Inductive Logic and the Justification of Induction

Patrick Maher (p-maher@uiuc.edu)
Department of Philosophy, University of Illinois at Urbana-Champaign

1. Philosophical method

Philosophy is a subject in which almost everything is disputed and very little progress is made. Why is it like that? I think the chief reason is that philosophers are concerned with vague and ambiguous concepts such as justification, knowledge, justice, virtue, free will, and personal identity. Because these concepts are vague and ambiguous, many of the questions that philosophers ask about them have no answer that is either true or false. We can describe this by saying that many questions philosophers ask are meaningless; they are what the logical empiricists called pseudo-questions. In view of this, it is not surprising that philosophers' work is so unproductive.

How can we make philosophy more productive? The approach I favor is to replace vague and ambiguous concepts by precise ones. The replacement concepts, besides being precise, need to be useful for at least some of the purposes for which we use the original concepts. Once a precise replacement concept has been specified we will be able to give definite answers to the questions that philosophers now argue about.

There are many different precise concepts that could be used to replace a vague and ambiguous concept. In choosing between these different possible replacements we should look for the one that best serves our purposes; thus the choice is a pragmatic one and different people can make different choices without either having made a mistake. Although we may cite the considerations that make us favor one replacement rather than another, we are not thereby disagreeing with people who favor a different replacement.

I've motivated this way of doing philosophy by presenting it as a method for replacing meaningless questions by meaningful ones. However, that is not the only thing to be said in its favor. Another thing to be said for it is this: The concepts about which philosophers argue are ones that have been handed down to us through a long history of use and misuse and modification under pressure of changing theories. I see no reason to think that the resulting concepts are the ones best suited to our purposes; we probably could do better with new concepts that are designed to serve our purposes. An ancient building that has been successively modified over the centuries is usually not the best for modern purposes and the same is likely to be true of concepts.

In addition, there is more scope for creativity in philosophy done the way I am advocating than in philosophy as it is usually practiced. That is because, on the approach I am advocating, we can design new concepts that serve our purposes instead of just exploring old ones that have been handed down to us. Doing philosophy this way also allows for beautiful and elegant theories, whereas philosophy done the usual way tends to be far from that.

It might be argued that philosophy as it is usually practiced is not really so different to what I am proposing. A philosopher who asserts that certain things are just or unjust, for example, might be construed as proposing that some of the vagueness or ambiguity in the concept of justice be resolved in a particular way. And isn't this just the proposal of a new and more precise concept, as I advocate? No, it isn't. First, I advocate presenting new proposals explicitly as proposals whereas the philosopher is here construed as making a proposal that is misrepresented as a claim with a truth value. Second, on the approach I am advocating the proposed concept should be given a precise definition; the philosopher in my example is not presenting a definition, just some ad hoc attitudes. In other cases philosophers do make claims that can be construed as definitions but these are usually vague, not precise.[1] Third, I hold that it would often be best to replace existing concepts with entirely different ones, something that is not within the scope of philosophers who claim to be analyzing existing concepts.

2. Application to induction

I will now illustrate my way of doing philosophy by applying it to the justification of induction.

Philosophers argue about whether beliefs concerning unobserved matters of fact are ever justified. A few[2] claim that Hume proved that such beliefs are never justified but most reject this conclusion and some[3] deny that Hume was even trying to prove that. Those who think that such beliefs are justified offer various arguments to support their position but these arguments seldom persuade anyone else. Here we have a classic philosophical problem in which almost everything is disputed and very little progress is made due to the vagueness and ambiguity of the relevant concept, in this case the concept of justified

[1] For an example, see the discussion of reliabilism in Section 4.3.
[2] Popper (1959) and Howson (2000), for example.
[3] Beauchamp and Rosenberg (1981) and Owen (1999), for example.

belief. My plan is to pursue a more productive approach by formulating a precise replacement for the concept of justified belief.

The concept of justified belief is used for a variety of purposes, not all of which could be served by any univocal concept. One of those purposes is to guide decision making; which decision we think it best to make in a given decision problem depends in general on what beliefs we take to be justified. I want my replacement for the concept of justified belief to also serve this purpose (and to serve it better than the original concept because of its greater clarity and precision). One of the other purposes for which the concept of justified belief is used is to praise or blame people; it seems to me that epistemologists focus on this purpose but I find that much less interesting and my replacement concept is not intended to be used for that purpose.

My replacement concept will be a mathematical function. This function will be defined by specifying, for pairs of propositions A and B, a number which is denoted $p(A|B)$. This number $p(A|B)$ will serve as a replacement for the old concept of the degree of certainty in A that is justified when the total evidence is B. Because the values of the function are determined by its definition, this replacement concept will be perfectly precise.

The assertion that $p(A|B)$ has some value, say r, will be a purely descriptive claim that is demonstrably either true or false. On the other hand, we are making a normative claim when we say that a person with total evidence B is justified in being certain of A to degree r. Thus my replacement concept will not mean the same as the concept it replaces. Despite this difference, my replacement concept can serve the purpose of guiding decision making because we can use the numbers $p(A|B)$ as an input to a rule for making decisions. I will propose a specific rule in Section 6.

Once the function p has been defined, the philosopher's question of whether beliefs about unobserved matters of fact are ever justified gets replaced by questions like the following: If A is a proposition about the properties of certain individuals, and B is a proposition about the properties of different individuals, is it ever the case that $p(A|B)$ is high? Also, if X is a logically true proposition, is it ever the case that $p(A|B)$ is greater than $p(A|X)$? We can answer these questions simply by consulting the definition of p. I think that, for any function p that anyone would find useful for guiding decision-making, the answer to these questions would be affirmative.

3. Inductive logic

I will now give a stipulative definition of inductive logic. It is the subject that is concerned with finding functions that are satisfactory replacements for the concept of justified belief and with investigating the properties of such functions.

Typical activities in inductive logic include the formulation of precise conditions that represent desiderata for such a function and the search for functions that satisfy those conditions. The choice of the conditions is a pragmatic one that different people might make differently. But once some conditions have been formulated, the question of what functions satisfy those conditions is purely mathematical.

The proposal that I am making in this paper can now be put this way: Let us abandon the usual sorts of philosophical arguments about the justification of induction and instead do inductive logic.

4. Comparisons

In this section I will compare my proposal with some other approaches to the justification of induction.

4.1. Carnap

There are important similarities between my proposals and various aspects of Carnap's work. This is not a coincidence; my views on philosophical method, presented in Section 1, are influenced by Carnap's views expressed in *Logical Syntax of Language* (1937) and elsewhere. Also my definition of inductive logic is close to one given by Carnap (1971, sec. 6).

But there are also aspects of Carnap's work that are inconsistent with my proposals. One such aspect is Carnap's methodology of explication. What Carnap calls explication involves replacing a vague concept with a more precise one and, to that extent, it is like the method I favor. However, Carnap thought that the replacement concept should be similar to the one that it replaces (1950, p. 7) whereas I do not require this; I merely require that the replacement concept serve some of the purposes of the original concept. Carnap also thought that it was important to clarify the concept one is replacing before attempting to find a replacement (1950, pp. 4f.) but I think such an attempt at clarification is a messy business that is difficult to get right and not worth the effort. The difficulty of getting it right is well illustrated by Carnap's own efforts to clarify the concept of degree of confirmation; he later conceded that "the paraphrases and informal explanations that I gave . . . are often ambiguous and may sometimes even be misleading" (1950, pp. xviif.). It is not worth the effort because we are not attempting to make the new concept similar to the one it replaces, only to have the new concept serve some of the purposes that the old one serves.[4]

Carnap, especially in his later work on inductive logic, also offered purported arguments that various things were or were not rational (see, for example, Carnap, 1971). Here he was not just attempting to clarify, but actually attempting to use, a vague and ambiguous pretheoretic concept. As a result he was trying to answer questions that he himself earlier recognized to be pseudo-questions. The methodology I have proposed does not attempt to answer pseudo-questions.

4.2. Subjective Bayesianism

Subjective Bayesians hold that people's degrees of certainty ought to satisfy the laws of the probability calculus. Some also require that later degrees of certainty should be related to earlier ones in a certain way. Apart from these constraints, though, subjectivists allow that people may have whatever degrees of certainty they like. So, for example, your degree of certainty that the sun will rise tomorrow can be anywhere from zero to one provided that your other degrees of certainty are in accord with the value you choose. De Finetti was one of the first and most uncompromising subjectivists but most of those who now work on probabilistic confirmation theory are subjectivists.

Subjectivists standardly dismiss Carnap's inductive logic as a misguided attempt to fix one particular probability function out of a whole class that could equally well be chosen. So they might similarly object to any specific function that I propose as a replacement for the concept of justified belief. However, I have said that the choice of a precise replacement for a concept is a pragmatic matter and different people may make different decisions without either having made a mistake. Thus the approach I am advocating incorporates a considerable degree of what might be called subjectivism.

Nevertheless, the approach I am advocating differs from subjectivism. Subjectivists do not attempt to define a numerical function, like the one I aim to define, that can be used to guide decision making. Instead they say that we should each use our own degrees of certainty to guide decision making. I think this is unsatisfactory because we rarely have degrees of certainty or, if we do, we doubt whether they are worth

[4] Recently Frank Jackson (1998) has defended conceptual analysis as an important philosophical task. I reject conceptual analysis for the same reasons that I reject Carnap's methodology of explication.

much. So I think there is a need for a function that provides a considered way of arriving at numbers that take account of the available evidence and that can be used as degrees of certainty. Although the choice of such a function is a pragmatic matter about which people may differ, there is no reason why large numbers of people could not agree in adopting one particular function.

Perhaps one reason that subjectivists do not think this is a worthwhile project is that they are impressed by the fact that in real life we are unable to state all our relevant evidence, either because there is too much of it, or because we have forgotten some of it, or because it is not fully verbalizable. I grant these facts and hence that, if E represents our total evidence relevant to some proposition A, we cannot expect to be able to calculate $p(A|E)$ simply because we cannot formulate E. On the other hand, we could formulate a proposition, D say, that represents an important part of our evidence and calculate $p(A|D)$. This can be regarded as an approximation to what $p(A|E)$ should be. Similarly, Galileo's laws of falling bodies are useful in many contexts even though they ignore air resistance and the frictions involved in rolling balls down inclined planes. The fact is that we do often argue about what the evidence supports and we cite evidence that we can state in giving such arguments; the sort of function I am proposing to define simply tries to codify the principles involved in such arguments.

4.3. Reliabilism

Many epistemologists hold that a belief is justified if and only if it was formed by a reliable process (Goldman, 1979). From my perspective this is most charitably construed as a proposal to replace the pretheoretic concept of a justified belief with the concept of a belief that is obtained by a reliable process. So construed, reliabilism is an alternative to my proposed replacement for the concept of justified belief.

One problem, which has been noted by others, is that "the" process by which a belief was formed can be identified in many different ways. Another problem is that the concept of a method's reliability is not perfectly clear and, depending on how it is defined, it may not always exist. Furthermore, even if we have identified the relevant method and measured its reliability, it has not been specified what level of reliability is required for a method to count as "reliable". For these reasons the reliabilist's replacement concept is vague and ambiguous whereas I am advocating the use of precise concepts.

Suppose, though, that we had a precise concept of the reliability of the method by which a belief was formed. The question then would be whether this proposal serves our purposes better than the proposal

I have outlined. This depends in part on what our purposes are and different people may have different purposes. What most interests me is to find a concept that is useful in guiding decision making. Reliabilism is poorly suited to this role for three reasons.

First, the concept of the reliability of a method seems to be related in some way to the ratio of truths generated by the method, and this presupposes that the method generates the sort of beliefs that can be true or false, namely categorical beliefs. On the other hand, what we need to guide decision making is not categorical beliefs but rather degrees of certainty and these aren't true or false. Second, even if a belief fits the evidence well, it need not have been formed by a reliable process, since it may have been formed for reasons other than the evidence; hence reliabilists deem some beliefs that fit the evidence well to be unjustified. On the other hand, beliefs that fit the evidence well are precisely the ones we want to use in making decisions. Third, the reliability of the process by which a belief was formed is in general unknown and so a procedure for guiding decision making cannot refer to it.

In short, reliabilism is not only vague and ambiguous but also unsuitable if one's purpose is to find a concept that is useful for guiding decision making.

5. Formalization

I've proposed replacing the concept of justified belief with a mathematical function that assigns numbers to pairs of propositions. The problem now is to find a suitable function. In the remainder of this paper I will take the first steps in that direction.

First, let me set out the formal framework that I will use. I start with the concept of a *state*, which is a way that things might be that is fully specific in all respects that are of interest in the present context. For example, if we are interested in tomorrow's weather, a state is one way that tomorrow's weather might be, specified as fully as matters to us. We now select a *set* of states in such a way that it is logically necessary that one and only one state is actual. This set of states will be denoted X.

For any subset A of X there is a corresponding proposition which asserts that the real state is in A. It is convenient to not distinguish between the set and the corresponding proposition; thus if A is a subset of X I call A a proposition. The following facts about propositions are worth noting:

1. X is a proposition that is logically true. That is because it is logically necessary that the real state is in X.

2. $\emptyset$ (the empty set) is a proposition that is logically false. That is because $\emptyset$ contains nothing and hence cannot contain the real state.

3. $A \cup B$ (the union of A and B) represents the disjunction of propositions A and B.

4. $A \cap B$ (the intersection of A and B) represents the conjunction of propositions A and B.

5. $\bar{A}$ (the complement of A, that is, the set of states in X that are not in A) represents the negation of the proposition A.

Although every subset of X may be considered to be a proposition, it is sometimes convenient to focus attention on only some of these subsets. I will denote the set of propositions that are being considered by $\mathcal{X}$. I assume that $\mathcal{X}$ contains X and is closed under complementation and countable union.[5]

In addition to the set X of states we can suppose there is a set Y that contains all the possible consequences that might be realized in the present context. An *act* can then be regarded as a function from X to Y.[6] Thus if f is an act and $x \in X$ then $f(x)$ represents the consequence that would be realized if f is chosen and the actual state is x.

6. Expected utility

I said that I intend the function p to be useful for guiding decisions. I also said that such guidance can be effected by using the numbers $p(A|B)$ as an input to a rule for making decisions. I will now propose a specific rule for this.

This rule requires that, in addition to the function p, we choose a *utility function*. This function, to be denoted u, will be a precise replacement for the vague and ambiguous concept of how good a consequence is. As with p, my choice of a specific u would be motivated by its intended role of guiding decision making.

Suppose that f is an act and $E \in \mathcal{X} - \{\emptyset\}$. Let the possible consequences of f be $a_1, \ldots, a_n$, where $a_i \neq a_j$ for $i \neq j$. Also, for each

[5] In mathematical terms, I am assuming that $\mathcal{X}$ is a σ-algebra (also known as a σ-field).

[6] I also require that if f is an act then, for each $y \in Y$, $\{x \in X\colon f(x) = y\} \in \mathcal{X}$.

$i = 1, \ldots, n$, let $A_i = \{x \in X\colon f(x) = a_i\}$. Then I define the *expected utility of f given E* to be:

$$EU(f|E) = p(A_1|E)u(a_1) + \cdots + p(A_n|E)u(a_n).$$

A parallel definition holds when f has infinitely many possible consequences.

The rule that I propose for making decisions is to choose an act f such that if E is the total evidence then, for any available act g, $EU(f|E) \geq EU(g|E)$. An illustrious line of authors from Arnauld (1662) to the present have agreed that the rule for making decisions should have this form. The rule takes account of the relevant considerations in a way that seems intuitively satisfactory to many and it also has attractive formal properties.[7] These are my motivations for proposing this rule.

7. Probability

With this background we can now turn to the specification of the function p. I will approach the specification of this function by imposing several constraints that it will be required to satisfy. I will call these constraints *axioms*.

Suppose f and g are two acts and that, in every state compatible with the evidence, the consequence obtained from choosing f is least as good as the consequence obtained from choosing g. Then most people, including me, would say that f is at least as good a choice as g. Expressing this view in our more precise concepts gives:

AXIOM 1. *If f and g are acts, $E \in \mathcal{X} - \{\emptyset\}$, and $u(f(x)) \geq u(g(x))$ for all $x \in E$, then $EU(f|E) \geq EU(g|E)$.*

Suppose now that, in every state compatible with the evidence, the consequence obtained from choosing f is *better* than the consequence obtained from choosing g. Then most people, including me, would say that f is a better choice than g. Expressing this view in our more precise concepts gives:

AXIOM 2. *If f and g are acts, $E \in \mathcal{X} - \{\emptyset\}$, and $u(f(x)) > u(g(x))$ for all $x \in E$, then $EU(f|E) > EU(g|E)$.*

[7] For some of the formal properties, and a defense of their attractiveness, see the discussion of transitivity, normality, and independence in Maher (1993, chs. 2 and 3).

If we multiplied all values of p by some positive number, the ordering of acts by expected utility would be unaffected and so the role of p in guiding decision making would not be affected. In particular, if $p(E|E) > 0$ we could multiply all values of p by $1/p(E|E)$, in which case $p(E|E)$ becomes 1. Therefore, as a purely conventional decision, I adopt:

AXIOM 3. *If $E \in \mathcal{X} - \{\emptyset\}$ and $p(E|E) > 0$ then $p(E|E) = 1$.*

These axioms imply the following theorem. (Proofs of theorems are given in Section 10.)

THEOREM 1. *Let $E \in \mathcal{X} - \{\emptyset\}$.*

(i) If $A \in \mathcal{X}$ then $p(A|E) \geq 0$.

(ii) If $A_1, A_2, \ldots \in \mathcal{X}$ and $A_i \cap A_j = \emptyset$ for $i \neq j$ then

$$p(A_1 \cup A_2 \cup \ldots |E) = p(A_1|E) + p(A_2|E) + \ldots.$$

(iii) $p(E|E) = p(X|E) = 1$.

This theorem can be summarized by saying that, for each nonempty proposition E, $p(\cdot|E)$[8] is a countably additive probability function on $\mathcal{X}$ with $p(E|E) = 1$.

There are many arguments in the literature that attempt to prove that rational degrees of certainty ought to satisfy the axioms of probability; for example, there are Dutch book arguments and representation theorems. The derivation of the probability axioms that I have just given is significantly different to those arguments. To start with, the conclusion is different. I am not making any claim about rational degrees of certainty; rationality is one of those vague and ambiguous concepts that I would replace with something precise. Second, the formal structure of the argument differs from both Dutch book arguments and representation theorems. Dutch book arguments for the additivity property make the objectionable assumption that bets that are severally acceptable are also jointly acceptable (Maher 1993, ch. 4; 1997); my derivation makes no such assumption. Representation theorems (such as in Maher 1993, ch. 8) require strong structural assumptions and heavy mathematics to establish their result whereas the derivation given here requires neither. These differences are made possible by the fact that I am arguing for a different conclusion.

[8] By $p(\cdot|E)$ I mean the function that maps each $A \in \mathcal{X}$ onto $p(A|E)$.

8. Conditionalization

Suppose that f would be a better choice than g when the evidence is E and also when the evidence is $\bar{E}$. In that case I would say that f is also the better choice when there is no evidence. This is a version of what Savage called the *sure-thing principle*. Savage (1954, p. 21) gave the following illustration of this principle:

> A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant to the attractiveness of the purchase. So, to clarify the matter for himself, he asks whether he would buy if he knew that the Republican candidate were going to win, and decides that he would do so. Similarly, he considers whether he would buy if he knew that the Democratic candidate were going to win, and again finds that he would do so. Seeing that he would buy in either event, he decides that he should buy, even though he does not know which event obtains.

Expressed precisely in our concepts, the principle is:

AXIOM 4. *If f and g are acts, $E \in \mathcal{X} - \{\emptyset\}$, $EU(f|E) \geq EU(g|E)$, and $EU(f|\bar{E}) \geq EU(g|\bar{E})$, then $EU(f|X) \geq EU(g|X)$.*

We now get:

THEOREM 2. *If $A, E \in \mathcal{X}$ and $p(E|X) > 0$ then*

$$p(A|E) = p(A \cap E|X)/p(E|X).$$

If we translate this into the vague and ambiguous terms that p is meant to replace, what this theorem says is that the degree of certainty in A that is justified when the total evidence is E equals the degree of certainty in $A \cap E$ that is justified when there is no evidence, divided by the degree of certainty in E that is justified when there is no evidence.

Many subjective Bayesians endorse a principle of *conditionalization*, which says that if you learn E and nothing else then your degree of certainty in any proposition A ought to equal your prior degree of certainty in $A \cap E$ divided by your prior degree of certainty in E. Theorem 2 is the analog, in my concepts, of this principle of conditionalization.

Bayesians have attempted to justify the principle of conditionalization with a variety of arguments, such as diachronic Dutch book arguments, but these arguments have also been extensively criticized.[9] My derivation of Theorem 2 is quite different to any of the published

[9] For some of the literature, see Maher (1993, ch. 5).

arguments for conditionalization; so far as I know, nobody has previously derived conditionalization from the sure-thing principle. As with my derivation of the axioms of probability, this difference depends on the fact that I am not arguing for the same conclusion. I don't say that a person would be irrational to revise degrees of certainty in a certain way; I think that claim is meaningless. What I say is that any function that I would accept as a precise replacement for the concept of justified degree of certainty will have prior and posterior probabilities related according to the usual formula (given in Theorem 2).

9. A priori probabilities

I have been pursuing the project of defining a function p that can serve as a precise replacement for the concept of justified degrees of certainty. Theorem 1 tells us that p must be such that $p(\cdot|E)$ is a countably additive probability function. Theorem 2 tells us that p must also satisfy the usual rule for conditional probabilities.

To proceed further it is useful to now focus on determining the values of $p(\cdot|X)$. These values are called *a priori probabilities* because they correspond to situations where there is no evidence. There are two reasons for focusing on a priori probabilities at this point.

1. If we can specify values of $p(\cdot|X)$ then, by Theorem 2, we can derive the values of $p(\cdot|E)$ for all E such that $p(E|X) > 0$. Thus, determining the a priori probabilities is sufficient to determine all the values of p that we are likely to need. This is not true for a posteriori probabilities.

2. Since $p(\cdot|X)$ is meant to correspond to the degree of certainty that is justified when there is no evidence, it is natural to require $p(\cdot|X)$ to satisfy various symmetry conditions. For example, we can require that the values of p be unchanged when the names of individuals are permuted. By contrast, once we have relevant evidence these symmetry conditions will no longer be plausible. Such symmetries makes it easier to fix the values of p.

Subjective Bayesians sometimes suggest that a priori probabilities are meaningless. What they sometimes allude to in support of this is that we cannot acquire concepts without at the same time acquiring a considerable amount of relevant evidence. But even if we cannot in fact acquire concepts without having relevant evidence, this is a contingent fact about us and it is not incoherent to suppose, counterfactually, that we have some concepts without having any relevant evidence. Even concept empiricists like Locke and Hume did not deny that innate ideas are conceivable.

There remains the question of how we are to assign a priori probabilities. In part this is done based on intuitive judgments about what kinds of values we would want to use if we lacked any relevant evidence. Most people seem to have such intuitive judgments. Another way to assign a priori probabilities is by choosing ones that lead to the most satisfactory a posteriori probabilities. Approached this way, we do not need to attribute any intuitive meaning to a priori probabilities in order to assign values to them; we can think of them as theoretical entities whose significance consists in the a posteriori probabilities that we derive from them and use in real situations. In practice this method is usually combined with the previous one but someone who can't make sense of the supposition of no evidence may use this method exclusively.

So much for the importance and coherence of assigning a priori probabilities. There remains the real work of actually finding satisfactory numerical assignments. This is a project that Carnap devoted several decades to and on which I have worked in recent years (Maher 2000; 2001). That is too large a topic to describe here but I hope that what I have said in this paper shows that the project is not misguided, contrary to what most philosophers today believe.

10. Proofs

10.1. Proof of Theorem 1

10.1.1. *Condition (i)*

Let $A \in \mathcal{X}$ and suppose that $p(A|E) < 0$.

Let a, b, and c be three different consequences with $u(a) > u(b)$. Let f and g be the acts such that

$$f(x) = a \text{ and } g(x) = b \text{ for all } x \in A;$$
$$f(x) = g(x) = c \text{ for all } x \in \bar{A}.$$

Now $p(A|E)u(a) < p(A|E)u(b)$ and so

$$\begin{aligned} EU(f|E) &= p(A|E)u(a) + p(\bar{A}|E)u(c) \\ &< p(A|E)u(b) + p(\bar{A}|E)u(c) \\ &= EU(g|E). \end{aligned}$$

This is a violation of Axiom 1. Hence the supposition that $p(A|E) < 0$ is false, that is, $p(A|E) \geq 0$.

10.1.2. *A finiteness condition*
Let $A_1, A_2, \ldots \in \mathcal{X}$ and $A_i \cap A_j = \emptyset$ for $i \neq j$. Suppose that

$$\sum_{i=1}^{\infty} p(A_i|E) = \infty. \tag{1}$$

Let $a_1, a_2, \ldots, b$, and c be different consesquences with $0 < u(a_i) \leq u(b)$ for all i. Let f and g be the acts such that

$$f(x) = a_i \text{ and } g(x) = b \text{ for all } x \in A_i;$$
$$f(x) = g(x) = c \text{ for all } x \in \cap_{i=1}^{\infty} \bar{A}_i.$$

Then by (1) we have $EU(f|E) = \infty > EU(g|E)$, which violates Axiom 1. Hence (1) is false and so $\sum_{i=1}^{\infty} p(A_i|E)$ is finite.

10.1.3. *Condition (ii)*
Let $A_1, A_2, \ldots \in \mathcal{X}$ and $A_i \cap A_j = \emptyset$ for $i \neq j$. I have just shown that $\sum_{i=1}^{\infty} p(A_i|E)$ is finite. Now suppose that

$$p(\cup_{i=1}^{\infty} A_i|E) < \sum_{i=1}^{\infty} p(A_i|E) \tag{2}$$

and let

$$\varepsilon = \sum_{i=1}^{\infty} p(A_i|E) - p(\cup_{i=1}^{\infty} A_i|E).$$

Thus $\varepsilon > 0$. Let $a, b_1, b_2, \ldots$, and c be different consequences. If $\sum_{i=1}^{\infty} p(A_i|E) > 0$, let these consequences be such that, for all i,

$$u(a) \geq u(b_i) > u(a) \left[1 - \frac{\varepsilon}{\sum_{i=1}^{\infty} p(A_i|E)}\right].$$

If $\sum_{i=1}^{\infty} p(A_i|E) = 0$, let these consequences be such that $u(a) > 0$ and $u(a) \geq u(b_i)$ for all i. Let f and g be the acts such that

$$f(x) = a \text{ and } g(x) = b_i \text{ for all } x \in A_i;$$
$$f(x) = g(x) = c \text{ for all } x \in \cap_{i=1}^{\infty} \bar{A}_i.$$

Then if $\sum_{i=1}^{\infty} p(A_i|E) > 0$ we have

$$\begin{aligned} EU(f|E) - EU(g|E) &= p(\cup_{i=1}^{\infty} A_i|E)u(a) - \sum_{i=1}^{\infty} p(A_i|E)u(b_i) \\ &= \left[\sum_{i=1}^{\infty} p(A_i|E) - \varepsilon\right] u(a) - \sum_{i=1}^{\infty} p(A_i|E)u(b_i) \end{aligned}$$

$$< \left[\sum_{i=1}^{\infty} p(A_i|E) - \varepsilon\right] u(a) - \sum_{i=1}^{\infty} p(A_i|E)\left[1 - \frac{\varepsilon}{\sum_{i=1}^{\infty} p(A_i|E)}\right] u(a)$$
$$= 0.$$

If $\sum_{i=1}^{\infty} p(A_i|E) = 0$ we have

$$\begin{aligned} EU(f|E) - EU(g|E) &= p(\cup_{i=1}^{\infty} A_i|E)u(a) \\ &= -\varepsilon u(a) \\ &< 0. \end{aligned}$$

It follows from (i) that either $\sum_{i=1}^{\infty} p(A_i|E) > 0$ or $\sum_{i=1}^{\infty} p(A_i|E) = 0$. Hence $EU(f|E) < EU(g|E)$. Since $u(f(x)) \geq u(g(x))$ for all x, this is a violation of Axiom 1. Therefore the supposition (2) is false.

A similar argument shows that we also cannot have

$$p(\cup_{i=1}^{\infty} A_i|E) > \sum_{i=1}^{\infty} p(A_i|E).$$

Hence (ii) is true.

10.1.4. *Condition (iii)*
Let a, b, and c be three different consequences with $u(a) > u(b)$. Let f and g be the acts such that

$$f(x) = a \text{ and } g(x) = b \text{ for all } x \in E;$$
$$f(x) = g(x) = c \text{ for all } x \in \bar{E}.$$

Then

$$EU(f|E) - EU(g|E) = p(E|E)[u(a) - u(b)].$$

But $u(f(x)) > u(g(x))$ for all $x \in E$ and so, by Axiom 2, $EU(f|E) > EU(g|E)$. Hence $p(E|E) > 0$. So, by Axiom 3, $p(E|E) = 1$.

Now, with a, b, and c as before, let f and g be the acts such that

$$f(x) = g(x) = c \text{ for all } x \in E;$$
$$f(x) = a \text{ and } g(x) = b \text{ for all } x \in \bar{E}.$$

Then

$$EU(f|E) - EU(g|E) = p(\bar{E}|E)[u(a) - u(b)].$$

But by Axiom 1, $EU(f|E) = EU(g|E)$; hence $p(\bar{E}|E) = 0$. So

$$\begin{aligned} p(X|E) &= p(E|E) + p(\bar{E}|E), \text{ by (ii)} \\ &= 1 + 0 = 1. \end{aligned}$$

10.2. Proof of Theorem 2

Let $A, E \in \mathcal{X}$ with $p(E|X) > 0$. Suppose that

$$p(A|E) > p(A \cap E|X)/p(E|X). \qquad (3)$$

Let a, b, and c be three consequences such that $u(a) = p(\bar{A}|E)$, $u(b) = p(A|E)$, and $u(c) = 0$. Let f and g be the acts such that

$$\begin{aligned} f(x) &= \left\{ \begin{array}{l} a \text{ if } x \in A \cap E \\ c \text{ otherwise} \end{array} \right\} \\ g(x) &= \left\{ \begin{array}{l} b \text{ if } x \in \bar{A} \cap E \\ c \text{ otherwise} \end{array} \right\}. \end{aligned}$$

Then

$$\begin{aligned} EU(f|E) &= p(A \cap E|E)u(a) \\ &= p(A|E)p(\bar{A}|E) \\ &= p(\bar{A} \cap E|E)u(b) \\ &= EU(g|E). \qquad (4) \end{aligned}$$

If $\bar{E} = \emptyset$ then $E = X$ whence $p(A \cap E|X)/p(E|X) = p(A|X) = p(A|E)$, contrary to (3); hence $\bar{E} \neq \emptyset$. Also $f(x) = g(x) = c$ for all $x \in \bar{E}$. So by Axiom 1,

$$EU(f|\bar{E}) = EU(g|\bar{E}). \qquad (5)$$

We also have

$$\begin{aligned} p(\bar{A}|E) &= 1 - p(A|E) \\ &< 1 - p(A \cap E|X)/p(E|X) \\ &= p(\bar{A} \cap E|X)/p(E|X). \qquad (6) \end{aligned}$$

So we get

$$\begin{aligned} EU(f|X) &= p(A \cap E|X)u(a) \\ &= p(A \cap E|X)p(\bar{A}|E) \\ &< p(A \cap E|X)p(\bar{A} \cap E|X)/p(E|X), \text{ by (6)} \\ &< p(\bar{A} \cap E|X)p(A|E), \text{ by (3)} \\ &= p(\bar{A} \cap E|X)u(b) \\ &= EU(g|X). \qquad (7) \end{aligned}$$

Equations (4), (5), and (7) together violate Axiom 4; hence the supposition (3) is false.

If instead we were to suppose that

$$p(A|E) < p(A \cap E|X)/p(E|X)$$

then the inequalities (6) and (7) would be reversed but that also violates Axiom 4. Hence $p(A|E) = p(A \cap E|X)/p(E|X)$.

References

Arnauld, A.: 1662, *La logique, ou, L'art de penser*. Paris: Chez Charles Savreux.

Beauchamp, T. L. and A. Rosenberg: 1981, *Hume and the Problem of Causation*. Oxford: Oxford University Press.

Carnap, R.: 1937, *The Logical Syntax of Language*. London: Routledge and Kegan Paul. Translated by Amethe Smeaton.

Carnap, R.: 1950, *Logical Foundations of Probability*. Chicago: University of Chicago Press. Second edition 1962.

Carnap, R.: 1971, 'Inductive Logic and Rational Decisions'. In: R. Carnap and R. C. Jeffrey (eds.): *Studies in Inductive Logic and Probability*, Vol. 1. Berkeley: University of California Press, pp. 5–31.

Goldman, A. I.: 1979, 'What is Justified Belief?'. In: G. S. Pappas (ed.): *Justification and Knowledge*. Dordrecht: Reidel, pp. 1–23.

Howson, C.: 2000, *Hume's Problem: Induction and the Justification of Belief*. Oxford: Oxford University Press.

Jackson, F.: 1998, *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford: Oxford University Press.

Maher, P.: 1993, *Betting on Theories*. Cambridge: Cambridge University Press.

Maher, P.: 1997, 'Depragmatized Dutch Book Arguments'. *Philosophy of Science* **64**, 291–305.

Maher, P.: 2000, 'Probabilities for Two Properties'. *Erkenntnis* **52**, 63–91.

Maher, P.: 2001, 'Probabilities for Multiple Properties: The Models of Hesse and Carnap and Kemeny'. *Erkenntnis*. Forthcoming.

Owen, D.: 1999, *Hume's Reason*. Oxford: Oxford University Press.

Popper, K. R.: 1959, *Logic of Scientific Discovery*. New York: Harper and Row. Second edition, 1968.

Savage, L. J.: 1954, *The Foundations of Statistics*. New York: John Wiley. Second edition, Dover 1972.