SECTION

V

FOUNDATIONS OF PROBABILITY AND INDUCTION

THE AIM OF INDUCTIVE LOGIC

RUDOLF CARNAP

University of California, Los Angeles, California, U.S.A.

By inductive logic I understand a theory of logical probability providing rules for inductive thinking. I shall try to explain the nature and purpose of inductive logic by showing how it can be used in determining rational decisions.

I shall begin with the customary schema of decision theory involving the concepts of utility and probability. I shall try to show that we must understand "probability" in this context not in the objective sense, but in the subjective sense, i.e., as the degree of belief. This is a psychological concept in empirical decision theory, referring to actual beliefs of actual human beings. Later I shall go over to rational or normative decision theory by introducing some requirements of rationality. Up to that point I shall be in agreement with the representatives of the subjective conception of probability. Then I shall take a further step, namely, the transition from a quasi-psychological to a logical concept. This transition will lead to the theory which I call "inductive logic".

We begin with the customary model of decision making. A person X at a certain time T has to make a choice between possible acts $A_1, A_2, \cdots$. X knows that the possible states of nature are $S_1, S_2, \cdots$; but he does not know which of them is the actual state. For simplicity, we shall here assume that the number of possible acts and the number of possible states of nature

---

The author is indebted to the National Science Foundation for the support of research in inductive probability.

are finite. X knows the following: if he were to carry out the act A_m and if the state S_n were the actual state of nature, then the outcome would be $O_{m,n}$. This outcome $O_{m,n}$ is uniquely determined by A_m and S_n; and X knows how it is determined. We assume that there is a utility function U_x for the person X and that X knows his utility function so that he can use it in order to calculate subjective values.

Now we define the *subjective value* of a possible act A_m for X at time T:

(1) DEFINITION.

$$V_{X,T}(A_m) = \sum_n U_X(O_{m,n}) \times P(S_n),$$

where $P(S_n)$ is the probability of the state S_n, and the sum covers all possible states S_n.

In other words, we take as the subjective value of the act A_m for X the *expected utility* of the outcome of this act. (1) holds for the time T before any act is carried out. It refers to the contemplated act A_m; therefore it uses the utilities for the possible outcomes $O_{m,n}$ of act A_m in the various possible states S_n. [If the situation is such that the probability of S_n could possibly be influenced by the assumption that act A_m were carried out, we should take the conditional probability $P(S_n|A_m)$ instead of $P(S_n)$. Analogous remarks hold for our later forms of the definition of V].

We can now formulate the customary *decision principle* as follows:

(2) Choose an act so as to maximize the subjective value V.

This principle can be understood either as referring to *actual* decision making, or to *rational* decisions. In the first interpretation it would be a psychological law belonging to *empirical* decision theory as a branch of psychology; in the second interpretation, it would be a normative principle in the theory of *rational* decisions. I shall soon come back to this distinction. First we have to remove an ambiguity in the definition (1) of value, concerning the interpretation of the probability P. There are several conceptions of probability; thus the question arises which of them is adequate in the context of decision making.

The main conceptions of probability are often divided into two kinds, objectivistic and subjectivistic conceptions. In my view, these are not two incompatible doctrines concerning the same concept, but rather two theories concerning two different probability concepts, both of them legitimate and useful. The concept of *objective* (or statistical) *probability is* closely connected with relative frequencies in mass phenomena. It plays an important role in mathematical statistics, and it occurs in laws of various branches of empirical science, especially physics.

The second concept is *subjective* (or personal) *probability.* It is the probability assigned to a proposition or event H by a subject X, say a person or a group of persons, in other words, the degree of belief of X in H. Now it seems to me that we should clearly distinguish two versions of subjective probability, one representing the *actual* degree of belief and the other the *rational* degree of belief.

Which of these two concepts of probability, the objective or the subjective, ought to be used in the definition of subjective value and thereby in the decision principle? At the present time, the great majority of those who work in mathematical statistics still regard the statistical concept of probability as the only legitimate one. However, this concept refers to an objective feature of nature; a feature that holds whether or not the observer *X* knows about it. And in fact, the numerical values of statistical probability are in general not known to *X*. Therefore this concept is unsuitable for a decision principle. It seems that for this reason a number of those who work in the theory of decisions, be it actual decisions or rational decisions, incline toward the view that some version of the subjective concept of probability must be used here. I agree emphatically with this view.

The statistical concept of probability remains, of course, a legitimate and important concept both for mathematical statistics and for many branches of empirical science. And in the special case that *X* knows the statistical probabilities for the relevant states S_n but has no more specific knowledge about these states, the decision principle would use these values. There is general agreement on this point. And this is not in conflict with the view that the decision principle should refer to subjective probability, because in this special situation the subjective probability for *X* would be equal to the objective probability.

Once we recognize that decision theory needs the subjective concept of probability, it is clear that the theory of *actual* decisions involves the first version of this concept, i.e., the *actual* degree of belief, and the theory of *rational* decisions involves the second version, the *rational* degree of belief.

Let us first discuss the theory of *actual* decisions. The concept of probability in the sense of the *actual* degree of belief is a psychological concept; its laws are empirical laws of psychology, to be established by the investigation of the behavior of persons in situations of uncertainty, e.g., behavior with respect to bets or games of chance. I shall use for this psychological concept the technical term *"degree of credence"* or shortly "credence". In symbols, I write '$Cr_{X,T}(H)$' for "the (degree of) credence of the proposition *H* for the person *X* at the time *T*". Different persons *X* and *Y* may have different credence functions $Cr_{X,T}$ and $Cr_{Y,T}$. And the same person *X* may have different credence functions Cr_{X,T_1} and Cr_{X,T_2} at different times T_1 and T_2; e.g., if *X* observes between T_1 and T_2 that *H* holds, then $Cr_{X,T_1}(H) \neq Cr_{X,T_2}(H)$. (Let the ultimate possible cases be represented by the points of a logical space, usually called the probability space. Then a proposition or event is understood, not as a sentence, but as the range of a sentence, i.e., the set of points representing those possible cases in which the sentence holds. To the conjunction of two sentences corresponds the intersection of the propositions.)

On the basis of credence, we can define *conditional credence,* "the credence of *H* with respect to the proposition *E"* (or "··· given E"):

(3) DEFINITION.

$$Cr'_{X,T}(H|E) = \frac{Cr_{X,T}(E \cap H)}{Cr_{X,T}(E)},$$

provided that $Cr_{X,T}(E) > 0$. $Cr'_{X,T}(H|E)$ is the credence which H would have for X at T if X ascertained that E holds.

Using the concept of credence, we now replace (1) by the following:

(4) DEFINITION.

$$V_{X,T}(A_m) = \sum_n U_X(O_{m,n}) \times Cr_{X,T}(S_n),$$

As was pointed out by Ramsey, we can determine *X*'s credence function by his betting behavior. A bet is a contract of the following form. *X* pays into the pool the amount *u*, his partner *Y* pays the amount *v*; they agree that the total stake $u+v$ goes to *X* if the hypothesis *H* turns out to be true, and to *Y* if it turns out to be false. If *X* accepts this contract, we say that he bets on *H* with the total stake $u+v$ and with the betting quotient $q = u|(u+v)$ (or, at odds of *u* to *v*). If we apply the decision principle with the definition (4) to the situation in which *X* may either accept or reject an offered bet on *H* with the betting quotient *q*, we find that *X* will accept the bet if *q* is not larger than his credence for *H*. Thus we may interpret $Cr_{X,T}(H)$ as the highest betting quotient at which *X* is willing to bet on *H*. (As is well known, this holds only under certain conditions and only approximately.)

Utility and credence are psychological concepts. The utility function of *X* represents the system of valuations and preferences of *X*; his credence function represents his system of beliefs (not only the content of each belief, but also its strength). Both concepts are theoretical concepts which characterize the state of mind of a person; more exactly, the non-observable micro-state of his central nervous system, not his consciousness, let alone his overt behavior. But since his behavior is influenced by his state, we can indirectly determine characteristics of his state from his behavior. Thus experimental methods have been developed for the determination of some values and some general characteristics of the utility function and the credence function ("subjective probability") of a person on the basis of his behavior with respect to bets and similar situations. Interesting investigations of this kind have been made by F. Mosteller and P. Nogee [13], and more recently by D. Davidson and P. Suppes [4], and others.

Now we take the step from empirical to *rational decision theory.* The latter is of greater interest to us, not so much for its own sake (its methodological status is in fact somewhat problematic), but because it is the connecting link between empirical decision theory and inductive logic. Rational decision theory is concerned not with actual credence, but with *rational* credence. (We should also distinguish here between actual utility and rational utility; but we will omit this.) The statements of a theory of this kind are not found by experiments but are established on the basis of requirements of rationality; the formal procedure usually consists in deducing theorems from axioms which are justified by general considerations of rationality, as we shall see. It seems fairly clear that the probability concepts used by the following authors are meant in the sense of rational credence (or rational credibility, which I shall explain presently): John Maynard Keynes (1921), Frank P. Ramsey (1928), Harold Jeffreys (1931), B. O. Koopman (1940), Georg Henrik yon Wright (1941), I. G. Good (1950), and Leonard J. Savage (1954). I am inclined to include here also those authors who do not declare e*X*plicitly that their concept refers to rational rather than actual beliefs, but who accept general a*X*ioms and do not base their theories on psychological results. Bruno De Finetti (1931) satisfies these conditions; however, he says explicitly that his concept of "subjective probability" refers not to rational, but to actual beliefs. I find this puzzling.

The term "subjective probability" seems quite satisfactory for the actual degree of credence. It is frequently applied also to a probability concept interpreted as something like rational credence. But here the use of the word "subjective" might be misleading (comp. Keynes [9, p. 4] and Carnap [1, § 12A]). Savage has suggested the term "personal probability".

Rational credence is to be understood as the credence function of a completely rational person *X*; this is, of course, not any real person, but an imaginary, idealized person. We carry out the idealization step for step, by introducing *requirements of rationality* for the credence function. I shall now explain some of these requirements.

Suppose that *X* makes n simultaneous bets; let the ith bet ($i = 1,\ldots., n$) be on the proposition H_i with the betting quotient q_i: and the total stake s_i;. Before we observe which of the propositions H_i are true and which are false, we can consider the *possible* cases. For any possible case, i.e., a logically possible distribution of truth-values among the *Hi,* we can calculate the gain or loss for each bet and hence the total balance of gains and losses from the n bets. If in *every* possible case *X* suffers a net loss, i.e., his total balance is negative, it is obviously unreasonable for *X* to make these n bets. Let *X*'s credence function at a given time be *Cr.* By a (finite) betting system in accordance with *Cr* we mean a finite system of n bets on n arbitrary propositions H_i $(i = 1,\ldots, n)$ with n arbitrary (positive) stakes s_i, but with the betting quotients $q_i = Cr(H_i)$.

(5) DEFINITION. *A function Cr is coherent if and only if there is no betting system in accordance with Cr such that there is a net loss in every possible case.* For *X* to make bets of a system of this kind would obviously be unreasonable. Therefore we lay down the *first requirement* as follows:

R1. *In order to be rational, Cr must be coherent.*

Now the following important result holds:

(6) A function *Cr* from propositions to real numbers is coherent if and only if *Cr* is a normalized probability measure.

(A real-valued function of propositions is said to be a probability measure if it is a non-negative, finitely additive set function; it is normalized if its value for the necessary proposition is 1. In other words, a normalized probability measure is a function which satisfies the basic axioms of the calculus of probability, e.g., the axioms I through V in Kolmogoroff's system [10, § 1].)

The first part of (6) ("... coherent if ...") was stated first by Ramsey *[15]* and was later independently stated and proved by De Finetti [5]. The much more complicated proof for the second part ("... only if ...") was found independently by John G. Kemeny [8, *p. 269]* and R. Sherman Lehman *[12, p. 256]*.

Let *Cr'* be the conditional credence function defined on the basis of *Cr* by (3). *As* ordinary bets are based on *Cr,* conditional bets are based on *Cr'.* The concept of coherence can be generalized so as to be applicable also to conditional credence functions. (6) can then easily be extended by the result that a conditional credence function *Cr' is* coherent if and only if *Cr' is* a normalized conditional probability measure, in other words, if and only if *Cr'* satisfies the customary basic axioms of conditional probability, including the general multiplication axiom.

Following Shimony [17], we introduce now a concept of coherence in a stronger sense, for which I use the term "strict coherence":

(7) DEFINITION. A *function Cr is strictly coherent if and* only *if Cr is coherent and there is no (finite) system of bets in accordance with Cr on molecular propositions such that the result is a net loss in at least one possible case, but not a net gain in any possible case.*

It is clear that it would be unreasonable to make a system of bets of the kind just specified. Therefore we lay down the *second requirement:*

R2. In order to be rational, a credence function must be strictly coherent. We define *regular credence function* (essentially in the sense of Carnap [1, § 55A]):

(8) DEFINITION. A *function Cr is regular if and only if Cr is a normalized probability measure and, for any molecular proposition H, Cr (H) = 0 only a if H is impossible.*

By analogy with (6) we have now the following important theorem; its first part is due to Shimony, its second part again to Kemeny and Lehman:

(9) A function *Cr* is strictly coherent if and only if *Cr is* regular.

Most of the authors of systems for subjective or logical probability adopt only the basic axioms; thus they require nothing but coherence. A few go one step further by including an axiom for what I call regularity; thus they require in effect strict coherence, but nothing more. Axiom systems of both kinds are extremely weak; they yield no result of the form *"P(H|E) = r",* except in the trivial cases where *r* is 0 or 1. In my view, much more should be required.

The two previous requirements apply to any credence function that holds for *X* at any time *T* of his life. We now consider two of these functions, Cr_n for the time T_n and Cr_{n+1} for a time T_{n+1} shortly after T_n. Let the proposition *E* represent the observation data received by *X* between these two time points. The *third requirement* refers to the transition from Cr_n. to Cr_{n+1}:

R3. (a) *The transformation of Cr_n into Cr_{n+1} depends only on the proposition E.*

(b) More specifically, Cr_{n+1} is determined by Cr_n and E as follows: for any H, $Cr_{n+1}(H) = Cr_n(E \cap H)|Cr_n(E)$ (hence $= Cr'_n(H|E)$ by definition (3)).

Part (a) is of course implied by (b). I have separated part (a) from (b) because *X*'s function *Cr* might satisfy (a) without satisfying (b). Part (a) requires merely that *X* be rational to the extent that changes in his credence function are influenced only by his observational results, but not by any other factors, e.g., feelings like his hopes or fears concerning a possible future event *H,* feelings which in fact influence the beliefs of all actual human beings. Part (b) specifies exactly the transformation of Cr_n into Cr_{n+1}; the latter is the conditional credence Cr'_n with respect to *E.* The rule (b) can be used only if $Cr_n(E) \neq 0$; this condition is fulfilled for any possible observational result, provided that Cr_n satisfies the requirement of strict coherence.

Let the proposition E_{n+2} represent the data obtained between T_{n+1} and a later time point T_{n+2}. Let Cr_{n+2} be the credence function at T_{n+2} obtained by R3b from Cr_{n+1} with respect to E_{n+2}. It can easily be shown that the same function Cr_{n+2} results if R3b is applied to Cr_n with respect to the combined data $E_{n+1} \cap E_{n+2}$. In the same way we can determine any later credence function Cr_{n+m} from the given function Cr_n either in *m* steps, applying the rule R3b in each step with one datum of the sequence $E_{n+1}, E_{n+2}, \ldots . E_{n+m}$, or in one step with the intersection $\bigcap_{p=1}^{m} E_{n+p}$. If *m* is large, so that the intersection contains thousands of single data, the objection might be raised that it is unrealistic to think of a procedure of this kind, because a man's memory is unable to retain and reproduce at will so many items. However, since our goal is not the psychology of actual human behavior in the field of inductive reasoning, but rather inductive logic as a system of rules, we do not aim at realism. We make the further idealization that *X* is not only perfectly rational but has also an infallible memory. Our assumptions deviate from reality very much if the observer and agent is a natural human being, but not so much if we think of *X* as a robot with organs of perception, data processing, decision making, and acting. Thinking about the design of a robot will help us in finding rules of rationality. Once found, these rules can be applied not only in the construction of a robot but also in advising human beings in their effort to make their decisions as rational as their limited abilities permit.

Consider now the whole sequence of data obtained by *X* up to the present time T_n: $E_1, E_2, \ldots, E_n$. Let K_{X,T_n} or, for short, K_n be the proposition representing the combination of all these data:

(10) DEFINITION.

$$K_n = \bigcap_{i=1}^{n} E_i.$$

Thus K_n represents, under the assumption of infallible memory, the total observational knowledge of X at the time T_n. *Now* consider the sequence of X's credence functions. In the case of a human being we would hesitate to ascribe to him a credence function at a very early time point, before his abilities of reason and deliberate action are sufficiently developed. But again we disregard this difficulty by thinking either of an idealized human baby or of a robot. We ascribe to him a credence function Cr_1 for the time point T_1; $Cr,$ represents X's personal probabilities based upon the datum E_1 as his only experience. Going even one step further, let us ascribe to him an *initial credence junction* Cr_0 for the time point T_0 before he obtains his first datum E_1. Any later function Cr_n for a time point T_n is uniquely determined by Cr_0 and K_n:

(11) For any H, $Cr_n(H) = Cr'_0(H|K_n)$, where Cr'_0 is the conditional function based on Cr_0.

$Cr_n(H)$ is thus seen to be the *conditional initial credence of* H given K_n.

How can we understand the function Cr_0? In terms of the robot, Cr_0 is the credence function that we originally build in and that he transforms step for step, with regard to the incoming data, into the later credence functions. In the case of a human being X, suppose that we find at the time T_n his credence function Cr_n. Then we can, under suitable conditions, reconstruct a sequence $E_1, \ldots, E_n$, the proposition K_n, and a function Cr_0 such that (a) $E_1, \ldots, E_n$ are possible observation data, (b) K_n is defined by (10), (c) Cr_0 satisfies all requirements of rationality for initial credence functions, and (d) the application of (11) to the assumed function Cr_0 and K_n would lead to the ascertained function Cr_n. We do not assert that X actually experienced the data $E_1, \ldots, E_n$, and that he actually had the initial credence function Cr_0, but merely that, under idealized conditions, his function Cr_n, could have evolved from Cr_0 by the effect of the data $E_1, \ldots, E_n$.

For the conditional initial credence (Cr'_0) we shall also use the term *"credibility*" and the symbol *'Cred'. As* an alternative to defining *'Cred'* on the basis of 'Cr_0', we could introduce it as a primitive term. In this case we may take the following universal statement as the main postulate for the theoretical primitive term *'Cred'*:

(12) Let *Cred* be any function from pairs of propositions to real numbers, satisfying all requirements which we have laid down or shall lay down for credibility functions. Let H and A be any propositions (A not empty). Let X be any observer and T any time point. If X's credibility function is *Cred* and his total observational knowledge at T is A, then his credence for H and T is $Cred(H|A)$.

Note that (12) is much more general than (11). There the function *Cred* (or Cr'_0) was applied only to those pairs H, A, in which A is a proposition of the sequence $K_1, K_2, \ldots,$ and thus represents the actual knowledge of X at some time point. In (12), however, A may be any non-empty proposition. Let A, be a certain proposition which does not occur in the sequence $K_1, K_2, \ldots,$ and H_1 some proposition. Then the statement

$$Cr_T(H_1) = Cred(H_1|A_1)$$

is to be understood as a counterfactual conditional as follows:

(13) If the total knowledge of X at T had been A_1, then his credence for H_1 at T would have been equal to $Cred\ (H_1|A_1)$.

This is a true counterfactual based on the postulate (12), analogous to ordinary counterfactuals based on physical laws.

Applying (12) to X's actual total observational knowledge $K_{X,T}$ at time T, we have:

$$\text{(14) For any } H,\ Cr_{X,T}(H) = Cred_X(H|K_{X,T}).$$

Now we can use credibility instead of credence in the definition of the subjective value of an act A_m, and thereby in the decision rule. Thus we have instead of (4):

(15) DEFINITION.

$$V_{X,T}(A_m) = \sum_n U_X(O_{m,n}) \times Cred_X\ (S_n|\ K_{X,T}),$$

(If the situation is such that the assumption of A_m could possibly change the credence of S_n, we have to replace '$K_{X.T}$' by '$K_{X,T} \cap A_m$.', see the remark on (1).)

If *Cred* is taken as primitive, *Cr.* can be defined as follows:

(16) DEFINITION. *For any* H, $Cr_o(H) = Cred\ (H|Z)$, *where* Z *is the necessary proposition (the tautology).*

This is the special case of (12) for the initial time T_o, when X's knowledge K_o is the tautology.

While $Cr_{X,T}$ characterizes the momentary state of X at time T with respect to his beliefs, his function $Cred_X$ is a trait of his underlying permanent intellectual character, namely his permanent disposition for forming beliefs on the basis of his observations.

Since each of the two functions Cr_0 and *Cred is* definable on the basis of the other one, there are two alternative equivalent procedures for specifying a basic belief-forming disposition, namely either by Cr_0 or by *Cred.*

Most of those who have constructed systems of subjective or personal probability (in the narrower sense, in contrast to logical probability), e.g., Ramsey, De Finetti, and Savage, have concentrated their attention on what we might call "adult" credence functions, i.e., those of persons sufficiently developed to communicate by language, to play games, make bets, etc., hence persons with an enormous amount of experience. In empirical decision theory it has great practical advantages to take adult persons as subjects of investigation, since it is relatively easy to determine their credence functions on the basis of their behavior with games, bets, and the like. When I propose to take as a basic concept, not adult credence but either initial credence or credibility, I must admit that these concepts are less realistic and remoter from overt behavior and may therefore appear as elusive and dubious. On the other hand, when we are interested in *rational* decision theory, these concepts have great methodological advantages. Only for these concepts, not for credence, can we find a sufficient number of requirements of rationality as a basis for the construction of a system of inductive logic.

If we look at the development of theories and concepts in various branches of science, we find frequently that it was possible to arrive at powerful laws of great generality only when the development of concepts, beginning with directly observable properties, had progressed step by step to more abstract concepts, connected only indirectly with observables. Thus physics proceeds from concepts describing visible motion of bodies to the concept of a momentary electric force, and then to the still more abstract concept of a permanent electric field. In the sphere of human action we have first concepts describing overt behavior, say of a boy who is offered the choice of an apple or an ice cream cone and takes the latter; then we introduce the concept of an underlying momentary inclination, in this case the momentary preference of ice cream over apple; and finally we form the abstract concept of an underlying permanent disposition, in our example the general utility function of the boy.

What I propose to do is simply to take the same step from momentary inclination to the permanent disposition for forming momentary inclinations also with the second concept occurring in the decision principle, namely, personal probability or degree of belief. This is the step from credence to credibility.

When we wish to judge the morality of a person, we do not simply look at some of his acts, we study rather his character, the system of his moral values, which is part of his utility function. Single acts without knowledge of motives give little basis for a judgment. Similarly, if we wish to judge the rationality of a person's beliefs, we should not simply look at his present beliefs. Beliefs without knowledge of the evidence out of which they arose tell us little. We must rather study the way in which the person forms his beliefs on the basis of evidence. In other words, we should study his credibility function, not simply his present credence function. For example, let *X* have the evidence *E* that from an urn containing white and black balls ten balls have been drawn, two of them white and eight black. Let *Y* have the evidence *E'* which is similar to *E,* but with seven balls white and three black. Let *H* be the prediction that the next ball drawn will be white. Suppose that for both *X* and Y the credence of *H is* ⅔. Then we would judge this same credence ⅔ to be unreasonable for *X*, but reasonable for *Y*. We would condemn a credibility function *Cred* as non-rational if $Cred(H|E) = ⅔$; while the result $Cred(H|\ E') = ⅔$ would be no ground for condemnation.

Suppose *X* has the credibility function *Cred,* which leads him, on the basis of his knowledge K_n at time T_n to the credence function Cr_n, and thereby, with his utility function *U*, to the act A_m, . If this act seems to us unreasonable in view of his evidence K_n and his utilities, we shall judge that *Cred* is non-rational. But for such a judgment on *Cred* it is not necessary that *X* is actually led to an unreasonable act. Suppose that for *E* and *H* as in the above example, K_n contains *E* and otherwise only evidence irrelevant for *H.* Then we have $Cr_n\ (H) = Cred\ (H|K_n) = Cred\ (H|E) = ⅔$; and this result seems unreasonable on the given evidence. If *X* bets on *H* with betting quotient ⅔, this bet is unreasonable, even if he wins it. But his credence is anyway unreasonable, no matter whether he acts on it or not. It is unreasonable because there are possible situations, no matter whether real or not, in which the result $Cred\ (H|E) = ⅔$ would lead him to an unreasonable act. Furthermore, it is not necessary for our condemnation of the function *Cred* that it actually leads to unreasonable *Cr*-values. Suppose that another man *X'* has the same function *Cred,* but is not led to the unreasonable *Cr*-value in the example, because he has an entirely different life history, and at no time is his total knowledge either *E* or a combination of *E* with data irrelevant for *H.* Then we would still condemn the function *Cred* and the man *X'* characterized by this function. Our argument would be as follows: if the total knowledge of *X'* had at some time been *E,* or *E* together with irrelevant data, then his credence for *H* would have had the unreasonable value ⅔. The same considerations hold, of course, for the initial credence function Cr_o corresponding to the function *Cred;* for, on the basis of any possible knowledge proposition *K,* Cr_0 and *Cred* would lead to the same credence function.

The following is an example of a requirement of rationality for Cr_0 (and hence for *Cred)* which has no analogue for credence functions. As we shall see later, this requirement leads to one of the most important axioms of inductive logic. (The term “individual” means “element of the universe of discourse”, or “element of the population” in the terminology of statistics.)

R4. *Requirement of symmetry. Let a_i and a_j be two distinct individuals. Let H and H' be two propositions such that H' results from H by taking a_j for a_i and vice versa. Then Cr_0 must be such that $Cr_o(H) = Cr_o(H')$.* (In other words, Cr_o must be invariant with respect to any finite permutation of individuals.)

This requirement seems indispensable. *H* and *H'* have exactly the same logical form; they differ merely by their reference to two distinct individuals. These individuals may happen to be quite different. But since their differences are not known to *X* at time T_o, they cannot have any influence on the Cr_0 -values of *H* and *H'*. But suppose that at a later time T_n, *X*'s knowledge K_n contains information *E* relevant to *H* and *H'*, say information making *H* more probable than *H'* (as an extreme case, *E* may imply that *H* is true and *H'* is false). Then *X*'s credence function Cr_n at T_n will have different values for *H* and for *H'*. Thus it is clear that R4 applies only to Cr_0, but is not generally valid for other credence functions Cr_n $(n > 0)$.

Suppose that *X* is a robot constructed by us. Because *H* and *H'* are alike in all their logical properties, it would be entirely arbitrary and therefore inadmissible =or us to assign to them different Cr_o-values.

A function Cr_0 *is* suitable for being built into a robot only if it fulfills the requirements of rationality; and most of these requirements (e.g., R4 and all those not yet mentioned) apply only to Cr_0 (and *Cred)* but not generally to other credence functions.

Now we are ready to take the step to *inductive logic.* This step consists in the transition from the concepts of the Cr_0-function and the *Cred*-function of an imaginary subject *X* to corresponding purely logical concepts. The former concepts are quasi-psychological; they are assigned to an imaginary subject *X* supposed to be equipped with perfect rationality and an unfailing memory; the logical concepts, in contrast, have nothing to do with observers and agents, whether natural or constructed, real or imaginary. For a logical function corresponding to Cr_0 , I shall use the symbol '*M*' and I call such functions (inductive) measure functions or M-functions; for a logical function corresponding to *Cred,* I shall use the symbol 'C' and I call these functions (inductive) confirmation functions or C-functions. I read *'C(H|E)'* as "the degree of confirmation (or briefly "the confirmation") of *H* with respect to *E"* (or: ". . . given E"). An M-function is a function from propositions to real numbers. A C-function is a function from pairs of propositions to real numbers. Any M-function M is supposed to be defined in a purely logical way, i.e., on the basis of concepts of logic (in the wide sense, including set-theory and hence the whole of pure mathematics). Therefore the value M (*A*) for any proposition *A* depends merely on the logical (set-theoretic) properties of *A* (which is a set in a probability space) but not on any contingent facts of nature (e.g., the truth of *A* or of other contingent propositions). Likewise any C-function is supposed to be defined in purely logical terms.

Inductive logic studies those M-functions which correspond to rational Cr_o-functions, and those C-functions which correspond to rational *Cred*-functions. Suppose M is a logically defined M-function. Let us imagine a subject *X* whose function Cr_o corresponds to M i.e., for every proposition *H,* $Cr_o(H) = M(H)$. If we find that Cr_0 violates one of the rationality requirements, say R4, then we would reject this function Cr_0, say for a robot we plan to build. Then we wish also to exclude the corresponding function M from those treated as admissible in the system of inductive logic we plan to construct. Therefore, we set up axioms of inductive logic about M-functions so that these axioms correspond to the requirements of rationality which we find in the theory of rational decision making about Cr_0 -functions.

For example, we shall lay down as the basic axioms of inductive logic those which say that M is a non-negative, finitely additive, and normalized measure function. These axioms correspond to the requirement R1 of coherence, by virtue of theorem (6). Further we shall have an axiom saying that M is regular. This axiom corresponds to the requirement R2 of strict coherence by theorem (9).

Then we shall have in inductive logic, in analogy to the requirement R4 of symmetry, the following:

(17) AXIOM OF SYMMETRY. M *is invariant with respect to any finite permutation of individuals.*

All axioms of inductive logic state relations among values of M or C as dependent only upon the logical properties and relations of the propositions involved (with respect to language-systems with specified logical and semantical rules). Inductive logic is the theory based upon these axioms. It may be regarded as a part of logic in view of the fact that the concepts occurring are logical concepts. It is an interesting result that this part of the theory of decision making, namely, the logical theory of the M-functions and the C-functions, can thus be separated from the rest. However, we should note that this logical theory deals only with the abstract, formal aspects of probability, and that the full meaning of (subjective) probability can be understood only in the wider context of decision theory through the connections between probability and the concepts of utility and rational action.

It is important to notice clearly the following distinction. While the *axioms* of inductive logic themselves are formulated in purely logical terms and do not refer to any contingent matters of fact, the *reasons* for our choice of the axioms are not purely logical. For example, when you ask me why I accept the axiom of symmetry (17), then I point out that if X had a Cr_0 function corresponding to an M-function violating (17), then this function Cr_0 would violate R4, and I show that therefore X, in a certain possible knowledge situation, would be led to an unreasonable decision. Thus, in order to give my reasons for the axiom, I move from pure logic to the context of decision theory and speak about beliefs, actions, possible losses, and the like. However, this is not in the field of empirical, but of rational decision theory. Therefore, in giving my reasons, I do not refer to particular empirical results concerning particular agents or particular states of nature and the like. Rather, I refer to a *conceivable* series of observations by X, to conceivable sets of possible acts, of possible states of nature, of possible outcomes of the acts, and the like. These features are characteristic for an analysis of *reasonableness* of a given function Cr_0, in contrast to an investigation of the *successfulness* of the (initial or later) credence function of a given person in the real world. Success depends upon the particular contingent circumstances, rationality does not.

There is a class of axioms of inductive logic which I call *axioms of invariance.* The axiom of symmetry is one of them. Another one says that M is invariant with respect to any finite permutation of attributes belonging to a family of attributes, e.g., colors, provided these attributes are alike in their logical (including semantical) properties. Still another one says that if E is a proposition about a finite sample from a population, then $M(E)$ is independent of the size of the population. These and other invariance axioms may be regarded as representing the valid core of the old *principle of indifference* (or principle of insufficient reason). The principle, in its original form, as used by Laplace and other authors in the classical period of the theory of probability, was certainly too strong. It was later correctly criticized by showing that it led to absurd results. However, I believe that the basic idea of the principle is sound. Our task is to restate it by specific restricted axioms.

It seems that most authors on subjective probability do not accept any axioms of invariance. In the case of those authors who take credence as their basic concept, e.g., Ramsey, De Finetti, and Savage, this is inevitable, since the invariance axioms do not hold for general credence functions. In order to obtain a stronger system, it is necessary to take as the basic concept either initial credence or credibility (or other concepts in terms of which these are definable).

When we construct an axiom system for M, then the addition of each new axiom has the effect of excluding certain M-functions. We accept an axiom if we recognize that the M-functions excluded by it correspond to non-rational Cr_0-functions. Even on the basis of all axioms which I would accept at the present time for a simple qualitative language (with one-place predicates only, without physical magnitudes), the number of admissible M-functions, i.e., those which satisfy all accepted axioms, is still infinite; but their class is immensely smaller than that of all coherent M-functions. There will presumably be further axioms, justified in the same way by considerations of rationality. We do not know today whether in this future development the number of admissible M-functions will always remain infinite or will become finite and possibly even be reduced to one. Therefore, at the present time I do not assert that there is only one rational Cr_0-function.

I think that the theory of the M- and C-functions deserves the often misused name of *"inductive logic"*. Earlier I gave my reasons for regarding this theory as a part of logic. The epithet "inductive" seems appropriate because this theory provides the foundation for inductive reasoning (in a wide sense). I agree in this view with John Maynard Keynes and Harold Jeffreys. However, it is important that we recognize clearly the essential form of inductive reasoning. It seems to me that the view of almost all writers on induction in the past and including the great majority of contemporary writers, contains one basic mistake. They regard inductive reasoning as an *inference* leading from some known propositions, called the premisses or evidence, to a new proposition, called the conclusion, usually a law or a singular prediction. From this point of view the result of any particular inductive reasoning is the *acceptance* of a new proposition (or its rejection, or its suspension until further evidence is found, as the case may be). This seems to me wrong. On the basis of this view it would be impossible to refute Hume's dictum that there are no rational reasons for induction. Suppose that I find in earlier weather reports that a weather situation like the one we have today has occurred one hundred times and that it was followed each time by rain the next morning. According to the customary view, on the basis of this evidence the "inductive method" entitles me to accept the prediction that it will rain tomorrow morning. (If you demur because the number one hundred is too small, change it to one hundred thousand or any number you like.) I would think instead that inductive reasoning about a proposition should lead, not to acceptance or rejection, but to the assignment of a number to the proposition, viz., its C-value. This difference may perhaps appear slight; in fact, however, it is essential. If, in accordance with the customary view, we accept the prediction, then Hume is certainly right in protesting that we have no rational reason for doing so, since, as everybody will agree, it is still possible that it will not rain tomorrow.

If, on the other hand, we adopt the new view of the nature of inductive reasoning, then the situation is quite different. In this case *X* does not assert the hypothesis *H* in question,. e.g., the prediction "it will rain tomorrow"; he asserts merely the following statements:

(18) (a) At the present moment T_n, the totality of *X*'s observation results is K_n.
(b) $C(H|K_n = 0.8$.
(c) $Cred_X(H|K_n = 0.8$.
(d) $Cr_{X,T_n}(H) = 0.8$.

(a) is the statement of the evidence at hand, the same as in the first case. But now, instead of accepting *H, X* asserts the statement (c) of the *Cred*-value for *H* on his evidence. (c) is the result of *X*'s inductive reasoning. Against this result Hume's objection does not hold, because *X* can give rational reasons for it. (c) is derived from (b) because *X* has chosen the function C as his credibility function. (b) is an analytic statement based on the definition of C. *X*'s choice of C was guided by the axioms of inductive logic. And for each of the axioms we can give reasons, namely, rationality requirements for credibility functions. Thus C represents a reasonable credibility function. Finally, *X*'s credence value (d) is derived from (c) by (14).

Now some philosophers, including some of my empiricist friends, would raise the following objection. If the result of inductive reasoning is merely an analytic statement (like (b) or (c)), then induction cannot fulfill its purpose of guiding our practical decisions. As a basis for a decision we need a statement with factual content. If the prediction *H* itself is not available, then we must use a statement of the *objective* probability of *H.* In answer to this objection I would first point out that *X* has a factual basis in his evidence, as stated in (a). And for the determination of a rational decision neither the acceptance of *H* nor knowledge of the objective probability of H is needed. The rational subjective probability, i.e., the credence as stated in (d), is sufficient for determining first the rational subjective value of each possible act by (15), and then a rational decision. Thus in our example, in view of (b) *X* would decide to make a bet on rain tomorrow if it were offered to him at odds of four to one or less, but not more.

The old puzzle of induction consists in the following dilemma. On the one hand we see that inductive reasoning is used by the scientist and the man in the street every day without apparent scruples; and we have the feeling that it is valid and indispensable. On the other hand, once Hume awakens our intellectual conscience, we find no answer to his objection. Who is right, the man of common sense or the critical philosopher? We see that, as so often, both are partially right. Hume's criticism of the customary forms of induction was correct. But still the basic idea of common sense thinking is vindicated: induction, if properly reformulated, can be shown to be valid by rational criteria.

REFERENCES

[1] CARNAP, R. *Logical foundations of probability.* Chicago, 1950.

[2] CARNAP, R. *The continuum of inductive methods.* Chicago, 1952.

[3] CARNAP, R. Inductive logic and rational decisions. (This is an expanded version of the present paper.) To appear as the introductory article in: *Studies in probability and inductive logic, Vol. 1,* R. Carnap, ed. Forthcoming.

[4] DAVIDSON, D., and P. SUPPES. *Decision making: An experimental approach.* Stanford, 1957.

[5] DE FINETTI, B. La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré, Vol.* 7 (1937), pp. 1-68.

[6] GOOD, 1. *J. Probability and the weighing of evidence.* London and New York, 1950.

[7] JEFFREYS, H. *Theory of probability.* Oxford (1939), 2nd ed. 1948.

[8] KEMENY, J. Fair bets and inductive probabilities. *Journal of Symbolic Logic,* Vol. 20 (1955), pp. 263-273.

[9] KEYNES, J. M. *A treatise on probability.* London and New York, 1921.

[10] KOLMOGOROFF, A. N. *Foundations of the theory of probability.* New York (1950), 2nd ed. 1956.

[11] KOOPMAN, B. O. The bases of probability. *Bull. Amer. Math. Soc., Vol.* 46 (1940), pp. 763-774.

[12] LEHMAN, R. S. On confirmation and rational betting. *Journal of Symbolic Logic,* Vol. 20 (1955), pp. 251-262.

[13] MOSTELLER, F. C., and P. NOGEE. An experimental measurement of utility. *Journal of Political Economy, Vol.* 59 (1951), pp. 371-404.

[14] NEUMANN, J. VON, and O. MORGENSTERN. *Theory of games and economic behavior.* Princeton (1944), 2nd ed. 1947.

[15] RAMSEY, F. P. *The foundations of mathematics and other logical essays.* London and New York, 1931.

[16] SAVAGE, L. J. *The foundations of statistics.* New York and London, 1954.

[17] SHIMONY, A. Coherence and the axioms of confirmation. *Journal of Symbolic Logic,* Vol. 20 (1955), pp. 1-28.

[18] WRIGHT, G. H. VON. *The logical problem of induction.* Oxford and New York (1941), 2nd ed., 1957.