Jeffrey Conditionalization and Scoring Rules

Jeff Dunn

DRAFT: September 6, 2017

1 Introduction

Suppose agents have credence functions, which satisfy probabilism, and which are defined over some set of worlds, $\mathscr{W}$. Distinguish two sorts of learning experiences such an agent could undergo. In the first kind of situation, the agent knows she will receive evidence that tells her that some one cell of a partition over $\mathscr{W}$ is true. In the second kind of situation, the agent knows that she will receive evidence that tells her that the weights she currently assigns to a partition over $\mathscr{W}$ need to be changed. For a simple example of the first sort of situation, we can imagine an agent who is about to learn whether E or $\overline{E}$. For a simple example of the second sort of situation, we can imagine an agent who is about to learn that her distribution of credence to E and $\overline{E}$ needs to be altered. Call the first situation a *conditionalization scenario* and the second situation a *Jeffrey scenario*.

The reason for these names is that many think that in a conditionalization scenario, the agent should plan to update via conditionalization, where conditionalization recommends that $c_{new}(X) = c_{old}(X|E)$ for all propositions X where E corresponds to the cell in the partition that was learned. In a Jeffrey scenario, on the other hand, many think that the agent should plan to update via Jeffrey conditionalization, where Jeffrey conditionalization recommends that $c_{new}(X) = \sum_i c_{old}(X|E_i) \times \mathbf{E}_i$ for all propositions X where $\mathbf{E}_i$ is the newly learned weight on the i^{th} cell of the evidence partition.

2 Conditionalization, Jeffrey Conditionalization, and Accuracy

How does the argument for planning to update via conditionalization go? The idea is this. The expected inaccuracy of an updating plan can be understood as the expected inaccuracy of the credence function that is the result of that plan, in the situations where that plan would be adopted. So, if $\mathscr{E}$ is the evidence partition, then let $R_{\mathscr{E}}$ be the updating plan. $R_{\mathscr{E}}$ is a function that outputs a particular credence function for every element of the partition. In particular, if E_w is the element of partition $\mathscr{E}$ that is true at w, then let R_{E_w} be the credence function that the plan $R_{\mathscr{E}}$ outputs at w. Let $\mathfrak{S}(c, w)$ be a measure of the inaccuracy of a credence function, c, in world, w. Hence, the expected inaccuracy of the plan $R_{\mathscr{E}}$ as viewed from the agent's current credence function, c, is: $\sum_w c(w)\mathfrak{S}(R_{E_w}, w)$.

Let's think how this works in a simple case. Suppose there are four worlds and the evidence partition consists of two elements, E and $\overline{E}$. Let's suppose that $w_1, w_2 \in E$ and $w_3, w_4 \in \overline{E}$. Now, what does an inaccuracy-minimizing $R_{\mathscr{E}}$ look like? The first thing to note is that it must be the case that $R_{E_{w_1}} = R_{E_{w_2}}$ and $R_{E_{w_3}} = R_{E_{w_4}}$, since part of what it is to be an updating plan is to give the same instructions for any worlds within a cell of a partition. For ease of reference, then, let R_E be the credence function that the plan recommends for w_1 and w_2 and $R_{\overline{E}}$ be the credence function that the plan recommends for w_3 and w_4. Thus, we want to find the R_E and $R_{\overline{E}}$ that minimize:

$$c(w_1)\mathfrak{S}(R_E, w_1) + c(w_2)\mathfrak{S}(R_E, w_2) + c(w_3)\mathfrak{S}(R_{\overline{E}}, w_3) + c(w_4)\mathfrak{S}(R_{\overline{E}}, w_4)$$

Note first that the inaccuracy minimizing plan must be such that R_E assigns E credence 1 and $R_{\overline{E}}$ assigns $\overline{E}$ credence 1. This is because at both w_1 and w_2 E is true, and so assigning it anything less than 1 would be more inaccurate. A similar thing is true for w_3 and w_4 with respect to $\overline{E}$. How should R_E and $R_{\overline{E}}$ treat other propositions? Greaves & Wallace prove that to minimize inaccuracy they should yield the verdicts of conditionalization. But for now, note something about the expected inaccuracy of this updating plan is calculated. In particular, we do not consider how R_E is expected to do from the perspective of every world the agent assigns credence to. Rather, we are concerned with is how well R_E is expected to do, only from the perspective of the E-worlds the agent assigns credence to. And a similar thing, of course, is true for $R_{\overline{E}}$: this plan is assessed only from the perspective of the $\overline{E}$-worlds the agent assigns credence to.

Noticing this is important, for we can then redescribe how the inaccuracy minimization argument for conditionalization goes in such a way that will make it more amenable to Jeffrey scenarios. Here is that redescription:

> Suppose I have credence function c now and we have specified an

> evidence partition, $\mathscr{E}$. We want to know what credence function, c', I should plan to transition to, supposing I learn that one of the elements of $\mathscr{E}$ is true. Here's how we do it. I should transition to the credence function, c' that (a) assigns the true element $E \in \mathscr{E}$ credence 1, and (b) has the lowest expected inaccuracy, from the perspective of c, ignoring the summands in this expectation that are weighted by $c(w_i)$ for $w_i \notin E$.

This redescription yields exactly the same results as the plan-based procedure mentioned above. And Greaves & Wallace's proof tells us that the c' that minimizes inaccuracy according to this procedure just is $c'(X) = c(X|E)$.

Now, consider how to adapt this procedure to Jeffrey scenarios where one learns not that an element of a partition is true, but rather that the evidence partition itself should be assigned different weights. For instance, suppose $c(E) = 0.3$ and $c(\overline{E}) = 0.7$, but I subsequently learn that E should be assigned a weight of $\mathbf{E}$ and $\overline{E}$ a weight of $\overline{\mathbf{E}}$. For instance, it may be that $\mathbf{E} = 0.6$ and $\overline{\mathbf{E}} = 0.4$. How do we evaluate updating plans in this kind of scenario, from the perspective of minimizing inaccuracy? The usual way to think about it is as follows: I should choose the credence function c' that (a) satisfies the evidential constraint (in this case, that $c'(E) = 0.6$ and $c'(\overline{E}) = 0.4$), and (b) has the lowest expected accuracy from the perspective of c. Notice that we don't ignore any summands because no worlds are eliminated by this evidence. If we follow this procedure, we get very different results for different scoring rules. If we use the log score, we get Jeffrey Conditionalization. If we use the Brier score, however, we get a very different updating rule (see Leitgeb & Pettigrew for details).

However, I think the procedure described in the previous paragraph is not quite right. The method for how we deal with Jeffrey scenarios should have some continuity with how we think of the more simple conditionalization scenarios. In the conditionalization scenarios, when E is learned, we look at the expected inaccuracy of the new credence function c' weighted only by $c(w_i)$ for $w_i \in E$. Why do we do that? Well, suppose you're going to do option A if w_1 or w_2 and B if w_3 or w_4. Once you learn that you are in w_1 or w_2, it doesn't matter how good A would be if you were in w_3 or w_4. Consider an example: suppose doubling my bet is a good idea if an ace is the next card dealt, but a bad idea otherwise. Once I learn that an ace is dealt, it doesn't matter to me in evaluating the decision to double my bet that doubling it would be a bad idea had an ace not come up.

Now, consider how this translates to Jeffrey scenarios. In evaluating R_E in the conditionalization scenario we completely ignore summands $c(w_i)$ for $w_i \notin E$. This is because we know that these $c(w_i)$ represent a mistaken view about what could happen. In the Jeffrey scenario, we should do something similar. Suppose, again, that $c(E) = 0.3$ and $c(\overline{E}) = 0.7$, but I subsequently learn that E should be assigned a weight of 0.6 and $\overline{E}$ a weight of 0.4. Then though I do not ignore any summands in the expected inaccuracy calculation completely, I should rescale the importance I give to them. In this scenario, for instance, when I come to calculate the expected inaccuracy of various possible credence functions I could transition to, I should weigh more heavily the influence of $c(w_i)$ for $w_i \in E$ compared to the $c(w_j)$ for $w_j \in$

$\overline{E}$. This is because I've just learned that c was assigning too much credence to the $\overline{E}$-worlds compared to the E-worlds. This will make our procedure for determining updating rules in Jeffrey scenarios continuous with what we do when propositions are learned in conditionalization scenarios.

Applied to the Jeffrey scenario we've been considering, that is, we want to choose the c' that respects the evidence concerning E and $\overline{E}$ and that minimizes:

$$\alpha \sum_{w_i \in E} c(w_i) \times \mathfrak{S}(c', w_i) + \beta \sum_{w_j \in \overline{E}} c(w_j) \times \mathfrak{S}(c', w_j)$$

where α and β are the appropriate weights. A natural way to assign these weights is as follows: $\alpha = \mathbf{E}/c(E)$ and $\beta = \overline{\mathbf{E}}/c(\overline{E})$. One reason this is natural is that it converges on the procedure for learning the proposition E in a conditionalization scenario as $\overline{\mathbf{E}} \to 0$, since in that case we completely ignore the summands for $c(w_j)$ for $w_j \in \overline{E}$. Further, $\mathbf{E}/c(E)$ is a natural way to measure how much more influence you need to give to different worlds in light of the information learned in a Jeffrey scenario. If $\mathbf{E}/c(E) > 1$ then you need to give those E-worlds more influence; if $\mathbf{E}/c(E) < 1$ you need to give E-worlds less influence. And if $\mathbf{E}/c(E) = 1$, this dictates no change in credence, which is just what we want (since nothing is learned).

The interesting thing is that if we set up things this way, then we get the result that every strictly proper scoring rule leads to Jeffrey Conditionalization. To see this, consider the simplest case where the evidence is $\mathbf{E}$, $\overline{\mathbf{E}}$. We want to choose the c' such that $c'(E) = \mathbf{E}$, $c'(\overline{E}) = \overline{\mathbf{E}}$, and that minimizes:

$$\mathbf{E}/c(E) \sum_{w_i \in E} c(w_i)\mathfrak{S}(c', w_i) + \overline{\mathbf{E}}/c(\overline{E}) \sum_{w_i \in \overline{E}} c(w_j)\mathfrak{S}(c', w_j).$$

Because $c'(E) = \mathbf{E}$ and $c'(\overline{E}) = \overline{\mathbf{E}}$, this is equivalent to:

$$\sum_{w_i \in E} c'(E)c(w_i)/c(E)\mathfrak{S}(c', w_i) + \sum_{w_j \in \overline{E}} c'(\overline{E})c(w_j)/c(\overline{E})\mathfrak{S}(c', w_j)$$

Since $\mathfrak{S}$ is strictly proper, the c' that minimizes this sum is the function such that

$$c'(w_i) = \frac{c'(E)c(w_i)}{c(E)}$$

$$c'(w_j) = \frac{c'(\overline{E})c(w_j)}{c(\overline{E})}$$

Now, notice that if we Jeffrey conditionalize on $\mathbf{E}$, $\overline{\mathbf{E}}$, then for any world, w:

$$c'(w) = c'(E)c(w|E) + c'(\overline{E})c(w|\overline{E})$$

For worlds $w_i \in E$, $c(w_i|\overline{E}) = 0$ and for worlds $w_j \in \overline{E}$, $c(w_j|E) = 0$. Hence:

$$c'(w_i) = c'(E)c(w_i|E) = \frac{c'(E)c(w_i \wedge E)}{c(E)}$$

$$c'(w_j) = c'(\overline{E})c(w_j|\overline{E}) = \frac{c'(\overline{E})c(w_i \wedge \overline{E})}{c(\overline{E})}$$

Because all the w_i are wholly contained in E and all the w_j are wholly contained in $\overline{E}$, $c(w_i \wedge E) = c(w_i)$ and $c(w_j \wedge \overline{E}) = c(w_j)$. Hence we have:

$$c'(w_i) = \frac{c'(E)c(w_i)}{c(E)}$$

$$c'(w_j) = \frac{c'(\overline{E})c(w_j)}{c(\overline{E})}$$

This suffices to show that the credence function that minimizes expected inaccuracy—appropriately weighted—relative to any proper score in a Jeffrey scenario is the credence function that results from Jeffrey Conditionalization.