Essential Vagueness: Two Models, One Simple Truth

Patrick Grim

Forthcoming in Ali Abasenezhad and Otavio Bueno, eds., *On the Sorites*, Springer

I. The Simple Truth of the Sorites

What the Sorites has to tell us is a simple truth regarding our categories. It appears to saddle us with something other than a simple truth—something worse, a contradiction or a problem or a paradox—only when we insist on viewing it through a discrete logic of categories.

Discrete categories and discrete logic are for robots. We aren't robots, and the simple truth is that we don't handle categories in the way any discrete logic would demand. For us non-robots, what the Sorites has to offer is a straightforward truth regarding how incapable robots and their logic are of handling categories like ours.

Categories come first. Discrete logic comes later.[1]

Categories are the core of conceptualization. Conceptualization is not merely something we do, it is something we can't do without, and we need categories in order to do it. But given the kinds of beings we are, our needs are rough and ready. Rough and ready categories satisfy those needs perfectly—in fact *only* rough and ready categories *could* satisfy those needs. In order to allow the kinds of things that kinds of beings like us do, the requirement is that categories must be clearly applicable in certain cases, and clearly inapplicable in certain other cases. 'Clearly applicable' means clear to us. That's the only requirement. That is all our categories need.

It need not be that for every case a category is discretely applicable or discretely something else. It need not be that for every case a category is either (check one) □ applicable or

□ something else: not applicable, non-applicable, or appropriately met with silence. Those are the demands of a discrete logic, not demands of categories per se. Nor need it be the case that for every case a category is either applicable or something else to some specific degree. That is a demand of a different discrete logic, but a logic that is discrete nonetheless. What makes a logic discrete is not the character of the 'something else' but the fact that we have a forced choice either/or between (check one) □ applicable or □ something else. A something else that is 'both applicable and not applicable' still gives us a discrete logic. A something else that is 'neither applicable nor non-applicable' does the same. The demands of any discrete logic are demands above and beyond the demands of categories per se. In order to operate, categories must be clearly applicable in certain cases and clearly inapplicable in certain others. 'Clearly applicable' means clear to us. That is all that they require.

Consider how our categories, conceptual or linguistic, are learned. We learn cases in which they clearly apply. 'Truck.' 'Dog.' We are corrected in cases in which they clearly do not apply. 'No, Johnny, that's not a truck.' Occasionally we may be given something subtler—a warning that we have strayed, but not too far. 'Well, Johnny (laughs) I guess that's sort of a truck. But you don't often see one made with two legs…' The laughter is appropriate, an indication that we have in some way strayed from the clear cases of applicability and inapplicability for which we have such a category to begin with.

A discrete logic—more specifically, a discrete semantics—is one that demands a semantic value for every case. There is some reservoir of distinct semantic values for propositions, and each proposition gets one. Structures of distinct elements can have any of various forms, and discrete logics have explored many of those forms. The reservoir may be beyond the size of sets. One of the semantic values may be 'no semantic value.' Gapped, glutted, classical, intuitionistic, intensional, and various forms of ascending structures are all discrete logics. Those are for various types of robots. We aren't any type of robot, and when it comes to our categories discrete logics merely need us astray.[2] All our categories require are cases of clear applicability and cases of clear inapplicability. That requirement carries no commitment to some semantic value appropriate to every case. There is indeed no commitment to anything about 'every case' at all.

II. Essential Vagueness

There have been many marvelously clever attempts to 'address' the sorites by reshaping a discrete semantics so that we can do some kind of justice to our conceptual categories.[3] None of these is entirely successful, and none can be. The simple truth of the sorites is that our categories don't operate in accord with *any* discrete semantics. What the sorites has to tell us is that models of conceptualization built in terms of a discrete semantics will fail. But if no discrete semantics can offer an appropriate model, what can?

Categories come first: they are conceptually primary and come first in terms of what is learned. The sharp edges and quantification of a discrete logic come much later: they are conceptually secondary and come significantly later.

In the terms of any discrete logic our fundamental conceptual categories will always appear vague. In their own terms, for our normal purposes, our categories are perfect as they are; 'vague' may be a derogatory epithet that comes from the perspective of a later logic. From that later perspective, however, many of our fundamental categories must be recognized as *essentially* vague. Any account of our categories that replaces vagueness with precision will distort them by that token alone. The only way to build a successful account of our categories is to incorporate essential vagueness in that account itself.

In the following sections, with the sorites in mind, I sketch two attempts at more appropriate models of categories and categorization. It is important to emphasize that both of these are offered *only* as models. Both attempt to give a better picture of how our categories actually function, but like all models everywhere—including all logical models—what they offer is only a picture. The first model may be nicer in some regards, in that it may make the sorites understandable in terms closer to the kinds of logic with which we are familiar. But it has important limitations: it may not be adequate for all cases, and strays at important points from the psychological reality of categorical thinking. The second model is intended to fit closer with the psychological reality of our categories, but may lose some of the charm of formalization in the process.

III. The Buried Quantifier

Any successful account of our categories must build essential vagueness in at the base. In a first approach it is built in as a hidden vague quantifier.[4]

The sorites is standardly run on monadic predicates such as 'bald,' 'tall', and 'old,' for which there are clear comparatives or graded adjectives: 'balder', 'taller', and 'older.' It is not the monadic predicates but the comparatives that form the core of this first account.[5]

Philosophers have long been obsessed with just two quantifiers—'all' and 'some.' Linguists recognize far more, including 'a few,' 'several', 'many', 'lots of' and 'almost all.' None of these is reducible to expressions in terms of $\forall x$ or $\exists x$, and none are specifiable in terms of a precise number of percentage. These are essentially *vague* quantifiers.

It is another vague quantifier that is employed in this first model. 'Zx' will be used to represent a quantifier that might be expressed as:

For the great bulk of x's…

For the overwhelming majority of x's…

It is quite generally true of x's that…

It is characteristic of x's that…

For a large percentage of x's…

With relatively few exceptions among x's…[6]

Could I do so, standing there by your elbow, I would introduce Zx ostensively, as it might be introduced in teaching a language. With piles of rice, or handfuls of beans, or gaggles of gerbils I am confident I would be able to introduce you to the essentially vague quantifier Zx in a short period of time.

In this first model, the core vagueness of our categories is captured by understanding our monadic predicates derivatively in terms of comparatives and a vague buried quantifier. Someone is 'bald' just in case they are balder than the great bulk of the comparison class. They are 'tall' if they are taller than the overwhelming majority of those in the comparison class. Someone is 'old' if it is quite generally true that they are older than the others. Using Bx for 'x is bald' and xBy for 'x is balder than y,' we can outline the monadic predicate 'bald' as:

$Bx \equiv Zy(xBy)$

The quantifier Zx is indefinite in two ways that very intuitively model our use of categories. There is, first of all, no answer to the question 'How many (or what proportion) constitute "the great bulk of"?' Like the quantifiers 'many', 'lots of', 'almost all', and 'a few,' Zx does not correlate with any definite number or proportion of a population. Like our categories, it is essentially imprecise.

Zx is also indefinite as to the comparison class. Balder than the great bulk of whom? Balder than the overwhelming majority of what group? These aspects of use are functions of context, and can only function in context. Because ‘bald’ is modeled in terms of Zx over a comparison class, ‘Bill is bald’ will be true only with reference to a particular comparison class. This holds for our categories in general: the hirsute man at the bald men’s convention may be the bald man at the hirsute men’s convention. The old kid in kindergarten is 6. The young guy in the nursing home is 60.

The comparison class, it should be noted, need not be set by present or actually existing individuals: a comparison class may remain contextually salient despite the demise of a large number of its members, for example. Were we to kill off all men over 5’8”, the class of recently existing men might remain as the comparison class. Context can also target comparison classes of normal individuals, ideal subjects, or a range of mere possibilities.[7]

Zx is an essentially vague quantifier. For such a quantifier it is clear that the logic appropriate to ∀x or ∃x will fail. Consider a forced march sorites using a comparison class of men numbering in the millions, the relation ‘x is balder than’ mapped onto that class, and ‘x is bald’ understood as ‘balder than the overwhelming bulk of’ those men. As we step down the line of candidates from the most bald to the least, will there be a particular point at which we pass from an individual who is ‘balder than the overwhelming bulk of’ the members of the group to a next individual who is not ‘balder than the overwhelming bulk?’ Clearly not: there will be no transition step such that:

∃x∃z(Zy(xBy) & z is next in line behind x & ~Zy(zBy))

The category ‘bald’ dictates no step in the forced march at which we transition ‘bald’ to ‘not bald.’ On this model, there is no such step precisely because there is no step at which we go from ‘balder than the great bulk of’ to ‘not balder than the great bulk of.’

If we deny the existence of a transition step, are we not forced to hold the classical equivalent of its negation,

$\forall x \forall z(Bx\ \&\ z \text{ has one more hair than } x \rightarrow Bz)$?

No. Quantifier Negation holds only on the assumption of excluded middle, which here comes down to the assumption that $\forall x(Bx \vee {\sim}Bx)$. But on the current account this abbreviates $\forall x(Zy(xBy) \vee {\sim}\, Zy(xBy))$. It is clear from the nature of Zx that excluded middle in this form simply will not hold. Excluded middle fails for *Zy(xBy)* and thus for *Bx*, and Quantifier Negation fails with it.

A modeling in terms of a buried vague quantifier seems to capture a number of important features of categories as we really use them. It also offers a nice formal take on the standard formulation of the sorites. The model makes it clear that the standard induction principle will be false, and makes it clear why it will be false:

$\forall x \forall z(Bx\ \&\ z \text{ has one more hair than } x \rightarrow Bz)$

But if the induction principle is false, why are we so tempted to think it is true (Fara 2000)?

As indicated above, fuzzy logic remains a discrete logic. But one of the appealing aspects of a fuzzy logic is that the induction principle, though not true, is very close to being true.[8] The buried quantifier offers a similar prospect. Although the induction principle is false, and must be so, it has a close relative that may well be true. Phrased in terms of our vague quantifier:

$Zx \forall z(Bx\ \&\ z \text{ has one more hair than } x \rightarrow Bz)$

All that has changed is the first quantifier. Unlike the induction principle, this is not a universal quantification over x at all. It is an essentially vague quantification: for the great bulk of cases, the overwhelming majority of cases, if x is bald and z has one more hair then z is bald as well. Because it holds for 'the great bulk of cases', it may be quite generally usable as a rule of thumb in any isolated case. But because it falls short of a full universal quantification it doesn't saddle us with a paradoxical sorites.

IV. Limitations of the Buried Quantifier

The buried quantifier offers an attractive model of some of the phenomena of categorization, with a nice payoff in terms of understanding where the sorites goes wrong in discrete logic. But there are also reasons to think that it cannot be adequate as a complete account.

There are cases that seem to run contrary to the account as presented. Kamp (1975) notes that most English cars are small, a fact that seems to turn on salient size categories for rather than on sizes of the vast majority of cars. Fara (2000) constructs a case in which a teacher divides a class between an A team of taller children and a B team of shorter children, after which the A children are naturally spoken of the tall kids and the B children as the short kids—despite the fact that division might have been made at a different point and despite the relative numbers on the two teams.[9]

The essential *vagueness* of the buried quantifier seems right. But there may aspects of *salience* in categorization which a quantifier like 'greater than the vast bulk of' or 'greater than the vast majority of' imperfectly captures because salience is not always even vaguely a matter of number or proportion. Categorization may rely on salience in ways that a vague quantifier will fail to capture precisely because it remains a quantifier.

There is another criticism that goes along with this one. We do not seem to use our categories in terms of quantifiers, however deeply buried, vague or not. The psychological evidence seems to be that our cognitively fundamental generalizations are generics such as

Birds fly

Ducks lay eggs

Ticks carry lime disease

Lions have manes

These are default generalizations, defeasible with further information—birds that are penguins do not fly—keyed to features taken to be salient for a kind. Most interestingly, the use of generics flies in the face of quantification: not all ducks lay eggs, nor do the vast majority, because it is only female ducks who lay eggs. It is only male lions that have manes. The third claim is taken to be true despite the fact that it is very small proportion of ticks that carry lime disease (Krifka et al 1995; Leslie 2007, 2008, 2012).

However categories are *learned*, moreover, they do not seem to be learned in terms of quantifiers, however deeply buried, vague or not. Infants in their first year are able to use generic categories in the sense above, dealing with new instances in terms of past cases and generic generalizations (Baldwin, Markman, & Melartin, 1993; Graham, Kilbreath & Welder, 2001). By the age of 3 or 4, children produce generics and draw inferences from them much as adults do (Gelman, Goetz, Sarnecka & Flukes 2008). But children of such an age do not handle kind-wide quantification in the way adults do (Hollander, Gelman, and Star 2002). Particularly telling against a buried quantifier account is the fact that 'most' seems very difficult to process at an early age (Papafragou and Schwartz 2005/2006).

Categories come first. Quantification, even vague quantification, comes later. Can we build a model that is closer to the psychological reality of category learning and use?

V. The Buried Relation

In this second model, monadic predicates like 'bald' are again analyzed in terms of something non-monadic. But here what is buried is not a quantifier but an essentially vague relation. Buried beneath all of our categorization, on this second model, is a relation of *similarity*. It is that relation that is crucial to our learning of categories. It is that buried relation that is central tour use of categories and to conceptualization in general. And that relation is essentially vague.

Part of the attempt in this second model is to construct an account that accords better with the psychological literature. How do people learn and employ categories—categories like 'bald' and 'tall,' for example? There are a number of theories going, but the two major contenders are prototype and exemplar theories.[10] Both function explicitly in terms of a concept of similarity.

In prototype theory, a single abstracted prototype forms the core of a category, with application of that category learned and applied in terms of salient similarity to the prototype (Posner & Keele 1968, 1970; Reed, 1972; Rosch 1975, 1983; Lakoff 1990; Smith & Minda, 2001; Taylor 2003; Gärdenfors 2004, 2017; Hampton 2006, 2017). In exemplar theory, which I will emphasize here, it is not a single abstracted prototype but a set of very concrete exemplars that are taken as the core of a category, stored with all details in place. Categories are learned and applied in terms of salient similarity to those exemplars (Medin & Schaffer 1978; Nosofsky1986, 2014; Estes 1994; Storms, De Boeck & Ruts 2000; Rouder & Ratcliff 2006; Reisberg, 2015). Often categories are formed and applied in terms of salient similarity to sets of both exemplars and counter-exemplars: core exemplars of things to which the category applies and of things to which it does not.

A category is learned first and foremost by acquiring a set of exemplars. It is often learned as a set of contrasts, each of which will have exemplars. 'Bald' may have a set of exemplars, 'not bald' another set. 'Red,' 'black,' and 'white' may all have sets of exemplars. From there on category use is dictated by the core concept of similarity: It is elements that are saliently similar to the exemplars for which the same term is used. It is when a new example is similar enough to the exemplars of bald that is categorized as bald, when it is similar to 'not bald' exemplars instead that it is categorized as 'not bald.' Hidden beneath the apparent monadic 'bald', on such an account, is a buried relation: 'bald' functions as 'is saliently similar to exemplars a, b, and c,' where a, b, and c are the exemplars for the term:

$$Bx \equiv Sx\{a,b,c\}$$

In exemplar theory, it should be emphasized, a, b, and c are not predicate-defined. Exemplars are stored lock stock and barrel, with all properties in place. They are property multi-dimensional, but those properties need not be named or even conceptualized. It is full-blooded full-property representatives that moor the end of a concept. Similarity judgments will be multi-dimensional as well, and one may decide later that someone really is bald on the basis of an element of similarity to exemplars that did not originally occur to one: it is the pattern of hair distribution among the exemplars that matters, perhaps, rather than the number of hairs. In the psychological literature it is put forward as a virtue of exemplar accounts that relevant properties may not themselves be conceptualized or distinguished, and may not even be 'linear' (Ashby & Maddox 2005).

The philosophical history of similarity is a sketchy one.[11] Nelson Goodman complained that appeals to similarity are promiscuous: any two things will be similar in some regard. "…'[I]s similar to' functions as little more than a blank to be filled" (Goodman 1972). Any proper analysis, he insisted, must give a reductive definition of the similarity in question: similar with respect to what predicate?

In the current model similarity is taken as a primitive, unanalyzed in other terms, much as in its application between possible worlds in David Lewis's *Counterfactuals* (1973).[12] Because it cannot be explained in more basic terms, Quine rejects any general notion of similarity as "logically repugnant," unsuitable for any mature science (Quine 1969, Weisberg 2013). That is precisely the point. Categories come first, conceived in primitive similarity. The discrete logic appropriate to a mature science comes much later. In the spirit of Goodman and Quine, it is tempting to say that two things are similar if they share a predicate: if both are P. Or that two things are similar if they share many predicates: both are P, Q, and R.[13] If two things are similar, and if our language is rich enough, this will be true: similar objects will share predicates. But on the current analysis it is the latter that is derivative of the former, not the other way around. The two things are similar to begin with, in the full primitive and exemplar-based sense. Predicate sharing is derivative, relative to the richness and extent of gerrymandering in the language in which those predicates appear.[14]

No-one is tempted to think that the relation of similarity is other than vague. No-one is tempted to think that everything we might name is either similar to a dead lizard or not, for example. There's the lizard. Here's a concrete statue of a duck. Quick, now, yes or no: is the statue similar to the dead lizard or not? The primitive relation of similarity forms the essentially vague core of the second model.

An essentially vague relation operating beneath our monadic categories, tied for each to unnamed exemplars, accords both with the available psychological and linguistic literature and explains many of the familiar features of vague predicates. Vague categories can be multi-dimensional—as are 'nice' and 'nasty,' 'wise' and 'reckless'—just as is the underlying concept of similarity. Vague categories can be salient-sensitive (and not merely 'great bulk of') sensitive, just as the underlying concept of similarity can be salient-sensitive. Vague categories are often context- and purpose- relative, just as is the underlying concept of similarity.

There is no more temptation to think that salient similarity will be discretely quantized than that ‘the great bulk of’ will be. Salient similarity seems to match the virtues claimed for a buried quantifier account, but without some of the drawbacks. We can expect the set of exemplars to be relative to a comparison class set by context. Exemplars of ‘bald’ in the context of the bald man’s convention will be different from those at the hirsute man’s convention. Salient similarity will shift accordingly: the bald man at the hirsute man’s convention may be the hirsute man at the bald man’s convention. Here, however, salience is set by exemplars and need not be set by a quantificational ‘vast bulk of,’ however vague. In terms of salient similarity set to a group of exemplars, Kamp’s English cars may be small and the children in Fara’s group A may be tall regardless of their quantificational proportions.

Similarity, of course, comes in shades of more or less: things can be judged to be more or less similar to a set of exemplars. The result is an ordinal ranking, or something like it. It does not suggest cardinality: something is not similar 0.7 to a given object, as opposed to similar 0.8. Relative similarity generates ‘something like’ an ordinal ranking in that it does come with a metric, but that metric is itself vague. One thing may be *much* more similar to a given object than something else. It is as if we can vaguely indicate large and small distances along a spectrum of similarity shades. But one thing is never precisely twice as similar to a given object as another.

What kind of analysis does an exemplar model offer for a forced march sorites?

Our exemplars, we presume, will be at one end. As we step down the line of candidates we find them progressively less similar to our exemplars. Because ‘bald’ is a matter of salient similarity, they become progressively less bald. But note that at each step we evaluate 'bald' in terms of salient distance from our original exemplars. Each step is judged in terms of distance from the far end. This is distinctly different than the pattern of reasoning used in the standard sorites argument. There we are urged to judge each step in terms of some minimal distance from *the previous step*. If x is B, and y is close to x, y just be B. But if y is B, and z is close to y… The similarity judgments of the present account, in contrast, are always moored to the initial set of exemplars at the extreme end.

As we step down the line of candidates ordered from the most bald to the least, we will find them progressively less similar to our exemplars. But precisely because what is at issue is a *progressive* distancing, there will be no discrete point at which we pass from a discrete case of 'saliently similar' to one of 'saliently dissimilar.' As noted above, categories are often learned in terms of exemplars and counter-exemplars: exemplars for both 'bald' and 'not-bald.' In asking whether a candidate in the middle is bald or not, then, one may not be considering one predicate but two, not similarity to one set of exemplars but two or in some cases more. Salient similarity may therefore fade out from both ends, or may form an interference pattern: a candidate may resemble both exemplar sets in important respects, or resemble neither set enough to determine the case. As emphasized in a number of contextual analyses of the sorites that have emphasized shifting judgments, it may also be that 'saliently similar' shifts in context as we move down the line (Raffman 1994, 1996, Shapiro 2003, Soames 2003).

Because of the essential vagueness of salient similarity *S*, there will then be no transition step such that:

$$\exists x \exists y (Sx\{a,b,c\}) \ \& \ y \text{ is next in line behind } x \ \& \sim Sy\{a,b,c\}))$$

There is no step at which our categories go from 'bald' to 'not bald'. On this model, that is precisely because there is no step at which we go from 'saliently similar' to our class of exemplars to 'saliently dissimilar.'

But if we deny the transition step, won't we have to affirm this classical equivalent?

$\forall x \forall z(Bx \ \& \ z \text{ has one more hair than } x \rightarrow Bz)$?

No, for the same reasons cited in the case of the vague quantifier. Quantifier Negation holds only on the assumption of excluded middle, which here comes down to the assumption that $\forall x(Sx\{a,b,c\} \vee \sim Sx\{a,b,c\})$, and it is clear from the essentially vague nature of salient similarity *S* that excluded middle in this form simply will not hold. Excluded middle fails for $Sx\{a,b,c\}$ and thus for *Bx*, and Quantifier Negation fails with it.

An important virtue of the vague buried quantifier is that it offers us a plausible replacement for the induction principle. On that account it is not and cannot be true that

$\forall x \forall z(Bx \ \& \ z \text{ has one more hair than } x \rightarrow Bz)$.

But there is something close which might well be true:

$Zx \forall z(Bx \ \& \ z \text{ has one more hair than } x \rightarrow Bz)$.

Can the current account offer a corresponding explanation for why, despite its falsity, the induction principle appears so plausible (Fara 2000)?

Just as the buried quantifier account appeals to a vague second-order quantification over cases, a buried relation account can appeal to a second-order application of a similarity relation. Similarity is not transitive. But there will be cases—indeed many cases—in which the field of salient similarity will be such that two neighboring items f and g will both be saliently similar to a set of exemplars. *f* will be very saliently similar to our exemplars, perhaps, *g* will be just a hair less similar, and as a result g will fall easily within the context-defined range of salient similarity:

Bf & g has one more hair than f → *Bg*

Consider now a handful of cases in which this clearly holds. Take those as our inferential exemplars. There will then be range of cases that are more or less similar to *those* exemplars. The smaller the 'one more hair' distance, the closer those cases will be to our exemplars. The more salient the similarity of *f* to begin with, the closer those cases will be.

The result is a categorical pattern of salient similarity repeated on a second level: on the category of cases in which one can infer baldness, say, from the baldness of a close case.

Discussion of the psychology of categorization above emphasized the role of generics. We do not seem to use our categories in terms of quantifiers. The psychological evidence is that our cognitively fundamental generalizations are not quantificational but default and defeasible generalizations such as

Birds fly

Ducks lay eggs

Lions have manes

Given a second-order application of salient similarity to our exemplars of legitimate inference in close cases, it is the generic character of our categories that offers an explanation for why the induction principle proves so tempting. What is true in the case of the sorites is not the universal generalization that is the discrete logician's rendering of the induction principle. What is true instead is a generic, which we might mark as such:

[Gen] If one man is bald and another differs by just one hair, he too is bald.

This is a true generalization—a true generic or categorical generalization. But it is only misconstrued if taken as a universal quantification. What it generalizes are characteristic, typical, salient, or representative cases of comparison. It remains a true generalization precisely because of those characteristic, typical, salient, or representative cases…just as it is true that ducks lay eggs, birds fly, and lions have manes.

VI. Two Models, One Simple Truth

There is a simple truth that the sorites is trying to tell us: that our categories do not and cannot function in the manner imposed by any discrete semantics.

Discrete logic is for robots. We aren't robots, and the simple truth is that we don't handle categories in the way any discrete logic would demand. For us non-robots, what the sorites has to offer is a straightforward truth regarding how incapable robots and their logic are of handling categories like ours.

The simple truth is that our fundamental categories are essentially vague. Is vagueness an essential property of categories *per se*? I see no reason to think so. Could there be a different set of categories which *would* obey a fully discrete or even classical logic? I see no reason to think not. That is what Goodman and Quine seemed to envisage as the language demanded by a mature science. That is what the epistemicists seem to think our categories must actually be like, applying or failing to apply independently of us and beyond our ken. That is what Wheeler and Quine attempted to construct by abandoning swizzle sticks in favor of 2-million-atom swizzle sticks (Wheeler 1979, Quine 1981).

I see nothing against such categories except that they are not ours, that they are quite generally useless for the purposes of creatures like us, and that they are in fact radically unapproachable from the categories with which we actually operate (Grim 1982). Our categories are of a kind appropriate to our conceptual needs, learned and applied in ways appropriate for creatures like us. The others are categories appropriate for robots. They are a discrete logician's attempt to break all conceptualization on the rack of a discrete logic. The simple truth is that our normal categories don't work in terms of the constructs of such a logic. Don't. Shouldn't. Can't.

If we are to learn the lesson of the sorites and not conceive of our categories and conceptualization in the terms of a discrete logic and discrete semantics, how are we to conceive of them? My attempt here has been to offer two models for a different approach that recognizes the essential and ineliminable vagueness of our categories. In one case an essential element of vagueness is modeled in terms of a vague buried quantifier. In another, with perhaps a better claim to psychological realism, it is modeled as the vague buried relation of similarity.

Both remain models of conceptualization, intended to retain at least some of the appeal of formalization. Both remain models in very much the sense that all logic has always been a model of conceptualization. The hope is that they may offer a route to better models…models that have learned the simple truth of the sorites.

References

Ashby, F. G. & W. T. Maddox (2005). "Human Category Learning," *Annual Review of Psychology* 56: 149-178.

Attneave, F. (1950). "Dimensions of Similarity," *American Journal of Psychology* 63, 5`6-556.

Baldwin, D. A., E. Markman & R. Melatrin (1993). "Infants' Ability to Draw Inferences about Nonobvious Object Properties: Evidence from Exploratory Play," *Child Development* 64, 711-728.

Barsalou, L. (2010). “Grounded Cognition: Past, Present, and Future.” *Topics in Cognitive Science* 2, 716-724.

Bruner, J. S., J. Goodnow, & G. Austin (1956). *A Study of Thinking*. New York: Wiley.

Estes, W. K. (1994). *Classification and Cognition*. New York: Oxford University Press.

Fara, D. G. (2000). “Shifting Sands: An Interest-relative Theory of Vagueness.” *Philosophical Topics* 28, 45-81.

Field, H. (2003). “The Semantic Paradoxes and the Paradoxes of Vagueness,” in JC Beall, ed., *Liars and Heaps*. Oxford: Clarendon Press.

Fine, K. (1975). “Vagueness, Truth and Logic.” *Synthese* 54, 235–59.

Forbes, G. (1994). *Modern Logic*. New York: Oxford University Press.

Gärdenfors, P. (2004). *Conceptual Spaces: The Geometry of Thought*. Cambridge MA: MIT Press.

Gärdenfors, P. (2017). *The Geometry of Meaning: Semantics Based on Conceptual Spaces*. Cambridge MA: MIT Press.

Goodman, N. (1972). “Seven Strictures on Similarity.” In Goodman, *Problems and Projects*. Indianapolis: Bobs-Merrill.

Graham, S. A., C. S. Kilbreath, & A. N. Welder (2001). “Words and Shape Similarity Guide 13-Month-olds’ Inferences about Nonobvious Object Properties.” In J. D. Moore & K. Stenning, eds., *Proceedings of the Twenty Third Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum, pp. 352-357.

Gelman, S. A., P. J. Goetz, B. S. Sarnecka, & J. Flukes (2008). “Generic Language in Parent-Child Conversations.” *Language Learning and Development* 4, 1-31.

Grim, P. (1982). "What Won't Escape Sorites Arguments," *Analysis* 42 (1982), 38-43.

Grim, P. (1997). "A Fuzzy Fairly Happy Face." New York Vagueness Group, http://www.pgrim.org/pgrim/fuzzy.htm

Grim, P. (2005). "The Buried Quantifier: An Account of Vagueness and the Sorites," *Analysis* 65, 95-104.h

Hampton, J. A. (2006). "Concepts as Prototypes." In B. H. Ross, ed., *The Psychologty of Learning and Motivatioon: Advances in Research and Theory* vol. 46. Amsterdam: Elsevier, pp. 79-113.

Hampton, J. A. (2011). "Conceptual Combination and Fuzzy Logic," in R. Belohlavek & G. J. Klir, *Concepts and Fuzzy Logic*, Cambridge MA: MIT Press, pp. 209-231.

Hampton, J. A. (2017). "Compositionality and Concepts." In J. A. Hampton & Y Winter, eds., *Compositionality and Concepts in Linguistics and Psychology*. London: Springer, pp. 95-122.

Heim, I. & A. Kratzer (1998). *Semantics in Generative Grammar*. Oxford: Blackwell.

Hollander, M. A., S. A. Gelman, & J. Star (2002). "Children's Interpretation of Generic Noun Phrases." *Developmental Psychology* 38, 883-894.

Kamp, J. A. W. (1975). "Two Theories about Adjectives." In E. L. Keenan, ed., *Formal Semantics of Natural Language*. Cambridge: Cambridge University Press, pp. 123-155.

Krifka, M., F. J. Pelletier, G. Carlson, A. ter Meulen, G. Chierchia & G. Link (1995). "Genericity: An Introduction." In G. Carlson & F. J. Pelletier, eds, *The Generic Book*. Chicago: University of Chicago Press, pp. 1-124.

Komatsu, L. (1992). "Recent Views of Conceptual Structure." *Psychological Bulletin* 112, 500-526.

Lakoff. G. (1990). *Women, Fire, and Dangerous Things: What Categories Reveal About the*

Mind. Chciago: Univ. of Chicago Press.

Laurence, S. & E. Margolis (1999). "Concepts and Cognitive Science." In S. Laurence & E. Margolis, eds., Concepts: Core Readings. Cambridge MA: MIT Press, pp. 3-82

Leslie, S. J. (2007). "Generics and the Structure of the Mind." *Philosophical Perspectives* 21, 375-405.

Leslie, S. J. (2008). "Generics: Cognition and Acquisition." *The Philosophical Review* 117: 1-49.

Leslie, S. J. (2012). "Generics Articulate Default Generalizations." *Recherches Linguistiques de Vincennes* 41, 25-45.

Lewis, D. (1973). *Counterfactuals*. Oxford: Basil Blackwell and Cambridge, MA: Harvard University Press.

Lewis, D. (1975). "Adverbs of Quantification." In E. Keenan, Ed., *Formal Semantics of Natural Language*. Cambridge: Cambridge Univ. Press, pp. 3-15.

Lewis, D. (1979). "Counterfactual Dependence and Time's Arrow." *Noûs* 13, 455-476.

Machery, E. (2009). *Doing Without Concepts*. New York: Oxford Univ. Press.

Medin, D. L. & M. M. Schaffer (1978). "Context Theory of Classification Learning." *Psychological Review* 85, 207-238.

Murphy, G. (2002). *The Big Book of Concepts*. Cambridge, MA: MIT Press.

Nosofsky, R. M. (1986). "Attention, Similarity, and the Identification Categorization Relationship. *Journal of Experimental Psychology: Gen*. 115, 39-57.

Nosofsky, R. M. (2014). "The Generalized Context Model: An Exemplar Model of Classification." In M. Pothos & A. Wills, eds., *Formal Approaches to Categorization*. New York: Cambridge University Press, pp. 18-39.

Papafragou, A. & N. Schwaz (2005/2006). "*Most* wanted" *Language Acquisition* 13, 207-251.

Posner, M. I. & S. W. Keele (1968). "On the Genesis of Abstract Ideas." *Journal of Experimental Psychology* 77, 353-363.

Posner, M. I. & S. W. Keele (1970). "Retention of Abstract Ideas." *Journal of Experimental Psychology* 83, 304-308

Quine, W. V. O. (1969). "Natural Kinds." In Quine, *Ontological Relativity and Other Essays*. New York: Columbia University Press.

Quine, W. V. O. (1981) "What Price Bivalence?" *Journal of Philosophy* 77, 90-5.

Raffman, D. (1994). "Vagueness without Paradox." *Philosophical Review* 103: 41-74.

Raffman, D. (1996). "Vagueness and Context Relativity." *Philosophical Studies* 81, 175-192.

Reed, S. K. (1972). "Pattern Recognition and Categorization." *Cognitive Psychology* 3, 382-407.

Reisberg, D. (2015). *Cognition: Exploring the Science of the Mind*. New York: W. W. Norton.

Rouder, J. N. & Ratcliff, R. (2006). "Comparing Exemplar- and Rule-based Theories of Categorization." *Current Directions in Psychological Science* 15, 9-13.

Rosch, E. (1975). "Cognitive Representations of Semantic Categories." *Journal of Experimental Psychology*: General 104, 192-233.

Rosch, E. (1983). "Prototype Classification and Logical Classification: The Two Systems," in E. K. Scholnick, ed., *New Trends in Conceptual Representation: Challenges to Piaget's Theory?* Hillsdale NJ: Lawrence Erlbaum, pp. 73-86.

Shapiro, S. (2003). "Vagueness and Conversation." In JC Beall, ed., *Liars and Heaps: New Essays on Paradox.* Oxford: Clarendon Press, pp. 39-72.

Shepard, R. N. (1980). "Multidimensional Scaling, Tree-fitting, and Clustering." *Science* 210, 390-398.

Smith, E. E. & D. L. Medin (1981). *Categories and Concepts*. Cambridge, MA: Harvard University Press.

Smith, J. D. & J. P. Minda (2001). "Journey to the Center of the Category: the Dissociation in Amnesia between Categorization and Recognition," *Journal of Experimental Psychology: Learn. Mem. Cogn.* 27, 984-1002.

Soames, S. (2003). "Higher-Order Vagueness for Partially Defined Predicates." In JC Beall, ed., *Liars and Heaps: New Essays on Paradox.* Oxford: Clarendon Press, pp. 128-150.

Storms, G., P. De Boeck & W. Ruts (2000). Prototype and Exemplar-based Information in Natural Language Categories." *Journal of Memory and Language* 42, 51-73.

Taylor, J. R. (2003). *Linguistic Categorization*. New York: Oxford Univ. Press.

ten Brinke, W., D. M. Squire, & J. Bigelow (2004). "Similarity: Measurement, Ordering, and Betweeness," in M. G. Negoita, R. J. Howlett, & L. C. Jain, eds., *Knowledge-Based Intelligent Information and Engineering Systems. Lecture Notes in Computer Science* Vol. 3214, pp. 996-1002.

Tversky, A. (1977). "Features of Similarity". *Psychological Review* 84, 327-352.

Tversky, A., & I., Gati (1987). "Studies of Similarity." In E. Rossch & B. Lloyd, eds., *Cognition and Categorization*. Hillsdale, NJ: Erlbaum.

van Fraassen, B. C. (1966). "Singular Terms, Truth-value Gaps, and Free Logic.' *Journal of Philosophy* 63, 481–95.

Weisberg, M. (2013). *Simulation and Similarity: Using Models to Understand the World*. New York: Oxford University Press.

Wheeler, S. C. (1979). "On That Which Is Not." *Synthese* 41, 55-73.

[1]Some of the psychological literature to this effect is touched on in section IV: Baldwin, Markman & Melartin, 1993; Graham, Kilbreath & Welder, 2001, Gelman, Goetez, Sarnecka & Flukes, 2008.
[2] A range of effects in the literature "all point to the fact that the human conceptual systems is not based on a firm grounding in logic (classical or otherwise) but instead has a different design with different purposes" (Hampton 2011, 224).
[3] Nobody asked, but among my candidates for the most clever are the three F's: Field 2003, Fine 1975, and van Fraaassen 1966.
[4] For a more complete treatment see Grim 2005.
[5] It is a virtue of an account based in comparatives that it seems to run in the right direction, deriving 'bald' from 'balder' rather than the other way around. Indeed, were 'bald' a monadic predicate obedient to the law of the excluded middle, it is unclear how any account could derive 'balder' from 'bald'. As Kamp notes, "It is quite obvious that if adjectives were ordinary predicates no such transformation could exist. How could we possibly define the relation *x is bigger than y* in terms of nothing more than the extension of the alleged predicate *big*?" (Kamp 1975, 127). Further support for the centrality of comparatives comes from the fact that it is so much more difficult to try to construct a sorites using vague nouns rather than adjectives: religion, politics, society. Such an attempt requires that these somehow be made gradable in ways that introduce comparatives.
[6] An adverbial interpretation might also be possible: x's are characteristically... or x's for the most part are... A useful comparison may be David Lewis's treatment of 'unselective quantifiers' (Lewis 1975).
[7] Much of the resistance to accounts grounded in comparatives stems from Kamp's classic (1975), but much of that opposition is founded on a short-sighted limitation to comparison classes of actual individuals or set by comparison nouns rather than the context as a whole.
[8] Using a Łukasiewicz conditional $/p \rightarrow q/ = \min[1, 1-/p/ + /q/]$, or equivalently $/p \rightarrow q/ = 1$ if $/p/ < /q/$, $1-\text{abs}[/p/-/q/]$ otherwise, all instances of the induction principle will be very true where two case of bald are close and thus claims p and q regarding them are very close in value. With progressive reapplication down a forced march, the conditional will nonetheless leak enough truth that one is not compelled to say that it is very true that a candidate sufficiently far down the line is bald (Forbes 1994, Grim 1997).
[9] The relevant comparison class need not be dictated by the noun at issue—cars or children, for example. Heim & Kratzer paint a scenario in which in the context of giant monsters 'Jumbo doesn't stand a chance...he's just a tiny elephant,' even though Jumbo may in fact be quite large among elephants.
[10] Prototype and exemplar theories have largely eclipsed classical rule-based theories of categorization, which assume that every category is represented by a set of necessary and sufficient features (Bruner, Goodnow & Austin 1956; Smith & Medin 1981, Ashby & Maddox 2005)—an older approach more in line with the approach one would expect from a discrete logic. For helpful reviews of the territory, including approaches beyond my scope here, see also Komatsu 1992, Laurence & Margolis 1999, Murphy 2002, Machery 2009 and Barsalou 2010.
[11] There is no entry for 'similarity' in the Stanford Encyclopedia of Philosophy, for example.
[12] Lewis attempts an algorithm for similarity in Lewis 1979, though this is explicitly limited to the case of worlds within the theory.
[13]Or have more predicates or features in common than they lack, appropriately weighted. Attempts to analyze similarity set-theoretically appear in both the psychological literature (Attneave 1950, Shepard 1980, Tversky 1977, Tversky & Gati 1987) and in a recent philosophical application (Weisberg 2013). From the perspective of computational retrieval, however, a recent review concludes that "there are reasons to doubt that any metric similarity measure could appropriately capture the nature of human similarity judgments" (ten Brinke, Squire & Bigelow 2004). From the perspective outlined here, any account which analyzes similarity in terms of features, predicates, or properties in common is fundamentally misguided. Features, properties, or applicable predicates represent categories that themselves depend on prior similarity judgments. Without basic similarity judgments one could not even form the feature categories any such account would demand.
[14] The general concept of a relation between two things, it should be noted, like the specific relation of similarity, remains basic: relations are not to be analyzed, for example, in terms of sets of ordered pairs. Appeal to standard set theory would immediately saddle us with all the problems of a discrete semantics. It may also be that in the end it is a comparative notion of similarity that lies deepest: not 'x is similar to y' but 'x is more similar to y than z is.' Such a move might unify the two models offered here. The fourth following paragraph is relevant to this point.