Understanding Taxonomies of Ecosystems: a Case Study

Alexandre Sorokine[1] and Thomas Bittner[2]

[1] Geography Department, University at Buffalo sorokine@buffalo.edu
[2] Institute of Formal Ontology and Medical Information Systems, Leipzig thomas.bittner@ifomis.uni-leipzig.de

Abstract

This paper presents a formalized ontological framework for the analysis of classifications of geographic objects. We present a set of logical principles that guide geographic classifications and then demonstrate their application on a practical example of the classification of ecosystems of Southeast Alaska. The framework has a potential to be used to facilitate interoperability between geographic classifications.

1 Introduction

Any geographic map or a spatial database can be viewed as a projection of a classification of geographic objects onto space [1]. Such classification can be as simple as a list of objects portrayed on the map or as complex as a multi-level hierarchical taxonomy such as used in the areas of soil or ecosystem mapping [2]. However, in any of these cases classification is predicated on a limited set of rules that ensure its consistency within itself and what it is projected on. Classifications of geographic objects, if compared to classifications in general, have certain peculiarities because geographic objects inherit many of their properties from the underlying space. Classifications of geographic objects typically manifest themselves as map legends.

The goal of this paper is to develop a formalized framework for handling of the structure of and operations on classifications of geographic objects. There are three groups of purposes for development of this framework: (1) such a framework would allow better understanding of the existing classification systems and underlying principles even for non-experts, (2) the framework can provide useful tips for developing new classification systems with improvements in terms of consistency and generality, (3) the framework would allow

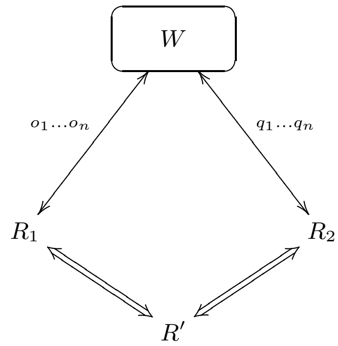


Fig. 1. Interoperability through representation models

more flexibility for achieving interoperability and fusion between datasets employing non-identical classification systems.

Each classification is a representation (representations are denoted by letters R_1 and R_2 on Fig. 1) of the real world W that was created using a unique and finite chain of operations ($o_1 \ldots o_n$ and $q_1 \ldots q_n$ respectively). The goal of our research is to outline operations $o_1 \ldots o_n$ and $q_1 \ldots q_n$ in a clear formal, understandable and non-ambiguous manner. This knowledge will allow us to build a new representation R' that would be able to accommodate both R_1 and R_2 thus achieving interoperability (shown on the diagram as double arrows) between them and possibly with other representations.

2 Classification of Ecological Subsections of Southeast Alaska

To demonstrate our theory we will use the classification of ecological subsection of Southeast Alaska [3] as a running example (Fig. 2). This classification was developed by the USDA Forest Service and it subdivides the territory of Southeast Alaska into 85 subsections that represent distinct terrestrial ecosystems. The purpose of the classification is to provide a basis for practical resource management, decision making, planning, and research.

The classification has three levels that are depicted in Table 1. The first level (roman numerals in Table 1) subdivides the territory into three terrain classes: active glacial terrains, inactive glacial terrains and post-glacial terrains. At the next level (capital letters in Table 1) the territory is subdivided according to its physiographic characteristics. The third level of the classification (numbers in Table 1) divides territory by lithology and surface deposits.

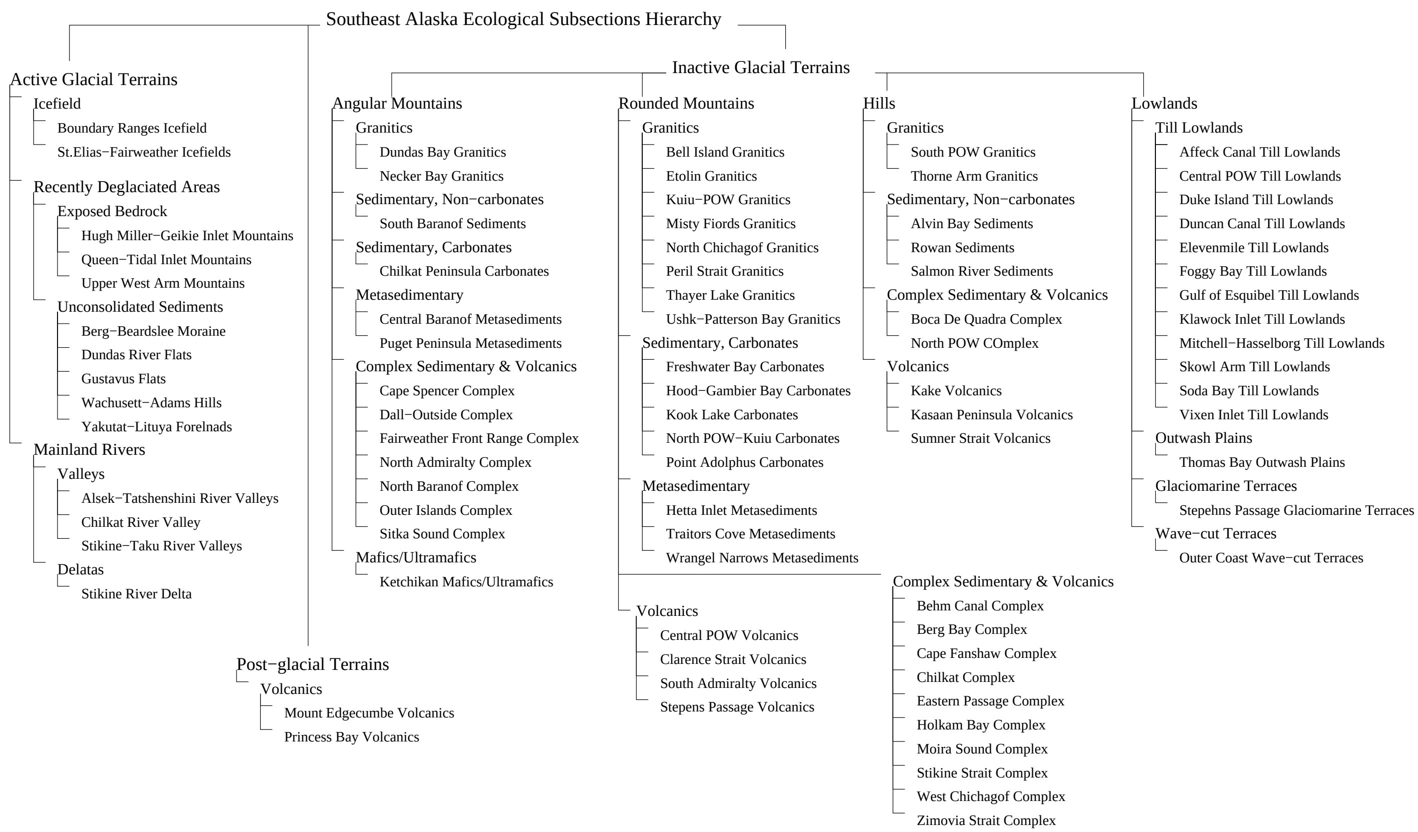


Fig. 2. Southeast Alaska Ecological Subsection Hierarchy [3, pp. 22–23]

Table 1. Hierarchical Arrangement of Subsections [3, Table 2, p. 16]

- I. Active Glacial Terrains
  - A. Icefields
  - B. Recently deglaciated areas
    - 1. Exposed Bedrock
    - 2. Unconsolidated sediments
  - C. Mainland rivers
    - 1. Valleys
    - 2. Deltas
- II. Inactive Glacial Terrains
  - A. Angular Mountains
    - 1. Granitics
    - 2. Sedimentary, Noncarbonates
    - 3. Sedimentary, Carbonates
    - 4. Meta-sedimentary
    - 5. Complex sedimentary & volcanics
    - 6. Mafics/Ultra-mafics
  - B. Rounded Mountains
    - 1. Granitics
    - 2. Sedimentary, Carbonates
    - 3. Meta-sedimentary
    - 4. Complex sedimentary & volcanics
    - 5. Volcanics
  - C. Hills
    - 1. Granitics
    - 2. Sedimentary, Carbonates
    - 3. Meta-sedimentary
    - 4. Complex sedimentary & volcanics
    - 5. Volcanics
  - D. Lowlands
    - 1. Till Lowlands
    - 2. Outwash Plains
    - 3. Glaciomarime Terraces
    - 4. Wave-cut Terraces
- III. Post-glacial Terrains
  - A. Volcanics

Roman numerals	Terrains classes
Capital letters	Physiographic classes
Numbers	Geologic classes

3 A Formal Theory of Classes and Individuals

In this section we will introduce logical theories that are needed to formalize relations behind ecosystem classifications and demonstrate their application using the classification of ecological subsection of Southeast Alaska (Table 1 and Fig. 2) as a running example. Formalization of the theories will be presented using first order predicate logic with variables $x, y, z, z_1, \ldots$ ranging over individuals and variables $u, v, w, w_1, \ldots$ ranging over classes. Predicates always begin with a capital letter. The logical connectors $\neg, =$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\equiv$ have their usual meanings: not, identical-to, and, or, ‘if …then’, ‘if and only if’ (iff), and ‘defined to be logically equivalent’. We write (x) to symbolize universal quantification and $(\exists x)$ to symbolize existential quantification. Leading universal quantifiers are assumed to be understood and are omitted.

Strict distinction between classes and individuals is one of the cornerstones of our theory. Typically classifications and map legends show only classes. However, the diagram on Fig. 2 shows a mix of classes and individuals. In our understanding ecological subsections, i.e., such entities as “Behm Canal Complex”, “Summer Strait Volcanics”, “Soda Bay Till Lowlands” and others leafs of the hierarchy, are individuals. All other entities that are not leafs (e.g., “Active glacial terrains”, “Granitics”, “Volcanics”, etc.) are classes. In the same sense Table 1 shows only classes of the classification.

3.1 The Tree Structure of Classes

Examples of classes are the class *human being*, the class *mammal*, the class *ecosystems of the polar domain*, the class *inactive glacial terrains*, etc. Classes are organized hierarchically by the *is-a* or the *subclass* relation in the sense that a male human being *is-a* human being and a human being *is-a* mammal, or, using our example, "Exposed Bedrock" is-a "Deglaciated Area". In the present paper the is-a or subclass relation is denoted by the binary relation symbol $\sqsubseteq$ and we use symbol $\sqsubset$ for the proper subclass relation. We will write $u \sqsubseteq v$ to say that class u is involved in the subclass relation with class v. Also we will call v a superclass of u if the relation $u \sqsubseteq v$ holds.

The proper subclass relation is asymmetric and transitive (ATM1–2). It very closely corresponds to the common understanding of the is-a (kind-of) relations:

$$\text{(ATM1) } u \sqsubset v \rightarrow \neg v \sqsubset u$$
$$\text{(ATM2) } (u \sqsubset v \wedge v \sqsubset w) \rightarrow u \sqsubset w$$

Axiom ATM1 postulates that if u is a proper subclass of v then v is not a proper subclass of u. Transitivity (ATM2) implies that all proper subclasses of a class are also proper subclasses of the superclass of that class. In our example (Table 1) class "Exposed Bedrock" is a proper subclass of "Recently Deglaciated Areas" that in turn is a proper subclass of "Active Glacial Terrains". Due to the transitivity of the proper subclass relation we can say that class "Exposed Bedrock" is also a proper subclass of the class "Active glacial terrains".

Next we define the relations of subclass $\sqsubseteq$ as $D_{\sqsubseteq}$. Unlike proper subclass, subclass relation allows for a class to be a subclass of itself:

$$D_{\sqsubseteq} \; u \sqsubseteq v \equiv u \sqsubset v \vee u = v$$

One can then prove that the subclass relation ($\sqsubseteq$) is reflexive, antisymmetric and transitive, i.e., a partial ordering.

Class overlap ($O_{\sqsubseteq}$) is defined as $D_{O_{\sqsubseteq}}$:

$$D_{O_{\sqsubseteq}} \; O_{\sqsubseteq} \; uv \equiv (\exists w)(w \sqsubseteq u \wedge w \sqsubseteq v)$$

Classes overlap if there exists a class that is a subclass of both classes, e.g., in Table 1 class "Icefields" overlaps with class "Active Glacial Terrains".

We now add the definitions of a root class and an atomic class (atom). A class is a root class if all other classes are subclasses of it (D_{root}). A class is an atom if it does not have a proper subclass (D_A):

$$\begin{array}{lll} D_{root} & \text{root } u & \equiv (\forall v)(v \sqsubseteq u) \\ D_A & A\ u & \equiv \neg(\exists v)(v \sqsubset u) \end{array}$$

In our example the root class of the classification would be a class of all Southeast Alaska ecological subsections (Fig. 2). Geologic classes designated with numbers in Table 1 are atoms because they do not have any proper subclasses. In practice in many classifications root classes are not specified explicitly however their existence has to be implied. For example, Table 1 does not contain a root class but it can be inferred from the context that the root class is "Southeast Alaska ecological subsections".

Since the hierarchy formed by classes of Southeast Alaska ecological subsections are the result of a scientific classification process we can assume that the resulting class hierarchy forms a tree. We are justified to assume that scientific classifications are organized hierarchically in tree structures since scientific classification employs the Aristotelean method of classification.

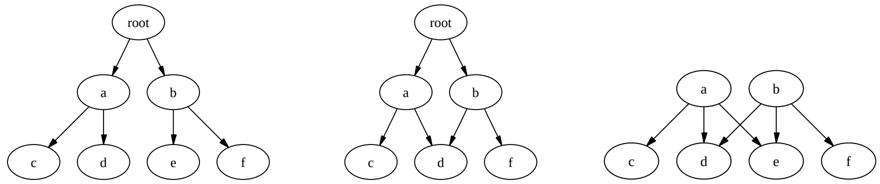
As [4] point out, in the Aristotelian method the definition of a class is the specification of essence (nature, invariant structure) shared by all instances of that class. Definitions according to Aristotle's account are specified by (i) working through a classificatory hierarchy from the top down, with the relevant topmost node or nodes acting in every case as undefinable primitives. The definition of a class lower down in the hierarchy is then provided by (ii) specifying the parent of the class (which in a regime conforming to single inheritance is of course in every case unique) together with (iii) the relevant differentia, which tells us what marks out instances of the defined class or species within the wider parent class or genus, as in: human = rational animal, where rational is the differentia (see also [5] for more details.)

Now we have to add axioms that enforce tree structures of the form shown in Fig. 3(a) and which rule out structures shown in Figs. 3(b) and 3(c). These additional axioms fall into two groups, axioms which enforce the tree structure and the finiteness of this structure respectively. We start by discussing the first group.

Firstly, we demand that there is a root class (ATM3). Secondly, we add an axiom to rule out circles in the class structure: if two classes overlap then one is a subclass of the other (ATM4). This rules out the structure in Fig. 3(b) and also it is very much true for our running example: all overlapping classes in Table 1 are subclasses of each other. Thirdly, we add an axiom to the effect that if u has a proper subclass v then there exists a class w such that w is a subclass of v and w and u do not overlap (ATM5). This rules out cases where a class has a single proper subclass or a chain of nested proper subclasses. Following [6] we call ATM5 the *weak supplementation principle*.

$$
\begin{array}{ll}
\text{(ATM3)} & (\exists u)\mathrm{root}(u) \\
\text{(ATM4)} & O_{\sqsubseteq}\ uv \rightarrow (u \sqsubseteq v \vee v \sqsubset u) \\
\text{(ATM5)} & u \sqsubset v \rightarrow (\exists w)(w \sqsubset u \wedge \neg O_{\sqsubseteq}\ wv)
\end{array}
$$

Upon inspection, the classification in Table 1 violates the weak supplementation principle (axiom ATM5) because the class of "Post-glacial terrains" has only a single subclass "Volcanics". For this reason classification on Table 1 is not a model of our theory. We will discuss this case in detail in Sect. 4 and show that the underlying classification does satisfy our axioms but that additional operations have been performed on these structures.



(a) Proper tree (b) Overlaps (c) Multiple roots

Fig. 3. Trees (a) and non-trees ((b) and (c))

Using $\mathrm{D}_{\mathrm{root}}$, the antisymmetry of $\sqsubseteq$, and ATM4 we can prove uniqueness of the root class. This rules out the structure shown in Fig. 3(c).

The second group of axioms that characterizes the subclass relation beyond the properties of being a partial ordering are axioms which enforce the finiteness of the subclass-tree. ATM6 ensures that every class has at least one atom as subclass. This ensures that no branch in the tree structure is infinitely long. Finally ATM7 is an axiom schema which enforces that every class is either an atom or has only finitely many subclasses. This ensures that class trees cannot be arbitrary broad.

$ATM6$ $(\exists y)(A\ y \wedge y \sqsubseteq x)$
$ATM7$ $\neg A\ y \rightarrow (\exists x_1, \ldots, x_n)((\bigwedge_{1\leq i\leq n} x_i \sqsubset y) \wedge (z)(z \sqsubset y \rightarrow \bigvee_{1\leq i\leq n} z = x_i))$

Here $(\bigwedge_{1\leq i\leq n} x_i \sqsubset y)$ is an abbreviation for $x_1 \sqsubset y \wedge \ldots \wedge x_n \sqsubset y$ and $\bigvee_{1\leq i\leq n} z = x_i$ for $x_1 = z \vee \ldots \vee x_n = z$.

3.2 Classes and Individuals

Classes and individuals are connected with the relationship of instantiation *InstOf* xu which first parameter is an instance and which second parameter is a class. *InstOf* xu then is interpreted as 'individual x instantiates the class u'. From our underlying sorted logic it follows that classes and individuals form disjoint domains, i.e., there cannot exist a entity which is a class as well as an individual. Therefore instantiation is irreflexive, asymmetric, and non-transitive.

In terms of our theory each individual (an ecological subsection) instantiates a class of the subsection hierarchy. A single class can be instantiated by several individuals. For example, we can say that individuals "Behm Canal Complex", "Berg Bay Complex" and others instantiate the class of "Rounded Mountains".

Axiom (ACI1) establishes the relationships between instantiation and the subclass relation. It tells us that $u \sqsubseteq v$ if and only if every instance of u is also an instance of v.

$$ACI1\ (u \sqsubseteq v \leftrightarrow (x)(InstOf\ xu \rightarrow InstOf\ xv))$$
$$TCI1\ (u = v \leftrightarrow (x)(InstOf\ xu \leftrightarrow InstOf\ xv))$$

From (ACI1) it follows that two classes are identical if and only if they are instantiated by the same individuals.

Finally we add an axiom that guarantees that if two classes share an instance then one is a subclass of the other (AI2).

$$AI2\ (\exists x)(InstOf\ xu \wedge InstOf\ xv) \rightarrow (u \sqsubseteq v \vee v \sqsubset u)$$

AI2 can be illustrated using the following example: classes "Inactive Glacial Terrains" and "Rounded Mountains" share instance "Kook Lake Carbonates" (Fig. 2) and "Rounded Mountains" is a subclass of "Inactive Glacial Terrains".

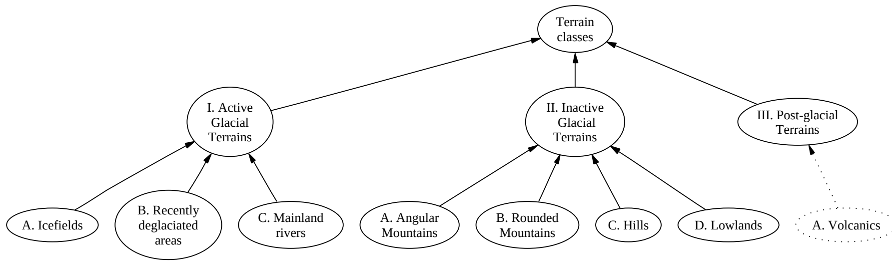
4 Applying the Theory to Multiple Classifications

As it was mentioned in Sect. 3.1, Southeast Alaska ecological subsections hierarchy (Fig. 2 and Table 1) does not satisfy one of the axioms of our classification theory: the weak supplementation principle (ATM5). It is easy to notice that at the third level of the classification on (Table 1) contains repeating classes. For example, class "Granitics" can be found under the classes "Angular Mountains", "Rounded Mountains" and "Hills" in the class "Inactive Glacial Terrains". Given this, it is possible to interpret classification on Fig. 2 as a *product of two independent classification trees*: classification of terrains (terrain classes and physiographic classes in Table 1) and classification of lithology and surface geology (geologic classes in Table 1). The hierarchies of classes that represented these two separate classifications are shown on Fig. 4(a) and Fig. 4(b) respectively. The product of these classifications is depicted in Table 2, with terrain classes as columns and geologic classes as rows. Each cell of the table contains the number of individuals that instantiate corresponding classes of both hierarchies.

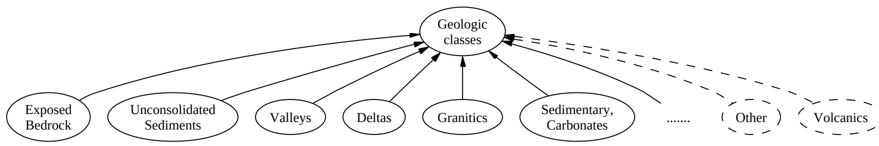
Class hierarchies on Fig. 4 contain two differences from the original hierarchy in Table 1. The class "Volcanics" that violates the weak supplementation principle was moved from the terrains hierarchy into the geologic classes hierarchy (Fig. 4 and Table 2). This is a more natural place for this class because there is already a class with an identical name.

Another problematic class is "Icefields". It is an atomic class and does not have any subclasses. Also it does not represent any geologic class. To be able to accommodate class "Icefields" in the product of classifications we have added a new class "Other" to the hierarchy of geologic classes (Fig. 4(b) and Table 2).

By using a product of two classifications we have avoided the problem of having a class with a single proper subclass. The remaining part of this section describes how a product of two or more classifications can be formalized.



(a) Terrain classes (class "Volcanics" violates the weak supplementation principle)



(b) Geologic classes (some classes are not shown, class "Volcanics" was added from Terrain classes hierarchy)

Fig. 4. Classification trees

4.1 From Theory to Models

The theory presented in the previous section gives us a formal account of what we mean by a classification tree and by the notion of instantiation. In this section we now consider set-theoretic structures that satisfy axioms given above. This means that we interpret classes as sets in such a way that the instance-of relation between instances and classes is interpreted as the element-of relation between an element and the set it properly belongs to and we interpret the is-a or subcell relation as the subset relation between sets. Sets satisfying our axioms then are hierarchically ordered by the subset relation then can be represented using directed graph structures in such a way that sets are nodes on the graph. Formally a graph is a pair $\mathcal{T} = (N, E)$ where N is a collection of nodes and E is a collection of edges. Let n_i and n_j nodes be nodes in N corresponding to the sets i and j then we have a directed edge e in E from n_i to n_j if and only if the set i is a subset of the set j. Since the sets we consider are assumed to satisfy the axioms given in the previous section it follows that the directed graph structures constructed in this way are trees and we call them *classification trees*.

We now are interested in classification trees, operations between them, and the interpretation of those operations in our running example. In particular

Table 2. The product of terrain and geology classifications

	Terrain classes							
Glaciation phases	Active Glacial Terrains			Inactive Glacial Terrains				Post-Glacial Terrains
Physiographic classes / Geologic classes	Icefields	Recently Deglaciated Areas	Mainland Rivers	Angular Mountains	Rounded Mountains	Hills	Lowlands	
Other	2							
Exposed Bedrock		3						
Unconsolidated Sediments		5						
Valleys			3					
Deltas			1					
Granitics				2	8	2		
Sedimentary, Noncarbonates				1		3		
Sedimentary, Carbonates				1	5			
Meta-sedimentary				2	3			
Complex Sedimentary and Volcanics				7	10	2		
Mafics, Ultramafics				1				
Volcanics					4	3		2
Till Lowlands							12	
Outwash Plains							1	
Glaciomarine Terraces							1	
Wave-cut Terraces							1	

we will use the notion of classification tree in order to formalize the notion of product between classification discussed in the introduction of this section.

Set theory allows us to also form higher order sets, i.e., sets of sets. In what follows we will consider levels of granularities that are sets of sets in the given interpretation. For example, the set of all leafs in a classification tree is a level of granularity. Since in our interpretation nodes in the tree structure correspond to sets levels of granularity are sets of sets. Below we will introduce the notion of a *cut* in order to formalize the notion of level of granularity. Notice that in the case of higher order sets the element-of relation is not interpreted as an instance-of relation.

4.2 Cuts

Classification trees can be intersected at their cuts. To define a cut (δ) let us take a tree $\mathcal{T}$ constructed as described above and let N be the set of nodes in this tree.

Definition 1. *A* cut δ *is a* cut *in the tree-structure* $\mathcal{T}$ *is a subset of* N *defined inductively as follows [7, 8]:*

(1) $\{r\}$ *is a cut, where* r *is the root of the tree;*
(2) For any class z *let* $\mathsf{d}\ (z)$ *denote the set of immediate subclasses of* z *and let* C *be a cut with* $z \in C$ *and* $\mathsf{d}\ (z) \neq \emptyset$*, then* $(C - \{z\}) \cup \mathsf{d}\ (z)$ *is a cut.*

For example, the hierarchy of terrain classes (Fig. 4(a) without class “Volcanics”) has five different cuts. By Definition 1 the root class “Terrain classes” is a cut. If we assume that class “Terrain classes” is z and C is a cut then $z \in C$. Immediate descendants $\mathsf{d}\ (z)$ of the root class z are classes $\mathsf{d}\ (z) = \{$ “I. Active Glacial Terrains”, “II. Inactive Glacial Terrains”, “III. Post-glacial Terrains”$\}$. Then $\mathsf{d}\ (z)$ will be the next cut because in this case $(C - \{z\}) = \emptyset$ thus $(C - \{z\}) \cup \mathsf{d}\ (z)$ leaves us with $\mathsf{d}\ (z)$. Repeated application of Definition 1 to the hierarchy of terrain classes (Fig. 4(a)) results if five cuts that are listed in Table 3.

Table 3. Examples of cuts in the hierarchy of terrain classes on Fig. 4(a)

1. the root class “Terrain classes”
2. “I. Active Glacial Terrains”, “II. Inactive Glacial Terrains”, “III. Post-glacial Terrains”
3. “A. Icefields”, “B. Recently deglaciated areas”, “C. Mainland rivers”, “A. Angular Mountains”, “B. Rounded Mountains”, “C. Hills”, “D. Lowlands”, “III. Post-glacial Terrains”
4. “I. Active Glacial Terrains”, “A. Angular Mountains”, “B. Rounded Mountains”, “C. Hills”, “D. Lowlands”, “III. Post-glacial Terrains”
5. “A. Icefields”, “B. Recently deglaciated areas”, “C. Mainland rivers”, “II. Inactive Glacial Terrains”, “III. Post-glacial Terrains”

Using Definition 1 and (ATM1–7) one can prove that the classes forming a cut are pair-wise disjoint and that cuts enjoy a weak form of exhaustiveness in the sense that every class–node in N is either a subclass or a superclass of some class in the cut δ at hand [7].

4.3 Joining Classification Trees

Let δ_1 and δ_2 be cuts in two different classification trees $\mathcal{T}_1$ and $\mathcal{T}_2$. Cuts are sets that are composed of classes that satisfy conditions on Definition 1. Let

$\delta_1 = \{u_1, u_2, \dots, u_n\}$ and $\delta_2 = \{v_1, v_2, \dots, v_m\}$. The cross-product $\delta_1 \times \delta_2$ of these cuts can be represented as a set of pairs that can be formed by classes in δ_1 and δ_2:

$$\delta_1 \times \delta_2 = \left\{ \begin{matrix} (u_1, v_1) & (u_1, v_2) & \cdots & (u_1, v_m) \\ (u_2, v_1) & (u_2, v_2) & \cdots & (u_2, v_m) \\ \vdots & \vdots & \ddots & \vdots \\ (u_n, v_1) & (u_n, v_2) & \cdots & (u_n, v_m) \end{matrix} \right\}$$

The product of two classification trees on Fig. 4 is depicted in Table 2. A product of two classification trees (or their cuts) will produce $N = n \times m$ pairs where n and m are the number of classes in the respective cuts. Most likely N will be greater than the number of classes that can be instantiated by the instances. In our example most of the cells of the Table 2 are empty indicating that there are no individuals that instantiate classes in either classification tree. The reason for it is that certain higher-level classes do not demonstrate as much diversity on the studied territory as other classes do. For example, class “Post-glacial terrains” is represented with only a single subclass “Volcanics” while class “Inactive glacial terrains” contains a total of 20 subclasses (Table 1).

To remove empty pairs of classes one has to *normalize* the product of two classifications, i.e., one has to remove the pairs of classes that do not have instances. Normalized product $\delta_1 \star \delta_2$ can be formally defined as the cross product of levels of granularity which yields only those pairs which sets have at least one element in common ($D_\star$).

$$D_* \; \delta_1 \star \delta_2 \leftrightarrow (u_i, v_i) \in \delta_1 \times \delta_2 \wedge (\exists x)(x \in u_i \wedge x \in v_i)$$

In this case each individual instantiates several classes each belonging to a different classification tree. In our example each individual instantiates one class from the classification tree of terrains and another class from the tree of geologic classes. For instance, ecological subsection “Thorne Arm Granitics” instantiates class “Hills” from the terrains classification tree and class “Granitics” from geologic classes.

5 Conclusions

In this paper we have presented an ontological framework to dissect and analyze geographic classifications. The framework is based on a strict distinction between the notion of a class and the notion of an individual. There are two types of relations in the framework: subclass relation defined between classes and instantiation relation defined between individuals and classes. Subclass relation is reflexive, antisymmetric and transitive. Classes are organized into classifications that form finite trees (directed acyclic graphs). The latter is achieved by requiring a classification to have a root class, prohibiting loops and classes with a single proper subclass.

We have demonstrated how a practical classification can be built using outlined principles on the example of Southeast Alaska ecological subsection hierarchy. Practical classification may require to employ additional operations such as a product of classifications and removal of some classes due to redundancy.

Even though most of the operations in our approach would seem obvious for geographers and ecologists, such operations have to be outlined explicitly if the goal of interoperation of two datasets is to be achieved or the information contained in classifications is to be communicated to non-experts in the area.

The formalized theory presented above can be used to facilitate interoperability between geographic classifications. Interoperability between classifications can be achieved by creating a third classification capable of accommodating of both of the original classifications. The formal theory of classes and individuals can be used to mark out the elements of classifications such as classification trees, cuts, products and normalized products. Original classifications have to be decomposed into these elements and then these elements have to be reassembled into a new and more general classification.

In a hypothetical example to interoperate Southeast Alaska ecological subsection with a similar hierarchy for some other region, we must first perform the operations outlined in Sect. 3 (mark out classification trees, cuts and their products) for both classifications. Then the trees from the different classifications would have to be combined. Territories with dissimilar geologic histories would posses different sets of classes and more diverse territories will have a larger number of classes. For instance, let us assume that the second hierarchy in our example was developed for a territory only partly affected by glaciation. Glaciation-related classes from Southeast Alaska are likely to be usable in that territory too. However, classifications for that territory will contain many classes that would not fit into the class trees specific for Southeast Alaska. Those classes would have to be added to the resulting classification trees. Most of these would fall under the "Post-glacial Terrains" class of Southeast Alaska hierarchy. Combined classification trees must satisfy axioms ATM1–7. Finally, a normalized product of class trees will have to be created. This methodology still awaiting testing in practical context.

Acknowledgments

The authors would like to thank Dr. Gregory Nowaki from the USDA Forest Service for useful and detailed comments on our paper. Dr. Nowaki is one of the leading developers of the classification of Southeast Alaska ecological subsections [3]. Support for the second author from the Wolfgang Paul Program of the Alexander von Humboldt Foundation and from National Science Foundation Research Grant BCS–9975557: Geographic Categories: An Ontological Investigation, is gratefully acknowledged

References

[1] E. Lynn Usery. Category theory and the structure of features in geographic information systems. *Cartography and Geographic information systems*, 20(1):5–12, 1993.

[2] David T Cleland, Peter E Avers, W Henry McNab, Mark E Jensen, Robert G Bailey, Thomas King, and Walter E Russell. National hierarchical framework of ecological units. In Mark S. Boyce and Alan W. Haney, editors, *Ecosystem Management Applications for Sustainable Forest and Wildlife Resources*, pages 181–200. Yale University Press, New Haven, CT, 1997.

[3] Gregory Nowaki, Michael Shephard, Patricia Krosse, William Pawuk, Gary Fisher, James Baichtal, David Brew, Everett Kissinger, and Terry Brock. Ecological subsections of southeastern Alaska and neighboring areas of Canada. Technical Report R10-TP-75, USDA Forest Service, Alaska Region, October 2001.

[4] Barry Smith, Jacob Khler, and Anand Kumar. On the application of formal principles to life science data: A case study in the gene ontology. In E Rahm, editor, *Database Integration in the Life Sciences (DILS 2004)*. Springer, Berlin, 2004. forthcoming.

[5] Harold P Cook and Hugh Tredennick. *Aristotle: The Categories, On Interpretation, Prior Anlytics*. Harvard University Press, Cambridge, Massachusetts, 1938.

[6] Peter Simons. *Parts, A Study in Ontology*. Clarendon Press, Oxford, 1987.

[7] P. Rigaux and M. Scholl. Multi-scale partitions: Application to spatial and statistical databases. In M. Egenhofer and J. Herring, editors, *Advances in Spatial Databases (SSD'95)*, number 951 in Lecture Notes in Computer Science, pages 170–184. Springer-Verlag, Berlin, 1995.

[8] T. Bittner and J.G. Stell. Stratified rough sets and vagueness. In W. Kuhn, M. Worboys, and S. Timpf, editors, *Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science. International Conference COSIT'03*, pages 286–303. Springer, 2003.