

A taxonomy of partitions

Thomas Bittner and Barry Smith

Department of Computer Science, Northwestern University, bittner@cs.nwu.edu
Department of Philosophy, Center for Cognitive Science and NCGIA, State University of
New York, Buffalo, phismith@buffalo.edu

Abstract: In this paper we propose a formal theory of partitions (ways of dividing up or sorting or mapping reality) and we show how the theory can be applied in the geospatial domain. We characterize partitions at two levels: as systems of cells (theory A), and in terms of their projective relation to reality (theory B). We lay down conditions of well-formedness for partitions and we define what it means for partitions to project truly onto reality. We continue by classifying well-formed partitions along three axes: (a) degree of correspondence between partition cells and objects in reality; (b) degree to which a partition represents the mereological structure of the domain it is projected onto; and (c) degree of completeness and exhaustiveness with which a partition represents reality. This classification is used to characterize three types of partitions that play an important role in spatial information science: cadastral partitions, categorical coverages, and the partitions involved in folk categorizations of the geospatial domain.

1 Introduction

Imagine that you are standing on a bridge above a highway checking off the makes and models of the cars that are passing underneath. Or that you are a postal clerk dividing envelopes into bundles or a laboratory technician sorting samples of bacteria into species and sub-species. Or imagine that you are making a list of the fossils in your museum or of the guests in your hotel on a certain night. In each of these cases you are employing a certain grid of cells, and you are recognizing certain objects as being located in those cells. We shall call such a grid of cells a *partition*, and we shall argue that partitions are involved in all listing, sorting, cataloguing and mapping activities.

Partitions are ways of structuring and understanding reality. Some partitions are flat: they amount to nothing more than a mere list. Other partitions are hierarchical: they consist of cells and subcells, the latter being contained within the former. Some partitions are built in order to reflect independently existing divisions on the side of objects in the world (the subdivision of the animal kingdom into species, the subdivision of heavenly bodies into galaxies, stars, planets, moons). Other partitions – for example the partitions created by nightclub doormen or electoral redistricting commissions – are themselves such as to create the necessary divisions on the side of their objects, and sometimes they create those very objects themselves.

In Smith and Brogaard (2000) the notion of partition is introduced as a generalization of David Lewis's (1991) conception of classes as the mereological sums of their singletons. Given its set-theoretical roots, our basic formal ontology of

partitions will have two parts: A. a theory of the relations between cells and the partitions in which they are housed, and B. a theory of the relations between cells and objects in reality. The counterpart of A in a set-theoretic context would be the study of the relations among subsets of a single set; the counterpart of B would be the study of the relations between sets and their members.

Division into units, counting and parceling out, listing, sorting, pigeonholing, cataloguing are activities performed by human beings in their traffic with the world. Partitions are the cognitive devices designed and built by human beings to fulfill these various purposes. As will be clear from what follows, the notion of partition that is hereby implied is only distantly related to the more familiar notion of a partition defined in terms of equivalence classes.

The paper is structured as follows. We start with a discussion of properties of partitions as systems of cells in the sense of theory A. We then consider partitions in their projective relation to reality in the sense of theory B. This provides us with the means to define what it means to say that a partition truly projects onto reality. In Section 4 we provide a classification of partition by characterizing various properties of the correspondence between partition and reality. We consider three classes of partitions that are relevant in the geographic domain in Section 5.

2 Partitions as system of cells

2.1 Partitions

Theory A studies properties partitions have in virtue of the relations between and the operations performed upon the cells from out of which they are built. All partitions involve cells arranged together in some sort of structure. This structure is intrinsic to the partition itself, independently of the objects which may be located in its cells. Cells in partitions may be nested one inside another in the way in which, for example, the species *crow* is nested inside the species *bird* which in turn is nested inside the genus *vertebrate* in standard biological taxonomies. We say that one cell, z_1 , is a subcell of another, z_2 , if the first is contained in the latter ('Cell' is 'Zelle' in German).

We write $z_1 \subseteq z_2$ in order to designate this relationship. The subcell relation \subseteq is a reflexive, transitive, and antisymmetric relation. Reflexivity means that each cell is, trivially, a subcell of itself. The property of antisymmetry specifies the relationship between identity and the subcell relation: if $z_1 \subseteq z_2$ and $z_2 \subseteq z_1$ then $z_1 = z_2$. Transitivity means that if $z_1 \subseteq z_2$ and $z_2 \subseteq z_3$ then $z_1 \subseteq z_3$. The nestedness of cells inside a partition yields chains of cells satisfying $z_1 \supseteq z_2 \supseteq \dots \supseteq z_n$. We shall demand, as the first of several master conditions which we shall impose on all partitions, that:

MA1: The chains $z_1 \supseteq z_2 \supseteq \dots \supseteq z_n$ in a partition are always finite.

We shall call the cells at the ends of such chains *minimal*. MA3 leaves open the issue as to whether partitions themselves are finite: thus it does not rule out the possibility that a given cell within a partition might have infinitely many immediate subcells.

Following Smith (1991) we can define the partition-theoretic sum and product of cells within partitions as follows. The partition-theoretic sum $z = z_1 \cup z_2$ of two cells

in a partition as a \subseteq -minimal cell satisfying a constraint to the effect that $z_1 \subseteq z$ and $z_2 \subseteq z$. Notice that this partition-theoretic sum is distinct from the mereological sum of cells. (Relative to the standard geopolitical partition of the surface of the globe, the partition-theoretic sum of Malta and Cyprus is: the British Commonwealth.) The partition-theoretic product, $z = z_1 \cap z_2$, of two cells is defined only if z_1 and z_2 are not mereologically disjoint. If it is defined, then it yields a largest subcell shared in common by z_1 and z_2 .

Every partition A ('partition' is 'Aufteilung' in German) has a maximal cell. We define the a maximal cell of a partition as:

$$M(z_1, A) \equiv Z(z_1, A) \text{ and } \forall z: Z(z, A) \rightarrow z \subseteq z_1$$

where 'Z(z, A)' means that z is a cell in the partition A. (In what follows we shall omit the predicate Z(z, A) in cases where it is clear that we are talking about cells within some fixed partition A. In addition, initial universal quantifiers will henceforth be taken as understood.) We now define the root of a partition as the unique maximal cell:

$$r(A)=z \equiv M(z, A) \text{ and } \forall z_1: M(z_1, A) \rightarrow z = z_1$$

and demand that every partition has a root:

$$\text{MA2: } \exists z: r(A)=z$$

This root is such that all the cells in the partition are included in it as subcells. MA2 implies that there are no partitions which are empty *tout court* in that they have no cells at all.

Every pair of distinct cells in a partition stand to each other within the partition either in the subcell relation or in the relation of mereological disjointness. Thus if two cells overlap, then only because one is a subcell of the other:

$$\text{MA3: } \exists z: (z = z_1 \cap z_2) \rightarrow z_1 \subseteq z_2 \text{ or } z_1 \supset z_2.$$

2.2 Trees

Philosophers since Aristotle have recognized that nested partitions can be represented as branching structures which mathematicians nowadays called trees. We here take the possibility of such representation as a constraint on what will be allowed to count as a partition within the terms of our theory.

$$\text{MA4: } \text{Partitions can be represented as trees.}$$

Trees are special sorts of graphs (they are directed graphs without cycles). That a tree is without cycles means that, if we move along its edges in such a way as to follow the direction of these edges, then we will always move down the tree and in such a way that, however far we travel, we will never return to the point from which we started.

The duality of trees and partitions can be proved by showing that every tree can be transformed into a structure, analogous to a Venn diagram, in which partition cells are represented as topologically simple and regular non-intersecting regions of the plane. Conversely every array of non-intersecting, possibly nested rings in the plain can be transformed into a tree in such a way that each ring is represented by a node in the tree, and each directed link in the tree represents an *immediately*

contains relation between the corresponding pair of nested rings. We can for illustrative purposes think of each partition as such a planar map.

2.3 Partition-theoretic sum and maximal cells

When we consider a partition as a Venn diagram, then minimal cells correspond to smallest regions within the diagram. It is important to recognize that it is not in general the case that every mereological sum of such regions has a corresponding cell in the partition. There is no cell in our standard biological partition of the animal kingdom consisting of rabbits together with jellyfish, and there is no cell, in our standard geopolitical partition of the surface of the globe consisting of Hong Kong together with Algeria. It is for this reason that the mereological sum of cells and the partition-theoretic sum of cells defined above are distinct operations. Since, by MA2, partitions always have a root cell, the partition-theoretic sum is always defined. However the results produced by this operation are not necessarily identical to those produced by the operation of taking standard mereological sums. We can however assert that the mereological sum is in every case a part of the partition-theoretic sum. We can think of \subseteq as the result of restricting mereological parthood relation, \leq , to natural units within a mereological whole.

Consider a collection of subsets of some set forming a partition of this set in the standard mathematical sense. These subsets (equivalence classes) are (1) mutually exhaustive and (2) pairwise disjoint. An analogue of condition (2) holds for minimal cells in our present framework, since minimal cells are always mereologically disjoint. Our non-minimal cells correspond to mereological sums of equivalence classes in the standard mathematical framework. Condition (1) does not necessarily hold within the framework of partition theory however. This is because, even where the partition-theoretic sum of minimal cells is identical to the root cell, the minimal cells still do not necessarily exhaust the partition as a whole. Consider a partition with just three cells labelled: *people*, *Bill*, and *Hillary*.

3 Partitions in their projective relation to reality

3.1 Projection

Partitions are more than just systems of cells. They are constructed to serve as pictures or maps of reality. In some cases the cells of a partition project but there are no objects for them to project onto. (Consider the partition cataloguing Aztec gods.) Here, however, we are interested primarily in partitions which do not project out into thin air in this way. We write 'P(z, o)' as an abbreviation for: cell z is projected onto object o. In what follows we shall assume that a unique projection is defined for each partition. In a more general theory we can allow projections to vary with time while the partition remains fixed (for example when we use a territorial grid of cells to map the presence of birds of given species in given areas from one moment to the next).

The theory of partitions allows us to employ a very general reading of the term 'object'. An object is any portion of reality: an individual, a part of an individual, a class of individuals (for example a biological species), a spatial region, a political unit (county, polling district, nation), or even (for present purposes) the universe as a

whole. An object in the partition-theoretic sense is everything onto which some cell of a partition can project.

Objects can be either of the bona fide or of the fiat sort (Smith 1995). Bona fide objects exist independently of human partitioning activity. They are, simply, recognized by partition cells. Fiat objects are objects created by human partitioning activity. Hence the corresponding partition cells do not only recognize their fiat objects – the latter are in fact created through the very projection of partition cells onto a corresponding portion of reality. Examples are the States of Wyoming and Montana.

3.2 Location

We are interested in partitions that succeed – which is to say, in partitions that are *true* of the corresponding portion of reality. That not all partitions are true follows from the fact that the relation $P(z, o)$ is marked by a certain taint of subjectivity and by a certain possibility of failure. The relation ‘ $L(o, z)$ ’ abbreviates: object o is located at cell z . $L(o, z)$ is free of taints of failure. It is determinately true, of every object o , and of every cell z , that o is either located, or not located, in z . When projection succeeds, then location is what results. Projection and location thus correspond to the two ‘directions of fit’ – from mind to world and from world to mind – between an assertion and the corresponding truthmaking portion of reality. (Searle 1983, Smith 1999)

Location presupposes projection: an object is never located in a cell unless through the projection relation associated with the relevant partition. Thus

$$\text{MB1} \quad L(z, o) \rightarrow P(o, z).$$

In the case where no errors have been made in the construction of a partition, $L(o, z)$ holds if and only if $P(z, o)$. This is because, in such a case, if a partition projects a given cell onto a given object, then that object is indeed located in the corresponding cell. Very many partitions – from automobile component catalogues to our maps of states and nations – have this quality without further ado, and it is such partitions upon which we shall concentrate in what follows. Such partitions are *transparent* to the corresponding portion of reality. In this case projection and location are converse relations with respect to the partition in question. Formally we write:

$$\text{DTr:} \quad \text{Tr}(A) \equiv \forall z \forall o: Z(z, A) \rightarrow (P(z, o) \rightarrow L(o, z))$$

We now impose a further master condition to the effect that:

$$\text{MB2:} \quad \text{All partitions are transparent in the sense of DTr.}$$

MB2 ensures that objects are actually located at the cells that project onto them. Notice however that a transparent partition, according to our definition, may still have empty cells. (Think of the Periodic Table, which leaves empty cells for chemical elements of types which have yet to be detected.) MB2 tells us only that, if a cell in a partition projects upon some object, then that object is indeed located in the corresponding cell. It does not tell us what happens in case a cell fails to project onto anything at all. DTr is correspondingly only a first step along the way towards a definition of *truth* for partitions.

Following Smith and Brogaard (2000) we define the notion of *recognition*, of an object by a cell, as follows. An object o is *recognized* by the cell z if and only if z is *projected* onto o and the object o is actually located at z . Formally we write:

DR1: $R(z, o) \equiv P(z, o) \text{ and } L(o, z)$

More generally, a partition recognizes an object if it has a cell that recognizes the object.

DR2: $R(A, o) \equiv \exists z: Z(z, A) \text{ and } R(z, o)$

The relation defined in DR2 is the partition-theoretic analogue of the standard set-membership relation. Since partitions are cognitive devices, and cognition is not omniscient, it follows that no partition is such as to recognize all objects. Partitions have limited domains. In particular, partitions have limitations on their domains which have to do with the phenomenon of granularity. This will be discussed in Section 3.4

3.3 Recognizing and preserving mereological structure

That upon which a partition is projected is a certain domain of objects in reality. We shall conceive the domain of a partition as a mereological sum. It is, as it were, the total mass of stuff upon which the partition sets to work: thus it is stuff prior to any of the divisions or demarcations effected by the partition itself. The domains of partitions will comprehend not only individual objects and their constituents (atoms, molecules, limbs, organs), but also groups or populations of individuals (for example biological species and genera, battalions and divisions, archipelagos and diasporas) and their constituent members. We will see below that spatial partitions, for example categorical coverages (Chrisman 1982), are one important family of partitions in our more general sense. We shall use the notation $D(o, A)$ in order to denote that the object o is a part of the domain of the partition A .

Partitions – think again of Venn diagrams – reflect the basic part-whole structure of reality through the fact that the cells in a partition are themselves such as to stand in the relation of part to whole. This means that, given the master conditions expressed within the framework of theory A above, partitions have at least the potential to reflect the mereological structure of the domain onto which they are projected. And in felicitous cases this potential is realized.

We say that the cells z_1 and z_2 *reflect* the mereological relationship between the objects onto which they are projected if and only if the following holds:

DR3: $RS(z_1, z_2) \equiv \forall o_1, o_2: (R(z_1, o_1) \text{ and } R(z_2, o_2)) \rightarrow (z_1 \subseteq z_2 \rightarrow o_1 \leq o_2)$.

This means that if z_1 is a subcell of z_2 then any object recognized by z_1 is a part of any object recognized by z_2 . A partition reflects the mereological structure of the domain it is projected onto if and only if each pair of cells recognizes in this way the mereological structure on the side of their objects:

DR4: $RS(A) \equiv \forall z_1, z_2: (Z(z_1, A) \text{ and } Z(z_2, A)) \rightarrow RS(z_1, z_2)$

We then impose a new master condition:

MB3 All partitions are *structure reflecting* in the sense of DR4.

What this means is that all partitions are such that if one cell is a subcell of another, then any object recognized by the first cell is a part of any object recognized by the second.

MB3 is still very weak. It does little more than ensure that partitions which satisfy it do not misrepresent the mereological relationships between their objects. But partitions might still be blind to (trace over) such relationships. In order to see

what this involves, we need to take a closer look at the relations \subseteq and \leq . Both represent *partial* rather than total orderings. This means that the axioms $\forall x, y: (x \leq y \text{ or } y < x)$ and $\forall z_1, z_2: (z_1 \subseteq z_2 \text{ or } z_2 \subseteq z_1)$ do not hold. There may be objects (or cells) that do not stand in the relations \leq (or \subseteq) to each other at all.

Objects that do not stand to each other in the part-whole relation are either disjoint or they are such as to overlap mereologically. On the partition side the interpretation of the absence of a part-whole relation between two cells is somewhat different. If z_1 and z_2 do not stand in the relation \subseteq to each other in a given partition, then this means ‘the partition does not know (or does not care) how z_1 and z_2 are related.’ We note that the minimal cells in a partition do not stand in the relation \subseteq to each other. From this we are entitled to infer nothing at all about the mereological relations among the corresponding objects.

Consider, for example, a partition that contains cells that recognize John and his arm, i.e., $L(\text{John}, z_1)$ and $L(\text{John's arm}, z_2)$. Then cell z_1 need not be a proper subcell of the cell z_2 . Partitions may trace over mereological relationships between the objects they recognize, but MB3 is strong enough to ensure that, if a partition tells us something about the mereological relationships on the side of the objects which it recognizes, then what it tells us is true.

Consider a domain consisting of two regions, x and y , that properly overlap, i.e., $x \wedge y = v$ and $x \wedge y \neq x$ and $x \wedge y \neq y$ where \wedge denotes mereological intersection in the realm of objects. Consider now a partition that recognizes x and y , i.e., $R(z_1, x)$ and $R(z_2, y)$. Assume further that z_1 and z_2 do not stand in any subcell relation to each other, i.e., their partition-theoretic intersection is empty. We can now distinguish two cases that still satisfy our transparency condition (DTr): (1) our partition does not recognize v ; (2) it recognizes v but traces over its mereological relationships to x and y . At the level of theory A we have explicitly excluded the possibility that cells which are not subcells of each other overlap (MA3). This reflects the fact that the tree structure of our partitions rules out cycles (diamonds) in their graph-theoretic representation. This condition is satisfied in case (1) as well as in case (2).

3.4 Granularity

The cells of a partition carry with them the feature of granularity. They recognize only single whole units. If a partition recognizes not only wholes but also one or more parts of such wholes, then this is because there are additional cells in the partition which do this recognizing job. Consider, for example, a partition that recognizes human beings, i.e., it has cells that project onto John, Mary, and so forth. This partition does not recognize parts of human beings – such as John’s arm or Mary’s shoulder – unless we add extra cells for this purpose. If a partition recognizes wholes and their parts, then it is not necessarily the case that it also reflects the mereological relationships between the two (as discussed above).

The theory of partitions inherits from mereology the feature that it is consistent with both an axiom to the effect that atoms exist and with the negation of this axiom. The theory thus enables us to remain neutral as to the existence of any ultimate simples in reality from out of which other objects would be constructed via summation. This is due to the fact that partitions are by definition *top-down* structures. The duality with trees puts special emphasis on this aspect: we trace

down from the root until we reach a leaf. A leaf need not necessarily be an atom in the sense that it projects upon something which has no further parts. The fact that there are leaves simply indicates that the partition does not care about what lies beneath a certain level of granularity. An object located at a minimal cell is an atom only relative to the partition involved.

Partitions are cognitive devices which have the built-in capability to recognize objects and to reflect certain features of the latter's mereological structure. They have two ways of tracing over or ignoring mereological structure: (1) tracing over mereological relations between the objects (wholes) which they recognize; (2) tracing over parts. (2) is (unless atomism is true) manifested by every partition, for partitions are in every case *coarse grained*.

4 Varieties of transparent partitions

In this section we discuss some of the more fundamental varieties of those partitions which satisfy the master conditions given above. We classify such partitions along three essentially orthogonal axes: (a) degree of correspondence to objects; (b) degree of structural fit; and (c) degree of completeness.

4.1 Functionality constraints

Partitions which possess the maximum degree of correspondence to objects must first of all be such as to rule out ambiguity on the side of their cell labels (or on the side of whatever it is in virtue of which projection is effected). This means that they must be such that their associated projection is a *functional* relation:

$$\text{CFP: } P(z_1, o_1) \text{ and } P(z_2, o_2) \rightarrow (o_1 = o_2 \rightarrow z_1 = z_2)$$

For partitions satisfying CFP, cells are projected onto single objects (one rather than two). We can then use a functional notation, $p(z)=o$, instead of $P(z, o)$, for projection. Notice, though, that p might still be a partial function, since CFP does not rule out the case where there are empty cells. An example of a partition not satisfying CFP is the partition created by a lazy schoolboy studying the history of the Civil War in England. This has one cell labeled 'Cromwell' – and so it does *not* distinguish between Oliver and his son Richard. Another example might be the partition utilized by those who talk of 'China' as if the Republic of China and the People's Republic of China were one object.

Consider a partition labeled 'heavenly bodies' and having just three minimal cells labeled 'The Morning Star' 'The Evening Star', and 'Venus', respectively. As we know, all three cells project onto the *same* object. Yet even so, it is still perfectly consistent with our definitions that this partition is true – that its distinct cells truly, though unknowingly, recognize the same object; for these cells are minimal, and thus neutral about the possible mereological relations obtaining on the side of that onto which they project. It is not unusual that we give different names (or coordinates, or class-labels) to things in cases where we do not know that they are actually the same. A good partition, though, should clearly be one in which such errors are avoided.

Partitions manifesting the highest degree of correspondence to objects must, in other words, be ones in which location is a *functional* relation:

$$\text{CFL: } L(o_1, z_1) \text{ and } L(o_2, z_2) \rightarrow (z_1 = z_2 \rightarrow o_1 = o_2)$$

In partitions that satisfy CFL, location is a (possibly partial) function, i.e., objects are located at single cells (one rather than two). The location function is partial, since no partition is omniscient. In transparent partitions satisfying CFP and CFL, projection and location are inverse relations wherever both are defined.

4.2 Structural constraints

We required of true partitions that they reflect the mereological structure of the domain they recognize. Remember that such reflection is to be understood in such a way that it leaves room for the possibility that a partition is merely neutral about (traces over) some aspects of the mereological structure of its target domain. Taking this into account, we can order partitions according to the degree to which they actually represent the mereological structure on the side of the objects onto which they are projected. At the one extreme we have: (1) partitions that completely reflect the mereological relations holding between the objects they recognize. At the other extreme are (2): partitions that completely trace over the mereological structure of the objects they recognize. Between these two extremes we have partitions that reflect some but not all of the mereological structure of the objects they recognize.

Under heading (1) are those true partitions which satisfy the weak converse of MB3, which means that if o_1 is part of o_2 , and if both o_1 and o_2 are recognized by the partition, then the cell at which o_1 is located is a subcell of the cell at which o_2 is located. Formally we can express this as follows:

$$\text{CS1 } o_1 \leq o_2 \text{ and } L(o_1, z_1) \text{ and } L(o_2, z_2) \rightarrow z_1 \subseteq z_2$$

We call partitions satisfying CS1 *mereologically monotonic*. An example of a mereologically monotone partition is the Dewey catalogue system used in libraries in order to organize the books in stock.

4.3 Completeness and exhaustiveness

So far we have allowed partitions to contain empty cells, i.e., cells that do not project onto any object. We now consider partitions which satisfy the constraint that every cell recognizes some object:

$$\text{CC: } Z(z, A) \rightarrow \exists o: R(z, o)$$

We say that partitions that satisfy CC *project completely*. Notice that this condition is independent of the functional or relational character of projection and location. Of particular interest, however, are partitions that project completely and in such a way that projection is a total function. These are partitions which satisfy both CFP and CC.

So far we have accepted that there are objects in our target domain that are not located at any cell. Those partitions are often not very satisfying: governments want *all* their subjects to be located in some cell of their partition of taxable individuals. They want their partitions to satisfy a completeness constraint to the effect that every object in the domain is indeed recognized. In this case we say location is complete. Alternatively we say that the partition *exhausts* its domain. Unfortunately we cannot use

$$(*) \quad D(o, A) \rightarrow \exists z: Z(z, A) \text{ and } R(z, o)$$

saying that if some object o is part of the domain of the partition A then there is a cell z in A that recognizes o to capture the desired constraint. This is because, for example, the tax authorities do not (as of this writing) want to tax the separate molecules of their subjects.

We believe that it will be necessary to promote several restricted forms of exhaustiveness, each one of which will approximate in different ways to the (unrealizable) condition of unrestricted exhaustiveness expressed in (*). To see how one such exhaustiveness condition might work in first (schematic) approximation, let us introduce a sortal predicate ϕ that singles out the kinds of objects our taxation partition is supposed to recognize (for example, human beings rather than parts of human beings). We now demand that the taxation partition recognize all of those objects in its domain which satisfy ϕ :

$$CE \quad D(o, A) \text{ and } \phi(o) \rightarrow \exists z: Z(z, A) \text{ and } R(z, o).$$

Think of CE as asserting the completeness of one partition *relative to* another, the ϕ -totalizer partition, which consists exclusively of minimal cells in which all and only the objects satisfying ϕ are located. Note that the idea underlying CE is closely related to the idea of granularity. An alternative means of formulating an exhaustiveness condition like CE, which constrains what can be located at minimal cells, is via a restriction on object size.

4.4 Redundancy

Partitions are natural cognitive devices and the designers and users of such devices build them in such a way that they will serve practical purposes. This means that they will normally avoid certain sorts of redundancy. Here we distinguish two sorts of redundancy: correspondence redundancy and structural redundancy.

Firstly, we have remarked already that necessarily empty cells (cells whose labels tell us *ex ante* that no objects can be located within them) are one type of redundancy of correspondence. Another such type would be involved if a partition were to contain two distinct cells whose labels tell us, again *ex ante*, that they must necessarily project upon the very same object. Thus, trivially, a partition should not contain two cells with identical labels. Partitions that avoid redundancy of correspondence satisfy:

$$CR1: \quad CFP \text{ and } CFL \text{ and } CC$$

The second sort of redundancy is not quite so trivial. Consider a partition with a cell labeled *vertebrates* which occurs as a subcell of the cell labeled *chordates* in our standard biological classification of the animal kingdom. Almost all chordates are in fact vertebrates. Suppose (for the sake of argument) that biologists were to discover that all chordates must be vertebrates. Then in order to avoid structural redundancy they would collapse into one cell the two cells of chordates and vertebrates which at present occupy distinct levels within their zoological partitions. A constraint designed to rule out such structural redundancy would be:

$$CR2: \quad \text{A cell in a partition never has exactly one immediate descendant.}$$

From this it follows that a one-celled partition is identical with its cell. (This rules out partition-theoretic analogues of the set-theorist's $\{\{a\}\}$.)

5 Partitions of geographic space

Partitions are, we repeat, natural cognitive devices. We assume that most of them are true in the sense that they are transparent and structure reflecting (they satisfy all of the master conditions MA1-4 and MB1-3 above). If we imagine the system of cells of a partition as being ranged over against a system of objects, with objects located in all the cells of the partition (under a certain relation of projection), then in the best case we have a partition that is mereologically monotone (CS1), projects completely (CC) and exhaustively (CE), and establishes a functional relationship to the domain it maps (CFP, CFL). Such ideal partitions are thereby also free of redundancy and hence CR1 and CR2 hold. We find examples of such perfection above all in the abstract, fiat domains of databases and spatial subdivisions.

In what follows we discuss cadastral maps, which come close to representing partitions which are perfect in the sense defined. We then move on to discuss categorical coverages which fall short of this sort of exact fit between partition and the corresponding objects in reality. Finally we discuss the 'folk' categorizations of geographic reality.

5.1 The perfect cadastre

The perfect cadastre is what exists in the databases of cadastral authorities. It is what you see when you examine cadastral maps. You see mathematically exact lines that separate land parcels. In the remainder we assume for the sake of simplicity that the cells on the map project onto corresponding parcels in reality (that the map contains no errors). We assume also that parcels are recognized by minimal cells of the cadastral partition. Partition cells are represented, for example, by entries in the German '*Grundbuch*' or in its computational equivalents. There are very strict rules for inserting, deleting, or changing cells in this partition, rules standardly expressed in the form of laws. These rules strive to guarantee that the cadastral partition has the ideal properties set forth above.

Land parcels are fiat objects. They are created through the projection of the cells in a cadastre onto reality itself. This is a *geodetic* projection of a sort which is described by a small number of axioms. It is mathematically well defined and can even (within certain limits) be computed. This projection imposes fiat boundaries onto reality in the same way that the plotter draws the lines on a cadastral map.

The projection (in our partition-theoretic sense) has the following properties: Cadastral partitions are true in that sense that cells correctly recognize objects, i.e., $P(z,o) \leftrightarrow L(o,z)$. Projection and location are functional relations, i.e., one cell projects onto one land parcel and one parcel is located at one cell. Cadastral partitions are CE-complete. The intuition underlying this thesis is that 'there are no no-mans-lands', which means: no zones within the domain of the cadastral partition that are assigned to no cell within the partition itself. Cadastres satisfy also CC-completeness, in that they satisfy a constraint to the effect that every piece of land is registered. These properties are ensured by law and extensive training of those who are charged with the task of maintaining the cadastre.

Cadastral partitions are (like most tax partitions) flat, in the sense that they recognize no mereological structure on the side of their objects. A tax partition recognizing families as well as individuals, and a cadastral partition recognizing multi-parcel estates as well as separate single parcels, may however embody some mereological structure. Consequently, cadastre partitions are not mereologically monotone, i.e., CS1 does not necessarily apply. They have, however, the property that they also recognize some mereo-topological structure in that sense that two cells are adjacent if and only if the corresponding land parcels are neighbors.

5.2 Categorical coverages

Area-class maps (W. Bunge 1966) or categorical coverages (Chrisman 1982) belong to a type of thematic maps that show the relationship of a property or attribute to a specific geographic area. A prototypical example of a categorical coverage is the land use map, in which a taxonomy of land use classes is determined (e.g., residential, commercial, industrial, transportation) and the specific area (zone) is then evaluated along the values of this taxonomy (Volta and Egenhofer 1993). Another prototypical example is soil maps, which are based on the classification of soil covering the surface of the earth (clay, slit, sand, ...). The zones of a categorical coverage are a jointly exhaustive and pair-wise disjoint subdivision of the relevant space (Beard 1988).

Using the notation introduced in this paper, there are two partitions involved in categorical coverages: the partition of an attribute domain (e.g., land use or soil) and the corresponding partition of the surface of the earth into zones. The projection of a partition onto the attribute domain creates a partition that is then in turn projected onto the surface of the earth. Both of the partitions involved satisfy all of the master conditions set forth above. The close relationship between the two partitions has been discussed for example by Beard (1988) and Frank *et al.* (1997). The same reciprocal relationship is illustrated in the way in which every map (a partition of space) stands to its *legend* (a partition of the attribute domain represented on the map).

To begin, consider the spatial component of a categorical coverage, i.e., the partition of some part of the surface of the earth. As already discussed in Smith and Brogaard (2000) spatial partitions enjoy peculiarly nice properties. Using the notions introduced in this paper we are now able to specify four of these properties more precisely:

Firstly, the partition is complete in that sense that there are no empty cells (CC). Secondly the minimal cells of the spatial partition *exhaust* a certain domain (a part of the surface of the earth) in the sense of CE. Consequently the root of the partition recognizes the mereological sum of all regions (zones) recognized by its cells. Thirdly, the correspondence between the cells in the partition of the spatial component of a categorical coverage and the zones it recognizes is one-one and onto. The fact that projection and location are here total, functional and mutually inverse is exploited extensively in the formalization and representation of categorical coverages (e.g. Frank *et al.* 1997, Erwig and Schneider 1999, Bittner and Stell 1998). Fourthly, as mentioned above, spatial partitions recognize some of the mereotopological structure of their domains even though they fall short of being mereologically monotone in the sense of CS1.

Due to these strong properties of their spatial component and the close relationship between the spatial and attribute components of categorical coverages, their partition of the pertinent attribute domain also satisfies the following rather strong constraints:

Firstly, the partition of the attribute domain is exhaustive relative to the spatial component: every minimal cell in the spatial partition has a corresponding minimal cell in the attribute partition. Consequently, the partition of the attribute domain exhausts the domain of all cases that actually occur in the region covered by the corresponding spatial partition. For example, if our spatial partition projects onto a desert, then the corresponding partition of soil types needs to be exhaustive for the different types of sand that occur in this area but it does not need to contain a cell labeled 'clay'. Secondly, projection and location need both to be functional, otherwise the regions carved out on the spatial side would not be jointly exhaustive and pairwise disjoint. Both functions may however be partial, as long as they are exhaustive relative to the pertinent spatial component.

Partitions of attribute domains are not necessarily limited to partitions consisting only of minimal cells (and one root cell). Consider a partition of the attribute domain Land-Use / Land-Coverage. There might be, for example, a non-minimal cell labeled 'Agricultural' in this partition with subcells 'Cultivated Cropland', 'Pasture', 'Livestock', and 'Poultry' (Frank *at al.* 1997). Hierarchical partitions of attribute domains are often created by refinement, i.e., we start with a root cell recognizing the attribute domain as a whole and add layers of subcells in such a way that the mereological sum of everything that is recognized by the cells of one layer is recognized also by the root cell. Consider for example a partition of the attribute domain 'Rainfall in inches'. There might be a layer of cells recognizing values falling within one or other of the three intervals $[0, 5)$, $[5, 10)$, $[10, \infty)$, together with a more refined layer recognizing values in: $[0, 2.5)$, $[2.5, 5)$, $[5, 7.5)$, $[7.5, 10)$, $[10, \infty)$.

Hierarchical partitions of the attribute domain create potentially hierarchical partitions of the spatial domain. Notice that the spatial component of hierarchical categorical coverages is not necessarily non-redundant in the sense of CR2. In the spatial component of a hierarchical categorical coverage 'Land Usage (Chicago)' there might be one single region that is recognized by both the cells 'Agricultural' and 'Cultivated Cropland'. In this case location is not a function. Technically the problem is dealt with by always using the most specific cell (the one farthest away from the root). The latter then exhaust space at every level of resolution. However, again, hierarchical spatial partitions are not mereologically monotone.

It is important to see that the regularity of these partition structures is due to the fact that the objects recognized by the given partitions are *fiat* objects carved out by the projecting partitions themselves. For example, in the categorical coverage for soil types there are certainly bona fide differences between sand and solid rock but the distinction between the many soil-types in between are certainly of the fiat sort. They are carved out or created by imposing a partition onto the attribute domain 'Soil on the surface of Earth'. This partition creates a spatial partition that, in being projected onto the surface of the earth, imposes fiat boundaries demarking 'categorical zones'. Those boundaries sometimes coincide with bona fide boundaries in reality but in most cases they do not.

5.3 A partition recognizing water bodies

We discussed spatial partitions or attribute partitions that induce spatial partitions. Those partitions are characterized by a high degree of structure and order not only due to the fact that they are spatial subdivisions but also due to the fact that there are well defined and strict rules (of scientific methodology or of law) which govern their construction and projection. In general partitions are much less well structured.

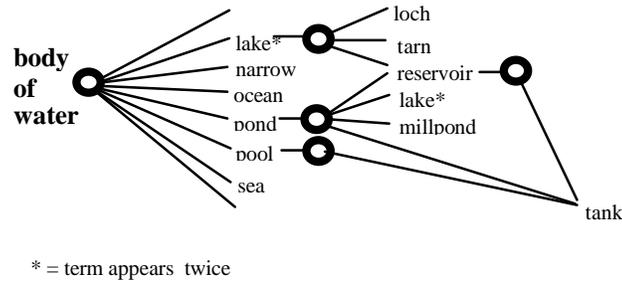


Figure 1: Ontology of Water Bodies and Related Entities, based on Definitions in the American Heritage Dictionary (taken from Smith and Mark 1999)

Smith and Mark 1999 analyzed the partition of water bodies and related entities which can be extracted from the definitions contained in the *American Heritage Dictionary*. The graph-theoretic representation of this partition is given in Figure 1. If we analyze this graph, then we can see easily that it is not a tree, since it contains cycles (e.g., pond, tank, reservoir, pond). We also can see that there are two cells labelled 'lake'. The latter clearly indicates that location is not a function relative to this partition.

We hypothesize that there are special features of dictionary definitions as hitherto compiled in virtue of which their underlying taxonomies appear to deviate from the tree structure. Moreover Guarino and Welty (2000) have shown how such taxonomies can very easily be reconstituted as trees in systematic fashion. This gives us some confidence that the ideas presented above may provide a framework for the construction of more coherent taxonomies for use in dictionaries and data standards in the future.

6 Summary and conclusions

This paper is a contribution to the formal ontology of partitions. We defined master conditions that need to be satisfied by every partition. These master conditions fall into two groups: (A) master conditions characterizing partitions as systems of cells, and (B) master conditions describing partitions in their projective relation to reality.

At the level of theory (A) partitions are systems of cells that are partially ordered by the partition-theoretic subcell relation. Such systems of cells are such that they can be always represented as trees, i.e., they are finite, have a unique maximal element, and they do not have cycles in their graph representation. But partitions are more than just systems of cells. They are cognitive devices that are directed towards reality. At the level of theory B we take this feature into account and characterize partitions using the relations of projection and location. Cells in partitions are

projected onto objects in reality. Objects are located at cells when projection succeeds. We then say that a partition *recognizes* the objects that are located at its cells.

Partitions do not only recognize objects, they are also capable of reflecting the mereological structure of the objects they recognize in the mereological structure of their cell array defined in theory A. This does not mean that all partitions actually do reflect the mereological structure of the objects they recognize, however, for it is an important feature of partitions that they are also capable of tracing over mereological structure. There are, for example, large classes of partitions that simply list objects without caring about how these objects hang together mereologically. In fact there are only very few partitions that completely represent the mereological structure of the objects they recognize.

Our discussion of granularity showed that partitions have two ways of tracing over mereological structure: (1) by tracing over mereological relations between the objects (wholes) which they recognize; (2) by tracing over parts. The tracing over of parts is (unless mereological atomism is true) a feature manifested by every partition, for partitions are in every case *coarse grained*.

We classified partitions along three essentially orthogonal axes: (a) degree of correspondence to objects; (b) degree of structural fit; and (c) degree of completeness and exhaustiveness. Along axis (a) we characterize partitions according to properties that the relations of projection and location have or lack: both may be functional relations, meaning that cells never project onto more than one object and that objects are never located at more than one cell. Along axis (b) we characterize partition according to the degree to which they reflect or trace over the mereological structure of the objects they recognize. We distinguished mereological monotony, mereological ignorance, and partitions that reflect some but not all mereological structure. Along axis (c) we characterized partitions by considering conditions under which every cell in a partition is projected onto some object and under which every object in a given domain is located at some cell. If projection or location relations are also functions, then we can recover in these terms the usual mathematical distinctions between partial, total, onto, and into mappings.

Three classes of partitions play an important role in spatial information science: cadastral partitions, categorical coverages, and partitions that we find in folk categorizations of geographic reality. We showed that partition theory can be used as a framework in whose terms the different properties of those partitions can be specified. Partition theory thus provides a formal-ontological foundation for understanding human partitioning activity in the geographic domain.

Acknowledgements

This work was supported in part by DARPA under the Command Post of the Future program and the National Science Foundation under the Research on Learning and Education program. Support from the American Philosophical Society, and from the NSF (Research Grant BCS-9975557: "Geographic Categories: An Ontological Investigation") is also gratefully acknowledged.

Bibliography

Beard, K. 1988 "Multiple representations from a detailed database: A scheme for automated generalization." Ph.D. thesis, University of Wisconsin, Madison

- Bittner, T. and Stell, J. G. 1998 "A Boundary-Sensitive Approach to Qualitative Location," *Annals of Mathematics and Artificial Intelligence*, 24, 93–114.
- Bunge, W. 1966 *Theoretical Geography*. Lund, Sweden.
- Casati, R. and Varzi, A.C. 1995 "The Structure of Spatial Location", *Philosophical Studies*, 82, 205-239
- Casati, R. and Varzi, A. C. 1999 *Parts and Places*, Cambridge, MA: MIT Press.
- Chrisman, N. 1982 "Models of Spatial Analysis Based on Error in Categorical Maps." Ph.D. thesis, University of Bristol, England
- Erwig, Martin and Schneider, Markus 1999 "The Honeycomb Model of Spatio-Temporal Partitions," *International Workshop on Spatio-Temporal Database Management* (Lecture Notes in Computer Science 1678), Berlin: Springer, 39-59.
- Frank, A., Volta, G., and McGranaghan, M. 1997 "Formalization of families of categorical coverages," *IJGIS* 11, 3, 214-231
- Guarino N. and Welty, C. 2000 "Ontological Analysis of Taxonomic Relationships," to appear in A. Laender and V. Storey, (eds.), *Proceedings of ER-2000: The 19th International Conference on Conceptual Modeling* (Lecture Notes in Computer Science), Berlin/New York: Springer-Verlag.
- Lewis, D. 1991 *Parts of Classes*, Oxford: Blackwell.
- Searle, J. R. 1983 *Intentionality. An Essay in the Philosophy of Mind*, Cambridge: Cambridge University Press.
- Smith, B. 1991 "Relevance, Relatedness and Restricted Set Theory", in G. Schurz and G. J. W. Dorn (eds.), *Advances in Scientific Philosophy. Essays in Honour of Paul Weingartner*, Amsterdam/Atlanta: Rodopi, 1991, 45–56.
- Smith, B. 1995 "On Drawing Lines on a Map", in Andrew U. Frank and Werner Kuhn (eds.), *Spatial Information Theory. A Theoretical Basis for GIS* (Lecture Notes in Computer Science 988), Berlin/Heidelberg/New York, etc.: Springer, 475–484.
- Smith, B. 1999 "Truthmaker Realism", *Australasian Journal of Philosophy*, 77 (3), 274–291.
- Smith, B. 2001 "True Grid", paper submitted to COSIT 2001.
- Smith, B. and Brogaard, B. 2000 "Quantum Mereotopology," in *Spatial and Temporal Granularity. Papers from the AAAI Workshop* (AAAI Technical Report WS-00-08), Menlo Park: AAAI Press, 2000, 25–31.
- Smith, B. and Brogaard, B. 2001 "A Unified Theory of Truth and Reference," *Logique et Analyse*.
- Smith, B. and Mark, D. M. 1999 "Ontology with Human Subjects Testing: An Empirical Investigation of Geographic Categories," *American Journal of Economics and Sociology*, 58: 2, 245–272.
- Volta, G. and Egenhofer, M. 1993 "Interaction with GIS Attribute Data Based on Categorical Coverages." in: Frank, A. and Campari, I. (eds.) *Conference on Spatial Information Theory*, Proceedings. (Lecture Notes in Computer Science, 716)