Carving Up Space: Existential Axioms for a Formal Theory of Spatial Regions*

Brandon Bennett†
Division of Artificial Intelligence
School of Computer Studies
University of Leeds, Leeds LS2 9JT, England
brandon@scs.leeds.ac.uk

July 11, 1995

Abstract

In this paper I investigate how one might arrive at a set of existential axioms that would specify a *complete* 1st-order theory of spatial regions. In such a theory all formulae would be true or false (given that constants are taken as existentially quantified with widest scope). The RCC spatial logic based on the relation of *connectedness*, C, will be taken as the starting point. This theory is not complete, since it does not guarantee the realisation of all possible configurations of regions. By adding extra existential axioms one would aim to ensure that the existence of any possible configuration is entailed by the theory. Thus the ontology of the theory would be strengthened from being relative to contingent facts (asserted in conjunction with the theory) to an absolute ontology characterising the domain of spatial regions as having a specific structure.

This is only an exploratory paper. I do not attempt to give an axiom set but only to shed some light on fundamental principles and problems relating to existential axioms in a spatial theory. I suggest a method of eliciting such axioms by considering the operations involved in constructing diagrams illustrating possible configurations of regions.

Keywords: space, logic, existence, completeness.

1 Introduction

Geometry and Topology are both highly developed fields of mathematics. In both areas, the formal theories developed take *points* (or in the case of incidence geometry points and lines) as primitive elements from which objects corresponding to regions are constructed set-theoretically. From the point of view of knowledge representation and automated reasoning this often leads to difficulties. One problem is that the most natural and useful way of presenting many kinds of spatial information is in terms of relationships that hold between regions of space or the bodies that occupy those regions. Another is that the use of set theory leads to highly intractable formal systems.

Although region or solid based formalisms have received relatively little attention a number of significant theories have been produced. de Laguna (1922) gives a very expressive though not fully-formalised theory based on the single primitive relation '*x can connect y and z*'. Whitehead (1929) outlined how many relationships between solids could be defined in terms of the relation '*x is connected to y*'. Leonard and Goodman (1940) give a

*This work was partially supported by the CEC under the Basic Research Action MEDLAR 2, Project 6471 and by the SERC under grants GR/G38652 and GR/H 78955.
†Substantial contributions from Dr A.G. Cohn and Dr N.M. Gotts are gratefully acknowledged.

formal theory based on the relation '*x is discrete from y*'. Building on Leśniewski's theory of *mereology* (an alternative to set theory based on the *part/whole* relation) by introducing a new *sphere* primitive, Tarski (1956) gave a theory of the 'geometry of solids', which is embedded, by means of definitions, into his axiomatisation of elementary Euclidean geometry (Tarski 1959). More recently, Clarke (1981, 1985) presented a calculus based on Whitehead's relation of 'connectedness' and this has been taken up and modified by Randell, Cohn and Cui (1989, 1992).

The under-development of this area consists not so much in a lack of theories but rather that the consequences and interconnections of the various theories have not been fully explored. A major reason for the difficulties encountered in evaluating and comparing the various theories is the lack of meta-mathematical and model-theoretic results about the calculi: the systems have for the most part been presented as uninterpreted calculi, with models being suggested only to give some intuitive understanding of the primitive concepts.[1] In view of this fact, the short paper 'Connection Structures' by Biacino and Gerla (1991) is very significant because it clearly reveals the relationship between the theories of Leonard and Goodman and of Clarke to the well-known mathematical structures of Boolean algebra and (ortho-complemented lattice).

Amongst all these theories, that of Tarski (1956) is the only one that is presented as *complete* and *categorical* (w.r.t. denumerable models). The completeness[2] means that any formula, containing only the vocabulary of the theory, can be proved either true or false (which means that the system is also decidable) and the categoricity means that every (denumerable) model is isomorphic to the intended model in which the primitives of the theory are interpreted with respect certain to geometrical and topological properties of three-dimensional Euclidean space.

A particular weakness in all the theories is the lack of attention to existential axioms. The universal properties of primitives such as 'connectedness' or the 'part/whole' relation seem to be rather more obvious than are statements guaranteeing the existence of regions exhibiting specific properties and configurations. Existential axioms require us to make choices about what counts as a region and to be definite about the domain of regions and its structure. Accordingly they are essential in rendering a theory categorical and thus fixing a single model modulo isomorphism (and denumerability if the theory is 1st-order[3]).

Tarski's theory is only made categorical by indirect means: firstly the notions of *point*, *equidistance* and *betweenness* are introduced by a series of definitions; then it is stipulated that these defined concepts obey the axioms of Euclidean geometry (Tarski 1959). He admits that the resulting system is not ideal:

> The postulate system given above is far from simple and elegant; it seems very likely that this postulate system can be essentially simplified by using intrinsic properties of the geometry of solids. (Tarski 1956)

He then gives an example of how an axiom stated indirectly in terms of points can be replaced by a simple existential axiom concerning the primitive notion of *sphere*.

In this paper I investigate existential axioms for the theory of spatial regions of Randell, Cui and Cohn (1992) (henceforth the RCC theory). This theory already contains existential axioms which, in contrast to Tarski's, are direct and strictly 1st-order. They are also more sophisticated than the theories of Leonard and Goodman (1940) and Clarke (1981, 1985), in that care has been taken to avoid problems (e.g. collapse of 'connection' to 'overlapping') which occur in these earlier theories. Nevertheless, it is unlikely that the RCC axioms provide a complete and categorical theory and it is not clear exactly what class of structures are characterised.

---

[1] It is worth noting that model theoretic properties of calculi of temporal intervals are much better understood than those of spatial formalisms. Allen's (1981) interval calculus has been thoroughly investigated by Ladkin (1987).

[2] The sense of 'completeness' here differs form the (more recently adopted) notion of the 'completeness' of the proof rules of a logic w.r.t. some semantics, which means that all entailments that are valid according to the semantics are derivable using the proof rules of the logic.

[3] If a 1st-order theory has a denumerable model it will also have non-denumerable models which cannot be excluded by any 1st-order axiom. In considering whether a 1st-order theory is categorical these can be discounted. A theory in which all denumerable (or smaller) models are isomorphic is more precisely called $\aleph_0$-categorical

Relation	*interpretation*	*Definition of* $R(x,y)$
$\mathsf{DC}(x,y)$	x is disconnected from y	$\neg\mathsf{C}(x,y)$
$\mathsf{P}(x,y)$	x is a part of y	$\forall z[\mathsf{C}(z,x) \rightarrow \mathsf{C}(z,y)]$
$\mathsf{PP}(x,y)$	x is a proper part of y	$\mathsf{P}(x,y) \wedge \neg\mathsf{P}(y,x)$
$\mathsf{EQ}(x,y)$	x is identical with y	$\mathsf{P}(x,y) \wedge \mathsf{P}(y,x)$
$\mathsf{O}(x,y)$	x overlaps y	$\exists z[\mathsf{P}(z,x) \wedge \mathsf{P}(z,y)]$
$\mathsf{DR}(x,y)$	x is discrete from y	$\neg\mathsf{O}(x,y)$
$\mathsf{PO}(x,y)$	x partially overlaps y	$\mathsf{O}(x,y) \wedge \neg\mathsf{P}(x,y) \wedge \neg\mathsf{P}(y,x)$
$\mathsf{EC}(x,y)$	x is externally connected to y	$\mathsf{C}(x,y) \wedge \neg\mathsf{O}(x,y)$
$\mathsf{TPP}(x,y)$	x is a tangential proper part of y	$\mathsf{PP}(x,y) \wedge \exists z[\mathsf{EC}(z,x) \wedge \mathsf{EC}(z,y)]$
$\mathsf{NTPP}(x,y)$	x is a nontangential proper part of y	$\mathsf{PP}(x,y) \wedge \neg\exists z[\mathsf{EC}(z,x) \wedge \mathsf{EC}(z,y)]$

Table 1: Relations definable in terms of C

2 The RCC Theory

The RCC formalism (Randell et al. 1992) (when I refer to RCC I always refer to the theory presented in that paper) is an axiomatisation of certain spatial concepts and relations in classical 1st-order predicate calculus[4]. The basic theory assumes just one primitive dyadic relation: $\mathsf{C}(x,y)$ read as 'x connects with y'. Individuals can be interpreted as denoting spatial regions. The C relation is reflexive and symmetric, which is ensured by the following two axioms:

$$\forall x \mathsf{C}(x,x) \qquad \textbf{(Cref)}$$

$$\forall xy[\mathsf{C}(x,y) \rightarrow \mathsf{C}(y,x)] \qquad \textbf{(Csym)}$$

Using C, a basic set of dyadic relations are defined:

The relations: P,PP,TPP and NTPP, being non-symmetrical, support inverses. For the inverses the notation Φ^{-1} is used, where $\Phi \in \{\mathsf{P}, \mathsf{PP}, \mathsf{TPP}, \mathsf{NTPP}\}$. These relations are defined by definitions of the form $\Phi^{-1}(x,y) \equiv_{\text{def}} \Phi(y,x)$. Of the defined relations, DC,EC,PO,EQ,TPP,NTPP, TPP^{-1} and NTPP^{-1} have been proven to form a jointly exhaustive and pairwise disjoint set.

The theory also contains a number of quasi-Boolean functions whose definitions can be regarded as existential axioms. These will be discussed below.

3 Identity and Extensionality

Before considering questions of existence of regions we need a clear idea of the conditions for identity and individuation of regions.

Axiomatic theories (particularly those which seek to characterise a single primitive relation), often contain some kind of *axiom of extensionality*. This is an axiom which asserts that identity of any two objects follows from their indiscernibility with respect to some property. Thus in set theory we have:

$$\forall x \forall y[\forall z[z \in x \leftrightarrow z \in y] \rightarrow (x = y)]$$

Such axioms can be regarded as strengthenings of Leibniz' principle of the *identity of indiscernibles*. This principle is the left-to-right component of a second order axiom which can be regarded as defining identity:

$$\forall x \forall y[\ \forall \Psi[\Psi(x) \leftrightarrow \Psi(y)]\ \leftrightarrow\ x = y\ ]$$

[4]In fact — for mainly technical reasons — a *sorted* 1st-order logic is employed; but in the simplified version of the theory presented here we need only one sort — that of *region*

Rather than requiring objects to be indiscernible with respect to all properties, we may require only that they cannot be distinguished in terms of a family of properties formed by (partially) instantiating some relation (over the universe of objects). The idea behind this specialisation of the axiom is that this family of properties is regarded as fixing all properties expressible in the theory. In the RCC calculus of regions the obvious axiom of extensionality would be:

$$\forall x \forall y [\forall z [\mathsf{C}(x,z) \leftrightarrow \mathsf{C}(y,z)] \rightarrow (x = y)] \qquad (\mathbf{Cext})$$

This states that if two regions x and y cannot be distinguished by some instance of $\mathsf{C}(\ldots, z)$ (i.e. we cannot find any region z such that $\mathsf{C}(x,z)$ does not have the same truth-value as $\mathsf{C}(y,z)$) they must be the same region. The force of this axiom is to claim that C is the defining relation for regions: regions can only be distinct if they differ with respect to their connectedness with other regions.

In the RCC theory we can derive something very similar to the axiom of extensionality. From the definitions of $\mathsf{EQ}(x,y)$ and $\mathsf{P}(x,y)$ given above we can very easily show that:

$$\forall x \forall y [\forall z [\mathsf{C}(x,z) \leftrightarrow \mathsf{C}(y,z)] \leftrightarrow \mathsf{EQ}(x,y)] \qquad (\mathbf{CEQ})$$

However, since the 'EQ' symbol is introduced by definition, this derived formulae does not have the force of the axiom of extensionality because 'EQ' need not necessarily have the properties of logical equality. Hence, the derivation does not show that an axiom of extensionality is redundant in the RCC calculus. What it shows rather is that if we take the equivalence

$$\forall x \forall y [(x = y) \ \leftrightarrow \ (\mathsf{P}(x,y) \wedge \mathsf{P}(y,x))] \qquad (\mathbf{P} =)$$

as an axiom rather than a definition and assume that the symbol '=' is to have its usual logical properties, then this formula is equivalent to **Cext** and can thus serve as an axiom of extensionality for the RCC theory.

4 Existence

A complete theory of spatial regions must not only provide us with a means of stating and reasoning about relationships holding between spatial regions. It must also provide us with some kind of ontology of these regions. This will be needed if we are interested in questions regarding the existence of regions.

Some of the definitions of relations in terms of C involve existential quantifiers. This quantification guarantees the existence of certain regions in order to *witness* the fact that some relation holds between other regions. For example, if we know that two regions overlap, this necessitates the existence of a region which is part of both regions. However, such existential commitment is contingent on certain kinds of fact being asserted. No existential facts follow *a priori* from the axioms imposed on C (symmetry and reflexivity).

In order to formulate the existential component of a theory of space we must have some idea of what regions we would like to exist. In the next few sections I investigate what regions we might expect to exist and present axioms which ensure that such regions must exist.

4.1 Demarcation and Existence

The domain of geometry can be characterised by reference to the physical procedure of constructing geometrical figures. More specifically, the domain of 2-dimensional elementary Euclidean geometry can be regarded as those configurations of points (and possibly also lines) constructible on a plane with the aid of a ruler and compass (see (Tarski 1959)). The axioms of elementary geometry have been arrived at by considering basic operations in the construction of figures. For example, given any two distinct points one can introduce a new point which lies between (using the ruler) or is equidistant from (using the compass) the initial two points. In this section I shall indicate how a similar analysis of the way in which configurations of regions may be drawn can lead us to axioms for the existence of spatial regions. My examination will be confined to 2-dimensional regions but, with a little more imagination and considerably more technical complication, this approach could equally well be applied to 3-dimensional space.

Suppose I demarcate a number of regions by drawing their boundaries on a piece of paper and I name each new region as I draw it. There is a sense in which the number of regions I create is greater than the number of regions which I have explicitly named. There are areas of the paper which are demarcated indirectly. For instance, in outlining any region I automatically create its complement, a region consisting of all of space (or all of the paper) except that which is outlined. And, for any two regions (connected or not) which I demarcate I can consider their combined space as another region (one might call this their 'sum' or 'fusion'). One might then say that any bounded region of the paper constitutes a region.[5]

In constructing a categorical theory of spatial regions we clearly cannot be concerned with any particular diagram showing a configuration of regions but rather with all possible such constructions. Thus regions exist whether or not we have actually demarcated them on a piece of paper. Nevertheless, it seems clear that the existence of a particular finite configuration of regions must mean that it would be possible to construct a figure demarcating those regions; and it is in this sense that the possibilities for constructing figures correspond to existential axioms concerning regions.

The RCC theory contains a set of definitions of *quasi-Boolean* functions from regions to regions. These functions can be seen as generating new regions from old in accord with the idea that boundedness in some figure ensures existence:

QBfun1) $\forall x \forall y \forall z[\mathsf{sum}(x,y) = z \;\leftrightarrow\; \forall w[\mathsf{C}(z,w) \leftrightarrow [\mathsf{C}(w,x) \vee \mathsf{C}(w,y)]]]$

QBfun2) $\forall x \forall \mathbf{y}[\mathsf{compl}(x) = \mathbf{y} \;\leftrightarrow\; \forall z[(\mathsf{C}(z,\mathbf{y}) \leftrightarrow \neg\mathsf{NTPP}(z,x)) \wedge (\mathsf{O}(z,\mathbf{y}) \leftrightarrow \neg\mathsf{P}(z,x))]]$

QBfun3) $\forall x \forall y \forall \mathbf{z}[\mathsf{prod}(x,y) = \mathbf{z} \;\leftrightarrow\; \forall w[\mathsf{C}(w,\mathbf{z}) \leftrightarrow \exists v[\mathsf{P}(v,x) \wedge \mathsf{P}(v,y) \wedge \mathsf{C}(w,v)]]]$

QBfun4) $\forall x \forall y[\mathsf{diff}(x,y) = \mathsf{prod}(x, \mathsf{compl}(y))]$

These formulae may be regarded as defining certain functions; but are not purely definitional because functions in 1st-order logic carry existential import: every term must denote some individual in the domain. Thus, e.g., the sum definition guarantees that for every two regions there exists a unique region (their sum) which is connected to all and only those regions connected to either of the original two regions.

This existential import means that all the functions apart from sum give rise to a problem because they are not total. For instance, if two regions are disjoint then there is no (non-empty) region which is their prod. One way round this is to introduce a *null* region as a possible value of the functions. However, care must be taken since in the definitions of topological relations it was assumed that quantifiers ranged over only non-empty regions (e.g. two regions overlap if they have a mutual *non-empty* part).

This difficulty is overcome in Clarke's theory (1981) by restricting permitted instantiations of universal quantifiers to individual variables or terms which have been proved equal to some individual variable. Randell et al. (1992) use a *sorted logic* (see e.g. (Cohn 1987)) to achieve essentially the same effect. Whilst the sorted solution requires more additional formal apparatus than Clarke's it has certain advantages: sorted logics are now well known (especially to computer scientists); sortal information can be effectively handled by automated theorem provers and can even cut down search space and proof length (Cohn 1987); the flexibility of a sorted formalism may be useful in extending and refining the theory. (An alternative solution would be to employ *free* logic, in which quantification does not necessarily carry existential import. The use of such a formalism for describing partial functions in the context of a theory of the 'part-whole' relation is explored in (Simons 1987).)

Space does not permit a full presentation of a sorted theory but for the present purposes a much simplified version will suffice. We stipulate that the quantified variables appearing in bold font in **QBfun2–3** range over non-empty regions and also the null region, whereas all other variables range only over non-empty regions. So, the range of the quasi-Boolean functions includes the null-region even though this is excluded from the domain of quantification in the C axioms and definitions.

The quasi-Boolean functions provide a notation that is very convenient for specifying certain spatial properties; but for the purpose of establishing existential axioms for the domain of (non-empty) regions it is perhaps less than fully satisfactory. The functions enforce an existential commitment which is too strong and then must be reduced by allowing that their value may sometimes be the null region, a pseudo-object, which doesn't really exist. However, the problematic null region may be eliminated entirely if, in place of the function defining axioms, we have existential axioms with suitable preconditions:

[5] We have to be a little bit careful here since we may want infinite, unbounded regions to exist (e.g. we may want a 'universal region' which is unbounded). But for present purposes we can think of all regions being part of some finite universal region which is bounded (e.g. the whole of a piece of paper).

exist1) $\forall x \forall y[\exists! z[\forall w[\mathsf{C}(z,w) \leftrightarrow [\mathsf{C}(w,x) \vee \mathsf{C}(w,y)]]]]$

exist2) $\forall x[\exists y[\neg \mathsf{C}(x,y)] \leftrightarrow \exists! y[\forall z[(\mathsf{C}(z,y) \leftrightarrow \neg \mathsf{NTPP}(z,x)) \wedge (\mathsf{O}(z,y) \leftrightarrow \neg \mathsf{P}(z,x))]]]$

exist3) $\forall x \forall y[\mathsf{O}(x,y) \leftrightarrow \exists! z[\forall w[\mathsf{C}(w,z) \leftrightarrow \exists v[\mathsf{P}(v,x) \wedge \mathsf{P}(v,y) \wedge \mathsf{C}(w,v)]]]]$

exist4) $\forall x \forall y[\neg \mathsf{P}(x,y) \leftrightarrow \exists! z \forall w[\mathsf{C}(z,w) \leftrightarrow \exists v[\mathsf{P}(v,w) \wedge \mathsf{C}(v,x) \wedge \neg \mathsf{NTPP}(v,y)]]]$

These axioms correspond to **QBfun1-4**: the variable that in those axioms was constrained to be equal to some function is now bound by the unique existential quantifier $\exists!$. Also, in the case of **exist2-4** an additional condition is added to limit the existential claim of the axiom. For example, in **exist3** the condition $\mathsf{O}(x,y)$ ensures that a unique 'product' region exists if and only if x and y overlap. **exist4** was arrived at by informally considering the properties of the 'difference' of two regions rather than being derived from **QBfun4**. Since diff is defined immediately from the functions prod and compl **QBfun4** contains no existential commitment of its own; so one would expect that correspondingly **exist4** should be derivable from the other axioms (this has not yet been verified).

The uniqueness of 'sum' regions follows from the axioms of extensionality, **Cext** or P=, so if one of these is adopted the '!' can be dropped from **exist1**. However it can conversely be shown (by supposing $x = y$ in **exist1**) that the uniqueness of sums entails extensionality so **exist1** enforces extensionality as well as guaranteeing the existence of a sum of every pair of regions.

5 Carving Up Space

The quasi-Boolean functions do not tell us how to actually construct figures; they merely ensure that given a figure demarcating a number of regions, certain other (derived) regions must also have been bounded. But the RCC theory also contains an additional existential axiom guaranteeing that every region has a non-tangential proper part:

$$\forall x \exists y[\mathsf{NTPP}(y,x)] \qquad \textbf{(NTPP)}$$

This axiom differs from those involving the quasi-Boolean functions in that it serves to introduce not only new regions but completely new boundaries. Clearly such *carve-up* axioms will be needed if we want to be able to construct all possible figures by decomposing regions into parts in all possible ways that are distinguishable in terms of the theory. In the remainder of this section I shall look at a number of ways in which additional boundaries may be introduced into a figure representing some configuration of regions.

I shall assume that we shall always be dealing with a configuration made up of a number of designated regions which are topologically simple: i.e. they are self-connected, are not made up of parts connected only at a single point and do not contain holes. These conditions can all be defined in terms of the C connective. Such definitions are rather complex and will not be given here. A detailed discussion can be found in (Gotts 1994). Starting from some configuration of simple regions we shall then be concerned with possible divisions of one of the regions into two simple parts parts.

The assumption of topological simplicity of the regions involved greatly simplifies the intellectual manageability of the problem. In particular, it means that, when we consider a diagram representing some configuration of regions, this diagram can be regarded as encompassing all topologically equivalent diagrams. If we allowed that regions might be disconnected (or have some other topological complexity), then an example diagram would have to be thought of as also representing topologically distinct situations in which one or more of the designated regions were split into multiple parts. This would greatly reduce the usefulness of analysing digrams as a means to finding existential axioms. Moreover, the restriction does not reduce the generality of the existential axioms, since regions with more complex topologies can be regarded as constructed, by means of sums and complements, from a number simple parts; and the existence of such sums and complements would be guaranteed by either the axioms **QBfun1-2** or **exist1-2** given above.

Starting with a single simple region there are three ways in which it can be divided into two self-connected parts:

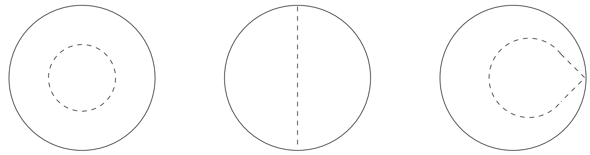

Figure 1: 3 intrinsicly different ways of carving up a region

The possibility of the first case is guaranteed by the **NTPP** axiom in the RCC theory, however the other two possible carve-ups are not. The fact that the theory allows models in which there are regions with no tangential proper parts may be considered a shortcoming of the axioms. This could be remedied by adding the following additional axiom:

$$\forall x \exists y [\mathsf{TPP}(y, x)] \qquad (\mathbf{TPP})$$

As it stands **TPP** does not distinguish between the second and third cases shown in fig. 1; but such a distinction could be made by means of the 'firm tangential part' (FTPP) relation defined in terms of C by Gotts (1994).

Tony Cohn has suggested that a full set of existential axioms might be arrived at by considering successive carve-ups starting from a single undifferentiated region. Thus, for each of the three cases of fig. 1 we would consider all the ways in which an additional division (into two self-connected parts) of one of the regions could be made. We have investigated each of the first two cases but space does not permit description of both, so I present only the analysis of the middle case. From this we can derive the following nine topologically distinct situations:[6]

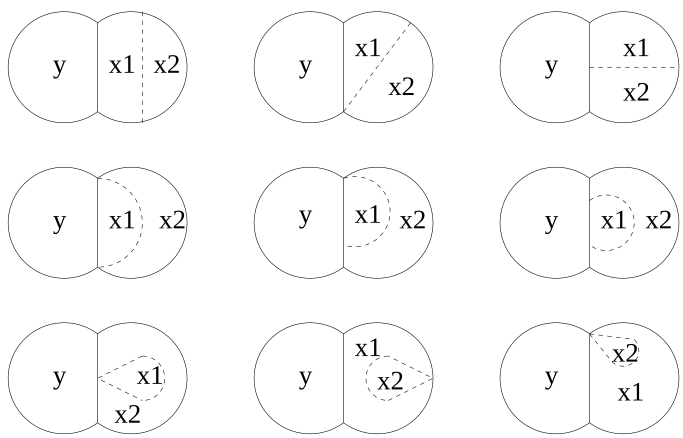

Figure 2: 9 ways in which one of a pair of self-connected EC regions may be split into two self-connected parts

Each of the nine cases corresponds to an existential axioms which can be stated relatively straightforwardly in the RCC system, although one has to make use of quite a few properties and relations, whose definitions in terms of C are quite involved. I shall employ the predicate $\mathsf{WCON}(x)$ (x is 'well-connected' (Gotts 1994)) to assert that x is topologically simple; and the relation $\mathsf{FEC}(x, y)$ to assert that two regions are 'firmly externally connected' (i.e. connected at a boundary segment rather than just at a point) — FEC is straightforwardly definable from Gotts's relation FTPP ('firmly tangential proper part' (Gotts 1994)). The following definitions and axioms should be regarded as first approximations, since it is likely that modifications and refinements will be needed.

It will be helpful to define first the condition of two regions being each topologically simple and being externally connected along a boundary segment:

$$\mathsf{WCONFEC}(x, y) \equiv_{\text{def}} (\mathsf{WCON}(x) \wedge \mathsf{WCON}(y) \wedge \mathsf{FEC}(x, y))$$

and the relation which holds between a region and two well-connected parts into which it is split:

$$\mathsf{SPLIT}(x, y, z) \equiv_{\text{def}} (x = \mathsf{sum}(y, z) \wedge \mathsf{WCON}(y) \wedge \mathsf{WCON}(z) \wedge \mathsf{TPP}(y, x) \wedge \mathsf{TPP}(z, x) \wedge \mathsf{EC}(y, z))$$

[6] We can disregard the case in which $x1$ is an NTPP of $x2$ because this possibility would already be ensured by the standard **NTPP** axiom.

These relations allow fairly concise specification of existential axioms for each situation. For example the top left case corresponds to the following:

$$\forall x \forall y[\mathsf{WCONFEC}(x,y) \rightarrow \exists x1 \exists x2[\mathsf{SPLIT}(x,x1,x2) \wedge \mathsf{EC}(x1,y) \wedge \mathsf{DC}(x2,y)]$$

and the top middle case to:

$$\forall x \forall y[\mathsf{WCONFEC}(x,y) \rightarrow \exists x1 \exists x2[\mathsf{SPLIT}(x,x1,x2) \wedge \mathsf{WCON}(\mathsf{sum}(x1,y)) \wedge \mathsf{EC}(x2,y) \wedge \neg\mathsf{WCON}(\mathsf{sum}(x2,y))]$$

In fact all nine cases can be distinguished by means of simple RCC relations and/or conditions of 'well-connectedness' or otherwise of certain sums of regions.

If this method of analysing successive decompositions is to provide a complete set of existential axioms, we need to be able to show that, once we have considered a sufficient number of cases, then any carve-up of a more complex configuration can be accounted for in terms of a cave-up of one of its simpler sub-parts. Although I have no compelling reason to believe this, it seems quite plausible. However, if it is not the case, then the analysis of possible bisections given in the next section may help complete the picture.

5.1 Describing Carve-ups in Terms of the Ends of New Boundaries

An alternative approach to specifying how additional boundaries my be introduced into arbitrarily complex configurations of (topologically simple) regions is to note that: whenever we split a region into two tangential parts, it is only the location of the end-points of the new bisecting line, relative to surrounding regions, that determines the topology of the resulting configuration.

For example, if a region A is externally connected to two regions B and C then A can be carved into two pieces A_1 and A_2 in such a way that A_1 and A_2 are both connected to B and to C (see figure 3a). In other words A can be split from one boundary section to another, where these boundary sections are defined by external connection to some other region. Alternatively may want to carve a region up in an even more specific way. If a region is externally connected to two other regions, such that all three regions are mutually connected at a point, then we may want to bisect the region at exactly that boundary point (see figure 3b). The other end of the bisection may be either at another such point or on a section of boundary defined by an external connection (as in the previous type of carve-up).

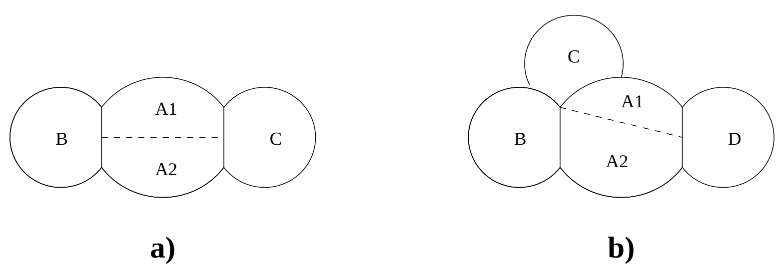

Figure 3: Two ways to carve up a region relative to surrounding regions

If we bisect a region starting from a section rather than a point, this would not necessarily exclude the case where the boundary does in fact join up with an extreme point of the boundary section. Hence we may want to specifically state that such a section should be joined at a mid-point. This would enable us to ensure the existence of regions created by both types of bisection.

It should be noted that this approach does not obviate the need for axioms of the kind described in the last section, because, in order to be able to describe bisections in terms of end-points, we need to first be able to generate configurations as complex as the initial situations shown in fig. 3 . Nevertheless, once configurations of sufficient complexity have been generated, the consideration of possible end-points for further bisections may be a better way to ensure completeness of the existential import of an axiom set.

6 How to Create the Universe

So far we have looked at how to construct new regions from the boundaries of other regions and how to carve regions up to produce new regions with new boundaries. But these forms of region generation will not get us very far unless we have some regions to combine or to chop up. The RCC theory ensures that there is a universal region, which is connected to all regions:

$$\mathsf{us} =_{\text{def}} \iota y[\forall z[\mathsf{C}(z,y)]] \qquad \textbf{(us)}$$

We must consider whether starting from such a region and applying the (relative) existential axioms (which generate regions by combination and dissection of regions whose existence has already been secured) guarantees the existence of all possible configurations of regions.

A number of questions arise immediately. In particular:

1. Is the universe infinite or finite?
2. Is the universe bounded or unbounded?

(Rather than attempting to actually answer these questions one might prefer to find alternative axioms characterising each of these possibilities.) These possibilities lead immediately to differences in the way in which the universe can be decomposed into sub regions. For example, it seems that there are only two ways in which a single boundary can be introduced into an undifferentiated universal region but the effect of such a division depends on whether the universe is finite or infinite:

1. A closed boundary can be introduced which divides the universe into an inner region and an outer region. If the universe is infinite then the inner region will be finite whilst the outer is infinite.
2. A boundary can be introduced which divides the universe into two regions, neither of which is surrounded by the other. If the universe is infinite both of these regions will be infinite.

Hence it is clear that, in determining possible configurations of regions by the method of construction of figures, special considerations must be given to the starting point of such a construction. Specifying the nature of a universal region and the dissections which it can undergo is one way of fixing this starting point. A number of issues concerning the topology of the whole universe and how this relates to problems of defining topological properties of regions within the universe have been examined by Gotts (1994).

7 Alternative Approaches

There are a number of alternative ways one could try to arrive at a complete theory of spatial regions. Following an approach which has been successfully employed in the analysis of temporal relations (Ladkin 1987) one can formulae a theory of spatial relations in terms of *Relation Algebra*. Existential commitment is then implicit in the *composition* operation on relations. This was discussed by me in (Bennett 1994) and more recently I have constructed a fully formal relation algebra based on the C relation (the status of this algebra w.r.t. the original 1st-order theory is currently under investigation).

Alternatively one can aim for completeness without directly considering existential commitment: (Bennett 1995) suggests how a decision procedure for the RCC calculus might be achieved through a *quantifier elimination* procedure.

8 Conclusion

I have briefly described a number of formal systems for representing and reasoning about spatial regions and solid objects, and have identified the insufficient attention to existential axioms as a weakness in these theories. I have suggested a method of eliciting existential axioms by examining and formalising the process of construction of figures illustrating configurations of regions. This method is analogous to the derivation of axioms of elementary geometry from a consideration of possible constructions with ruler and compass.

Further work needs to be done to establish a complete set of modes of figure construction. These operations could then be formalised in terms of the RCC calculus to yield existential axioms which could be added to the theory. One would expect that the resulting system would be *complete* in the sense that all formulae containing only logical symbols and relations/functions defined in the theory would be provable either true or false.

References

Allen, J. F.: 1981, An interval-based representation of temporal knowledge, *Proceedings 7th IJCAI*.

Bennett, B.: 1994, Some observations and puzzles about composing spatial and temporal relations, *in* R. Rodríguez (ed.), *Proceedings ECAI-94 Workshop on Spatial and Temporal Reasoning*.

Bennett, B.: 1995, Towards a decision procedure for the RCC theory of spatial regions, *Proceedings of AISB95 workshop — Automated Reasoning: bridging the gap between theory and practice*.

Biacino, L. and Gerla, G.: 1991, Connection structures, *Notre Dame Journal of Formal Logic*.

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.: 1987, A more expressive formulation of many sorted logic, *Journal of Automated Reasoning* **3**, 113–200.

de Laguna, T.: 1922, Point, line and surface as sets of solids, *The Journal of Philosophy* **19**, 449–461.

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.

Ladkin, P.: 1987, *The Logic of Time Representation*, PhD thesis, University of California, Berkeley. Kestrel Institute report KES.U.87.13.

Leonard, H. S. and Goodman, N.: 1940, The calculus of individuals and its uses, *Journal of Symbolic Logic* **5**, 45–55.

Randell, D. A., Cui, Z. and Cohn, A. G.: 1992, A spatial logic based on regions and connection, *Proc. 3rd Int. Conf. on Knowledge Representation and Reasoning*, Morgan Kaufmann, San Mateo, pp. 165–176.

Randell, D. and Cohn, A.: 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.

Simons, P.: 1987, *Parts: A Study In Ontology*, Clarendon Press, Oxford.

Tarski, A.: 1956, Foundations of the geometry of solids, *Logic, Semantics, Metamathematics*, Oxford Clarendon Press, chapter 2. trans. J.H. Woodger.

Tarski, A.: 1959, What is elementary geometry?, *in* L. Brouwer, E. Beth and A. Heyting (eds), *The Axiomatic Method (with special reference to geometry and physics)*, North-Holland, Amsterdam, pp. 16–29.

Whitehead, A. N.: 1929, *Process and Reality*, The MacMillan Company, New York.