A new resolution of the Judy Benjamin problem

⋆ ⋆ ⋆

Igor Douven
Institute of Philosophy
University of Leuven

Jan-Willem Romeijn
Faculty of Philosophy
University of Groningen

Contents

1 The Judy Benjamin problem

In an example by van Fraassen [1981], Judy Benjamin is dropped in an area divided into Red (R) and Blue ($\neg R$) and into Second Company (S) and Headquarters ($\neg S$) sections. She assigns equal probability to all quadrants Q_i.

	S	$\neg S$
R	Q_1	Q_2
$\neg R$	Q_3	Q_4

She then receives this radio message: “I can’t be sure where you are. If you are in Red territory, the odds are 3 : 1 that you are in Headquarters area.” How should she adapt her probabilities?

Intuitive desiderata for a solution

D1: Her conditional probability for being in $\neg S$ given that she is in R should be three times her conditional probability for being in S given that she is in R;

D2: none of her conditional probabilities given any proposition in $\{\neg R, R \wedge S, R \wedge \neg S\}$ should change; and

D3: the probability of being in R should not change: no information has been received that would seem relevant to that.

Distance minimization

Various “distance minimization rules” have been considered in relation to the Judy Benjamin problem.

The idea is that the information contained in the radio message imposes a specific constraint on the probability assignment over the segments Q_i and that Judy’s new probability function should be the one that satisfies this constraint and otherwise deviates as little as possible from her old one.

The phrase “as little as possible” is then spelled out in terms of a distance function.

Relative entropy distance minimization
The best studied distance function is relative entropy:

$$\mathrm{RE}(\Pr, \Pr') = \sum_i \Pr'(Q_i) \log \frac{\Pr'(Q_i)}{\Pr(Q_i)},$$

Using this, we can look for the closest new probability assignment that satisfies the constraint:

$$\Gamma = \left\{ \Pr : \frac{\Pr(Q_2)}{\Pr(Q_1)} = 3 \right\}, \qquad \Pr_{\mathrm{new}} = \{\Pr \in \Gamma : \mathrm{RE}(\Pr, \Pr_{\mathrm{old}}) \text{ minimal}\}.$$

Two worries

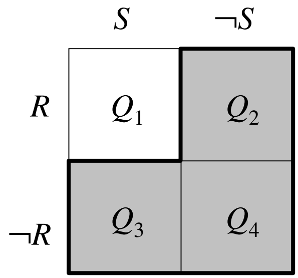
1. All the distance minimization rules that have been discussed in the context of the Judy Benjamin problem violate desideratum D3: Judy's probability for R changes after hearing the radio message;

2. each of these rules leads Judy to assign a *different* probability to R, but there are no principled grounds for choosing between the rules.

The second worry is considered to be still an open problem. Van Fraassen has tried to explain away the first.

Worry 1 removed?

If Judy determines her new probability for *R* by means of any of the distance minimization rules, it decreases; so D3 is violated.



But according to van Fraassen, the intuition underlying that desideratum is not to be trusted anyway. For consider the limiting case: if Judy learns “If in Red, then in Headquarters, period,” the decrease in the probability of *R* is a matter of course. Or is it?

2 Updating on conditionals

Van Fraassen is supposing that upon learning a conditional, one should conditionalize on the corresponding material conditional. An attractively simple picture of updating on conditionals: is it correct?

There is little one can say about conditionals that is *not* controversial. Do conditionals have truth conditions? If yes, what are they? If no, how do we account for compounds of conditionals?

The question how we ought to update on a conditional has even been largely ignored in the philosophical literature. (Has the question how people actually update on a conditional been equally ignored in the psychological literature?)

Conditionalizing on material conditionals

It is a well-known fact that

If (i) $0 < \Pr(A) < 1$, (ii) $0 < \Pr(B)$, and (iii) $\Pr(B|A) < 1$, then $\Pr(A|A \supset B) < \Pr(A)$.

But consider this example: Sarah and Marian have arranged to go for sundowners at the Westcliff hotel tomorrow. Sarah feels there is some chance that it will rain, but thinks they can always enjoy the view from inside. To make sure, Marian consults the staff at the Westcliff hotel and finds out that in the event of rain, the inside area will be occupied by a wedding party. So she tells Sarah: "If it rains tomorrow, we cannot have sundowners at the Westcliff." Upon learning this conditional, Sarah sets her probability for sundowners *and* rain to zero, *but she does not optimistically reduce her probability for rain after learning the conditional!*

Comments

This is not the end of the material conditionals account:

- advocates of the material conditional account are not necessarily committed to conditionalization;
- even if they are, they can legitimately claim that in learning a conditional we come to know more than the corresponding material conditional. (For instance, we may in addition come to know that there is a certain evidential relationship between antecedent and consequent.) And $\Pr(A \mid A \supset B) < \Pr(A)$ is compatible with $\Pr(A \mid (A \supset B) \wedge C) \geqslant \Pr(A)$.

Still, we cannot uncontroversially defend the consequences of a minimum relative entropy update in the Judy Benjamin case by referring to conditionalization on a material conditional.

An alternative proposal

The bulk of the proposal concerns the kind of case for which, pre-theoretically, generalized versions of the desiderata D1–D3 hold (so in particular cases in which the learning of a conditional is or would be irrelevant to one's probability for the conditional's antecedent, as in the Judy Benjamin case).

The update rule for this kind of case consists of two parts. This is the first:

> After learning "If A, then the odds for $B_1, \ldots, B_n$ are $c_1 : \cdots : c_n$," where $\{\neg A, A \wedge B_1, \ldots, A \wedge B_n\}$ is a partition, a person should set her probability for B_i conditional on A equal to $c_i \big/ \sum_{j=1}^{n} c_j$, for all i.

For instance, after learning the radio message, Judy should set her conditional probability for being in Headquarters Company area given that she is in Red territory equal to 3/4 and her probability for being in Second Company area given that she is in Red territory equal to 1/4.

Adams conditioning

The second part of the update rule consist of a proposal made in Bradley [2005] in the context of preference kinematics, which is called *Adams conditioning*:

> Given a partition $\{U_0, U_1, ... U_n\}$, and supposing we obtain new probabilities $\mathrm{Pr}_{\mathrm{new}}(U_i)$ for $i = 1, ..., n$, the new probability $\mathrm{Pr}_{\mathrm{new}}$ must be as follows:
>
> $$\mathrm{Pr}_{\mathrm{new}}(C) \;=\; \mathrm{Pr}_{\mathrm{old}}(C \,|\, U_0)\,\mathrm{Pr}_{\mathrm{old}}(U_0) + \sum_{i=1}^{n} \mathrm{Pr}_{\mathrm{new}}(U_i)\,\mathrm{Pr}_{\mathrm{old}}(C \,|\, U_i).$$

It follows from a theorem proven by Richard Bradley that if we update on conditionals whenever D1–D3 hold pre-theoretically, then our posterior probability function *will* satisfy these desiderata.

Adams vs. Jeffrey I

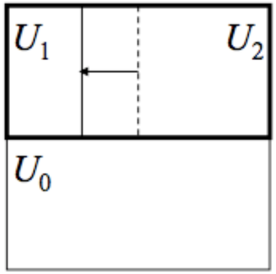
Information does not always come in neat propositional packages. Richard Jeffrey devised a rule for updating a probability assignment on new information captured by a probability assignment over a partition of possible events.

$$\mathrm{Pr}_{\mathrm{new}}(C) = \sum_i \mathrm{Pr}_{\mathrm{new}}(Q_i)\,\mathrm{Pr}_{\mathrm{old}}(C \mid Q_i).$$

Jeffrey's rule does not tell us how we can obtain this probability assignment over the partition of Q_i, other than that it stems from our observation and experience.

Adams vs. Jeffrey II

Say that we learn "If R, then the odds for $\neg S : S$ are $q_1 : q_2$", and that, upon learning this, we do not want to adapt our degree of belief $\Pr(R) = r$.



We can achieve this by applying Jeffrey conditionalization to the partition of events $\mathcal{U} = \{U_0, U_1, U_2\} = \{\neg R, R \wedge \neg S, R \wedge S\}$ using the odds, ${}^{(1-r)}\!/_{r}(q_1 + q_2) : q_1 : q_2$.

Adams vs. Jeffrey III

So, Jeffrey conditionalization solves the Judy Benjamin problem in a way that respects D1–D3. This is a surprise, as the problem had been taken to motivate the search for update mechanisms other than Bayes's and Jeffrey's rules.

On the other hand, once we know that Adams conditioning helps us solve the problem, it is not really a surprise that Jeffrey conditioning does: the former is just a special case of Jeffrey's rule; the only difference is that in Adams conditioning, the probability of one of the elements is "hardwired" to be invariant.

Adams vs. Jeffrey IV

Thus we can choose to. . .

- take the invariance of the probability of the antecedent as an explicit part of the input to the update rule, as for Jeffrey's rule. We may then derive the required constraint from the context of the example cases.
- take the invariance of the probability of the antecedent as implicit to the update rule itself. Based on the context we may then decide that Adams conditioning is applicable.

The difference between these two ways of updating is of little consequence. The boundary between criteria for applicability and input seems vague.

3 A distance function for Adams conditioning

The original approach to the Judy Benjamin problem had been to look for some distance minimization rule. While we can do without such a rule, we may ask whether there is one that yields the results of Adams conditioning.

It turns out that minimizing the *inverse relative entropy* distance exactly yields the required results.

$$\mathrm{IRE}(\mathrm{Pr}, \mathrm{Pr}') = \sum_i \mathrm{Pr}(U_i) \log \frac{\mathrm{Pr}(U_i)}{\mathrm{Pr}'(U_i)}$$

(RE and IRE are genuinely different functions; they are not symmetric.)

Remark
Note that IRE minimization is not just formally very close to RE minimization, but also conceptually: where the latter has you select the probability function that is RE-closest to your present probability function *as seen from your current perspective*, IRE minimization has you select the probability function that is RE-closest to your present probability function *as seen from the perspective you will have after adopting the probability function to be selected*.

4 Distance minimization generalized

In the cases of Sarah and Judy, learning the conditional was pre-theoretically irrelevant to the probability for the conditional's antecedent. These cases are covered by our proposal. But not all cases are like this:

A jeweller has been shot in his store and robbed of a golden watch. However, it is not clear at this point what the relation between these two events is; perhaps someone shot the jeweller and then someone else saw an opportunity to steal the watch. Kate thinks there is some chance that Henry is the robber (R). On the other hand, she strongly doubts that he is capable of shooting someone, and thus, that he is the shooter (S). Now the inspector, after hearing the testimonies of several witnesses, tells Kate: "If Henry robbed the jeweller, then he also shot him." As a result, Kate becomes more confident that Henry is *not* the robber, while her probability for Henry having shot the jeweller does not change.

More on Kate and Henry
Kate's case can be accommodated by a sort of counterpart of Adams conditioning that keeps the probability of the *consequent* invariant.

But what if the update leads to a conflict between the probabilities of the antecedent and consequent? Perhaps Kate cherishes the idea that Henry is not a murderer, but on the other hand she realizes that he was in need of some fast cash and might therefore well be the robber.

She must try to find a trade-off between maintaining a high probability of Henry's being the robber and maintaining a low probability of his having shot the jeweller, and she must do so under the constraint that he cannot have done the former without having done the latter.

A large class of distance functions

To accommodate a trade-off between antecedent and consequent, we may use a Hellinger distance and supplement it with weights. Where $\{Q_1, \ldots, Q_n\}$ are the strongest consistent propositions in the algebra, and $w_i \in \mathbb{R}^+ \cup \{\omega\}$ is the weight assigned to Q_i, the rule says that we ought to minimize the following function:

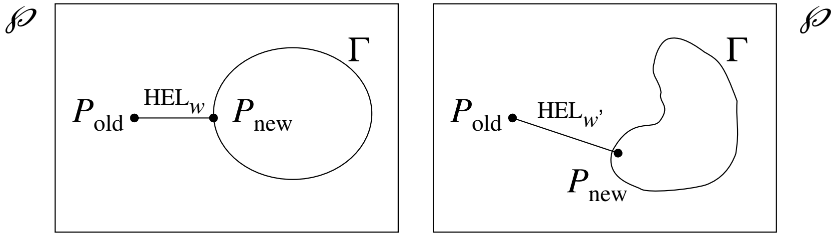
$$\mathrm{EE}(\mathrm{Pr}, \mathrm{Pr}') = \sum_{i=1}^{n} w_i \left(\sqrt{\mathrm{Pr}(Q_i)} - \sqrt{\mathrm{Pr}'(Q_i)}\right)^2 .$$

The higher w_i, the more resistance to deviations in the probability $P(Q_i)$. In other words, the weights indicate how much you hate to change the probability for a given cell of the partition. In still other words, the values of the weights express epistemic entrenchment.

Two remarks

The rule of EE minimization can be applied quite generally, also to the kind of conditionals for which we proposed the Adams type update. Indeed, that update rule is a limiting case of the EE rule.

The rule can be visualized as follows:



Numerical example

Setting the odds $\Pr_{\text{new}}(Q_2) : \Pr_{\text{new}}(Q_1)$ to $3 : 1$ and to $50 : 1$ respectively, setting $w_2 = w_4 = 1$, and varying $w = w_1 = w_2$, we obtain the following updated probability assignments.

		Probability			
Odds	Weight	$Q_1 = R \wedge S$	$Q_2 = R \wedge \neg S$	$Q_3 = \neg R \wedge S$	$Q_4 = \neg R \wedge \neg S$
-	-	0.10	0.70	0.10	0.10
3	1	0.53	0.18	0.15	0.15
	5	0.21	0.07	0.13	0.60
	100	0.10	0.03	0.10	0.76
50	1	0.47	0.01	0.26	0.26
	5	0.15	0.00	0.13	0.72
	100	0.10	0.00	0.10	0.79

Is more needed? I

The weights a person is supposed to assign to the relevant propositions will not come out of thin air but may be assumed to be interconnected with (even if presumably not fully determined by) things she believes; nor will these weights remain fixed once and for all but will, plausibly, themselves change in response to things the person learns.

Our proposal is silent on both of these issues. To properly address them, we may well have to go beyond our current representation of epistemic states in terms of degrees of belief plus weights.

Is more needed? II

But perhaps we do have to rely on our own judgment in assigning weights.

Something very similar is already the case for the kind of uncertain learning events that Jeffrey's rule was devised for: there is no rule telling us how a glimpse of a tablecloth in a poorly lit room is to change our assignment of probabilities to the various relevant propositions concerning the cloth's color.

The point may be more general still. Bradley [2005]: "[Even Bayes's rule] should not be thought of as a universal and mechanical rule of updating, but as a technique to be applied in the right circumstances, as a tool in what Jeffrey terms the 'art of judgment'." In the same way, determining and adapting the weights EE supposes may be an art, or a skill, rather than a matter of calculation or derivation from more fundamental epistemic principles.

5 Discussion

- Conditionalization on the material implication is not necessarily a good account of the learning of conditional information.
- Hence, the fact that relative entropy minimization affects the probability of the antecedent cannot be defended by reference to conditionalization on a material conditional.
- If we gather the constraints imposed by the Judy Benjamin story, they pin down a complete probability assignment over a partition, and we can apply Jeffrey's rule of updating.
- Alternatively, we can apply Adams conditioning, using an incomplete probability assignment over a partition as input. The further constraint then appears as a condition of applicability.

Discussion (continued)

- Minimizing the distance function IRE gives the same result as Adams conditionalization.
- We can define a whole class of distance functions, each of them associated with different epistemic entrenchments for the probabilities of the elements of the partition.
- In the face of this plethora of update rules, we note that picking the appropriate rule for a given occasion may be an art rather than a matter of appealing to a further rule or rules.

Thank you