. McKee, G.C. Dean eds.), 231-249, Pitman, London.
- SMITH, B., 1993. "Ontology and the logistic analysis of reality", Preprint.
- THOM, R., 1988. *Esquisse d'une Sémiophysique*, Paris, InterEditions.
- WHITE, G., 1993. "Mereology, Combinatorics, and Categories", Preprint.

A Mereotopological Definition of ‘Point’*

Carola Eschenbach

Universität Hamburg, FB Informatik, AB WSV

Vogt-Kölln-Str. 30

D-22527 Hamburg, Germany

eschenbach@informatik.uni-hamburg.de

Abstract

Usually, topology is formalised on the basis of set-theoretic notions. But mereology, the formal theory of ‘part-of’ and related concepts, is a suitable alternative to set theory for this purpose. Thus, formal approaches interrelating mereological and topological notions (‘mereotopological approaches’) also have a long tradition.

In this paper, a mereotopological definition of ‘point’ is introduced, based on the topological primitive of ‘region’. It is shown that this definition is general enough, such that it allows definitions of all the usual separation properties. In contrast to other proposals in similar frameworks, the relation between points and regions is assumed to be the mereological relation of part-of. In this framework a topological treatment of granularity is possible, in which points at a coarser level of granularity are topologically structured, when analysed on a finer level.

Thus, mereotopology turns out not to be a mere terminological variant of point-set topology but to contribute to the foundations of theories of several domains of interest in cognitive science which exhibit topological structure.

1 Introduction

Motivated by the interest in basically spatial concepts such as ‘self-connectedness’, ‘separation’, ‘boundary’, ‘exterior’ and ‘interior’, point-set topology investigates the properties of sets of points, functions between sets of points, their properties and relations etc. This leads to questions such as whether space is a (structured) *set* of points or whether its structure can be adequately represented in terms of sets of points. Since there are at least epistemological reasons to give a negative answer to the former question, one might search for an alternative framework for studying topological concepts. The framework should both allow the definition of

*Thanks to Christopher Habel, Barry Smith, Laure Vieu, and the participants at the workshop on ‘Topological Foundations of Cognitive Science’ for comments on earlier versions of this paper.

topological structures without points and not require topological entities to be sets. Classical Mereology, the formal theory of 'part-of' and related concepts, is a sound and suitable basis for this (cf. Tarski 1956, Smith 1993, Varzi 1993, Eschenbach and Heydrich (to appear)). We will call approaches interrelating mereological and topological notions 'mereotopological' and those interrelating mereological and geometrical notions 'mereogeometrical approaches'. The family of such approaches has a long tradition in discussions of the foundations of mathematics (cf. Huntington 1913, de Laguna 1922, Whitehead 1929, Menger 1940, Clarke 1981, 1985).

Although point-set topology is based on the assumption of the existence of points it is not able to contribute to the elucidation of the notion of 'point'. Taking points for granted, it allows for the study of the structure of families of sets of points built up on higher levels. In contrast to this, mereotopological approaches are able to study both the notion of 'point' and topological structures not based on points. Table 1 gives a classification of mereotopological and mereogeometrical approaches with respect to the definition of the notion of 'point'.

Points are	individuals	classes of individuals	sequences of individuals	classes of sets of individuals
Definition is				
mereological	Euclid Smith 1993			
topological	Eschenbach	de Laguna 1922 Clarke 1985	Menger 1940	Whitehead 1929
geometrical	Huntington 1913	Tarski 1956		

Table 1: Classification of mereotopological definitions of 'point'.

The outline of the paper is as follows: The next two sections will give a brief introduction to Classical Mereology and to the region-based approach of mereotopology. After this, a short discussion concerning the nature of points will prepare the region-based definition of 'point' proposed in section 4. Two sorts of consequence of this definition will be discussed. First, it will be shown that the definition does not restrict the topological structure to topological spaces enjoying certain separation properties only.¹ Second, the discussion of topological structure is related to the discussion of distinguishing different levels of granularity (Hobbs 1985, Habel 1991, 1993) for the representation of space. It will be argued that the approach presented here allows for modelling space in a way, such that an object—on one level of granularity—can be modelled as a point (without any internal topological structure) and—on another level of granularity—as a region with an internal topological structure. Finally, this definition will be contrasted with region-set definitions of 'point' to be found in the literature.

¹Thus, based on this definition it is possible to differentiate between topological structures without points and topological structures in which points are not separated from each other.

2 Classical Mereology

As argued by Eschenbach and Heydrich (to appear), Classical Mereology (CM) is a formal theory of the relation of 'part-of' and its conceptual relatives 'overlap', 'discreteness' and 'sum', and does not give rise to any concept of integrity or being a whole. CM as introduced by Leśniewski (1916, 1927–30, 1983) is applicable in each and every domain, while concepts of integrity are dependent on the particular features of the structure exhibited by a specific domain. The enrichment of CM by topological notions is an elementary possibility of introducing means for establishing notions of integrity. The concept of self-connectedness, e.g., is a fundamental concept of integrity specifiable on the basis of topological notions. The interrelation of mereological and topological concepts will thus allow for approaching a genuine theory of 'part-whole-relations'.

There have been several approaches to topological or geometrical structures which are not formulated in the framework of point-set topology. Although most of them involve some notion of 'parthood', 'containment' or 'inclusion', most of them are not based on CM. The interrelation of mereological and topological notions can be accomplished in different ways (see Varzi 1993). In several cases, the mereological apparatus needed is gained on the basis of a topological primitive (see de Laguna 1922, Whitehead 1929, Clarke 1981). Consequently, the mereological structure involved in these formalisms need not obey the laws of CM.

Here we will proceed by introducing CM first and using its notion as a basis for defining topological notions (cf. Tarski 1956, Smith 1993, Eschenbach and Heydrich (to appear)). The mereological primitive employed is the binary relation of 'discreteness'. The axiomatisation is close to the one given by Leonard and Goodman (1940).²

Primitive notion: *discreteness* (*l*)

Definitions:

- [D1] x is *part of* y iff x is discrete from everything y is discrete from.
 $(x < y \equiv_{df} \forall z [z \downarrow y \supset z \downarrow x])$
- [D2] x is a *proper part of* y iff x is part of y and y is not part of x .
 $(x \ll y \equiv_{df} x < y \wedge \neg(y < x))$
- [D3] x and y *overlap* iff they have a common part. $(x \circ y \equiv_{df} \exists z [z < x \wedge z < y])$
- [D4] x is the *sum* of some entities iff x is discrete from exactly those entities which are discrete from each of them.
 $(\sigma z [\Phi(z)] \equiv_{df} \iota x [\forall y [y \downarrow x \equiv \forall z [\Phi(z) \supset y \downarrow z]]], x + y \equiv_{df} \sigma z [z = x \vee z = y])$
- [D5] x is the *product* of some entities iff x is the sum of all their common parts.
 $(\pi z [\Phi(z)] \equiv_{df} \sigma y [\forall z [\Phi(z) \supset y < z]], x \cdot y \equiv_{df} \sigma z [z < x \wedge z < y])$
- [D6] x is an *atom* iff it has no proper part. $(At(x) \equiv_{df} \neg \exists y [y \ll x])$
- [D7] The (*mereological*) *complement* of x is the sum of all entities discrete from x . $(x^{-1} \equiv_{df} \sigma y [y \downarrow x])$

²The logical framework needed for this axiomatisation is that of second order logic with descriptions and identity, including some means for treating non-referring expressions. Plural quantification—as discussed by Lewis (1991) with respect to its ontological advantages—is a useful alternative to quantification over predicates.

- [D8] A mereological structure is *atom-free* iff there is no atom.
 (Atom-free \equiv_{df} $\neg \exists x [\text{At}(x)]$)

Axioms:

- [A1] x and y are discrete iff x and y do not overlap. ($x \wr y \equiv \neg x \circ y$)
 [A2] If x is part of y and y is part of x , then x and y are identical.
 ($x < y \wedge y < x \supset x = y$)
 [A3] For any entities, their sum exists. ($\exists x [\Phi(x)] \supset \exists y [y = \sigma x [\Phi(x)]]$)

Since the topic of this paper is not CM, details of the structure defined by [A1–3] are here left aside.³ For the sake of the discussion of points, four lemmata on properties of atoms should however be mentioned.

Lemma 1: Atom x is part of y iff x overlaps y . ($\forall x, y [\text{At}(x) \supset (x < y \equiv x \circ y)]$)

Lemma 2: x is an atom iff x is the product of all entities it overlaps.
 ($\forall x [\text{At}(x) \equiv x = \pi y [x \circ y]]$)

Lemma 3: Atoms are equal iff they overlap. ($\forall x, y [\text{At}(x) \wedge \text{At}(y) \supset (x = y \equiv x \circ y)]$)

Lemma 4: A mereological structure is atom-free iff every entity overlaps two discrete entities.

(Atom-free $\equiv \forall x [\exists y, z [x \circ y \wedge x \circ z \wedge y \wr z]]$)

3 Region-based topology

Based on the primitive notion of ‘region’, fundamental topological notions such as ‘boundary’, ‘open region’, ‘closed region’, ‘closure’, ‘interior’, ‘exterior’, ‘separation’, ‘self-connectedness’ etc. can be defined. ‘Region’ is a unary predicate, not to be understood as implying any specific dimensionality. Thus, it can be true of entities of the highest dimension of the topological structure under consideration. E.g., in studying the topological structure of space, ‘region’ can be true of three-dimensional entities, and in the case of time, it can be true of one-dimensional ones. Notice that point-set topology has no notion of region.⁴ As shall become clear by the following definitions, region-based topology deals with regions and their parts only.⁵

³The structure is a Boolean algebra which is complete with the exception that there is not anything discrete from everything, nothing which is the sum of no entities, and no product of discrete entities, in short: there is no empty individual. Those who are not familiar with CM might, for the time being, imagine the structure defined to be the power-set of some set deprived of the empty set; sum corresponding to union, product to intersection, part-of to subset, overlap to having a non-empty intersection, discreteness to having no common non-empty subset, and atoms to singleton sets. But be aware that CM is more general than this, since it allows for structures without atoms (cf. e.g. Simons 1987), while the interpretation offered here brings atoms with it. Lewis (1991) presents an elaborate discussion of the interrelation between CM and set-theory.

⁴Theorem 19 shows that regions correspond to point-sets which are subsets of the closure of their interior. The union of an open set and a subset of its boundary can therefore be considered a region. Based on the interpretation offered in footnote 3, standard topology can thus be thought of as a specific case of mereotopology.

⁵Parts of regions are called topological entities here. As will be obvious, region-based topology is restricted to that part of the universe which consists of topological entities only. When necessary,

Primitive notion: *region* (\mathcal{R})

Definitions:

- [D9] x is a *topological entity* iff x is part of a region. ($\mathcal{T}(x) \equiv_{df} \exists y [\mathcal{R}(y) \wedge x < y]$)
- [D10] The *topological universe* is the sum of all regions. ($\mathcal{U}_{\mathcal{R}} =_{df} \sigma y [\mathcal{R}(y)]$)
- [D11] Regions x and y are *internally connected* iff they have a common part which is a region. ($x \star y \equiv_{df} \mathcal{R}(x) \wedge \mathcal{R}(y) \wedge \exists z [\mathcal{R}(z) \wedge z < x \wedge z < y]$)
- [D12] Regions x and y are *externally connected* iff they overlap but are not internally connected. ($x \times y \equiv_{df} \mathcal{R}(x) \wedge \mathcal{R}(y) \wedge x \circ y \wedge \neg(x \star y)$)
- [D13] Region x is *open* iff it is not externally connected to any region.
($op(x) \equiv_{df} \mathcal{R}(x) \wedge \neg \exists y [x \times y]$)
- [D14] The *interior* of x is the sum of the open regions which are part of x .
($x^{\circ} =_{df} \sigma y [op(y) \wedge y < x]$)
- [D15] y is *adherent to* x iff every open region which overlaps y overlaps x .
($y \triangleleft x \equiv_{df} \forall z [op(z) \wedge z \circ y \supset z \circ x]$)
- [D16] The *closure* of x is the sum of all topological entities adherent to x .
($x^c =_{df} \sigma y [\mathcal{T}(y) \wedge y \triangleleft x]$)
- [D17] The topological entity x is *closed* iff it is identical to its closure.
($cl(x) \equiv_{df} (x = x^c)$)
- [D18] The *topological complement* of x is the sum of all topological entities discrete from x . ($x^{-t} =_{df} \sigma y [\mathcal{T}(y) \wedge y \downarrow x]$)
- [D19] The *boundary* of x is the product of its closure and the closure of its topological complement. ($x^b =_{df} x^c \cdot (x^{-t})^c$)
- [D20] The topological entities x and y are *separated* iff x is discrete from the closure of y and y is discrete from the closure of x .
($x \parallel y \equiv_{df} \mathcal{T}(x) \wedge \mathcal{T}(y) \wedge x^c \downarrow y \wedge x \downarrow y^c$)
- [D21] The topological entity x is *self-connected* iff it is not the sum of two separated topological entities. ($con(x) \equiv_{df} \mathcal{T}(x) \wedge \neg \exists z, y [z + y = x \wedge z \parallel y]$)
- [D22] y is an *inner part* of the topological entity x iff y is part of an open region z which is part of x . ($y <_i x \equiv_{df} \mathcal{T}(x) \wedge \exists z [op(z) \wedge y < z \wedge z < x]$)
- [D23] If x is a topological entity, then y is a *dangling part* of x iff y is a proper part of x , x has a topological complement z , and y is an inner part of the closure of z . ($y <_{dg} x \equiv_{df} \mathcal{T}(x) \wedge y \ll x \wedge \exists z [z = x^{-t} \wedge y <_i z^c]$)
- [D24] A topological space is *grounded on closed entities* iff every region is the sum of the closed entities which are part of it.
($GoCE \equiv_{df} \forall x [\mathcal{R}(x) \supset x = \sigma y [cl(y) \wedge y < x]]$)

Different restrictions on the primitive of 'region' lead to different topologies just as different restrictions on 'open set', 'neighbourhood' or 'closure' in classical topology do. The axioms [A4–6] seem to be essential if the structure is to be called 'topological' at all.

Axioms:

- [A4] Every region has an open region as a part. ($\forall x [\mathcal{R}(x) \supset \exists y [op(y) \wedge y < x]]$)

quantification is restricted explicitly to allow for an easy embedding of topology into a more comprehensive ontological framework.

- [A5] For any regions their sum is a region. $(\forall x [x = \sigma y [\mathcal{R}(y) \wedge y < x] \supset \mathcal{R}(x)])$
 [A6] The product of any two overlapping open regions is an open region.
 $(\forall x, y [op(x) \wedge op(y) \wedge x \circ y \supset op(x \cdot y)])$

Example 1: Let a, b, c be three mutually discrete entities and $a, b, a+b, a+c, b+c, a+b+c$ the regions. According to the definitions given above, $a+b+c$ is the topological universe and any part of it is a topological entity. $a+c$ and $b+c$ are externally connected, while $a+c$ and $a+b$ are internally connected. $a, b, a+b, a+b+c$ are open regions, and $c, a+c, b+c, a+b+c$ are closed entities. $a, b, a+b+c$ are the interiors of $a+c, b+c$ and $a+b+c$, respectively. c is adherent to $a, b, c, a+b, a+c, b+c, a+b+c$, but has nothing but itself adherent to it. $a+c, b+c, a+b+c$ are the closures of a, b and $a+b$, respectively. c is the boundary of $a, b, c, a+c, b+c$ and $a+b$, but not of $a+b+c$. $a+b$ is not self-connected, since a and b are separated. But neither a and $b+c$, nor b and $a+c$, nor c and $a+b$ are separated. Therefore $a+b+c$ is self-connected.

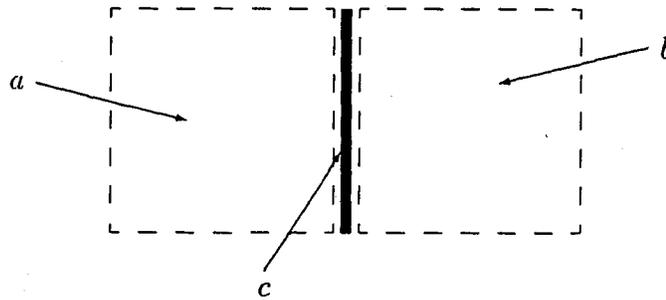


Figure 1: Example 1.

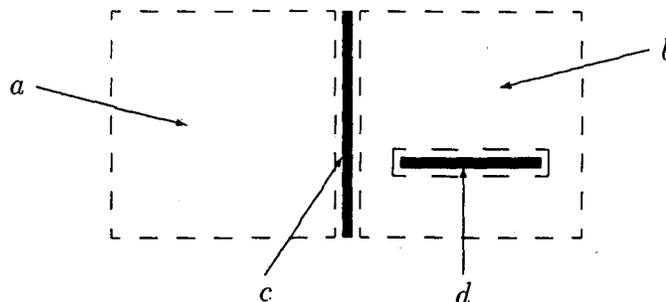


Figure 2: Example 2.

Example 2: Let a, b, c, d be four mutually discrete entities and $a, b, a+b, a+c, b+c, b+$

$d, a+b+c, b+c+d, a+b+c+d$ be the regions. $a+b+c+d$ is the topological universe and any part of it is a topological entity. $a, b, a+b, b+d, a+b+d, a+b+c, a+b+c+d$ are open regions and $c, d, a+c, b+c+d, a+b+c+d$ are closed entities. $a, b, b+d, a+b+c$ are the interiors of $a+c, b+c, b+d+c$ and $a+b+c$, respectively. c is adherent to $a, b, c, a+b, a+c, b+c, a+b+c$, but has nothing but itself adherent to it. d is, e.g., adherent to $b, d, a+b, b+c, a+b+c$, but has nothing but itself adherent to it. $a+c, b+c+d, a+b+c+d$ are the closures of a, b and $a+b$, respectively. c is the boundary of $a, b+d, c, a+c, b+c+d$ and $a+b+d$. d is the boundary of d and $a+b+c$. $c+d$ is the boundary of $b, a+b$ and $b+c$. a and b are separated and, thus, $a+b$ is not self-connected. But neither a and $b+c$, nor b and $a+c$, nor c and $a+b$ are separated. Therefore $a+b+c$ is self-connected. b and d are inner parts of $b+d$ and d is a dangling part of $a+d$.

The structure defined by these axioms is in many respects similar to classical point-set topology, although, as standardly in approaches based on Classical Mereology, there is nothing like an empty topological entity (the mereological structure does not include an empty object). The mereotopological counterparts of well known topological properties and interrelations are otherwise valid.

Theorem 5: The topological universe is open and closed. $(op(\mathcal{U}_{\mathcal{R}}) \wedge cl(\mathcal{U}_{\mathcal{R}}))$

Theorem 6: The sum of open regions is open. $(\exists x [\Phi(x)] \wedge \forall x [\Phi(x) \supset op(x)] \supset op(\sigma x [\Phi(x)]))$

Theorem 7: x is an open region iff x is identical to its interior. $(\forall x [op(x) \equiv x = x^{\circ}])$

Theorem 8: The sum of two closed entities is closed. $(\forall x, y [cl(x) \wedge cl(y) \supset cl(x+y)])$

Theorem 9: If x is the product of closed entities, then x is closed.

$(\exists y [\Phi(y)] \wedge \forall y [\Phi(y) \supset cl(y)] \supset \forall x [x = \pi y [\Phi(y)] \supset cl(x)])$

Theorem 10: The closure of x is the product of all closed topological entities x is part of. $(\forall x [x^c = \pi y [cl(y) \wedge x < y]])$

Theorem 11: The interior of x is part of x . $(\forall x, y [y = x^{\circ} \supset y < x])$

Theorem 12: If x is part of the topological entity y , then x is adherent to y . $(\forall x, y [T(y) \wedge x < y \supset x \triangleleft y])$

Theorem 13: If x is a topological entity, then x is part of the closure of x .

$(\forall x [T(x) \supset x < x^c])$

Theorem 14: If x is a topological entity and y the topological complement of x , then x is open iff y is closed, and x is closed iff y is open.

$(\forall x, y [T(x) \wedge y = x^{-t} \supset (op(x) \equiv cl(y)) \wedge (cl(x) \equiv op(y))])$

Theorem 15: If the topological space is grounded on closed entities, then y is adherent to x iff no part of y is separated from x .

$(GoCE \supset \forall x, y [y \triangleleft x \equiv \neg \exists z [z < y \wedge z \parallel x]])$

Notice that regions need not be self-connected. The following theorems are meant to shed some light on what can be conceived of as a region. In essence, a region is an entity whose boundary is (entirely) adherent to its interior and is not too thick (i.e. has no region as a part).

Lemma 16: If there is a region, then the topological universe is a region.

$$(\exists x [\mathcal{R}(x)] \supset \mathcal{R}(\mathcal{U}_{\mathcal{R}}))$$

Lemma 17: x is a topological entity iff x is part of the topological universe.

$$(\forall x [\mathcal{T}(x) \equiv x < \mathcal{U}_{\mathcal{R}}])$$

Lemma 18: Every region is a topological entity. $(\forall x [\mathcal{R}(x) \supset \mathcal{T}(x)])$

Theorem 19: If x is a region, then it has no dangling parts.

$$(\forall x [\mathcal{R}(x) \supset \neg \exists y [y <_{dg} x]])$$

Theorem 20: If x is a region and y the boundary of x , then y has no region as a part. $(\forall x, y [\mathcal{R}(x) \wedge y = x^b \supset \neg \exists z [\mathcal{R}(z) \wedge z < y]])$

4 Points

Euclid's definition (a point is that which has no (proper) part) is what first comes to mind when we are asked what is a point.⁶ This definition is a purely mereological one, and there already is a place for this kind of entity in the above: they are atoms. But although 'point' should be defined in a way that is based on topological notions, it is worth comparing atoms and their properties to points.

Atom are those entities which do not exhibit any structure at all: They do not have any proper part and do not overlap another atom nor any entity they are not a part of. Points are entities which do not exhibit any *topological* structure: They do not have a proper part which is a region or a point, they do not overlap another point nor overlap any region they are not in.

An alternative definition of 'point' is a (geo)metrical one: A point is what is not extended. This definition is based on a metrical—and not a purely topological—concept. Since geometrical and metrical structures are richer than topological ones in the sense that topological concepts can be defined in (geo)metrical terms, it is worth looking for a basically topological definition.⁷

A definition of 'point' is appropriate if it is in accordance with the intuitive notion we have. An intuitively clear mereotopological definition of 'point' is one which takes points to be regions which have no region as proper part. Such regions do not exhibit any topological structure. A corresponding idea is expressed by Huntington in the context of mereogeometry.⁸ But there are topological structures which do not involve such regions, and therefore it seems worth search for a more general definition.

The relation between points and regions is often specified by the preposition *in*. This relation can be understood as being the mereological relation of part-of, which makes points topological entities. In addition, it seems to be sensible to

⁶Cf. Smith (1993), a recent approach in mereotopology along this line.

⁷Although at this stage of discussion it is not possible to show that the definition presented here fulfils this requirement, it should be noted that an appropriate definition of 'point' should lead to the validity of: If x is a point with respect to the topological structure arisen by a metrical space (or induced by the metric), then x is not extended with respect to the underlying metric.

⁸Taking *sphere* and *containment* as primitive notions, Huntington (1913) defines points to be spheres which do not contain a(nother) sphere.

assume that every point is in at least one region. This makes points ontologically dependent on regions. An identity criterion for points can also be based on region: Points which are in exactly the same regions are equal. As said before, although points are topological entities, they do not exhibit any topological structure. These requirements are met by definition [D25].

Definitions:

- [D25] x is a *point* iff x is a topological entity, x is part of every region it overlaps, and every topological entity which is part of any region x overlaps is part of x .
 $(PT(x) \equiv_{df} T(x) \wedge \forall z [T(z) \supset (z < x \equiv \forall y [\mathcal{R}(y) \wedge x \circ y \supset z < y]])]$
- [D26] x is *in* y iff x is a point, y a region, and x is part of y .
 $(IN(x, y) \equiv_{df} PT(x) \wedge \mathcal{R}(y) \wedge x < y)$
- [D27] z is a *neighbourhood* of point x iff z is a topological entity and x is in an open region y which is part of z .
 $(\mathcal{N}(z, x) \equiv_{df} PT(x) \wedge T(z) \wedge \exists y [op(y) \wedge IN(x, y) \wedge y < z])]$
- [D28] A topological space is *point-free* iff there is no point in the topological universe. $(Pfr \equiv_{df} \neg \exists x [PT(x) \wedge IN(x, \mathcal{U}_{\mathcal{R}})])]$
- [D29] A topological space is *grounded on points* iff every region is the sum of the points in it. $(GoP \equiv_{df} \forall y [\mathcal{R}(y) \supset y = \sigma x [PT(x) \wedge IN(x, y)])]$

The topological structure defined by [A4-6] is not restricted in any way with respect to the existence of points. The structure may allow for topological differentiation without limit. Point-free topology is possible⁹, as well as topology which includes points but is not grounded on them.

That points can have parts is shown in the following example.

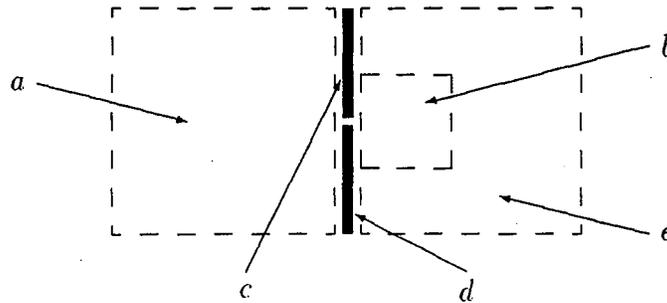


Figure 3: Example 3.

Example 3: Let a, b, c, d, e be five mutually discrete entities, $a, b + e, a + b + e, a + c + d, b + c + d + e, a + b + c + d + e$ the regions. $a, b + e, a + b + e, a + b + c + d + e$ are open

⁹A simple case of point-free topology is an atom-free mereological structure in which every entity is a region.

regions and $a + c + d, b + c + d + e, a + b + c + d + e$ the closures of $a, b + e, a + b + e$ respectively. $c + d$ is the closure of c, d and itself and $c + d$ is the boundary of $a, b + e, a + c + d, b + c + d + e, a + b + e$. In this structure, there are three points: $a, b + e, c + d$, and b, c, d, e are proper parts of points.

Theorems 21 and 22 show that Huntington's (1913) definition of 'point' captures a special case of the definition presented here.

Theorem 21: If x is a region which has no region as proper part, then x is a point.
 $(\forall x [\mathcal{R}(x) \wedge \neg \exists y [\mathcal{R}(y) \wedge y \ll x] \supset PT(x)])$

Theorem 22: If x is a region which does not overlap any region it is not part of, then x is a point. $(\forall x [\mathcal{R}(x) \wedge \neg \exists y [\mathcal{R}(y) \wedge x \circ y \wedge \neg x < y] \supset PT(x)])$

The lack of topological structure of points is expressed in theorems 23, 24, 28 and 29. The following theorems show the similarity of points and atoms and that the definition meets the analysis of the intuitive notion of 'point'.

Theorem 23: Point x is in region y iff x overlaps y .
 $(\forall x, y [PT(x) \wedge \mathcal{R}(y) \supset (IN(x, y) \equiv x \circ y)])$

Theorem 24: The topological entity x is a point iff x is the product of all regions it overlaps. $(\forall x [PT(x) \equiv (x = \pi y [\mathcal{R}(y) \wedge x \circ y])])$

Theorem 25: Points are equal iff they overlap.
 $(\forall x, y [PT(x) \wedge PT(y) \supset (x = y \equiv x \circ y)])$

Theorem 26: A topological space is point-free iff every topological entity overlaps two discrete regions. $(Pfr \equiv \forall x [T(x) \supset \exists y, z [x \circ y \wedge x \circ z \wedge y \wr z]])$

Theorem 27: Points are equal iff they are in exactly the same regions.
 $(\forall x, y [PT(x) \wedge PT(y) \supset (x = y \equiv \forall z [\mathcal{R}(z) \supset (IN(x, z) \equiv IN(y, z))])])$

Theorem 28: If x is a point and y part of x , then x is part of the closure of y .
 $(\forall x, y [PT(x) \wedge y < x \supset x < y^c])$

Theorem 29: Every point is self-connected. $(\forall x [PT(x) \supset con(x)])$

That no atoms are part of the topological universe does not mean that there are no points in it. Just take any example given and assume that the basic, non-analysed objects are infinitely dividable. On the other hand, that there are no points in the topological universe means that there are no atoms which are part of it. The validity of theorem 31 poses a serious restriction on the possibility of presenting examples of point-free topological structures. First, topologies derived from finitely many entities (atoms) are grounded on points. Second, if you think of the domain of Classical Mereology as sketched in footnote 3, standard topology can be thought of as mereotopology. But since singleton sets are atoms with respect to this structure, in this domain one will always get atomic structures and, accordingly, topological structures based on points.

Theorem 30: If the topological entity x is part of every region it overlaps, then x is part of a point. $(\forall x [T(x) \wedge \forall z [\mathcal{R}(z) \wedge x \circ z \supset x < z] \supset \exists y [PT(y) \wedge x < y]])$

Theorem 31: Every atom which is a topological entity is part of a point.
 $(\forall x [T(x) \wedge At(x) \supset \exists y [PT(y) \wedge x < y]])$

Theorems 32 to 35 show that the relation between points and their neighbourhoods corresponds to the classical topological one.

Theorem 32: A point is inner part of every one of its neighbourhoods.

$$(\forall x, y [PT(x) \wedge \mathcal{N}(y, x) \supset x <_i y])$$

Theorem 33: Every topological entity which has a neighbourhood of point x as a part is a neighbourhood of x . $(\forall x, y [PT(x) \wedge \exists z [\mathcal{N}(z, x) \wedge z < y] \supset \mathcal{N}(y, x)])$

Theorem 34: The product of two neighbourhoods of a point x exists and is a neighbourhood of x . $(\forall x, y, z [PT(x) \wedge \mathcal{N}(y, x) \wedge \mathcal{N}(z, x) \supset \exists w [w = y \cdot z \wedge \mathcal{N}(w, x)])]$

Theorem 35: Every neighbourhood z of a point x has a neighbourhood y of x as a part, such that z is neighbourhood of every point which is part of y .

$$(\forall x, z [PT(x) \wedge \mathcal{N}(z, x) \supset \exists y [\mathcal{N}(y, x) \wedge y < z \wedge \forall w [PT(w) \wedge w < y \supset \mathcal{N}(z, w)])]])$$

In the approach presented, the relation between points and regions or boundaries is the mereological relation of part-of. But the definition of 'point' does not imply that every point is, e.g., a part of at least one boundary. In addition, the question whether every region or boundary is exclusively built up from points (i.e. is the sum of the points which are part of it) might be discussed. Therefore, it is possible to study whether a topological structure allows for something other than points (i.e. for something not having any point as part or, even stronger, something not overlapping any point).

5 Points and Separation

Theorem 27 suggests that points are always separated in a sense analogous to the condition on T_0 -spaces in point-set topology.¹⁰ From the definition of 'point' it might be suspected that points are even separated according to the demands of T_1 -spaces.¹¹ To show that this is not the case, first the definitions of the separation properties have to be transferred to region-based topology. It is important to notice that the mereotopological analogue of the condition on normal spaces in region-based topology is independent from what points are and whether any exist.

Definitions:

[D30] A topological space is a T_0 -space iff for every two (distinct) points there is an open region z such that exactly one of them is in z .

$$(T_0 \equiv_{df} \forall x, y [PT(x) \wedge PT(y) \wedge x \neq y \supset \exists z [op(z) \wedge (IN(x, z) \neq IN(y, z))]])$$

[D31] A topological space is a T_1 -space iff for every two (distinct) points x, y there is an open region z such that x is in z and y is not in z .

$$(T_1 \equiv_{df} \forall x, y [PT(x) \wedge PT(y) \wedge x \neq y \supset \exists z [op(z) \wedge IN(x, z) \wedge \neg IN(y, z)])]$$

¹⁰A point-set topological space is a T_0 -space iff for every two (distinct) points there is an open set z such that exactly one of them is in z .

¹¹A point set topological space is a T_1 -space iff for every two (distinct) points x, y there is an open set z such that x is in z and y is not in z .

- [D32] A topological space is a T_2 - (or *Hausdorff*-)space iff for every two (distinct) points x, y there are two discrete open regions u, z such that x is in z and y is in u .
 $(T_2 \equiv_{df} \forall x, y [PT(x) \wedge PT(y) \wedge x \neq y \supset \exists u, z [op(z) \wedge op(u) \wedge u \uparrow z \wedge IN(x, z) \wedge IN(y, u)])]$
- [D33] A topological space is *regular* iff for every closed entity y and point x not in y there are two discrete open regions u, z such that x is in z and y is part of u .
 $(Regular \equiv_{df} \forall x, y [PT(x) \wedge cl(y) \wedge \neg IN(x, y) \supset \exists u, z [op(z) \wedge op(u) \wedge u \uparrow z \wedge IN(x, z) \wedge y < u]])]$
- [D34] A topological space is a T_3 -space iff it is a T_1 -space and regular.
 $(T_3 \equiv_{df} T_0 \wedge Regular)$
- [D35] A topological space is *normal* iff for any two discrete closed entities y, x there are two discrete open regions u, z such that x is part of z and y is part of u .
 $(Normal \equiv_{df} \forall x, y [cl(x) \wedge cl(y) \wedge x \uparrow y \supset \exists u, z [op(z) \wedge op(u) \wedge u \uparrow z \wedge x < z \wedge y < u]])]$
- [D36] A topological space is a T_4 -space iff it is a T_1 -space and normal.
 $(T_4 \equiv_{df} T_0 \wedge Normal)$

Points need not be separated since they might be in exactly the same *open* regions while there is a non-open region only one of them is in. Whether a topological structure is a T_0 - or T_1 -space depends on whether or not we add the condition that points are closed entities.

Theorem 36: There are mereotopological spaces which are not T_0 -spaces.

Proof of 36: Let a, b, c, d be four mutually discrete entities, $a, b, a + b, a + c, a + d, a + c + d, b + c + d, a + b + c + d$ the regions. In this structure, there are four points: a, b, c, d . $a + b + c + d$ is the only open region c or d are in. Thus, c and d are in exactly the same open regions and therefore not separated. Q.E.D.

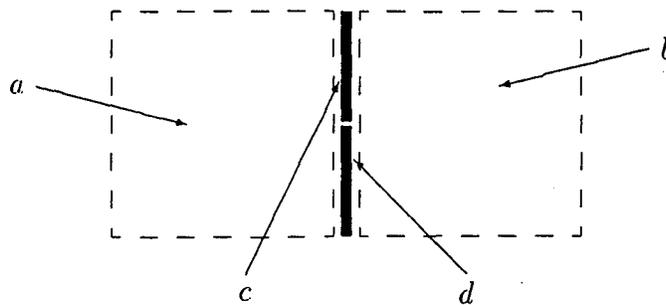


Figure 4: Proof of theorem 36.¹²

¹²The figure might be misleading, since it suggests that the relation between a and c and d is

Theorem 37: If every point of a topological space is closed, then it is a T_1 -space. $(\forall x [PT(x) \supset cl(x)] \supset T_1)$

Theorem 38: If a T_1 -space is grounded on points, then every point is closed. $(T_1 \wedge \text{GoP} \supset \forall x [PT(x) \supset cl(x)])$

Theorem 39: If every point of a topological space is closed, then the topological space is a T_3 -space iff it is regular. $(\forall x [PT(x) \supset cl(x)] \supset (T_3 \equiv \text{Regular}))$

Theorem 40: If every point of a topological space is closed, then the topological space is a T_4 -space iff it is normal. $(\forall x [PT(x) \supset cl(x)] \supset (T_4 \equiv \text{Normal}))$

6 Topological levels

In several areas of cognitive science it has been noted that there is a need to take different levels of granularity or refinement into account (Hobbs 1985, Habel 1991, 1993). Knowledge and belief about several domains can be organised according to which level of detail is needed for expressing it. To give an example from the area of knowledge about space: Maps of countries present (smaller) cities as points (without internal structure), while maps of these cities show a lot of structure, and, in turn, present houses, blocks and (smaller) parks as unstructured. The question addressed in this paragraph is, how mereotopological structures on different levels of granularity can be interrelated. The introduction of different levels of granularity should allow for topological entities which are points on the coarser level, while exhibiting topological structure on the finer level.

The comparison of different topological structures of one domain can also be done in the framework of point-set topology.¹³ Although there might be immense differences between different structures of one domain, all structures have the same notion of point. This holds because what is a point in classical topology is only dependent on what the domain is. In contrast to this, mereotopology allows for topological entities which are points on the coarser level and structured regions on a finer level of granularity.

The mereotopological structure assumed as basis for defining the topological structure is domain independent and also independent of levels of granularity. If an entity has parts, than it has these parts, independently of whether they are of interest on the level of granularity under consideration or not. Different levels of granularity in topologically structured domains must therefore be differentiated on the basis of the topological notions. In the region-based approach to mereotopology proposed here, one may get different sets of points by introducing different notions of 'region'. (In what follows, the subscripts c and f are used to refer to the topological structures imposed by \mathcal{R}_c and \mathcal{R}_f .)

There are different ways in which systems of regions can be interrelated. The interrelation between different topological levels in one system of granularity should

the same as that between b and c and d . That this is not the case should be clear from the explicit definition.

¹³A topological structure is called 'coarser' than another, if every set which is open wrt. the former is open wrt. the latter structure.

obey axiom [A7], which forces the topological universe of the different levels of granularity to be invariant, and axiom [A8], which motivates the terms 'finer' and 'coarser' level of granularity.

Axioms:

[A7] The topological universe is the same on all levels of granularity. ($\mathcal{U}_{\mathcal{R}_f} = \mathcal{U}_{\mathcal{R}_c}$, $\sigma_y[\mathcal{R}_f(y)] = \sigma_y[\mathcal{R}_c(y)]$)

[A8] Any region at the coarser level is a region at the finer level. ($\forall x[\mathcal{R}_c(x) \supset \mathcal{R}_f(x)]$)

As a consequence of these axioms, the following theorems are provable.

Theorem 41: Any topological entity on the finer level is a topological entity on the coarser level and vice versa. ($\forall x[\mathcal{T}_f(x) \equiv \mathcal{T}_c(x)]$)

Theorem 42: If x is the topological complement of y on one level, then x is the topological complement of y on the other level of granularity. ($\forall x, y[x = y^{-f} \equiv x = y^{-c}]$)

Theorem 43: If x and y are internally connected on the coarser level, then they are internally connected on the finer level. ($\forall x, y[x \star_c y \supset x \star_f y]$)

Theorem 44: If x is a point on the finer level, then x is part of a point at the coarser level. ($\forall x[PT_f(x) \supset \exists y[PT_c(y) \wedge x < y]]$)

The topological granularity system allows for many different kinds of structures. It is, e.g., not provable, that boundaries on the finer level are less (or 'thinner') than those on the coarser level. It is also possible that an open region on the coarser level is not open on the finer level. Thus, the relation between topological structures as defined by [A7] and [A8] differs from the relation 'finer-coarser' as standardly defined on the basis of classical topology. The relation corresponding to the classical one is gained by the addition of one of the following statements, which are equivalent to each other in the given context, as an axiom.

- Every open region on the coarser level is open on the finer level. ($\forall x[op_c(x) \supset op_f(x)]$)
- Every closed entity on the coarser level is closed on the finer level. ($\forall x[cl_c(x) \supset cl_f(x)]$)
- If y is the interior of x at the coarser level, then y is part of the interior of x at the finer level. ($\forall x, y, z[y = x^{oc} \wedge z = x^{of} \supset y < z]$)
- If y is adherent to x at the finer level, then y is adherent to x at the coarser level. ($\forall x, y[y \triangleleft_f x \supset y \triangleleft_c x]$)
- If y is the closure of x at the finer level, then y is part of the closure of x at the coarser level. ($\forall x, y, z[y = x^{cf} \wedge z = x^{cc} \supset y < z]$)
- If y is the boundary of x at the finer level, then y is part of the boundary of x at the coarser level. ($\forall x, y, z[y = x^{bf} \wedge z = x^{bc} \supset y < z]$)

As a consequence, one also gets the validity of the following statements:¹⁴

¹⁴If the topological structure at the coarser level is grounded on closed entity, these statements are also equivalent to the preceding ones.

- If x and y are separated on the coarser level, then x and y are separated on the finer level. $(\forall x, y [x \parallel_c y \supset x \parallel_f y])$
- If x is self-connected on the finer level, then x is self-connected on the coarser level. $(\forall x [con_f(x) \supset con_c(x)])$

The region-based approach to mereotopology thus allows for a more general relation of comparison of topological structures than point-set topology. This relation can be interpreted in terms of distinguishing between different level of granularity. Since the definition of point proposed here is relative to the mereotopological structure, points on one level of granularity need not to be points on another one.

7 Region-set approaches to 'point'

Several approaches to the definition of the notion of 'point' assume points to be sequences or sets of regions or even sets of sets of regions (cf. de Laguna 1922, Menger 1940, Whitehead 1929, Tarski 1956, Clarke 1985) and thereby introduce an additional level into the ontological structure. We will call such definitions, which are motivated by the idea that points are more abstract than regions, 'region-set' definitions. In contrast to this, the definition of 'point' given in section 4 leads to points being as concrete as regions are. This allows assumptions such as cities or other concrete objects to be points. A translation of Menger's (1940, pp. 91) definition to the nomenclature of region-based topology is presented here as an example of region-set approaches to point definitions.

Definitions:

- [D37] Region x is *completely contained* in region y iff the closure of x is an inner part of y . $(x \ll_i y \equiv_{df} \mathcal{R}(x) \wedge \mathcal{R}(y) \wedge x^c \ll_i y)$
- [D38] Region x is *disjoint from* region y iff there is no region completely contained in both of them. $(x \dagger y \equiv_{df} \mathcal{R}(x) \wedge \mathcal{R}(y) \wedge \neg \exists z [z \ll_i x \wedge z \ll_i y])$
- [D39] A sequence of regions, x_1, x_2, \dots , is *strictly decreasing* iff x_{k+1} is completely contained in x_k for each k . $(str-decr((x_i)) \equiv_{df} \forall k [x_{k+1} \ll_i x_k])$
- [D40] An *M-point* is a strictly decreasing sequence of open regions, x_1, x_2, \dots , such that for any open region y which does not completely contain any of the x_k , and any open region z completely contained in y , z is disjoint from almost all x_k . $(M-PT((x_i)) \equiv_{df} str-decr((x_i)) \wedge \forall k [op(x_k)] \wedge \forall y, z [op(y) \wedge \neg \exists k [x_k \ll_i y] \wedge op(z) \wedge z \ll_i y \supset \exists k [z \dagger x_k]])$
- [D41] An M-point, x_1, x_2, \dots , is said to *lie in* region y iff y contains completely a x_k (and consequently almost all x_k). $(M-IN((x_i), y) \equiv_{df} M-PT((x_i)) \wedge \mathcal{R}(y) \wedge \exists k [x_k \ll_i y])$
- [D42] The M-points x_1, x_2, \dots and y_1, y_2, \dots are called *equal* iff each x_i contains a y_j (and consequently almost all y_j) completely and each y_i contains a x_k (and consequently almost all x_k) completely. $((x_i) =_M (y_i) \equiv_{df} M-PT((x_i)) \wedge M-PT((y_i)) \wedge \forall i [\exists j, k [y_j \ll_i x_i \wedge x_k \ll_i y_i]])$

As Menger explicitly states, the main motivation for giving this definition is to build up an analogy between topology and arithmetic. This definition parallels a

well known way of defining real numbers as sequences of rational intervals. But the parallelism is only superficial. In addition to the acceptance of set-theory, the ontological basis of defining reals in arithmetic is the existence of rationals. As a consequence, there is a canonical embedding of the rationals in the set of reals. But in the case M-points there seem to be no correlates for the rationals, and no idea of how such correlates might look.

One problem facing Menger's approach is that of identity of M-points. Menger defines identity of M-points differently from identity of sequences ([D42]). In addition, if there are two different M-points, then there are two discrete open regions such that exactly one of the points lies in either. This means that M-points are separated according to the conditions of T_2 -spaces (see definition [D32]). Therefore definition [D40] restricts the range of structures which can be studied.

It seems to be more fruitful to distinguish between, on the one hand, points and other topological entities and, on the other hand, sequences of topological entities and their properties with respect to convergence. The definition proposed in section 4 allows for a rejection of the existence of points without denying the existence of M-points. The distinction between these two levels corresponds to the distinction between rational and real numbers.

In general, region-set approaches avoid the use of set theory for the discussion of regions, but introduce set theory for the definition of 'point'. The basic character of points—they do not exhibit any structure—is not reflected by these definitions. And finally, consequences of claims such as 'there are (no) points with such and such property' are in combination with this kind of definition not easy to understand.¹⁵

8. Conclusion

The region-based approach to mereotopology is one of the approaches interrelating topological and mereological notions in theories about regions, boundaries and connection. As we have seen, it allows for a treatment of points as entities of the same ontological level as regions. Thus, this approach gives rise to a pure topological notion of '*concrete points*' without requiring regions or boundaries to consist of points (only). The definition does not restrict the range of possibilities concerning the separation properties of points, since it allows for the existence of non-separated points. Like other mereotopological approaches, this calculus can be enriched by a definition of '*abstract points*', which might be sets or sequences of regions, such that the interrelation of concrete and abstract points can be studied.

Assuming points and boundaries to be parts of regions permits the formalization of structural interrelations between different perspectives on a domain as discussed here with respect to the possibility of different levels of granularity or refinement. As we saw, the mereotopological approach presented generalizes the standard approach in two respects. On the one hand, an object can be a point on the coarser level while being a structured region on the finer level, which is excluded in the standard

¹⁵In Eschenbach (in preparation) the approach of Clarke (1985) is discussed. The main problem of this approach (reduction of topology to Classical Mereology) can be claimed to result from the difficulty of understanding such assumptions.

approach by assuming points to be on another ontological level than open sets etc. On the other hand, defining the relation of coarser and finer on the basis of the new primitive 'region' yields a more general relation than is defined in the standard approach on the basis of the primitive 'open set'.

As this discussion has shown, mereotopology is not a mere terminological variant of classical topology but may contribute to the foundations of the theory of space, time and other domains of interest in cognitive science which exhibit mereological and topological structures.

References

- [1] Clarke, B.L. (1981): 'A calculus of individuals based on "connection"'. *Notre Dame Journal of Formal Logic* 22, 204-218.
- [2] Clarke, B.L. (1985): 'Individuals and points'. *Notre Dame Journal of Formal Logic* 26, 61-75.
- [3] Eschenbach, C. (in preparation): 'A note on Clarke's points'. Ms., Hamburg.
- [4] Eschenbach, C. & W. Heydrich (to appear): 'Classical mereology and restricted domains'. To appear in: N. Guarino & R. Poli (eds.): *Formal Ontology in Conceptual Analysis and Knowledge Representation*, Kluwer. (Sorter versions in: N. Guarino & R. Poli (eds., 1993): *Proceedings of the International Workshop on Formal Ontology in Conceptual Analysis and Knowledge Representation*. Ladseb-CNR Report 01/93, Padova, Italy. 205-217; and in: R. Casati & G. White (eds., 1993): *Philosophy and the Cognitive Sciences*. The Austrian Ludwig Wittgenstein Society, Kirchberg am Wechsel. 219-223.)
- [5] Habel, Ch. (1991): 'Hierarchical representations of spatial knowledge: Aspects of embedding and granularity'. Paper presented at the Second International Colloquium on Cognitive Science (ICCS-91).
- [6] Habel, Ch. (1993): 'Representing Space and Time: Discrete, Dense or Continuous? - Is that the Question?' Zentrum für Interdisziplinäre Forschung, Universität Bielefeld, Forschergruppe *Mental Models in Discourse Processing*. Report 6/93.
- [7] Hobbs, J.R. (1985): 'Granularity'. *Proc. 9th IJCAI*. 432-435.
- [8] Huntington, E.V. (1913): 'Postulates for abstract geometry'. *Mathematische Annalen* 73. 522-559.
- [9] Laguna, T. de (1922): 'Point, line and surface as sets of solids'. *The Journal of Philosophy* 19. 449-461.
- [10] Leonard, H.S. & N. Goodman (1940): 'The calculus of individuals and its uses'. *Journal of Symbolic Logic* 5. 45-55.

- [11] Leśniewski, S. (1916) 'Podstawy ogólnej teorii mnogości I'. [Foundations of a General Theory of Manifolds.] *Prace Polskiego Kola Naukowego w Moskwie, Sekcja matematyczno-przyrodnicza*, 2. Moscow.
- [12] Leśniewski, S. (1927-30) 'O Podstawach Matematyki'. *Przegląd Filozoficzny* 30 (1927) 164-206; 31 (1928) 261-91; 32 (1929) 60-101; 33 (1930) 75-105, 142-70.
- [13] Leśniewski, S. (1983) 'On the foundations of mathematics'. [Abridged English translation of 'O Podstawach Matematyki'] *Topoi* 2. 7-52.
- [14] Lewis, D. (1991): *Parts of Classes*. Basil Blackwell: Oxford.
- [15] Menger, K. (1940): 'Topology without points'. *Rice Institute Pamphlets* 27. 80-107.
- [16] Simons, P. (1987): *Parts. A Study in Ontology*. Clarendon Press: Oxford.
- [17] Smith, B. (1993): 'Ontology and the logistic analysis of reality'. In: G. Häfliger & P.M. Simons (eds.): *Analytic Phenomenology*. Kluwer: Dordrecht, Boston, London. 223-245. (Preliminary version in: N. Guarino & R. Poli (eds., 1993): *Proceedings of the International Workshop on Formal Ontology in Conceptual Analysis and Knowledge Representation*, Ladseb-CNR Report 01/93, Padova, Italy. 51-68.)
- [18] Tarski, A. (1956): *Logic, Semantics, Metamathematics*. Oxford University Press. (trans. by J.H. Woodger.)
- [19] Varzi, A.C. (1993): 'On the boundary between mereology and topology'. To appear in: R. Casati, B. Smith & G. White (eds.): *Philosophy and Cognitive Science*. Proceedings of the 16th International Wittgenstein Symposium, Vienna: Hoelder-Pichler-Tempsky (preprinted as IRST Technical Report 9308-04, Trento.)
- [20] Whitehead, A.N. (1929): *Process and Reality*. The Macmillan Co.: New York. Cited after: *Process and Reality: Corrected Edition* (1978): The Free Press, Macmillan Pub. Co.: New York.

Discreteness, Finiteness, and the Structure of Topological Spaces

Christopher Habel

University of Hamburg, Computer Science Department

Doctoral Program in Cognitive Science

Vogt-Kölln-Str. 30, D-22527 Hamburg, Germany

habel @informatik.uni-hamburg.de

1. The Concept of Representation

One of the fundamental assumptions of Cognitive Science is that the concept of representation is central for investigating and explaining cognition.

“According to the representational theory of the mind, [...] mental states are representational states, and mental activity is the acquisition, transformation and use of information [...]” (Sterelny 1990, p. 19)

In agreement with the Physical Symbol System Hypothesis (Newell & Simon 1976), “A physical symbol system has the necessary and sufficient means for general intelligent action.”, cognition can be seen as computation (Pylyshyn 1984), whereas the behavior of these cognitive, computational systems depends on the internal representations which it, in turns, effects. Since the properties of the representation influence the computation, and the demands of the processes determine the resulting representations, in studying cognitive processes the system of mutual constraints between representations and processes has to be taken into account.

In the present paper I will address properties of representational systems concerning the question of “finite vs. infinite nature of representations”. To start, let us have a look at formal, i.e. mathematical, symbol systems, such as Turing Machines. The behavior of a Turing Machine is determined by two types of representations: The current states and the current inscription on the tape; both are finite in the sense that they can be coded by finite means: only a finite part of the Turing tape contains symbols distinct from the “blank”.¹ As Chomsky argues, the mathematical concepts of grammars and automata are fundamental for understanding “competence as a system of generative processes”, i.e. “how a language can (in Humboldt’s words) “make infinite use of finite means”” Chomsky (1965, p. 4ff.).

Questions of finiteness are discussed in Cognitive Science not only from a formal

I would like to thank Carola Eschenbach for her helpful comments on earlier drafts of the manuscript.

¹ Similar finiteness conditions hold for other types of symbol systems investigated in the mathematical theory of formal grammars and automata.

mathematical point of view but also from an empirical perspective. Johnson-Laird postulates with respect to “mental models”, a specific type of internal representations:

“The principle of finitism: A mental model must be finite in size and cannot directly represent an infinite domain.” (Johnson-Laird 1983, p.398)²

Evidence from cognitive psychology and neuropsychology suggests that mental images are restricted with respect to size and resolution: The *visual buffer*, the medium for the realization of mental images, has limited spatial extent and limited resolution (Kosslyn 1980, p. 139). The visual buffer can represent objects in multiple – but limited – scales; by the multiscaled property it is possible to resolve the conflict between scope and resolution (Kosslyn 1994, pp. 95–104).

Since “mental models” (Johnson-Laird 1983) as well as “depictive representations” or “quasi-pictorial representations” (Kosslyn 1984, 1994) are non-propositional, often characterized as “analog”, it is important to mention that finiteness constraints are not restricted to non-propositional representations but also concern propositional representations. Johnson-Laird and Kosslyn motivate and justify the finiteness assumptions on the limitations of the underlying mental machinery, the brain in general or, more specifically, the visual buffer in the case of mental images.

The topic of the present paper is a question emerging from the finiteness assumptions: “How can the non-finite space-time reality be represented in a finite medium?” To get a better grasp of this problem, which I called *the continuity - finiteness dilemma* in Habel (1993), we need to examine the concept of *representation*. Following Palmer (1978, pp. 263ff.), “representation situations” can be characterized by the represented (\mathcal{W}_1) and the representing world (\mathcal{W}_2), their internal structures, which can be seen as relational systems in the sense of formal model theory, and a partial mapping, ρ from \mathcal{W}_1 to \mathcal{W}_2 , which is called the representational mapping.

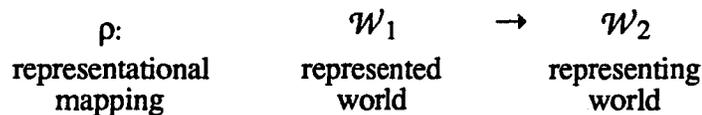


Fig. 1: A representation situation

Applying this general representation scheme (Fig. 1) to mental representations leads to a more complex representation situation (Fig. 2), in which four worlds and three representational mappings are involved. According to the representational theory of mind, an essential module of any cognitive system is a mental model (\mathcal{W}_2) of the real world (\mathcal{W}_0). The finiteness constraints hold on this representation, \mathcal{W}_2 . Cognitive Science investigates also second-level representations, namely representations (\mathcal{W}_3) of mental representations (\mathcal{W}_2); the finiteness of \mathcal{W}_3 is inherited from \mathcal{W}_2 . Note that \mathcal{W}_2 is representation as well as representandum. On the other hand, in sciences (partial) mathematical models of the real world (\mathcal{W}_1) are investigated, such as formal models of space-time in physics.

² Furthermore, there exist a *constraint of quantity isomorphism*, which can informally be described by: There is a one-to-one correspondence between the entities of the world (real or fictional) and the entities in the mental model. This constraint is the consequence of Johnson-Laird's (1983) principles of *structural identity* (p. 419) and that of *set formation* (p. 429).

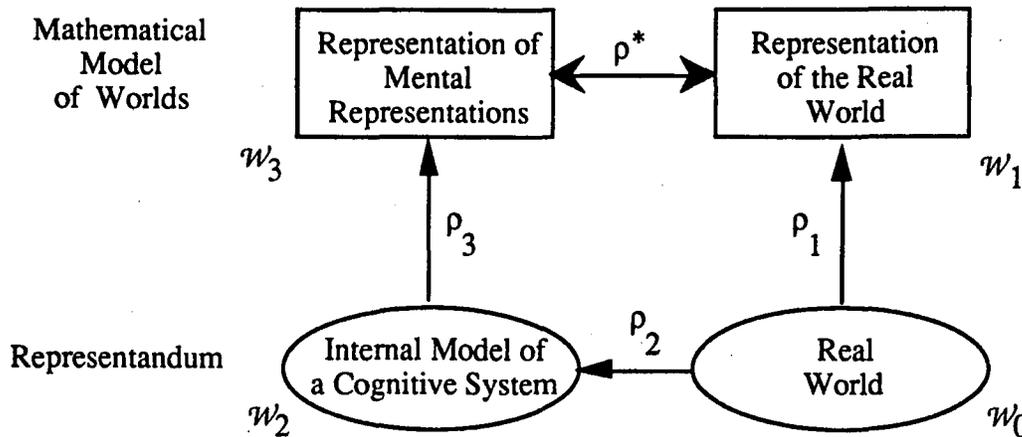


Fig. 2: The representation situation for "mental representations"³

The *continuity - finiteness dilemma* concerns the relation given by ρ_2 between \mathcal{W}_0 and \mathcal{W}_2 as well as the correspondence ρ^* between \mathcal{W}_1 and \mathcal{W}_3 , which is induced by the representational mappings ρ_1 , ρ_2 and ρ_3 in the "representation diagram" (Fig. 2).

2. Towards a Cognitive Representation of Time and Events

Since the domains of time and of events are strongly connected⁴ – especially via a canonical function from events to time entities, which assigns to each event e its "running time" $rt(e)$ – I will discuss the two domains in parallel: As empirical data for human common-sense understanding of time, I will use natural language discourse on events.

2.1. Basics of Temporal Ontology

I will start with a short survey on the most relevant options for time structures, which have been discussed in philosophy of science, logics and linguistics, namely discrete, dense and continuous time.⁵ The difference between these types concerns the *size* as well as the *neighborhood* and *separation* properties of the time structure, i.e. it is a genuine topological differentiation.

The common axiomatic basis of time structures is given by a set of conditions, which characterize the order of time. They are independent of the choice between discreteness, density or continuity.

³ Paivio's (1986, pp. 18ff.) discussion of Palmer's (1978) analysis of cognitive representations contains three levels of representation, namely \mathcal{W}_0 , \mathcal{W}_2 and \mathcal{W}_3 of Fig. 2; "non-cognitive" representations, \mathcal{W}_1 , (e.g. those of physics) are not discussed by Paivio.

⁴ In the present paper "events" is used in a wide sense that is synonymous to Bach's (1986) *eventuality*. Thus, this notion includes *events* (in a narrow sense), *processes* and *states*. On detailed discussions of the connection between *time* and *events* see Kamp (1979), who constructs time from events, and Krifka (1987), who presupposes the independent existence of the two domains, which are connected by the *running-time* function "rt".

⁵ In the following, I take the perspective of point structures, in contrast to interval or period structures. This opposition is neutral with respect to the main topic discussed here. In the description of time structures and in the formulation of the axioms, I follow - with minor revisions - van Benthem (1991).

- [TRANS] $\forall x, y, z: (x < y \ \& \ y < z \rightarrow x < z)$
 [IRREFL] $\forall x: \neg x < x$
 [LIN] $\forall x, y: (x = y \vee x < y \vee y < x)$

Now it is possible to differentiate between types of time structure concerning size and topological structure, by further axioms:

Discrete time, with \mathbb{Z} as a distinguished model, is characterized by the discreteness axioms, which also characterizes the property of step-wise succession: each element has an unique immediate successor and predecessor.⁶

- [D.SUCC] $\text{succ}(x, y) \equiv_{\text{def}} x < y \ \& \ \neg \exists u: (x < u \ \& \ u < y)$
 [DISCR] $\forall x, y: (x < y \rightarrow \exists z: \text{succ}(x, z) \ \& \ \exists z: \text{succ}(z, y))$

Dense time, with \mathbb{Q} as a distinguished model, is characterized by the density axiom:

- [DENS] $\forall x, y: (x < y \rightarrow \exists z: x < z < y).$

For each pair of elements there is a further one in between them. The density axiom prevents the existence of an immediate successor or predecessor.

Continuous time, with \mathbb{R} as a distinguished model, is characterized by the continuity axiom:⁷

- [CONT] $\forall \Phi: [[\forall x, y: ((\Phi x \ \& \ \neg \Phi y) \rightarrow x < y) \ \& \ \exists x: \Phi x \ \& \ \exists x: \neg \Phi x]$
 $\rightarrow \exists z: (\forall u (z < u \rightarrow \neg \Phi u) \ \& \ \forall u: (u < z \rightarrow \Phi u))]$

The conceptual complexity of the three different types of time structure is reflected by their formal properties with respect to their definability (in first-order or higher-order logic) and cardinality:

	distinguished model	cardinality	definable in
non-finite discrete time	\mathbb{Z}	\aleph_0	first-order logic
dense (non-continuous) time	\mathbb{Q}	\aleph_0	first-order logic
continuous time	\mathbb{R}	2^{\aleph_0}	higher-order logic

Despite the “complexity leap” from density to continuity, it is standard practice in physics to see time as a structure isomorphic to the real numbers, \mathbb{R} (cp. Grünbaum, 1973). Newton-Smith, an advocate of continuous time, describes the situation in physics as follows:

“While time can coherently be supposed to be discrete we have no good reasons for taking seriously the hypothesis that it is so. For no one has been able to produce viable physical theories that treat time as discrete. Indeed, all mainline physical theories represent time by a parameter ranging over the real numbers and in so doing treat time as continuous. Interestingly, ... we can

⁶ Note that the condition “ $x < y$ ” is needed for the specific cases of the beginning or end of time in finite temporal structures.

⁷ Φ is a variable over first-order predicates; in the axiom of continuity – via quantification of predicates – the Dedekind-cut property of real number is postulated for time, too. For details, see van Benthem (1991, pp. 29ff.) Further, note that continuity implies density.

construct equally viable alternatives to these physical theories in which time is treated as merely dense and not continuous." Newton-Smith (1980, p. 121)

Since there are physically adequate alternatives to continuous time structures, an interesting question is, why there is a common preference to "real valued time" in the sciences. A preliminary answer is: The mathematics of the real numbers, the classical calculus, is "easier" than its rational number alternatives. The continuity axiom guarantees the existence of Dedekind-cut elements within the set in question, e.g. in \mathbb{R} . Consider the predicate $\Phi x \equiv_{\text{def}} x^2 < 2$ on \mathbb{R}^+ , then [CONT] gives rise to the existence of an entity z with: $\forall u (z < u \rightarrow 2 \leq u^2)$ & $\forall u (u < z \rightarrow u^2 < 2)$; the conventionalized description of z is " $\sqrt{2}$ ". This Dedekind-cut property is the cause for *convergence* in the basic set, i.e. that the limit of a sequence if it exists, is element of the same set as the members of the sequence. (This exactly distinguishes \mathbb{Q} from \mathbb{R} .) The existence of limits is essential for the real valued calculus: the classical theories of differential and integral equations, which are the core of the mathematical inventory of physics, depend on the existence of limits. Often, the non-continuous analog of a solvable differential equations, the difference equations, do not have a solution. From this point of view, giving preference to \mathbb{R} over \mathbb{Q} can be seen as a primarily technical decision which is not justified by theoretical reasons or empirical evidence.

This prejudice to a continuous time structure, which is common in the sciences, has widely been adopted in linguistics (and cognitive science in general). As an example of such an adoption, I refer to Kamp & Reyle (1993): Although in the beginning of the chapter on "Tense and Aspect" they shortly discuss the different options of time structures, they follow the real-numbers path after the following "justification":

"What counts as a natural property of time is a matter that cannot be easily decided. However, we regard it as uncontroversial that the natural properties of time, whatever they may be, do not *exclude* the structure of the real numbers." Kamp & Reyle (1993, p. 491, footnote 2.)

In contrast to common practice in semantics, which widely neglects the discussion on the structure of time or sees this topic as a peripheric problem, I consider the question of "finite vs. infinite" or "discrete vs. dense/continuous" time structures to be central for a cognitive approach to mental representations, namely to determine adequate "topological structures" of representational systems.

2.2. Systems of Discrete Time Structures

It is common to use phrases as "(and) next" or "immediately after" in English or "direkt/unmittelbar (da)nach" in German, by which one event is connected to its successor,⁸ see the following examples from Collins COBUILD (1987):

Next, we did the Merchant of Venice...

Immediately I finish the show I get changed and go home.

She finished her cigarette, then lit another one immediately.

A first analysis of sentences of this type leads to the following structure on the represen-

⁸ Motivated by such phenomena von Wright (1965) developed a time logic based on a modal connective T , which is to be read as *and next*.

tation level: $SUCC(e_1, e_2) \approx$ "e₂ succeeds immediately e₁". By the properties of immediate succession the time and event structure are constrained to be discrete (see sect. 2.1, especially [DISC]).

On the other hand, it is possible to have a discourse structure explicitly mentioning an event and its immediate successor in the first sentence, and referring to a third event between the two immediately succeeding events in the second sentence, as in:

- (1) i. Immediately after graduation she started university education.
 Before she started university education, she worked N weeks in a factory.
- ii. Immediately after graduation she was climbing in the Alps.
 Before she went to the Alps, she worked N weeks in a factory.

In an informal preexperiment these sentences were presented to subjects (students of computer science); especially the parameter N was varied. I will briefly summarize the results: the subjects valued the discourse as fully acceptable, if N was small enough.⁹ To sum up, examples from natural language seem to lead to the following unsatisfiable pair of properties for the structure of cognitive time: the time structure holds the properties of step-wise succession and of density, which are incompatible.¹⁰

The conflict between immediate succession / discreteness on the one hand, and density on the other hand, vanishes, when a cognitive, process-oriented point of view is taken. According to the idea of - what I call - "proxy-semantics", an internal representation of proxies for the entities spoken about is constructed during the process of discourse comprehension.¹¹ The most important break with traditional semantics, such as Montague semantics, is that the introduction of discourse entities into the representations and the "maintenance" of these proxies is an essential part of semantics. From this point of view, the situation of example (1) can be schematized by Figure 3:

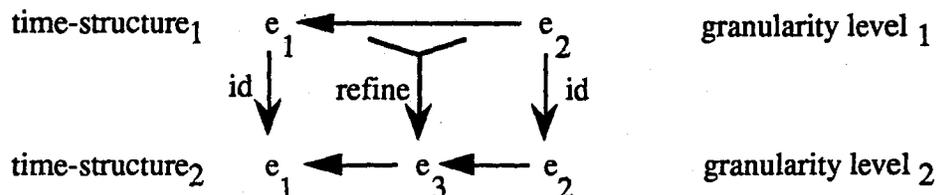


Fig. 3: Development of time-structure: change into a finer granularity-level
 ("←" \approx " _ is predecessor of _")

The development of time-structures, namely from TS_1 to TS_2 , can be seen as a change from a coarser level of *detail* to a finer one; these levels I call - in the spirit of Hobbs' (1985) seminal paper - granularity levels: But, saying that a finer granularity level was introduced does not imply a withdrawal of the coarser one. Instead of assuming a process, in which at any time one and only one time-structure exist, namely the structure

⁹ "Small enough" is a gap of 8-10 weeks between the reference event and its immediate successor in the university case and approximately of 3-4 weeks in the climbing example.

¹⁰ Similar observations and analyses can be found at several places in the literature, see for instance Kamp (1979, p. 408f.) and Link (1987, p. 248f.).

¹¹ Representative for proxy-semantics in the logico-linguistic tradition is Discourse Representation Theory (see Kamp & Reyle 1993), for the artificial intelligence / cognitive science type compare Referential Nets (Habel 1986).

refined as far as possible by the information given by the discourse, I propose the simultaneous existence of several time-structures organized in a hierarchy of granularity levels.¹² As usual in embedding structures there are “identification links” (id), which describe the “trans-level identities”, and “refinement links” (refine), which describe the “immediate precursor constellations” (IPC), which are refined to “chains of IPCs”.

According to the finiteness assumptions postulated in section 1, I assume a time-structure hierarchy, in which each granularity level is discrete and finite. The embedding or refinement faculty of the cognitive system induces a property on the hierarchical time structure which I call *density in intensio*. This notion is motivated by Church's (1941) terminology describing the fundamental dichotomy of perspectives on functions:

- function in intension, where a function is seen as a rule of correspondence, which specifies for a given entity (the argument) another entity (the value).
- function in extension, where a function is seen as set of pairs consisting of an argument and the corresponding value.

From a computational point of view: Functions in extension correspond to fixed structures, e.g., tables; thus, to compute a function-in-extension means to look up the table. In contrast to this, the computation of a function-in-intension is real computation, namely the application of the rule of correspondence mentioned above.

With respect to mental representations there are also different ways to specify or characterize time-structures: (α) by listing all the time entities, (β) by stating a property which an entity must have to be qualified as entity in the structure, and (γ) by defining rules of construction which generate the entities in the structure. (α) corresponds to functions in extension, (β) and (γ) to functions in intension. In this sense the “density of the time structure” is not extensional; The process of repeatable application of the construction rules (of embedding and refinement) leads to an intensional characterization of the time-structure following the density principle [DENS].

The perspective emphasizing the process of constructing time-structures does not require the prior existence of a time-structure. This contrast corresponds to the opposition – see Fig. 2 – between the real world, \mathcal{W}_0 , and the models of physics, \mathcal{W}_1 , on the one hand, and internal models of cognitive systems, \mathcal{W}_2 , and their representations, \mathcal{W}_3 , on the other hand. The “incompatibility” between the time-theories from physics and those needed for cognitive science vanishes if considering the just mentioned opposition. But it is not enough to understand the reason of incompatibility: cognitive science in general, and linguistics in particular, should look for alternatives to explicitly, extensionally given dense and continuous time-structures.¹³

¹² The assumption of a hierarchy of time-structures – instead of one maximally fine structure – is motivated as an instance of a general principle of mental representations, namely the simultaneous accessibility to information of different degrees of detail. For the domain of discourse comprehension and production this principle is empirically well supported; see as an example Kintsch & van Dijk (1978), especially on levels of details in summarizing.

¹³ Landman (1991, pp. 136ff.) discusses refinement problems similar to those presented in the present paper. He argues in the opposite direction: against finite structures and in favor of dense models of time. Since Landman's argumentation is not focused on phenomena of cognitive processes, his orientation to extensional presupposed dense time structures is not surprising.

3. Topology and Granularity Systems of Time

The topology of a time-structure is the system of temporal entities and neighborhoods of such entities meeting the constraints given by the axioms of ordering.¹⁴ If a system of granularity levels is taken into account – as proposed in sect. 2 – it is necessary to determine the properties of “proper refinement”. Note that the examples given above have been concerned with “nearly ideal situations”; now, I will indicate how some more complex, but more realistic examples, have to be dealt with.

Given two time-structures $(TS_1, <_1)$ and $(TS_2, <_2)$, seen as time-graphs TG_i , i.e. TS_i is the set of nodes of TG_i and the edges of TG_i are given by $SUCC_i(e_1, e_2)$ instances, a *refinement* is a pair of relations (ρ_e, ρ_o) , where:¹⁵

- $\rho_e \subset TS_1 \times TS_2$ is a left-total relation, describing the embedding constellation of the time *entities*, i.e. the nodes of the time-graphes,
- $\rho_o \subset (TS_1 \times TS_1) \times (TS_2 \times TS_2)$ is a left-total relation, describing the embedding constellation of the *ordering* constellations, i.e. the *SUCC* instances represented by the vertices of the time-graphes.

(From the perspective of graph refinement, the relations ρ_e and ρ_o can be seen as refinement operations; therefore I use an assignment notation symbolized with “ \rightarrow ”.) The situations verbally introduced by (1) and depicted in Fig. 3. are of the most basic refinement type, that of *splitting a succession*, by introducing a new event. It can be described as follows:

$$\begin{aligned} TS_1 &= \{e_1, e_2\}, & TS_2 &= \{e_1, e_2, e_3\} \\ \rho_e &: e_1 \rightarrow e_1, e_2 \rightarrow e_2 \\ \rho_o &: SUCC_1(e_1, e_2) \rightarrow SUCC_2(e_1, e_3), SUCC_2(e_3, e_2) \end{aligned}$$

In contrast to this we have *splitting of entities*,¹⁶ for example verbally described by:

(2) Immediately after graduation she started university education.

She enrolled at the MIT and moved to Boston.

I will only mention the relevant parts of the refinement here:

$$\begin{aligned} TS_1 &= \{e_1, e_2\}, & TS_2 &= \{e_1, e_2, e_3, e_4\} \\ \rho_e &: e_1 \rightarrow e_1, e_2 \rightarrow \{e_3, e_4\} \\ \rho_o &: SUCC_1(e_1, e_2) \rightarrow SUCC_2(e_1, e_3), SUCC_2(e_3, e_4) \end{aligned}$$

The splitting of entities by ρ_e , i.e. the decomposition of events or time entities, induces

¹⁴ The notion “system of neighborhoods” is an abbreviation for the axiomatic system with respect to the basic topological concepts of a time-structure, which I can not discuss in detail here. See van Benthem (1991) on “points and periods” in time-structures, and Eschenbach (1994) on a mereotopological definition of “point”, which is appropriate to handle granularity structures.

¹⁵ Since time-structures and eventuality-structures are canonically connected, I will discuss in the following only one type of structures, namely those of time.

The use of set-theoretic notions for describing the approach does not imply that time- or event-structures are seen here as set-theoretic entities. A mereotopological analysis of the phenomena in question – based on Eschenbach (1994) – is in preparation.

¹⁶ Splitting entities in the way I describe here presupposes an ontology of time and events in which it is possible to analyze points at different levels of granularity. Or, in other words, points of one level can be refined to structured entities on a finer level. See Eschenbach (1994) for an mereotopological approach, which satisfies this demand.

for the refinement of ordering ρ_0 that the newly introduced entities have to be integrated into the temporal ordering, as depicted in Fig. 4.

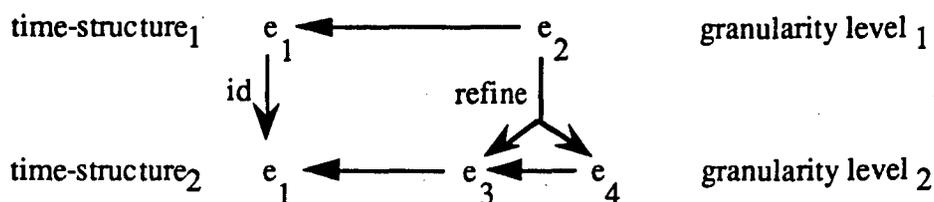


Fig. 4: Splitting of entities

Now I come back to conditions of proper refinement; the splitting-of-events example (2) gives rise to refinement constraints as

$$[\text{PROPREFINE-1}] \quad \forall e_1, e_2 (e_1 <_1 e_2 \rightarrow \forall e_3, e_4 : \\ (\rho_e(e_1, e_3) \ \& \ \rho_e(e_2, e_4) \rightarrow e_3 <_2 e_4))$$

There is an analog constraint for time-graphs formulated with respect to graph refinement; i.e. considering the replacement of nodes and vertices of TG_1 by graphs leading to TG_2 . These constraints serve to guarantee that in time- or event-splitting the ordering of time is maintained.

4. Conclusion

From a cognitive point of view, internal representations are finite, and therefore time (and eventualities) as part of the internal representations of a cognitive system have a discrete structure. On the other hand, repeatable refinement, which can be seen as leading to density in intension, is a central faculty in cognition.

Refinement processes operate on hierarchical systems of time-structures, which can be interpreted as granularity systems. From a formal point of view, these granularity systems are a non-trivial extension of standard topology.¹⁷

References

- Bach, Emmon (1986). The algebra of events. *Linguistics and Philosophy*, 9, 5–16.
- van Benthem, Johan (1991; 2nd ed.). *The Logic of Time*. Dordrecht: Reidel.
- Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, MA.: MIT Press.
- Collins COBUILD English Language Dictionary*. (1987). London & Glasgow: Collins.
- Eschenbach, Carola (1994). A mereotopological definition of 'point'. In C. Eschenbach, C. Habel & B. Smith (Eds.), *Topological Foundations of Cognitive Science*, Papers from the Workshop at the FISL-CS, Buffalo, NY. July 9-10, 1994. Bericht

¹⁷ The case of space gives a wide class of further evidence for discrete representations, structured in granularity systems. The argument for mereotopology and against set-theoretic topology holds as well as in the time / event domain.

- Nr. 37 des Graduiertenkolleg *Kognitionswissenschaft* Hamburg.
- Grünbaum, Adolf (1973; 2nd ed.). *Philosophical Problems of Space and Time*. Dordrecht: Reidel.
- Habel, Christopher (1986). *Prinzipien der Referentialität*. Berlin: Springer-Verlag.
- Habel, Christopher (1993). *Representing Space and Time: Discrete, Dense or Continuous? – Is that the Question?* (Research Group “Mental Models in Discourse Processing”. Technical Report 6/93). Center for Interdisciplinary Research, University of Bielefeld.
- Hobbs, J. R. (1985). Granularity. *Proceedings of 9th IJCAI*, 432-435.
- Johnson-Laird, Philip N. (1983). *Mental Models*. Cambridge: Cambridge University Press.
- Kamp, Hans (1979). Events, instants and temporal reference. In R. Bäuerle, U. Egli, & A. von Stechow (Eds.), *Semantics from Different Points of View* (pp. 376-417). Berlin: Springer.
- Kamp, Hans & Reyle, Uwe (1993). *From Discourse to Logic*. Dordrecht: Kluwer.
- Kintsch, Walter / van Dijk, Teun (1978). Towards a model of discourse comprehension and production. *Psychological Review*, 85. 363-394.
- Kosslyn, Stephen M. (1980). *Image and Mind*. Cambridge, MA.: Harvard UP.
- Kosslyn, Stephen M. (1994). *Image and Brain*. Cambridge, MA.: MIT-Press.
- Krifka, Manfred (1987): Nominal reference and temporal constitution. In J. Groenendijk, M. Stokhof & F. Veltman (Eds.), *Proceedings of the Sixth Amsterdam Colloquium* (pp. 153-173).
- Landman, Fred (1991). *Structures for Semantics*. Dordrecht: Reidel.
- Link, Godehard (1987): Algebraic semantics of event structures. In J. Groenendijk, M. Stokhof & F. Veltman (Eds.), *Proceedings of the Sixth Amsterdam Colloquium* (pp. 143-162).
- Newell, Allen & Simon, Herbert A. (1976) Computer science as empirical inquiry. *Communications of the ACM*, 19, 113-126.
- Newton-Smith, W. H. (1980). *The Structure of Time*. London: Routledge & Kegan Paul.
- Paivio, Allan (1986). *Mental Representation: A dual coding approach*. New York: Oxford University Press.
- Palmer, Stephen E. (1978). Fundamental aspects of cognitive representations. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 259-303). Hillsdale, NJ.: Erlbaum.
- Pylyshyn, Zenon (1984). *Computation and Cognition*. Cambridge, MA.: MIT-Press.
- Sterelny, Kim (1990). *The Representational Theory of Mind*. Oxford: Basil Blackwell.
- von Wright, Georg Henrik (1965). And next. *Acta Philosophica Fennica*, 18. 293-304.

Mass reference and the geometry of solids*

Almerindo E. Ojeda
University of California at Davis

To account for the semantics of mass nouns, models have been proposed which provide atomless mereologies. The class of atomless mereologies is, however, prodigiously rich and varied. For indeed, if k is an infinite cardinal, then there are exactly 2^{2^k} atomless mereologies of cardinality 2^k no two of which are isomorphic. The purpose of this paper is to claim that natural reference to masses calls for only one of these structures —that of Tarski's mereology of solids. It follows from this proposal that (i) the mass domain has the power of the continuum, that (ii) the intuitive analyses of *space* and *time* can be unified, that (iii) we can account for the fact that natural languages routinely locate *all* masses in space, that (iv) all mass nouns have the same semantic structure, that (v) mass noun denotations indeed have the same structure as the entire mass domain, and that (vi) every partition of a mass is denumerable.

1. Introduction

To account for the semantic properties of mass terms, semanticists have generally assumed sets of infinitely divisible masses which are related as parts are related to wholes.¹ To be more precise, they have proposed models which provide a set M and a binary relation \leq which jointly verify the conditions in (1).

- (1) TRANSITIVITY: For all $x, y, z \in M$: $x \leq y$ and $y \leq z$ jointly imply that $x \leq z$.
COMPLETENESS: For all $N \subseteq M$: If N is nonempty, then N constitutes exactly one element of M .
ATOMLESSNESS: For all $x \in M$: there exists some $y \in M$ such that both $y \leq x$ and $y \neq x$.

The notion of constitution involved in the condition of completeness is defined in (2), while the notion of overlap involved in the notion of constitution is defined in (3).

* I am indebted to Donald Monk, Carlos Borges, Steven Lapointe, Alexandre Turull, Joel Friedman, Edward Keenan, Fred Landman, and Melven Krom for useful exchanges about the contents of this paper. I am also grateful to Barry Smith and the participants at the Symposium on the Topological Foundations of Cognitive Science (State University of New York at Buffalo, July 9-10, 1994), where this paper was first presented.

¹ See for example Wald (1977), Link (1983), Krifka (1987), and Ojeda (1993). Similar assumptions can be found in Sharvy (1983), Bunt (1985), and Lønning (1987).

- (2) Some $N \subseteq M$ constitutes some $m \in M$ if and only if
- (i) For all $x \in N$: $x \leq m$.
 - (ii) For all $x \leq m$: there exists some $n \in N$ such that x overlaps n .
- (3) Some $x \in M$ overlaps some $y \in M$ if and only if there is some $z \in M$ such that both $z \leq x$ and $z \leq y$.

In short, semanticists have proposed models which provide *atomless mereologies* $\langle M, \leq \rangle$.² We will say that M is a *set of masses*, that \leq is a *relation of mass inclusion*, and that $\langle M, \leq \rangle$ is a *mereology of masses*.

Having availed themselves of models with mereologies of masses, semanticists proceed to interpret a mass noun relative to a model as a principal ideal of the mereology of masses provided by the model (a *principal ideal* of a mereology of masses can be defined as the set of masses included in a particular mass). To illustrate, the mass noun *water* would denote, relative to a suitable model, the set of masses included in a particular mass —the mass of water provided by the model.

The proposal that the masses of natural linguistic discourse form an atomless mereology is hardly a trivial one. Yet, the class of atomless mereological structures is truly enormous. For indeed, if k is any nondenumerable cardinal, then the number of structurally distinct atomless mereological structures with k elements is an exponential function of k (Section 2). The question thus arises as to whether natural language semantics in fact needs the entire class of atomless mereological structures. The purpose of this paper is to answer this question in the negative and to argue that natural reference to masses calls for only *one* such structure —the structure of the mereology of solids developed by Tarski in his work on the foundations of geometry. Section 3 describes this mereology from several points of view.

The claim that the mass domain has the structure of the mereology of solids provides a complete structural characterization of the mereology of masses. It thus allows us to infer a number of important facts about mass reference. We may infer, for example, that the mass domain has the power of the continuum, so a universe of discourse will have exactly as many masses as there are points in space. Since mass nouns denote principal ideals and every mass in a model generates one, a model can provide no more than 2^{\aleph_0} distinct mass noun interpretations (Section 4).

We may also infer that the mass domain is 'homogeneous', which is to say that every principal ideal of the mass domain constitutes a mereology with the same structure as the mass domain as a whole. Since mass nouns denote principal ideals, every mass noun will have the same structure as the entire mass domain, and all mass nouns will have the same semantic structure. The structure of Tarski's mereology of solids will thus be reproduced inside each and every mass noun denotation (Section 5).

The claim that the mass domain has the structure of the mereology of solids furthermore implies that the mereology of masses satisfies the 'countable chain

² See Simons (1987) for a comprehensive study of mereologies.

condition'. This means that no mass may have a nondenumerable set of mutually disjoint and jointly exhaustive parts. If the taxonomic uses of mass nouns (cf. sentences like *this store carries all three wines*) involve sets generated by partitions of masses,³ then a denumerable bound exists on the cardinalities of these sets (Section 6).

We conclude this introduction by describing the close connection that exists between mereologies and complete Boolean algebras. This connection will be crucial for the sequel, as it will justify the application of the theory of Boolean algebras to our study of the mereology of masses. The connection between mereologies and complete Boolean algebras can be readily described by the following procedures (Tarski 1956b, 333f).

- (4) a. If we take a mereology $\langle M, \leq \rangle$, add to M an element not already in M , and redefine \leq so the added element bears the redefined relation to all the elements of M , then we have a complete Boolean algebra.
- b. Conversely, if we take a complete Boolean algebra $\langle A, \leq \rangle$, remove from A its null element (i.e. the element which bears \leq to all the elements of A), and redefine \leq so that it does not relate the deleted element to any element of the diminished A , then we have a mereology.⁴

Now, if $\langle A, \leq \rangle$ is a complete Boolean algebra, it is customary to use A^+ to refer to the set of nonnull elements of A . We may therefore use $\langle A^+, \leq \rangle$ for the mereology derived from $\langle A, \leq \rangle$ by (4).

2. The variety of atomless mereologies

Our goal in this section is to give a precise idea of the variety atomless structures exhibit. To do so we will need a definition of mereological isomorphism and a lower bound on the cardinality of atomless mereologies. The definition is as follows.

- (5) A mereology $\langle M, \leq_M \rangle$ and a mereology $\langle N, \leq_N \rangle$ are said to be *isomorphic* if and only if there exists a bijection $\Phi: M \rightarrow N$ such that for all $x, y \in M$,

$$x \leq_M y \text{ if and only if } \Phi(x) \leq_N \Phi(y).$$

To indicate in writing that a mereology $\langle M, \leq_M \rangle$ is isomorphic to a mereology $\langle N, \leq_N \rangle$ we will use $\langle M, \leq_M \rangle \cong \langle N, \leq_N \rangle$.

³ See Ojeda (1993, Chapter 5).

⁴ To simplify exposition we will here and henceforth identify a Boolean algebra with its canonical partial order. Nothing of substance is lost in doing so, since Boolean algebras and their canonical partial orders are interdefinable (Monk 1989, 15). Notice that mereologies are free structures in the sense that they may contain no elements. Naturally, a mereology will contain no elements if and only if it corresponds, by (4), to the trivial Boolean algebra of one element (Monk 1989, 9).

Two mereologies are thus isomorphic if and only if their elements can be placed in a one to one correspondence which preserves their mereological relations. Since a mereological structure is completely determined by its relation, the notion of mereological isomorphism in (5) adequately formalizes our intuitive notion of identity with respect to mereological structure.

We now need a lower bound on the cardinality of an atomless mereology. To provide one, let $\langle M, \leq \rangle$ be any atomless mereology. If M contains any elements whatsoever, then it will contain infinitely many of them. For indeed, suppose M contains some element m . It follows from the atomlessness of $\langle M, \leq \rangle$ that m will contain, as parts, an infinitely descending chain of masses, each different from the others. Hence, M will be infinite if it contains any element whatsoever (= if it is not empty). But then we can show that if M is not empty, then it will even have a nondenumerable infinity of elements. In other words, we can prove the following fact.⁵

(6) Every nonempty atomless mereology has at least 2^{\aleph_0} elements.

We are finally in a position to evaluate structural diversity in the class of atomless mereologies. It is given by (7).

(7) If k is an infinite cardinal, then there are exactly 2^{2^k} atomless mereologies of cardinality 2^k no two of which are isomorphic.

Thus, since \aleph_0 is an infinite cardinal, there will be $2^{2^{\aleph_0}}$ structurally distinct atomless mereologies of cardinality 2^{\aleph_0} . But 2^{\aleph_0} is, as we have seen, only the smallest possible cardinality for a nonempty atomless mereology. In addition, there will be $2^{2^{2^{\aleph_0}}}$ structurally distinct atomless mereologies of cardinality $2^{2^{\aleph_0}}$; $2^{2^{2^{2^{\aleph_0}}}}$ structurally distinct atomless mereologies of cardinality $2^{2^{2^{\aleph_0}}}$ —and so on through the endless hierarchy of infinite cardinalities.⁶

⁵ For, given (4), the proposition follows from (i) the infinite cardinality of atomless mereologies, (ii) the fact that complete Boolean algebras of infinite cardinality have at least 2^{\aleph_0} elements (Monk 1989, 40), and (iii) the fact that $2^{\aleph_0} - 1 = 2^{\aleph_0}$ (so the removal of the null element will have no effect whatsoever on cardinality).

⁶ As for proof of (7), consider the following. There are only three kinds of Boolean algebras: the atomless, the atomistic, and the Cartesian products of the two. Furthermore, for each cardinal k , there can only be one complete atomistic Boolean structure of cardinality 2^k , namely the atomistic structure with k atoms. Hence there is just one complete atomistic Boolean structure of cardinality 2^k , where k is an infinite cardinal. Next, let a be the number of complete atomless Boolean structures of cardinality 2^k . It follows that there can only be ak complete product structures of cardinality 2^k (they are the products of the atomistic structures with 1, 2, ..., k atoms times each of the a atomless structures). Now comes the crucial step: for each infinite cardinal k there are exactly 2^{2^k} complete Boolean structures of cardinality 2^k (Monk 1989, 482). Hence, by the previous calculations, $1 + a + ak = 2^{2^k}$. This means

It should be clear that the prodigious variety of atomless mereologies reported in (7) reckons only differences in structure, while altogether ignoring differences in content. In other words, (7) recognizes only the different ways in which the binary relation \leq of a mereology $\langle M, \leq \rangle$ can be defined, setting aside the differences M proper may bring about. Such differences cannot begin to be assessed, however, unless and until a theory is developed concerning the kinds of sets M can be —that is in addition to being able to serve as the domain of an atomless mereology. The proposals to be made in this paper will do nothing to restrict the diversity in content such an important theory would presumably provide.

3. The mereology of solids

In an address delivered at the First Polish Mathematical Congress of 1927, Tarski proposed a new foundation for the geometry of solids —one which inverted the logical structure of geometrical concepts by taking solids as primitive while defining points, lines, and surfaces, in terms of solids.⁷ To do so he began by developing a mereology as follows. Suppose we were given the set of solids of three dimensional Euclidean space and the relation of solid inclusion. Say that two solids overlap if and only if there is a solid which they both include. Say also that some set of solids constitutes a particular solid if and only if (i) every solid in the set is included in the particular solid and (ii) every solid included in the particular solid overlaps at least one of the solids in the set. Suppose, finally, that the relation of solid inclusion is transitive, and that every nonempty set of solids, even if spatially disconnected, constitutes one and only one solid. It follows from these assumptions and definitions that the set of solids of three dimensional Euclidean space forms a mereology when taken in conjunction with the relation of solid inclusion. Call this the *mereology of solids*.⁸

Notice that the mereology of solids is atomless, as every solid will include solids other than itself. To develop an intuition for this fact, imagine a sphere of unit radius. It will include a concentric sphere with radius 1/2. This sphere will in turn include a concentric sphere of radius 1/4, which will in turn include a sphere of radius 1/8, and so on without end (the point lying at the center of all these spheres is not itself a sphere or even a solid, but only a point).⁹

Now, if we are willing to forgo of intuition, then the mereology of solids can be formulated in ordinary, noninverted, geometry —one where points, not solids, are primitive (Tarski 1956a). When it is thus formulated, the mereology of solids becomes the *mereology of regular open sets (of three dimensional Euclidean space)*. Here we

that $a = 2^{2^k}$, so there are 2^{2^k} complete *atomless* Boolean structures of cardinality 2^k . The proposition in (7) now follows from (4).

⁷ Such geometries were also investigated by Whitehead (1919; 1920), De Laguna (1922), and Nicod (1924). The idea of taking solids as the primitives of geometry can be traced back to Leibniz (cf. Mach 1906, 50).

⁸ See Tarski (1956a).

⁹ Formally, points are *sets* of solids in Tarski's system. Hence they cannot be solids.

suppose we are given the set of points of three dimensional Euclidean space. Call this set R^3 . The mereology of regular open sets can now be constructed as follows. Let $RO(R^3)$ be the family of regular open sets of R^3 and let \subseteq continue to be the relation of set inclusion.¹⁰ Since R^3 is a topological space, the pair $\langle RO(R^3), \subseteq \rangle$ is a complete Boolean algebra (Monk 1989, 26f). But the null element of this algebra turns out to be the empty set. Hence, by (4), the pair $\langle RO(R^3)^+, \subseteq \rangle$ is a mereology in which $RO(R^3)^+$ is the family of *nonempty* regular open sets of R^3 .

The notions of overlap and constitution in the mereology of regular open sets come out as follows. Two elements of $RO(R^3)^+$ overlap if and only if they have a nonempty intersection. A given element of $RO(R^3)^+$ constitutes a given set of elements of $RO(R^3)^+$ if and only if the given element is the regularized union of the given set.¹¹ Finally, it should be clear that $\langle RO(R^3)^+, \subseteq \rangle$ is atomless, as any element of $RO(R^3)^+$ will always contain another within it — say an open sphere.¹²

But let us return now to the *algebra* $\langle RO(R^3), \subseteq \rangle$ of regular open sets of R^3 . Notice that it can be reconstructed as follows. We may assume that R^3 can be placed in a one to one correspondence with triples of real numbers. Place R^3 in any such correspondence. Consider then the set of spheres of R^3 whose centers and radii are completely described by rational numbers. Suppose we were now to “peel off” the surfaces of these spheres. Suppose, that is, that we were interested only in the interiors of our “rational spheres”. Collect all these sphere interiors in a set. Call this set B . Since the rational numbers are denumerable, so will B . Consider now the subalgebra of $\langle RO(R^3), \subseteq \rangle$ generated by B . It turns out to be a Boolean algebra which is both atomless and denumerable, but not complete. It will be atomless because every sphere interior will contain nonempty sphere interiors other than itself. It will be denumerable because the subalgebra of $\langle RO(R^3), \subseteq \rangle$ generated by B happens to be the closure of denumerably many sphere interiors under *finite* regularized union. It is not complete because it is not closed under *infinite* regularized unions. Suppose, then, that we were to make our algebra complete. Suppose, that is, that we were to add to our subalgebra the *infinite* regularized unions of the rational sphere interiors. It turns out that we would obtain $\langle RO(R^3), \subseteq \rangle$ itself! This reconstruction of $\langle RO(R^3), \subseteq \rangle$ from a denumerable

¹⁰ An *open sphere* is a set of points of R^3 whose distance from a given point of R^3 is less than some positive real number. A subset of R^3 is *open* if and only if each one of its points is contained in an open sphere which is in turn contained in the subset. A subset of R^3 is *closed* if and only if it is not open. The *closure* of a subset of R^3 is the intersection of all the supersets thereof which are closed. The *interior* of a subset of R^3 is the union of all the subsets thereof which are open. Finally, a *regular open set* of R^3 is a subset of R^3 which coincides with the interior of its own closure (Monk 1989, 1242). The regular open sets have been intuitively described as those open sets which have no “cracks” or “pinholes” (Bell 1985, 3).

¹¹ Here the *regularization* of a set of points is simply the interior of its closure (see preceding footnote). Thus, the *regularized intersection* of a family of regular open sets is the interior of the closure of their intersection, while the *regularized union* of a family of regular open sets is the interior of the closure of their union.

¹² Open spheres have been defined in footnote 10 above.

base will play a crucial role in the arguments for the central thesis of this paper, as we will now see.

4. The mereology of masses

We are finally in a position to state precisely the main claim of this paper. It is that the mereology of masses that semanticists have invoked in their interpretation of the mass nouns of natural language is isomorphic to the mereology of solids, and hence to its correlate in ordinary geometry —the mereology of regular open sets of three dimensional Euclidean space. More succinctly, our hypothesis is (8), where $\langle M, \leq_M \rangle$ is the mereology of masses and $\langle S, \leq_S \rangle$ is the mereology of solids.

$$(8) \quad \langle M, \leq_M \rangle \cong \langle S, \leq_S \rangle$$

Since isomorphic mereologies are structurally indistinguishable, the mereology of solids (or the mereology of regular open sets, if you prefer) will provide us with a complete characterization of the mereological structure which articulates the masses of any universe of discourse which has them.

To begin with, (8) will provide us with a precise measure of the size of the mereology of masses. For, notice that the mereology of solids has exactly 2^{\aleph_0} elements.¹³ Hence by (8),

$$(9) \quad \text{The mereology of masses has exactly } 2^{\aleph_0} \text{ elements.}$$

It should be noted that (6) and (9) imply that the mereology of masses has the smallest cardinality an atomless mereology may have. The claim in (8) thus places the mereology of masses among the simplest atomless mereologies —at least as far as cardinality is concerned.

Furthermore, if grammars are finitary formalisms, then the result in (9) implies that no grammar can provide individual names for all the masses of a universe of natural discourse. It should be clear that finitary grammars, even the ones which have recursive procedures at their disposal, can provide at most \aleph_0 individual constants.

¹³ Let $\langle A, \leq \rangle$ be a Boolean algebra. Some subset B of A^+ is *dense* in $\langle A, \leq \rangle$ if and only if for every $a \in A^+$ there is some $b \in B$ such that $b \leq a$ (Monk 1989, 54). Now, notice that the set of rational spheres is both denumerably infinite and dense in the algebra of regular open sets of R^3 (or in the algebra of solids). But if some set B is dense in some Boolean algebra $\langle A, \leq \rangle$, then A has at most $2^{|B|}$ elements (Monk 1989, 55). The algebra of regular open sets of R^3 (or the algebra of solids) thus has *at most* 2^{\aleph_0} elements. Hence, by (4), so does the mereology of solids. But (6) stated that every atomless mereology has *at least* 2^{\aleph_0} elements. The mereology of solids must therefore have *exactly* 2^{\aleph_0} elements.

5. The interpretation of mass nouns

The claim in (8) allows us to provide intuitive interpretations for the individual mass nouns of language. Consider first the mass noun *space*. It seems natural to claim that the noun *space* denotes the set of regions of space. But the regions of space are the solids of space. The particular noun *space* would thus denote a mereology which is isomorphic to the entire mereology of masses.

Let us suppose next that we were to suppress from the world everything but the water it contained. We would obtain a vast region of space filled with the contents of oceans, lakes, rivers, pipes, pitchers, glasses, drops, and so on. Notice now that this region may be regarded as a solid. If, as is generally assumed, the noun *water* denotes the set of portions of water in the world, then this noun will be interpreted as the set of portions of water occupying this solid. This set of portions clearly forms a mereology when taken in conjunction with the relation of inclusion (on the set of portions of water). In fact, this mereology would be isomorphic to the mereology of solids which are occupied by water. The common noun *water* thus denotes a mereology which is isomorphic to a particular mereology of solids —the mereology of solids occupied by water.

But it might be objected that the mereology of solids is atomless while the mereology of water is atomistic, its atoms being the H₂O molecules. Such an objection would miss, however, the essential nature of mass nouns —the fact that they fail to provide a criterion for the individuation of their reference (Quine 1960, 91). For indeed, it is the atomlessness of mass nouns that distinguishes them from their count counterparts and accounts for the fact that only the latter can withstand pluralization, and modification in terms of number, size, and shape:¹⁴

- (10) a. There are kings here.
 b. I saw a/one king.
 c. This is a big/tall king.

- (11) a. ?There are waters here.
 b. ?I saw a/one water.
 c. ?This is a big/tall water.

In fact, natural languages treat (their nouns for) *water* the same way they treat (their nouns for) *space* and *time*, where we would be hard put to find atoms. We conclude that the objection that solids and portions of water differ with respect to atomicity is not based on the semantic analysis of the noun *water* in ordinary English. Rather, it is based on the chemical analysis of water. Since the former yields the result that the denotation

¹⁴ The sentences in (11) are of course well formed if reference to natural kinds or standard units of water is intended. It is arguable, however, that when used in this way, *water* is not a mass noun anymore. Here water is individuated or atomized into kinds or units. We shall return below to the 'taxonomic' readings of reinterpreted mass nouns.

of the noun *water* is atomless, semanticists must interpret this noun in terms of an atomless mereology (while remaining neutral as to whether water, the chemical substance, is atomistic or atomless).

In any event, what has been said for *water* can be said for *earth, air, fire*, and indeed all mass nouns whose extensions are localized in space. Every mass noun whose extension is localized in space denotes a mereology which is isomorphic to a mereology of solids —the mereology of solids occupied by this extension.¹⁵ But then, as far as natural languages are concerned, *all* mass nouns can be localized in space. For, notice that we can in fact speak of the love in my heart, the anger in the inner city, the beauty in a canvas, the power in a motor, the weather in the tropics, and the time in the classroom. If every mass noun can be interpreted as a set of situated parts, then every mass noun denotation is isomorphic to the mereology of solids occupied by its extension.

Now, I am sure it could be argued that love is not located in my heart in the same way that a portion of water is contained in a bowl. This fact, however, seems to be irrelevant to actual linguistic usage, which uses the same grammatical devices to locate water in a bowl and love in my heart —for example a locative prepositional phrase headed by the preposition *in*. And this is certainly not an idiosyncrasy of English. In fact, I would be surprised if a natural language were found which used systematically different grammatical means to describe the location of material and immaterial entities.

But it is possible to provide further support for the claim that all mass nouns, even immaterial ones, are isomorphic to a mereology of solids. Consider the noun *time*. Intuitively, it denotes not the solids of a space, but rather the (series of) intervals of a line, namely the linear and continuous series of moments of time.¹⁶ So let us take this intuition about the semantics of *time* seriously and assume that the noun *time* indeed denotes the set of series of open intervals of a line R . This set forms a mereology with the relation of inclusion on the interval series of R —the *mereology of series of time intervals*. But this mereology is the intuitive counterpart of $\langle RO(R)^+, \subseteq \rangle$, the mereology of the regular open sets of R under subset inclusion. By (4), we may turn this mereology into a complete Boolean algebra by adding to it the empty interval. If we do so we would obtain $\langle RO(R), \subseteq \rangle$, the algebra of the regular open sets of R . Interestingly, this algebra can be reconstructed as follows.

We may assume that R can be placed in a one to one, order preserving, correspondence with the line of real numbers. Place R in any such correspondence. Consider then the set of open intervals of R whose endpoints correspond to rational numbers. Call this set B . Since the rational numbers are denumerable, so is B . Construct now the set of *finite*, regularized, unions of the intervals of B . This set forms

¹⁵ If solids correspond to open sets, then no mass would contain its boundaries. This surprising conclusion seems to be supported by linguistic fact, as it is possible to say of two masses that they touch without overlapping. If masses were to contain their boundaries, the density of space would render true mass contact (= contact without overlap) impossible.

¹⁶ It should be clear that a time is a series of intervals rather than a single interval (cf. *my time at the piano, your time in bed*, and so on). This of course does not force times to be scattered individuals, as a series may contain just one term. It only allows a time, like a solid, to be scattered.

a Boolean algebra when taken in conjunction with the relation of subset inclusion of interval series. In fact, it forms a Boolean algebra which is both atomless and denumerable, but not complete. The algebra is atomless because every open interval will contain nonempty open intervals other than itself. It is denumerable because it has been designed to contain only the *finite* regularized unions of denumerably many open intervals. It is not complete, however, as it does not contain the *infinite* regularized unions of the intervals in B . Suppose, then, that we were to complete this algebra. We would obtain $\langle RO(R), \subseteq \rangle$.

The preceding reconstruction of $\langle RO(R), \subseteq \rangle$ is interesting because it allows us to show that this algebra is isomorphic to algebra $\langle RO(R^3), \subseteq \rangle$ of regular open sets of R^3 . For, notice first that any two Boolean algebras which are atomless and denumerable are isomorphic (Monk 1989, 74). Notice further that all the completions of any given Boolean algebra are isomorphic as well (Monk 1989, 60). Now, since $\langle RO(R), \subseteq \rangle$ is the completion of the atomless denumerable algebra of rational intervals of R and $\langle RO(R^3), \subseteq \rangle$ is the completion of the atomless denumerable algebra of rational spheres of R^3 , then $\langle RO(R), \subseteq \rangle$ and $\langle RO(R^3), \subseteq \rangle$ must be isomorphic to each other.

Let us delete now the null elements from $\langle RO(R), \subseteq \rangle$ and $\langle RO(R^3), \subseteq \rangle$. By (4) we obtain two mereologies, namely the mereology $\langle RO(R)^+, \subseteq \rangle$ of regular open sets of R , and the mereology $\langle RO(R^3)^+, \subseteq \rangle$ of regular open sets of R^3 . Naturally, these two mereologies will be isomorphic. Hence, so will their intuitive counterparts, namely the mereology of series of time intervals and the mereology of solids.

It follows that the intuitive interpretation of space as the mereology of solids is isomorphic to the intuitive interpretation of time as the mereology of series of time intervals. Two seemingly disparate linguistic intuitions can thus be unified by the claim in (8).

It should go without saying that this does not mean that series of time intervals are solids or that solids are series of time intervals. It just means that solids and series of time intervals have the same mereological structure. The claim in (8) thus provides a frame in which the semantics of *space* and *time* can be unified; it supports a unified account of the semantics of *space* and *time*.

The isomorphism between $\langle RO(R), \subseteq \rangle$ and $\langle RO(R^3), \subseteq \rangle$ is interesting also because Bunt (1985, 66ff) used $\langle RO(R), \subseteq \rangle$ as a model for his *continuous ensembles*—the structures he proposed to interpret mass nouns. The mereologies of masses we have proposed are, therefore, (the nonnull portions of) some continuous ensembles. It should be added, however, that Bunt attached no importance to the particular structure of $\langle RO(R), \subseteq \rangle$. For him $\langle RO(R), \subseteq \rangle$ was just an example of a continuous ensemble—and thus a means for establishing the *consistency* of his theory of continuous ensembles. But any other of the myriad atomless and complete Boolean algebras could have served the same end.

6. The homogeneity of mass reference

The claim in (8) has allowed us to interpret every mass noun, material or otherwise, as a mereology which is isomorphic to a mereology of solids —namely the mereology formed by all the solids contained in a particular solid (and the relation of solid inclusion). The question thus arises as to whether we can characterize the structures of these mereologies. It is to this question that we will now turn.

First, let us point out that the mereology formed by all the solids contained in some solid s is a rather natural object. It may be called *the relativization of the mereology of solids to s* . Next, let us say that a mereology is *homogeneous* if and only if it is isomorphic to each one of its relativizations. It should be clear that homogeneity is hardly a trivial property of mereologies. No mereology with more than one atom can, for example, be homogeneous (Monk 1989, 681).¹⁷ Yet, the mereology of solids is demonstrably homogeneous.¹⁸ It is therefore isomorphic to each one of its relativizations. It follows that the mereology formed by all the solids contained in a particular solid is isomorphic to the mereology of solids as a whole.

Now, since we have argued that every mass noun denotes a mereology of masses which is isomorphic to the mereology of solids occupied by these masses, every mass noun denotes a mereology which is isomorphic to the entire mereology of solids. Since the latter is by (8) isomorphic to the mereology of masses, a further application of the transitivity of isomorphy entails that

- (12) Every mass noun denotes a mereology which is isomorphic to the mereology of masses as a whole.

But the relation of isomorphism is symmetric as well as transitive. Hence,

- (13) All mass denotations will be isomorphic to each other.

And as corollary of the claim in (8) and the homogeneity of the mereology of solids we have that

- (14) The mereology of masses is homogeneous.

Finally, since every element of the denotation of a mass noun determines a relativization of its own, the mereology of any portion of a mass will be isomorphic to the entire

¹⁷ An atom of a mereology $\langle M, \leq \rangle$ is any $m \in M$ such that, for all $n \in M$, $n \leq m$ implies that $m \leq n$.

¹⁸ Given the connections between $\langle S, \leq \rangle$, $\langle RO(R^3)^+, \leq \rangle$, and $\langle RO(R^3), \leq \rangle$, the claim follows given that (i) every Boolean algebra which is atomless and denumerable is homogeneous (Monk 1989, 682), that (ii) the completion of every homogeneous Boolean algebra is homogeneous (Monk 1989, 682), and (iii) that $\langle RO(R^3), \leq \rangle$ is the completion of a Boolean algebra which is atomless and denumerable (see Section 3).

mereology of masses. All masses therefore have the same mereological structure. Mass is 'self-similar' with respect to mereological structure. In a rather precise sense, then, mass nouns refer homogeneously.

7. The denumerability of mass partitions

The corollary in (14) and the result on the cardinality of the mereology of masses presented in (9) imply that every mass in the mereology of masses will contain 2^{\aleph_0} masses. It is therefore interesting to show that the main claim of this paper entails that no mass in the mereology of masses may have more than \aleph_0 nonoverlapping parts.

To make this point, let us say that a subset of a mereology is *pairwise disjoint* if no two elements thereof overlap. Let us furthermore say that a mereology satisfies *the countable chain condition* if and only if all its pairwise disjoint subsets are countable. Now, it is easy to show that the mereology of solids satisfies the countable chain condition.¹⁹ This means that, if the claim in (8) holds, then

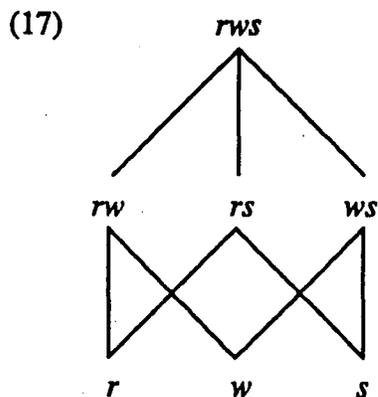
(15) The mereology of masses satisfies the countable chain condition.

In addition to its intrinsic interest for a theory of the mereology of masses, (15) imposes interesting bounds on the taxonomic reinterpretations of mass nouns. Consider for example a sentence like (16).

(16) This store carries all three wines: red, white, and rose.

As can be readily seen, the typically mass noun *wine*, seems to have been pluralized in this sentence. Interestingly, however, what has been pluralized in this sentence is not 'wine', but rather 'kind or variety of wine'. One way to account for this fact is to observe that the mereology of wine contains a subset which partitions the wine in the world in three: the portion of wine containing all the red wine, the portion of wine containing all white wine, and the portion of wine containing all rose wine. But the mereology of wine also contains all the combinations of these three kinds. In other words, it contains the substructure in (17).

¹⁹ Since R^3 has a countable base, the algebra of regular open sets of R^3 satisfies the countable chain condition (cf. Halmos 1963, 61). Hence so does the algebra of solids and its positive counterpart, the mereology of solids.



We may now claim that this substructure is what *wines* denotes in (16).²⁰

In light of the preceding discussion, the relevance of (15) for an account of the taxonomic reinterpretations of mass nouns should be clear. If a taxonomic reinterpretation denotes a structure based on the partition of a mass, then (15) imposes an upper bound on the cardinality of such a partition; it has to be denumerable. Furthermore, since the substructure denoted by a reinterpreted mass plural is generated by closing a basic partition under the operation of constitution defined in (2), the postulate of completeness in (1) imposes an upper bound on the size of this substructure; it cannot exceed the power of the continuum.

References

- Bell, John (1985) *Boolean-valued models and independence proofs in set theory* [= *Oxford Logic Guides*, 12]. Oxford, Clarendon Press.
- Bunt, Harry (1985) *Mass terms and model-theoretic semantics*. Cambridge, Cambridge University Press.
- De Laguna, Theodore (1922) "Point, line, and surface, as sets of solids". *The Journal of Philosophy* 19: 449–461.
- Halmos, Paul (1963) *Lectures on Boolean algebras* [= *Van Nostrand Mathematical Studies I*]. Princeton, Van Nostrand.
- Krifka, Manfred (1987) "Nominal reference and temporal constitution: towards a semantics of quantity" [= *Forschungsstelle für natürlich-sprachliche Systeme, Bericht 17*]. Universität Tübingen, Biesingerstr. 10, D-7400, Tübingen.
- Link, Godehard (1983) "The logical analysis of plurals and mass terms: a lattice-theoretical approach". In R. Bäuerle et al. (Eds.) *Meaning, use, and interpretation of language*. Berlin, Walter de Gruyter.
- Lønning, Jan Tore (1987) "Mass terms and quantification". *Linguistics and Philosophy* 10: 1–52.
- Mach, Ernst (1906) *Space and geometry*. La Salle, Open Court.
- Monk, Donald (1989) *Handbook of Boolean Algebras*. Three volumes. Edited by J. Donald Monk with the cooperation of Robert Bonnet. Elsevier, Amsterdam.

²⁰ See Ojeda (1993, Chapter 5) for more extended discussion.

- Nicod, Jean (1924) *La géométrie dans le monde sensible*. Paris, Alcan.
- Ojeda, Almerindo (1993) *Linguistic individuals* [= *CSLI Lecture Notes Series 31*]. Stanford, Center for the Study of Language and Information, Stanford University.
- Quine, Willard Van Orman (1960) *Word and object*. Cambridge, MIT Press.
- Sharvy, Richard (1980) "A more general theory of definite descriptions". *The Philosophical Review* 89: 607–624.
- Simons, Peter (1987) *Parts. A study in ontology*. Oxford, Clarendon Press.
- Tarski, Alfred (1956a) "Foundations of the geometry of solids". In *Logic, semantics, metamathematics*. Papers from 1923 to 1938 by Alfred Tarski. Translated by J. H. Woodger. Second edition, edited and introduced by John Corcoran, 1983. Indianapolis, Hackett Publishing Company, pp. 24–29.
- Tarski, Alfred (1956b) "On the foundations of Boolean algebra". In *Logic, semantics, metamathematics*. Papers from 1923 to 1938 by Alfred Tarski. Translated by J. H. Woodger. Second edition, edited and introduced by John Corcoran, 1983. Indianapolis, Hackett Publishing Company, pp. 320–341.
- Wald, Jan (1977) *Stuff and words: a semantic and linguistic analysis of nonsingular reference*. Ph. D. dissertation, Brandeis University.
- Whitehead, Alfred (1919) *An enquiry concerning the principles of natural understanding*. Cambridge, Cambridge University Press.
- Whitehead, Alfred (1920) *The concept of nature*. Cambridge, Cambridge University Press.

Defining A 'Doughnut' Made Difficult

N M Gotts

Division of AI, School of Computer Studies,
University of Leeds, England
email:ngotts@scs.leeds.ac.uk

Abstract

Qualitative descriptions of spatial properties and relationships, and qualitative spatial reasoning, are of fundamental importance in human problem-solving: even where we use quantitative approaches, these depend on an accompanying qualitative representation. The work described is aimed at formalising some topological aspects of qualitative spatial description and reasoning. The paper continues the work of Randell, Cohn and Cui on region-based qualitative representations of spatial properties and relations, built on the 'logic of connection' developed by Clarke.

It is shown how taxonomies of the topological properties and relationships of spatial regions can be developed, using the single primitive 'C', where 'C(x, y)' indicates that regions x and y are 'connected', meaning that their closures share at least one point. This is done by considering a specific task: deciding whether or not a region has the topology of a 'doughnut', or solid torus. A range of 'near misses' and doubtful cases is discussed, and ways of distinguishing these from the 'target', the solid torus, are considered. It is shown how the task could be performed given certain assumptions about the target region and about regions in general. These assumptions are then progressively relaxed; as this is done, the task requires the definition of successive layers of terminology, all derived ultimately from 'C', and providing a basis for successively broader taxonomies of topological properties and relationships.

1 Introduction

The title of this paper may call for a brief explanation. The 'doughnut' of the title is the topological entity also known as a 'solid torus', illustrated in figure 1. The paper investigates the task of defining this topological object, or distinguishing it from all topologically distinct objects, in terms of a particular formal system for qualitative representation of spatial properties and relations developed at the University of Leeds over the past few years and known as 'RCC-theory' (Randell, Cui and Cohn 1992, Cohn, Gotts, Randell, Cui, Bennett and Gooday 1993, Cohn, Randell and Cui 1994). How difficult this task is, of course, depends upon the range of possibilities

that must be excluded. The approach taken is to investigate the task at first within a fairly restricted set of possibilities, and then to make the task increasingly difficult by successive expansions of the set of alternatives to be excluded.

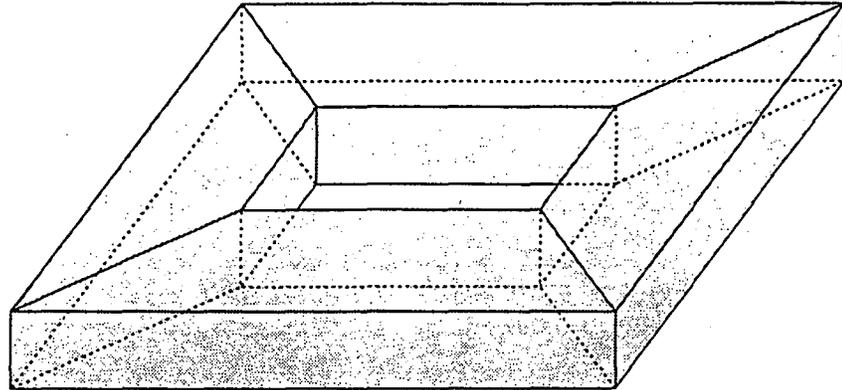


Figure 1: The Target: a 'Doughnut', or Solid Torus

It is the topology of the object shown in figure 1 that is of interest, not its other spatial properties, nor what it is made of: the doughnut pictured there is rectilinear simply because this was the easiest way to draw it with the graphics package used, and the internal partitions shown by dotted lines are included for clarity only. Many everyday objects (including rings and bracelets, sewing needles, pipes and tubes, nuts and washers, and links in chains) are 'doughnuts' in the topological sense used here, and in many cases their functionality depends on this topology. As figure 2 shows, topological doughnuts can be of widely varying shapes (the object shown at the far right is intended to be 3-dimensional, and could be a doughnut in the everyday sense).

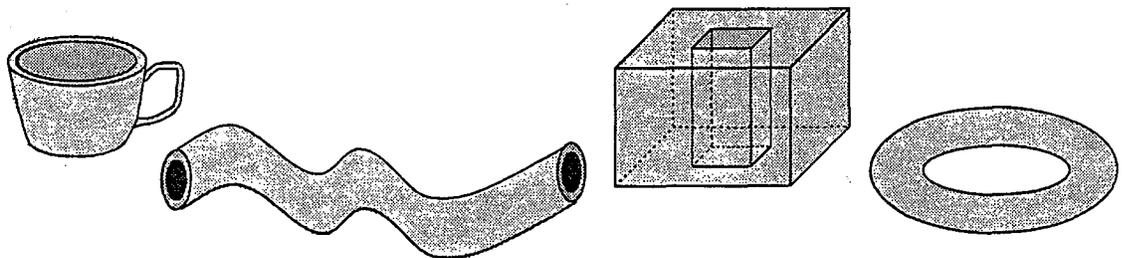


Figure 2: Doughnut Variations

This paper deals with the same set of issues as (Gotts 1994), but includes significant revisions of the line of argument developed there, and is intended to clarify some aspects of the approach taken, in part by the use of a larger number of figures. Following this introduction, section 2 deals with questions of motivation and background: in brief, why the problem tackled is of interest. Section 3 is essentially a description of the problem, explaining why it is difficult. Section 4 explains RCC-theory, as far as is required for the purposes of this paper, while section 5 is the core of the paper, systematically exploring the problem of defining a doughnut in

terms of RCC-theory, using what can be called the 'mystery region' approach. This is a kind of thought experiment: imagine that a computational system is asked to determine whether a 'mystery region' of space has the topology of a doughnut, by asking questions couched solely in terms of the 'C' primitive of RCC-theory. Section 6, ending the paper, briefly summarizes what has been said, and describes the intended direction of future work.

2 Motivation and Background

RCC-theory is so named after the investigators who developed it (D. Randell, A.G. Cohn and Z. Cui), but 'RCC' can also stand for 'Region-Connection Calculus': it takes extended regions rather than points as fundamental, and builds on the primitive relation of 'connection', symbolised C, between a pair of regions. $C(x, y)$ is read ' x is connected with y '. RCC's work is based on Bowman Clarke's (Clarke 1981, Clarke 1985) 'calculus of individuals based on connection', but considerations of computational efficiency and plausibility discussed below, and in more detail in (Randell et al. 1992) have led them to modify Clarke's calculus.

The importance of finding good ways to represent *qualitative* spatial properties and relations can be better appreciated by noticing that even spatial representations with high quantitative content, such as engineers' or architects' drawings, require the identification of distinct objects or regions, to which particular pieces of quantitative spatial information (lengths, widths, volumes, angles) are attached; such drawings must successfully depict a spatial *structure*, based on relations of inclusion, adjacency, linkage and so forth, if the measurements they also include are to be any use. Topology, with which this paper is concerned, is a fundamental aspect of spatial structure, but not the whole of it. The four doughnuts of figure 2 are topologically identical, but differ in their structure (the ways they can be divided into parts and how those parts inter-relate).

In recent papers (Cohn, Randell, Cui and Bennett 1993, Cohn et al. 1994), RCC's main concerns have moved toward computational issues, and the exploitation of an additional primitive concerning convexity introduced in (Randell and Cohn 1989). The purpose of this paper, by contrast, is to investigate the *expressive power* of the C predicate: how much can be expressed using this single predicate? This is most easily done in terms of specific tasks, and that of specifying the topological properties of a solid torus in terms of C alone was chosen as non-trivial but possibly achievable (as will be seen, it is not quite achieved here).

The work reported is related to that of Casati and Varzi on holes (Casati and Varzi 1994), but shows that two of their three types of holes (tunnels and internal voids) can be described in terms of C — these authors use C, but introduce the relationship between a hole and its 'host' as an additional primitive. Vieu (Vieu 1993) and her colleagues have also used Clarke's work as a basis for spatial reasoning.

In representing either spatial or temporal relations, a decision arises: whether to treat points, extended regions, or both as primitives (in the temporal case, points and regions are often called instants and intervals respectively). In a point-based

approach, regions are defined as sets of points; in a region-based approach, points may be defined in terms of sets of regions, or omitted altogether. An advantage of point-based representations is that they can use the extensive work of mathematicians on point-set topology. Three countervailing advantages have been suggested for a region-based ontology: that regions are somehow 'closer to perception' and therefore psychologically primitive; that expressing spatial relations in terms of regions is a useful form of abstraction, paralleling the abstraction from the real-number line to landmark values and intervening intervals exploited in work on qualitative physics (Weld and De Kleer 1990); and that the point-based approach gives rise to counterintuitive distinctions such as that between closed and open regions (those that do and do not include their own boundaries), and weird constructions such as space-filling curves that cross themselves at every point. The first of these claims is dubious, taking into account the constructive nature of perception: regions, as much as points, are commonsense-theoretical entities in terms of which we interpret information from the senses. However, the second and third give sufficient reason to investigate the region-based alternative.

So far as the second is concerned, the descriptions of spatial configurations we use in everyday life, and even in technical contexts, are frequently expressed in terms of relations between extended regions, whose extension in terms of points is left unspecified, and may indeed be unspecifiable. Work in progress at Leeds, reported in the paper by Cohn and Gotts in this volume, indicates that spatial indeterminacy and vagueness may be better represented using regions rather than points as a basis, and develops a representation of regions with indeterminate boundaries based on RCC-theory.

Turning to the claim that point-based approaches have counter-intuitive consequences, it seems odd — as noted in (Randell et al. 1992) — that two regions can be distinct, yet take up exactly the same part of space, as an open region and its closure do if we allow regions to be either open or closed sets of points. Also, if we take point-sets as fundamental we must allow for those that are neither closed nor open, including *some but not all* of their boundary points, and for oddities such as 2-dimensional point-sets including all those points within an area except those with rational coordinates. Furthermore, if we consider a physical object as occupying a region, is that region closed or open? In the case of a solid object it might seem intuitive to regard the occupied region as containing its boundary, and thus as closed. However, if two such objects touch, this suggests either that one is open and one closed (at least at their common boundary), or that both are closed and share their boundary points. Neither solution is obviously wrong, but both give rise to unease. In RCC-theory's region-based approach, this problem is avoided by dropping the open/closed distinction (which Clarke's system retained).

Advocacy of region-based rather than point-based spatial and temporal representation has connections with a broader opposition to the view that set theory and predicate logic provide an adequate basis for the formal representation of the world. Those logicians and philosophers who have worked on the alternative or supplementary approaches ('mereology' or 'calculus of individuals') include Whitehead (Whitehead 1929), Leśniewski (originator of the term 'mereology'), Tarski (Tarski

1956), Leonard and Goodman (Leonard and Goodman 1940), Clarke (referred to above), and recently Simons (Simons 1987), Varzi (Varzi 1993), and Smith (Smith 1994). Simons reviews much of the earlier work in this area.

'Individuals' in the sense used in 'calculus of individuals', are whatever singular things we take the world to contain. In this wide context it has been argued (Varzi 1993) that we should allow entities to have the same spatio-temporal extent without being identical or having any parts in common. In the current context, where we are concerned with relations between spatial regions, we shall assume an extensional approach: if exactly the same spatial relations apply to region x as to region y , then x and y are the same region.

3 Description of the Problem

In this section, we look at the problem of 'defining a doughnut' in more detail. We adapt an approach pioneered by Winston (Winston 1975, Ch.5), that of the 'near miss'. Winston advocated the use of both examples and counterexamples of concepts such as 'arch' to teach a vision program to interpret a blocks-world picture correctly: the counterexamples used, however, were intended to be 'near misses': objects that have most of the properties of an arch, but not all. Similarly, since we want our imaginary computational system to decide whether a 'mystery region' is or is not a doughnut, we need to think about the types of region that might confuse it, and indeed, to decide for ourselves exactly what we do and do not want to count as a doughnut. (Note that we are not — at least at present — thinking of a system that can distinguish doughnuts from non-doughnuts visually; rather, we want the system to find out whether or not we have a doughnut in mind by asking appropriate questions.)

Each of the next six figures shows a set of 'near misses'. Finally, we have a figure showing a selection of doubtful cases, where it is less clear that we want to exclude the regions pictured from our concept of a doughnut.

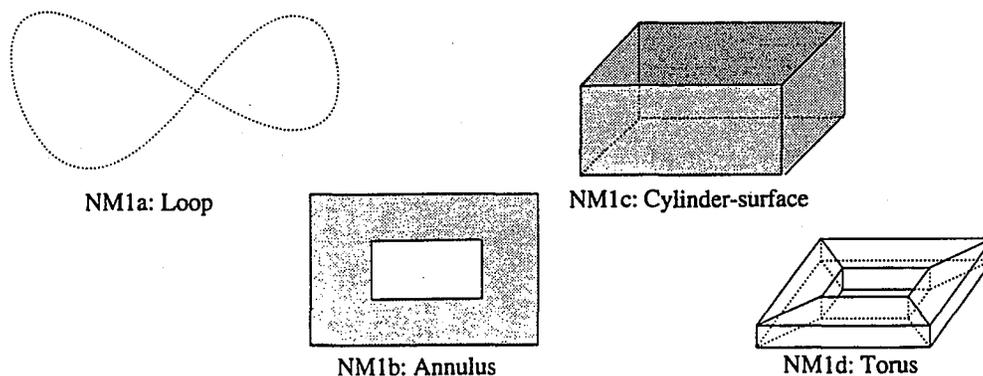


Figure 3: Near Misses: 1

All the examples of figure 3 fail to be doughnuts because they are of the wrong dimensionality. The top left figure (NM1a) is intended to be a simple closed curve

embedded in 3-dimensional space: imagine a loop of string or wire made thinner and thinner until it reaches zero thickness. The annulus (NM1b) and cylinder surface (NM1c) are topologically identical. The torus (NM1d) is of course the *surface* of the solid doughnut which we want to be able to identify. All these near misses (the set will be referred to collectively as 'NM1') share the doughnut's property of having a single hole through them, allowing two opposite directions of travel around the hole to be defined, and in some cases permitting them to be *linked* to others of their kind. This is important: the doughnut, annulus and loop could be successive *idealizations* or simplified models of an object or region such as a road circuit, a moat, or a link in a chain.

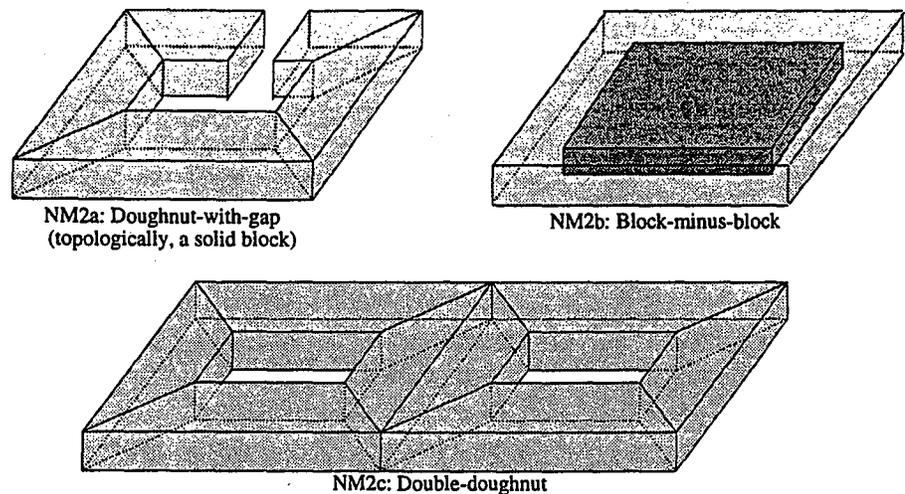


Figure 4: Near Misses 2

The three regions depicted in figure 4, and the four shown in figure 5 (NM2 and NM3) are all homogeneously 3-dimensional regions (unlike some of those in later figures) that differ from the doughnut itself in quite straightforward ways. Nevertheless, they do raise some points of interest. The 'doughnut with gap' may, if the gap is narrow enough, be *functionally* equivalent to a doughnut in some contexts (think of current jumping a small gap in a metal doughnut, or of a link in a heavy metal chain). Topologically, however, it is equivalent to a solid block (normally the term 'solid ball' would be used for this simplest of finite-diameter 3-dimensional topological objects; 'solid block' is used here simply because all objects or regions are shown as rectilinear). The 'double doughnut' shares the potential to link with others of its kind, and other functional properties, with the doughnut itself. The 'block-minus-block' provides an opportunity to explain the convention used for representing hollow objects: the internal void is shaded in dark grey and shown as opaque, the solid volume remaining is shaded in light grey and shown as transparent. It also points up an interesting relationship between 2- and 3-dimensional regions. Frequently, one can identify 3-dimensional 'counterparts' of 2-dimensional entities: a solid ball or block is in this sense the counterpart of a disc, and (moving beyond topology to metric aspects of geometry), a cube the counterpart of a square. Both the doughnut and the block-minus-block, however, have a reasonable claim to

be the 3-dimensional counterpart of the annulus: both can be described as a block that 'hosts' a single hole (c.f. (Casati and Varzi 1994)), just as an annulus can be described as a disc that hosts a single hole; but in three dimensions, a topologically significant hole (as opposed to a mere surface depression) can be either a hole right through the object (the doughnut), or an internal void (the block-minus-block). In one important respect, the block-minus-block has a better claim to be the counterpart of the annulus than the doughnut: like the annulus, it has two separate boundaries. In terms of connectivity, however, the opposite is true: in both the annulus and the doughnut, one can embed a simple closed curve that cannot be shrunk to a point while remaining in the embedding figure — which is not the case for the block-minus-block.

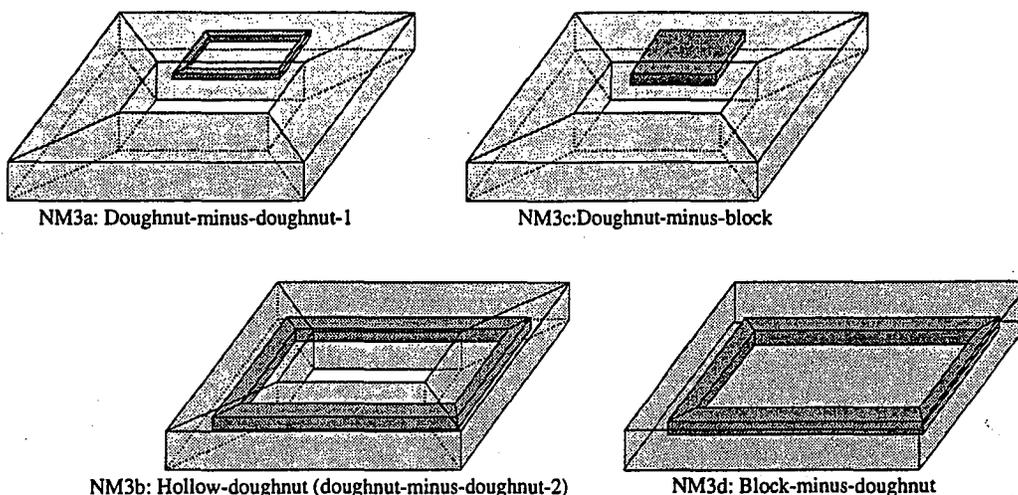


Figure 5: Near Misses 3

All the NM3 near misses have at least one toroidal boundary surface, like that of the doughnut — but all have two boundaries, outer and inner, like the annulus and block-minus-block. Note the two different ways shown of 'subtracting' one doughnut from another, making a doughnut with a 'doughnutiform' internal void. There are in fact infinitely many different ways of doing this: the internal void may go more than once round the hole through the containing doughnut, and may be knotted in an infinite number of different ways.

In figures 6 and 7, we see objects or regions which are 3-dimensional, but not homogeneously so (we will go into this distinction in more detail in section 5; for now, we just note that all of the regions shown in these figures have anomalous points on their surfaces, where the surface does *not* simply divide the shaded region from the rest of space). All the regions in figure 6 (NM4) in some sense 'approach' the doughnut-with-gap of NM2a. Those on the right seem in some way more 'deviant' than those on the left (having lower-dimensional parts rather than just 1-dimensional or 0-dimensional — pointlike — connections between parts), and in fact will turn out not to be 'regions' at all in the intended interpretation of RCC-theory (see section 5). All the NM4 near misses, like the NM1 objects, can host 'unshrinkable' loops, if we allow these to include surface as well as interior points. Notice that attaching one

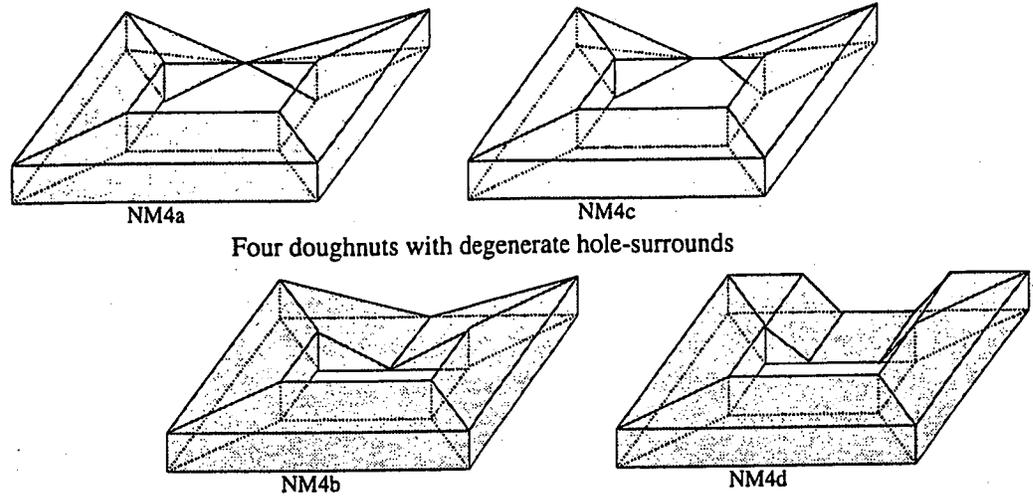


Figure 6: Near Misses 4

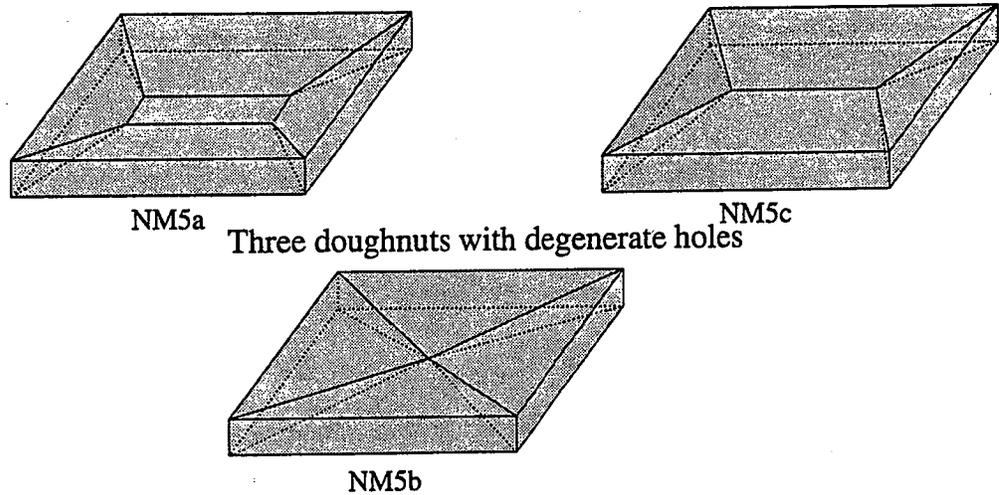


Figure 7: Near Misses 5

of these entities to a genuine doughnut (in the way that two doughnuts are attached to produce the double doughnut of NM2) gives another range of near misses, while attaching the doughnut-with gap to a doughnut in this way produces a doughnut.

Turning to figure 7 (NM5), NM5a seems more 'deviant' than the others, and again will turn out not to be an RCC-theory region at all in our chosen interpretation. Attaching NM5b to the middle of a larger flat block produces (topologically speaking) a block minus a block, with the two blocks having one common surface point: a problem region mentioned in section 5 of the paper. When embedded in Euclidean 3-space (E^3), all NM5 objects share the property that there are *interior* loops that cannot be continuously shrunk to a point while keeping the shrinking loop within the interior throughout the shrinking process.

The final near miss, shown in figure 8 (NM6), is E^3 minus a doughnut: all of an infinite 3-D space *other* than the opaque doughnut. Like the doughnut itself, it has just one surface, which is toroidal.

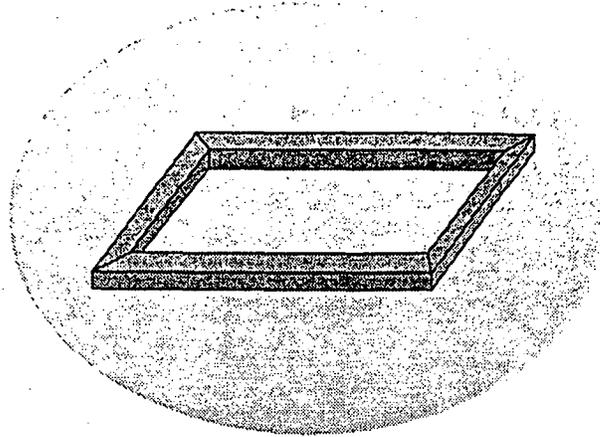
NM6: $\overset{3}{E}$ minus doughnut

Figure 8: Near Misses 6

Finally in this section, we come to a set of 4 doubtful cases. We will see that of the four, all but the 'doughnut with knotted hole' are in fact topologically equivalent to the doughnut. However, the *complements* of the other three objects, when they are embedded in E^3 , do differ topologically from the corresponding complement of the ordinary doughnut. This illustrates the distinction between *intrinsic* and *extrinsic* topological properties of a region: the latter, but not the former, are dependent on how the entity is embedded in a containing space.

4 'Connection' Between Regions as a Basis for Spatial Representation and Reasoning

RCC's basic theory, as presented in (Randell et al. 1992, section 4) uses two axioms establishing that C is reflexive and symmetric:

$$\forall x C(x, x) \\ \forall x, y [C(x, y) \rightarrow C(y, x)],$$

plus metalinguistic definitions of quasi-Boolean functions¹ guaranteeing the existence of complements of regions, and of the sum, product and difference of ordered pairs of regions (within restrictions explained below). Additional relations are defined in terms of C: DC(x, y) (x is disconnected from y), P(x, y) (x is part of y), PP(x, y) (x is a proper part of y), EQ(x, y) (x coincides with y), O(x, y) (x overlaps with y), PO(x, y) (x partially overlaps with y), DR(x, y) (x is discrete from y), EC(x, y) (x is externally connected with y), TPP(x, y) (x is a tangential proper part of y), and NTPP(x, y) (x is a non-tangential proper part of y). P is nonsymmetric, and PP, TPP and NTPP are asymmetric; their inverses will be symbolised here as PI, PPI, TPPI and NTPPI. The eight relations DC, EC, PO, TPP, NTPP, TPPI,

¹Quasi-Boolean rather than Boolean because there is no null region.

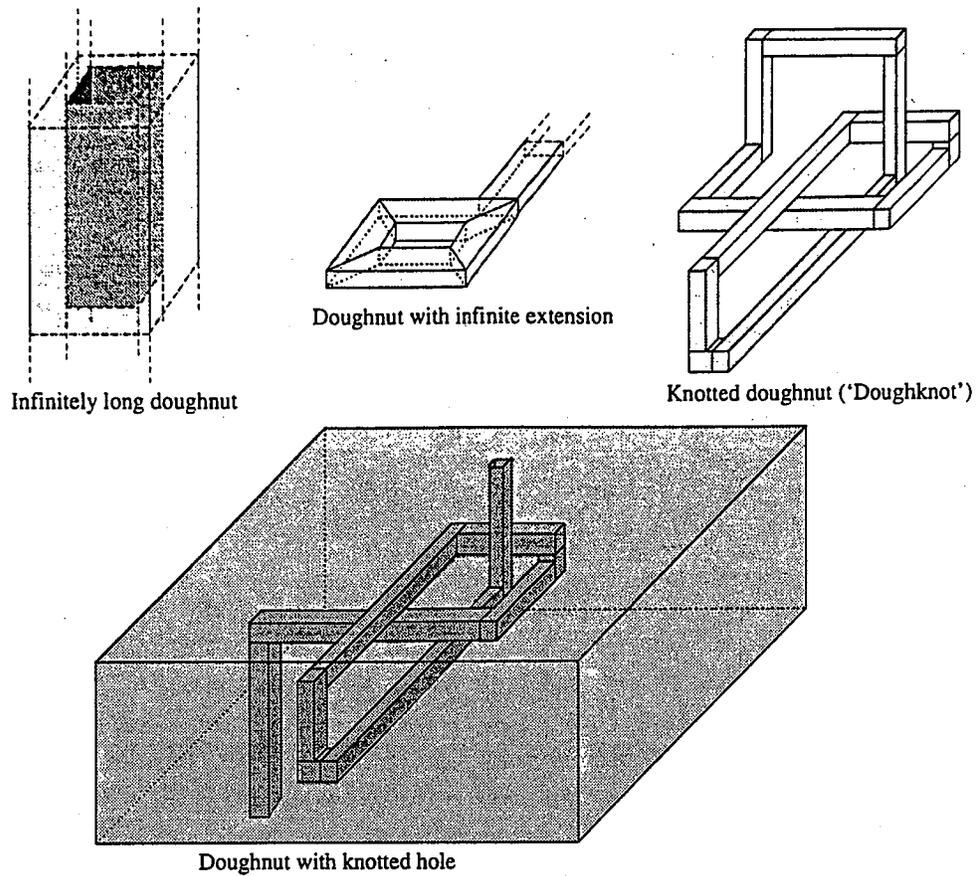


Figure 9: Doughnut Or Not?

NTPPI and EQ constitute a pairwise exclusive and jointly exhaustive set of 'base relations': exactly one of the eight must hold between an ordered pair of regions.

The definitions of the additional relations used in this paper are given below; the remainder can be found in (Randell et al. 1992).

$$\begin{aligned}
 DC(x, y) &\equiv_{def} \neg C(x, y) \\
 P(x, y) &\equiv_{def} \forall z [C(z, x) \rightarrow C(z, y)] \\
 PP(x, y) &\equiv_{def} P(x, y) \wedge \neg P(y, x) \\
 EQ(x, y) &\equiv_{def} P(x, y) \wedge P(y, x) \\
 O(x, y) &\equiv_{def} \exists z [P(z, x) \wedge P(z, y)] \\
 EC(x, y) &\equiv_{def} C(x, y) \wedge \neg O(x, y) \\
 TPP(x, y) &\equiv_{def} PP(x, y) \wedge \exists z [EC(z, x) \wedge EC(z, y)] \\
 NTPP(x, y) &\equiv_{def} PP(x, y) \wedge \neg \exists z [EC(z, x) \wedge EC(z, y)]
 \end{aligned}$$

The universal region (U_s), and the quasi-Boolean functions, introduced via explicit metalinguistic definitions in (Randell et al. 1992), are here defined implicitly using additional object-language axioms, which have the same effect: asserting the existence of a region (given that of one or more others, except in the case of the U_s axiom). The functions concerned are $compl(x)$ (the region-complement of x , de-

defined as a region only when x is not Us); $sum(x, y)$ (the region-sum of x and y); $prod(x, y)$ (the region-product or intersection of x and y , defined as a region only when $O(x, y)$); and $diff(x, y)$ (the region-difference of x and y , defined as a region only when $\neg P(x, y)$). The complications arising because applying one of these functions to certain arguments does not produce a region are dealt with using a sorted logic, LLAMA (Cohn 1992). The functions $compl(x)$, $diff(x, y)$ and $prod(x, y)$ are partial over the domain of regions, but are rendered total within LLAMA by introducing a sort NULL, disjoint from the sort REGION, and specifying sortal restrictions on the functions' arguments. The axioms below depend on the assumption that these restrictions are applied.

$$\begin{aligned} & \forall x [C(x, Us)] \\ & \forall x, y [[C(y, compl(x)) \equiv \neg NTPP(y, x)] \wedge [O(y, compl(x)) \equiv \neg P(y, x)]] \\ & \forall x, y, z [C(z, sum(x, y)) \equiv C(z, x) \vee C(z, y)] \\ & \forall x, y, z [C(z, prod(x, y)) \equiv \exists w [P(w, x) \wedge P(w, y) \wedge C(z, w)]] \\ & \forall x, y [NULL(prod(x, y)) \equiv DR(x, y)] \\ & \forall x, y, z [C(z, diff(x, y)) \equiv C(z, prod(x, compl(y)))] \end{aligned}$$

Finally, the following axiom is added, to establish that all regions have an NTPP, ruling out 'atomic' regions:

$$\forall x \exists y [NTPP(y, x)].$$

This suffices to prove that all regions have an infinite number of NTPPs.

5 Defining a 'doughnut' in terms of connection

Given a 'mystery region' r , what questions could a system that knows about nothing other than C ask to determine whether r is a doughnut, distinguishing it from the near misses illustrated, and from any others? This depends, as already noted, on the range of possibilities from which it must be picked out. We can exclude a near miss in 3 ways: by adding more axioms; by specifying conditions, expressed in terms of C , that a doughnut must meet but other regions need not; or by placing limitations on the permitted range of *interpretations* of the axioms.

Throughout this paper, the universal region Us is assumed to be an N -dimensional manifold (i.e., every point in it has a neighbourhood topologically equivalent to an N -dimensional disc, N being the same for all points), and each region is assumed to be an open set of points within Us , equal to the interior of its closure (a 'regular set')². We also assume that the closure of the open set of points (the *region-closure*) is *triangulable*. Two regions are connected — $C(x, y)$ — iff their region-closures share at least one point. We will call this restriction on the intended interpretation

²There are slightly different interpretations of 'region' which give an isomorphic set of regions — e.g. identify a region with the set of point-sets giving the same result when interior and then closure are taken.

of RCC-theory *assumption 0*. All the near misses of NM1, NM4c, NM4d, and NM5a are thereby excluded, as they do not correspond to regular open sets.

In addition to the fundamental assumption above, we initially adopt the following set of assumptions:

- 1. U_s , the universal region, is an infinite 3-dimensional Euclidean space (E^3).
- 2. The region-closure of each region is a *locally Euclidean space*: one in which each point has a neighbourhood topologically equivalent either to an N -dimensional disc, or to half of an N -dimensional disc, with N being the same for all points³.
- 3. The 'mystery region', region r , is of finite diameter.

These assumptions will be progressively weakened in the course of this section.

NM4c, NM4d, NM5a, and all those in NM1, would be excluded from the set of regions by the combination of the non-atomic axiom and assumption 1, even if they were not already excluded by assumption 0. NM4a's and NM4b's interiors are manifolds, but their closures are not locally Euclidean spaces; and the same is true of NM5b and c. NM6 and two of the doubtfuls are excluded by condition 3. This leaves just those near misses in NM2 and NM3, and the doughknot and doughnut-with-knotted-hole in figure 9. How, if at all, can we distinguish the doughnut from each of these by using C ?

So far as the NM2 cases are concerned, in conventional topological terms these all differ from the doughnut in connectivity. Using C as a starting point, we can define a similar concept, 'finger-connectivity', which will serve to distinguish all three NM2 near misses from the doughnut.

The first step is to define a self-connected (CON) region, one which cannot be divided into two DC parts. We can define the predicate CON thus:

$$\text{CON}(x) \equiv_{\text{def}} \forall y \forall z [\text{EQ}(x, \text{sum}(y, z)) \rightarrow C(y, z)].$$

(In terms of point-set topology, with the interpretation of 'region' used here, this means the *region-closure* is self-connected.) We can also define the 'separation-number' of a region: the minimum number of CON parts into which it can be divided (upper-case letters are used to stand for variables ranging over the natural numbers, and '+' has its normal arithmetical meaning when applied to these numbers):

$$\begin{aligned} \text{SEPNUM}(r, 1) &\equiv_{\text{def}} \text{CON}(r) \\ \text{SEPNUM}(r, N+1) &\equiv_{\text{def}} \exists s, t [\text{EQ}(r, \text{sum}(s, t)) \wedge \text{DC}(s, t) \wedge \text{CON}(s) \wedge \text{SEPNUM}(t, N)]^4. \end{aligned}$$

³A 1-dimensional disc is an open line-segment; a 3-dimensional disc is an open solid ball or block.

⁴Being recursive, this does not have the form of a normal definition, and, strictly speaking, requires an axiomatisation of the natural numbers to be added to RCC-theory. However, wherever SEPNUM is used in this paper, it could if required be 'unpacked' to a form using only C , and variables ranging over regions: we could define SEPNUM1, SEPNUM2 and so forth, as far as we required. The same is true for other terms defined using SEPNUM.

Finger-connectivity is then defined, using SEPNUM, in terms of the possible *dissections* of a CON region (for convenience, we here define the finger-connectivity of a non-CON region as 0; in (Gotts 1994) it was left undefined). A dissection is a division of a region into a finite number of CON parts, which are non-overlapping, and jointly exhaustive of the region. The *dissection-graph* corresponding to a dissection is defined as follows: a node of the graph corresponds to each piece of the dissection, and two nodes are joined (by a single link) iff the corresponding parts are connected. The finger-connectivity of a CON region identifies the largest of a specific family of dissection-graphs which it can host. This family is of graphs with $N + 2$ nodes ($N \geq 1$), with two of the nodes distinguished from the rest. These two are not linked, while each of the other N nodes has a link to each of them, but none to each other. The dissection-graph can be drawn like two N -fingere hands with pairs of corresponding fingertips touching (figure 2). A line-segment, disc or solid ball has finger-connectivity 1, a circle, annulus, torus or doughnut finger-connectivity 2, a 2-hole torus or doughnut 3, and so forth. Formally, FCON is defined thus:

$$\begin{aligned}
 \text{FCON}(r, 0) &\equiv_{def} \neg \text{CON}(r) \\
 \text{FCON}(r, N) &\equiv_{def} \text{CON}(r) \wedge \\
 &\exists a, x, b [\text{EQ}(r, \text{sum}(a, \text{sum}(x, b))) \wedge \text{DC}(a, b) \wedge \text{EC}(a, x) \wedge \text{EC}(x, b) \wedge \text{SEPNUM}(x, N)] \wedge \\
 &\neg \exists a, y, b [\text{EQ}(r, \text{sum}(a, \text{sum}(y, b))) \wedge \text{DC}(a, b) \wedge \text{EC}(a, y) \wedge \text{EC}(y, b) \wedge \\
 &\text{SEPNUM}(y, N + 1)].
 \end{aligned}$$

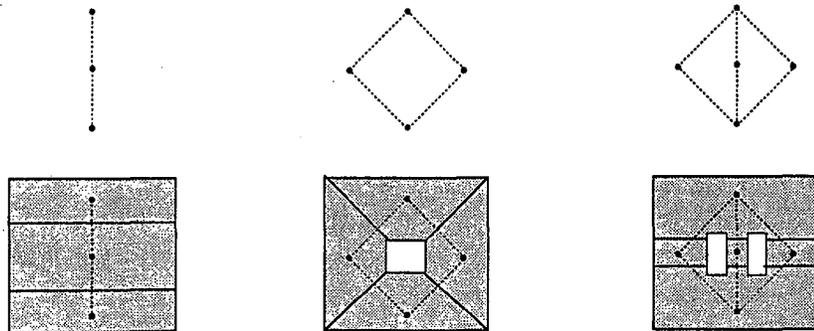


Figure 10: Dissection-Graphs And Dissections: Finger-Connectivities 1, 2 and 3

Of the near misses in NM2, the block-minus-block and doughnut-with-gap will not host the second of the three dissection-graphs (they are of FCON 1), and the double doughnut is of FCON 3). In NM3, the block-minus-doughnut, doughnut-minus-block, and hollow doughnut are all FCON 2 — like the doughnut — but doughnut-minus-doughnut-1 is of FCON 3. All the doubtful cases of figure 9 are FCON 2. So how can we eliminate the remaining three near misses from NM3? Can we eliminate the doubtful cases?

Under assumptions 1-3 above, we can do so by defining a predicate that specifies

how many separate boundaries two EC regions share⁵:

$$\begin{aligned} \text{SBNUM}(r, s, N) \equiv_{def} & \text{EC}(r, s) \wedge \\ & \exists x[\text{PP}(x, r) \wedge \text{DC}(\text{diff}(r, x), s) \wedge \text{SEPNUM}(x, N)] \wedge \\ & \neg \exists y[\text{PP}(y, r) \wedge \text{DC}(\text{diff}(r, y), s) \wedge \text{SEPNUM}(y, N + 1)]. \end{aligned}$$

As a special case, we can define the number of boundaries a region (other than Us) shares with its complement:

$$\text{CBNUM}(r, N) \equiv_{def} \text{SBNUM}(r, \text{compl}(r), N).$$

All the objects pictured in NM3, if embedded in E^3 , have CBNUM 2; we can define a doughnut as a region with FCON 2 and CBNUM 1.

Weakening assumptions 1-3 will show how the doughnut can be distinguished among successively larger classes of possibilities. The process is not completed within this paper. To start with, replace assumption 1 with assumption 1a:

1a. Us is E^N for some natural number N .

In terms of the near misses, this means we have to find ways of excluding the loop, cylinder-surface, annulus and torus of NM1. Because the axioms ensure that every region has an NTPP, all regions must have the same dimensionality as Us , given that the latter is a manifold; but how could we discover the dimensionality of Us if not told it? In (Gotts 1994), distinctions between regions of different dimensionality were made by specifying the properties of boundaries between EC parts of a region: this approach turned out to be a complicated one, partly but not wholly because RCC-theory does not allow direct reference to boundaries, or other entities of lower dimensionality than the regions in any given model of the theory. Here, an alternative approach is taken, but it is admitted that further investigation is still required: the approach has not been proved correct.

This approach is based on one of the conventional topological definitions of dimension (see for example (Armstrong 1979, section 9.5)), which uses the ideas of a *finite open cover* of a space, and of one such cover being a *refinement* of another. Roughly speaking, a finite open cover of a space is a finite set of subspaces whose union is the whole space, and one such cover is a refinement of another if each member of the first is part of a member of the second. A Hausdorff space (any Euclidean space is a Hausdorff space) is N -dimensional if any finite open cover has a refinement consisting of $N + 2$ subspaces which do not *all* share a common point, and this is not so for any smaller N .

The RCC-based adaptation of this definition proposed here uses a specialisation of the PO relation: *layered* partial overlap (LPO), in which the 'pure' (non-overlapping) parts of the two regions are DC from each other:

⁵We leave it undefined for pairs of regions that are not EC.

$$\text{LPO}(x, y) \equiv_{\text{def}} \text{PO}(x, y) \wedge \text{DC}(\text{diff}(x, y), \text{diff}(y, x)).$$

We define an 'LPO cover' of a region as a finite set of subregions that together include the whole of the region, such that any two members of the set are related either by DC, or by LPO. The upper and lower halves of figure 11 show two ways a square can be given an LPO cover. The left part of each half of the figure shows the disassembled pieces of an LPO cover, the right the fully assembled square (with overlaps of two, three or four components shown in successively darker shades), and the centre the way in which the lower left component relates to the upper right and upper left when the square is assembled.

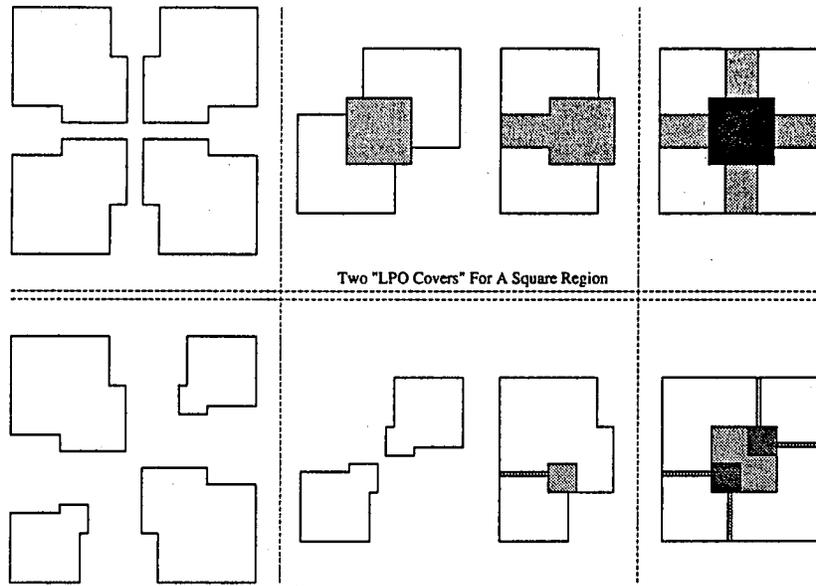


Figure 11: Two 'LPO Covers' For A Square Region

It is conjectured that, in general, for an N -dimensional region, any LPO cover with $N + 2$ members will have a 'one-to-one refinement' (one set of regions is a 'one-to-one refinement' of another if a one-to-one correspondence between the two sets can be established such that each of the first set is a P of the corresponding member of the second), such that there is no region that is a common part of all regions in the refinement; and that this will not be the case for any smaller N . In figure 11, the lower LPO cover is a refinement of the upper, and it can be seen that, unlike the upper, the four regions in the lower LPO cover have no common part. If this conjecture is correct, we could define a predicate DIM1 ('is 1-dimensional') as follows:

$$\begin{aligned} \text{DIM1}(r) \equiv_{\text{def}} & \\ & [\forall x, y, z [\text{EQ}(r, \text{sum}(\text{sum}(x, y), z)) \wedge \\ & [\text{DC}(x, y) \vee \text{LPO}(x, y)] \wedge [\text{DC}(x, z) \vee \text{LPO}(x, z)] \wedge [\text{DC}(y, z) \vee \text{LPO}(y, z)]] \rightarrow \\ & \exists x_1, y_1, z_1 [\text{EQ}(r, \text{sum}(\text{sum}(x_1, y_1), z_1)) \wedge \\ & \text{P}(x_1, x) \wedge \text{P}(y_1, y) \wedge \text{P}(z_1, z) \wedge \\ & [\text{DC}(x_1, y_1) \vee \text{LPO}(x_1, y_1)] \wedge [\text{DC}(x_1, z_1) \vee \text{LPO}(x_1, z_1)] \wedge [\text{DC}(y_1, z_1) \vee \text{LPO}(y_1, z_1)] \wedge \\ & [\neg \exists q [\text{P}(q, x_1) \wedge \text{P}(q, y_1) \wedge \text{P}(q, z_1)]]]. \end{aligned}$$

Definitions of DIM2, DIM3, and so forth could be constructed analogously⁶. Our doughnut-testing system could therefore begin investigating whether the mystery region is a doughnut by asking whether it is 3-dimensional. If not, it is not a doughnut; if so, the 'FCON 2, CBNUM 1' criterion can be used as before.

We can take the weakening of condition 1 a step further, substituting for 1a:

1b. U_s is an N -dimensional orientable manifold.

We insist that U_s be orientable because, so far, orientability has not been characterised in terms of C , nor is it known whether this is possible.

If 1b replaces 1a, r may not be identifiable via the same set of questions as when U_s was assumed to be some E^N . For example, assume that U_s itself is a doughnut, and that r is U_s . The criterion given above will not work correctly, because r has CBNUM 0. Nor is this the only case where such problems would arise. Suppose r and U_s are both doughnuts, topologically speaking, but the relationship of r to U_s is equivalent to that of one doughnut in the middle of a stack of doughnuts: r will have CBNUM 2, and so will fail the test. In considering the near misses of section 3, we can no longer rely on eliminating the three in NM3 that are of FCON 1 by looking at the CBNUM property: in each case, U_s might be precisely E^3 minus that region's internal void.

The problem is that the CBNUM is an *extrinsic* property of a region: that is, it is dependent on the way the region is embedded in U_s . We need an *intrinsic* definition of boundary-number — one expressed solely in terms of the way parts of a region relate to each other. No intrinsic C -based definition of boundary-number (IBNUM) known to apply *whatever* the dimensionality of r has yet been found. However, a C -based definition that works at least for regions with three or fewer dimensions can be presented; it depends on the prior definition of two other relations: FTTP(x, y) (read ' x is a firmly-tangential proper part of y '), and ITTP(x, y) (' x is an intrinsically tangential proper part of y ').

Intuitively, 'ITTP(x, y)' means that x is a proper part of y , and is not completely surrounded by $\text{diff}(y, x)$ — the part of y that is not also part of x . If U_s is E^3 , then a region x consisting of an infinitely long solid cylinder stretching (for example) the entire length of the x -axis is an NTPP of U_s — there are no regions EC to U_s , so there cannot exist a region that is EC to both U_s and x — but x is also an ITTP of U_s , because there are two directions (along the x -axis in either direction) in which x extends as far as U_s (i.e., infinitely far). A TPP is always an ITTP, and an INTTP (intrinsically non-tangential proper part — a PP that is not an ITTP) is always an NTPP, but the converses of these statements do not hold. If x is of finite diameter, and y is E^N , then $\text{TPP}(x, y) \equiv \text{ITTP}(x, y)$ will always be true, but as we have just seen, this equivalence will not always hold if x and y are both of infinite diameter; nor need it do so if U_s itself is of finite diameter.

The best definition of ITTP yet found relies on the prior definition of the FTTP

⁶A general DIM predicate, with two arguments, a region and a natural number, would be much more difficult to define.

relation. Intuitively, if x and y are both of finite diameter, then an N -dimensional x is an FTTP of y iff it is an PP of y and shares an $N-1$ -dimensional patch of boundary with y , but we need, and can produce, a definition that extends to regions of infinite diameter (at least, for regions of 3 or fewer dimensions). Figure 12 illustrates the relationship between the TPP/NTTP distinction of basic RCC-theory on the one hand, and the ITTP/INTTP distinction (top) and the FTTP relation (bottom) on the other.

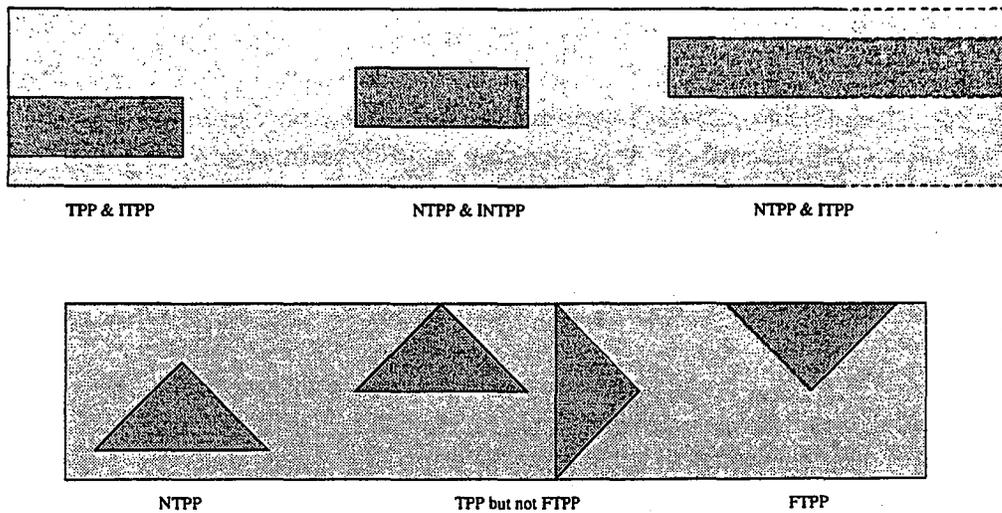


Figure 12: TPPs and ITTPs (top); TPPs and FTTPs (bottom)

The formal definition of FTTP is as follows:

$$\begin{aligned}
 \text{FTTP}(x, y) \equiv_{\text{def}} & \text{PP}(x, y) \wedge \\
 & \exists u, v, w [P(w, x) \wedge [\text{EQ}(\text{sum}(u, v), w)] \wedge \text{EC}(u, v) \wedge \\
 & \text{DC}(v, \text{diff}(y, w)) \wedge \text{FCON}(w, 1) \wedge \text{FCON}(v, 1) \wedge \text{FCON}(u, 1) \wedge \\
 & \neg \exists r, s, t [\text{EQ}(\text{sum}(s, t), u)] \wedge \text{PP}(r, s) \wedge \text{EC}(s, t) \wedge \\
 & \text{DC}(t, \text{diff}(s, r)) \wedge \text{FCON}(r, 2) \wedge \text{FCON}(s, 1) \wedge \text{FCON}(t, 1) \wedge \\
 & \neg \exists q [\text{PP}(q, r) \wedge \text{DC}(t, \text{diff}(s, q)) \wedge \text{FCON}(q, 1)]]].
 \end{aligned}$$

What this says is that x is an FTTP of y iff it is a PP of y , and has a part, w , which has FCON 1, and can itself be dissected into two smaller FCON 1 regions, one of which (v) is DC from $\text{diff}(y, w)$, while the other (u), which does touch $\text{diff}(y, w)$, meets a somewhat complicated condition, asserting that a particular type of 3-piece dissection of u (into t , r , and $\text{diff}(s, r)$) cannot exist. Any 1-dimensional or 2-dimensional region of FCON 1 meets this condition, but some 3-dimensional FCON 1 regions — those which would have one or more internal voids if embedded in E^3 — do not. The block-minus-block of NM2 is precisely such an FCON 1 region. If we consider 'subtracting' one block from another in such a way that the residue is also FCON 1, this residue may itself be a solid block (if we simply cut the original in half, for example), or it may be a block-minus-block. If the original block is itself a PP of some larger region, and we want to cut it into two FCON 1 parts so that one of them is DC from the rest of the larger region, the other *must* have an

internal void (be a block-minus-block) if the original block is *not* an FFTP of the larger region. It is thought that the definition of an FFTP could be extended to four or more dimensions in a similar fashion, but the details have not been worked out.

With the FFTP definition in hand, we can define an ITPP as follows:

$$\begin{aligned} \text{ITPP}(x, y) &\equiv_{\text{def}} \text{PP}(x, y) \wedge \\ &\forall w[\text{PP}(x, w) \wedge \text{PP}(w, y) \wedge \text{DC}(x, \text{diff}(y, w)) \rightarrow \text{FFTP}(w, y)] \end{aligned}$$

— that is, any PP of y that ‘envelopes’ x , shielding it from the rest of y , must be an FFTP of y .

Before we achieve our goal of defining intrinsic boundary number for regions of three or fewer dimensions, we will need one more intermediate definition, of the predicate MAX-P (Cohn et al. 1994). MAX-P(x, y) means that x is a CON P of y , while any PPI of x which is a P of y is not CON:

$$\text{MAX-P}(x, y) \equiv_{\text{def}} \text{CON}(x) \wedge \text{P}(x, y) \wedge \neg \exists z[\text{PP}(x, z) \wedge \text{P}(z, y) \wedge \text{CON}(z)].$$

The IBNUM of a region r of three or fewer dimensions can now be defined as the greatest SEPNUM of any s such that PP(s, r), each MAX-P of s is an ITPP of r , and any ITPP of r connects with s (each MAX-P will ‘cover’ one of the boundaries):

$$\begin{aligned} \text{IBNUM}(r, N) &\equiv_{\text{def}} [N = 0 \wedge \neg \exists s[\text{ITPP}(s, r)]] \vee \\ &[\exists s[\text{ITPP}(s, r)] \wedge \\ &\exists x[\text{SEPNUM}(x, N) \wedge \\ &\quad \forall z[\text{MAX-P}(z, x) \rightarrow \text{ITPP}(z, r)] \wedge \forall t[\text{ITPP}(t, r) \rightarrow \text{C}(t, x)]] \wedge \\ &\neg \exists y[\text{SEPNUM}(y, N + 1) \wedge \\ &\quad \forall z[\text{MAX-P}(z, y) \rightarrow \text{ITPP}(z, r)] \wedge \forall t[\text{ITPP}(t, r) \rightarrow \text{C}(t, y)]]] \end{aligned}$$

A doughnut can then be defined as a 3-dimensional region with FCON 2 and IBNUM 1.

We now turn to the possibility of weakening assumption 2: that the region-closure of any region forms a locally Euclidean space. This means that we have to find C-based criteria for eliminating those near misses in NM4 and NM5 that are not excluded by assumption 0 (that U_s is a manifold, and all regions correspond to regular open sets of points): that is, we must exclude NM4a, NM4b, NM5b and NM5c. These near misses do correspond to regular open sets of points, but the closures of these open sets are not locally Euclidean spaces: some of their boundary points do not have a neighbourhood topologically equivalent to half a 3-dimensional disc. (Another way of expressing this is to say that the *boundaries* of these regions are not locally disclike everywhere.) NM5b and c, in fact, have FCON 1, but putting another hole right through either of them would produce an FCON 2 region with IBNUM 1; NM4a and b are of FCON 2 and IBNUM 1 already. How can our doughnut-testing system eliminate these anomalous regions, and all others of similar type?

It is not difficult to define a predicate, ICON (for interior-connected), such that ICON(r) means that r does not divide into two or more parts which are only con-

nected because their closures share a point:

$$\text{ICON}(x) \equiv_{\text{def}} \forall y, z[\text{INTPP}(y, x) \wedge \text{INTPP}(z, x) \rightarrow \exists w[\text{CON}(w) \wedge \text{INTPP}(y, w) \wedge \text{INTPP}(z, w) \wedge \text{INTPP}(w, x)]]$$

However, $\text{ICON}(r)$ does not rule out all regions with anomalous boundaries, and in particular does not exclude any of the cases illustrated in figures 6 and 7.

We can define a more restrictive predicate, which we call WCON (for 'well-connected'), which appears to deal with all poorly-connected regions — that is, it is believed a 2- or 3-dimensional WCON region will always have a boundary with a locally linear or disclike topology respectively. Given this definition, we can remove assumption 2, and then have the choice of incorporating the conditions ensuring that a region is WCON as an additional axiom, or allowing both WCON and non- WCON regions while being able to distinguish the two. In order to define a WCON region formally, an intermediate definition will be necessary, but figure 13 should make it easy to remember the differences that are intended to exist between the three predicates CON , ICON , and WCON .

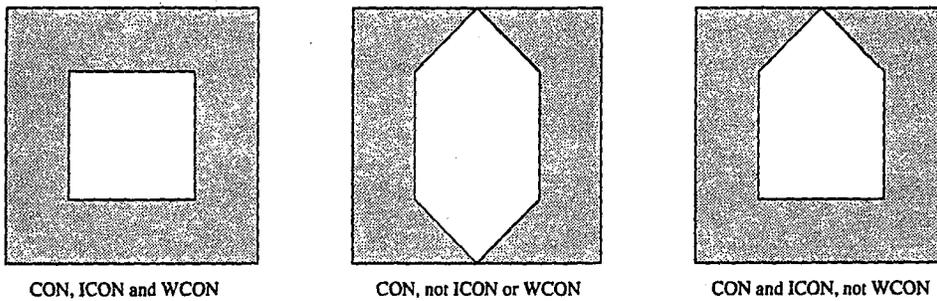


Figure 13: Types of CON Region

An FCON 1 , IBNUM 1 region s will be called a 'superficial proper part' of another region r — $\text{SPP}(s, r)$ — iff s is an ITPP of r , and for any INTPP of r , say t , there is another INTPP of r , u , such that $\text{diff}(s, \text{sum}(t, u))$ is also an FCON 1 region with IBNUM 1 . Intuitively, what this means is that if we remove any part of an SPP that does not touch the boundary of the containing region, we can always remove more of the SPP , again without touching the containing region's boundary, in such a way as to leave a remainder which is again FCON 1 and IBNUM 1 . Examples and counterexamples of SPPs of a 2-dimensional region are illustrated in figure 14. Formally:

$$\begin{aligned} \text{SPP}(s, r) \equiv_{\text{def}} & \text{FCON}(s, 1) \wedge \text{IBNUM}(s, 1) \wedge \\ & \text{ITPP}(s, r) \wedge \forall t[\text{INTPP}(t, r) \rightarrow \exists u[\text{INTPP}(u, r) \wedge \\ & \text{FCON}(\text{diff}(s, \text{sum}(t, u)), 1) \wedge \text{IBNUM}(\text{diff}(s, \text{sum}(t, u)), 1)]] \end{aligned}$$

It is thought that this same class of FCON 1 , IBNUM 1 ITPPs could also be picked out by specifying that s and r share only one patch of intrinsic boundary,

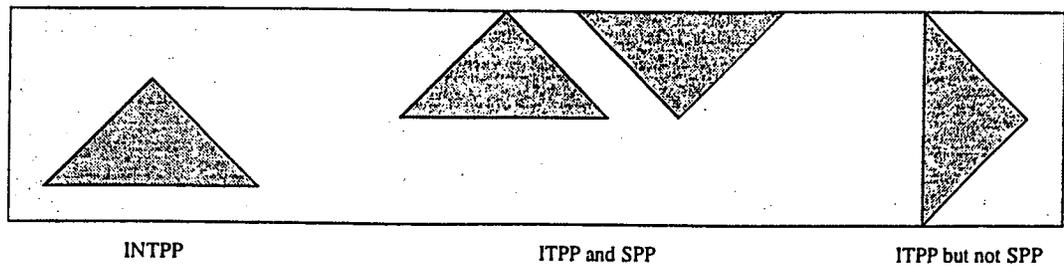


Figure 14: SPPs and ITPPs

and if s and r are N -dimensional, this patch is a point, or an M -dimensional disc where $M < N$. However, this condition on the form of the patch is redundant if $N \leq 2$. Also, this alternative means of specifying the class of SPPs is more difficult to express in terms of C .

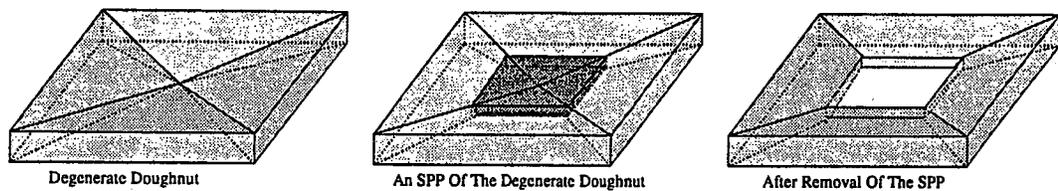


Figure 15: Removing An SPP From A Degenerate Doughnut

Removing an SPP from an ordinary doughnut cannot leave a remainder with a different FCON. For the examples of NM4 and NM5, however, there are SPPs, the removal of which would leave a remainder with a different FCON, as shown in figure 15 for NM5b. At one stage it was thought that this could be used as a way to exclude all poorly-connected regions. However, there are regions which do not have such an SPP, but which are poorly-connected. One such is a solid block with a smaller solid block removed from it in such a way that the surfaces of the two share a single point. Topologically, such a region could be created by attaching the region at the left of figure 15 to the top of a cuboid, so that the indentation in the bottom becomes an interior hole, meeting the outside at a single point. Such a region has FCON 1 and IBNUM 1, like a solid block, and removing an SPP cannot change this.

However, a WCON region *can* be defined as one with no SPPs of a certain sort. Specifically, a WCON region r has no SPP s which *itself* has an SPP t such that $DC(t, \text{diff}(r, s))$ and such that $\text{diff}(s, t)$ is not an FCON 1 region:

$$\text{WCON}(r) \equiv_{def} \forall s, t [\text{SPP}(s, r) \wedge \text{SPP}(t, s) \wedge DC(t, \text{diff}(r, s)) \rightarrow \text{FCON}(\text{diff}(t, s), 1)].$$

Now let us briefly consider what happens if we drop assumption 3: that the mystery region r is of finite diameter. Both the 'infinite doughnut' and the 'doughnut with infinite protruberance' of figure 9 are of FCON 2 and IBNUM 1, like the doughnut itself. Indeed, both these regions are topologically equivalent to a finite doughnut: they can be transformed into the doughnut by a one-to-one continuous mapping with a continuous inverse, just as (for example), a finite line segment can be transformed

into an infinite line: the distinction between spaces of finite and infinite diameter is *not* a topological one, and our C-based system is like conventional point-set topology in failing to distinguish them. Any dissection of the doughnut could be replicated for either of these infinite-diameter regions, simply by 'stretching' some or all of the regions in the dissection by an infinite factor. If we *know* that U_s is E^3 , we can distinguish the 'true' doughnut from either of these 'imposters', because they will be ITPPs of U_s and the true doughnut will not; and if we know that U_s is itself of finite diameter, the imposters cannot fit inside it; but if U_s is itself finite in some directions and infinite in others (e.g., an infinite cylinder), or if the doughnut-testing system is not given this piece of information (which cannot itself be expressed in terms of C), the system can only ask whether region r is an ITPP of U_s . If the answer is 'no', it can rule out the two imposters; but if it is 'yes', no further progress can be made. Since the two regions *are* topologically equivalent to a doughnut, however, we may decide to allow them to *be* doughnuts.

The same approach could be taken with regard to the 'doughknot', also shown in figure 9. Again, this region is topologically equivalent to the ordinary doughnut; it differs only in the way it is embedded in its containing 3-dimensional space. *Expressing* this extrinsic difference in terms of C would, it is thought, be by no means easy; and the *general* problem of determining whether two knots (embeddings of a closed curve in E^3) are equivalent remains an unsolved question in topology. Since the doughknot is not *intrinsically* different from the ordinary doughnut, we might simply accept it (and all more complex knottings of the doughnut) *as* doughnuts; but the last region shown in figure 9, the 'doughnut-with-knotted-hole', is *not* topologically equivalent to the ordinary doughnut. I am indebted to Achille Varzi for pointing out that I had not shown how to distinguish this from the ordinary doughnut using C; whether this can be done remains an open question, but at present the doughnut-testing system will be 'fooled' by such a region, which will pass all the tests for doughnuthood it can apply. What is more, the knot in the hole need not be the trefoil knot shown in figure 9: any one of the infinite number of distinct knots known to exist could take its place, and generate another topologically non-equivalent region.

6 Discussion: how far can C take us?

It has been shown that C can be the basis of rich topological classifications of regions and the spatial relations between them. The search for a definition of the doughnut has led to the exploration of a wide range of possibilities; and brought out the distinction between the intrinsic and extrinsic topological properties of a region, and the need to examine the range of possible models of RCC-theory.

Which of the informally-stated assumptions considered could reasonably be made under which circumstances deserves a brief comment. If the regions we wish to consider correspond to the space occupied by solid or liquid bodies, then we can assume that they are of finite diameter, and that they are triangulable. It might also seem that we could assume each region's closure to be 3-dimensional, but this is

to ignore the role of *idealization* in spatial reasoning: for many purposes it is useful to regard a piece of paper as 2-dimensional, or a rope as 1-dimensional. Similarly, if we wish to consider something with a cellular structure, such as a piece of plant or animal tissue, or a foam, we may need to consider regions which are best idealized as consisting of one or more 3-dimensional 'lobes' joined at points or along lines rather than at surfaces — contrary to the assumption that regions have closures which are locally-Euclidean spaces.

The exploration of taxonomies of spatial properties and relations based on C is far from complete. Separation-number, finger-connectivity and intrinsic-boundary-number do not distinguish all non-homeomorphic regions, even among manifolds whose closures are locally-Euclidean spaces. In general, this problem is unsolvable (it is known that there is no general method for determining whether two N -dimensional manifolds are homeomorphic for $n > 3$) (Stillwell 1980, p.5). Even for finite-diameter manifolds embeddable in E^3 , much work on C -based classification remains — as the example of the doughnut with knotted hole demonstrates. In the 2-dimensional case, a way of expressing *orientability* will be sought. In the 3-dimensional case, there are several levels of complexity to consider. First, there may be any number of separate maximal CON parts. A CON region can have any number of surfaces, and each such surface may be a sphere or N -hole torus. So the number and nature of the region's surfaces give a first layer of classification for CON regions, and we can classify non-CON regions by the 'bag'⁷ of such CON regions they comprise.

However, this is not the end of the matter, as we have already seen from the case of the knotted hole: if a surface of a 3-dimensional region is an N -hole torus, it may be embedded in 3-space in an infinite number of different ways. What is more, if there are two or more such surfaces, they may be *linked* (i.e., so arranged in space that a sphere cannot be placed so that one of them is on one side of it and the rest on the other) in an infinite number of ways. Consider the case, illustrated in figure 5, of a CON region with two surfaces, each a 1-hole torus (i.e. a doughnut minus a doughnut). The inner surface may be configured relative to the outer so that a sphere could be interposed between the two, or it may be wrapped one or more times round the hole 'through' the doughnut. This example shows that the FCON of a CON region cannot be calculated simply from those of its boundaries: if a sphere can be interposed between the two surfaces (as in NM3a) the FCON of the solid is 3; if not, it is 2 (at least in the simplest case, shown in NM3b, where the inner torus wraps around the hole through the doughnut once, producing a 'hollow doughnut'; more complex cases have still to be checked).

If a solid has two or more toroidal inner surfaces (whether these are simple — FCON 2 — tori, or are of higher connectivity), a new type of complexity arises, as two or more of these inner surfaces can be linked in a further infinite variety of ways, not possible when one of the surfaces is the outer surface of the region. Finally, two or more of the components of a non-CON region can also be linked — in the same set of possible ways, although in this case the topological distinctions between the

⁷A 'bag' in this sense is like a set, but elements can be repeated any number of times — just as a real bag may contain a number of objects, of which some may be indistinguishable from each other.

resulting regions are extrinsic, not intrinsic.

In the immediate future, investigation of the relationship between RCC's approach and point-set topology will continue, and formal proofs of the assertions made in this paper will be sought. Further work may be done on ways of classifying such multi-surface 3-dimensional regions, drawing on mathematicians' work on algebraic or combinatorial topology — e.g. (Stillwell 1980, Fuks and Rokhlin 1984) — and exploring further the notion of a dissection-graph: if two CON regions have the same finger-connectivity but are not topologically equivalent, can other dissection-graphs be used to distinguish them?

Beyond this, integrating topological and other qualitative aspects of spatial properties and relations is a major task. The work described in this paper needs to be integrated with that on convexity and inside/outside relations done by RCC, and with the work on conceptual neighbourhoods done by Freksa (Freksa 1992), RCC (Cohn et al. 1994) and others.

In conclusion, what are the advantages and disadvantages of the general approach to spatial representation and reasoning developed by RCC and continued here: that of working from a minimal set of primitives and axioms, exploiting their potential as far as possible before adopting any more? The advantages are several: such an approach has a mathematical and philosophical elegance absent from more complex systems of representation, discourages *ad hoc* additions to the system to meet unconsidered problems, and should ensure thorough familiarity with its properties and implications. It should be relatively simple to interface a system of spatial representation using a small number of primitives and axioms with another system — such as a vision module or geographical database. Similarly, it should be easier to investigate in depth the relationship of RCC's approach to point-set topology than would be the case for a more complex system. On the other hand, the work reported has made clear to the author that expressing what are intuitively quite simple concepts — such as the topological properties of a doughnut — in terms of a single primitive and a few axioms is neither easy nor free of pitfalls. In particular, inability to refer directly to dimensionality, to boundaries, and to the conceptual links between these concepts, has given rise to considerable difficulties. Only by trying to do as much as we can with as little as possible, however, are we likely to discover what representational primitives are likely to be most useful in spatial and qualitative reasoning.

Acknowledgements

The support of the EPSRC (formerly SERC) under grant no. GR/H 78955 is gratefully acknowledged. This work was also partially supported by the CEC under the Basic Research Action MEDLAR 2, Project 6471. Thanks are also due to Tony Cohn, John Gooday and Brandon Bennett for useful comments.

References

- Armstrong, M. A.: 1979, *Basic Topology*, McGraw-Hill, London.
- Casati, R. and Varzi, A.: 1994, *Holes and Other Superficialities*, MIT Press, Cambridge, MA.
- Clarke, B. L.: 1981, A calculus of individuals based on connection, *Notre Dame Journal of Formal Logic* **23**(3), 204-218.
- Clarke, B. L.: 1985, Individuals and points, *Notre Dame Journal of Formal Logic* **26**(1), 61-75.
- Cohn, A. G.: 1992, Completing sort hierarchies, *Computers and Mathematics with Applications* **23**(6-9), 477-491.
- Cohn, A. G., Gotts, N. M., Randell, D. A., Cui, Z., Bennett, B. and Gooday, J. M.: 1993, Exploiting temporal continuity in temporal calculi, Specialist Meeting on Time in Geographic Space, Lake Arrowhead, California.
- Cohn, A. G., Randell, D. A. and Cui, Z.: 1994, Taxonomies of logically defined qualitative spatial relations, in N. Guarino and R. Poli (eds), *Formal Ontology in Conceptual Analysis and Knowledge Representation*, Kluwer. To appear.
- Cohn, A. G., Randell, D. A., Cui, Z. and Bennett, B.: 1993, Qualitative spatial reasoning and representation, in N. P. Carreté and M. G. Singh (eds), *Qualitative Reasoning and Decision Technologies*, CIMNE, Barcelona, pp. 513-522.
- Freksa, C.: 1992, Temporal reasoning based on semi-intervals, *Artificial Intelligence* **54**, 199-227.
- Fuks, D. B. and Rokhlin, V. A.: 1984, *Beginner's Course in Topology: Geometric Chapters*, Springer-Verlag. Translated from the Russian by Andrei Iacob.
- Gotts, N. M.: 1994, How far can we 'C'? defining a 'doughnut' using connection alone, in J. Doyle, E. Sandewall and P. Torasso (eds), *Principles of Knowledge Representation and Reasoning: Proceedings of the 4th International Conference (KR94)*, Morgan Kaufmann, San Francisco, CA.
- Leonard, H. S. and Goodman, N.: 1940, The calculus of individuals and its uses, *Journal of Symbolic Logic* **5**, 45-55.
- Randell, D. A. and Cohn, A. G.: 1989, Modelling topological and metrical properties of physical processes, in R. Brachman, H. Levesque and R. Reiter (eds), *Proceedings 1st International Conference on the Principles of Knowledge Representation and Reasoning*, Morgan Kaufmann, Los Altos, pp. 55-66.
- Randell, D. A., Cui, Z. and Cohn, A. G.: 1992, A spatial logic based on regions and connection, *Proceedings 3rd International Conference on Knowledge Representation and Reasoning*, Morgan Kaufmann, San Mateo, pp. 165-176.

- Simons, P.: 1987, *Parts: A Study In Ontology*, Clarendon Press, Oxford.
- Smith, B.: 1994, Ontology and the logistic analysis of reality, in N. Guarino and R. Poli (eds), *Formal Ontology in Conceptual Analysis and Knowledge Representation*, Kluwer. To appear.
- Stillwell, J.: 1980, *Classical Topology and Combinatorial Group Theory*, Springer-Verlag, New York.
- Tarski, A.: 1956, Foundations of the geometry of solids, *Logic, Semantics, Metamathematics*, Oxford Clarendon Press, chapter 2. trans. J.H. Woodger.
- Varzi, A. C.: 1993, On the boundary between mereology and topology, *Proceedings 16th Wittgenstein Symposium*.
- Vieu, L.: 1993, A logical framework for reasoning about space, in A. U. Frank and I. Campari (eds), *Spatial Information Theory: a Theoretical Basis for GIS*, Vol. 716 of *Lecture notes in computer science*, Springer-Verlag, pp. 25-35. Proceedings of COSIT'93, Elba, Italy, September 1993.
- Weld, D. S. and De Kleer, J. (eds): 1990, *Readings in Qualitative Reasoning About Physical Systems*, Morgan Kaufman, San Mateo, Ca.
- Whitehead, A. N.: 1929, *Process and Reality*, The MacMillan Company, New York.
- Winston, P. H. (ed.): 1975, *The Psychology of Computer Vision*, McGraw-Hill.



A Theory of Spatial Regions with Indeterminate Boundaries

A G Cohn and N M Gotts
Division of AI, School of Computer Studies,
University of Leeds, Leeds, LS2 9JT, UK.
email: {agc,ngotts}@scs.leeds.ac.uk.

Abstract

The paper proposes considers the problem of representing and reasoning about spatial regions with undetermined boundaries. First we build a first order theory of such regions and then propose a possible translation this theory into an adaptation of 'RCC-theory', a region-based system for representing qualitative spatial relations developed over the last few years (Randell, Cui and Cohn 1992, Cohn, Randell and Cui 1994). The proposed translation is referred to as the 'egg-yolk' representation: a region with undetermined boundaries (a 'vague region') is represented by a pair of concentric regions with determinate boundaries ('crisp regions'), which provide limits (not necessarily the tightest limits possible) on the range of indeterminacy.

1 Introduction

The topic of this paper¹ is how best to deal with vagueness in spatial representation and reasoning, particularly within the framework of 'RCC-theory', (Randell, Cui and Cohn 1992, Cohn et al. 1994), which provides a representation of topological properties and relations in which regions rather than points are taken as primitive. We are concerned here with regions having vague or indeterminate boundaries but with a known location, not with vaguely located entities.

Many of the spatial regions we consider in everyday contexts do not have precise boundaries: consider urban areas, clouds of gas, galaxies, and habitats of particular plants. Such 'regions' will be called 'vague' or 'NonCrisp' here, in contrast to 'Crisp' regions: those with precisely-defined boundaries. RCC-theory as originally devised has no means of representing and reasoning about NonCrisp regions.

What properties do we require of a treatment of spatial vagueness? First, it should be logically consistent. Second, it should so far as possible respect our intuitions: the more we are able to express the kinds of things we want to say about vague regions, the better. Third, it should be computationally tractable, both in

¹This paper is a revised version of (Cohn and Gotts 1994).

itself, and when combined with spatial information expressed in precise terms. Finally, if it can be linked to an existing treatment of precise spatial information in a straightforward fashion, so much the better.

In the first part of this paper we briefly review a first order theory of regions with crisp, well defined, boundaries. Then we present an axiomatisation of regions with undetermined boundaries. Finally, we attempt to build a system for representing and reasoning about vague regions on the basis of an existing representation for Crisp ones and show how it can be related to the above axiomatisation. This approach is referred to as the 'egg-yolk' representation, for reasons which will become clear. We do not claim to have produced a complete solution to the problems of representing regions with undetermined boundaries, but believe we have made considerable progress toward one.

2 The RCC-theory of Spatial Regions

The focus of research on spatial representation and reasoning at Leeds has been to evaluate, extend and implement a theory ² of space and time based upon Clarke's (Clarke 1981, Clarke 1985) 'calculus of individuals based on connection', and expressed in the many-sorted logic LLAMA (Cohn 1987) ³. Our revised and extended theory, now known as 'RCC-theory' has been developed in a series of papers, including (Randell and Cohn 1989), (Randell 1991), (Randell and Cohn 1992), (Randell, Cohn and Cui 1992), (Cohn et al. 1994), (Bennett 1994), and (Gotts 1994). The most distinctive feature of Clarke's 'calculus of individuals', and of our work, is that extended *regions* rather than points are taken as fundamental, and it is partly this feature which makes RCC-theory a promising basis for a treatment of vague regions. Our formal theory supports regions having either a spatial or a temporal interpretation (temporal 'regions' are periods of time). Informally, these regions may be thought of as infinite in number, and 'connection' may be any relation from external contact (touching without overlapping) to spatial or temporal identity. Spatial regions may have one, two, three, or even more than three dimensions, but in any particular model of the formal theory, all regions are of the same dimensionality. Thus, if we are concerned with a two-dimensional model, such as one in which regions are areas of land, the boundary lines and points at which these regions meet are not themselves considered regions. Indeed, they cannot be referred to *directly* within RCC-theory, in which they are all assigned to sort NULL, which is disjoint from sort REGION. However, as (Gotts 1994) shows, a great deal can be specified about the properties and relations of such lower-dimensional boundary entities while referring explicitly only to relations between the higher-dimensional regions they bound.

²We use the word 'theory' in its logical/mathematical sense, meaning a set of formal axioms which specify the properties and relations of a collection of entities, not in the natural scientist's sense of an empirically testable explanation of observed regularities.

³In a many-sorted logic, the universe of discourse is divided into subsets called 'sorts', and quantifiers may be restricted to these sorts. Function and predicate symbols are defined only on particular combinations of argument sorts.

The basic part of the formal theory assumes a primitive dyadic relation: $C(x, y)$, read as 'x connects with y' (where x and y are regions). Two axioms are used to specify that C is reflexive and symmetric. C can be given a topological interpretation in terms of points incident in regions. In this interpretation, $C(x, y)$ holds when the topological *closures* of regions x and y share at least one point.

Using the relation C , further dyadic relations are defined. In particular a set of eight 'base relations' can be defined, of which one and only one will hold between a given pair of regions (these relations thus form a jointly exhaustive and pairwise disjoint or JEPD set). These are illustrated in the upper part of figure 1. The eight are: DC (DisConnected), EC (Externally Connected), PO (Partially Overlapping), TPP (Tangential Proper Part), NTPP (Non-Tangential Proper Part), EQ or = (Equal), TPPI (Tangential Proper Part Inverse), NTPPI (Non-Tangential Proper Part Inverse).

This set of eight base relations, referred to as the 'RCC-8' relations, is used in much of our work on qualitative spatial relations. In this paper, a smaller JEPD set of relations will be used: RCC-5, whose relation to RCC-8 is illustrated in figure 1⁴. RCC-5 consists of the relations {PO, PP, EQ, PPI, DR} (Partially Overlapping, Proper Part, Equal, Proper Part Inverse, and Distinct Regions). Considered relative to RCC-8, RCC-5 lumps together TPP and NTPP as PP, TPPI and NTPPI as PPI, and EC and DC as DR. Within RCC-5, C need no longer be available as a primitive: we can use PP (among other possibilities) as a primitive from which to define all the other RCC-5 relations.

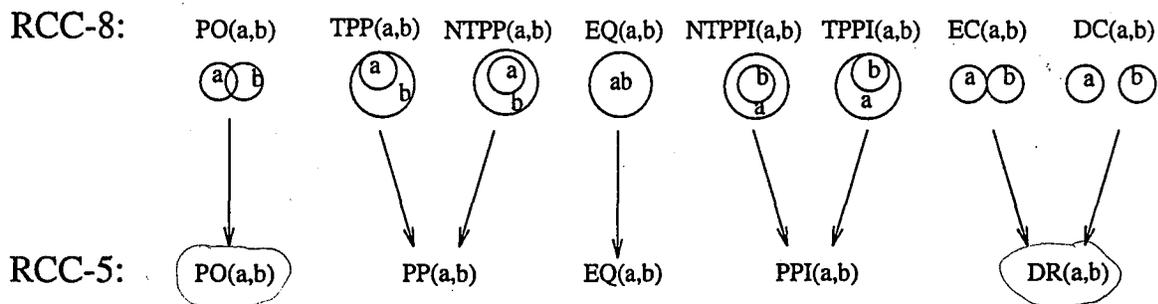


Figure 1: The RCC-8 and RCC-5 sets of JEPD relations between region-pairs

3 Describing and Reasoning about Vague Regions

We attempt in this paper to extend our connection-based approach to qualitative spatial representation and reasoning to regions with indeterminate boundaries: vague or 'NonCrisp' regions. What kinds of things do we want to be able to say about vague regions? How do we want vague regions to relate to 'Crisp' regions, and to each other? What formalism will most efficiently and intelligibly express these relations?

⁴Although it is planned to extend the work, using the RCC-8 relations as a basis, in the near future.

We want to be able to say at least some of the same sorts of things about vague regions as about Crisp ones: that one contains another (southern England contains London, even if both are thought of as vague regions), that two overlap (the Sahara desert and West Africa), or that two are disjoint (the Sahara and Gobi deserts). In these cases, the two vague regions represent the space occupied by distinct entities, and we are interested in defining a vague area corresponding to the space occupied by either of the two, by both at once, or by one but not the other.

We may also want to say that one vague region is a 'crisper' version of another. For example, we might have an initial (vague) idea of the area inundated in a flood, and then receive information from a systematic survey, reducing the imprecision in our knowledge. In this case, the vagueness of the vague regions is a matter of our ignorance; in others, it appears intrinsic: consider an informal geographical term like 'southern England'. Here, the uncertainty about whether particular places (somewhere to the north of London but to the south of Birmingham) are included cannot be resolved by evidence, unless we simply decide to take the majority view across a particular sample of opinion as decisive. There is a degree of arbitrariness about any particular choice of an exact boundary, and for most purposes, none is required. But *if* we decide to define a more precise version of such an informal term for some purpose, the choice of a precise definition is by no means *wholly* arbitrary: generally, we can distinguish between more and less 'reasonable' choices of more precise description. We need to keep the distinction between ignorance-based and intrinsic vagueness in mind, but many of the problems of representation and reasoning are the same in both situations, and the two are often found together: consider the area inhabited by a particular bird species or subspecies. Here, our knowledge may indeed be limited, but in addition, just where the limits lie may depend on whether we count areas sometimes strayed into on the fringes of the bird's range, or where we decide to say that one species or subspecies ends and another begins.

We can begin by considering what we may want to say, and what we want to hold true, when we consider relations between alternative vague-region versions of an entity's spatial extent. We will need a name for the relation between two estimates of which one is a refinement or 'crisping' of the other, which we could express $X \prec Y$, read: 'region X is a crisping of vague region Y ' (we will use upper-case letters for variables ranging over both Crisp and NonCrisp regions, in contrast to the lower-case letters used for variables ranging only over the normal, Crisp regions of RCC-theory).

What sort of properties do we want \prec to have? First, \prec should be asymmetric (if one NonCrisp region is a crisper version of a second, then the second cannot simultaneously be a crisper version of the first), irreflexive (a NonCrisp region is not a crisping of itself) and transitive (if NonCrisp region A is a crisping of NonCrisp region B , and B of C , then A is a crisping of C):

$$(A1) \forall X, Y [X \prec Y \rightarrow \neg Y \prec X]$$

$$(A2) \forall X, Y, Z [(X \prec Y \wedge Y \prec Z) \rightarrow X \prec Z]$$

A1 ensures that \prec is asymmetric and hence irreflexive, A2 that it is transitive.

It will be convenient to define also a relation \preceq , read as 'either X is a crisping of Y , or X and Y are equal'.

$$(D1) X \preceq Y \equiv_{def} X \prec Y \vee X = Y$$

This raises the question of how best to deal with identity of non crisp regions: should it be presupposed, or it could be defined in terms of \prec thus:

$$X = Y \equiv_{def} \forall Z [X \prec Z \equiv Y \prec Z].$$

(X and Y are equal iff X is a crisping of any Z iff Y is a crisping of that Z.) Note that this assumes that there is at most one distinct NonCrisp regions with no decrispings.⁵

Further relations can also be defined in terms of \prec , including $X \succ Y$ ('X is a vaguer version of Y') and MA(X, Y) ('X and Y are mutual approximations'):

$$(D3) X \succ Y \equiv_{def} Y \prec X$$

$$(D4) MA(X, Y) \equiv_{def} \exists Z [Z \preceq X \wedge Z \preceq Y].$$

The first of these simply defines \succ as the inverse of \prec . The definition of MA means that two (noncrisp) regions are 'mutually approximate' if they are equal, or share a common crisping, which is intuitively reasonable. We can also define a predicate Crisp:

$$(D5) Crisp(X) \equiv_{def} \neg \exists Y [Y \prec X]$$

and, correspondingly, NonCrisp:

$$(D6) NonCrisp(X) \equiv_{def} \exists Y [Y \prec X].$$

This definition of a Crisp NonCrisp region simply specifies that it is one with no crispings. Using \prec and Crisp, we can define the relation $\prec\prec$ ('X is a complete crisping of Y'):

$$(D7) X \prec\prec Y \equiv_{def} X \prec Y \wedge Crisp(X).$$

If one NonCrisp region is a crisping of another, we would surely like there to be an alternative crisping, incompatible with the first: what else could it mean to substitute a crisper version of a vague region, if no alternative possibilities were thereby excluded? One way of ensuring this is to add the axiom:

$$(A3) \forall X, Y [X \prec Y \rightarrow \exists Z [Z \prec Y \wedge \neg MA(X, Z)].$$

There is an interesting and largely unexplored set of parallels between the relations of crisping and decrisping as described here on the one hand, and part/whole relations (such as we find in RCC-5) on the other. We can draw a correspondence between \prec and the relation PP (Proper Part), and hence between the pairs \preceq and P and \succ and PPI. MA corresponds to O ('Overlap', meaning that any relation other than DR holds between two REGIONS: they have some REGION as common part). Axioms A1 and A2 correspond to fundamental axioms establishing the properties of the proper part relation, while A3 corresponds to the 'Weak Supplementation Principle' (Simons 1987)⁶, which in the terms used here can be expressed:

$$(WSP) \forall x, y [PP(x, y) \rightarrow \exists z [PP(z, y) \wedge DR(x, z)].$$

Can we carry this parallel further? We could add, to the axioms we have so far, a requirement that the existence of at least one common crisping of a pair of

⁵If we had defined equality as: $X = Y \equiv_{def} \forall Z [Z \prec X \equiv Z \prec Y]$ instead, then note that this would require that there would be at most one NonCrisp region with no crispings, which seems intuitively unreasonable.

⁶A version of this theorem also applies to RCC-8 and hence RCC-5, see (Randell, Cui and Cohn 1992).

NonCrisp regions implies the existence of a 'vaguest common crising' (VCC), such that if Z is the VCC of X and Y , any other common crising of X and Y will also be a crising of Z .

$$(A4)\forall X, Y[\exists Z[Z \prec X \wedge Z \prec Y] \rightarrow \exists Z\forall W[W \preceq Z \equiv W \preceq X \wedge W \preceq Y]].$$

This corresponds to a mereological axiom guaranteeing the existence of a 'maximal common part' (or 'product') for any two overlapping entities⁷. Axioms corresponding to the four specified so far define what Simons calls a 'Minimal Extensional Mereology' (a mereology being a system of part/whole or part-of relations).

The existence of the analogue to the idea of a binary *product* described above suggests we might look for a parallel to the concept of a binary sum as well; however we will not pursue this further here.

We might also find it useful to be able to describe relations between a vague region and a completely crisp one (i.e., a normal region of the type dealt with in RCC-theory); in particular, the relation between a crisp and a vague region where the former can be regarded as a '*complete crising*' (\prec) of the latter. In fact, we can use the possible relations between complete crispings of two vague regions representing *different* spatial entities in classifying the possible relations between such pairs of vague regions, as we will now demonstrate.

Obviously, depending on the initial configuration of the two vague regions, various possibilities might arise. For example if the two NonCrisp regions are sufficiently far apart — for example if we consider both Leeds (England) and Buffalo (New York State) to be NonCrisp — then it seems reasonable to insist that for such a configuration the relationship between any complete crising of both regions would be DR (see figure 2(a)). Similarly, it would seem reasonable that however we made Buffalo and the USA crisp, then Buffalo would always be a proper part of the USA (figure 2(b)). In other cases things might not be so clear cut: consider two species with core habitats close to but distinct from each other, but which may encounter each other in an intermediate zone marginal for both (2(c)). Here, the relationship between a pair of complete crispings of the two regions might be either DR or PO. In general, there will be some set of (one or more) RCC-5 relations each of which may hold between complete crispings of a pair of vague regions. This set of possible relations can itself be regarded as a relation holding between the two NonCrisp regions.

Which sets of RCC-5 base relations can 'hold' in this way? It seems reasonable to assume that each such set will be a *conceptual neighbourhood* (Freksa 1992, Cohn et al. 1994) of the RCC-5 set of relations. Given a set of JEPD binary relations, such as RCC-5 or RCC-8, we can call any pair of them *immediate conceptual neighbours* if each can be transformed into the other by a process of gradual, continuous change which does not involve passage through any third relation. In the case of RCC-5 (figure 3), all the other four relations are immediate conceptual neighbours of PO; and PP and PPI are immediate conceptual neighbours of EQ; but there are no other pairs of immediate conceptual neighbours.

Any subset of such a set of JEPD relations is a *conceptual neighbourhood* if any pair belonging to the subset are themselves immediate conceptual neighbours, or

⁷A different kind of 'product' between pairs of NonCrisp regions will be defined in the next section

p. 133
fig 1

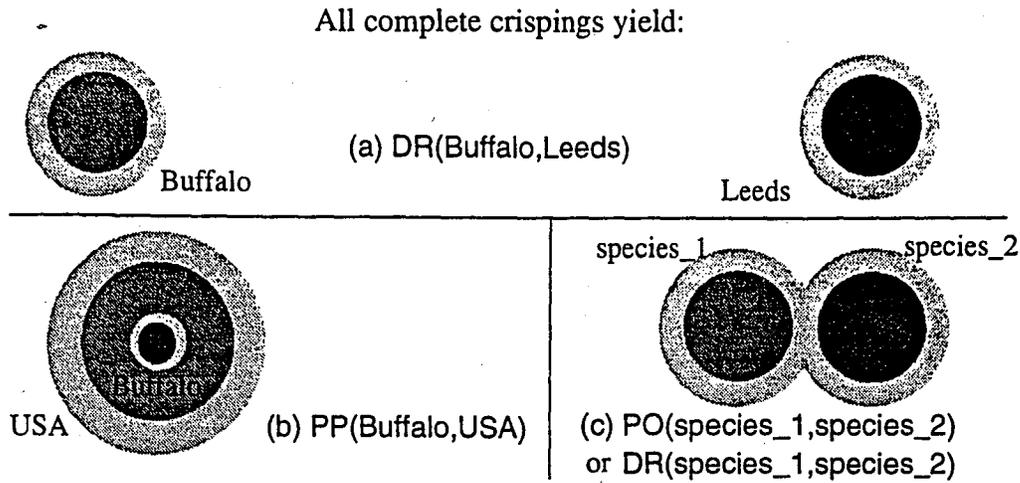


Figure 2: What relationships are there between complete crispings of these regions?

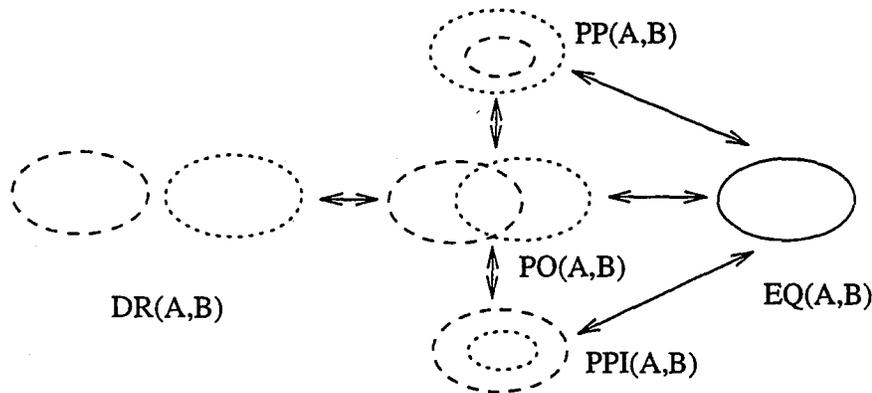


Figure 3: Immediate conceptual neighbours among the RCC-5 relations

can form the ends of a 'chain' of relations in which each adjacent pair are immediate conceptual neighbours. There are exactly 23 possible conceptual neighbourhoods in the RCC-5 set of relations: 5 of rank 1 6 of rank 2, 7 of rank 3, 4 of rank 4 and 1 of rank 5:

- {DR}, {PO}, {PP}, {PPI}, {EQ},
- {DR,PO}, {EQ,PP}, {EQ,PPI}, {PO,PP}, {PO,PPI}, {EQ,PO},
- {DR,PO,PP}, {DR,PO,PPI}, {PO,PP,EQ}, {PO,PPI,EQ}, {PP,PPI,EQ},
- {DR,PO,EQ}, {PO,PPI,PP},
- {DR,PO,PP,PPI}, {EQ,PO,PP,PPI}, {DR,PO,PPI,EQ}, {DR,PO,PP,EQ},
- {DR,EQ,PO,PP,PPI}.

Can all of these arise as sets of possible relations between the complete crispings of pairs of NonCrisp regions? We will argue the contrary. First consider {EQ}: if this were the only possible relationship between complete crispings of two regions, the regions must surely have been crisp (and EQ) in the first place. We believe all the other rank 1 conceptual neighbourhoods are possible: we have already effectively

argued for DR above when discussing Leeds and Buffalo (figure 2a); for PP, and PPI just imagine one very small NonCrisp region right in the centre of a very much larger one (figure 2b) while for PO consider two large regions each considerably overlapping the other but still with a significant subregion not part of the other (like the Sahara and West Africa).

Now consider rank 2 conceptual neighbourhoods; the only one we would want to argue is not possible is {EQ,PO}: given that the two regions are similar enough to crisp to EQ and dissimilar enough to crisp to PO then why should they not crisp to PP or PPI? Consider a pair of NonCrisp regions X and Y , which have at least one pair of complete crispings of that are EQ, and another pair that are PO. Consider a pair of these complete crispings of X and Y that are EQ. To reach a pair of crispings that are PO, either both of the members of the EQ pair must be enlarged (with their areas of enlargement not completely coinciding), or both must be shrunk (with their areas of shrinkage not completely coinciding). In the first case, enlarging only the complete crispering of X will give a PPI pair, enlarging only the complete crispering of Y will give a PP pair. If the PO pair is reached by shrinking the members of the EQ pair, shrinking only the complete crispering of X will give a PP pair, and shrinking only the complete crispering of Y a PPI pair. So if complete crispings can be chosen independently for any two NonCrisp regions, {EQ,PO} is not among those sets of possible relations that can occur between regions⁸. All the other rank 2 conceptual neighbourhoods seem reasonable: {DR,PO} is possible if two NonCrisp regions slightly overlap each other as in figure 2c — the crispering might retain the overlap or result in discrete regions; for {EQ,PP}, consider an island, with vague boundaries owing to tides or erosion, and some part of the island where a particular plant could live — the latter region might sometimes be equal and sometimes a proper part of the former but could never be in a PPI or PO relation to it. The duality of PP and PPI means that {EQ,PPI} is also reasonable. We want to allow {PO,PP} in order to cope with a small region which is somewhere around the edge of a very much larger region — the former region might be a part or might partially overlap once the two regions are completely crisped; again duality forces us to allow {PO,PPI} as well.

Similar arguments lead us to accept just two rank 3 conceptual neighbourhoods: {DR,PO,PP} and {DR,PO,PPI}; one rank 4 neighbourhood: {EQ,PO,PP,PPI} and the sole rank 5 neighbourhood {DR,EQ,PO,PP,PPI} (it seems clear that this should be allowed since if we had two very vague regions superimposed then surely any relationship might result depending on the crispering of each region).

This gives 13 out of the 23 possible conceptual neighbourhoods of RCC-5 relations as possible sets of relations between the complete crispings of a pair of NonCrisp regions.

{DR}, {PO}, {PP}, {PPI},

⁸There may be applications where this independence assumption would be unjustified: for example, if we wanted to permit vague regions to have a precise, fixed size, with only their position being vague. Here, {EQ,PO} might be the full set of alternatives for complete crispings of some pairs of vague regions. However at the beginning of this paper we decided to not to consider locational vagueness.

{DR,PO}, {EQ,PP}, {EQ,PPI}, {PO,PP}, {PO,PPI},
 {DR,PO,PP}, {DR,PO,PPI},
 {EQ,PO,PP,PPI},
 {DR,EQ,PO,PP,PPI}}.

This constitutes a set of 13 JEPD possible relations between vague regions (compared to five in RCC-5). Can we refine this classification any further? We believe so. The crucial observation is that if one of a pair of vague regions is made completely crisp, then, in general, not all of the relationships that were originally possible will still remain so. However rather than explore this in detail now, we will return to this analysis after presenting the 'egg-yolk' interpretation of vague regions, and perform the analysis within that interpretation⁹.

4 The Egg-Yolk Theory

Lehmann and Cohn (Lehmann and Cohn 1994) suggest using two (or possibly more) concentric subregions, indicating degrees of 'membership' in a vague region. (In the simplest, two-subregion case, the inner subregion is referred to as the 'yolk', the outer as the 'white', and the inner and outer subregions together as the 'egg'.) This egg-yolk approach is proposed in the context of the problem of integrating heterogeneous databases, where the notions of 'regions' and 'spatial relations' are used metaphorically to represent sets of domain entities, and relations between these sets. Is it a satisfactory approach to spatial vagueness in a literal sense? How does it relate to the crisp/vague distinction outlined above?

The egg-yolk formalism developed in (Lehmann and Cohn 1994) allows for just 5 base relations (corresponding to the RCC-5 set discussed in section 2: DR, PO, PP, PPI and EQ) between any egg-egg or yolk-yolk pair, or any egg and the yolk belonging to another egg (a yolk is always a PP of its own egg). This choice of a set of base relations is not an intrinsic feature of the approach, but corresponds to the demands of the database interpretation, where the concept of two sets of entities being 'externally connected' (EC), and the distinction between tangential and nontangential proper parts are not useful. The RCC-5 set produces 46 possible relations between a pair of egg-yolk pairs¹⁰, as shown in figure 4. Can these 46 possibilities be identified with the possible relations between complete crispings of two NonCrisp regions? We believe they can.

At first glance, there is one apparent problem with the egg-yolk approach: the most obvious interpretation is that it simply replaces the precise dichotomous division of space into 'in the region' and 'outside the region' of the basic RCC theory by an equally precise trichotomous division into 'yolk', 'white' and 'outside' — and this appears contrary to our intuitions about how vagueness 'works'. What we appear to need if we want to reflect the way people represent and reason with vague regions

⁹We hope also to perform the complete analysis independently of the egg-yolk interpretation, but we have not done so yet.

¹⁰There are 42 in (Lehmann and Cohn 1994), but this is because the database application considered there makes it unnecessary to distinguish different yolk-yolk relations if the eggs are EQ.

penumbra - non-CRISP

is something different: not only is there a 'doubtful' or 'borderline' zone around the edges of a vague region, but that zone itself does not have precise boundaries. Is there a way of using the egg-yolk formalism that is consistent with this? What is the relationship between the intuitive idea of a vague region, and an egg-yolk pair of nested crisp regions?

We are currently working on a self-contained axiomatisation of reasoning about relations between vague regions, which can then be modelled or interpreted within the egg-yolk formalism. Here, we can only give an informal characterisation of what this approach involves. The 'egg' and 'yolk' of an egg-yolk pair are taken to represent conservatively-defined limits on the possible 'complete crispings' or precise versions of a vague region: any acceptable complete crising must lie between the inner and outer limits defined by yolk and egg, but we do *not* assert that any crisp region meeting this constraint should be considered an acceptable complete crising of the vague region. We leave undefined what additional conditions, if any, must be met by such a complete crising. Thus, in this use of the egg-yolk formalism we do not claim that the borders of egg and yolk represent precise limits of a NonCrisp region's penumbra of vagueness, but rather that the entire penumbra (and perhaps more) lies between these limits. This gives us the kind of indefiniteness in the extent of vagueness, or 'higher-order vagueness', that intuition demands. Consider the vague region 'beside my desk'. There are some precisely defined regions, such as a cube 10cm on a side, 5cm from the right-hand end of my desk, and 50cm from the floor, that are undoubtedly contained within any reasonable complete crising of this NonCrisp region. There are others, such as a cube 50m on a side centred at the front, top right-hand corner of the desk, that contain any such reasonable complete crising. These two could correspond to the 'yolk' and 'egg' of an egg-yolk pair corresponding to the vague region 'beside my desk'; but there are precisely defined regions including the smaller and lying within the larger of the pair that would *not* be reasonable complete crispings of 'beside my desk'. We do not and need not specify *exactly* where the limits of acceptability lie.

The set of possible egg/yolk configurations is shown in figure 4¹¹. Configuration 1, given our interpretation of vague regions in terms of egg-yolk pairs of RCC regions, shows a pair of NonCrisp regions such that *any* pair of complete crispings of the two are DR (as in figure 2a). Configuration 2 (compare figure 2c) can be interpreted as a pair of NonCrisp regions X (dashed) and Y (dotted), such that it may be possible to choose complete crispings of the two to be DR or PO, and for any complete crising of X or Y, a DR complete crising of the other may be available, but there are some complete crispings of each for which no PO complete crising of the other can be found. Similarly, although configuration 9 has the same set of possible RCC-5 relations between any complete crispings of the vague regions involved, only PO is always reachable. These two configurations are illustrated in

¹¹Note that although the eggs and yolks are shown as one piece regions, there is of course nothing to stop them being multipiece, which might be useful in modelling some situations — e.g. one might want to consider that the Tower of London, Buckingham Palace and Westminster Palace and Abbey were always part of London whatever else was, and could thus form a trio of disconnected yolks for the London egg.

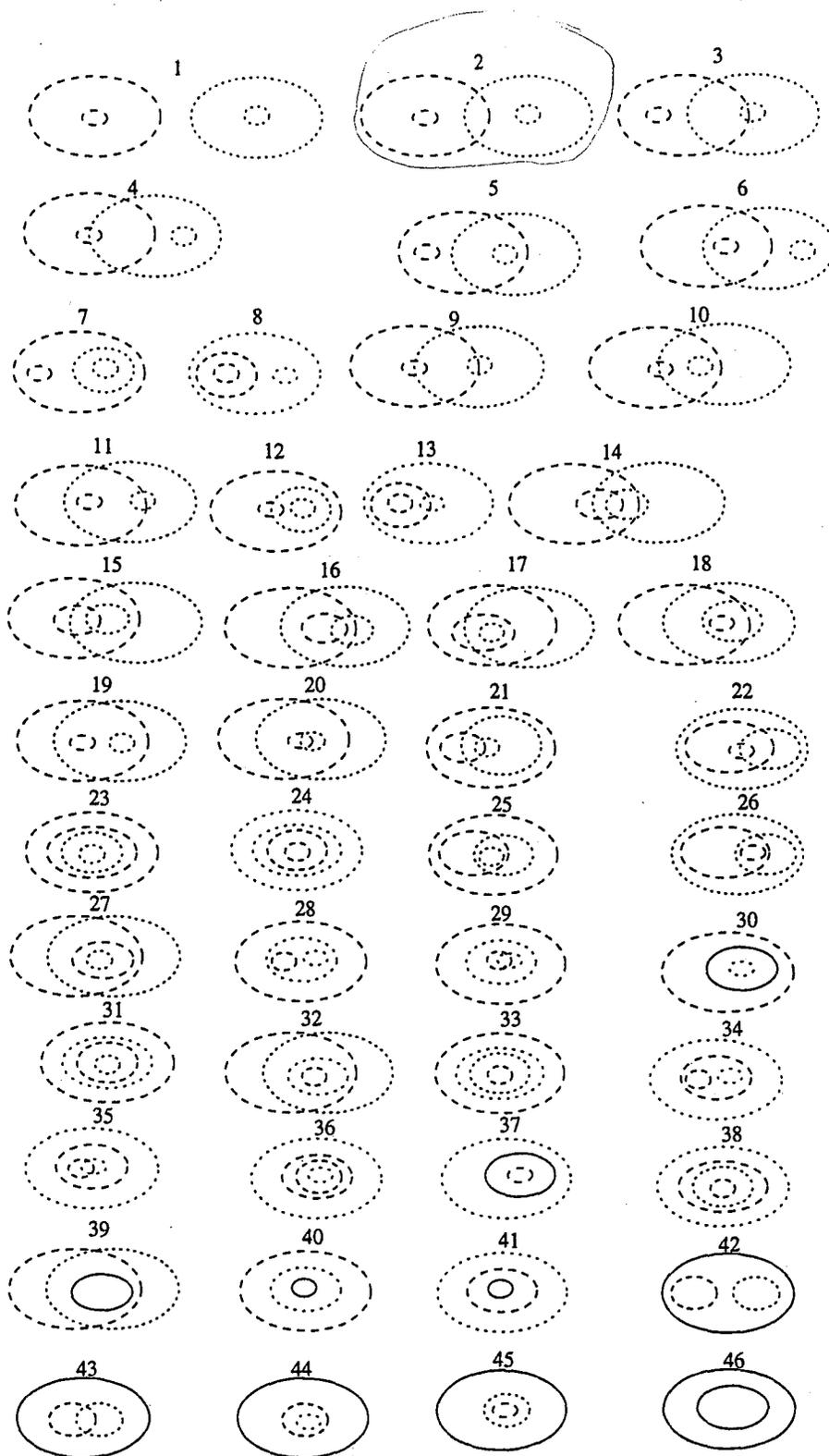


Figure 4: The 46 possible relationships between two egg-yolk pairs

figure 5. A complete analysis of the possible sets of alternative complete crispings

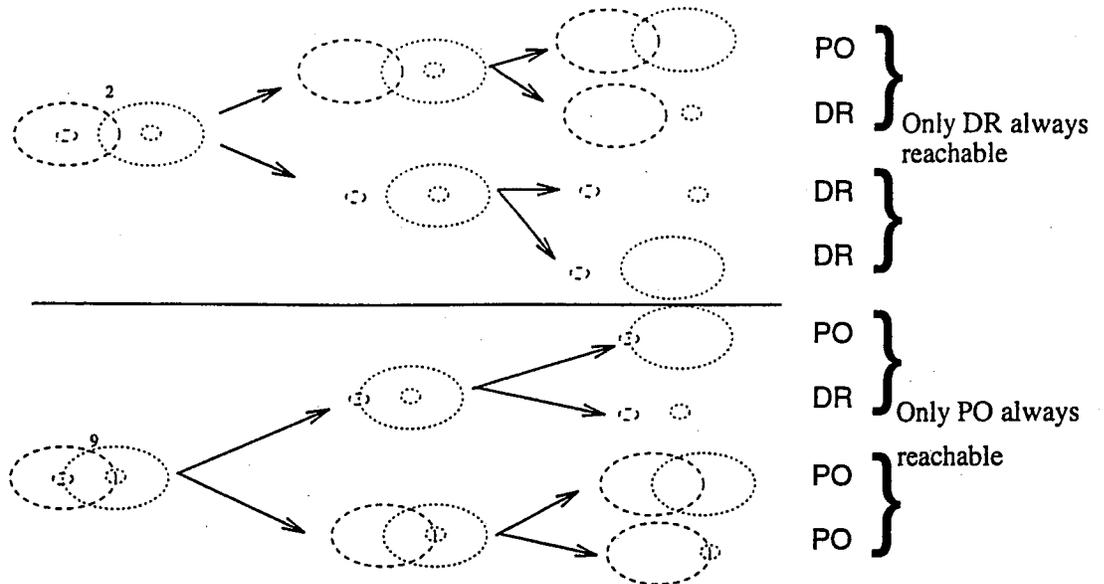


Figure 5: The reachable RCC-5 relations from configurations 2 and 9.

of egg-yolk pairs yields the result that the set of possibilities for a given pair of egg-yolk pairs always coincides with one of the 13 conceptual neighbourhoods referred to in the previous section. Depending on how it is done, as noted above, completely crisping one region may reduce the range of possibilities that can *then* be created by completely crisping the other.

An analysis taking account of the sets of possibilities that remain *however* either the dotted or the dashed egg-yolk pair is crisped, as well as on the full set of possibilities available before either is crisped, distinguishes all 46 region pairs except for the pair 39-43. We are still investigating how this pair can best be distinguished. One approach is to list, not just those RCC-5 relations that can always be reached however X is completely crisped, or however Y is completely crisped, but all *minimal sets* of RCC-5 relations of which at least *one* can always be reached under the same circumstances. In the case of configurations 39 and 43, no single RCC-5 relation between complete crispings always remains possible, once either the dotted or the dashed vague region is completely crisped. In configuration 39, however, any complete crisping of the dashed region will leave at least one from each of the following sets of relations reachable: $\{EQ, PO\}$, $\{EQ, PPI\}$, $\{PP, PO\}$, $\{PP, PPI\}$. In the case of configuration 43, the corresponding sets of relations are: $\{EQ, PO\}$, $\{EQ, PP\}$, $\{PPI, PO\}$, $\{PP, PPI\}$ (the second and third sets differ, with PP replacing PPI and vice versa). If we consider crisping the dotted region first, we find that the list of minimal sets of relations for configuration 39 is the same as that found for configuration 43 when crisping the dashed region first; and conversely, the list of minimal sets of relations for configuration 43 is the same as that found for configuration 39

when crisping the dashed region first¹². The situation is illustrated in figure 6.

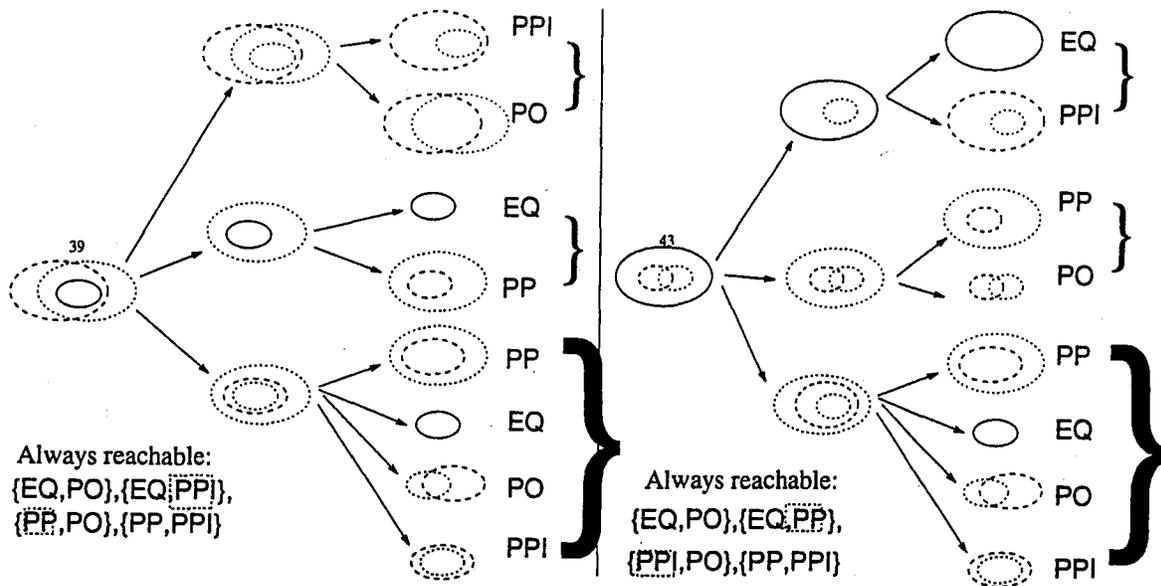


Figure 6: Distinguishing between configurations 39 and 43. The differences in the reachability sets (which list all the minimal sets of RCC-5 relations such that at least one relation is always reachable) are highlighted in grey.

With as many as 46 separate configurations or relations, it may often be helpful to cluster them. There are two conceptually different, but as it turns out, effectively identical ways of structuring the data. The first is to use the idea of possible RCC-5 relations between complete crispings of the configurations, as outlined earlier. This clusters the relations into 13 groups (see figure 7). The arrows between the groups represent the fact that one set of complete crispings is a subset of another (thus the singleton sets have no outward arrows). Alternatively, one can compute the sets of configurations which are mutual *partial* crispings of each other, i.e. for any two elements in such a set, each configuration can be partially crisped into any of the others, by expanding one or both yolks and/or shrinking one or both egg, but leaving each egg with a distinct yolk and white. It turns out that one obtains exactly the same set of 13 sets of configurations; in this case one can interpret the arrows of figure 7 as indicating that each configuration in the set the arrow stems from can be crisped to any relation in the set it points to. It is also fairly easy to find linguistic labels for the 13 clusters of configurations, based on the set of possible complete crispings. For example, the set {DR,PO,PPI}, as it applies to two egg-yolk regions X and Y , could be summed up by asserting that X is *not* a part of Y , and the set {DR,PO} as 'Neither region is part of the other'.

¹²It is interesting to note that these sets are essentially the minimal hitting sets of Reiter (Reiter 1987).

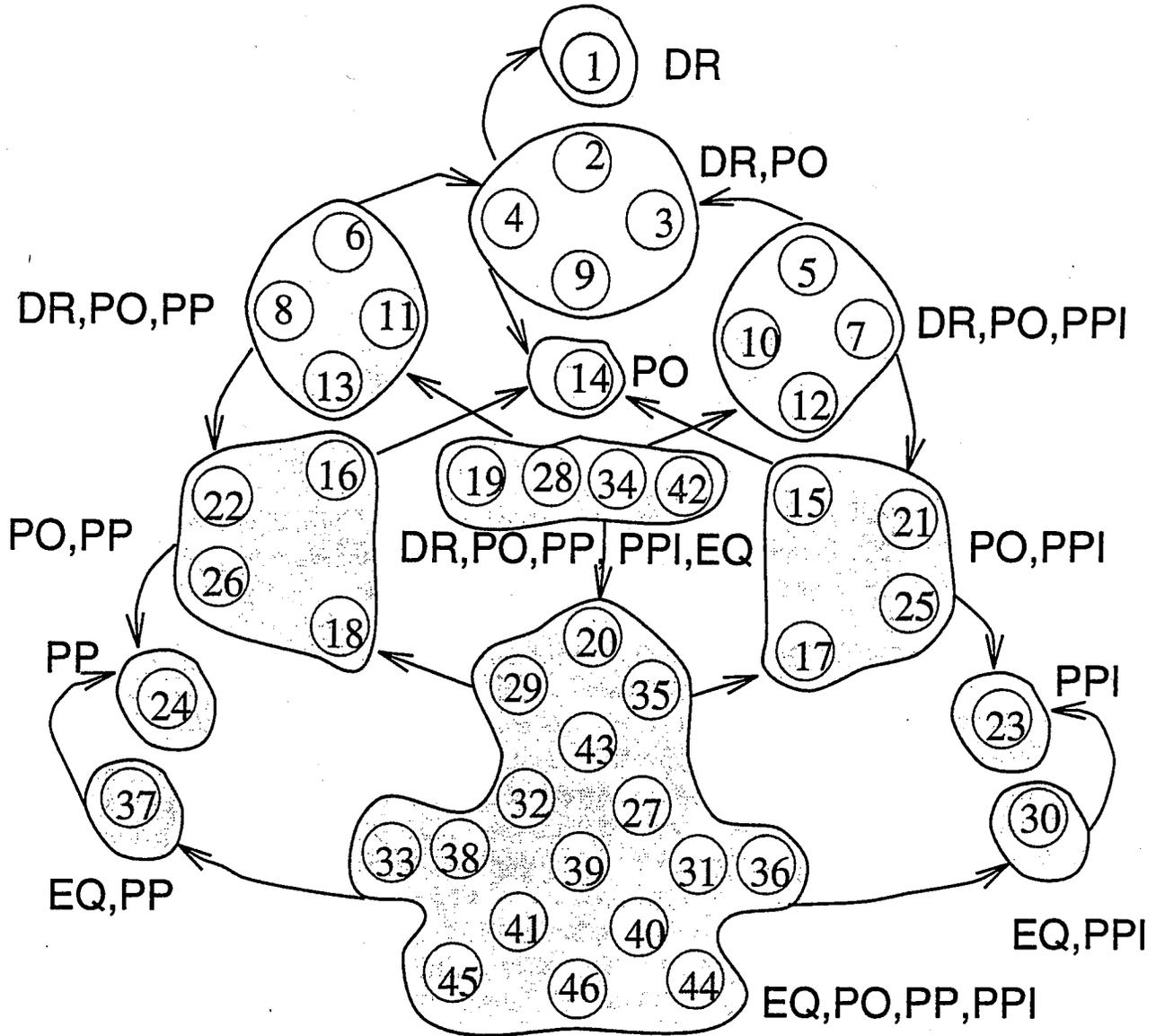


Figure 7: Clustering the 46 relations. Each group represents either a set of mutually crispable relations, or configurations whose complete crispings have the same RCC-5 relations. The arrows represent the crispings relationships or a subset relationship between the sets of complete crispings.

5 Related Work

Probably the closest piece of work to ours is that of Clementini and Di Felice (Clementini and Di Felice 1994) which was written simultaneously and entirely independently in response to the call for contributions to the GISDATA workshop on regions with undetermined boundaries which prompted our work. However, whereas our work is based on our previous logical formulation of spatial representation and reasoning, theirs is based on the point set theoretic approach inspired by Egenhofer and various co-workers (Egenhofer and Franzosa 1991, Egenhofer 1991, Egenhofer and Herring 1991, Egenhofer and Al-Taha 1992). They have generalised the '9-intersection' approach of Egenhofer and Herring (Egenhofer and Herring 1991). This classifies relations between pairs of 2D regions according to whether or not each of the interior, boundary and exterior of one region intersect (share any points with) each of the interior, boundary and exterior of the other — giving 9 possible intersections to consider in all.

If a pair of 2D regions with normal 1D boundaries are considered within this approach, then just 8 relations are obtained (Egenhofer 1991), corresponding precisely to those of RCC-8. However if we consider two 2D regions, each with a 2D boundary-zone around its edge, 44 possible relations are obtained. These correspond to our 46 egg-yolk relations: there are two less because the 9-intersection model as employed by Clementini and Di Felice cannot distinguish between 30 and 31, or between 37 and 38. In both these cases, the 9-intersection gives the same result for the two configurations. In particular, the boundary-zones of each pair of regions considered intersect, and no attempt is made to use the *dimensionality* of the intersection in their paper.

Clementini and Di Felice also eliminate 19, 28, 34 and 42 because they assume that the region of indeterminacy is very small in relation to the entire region. Figure 8 illustrates these four configurations: in each case each region's 'yolk' lies within the other's white.

Notice that these 4 regions form a cluster in our analysis, which seems a nice result: leaving out a set of regions which form such a cluster appears more reasonable than leaving out part of such a cluster. The set of complete crispings associated with this cluster is the entire RCC-5 set: for this set of configurations, nothing can be determined about the possible complete crispings.

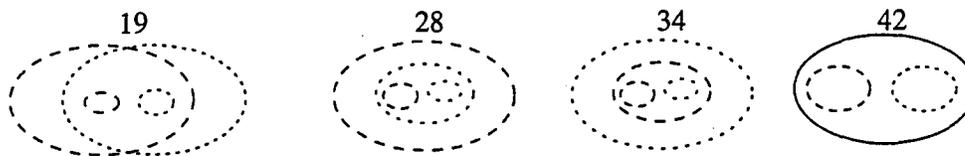


Figure 8: The four relations not included by Clementini and Di Felice.

Clementini and Di Felice do not use any notion similar to that of one region being a crisper version of another. Indeed, one can imagine they might not want to

use such a notion, since it appears that they are principally interested in modelling relations between pairs of regions, each representing a geographic entity with a (relatively narrow) 'transitional zone' around its edge. They group their relations into clusters, but this is done in a relatively adhoc manner, by an informal analysis of the relations and the assignment of intuitively reasonable names to the clusters they identify thereby. Their clustering (our numbering scheme is retained for consistency) is displayed in figure 9. Notice that the links in their diagram are 'closest topological distance' (ie minimum difference in the 9-intersections) rather than our 'is crisper than'. It is not certain what intuitive basis there is for the 'closest topological distance' relation. The two approaches produce similar but not identical clusterings.

6 Work In Progress

This paper reports work that is still in progress as we write. This section sketches some of the directions in which we are developing the ideas discussed.

The axiomatisation presented earlier makes relatively few commitments. There are however stronger versions which might be desirable in particular circumstances. For example, we might want to insist that crisping was 'dense', i.e. that:

$$\forall X, Y [X \prec Y \rightarrow \exists Z [X \prec Z \wedge Z \prec Y]]$$

or that any vague region could be made crisper:

$$\forall X [\text{NonCrisp}(X) \rightarrow \exists Y [\text{NonCrisp}(Y) \wedge Y \prec X]]$$

Notice that this would entail that any region had an infinite number of crispings.

Or we might want insist that every vague region could be made completely crisp:

$$\forall X [\text{NonCrisp}(X) \rightarrow \exists Y [Y \prec X]]$$

Analogously to axiom (A3) which ensured alternative crispings, we might want to insist that there exist alternative decrispings:

$$\forall X, Y [X \prec Y \wedge \exists W [Y \prec W] \rightarrow \exists Z [\prec Z \wedge \neg Y \prec Z \wedge \neg Z \prec Y]]$$

However the analogue of the existence of complete crispings, would be the existence of a 'complete decrisping' which would presumably entail the existence of a single very vague region which every other region would be crisper than.

$$\exists Z \forall W [W \neq Z \rightarrow W \prec Z]$$

Other axioms are of course possible, for example ones describing the notion of a crispest common decrisping analogous to a 'binary sum' mentioned earlier.

Another possibility not currently entertained by our current axiomatisation is whether the degree of vagueness about the boundaries of a region extend to the limiting case of the region having null extent (i.e. not in fact be a region after all). This would of course increase the number of possible egg-yolk relationships from 46. We are currently considering exactly how such an extension might be correctly effected.

The interpretation of vague regions in terms of pairs of limiting RCC-regions, appears to have considerable promise in reconciling our intuitions about spatial vagueness with a consistent formal treatment. However, there is at least one aspect of vagueness which this approach, at first sight, may appear to ignore. Suppose we want to consider the relationship between a hill and the neighbouring valley,

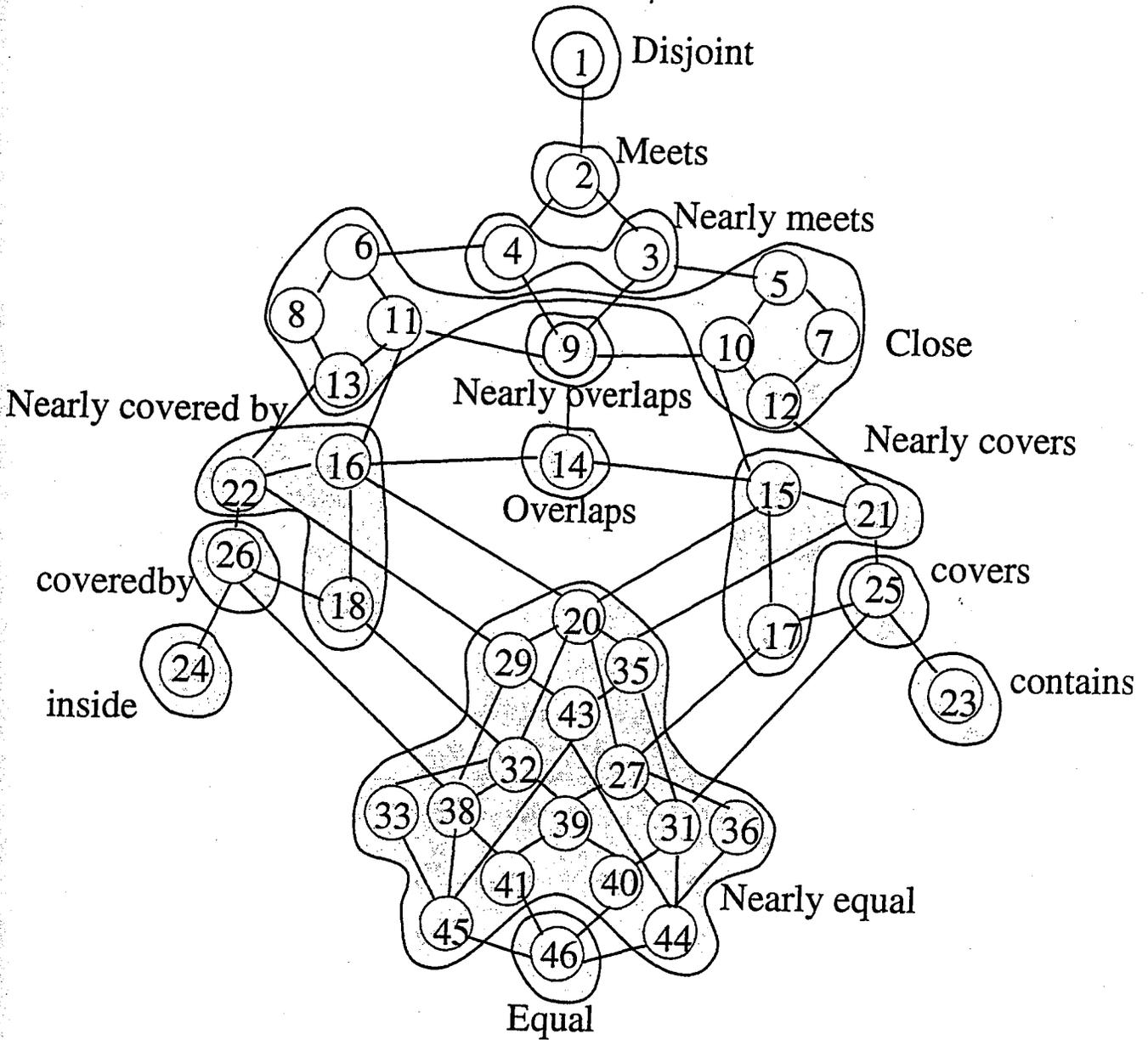


Figure 9: The clusters of Clementini and Felice (our numbering system). There are only 40 nodes, since 19, 28, 34, 42, 30 and 37 are not considered in their theory.

or a galaxy and the surrounding intergalactic space. It may not be possible to specify precisely *where* the boundary lies, but surely the two regions in each case are most naturally considered to be EC (Externally Connected, or touching without overlapping) — in terms of RCC-theory. In such a case, the choice of a complete cringing for one NonCrisp region limits our range of choices for another NonCrisp region (the two are nonindependent). We need to distinguish between the ‘within-model’ relation \prec — one region being a cringing of another — from the relation between more and less sharply-defined models of a real-world spatial configuration. Given a model of the hill-and-valley relationship in which the two are represented as NonCrisp regions, we need to be able to specify that in any more sharply-defined model in which the hill and valley are represented as Crisp regions, these will be EC to each other — that is, will meet along a common boundary but not overlap. We believe the axiomatisation of the relations between vague regions independent of the egg-yolk model we are developing will provide us with the necessary formal apparatus to express such constraints. This axiomatisation, together with other formal aspects of the work, is intended to appear in a forthcoming technical report.

We may also investigate variant notions of cringing where for example, cringing would not be allowed to change topology (e.g. the number of maximal one-piece subregions) or shape (as measured by concavities — see (Cohn et al. 1994)). Another aspect to be considered is temporal versus atemporal crings: the indefiniteness of a hill/valley might perhaps be considered essentially atemporal whilst a river might be considered essentially temporal (at any time the river is well defined, but it may change (though erosion or flooding) over time).

7 Conclusions

The analysis presented here provides an interesting and useful way of representing and reasoning about regions with indefinite boundaries. We have informally specified some properties which we believe such regions have, and then formally axiomatised them. We have also proposed a technique for representing such vague regions as pairs of traditional crisp regions — the advantage of this approach if it carries though fully is that reasoning about vague regions can be performed just by translating to RCC-theory and then using standard tools (e.g. composition tables) without having to build special purpose reasoning mechanisms for the calculus of regions with indefinite boundaries.

8 Acknowledgements

We gratefully acknowledge the support of the EPSRC under grant GR/H/78955 and also the CEC under the Basic Research Action MEDLAR 2, Project 6471. The first author had useful discussions with Eliseo Clementini. We also acknowledge the comments of various anonymous referees and of the participants at the Workshop on the Topological Foundations of Cognitive Science, at the First International Summer Institute in Cognitive Science, Buffalo, USA, July 1994.

References

- Bennett, B.: 1994, Spatial reasoning with propositional logics, in J. Doyle, E. Sandewall and P. Torasso (eds), *Principles of Knowledge Representation and Reasoning: Proceedings of the 4th International Conference (KR94)*, Morgan Kaufmann, San Francisco, CA.
- Clarke, B. L.: 1981, A calculus of individuals based on connection, *Notre Dame Journal of Formal Logic* **23**(3), 204-218.
- Clarke, B. L.: 1985, Individuals and points, *Notre Dame Journal of Formal Logic* **26**(1), 61-75.
- Clementini, E. and Di Felice, P.: 1994, An algebraic model for spatial objects with undetermined boundaries, in P. Burroughs and A. M. Frank (eds), *Proceedings, GISDATA Specialist Meeting on Geographical Entities with Undetermined Boundaries*, Taylor Francis. To appear.
- Cohn, A. G.: 1987, A more expressive formulation of many sorted logic, *Journal of Automated Reasoning* **3**, 113-200.
- Cohn, A. G. and Gotts, N. M.: 1994, The 'egg-yolk' representation of regions with indeterminate boundaries, in P. Burroughs and A. M. Frank (eds), *Proceedings, GISDATA Specialist Meeting on Geographical Objects with Undetermined Boundaries*, Francis Taylor. To appear.
- Cohn, A. G., Randell, D. A. and Cui, Z.: 1994, Taxonomies of logically defined qualitative spatial relations, in N. Guarino and R. Poli (eds), *Formal Ontology in Conceptual Analysis and Knowledge Representation*, Kluwer. To appear.
- Egenhofer, M.: 1991, Reasoning about binary topological relations, in O. Gunther and H. J. Schek (eds), *Proceedings of the Second Symposium on Large Spatial Databases, SSD'91 (Zurich, Switzerland). Lecture Notes in Computer Science 525*, pp. 143-160.
- Egenhofer, M. and Franzosa, R.: 1991, Point-set topological spatial relations, *International Journal of Geographical Information Systems* **5**(2), 161-174.
- Egenhofer, M. and Herring, J.: 1991, Categorizing binary topological relationships between regions, lines and points in geographic databases, *Technical report*, Department of Surveying Engineering, University of Maine.
- Egenhofer, M. J. and Al-Taha, K. K.: 1992, Reasoning about gradual changes of topological relationships, in A. U. Frank, I. Campari and U. Formentini (eds), *Theories and Methods of Spatio-temporal Reasoning in Geographic Space*, Vol. 639 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, pp. 196-219.
- Freksa, C.: 1992, Temporal reasoning based on semi-intervals, *Artificial Intelligence* **54**, 199-227.

- Gotts, N. M.: 1994, How far can we 'C'? defining a 'doughnut' using connection alone, in J. Doyle, E. Sandewall and P. Torasso (eds), *Principles of Knowledge Representation and Reasoning: Proceedings of the 4th International Conference (KR94)*, Morgan Kaufmann, San Francisco, CA.
- Lehmann, F. and Cohn, A. G.: 1994, The EGG/YOLK reliability hierarchy: Semantic data integration using sorts with prototypes, *Proceedings, Conference on Information Knowledge Management*, ACM. To appear.
- Randell, D. A.: 1991, *Analysing the Familiar: Reasoning About Space and Time in the Everyday World*, PhD thesis, University of Warwick.
- Randell, D. A. and Cohn, A. G.: 1989, Modelling topological and metrical properties of physical processes, in R. Brachman, H. Levesque and R. Reiter (eds), *Proceedings 1st International Conference on the Principles of Knowledge Representation and Reasoning*, Morgan Kaufmann, Los Altos, pp. 55-66.
- Randell, D. A. and Cohn, A. G.: 1992, Exploiting lattices in a theory of space and time, *Computers and Mathematics with Applications* **23**(6-9), 459-476. Also appears in "Semantic Networks", ed. F. Lehmann, Pergamon Press, Oxford, pp. 459-476, 1992.
- Randell, D. A., Cohn, A. G. and Cui, Z.: 1992, Naive topology: Modelling the force pump, in P. Struss and B. Faltings (eds), *Advances in Qualitative Physics*, MIT Press, pp. 177-192.
- Randell, D. A., Cui, Z. and Cohn, A. G.: 1992, A spatial logic based on regions and connection, *Proceedings 3rd International Conference on Knowledge Representation and Reasoning*, Morgan Kaufmann, San Mateo, pp. 165-176.
- Reiter, R.: 1987, A theory of diagnosis from first principles, *Artificial Intelligence* **32**, 57-95.
- Simons, P.: 1987, *Parts: A Study In Ontology*, Clarendon Press, Oxford.

MEREOTOPOLOGICAL CONSTRUCTION OF TIME FROM EVENTS

Fabio Pianesi and Achille C. Varzi

Istituto per la Ricerca Scientifica e Tecnologica (IRST)
I-38050 Povo (Trento), Italy

ABSTRACT. Recent work in temporal reasoning tends to reduce events to mere time intervals, or intervals *cum description*. In this paper we follow the opposite strategy, arguing that the formal connection between the way events are perceived to be ordered and the underlying temporal dimension is essentially that of a construction of a linear ordering from the mereotopological properties of an orientable structure including events as *bona fide* individuals. The account is expressed as a first-order theory using the primitives of parthood and boundary and can be extended—we venture to claim—to provide a similar treatment of spatial entities such as physical bodies and masses.

1. INTRODUCTION

We are used to regarding actions and other events, such as Brutus' stabbing of Caesar or the sinking of the Titanic, as occupying intervals of some underlying linearly ordered temporal dimension. This attitude is so natural and compelling that one tends to disregard the obvious difference between time intervals and actual happenings in favor of the former. Much recent work in temporal reasoning can in fact be viewed as embodying this reductionist approach, focusing on the sort of interval algebra that seems to be required for an adequate representation of time structures while ignoring the realm of events as such. This applies, for instance, to the systems developed under the impact of Allen's work

This is a revised and extended version of a paper appeared in the *Proceedings of the 11th European Conference on Artificial Intelligence* (A. G. Cohn ed.), Chichester: John Wiley & Sons, 1994, pp. 396–400. Section 2, Section 3 and the Appendix also draw from material originally published in the *Proceedings of the 9th Amsterdam Colloquium* (P. Dekker and M. Stokhof eds.), Amsterdam: Institute for Logic, Language and Computation, 1994, pp. 527–546. We are thankful to participants in both conferences for helpful discussion. We are also indebted to Roberto Casati, Gennaro Chierchia, Carola Eschenbach, Alessandra Giorgi, Fausto Giunchiglia, Jean Petitot, Chris Pinon, Barry Smith, and Graham White for illuminating comments on earlier drafts.

[1981, 1984], where events become mere “intervals cum description”. One good reason for doing so is of course that interval algebra is rather simple and yet rich enough to support the modelling of a variety of temporal reasoning problems. However, even from the perspective of artificial intelligence the point of talking about time is to talk about what actually happens or might happen at some time or another. And since different events overlap in so many different ways, we eventually need to deal with the full picture whether our concern be with planning, natural language processing, common sense metaphysics, or what have you.

This raises two questions. The first is whether we can actually go beyond time, as it were, i.e., whether we can take events as *bona fide* individuals and deal with them directly, just as we can deal with spatial entities such as physical bodies or masses without confining ourselves to their spatial representations. Philosophically this is a rather controversial issue, and ties in with a number of unsettled problems concerning e.g. the definition of adequate identity and individuation criteria for events (see Bennett [1988] and Pfeifer [1989] for an overview, and Casati & Varzi [1994b] for comprehensive bibliography). The second question is whether we can actually do without time, i.e., whether we can dispense with time intervals as an independent ontological category and focus only on actual happenings, in crude opposition to the standard form of reductionism mentioned above—in short, whether we can construe time from events. One suggestion in this direction is of course the classical account of Russell [1914, 1936], where time instants are construed as maximal sets of pairwise simultaneous (or partially simultaneous) events. This treatment is echoed in various later accounts, from Whitehead [1929] or Walker [1947] to Kamp [1979] and van Benthem [1983], where events are taken as primary entities inducing periods as secondary (and instants as merely tertiary) entities. More recently, Thomason [1989] has argued that the mathematical connection between the way events are perceived to be ordered and the underlying temporal dimension is essentially that of a free construction (in the category-theoretic sense) of linear orderings from event orderings, induced by the binary relation “wholly precedes”. In Pianesi & Varzi [1994] we put forward an alternative account based entirely on the mereological and topological characterization of event structures. The characterization was given on independent grounds—among other things to outline an answer to the first question above—and the possibility of reconstructing the temporal dimension directly from events was noted as an aside. In this paper we take on that result and develop it further with specific reference to both its formal ontological underpinnings and its relevance to the analysis of temporal discourse and reasoning.

2. COMBINING MEREOLGY AND TOPOLOGY

A major role of our general formal ontological background is played by the relation of parthood. Thus we sympathize with the view that mereology—the theory of parts and wholes, as rooted in the work of Leśniewski [1916] and Leonard & Goodman [1940]—provides a resourceful alternative to set theory for the formal analysis of common-sense reality. Specifically with respect to events, this view goes back to Whitehead [1919] and

has been defended e.g. by Thomson [1977] and, more recently, by Moltmann [1991] and Franconi *et al.* [1993]. At the same time, reasoning about the common-sense world also shows that a purely mereological outlook is too tight unless integrated with concepts and principles of a topological nature (Tiles [1981], Simons [1987, 1991a], Bochman [1990], Smith [1994a, 1994b]). For instance, mereologically there is no way to distinguish between a self-connected whole, such as a stone or a whistle, and a scattered entity made up of several disconnected parts, such as a broken glass, a soccer tournament, or Brutus' repeated stabbing of Caesar. Moreover, mereology alone cannot account for some very basic spatio-temporal relations among the entities of ordinary discourse, such as the relationship between an object and its surface, or the relation of continuity between two successive events, or the relation of something being inside, abutting, or surrounding something else. All of these are phenomena that run afoul of plain part-whole relations, and their systematic account requires a topological machinery of some sort.

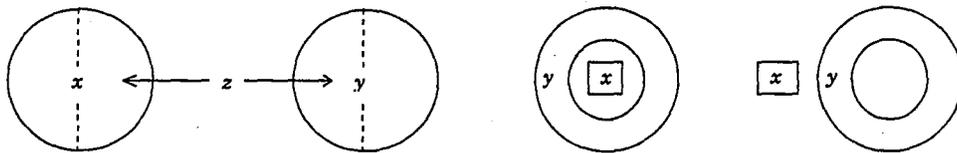


Figure 1. Left: pure mereological reasoning cannot do justice to the topological distinction between two discs x and y (on the one hand) and the scattered sum z of their facing halves (on the other). Right: the two patterns involve no mereological difference and topology is needed to keep track of the opposition in terms of inclusion of the square, x , inside the doughnut, y .

There are actually various ways in which the two domains of mereology and topology can be combined together (see Eschenbach & Heydrich [1993] and Varzi [1994] for a first assessment). One can see them as two independent provinces (following in the footsteps of *inter alia* Tiles [1981], Lejewski [1982], and Smith [1994a]); or one may grant priority to topology and characterize mereology derivatively, for instance defining parthood in terms of topological connection (as in Clarke [1981, 1985]). The latter approach is apparently more popular in AI, and has been applied e.g. to spatio-temporal reasoning (Randell & Cohn [1989, 1992], Randell *et al.* [1992a, 1992b]) and to the analysis of spatial prepositions in natural language (Vieu [1991], Aurnague & Vieu [1993]). Indeed this approach proves fit to account for a fair deal of mereotopological reasoning if we confine ourselves to an ontology of temporal intervals and/or spatial regions. If, however, we are to take an open-faced attitude towards real world entities and actual happenings (without identifying them with their respective spatio-temporal co-ordinates), then the reduction of mereology to a distinguished chapter of topology seems hardly tenable, as different entities can occupy exactly the same spatio-temporal regions (see Doepke [1982] or Simons [1985, 1986]). An object can be wholly located inside a hole, hence totally connected with it, without bearing any mereological relation to the hole itself (Casati & Varzi [1994a]). Or two events may have exactly the same topological connections and yet be mereologically distinct, as with the rotating and the becoming warm of a metal ball that is simultaneously rotating and becoming warm (example from Davidson [1969]).

In general, therefore, we are inclined to favor the first option mentioned above, treating mereology and topology as conceptually independent domains. Formally this is reflected in our use of two distinct primitives, viz. a pure mereological notion of *part* and a pure topological notion of *boundary*. Several other sets of primitives could of course do the job. However, our choice is not entirely arbitrary, at least for the topological part. In principle our favored strategy for combining mereology and topology can also be appreciated in some linguistics-oriented work on tense and aspect, such as Kamp's [1979], Bach's [1986], or Link's [1987] algebraic semantics for event structures, where a mereological relation defined on temporal entities is typically matched with an independent ordering of temporal precedence. For our present purposes, however, such a course would have a flavor of circularity. If the point is to provide a construction of time from the fundamental mereotopological features of event structures, it is essential that we take these features to be of a most general, time-independent nature. And in this sense reference to parts and boundaries appears to be a natural option both ontologically (Chisholm [1984], Smith [1992]) and from a cognitive standpoint (Jackendoff [1991]).

3. THE FORMALISM

3.1 Preliminaries. The entire mereotopological machinery can be developed within a first-order language with identity and descriptions. We shall use ' \neg ', ' \wedge ', ' \vee ', ' \rightarrow ', and ' \leftrightarrow ' as connectives for negation, conjunction, disjunction, material implication, and material equivalence respectively; ' \forall ' and ' \exists ' for the universal and existential quantifiers; and ' ι ' for the definite descriptor. To simplify readability, we shall rely on standard conventions minimizing the use of parentheses. In particular, we shall assume the five connectives to bind their arguments in decreasing order of strength as listed, so that negation binds the strongest and equivalence the weakest.

Note that the description operator will not have any use *per se*. However, it will play a crucial role in the definition of the term-forming operator of general sum by means of which several basic mereological and topological notions will be characterized. (Alternatively one could use set variables along with a set fusion operator, as in Tarski [1937] or Leonard & Goodman [1940], but this method would introduce additional complications and would be in contrast with the above-mentioned foundational outlook on mereology.) Since this sum operator may be undefined for some arguments, the underlying logical apparatus requires some means of accounting for the possibility of non-denoting expressions.

This of course can be done in a number of different ways. A preferred option is a free quantification theory with a supervaluational semantics (as in Bencivenga [1980] or Varzi [1983]). This would have the advantage of entertaining the unconditional acceptance of various "natural" principles that on a standard treatment *à la* Russell [1905], where descriptions are handled indirectly as improper symbols, would only be assertable under restriction. The following are some interesting cases in point (where 'Q' is any predicate and ' ϕ ' an arbitrary formula):

- (1) $Q(\iota x Q(x))$
- (2) $y = \iota x(x=y)$
- (3) $\iota x Q(x) = \iota x Q(x)$
- (4) $\phi(\iota x Q(x)) \leftrightarrow \phi(\iota x Q(x))$
- (5) $\exists y(y=\iota x \phi x) \leftrightarrow \exists y \forall x(\phi x \leftrightarrow x=y)$
- (6) $\forall y(y=\iota x \phi x \leftrightarrow (\phi y \wedge \forall x(\phi x \rightarrow x=y)))$

Each of these principles has been taken seriously by most free logicians ever since the pioneering work of Hintikka [1959] (see Lambert [1987] for an overview), and only a supervaluational semantics appears to be adequate to account for them in a systematic way. Unfortunately, however, this account is seriously defective from a computational perspective, as the set of valid principles is known to be not recursively enumerable.

In the face of this, among the several alternatives available (including Lesniewski's [1916] original approach) we find it convenient to rely instead on the minimal theory stemming from Lambert [1962], consisting in assuming (6) as the only specific principle for descriptive expressions. This theory is "minimal" insofar as it only captures the logic of ' ι ' with respect to descriptions that are denotationally successful, leaving open the issue of what principles should continue to hold in the presence of referential gaps. It does, however, allow us to treat descriptive terms as genuine singular terms, and this is all we need for our present purposes. In fact, with slight modifications of the original argument of Van Fraassen & Lambert [1967], it can be shown that a system obtained by adding axiom (6) to a free quantification theory is complete with respect to a semantics using partial models with classically saturated bivalent valuations. For definiteness, in the following we shall assume the underlying free quantification theory to be obtained from the classical predicate calculus by replacing the principle of Universal Instantiation with its universal closure:

- (7) $\forall y(\forall x \phi x \rightarrow \phi y)$.

Though too weak to validate (1), (2), or other specific principle for descriptions, this system is strong enough to secure the validity of (3), (4), and any other propositional or quantificational law for singular terms.

This minimalistic strategy has already been exploited in the context of mereological theorizing by Simons [1991b], and will prove sufficient also for our purposes in spite of their failing to reveal the whole truth. In any case, to facilitate comparisons we shall try to highlight those points where the choice of a different strategy may affect the theory.

3.2 Mereology. We symbolise the primitive mereological relation of parthood by ' P ', so that ' $P(x, y)$ ' reads " x is (a) part of y ". Derived notions, such as identity, overlapping, and the like, or the operations of sum, product, difference, etc. can be immediately defined:

- | | | | | |
|-----|-----------|----------|--------------------------------------|---------------------------|
| DP1 | $x=y$ | $=_{df}$ | $P(x, y) \wedge P(y, x)$ | x is identical with y |
| DP2 | $O(x, y)$ | $=_{df}$ | $\exists z (P(z, x) \wedge P(z, y))$ | x overlaps y |
| DP3 | $X(x, y)$ | $=_{df}$ | $O(x, y) \wedge \neg P(x, y)$ | x crosses y |

DP4	$PO(x, y)$	$\text{=df } X(x, y) \wedge X(y, x)$	x properly overlaps y
DP5	$PP(x, y)$	$\text{=df } P(x, y) \wedge \neg P(y, x)$	x is a proper part of y
DP6	$\sigma x \phi x$	$\text{=df } \iota x \forall y (O(y, x) \leftrightarrow \exists z (\phi z \wedge O(z, y)))$	sum of all ϕ ers
DP7	$\pi x \phi x$	$\text{=df } \sigma x \forall z (\phi z \rightarrow P(x, z))$	product of all ϕ ers
DP8	$x+y$	$\text{=df } \sigma z (P(z, x) \vee P(z, y))$	sum of x and y
DP9	$x \times y$	$\text{=df } \sigma z (P(z, x) \wedge P(z, y))$	product of x and y
DP10	$x \sim y$	$\text{=df } \sigma z (P(z, x) \wedge \neg O(z, y))$	difference of x and y
DP11	$\sim x$	$\text{=df } \sigma z (\neg O(z, x))$	complement of x
DP12	U	$\text{=df } \sigma z (z=z)$	universe

(In DP6-DP7, ' ϕx ' stands for any formula and ' ϕz ' for the result of substituting each free occurrence of ' x ' in ϕ by occurrences of ' z '.) Note that the functors/terms based on DP3 may be partially defined (i.e., correspond to an improper description) unless we go with the fiction of a null individual that is part of everything (as in Martin [1965, 1978]). For instance, non-overlapping entities will have no product and the universe will have no complement. This of course introduces a significant deviation in the obvious correspondence between mereological operations and the standard set-theoretical operations of union, intersection, difference, etc.

As purely mereological axioms we assume the following two, along with the standard axioms for identity:

- AP1 $P(x, y) \leftrightarrow \forall z (O(z, x) \rightarrow O(z, y))$
 AP2 $\exists x \phi x \rightarrow \exists x \forall y (O(y, x) \leftrightarrow \exists z (\phi z \wedge O(z, y)))$

AP1 secures that parthood is an extensional partial ordering while AP2 guarantees that every satisfied condition ϕ picks out an entity consisting of all ϕ ers. This yields a classical mereology as usually understood, corresponding to a Boolean algebra with zero deleted (Tarski [1935]; cp. Eberle [1970] and Simons [1987]). A sample selection of theorems relevant to the following is listed below:

- TP1 $P(x, x)$
 TP2 $P(x, y) \wedge P(y, z) \rightarrow P(x, z)$
 TP3 $x=y \leftrightarrow \forall z (P(z, x) \leftrightarrow P(z, y))$
 TP4 $x=y \leftrightarrow \forall z (P(x, z) \leftrightarrow P(y, z))$
 TP5 $x=y \leftrightarrow \forall z (O(x, z) \leftrightarrow O(y, z))$
 TP6 $PP(x, y) \rightarrow \exists z (P(z, y) \wedge \neg O(z, x))$
 TP7 $\neg P(x, y) \rightarrow \exists z (P(z, x) \wedge \neg O(z, y))$
 TP8 $\exists x \phi x \rightarrow \exists y (y = \sigma x \phi x)$

It may be worth recalling that none of these principles is uncontroversial. For instance, since Rescher [1955] several authors have had misgivings about such straightforward consequences of AP1 as TP2, expressing the transitivity of 'P' (see e.g. Cruse [1979], Winston *et al.* [1987], Moltmann [1991]), or TP3-TP5, expressing its extensionality (Wiggins [1979], Simons [1987]). In both cases, however, the objections involve reasoning about the variety of part-whole relations that may be distinguished (e.g., between components and complex, or quantity and mass) and may therefore be disre-

garded as long as we remain at a sufficiently general level of analysis. Even such an apparently innocent consequence of AP1 as the reflexivity of 'P' (TP1) might on some conditions be objected to, particularly insofar as the logical background that we are assuming allows for non-denoting terms. For instance, in this regard Simons [1991b] suggests applying the falsehood principle of Fine [1981] to deny that 'P(x, x)' can be true when x does not exist. However, such a stance would introduce a disturbing asymmetry between parthood and other basic predicates such as identity—of which parthood is a generalization—unless we are also ready to make a self-identity statement ' $x=x$ ' depend on the existence of x . This is in contrast with our general attitude towards free logic, which in fact follows the popular policy of assuming a standard identity theory. Therefore, also this type of objection will be disregarded in the following.

The second axiom is not uncontroversial either. For one thing, in the presence of AP1 it implies the so-called "supplementation" principles expressed by TP6-TP7, which some authors in the tradition of Brentano's *Kategorienlehre* [1933] found reason to deny (see e.g. Chisholm [1978]). There are actually cases where restrictions on the domain of interpretation might involve violating such principles (a disc with a disc removed is not a disc), but this is already a matter of *material* ontology and need not concern us for the moment. As a matter of generality, the existence of a remainder between a whole and a proper part can hardly be denied: otherwise it would be possible for an object or event to have a single proper part, and that goes against any intuitive understanding of the very meaning of 'part'. Likewise, AP1 has often been disputed for having counter-intuitive instances when ϕ is true of scattered or otherwise ill assorted entities or events, such as the totality of red things, or Brutus' birth and his stabbing of Caesar (see e.g. the early criticisms of Lowe [1953]). From a purely mereological perspective, however, this criticism is also off target. If you already have some things, allowing for their sum is no further commitment: the sum *is* those things (Lewis [1991]; that the sum is always *uniquely* defined is guaranteed by TP8). In any case, one may feel uncomfortable with treating unheard-of mixtures as individual wholes; but which wholes are more "natural" than others is not a mereological issue. As noted above, the question of what constitutes a natural, integral whole cannot even be formulated in mereological terms: it is precisely here that topology—the theory of boundaries—comes in.

3.3 Topology. The primitive topological notion of boundary is symbolized by 'B', so that 'B(x, y)' reads " x is a boundary in y ". (Following Chisholm [1984], we say "boundary *in*" (rather than *of*) to avoid a reductive interpretation of boundaries as *maximal* boundaries. In general, any boundary *in* something is a boundary *of* some part thereof.) Some useful derived notions can be defined as follows:

DB1	$b(x)$	$=_{df} \sigma z (B(z, x))$	maximal boundary of x
DB2	$c(x)$	$=_{df} x + b(x)$	closure of x
DB3	$i(x)$	$=_{df} x \sim b(x)$	interior of x
DB4	$e(x)$	$=_{df} \sim x \sim b(x)$	exterior of x
DB5	$Cl(x)$	$=_{df} x = c(x)$	x is closed
DB6	$Op(x)$	$=_{df} x = i(x)$	x is open

DB7	$T(x, y) =_{df} \exists z (P(z, x) \wedge B(z, y))$	x is a tangential part of y
DB8	$l(x, y) =_{df} \exists z (Op(z) \wedge P(x, z) \wedge P(z, y))$	x is an interior part of y
DB9	$C(x, y) =_{df} O(c(x), y) \vee O(c(y), x)$	x is connected with y
DB10	$EC(x, y) =_{df} C(x, y) \wedge \neg O(x, y)$	x externally connects to y
DB11	$ST(x, y) =_{df} \forall z (l(x, z) \rightarrow X(z, y))$	x straddles y
DB12	$Cn(x) =_{df} \forall y \forall z (x=y+z \rightarrow C(y, z))$	x is self-connected
DB13	$CE(x, y) =_{df} Cn(x) \wedge P(y, x)$	x is a convex extension of y
DB14	$CC(x, y) =_{df} CE(x, y) \wedge \forall z (PP(z, x) \rightarrow \neg CE(x, y))$	x is a convex closure of y

Again, these functors, predicates, and relations may not be defined for certain arguments, as they may involve improper (e.g., denotationless) definite descriptions.

Note that nothing in these definitions implies that boundaries are always parts of the entities they bound. In fact we accept the standard topological distinction between open and closed entities, allowing for entities with external boundaries. For example, in Pianesi & Varzi [1994] we propose a characterization of the standard classification of event types (Vendler [1957]) by treating processes (such as *climbing the mountain*) as non-closed events, accomplishments (*having climbed the mountain*) as closed processes, and achievements (*reaching/having reached the top*) as parts or the corresponding boundaries.

In this regard, our basic axiomatization is in line with the familiar Kuratowski axioms [1922], which we assume in the following form:

AB1	$P(x, c(x))$
AB2	$c(c(x)) = c(x)$
AB3	$c(x+y) = c(x) + c(y)$

This gives us a straightforward reformulation of much of standard topology based on mereology instead of set theory, provided only that we take care in handling undefined operators. In particular, it follows immediately that the relation of connection is reflexive and symmetric and that boundaries are always symmetrical, in the sense that every boundary of an entity is also a boundary of the entity's complement (see Chisholm [1992/93] and Smith [1994b] for a discussion of non-symmetric boundaries). Here is a list of further theorems that can be proved from AB1–AB3 and that will be used in the following developments:

TB1	$B(x, y) \wedge B(y, z) \rightarrow B(x, z)$
TB2	$P(x, y) \wedge B(y, z) \rightarrow B(x, z)$
TB3	$l(x, y) \wedge P(y, z) \rightarrow l(x, z)$
TB4	$P(x, y) \wedge l(y, z) \rightarrow l(x, z)$
TB5	$B(x, y) \leftrightarrow \forall z (P(z, y) \rightarrow ST(z, y))$
TB6	$B(x, y) \leftrightarrow \forall z (P(z, y) \rightarrow T(z, y))$
TB7	$\forall x (\phi x \rightarrow B(x, y)) \rightarrow B(\sigma x \phi x, y)$

The last of these theorems is particularly important, as it shows that boundaries are closed under general sum and are therefore closed under all mereological properties. Following Smith [1994a], however, we also wish to capture some further common-sense intuitions that go beyond the repertoire of standard topology. In particular, we need at

least a rendering of the intuitive Aristotelian-Brentanian idea that boundaries are “parasitic” entities, i.e., cannot exist independently of the larger entities they bound (Brentano [1976]; cp. also Chisholm [1989], Bochman [1990], Smith [1992], White [1993]). This stands in opposition to the standard set-theoretic conception of boundaries as sets of ordinary, ontologically independent points. More specifically, we assume that every self-connected boundary is a boundary part of some larger self-connected entity with a non-empty interior:

$$AB4 \quad Cn(x) \wedge \exists y B(x, y) \rightarrow \exists z \exists w (Cn(z) \wedge B(x, z) \wedge P(x, z) \wedge I(w, z))$$

It is understood that further principles would be needed to obtain at least a rough approximation of the folk theory of spatio-temporal continua. Here, however, we shall content ourselves with AP1-AP2 and AB1-AB4, regarding this as a minimal theory for general purposes.

It is worth observing that, within these limits, the main connection between our mereological and topological notions is neatly expressed by the following theorem:

$$TPC \quad P(x, y) \rightarrow \forall z (C(z, x) \rightarrow C(z, y))$$

As already hinted at above, systems in the tradition of Clarke [1981] also assume the converse of this principle, with the effect of reducing mereology to topology. By contrast, the possibility that topologically connected entities bear no mereological relationship to one another leaves room for a much richer taxonomy of basic mereotopological relations than usually recognized (cp. Varzi [1993] and Cohn *et al.* [1993]). For instance, the customary relations of connection, overlapping, parthood, and interior parthood introduced above, and common to most known systems, can be integrated by the following (Varzi [1993]):

- | | | |
|------|--|----------------------------|
| DB15 | $E(x, y) =_{df} \forall z (C(z, x) \rightarrow C(z, y))$ | x is enclosed in y |
| DB16 | $S(x, y) =_{df} \exists z (E(z, x) \wedge E(z, y))$ | x is superimposed on y |
| DB17 | $A(x, y) =_{df} C(x, y) \wedge \neg S(x, y)$ | x abuts y |
| DB18 | $W(x, y) =_{df} E(x, I(y))$ | x is within y |

Evidently S is implied by O , E by P , and W by I . The resulting taxonomy of relations is depicted in the following picture.

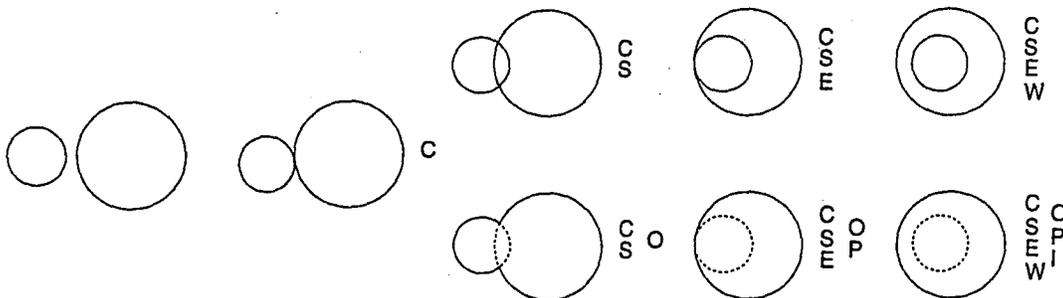


Figure 2. Some basic mereotopological relations exploiting the distinction between mereological parthood (overlapping, etc.) and mere topological enclosure (superposition, etc.).

4. EVENT STRUCTURES

Let us now see how our general mereotopological framework can be specialized to a domain of events. As already pointed out, on a rather popular conception events are regarded as intrinsic temporal entities, i.e., entities that stem, so to speak, from the temporal dimension. In forms ranging from the strong reductionist view that events are nothing but temporal intervals to the weaker forms that take events as primitive entities endowed with some primitive temporal relation defined on them (the strict ordering of Kamp [1979] or Thomason [1989]), this view has been a predominant approach in AI, natural language semantics, and knowledge representation. By contrast, in the following we will show that events can be given an independent characterization that satisfies our intuitions about their mutual relations and derives the whole of time, temporal relations included, by imposing suitable restrictions on the underlying mereotopological apparatus. More precisely, we will show that the formal connection between the way events are perceived to be ordered and the underlying temporal dimension is essentially that of a construction of a linear ordering from the mereotopological properties of an oriented domain structure in which events are included as *bona fide* individuals. In other words, time is a by-product of what we call an *event structure*.

4.1 Divisors. Our characterization of event structures relies on the auxiliary concept of a *divisor*. Intuitively, a divisor is a somewhat complex event that separates the entire domain into two disconnected parts, thus making it possible to choose one part as corresponding to the sum total of all events that temporally precede the divisor, and the other as the sum of those events that follow it. (Of course, the choice of precedence vs. following will be arbitrary, as long as successive choices for different divisors be done in a consistent way: if we think of an event domain as comprising the totality of all happenings—past and future—there is no *a priori* way to fix the temporal orientation.) Formally, this is obtained with the help of both mereological and topological notions by requiring every divisor to split all of its neighbourhoods into disconnected parts:

$$\text{DE1} \quad D(x) =_{\text{df}} \forall y (I(x, y) \rightarrow \neg Cn(y \sim x))$$

That this is essentially an analogue of the 1-codimensionality property that in standard set theory characterizes points with respect to the line, lines with respect to the plane, and so on (see White [1993]).

Note that a divisor is itself an event, though not every event is a divisor. We may think of a divisor as a sort of cross event made up of all that happens during a certain “period”. By contrast, an action such as Brutus’ stabbing of Caesar, or an incident such as the sinking of the Titanic, are “local” events: many other actions and events occurred at the same time, but in different places. This is of course an enrichment arising from our rejection to identify events with the intervals that they occupy. Events are fully-fledged entities, though we need not for this reason see them as endowed with a multidimensional part structure distinguishing for instance between spatial and temporal mereological relations (see Moltmann [1991] for a proposal in this direction).

Note also that if we refer to divisors for the purpose of characterizing the notions of past and future, these latter will be deprived of any absolute meaning and become relative notions: given any divisor x , the suggestion is to interpret the events on one side of x as past events, and the ones on the other side as future events, *relative to x* . There is no past or future except with respect to some division of the whole of history. As will become clearer shortly, a major business that we need to face is therefore to guarantee that such a relativistic account be nevertheless suited to the task; that is, we must make sure that the all divisors partition the past from the future in a coherent way. (For instance, an event that lies in the past relative to a given divisor x must also lie in the past relative to any divisor y that lies in the future of x .)

We also want our construction to allow for a certain degree of control on its grain. Intuitively, the idea is that any two events that are part of the same divisor should count as simultaneous. However, for that purpose we need some means of associating each event in the domain with the "right" divisor, as it were, viz. the *minimal* one containing it. (Otherwise, for any two events x and y we could pick out a sufficiently large divisor z containing both.) Moreover, the predicate 'D' picks out the distinguishing property of a divisor relative to the full mereotopological structure of the given domain of events, but the fact remains that any domain admits of an indefinite, potentially infinite number of divisors, only some of which can (or need) be collected in a temporally coherent structure. In other words, not all divisors can or need be considered together. Thus, every domain can be associated with various dividing devices δ (including divisors that ignore some mereological distinctions, treating e.g. as atomic events which are in fact mereologically complex), giving rise to a variety of event structures. Furthermore, the possibility of varying the grain itself may be a welcome one. This can be helpful, for instance, to account for the various degrees of precision that natural language permits when talking about events and time.

4.2 Structures. Given all of this, we define an *event structure* quite generally as a triple $\langle \mathcal{E}, \delta, f \rangle$ made up of a domain of events, \mathcal{E} , a dividing device δ , and a function f of orientation. The conditions are as follows. First, we assume \mathcal{E} to be mereotopologically self-connected:

$$\text{AE1} \quad \forall z(O(z, x) \vee O(z, y)) \rightarrow C(x, y)$$

We are here taking \mathcal{E} as a domain closed under all mereological and topological principles set forth in the previous section, understanding individual variables to range over this domain. Thus, condition AE1 could also be written as

$$\text{AE1}' \quad C_n(U)$$

which effectively corresponds to the statement that the whole universe is self-connected. This is stipulative, but reflects the idea that there are no gaps in history: something is always happening, whether remarkable or not. (Compare the discussion in Newton-Smith [1980]. As a matter of fact, if the universe consisted of two or more separate parts, one could still apply the reasoning below to each of them, ending up with a family of discon-

nected worlds each having its own temporal ordering. However here we shall not pursue this possibility to keep things simple and intuitively straightforward.)

Second, δ is a divisor-specific condition incorporating the above-mentioned requirements of minimality and granularity. It satisfies the following axioms:

- AE2 $\delta(x) \rightarrow D(x)$
 AE3 $\sigma x \delta(x) = \sigma x D(x)$
 AE4 $\forall x (\phi x \rightarrow \delta(x)) \rightarrow \delta(\pi x \phi x)$
 AE5 $\forall x (\phi x \rightarrow \delta(x)) \rightarrow \forall z (CC(z, \sigma x \phi x) \rightarrow \delta(z))$

In other words, the events fulfilling δ are all divisors (AE2), covering the entire domain of divisors (AE3) and such that the product or connected sums of every number of them is itself a divisor (AE4–5). These latter conditions are illustrated in Figure 3. In set-theoretic terms, they have the effect of making divisors into a closure system. The corresponding closure operator associates each event x with the *smallest divisor* containing x . It can be defined as follows:

$$DE2 \quad d(x) =_{df} \pi z (\delta(z) \wedge P(x, z))$$

That this is indeed a closure operator is guaranteed by the following theorems, which are the analogues of the usual increasing, idempotency, and monotonicity conditions:

- TE1 $P(x, d(x))$
 TE2 $d(d(x)) = d(x)$
 TE3 $P(x, y) \rightarrow P(d(x), d(y))$

Finally, the last term in the definition of an event structure is a (possibly partial) map $f: \mathcal{E} \rightarrow \mathcal{E}$ closed under the following mereological conditions:

- AE6 $f(x) = f(d(x))$
 AE7 $f(x) + f'(x) = \sim d(x)$
 AE8 $P(x, f(y)) \rightarrow P(f(x), f(y))$
 AE9 $\delta(x) \wedge \delta(y) \wedge O(y, f(x)) \wedge O(y, f'(x)) \rightarrow P(x, y)$

where in general

$$DE3 \quad f'(x) =_{df} \sigma z (P(x, f(z))).$$

As we shall see below, our approach will be to think of f as a function of temporal orientation associating each event in the given domain with the totality of events that precede it. Correspondingly, f' will map each event to its future events, and the above axioms will secure that these two mappings be coherent throughout the domain. (As already noted, this choice is purely conventional. We could as well turn things around, taking f and f' to indicate future and past events, respectively.) In particular, AE6–AE7 relate this intended reading of f and f' to the intuitive interpretation of the divisor operator δ ; AE8 relates it to the underlying mereological setting; and AE9 relates it to both, making sure that all divisors operate in a parallel fashion, i.e. have a uniform temporal orientation. The behavior of these functions is illustrated in figures 4 and 5 below. (Note that DE3 guaran-

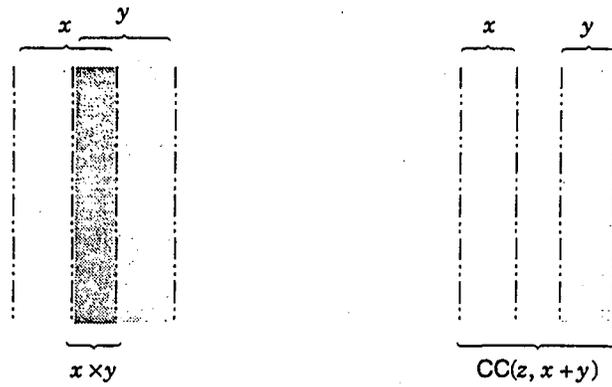


Figure 3. Divisors are closed under product (left) and connected sum (right).

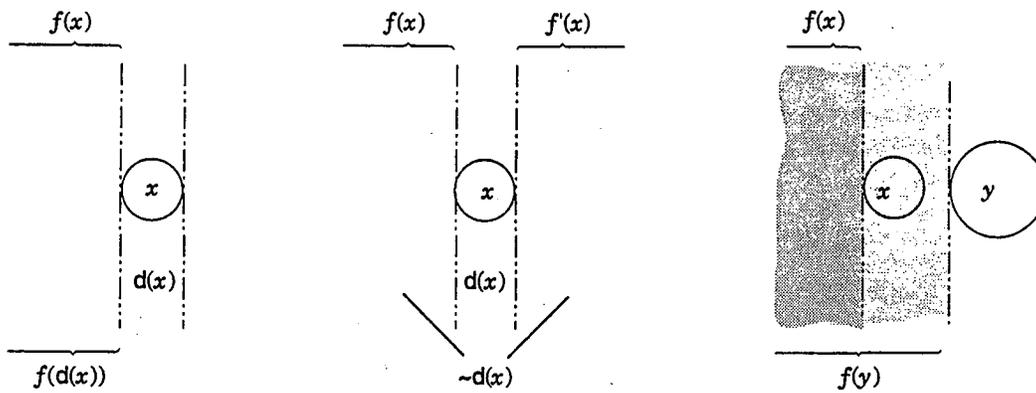


Figure 4. Pictorial representation of the mereological relations expressed by axioms AE6 (left), AE7 (middle), and AE8 (right).

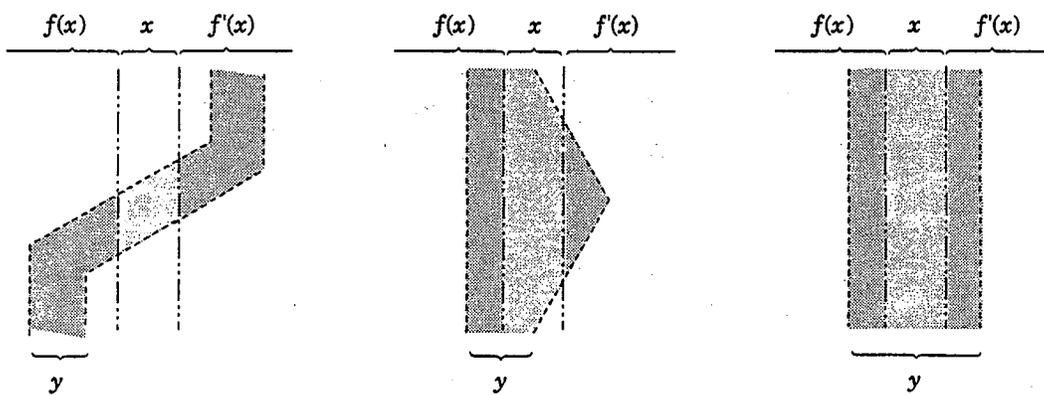


Figure 5. The patterns on the left and in the middle violate, whereas the one on the right fulfills, the requirement that a divisor y cannot overlap both "past" and "future" of another divisor x unless the former includes the latter (AE9).

tees perfect duality between f and f' ; thus, just like AE7 and AE9 are symmetric with respect to these two functions, likewise we can immediately prove analogues of AE6 and AE8 with f' in place of f :

- TE4 $f'(x) = f'(d(x))$
 TE5 $P(x, f'(y)) \rightarrow P(f'(x), f'(y))$

Similar facts will always hold in the following, though we shall generally confine ourselves to spelling out facts and principles relative to f .)

Further constraints on \mathcal{E} , δ , or f can of course be added. For instance, one may want to rule out the possibility that there be atomic events (i.e., events with no proper parts) that are open. Or one may want to impose stronger conditions on the granularity of δ (such as the requirement that an event's divisor be never "thicker", i.e., of longer duration than the event itself) or on the self-connectedness of \mathcal{E} (such as density or, alternatively, discreteness: the former asserts that between two successive events there is always a third one; the latter amounts to the opposite requirement that every event has some immediate successor and some immediate predecessor). In our formalism these conditions would correspond to the following postulates respectively (the last two corresponding to the two sides of discreteness):

- AE10 $Op(x) \rightarrow \exists y PP(y, x)$
 AE11 $O(x, d(y)) \leftrightarrow O(d(x), d(y))$
 AE12 $P(x, f(y)) \rightarrow \exists z (P(x, f(z)) \wedge P(z, f(y)))$
 AE13 $\exists y P(x, f(y)) \rightarrow \exists z (P(x, f(z)) \wedge \neg \exists w (P(x, f(w)) \wedge P(w, f(z))))$
 AE13' $\exists y P(x, f'(y)) \rightarrow \exists z (P(x, f'(z)) \wedge \neg \exists w (P(x, f'(w)) \wedge P(w, f'(z))))$

We shall not address these issues here. For our purposes, let us only stress once again that the operators and mappings introduced above may possibly be undefined for some arguments, and that different models may therefore be obtained according to the underlying logic.

5. THE CONSTRUCTION OF TIME

As mentioned above, our purpose is to consider f as a function of temporal orientation (relative to a given event structure $\langle \mathcal{E}, f, \delta \rangle$). To this end, note first of all that the following hold for all events x and y in the given domain \mathcal{E} :

- TE6 $P(x, f(y)) \rightarrow \neg O(x, f'(y))$
 TE7 $P(x, f(y)) \leftrightarrow P(y, f'(x))$
 TE8 $P(x, y) \rightarrow P(f(y), f(x))$
 TE8' $P(x, y) \rightarrow P(f'(y), f'(x))$

As illustrated in figure 6 below, the first of these conditions reflects the idea that past and future (relative to a given event x) do not overlap; the second principle reflects the idea that whatever happens before a certain event y must be such as to include y among its future events and, viceversa, whatever happens after an event x must include x among its past events; while the third and fourth principles express the idea that whenever an event is included in another, the past and future of the latter must be included in the past and future of the former, respectively.

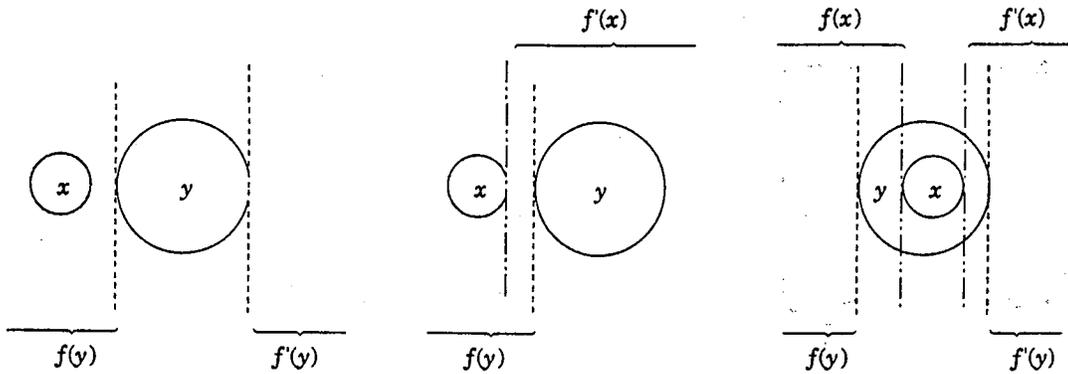


Figure 6. Pictorial representation of the mereological relations expressed by TE6 (left), TE7 (middle) and TE8–8' (right).

As we have it, of course, only events that have a (minimal) divisor can be fully matched for precedence, since f or f' may otherwise be undefined. This is an important consequence of AE7:

$$\text{TE9} \quad \exists y (y=f(x)) \wedge \exists y (y=f'(x)) \rightarrow \exists y (y=d(x))$$

In particular, it follows that $f(f(x))$ and $f'(f'(x))$ are always undefined. (For what would it *mean* to say that “the past” and “the future” have a past and future of their own?) Furthermore, note that $f'(x)$ is not defined whenever $f(x)$ is defined (unless $d(x)$ is already defined) and vice versa. Accordingly, when considering properties such as those in TE6–TE8', one should always keep in mind that they may come out undefined in the case of events infinitely reaching backwards or forwards. Alternatively, we can think of such events as having a degenerate future or a degenerate past, respectively. This is captured in the following equivalences:

$$\text{TE10} \quad f(x) = \sigma z (P(x, f'(z))) = \pi z (\exists y PP(y, x) \wedge z=f(y))$$

$$\text{TE10}' \quad f'(x) = \sigma z (P(x, f(z))) = \pi z (\exists y PP(y, x) \wedge z=f'(y))$$

Thus, f yields the “limit” past of an event, which tends to nil in the case of events infinitely reaching backwards—and likewise for f' (future).

This provides grounds for the intended interpretation of the relative properties expressed by TE6–TE8'. Two further interesting theorems are the following:

$$\text{TE11} \quad f'(f(x)) = \sim(f(x))$$

$$\text{TE11}' \quad f(f'(x)) = \sim(f'(x))$$

These theorems, in a sense, generalize AE7 yielding the following corollaries:

$$\text{TE12} \quad P(x, f'(f(x)))$$

$$\text{TE12}' \quad P(x, f(f'(x)))$$

Together with TE8 and TE8', these ensure that f and f' behave as a pair of Galois connection in \mathcal{E} . Moreover, it is worth observing that composing f and f' induces corresponding topological structures. For let $g=ff'$; then we can prove:

- TE13 $P(x, g(x))$
 TE14 $g(g(x)) = g(x)$
 TE15 $g(x+y) = g(x) + g(y)$
 TE16 $P(x, y) \rightarrow P(g(x), g(y))$

(likewise for $g=f'f$). Thus, the interaction of f and f' yields a well-behaved topological closure operator. (This may correspond to the topology of left (ff') or right ($f'f$) unbounded temporal intervals.) Particularly, TE14 shows that their composition reaches a fixed point immediately after two applications.

All these facts, then, allow us to think intuitively of \mathcal{E} as a domain of events and f as a function of *temporal orientation*, as desired. In particular, we can say that an event x *temporally precedes* (wholly) an event y iff x is part of $f(y)$, while x and y *temporally overlap* iff they have some parts whose minimal divisors overlap:

- DE6 $TP(x, y) \text{ =df } P(x, f(y))$
 DE7 $TO(x, y) \text{ =df } \exists z \exists w (P(z, x) \wedge P(w, y) \wedge O(d(z), d(w)))$

(In DE6, we can of course use the same definiens to introduce the converse relation of y *temporally following* x ; by TE7, this will be tantamount to DE8:

- DE8 $TF(y, x) \text{ =df } P(y, f'(x))$.

As for DE7, reference to parts is necessary since x and y may not have a divisor.) Note that parthood excludes precedence (following) whereas overlap implies temporal overlap:

- TE17 $P(x, y) \rightarrow \neg TP(x, y)$
 TE18 $O(x, y) \rightarrow TO(x, y)$.

Note also that these relations allow one to introduce an entire family of additional notions, in analogy with the basic mereological setting. For instance, one can define relations of *temporal inclusion*, *coincidence*, *straddling*, etc. in an obvious way:

- DE9 $TI(x, y) \text{ =df } \forall z (TO(x, z) \rightarrow TO(y, z))$
 DE10 $TC(x, y) \text{ =df } TI(x, y) \wedge TI(y, x)$
 DE11 $TS(x, y) \text{ =df } TO(x, y) \wedge \neg TI(x, y)$.

These can easily be seen to behave in analogy to the corresponding mereological relations of parthood, identity, crossing, etc. (for instance, TI is a partial ordering and TC an equivalence relation). In fact, it is apparent that in the case of divisors, these temporal relations reduce to their basic mereological correlates:

- TE19 $\delta(x) \wedge \delta(y) \rightarrow TO(x, y) \leftrightarrow O(x, y)$
 TE20 $\delta(x) \wedge \delta(y) \rightarrow TC(x, y) \leftrightarrow x=y$
 TE21 $\delta(x) \wedge \delta(y) \rightarrow TI(x, y) \leftrightarrow P(x, y)$
 TE22 $\delta(x) \wedge \delta(y) \rightarrow TS(x, y) \leftrightarrow X(x, y)$

This should shed some light on the temporal dimension implicit in our event structures, where non-dividing events can be (partially) simultaneous without necessarily bearing mereological relations to one another.

Finally, and most importantly, using the above notions we are now ready to prove all of the following:

- TE23 $TO(x, x)$
- TE24 $TO(x, y) \rightarrow TO(y, x)$
- TE25 $TP(x, y) \rightarrow \neg TO(x, y)$
- TE26 $TP(x, y) \rightarrow \neg TP(y, x)$
- TE27 $TP(x, y) \wedge TP(y, z) \rightarrow TP(x, z)$
- TE28 $TP(x, y) \wedge TO(y, z) \wedge TP(z, t) \rightarrow TP(x, t)$
- TE29 $TP(x, y) \vee TP(y, x) \vee TO(x, y)$

These are the mereological counterparts of the seven axioms for strict linear orders employed by Kamp [1979] in his construction of time instants out of ordered events. Counterparts of the axioms for interval structures of van Benthem [1983] can be proved in a similar way (e.g. by taking divisors as the counterparts of intervals and then reasoning in terms of the relations of precedence and temporal overlapping as introduced above).

Thus, although silent about time, event structures ultimately permit a full retrieval of the temporal dimension. Time is not a primary notion — be it a purely relational one or in the form of an independent ontological category (intervals or instants). It is a by-product of the possibility of orienting the domain of all happenings.

APPENDIX. PUNCTUAL EVENTS

Event structures are strongly dependent on the choice of a specific condition δ : divisors are events that separate the past from the future; but what exactly is to count as a relevant divisor, i.e. what is the granularity of the temporal segmentation, is not fixed once and for all. Actually, one may observe that a structure's orientation function f —and, consequently, its correlate f' —is completely determined by δ , free choice being limited to associating it with the totality of preceding events or with the totality of events that follow. Thus, varying the conditions on the component δ is the main formal instrument at our disposal to approach different issues within the uniform framework provided by event structures.

One notion where this variability becomes crucial is that of punctuality. Intuitively, punctual events are instantaneous, i.e., do not extend over any time *interval*: they are *located in* time but do not *take up* time. As hinted at in Section 3.3, these include for instance boundary events traditionally classified as “culminations” or “achievements” (Vendler [1957]). Within the proposed setting, however, this does not amount to a requirement of atomicity: what counts as instantaneous, as opposed to extended in time, depends entirely on the temporal structure that is induced by the domain of events on which the discourse is being interpreted—that is, ultimately, on the choice of the divisor condition δ in the relevant event structure. Thus, punctuality must also be accounted for in relative terms. Punctuality requires some sort of minimality, in the intuitive sense that a punctual event cannot accommodate more structured ones. However, we need not consider the distinction between instants and intervals, and more generally any distinction

based on some absolute notion as size or duration, to be the relevant parameters (contrary to a rather standard practice following Kamp [1979]). We also need not impose any specific axioms for characterizing punctuality (also contrary to Kamp's approach). Rather, the distinguishing properties of punctual events and instant algebras will be derived from more basic aspects of our constructions.

As a first step, let us introduce the notion of a *minimal divisor* (relative to a given event structure $\langle \mathcal{E}, \delta, f \rangle$):

$$\text{DA1} \quad \text{MD}(x) =_{\text{df}} \delta(x) \wedge \forall y(P(y, x) \rightarrow \neg\delta(y))$$

Thus, a divisor x is minimal iff it does not contain other divisors (relative to the same structure). As a consequence, every event which is part of such an x has x as its divisor:

$$\text{TA1} \quad \text{MD}(x) \wedge P(y, x) \rightarrow d(y)=x$$

In a sense, "temporal" differences are neglected inside a minimal divisor; any events that are parts of such a divisor are co-temporaneous:

$$\text{TA2} \quad \text{MD}(x) \wedge P(y, x) \wedge P(z, x) \rightarrow \text{TC}(y, z)$$

More generally, we have

$$\text{TA3} \quad \text{MD}(x) \wedge \text{TO}(y, x) \wedge \text{TO}(z, x) \rightarrow \text{TO}(y, z)$$

$$\text{TA4} \quad \text{MD}(x) \wedge P(w, x) \wedge \text{TO}(y, w) \wedge \text{TO}(z, w) \rightarrow \text{TO}(y, z).$$

Thus, if two events temporally overlap a minimal divisor (or a part thereof) they temporally overlap each other. Vice versa, we have that:

$$\text{TA5} \quad \delta(x) \wedge \forall y \forall z(\text{TO}(y, x) \wedge \text{TO}(z, x) \rightarrow \text{TO}(y, z)) \rightarrow \text{MD}(x).$$

Putting TA3 and TA5 together, the fundamental properties characterizing punctual events according to Kamp [1979] can be shown to hold of minimal—and only minimal—divisors. We can then propose the following definition for punctual events:

$$\text{DA2} \quad \text{PE}(x) =_{\text{df}} \text{MD}(d(x)).$$

Thus, punctual events are not just—and not necessarily—atomic events, i.e. events with no proper parts (though of course every atomic event is punctual regardless of δ). Rather, they are events whose internal structure is irrelevant for the purpose of temporal distinctions.

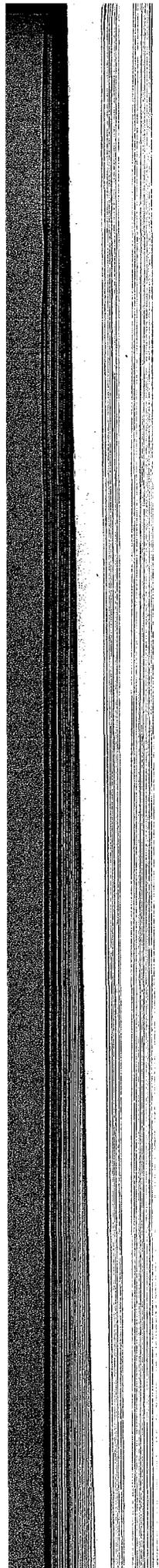
Moreover, a major advantage of this notion of punctuality (over Kamp's, for instance) is that it is relativized with respect to the particular event structure at hand (and, ultimately, with respect to the particular divisor condition δ), while retaining Kamp's basic insights. By changing δ , events previously treated as punctual may become non-punctual, in that their internal temporal structure is made available, and vice versa. On the other hand, once we declare a given event to be punctual, we place restrictions on the kind of structures it can be an element of and, eventually, on the specific δ 's that are available.

REFERENCES

- Allen J. F., 1981, 'An Interval-Based Representation of Temporal Knowledge', *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, Vancouver: IJCAI [Morgan Kaufmann], Volume 1, pp. 221-26.
- 1984, 'Towards a General Theory of Action and Time', *Artificial Intelligence*, 23, 123-54.
- 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.
- Bach E., 1986, 'The Algebra of Events', *Linguistics and Philosophy* 9, 5-16.
- Bencivenga E., 1980, 'Free Semantics for Definite Descriptions', *Logique et Analyse* 23, 393-405.
- Bennett J., 1988, *Events and their Names*, Oxford: Clarendon Press.
- Benthem J. van, 1983, *The Logic of Time*, Dordrecht: Kluwer (2nd ed. 1991).
- Bochman A., 1990, 'Mereology as a Theory of Part-Whole', *Logique et Analyse*, 129-30, 75-101.
- Brentano F., 1933, *Kategorienlehre*, hrsg. von A. Kastil, Leipzig: Meiner (Eng. trans. by R. M. Chisholm and N. Guterman, *The Theory of Categories*, The Hague: Nijhoff, 1981).
- 1976, *Philosophische Untersuchungen zu Raum, Zeit und Kontinuum*, hrsg. von S. Körner und R. Chisholm, Hamburg: Meiner (Eng. trans. by B. Smith, *Philosophical Investigations on Space, Time and the Continuum*, London: Croom Helm, 1988).
- Casati R., Varzi A. C., 1994a, *Holes and Other Superficialities*, Cambridge, MA, and London: MIT Press/Bradford Books.
- 1994b, *Events. An Annotated Bibliography*, Milano: CUEM.
- Chisholm R. M., 1976, *Person and Object. A Metaphysical Study*, London: Allen & Unwin.
- 1978, 'Brentano's Conception of Substance and Accident', in R. M. Chisholm and R. Haller (eds.), *Die Philosophie Brentanos*, Amsterdam: Rodopi, pp. 197-210.
- 1984, 'Boundaries as Dependent Particulars', *Grazer Philosophische Studien* 10, 87-95.
- 1992/93, 'Spatial Continuity and the Theory of Part and Whole. A Brentano Study', *Brentano Studien* 4, 11-23.
- Clarke B. L., 1981, 'A Calculus of Individuals Based on "Connection"', *Notre Dame Journal of Formal Logic* 22, 204-18.
- 1985, 'Individuals and Points', *Notre Dame Journal of Formal Logic* 26, 61-75.
- Cohn A. G., Randell D. A., Cui Z., 1993, 'A Taxonomy of Logically Defined Qualitative Spatial Regions', in N. Guarino and R. Poli (eds.), *International Workshop on Formal Ontology in Conceptual Analysis and Knowledge Representation*, Padova: Ladseb-CNR, pp. 149-58.
- Cruse D. A., 1979, 'On the Transitivity of the Part-Whole Relation', *Journal of Linguistics* 15, 29-38.
- Davidson D., 1969, 'The Individuation of Events', in N. Rescher (ed.), *Essays in Honor of Carl G. Hempel*, Dordrecht: Reidel, pp. 216-34.
- Doepke F. C., 1982, 'Spatially Coinciding Objects', *Ratio* 24, 45-60.
- Eschenbach C., Heydrich W., 1993, 'Classical Mereology and Restricted Domains', in N. Guarino and R. Poli (eds.), *International Workshop on Formal Ontology in Conceptual Analysis and Knowledge Representation*, Padova: Ladseb-CNR, pp. 205-17.
- Fine K., 1981, 'Model Theory for Modal Logic, Part III: Existence and Predication', *Journal of Philosophical Logic* 10, 293-37.
- Franconi E., Giorgi A., Pianesi F., 1993, 'Tense and Aspect: A Mereological Approach', in *Proceedings of the 13th International Joint Conference on Artificial Intelligence*, Volume 2, Chambéry: IJCAI [Morgan Kaufmann], pp. 1222-1228.
- Hintikka K. J. J., 1959, 'Towards a Theory of Descriptions', *Analysis* 19, 79-85.
- Jackendoff R., 1991, 'Parts and Boundaries', *Cognition* 41, 9-45.
- Kamp H., 1979, 'Events, Instants, and Temporal Reference', in R. Bäuerle, U. Egli and A. von Stechow (eds.), *Semantics from Different Points of View*, Berlin/Heidelberg: Springer-Verlag, pp. 376-417.

- Kuratowski K., 1922, 'Sur l'opération A⁻¹ d'analyse situs', *Fundamenta Mathematica*, 3, 182-99.
- Lambert, K., 1962, 'Notes on E! III: A Theory of Descriptions', *Philosophical Studies* 13, 51-59.
- 1987, 'On the Philosophical Foundations of Free Description Theory', *History and Philosophy of Logic* 8, 57-66.
- Lejewski C., 1982, 'Ontology: What's Next?', in W. Leinfellner, E. Kraemer and J. Schank (eds.), *Language and Ontology. Proceedings of the 6th International Wittgenstein Symposium*, Vienna: Hölder-Pichler-Tempsky, pp. 173-85.
- Leonard H. S., Goodman N., 1940, 'The Calculus of Individuals and Its Uses', *Journal of Symbolic Logic* 5, 45-55.
- Leśniewski S., 1916, *Podstawy ogólnej teorii mnogości. I*, Moskow: Prace Polskiego Koła Naukowego w Moskwie, Sekcja matematyczno-przyrodnicza (Eng. trans. by D. I. Barnett, 'Foundations of the General Theory of Sets. I', in S. Leśniewski, *Collected Works*, ed. by S. J. Surma, J. Szrednicki, D. I. Barnett, F. V. Riskey, Dordrecht/Boston/London: Nijhoff, 1992, Vol. 1, pp. 129-73).
- Lewis D. K., 1991, *Parts of Classes*, Oxford: Basil Blackwell.
- Link G., 1987, 'Algebraic Semantics for Event Structures', in J. Groenendijk, M. Stokhof, F. Veltman (eds.), *Proceedings of the Sixth Amsterdam Colloquium*, Amsterdam: Institute for Logic, Language and Computation, pp. 243-62.
- Lowe V., 1953, 'Professor Goodman's Concept of an Individual', *Philosophical Review* 62, 117-26.
- Martin R. M., 1965, 'Of Time and the Null Individual', *The Journal of Philosophy* 62, 723-36.
- 1978, *Events, Reference, and Logical Form*, Washington: Catholic University of America Press.
- Moltmann, F., 1991, 'The Multidimensional Part Structure of Events', in A. L. Halpern (ed.), *Proceedings of the Ninth West Coast Conference on Formal Linguistics*, Stanford: Center for the Study of Language and Information, pp. 361-78.
- Newton-Smith W. H., 1980, *The Structure of Time*, London/Boston/Henley: Routledge & Kegan Paul.
- Pfeifer K., 1989, *Actions and Other Events: The Unifier-Multiplier Controversy*, New York: Peter Lang.
- Pianesi F., Varzi A. C., 1994, 'The Mereology of Event Structures', in P. Dekker and M. Stokhof (eds.), *Proceedings of the 9th Amsterdam Colloquium*, Amsterdam: Institute for Logic, Language and Computation, 1994, pp. 527-546.
- Randall D. A., 1991, *Analysing the Familiar: Reasoning about Space and Time in the Everyday World*, University of Warwick: PhD Thesis.
- Randall 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: Morgan Kaufmann, pp. 357-68.
- 1992, 'Exploiting Lattices in a Theory of Space and Time', *Computers and Mathematics with Applications* 23, 459-76.
- Randall D. A., Cui Z., Cohn A. G., 1992a, 'An Interval Logic of Space Based on "Connection"', in B. Neumann (ed.), *Proceedings of the 10th European Conference on Artificial Intelligence*, Chichester: John Wiley & Sons, pp. 394-98.
- 1992b, 'A Spatial Logic Based on Regions and Connections', in B. Nebel, C. Rich and W. Swartout (eds.), *Principles of Knowledge Representation and Reasoning. Proceedings of the Third International Conference*, Los Altos: Morgan Kaufmann, pp. 165-76.
- Rescher N., 1955, 'Axioms for the Part Relation', *Philosophical Studies* 6, 8-11.
- Russell B. A. W., 1905, 'On Denoting', *Mind* 14, 479-93.
- 1914, *Our Knowledge of the External World*, London: Allen and Unwin.
- 1936, 'On Order in Time', *Proceedings of the Cambridge Philosophical Society* 32, 216-28.
- Simons P. M., 1985, 'Coincidence of Things of a Kind', *Mind* 94, 70-75.
- 1986, 'Tractatus Mereologico-Philosophicus?', *Grazer Philosophische Studien* 28, 165-86.
- 1987, *Parts. A Study in Ontology*, Oxford: Clarendon.
- 1991a, 'Whitehead und die Mereologie', in M. Hampe and H. Maassen (eds.), *Die Gifford Lectures*

- und ihre Deutung. *Materialien zu Whiteheads "Prozess und Realität"*, vol. 2, Frankfurt: Suhrkamp, pp. 369-88.
- 1991b, 'Free Part-Whole Theory', in K. Lambert (ed.), *Philosophical Applications of Free Logic*, Oxford/New York: Oxford University Press, pp. 285-306.
- Smith B., 1992, 'Characteristica Universalis', in K. Mulligan (ed.), *Language, Truth and Ontology*, Dordrecht/Boston/London: Kluwer, pp. 50-81.
- 1994a, 'Ontology and the Logistic Analysis of Reality', in G. Häfziger and P. M. Simons (eds.), *Analytic Phenomenology*, Dordrecht/Boston/London: Kluwer, pp. 223-245.
- 1994b, 'The Formal Ontology of Space', to appear in L. Hahn (ed.), *Roderick Chisholm* (Library of Living Philosophers), La Salle: Open Court.
- Tarski A., 1935, 'Zur Grundlegung der Booleschen Algebra. I', *Fundamenta Mathematicae*, 24, 177-98 (Eng. trans. by J. E. Woodger, 'On the Foundation of Boolean Algebra', in A. Tarski, *Logics, Semantics, Metamathematics, Papers from 1923 to 1938*, Oxford: Clarendon Press, 1956, pp. 320-41).
- 1937, 'Appendix E', in J. E. Woodger, *The Axiomatic Method in Biology*, Cambridge: Cambridge University Press.
- Thomason S. K., 1989, 'Free Construction of Time from Events', *Journal of Philosophical Logic* 18, 43-67.
- Thomson J. J., 1977, *Acts and Other Events*, Ithaca: Cornell University Press.
- Tiles J. E., 1981, *Things That Happen*, Aberdeen: Aberdeen University Press.
- Van Fraassen B. C., Lambert K., 1967, 'On Free Description Theory', *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 13, 452-72.
- Varzi A. C., 1983, 'Free Semantics: Supervaluations at the Predicate Level', in P. Weingartner and J. Czermak (eds.), *Epistemology and Philosophy of Science. Proceedings of the 7th International Wittgenstein Symposium*, Vienna: Hölder-Pichler-Tempsky, 1983, pp. 35-38.
- 1993, 'Spatial Reasoning in a Holey World', in P. Torasso (ed.), *Advances in Artificial Intelligence. Proceedings of the 3rd Congress of the Italian Association for Artificial Intelligence*, Berlin/Heidelberg: Springer-Verlag, pp. 326-336.
- 1994, 'On the Boundary Between Mereology and Topology', in R. Casati, B. Smith and G. White (eds.), *Philosophy and the Cognitive Sciences. Proceedings of the 16th International Wittgenstein Symposium*, Vienna: Hölder-Pichler-Tempsky, pp. 419-438.
- Vendler Z., 1957, 'Verbs and Times', *The Philosophical Review* 66, 143-60.
- Vieu L., 1991, *Sémantique des relations spatiales et inférences spatio-temporelles: Une contribution à l'étude des structures formelles de l'espace en Langage Naturel*, Université Paul Sabatier de Toulouse: PhD Thesis.
- Walker A. G., 1947, 'Durées et instants', *Revue Scientifique* 85, 131-34.
- White G., 1993, 'Mereology, Combinatorics, and Categories', Preliminary Report for the SNF Project 11.31211.91, Schaan.
- Whitehead A. N., 1919, *An Enquiry Concerning the Principles of Human Knowledge*, Cambridge: Cambridge University Press.
- 1929, *Process and Reality. An Essay in Cosmology*, New York: Macmillan.
- Wiggins D., 1979, 'Mereological Essentialism: Asymmetrical Essential Dependence and the Nature of Continuants', *Grazer Philosophische Studien* 7, 297-315.
- Winston M., Chaffin R., Herrmann D., 1987, 'A Taxonomy of Part-Whole Relations', *Cognitive Science* 11, 417-44.



1. Introduction and Motivation

1.1. Mereology and set theory

It is interesting that Encyclopedia Britannica [1] classifies mereology under "Logic, Applied", and set theory under "Logic, Formal", although the motivations and origins of both theories are similar, and so is their subject matter. That is, both of them appeared as an answer to the Russell paradox in naive set theory (cf. Section 3.1, [20], or almost any introductory textbook in mathematical logic), and both of them address the problem of how an object can be decomposed or put together from smaller pieces. Roughly speaking, mereology is a theory of parts and wholes; i.e. it tries to axiomatize the properties of relation *a is-a-part-of b*. For example, one of its axioms postulates that *a is-a-part-of b*, e.g. between a wheel and a car, holds if and only if any item that is disjoint from *b* is also disjoint from *a* (disjointness is assumed to be a primitive relation).

Set theory tries to address the questions of when an element is a member of a class -- for instance, of all barbers that do not shave themselves, and how one can form new classes of objects. One of its axioms (comprehension) says that given a set one can form another set consisting of those elements of the set which have a certain property. For example, from the set of all people we can form a new set, consisting of those who are barbers. (Intuitively, a set is a "small" class whose existence will not result in a paradox).

The formal apparatus of set theory seems to be better developed than that of mereology, due chiefly to the influence of Goedel and Cohen. But in AI neither of the two theories has had much of an impact, although certainly more people are familiar with set theory, and some attempts have been made to make it useful ([18], [19] and [24]). For a discussion of set theory in the AI context see [16]; see also [11] for an introduction to set theory, and [14] and [7] for a coverage of more advanced topics; the latter should also give the reader an idea about the influence of Cohen and Goedel on mathematical logic.

In this paper we argue that both set theory and mereology have specific roles to play in the theory and practice of multimedia. They both provide us with a language to talk about multimedia objects. Notice that, say, first-order logic would be too weak for that, because it does not provide an axiomatic description of the structure of objects; and any useful theory for multimedia applications must contain some specific axioms that deal with structural properties, e.g. of images. As we observe further in the paper, the two natural candidate languages play complementary roles: one is used to index (set theory), the other to describe knowledge about indexes (mereology). This observation leads to a simple, yet realistic, formal theory of storage and retrieval of multimedia material, e.g. video clips; the theory combines set theory with mereology, and is motivated by practical needs of multimedia authoring. (Clearly, its existence also challenges the above classification of set theory and mereology by the Britannica).

Computational Mereology: A Study of Part-of Relations for Multimedia Indexing

Wlodek Zadrozny and Michelle Kim

IBM Research, T. J. Watson Research Center
P.O. Box 704, Yorktown Heights, NY 10598
wloz@watson.ibm.com

Abstract

How to index or retrieve multimedia objects is by no means obvious, because the computer can retrieve right multimedia material only if it reasons about its content. We show that it is possible to write formal specifications of this reasoning process using set theory and mereology.

We discuss the theoretical consequences of trying to use mereology and set theory for multimedia indexing and retrieval. We reexamine the roles of mereology and set theory in knowledge representation. We conclude that both commonsense set theories and mereologies should play the role of constraining databases of arbitrary multimedia objects, e.g. video clips. But although both should be viewed as database constraints, we argue that part-of hierarchies should be used to encode relatively permanent background knowledge, elsewhere named the *referential level*, while member-of hierarchies should describe arbitrary multimedia records. We also propose a language and a set of axioms, SetNM, for natural mereologies with sets. A multimedia indexing system can then be viewed as a particular SetNM theory.¹

¹ This paper has been accepted for publication in *Annals of Mathematics and Artificial Intelligence*.

With regard to the practice, a subset of the indexing system described here has been implemented as part of the IBM and New England Medical Center project for building a home health-care prototype system for children with leukemia (cf. [2], [15]). The home care system makes extensive use of multi-media objects such as video clips, audio clips, images, text to provide information and guidance. By using member-of relations extensively and lexical hierarchies, current system allows the users to find information without their giving exact descriptions of what they want to know. Adding part-of relations is being considered.

1.2. Describing objects and pictures

With emerging multimedia technologies, the problem of managing large amounts of data has become increasingly difficult. In addition to the traditional text-based data, a multimedia computer system has to store, and manipulate very large non-textual databases, such as image, video, and audio databases. Many approaches have been proposed in the past for representing and analyzing non-textual data; movement analysis using a notation for dance (labanotation, cf.e.g. [1]) is one such example. When applied to multimedia information fragments (e.g. motion video segments) such approaches may be useful capturing certain properties, but their utility is limited.

We need a more general means of querying these multimedia databases. In particular a database scheme should allow users to search through databases by partial match rather than by exact match. Furthermore, it should be possible to search such databases by describing the desired records (e.g. video clips) in words. Since vision systems are far away from being able to describe pictures, automated analysis of content of such videos is out of the question. What remains as the only reasonable possibility is to have a human index multimedia objects, and in the process of retrieval try to match stored descriptions with queries. Such a matching must take into account several factors which we describe in the following sections. But the first and obvious problem is that words cannot completely describe the content of a picture. For instance, anyone creating an index of a video clip containing the scene of a police car chasing a red Porsche would *not* include the information that you can see the wheels of the cars turning. Still, if you want to find videos containing wheels, you would like your system to know that cars have wheels, i.e. that a "wheel" is a "part-of" a "car".

1.3. The role of reasoning in indexing and retrieval

By the above arguments, reasoning is necessary for retrieval, because often it is necessary to change the original query, e.g. by adding a missing piece of information. As another example consider the fact that a "lake" might appear in a picture of a "sailboat", and getting to "lake" from "sailboat" via e.g. "boat" and "water" should be based on some form of reasoning. The same is true for storage. A picture depicting a "sailboat" does not have to be also indexed by "sail", "mast", and all possible details that can be easily found given a database of facts about how boats are built, for example based on a pictorial dictionary such as [9].

These examples suggest that we have to take into account two basic factors: the content of a picture, and background knowledge about the concepts used to describe the content. Regarding background knowledge, we can see two aspects of it. Namely, the general world knowledge about relations between objects, which allows us for instance connect "lake" and "sailboat", and the knowledge of the structure of everyday objects which brings the connection between "sailboat", "sail", and "mast". In this paper we restrict our attention to the descriptions of content and structure of objects. The issue of describing knowledge about relations between objects has been discussed at length in [26] and [27], and the formal model developed there (and briefly described in Section 3) is applicable in the present context and can accommodate the new theory we are about to propose.

Summarizing the above arguments, we can see that how to index or retrieve multimedia objects is by no means obvious, since the computer can retrieve right multimedia material only if it reasons about its content. What we want to show is that it is possible to write formal specifications of this reasoning process using set theory and mereology. As we shall see, this requires changing the shape of some axioms. But as a reward for this work, we get a well-motivated logical system in which both theories coexist and complement each other, a first system of this kind. Also, in the course of developing the axioms, in addition to illustrating the respective roles of both theories, we present a set of canonical and practical examples emphasizing the role of mereology.

2. Defining part-of relations

2.1. *How to describe a picture*

Although the theory of part-of relations is interesting for its own sake, we will not try to create one that would be useful for all possible purposes. At this point, we need a part-of hierarchy that could be used for multimedia indexing and retrieval. This does not mean, however, that we do not hope to learn something interesting and more general about such hierarchies. Let us first consider the problem of encoding background knowledge about the structure of everyday objects. As it turns out, trying to simply state information about parts of almost any objects introduces some interesting problems. For example, we may want to create a description of a first-aid kit, as in Figure 1.

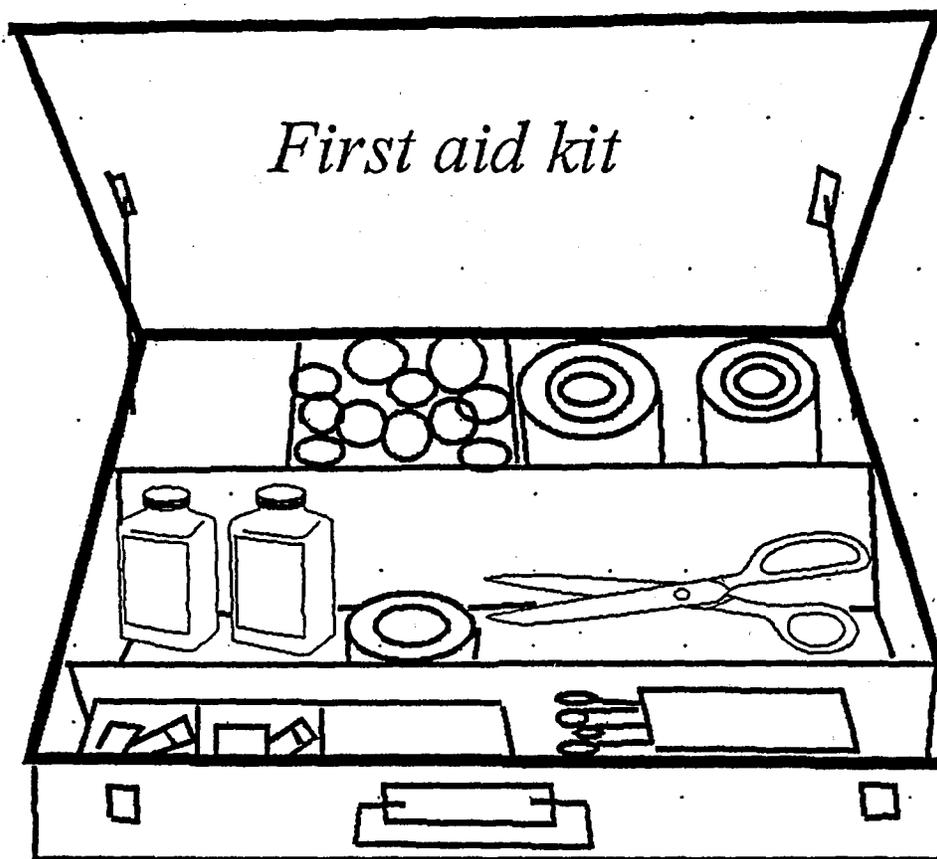


Figure 1. A first-aid kit

Now consider scissors, which are part of the the kit. The scissors in the kit could be coindexed with another picture of scissors which contains the information that "edge", "pivot", "blade", "shank" and "handle" are its parts (e.g. as on p. 487 of [9], which we can imagine having in a machine-usable form). Is now "blade" a part of "first-aid-kit"? In some circumstances "blade" might be considered a part of "first-aid-kit" (e.g. if a sharp edge is needed), but normally it would not be. We notice that we can have problems with the transitivity of the part-of relation. We solve it by assuming the transitivity of the part-of relation (hence "blade" becomes a part of "first-aid-kit"), although we can imagine systems where a part-of relation is not transitive.

Yet, there are two other, slightly less noticeable, but important and novel, aspects of the problem of encoding the background knowledge about the structure of objects. They have to do with the difficulty of describing pictures by a simple part-of relation. First, despite it not being the case on the picture, we would like to say that "syringe"

can be part-of "first-aid-kit" (for instance when someone is allergic to bee stings). Secondly, even though on the picture "cotton-applicator" is visible, in reality, it will most likely be hidden in a box, and hence invisible. We propose a solution to these problems which consists of parameterizing the part-of relation. To this end, we introduce two kinds features that are important for multimedia applications: (1) The first attribute is *iconic*(*Type*,*Value*) with *Type* ranging over audible, visible, tactile, X-ray, smell(?). The intended meaning of *iconic*(*visible*,1) is that we have a "visible" instance of the part-of relation, such as "wheel" being a visible part-of "car", or "cotton-applicator" being an invisible part of "first-aid-kit" (with the value of the parameter changed from 1 to 0). By the same token, *iconic*(*X-ray*,0) would mean that an object is not visible on an X-ray. (2) The second attribute, *necessary*(*_*), describes whether an object is a necessary (essential) part of another object. Most likely, these are all default attributes, and can be overridden in particular cases. Obviously this list of possible features is not considered exhaustive, and we would like to state as an open problem giving some theoretically satisfying characterization of possible parameters of the part-of relation.

Under this convention we obtain the following (partial) formalization of the structure of a first-aid kit and scissors:²

```

/* first-aid-kit */
pof( syringe      ,first-aid-kit  ,[iconic(v,1) , necessary(0) ]).

pof( scissors    ,first-aid-kit  ,[iconic(v,1) , necessary(1) ]).
pof( bandage     ,first-aid-kit  ,[iconic(v,1) , necessary(1) ]).
pof( aspirin     ,first-aid-kit  ,[iconic(v,1) , necessary(1) ]).
pof( rubbing-alcohol ,first-aid-kit  ,[iconic(v,1) , necessary(1) ]).
pof( cotton-swab ,first-aid-kit  ,[iconic(v,1) , necessary(1) ]).
pof( needle      ,first-aid-kit  ,[iconic(v,0) , necessary(1) ]).
pof( cotton-applicator,first-aid-kit  ,[iconic(v,0) , necessary(1) ]).

/* scissors; v stands for "visible" */
pof( edge        , scissors    ,[iconic(v,1) , necessary(1) ]).
pof( blade       , scissors    ,[iconic(v,1) , necessary(1) ]).
pof( pivot       , scissors    ,[iconic(v,1) , necessary(1) ]).
pof( handle      , scissors    ,[iconic(v,1) , necessary(1) ]).

```

Now we can easily define the regular part-of relation as a two-argument predicate, represented here as an infix operator *pof*

```

Arg1 pof Arg2 :-
    pof(Arg1,Arg2,[iconic(_,_), necessary(_)]).

```

and the "visible"-part-of as

```

Arg1 v-pof Arg2 :-
    pof(Arg1,Arg2,[iconic(v,1), necessary(_)]).

```

² The word *iconic* -- refers to a simplified version of Peirce's classification of signs: icon, index, symbol ([17] vol. VIII. pp.220-245).

where “_” is the “don’t care” symbol in our version of Prolog (cf. [8] for an introduction to Prolog). Note that there is no possibility of confusing *prof* with *prof(x,y,z)* -- the first one has two arguments, the other three.

2.2. From a part-of relation to a mereology

What we have seen is a database of facts defining the part-of relations. But is it mereology?

Mereology was introduced by Lesniewski (cf. [1] and [21]) as an answer to the Russell paradox. The simplest system of mereology is formulated in a first order language with the primitive \leq (is a part-of). Other relations such as $|$ (is disjoint from) can be defined using \leq , e.g. as $x|y \Leftrightarrow \neg(\exists z)[z \leq x \ \& \ z \leq y]$. The relation \leq should satisfy three postulates.

- Anti-symmetry

$$(\forall x)(\forall y)[x \leq y \ \& \ y \leq x \rightarrow y = x]$$

- Disjointedness of parts

$$(\forall x)(\forall y)[x \leq y \Leftrightarrow (\forall z)[z|x \rightarrow z|y]]$$

- An axiom-scheme for sum

$$(\exists x)[P(x)] \rightarrow (\exists y)(\forall z)[\neg(z|y) \Leftrightarrow \exists z'[P(z') \ \& \ \neg(z|z')]]$$

We used symbol \leq to encode the part-of relation. In general, we would like to separate the discussion of formal properties of part-of relations from the discussion of a particular computer implementation. We do it, by using the symbols \leq , $<$, $<_i$ etc. for the former, and the symbols *prof* and *prof_i* for the latter.

Notice that according to these axioms an object is a part of itself, i.e. $x \leq x$. Moreover, the part-of relation \leq is transitive. The axioms of sum says that for any property *P* we can form a sum of all parts that have this property. In particular it follows that one can make finite sums taking $P(x) = y_1 \leq x \vee \dots \vee y_k \leq x$.

Obviously, the facts about first-aid kits together with the definition of *prof* by a Prolog clause do not form mereology as as described by the axioms above or by the various axiomatic systems discussed in [21]; e.g. because the part-of relations *prof(...)* are not transitive or do not satisfy the axiom of sum. On the other hand, we have stated some facts about the *prof* relations, and implicitly described their model -- the minimal model of the above Prolog program. We could then rephrase the question as: is this a model of mereology? Can we make a mereology out of it? The answer depends on our expectations.

Models of mereologies are partial orderings. For a full classical mereology a model is a complete Boolean algebra without 0; hence, in the finite case, it has $2^n - 1$ elements (e.g. [7] and [21], p.25). Clearly, given *n* “atomic” parts (i.e. parts without proper subparts) in any realistic scenario, we would not like to have to explicitly represent all

$2^n - 1$ elements. At this point we should reexamine the *pof* relations of Section 2, with the view that we could be able to formulate a system of *natural mereology* in which it would be possible to adequately specify the behavior of a multimedia retrieval system. Thus, if we look at those examples, we can see that we should take as a model a collection of finite partial orders (which itself is a partial order). That way we get e.g. the transitivity. For convenience we can assume the existence of the top element 1, intuitively corresponding to unspecified part containing any collection of objects as sub-parts (i.e. 1 is the answer if there is no real object that satisfies a query).

The really important question is what to do with the sums. Classical mereology insists on the existence of arbitrary sums of parts: any sum of parts of an entity is a part of it. Clearly, this cannot be taken literally -- imagine representing the set of all parts of a jumbo-jet; we would be facing the combinatorial explosion of parts. On the other hand, there are good arguments for allowing (arbitrary?) sums, e.g. in the context of retrieval. To see it, consider the following query: "find an object that has [a,b,c] as a subpart". What if a, b, c, and d are parts of A that are listed in the database, but there is no part consisting of a, b and c? Then, to retrieve the object A, the system would need a rule sanctioning that kind of deduction, i.e. a version of the axiom of sums.

We return to the question of what does the database of *pof* relations represent? We believe it should be viewed as a model of a mereology. There is no reason to insist on one mereology; for instance for a very small database the axiom of sum can be valid, because the combinatorial explosion of parts can be contained. But is there a relationship between the axioms of a mereology and this model? Yes. The axioms should be viewed as database constraints. Hence mereology can be viewed as a theory of constraints for databases of the kind we have described above. Then, if the axioms of mereology are viewed that way, they should be reexamined in this new context, which we will do in Section 4.

In [24] we have shown that it is possible to formulate various commonsense set theories. We argued that the property of being a set could be relative to a theory; and we introduced the notion of "actually constructed" sets. This was done in order to show that certain common-sense intuitions about cardinalities and well-orderings can be expressed in such a framework. Namely, we constructed models in which some sets were well-ordered without having a cardinality, and vice versa -- models in which there were sets with known number of elements, but which could not be well ordered. This paper can be viewed as complementing the arguments presented there. Namely, if axioms of set theory shape our view of elements in a member-of relations, the same is true about mereology defining possibilities and problems in our dealings with the part-of relations.

3. The role of set theory

In this paper we propose that mereology be used to describe (and constrain descriptions of) non-arbitrary collections of objects, while set-theory should be used for arbitrary collections of elements. In this section we first introduce the concepts of ax-

iomatic set theory and then sketch how they can be applied to describe the content of multimedia objects.

3.1. Axiomatic set theory

Set theory describes sets, i.e. collections of abstract objects. But it does not allow all possible collections of objects to exist. It describes sets by dividing abstract collections into "consistent" and "inconsistent" ones. Thus $\{X: \neg (X \in X)\}$ (the Russell class of X's which are not members of themselves) is such an inconsistent collection, and its existence is forbidden by the dominant set theories. The discovery of inconsistent, formally defined collections of objects led to the development of axiomatic set theories and Lesniewski's mereology.

The theory of Zermelo and Fraenkel, denoted by ZF or ZFC (C standing for the axiom of choice) is the most important axiomatic set theory. In this theory, the existence of a set, such as the set of all natural numbers, is derived from axioms. These axioms are formulated in the standard first order language containing additionally the relation of membership \in . For instance, the axiom of infinity says:

$$(\exists x)[\emptyset \in x \ \& \ (\forall y \in x)[y \cup \{y\} \in x]]$$

By rejecting some of the axioms of ZFC, in particular the axiom of infinity, and slightly reformulating the other axioms, we obtain a weaker set theory -- the theory of Kripke and Platek or KPU. Barwise and Perry [6] mention KPU as a possible basis for their situation semantics (cf. [10] for an introduction). In KPU, individuals without elements, called *points*, *atoms*, or *urelements* (primitive elements) are allowed to exist (while in ZFC the only set without members is the null set \emptyset). The KPU set theory is expressed with the language of set theory containing the additional predicate *Atom*(.). It differs from ZFC in that it does not have the axioms of powerset, substitution¹ and infinity. (But Σ - replacement, which is a weak version of the axiom of substitution is provable in KPU). Moreover the axiom of subsets (separation) applies only to formulas with restricted quantifiers -- Δ_0 - formulas. (Δ_0 formulas have all quantifiers restricted; for instance, $\forall s \exists t P(s,t)$ is not a Δ_0 formula, while $(\forall s \in a)(\exists t \in b)P(s,t)$ is Δ_0). Also the axioms of choice and regularity have slightly different forms (for details, cf. [4])

Axioms of KPU (with Powerset)

A1. Extensionality:

$$(a) \text{Atom}(a) \rightarrow \neg (s \in a)$$

$$(b) \neg \text{Atom}(a) \ \& \ \neg \text{Atom}(b) \rightarrow [\forall s (s \in a \Leftrightarrow s \in b) \rightarrow a = b]$$

$$\text{A2. Pairing: } \forall s \forall t \exists u [u = \{ s, t \}]$$

$$\text{A3. Empty set: } \exists s [\neg \text{Atom}(s) \ \& \ (\forall t) \neg (t \in s)]$$

$$\text{A4. Union(binary): } \forall s \forall t \exists u [u = (s \cup t)]$$

A5. Intersection (binary): $\forall s \forall t \exists u [u = (s \cap t)]$

A6. Difference (binary): $\forall s \forall t \exists u [u = (s - t)]$

A7. Powerset: $\forall s \exists t [t = P(s)]$

A8. Union: $\forall s \exists t \forall u [u \in t \Leftrightarrow \exists u' [u' \in s \ \& \ u \in u']]$

Additionally, the set theory should satisfy the following schemes:

A9. Regularity (\in - induction): $\forall s [(\forall t \in s) P(t) \rightarrow P(s)] \rightarrow \forall s P(s)$

where P is any formula.

A10. Δ_0 -Separation (Subsets): $\forall s \exists t \forall u [u \in t \Leftrightarrow u \in s \ \& \ P(s)]$

where P is a Δ_0 - formula ¹.

A11. Δ_0 - Collection: $\forall s \in a \exists t P(s,t) \rightarrow \exists b (\forall s \in a) (\exists t \in b) P(s,t)$

where P is a Δ_0 - formula.

In addition, the obvious axioms defining the operators $P, \{ \dots \}, \cup, \cap$ and $-$ are needed.

The axioms of the above theory are not independent, for instance A8 and A2 imply A4. But they form three groups of different status: We have no doubt that any set theory should satisfy the axioms A1-A6 (group 1). We are not sure that powersets and unions (A7 and A8 -- group 2) should always exist. (But notice that the powerset of a finite set can always be constructed by a finite composition of pairing and binary union). One can also argue how strong/weak should be the schemes of regularity, separation, choice and replacement/substitution (group 3). The axiom of regularity forbids \in -chains: $s \in s, s \in s_1 \in \dots \in s_n \in s$.

3.2. Set theory for multimedia indexing

How do we imagine the division of labor between set theory and mereology in the context of multimedia applications? We would use part-of relations to describe a visual dictionary, since objects depicted there have non-arbitrary relations; and use the member-of relation to describe a video clip or a picture. Thus, indexes of a multimedia data would be sets encoding their attributes. For example, a picture of the mess on my desk could have been stored and indexed by the following facts (*mof* -- stands for "member-of").

mof(cup, mess-on-the-desk).
 mof(book, mess-on-the-desk).
 mof(papers, mess-on-the-desk).

...

I could retrieve it by formulating a query about, say, a cup and a book. But in a more interesting scenario I would like to retrieve it by formulating a query about a handle, and to use the facts that "cup" member-of "mess-on-the-desk" and "handle" part-of "cup" during the processing of the query. Therefore, for retrieval, the query would have to make reference to background knowledge about everyday objects; a formal model of such background knowledge, i.e. situation independent collections of facts, was presented in [26] under the name *referential level*. Under the related scheme that we propose here, in the simplest case, the background knowledge would consist of a database of $pof(x,y,z)$ facts. But when more than one decomposition of objects into parts is needed it would consist of a database of *collections* of $pof(x,y,z)$ facts. This would happen for instance when it is necessary to reason about assembling a bicycle from a kit, where its part-of decomposition for the purpose of assembly does not agree with the typical description of bicycle parts, and both are needed for a successful accomplishment of the task. Reasoning with background knowledge as consisting of sets of theories is discussed in detail in [22], [23], [26] and [27], and we refer the reader there.

Thus the basic idea behind our proposal should be clear now: a commonsense set theory would constrain the object level, while mereology would do the same for the background knowledge about what can be a part of what. But, as we did for mereology, we can ask the obvious question of how much set theory is actually needed to ensure the properties of sets we care about. For example, maybe we only need one level of braces $\{, \}$, which is sufficient to encode the *mess-on-the-desk* as $\{cup, book, papers, \dots\}$.

While we cannot exclude a possibility of using different sets of axioms describing properties of the membership relation, there are convincing arguments that we should allow arbitrary sets of objects, and hence the deeper nesting of $\{\dots\}$. These deeper nesting are useful: with sets we can encode predicate-argument relations. For instance, a video in which a dog bites John can be indexed by $\{bite\{dog\{john\}\}\}$, while $\{bite\{john\{dog\}\}\}$ is the sequel in which John bit off the dog's ear. A similar encoding can be done for modification, to encode *John bit off the ear of the dog who bit him*. Then, since there is a priori no limit on the depth of modification, we should not put an a priori limit on the depth of nesting.

4. The formal system SetNM: Natural mereology with finite set theory

The language of SetNM is a first order language containing the primitive relations $<$, and \in ; two types: PART and SET; constants 1 and \emptyset ; an infinite number of variables of type PART written as $x_{(i)}$, $y_{(i)}$, $z_{(i)}$; an infinite number of variables of type SET written as $s_{(i)}$, $t_{(i)}$, $u_{(i)}$; an infinite number of predicate names; and equality $=$. In order to simplify the notation we will refer to to all $<$, relations by $<$.

Axioms for Natural Mereology with Sets

The relation $<$ is defined on PART \times PART; the relation \in on SET \times SET. The axioms for Natural Mereology with Sets consists of three groups:

1. The axioms of KPU as described in Section 3.
2. Axioms defining properties of the part-of relations, as described below.
3. Axioms relating parts and sets.

We begin with the last group.

- Any part is a set (but not necessarily vice versa)

$$x < y \rightarrow \exists s[s = x]$$

- The set of atoms is identified with the set of parts, i.e. $U = \text{PART}$:

$$(\forall s)[\text{Atom}(s) \rightarrow (\exists x)[s = x]]$$

$$(\forall x)[\text{Atom}(x)]$$

Now we list the axioms defining the basic properties of $<$.

- $<$ is a strict partial order on $\text{PART} \times \text{PART}$

$$(x < y) \rightarrow \neg(y < x)$$

$$(x < y \ \& \ y < z) \rightarrow (x < z)$$

- 1 is the unique top element of $<$

$$x < 1$$

$$\neg(1 < x),$$

$$(\forall x)[(\forall y \neq 1)(y < x) \rightarrow x = 1]$$

- Weak Supplementation Principle (p.28 of [21])

$$(\forall x)(\forall y)[y < x \rightarrow (\exists z)[z < x \ \& \ z|y]]$$

- Existence of the least upper bound

$$(\forall x_1)(\forall x_2)(\exists y)(\forall z)[y < z \Leftrightarrow x_1 < z \ \& \ x_2 < z]$$

We end this section with a discussion of the axioms of SetNM. It is easy to see that SetNM is a modification of set theory. It is obtained by postulating that in addition to the membership relation a part-of relation with certain properties is defined on some sets. The question on which sets it should be defined can be a matter of discussion. Thus, although in the axioms that relate sets and parts the first axiom is a consequence of the second one, we decided to list both of them, since we can imagine for instance a mereology in which atoms are identified with parts which do not have proper sub-parts.

We do not have extensionality for PART's, although we have it for sets. In the latter case, the motivation being that indexes of two different multimedia objects must be distinct. Extensionality for PART is not assumed, because we do not want to exclude different names for differently arranged collections of the same parts, e.g. a bicycle and a bicycle kit. The weak supplementation principle eliminates models of mereology where all parts overlap each other. It says that if an individual has a proper part, it

must have another part disjoint from the first one. Notice that in set theory a set and its singleton are always distinct, i.e. $s \neq \{s\}$.³

Finally, we do not assume that sums exist. That is, not all collections of parts form an individual in PART. But we do assume existence of the least upper bounds. That is, for any finite collection of parts, there is an individual in PART of which all members of the collection are a subpart. (The finite case is an immediate consequence of the the existence of the least upper bound of two individuals). We believe that the existence of least upper bounds is sufficient for dealing with decompositions of objects into parts. Notice that if a sum exists, then so does a least upper bound, but not vice versa. For example, in a simple model consisting of of five objects, the least upper bound of any two elements of $A = [a, b, c, d]$ might be A , but no sum of any two elements exists (cf. the discussion in Section 2.2). If a stronger axiom of sums is needed, in [21] we can find a discussion of its many variants.

5. SetNM theories as formalizations of multimedia indexing schemes

In order to describe the formal model of indexing we are proposing we have to construct a model of SetNM. This is easily done. We begin with the mereologies. We assume we are given a set $R(P)$ of background knowledge facts about part-of relations, perhaps organized by topics, and an active database $DB(P)$ of relevant part-of relations which is a subset of a set $R(P)$. We assume it is a collection of ground atoms. In the simplest case $DB(P) = R(P)$ is a set of atoms of type $x \prec_i y$, e.g. $WHEEL \prec_i CAR$ ("a wheel is a visible part of a car"). More generally, $DB(P)$ can be computed from $R(P)$ by the methods of [26] and [27]. At this point, the precise method of obtaining $DB(P)$ is not important. However, it is important that the part-of relations described in $DB(P)$ satisfy the axioms of mereology described in Section 4. The set PART(P) consisting of names of objects that appear in $DB(P)$ will be our set of atoms.

Since we are interested in models of indexing and retrieval, we must also be given a set of indexes I , and obviously we must assume that each index is a set made of atoms contained in PART(P). Informally, the set I is a set of descriptions of multimedia data. In contrast with PART(P), which satisfies the axioms of mereology, the set I does not have to satisfy the axioms of set theory (because there is no reason why a union or an intersection of two indexes should be an index). Those axioms have to be satisfied by the set of potential indexes, M , which includes I . M is defined as the smallest model of KPU that contains the set PART(P) of constants appearing in $DB(P)$ as the set of its atoms. Then, by the definition, the structure

$$\langle M, \in, \prec_i \rangle$$

is a model of SetNM. (Notice that we have to trivially extend the domains of the part-of relations to all the urelements of M).

³ A theory of non-well founded sets where $s = \{s\}$ is possible is described in [3], and its applications to semantics of natural language are discussed in [5] and [25].

A query Q is an open formula in the language $L(\in, <_i)$. An answer is a substitution θ of indexes (members of I) for all free variables of Q such that

$$\langle M, I, \in, <_i \rangle \models Q[\theta]$$

The queries are fairly expressive. E.g. one can ask for a set of indexes such that they contain some unspecified common entity. But, in applications we played with, one typically asks for an index with certain properties; this corresponds to restricting queries to formulas with just one free variable. To see how it works, let us briefly discuss this model of indexing and retrieval.

Examples: Consider a short video clip about preparation for a bicycle trip. It consists of just two cuts, one showing a bicycle and a tent, and the other the content of a first-aid kit. Because we have previously stored the information about parts of bicycles, tents, and first-aid kits, the clip can simply be indexed as $\{\{bicycle, tent\}, fa - kit\}$.

What are the queries that would produce this index as an answer? The simplest query is the formula $fa - kit \in Y$, which reads "find an index which contains a *fa - kit*". A slightly more interesting one would be $aspirin < X \ \& \ X \in Y$, which can be translated as "find an index which contains aspirin as a part of something"; here this something or X is of course the *fa - kit*. Now, notice that the query $aspirin \in Y$ would not produce our set as an answer; neither would we get it using the formula $bicycle \in Y$, because $bicycle \in \{bicycle, tent\}$ and $bicycle \notin \{\{bicycle, tent\}, fa - kit\}$.

However, our query language is expressive enough to encode more forgiving relations between objects and indexes. We can define the relation $s \in^* t$ to hold if s is a member of the transitive closure of t , i.e. if either if s is a member of t , or s is a member of a member of t , or s is a member of a member of a member of t , and so on. The existence of the transitive closure of any set is provable from the axioms. (By the axiom of regularity, there can be only finitely many layers of membership for any set, so the relation can be easily expressed in an always terminating computer program). Then, we can ask for "an index with a bicycle" by writing $bicycle \in^* Y$.

Notice the influence of the structure of indexes on our ability to control the precision and recall. By saying that $\{a, \{b\}\}$ is not $\{a, b\}$ we increase the precision of queries. But by using the membership in the transitive closure of indexes, which roughly(!) corresponds to identifying all the sets $\{a, \{b\}\}$, $\{\{a\}, \{b\}\}$, $\{\{\{a\}\}, \{b\}\}$, and the like, with the set of their atoms $\{a, b\}$, we increase the recall.

6. Some open problems

There are a number of interesting problems that arise when we try to apply mereology to the task of indexing and retrieving of multimedia objects. Some of them, like having to deal with necessary part-of's are not new, just acquire a slightly different meaning; others, like dealing with visible part-of's are new. There are problems we have not discussed so far, but which can be very relevant, since they have to be solved in prac-

tice. They can be grouped according to categories, and we give an example in each category. The categories are named after the source of the problems.

Natural language ambiguity

In our example, the first-aid kit contains a needle. If it contains a syringe, we have

```
needle pof first-aid-kit
needle pof syringe
```

but clearly we are talking about different kinds of needles. Hence the problem of ambiguity (cf. e.g. [12]). The natural question arises: since we cannot assume that names refer to the same types of entities (although we can assume that they are related by a kind-of relation), what are the consequences of this fact for the logic of the part-of relations?

The content of the database

The next question is what do the tokens in the database represent. For instance, we have

```
pof( edge , scissors, [iconic(v,1), necessary(1)])
```

but there are two edges on the picture. There may be various ways of describing what is really there, e.g. using multisets or rather multiparts to represent objects with many parts of the same kind. But notice that for the purpose of indexing and retrieval the information about there being two edges might not be necessary.⁴ That is, we can read

```
edge pof scissors
```

as "for any object named 'scissors' there is an object named 'edge' which is a part of it". This is also interesting if we compare it with

```
edge mof scissors
```

(defined as $mof(edge,scissors)$) under the assumption that sets represent arbitrary collections of objects. Here, the expression reads "there is an object named 'scissors' and an object named 'edge' which is a part of it". Notice the difference in the quantifiers: $\forall\exists$ vs. $\exists\exists$.

What does it mean that a database satisfies some constraints?

We have said that the axioms of mereology can be viewed as constraints on multimedia databases. But a database can satisfy such constraints in more than one way. For instance, the transitivity of the *pof* relation can be either assured by adding enough statements of the form

```
x pof y
```

so that it is implicitly true, or by adding a rule such as

```
X pof Z :- X pof Y , Y pof Z .
```

⁴ If it is, one might add cardinality as another parameter of the *pof(...)* relation.

On the other hand, for a Prolog implementation the axiom of extensionality either for SETs or for PARTs (stating that objects with identical members/parts are equal) cannot be that easily formulated as a rule of a program, and hence would probably have to be regarded as a higher order constraint. However, if our programming language is a version of CLP [13], such a rule could be formulated. Notice also, we do not have to postulate this axiom for SETs if we want to allow the same description to be used for different multimedia objects. Intuitively, in that case, a description represents partial information about an object.

Other problems

Within the context we described there are certainly other interesting problems. For instance, with videos we have to confront the problems mentioned on pp.173-251 of [21] of creating a mereology that contains a time variable. And, as usual, we can ask about computational aspects of various formalizations, e.g. worry about complexity issues.

7. Conclusions

We have presented a system of computational mereology and set theory. We call it "computational" because it addresses some issues of computation, such as restricted forms of the axiom of sums, different kinds of part-of relation (visible, x-ray, ...), and because some of the ideas presented here have actually been implemented in a system. The other reason for calling it "computational" is to contrast it with the classical mereologies, which are not computational, e.g. because the axiom of sums causes the combinatorial explosion of parts. More important, they are not computational, because they were presented as axiomatic theories, and not knowledge representation schemes. In contrast, in our theory, mereology is used to specify how background knowledge about objects and their parts should be used in reasoning about multimedia indexes, and the emphasis is put precisely on knowledge representation issues.

Since the work we are presenting in this paper is the first of the kind, all conclusions must be regarded as tentative. For instance, with respect to the axiom of sums, if we view axioms of mereology as constraints, then perhaps this axiom should only apply to databases containing descriptions of objects with not too many elements. In practice, there seem to be a natural way for describing those subclasses of multimedia databases for which given axioms should hold. But also notice that there may be a less orthodox approach which could be universally applicable, even though it might partially disagree with the view presented in this paper. Namely, given a database, we can make a particular set of axioms true (for the purpose of search) if according to information retrieval criteria such an augmented database can be successfully queried. That is, the inclusion, e.g. of the axioms of sum should depend on its contribution to recall and/or precision. Ditto for commonsense set theory. In such a case, more work is required, e.g. to empirically validate the influence of various axioms on recall and precision.

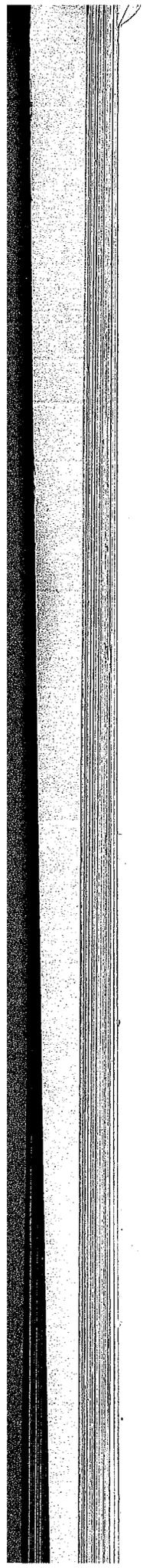
An interactive multimedia system is driven by human users, who would welcome in it some built-in intelligence. We suggested that a consequence of this fact might be the need to use set theories and mereologies for indexing and retrieval, and allowing users to search through databases by partial rather than exact match. The parameterized *pof* relations and *mof* relations can be used to facilitate such a search by similarity. Namely, in our experience, it is possible to use both relations for for multimedia indexing, and to view mereologies and set theories as constraints on databases of indexes. The contribution of the paper lies in identifying the formal apparatus that can be used for the purpose of indexing, and, hopefully, in opening a new avenue of research. The symmetry between mereology (-ies) and set theory (-ies) -- namely, that the latter should be used to encode arbitrary multimedia objects, while the former the relatively permanent background knowledge -- is an intuitively valid justification for the approach. From the formal point of view, having two structural relations on individuals should lead to as much interesting mathematics as playing with the two principal arithmetic operations, plus (+) and times (\times).

References

- [1] *Encyclopedia Britannica*.
Encyclopedia Britannica, Inc., Chicago, IL, 1984.
- [2] Hospitals Attack a Crippler: Paper.
Business Week, (Feb. 21):104-106, 1994.
- [3] P. Aczel.
Non-well-founded Sets.
CSLI Lecture Notes No.14, Center for the Study of Language and Information, Stanford, CA, 1988.
- [4] J. Barwise.
Admissible Sets and Structures.
Springer-Verlag, New York, NY, 1975.
- [5] J. Barwise and J. Etchemendy.
The Liar: An Essay on Truth and Circularity.
Oxford University Press, New York, NY, 1987.
- [6] J. Barwise and J. Perry.
Situations and Attitudes.
MIT Press, Cambridge, MA, 1983.
- [7] J. Barwise(Ed.).
Handbook of Mathematical Logic.
North-Holland, New York, NY, 1977.

- [8] W. Clocksin and C. Mellish.
Programming in Prolog.
Springer-Verlag, New York, NY, 1981.
- [9] J. Corbeil.
The Stoddart Visual Dictionary.
Stoddart, Toronto, Canada, 1986.
- [10] K. Devlin.
Logic and Information.
Cambridge University Press, Cambridge, UK, 1991.
- [11] P. Halmos.
Naive Set Theory.
Van Nostrand, New York, NY, 1960.
- [12] G. Hirst.
Semantic interpretation and the resolution of ambiguity.
Cambridge University Press, Cambridge, Great Britain, 1987.
- [13] J. Jaffar and M. Maher.
Constraint Logic Programming: a Survey.
Journal of Logic Programming, (10), 1994.
- [14] T. J. Jech.
Set Theory.
Academic Press, New York, NY, 1978.
- [15] M. Kim.
Guardian: A Home-Care Consulting System for Children with Leukemia.
IBM T.J. Watson Research Center, Yorktown Heights, New York. 1994.
RC XXXXX.
- [16] M. Pakkan and V. Akman.
Issues in commonsense set theory.
Dept. Comp. Engin. and Inform. Sci., Bilkent University, Ankara, Turkey. 1992.
- [17] C. Peirce.
Collected Papers of Charles Sanders Peirce.
Belknap Press of Harvard University Press, Cambridge, MA, 1984.
- [18] D. Perlis.
Circumscribing with sets.
Artificial Intelligence, 31:201-211, 1987.
- [19] D. Perlis.
Commonsense set theory.
in P. Maes and D. Nardi, editor, *Meta-Architecture and Reflection*, pages 87-98, North-Holland, Amsterdam, 1987.

- [20] J. Shoenfield.
Mathematical Logic.
Addison Wesley, New York, New York, 1967.
- [21] P. Simons.
Parts: A Study in Ontology.
Clarendon Press, Oxford, UK, 1987.
- [22] W. Zadrozny.
Intended Models, Circumscription and Commonsense Reasoning.
Proc. Intern. Joint Conf. on Artificial Intelligence, IJCAI-87, pages 909-916, Milan, Italy, 1987.
- [23] W. Zadrozny.
A Theory of Default Reasoning.
Proc. Sixth National Conf. on Artificial Intelligence, AAAI-87, pages 385-390, Seattle, Washington, 1987.
- [24] W. Zadrozny.
Cardinalities and well orderings in a commonsense set theory.
in Ronald J. Brachman, Hector J. Levesque, and Raymond Reiter, editor, *Proc. First Intern. Conf. on Principles of Knowledge Representation and Reasoning*, pages 486-497, Morgan Kaufmann, San Mateo, CA, 1989.
- [25] W. Zadrozny.
On compositional semantics.
Proc. Coling'92, pages 260-266, 1992.
Nantes, France.
- [26] W. Zadrozny.
Reasoning with background knowledge -- a three-level theory.
Computational Intelligence, 10(2):150-184, 1994.
- [27] W. Zadrozny and K. Jensen.
Semantics of Paragraphs.
Computational Linguistics, 17(2):171-210, 1991.



Information about the authors

Anthony Cohn is a Reader in Automated Reasoning in the Division of AI, School of Computer Studies at the University of Leeds. He leads a research group working on Automated Reasoning with a particular focus on qualitative spatial reasoning. He holds a number of grants in the area, and has published widely. He has been Chairman of the UK AI Society AISB and also of the European Coordination Committee on AI; he was Programme Chair of the European AI Conference ECAI94.

Carola Eschenbach studied Computer Science and Mathematics at the University of Hamburg. In her thesis she worked on natural language expressions of measurement and counting. Currently, she is working in the section Wissens- und Sprachverarbeitung (knowledge and language processing) of the Computer Science department in Hamburg. She is member of the graduate program in Cognitive Science at Hamburg.

Nick Gotts has a background in psychology and artificial intelligence, and has worked in varied fields including wayfinding, medical expert systems and the simulation of spiders' web-building skills. He is currently at the School of Computer Studies, University of Leeds, where his work concerns formal systems for qualitative spatial representation and reasoning.

Christopher Habel is Professor of Computer Science (and Adjunct Professor of Linguistics) at the University of Hamburg and Director of the Doctoral Program in Cognitive Science. His research focuses on the representation of knowledge and semantics of natural language.

Dr. Michelle Y. Kim is a research staff member at the IBM Thomas J. Watson Research Center. Her research interests include artificial intelligence and multi-media computing.

Almerindo E. Ojeda is Associate Professor of Linguistics at the University of California at Davis. His research interests lie with natural language syntax (discontinuous constituency) and semantics (the semantics of individuation). He is a coeditor of *Syntax and Semantics 20: Discontinuous Constituency* and the author of *Linguistic Individuals*.

Jean Petitot, born in 1944, is a mathematician and a philosopher interested in dynamical modeling in cognitive sciences and naturalized phenomenology. He is currently "Directeur d'Etude" at the "École des Hautes Etudes en Sciences Sociales" in Paris where he manages the "Seminar of Philosophy of Mathematics" and the "Center for Cognitive and Semiotic Modeling". He is also a member of the CREA (Applied Epistemology Research Center) at the École Polytechnique. He is the author of "Morphogénèse du Sens" (Paris, PUF, 1985), "Les Catastrophes de la parole: de Roman Jakobson a René Thom" (Paris, Maloine, 1985), "La philosophie

transcendantale et le problème de l'objectivité" (Paris, Osiris, 1992), "Physique du Sens" (Paris, CNRS, 1992), and more than 150 papers.

Fabio Pianesi (Folign, Italy, 1956) studied psychology and computer science at the Universities of Padova and Rome. At the moment he is a researcher at the Institute for Scientific and Technological Research (IRST), in Trento, Italy. His main interests are in the area of computational and formal linguistics (syntax and semantics). Recent publications include papers at IJCAI-93 (with E. Franconi and A. Giorgi), at ECAI-94 (with A. Varzi), at the 9th Amsterdam Colloquium (with A. Varzi) and at various conferences on formal and generative linguistics.

Barry Smith is Professor of Philosophy and Member of the Center for Cognitive Science at the State University of New York at Buffalo. He is the author of *Austrian Philosophy* (Chicago 1994) and of some 150 scientific papers in formal ontology and related fields. He is the editor of *The Monist: An International Quarterly Journal of General Philosophical Inquiry*, and of *Parts and Moments: Studies in Logic and Formal Ontology* (Munich 1982). He is co-editor of *Handbook of Metaphysics and Ontology* (2 vols., Munich 1991), *Philosophy and the Cognitive Sciences* (Vienna 1994) and *The Cambridge Companion to Husserl* (Cambridge 1995).

Achille C. Varzi (Galliate, Italy, 1958) studied philosophy at the university of Toronto, Ont., and is presently with the Institute for Scientific and Technologic Research (IRST) in Trento, Italy. Recent publications include *Holes and Other Superficialities* (with R. Casati, MIT Press, 1994), an annotated bibliography on *Events* (also with R. Casati, CUEM, 1994) and papers on philosophical logic, formal semantics, formal ontology, and commonsense reasoning.

Graham White works in logic and the philosophy of computer science. He recently completed a research project on the philosophy of artificial intelligence for the Swiss National Research Foundation.

Wlodek Zadrozny manages the Foundations of AI group at IBM T.J. Watson Research Center. The group is currently building multimodal dialog systems. His principal interests are semantics of natural language and knowledge representation. One of his recent results is a proof that compositionality is vacuous, i.e. that the postulate of compositionality does not constrain theories of language. His work on dialog deals with semantic complexity and the merging of semantic, syntactic and pragmatic information in a grammar of constructions.

Wojciech Żelaniec, born 1959 in Gdańsk, Poland, studied mathematics and philosophy in Gdańsk, Lublin and Germany. Since 1990 has occupied various teaching, research and editorial positions in Poland, Liechtenstein and the USA. Main areas of interest: ontology and epistemology, mathematical methods in philosophy, social philosophy.