terization of defensibility. Briefly, I avoid this problem by relativizing epistemic consistency to a given person *a* (more precisely, to the epistemic universe *Ua*) and revising the formal characterization of epistemic consistency. This formal characterization is given at the beginning of the appendix to this paper.

6 See H. E. Kyburg, Jr., 'Conjunctivitis', in this volume. Kyburg distinguishes 'weak' from 'strong' versions of several principles governing logical properties. In his discussion, these principles apply to sets of *reasonably accepted* sentences; I have taken the liberty of adapting his terminology to the sets I am calling 'epistemic universes'.

7 The concept of a doxastic alternative is borrowed from Hintikka [12], p. 49.

8 I am grateful to my colleague Brian Chellas for pointing out the potential of (C.dens).

9 Henry E. Kyburg, Jr. [17], pp. 197–8.

10 The term 'justified' that occurs in these principles is one that we have not dealt with in this paper. It is a term often used in discussions of rational belief, however. My argument is based partly on an ambiguity in this term and I have deliberately used it in order to keep my presentation of the lottery paradox as innocent as possible.

11 This suggestion is made, for example, by Schick, *op. cit.*

12 The relationship of accessibility is defined in the Appendix to this paper.

13 It is worth pointing out again that my claim here depends on the schematic explication of Reasonable given in Section I.B. If I were to change the definition of 'sustains' so that k can be *equal to* or greater than .5, then it seems doubtful that suspension of belief would ever be obligatory.

HENRY E. KYBURG, JR.

CONJUNCTIVITIS*

I

Consider a set S of statements that may be taken to represent an idealized body of scientific knowledge. Let s_1 and s_2 be members of S. Should we regard the conjunction of s_1 and s_2, also as a member of S? It is tempting to answer in the affirmative, and a number of writers, whose systems we shall consider below, have indeed answered this way. An affirmative answer is conjunctivitis, which may be expressed by the following principle:

The Conjunction Principle: If S is a body of reasonably accepted statements, and s_1 belongs to S and s_2 belongs to S, then the conjunction of s_1 and s_2 belongs to S.

This principle is clearly equivalent to the following principle.

The Conjunctive Closure Principle: If S is a body of reasonably accepted statements, then the conjunction of any finite number of members of S also belongs to S.

Already the intuitive plausibility of the conjunction principle begins to fade; while it seems reasonable enough to want to accept the conjunction of two relatively elementary statements that are individually acceptable, it seems quite unreasonable to accept all the enormously long conjunctions of elements in S. But the reasonableness or unreasonableness of the principle will depend, of course, on what other principles one also accepts.

One principle which is, so far as I know, universally accepted[1], is the principle that anything entailed by a member of S should also be a member of S. If it is reasonable to accept s_1, and s_1 entails s_2, then it is reasonable to accept s_2. I shall call this the weak deduction principle.

The Weak Deduction Principle: If S is a body of reasonably accepted statements, and s_1 belongs to S, and $s_1 \supset s_2$ is a theorem of our underlying logic, then s_2 belongs to S.

Another principle which I am sure is universally accepted is that the set of reasonably accepted statements S should contain no contradictions.

Whether or not it is psychologically possible to believe a contradiction (with practice perhaps it is), we do not want to regard it as rational. This I shall call the weak consistency principle.

The Weak Consistency Principle: If S is a body of reasonably accepted statements, then there is no member of S that entails every statement of the language.

In 1961 I offered an argument from the weak consistency principle and the weak deduction principle to the denial of the conjunction principle. This argument has come to be called the lottery paradox, and has engendered a number of principles designed to restrict the contents of S in such a way that all three of the principles mentioned so far hold.

The argument is this: Consider a fair lottery with a million tickets. Consider the hypothesis, 'ticket number 7 will not win'. Since this is, by hypothesis, a fair lottery, there is only one chance in a million that this hypothesis is false. Surely, I argued, this is reason enough to accept the hypothesis. But a similar argument would provide reason to accept the hypothesis that ticket i will not win, no matter what ticket number i may be. By the conjunction principle, we obtain from 'ticket 1 will not win' and 'ticket 2 will not win', the statement 'neither ticket 1 nor ticket 2 will win'; from the last statement, together with the statement 'ticket 3 will not win', by the conjunction principle, 'neither ticket 1, nor ticket 2, nor ticket 3 will win'; and so on, until we arrive at the reasonable acceptance of a long conjunction which can be briefly expressed as: 'For all i, if i is a number between one and a million, inclusive, ticket i will not win'. But we may also suppose that S contains the statement that the lottery is fair; and this statement entails the statement: 'For some i, i is a number between one and a million inclusive, and ticket i will win'. By the weak deduction principle we must therefore include this latter statement in S. By the conjunction principle we must therefore include the conjunction of the universally quantified statement and the existentially quantified statement in S. But this conjunction is an explicit contradiction from which any statement will follow in violation of the weak consistency principle. I concluded that it was worth while to hang onto the weak deduction principle and the weak consistency principle, and therefore that the conjunction principle should be abandoned.

Quite a number of people, finding the conjunction principle more plausible than I do, have attempted to spike this argument here or there.

One of the earliest attempts was made by Salmon [14], who suggested that one ought not to accept particular statements (such as 'Ticket 7 will not win the lottery'), but restrict one's acceptances to general statistical or universal generalizations. Since a number of writers have followed Salmon in this ploy, it is worth stating a statistical version of the same argument.[2]

Consider a finite population P of entities, each of which either has or lacks a certain quality Q. We draw a random sample (in any sense of 'random' you choose) of n of the P's. A certain proportion of the members of the sample, f, have the property Q; we know that in the parent population P some unknown proportion p have the property Q. Now consider hypotheses of the form 'p lies in the interval i'; for example: 'p lies in the interval $(f-.1, f+.1)$', 'p lies in the interval $(f-.0001, f+.675)$', etc. There are a number of principles of inference that one might adopt for arriving at acceptable statistical statements of this form. I shall consider two, though what I say will apply to other principles as well. Let us call them the Bayesian Acceptance Principle and the Classical Acceptance Principle.[3] In accordance with the Bayesian Principle, we shall accept a statistical hypothesis if its posterior probability is greater than $1-\varepsilon$, i.e., if the probability of its negation is less than ε. According to the Classical Principle, we will accept a hypothesis provided the probability of rejecting it by mistake is less than ε. Since the argument is slightly different in the two cases, I shall treat them separately.

Bayesian case: There are any number of intervals i such that the hypothesis '$p \in i$' is acceptable, under the assumption that the prior distribution of p is continuous between 0 and 1. Let i_c be the intersection of all these intervals. By the conjunction principle, '$p \in i_c$' is acceptable. Again under the assumption of continuity, it is possible to divide any interval i and in particular i_c into a finite number of subintervals $i_1, i_2, \ldots, i_m$, such that the posterior probability of '$p \in i_k$' is less than ε, for all k, $1 \leqslant k \leqslant m$. But this is just to say that the posterior probability of '$\sim p \in i_k$' is greater than $1-\varepsilon$ for all k, $1 \leqslant k \leqslant m$, and thus that the hypothesis '$\sim p \in i_k$' is acceptable for all k, $1 \leqslant k \leqslant m$. The conjunction principle then entails that '$\sim p \in i_c$' is acceptable, in virtue of the fact that $i_c = \bigcup i_k$. Thus by the conjunction principle we have '$p \in i_c \,\&\, \sim p \in i_c$' in our body of acceptable statements in violation of the weak consistency principle.

Classical case: Again there are any number of intervals i such that the probability that we will falsely reject the hypothesis '$p \in i$' when we observe f is less than ε. To be more precise (and more classical), to each of these intervals i_j there will correspond an interval E_j, such that if we reject the hypothesis '$p \in i_j$' if and only if the observed frequency f falls outside the interval E_j, then we will falsely reject the hypothesis no more than ε of the time. Of those hypotheses '$p \in i_j$' such that in point of fact f falls in the corresponding interval E_j, we say that they 'are not rejected at the ε level of significance'. In particular, to each hypothesis of the form '$p \in (a, b)$' there will correspond a test interval $(a-d, b+e)$ (it will always include the *closed* interval $[a, b]$), such that if we reject the hypothesis if and only if we observe a value of f not falling in the test interval, we shall falsely reject it less than ε of the time. Consider two hypotheses '$p \in (a, f)$' and '$p \in (f, b)$'. At any level of significance, the value f of the observed frequency will fall within the test interval corresponding to each of these hypotheses.[4] Thus at any level of significance we will accept both the hypothesis '$p \in (a, f)$' and the hypothesis '$p \in (f, b)$'. By the conjunction principle, we must then accept their conjunction. But there is no number that belongs both to (f, b) and (a, f); again we find a violation of the weak consistency principle.

II

Among those who have adopted principles of acceptance satisfying the principle of conjunction, the weak principle of consistency, and the weak principle of deduction, are included Hempel [3], Hintikka [7], Lehrer [12], and Levi [13]. In the ensuing sections I shall consider the principles proposed by each of these writers in turn, both from the viewpoint of strength and from the viewpoint of plausibility. We begin with Hempel. In his well known paper, 'Deductive-Nomological vs. Statistical Explanation', Hempel lays down certain "necessary conditions of rationality in the formation of beliefs". I shall discuss them one by one.

(CR1) Any logical consequence of a set of accepted statements is likewise an accepted statement; or, K contains all logical consequences of any of its subclasses.

An obviously equivalent formulation is the following:

The Principle of Deductive Closure: The set S is closed under deduction.

It should be observed that this principle is not equivalent to the weak deduction principle. Hempel's criterion entails the weak deduction principle, but the converse does not hold. For example, the set: {'S', '$S \supset P$', '$S \vee \phi$', '$\sim P \supset \sim S$', '$S \& T$', 'T',...} of which 'P' is not a member satisfies the weak principle, but not Hempel's strong principle. However, the conjunction principle, together with the weak deduction principle, are equivalent to the principle of deductive closure:

THEOREM 1: Principle of deductive closure≡(principle of conjunction and weak deduction principle)

PROOF: If S satisfies the principle of deductive closure, then if s_1 and s_2 belong to S, so do all their consequences; among their consequences is the conjunction of s_1 and s_2. Thus the principle of conjunction. And if S satisfies the principle of deductive closure, then if s_1 belongs to S, so does any consequence of s_1. Thus the weak principle of deduction. Now suppose S satisfies both the principle of conjunction and the weak deduction principle; then suppose that $s_1, \ldots, s_n$ each belong to S, and that s follows from $s_1, \ldots, s_n$ as premises. By the principle of conjunction, the conjunction of $s_1, \ldots, s_n$ belongs to S; by the standard deduction theorem, plus the weak deduction principle it then follows that s belongs to S.

What is questionable about the strong principle of deductive closure is, I think, precisely the principle of conjunction. This may involve a matter of intuition: I simply don't believe that everything I believe on good grounds is true, and I think it would be irrational for me to believe that. Other people seem to think the opposite. I suspect that at root there is a confusion of quantifiers: of everything that I believe, it is correct to say that I believe it to be true; but it is not correct to say that I believe everything I believe to be true. In symbols:

(1) (x) (I believe $x \supset$ I believe x to be true)
(2) I believe (x) (I believe $x \supset x$ is true).

Statement (1) seems true, statement (2) false.

Hempel's second condition of rationality is a consistency condition

(CR2) The set K of accepted statements is logically consistent.

Although it might be possible to construe this as the weak principle of consistency above, it is more likely that Hempel has in mind a strong principle of consistency:

Strong Principle of Consistency: If S is a body of reasonably accepted statements, then there is no finite subset of S, $s_1, \ldots, s_n$, such that every statement of the language follows from $s_1, \ldots, s_n$ as premises.

Again conjunction plays an important role. In the presence of the conjunction principle, the strong principle of consistency and the weak principle of consistency are equivalent.

THEOREM 2: Conjunction principle $\supset$ [strong principle of consistency $\equiv$ weak principle of consistency].

Again one can question the plausibility of the strong principle. I probably cannot believe a contradiction, or act on one. But I can certainly believe, and even act on, each of a set of statements which, taken conjointly, is inconsistent. Indeed, when I lend my moral support to a lottery, without buying a ticket, this is one way to describe what is going on; though a more adequate and complete description would involve expectation.

Hempel's third criterion is, he claims, "simply a restatement of the requirement of total evidence" (p. 151).

(CR3) The inferential acceptance of any statement h into K is decided on by reference to the total system K.

This criterion, though it is not strictly relevant to the problem of conjunctivitis, is worth a comment or two. It is clearly in conflict with the principles of fallibilism and empiricism which underly much of what Hempel writes. Surely it is a consequence of these principles that even if a statement s becomes a part of our body of reasonable beliefs, we will, if the evidence begins to go the other way, be ready to reject it. But if we incorporate the statement h into the body of beliefs K, then, whatever else we add to that body of beliefs, its probability, relative to that body of beliefs, is going to be unity. Once accepted, no added evidence can ever render h improbable. The suggestion – which requires a great deal of development – is that we shall have to keep our evidential base separate from the body of reasonable beliefs erected on that base. There is also a new path to inconsistency opened by this principle. Let us suppose that there is a set P, 99% of which are Q_1; that 99% of the Q_1 are Q_2; ... that 99% of the Q_n are R. There is nothing inconsistent in supposing this, and that at the same time 99% of the P are not-R. But there is obviously something inconsistent about accepting the statement that an arbitrary P is not-R (since it is a P and that's all we know about it), and accepting the statement that an arbitrary P is R (since it is a P and that's all we know about it).

Hempel finally proposes a measure of epistemic utility (related to the content of a statement in terms of a logical measure function) according to which the principle of maximizing utility leads to the rule:

> Tentative rule for inductive acceptance: Accept or reject h, given K, according as $c(h, K) > 1/2$ or $c(h, K) < 1/2$; when $c(h, K) = 1/2$, h may be accepted, rejected, or left in suspense (p. 155).

It is clear that this rule is not only, as Hempel puts it, "too liberal", but leads directly, through the strong principle of deduction, to inconsistency, provided only that there are three hypotheses, jointly exhaustive, whose probabilities, relative to k are all less than a half.

III

A recent paper by Hilpinen and Hintikka develops an inductive acceptance rule that is demonstrably consistent, and satisfies a number of criteria similar to those discussed above. Their system uses Hintikka's 1965 system of inductive logic [4], in which non-zero degrees of confirmation come to be assigned to general sentences. This system of inductive logic is applicable in principle to all first order languages (without identity), but is developed in detail only for monadic languages. In point of fact, the system developed by Hilpinen and Hintikka satisfies the strong consistency principle, and the strong principle of deductive closure. Put in terms of the most finely articulated statements above: their system satisfies the weak deduction principle, the weak consistency principle, and the conjunction principle. (Since these three principles are independent, it seems best to refer to them separately.)

Since Hintikka's system is not as well known as it should be, a brief review of its features may be helpful here. Consider a language L, containing k primitive monadic predicates 'P_i'. By means of these predicates, one can characterize $K = 2^k$ kinds of individuals, corresponding to Carnap's 2^k Q-predicates. For the sake of simplicity we suppose that instantiation of each of the Q-predicates is logically possible, or in Carnap's terms, that the primitive predicates are logically independent. These Q-predicates are called by Hintikka 'attributive constituents'. A *constituent* consists of a specification, for each of the K Q-predicates or attributive constituents, of whether or not it is instantiated in the world. There are 2^K-1 possible different constituents (because it is logically false that *no* constituent be instantiated in a non-empty universe).

There are various ways of assigning a priori probabilities to the constituents. One might take the probability of a constituent to be proportional to $(w/K)^\alpha$, where α is an arbitrary constant, and w is the number of Q-predicates that are alleged to be instantiated by the constituent. For our purposes the simpler assignment of equal probabilities to each constituent (also worked out by Hintikka and Hilpinen) will suffice.

Let e be a sentence that asserts, for each member of a sample of n individuals, and for each of our primitive predicates 'P_i', either that that individual has the property P_i or that it has the property $\sim P_i$. Since we may order the Q-predicates in an arbitrary way there is no loss of generality in supposing that our sample of n individuals provides instantiation of the first c Q-predicates. The constituents that are consistent with our evidence all have the form:

$$(\exists x)\, Q_1(x)\ \&\ (\exists x)\, (Q_2(x))\ \&\cdots\&\ (\exists x)\, Q_c(x)\ \&\ (\exists x)\, Q_{i_1}(x)\ \&\ (\exists x)\, Q_{i_2}(x)\ \&\cdots\&\ (\exists x)\, Q_{i_m}(x)\ \&\ (x)\, (Q_1(x) \vee Q_2(x) \vee \cdots Q_c(x) \vee Q_{i_1}(x) \vee \cdots \vee Q_{i_m}(x)),$$

where $i_j < i_k$ whenever $j < k$, and $i_1 > c$.

Let C_c be the constituent that asserts that just those Q-predicates instantiated by our evidence are exemplified in the universe (i.e., C_c='$(\exists x)\, Q_1(x)\ \&\cdots\&\ (\exists x)\, Q_c(x)\ \&\ (x)\, (Q_1(x) \vee \cdots \vee Q_c(x))$', and let C_w be any other constituent consistent with our evidence. Then it is possible to show:

(1) $P(C_c, e) > P(C_w, e)$

(2) $\lim_{n\to\infty} P(C_c, e) = 1$

(3) $\lim_{n\to\infty} P(C_w, e) = 0.$

One further fact is important:

(4) Every consistent general sentence h of L_k can be transformed into a disjunction of constituents; thus $P(h, e) = \sum_i P(C_i, e)$, where the summation is extended over all those constituents in the disjunction equivalent to h.

It is possible to show ((2) and (3) give an intuitive justification) that for given $\varepsilon < 1/2$, we can calculate an integer n_0 such that if $n > n_0$, one and only one constituent will have a probability, relative to e, greater than $1-\varepsilon$. This constituent will of course be C_c.

We now adopt the following for our rule of acceptance: Accept a *general* statement h, given evidence e, if and only if: the probability of h is greater than $1-\varepsilon$, and more than n_0 objects have been examined. Formally (p. 11):

(D.Ac) $Ac(h, e) =_{\text{Df}}$ (i) $P(h, e) > 1 - \varepsilon$, where $0 < \varepsilon \leqslant 0.5$
(ii) $n > n_0$.

In virtue of (4) a hypothesis h can have a probability greater than $1-\varepsilon$ only if the constituent C_c appears in its distributive normal form; thus every hypothesis that is acceptable must be consistent with C_c; and thus they must be jointly consistent. The strong consistency condition, and thus the weak consistency condition also, is satisfied.

The principle of conjunction is also satisfied. If h_i and h_j are acceptable, then C_c must occur in the distributive normal form of each of them, and thus C_c will occur in the distributive normal form of their conjunction. The probability of their conjunction will therefore be greater than $1-\varepsilon$, and their conjunction will therefore be acceptable.

The principle of deductive closure so far fails, however, even in its weak form. Let '$(x)A(x)$' be an acceptable generalization; one of its deductive consequences is '$A(a)$'. But there is no clause in (D.Ac) that will allow us to accept '$A(a)$'.

Deductive closure does hold, however, for *general* statements. Suppose that the general statements $h_1, \ldots, h_n$ are all acceptable, and that they entail a factual general statement h. Since C_c occurs in the distributive normal form of each h_i, C_c must also occur in the distributive normal form of h itself. Thus h must be acceptable, and, so far as general statements h (containing no individual constants) are concerned, the strong principle of deductive closure is satisfied.

THEOREM 3: If K is the set of all those statements h such that $Ac(h, e)$, then K satisfies the conjunction principle and the strong consistency principle, but neither the strong nor the weak deduction principle. The strong and weak deduction principles are satisfied if we restrict their range of application to completely general sentences.

The system so far provides us with no way of accepting singular statements of the form $M(a_i)$, where M is a molecular predicate equivalent to a disjunction of Q-predicates, and a_i an individual constant. Hintikka and Hilpinen show that in order to preserve consistency, our rule for the acceptance of singular predictions must be:

(D.Ac sing) A singular hypothesis $(A(a_i)$ is acceptable if and only if the generalization $(x)A(x)$ is acceptable (p. 18).

Let A_c be the disjunction of the c Q-predicates in C_c. '$(x)A_c(x)$' is acceptable, and the probability of the conjunction of any arbitrary number of statements of the form $A_c(a_i)$ is at least $1-\varepsilon$. Indeed:

$$\lim_{r\to\infty} P[A_c(a_1)\,\&\cdots\&\,A_c(a_r)] = P(C_c, e)$$

Let us look at consistency and deductive closure. Suppose the singular statements '$A_1(a_1)$', '$A_2(a_2)$',...'$A_n(a_n)$' are acceptable as being instances of acceptable universal generalizations, and that the compound statement S is deducible from them. Since '$(x)A_c(x)$' is the strongest universal generalization, '$A_1(a_1)$' is deducible from '$A_c(a_1)$', '$A_2(a_2)$' is deducible from '$A_c(a_2)$' etc., and the statement S is deducible from the conjunction '$A_c(a_1)\,\&\,A_c(a_2)\&\cdots\&A_c(a_n)$', therefore S is deducible from '$(x)A_c(x)$'. This ensures that the probability of S will be at least $1-\varepsilon$. But we have no grounds for either asserting or denying that S is acceptable; acceptability has been defined only for completely general statements, and for singular statements of the form $A(a)$. Deductive closure thus far fails. Acceptability in general, however, would presumably be defined in this way:

(Acc) A statement S is acceptable if and only if it belongs to every class K of statements, closed under deduction, containing e, such that each member of K that is completely general is acceptable by (D.Ac), and such that each statement acceptable by (D.Ac) is a member of K.

The separate principle (D.Ac sing) is deducible from this principle.

THEOREM 4: If K is the set of all statements h, such that h is acceptable by the principle (Acc), then K satisfies the conjunction principle, the strong consistency principle, and the strong deduction principle.

PROOF:

(a) Conjunction Principle: The conjunction of s_1 and s_2 belongs to every class K closed under deduction to which both s_1 and s_2 belong, and therefore to every class K closed under deduction and meeting further requirements as well to which both s_1 and s_2 belong.

(b) The Strong Consistency Principle: By hypothesis e is consistent, and consistent with the strongest acceptable generalization, '$(x)Ac(x)$', Other statements are obtained only by deduction, but deduction cannot introduce inconsistency.

(c) The Strong Deduction Principle: Since the classes K are closed under deduction the deduction principle is satisfied automatically.

The set of statements characterized as acceptable by this principle consists essentially of C_c and e, together with all of their deductive consequences. This set is essentially the set of deductive consequences of a single statement. We shall find this to be characteristic of those acceptance rules for which deductive closure holds, and which are demonstrably consistent.

There are certain shortcomings to the system of Hintikka and Hilpinen. In the first place, as it stands, it does not allow us to take account of the relative frequencies with which the Q-predicates are exemplified. In a similar system proposed by Hintikka [6] the probability that an individual a will have a certain molecular property M will depend on the relative frequencies with which Q-predicates have been observed to be exemplified; but in the system under discussion we can never accept $A(a)$ unless the universal generalization $(x)M(x)$ is acceptable. This is a general feature of these systems, and necessarily so. All of these systems accept the principle of conjunction, and it follows directly from the principle of conjunction that if an arbitrary predictive inference of the form '$A(a)$' is acceptable, then its universal generalization '$(x)A(x)$' is acceptable, at least when restricted in scope to the unobserved part of the universe of discourse. This should be stated formally.

THEOREM 5: If the principle of conjunction is accepted, and if for any arbitrary individual a among the unobserved individuals, the singular predictive inference '$A(a)$' is acceptable, then the universal generalization '$(x)A(x)$' is acceptable, when restricted in scope to unobserved individuals.

PROOF: Let the evidence e mention only k individuals, as failing to satisfy the predicate 'A'; let these individuals be $a_1 \ldots a_k$ (k may be 0). Then if the principle of conjunction is accepted, then if in general any singular predictive inference of the form '$A(a_i)$' $(i>k)$ is acceptable, then every finite conjunction of statements of this form is acceptable, and the universal generalization '$(x)\,(\sim x \in \{a_1, a_2, \ldots, a_k\} \supset (A(x)))$' is acceptable in a finite language. For an infinite language, we need only note that '$A(a_{k+1})$' is acceptable, and that if the restricted universal generalization is acceptable for all individuals up to the nth, then by the principle of conjunction it is acceptable also for all individuals up to the $(n+1)$st. A plausible induction principle yields the conclusion. In systems such as Hintikka's without identity we cannot express a generalization that takes account of exceptions; but the same result follows in those systems for cases when $k=0$.

The problem of extending the system of Hintikka and Hilpinen to a full first order logic is very knotty. Some steps have been taken in this direction by Hintikka, who has defined constituents quite generally. Hintikka's approach to inductive logic is applicable in principle to all first order logics, though the definition of (Acc) would obviously have to be enormously complicated even to deal with a language containing two-place predicates. Tuomela [15] has begun the attempt to construct an inductive logic for monadic languages with identity. But in virtue of the fact that the essence of the acceptance rule is that we are directed to accept a single statement $(C_c \& e)$, together with all of its deductive consequences, any such system will be open to objections of the sort that will be applied below to proposals of Levi and Lehrer.

Finally, it should be observed that in order to do statistical inference, we need a general higher order logic. In particular, we need to be able to speak of the set of n-membered subsets of a set S, if not in general, at least for sets of fairly high order. We also need a language rich enough for measure theory. Such a language, of course, is enormously more powerful than anything hitherto considered by Hintikka, Hilpinen, or Tuomela. To be sure one must start somewhere. But it is difficult to see what principles, analoguous to those adopted for the monadic predicate calculus, could be used to avoid the statistical versions of the lottery paradox described above.

IV

The system of inductive acceptance described by Levi [13] has the overwhelming advantage of being applicable to very rich languages. It is thus the sort of system we can apply in the kinds of circumstances that we can actually find ourselves in. It can be applied, for example, to the problem

of accepting statistical hypotheses on given evidence. Two new concepts must be explained before the rule can be stated. The most important and most novel concept is that of an *ultimate partition.* Levi argues that the inductive inference maker does not conduct his inquiries in a vacuum; he does so in a context determined in part by a felt need, a problem which he is seeking to solve. Thus before the inquiry starts, an investigator has an idea, which may be quite clear, or, unanalyzed, may be rather confused, as to what would constitute a relevant answer to his problem. These answers can be related to relatively atomic relevant answers, which constitute an ultimate partition, U in the following way. Following Levi, we suppose that there is a certain statement b which represents background information, not up for test in the inquiry, and a certain statement e, which represents the body of evidence of the inquiry. The ultimate partition U_e is a (usually finite) set of sentences in L (the language of the inquiry, which may be as rich as you please), such that each element of U_e is consistent with b and e, and such that the conjunction of b and e entails (i) that some member of U_e is true, and (ii) that at most one member of U_e is true, and (iii) every relevant answer is logically equivalent to the disjunction of zero or more members of U_e, where we understand the disjunction of zero members of U_e to be the conjunction (inconsistent with b and e) of all the members of U_e. The set of sentences M_e represents the canonical standardized list of relevant answers. For an ultimate partition containing n members, M_e is the set of 2^n statements, constructed by forming the disjunction (in alphabetical order) of m $(1 \leqslant m < n)$ elements of U_e, and adding to that list S_e, the disjunction of all the members of U_e and C_e, the conjunction of all the members of U_e. The subscript 'e' reflects the evidence e. An initially ultimate partition would be designated by 'U', and its set of ultimate answers in canonical form by 'M'. Given some evidence e, the initially ultimate partition would be reduced by the deletion of any elements of U that were inconsistent with e. This produces a truncated ultimate partition U_e, and leads to a new (and correspondingly truncated) set of relevant answers M_e.

The other crucial concept is content. Content is defined relative to ultimate partitions. Each element of the initial ultimate partition is taken to have the same content, on the grounds that any difference in content would lead to a finer ultimate partition. To arrive at the conditional content of a hypothesis H, given certain evidence e, one merely applies the same principle to the truncated ultimate partition: one takes each element of U_e to have the same content. This leads to the conclusion that in general,

> cont$(H, e) = m/n$, where n is the number of elements in U_e, and m is the number of elements in U_e that are inconsistent with H (p. 70).

With these two concepts at hand, we can state Levi's Inductive Acceptance Rule (p. 86).

Rule A: (a) Accept b & e and all its deductive consequences.

(b) Reject all elements a_i of U_e, such that $p(a_i, e) < q$ cont $(\sim a_i, e)$ i.e., accept the disjunction of all unrejected elements of U_e as the strongest element in M_e accepted via induction from b & e.

(c) Conjoin the sentence accepted as strongest via induction according to (b) with the total evidence b & e and accept all deductive consequences.

(d) Do not accept (relative to b, e, U_e, the probability distribution, and q) any sentences other than these in your language.

The number q referred to in the acceptance rule may be construed as an index of boldness; it ranges from 0 to 1, and the larger it is, the less cautious will one be in accepting statements not entailed by the background knowledge and evidence b & e. The number q reflects the "relative importance of the two desiderata: truth and relief from agnosticism".

As it stands, rule A obviously leads to the acceptance of a set of statements satisfying the strong principle of deductive closure (indeed Levi takes the condition of deductive cogency, as he calls it, as a condition of adequacy), and thus also the conjunction principle; the strong consistency principle is satisfied because q must be less than or equal to 1, and the probability level at which elements of U_e are rejected must therefore be less than $1/n$, where n is the number of elements of U_e. Even if q is taken to be 1, and the number of elements n of U_e is taken to be 2, the inequality in (b) preserves us from inconsistency. The general principle is the same as that embodied in Hintikka's system already discussed: what is accepted is essentially a single (strongest) hypothesis H, together with its deductive consequences.

Unlike Hintikka's system, in Levi's system a high probability is not necessary for acceptance. Levi's system has a rule of rejection which is not purely probabilistic, but is dependent on content (as determined by the number of elements in the ultimate partition) and on the index of caution q. One may perfectly well end up accepting a proposition on the evidence whose probability relative to that evidence is less than 1/2, so long as it is more probable than any competing alternative. The canonical example is that in which there are three elements in U_e, each (therefore) having content 1/3, q is 1, and the probabilities of a_1, a_2 and a_3 are respectively 0.4, 0.3, and 0.3. We are directed to accept a_1 and all the deductive consequences of $a_1 \& b \& e$. Strong consistency is nevertheless preserved.

While Hintikka's system appears to be the prototype of a global approach to problems of belief and acceptance, and thus to be limited by the choice of a language, and open to the difficulties of attempting to develop a similar system for richer languages, Levi's system is frankly local and problem dependent. Another way of construing the relation is to say that for Hintikka, the language we use determines the ultimate partition, while for Levi the ultimate partition is determined both by a language and by a particular problem; or perhaps it could be put this way: the partition is determined by a particular problem together with the language in which we represent that problem to ourselves.

The fact that Levi's rule of acceptance is relativized to a given language L, background knowledge b, evidence e, and a probability distribution P, raises no eyebrows; we surely expect that what a plausible acceptance rule will dictate will depend on these factors. That what we accept should depend in some way on how cautious or bold we are being, as expressed by the number q, also seems reasonable enough. The important question concerns the relativization to an ultimate partition. Given the language, b, e, the probability distribution and q, it is clear that different ultimate partitions may lead to the acceptance of different sets of statements. It would be blatantly contrary to the whole pragmatic spirit of Levi's approach to suppose that there is some special, preferred, universal, ultimately ultimate partition. All ultimate partitions must be treated on a par, though at a given time, under given circumstances, we may not *consider* or *think about* more than one. But we could consider several. Levi asks rhetorically, "...why is it impossible for conditions K and K'

to prevail at the same time, such that a man believes a deductively consistent and closed system of sentences Γ based on K and simultaneously a deductively consistent and closed system of sentences Γ' based on K'? To be sure he will believe, and believe rationally, all sentences in the set $\Gamma \cup \Gamma'$, and this set may very well be neither consistent nor [deductively] closed. But ... why should this be objectionable?" (p. 94).

One answer to this rhetorical question might consist in quoting the arguments Levi aduces elsewhere in his book in favor of the Principle of Deductive Cogency. Whatever reason there is for demanding that the set Γ based on K should be deductively consistent and closed, are these not also reasons for demanding that $\Gamma \cup \Gamma'$ be deductively consistent and closed? But I rather agree with Levi that there is nothing at all objectionable about a man believing rationally all the sentences in the set $\Gamma \cup \Gamma'$, where this set is neither consistent nor deductively closed. Indeed, since for any hypothesis that has a high probability, we can construct an ultimate partition which will lead to the acceptance of that hypothesis, we are in essentially the same state, so far as statements with high probability are concerned, as we are in one of my rational corpora. Levi points out (p. 95) that his Rule A "... requires that this set [of rationally accepted sentences] must be divisible into subsets which are consistent and closed relative to the total evidence and the ultimate partitions detached at that time". But since we can detach any partition we want, this latter requirement is empty; and it is trivial to divide any set of sentences satisfying my conditions for a rational corpus into subsets that are deductively cogent. Indeed, it is trivial to do this for any purely probabilistic rule of acceptance.

THEOREM 6: If $S \in K$ if and only if $\text{Prob}(S, e) > r$, then K may be divided into subsets which are consistent and closed – i.e., for which the strong consistency principle and the strong deduction principle are satisfied.

PROOF: For each sentence S of K, let K_S consist of S, together with all of its deductive consequences. The K_S are the required division. If $T \in K_S$, then T is entailed by S. But if S entails T, the probability of T cannot be less than that of S (on *any* interpretation of probability) and so T belongs to K. Conversely if $S \in K$, then $S \in K_S$. So $K = \bigcup K_S$. The sets K_S are deductively closed; they satisfy the Conjunction Principle. It remains to show that they are consistent. An inconsistency can appear only if it is entailed by some sentence S of K. But the probability of an inconsistency (on any interpretation of probability) is 0, and if S entails it, the probability of S cannot be greater than 0; and so S cannot be a member of K on purely probabilistic grounds.

Let us examine the set of statements that might come to be accepted (by means of some ultimate partition or other) in Levi's system. Levi points out that this set of statements will not in general satisfy either the strong deduction principle or the strong consistency principle. It is interesting to observe that these sets of statements do satisfy the weak deduction principle and the weak consistency principle.

THEOREM 7: If Γ_ω is the set of all those statements S such that there is an ultimate partition U_e, relative to which, and b, e, and q, S comes to be accepted by Rule (A), then Γ_ω satisfies the weak deduction principle and the weak consistency principle.

PROOF: Every statement that comes to be accepted in this set of statements is accepted originally in relation to some ultimate partition. But Rule A demands that when S is accepted relative to some ultimate partition, all of the deductive consequences of $S \& b \& e$ should also be accepted; thus the weak deduction principle. As for consistency, if S is inconsistent in itself, it can never come to be accepted by Rule A, and thus cannot be a member of the set of accepted statements relative to any ultimate partition.

There is one further kind of consistency we might ask about. The pair of statements $(S, \sim S)$ is perfectly consistent in the weak sense – i.e., it contains no inconsistent statement among its elements – and yet seems rather flatly wrong, somehow. What I shall call the Principle of Pairwise Consistency stipulates that a body of accepted statements should not contain any such pairs of statements.

Principle of Pairwise Consistency: If K is a body of reasonably accepted statements, then for no statement S of the language is it the case that both S and the denial of S belong to K.

In a similar manner we may define for every n a principle of n-wise consistency:

Principle of n-wise Consistency: If K is a body of reasonably accepted statements, then for no set of statements $s_1, s_2, \ldots, s_{n-1}$ is it the case that each of $s_1, s_2, \ldots, s_{n-1}$, and $\sim (s_1 \& s_2, \ldots, s_{n-1})$ is a member of K.

We should first observe that any purely probabilitistic acceptance rule (with acceptance level greater than 1/2) satisfies the Principle of Pairwise Consistency, and any purely probabilistic acceptance rule with acceptance level greater than $n/(n+1)$ satisfies the Principle of n-wise Consistency.

THEOREM 8: If $S \in K$ if and only if $\text{Prob}(S, e) > r$ (where $r > 1/2$), then K satisfies the Principle of Pairwise Consistency. If $S \in K$ if and only if $\text{Prob}(S, e) > n/(n+1)$ then K satisfies the Principle of n-wise Consistency.

PROOF: For ordinary point probabilities, the probability of S is one minus the probability of $\sim S$, so both probabilities cannot be greater than 1/2. For my interval probabilities, it is possible to show that $\text{Prob}(S) = (p, q)$ if and only if $\text{Prob}(\sim S) = (1-q, 1-p)$; so the same argument goes through.

$$\text{Prob} \sim [S_1 \;\&\; S_2 \;\&\cdots\&\; S_{n-1}] = \text{Prob}[\sim S_1 \vee \sim S_2 \vee \cdots \vee \sim S_{n-1}]$$

$$\text{Prob}[\sim S_1 \vee \sim S_2 \vee \cdots \vee \sim S_{n-1}] \leqslant \text{Prob}(\sim S_1) + \text{Prob}[\sim S_2 \vee \sim S_2 \vee \cdots \sim S_{n-1}]$$
$$\text{Prob}(\sim S_1) + \text{Prob}(\sim S_2) + \cdots + \text{Prob}(\sim S_{n-1})$$

$$\text{Prob}(S_i) \geqslant n/(n+1);\ \text{Prob}(\sim S_i) \leqslant 1 - n/(n+1) = 1/(n+1)$$

$$\text{Prob} \sim [S_1 \;\&\; S_2 \;\&\cdots\&\; S_{n-1}] \leqslant (n-1)\ 1/(n+1) = (n-1)/(n+1) < n/(n+1).$$

Therefore $\sim [S_1 \;\&\; S_2 \;\&\cdots\&\; S_{n-1}]$ is not probabilistically acceptable.

Now let us observe that Levi's general sets of rationally accepted statements, although they do in fact satisfy the weak consistency principle, do not satisfy the Pairwise Consistency Principle. Consider a three ticket biased lottery, in which ticket 1 has the probability 0.4 of winning, and the tickets 2 and 3 have the probability 0.3 of winning. Relative to the ultimate partition: {Ticket 1 wins, ticket 2 wins, ticket 3 wins}, we will be able to accept the statement 'ticket 1 wins' together with all its deductive consequences. Relative to the ultimate partition {ticket 1 wins, ticket 1 does not win}, we will be able to accept the statement 'ticket one does not win' together with all of its deductive consequences. Thus in the union of the statements accepted relative to each of these ultimate partitions, we will find both a certain statement and its denial (We assume $q = 1$.)

It is interesting to observe that while Levi takes me to task for abandoning the rule of conjunction in my system (and with it the strong deduction principle and the strong consistency principle), the set of rationally accepted beliefs that he comes up with not only abandons all three of these principles (conjunction, strong deduction, strong consistency) but also fails to satisfy the pairwise consistency principle, which my own system, like that of any other purely probabilistic system, does satisfy. It is also interesting to note that the requirements of Rule A, construed locally and not globally, are satisfied by *any* probabilistic rule of acceptance (including mine).

V

Keith Lehrer discusses a purely probabilistic rule of inductive inference, and what amounts to the strong consistency condition. He shudders with horror at what he regards as an abandonment of consistency, and attempts to provide a rule of inductive inference which will satisfy the three strongest principles: the conjunction principle, the strong principle of consistency, and the strong principle of deduction. The principle he comes up with contains as parameters P, an appropriate probability function, e, a body of evidence, and L, a formal language. There is no parameter corresponding to degree of caution, or level of practical certainty. There is no relativization to a given set of hypotheses. The rule is:

RDI: $D(k, e)$ [k is directly inducible from e] if, for any h, if it is not the case that $\ulcorner e \;\&\; k \urcorner \vdash h$, then $P(k, e) > P(h, e)$.

IR: $I(h, e)$ [h is inducible from e] if and only if h is a member of a set I_e such that $I_e = I_1 \cup I_2 \cup \cdots \cup I_m$, [where]
$I_1 =$ the set of hypotheses h_j such that $D(h_j, e)$,
and letting C_n be a conjunction of the members of I_n that is logically equivalent to I_n,
$I_2 =$ the set of hypotheses h_j such that $D(h_j, C_1 \;\&\; e)$
$I_3 =$ the set of hypotheses h_j such that $D(h_j, C_2 \;\&\; e)$
...
$I_m =$ the set of hypotheses h_j such that $D(h_j, C_{m-1} \;\&\; e)$.

We may elucidate the 'logical equivalence' of a statement (C_n) and a set of statements (I_n) in the obvious way: C_n is derivable from the set of statements I_n, and every member of I_n is derivable from C_n. It might be wondered if, for every I_n, there is a finite conjunction of members of I_n from which every member may be derived. In a finite language this is the case; in an infinite language C_n may turn out to be an infinite conjunction. It is possible to prove that the set of statements induced by IR from consistent evidence satisfies the conjunction principle (if h and k are inducible from e by IR, so is their conjunction); that this set of statements is deductively closed (if $h_1, \ldots, h_n$ are inducible from e by IR, and $\ulcorner h_1, \ldots, h_n \urcorner \vdash k$, then k is inducible from e); and that it is strongly consistent (it contains no set of statements $h_1, \ldots, h_n$ such that $\ulcorner h_1, \ldots, h_n \urcorner \vdash \ulcorner k \& \sim k) \urcorner$.

But we have achieved our goal of an inductive rule that satisfies these strong principles at essentially the same cost as that paid for Hintikka and Hilpinen's rule, namely: there is essentially only *one* hypothesis that we may induce from given evidence. Anything else we are allowed to induce will turn out to be merely an implicate of the evidence and that one strongest hypothesis. Indeed the situation for Lehrer's system is even stranger; at a given level we cannot induce all the deductive consequences of the strongest hypothesis we can induce, but only a string of implicates, of which each implies *all* the statements lower in the string.

LEMMA: If h belongs to I_m, then h belongs to I_{m+1}.

PROOF: I_{m+1} is the set of hypotheses h_j such that $D(h_j, C_m \& e)$, i.e., such that for any h, if it is not the case that $C_m \& e \& h_j$ entails h, then $P(h_j, C_m \& e) > P(h, C_m \& e)$. But $C_m \& e$ does entail h, so h is directly inducible from $C_m \& e$.

THEOREM 9: If h and k are directly inducible from e, then either $k \& e$ entails h or $h \& e$ entails k.

PROOF: Suppose that $k \& e$ does not entail h. Then $P(k, e) > P(h, e)$. Suppose that $h \& e$ does not entail k. Then $P(h, e) > P(k, e)$. Therefore either $k \& e$ entails h or $h \& e$ entails k.

Observe that this means that if h and k are any two members of C_n, then either $C_{n-1} \& e \& k \vdash h$ or $C_{n-1} \& e \& h \vdash k$.

(Now, incidentally, we see why the conjunction principle holds. If h and k are directly inducible from e, then so is their conjunction, simply because their conjunction is equivalent, given e, to one of them alone: i.e., either e entails $k \equiv h \& k$, or e entails $h \equiv h \& k$.)

We now come to the main theorem regarding Lehrer's system, which is that there is essentially only *one* statement inducible from e. In order to show this rigorously for infinite languages we must allow ourselves infinite conjunctions; two observations on this move are in order. It seems to be only in pathological cases that the move to infinite conjunctions is required. This move must be allowable in Lehrer's system, since he says to let "C_n be a conjunction of members of I_n that is logically equivalent to I_n", and in the pathological cases no finite conjunction will do. I suspect, however, that worry about the infinite case is academic.

THEOREM 10:

(a) For every n, there is a statement, k_n, which is either inducible from C_{n-1} and e, or which is an infinite conjunction, each member of which is inducible from C_{n-1}, such that every member of I_n is entailed by k_n and e.

(b) There is a statement, k_s, such that every statement inducible from e is entailed by e and k_s, and such that k_s is inducible from e, or is an infinite conjunction, each member of which is inducible from e.

PROOF: For every two statements in I_n, either the first, in conjunction with e and C_{n-1}, entails the second, or the second, in conjunction with e and C_{n-1}, entails the first. This relation gives a partial ordering of the hypotheses inducible from C_{n-1} and e. Either (as would in general be true) there is a hypothesis k such that every element of I_n is derivable from $k \& e \& C_{n-1}$, and then k is the k_n whose existence is asserted in (a), or else there will be an infinite sequence of hypotheses inducible from $C_{n-1} \& e$, $k_1 \ldots, k_m \ldots$, such that $k_i \& C_{n-1} \& e$ entail k_{i-1}, but $k_{i-1} \& C_{n-1} \& e$ do not entail k_i. In this case we let k_n be the infinite conjunction of the hypotheses k_i. (Observe that in this case, C_n will have to be an infinite conjunction, too. We may, of course take k_n as C_n itself in this case.) Therefore $k_n \& e \& C_{n-1}$ does entail h.

In a similar way, either there is an n^* such that for n and m greater than n^*, C_n and e entail C_m and e and conversely, or else, for every n, C_{n+1} and e entail C_n, but C_n and e do not entail C_{n+1}. In the former case k_n is the k_s of (b). In the latter case, there is an infinite conjunction, C^*, such that every member of C^* is inducible from e, and every finite conjunction of members of C^* is inducible from e, which is such that for any statement h whatever, if h is inducible from e, h is deducible from C^*; namely the conjunction of all k_n.

It is perhaps worth observing that the conjunction principle does not hold for infinite conjunctions. Consider a sequence of hypotheses, $k_1, \ldots, k_n$ each of which, with e, entails the preceding hypotheses, but is not entailed by them (with e). The probability of the conjunction $k_1 \& k_2 \& \cdots \& k_n$, given e, is just the probability of k_n given e. The probability of the infinite conjunction $K_1 = k_1 \& k_2 \& \ldots, k_n \ldots$ is thus just the $\lim_{n \to \infty} P(k_n, e)$. Each of the k_i will be inducible provided that for any h not entailed by $k_i \& e$, $P(k_i, e) > P(h, e)$. But $P(h, e)$ could be $\lim_{n \to \infty} P(k_n, e)$ (since

$P(k_i, e) > \lim_{n\to\infty} P(k_n, e)$), and h might not be entailed even by e and the infinite conjunction of the k_i. Thus the infinite conjunction would not be inducible, even though each of its members was. Thus k_n and k_s may not themselves be acceptable.

VI

The system of Hintikka and Hilpinen and the system of Lehrer are unsatisfactory for essentially the same reason. They boil down to the claim that given a language and a body of evidence, there is essentially just one strongest statement that can be accepted. This approach to induction is global with a vengeance. It suggests that as scientists or even as people we do not induce hypothesis by hypothesis, but that induction consists in principle of inducing at each stage of inquiry – i.e., for each body of evidence e – a single monumentally complex conjunctive statement. Observe that we cannot even consider parts of the complex hypothesis in isolation from other parts. Although the evidence may have the form $e_i \& e_j$, and e_i may be utterly irrelevant to h_j, the fact that h_j is inducible from e_j will have no bearing at all on whether h_j is inducible from $e_i \& e_j$. A hypothesis h_i, not entailed by h_j and e_i and e_j may always turn out to be more probable on $e_i \& e_j$ than h_j is. Indeed, one may wonder if the exceedingly high confirmation of the hypothesis that the speed of light is finite will not preclude the acceptance of *any* hypothesis concerning the cause of cancer, the existence of life on Mars, or the amount of inflation to be anticipated in the coming year.

It may be argued that any global system in which the conjunction principle is satisfied will suffer from these shortcomings. In any such system there will be a statement from which every acceptable statement follows, and which is either acceptable itself, or is an infinite conjunction each finite conjunct of which is acceptable.

THEOREM 11: If the conjunction principle is satisfied for an inductive acceptance rule in the language L, then there is a statement C^* in L, or an infinite conjunction C^* every component of which is in L, such that if h is a finite statement, h is inductively acceptable if and only if h is entailed by C^* and the basic evidence e.

PROOF: Entailment given e provides a partial ordering of the statements in L. Suppose first that there are two statements in L, h_1 and h_2, each of which is inductively acceptable, and neither of which is ranked above the other in the partial ordering. By the conjunction principle their conjunction must also be acceptable, and by the partial ordering, their conjunction must be ranked higher than either one alone. Either there is an acceptable hypothesis of maximum rank (from which all the other hypotheses then follow, given e) or else for every hypothesis h, there is a stronger one h' which, conjoined with e, entails h, but which is not entailed by h conjoined with e. If this is the case, an argument like that of a preceding theorem will give a C^* satisfying the conditions of the theorem.

Given the conjunction principle, as I showed earlier, it also follows that the strong and weak consistency conditions are equivalent, and that the strong and weak deduction principles are equivalent. Thus I think it is appropriate to focus on the conjunction principle as a source of the peculiarities to which such systems as those we have looked at give rise.

It is difficult to give an argument against the conjunction principle, partly because it is so obvious to me that it is false, and partly because it is so obvious to certain other people that it is true. The most persuasive arguments perhaps are those which stem from the last theorem presented; it seems preposterous to suppose that all of our inductive knowledge has to be embodiable in a single fat statement. It seems too limiting to say that I have to believe the conjunction of everything I have a right to believe (there cannot be very much, then, that I have a right to believe), and it seems even more unreasonable to claim to have a right to believe the conjunction of everything I have a right to believe. Although I claim to have good reasons for believing every statement I believe, I claim also to have good reasons for believing that some of those statements are false. I think both of those claims are perfectly sound; and if they are, the conjunction principle is false.

VII

The system of rational beliefs I have developed elsewhere accepts the failure of the principle of conjunction. It is an attempt at a global theory, and for a global theory the conjunction principle seems flatly false. Having abandoned the principle of conjunction, it becomes possible to distinguish strong and weak deductive closure, strong, weak, pairwise and n-wise consistency. In what follows I will briefly characterize a simplified version of the original system, freed, of course, from the original inconsistency.

We begin (as always) with a language L; we suppose it to contain a set theory, terms denoting operations, properties, relations, etc., which may be of a theoretical as well as of an observational character. We let B denote a set of statements that are accepted as evidence or background information. In a given context B represents the basic rational corpus, or body of knowledge. We suppose, by hypothesis, that B is pairwise consistent, and satisfies weak deductive closure. B thus contains all the theorems of our language. To be sure, if we pick the wrong axioms for set theory – inconsistent ones – in our language, the set B will be empty. But we must always suppose, so long as we know no better, that the language we speak is consistent. We do not suppose that we believe in any active or behavioristic sense every statement in B; rather we say that the contents of B are what we are *committed* to believing.

AXIOM I: $S \in B \supset \sim nS \in B$ ('nS' denotes the denial, in L, of S).

AXIOM II: Thm S cd $T \& S \in B \supset T \in B$ ('S cd T' denotes the conditional whose antedédent is S, and whose consequent is T; 'Thm S cd T' says that the statement S cd T is a theorem of L).

There is no need here for defining the probability relation; there are certain properties of that relation I shall refer to, which I shall state as axioms. It should be observed that on the basis of a definition of probability like that I have provided elsewhere, these axioms turn out to be theorems. Probability I take to be relative to a body of knowledge or rational corpus B; it is a relation that holds between a statement S, the rational corpus B, and a pair of fractions p and q. We say, relative to B, the probability of S is the pair of fractions (p, q), and we symbolize this assertion: 'Prob$_B$ (S, p, q)'. It should be observed, not as part of the formal development here, but simply as background information, that on the definition I have offered, every probability statement is based on a *known* statement concerning relative frequencies, i.e., that if the relation Prob$_B(S, p, q)$ holds, there is as a member of B some corresponding statistical statement asserting that a certain relative frequency or measure lies between the ratios denoted by p and q. The properties of probability that we shall need are the following:

AXIOM III: Prob$_B(S, `p', `q') \supset$ Prob$_B(nS, `1-q', `1-p')$.

AXIOM IV: [Thm S cd $T \&$ Prob$_B(S, p, q) \&$ Prob$_B(T, p', q')] \supset$ Thm p' gr p ('p' gr p' denotes the statement in abbreviated notation consisting of the fraction p', followed by '>', followed by the fraction p).

AXIOM V: $S \in B \supset$ Prob$_B(S, `1/1', `1/1')$.

Our final axiom requires an auxiliary notion, that of a biconditional chain in B. Since we have not assumed deductive closure in B, it is perfectly possible that $S\,b\,T$ (the biconditional whose antecedent is S and whose consequent is T) is a member of B and that $T\,b\,R$ is a member of B, when $S\,b\,R$ is not a member of B. But we want to say that S and R are related by a biconditional chain in B anyway. In general we shall say that S and R are related by a biconditional chain in B, in symbols, $S\,bc_B\,T$, when $S\,b\,T$ belongs to every set of statements containing $P\,b\,Q$ whenever $P\,b\,Q$ belongs to B, and containing $P\,b\,R$ whenever it contains both $P\,b\,Q$ and $Q\,b\,R$. Formally:

DEFINITION 1: $S\,bc_B T =$ df$(K)((P)\ (Q)\ (R)\ ((P\,b\,Q \in B \supset P\,b\,Q \in K) \& (P\,b\,R \in K \& P\,b\,Q \in\ \supset P\,b\,Q \in K)) \supset S\,b\,T \in K)$.

The final axiom simply says that any two statements related by a biconditional chain have essentially the same probability.

AXIOM VI: $(S\,bc_B\,T \&$ Prob$_B(S, p, q) \&$ Prob$_B(T, p', q')) \supset$ (Thm p id p' $\&$ Thm q id q') ('x id y' denotes the statement [in abbreviated notation] consisting of x followed by '=' followed by y).

We are now in a position to characterize the set of statements B_r which may be induced from B, at the level r. Note that we cannot simply include a statement in B itself on the grounds that its probability is at least r, unless we take r to be 1, for the probability of S, relative to B, can be less than one only if S is not a member of B.

THEOREM 12: (Prob$_B(S, `p', `q') \&$ Thm '$1 > p$'$) \supset \sim S \in B$ (Axioms V and VI).

Let us define B_r to be the set of statements whose probability, relative to B, is at least r, where r denotes a ratio greater than a half.

DEFINITION 2: $B_r =$ df$\{S$: Prob$_B(S, p, q) \&$ Thm p gr $r\}$.

B_r is thus a set of statements accepted on purely probabilistic grounds. We can show that B_r satisfies the weak consistency principle, the weak deduction principle, and the pairwise consistency principle, which, recall, failed for Levi's general system.

THEOREM 13: $(S \in B_r \&$ Thm S cd $T) \supset T \in B_r$ (Axiom IV, D-3).

THEOREM 14: $\sim (\exists S)\ (S \in B_r \& (T)$ (Thm S cd T)).

PROOF: If (T) (Thm S cd T), then Thm nS. Thus $nS \in B$. Prob$_B(nS$, '1', '1'). By axiom III, then, Prob$_B(S$, '0', '0'). By the (hypothetical) consistency of B, we have $\sim$Thm '0' gr r, and thus $\sim S \in B_r$.

THEOREM 15: (Pairwise Consistency and n-wise consistency for n such that $n/(n+1) < r$)

$$(S)\,(S \in B_r \supset \sim nS \in B_r) \text{ (Axiom III).}$$
$$(S_1)\,(S_2) \ldots (S_{n-1})\,[(S_1 \in B_r \,\&\, S_2 \in B_r \,\&\cdots\&\, S_{n-1} \in B_r) \supset \sim n(S_1 \text{ cj } S_2 \ldots \text{ cj } S_{n-1}) \in B_r]$$

(where S cj T represents the conjunction of S and T).

Lehrer makes a point of the consistency of what is induced with the evidence on the basis of which it is induced. This condition is satisfied here for single statements given that B is deductively closed, which may not be an unreasonable supposition for 'observation statements'.

THEOREM 16: $(S)(B \vdash S \supset S \in B) \supset (S)\,(S \in B_r \supset \sim (B, S \vdash \text{contradiction}))$.

PROOF: If $B, S \vdash$ contradiction, then $B \vdash nS$. By the hypothesis of the theorem, then, $nS \in B$, and $\text{Prob}_B(S, \text{'0'}, \text{'0'})$, and $\sim S \in B_r$.

As I pointed out earlier, there are parts of any system like this which are strongly consistent and deductively closed. The relation between statements S and T, Thm S cd T, provides a partial ordering of the elements of B_r. Let us define a strongest acceptable statement, in symbols STR, to be a statement such that no other statement in B_r bears the relation in question to it, unless it also bears that relation to the statement. Thus:

DEFINITION 3: $STR_{B_r}\, S =df\, S \in B_r \,\&\, (T)\,((T \in B_r \,\&\, \text{Thm } T \text{ cd } S) \supset \text{Thm } S \text{ cd } T)$.

The set of consequences of any strongest statement in B_r satisfies the strong deduction principle, the strong consistency principle, and (!) even the conjunction principle.

THEOREM 17: $STR_{B_r}\, S \supset [(T)\,(\text{Thm } S \text{ cd } T \equiv T \in A) \supset (\sim A \vdash \text{contradiction} \,\&\, (R)\,(A \vdash R \supset R \in A) \,\&\, (R)\,(T)\,((R \in A \,\&\, T \in A) \supset R \text{ cj } T \in A))]$.

Perhaps, in these terms, it is the fact that most of our beliefs are not strongest beliefs that has led people to feel that our beliefs belong to systems of beliefs which satisfy the conjunction principle.

There are a number of other theorems we can prove. For example, we can prove that if a complex statement S is like a complex statement T, except for containing occurrences of the statement P where T contains occurrences of the statement Q, and if P and Q are connected by a biconditional chain in B, then S will be a member of B_r if and only if T is a member of B_r. In a similar fashion it is possible to define the concept of an identity chain in B. Then it is possible to prove that if S is like T except for containing occurrences of the term p where T contains occurrences of the term q, and p and q are connected by an identity chain in B, then S will be a member of B_r if and only if T is a member of B_r. Even without complete deductive closure, there is a lot that can be shown to hold in B_r.

The issue is only whether or not there is a single strongest statement in B_r – i.e., whether there is a statement S^* such that $(T)\,(T \in B_r \supset \text{Thm } S^* \text{ cd } T)$. When stated thus baldly, the answer is obvious; it is gratuitous to suppose that there is any such statement. Indeed the supposition that there is is one of the secondary symptoms of the disease I have called conjunctivitis.

University of Rochester

BIBLIOGRAPHY

[1] R. B. Braithwaite, *Scientific Explanation*, Cambridge 1953.
[2] Gilbert Harman, 'Detachment, Probability and Maximum Likelihood', *Nous* **1** (1967) 401–11.
[3] Carl Hempel, 'Deductive-Nomological vs. Statistical Explanation', in *Minnesota Studies in the Philosophy of Science*, III (ed. by H. Feigl and G. Maxwell), Minneapolis 1962.
[4] Jaakko Hintikka, 'Towards a Theory of Inductive Generalization', in *Proc. 1964 International Congress for Logic, Methodology, and Philosophy of Science* (ed. by J. Bar-Hillel), Amsterdam 1965.
[5] Jaakko Hintikka, 'Distributive Normal Forms in First-Order Logic', in *Formal Systems and Recursive Functions* (ed. by Crossley and Dummett), Amsterdam 1965.
[6] Jaakko Hintikka, 'A Two-Dimensional Continuum of Inductive Methods', in *Aspects of Inductive Logic* (ed. by J. Hintikka and P. Suppes), Amsterdam 1966.
[7] Jaakko Hintikka and Risto Hilpinen, 'Knowledge, Acceptance, and Inductive Logic', in *Aspects of Inductive Logic* (ed. by J. Hintikka and P. Suppes), Amsterdam 1966.
[8] Richard Jeffrey, *The Logic of Decision*, New York 1965.
[9] Henry E. Kyburg, Jr., *Probability and the Logic of Rational Belief*, Middletown 1961.
[10] Henry E. Kyburg Jr., 'Probability and Randomness', *Theoria*, **29** (1963).
[11] Henry E. Kyburg Jr., 'Probability, Rationality, and Rule of Detachment', in *Proc. 1964 Congress for Logic, Methodology and Philosophy of Science* (ed. by J. Bar-Hillel), Amsterdam 1965.
[12] Keith Lerher, 'Induction, Reason, and Consistency', forthcoming in *Minnesota Studies in the Philosophy of Science*.
[13] Isaac Levi, *Gambling With Truth*, New York 1967.
[14] Wesley Salmon, private correspondence.
[15] Raimo Tuomela, 'Inductive Generalization in an Ordered Universe', in *Aspects of Inductive Logic* (ed. by J. Hintikka and P. Suppes), Amsterdam 1966.

REFERENCES

* Much of the research on which this paper is based has been supported by the National Science Foundation, through grants 708, 1179, and 1962.

[1] Except in one of Keith Lehrer's systems, described in this volume.

[2] A muddier version of this argument was presented in [11]; a cleaned-up version is mentioned by Harman in [2].

[3] These principles are not essential parts of the Bayesian or Classical statistical theory. One can develop the theory of statistical inference without considering the question of acceptance one way or the other. The classical theory requires us to *reject* certain hypotheses, but it is hardly necessary to point out (as statisticians of this persuasion inevitably do) that to reject a statement is not (necessarily) to accept it. Bayesian theory is sometimes coupled with a philosophy according to which one *never* accepts any *hypothesis*.

[4] The test interval for '$p \in (a, f)$' will include the closed interval $[a, f]$ and the test interval for '$p \in (f, b)$' will include the closed interval $[f, b]$.

GILBERT H. HARMAN

INDUCTION*

A Discussion of the Relevance of the Theory of Knowledge to the Theory of Induction (with a Digression to the Effect that neither Deductive Logic nor the Probability Calculus has Anything to Do with Inference)

I

In 1963 Edmund Gettier demonstrated that knowledge is not simply justified true belief[1]; and what has been learned in the resulting discussion[2] has important implications for a theory of reasoning. This paper describes some of those implications; more generally, it attempts to show how theories of knowledge and reasoning must be adapted to each other if one is to achieve a unified theory of both.

An obvious connection between one's theory of knowledge and one's theory of reasoning is that one can take reasoning to be warranted or valid if it could give a person knowledge. For example, a detective comes to know who the murderer is by reasoning from circumstantial evidence: in such a case his reasoning can be said to be valid or warranted. Whether a person knows something by reasoning depends (in part) on whether his reasoning justifies his belief. In the language of inductive logic, knowledge depends on whether reasoning justifies acceptance of one's conclusion. Epistemologists refer to principles that warrant belief where logicians refer to rules of acceptance.

If we thus approach inductive logic from the theory of knowledge, we will want it to provide a *strong rule of acceptance.* Roughly speaking, such a rule tells one that, given certain evidence, one may infer, accept, or believe nonprobabilistic conclusions which may be used as part of the evidence in further reasoning or inference. That is oversimplified, since sometimes one should *reject* something previously accepted. More precisely, prior to inference one accepts a set of propositions which serve as premises or evidence; inference leads to a modification of the set either by the acceptance of further propositions or by the rejection of propositions previously accepted; and one can then use the revised set as a basis for future inference.

A person can come to know something by inference only if rules of acceptance authorize him to accept it. The following test of proposed