The Void in Hydro Ontology

Torsten HAHMANN [a], Boyan BRODARIC [b]

[a] *Department of Computer Science, University of Toronto, Toronto, ON, Canada*
[b] *Geological Survey of Canada, Natural Resources Canada, Ottawa, ON, Canada*

Abstract. Voids are extremely important to water science, because their size and connectivity determines the storage and flow of water both above and below the ground surface. While previous formal theories about voids strictly consider holes hosted inside objects, we generalize voids to also include spaces between objects, and distinguish voids in macroscopic objects from those occurring microscopically in an object's matter. These notions are axiomatized in first-order logic as an extension of the DOLCE ontology, and are applied to key aspects of hydrology and hydrogeology, laying the groundwork for a foundational hydro ontology.

Keywords. hydro ontology, hydrogeology, mereotopology, constituency, physical void, hole, gap, pore space, DOLCE

1. Introduction

Water is a natural resource necessary for human life. It is found in the atmosphere, on the ground surface, and in the subsurface. Environmental conditions drive the movement of water between these spheres, as reflected in the well-known water cycle. Many scientific and social issues require a sophisticated understanding of the water cycle, including climate change, flood risk, and groundwater contamination. Such issues also require increased access to large volumes of water data, which are starting to be provided by Spatial Data Infrastructures (SDI) [16] in countries such as the USA, Canada, Australia, and Europe. However, at present there exist multiple SDI data standards emerging for both the surface and subsurface water domains, and these are being developed somewhat independently leading to incomplete and incompatible representations, and a lack of coupling between them. This hinders water cycle modelling, which requires an integrated approach to water data, and signals a need for the development of a reference ontology to bridge and disambiguate conceptual differences. Although the water cycle provides boundary entities that straddle the surface and subsurface domains, e.g. the notion of 'baseflow' as the discharge of water from subsurface to surface, a strict focus on boundary entities ignores shared foundational aspects, such as a container schema which represents water contained in holes or gaps. Thus, an important component of a reference hydro ontology is the unified representation of boundary entities and common foundational entities. In this paper we begin to develop such a representation, by building a formal theory for specific aspects of the physical containment of water. In particular, we classify and characterize empty spaces – voids – that can be filled with water, and show how the resulting distinctions can be used to help characterize key water entities, such as an aquifer or lake.

This work makes the following original contributions: it generalizes formal theories of holes to voids, to include both spaces within objects (holes) and between ob-
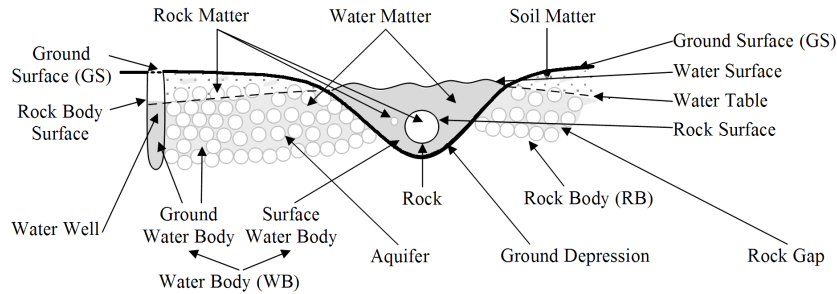
Figure 1. Examples of physical objects (rock body, water body, aquifer, rock) and physical voids (water well, ground depression, gaps between rocks) in hydrogeology. Different kinds of matter are shown at the top, and various features are illustrated on the right (ground surface, water surface, rock surface, water table).

jects (gaps); it provides a physical characterization of voids, distinguishing and relating voids hosted by physical objects and the matter that constitutes them, leading to a refined taxonomy of voids; it specializes the DOLCE [17] foundational ontology with physical voids and associated relations, and in doing so it formalizes some key entities in hydro ontology using first-order logic, laying a basis for its further development. In this sense, this work is very much an ontology engineering task, extending an existing foundational ontology through the refinement of some fundamental distinctions about voids.

The paper is organized as follows: Sec. 2 discusses semantic issues in water data standards that motivate this work, and Sec. 3 identifies gaps in related work on voids. In sections 4, 5, and 6 we develop the formal theory on physical voids, focusing respectively on spatial regions, physical objects, and physical voids, with application mainly to hydrogeology. Sec. 7 concludes with a brief summary and a note about future directions.

2. Ontological Issues in Groundwater Data Standards

The lack of surface-subsurface integration, and the occurrence of semantic incompatibilities, in present water data standards can be demonstrated by comparing emerging groundwater data schema: the INSPIRE schema [15] is a Europe-wide initiative, and the Groundwater Markup Language (GWML) [3] is mainly used in North America. Comparison of the schemas reveals the following issues:

• **Semantic ambiguity:** it is unclear whether the INSPIRE GroundWaterBody refers to a specific amount of water, which might change location, or to the object consisting of water but fixed to a specific location? E.g. the water body in the Ogallala aquifer of the US Great Plains is a timeless entity tied to the location of the aquifer, whereas the specific amounts of water that constitute the water body change over time – they can enter or leave the aquifer. The related GWML GroundwaterBody is clearer, as it distinguishes a water body object from the matter that comprises it, but this still leaves in question the compatibility of the INSPIRE and GWML entities.

• **Semantic incompleteness:** an aquifer is characterized similarly in both INSPIRE and GWML, however both UML representations are incomplete, capturing only a fragment of the intended meaning evident in the accompanying text, i.e. both indicate that an aquifer is a rock body, but do not capture the fact that an aquifer is wet, porous, permeable, and yields water to wells. The expressivity limits of UML are partially responsible, but in the end the conditions expressed in the text are not completely formalized.

• **Semantic granularity:** elements of the containment schema are quite general in GWML: GWML Reservoir denotes the sum of fillable empty spaces in a rock body, but voids in general are not further differentiated; thus, it is impossible to distinguish large holes such as caves from minute gaps such as spaces between rock grains. The INSPIRE schema lacks voids altogether. This highlights a significant difference between GWML and INSPIRE: a GWML aquifer is a rock body that hosts a space filled by a water body composed of groundwater, and in INSPIRE it is a rock body that does not host any space nor contain groundwater directly, rather it contributes to a whole that consists of the rock body and water body as parts.

• **Surface-subsurface disconnection:** the groundwater schemas are largely disconnected from relevant surface water schemas, with neither boundary nor shared entities represented. The INSPIRE schemas are an exception, in that a specialization of GroundWaterBody exists in both the surface and subsurface schemas, but with different parent entities in each schema, leading to further ambiguities about the intended meaning of the entity.

3. Hydro Ontology and Related Work

Key to our approach is the notion that a containment schema is central to hydro ontology. In the schema, a physical container hosts a void in which water can be stored and through which it can flow. Voids are considered here to be physical entities devoid of the hosting body's matter, but that can be filled by other matter such as gases, liquids, or solids. The nature of the containers and voids is different for surface and subsurface water entities, as illustrated in Figure 1. Surface water is stored and transmitted in depressions in the ground surface, such as in lakes or rivers. Subsurface water exists in the gaps between unconsolidated materials such as sand or gravel, in the gaps between the grains and crystals that make up consolidated rock bodies, and in the spaces between rock bodies or cavities within them, such as caves. The size and connectivity of voids determines important qualities associated with the storage and flow of water in both the surface and subsurface, and these constitute perhaps the most important attributes for water science.

As the focus of this paper is on voids and their relation to hosting containers, in a hydro context, the relevant related work includes formal theories of holes, places, and other hydro entities. The seminal formal theory on holes is developed by Casati & Varzi [6]. Holes are self-connected empty spaces, classified as cavities (e.g. caves), tunnels (e.g. donut holes), or hollows (e.g. canyons), and their hosts are non-scattered wholes. Holes are thus appropriate to represent surface water hollows as well as major subsurface cavities, but gaps between objects are not included, a critical shortcoming in the representation of subsurface voids. The relations between voids at different physical scales is also not considered, such as between macroscopic voids in an object and the microscopic voids in the matter constituting the object; e.g. between a canyon containing a lake and the tiny spaces in the aquifer materials beneath the lake, as shown in Fig. 1. In other work, formal representation of holes is also discussed in relation to relative places and surface features, such as cracks, but these are less relevant to hydro ontology [10,13]. Voids are typically not represented in foundational ontologies, though most such as DOLCE possess a general category for dependent entities [17]. Existing work on hydro ontology focuses on the physical container for surface water – to enable identification and classification [14,18], on vocabularies mainly for water constituents [1,2], and on basic hydrogeology entities [5,19], but voids are not considered in any substantial way.

4. Spatial Regions

A central goal of this paper is to represent aspects of voids from primarily a hydrogeological perspective. For that purpose, we are only concerned here with enduring entities that are located in physical space such as rock formations, sediments, and various kinds of water-related bodies such as rivers, lakes, groundwater, aquifers, and wells. Perdurants such as processes, plus non-physical entities, are out of scope, and while dependent qualities such as volume or depth are important to water science, they are of secondary concern here and are not considered in this work.

A key notion here is that of a spatial region, which we discuss from a geometrical, indeed mereotopological, perspective. We distinguish the physical space populated by real physical entities from an abstract space populated by spatial regions, which are of purely geometrical and topological nature. This distinction allows us to consider abstract space as a mathematical-logical construct, which provides flexibility in spatial operations. To map entities located in physical space to their associated abstract spatial regions, we reuse the region function $r(x)$ from layered mereotopology [8,9][1]. The range of the region function defines the DOLCE category of 'spatial region' S (S1, S2, S-T1). We refer to entities of this category henceforth simply as 'regions'. Throughout, all axioms and definitions are assumed to be implicitly universally quantified.

(S1) $S(r(x))$ (the range of the region function are spatial regions)
(S2) $S(x) \leftrightarrow x = r(x)$ (spatial regions are their own region)
(S-T1) $r(r(x)) = r(x)$ (region function idempotent; from S1 and S2)

Regions are related to each other by spatial inclusion $\subseteq_r$ $(S3)$[2]. We say $x \subseteq_r y$ iff the spatial region x is a subregion of y. $\subseteq_r$ is the dimension-independent mereological relation (S4, S5, S6) adapted from [11,12], with the original axiom numbering included in parentheses. For convenience we maintain $ZEX(x)$ to denote a unique zero region of no extent and no location (S7). While $\subseteq_r$ is restricted to regions, it can be extended as $\subseteq$ to non-regions (S8) so that all subsequently defined relations equally apply to regions and non-regions, unless otherwise noted. We further define proper spatial inclusion $\subset$ (S9) and contact C (C-D) in terms of $\subseteq$. If an entity x is in contact with a proper subset of the entities some y is in contact with, x must be properly spatially included in y (S10).

To compare regions dimensionally, we reuse the axiomatization of the primitive relation $x <_{\dim} y$, meaning 'x has a lower dimension than y', from [11]. Note that, e.g., a surface (no thickness) is always considered to have dimension two even when it curves in 3D space, because it takes the dimension of the smallest embedding topological space. We also use the defined relations $\leq_{\dim}$ (lower or equal dimension) and $\prec_{\dim}$ (preceding, i.e. next-lowest dimension). For simplicity, we say that a non-region has the same dimension as the region it occupies (S11), thus extending the scope of $<_{\dim}$ to non-regions.

(S3) $x \subseteq_r y \rightarrow S(x) \wedge S(y)$ (region inclusion)
(S4) $S(x) \wedge \neg ZEX(x) \leftrightarrow x \subseteq_r x$ (C-A1: $\subseteq_r$ reflexive)
(S5) $x \subseteq_r y \wedge y \subseteq_r x \rightarrow x = y$ (C-A2: $\subseteq_r$ antisymmetric)
(S6) $x \subseteq_r y \wedge y \subseteq_r z \rightarrow x \subseteq_r z$ (C-A3: $\subseteq_r$ transitive)

---

[1]In the context of DOLCE, the function $r(x)$ can be see as a function returning a 'spatial quale' that is related to the entity x at the predetermined time point by the 'spatial location' quality.

[2]Spatial inclusion is a simplified version of DOLCE's temporary spatial inclusion relation $\subseteq_{s,t}$

(S7) $ZEX(x) \rightarrow S(x) \wedge \forall y\, [x \nsubseteq_{\mathrm{r}} y \wedge y \nsubseteq_{\mathrm{r}} x]$ (C-A4: no inclusion for ZEX)

(S8) $x \subseteq y \equiv r(x) \subseteq_{\mathrm{r}} r(y)$ (spatial inclusion for regions and non-regions)

(S9) $x \subset y \equiv x \subseteq y \wedge y \nsubseteq x$ (proper spatial inclusion)

(C-D) $C(x, y) \equiv \exists z\, [z \subseteq x \wedge z \subseteq y]$ (contact: a shared entity exists)

(S10) $\neg ZEX(x) \wedge \forall z\, [C(z, x) \rightarrow C(z, y)] \wedge \exists z\, [C(z, y) \wedge \neg C(z, x)] \rightarrow x \subset y$

                               (C-A5: C strictly monotone implies proper spatial inclusion)

(S11) $x =_{\dim} r(x)$ (dimension of a non-region is the dimension of its occupied region)

We define a (proper) region part to be a (properly) contained entity of equal dimension (EP-D, EPP-D). Note that P and PP do not function as general mereological predicates here, that is, they do not capture part-whole relations between real physical entities and their parts (though spatial parthood P is often a prerequisite for such a relation). Each entity can only have parts of equal dimesion, for example, the sum of multiple points is always assumed to be of the same dimension as a single point. But regions of different dimensions can coexist is a single model. In particular, a region may spatially include other regions of lower dimensions. E.g., a 3D region can include other 3D regions, but also 2D areas, 1D line segments, or 0D points. S12 governs the interaction between spatial inclusion and relative dimension.

Essentially, the axioms for $<_{\dim}$ from [11] ensure that the relative dimension relation forms a linear discrete order over classes of entities of equal dimension. The linear order is bounded in either direction, that is, a lowest dimension (that of the zero region) and a maximal dimension must exist. We denote the entities of minimal non-zero dimension and of maximal dimension by $MinDim(x)$ and $MaxDim(x)$.

Two kinds of contact from [11,12] are relevant here: partial overlap $PO(x, y)$ holds when x and y share a part ('partial' for overlapping in a part) of equal dimension (PO-D), while superficial contact $SC(x, y)$ holds when x and y only share entities of lower dimensions (SC-D). A third type of contact, incidence, can also be described but is not necessary here. Using partial overlap, we enforce strong supplementation for parts (S13).

(EP-D) $P(x, y) \equiv x \subseteq y \wedge x =_{\dim} y$ (equi-dimensional parthood)

(EPP-D) $PP(x, y) \equiv P(x, y) \wedge r(x) \neq r(y)$ (equi-dimensional proper parthood)

(S12) $x \subseteq y \rightarrow x \leq_{\dim} y$ (CD-A1: inclusion requires lower or equal dimension)

(PO-D) $PO(x, y) \equiv \exists z\, [P(z, x) \wedge P(z, y)]$ (contact: partial overlap)

(SC-D) $SC(x, y) \equiv \exists z[z \subseteq x \wedge z \subseteq y] \wedge \forall z[z \subseteq x \wedge z \subseteq y \rightarrow z <_{\dim} x \wedge z <_{\dim} y]$

                                               (contact: superficial contact)

(S13) $\neg ZEX(x) \wedge \neg ZEX(y) \wedge \neg P(y, x) \rightarrow \exists z\, [P(z, y) \wedge \neg PO(z, x)]$

                                         (EP-E2: strong supplementation)

We further assume that the set of regions of maximal dimension (3D regions in DOLCE) is mereologically closed, i.e., that for every pair of spatial regions x and y with $MaxDim(x)$ and $MaxDim(y)$, the intersection $x \cdot y$, the difference $x - y$, and the sum $x + y$ are defined (they may yield the zero region). Moreover, we assume that a universal region S_{u} of maximal dimension exists with $S(S_{\mathrm{u}}) \wedge MaxDim(S_{\mathrm{u}}) \wedge \forall x[x \subseteq S_{\mathrm{u}}]$. Then the complement $x' = S_{\mathrm{u}} - x$ is defined for any entity x of maximal dimension. Note that these closure operations apply only to regions; we do not force the set of physical entities to be closed in the same way. The details of the definitions of the mereological closure operations are not important here[3], it is sufficient that they allow us to define

---

[3]The complete axiomatization can be found at `www.cs.toronto.edu\~torsten\hydro`

the well-known notion of (self-)connectedness (Con-D) and the lesser-known notion of interior-connectedness (ICon-D) (meaning the interior of $x + y$ is a single piece; also known as strong self-connectedness) defined in terms of strong connection (C_S-D) [4]. Two entities are strongly connected (short: *s-connected*), $C_S(x, y)$, if they are superficially connected by sharing an entity of the next lowest dimension, $\prec_{\dim}$. For example, two 3D bodies are s-connected if they touch in a 2D surface, but not if they only touch in a line segment or in points. For our purposes here, we assume the universal region S_u to be interior-connected, i.e., $ICon(S_u)$, which further ensures that any $x \neq S_u$ is superficially connected to its complement x', i.e., $SC(x, x')$.

(C_S-D) $C_S(x, y) \equiv SC(x, y) \wedge x =_{\dim} y \wedge r(x) \cdot r(y) \prec_{\dim} x$ (strongly connected)
(Con-D) $Con(x) \equiv \forall y[PP(y, x) \rightarrow C(y, r(x) - r(y))]$ (self-connectedness)
(ICon-D) $ICon(x) \equiv \forall y[PP(y, x) \rightarrow C_S(y, r(x) - r(y))]$ (interior-connectedness)

Finally, we supplement our theory of spatial regions by a primitive convex hull function $ch(x)$. For a detailed axiomatization of ch we refer to [7,9]; here we only provide the axioms needed to fit ch into our spatial theory, ensuring that the convex hull is always a region (CH1), is equally defined for non-regions (CH2), and is interior-connected (CH3).

(CH1) $S(ch(x))$ (convex hull ch is a spatial region)
(CH2) $ch(x) = ch(r(x))$ (ch defined with respect to occupied regions)
(CH3) $\neg ZEX(x) \rightarrow ICon(ch(x))$ (ch is interior connected)

5. Physical Endurants

In addition to spatial regions, we also consider entities located in physical space, referred to as physical endurants, PED, in DOLCE[4]. Indeed, we restrict ourselves to these two disjoint categories of entities (P1). In DOLCE, it is assumed that physical endurants are 'real' in the sense that they occupy a spatial region of maximal dimension (P2) and are constituted by matter. E.g. in 3D space, every physical entity must be 3D[5] and occupy a non-zero region (P-T1). While we can talk about lower-dimensional abstractions such as lines in abstract geometrical space (the category S), those abstractions have no physical equivalent in PED.

DOLCE distinguishes three disjoint subcategories of physical endurants (P3, P4): physical objects POD (e.g. a body of water), amounts of matter M (e.g. the water that constitutes a body of water), and features F (e.g. the water surface). Moreover, the hydrogeological objects considered here are all non-agentive physical objects $NAPO$, i.e., physical objects that do not act by themselves or pursue goals (P5).

(P1) $PED(x) \leftrightarrow \neg S(x)$ (regions and physical endurants exhaustive and disjoint)
(P2) $PED(x) \rightarrow MaxDim(x)$ (PEDs occupy space of codimension 0)
(P-T1) $PED(x) \rightarrow \neg ZEX(r(x))$ (from P2, D-D5, D-A5)
(P3) $PED(x) \leftrightarrow POD(x) \vee M(x) \vee F(x)$ (exhaustive categories of PED)
(P4) $\neg[POD(x) \wedge M(x)] \wedge \neg[POD(x) \wedge F(x)] \wedge \neg[M(x) \wedge F(x)]$
                      (physical objects, amount-of-matters, and features are disjoint)
(P5) $NAPO(x) \rightarrow POD(x)$ (non-agentive physical objects)

---

[4] An essential criteria for something to be an endurant is that its parts are wholly present at any point in time. This requires any proposition about an endurant to be relative to a timepoint. For simplicity we omit the time reference here; our axiomatization can be thought of as capturing an extract of the world at a fixed timepoint.

[5] Our axiomatization works with any number of dimensions; we can deal with objects in 2D such as in Fig. 1.

5.1. Features

Physical features, F, depend on other physical endurants as 'host' (P6, P7; cf. [6]), the hosting relation being asymmetrical (P8). Physical features must not be confused with geometric abstractions such as boundaries, which may be captured as lower-dimensional spatial regions. In DOLCE, physical features are specialized as relevant parts RPF, e.g. bumps, edges, surfaces, or dependent places DPF, e.g. shadows and holes (P9, P10). The main difference is that RPFs are constituted by their host's matter and are thus a spatial part thereof (P11), while DPFs cannot overlap their host (P12).

(P6) $hosts(x, y) \rightarrow PED(x) \wedge F(y)$ (only features are hosted; the host being a PED)
(P7) $F(x) \leftrightarrow \exists y[hosts(y, x)]$ (any feature must be hosted)
(P8) $hosts(x, y) \rightarrow \neg hosts(y, x)$ ($hosts$ relation asymmetric)
(P9) $F(x) \leftrightarrow RPF(x) \vee DPF(x)$ (RPF and DPF are exhaustive classes of features)
(P10) $\neg RPF(x) \vee \neg DPF(x)$ (RPF and DPF are disjoint classes)
(P11) $hosts(x, y) \rightarrow [RPF(y) \leftrightarrow P(y, x)]$ (RFP is part of its host)
(P12) $hosts(x, y) \rightarrow [DPF(y) \leftrightarrow \neg PO(y, x)]$ (DPF does not overlap its host)

5.2. Matter

DOLCE further distinguishes a physical object from the matter that constitutes it, e.g. a body of rock and some rock matter, or a water body and some amount of water. Rock matter can vary in its degree of consolidation from unconsolidated material, such as sand or gravel, to consolidated material composed of grains or crystals, such as sandstone or granite. Soil is a mixture of rock, water, organic matter, and gases, and water is primarily H_2O with other suspended or dissolved materials, most notably rock or soil matter. The specific amount of matter that constitutes a physical object can change over time, e.g. the water in a surface water body changes due to evaporation, precipitation, and water flow, but the object itself endures wholly at every timepoint of its existence: the water body that forms Lake Ontario persists even when its water matter is completely exchanged. Here, we restrict ourselves to a single timepoint, and limit constitution to be a relation between some matter and a physical object or its relevant-part feature.

We use a binary version of DOLCE's direct constituency $DK(x, y, t)$ [17] fixed to some t. DOLCE's relation DK requires a physical endurant and its matter to be spatially coincident (at time t), that is, $r(x) = r(y)$, which implies maximal coverage for matter, e.g. the sum of all solids, gaseous, and liquid matter in an object. Here we are mainly interested in the *entirety* of an object's *primary* matter. For example, we may ignore the air in an aquifer. Hence, our weakened constitution relation DK_1 uses $x \subseteq y$ (DK$_1$-D). Moreover, DK_1 is restricted to the first step in physical scale, i.e., the direct constituency of a physical object to its matter, such as a rock to granite, as opposed to, e.g., the atomic or molecular constituency of an amount of matter, such as some granite to its chemical composition. Therefore, DK$_1$ uniquely identifies the entire matter that constitutes an object (P13), rendering DK_1 irreflexive and asymmetric.

(DK$_1$-D) $DK_1(x, y) \rightarrow M(x) \wedge [POD(y) \vee RPF(y)] \wedge x \subseteq y$
                    (primary direct constitution of an object or relevant-part feature by matter)
(P13) $DK_1(x, y) \wedge DK_1(z, y) \rightarrow x = z$ (an object's constituent matter is unique)

Now we can state that every water body WB is a $NAPO$ (non-agentive physical object) that consists of some amount of water if it has some constituent at all. This allows for notions such as a 'dry lake' or 'dry aquifer' – physical objects that can be without water at some time point. In contrast, any rock body RB must be constituted by some amount

of matter, which can only be rock matter[6], whereas soils are minimally composed of rock matter or organic matter, but might have other stuff as well[7].

$$Water(x) \vee RockMatter(x) \vee OrganicMatter(x) \rightarrow M(x)$$

$$Soil(x) \rightarrow M(x) \wedge \exists y[P(y,x) \wedge (RockMatter(y) \vee OrganicMatter(y)]$$

$$WB(x) \rightarrow NAPO(x) \wedge \forall y[DK_1(y,x) \rightarrow Water(y)]$$

$$RB(x) \equiv NAPO(x) \wedge \exists y[DK_1(y,x)] \wedge \forall y[DK_1(y,x) \rightarrow RockMatter(y)]$$

Aquifers are more complex in that an aquifer consists of some porous rock matter with its accessible spaces completely filled with an amount of water that can flow to wells[8]. To begin, we define a so-called HydroRockBody, which consists of two parts, a rock body and a water body, whose sum region overlaps exactly the HydroRockBody's region.

$$HydroRockBody(aq) \rightarrow NAPO(aq) \wedge \exists rb, wb\big[r(aq) = r(rb) + r(wb) \wedge$$
$$RB(rb) \wedge WB(wb)\big] \tag{AQ1}$$

AQ1 ensures that the rock body and water body that constitute a HydroRockBody cannot overlap. However, AQ1 is by no means a complete definition: we have not expressed that the water body only occupies spaces within the rock body. To do so, we must represent such 'void' spaces, which are the focus of the remainder of our paper.

6. Physical Voids

A physical void is, intuitively, a physical feature whose region is not occupied by the region of its physical host. Examples are holes of maximal dimension such as caves or canyons, which stand in contrast to cracks that we model as lower-dimensional [13]. According to Casati & Varzi [6] a hole can only exist if a physical endurant's region is strictly smaller than its convex hull. The same applies to a void. We will capture this necessary condition by first defining an abstract spatial notion of a void region – a spatial region not occupied by a specific physical entity (the host) but inside the host's convex hull (VS-D). Void regions rely on a specific host and are disjoint only from the host's occupied region. The void regions of a given host are the regions in which voids can possibly be spatially located. The void regions of a single host are mereologically closed under intersections and sums.

Physical voids (in short: 'voids') are real physical endurants spatially located in a void region (V1). We use the relation of 'hosting a void' ($hosts\text{-}v$) as a primitive relation between a void and its host, in the spirit of [6], who use the primitive relation of 'hosting a hole'. Thereby we generalize holes to voids, but do not address the open question about which void regions have physical void counterparts. A more thorough discussion of the challenges involved in identifying physical holes and voids is offered in [6].

It is reasonable to assume that only void regions that are s-connected to their host qualify as the regions for voids (V1). A void (V-D), may also be simple or complex depending on whether it is interior-connected (V_S-D, V_C-D). We also require a complex void to be composed of simple voids, the interior-connected parts of the complex void

---

[6]Note that on the next level of constitution, which is not captured by DK_1 and is beyond the scope of this work, the constituent of an amount of rock matter may include, amongst other things, organic components.

[7]see: Soil Science Society of America, Glossary of soil science terms, https://www.soils.org/publications/soils-glossary/

[8]see: US Geological Survey, Aquifer Basics, http://water.usgs.gov/ogw/aquiferbasics/

(V2). These axioms and definitions, together with the earlier restriction of the *hosts* relation (P12), entail that voids are dependent place features (V-T1).

(VS-D) $VS(x,y) \leftrightarrow PED(x) \wedge S(y) \wedge y \subset ch(x) \wedge \neg PO(y,x)$ (void region: spatial subregion of a physical endurant's convex hull not overlapping the endurant's region)

(V1) $hosts\text{-}v(x,y) \rightarrow hosts(x,y) \wedge VS(x,r(y)) \wedge C_S(x,y)$ (hosting a void)

(V_S-D) $V_S(y) \equiv ICon(y) \wedge \exists x[hosts\text{-}v(x,y)]$ (simple void is interior-connected)

(V_C-D) $V_C(y) \equiv \neg ICon(y) \wedge \exists x[hosts\text{-}v(x,y)]$ (complex void not interior-connected)

(V-D) $V(x) \leftrightarrow V_S(x) \vee V_C(x)$ (a void is a simple or complex void)

(V2) $hosts\text{-}v(x,y) \wedge V_C(y) \wedge PO(z,y) \rightarrow \exists v[hosts\text{-}v(x,v) \wedge V_S(v) \wedge PO(z,v)]$ (anything overlapping a complex void overlaps a simple void of the same host)

(V-T1) $V(y) \rightarrow DPF(y)$ (voids are dependent place features)

As an additional restriction, a void cannot be hosted by other voids (V3), though all other kinds of physical endurants may host voids. In particular, non-void dependent places, such as shadows, can host voids, e.g. there can be a hole in a shadow. V4 postulates that every void is hosted by some non-feature, that is, if a void is hosted by some feature, the host of that feature also hosts the void. For example, a void hosted by a surface is also hosted by the object or matter of which it is a surface. Conversely, a void in an object or in some matter must also be a void in a surface thereof (the 'void lining', compare [6]), which is a relevant part feature of the object or matter (V5).

V6 might be more controversial: while we allow voids to be spatially included in one another, we do not allow them to overlap without spatial inclusion, in order to keep things neat. For example, a canyon may be a void in the ground surface, but it may fully include a smaller void at the bottom of the canyon in which a river flows, cf. Fig. 2(d). Thus, voids do not necessarily occupy *maximal* interior-connected void regions. The reasons for this choice will become clear in the following subsections, the intuition behind it is illustrated by the rightmost example in Fig. 4.

(V3) $hosts(x,y) \wedge V(y) \rightarrow \neg V(x)$ (voids cannot host voids)

(V4) $hosts\text{-}v(x,y) \wedge RPF(x) \rightarrow \exists z[hosts(z,x) \wedge \neg F(z) \wedge hosts\text{-}v(z,y)]$
(every void hosted by a relevant part feature is also hosted by that feature's host)

(V5) $hosts\text{-}v(x,y) \wedge \neg F(x) \rightarrow \exists z[hosts(x,z) \wedge RPF(z) \wedge hosts\text{-}v(z,y)]$
(every void is hosted by some relevant part: the surface of its host)

(V6) $hosts\text{-}v(x,y) \wedge hosts\text{-}v(x,z) \wedge PO(y,z) \rightarrow y \subseteq z \vee z \subseteq y$
(overlapping voids of the same host must be related by spatial inclusion)

Voids are not necessarily preserved by parthood: the void v in Figure 2(a) is not a void hosted by the physical endurant occupying $r(x) + r(y)$ because x and y are reciprocal fillers, though v is still a void in the endurant's part x. V7 and V8 capture weaker conditions under which voids must exist in parts and wholes, strengthening A2.4 of [6]. These conditions are illustrated in Fig. 2(a-c).

(V7) $hosts\text{-}v(x,v) \wedge P(x,y) \wedge PED(y) \wedge \neg DPF(y) \wedge v \nsubseteq y \rightarrow \exists u[r(v) - r(y) \subseteq r(u) \wedge hosts\text{-}v(y,u)]$ (if a non-dependent physical endurant y with part x, which hosts void v, does not completely fill v, then $r(v) - r(y)$ must be in some void u of y)

(V8) $hosts\text{-}v(x,v) \wedge P(y,x) \wedge PED(y) \wedge \neg DPF(y) \wedge PO(v,ch(y)) \rightarrow \exists u[r(u) = r(v) \cdot ch(y) \wedge hosts\text{-}v(y,u)]$ (if void v in x overlaps the convex hull of part y of x, then y hosts a void u that occupies region $r(v) \cdot ch(y)$)
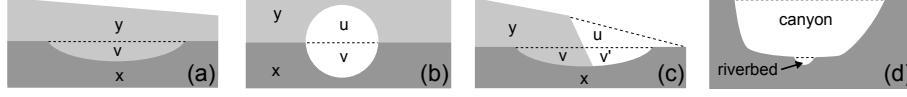
Figure 2. (a) The void v in x is not a void occupying $r(x) + r(y)$: it is completely filled by y. (b) The void with region $r(v) + r(u)$ in the physical endurant occupying $r(x) + r(y)$ has a part u with $r(u) = (r(v) + r(u)) \cdot ch(y)$ that is a void in y and a void in the entity occupying $r(x) + r(y)$. (c) The void with the region $r(v) + r(v')$ in x is not completely filled by the endurant occupying $r(x) + r(y)$: $r(x) + r(y)$ only fills v but not v'. Thus $r(v')$, a subregion of $r(v) + r(v')$, is part of the void region $r(v') + r(u)$ in the entity occupying $r(x) + r(y)$. (d) The riverbed void is spatially included in the canyon void.

6.1. Holes and gaps

While we distinguish between a simple and a complex void based on the interior-connectedness of the void, the connectedness of a void's host has yet to be examined. If the host of a void is interior-connected, we call the void a 'hole' (V9, Hole-D; following [6] while strengthening connectedness to interior-connectedness). If the host of a void is not interior-connected, i.e. disconnected or only connected in lower-dimensional spatial regions, we call the void a 'gap' (V10, Gap-D). Intuitively, a gap is the space between the parts of a scattered host, such as the gap(s) between individual pebbles in a gravel pit. In hydrogeology, gaps are most prominent in rock matter, because though a rock body may appear solid, its matter consisting of individual grains or crystals is often not s-connected, i.e., not fused together in the sense that some individual grains or crystals may only be connected at edges, leaving gaps (pores) that can be filled with water.
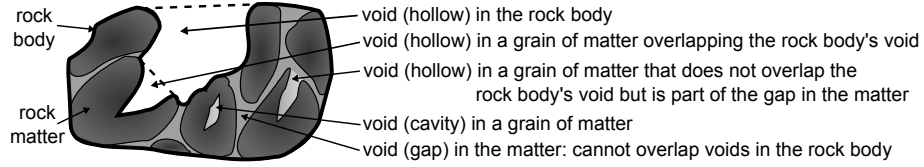
It is easily verified that for any specific host $hosts\text{-}h$ and $hosts\text{-}g$ are disjoint and exhaustive subrelations of $hosts\text{-}v$. Then holes and gaps are exhaustive categories of voids (V-T2), but some voids might be gaps and holes with respect to different hosts.

(V9) $hosts\text{-}h(x,y) \equiv hosts\text{-}v(x,y) \wedge ICon(x)$ (non-scattered host of a void)
(Hole-D) $Hole(y) \equiv \exists x[hosts\text{-}h(x,y)]$ (hole has a non-scattered host)
(V10) $hosts\text{-}g(x,y) \equiv hosts\text{-}v(x,y) \wedge \neg ICon(x)$ (scattered host of a void)
(Gap-D) $Gap(y) \equiv \exists x[hosts\text{-}g(x,y)]$ (gap has a scattered host)
(V-T2) $V_S(x) \leftrightarrow Gap(x) \vee Hole(x)$ (gap and hole exhaustive classes of simple voids)

6.2. Voids in objects and in their constituting matter

For hydrogeology it is not only important to distinguish holes from gaps, but also to distinguish macroscopic voids in an object (POD or RPF) from microscopic voids in its constituting matter. This clearly distinguishes two notions of voids that are often confused in natural language: e.g. 'a hole in the limestone' typically refers to a hole in the rock body constituted by some limestone, such as a cave, and not to the microscopic spaces between the grains of the limestone.

Thus far, voids in an object can also be voids in its matter, e.g. a cave in a rock body is also a cave in its limestone, but the converse does not hold, in that microscopic voids in matter are not voids in the object because they spatially overlap with the object's body (compare DK_1-D). To isolate the microscopic voids in an object's matter, we define 'pore space' as the voids in matter that do not overlap voids in the associated object (V11). Then, a cave in a rock body is not considered to be part of the rock body's pore space. The 'void space' of an object is its pore space together with the regions occupied by the object's voids (V12). Because $hosts\text{-}v$ is a primitive relation, the definitions of

rock body — void (hollow) in the rock body
— void (hollow) in a grain of matter overlapping the rock body's void
— void (hollow) in a grain of matter that does not overlap the rock body's void but is part of the gap in the matter
rock matter — void (cavity) in a grain of matter
— void (gap) in the matter: cannot overlap voids in the rock body

Figure 3. Examples of voids in an rock body, its matter (dark grey), and individual grains of matter. Light grey areas are gaps in rock matter, and dashed lines are some void openings. The gaps cannot overlap voids in the object. Likewise, while individual grains (physical objects) may have voids, their matter cannot host voids that would overlap the rock body's void. Then the pore space of the rock body only consists of the voids in the body's rock matter, which is guaranteed to not overlap the void space of the rock body (white area).

pore and void space presuppose an identification of all 'voids'. To properly appreciate V11 and V12, recall that PO applies equally to regions and non-regions and that strong supplementation (S13) ensures that the extension of PO uniquely identifies a region; thereby V11 and V12 effectively capture sums.

While the pore and void space of a physical endurant are spatial regions, they manifest themselves in (simple or complex) voids hosted by the endurant's matter (V13, V14). It is entailed that matter as well as dependent places have no pore space (V-T3), which is a property of an object, though they can have void space.

(V11) $PO(v, porespace(o)) \leftrightarrow \exists m \big[DK_1(m, o) \land \forall u[hosts\text{-}v(o, u) \to \neg PO(v, u)] \land \exists u[hosts\text{-}v(m, u) \land PO(v, u)] \big]$ (pore space of an object overlaps any region that overlaps some void in the matter and does not overlap the voids in the object)

(V12) $PO(v, voidspace(o)) \leftrightarrow PO(v, porespace(o)) \lor \exists u[hosts\text{-}v(o, u) \land PO(v, u)]]$
(void space of an object comprises its pore space and all its voids)

(V13) $\neg ZEX(porespace(o)) \to \exists v, m[r(v) = porespace(o) \land hosts\text{-}v(m, v) \land DK_1(m, o)]$ (non-empty pore space is the region of a void in the object's matter)

(V14) $\neg ZEX(voidspace(o)) \to \exists m, v[r(v) = voidspace(o) \land hosts\text{-}v(m, v) \land DK_1(m, o)]$ (non-empty void space is the region of a void in the object's matter)

(V-T3) $M(x) \lor DPF(x) \to ZEX(porespace(x))$
(matter and dependent places have no pore space)

For example, caves in a porous rock body are not part of its pore space, while tiny gaps and holes in its matter are part of its pore space. More generally, we can capture porous objects such as sponges with this approach. The notion of void space also allows us to refine the definition of a HydroRockBody (AQ2), to indicate its water (partially) fills voids in the rock body or the rock body's matter; and to define a (water) reservoir as a void – which must exist by V14 – that occupies the entire void space of some rock body.

$$PorousObject(x) \equiv POD(x) \land \neg ZEX(r(porespace(x))$$

$$HydroRockBody(aq) \to NAPO(aq) \land \exists rb, wb \big[r(aq) = r(rb) + r(wb) \land$$
$$RB(rb) \land WB(wb) \land r(wb) \subseteq voidspace(rb) \big] \quad \text{(AQ2)}$$

$$Reservoir(wr) \equiv V(wr) \land \exists rb[RB(rb) \land r(wr) = voidspace(rb)]$$

6.3. Cavities, tunnels, hollows, and caverns

In our final step, we categorize voids by their opening(s), i.e. their connectivity to regions not occupied by the host, such as other entities' regions and the host's void regions. The central concept of the 'opening of a void' (V15) is defined as the lower-dimensional inter-
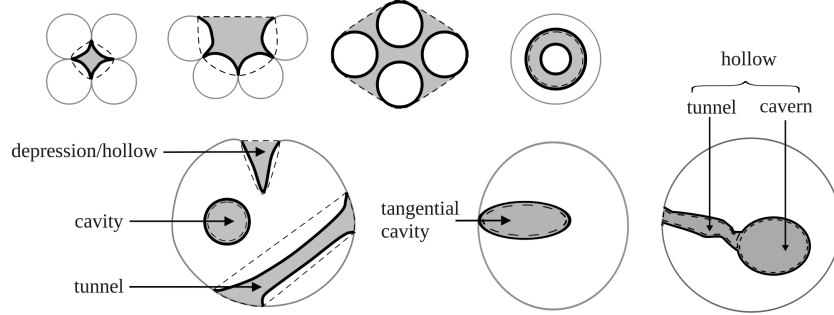
Figure 4. Examples of gaps (top) and holes (bottom), which are shown as grey areas within the convex hull (dashed lines) of their hosts (white). Thick solid lines show the surface of a host that also hosts the void.

section between the void's region and the complement of the sum of the void's and host's regions. This opening is a purely abstract spatial region of non-maximal dimension.

A cavity (V16, CAV-D) has no opening (internal cavity; V17) or has a degenerate opening that is not a surface but is a point or line (tangential cavity; V18). Hollows are depressions in an interior or exterior surface and have exactly one interior-connected surface opening (V19, HOL-D), either to the outside (external hollow; V20) or to another void of the same host (cavern; V21). Tunnels or, more generally, tunnel systems have openings that consist of multiple not s-connected pieces (V22, TUN-D). Hollows and tunnels are not required to be maximal interior-connected voids: we want to maintain the flexibility to allow, for example, a hollow to consist of a tunnel leading to a cavern, as in the rightmost example in Fig. 4. Consider Fig. 2(c) as another example: there are good reasons to call v' a void of the entity occupying $r(x) + r(y)$ – even though v' is not maximal (the void region $r(v') + r(u)$ is maximal), $r(v')$ is the greatest void region reasonably occupied by a water body such as a lake or river. In contrast, $r(u)$ is a void region but u is likely not considered a void. To correctly capture v' as an 'external hollow', V20 treats void regions not occupied by the object's voids as external. Note that cavities are implicitly required to be maximal voids. Any specific simple void in a host must be either a cavity, tunnel, or hollow and only one of those (V-T4, V-T5).

(V15) $hosts\text{-}v(x, v) \rightarrow op(x, v) = r(v) \cdot (r(x) + r(v))'$
      (definition of the opening of a void v as the boundary not shared with its host x)
(V16) $cav(x, y) \equiv hosts\text{-}v(x, y) \wedge op(x, y) \nprec_{\dim} x$ (cavity-hosting)
(CAV-D) $CAV(y) \equiv \exists x[cav(x, y)]$ (cavity: void without proper opening)
(V17) $cav_{\mathrm{i}}(x, y) \equiv cav(x, y) \wedge ZEX(op(x, y))$ (internal cavity has no op.)
(V18) $cav_{\mathrm{t}}(x, y) \equiv cav(x, y) \wedge \neg ZEX(op(x, y))$ (tangential cavity has degenerate op.)
(V19) $hol(x, y) \equiv hosts\text{-}v(x, y) \wedge op(x, y) \prec_{\dim} x \wedge ICon(op(x, y))$ (hollow-hosting)
(HOL-D) $HOL(y) \equiv \exists x[hol(x, y)]$ (hollow: void with single $ICon$ opening)
(V20) $hol_{\mathrm{e}}(x, y) \equiv hol(x, y) \wedge \exists z[P(z, op(x, y)) \wedge \forall u[hosts\text{-}v(x, u) \rightarrow \neg PO(z, u)]]$
    (external hollow: a part of the opening is not shared with any other void of the host)
(V21) $cavern(x, y) \equiv hol(x, y) \wedge \forall z[P(z, op(x, y)) \rightarrow \exists u[hosts\text{-}v(x, u) \wedge PO(z, u)]]$
      (cavern: internal hollow that only opens to other voids of its host)
(V22) $tun(x, y) \equiv hosts\text{-}v(x, y) \wedge op(x, y) \prec_{\dim} x \wedge \neg ICon(op(x, y))$ (tunnel-hosting)
(TUN-D) $TUN(y) \equiv \exists x[tun(x, y)]$ (tunnel system: void with not $ICon$ opening)

(V-T4) $[\neg cav(x,y) \wedge \neg hol(x,y)] \vee [\neg cav(x,y) \wedge \neg tun(x,y)] \vee [\neg hol(x,y) \wedge \neg tun(x,y)]$

(cavity, tunnel, and hollow are pairwise disjoint with respect to a specific host)

(V-T5) $hosts\text{-}v(x,y) \wedge V_\text{S}(y) \leftrightarrow cav(x,y) \vee tun(x,y) \vee hol(x,y)$

(cavity, tunnel, and hollow are exhaustive categories of simple void)

Note that the distinction between holes and gaps is independent of the distinction between cavities, tunnels, and hollows as Figure 4 demonstrates: a gap (top examples) can form a cavity (left- and rightmost), a tunnel (second from right), or a hollow (second from left), while a hole (bottom left and center examples) can also be any of those.

The portion of a physical object's interior-connected void or pore space that has external openings is called its *connected* void or pore space (V23, V24).

(V23) $PO(v, con\text{-}voidspace(o)) \leftrightarrow \exists u[PO(v,u) \wedge ICon(u) \wedge u \subseteq voidspace(o) \wedge$
$\quad C_\text{S}(u, (r(o) + voidspace(o))')]$ (connected void space has external opening)

(V24) $PO(v, con\text{-}porespace(o)) \leftrightarrow \exists u[PO(v,u) \wedge ICon(u) \wedge u \subseteq porespace(o) \wedge$
$\quad C_\text{S}(u, (r(o) + porespace(o))')]$ (connected pore space has external opening)

Finally, in AQ3 we can distinguish a groundwater body ($GroundWB$), such as the water body in an aquifer, from a surface water body ($SurfaceWB$), such as found in a river or lake, assuming GS denotes the ground surfaces – distinguished relevant parts of some non-agentive physical objects. We also further refine a HydroRockBody: the hosted water body must be a groundwater body and its region must be located not just in the voidspace of the hosting rock body, but more precisely in its connected void space.

$$GS(gs) \rightarrow RPF(gs) \wedge \exists o[NAPO(o) \wedge hosts(o, gs)]$$
$$GroundWB(wb) \rightarrow WB(wb) \wedge \exists rb, gs[RB(rb) \wedge hosts(rb, gs) \wedge GS(gs) \wedge$$
$$r(wb) \subseteq voidspace(rb) \wedge \forall v[hol_\text{e}(rb, v) \rightarrow \neg PO(wb, v)]]$$
$$SurfaceWB(wb) \rightarrow WB(wb) \wedge \exists gs[hol_\text{e}(wb, gs) \wedge GS(gs)] \qquad \text{(AQ3)}$$
$$HydroRockBody(aq) \rightarrow NAPO(aq) \wedge \exists rb, wb\big[r(aq) = r(rb) + r(wb) \wedge$$
$$RB(rb) \wedge GroundWB(wb) \wedge r(wb) \subseteq con\text{-}voidspace(rb)\big]$$
$$HydroRockBody(aq) \leftarrow Aquifer(aq) \vee Aquitard(aq) \vee Aquiclude(aq)$$

We can further define saturated and unsaturated HydroRockBodies by changing the last term in the HydroRockBody condition from $r(wb) \subseteq con\text{-}porespace(aq)$ to either $r(wb) = \ldots$ (saturated) or $r(wb) \subset \ldots$ (unsaturated). Specific kinds of HydroRock-Bodies are aquifers, aquitards, and aquicludes. While aquifers have a high percentage of permeable void space, i.e. connected void space with openings large enough to permit water flow, aquitards and aquicludes have little or no permeable void space, respectively. However, within our framework we cannot express percentages of permeable void space.

7. Conclusion

Voids are extremely important to water science, but they are complex things. They can be as large as caves, or as minute as the gaps between rock grains, and they can be connected or disconnected; significantly, their particular configuration can impede or enhance the storage of water, as well as its flow. In this paper we generalize the notion of hole from [6] to voids, add gaps as spaces between objects, and provide a physical characterization that delineates voids in macroscopic objects from those occurring microscopically in

an object's matter. The developed axiomatization in first-order logic specializes and extends the DOLCE foundational ontology, and is applied to the domain of hydrogeology, providing a basis for a foundational hydro ontology.

A known limitation of the work is the narrowness of our direct constitution relation, which limits constitution to one step in physical scale, between an object and its matter. While voids at finer physical scales are of secondary importance to the storage and flow of water, they should nonetheless be included for completeness. Another outstanding question is the identification of physical voids, an issue that seems difficult to completely resolve. Further work to advance hydro ontology is in progress, including containment notions to allow water to fill voids, perdurant notions to permit water to flow through voids, the definition of key water entities such as lakes and rivers using these notions, and the transfer of them to DOLCE.

8. Acknowledgements

References

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

[2] L. Bermudez and M. Piasecki. Hydrologic Ontology for the Web. *American Geophysical Union Fall Meeting*, abstract U41B-0016, 2003.

[3] E. Boisvert and B. Brodaric. GroundWater Markup Language (GWML) – Enabling Groundwater Data Interoperability in Spatial Data Infrastructures. *Journal of Hydroinformatics*, 14(1):93–107, 2012.

[4] S. Borgo, N. Guarino, and C. Masolo. A Pointless Theory of Space Based on Strong Connection and Congruence. In *Principles of Knowledge Representation and Reasoning (KR-96)*, 220–229, 1996.

[5] B. Brodaric and F. Probst. Enabling Cross-Disciplinary E-Science by Integrating Geoscience Ontologies with DOLCE. *IEEE Intelligent Systems*, 24(1):66-77, 2009.

[6] R. Casati and A. Varzi. *Holes and other Superficialities*. MIT Press, 1994.

[7] A. Cohn, B. Bennett, J. Gooday, and N. Gotts. RCC: A Calculus for Region Based Qualitative Spatial Reasoning. *GeoInformatica*, 1:275–316, 1997.

[8] M. Donnelly. Layered Mereotopology. In *Int. Joint Conf. on Artif. Intell. (IJCAI-03)*, 1269–1274, 2003.

[9] M. Donnelly. Containment Relations in Anatomical Ontologies. In *American Medical Informatics Association Symp. (AMIA 2005)*, 206–210, 2005.

[10] M. Donnelly. Relative Places. *Applied Ontology*, 1:55–75, 2005.

[11] T. Hahmann and M. Grüninger. Multidimensional Mereotopology with Betweenness. In *Int. Joint Conf. on Artif. Intell. (IJCAI-11)*, 906–911, 2011.

[12] T. Hahmann and M. Grüninger. A Naïve Theory of Dimension for Qualitative Spatial Relations. In *Symp. on Logical Formalizations of Commonsense Reasoning at AAAI-SS'11*. AAAI Press, 2011.

[13] T. Hahmann and M. Grüninger. Detecting Physical Defects: A Practical 2D-study of Cracks and Holes. In *AAAI-SS'09: Benchmarking of Qualitative Spatial and Temp. Reasoning Systems*. AAAI Press, 2009.

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

[15] INSPIRE Thematic Working Group Geology. D2.8.II.4 INSPIRE Data Specification on Geology – Draft Guidelines. Tech. report, INSPIRE, 2011.

[16] I. Masser. *Building European Spatial Data Infrastructures*, 2nd ed., ESRI Press, 2010.

[17] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, and A. Oltramari. Wonderweb deliverable D18 - ontology library (final report). Tech. report, ISTC-CNR, Trento, 2003.

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

[19] A. Tripathi and H. Babaie. Developing a Modular Hydrogeology Ontology by Extending the SWEET Upper-Level Ontologies. *Computers & Geoscience*, 34(9):102–1033, 2008.