1. INTRODUCTION

Scientific progress often depends on the introduction of theories that have not previously been entertained. Indeed, practically every currently accepted scientific theory was first introduced at some period in recorded history. Scientific education recapitulates this development in the life of the individual, students being introduced to theories they had never previously contemplated.

Some people seem to think that Bayesian theory requires a person to give probabilities to every proposition that could ever be formulated. If that were true then new theories would never arise for Bayesian agents; the fact that new theories do arise for real agents would then be an embarrassment for Bayesian theory. However, the thought that gives rise to this putative embarrassment is mistaken; Bayesian theory does not require a person to give probabilities to every proposition that can be formulated. Standard Bayesian theory assumes that the propositions to which a person gives probabilities form an algebra, that is, a set of propositions that is closed under union and complementation; it is not required that this algebra include every proposition that can be formulated.

Nevertheless, new theories do present Bayesians with two genuine problems.

One problem is how to determine the probability that should be assigned to a new theory. If we suppose that the probabilities of existing theories should not change merely due to the introduction of a new theory, then the probability of a newly introduced theory must be somewhere between 0 and $r$, where $r$ is the prior probability that none of the existing theories is true. But how is one to choose the probability in that
interval? To say that anything is permissible does not seem reasonable, since some newly introduced theories are more plausible than others; but how can a Bayesian give contentful advice here?

It has been claimed that the introduction of new theories changes scientists' probabilities for existing theories. This raises the second problem: Can a Bayesian say what the probabilities of existing theories should be after a new theory is introduced?

Earman has argued that the paradigm Bayesian theory of learning, namely conditionalization, cannot solve these problems.

Conditionalizing (in any recognizable sense of the term) on the information that just how a heretofore unarticulated theory $T$ has been introduced is literally nonsensical, for such a conditionalization presupposes that prior to this time there was a well-defined probability for this information and thus for $T$ (Earman 1992, p. 196).

Earman goes on to argue that shifts in probability, induced by the recognition of new possibilities, are ubiquitous even in daily life; he concludes that violations of conditionalization must be allowed to occur regularly. The result is that Bayesian learning theory is essentially vacuous: “All that remains of Bayesianism in its present form is the demand that new degrees of belief be distributed in conformity with the probability axioms” (p. 198).

The purpose of this paper is to dispel such pessimism. I will show that Bayesian theory “in its present form” contains natural solutions to the two problems I have posed. Thus I will show that Bayesian theory does give guidance regarding the probability that should be assigned to a new theory, and regarding the probabilities that the old theories should have after a new theory is introduced.

2. CONDITIONALIZATION

It should be clear that there are ways of referring to theories that have not yet been formulated. In fact, I have already referred several times to the class of such theories. Individual elements in that class can also be referenced in various ways. One way is via the temporal order in which the theories are proposed; thus we can speak of the next theory that is proposed in some area, the one that is proposed after that, and so on.
Since we can refer to not-yet-formulated theories, we can also have preferences regarding bets on them. For example, I might prefer the option of receiving a dollar if the next theory formulated is false, to the option of receiving the dollar if that theory is true. It follows that we can have subjective probabilities for theories that have not yet been formulated.  

We can also talk about the properties that a not-yet-formulated theory may have. For example, the claim that the next theory to be formulated will be very simple is one that we can understand before the theory is formulated. And again, we can have subjective probabilities for propositions of this sort. For example, I might think it unlikely that the next theory to be formulated will be simpler than any existing theory.

Suppose, then, that $A_1, A_2, \ldots, A_k$ are the theories on some subject that have been formulated to date, that $A_{k+1}$ denotes the next theory that will be formulated, $A_{k+2}$ the one after that, and so on. Suppose that there is a measure $m(A_i)$ of the theoretical virtues of $A_i$ that we regard as relevant to its prior probability. For example, $m(A_i)$ might be taken to be a measure of the simplicity of $A_i$. Assume that once $A_i$ has been formulated we can determine the value of $m(A_i)$ with certainty; prior to that there will be uncertainty about the value of $m(A_i)$. I am not suggesting that these assumptions are realistic, but merely make them in order to be able to give a relatively concrete example; they will be relaxed in the following section.

Let the $A$-algebra be the smallest algebra that contains all $A_i$ and all propositions of the form $m(A_i) = r$ for some real number $r$. We can assume that the propositions in the $A$-algebra all have a subjective probability, even though $A_{k+1}, A_{k+2}, \ldots, A_{k+2}$, have not yet been formulated.

Now suppose that $A_{k+1}$ is formulated, and that the value of $m(A_{k+1})$ turns out to be $r$. Let $p$ be my probability function before $A_{k+1}$ was formulated, and let $q$ be the probability function I have afterwards. It may be the case that, on the $A$-algebra, $q(\cdot) = p(\cdot | m(A_{k+1}) = r)$; that is to say, this identity holds if the `$\cdot$' is replaced by any proposition in the $A$-algebra. If that is so, then my posterior probabilities on the $A$-algebra are just my prior probabilities conditioned on $m(A_{k+1}) = r$. 
For this shift to accord with the principle of conditionalization, it needs to be the case that \(m(A_{k+1}) = r\) is all that was learned when \(A_{k+1}\) was formulated. According to the proposal I made in Maher (1993, p. 121), this means that if \(R_q\) is the proposition that I come to have probability function \(q\), then \(p(\cdot | R_q) = p(\cdot | m(A_{k+1}) = r)\). We can perfectly well suppose that this identity holds for every proposition in the \(A\)-algebra; and then, so far as the \(A\)-algebra is concerned, \(m(A_{k+1}) = r\) does capture all that I learned when shifting from \(p\) to \(q\). Thus it is possible for the shift in probabilities on the \(A\)-algebra to accord with the principle of conditionalization.

As an example of how such conditionalization could proceed, suppose that for all real numbers \(r_j\) I have

\[
p(A_i | \cap_j [m(A_j) = r_j]) = \frac{\sum r_i}{r_j}.
\]

This implies that

\[
(1) \quad p(A_i) = E_p \frac{m(A_i)}{\sum_j m(A_j)}
\]

where \(E_p\) is expected value calculated using \(p\). It also implies that

\[
(2) \quad p(A_i | m(A_{k+1}) = r) = E_q \frac{m(A_i)}{\sum_j m(A_j)}
\]

where \(q\) is the probability function obtained from \(p\) by conditioning on \(m(A_{k+1}) = r\).

A very simple numerical illustration: Suppose that \(A_1\) is the only theory that has been formulated, and that \(m(A_1) = 5\). I am sure that the true theory is either \(A_1\), \(A_2\), or \(A_3\). Since \(A_2\) and \(A_3\) have not yet been formulated I do not yet known \(m(A_2)\) or \(m(A_3)\), but I have

\[
\begin{align*}
p[m(A_2) = 1 \text{ and } m(A_3) = 1] &= 0.7 \\
p[m(A_2) = 10 \text{ and } m(A_3) = 1] &= 0.2 \\
p[m(A_2) = 1 \text{ and } m(A_3) = 10] &= 0.09 \\
p[m(A_2) = 10 \text{ and } m(A_3) = 10] &= 0.01.
\end{align*}
\]
PROBABILITIES FOR NEW THEORIES

Then (1) gives
\[
p(A_i) = \frac{5}{5+1+1}(0.7) + \frac{5}{5+10+1}(0.2) + \frac{5}{5+1+10}(0.09) + \frac{5}{5+10+10}(0.01)
\]
= 0.59, to 2 decimal places.

Similar calculations for \(A_2\) and \(A_3\) give
\[
p(A_2) = 0.23; p(A_3) = 0.18.
\]

Now suppose \(A_2\) is formulated and I find that \(m(A_2) = 10\). Letting \(q\) denote \(p\) conditioned on the information that \(m(A_2) = 10\), we have
\[
q[m(A_2) = 10 \text{ and } m(A_3) = 1] = \frac{0.2}{0.2+0.01} = 0.95
\]
\[
q[m(A_2) = 10 \text{ and } m(A_3) = 10] = \frac{0.2}{0.2+0.01} = 0.05.
\]

Then by (2) we obtain
\[
q(A_1) = 0.31; q(A_2) = 0.61; q(A_3) = 0.08.
\]

Note that the probability of the existing theory \(A_1\) has changed due to the discovery of the new theory \(A_2\). The probability of \(A_2\) has increased because this theory turned out to be simpler than expected, and this has caused the probability of \(A_1\) and \(A_3\) to decrease. If \(m(A_2)\) had been 1 then the probability of \(A_1\) and \(A_3\) would have increased.

I have now shown, both by general argument and by example, that the probabilities of propositions in the \(A\)-algebra can be updated by conditionalization. But we still have to deal with the fact that I have also extended my probabilities to include a proposition that was not in the \(A\)-algebra, namely the proposition that expresses the content of the theory \(A_{k+1}\). Let \(B_{k+1}\) be this proposition and let the \(B\)-algebra be the smallest algebra containing the \(A\)-algebra together with \(B_{k+1}\). The probability function \(q\) that I come to have after \(B_{k+1}\) is introduced is defined for all propositions in the \(B\)-algebra (we can assume), though \(p\) is not defined for some of these propositions, including \(B_{k+1}\). So clearly \(q\) on the \(B\)-algebra is not obtained from \(p\) by conditionalization. However, once \(B_{k+1}\) has been formulated, I have learned that \(A_{k+1} = \ldots\)
$B_{k+1}$, so that in the $B$-algebra $B_{k+1}$ and $A_{k+1}$ differ only on a set of $q$-probability 0. Thus the probability of any proposition in the $B$-algebra must be unchanged if all occurrences in it of $B_{k+1}$ are replaced by occurrences of $A_{k+1}$; and the proposition obtained by this substitution is in the $A$-algebra. So *probabilities on the $B$-algebra are fixed by the probabilities on the $A$-algebra*. Since the probabilities on the $A$-algebra are updated by conditionalization, it follows that even the probabilities on the $B$-algebra are determined by conditionalization.

I will conclude this section by reviewing the argument of Earman's that was quoted in the preceding section. It may be set out more explicitly as follows.

1. One can condition on the information that $T$ was introduced only if, before $T$ was introduced, one has a probability for the proposition that $T$ would be introduced.

2. One can have a probability for the proposition that $T$ would be introduced only if one knows what $T$ asserts.

3. It is impossible to know what $T$ asserts before it is introduced.

$\therefore$ 4. It is impossible to condition on the information that $T$ was introduced.

$\therefore$ 5. Changes in probability due to the introduction of $T$ cannot be by conditionalization.

If $T$ is a designation of the theory that I can use before the theory is introduced, such as $A_{k+1}$, then premise (2) is false. If $T$ is a statement of the content of the theory – $B_{k+1}$ in my discussion – then the premises of the argument are correct and (4) follows validly from them. However, (5) does not follow, as my discussion and example show. Changes in probability on the $A$-algebra can be by conditionalization. Further, while the probabilities of propositions that appear only in the $B$-algebra cannot be obtained by conditionalization alone, they are fixed by conditionalization on the $A$-algebra together with the learning experience that prompts the extension of the probability function to the $B$-algebra.
The use of conditionalization in the preceding section depended on having a function \( m \) that completely specifies the factors deemed relevant to the prior probability of a theory. While inductive logicians have attempted to define such a function, I share the widespread view that no proposal made to date is adequate.\(^4\) Thus while the preceding section demonstrates the theoretical possibility of updating by conditionalization when a new theory is introduced, I do not think the method used there is one that we can apply in the real world.

Is there some other way that conditionalization could be used to update probabilities when a new theory is introduced? Skyrms (1980) argued that a wide range of learning situations can be brought under the conditionalization model. This is done by supposing that the learning situation fixes the new probabilities for the elements of some partition \( C_1, C_2, \ldots, \) and that the probabilities of the other propositions are obtained by conditioning on the value of the new probabilities of the \( C_i \). That is to say, one conditions on propositions of the form \( \cap_i [q(C_i) = r_i] \).

If we were to try to use this device to deal with the introduction of new theories, I suspect that the \( C_i \) would need to specify truth values for the theories \( A_1, A_2, \ldots \). But the new probabilities of the \( C_i \) are not obtained by conditionalization, since they must be fixed before conditionalization can be applied. Thus Skyrms's device does not provide support for the view that when a new theory is introduced, any revisions of the probability of it and competing theories will be by conditionalization.

I therefore believe that we need to allow for a more general model than conditionalization if we are to deal with realistic situations in which a new theory is introduced. The standard Bayesian move at this point is to invoke Jeffrey's (1965) probability kinematics; however, in the present situation probability kinematics faces the same difficulty noted in the preceding paragraph. Probability kinematics begins by supposing there is a partition \( C_1, C_2, \ldots, \), whose elements have their probability fixed by the learning experience, and it says that the new probability function \( q \) should satisfy

\[
q(\cdot) = \sum_i p(\cdot | C_i) q(C_i).
\]
But it seems likely that if the learning experience is the discovery of a new theory, then the \( C_i \) would need to specify truth values for this theory and its competitors. In that case, (3) tells us nothing about how to assign probabilities to these theories, other than that it should be done coherently. If this were all that could be said, Earman's pessimism would be vindicated.

Fortunately, we can do better. The key is go back to first principles. The fundamental Bayesian principle of maximizing expected utility says that if one's probabilities and utilities are rational, then an act is rational just in case it maximizes expected utility. An "act" here is any alternative that is subject to normative evaluation; it need not be directly subject to the will (Maher 1993, sec. 6.3.3). Use of a method for revising probabilities when a new theory is introduced is an act in this sense. Thus the fundamental Bayesian rule for revising probabilities when a new theory is introduced is to use a method that maximizes expected utility.

Let \( a \) be a method of revising probabilities when a new theory is introduced. I may not be able to describe \( a \) more precisely than to say that it consists of deliberating in a certain way and adopting the probability function that then comes to seem reasonable. Still, we can calculate the expected utility of \( a \). Suppose that after the new theory is introduced I will be faced with a further decision that involves choosing an act from the set \( B \). Suppose that there is a unique act \( b_q \) that I will choose from \( B \) if my probability function at that time is \( q \). Let \( S \) be a set of states that determine the utility of any act \( b \in B \), and let these states be causally independent of \( a \). Let \( Q \) be the set of probability functions that I might come to have after the new theory is introduced. Suppose I am sure there is a fact about what \( q \in Q \) I will acquire if I choose \( a \); that is, I am sure that for some \( q \in Q \), the counterfactual conditional \( a \rightarrow R_q \) is true. I will use the notations \( s \cap a \rightarrow R_q \) to denote that state \( s \) and counterfactual conditional \( a \rightarrow R_q \) both obtain. Propositions of the form \( s \cap a \rightarrow R_q \) are causally independent of \( a \). I will assume that \( a \) does not influence utility except via its influence on \( q \) and thereby on the choice made from \( B \); in that case propositions of the form \( s \cap a \rightarrow R_q \) determine the utility that will be obtained from choosing \( a \). Thus the expected utility of \( a \) can be calculated using as states the partition of
propositions of the form \( s \cap a \rightarrow R_q \), giving

\[
\mathcal{E}(a) = \sum_{s \in S} \sum_{q \in Q} p(s \cap a \rightarrow R_q)u(s \cap b_q).
\]

As a simple example of how this formula can be applied, I will rework the example of the preceding section, dispensing with the function \( m \) that there allowed conditionalization to be used. As before, \( A_1 \) is the only theory that has been formulated, and I am sure that the true theory is either \( A_1, A_2, \) or \( A_3 \). My probabilities for these theories are

\[
p(A_1) = 0.59; \quad p(A_2) = 0.23; \quad p(A_3) = 0.18.
\]

Let \( a \) be some method of revising probabilities when a new theory is introduced. For example, \( a \) might be the method of just reflecting directly on what probabilities now seem reasonable; or it might be some more structured method. Suppose I am sure that there are only two probability functions that I could acquire by using \( a \) when \( A_2 \) is introduced; these are \( q_1 \) and \( q_2 \), with

\[
q_1(A_1) = 0.31; \quad q_1(A_2) = 0.61; \quad q_1(A_3) = 0.08;
\]

\[
q_2(A_1) = 0.67; \quad q_2(A_2) = 0.13; \quad q_2(A_3) = 0.20.
\]

Suppose that after \( A_2 \) is introduced I will be offered an even-money bet on \( A_2 \) for $1. Then \( B = \{ b_1, b_2 \} \), where \( b_1 \) is acceptance of the bet and gives $1 if \( A_2 \) and \(-$1 otherwise, while \( b_2 \) is rejection of the bet and is certain to leave my wealth unchanged. We can then take the set \( S \) of states to be \( \{ A_2, \hat{A}_2 \} \), since these determine the outcome of the bet. Then the formula of the preceding paragraph gives that

\[
\mathcal{E}(a) = p(A_2 \cap a \rightarrow R_{q_1})u($1) + p(A_2 \cap a \rightarrow R_{q_2})u($0) + p(\hat{A}_2 \cap a \rightarrow R_{q_1})u($-1) + p(\hat{A}_2 \cap a \rightarrow R_{q_2})u($0).
\]

Suppose utilities are linear with money, so that we can set \( u($1) = 1,\)
\( u($0) = 0, \) and \( u($-1) = -1. \) Suppose also that

\[
p(A_2 \cap a \rightarrow R_{q_1}) = 0.13; \quad p(\hat{A}_2 \cap a \rightarrow R_{q_1}) = 0.08.
\]
Substituting these values in the equation for $E(a)$ gives

$$E(a) = 0.13 - 0.08 = 0.05.$$  

That is, the expected value of $a$ is 5 cents. It is small because I will bet only if I come to have probability function $q_1$, and the probability of that is only 0.21; furthermore, if I do bet, there is still a probability of $0.08/0.21 = 0.38$ of losing the bet.

So it makes sense to talk about the expected utility of different methods of revising probabilities when a new theory is introduced. Furthermore, this approach will work even in purely cognitive contexts, where no monetary bets or other practical applications are envisioned; for the decision to accept or reject a theory can itself be regarded as a kind of bet, albeit one with cognitive rather than pragmatic utility (Maher 1993, ch. 6).

I will now say something about the sorts of methods of revising probability that maximize expected utility.

In Maher (1993, pp. 116–120) I showed that, other things being equal, expected utility is maximized by using a method that is sure to produce a shift that satisfies Reflection. While I was not there considering the introduction of new theories, the result carries over into the present context. So provided other things are equal, a rational method for revising probabilities when a new theory is introduced will be a method that is sure to produce a shift that satisfies Reflection. The following examples illustrate the application of this result.

Suppose that I determine my new probabilities after a theory is introduced by a certain process of deliberation. If $q$ is a probability function that I might settle on after this deliberation, and if I know that I have engaged in such deliberation after I have done it, then $q$ will give probability 1 to the proposition that I have engaged in such deliberation. If I trust the deliberate process, the shift will then satisfy Reflection. So updating probabilities by a suitable process of deliberation is consistent with the principle of maximizing expected utility.

Alternatively, I might forgo deliberation and assign a probability to the new theory using a random number generator, adjusting other probabilities as necessary to maintain coherence. If $q'$ is a probability
function that I might acquire in this way, then \( p(A_{k+1} | R_q) = p(A_{k+1}) \), which will not in general equal \( q'(A_{k+1}) \). So this method produces shifts that violate Reflection and will not in general maximize expected utility.

Producing shifts that satisfy Reflection is – when other things are equal – only a necessary condition for the rationality of a method, not a sufficient condition. It is therefore important to consider other factors that influence the expected utility of using a method. Of these other factors, the main one to be noted here is that the expected utility of a method of revising probabilities tends to become greater as \( Q \), the set of probability functions that could result from \( A \), becomes larger and more diverse. In particular, if \( A_1 \) could result in any one of several probability functions, and \( A_2 \) necessarily results in a single probability function, then \( \mathcal{E}(A_1) \geq \mathcal{E}(A_2) \), provided both methods are sure not to produce a shift that violates Reflection. This is proved in Maher (1990).

Deliberating longer and harder would normally have the potential to produce a larger and more diverse set of posterior probability functions than more superficial deliberation. So one application of the result in the preceding paragraph is that longer and harder deliberation generally is worth more than superficial deliberation. Of course, we need to balance this value against the additional costs involved.

So far I have been talking about deliberation in a very abstract way. Partly that is because reasonable methods of deliberation are hard to specify precisely. However, it may be worth remarking on a method Bayesians often use when a new theory is introduced: One imagines what probability one would have given to the new and existing theories if all these theories had been formulated before the evidence that is used to discriminate between them was known; one then updates these counterfactual prior probabilities on that evidence, thus arriving at a current probability function. If the resulting probability function seems reasonable in the light of this calculation it is adopted; if not, one reconsiders the judgments that led to it until, hopefully, one achieves a state of reflective equilibrium.

It has often been noted that Bayesians make use of considerations of counterfactual priors for newly introduced theories. I think this has sometimes been interpreted as showing that Bayesians deny the reality of newly introduced theories. On the approach I am outlining here, that
is a misinterpretation of the purpose of this Bayesian device. It is not an attempt to deny reality, but simply a method for revising probabilities which we feel produces good results; i.e., it can produce a wide variety of posterior probability functions, and the shift to any one of them satisfies Reflection.

4. CONCLUSION

Contrary to what has been widely supposed, Bayesian theory deals successfully with the introduction of new theories that have never previously been entertained. The theory enables us to say what sorts of method should be used to assign probabilities to these new theories, and it allows that the probabilities of existing theories may be modified as a result.6

NOTES

1 I treat propositions as sets of states, so that set-theoretic union and complementation correspond to the logical operations of disjunction and negation, respectively.
2 Chihara (1987, sec. 5) and Earman (1992, pp. 196ff.) take this position and support it with examples. However, in their examples the shift in the probability of existing theories is plausibly due to the confirmation of the new theory after its introduction; such a case can obviously be handled by standard Bayesian learning theory once the new theory has been assigned a probability. The interesting case is the one in which the introduction of a theory itself alters the probability of existing theories; in the next section I will show that this can occur.
3 For an exposition and defense of the interpretation of subjective probability assumed here, and its application to scientific theories, see Maher (1993).
4 One proposal, endorsed by Dorling, would define my function \( m \) by \( m(A) = 2^{-K(A)} \), where \( K(A) \) is the number of bits required to represent \( A \) as a prefix-free program "in the theorist's internal programming language" (Dorling 1991, p. 199). However, I doubt that "the theorist's internal programming language" is a well-defined entity. Dorling asserts that "[o]ne of the fundamental theorems of [complexity] theory shows that for any reasonably non-trivial theories their relative priors so assigned are negligibly dependent on the choice of original programming language" (loc. cit.). The theorem I think he is alluding to (Li and Vitanyi 1989, p. 170) says that the number of bits needed to encode a program is unique up to a positive constant, and that strikes me as practically no uniqueness at all. Furthermore, this whole approach rests on a positivist assumption that
I think is false, namely that what a scientific theory asserts is that a particular sequence of observations will be obtained.

For those who want to correlate the present treatment with the one in the preceding section: \( q_1 \) corresponds to \( p(lm(A_2) = 10) \), \( a \rightarrow R_q \) corresponds to \( m(A_2) = 10 \). \( q_2 \) corresponds to \( p(lm(A_2) = 1) \), and \( a \rightarrow R_q \) corresponds to \( m(A_2) = 1 \).

I thank Brad Armendt for helpful comments on this paper.

REFERENCES


Department of Philosophy
University of Illinois at Urbana-Champaign
Urbana, IL 61801
USA