Noûs

NOÛS 49:1 (2015) 197–200
doi: 10.1111/nous.12057

In Defense of IP: A Response to Pettigrew[1]

J. T. Ismael
University of Arizona

Back in 2008 I argued that Lewis misstated his own Principal Principle when he formalized it. His verbal statement said simply that if one knows what the chance of e is then one should (barring magical information from the future) adopt it as one's credence. There was nothing in that statement about conditionalization on the truth of ur-chance functions. And I pointed out that ur-chance functions of the kind one finds in physics do not come with assignments of probability to propositions asserting the truth of ur-chance functions. Nor does anything in their epistemic use require that they be extended to deliver such assignments. Then I introduced IP as a natural generalization of PP that said that if you don't know what the correct ur-chance function is, you obtain your credences from a weighted mixture of the chances assigned by epistemically possible ur-chance functions.[2]

My response to Pettigrew's criticism of my proposal can be quick, since he has acknowledged that the problems that he adduces for IP don't arise unless ur-chance functions are extended to assign probabilities to propositions of the form [C_{ch}]. I will give a brief defense of the epistemology implicit in IP, motivate the restriction against extending ur-chance functions, and point out a mistake in his argument that a reductionist about chance must reject the restriction.

First a bit of terminology:

We distinguish first-order credences from second-order credences. **First order credences** assign probabilities to physical events. **Second-order credences** assign probabilities to propositions about what credences one ought to have. A **pure credence function** is a first-order credence function that represents the epistemic state of a *non-reflective subject*.

Mixtures are weighted sums of first-order unconditional credences assigned by reflective subjects who are uncertain which ur-chance function is correct.

Weights are the credences that subjects assign to credence functions.

IP*: The principle that a reflective agent R with second-order uncertainty applies to form credences is

$$b_t(A) = \sum \beta_t ch_t(A)$$

Where the ch_i is an epistemically possible pure credence function at t, $ch_t(A)$ is the t-chance assigned to A, and $\beta_t = b_t(C_{ch})$.[3]

IP says that a reflective subject who is unsure about which ur-chance function is correct (or 'epistemically ideal') should divide his credences between those *pure* credence functions that might be epistemically ideal by his lights in a way that reflects the probability he assigns to their being ideal. Beliefs about the relative reliability of first-order credence functions are formed by *reflective* agents dividing credence among pure credence functions.

Motivating the Restriction

So long as credence functions for reflective subjects are representable as mixtures of pure credence functions, the problems that Lewis ran into with undermining, as well as the kinds of problems that Pettigrew derives from IP without the restriction, don't arise. I view all of those problems as simply reinforcing the need for the restriction. And I view them as confirming the insight that problems arise for principles relating first order uncertainty about matters of fact and second order uncertainty about what degrees of belief one ought to have in the presence of first-order uncertainty because of indiscriminate mixing of first and second-order credences. IP* separates first-order credences from second-order credences firmly and builds credence functions for believers out of first-order credences by adding beliefs about the correctness (or epistemic ideal-ness) of those first-order credence functions. That accords with what we do intuitively in everyday contexts. If you are wondering whether to bring an umbrella and consult three weather stations, each assigning a probability for rain tomorrow, you combine their opinions by forming a mixture weighing each in proportion to your beliefs about how reliable they are.[4]

Pettigrew thinks that restricting the ch_t's to pure credence functions is not an option open to reductionists because the reductionist is committed to an entailment from descriptions of total history to propositions of the form $[C_{ch}]$, and so he thinks that if chance functions are defined on total histories, for the reductionist, they must be defined for ur-chance functions themselves. He writes:

> [T]he frequentist's ur-chance functions are certainly able to assign chances to entire histories of non-modal facts . . . So too are chances given by the best-system analysis. On that account, the chances are whatever our best theory says they are. The goodness of a theory is a function of its strength, simplicity, and fit, where the fit of a theory is the chance that it assigns to the entire history of the actual world. Thus, in order for a theory to even be a candidate for the best theory, the chance functions that it posits must be defined on this history. And if this is the case, then there seems every reason to believe that they will be defined on other possible, nonactual histories . . . .So it seems that, in order for the best-system analysis or the frequentist account to work, chances must be defined for chance hypotheses, as required by our arguments above.

But that is not correct. Let H be a total history for w and let's call the proposition *that* H is a total history $H^{\mathbf{total}}$. There is no entailment from H to $H^{\mathbf{total}}$, and that is sufficient to show that a frequentist or reductionist of Lewisian stripe will never be forced by her historical knowledge to assign a probability to a proposition of the form $[C_{ch}]$. So long as, for any time t, she assigns a positive probability to the world continuing indefinitely after t she will not by her own lights be in a position to form a second-order belief about which ur-chance function is correct. Indeed, that was crucial to the resolution IP provided to the problems that Lewis ran into with undermining. So long as there are epistemically possible futures among which one is dividing expectation that would undermine any firm belief about what the chances are (a run of frequencies or a sequence of events that would change the balance of simplicity, strength), IP will not dictate an assignment of a probability of 1 to any ur-chance function whatsoever, and *a fortiori* it will not assign a probability of 1 to a modest ur-function.

Auxiliary Questions

Where do the weights in IP come from? If pure credence functions do not assign themselves probabilities, the weights cannot come from theories of chance. One of the virtues of IP* is that it separates the probabilities derived from reasoning within the scope of a theory from the probabilities we apply in reasoning about which theory is correct, and makes it clear that the latter must have a separate source. There is an interesting connection to the problem of theory choice in the philosophy of science. The problem of finding objective weights is effectively the problem of theory choice in science. The reflective agent choosing between different conditional credence functions recapitulates that of the theorist choosing between different global theories since physical theories embody pure chance functions. In both cases, we have a presumably rational epistemic agent who needs an external source for assessing the reliability of theories that embody different pure conditional probability functions. How should he proceed? Assuming the evidence comes in the form of information about local matters of particular past fact, the theorist can knock out theories and renormalize as predictions are falsified, but how does he divide opinion among the remaining theories? The Bayesian solution is that the weights come from priors, but they at least cannot be rationally dictated by the theories themselves.

How much can we weaken the restriction without landing in trouble? How much can we enrich pure credence functions with second order probabilities and apply IP without landing in trouble. The following constraint is at least necessary; a pure credence function cannot be *absolutely* immodest, which is to say that it can be modest conditional on limited evidence, but cannot assign itself any probability less than 1 conditional on all evidence.

Notes

[1] I owe a very special thanks to Richard Pettigrew for his wonderful work on this, which certainly deepened my own thinking about it.

[2] I called IP the 'Generalized Principal Principle' and referred to ur-chance functions, following Lewis, as 'theories of chance'. Here I follow Pettigrew's terminology.

[3] IP* is just IP with the restriction to pure credence functions more explicitly stated. Pettigrew's expression of IP conditionalizes on E_t. Mine does not. It makes no difference. So long at E_t contains

no evidence from the future, the chances screen off E_t, because they screen off any historical evidence. "A Modest Proposal About Chance", *Journal of Philosophy*, 108 (8), pp. 416–442, 2011.

[4] And this is true even if they have a determinate degree of confidence in their own opinion. It is perfectly rational for Suzi and Bob to form different mixtures from the same set of recommendations by experts who are fully confident of their own opinions, because they disagree the relative reliability of the recommenders. That by itself ought to show that the weights have to have a separate source.