

COMPLEX TOPOGRAPHIC FEATURE ONTOLOGY PATTERNS

Dalia Varanka and Thomas J. Jerris

U.S. Geological Survey, 1400 Independence Road, Rolla, MO, USA – (dvaranka, tjerris)@usgs.gov

KEYWORDS: ontology patterns, complex features, topographic features, linked-data model

ABSTRACT:

Semantic ontologies are examined as effective data models for the representation of complex topographic feature types. Complex feature types are viewed as integrated relations between basic features for a basic purpose. In the context of topographic science, such component assemblages are supported by resource systems and found on the local landscape. Ontologies are organized within six thematic modules of a domain ontology called Topography that includes within its sphere basic feature types, resource systems, and landscape types. Context is constructed not only as a spatial and temporal setting, but a setting also based on environmental processes. Types of spatial relations that exist between components include location, generative processes, and description. An example is offered in a complex feature type 'mine.' The identification and extraction of complex feature types are an area for future research.

1. INTRODUCTION

Complex feature types are assemblages of multiple basic feature components that depend on each other for functional or some other meaningful purposes. Assemblages may be multiple occurrences of a single feature type; for example, expanses of tree area form woodlands. Other complex feature types can be combinations of different feature types and associated geometries, such as the control tower points, runway lines, and building areas of airports. Complexes such as these are difficult to define in current remote sensing data collection modalities because of the blurring of multi-band signatures or the overlap of features when viewed from above in aerial photographs. Traditional geometric data models, such as points, lines, and areas, are well suited to represent simple, basic feature types, such as 'lake,' 'road,' or 'location point,' but most geographic information systems GIS cannot flexibly enable geometric feature combinations. Basic features can be associated in GIS, however, as simple or complex relationship classes or achieved through complex processing, such as map algebra (ESRI).

This study examines the potential of using ontology design patterns (ODP) to represent complex topographic feature types. The broad meaning of ontology is to model a perspective of real-world cognition. ODP map user perspective models to existing models of the database. ODP are formal complex assemblages of feature concepts and their interrelations in the linked triple data format for reuse in multiple applications (Dago and others, 2005; ontologydesignpatterns.org). Identical nodes of triples link together or are linked by relations to form network graphs in a data model called Resource Description Framework, or RDF (W3C, 2010). Ontology patterns are appropriate models for topographic geosemantic data because topographic data primarily serve as a base for diverse applications by the public and scientists; no specific single use is determined; the design of patterns enables their reuse as baseline applications. An example of an ontology pattern appears in Figure 1.

Maps are sometimes considered to be geospatial ontologies, but complex features on maps are subject to the constraints of cartographic scale. As topographical data changed media from paper maps to digital databases, features became identified as segments of a data model, and not as entities within the representational context of places on the map. ODP can model complex features and their relation to their context. Their representation is constrained by operational scale, scale meant to represent geographical processes (Lam and others, 2005).

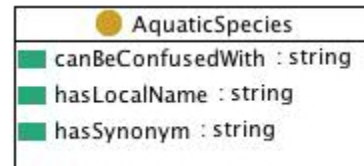


Figure 1. SpeciesNames is an ODP developed by Aldo Gangemi (2010) to express the terminological variants and the conceptual similarity that can be sources of confusion between species.

The approach that is described in this paper is to show how ontology patterns can be built for topographic data using simple feature types categorized taxonomically within thematic domain modules of a central ontology called 'Topography.' For example, terrain, surface water, and ecological regimes are broad thematic modules that interrelate to form a concept of topography defined as the character of the surface of the earth. Basic features are related across modules with different spatial relation types, building contexts of different operational scales. An example is devised using a complex feature type, 'mine.' The paper ends with a discussion of some implications of the ODP model for geographic information handling.

2. ONTOLOGIES FOR TOPOGRAPHIC COMPLEX FEATURES

2.1 Complex Features

For ontology design purposes, the ontology of topography is a primary-theory ontological perspective in which geospatial features are considered real and rooted in human common-sense experience of the world that is perceived in order to function in everyday life (Smith and Mark, 2003). Feature properties may vary with cultural perspectives or individual perceptions, but an assumption of real objects accompanies ontological concepts. Space can be represented in various ways, such as geometric systems and coordinates with each entity assigned a location, or employing relative ways, such as topologic or process-based location. Some data are organized as seamless fields to be downloaded from servers. When downloaded, the data are partitioned into regional extents that are expressed in terms of coordinates, but users usually add relative spatial relations to other localized features resembling objects. This process resembles the ontology of the field/object relations (Mark and Smith, 2004).

Spatial entity types are mixed between material, such as topographic features, and immaterial, such as science processes and descriptions. Feature boundaries are mostly clearly marked as a result of technological constraints, but fuzzy as conceptual entities. Space and time can be differentiated, as in static representations such as maps, or they can be combined, as in hydrological modeling. Distinctions between physical and socio-economic entity types are not considered here because of multiple perspectives on topography; for example, a garden can be considered a physical, ecological entity type, or a socio-economic asset to a residence. A clearer distinction can be made between bona fide and fiat entity types (Casati and Varzi, 1999). Bona fide entities are generated by physical forces and fiat types are determined by humans. Complex features involve both types.

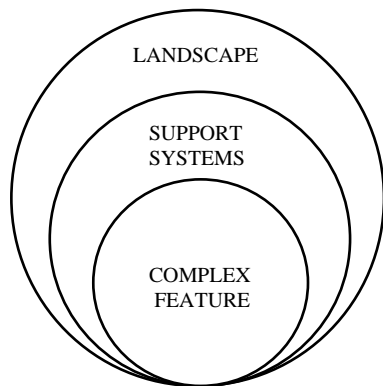


Figure 2. Complex features type, supporting systems, and affected landscape extent

To better serve advanced analyses, such as those required by science or intelligence, ontologies integrate discrete data for complex systems analysis by leveraging the inference, or automatic connection between elements, that the linked-data model of triples and network graphs allows. Ontologies of

complex features control the connection of elements between classification systems by applying logical reasoning through the use of Web Ontology Language (OWL). In addition to the feature components, connections are drawn to broader supporting systems, such as infrastructure or resource systems. All complex features occur on the landscape, the representation of which is commonly omitted in conceptual models (Figure 2). The inter-relation of these three levels of elements, features, systems, and landscape, build a context of place, a significant aspect of complex features. The context of place is conceptualized in ontology patterns of interrelated topographic domains. To model landscape processes, cartographic abstraction is expanded and extended in linked data systems using physical and ecological interpretation to simulate normally linguistic transformations.

Activity plays a critical role in topography (Kuhn, 2001). The linked-data model leverages dynamic information found in social networks and behavior. Intelligence applications of complex features include uncommon feature types, such as 'person' or 'place name' (National Geospatial-Intelligence Agency, 2007). The inclusion of humans as components in feature model research, with attributes such as cultural behavior, social interaction, and family life, combines social networking and geographical knowledge and extends the complexity of topographic ODP (Gruber, 2007).

A complex feature can display necessary and sufficient conditions for the meaning of its existence or it can display an extensive number of components of variable types and functional roles. Details are typically driven by a specific context, and share a common purpose or objective. Geographical context can be defined simply as the representational time and place of a particular setting. One common way to represent geographical scale is to use the classification taxonomy as a form of greater or lesser detailed representation with few or more instances. For richer modeling, operational scale will depend on the spatial relation properties modeled in the ontology and the semantics they convey.

2.2 Topographic Science Ontologies

Six subject domains fall within the scope of an ontology of topographic science called 'Topography': Terrain, Surface Water, Ecological Regimes, Built-Up Areas, Divisions, and Events. These modules were developed using a top-down approach of topographic science knowledge and a bottom-up approach using feature type terms based on standards developed by the USGS and its partners. Relations between these resources use the Open Geospatial Consortium standard spatial relation terms based on the 9-Intersection model (Herring, 2006). Three general types of properties relate feature types; locators, generators, and descriptors to reflect the spatial locations, generative processes, descriptions of subjects and objects of triples (Table 1).

Subjects	Objects			
	F	L	G	D
featureType - F	1	2	3	4
locator - L	5	1	5	-
generator - G	5	2	1	6
descriptor - D	7	-	3	1

Table of Relation Codes			
owl:sameAs	1		
locatedAt	2	locationOf	5
generatedBy	3	generates	6
describedAs	4	describes	7

Table 1. Relation codes linking triple data subjects and objects in a cross-matrix. The numbers in the table refer to codes identified in the Table of Relation Codes.

The design of the ontology pattern and their level of detail must be sensitive to the wide range of landscape characterizations represented by topographic data (Brewer and others, 2009).

2.3 An Example of a Complex Feature Type: Mine

Our example of a complex feature type ontology is based on the feature type 'Mine.' By definition, a mine is an excavation of the earth for extracting minerals (WordNet). A mine is always located at the site of the mineral or metal deposit and will always have equipment for moving earth material. A taxonomy of mines is made up of 4 cross-linked factors; a mine may be at the surface or underground, and may extract metals or industrial minerals. These would be the necessary and sufficient conditions of the particular feature, mine (Figure 3). Without these conditions the feature would not exist, and does not depend on associated components, such as infrastructure, to still be called a mine.

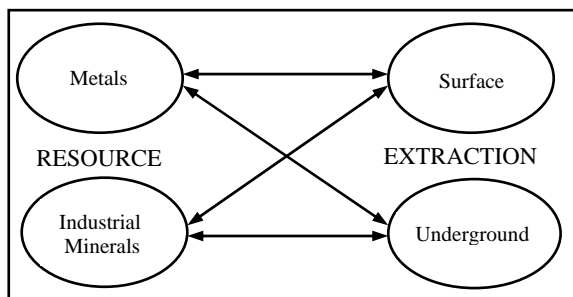


Figure 3. Necessary and sufficient conditions of a complex feature class, mine

Beyond these criteria, infrastructure and labor could be regarded as secondary conditions. It could be argued that a mine must be controlled by humans and that resources are extracted for consumption by a greater market, not all of which would happen directly at the mine site. Workers may reside adjacent to the mine site or commute. Developed areas may be

far from the mine so that transportation may be provided by the company bus or private cars may use the road system where the mine is located near residential areas. Roads may be rough, but navigable by large vehicles. Transportation supports labor and extraction, and takes a form depending on the landscape, which varies from place to place. The linkages between these concepts are largely geographic.

An additional set of prevailing conditions apply to the type of resources and their extraction method, processing and waste. Surface and underground mines vary in the degree of their disturbed area and whether a vertical, horizontal, or diagonal shaft is present. The common features between these types may include buildings, conveyors, power lines, communications, and transportation options, such as railroads or barges. Industrial minerals, such as crushed limestone for example, does not require extensive processing facilities at the mine location. Metal ores require more processing techniques to obtain the desired product. The processing of the resources can leave waste deposits at the mine site, or they may be transported to other locations depending on economics, policies, or regulations.

Linked data enhances the potential to articulate features to the landscape context as well. For example, the proximity of mines to water can create environmental quality concerns. Mining operations can be detrimental to surface and subsurface waters when quantities of the ore leach into those waters after they've been extracted from the host rock. The minerals may be toxic to water- and plant-life and can even reach the water table where water-wells may be located. Such potential risks can be modeled as part of the topographic network by linking the semantic identifiers of 'mine' and the hydrologic network of its area.

The example of 'mine' as a complex feature type indicates how complex feature types can be clarified through the use of semantic graphs (Figures 4 and 5). In relational databases such as GIS, mines are typically represented as points with latitude and longitude. In remote sensing, the visual identification of mines from pan-chromatic and multi-spectral imagery requires expert interpretation and risks false-positives. Mines could be mistaken for some farming activities, geologic hazards, or even new construction sites. This ambiguity may even be exploited for misidentification. Actual mines, however, have component features with individual identifiers, such as names, id numbers, or relative identities, for example as parts of larger complexes. A linked-data approach draws on that specific information to disambiguate the components because each component carries a universal resource identifier (URI), a string of characters used to identify the resource on the Internet (Mealling and Denenberg, 2002). By applying complex feature ontology to other data sources, misidentification may still exist, but the feature is flagged for further investigation for geographical information extraction.

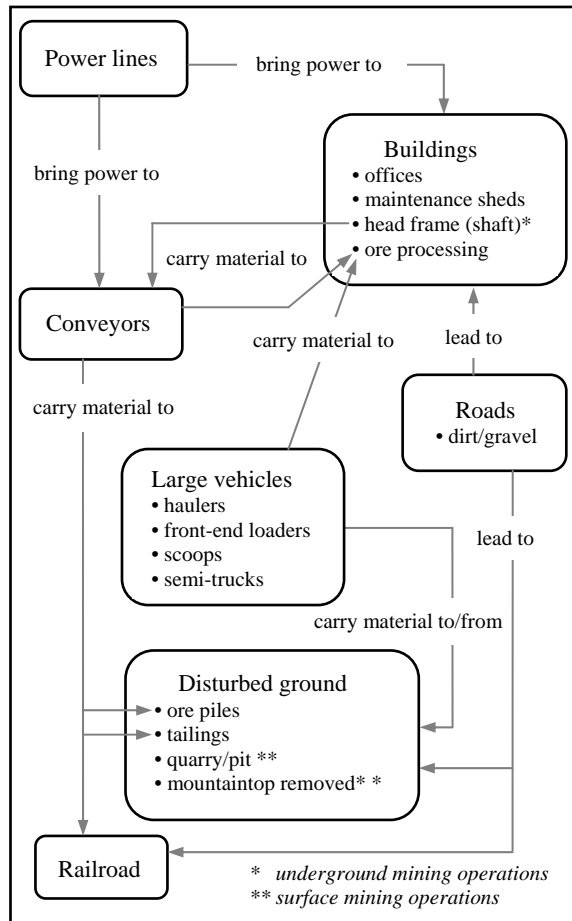


Figure 4. Semantic graph for mining operations.

The component features of a mine are geometric combinations of points (such as buildings), line (such as roads), and areas (such as disturbed ground) (Figure 4). Some features are optional, such as the mining shaft, and features are temporal, such as the movement of vehicles. Mine, as a feature class, involves an expanding context of feature class, support systems such as to railroads, and landscape, as the quarry/pit or mountain top removal links the mine directly to the land. The semantics of feature relations (semantic properties) are repeatable and refer back to basic and necessary conditions for the semantic meaning of the complex feature class 'mine'; that is, the movement of material and the power needed for its objective.

Complex features such as mines exist as sub-graphs in the Semantic Web. That means, they are subsets of the broader network of the Semantic Web. The components of a complex feature would have a high relevance value and strong association with each other and would be similar to each other in context (Ramakrishnan and others, 2005). To extract complex features, their subgraphs must be extracted as a unit or within the expanding context of the concept shown in Figure 2, feature, systems, and landscape. The exploration of these complex relationships is called semantic analytics (Perry and others, 2009). The semantic analytics for complex feature identification and extraction depends on techniques such as

subgraph isomorphism, but requires further research (Ullmann, 1976).

3. CONCLUSIONS

Complex features have component features with functional relations to each other and to outside systems. In this way, the complex feature ontologies extend from landscape and cartographic categorization to cross-integrate and connect a wide range of landscape features and processes across regions of the United States.

4. ACKNOWLEDGMENTS

The authors thank Cheryl M. Seeger, Missouri Department of Natural Resources, for her valuable comments about this work.

5. REFERENCES

- Brewer, C.A., Buitenfield, B.P, and Usery, E.L, 2009. Evaluating Generalizations of Hydrography in Differing Terrains for *The National Map* of the United States. In: Proceedings, 24th International Cartographic Conference, Santiago, Chile, 2009.
- Casati, R. and Varzi, A., 1999. *Parts and Places, the Structures of Spatial Representation*. Cambridge, Mass, MIT Press.
- Dago, E., Blomqvist, E., Gangemi, A., Montiel, E., Nikitina, N., Presutti, V., Villazon-Terrazas, B., 2005. Pattern based ontology design: methodology and software support. NoOn: Lifecycle Support for Networked Ontologies, Integrated Project (IST-2005-027595). NeOn-project.org.
- ESRI, 2010. ArcGIS. <http://www.esri.com/products/index.html>.
- Gangemi, A., 2010. Submissions:SpeciesNames. NeOn.org. <http://ontologydesignpatterns.org/wiki/Submissions:SpeciesNames>
- Gruber T., 2007. Ontology of Folksonomy: A Mash-up of Apples and Oranges. *International Journal on Semantic Web & Information Systems* 3 (2) <http://tomgruber.org/writing/ontology-of-folksonomy.htm>.
- Herring JR (ed), 2006. Open GIS implementation specification for geographic information – Simple feature access – Part 1: Common architecture. Open Geospatial Consortium Inc., OGC 06-103r3, Wayland, Mass
- Kuhn, W. 2001. Ontologies in support of activities in geographical space. *International Journal of Geographic Information Systems* 15(7): 613 – 631.
- Lam, N., Catts, D., Quattrochi, D., Brown, D., and McMaster, R., 2005. Scale: In: McMaster, R.B., and Usery, E.L. (eds.) *A Research Agenda for Geographic Information Science*: Boca

Raton, Florida, University Consortium for Geographic Information Science, CRC Press, p. 93-128.

Mealling M., Denenberg R. (eds), 2002. Report from the Joint W3C/IETF URI Planning Interest Group: Uniform Resource Identifiers (URIs), URLs, and Uniform Resource Names (URNs): Clarifications and Recommendations. Request for Comments 3305. The Internet Society, Reston, Virginia

National Geospatial-Intelligence Agency, 2007. Complicated Features Workshop Conference Report. Warrenton, Virginia.

OntologyDesignPatterns.org, 2010. Semantic Web portal, NeOn Project, http://ontologydesignpatterns.org/wiki/Main_Page.

Perry, M., Sheth, A., Arpinar, I.B., and Hakimpour, F., 2009. Geospatial and Temporal Semantic Analytics. In: Karimi, H.A. (ed) *Handbook of Research on Geoinformatics*. Hershey, Information Science Reference.

Ramakrishnan, C., Milnor, W.H., Perry, M., and Sheth, A., 2005. Discovering Informative Connection Subgraphs in Multi-relational Graphs. *SIGKDD Explorations*, 7 (2), pp. 56 – 63.

Smith, B., and Mark, D., 2003. Do Mountains Exist? Towards an Ontology of Landforms. *Environment and Planning*, 30 (3), pp. 411-427.

Ullman, J.R. 1976. An Algorithm for Subgraph Isomorphism. *Journal of the Association for computing Machinery* 23(1): pp. 31 – 42.

W3C, 2010. Semantic Web. World Wide Web Consortium. <http://www.w3.org/standards/semanticweb/>

WordNet. WordNet, A lexical database for English. Princeton University. <http://wordnet.princeton.edu/>.