DAVID MILLER

The Measure of All Things

In this paper[1] I present no results, only questions; and, I hope, difficulties.

1. Central to all theories of induction is an insistence on instances. Instances of a universal theory are needed to confirm it. The problem is: how many?

Hume of course showed that no quantity of instances could verify a universal generalization. In his *Abstract* (1740, p. 15) he argued further that no number of instances could make a universal generalization even probable. The former result is not seriously contested these days; if we can judge by the continued interest in probabilistic inductive logic, the latter is. Since the argument is as good as the same in each case, this is somewhat puzzling.

Perhaps one difference is that it actually seems to be possible, for very simple languages, to give probability measures according to which universal generalizations receive high probabilities (or 'degrees of confirmation') relative to finite evidence. Certainty, it appears, cannot be so achieved.

This is really rather a rough way of putting the matter; shortly I shall smooth it out a little. But for the moment I just want to say that I think the difference is no difference. Just as there is not a certain number of instances that will make a hypothesis certain, so there is not a reasonable number that will make it reasonable.

[1] The present paper has little in common — beyond sameness of spirit — with the paper read at Minnesota in 1968. Most of what I said there either was unoriginal or is no longer my opinion. In the latter category I include all the doubts I expressed on whether Popper's noninductive methodology could handle problems of pragmatic or practical preference. I should now like unreservedly to withdraw all such doubts; the problems seem to me now susceptible of an elegant and simple solution. See Popper, ch. 1 (1972) and especially sec. 14 (1974).

2. It has undoubtedly been acknowledged that, so abundant is the supply of probability measures, there is no real difficulty in locating one which makes any hypothesis, unrefuted by particular evidence, as probable as one likes thereon. The trouble comes in making the measure even tolerably plausible. For it has been appreciated also that the choice of one probability measure over all others is hardly to be justified by pure logic alone. (This is obvious in the case of subjectivists and has been made explicit by such writers as Carnap (1952, sec. 18) and Hintikka (1970, pp. 23–24).) To this extent, anyway, Hume's extension of Hume's argument to the probabilistic case has not been utterly disregarded. Levi sums up a fairly popular position well when he says, "The skeptical despair that threatens quests for global justification need not infect efforts at local justification" (1967, p. 6). Likewise, Hacking remarks, "Objective *local* inductive logic is already a reality. It is practiced by statisticians" (1971, p. 610).

Such admissions are surely sufficient to refute inductive logic's claim (if it makes the claim) to be a viable theory of knowledge. But never mind. Let us grant local justification, or local inductive logic, as a possible activity (without identifying it with statistics); grant, that is, that we often make a (nonlogical) choice of 'logical' probability measure and in terms of it perform inductive inferences. Even so, I maintain, inductive methods are next to useless. Hume's arguments cannot be so readily dismissed.

In most cases to date high probabilities for universal theories have been taken as a goal of inductive methods. (Kyburg, 1961, Hintikka and Hilpinen, 1966, and Hilpinen, 1968, are examples.) Sometimes, however, acceptance according to some rule or other is the primary goal, and high probabilities, though valued, enter only indirectly (Levi, 1967). In apparent opposition to both groups has been Carnap, who has on occasion been less than a champion either of universal laws or of acceptance rules. But it is clear that he attaches some significance to both, especially if purely theoretical issues alone are at stake (Carnap, 1966, pp. 258–60). If high probability means nothing for a theory, it is indeed somewhat hard to see the point of probabilistic inductive logic. And this holds despite the fact that very often high probability is hardly deemed on its own a satisfactory understudy for certainty. After all, to the extent that the measure involved is not 'logical', to that extent what Carnap calls the 'logical factor' may need overruling by the 'empirical factor'. Repetition or enumeration of instances is supposed to be as valuable in strengthening weight of evidence, in making probability judgments reliable, as in boosting probability; though admittedly in some cases the two jobs seem to merge into one (Levi, 1967, ch. 9; Hilpinen, 1968). In accordance with all this I shall hereafter suppose that high probabilities are esteemed, without inquiring into the reason for such treatment.

3. The case against inductive inferences which aim at certainty can be split in two. (1) To render certain a universal statement we would have to examine every object in the universe. (2) Even if we did this, it would not be enough; we would still have to ensure that we *had* examined every object — and this involves another universal statement.

To put it round the other way: even were we to know exactly what the universe consisted of, we would still have to investigate every object in it in order to confer certainty on generalizations like "All *A*'s are *B*'s." It is not surprising, therefore, that if we permit ourselves the luxury of concluding "All *A*'s are *B*'s" after observing just $a_0, \ldots, a_{n-1}$ as 'positive instances', then we can conclude also that $a_0, \ldots, a_{n-1}$ exhaust the universe of discourse. More precisely, if the rule "x_0 is *A* and *B*; x_1 is *A* and *B*; . . . ; x_{n-1} is *A* and *B*; *therefore* all *A*'s are *B*'s" is allowed to be inductively valid for sufficiently large *n*, then, by substituting a tautological property for *A* and "observed" for *B*, we can conclude that beyond the observed objects $a_0, \ldots, a_{n-1}$ the universe is quite empty. This argument is basically due to Goodman.[2]

If our conclusions are supposed to be certain, the point of the argument is indeed obvious. For if we cannot be certain that all the universe has been inspected, neither can we be certain that there does not exist an *A* which is not a *B*. Thus it is just not true that deduction is the special case of induction that obtains when all objects in the universe have been observed. I would contend that there is, likewise, no process of 'near-deduction' that obtains when nearly all objects have

[2] It is usually presented as a criticism of the rule "All observed *A*'s are *B*'s; *therefore* all *A*'s are *B*'s." But this rule — having, as it has, a universal premise — can only be applied after application of a rule like that given in the text. It is therefore less fundamental. The unconvincing form of the rule in the text — symmetric in A and B in the premises and asymmetric in the conclusion — is a symptom of the nonexistence of such things as positive instances. This objection is perhaps too well known to need repetition; see, however, Popper's reply to Quine (Popper, 1974).

been observed — we need to know the size of the universe in this case too. On the other hand, suppose we permit ourselves the near-luxury of concluding with near-certainty "All *A*'s are *B*'s" after observing just $a_0, \ldots, a_{n-1}$ as 'positive instances'. Can we conclude also (with near-certainty) that $a_0, \ldots, a_{n-1}$ exhaust the universe of discourse?

Probabilistic replicas of the argument I have ascribed to Goodman will be the main concern of this paper. In this section we deal with some informal replicas; in later sections with formal ones.

No assumption is made about the size of the universe except that it is no smaller than the number n of individuals $a_0, \ldots, a_{n-1}$ we have observed. All are 'positive instances'. We conclude, "Probably all *A*'s are *B*'s"; or " 'All *A*'s are *B*'s' is probably true"; or "The probability of 'All *A*'s are *B*'s' relative to the evidence $=$ r." By the same token, may we not conclude also that probably everything has been observed; or that "All things have been observed" is probably true; or that the probability of "All things have been observed," relative to the evidence, is equal to r?

In the first two cases it is hard to know how to reply. "Probably all *A*'s are *B*'s" and its metalinguistic analogue " 'All *A*'s are *B*'s' is probably true" are often cited, in a typically vague manner, as typical inductive conclusions. There is no obvious reason why the conclusion that everything has been observed should be inadmissible, except that it is unexpected. Indeed, it is disastrous; for as soon as we can conclude that probably all *A*'s are *B*'s, after observing some, we are entitled to conclude that probably our conclusion is vacuous anyway — that is, it probably tells us no more than we know already. But I shall not say more here about these informal inductions; there is indeed very little more to say.

It is with examples such as the metric example that much current work in induction is concerned. Here, at least, something can be said. Remember, we have a conclusion, "The probability of 'All *A*'s are *B*'s' relative to the evidence $=$ r." Can we hope that the probability of "All things have been observed" is likewise r, or near r?

In most systems on the market it would appear that we cannot, for reasons that I shall sketch in the next section. But the question still remains: If high probabilities can be obtained reliably only after a large number of instances have been observed, are we not in danger of making it also highly probable that the universe has been exhausted? (In this case the induction is hardly worth making.) Or, more generally: How many instances are needed before we can make reliable inferences? Too many?

4. Henceforth I shall restrict myself to the systems of Carnap and Hintikka, and to related systems (such as Levi's treatment of universal generalization, 1967, ch. 13). My detailing will be extremely sketchy.

We are concerned with elementary languages containing the usual logical trappings, identity, individual constants, and monadic predicate symbols. (The systems discussed in any detail by Hintikka do not include the sign of identity; but there is a generalization developed by Hilpinen (1966) which does.) Identity does not in fact have much to do in most of the systems under consideration, since no two individual constants (names) are allowed to name the same object. Indeed, the only permissible interpretation for a language with N names is a domain of N individuals, each named by exactly one name. (The case of an infinite universe is handled by considering what happens when N grows without limit.) Consequently universal statements are, within the language, deemed logically equivalent to N-fold conjunctions of instantial statements.

Evidence statements are intended to be singular statements, containing predicate symbols and individual constants. Though many different ways have been invented of supplying an initial probability distribution, all of them lead to some relative probabilities' depending on the number of names in the language. Certainly that number is fixed in advance; so the size of the universe is known — assumed, that is to say — in advance.

In formal systems like these such an assumption is hardly enough, as it was in our more informal treatment above, to make the sentence "All things have been observed" in any way acceptable. For normally there is no predicate "observed." And if it does exist as primitive it can expect no special (formal) relationship with evidential statements. "All things have been observed" can just as well attain, or fail to attain, a high probability as can "All things are blue," given the right sort of evidence. On the other hand, if we impose the (methodological) requirement of total evidence any actual evidence will be weighed down with positive occurrences of "observed," so that "All things have been observed" will be as well confirmed as any universal sentence. This, I suppose, is why workers in inductive logic have chosen not to feature "observed" as a primitive predicate (see also p. 358 below).

Given the sign of identity we can of course get round this lack of expressibility, more or less. If $a_0, \ldots, a_{n-1}$ are the objects that have been observed, the sentence $(x)(x = a_0 \vee \ldots \vee x = a_{n-1})$ can be used to say that everything has been observed (though it is not equivalent to it). But this statement will stick at probability zero so long as $n < N$, the preordained cardinality of the universe, whatever the evidence. For given the interpretation of universal statements it states a contradiction.

It would look, then, as if the simple version of Goodman's paradox outlined above cannot be revived in most influential systems of inductive logic. This is hardly because the systems are too cleverly designed; they can scarcely be said to *surmount* the criticism, only to *avoid* it. Is it possible to make these systems more susceptible to this line of attack, whether or not the attack succeeds?

Crucial, of course, is the assumption that the size of the universe, and even its composition, can be known in advance. It must be tempting for an inductivist to pigeonhole this assumption among the 'global' ones that are agreed to be necessary and to treat inquiries as 'local' only when such an assumption has already been made. But this, I suspect, is asking a bit much. To be sure, the universe whose elements we are supposed to know about in advance is not in general the universe. It may be an urnful of balls; it may be a pack of cards. (In this case we shall surely be prepared to acknowledge a previous acquaintance.) But the universe is quite often something like the set of future tosses with a coin, a set whose cardinality seems far from fixed.[3] It does seem that an inductive logic that cannot cope with such universes may be an inductive logic without a future.

5. I have yet to offer any details of the way in which probability is distributed in the systems of Carnap and Hintikka. Carnap calls a *state-description* a sentence which is consistent and otherwise maximally strong; it tells of every object exactly what properties it has. (Quantifiers can be ignored, since all quantified sentences are equivalent to unquantified ones.) We also ignore, of course, differences in order of conjuncts in state-descriptions. Call two state-descriptions isomorphic

[3] This oversimplifies. See my "What's in a Numeral?" (unpublished).

if each can be obtained from the other by a permutation of individual constants. Given a state-description, we call the *structure-description* to which it belongs the disjunction of all state-descriptions isomorphic to the given state-description. Structure-descriptions of course also contain individual constants, and Carnap initially distributes his probability measure over structure-descriptions (c^* does it equally); it is then subdivided, within each structure-description, among state-descriptions. Right from the start, then, many probabilities depend considerably on the size of the universe. Carnap succeeds in softening somewhat this dependence by his requirement C6 (1952, p. 13) that probabilities of singular sentences on singular evidence be the same, whatever the size of the universe. But as Hintikka notes, "The main drawback of . . . Carnap's procedure seems . . . to be its dependence on the domain of individuals which the language in question presupposes. . . . In most applications, however, this domain is largely unknown. In general it seems . . . perverse to start one's inductive logic from the assumption that one is in some important respect already familiar with the whole of one's world . . ." (1965, p. 279). Hintikka's systems avoid this dependence in the early stages, but not for long. For his initial handout of probability is over what he calls *constituents*, which are the strongest consistent sentences of the language which contain no individual constants. Each structure-description can be treated as a disjunct in just one constituent, so structure-descriptions (and thereafter state-descriptions) receive their absolute probabilities according to some subdistribution (usually a uniform one in each case). Thus it is far from true that the absolute probabilities are in general independent of *N*. What is true is that the 'possible states of the world' over which we initially distribute probabilities are independent of N; this Hintikka acknowledges (see 1965, p. 282, n. 19).

Like Hintikka I feel that no dependence of this sort is particularly desirable. How, however, can it be avoided in systems akin to Carnap's and Hintikka's? For, first of all, if we don't agree on some set of individual names then we have no determinate set of state- or structure-descriptions. In Carnap's system this means that no probabilities can be assigned at all; in Hintikka's no probabilities on singular evidence. (I exclude, of course, the trivial cases.) Or, rather, probabilities can be assigned (using "*N*" as a variable for the power of the universe) but there is no way of comparing them with one another. What is more, if our set of names is not finite, state- and structure-descriptions cease to be well formed — because infinitely long. And secondly, if the names are in any sense 'proper names' we cannot suppose that we use them without being familiar with (or perhaps 'acquainted with') their nominata. Thus the very use of names seems to presuppose familiarity with members of the universe; so is it worth trying to fight against it?[4]

Both these points can, in my opinion, be answered. It is possible to develop systems, similar to those already discussed, in which the size of the universe is in no sense 'fixed' in advance. It is possible, therefore, to produce systems in which the query "What is the probability that the universe has been exhausted, given the evidence?" does not always, when it hasn't been exhausted, give the answer zero.

To start with suppose that the universe is finite, though its cardinality is otherwise secret. A way around the first point is then almost obvious. In both Carnap's and Hintikka's approaches the crucial handout of probability is among structure-descriptions. So suppose that we have examined two objects. We consider to start with the set S_2 of all structure-descriptions in a universe with two names, a_0, a_1; call p_2 the measure we choose to put on S_2. Now add to S_2 all structure-descriptions involving those names and one other, a_2. This gives us a larger set of structure-descriptions, S_3, but one over which we can distribute probabilities as before. In terms of such a distribution, p_3, it makes good sense to ask for the probability, given the evidence, of the sentence $(x)(x = a_0 \vee x = a_1)$. For this is just asking for the measure, according to p_3, of the state-descriptions, consistent with the evidence, which contain only the names a_0 and a_1. It seems not unreasonable to demand that p_3, the distribution over S_3, be a genuine extension of p_2, so that the latter probabilities are obtainable from the former by relativization with respect to the sentence $(x)(x = a_0 \vee x = a_1)$; for any h not containing the name a_2 we have $p_2(h) = p_3(h, (x)(x = a_0 \vee x = a_1))$. This means that when, more generally, we add further names a_3, a_4, . . . , the probability of $(x)(x = a_0 \vee x = a_1)$ with respect to the successive measures p_3, p_4, . . . , will, if anything, decrease. As the number of individual names increases indefinitely, the probability of $(x)(x =$

[4] This paragraph is a trifle naive. Hintikka's later system (1966) enables probabilities on singular evidence to be assigned in an infinite universe. There is, naturally, no need to mention the universe's cardinality. But what the state- and structure-descriptions are is left a little vague. See the discussion by Hilpinen (1968, pp. 53–58).

$a_0 \vee x = a_1$) will therefore tend to a limit from above. Further conditions may be required on the sequence of measures to ensure that other probabilities have limits as the number of names increases. Then we can take the probability of a sentence as the limit of the probabilities it receives under the sequence of measures $p_2, \ldots$ outlined above.

What this amounts to is a distribution over all possible finite cardinalities for the universe. That is, we assign to each sentence of the form $(x)(x = a_0 \vee \ldots \vee x = a_{n-1})$ a positive probability. Should we wish to allow the possibility of an infinite universe, we need only take care that the limit of the probabilities above is not 1. Of course, the infinitude of the universe cannot be stated in the language, so this gambit will not assist us in calculating probabilities on the assumption that the universe is infinite. But this is a separate problem.

A method such as this for avoiding any decision on the size N of the universe is intuitively worthless as long as the names $a_0, \ldots$ are imagined to be proper names of the usual sort. For it is quite obvious that no account has been taken above of any structure-descriptions that contain, for example, just the names a_0 and a_2. We have required that the names featured in any structure-description form *an initial segment of the list of all names*. An argument due to Popper, however, has indicated that most systems of inductive logic are anyway unsatisfactory if they employ proper names like "John" and "Mary" (Popper, 1967). And a reply to Popper by Jeffrey suggested what can be done about it. Popper's point can be put as follows. Suppose that there are two buttons in a box, Udolpho and Warren. If we observe that one is amber, without 'observing' its name, we have to formulate our evidence as "Udolpho is amber or Warren is amber." Probabilities calculated relative to evidence so formulated will differ in general from probabilities calculated relative to "Udolpho is amber" or to "Warren is amber." Yet the information that the selected button is in fact Udolpho (or that it is in fact Warren) should hardly make any difference to the probabilities (for example, to the probability that both buttons are amber). Jeffrey responds (1968) by remarking that in such circumstances to choose a name like "Udolpho" is to ask for trouble; instead we should make use of names like "the observed button," "the other button." Our evidence can then read "The observed button is amber"; there is no need to consider the weaker sentence "Either the observed button is amber or the other button is amber." Thus no problem arises.

In our more general case we must therefore take a_0, . . . not as orthodox proper names, but as some such names as "the first object observed," "the second object observed," and so on. The supposition that objects are observed one at a time is perhaps slightly fanciful, but hardly outrageous compared with the other idealizations that inductive logic asks us to concede. But once we allow it, most of the problems caused above by proper names are at an end.

For example, there seems no longer much to be said for the assumption that every object is endowed with a name. Indeed, since the universe is not known, neither are the objects in it. Names become attached to particular objects ('fixed' or 'known' objects) only when these objects swing into sight and become observed (but see fn. 7 below) — or anyway when they become identified; up to then, the most that we can say of, for example, the name "the tenth observed object" is that it is the name of the tenth observed object (if there is or will be one). And since we are hardly going to be persuaded a priori that all objects have been or will be observed, we must concede that some of them may never succeed in getting named in this way. (Of course, persuasion a posteriori that all objects have been observed is just what I want to argue that inductive logic ought to lead to.) Note, however, that such names as "the first unobserved object" are also possible and indeed are needed for singular prediction. But not all things can be named even in this manner, unless *the unobserved part of the universe can be well ordered* — a postulate that I would not particularly want to make. (See my "What's in a Numeral?") It must be admitted, however, that structure-descriptions and state-descriptions lose much of their importance in a scheme like this. For if some objects are anonymous there is not much of a sense in which a structure-description expresses a possible world or type of world. The same is of course not true of Hintikka's constituents. One way to accommodate this deficiency is to distribute whatever probability is allotted to a constituent *not* over the structure-descriptions it involves (it doesn't involve structure-descriptions), but over 'truncated structure-descriptions'. By a truncated structure-description, I hasten to explain, I mean a sentence lying half way between a constituent and a structure-description: a sentence which attributes a suitable galaxy of properties to the named objects, but for the rest of the universe merely says which properties are instantiated (*if any*). Since all observed objects have names, relative probabilities on evidence are hereby determined. (Note that there is no obvious way of solving this problem in a system like Carnap's which probes no deeper than structure-descriptions.)

Let us take a very simple example. We have three individual names, a_0, a_1, a_2. Two monadic predicates P_0, P_1 produce four Q-predicates Q_0, Q_1, Q_2, Q_3. If the sign of identity is absent from the language then the following is a fairly typical constituent:

$$(\exists x)Q_1x \,\&\, (\exists x)Q_2x \,\&\, (x)(Q_1x \vee Q_2x).$$

It asserts that Q_0 and Q_3 are the Q-predicates that the universe omits. Now consider the sentence

$$Q_1a_0 \,\&\, Q_2a_1 \,\&\, Q_1a_2 \,\&\, (\exists x)Q_3x \,\&\, (x)(Q_1x \vee Q_2x \vee Q_3x);$$

this tells us which Q-predicates the named objects a_0, a_1, a_2 subscribe to, and of the rest of the universe is no more particular than is a constituent. And in fact it is actually less particular, for it does not tell us whether or not there is any object besides a_1 with the property Q_2. To get what we want, therefore, we shall have to introduce the sign of identity. (Of course, we need it for other reasons too; and I excluded it above only for simplicity's sake.) What would be a fairly typical constituent when we enrich this miniature language with the identity sign?

Hilpinen (1966, p. 134) has pointed out that we need to limit the maximum number of layers of quantifiers in a sentence (its *depth*) if we are to solve this problem. Suppose that it is limited to 2. Then a constituent might be something like

$$(\exists x)(Q_1x \,\&\, (\exists y)(y \neq x \,\&\, Q_1y)) \,\&\, (\exists x)(Q_2x \,\&\, (y)(Q_2y \rightarrow y = x)) \,\&\, (x)(Q_1x \vee Q_2x).$$

This tells us that Q_1 and Q_2 are still the only Q-predicates admitted by the universe, that at least two things satisfy Q_1, and that one thing only satisfies Q_2. Now a truncated state-description would be some sentence like the following:

$$\begin{aligned} Q_1a_0 \,\&\, Q_2a_1 \,\&\, Q_1a_2 \,\&\, &(\exists x)(x \neq a_0 \,\&\, x \neq a_2 \,\&\, Q_1x \,\&\, (\exists y)(y \neq x \,\&\, Q_1y)) \\ &\&\, (\exists x)(x \neq a_1 \,\&\, Q_2x \,\&\, (y)(Q_2y \rightarrow y = x)) \\ &\&\, (x)(Q_1x \vee Q_2x). \end{aligned}$$

It tells us the Q-predicates of the observed objects a_0, a_1, a_2, and of the rest of the universe it tells us how the Q-predicates are distributed, as well as is possible with two levels of quantifiers only.

This is a truncated state-description. A truncated structure-description is the disjunction of all truncated state-descriptions that are identical up to the permutation of individual names. It is over these sentences that I would propose distributing the measure already accorded (by some procedure akin to Hilpinen's) to the constituent cited just above. We can then make a further, presumably uniform, subdistribution among truncated state-descriptions. In this way we will generate a sequence of measures p_0, . . . Since all observed objects have names, this is enough to determine all the relative probabilities that we could ever need. Of course there are problems; for example, what depth can we allow our constituents to achieve? It would be reasonable, I think, to have this increase with the number of names in the language. But then what further restrictions would be needed to ensure that our sequence of measures has the required limiting properties? And so on.

I shall not discuss these problems, since we have already taken ourselves rather far from conventional systems of inductive logic; and it is not my task, even less my desire, to develop a new system. What I want to return to is the original problem of this paper, the probabilistic replication of Goodman's argument stated on p. 352 above.

6. We have, therefore, a sequence of measures defined over an ever expanding set of structure-descriptions — those involving just a_0 and a_1, those involving just these two or alternatively a_0, a_1, and a_2, and so on. So suppose that we have observed n objects. Whatever objects they are, their names will be $a_0, \ldots, a_{n-1}$. What now is the probability, given this evidence, of the sentence $(x)(x = a_0 \vee \ldots \vee x = a_{n-1})$; the probability, that is, that every object is among those observed? It is clear that this will depend a good deal on the sequence of measures p_0, . . . employed; and it will be small, if not zero, unless the structure-descriptions with few individual names (at least n, of course, for the evidence demands that) are very heavily weighted in comparison with the others. The critical point is, To what extent will universal statements suffer in the same way? Can we ensure that most of them will not have negli-

gible probabilities too? To be sure, $(x)(x = a_0 \vee \ldots \vee x = a_{n-1})$ will be stronger, so not more probable, than any other universal statement; but it may be that it will be only marginally less probable.

This is all too vague to amount to anything like a direct criticism. Yet what it seems must be true is this. Suppose that we attach no credence at all to the view that the universe is infinite. Then if $a_0, \ldots, a_{n-1}$ have been observed, for any $\epsilon > 0$ there is some finite number k such that, on the evidence, the sentence "There are no more than k objects other than $a_0, \ldots, a_{n-1}$" has probability in excess of $1 - \epsilon$. If, therefore, high probability means anything, it must mean something for this hypothesis. If, for example, we have an acceptance rule under which sufficiently probable hypotheses are acceptable (subject perhaps to other conditions), then this hypothesis is acceptable. If the universe is assumed to be finite, it turns out that we can put a bound on it. On the other hand, if we consider the possibility of an infinite universe — as we surely should — two cases can arise, depending on how seriously we take the possibility. If it is an outside one, so that the limit as m grows of the probabilities of the sentences $(x)(x = a_0 \vee \ldots \vee x = a_{m-1})$ is greater than $1 - \epsilon$, then what has just been said goes through in just the same way; we can put a finite bound on the size of the universe. If the opposite obtains, this cannot be done. But then, can we be sure that any universal statement can reach a probability as great as $1 - \epsilon$? It seems to me that we cannot. For since the infinitude of the universe is not statable in the language, how are we to calculate the probability of any universal statement relative to the assumption that the universe is of infinite power? And without this contribution we are lost. (Note, incidentally, that the talk of probability in this paragraph can be interpreted as either absolute or relative probability; and that when there is clearly evidence available I take it for granted that the probability is relative. It is easily seen that it hardly matters at all what the evidence is.)

Thus if an inductive logic of this sort is to be of any use, it must allow us to put a finite bound on the universe. I don't see inductivists relishing this conclusion. But they will of course point out that there is no inductive logic of this sort in existence and perhaps that I have made plain the unattractiveness of such a system. But if we bar from the object language statements about the size of the universe, how are we to judge whether we might not have observed too many instances? It is to this question that we finally turn.[5]

7. Against the view that inductive methods are next to useless enthusiasts may cite the following partial successes. Suppose that we assume in advance (a global assumption) that the universe is of size N. Then it does seem possible for universal hypotheses to receive gratifyingly high probabilities when the number of observed objects n is a small fraction of N.[6] Thus informative conclusions can be reached despite the fact that only a small part of the universe has been inspected. There is no reason to suppose that this is not true when the size of the universe is known to lie between certain limits, though not known exactly.

My criticism here — it is really the same criticism over again — is that if the conditions envisaged really obtain, then the n objects observed are not a 'small part' of the universe. One hundred buttons need not be a 'small sample' from a box of ten million buttons; they may amount to almost all of them. Likewise, one hundred buttons may make up almost all of even a countable infinity of buttons. *It all depends on how we measure the universe.* To be sure, if the universe is finite, size N say, the most natural way of measuring a subset of size n is to allot it n/N. This uniform distribution is the only one under which every subset of any given cardinality is accorded the same measure; in particular, the only one in which all individuals in the universe are treated alike. *In an infinite universe no measure satisfies these conditions.* But this is not the only way of measuring the universe; and though it seems to me as good a way as any other, I don't think that an inductivist ought to think that.

For *it follows from the inductivist point of view that the individuals in the universe should not be treated alike.* In particular, induction treats an *observed* object as in some way of deeper significance than an *unobserved* one. If only one object has been observed the believer in induction thinks that he has already learned something about the rest of the universe; consequently, the first observed object ought to

[5] The lack of precise — or anyway precisely stated — results in this section is no doubt deplorable. But I hope that I have given sufficient indication of the hazards any system developed along the lines suggested would encounter.

[6] For doubts, and discussions, see my "The Acceptance World."

attract greater measure than any other does (though not necessarily more than do all the others put together). The second object observed deserves for the same reason to be the next most heavily weighted, and so on down the line.[7] It seems reasonable that the rate of decrease should satisfy the following: when we reach a number n of observations after which we feel entitled to ignore the rest of the universe, and come to a conclusion, then the sum of the measures of the first n objects should be substantial. (This can only be rough, since n is likely to vary for hypotheses of different strength; this happens in Hilpinen (1968), for example. But it may be possible to make a refinement by taking into account utility considerations.)

In other words it seems to me reasonable that if a theory manages to attain so high a probability as to be acceptable, then a very considerable part of the universe should have been examined. Naturally, I do not expect an inductivist to feel the same. But what measure for the universe *will* he use? Some independent argument for using a uniform measure has to be provided. Certainly it should not be uncritically assumed that the uniform measure is the 'natural' way of measuring the universe. There is no more a 'natural' (or 'correct') way of measuring the universe than there is a 'natural' or 'correct' way of measuring probability. Indeed, if the universe is assumed to be infinite, then the uniform distribution — anyway not countably additive — would assign zero measure to each element. Thus in this case *every finite sample would constitute a negligible part of the universe.* In such a situation it is difficult to see why we should take its message so seriously.

If anything, I should have thought, our probability measure and our measure of the universe should harmonize in the way suggested. Then

[7] It may be said that this is no proper measure, since we do not know which object the first one will be. So which object is it that gets the largest measure? The answer is: the first one observed. We have to get used to the idea that any other names owned by the individuals in the universe are incidental and that the objects are identifiable by the names suggested on p. 358. (For a fuller defense see my "What's in a Numeral?") But this is not especially bizarre. Think of "the first card drawn," "the first throw of the coin," and so on.

That not all individuals are treated alike in inductive logic may seem to be contradicted by Carnap's condition C7 (1952, p. 14). But if we use the sort of names suggested here, perhaps we will be more inclined to give up this condition of Carnap's. I cannot discuss here the complications that this would lead to. (It is interesting that in the Kolmogorov/Martin-Löf theory of randomness an attribute with a very low frequency or probability in a Bernoulli sequence *cannot* occur at the first trial, though it can of course occur at others. See Martin-Löf, 1966, p. 616.)

inductive methods would be more or less fruitless. But if there is no such harmony, there is certainly a need for an argument why there is not. Yet even the recognition that such an argument is lacking is sadly lacking itself. This is the current state of probabilistic inductive logic.

In conclusion I say just this. I have attempted to spell out the implications of the view, acknowledged but not apparently taken sufficiently seriously by Carnap (1966, pp. 250–51) and Hintikka (1970, p. 9), that *high relative probability means low relative content*; that a universal statement that is rendered highly probable by evidence has, relative to that evidence, almost nothing of interest to say. I have suggested, tentatively enough, I hope, that this lack of content is revealed in the paucity of the unobserved part of the universe — as it should appear to anyone who assigns such high probabilities. And that if one wants to treat what has not been observed as still of interest, one must give up after all the craving for high probabilities.

REFERENCES

Carnap, R. (1952). *The Continuum of Inductive Methods*. Chicago: University of Chicago Press.

Carnap, R. (1966). "Probability and Content Measure," in P. K. Feyerabend and G. Maxwell, eds., *Mind, Matter, and Method*. Minneapolis: University of Minnesota Press. Pp. 248–60.

Hacking, I. (1971). "The Leibniz-Carnap Program for Inductive Logic," *Journal of Philosophy*, vol. 68, no. 19, pp. 597–610.

Hilpinen, R. (1966). "On Inductive Generalization in Monadic First-Order Logic with Identity," in J. Hintikka and P. Suppes, eds., *Aspects of Inductive Logic*. Amsterdam: North Holland. Pp. 133–54.

Hilpinen, R. (1968). *Rules of Acceptance and Inductive Logic*, in *Acta Philosophica Fennica*, vol. 22. Amsterdam: North Holland.

Hintikka, J. (1965). "Towards a Theory of Inductive Generalization," in Y. Bar-Hillel, ed., *Logic, Methodology and Philosophy of Science*. Amsterdam: North Holland. Pp. 274–88.

Hintikka, J. (1966). "A Two-Dimensional Continuum of Inductive Methods," in J. Hintikka and P. Suppes, eds., *Aspects of Inductive Logic*. Amsterdam: North Holland. Pp. 113–32.

Hintikka, J. (1970). "On Semantic Information," in J. Hintikka and P. Suppes, eds., *Information and Inference*. Dordrecht: Reidel. Pp. 3–27.

Hintikka, J., and R. Hilpinen (1966). "Knowledge, Acceptance, and Inductive Logic," in J. Hintikka and P. Suppes, eds., *Aspects of Inductive Logic*. Amsterdam: North Holland. Pp. 1–20.

[Hume, D.] (1740). *An Abstract of a Book lately Published entituled A Treatise of Human Nature, &c.* London: C. Borbet.

Jeffrey, R. C. (1968). "The Whole Truth," *Synthese*, vol. 18, pp. 24–27.

Kyburg, H. E., Jr. (1961). *Probability and the Logic of Rational Belief*. Middletown, Conn.: Wesleyan University Press.

Levi, I. (1967). *Gambling with Truth*. New York: Knopf.
Martin-Löf, P. (1966). "The Definition of Random Sequences," *Information and Control*, vol. 9, pp. 602–19.
Miller, D. (forthcoming). "The Acceptance World," *Theoria*.
Miller, D. (unpublished). "What's in a Numeral?"
Popper, K. R. (1967). "The Mysteries of Udolpho," *Mind*, n.s. vol. 76, no. 301, pp. 103–10.
Popper, K. R. (1972). *Objective Knowledge*. London: Oxford University Press.
Popper, K. R. (1974). *Replies*, in P. A. Schilpp, ed., *The Philosophy of Karl Popper*. La Salle, Ill.: Open Court.