This Magic Moment: Horwich on the Boundaries of Vague Terms

Consider the following argument:

(1) Bertrand Russell was old at age 3×10^{18} nanoseconds (that's about 95 years)

(2) He wasn't old at age 0 nanoseconds

(3) So there is a number N such that he was old at N nanoseconds and not old at k nanoseconds for any $k<N$.

Presumably he was old for $k>N$ as well as for $k=N$. Given this, (3) says that there's a sharp boundary between when he was old and when he wasn't. (You could of course make it sharper than a nanosecond, by further division of the unit.)

Many people find the conclusion extremely counterintuitive; but premises (1) and (2) seem incontrovertible, and (3) follows from them using the least number principle.

Could it be that there's a problem with the unrestricted application of the least number principle to vague or otherwise indeterminate concepts? Paul Horwich considers this in "The Sharpness of Vague Terms",[1] but says

(I) that such a position is an act of desperation, and

(II) that when (3) is properly understood it shouldn't seem particularly counter-intuitive.

I will consider these claims in reverse order.

1. Are sharp boundaries counterintuitive?–I One worry about (3) is that our inability to know which number is the critical number N (i.e. which moment is the magic moment) doesn't seem at all like other cases of inevitable ignorance (e.g. the impossibility of knowing certain details of what's going on inside a specific black hole). Can we do justice to the evident difference?

Horwich thinks we can do justice to the difference. To this end he offers the following account of our inability to know the location of the boundary:

(a) Learning to employ 'old' involves something like acquiring a pattern of conditional degrees of inclination to accept 'S is old' or 'S is not old' on the basis of beliefs or assumptions about "underlying parameters", e.g. the time since S's birth.

(bi) This learned pattern of conditional degrees of inclination doesn't include any stable inclination to apply either 'S is old' or 'S is not old' when the assumed value of the "time since birth" parameter is in a mid-range. (We might have unstable inclinations.) (bii) The learned pattern may even include a prohibition against having a stable inclination to accept either 'S is old' or 'S is not old' in these cases. (This is of course compatible with our stably accepting the disjunction 'S is either old or not old', and Horwich thinks we should stably accept that.)

(c) Were we to substantially change the inclinations, say by starting to stably call people known to be in the mid-range 'old', this would constitute a change of meaning in the term.

(d) Knowledge requires stable belief, not based on error about the values of the underlying parameters; so one couldn't know of someone in the midrange that he is old, on the

[1] Chapter 4 of *Reflections on Meaning* (Oxford 2005).

current meaning of 'old'.

Let's grant that this explains why *we* can't know the the location of the borderline for our term 'old'. (I have doubts about the way that the notion of change of meaning is being deployed, but this isn't the place to pursue them.) But one of the things that seems to separate our ignorance of the borderline from ignorance of the details inside a black hole is that in the latter case there seems to be no conceptual incoherence in supposing an omniscient god who does know the details inside the black hole; whereas *not even an omniscient god could know "the location of the boundary between the old and the non-old"*. More generally, the following seems to be part of our conception of vagueness:

(*) Not only can't *we* know any claim that purports to give "the location of the boundary", no superior intelligence (god, Martian, whatever) could know such a claim either, even if that being thought in a system of representation very unlike ours.

At first blush it would seem that Horwich would have to deny the italicized claim, and reject (*). For if there is a unique critical number *N* (i.e. a unique *N* such that Russell was old at *N* nanoseconds and not old a nanosecond before), then there's a truth about what that number is, and so a fact about what that number is, and an omniscient being by definition knows all the facts. Some might try to dispute that with a highly "inflationary" notion of truth or fact, but certainly Horwich can't. So it initially surprised me to find Horwich asserting (*); asserting, indeed, that there is a *conceptual incoherence* in supposing that a god or a Martian knows the boundary of 'old'. How can he consistently say that? The answer, I think, is that Horwich is implicitly telling us that the concept of an omniscient being is conceptually incoherent: since it is conceptually incoherent to suppose that anyone knows the location of the boundary, and since an omniscient being would have to know the location, omniscience is conceptually incoherent. Postulating an omniscient being is like postulating a square circle. (A surprising contribution to the philosophy of religion!)

So far so good. But if there really is a critical number *N*, how are we to argue for (*)? (a)-(d) above explain (let's assume) why *we* can't know the value of *N*, but why couldn't someone else who thought or spoke only in a different language? Horwich's argument for this appeals to a version of deflationism, according to which

(D) the only way for someone (e.g. god or Martian) to judge that one of our terms α is true of an object k is to first argue that our term α means the same as some term β of his own language, and then argue that his own term β applies to k.

So if his term β really does mean the same as our term 'old', it is governed by the same rule for belief, and so he is prohibited from stably applying either β or its negation to the mid-range case. And so he can have no stable view about whether our term α applies to that case.

This is an ingenious argument, but I wonder whether the principle (D) on which it relies is believable. Suppose I hear mathematicians employing a word 'borniforous' that's completely unfamiliar to me, and hear them say that borniforous things are mathematical objects of some sort. It seems to follow from (D) that until I'm in a position to come up with a synonymous expression that I understand, I can't legitimately believe that their word doesn't apply to snails. That seems absurd.

I'm hesitant to rest too much on this critique: the question of how a deflationist should treat predicates in other languages that are untranslatable into ours is difficult, and perhaps the correct thing to say about it will allow for some response to the worry not far from the one Paul gives. For instance, maybe understanding a foreign term requires *either* translating it into a pre-existing term of one's language *or* incorporating it directly into one's language, and maybe Horwich's argument can be extended to take the second disjunct into account. I will not pursue this.

2. Are sharp boundaries counterintuitive?–II I turn to a quite different worry. Suppose we grant that Horwich has explained why no one, no matter how superior to us, can have *knowledge of* or *stable belief in* the location of the alleged sharp boundary between old and non-old. Still, that isn't enough to explain away reasonable doubts about there being such a boundary.

For instance, imagine someone, Roger, who thinks his life will go better if the number he chooses for his bank account password ends in the same digit as the last digit of the critical number N than it would if he had chosen a different last digit. (He realizes he'll never know if he made the right choice.) Compare Roger to Sam, who thinks his life will go better if the number he chooses for his bank account password ends in the 17^{th} significant digit of the Centigrade temperature at the currently hottest point in the interior of the sun. Sam's belief is thoroughly irrational, but Roger's intuitively seems even worse: it seems based on a kind of conceptual confusion about vagueness. It's hard to see how one could explain this on Horwich's account; certainly showing the impossibility of knowledge of the location of the boundary, or stable belief about it, doesn't address this.

There are plenty of other examples to the same effect. Imagine a person who knows his own age very precisely (and also knows his own level of physical decrepitude, and any other age-related facts of clear interest), but who *wonders whether* he's passed the alleged critical point that marks the unknown boundary between the old and non-old, and *hopes that* he hasn't. It isn't enough to say that his wondering about this is idle, and that his hope is irrational since it can't affect anything of interest to him; that's true of someone who wonder's or hopes about the 17^{th} significant digit of the Centigrade temperature at the currently hottest point in the interior of the sun, but as above, the critical number case seems different in being conceptually confused. (Indeed, it isn't clear that irrationality as opposed to conceptual confusion is involved in the case of the critical age. Many people idly wonder or hope about things of no independent interest: consider a dying man who wonders whether the Red Sox will win the pennant, and hopes that they will, though he knows he won't be around to savor it. For a believer in a sharp boundary, why should wondering and hoping about whether he's passed the boundary be any more irrational than that?)

A third example involves moral attitudes. To make the example somewhat realistic, it's better to switch from "the nanosecond where Russell became old" to "the nanosecond where a given life begins". Obviously the argument carries over to this case; that is, classical logic supplies an argument that there is a precise moment at which life begins. Given this, it is hard to see why someone shouldn't attach moral significance to that moment, and be deeply troubled by the thought that the magic moment might be earlier than he'd assumed (say at conception rather than at birth), making some action that he's already performed at a point in between deeply immoral. But it seems to me that this attitude, while sensible according to a supernaturalist metaphysics according to which at some unknown point our bodies are (instantaneously) infused by "vital substance", is not a sensible attitude for anyone who regards the question of when life begins as vague. And it seems hard to explain why this should be, if there is a true answer to the question "At which moment does life begin?".

Horwich sometimes says that the answers to questions like whether the critical number N is odd or even are "indeterminate", but this is no help: for him this just means that the answers aren't stably believable, so it in no way reduces the worries raised in this section. (This is in contrast to the role of indeterminacy in views which restrict excluded middle: there, asserting of a specific claim that it is indeterminate commits one to rejecting the corresponding instance of excluded middle.)

I'm sure there's nothing in what I've said that a defender of sharp boundaries would be unable to swallow. I do think, though, that there's something quite counterintuitive in the postulation of sharp boundaries; an explanation of why knowledge of or stable belief in their location is conceptually impossible does not make the counterintuitiveness go away.

3. Act of desperation? What is required to avoid the conclusion that there is a sharp line separating when Russell was old from when he wasn't? We certainly don't need to completely jettison the classical least number principle

(CLNP) $\exists nF(n)$ implies $\exists N[F(N) \wedge (\forall k<N)\neg F(k)]$.

We need only to weaken it slightly, in a way that allows it to have significant application even if excluded middled isn't assumed to hold generally:

(GLNP) $\exists n[F(n) \wedge (\forall k<n)(F(k) \vee \neg F(k))]$ implies $\exists N[F(N) \wedge (\forall k<N)\neg F(k)]$.

('G' is for 'generalized'.) When F obeys excluded middle, (GLNP) reduces to (CLNP). Presumably excluded middle holds throughout mathematics, and indeed whenever vagueness and related phenomena are not at issue. This suffices to explain why in dealing with precise language (or language that can be taken to be precise for the practical purposes at hand), (CLNP) can be assumed. But when vagueness is at issue, we can avoid the argument for sharp boundaries by restricting excluded middle where vague concepts are concerned, and recognizing that if we do so then the appropriate form of least number principle is (GLNP).

To put it in a slightly sloppy but picturesque manner, the idea is that for numbers *n* in a certain range, the claim that Russell was old at *n* nanoseconds is "fuzzy": it's inappropriate to assume that at that age *he was either old or not old*. (The range where this is fuzzy is *itself* fuzzy; this is crucial to the plausibility of the approach, and I will discuss it in a moment.) Given this, it's inappropriate to assume that there is a first *n* at which he was old. For to say of any given *n* that it is the first is to say that he was old at *n* nanoseconds of age and not old at $n-1$, and this will be fuzzy at best: it will be fuzzy if *n* or $n-1$ falls into the fuzzy region, false if they both fall outside it. (If it's fuzzy whether one of them falls into the fuzzy region, then it will be fuzzy whether the claim that *n* is the first one at which Russell was old is fuzzy: we will have higher order fuzziness. But in any case, the claim that *n* is the first won't be clearly true.)

Supervaluationists argue that even if all claims of form '*n* is the first natural number such that Russell was old at *n* nanoseconds of age' are at best fuzzy, still the existential generalization is clearly true. It seems to me that such a view does not avoid positing sharp boundaries, though I will not pause to argue that. If one wants to avoid positing sharp boundaries, one should take an existential generalization of claims that are fuzzy at best to be fuzzy. So the view is that it is inappropriate to assert 'There is a first natural number *n* such that Russell was old at *n* nanoseconds of age'. (That doesn't mean one ought to assert that there is no first nanosecond at which he was old. Negations of fuzzy claims are fuzzy too, hence inappropriate to assert.)

Views of this sort obviously raise a huge number of questions, and it is not my purpose here to discuss any of them in detail. Indeed, I won't discuss most of them at all, but I do want to mention one worry very briefly. The worry is that an approach like this can avoid a sharp line between the old and the not-old only by introducing other sharp lines, e.g. between one might call the *determinately* old and the *not determinately* old. The thought is that even if the law of excluded middle doesn't apply to the predicate 'old', it must apply to the predicate 'determinately old'; in which case there must be a first nanosecond at which Russell is determinately old. If so, little progress would have been made.

My answer to this is that on any reasonable way of introducing the notion of determinately old, excluded middle cannot be assumed to hold of it. Consider a few representative attempts:

(A) Russell is determinately old at those nanoseconds for which (i) he's old and (ii) it is true that he's either old or not old;

(B) Russell is determinately old at those nanoseconds for which (i) he's old and (ii) it is appropriate to assume that he's either old or not old;

(C) Russell is determinately old at those nanoseconds for which it is neither the case that he is not old nor that it's "fuzzy" whether he's old.

But on (A), there's no distinction between 'determinately old' and 'old'. As Horwich rightly insists, 'True(<p>)' is equivalent to 'p'; so A(ii) is equivalent to 'he's either old or not old', which is strictly weaker than A(i); so when conjoined with A(i) one just gets A(i), i.e. 'he's old'.

With (B), adding (ii) does produce a genuine strengthening. But given that 'appropriate' is obviously vague, there's still no reason to think that 'determinately old' has sharp boundaries.

The situation with (C) is similar to that of (B): it's unclear how exactly to explain "fuzzy", but it seems like however one explains it, it's bound to itself be vague.

Still, there's a substantial worry: that we could produce a sharp border by iterating a "non-fuzziness" operator into the transfinite. That is, why doesn't the sequence

> old;
>
> old and not fuzzy whether old;
>
> old, not fuzzy whether old, and not fuzzy whether fuzzy whether old;
>
> and so forth

collapse by level ω or by some higher transfinite level γ? If this were to happen—and it *does* happen in many standard proposals for non-classical logics for vagueness, e.g. the Lukasiewicz continuum-valued logic—then that would be a disaster. For then there would be a number N such that Russell was determinately$^{\gamma}$ old at nanosecond N but not determinately$^{\gamma}$ old a nanosecond before; we'd have a sharp boundary for 'determinately$^{\gamma}$ old', so why not just take this as the sharp boundary for 'old'?

If the determinately operator collapses to bivalence in this way, nothing would be gained by going non-classical. And it is a delicate matter to get a non-classical logic of vagueness in which such a collapse is avoided. Still, it can be done: there are reasonable logics of vagueness in which such a collapse never occurs.

Admittedly, a non-classical logic appropriate to vagueness is somewhat complicated. Given that almost every term is somewhat vague, wouldn't the non-classical approach make proper reasoning about ordinary subjects difficult? I think this worry is exaggerated. It might be useful to compare the case to geometric reasoning. We all know that space is not quite Euclidean, and indeed fails to be Euclidean in a quite complicated way; nonetheless, we are safe in using Euclidean reasoning except in special contexts, because the error involved in doing so is so slight. That is the policy I recommend for logic: reason classically, except for those situations where there is reason to think that the errors induced by such reasoning are significant. Situations where we derive boundaries for vague terms look like just the sort of situation to worry about!

4. Broadening the range of considerations. How do we decide between a classical logic approach to vagueness, which must postulate sharp borders, and a non-classical approach that avoids this but complicates the logic? It's a matter of weighing costs and benefits. I haven't tried to argue that the weight of the benefits is on the non-classical side: that would be a big task. Rather, I've just tried to argue that the non-classical approach is not without motivation (especially if it avoids the danger of collapsing determinately operators). I'd like to conclude by mentioning an additional item on the nonclassical side of the ledger: Berry's paradox.

Say that a 1-place formula *F(x)* of English is *uniquely true of* an object *c* if it is true of *c* and of nothing else. Let an *S-formula* be a formula of English with less than a thousand symbols. Then

> 1. There are only finitely many S-formulas; since there are infinitely many natural numbers, *there must be natural numbers that no S-formula is uniquely true of.*

So, by the least number principle,

> 2. There is a smallest natural number M such that no S-formula is uniquely true of it.

But 'x is the smallest natural number such that no formula of English with less than a thousand symbols is uniquely true of it' is an S-formula. So

> 3. 'x is the smallest natural number such that no formula of English with less than a thousand symbols is uniquely true of it' isn't uniquely true of M;

that is, it isn't uniquely true of the smallest natural number such that no formula of English with fewer than 1000 symbols is uniquely true of it. In other words,

> 3*. 'x is the smallest natural number such that no formula of English with fewer than 1000 symbols is uniquely true of it' either
>
> (i) *isn't* true of the smallest natural number such that no formula of English with fewer than 1000 symbols is uniquely true of it,

or (ii) *is* true of things *other than* the smallest natural number such that no formula of English with fewer than 1000 symbols is uniquely true of it.

Either option is thoroughly counterintuitive, and a gross violation of the basic "true-of" schema

(T) 'F(x)' is true of c if and only if F(c).

Horwich often emphasizes the centrality and importance of the truth-of schema, and for good reason. But we see that the unrestricted least number principle forces a violation of that schema. This is a substantial consideration in favor of restricting the least number principle in the context of vagueness, and hence in favor of restricting the law of excluded middle which underlies it.

Horwich takes the opposite stance, of restricting not excluded middle but the truth-of schema. But that has a high cost. Let's look at the point of the notions of truth and truth-of. Sticking to truth for simplicity: suppose I forget the details of what a Doomsayer said yesterday, but remember the gist well enough to conclude:

> If everything he said yesterday is true, then we're in trouble.

On the assumption that what he said was $p_1,\ldots,p_n$, this had better be equivalent to

> If p_1 and ... and p_n, then we're in trouble.

This requires the intersubstitutivity of True(<*p*>) with *p* in extensional contexts. Given the very minimal law $p \leftrightarrow p$, this yields the truth schema True(<*p*>)$\leftrightarrow p$. (The situation with 'true of' is similar.)

Restricting intersubstitutivity restricts the ability to generalize in a reasonable way, leading to extreme pathologies in theories of truth that reject intersubstitutivity or the truth schema. But we can keep the intersubstituvity principle and the truth schema unrestricted if we weaken excluded middle (and more or less equivalently, the least number principle); similarly for truth-of. We can do this in a way that allows for fully classical reasoning when no "ungrounded" uses of 'true' are present (and when vagueness isn't at issue): for instance, we can accept classical reasoning within mathematics without restriction.

And the logics that keep the truth schema (and the intersubstitutivity of True(<*p*> with *p*) seem to be fully suited to deal with vague and indeterminate concepts in the way sketched earlier. I think this is no accident: there's a strong intuitive connection between the Sorites example (Russell is old) and the Berry paradox example. Moreover, the obstacles that must be overcome in getting a logic that adequately handles vagueness and the semantic paradoxes are pretty much the same in both cases: for instance, in both cases we must make sure that no determinacy operator collapses to bivalence when iterated. (This is why the Lukasiewicz logic fails both as a logic for vagueness and as a logic for the semantic paradoxes.)

That there is a connection here is of no surprise. Vague concepts and 'true' seem species of indeterminate concepts. 'True' initially seems determinate, because it seems that the truth schema True(<*p*>)↔*p* settles its extension. But once we reflect on "ungrounded" sentences (such as Truthteller sentences, which assert their own truth; and Liar sentences, which assert their own *un*truth), we see that this is an illusion. This connection makes it natural to use the same logic for such "ungrounded" applications of 'true' as for vague predicates.

That there is a link between the semantic paradoxes and the paradoxes of vagueness is perhaps further suggested by another paradox of the same ilk, which seems to have ties to both. (I think I first heard of it many years ago in Martin Gardner's *Scientific American* column.) Some natural numbers aren't very interesting. So there must be a smallest one that isn't very interesting. The smallest one that isn't very interesting! What an interesting number! Contradiction. (In case anyone is tempted to regard this as a proof that every natural number is very interesting, it's worth remarking that an analogous proof using the classically correct least ordinal principle yields that every ordinal number is very interesting. Since for any cardinal number c, there are more than c ordinal numbers, this seems quite surprising!)

Another kind of paradox that suggests a connection is what Sorenson calls a "no-no" paradox: Person A asserts that what person B is saying is not true, at the same time that person B says that what person A says isn't true. Classically, either what A says is true and what B says isn't, or vice versa; and yet A and B seem symmetrically placed. (We might even imagine that A and B are Doppelgangers in a completely symmetric universe; in which case we have a failure of truth to supervene on non-semantic facts.) Intuitively this is a kind of underdetermination reminiscent of vagueness, and the paradox arises only from the supposition of excluded middle.

To summarize, I think there is considerable pressure in the vagueness case to slightly weaken the logic so as to avoid postulating counterintuitive boundaries, and even more pressure in the semantic paradox case to weaken the logic in the same way to enable us to keep the truth and truth-of schemas. These two pressures to weaken the logic are, I think, mutually reinforcing, and succumbing to this joint pressure is not the desperate measure that Horwich suggests it is.[2]

Hartry Field
New York University

[2] Thanks to Paul Horwich for helpful discussions of these issues.