Order Dependence and Jeffrey Conditionalization*

Daniel Osherson

March 2, 2002

Listening to the radio I hear a forecast for rain but I'm not sure whether it comes from the chief meteorologist or from his unreliable deputy. The following distribution P_1 of probability emerges in my mind.

(1)

state	probability
RC	.4
$R\bar{C}$	.1
$\bar{R}C$	.2
$\bar{R}\bar{C}$	.3

where

R = It rains today.
C = The chief was speaking.

So, $P_1(R) = .5$.

A glance at the sky raises my probability of rain to .7. As it happens, the conditional probabilities of each state given rain remain the same, and similarly for their conditional probabilities given no rain. As Jeffrey (1983, Ch. 11) points out, my new distribution P_2 is therefore fixed by the law of total probability. For example, $P_2(RC) = P_2(RC \mid R)P_2(R) + P_2(RC \mid \bar{R})P_2(\bar{R}) = P_1(RC \mid R)P_2(R) + P_1(RC \mid \bar{R})P_2(\bar{R}) = (.8)(.7) + (0)(.3) = .56$. Similar calculations fill in the rest of my new distribution as follows.

(2)

state	probability
RC	.56
$R\bar{C}$	.14
$\bar{R}C$	.12
$\bar{R}\bar{C}$	.18

Hence, $P_2(C) = .68$.

Now I hear another snippet from the radio which lowers my probability of the chief speaking to .2. Again, the conditional probabilities of states

*Thanks to Sherri Roush for very helpful comments. Correspondence to D. Osherson, MS-25, Rice University, P.O. Box 1892, Houston TX 77251-1892. Email: osherson@rice.edu.

given this event (or its absence) don't change. Jeffrey's rule shows my new distribution to be:

(3)

state	probability
RC	.165
$R\bar{C}$	.350
$\bar{R}C$	.035
$\bar{R}\bar{C}$	.450

Things could have happened in reverse order, of course. Suppose that I first heard a new snippet and lowered my probability of C to .2. Jeffrey's principle would lead from (1) to:

(4)

state	probability
RC	.133
$R\bar{C}$	.200
$\bar{R}C$	.066
$\bar{R}\bar{C}$	.6

If I next glanced at the sky and raised the probability of rain to .7 then (4) would give way to:

(5)

state	probability
RC	.280
$R\bar{C}$	.420
$\bar{R}C$	.030
$\bar{R}\bar{C}$	.270

Horror! Distributions (3) and (5) don't match. The mere order in which Jeffrey's rule was applied lands me in different epistemological positions. The rule has occasionally been questioned on this basis, most recently by Döring (1999) who calls such order effects "an embarrassment."[1]

The embarassment is relieved, however, by distinguishing between data and probability assessments. Let sky_{12} be the patch of sky I saw in the transition from P_1 to P_2, and let $snip_{23}$ be the snippet of radio I heard in the transition from P_2 to P_3. Likewise, let sky_{45} and $snip_{14}$ be their counterparts in the second story. If we could tell the tale in such a way that $sky_{12} = sky_{45}$ and $snip_{23} = snip_{14}$ are evident then there would be paradox in

[1] See also Field (1978) whose alternative proposal for belief dynamics is meant to avoid "asymmetries." Although symmetric, Field's method fails to be idempotent inasmuch as two exposures to the very same data can have a different effect than just one (not the case for Jeffrey's rule). See Garber (1980).

the air for two reasons. First, the order in which the two data are assimilated seems accidental so shouldn't matter. Second, the order *doesn't* matter if we conceive of the data's impact as issuing from conditionalization on each datum in turn, as described in Pearl (1988). (Conditionalization is not sensitive to order.) In contrast, if $sky_{12} \neq sky_{45}$ and $snip_{23} \neq snip_{14}$ are plausible then the difference between P_3 and P_5 can be attributed to the different data that led to each from P_1.

In fact, there is good reason *not* to believe in the equation of sky_{12} with sky_{45} and $snip_{23}$ with $snip_{14}$. When I experienced sky_{12}, my probability for rain was one-half whereas it was one-third when I experienced sky_{45}. Yet both patches led me to assign probability .7 to rain. This suggests that sky_{45} showed darker clouds than did sky_{12}. More formally, suppose that I were willing and able to incorporate sky_{12} and sky_{45} into my event-space. Then it would be reasonable for me to calculate the probabilities of rain via

$$P_2(R) = P_1(R \mid sky_{12}) = \frac{P_1(sky_{12} \mid R) \times P_1(R)}{P_1(sky_{12})} = \frac{P_1(sky_{12} \mid R) \times \frac{1}{2}}{P_1(sky_{12})}$$

and

$$P_5(R) = P_4(R \mid sky_{45}) = \frac{P_4(sky_{45} \mid R) \times P_4(R)}{P_4(sky_{45})} = \frac{P_4(sky_{45} \mid R) \times \frac{1}{3}}{P_4(sky_{45})}.$$

What can be said about $P_1(sky_{12} \mid R)/P_1(sky_{12})$ and $P_4(sky_{45} \mid R)/P_4(sky_{45})$? My long experience viewing the sky can be counted on to overwhelm the impact of a radio snippet on my prior probabilities for patches. Hence, for any given patch $\boldsymbol{p}$, $P_4(\boldsymbol{p}) \approx P_1(\boldsymbol{p})$. Likewise, the snippet will have minimal impact on the conditional probability of observing $\boldsymbol{p}$ given that it rains today, so $P_4(\boldsymbol{p} \mid R) \approx P_1(\boldsymbol{p} \mid R)$. If sky_{12} were identical to sky_{45}, we could therefore expect $P_1(sky_{12} \mid R)/P_1(sky_{12}) \approx P_4(sky_{45} \mid R)/P_4(sky_{45})$, hence $P_5(R) < P_2(R)$. But in fact, our stories set $P_5(R) = P_2(R) = .7$. Similar remarks apply to $snip_{23}$ and $snip_{14}$; if they were identical then $P_4(C) \neq P_3(C)$, contrary to fact.

None of this is undeniable. Someone might hold out, for example, against $P_4(\boldsymbol{p} \mid R) \approx P_1(\boldsymbol{p} \mid R)$. But our assumptions are certainly reasonable. So it is reasonable to suppose that that $sky_{12} \neq sky_{45}$ and $snip_{23} \neq snip_{14}$.

Of course, the whole point of Jeffrey's rule is to circumvent appeal to ephemera like sky-patches and radio-snippets as events that support conditioning. For, the rule allows distributions to evolve solely on the basis of changes to the probabilities of familiar events (provided the relevant conditional probabilities stay constant). Conditioning on sky_{12} and sky_{45} nonetheless remains an option for a rational agent willing to specify the needed likelihood functions and priors. The foregoing calculations thus show that $sky_{12} \neq sky_{45}$ and $snip_{23} \neq snip_{14}$ are plausible claims. Since they provide a sufficient explanation for the disparity between P_3 and P_5, Jeffrey's rule is not impugned.

References

Döring, F. (1999). Why Bayesian Psychology is Incomplete. *Philosophy of Science*, 66:S379–S389.

Garber, D. (1980). Field and Jeffrey Conditionalization. *Philosophy of Science*, 47:142 – 145.

Jeffrey, R. C. (1983). *The Logic of Decision (2nd Edition)*. Chicago IL, The University of Chicago Press.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo CA, Morgan Kaufmann.