Dealing with Conceptual Indeterminacy: A Framework based on Supervaluation Semantics

Lucía Gómez Álvarez[1] and Brandon Bennett[1]

University of Leeds, Leeds LS2 9JT, United Kingdom
{sc14lga,B.Bennett}@leeds.ac.uk

Abstract. It is widely accepted that most natural language terms do not have precise universally agreed definitions that fix their meanings. Instead, humans use terms in a variety of ways that adapt to different contexts and points of view.

In this paper we present a framework based on Supervaluation Semantics for interpreting languages in the presence of *semantic variability*. This work builds on supervaluationist accounts, which explain linguistic indeterminacy in terms of a collection of possible precise interpretations of the terms of the language. We extend the basic supervaluation semantics by adding the notion of *standpoint*. A multi-modal logical language for describing standpoints is presented. The language includes a modal operator $\Box_s$ for each standpoint s, such that $\Box_s\,\phi$ means that proposition ϕ is unequivocally true according to standpoint s — i.e. ϕ is true at all precisifications compatible with s. We show how it can be used to represent logical properties and connections between alternative ways of describing a domain and different accounts of the semantics of terms.

Keywords: Supervaluation Semantics · Modal logics · Vagueness · Multi-modal logics.

1 Introduction

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 [10].

This phenomenon has been approached in many ways within a highly interdisciplinary area of research, from philosophy to linguistics, cognitive science and artificial intelligence. Philosophical interest on this issue dates well back to Ancient Greek philosophy [3, 4] and has often been studied as vagueness in natural language. One traditional view is that imprecision of terms presents an obstruction to good philosophy [12], which should be circumvented by establishing precise definitions. But views of human language becoming prominent in the 20th century (e.g. [14, 24]) have tended to accept semantic heterogeneity as a fundamental feature of human communication[1] that improves the adaptability of terms to diverse contexts. Numerous logical theories of vagueness have been proposed to model different aspects of it. To date, the most popular approaches are based on *many-valued logics* [23], *supervaluation semantics* [11] and mainly *fuzzy logic* [26, 25]. 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 of sorites vagueness. However, fuzzy sets do not fully characterise the different precise overlapping meanings that a term can adopt, which can be sharp but diverse and relevant in different contexts, and fails to incorporate *penumbral connections* [17] among them.

The current research is mainly concerned with the problem of reasoning within a scenario of semantic heterogeneity. Instead of aiming at representing the full variation of certain natural language terms (e.g. via a fuzzy set), we focus on providing a means to model scenarios in which certain interpretations or standpoints coexist and investigate reasoning within such settings. In this way, our framework can be seen as complementary rather than rival to others.

Furthermore, ontology matching techniques are mainly conceived to 'solve' heterogeneous situations instead of 'representing' them. Our intention is to supplement the existing work on the topic of conceptual indeterminacy by providing explicit support for semantic variation within the representation framework. Nevertheless, as discussed in the future work section of our paper, we consider that ontology matching techniques could have a huge potential to populate supervaluated ontologies as well as to guide both logic and heuristic reasoning within them.

The rest of the paper is organised as follows. It first presents a motivation example in chapter 2. We then introduce the supervaluationist account vagueness in chapter 3, overview the core ideas of standpoint semantics in chapter 4 and give a formal specification of the standpoint logic in chapter 5. Following this, we illustrate the expressive capabilities of the logic (chapter 6) and discuss its application to combining ontologies (chapter 7). We revisit the use case in chapter 8 and conclude with some final remarks and indication of future work.

2 A Motivating Example

How much forest is there in the world? is a surprisingly difficult question to answer [1]. A broad range of forest concepts and definitions [19] have been specified for different purposes, leading to discrepancies of estimates due to the different conceptual and methodological approaches [13]. This has been recognised to be one of the key challenges [7] preventing the use of the more and more datasets (e.g. [15]) and portals (e.g. Global Forest Watch) that have emerged for enabling

[1] Some research suggests that it may enable more efficient communication [27, 21].

the tracking, comparing and understanding of the available information on land use-cover and global forest extent.

We consider the situation of a decision maker needing to understand a complex scenario in which data and conceptual models coming from different contexts need to be critically analysed. We consider five examples of inferences relevant to the aspects in which the diversity of forest definitions has been reported to pose challenges for the acquisition of global knowledge [7]. For the current purpose we set the use case of a stakeholder using one of the portals (e.g. GFW) and the proposed framework to reason about the following queries:

Q1. Is area a consistently a forest for all reasonable forest interpretations or is it only *arguably* a forest? What discrepancies are there?
Q2. Area a has been degraded from the status of original 'intact forest' and its status is now contested. How is it described according to other views?
Q3. Are there inconsistencies among the data or only different data classifications? E.g. Are there areas reported to satisfy some criteria that is inconsistent with the way they are classified by others?
Q4. What can we say about area a that is differently classified under different interpretations of terminology?

The scenario is not dynamic or informal. On the contrary, the agents that set their standpoints are institutions and/or scientific communities that produce and distribute data relevant to specific contexts. E.g. An area in Iran may have been considered a forest by a national agency while according to a global definition it may not be. We shall return to this example towards the end of the paper, once we have presented our *logic of standpoints*, and consider how our representation can be used to express the requirements of these queries.

3 Supervaluation Semantics

In supervaluationist accounts of linguistic indeterminacy the meanings of vague terms are explained in terms of a collection (a set or some more structured ensemble) of possible precise interpretations (often called precisifications). An early proposal that vagueness can be analysed in terms of multiple precise senses was made by Mehlberg [20] and a formal model based on a multiplicity of classical interpretations was applied to the analysis of vagueness by Fine [11]. A similar approach was proposed by [16], which has been influential among linguists.

Apart from providing a general framework for specifying a semantics of vagueness, the supervaluationist idea is also attractive in that it can account for penumbral connection [11], which many believe to be an essential ingredient of an adequate theory of vagueness. This is the phenomenon whereby logical laws (such as the principle of non-contradiction) and semantic constraints (such as mutual exclusiveness of two properties – e.g. '... is red' and '... is orange') are maintained even for statements involving vague concepts. The solution, in a nutshell, being that, even though words may have multiple different interpretations, each admissible precisification of a language makes precise all vocabulary in a way that ensures mutual coherence of the interpretation of distinct but semantically related terms.

Moreover, such frameworks offer a natural treatment of what we call *conceptual vagueness*[2], which arises when there is a lack of clarity on which attributes or conditions are essential to the meaning of a given term, so that it is controversial how it should be defined. This kind of conceptual indeterminacy underlies the controversy about whether to define forest in terms of land cover or land use, and this case is a good illustration of the overlapping of the applicability of the two interpretations, since the use to which land can be put depends to a large extent on the material and ecological properties of its land cover.

4 The Logic of Standpoints

We now overview the ideas of *standpoint semantics*, which is essentially an elaboration of supervaluation semantics introduced in [6]. However, whereas that paper is concerned with anchoring precisifications (and hence standpoints) to objective properties of the world, this work is mainly aimed at modelling and reasoning with different partial interpretations of the language (i.e. different standpoints).

A *precisification* in standpoint semantics is identified with a precise interpretation of the language, such that for any state of the world there is a unique extension for every predicate. This notion differs from Fine's terminology, where a precisification need not be completely precise, but leave certain propositions indeterminate. Fine's precisifications form a partial order where p_2 is more precise than p_1 if all propositions with a definite truth value in p_1 have the same truth value in p_2 but some propositions that are indeterminate in p_1 have a determinate truth value in p_2.

The standpoint semantics [6] uses an almost equivalent but conceptually simpler model in which a partially determinate interpretation is called a *standpoint* and is modelled by the set of all (fully determinate) precisifications that are consistent with the (partially determinate) standpoint. Since they are modelled as sets, standpoints form a lattice under the subset relation; and, when one standpoint is a subset of another, we may also regard it as more precise, since it rules out certain interpretations of the language.

Standpoints are collections of precisifications. They are typically organised by constraints (axioms and threshold limitations) that pick out a corresponding set of admissible precisifications that satisfy these constraints. For example, a standpoint might be constrained by the definition **E1**, but the value of the threshold $\mathsf{t\_tall}$ might not be fixed although constrained to lie between certain values **E2**.

E1. $\forall x[\mathsf{Tall}(x) \leftrightarrow \mathsf{height}(x) > \mathsf{t\_tall}]$

[2] The distinction between *sorites vagueness* (the applicability of a predicate depends on measurable parameters but their thresholds are undetermined) and *conceptual vagueness* is introduced in [6].

E2. $\mathsf{t\_tall} > 175cm \wedge \mathsf{t\_tall} < 185cm$

Such values may not be explicitly given, since they could be inferred from assertions associated with the standpoint. For instance, if a person of height 175cm is asserted not to be tall, the threshold for tallness must be greater than 175cm.

5 The Formal Language Standpoint Logic (SL)

The formalism that we present here goes significantly beyond [6] in explicitly representing standpoints within the logic of the object language by means of modal operators (previously standpoints were handled by auxiliary semantic apparatus).

Standpoint logic is a multimodal logic with partially ordered modalities. Standpoints are modelled as a set of modal operators: If s is a standpoint and $\square_s \phi$ then ϕ is the case for all precisifications which are compatible with the standpoint s. The partial ordering encodes the subset relation which may hold between two standpoints s and s'.

5.1 Syntax

Our formal language $\mathcal{L}_S$ is an extension of classical first-order calculus including numerical symbols and comparison relations ($=$ and $<$) as well as the usual boolean operators and quantifiers.

Vocabulary. The non-logical symbols of the language are specified by a *vocabulary*, which is a tuple of the form:

$$\mathbb{V} = \langle \mathcal{N}, \mathcal{X}, \mathcal{R}, \mathcal{F} \rangle \ ,$$

where $\mathcal{N}$ is a set of nominal constants, $\mathcal{X}$ is a set of nominal variables, $\mathcal{R} = (\mathcal{R}_1, \cup \ldots \cup \mathcal{R}_n \cup \ldots)$ is the set of predicate/relation symbols, whose subsets $\mathcal{R}_n$ are the sets of n-ary predicate symbols and $\mathcal{F} = (\mathcal{F}_1, \cup \ldots \cup \mathcal{F}_n \cup \ldots)$ is the set of function symbols, subsets $\mathcal{F}_n$ being the sets of n-ary function symbols.

Terms. The language has two types of terms: one type refer to individual entities and the other refer to numerical magnitudes:

- $\mathcal{T}_n = \mathcal{N} \cup \mathcal{X}$ is the set of *nominal terms* of the language.
- $\mathcal{T}_m = \mathcal{T}_{\mathbb{D}} \cup \{ f(\tau_1, \ldots, \tau_n) \mid \ f \in \mathcal{F}_n \wedge \tau_1, \ldots \tau_n \in \mathcal{T}_n \}$ is the set of *magnitude terms*.

The set $\mathcal{T}_m$ includes the set $\mathcal{T}_{\mathbb{D}}$ of decimal numerals, as well as terms formed by applying function symbols to nominal terms, which give the value of some scalar property of an entity (e.g. height) or tuple of entities (e.g. the distance between two entities).

Atomic Propositions. The language has the following forms of atomic proposition:

- $R(\tau_1, \ldots \tau_n)$, where $\tau_1, \ldots \tau_n \in \mathcal{T}_n$,
- $\tau_1 = \tau_2$, where $\tau_1, \ldots \tau_2 \in (\mathcal{T}_n \cup \mathcal{T}_m)$,
- $\tau_1 \leq \tau_2$, where $\tau_1, \ldots \tau_2 \in \mathcal{T}_m$

$R(\tau_1, \ldots \tau_n)$ asserts that relation R holds of the nominal terms $\tau_1, \ldots \tau_n$ (which are named entities and/or quantified variables) and $\tau_1 = \tau_2$ is the usual equality relation, that can hold either between named entities and/or variables.

Complex Propositions. For any $\phi, \psi \in \mathcal{L}_S$, and $x \in \mathcal{X}$ the following complex propositions are also in $\mathcal{L}_S$:

- $\neg\phi$, $(\phi \wedge \psi)$, $(\phi \vee \psi)$, $(\phi \rightarrow \psi)$, $(\phi \leftrightarrow \psi)$, $\forall x[\phi]$, $\exists x[\phi]$ — the standard boolean operators and first-order quantifiers,
- $\Box_s \phi$ meaning ϕ is true in standpoint s, i.e. in all precisifications compatible with standpoint s.
- $\Box_* \phi$ — meaning ϕ is true in all precisifications. $\Box_* \phi$ is a special case of $\Box_s \phi$ where $*$ is the standpoint containing the set containing all precisifications.

$\mathcal{L}_S$ is the smallest set containing all atomic propositions and all complex propositions formed by these constructions.

We can easily define other useful operators such as $\Diamond_s \phi \ \equiv_{def} \ \neg \Box_s \neg\phi$, i.e. according to standpoint s ϕ *may* be considered true; $\mathcal{I}_s \phi \ \equiv_{def} \ (\Diamond_s \phi \wedge \Diamond_s \neg\phi)$, i.e. the truth of ϕ is indeterminate with respect to standpoint s, or $\mathcal{D}_s \phi \ \equiv_{def} \ (\Box_s \phi \vee \Box_s \neg\phi)$, i.e. according to standpoint s ϕ has a determinate truth value.

5.2 Semantics

A standpoint semantics interpretation structure is a tuple

$$\mathbb{S} = \langle P, D, \mathbb{V}, \mathcal{S}, \sigma, \delta, \rho \rangle \ ,$$

where:

- P is the set of precisifications,
- D is a non-empty set, the domain of individuals,
- $\mathbb{V} = \langle \mathcal{N}, \mathcal{X}, \mathcal{R}, \mathcal{F} \rangle$ is a vocabulary,
- $\mathcal{S}$ is the set of standpoint symbols.
- $\sigma : \mathcal{S} \rightarrow (2^P/\emptyset)$ is a function mapping each standpoint symbol to a non-empty set of precisifications,
- $\delta = \delta_n \cup \delta_m$, where $\delta_n : \mathcal{T}_n \rightarrow D$ maps each nominal term to an element of the domain of individuals and $\delta_m : \mathcal{T}_m \rightarrow \mathbb{Q}$ maps each magnitude term to a real number.
- $\rho = (\rho_1 \cup \ldots \cup \rho_n \ldots)$, where $\rho_n : \mathcal{R}_n \times P \times D^n \rightarrow \{\mathbf{t}, \mathbf{f}\}$. So ρ maps, each n-ary predicate, precisification and n-tuple of individuals to a truth value.

With respect to the interpretation structure, formulae are interpreted as follows:

- $[\![r(\tau_1, \dots \tau_n)]\!]^p_{\mathbb{S}} \;=\; \rho(r, p, \langle\delta(\tau_1), \dots \delta(\tau_n)\rangle))$,
- $[\![(\tau_1 = \tau_2)]\!]^p_{\mathbb{S}} \;=\; \mathbf{t}$ if $\delta(\tau_1) = \delta(\tau_2)$, else $= \mathbf{f}$,
- $[\![(\tau_1 \leq \tau_2)]\!]^p_{\mathbb{S}} \;=\; \mathbf{t}$ if $\delta(\tau_1) \leq \delta(\tau_2)$, else $= \mathbf{f}$,
- $[\![\neg\phi]\!]^p_{\mathbb{S}} \;=\; \mathbf{t}$ if $[\![\phi]\!]^p_{\mathbb{S}} = \mathbf{f}$, else $= \mathbf{f}$,
- $[\![\phi \wedge \psi]\!]^p_{\mathbb{S}} \;=\; \mathbf{t}$ if $[\![\phi]\!]^p_{\mathbb{S}} = \mathbf{t}$ and $[\![\psi]\!]^p_{\mathbb{S}} = \mathbf{t}$, else $= \mathbf{f}$,
- $[\![\forall x[\phi]]\!]^p_{\mathbb{S}} \;=\; \mathbf{t}$ if ($[\![\phi]\!]^{pw}_{\mathbb{S}'} = \mathbf{t}$, for every interpretation structure $\mathbb{S}' = \langle P, D, \mathbb{V}, \mathcal{S}, \sigma, \delta', \rho\rangle$, such that δ' is identical to δ, except that $\delta'(x)$ *may* have a different value from $\delta(x)$), else $= \mathbf{f}$,
- $[\![\Box_s \phi]\!]^p_{\mathbb{S}} \;=\; \mathbf{t}$ if $[\![\phi]\!]^{p'w} = \mathbf{t}$ for all $p' \in \sigma(s)$, else $= \mathbf{f}$. Note that this includes the case there $s = *$.

5.3 Proof Theory

We now give a set of axioms for the logic, which capture significant aspects of the semantics. We believe all our axioms are sound with respect to the semantics, but we do not claim they are complete. Establishing a proof system that we can prove sound and complete is the subject of ongoing work.

Being built upon an underlying classical logic we have that ϕ is a theorem of $\mathcal{L}_s$ if it is an axiom of first-order logic (where we treat all modal sub-formulae as atomic propositions).

Modal Axioms. In the following axioms hold for any of the operators $\Box_{s_i}$ and the special case $\Box_*$.

AN $\Box_{\mathrm{s}}\, \phi$ for any theorem ϕ

AK. $\Box_{\mathrm{s}}(\phi \rightarrow \psi) \rightarrow (\Box_{\mathrm{s}}\, \phi \rightarrow \Box_{\mathrm{s}}\, \psi)$

AT. $\Box_* \phi \rightarrow \phi$

AD. $\Box_s \phi \rightarrow \Diamond_s \phi$

A4. $\Box_{\mathrm{s}}\, \phi \rightarrow \Box_{\mathrm{s}} \Box_{\mathrm{s}}\, \phi$

A5. $\Diamond_{\mathrm{s}}\, \phi \rightarrow \Box_{\mathrm{s}} \Diamond_{\mathrm{s}}\, \phi$

A1. $\Box_s \phi \rightarrow \Box_{s'} \phi$ for $(s' \preceq s)$ [3]

A2. $\Box_s \phi \leftrightarrow \Box_* \Box_s \phi$

A3. $\Box_* \phi \rightarrow \Box_s \phi$

Thus, the $\Box_*$ operators is an **S5** modality, whereas the rest of $\Box_s$ operators satisfy the axioms of the modal logic **KD45**.

Axiom **A1** can also be expressed in the form $\Diamond_{s'} \phi \rightarrow \Diamond_s \phi$ for $(s' \preceq s)$ and it captures the partial order between standpoints, by ensuring that any proposition considered definite in a given standpoint, is also considered definite in any sharper standpoint. **A3** is easily derivable from **A1**. **A2** captures the property of the semantics whereby once we apply a standpoint operator to a propositional formula, we will get a proposition that is either true at all precisifications or false at all precisifications. This is like in the well-known modal logic **S5**, where $\Box \phi \rightarrow \Box\Box \phi$ and $\Diamond \phi \rightarrow \Box \Diamond \phi$. By combining **A2** and **A3**, we get:

T1. $\Box_{s'} \Box_s \phi \leftrightarrow \Box_s \phi$

[3] $s' \preceq s$ asserts that s' is at least as sharp as s i.e. $\sigma(s_1) \subseteq \sigma(s_2)$; every proposition that is unequivocally true in s is also true in s'.

Quantifier Axioms. The universal quantifier satisfies its classical axioms, which are covered by axiom **C** above. What about axioms specifying the interaction between the modal operators and quantification?

Since our semantics is based on a single domain of individuals, its models will satisfy the Barcan formula:

A7. $\forall x[\square_{\mathrm{s}}\, \phi(x)] \rightarrow\ \square_{\mathrm{s}}\, \forall x[\phi(x)]$

Issues Regarding the Domain of Quantification. One can argue that if we change the precisification according to which the world is classified, the set of entities is likely to change. For example, under one precisification, a particular tree-covered area might be determined to form a single forest under one precisification, whereas under another it might be determined as consisting of two forests separated by a band of heathland. However, one could take the contrary view that, the set of entities can be regarded as the same even though their classification changes. This is consistent with a *de dicto* view of vagueness, in which it is only linguistic descriptions that are vague, not the objects that they describe. In fact, we believe that our semantics is inadequate to adequately express correspondencies between entities at different worlds and/or precisifications, e.g. we cannot say that different tree covered areas correspond to the *same* forest seen from different standpoints, which agree that there is one forest within a given region, but disagree about its extent.

6 Expressive Capabilities of $\mathcal{L}_S$

We now examine the expressive capabilities of our standpoint logic and illustrate them with a variety of examples.

Judgements Regarding Graded Predicates. Problems of formalising reasoning with vague predicates are the subject of a huge amount of debate in the philosophical literature. Here we do not go into detail about the issues involved, but just give a simple example of how our formalism can deal with some aspects of this problem in relation to attributions of the typical graded adjectives 'tall' and 'short' for two standpoints $s1$ and $s2$ where $s_2 \preceq s_1$.

E3. $\square_* \neg\exists x[\mathsf{Tall}(x) \wedge \mathsf{Short}(x)]$
E4. $\square_*[\boldsymbol{height}(\mathsf{tara}) = 186cm]$
E5. $\square_*[\boldsymbol{height}(\mathsf{simon}) = 160cm]$
E6. $\square_{s_1}[\forall x[\mathsf{Tall}(x) \leftrightarrow \boldsymbol{height}(x) > \mathsf{t\_tall}]$
E7. $\square_{s_1}[\mathsf{t\_tall} < 185cm]$
E8. $\square_{s_1}[\mathsf{Short}(\mathsf{simon})]$
E9. $\square_{s_2}[\mathsf{t\_tall} = 180cm]$

All precisifications must satisfy the given penumbral connection axiom **E3**: nothing can be both Tall and Short. Thus this formula will be true in all standpoints. Formulae **E4** and **E5** express objective, non-vague facts that are taken to be true for all precisifications. Given **E3**, **E4** and **E7** we know that $\square_{s_1}[\mathsf{Tall}(\mathsf{tara})]$.

Because we know that s_2 is at least as sharp as s_1, by using axiom **A1** we can infer that $\Box_{s_2}[\mathsf{Tall}(\mathsf{tara})]$. Note that $s_2 \preceq s_1$ implies that at least those definite tall cases in s_1 are also definite in s_2, so the threshold for tallness in s_2 must be equal or lower than that of s_1. Also, with **E8**, we know $\Box_{s_1}[\mathsf{t\_tall} > 160cm]$.

Conceptual Variation. In $\mathcal{L}_S$ it is easy to stipulate that a particular definitional axiom holds for a particular standpoint. In the domain of forests we might have:

E10. $\Box_{s_1} \forall x[\mathsf{Plant}(x) \leftrightarrow \mathsf{Embryophyta}(x)]$
E11. $\Box_{s_1} \forall x[\mathsf{Tree}(x) \leftrightarrow (\mathsf{Plant}(x) \wedge \mathsf{Woody}(x) \wedge \mathsf{Tall}(x))]$

In another standpoint s_2 we might have:

E12. $\Box_{s_2} \forall x[\mathsf{Plant}(x) \leftrightarrow \mathsf{Organism}(x) \wedge \neg\mathsf{Animal}(x)]$

E10 and **E12** express inconsistent definitions (a mushroom would be a Plant accordint to s_2 but not according to s_1) but the formulae can coexist consistently within $\mathcal{L}_S$ because the definitions are within different standpoint operators.

Penumbral Connections. Supervaluation semantics regards precisifications as applying to the whole language. This is because the meaning of related vague concepts cannot always vary independently. We do not allow precisification in which related concepts are interpreted inconsistently (e.g. someone can be both 'tall' and 'short'). Modelling this kind of *penumbral connection* is one of the main motivations of supervaluation semantics [11]. Our standpoint logic enables us to impose penumbral connections between concepts both for what we consider to be all reasonable interpretations, using the $\Box_*$ operator, and also from the point of view of some particular standpoint s, by means of the $\Box_s$ operator. We have already seen in **E3**, a formula expressing a condition that two predicates are mutually exclusive over all precisifications.

It may also be useful to specify that concepts are semantically independent within a standpoint (or globally). This may allow for the modularisation of a vocabulary into sets of terms that do not impinge on each other (e.g. concepts relating to forests can be modelled without worrying about mountains or buildings). We say that two vague propositions ϕ and ψ are *penumbrally independent* when every sense in which ϕ can be interpreted is compatible with every sense which ψ can be interpreted. In $\mathcal{L}_S$ logic we can express this by:

$$\begin{array}{lcl} (\ (\Diamond_s \phi \wedge \Diamond_s \psi) \rightarrow \Diamond_s(\phi \wedge \psi) & \wedge & (\Diamond_s \phi \wedge \Diamond_s \neg\psi) \rightarrow \Diamond_s(\phi \wedge \neg\psi) \wedge \\ \ \ (\Diamond_s \neg\phi \wedge \Diamond_s \psi) \rightarrow \Diamond_s(\neg\phi \wedge \psi) & \wedge & (\Diamond_s \neg\phi \wedge \neg \Diamond_s \psi) \rightarrow \Diamond_s(\neg\phi \wedge \neg\psi)) \end{array}$$

Concept sharpness. We can say that a particular predicate R is extensionally sharp in a given standpoint s with a formula such as $\forall\overline{x}\mathcal{D}_s[R(\overline{x})]$, where $\overline{x}$ stands for n different universally quantified variables and where n is the arity of R.

Moreover, we can express the condition that 'standpoint s' has an extensionally sharper interpretation of concept C than standpoint s', which we represent by the form $(s' \preceq s) : \lambda\overline{x}[\mathsf{C}(\overline{x})]$. This is just syntactic sugar for an actual formula of $\mathcal{L}_S$ that can be defined as follows:

$$(s' \preceq s) : (\lambda\overline{x}[\mathsf{C}(\overline{x})]) \equiv (\ \forall x[\Box_s \mathsf{C}(\overline{x}) \rightarrow\ \Box_{s'} \mathsf{C}(\overline{x})] \wedge \forall x[\Box_s \neg\mathsf{C}(\overline{x}) \rightarrow\ \Box_{s'} \neg\mathsf{C}(\overline{x})]\ )$$

7 Combining Unaligned Ontologies

In the process of defining an ontology, there may be a number of different meanings attached to certain terms. Moreover, users may hold that having all these meanings is useful, as they are relevant for different contexts [9, 8]. While the analysis of [9] proposes a descriptive strategy to deal with the semantic variability of terms, we suggest that 'supervaluating' concepts preserves the advantages of the former strategy while: (a) making explicit the fact that such series of definitions are linked to a single vague concept; and (b) making further inferences possible through the framework proposed here.

We thus see the potential for *supervaluated ontologies* by design, that are not the fruit of the integration of different models but a direct formalisation of a domain in which the semantic variablility of its terms is sufficiently meaningful and/or relevant to be represented in the ontology.

Embedding Ontologies within $\mathcal{L}_S$. Moreover, our language $\mathcal{L}_S$ provides many ways in which the relationships between terms in different ontologies can be described and we are still investigating which of these are most practical. However, we now outline some possibilities. Any given ontology $\mathcal{O}_i$ is identified with a set of formulae incorporating both semantic constraints (its *T-Box*) and facts (its *A-Box*). We also associate $\mathcal{O}_i$ with a standpoint s_i. Our aim is to construct a theory Θ in $\mathcal{L}_S$ in which we embed two or more unaligned ontologies.

We assume that every formula of the ontology is unequivocally true with respect to that standpoint, so an ontology's standpoint can only be indefinite about formulae that are not explicitly stated or implied by its explicit statements.

We suggest that when interpreting and combining ontologies it will be useful to also create a general *umbrella ontology* $\mathcal{O}_*$. This will contain a set of constraining formulae, which we would expect to hold in any reasonable interpretation of the vocabulary under consideration. Thus, for each $\psi \in \mathcal{O}_*$ we have $\Box_* \psi \in \Theta$. This means that the standpoint of each ontology will inherit any penumbral connections that we consider to be essential, but may have been omitted from the ontology itself.

Inter-Ontology Concept Alignment and Interaction. We address the general and common case where the ontologies are poorly aligned both in that: (1) only some of the terms in each ontology have a closely corresponding counterpart in the other ontology; and (2) even where a concept in one ontology has a 'matching' counterpart in the other, these counterparts are not exactly equivalent. Such counterparts may differ either in terms of explicit definitions or axioms within the ontology itself, or in terms the criteria that have been used to determine the instances of a predicate (for instance a particular settlement might be classified as a Forest in one ontology but not in the other (e.g. because that ontology requires that a forest must have a certain minimal area), where it is instead considered to be an instance of Woodland. We assume that some form of alignment and matching of conceptual terms has already been carried out.

E13. $\forall x[\Box_{s_1} \mathsf{C}(x) \leftrightarrow \Box_{s_2} \mathsf{D}(x)]$ (equivalence)
E14. $(s' \preceq s) : (\lambda\overline{x}[\mathsf{C}(\overline{x})])$ (sharpening)
E15. $\forall x[\neg(\Box_{s_1} \mathsf{C}(x) \wedge \Box_{s_2} \mathsf{D}(x))]$ (exclusion)
E16. $\forall x[\Box_{s_1} \mathsf{C}(x) \rightarrow \Box_{s_2} \mathsf{D}(x)]$ (subsumption)
E17. $\exists x[\Box_{s_1} \mathsf{C}(x) \wedge \Box_{s_2} \mathsf{D}(x)]$ (strong overlap)
E18. $\exists x[\Diamond_{s_1} \mathsf{C}(x) \wedge \Diamond_{s_2} \mathsf{D}(x)]$ (weak overlap)

For present purposes we simply assume that certain entities can be identified between ontologies. In practice, identifying objects between different ontologies may be non-trivial and involve complex issues. These issues are related to the problems of establishing correspondences between entities at different possible worlds and/or precisifications.

8 Reasoning about Forests

As described in the *Motivation Example*, for the current purpose we set a scenario in which a stakeholder is using the proposed framework to analyse the conceptual variation between the forest concepts of the different data sources of an online platform (e.g. Global Forest Watch). In this section we show how the answers to the questions set in the *Motivating Example* section may be inferred.

Q1. Is area a consistently forest for all forest interpretations? Otherwise, is it arguably a forest?:
$\Box_*[\mathsf{Forest}(\mathsf{a})]$, $\quad \Diamond_*[\mathsf{Forest}(\mathsf{a})] \wedge \neg \Box_*[\mathsf{Forest}(\mathsf{a})]$
Q2. A degraded area a stopped being classed as Intact Forest[4] and is now only arguably a forest. What is it according to the other views?
$\forall s[\Box_s \neg\mathsf{Forest}(\mathsf{a}) \rightarrow \Diamond_s \mathsf{Landtype}(\mathsf{a})]$
Q3. What can we infer from an area that satisfies some but not all the interpretations?
Consider an example where we have three standpoints, s_1, s_2 and s_3 where $s_3 \preceq s_2$, two potential forests p and q and some facts:

E24. $\Box_{s_1} \mathsf{Forest}(\mathsf{p})$, $\Box_{s_2} \neg\mathsf{Forest}(\mathsf{p})$
E25. $\Box_{s_1}[\forall x[\mathsf{Forest}(x) \leftrightarrow (\boldsymbol{canopy}(x) > \mathsf{t\_can})]$, $\quad \Box_{s_1}[\mathsf{t\_can} > 30]$
E26. $\forall x[\Box_{s_2} \mathsf{Forest}(x) \leftrightarrow (\Box_{s_1} \mathsf{Forest}(x) \wedge \boldsymbol{av\_tree\_height}(x) > 2m)]$

[4] IFL are unfragmented areas with at least 50,000ha and 10Km width. These were then mapped from Landsat satellite imagery for the year 2000. [22]

E27. $\Box_*[\forall x \mathsf{Forest}(x) \leftrightarrow \neg\mathsf{Savannah}(x)]$
E28. $\Box_{s_3}$ $\mathsf{Savannah}(\mathsf{q})$, $\boldsymbol{canopy}(\mathsf{q}) = 35$

We can infer the following:

E29. $\boldsymbol{av\_tree\_height}(\mathsf{p}) < 2m$
E30. $\Box_{s_2} \neg\mathsf{Forest}(\mathsf{q})$
E31. $\neg\Box_{s_1} \mathsf{Forest}(\mathsf{q})$
E32. $\Box_{s_3}[\mathsf{t\_can} > 35]$

Q4. Are there inconsistencies among the data or only different data classifications? e.g. Are there areas reported to satisfy some criteria that don't fit the classifications? Hypothesis: temporal clearings?

E33. $\boldsymbol{canopy}(\mathsf{p}) = 25$

Consistency is ensured within precisifications and thus the system would show any inconsistent standpoint s to be $\emptyset$. In this case **E24**, **E25** and **E33** imply that $\Box_{s_1}[\mathsf{t\_can} < 25]$, which is inconsistent with **E25** in $\mathsf{s1}$.

9 Conclusions and Further Work

We have presented a formal language which we believe is well-suited for describing the variability of vague concepts and also for expressing relationships between different points of view regarding concept meaning in different contexts. Development of our system is ongoing, and many questions remain regarding its generality and practical applicability.

We have given both a semantics and a set of axioms for our formal language. In future work we will provide alternative Kripke style semantics with partially order operators, along the lines of [5, 2] and aim to establish completeness (possibly with addition of further axioms) via a Henkin-style proof that there is a model for every consistent set of formulae. However, given its complexity, we do not envisage our axiom set being used as practical inference mechanism. Rather, it would serve as a framework within which one could define more limited sublanguages, suitable for particular data interpretation tasks.

With regard to applications of the theory we plan to develop its use for combining non-aligned ontologies and to test this on specific examples, particularly concentrating on the forestry domain, which provides many interesting scenarios and challenges. Regarding developing the underlying semantics, there are some very interesting and tricky issues concerning correspondences between individuals identified with respect to different standpoints and/or possible worlds.

References

1. Achard, F.: Forest definition and extent. In: Vital Forest Graphics, pp. 6–9. UNEP/Earthprint (2009)
2. Allwein, G., Harrison, W.L.: Partially-ordered Modalities. Advances in Modal Logic (AiML) (2010)

3. Aristotle: On sophistical refutations [and] On coming-to-be and passing away. Loeb classical library, Greek authors (1955)
4. Atherton, C.: The Stoics on ambiguity. Cambridge classical studies (1993)
5. Beklemishev, L.D., Fernández-Duque, D., Joosten, J.J.: On Provability Logics with Linearly Ordered Modalities. Studia Logica **102**(3), 541–566 (6 2014)
6. Bennett, B.: Standpoint semantics: a framework for formalising the variable meaning of vague terms. In: Cintula, P., Fermüller, C., Godo, L., Hájek, P. (eds.) Understanding Vagueness — Logical, Philosophical and Linguistic Perspectives, pp. 261–278. Studies in Logic, College Publications (2011)
7. Böttcher, H., Herrmann, L., Herold, M., Romijn, E.: Independent Monitoring: Building trust and consensus around GHG data for increased accountability of mitigation in the land use sector (2018)
8. Buttigieg, P.L., Pafilis, E., Lewis, S.E., Schildhauer, M.P., Walls, R.L.: The environment ontology in 2016: bridging domains with increased scope, semantic density, and interoperation. Journal of Biomedical Semantics **7**(1), 57 (12 2016)
9. Carrara, M., Garbacz, P., Vermaas, P.E.: If engineering function is a family resemblance concept: assessing three formalization strategies. Applied Ontology **6**(2), 141–163 (2011)
10. Egré, P., de Gardelle, V., Ripley, D.: Vagueness and Order Effects in Color Categorization. Journal of Logic, Language and Information **22**(4), 391–420 (12 2013)
11. Fine, K.: Vagueness, truth and logic. Synthese **30**(3-4), 265–300 (1975)
12. Frege, F.L.G.: On Sense and Reference. In: Translations from the Philosophical Writings of Gottlob Frege (1970)
13. Grainger, A.: Difficulties in tracking the long-term global trend in tropical forest area. Proceedings of the NAS of the USA **105**(2), 818–823 (2008)
14. Grice, H.: Logic and Conversation. Unpublished manuscript (1975)
15. Hansen, M.C., Al, E.: High-Resolution Global Maps of 21st-Century Forest Cover Change. Science **850**(27), 123–134 (2013)
16. Kamp, H.: Two theories about adjectives. In: Keenan, E. (ed.) Formal Semantics of Natural Language. Cambridge University Press, Cambridge, England (1975)
17. Kamp, H., Partee, B.: Prototype theory and compositionality. Cognition **57**(2), 129–191 (1995)
18. Keefe, R., Smith, P.: Vagueness: a Reader. MIT Press (1996)
19. Lund, H.G.: Definitions of Forest, Deforestation, Afforestation and Reforestation. Forest Information Services pp. 1–159 (2007)
20. Mehlberg, H.: The Reach of Science. University of Toronto Press (1958), extract on Truth and Vagueness, pp. 427-55, reprinted in [18]
21. Piantadosi, S.T., Tily, H., Gibson, E.: The communicative function of ambiguity in language. Cognition **122**(3), 280–291 (3 2012)
22. Potapov, P., Hansen, M.C., Laestadius, L., Turubanova, S., Yaroshenko, A., Thies, C., Smith, W., Zhuravleva, I., Komarova, A., Minnemeyer, S., Esipova, E.: The last frontiers of wilderness: Tracking loss of intact forest landscapes from 2000 to 2013. Science Advances **3**(1), e1600821 (1 2017)
23. Tarski, A.: Logic, semantics, metamathematics: papers from 1923 to 1938. Hackett Publishing (1980)
24. Wittgenstein, L.: Philosophical investigations (2009)
25. Zadeh, L.: Fuzzy sets. Information and Control **8**(3), 338–353 (1965)
26. Zadeh, L.: Fuzzy Logic and Approximate Reasoning. D. Reidel Publishing Company **30**, 407–428 (1975)
27. Zipf, G.K.: Human Behavior and the Principle of Least Effort. Addisson-Wesley Press, Cambridge (1949)