

ORIGINAL ARTICLE

Against Radical Credal Imprecision

Susanna Rinard

University of Missouri – Kansas City

A number of Bayesians claim that, if one has no evidence relevant to a proposition P , then one's credence in P should be spread over the interval $[0, 1]$. Against this, I argue: first, that it is inconsistent with plausible claims about comparative levels of confidence; second, that it precludes inductive learning in certain cases. Two motivations for the view are considered and rejected. A discussion of alternatives leads to the conjecture that there is an in-principle limitation on formal representations of belief: they cannot be both fully accurate and maximally specific.

Keywords Bayesian; Bayesianism; conditionalization; credal imprecision; higher-order vagueness; imprecise credences; imprecise probability; mushy credence; supervaluationism; supervaluationist

DOI:10.1002/tht3.84

Introduction

Suppose you know of two urns, GREEN and MYSTERY. You are certain that GREEN contains only green marbles, but have no information relevant to the colors of the marbles in MYSTERY. A marble will be drawn at random from each. How confident should you be that the marble drawn from GREEN will be green? You should be certain, or very nearly so. How confident should you be that the marble drawn from MYSTERY will be green? You shouldn't be anywhere near certain. So, you should be more confident that the marble from GREEN will be green than that the marble from MYSTERY will be green.

I will use this observation as the basis of an argument against the *wide interval view*, which says (roughly) that, in a case of no evidence, rationality requires credence to be spread over $[0, 1]$. This view is endorsed by a number of Bayesians, including Joyce (2005), Weatherson (2007), Kaplan (1996), and Walley (1991). A second argument against the view will be that in certain cases, it renders inductive learning impossible. Two motivations for the view will be discussed and found wanting. Consideration of alternatives will lead to the conjecture that there are significant in-principle limitations on any attempt at the formal representation of belief.

The set of functions model

Traditionally, Bayesians have held that the doxastic state of a rational agent can be represented by a single probability function. The value assigned to a proposition P by this function is the agent's *credence* in P , a measure of her subjective level of confidence.

Correspondence to: E-mail: srinard@gmail.com

However, some Bayesians object that our levels of confidence are not—and should not be—so fine-grained.¹ They propose to model an agent's doxastic state with a *set* of probability functions, called the agent's representor. (See, e.g., Jeffrey (1983), van Fraassen (1990), Hajek (2003), and Joyce (2005, 2010)). This model can be seen as analogous to the supervaluationist approach to vagueness. On supervaluationism, a proposition is determinately true if true according to all admissible precisifications of the term in question; if true according to some, but not all, admissible precisifications, it is indeterminate whether it is true.

Analogously, we can see each probability function as one possible precisification of the agent's doxastic state; the *admissible* precisifications are just those functions that are members of the agent's representor. So, a proposition about an agent's doxastic state is *determinately* true if true according to *every* function in the agent's representor. If a proposition is true according to *some, but not all*, functions in the representor, it is *indeterminate* whether it is true. (A proposition *P* about an agent's doxastic state is true *according to* a function *Pr* just in case, if the agent's doxastic state were properly represented by the single function *Pr*, *P* would be true. For example, if $\text{Pr}(P) = .8$, it is true *according to Pr* that the agent's credence in *P* is .8.)

For example, consider the proposition, LUCKY, that someday I will find a four-leafed clover. I am more confident that a fair coin will land heads than that LUCKY is true, so every function in my representor must have $\text{Pr}(\text{LUCKY}) < .5$. However, I am more confident of LUCKY than that pigs fly, so every function must have $\text{Pr}(\text{LUCKY}) > \text{Pr}(\text{pigs fly})$. I regard LUCKY as independent of the proposition that the number of stars is even, so all functions have $\text{Pr}(\text{LUCKY}) = \text{Pr}(\text{LUCKY} | \text{number of stars is even})$. However, there is no number *r* such that it's determinate that my level of confidence in LUCKY is exactly *r*. It's indeterminate whether my level of confidence in LUCKY is .09992, or .10001, or .1, etc. So, at least one function in my representor has $\text{Pr}(\text{LUCKY}) = .09992$; another has $\text{Pr}(\text{LUCKY}) = .10001$; etc.

The representor determines a spread of values for each proposition. I assume that this is always an interval, and I will use “agent *A* has credence [*c*, *d*] in *P*” to mean that for every *r*, *r* is a member of [*c*, *d*] if and only if there is some function in *A*'s representor with $\text{Pr}(P) = r$.² Importantly, these interval-valued “credences” do not have all of the properties we normally associate with credences. For example, we would normally say that an agent is more confident of *P* than *Q* just in case her credence in *P* is greater than her credence in *Q*. However, this doesn't work for interval-valued credences. First, it's not clear that the “greater than” relation is always defined for intervals. Second, there can be cases in which an agent has the same interval-valued credence for both *P* and *Q*, but she is more confident of *P* than *Q*, because each function in her set has $\text{Pr}(P) > \text{Pr}(Q)$.³ On the supervaluationist interpretation, An agent is determinately more confident of *P* than *Q* if and only if every function in her representor has $\text{Pr}(P) > \text{Pr}(Q)$.

This paper will presuppose the supervaluationist interpretation of the set of functions model in arguing against the wide interval view. This interpretation is a standard one; it is endorsed by many proponents of the model (including van Fraassen (1990), Hajek (2003), and Joyce (2010)). I am not aware of any other existing interpretation that would

vindicate the wide interval view. However, a committed defender of this view may seek to respond to my arguments by attempting to develop such an alternative interpretation. Although I suspect there is no tenable alternative that is compatible with the wide interval view, space constraints prevent me from defending that claim here. In this paper, I will rest content with the conditional claim that *if* we assume the supervaluationist interpretation, then the wide interval view must be rejected.

An argument against the wide interval view

Recall the two urns GREEN and MYSTERY. An agent—let's call her Candace—is certain that GREEN contains only green marbles, but has no information about the colors of the marbles in MYSTERY. As noted above, Candace should be determinately more confident of G-green (the marble drawn from GREEN will be green) than M-green (the marble drawn from MYSTERY will be green). So, if she is rational, every function in her representor will have $\text{Pr}(\text{G-green}) > \text{Pr}(\text{M-green})$. It follows that no function in her representor has $\text{Pr}(\text{M-green}) = 1$. So—crucially—her credence in M-green is not $[0, 1]$, since this interval contains 1, a value not assigned to M-green by any function in her representor.

So, the fact that Candace should be more confident of G-green than M-green undermines the wide interval view. Many proponents of the set of functions model endorse some version of this view. For example, Joyce (2005) holds that Candace's credence in M-green should be $[0, 1]$ (he discusses a case identical in all relevant respects). Kaplan (1996) claims that if one has no information concerning the truth of P, one's credence in P should be $[0, 1]$; Walley (1991, 228) has a similar view. Since Candace has no evidence concerning M-green, Kaplan and Walley would both agree with Joyce (2005) that her credence in this proposition should be $[0, 1]$. Weatherson (2007) holds that an agent with no empirical information should have credence $[0, 1]$ for all propositions that are not a priori certain.⁴

Let's consider in more detail the crucial claim that Candace should be determinately more confident of G-green than M-green. I motivated this by saying that she should be certain of G-green—or very nearly so—but nowhere near certain of M-green.

I don't think anyone would object to the assumption that she should be nowhere near certain of M-green. After all, she has absolutely no information concerning the colors of the marbles in MYSTERY. They could be all red; or all black; or half yellow half lavender; etc. It would clearly be irrational for her to be certain, or almost certain, that the marble will be green.

What about the assumption that she should be certain, or very nearly so, of G-green? I stipulated that Candace is certain that GREEN contains only green marbles, but one might object that a rational agent would never be certain of any empirical proposition. I can grant this, though, while stipulating that Candace has the best possible evidence for the claim that GREEN contains only green marbles. Surely there is some possible evidence given which she should be nearly certain of this, and so nearly certain that the next drawn marble will be green. If so, then she should be determinately more confident of G-green than M-green.⁵

Miriam Schoenfield (2012) has independently developed a similar argument against the wide interval view. She points out that, if we have no empirical evidence, we should be more confident of P than Q if Q entails P but P does not entail Q . So every function in our set should have $\Pr(P) > \Pr(Q)$, which means we cannot assign $[0, 1]$ to both P and Q . (The function with $\Pr(P) = 0$ would violate $\Pr(P) > \Pr(Q)$.)

At this point, a defender of the wide interval view may retreat to the position that the rational credence for Candace to have in M -green is the *open* interval $(0, 1)$. However, a slightly modified version of the argument undermines this proposal as well.

Consider a third urn, GREENEXCEPT. Candace knows that this urn contains exactly 1 million marbles, all of which are green, except one, which is white. Let GE -green be the proposition that the next marble drawn at random from this urn will be green. Given Candace's background knowledge, her credence in GE -green should be exactly $999,999/1,000,000$. So, if she is rational, every function in her representor will have $\Pr(GE\text{-green}) = 999,999/1,000,000$. She is very nearly certain that GE -green is true.

Since she is far from certain of M -green, she is determinately more confident of GE -green than she is of M -green. So all functions in her representor have $\Pr(GE\text{-green}) > \Pr(M\text{-green})$; so they all have $\Pr(M\text{-green}) < 999,999/1,000,000$. But then her credence in M -green is not $(0, 1)$, since this interval contains some (in fact, infinitely many) numbers greater than $999,999/1,000,000$. I conclude that Candace's credence in M -green should be neither $[0, 1]$ nor $(0, 1)$.

Another problem for the wide interval view

The wide interval view faces another problem: it renders inductive learning impossible in certain cases. (This is also noted in Weatherson (2007) and Joyce (2010).) For example, consider an urn about which you know only the following: either all the marbles in the urn are green ($H1$), or exactly one tenth of the marbles are green ($H2$). You have no information concerning which of these two hypotheses is true; so, in accordance with the wide interval view, you have credence $[0, 1]$ in each. You then learn that a marble drawn at random from the urn is green (E). On the set of functions model, updating proceeds by individually conditionalizing each function in your representor on your new evidence.⁶ Each function in your representor will have its posterior probability for $H1$ determined by Bayes' rule as follows: $\Pr(H1|E) = \Pr(E|H1)\Pr(H1) / [\Pr(E|H1)\Pr(H1) + \Pr(E|H2)\Pr(H2)]$. In this case, all the functions in your representor agree on the likelihoods, as they are fixed by objective chances in accordance with the Principal Principle: $\Pr(E|H1) = 1$ and $\Pr(E|H2) = 1/10$. Substituting $1 - P(H1)$ for $P(H2)$ and simplifying yields $\Pr(H1|E) = \Pr(H1) / [1/10 + 9/10\Pr(H1)]$. $\Pr(H1|E) = 1$ when $\Pr(H1) = 1$, and $\Pr(H1|E) = 0$ when $\Pr(H1) = 0$. Moreover, since this function is increasing and continuous, for every r_2 strictly between 0 and 1, there is some r_1 strictly between 0 and 1 such that when $\Pr(H1) = r_1$, $\Pr(H1|E) = r_2$ (see Figure 1).

This means that after conditionalizing each function in your representor on E , you have exactly the same spread of values for $H1$ that you did before learning E —namely, $[0, 1]$. Moreover—and this is what is most problematic—your credence for $H1$ will

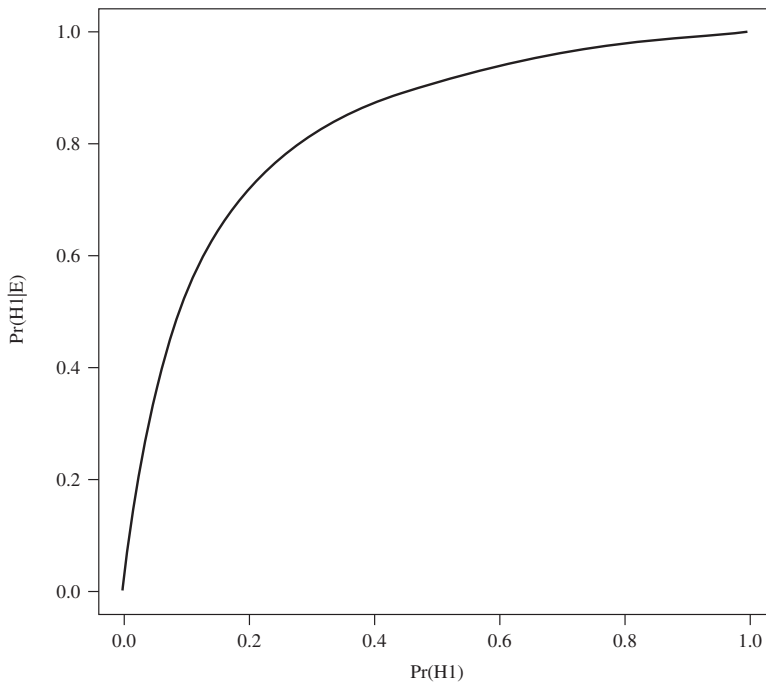


Figure 1: Graph of $\text{Pr}(H1|E) = \text{Pr}(H1) / [1/10 + 9/10\text{Pr}(H1)]$. $y = \text{Pr}(H1|E)$, $x = \text{Pr}(H1)$.

remain at $[0, 1]$ no matter how many marbles are sampled from the urn and found to be green. Suppose, for example, you learn that four billion marbles were sampled (with replacement) from the urn, and all were found to be green. In light of this strong evidence for $H1$, you should be fairly confident that $H1$ is true. At the very least, you should be more confident of $H1$ than its negation (which in this case is $H2$), so your level of confidence in $H1$ should be greater than .5. However, we can show, simply by iterating the argument in the previous paragraph, that your credence in $H1$ will still be $[0, 1]$, even after conditionalizing every function in your representor on the evidence that four billion marbles were sampled from the urn and found to be green.

As before, retreating to the claim that the interval should be $(0, 1)$ will not solve the problem. The considerations above also show that if your initial credence in $H1$ is $(0, 1)$, it will remain there. It will be impossible for you to become confident in $H1$, no matter how many marbles are sampled and found to be green.

What made the wide interval view initially plausible?

I have now given two arguments against the wide interval view. But in order to make the case as compelling as possible, we should consider what reasons one might have had for adopting the view in the first place. I will discuss two possible motivations.

The first can be summed up by the slogan “no evidence, no constraints.” If you have no evidence relevant to some proposition P , what grounds could you have for placing any constraints on admissible probability functions with respect to their treatment of P ? And if you have no such constraints, then for every r between 0 and 1, your representor will contain a function with $\Pr(P) = r$.

As we have seen, this line of thought is mistaken. It is plausible that which constraints are appropriate is determined (at least in part) by the state of your evidence concerning P , but it doesn't follow that lack of evidence results in no constraints. For propositions like *M-green*, the proper response to lack of evidence is uncertainty—to be far from certain. As we have seen, that places a constraint on admissible probability functions.

The second motivation appeals to our ignorance of objective chances. One familiar connection between chance and credence is codified in the Principal Principle, which says (roughly) that if you know the objective chance of P is r , then every function in your representor should have $\Pr(P) = r$. This might lead you to suppose that if you *don't know* whether P 's chance is r , then your representor should contain at least one function with $\Pr(P) = r$. After all, why rule out a function that assigns to P a number that, for all you know, matches E 's objective chance? In the case of *M-green*: for any number r between 0 and 1, for all you know, the percentage of green marbles in MYSTERY is r , and so for all you know, the objective chance of *M-green* is r .

There is a kernel of truth here: for every such r , you should not rule out the possibility that the objective chance of *M-green* is r . However, to exclude from your representor functions according to which $\Pr(\text{M-green}) = r$ is *not* to rule out the possibility that the *objective chance* of *M-green* is r . To rule out r as a possible chance for *M-green* would be to have credence zero in the proposition “the objective chance of *M-green* is r .” Eliminating functions from your representor with $\Pr(\text{M-green}) = r$ does not require you to have credence zero in the proposition “the chance of *M-green* is r .”

Moreover, both Joyce (2010) and Weatherson (2002) explicitly reject the claim that your representor must contain a function with $\Pr(E) = r$ for every r that, for all you know, is the objective chance of E . Joyce's reason is as follows (the case is modified slightly): Suppose you are presented with two identical-looking urns. One of them, you are told, contains only blue balls; the other contains only red balls. A fair coin is flipped; if it lands heads, the all-blue urn is selected; if tails, the all-red urn is selected. You are ignorant of the outcome of the coin toss, but are then presented with whichever urn was selected. A ball will be drawn from the selected urn. What is your credence in the proposition, B , that it will be blue? Joyce holds that in such a case, your credence should be precisely .5; that is, every function in your representor should have $\Pr(B) = .5$. This is so even though, for all you know, the objective chance of B is 1; and, for all you know, the objective chance of B is 0.

Alternatives to the wide interval view

Suppose, then, that we reject the wide interval view. What credence *should* Candace have in *M-green*, given that we've ruled out both $[0, 1]$ and $(0, 1)$? I won't give a detailed answer, but will make some brief remarks.

One option is to return to the traditional Bayesian view that any rational agent can be represented by a *single* probability function. As we saw earlier, though, this view suffers from the problem of false precision. Alternatively, one might retain the set of functions model, and represent Candace's credence in M-green by some interval other than $[0, 1]$ or $(0, 1)$. However, this view faces its own version of the problem of false precision. What are the endpoints of that interval? Any *particular* choice seems arbitrary. To specify one unique interval would be to claim that for every real number r , either r is determinately *not* Candace's credence (those numbers outside the interval), or it is *indeterminate* whether r is Candace's credence (those numbers inside the interval). Some numbers are easy to classify. 0 is determinately not Candace's credence; likewise for .00000001. But what is the lowest number r such that it's indeterminate whether r is her credence? The problem is that there isn't one. There is no sharp cut-off between numbers that are determinately not her credence, and those such that it's indeterminate whether they are.

This is analogous to the problem of higher-order vagueness, which arises for those who try to classify each shade on the color spectrum into one of three categories: determinately red, indeterminately red, or determinately not red. The problem is that there is no sharp cut-off between these categories. There isn't a *first* shade that is indeterminately red, just as there isn't a *lowest* number that is indeterminately Candace's credence.

One alternative approach to color vagueness involves many-valued logics, on which each shade is assigned a precise numerical degree of redness. The analogous approach would be to represent a doxastic state with a *fuzzy* set of probability functions. Different functions are assigned different precise degrees of membership in the set. Unfortunately, the problem remains. What is the first shade with a degree of redness less than 1? There is no good answer. What is the lowest number r such that $\Pr(\text{M-green}) = r$ according to some function whose degree of membership in the set is greater than 0? Again, there is no good answer.

All three options I've mentioned so far suffer from some version of the problem of false precision. I will now describe an alternative approach that avoids this problem. The cost is that we must give up on the idea that formal methods can provide a maximally specific accurate representation of an agent's doxastic state.

First, note that *some* ordinary descriptions of doxastic attitudes are *more specific* than others. For example, saying that you are *almost certain* of P is more specific than saying that you *believe* P. Both descriptions are accurate; one is more specific than the other.

We can give a perfectly accurate, though rather unspecific, description of the doxastic state that Candace should be in, if we say that she is more confident of M-green than any proposition in which she has credence determinately less than .00001, and less confident of M-green than any proposition in which she has credence determinately greater than .99999. We can even use the interval $[\text{.00001}, \text{.99999}]$ to characterize her doxastic state, if this is understood as equivalent to the description just given. This use of interval notation is importantly different from that used elsewhere in this paper. On the interpretation used elsewhere, to assign the interval $[c, d]$ to P is to say two things: first, that every number outside $[c, d]$ is determinately *not* the agent's credence; second, that for every number r *inside* $[c, d]$ it's indeterminate whether the agent's credence is

r . As we have seen, descriptions of this kind suffer from the problem of false precision, analogous to the problem of higher-order vagueness.

On the new interpretation proposed here, the assignment of $[c, d]$ to P asserts something weaker, namely, *only* the first of the two claims just mentioned: that every number outside $[c, d]$ is determinately not the agent's credence. We do *not* interpret this as also saying that for every r inside $[c, d]$, it's indeterminate whether the credence is r . We simply remain silent on that question. The interval may contain some numbers that are determinately not the agent's credence, or even a number that determinately is the agent's credence.⁷

On this interpretation, there will be multiple intervals that accurately characterize an agent's doxastic attitude toward a single proposition P . For example, as just noted, the attitude it would be rational for Candace to take towards M -green is accurately characterized by $[\text{.00001}, \text{.99999}]$, since for each r outside this interval, it is determinate that her credence is not r . It is also accurately characterized by $[\text{.000011}, \text{.99998}]$, since this interval also excludes only numbers that are determinately not her credence. Both of these descriptions are perfectly accurate; the second is more specific than the first, but neither is maximally specific.

On this approach we gain accuracy, but it comes at the cost of specificity. I conjecture that this trade-off is *in-principle unavoidable*: that it is simply not possible to provide a maximally specific yet fully accurate formal representation of belief.

Conclusion

I have presented two arguments against the wide interval view. First, I pointed out that this view is inconsistent with the observation that extreme ignorance with respect to P requires one to be far from certain that P is true. Second, I noted that in some cases, the view renders inductive learning impossible. I then considered two motivations for the view, and found them both to be unsuccessful. Finally, a discussion of alternatives to the view led me to conjecture that it is not possible to provide a formal representation of belief that is both fully accurate and maximally specific.

Acknowledgments

Many thanks to Adam Elga, Brie Gertler, Andrew Graham, Daniel Greco, Alan Hajek, Ned Hall, Brian Hedden, Sarah Moss, Paul Schoenfield, Roger White, and two anonymous reviewers for *Thought*. Special thanks to Vann McGee for answering a technical question. I am deeply indebted to Miriam Schoenfield for very helpful feedback on multiple drafts.

Notes

- 1 One motivation for imprecise credences concerns *thinness* of evidence. However, sharp credences can be required even when evidence is maximally thin. Consider "If a fair coin is tossed, it will land heads." Someone with no evidence should have credence precisely .5 in this proposition. Moreover, as Schoenfield (2012) points out, credal imprecision can arise

due to the *complexity* of even a substantial body of evidence. Thinness of evidence is neither necessary nor sufficient for credal imprecision.

- 2 Note that this is compatible with the point emphasized in Joyce (2010) that an agent's interval for P does not give a *complete* representation of their doxastic attitude towards P (this is done by the entire set of functions itself), but rather describes one feature of it.
- 3 This can happen if you assign some *open* interval, to both P and Q.
- 4 In the text I stipulate that the agent has some empirical information (concerning the contents of GREEN), and so the argument may not seem to touch Weatherson's version of the view. However, a modified version does undermine Weatherson's view: we can stipulate instead that the agent has no empirical information, and replace G-green with a simple tautology, like $P \rightarrow P$.
- 5 Additional support for this claim may be garnered from the fact that a rational agent should clearly prefer to receive a gift of \$10,000 if G-green is true than to receive a gift of \$10,000 if M-green is true.
- 6 Although this is the standard view of updating, it is not universal. Weatherson (2007) denies it, partly because of the updating problem that results when combined with the wide interval view.
- 7 We can even use the formalism of sets of probability functions, as long as we are careful to interpret it differently, as follows: a set of functions counts as accurate just in case the properties shared by all functions in the set correspond to determinate facts about the agent's doxastic state. There will be multiple sets that satisfy this requirement. Some will be more specific than others.

References

- van Fraassen, B. "Figures in a Probability Landscape," in *Truth or Consequences*, edited by M. Dunn and A. Gupta. Dordrecht: Kluwer, 1990.
- Hajek, A. "What Conditional Probability Could Not Be." *Synthese* 137.3 (2003): 273–323.
- Jeffrey, R. C. "Bayesianism with a Human Face," in *Minnesota Studies in the Philosophy of Science. Testing Scientific Theories*, Vol. 10, edited by J. Earman. Minneapolis: University of Minnesota Press, 1983.
- Joyce, J. "How Probabilities Reflect Evidence." *Philosophical Perspectives* 19.1 (2005): 153–78.
- Joyce, J. "A Defence of Imprecise Credences in Inference and Decision Making." *Philosophical Perspectives* 24.1 (2010): 281–323.
- Kaplan, M. *Decision Theory as Philosophy*. Cambridge: Cambridge University Press, 1996.
- Schoenfield, M. "Chilling Out on Epistemic Rationality." *Philosophical Studies* 158.2 (2012): 197–219.
- Walley, Peter. *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall, 1991.
- Weatherson, B. "Keynes, Uncertainty, and Interest Rates." *Cambridge Journal of Economics* 26.1 (2002): 47–62.
- Weatherson, B. "The Bayesian and the Dogmatist." *Proceedings of the Aristotelian Society* 107 (2007): 169–85.