Epistemically Vague Boundaries without Margins of Error

Robert C. Koons

Professor of Philosophy, University of Texas

I propose to give a solution to the sorites paradox with respect to monadic predicates expressing sensible qualities that is squarely within the epistemicist camp of Williamson (1994, 1996), Sorenson (1988, 2001), and Rescher (2009). In particular, I will address what appears to be the Achilles' heel of that account: namely, the problem of giving a plausible story about how an exact boundary around the extension of such predicates can be determined by the cooperation of the world and our linguistic practices. This solution will rely on recent work on the metaphysics of properties and thus will provide a *metaphysical*, rather than epistemological, defense of epistemicism. I will focus on the sorites paradox as the entrance point into the problem.

1. What is the *Sorites* problem?

A paradox is a set of propositions that are jointly inconsistent but individually so plausible as to be nearly undeniable. Paradoxes challenge us because they seem to force us to choose which premise to reject, when we are reluctant to reject any of them.

The so-called *sorites* paradox takes the following form:

(S1) There are a finite number of cases x(1) through x(n) such that:

(C1) x(1) is not F,

(C2) x(n) is F, and

(C3) for all k such that $0 < k < n$, if x(k) is not F, then x(k+1) is not F.

This is obviously an inconsistent proposition, given some very elementary facts about arithmetic, but in order to constitute a paradox, we would need to find a predicate F and a finite set of cases x(1) through x(n) of such a kind that each of (C1), (C2), and (C3) are virtually undeniable. What reason do we have for believing in such things?

In fact, the claim that there is a *sorites* paradox seems to depend upon a confusion of metaphysics and epistemology. The following proposition is very plausible:

(S2) There are a finite number of cases x(1) through x(n) such that:

(C1) x(1) is known not to be F, and so not known to be F,

(C2) x(n) is known to be F, and

(C3) for all k such that 0 < k < n, if x(k) is not known to be F, then x(k+1) is not known to be F.

However, (S2) is not a paradox, because it is not inconsistent. What (S2) implies is that there is at least one finite series of such a kind that x(m) in that series is F even though it is not known that x(m) is F. That is, (S2) simply implies that the relevant (unnamed) knower is not omniscient with respect to being F. That doesn't seem surprising, no matter what property 'F' is supposed to express. Some have supposed that phenomenal properties of sense-data are the sort of properties about which the subject of experience must be omniscient in this way, but this is surely too strong a thesis to saddle such a sense-data theory with.

Believers in phenomenal properties have two ways, both quite plausible, for avoiding the (S2) paradox. First, they should be satisfied with a relatively weak thesis about seemings: namely, the claim that if some sense-datum seems to be F (where 'F' expresses a phenomenal property), then it is F. This doesn't entail that if it is F, it must seem to be F, nor that if it isn't F, it must seem not to be F. Second, phenomenal-property theorists can legitimately posit a gap between seemings and knowledge. They needn't accept that whenever a sense-datum appears to be F it is known by the subject to be F. At most, they should suppose that, if some datum appears to be F, then there is some phenomenal property G, such that being F entails being G and such that the subject knows the datum to be G. If I experience a datum that is some specific shade of phenomenal color (like crimson), I need not know that it is crimson, since I might lack the relevant concept, but I should surely know something about its phenomenal color.

We might try replacing actual knowledge in S2 with some sort of ideal knowability:

(S3) There are a finite number of cases x(1) through x(n) such that:

(C1) it is not knowable that x(1) is F,

(C2) it is knowable that x(n) is F, and

(C3) for all k such that $0 < k < n$, if it is not knowable that x(k) is F, then it is not knowable that x(k+1) is F.

Before reconsidering the case of phenomenal properties, let's consider whether (S3) poses any paradox with respect to natural physical properties. Consider, for example, physical length or temporal duration. Let 'F' be 'is less than m picometers long' or 'is less than m picoseconds in duration'. There are cases that we know not to be F (the width of the Milky Way galaxy or the duration of a Mahler symphony) and other cases we know not to be F (microscopic cases like the width of a proton or the duration of a gamma ray oscillation). Nonetheless, there will be a finite series falsifying (C3) of (S3), because there are differences in length or duration that are in principle undetectable by us (certainly given our current capacities and technology).

Why is the falsity of (C3) in such cases so unproblematic? Presumably, it has something to do with the fact that we think that some of our physical concepts, like the picometer or picosecond, manage to "cut nature at the joints" (to use Plato's phrase from the *Timaeus*) in such a way that far outruns our capacities for measurement and discrimination. This is unsurprising so long as there are certain properties that are naturally isolated from other properties of the same type. For example, Planck's length is a very determinate quantity, naturally isolated from other lengths by virtue of its playing a highly privileged role in the laws of nature. The speed of light is a similarly privileged quantity. No other quantity of speed in its neighborhood plays a similar role in the laws of nature. To avoid the *sorites* paradox, all we have to do is to assume that the concept F manages to express one of these isolated properties. This seems a reasonable hypothesis in many cases: we could imagine that our concepts or our predicates manage to express such a property by being causally connected to particular instances of it. Or we might suppose that isolated properties act as a kind of *reference magnet* (or *intension magnet*), in the sense that the best interpretation of our conceptual and linguistic activity would be one that assigns such isolated properties (or properties that can be relatively simply defined in terms of isolated properties) to our concepts.

Consequently, if S3 is to provide a paradox, there must be some concepts that express properties that are not isolated in this way--that belong to a domain or determinable property forming a continuum in which there are no natural joints.

2. Theories of Properties

On some theories of properties, this is impossible. For example, according to David Armstrong's "sparse" theory of universals, all properties are universals, and all universals are isolated. However, such a severe solution to the problem introduces some serious epistemological difficulties. What, in particular, are we to do with our concepts of the sensible qualities, both primary (size, shape, orientation) and secondary (color, warmth, texture, pitch, volume)? One option would be to adopt a global error theory about such concepts: they simply do not express properties, and all propositions expressed by means of them are false. However, such an error theory amounts to a kind of radical skepticism about our observational knowledge, which would make the problem of grounding any of our scientific knowledge of the world's real properties quite daunting.

So, let's suppose that at least some of our concepts express non-isolated properties, properties that occupy points in a space of properties that is at least dense, if not continuous, and within which there are no natural boundaries. Such properties are neither privileged nor definable in terms of privileged properties. How, then, can the intensions of such concepts be determined?

A promising strategy is to borrow an idea from resemblance or natural-class nominalism. This strategy is still available to us, even if we choose to be realists about universals, just so long as we do not assume a very sparse theory of universals, that is, so long as we don't assume that all universals are isolated. Suppose, for example, that there is a universal corresponding to each determinate shade of color. If there are non-denumerably many such universals, forming a three-dimensional space of colors without internal boundaries, then the problem of associating an intension with a color concept cannot be solved by simply finding the one isolated universal instantiated in the normal or paradigm cases to which the concept has been applied, since *ex hypothesi* there is no such single universal, but rather a large number of relevant universals scattered across an undivided continuum of shades.

Contemporary nominalists, like David K. Lewis (Lewis 1983, pp. 347-8) or Gonzalo Rodriguez-Pereyra (2002), have proposed that concepts be associated with natural classes, which can be defined in terms of a metaphysically privileged resemblance relation. (We'll set aside any worries about whether this relation itself requires a universal.) To use Lewis's proposal, a class C is natural if and only if there is some class D such that, for every x, x and the members of C resemble each other and do not likewise resemble any member of D if and only if x is a member of C. I believe that it is more useful to think, as Rodriguez-Pereyra does, in terms of a set of degrees of resemblance. We can then say that a class C is natural if and only if it contains all of the cases that resemble each of some subset of C (the paradigms) to at least some definite degree. Equivalently, we can say that class C is natural if and only if there is some degree d and some class E, such that C includes all and only those things that resemble *all* members of E to degree d.

In order to further simplify things, I will assume that there is a single three-term relation: *x resembles y to degree z*. I am also going to assume that there is a single class of intensities that can play the role of being the third term in this relation, and I will assume that the intensities can be linearly ordered by a single *greater-than* relation, and that the set of intensities is continuous, in the sense that for every series of intensities bounded above there is a least upper bound, and for every series of intensities bounded below there is a greatest lower bound. I will also assume that both the three-term resemblance relation and the order relation on intensities are isolated properties--i.e., that they are each metaphysically privileged in a way that no relation with a similar intension is. If we didn't assume this, then we would face an infinite regress in accounting for the intension of any concept.

Given any determinable property with its own range of determinate properties, we can, by selecting two or more paradigmatic determinate properties, use the resemblance relation to impose a set of sharp boundaries upon the range. Again, for simplicity's sake, let's assume that we are going to define a pair of contrary properties by means of a pair of paradigms. For example, we could define the concepts of light-colored and dark-colored by means of the paradigms of pure white and pure black. A case is light-colored if and only if it is more similar to white than it is to black, and it is dark-colored if and only if it is more similar to black than it is to white. Such a definition will divide all cases into two or three kinds: the light-colored, the dark-colored, and (possibly) the borderline cases. A borderline case would resemble black to exactly the same degree to which it resembles white. The borderline cases are exceptional, in the sense of forming a set of measure zero, given the continuity of the class of degrees. Nearly all cases will fall on one side or the other of the boundary.

Let's call a concept whose intension is grounded in this way a 'resemblance-based' concept. In order to grasp such a concept fully, I must know two things: the relevant paradigm cases, and the intension of comparative resemblance relation. Knowledge of truths involving such a concept requires at least a partial grasp of the concept. It is reasonable to suppose that such partial grasping of a concept is possible only if the full grasping is also possible. So, let's turn our attention to the case of an ideal wielder of the relevant resemblance concepts: one who both knows the relevant paradigm cases and whose sensory and cognitive apparatuses are fully attuned or calibrated to the privileged resemblance relation.

In such an ideal case, appearance of resemblance and reality of resemblance would perfectly coincide:

(CAR-Pos) x appears to resemble y more than z iff x resembles y more than z.

It also seems reasonable, at least prima facie, to assume that the ideal subject's capacity for knowledge coincides with appearances:

(CKA-Pos) It is knowable that x resembles y more than z iff x appears to resemble y more than z.

I've labeled these theses as CAR-Pos and CKA-Pos to distinguish them from the corresponding theses involving negative facts:

(CAR-Neg) x appears not to resemble y more than z iff x does not resemble y more than z.

(CKA-Neg) It is knowable that x does not resemble y more than z iff x appears not to resemble y more than z.

The first of these negative theses is quite plausible, since we can assume a kind of law of excluded for appearances: if x, y, and z are all apparent, and x does *not appear* to resemble y more than z, then x appears *not to resemble* y more than z. Such a law of excluded middle does not seem at all plausible for knowability (for reasons I will discuss in section 5 below).

Putting (CAR-Pos) and (CKA-Pos) together, we arrive at the coincidence of knowability and reality:

(CKR-Pos) it is knowable that x resembles y more than z iff x resembles y more than z.

The CKR-Pos thesis comes closer to making S3 into a genuine paradox, since both theses are plausible and yet they appear on first glance to be jointly inconsistent, on the assumption that F is a resemblance-based concept. Here is an expanded version of S3, incorporating my analysis of resemblance-based concepts.

(S3′) There are a finite number of cases x(1) through x(n) such that:

(C1) it is not knowable that x(1) is F, because x(1) resembles some contrary paradigm more than it resembles the F-paradigm,

(C2) it is knowable that x(n) is F, i.e., it is knowable that x(n) resembles the F-paradigm more than it resembles any contrary paradigm, and

(C3) for all k such that 0 < k < n, if it is not knowable that x(k) is F, then it is not knowable that x(k+1) is F, i.e., if it is not knowable that x(k) resembles the F-paradigm more than any contrary paradigm, then it is also not knowable that x(k+1) resembles the F-paradigm more than any contrary paradigm.

CKR-Pos requires that, if any x(k+1) in the series does resemble the F-paradigm more than any contrary paradigm, then it is knowable that x(k+1) resembles the F-paradigm more than any contrary paradigm, no matter what the case is with respect to x(k). This would appear to bring CKR-Pos in conflict with clause C3 of S3′.

In fact, however, CKR-Pos and S3′ are not logically inconsistent, so long as in the case of each finite series satisfying the conditions in S33, there is some case x(m) such that x(m) is a *borderline case*, a case that resembles the F-paradigm and some contrary paradigm to *exactly the same degree*. In such cases, x(m) does not resemble the one paradigm more than the other, and so CKR-Pos is silent on what is knowable about such a case.

In a final effort to create a paradox, we might have a go at filling this gap. It seems very plausible that no subject, not even a perceptually and cognitively ideal subject, could demarcate precisely the boundary between the F and non-F cases. The ability so to demarcate resemblance-based concepts would be truly superhuman, not merely an idealization of our actual abilities. Let’s call this the No-Demarcation thesis:

(ND) It is not knowable which cases lie on the borderline between two resemblance-based properties.

Do S3′, CKR-Pos, and ND form an inconsistent triad? Not quite, but they might be thought to pose a serious problem. The combination of S3 and CKR-Pos entails that there are cases that cannot be known to be F and cannot be known to fall under any contrary property. The thesis ND simply demands that these cases not be knowable as borderline cases. WE could get a contradiction if we could add either CKR-Neg or negative "introspection" with respect to knowability:

(NIK) If it is not knowable that p, then it is knowable that it is not knowable that p.

However, neither CKR-Neg nor NIK are especially plausible principles. Nonetheless, we are owed some explanation of why we cannot identify the borderline cases, given our capacity (under ideal conditions) of correctly classifying all of the non-borderline cases. It is one thing to show that there is no plausible argument for thinking that we are able to find the sharp boundary between contraries; it is quite another to explain why the discovery of such sharp boundaries is impossible. It is to this second task that we now turn.

3. The Margin-of-Error Theory

The most popular solution to this problem is the margin-of-error theory of Roy Sorenson (1988) and Timothy Williamson (1990). On this view, it is knowable that a case x(m) is F only if x(m) lies some finite distance from the F-boundary. Even if x appears to be F, it cannot be known to be F if it lies too close (in terms of resemblance) to cases that are not F. This account involves rejecting what I've called CKA-Pos:

(CKA-Pos) it is knowable that x resembles y more than z iff x appears to resemble y more than z.

According to the margin-of-error theorists, the right-to-left implication holds only in those cases in which the distance in apparent resemblance lies above some threshold of discrimination, a margin for error.

As I noted in 1994 (Koons 1994, p. 444), the margin-of-error theory has an unacceptably strong consequence concerning the possibility of fully reflective knowledge. Where F is a resemblance-based concept, knowing that something is F requires a margin of error. But then knowing that I know that something is F requires a margin of error that is roughly twice as great, since I have to verify, not only that the thing is F but that it is far enough from the nearest non-F cases to be a suitable candidate for knowledge. Similarly, each level of reflective knowledge requires another roughly equal increment of safety. Consequently, for every case of the satisfaction of a resemblance-based concept, there will be some number n such that reflective knowledge to level n is not possible. Thus, infinitely reflective knowledge is never possible: one is never in a position to know that one knows that one knows, etc., to any arbitrary number of iterations.

But this means that common knowledge is impossible, since a fact that p is common knowledge within a group only if everyone knows that everyone knows that... everyone knows that p, to any arbitrary number of iterations. Such common knowledge plays an important role in a number of game-theoretic and linguistic explanations. Ruling it out is an unacceptably high cost of the margin-of-error theory.

In addition, it seems implausible in itself. Consider a paradigm case of redness, for example. It seems plausible that I can know that I know that it is red, and so on, to any number of iterations of knowledge. Its redness is so obvious that no doubt can creep in at any level of reflection.

4. The Indiscriminability Account

Crispin Wright's proposed solution to the problem appeals to the fact of pairwise indiscriminability (Wright 1975). Two cases are indiscriminable if it is impossible to tell, simply by comparing one to the other, whether they agree or disagree in appearance. Let's call cases of indiscriminability cases of 'matching':

> (Matching) Cases x and y match iff it is not knowable (by simple, case-to-case comparison) whether they agree or disagree in appearance.

It is quite plausible to think that if two cases are close enough in appearance, then they match in this sense, even when we suppose ideal capacities of perception and cognition. The *sorites* series demonstrate that matching is a non-transitive relation: it can happen that cases x and y, y and z match, but x and z do not match. If we assume that we cannot discriminate matching cases, then we would have an explanation for the No Demarcation thesis. We cannot identify borderline cases because there are matching non-borderline cases, and we cannot discriminate between matching cases.

However, as Russell, Goodman, and Burgess have pointed out (Russell 1940, Goodman 1951, Burgess 1990), it is simply false that we cannot discriminate between matching cases. Since matching is a non-transitive relation, we can distinguish between matching cases that differ in appearance by discovering which other cases they match and do not match. Let's say that two cases 'super-match' just in case they match exactly the same cases;

> (Super-matching) Cases x and y super-match iff for all z, x matches z iff y matches z.[1]

[1] Burgess (1990) argues convincingly that, since observable qualities are extended in space and time, super-matching should be defined more strictly: x super-matches y iff for all z, z matches any part of x iff z matches any part of y.

It is an immediate consequence of this definition that super-matching is transitive (in fact, it has to be an equivalence relation). We can assume that if two cases differ at all in appearance, then they do not super-match. The difference in appearance will make some difference to which possible third cases each matches.

(Revelation) If cases x and y differ in appearance, then they do not super-match.

Suppose that we have two matching cases, x and y, where x is a true F-borderline case and y is not. Let's further suppose that y is F given Revelation, x and y do not super-match. Hence, there is some case z such that either x matches z and y does not, or y matches z and x does not. We can assume that if two cases do not match, it will be clear that one more closely resembles some paradigm than does the other. So, if x matches z but y does not, this must be because y more closely resembles the F-paradigm than z does. Since z and x match, this means that y must resemble the F-paradigm more than x does. Similarly, if y matches z and x does not, this must be because z more closely resembles the F-paradigm than x does, and, since y and z matches, this means once again that y more closely resembles the F-paradigm than x does.

Thus, we can use super-matching to distinguish borderline cases from non-borderline cases. We still need an explanation of the No Demarcation Fact.

5. Semi-Decidability

Let's return to the question of CKA-Pos versus CKA-Neg. Let's suppose that case x is somewhere near the borderline between being F and being G (where G corresponds to one of the competing paradigms with respect to which the intension of F is determined). To judge whether x is F or G or neither is a matter of comparing two degrees or intensities of resemblance: that between x and the F-paradigm and that between x and the G-paradigm. Compared simply to one another, we may have a case of matching: we may be unable to tell which degree is greater, and so we may be initially in doubt whether x is an F or not.

However, matching does not exhaust our epistemic resources in the ideal case. We can also appeal to Russell-Goodman super-matching. We can try to find a case of resemblance that matches one of the two degrees and not the other. If this third case matches the distance between x and the F-paradigm but is clearly greater than the distance between x and the G-paradigm, then we can conclude that x is G. If the third case matches the distance between x and the F-paradigm but is clearly less than the distance between x and the G-paradigm, we can conclude that x is F. If x is not precisely midway between the F and the G-paradigms, and if Revelation is true, there will be some case which enables us to discern on which side of the F-G boundary x lies. This vindicates CKA-Pos and CKR-Pos: any case of being F can be known to be such.

However, the epistemic situation is very different for a case lying on the boundary between F and G. We will never be able to verify that the two distances are exactly the same: we can only repeatedly fail to verify that they are different. In other words, being a borderline case is only *semi-decidable*: we can know that something is not on the borderline if it is not, be we can never know that something is on the borderline if it is. Since borderline cases will be neither F nor G, CKA-Neg will fail for any resemblance-based concept. This provides us with an explanation of the No Demarcation thesis, and it does so without threatening in anyway the possibility of infinitely iterated or common knowledge.

In addition, this explanation does not threaten to ramify into a global skepticism about empirical facts, since the measure of borderline cases is zero and the measure of non-borderline cases is one. We have good reason to be confident, a priori, that we will never actually encounter a borderline case, and so we can legitimately treat the ideal case as one of fully grasping the resemblance-based concepts, despite the semi-decidability of the borderline cases. A sufficiently close approximation to this ideal can serve as the standard for justification and even for knowledge.

The margin-of-error theory suffers from both of these defects. Instead of supposing that at each stage of inquiry, there are some unrecognized cases of F, the margin-of-error theorist supposes that there are some cases of F that are unrecognized at every stage of inquiry (since they lie within some margin of safety around the boundary). As I have argued, this makes common knowledge impossible. It also entails that it is metaphysically impossible to grasp any resemblance-based concept fully. If we cannot grasp these concepts fully, we have no way of judging just how wide the margin for error needs to be. Hence, we have a presumptive and indefeasible doubt about any of our judgments involving resemblance-based concepts, resulting in a global skepticism about observational knowledge.

I have shown that the Russell-Goodman criterion of super-matching can be applied to explain how an ideal wielder of a resemblance-based concept can successfully classify all positive instances of that concept, without being able to demarcate a precise boundary around that concept. This did not require identifying color-concepts with "Goodman-shades," very fine-grained concepts that pick out a single, determinate shade of color. Instead, I used a realistic model making use of paradigms and a comparative resemblance relation to carve a finite number of coarse-grained intensions from a continuum of determinate properties.

A final observation about the limitations of this result. I can explain the impossibility of discovering the sharp boundary between contrary sensible qualities, but my account does not apply to delimiting the parts of apparently vague objects, like Mt. Everest. Suppose, for example, that we assumed that Mt. Everest is identical to that candidate precise collection of atoms that most resembles our paradigm of a mountain, say, Mt. Fuji. The problem is that all such paradigm mountains are themselves vague. In the case of sensible qualities, in contrast, it was plausible to suppose that there existed qualitatively isolated paradigms, like pure white or black.

We might try to extend the solution by applying it to the material composition relation itself: some atoms joingly *compose something* if and only if they stand in a relation that more closely resembles paradigms of composition than it resembles the paradigms of non-composition. However, this cannot work, since material composition is itself a metaphysically isolated relation – one not, therefore, amenable to such a resemblance-theoretic account. In addition, we would still face the problem of finding non-vague paradigms of composition.

It is not clear, however, that epistemicism has a problem to solve in the vicinity of the vagueness of composition in the first place, since there is no reason to think that sensible *particulars* (in contrast to sensible *qualities*) fully reveal themselves in our experience. We are not surprised to learn that our sensory experience largely ignorant of the internal constitution (including the precise limits of the inclusion of microscopic parts) of the things we perceive. If we were to embrace purely phenomenal entities, like sense data as traditionally conceived, then we might face the threat of further sorites paradoxes concerning their spatial extent. Sorites arguments might well be taken as further evidence against the existence of such entities.

Robert C. Koons
University of Texas

Bibliography

Burgess, John A. (1990), "Phenomenal qualities and the nontransitivity of matching," *Australasian Journal of Philosophy* 68:206-20.

Burgess, John A. (2001), "Vagueness, epistemicism, and response-dependence," *Australasian Journal of Philosophy* 79:507-24.

Goodman, Nelson (1951), *The Structure of Appearance* (Cambridge, Mass.: Harvard University Press).

Koons, R. C. (1994), "A New Solution to the Sorites Paradox," *Mind* 103:439-449.

Lewis, David K. (1983), "New Work for a Theory of Universals," *Australasian Journal of Philosophy* 61:343-77.

Rescher, Nicholas (2009), *Unknowability: an enquiry into the limits of knowledge* (New York: Lexington Books), chapter 7.

Rodriguez-Pereyra, Gonzalo (2002), *Resemblance Nominalism: A Solution to the Problem of Universals* (Oxford: Clarendon Press).

Russell, Bertrand (1940), *Inquiry into Meaning and Truth*, London: G. Allen and Unwin.

Sorenson, Roy (1988), *Blindspots*, Oxford: Oxford University Press.

Sorenson, Roy (2001), *Vagueness and Contradiction*, New York: Oxford University Press.

Williamson, Timothy (1992), *Identity and Discrimination*, Oxford: Blackwell.

Williams, Timonty (1996), "What Makes it a Heap"? *Erkenntnis* 44: 327-339.

Wright, Crispin (1975), "On the coherence of vague predicates," *Synthese* 30:325-65.