An Ontology Design Pattern for Surface Water Features

Gaurav Sinha[1], David Mark[2], Dave Kolas[3], Dalia Varanka[4], Boleslo E. Romero[5], Chen-Chieh Feng[6], E. Lynn Usery[4], Joshua Liebermann[7], and Alexandre Sorokine[8]

[1] Department of Geography, Ohio University, Athens, OH, USA
[2] Department of Geography, University at Buffalo, Buffalo, NY, USA
[3] Raytheon BBN Technologies, Columbia, MD, USA
[4] U.S. Geological Survey, Rolla, MO, USA
[5] Department of Geography, University of California, Santa Barbara, CA, USA
[6] Department of Geography, National University of Singapore, Singapore
[7] Tumbling Walls LLC, 85 High Street, Newton, MA 02464
[8] Oak Ridge National Laboratory, Oak Ridge, TN, USA

Abstract. Surface water is a primary concept of human experience but concepts are captured in cultures and languages in many different ways. Still, many commonalities exist due to the physical basis of many of the properties and categories. An abstract ontology of surface water features based only on those physical properties of landscape features has the best potential for serving as a foundational domain ontology for other more context-dependent ontologies. The *Surface Water* ontology design pattern was developed both for domain knowledge distillation and to serve as a conceptual building-block for more complex or specialized surface water ontologies. A fundamental distinction is made in this ontology between landscape features that act as containers (e.g., stream channels, basins) and the bodies of water (e.g., rivers, lakes) that occupy those containers. Concave (container) landforms semantics are specified in a *Dry* module and the semantics of contained bodies of water in a *Wet* module. The pattern is implemented in OWL, but Description Logic axioms and a detailed explanation is provided in this paper. The OWL ontology will be an important contribution to Semantic Web vocabulary for annotating surface water feature datasets. Also provided is a discussion of why there is a need to complement the pattern with other ontologies, especially the previously developed *Surface Network* pattern. Finally, the practical value of the pattern in semantic querying of surface water datasets is illustrated through an annotated geospatial dataset and sample queries using the classes of the *Surface Water* pattern.

1 Introduction and Motivation

Surface water refers to water that exists on a surface at a greater mass than just detectable moisture on the earth's surface. It is a critical natural resource for life on earth, and a primary category of environmental reality within the realm of human experience. Yet, given the immensely rich and varied contexts of our experiences, it is not surprising that features associated with surface water (and the landscape, in general), are perceived and lexicalized quite differently, depending on which surface water

M. Duckham et al. (Eds.): GIScience 2014, LNCS 8728, pp. 187–203, 2014.

characteristics are recognized and emphasized by a culture or language. Similar, though less extreme, differences are also found among scientists and other professions. The innovative multidisciplinary field called *Ethnophysiography* has emerged recently for exploring these nuances arising from the intersection of language, culture, and cognition as they affect the interpretation of the landscape [21-22].

The variability implies that there exist several ways for classifying and relating features, and that there may be loss of information due to semantic interoperability between different conceptual systems. Researchers have explored topographic gazetteers and spatial data standards from countries across the world, such as GNIS[1], SDTS[2], and Geonet Names Server[3] (all from USA), INSPIRE[4] (Europe), TTDMS[5] (Taiwan), and topographic map standards from the Russian Federation (as discussed in [8]), to show the varied way topographic concepts are understood and formalized [5, 8, 20, 24]. Even formally developed systems such as WordNet[6] and EnvO[7], or SWEET[8] are inconsistent with each other—and with common sense conceptualizations of geospatial phenomena. Part of the problem is that different aspects of the landscape may be preferentially paid attention to in different cultures, languages and professions [31-34]. Hence, categories for landscape, including surface water, may be difficult to generalize, and terms may not have one-to-one correspondence with terms in other languages [21-22, 34].

Such dissimilarities should not, however, distract from the fact that people do communicate successfully across cultural and linguistic barriers and standards and ontologies can also be rendered interoperable (albeit with some information loss). For example, comparison of Russian, Taiwanese (Mandarin language) and US (English) geospatial standards revealed several terrain and hydrographic qualities, relations and categories that are shared and can be used for concept matching between the national geospatial standards [8]. As another example, the ambitious European INSPIRE spatial data infrastructure initiative can cater to the variation across European countries by capturing localized, country-specific geographic semantics in separate microtheories, and then allow inference of a subset of shared conceptualizations to enable semantic interoperability at the global European level [5]. Yet, such interoperability driven studies only hint at what may be some of the underlying principles of different conceptualization systems. For a comprehensive understanding, substantial research on geographic cognition [26], nature of geographic categories [31], and naïve geography [6] will be needed to discover general principles. It is safe to assume that such theories are unlikely to emerge anytime soon, since many languages, cultures and contexts would have to be investigated to identify truly stable categories and properties.

---

1 Geographic Names Information System (GNIS): `http://geonames.usgs.gov`
2 Spatial Data Transfer Standard (SDTS): `http://mcmcweb.er.usgs.gov/sdts`
3 GEOnet Names Server (GNS): `http://earth-info.nga.mil/gns/html`
4 INSPIRE Directive: `http://inspire.jrc.ec.europa.eu`
5 Taiwan Topographic Map Data Standard (TTMDS): `http://fas.harvard.edu/chgis/work/coding/feat_types_tw.htm`
6 Wordnet: A Lexical Database for English: `http://wordnet.princeton.edu`
7 Environmental Ontology: `http://environmentontology.org`
8 SWEET ontologies: `http://sweet.jpl.nasa.gov`

In the shorter term, scientific ontologies and well-known theoretical frameworks, including Gibson's theory of environmental affordances [10], Horton's primary theory [16], Lakoff and Johnson's ideas of experiential realism and embodied cognition [18-19], and Hayes' naïve physics manifesto [14, 15] offer reasonable justification for making some basic assumptions about how people experience the world. The consensus from these theoretical frameworks seems to be that there exist some physical precepts of the landscape that all human beings (with similar physical sensory capabilities) are able to perceive and experience, irrespective of their background and context. Identifying the minimal components of that *commonly experienced* landscape would allow the design of foundational landscape ontologies [29]. Such ontologies will serve best if conceptually grounded in basic, easily generalizable experiences of physical reality. This paper contributes to this research agenda by presenting a new *Surface Water* ontology that captures the essential semantics of discrete surface water features and their physical connection to the earth's surface. The semantics were primarily derived from considerations of physically observable properties and features in the landscape.

The success of any foundational domain ontology rests on it being relatively abstract and sparse in terms of how many categories and properties it specifies to avoid over-commitments and be useful across domains. It should also be easily extensible across geographic scales, and provide clear criteria for adding more specialized domain concepts. One popular and effective semantic engineering solution is to create small ontology design patterns to specify conceptualizations pertaining to a particular slice of a domain [9]. The *Surface Water* ontology presented here is such a small, easily comprehensible, and generalized ontology design pattern focused on some simple concepts from the domain of surface water. It is intended to serve both as a conceptual building-block for guiding the design of more complex surface water ontologies, and also as a self-contained knowledge representation unit capturing essential surface water semantics reusable in any domain with equal validity.

The fundamental design principle used for this pattern is to explicitly separate landscape features that act as containers for water to flow or collect, from the flowing and standing bodies of water that occupy those containers. The categories modeled by this pattern are abstract enough to function as "meta" categories closely corresponding to (but not equivalent to) basic categories encountered in hydrology, a field-observation driven geoscience domain that already offers a stable system of surface water feature types, and also to categories often encountered in natural languages. The pattern's categories reflect distinctions driven by observable physical properties (e.g., shape, size, depth, flow of liquids), and thus they are compatible with Horton's primary theory [16]. That leads to the surmise that such distinctions should also be inducing recognition of similar categories, albeit with additional properties and at different conceptual granularities, in most cultures and natural languages.

The *Surface Water* pattern captures semantics that arise from the object view of the surface water domain, but cannot capture non-channelized flows directly on the surface during floods and runoff. Hence, this paper also includes a brief discussion of another pattern (developed earlier by the authors), called the *Surface Network* pattern which is based on a well-known theory for conceptualizing any surface as an abstract network of simple shape elements [4, 25]. The *Surface Network* pattern can help incorporate surface flow semantics *anywhere* on the surface, not just in channels or basins. For complete representation of surface water semantics, several other intuitive object, network and field based ontologies will also be needed.

The rest of this paper is organized as follows. Section 2 presents the methodology and conceptual motivation for designing the pattern. The conceptual foundations and all the axioms of the OWL ontology are presented in Section 3. The practical utility of the pattern for semantic querying and annotation is discussed in Section 4. Section 5 briefly discusses how the *Surface Network* pattern applies to the domain of surface water, and Section 6 wraps up the paper with some general conclusions.

2 Pattern Design

2.1 Methodology

The *Surface Water* ontology pattern is supposed to function as a core surface water domain ontology (complemented by the *Surface Network* pattern), that is sufficiently abstract to be applied to more specific geospatial ontology applications. As mentioned above, ontology design patterns are small ontologies capturing essential, reusable qualities of a theme, and acting as building blocks in more complex ontologies. They reduce duplicated work and the core elements of the pattern facilitate data integration since they are designed to remain consistent when reused within different applications [9]. A key requirement for pattern design is that both domain and ontology engineering knowledge experts need to understand each other's perspectives. An increasing number of patterns are being designed at Geo-Vocabulary Camps (GeoVoCamps), which are a bottom-up, participatory approach to pattern design, achieved through 2-3 day working sessions of domain experts and ontology engineers to discuss and implement patterns for the geospatial domain. Philosophically motivated debates, and extensive discussions about the practical scope of the pattern and which domain entities and properties should be selected, characterize the GeoVoCamp workshops. Semantic engineering principles and implementation method determine the final form of the pattern which is generally available online and sometimes also documented as research publications [3, 17, 30].

The *Surface Water* pattern was developed at GeoVoCampDC2013[9] by the authors of this paper, most of whom also worked together to develop the *Surface Network* pattern at an earlier GeoVoCampSOCoP2012[10]. Both workshops were organized by members of Spatial Ontology Community of Practice (SOCoP).[11] There is no single authoritative resource that can be cited for the *Surface Water* pattern. As is the case

[9] GeoVocampDC2013: `http://vocamp.org/wiki/GeoVoCampDC2013`
[10] GeoVocampSOCoP2012: `http://vocamp.org/wiki/GeoVoCampSOCoP2012`
[11] Spatial Ontology Community of Practice (SOCoP): `http://socop.org`

for most patterns, the insights came from the collective research and practice of the authors. A wide variety of resources on surface water concepts was also known to them and considered more than sufficient as background knowledge for making decisions about both surface water domain and pattern design issues. The following sources of knowledge informed the *Surface Water* pattern design: natural language texts, multilingual dictionaries, encyclopedias, geospatial data standards (GNIS, SDTS), geoscientific reference texts, lexical databases (NGA GNS, WordNet), geoscience ontologies (SWEET, EnvO) and prior geographic and formal ontology research on scientific, legal, and folk concepts of surface water or closely related concepts [1-2, 5, 7-8, 12-13, 20-22, 24, 28-34].

2.2 Conceptual Background

A pattern needs to be generic enough to find recurring use in diverse contexts [9]. A well-established method for designing and motivating patterns is identification of a set of competency questions that refine the general use case and illustrate the types of semantic queries that can be addressed by implementing the pattern in more domain-specific contexts [11]. Some typical questions that best illustrate the generality and scope of the *Surface Water* pattern in a wide variety of contexts are listed below.

Q1. "Find all standing water bodies that are completely located in region X."
Q2. "Find all direct tributaries flowing into river X."
Q3. "Find all types of streams that originate from and also terminate in a basin."
Q4. "Find all valleys draining into a lake X."
Q5. "Find all streams which drain into lakes that do not fill their basins."

These queries can be relevant in a wide range of domains such as topographic mapping and querying, hydrological analysis, digital terrain analysis, pollution transport modeling, navigation, habitat analysis, natural resource conservation, disaster planning etc.. For example, a water body is abstract enough to resolve to different entity types (lakes, ponds, reservoirs) in different contexts, including different geographic scales (Q1). Tributaries of a river could be queried for determining navigation, or tracing pollution pathways, or to assess stream volumes (Q2). Streams, rivers, creeks, runs and many other flowing water features can be all treated as specializations of a single abstract type of channelized flow, and yet be distinguished from each other when needed in different contexts (Q3). Similarly, hydrographic, terrain, and other databases can be integrated and queried collectively by creating ontologies that explicitly capture the physical relationship between the land surface, (concave) land forms, and surface water (Q3-5). These types of competency questions helped identify the essential classes and properties of the *Surface Water* pattern.

3 Formalization of the Surface Water Pattern

3.1 General Principles

The following conceptualization underlies the *Surface Water* pattern: *There are locations or regions on the surface of the earth that host concave landforms, many of which interconnect, and act as containers for water to collect and/or flow through in dominant amounts under the influence of gravity.* The pattern, therefore, distinguishes between the terrain feature that acts as a container and the body of water, both considered to be overlapping in the same space. This is the most fundamental idea recognized in this pattern. The pattern is, thus, divided into two conceptual parts or modules: *Dry* and *Wet* to capture the two types of semantics separately, and to allow focused specializations in the future. The *Dry* module captures the semantics of concave landscape features (channel, depression and interface), which can exist regardless of the presence of surface water, but do act as containers for sustained water flow and storage. The *Wet* module is dependent on and reuses the *Dry* module features to capture the semantics of hydro features (stream segment, water body, and fluence) that occupy the features whose semantics are defined in the *Dry* module. Note that there are several types of snow and ice formations that may not be properly addressed by this pattern, which has been designed for the typical cases of liquid surface water features and contained in channels and depressions. The classes and properties within the *Dry* and *Wet* modules are formally encoded using Web Ontology Language (OWL), and available online.[12, 13] All semantics of the modules and how they interrelate is also captured schematically in Fig. 1.

Some general issues related to pattern design and how they are discussed in this paper need to be clarified as well. First and foremost, Description Logic (DL) notation is used for presenting axioms in this paper since it is much more compact than OWL. Names of properties are simplified to not begin with “has”, but they should be easy to track because class names begin with capital letters, and property names begin with small letters. Global domain and range declarations over properties are not used because that is known to reduce interoperability—all domain-range declarations for properties are defined only in the context of specific classes. All classes of the pattern are declared to be pairwise disjoint because they do not cover overlapping categories. Disjointness declaration is the recommended practice in OWL for improving inference about domain concepts. The DL axioms for disjointness are not presented below for lack of enough space. Property axioms are also not included for space constraints, but their intended purpose should be evident from Fig. 1, and the discussion of the axioms below.

---

[12] *Dry* module URI:
`http://purl.org/geovocamp/ontology/SurfaceWater_Dry`

[13] *Wet* module URI:
`http://purl.org/geovocamp/ontology/SurfaceWater_Wet`

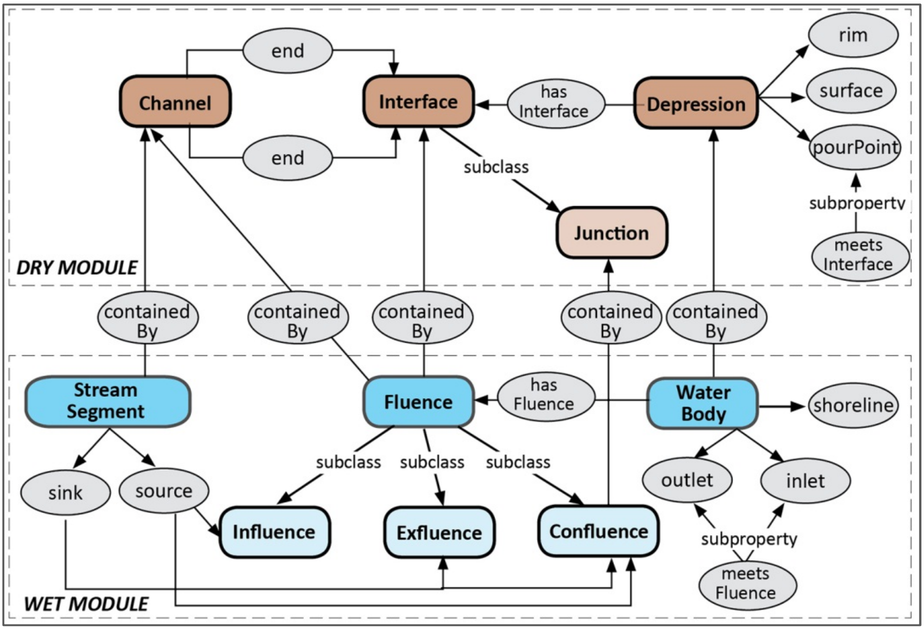

Fig. 1. Surface *Water* pattern's *Dry* and *Wet* module classes (brown/blue) and properties (grey)

3.2 Dry Module Semantics

The *Dry* module has two primary classes, *Channel* and *Depression*, while a third supporting class *Interface* formalizes concepts of spatial connections between surface water features in the terrain. These classes represent three-dimensional terrain features, where water occurs dominantly and their purpose is to be a foundation for specifying the semantics of classes of the *Wet* module.

Channel. The *Channel* class captures the semantics of a linear conduit with two ends, which is located on or is a natural part of the earth's surface, and as a consequence of its shape, it acts as a container where water can collect and flow in dominant amounts between the two ends of the conduit. Specifying that a channel has exactly two ends, each of which is formally represented through the *interface* class is considered sufficient to support flow semantics (specified in the *Wet* module). Axiom A1 encodes this logic. There are some properties (*lowerEnd*, *upperEnd*, *bed*, and *bank*) that are only included in the pattern to support future specializations of the *Channel* class, but not used in this pattern to maintain its generality. For example, the distinction between upper and lower ends of a channel is not made since it would preclude channels where flows reverse temporarily. Similarly, only well-defined channels may be deemed to have a bed and bank, but not many minor channels transporting thin rivulets of water.

$$Channel \sqsubseteq (\leq 2 end.Interface \sqcap \geq 2 end.Interface) \quad \text{(A1)}$$

Interface. An *interface* is a conceptual abstraction to represent "transition" locations at the end of channels or on the boundary of depressions. The most common use of the interface class will be to represent the two ends of a channel. While axiom A1 restricts that a channel's end can only be an interface, axiom A2 further restricts the interface to be the end of only a channel, and nothing else. An interface that represents the physical merger or bifurcation of channels, or the merger of a channel and depression, is a *Junction* (A3). When an interface represents the end of just a single channel, or a junction between a channel and a depression, it can be thought of as the cross sectional 2D planar area through which flow would enter or exit the channel. An interface (junction) involving three channel ends will always be a volume, since channels are assumed to have some depth. A channel can have only two interfaces, at each of its two ends, but a depression can have any number of interfaces, including none at all, as specified later (A7). Note that axioms A1 and A3 together also imply that a channel is atomic, in that it cannot contain another channel (i.e., no sub-channels are possible).

$$Interface \sqsubseteq \forall endOf.Channel \tag{A2}$$

$$Junction \sqsubseteq Interface \tag{A3}$$

Depression. A depression is a concavity in the earth's surface that is surrounded by higher ground all around and which can contain water by virtue of its shape and material surface. Depressions can be as large as ocean basins or as small as holes found in a channel bed. A depression is defined as being spatially enclosed by or having an upper bound marked by its rim (A4). The rim is the highest elevation line (a contour) that encloses the depression. Functionally, the rim denotes the level below which water can stay contained in the depression without overflowing. Formal specification of this definition of rim would require specification of multiple mathematical and spatial concepts, and is beyond the scope of this simple pattern. A depression also has the property of having exactly one pour point (A5), which marks the lowest location from where water would exit naturally if the depression was maximally filled, up to the level of its rim. The pour point is at the same elevation as and touches the rim (not formalized in OWL). A depression must also have a surface so that it can support a water body (A6). This means that this pattern allows only those individual depressions, whose surfaces allow containment of water bodies to be members of this class. Finally, the *meetsInterface* property is used to specify that a depression is connected to other channels, the outside region, or in rare cases, to another depression, through only *interfaces* (A7). As illustrated in Fig. 1, the *pourPoint* property is also declared as a subproperty of *meetsInterface*, since it must connect the depression to an interface which will then connect to other features or the region outside of the depression. The pour point is always on the rim, but interfaces with some channels which bring inflow may be on the rim or below it inside the depression (not formalized in OWL).

$$Depression \sqsubseteq (\leq 1 rim \sqcap \geq 1 rim) \tag{A4}$$

$$Depression \sqsubseteq (\leq 1 pourPoint \sqcap \geq 1 pourPoint) \tag{A5}$$

$$Depression \sqsubseteq (\leq 1 surface \sqcap \geq 1 surface) \tag{A6}$$

$$Depression \sqsubseteq \forall meetsInterface.Interface \tag{A7}$$

3.3 Wet Module Semantics

The *Wet* module reuses the classes of the *Dry* module to specify the semantics of surface water flow and collection in channels and depressions, respectively. There are three classes in the *Wet* module: *fluence, stream segment,* and *water body*. Instances of the latter are always contained by the channel and depression, respectively, while fluences may be contained in interfaces, junctions or channels.

Fluence. A fluence is the transitional water entering or leaving a stream segment or a water body. For stream segments within channels, the fluence can be either within a channel interface if flow starts or ends at the channel end, or outside the interface and within the channel, if flow starts or ends not at the end, but within a channel somewhere. For a water body, the fluence can either within be the interface to a channel or inside the depression containing the water body. The fluence class is further specialized through three (pairwise disjoint) subclasses: *influence*, *exfluence* and *confluence* (A8-A10) to capture all the ways flow can start or end for a stream segment or enter or exit from a water body. The influence is the source, the exfluence, the sink, and the confluence can be both the source and sink of (different) stream segments or a stream segment and a water body. The confluence is a type of fluence signifying water merging or transitioning from one stream segment into another stream segment or water body, or from a water body to a stream segment. It is always contained within the junction of channels, since stream segments must exit a channel to meet other stream segments (A11). If the stream flows through the entire channel, then the influence and exfluence are contained in the interfaces at the end of the channel, otherwise they are contained within the channel somewhere (A12-A13). Note that axioms A8-A13 do not specifically preclude the *fluence* classes from being re-used to cover non-channelized flow semantics. However, this pattern was designed to focus only on channelized flows, and broader than intended interpretation of the semantics is not recommended.

$$Influence \sqsubseteq Fluence \tag{A8}$$

$$Exfluence \sqsubseteq Fluence \tag{A9}$$

$$Confluence \sqsubseteq Fluence \tag{A10}$$

$$Confluence \sqsubseteq (\leq 1 containedBy.Junction \sqcap \geq 1 containedBy.Junction) \tag{A11}$$

$$Influence \sqsubseteq (\leq 1 containedBy.Interface \sqcap \geq 1 containedBy.Interface) \sqcup (\leq 1 containedBy.Channel \sqcap \geq 1 containedBy.Channel) \tag{A12}$$

$$Exfluence \sqsubseteq (\leq 1 containedBy.\ Interface \sqcap \geq 1 containedBy.Interface) \sqcup (\leq 1 containedBy.Channel \sqcap \geq 1 containedByChannel) \tag{A13}$$

Stream Segment. A stream segment is contained (and flows) within the channel (A14). Every stream segment has two flow related properties: a *source* and a *sink* (A15-A16), which mark the inflow and outflow ends of a stream segment. Flow as a process is too complex to be explicitly formalized in OWL. Instead, it is implied indirectly as directed from the source to the sink. Stream segments can only meet at a confluence. Note that multiple sources and sinks of merging or diverging streams are all ‘resolved’ to the *same* physical confluence within the junction that contains it. The use of source and sink properties prevents the unrealistic situation of multiple coincident fluence and interface entities, when stream segments meet. If a stream segment’s source receives water from another stream segment or water body, a source will be a confluence, or if the source is neither a stream segment or water body it will be an influence, but never an exfluence (A17). Similarly, a stream segment’s sink can never be an influence, and only either a confluence (if the stream segment loses water to another stream segment or water body) or exfluence, otherwise. The axioms do not preclude the possibility of more than one stream segment contained in a channel, if a stream segment does not flow end to end, but it should be a rare possibility, if at all.

$$StreamSegment \sqsubseteq (\leq 1 containedBy.Channel \sqcap \geq 1 containedBy.Channel) \quad \text{(A14)}$$

$$StreamSegment \sqsubseteq (\leq 1 source \sqcap \geq 1 source) \quad \text{(A15)}$$

$$StreamSegment \sqsubseteq (\leq 1 sink \sqcap \geq 1 sink) \quad \text{(A16)}$$

$$StreamSegment \sqsubseteq (\leq 1 source.(Influence \sqcup Confluence) \sqcap \geq 1 source.(Influence \sqcup Confluence)) \quad \text{(A17)}$$

$$StreamSegment \sqsubseteq (\leq 1 sink.(Exfluence \sqcup Confluence) \sqcap \geq 1 sink.(Exfluence \sqcup Confluence)) \quad \text{(A18)}$$

Water Body. This class represents a standing collection of water contained within a depression (A19). The water is contained due to the (impermeable) depression surface and between the depression’s surface anywhere up to the rim of the depression. Every water body, therefore, has a shoreline (A20). The shoreline achieves the highest level of the rim only when the depression is full (e.g. a full lake basin), otherwise the shoreline is at a lower level (typical in arid areas). This relationship between the shoreline and rim is not specified explicitly in OWL because it would need more complex axioms and incorporation of too many extra mathematical and spatial properties and entities. A water body meets stream segments that flow into or out of the water body. The *meetsFluence* property specifies that a water body connects to other bodies of water only through fluences (A21). If the water body fills a depression completely, it has an *outlet*, otherwise not (A22). The outlet can be either a confluence or exfluence (A23). It will be considered a confluence if the water body flows out to form a stream segment contained in a channel. Otherwise, if the water body loses water through the outlet in such a way that no sustained stream segment and channel are found connecting to the interface representing the pour point, then the outlet is considered to be an exfluence. A water body may have any number of *inlets*, including none at all, depending on how many stream segments or other discrete sources (e.g., underground springs) introduce net inflow into the water body. If an *inlet* exists, then it is a confluence if the inflow is from a stream segment in a channel, or an influence if its flow is not confined in a channel (A24). As shown in Fig. 1, *outlet* and *inlet* are also subproperties of *meetsFluence* because the inlet and outlet of a water body will always be represented by a fluence.

$$WaterBody \sqsubseteq (\leq 1 containedBy.Depression \sqcap \geq 1 containedBy.Depression) \quad (A19)$$

$$WaterBody \sqsubseteq (\leq 1 shoreline \sqcap \geq 1 shoreline) \quad (A20)$$

$$WaterBody \sqsubseteq \forall meetsFluence.Fluence \quad (A21)$$

$$WaterBody \sqsubseteq \leq 1 outlet \quad (A22)$$

$$WaterBody \sqsubseteq \forall outlet.(Exfluence \sqcup Confluence) \quad (A23)$$

$$WaterBody \sqsubseteq \forall inlet.(Influence \sqcup Confluence) \quad (A24)$$

3.4 Discussion

The *Surface Water* pattern was designed as a minimalist ontology modeling, and should not be expected to address all possible surface water types and cases. The designed classes cover a limited set of categories that are likely to be widely, if not universally, shared by most people. These categories are only supposed to serve as basic building blocks, similar to foundation ontology categories, for more specialized and context dependent categories. For example, semantics of braided streams may require a more complex channel type to be introduced. Wetlands (e.g., marsh, swamp, fen) forming over permeable lands and/or not contained in depressions may be modeled as another unrelated pattern, or as an extension to this pattern (e.g., as a special type of water body that supports substantial vegetation and/or is characterized by certain soil types). This pattern is also not designed to support semantics of processes that lead to changes in the physical properties of surface water features due to action of water contained inside. However, the pattern's classes should be able to describe the instances of surface water features at different times or stages of evolution, which will allow sharing information about the spatiotemporal behavior (albeit only in terms of snapshot states) of specific features and/or geographic areas where the features are located.

4 Applications of the Surface Water Ontology Pattern

4.1 Aligning Geo-Databases and Annotating Mapped Features

The *Surface Water* pattern has theoretical value as an encapsulation of fundamental domain categories and properties. The pattern is also an ontology for the Semantic Web, and a practical guide for making hydro-GIS datasets interoperable at a generalized level. Surface water features get represented in different ways in geospatial databases depending on intended use. A primary use of this pattern would be to make databases interoperable by interpreting the features as instances of the basic categories of the *Surface Water* pattern. As an example, Fig. 2 shows mapped representations of real world surface water features for a small study area. The left diagram in Fig.2 maps instances of Dry module classes, and the right diagram maps instances of the Wet module classes for the same study area. In this example, all depressions contain water bodies, and all junctions contain confluences. However, not all channels contain stream segments, and not all interfaces contain fluences. All fluences are confluences, except one which is an influence that can be inferred by visual comparison of the left and right diagrams to be located not within the interface at the channel's end, but within the channel itself. This implies that the stream segment does not fully traverse the channel end to end, but instead has a source starting somewhere downstream from the upper (flow) end of the channel. This example, thus, clearly underscores the importance of distinguishing between spatially overlapping instances of different surface water feature categories.

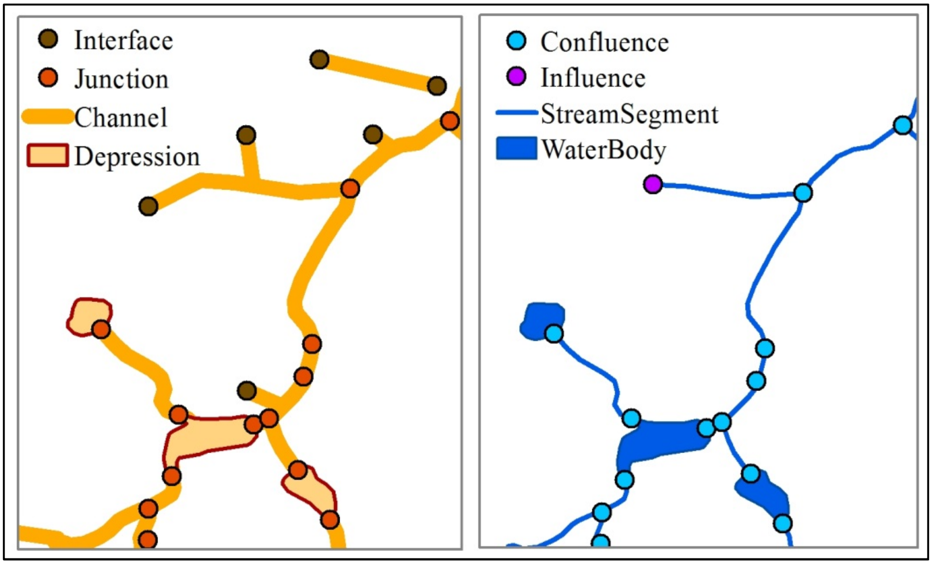

Fig. 2. An example of how surface water features can be described and mapped as instances of classes defined in the *Dry* (left) and *Wet* (right) modules of the *Surface Water* pattern

4.2 Querying Geo-Databases

A quick test of practicability is to check if the pattern is useful in answering the competency questions listed in Section 2. The queries listed below confirm and show how the pattern can be used to construct domain queries using terms (italicized below) corresponding to the pattern's categories and their properties. The returned features can be references to instances of surface water features shown above in Fig. 2.

Q1. "Find all *water bodies* which are *containedBy depressions* whose *rims* are completely contained within the spatial extent of region X."

Q2. "Find all *stream segments* whose *sinks* are *confluences* with any *stream segment* with name X."

Q3. "Find all stream segments whose *sources* and *sinks* are *confluences containedBy junctions* that are *interfaces* with a *depression*."

Q4. "Find all *channels* with *junctions* which *contains confluences* with the *water body* named X."
Q5. "Find all *stream segments* having *confluences* with *water bodies* that do not have an *outlet*."

As can be seen, semantics related to *both* bodies of water and their containing landforms are needed to correctly frame the queries. To make these sorts of queries possible, especially on the Semantic Web, we are also developing software to convert hydro-GIS datasets into RDF triples (standard graph data model for the Semantic Web), which can then be queried using the GeoSPARQL[14] semantic query language. Fortunately, the GeoSPARQL ontology also offers built-in support for geospatial data querying, as will be needed for most surface water datasets. There are, obviously, many other questions of interest for the domain that are not answerable with this simple pattern, and many complex geospatial queries that cannot be handled by GeoSPARQL. Also, data on "dry" stream channels and depressions are not as common as datasets that map only streams and water bodies. Establishing topological connections to find junctions and fluences, and distinguishing the types of fluences based on flow direction is only possible with advanced geospatial data models.

5 Integration with Other Ontologies

Several sources informed the design of the *Surface Water* pattern. The formalization of voids in hydrogeology [12] was identified as a mid-level foundational ontology before subsequently designing the *Surface Water* pattern, especially because it formalizes the container schema, and is itself aligned with the DOLCE foundational ontology [23]. However, it presented some immediate problems due to formalization in Common Logic and a focus on hydrogeology. These are not insurmountable issues, but still prevented a quick alignment process. Additionally, the immediate goal of this work is to integrate the *Surface Water* and *Surface Network* patterns since the latter addresses surface flow semantics not encoded in the former. The *Surface Network* pattern is unlikely to integrate well with the ontology of hydrogeological voids [12], which is another reason for not yet investing in alignment with the ontology of voids. In the following paragraphs the discussion is focused on the *Surface Network* pattern's utility as an abstract and reusable surface water domain ontology.

A surface network mathematically describes the global spatial shape of a (twice differentiable, smooth) surface in terms of a topological network of critical points (peaks, saddle points, and pits) and lines (ridge, course, slope, and contour lines), which together provide a generalized representation of the global surface shape [4, 25, 27]. The theory also includes two types of areal districts, hills and dales, which are bound by course lines and ridge lines, respectively, and exhaustively partition the surface into morphological parts, independent of each other. Three other areal feature types can be recognized, although only implicitly referred to in the literature: territory (area of overlap between exactly one hill and one dale), hilltop (enclosing area around

[14] GeoSPARQL: `http://www.opengeospatial.org/standards/geosparql`

a peak down to the contour of the highest saddle point connected to the peak via a ridge line), and basin (enclosing area around a pit up to the contour of the lowest saddle point connected to the pit via a course line). Surface networks offer a discrete, feature based abstraction of surfaces, which themselves lack searchable objects. This inspired the design of the *Surface Network* and the *Geospatial Surface Network* patterns, the former for topological abstractions of the surface, and the latter extending the former with support for metric geospatial surfaces. The two patterns are available as OWL ontologies[15, 16] and detailed in another manuscript [30].

These patterns are relevant to surface water semantics and are mentioned here because they formalize surface shape semantics. Intuitively, the earth's surface already provides the potential for surface water collection and flow. Water flows along channels under the influence of gravity, collects and flows along lines of steepest descent, often forming well-developed stream channels, and at other times just downhill anywhere on the surface, as during a rainfall event or a flood. Areas that drain together to a pit or basin form drainage basins (i.e., watersheds), and are bound and demarcated by drainage divides (i.e., ridges). Water naturally flows toward the lowest points available in a drainage basin, where it starts to collect and fill basins, which are the lowest areas around the pit within the drainage basin.

Keeping the above statements in mind, it emerges that if the earth's surface is abstracted as a surface network, its shape elements will easily capture the above semantics, albeit at an abstract level. Some shape elements correspond closely to categories of the *Dry* module of the *Surface Water* pattern, and some others capture additional surface semantics. For example, a surface network basin would be equivalent to a depression, the pit will model the lowest point in a depression, the pale will correspond exactly to the pour point of a depression, and the contour passing through the pale would be the rim of the depression. Surface networks also add concepts of drainage basins (dales) and drainage divides (ridge lines) missing in the *Surface Water* pattern, and course lines which abstract locations where water flows consistently, and therefore are conceptually quite similar to channels. The *Geospatial Surface Network* can be used to conceptualize hydrological stream networks, and non-channelized overland/sheet flows of water—semantics which are missing from the *Surface Water* pattern.

The reason for designing a *Surface Water* pattern, separate from the *Surface Network* pattern is that the mathematically abstract surface network elements do not correspond perfectly with the features of the real world. Because course lines must extend strictly between saddle points to pits, only a subset of flow channels can be made to correspond to course lines. Also, course lines technically never meet, so channel junctions cannot be modeled without making some theoretical adjustments to surface network theory. Furthermore, because pits and basins are often generalized in dry land focused digital elevation models, pits, depressions and the pales and channels connected to them practically never get recognized as part of surface networks. Still,

---

[15] *Surface Network* OWL pattern URI: `http://purl.org/geovocamp/ontology/SurfaceNetwork`

[16] *Geospatial Surface Network* OWL pattern URI: `http://purl.org/geovocamp/ontology/GeospatialSurfaceNetwork`

there are some benefits to using the *Surface Network* pattern as outlined above, and their possible integration into a larger pattern. The authors are currently planning to address the integration and re-evaluation of both patterns in a future GeoVoCamp workshop.

6 Conclusions

Ontology patterns allow us to specify our concepts in knowledge modules and force us to extract the most essential concepts of a domain. The *Surface Water* pattern was created at a GeoVoCamp, by domain experts and knowledge engineers. The pattern follows general design principles and includes only highly generalized categories based on physically observable characteristics. Since the pattern is simple and formalized in OWL, it can be used on the Semantic Web to share surface water feature datasets. The pattern should be easily reusable in different domains to annotate and implement semantic querying of surface water feature datasets. It can also be used to implement GIS data models that would be compatible with the naïve geography approach to GIS design and querying. As discussed above, the pattern is useful for intuitive querying and mapping of terrain and hydrographic GIS datasets.

The *Surface Water* pattern is abstract to be generalizable. On the other hand, that also means that it must be combined and or extended with other topography related ontologies, and also aligned with other foundational ontologies, to help realize its true potential. In that respect, the separation of the *Dry* and *Wet* modules is well suited conceptually to link and develop specialized ontologies for terrain and surface water in tandem. The *Dry* module should be specialized to add surface water semantics pertaining to morphology (e.g., size, shape, topology) and terrain composition (e.g., based on types of rocks, soils, vegetation), while the *Wet* module should be specialized to capture categories and properties related specifically to hydrologic characteristics (e.g. flow volume, flow frequency, source, and quality). Finally, it should be noted that although this pattern is designed and discussed for surface water semantics, it should be usable for modeling basic semantics of flow possibility of other liquids (e.g., polluted plumes, lava flows) on terrain or other physical surfaces, including that of other planets.

Acknowledgments. This manuscript has been prepared in part by Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, Tennessee 37831-6285, managed by UT-Battelle, LLC for the U.S. Department of Energy under contract no. DEAC05-00OR22725. Accordingly, the publisher, by accepting the article for publication acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for U.S. Government purposes. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government. The authors are grateful to Gary-Berg Cross and Spatial Ontology Community of Practice (SOCoP) for organizing GeoVoCampDC2013, and the USGS and three anonymous reviewers for providing comments.

References

1. Bennett, B.: Application of Supervaluation Semantics to Vaguely Defined Spatial Concepts. In: Montello, D.R. (ed.) COSIT 2001. LNCS, vol. 2205, pp. 108–123. Springer, Heidelberg (2001)
2. Bromhead, H.: Ethnogeographical Categories in English and Pitjantjatjara/Yankunytjatjara. Language Sciences 33, 58–75 (2011)
3. Carral, D., Scheider, S., Janowicz, K., Vardeman, C., Krisnadhi, A.A., Hitzler, P.: An Ontology Design Pattern for Cartographic Map Scaling. In: Cimiano, P., Corcho, O., Presutti, V., Hollink, L., Rudolph, S. (eds.) ESWC 2013. LNCS, vol. 7882, pp. 76–93. Springer, Heidelberg (2013)
4. Cayley, A.: On Contour Lines and Slope Lines. London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 18, 264–268 (1859)
5. Duce, S., Janowicz, K.: Microtheories for Spatial Data Infrastructures – Accounting for Diversity of Local Conceptualizations at a Global Level. In: Fabrikant, S.I., Reichenbacher, T., van Kreveld, M., Schlieder, C. (eds.) GIScience 2010. LNCS, vol. 6292, pp. 27–41. Springer, Heidelberg (2010)
6. Egenhofer, M., Mark, D.M.: Naïve Geography. In: Kuhn, W., Frank, A.U. (eds.) COSIT 1995. LNCS, vol. 988, pp. 1–15. Springer, Heidelberg (1995)
7. Feng, C.-C., Bittner, T., Flewelling, D.M.: Modeling Surface Hydrology Concepts with Endurance and Perdurance. In: Egenhofer, M., Freksa, C., Miller, H.J. (eds.) GIScience 2004. LNCS, vol. 3234, pp. 67–80. Springer, Heidelberg (2004)
8. Feng, C.-C., Sorokine, A.: Comparing English, Mandarin, and Russian Hydrographic and Terrain Categories. International Journal of Geographical Information Science (2013) (in press), doi:http://dx.doi.org/10.1080/13658816.2013.831420
9. Gangemi, A., Presutti, V.: Towards a Pattern Science for the Semantic Web. Semantic Web 1(1-2), 61–68 (2010)
10. Gibson, J.J.: The Ecological Approach to Visual Perception, Houghton Mifflin, USA (1979)
11. Gruninger, M., Fox, M.S.: The Role of Competency Questions in Enterprise Engineering. In: Proceedings of the IFIP WG5.7, Workshop on Benchmarking – Theory and Practice, pp. 212–221 (1994)
12. Hahmann, T., Brodaric, B.: The Void in Hydro Ontology. Frontiers in Artificial Intelligence and Applications: Formal Ontology in Information Systems 239, 45–58 (2012)
13. Hart, G., Dolbear, C., Goodwin, J., Feliciter, L.: Using Structured English to Represent a Topographic Hydrology Ontology. In: OWL Experiences and Directions Workshop (2007)
14. Hayes, P.: The Naïve Physics Manifesto. In: Michie, D. (ed.) Expert Systems in the Micro-Electronic Age, pp. 242–270. Edinburgh University Press (1979)
15. Hayes, P.: Naive Physics I: Ontology of Liquids. In: Hobbs, J., Moore, R. (eds.) Formal Theories of the Commonsense World, pp. 71–108. Ablex, Norwood (1985)
16. Horton, R.: Tradition and Modernity Revisited. In: Hollis, M., Lukes, S. (eds.) Rationality and Relativism, pp. 201–260. Basil Blackwell, Oxford (1982)
17. Hu, Y., Janowicz, K., Carral, D., Scheider, S., Kuhn, W., Berg-Cross, G., Hitzler, P., Dean, M., Kolas, D.: A Geo-Ontology Design Pattern for Semantic Trajectories. In: Tenbrink, T., Stell, J., Galton, A., Wood, Z. (eds.) COSIT 2013. LNCS, vol. 8116, pp. 438–456. Springer, Heidelberg (2013)
18. Lakoff, G.: Women, Fire, and Dangerous Things: What Categories Reveal about the Mind. University of Chicago Press, Chicago (1987)

19. Lakoff, G., Johnson, M.: Philosophy in the Flesh. University of Chicago Press, Chicago (1986)
20. Mark, D.M., Smith, B.: A Science of Topography: From Qualitative Ontology to Digital Representations. In: Bishop, M.P., Shroder, J.F. (eds.) Geographic Information Science and Mountain Geomorphology, pp. 75–100. Springer-Praxis, Chichester (2004)
21. Mark, D.M., Turk, A.G., Stea, D.: Progress on Yindjibarndi Ethnophysiography. In: Winter, S., Duckham, M., Kulik, L., Kuipers, B. (eds.) COSIT 2007. LNCS, vol. 4736, pp. 1–19. Springer, Heidelberg (2007)
22. Mark, D.M., Turk, A.G.: Landscape Categories in Yindjibarndi: Ontology, Environment, and Language. In: Kuhn, W., Worboys, M.F., Timpf, S. (eds.) COSIT 2003. LNCS, vol. 2825, pp. 28–45. Springer, Heidelberg (2003)
23. Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A.: WonderWeb Deliverable D18 Ontology Library (final) (2003), `http://wonderweb.semanticweb.org/deliverables/documents/D18.pdf` (last accessed May 15, 2014)
24. Masser, I.: Building European Spatial Data Infrastructures, 2nd edn. ESRI Press, Redlands (2010)
25. Maxwell, J.C.: On Hills and Dales. The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science 40, 421–427 (1870)
26. Montello, D., Freundschuh, S.: Cognition of Geographic Information. In: McMaster, R., Usery, L.E. (eds.) A Research Agenda for Geographic Information Science, pp. 61–91. CRC Press (2005)
27. Rana, S.S. (ed.): Topological Data Structures: An Introduction to Geographical Information Science. John Wiley & Sons, Ltd., Chichester (2004)
28. Santos, P., Bennett, B., Sakellariou, G.: Supervaluation Semantics for an Inland Water Feature Ontology. In: Proceedings of International Joint Conference on Artificial Intelligence, IJCAI 2005, pp. 564–569 (2005)
29. Sinha, G., Mark, D.M.: Toward a Foundational Ontology of the Landscape. In: Fabrikant, S.I., Reichenbacher, T., Kreveld, M., Schlieder, C. (eds.) GIScience 2010, Extended Abstracts, Zürich, Switzerland, September 14-17 (2010)
30. Sinha, G., Kolas, D., Mark, D.M., Romero, B.E., Usery, L.E., Berg-Cross, G., Padmanabhan, A.: Surface Network Ontology Design Patterns for Linked Topographic Data. Manuscript submitted to Semantic Web for review
31. Smith, B., Mark, D.M.: Geographic Categories: An Ontological Investigation. International Journal of Geographical Information Science 15 (7), 591–612 (2001)
32. Taylor, M., Stokes, R.: Up the creek. What is Wrong with the Definition of a River in New South Wales? Environment and Planning Law Journal 22(3), 193–211 (2005)
33. Taylor, M.P., Stokes, R.: When is a River not a River? Consideration of the Legal Definition of a River for Geomorphologists Practising in New South Wales, Australia. Australian Geographer 36(2), 183–200 (2005)
34. Wellen, C.C., Sieber, R.E.: Toward an Inclusive Semantic Interoperability: The Case of Cree Hydrographic Features. International Journal of Geographical Information Science 27(1), 168–191 (2013)