



Hypercomparatives

Author(s): Adam Morton

Reviewed work(s):

Source: *Synthese*, Vol. 111, No. 1 (Apr., 1997), pp. 97-114

Published by: [Springer](#)

Stable URL: <http://www.jstor.org/stable/20117617>

Accessed: 11/10/2012 16:07

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Springer is collaborating with JSTOR to digitize, preserve and extend access to *Synthese*.

<http://www.jstor.org>

ADAM MORTON

HYPERCOMPARATIVES*

ABSTRACT. In natural language we rarely use relation-words with more than three argument places. This paper studies one systematic device, rooted in natural language, by which relations of greater adicity can be expressed. It is based on a higher-order relation between 1-place, 2-place, and 4-place relations (and so on) of which the relation between the positive and comparative degrees of a predicate is a special case. Two formal languages are presented in this connection, one of which represents the language of communication and the other the contextual information against which the first language is interpreted. A semantical theory is described, which treats the two languages in an interdependent way. Logical consequence is non-compact. Connections with issues about vagueness are made.

We can only say and comprehend quite simple things. Beyond a certain point syntactical complexity defeats us. So it is not surprising that natural languages have evolved ways of expressing complex thoughts with simpler linguistic expressions, typically by clever use of context. In this paper I discuss one of the ways in which we keep our syntax manageable: we avoid predicates of very great adicity. 3-place predicates are fairly uncommon in ordinary language, and 4-place predicates are definitely rare. But these are pretty small numbers and we think many thoughts that can naturally be expressed in terms of n -place predicates for larger n . For example, I argue below that many ascriptions of color implicitly refer to a 4-place relation, although the color predicates are typically monadic. Somewhat different is the progression: wants, prefers, prefers-by-more. In natural language we say 'a wants o '; economists favor 'a prefers o_1 to o_2 '; and there are also preferences which are best expressed in terms of 'a prefers o_1 to o_2 by more than she prefers o_3 to o_4 '. (See Morton 1991, 1994).

One way in which we manage to express some of these thoughts without too complex a syntax is by asserting that objects are related by some relation R , in a context in which saying this expresses the claim that these *and others* are related by another relation R' . R' usually bears some higher-order relation to R , which is determined by the context. For example one might say 'a is equivalent to b', meaning that a bears to b an equivalence relation which also relates other objects under discussion. (Sometimes, for integers 'n is equivalent to m' will mean ' $n \equiv m \pmod{p}$ ' for some

contextually specified p .) Or one might say ‘ a is to the right of b ’, meaning ‘from point c the rotation from line ca line cb is anticlockwise in the plane orthogonal to direction δ ’. I shall discuss one particular higher-order relation, the H relation, which can hold between relations of different adicities. It holds between ‘ x is P ’ and ‘ x is P -er than y ’, and between this second relation and ‘ x is P -er than y by more than z is P -er than z ’. And so on. I call it the H -relation because the four place relation just mentioned is a generalization of the comparative, of the predicate P . This approach thus contrasts with approaches such as those of Klein (1980) which use concepts which correspond to the positive degree of adjectives and Kamp (1975, 1981) which use the comparative, since it analyses a relation which holds between positive, comparative, hypercomparative, and so on, none of which is basic.

My method is to construct two formal languages. One models a language in which communication is carried out, and the other models the contextual information relative to which sentences of the first language are evaluated. This two language approach captures the inferential patterns of a variety of English idioms and models some kinds of vagueness. It turns out that logical consequence for the system is non-compact, and therefore does not have a complete axiomatization. This fact is relevant to the analysis of sorites paradoxes.

1. THE LANGUAGES

The first language, \mathbf{H} consists of a first-order predicate calculus augmented with two additional operators, C and M . C doubles the number of argument places of a predicate, and its intended function is to transform positives into comparatives into hypercomparatives. M is a predicate modifier interpretable as ‘much’ or ‘very’. If P^1x is interpreted as x is red then CP^1xy is interpreted as x is at least as red as y , MP^1x as x is very red, MCP^1xy as x is much more red than y , and CCP^1xyzw as x is more red than y by at least as much as z is more red than w . (Or, more idiomatically, *the difference between the degrees to which x and y are red is at least as great as that between the degrees to which z and w are red*. I shall sometimes, when no confusion will result, say just x is P -er than y or x is P -er than y by more than z is P -er than w .) Note that claims expressed in terms of the hypercomparative, though unnatural-sounding, are often more unambiguously true or false than those involving just the positive or comparative. Thus it may be a pretty moot point whether a particular shade of purple-red is red, and also not at all clear whether that shade of purple-red is more red than a shade of orange-red. But it is a definite fact

that one shade of purple-red is more red than another by more than a given shade of orange-red is more red than another.

Sentences like these can also be formalized in a very different way: *x is red* as $\exists\delta Px\delta$; *x is more red than y* as $\exists\delta\exists\mu(Py\delta \& Py\mu \& \delta > \mu)$, and *x is more red than y by more than z is more red than w* as $\exists\delta\mu\exists\nu\exists\sigma (Px\delta \& Py\mu \& Pz\nu \& Pw\sigma \& (\delta, \mu) >' (\nu, \sigma))$. Here the variables δ, μ, ν, σ range over degrees of redness and $>$ and $>'$ are orderings over degrees and pairs of degrees. (See the H_n orderings defined in Section 3 below.) The second formal language I shall use, **I**, expresses these relations between degrees in essentially these terms. The language of degrees and orderings is more powerful than that of positive and comparative but further from the appearance of natural language. I shall use sentences of **I** to express contextual information against which sentences of **H** are evaluated. (For I-style formalization in a linguistic context see Klein 1980; Stechow 1983 or in a philosophical context Williams and Morton 1984; Clark 1984.)

H represents the pure idiom of positive and comparative (and hypercomparative). It does not have quantifiers ranging over degrees. But it does have two predicate modifiers, the comparative-forming operator *C* and the 'more/very' operator *M*. **H** contains an infinite stock of one-place predicates, P^n , infinitely many variables v_n , $\&$, \neg , \forall , *M*, and *C*. (There are no individual constants, nor are there n -place primitive predicates for $n > 1$.) Take \exists and \supset to be defined as usual. The formation rules are as for any predicate calculus augmented with the rules:

If *S* consists of a primitive predicate preceded by a string of n *C*'s then *S* is a 2^n -place predicate.

If *S* consists of a primitive predicate preceded by a string of n *C*'s preceded by a single *M* then *S* is a 2^n -place predicate.

To give these the intended senses, think of e.g. ' $\forall v_1\forall v_2\exists v_3\exists v_4 CCPv_1v_2v_3v_4$ ' as saying that for all v_1 and v_2 there are v_3 and v_4 such that v_1 is *P*-er than v_2 by at least as much as v_3 is *P*-er than v_4 . And think of ' $MCPv_1v_2$ ' as saying that v_1 is much *P*-er than v_2 . (I take 'very' to be the 1-place variant of 'much'. So ' MPv_1 ' says that v_1 is very *P*.)

In referring to sentences of **H** I shall sometimes use a vector notation to save subscripts and dots. Thus instead of $\forall v_1\forall v_2 \dots \forall v_{2^n} C \dots CP^m(v_1, \dots, v_{2^n})$ (n *C*'s) I shall write $\forall \mathbf{v} C^{(n)}P^m(\mathbf{v})$. That is, I shall indicate a string of n variables by a boldface letter, a string of n *C*'s by $C^{(n)}$, and let the number of variables be deduced from context (where it is so deducible, that is). I shall write (\mathbf{v}, \mathbf{w}) for the concatenation of two equally long strings of variables.

In working out a semantics for **H** what validities should we aim at? Here are some obvious ones:

- All instances of valid schemata of first order logic (including cases where a predicate involving C or M is substituted for a schematic letter).
- All instances of the following

- (i) $(C^{(m)}P(\mathbf{v}) \ \& \ \neg C^{(m)}P(\mathbf{w})) \supset C^{(m+1)}P(\mathbf{v}, \mathbf{w})$
 “ x at least as P as y and z not at least as P as w entails x more P than y by at least as much as z more P than w ”.
- (ii) $(C^{(m)}P(\mathbf{s}, \mathbf{t}) \ \& \ C^{(m)}P(\mathbf{t}, \mathbf{u})) \supset C^{(m)}P(\mathbf{s}, \mathbf{u})$
 “ x at least as P as y and z at least as P as w entails x at least as P as z ”.
- (iii) $C^{(m)}P(\mathbf{s}, \mathbf{s})$
 “ x is at least as P as x ”.
- (iv) $(C^{(m)}P(\mathbf{s}, \mathbf{t}) \ \& \ C^{(m)}P(\mathbf{t}, \mathbf{u}) \ \& \ C^{(m)}P(\mathbf{u}, \mathbf{v})) \supset C^{(m+1)}P(\mathbf{s}, \mathbf{v}, \mathbf{t}, \mathbf{u})$
 “ x at least as P as y and y at least as P as z and z at least as P as w entails x more P than w by more than y is more P than z ”.
- (v) $(C^{(m)}P(\mathbf{v}, \mathbf{w}) \ \& \ MC^{(m-1)}P(\mathbf{w})) \supset MC^{(m-1)}P(\mathbf{v})$
 “ x at least as P as y and y very P entails x very P ”.
- (vi) $(MC^{(m)}P(\mathbf{u}) \ \& \ \neg MC^{(m)}P(\mathbf{v})) \supset \neg C^{(m+1)}P(\mathbf{v}, \mathbf{u})$
 “ x very P and y not very P entails not x at least as P as y ”.

(The statement in quotes beneath each formula is a crude version of a special case, to help in reading it.)

Note how the last two of these capture validities about “very” and “much” which are intuitively important in our grasp of its meaning and which depend essentially on the relation of positive to comparative.

The language **I** has a different syntax, that of a two-sorted quantification theory with quantifiers over both individuals and degrees. **I** is a two-sorted first-order language with dyadic predicates P^1, P^2, \dots , and 2^n -place predicates \geq_m^n for $n, m = 0, 1, \dots$. The P^m correspond to the P^m of **H**; the intended sense of $P^m a \delta$ is that a has P^m (of **H**) to degree δ . And each \geq_m^n corresponds to $C^{(n)}P^m$: $\delta \geq_m^n \delta'$ when δ, δ' represent the degrees of (the constituents of) \mathbf{x}, \mathbf{x}' and $C^{(n)}P^m \mathbf{x} \mathbf{x}'$. The variables are x, y, z, z', z'', \dots , ranging over individuals (objects in a specified domain) and δ, δ', \dots , ranging over degrees. The arguments of P^m must be an individual variable and a degree variable, in that order, and the arguments of \geq_m^n must be degree variables. Note that \geq_m^0 is well formed. It represents the limiting case of a 2^0 -place ordering, that is, a set. The intended sense of $\geq_m^0 \delta$ is that δ is a degree for which P^m holds. The syntax is as for any first-order two-sorted language with these primitives and these restrictions.

(For definiteness' sake assume that the primitive logical symbols are $\&$, \neg , \forall , as for **H**).

2. SEMANTICS: INFORMALLY

The semantics is aimed primarily at **H**. But the best, or at any rate the most interesting, way to do this is to treat **H** and **I** simultaneously. The aim of a semantics for **H** is to define truth-in-a-model in terms of the degrees to which objects in a domain satisfy **H**'s predicates. So we must assign to every pair of an object and a predicate a degree. The important fact about degrees is that they are ordered: the model has to determine orderings between them in terms of which the extension of the $C^{(m)}P^{(n)}$ be determined (the H_n orderings of Section 3 below). One might at first think that this can be done in terms of a single ordering which holds between a pair of degrees when the degree to which the first falls under the predicate in question exceeds that to which the second does. And in fact quite a lot about a predicate can be expressed in terms of this ordering. In particular, different kinds of predicates have different kinds of orderings. The degrees of some predicates ('is pregnant') are simple two-element Boolean algebras. And others ('is intelligent') have complex infinite orderings riddled with incomparabilities. One might hope that the character of the ordering of a predicate's degrees has some bearing on the syntactical category (adjective, attributive adjective, sortal or non-sortal noun, etc.) that expresses it in English. (For related issues see Aqvist 1981; Vendler 1967.)

But in fact we need more complex information than can be given with a two-place ordering. Consider 'red' again. 'Is redder than' is vague: before there can be an answer to the question whether a given shade of purple-red is redder than one of orange-red a particular ordering for 'redder' has to be chosen from those allowed by the meaning of 'red'. What all the permissible orderings have in common is that they are consistent with the facts about when one shade is redder than another by more than a third is more than a fourth. (We could take the position of a pair of shades (δ , σ) in the ordering of pairs by difference of redness to be given by the size of the angles between the radii to δ and σ fixed point in the color solid, say half-way along the line from the most to the least saturated red.) In accordance with this, the semantics for **H** is based on a natural (but not incontestable) assumption:

The grounding assumption: for every predicate P there is an integer m , a set of abstract degrees Δ and a partial ordering of 2^{m-1} -tuples of elements of Δ associated with members of the domain, such that $C^{(m)}P$

holds between 2^m elements of the domain iff their degrees are related by the ordering.

The problem then is to specify the conditions under which $C^{(n)}P$ holds for n not equal to the grounding value m . (This is discussed in the next section.) The grounding assumption is not obviously true for all predicates. It is most plausible for predicates ascribing sizes, shapes, colors, and other perceptual properties. It is least plausible for cluster concepts. It represents an assumption that vagueness is only so deep: ‘underneath’ every predicate there is an ordering of its degrees, though perhaps a very complex one, that is perfectly precise. Even making this assumption it is not clear how surface vagueness is to be tamed and contextuality represented. What follows is one way of doing it.

An interpretation for the language will specify a domain D of objects plus a set Δ of degrees. The connection of a predicate with the objects satisfying it comes indirectly: to each predicate and each object a degree is assigned, and then each predicate will have an extension (roughly) relative to each degree. A sequence of individuals cannot simply satisfy an open sentence, on this approach. For individuals fall under predicates to various degrees, so that the most we can say is that a sequence satisfies an open sentence when the thresholds for possession of the atomic predicates are set at specific points. Even this is sometimes more than can be assumed, since there will often be a choice of ordering relations for a predicate so that satisfaction will be relative not only to the choice of a threshold but also to the choice of the ordering on which the threshold lies.

Consider for example how we may set thresholds for a predicate like “red” given in terms of its hypercomparative, with a consequent choice of orderings for its comparative. For any two degrees (e.g. hues) δ and σ it may be indeterminate whether one dominates the other or whether they are incomparable. But suppose it is arbitrarily decided not only that $\delta \geq \sigma$ but that the pair (δ, σ) is a threshold case, such that the first is minimally redder than the second. Then the class of dyadic orderings is restricted in two ways. First, to those in which $\delta \geq \sigma$. And second, to those in which the separation of pairs is at least as great as that between δ and σ , that is, such that $\tau \geq \mu$ only when $(\tau, \mu) \geq' (\delta, \sigma)$. (Thus \geq' is an ordering of pairs which is to P -er as \geq is to P .) Note that if one has narrowed the choice of interpretations of the comparative in this way one will still need to pick a threshold-point to fix the interpretation of the positive. (Having fixed upon a subordering of the color solid to be “redder than” one still needs to find a point on it to divide red from not-red.) The converse is not generally true: to fix the interpretation of a comparative one does not need to have determined a threshold-point for the positive. And in general

$C^{(m)}P$ requires thresholds to be fixed for all degrees between m and the n for which the interpretation gives an unambiguous extension.

So something more complex than a simple threshold is often needed. The best way to give complex information is with a structured proposition, and here the language **I** enters. I shall specify truth for sentences of **H** relative to a model M and an arbitrary sentence of **I**. This sentence can specify a threshold, but it can also specify more complex information.

3. CONSTRAINTS ON THE H-RELATION

Suppose that we are given an ordering of pairs, $(\delta, \sigma) \geq (\tau, \mu)$, where $\delta, \sigma, \tau, \mu$ range over degrees of a predicate P , and \geq is a pre-ordering (transitive and reflexive). What constraints does this put on the choice of the corresponding ordering $\delta \geq \sigma$? (We know when x is P -er than y by more than z is P -er than w , and we need to know the possibilities for x is P -er than y .) In this section I state some relations between orderings of single degrees and orderings of pairs of degrees (and in general between orderings of n -tuples and orderings of $2n$ tuples) which hold when the two orderings correspond to the comparative and hypercomparative of the same predicate. In stating them it is important to keep in mind which ordering is the fixed basis which is constraining the other. Thus we might begin with an ordering \geq of degrees which gave the basic facts about the structure of the predicate (see the grounding assumption of the previous section) and want to constrain orderings \geq^+ of pairs of degrees which are to determine the truth of hypercomparative assertions in terms of its comparative. Or we might begin with an ordering \geq of pairs of degrees and want to constrain orderings \geq^- of degrees which are to determine the truth of comparative assertions in terms of the hypercomparative. Given an ordering of degrees and an ordering of pairs of degrees the constraints will be different, depending on which one is taken as the determining one and which the determined.

One obvious constraint says that if a pair of degrees is contained within another pair, in the ordering of degrees, then their separation cannot be larger, as measured by the derived ordering of pairs. Formally

$$(C1) \quad \text{if } \delta > \sigma_1 > \tau \ \& \ \delta > \sigma_2 > \tau \text{ then not } (\sigma_1, \sigma_2) >^+ (\delta, \tau)$$

(where $>$, and below \cong , are defined in terms of \geq in the obvious way.) σ_1 and σ_2 lie between δ and τ , so clearly their separation is less. (The

formulation allows that σ_1 and σ_2 may be incomparable in \geq .) Another very natural constraint is

- (C2) if $(\delta, \sigma) > (\sigma, \delta)$ then $\delta \geq^- \sigma$ and if $\delta > \sigma$ then $(\delta, \sigma) >^+ (\sigma, \delta)$ and if $\delta \cong \sigma$ then $(\delta, \sigma) \cong^+ (\sigma, \delta)$.

Given \geq between single degrees this determines only the ordering in \geq^+ of pairs of the form (δ, σ) , (σ, δ) . And given \geq between pairs it determines only the ordering in \geq^- of δ and σ when one of (δ, σ) and (σ, δ) is strictly greater than the other. (C2) would determine the ordering of pairs from the ordering of degrees if the *if* were an *iff*. But that, though at first sight plausible, ignores the possibility that when (δ, σ) and (σ, δ) are incomparable in the ordering of pairs different orderings of single degrees compatible with that pair ordering may rank x and y differently. (Think again of borderline shades of red.) Similar considerations show that the + and – superscripts are necessary in (C1), (C3), and (C4).

- (C3) if $(\delta, \sigma) > (\tau, \mu)$ and $\tau \geq^- \mu$ then $\delta \geq^- \sigma$ and if $(\delta, \sigma) >^+ (\tau, \mu)$ and $\tau > \mu$ then $\delta > \sigma$.

(C1)–(C3) should be augmented with a self-explanatory ‘book-keeping’ principle.

- (C4) if $\delta > \sigma$ then for all τ, μ $(\tau, \mu) \geq^+ (\sigma, \delta)$.

These are pretty weak constraints. A slightly stronger constraint is provided by the thought that if $\delta > \sigma > \tau$ then $(\delta, \tau) > (\delta, \sigma)$. In fact this condition is not obviously true except in the case where δ , σ , and τ are separated by a finite chain of minimal differences. For in other cases there might be infinitely many elements between δ , σ , and τ , in which case it is far from obvious that the relative separation of (δ, τ) and (δ, σ) is determined just by the ordering of δ , σ , τ . (Metric facts go beyond topological facts.) But we should ensure that e.g. steps 999 and 1 in an ω series are separated by more than steps 2 and 1. Thus the following constraint. To state it I define a \geq -preserving map to be a function Φ such that for all σ, τ $\Phi\sigma \geq \Phi\tau$ iff $\sigma \geq \tau$.

- (C5) if $\delta > \sigma$ and $\tau > \mu$ and for all \geq -preserving maps for which $\delta \geq \Phi\tau > \Phi\mu \geq \sigma$, it is not the case that $\delta \cong \Phi\tau$ and $\sigma \cong \Phi\mu$, then $(\delta, \sigma) > (\tau, \mu)$.

(There are no +, – superscripts here, because the constraint applies in both directions.)

(C1)–(C5) generalise in an obvious way to relations between $\geq^{(n)}$ and $\geq^{(n+1)}$. For example (C1) generalizes to

$$\text{if } \delta \geq^{(n)} \sigma_1 \geq^{(n)} \tau \text{ \& } \delta \geq^{(n)} \sigma_2 \geq^{(n)} \tau \text{ then not } (\sigma_1, \sigma_2) >^{(n+1)+} (\delta, \tau)$$

where $\delta, \sigma_1, \sigma_2, \tau$ are 2^{n-1} -tuples of degrees.

In order to be perfectly explicit about these orderings let me define a H_n ordering to be a relation R between 2^n members of some domain, such that the corresponding dyadic relation \geq , which holds between $(\delta_1, \dots, \delta_{n-1})$ and $(\sigma_1, \dots, \sigma_{n-1})$ when $(\delta_1, \dots, \delta_{n-1}, \sigma_1, \dots, \sigma_{n-1})$ is in R , is a pre-ordering. I shall often notate a H_n ordering as if it were the corresponding pre-ordering, thus the $\geq^{(n)}$ expressions above.

We can thus define a relation which holds between a basic ordering $\geq^{(n)}$ and a derived ordering $\geq^{(m)}$ when they are joined by a chain of orderings satisfying (C1)–(C5) taken thus generalized. (the chain going in the + direction if $m > n$ and the – direction if $n > m$). Call this the H -relation. It plays a central role in the semantics below.

An interesting consequence of these constraints concerns the characterization of standard and non-standard models. Take a sequence of order type $\omega(\omega + \omega^*)$ (a typical non-standard model of arithmetic). It is not easy to say in unproblematic, for example first-order, terms what the difference between this and a standard ω -sequence is, although they are intuitively very different. But thinking in terms of orderings of pairs as well as of elements the two sequences are different in several ways. For example the $\omega(\omega + \omega^*)$ sequence, but not the ω sequence, will intuitively satisfy the following sentence

$$\begin{aligned} \exists \delta \exists \sigma \forall \tau ((\tau < \delta \text{ \& } (\delta, \tau) \geq (\sigma, \delta)) \supset \exists \mu ((\delta, \mu) \\ \geq (\sigma, \delta) \text{ \& } \delta > \mu > \tau)). \end{aligned}$$

The sentence in effect says that there are infinite separations; there is an object δ such that for each individual preceding δ by more than a certain separation there is another succeeding it but preceding δ and also separated from δ by at least that separation. (I said ‘intuitively’, because there are also orderings of pairs consistent with the $\omega(\omega^* + \omega)$ ordering which do not make this sentence true. But they are not the orderings that express the standard sense of separation in the ordering.) A formulation like this will be used in Section 5 below, in the incompleteness proof.

4. SEMANTICS: FORMALLY

An interpretation for \mathbf{H} is a 4-tuple $\langle D, \Delta, I, J \rangle$. D is the domain for quantifiers of \mathbf{H} ; Δ is a set from which the degrees to which objects satisfy predicates are taken; I assigns orderings between degrees to predicates; and J assigns degrees in Δ to pairs of a member of D and a predicate.

Δ has to contain a partially ordered set of degrees for each predicate. But there is no need to postulate a *separate* set for each predicate in the language. For often different predicates can be related to the same degrees. For example both “big” and “small” have sizes as degrees and all color predicates have hues. I shall assume that all of the predicates of \mathbf{H} have degrees which come from a single set Δ . Even if the stock of atomic predicates have very varied intended interpretations, degrees for all of them can be found in Δ as long as a rich enough variety of H_n orderings can be constructed over it.

The satisfaction relation is thus $\sigma, M, \mathbf{t} \models s$ meaning: sequence σ of objects from D satisfies closed or open sentence s of \mathbf{H} under interpretation M relative to contextual information provided by (closed) sentence \mathbf{t} of \mathbf{I} .

For each primitive predicate P^n , $I(P^n)$ is a H_m ordering between $(m - 1)$ -tuples of members of D , representing some hypercomparative of P (or P itself if $m = 0$). I shall refer to the relevant m for a given P^n as $I(n)$. For each member d of D and each P^n , $J(d, P^n)$ is a member of Δ , representing the degree to which d has P^n . For any variable v_n , $I(v_n)$ is a member of D .

Sentences of \mathbf{I} can also be evaluated relative to a model M . Given an interpretation $\langle D, \Delta, I, J \rangle$ of \mathbf{H} let a H -to- I model be a two-sorted first-order model $\langle D, \Delta, I' \rangle$ for \mathbf{I} , in which I' assigns to each P^m of \mathbf{I} the relation $\{(x, J(x, P^m)); x \in D\}$, assigns to each \geq_m^n some H_n ordering over Δ , assigns a member of D to each free individual variable and a member of Δ to each degree variable, and is otherwise like I .

I shall assume that the orderings assigned by I and I' are such that they contain a least upper bound and a greatest lower bound for every sequence of degrees ordered by the orderings, and that Δ is rich enough to contain these least and greatest bounds. Call this the *completeness assumption* on I and Δ .

Now to define $\sigma, M, \mathbf{t} \models s$. There are two cases for the basis clauses.

(a) s is $C^{(m)}P^n(x)$. For simplicity I shall assume that \mathbf{x} is x_1, x_2, \dots . And I shall write ‘ $J(s)$ ’ for ‘ $J(\sigma(s), P^n)$ ’. Note that \mathbf{t} is redundant when $m = I(n)$.

If $m = I(n)$ then $\sigma, M, \mathbf{t} \models C^{(m)}P^n(\mathbf{x})$ iff $(J(1), \dots, J(m)) \in I(P^n)$.

If m is not $I(n)$ then $\sigma, M, \mathbf{t} \models C^{(m)}P^n(\mathbf{x})$ iff $(J(1), \dots, J(m)) \in R$ for all relations R in D^m such that

- (i) R bears the H -relation to $I(P^n)$ and
(ii) there is a H -to- I model for M which assigns R to \geq_n^m and which makes \mathbf{t} true.

(b) s is $MC^{(m)}P^n(x)$. M is defined in terms of C . The idea is to say that e.g. x is very P if it is ‘more than half way up’ in the scale of P -ness: for every degree above the degree to which x is P there is another as far below it. This might not seem to need a clause in the semantics, for one might make $MC^{(m)}P^n(\mathbf{x})$ a notational abbreviation for

$$(\mathbf{y})(C^{(m+1)}P^n(\mathbf{y}, \mathbf{x})) \supset \exists \mathbf{z}(C^{(m+1)}P^n(\mathbf{x}, \mathbf{z}) \ \& \ C^{(m+2)}P^n(\mathbf{x}, \mathbf{z}, \mathbf{y}, \mathbf{x}))$$

But the problem would then arise that in a model of H there may not actually be many objects satisfying P , and we would want to allow that, for example, only six things satisfy P and they are all very P . So clause (b) captures the same sense but working in terms of degrees in Δ rather than objects in D . (The same kind of consideration is relevant to Wheeler’s (1972) definition of ‘very tall’ as ‘tall among the tall things’: one wants it to be possible that all the actually tall things are very tall.) Therefore a better formulation is:

$\sigma, M, \mathbf{t} \models MC^{(m)}P^n(\mathbf{x})$ iff
if a 2^{m+1} -membered sequence δ of members of Δ is such that $\delta_i = \sigma(i)$ for $i \leq n$ and $\delta \in R^{m+1}$ then there is a 2^{m+1} -membered sequence δ' such that $\delta'_i = \sigma(i)$ for $i > n$ and $\delta' \in R^{m+1}$ and $(\delta, \delta') \in R^{m+2}$, for all R^{m+1}, R^{m+2} bearing the H relation to $I(P^n)$ and to each other.

Standard recursion clauses for $\&$, \neg , \forall , complete the definition. Note how (b) makes ‘very’ and ‘much’ instances of the same concept, as seems intuitively right. In doing so it has to refer to three H -related relations. Note also that this analysis makes very/much independent of the context information \mathbf{t} . Whether something is very tall depends only on where its height lies among the possible heights. These two features are independent: there are variant definitions of very/much which allow for contextually set borderlines, but they are not needed to make the points I am making in this paper.

5. VALIDITIES AND INCOMPLETENESS

Call s a *semantical consequence* of H, \mathbf{T} , where H is a set of sentences of \mathbf{H} and \mathbf{T} is a set of sentences of \mathbf{I} , if for all \mathbf{t} in \mathbf{T} and h in H $\sigma, M, \mathbf{t} \models s$

holds for all σ over the domains of all models M such that $\sigma, M, \mathbf{t} \models h$. Write this $H, \mathbf{T} \models s$. Call s *valid* if $H, \mathbf{T} \models s$ for all H , and all satisfiable \mathbf{T} . All the members of the list in Section 2 are valid according to this definition. Thus (i)–(iv) are immediate consequences of the fact that the definition constrains the $C^{(m)}$ to satisfy the conditions on the H -relation. (v), (vi) follow immediately from clause (b) of the definition of satisfaction. In fact, (i)–(vi) exploit very little of the resources of the definition. They are not for example sensitive to the interaction of contextual information and logical form involved in the semantics.

Although the semantics for \mathbf{H} construes it as sound with respect to the validities listed it does not have a complete axiomatization. For semantical consequence for \mathbf{H} , given the semantics, is non-compact. Consider the relation between the infinite set of premises \mathbf{P}_n , the conclusion \mathbf{C} , and the contextual information \mathbf{t} below.

$$\mathbf{P}_1 : \quad \forall v_1 \forall v_2 (\underline{P}'v_1v_2 \vee \underline{P}'v_2v_1 \vee \underline{P}'v_2v_1)$$

$$\mathbf{P}_n (n > 2) : \quad \exists v_1 \dots \exists v_n \exists y \forall z (\underline{P}'vv_2 \& \dots \& \underline{P}'v_{n-1}v_n \& \underline{P}'v_ny)$$

$$\mathbf{t} : \quad \forall \delta (\forall \sigma (\sigma > \delta \supset \exists \tau (\sigma > \tau > \delta)) \supset \forall \sigma \forall \tau ((\sigma > \delta \& \tau > \delta) \supset (\sigma, \delta) \cong (\tau, \delta))$$

$$\mathbf{C} : \quad \exists x \exists y \exists z (\underline{P}'xy \& \underline{P}'yz \& \underline{P}''xzyz)$$

(I have used a greater variety of variables than the official syntax of \mathbf{H} and \mathbf{I} permits, for the sake of clarity. And I have written P for P^1 and $>$ for $>_1$, $\underline{P}xy$ for $Pxy \& \neg Pyx$, and $\underline{P}=\underline{P}$ for $Pxy \& Pyx$.)

The \mathbf{P}_n are premises. \mathbf{P}_1 says that P' forms a complete ordering. Each successive \mathbf{P}_n says that there are n objects related by \underline{P}' , and there is another object to which all n objects bear \underline{P}' . The conclusion \mathbf{C} says that there are two pairs among the objects in the domain which have equal separations although one of the pairs is contained within the other as ordered by \underline{P}' . The contextual information \mathbf{t} requires that if there are degrees in limit ordinal position, i.e. with no immediate predecessor and infinitely many non-immediate predecessors, then the separations between such degrees and all preceding degrees are the same. (I shall say of both degrees and objects that one precedes the other when it bears $>$ to it.)

$$\mathbf{P}_1, \dots, \mathbf{P}_n, \mathbf{t} \models \mathbf{C} \text{ is not true for any finite } n. \text{ But } \{\mathbf{P}_n : n = 1, 2, \dots\}, \mathbf{t} \models \mathbf{C}.$$

Proof: The \mathbf{P}_n successively build up a description of a model of order type at least $\omega + 1$. \mathbf{P}_1 requires that the ordering be complete. Each finite

subset of P_i is true in a model of exactly n elements, the n elements required to instantiate the variables $x_1 \dots x_n$ in the greatest P_n in the set. This ordering is consistent with an ordering for P'' which falsifies **C** (for example the ordering which gives all the x_i and y equal separations, and thus by **C5** equates no two pairs (x, y) (x, z)). On the other hand each model of all the P_i will consist of infinitely many individuals x_1, x_2, \dots ordered by P' , instantiating the x_i , all preceding an individual y instantiating y . Consider the sequence of degrees in P of the x_i . They are ordered by $>$ and have a least upper bound according to the completeness assumption on orderings of Section 4. This least upper bound δ satisfies the antecedent of **t** and thus by **t** the separation between it and x_1 and x_2 are identical. So all orderings for P are constrained by (a)(ii) of the satisfaction definition to make **t** true they will all make these separations identical, and so **C** will be satisfied.

As a result, given that the system is sound, we may conclude that it is incomplete. Although sets of validities like those of Section 2 will entail at least some of the sentences true under all interpretations, no recursive derivation procedure will connect all finite sets of premises and all finite sets of context information sentences with all sentences they semantically entail. Compactness fails largely because of condition **C5** on the H relation, in combination with the assumption of a fairly rich set of degrees. Though this is a fairly mild condition, alternatives are conceivable. But it seems to me very unlikely that any relation between positive, comparative, and hypercomparative which respects intuitions both about ordering and about cardinality will have very different effects.

6. INDICES, INFORMATION, VAGUENESS

The incompleteness of **H** is interesting, but it is not the most important thing about the semantics of Sections 2 and 4. The most important thing there is the way in which contextual information is specified in full sentences instead of in indices. The argument for incompleteness depends on the interaction between contextual information sentences and premise-and-conclusion sentences. (The method has a general resemblance to supervaluation in that truth requires that a sentence be satisfied by all relations between $>$ and $>'$ that satisfy the *a priori* constraints and the contextual information.)

A treatment of comparatives and degrees should say something about vagueness. The groundedness assumption of Section 2 in effect says that every predicate has a perfectly precise hypercomparative at some level. And in terms of these non-vague extensions vague predicates of higher or

lower comparativity are interpreted (relative to context). At these levels versions of the standard symptoms of vagueness can appear. The truth values of sentences depend on how borderlines are specified. The effects of some borderline specifications reveal the underlying hypercomparatives of the predicates in question. For example the assumption that a particular purple-red object is redder than a particular orange-red one can determine the borderline for 'redder' in such a way that a third object is, in context, redder than a fourth. In the formal system presented here failure to specify borderlines results in falsehood, but it is not hard to see how truth value gaps could appear in a richer situation, in which for example the contextual information sentences contain indexicals with undetermined reference or refer to one another in ungrounded loops.

But one important feature of vagueness does not always appear in a standard guise. Higher order vagueness – the indeterminacy of which cases are borderline – disappears as a semantic phenomenon, and reappears as an epistemic one. Where a borderline occurs is typically not knowable in the absence of contextual information, but even given such information deducing which objects lie on which sides of the borderline may be limitingly impossible. Sorites paradoxes look rather different in this light. (On higher order vagueness see Wright 1992, and chapters 4 and 5 of Williamson 1994.)

The crucial fact is the variety of possible contextual information (the t in $\sigma, M, t \models s$). Some information is *specific*, determining a borderline – for example ' o is tall and nothing less tall than o is'. But most often the information is not specific. There are three importantly different ways in which the information can fail to be specific. Some information is *indefinite*, entailing the existence but not the position of borderlines – for example 'there is a degree between that of o and that of n which divides tall from non-tall'. Some information is *non-committal* specifying nothing about borderlines – for example 'if o is tall then p is'. Some information is *buck-passing*, specifying borderlines which are definite only relative to yet further information – for example 'tall people are many centimetres taller than o '.

Different natural language predicates have different typical contextual information. At one extreme there are minimally vague predicates with very determinate borderlines. An example is the predicate 'first class' as applied to British final exam performance. One way of understanding it is in terms of an underlying comparative 'better mark than' defined over percentage grades. Then a standard convention interprets all marks from 70 up as non-borderline firsts, all marks from 68 down as non-borderline non-firsts, and 69 as borderline. There is thus no indeterminacy in the placement

of borderlines. At the other extreme there are Sorites-inducing predicates like 'is bald'. This is very plausibly related to an underlying fairly precise comparative 'is more bald than'. Then typical contextual information almost compulsorily includes the buck-passing clause 'someone is bald if they do not have many hairs on their head'. And then 'many' has a very complex context-dependence, being affected in part by the purposes of the speakers. Among their purposes may be something involving baldness, leading back in a potentially ungrounded circle. Moreover the contextual information can be very varied; it may specify that particular people are and are not bald, and it may contain facts about baldness that everyone takes as common knowledge, for example that anyone with twenty or fewer hairs on their head is bald, no matter what, and anyone with more than a million is not.

Where contextual information is not buck-passing, borderlines are not vague either. And to that extent there is – on this account – no second-order vagueness. But borderlines may be under-specified, by indefinite or non-committal information. The effects of this can be much like that of vagueness. For given contextual information the truth value (if any) of a particular predication depends on the consequences of that information. Consequences are hard to work out, and where the information is vague or non-committal there is a problem knowing what information to reason from. Moreover, if the logic which determines truth values is like **H** then it is incomplete. So we may not be able to rely on mechanical deduction. And when contextual information is buck-passing drawing consequences may be even harder, because of ungrounded cycles.

Though the effects of vagueness and of under-specification are similar there is a crucial difference: vagueness is a semantic matter, while under-specification is an epistemic one. For the kinds of predicates this paper is concerned with, then, Sorites paradoxes are likely to be resolved by exploiting the difference between truth and knowledge. To end the paper, let me sketch how this must run. (For a more thoroughly semantic treatment see Cleave 1987, and for an incisive analysis of the difference between epistemic and semantic factors in vagueness see Williamson (1992) and (1994).)

Sorites paradoxes are generated by two assumptions, *poisoned induction* and *attainability*. An analysis of the Sorites paradox must provide an interpretation on which they are consistent, but on which the contradiction does in fact *seem* to flow from them. That is, the analysis must show why the derivation is so plausible, without making it actual. (See Peacocke 1981, Rolf 1980.)

Poisoned induction says that the value of a vague predicate does not change with small enough increments on some dimension. One more grain doesn't make a heap. And attainability says that enough such increments will lead from a case where the predicate clearly applies to one where it clearly does not.

There is no problem construing attainability so that it is consistent with vagueness. For example in the semantics of Sections 4 and 6 if for an interpretation M contextual information sets a P /not- P borderline at δ , then $P(a)$ and $\neg P(b)$ will hold whenever $J(a, P) > \delta$ and $J(b, P) < \delta$ (where J is the degree-assigning function of the semantics for \mathbf{H}). A plausible construal of poisoned induction goes deeper into the features of the present analysis. It can be stated as follows, where f is a function that takes one 'slightly' higher up the ordering of degrees:

- (PI) if for given M, σ, \mathbf{t}, S is a finite statement of the content of M and t such that $P(a)$ may be deduced from S and b is slightly more P than a , then it is not the case that $\neg P(b)$ can be deduced from S .

Using a plausible interpretation of 'slightly' and writing \vdash for some relation of deductive consequence this could be expressed as

- (PI') If $M, \mathbf{T} \vdash P(a)$ and $M, \mathbf{T} \vdash \neg MCP(b, a)$ then not $M, \mathbf{T} \vdash \neg P(b)$.

The semantics of 2 and 4 can satisfy (PI') for many M, \mathbf{T}, P for many reasonable \vdash . Given the incompleteness of \models , (PI') is almost inevitable, in fact. And the informal grounds for (PI) are very strong: information that tells you that there is a borderline, and that one object is P , need not tell you whether another fairly similar object is P . (PI') is itself vague, in that deduction relations \vdash come in stronger or weaker forms, approaching nearer or less near to \models . Accepting this vagueness as inevitable, there is another formulation which depends less on the incompleteness of the logic.

- (PI') If $M, \mathbf{T} \vdash_n P(a)$ and $M, \mathbf{T} \vdash_n \neg MCP(b, a)$ then not $M, \mathbf{T} \vdash_n \neg P(b)$.

where \vdash_n means 'can be deduced using deduction relation \vdash in n steps'.

Stronger and subtler forms of poisoned induction can be formulated. And some of them come nearer to capturing the intuitions that make sorites paradoxes so gripping for predicates such as 'heap' and 'bald'. The semantics of this paper allow one to describe, but not to formalize, one such.

Formulate poisoned induction as follows

(PI⁺) if $\sigma, M, \mathbf{t}(t) \models P(a)$ and $\sigma, M, \mathbf{t}(t^+) \models \neg MCP(b, a)$ then $\sigma, M, \mathbf{t}(t^+) \models P(b)$

where (i) $\mathbf{t}(t)$ is the context specifying information at the time at which ‘ $P(a)$ ’ is uttered and $\mathbf{t}(t^+)$ that at an immediately later time, and (ii) it is not known that $\sigma, M, \mathbf{t}(t) \models \neg P(b)$.

This is consistent with both the existence of precise borderlines *at each moment* and the attainability principle. (It is also of course consistent with the existence of underspecified borderlines at each moment.) For it represents a form of what David Lewis (1983) calls ‘accommodation’: when ‘ $P(a)$ ’ is uttered the context changes in such a way that ‘ a ’ instead of denoting a near-to-the-edge instance of P is now a more central case. (a is the same; the edge moves.)

And, to my mind, that is the real solution to the paradox. For predicates such as ‘bald’ the bare contextual information – the information that does not include what sentences have been asserted recently – is not enough to determine very many truth values. So we rely on additional information about what has been asserted, and principles of accommodation. But then the assertion that $P(a)$ changes the context in which the assertion of $P(b)$ is evaluated. The informal ideas of this paper allow one to state this solution in outline. But the formal system cannot represent it. To do that one would have to incorporate into the context-specifying language **I** time references, epistemic operators, and something like David Kaplan’s logic of indexicals (Kaplan 1989). Ideally one would combine a logic of indexicals, the ideas of this paper, the context-sensitive models of Kamp (1981), and the ‘logic of clarity’ of Williamson (1994). That is a project worth undertaking. But what makes the project even describable is a simple methodological point: specify contexts with sentences rather than indices.

NOTE

* I have been helped by comments from John Broome, Chris Swoyer, John Cleave, Stephan Körner, and Christopher Williams. The referees’ comments were extremely detailed and constructive.

REFERENCES

Aqvist, L.: 1981, ‘Predicate Calculi with Adjectives and Nouns’, *Journal of Philosophical Logic* **10**, 1–26.

- Clark, M.: 1984, 'Degrees of Comparison', *Analysis* **44**, 178–80.
- Cleave, J.: 1987, 'Logic and Inexactness', in J. Szrdnicki (ed.), *Stephan Körner – Philosophical Analysis and Reconstruction*, Nijhoff, Dordrecht.
- Fine, K.: 1975, 'Vagueness, Truth, and Logic', *Synthese* **30**, 265–300.
- Kaplan, D.: 1989, 'Demonstratives', in J. Almog, J. Perry, and H. Wettstein (eds.), *Themes from Kaplan*, Oxford University Press, New York.
- Kamp, H.: 1981, 'The Paradox of the Heap', in U. Mönnich (ed.), *Aspects of Philosophical Logic*, Reidel, Dordrecht.
- Klein, E.: 1980, 'A Semantics for Positive and Comparative Adjectives', *Linguistics and Philosophy* **4**, 1–45.
- Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky: 1971, *Foundations of Measurement*, Vol. 1, Academic Press, New York.
- Lewis, D.: 1983, 'Scorekeeping in a Language Game', in *Philosophical Papers*, Vol. 1, Oxford University Press, Oxford.
- Morton, A.: 1991, *Disasters and Dilemmas*, Basil Blackwell, Oxford.
- Morton, A.: 1994, 'Two Places Good Four Places Better', *Proceedings of the Aristotelian Society*, Suppl. **68**, 187–98.
- Peacocke, C.: 1981, 'Are Vague Predicates Incoherent?', *Synthese* **46**, 121–41.
- Rolf, B.: 1980, 'A Theory of Vagueness', *Journal of Philosophical Logic* **9**, 315–25.
- Stechow, A. von: 1983, 'Comparing Semantic Theories of Comparison', *Journal of Semantics* **3**, 1–77.
- Vendler, Z.: 1967, 'The Grammar of Goodness', in *Linguistics in Philosophy*, Cornell University Press, Ithaca.
- Wheeler, S.: 1972, 'Attributives and their Modifiers', *Nous* **6**, 310–34.
- Williams, C. J. F. and A. Morton: 1984, '"Comparatives" and "Comparatives and Degrees"', *Analysis* **44**, 15–20.
- Williamson, T.: 1992, 'Inexact Knowledge', *Mind* **101**.
- Williamson, T.: 1994, *Vagueness*, Routledge, London.
- Wright, C.: 1992, 'Is Higher Order Vagueness Coherent?', *Analysis* **52**, 129–39.

Department of Philosophy
 University of Bristol
 Bristol BS8 1TB
 United Kingdom