Constructing Geo-Ontologies by Reification of Observation Data

Benjamin Adams
Department of Computer Science
University of California, Santa Barbara
Santa Barbara, CA, USA 93106-5110
badams@cs.ucsb.edu

Krzysztof Janowicz
Department of Geography
University of California, Santa Barbara
Santa Barbara, CA, USA 93106-4060
jano@geog.ucsb.edu

ABSTRACT

The semantic integration of heterogeneous, spatiotemporal information is a major challenge for achieving the vision of a multi-thematic and multi-perspective Digital Earth. The Semantic Web technology stack has been proposed to address the integration problem by knowledge representation languages and reasoning. However approaches such as the Web Ontology Languages (OWL) were developed with decidability in mind. They do not integrate well with established modeling paradigms in the geosciences that are dominated by numerical and geometric methods. Additionally, work on the Semantic Web is mostly feature-centric and a field-based view is difficult to integrate. A layer specifying the transition from observation data to classes and relations is missing. In this work we combine OWL with geometric and topological language constructs based on similarity spaces. Our approach provides three main benefits. First, class constructors can be built from a larger palette of mathematical operations based on vector algebra. Second, it affords the representation of prototype-based classes. Third, it facilitates the representation of classes derived from machine learning classifiers that utilize a multi-dimensional feature space. Instead of following a *one-size-fits-all* approach, our work allows one to derive contextualized OWL ontologies by *reification* of observation data.

Categories and Subject Descriptors

I.2.4 [**Artificial Intelligence**]: Knowledge Representation Formalisms and Methods

General Terms

Theory, Design, Human Factors

Keywords

Ontology, Context, Conceptual Spaces, Observations, Provenance

1. INTRODUCTION

In the past, description logics-based knowledge representation languages have been often misunderstood as a replacement to numerical and statistical modeling. Instead, ontologies are best understood as a communication and exchange layer. While the geosciences rely on observations, models, and simulations to answer complex scientific questions such as the impact of global change, the majority of results are communicated and exchanged using categories. For instance, we study deforestation, mountain microclimate, or rural economics by referring to *forests*, *mountains*, or *towns*. However, the definition of these terms differs between domains and even within information communities. Meta-studies have shown that data combined on the category level is, in fact, often incommensurable due to the hidden ambiguities in the definition of the used categories. For instance, research addressing the question of how *green* different sources of renewable energy are, has used different measures, criteria, and weights to define *greenness*; see also [35] for a public policy perspective. Whether the destruction of rare ecosystems by large solar power plants reduces the greenness of solar energy or not differs between studies. Even more, how do we arrive at an objective quantification for the value of a specific ecosystem to determine weights for our measures? By restricting the interpretation of the used terms towards their intended meaning [25, 29], *ontologies* can assist in uncovering such incompatibilities and foster semantic interoperability.

The ability to formally describe and infer the semantics of terms, their combinations, and their relationships using a knowledge representation language is a fundamental component of the Semantic Web. Description logics, which form the basis for the Web Ontology Language (OWL), presuppose the representation of a class as a set of individuals that share certain properties; i.e., it considers the necessary and sufficient features required for membership in a class. However, as argued by Sheth at al. [46], limiting ontologies by restricting them to a description logics-based formalization will artificially limit their adaptation and suitability to model complex domain knowledge as described above. Consequently, various extensions to knowledge representation languages such as OWL have been proposed within the last years. For instance, Bouquet et al. [6] proposed C-OWL to define contextualized, local ontologies and connect them via *bridge rules*, while Stoilos et al. [49] propose various fuzzy extensions to OWL. A review of approaches to handle uncertainty and vagueness on the Semantic Web was recently published by Lukasiewicz and Straccia [31].

Research from cognitive science shows ample evidence of categories with graded membership based on the *similarity* of entities to one or multiple prototypes [42, 23]. One result of this finding is that similarity-based classification has become a common approach in machine learning algorithms [37]. Further, Gärdenfors [18] has argued that the standard set-theoretic concept combination operations central to OWL do not fully capture the way how humans

combine the meanings of terms in natural languages. He proposed the use of a geometric and topological framework called *conceptual spaces* that allows to define semantic operations, such as concept combination, as mathematical operations on the points and regions in a semantic space. Central to this work is that similarity measurement, encoded as a weighted distance function in this space, is a fundamental component to the learning of new properties and complex concepts [19]. Recently, Raubal and Adams proposed to use Gärdenfors' notion of conceptual spaces to personalize search on the Semantic Web [40].

Measuring semantic similarity between concepts or individuals defined by an ontology has been successfully applied in GIScience and also the Semantic Web within the last years. Application areas range from semantic interoperability and ontology matching to understanding uncertainty in geographic feature types and land use classes to semantics-enabled geographic information retrieval [15, 39, 3, 10, 28]. However, so far, similarity has been applied as a measurement over pre-designed ontologies, i.e. as *a posteriori* measures on the representations. For instance, common techniques involve measuring similarity as function of edge distance in subsumption hierarchies. As a result, similarity reasoning is not used to construct ontologies but to quantify differences in or between existing ontologies.

While geometric representations, e.g., based on conceptual spaces, offer an intuitive approach to semantic similarity as analogy to spatial distance, they often lack clear means for combining the concepts they represented and making inferences on those combinations. For example, it is not clear how one might take a shape descriptor for a prototypical leaf image and combine it with a shape descriptor for round objects to infer a representation of a round leaf class. In contrast, description logics do provide mechanisms for inferring the semantics of concept combinations but rely on the set-theoretic intersection, union, and negation operations, which are insufficient for handling many kinds of concept combinations that humans perform, such as in the previous example. One reason is that combination is not symmetric; there is a modifier concept and a modified concept. And more important the semantics of the modifier changes depending on the concept being modified. For example, the semantics of round are quite different when attached to leaf as opposed to ball. Gärdenfors argues that the geometric approach gives us better mechanisms for reasoning over these kinds of concept combinations [17].

Based on Couclelis' framework of *semantic contraction* [9], we propose to use conceptual spaces with similarity-based class constructors as meta-language to create contextualized OWL ontologies. To do so, we present a hybrid, dual layered approach that combines OWL 2 with a logical formalism for conceptual spaces called OWL 2 $\mathcal{CS}$. To combine these formalisms we define a function, which we call a *reification function* [27, 43], that maps from the domains of points, regions, and dimensions within a conceptual space to the range of instances, properties, and classes in an OWL 2 ontology. Following our approach, concepts can be represented as geometric regions and points in a structured set of multidimensional metric spaces [17, 2]. In contrast to previous work [6, 49, 31], we do not argue for more expressive extensions to OWL, but propose an architectural solution. We argue that geometric approaches are most suitable on levels 2 (*observables*) and 3 (*similarities*) of Couclelis' framework, i.e., on the level of *fields* – before objects are introduced [24]. Contextualized ontologies can be derived from conceptual spaces based on reification; cp. level 4 and

5 of the semantic contraction framework. As these ontologies are specified in OWL, they are decidable and compatible with the vast spectrum of Semantic Web technologies. The conceptual space-based meta-level acts as provenance information [51] and fosters interoperation by semantic translations; see Figure 1. Just recently, the report from the Climate Knowledge Discovery Workshop 2011 underpins the need for a provenance-enabled framework that helps to integrate numerical models with ontologies [7].
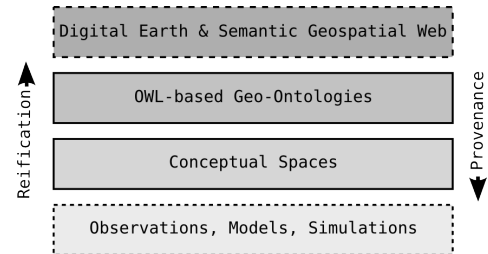

Figure 1: Observation-driven ontology engineering: Observations and numerical models act as the basis for representing knowledge by conceptual spaces. Using a reification function and a threshold-based similarity class connective, these representations are mapped to contextualized OWL ontologies for use on the Semantic Geospatial Web [14]. To understand how these ontologies were derived, the classes can be re-reified to conceptual spaces again.

The remainder of the paper is structured as follows. First, we introduce preliminaries required for the understanding of our work and highlight related research. Next, in section 3, we describe a formal theory of conceptual spaces. Section 4 specifies the integration of our work with OWL. In section 5, we discuss the creation of contextualized OWL ontologies by reification of regions in conceptual spaces. Finally, we conclude our work and highlight directions of further research.

2. PRELIMINARIES
In this section we provide preliminary information on the Web Ontology Language, conceptual spaces, and related works.

2.1 OWL 2
OWL 2 is a Semantic Web language to represent domain knowledge in the form of ontologies specified by axioms, entities, and expressions. Axioms are propositions that state certain facts about the world are true, and can be used to deduce other facts through logical implication. Entities are the atomic components of these axioms and consist of individuals, classes, and properties. Individuals represent objects in the world (though they can be abstract non-physical objects), classes represent sets of objects, and properties, called roles, represent binary relations. A distinction is made between object properties, which represent a relation between an object and another object, and datatype properties, which relate a data value to an object. These entities can be combined using logical connectives, such as conjunction and disjunction, into expressions. The type of class and property connectives available for use in a particular OWL profile has a direct impact on the decidability and complexity of this language.

The model theoretic semantics of OWL 2 are very closely based on the semantics of the $\mathcal{SROIQ}$ description logic with the addition of concrete datatypes and punning where individuals, classes, and properties have a direct correspondence to instances, concepts,

and roles, respectively, in $\mathcal{SROIQ}$ [38]. The *signature* or *vocabulary* of an OWL 2 ontology is defined as a 7-tuple $V = (V_C, V_{OP}, V_{DP}, V_I, V_{DT}, V_{LT}, V_{FA})$ where V_C is a set of classes, V_{OP} is a set of object properties, V_{DP} is a set of data properties, V_I is a set of individuals, V_{DT} is a set of datatypes, V_{LT} is a set of literals (i.e., data values), and V_{FA} is a set of constraining facet and value pairs. This vocabulary is defined over a datatype map, which is a set of datatypes that are mostly derived from the XML Schema datatypes, such as xsd:integer and xsd:float. Without going into any details here, the semantics of OWL is given by the interpretation function $.^{\mathcal{I}}$ that maps expressions to sets, e.g., $C \sqcap D = C^{\mathcal{I}} \cap D^{\mathcal{I}}$ for the conjunction of classes; see OWL 2 Overview for more information: `http://www.w3.org/TR/owl2-overview/`.

There exist a number of different syntaxes for OWL that are equally expressible, although all OWL reasoners are required to support the RDF/XML serialization of the language. We use the functional-style syntax to describe OWL 2 $\mathcal{CS}$, simply because it is the most efficacious in terms of both readability and space constraints. A future work will be to fully specify an RDF/XML serialization for OWL 2 $\mathcal{CS}$ by combining OWL 2 with a suitably extended RDF/XML version of the Conceptual Space Markup Language (CSML) defined in previous work [1].

2.2 Conceptual Spaces

Conceptual spaces are a geometric and topological framework for knowledge representation proposed by Gärdenfors [17]. A conceptual space is a mathematically defined space composed of quality dimensions, which is further structured by the organization of the dimensions into disjoint groups called domains. An example of a domain is *color* with three quality dimensions *hue*, *value*, and *saturation*. This organization is exploited for semantic reasoning such as concept combination. Quality dimensions within a domain are integral dimensions that are separable from dimensions in other domains. This notion of integral dimensions corresponds to the use of the term in psychological research whereby an object cannot have a value for a dimension without also having a value for the dimensions with which it is integral [20]. For example, the color of an object cannot be described fully without having a value for all three dimensions: *hue*, *value*, and *saturation*.

In conceptual spaces, a property is defined as a region in one domain, e.g., the property *red* is represented by a region in the color domain. Concepts are represented as sets of regions (i.e., properties) along with information indicating how the properties are correlated. Objects in a conceptual space are represented by a set of quality values, $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$. Gärdenfors and Williams [19] showed how to reason about categories represented in a conceptual space using the region connection calculus. Furthermore, they illustrated that conceptual spaces can be used to model context-sensitive, non-monotonic categorization. Adams and Raubal [2] presented a formal algebra that describes operations for concept combination on concepts and properties represented in a such a space. Techniques for learning regions include Voronoi tessellation, nearest neighbor, and clustering algorithms [11, 37].

The terms *instance*, *concept*, *property*, and *domain* have different meanings in OWL and conceptual spaces. We will use the convention of prefixing those terms with $\mathcal{CS}$ when referring to conceptual spaces. Without appellation those terms should be assumed to have their standard meaning in description logics. For example, a "$\mathcal{CS}$ property" is a property represented as a region in a conceptual space domain whereas a "property" is a property in an OWL ontology.

2.3 Related works

Sheremet et al. [45] have developed a logic for concepts and similarity that combines the $\mathcal{ALCQIO}$ description logic with a similarity logic SL. The work presented here is different in that it imposes a conceptual space structure upon the similarity spaces. This structure has the benefit that it has formal operations for concept combination, and that convexity of regions in the space plays a principal role in the interpretation of its semantics. The formulation of conceptual space theory described below expands on previous work on a conceptual space algebra [2]. It also draws from previous work on the development of hybrid logics that contain both description logic and metric space logic components [30], although the $\mathcal{CS}$ logic has a more complex structure than their MS logic. For an examination of the relation between conceptual spaces and frames and how one might recover a frame model from a conceptual space model see [52]. Dietze et al. have proposed a conceptual space layer to aid context aware discovery of Semantic Web services [12]. In our work, the primarily concern is the use of conceptual representations to facilitate class construction and learning. The approach presented here is akin to Goguen's [22] Unified Concept Theory for combining symbolic and geometric representations through relations, though our approach is not based on category theory.

3. CONCEPTUAL SPACE THEORY

In this section we define the elements of a conceptual space and present a motivating example. The theory builds upon the metric conceptual space algebra [2] but uses a different symbology to better align with description logics. The section also describes a prototype as a relation between a concept and an instance, and introduces a formal representation of correlations between dimensions.

3.1 Motivating example

Mountain is a geographic feature type that figures prominently in people's conceptualizations of geographic space [34]. However, although *Mountain* is an archetypal type there is no universal, formal definition for mountain [48]. In fact, mountains are objects without clearly defined boundaries and the necessary and sufficient features for membership in the mountain class are highly location-dependent. This *localness* of the mountain concept is illustrated by the map of the world's "principal" mountains and rivers shown in Figure 2 [36]. In this map from 1846 the relative heights of mountains and lengths of rivers are shown schematically. Implicitly, the definition of a "principal" is based on height but also varies greatly by region. Mountains from England are illustrated even though all of those principal mountains are significantly shorter than many Asian mountains that are not shown. These differences stem not only from cultural differences but also the inherent context-dependence of geographic concepts. Traditionally, GIScientists have avoided representations that operate on mountain objects in lieu of operating on field based elevation data and using related methods. However, to realize a Digital Earth that interfaces with humans, it is imperative to create ontologies that define semantics of terms such as *Mountain* using links to observation data. To illustrate the elements of the $\mathcal{CS}$ theory, we discuss how observation data about mountains can be used to create different, context-dependent *Mountain* classes for geo-ontologies.

3.2 Definitions

Definition 1. $\mathcal{CS}/\mathcal{D}$. A $\mathcal{CS}$ domain ($\mathcal{CS}/\mathcal{D}$) is a bounded metric space, $\langle M, d \rangle$, where M is a set of points and d is a distance function (1).

$$d : M \times M \to \{0\} \cup \mathbb{R}_+. \qquad (1)$$

Figure 2: Heights of the *principal mountains* in the world; see *http://bit.ly/mecWLs* for a high resolution image.

A set of named quality dimensions ($\mathcal{CS}/\mathcal{Q}$) form the basis of a $\mathcal{CS}/\mathcal{D}$.

Example 1. For mountain features, examples of observable quality dimensions include *altitude*, *relative relief*, *slope steepness*, *prominence*, *topographic isolation*, and *land volume* [21]. These values are recoverable for any location from a digital elevation model or other type of field-based representation. See Figure 3 for an illustration of a domain. Further domains may be defined based on vegetation, climate information, shape, and so forth.
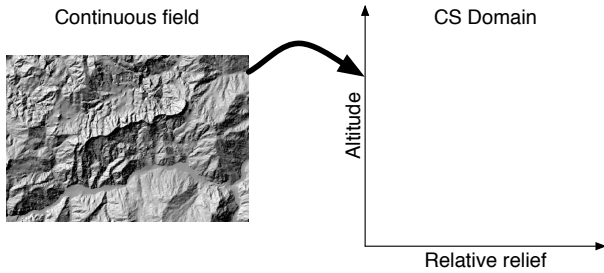

Figure 3: A sample domain for representing properties of peaks drawn from continuous field.

A context-based semantic distance function on a $\mathcal{CS}/\mathcal{D}$ composed of continuous $\mathcal{CS}/\mathcal{Q}$s is defined as a weighted Minkowski distance metric (2).

$$d(\mathbf{x}, \mathbf{y}, \mathbf{w}) = \left(\sum_{i=1}^{n} w_i (x_i - y_i)^p \right)^{1/p} \qquad (2)$$

where $\mathbf{x}$ and $\mathbf{y}$ are vectors of quality dimension values for an n dimensional domain and $\mathbf{w}$ is a vector of weights on the dimensions. The order of the distance, p, will typically be 2 (Euclidean) or 1

(Manhattan). Similarity and semantic distance are dual concepts, and a measure of similarity in a conceptual space is an (often exponentially) decaying function of distance [44]. The choice of an appropriate distance metric and its corresponding similarity function is application dependent. Given that a $\mathcal{CS}/\mathcal{D}$ is metric it satisfies the axioms of identity, symmetry, and triangle inequality. However, in accordance with the literature, the latter two are relaxed by the introduction of saliency weights on the dimensions.

It follows from this definition that domains and quality dimensions present a means for ordering property values. Although conceptual spaces allow for nominal quality dimensions, due to our hybrid approach there is no benefit of having dimensions without an ordering (as this reduces the expressiveness to set theory-based representations). From an engineering perspective it often makes sense to define the quality dimensions as normalized dimensions, from [0,1] or using a z-score [39]. This normalization has the benefit of making the semantic distance measure scale independent. For simplicity, we will assume a [0,1] normalization for every dimension, though this is not a requirement.

Definition 2. $\mathcal{CS}/\mathcal{P}$. A $\mathcal{CS}$ property ($\mathcal{CS}/\mathcal{P}$) is a region in a $\mathcal{CS}/\mathcal{D}$, where a region is defined as a star-convex set of points in the space.

For semantic reasoning we place a special emphasis on the subset of CS properties that satisfy a given *betweenness* relationship for all points $i \in \mathcal{CS}/\mathcal{P}$, which are denoted as *natural* CS properties ($\mathcal{CS}/\mathcal{NP}$). A strong definition of betweenness defines a $\mathcal{CS}/\mathcal{NP}$ as a convex region in $\mathcal{CS}/\mathcal{D}$, such that for all i and $j \in \mathcal{CS}/\mathcal{NP}$ and all $x \in [0, 1]$ the point $(1 - x)i + xj \in \mathcal{CS}/\mathcal{NP}$. A $\mathcal{CS}/\mathcal{P}$ that is not a $\mathcal{CS}/\mathcal{NP}$ is termed a *refined* property, $\mathcal{CS}/\mathcal{RP}$.

Let p and q be $\mathcal{CS}/\mathcal{NP}$s. Following the definition of convexity, $p \cap q$ is also a $\mathcal{CS}/\mathcal{NP}$. If $p \subseteq q$ then p is a subProperty of q. For all $\mathcal{CS}/\mathcal{RP}$ there exists a set of $\mathcal{CS}/\mathcal{NP}$s, N, such that $\bigcup_{i \in N} i = \mathcal{CS}/\mathcal{RP}$. Computationally efficient approximations of $\mathcal{CS}/\mathcal{NP}$ are hypersphere or polytope representations. A hypersphere $\mathcal{CS}/\mathcal{NP}$ can be constructed from a point and distance pair. Polytope representations can be constructed from convex hull, Voronoi tessellation, and other numerical classification algorithms.

Example 2. Compositional properties such as *tall* and *high* are represented as regions in a domain as well, though they entail different inferences from non-compositional properties when combined with other classes; see next section for more explanation. Classes defined by such properties are known as contrast classes [17]. Figure 4 shows how one can represent *tall* and high as regions in the domain from the previous example. Note that in this representation *high* depends solely on *altitude* while *tall* depends on *relative relief*. It should be stressed that these definitions are simplifications, and just like in any forms of ontology development the choice of property definitions in a conceptual space is application dependent. More complicated properties such as *visually impressive* can be defined as regions based on a functional such as the Omnidirectional Relief and Steepness (ORS) measure [13].

Definition 3. $\mathcal{CS}/\mathcal{I}$. A $\mathcal{CS}$ instance ($\mathcal{CS}/\mathcal{I}$) is a set of points $\{i_1, i_2, \ldots, i_n\}$ such that every point is in a distinct $\mathcal{CS}/\mathcal{D}$.

From Definitions 2 and 3 we see that a $\mathcal{CS}/\mathcal{P}$ consisting of one point is semantically equivalent to a $\mathcal{CS}/\mathcal{I}$ in one domain. Following [2] the distance between two instances (Equation (3)) can
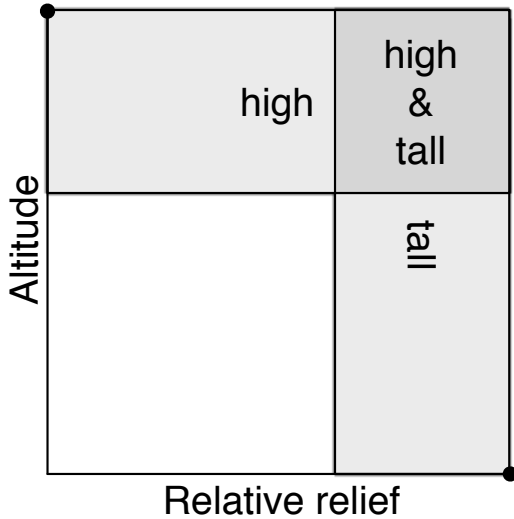
Figure 4: Sample properties.

be defined as a weighted sum of the intradomain point distances calculated using Equation (2).

$$d(\mathcal{CS}/\mathcal{I}_1, \mathcal{CS}/\mathcal{I}_2) = \sum_{i=1}^{D} k_i d(\mathbf{x_i}, \mathbf{y_i}, \mathbf{w_i}),\qquad(3)$$

where D is the number of domains shared by the instances and k_i is a salience weight on the ith domain. Imposing weights on the domains allows one to emphasize some properties over others depending on the context.

Example 3. Individual mountain objects are represented as sets of points based on observation data; see Figure 5 for an example of points in one domain. For a given mountain one can represent other properties in additional domains as described above.
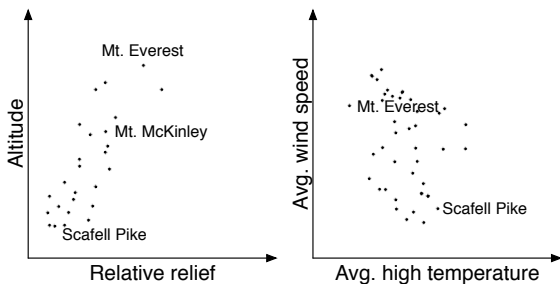

Figure 5: Sample instances.

Definition 4. $\mathcal{CS}/\mathcal{C}$. A $\mathcal{CS}$ concept ($\mathcal{CS}/\mathcal{C}$) is a tuple $\langle P, Cor \rangle$ where P is a finite set of $\mathcal{CS}$ properties $\{p_1, p_2, \ldots, p_m\}$, and Cor is a cross covariance matrix. Every property in P is a region in a distinct $\mathcal{CS}/\mathcal{D}$. Cor describes the pairwise correlations between all the quality dimensions in the set of $\mathcal{CS}/\mathcal{D}$s covered by P.

A concept is a generalization of a property extended across multiple domains. An instance I is an instanceOf a concept if for all points i in I there exists a property p in P, such that i is contained

within the region of p. Likewise, one concept has an isA relationship to another concept if the former's regions are all contained within the latter. The conditions for instanceOf and isA are not necessary and sufficient, because subsumption reasoning for concepts and instances can be more complicated in a conceptual space. For example, concept combination can be non-monotonic; see below. Therefore, the subsumption reasoning must be extended with the notion of a similarity-based subsumption as well.

Intuitively, this means that if we repeatedly combine a concept with a modifier the new concepts will change from being sub-concepts of the original concept to not being subsumed. For example, if we modify the *lake* concept with the property *salinity* it might still be a lake, but if we modify it again to be the size of an ocean then it will no longer be a lake. Because dimensions are weighted for similarity measurement, combinations that involve changing quality values of highly salient properties will more likely make the new concept not subsumed by the original. In the previous example, size is more salient than salinity.

Example 4. The concept of a mountain is different depending on where you are in the world. The tallest mountains in England would scarcely be foothills in Himalayas. In a conceptual space a mountain concept (i.e., its set of necessary properties) is defined based on observation data and different data (e.g., from different locales) will produce different representations. For example, by finding the convex hull of mountain instances in England we can represent the local mountain concept for England and likewise for Asia; see Figure 6. This approach is related to the methodology of Bennett et al [5].
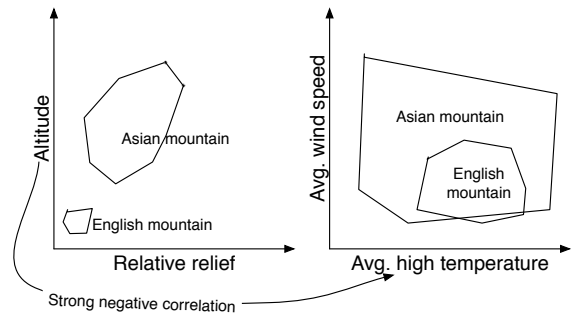

Figure 6: Sample mountain concepts.

Definition 5. **Prototype**. The prototype relation is a binary relation on the set of concepts and instances, $\langle C \times I \rangle$ with the restriction that all points in I must be elements of a property of C.

It may be convenient to define the prototype as the centroids of the property regions, but other points may be appropriate as well. For example, a good representation of a prototypical mountain may not be the "average" mountain based on a set of observations. Similar to trees, a prototypical mountain corresponds to an extreme value along an "ideal" quality dimension such as *altitude* [32].

Definition 6. **instanceOf**. Let I be a $\mathcal{CS}/\mathcal{I}$ and C be a $\mathcal{CS}/\mathcal{C}$. I is an instanceOf C iff it meets the above criteria in Definition 4 OR the weighted distance between I and the prototype of C is less than or equal to α, a scalar similarity threshold.

Definition 7. **isA**. Let C_1 and C_2 be $\mathcal{CS}/\mathcal{C}$s. C_1 isA C_2 iff it meets the above criteria in Definition 4 OR the distance between C_1 and C_2 is less than or equal to β, a scalar similarity threshold.

The simplest definition of distance between two concepts is the weighted distance between the prototypes of the two concepts. An alternative measure is the weighted sum of Hausdorff distances between properties. The Hausdorff distance is a measure of how close two sets are in a metric space and has found wide use in many applications such as image similarity [26]. Let X and Y be two sets it is defined in Equation (4):

$$d_{\mathrm{H}}(X,Y) = \max\{ \sup_{x \in X} \inf_{y \in Y} d(x,y), \sup_{y \in Y} \inf_{x \in X} d(x,y) \} \quad (4)$$

If X and Y are convex sets the Hausdorff distance is equal to the Hausdorff distance of the boundaries of X and Y [50]. Let $P = \{p_1, p_2, \ldots, p_n\}$ and $Q = \{q_1, q_2, \ldots, q_n\}$ be sets of $\mathcal{CS}/\mathcal{P}$ for two concepts described in n domains. The distance between P and Q is shown in Equation (5): sections

$$\sum_{i}^{n} w_i d_{\mathrm{H}}(p_i, q_i) \quad (5)$$

This assumes that P and Q are defined for the same domains.

As previously discussed, since instanceOf and isA relations are defined in part by a weighted similarity measurement, the determination of whether or not a subsumption relationship exists in any given situation is context-dependent.

Definition 8. A $\mathcal{CS}$ knowledge base ($\mathcal{CS}/\mathcal{K}$) is defined as an ordered triple $\langle D, I, C \rangle$, where D is a finite set of $\mathcal{CS}$ domains, I is a finite set of $\mathcal{CS}$ instances, and C is a finite set of $\mathcal{CS}$ concepts.

We will use the short hand term *conceptual space* to mean a $\mathcal{CS}/\mathcal{K}$, though in fact a knowledge base can contain not only multiple spaces (domains), it also specifies the contents of those spaces.

4. INTRODUCING OWL 2 $\mathcal{CS}$

In this section we introduce OWL 2 $\mathcal{CS}$, an approach for combining the components of a $\mathcal{CS}/\mathcal{K}$ with an OWL ontology into a common serialization. In addition to named classes, properties, and instances, an OWL 2 $\mathcal{CS}$ file has named $\mathcal{CS}$ domains, $\mathcal{CS}$ properties, $\mathcal{CS}$ instances, and $\mathcal{CS}$ concepts. The $\mathcal{CS}$ structures can be described using the Conceptual Space Markup Language (CSML), which allows to assert a unique IRI for each element in a conceptual space [1]. An OWL 2 $\mathcal{CS}$ file therefore is an XML serialization with both OWL and CSML elements. Due to space restrictions we will use an equivalent functional-style syntax in place of CSML (see table 1). In the examples shown a region in a domain is specified as a system of linear inequalities using the `LinearInequalities` function, which has the semantics of being an anonymous property. In place of any `LinearInequalities` construct, a named property can also be used.

4.1 Concept constructors

One of the main advantages of the conceptual space theory is that one can define concept constructors that go beyond set intersection and union. The construction of a concept from the combination of two other concepts in a conceptual space takes the form of one of several possible mathematical operations. Adams and Raubal [2] defined three such operations: 1) property-concept combination, 2) concept-concept combination, and 3) contrast class-concept combination. These three operations take as input a modifier in the form

of one or more regions (i.e., a property or concept) and a concept to be modified. Given these inputs it outputs a new concept. For 1) and 3) the resulting concept is only modified in one domain. These combinations correspond roughly to adjective-noun combinations like *long river*. A concept-concept combination will combine concepts across more than one domain with a salience indicator on the domains that indicates in which domains the properties should change. This combination corresponds roughly to noun-noun combinations like *desert town*.

Contrast classes are properties that are combined by taking the shape of the contrast class and projecting onto the property of the modified concept in that domain. Refer to [2] for a detailed mathematical explication on how this and other combination operations work. In the example of *long river*, the *long* property is a region specifying a range along the upper part of the length dimension. See Figure 7 for an illustration for how the semantics of the modifying property, such as *long* or *short*, is inferred differently depending on the concept that it is modifying. In that example *long British river* is more similar to a *short North American river* than *long North American river*. The similarity between these properties is formally measured in terms of the Hausdorff distance as described in the previous section. In fact, when considering only the length dimension *long British river* **isA** *short N. American river*, since it is a subset. However, when including other properties of rivers, such as their locations this is no longer the case. In the following section we show how by specifying a context in terms of salience weights on the dimensions we can generate different **isA** relations.
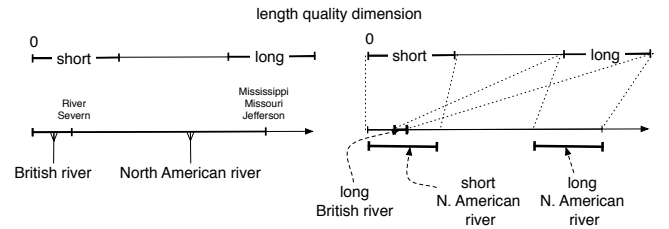

Figure 7: Illustration of contrast class-concept combination (only length dimension shown).

CSML does not as of yet have XML serializations for defining $\mathcal{CS}$ concept constructions, and due to space constraints we present these operations in a functional-style notation for OWL 2 $\mathcal{CS}$. We define the EquivalentConcept constructor for the $\mathcal{CS}$ entities to be an analog to the EquivalentClass constructor in OWL 2:

```
EquivalentConcept(:salineLake
    PropertyConceptCombination(
        :saline :lake ) )

EquivalentConcept(:desertTown
    ConceptConceptCombination(
        :desert (:climate)
        :town ) )

EquivalentConcept(:tallMountain
    ContrastClassConceptCombination(
        :tall :mountain ) )

EquivalentConcept(:longRiver
    ContrastClassConceptCombination(
        :long :river ) )
```

$\mathcal{CS}$ element	Example functional notation
Domain	`Domain(:fromElevationData` `  QualityDimensions( (:altitude 0.0 1.0) (:relativeRelief 0.0 1.0) ) )`
Property	`Property(:tall :fromElevationData` `  (LinearInequalities( (>= :relativeRelief 0.7) (>= :altitude 0.0) ) )`

Instance	`Instance(:scafellPike`
	`  (:fromElevationData (:altitude 0.11) (:relativeRelief 0.14) )`
	`  (:fromClimateData (:avgWindSpeed ... ) ... ) ... )`

Concept	`Concept(:britishMountain`
	`  (:fromElevationData (LinearInequalities ( ... ) ) )`
	`  (:fromClimateData (LinearInequalities( ... ) ) ) ...`
	`  CovarianceMatrix ( [0.07 0.42 -0.13 ... ] [ ... ]  ) )`

Table 1: Functional notation for declaration of $\mathcal{CS}$ elements

In the above examples **:saline**, **:tall**, and **:long** are $\mathcal{CS}/\mathcal{P}$s; **:lake**, **:desert**, **:town**, **:mountain**, and **:river** are $\mathcal{CS}/\mathcal{C}$s; and **:climate** is a $\mathcal{CS}/\mathcal{D}$. The concept-concept combination shown indicates that we want to combine the **:climate** property only of **:desert** with the **:lake** concept. The new concepts **:salineLake**, **:desertTown**, **:tall-Mountain**, and **:longRiver** are all $\mathcal{CS}/\mathcal{C}$s, which means that they will be written in CSML as sets of convex regions represented as systems of linear inequalities. It should be noted that these operations as defined in [2] do not incorporate the correlation information as part of each $\mathcal{CS}/\mathcal{C}$s description.

4.2 Different Approach to Semantics

Conceptual spaces are a framework for knowledge representation that is based in part on the paradigm of cognitive semantics [16]. A key aspect of cognitive semantics is that the meaning of a term is a conceptual structure (in the mind of an agent) and this meaning comes prior to any notion of truth assignments on sentences built with these terms. This approach to semantics is different from model-theoretic semantics, which defines meaning in terms of truth-values in possible worlds [8]. By creating a hybrid representation that builds conceptual spaces on top of description logics, however, we must on some level reconcile these two approaches to semantics. The approach we take is that an OWL reification function is a fixing of the cognitive semantic representation to a possible world where truth-values are assigned. In the case that the conceptual regions in a conceptual space are learned from observations of phenomena in the world, we can say that an OWL reification of the conceptual space is a model-theoretic interpretation of the geometric and topological configuration of the regions learned in the space.

5. REIFICATION

The reification of an OWL 2 $\mathcal{CS}$ ontology to an OWL 2 ontology maps the entities in the $\mathcal{CS}$ layer into strict OWL 2 entities. For a concepts defined by prototypical instances, this translation results in a loss of information because the boundaries of classes must be defined for a particular similarity threshold and the notion of a prototype is lost. In addition, no information about correlations is saved. Reification, therefore, freezes the ontology into a particular contextual stance. The advantage is that it results in an ontology that can be reasoned with by existing semantic reasoners and which

is decidable. Ideally, from an ontology engineering perspective, the relation is maintained between the reified ontology and the original OWL 2 $\mathcal{CS}$, which affords the ability to do decidable reasoning but still access the more fine-grained representation for similarity reasoning and more complex concept combinations.

The following describes a suggested correspondence of OWL entities to elements of a conceptual space model. Note, that alternate reification functions can be defined. Classes correspond to $\mathcal{CS}$ concepts, object properties correspond to $\mathcal{CS}$ quality dimensions and domains, and individuals correspond to $\mathcal{CS}$ instances and $\mathcal{CS}$ properties. The latter may seem counterintuitive; however, a $\mathcal{CS}$ property is best considered as a property *value* whereas the object properties (in OWL terminology) that relate $\mathcal{CS}$ instances to $\mathcal{CS}$ properties are derived from the domain structure. In fact, the compressing of points and regions in a conceptual space into one type (OWL individual) is consequent from the semantic contraction from fields to objects.

5.1 Mapping rules

A $\mathcal{CS}$ to OWL 2 mapping, M, shown in Equation (6) is a function from a $\mathcal{CS}/\mathcal{K}$ (i.e., a set of domains, instances, and concepts and their constituent properties) into an OWL 2 *vocabulary*.

$$M : \mathcal{CS}/\mathcal{K} \to V \qquad (6)$$

In this section we present some basic rules for performing this translation for different circumstances. We use the symbology of $\rho(\cdot)$ to represent the *reification* function that generates OWL axioms from the elements of a conceptual space. A *conceptual space provenance* function is the inverse of the reification function $\rho^-(\cdot)$ mapping OWL entities to a conceptual space elements. Since both the OWL and $\mathcal{CS}$ elements (encoded in CSML) are identified by unique IRIs, a relation can be stored mapping the translated OWL entities back to the appropriate element at the $\mathcal{CS}$ layer. In the translated OWL ontology this is an defined as an object property, **:hasConceptualSpaceProvenance** that is an inverse of the **:hasReification** property. This way semantic similarity reasoners and other tools can be developed that utilize both the $\mathcal{CS}$ and OWL layers as needed.

$\mathcal{CS}/\mathcal{D} \to$ *set of transitive object properties*

A domain is a set of quality dimensions and each dimension in a domain is translated into a pair of inverse object properties that are transitive, **:hasMore*Dimension*** and **:hasLess*Dimension***. For example, the *length* dimension generates **:hasMoreLength** and **:hasLessLength** object properties. Let Q_{CS} be a CS/Q, $\rho(D_{CS}) = p \in V_{OP}$. In addition, the domain is defined as an **AsymmetricObjectProperty(:has*Domain*Value)**. Let D_{CS} be a CS/D, $\rho(D_{CS}) = p \in V_{OP}$. This property is a relation between individuals created from data observation instances and individuals created from named CS/Ps.

$CS/P \rightarrow individual$

When mapping an OWL 2 CS hybrid model into an OWL 2 ontology, CS regions are translated into individuals. Let P_{CS} be a CS/P, $\rho(P_{CS}) = i \in V_I$. Since there are infinitely many possible regions in a conceptual space only named property regions and the unnamed regions defined as parts of CS concepts are actually translated. For each named $\rho(P_{CS})$ a new class is also defined as

EquivalentClass(:*Property*Objects
ObjectSomeValuesFrom(:has*Domain*Value $\rho(P_{CS})$)).

For all $P'_{CS} \subseteq P_{CS}$:

subClassOf(:*Property*'Objects :*Property*Objects).

$CS/I \rightarrow object\ property\ assertions$

Let I_{CS} be a CS/I, $\rho(I_{CS}) = i \in V_I$. For every point defined as part of a CS/I a set of object property assertions is made. Two different approaches can be taken for these assertions depending on the application: a similarity-based approach or an inclusion-based approach. The similarity-based approach is taken when the properties are defined primarily by their prototype instance and is context dependent based on weights, w, on the dimensions.

Algorithm 1 Similarity-based instance to object property assertion

  for all points $\in CS/I$ **do**
    for all CS/P in domain of point **do**
      if d(point, prototype of property, w) $< a$ **then**
        ObjectPropertyAssertion(:has*Domain*Value $\rho(I_{CS})$
        $\rho(P_{CS}))$
      end if
    end for
  end for

Rather than doing a test based on distance, the inclusion-based approach tests if a point falls within a CS property region. The inclusion-based approach is ultimately based on similarity because the regions are constructed geometrically utilizing a distance metric, but the thresholds are predefined in the CS layer rather than during the reification stage. As discussed in [2] if a property region is defined as a system of n linear inequalities then this point inclusion test is $O(n)$.

$CS/C \rightarrow class$

Let C_{CS} be a CS/C, $\rho(C_{CS}) = c \in V_C$. For any mapping from a CS/C to an OWL class, the set of salient domains must be specified. For example, while there can be different types of observation measurements associated with a mountain concept that are represented in separate domains within a conceptual space, only a subset of the CS properties will be necessary properties for a given reification. The new class reification can be defined either in terms of

property restrictions on properties reified from domains (as above) or as a closed class that is an enumeration of all the $\rho(I_{CS})$ individuals reified from CS instances that satisfy the **instanceOf** relation described in Section 3.

The new class defined as a property restriction on this set of CS properties takes the following form:

EquivalentClasses($\rho(C_{CS})$ ObjectIntersectionOf(
ObjectSomeValuesFrom(:has*Domain1*Value $\rho(P_{CS})$)
ObjectSomeValuesFrom(:has*Domain2*Value $\rho(P'_{CS})$) ...))

When reified as an enumeration of individuals the new class is defined based on a context-dependent weighted similarity measure. **EquivalentClasses($\rho(C_{CS})$ ObjectOneOf($\rho(I_{CS_1})$ $\rho(I_{CS_2})$ ...))**, where d(I_{CS_i}, C_{CS}, w) $< a$. Similarly, the similarity-based **isA** relation described in Section 3 can be directly translated into **subClassOf($\rho(C'_{CS})$ $\rho(C_{CS})$)**.

Same individual

In a conceptual space the notion of identity is supervened on similarity defined as a function of distance. Two instances that are defined by the same points are considered as identical instances because their semantic distance is 0. Given that there might be errors in the observation measurements one can define a rule for mapping instances to a **SameIndividual($\rho(i_1)$ $\rho(i_2)$)** axiom if $d(i_1, i_2, w) < \alpha$, where α is a fixed constant error threshold.

Combining reified classes with OWL-defined assertions

Using the mountain example from previous sections one can see how a *BritishMountain* CS concept can be reified into a **:BritishMountain** class using a property restriction derived from the height dimension. However, British mountains also have the property restriction of being on the island of Great Britain, which can be easily represented using a object property restriction in OWL. The **:BritishMountain** class therefore is described as an intersection of the reified class and additional restrictions. When calculating the semantic similarity of the **:BritishMountain** class to another representation of mountain, not only can the semantic similarity reasoner rely on feature-based similarity using subsumption relationships but also for those properties with a conceptual space provenance it can use a more refined similarity measure based on distances within the conceptual space.

6. CONCLUSION

In this paper, we propose a methodology to derive context-specific geo-ontologies by reifying geometric representations; i.e., points and regions in a conceptual space. In contrast to related work, we do not argue for the introduction of more expressive description logics which comes at the price of increased worst-case complexity or loss of decidability [4], but for an architectural solution. Following Sheth et al. [46], knowledge representation on the Semantic (Geospatial) Web should not be restricted to description logics only. Instead, we propose conceptual spaces as an additional layer during ontology engineering that is closer to numerical models used for classification in the geosciences. Depending on the application area, these geometric representations are reified using a similarity-based class constructor and context dependent thresholds to plain OWL classes. Scholars interested in exchanging data on the category level can track down how these classes were constructed and uncover incompatibilities that are not visible on the class level.

This approach has a number of benefits. First, it affords not only the standard Semantic Web connectives such as intersection, union, and complement but also more complex combinations based on geometric and topological operations. Second, the geometric representation allows one to describe graded categories that are constructed based on similarity to a prototypical exemplar of the category, which has direct relevance to ranking in semantic search. Third, it creates a direct bridge to the myriad machine learning categorization algorithms based on similarity measures. Following ideas from conceptual space theory, we impose a structure (with respect to convexity, domains, and integral versus separable dimensions) that allows for automated logical inferences about semantics based on the configurations of the regions in the spaces. Finally, the resulting geo-ontologies can be expressed in smaller OWL 2 profiles, e.g., $\mathcal{EL}^{++}$, compared to approaches that propose to move uncertainty and vagueness into the immediate representation language. Successful large-scale ontologies from the life sciences, such as GALEN, demonstrate that in many cases expressivity is not the main concern [41]. Constructing ontologies as reifications of observation data is also crucial for ontology personalization. In context of mobile computing and location-based services, our approach can be extended to construct *ontologies to go*, i.e., automatically calibrate class definitions to the current location of the user.

Besides GIScience, there are several application areas that would benefit from the ability to describe classes in ontologies using a geometric approach. Two examples are multimedia and the Sensor Web. In the first case, classification is often based on low-level features that are closely grounded to perception (e.g., shape, color, and texture). A common method used by the machine learning and computer vision is to represent prototypical exemplars of classes of these types as points in a multi-dimensional feature space [33]. The Semantic Sensor Web aims to represent the semantics of aggregate patterns derived from observations distributed over a geographical space [47]. Often these data take the form of a time series and classification algorithms compare individual observations to a prototypical exemplar using a distance metric. In both cases properties of individuals are more naturally represented as vectors of quality values than symbolic relations.

While we have applied similarity measures and conceptual spaces to various use cases in previous work, the reification framework has to be separately evaluated with domain experts in the near future.

Acknowledgments

7. REFERENCES

[1] B. Adams and M. Raubal. Conceptual space markup language (CSML): Towards the cognitive semantic web. In *International Conference on Semantic Computing (ICSC 2009)*, pages 253–260. IEEE Computer Society, 2009.

[2] B. Adams and M. Raubal. A metric conceptual space algebra. In K. Hornsby, C. Claramunt, M. Denis, and G. Ligozat, editors, *Spatial Information Theory*, volume 5756 of *Lecture Notes in Computer Science*, pages 51–68. Springer Berlin / Heidelberg, 2009.

[3] O. Ahlqvist and A. Shortridge. Characterizing land cover structure with semantic variograms. In A. Riedl, W. Kainz, and G. Elmes, editors, *Progress in Spatial Data Handling – 12th International Symposium on Spatial Data Handling*, pages 401–415. Springer, 2006.

[4] F. Baader, C. Lutz, and A.-Y. Turhan. Small is again beautiful in description logics. *KI*, 24(1):25–33, 2010.

[5] B. Bennett, D. Mallenby, and A. Third. An ontology for grounding vague geographic terms. In *Formal Ontology in Information Systems - Proceedings of the Fifth International Conference (FOIS 2008)*, volume 183, pages 280–293. IOS Press, 2008.

[6] P. Bouquet, F. Giunchiglia, F. Harmelen, L. Serafini, and H. Stuckenschmidt. C-OWL: Contextualizing ontologies. In D. Fensel, K. P. Sycara, and J. Mylopoulos, editors, *The Semantic Web – Proceedings of the Second International Semantic Web Conference (ISWC 2003), October 20-23, Sanibel Island, FL, USA*, volume 2870 of *Lecture Notes in Computer Science*, pages 164–179. Springer, 2003.

[7] R. Budich, P. Fox, A. Ganguly, J. Kinter, P. Nyberg, and T. Weigel. Climate knowledge discovery workshop report. Technical report, June 2011.

[8] C. C. Chang and H. J. Keisler. *Model Theory, Studies in Logic and the Foundations of Mathematics (3rd ed.)*. Elsevier, 1990.

[9] H. Couclelis. Ontologies of geographic information. *International Journal of Geographic Information Science*, 24:1785–1809, 2010.

[10] I. Cruz and W. Sunna. Structural alignment methods with applications to geospatial ontologies. *Transactions in GIS*, 12(6):683–711, 2008.

[11] M. de Berg, M. van Kreveld, M. Overmars, and O. Schwarzkopf. *Computational Geometry: Algorithms and Applications*. Springer-Verlag, second edition, 2000.

[12] S. Dietze, A. Gugliotta, and J. Domingue. Conceptual situation spaces for semantic situation-driven processes. In M. Hauswirth, M. Koubarakis, and S. Bechhofer, editors, *Proceedings of the 5th European Semantic Web Conference*, LNCS, Berlin, Heidelberg, June 2008. Springer Verlag.

[13] E. Earl and D. Metzler. A new topographic functional. `http://www.peaklist.org/spire/theory/paper.pdf`, 2011.

[14] M. Egenhofer. Toward the semantic geospatial web. In *GIS '02: Proceedings of the 10th ACM international symposium on Advances in geographic information systems*, pages 1–4, New York, NY, USA, 2002. ACM.

[15] M. Gahegan and B. Brodaric. Examining uncertainty in the definition and meaning of geographical categories. In G. J. Hunter and K. Lowell, editors, *5th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, 2002.

[16] P. Gärdenfors. Some tenets of cognitive semantics. *Pragmatics & beyond. New series*, 55:19–36, 1999.

[17] P. Gärdenfors. *Conceptual Spaces: The Geometry of Thought*. A Bradford Book. MIT Press, 2000.

[18] P. Gärdenfors. How to make the semantic web more semantic. In A. C. Varzi and L. Vieu, editors, *Formal Ontology in Information Systems: proceedings of the third international conference (FOIS-2004)*, volume 114 of *Frontiers in Artificial Intelligence and Applications*, pages 17–34. IOS Press, 2004.

[19] P. Gärdenfors and M.-A. Williams. Reasoning about categories in conceptual spaces. In B. Nebel, editor, *IJCAI*, pages 385–392. Morgan Kaufmann, 2001.

[20] W. Garner. *The processing of information and structure*. L. Erlbaum Associates, 1974.

[21] A. J. Gerrard. *Mountain Environments: An Examination of the Physical Geography of Mountains*. The MIT Press, 1990.

[22] J. A. Goguen. What is a concept? In F. Dau, M.-L. Mugnier, and G. Stumme, editors, *Proceedings of the 13th International Conference on Conceptual Structures (ICCS 2005)*, volume 3596 of *Lecture Notes in Computer Science*, pages 52–77. Springer, 2005.

[23] R. L. Goldstone and J. Son. Similarity. In K. Holyoak and R. Morrison, editors, *Cambridge Handbook of Thinking and Reasoning*, pages 13–36. Cambridge University Press, 2005.

[24] M. F. Goodchild, M. Yuan, and T. J. Cova. Towards a general theory of geographic representation in gis. *International Journal of Geographic Information Science*, 21:239–260, January 2007.

[25] N. Guarino. Formal Ontology and Information Systems. In N. Guarino, editor, *International Conference on Formal Ontology in Information Systems (FOIS1998)*, pages 3–15. IOS Press, Trento, Italy, 1998.

[26] D. P. Huttenlocher, G. A. Klanderman, and W. Rucklidge. Comparing images using the hausdorff distance. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(9):850–863, 1993.

[27] K. Janowicz, B. Adams, and M. Raubal. Semantic referencing - determining context weights for similarity measurement. In S. I. Fabrikant, T. Reichenbacher, M. J. van Kreveld, and C. Schlieder, editors, *GIScience*, volume 6292 of *Lecture Notes in Computer Science*, pages 70–84. Springer, 2010.

[28] K. Janowicz, M. Raubal, and W. Kuhn. The semantics of similarity in geographic information retrieval. *Journal of Spatial Information Science*, (2):29–57, 2011.

[29] W. Kuhn. Semantic Engineering. In G. Navratil, editor, *Research Trends in Geographic Information Science*, pages 63–74. Springer, 2009.

[30] O. Kutz, F. Wolter, and M. Zakharyaschev. A note on concepts and distances. In *Proceedings of DL2001*, volume 49 of *CEUR-WS*, pages 113–121, 2001.

[31] T. Lukasiewicz and U. Straccia. Managing uncertainty and vagueness in description logics for the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(4):291 – 308, 2008.

[32] E. B. Lynch, J. D. Coley, and D. L. Medin. Tall is typical: Central tendency, ideal dimensions, and graded category structure among tree experts and novices. *Memory and Cognition*, 28(1):41–50, 2000.

[33] B. Manjunath, P. Salembier, and T. Sikora. *Introduction to MPEG-7: multimedia content description interface*. John Wiley & Sons Inc, 2002.

[34] D. M. Mark, B. Smith, and B. Tversky. Ontology and geographic objects: An empirical study of cognitive categorization. In *International Conference on Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science*, COSIT '99, pages 283–298, London, UK, 1999. Springer-Verlag.

[35] P. Menanteau, D. Finon, and M.-L. Lamy. Prices versus quantities: choosing policies for promoting the development of renewable energy. *Energy Policy*, 31(8):799 – 812, 2003.

[36] S. A. Mitchell. Heights of the principal mountains in the world. `http://www.davidrumsey.com`, 1846.

[37] T. M. Mitchell. *Machine Learning*. McGraw-Hill Series in Computer Science. WCB/McGraw-Hill, Boston, MA, 1997.

[38] B. Motik, P. F. Patel-Schneider, and B. C. Grau. Owl 2 web ontology language: Direct semantics. World Wide Web Consortium, Working Draft WD-owl2-semantics-20081202, December 2008.

[39] M. Raubal. Formalizing conceptual spaces. In A. C. Varzi and L. Vieu, editors, *Formal Ontology in Information Systems, Proceedings of the Third International Conference (FOIS 2004)*, pages 153–164. IOS Press, 2004.

[40] M. Raubal and B. Adams. The semantic web needs more cognition. *Semantic Web*, 1(1-2):69–74, 2010.

[41] A. Rector and J. Rogers. Ontological and practical issues in using a description logic to represent medical concept systems: Experience from galen. In P. Barahona, F. Bry, E. Franconi, N. Henze, and U. Sattler, editors, *Reasoning Web*, volume 4126 of *Lecture Notes in Computer Science*, pages 197–231. Springer Berlin / Heidelberg, 2006.

[42] E. Rosch. Principles of categorization. In E. Rosch and B. B. Lloyd, editors, *Cognition and Categorization*, pages 27–48. Lawrence Erlbaum Associates, Hillsdale (NJ), USA, 1978.

[43] S. Scheider, F. Probst, and K. Janowicz. Constructing bodies and their qualities from observations. In A. Galton and R. Mizoguchi, editors, *Formal Ontology in Information Systems, Proceedings of the Sixth International Conference, FOIS 2010*, pages 131–144. IOS Press, 2010.

[44] R. N. Shepard. Toward a universal law of generalization for psychological science. *Science*, 237(4820):1317, 1987.

[45] M. Sheremet, D. Tishkovsky, F. Wolter, and M. Zakharyaschev. A logic for concepts and similarity. *J. Log. Comput.*, 17(3):415–452, 2007.

[46] A. Sheth, C. Ramakrishnan, and C. Thomas. Semantics for the semantic web: the implicit, the formal and the powerful. *International Journal on Semantic Web and Information Systems*, 1:1–18, 2005.

[47] A. P. Sheth, C. Henson, and S. S. Sahoo. Semantic sensor web. *IEEE Internet Computing*, 12(4):78–83, 2008.

[48] B. Smith and D. M. Mark. Do mountains exist? Towards an ontology of landforms. *Environment and Planning B (Planning and Design)*, 30:411–427, 2003.

[49] G. Stoilos, G. B. Stamou, and J. Z. Pan. Fuzzy extensions of owl: Logical properties and reduction to fuzzy description logics. *Int. J. Approx. Reasoning*, 51(6):656–679, 2010.

[50] M. D. Wills. Hausdorff distance and convex sets. *Journal of Convex Analysis*, 14(1):109–117, 2007.

[51] S. Zednik, P. Fox, and D. L. McGuinness. System transparency, or how I learned to worry about meaning and love provenance! In D. L. McGuinness, J. Michaelis, and L. Moreau, editors, *Provenance and Annotation of Data and Processes - Third International Provenance and Annotation Workshop, IPAW 2010*, volume 6378 of *Lecture Notes in Computer Science*, pages 165–173. Springer, 2010.

[52] F. Zenker. From features via frames to spaces: Modeling scientific conceptual science without incommensurability or aprioricity. `http://www.frankzenker.de/downloads/Zenker%202010%20From%20Features%20via%20Frames%20to%20Spaces.pdf`.