Geographic Regions as Brute Facts, Social Facts, and Institutional Facts

Daniel R. Montello

1. What Are Geographic Regions?

Geographic regions (henceforth, regions for short) are spatially extended pieces of (near) earth surface that share some aspect of similarity across their extents. The most fundamental aspect of similarity shared within a region is “locational similarity”—proximity. But regions nearly always share common thematic content or activity too. That is, regions are defined not just spatially but according to the human and natural entities or processes occurring there. Regionalization is spatial categorization, and importantly, this spatiality is literal rather than metaphorical, as it is with most other category systems. The modifier “geographic” implies a region of space that is prototypically two-dimensional (though often bumpy or uneven) and at or near the earth surface. Volumetric exceptions exist underground and in the atmosphere, but they are rare given the geographer’s interest in the earth as the home of human habitation. Also because of the human-centered focus of geography, the human and natural themes that provide content for regions are usually phenomena at human-centered spatial and temporal scales—not too fast or small, not too slow or large (Montello 2001).

The fact that geographic regions typically encompass places that are close together is interesting and conceptually nontrivial because it leads to a generalization about the shape of regions. Regions tend to be...
“compact”—spatially contiguous and unfragmented areas that extend roughly equally in all directions (i.e., roughly circular). This locational similarity within regions arises for an intriguing variety of reasons. In fact, closer places are more similar, on average. This principle (famously called the “First Law of Geography” by Tobler 1970) holds true as a generalization because more interaction occurs between closer places, which in turn holds true because interaction (movement of air masses, distribution of newspapers) has less “friction” (air pressure, delivery cost) to overcome between closer places. But people also tend to perceive and conceive of closer things as being more similar. A lower-level perceptual example is provided by the Gestalt law of proximity as a basis for figural organization in visual perception; a more high-level cognitive example is provided by the tendency people have to interpret closer graphic elements as representing more similar entities in graphical spatializations of databases—the “distance-similarity metaphor” (Fabrikant, Montello, and Mark 2006; Montello et al. 2003). It’s also been shown that people care more about closer objects and events (Ekman and Bratfisch 1965). Furthermore, compact regions are more useful intellectually and administratively. The efficiency of resource distribution in compact countries as compared to elongated, prurient, or fragmented countries provides a clear example of the latter. “Remote” means away from the rest of a region; severe remoteness does not occur in compact regions. The notion that regions are naturally or rationally compact, at the very least contiguous, was in fact expressed during mid-twentieth-century debates in the U.S. in opposition to giving statehood to Alaska and Hawaii (Jones 1959). This notion is also reflected in laws about the proper design of electoral districts and legal challenges over the acceptability of noncompact “gerrymandered” districts (Forest 2001).

As a subset of categorical thinking, regional thinking is fundamental to the human conceptualization of built and natural landscapes; the universal verbal labeling of places and features on the earth reflects this. It has been fundamental specifically to geographers, too. In the intellectual toolbox of the geographer, the concept of regions has been one of the basic tools—a standard claw hammer rather than a cooper’s howel. The region concept has always provided service to geographers, whether they worked primarily in natural language, formal language, or spatial language. 2 It has helped geographers describe their domain of interest, the (near) earth surface that supports its natural and human phenomena; it has also helped geographers in their attempts to achieve the other traditional scientific goals of prediction, explanation, and control (Tobler [1973] discussed an analogy between regionalization and time-series analysis, pointing out how in both cases continua are discretized for statistical analysis). Regions still serve these functions and will likely continue to do so, even in the digital world of the Internet, virtual reality, and geographic information systems; Stubbjär (2001) provides several examples that show how the advent of digital geographic information systems make an understanding of the ontology of regions even more important. Still today, Kimble’s (1951) claim that geography is the “study of regions” is viable (his criticisms concerning the lack of agreement about specific region identities notwithstanding).

The ontology of geographic regions may be further articulated by considering a taxonomy of region types (originally presented in Montello 2003), a taxonomy that incorporates and hopefully clarifies traditional conceptualizations of regions by academic geographers. The taxonomy is based on the information and procedures that people, including geographers, use to identify the regions. The taxonomy consists of four types of geographic regions: administrative, thematic, functional, and cognitive. Administrative regions are created by legal or political action, by decree or negotiation. These include regions based on property ownership (i.e., cadastral regions or “real estate”) as well as regions based on bureaucratic or governmental control, such as census tracts, provinces, and countries. Thematic regions are formed by the systematic measurement and mapping of one or more observable content variables or themes; those based on more than one theme are multivariate thematic regions. Thematic regions express where particular entities or properties exist, whether natural (rainfall, pine trees) or human (languages, crops) in origin. Functional regions are based on patterns of interaction across space (and time) between separate locations on the earth. Spatial interaction is the movement of matter or energy from place to place: people, commodities, water, seeds, earthquake tremors. Patterns in transmitted matter or energy earth’s surface, I do not intend to restrict the term “region” to a specific scale. Geographers have often spoken of a “regional scale” (e.g., Meyer et al. 1992; Richardson 1992; Stubbjär 2001), a term meant to exclude smaller entities like places and larger entities like continents. I consider all of these entities to be examples of regions.

I consider all of these entities to be examples of regions.

3. People commonly use the same region label at different times to refer to different types of regions, a phenomenon that might be termed region polymorphism. For example, the name “California” is typically used to refer to an administrative region but at other times could refer to a thematic region, a functional region, or a cognitive region (a “California state of mind”).

1. The tendency toward compactness is strictly true only on an isotropic plain, a flat twodimensional surface that has no barrier or facilitation to interaction differentially in different directions. On the real earth, oceans, mountains, uneven population distributions, transportation networks, and so on lead to nonisotropic surfaces that help create noncompact regions. However, the modifications these factors impose on compactness can be accounted for analytically in order to predict more accurately actual region shapes (e.g., Haggerty 2001).

2. Aside from the fact that geographic regions concern scales of human activities on the
can encode information, the basis for communication, which is thus a form of spatial interaction. As with thematic regions, multivariate functional regions may be based on a combination of more than one type of interaction. Finally, cognitive regions are produced by people's informal perceptions and conceptions of the earth surface, based on direct experience, maps, hearsay, and so forth. Examples include downtown and the Midwest (as these concepts are used informally by lay people); another would be our special meeting place in the woods out back.

2. Boundaries and the Ontological Status of Regions

The nature of the boundary between the inside and outside of regions is central to an ontological exposition of regions (Jones 1959; Kristof 1959). Boundaries are prototypically thought of as lines but often are not lines. Instead, they are two-dimensional bands or "transition zones." The two-dimensionality of region boundaries is usually due to their vagueness (Mark and Csillag 1989). Vagueness means that a boundary is less than absolutely precise in its second dimension; a perfectly precise region boundary is a one-dimensional line. A variety of near synonyms for vagueness include imprecision, indeterminacy, ill-boundedness, gradation, error, uncertainty, and fuzziness. There are several potential causes for boundary vagueness:

a) measurement error or imprecision (measurement vagueness)
b) boundary changes over time (temporal vagueness)
c) alternative variable combinations in multivariate regions (multivariate vagueness)
d) disagreement about boundary locations (contested vagueness)
e) fundamentally vague concepts of reality (conceptual vagueness)

These are quite distinct causes of vagueness, with very different practical and philosophical implications (e.g., Burrough and Frank 1996). With respect to the ontology of regions, two distinctions suggested by this list are perhaps most important. One is between contested vagueness, which allows single persons or groups to identify regions with precise boundaries (just not the same ones), and the other causes, which produce "universally held" vagueness (time leads to universally held vagueness whenever a sufficient interval of time to allow boundary changes has elapsed). A second critically important distinction in the list is between the first four causes and the last, conceptual vagueness. The first four may always or typically lead to vagueness in practice even though they need not in theory. Conceptual vagueness, in contrast, necessarily leads to vagueness even in theory.

The four types of regions from the taxonomy tend to differ in the extent and nature of their boundary vagueness. For example, all four region types, not just administrative regions, may have boundaries that are vague in the contested sense, though only contested vagueness over administrative boundaries tends to lead to war. In contrast, only administrative regions typically have boundaries that are conceptually precise. And this is not just an invention of modern Western thinking; Jones (1959) cites several ethnographic instances of hunter-gatherers precisely demarcating their territories with fences or rivers or other markers. This suggests the special status of administrative regions. For the other three region types, the degree to which places on earth have the region's characteristic property or properties typically varies more or less continuously, and only a partially arbitrary decision can locate boundaries. For example, it would be impossible in practice to ever measure all of the natural variables relevant to defining an ecosystem region at every location at the same moment in time with perfect accuracy and precision, and if you could, you would not get a completely contiguous region anyway. But precisely delimiting an ecosystem region cannot generally be done even in theory. The ecosystem region is simply not defined precisely enough as a concept to produce crisp boundaries, whatever the measurement fantasies entertained. Similarly, cognitive and functional regions are typically fundamentally vague, with every crisp representation a fiction to some extent. There is not in principle, let alone in fact, a precise boundary around the Bible Belt or the region where people receive the London Times (there must be at least one person in almost every corner of the world, not just in the Commonwealth, that receives that paper). This is not to say there is no transition in reality corresponding to the boundaries of these regions, only that the transition which really exists is really not sharp. Nor does boundary vagueness nullify the fact that the utility of region systems depends in part on their correspondence to states of the real world. I revisit the question of the reality- correspondence of regions below.

Administrative regions are thus unusual in being the only one of the four region types that typically has potentially precise boundaries (see...
need to be for a particular purpose, which has often been relatively imprecise in practice (Kristof 1959). There are numerous administrative boundaries in the world that are vague, but only because no one has yet cared to make them precise.

Brute facts distinguishes brute facts from social facts (see also Smith and Searle 2003). In much of his writing, Searle is particularly interested in a subset of social facts he calls institutional facts. These are collective intentional facts that have deontic powers resulting from the assignment of status functions, functions that create obligations, rights, and duties. Deontic refers to the modal logic of obligation and permissibility—what one may or may not do, or ought or ought not do. Classic examples of institutional facts are money and legal codes. A cadastral region (i.e., real estate) is a good example of an institutional fact. Land ownership conveys certain obligations and permissions on its owners. Philosophically, this is interesting because the status functions of real estate derive from a social contract, not from any physical reality such as a barbed-wire fence surrounding the perimeter (although such a fence may well be in place).


In *The Construction of Social Reality*, the philosopher John Searle (1995) distinguishes brute facts from social facts (see also Smith and Searle 2003). *Brute facts* are facts independent of human intentionality, often but not always facts about physical reality. A stone is a brute fact, as is a cloud, but so is an automobile, even though its genesis as an automobile is anthropic. In contrast, *social facts* fundamentally depend on human intentionality for their existence. An exposition of the ontology of regions is a social fact, as is a gathering of scholars discussing ontology at a conference (though a spatial agglomeration of humans is a brute fact). One may note a further distinction between social facts that are idiosyncratic to an individual person’s cognition and those that are shared among a collection of individuals. Thus, social facts might better be termed “intentional facts,” which can be further subdivided into “individual” and “collective” intentional facts (Smith 2003). Intentionality, in the specific philosophical sense meant here, concerns the fact that certain entities apparently have the property of “aboutness” or “directedness” (Dennett and Haugeand 1987; Searle 1983; Siewert 2002). That is, they necessarily refer to some object or event, actual or imagined, outside themselves. To many scholars, intentionality has been seen as a defining property of mental states. An example would be a belief that a region of some type exists—the belief is about something outside itself, in this case a portion of the earth’s surface.

In much of his writing, Searle is particularly interested in a subset of social facts he calls institutional facts. These are collective intentional facts that have deontic powers resulting from the assignment of status functions, functions that create obligations, rights, and duties. Deontic refers to the modal logic of obligation and permissibility—what one may or may not do, or ought or ought not do. Classic examples of institutional facts are money and legal codes. A cadastral region (i.e., real estate) is a good example of an institutional fact. Land ownership conveys certain obligations and permissions on its owners. Philosophically, this is interesting because the status functions of real estate derive from a social contract, not from any physical reality such as a barbed-wire fence surrounding the perimeter (although such a fence may well be in place).

Thus, when we consider the identity of geographic regions as brute facts, social facts, or institutional facts, we see that administrative regions (including cadastral regions) are clearly not only collective intentional (social) facts but also institutional facts. Their institutional nature makes administrative regions different than other region types. Searle’s (1995) heuristic formulation for the definition of institutional facts—*X counts as Y in context C*—applies to administrative regions. This seems true no matter whether the boundaries that designate such regions are, in Smith’s terms, fiat or bona fide or some combination. In his reply to Smith (Smith and Searle 2003), Searle questions Smith’s claim that there are fiat administrative boundaries that are not brute facts. On the surface, Smith’s claim appears quite correct. Administrative boundaries often exist that are not directly marked in any way in the world. For example, the Law of the Seas...
recognizes “ocean economic zones” as (administrative) regions that extend offshore from sovereign lands. Dominion over these zones is an institutional fact lacking physical reality or markers. Another example is provided by so-called geometric boundaries, flat boundaries that are “artificially” straight, quite common for administrative boundaries in many parts of the U.S. (Jones 1959 claims that the papal line of demarcation in South America was the first one). As I argued at greater length in Montello (2003), however, the flat or “brute-free” nature of boundaries should not be exaggerated. Administrative boundaries are always located relative to some brute fact; even administrative boundaries that are unmarked by physical objects are identified by surveying from other marker objects or from absolute coordinate systems based on physical aspects of the earth’s geodesy, such as the equator, the Greenwich Observatory, the edge of a landmass, or satellite positions and signals.

In contrast to Searle’s doubts, I have more concern about Smith’s claim that any administrative boundaries are solely brute facts, even if he were correct to claim that some are not brute facts. In fact, even when physical features mark administrative boundaries (quite common at various times and places), or have served historical roles in establishing boundaries, the features are generally not the boundaries, only markers for the boundaries. “The boundary . . . has no life of its own, not even a material existence. Boundary stones are not the boundary itself . . . only its visible symbols” (Kristof 1959, 272). The area in the American Southwest known as Four Corners exemplifies this point nicely. This location, where the boundaries of four states meet at a point, is marked on the ground with a plaque showing four borderline segments. Although tourists are seemingly impelled to straddle these line segments, they are not in fact the boundaries—they denote the boundaries. When a vandal moved the plaque several feet on January 25th of 2004, Reuters’ news service reported that “a debate has ensued as to whether this means the state borders have been shifted.” The story does not make clear whether this debate is uninformed or tongue-in-cheek.

What about other geographic regions—what kinds of “facts” are they? Thematic and functional regions perhaps seem to be examples of brute facts only. But that is not the case, because these types of regions are always created as a product of human intentionality. The common existence of multivariate vagueness makes this clear; different maps of vegetation regions vary because different vegetation scientists (such as biographers) choose to include different themes, and measure them differently, in their definitions and delineations of the regions. Thematic and functional regions definitely are about something, namely the earth surface, so they are intentional (social) facts. At the same time, they have a physical, brute reality.

Cognitive regions too are clearly intentional facts. As we noted above, intentional facts may be individual or collective; similarly, cognitive regions may be quite idiosyncratic or shared across a collection of individuals. “The Bible Belt” is relatively shared; “my special place where I daydream by the river” is relatively idiosyncratic. Of course, the two blend into one another: Aggregated individual cognition helps make culture, and instantiated cultural cognition helps make individual minds. But whether individual or social, recognizing the intentional nature of cognitive regions does not deny that they too depend on physical brute facts about the earth surface, not just human ideas. Reality is a major basis for human experience and conceptual structure. The bend in the river that makes my spot so apt for its purpose is really there.

Recognizing the role of intentionality in the identification of geographic regions is critical to understanding their ontology. Intentionality confers the property of semiotic reference, defined and understood either idiosyncratically by an individual person or socially by a collection of individuals. Semiotics is the study of sign systems (Berger 1999; MacEachren 1995; Sebeok 2001). According to Peirce’s conceptualization, sign systems consist of sign-vehicles (marks, patterns, physical objects, or events), interpretants (concepts, semantic interpretations), and referents (real or imagined entities or properties that are connected to the interpretant by the sign-vehicle). Alternatively, the conceptualization of semiotics provided by Saussure referred only to sign-vehicles, which he called signifiers, and interpretants, which he called signifieds; in other words, he omitted the referent.

All geographic regions are thus, in part, semiotic entities. This is clear in the case of administrative regions. Searle uses the term “status indicator” to refer to outward signs of deontic status; these are sign-vehicles for institutional facts (“Private property—No Trespassing” is a good example). This is reminiscent of Peirce’s term “legisign,” the case wherein a law acts as a sign-vehicle. But thematic, functional, and cognitive regions are also semiotic. The property of intentionality, not the property of deontic reference, makes regions semiotic entities. Intentionality creates a mapping of social facts onto brute facts; this mapping can be many to one, one to one, or one to many. Furthermore, using Searle’s terminology, there are freestanding “Y-terms” that are themselves sign-vehicles but with referents that are not brute facts (though it is likely that all such freestanding terms can ultimately be connected to brute facts, real or imagined). An example would be small red hexagons on a map that stand for stop signs at street

6. Cognitive regions shared by cultural subgroups of nonexperts have traditionally been called vernacular regions.
corners, which are themselves sign-vehicles. But geographic regions do not involve freestanding Y-terms, because real regions (not imaginary or literary regions) refer to parts of the earth’s surface.

4. Conclusions: Regions as Social Facts and Brute Facts

Above, I described a taxonomy of geographic regions said to be based on the way regions are identified. One might instead say the taxonomy is based on the way regions are created. Such a usage points to the semiotic identity of regions. But this is not to say that regions, even administrative or cognitive regions, are merely conceptualizations. They are also brute facts. Real estate is made of soil and rock as well as deontic prerogative. To ask whether regions are brute facts or intentional facts is misleading (see the exchange in Smith and Searle 2003). It creates paradoxes by insisting on a dualism in which social or even institutional facts are claimed to be exclusively physical truths or conceptual truths. All geographic regions are the product of mind-body interactionism (a swineherd and an accountant discussed this further in Montello 2003). They are not just idealistic entities, nor naïvely realistic entities, but semiotically realistic entities. Semiotic realism requires sentient intentionality in a real world of differentiated physical reality.

Even administrative regions always refer to a brute reality. Administrative regions are special in being the one type of region that is institutional, as Searle defines it, but that does not make them idealistic chimera. It is true that the physicality of administrative regions (or any other type of region) is not a sufficient condition for their identification, but it is a necessary condition. Intentionality leads to alternative regionalizations of the same piece of earth surface, but there is still only one “tract o’ land.” Regions exist in laws and regulations, in minds and cultures, and in the real world that must ultimately be a concern of an animal species that has a natural existence.

7. A good argument can be made that there really are no pure brute facts as Searle defines them. A stone, a cloud, and a tree obviously have physical, intention-independent existence—but not as a stone, a cloud, or a tree per se. Only human intentionality uses the property of size to distinguish a stone from a sand grain or a mountain. Without human intentionality (including perception), thresholds in the continuously varying density of water droplets in the atmosphere would not appear as the boundaries of discrete clouds. In the context of regions, Kristof stated it thusly: “Not only boundaries but all limits ascribed to an area...are always subjective. They are defined anthropocentrically” (1959, 276-77).


