FACING UP TO THE SORITES PARADOX

Terry Horgan

University of Memphis

The ancient sorites paradox is traditionally attributed to Eubulides, a contemporary of Aristotle and a member of the Megarian school, who is also credited with inventing the liar paradox. The sorites paradox figures centrally in most discussions of vagueness in philosophy and in logic. In my view, it has profound implications for metaphysics and semantics, as well as for logic. In this paper I will briefly explain why I think so, in a way that draws upon my other writings on vagueness.[1] The paper also will constitute a brief, opinionated, overview of the pertinent intellectual landscape, reflecting my own philosophical perspective on the issues.[2]

1.   Troubles for Standard Two-Valued Logic.

Let 'B(n)' abbreviate 'a man with n hairs on his head is bald'. Here is a familiar quantificational version of a sorites-paradoxical argument:

$(n)[B(n) \supset B(n+1)]$

$B(0)$

∴ $B(10^{17})$

Although the conclusion of this argument is surely false, it is hard to find a satisfying diagnosis of how the argument goes wrong if we confine ourselves to standard two-valued logic. For, since the inference is classically valid, and the second premise is clearly true, the only apparent option is to reject the first premise by asserting its negation:

(1) $\sim(n)[B(n) \supset B(n+1)]$.

But this statement is logically equivalent to

(2) $(\exists n)[B(n) \ \& \sim B(n+1)]$.

And this latter statement, as interpreted within standard two-valued logic, asserts the existence of a determinate one-hair transition point between baldness and non-baldness. Yet it seems essential to the vagueness of 'bald' that there be no such transition point. How, then, to diagnose such paradoxical sorites arguments?

One approach, which has the virtue of preserving standard two-valued logic, is epistemicism.[3] On this view, statement (2) is true, and this is because there really is a determinate transition-point between baldness and non-baldness; yet this transition point is unknowable by finite minds like ours. The essence of vagueness, says the epistemicist, is a particular kind of unknowability of determinate conceptual boundaries; vague concepts always have such boundaries, but we can never know where they are. But most philosophers and logicians who have reflected on vagueness, myself included, find epistemicism just impossible to believe.

So for those who reject epistemicism, it appears hard to deny that an adequate diagnosis of the sorites paradox will require some kind of repudiation of standard two-valued logic. What seems needed is a non-standard logic of vagueness which does one or both of these things: (i) provides some way of rejecting the first premise in the above sorites argument without commitment to a sharp transition from baldness to non-baldness; or (ii) renders such sorites arguments invalid. There are various proposed approaches to the logic of vagueness, and there is no consensus about which is best. Among philosophers, and among logicians in philosophy departments, probably the most popular approach is supervaluationism, which I will say more about below.[4] (Among computer scientists and mathematicians who work in logic, probably the most popular view is so-called "fuzzy logic."[5] More on this below too.)

2.   The Metaphysics of Vagueness: Two Approaches.

There are two broad metaphysical approaches to vagueness. One approach affirms so-called <u>ontological</u> vagueness, i.e., vagueness in the mind-independent, discourse-independent, world. More specifically, it affirms the existence of genuine <u>objects</u> and <u>properties</u> that are

vague. The other approach denies ontological vagueness, and asserts that vagueness is entirely a matter of language (and of thought-content). On the non-ontological view, vagueness is a purely semantic affair, akin to phenomena like ambiguity and semantic indeterminacy.[6] On this metaphysical issue, too, there is no consensus. But probably the more popular view is the non-ontological conception of vagueness.

If one does adopt the non-ontological view, in combination with some alternative to standard two-valued logic, then what happens to the notion of <u>truth</u>? In effect, truth (for vague discourse) gets treated not as direct language-world correspondence, but rather as <u>indirect</u> correspondence between vague language (and vague thought-content) on one hand, and non-vague reality on the other. Supervaluationism, when combined with the denial of ontological vagueness, provides an example. Instead of construing truth directly, by appeal to a single intended interpretation that assigns vague objects as the referents of singular terms and assigns vague properties (or vague extensions) to predicates, the supervaluationist construes truth simpliciter in an indirect way: viz., as truth in all permissible interpretations--there being more than one permissible interpretation, each of which "precisifies" the statement's vague vocabulary in a specific way. The language/world correspondence relation is now indirect because truth is not grounded in straightforward referential links between sub-sentential vocabulary on one hand, and genuine vague objects and vague properties on the other hand. (Falsity, likewise, is indirect non-correspondence, for the same reason.) In the case of the statement that Mt. Whitney is tall, for instance, the indirect correspondence that constitutes truth is essentially this: under all ways of precisifying the expressions 'Mt. Whitney' and 'tall' that are consistent (apart from their precision) with the semantic standards governing these terms, the object assigned to 'Mt. Whitney' instantiates the property assigned to 'tall'.[7]

Although supervaluationism provides one possible implementation of the notion of indirect correspondence, the core distinction between direct and indirect correspondence can be formulated more generically. I myself, in a number of writings, have proposed construing truth

as <u>semantically correct assertibility</u>, under contextually operative semantic standards (and falsity as semantically correct deniability, under such standards.)[8] The semantic standards conspire with goings-on in the mind-independent, discourse-independent, world to yield truth. In the limit case, the applicable standards are <u>referentially strict</u>: they require referential linkages connecting a statement's basic subsentential constituents to real objects and real properties. In this limit case, truth (i.e., semantically correct assertibility) is direct correspondence. But the contextually operative semantic standards can also work in such a way that the requisite goings-on in the world need not involve entities or properties answering directly to the statement's basic subsentential constituents. In such cases, truth is indirect correspondence.[9]

In short, approaches to vagueness that (i) repudiate ontological vagueness, (ii) embrace some alternative to standard two-valued logic, and (iii) treat various non-negative statements employing vague language as true, are effectively construing truth as indirect correspondence. Moreover, indirect correspondence is naturally understood as semantically correct assertibility, under assertibility standards that are not referentially strict. From this perspective, vague statements are frequently true, under contextually operative semantic standards, but they do not carry genuine ontological commitment to any vague objects or vague properties (under the operative standards)—a good thing, given the denial of ontological vagueness.

Those who posit ontological vagueness, on the other hand, are under no immediate pressure to construe truth as indirect correspondence. Instead, they can regard truth as direct correspondence between vague language (and vague thought-content) on one hand, and vague reality on the other. Typically, believers in ontological vagueness will also espouse some approach to the logic of vagueness that departs from standard two-valued logic; for, there remains an apparent need for such a departure in order to disarm the sorites paradox without commitment to sharp transitions between categories like bald and non-bald. So, for believers in ontological vagueness, reality itself—the instantiation of genuine, vague, properties by genuine, vague, objects—conforms to some kind of non-standard logic.

One way to implement this approach is to posit ontological vagueness and also embrace some version of so-called "fuzzy logic." Here the key idea is that genuine vague properties can be instantiated not merely in an all-or-nothing way, but also <u>partially</u>, in specific numerical degrees from zero to one. Likewise, genuine vague objects can possess spatial or temporal parts not merely in an all-or-nothing way, but in specific numerical degrees. Language-world correspondence is still direct, on this account, but comes in degrees.[10]

3.   The Impossibility of Ontological Vagueness.

So far I have described two broad, non-epistemicist, approaches to vagueness. One approach repudiates ontological vagueness, rejects standard two-valued logic, and treats truth as indirect correspondence. The other approach embraces ontological vagueness, rejects standard two-valued logic, and treats truth as direct language-world correspondence. In my own view, careful attention to the nature of vagueness settles this debate in favor of the first alternative. Matters of logic, closely connected to the sorites paradox, now enter the discussion in a new and different way, and establish the impossibility of ontological vagueness. Let me explain.

It is widely believed that an essential attribute of vagueness is what Mark Sainsbury 1990 calls <u>boundarylessness</u>, a feature that can be characterized by reference to sorites sequences associated with vague terms. Consider a vague term—say, 'heap'—and consider a sorites sequence involving the given term in which the initial statement is true and the final statement is false—say, a series of statements successively predicating the vague term 'heap' first to a pile of sand with 1 billion grains, then to an object produced by removing just one grain, then to an object produced by removing yet another single grain, and so forth down to a statement predicating 'heap' to a single grain of sand. To say that vagueness is boundarylessness is to say that in such a sequence, (i) initially there are true statements (with each predecessor of any true statement being true); (ii) later there are false statements (with each successor of a false statement being false); and (iii) there is no determinate fact of the matter about the transition from true

statements to false ones. Condition (iii) I call <u>robustness</u>. It requires not only that there be no determinate <u>abrupt</u> transition from true statements to false ones, but also that the truth/falsity transition should involve no determinate semantic transitions at all; thus it also precludes, for instance, an overall true-to-false transition involving first a determinate abrupt transition from truth to the semantic status "indeterminate whether true or false," and later another determinate abrupt transition from this in-between status to falsehood.[11]

If one considers what it would take to fully accommodate boundarylessness—that is, accommodate it in a way that thoroughly eschews arbitrary semantic transitions of any kind— one finds that, for the successive statements in a sorites sequence, there are semantic requirements in play that cannot be simultaneously satisfied. Boundarylessness has two conceptual poles. On one hand there is an individualistic pole, applicable to individual pairs of adjacent statements in a sorites sequence: viz., for any pair of adjacent statements, the two statements must have the same semantic status (truth, falsity, indeterminateness, or whatever). Otherwise there would a determinate semantic transition between them, contrary to the claim that there is no determinate fact of the matter about semantic transitions in the sequence. On the other hand, there is also a collectivistic pole in the notion of boundarylessness, applicable globally with respect to a sorites sequence as a whole. Two collectivistic requirements apply. First, it is impermissible to iterate indefinitely the individualistic-pole requirement for successive adjacent pairs of statements, in the manner of paradoxical sorites arguments. Second, there is simply no determinate collective assignment of semantic status to all the statements in a sorites sequence. These individualistic and collectivistic requirements cannot be jointly satisfied; for, the only way that a sorites sequence could fully conform to the individualistic pole would be for every statement in the sequence to have the same semantic status.   (This is the lesson of the sorites paradox, which emanates directly from the individualistic pole of boundarylessness.) So boundarylessness is logically incoherent, in a specific way: it imposes mutually unsatisfiable semantic standards upon vague discourse.

The logical incoherence of vagueness is <u>generic</u>, in the following sense: it is not directly linked to, and does not presuppose, any particular approach to the logic of vagueness. Debates about the specific logical principles governing vague discourse are largely independent of the generic logical incoherence of vagueness—a theme to which I will return shortly. Moreover, the specific kind of generic incoherence exhibited by vagueness needs to be distinguished from a stronger, and highly malevolent, kind of generic logical incoherence. Vagueness does involve <u>weak</u> generic logical incoherence—viz., the presence of mutually unsatisfiable semantic standards governing vague discourse (and vague thought-content). But this does not necessarily bring in its wake <u>strong</u> generic logical incoherence—viz., commitment to individual statements that are logically contradictory, such as statements of the form $\Phi \ \& \sim\Phi$. On the contrary, I maintain that vagueness can, and does, possess weak generic logical incoherence without possessing the strong kind—another theme to which I will return shortly.

So weak logical incoherence is a feature of the contextually operative semantic standards governing vague discourse, in ordinary contexts of usage. Insofar as these semantic standards are not referentially strict, and thus only require indirect language-world correspondence, there is no particular problem about this (provided that the semantic standards are not also strongly logically incoherent). However, the <u>world</u> cannot be logically incoherent, even in the weak way: it cannot have features that are the ontological analogues of mutually unsatisfiable semantic standards.[12] Hence there cannot be ontological vagueness. Therefore, barring epistemicism, the only viable general approach to vagueness is the one that conceives it non-ontologically, and construes truth as indirect correspondence between vague language and non-vague reality. The sorites paradox is at the heart of the matter in the reasoning that has led up to this conclusion, since the paradox directly reflects the weak generic logical incoherence of vagueness.

4.   Logical Discipline and Weak Logical Incoherence.

How is it that the semantic standards normally governing vague discourse can incorporate

weak generic logical incoherence without the strong kind? Briefly, my story goes as follows. Incompatible individualistic and collectivistic semantic requirements are indeed in force insofar as vague discourse exhibits boundarylessness. That is, no requirement is defeated by any others, in the sense of having defeasibility conditions that are satisfied by the presence of the competing and incompatible requirements. But these competing requirements are not on a par with one another either. The collectivistic-pole requirements <u>dominate</u> the individualistic-pole requirements without defeating them; that is, to the extent that the requirements conflict, truth is determined by the collectivistic-pole requirements. In practice, this means that paradoxical sorites arguments are to be eschewed; it also means that one must not acknowledge the existence of any determinate semantic transitions (even unknown or unknowable ones) in a sorites sequence. (Semantic status still must conform partially to individualistic-pole requirements, however. For instance, it is never the case, for any specific pair of adjacent statements in a sorites sequence, that the two statements differ in semantic status.[13]) So the semantic standards governing vague discourse are <u>logically disciplined</u>, in virtue of the dominance (without defeat) of collectivistic-pole requirements. Because of this logical discipline, no logically incoherent statement is true, under those standards; strong generic logical incoherence is avoided.

5.   Transvaluationism and Its Potential Implementations.

<u>Transvaluationism</u> is my name for the general approach to vagueness I have been describing. Transvaluationism makes two fundamental claims. First, vagueness is weakly logically incoherent without being strongly logically incoherent. Second, vagueness is viable, legitimate, and indeed essential in human language and thought; its weak logical incoherence is benign rather than malevolent. Just as Nietzsche held that one can overcome nihilism by embracing what he called the transvaluation of all values, transvaluationism asserts that vagueness, although logically incoherent in a certain way, can and should be affirmed and embraced, not nihilisticaly repudiated.[14]

If vagueness is really boundarylessness, as it certainly appears to be, then, since boundarylessness involves disciplined weak logical incoherence, an adequate treatment of vagueness will have to be some version of transvaluationism. Moreover, since the weak logical incoherence involved is generic, transvaluationism itself is a fairly generic approach, potentially open to further development and articulation in a variety of different ways. Numerous details about the logic and semantics of vagueness remain open within the generic conception, and might get handled differently in different versions.[15] But regardless of how the details go, any account of vagueness that seriously comes to grips with boundarylessness must be a version of transvaluationism—whether its proponents acknowledge this fact or not. Much of the recent literature on so-called "higher-order vagueness"—i.e., vagueness of categories like truth, falsity, and the category "neither true nor false"—is best viewed as falling within the transvaluationist framework. In effect, specific proposals concerning higher-order vagueness amount to suggested strategies for implementing the dominance-without-defeat of collectivistic semantic standards over individualistic ones.[16]

The weak generic logical incoherence that any such proposal must take on board, at least implicitly, will inevitably reveal itself when one considers what the advocate of the particular proposal will be forced to say when confronted with what I call a "forced march" through a sorites sequence. Consider, for instance, a sorites sequence for baldness: $B(0), B(1), ..., B(10^{17})$. A forced march through this sequence is a series of questions, with respect to each successive statement, "Is it true?" Each of the questions is perfectly meaningful. And for no two successive questions could it be correct to give different answers; for, that difference would mark a determinate semantic transition, contrary to the nature of vagueness. So the only thing to do, when confronted with the prospect of forced-march querying, is to refuse steadfastly to play that question-and-answer game. Instead of taking the forced march, adopt a Zen attitude: be tranquilly silent in the face of those persistent queries, in the knowledge that no complete set of answers is semantically correct.[17] This is the right thing to do, because it reflects the dominance

of collectivistic-pole semantic requirements over individualistic-pole requirements. But although this refusal to take the forced march is entirely appropriate as a tactic for avoiding commitment to any logically contradictory statements, it would be self-deception to think that such an avoidance tactic somehow eliminates the weak logical incoherence of vagueness. The individualistic-pole requirements are still in force, even though they are dominated by the logically incompatible collectivistic-pole requirements; for, the respective queries in the forced march are all still meaningful and each still demands the same answer as its predecessor, even though it is proper and respectable to duck those individualistic semantic requirements by refusing to take the forced march. The forced march is essentially just the sorites paradox itself, with our noses rubbed in it.[18] What it reveals is the weak generic logical incoherence of vagueness.

6. Conclusion.

To summarize: Although vagueness is a legitimate and essential feature of human language and thought, it is weakly generically logically incoherent; yet it is not strongly generically logically incoherent, because it exhibits logical discipline involving the dominance, without defeat, of collectivistic semantic standards over individualistic ones. This generic approach to vagueness, which I call transvaluationism, has several important consequences. First, there cannot be ontological vagueness. Second, truth, for vague discourse, is indirect correspondence between vague language and non-vague reality. Third, any adequate account of the logic of vagueness will incorporate the weak generic logical incoherence of boundarylessness; it will be an implementation of transvaluationism, rather than an alternative to it.

I remarked earlier that indirect correpondence is best understood as semantically correct assertibility, under contextually operative semantic standards that are not referentially strict. I also said that we sometimes employ language under limit-case, referentially strict, direct-correspondence, semantic standards. These are the standards appropriate for serious ontological inquiry. When they are in play, so is classical two-valued logic. Under this limit-case use of

language, sorites reasoning can be correctly employed to construct <u>reductio ad absurdum</u> arguments against the existence of vague properties or vague objects—including, of course, not only mountains and clouds, but also tables, chairs, and (regrettably) persons.[19] Fortunately, however, ordinary uses of vague language are not governed by limit-case semantic standards, and normally the contextually operative standards conspire with non-vague reality to render much of our vague discourse true.[20]

---

[1] See Horgan 1994, 1995, 1998b, in press. Other papers of mine discussing vagueness are cited therein.

[2] For more extended overviews, including bibliographical information, see Tye 1995 and the editors' introductory essay in Keefe and Smith 1996. For a thorough critical overview, including rich bibliographical information, see Williamson 1994.

[3] For articulation and defense of epistemicism, see Williamson 1994 chapters 7 and 8, and the additional sources cited there.

[4] The <u>locus classicus</u> of the supervaluationist theory of vagueness is Fine 1975.

[5] The <u>locus classicus</u> of the fuzzy-logic treatment of vagueness is Zadeh 1975.

[6] A mixed view is possible too, asserting that some vagueness is ontological and some is purely semantic; but I will ignore this hybrid position, for simplicity.

[7] Supervaluationism preserves standard logic's theorems and inference rules, but not its bivalent semantics. A statement applying a vague predicate to a borderline instance will be neither true nor false, since it will come out true under some permissible interpretations and false under others. With respect to the sorites paradox, the supervaluationist affirms both statement (1) and statement (2) above. But under supervaluationist semantics, statement (2) can no longer be understood as asserting a determinate transition-point between baldness and non-baldness. Rather, (2) is true because it comes out true under all permissible interpretations of 'bald'.

[8] I call this general approach to truth "contextual semantics." It is explored and developed in Horgan 1995, 1998a, and in various other papers of mine cited therein (some co-authored with Mark Timmons). The overall defense of contextual semantics in these papers proceeds largely by appeal to metaphysical and epistemological considerations orthogonal to matters of vagueness. (For awhile I called the view

"language-game semantics, then "psychologistic semantics.") The relevant notion of semantic correctness has nothing to do with matters of etiquette. A statement can be semantically correct, in the relevant sense, even if it would be impolite, impolitic, or otherwise inappropriate to utter it. Semantic correctness is also distinct from epistemic warrant; a statement can be epistemically warranted but semantically incorrect, and can be semantically correct but epistemically unwarranted.

[9] Contextual semantics allows for contextual variation in semantic standards not only across different domains of discourse, but even within a given domain. In philosophical contexts of inquiry where serious ontological issues are at stake, for instance, typically the standards governing a local mode of discourse become referentially strict--or anyway, closer to referentially strict than they normally are. One way to signal such a shift of score in the language game is by means of emphasis terms, thus: "Does Mt. Whitney exist? Of course! Does it really exist? No!".

[10] Fuzzy logic typically treats all real numbers from zero to one as truth values. Consider, for instance, the predicate 'B(n)' introduced earlier, and the sequence of statements B(0), B(1), ..., B(10^{17}). Those who address vagueness via fuzzy logic will say that the numerical truth values of these successive statements gradually diminish from one (outright truth) to zero (outright falsity) as we progress through the sequence, and that this is the key to blocking the above paradoxical sorites argument without positing any transition-point between bald and non-bald. For an approach that embraces ontological vagueness but weds it to a logic significantly different from fuzzy logic, see Tye 1990, 1994.

[11] For convenience of exposition, here I discuss boundarylessness metalinguistically, in terms of statemements and their semantic status. But the same core idea applies equally well at the first-order level of description. Consider a sorites sequence consisting of the respective sand conglomerations themselves. To say that heaphood is boundarylessness is to say that (i) initially in this sequence there are heaps (with each predecessor of a heap being a heap); (ii) later there are non-heaps (with each successor of a non-heap being a non-heap); and (iii) there is no determinate fact of the matter about the transition from heaps to non-heaps.

[12] For example, there cannot be a genuine property H (for 'heaphood'), and a sequence of sand conglomerations each of which has one fewer grain than its predecessor, such that (i) initially in the sequence there are instances of H (with each predecessor of an H instance being an H instance), (ii)

eventually there are non-H instances (with each successor of a non-H instance being a non-H instance), and (iii) for each pair of successive piles in the sequence, either both are H instances, or both are non-H instances, or both are neither.

[13] Does this mean that under the correct collective assignment of semantic status to all the statements in a sorites sequence, no two adjacent statements differ in semantic status? No. According to the collectivistic-pole requirements, there is no correct collective assignment of semantic status to all the statements in the sequence.

[14] One reason I call my position transvaluationism is to emphasize that it is not a species of what Williamson 1994 calls nihilism—the view that "vague expressions are empty; any vaguely drawn distinction is subverted" (p. 165). Another reason is to emphasize the need for a "transvaluation of all truth values," so to speak—i.e., the need to transcend the impossible goal of finding some logically coherent, semantically correct, collective assignment of semantic status to all the statements in a sorites sequence. The proper goal for a semantics of vagueness, rather, is to provide an adequate account of the normative standards governing semantically correct assertoric practice.

[15] Perhaps transvaluationism can even be implemented by standard two-valued logic, employed in a way that respects in practice the logically disciplined weak generic incoherence of vagueness. Concerning our accommodation of vagueness, Quine 1995 remarks, "What I call my desk could be equated indifferently with countless almost coextensive aggregates of molecules, but I refer to it as a unique one of them, and I do not and cannot care which. Our standard logic takes this...in stride, imposing a tacit fiction of unique though unspecifiable reference" (p. 57).

[16] One salient example, discussed in Horgan 1998b section 4, is what I call "iterated supervaluationism." The core idea of this approach is that the metalanguage for stating supervaluationist semantics is itself vague, and thus it too is subject to a supervaluationist treatment in a meta-meta-language; and so on, all the way up the metalinguistic hierarchy. More specifically, the metalinguistic expression 'permissible interpretation' is itself vague, and thus has various permissible interpretations itself in a meta-meta-language; and so on up. I myself am inclined to prefer a different approach to the logic of vagueness, described in Horgan 1994, sections 2 and 3.

[17] The phrase 'Zen attitude', suggested by Matjaz Potrc, seems to me to capture well the spirit of my

recommended approach—more so than my use, in earlier papers, of sado-masochistic imagery (the dominatrix, cracking her whip) in connection with the theme of dominance without defeat.

[18] As my colleague John Tienson once remarked to me.

[19] Peter Unger was right! See Unger 1979a, 1979b.

[20] I thank Robert Barnard, Matjaz Potrc, John Tienson, and Mark Timmons for helpful comments and discussion.