Vagueness and the Sorites

Robert van Rooij

1 Partial logic

Definition 1.
Let $\mathcal{L}$ be a predicate logical language. A (partial) model $\mathcal{M}$ for $\mathcal{L}$ is an ordered pair $\langle \mathcal{D}, \mathcal{I} \rangle$, where

1. *$\mathcal{D}$ is a non-empty set, the* domain *of the model.*
2. *For all individual constants it holds that $\mathcal{I}(a)$ is an element of $\mathcal{D}$. We assume for simplicity that for each object d of $\mathcal{D}$ there exists at least one individual constant a such that $\mathcal{I}(a) = d$.*
3. *for all n-ary predictate P we assume: $\mathcal{I}(P)$ is a partial function from $\mathcal{D}^n$ in $\{0, 1\}$.*

Definition 2.
Let $\mathcal{M} = \langle \mathcal{D}, \mathcal{I} \rangle$ and $\mathcal{M}' = \langle \mathcal{D}, \mathcal{I}' \rangle$ be two partial models with both the same domain. $\mathcal{M}'$ is a refinement *of $\mathcal{M}$ iff the following conditions are met:*

(a) If P is n-ary and $\mathcal{I}(P)(\langle d_0 \ldots d_n \rangle) = 1$, then $\mathcal{I}'(P)(\langle d_0 \ldots d_n \rangle) = 1$;

(b) If P is n-ary and $\mathcal{I}(P)(\langle d_0 \ldots d_n \rangle) = 0$, then $\mathcal{I}'(P)(\langle d_0 \ldots d_n \rangle) = 0$.

The principle of stability now says that if sentence φ is true/false in a model $\mathcal{M}$, then φ has to stay true/false if $\mathcal{M}$ is getting more precise. Formally, let $\mathcal{M}' = \langle \mathcal{D}, \mathcal{I}' \rangle$ be a refinement of $\mathcal{M} = \langle \mathcal{D}, \mathcal{I} \rangle$. Then it has to be the case that for all φ:
(i) If $\mathcal{V}_{\mathcal{M}}(\varphi) = 1$, then $\mathcal{V}_{\mathcal{M}'}(\varphi) = 1$.
(ii) If $\mathcal{V}_{\mathcal{M}}(\varphi) = 0$, then $\mathcal{V}_{\mathcal{M}'}(\varphi) = 0$.

Definition 3.
Let $\mathcal{L}$ be a language of predicate logic and $\mathcal{M} = \langle \mathcal{D}, \mathcal{I} \rangle$ a partial model for $\mathcal{L}$. Then

- $\mathcal{V}_{\mathcal{M}}(Pa_0 \ldots a_n) = 1$ *iff* $\mathcal{I}(P)(\langle \mathcal{I}(a_0), ..., \mathcal{I}(a_n) \rangle) = 1$
  $\mathcal{V}_{\mathcal{M}}(P(a_0 \ldots a_n) = 0$ *iff* $\mathcal{I}(P)(\langle \mathcal{I}(a_0), ..., \mathcal{I}(a_n) \rangle) = 0$

- $\mathcal{V}_\mathcal{M}(\neg\varphi) = 1$ *iff* $\mathcal{V}_\mathcal{M}(\varphi) = 0$
  $\mathcal{V}_\mathcal{M}(\neg\varphi) = 0$ *iff* $\mathcal{V}_\mathcal{M}\varphi) = 1$

- $\mathcal{V}_\mathcal{M}(\varphi \wedge \psi) = 1$ *iff* $\mathcal{V}_\mathcal{M}(\varphi) = 1$ *and* $\mathcal{V}_\mathcal{M}(\psi) = 1$
  $\mathcal{V}_\mathcal{M}(\varphi \wedge \psi) = 0$ *iff* $\mathcal{V}_\mathcal{M}(\varphi) = 0$ *or* $\mathcal{V}_\mathcal{M}(\psi) = 0$

- $\mathcal{V}_\mathcal{M}(\varphi \vee \psi) = 1$ *iff* $\mathcal{V}_\mathcal{M}(\varphi) = 1$ *or* $\mathcal{V}_\mathcal{M}(\psi) = 1$
  $\mathcal{V}_\mathcal{M}(\varphi \vee \psi) = 0$ *iff* $\mathcal{V}_\mathcal{M}(\varphi) = 0$ *and* $\mathcal{V}_\mathcal{M}(\psi) = 0$

- $\mathcal{V}_\mathcal{M}(\varphi \rightarrow \psi) = 1$ *iff* $\mathcal{V}_\mathcal{M}(\varphi) = 0$ *or* $\mathcal{V}_\mathcal{M}(\psi) = 1$
  $\mathcal{V}_\mathcal{M}(\varphi \rightarrow \psi) = 0$ *iff* $\mathcal{V}_\mathcal{M}(\varphi) = 1$ *and* $\mathcal{V}_\mathcal{M}(\psi) = 0$

- $\mathcal{V}_\mathcal{M}(\exists x\varphi) = 1$ *iff* $\mathcal{V}_\mathcal{M}([^a/x]\varphi) = 1$ *for at least one constant a*
  $\mathcal{V}_\mathcal{M}(\exists x\varphi) = 0$ *iff* $\mathcal{V}_\mathcal{M}([^a/x]\varphi) = 0$ *for each contstant a*

- $\mathcal{V}_\mathcal{M}(\forall x\varphi) = 1$ *iff* $\mathcal{V}_\mathcal{M}([^a/x]\varphi) = 1$ *for each constant a*
  $\mathcal{V}_\mathcal{M}(\forall x\varphi) = 0$ *iff* $\mathcal{V}_\mathcal{M}([^a/x]\varphi) = 0$ *for at least one constant a.*

2 Supervalutation

Definition 4.
Let $\mathcal{L}$ be a language of predicate logic. A supermodel $\mathcal{M}$ for $\mathcal{L}$ is an ordered pair $\langle \mathcal{D}, \mathcal{J} \rangle$, where

1. *$\mathcal{D}$ is a non-empty set, the* domain *of the modela;*

2. *$\mathcal{J}$ a (non empty) set of (partial) interpretations such that*

   (a) *every $\mathcal{I} \in \mathcal{J}$ assigns to every constant a an element $\mathcal{I}(a)$ of $\mathcal{D}$;*

   (b) *every $\mathcal{I} \in \mathcal{J}$ assigns to each n-ary predicate P a partial function $\mathcal{I}(P)$ from $\mathcal{D}^n$ in $0, 1$.*

   In addition, the following restriction must be obeyed:

   (a) *for all $\mathcal{I}, \mathcal{I}' \in \mathcal{J}$ and each individual constant a it holds that $\mathcal{I}(a) = \mathcal{I}'(a)$;*

   (b) *there is a $\mathcal{I}_0 \in \mathcal{J}$ such that for all $\mathcal{I} \in \mathcal{J}$ it is the case: $\mathcal{I}$ is a refinement of $\mathcal{I}_0$;*

   (c) *for all $\mathcal{I} \in \mathcal{J}$ there is a $\mathcal{I}' \in \mathcal{J}$ such that $\mathcal{I}'$ is a refinement of $\mathcal{I}$.*

Definition 5 (Truth Definition).
Let $\mathcal{M} = \langle \mathcal{D}, \mathcal{J} \rangle$ be a supermodel, and φ a sentence. By definition it holds that:

- $\mathcal{W}_{\mathcal{M}}(\varphi) = 1$ *iff* $\mathcal{V}_{\langle\mathcal{D},\mathcal{I}\rangle}(\varphi) = 1$ *for* all *total* $\mathcal{I} \in \mathcal{J}$.
- $\mathcal{W}_{\mathcal{M}}(\varphi) = 0$ *iff* $\mathcal{V}_{\langle\mathcal{D},\mathcal{I}\rangle}(\varphi) = 0$ *for* all *total* $\mathcal{I} \in \mathcal{J}$.
- *In all other cases* $\mathcal{W}_{\mathcal{M}}(\varphi)$ *is undetermined.*

3 Dummett's Diagnose

Variants of the Sorites are easy to come by. you need is a vague predicate P, preferably one with a comparative form, the use of which is guided by the following rule:

> If two objects are observationally indistinguishable in the respects relevant to P, then either both satisfy P or else neither of them does.

Following Kamp [1981], we will refer to this principle, which expresses the Equivalence of Observationally Indistinguishable objects, as EOI.

With such a vague predicate P, one can nearly always associate a domain $\mathcal{D}$ and a relation $\mathcal{R}$ with the following properties:

1. $\mathcal{D}$ is nonempty; the objects in $\mathcal{D}$ are similar in kind and (therefore) comparable in the respects relevant to P.
2. $\mathcal{R}$ is the irreflexive and transitive relation on $\mathcal{D}$ consisting of the pairs $\langle d, e\rangle$ of which the first member d is observationally *more* P than the second member e.
3. There are objects d and e in $\mathcal{D}$ such that
   (a) P clearly applies to d;
   (b) P clearly does not apply to e;
   (c) There is a finite sequence of objects $d = d_0, d_1, \ldots, d_{n-1}, d_n = e$ in $\mathcal{D}$ such that for any two successive objects d_i and d_{i+1} in this sequence, neither $d_i\mathcal{R}d_{i+1}$, nor $d_{i+1}\mathcal{R}d_i$.

Let us write $x\mathcal{E}y$ iff neither $x\mathcal{R}y$ nor $y\mathcal{R}x$. Whenever $x\mathcal{E}y$, x and y are observationally indistinguishable in the respects relevant to the predicate P, and therefore it is tempting to read '$x\mathcal{E}y$' as 'x is observationally *just as* P as y'. We will not resist this temptation, even though we are aware that $\mathcal{E}$ need not be an equivalence relation. That is, $\mathcal{E}$ is reflexive and symmetric, but not in general transitive. It may very well occur that there is no discriminable difference between the objects x and y — x is observationally just as P as y — and no discriminable difference between the objects y and z — y is observationally just as P as z — whereas x and z *can* be discriminated — x is observationally more P than z. It is easy to see now how the paradox can arise. The principle EOI forces us to use the predicate P in a way that would only be coherent if the relation $\mathcal{E}$ were transitive.

For, what else does EOI express but:

> (EOI 1) For any $x, y \in \mathcal{D}$, if $x\mathcal{E}y$, then, if P applies to x, P applies to y.

So, you cannot assign the predicate P to d_0 without having to assign it to $d_1, d_2, d_3, \ldots$, and, finally, to d_n as well. EOI forces you to do so, even though $\mathcal{E}$ is not transitive: d_0 and d_n could be so far apart — d_0 has got no hairs, d_n has 150,000 hairs; the temperature of d_0 is 2^o C, the temperature of d_n is 80^o C — that it would seem perfectly all-right to say that d_0 is P — bald, or cold — but that d_n is not. From the above it will be clear that we are in a predicament: either we accept EOI, in which case we are forced to conclude that the use of at least some observational predicates is intrinsically inconsistent, or we do not accept EOI, in which case it seems to follow that there are no truly observational predicates.

Dummett is inclined to choose the first horn of this dilemma, and there is something attractive in this point of view. After all, normally we are dealing with just a few objects, all very well discernable from each other. In those circumstances EOI does *not* give rise to inconsistency; normally, it serves its purpose quite well. Only in exceptional situations — i.e. when we are confronted with sequences of objects as described above — things go wrong. But then, vague predicates like 'cold' and also 'bald' and 'short' and 'heap' are not meant to be used in those situations; what we should use there is other, finer tools; we should no longer talk in terms of 'cold', 'short' etc., but in terms of degrees centigrade or millimeters or grains of sand. The Sorites Paradox typically arises when the coarse tools of vague predicates are used *side by side* with these finer tools, i.e. when one starts saying things like 'if someone with *23567* hairs on his head is not *bald*, then neither is someone with *234566* hairs on his head'. As such, it merely reflects that one should not use the coarse tools in circumstances where only other, more delicate, ones are applicable. [1]

Dummett's position seems a strong one, and, if a choice between the true observational character of at least some predicates and the consistency of lan-

[1]Note that this solution of the Sorites Paradox is completely in line with Wittgenstein's *Philosophische Untersuchungen*. See in particular sections 85-87: 'A rule stands like a sign-post . . . The sign-post is in order–if, under normal circumstances, it fulfils its purpose'.

guage really cannot be avoided, it is the only position an empiricist can take. But there is a way out. One can loosen the ties between vague predicates and observation, without having to cut them. How this can be done will be explained right now. [2]

Let us return to the observational relation $\mathcal{R}$. The problem is that its resolution is too poor: a basketball player x can *in fact* be shorter than a basketball player y, without *observationally* being so. Or, to phrase it more carefully, it may be that x cannot be seen to be shorter than y by the naked eye, while a less crude measurement shows that he is. Our powers of discrimination are limited; that is why observational equivalence is not transitive. We can, however, define a relation $\mathcal{R}^{\mathcal{D}}$ from $\mathcal{R}$ which, as it were, increases the resolution:

Definition 6.
Let $\mathcal{R}$ be an irreflexive binary relation on a domain $\mathcal{D}$. Define the sharpening $\mathcal{R}^{\mathcal{D}}$ *of* $\mathcal{R}$ relative to $\mathcal{D}$ *by:* $x\mathcal{R}^{\mathcal{D}}y$ *if and only if*
either (i) $x, y \in \mathcal{D}$ *and there is some* $z \in \mathcal{D}$ *such that* $z\mathcal{R}y$ *and not* $z\mathcal{R}x$*,*
or (ii) $x, y \in \mathcal{D}$ *and there is some* $z \in \mathcal{D}$ *such that* $x\mathcal{R}z$ *and not* $y\mathcal{R}z$*.*

Suppose you are confronted with three objects d_1 and d_2, and d_3; you cannot discern any difference in length between d_1 and d_2; you cannot discern any difference in length between d_2 and d_3 either; but you can discern a difference in length between d_1 and d_3: d_1 is observationally shorter than d_3. Wouldn't you then conclude that there *must* be a difference in length between d_1 and d_2 even though you cannot see it by the naked eye? Wouldn't you be inclined to say then that d_1 is *in fact* shorter than d_2, even though this is not observationally so? The following facts, the proofs of which are left as an *exercise*, are noteworthy:

- $\mathcal{R} \subseteq \mathcal{R}^{\mathcal{D}}$;
- $\mathcal{R}^{\mathcal{D}}$ is irreflexive;
- $\mathcal{R}^{\mathcal{D}}$ is transitive if $\mathcal{R}$ is.

Hence, if $\mathcal{R}$ is the relation 'being visibly shorter than', we are really allowed to think of $\mathcal{R}^{\mathcal{D}}$ as a sharpening of that relation.

- $(\mathcal{R}^{\mathcal{D}})^{\mathcal{D}} = \mathcal{R}^{\mathcal{D}}$.

Hence, if we were to call an irreflexive relation *sharp* if sharpening it in the manner described above leaves it unchanged, then we could prove $\mathcal{R}^{\mathcal{D}}$ to be sharp.

[2]The basic ideas goes back to Russell and Goodman.

- If $\mathcal{D} \subseteq \mathcal{D}'$, then $\mathcal{R}^{\mathcal{D}} \subseteq \mathcal{R}^{\mathcal{D}'}$.

Actually, the resolution may *increase* if elements are added to $\mathcal{D}$: it is very well possible that two objects that cannot be discriminated within $\mathcal{D}$, can be discriminated within $\mathcal{D}'$.

- Define a relation $\mathcal{E}_{\mathcal{D}}$ from $\mathcal{R}^{\mathcal{D}}$ in a way analogous to the way $\mathcal{E}$ is defined from $\mathcal{R}$: $x\mathcal{E}_{\mathcal{D}}y$ iff neither $x\mathcal{R}^{\mathcal{D}}y$ nor $y\mathcal{R}^{\mathcal{D}}x$. Then we find that even though $\mathcal{E}$ is not necessarily transitive, $\mathcal{E}_{\mathcal{D}}$ is. Indeed, $\mathcal{E}_{\mathcal{D}}$ is an *equivalence* relation.

So, even if you do not want to call $\mathcal{R}^{\mathcal{D}}$ sharp — after all, adding elements to $\mathcal{D}$ may make it sharper, the least you will have to say is that it is sharp *enough* to avoid the kind of problems we got with $\mathcal{R}$. That is, if we replace the formalization (1) of the principle EOI given above by the following new formalization:

> (EIO 2) For any $x, y \in \mathcal{D}$, if $x\mathcal{E}_{\mathcal{D}}y$ and P applies to x, then P applies to y,

contradictions need no longer arise. Suppose you are presented with a series of objects as described in the beginning of this section. Given the first formulutuion of te principle of EIO (EIO 1), you cannot assign the predicate P to d_0 without having to assign it to $d_1, d_2, d_3, \ldots$, and finally to d_n as well. But given the second formulation (EIO 2), you need not be led into contradicting yourself. Clearly, there must be three objects d_i, d_{i+1}, and d_k in the series such that $d_i\mathcal{R}d_k$, but not $d_{i+1}\mathcal{R}d_k$. Choose such d_i, d_{i+1}, and d_k,and you can deny the truth of the premise that if d_i is P, d_{i+1} is — and this without giving up the principle EOI: d_i and d_{i+1} are *observationally* distinguishable in a respect relevant to P, since the purely observational predicate $x\mathcal{R}d_k$ applies to d_i, but not to d_{i+1}.

Are our new relations $\mathcal{R}^{\mathcal{D}}$ and $\mathcal{E}_{\mathcal{D}}$ really observational? Well, they are clearly not *directly* observational like the relation $\mathcal{R}$. In many cases it is impossible to decide on the basis of *direct* observation whether $\mathcal{R}^{\mathcal{D}}$ (or $\mathcal{E}_{\mathcal{D}}$) applies to a pair of objects or not. This is because in the definition of $\mathcal{R}^{\mathcal{D}}$ (and hence in the definition of $\mathcal{E}_{\mathcal{D}}$) a quantification occurs. But, while $\mathcal{R}^{\mathcal{D}}$ and $\mathcal{E}_{\mathcal{D}}$ are not directly observational, they are certainly observational in the sense that they are defined from a directly observational relation $\mathcal{R}$ by means of logic alone. Every statement containing references to $\mathcal{R}^{\mathcal{D}}$ or $\mathcal{E}_{\mathcal{D}}$ can be translated into an equivalent expression in which no reference is made to relations other than $\mathcal{R}$. Therefore, if (1) is replaced by (2), the links between the predicate P and direct observation are not severed; the principle that the use of P should be guided by *mere* observation is just replaced by the principle that it should be guided by observation *aided by reason.*

Side Remark

In the above, we have not made any use of the properties of the relation $\mathcal{R}$, except of its irreflexivity; so the method is very general. But in many cases — and here the relation 'being visibly shorter than' may serve as an example — $\mathcal{R}$ will satisfy certain extra principles, which have been stated in **?**:

Definition 7 (Luce's Axioms).L1 *For no $x \in \mathcal{D}, x\mathcal{R}x$;*

L2 *For any $w, x, y, z \in \mathcal{D}$, if $w\mathcal{R}x$ and $y\mathcal{R}z$, then either $w\mathcal{R}z$ or $y\mathcal{R}x$;*

L3 *For any $x, y \in \mathcal{D}$, if some $z \in \mathcal{D}$ is such that $x\mathcal{R}z$ and $z\mathcal{R}y$, then every $z \in \mathcal{D}$ will be such that either $x\mathcal{R}z$ or $z\mathcal{R}y$.*

Note that **L1** and **L2** imply transitivity. Think of $\mathcal{R}$ as the relation 'being visibly shorter than' on the domain of basketballplayers. Then one way to see what these axioms mean is to assume that at every moment there is a real number c such that you judge a basketball player x to be shorter than a basketballplayer y just in case y's length exceeds x's length by c millimeters or more. So, instead of $x\mathcal{R}y$ we can write $length(x) + c \leq length(y)$. It is easy to check that this last relation satisfies Luce's Axioms. Once you accept the validity of Luce's Axioms for the relation 'being visibly shorter than', you get certain desirable properties for the sharpening of this relation in the bargain. The relation $\mathcal{R}^{\mathcal{D}}$ turns out to be almost-connected, i.e. for any $x, y, z \in \mathcal{D}$ the following holds: if $x\mathcal{R}^{\mathcal{D}}y$, then $x\mathcal{R}^{\mathcal{D}}z$ or $z\mathcal{R}^{\mathcal{D}}y$. Therefore, the $\mathcal{E}_{\mathcal{D}}$-equivalence classes are *linearly* ordered by the relation that holds between two $\mathcal{E}_{\mathcal{D}}$-equivalence classes $\{x \in \mathcal{D} | x\mathcal{E}_{\mathcal{D}}d\}$ and $\{x \in \mathcal{D} | x\mathcal{E}_{\mathcal{D}}e\}$ just in case $d\mathcal{R}^{\mathcal{D}}e$. [3]

4 Contextual Resolution

So far we have paid hardly any attention to what — apart from vagueness — is the most salient feature of the predicates that can be used in a Sorites type of argument, namely their context dependency. At least two different kinds of

[3] To be more precise, the structure $\langle \mathcal{D}, \mathcal{R}^{\mathcal{D}}, \mathcal{E}_{\mathcal{D}} \rangle$ is what Hempel has called a *quasi-series*:

1. $\mathcal{R}^{\mathcal{D}}$ is irreflexive and transitive;
2. $\mathcal{E}_{\mathcal{D}}$ is an equivalence relation;
3. for any $x, y \in \mathcal{D}$, either $x\mathcal{R}^{\mathcal{D}}y$, or $y\mathcal{R}^{\mathcal{D}}x$, or $x\mathcal{E}_{\mathcal{D}}y$

Now, for each $d \in \mathcal{D}$, define $[d]$ to be $\{x \in \mathcal{D} | x\mathcal{E}_{\mathcal{D}}d\}$ and define: $[d] < [e]$ iff $d\mathcal{R}^{\mathcal{D}}e$. This is well-defined. Set $\mathcal{D}/\mathcal{E}_{\mathcal{D}} = \{[d] | d \in \mathcal{D}\}$; then the structure $\langle \mathcal{D}/\mathcal{E}_{\mathcal{D}}, < \rangle$ is a linear order. Such linear orderings form the basic input for any form of quantitative measurement.

context-dependency are involved. The first kind results from the phenomenon that the question whether or not a certain object may be called P (tall, warm, yellow), does not only depend on the object itself, but also on the set of objects it is compared with. What is tall, or warm, or yellow in one context can be short, or cold or orange in another. For example, if you were asked to select the short strokes from the right picture below, you might very well want to include stroke b, whereas in the left picture, the very same stroke b may be said to be one of the long strokes.

Note that it is not the fact that the capacities of our senses are limited and the resulting poor resolution of observational relations that causes this behaviour of vague predicates. The relative lenghts of the strokes drawn are clear enough, still stroke b may be called long in one picture but not in the other.

We may contrast this property of vague predicates with a second kind of context-dependency that does result from the fact that the resolution of an observational relation may depend on the domain of discourse. Suppose that you are presented with two different colour patches a and b not visibly differing in shade (not $a\mathcal{R}b$). Further suppose that a and b are the only objects in your domain of discourse. You might be quite at a loss to answer the question whether a is red or not, but even so EOI forces you to accept that if a is red, b is red too. This is because $a\mathcal{E}_{\{a,b\}}b$ holds. Now, let's change the domain of discourse and add a third individual c to it, such that a and c are, but b and c are not, visibly distinguishable. As soon as this addition is made, a and b are no longer equivalent (not $a\mathcal{E}_{\{a,b,c\}}b$ and so the conditional is no longer forced upon you. Resolution generally depends on the number of comparable objects; the smaller this number, the less resolution. In the worst case, when there are only two objects in the domain, we are at the level of directly observational relations: $a\mathcal{E}_{\{a,b\}}b$ iff $a\mathcal{E}b$.

It will now be clear what our amendment to Dummett's analysis is: Dummett is right in saying that if there is no discriminable difference between two objects, then either P applies to both or to neither of them; but it should be added that this holds only in contexts where no objects other than these two are at stake. If the reference group gets larger the objects may *become* distinguishable.

In ordinary discourse domains of evaluation seldomly stay fixed. Objects are unceasingly being introduced into these domains and taken out again. This can happen either by means of linguistic acts like the uttering of a sentence or by non-linguistic acts like pointing at things, covering or uncovering them, etcetera. One may exploit this phenomenon to obtain versions of the Sorites paradox. The following is a quotation from **?**:

In front of us is a large screen. Its extreme left is green, its extreme right yellow, and there is a gradual transition from the one colour to the other. The screen is subdivided into many small squares, so small that each square appears to have a uniform hue and moreover the colours of no two adjacent squares can be distinguished by sight. Compare the following two experiments.
1) We are both facing the screen which is entirely visible to you. I begin by pointing at a little square on the extreme left and ask you what its colour is. If you are not colourblind you will surely answer: 'green'. I then point to the adjacent little square on the right and ask the same question. Probably you will again say 'green' Then I point at the square to the right of this one, and so on. After a while your answers 'green' will become hesitant, increasingly so, until the point is reached where you either say: 'Now I don't know what to say any more', or else some such thing as 'this one really looks more like yellow'.
2) This time the big screen is completely covered. I ask the same question about the same squares in the same order. But now I proceed as follows. When I ask my first question I *only* uncover the first little square. After you have answered I reveal the square next to it. Then, after your second answer I cover the first square and uncover the third; after you have made your third reply, I cover the second and uncover the fourth, etc.

How will the outcomes of these two experiments compare? In all likelyhood you will carry on 'green ' for a longer time in the second trial than in the first.

With the apparatus developed thus far, we can explain what is happening here. In the first experiment, where you have all the patches before you, the domain of discourse is relatively large and resolution is accordingly good. As soon as you start to feel unsure about the greenness of the square currently being pointed at, you will have no trouble in finding a square differentiating it from its predecessor. You may therefore feel free to deny the present's square being green without any violation of the principle EOI. In the second experiment, on the other hand, the differentiating objects are systematically being made unavailable to you, and EOI forces you to keep on saying 'green', or so it seems. Let us to put the above considerations together in a definition, which spells out what constraints there are if you have to decide to which objects in a certain set C - this set determines the *context* — the predicate P applies and which objects in C it does not.

Definition 8.

Let $\mathcal{D}$ and $\mathcal{R}$ be like above and let C be any subset of $\mathcal{D}$. An admissable interpretation *for P in C is any function $\mathcal{I}_C(P)$ from C into $\{0,1\}$ that meets the following conditions:*

(EOI) *If $x\mathcal{E}_C y$, then $\mathcal{I}_C(P)(x) = \mathcal{I}_C(P)(y)$;*

(MON) *If $x\mathcal{R}^C y$ and $\mathcal{I}_C(P)(y) = 1$, then $\mathcal{I}_C(P)(x) = 1$;*

(DIS) *If there are $x, y \in C$ such that $x\mathcal{R}y$, then there are $x, y \in C$ such that $\mathcal{I}_C(P)(x) \neq \mathcal{I}_C(P)(y)$.*

Note that $\mathcal{I}_C(P)$ is a total function. Hence, if $x\mathcal{R}^C y$ and $\mathcal{I}_C(P)(x) = 0$, then $\mathcal{I}_C(P)(y) = 0$.

The definition leaves a lot of freedom, which is part of the reason why we call a predicate P whose use is guided by these principles vague.

Cross contextual constraints

Once you have chosen an (admissable) interpretation for P in C, this puts some further constraints on the decisions you can take when some *new* objects are added to C. You cannot start all over again when C is extended to B; not any old admissable interpretation of P in the context B is coherent with the interpretation you have chosen in context C.

More precisely:

Definition 9.
Let $C \subseteq B$, $\mathcal{I}_C(P)$ an admissable interpretation of P over C, and $\mathcal{I}_B(P)$ an admissable interpretation of P over B. Then $\mathcal{I}_B(P)$ is coherent with *$\mathcal{I}_C(P)$ iff the following two conditions are fulfilled:*

(+ -) *There are $d \in C$ such that $\mathcal{I}_C(P)(d) = 1$ and $\mathcal{I}_B(P)(d) = 0$ only if there are $e \in B \setminus C$ such that $\mathcal{I}_B(P)(e) = 1$;*

(- +) *There are $d \in C$ such that $\mathcal{I}_C(P)(d) = 0$ and $\mathcal{I}_B(P)(d) = 1$ only if there are $e \in B \setminus C$ such that $\mathcal{I}_B(P)(e) = 0$.*

Examples

Let $\mathcal{D} = \{1,2,3,4,5,6,7,8,9,10,11,12\}$, and set $d\mathcal{R}e$ iff $d < e$. Consider the sets $C = \{1,3,4,7,9,10\}$, and $B = \{1,3,4,7,9,10,12\}$ The following are all admissable interpretations for the predicate *small.* (Here we set a number in boldface when it is small according to the interpretation in question).

$$\mathcal{I}_C(P) : \{\mathbf{1,\ 3}, 4, 7, 9, 10\}$$

$$\mathcal{I}'_C(P) : \{\mathbf{1,\ 3,\ 4}, 7, 9, 10\}$$

$$\mathcal{I}_B(P) : \{\mathbf{1,3}, 4, 7, 9, 10, 12\}$$

$$\mathcal{I}'_B(P) : \{\mathbf{1,3,4}, 7, 9, 10, 12\}$$

Starting from $\mathcal{I}_C(P)$, both $\mathcal{I}_B(P)$ and $\mathcal{I}'_B(P)$ are coherent, but starting from $\mathcal{I}'_C(P)$, only $\mathcal{I}'_B(P)$ is. That is, if you have first decided that 1, 3, and 4 are the small numbers in the set 1,3,4,7,9,10, you are not allowed , when 12 is added to this set, to call only 1 and 2 small.

Let $B \subseteq C$, $\mathcal{I}_B(P)$ an admissable interpretation of P in context B, and $\mathcal{I}_C(P)$ an admissable interpretation of P in context C that is coherent with $\mathcal{I}_B(P)$. It will be clear that the rules given above do not guarantee that

$$\{d \in B | \mathcal{I}_B(P)(d) = 1\} \subseteq \{d \in C | \mathcal{I}_C(P)(d) = 1\}$$

An informal example, which shows that this is how it should be, is easy to find: take B the set of basketballplayers, C the set of human beings, and P the predicate ‘short’. All basketballplayers are human beings, but not all short basketplayers are short human beings.

We are ready now to specify a truthdefinition for a language $\mathcal{L}$ that has as its logical vocabulary $\neg$ (for negation), and $\rightarrow$ (for implication), and as its non-logical vocabulary *one* vague predicate P, and individual constants $a_0, a_1, a_2, \ldots$.

Definition 10.
A coherent model for this language $\mathcal{L}$ is a triple $\mathcal{M} = \langle \mathcal{D}, \mathcal{R}, \mathcal{I} \rangle$ where

1. *$\mathcal{D}$ and $\mathcal{R}$ have the properties described before;*
2. *$\mathcal{I}$ is a function that assigns*
   (a) *an element of $\mathcal{D}$ to every individual constant a;*
   (b) *an admissable interpretation $\mathcal{I}_C(P)$ for P in C to each C; whenever $C \subseteq C'$, $\mathcal{I}_{C'}(P)$ must be coherent with $\mathcal{I}_C(P)$.*

We will say that a context C is *decisive* for a sentence φ if and only if $\mathcal{I}(a) \in C$ for every individual constant a occurring in φ. The truthdefinition is given in two steps. First it is explained what it means for a sentence φ to be true in a context C that is decisive for φ.

- $\mathcal{V}_\mathcal{M}(P(a), C) = 1$ iff $\mathcal{I}_C(P)(\mathcal{I}(a)) = 1$;
- $\mathcal{V}_\mathcal{M}(\neg\varphi, C) = 1$ iff $\mathcal{V}_\mathcal{M}(\varphi, C) = 0$;

- $\mathcal{V}_{\mathcal{M}}(\varphi \to \psi, C) = 1$ iff $\mathcal{V}_{\mathcal{M}}(\varphi, C) = 0$ or $\mathcal{V}_{\mathcal{M}}(\psi, C) = 1$.

In the second step we want to extend the above to contexts that are not decisive.

Let Δ be a set of sentences, and $\mathcal{M} = \langle \mathcal{D}, \mathcal{R}, \mathcal{I} \rangle$ a coherent model and $C \subseteq \mathcal{D}$. Set

$C_\Delta = C \cup \{\mathcal{I}(a) | a$ occurs somewhere in the sentences of $\Delta\}$.

Notice that C_Δ is the minimal extension of C that is decisive for all sentences in Δ.

Now we can stipulate that for a given model $\mathcal{M}$ an arbitrary context C and an arbitrary set Δ of sentences the following will hold:

$$\mathcal{V}_{\mathcal{M}}(\Delta, C) = 1 \text{ iff for every } \varphi \in \Delta, \mathcal{V}_{\mathcal{M}}(\varphi, C_\Delta) = 1$$

Presumably, you wonder why the second step of the truthdefinition is formulated not for sentences but for sets of sentences. The following example should help to you to appreciate this manoeuvre.

Example

Consider any model $\mathcal{M} = \langle \mathcal{D}, \mathcal{R}, \mathcal{I} \rangle$ such that

1. $\mathcal{D} = \{d_0, d_1, \ldots, d_{1000}\}$;
2. for any i, j, $d_i \mathcal{R} d_j$ iff $i < j + 50$;
3. $\mathcal{I}(a_i) = d_i$ for every i;
4. If $d_0 \in C$, then $\mathcal{I}_C(P)(d_0) = 1$; if $d_{1000} \in C$, then $\mathcal{I}_C(P)(d_{1000}) = 0$

Then the following holds:

$\mathcal{V}_{\mathcal{M}}(\{P(a_i) \to P(a_{i+1})\}, \varnothing) = 1$ for every $i < 1000$;

$\mathcal{V}_{\mathcal{M}}(\{P(a_i) \to P(a_{i+1}) | i < 1000\}, \varnothing) = 0$.

In other words, a speaker is forced to accept each of the sentences $P(a_i) \to P(a_{i+1})$ taken in isolation. Still, this does not mean that he or she has to accept all these sentences taken together. This is the key to our solution of the Sorites. When a decision has to be taken concerning, say, $P(a_{27}) \to P(a_{28})$ taken in isolation, the objects d_{27} and d_{28} are to be compared in a context — $\{d_{27}, d_{28}\}$ to be precise — in which they are observationally indistinguishable. But when the whole set $\{P(a_i) \to P(a_{i+1}) | i < 1000\}$ is at stake, d_{27} , and d_{28} will be distinguishable, because in that case they must be compared with all other objects in $\mathcal{D}$.

Definition 11.

$\Delta \models \varphi$ iff for every coherent $\mathcal{M} = \langle \mathcal{D}, \mathcal{R}, \mathcal{I} \rangle$ and $C \subseteq \mathcal{D}$ the following holds: if $\mathcal{V}_{\mathcal{M}}(\Delta, C_{\Delta \cup \{\varphi\}}) = 1$, then $\mathcal{V}_{\mathcal{M}}(\varphi, C_{\Delta \cup \{\varphi\}}) = 1$.

Fact: $\Delta \models \varphi$ iff Δ/φ is classically valid.

Conclusion: The Sorites argument is valid. Moreover, each of its premises, taken separately, is true. Still, the conclusion does not follow since taken all together, the premises are false.