Comments on David Chapman's Essay, "Probability Theory Does Not Extend Logic"

Kevin S. Van Horn

July 23, 2016

1 Propositional versus Predicate Calculus

I agree with Chapman that probability theory does not extend the predicate calculus. I had thought this too obvious to mention, but perhaps it needs emphasizing for people who haven't studied mathematical logic. Jaynes, in particular, was not versed in mathematical logic, so when he wrote about "probability theory as extended logic" he failed to properly identify *which* logic it extended. I disagree with Chapman's insistence that "logic" means "the predicate calculus." There are, in fact, a variety of logics aside from the predicate calculus: the (classical) propositional calculus, intuitionistic logics, temporal logic, etc. In a technical setting the term "logic," with no qualifier, is ambiguous. So the title of the esssay is misleading; probability theory does, in fact, extend *a* logic, the propositional calculus.

Probability theory and the predicate calculus both extend the propositional calculus, in different directions: probability theory lets you reason about degrees of credibility, and the predicate calculus adds quantification ("for all x" and "there exists x"), which is needed to reason about infinite domains. I agree with Chapman that a complete theory of rationality needs both, but disagree as to why.

Chapman claims that "[the predicate calculus] is capable of expressing complex relationships among different objects, and probability theory is not," adding that "propositional calculus cannot talk about objects at all." This is not really true. You can certainly include in your set of atomic propositions expressions such as "$1 < 2$" or "$(3 + 5) > 10$." The important difference between the propositional calculus and the predicate calculus and propositional calculus is that the latter allows you to reason about infinite domains. On *finite* domains the propositional calculus and predicate calculus have equal expressive power and proof power; the latter merely provides a notation that is more concise for expressing certain kinds of statements, and allows shorter proofs in some cases.

For example, consider the syllogism Chapman discusses: Socrates is a man; all men are mortal; therefore Socrates is mortal. In the predicate calculus this would be expressed as

$$\text{isMan}\,(\text{Socrates})\ \& \\ (\forall x.\text{isMan}(x) \rightarrow \text{isMortal}(x))$$

from which we would conclude isMortal (Socrates). If you had a finite domain restricted to the four entities Socrates, Archimedes, Zeus, and Awesomeness, you could express the above in the propositional calculus as

$$\begin{gathered} \text{isMan}\,(\text{Socrates})\ \& \\ (\text{isMan}\,(\text{Socrates}) \rightarrow \text{isMortal}\,(\text{Socrates}))\ \& \\ (\text{isMan}\,(\text{Archimedes}) \rightarrow \text{isMortal}\,(\text{Archimedes}))\ \& \\ (\text{isMan}\,(\text{Zeus}) \rightarrow \text{isMortal}\,(\text{Zeus}))\ \& \\ (\text{isMan}\,(\text{Awesomeness}) \rightarrow \text{isMortal}\,(\text{Awesomeness}))\,, \end{gathered}$$

from which you could also conclude isMortal (Socrates). Note that this trick won't work on infinite domains, as you would need to AND together an infinite number of propositions.

My view is that the role of the predicate calculus in rationality is in model building. It gives us the tools to create mathematical models of various aspects of our world, and to reason about the properties of these models. The predicate calculus is indispensable for doing mathematics. On the other hand, when considering statements about specific things that may have happened, or that we may experience, or data we have acquired, we are in a finitary realm where the propositional calculus (and its extension to probability theory) are adequate. We are talking about specific, concrete instances of possible phenomena, we have only a finite number of observations, and we can experience only a finite portion of the world. A universally quantified statement of the form "for all x, $R(x)$ is true" is something that we can never observe nor experience.

2 Cox's Theorem

Chapman writes, "Technically, Cox's proof was simply wrong." This is untrue; it is *incomplete*, not *wrong*. Cox did not explicitly state all of his assumptions, but if you read his paper carefully they're straightforward to pick out. I speak from experience here.

Chapman complains, "Philosophically, it is unclear that all his requirements were intuitive. For example, the proof requires negation to be a twice-differentiable function." Modern versions of Cox's Theorem don't require this. They have the much weaker requirement that negation be a nonincreasing function: if the plausibility of A increases, the plausibility of not A should either stay the same or decrease. If you want to explore all the assumptions used in a modern version of Cox's Theorem, and the justifications for those assumptions, check out the "tour" of Cox's Theorem I wrote in 2003, available at `http://ksvanhorn.com/bayes/Papers/rcox.pdf`.

What is "intuitive" and what is not is certainly something that varies from one person to another. When you look at proposed alternatives to probability theory, however, the important point of departure seems to be the assumption, used in Cox's Theorem and variants, that the credibility of a proposition can be represented as a *single* number, rather than requiring, say, a pair of numbers as in Dempster-Shafer belief-function theory. I'm not aware of any remotely viable alternative to probability theory as a logic for reasoning about degrees of credibility that represents credibility as a single number. So I think it reasonable to say that Cox's Theorem, in its modern form, gives a good argument that

- probability theory is *a* logic of credible reasoning, and
- probability theory is the *simplest* viable logic of credible reasoning.

Being able to represent a degree of credibility as a single number is important in applications to instrumental rationality. The whole idea of optimizing expected utility depends on credibility being a single number. If instrumental rationality is your focus, it seems hard to avoid basing your epistemology on the laws of probability.

You might also want to check out de Finetti's Dutch Book arguments for probabilities as subjective degrees of belief. He shows that, if your degrees of belief don't follow the laws of probability, then you can be made to accept a bet that you are certain to lose. These arguments implicitly assume that "degree of belief" or "degree of credibility" is a single number, and not a pair of numbers or something more complex. A weakness of Dutch Book arguments is that they assume your degree of credibility is proportional to the price you would be willing to pay for a bet. This mixes up epistemology and decision making from the get go, and epistemology is arguably the more fundamental of the two, something that you should be able to discuss independently of applications to decision making.

Finally, I am currently writing up an alternative to Cox's Theorem that achieves the same result with different and, I think, much harder to dispute assumptions. My approach doesn't assume that degrees of credibility are single numbers, it does not assume that the credibility of a compound proposition $\neg A$ or $A \wedge B$ can be decomposed in any particular way, nor does it require the continuity assumptions of existing variants of Cox's Theorem.

3 The Iron Grip of Frequentism

Chapman seems wedded to viewing probabilities as frequencies, and *only* as frequencies. This comes through when he writes

> Suppose two events are "independent": approximately, there is no causal connection between them.

Thinking of probabilities as frequencies may lead you to think of probabilistic dependence as having to do with causal connections of some sort (perhaps a common cause), but when dealing with epistemic probabilities this need not be the case. For example, if you are considering two coin flips from the same coin, which may be either a coin biased 60/40 towards heads, or a coin biased 60/40 towards tails, then the coin flips are independent events if you know which coin is being used, but *dependent* events if you do not. (The result of the coin flip gives you evidence as to which coin it is, which then affects your probabilities for the next flip.)

Chapman also writes,

> Statistical inference is based on probability theory, and enables reasoning from specifics to generalities in some cases. It is not just probability theory, though;

—a statement which is only true of *frequentist* statistical inference, and is flat-out wrong for *Bayesian* statistical inference, which *is* in fact just probability theory. Chapman has such a determinedly frequentist mindset that he can't seem to even acknowledge the existence of a non-frequentist statistics.

Most tellingly, Chapman writes,

> Probability theory doesn't work when you have inadequate information. Implicitly, it demands that you always have complete confidence in your probability estimate. . .

which is a statement that only makes sense if you are talking about probabilities as frequencies. To speak of probability *estimates* in which one can have "complete confidence" is to assume that these probabilities are properties of the system under consideration. But we are talking about *epistemic* probabilities here, which are a function only of the *information* you have available. Chapman is falling prey to what Jaynes calls the Mind Projection Fallacy: treating an internal property of one's mind as if it were a physical property of an external system under analysis. His argument here against probability theory as a logic of justiable strength of belief is circular, in that his criticism assumes from the get-go that probabilities are *not* epistemic, do not measure justifiable strength of belief, but are rather frequencies.

By the way, this same comment applies to the alternative systems for reasoning about uncertainty he mentions. He links to a Wiki page on "imprecise probability". . . but the motivation for all of these systems is an assumption that probabilities are frequencies that are objective physical properties of the system under consideration, and therefore one may be unsure of what the "actual" probabilities are. The motivation for considering these alternatives collapses once you allow for epistemic probabilities.