A Note on Jeffrey Conditionalization

Author(s): Hartry Field
Reviewed work(s):
Source: *Philosophy of Science*, Vol. 45, No. 3 (Sep., 1978), pp. 361-367
Published by: The University of Chicago Press on behalf of the Philosophy of Science Association
Stable URL: http://www.jstor.org/stable/187023
Accessed: 22/08/2012 00:53

A NOTE ON JEFFREY CONDITIONALIZATION*

HARTRY FIELD

University of Southern California

Bayesian decision theory can be viewed as the core of psychological theory for idealized agents. To get a complete psychological theory for such agents, you have to supplement it with input and output laws. On a Bayesian theory that employs strict conditionalization, the input laws are easy to give. On a Bayesian theory that employs Jeffrey conditionalization, there appears to be a considerable problem with giving the input laws. However, Jeffrey conditionalization can be reformulated so that the problem disappears, and in fact the reformulated version is more natural and easier to work with on independent grounds.

Bayesian decision theory can be viewed as the core of a psychological theory for certain idealized agents. It is only the *core* of a psychological theory: it must be supplemented by input and output laws. The Bayesian law of conditionalization is not itself an input law: it says that if P is a person's belief function at time t, then the person's belief function P' at time t' will be given by

$$(1)\quad P'(A) = P(A \wedge E_1 \wedge \ldots \wedge E_n)/P(E_1 \wedge \ldots \wedge E_n)$$

for *some* sentences $E_1, \ldots, E_n$; but it does not say *which* sentences $E_1, \ldots, E_n$ the change in belief function will originate from. Intuitively, Bayesians think of the change in belief function as a result of the agent's observations between t and t'; intuitively, $E_1, \ldots, E_n$ are the sentences *directly affected* by these observations. What we need, then, is a law that tells us which sentences in an agent's language will be directly affected by which kinds of sensory stimulation. Such a law might be given by characterizing the sentences in the agent's language that count as observation sentences, and defining a function that assigns to each observation sentence a class of possible sensory stimulations that would directly affect it.[1] (We would have to demand that no sensory stimulation directly affect more than finitely many observation sentences, if the form of (1) is to be preserved.) In effect,

*Received August, 1977.

[1]This would give a deterministic input law; non-deterministic ones are also imaginable, of course, but will not be considered here.

[2]Several changes *each of which directly affect only one observation sentence*, that is.

Philosophy of Science, 45 (1978) pp. 361–367.

then, such a law would correlate with each possible sensory stimulation a finite set of observation sentences, viz. those that the stimulation would directly affect.

The Bayesian conditionalization law (1) is very much of an idealization: a consequence of it is that after the observation, the directly affected sentences $E_1, \ldots, E_n$ will have a probability 1. (And they will never lose this probability, if beliefs change *only* by conditionalization.) It is natural to try to reduce the level of idealization: we might keep the idea that observation *directly affects* certain sentences, but give up the idea that the directly affected sentences have to acquire a probability of 1. In trying to work out this idea we must remember that *there are two laws that need to be generalized: we need to generalize not only the law of conditionalization (1), but also the input law.* This raises some questions about Richard Jeffrey's method of generalizing (1) in Ch. 11 of *The Logic of Decision* ([1]).

Before looking at the problem of generalizing (1), let's look at the problem of generalizing the input law. Our earlier input law correlated with each possible sensory stimulation a finite set of observation sentences, viz. those that the stimulation would directly affect. So it is natural to think that the generalized input law would correlate with each possible sensory stimulation both a (finite) set of observation sentences that that stimulation directly affects, and also an assignment to each such observation sentence of a number that represents the degree to which the stimulation affects it. (Call this number the *input parameter* associated with the observation sentence.) This way of generalizing the input law imposes constraints on the generalization of (1). For the new probability function should be determined by the old one together with a list of which observation sentences have been directly affected and how much each one has been affected; and how much each one has been affected must be measured by the input parameter, in order to make use of the generalized input law. In other words, the generalization of (1) should (in the case where $n = 1$, i.e. where only one observation sentence is directly affected) be of the form

(2) $P' = \Phi(P, E, \alpha)$.

where E is the observation sentence directly affected and α is the input parameter representing the degree to which E has been affected.

Jeffrey has provided a generalization of (1); but it does not have the form (2), and this raises a serious question as to how to supplement it with an input law. Jeffrey's generalization (again restricting our attention to the case where $n = 1$) is

$$(3)\ P'(A) = qP(A \wedge E)/P(E) + 1 - q)P(A \wedge {\sim}E)/P({\sim}E)$$

for some sentence E and some number q in the interval [0,1]. From (3) it follows (given the laws of probability) that the probability of E after the observation is q. This means that q cannot be an input parameter, given the definition of 'input parameter' above: for it is clear that the probability q which I attach to an observation sentence E after I have been subjected to a sensory stimulation will depend not only on the sensory stimulation but also on the probability I attached to E before the stimulation. Because of this, no simple input law of the sort envisaged before could relate the value of q to the kind of sensory stimulation imposed. Instead, the input law would have to be more complicated: it would have to represent q as a function of the prior probability p of E as well as of the sensory stimulation that the agent was subjected to. In fact, it is hard to see how to come up with an input law of this more complicated form except by introducing a number α that represents the effects that the sensory stimulation has *by itself* (independently of the value of p), and then representing q as a function of α and p together. If we did that—if we introduced an input parameter α and found a function ψ such that $q = \psi(\alpha, P(E))$—then by combining this last equation with Jeffrey's law (3) we would get a law of the form (2). In other words, the only obvious way to supplement Jeffrey's law (3) with an input law is to reparameterize it, using an input parameter in place of q. And Jeffrey has not shown us how to do that.

What is Jeffrey's attitude towards this problem? I am not sure. In "Carnap's Empiricism" ([2]) Jeffrey quotes an exchange of letters in the late 1950's between himself and Carnap, in which Carnap makes a point similar to the point made above. Carnap puts the point by saying that we ought to generalize the process of conditionalization by starting from the evidence sentence E together with a number α which "is to indicate the subjective certainty of the sentence on the basis of the observational experience" ([2], p. 42); and he points out that Jeffrey's parameter does not serve the role of α, since it depends on the prior probability of E. As I have said, this is similar to the point I have made above; but there are several differences. First of all, Carnap speaks of α as if it, like q, could be regarded as a probability; this seems to me to be quite unwarranted, and I suspect that it was because of this unwarranted assumption that Carnap "very soon . . . found some difficulties [in his attempt to formulate a law of form (2)] and did not know how to overcome them" ([2], p. 42). A second difference is that Carnap nowhere explicitly mentions the problem on which I have focused, the problem of giving a complete psychological theory for a Bayesian agent; and he nowhere explicitly mentions what I have called input laws. Carnap's criticism of Jeffrey is that there are cases where a person ought *if he is rational* to come to attach a high probability q to a directly affected sentence E, but that nothing in Jeffrey's constraints requires him to do so. I do not think that this way of putting the problem makes the problem very persuasive. In any case, it did not persuade Jeffrey; I hope that my own presentation, in terms of the problem of giving a complete psychological theory, makes the point more compelling.

It should be clear from my remarks above that the need of generalizing the input law as well as the law of conditionalization does not force us to conclude that Jeffrey's generalization (3) is incorrect in the sense that it leads to the wrong transformations of belief functions; and I do not so conclude. What I *do* conclude is only that we need to re-parameterize Jeffrey's transformation law, using a parameter α that can be viewed as in some sense a function of sensory stimulation alone. In the rest of this note I will describe a re-parameterization that I think does the job. In order to argue that it does in fact do the job one would have to argue that the parameter α that it involves is indeed an input parameter, and I am not sure how to argue this in any detail. (I will present one piece of evidence for the claim, however.) What I *will* do is to show that the reformulation I advocate has advantages over Jeffrey's formulation even independently of the input problem.

The parameter that I would use to redescribe transformation (3) is

$$(4)\ \alpha =_{\mathrm{df}} (1/2) \log ((q/p)/((1-q)/(1-p)))$$

where p is $P(E)$ and q is $P'(E)$. (The (1/2) log is of course just in there for reasons of scale; the really essential thing is the fraction that follows them.) Solving for q in terms of α, we get

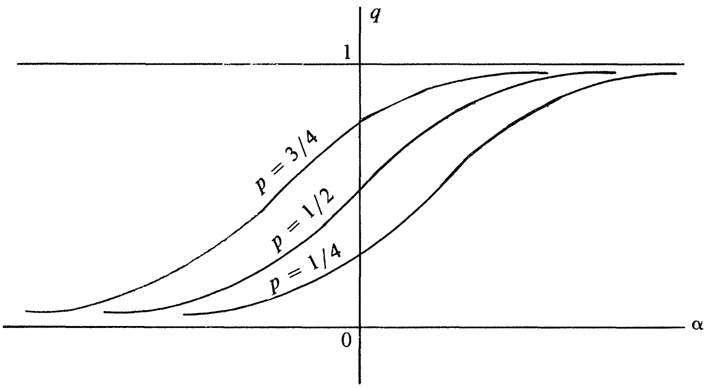
$$(5)\ q = (pe^{\alpha})/(pe^{\alpha} + (1-p)e^{-\alpha})$$

This gives q as a function ψ of α and of p, as required several pages back. The graph at the top of page 365 shows what q looks like as a function of α, for several fixed values of p.

Qualitatively, this is pretty much what we'd expect if α were an input parameter.

If we substitute the formula (5) for q into (3), we get the desired reformulation of Jeffrey's generalization of conditionalization:

$$(6)\ P'(A) = (e^{\alpha} P(A \wedge E) + e^{-\alpha} P(A \wedge {\sim}E))/(e^{\alpha} P(E) + e^{-\alpha} P({\sim}E))$$



Note that (by the definition (4)) α can be any real number. Note also that if $\alpha = 0$, (b) gives us that $P'(A) = P(A)$; in other words, the probability function does not change, the sensory stimulation was completely uninformative. The corresponding value of the Jeffrey parameter q in the case of an uninformative sensory stimulation is p. The fact that the criterion of a sensory stimulation being uninformative is expressed in the reformulation by the condition that $\alpha = 0$ (rather than by any condition that involves p) seems to me to be some evidence that α is the desired input parameter.

Strict conditionalization fits into the formulation (6) as a limit: as α approaches $+\infty$, $P'(A)$ approaches $P(A \wedge E)/P(E)$; as α approaches $-\infty$, $P'(A)$ approaches $P(A \wedge \sim E)/P(\sim E)$. In Jeffrey's parameterization, strict conditionalization is a special case rather than a limiting case; it corresponds to the values $q = 1$ and $q = 0$ respectively. I suspect that the fact that strict conditionalization is not a special case of (6) is no loss—I suspect that strict conditionalization *should* be regarded as an oversimplification that can't ever really arise—but if you want to allow it, you can allow change to occur by (1) as well as by (6).

What I now want to show is that formulation (6) of Jeffrey's law (3) has some very nice properties, independent of the question of input laws. Consider what happens when we change our probability functions *several times in succession*. Suppose that first P changes to P', as a result of a sensory stimulation that directly affects E; then P' changes to P'', as a result of another stimulation that directly affects E'. If we express the combined result of both changes in terms of the Jeffrey parameters q and q' of the individual changes, we get a very complicated law; moreover, it is an *asymmetric* law, in that the simultaneous interchange of E with E' and q with q' very much affects the result. If however we express the combined result in terms of the parameters α and α' of the individual changes, we get a law that is both simple and symmetric:

$$(7)\ P''(A) = (e^{\alpha+\alpha'}P(A \wedge E \wedge E') + e^{\alpha-\alpha'}P(A \wedge E \wedge \sim E') + e^{-\alpha+\alpha'}P(A \wedge \sim E \wedge E') + e^{-\alpha-\alpha'}P(A \wedge \sim E \wedge \sim E'))/ (e^{\alpha+\alpha'}P(E \wedge E') + e^{\alpha-\alpha'}P(E \wedge \sim E') + e^{-\alpha+\alpha'}P(\sim E \wedge E') + e^{-\alpha-\alpha'}P(\sim E \wedge \sim E'))$$

The increase in simplicity is even more dramatic for sequences of more than two changes. Thus the α parameter is much easier to work with than is Jeffrey's parameter q.

The reformulation I have given can be extended to the case of sensory stimulations that directly affect more than one observation sentence. In this case Jeffrey's law of change is more complicated than (3). Let $E_1, \ldots, E_n$ be the observation sentences directly affected; and let $F_1, \ldots, F_k$ (where $k = 2^n$) be the sentences of the form $G_1 \wedge \ldots \wedge G_n$, where each G_i is either E_i or $\sim E_i$. Then Jeffrey's generalization of (3) is as follows:

$$(3')\ P'(A) = \Sigma q_i P(A \wedge F_i)/P(F_i),$$

where each q_i is in (0,1) and $\Sigma q_i = 1$. The analogous generalization of (6) is

$$(6')\ P'(A) = \Sigma e^{\alpha_i} P(A \wedge F_i)/\Sigma e^{\alpha_i} P(F_i)$$

where each α_i is a real number and $\Sigma \alpha_i = 0$. To see this, generalize (4) as

$$(4')\ \alpha_i =_{df} (1/k) \log \prod_j ((q_i/p_i)/(q_j/p_j))$$

or equivalently,

$$\alpha_i = \log ((q_i/p_i)/(\Pi (q_j/p_j))^{1/k}).$$

(Here p_i is $P(F_i)$ and q_i is $P'(F_i)$). It follows that

$$\Sigma \alpha_i = \sum_i \log ((q_i/p_i)/(\Pi (q_j/p_j))^{1/k})$$
$$= \log \prod_i ((q_i/p_i)/(\Pi (q_j/p_j))^{1/k})$$
$$= \log ((\Pi (q_i/p_i))/(\Pi (q_j/p_j))) = \log 1 = 0,$$

as desired. It also follows that

$$e^{\alpha_i}/e^{\alpha_j} = (q_i/p_i)/(q_j/p_j),$$

so that

$$q_i/q_j = p_i e^{\alpha_i}/p_j e^{\alpha_j}$$

Therefore $\sum_i (q_i/q_j) = \sum_i (p_i e^{\alpha_i}/p_j e^{\alpha_j})$. Since $\Sigma q_i = 1$,

this gives us

$$(5') \quad q_j = p_j e^{\alpha_j}/\Sigma p_i e^{\alpha_i}$$

Substituting this into (3′), we get (6′), as desired.

Note that the law (7) for successive changes[2] is just like the law (6′) for changes that directly affect several E_i's at once, except that for successive changes there is a restriction on the form that the α_i's can take: here the 2^n α_i's meet an independence condition which says that there are n numbers $\alpha, \alpha', \ldots, \alpha^{(n)}$ such that the α_i's associated with $E_1 \wedge \ldots \wedge E_n$, $E_1 \wedge \ldots \wedge \sim E_n$, etc. are $\alpha + \alpha' + \ldots + \alpha^{(n-1)} + \alpha^{(n)}$, $\alpha + \alpha' + \ldots + \alpha^{(n-1)} - \alpha^{(n)}$, etc..

REFERENCES

[1] Jeffrey, R. *The Logic of Decision*. New York: McGraw-Hill, 1965.
[2] Jeffrey, R. "Carnap's Empiricism." In Maxwell and Anderson (eds.), *Minnesota Studies in the Philosophy of Science*, vol VI. Minneapolis: University of Minnesota Press, 1975.