Talking about Forests: an Example of Sharing Information Expressed with Vague Terms

Lucía Gómez Álvarez [a], Brandon Bennett [a] and Adam Richard-Bollans [a]
[a] *University of Leeds, The United Kingdom*

Abstract. Most natural language terms do not have precise universally agreed definitions that fix their meanings. Even when conversation participants share the same vocabulary and agree on taxonomic relationships (such as subsumption and mutual exclusivity, which might be encoded in an ontology), they may differ greatly in the specific semantics they give to the terms.

We illustrate this with the example of 'forest', for which the problematic arising of the assignation of different meanings is repeatedly reported in the literature. This is especially the case in the context of an unprecedented scale of publicly available geographic data, where information and databases, even when tagged to ontologies, may present a substantial semantic variation, which challenges interoperability and knowledge exchange.

Our research addresses the issue of conceptual vagueness in ontology by providing a framework based on supervaluation semantics that explicitly represents the semantic variability of a concept as a set of admissible precise interpretations. Moreover, we describe the tools that support the conceptual negotiation between an agent and the system, and the specification and reasoning within standpoints.

Keywords. concept negotiation, supervaluation, standpoint, forest, GIS, vagueness

1. Introduction

Since the shift in philosophy of language from logical positivism to behaviourism and pragmatism, it is widely accepted that most natural language terms do not have precise universally agreed definitions that fix their meanings. Even when conversation participants share the same vocabulary and agree on taxonomic relationships (such as subsumption and mutual exclusivity, which might be encoded in an ontology), they may differ greatly in the specific semantics they give to the terms in a particular situation. Moreover, except for certain technical terms, individuals do not hold permanent and precise interpretations of the meaning of terms themselves [1].

Humans generally cope with these imprecisions of language by using context and other pragmatic information to narrow the semantic variability of the terms. We say, they 'take a standpoint' on the semantics of the terms. If conflicts occur during the conversation, participants may adapt dynamically their standpoints in order to maintain the cooperation principles (Grice [2]) and achieve successful information exchange. In settings where precision is necessary, such as in scientific or policy making domains, an explicit concept negotiation may be needed if the participants expect or detect conflicting standpoints.

To support human-machine interactions, ontologies are aimed at providing a common vocabulary in which shared knowledge can be represented [3]. However, in most ontology-driven systems, concepts have rigid and static semantics that do not take account of vagueness or context dependence, and in this respect fail to reproduce the conditions of human conversation. Meanwhile, the extensive and increasing amount of data accessible through the internet encourages research on the remaining core challenges facing ontology construction, among them semantic heterogeneity.

In this paper we take the example of negotiating the interpretation of the term 'forest', and propose a framework based on supervaluation semantics to allow for semantic variability within the ontology. We first introduce what we understand as vagueness and how it affects geographical objects in general, followed by the specific case of forest. We then summarise different approaches to vagueness and how they relate to our research. We propose a framework based on standpoint semantics and its tools for supporting share of information, concept negotiation and querying and reasoning within an agent's standpoint.

2. Vagueness and Knowledge Sharing in a Geographic Ontology

Vagueness is pervasive in language [4] and arises whenever a concept or linguistic expression admits of borderline cases of application [5]. We adopt here the distinction proposed in [6] between *sorites* vagueness and *conceptual* vagueness. While *Sorites vagueness* occurs when the applicability of a predicate depends on specific measurable parameters but their thresholds are undetermined, *conceptual vagueness* arises when there is a lack of clarity on which attributes or conditions are essential to fix the meaning of a given term, so that it is controversial how it should be defined [7,6]. This is different from 'simple ambiguity', where a term has more than one distinct meaning.

Moreover, the geographic domain is particularly characterised by objects lacking *bona fide* boundaries and by the inherent vagueness in the definition of geographic descriptors, frames of reference and context [8,9]. When thinking about geographic space, people typically employ several different concepts, and change between them frequently depending on the scale and the perceptual and geometrical properties of the space [10] in addition to the contextual and pragmatic information.

While the value of ontologies is now well established [11], their support of vagueness and semantic heterogeneity[1] remains challenging. One of the main advantages of ontologies is that they improve the interoperability acting to enforce a consensus view reached by a community regarding a certain domain [7]. This is done by formalising the semantics of the terminology of the domain of discourse in some logic formalism. Typically this process involves the cooperation of domain experts which results in a unified decision on the formalization of the semantics of the terminology. However, as a result of the semantic heterogeneity and vagueness of the concepts to define, strong semantic commitments favour specific interpretations of language and involve a loss of generality, thus restricting the opportunities for interoperability. On the other hand, approaches with shallower semantics rely fundamentally on taxonomic relationships, such as subsumption and mutual exclusivity, thus leaving uncertainty on the specific semantics of instances of these terms and potentially compromising the sound reuse of information.

[1] Occurs when ontologies, schemas or datasets of the same domain present differences in meaning and interpretation of categories and/or data values, thus challenging interoperability.

2.1. *The Case of Forests*

In this research we examine the term ‘forest’, for which a broad range of definitions (more than 600 were reported in [12]) have been specified for different purposes. Beyond the spatio-temporal context dependence and fuzziness of their boundaries, many factors contribute to the semantic variability of ‘forest’. They arise from: fundamental differences between the *land cover* / *land use*[2] perspectives, specific uses of the term in different disciplines (ecology, forestry, environmental science, ...) and for different management objectives [14], pragmatic differences between conceptualizations created for science and policy [15], different aspects involved in classifying land as opposed to individuating and demarcating geographic objects [7], and the modelling from endurant or perdurant perspectives on fundamental ontology [16,14].

This scenario poses challenges both for the acquisition of global forest knowledge, particularly its extent and spatial distribution [17,18,19], and for the sound reuse of information and knowledge extraction from the many resources that are increasingly publicly available [20,21]. Moreover, different pieces of research point at cases of cross-disciplinary information reuse among semantically non-interoperable datasets in science, resulting in misleading results [15,14]. At the same time, the option of a centralised and standard definition of ‘forest’ (the definition of the FAO[3] plays that role *de facto*) only supports the discourse of certain institutions and is reported to be unsuitable to capture the needs of different contexts, disciplines and agents [15,14]. In these cases, the lack of recognition for alternate definitions may distort the understanding of certain scenarios such as in the case of contested spaces [16].

Despite the call for addressing the variability of interpretations and definitions of ‘forest’ in the literature [19,14,15,20], most public ontologies containing the concept ‘forest’ either avoid any semantic commitments beyond those embedded in the taxonomies or reduce them to the FAO definition. In the best scenarios, the forest concept has a fuzzy boundary thus not committing to that fixed by the FAO. In this paper we propose a framework that explicitly recognises a variety of acceptable definitions of ‘forest’ and we outline how such a system can support different interactions in a context of semantic heterogeneity. Specifically, we support agents to discover the semantic heterogeneity of a concept in the ontology, to analyse it and to take their standpoint. Subsequently, they can reason and query the ontology according to their standpoint.

3. Approaches to Vagueness

Our intent is to provide a novel way of tackling the prior issues by approaching them from the perspective of *vagueness in ontology*, more specifically conceptual vagueness. In this section, we briefly review some approaches to vagueness and concept creation and negotiation, and how they relate to our research on forest definitions.

The most broadly used logic-based techniques to model vagueness in information systems are multivalued logics [22], in particular fuzzy logic [23]. Fuzzy logic works by assigning *degrees* of truth to statements rather than making truth valuation a binary choice. As a result, this approach provides a reasonably intuitive model for *sorites vagueness*[7], assigning a gradually increasing value to borderline cases as they transition from

---

[2] While the former defines forest in terms of the ecological layer and the physical characteristics of the land, the latter does it with regard to the purpose to which the land is put to use by humans [13].

[3] The Food and Agriculture Organization of the United Nations.

less to more likely. However, fuzzy sets do not fully characterise the different precise overlapping meanings that a term can adopt, which can be sharp but diverse, and fails to incorporate penumbral connections [24] among them. For these reasons, we consider that Fuzzy logic and similar formalisms are not suitable for the problem under consideration.

Conceptual spaces (Gardenfors, [25]), have received great attention in the last years. Although they are not a formal theory of vagueness, they are relevant for this research as they deal with concept formation and representation, expressed within a geometrical space. A concept, say vanilla flavour, is then defined as some region of a four dimensional space of taste, where the dimensions are salt, sour, sweet and bitter. This approach shows interesting features for concept comparison through geometrical operations within a specific n-dimensional space, conceptual adaptation to context (location) in GIS [26] and an account of prototypicality and borderline cases [27]. However, it remains unclear how to tackle the issue of conceptual vagueness, as every concept is defined within a fixed set of quality dimensions; choosing the relevant ones and thus describing the underlying conceptual space for complex and ambiguous categories is not trivial [25].

Semantic heterogeneity and concept negotiation in information systems have been mainly investigated as a separate phenomenon from that of concept creation and vagueness. Instead, they are approached as a phenomenon that *arises* in the context of the need for interoperability of two systems, ontologies or agents, and the necessary conciliations for their successful interoperation. In the area of ontology matching, a wide literature review is provided in [28]. It is our intention to provide a complementary approach to the existing work on the topic, based on the explicit support for semantic variation within an ontology. Thus, the framework aims to support the semantic negotiation of the meaning of its concepts, by providing agents with expressive power to represent their interpretation of the terms of the ontology both when instantiating its concepts and when querying and reasoning within it.

4. A Framework Based on Standpoint Semantics

Standpoint Semantics is our theoretical framework for representing, interpreting and reasoning about information expressed using vague terminology. This theory is an elaboration of the Supervaluation approach, in which the semantics of a vague language is modelled by a set of precise interpretations called precisifications, where each precisification corresponds to a precise and coherent interpretation of all vocabulary of the language.

4.1. Supervaluation Semantics

Consider a formal language based on classical first order logic, with a vocabulary $\mathscr{V} = \mathscr{N} \cup \mathscr{C} \cup \mathscr{R}$, where $\mathscr{N}$ is a set of name symbols (these may be used as constants or as variables when used with a quantifier), $\mathscr{C}$ a set of unary concept terms and $\mathscr{R}$ a set of binary relation symbols. The set $\mathscr{L}$ of formulae of the language is the smallest set that contains Ca and Rab for every $C \in \mathscr{C}$ every $R \in \mathscr{R}$ and every $a, b \in \mathscr{N}$, and, for every φ and ψ in $\mathscr{L}$, formulae $\neg\varphi$, $\varphi \wedge \psi$ and $\varphi \vee \psi$ are also in $\mathscr{L}$.[4]

We define an interpretation structure $\mathscr{I} = \langle P, D, \mathscr{V}, \delta \rangle$, where:

- P is a set of precisifications,
- D is a set of individuals (the domain of discourse),

[4] Here we omit the implication symbol '$\rightarrow$' from the syntax. But it can easily be defined by $\varphi \rightarrow \psi \equiv_{def} \neg\varphi \vee \psi$.

- δ is a denotation function, which, for each precisification and each vocabulary symbol in $\mathscr{V}$, gives the semantic value of that symbol. It is a compound of the following sub-functions:
  - — $\delta_{\mathscr{N}} : (P \times \mathscr{N}) \to D$ maps name symbols to elements of the domain,
  - — $\delta_{\mathscr{C}} : (P \times \mathscr{C}) \to 2^D$ maps concept terms to subsets of the domain,
  - — $\delta_{\mathscr{R}} : (P \times \mathscr{R}) \to 2^{D \times D}$ maps relation symbols to sets of ordered pairs of elements of the domain.

The only difference from the usual classical logic semantics is that the denotation of each symbol is relative to a precisification $p \in P$, which determines a precise interpretation of that predicate. We write $p \Vdash_{\mathscr{I}} \varphi$ to mean that the formula φ is true for precisification p, in interpretation structure $\mathscr{I}$. This determines the truth of propositions as follows:

- $p \Vdash_{\mathscr{I}} Ca$ iff $\delta(p,a) \in \delta(p,C)$
- $p \Vdash_{\mathscr{I}} Rab$ iff $\langle \delta(p,a), \delta(p,b) \rangle \in \delta(p,R)$

- $p \Vdash_{\mathscr{I}} \neg\varphi$ iff it is not the case that $p \Vdash_{\mathscr{I}} \varphi$
- $p \Vdash_{\mathscr{I}} \varphi \wedge \psi$ iff $p \Vdash_{\mathscr{I}} \varphi$ and $p \Vdash_{\mathscr{I}} \psi$
- $p \Vdash_{\mathscr{I}} \varphi \vee \psi$ iff $p \Vdash_{\mathscr{I}} \varphi$ or $p \Vdash_{\mathscr{I}} \psi$
- $p \Vdash_{\mathscr{I}} \forall x[\varphi]$ iff $p \Vdash_{\mathscr{I}'} \varphi$ for every interpretation structure $\mathscr{I}' = \langle P, D, \mathscr{V}, \delta' \rangle$ that is identical to $\mathscr{I}$, except that the value of $\delta'(p,x)$ may be any element of D.[5]

Notice that this semantics not only accounts for variability in the meaning of vocabulary terms but also allows one to enforce dependencies between the terms occurring in a formula. For example, precisifications might vary in the set of people considered tall. If precisification p_1 sets a greater height threshold than p_2 for applicability of the vague property Tall, we might have $\delta(p_1, \mathsf{Tall}) = \{\mathsf{sally}, \mathsf{tom}\}$, $\delta(p_2, \mathsf{Tall}) = \{\mathsf{sally}, \mathsf{tom}, \mathsf{uli}\}$. Hence, we would have $p_1 \Vdash_{\mathscr{I}} \mathsf{Tall}(\mathsf{uli})$ and $p_2 \Vdash_{\mathscr{I}} \neg\mathsf{Tall}(\mathsf{uli})$. But no precisification can make the formula $\mathsf{Tall}(\mathsf{uli}) \wedge \neg\mathsf{Tall}(\mathsf{uli})$ true. Moreover, we can also enforce constraints between different terms in a formula. For instance, we might require that the denotations of Tall and Short are disjoint for all precisifications. Thus, no formula of the form $\mathsf{Tall}(x) \wedge \neg\mathsf{Short}(x)$ would be true in any precisification.

The supervaluation semantics allows us to to augment the logical language with operators that are interpreted relative to the set of precisifications. In particular, we can specify a semantics for $\mathbf{U}\varphi$, meaning that φ is *unequivocally* true, and $\mathbf{S}\varphi$, meaning that φ is *in some sense* true:

- $p \Vdash_{\mathscr{I}} \mathbf{U}\varphi$ iff $q \Vdash_{\mathscr{I}} \varphi$ for every precisification $q \in P$,
- $p \Vdash_{\mathscr{I}} \mathbf{S}\varphi$ iff $q \Vdash_{\mathscr{I}} \varphi$ for some precisification $q \in P$.

4.2. Standpoint Semantics

Standpoint semantics adds detail to the basic supervaluation approach in several ways. The semantic choices that determine each particular precisification are modelled explicitly. These consist of choices of (a) threshold values that determine the applicability of 'sorites vague' predicates such as 'tall' and 'bald'; and (b) choices of definitions that resolve conceptual ambiguities, such as which kinds of vegetation species can be considered as constituting a forest.

[5]Here, for simplicity, we are assuming that the domain of entities will be the same for every precisification. This is not plausible in general.

Standpoint semantics explicitly models the variability of vague terminology in relation to precise observable measurements and properties. Such measurements could be heights, weights, distances and the properties might relate to physical composition, topological relationships. This is the kind of information one might store in a database (or compute directly from the information in a database). Of course, in practice such information might be inaccurate for various reasons such as limitations of measuring equipment or errors in data capture. But these issues are not due to vagueness of terminology, so in formulating the standpoint theory we need not worry about the correctness of the measurement data: we take it as we find it.

4.2.1. A Sublanguage of Objective Observables

We now specify a classical first order language $\mathscr{L}_o$ for describing entities in terms of objective measurements, properties and relations. Let its vocabulary be the set $\mathscr{V}_o = \mathscr{N} \cup \mathscr{D} \cup \mathscr{F} \cup \mathscr{G} \cup \mathscr{C} \cup \mathscr{R}$, where $\mathscr{N}$ is a set of names, $\mathscr{D}$ is a set of numerical expressions (e.g. standard decimals), $\mathscr{F}$ is a set of unary function symbols, $\mathscr{G}$ is a set of binary function symbols, $\mathscr{C}$ is a set of (precise) unary concepts and $\mathscr{R}$ is a set of (precise) binary relations. So in a forestry related domain a vocabulary might be something like:

$$\langle\{\mathsf{tree1},...,\mathsf{boris},...\},\{\mathit{height},\mathit{radius}\},\{\mathit{distance}\},\{\mathsf{Oak},\mathsf{Beech},...\},\{\mathsf{Owns},...\}\rangle\ .$$

Here we have identifiers for individual trees, names of people, unary functions giving height and (canopy) radius measurements of the trees, a binary function giving the distance between any two trees, predicates specifying the species of trees and an owenership relation between trees and people.

In addition to the predications, *Ca* and *Rab*, of $\mathscr{L}$, the vocabulary of $\mathscr{V}_o$ enables us to include in $\mathscr{L}_o$ atomic formulae of the forms $\tau_1 = \tau_2$ and $\tau_1 \leq \tau_2$, where each of τ_1 and τ_2 can be either a functional term (of the form $f(a)$ or $g(a,b)$) or a numerical expression (e.g. in standard decimal notation). So we can write formulae such as:

$$\mathit{distance}(\mathsf{tree1},\mathsf{tree2}) = 23.5 \quad \text{or} \quad \mathit{height}(\mathsf{tree6}) \leq \mathit{height}(\mathsf{tree57}).$$

Like $\mathscr{L}$, the language $\mathscr{L}_o$ is also closed under combination by Boolean connectives and the quantification operators. An interpretation of $\mathscr{L}_o$ is determined by a structure $\mathscr{I}_o = \langle D, \mathscr{V}_o, \delta_o\rangle$, where

- D is a domain of entities,
- δ_o maps the vocabulary in $\mathscr{V}_o$ to their semantic denotations and is a compound of the following sub-functions:

—	$\delta_{\mathscr{N}} : \mathscr{N} \to D$	maps names to elements of the domain,
—	$\delta_{\mathscr{D}} : \mathscr{D} \to \mathbb{Q}$	maps numerical expressions to rationals,
—	$\delta_{\mathscr{F}} : \mathscr{F} \to (D \to \mathbb{Q})$	maps unary function symbols to functions from elements of the domain to rationals,
—	$\delta_{\mathscr{G}} : \mathscr{G} \to ((D \times D) \to \mathbb{Q})$	maps binary function symbols to functions from pairs of elements of the domain to rationals,
—	$\delta_{\mathscr{C}} : \mathscr{C} \to 2^D$	maps concept terms to subsets of the domain,
—	$\delta_{\mathscr{R}} : \mathscr{R} \to 2^{D \times D}$	maps relation symbols to sets of ordered pairs of elements of the domain.

The truth conditions of the language $\mathscr{L}_o$ relative to interpretation $\mathscr{I}_o$ are as follows:

- $\Vdash_{\mathscr{I}_o} Ca$ iff $\delta(a) \in \delta(C)$
- $\Vdash_{\mathscr{I}_o} Rab$ iff $\langle\delta(a),\delta(b)\rangle \in \delta(R)$

- $\Vdash_{\mathscr{I}_o} \tau_1 = \tau_2$ iff $\delta(\tau_1) = \delta(\tau_2)$
- $\Vdash_{\mathscr{I}_o} \tau_1 \leq \tau_2$ iff $\delta(\tau_1) \leq \delta(\tau_2)$

The interpretation of formulae formed by the Boolean connectives and quantifiers will be exactly the same as for standard first order logic, so we do not specify it here.

4.2.2. Adding Vague Predicates

We now specify a language $\mathscr{L}_v$ in which we can specify possible precise interpretations of vague predicates. In adition to $\mathscr{V}_o$, we also have the vocabulary $\mathscr{V}_v = \mathscr{C}_v \cup \mathscr{R}_v \cup \mathscr{T}_v$. where $\mathscr{C}_v$ is a set of vague unary predicates, $\mathscr{R}$ a set of vague binary relations $\mathscr{T}_v$ a set of *threshold* variables. The specification is technically complex, although it has a relatively simple informal explanation. Each precisification determines a mapping of vague conceptual terms and relations to possible definitions and also determiens a valuation of threshold variables that fix the limits of applicability of vague graded predicates. Thus each precisification provides a translation from formulae containing vague terms of $\mathscr{V}_v$ into formulae containing only the precise objective vocabulary of $\mathscr{V}_o$.

Let $\mathscr{L}_d$ be the set of possible defining formulae. These are any well formed formulae over the vocabulary $\mathscr{V}_o \cup \mathscr{T}_v$, where as well as measurement functions and decimals, the terms that can occur in atomic formulae of $\mathscr{L}_d$ formed with the $=$ and $\leq$ relations also include threshold variables. So for example, we could have $\mathit{height}(x) \leq \mathsf{thresh1}$. We write $\mathscr{L}_d^n$ to refer to the subset of $\mathscr{L}_d^n$ containing those formulae that have exactly n distinct free variables — i.e. formulae in which exactly n symbols in $\mathscr{N}$ occur outside the scope of any quantifier.

We now give an explicit specification of the particular interpretation associated with a precisification. Let each precisification p be associated with a tuple $\langle p_T, p_C, p_R \rangle$, where:

- $p_T : \mathscr{T}_v \to \mathbb{Q}$ maps each threshold variable to a number,
- $p_C : \mathscr{C}_v \to \mathscr{L}_d^1$ maps each vague unary predicate to a precise unary definition,
- $p_R : \mathscr{R}_v \to \mathscr{L}_d^2$ maps each vague binary predicate to a precise binary definition.

So for example, for precisification p we could have:

$p_C(\textbf{Tree}) = \mathsf{Plant}(x) \wedge \mathsf{Woody}(x) \wedge \mathsf{min\_tree\_height\_thresh} \leq \mathit{height}(x)$

and $p_T(\mathsf{min\_tree\_height}) = 10$, so precisification p defines **Tree** to be a woody plant of height greater than or equal to 10 (metres).

We can now give an interpretation function for the full language $\mathscr{L}_v$ in which formulae containing vague terms are interpreted relative to a precisification, which maps them into precise formulae of $\mathscr{L}_o$. The interpretation is specified by $\mathscr{I}_v = \langle D, \mathscr{V}_o, \delta_o, P, \mathscr{V}_v, \delta_v \rangle$, where the interpretation of symbols in $\mathscr{V}_o$ is given as for the structure $\mathscr{V}_o$ given above, and the interpretation of formulae in $\mathscr{L}_v$ is given by:

- $p \Vdash_{\mathscr{I}_v} \varphi$ is evaluated as $\Vdash_{\mathscr{I}_o} \varphi$ iff φ does not contain any symbol in $\mathscr{V}_v$,
- $p \Vdash_{\mathscr{I}_v} \tau_1 = \tau_2$ iff $\delta'(p, \tau_1) = \delta'(p, \tau_2)$, where $\delta'(p, \tau) = \delta_o(\tau)$ for all terms in the precise language $\mathscr{L}_o$ (for whose interpretation p is not relevant) and $\delta'(p, \tau) = p_T(p, \tau)$, where τ is a threshold variable (i.e. $\tau \in \mathscr{T}_v$),
- $p \Vdash_{\mathscr{I}_v} \tau_1 \leq \tau_2$ iff $\delta'(p, \tau_1) \leq \delta'(p, \tau_2)$, where $\delta'(\tau)$ is defined the same way as in the previous clause,
- $p \Vdash_{\mathscr{I}_v} \mathbf{C}a$ iff $\Vdash_{\mathscr{I}_o} p_C(\mathbf{C})(a/x)$, where $p_C(\mathbf{C})(a/x)$ is the result of substituting a for the free variable x occurring in the precise definition of $\mathbf{C}$ determined by the function p_C of precisification p,

- $p \Vdash_{\mathscr{I}_v} \mathbf{R}ab$ iff $\Vdash_{\mathscr{I}_o} p_R(\mathbf{R})(a/x, b/y)$, where $p_R(\mathbf{R})(a/x, b/y)$ is the result of substituting a and b for the free variables x and y, in order of their occurrence in the definition of $\mathbf{R}$ determined by the function p_R of precisification p,

5. Towards Sharing Information Linked to Different Precisifications

In this section we explore the potential of the proposed system to enable communication between agents holding diverse standpoints, by providing a shared ontology that supports vague semantics. Our framework enables agents to both link data to the ontology according to their precise interpretation and also to query and reason within the ontology according to a specific standpoint, which can be modified during the interaction. In a context of sharing information, the proposed framework offers tools that enable both the analysis of the relations that hold between precisifications and the execution of modifying operations such as calculating the intersection or union of a pair of precisifications. This, together with information on the instances satisfying these precisifications, serves as a support to the agent for the specification of sound standpoints that guarantee integrity and enable interoperation with the ontology.

5.1. Relations

Assuming that a concept negotiation is necessary, we support the analysis of the relations that hold between precisifications. We identify five main relations, analogous to RCC5, to be inferred by the system. These are:

1. *Equivalence*: $p_1 \leftrightarrow p_2$. Anything that is a 'forest' in p_1 must be a 'forest' in p_2 and vice versa.
2. *Subsumption*: $p_2 \rightarrow p_1$. Anything that is a 'forest' in p_2 must be a 'forest' in p_1.
3. *Inverse subsumption*: $p_1 \rightarrow p_2$. Anything that is a 'forest' in p_1 must also be in p_2.
4. *Disjunction*: $p_1 \rightarrow \neg p_2$. No entity can be a 'forest' in both p_1 and p_2
5. *Overlap*: *None of the previous (NP)* relations hold between p_1 and p_2. Thus there may be entities that are classified as 'forest' in both p_1 and p_2, although neither precisification has a stronger definition that the other.

Relations between precisifications show, to some degree, the connections among them. While overlap is the most common scenario, other relations such as subsumption and disjunction provide valuable information to the agent, to the point of interoperation becoming trivial (such as when information is linked to a definition that is subsumed by the agent's definition) or not feasible because of explicitly conflicting commitments (in the case of disjoint definitions).

It must be noted that the relation between precisifications is purely logical, i.e. is a relationship that holds between the formal representation of two interpretations of the semantics of the ontology, $p1$ and $p2$, which is not dependent on the real world objects that may instantiate them. Consequently, the relation between the logical relations of the precisifications and that of the sets of the real instances (in this world) may not match up. However, the former do constrain the possibilities of the latter. This is shown in Table 2 and discussed in more detail in Section 5.4.

Table 1. The operations (columns) than can be performed between $p1$ and $p2$ and the result with respect to the relations holding between them

Operations		**Logic definitions**				
		$p_1 \leftrightarrow p_2$	$p_2 \rightarrow p_1$	$p_1 \rightarrow p_2$	$p_1 \rightarrow \neg p_2$	NP
Union	$p1 \cup p2$	$p1, p2$	$p1$	$p2$	$p1 \vee p2$*	$p1 \vee p2$
Intersection	$p1 \cap p2$	$p1, p2$	$p2$	$p1$	$\emptyset$	$p1 \wedge p2$
Complement	$p2 \setminus p1$	$\emptyset$	$\emptyset$	$p2 \wedge \neg p1$*	$p2$	$p1 \wedge \neg p2$
Symmetric difference	$p1 \Delta p2$	$\emptyset$	$p1 \wedge \neg p2$	$p2 \wedge \neg p1$	$p1, p2$	$p1 \oplus p2$

5.2. Operations

In order to support both the analysis of the differences and variation among the existing instantiations of the concepts of the ontology and also to facilitate the specification of standpoints that comprise the appropriate subset of precisifications, we define four basic operations between precisifications.

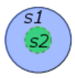
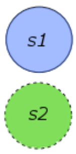
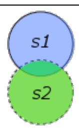
1. *Union:* $p1 \cup p2$. All the instances that satisfy either $p1$ or $p2$ (or both). E.g. Both $p1$ and $p2$ are considered admissible precisifications for a specific use of the ontology by a certain agent.
2. *Intersection:* $p1 \cap p2$. All the instances that satisfy both $p1$ and $p2$. E.g. An agent is interested on the intersection between their own interpretation $p1$ and a precisification in the ontology $p2$.
3. *Complement:* $p2 \setminus p1$. All instances that satisfy $p2$ but not $p1$. E.g. An agent finds $p1$ conflicting with its interpretation and wants to prevent instances of $p1$ from his query on $p2$.
4. *Symmetric difference:* $p1 \Delta p2$. All instances that satisfy either $p1$ or $p2$ but not both. E.g. An agent is interested in exploring the borderline scenarios in which $p1$ and $p2$ disagree.

Table 1 shows the analysis of these operations with respect to the relations holding between $p1$ and $p2$, which allows for the simplification of the result in some scenarios and shows the operation to be trivial or the empty set (∅) in others. In the table, two cases are marked with a star (*), the union of two disjoint precisifications and the complement of a precisification $p1$ subsumed in $p2$. Albeit legal, both operations are potentially problematic and the agent should be aware of the explicit inconsistencies between $p1$ and $p2$ in the former and the subsumption of $p1$ in $p2$ in the latter.

5.3. Applications in the Context of Interoperability and Knowledge Sharing

In the context of knowledge sharing, operations between precisifications are expected to be used for two main purposes, namely concept negotiation and specification of the standpoints. In practice, concept negotiation not only involves the analysis of the objective meaning and ontological commitments formalised on the different precisifications, but also of the real world implications of such commitments. This analysis is possible by querying the ontology with the previous operations, as the agent can gain insight on which instances fall within the borderline scenarios, which instances comply with all the desired precisifications and so on. In the case of GIS, this is enhanced because a spa-

Table 2. Relations between the logical relations among two definitions, $p1$ and $p2$ (columns), and the RCC5 spatial relations between the projections of the instances satisfying those definitions, $s1$ and $s2$ (rows).

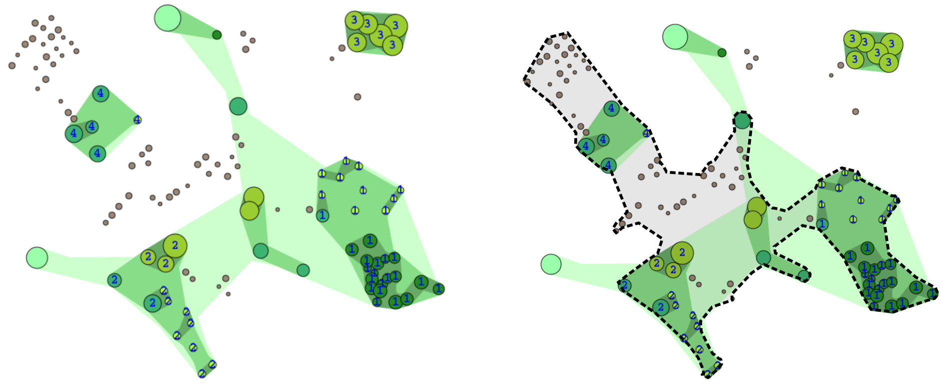
Spatial extension	**Logic definitions**				
	$p_1 \leftrightarrow p_2$	$p_2 \rightarrow p_1$	$p_1 \rightarrow p_2$	$p_1 \rightarrow \neg p_2$	NP
	Expected	The weakness of $p1$ with respect to $p2$ does not manifest in any instance.	The weakness of $p2$ with respect to $p1$ does not manifest in any instance.	Impossible	The only instances lay in the intersection of $p1$ and $p2$.
	Impossible	Expected	Impossible	Impossible	The only instances of p2 lay in the intersection of p2 and p1.
	Impossible	Impossible	Expected	Impossible	The only instances of p1 lay in the intersection of p2 and p1.
	Impossible	Impossible	Impossible	Expected	Despite the definitions overlapping, there are no known instances on the intersection between them.
	Impossible	Impossible	Impossible	Impossible	Expected

tial projection is available, thus providing powerful means for understanding the way definitions perform in specific areas of interest.

Moreover, once the process of concept negotiation is complete for a particular use of the ontology, operations can be used to support the formalization of the agent's standpoint $s1$. During further interactions of the agent with the ontology, queries will be associated with this standpoint. $s1$ acts by restricting the admissible precisifications of the ontology to those which are consistent with $s1$ and enabling reasoning and information retrieval within the subset.

5.4. Interpreting the Spatial Projections of Precisifications/Standpoints

In the case of geographical information, the spatial projection of the instances can be particularly relevant to complement the analysis of the relations that hold between precisifications. One might expect that the possible relations holding between precisifications (equivalence, subsumption, inverse subsumption, overlap and disjunction) would map to the analogous RCC5 relations (equal, proper part, inverse proper part, partial overlap and discrete) between the total spatial area covered by sets of instances satisfying each precisification. However, that does not necessarily need to be the case. Instead, relations between precisifications only restrict the space of possibilities but leave some variation open.



(a) (b)

Figure 1. Example forest diagrams. Figure 1(a) shows four numbered forests with stricter and loser extensions. Figure 1(b) shows the overlap with a different definition of forest

In Table 2 we show the possible relations holding in the different scenarios. While the analogous to RCC5 is the expected relation to hold, other possibilities are allowed. These may be symptomatic of different circumstances. Take, for example, the case where the spatial projection of $p1$ is equal to that of $p2$. It may be the case that $p1$ and $p2$ are equivalent. $p1$ could also be weaker than $p2$ by allowing some more scenarios, but they never manifest in the available data about the state of the world. Even, it could be that $p1$ and $p2$ describe forest in a logically different way (e.g. one uses tree proximity and the other canopy cover) that is highly correlated, and therefore both map to the same objects. In other cases it may be due to mere exemplar clustering[6]. Finally, even when the expected spatial relation holds between the projections, it is expected to be useful for the agent to know how populated are the intersections or borderline scenarios of the definitions under consideration.

6. Examples from Implementation Prototype

The diagrams in Figure 1 give some output examples from our prototype standpoint semantics based software, which enables visualisation of forest data in accordance with different standpoints. In Figure 1(a) we see forest extensions according to three different precisifications in different shades of green. The darker green corresponds to a stricter definition, which requires trees to be closer together to count as constituents of a forest. None of these precisifications consider the shrubs (small brown discs) to be forest constituents. In Figure 1(b) another precisification (demarcated by a black dashed line) is overlayed over the map. According to this precisification the shrubs are counted as forest constituents, so the forest takes on a very different shape, overlapping the extensions associated with the other precisifications.

[6]In ordinary situations, objects that exhibit one property, will very often also exhibit another property and vice versa, even though there is no necessary connection between the properties. The cause may be because of patterns and regularities that are essentially contingent [6].

7. Conclusion

It is widely agreed that, while domain specific ontologies can be used within a community of users for achieving a consensus about conceptualizing, structuring and sharing domain knowledge [29], it is unrealistic to expect that a fixed interpretation could satisfy the needs of different communities in real-world applications. In that scenario, semantic heterogeneity challenges the successful interoperation of systems and agents within decentralized environments. In this paper we have analysed the needs of the multidisciplinary forest community. The semantics of the key concepts of their ontology need to be negotiated depending on the context of use, and systems should be semantically rich to guarantee the sound reuse of information between different domains.

We present a novel approach, consisting on a framework based on ‘Standpoint Semantics’, that makes explicit the semantic variability of the terms of an ontology. We suggest that our framework can support agents with different interpretations and standpoints on their interoperation with a common ontology. Our aim is to provide a description of such a theory and to outline the particular features that can support the conceptual negotiation between agents.

Further work remains to be done to provide a fully operable standpoint semantics ontology, so, in a sense, this paper is only a preliminary step. However, we expect to open the possibility for a complementary approach to the research on semantic alignment in distributed systems. We consider that, in the current context, it is key to support ontologies that can both bring together the views of interdisciplinary domains and that are expressive enough to prevent reasoning and inference from what can be inconsistent interpretations. Moreover, we believe that the explicit acknowledgement of the semantic variability of the terms of the ontology is key for that purpose.

References

[1] P. Egré, V. de Gardelle, and D. Ripley, “Vagueness and Order Effects in Color Categorization,” *Journal of Logic, Language and Information*, vol. 22, no. 4, pp. 391–420, 2013.

[2] H. P. Grice, “Logic and conversation,” in *Syntax and Semantics, Vol. 3, Speech Acts*, pp. 41–58, New York: Academic Press, 1975.

[3] T. R. Gruber, “A translation approach to portable ontology specifications,” *Knowledge Acquisition*, vol. 5, pp. 199–220, 6 1993.

[4] L. C. Burns, *Vagueness : an Investigation into Natural Languages and the Sorites Paradox*. Springer Netherlands, 1991.

[5] A. C. Varzi, “Vagueness,” *Encyclopedia of Cognitive Science*, pp. 459–464, 2003.

[6] B. Bennett, “Modes of Concept Definition and Varieties of Vagueness,” *Applied Ontology*, vol. 1, no. 1, pp. 17–26, 2005.

[7] L. Gómez Álvarez and B. Bennett, “Classification, Individuation and Demarcation of Forests: formalising the multi-faceted semantics of geographic terms,” in *13th International Conference on Spatial Information Theory*, Leibniz International Proceedings in Informatics, 2017.

[8] P. Agarwal, “Ontological Considerations in GIScience,” *International Journal of Geographical Information Science*, vol. 19, no. 5, pp. 501–536, 2005.

[9] B. Smith and D. M. Mark, “Ontology and geographic kinds,” in *International Symposium on Spatial Data Handling*, 1998.

[10] M. J. Egenhofer and D. M. Mark, “Naive Geography,” in *International Conference on Spatial Information Theory*, pp. 1–15, Springer Berlin Heidelberg, 1995.

[11] N. Guarino, “Formal Ontology and Information Systems,” *Formal ontology in information systems: Proceedings of the first international conference*, vol. 46, pp. 81–97, 1998.

[12] H. G. Lund, “Definitions of Forest, Deforestation, Afforestation and Reforestation,” *Forest Information Services*, pp. 1–159, 2007.

[13] W. Meyer and B. Turner, *Changes in land use and land cover : a global perspective*. Cambridge University Press, 1994.

[14] R. Chazdon, P. Brancalion, and L. Laestadius, "When is a forest a forest? Forest concepts and definitions in the era of forest and landscape restoration," *Ambio*, vol. 45, no. 5, pp. 538–550, 2016.

[15] A. Grainger, "Uncertainty in the construction of global knowledge of tropical forests," *Progress in Physical Geography*, vol. 34, no. 6, pp. 811–844, 2010.

[16] T. Wong, C. O. Delang, and D. Schmidt-Vogt, "What is a forest? Competing meanings and the politics of forest classification in Thung Yai Naresuan Wildlife Sanctuary," *Geoforum*, vol. 38, pp. 643–654, 2007.

[17] D. Fuller, "Tropical forest monitoring and remote sensing: A new era of transparency in forest governance?," *Singapore Journal of Tropical Geography*, vol. 27, no. 1, pp. 15–29, 2006.

[18] C. Giri, Z. Zhu, and B. Reed, "A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets," *Remote Sensing of Environment*, vol. 94, no. 1, pp. 123–132, 2005.

[19] H. G. Lund, "What is a forest? Definitions do make a difference: An example from Turkey," *Avrasya Terim Dergisi*, vol. 2, no. 1, pp. 1–8, 2014.

[20] A. Comber, P. Fisher, and R. Wadsworth, "What is land cover?," *Environment and Planning B: Planning and Design*, vol. 32, no. 2, pp. 199–209, 2005.

[21] A. Comber, P. Fisher, and R. Wadsworth, "You know what land cover is but does anyone else? an investigation into semantic and ontological confusion," *International Journal of Remote Sensing*, vol. 26, no. 1, pp. 223–228, 2005.

[22] A. Tarski, *Logic, semantics, metamathematics: papers from 1923 to 1938*. Hackett Publishing, 1980.

[23] L. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.

[24] H. Kamp and B. Partee, "Prototype theory and compositionality," *Cognition*, vol. 57, no. 2, pp. 129–191, 1995.

[25] P. Gardenfors, "Conceptual spaces as a framework for knowledge representation," *Mind and Matter*, vol. 2, no. 2, pp. 9–27, 2004.

[26] B. Adams and K. Janowicz, "Constructing geo-ontologies by reification of observation data," in *Proceedings of the 19th ACM SIGSPATIAL- GIS '11*, p. 309, 2011.

[27] I. Douven, L. Decock, R. Dietz, and P. Égré, "Vagueness: A Conceptual Spaces Approach," *Journal of Philosophical Logic*, vol. 42, no. 1, pp. 137–160, 2013.

[28] Y. Kalfoglou and M. Schorlemmer, "Formal Support for Representing and Automating Semantic Interoperability," in *European Semantic Web Symposium*, pp. 45–60, Springer Berlin Heidelberg, 2004.

[29] A. Maedche, B. Motik, N. Silva, and R. Volz, "MAFRA—a mapping framework for distributed ontologies," in *International Conference on Knowledge Engineering and Knowledge Management*, pp. 235–250, Springer Berlin Heidelberg, 2002.