SPECIAL SECTION: ONTOLOGICAL ISSUES FOR THE NATIONAL MAP

On the Integration of Regional Classification and Delineation Systems into The National Map

Thomas Bittner

Departments of Philosophy and Geography & National Center for Geographic Information and Analysis (NCGIA) / State University of New York at Buffalo / NY / USA

Abstract

Many of the qualities that characterize geographic regions are vague and granular in their nature. In many quality-based classification and delineation systems for geographic regions, therefore, there is a trade-off between the possible precision of the quality-based delineation and the scientific sophistication of the quality-based classification of geographic regions. This poses a dilemma for the US Geological Survey’s National Map, whose purpose is to provide various integrated classification and delineation systems that can serve a wide range of users. Some users need precise delineation systems, while others need sophisticated classification systems. Many users are required to use and to produce data that are not affected by the above trade-off and that can be integrated in consistent ways. This article discusses an ontology-based solution to this problem, presented in the specific context of systems for classifying and delineating eco-regions and eco-zones.

Keywords: ontology, classification and delineation systems, eco-regions, eco-zones, biomes, National Map

Résumé

La plupart des qualités qui caractérisent les régions géographiques sont plutôt vagues et modulaires. Par conséquent, dans de nombreuses classifications basées sur la qualité et de nombreux systèmes de délimitation pour les régions géographiques, on fait un compromis entre la précision possible de la délimitation basée sur la qualité et la précision scientifique de la classification basée sur la qualité des régions géographiques. Ceci pose un dilemme pour la National Map du US Geological Survey, dont l’objectif est de fournir des systèmes intégrés de classification et de délimitation à un large éventail d’utilisateurs. Certains utilisateurs ont besoin de systèmes de délimitation précis, tandis que d’autres ont besoin de systèmes de classification sophistiqués. De nombreux utilisateurs utilisent et produisent des données non touchées par ce compromis et pouvant être intégrées avec une certaine constance. Dans l’article, on parle d’une solution basée sur l’ontologie, et on la présente dans le contexte des systèmes visant la classification et la délimitation des régions et des zones écologiques.

Mots clés : ontologie, systèmes de classification et de délimitation, régions écologiques, zones écologiques, biomes, National Map

1. Introduction

The aim of the US Geological Survey’s National Map (TNM) is to provide geographic information to broad groups of users ranging from researchers in the life sciences to administrators and decision makers in various government planning and management agencies (USGS 2009). Among TNM’s aims is to provide geographic information in the form of “area-class maps” (Bunge 1962) or “categorical coverage maps” (Chrisman 1982, 1997). Such maps are popular representations of classification and delineation systems for geographic
regions that are based on qualities that characterize such regions, including climate, landforms, and vegetation (Hartshorne 1939; Lewis 1982). Within the framework of TNM, such systems are required to deliver a sophisticated degree of classification as well as the highest possible degree of precision in delineating geographic regions. Precise delineation is desirable for many administrative purposes (Omernik 2004); a sophisticated and detailed classification system is desirable for scientific purposes, for data integration (Schuurman and Leszczynski 2006), and to facilitate automated reasoning (Yetongnon and others 2006) and the integration of classification-based data into the Geo-semantic Web (Abdelmoty and others 2005; Semantic Web 2009).

Unfortunately, a precise quality-based delineation of geographic regions relies to a significant extent on local combinations or patterns of qualities that distinguish neighbouring regions (Omernik 2004). By contrast, quality-based classification systems for geographic regions rely on non-local combinations or patterns of qualities that differentiate classes (or universals) independently of the specific location of their instances. In this article, based on work reported elsewhere (Bittner 2009), the eco-region classification and delineation systems of the US Forest Service (USFS 1995) and the US Environment Protection Agency (EPA 1997) are used to illustrate the trade-off between the local character of the precise quality-based delineation of geographic regions and the non-local character of their quality-based classification.

Consider the map displayed in Figure 1(a). This map depicts some of the Level III eco-regions of the EPA system (non-bold black boundaries) and corresponding Section eco-regions of the USFS system (bold black boundaries). Both EPA and USFS systems delineate eco-regions of comparable scale; in addition, both systems delineate eco-regions based on similar qualities, including the distribution patterns of plants and animals as well as patterns of climatic, soil, and similar land-surface- and climate-based qualities. Thus, there seem to be good reasons to contemplate the notion that the boundaries of the eco-regions delineated by the two systems coincide at least roughly. However, Figure 1(a) shows that although the delineations of EPA Level III eco-regions and USFS Section eco-regions are similar in some obvious but hard-to-specify way, the boundaries do not coincide at all. The EPA’s delineation appears to be more precise and detailed than the rather coarse and highly generalized USFS delineation. As discussed by Gerard McMahon and others (2001), non-coinciding boundaries of regions in classification and delineation systems that are based on similar qualities are a common problem.

An analysis of the EPA system shows that its more precise and detailed delineation is due to its extended use of local qualities and quality patterns, which, in addition, are often specified relative to qualities of neighbouring regions (e.g., lower/higher precipitation, less relief, or more irregular; EPA 2002; Bittner 2009). Clearly, such local quality patterns and quality differences characterize how specific regions are situated in their specific environments; they do not necessarily determine what makes a given geographic region an instance of a broader class (or universal).

By contrast, in the USFS system the focus is more on a classification system and general spatial nesting structures of the identified classes (universals). Figure 1(b) is a graph representing the spatial nesting of classes identified in the USFS system. In such a classification system, the focus is on qualities and quality patterns that characterize geographic regions independently of their specific surroundings and that determine to which larger class (universal) a particular geographic region belongs. Because of this focus on non-local quality patterns for classification purposes, the delineation of eco-regions is rather coarse (as is clearly visible in Figure 1(a)). Thus, the USFS system offers sophisticated and detailed quality-based classification at the cost of precise quality-based delineation; by contrast, the EPA system provides precise quality-based delineation at the cost of a sophisticated quality-based classification system.

Because of the fundamentally vague and granular nature of many of the qualities that characterize geographic regions (Mark and Csillag 1989; Ahlqvist 2005; Waterton 2002; Bailey 2004; Omernik 2004; Kronenfeld 2003), there is no easy way to overcome the trade-off between the local character of the precise quality-based delineation of geographic regions and the non-local character of the quality-based classification of geographic regions (Bittner 2009). In what follows, one possibility to overcome this trade-off is discussed. The aim is to build a classification and delineation system that includes both precise quality-based delineation and sophisticated and non-local quality-based classification. The underlying idea is to integrate the USFS and EPA systems into a single system that is similar in structure to the system used by the World Wildlife Fund (WWF; Olson and others 2001). The resulting system inherits the best aspects of all three of the original systems and could be integrated in TNM to meet the needs of all TNM users.

A previous ontology-based analysis of quality-based classification and delineation systems for geographic regions (Bittner 2009) provides the theoretical framework for this article. It also provides the basis for a symbolic, logic-based representation that may facilitate the integration of TNM into the Geo-semantic Web (Semantic Web 2009).

2. Ontological Properties of Classification and Delineation Systems

To integrate the three systems, shared structural properties are formalized using the top-level ontology of (Bittner, Donnelly, and Smith 2009; Bittner 2009). Here relevant
top-level terms are introduced, using the USFS and EPA systems as examples. Particularly important definitions are also summarized in Table 1.

2.1 MEREOLOGY

Consider the maps of the eco-regions of North America shown in Figure 2. Each map represents a collection of eco-region particulars of a certain size range that are parts of the North American continent (NAC). The symbols LI, LII, and LIII are used to designate the collections whose members are depicted on the respective maps. The collection LI partitions the continent of North America (NAC): the members of LI jointly sum up to NAC, and no two distinct members of LI overlap. Similarly, the collection LII is a more fine-grained partition of NAC. The collection
Table 1. Definitions for basic categories and relations (Bittner 2009; Bittner and others 2009)

<table>
<thead>
<tr>
<th>Category/Relation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Particulars</td>
<td>Independent continuants, i.e., entities that endure through time while undergoing different kinds of changes (you, me, …)</td>
</tr>
<tr>
<td>Qualities</td>
<td>Specific properties or qualities that inhere in particulars (your height, my weight, …)</td>
</tr>
<tr>
<td>Collections</td>
<td>Particulars grouped in an arbitrary set-like manner; collections are finite and non-empty ({you, me}, …)</td>
</tr>
<tr>
<td>Universals</td>
<td>Particulars grouped in more restricted ways that often result in tree-like structures (human being, mammal, vertebrate, …); universals and collections have very different temporal properties (Bittner and others 2009)</td>
</tr>
<tr>
<td>Quality universals</td>
<td>Particular qualities grouped in rather restricted ways that often result in tree-like structures (weight, height, …)</td>
</tr>
<tr>
<td>Partition</td>
<td>Collection ( p ) partitions the particular ( y ) if and only if the members of ( p ) jointly sum up to ( y ) and no two distinct members of ( p ) overlap</td>
</tr>
<tr>
<td>( \uparrow )-p-included</td>
<td>Collection ( p ) is upwardly partonomically included (( \uparrow )-p-included) in collection ( q ) if and only if every member of ( p ) is a part of some member of ( q )</td>
</tr>
<tr>
<td>Sub-universal of</td>
<td>Universal ( c ) is a sub-universal of universal ( d ) if and only if every instance of ( c ) is an instance of ( d )</td>
</tr>
<tr>
<td>( \uparrow )-u-part-of</td>
<td>Universal ( c ) is a ( \uparrow )-u-part-of universal ( d ) if and only if every instance of ( c ) is a part of some instance of ( d )</td>
</tr>
</tbody>
</table>

Homogeneous: Let \( Q_1, \ldots, Q_n \) be variables ranging over quality universals. A geographic region \( x \) is homogeneous with respect to the quality pattern \( Q = \langle Q_1, \ldots, Q_n \rangle \) (\( Q \)-homogeneous) if and only if for every \( j \) such that \( 1 \leq j \leq n \) it holds that the sum of all the parts of \( x \) that are regions of geographic scale and in which inheres an instance of \( Q_j \) is roughly the same size as \( x \). In addition, it holds that the same quality does not appear twice in a quality pattern (i.e., if \( Q_j = Q_k \) then \( j = k \)).

Maximally homogeneous: Region \( x \) is maximally homogeneous with respect to the quality pattern \( Q = \langle Q_1, \ldots, Q_n \rangle \) (maximally-\( Q \)-homogeneous) if and only if \( x \) is \( Q \)-homogeneous and all \( Q \)-homogeneous regions that overlap \( x \) are parts of \( x \).

Quality pattern: Let \( Q_1 = \langle Q_1, \ldots, Q_n \rangle \) and \( Q_2 = \langle Q_1', \ldots, Q_m' \rangle \) be quality patterns. \( Q_1 \) is a sub-pattern of \( Q_2 \) (\( Q_1 \subseteq Q_2 \)) if and only if for all \( Q_k (1 \leq k \leq n) \) there is a \( Q_k' (1 \leq j \leq m) \) such that \( Q_k = Q_k' \); \( Q_1 \) is a proper sub-pattern of \( Q_2 \) (\( Q_1 \subset Q_2 \)) if and only if \( Q_1 \subseteq Q_2 \) and not \( Q_2 \subset Q_1 \).

Genus- or species-pattern: Let \( QS = \langle Q_1, \ldots, Q_n \rangle \) and \( QG = \langle Q_1', \ldots, Q_m' \rangle \) be quality patterns such that if \( Q_j \) is a proper sub-universal of \( Q_j' \) for \( 1 \leq j \leq n \), then \( QG \) is a genus-pattern of \( QS \) and \( QS \) is a species-pattern of \( QG \).

LIII partitions the continental United States and is complemented by similar partitions to the north (Wiken 1986; Marshall and Schut 1999) and to the south (Rzedowski 1978). The collections LI, LII, and LIII are spatially nested (\( \uparrow \)-p-included) within one another; that is, every member of LII is a part of some member of LI. Similarly, LIII is \( \uparrow \)-p-included in LII, LIII is \( \uparrow \)-p-included in LI, and so on. Specific kinds of partitions also include exhaustive tessellations and other exhaustive subdivisions of space, including political subdivisions.

2.2 Qualities and Quality Patterns

Geographic regions are characterized by qualities. Relevant for this presentation are land-surface, climate, biogeographic, and biomic qualities. Land-surface qualities are determined by the distribution pattern of plants, animals, soils, landforms, and so on (Omernik 1987). Climate qualities are determined by the seasonal fluxes of energy and moisture and the resulting broad climatic similarities, similar average weather, and climax vegeta-
tion (Koeppen 1931; Bailey 1983; AMS 2007). Biomic qualities are determined by distinctive plant and animal groups (e.g., temperate grasslands and savannas; Olson and others 2001). Bio-geographic qualities are determined by historical and evolutionary distribution patterns of plants and animals (Udvardy 1975; Olson and others 2001). Some of these qualities are interrelated; for example, biomic qualities depend on climate qualities. Relevant classes of geographic qualities (quality universals) and taxonomic (sub-universal) relations between them are depicted in Figure 3. Particular instances of such quality universals inhere in specific geographic regions.

In quality-based classification and delineation systems, geographic regions are characterized by quality patterns (or mosaics of qualities; Wiken 1986; Omernik 1995). Here quality patterns are understood as ordered tuples of quality universals. Some relevant quality patterns whose qualities jointly characterize ecological and bio-geographic regions are displayed in Table 2 (Omernik 1987, 1995; Bailey 1983; Olson and others 2001). Such quality patterns are formed by the land-surface, climate, biomic, and bio-geographic quality universals shown in Figure 3.

To capture the granular and scale-dependent ways in which quality patterns characterize geographic regions, the technical notion of homogeneity (Bittner 2009) is used (see Table 1.) Consider, for example, the geographic regions Central Great Plains (CGP) and Flint Hills (FH), labelled respectively “9.4.2” and “9.4.4” in Figure 1(a). CGP and FH are homogeneous with respect to the quality patterns Q-GCP and Q-FH (Omernik 1987), specified in Table 2, respectively. That is, the region CGP is homogeneous with respect to all the quality universals that form the tuple Q-GCP. One quality universal of the pattern Q-GCP is Irregular plain. That CGP is homogeneous

Figure 2. Eco-regions of North America and the continental United States (EPA) at different scales (EPA, 1997).
with respect to the quality universal _Irregular plain_, however, does not mean that every part of geographic scale within _CGP_ has a land-surface form that is an instance of _Irregular plain_. There may be comparatively small (geographic-scale) regions in which different land-surface forms are instantiated (for details see Bittner 2009; specific examples of (in)homogeneities can be found in Bailey 1983, 2004; Omernik 2004; and Olson and others 2001.)

In quality-based classification and delineation systems, regions that are maximally homogeneous with respect to certain quality patterns are of particular importance (Bittner 2009). For example, all members of the collection LIII are maximally homogeneous with respect to some quality pattern (Omernik 1987; EPA 1997, 2002); in particular, the regions _CGP_ and _FH_ are maximally homogeneous with respect to the quality patterns _Q-GCP_ and _Q-FH_ shown in Table 2. The boundary between _CGP_ and _FH_ marks a discontinuity between the maximally _Q-GCP_-homogeneous region _CGP_ and the maximally _Q-FH_-homogeneous region _FH_. Further, _CGP_ is maximally homogeneous with respect to the quality _Irregular plain_ and _FH_ is maximally homogeneous with respect to the quality _Open hills_.

Of course, the exact location of the boundary that separates a region that is maximally homogeneous with respect to the quality _Irregular plain_ from a region that is maximally homogeneous with respect to the quality _Open hills_ is subject to vagueness (Rossum and Lavin 2000; Omernik 2004); that is, there is a wide range of equally

Figure 3. Land-surface, climate, biomic, and bio-geographic quality universals (Omernik 1987; Koeppen 1931; Bailey 1983; Udvardy 1975; Olson and others 2001).
good candidates for the boundary location (Keefe and Smith 1996; Bittner and Smith 2003). The same is true for the conjunction of all the qualities in Q-GCP and Q-FH. For this reason, additional local qualities and quality patterns are needed to identify more precisely the location of the boundary between CGP and FH (Bittner 2009).

As pointed out above, such local qualities and quality patterns are often specified relative to qualities of neighbouring regions (e.g., lower/higher precipitation, less relief, or more irregular; EPA 2002).

### 2.3 QUALITY PATTERN AND CLASSIFICATION

The quality-based classification of eco-region universals can be formalized in terms of maximal homogeneity with respect to quality patterns and in terms of two kinds of relations between quality patterns: the sub-pattern relation and the genus–species relation. Roughly, a sub-pattern of the quality pattern Q is a sub-tuple of Q. A species-pattern of the quality pattern Q is formed by sub-universals of the quality universals in Q, and a genus-pattern of the quality pattern Q is formed by super-universals of the quality universals in Q (see Table 1 for more precise definitions). For example, among the quality patterns Q-Do, Q-Di, Q-P, and Q-S shown in Table 3, the sub-pattern relation holds, as indicated at the bottom of Figure 4(a). In Tables 2 and 3, the quality patterns Q-GCP and Q-FH are both species-patterns of the genus-pattern Q-L (which is a sub-pattern of Q-S).

### Table 2. Some (species) quality patterns that characterize geographic regions of the North American continent (NAC) (Bailey 1983; Omernik 1987; Olson and others 2001; Bittner 2009)

<table>
<thead>
<tr>
<th>Quality Pattern for the Ecoregion</th>
<th>Symbol</th>
<th>Constituting Quality Universals (see Figure 3)</th>
<th>Genus Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prairie Division (of NAC)</td>
<td>Q-Pra</td>
<td>&lt;Humid Temperate Climate Regime, Prairie Climate Type, Lowland&gt;</td>
<td>Q-Di</td>
</tr>
<tr>
<td>Dry Domain (of NAC)</td>
<td>Q-Dry</td>
<td>&lt;Dry Climate Regime&gt;</td>
<td>Q-Do</td>
</tr>
<tr>
<td>Humid Temperate Domain</td>
<td>Q-HTe</td>
<td>&lt;Humid Temperate Climate Regime&gt;</td>
<td>Q-Do</td>
</tr>
<tr>
<td>Humid Tropical Domain</td>
<td>Q-HTr</td>
<td>&lt;Humid Tropical Climate Regime&gt;</td>
<td>Q-Do</td>
</tr>
<tr>
<td>Polar Domain</td>
<td>Q-Pol</td>
<td>&lt;Polar Climate Regime&gt;</td>
<td>Q-Do</td>
</tr>
<tr>
<td>Central Great Plains</td>
<td>Q-GCP</td>
<td>&lt;Irregular plains, Bluestem grama prairie, Cropland, Dry Mollisols&gt;</td>
<td>Q-L</td>
</tr>
<tr>
<td>Flint Hills</td>
<td>Q-FH</td>
<td>&lt;Open Hills, Bluestem prairie, Subhumid grassland, Mollisols&gt;</td>
<td>Q-L</td>
</tr>
<tr>
<td>Canadian Low Arctic Tundra</td>
<td>Q-NaTu</td>
<td>&lt;Nearctic Realm, Tundra&gt;</td>
<td>Q-238</td>
</tr>
<tr>
<td>Alaskan North Slope</td>
<td>Q-NaTu</td>
<td>&lt;Nearctic Realm, Tundra&gt;</td>
<td>Q-238</td>
</tr>
</tbody>
</table>

### Table 3. Genus quality pattern formed by quality universals shown in Figure 3 (Bittner 2009; Olson and others 2001)

<table>
<thead>
<tr>
<th>Quality Pattern Type</th>
<th>Symbol</th>
<th>Constituting Quality Universals (see Figure 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Land surface</td>
<td>Q-L</td>
<td>&lt;Land-surface form, Climax Plant Formation, Land use, Soil type&gt;</td>
</tr>
<tr>
<td>Climate</td>
<td>Q-Di</td>
<td>&lt;Climate regime&gt;</td>
</tr>
<tr>
<td>Climate + Land surface</td>
<td>Q-S</td>
<td>&lt;Climate regime, Climate type, Elevation, Climax vegetation&gt;</td>
</tr>
<tr>
<td>Biomic</td>
<td>Q-BI</td>
<td>&lt;Biomic quality&gt;</td>
</tr>
<tr>
<td>Bio-geographic</td>
<td>Q-BG</td>
<td>&lt;Biogeographic eco-region quality&gt;</td>
</tr>
<tr>
<td>Biomic + biogeographic</td>
<td>Q-238</td>
<td>&lt;Biomic quality, Bio-geographic quality&gt;</td>
</tr>
</tbody>
</table>
The eco-region classification of the USFS (Bailey 1983) is a prototypical example for distinguishing universals in terms of maximal homogeneity with respect to quality pattern and in terms of sub-pattern and genus–species relations between quality patterns (Bittner 2009).

The quality pattern $Q-S$ (see Table 3) is the defining pattern for the eco-region universal Section. That is, all instances of Section are maximally homogeneous with respect to some species-pattern of $Q-S$. Instances of the universals Province, Division, and Domain are maximally homogeneous with respect to some species-pattern of, respectively, $Q-P$, $Q-Di$, and $Q-Do$ – all of which are sub-patterns of $Q-S$. For example, instances of Domain are maximally homogeneous with respect to the species-pattern $Q-Do$. Since Climate regime is the only quality of $Q-Do$, all instances of Domain are maximally homogeneous with respect to one of the proper sub-universals of Climate regime shown in Figure 3. In particular, the quality pattern $Q-Dry$ is a species-pattern of $Q-Do$, and the quality pattern $Q-Do$ is a sub-pattern of $Q-S$.

The boundaries that separate regions that are maximally homogeneous with respect to distinct quality patterns are subject to vagueness. Since all eco-region universals in the USFS system are identified in terms of sub-patterns and species-patterns of the non-local quality pattern $Q-S$, no additional (local) qualities can be taken into account to delineate more precisely the location of the boundaries between regions with distinct quality patterns. For this reason, the resulting delineation is relatively imprecise. The high degree of vagueness affecting the boundaries is reflected in Figure 1(a) in the highly generalized ways in which boundaries are drawn.

3. The Classification and Delineation System of the WWF

The classification and delineation system of the WWF (Olson and others 2001) is one that – to a certain degree – avoids the disadvantages of two extreme positions: (a) focusing on classification based on non-local quality patterns at the cost of precision in delineation, and (b) focusing on delineation based on local quality patterns at...
the cost of the classification. This is achieved by a bottom-up approach in which members of a relatively fine and precisely delineated partition are aggregated according to non-local quality patterns.

3.1 LOCAL QUALITY PATTERNS FOR DELINEATION AT THE MOST DETAILED SCALE

The WWF system introduces a collection of 825 geographic regions (WWF_867). This collection forms a partition of the land surface of the Earth at the most detailed scale. In North America, WWF_867 includes the EPA collection LIII, a collection of eco-regions of Canada (Wiken 1986; Marshall and Schut 1999), and a collection of eco-regions of Mexico (Rzedowski 1978) as sub-collections. In its focus on local quality patterns to achieve a precise delineation, this Canadian system is similar to the EPA system (Wiken 1986; Omernik 1995). In fact, the underlying methodology was pioneered in an Environment Canada technical report (Wiken 1986).

Since the WWF system incorporates both EPA collection LIII and the corresponding Canadian delineation, it embraces the focus on precision in delineation based on local quality patterns at the most detailed scale. Because of this focus on local qualities to delineate regions precisely, there is no single quality pattern Q for WWF_867 such that all members of WWF_867 are maximally Q-homogeneous. Consequently, there is no single genus-pattern such that non-trivial sub-universals (universals with more than one instance) of geographic region (Georegion) can be identified in terms of sub-pattern and genus-species pattern of Q.

3.2 BIOME AND ECO-ZONE UNIVERSALS

The definitions of the Ecozone (also called Biogeographic Realm) and Biome universals of the WWF system are based on relations between the non-local quality patterns Q-BG, Q-BI, and Q-238, shown in Table 3. These quality patterns, in turn, are formed by the biomic and bio-geographic quality universals shown in Figure 3. In terms of homogeneity with respect to quality pattern, one can define Ecozone:

**Definition 1:** Particular x is an instance of the universal Ecozone (or Biome) if and only if (a) x is a geographic region that is the sum of members of the collection WWF_867 and (b) x is maximally homogeneous with respect to some species-pattern of Q-BG (or Q-BI).

Instances of proper sub-universals of Ecozone (or Biome) are sums of members of the collection WWF_867 that are maximally homogeneous with respect to one specific immediate species-pattern of Q-BG (or Q-BI). For example, every instance of Nearctic Ecozone is a sum of members of WWF_867 that, as a whole, is maximally homogeneous with respect to the quality (pattern) of Tundra climate type. Similarly, every instance of Tundra Ecoregion is a sum of members of WWF_867 that, as a whole, is maximally homogeneous with respect to the quality (pattern) ‘Tundra climate type’.

At an intermediate scale, the universal EcoBiome is identified:

**Definition 2:** Particular x is an instance of the universal EcoBiome if and only if (a) x is a geographic region that is a sum of members of the collection WWF_867 and (b) x is maximally Q-homogeneous with respect to some immediate species-pattern of Q-238 (see Table 3).

The WWF system does not use the term ‘EcoBiome’; the term is used here to emphasize that instances of EcoBiome are maximally homogeneous with respect to quality patterns that are formed by biomic and bio-geographic quality universals. For example, the regions that in the WWF system are called ‘Alaskan North Slope Coastal Tundra’ and ‘Canadian Low Arctic Tundra’ are instances of EcoBiome; both are maximally homogeneous with respect to the quality pattern Q-NaTu, a species-pattern of the genus-pattern Q-238 (see Tables 2 and 3). The universal EcoBiome currently has 238 instances; for this reason, the collection of all current instances of the universal EcoBiome is called WWF_238.

Ecozone, Biome, and EcoBiome are all distinct universals and sub-universals of the universal Georegion, as depicted in Figure 5(c).

3.3 SPATIAL NESTING

Instances of what I have identified here as the universal EcoBiome are smaller than instances of the universals Ecozone and Biome but (in most cases) larger than members of the collection WWF_867. Let EZC be the collection of all present instances of the universal Ecozone, and let BC be the collection of all instances of the universal Biome. The collections EZC, BC, WWF_238, and WWF_867 all partition the terrestrial surface of the Earth (Olson and others 2001; see Figure 5(a)); moreover, WWF_867 is immediately |p-included in WWF_238, and WWF_238 is immediately |p-included in both BC and EZC. BC and EZC are not |p-included in one another, since some of their members partially overlap.

4. Toward an Integrated System

The key insight to be gained from an analysis of the WWF system is that the trade-off between the need for non-local quality patterns for classification purposes and the need for local quality patterns to enable precise delineation can be overcome: in a bottom-up approach, in which the members of a precise, local-quality-based partition (the base partition) are aggregated (summed up) to larger regions with respect to non-local quality patterns, the aggregated regions form coarser partitions that p-include the
base partition. Using this insight, it is now quite natural to integrate the USFS, EPA, and WWF systems.

4.1 THE BASE PARTITION

Let $LIII_{NAC}$ be the collection of members of $WWF_{867}$ that are parts of the North American continent. $LIII_{NAC}$ is a natural choice for the base of an integrated classification and delineation system for eco-regions and eco-zones of The National Map: thanks to its extensive use of local quality patterns, the delineation is quite precise and is useful for administrative and management purposes (McMahon and others 2001; Olson and others 2001). That is, at least within the continental United States, it is possible to minimize the vagueness of the locations of the boundaries separating the various eco-regions at this scale by using local quality patterns and then crisping the remaining degrees of vagueness by fiat (Smith 1995) in a way agreed upon by the EPA, the WWF, and the USFS (McMahon and others 2001; Olson and others 2001).

It follows that the EPA collections $LI$ and $LII$, as well as the universals $Biome$, $Ecozone$, and $EcoBiome$ of the original WWF system, are automatically part of the integrated system.

4.2 INTEGRATING THE USFS SYSTEM

All members of $LIII$ are maximally homogeneous with respect to some species-pattern of the land-surface quality pattern $Q-L$ shown in Table 3 (Omernik 1987; see Table 2). That is, the delineation of regions of $LIII$ can be considered a more precise version of the original delineation of instances of Section. The increased precision is based on additional local quality patterns, as discussed above. Thus, in an integrated system, the instances of Section in the United States can be identified with the corresponding members of $LIII$. Instances of the universals $Domain$, $Division$, and $Province$ are identified using the quality patterns in the original USFS definitions, as depicted in Figure 4. In addition, these quality patterns serve as criteria for the aggregation of members of the base partition:

Definition 3: Particular $x$ is an instance of the universal $Province$ (or $Division$ or $Domain$) if and only if (a) $x$ is the sum of members of the collection $LIII_{NAC}$ and (b) $x$ is maximally homogeneous with respect to some species of the genus quality pattern $Q-P$ (or $Q-Di$ or $Q-Do$).

For example, every instance of the universal $Province$ is a sum of all the members of the collection $LIII$ that are maximally homogeneous with respect to some proper sub-universal of Climate regime. Proper sub-universals of $Province$, $Division$, and $Domain$ are defined as sums of members of $LIII$ that are maximally homogeneous with respect to a single species-pattern of $Q-P$, $Q-Di$, and $Q-Do$ respectively. Thus, sub-universals of $Domain$ are defined as displayed in Figure 4(b).

The relationships between the sub-universal relation among eco-region universals and the relations between quality patterns that serve as criteria for the summation of regions from the base partition are depicted in Figure 5(c). This taxonomic structure and the corresponding relations between quality patterns obviously correspond to the structures depicted in Figure 4.

4.3 SYMBOLIC REPRESENTATION

An important requirement for integrated classification and delineation systems for $TNM$ is that such a system not be represented only as traditional area-class maps or categorical coverage maps using conventional GIS data structures. To facilitate automated reasoning and integration into the Geo-semantic Web (Semantic Web 2009), classification and delineation systems also need to be represented symbolically in a logic-based framework. Such symbolic representations should embrace standards such as the Ontology Web Language, or OWL (Baader and others 2002; Horrocks and others 2003; Knublauch and others 2004).

The ontology-based methodology demonstrated here facilitates such symbolic representations: particulars and collections can be represented easily as individuals and enumerated classes in OWL. Universals can be represented as classes that form taxonomic trees. The relations part of, $\sqsubseteq$-u-part-of, $\sqsubseteq$-p-included-in, partition of, and so on can be represented in OWL in ways demonstrated elsewhere (Bittner and Donnelly 2005; Smith and others 2007). In particular, the important sub-universal of, $\sqsubseteq$-u-part-of, $\sqsubseteq$-p-included-in, and partition of relations can be represented very naturally in OWL as graph-like structures. Graph-like structures of the same type are used here to represent an integrated eco-region, biome, and eco-zone classification and delineation system for North America.

Let $EZC_{NAC}$, $BC_{NAC}$, and $WWF_{238}_{NAC}$ be the collections of eco-regions that have as members all those members of $EZC$, $BC$, and $WWF_{238}$, respectively, that are parts of the North American continent. Similarly, let $Dom_{NAC}$, $Div_{NAC}$, $Pr_{NAC}$, and $Sec_{NAC}$ be the collections of eco-regions that have as members all those instances of $Domain$, $Division$, $Province$, and Section, respectively, that are parts of the North American continent.

If the North American continent is identified with the Nearctic eco-zone of the WWF system, then the collection $EZC_{NAC}$ has a single member – the Nearctic Ecozone. In Figure 5(a), the graph of the relation partition-of$_{NAC}$ represents the fact that all these collections partition the North American continent.

The graph of the relation $\sqsubseteq$-p-included-in$_{NAC}$ in Figure 5(b) shows that eco-regions of sub-regional scale form the basis of the delineation (i.e., $LIII_{NAC} = Sec_{NAC}$). Since the different systems form geographic regions at
coarser scales using different kinds of quality patterns, there are different hierarchical subdivisions that do not necessarily coincide. That is, each of the partitions \( Pr_{NAC} \), \( LII \), and \( WWF_{238} \) may have members whose boundaries lie skew to the boundaries of members of the other two partitions. This is the reason for the lattice structure in Figure 5(b).

Consider the relationship between the spatial nesting of \( Georegion \) universals and the aggregation criteria used in their definitions. If \( Georegion \) universal \( E_i \) is a \( \sqsubset \)-part-of \( E_j \), then the quality pattern that serves as aggregation criterion in the definition of \( E_j \) is a sub-pattern of the quality pattern that serves as aggregation criterion in the definition of \( E_i \). This can be verified by inspecting the relations between the quality pattern in Figure 5(c) and the partonomic inclusion lattice in Figure 5(b).

5. Conclusions

A quality-based classification and delineation system with a precise delineation component and a sophisticated classi-

Figure 5. Symbolic representation of an integrated classification and delineation system: (a) partitions of the terrestrial surface of the North American continent. (b) graph of the relation \( \sqsubseteq \)-part-of \( NAC \); (c) quality-based classification and summation in the integrated system (bold arrows represent the relation \( \sqsubseteq \)-universal-of; dashed arrows indicate that all instances of the universal pointed to are sums of members/instances of the origin of the arrow; a dotted line connecting a quality pattern to a universal indicates that this quality pattern is used as a differentia in the definition of the corresponding universal).
The National Map

The underlying ontology-based framework not only simplifies the integration but also provides a formal language for the symbolic representation of quality-based classification and delineation systems. The top-level notions summarized in Table 1 provide a general framework for the formalization and symbolic representation of any kind of quality-based classification and delineation system for geographic regions. For this reason, the special case of eco-region and eco-zone classification and delineation system may provide a blueprint for how to represent quality-based classification and delineation systems in The National Map in general.

Acknowledgements

Financial support for the author to attend the UCGIS workshop Building an Ontology for The National Map is gratefully acknowledged.

Author Information

Thomas Bittner, PhD, is Assistant Professor in the Departments of Philosophy and Geography, State University of New York at Buffalo, NY USA. E-mail: bittner3@buffalo.edu. He is also a research scientist at the New York State Center of Excellence in Bioinformatics and Life Sciences and the National Center for Geographic Information and Analysis (NCGIA). His area of specialization is formal ontology and its applications in geography, GIScience, bio-informatics, and other fields. His current research focuses applying formal ontology, symbolic logic, and qualitative representation and reasoning techniques to vague geographic phenomena.

References


