Imprecise Probability and Higher Order Vagueness

Susanna Rinard
Harvard University
July 10, 2014

Abstract

There is a trade-off between specificity and accuracy in existing models of belief. Descriptions of agents in the tripartite model, which recognizes only three doxastic attitudes—belief, disbelief, and suspension of judgment—are typically *accurate*, but not sufficiently *specific*. The orthodox Bayesian model, which requires real-valued credences, is perfectly *specific*, but often *inaccurate*: we often lack precise credences. I argue, first, that a popular attempt to fix the Bayesian model by using sets of functions is also inaccurate, since it requires us to have interval-valued credences with *perfectly precise* endpoints. We can see this problem as analogous to the problem of higher order vagueness. Ultimately, I argue, the only way to avoid these problems is to endorse *Insurmountable Unclassifiability*. This principle has some surprising and radical consequences. For example, it entails that the trade-off between accuracy and specificity is in-principle unavoidable: sometimes it is simply impossible to characterize an agent's doxastic state in a way that is both *fully accurate* and *maximally specific*. What we *can* do, however, is improve on both the tripartite and existing Bayesian models. I construct a new model of belief—the *minimal model*—that allows us to characterize agents with much greater specificity than the tripartite model, and yet which remains, unlike existing Bayesian models, perfectly accurate.

0.   Introduction

Much traditional epistemology employs a tripartite model of belief, which recognizes three doxastic attitudes: belief, disbelief and suspension of judgment. However, this model is too coarse-grained. Its descriptions of agents are typically *accurate*, but they are not sufficiently *specific*. For example, one may believe both P1 and P2, and yet be more confident of one than the other. The tripartite model is blind to these differences in confidence, and yet they are crucially important (for example, in the explanation of action).

Observations like these are often taken to motivate the orthodox Bayesian model, which recognizes uncountably many different doxastic attitudes: one for each real number between 0 and 1. However, this model suffers from the opposite problem. It is too fine-grained. Descriptions of agents in this model are perfectly *specific*—reflecting even the subtlest differences in confidence—but this comes at the cost of *accuracy*. Sometimes we lack precise, point-valued credences.

The tripartite model, then, has accuracy without specificity; the orthodox Bayesian model has specificity without accuracy. One aim of this paper is to present powerful reasons for the claim that this trade-off between accuracy and specificity in the

representation of belief is *in principle unavoidable*: it is simply not possible to represent belief in a way that is both *fully accurate* and *maximally specific*.

What we *can* do, however, is improve on both the tripartite and orthodox Bayesian models. Another aim of the paper is to construct and promote a new model of belief—the *minimal model*—which allows for much greater specificity than the tripartite model, and yet which remains, unlike the orthodox Bayesian model, perfectly accurate.

I begin in section 1 by discussing a popular attempt to improve on the orthodox Bayesian model by using a *set* of functions to represent an agent, rather than a *single* function. It will turn out that this set-of-functions model suffers from a slightly different version of the same sort of inaccuracy; specifically, it requires interval-valued credences with *perfectly precise* endpoints. An analogy between credal imprecision and vagueness will reveal that the set-of-functions model is analogous to supervaluationism, and the problem of interval endpoint precision analogous to the problem of higher-order vagueness. We'll see that other approaches to vagueness inspire analogous models of belief, but many suffer from analogs of some version of the higher-order vagueness problem.

In section 2 I argue that the only way to avoid these problems is by endorsing a surprising and radical principle I call Insurmountable Unclassifiability. I then ask what can be said about the representation of belief if this principle is true. As a preliminary to answering, in section 3 I present a novel way of using sets of functions, and intervals, to characterize doxastic states. This *minimal model* allows us to give multiple accurate characterizations of the same agent at different levels of specificity.

The minimal model has two primary virtues. First, it constitutes an improvement over the tripartite model and both existing Bayesian models, in that it allows for much greater specificity than the tripartite model, and yet, unlike existing Bayesian models, remains fully accurate. Second, it is neutral on the issues that divide defenders of different views on credal imprecision and higher order vagueness. Thus, it provides a framework within which we can make progress on a wide range of questions in epistemology and decision theory without having to first take a stand on these difficult and controversial issues.

In section 4 I make use of the minimal model in showing that if Insurmountable Unclassifiability is true, then the trade-off between accuracy and specificity in the representation of belief is *in principle unavoidable*: it is impossible to characterize an agent's doxastic state in a way that is both fully accurate and maximally specific.

Insurmountable Unclassifiability is a radical and mysterious view, however, as becomes apparent in section 5, where I discuss the forced march scenario. In section 6 I display some advantages of the view by showing how it can dissolve two challenging problems: puzzle cases for the Principle of Indifference, and diachronic decision problems for agents who lack precise credences.

Ultimately I do not come down one way or the other on whether we should accept this fascinating view. My aim is only to argue that there is a powerful reason in favor of it—namely, that it is necessary for avoiding the analogous problems of inaccuracy due to overprecision in the representation of belief, and higher order vagueness—and to begin exploring some of its consequences.

1.   The Vagueness Analogy

In this section, I begin by spotlighting the problem faced by the orthodox Bayesian model. Then I discuss a popular attempt to fix this problem by using a *set* of functions, rather than a *single* function. I point out that this set-of-functions model faces a problem very similar to the one that undermined the single-function model. Pursuing an analogy between credal imprecision and vagueness helps us gain deeper insight into the nature of these problems. In particular, we can see the set-of-functions model as analogous to the supervaluationist approach to vagueness, and the problem facing this model as analogous to a well-known problem for supervaluationism, sometimes called the *problem of higher-order vagueness*. I then briefly review two other approaches to vagueness: an infinite hierarchy of borderline cases, and many-valued logics. In each case, we see that it would be possible to construct an analogous model of belief, but that the resulting model would suffer from a version of the higher-order vagueness problem.

The orthodox Bayesian model represents the doxastic state of an agent by a function which assigns to each proposition some real number between 0 and 1. However, our confidence levels are not always that precise. For example, consider LUCKY, the proposition that you will find a four-leaf clover tomorrow. I could inquire about your level of confidence in LUCKY, and demand that you choose exactly one real number between 0 and 1, precise down to the millionth decimal place (and beyond)—but there would be an element of arbitrariness in any choice you might make.[1,2]

One popular attempt to fix this problem involves using a *set* of functions to represent each agent, rather than a *single* function. (Proponents include Jeffrey (1983), van Fraassen (1990), and Joyce (2005, 2010).) This set generates interval-valued credences; one's credence in P is the interval which contains all and only the real numbers r such that some function in one's set has $\Pr(P) = r$.

However, this model faces a problem much like the one that undermined the single-function model. The problem of arbitrariness resurfaces, just in a slightly different form. What, exactly, is the upper endpoint of your interval-valued credence for LUCKY? .0001? .00009? .000121? Again, any particular number seems arbitrary.[3]

We can gain a deeper understanding of these problems of arbitrariness by pursuing an analogy between credal imprecision and vagueness. To begin, I'll define a predicate MC. Roughly, MC applies to numbers that represent credences greater than your level of confidence in LUCKY. For a more precise definition, first, let B[r] say that

---

[1] Some insist that despite appearances to the contrary, every agent does in fact have a precise credence in every proposition. Arguing against this view is beyond the scope of this paper. Those who are sympathetic to it may interpret the paper as aiming to defend the conditional claim that the conclusions pursued here will follow, *if* in fact we sometimes lack precise credences.

[2] One might object to my presupposition that this inaccuracy in the orthodox Bayesian model is problematic. After all, simplifying idealizations are commonly used with great success in models in science. My reply is simple: whether an idealization is problematic depends entirely on the purpose to which the model is put. For example, if the goal is prediction, an idealization might be unproblematic (and even beneficial). In this paper, however, I evaluate models according to their ability to satisfy a simple curiosity about the actual facts—the truth, the whole truth, and nothing but the truth—about the nature of our doxastic states. Relative to this goal, any inaccuracy in the way a model represents our doxastic states is automatically a shortcoming.

[3] The problem of interval endpoint arbitrariness is also raised (though not under that name) in Sturgeon (2008, 158) and Maher (2006), and discussed in Kaplan (2010).

a coin with bias *r* towards heads will land heads on the next toss. For example, B[.7] says that a coin with bias .7 towards heads will land heads on the next toss. Assume that, in accordance with rationality, for all *r*, your credence in B[r] is *r*. Then we can define MC (for *more confident*) as the predicate that applies to a number *r* just in case you're more confident of B[r] than LUCKY.

MC is analogous to paradigm vague predicates like TALL. In each case there is a spectrum such that the predicate clearly applies at one end, and clearly fails to apply at the other, but there does not seem to be a sharp boundary. Both seem to admit of borderline cases, and give rise to sorites-style paradoxes.

Importantly, there is a close connection between our lack of a precise credence in LUCKY and the vagueness of MC.[4] If we had some precise credence *c* in LUCKY, there would be a sharp boundary between MC and not-MC: all numbers greater than *c* would have MC, and all numbers equal to or less than *c* would have not-MC. A natural thought, then, is that credal imprecision is of a piece with the vagueness of predicates like MC.

The set-of-functions model is an attempt to account for our lack of precise credences, and, thereby, an attempt to account for the vagueness of predicates like MC. We can interpret this model along supervaluationist lines: each function in your set is one admissible precisification of your doxastic state. Functions excluded from your set are *in*admissible precisifications. A proposition about your doxastic state is *determinately* true if true according to all functions in your set. It's *indeterminate* if true according to some, but not all, functions in your set.

For example, if all functions in your set have Pr(B[.9]) > Pr(LUCKY), it's determinate that you're more confident of B[.9] than LUCKY, i.e., it's determinate that MC[.9] is true. If some functions in your set have Pr(B[.0001]) > Pr(LUCKY) while others do not, it's indeterminate whether you're more confident of B[.0001]) than LUCKY, i.e., it's indeterminate whether MC[.0001] is true. So, on the supervaluationist interpretation, we can see the set-of-functions model as an attempt to account for the vagueness of MC by introducing a third category, *indeterminately* MC, just as the supervaluationist tries to account for the vagueness of TALL by introducing a third category, *indeterminately* TALL.[5]

Supervaluationism faces a well-known problem: it requires a sharp boundary where, intuitively, there shouldn't be one, namely, between the *determinate* and *indeterminate*. This is the so-called problem of higher-order vagueness. On the supervaluationist interpretation of the set-of-functions model, the problematic requirement of precise endpoints for interval-valued credences is just an instance of this more general problem. A sharp upper endpoint for one's interval-valued credence in LUCKY would constitute a sharp boundary between the numbers to which MC *determinately* applies and those to which it *indeterminately* applies. But there is no sharp boundary here.[6]

---

[4] For simplicity I will speak as if MC is vague, but strictly speaking I rely only on the claim that MC is structurally analogous to paradigm vague predicates.

[5] Supervaluationist approaches to vagueness are defended in Fine (1975) and Keefe (2000), among others.

[6] Unfortunately, the question of how to interpret the set-of-functions model has received very little attention in the literature, so although some explicitly endorse the supervaluationist interpretation (including van Fraassen (1990, 2005, 2006) and Hajek (2003)), it is unclear how widespread this interpretation is, and what alternatives might look like. One clearly incompatible interpretation appears in Schoenfield (2012) and Kaplan (2010): if some functions in your set have Pr(A) > Pr(B), while others have Pr(A) < Pr(B), and

As the literature has shown, it is extremely difficult to do justice to higher order vagueness. Below I briefly review a couple of prominent attempts and why they are unsatisfactory. Because of the analogy between credal imprecision and vagueness, in each case there is an analogous model of belief that would face analogous problems.

First, some postulate an infinite hierarchy of borderline cases, borderline borderline cases, etc. For example, indeterminately tall heights are first-order borderline cases, but there are also second-order borderline cases: heights that are borderlines of the distinction between determinately tall and indeterminately tall. For every natural number n, there are nth-order borderline cases. This strategy could be employed in an attempt to fix the problem of precise endpoints for interval-valued credences: we would account for the lack of a sharp line between determinately MC and indeterminately MC by postulating borderline cases of that distinction, and so on up the hierarchy.

However, there is a compelling objection to this approach. Consider paradigm tall heights that aren't borderline cases at *any* level. Call these *absolutely* tall. If every height must be classified as either absolutely tall, or absolutely not tall, or, for some natural number n, a borderline case of nth order, then there will be a sharp cut-off between those classified as absolutely tall and those classified some other way. But there should not be a sharp cut-off here.[7] Similarly, if every real number between 0 and 1 must be classified as either absolutely MC, or absolutely not MC, or, for some natural number n, a borderline case of nth order, then there will be a sharp cut-off between the numbers that are absolutely MC and those that are not. But there is no sharp boundary here.

A different class of views on vagueness involves multiple *degrees* of truth.[8] Some postulate uncountably many: one for each number between 0 and 1, inclusive. This view also faces the problem of higher order vagueness. If we assign some precise degree of truth to every proposition of the form TALL(r), there will be a sharp cut-off between the numbers such that TALL(r) is true to degree 1, and the numbers such that TALL(r) is true to some degree less than 1. But there should not be a sharp cut-off here.

The analogous model of belief, on which each proposition of the form MC(r) is assigned some precise degree of truth, suffers from the same problem: it requires a sharp cut-off between those numbers for which MC(r) is true to degree 1, and those for which MC(r) is true to some degree less than 1. But again, there is no sharp boundary here.

2. Insurmountable Unclassifiability

The orthodox Bayesian model of belief, as we have seen, is inaccurate. The nature of the inaccuracy is that it requires point-valued credences, which, as we saw in the previous section, amounts to the requirement that there be sharp lines where, intuitively, there aren't any—such as between MC and not-MC. (MC(r) says, recall, that

still others Pr(A) = Pr(B), then it's false that you're more confident of A than B; false that you're less confident of A than B; and false that you're equally confident of A than B. Your attitude towards A and B falls into a fourth category. (This is analogous to a view in value theory defended, among others, by Ruth Chang (see, for example, Chang (2002)), on which it can be the case that A is neither more valuable than B, nor less valuable, nor equally valuable.) This interpretation requires, implausibly, a sharp boundary between the numbers that have MC, and those that fall into the fourth category.

[7] Similar problems are presented in Sainsbury (1991) and Wright (2009).
[8] See, for example, Smith (2008) and Zadeh (1975).

you are more confident of B[r]—that a coin with bias *r* towards heads will land heads on the next toss—than LUCKY (that you will find a four-leaf clover tomorrow).)

So, we cannot use the two categories *MC* and *not-MC* to give an exhaustive classification of all numbers between 0 and 1. In the previous section we surveyed a number of different attempts to reach a categorization that *is* genuinely exhaustive by *refining* this initial two-way distinction. Specifically, each attempt involved adding more (in some cases, infinitely more) intermediate categories. But all these attempts foundered on the same sort of problem: each ended up requiring sharp lines where, intuitively, there shouldn't be any.

In fact, we can identify an even more substantive commonality among the sources of failure of these different attempts. In each case, the proposed model failed because it ended up requiring a sharp line between those numbers for which MC(r) is *as true as possible*, and those for which it is not. For example, consider the supervaluationist framework. Here, to be determinately true is to be as true as possible, and to be indeterminately true is to fail to be as true as possible. So the sharp line, required by this framework, between the determinate and the indeterminate amounts to a sharp line between what's as true as possible and what's not. Postulating an infinite hierarchy of borderline cases doesn't help; we still have a sharp line between what's as true as possible (in this case, when MC(r) is *absolutely* true, not borderline at *any* level) and what's not (when MC(r) is either a borderline case at *some* level, or absolutely false). Many-valued logics require a sharp line between those numbers for which MC(r) is true to degree 1 (as true as possible), and those true to some degree less than 1 (not as true as possible).

The error that unifies these models, then, is their joint commitment to a sharp line between what's true as possible and what's not. That there is no sharp line here imposes a severe restriction on our ability to classify the numbers between 0 and 1 according to their MC status. This restriction can be stated more precisely, as follows:

Insurmountable Unclassifiability: For any set of categories C with properties (1) and (2) (definitions to follow), it is not the case that every number between 0 and 1 can be classified into some category in C. (1) For some proper subset of C, any number r that falls into some category in that subset is, in virtue of being in that category, such that MC(r) is as true as possible. (2) Any number r that falls into some category in C in the complement of that proper subset is, in virtue of being in that category, such that it's not the case that MC(r) is as true as possible.

Insurmountable Unclassifiability is intended to capture the idea that, not only can we not classify all numbers along the two-way MC/not-MC distinction, but we also cannot classify all numbers according to any set of categories that is intended to *supersede* this two-way distinction, or *refine* it via the addition of (even infinitely many) more intermediate categories.[9] What is wrong the initial two-way distinction is not that there are too few categories into which to sort numbers. Rather, the fundamental problem is the exhaustiveness assumption. The mistake is to think that *every* number can be

---

[9] Sainsbury (1991) sketches a view that has much in common with Insurmountable Unclassifiability.

assigned to some category that completely captures its status with respect to MC. Insurmountable Unclassifiability denies this.[10]

Now, it is clear that *some* numbers are classifiable in categories that are members of a set with properties (1) and (2). What Insurmountable Unclassifiability denies is just that *all* numbers are so classifiable. That some, but not all, numbers are so classifiable might lead one to think that we can draw a sharp line between those that are, and those that aren't. But, as we will see, this, too, is not possible. That it's not begins to illustrate the truly *insurmountable* nature of the unclassifiability.

Suppose we were able to say, for each number, whether it was so classifiable or not. If so, then, for each classifiable number *r*, we can say whether or not MC(r) is as true as possible. (If we couldn't, then *r* wouldn't be in the classifiable category.) So, there would be a sharp two-way distinction between, on the one hand, the numbers that are classifiable, and such that MC(r) is as true as possible; and, on the other, those that are either unclassifiable, or, classifiable, but not as such that MC(r) is as true as possible. But that would constitute a sharp two-way classification between the numbers for which MC(r) is as true as possible, and those for which it is not! Each number in the first category—classifiable, and such that MC(r) is true as possible—is clearly such that MC(r) is true as possible. Of the numbers in the second category, those that are classifiable, but not such that MC(r) is true as possible, are, obviously, not such that MC(r) is true as possible. And—crucially—those in the second category that are not classifiable are clearly such that it's not the case that MC(r) is true as possible. After all, if MC(r) *were* as true as possible, then that number would fall into the first category. So Insurmountable Unclassifiability entails that, although some numbers are classifiable, and it's not the case that all numbers are classifiable, we cannot categorize each number as either classifiable or not.

As will become even more apparent later in the paper, Insurmountable Unclassifiability is quite a radical and mysterious view. In this section I have argued, though, that accepting it is absolutely necessary if we are determined to avoid problems of the sort that plagued the orthodox Bayesian model and the myriad attempts, discussed in the previous section, to improve upon it. That is, accepting Insurmountable Unclassifiability is absolutely necessary to avoid commitment to the existence of sharp lines where, intuitively, there shouldn't be any.

3.   The Minimal Model

I began this paper by noting a trade-off between accuracy and specificity in two popular models of belief. In the section following this one I will show that if Insurmountable Unclassifiability is true, then this trade-off is in principle unavoidable: it

---

[10] Those with supervaluationist inclinations might think the solution is to posit a third category, consisting of borderline cases of the distinction between what's as true as possible and what's not. But there can be no borderlines of this distinction. Suppose there were. Then there would be a distinction between the determinately true as possible, and the borderline true as possible. Consider an instance of the second category. Since it is determinately in the second category, rather than the first, it determinedly fails to be as true as it possibly could be. (It could be in the first category.) So what we would have is not *borderline* as true as possible, but rather, determinately *not* as true as possible. It is part of the nature of the distinction between what's as true *as possible*, and what's not, that there can be no borderlines of it. (Compare the argument in Broome (1997) for his Collapsing Principle and an argument in Barnes (1982), p. 55.)

is impossible to characterize belief in a way that is both fully accurate and maximally specific. In order to do this, I will first construct a new way of using interval notation and sets of functions to represent belief, which I call the *minimal model*. That is the aim of this section. As we will soon see, it turns out that this minimal model constitutes an improvement over both the tripartite model and existing Bayesian models: it allows for much greater specificity than the tripartite model, and yet remains, unlike existing Bayesian models, perfectly accurate.

First, notice that among *informal* characterizations of belief, some are *more specific* than others. For example, suppose I say you're *more confident of P than not*, and then elaborate that you're *nearly certain of P*. The second description is more specific than the first, but both are perfectly accurate. The new *minimal* interpretation of sets and intervals presented here allows us to give, in a similar fashion, multiple accurate characterizations, at different levels of specificity, of a single agent's doxastic state.

I'll start with intervals. First I'll give a definition that is helpful, but potentially misleading; then I'll do it more carefully. Helpful, but misleading: on the minimal interpretation, [c, d] accurately characterizes an agent's doxastic state toward H just in case *the agent's level of confidence in H is contained within [c, d]*. Note that on this interpretation there will always be multiple accurate intervals at varying levels of specificity. On reason is that, if [c, d] is accurate, then any larger (more inclusive) interval is also accurate. This is because, if one's level of confidence is contained within [c, d], then it is contained within any interval of which [c, d] is a subset. For example, since I am fairly confident that LUCKY is false, my doxastic attitude toward LUCKY is accurately characterized by [0, 1], [0, .8], [0, .5], and many others. On the minimal interpretation, each of these is perfectly accurate; some are more specific than others.

The definition given above is misleading insofar as it presupposes that the agent has a precise, point-valued level of confidence. The fix is: on the minimal interpretation, an interval [c, d] accurately characterizes an agent's doxastic state toward H just in case (1) the agent is *more* confident of H than any proposition in which her credence is *less* than *c*; and (2) the agent is *less* confident of H than any proposition in which her credence is *greater* than *d*. One final clarification: the conditions on the right-hand side must be *as true as possible* for the interval to count as accurate.

One noteworthy feature of this new use of intervals is that it is entirely compatible with the substantive views held by proponents of different views on credal imprecision, including the supervaluationist interpretation, an infinite hierarchy of borderline cases, many-valued logics, Insurmountable Unclassifiability, even the single-function model, etc. For example, if the supervaluationist has [c, d] as the agent's unique interval-valued credence in H, then that interval (as well as any more inclusive interval) automatically counts as accurate on the minimal interpretation. This is because the supervaluationist interpretation of the interval notation is logically stronger than the minimal interpretation of the same notation (hence the name *minimal*). For another example, if one has a point-valued credence *c* in H, then every interval of which *c* is a member will count as accurate on the minimal interpretation. What this makes salient is that *to characterize an agent's doxastic attitude using the new, minimal interval notation is to remain neutral on the issues that divide defenders of all different views on credal (im)precision.* This is because, on the minimal interpretation, to characterize an agent's attitude towards H with some interval is to remain completely neutral about the status of numbers *inside* that

interval. For any such number *r*, it may be that one is *more* confident of H than B[r]; or *less* confident; or *equally* confident; it may be indeterminate whether *r* is one's level of confidence in H; etc. For any possible status *r* might have, we remain completely neutral about whether *r* has that status. To describe one's attitude with an interval is to be committal only about numbers outside that interval.

We can give a new, minimal re-interpretation of sets of functions in the same vein. First pass: a set counts as accurate just in case every proposition about the agent's doxastic state that is true according to *all* functions in that set is true (as possible) of the agent. This is only a first-pass definition, though, because some propositions true according to all functions in the set are *not* true of the agent. For example, it is true according to every function in the set that you have a precise credence in LUCKY. (The functions agree that you have a precise credence; they just disagree about what it is.) But *that you have a precise credence in LUCKY* is precisely what we want to deny! We can get to the root of this problem by noticing that, although the existential claim *there is some real number r such that r is your precise credence in LUCKY* is true according to every function, there is no instance of that existential claim that is true according to every function; that is, there is no *r* such that *r is your precise credence in LUCKY* is true according to every function in the set. So, we can fix the problem by revising the definition as follows: First, let Z be the set of all propositions true according to all functions in the set. Generate Z- by removing from Z any proposition that is, or is equivalent to, some existential claim such that no instance of that existential claim is true according to every function in the set. Now the proper definition: on the minimal interpretation, a set counts as accurate just in case every proposition in Z- is true (as possible) of the agent.

To characterize an agent with a set, on this new, minimal interpretation, is to remain entirely neutral on the status of propositions about which different functions in the set disagree, just as the minimal interval notation remains neutral on the status of numbers inside the interval. There will always be multiple accurate sets, just as there are always multiple accurate intervals.   In addition, as before, this interpretation is compatible with different views on credal imprecision. For example, if S is the agent's unique set of functions, interpreted in the supervaluationist way, then S (and any more inclusive set) counts as an accurate description of that agent on the minimal interpretation. If an agent is best represented by a single credence function, any set containing that function is accurate.

This minimal model can easily accommodate traditional Bayesian approaches to a wide range of different issues, such as learning from experience, theory confirmation in science, decision theory, etc. It will typically be the case that if, on the single-function model, some proposition B is true of an agent if that agent's credence function has property Q, then, on the minimal model, B will be true of the agent if there is some set of functions, which accurately characterizes that agent, all of whose members have property Q. For example, on the minimal model we can say that E confirms H if there is some accurate set of functions, all of whose members have Pr(H|E) > Pr(H). The agent satisfies minimal synchronic requirements for rationality if there is some accurate set of functions, all of whose members conform to the axioms of probability. The agent rationally updates on evidence E, received between t1 and t2, just any case any set of

functions S2 that is accurate at t2 can be obtained from some set S1 that was accurate at t1 via conditionalizing each function in S1 on E. And so forth.

What I want to emphasize here, though, is that when it comes to the representation of belief, the minimal model is an improvement over both the tripartite model and existing Bayesian models. It allows for characterizations much more specific than those of the tripartite model. For example, if [.95, .96] accurately characterizes my attitude toward A, and [.97, .98] my attitude toward B, the model represents that I am more confident of B than A, even if I believe both. Yet it remains (unlike existing Bayesian models) perfectly accurate. It does not fall prey to the problem of arbitrariness, since accurate intervals are not taken to be *uniquely* accurate. It is perfectly accurate to characterize my attitude toward LUCKY with [0, .8], even though the endpoints are perfectly precise, because this characterization is not taken to be a unique best one. Other intervals, such as [0, .7] or [0, .6], may be equally accurate.

The only shortcoming of this model is that it typically provides only partial descriptions of the agent's doxastic state. To characterize one's doxastic state with [0, .7], for example, is not maximally informative, since narrower intervals may be equally accurate. Whether this shortcoming can be remedied is addressed in the next section.

4.   Insurmountable Unclassifiability Renders Impossible the Combination of Perfect Accuracy and Maximal Specificity in the Representation of Belief

In the previous section I constructed the *minimal model*, which allows us to give multiple accurate descriptions of an agent's doxastic state at different levels of specificity. It is fine to characterize one's doxastic state by giving a few descriptions of this kind, but doing so raises a natural question. What is the *most specific* accurate interval? For example, above I listed the following as accurate characterizations of my attitude toward LUCKY: [0, 1], [0, .8], and [0, .5]. Having done this, it is natural to ask about other intervals. What about [0, .3]? (0, .27)? [0, .2]? In particular, it is natural to wonder which interval is the *most specific* interval that is still accurate. After all, any information encoded in a less specific interval is also contained in a more specific interval—so why not just isolate the maximally specific accurate interval, identify it as such, and forget about the rest?

Interestingly, it turns out that if we are serious about avoiding counterintuitive sharp lines—i.e, if we accept Insurmountable Unclassifiability—then we must regard this very natural thought as deeply flawed. Suppose there were a most specific interval that accurately characterized my attitude towards LUCKY—say, [c, d]. Then any narrower interval would fail to be accurate. But then *d* would constitute a precise boundary where there shouldn't be one. For all numbers greater than *d*, MC(r) would be *as true as possible*; for all numbers equal to or less than *d*, MC(r) would fail to be as true as possible. But, as we have seen, Insurmountable Unclassifiability entails that we can't classify all numbers *r* according to this scheme. The upshot: if we accept Insurmountable Unclassifiability—the only way to avoid problems of arbitrariness and higher-order vagueness—then sometimes *there is no maximally specific, fully accurate characterization of one's doxastic state.*

This is a surprising and radical conclusion. It means that it is impossible to give a *complete* description of an agent's doxastic state; there is *no such thing* as a complete

description. We can give partial descriptions of an agent's doxastic state; some will be more specific than others. But we can never identify a particular description as *maximally specific*, or *complete*.[11]

5. Quietism and the Forced March

In order to draw out some further implications of Insurmountable Unclassifiability, consider a *forced march* scenario. Consider the following finite series of intervals. All intervals in the series share the same lower endpoint: 0. The upper endpoint of the first element of the series is 1. For each interval in the series, its upper endpoint is the result of subtracting some minuscule positive real number ε from the upper endpoint of its predecessor. The last element of the series is the first interval whose upper endpoint is equal to or less than 0.

Imagine going through the elements of this series one by one, and asking someone, for each interval, whether it is accurate concerning their level of confidence in LUCKY. At first, the obvious answer is *yes*. But at some point, the response must be something else. (Otherwise the person will answer *yes* to the last element of the series, which is clearly incorrect.) But what can they say?

It follows from Insurmountable Unclassifiability that in order to avoid error, the speaker is at some point required to perform a speech act that is non-committal in the following sense: the content of the speech act is compatible with its being the case that *yes* would have been a correct answer; and also compatible with its being the case that *yes* would have been an *incorrect* answer. Which speech act may have these features is an empirical question. Silence is a natural candidate—but only if it's understood *not* to implicate that the answer is not *yes*. The person might say "I'm bored of this. Let's go swimming!" or "Was that a Pileated Woodpecker that just flew by?"

This suggestion—that, at some point, one must switch to a non-committal speech act—may strike some as unsatisfying. Those who find it unsatisfying may try to stipulate that the subject will say *yes* if and only if *yes* would be a correct answer. Then speech acts like "Let's go swimming!" won't allow them to wriggle out of the question.

However, a defender of Insurmountable Unclassifiability should deny that such a stipulation can be made. On the assumption that, for each question, there is a fact of the matter about whether not the subject responded with *yes*, then the claim that they will say *yes* if and only if *yes* is a correct answer is equivalent to the claim that, for each interval,

---

[11] How might this affect other issues, such as updating, decision theory, etc.? In a sense, not at all—these are still handled in the minimal model exactly as described in the previous section. But Insurmountable Unclassifiability introduces a new wrinkle: there is now no *guarantee* that there will be a fact of the matter about whether the relevant conditions obtain (although there *may* always be a fact of the matter). For example, on the minimal model we have that *if* there is some accurate set of functions, all of whose members have $\Pr(H|E) > \Pr(H)$, then E confirms H. And if there is *no* accurate set of functions, all of whose members have $\Pr(H|E) > \Pr(H)$, then it's not the case that E confirms H. But since, according to Insurmountable Unclassifiability, it's not the case that every set is classifiable as either accurate or not, the door is now open to the possibility that it will not be settled whether E confirms H. It is important to emphasize, however, that Insurmountable Unclassifiability does not *require* that there sometimes fail to be a fact of the matter here. It is compatible with Insurmountable Unclassifiability that there is always a fact of the matter about whether E confirms H. The same goes for other issues, such as rational updating, minimal constraints on synchronic rationality, etc.

there's a fact of the matter about whether it's accurate or not. And the denial of this is a core commitment of Insurmountable Unclassifiability.

What we are left with, then, is a form of Quietism. For any subject of the forced march, it must be the case that there is some pair of adjacent intervals such that they answer *yes* to the question about one but give a different answer to the question about the other. That answer must be non-committal in the sense described above.

This section has only begun to explore the implications of Insurmountable Unclassifiability; many further questions remain. Already, though, we can see that it leads to some surprising, radical, and mysterious conclusions. Those who are serious about avoiding the problems of arbitrariness, higher-order vagueness, and counterintuitive sharp lines are in for some exciting times!

6. Dissolving Problems with Insurmountable Unclassifiability

My goal in this section is to highlight some of the virtues of Insurmountable Unclassifiability by showing how it dissolves two challenging problems: (1) puzzle cases for the Principle of Indifference; (2) diachronic decision problems for agents who lack precise credences.

The much-discussed *Principle of Indifference* (POI) says that if one has no more reason to believe A than B, and no more reason to believe B than A, then one's credence in A should equal one's credence in B.[12] In the orthodox Bayesian model, where credences are point-valued, this famously seems to lead to contradiction in some cases. For example, suppose a factory produces cubes of equal size.[13] You know only that the length (L) of a side of a cube is 2 feet or less. Plausibly, you have no more reason to believe $0 < L \leq 1$ than $1 < L \leq 2$, and vice versa; so, according to POI, you must have the same point-valued credence in each of these propositions, namely, ½. But now notice that your initial information is equivalent to the information that the area (A) of a side is 4 feet or less. Now it seems you have no more reason to believe any one of these four possibilities than any of the others: $0 < A \leq 1$; $1 < A \leq 2$; $2 < A \leq 3$; $3 < A \leq 4$. So, according to POI, you must have the same point-valued credence in each, namely, ¼. But this contradicts the recommendation of the first application of the principle, since $0 < L \leq 1$ is equivalent to $0 < A \leq 1$.

Some (including Joyce (2005) and Weatherson (2007)) have claimed that moving to the set-of-functions model—with the attendant move to interval-valued credences—solves the problem: that we can respect both verdicts of the POI by giving the same (unique, maximally specific) interval-valued credence to each member of the two-celled partition, and the same (unique, maximally specific) interval-valued credence to each member of the four-valued partition. However, as I have shown elsewhere [reference removed for blind review], this is coherent only if one's interval-valued credence in all six is [0, 1]. And, as I have argued elsewhere, (maximally specific) interval-valued credences with such extreme endpoints are not rationally permissible [reference removed for blind review].

This entire set of problems evaporates if Insurmountable Unclassifiability is adopted—specifically, if we give up the presupposition that we must have a maximally

---

specific, fully accurate interval-valued or point-valued credence in each proposition. First, we can use the minimal model to re-state the POI in a way compatible with Insurmountable Unclassifiability, as follows: if one has no more reason to believe A than B, and no more reason to believe B than A, then, if one is rational, then if some interval R is an accurate characterization of one's attitude toward A, then that same interval R is also an accurate characterization of one's attitude toward B, and vice versa. This principle does not lead to contradiction, even if we agree that you have no more reason to believe 0 < L ≤ 1 than 1 < L ≤ 2, *and* no more reason to believe any of the following than any other: 0 < A ≤ 1; 1 < A ≤ 2; 2 < A ≤ 3; 3 < A ≤ 4. For example, each of the following intervals can coherently accurately characterize your attitude toward each of these six propositions: [0, .75]; [.12, .8]; [.011, .673]; etc. With Insurmountable Unclassifiability in place, there is no longer any requirement to find a *maximally specific* accurate interval for each proposition—and it is *this* requirement that we should give up in light of the contradiction, not the innocuous principle that when you have no more reason to believe one thing than another, you shouldn't take different doxastic attitudes towards them.

Giving up this requirement also pulls the rug out from under an argument in Elga (2010) for the claim that rationality requires us to have a precise, point-valued credence in every proposition. Elga draws this conclusion after considering a number of different possible decision rules one might use if one had a (unique, maximally specific) interval-valued credence in some proposition. In each case, he argues that the rule is unacceptable. He infers that it is not rational to have a (unique, maximally specific) interval-valued credence in any proposition; and concludes from this that all rational credences are point-valued.

Once Insurmountable Unclassifiability is on the table, it's clear that the final step in Elga's argument is invalid. Elga does not consider the possibility that rationality might allow one to lack a point-valued credence and *also* lack a unique, maximally specific interval-valued credence.[14] I will argue that the decision theory that accompanies Insurmountable Unclassifiability gives the right results about the case on which, Elga claims, the view that we have a (maximally specific) interval-valued credence founders.

Elga supposes that the agent has a maximally specific credence in H of [.1, .8]. He then supposes the agent is offered some bet A, immediately followed by another bet B. If the agent accepts bet A, she will gain $15 if H is true but lose $10 if H is false. If she accepts bet B, she will lose $10 if H is true, but gain $15 if H is false. The thing to notice is that if the agent accepts *both* bets, she is guaranteed to gain $5. Of course, rejecting both guarantees $0. So it would be irrational to reject both bets. (We assume the agent cares only about money, etc.) But, Elga claims, no decision theory for (maximally specific) interval-valued credences can accommodate this. For example, according to one popular rule that has been frequently endorsed in the literature, rejecting bet A is permissible, and rejecting bet B is permissible as well.

Now, suppose that, in accordance with Insurmountable Unclassifiability, the agent in this case lacks a maximally specific interval-valued credence, but that [.1, .8] is an accurate characterization (in the minimal model) of their doxastic attitude. What decision-theoretic commitments would follow? As we have seen, the minimal model remains completely silent on matters about which different functions in some accurate set

---

[14] It is understandable that he does not consider this view, as it has heretofore not been a salient position in the debate.

disagree. In this case, since [.1, .8] is an accurate interval, there will be some accurate set which contains, for each value v in [.1, .8], some function with $Pr(H) = v$. Some functions in that set have the expected value of rejecting bet A greater than accepting, but others have the expected value of accepting greater than rejecting. There is no agreement among the members of this set about whether rejecting is permissible or not. So the minimal model remains completely silent on that question. The same is true for bet B. This is good news—the minimal model does not have the implausible consequence that rejecting each bet is permissible.

However, we might wonder whether we can also do justice to the intuition that rejecting both bets would be impermissible. In fact, we can! Although the minimal model does not prohibit the individual action of rejecting bet A; and does not prohibit rejecting bet B; it *does* prohibit the compound action of rejecting both bets. This is because every function in that accurate set agrees that the expected value of rejecting both bets is $0, and that the expected value of accepting both bets is $5; and so all functions agree that rejecting both bets is rationally impermissible. The minimal model yields exactly the result that Elga claimed no decision theory for non-precise credences could accommodate.[15]

7.   Conclusion

I began by noting a trade-off between accuracy and specificity in two commonly-employed models of belief. The tripartite model, which recognizes only three doxastic attitudes, allows us to characterize belief in a way that is *perfectly accurate*, but *not sufficiently specific*. The model is too coarse-grained. The orthodox Bayesian model, on the other hand, is too fine-grained. It represents agents in a way that is highly specific, but typically inaccurate. It requires point-valued credences in every proposition, but usually our doxastic attitudes are not this precise (nor does rationality require them to be). A popular modification of the Bayesian model—the set-of-functions model—suffers from a similar sort of problem. It requires a unique interval-valued credence, with perfectly precise endpoints, for each proposition. But our doxastic attitudes are usually not this precise either (nor does rationality require them to be).

I have shown that if Insurmountable Unclassifiability is true, then this trade-off between accuracy and specificity is in-principle unavoidable: it is simply not possible to characterize belief in a way that is both fully accurate and maximally specific. I gave a powerful reason in favor of this principle: it is, I argued, non-negotiable for those serious about avoiding the problem of inaccuracy due to overprecision in the representation of belief, and the analogous problem of higher order vagueness. Moreover, it can dissolve two challenging problems: puzzle cases for the Principle of Indifference, and diachronic decision problems for agents who lack point-valued credences.

---

[15] One thing this discussion reveals is that Elga's argument against some other decision rules is too quick. For example, what he calls permissive rules have the same verdict as the minimal model—that the compound action of rejecting both bets is impermissible—for similar reasons. However, this is complicated by the fact that according to these rules, rejecting bet A is permissible, as is rejecting bet B. So they are committed to the implausible result that a certain compound action is impermissible even though each component action is permissible. The combination of Insurmountable Unclassifiability and the minimal model has no such implausible commitment.

However, it has not been my aim in this paper to reach a final verdict on Insurmountable Unclassifiability. It will doubtless be controversial; some of its consequences are surprising and radical. For example, it entails that the subject in a forced march scenario is required, at some point, to respond in a way that is maximally non-committal on the question asked.

Regardless of whether or not we accept this fascinating and mysterious view, we can improve on both the tripartite model and the existing Bayesian models of belief. One aim of the paper has been to describe and promote a new model of belief—the *minimal model*—that enables us to use intervals, and sets of functions, to characterize belief in a way that is much more specific than the tripartite model, and yet which does not fall prey to problems of overprecision and so is perfectly accurate. A further virtue of the minimal model is that it remains neutral on the issues that divide defenders of different approaches to the problem of higher order vagueness. Thus, it provides a framework within which we can make progress on a wide range of different issues in epistemology and decision theory without first having to take a stand on this challenging and controversial problem.

References:

Barnes, J. (1982) Medicine, Experience, and Logic. In J. Barnes, J. Brunschwig, M. F. Burnyeat, and M. Schofield (eds.) Science and Speculation. Cambridge: Cambridge University Press.

Broome, J. (1997) Is Incommensurability Vagueness? In R. Chang (ed.) Incommensurability, Incomparability, and Practical Reason. Cambridge: Harvard University Press.

Chang, R. (2002) The Possibility of Parity. Ethics 112: 659-88.

Fine, K. (1975) Vagueness, Truth and Logic. Synthese 30: 265-300.

Hajek, A. (2003) What Conditional Probability Could Not Be. Synthese 137(3): 273-323.

Huemer, M. (2009) Explanationist Aid for the Theory of Inductive Logic. British Journal for the Philosophy of Science 60(2): 345-375.

Jeffrey, R. C. (1983) Bayesianism with a Human Face. In J. Earman (ed.) Minnesota Studies in the Philosophy of Science, volume 10, Testing Scientific Theories. Minneapolis: University of Minnesota Press.

Joyce, J. (2005) How Probabilities Reflect Evidence. Philosophical Perspectives, 19(1): 153-178.

Joyce, J. (2010) A Defence of Imprecise Credences in Inference and Decision Making. Philosophical Perspectives, 24(1): 281-323.

Kaplan, M. (2010) In Defense of Modest Probabilism. Synthese, 176(1):41-55.

Keefe, R. (2000) Theories of Vagueness. Cambridge: Cambridge UP.

Maher, P. (2006) Book Review: David Christensen. *Putting Logic in its Place: Formal Constraints on Rational Belief*. Notre Dame Journal of Formal Logic 47(1): 133-149 (2006).

Sainsbury, M. (1991) Is There Higher-Order Vagueness? Philosophical Quarterly 41(163):167-182.

Schoenfield, M. (2013) Permission to Believe: Why Permissivism is True and What it Tells Us about Irrational Influences on Belief. Nous 47(1).

Schoenfield, M. (2012) Chilling Out on Epistemic Rationality. Philosophical Studies 158(2): 197-219.

Smith, N.J.J., 2008. Vagueness and Degrees of Truth. Oxford: Oxford University Press.

Sturgeon, S. (2008) Reason and the Grain of Belief. Noûs 42(1):139-165.

van Fraassen, B. 1989. *Laws and Symmetry*. Oxford: Clarendon Press.

van Fraassen, B. (1990) Figures in a Probability Landscape. In M. Dunn and A. Gupta (eds.), Truth or Consequences. Dordrecht: Kluwer.

van Fraassen, B. (2005) Conditionalizing on Violated Bell's Inequalities. Analysis, 65(1):27-32.

van Fraassen, B. (2006) Vague Expectation Value Loss. Philosophical Studies 127(3).

Weatherson, B. (2007) The Bayesian and the Dogmatist. Proceedings of the Aristotelian Society, 107: 169 – 85.

White, R. (2014) Evidence Cannot Be Permissive. In M. Steup, J. Turri, and E. Sosa (eds.) Contemporary Debates in Epistemology, Second Edition p. 312-324.

White, R. (2010) Evidential Symmetry and Mushy Credence. In T.S. Gendler and J. Hawthorne (eds.) Oxford Studies in Epistemology, volume 3. Oxford, UK: Oxford University Press.

Wright, C. (2009) The Illusion of Higher-Order Vagueness. In R. Dietz and S. Moruzzi (eds.), Cuts and Clouds. Vaguenesss, its Nature and its Logic. Oxford University Press.

Zadeh, L. (1975) Fuzzy Logic and Approximate Reasoning. Synthese, 30: 407-428.