A taxonomy of granular partitions

Thomas Bittner and Barry Smith

Qualitative Reasoning Group, 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 granular 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 granular partitions at two levels: as systems of cells, and in terms of their projective relation to reality. We lay down conditions of well-formedness for granular partitions, and we define what it means for partitions to project transparently onto reality in such a way as to be structure-preserving. We continue by classifying granular partitions along three axes, according to: (a) the degree to which a partition represents the mereological structure of the domain it is projected onto; (b) the degree of completeness and exhaustiveness with which a partition represents reality; and (c) the degree of redundancy in the partition structure. This classification is used to characterize three types of granular 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 (a) a geologist classifying soil samples or (b) a spatial analyst classifying the raster pixels of a digital image or (c) a hotel manager making a list 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. In case (a) the cells are labeled, for example, ‘clay’ or ‘sand’ and the objects you are recognizing as located in these cells are your soil samples. In case (b) the cells are labeled with the names of vegetation classes, each class being made to correspond to a particular spectrum of frequencies in the pixel array, and the objects that are located within those cells are raster cells within the partition which is the pixel image. In case (c) the cells correspond to the rooms in your hotel, the objects are the individuals or groups who are, according to the hotel register, assigned to these rooms on any given night.

We shall call a grid of cells of the type used in these examples a *granular partition*, and we shall argue that granular partitions are involved in all listing, sorting, cataloguing and mapping activities.

Granular partitions are ways of structuring reality in order to make it more easily graspable by cognitive subjects such as ourselves. Some partitions are flat: they amount to nothing more than a mere list (case c). 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 and subspecies, the subdivision of heavenly bodies into galaxies, stars, planets, moons, etc.). Other partitions – for example the partitions created by 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. Some partitions involve the imposition of a layer of quasi-discreteness upon an underlying reality which in itself has the structure of a continuum.

In Smith and Brogaard (2000) the notion of granular partition was 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 granular 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 and cataloguing are activities performed by human beings in their traffic with the world. Granular 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 granular 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 granular partitions as systems of cells in the sense of theory (A). We then consider granular partitions in their projective relation to reality in the sense of theory (B). This provides us with the tools to define what it means to say that a granular partition projects onto reality in a transparent and structure-preserving way. We then provide a classification of granular partition by characterizing various properties of the correspondence between partition and reality, and we go on from there to discuss relationships between set theory, mereology, and the theory of granular partitions as alternative tools for the purposes of formal ontology. We conclude by considering three classes of partitions that have an important role to play in the geographic domain.

2 Granular partitions as system of cells

2.1 Cells and subcells

All granular partitions involve cells arranged together in some sort of structure. This structure is intrinsic to the partition itself, and obtains independently of whether there are objects located in its cells. Cells in granular partitions may be nested one inside another in the way in which species are nested within genera in standard biological taxonomies. Theory (A) studies properties granular partitions have in virtue of the relations between and the operations which can be performed upon the cells from out of which they are built. We say that one cell, z_1, is a subcell of another cell, 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, and we postulate as an axiom or master condition:

MA1: The subcell relation $\subseteq$ is reflexive, transitive, and antisymmetric.

Every granular partition A (‘partition’ is ‘*Aufteilung*’ in German) has a maximal cell defined as:

DMax: $Max(z_1, A) \equiv Z(z_1, A)$ 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 the condition Z(z, A) will be omitted in cases where it is clear that we are talking about cells within some fixed partition A. In addition, initial universal quantifiers will be taken as understood.) We now demand that

MA2: $\exists z: Max(z, A)$

which ensures that every granular partition has a maximal cell in the sense of DMax. From the antisymmetry of the subcell relation it follows that this cell is unique. This root cell, denoted r(A), is such that all the cells in the partition are included in it as subcells.

The nestedness of cells inside a partition yields chains of cells satisfying $z1 \supseteq z2 \supseteq \ldots \supseteq zn$. We shall call the cells at the ends of such chains minimal cells, and define:

DMin: $Min(z_1, A) \equiv Z(z_1, A)$ and $\forall z: Z(z, A) \rightarrow (z \subseteq z_1 \rightarrow z = z_1)$

Another important aspect of granular partitions is then:

MA3: Each cell in a partition is connected to the root by a finite chain.

MA3 leaves open the issue as to whether granular 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.

2.2 Partition-theoretic sum and product of cells

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 disjointness. In other words:

MA4: Two cells overlap only if one is a subcell of the other.

Or in symbols:

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

From MA3 and MA4 we can prove by a simple reductio that the chain connecting each cell of a partition to the root is unique.

Following Smith (1991) we can define the partition-theoretic sum and product of cells within granular partitions as follows. The partition-theoretic sum $z = z_1 \cup z_2$ of two cells in a partition is the $\subseteq$-minimal cell satisfying $z_1 \subseteq z$ and $z_2 \subseteq z$. 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 the largest subcell shared in common by z_1 and z_2.

2.3 Trees

Philosophers since Aristotle have recognized that the results of our sorting and classifying activities can be represented as those sorts of branching structures which mathematicians nowadays called trees. Trees are rooted directed graphs without cycles (Wilson and Watkins 1990). Every finite partition can be represented very simply as a rooted tree in such a way that the cells in the partition correspond to vertices in the tree and vertices are connected by an edge if and only if the corresponding cells stand to each other in an immediate subcell relation.

We can represent a partition not only as a tree but also as a simple sort of Venn diagram. In a Venn diagram partition cells are represented as topologically simple and regular regions of the plane. Our partitions are Venn diagrams within which regions do not intersect. In the remainder we will often think of partitions as such planar maps – they are Venn diagrams without overlapping – and the minimal cells correspond to the smallest regions within such diagrams.

3 Granular partitions in their projective relation to reality

3.1 Projection

Granular partitions are more than just systems of cells. They are built to serve as pictures or maps of reality. Granular partitions are systems of cells that project onto reality in something like the way in which a bank of flashlights projects onto reality when it carves out cones of light in the darkness. 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 granular 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 granular partition. For a more general discussion see (Bittner and Smith 2001).

The theory of granular partitions allows us to employ a very general reading of the term 'object'. An object in the partition-theoretic sense is everything onto which some cell of a partition can project: an individual, a part of an individual, a group or 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.

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 (highlighted) by partition cells. Fiat objects are objects created by our human partitioning activity. Hence it may be that the corresponding partition cells not only recognize their fiat objects but that the latter are in fact created through the very projection of partition cells onto the corresponding portion of reality. Examples are the States of Wyoming and Montana. For an extended discussion of the relationships between granular partitions and fiat objects see (Bittner and Smith 2001).

3.2 Location

When projection succeeds then the corresponding granular partition *represents* the corresponding portion of reality transparently and in such a way that mereological structure is preserved.

We write 'L(o, z)' as an abbreviation for: object o is located at cell 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

MB1 $L(o, z) \rightarrow P(z, o)$.

In the case where no errors have been made in the construction and the projection of a granular 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.

MB2 $P(z, o) \rightarrow L(o, z)$.

Very many granular partitions – from automobile component catalogues to our maps of states and nations – have this quality without further ado, and it is such granular partitions upon which we shall concentrate in what follows. Such granular 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:

DTr: $Tr(A) \equiv \forall z \forall o: P_A(z, o) \leftrightarrow L_A(o, z)$.

MB1 and MB2 jointly ensure 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.) MB1 and MB2 tell us only that, if a cell in a partition projects upon some object, then that object is indeed located in the corresponding cell. They do not tell us what happens in case a cell fails to project onto anything at all.

An object o is *recognized* by a cell z if and only if z is projected onto o and the object o is actually located at z. A partition recognizes a given object if and only if it has a cell that recognizes that object (Smith and Brogaard 2001). We shall sometimes use the term ‘recognition’ as a synonym for ‘transparent projection’ in what follows.

4 Functionality constraints

4.1 Projection is functional: The confused schoolboy

The notion of transparency is still very weak. Thus it is consistent with ambiguity on the side of the cells in relation to the objects they target, that is with the case where one cell projects onto two distinct objects. Consider the partition created by a lazy schoolboy studying the history of the Civil War in England, which has just one cell labeled ‘Cromwell’. Thus it does not distinguish between Oliver, the Lord Protector, and his son Richard. Or consider 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.

To eliminate such ambiguity we lay down a requirement to the effect that each partition must be such that its associated projection is a *functional* relation:

MB3: $P(z, o_1)$ and $P(z, o_2) \rightarrow o_1 = o_2$

For granular partitions satisfying MB3, cells are projected onto single objects (one rather than two).

4.2 Location is functional: The Morning Star and the Evening Star

Consider a partition having root cell labeled 'heavenly bodies' and three subcells labeled: 'The Morning Star', 'The Evening Star', and 'Venus', respectively. As we know, all three subcells project onto the same object. This partition is clearly somewhat barren; but it is still perfectly consistent with the conditions we have laid out thus far. Its distinct subcells truly, though unknowingly, project onto the same object. It is not unusual that we give different names (or coordinates, or class-labels) to objects 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.

Granular partitions manifesting the desired degree of correspondence to objects in this respect must accordingly be ones in which location, too, is a *functional* relation:

MB4: $L(o, z_1)$ and $L(o, z_2) \rightarrow z_1 = z_2$

In granular partitions that satisfy MB4, location is a function, i.e., objects are located at single cells (one rather than two). The location function is however partial, since partitions are not omniscient. As MB3 rules out co-location (overcrowding), so MB4 rules out co-projection (redundancy).

5. Structural mapping

MB1 and MB2 are, even when taken together with MB3 and MB4, still very weak. They thus represent only a first step along the way towards an account of correspondence to reality for granular partitions. Such correspondence will involve two further dimensions: of *structural mapping*, and of *completeness*. In the present section we address our attention to the topic of structural mapping.

5.1 Recognizing mereological structure

Each granular partition reflects the basic part-whole structure of reality through the fact that its cells 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, granular partitions have at least the potential to reflect the mereological structure of the relevant domain. 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:

DS1: $RS(z_1, z_2) \equiv L(o_1, z_1)$ and $L(o_2, z_2)$ and $z_1 \subset z_2 \rightarrow o_1 < o_2$

If z_1 is a proper subcell of z_2 then any object recognized by z_1 must be a proper 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 satisfies DS1:

DS2: $RS(A) \equiv \forall z_1,z_2$: $(Z(z_1, A)$ and $Z(z_2, A)) \rightarrow RS(z_1, z_2)$

We then impose a new master condition:

MB5: All granular partitions are *structure reflecting* in the sense of DS2.

Note that even MB5 is still very weak. Its effect is in a sense entirely negative: it merely ensures that granular partitions do not misrepresent the mereological relationships between their objects. But granular partitions might still be blind to (trace over) such relationships. *Minimal* cells might project onto objects which stand to each other in any one of the entire range of possible mereological relations (parthood, proper parthood, disjointness, and overlap). Pairs of cells z_1 and z_2 which do not stand to each other in the subcell relation are likewise neutral as to the mereological relations between their objects. This means that the corresponding partition does not know (or does not care) how o_1 and o_2 are related, which means that 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(John, z_1) and L(John's arm, z_2). Cell z_1 need not be a proper subcell of the cell z_2, because granular partitions may trace over mereological relationships between the objects they recognize. MB3 is however still 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.

5.2 The domain of a partition

That upon which a partition is projected, its domain, is a certain mereological sum of objects in reality. It is, as it were, the total mass of stuff upon which the partition sets to work: thus it is stuff conceived as it is prior to any of the divisions or demarcations effected by the partition itself. The domains of granular partitions might 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 parts or members. Granular partitions can be used to impose a division into discrete units upon continuous domains, for example through temperature or frequency bands. We shall see that maps of land use or soil type are another important family of granular partitions in the sense here advanced.

Formally we define the *domain* of a partition simply as the object onto which its root cell is projected:

DD $\quad D(A) = p(r(A))$

MB1–5 already ensure (a) that everything that is located at some cell of the partition is part of what is located at the corresponding root cell; and (b) that for each partition there can be only one such object. We now demand that every partition has a non-empty domain:

MB6 $\quad \exists x: x = D(A)$

We then say that a partition *represents its domain correctly* if and only if MA1–4 and MB1–6 hold.

5.3 Granularity

A granular partition is granular in virtue of the fact that it can recognize an object without recognizing all its parts. The theory of granular partitions can thus provide the basis for understanding the selective focus of our maps and classifications and above all their ability to trace over parts below a certain level. To impose a partition on a given domain of reality is to *foreground* certain objects and features in that domain and to trace over others.

Partition cells always project onto wholes. 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 the molecules in Mary's shoulder – unless we add extra cells for this purpose. And even if a partition recognizes wholes and their parts, then as we saw above it is not necessary that it also reflects the mereological relationships between the two.

The theory of granular 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 granular 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 in reality which has no further parts. The fact that there are leaves simply indicates that our partition does not care about what, on the side of reality, lies beneath a certain level of granularity. An object located at a minimal cell is an atom only relative to the partition which we bring to bear.

6 Varieties of granular partitions

In this section we discuss some of the more fundamental varieties of those granular partitions which satisfy the master conditions (MA1-4 and MB1-6) given above. We classify them according to: (1) degree of structural fit; (2) degree of completeness and exhaustiveness of projection; and (3) degree of redundancy.

6.1 Structural constraints

We required of granular 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 granular partitions according to the degree to which they do indeed succeed in representing the mereological structure on the side of the objects onto which they are projected. At the one extreme we have (1): granular partitions that completely reflect the mereological relations holding between the objects they recognize. At the other extreme are (2): granular partitions that completely trace over the mereological structure of the objects they recognize (except to the degree that they recognize them as part of the domain in question). Between these two extremes we have granular partitions that reflect some but not all of the mereological structure of the objects they recognize.

Under heading (1) are those granular partitions which satisfy the weak converse of MB5, 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:

$$\text{CM:} \quad L(o_1, z_1) \text{ and } L(o_2, z_2) \text{ and } o_1 < o_2 \rightarrow z_1 \subset z_2$$

We call granular partitions satisfying CM *mereologically monotonic*.

6.2 Projective completeness

So far we have allowed granular 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:

CC: $Z(z, A) \rightarrow \exists o: L(o, z)$

We say that granular partitions that satisfy CC *project completely*. These partitions are of particular interest since in this case projection is a total function.

6.3 Exhaustiveness

There may be objects in our target domain that are not located at any cell. The resulting granular partitions are not very satisfying: governments want all their subjects to be located in some cell of their partition of taxable individuals. They want their partition to satisfy an exhaustiveness constraint to the effect that every object in the pertinent domain is indeed recognized. But what does it mean to say that a partition exhausts its domain? Unfortunately we cannot capture this notion formally by using

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

which asserts 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. The tax authorities (as of this writing) do not want to tax the separate molecules of their subjects.

To formulate an acceptable alternative to (*) will be a difficult matter. In fact 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 can now demand that the taxation partition recognize all of those objects in its domain which satisfy φ:

CE_φ $o \leq D(A) \text{ and } \varphi(o) \rightarrow \exists z: Z(z, A) \text{ and } L(o, z)$.

Think of CE_φ 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. We will discuss examples of other such conditions in section 8.

6.4 Redundancy

Granular partitions are natural cognitive devices and the designers and users of such devices build them to serve practical purposes. This means that they will normally strive to avoid certain sorts of redundancy. One sort of redundancy – which we might call correspondence redundancy – is excluded already by condition CC. This consists in the presence of necessarily empty cells (cells whose labels tell us *ex ante* that no objects can be located within them).

But partitions can manifest also what we might call structural redundancy, and this 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:

CR: A cell in a partition never has exactly one immediate descendant.

7 Set theory, mereology, and granular partitions

7.1 Partition theory as an alternative to set theory and mereology

The theory of granular partitions is intended to serve, first of all, as an alternative to set theory both as a tool of formal ontology and as a framework for the representation of human common sense. Currently it is the naïve portion of set theory that is used in almost all work on common-sense reasoning and in related investigations of natural language semantics. Kinds, sorts, species are standardly treated as sets of their instances; subkinds as subsets of these sets. Set theory nicely does justice to the granularity that is involved in our sorting and classification of reality by treating objects as *elements* of sets, i.e. as single whole units within which further parts are not recognized.

But set theory also has its problems, not the least of which is that it supports no distinction between natural granular totalities (such as the species *cat*) and such *ad hoc* totalities as for example: {the moon, Napoleon, justice}. Set theory also has problems when it comes to dealing with the fact that biological species and similar entities may remain the same even when there is a turnover in their instances. For sets are identical if and only if they have the same members. If we model the species *cat* as the set of its instances, then this means that cats form a different species every time a new cat is born or dies. If, similarly, we model an organism as the set of its cells, then this means that it becomes a different organism whenever cells are gained or lost.

Set theory also has problems when it comes to dealing with relations between objects at different granularities. An organism is a totality of cells, but it is also a totality of molecules, and it is also a totality of atoms. Yet the corresponding *sets* are distinct, since they have entirely distinct members.

More recently, attempts have been made to solve some of these problems by using mereology as a framework for ontological theorizing. Mereology is better able to do justice in realistic fashion to the relations between wholes and their constituent parts at distinct levels of generality. All the above-mentioned totalities (of cells, molecules, atoms) can be recognized, when treated mereologically, as being one and the same. Mereology has one further advantage over set theory as a tool for the sort of middle-level ontological theorizing which the study of common-sense reasoning requires, namely that it does us not require that, in order to quantify over wholes of given sorts, one must first of all explicitly specify all the parts.

On the other hand, however, mereology, too, has its problems. Above all it does not have the machinery for coping with the phenomenon of granularity; for if we quantify over wholes in a mereological framework, then we thereby quantify over all the parts of such wholes, both known and unknown, at all levels of granularity. Mereology can mimic the advantages of set-theory in this respect only if we depart from realism and make the idealizing commitment to atomism. (Galton 1999) Set theory and mereology are then in practice indistinguishable, since each whole becomes isomorphic to a certain set of atoms.

The theory of granular partitions presented in this paper is the product of an effort to build a more realistic, and also a more general and flexible, framework embodying some of the strengths of both set theory and mereology while at the same time avoiding their respective weaknesses. At the formal level it assumes standard extensional mereology (Simons 1987) and adds the primitives and axioms of theories (A) and (B). It thereby avoids the disadvantages of the unrestricted part-of relation via the intermediate formal machinery of cells, which adds to mereology the features of selectivity and granularity.

7.2 Partition theory and set theory

Partition theory, as already noted, is a generalization of set theory understood in Lewis's sense. At the formal level there are some obvious similarities between sets and granular partitions: (a) the subcell relation and the subset relation are both partial orders (MA1); (b) the minimal chain condition (MA2) is the analogue of the set-theoretic *Begründungsaxiom*; (c) the existence of a root cell of which all subcells are parts corresponds to the conception of sets as containers; (d) the transparency and functionality of projection and location (MB1-4) reflect analogous features of the element-of relation.

At the same time there are a number of important differences between the two frameworks. Above all partition theory is designed to do justice to the fact that not all members of the powerset of a set are of interest in the sorts of natural contexts in which sorting and classifying occur. Partitions are cognitive artifacts. They comprehend only those subcell-cell relations which reflect some sort of natural inclusion relation – for example between a species and its genus – on the side of objects in the world. Some sets then have a structure which precludes them from being even considered as partitions in the sense defended here. Consider, for example, the set {{a, b}, {a, c}}. Since we have $\{a\} \subseteq \{a, b\}$ and $\{a\} \subseteq \{a, c\}$, any corresponding partition would violate MA4, the condition designed to exclude double counting.

8 Granular partitions of geographic space

Granular partitions are, we repeat, natural cognitive devices. We assume that the primary examples of partitions are transparent and structure reflecting (they satisfy all of the master conditions MA1–4 and MB1–6 above). If we imagine the system of cells of a partition as being ranged over against a system of objects, with all the cells of the partition being occupied by objects (under a certain relation of projection), then in the best case we have a partition that is mereologically monotone (CM) and such as to project completely (CC) and exhaustively (CE_φ) relative to some condition φ. Such ideal granular partitions are thereby also free of redundancy (CR). 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 granular 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.

8.1 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. We are here assuming 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 all and only parcels are recognized by the 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, by means of which we seek to guarantee that the cadastral partition has the ideal properties set forth above.

Land parcels are fiat objects. They are created (in no small part) through the very projection of the cells in the 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 transparent in the 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_φ-complete, where φ selects minimal cells that recognize pieces of land that are parcels. (Defining φ is a complicated matter of law, and currently there exist only informal definitions.) 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 do not contain empty cells, i.e., cadastral entries that do not correspond to any piece of land. These properties are (in the cases of interest to us here) ensured by law and by extensive training on the part of those who are charged with the task of maintaining the cadastre.

Cadastral partitions may recognize some mereological structure on the side of their objects. For example, a cadastral partition may recognize multi-parcel estates as well as separate single parcels. If a cadastral partition properly recognizes all the pertinent multi-parcel estates then it is mereologically monotone, i.e., CM holds. Cadastral partitions have the property that they recognize, too, some of the mereotopological structure on the side of their objects, in the sense that two cells are adjacent in the cadastre if and only if the corresponding land parcels are neighbors on the ground.

8.2 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 a 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 a classification of the soil covering the surface of the earth (into *clay*, *silt*, *sand*, etc.). The zones of a categorical coverage are a jointly exhaustive and pair-wise disjoint subdivision of the relevant region of space (Beard 1988).

There are in fact two reciprocally dependent granular partitions involved in categorical coverages. On the one hand is the partition of the attribute domain (e.g., of land use or of soil types); we can think of the attribute domain with which we start as a continuum, which is then partitioned into discrete bands in light of our practical purposes, capabilities of measurement, and so forth. On the other hand is the partition of the surface of the earth into corresponding zones. Both of these partitions satisfy all of the master conditions set forth above. The close relationship between the two 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 categorical map (a partition of space) stands to its *legend* (a partition of the attribute domain represented on the map).

Consider, first of all, the spatial component of a categorical coverage, which is a partition of some portion of the surface of the earth. Using the notions introduced in the foregoing we are now able to specify four properties of this partition more precisely as follows:

First, the partition is complete in the sense that there are no empty cells (CC). Secondly, the minimal cells of the partition *exhaust* a certain domain (a part of the surface of the earth) in the sense of CE_{φ}, where φ selects topologically simple and maximal regions that are of one or other of the soil types recognized by the partition of the attribute domain. Consequently the root of the partition recognizes the mereological sum of all the 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 coverges (e.g. Frank *et al.* 1997, Bittner and Stell 1998, Erwig and Schneider 1999). Fourthly, as in the case of cadastral maps, spatial partitions recognize the mereotopological structure of their domains in the sense that they are not only mereologically monotone in the sense of CM but also such that two cells in the spatial partition are adjacent if and only if the corresponding parcels are neighbors on the ground. This is the case because the geodetic transformations used to map features on the surface of the Earth onto planar maps preserve topological relations (assuming perfect transformations without error and modulo the feature of limited resolution). This implies that the part-of relation is also preserved by the given mappings. Spatial partitions can be considered as Venn diagrams and hence they can be transformed into a partition structure where the part-of relation becomes the subcell relation along the lines described above.

These properties of their spatial component and the close relationship between the spatial and attribute components of categorical coverages mean that the partition of the pertinent attribute domain also satisfies the following nice constraints:

First, it is exhaustive *relative* to the spatial component. Every minimal cell in the spatial partition (a topologically simple zone of homogeneous coverage) has a corresponding minimal cell in the attribute partition. This immediately follows from the definition of the selection predicate φ for minimal cells of the spatial component. 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 and which we find it important to distinguish, 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. The location function is partial if there exist soil types that are not recognized by the attribute partition and the projection is partial if there are empty cells in the attribute partition.

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 labeled *cultivated cropland*, *pasture*, *livestock*, and *poultry*. 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, ∞), together with more refined layers recognizing values in: [0, 2.5], [2.5, 5], [5, 7.5], and so forth.

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 CR. 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' where the second is a subcell of the first. In this case location is not a function since the region in question is located within both cells. Technically the problem is dealt with by letting only the most specific cell (the one farthest away from the root) project onto the region in question.

It is important to see that the regularity of the given partition structures is due to the fact that the objects recognized 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 of the fiat sort. They are created by imposing a partition onto the attribute domain 'soil on the surface of Earth'. (Smith and Mark 1999) This partition, on being projected, then creates as its target a spatial partition whose cells are separated by spatial fiat boundaries on the ground. The latter demarcate 'categorical zones', which are homogeneous at the level of granularity determined by the map. The given boundaries sometimes coincide with bona fide boundaries in reality, but in most cases they do not do so.

8.3 A folk categorization of water bodies

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

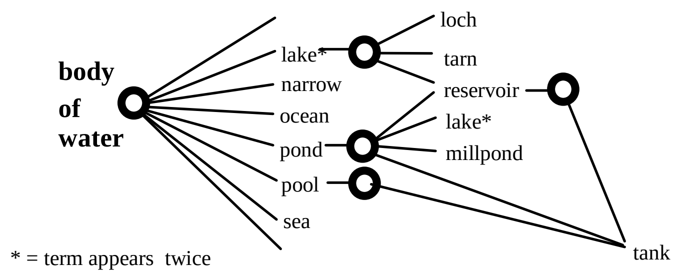

Figure 1: Ontology of Water Bodies and Related Entities, based on Definitions in the American Heritage Dictionary (taken from Smith and Mark 1999)

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 labeled 'lake'. The latter clearly indicates that location is not a function relative to this partition.

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

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.

Bittner, T. and B. Smith 2001 “Vagueness and Granular Partitions.” to appear in C. Welty and B. Smith (eds.), *Proceedings of FOIS-2001: The 2nd International Conference on Formal Ontology in Information Systems,* Sheridan Press.

Bunge, W. 1966 *Theoretical Geography*. Lund: Gleerup.

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,” *International Journal of Geographic Information Science*, 11: 3, 214-231

Galton, A. C. 1999 "The Mereotopology of Discrete Space", in C. Freksa and D. M. Mark (eds.) *Spatial Information Theory: Cognitive and Computational Foundations of Geographic Science* (Lecture Notes in Computer Science 1661), Berlin/New York: Springer, 251–266.

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.

Simons, P. M. 1987 *Parts: A Study in Ontology*. Oxford: Clarendon 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”, in this volume.

Smith, B. and Brogaard, B. 2000 “Quantum Mereotopology,” forthcoming in *Annals of Mathematics and Artificial Intelligence.*

Smith, B. and Brogaard, B. 2001 “A Unified Theory of Truth and Reference,” *Logique et Analyse*, in press.

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).

Wilson, R. J. and J. J. Watkins (1990). *Graphs – An Introductory Approach*. New York, John Willey and Sons, Inc.