Ontological Tools for Geographic Representation
Roberto Casati, Barry Smith, Achille Varzi

To cite this version:
Ontological Tools for Geographic Representation

Roberto Casati  
_CNRS, Aix-en-Provence (France)_

Barry Smith  
_Department of Philosophy and Center for Cognitive Science, SUNY Buffalo, NY_

Achille C. Varzi  
_Department of Philosophy, Columbia University, New York, NY_


**Abstract.** This paper is concerned with certain ontological issues in the foundations of geographic representation. It sets out what these basic issues are, describes the tools needed to deal with them, and draws some implications for a general theory of spatial representation. Our approach has ramifications in the domains of mereology, topology, and the theory of location, and the question of the interaction of these three domains within a unified spatial representation theory is addressed. In the final part we also consider the idea of non-standard geographies, which may be associated with geography under a classical conception in the same sense in which non-standard logics are associated with classical logic.

1. Introduction

This paper is a contribution to the ontology of geography and to the theory of spatial representation with special reference to spatial phenomena on the geographic scale. Geography presents an interesting trade-off between empirical issues on the one hand, and ontological issues on the other. What is a geographic entity? What is the relationship between a geographic entity and a physical territory? Can a geographic entity survive without a territory or without definite borders? Can it survive radical changes in its territory? Are there clear-cut identity criteria for geographic entities? To be sure, not everything is settled once clear definitions of ontological concepts such as national borders and national identity are provided. But further questions can-
not even be addressed without agreement as to the meanings of the fundamental terms at issue.

2. Basic Issues

We may classify the basic issues of a theory of spatial representation, broadly understood, into four main categories (with no pretension to completeness; see [8, 20, 26] for further material).

2.1 Regions of space and things in space. Common sense distinguishes not only objects and events, but also regions of space in which object are located and events occur. Must all of these types of entities (and the corresponding types of relations) be included as primitives within the domain of the theory? [3, 9] Or can we do without objects and events by conceiving them in terms of predicates assigned to corresponding regions? Can we do without regions by conceiving space, in Leibnizian fashion, as a merely relational order?

2.2 Absolutist vs relational theories of space. This last option has been the focus of an intense debate in philosophy [15], and has immediate consequences for a representational theory. It can be raised not only in relation to each single region of space, but also to space in its entirety. Does space exist as an independent individual (a sort of container) over and above all objects and spatial relations between objects, or are those relations the only facts of the matter concerning space?

2.3 Types of spatial entities. Different types of entities bear different types of relations to space [10]. For instance, material objects occupy the space in which they are located (they cannot share it with other material objects, as a car cannot share a parking space with another car); but immaterial objects—such as holes or shadows, and also processes and events—are less exclusive and can share locations with their peers [7, 11]. A complete descriptive account should address these differences.

2.4 Boundaries and vagueness. Possession of a boundary is one mark of individuality. A boundary separates an entity from its environment, and in the geographic domain the existence of a (complete) boundary is the first
criterion for the individuation of an autonomous entity (e.g., a politically independent unit). However, boundaries give rise to a number of ontological conundrums and may themselves be difficult to individuate [23, 27, 32]. Moreover, natural boundaries are often fuzzy or otherwise indeterminate, and an adequate theory must take this into account. It must also address the relations between boundaries in nature and those boundaries which exist as the products of legislation or administrative fiat. [5]

3. Geographic Ontology

Over and above the general issues listed above, there are some ontological problems that arise specifically in the geographic domain (see [6] for some relevant background).

3.1 Geographic objects. A general question concerns the nature of the entities geographers deal with [13, 19]. Common sense recognizes as its prototypes entities such as material objects, artifacts and people. The entities to which geographers refer—nations, neighborhoods, deltas, deserts—are of a different kind. The basic metaphysical question concerns the status of these entities. Are there geographic things? What kinds of geographic things are there? Two categories can be distinguished, corresponding to a traditional distinction between physical and human geography. On the one hand there are mountains, rivers, deserts. How are such entities individuated from each other? (Here issues of vagueness arise.) On the other hand there are socio-economic units: nations, cities, real-estate subdivisions—the spatial shadows cast by different sorts of systematically organized human activity [1, 2].

One extreme position on the existence of geographic objects would be strong methodological individualism: there are, on one version of this view, only people and the tables and chairs they interact with on the mesoscopic level, and no units on the geographic scale at all. At the opposite extreme is geographic realism: socio-economic units and other geographic entities exist over and above the individuals that they appear to be related to and have the same ontological standing as these. A more reasonable position is one or other form of weak methodological individualism: if geographic units exist as such, then they depend upon or are supervenient upon individuals. One
form of this position would accept both individuals and the behavioral settings in which individuals act. Larger-scale socio-economic units would then be accounted for in terms of various kinds of connections between such behavioral settings, illustrated for example by the command hierarchy of an army [16, 24].

More detailed questions here are: What processes produce socio-economic units? What ontology should we use for them? (What is a town square? What is a neighborhood?) How do they begin to exist? How do they evolve through time? How do they cease to exist? Can socio-economic units move? Can they have intermittent existence? Can they be resuscitated? Under what conditions can they merge or split? With what sorts of regions can socio-economic units be associated? With regions of any dimension $n$? With scattered regions?

3.2. Borders. A central notion in geographic representation is that of a border between two adjacent nations or counties or parcels of real estate. We may distinguish here two different types of borders or boundaries [19, 25]. Borders that correspond to qualitative physical differentiations or spatial discontinuities in the underlying territory (coastlines, rivers) we call *bona fide* boundaries; human-demarcation-induced borders we call *fiat* boundaries. Thus, the border of Australia is a bona fide boundary; the borders of Wyoming or Colorado, as well as most US county- and property-lines and the borders of postal districts, are examples of fiat boundaries. Correspondingly, we distinguish between fiat and bona fide objects depending on whether their boundaries are of the fiat or bona fide sort [18]. (Objects such as the North Sea whose boundaries include segments of both the fiat and bona fide kinds qualify as fiat objects.)

Most examples of fiat objects in the geographic world are correlated with two-dimensional regions on the surface of the globe. Examples of three-dimensional fiat objects are provided by the subterranean volumes of land to which mineral rights have been assigned, and also by the sectors and corridors in space established for the purposes of air traffic control. These may be quite complicated three-dimensional worms; they may intersect each other and they may have holes. On the other hand, insofar as an object whose boundary is not entirely of the bona fide variety counts as a fiat ob-
ject, many ordinary geographic entities, such as mountains, will also qualify as three-dimensional fiat objects. This is because the line which separates mountain and valley is a fiat line only (in fact a collection of fiat lines). In this sense, the fiat/bona fide distinction cross-cuts the superficially similar distinction between objects of physical and of human geography mentioned in §3.1.

4. Theoretical Tools

With these philosophical issues in the background, we may now distinguish three main theoretical tools that are required for the purposes of developing an overall formal theory of spatial representation that can help us to solve the problems of geographic representation. (Again, our presentation will be schematic. See [26] for related background material.)

4.1 Mereology. A major part of our reasoning about space involves mereological thinking, reasoning in terms of the part relation. Mereology represents a powerful alternative to set theory and appears to be especially suited for spatial and geographic representation. One can appreciate why this is so when one considers the question of identity criteria for geographic entities. Imagine a situation in which Italy sells Sicily to the United States. Is Italy the same state after the sale? If countries are construed as sets of regions in the mathematical sense of ‘set’, then one is forced to talk of two different countries in such a case, for sets are the same if and only if they have exactly the same elements. If, on the other hand, countries are mereological aggregates, then looser identity criteria for aggregates would allow one to accept the continued existence of geographic objects even after the loss of certain sorts of parts.

Formally, we may assume a mereology to be a first-order theory constructed around the primitive is a part of (interpreted so as to include identity as a limit case), which we symbolize as a binary predicate ‘≤’. In a full-blown account, this primitive should really be a three-place relation involving a temporal parameter [28], but we shall not consider such ramifications here. If we define overlap as the sharing of common parts:
\[ D \leq 1 \quad O(x, y) := \exists z (z \leq x \land z \leq y), \]

then the axioms for standard (non-tensed) mereology can be formulated as follows [17]:

\begin{align*}
A \leq 1 & \quad x \leq x \\
A \leq 2 & \quad x \leq y \land y \leq x \rightarrow x = y \\
A \leq 3 & \quad x \leq y \land y \leq z \rightarrow x \leq z \\
A \leq 4 & \quad \forall z (z \leq x \rightarrow O(z, y)) \rightarrow x \leq y \\
A \leq 5 & \quad \exists x (\phi x) \rightarrow \exists y \forall z (O(y, z) \leftrightarrow \exists x (\phi x \land O(x, z))).
\end{align*}

(Here and in the sequel initial universal quantifiers are to be taken as understood, and variables are to be conceived as ranging over all spatial entities.) Thus, parthood is a reflexive, antisymmetric, and transitive relation, a partial ordering. In addition, \( A \leq 4 \) ensures that parthood is extensional, whereas the schema \( A \leq 5 \) guarantees that for every satisfied property or condition \( \phi \) (i.e., every condition \( \phi \) that yields the value true for at least one argument) there exists an entity, the sum or fusion, consisting precisely of all the \( \phi \)ers. This entity will be denoted by ‘\( \sigma x (\phi x) \)’ and is defined as follows:

\[ D \leq 2 \quad \sigma x (\phi x) := \forall y \exists z (O(y, z) \leftrightarrow \exists x (\phi x \land O(x, z))). \]

With the help of this operator, other useful notions are easily defined. In particular, one can define the quasi-Boolean operators:

\begin{align*}
D \leq 3 & \quad x + y := \sigma z (z \leq x \lor z \leq y) \quad \text{sum} \\
D \leq 4 & \quad x \times y := \sigma z (z \leq x \land z \leq y) \quad \text{product} \\
D \leq 5 & \quad -x := \sigma z (\neg O(z, x)) \quad \text{complement}
\end{align*}

4.2 Location. A general theory of spatial location is needed over and above mereology in order to permit the investigation of the relation between a geographic entity and the region of space in which it is located. For this relation is not one of identity: a nation or a county is not identical with the spatial region it occupies. Italy can shrink or change its shape, but a spatial region necessarily has the shape and size it has. Moreover, two or more distinct geographic entities can share the same location at the same time—for instance, the city and the state of Hamburg. (Our common-sense talk of “locations” and “spatial regions” here should not be taken as expressing com-
mitment to an ultimately absolutist conception of space. The theory of location is in principle neutral with regard to the issues of § 2.2.)

Formally, the theory of location is conveniently axiomatized in terms of a primitive ‘L’ expressing exact location [9]. We write ‘Lxy’ for: ‘x is exactly located at region y’. With the help of mereology, the set of available locative relations can then be expanded as follows:

\[ DL1 \quad FL(x,y) := \exists z (z \leq y \land L(x,z)). \quad \text{full location} \]
\[ DL2 \quad PL(x,y) := \exists z (z \leq x \land L(z,y)). \quad \text{partial location} \]

Thus, Switzerland is fully located in the region where Europe is exactly located, while Turkey is only partially so. The basic axioms for L are:

\[ AL1 \quad L(x,y) \land L(x,z) \rightarrow y=z \]
\[ AL2 \quad L(x,y) \land z \leq x \rightarrow FL(z,y) \]
\[ AL3 \quad L(x,y) \land z \leq y \rightarrow PL(x,z) \]
\[ AL4 \quad L(x,y) \rightarrow L(y,y). \]

By AL1, a single entity cannot be exactly located at two distinct regions (hence, in particular, two distinct regions cannot be exactly co-located): L is a functional relation. AL2 and AL3 form the bridge from the theory of location to its mereological basis. AL4 guarantees that L behaves as a reflexive relation whenever it can: all (and only) those things at which something is located—i.e., on the intended interpretation, all spatial regions—are located at themselves. Thus, although we need not assume that everything is located somewhere (which would be a way of characterizing a world inhabited exclusively by spatial entities), we must inevitably ensure that this holds at least of all regions. Conversely, note that we are not assuming that every region is the region of something other than itself: not every region is a region at which something is located. (See § 5 below for more on this issue.)

AL1–AL4 form a minimal set of axioms. Now define a relation of co-incidence as holding between entities that share the same location (such as the city and state of Hamburg):

\[ DL3 \quad x \approx y := \exists z (L(x,z) \land L(y,z)). \quad \text{coincidence} \]

The coincidence relation \( \approx \) is easily seen to be an equivalence relation. In
addition, we assume two postulates to the effect that coinciding entities have coinciding parts and are closed under arbitrary sums:

$$\text{AL5} \quad y \approx z \land x \leq y \rightarrow \exists w(w \leq z \land x = w)$$

$$\text{AL6} \quad \exists y(\phi y) \land \forall y(\phi y \rightarrow x = y) \rightarrow x = \sigma y(\phi y).$$

Thus, in particular, if $x$ coincides with two entities $y$ and $z$, then it coincides also with their sum $y + z$.

4.3 Topology. One should not suppose that all geographic entities are connected, or of a piece. Slovakia or Wyoming have this property; Turkey and the United States do not. It is, however, impossible to account for this difference on the basis of a purely mereological framework [29]. Moreover, mereology alone cannot account for some very basic spatial relations, such as the relationship of continuity between two adjacent objects (countries, land parcels, postal districts), or the relation of one thing’s being entirely inside or surrounding some other thing. To provide a systematic account of such relations, which go beyond plain part-whole relations, and to ensure semantic transparency and computational efficiency of our intended general theory, we will require a topological machinery of one sort or another [14, 21, 30]. Finally, as mentioned in §3.3, the notion of a border between two adjacent lands or countries is crucial for geographic ontology, and this notion is topological in nature. However, the difference between fiat and bona fide boundaries will require two distinct sorts of topological theory in order to cope fully with these matters [25].

This is because the theory of bona fide boundaries corresponds formally to an ontology based on ordinary topology. Let ‘$B$’ denote this primitive boundary relation, so that ‘$B(x, y)$’ reads “$x$ is a bona fide boundary for $y$”. Using $\text{AL5}$, the (bona fide) closure of an entity $x$ is defined as the sum of $x$ with all its bona fide boundaries:

$$\text{DB1} \quad c(x) := x + \sigma z (B(z, x)).$$

(If $x$ has no bona fide boundaries, as for instance in the case of Wyoming, then we identify $c(x)$ with $x$.) The basic postulates for bona fide boundaries are then obtained by mereologizing the standard Kuratowski axioms for closure operators:
AB1 \[ x \leq c(x) \]
AB2 \[ c(c(x)) \leq c(x) \]
AB3 \[ c(x+y) = c(x)+c(y). \]

(Axiom AB1 is actually derivable from A≤1.) Connection is then defined in the obvious way:

\[
\text{DB2 } \quad C(x, y) := O(c(x), y) \lor O(c(y), x).
\]

and it follows that two discrete (i.e., non-overlapping) entities can be connected only if one of them is not closed (i.e., does not include all of its bona fide boundaries as parts).

Now, let ‘B*’ denote the primitive boundary relation for fiat boundaries, so that ‘B*(x, y)’ reads “x is a fiat boundary for y”. This relation is axiomatized as follows [23, 25]:

\[
\text{AB*1 } \quad B^*(x, y) \rightarrow x \leq y.
\]
\[
\text{AB*2 } \quad B^*(x, y) \land B^*(y, z) \rightarrow B^*(x, z).
\]
\[
\text{AB*3 } \quad x \leq y \land B^*(y, z) \rightarrow B^*(x, z).
\]

AB*1 requires that fiat boundaries are parts of the entities they bound (in contrast to bona fide boundaries, which may bound an entity from the outside). AB*2 and AB*3 ensure that B* is transitive and disjunctive: any point along the border of Wyoming is a boundary point of Wyoming; and every segment of the border is a boundary segment. Note that AB*1 rules out the possibility of introducing a significant analogue of the closure operator ‘c’ in the fiat world: the sum of an object with its fiat boundaries is in every case just the object itself. There is therefore no counterpart to the Kuratowski axioms in the theory of fiat boundaries. It also follows from AB*1 that fiat boundaries are not symmetric (whereas bona fide boundaries are: the bona fide boundary of an object is always the bona fide boundary of the object’s complement).

To characterize the relation of connection by fiat boundary—the sort of connection that occurs along the boundary between France and Germany—we now rely on the relation of coincidence. The idea, stemming from Brentano [4], is simply that there is a sui generis form of connection
which obtains between two adjacent entities whenever their fiat boundaries coincide at least in part:

\[ \text{DB}^*1 \quad \text{C}^*(x, y) := \text{O}(x, y) \lor \exists z \exists w (\text{B}^*(z, x) \land \text{B}^*(w, y) \land z \approx w) \]

Using this notion, self-connectedness can finally be characterized:

\[ \text{DB}^*2 \quad \text{Cn}^*(x) := \forall y \forall z (x = y + z \rightarrow \text{C}^*(y, z)). \]

This notion allows us to distinguish between self-connected and disconnected geographic entities. In addition, it also allows us to formulate suitable axioms to capture the idea that all boundaries are ontologically parasitic on (i.e., cannot exist in isolation from) their hosts, the entities they bound [12, 23]. This thesis—which stands opposed to the ordinary set-theoretic conception of boundaries as, effectively, sets of points, each one of which can exist though all around it be annihilated—can be expressed as follows. Let ‘IP’ and ‘IP*’ denote the relations of bona fide and fiat interior parthood:

\[ \text{DB}^3 \quad \text{IP}(x, y) := x \leq y \land \forall z (\text{B}(z, y) \rightarrow \neg \text{O}(x, z)). \]

\[ \text{DB}^*3 \quad \text{IP}^*(x, y) := x \leq y \land \forall z (\text{B}^*(z, y) \rightarrow \neg \text{O}(x, z)). \]

We now require, for self-connected boundaries, the existence of self-connected wholes which they are the boundaries of:

\[ \text{AB}^4 \quad \exists y \text{B}(x, y) \land \text{Cn}^*(x) \rightarrow \exists y \exists z (\text{Cn}^*(y) \land \text{B}(x, y) \land \text{IP}(z, y)). \]

\[ \text{AB}^*4 \quad \exists y \text{B}^*(x, y) \land \text{Cn}^*(x) \rightarrow \exists y \exists z (\text{Cn}^*(y) \land \text{B}^*(x, y) \land \text{IP}^*(z, y)). \]

5. Classical and Non-Classical Geographies

Let us now tentatively suggest a possible enrichment of our ontological instrumentarium. An analogy could help to fix our ideas here. In the nineteenth century a number of non-classical (non-Euclidean) geometries were put forward. These involved accepting some of the axioms of Euclidean geometry and rejecting others, say the axiom concerning parallels. Likewise, recent decades have witnessed a proliferation of non-classical logical theories which result from the rejection of one or other of the principles of classical logic (such as the law of excluded middle or the principle of double ne-
gation) or through a weakening or strengthening of classical logic that is obtained through deleting or adding specific logical constants. In a similar fashion, one might now consider the possibility of defining a classical geographic frame and setting this against certain non-classical geographies [22] which would arise by similar adjustments of associated constraints. Different Geographic Information Systems may then be viewed as stemming from the combination of a specific geography (classical or non-classical) with the necessary mereological, topological, and location-theoretic background. In this way a systematic order may be gained in the sphere of Geographic Information Systems, where alternative systems have thus far been constructed on an ad hoc basis.

5.1 Principles of classical geography. There is of course no such thing as “classical geography”, at least in the sense that there is no single universally recognized formulation that has a status analogous to that of Euclidean geometry or classical logic. But we can nevertheless proceed by putting forward some principles that seem plausible for a minimal characterization of geographic representation, and which are such that the violation of one or other of them produces intuitively incomplete representations.

We may broadly characterize a geography $G$ on a region $R$ as a way of assigning (via the location relation) geographic objects of given types to parts (subregions) of $R$. We shall use the letters ‘$a$’, ‘$b$’, ‘$c$’ as variables for geographic entities (nations, counties, districts, lakes, islands, etc.) or mereological combinations thereof, and ‘$x$’, ‘$y$’, ‘$z$’, as variables for regions of space (locations). Classical geography can now be defined by the following axioms:

$$\text{CG1} \quad \exists x L(a,x)$$
$$\text{CG2} \quad \exists a L(a,x)$$
$$\text{CG3} \quad L(a,x) \land L(b,x) \rightarrow a=b$$

By CG1, every geographic entity is located at some region (indeed this region is unique by AL1, the first location axiom). By CG2–CG3, every region has a unique geographic entity located at it.

The term ‘classical geography’ does not carry any normative claim. It simply describes a rather robust way of tiling regions in the presence of cer-
tain general constraints. Thus, a model of classical geography could be visualized as a set of directions for coloring maps. We have a fixed set of colors; every sub-region of the map has some unique color, and every color is the color of a unique region of the map. But the model makes no essential reference to maps. It can be applied straightforwardly to the terrain itself. For instance, we can generate a simple classical geography via the tiling which divides the earth’s surface into land and water. A somewhat more sophisticated classical geography covering the whole of the earth’s surface is obtained if we allow within our tiling a division between nations (including quasi-nations such as Antarctica), national waters and international waters.

5.2 Non-classical geographies. Consider now the effects of dropping one or more of the axioms of classical geography.

Dropping CG1 licences non-spatial geographic units. An example could be that of the Sovereign Military Hospitaler Order of St. John, or Poland during the Era of Partition. (A default assignment that preserves the axioms of classical geography would consist in considering some more or less arbitrarily chosen region as Ersatz-Poland—perhaps the headquarters of the Government in Exile in London—during the era in which the entity in question does not have any territory to call its own.)

Dropping CG2 will licence a map with gaps, i.e., regions that are assigned no unit. (A default assignment that preserves the axioms of classical geography would consist in considering all such regions as occupied by an object of the No Man’s Land type.)

Dropping CG3 will licence a map with gluts, i.e., regions that are assigned two or more competing units. (A default assignment that preserves the axioms of classical geography would consist in considering all such regions as occupied by objects of the Disputed Land type.)

One may also consider weakening AL1 so as to allow for the possibility of duplicates. Mainland China and Taiwan both claim to be the only China, but we cannot accept both claims if we assume AL1. (A default assignment that preserves the axioms of classical geography would here consist in considering the mereological sum of the competing regions as associated with a single spatial object.)
We might also add axioms to those of classical geography, for example an axiom to the effect that all geographic units are connected. We might finally consider how the properties of geographic boundaries relate to the axioms of classical geography. We shall say that a boundary is geometrically two-sided if it divides two adjacent units. In a classical geography, the geometric two-sidedness of any boundary is secured by the completeness of the tiling. This is no longer the case if non-classical geographies are considered. For instance, in a gappy geography the boundaries of objects at the edges of non-assigned zones will be one-sided only. And so, in a glutty geography, will be the boundaries of objects at the edges of zones assigned to more than one object.

6. Conclusion

The work to be done here is not simply that of assembling a collection of difficult cases for standard geography. Rather, we are concerned with the identification of the different ways of treating geographic structures. We hold that the systematic analysis of the basic structures involved in spatial representation is a first fundamental step towards the development of an ontologically and cognitively well-grounded understanding of space and spatial information science.

Acknowledgments. Thanks are due to the National Center for Geographic Information and Analysis (NCGIA) which has provided valuable support in our work on the present project.

References


