
Standpoint Semantics: A Framework for Formalising the Variable Meaning of Vague Terms

BRANDON BENNETT¹

Overview

The paper develops a formal model for interpreting vague languages in a setting similar to that of *supervaluation* semantics. Two modes of semantic variability are modelled, corresponding to different aspects of vagueness: one mode arises where there can be multiple definitions of a term involving different attributes or different logical combinations of attributes. The other relates to the threshold of applicability of a vague term with respect to the magnitude of relevant observable values.

The truth of a proposition depends on both the possible world and the *precisification* with respect to which it is evaluated. Structures representing both possible worlds and precisifications are specified in terms of primitive functions representing observable measurements, so that the semantics is grounded upon an underlying theory of physical reality. On the basis of this semantics, the acceptability of a proposition to an agent is characterised in terms of a combination of the agent's beliefs about the world and their attitude to admissible interpretations of vague predicates.

1 Introduction

The terminology of natural language is highly affected by vagueness. Except in specialised circumstances, there are no generally agreed criteria that precisely determine the applicability of our conceptual vocabulary to describing the world. This presents a considerable problem for the construction of a knowledge representation language that is intended to articulate information of a similar kind to that conveyed by natural language communication.

The fundamental idea of the *supervaluationist* account of vagueness, is that a language containing vague predicates can be interpreted in many different ways, each of which can be modelled in terms of a precise version of the language, which is referred to as a *precisification*. If a classical semantics is used to give a denotational valuation of expressions for each of these precise versions, the interpretation of the vague language itself is given by a *supervaluation*, which is determined by the collection of these classical valuations.

¹Partial support of the *Co-Friend* project (FP7-ICT-214975, www.cofriend.net) and the EU Framework 7 is gratefully acknowledged. The paper has been enhanced by suggestions made in the detailed comments of two anonymous reviewers.

The view that vagueness can be analysed in terms of multiple senses was proposed by Mehlberg [16], and a formal semantics based on a multiplicity of classical interpretations was used by van Fraassen [18] to explain ‘the logic of presupposition’. It was subsequently applied to the analysis of vagueness by Fine [8], and thereafter has been one of the more popular approaches to the semantics of vagueness adopted by philosophers and logicians. In the current paper we introduce a logic based on the essential idea of supervaluation semantics, but different in several respects from previous systems.

A major strength of the supervaluation approach is that it enables the expressive and inferential power of classical logic to be retained (albeit within the context somewhat more elaborate semantics) despite the presence of vagueness. In particular, necessary logical relationships among vague concepts can be specified using classical axioms and definitions. These analytic interdependencies (referred to by Fine as *penumbral* connections) will be preserved, even though the criteria of correspondence between concepts and the world are ill-defined and fluid.

Investigation of supervaluation semantics in the philosophical literature tends, as one might expect, to be drawn towards subtle foundational questions, such as those concerning the *sorites* paradox and second-order vagueness. By contrast, the purpose of the current paper is to flesh out the details of a particular variant of supervaluation semantics and to develop an expressive formal representation language that could be employed within information processing applications.

The development of the supervaluation idea in the current paper also departs somewhat from that proposed by Fine. In Fine’s theory, precisifications vary in their level of precision, so that one precisification may be a more precise version of another. This gives rise to a partial order on precisifications. Fine then proposes a semantics that takes account of this ordering and defines a notion of *super-truth* in terms of the precisification structure as a whole: super-truth corresponds to truth at all maximally precise and *admissible* precisifications (where ‘admissible’ means that a precisification is considered a reasonable interpretation of the language and is taken as a primitive notion). Moreover, Fine suggests that ‘truth’ in a vague language may be identified with this notion of super-truth.

By contrast, in the current paper, we take each precisification to be a maximally precise version of the language. And we consider truth primarily as a property of propositions that is relative to a particular precisification, rather than determined by the whole set of possible precisifications. However, we will also introduce the notion of a proposition holding relative to a *standpoint*, which is associated with a set of precisifications considered acceptable by some agent. Formally, this notion is somewhat akin to ‘super-truth’, except that instead of assuming a fixed set of admissible precisifications, we consider the set of admissible precisifications to be determined relative to a particular agent in a particular situation.

Like supervaluation semantics, standpoint semantics may be regarded as a rival to the *fuzzy logic* approach to semantic indeterminacy [19]. Whereas fuzzy logic explores this indeterminacy in terms of degrees of truth and non-classical truth functions, standpoint semantics focuses on truth conditions rather than truth values, and employs a notion of truth that is close to the classical view, although relativised to account for a variety of possible interpretations. Nevertheless, we believe that strong correspondences between standpoint semantics and fuzzy logic can be established. By introducing a

probability distribution over the space of precisifications, degrees of *acceptability* can be introduced, and it can be shown that the acceptability of vague conjunctions is governed by certain of the T -norms commonly used in fuzzy logic. Results in this area are beyond the scope of the present work, where we focus on the core model theory and employ a much simpler model of an agent's attitude to precisifications.

The formalism developed in the current paper takes ideas from previous theories proposed by Bennett [2, 4] and Halpern [9] and elaborates material presented in [5]. Halpern's paper analyses vagueness in terms of the subjective reports of multiple agents, but these play a similar role in his semantics to precisifications in the semantics proposed in this paper. Our approach also has some commonality with that of [13] and [14].

In [4] the semantics of vague adjectives is characterised in terms of their dependence on *relevant objective observables* (e.g. 'tall' is dependent on 'height'). One of the primary aims of the current paper is to provide a rigorous foundation for the notion of precisification, in which the interpretation associated with a precisification is explicitly defined in terms of choices made in imposing distinctions with regard to continuously varying properties manifest in possible states of the world. Bennett [2] proposed a two-dimensional model theory, in which the interpretations of propositions are indexed both by precisifications and possible worlds. Whereas a somewhat *ad hoc* relation of relevance between between vague predicates and observables was introduced in [4], the current paper makes a much more specific connection, in which thresholds occur explicitly in definitions of vague predicates. A concrete example of the use of this approach in an implemented computer system for processing geographic information can be found in [17] and [6].

The structure of the paper is as follows: in the next section we give an overview of the formal theory that will be developed, and consider some examples illustrating different kinds of vagueness. In Section 3 we specify a formal language that makes explicit the structure of both possible worlds and precisifications in terms of possible values of observable measurements. Section 5 gives a formal model of an agent's *standpoint* with respect to possible worlds that the agent considers plausible and precisifications that the agent considers admissible. We end with a consideration of further work and conclusions.

2 Preliminaries

Before getting into the details of our formal language and its semantics, we first clarify some aspects of our approach.

2.1 Comparison classes

To avoid confusion we briefly consider a phenomenon that is often associated with vagueness but will not be considered in the current paper. This is the relativity of the interpretation of vague adjectives to a given *comparison class*. For instance, when describing an object or creature as tall, we make a judgement based on the height of that object or creature. But in judging that a woman is tall we employ a different threshold of tallness from when describing a giraffe as tall. As in this example, the relevant comparison class is often determined by the count noun used to refer to the object, but it may sometimes be determined by the particular set of objects present in a given situation.

However, comparison class relativity is a side-issue that is not essential to vagueness itself. Even if we restrict attention to a definite class of individuals (say, adult males in Belgium) the adjective ‘tall’ is still vague. Similar remarks could be made about more general issues of context variability of the interpretation of terminology. If required, an explicit model of comparison class dependency of vague adjectives (perhaps similar to that given in [4]) could be incorporated into an extended version of our theory.

2.2 Distinguishing conceptual and sorites vagueness

An important feature of the proposed theory is that it makes a clear distinction between two forms of vagueness.

One type of vagueness arises where there is ambiguity with regard to which attributes or conditions are essential to the meaning of a given term, so that it is controversial how it should be defined. We call this *conceptual vagueness* (or ‘deep ambiguity’). A good example of conceptual vagueness is the concept of *murder*. Although in most cases there will be general agreement as to whether a given act constitutes murder, the precise definition is subtle and controversial. Judicial systems vary as to the stipulations they make to characterise the crime of murder. Thus one may debate whether murder requires malice or intent, whether the murderer must be sane, whether the victim must be unwilling etc. Moreover, even where conditions are stipulated in great detail, cases may arise that defy simple judgement.

A somewhat different kind of vagueness occurs when the criteria for applicability of a term depend on placing a threshold on the *required magnitude* of one or more variable attributes. For instance, we may agree that the appropriateness of ascribing the predicate ‘tall’ to an individual depends on the height of that individual, but there is no definite height threshold that determines when the predicate is applicable. We refer to this as *sorites vagueness*, since the essence of the sorites paradox is the indeterminacy in the number of grains required to make a heap.

So to summarise: in the case of *conceptual vagueness* there is indeterminism regarding which property or logical combination of properties is relevant to determining whether a concept is applicable, whereas with *sorites vagueness* the relevant properties are clear, but the degree to which these properties must be present is indefinite.²

2.2.1 Combined modes of vagueness

It should be emphasised that the two kinds of vagueness I have identified are not exclusive—a single word or phrase may be, and often is, imbued with both conceptual and sorites vagueness.

For example, [11] considers the conditions under which one might describe a person as ‘clever’. Here it is not clear what parameter or parameters are relevant to the attribution of cleverness. As Kamp suggests, quick wittedness and problem solving ability are both indications of cleverness, although one person might be considered more quick witted than another and yet less capable of problem solving.

Even the adjective ‘tall’, which is almost a paradigm case of sorites vagueness, is also to some extent affected by conceptual vagueness. This is because there is no universally agreed specification of exactly how a person’s height should be measured.

²This distinction was identified and analysed in [3], but no formal semantics was presented.

Perhaps we can agree that shoes should not count towards height and that hair should also be excluded; but what about a wart on the top of a person's head? What if a person stretches their neck or hunches their shoulders? Of course such factors rarely impinge on actual uses of the word 'tall', and if we really wanted an objective measure of height we could legislate that a particular measurement regime be used. Nevertheless, the very fact that one might need to carry out such legislation in order to get a reliable objective measurement of tallness, demonstrates that the natural language adjective 'tall' is subject to conceptual ambiguity.

The adjective 'bald' and the count noun 'heap', which are also ubiquitous in discussions of the sorites paradox clearly suffer from considerable conceptual ambiguity as well as threshold indeterminacy. Baldness is not only judged by the number of hairs on a head, but also where they are located—it is a matter of coverage, not only numerical quantity. Likewise, a heap is not just any collection of grains. Intuitively, heap-hood requires particular structural properties; but the exact nature of these is difficult to pin down.

Despite their close connection, there are significant differences in the type of semantic variability involved in the two kinds of vagueness—to reiterate: conceptual vagueness is indeterminacy in the attribute or combination of attributes that must be present, whereas sorites vagueness is indeterminacy in the degree to which a continuously varying attribute (or attributes) must be present. Hence, we believe that a semantics for vague languages is most clearly specified by separating the two modes.

2.3 Predication, observables and thresholds

Our semantics explicitly models the applicability of vague predicates in terms of thresholds applied to relevant observable measurements. In the simplest case we assume that our judgement of whether a predicate ϕ applies to object x depends only on the value of a single measurement $f(x)$ —the higher the value of $f(x)$, the more we are inclined to judge that $\phi(x)$ is true. Let $\tau(\phi)$ denote some reasonable threshold that we might set for the applicability of ϕ . Then $\phi(x)$ is judged to be true if $f(x)$ is greater than $\tau(\phi)$ and false if $f(x)$ is less than $\tau(\phi)$.

The case where we have $f(x) = \tau(\phi)$ presents a technical issue, in that there is no obvious basis to decide between assigning truth or falsity. This will be avoided by restricting the semantics so that a threshold value may not have the same value as any observable function.

2.4 Relating precisifications to cognitive attitudes

As well as relating precisifications to states of the world, we also model their relationship to the cognitive states of agents. We give an account of an agent's attitude to vague propositions in terms of a formalised notion of *standpoint*, which describes the agents belief state as well as the range of interpretations of vague terminology that they consider admissible. A *standpoint* will be modelled by a structure $\langle B, A, \Psi \rangle$, where: B is the set of possible worlds compatible with the agent's beliefs; A is the set of precisifications that are acceptable to the agent; and Ψ is a set of definitional theories that specify different ways in which the meaning of vague predicates can be represented in terms of some logical combination of threshold constraints. Hence, A models an agent's stand-

point with respect to sorites vagueness, while Ψ models the standpoint in relation to conceptual vagueness.

As is usual in supervaluation-based approaches, we assume that when describing a particular situation or appraising a given set of propositions, a language user employs a choice of thresholds that is consistent across usages of all concepts. Thus, where two or more concepts have some semantic inter-dependence, this will be maintained by consistent usage of thresholds. For example ‘tall’ and ‘short’ are (in a particular context) mutually exclusive and are dependent on the objective observable of height. Thus the height threshold above which a person to be considered ‘tall’ must be greater than the height threshold below which a person is considered ‘short’. However, we do not assume that an agent’s point of view is determined by a single precisification but rather by a set of accepted thresholds (or, as in our further work, by a probability distribution over the set of all precisifications).

3 A language of precise observables and vague predicates

In this section we define a general-purpose formal language that (despite bearing only a coarse correspondence to the structure and meaning of natural language) is intended to exhibit some fundamental principles that govern the phenomenon of vagueness.

A key idea underlying the construction of this formalism is that the language should contain two types of vocabulary:

- A precise vocabulary for describing the results of precise objective measurements of the state of the world.
- A vague vocabulary which is defined in terms of the precise vocabulary, relative to a valuation of certain *threshold parameters*, which may occur in the definitions.

3.1 Measurement structures

At the base of the semantics is a structure that represents the state of a *possible world* in terms of a valuation of *measurement functions*, which specify the results of observations applied to the entities of some domain.³

An *n-ary measurement function* over domain D is a function $\mu : D^n \rightarrow \mathbb{Q}$, with \mathbb{Q} being the set of rational numbers. Thus, \mathbb{Q}^{D^n} is the set of all *n-ary measurement functions* and $\bigcup_{n \in \mathbb{N}} \mathbb{Q}^{D^n}$ is the set of all measurement functions of any arity (with domain D).

A *measurement structure* is a tuple $\langle D, M, v_M, w \rangle$, where:

- D is a domain of entities;
- $M = \{\dots, f_i, \dots\}$ is a set of measurement function symbols;
- $v_M : M \rightarrow \mathbb{N}$, is a mapping from the symbols in M to the natural numbers, giving the arity of each function;
- $w : M \rightarrow \bigcup_{n \in \mathbb{N}} \mathbb{Q}^{D^n}$, such that if $v_M(f) = m$ then $w(f) \in \mathbb{Q}^{D^m}$, is a function mapping each *n-ary function symbol* to a measurement function from D^n to \mathbb{Q} .

³Credit is given to one of the anonymous reviewers for suggestions leading to a more precise formulation of measurement structures than had been given in the originally submitted version of this paper.

Since each assignment function w characterises the domain and function symbols, as well as determining a valuation of the measurement functions over the domain, we regard each w as representing a *possible world*. A possible world in this sense is an *arbitrary* valuation of the function symbols over the domain. The valuation need not respect physical laws with regard to possible combinations of measurable values, so, in so far as the observable functions are intended to correspond to actual kinds of measurements, such worlds could be physically impossible.

Given a domain D , a set of measurement function symbols M , and an arity specification function v_M , the set of worlds that can be specified in terms of these elements can be defined by:

$$\text{Worlds}(D, M, v_M) = \{w \mid \langle D, M, v_M, w \rangle \text{ is a measurement structure}\}.$$

This definition assumes that we have the same set of entities present in every possible world—i.e. we have *constant domains*. There are strong arguments that this condition is unrealistic for a domain of real physical objects. However, this issue is complex and tangential to the main concerns of this paper, and will not be addressed here.

A *measurement frame* is a structure that specifies all possible worlds determined by a given measurement structure:

$$\langle D, M, v_M, W \rangle,$$

where $W = \text{Worlds}(D, M, v_M)$.

3.2 A language of measurements and thresholds

Let us now consider the definition of a predicative language that can be interpreted relative to a measurement structure.

Let $\mathcal{L}(M, v_M, T, V)$ be the set of formulae of a first-order logical language⁴ whose non-logical symbols consist of: a finite set of measurement function symbols $M = \{f_1, \dots, f_k\}$, a finite set, $T = \{t_1, \dots, t_l\}$, of *threshold parameter* symbols, strict and non-strict inequality relations ($<$ and \leq), and a denumerable set $V = \{\dots, x_i, \dots\}$ of variable symbols. Every atomic formula of $\mathcal{L}(M, v_M, T, V)$ takes one of the forms:

A1. $f_j(x_1, \dots, x_n) \leq f_k(y_1, \dots, y_m)$

A2. $t_i \leq t_j$

A3. $t_i < f_j(x_1, \dots, x_n)$

A4. $f_j(x_1, \dots, x_n) < t_i$

where $n = v_M(f_j)$ and $m = v_M(f_k)$. Nested measurement functions are not allowed, since they operate on entities and their values are real numbers. Complex formulae are formed from atomic formulae by means of the standard truth-functional connectives and quantifiers over the variables (but not over the threshold parameters). $\mathcal{L}(M, v_M, T, V)$ includes formulae with free variables.

⁴In fact, this account does not depend on the specific details the language. I choose standard first-order logic for definiteness, but there could be reasons to use a more expressive language.

A realistic theory of observable measurements would model the fact that only certain measurements can be usefully compared. This could be achieved by classifying the observable functions into sorts according to the type of quantity that they designate and then restricting the syntax so that only observable values of the same type can be compared by means of the \leq relation, thus restricting the syntax of atoms of the form **A1**. If the observables correspond to the kinds of measurement used in Newtonian physics, then the sorts of observable measurement could be classified in terms the fundamental dimensions of *length* (L), *time* (T) and *mass* (M), and also combinations of these basic dimensions formed by products and reciprocals (for example *area* has type $L \times L$ and *velocity* has type L/T). But the classification and comparability of measurement types is tangential to our primary aim of formalising modes of vagueness; so, for present purposes, we make the simplifying assumption that all measurements are comparable.

Although the syntactic specification of our language $\mathcal{L}(M, v_M, T, V)$ allows arbitrary logical combinations of atoms, the different forms of atoms express different kinds of information, which one would not normally mix within the same proposition. In applying the language to formalising the semantics and use of particular observables, one would typically employ formulae of the following forms, which are homogeneous with respect to the kinds of atoms they contain:

Constraints on observables Formulae containing only atoms of the form **A1** (i.e. those that do not contain threshold parameters) can be regarded as expressing constraints on the physical structure of the world. However, the language considered here is too limited to express a fully-fledged physical theory. To specify such a theory we would need operators designating mathematical functions that govern the relationships between observables (for instance $\forall x[density(x) = mass(x)/volume(x)]$). This would require the language to be extended with an appropriate vocabulary of mathematical functions to operate upon and combine values of the observables.

Threshold constraints Formulae of the form **A2** express ordering constraints between thresholds. A typical example would be $t_short \leq t_tall$, stating that the threshold below which an individual is considered short is less than the threshold above which an individual is considered tall. Atoms of the form **A2** will not normally occur within more complex formulae. In applying our representation to describing the semantics of vague predicates, we have found many cases where it seems appropriate that a strict ordering be imposed between two different but related thresholds. In certain cases more complex ordering constraints may be required (such as in the specification for the colour purple, given below in Section 4).

Judgements A third class comprises those formulae containing atoms of forms **A3** and **A4**, by which the value of an observable measurement function is compared to a threshold parameter. These are perhaps the most significant type of formulae in the language as they play a crucial role in our account of vagueness. Simple examples include $t_tall < height(x)$ and $weight(x) < t_heavy$. These express judgements that a given measurable property of an object lies above or below a given threshold. In the next section we shall see how these thresholds are linked to vague predicates.

3.3 Predicate definitions

Each formula of $\mathcal{L}(M, v_M, T, V)$ defines a predicate of arity n , where n is the number of free variables in the formula. Hence, we can extend L by defining new predicate symbols by means of formulae of the form

$$\mathbf{PG.} \quad \forall x_1, \dots, x_n [R(x_1, \dots, x_n) \leftrightarrow \Phi(t_1, \dots, t_m, x_1, \dots, x_n)],$$

where $\Phi(t_1, \dots, t_m, x_1, \dots, x_n)$ is any formula in $\mathcal{L}(M, v_M, T, V)$ incorporating parameters t_1, \dots, t_m and with free variables x_1, \dots, x_n .

For instance, $\forall x [\text{Tall}(x) \leftrightarrow (t_{\text{tall}} \leq \text{height}(x))]$ is a typical example of a predicate defined in this way. Here, height is a measurement function and t_{tall} is a threshold parameter. An informal interpretation of this formula is that an entity is tall just in case its height is greater than or equal to the value of the parameter t_{tall} .

3.4 An extended language including defined predicates

The language $\mathcal{L}(M, v_M, T, V)$ of measurements and thresholds will serve as a basis for an extended language incorporating symbols for vague predicates.

Let $\mathcal{L}(M, v_M, T, V, R, v_R, N)$ be the language obtained by supplementing the vocabulary of $\mathcal{L}(M, v_M, T, V)$ with a set of predicate symbols R , such that the arity of each symbol is given by the function $v_R: R \rightarrow \mathbb{N}$ and a set of constant symbols N (which will denote objects of the domain). The set of atomic formulae is extended to include those of the form

$$\mathbf{A5.} \quad R_i(\alpha_1, \dots, \alpha_n),$$

where each $\alpha_i \in (V \cup N)$. The complete set of formulae of $\mathcal{L}(M, v_M, T, V, R, v_R, N)$ includes all formulae constructed from this extended set of atomic formulae by application of Boolean connectives and quantification over variables in V .

3.5 Predicate grounding theories

Given a language $\mathcal{L}(M, v_M, T, V, R, v_R, N)$, a *predicate grounding theory* for this language is a set of formulae of the form **PG**, containing one formula for each relation $R_i \in R$. Thus, the predicate grounding theory *defines* every relation in R in terms of a formula of the sub-language $\mathcal{L}(M, v_M, T, V)$.

Let Θ be the set of all predicate grounding theories for $\mathcal{L}(M, v_M, T, V, R, v_R, N)$. Since each of these grounding theories includes a definition of every predicate in the language, we can define a function,

$$\text{Def: } \Theta \times R \rightarrow \mathcal{L}(M, v_M, T, V),$$

such that $\text{Def}(\theta, R)$ is a formula with $v_R(R)$ free variables, which gives a possible definition of the relation R .

3.6 Parameterised precisification models

We now define a model structure to provide a semantics for $\mathcal{L}(M, v_M, T, V, R, v_R, N)$. The model incorporates a measurement frame together with mappings from the language

symbols onto elements of the frame. Specifically, a *parameterised precisification model* is a structure

$$\mathfrak{M} = \langle \mathcal{M}, R, v_R, N, V, T, \Theta, \kappa, \xi, P \rangle,$$

where

- $\mathcal{M} = \langle D, M, v_M, W \rangle$ is a measurement frame;
- R is a set of predicate symbols;
- $v_R : R \rightarrow \mathbb{N}$ gives the arity of each predicate symbol;
- N is a set $\{\dots, n_i, \dots\}$ of nominal constants;
- V is a set $\{\dots, x_i, \dots\}$ of variable symbols;
- T is a finite set $\{\dots, t_i, \dots\}$ of threshold parameter symbols;
- $\Theta = \{\dots, \theta_i, \dots\}$, where each θ_i is a predicate grounding theory for the language;
- $\kappa : N \rightarrow D$ maps nominal constants to entities of the domain;
- $\xi : V \rightarrow D$ maps variable symbols to entities of the domain;
- $P = \{p \mid p : T \rightarrow (\mathbb{R} \setminus \mathbb{Q})\}$, is the set of all mappings from threshold parameters to *irrational* numbers. (P is the set of precisifications.)

The assignment of irrational numbers to threshold parameters is primarily a technical means to ensure that every observable value is either greater or smaller than any threshold parameter. However, it can be motivated by regarding the values of thresholds as *cuts* (in roughly the same sense as in *Dedekind cut*) between two sets of rational numbers, where all values in one set are strictly lower than all values in the other set. (On the other hand, it may be preferable to specify the domains of observables and thresholds as constituting disjoint sub-domains of the rationals.)

3.7 Interpretation function

The semantic interpretation function, $[\chi]_{\mathfrak{M}}^{w,p,\theta}$, gives the denotation of any formula or term χ of the language relative to a given model \mathfrak{M} , a possible world $w \in W$, a predicate grounding theory $\theta \in \Theta$ and a precisification $p \in P$. The θ index models semantic indeterminacy arising from conceptual vagueness, whereas the p index models indeterminacy due to sorites vagueness.

To specify the interpretation function, the following auxiliary notations will be used:

- $\mathfrak{M} \stackrel{x}{\sim} \mathfrak{M}'$ means that models \mathfrak{M} and \mathfrak{M}' are identical except for their variable assignment functions ξ and ξ' . And moreover, these assignment functions are identical, except that they may differ in the value assigned to the variable x .
- $\text{Subst}([x_1 \Rightarrow \alpha_1, \dots, x_n \Rightarrow \alpha_n], \phi)$ refers to the formula resulting from ϕ after replacing each variable x_i by α_i .

The interpretation function can now be specified. The Boolean connectives and quantifier have their standard classical interpretation:

- $\llbracket \neg\phi \rrbracket_{\mathfrak{M}}^{w,p,\theta} = \mathbf{t}$ if $\llbracket \phi \rrbracket_{\mathfrak{M}}^{w,p,\theta} = \mathbf{f}$, otherwise = \mathbf{f} ;
- $\llbracket \phi \wedge \psi \rrbracket_{\mathfrak{M}}^{w,p,\theta} = \mathbf{t}$ if $\llbracket \phi \rrbracket_{\mathfrak{M}}^{w,p,\theta} = \mathbf{t}$ and $\llbracket \psi \rrbracket_{\mathfrak{M}}^{w,p,\theta} = \mathbf{t}$, otherwise = \mathbf{f} ;
- $\llbracket \forall x[\psi] \rrbracket_{\mathfrak{M}}^{w,p,\theta} = \mathbf{t}$ if $\llbracket \psi \rrbracket_{\mathfrak{M}'}^{w,p,\theta}$ for all \mathfrak{M}' such that $\mathfrak{M} \stackrel{x}{\sim} \mathfrak{M}'$, otherwise = \mathbf{f} .

The inequality relations are interpreted as follows, where γ_i and γ_j may each be either a threshold parameter or a measurement function term:

- $\llbracket \gamma_i \leq \gamma_j \rrbracket_{\mathfrak{M}}^{w,p,\theta} = \mathbf{t}$ if $\llbracket \gamma_i \rrbracket_{\mathfrak{M}}^{w,p,\theta}$ is less than or equal to $\llbracket \gamma_j \rrbracket_{\mathfrak{M}}^{w,p,\theta}$, otherwise = \mathbf{f} ;
- $\llbracket \gamma_i < \gamma_j \rrbracket_{\mathfrak{M}}^{w,p,\theta} = \mathbf{t}$ if $\llbracket \gamma_i \rrbracket_{\mathfrak{M}}^{w,p,\theta}$ is strictly less than $\llbracket \gamma_j \rrbracket_{\mathfrak{M}}^{w,p,\theta}$, otherwise = \mathbf{f} .

The value of measurement functions depends on the possible world in which the measurement is made; and hence, their interpretation depends on the w index:

- $\llbracket f(\alpha_1, \dots, \alpha_n) \rrbracket_{\mathfrak{M}}^{w,p,\theta} = w(f)(\langle \delta(\alpha_1), \dots, \delta(\alpha_n) \rangle)$,
where $\delta(\alpha) = \kappa(\alpha)$ if $\alpha \in N$ and $\delta(\alpha) = \xi(\alpha)$ if $\alpha \in V$.

Interpretation of the threshold parameters depends on the precisification index, p :

- $\llbracket t \rrbracket_{\mathfrak{M}}^{w,p,\theta} = p(t)$;

Finally, the interpretation of the defined predicate and relation symbols is dependent upon the grounding theory θ :

- $\llbracket R(\alpha_1, \dots, \alpha_n) \rrbracket_{\mathfrak{M}}^{w,p,\theta} = \llbracket \text{Subst}([x_1 \Rightarrow \alpha_1, \dots, x_n \Rightarrow \alpha_n], \text{Def}(\theta, R)) \rrbracket_{\mathfrak{M}}^{w,p,\theta}$.

On the basis of the interpretation function, a semantic satisfaction relation can be defined by

$$\mathfrak{M}, \langle w, p, \theta \rangle \Vdash \phi \quad \text{iff} \quad \llbracket \phi \rrbracket_{\mathfrak{M}}^{w,p,\theta} = \mathbf{t}.$$

This says that formula ϕ is true in model \mathfrak{M} , at world w and precisification p , with respect to predicate grounding theory θ .

The *interpretation set* of a proposition relative to a model \mathfrak{M} is given by:

$$\llbracket \phi \rrbracket_{\mathfrak{M}} = \{ \langle w, p, \theta \rangle \mid (\mathfrak{M}, \langle w, p, \theta \rangle \Vdash \phi) \}.$$

This is the set of world/precisification/grounding theory triples for which formula ϕ is evaluated as true.

We have now established the main result of this paper: we have defined a first-order language with a semantics that gives a special status to observable measurements and threshold parameters. The interpretation function for this language is such that each valuation of observable measurements corresponds to a possible world, and each valuation of threshold parameters corresponds to a precisification. In terms of this semantics, each propositional formula is interpreted as the set of possible world/precisification/grounding theory triples at which the proposition is considered to be true.

4 Penumbral connections

A major selling point of supervaluation-based accounts of vagueness is that they provide a framework within which one can model dependencies among vague predicates—i.e. *penumbral connections*. To use an example of Fine, there is a vague borderline demarcating the applicability of the terms ‘pink’ and ‘red’ but the terms are exclusive in that one would not normally describe an object as both pink and red. A primary motivation for the development of standpoint semantics was to refine and make more explicit the nature of such interdependencies.

The phenomenon of penumbral connection is of course controversial and is one of the key points of contention between supervaluationist approaches and fuzzy logics. Fuzzy logicians tend to the view that an object can be both pink and red to some degree, and consequently a proposition of the form $(\text{Pink}(x) \wedge \text{Red}(x))$ may also be true to some degree. By contrast, a supervaluationist would say that an object may be pink according to one precisification and red according to another but there is no precisification according to which the object is both pink and red, and hence the proposition $(\text{Pink}(x) \wedge \text{Red}(x))$ must be false (indeed *super-false*). In the present account, we lean more towards the supervaluationist account, although we do have leeway to accommodate some facets of the fuzzy viewpoint. Since our interpretation function evaluates propositions with respect to a predicate grounding theory (as well as a precisification) we may weaken dependencies between predicates by allowing that some grounding theories do not enforce them. But in the present paper we do not consider this possibility in detail.

Within the language $\mathcal{L}(M, v_M, T, V, R, v_R, N)$, penumbral connections are made explicit by predicate grounding definitions, and also by specifying ordering constraints between threshold parameters. We conjecture that this approach, is adequate to describe most, if not all, dependencies between vague predicates. In the case of predicates ‘tall’ and ‘short’, which have provided most of our examples, so far, their penumbral connection is straightforwardly represented by the definitions $\forall x[\text{Tall}(x) \leftrightarrow (\text{t_tall} \leq \text{height}(x))]$ and $\forall x[\text{Short}(x) \leftrightarrow (\text{height}(x) \leq \text{t_short})]$ and the threshold constraint $\text{t_short} < \text{t_tall}$.

As a more tricky example, we consider the vague colour terms *red*, *orange*, *pink* and *peach*. We may describe this four-fold categorisation in terms of divisions based on two threshold parameters, one concerning the observed *hue* of an object and the other concerning the *saturation* of the object’s observed colour (low saturation being characteristic of *pastel* colours such as pink and peach). However, a problem arises concerning the hue of a colour. Hue is normally measured in terms of a cyclical scale which runs from red, through the rainbow to violet and then back to red. This measure corresponds well with our perceptual experience of colour, in that we perceive colours as if they form a circle in which there is a continuous transition from blue through violet to red. In order to circumvent this we can adopt a measurement of hue which forms a linear scale based on the physical frequency spectrum of light. Thus, $\text{hue}(x)$ would give the value corresponding to the frequency within the visible spectrum that is most strongly reflected by object x . Using this measurement of hue, the colour predicates may be defined as follows:

$$\text{Red}(x) \leftrightarrow ((\text{hue}(x) < \text{t_red-orange}) \wedge (\text{t_pastel} < \text{saturation}(x)))$$

$$\begin{aligned} \text{Orange}(x) \leftrightarrow (& (\text{t_red-orange} < \text{hue}(x)) \wedge (\text{hue}(x) < \text{t_orange-yellow}) \\ & \wedge (\text{t_pastel} < \text{saturation}(x))) \end{aligned}$$

$$\text{Pink}(x) \leftrightarrow ((\text{hue}(x) < \text{t_red-orange}) \wedge (\text{saturation}(x) < \text{t_pastel}))$$

$$\text{Peach}(x) \leftrightarrow ((\text{t_red-orange} < \text{hue}(x)) \wedge (\text{hue}(x) < \text{t_orange-yellow}) \\ \wedge (\text{saturation}(x) < \text{t_pastel}))$$

And the threshold constraint ($\text{t_red-orange} < \text{t_orange-yellow}$) should also be specified.

This analysis faces a further complication if we consider a colour such as purple, where the boundary between a purple hue and a red hue could be interpreted as lying either near the high end of the scale or at the low end of the scale (if some reddish hues are regarded as purple). To account for this we must allow that the blue-purple threshold boundary can be either lower or higher than the purple-red boundary. In the first case the hue of a purple object must lie both above the blue-purple threshold *and* below the purple-red threshold. But in the second case (where the range of purples is regarded as wrapping round from the high end to the low end of the hue scale) an object will count as purple if its hue is *either* higher than the blue-purple threshold *or* lower than the purple-red threshold. Thus, we would get the following grounding definition (with no additional constraint on the ordering of t_blue-purple and t_purple-red):

$$\begin{aligned} \text{Purple}(x) \leftrightarrow & (((\text{t_blue-purple} < \text{t_purple-red}) \wedge \\ & (\text{t_blue-purple} < \text{hue}(x)) \wedge (\text{hue}(x) < \text{t_purple-red})) \\ & \vee \\ & ((\text{t_purple-red} < \text{t_blue-purple}) \wedge \\ & ((\text{t_blue-purple} < \text{hue}(x)) \vee (\text{hue}(x) < \text{t_purple-red})) \\ &)) \\ & \wedge (\text{t_pastel} < \text{saturation}(x)) \end{aligned}$$

5 Standpoints and proposition evaluation

In order to take account of an agent's comprehension of propositional information, we need to relate the agent's cognitive state to our formal semantics of propositions, which gives the meaning of a proposition in terms of an interpretation set. Two aspects of the cognitive state are clearly relevant: what the agent believes about the state of the world, and what the agent regards as an acceptable usage of terminology (especially vague predicates). Whether the agent considers a proposition to be true will depend on both these aspects.

Beliefs may be modelled either syntactically in terms of formulae expressing facts and theories that an agent regards as true, or semantically in terms of possible states of the world that an agent considers plausible. In the framework of classical logic, where each predicate has a definite meaning, the two perspectives are tightly linked, since any set of formulae determines a fixed set of possible worlds that satisfy that set. But if propositions can vary in meaning, according to different interpretations of vague predicates, the correspondence is more fluid: different interpretations will be true in different sets of possible worlds. Thus, in order to separate an agent's beliefs about

the world from their attitude to linguistic meanings, the beliefs must be modelled in a way that is not affected by linguistic variability. Hence, our model of belief is primarily based on sets of plausible possible worlds rather than theories. Of course the structure of a possible world will still be determined relative to a formal language, but this will be the limited language of observable measurement functions, which contains no predicates other than precise ordering relations.

In accordance with the interpretation function, $\llbracket \chi \rrbracket_{\mathfrak{M}}^{w,p,\theta}$, specified above, the agent's attitude to the meanings of vocabulary terms is modelled in terms of both the predicate grounding definitions and the choices of threshold values that the agent considers to be acceptable.

5.1 A formal model of a standpoint

We represent an agent's attitude by a structure that we call a *standpoint*, which characterises the range of possible worlds and linguistic interpretations that are plausible/acceptable to the agent. Formally, a standpoint is modelled by a tuple,

$$\langle B, A, \Psi \rangle,$$

where:

- $B \subseteq W$ is the agent's *belief set*—i.e. the set of possible worlds that are compatible with the agent's beliefs,
- $A \subseteq P$ is the agent's *admissibility set*—i.e. the set of precisifications that the agent considers to make reasonable assignments to all threshold parameters, and hence to be *admissible*,
- $\Psi \subseteq \Theta$ is a set of predicate grounding theories that characterises all possible definitions of ambiguous predicates that the agent regards as acceptable.

In this model, the belief state of an agent is characterised in purely *de re* fashion—that is in terms of states of the world rather than in terms of linguistic propositional expressions that are accepted as true. This belief model (due to Hintikka [10]) is relatively simple and clear, although may be criticised on the grounds that it treats agents as *logically omniscient*—they always believe all logical consequences of their beliefs. In further development it might be useful to introduce a richer belief theory within which one can distinguish an agent's *de re* beliefs, from explicit propositional belief (along the lines of [7]).

In order that A adequately models the set of precisifications that are acceptable to an agent, we may want to place restrictions on which subsets of P are regarded as legitimate admissibility sets. One plausible requirement is that the set of acceptable values that could be assigned to any given threshold parameter ought to lie in a range that is convex with respect to the ordering relation on the value domain. In other words, the possible values of a parameter include all intermediate values that lie between other possible values. For example, if one sometimes uses language on the basis of a threshold for tallness of 180 cm and at other times on the basis of 185 cm being the threshold, then one may argue that 182 cm or 183 cm must also be reasonable threshold choices. Counter to this, one might object that, in describing a particular state of the world involving

specific individuals, it may be natural to divide tall from short individuals only at certain points, because of the distribution of their heights—i.e. such a break is more natural if it falls in a gap in the height distribution of the individuals under consideration. The question of what conditions should be placed on A is a subject of ongoing work. But, even without any further constraints, our semantics already captures significant aspects of the interpretation of vague predicates.

The Ψ component of a standpoint allows one to model an agent's ambivalence with regard to the appropriate grounding definition for certain predicates. Thus, rather than specifying each grounding theory $\theta \in \Psi$ separately it would be more feasible to give a range of possible definitions for each conceptually vague predicate (or group of inter-dependent group of predicates). An acceptable grounding theory then corresponds to a selection for each predicate of a particular definition from a set of possible alternatives.

5.2 Truth with respect to a standpoint

For any given model \mathfrak{M} , we can now formally define the condition that a formula ϕ holds with respect to a particular standpoint $\langle B, A, \Psi \rangle$. Specifically, we define:

- $\mathfrak{M}, \langle B, A, \Psi \rangle \Vdash \phi$ iff $(B \times A \times \Psi) \subseteq \llbracket \phi \rrbracket_{\mathfrak{M}}$.

So ϕ holds for a standpoint if it is true at all worlds in the belief set for all admissible precisifications and all acceptable predicate grounding theories. In other words, the agent considers that, for any reasonable interpretation of ambiguous predicates and all choices of threshold parameters, ϕ is true in all possible worlds consistent with the agent's beliefs.

5.3 Weaker forms of assertion relative to a standpoint

Our standpoint semantics also enables us to specify a number of modal-like operators by means of which we can describe more ambivalent and/or less confident attitudes that an agent may have to a given proposition:

- $\mathfrak{M}, \langle B, A, \Psi \rangle \Vdash \text{CouldSay}(\phi)$ iff $(B \times \{p\} \times \{\theta\}) \subseteq \llbracket \phi \rrbracket_{\mathfrak{M}}$,
for some $p \in A$ and some $\theta \in \Psi$.
- $\mathfrak{M}, \langle B, A, \Psi \rangle \Vdash \text{CouldBe}(\phi)$ iff $(\{w\} \times A \times \Psi) \subseteq \llbracket \phi \rrbracket_{\mathfrak{M}}$, for some $w \in B$.
- $\mathfrak{M}, \langle B, A, \Psi \rangle \Vdash \text{CouldBeSay}(\phi)$ iff $\langle w, p, \theta \rangle \in \llbracket \phi \rrbracket_{\mathfrak{M}}$,
for some $w \in B$, $p \in A$ and $\theta \in \Psi$.

$\text{CouldSay}(\phi)$ asserts that for all worlds in the agent's belief set, ϕ is true in some admissible precisification for some acceptable grounding theory. This operator is used to characterise an assertion made in a context where an agent is fully confident that their beliefs relevant to ϕ are correct, but is unsure about the choice of words used to express ϕ . By contrast, $\text{CouldBe}(\phi)$ means that, for all reasonable interpretations of predicate definitions and thresholds, there is some world compatible with the agent's beliefs where ϕ is true. In this case the interpretation of the words used to express ϕ is taken to be uncontroversial, but the state of reality, which would determine whether ϕ is true, is uncertain. Finally, $\text{CouldBeSay}(\phi)$ indicates that there is some combination

of acceptable predicate definitions, threshold choices and a world state compatible with the agent's beliefs, according to which ϕ would be interpreted as true.

This distinction between the operators CouldBe and CouldSay is closely related to distinctions made by J.L. Austin [1], in his analysis of different ways in which the sense of a predicate may be related to the properties of an object to which it is applied. He introduced the idea of the *onus of match* between the sense of a word and the corresponding property of the object, and suggested that in some speech situations one is clear about the meaning of the word but unsure whether the object possesses the appropriate property, whereas in others one is clear about the properties of the object but unsure about whether the word adequately describes the object.

A number of other modalities could be defined. In the specifications for CouldSay(ϕ) and CouldBeSay(ϕ), the indices giving the precisification p and grounding theory θ are both allowed to vary (independently). But we could, for instance, define a modality M such that $M(\phi)$ is true iff $(B \times A \times \{\theta\}) \subseteq [\![\phi]\!]_{\mathfrak{M}}$ for some particular $\theta \in \Psi$ —i.e. there is some acceptable grounding theory, relative to which ϕ holds for every admissible precisification. Such a modality does not have an obvious informal interpretation, since in ordinary language we tend to conflate conceptual and sorites variability within the general phenomenon of vagueness; however, it may still provide an informative characterisation of an agent's attitude to a proposition.

6 Further work and conclusions

We have given an overview of a semantic framework within which various significant aspects of vagueness can be articulated. Although the theory developed so far is already quite complex, there are still many loose ends that would need to be tied up in order to provide a solid foundation for representing and reasoning with information expressed in terms of vague predicates. As such, the framework is intended to provide a platform upon which more practical knowledge representation languages can be developed. Such development would most likely involve both simplification of some parts of the formalism and elaboration of others.

In the current work, the focus has been on semantics. Although the syntax of a formal representation has been specified, nothing has been said about the proof theory governing valid inference within this system. In fact, since the basic language is essentially a variant of first-order logic, standard proof systems can be applied. However, if (as suggested in Section 5.3) the language is extended by modal operators to express the ambivalent truth status of vague propositions, then additional inference rules will be needed in order to take account of these operators. Investigation of the proof theory of such extended languages is the subject of ongoing work.

An obvious major extension to the theory would be to add probability distributions over the belief set and/or the set of admissible precisifications to model the relative plausibility of different possible worlds and the relative acceptability of different precisifications. Indeed, work has already been carried out on such an extension. Initial investigations seem to be fruitful and indicate that this approach may be fruitful as a means to explain the famous *sorites paradox* and related phenomena.

Another interesting direction for further work would be to study the assimilation of new information in terms of the transformation from an agent's initial standpoint to a

modified standpoint. Here the issue arises that when an agent receives information that is incompatible with their current standpoint, they must choose whether to modify their beliefs or to try to interpret the information from the point of view of a different standpoint. This study could extend to more general aspects of the exchange of information between two agents in the presence of vagueness. The notion of *context* is likely to be relevant to such an investigation.

For practical applications it may be convenient to replace the somewhat elaborate model theory we have given with a more standard first-order semantics. This would require that the semantic indices (worlds and precisifications) were in some way incorporated into the object language. This could be achieved in a similar way to how temporal logics are often treated within AI formalisms (e.g. by a Situation Calculus [15] style formulation such as $\text{Holds}(\phi, \langle w, p, \theta \rangle)$).

In summary, this paper has outlined the structure of a formal semantics for interpreting vague languages, that models both the definitional ambiguity of conceptual terms and the variability of thresholds used to determine their applicability. The framework characterises an explicit link between the thresholds governing predicate applicability and observable properties of the world. This link provides a basis for detailed semantic modelling of the modes of variability in the meanings of particular vague predicative terms. It also enables specification of complex penumbral connections between related terms. Additionally, the paper has suggested a model of the cognitive standpoint of an intelligent agent incorporating both a belief state and an attitude towards the interpretation of vague terms.

BIBLIOGRAPHY

- [1] John L. Austin. How to talk: some simple ways. *Proceedings of the Aristotelian Society*, 1953. Also in J.L. Austin Philosophical Papers, Clarendon Press, 1961.
- [2] Brandon Bennett. Modal semantics for knowledge bases dealing with vague concepts. In Anthony G. Cohn, Leonard Schubert, and S. Shapiro, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the 6th International Conference (KR-98)*, pages 234–244. Morgan Kaufman, 1998.
- [3] Brandon Bennett. Modes of concept definition and varieties of vagueness. *Applied Ontology*, 1(1): 17–26, 2005.
- [4] Brandon Bennett. A theory of vague adjectives grounded in relevant observables. In Patrick Doherty, John Mylopoulos, and Christopher A. Welty, editors, *Proceedings of the Tenth International Conference on Principles of Knowledge Representation and Reasoning*, pages 36–45. AAAI Press, 2006.
- [5] Brandon Bennett. Possible worlds and possible meanings: A semantics for the interpretation of vague languages. In *Commonsense 2011: Tenth International Symposium on Logical Formalizations of Commonsense Reasoning*, AAAI Spring Symposium, Stanford University, 2011. AAAI.
- [6] Brandon Bennett, David Mallenby, and Allan Third. An ontology for grounding vague geographic terms. In C. Eschenbach and M. Gruninger, editors, *Proceedings of the 5th International Conference on Formal Ontology in Information Systems (FOIS-08)*. IOS Press, 2008.
- [7] Ronald Fagin and Joseph Y. Halpern. Belief, awareness, and limited reasoning. *Artificial Intelligence*, 34:39–76, 1988.
- [8] Kit Fine. Vagueness, truth and logic. *Synthèse*, 30:263–300, 1975.
- [9] Joseph Y. Halpern. Intransitivity and vagueness. In *Principles of Knowledge Representation: Proceedings of the Ninth International Conference (KR-2004)*, pages 121–129, 2004.
- [10] Jaakko Hintikka. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press, Ithaca, N.Y., 1962.
- [11] Hans Kamp. Two theories about adjectives. In E.L. Keenan, editor, *Formal Semantics of Natural Language*. Cambridge University Press, Cambridge, England, 1975.
- [12] Rosanna Keefe and Peter Smith. *Vagueness: a Reader*. MIT Press, 1996.

- [13] Jonathan Lawry. Appropriateness measures: an uncertainty model for vague concepts. *Synthese*, 161:255–269, 2008.
- [14] Jonathan Lawry and Yongchuan Tang. Uncertainty modelling for vague concepts: A prototype theory approach. *Artificial Intelligence*, 173(18):1539–1558, 2009.
- [15] John McCarthy and Patrick J. Hayes. Some philosophical problems from the standpoint of artificial intelligence. In B. Meltzer and D. Mitchie, editors, *Machine Intelligence*, volume 4, pages 463–502. Edinburgh University Press, 1969.
- [16] Henryk Mehlberg. *The Reach of Science*. University of Toronto Press, 1958. Extract on Truth and Vagueness, pages 427–55, reprinted in [12].
- [17] Paulo Santos, Brandon Bennett, and Georgios Sakellariou. Supervaluation semantics for an inland water feature ontology. In Leslie Pack Kaelbling and Alessandro Saffiotti, editors, *Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI-05)*, pages 564–569. Edinburgh, 2005. Professional Book Center.
- [18] Bas C. van Fraassen. Presupposition, supervaluations and free logic. In Karel Lambert, editor, *The Logical Way of Doing Things*, Chapter 4, pages 67–91. Yale University Press, New Haven, 1969.
- [19] Lofti A. Zadeh. Fuzzy logic and approximate reasoning. *Synthese*, 30:407–428, 1975.

Brandon Bennett
School of Computing
University of Leeds
Leeds, LS2 9JT, UK
Email: B.Bennett@leeds.ac.uk