A Qualitative Account of Discrete Space

Anthony J. Roy and John G. Stell

School of Computing, University of Leeds
Leeds, LS2 9JT, UK
{anthonyr, jgs}@comp.leeds.ac.uk

Abstract. Computations in geographic space are necessarily based on discrete versions of space, but much of the existing work on the foundations of GIS assumes a continuous infinitely divisible space. This is true both of quantitative approaches, using $\mathbb{R}^n$, and qualitative approaches using systems such as the Region-Connection Calculus (RCC). This paper shows how the RCC can be modified so as to permit discrete spaces by weakening Stell's formulation of RCC as Boolean connection algebra to what we now call a connection algebra. We show how what was previously considered a problem—with atomic regions being parts of their complements—can be resolved, but there are still obstacles to the interplay between parthood and connection when there are finitely many regions. Connection algebras allow regions that are atomic and also regions that are boundaries of other regions. The modification of the definitions of the RCC5 and RCC8 relations needed in the context of a connection algebra are discussed. Concrete examples of connection algebras are provided by abstract cell complexes. In order to place our work in context we start with a survey of previous approaches to discrete space in GIS and related areas.

1 Introduction

Models of space are important in GIS. Traditionally two approaches are used: the raster approach, and the vector approach. The difference between the two approaches lies primarily in the assumptions made about the underlying model of space. In the case of the raster, a discrete space is assumed—the pixels representing a partition of "real" space. In the vector approach, the underlying space is assumed to be continuous (i.e., $\mathbb{R}^2$). This latter space is convenient from a theoretical viewpoint, but has deficiencies when applied to real data. Data collected on geographical regions is by its very nature discrete—there will either be error bounds or an imposed granularity (e.g., the pixel size of stored satellite images), which allows measurements to limited precision. When it comes to computing with this data, further error can creep in—the data are considered to be based in $\mathbb{R}^2$, but the computation can be done only at an approximation of this due to the finite nature of the computer. The relationship between the discrete space of computation and idealized continuous space is thus important for GIS.

Two main types of discrete space are reviewed here, graph based approaches based essentially on a set of discrete points or pixels, and cell complex ap-

M.J. Egenhofer and D.M. Mark (Eds.): GIScience 2002, LNCS 2478, pp. 276–290, 2002.

proaches, which seem a promising compromise between vector and raster approaches.

Another important area of interest in GIS is that of spatial relations. Spatial data is only of use to us if we can reason with it, and the relations between regions in space is one of the things it is important to be able to query. Various approaches to formalizing spatial relations have been investigated, notably Egenhofer's 4- and 9-intersection relations [1,2,3], and RCC5 and RCC8 systems based on the region connection calculus (RCC) of Cohn *et al.* [4].

The 4- and 9-intersection relations were developed with continuous space in mind, and the RCC with a mereotopological space in mind (though this has been shown to be equally valid when applied to standard topological formulations of $\mathbb{R}^n$[5]). We review Egenhofer and Sharma's [6] approach applying the 9-intersection relations to a discrete graph based model of space, and Winter and Frank's [7,8] approach of applying the 9-intersection relations to a cell complex model of space. RCC has not been successfully adapted to discrete space, due to the long-standing problem of an atomic region being logically both a part of, and disjoint from, its complement. In an early RCC paper[4], the authors discuss various solutions to this problem, but these all involve postulating additional primitive sorts in order to solve the problem. Galton[9] looks at a reformulation of the spatial relations based on RCC for a discrete graph based space.

In this paper we use an algebraic approach to the RCC, and generalize it so as to admit models of discrete space, as well as the traditional continuous models. A particular model for this approach is shown to be the set of closed sub-complexes of a cell complex, and this gives a discrete space, whilst giving plausible reasons why the fact that atomic regions are parts of their complements is not actually a problem. Before we discuss this new form of RCC, in Section 3.1, we provide a survey of existing work on discrete space in GIS and related areas.

2 Discrete Models of Space

In the following survey we look at two classes of discrete spatial models, graph based approaches and cell complex approaches. The former either take the form of point or pixel based space (nodes) with an adjacency relation (arcs), or realm based approaches using points (nodes) and non-intersecting line segments (arcs). The latter approaches use cell complexes to represent a partition of $\mathbb{R}^2$ in order to derive a topologically more appropriate model of space.

2.1 Graph-Based Approaches

The first class of approaches we look at here are those whose underlying structure can be considered as a graph. The approaches can be further subdivided by the style of approach taken, whether the graph structure is defined by an adjacency relation of some kind on a set of points, or whether it is defined by lines between points.

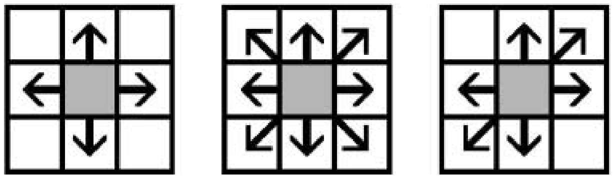

Fig. 1. 4,8 and 6-adjacency on $\mathbb{Z}^2$.

The former approach typically takes $\mathbb{Z}^n$ (or a subset of this) as its set of points (though arbitrary sets of points are of course possible), and defines an adjacency relation on these points in a uniform manner. Galton[9] allows for the adjacency relation to be any symmetric, reflexive relation on arbitrary sets of points (though *strict* adjacency may require the relation to be antisymmetric). However, in general adjacency relations are taken to be one of a few restricted cases. Common adjacency relations on sets of points arranged in a grid are 4-adjacency, 8-adjacency and 6-adjacency. Two points are 4-adjacent if they are vertical or horizontal neighbours in the grid, 8-adjacent if they are either 4-adjacent or are diagonal neighbours. 6-adjacency is a compromise of the two, and two points are 6-adjacent if they are 4-adjacent, or diagonal neighbours on a specified diagonal (see Figure 1).

Such approaches have various problems associated with them (for in depth discussion see [10]), in particular the *connectivity paradox*. This paradox is derived from the Jordan theorem, which states that any simple closed curve divides the plane into exactly two disconnected parts (the inside and outside of the curve)—ideally this theorem would also apply in a discrete model of space. However, using 4-adjacency on a set of points we get the situation where a simple closed curve (i.e., simple closed path in the adjacency graph) may divide the space into more than two disconnected regions, whilst in 8-adjacency we have a situation where a simple closed curve exists, but the space remains in one piece. Figure 2 illustrates these cases: the first diagram shows the disconnected interior of a curve in the 4-connected case; the second diagram shows the adjacency between interior and exterior of the closed curve in the 8-connected case. 6-adjacency does have the required properties to satisfy the Jordan theorem, however it leads to skewed boundaries, and a counter-intuitive asymmetry in the usual square grids. Another attempt at resolving such issues is to have separate adjacency relations for background (4-adjacency) and object (8-adjacency). Problems of course arise here as we need to keep track not only of set membership, but of adjacency, as adjacency relations will have to be updated as set membership changes. Problems also arise when we have situations of varying levels of set inclusion (such as greyscale)—where should the line between object and background be drawn?

A related problem is that of boundary. Classically, the boundary of a subset S of a space comprises the set of elements for which any open neighbourhood overlaps both S and its complement. When we try to transfer this concept to

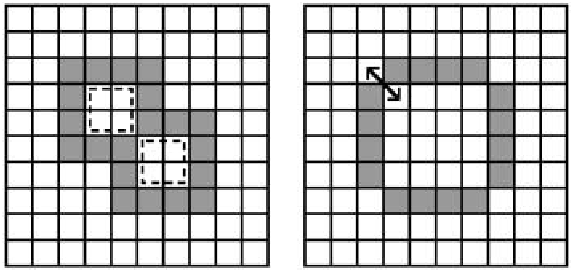

Fig. 2. 8-adjacency and 4-adjacency failing to satisfy the Jordan theorem.

the discrete model, the smallest neighbourhood of a point being those points adjacent to it, we get a broad (2 point wide) boundary (regardless of the type of adjacency).

The Dual Grid Approach. An alternative approach takes sets of points and lines between points as its basis, the points being nodes and the lines being arcs in the underlying graph. Work in this area has been developed by Schneider[11] in his Realm Theory, and more recently by Güting and Lema[12] in their Dual Grid theory.

In realm theory, a realm is defined to be a set of points, and line segments between those points, with the following restrictions: two line segments may not intersect, except at an end point; a point may not lie on a line, except at one of its end points; if a line segment is in the realm, then so are its two endpoints.

In the event that line segments are added to the realm such that they intersect with another line or point somewhere other than at its end points, then a redrawing algorithm is used to alter the points and line segments so that they satisfy the above requirements. This stage is known to be the main source of problems with this approach [12]. Problems include: the rewriting may change the relationships between points and line segments; the space overhead caused by the rewriting algorithm could be large if many objects overlap in one area; objects can change due to rewriting even if the user has no access to modifying objects.

Güting and Lema propose a different approach to the realm, consisting of a set of points with rational coordinates (numerator and denominator having integer values between $\pm 2^n$ and $\pm 2^m$, respectively), and line segments that must start and finish on a grid point, and can lie only on a small subset of possible lines (these "supporting lines" may only intersect at grid points). The allowable lines proposed by the authors are given by the formula $Ax + By = C$ where coefficients are integers and the following conditions are met: $|A|, |B| < \sqrt{2^{m-1}}$ and $|C| < \frac{2^{n-1}}{\sqrt{2^{m-1}}}$.

Regions and Operations on Regions. Given a model of space, we can talk about what subsets of the space should constitute *regions*. Various approaches are in the literature, from arbitrary subsets of points [9] to more complex definitions such as the work done by Egenhofer and Sharma [6].

This latter work restricts regions to consist of a 4-connected non-empty boundary (simple closed curve), with a non-empty interior, and interior and exterior are both 4-connected. Two additional conditions restrict the size of a region, such that it cannot be too small or too large. This essentially restricts regions to the digital analogue of the disc-like regions used in Egenhofer's previous work on topological relations between regions in $\mathbb{R}^n$[1].

Regions in realm theory are built up of cycles (in the underlying graph) of line segments. Face objects consist of a cycle, f, along with a collection of cycles, C, each lying inside f, representing holes in the face. Region objects are collections of faces. Hence in this theory regions can be potentially complex things, multipart entities with holes. The same formulation applies to the Dual Grid theory, since the Dual Grid is translated into a realm before regions and operations between these are defined.

Qualitative Relations on Regions. Two prominent sets of jointly exhaustive, pairwise disjoint (JEPD) qualitative spatial relations (i.e., sets of binary relations for which any pair of regions will satisfy exactly one of the relations) crop up in the literature time and again. These are Egenhofer's 9-intersection relations, looking at the intersections between the interior, boundary and exterior of two regions, and the RCC relations of Cohn *et al.* In the continuous case, both sets of relations can be shown to give a set of 8 relations, in the case of the 9-intersection relations these are restricted to disc-like regions (i.e., single piece regions with no holes). Both of these sets of relations have been looked at in relation to regions in discrete space.

Egenhofer and Sharma have applied the 9-intersection relations to their model of discrete space, and have shown that in their discrete model (note again here that the regions are restricted by definition to be disc-like regions) there are 16 possible JEPD relations between regions, the additional relations being possible due to the nature of the pixel-wide boundary. In the continuous case, if the boundary of a region intersects the interior of another region, then it is the case that they intersect in their interiors. In the discrete case this is no longer the case—two regions may have their boundaries intersecting each others interiors without the interiors intersecting at all. Note that in discrete models where boundaries are not represented, Egenhofer's 4-intersection relations apply to disc-like regions, giving a set of 5 JEPD relations.

Galton also looks at qualitative relations on regions, but in this case uses relations defined to be analogous to the RCC. In the discrete case the definition of part in terms of connection, as in (1) in Section 3.1 below, appears to cause problems, since everything connected to an atomic region is also connected to that regions complement—hence every atomic region is a part of its complement. Galton defines his relations directly on his specific model of space to reflect this problem, and arrives at a set of 8 JEPD relations analogous to the RCC8 (he calls these the RCC8D relations). Randell, Cui and Cohn discuss the problem of atomic parts[4], and come up with various proposed solutions by adding a new primitive to the theory, either *atom* or *point*. In this paper we avoid adding further primitives, and show in Section 3.2 that a simpler solution is possible.

2.2 Cell Complex Approaches

Cell complexes are becoming ever more popular as an alternative model of discrete space to the standard pixel/point approaches, primarily for their topological properties. In the above approaches, sets of points have a trivial topology, that is the open and closed sets in the topology are the same. Attempts at adding a notion of boundary to sets of points do not properly address this problem. Problems crop up such as a region and its complement do not share a boundary, boundaries have a non-zero area, and other such anomalies (see [10]).

Some have considered regarding the "cracks" between pixels as the boundary, as this would seem to solve many of the above problems. This approach is essentially adding a new entity to the model, and is a move toward using cell complexes. Cell complexes are essentially a collection of discrete objects of different dimension. In a two-dimensional picture for example we would have cells of 0, 1 and 2 dimensions. A formal definition is given in Section 5 below.

It should be noted that Kovalevsky has shown ([10]) that not only is a cell complex a topological space, but that any finite separable topological space is a cell complex. For this reason we know that cell complexes have a useful topology.

Much work has been done on the topological aspects of cell complexes, but little on qualitative spatial reasoning using cell complexes as a model of space. Winter and Frank[7] however propose such an approach, combining their *hybrid raster* model of discrete space (a cell complex with a regular grid structure) with the 9-intersections relations. They confirm that the same 8 relations hold between regions defined on the complex as are possible in $\mathbb{R}^2$, given that regions represented in the complex are suitably regular (no 1-dimensional spikes for example).

The approach we consider below looks at the problem from an algebraic point of view, and shows that the RCC8 relations can be axiomatized in an appropriate way such that both the continuous case and the discrete case (with particular models satisfying the correct algebraic structure) are both special cases of the theory, and both are interpreted in the intuitive way.

3 A Discrete Version of the RCC

3.1 The Region-Connection Calculus

The Region-Connection Calculus (RCC) [13] has been widely studied as a basis for qualitative spatial reasoning (QSR). RCC formalizes space by positing a set of regions, the properties of which are given by axioms. The axiomatization is intended to describe continuous, infinitely divisible space. One of the axioms requires that every (non-empty) region has a strictly smaller region lying within its interior. The issue of whether it is possible to modify the RCC so as to permit atomic regions, as would be required for discrete or finite spaces, has been raised in one of the early RCC papers [4]. However the modifications suggested there were substantial and it is debateable whether the resulting systems should count as discrete analogues of RCC or as quite different axiomatizations of discrete space.

In this section we discuss new results on modifying RCC to admit discrete spaces as models, while not excluding continuous models. We propose a system that is an approach to axiomatizing space based on regions but which does not force the space to be either discrete or continuous. We anticipate that axiomatizations in which space is required to be discrete can be constructed using our system as a basis. The usual RCC for continuous space is obtained by adding extra requirements to our axioms. We show that our axioms include one important example of discrete space: that based on the notion of Abstract Cell Complex (ACC).

There are several equivalent ways of formulating the original RCC; we follow Stell's approach [5] in terms of Boolean connection algebras. A BCA is, briefly, a Boolean algebra, A, together with a binary relation of connection satisfying the following axioms, where R denotes A without the bottom element, $\overline{R}$ is R without the top element, and x' is the complement of x.

(B1) C is symmetric.
(B2) C restricted to R is reflexive.
(B3) $\forall x \in \overline{R}\ \ \mathsf{C}(x, x')$
(B4) $\forall x, y, z \in R\ \ \mathsf{C}(x, y+z)$ iff $(\mathsf{C}(x, y)$ or $\mathsf{C}(x, z))$
(B5) $\forall x \in \overline{R}\ \ \exists y \in R\ \ \not\mathsf{C}(x, y)$

The final axiom B5 says that every region has some region that is not connected to it. This does not by itself force space to be continuous, but because the regions form a Boolean algebra, it follows that every region, r, has a region not connected to r', its complement. This is because given $r \in \overline{R}$ the region r' must also be in $\overline{R}$, so B5 ensures the existence of a region, y, not connected to r'. Since every region is connected to its complement, y must be strictly smaller that r and so every region has a smaller region within it. We see that the interaction between the Boolean complement and axiom B5 is what forces spaces that satisfy the axioms to be continuous.

It has already been noted that the complement operation provides some kind of obstacle to the generalization of RCC to spaces with atomic regions. In [4] it is explained how an atomic region would necessarily be part of its complement. This appears to be a contradiction—surely no region could be part of its complement. However this apparent contradiction can be avoided by weakening the properties of the complement operation. In Section 3.2 we use a dual pseudo-complement, and with this weaker type of complement atomic regions can actually be parts of their complements.

Even with this new interpretation of complement, it is not possible to maintain the relationship (1) between part and connection in the case of finitely many regions.

$$\mathsf{P}(x, y) \text{ iff } \forall z\ \ \mathsf{C}(z, x) \text{ implies } \mathsf{C}(z, y) \tag{1}$$

However, there are good reasons for expecting this relationship to fail to hold, and we discuss these below.

3.2 Modifying the Axioms to Permit Discrete Space

To investigate a discrete variant of RCC we weaken the notion of Boolean connection algebra, to what we call simply a connection algebra. This uses a weaker kind of lattice than the Boolean algebra used in Boolean connection algebra. Here we use what is known as a dual pseudo-complemented distributive lattice, or a dual p-algebra. A full definition can be found in [14], but the main point to note here is that there is a complement-like operation (the dual pseudo-complement), which satisfies $r + r^* = \top$, but not necessarily $r \cdot r^* = \bot$. This allows a region to be a part of its (dual pseudo-)complement.

Definition 1 (connection algebra). *A* connection algebra *is a dual p-algebra* $\langle A; \bot, \top, ^*, +, \cdot \rangle$ *equipped with a binary relation* C *satisfying the following axioms, where* $R = A - \{\bot\}$, *and* $\overline{R} = R - \{\top\}$.

(CA1) C *is symmetric.*
(CA2) C *restricted to* R *is reflexive.*
(CA3) $\forall x \in \overline{R}\ \ \mathsf{C}(x, x^*)$
(CA4) $\forall x, y, z \in R\ \ \mathsf{C}(x, y + z)$ *iff* $(\mathsf{C}(x, y)$ *or* $\mathsf{C}(x, z))$

There are no finite BCAs (once the trivial one or two element cases are excluded), but there are finite connection algebras, one family of which is discussed in detail in Section 5. In all the finite cases, the relationship (1) can no longer hold. That is, it can be proved that in any finite connection algebra there will be distinct elements a_1 and a_2 that are connected to exactly the same elements, so $\mathsf{EQ}(a_1, a_2)$ holds. This is equivalent to the relationship (1) always failing in a finite connection algebra. We do not regard this failure as a problem with the formulation of connection algebras. The relationship between part and connection is important for continuous space, which is what BCA's model, but finite and discrete spaces are approximations to continuous space. Approximations inevitably involve loss of information, and it seems that this loss of information can be used to explain why the relationship fails. To each continuous region, r, we can associate the set of all regions to which it is connected, $C(r)$. What can happen in passing from continuous regions to their approximations, is that two regions r_1 and r_2 are approximated as different elements, a_1 and a_2, of the connection algebra while the sets $C(r_1)$ and $C(r_2)$ are approximated by the same set of elements of the connection algebra. The intuition seems to be that in the continuous world there are enough regions about to be able to distinguish r_1 from r_2, but when we have only finitely many regions this is no longer the case.

4 Sets of Spatial Relations

It has already been noted that qualitative spatial relations are an important part of spatial reasoning. In this section we look in more depth at the Region-

Connection Calculus, and in particular JEPD sets of binary relations within the RCC.

Two sets are of interest here, the RCC5 relations, and the RCC8 relations. RCC5 consists of the relations $\{\mathsf{DR}, \mathsf{PO}, \mathsf{PP}, \mathsf{PPi}, \mathsf{EQ}\}$. These have the intuitive meanings of discrete regions, partial overlap, proper part, proper part inverse, and equality. Figure 3 illustrates these:

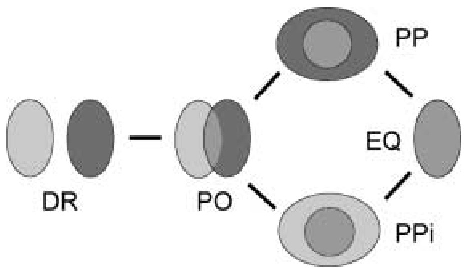

Fig. 3. The RCC5 relations.

The RCC5 relations are defined by:

$$\begin{aligned}
\mathsf{DR}(x,y) &\iff \neg\mathsf{O}(x,y)\\
\mathsf{EQ}(x,y) &\iff \mathsf{P}(x,y)\wedge\mathsf{P}(y,x)\\
\mathsf{PO}(x,y) &\iff \mathsf{O}(x,y)\wedge\neg\mathsf{P}(x,y)\wedge\neg\mathsf{P}(y,x)\\
\mathsf{PP}(x,y) &\iff \mathsf{P}(x,y)\wedge\neg\mathsf{P}(y,x)\\
\mathsf{PPi}(x,y) &\iff \mathsf{PP}(y,x)
\end{aligned}$$

These five relations use the definition of overlap given in (2) below.

$$\mathsf{O}(x,y) \iff \exists z\in R\ \ \mathsf{P}(z,x)\wedge\mathsf{P}(z,y) \tag{2}$$

In our generalized version of RCC, the elements of R may be boundaries of regions or parts of boundaries, as well as genuine regions. This means that a modification to the definition of overlap is required, since, in the case of closed regions, any region has its boundary as a part of it. Thus, the definition (2) would entail regions not just connected to their complements but overlapping them.

The solution to this problem is to require overlapping regions to share a region with non-empty interior; that is something that is not merely a boundary or part of one.

$$\mathsf{O}(x,y) \iff \exists z\in R\ \ \mathsf{P}(z,x)\text{ and }\mathsf{P}(z,y)\text{ and }z^*\neq\top \tag{3}$$

In the case that the underlying algebra is Boolean, this is equivalent to the definition in (2) since then $z^*\neq\top$ for any $z\in R$. There are many alternative ways of characterizing those entities that are not purely boundaries, for example $z \wedge z^* \neq z$. A stronger requirement would be $z^{**} = z$. Entities satisfying this cannot be boundaries, but additionally they must be regular, so excluding for example a closed disc with a one dimensional spike attached. The possibility of special status for entities that satisfy $z^{**} = z$ is mentioned in a similar algebraic setting by Lawvere [15, p10].

With overlap defined as in (3), the RCC5 relations form a JEPD set when restricted to entities that are not purely boundaries. However, if we consider two equal boundaries they will be both EQ and DR. The formulation of relations between boundaries and between boundaries and non-boundary elements is the subject of ongoing work, but we anticipate that different sets of JEPD relations may be required according to whether we are dealing with boundaries or not.

The RCC5 relations can be extended to the RCC8, by looking at whether regions are connected as well as whether they overlap. This essentially has the effect of splitting each of DR, PP and PPi into two, as seen in Figure 4.

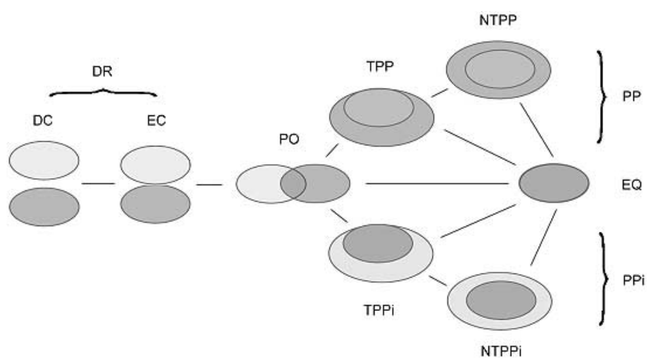

Fig. 4. The RCC8 relations, showing their relationship with the RCC5.

The definitions of the RCC8 relations are PO and EQ defined as in the RCC5 with the following additions:

$$\begin{aligned}
\mathsf{DC}(x,y) &\iff \neg\mathsf{C}(x,y)\\
\mathsf{EC}(x,y) &\iff \mathsf{C}(x,y) \wedge \neg\mathsf{O}(x,y)\\
\mathsf{TPP}(x,y) &\iff \mathsf{PP}(x,y) \wedge \exists z \mathsf{EC}(z,x) \wedge \mathsf{EC}(z,y)\\
\mathsf{TPPi}(x,y) &\iff \mathsf{TPP}(y,x)\\
\mathsf{NTPP}(x,y) &\iff \mathsf{PP}(x,y) \wedge \neg\exists z \mathsf{EC}(z,x) \wedge \mathsf{EC}(z,y)\\
\mathsf{NTPPi}(x,y) &\iff \mathsf{NTPP}(y,x)
\end{aligned}$$

Similarly to the RCC5 relations, the definitions as they stand are deficient when applied in the context of our connection algebras. Consider the case where we have a pair of regions one of which is a proper part of the other, and whose boundaries do not share any common parts. We would expect these to have the

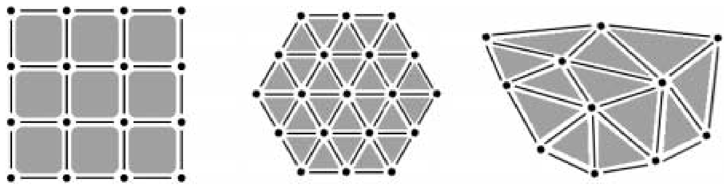

Fig. 5. Abstract Cell Complex with 0-, 1-, and 2-cells.

relation NTPP, when in fact the existence of boundary regions in our calculus means that they *also* have the relation TPP! Hence we again qualify the relations TPP and NTPP by ensuring that the externally connected region is a non-boundary region:

$$\mathsf{TPP}(x,y) \iff \mathsf{PP}(x,y) \wedge \exists z \mathsf{EC}(z,x) \wedge \mathsf{EC}(z,y) \wedge z^* \neq \top$$
$$\mathsf{NTPP}(x,y) \iff \mathsf{PP}(x,y) \wedge \neg\exists z \mathsf{EC}(z,x) \wedge \mathsf{EC}(z,y) \wedge z^* \neq \top$$

This set of relations again form a JEPD set when restricted to non-boundary regions.

5 Connection Algebras Constructed from Cell Complexes

This section introduces the notion of a *cell complex*, and shows how such a complex, with regions defined in a certain way, provides a finite model of a connection algebra. This gives us a concrete example of the axiomatic theory, and one that is a sound model of discrete space.

5.1 Abstract Cell Complexes

A cell complex is often thought of as a partition of space, where the elements (cells) of the partition are of varying dimension. The cells of highest dimension can be thought of as open subsets of the same dimension as the space, bounded by elements of lower dimension. Cells of dimension n are denoted n-cells.

The diagram in figure 5 provides a visualization of various cell complexes, a regular grid, a regular triangular grid and a triangulation. The 2-cells are shown as shaded regions, the 1-cells by lines and the 0-cells as solid discs.

The following definition formalizes the notion of a cell complex:

Definition 2 (Abstract Cell Complex). *An* abstract cell complex *(ACC)* $K = (E, B, dim)$ *is a set* E *along with an antisymmetric, irreflexive and transitive* bounding *relation* B*, on* E*, and a function* $dim : E \to \mathbb{N}$ *taking elements of* E *to their dimension. If* $(a, b) \in B$ *then we say* a bounds b*, or that* a *is a* face *of* b*. The ACC is required to satisfy the condition that whenever* a *bounds* b*, the dimension of* a *is strictly less than that of* b*.*

An element a such that $dim(a) = n$ is called an *n-dimensional* element, or *n-cell*. The dimension of a complex is the dimension, n, of its highest dimension cell, and is denoted an *n-complex*. A sub-complex S of a complex K is simply a subset of the elements of K along with the dimension and bounding relations restricted to that subset.

The definitions of closure and interior of a sub-complex formalize the ideas of adding all bounding cells, and of removing all boundary cells from a sub-complex. The closure of a sub-complex S of a complex $K = (E, B, dim)$, is the union of S and all cells bounding the cells in S: $cl(S) = \{x \in E \mid x \in S \text{ or } \exists y \in S\ (x, y) \in B\}$. Dually the interior of S is: $int(S) = \{x \in S \mid \forall y \in E \text{ if } (x, y) \in B \text{ then } y \in S\}$, i.e., the subset of S that contains all elements of S that only bound elements of S.

5.2 Cell Complexes as Dual P-Algebras

If we assume that any region that we are interested in will include its boundary, then we can define regions in a cell complex as the set of closed subcomplexes. We can show that such a set forms a dual p-algebra. Taking connection to be intersection at a 0-cell, we can also show that the resulting structure is a connection algebra.

Formally, Given any ACC, $K = (E, B, dim)$, its set of closed subsets, $\overline{K} = \{X \subseteq E : cl(X) = X\}$, forms a dual p-algebra $\langle \overline{K}; +, \cdot, ^*, \bot, \top \rangle$, where for any $x \in \overline{K}$, the dual pseudo-complement x^* is defined to be the closure of the set-theoretic complement of x, and where $+$ and $\cdot$ are the set theoretic union and intersection of sets of cells.

To define a connection relation on $\overline{K}$ we use the symmetric closure of the bounding relation writing $x \sim y$ iff x bounds y or y bounds x. Then connection is defined for $x, y \in \overline{K}$ by

$$\mathsf{C}(x, y) \iff \exists a \in x, b \in y\ x \sim y.$$

Note that in the ACC case this is equivalent to two regions sharing a 0-cell, since if $a \in x$ bounds $b \in y$, then $a \in y$, since x and y are both closed, and similarly any 0-cells bounding a will be in both x and y.

The pseudo-complement is illustrated by an example in figure 6. The figure shows a region (i.e., a closed subcomplex of a 2-dimensional ACC), as the set of cells contained wholly within the oval, and its dual pseudo-complement (indicated similarly). The key feature to note is that certain 1-cells and 0-cells lie both in the region and in its pseudo-complement, so that if r is the region then $r \wedge r^*$ is not empty.

When we consider the pseudo-complement of those closed regions with no interior, which include in particular the singleton 0-cells, and the sets of cells (of dimension less than n), which form the boundaries of regions having interiors, we find that if r is such a region then $r^* = \top$ and so $r^{**} = \bot$.

This illustrates how the apparent paradox of atomic regions being parts of their complements is resolved. Once we replace the Boolean complement by the

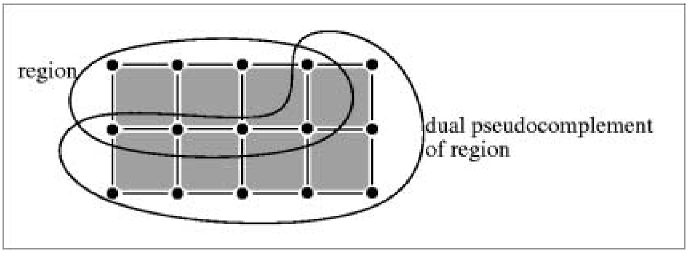

Fig. 6. Example of dual pseudo-complement on a cell complex.

dual pseudo-complement we find that r genuinely is part of r^* since r^* is the whole space. It is also significant that entities being parts of their complements in this sense is not just a feature of atomic regions—it applies once we admit boundaries.

6 Conclusions and Further Work

We have shown how to modify the Region-Connection Calculus so that its models include discrete spaces as well as the continuous ones ,which it already allows. The significance of the particular approach used is that the resulting system includes the existing RCC, unlike some earlier attempts to admit atomic regions. Our work also shows that the apparent contradiction of atomic regions being parts of their complements can be resolved by interpreting "complement" in an appropriate way. We have shown how to modify the definitions of the RCC8 relations for non-boundary elements of connection algebras. Evidence of the suitability of the notion of connection algebra has been provided by showing that the notion of discrete space based on abstract cell complexes yields a model of the theory.

There is much scope for further work. Current work is concerned with implementation of algorithms such as convex hull for ACCs, and convexity is discussed briefly below. Future theoretical work will examine what models exist for the modified RCC besides the one exhibited here. We also expect to extend the work to encompass discrete regions subject to vagueness and granularity.

Convexity

More expressive sets of RCC relations (such as the RCC15) can be defined by adding additional operators such as *convex hull.* We have begun to investigate various convex hull operators in discrete space, with a view to forming an axiomatization of such an operator. Possible axioms include the following given a dual p-algebra $\langle A; \bot, \top, ^*, +, \cdot \rangle$:

1. $\forall x \quad conv(x) = conv(conv(x))$ (Idempotent)
2. $\forall x \quad x \leq conv(x)$ (Increasing)

3. $\forall x, y \quad x \leq y \rightarrow conv(x) \leq conv(y)$ (Monotonic)
4. $\forall x, y, z \quad (y \leq conv(z + x) \wedge y \not\leq z) \rightarrow x \not\leq conv(y + z)$ (Anti-exchange)
5. $\forall x, y \quad conv(x) + conv(y) \leq conv(x + y)$ (Union)
6. $\forall x, y \quad conv(x) \cdot conv(y) = conv(x \cdot y)$ (Intersection)

Various definitions of convex hull operators in the particular case of cell complexes exist. Three contenders for a convex hull operator we shall call the *image hull*, the *open convex hull* an the *closed convex hull*. The *image hull* is simply the digital image of the convex hull of the set when considered embedded is continuous space. The open and closed convex hulls are the intersections of all open (resp. closed) halfplanes containing the set. All three operators are increasing, monotonic and have the union and intersection properties. However, the image hull is not idempotent, and it is unclear as yet which (if any) of the operators has the anti-exchange property.

Acknowledgements

The work is supported by EPSRC under the project: Digital Geometry and Topology: An Axiomatic Approach with Applications to GIS and Spatial Reasoning.

References

1. Egenhofer, M.J., Franzosa, R.: Point-set topological spatial relations. International Journal of Geographical Information Systems **5** (1991) 161–174
2. Egenhofer, M.J., Herring, J.R.: Categorizing binary topological relations between regions, lines and points in geographic databases. Technical report, Department of Surveying Engineering, University of Maine (1991)
3. Egenhofer, M.J., Franzosa, R.: On the equivalence of topological relations. International Journal of Geographical Information Systems **9** (1995) 133–152
4. Randall, D.A., Cui, Z., Cohn, A.G.: A spatial logic based on regions and connection. In Nebel, B., Rich, C., Swartout, W., eds.: Principles of Knowledge Representation and Reasoning. Proceedings of the Third International Conference (KR92), Morgan Kaufmann (1992) 165–176
5. Stell, J.G.: Boolean connection algebras: A new approach to the region-connection calculus. Artificial Intelligence **122** (2000) 111–136
6. Egenhofer, M., Sharma, J.: Topological relations between regions in $\mathbb{R}^2$ and $\mathbb{Z}^2$. In Abel, D., Ooi, B.C., eds.: Advances in Spatial Databases – Third International Symposium on Large Spatial Databases, SSD '93. Volume 692 of Lecture Notes in Computer Science., Springer–Verlag (1993) 316–336
7. Winter, S., Frank, A.U.: Topology in raster and vector representation. Geoinformatica **4** (2000) 35–65
8. Winter, S.: Topological relations between discrete regions. In Egenhofer, M.J., Herring, J.R., eds.: Advances in Spatial Databases. Volume 951 of Lecture Notes in Computer Science. Springer, Berlin (1995) 310–327
9. Galton, A.: The mereotopology of discrete space. In Freska, C., Mark, D.M., eds.: Spatial Information Theory. Volume 1661 of Lecture Notes in Computer Science., Berlin, Springer-Verlag (1999) 251–266

10. Kovalevsky, V.A.: Finite topology as applied to image analysis. Computer Vision, Graphics and Image Processing **46** (1989) 141–161
11. Schneider, M.: Spatial Data Types for Database Systems. Volume 1288 of Lecture Notes in Computer Science. Springer (1997)
12. Lema, J., Güting, R.: Dual grid: A new approach for robust spatial algebra implementation. GeoInformatica **6** (2002) 57–76
13. Cohn, A.G., Bennett, B., et al.: Qualitative spatial representation and reasoning with the region connection calculus. GeoInformatica **1** (1997) 275–316
14. Balbes, R., Dwinger, P.: Distributive Lattices. University of Missouri Press (1974)
15. Lawvere, F.W.: Introduction. In Lawvere, F.W., Schanuel, S.H., eds.: Categories in Continuum Physics. Volume 1174 of Lecture Notes in Mathematics., Springer-Verlag (1986) 1–16