8-1-2014

Toward A Foundational Hydro Ontology For Water Data Interoperability

Boyan Brodaric

Torsten Hahmann

11th International Conference on Hydroinformatics
HIC 2014, New York City, USA

TOWARD A FOUNDATIONAL HYDRO ONTOLOGY FOR WATER DATA INTEROPERABILITY

BOYAN BRODARIC (1), TORSTEN HAHMANN (2)
(1): Geological Survey of Canada, Ottawa, ON, Canada
(2): National Center for Geographic Information and Analysis, School of Computing and Information Science, University of Maine, Orono, ME, USA

Hydro ontologies are digital artifacts that contain representations of hydrological entities. They are being constructed to advance interoperability within and between various water data networks. Increased growth of such networks, however, is prompting integration of a greater variety of data, which requires a more general and principled approach to hydro ontology. One such approach is the development of a canonical reference ontology. This paper describes recent progress on the development of such an ontology for both the surface and subsurface hydro domains. Identified is a container schema that anchors the ontology, and also described is an initial representation of the ontology.

1. INTRODUCTION

Many scientific and social issues require an integrated approach to groundwater and surface water management. Applications in climate change, water accounting, and groundwater contamination are but a few examples. For these applications, relevant data is increasingly being obtained from online water data networks, but the surface and subsurface aspects of such systems are often disconnected, and this disconnection is being propagated to standards for data interoperability. For example, European data standards for surface water and groundwater are being developed in distinct themes, respectively hydrography and geology [8, 9]. Likewise, the Open Geospatial Consortium, in partnership with the World Meteorological Organization, is developing similar data standards [2, 3] that are as yet largely unconnected. While recognition of the need to unify data representations across domains is growing, there does not exist a common conceptualization. Foundational ontologies are a candidate for such a conceptualization, because they contain general entities that are universal across domains. However, work to date on hydro ontologies has taken neither a foundational nor unifying perspective, inasmuch as related work focusses on surface bodies [7, 11], subsurface features [12], or water quality parameters [1]. In this paper we describe recent progress on the development of a reference ontology for the hydro domain that takes a more foundational and unifying approach. Identified is a schema common to the surface and subsurface domains, consisting of four key concepts and three relations that anchor the ontology; also developed is an initial representation of the schema, cast as a specialization of the DOCLE foundational ontology [10]. Section 2 discusses methods used to develop and assess the ontology; Section 3 describes the hydro container schema and the associated ontology, evaluating the latter; and Section 4 provides a short summary and discusses some future directions.

2. METHODS – DEVELOPING A REFERENCE HYDRO ONTOLOGY

Development of an ontology meant to unify distinct domains is analogous to the alignment of diverse ontologies. One approach to such ontology alignment is the identification of contrast, boundary, and common concepts, which help demarcate the conceptual borders of domains. Contrast concepts encompass differences between domains, and are typically excluded from a reference ontology, because they are domain-specific. Boundary concepts conceptually straddle domains, emphasizing linkages between them, such as the connecting elements of the water cycle, e.g. baseflow and infiltration describe the flow of water between the surface and subsurface. Because boundary concepts serve to bridge domains, they are good candidates for inclusion in a reference ontology. Lastly, common concepts consist of generalizations that apply across domains, and can serve as a framework for them—they are therefore integral to a reference ontology.

In this paper we focus exclusively on identifying and representing concepts common to the subsurface and surface hydro domains, leaving boundary concepts to future work. Ontological analysis is used to identify the concepts, and some recently defined evaluation criteria are used to assess the results: intelligibility (are the results understandable to domain experts?), fidelity (is the domain accurately represented?), fitness (does the ontology satisfy its usage requirements?), and consistency (are the results internally consistent?)

3. RESULTS – A NASCENT REFERENCE HYDRO ONTOLOGY

The hydro ontology development is divided into two parts: (1) an ontological analysis of the subsurface and surface water domains using the hydro container schema as an overarching framework, and (2) representation of this schema within the DOLCE foundation ontology.

3.1 The hydro container schema: common hydro concepts

In the same way that the hydrologic cycle can be analyzed to provide boundary concepts, the hydro container schema can be analyzed to identify common hydrologic concepts. In this sense, the hydro container schema is a general description about how water can be contained by something else. Analyzing the subsurface and surface hydro domains from this perspective reveals the following shared concepts, as shown in Figure 1 and extended from [2]:

- *Container object*: is the object that contains water, e.g. the ground surrounding a river channel, or the subsurface rock body identified as an aquifer.
- *Container matter*: is the material that constitutes the container object, such as the rock material that makes up the wall of a river channel or that constitutes the aquifer.
- *Void*: is the space in the container that can be filled with water, e.g. the space in a river channel, or the pores or fractures in an aquifer.
- *Water body*: is the object that is constituted by water. It is differentiated from the water matter, because the matter will change over time: a specific river is always the same river, but its matter will change as quickly as it flows, e.g. the St. Lawrence River is a water body made up of various samples of water that flow through it over time.
- *Water matter*: is the water material that makes up a water body and flows within it.
- *Water flow*: denotes the flow of water within a water body. The flow concept is out of scope for this work.

The hydro container schema is not only characterized by these key entities, but also by three important relations between the entities: namely hosting, containment, and constitution:

- *Hosting*: refers here to the relation between a physical object and a specific hosted feature, often voids. Examples include a body of ground hosting a depression (a river channel), or an aquifer hosting pores or fractures. The void-hosting relation has been formalized for hydro ontology in [4].

- *Containment*: refers here to the relation in which something is inside a void, or surrounded by a physical container. For example, water is inside the ground depression of the river channel, inside the pores of an aquifer, or it is surrounded by the channel or aquifer. Containment relations have been formalized for hydro ontology in [5].
- *Constitution*: refers here to the relation between a physical object and the matter it is made of, e.g. the soil or rock matter that make up the walls of a river channel or comprise an aquifer, or the water matter that makes up a water body. The constitution relation has been formalized with hydro ontology in mind in [6].

Figure 2 illustrates these relations. Note that matter has been conflated into one entity, with the understanding that this entity encompasses all forms of matter, including rock matter, soil mater, and water matter. Figure 2 also specifies that a hydro container is essentially a physical object that serves as a conduit for water, and in natural settings consists of a material body wholly or partially occupying the ground surface or subsurface.

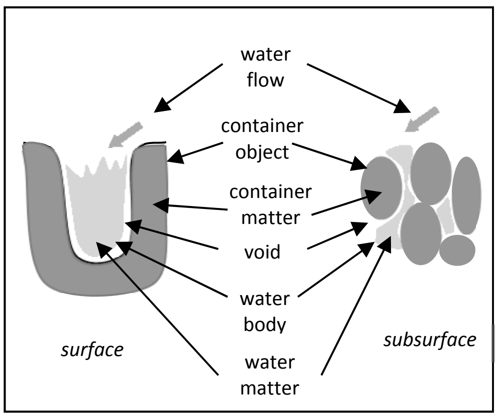

Figure 1: hydro container schema entities

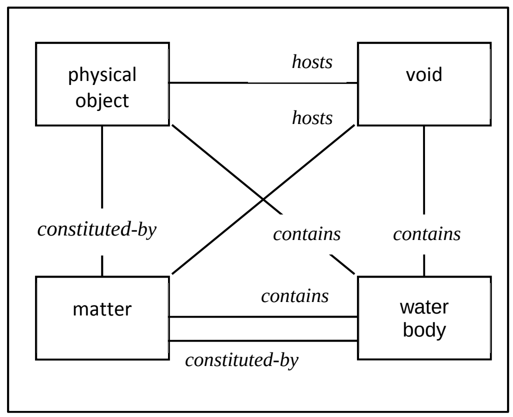

Figure 2: hydro container schema relations

3.2 DOLCE representation of the container schema

DOLCE is a foundational ontology that provides a conceptual superstructure—a suite of general kinds of entities and relations—that can be specialized by domain-specific elements, such as the common hydrologic concepts and relations from the hydro container schema. DOLCE not only provides a conceptual template for ontology design, it also promotes quality assurance through a focus on formal ontology development that helps avoid logical and conceptual inconsistencies. As shown in Figure 3, the hydro container schema (light boxes) fits naturally into DOLCE's structure (dark boxes). Matter, including water and rock matter, specialize DOLCE's `dol:Amount-of-Matter`, bodies of ground and water bodies specialize `dol:NA-Physical-Object`, and voids specialize `dol:Feature`. The DOLCE entities are characterized as follows:

- `dol:Physical-Endurant`: denotes physical entities of all kinds, defined as being located in space and fully present at any point in time during their existence. This latter property distinguishes them from process-like entities, which are never fully present at any particular time point, because their entirety evolves over time, e.g. an earthquake is never fully present at any point in time, unlike an aquifer. The three specializations are `dol:NA-Physical-Object`, `dol:Ampount-of-Matter`, `dol:Feature`.

- `dol:NA-Physical-Object`: denotes physical entities that are not intentional agents, but are ideal for representing containers such as rock bodies, as well other physical things they contain such as water bodies.
- `dol:Amount-of-Matter`: denotes a lump of stuff that makes up a physical object, such as water constituting a water body, or sandstone constituting an aquifer.
- `dol:Feature`: is a parasitic entity hosted by a physical entity, such as a canyon or river bank hosted by a body of ground. Voids, the spaces occupied by water, are clearly features hosted by some physical object.

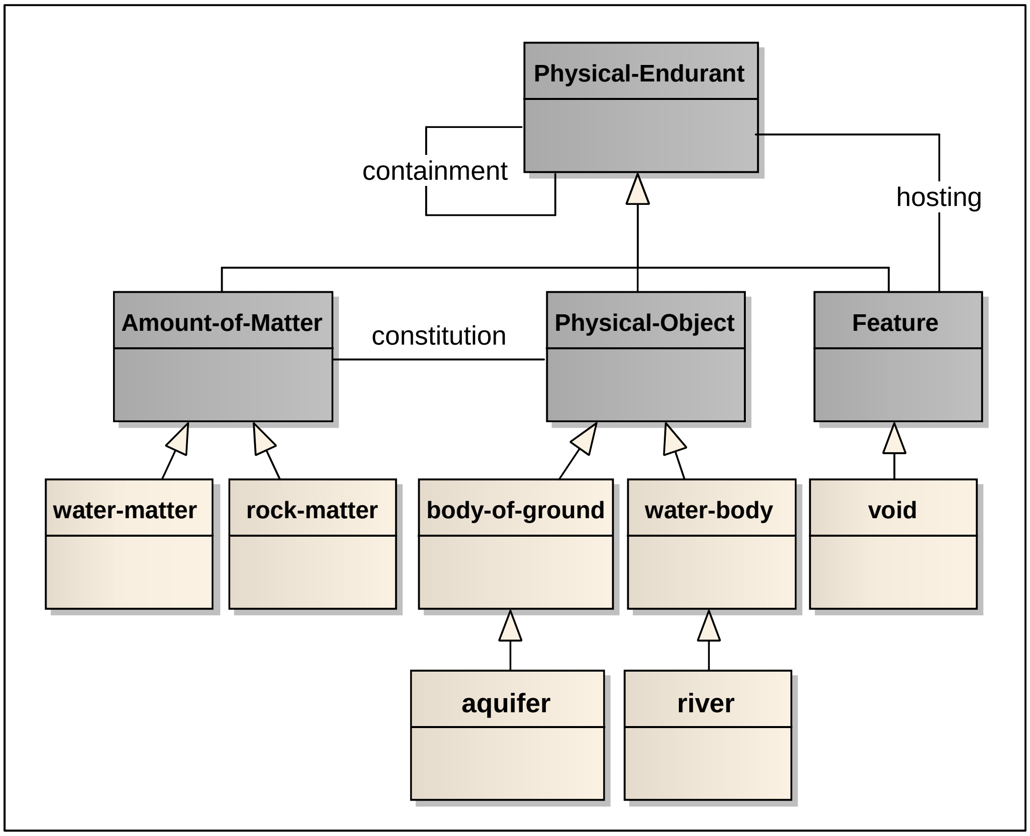

Figure 3: hydro container schema in the DOLCE foundational ontology

DOLCE's structure also enables a coarse representation of the relations from the hydro container schema: the hosting relation holds only between a physical endurant and a feature, the constitution relation only between a physical object and an amount of matter, while the containment relation can hold between arbitrary physical endurants, in the sense that physical objects and amounts of matter can enclose each other as well as voids, and likewise these entities can be inside voids. Note that for hydro ontology purposes we utilize an augmented form of these relations that enhance DOLCE's native offerings [4, 5, 6]. Also note that Figure 3 illustrates only a portion—the essential core—of the developed hydro ontology.

3.3 Evaluation of results

The hydro ontology developed from the container schema is evaluated in terms of intelligibility, fidelity, fitness, and consistency. Intelligibility is achieved through the use of domain-friendly concepts and their names with intuitive meanings, such as "water-body" or "body-of-ground". Fidelity is achieved through conformance with the hydro container schema, which we claim is a conceptual underpinning for the hydro domain. Fitness for use remains to be determined, as the ontology is still in development, though we hope to utilize it to mediate between surface water and groundwater data standards in online data interoperability. Consistency is enforced through the inherited DOLCE structure and via its logical encoding and testing. To further help verify consistency and fitness, the ontology is represented in Common Logic, which is rigorous enough to allow automated reasoning in response to competency questions. This will eventually help evaluate whether the ontology meets basic requirements, or has either desirable or unwanted implications, and it will also help determine their scope, i.e. whether they are (or should be) true within all or some of the ontology.

4. SUMMARY AND FUTURE DIRECTIONS

Ontological analysis, consisting of the examination of the logical first principles common to the subsurface and surface water domains, is used here to identify the hydro container schema as a conceptual foundation for hydro ontology development. The schema consists of four entities: a *physical object* that contains a *water body*, the *voids* hosted by the container and filled by the water body, and the *matter* that constitutes the container and water body. The schema also consists of three relations that bind these entities logically: *constitution* between a physical object and its matter, *hosting* of a void by a physical object, and *containment* of a physical object, matter, or void, by another. The most pressing conceptual issues involve assimilation of the flow concept, and eventual inclusion of atmospheric water. A key implementation issue is migration of the logical theory into a more prevalent ontology language such as OWL/RDF. The latter should be aided by the automated reasoning capabilities afforded by our Common Logic representation, which enables largely automated conversion to other formats that are less formal or expressive, e.g. to database or RDF schema. Finally, we also anticipate use of the ontology for interoperability across emerging groundwater and surface water data networks.

REFERENCES

[1] *Beran, B., Piasecki. M. (2009) Engineering New Paths to Hydrologic Data. Computers & Geosciences, 35(4):753–760*

[2] *Boisvert, E., Brodaric, B. (2012) GroundWater Markup Language (GWML) – enabling groundwater data interop. in spatial data infrastructures. J. Hydroinformatics 14(1):93–107.*

[3] *Dornblut, I., Atkinson, R. (2014) OGC HY_Features: a Common Hydrologic Feature Model. Open Geospatial Consortium Technical Report OGC 11-039r3, 55pp.*

[4] *Hahmann, T., Brodaric, B. (2012) The void in hydro ontology. In: Conf. on Formal Ontology in Inf. Systems (FOIS-12), IOS Press, 45–58.*

[5] *Hahmann, T., Brodaric, B. (2013) Kinds of full physical containment. In: Conf. on Spatial Inf. Theory (COSIT-13), Springer , 397-417.*

[6] *Hahmann, T., Brodaric, B. (2014) Voids and material constitution across physical granularities. Submitted to Conf. on Formal Ontology in Inf. Systems (FOIS-14) .*

[7] *Hart, G., Dolbear, C., Goodwin, J., Feliciter. L. (2007) Using Structured English to represent a Topographic Hydrology Ontology. In OWL Experiences & Directions Workshop.*

[8] *INSPIRE Thematic Working Group Geology (2013) D2.8.II.4 INSPIRE Data Specification on Geology – Draft Guidelines, v3.0 rc3. Tech. report, INSPIRE, 369pp.*

[9] *INSPIRE Thematic Working Group Hydrography (2009) D2.8.I.8 INSPIRE Data Specification on Hydrography – Guidelines, v3.0. Tech. report, INSPIRE, 175pp.*

[10] *Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A. (2003) WonderWeb Deliverable D18 – Ontology Library. Technical report, ISTC-CNR, Trento, 349pp.*

[11] *Santos, P., Bennett, B., Sakellariou, G. (2005) Supervaluation Semantics for an Inland Water Feature Ontology. In: Int. Joint Conf. on Artif. Intell. (IJCAI-05), 564–569.*

[12] *Tripathi, A., Babaie, H. (2008) Developing a modular hydrogeology ontology by extending the SWEET upper-level ontologies. Comp. & Geosci. 34(9):1022–1033.*