

OBJECTIVE PROBABILITY THEORY THEORY*

ABSTRACT. I argue that to the extent to which philosophical theories of objective probability have offered theoretically adequate *conceptions* of objective probability (in connection with such desiderata as causal and explanatory significance, applicability to single cases, etc.), they have failed to satisfy a *methodological* standard – roughly, a requirement to the effect that the conception offered be specified with the precision appropriate for a physical interpretation of an abstract formal calculus and be fully explicated in terms of concepts, objects or phenomena understood independently of the idea of physical probability. The significance of this, and of the suggested methodological standard, is then briefly discussed.

1.

Philosophical discussions on the topic of probability have mainly focused on two kinds of issues, the first having to do with the *concept* of probability and the second having to do with *methodological* standards which an interpretation of probability (i.e., a philosophical theory about the nature of probability) itself must satisfy if it is to be an adequate interpretation. As to the first kind of issue, philosophical theories of probability must endorse some more or less vague intuitions about what kind of thing probability is, and the conception of probability offered must accommodate the intuitions endorsed. For example, it's generally thought that probability must satisfy some sort of standard axiomatization, such as Kolmogorov's; it's often thought that physical probability must be objective in that probability values are correct or incorrect independently of anyone's state of knowledge or beliefs about the correctness of the values; on the other hand, it's also often thought that probability can only be a measure of our ignorance; it's generally thought that probability must have predictive significance and appropriately reflect certain causal features of the physical world; and it's generally thought that probability, whatever it is, must be applicable to the "the single case", in particular, in contexts of rational decision making and of probabilistic explanation. Any theory of probability which doesn't endorse or accommodate sufficiently many such intuitions wouldn't constitute an interpretation of *probability*, but rather of something else if

anything else. Much recent philosophical work on probability has been devoted to developing conceptions of probability that are sensitive to certain intuitions and to arguing that one or another proposal does *not* adequately accommodate some such intuitions. Now this is not to deny, of course, that there may be several different useful interpretations of probability. And I don't mean to assert that all of the above intuitions are relevant to every conception of probability. Rather, the point is just that at least one desideratum relevant to assessing the adequacy of a philosophical interpretation of probability is that the *concept* offered must be theoretically adequate in some appropriate sense.

I shall divide the conditions which an interpretation of probability must satisfy in order for it to be theoretically adequate into two parts. I shall call the condition that the interpretation of probability offered must satisfy some standard axiomatization the *condition of admissibility*. This follows the terminology of Salmon (1967, pp. 62–63), except that Salmon uses the term 'criterion' instead of 'condition'. The condition that the concept offered must be otherwise theoretically adequate I shall call the *condition of conceptual adequacy*. This condition roughly corresponds to Salmon's "criterion of applicability", the force of which, he points out, may be succinctly summarized by Bishop Butler's aphorism, "Probability is the very guide of life".

The second kind of issue in philosophical discussions of probability has to do with philosophical methodology and general standards, or conditions of adequacy, which a philosophical theory of probability must satisfy independently of the particular conception of probability proposed. Thus Suppes (1973, 1974) has recently criticized Popper's (e.g., 1959) propensity interpretation on the grounds that it does not *formally characterize* probability in terms of ideas *understood independently* of quantitative probability, supposing that any adequate interpretation of probability must do this regardless of the particular conception of probability offered, whether it be subjective, Bayesian, single case propensity, hypothetical limiting frequency, etc. And Suppes (1973), Kyburg (1974, 1978), and Giere (1976) have recently attempted to develop the propensity interpretation in such a way that it satisfies what Giere (1976) calls "Suppes' Demand for a Formal Characterization of Propensities".

In the second section of this paper, I will elaborate Suppes' demand, dividing it into two parts. The *condition of formal adequacy* will demand that any adequate interpretation of probability provide us with

a definition of its characteristic kind of axiomatized structure \mathcal{C} – one instance, \mathcal{S} , of which will be the interpretation's "intended model", as explained below – where certain features of such a structure must be "representable" in terms of a probability function P on an appropriate structure \mathcal{B} , generally a part of a characteristic structure \mathcal{C} , as explained below. It will be seen that satisfaction of this condition is importantly related to the testability of a theory of probability in connection with satisfaction of the condition of admissibility. The *condition of interpretation/idealization* will take seriously the part of Suppes' demand – not adequately appreciated, I think, in Suppes (1973), Kyburg (1974, 1978), and Giere (1976) – that probability be characterized in terms of things *understood independently* of quantitative probability. This part of Suppes' demand itself has two parts: first, that the things in terms of which probability is explicated be *understood*, and second, that they be understood *independently* of the concept of quantitative probability. The condition of interpretation/idealization will demand specification of the intended model \mathcal{S} , alluded to above and explained more fully below, and that \mathcal{S} be a model of, an *idealization* of, some understood concepts, objects or phenomena – ideally, features of the observable, empirically accessible world – so that those things constitute an *interpretation* of the constituents of \mathcal{S} which at least roughly obey the relevant axioms which \mathcal{S} obeys. The "understood" part of the second part of Suppes' demand will be satisfied if we "understand" that those things obey the relevant axioms, and the "independently" part of the second part of Suppes' demand will be satisfied if it is shown that \mathcal{S} , the axioms characterizing \mathcal{S} , and the part of the world thereby modeled can be studied and characterized without appeal to concepts of quantitative probability.

Thus, the various conditions of adequacy which I shall advance work *together* to ensure that a philosophical theory of probability which satisfies them all will be adequate. Indeed, just as satisfaction of formal adequacy will play an important role in ensuring the testability of a theory in connection with admissibility, so satisfaction of interpretation/idealization will play an important role in testing whether or not the theory has adequately identified the intended concept of probability, a concept in virtue of which conceptual adequacy may be satisfied. Also in section 2, we shall see the connection between the condition of interpretation/idealization and Salmon's criterion of ascertainability, according to which it must be possible, in principle, to

ascertain the values of probabilities.

In sections 3–6, I shall examine instances of what I take to be the four main kinds of interpretations of probability according to which probability is objective, or physical, with an eye towards the extent to which they satisfy the conditions of adequacy elaborated in the first two sections. I shall examine the actual limiting frequency conception, attributed to Von Mises (1957, 1964) and Reichenbach (1949), the hypothetical limiting frequency view as formulated by Kyburg (1974, 1978), the “long run” construal of Popper’s propensity interpretation (e.g., 1957, 1959), and Fetzer’s (1971, 1981) “single case” propensity view. All of these views can be formulated in such a way that they satisfy the conditions of formal adequacy and admissibility. What I shall argue is that none of them satisfies *both* the condition of conceptual adequacy *and* the condition of interpretation/idealization and that to the extent that they satisfy one of these two conditions, they fail to satisfy the other. I shall argue that as far as conceptual adequacy goes, the theories rank from better to worse roughly in the following order: single case propensity, long run propensity, hypothetical limiting frequency, actual limiting frequency. And I shall argue that with respect to interpretation/idealization, these theories rank roughly in the opposite order.

It is perhaps worth noting that this general kind of tension between the satisfaction of two desiderata of adequacy is, of course, not new in philosophy, nor even in the philosophy of probability. Fetzer (1974) has noted a tension between satisfaction of “epistemological criteria” (on which actual frequency conceptions of probability seem to be preferable to a single case propensity account) and “systematic considerations” (on which the propensity interpretation is preferable). In the philosophy of mathematics, Benacerraf (1973) argues that no adequate account of mathematical truth has allowed for an adequate account of mathematical knowledge, and vice versa – i.e., roughly, that reasonable theories of mathematical *truth* leave it unintelligible how we can obtain mathematical *knowledge* while reasonable *epistemologies* fail to show how the suggested “truth conditions” are really conditions of *truth*. While my condition of interpretation/idealization is more methodological than epistemological in character, these tensions are of the same general kind as the one I shall argue is present in the philosophy of objective probability. Perhaps closer to the tension I shall try to characterize in objective probability theory is one that can be found in the philosophical foundations of modal logic. While the analysis of possible worlds as

maximally consistent sets of sentences makes the conception of possible worlds very *clear*, that conception is clearly also *theoretically inadequate* as a result, in part, of the limited expressive power of any available language. On the other hand, the analysis of possible worlds as, say, “ways the world could have been”, while perhaps closer to the *theoretically intended* conception, would seem to be *methodologically unsound*, in that it would render the usual analysis of possibility and of counterfactuality in terms of possible worlds circular.

2.

Although the finite relative frequency interpretation of probability, endorsed by Russell (1948) and mentioned favorably by Sklar (1970), has been forcefully criticized by many philosophers as being conceptually inadequate in several important respects, its basic features are relatively simple, and it will serve well as an example in terms of which the conditions of interpretation/idealization and formal adequacy (whose satisfaction is independent of satisfaction of the conceptual adequacy requirement) can be explained. On this interpretation, roughly, the probability of an attribute A in a reference class B is the relative frequency of occurrences of A within B , where A and B are the *finite* classes of *actual* occurrences of events of the relevant kinds.

To be more precise about the interpretation, we may define finite relative frequency structures (*FRF*-structures) \mathcal{C} as follows. Where E is any finite class and F is the power set of E (i.e., the set of all subsets of E) and $\#$ is a function which assigns to any member of F its cardinality (i.e., the number of its elements), $\langle E, F, \# \rangle$ is an *FRF-structure*. (Alternatively, F may be *any* Boolean algebra of subsets of E .) Thus, an *FRF-structure* is any triple $\langle E, F, \# \rangle$ that satisfies certain axioms, which axioms will guarantee that F is 2^E and that $\#$ is the cardinality function. Such a structure is an example of a characteristic structure \mathcal{C} , alluded to in section 1, where *FRF*-structures are (part of) what finite relative frequentists might use to get their interpretation of probability to satisfy the condition of formal adequacy. To complete the demonstration that the finite relative frequency theory satisfies the formal adequacy requirement, we show that certain features of an *FRF-structure* can be *represented* by a structure $\langle \mathcal{B}, P \rangle$, where \mathcal{B} is a Boolean algebra and P is a probability function on \mathcal{B} . For an *FRF-structure* $\langle E, F, \# \rangle$, simply let \mathcal{B} be F (alternatively, *any* Boolean algebra of subsets of E) and, for A ,

$B \in \mathcal{B}$, let $P(A) = \#(A)/\#(E)$ and $P(A/B) = \#(A \cap B)/\#(B)$. From the axiomatization of $\langle E, F, \# \rangle$, it can easily be shown that P is a probability on \mathcal{B} , i.e., that P satisfies probability axioms, where the arguments of P are the elements of \mathcal{B} . Thus, the characteristic kind of formal structure \mathcal{C} for the finite relative frequency interpretation – *FRF*-structures – has been characterized, and it has been shown how certain features of an *FRF*-structure can be represented in terms of a probability on an appropriate structure determined (at least in part) by the *FRF*-structure. And it is just these two things – the *definition of the characteristic kind of structure* with the capacity to yield *probabilistic representation* – that must be given for the condition of formal adequacy to be satisfied.

Of course the specification of the characteristic kind of structure and the probabilistic representation of certain features of such structures does not by itself constitute an appropriate *interpretation* of probability. For these may be just abstract mathematical entities, where objective probability is supposed to apply to the physical world. In order to complete the interpretation, therefore, both the “intended model” – an *instance* of the characteristic *kind* of structure – and the intended interpretation of the intended model must be specified. For the finite relative frequency interpretation, this may be done as follows. An *FRF*-structure $\langle E, F, \# \rangle$, is the *intended model* if there is a one-to-one correspondence between E and some set of events (or trials) in the physical world. Of course the intended model may be relativized to “local contexts”, where there is a one-to-one correspondence between E and the set of events of “local interest”, e.g., the set of all throws of dice, or of a particular die, or the set of all married American men who apply for life insurance at the age of 50. Then the sets in F correspond to the relevant properties, e.g., role of a die coming up “6”, a person’s dying before the age of 75, etc. The *intended interpretation* of the intended model is just the one-to-one correspondence between the elements of E and the relevant events in the world, and thus between the sets in F and corresponding properties, where $\#$ is interpreted as the cardinality function on the interpretation of F .

Thus, the condition of interpretation/idealization is satisfied when the intended characteristic structure (the intended model) and the theory’s interpretation of the intended model are both given. The reason why the condition of interpretation/idealization is so-called is that it concerns specification of the intended model of the theory and the intended

relation between that structure and the world, where (i) the relevant features of the world are an *interpretation* of that structure and (ii) the (perhaps abstract) entities, relations, functions, etc., of that structure, together with the structure's axiomatization, serve as an *idealization* of the relevant part of the world.

Talk of "*the* intended model", of course, is not meant to imply that there is, literally, *exactly one* such structure: rather, there is supposed to be (for the global context and for each local context) just one structure *modulo isomorphism*, where isomorphic structures are identified with one another. Also, I suppose it would be possible to collapse the two parts of interpretation/idealization – i.e., the specification of the intended model and the establishment of an interpretation of the model – into just one part, where, for the finite relative frequency interpretation, E is identified with the relevant set of events or trials in the world and the sets in F are identified with the relevant properties, construing properties extensionally as sets. But it is nevertheless worthwhile to distinguish conceptually between the role of the intended model (a structure whose constituents may be abstract mathematical entities and whose role is, in part, to show satisfaction of the condition of formal adequacy) and the role of the (ideally, physical and observable) entities in the world which the components of the intended model are interpreted as – the things to which probability is supposed to apply. For, in the case of the finite relative frequency interpretation, it seems that sets are indeed a rather crude *idealization* of properties which works well for the purposes of that interpretation. And in the second place, for other interpretations of probability (e.g., decision theoretic foundations of rational subjective probabilities, on which see, e.g., Eells 1982), both the *constituents* of the intended model and the *axiomatization* of the intended model are quite clearly very crude idealizations of the real-world entities and phenomena which they are supposed to model. The distinction in question is analogous to the distinction between two kinds of interpretations of Euclid's axioms for geometry: 'point' and 'line' can be interpreted *abstractly* as mathematical points (e.g., as pairs of real numbers) and abstract mathematical lines (e.g., sets of mathematical points $\langle x, y \rangle$ that satisfy $mx + b = y$ for some real numbers m and b); or they could be interpreted *physically* as physical points and physical lines (e.g., the possible paths of light rays). Similarly, the probability function may be interpreted abstractly as a function on an abstract intended model \mathcal{S} , and then also physically as a function on the features of the world modeled by

\mathcal{S} , via the connection between those features and \mathcal{S} established by satisfaction of the condition of interpretation/idealization.

Leaving aside for now the question of the conceptual adequacy of the finite relative frequency theory of probability (which will be discussed in the next section along with the conceptual adequacy of the actual limiting frequency interpretation), note how satisfaction of the conditions of formal adequacy and interpretation/idealization work together to ensure satisfaction of admissibility and to effect what Suppes (1974) calls "systematic definiteness" of the interpretation. Let A and B be any properties – or sets of events – in whose probabilities we may be interested, or in probabilistic relationships between which we may be interested. On the finite relative frequency theory, we must, guided by a local or global context, construct the intended model $\langle E, F, \# \rangle$, where there are sets, say A' and B' , corresponding to A and B , such that $A', B' \in F$. According to the rules given in the probabilistic representation part of the satisfaction of the condition of formal adequacy, we get the structure $\langle \mathcal{B}, P \rangle$, P being a probability on \mathcal{B} , where \mathcal{B} includes both A' and B' . Then, on the finite relative frequency interpretation of probability, the probability of A , and of A given B – in symbols, $\text{prob}(A)$ and $\text{prob}(A/B)$ – are just $P(A')$ and $P(A'/B')$. Also, we know that the interpretation satisfies the condition of admissibility, since P satisfies probability axioms; this is ensured by the interpretation's satisfaction of formal adequacy.

It should be clear that the conditions of formal adequacy and of interpretation/idealization must be satisfied by any satisfactory interpretation of probability. As to formal adequacy, an interpretation of probability is, after all, supposed to be an interpretation of the function symbol that appears in some axiomatization of probability, and it is difficult to see how it could possibly be shown that a purported interpretation of that symbol satisfies the axioms unless some kind of formal structure \mathcal{B} , e.g., a Boolean algebra, is provided by the theory. And if probability is to be some feature of the world – some kind of physical probability or even degree of belief – then the structure \mathcal{B} cannot come from just anywhere: it must be related to the world in some appropriate manner. Some features of the world that are understandable, or at least capable of being studied, independently of probability must be identified. And for these features to be systematically related to the structure \mathcal{B} , these features must first be idealized and abstractly represented in terms of some structure \mathcal{S} characteristic of the inter-

pretation of probability in question, so that one can demonstrably infer that \mathcal{B} , together with a probability P on \mathcal{B} , represents the appropriate features of the intended model \mathcal{S} , and thus, indirectly, the appropriate features of the world. The general picture is as indicated in Fig. 1, where the concepts, objects and phenomena appropriate to some familiar interpretations of probability other than the finite relative frequency interpretation are indicated. Note that the brackets on the left overlap, indicating that the specification of the *characteristic kind* of structure pertains to formal adequacy, where identification of the *intended* model pertains to interpretation/idealization.

It is of some interest to compare the interpretation/idealization requirement with Salmon's "*criterion of ascertainability*":

This criterion requires that there be some method by which, in principle at least, we can ascertain values of probabilities. It merely expresses the fact that a concept of probability will be useless if it is impossible in principle to find out what the probabilities are. (1967, p. 64)

The condition of interpretation/idealization is intended, in part, to capture the idea that probabilities should be ascertainable, but in a weaker sense than Salmon's criterion seems to require. The condition is only intended to ensure that probability statements have "empirical interpretation" – or "empirical content" – in a weaker sense similar to the one assumed by some versions of the hypothetico-deductive model of science. Consider Fig. 1. The entities of the top box can be thought of as *observable* (or, at least "*pre-theoretical*", i.e., "*pre-probability-theoretical*") *entities*, and the laws that govern them as *lawlike empirical generalizations* expressed in terms of an *observational* (or "*pre-probability-theoretical*") *vocabulary*. The concept of probability, as it figures in the bottom box, can be thought of as a *theoretical concept* of the philosophical theory of probability in question, while the probability axioms, together with the mathematical principles that relate the probability concept to the intended structure \mathcal{S} , can be thought of as the *theoretical* or *internal principles* of the philosophical theory of probability, which principles are expressed in terms of a *theoretical vocabulary*. And finally, the principles of interpretation/idealization, symbolized by the arrows between the top and middle box, can be thought of as *bridge principles* which relate the observable (pre-theoretical) entities of the top box to the mathematical entities of the middle box, in terms of which the theoretical concept of probability of

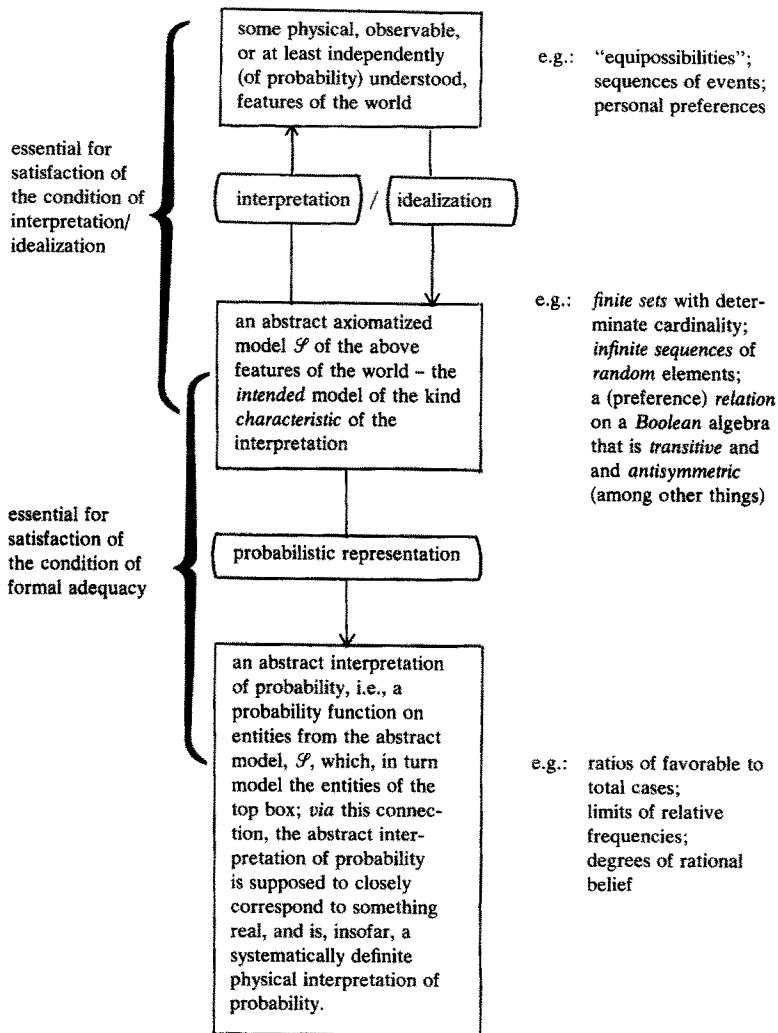


Fig. 1.

the bottom box is characterized *via* what I have been calling "probabilistic representation".

On the hypothetico-deductive model of science, of course, the bridge principles are supposed to function as *contextual* or *implicit* or *partial*

definitions of theoretical terms: they need not *completely* specify the meanings of the theoretical terms. Given this, and given that the principles of interpretation/idealization of a philosophical theory of probability are supposed to function much as the bridge principles of the hypothetico-deductive model of science, it should not be surprising if some interpretations of probability will not be able to specify empirical methods for ascertaining, even in principle, exact numerical values of all probabilities. The model of philosophical theories of probability presented in this section does not require that philosophical theories of probability be ascertainable in the stronger sense which Salmon seems to require. This would seem to be a virtue of initial neutrality between various philosophical theories of probability, for if the stricter version of ascertainability were insisted upon, it would seem that certain theories – certain propensity and dispositional accounts as well as certain Bayesian theories¹ – would be ruled out from the outset.

In concluding this section, I would like to emphasize again some important connections between the four conditions of adequacy suggested above and how they work together to ensure that an interpretation of probability that satisfies them all will be an adequate theory. We have seen that satisfaction of the condition of formal adequacy is required to demonstrate satisfaction of admissibility. Both formal adequacy and interpretation/idealization are required to show that the phenomena *in the world* upon which a given interpretation of probability focuses are indeed *probabilistic* phenomena in the sense that abstract probability theory applies to them. And finally, satisfaction of interpretation/idealization is supposed to identify precisely the intended concept and supply empirical (or, at least, “pre-probability-theoretical”) content – *cognitive significance* for positivists – for the *conception* of probability offered, without which it would seem that satisfaction of the condition of conceptual adequacy would be empty.

3.

The actual limiting frequency view of probability is a generalization of the finite relative frequency theory which is supposed to be applicable even if the relevant classes have infinite cardinality. My presentation of the characteristic kind of structure for the actual limiting frequency view will roughly follow that of Suppes' (1974), except for notation. An *ALF-structure* is any triple $\langle E, F, \# \rangle$ such that E is a sequence (by which

I shall understand any function whose domain is the natural numbers,² where I shall write ' E_i ' rather than ' $E(i)$ ', for the value of the function with argument i , F is the power set of the range of E (i.e., the set of all subsets of the set of possible values of E), and $\#$ is the binary function such that for any natural number n and element A of F , $\#(A, n)$ is the cardinality of the set $\{E_i: i \leq n \text{ and } E_i \in A\}$. (Alternatively, F may be taken to be any Boolean algebra of subsets of the range of E .) Probabilistic representation proceeds as follows in terms of relative frequencies. The *relative frequency of a set A ($\in F$) in the first n terms of E* is defined to be $\#(A, n)/n$. The *limiting relative frequency of A in E* is defined to be the limit of the relative frequency of A in the first n terms of E as n approaches infinity (if the limit exists), where the limit of a real-valued function $f(n)$ as n approaches infinity, $\lim_{n \rightarrow \infty} f(n)$, is defined to

be that number r (if any) such that for every $\epsilon > 0$, there is a natural number N_ϵ such that for all $n > N_\epsilon$, $|f(n) - r| < \epsilon$. Now let \mathcal{B} be any Boolean algebra of subsets of the range of E such that the limiting relative frequency of every element of \mathcal{B} in E exists. Then the structure $\langle \mathcal{B}, P \rangle$ is a probabilistic representation of the ALF-structure $\langle E, F, \# \rangle$, where for any A and B in \mathcal{B} ,

$$P(A) = \lim_{n \rightarrow \infty} \frac{\#(A, n)}{n},$$

and

$$P(A/B) = \lim_{n \rightarrow \infty} \frac{\#(A \cap B, n)}{\#(B, n)} = \frac{P(A \cap B)}{P(B)}.$$

An alternative approach (the one I shall have in mind in the sequel) would include in the axiomatization of ALF-structures the stipulation that all limiting frequencies of elements of F exist (Axiom of Convergence, or Limit Axiom), so that the set F could not in general simply be the power set of the range of E . This has the effect that for any ALF-structure $\langle E, F, \# \rangle$, there would be a uniquely characterizable probabilistic representation: the function P (defined above) on F itself. For conceptual adequacy, one might also want to include an Axiom of Randomness (Principle of the Excluded Gambling System), such as Von Mises', in the axiomatization of ALF-structures. In any case, it should be clear that the actual limiting frequency view of probability is admissible and formally adequate.³

For satisfaction of the condition of interpretation/idealization, the actual limiting frequency interpretation must specify an intended model, a particular ALF-structure, for the "global" context or any given "local" context. For the global context, one may insist on one-to-one correspondences between the natural numbers n and *times*, between the range of E and the set of past, present and future (actual) events, and between F and the relevant attributes. Then perhaps the natural number 1 could correspond to the "first" time, an element E_n of the sequence would correspond to what happens at the n th time, and so on. This assumes, of course, that the universe is temporally finite in the pastward direction, which assumption could be gotten around, though, by not insisting that E orders events temporally. That time is not dense is also assumed here. More plausibly, however, the limiting frequency view is applicable to local contexts, where the range of E is a set of events of local interest, e.g., tosses of a particular die, or of all dice, or human births, or applications for life insurance.

For the limiting frequency view of probability, unlike the finite relative frequency view, probability attaches to *ordered* sets of events, i.e., sequences of events whose order is given by the underlying sequence E of *all* events relevant to the local or to the global context. And it is clear that a solution to the problem of interpretation/idealization must specify an intended order in which the relevant events are to be taken. This is because, as long as there are infinitely many elements of E (or of a set B) that are elements of a set A and also infinitely many that are not elements of A , then the limit of the relative frequency of A 's in E (or in B) could be any number whatsoever between 0 and 1, inclusive, depending on the order in which the events are taken: in this case, *the order completely determines the probability*. Supposing that this is a problem for the actual limiting frequency view, then would it be a problem in connection with satisfaction of conceptual adequacy or a problem in connection with satisfaction of interpretation/idealization? The answer depends on the nature of the (perhaps more or less vague) *conception* of probability, whose theoretical adequacy must be certified if the theory is to satisfy the condition of conceptual adequacy, and which must be made precise and given empirical content if the theory is to satisfy the condition of interpretation/idealization. It is the job of interpretation/idealization to make that conception clear and precise (which, for the purposes of this paper, I am assuming may be theoretically adequate, or not, independently of its clarity and precision). So if the conception were vague

and noncommittal with respect to the order of the events, the problem described above would be one for interpretation/idealization: a solution to the problem of interpretation/idealization must then justify one order among the many possible as the intended one. But if the conception of probability offered were clear with respect to the intended order of events, then the actual limiting frequency theory would be conceptually adequate, or not, in part to the extent to which the order to which the conception is committed is itself theoretically justified. Since in all of the discussions of the application of probability construed as limiting frequency (which I have seen, at any rate) it is clear that the intended order of events is their temporal order, I shall assume that the *conception* of probability is clear about the intended order of the relevant events: it is their temporal order. Thus, the problem under discussion is not a problem for interpretation/idealization. Indeed, it seems that the actual limiting frequency interpretation fares quite well on the condition of interpretation/idealization: all of the relevant concepts – that of events, of temporal order, of cardinality, of sequences, of sets, of limits, etc. – are either fairly well understood or at least such as can be studied and understood independently of the concept of probability.

So the actual limiting frequency view (as understood for the purposes of this paper) will be conceptually adequate, or not, in part to the extent to which using the temporal order of events in calculating limiting frequencies is theoretically justified. I can think of no reasons in favor of or against using the temporal order rather than any other order, but the absence of reasons in either direction might itself suggest the argument that using the temporal order is *arbitrary*. Also, note that one effect of always using *one* order, such as the temporal order, of events is to make the probability of an attribute *A* – or the probability of an attribute *A* within a reference class *B* – invariant over time, where (although I have no natural suggestions along these lines) perhaps one way of accommodating the intuitive possibility that probabilities might *change* over time would be to use different orders of the relevant events at different times. The fact that, intuitively, probabilities seem to change over time (e.g., the probability of a person living more than 60 years has, intuitively, changed over time) – while it seems that, on limiting frequency conceptions, all probabilities are fixed for all time – will be recognized as a version of the reference class problem, or of the problem of the single case.

The frequentist response, of course, is that $P(A/B)$ – where B is the class of incidents of human births and A is the class of incidents of a being living more than 60 years – has *not* changed, but that different reference classes are appropriate for assessing probabilities with predictive significance at different times. Let C be the class of incidents of human births where the person will enjoy the results of modern medical advances throughout his life. Then $P(A/B \cap C)$ is one probability and $P(A/B \cap \bar{C})$ is another, where the first is more appropriate for predictive purposes today and the second would have been more appropriate, say, during the Dark Ages.

The problem of the single case for the actual limiting frequency theory of probability is usually formulated as follows. On the frequency view, probability attaches to (ordered) *collections* of events; but we are sometimes interested in the probability that, e.g., the *next* event will exemplify some attribute. The problem of the single case, then, asks how the frequency interpretation can apply to single events. And the general form of proposed solutions to the problem is to give a rule for choosing an appropriate reference class to which the single event belongs and to say that the probability of the relevant attribute within that reference class should be transferred to the single event in question. Hence, the problem is sometimes also called ‘the problem of the reference class’.

I do not believe that there is an adequate solution to the problem of the reference class for the actual limiting frequency view of probability. And in light of the fact that we would ideally like an explication of physical probability to have predictive and explanatory significance for single events (which may occur irreducibly probabilistically) and to have significance in connection with decision making in individual decision situations, this constitutes a serious limitation to the *conceptual adequacy* of the theory. Without going into much detail, let me summarize some of the considerations that have been, or could be, brought to bear against two proposed solutions to the problem.

Reichenbach’s solution to the reference class problem, of course, was to choose “*the narrowest class for which reliable statistics can be compiled*” (1949, p. 374). Thus, in relation to the above example, to assess the probability that a particular individual will live more than 60 years, it is better to use one of $B \cap C$ and $B \cap \bar{C}$ as the reference class than just B , depending on which class the individual in question belongs to, provided that reliable statistics – with respect to life span –

can be compiled for the two classes. It is also part of Reichenbach's solution that we do not narrow a class with respect to another class when the second class is "known to be irrelevant" (1949, p. 374); that is, if class D is known to be such that $P(A/B \cap C \cap D) = P(A/B \cap C)$, then we should not favor the use of the narrower class $B \cap C \cap D$ over use of $B \cap C$. Note two features of this solution. First, there is a subjectivist element in the solution, in that the choice of reference class depends on the reliability of our knowledge of the relevant statistics. Second, as Salmon (1971, p. 41) has pointed out, it is often the case that the more reliable the statistics, the broader the reference class must become, and the narrower the reference class, the less reliable the statistics become. This is in part because, for classes A , B , and C , $P(A/B \cap C)$ is not a function of – cannot be calculated from only – $P(A/B)$ and $P(A/C)$.

Now Reichenbach insisted, of course, that, literally speaking, probability applies only to sequences. And in connection with the single case, he says,

We are dealing here with a method of technical statistics; the decision for a certain reference class will depend on balancing the importance of the prediction against the reliability available. (1949, p. 375)

In a similar vein, Salmon suggests that, plausibly, Reichenbach

was making a distinction similar to that made by Carnap between the principles belonging to inductive logic and methodological rules for the application of inductive logic. The requirement of total evidence, it will be recalled, is a methodological rule for the application of inductive logic. Reichenbach could be interpreted as suggesting analogously that probability theory itself is concerned only with limit statements about relative frequencies in infinite sequences of events, whereas the principle for selection of a reference class stands as a methodological rule for the practical application of probability statements. (1971, p. 41)

But one may nevertheless insist that an account of physical probability is conceptually inadequate unless, on the account, probability applies objectively to single events, which is not implausible if one thinks, in the first place, that ideally what one would like to know in particular decision problems is the controlling objective probabilities, and, in the second place, that there can be correct and complete statistical explanations of particular events that may occur irreducibly probabilistically. Such an account of probability cannot let the values of probabilities depend on the incomplete state of our knowledge, which a

purely methodological account of single case probabilities must require.

But, in light of the distinction Salmon suggests Reichenbach might have had in mind, perhaps the actual limiting frequency view can be elaborated in such a way as to apply objectively to the single case. To get around being forced into a *methodological* context by the *incompleteness* of our state of knowledge, we simply envisage, for the purpose of remaining in a nonmethodological context of *explicating* single case objective probabilities, a hypothetical state of *complete* knowledge with respect to the relevant facts. This would seem to have the promise of eliminating the subjectivity of Reichenbach's solution, eliminating the practical conflict between reliability of statistics and narrowness of the reference class, while preserving the frequency conception of probability. I think that part of Fetzer's paper "Reichenbach, Reference Classes, and Single Case 'Probabilities'" (1977; see also his 1981, pp. 78–86) can be viewed as following up this idea and showing that it can only result in a trivialization of single case probabilities on the frequency view, where all such probabilities will turn out to be either 0 or 1.

Following up Reichenbach's general ideal that "the probability will approach a limit when the single case is enclosed in narrower and narrower classes, to the effect that, from a certain point on, further narrowing will no longer result in noticeable improvement" (1949, pp. 375–76), Fetzer defines an *ontically homogeneous reference class with respect to an attribute A and a trial (single event) x* (roughly) as a class *B* such that $x \in B$ and for all $B' \subseteq B$, $P(A/B) = P(A/B')$, and he suggests that on Reichenbachian principles, the appropriate reference class for *x* relative to *A* would be some ontically homogeneous reference class with respect to *A* and *x*. But which one? Presumably, the appropriate one would be the first one that one "reaches" in successively narrowing some initial candidate with respect to which Fetzer calls *permissible predicates* (i.e., predicates which are permissible for use in the description of a reference class): predicates which *do not imply the presence or absence of the relevant attribute, which are not satisfied by at most a finite number of things on logical grounds alone, and which are satisfied by at least one thing* (presumably the single event in question). And it is then argued, on the basis of the principle

If *x* and *y* are different events, then there is a permissible predicate *F* such that Fx and $\bar{F}y$,

that the set of permissible predicates satisfied by a single case x will not all be satisfied by any other single case y , so that the "appropriate" reference class turns out to be just $\{x\}$ and all single case probabilities turn out to be either 0 or 1, depending on whether the single event in question lacks or has the relevant attribute.

Note that it seems that Fetzer's argument assumes that conjunctions of permissible predicates are themselves permissible. This, of course, would need some argument, for it is obviously possible for two predicates, neither of which is satisfied by at most a finite number of events on logical grounds alone, to have a conjunction which is. For example, let F be the predicate "is the sinking of the *Titanic* or happens before the year 1900" and let G be the predicate "is the sinking of the *Titanic* or happens after the year 1900". As another example, let F specify just spatial coordinates and G a time. Perhaps the intent, however, is that permissible predicates must all be *dispositional* predicates of some kind, where his theory of dispositions (see especially his 1981, pp. 160–61 and 190–92) would somehow ensure that conjunctions of permissible predicates will be permissible. In any case, Fetzer's *main* point still holds, namely that on actual limiting frequency conceptions of probability, it is impossible to distinguish between factors that are statistically relevant *because of a "real causal" connection* and those which are statistically relevant *purely by coincidence*. And whether or not it is always possible to describe a given event *uniquely* in terms of permissible predicates is not so much at issue.

Suppose that in fact some event x is the only event in the course of the world's actual history that satisfies each of the predicates F_1, \dots, F_n , each assumed to be permissible in some correct sense of 'permissible'. Then the actual limiting frequency of the relevant attribute, say A , in the class, say B , of individuals that satisfy each of F_1, \dots, F_n is either 0 or 1, depending on whether x lacks or has attribute A . And it is surely possible that, for any i between 0 and n , there are *many* events which satisfy each of $F_1, \dots, F_{i-1}, F_{i+1}, \dots, F_n$. And, where, for each such i , B_i is the reference class of events that satisfy each of $F_1, \dots, F_{i-1}, F_{i+1}, \dots, F_n$, it is clearly also possible that the $P(A/B_i)$'s (on the actual limiting frequency interpretation) all differ from each other and from $P(A/B)$ (on the actual limiting frequency interpretation). And all this is consistent with *none of the F_i 's having anything physically to do with the presence or absence of A in x : it just happens that x is the only event in the course of the world's actual history that satisfies each of the F_i 's.*

And the fact that it is possible that there be *just one* event is not so much to the point. Suppose that in fact, in the entire course of the world's actual history, there will be only three instances of a well-balanced coin being fairly tossed when the moon, Mars and Halley's comet are in close opposition. Then the actual limiting frequency of a given such event resulting in tails-up will be either 0, $\frac{1}{3}$, $\frac{2}{3}$, or 1, yet, intuitively, it is absurd to conclude that the celestial configuration described introduces a *physical bias* into the trial, just because removal of any of the three factors (moon in opposition, Mars in opposition, Halley's comet in opposition) yields a limiting relative frequency of about $\frac{1}{2}$. Intuitively, it would seem that the *probability* of tails on any of the three tosses is (about) $\frac{1}{2}$, though of course the *actual frequency* of tails in the circumstances described is artificially limited to being one of the four values specified above.

Relative frequentists could respond to such examples in a number of ways. They could, for example, say that in order for an actual relative frequency to be the true single case probability, one must use a large enough reference class for which reliable statistics are available. Thus, Salmon (1971) urges that, instead of using the *narrowest* such reference class, we should use the *broadest* homogeneous such reference class to which the single case in question belongs, where a *homogeneous reference class for an attribute A* is defined to be a class for which it is impossible to effect a statistically relevant partition (with respect to A) without already knowing which elements of the class have attribute A and which do not. This takes seriously Reichenbach's idea that, "Classes that are known to be irrelevant for the statistical result may be disregarded" (1949, p. 374). But still, it would seem that as long as we are dealing with actual relative frequencies, such a partition could be statistically relevant just as a matter of coincidence, as when we partition the class of tosses of honest coins by whether or not they occur when the moon, Mars and Halley's comet are in close opposition. See Fetzer (1977, pp. 199–201; and 1981, pp. 91–92) for another kind of criticism of Salmon's approach and for further discussion.

On the other hand, one may simply insist upon the use of some infinite sequence of events which are similar to the single event in question in all relevant (causal) respects. But, first, this would require an explication of causal relevance prior to an explication of probability, where this would render circular recent attempts to explicate causality in terms of probability relations.⁴ And second, even if this could be

done, it is possible that there might be, say, only two or three events in the course of the world's actual history that are similar to the single event in question in all relevant respects.

A frequentist may respond to this last difficulty along the following lines, as Salmon (1979, pp. 11–12) has suggested that Reichenbach would (see his 1949, §34). Instead of considering the three tosses discussed above as members of some *actual* (and thus possibly finite) sequence, consider them as members of the sequence of tosses that *would* exist *were* we to toss the coin infinitely many times under the same circumstances, and then ask what the limiting relative frequency of tails would be in this sequence. Of course this is to abandon the idea that probability should be explicated in terms of sequences of *actual* events. In the next section, we look at the *hypothetical* limiting frequency interpretation, which attempts to specify appropriate principles for extending actual finite sequences (e.g., single element sequences) to hypothetical infinite sequences.

But as to the *actual* limiting frequency interpretation of probability, it seems correct to conclude that, while the theory is basically *adequate as far as interpretation/idealization* (and formal adequacy and thus also admissibility) goes, it is *conceptually inadequate* in that, although it may be argued to have appropriate predictive significance, it is incapable of characterizing the difference between genuine physical connections and merely historical coincidences and is, largely for this reason, incapable of applying appropriately to single events.

4.

In order to accommodate some of the difficulties discussed above in connection with actual relative frequency conceptions of probability, a hypothetical limiting frequency conception may be advanced, according to which the probability of an attribute *A* – or the probability of *B*'s being *A*'s – is equal to the limiting frequency of *A* in a hypothetical infinite extension of the actual (finite) sequence of events – or a hypothetical infinite extension of the actual (finite) sequence of *B*'s. Thus, $P(A)$ – or $P(A/B)$ – is supposed to be what the limiting frequency of *A* *would be* – or what the limiting frequency of *A* in the sequence of *B*'s *would be* – if the world's history *were* infinite – or if the sequence of *B*'s *were* infinite. This is the basic *conception* of probability on hypothetical limiting frequency views, which conception must be

given precision in a solution to the problems of formal adequacy and of interpretation/idealization if the interpretation is to be adequate.

Kyburg (1974, 1978) formulates semantics for the hypothetical limiting frequency view (which may be viewed as a solution to the problem of formal adequacy and part of a solution to the problem of interpretation/idealization) as follows. He begins with a first order language with identity and enough mathematical machinery to axiomatize the three place predicate S in such a way that ' $S(A, B, r)$ ' is true in a model $M = \langle U, R \rangle$ (where R is a set of relations on U and functions on U, U^2 , etc., and U contains at least the empty set, \emptyset) if and only if⁵

- (i) B is an infinite sequence of sets, B_1, B_2, \dots , where $B_i \subseteq B_j$ if $i < j$,
- (ii) A is a set,
- (iii) r is a real number, and
- (iv)
$$r = \lim_{i \rightarrow \infty} \frac{\#(A \cap B_i)}{\#(B_i)}.$$

The relations and functions of these models are assumed to be "compatible" in the sense that if $\langle U, R \rangle$ and $\langle U', R' \rangle$ are any two such models with $U \subseteq U'$, then (i) for every predicate symbol A of the language, $R(A) \subseteq R'(A)$ (where $R(A)$ and $R'(A)$ are the relations which R and R' assign to A) and (ii) for any k -place function symbol f of the language, if $x \in U^k$, then either $R(f)(x) = R'(f)(x)$ or $R(f)(x) = \emptyset$. (The reason for insisting that \emptyset is an element of every model is so that what would otherwise be a partial function may take \emptyset as a value where it would otherwise be undefined.)

The "actual world" is taken to be a particular model, $M^* = \langle U^*, R^* \rangle$. A *future model* is any model in which every (actually) true observation sentence pertaining to times up to the present is true. And a *lawful model* is any model in which all (actually) true universal (nonstatistical) physical laws are true. A model $M' = \langle U', R' \rangle$ is an *extension* of a model $M = \langle U, R \rangle$ if $U \subseteq U'$ and the relations and functions in R are the restrictions to U of the relations and functions in R' . Finally, for a term B interpretable as a sequence, a model M' is a *B-maximal extension* of a model M if M' is an extension of M and either B is infinite in M' or no extension of M' extends B . Thus, while all lawful future worlds may have finite histories, *B*-maximal extensions of such worlds may have infinite histories. Finally, truth conditions for hypothetical limiting frequency statements, ' $P(A/B) = r$ ', are given as follows:

' $P(A/B) = r$ ' is true (in the actual world) just in case ' $S(A, B, r)$ ' is true in every (or "almost every") B -maximal extension of every lawful future world.

This is not *exactly* the same semantics for hypothetical relative frequency statements as that given by Kyburg – indeed, he considers several variations – but it is close and captures all the features of the hypothetical limiting frequency interpretation which I wish to discuss.

I shall structure the discussion of the hypothetical limiting frequency view around the conditions of adequacy discussed in the first two sections of this paper. As to conceptual adequacy, the rough conception of probability offered was presented above, and all I have to say about the conceptual adequacy of the view is that it seems clearly superior to the actual limiting frequency conception in that it deals with the possibilities that the actual history of the world is *finite* (where the actual frequencies may, in some cases, be *ratios of small numbers* that don't faithfully represent the relevant features of the physical world) and that it may exhibit *coincidences*. It deals with the first possibility by envisioning hypothetical *infinite extensions* of the actual world's history and with the latter possibility by considering *many* extensions of the actual world's history, where, presumably by invoking the law of large numbers idea, this is taken to accommodate actual world coincidences of a global character (but see below on this). The solutions to the problems of formal adequacy and interpretation/idealization are then supposed to clarify this rough conception.

Recall that a solution to the problem of formal adequacy is supposed to identify the *characteristic kind of structure*, some features of which can be given *probabilistic representation*. I suggest that a characteristic structure for the hypothetical limiting frequency interpretation be understood to be of the following form (variants are possible, as discussed below): an *HLF-structure* is an (appropriately axiomatized – see below) sextuple, $\langle \mathcal{L}, \mathcal{M}, M^*, F, L, E \rangle$, where \mathcal{L} is a first-order language with axioms (at least for the three-place predicate S), \mathcal{M} is a set of models $M = \langle U, R \rangle$ of the kind defined above, $M^* \in \mathcal{M}$, F and L are subsets of \mathcal{M} , and E is a relation in $\mathcal{L} \times \mathcal{M} \times \mathcal{M}$. Then truth conditions for hypothetical limiting frequency statements could alternatively be given as follows:

' $P(A/B) = r$ ' is true (relative to a given *HLF-structure*) if and only if for every $M \in \mathcal{M}$ such that $F(M)$ and $L(M)$,

' $S(A, B, r)$ ' is true in every (or almost every) $M' \in \mathcal{M}$ such that $E(B, M, M')$.

Also, if "every (or almost every)" could be made precise in an appropriate way, it would be possible to infer, in light of these semantics, a finitely additive probability space from a structure characteristic of the theory, thus explicitly satisfying the probabilistic representation part of the condition of formal adequacy as formulated above. Of course the intended model of the hypothetical limiting frequency interpretation would have M^* correspond in some appropriate way to the actual world, where F is the set of future worlds (as described above), L is the set of lawful worlds (as described above), and E is the relation such that $E(B, M, M')$ is true in the intended model if and only if M' is a B -maximal extension of M (as described above). And characteristic structures would be axiomatized in such a way as to guarantee that the purely formal relations between, e. g., worlds M and B -maximal extensions of M would hold, e.g., the axioms should imply that if $E(B, M, M')$ is true in a characteristic structure, then $U \subseteq U'$, where $M = \langle U, R \rangle$ and $M' = \langle U', R' \rangle$.

Before further investigating the formal adequacy of the hypothetical limiting frequency theory, it is worth pointing out that *HLF*-structures could have been construed differently. For example, L could be construed as a binary relation on \mathcal{M} , where $L(M, M')$ holds in the intended *HLF*-structure if and only if all the universal laws that hold in M also hold in M' , and similarly for F . Or L could be thought of as a function from \mathcal{M} to subsets of \mathcal{L} , where, in the intended model, $L(M)$ is the set of universal laws true in M . Then the set of lawful-relative-to- M models could be identified in the obvious way – and similarly for F .

As far as formal adequacy goes, it seems that the only unclarity in the hypothetical limiting frequency theory is in connection with the phrase "every (or almost every)". Which is it? Without at least a specification of which it is, the truth conditions for ' $P(A/B) = r$ ' aren't definite, and it would not be possible to construct a finitely additive probability space from an *HLF*-structure in the light of the given semantics. Consider first the possibility of reading the phrase as "every". This would surely give us formal adequacy of the theory, but it would render *every* probability statement false in the intended model (i.e., given the intended meanings of 'future model', 'lawful model', etc.), thus rendering the interpretation inadequate in relation to interpretation/

idealization. Consider, for example, statements of the form 'The probability of tails on a "fair" toss of this coin is r ', and let us assume that this statement form – in symbols, ' $P(A/B) = r$ ' – yields a true statement (intuitively) just when ' $\frac{1}{2}$ ' is substituted for ' r '. Given the truth conditions for statements of hypothetical relative frequency stated above – and reading "every (or almost every)" as "every" – a statement ' $P(A/B) = r$ ' is true if and only if ' $S(A, B, r)$ ' is true in *every* B -maximal extension of every lawful future world. But surely there is *some* such world in which ' $S(A, B, 1)$ ' is true, i.e., in which the limiting relative frequency of tails is 1. And, as Skyrms says,

On the hypothesis that the coin has a propensity of one-half to come up heads on a trial and that the trials are independent, *each infinite sequence of outcomes is equally possible. If we look at all physically possible worlds, we will find them all*, including the outcome sequence composed of all heads. (1980, p. 32)

That is, there is also *some* B -maximal lawful future world in which ' $S(A, B, 0)$ ' is true. Thus, for *no* value of r is ' $S(A, B, r)$ ' true in *all* B -maximal extensions of lawful future worlds. So, let us abandon the "every" reading of "every (or almost every)" and consider now the "almost every" reading.

How are we to understand "almost every" in a precise way? We surely cannot take it to mean "all but a finite number", for, in the coin tossing example of the previous paragraph, if there is one world in which ' $S(A, B, r)$ ' is true for some value of r , then surely there are infinitely many such worlds. Fetzer and Nute (1979, 1980; see also Fetzer 1981, 56ff) have suggested the following way of making Kyburg's truth conditions precise on the "almost every" reading. Where M_1, M_2, \dots , is an infinite sequence of B -maximal extensions of lawful future worlds,

$$'P(A/B) = r' \text{ is true (in the actual world } M^*) \text{ if and only if} \\ \lim_{k \rightarrow \infty} \frac{\# \{M_i : i \leq k \text{ and } 'S(A, B, r)' \text{ is true in } M_i\}}{k} = 1.$$

Actually, this is a slight variant of Fetzer and Nute's suggestion, which is closer to Kyburg's formulation: the former assume that the worlds M_i are *themselves* future and lawful, despite their having *infinite* sequences of B 's (more on this general idea below).

There are several difficulties with this proposal. First, there are, presumably, at least continuum many B -maximal extensions of any lawful future world, where, again, ' B ' means 'this coin tossed' and ' A '

means 'tails': for each infinite sequence of heads and tails, there is at least one B -maximal extension of any lawful future world, and infinite sequences of heads and tails can be identified with functions from the natural numbers into {heads, tails}, of which there are continuum many. So the natural question at this point is, "On what principles do we select the *denumerably* long sequence M_1, M_2, \dots , of B -maximal extensions of lawful future worlds from the *nondenumerably* many such worlds?" Of course for any value of r , with $0 \leq r \leq 1$, there are infinitely many sequences of B -maximal extensions of any lawful future world such that, using any one of them, the truth conditions suggested will yield the truth of ' $P(A/B) = r$ '. This is because there are infinitely many infinite sequences of heads and tails for which the limiting relative frequency of tails is r , for any value of r , $0 \leq r \leq 1$.

Now *part* of the problem of selecting an appropriate sequence of B -maximal extensions of lawful future worlds would be solved if plausible principles governing which B -maximal extensions of lawful future worlds should be *elements* of such a sequence could be provided. Perhaps we should require that such B -maximal extensions of lawful future worlds themselves be lawful. (It seems that part of the intent of the definition of 'extension' is that all extensions of future worlds will themselves be *future*, for it is natural to assume that the sequences in R of a world $\langle U, R \rangle$ are sequences of events taken in their temporal order, so that all B -maximal extensions of lawful future worlds will automatically be future worlds.) But it would require an argument to establish that *any* B -maximal extension of *any* lawful future world is lawful, if $R(B)$ is supposed to be *infinite* in any B -maximal extension $\langle U, R \rangle$ of any lawful future world: conceivably, the universal laws true of the actual world might imply that the world will have a finite history, as some cosmological models predict. Of course Kyburg's definition of B -maximal extension has a clause in it to handle the case in which there is no extension which extends a finite sequence B . But it is not explained *why* there may be no such extension in some cases, e.g., whether or not it could be a matter of physical law. And note that whether or not the actual world's history is finite as a matter of physical law should not, intuitively, control whether or not some probabilities have irrational values.

But even if plausible principles for selecting a denumerable *set* of B -maximal extensions of lawful future worlds could be given, there would remain the problem of *ordering* the models in this set so as to obtain the infinite *sequence* of models required for the truth conditions

to be applicable. As long as there are infinitely many *not*- $S(A, B, r)$ -worlds as well as infinitely many $S(A, B, r)$ -worlds in a given set of B -maximal extensions of lawful future worlds, the truth of ' $P(A/B) = r$ ' on the suggested truth conditions will depend on the particular *order* in which the set of B -maximal extensions is taken, for any value of r , for this order will determine whether or not the main sequence in the truth conditions converges to 1. Note also that if there are infinitely $S(A, B, r)$ -worlds and infinitely many $S(A, B, s)$ -worlds in a given set of B -maximal extensions of lawful future worlds, there will be infinitely many orderings of the set which will yield the truth of ' $P(A/B) = r$ ' and also infinitely many that will yield the truth of ' $P(A/B) = s$ ', and this conditional statement holds for *any* values of r and s .

Now none of these difficulties with Fetzer and Nute's suggestion is a deep one for the problem of formal adequacy. Instead of taking \mathcal{M} to be a set of models in the *HFL*-structures, we could insist that \mathcal{M} be some infinite sequence of models. For probabilistic representation, simply take some largest Boolean algebra of terms (more precisely, equivalence classes of terms A, B under the relation $\mathcal{L} \vdash A = B$) for which Fetzer and Nute's truth conditions give probabilities. But that is an additional constituent of *HFL*-structures which has to be accommodated in characterizing the *intended* model of the theory, for satisfaction of the condition of interpretation/idealization. Note that there is no "natural" ordering of the *worlds*, whereas the actual limiting frequency theory is able to take advantage of the natural temporal order of *events*. The difficulties elaborated above are indeed deep and sticky problems for the hypothetical limiting frequency interpretation in connection with the condition of interpretation/idealization.

I conclude that while the hypothetical limiting frequency interpretation is *superior to the actual limiting frequency view as far as conceptual adequacy is concerned*, it is *inferior with respect to interpretation/idealization*.

5.

The theoretical advantage, discussed in the previous section, of the hypothetical limiting frequency conception over the actual limiting frequency conception was that the former deals with the possibilities that the history of the actual world is *finite* and that it *may exhibit*

coincidences that are not representative of the relevant physical features of the world. Thus, the artificial restriction of the actual limiting frequency of tails, in tosses of honest coins when the moon, Mars and Halley's comet are in close opposition, to the values $0, \frac{1}{3}, \frac{2}{3}$, and 1 results from the *finitude* of the relevant sequence; and the fact that the actual limiting frequency turned out to be, say, $\frac{1}{3}$, rather than, say, 0 or $\frac{2}{3}$, is a *coincidence that doesn't appropriately reflect the relevant physical facts*, e.g., the symmetry of the coin, the physical "honesty" of the toss, etc. By considering a hypothetical *infinite extension* of the actual sequence of such tosses, we do not artificially limit the possible values of the limiting frequency to the values $i/3$, for $i = 0, 1, 2, 3$. And considering *infinitely many such* hypothetical extensions of the actual sequence of just three tosses is supposed to accommodate the possibility that even in a lawful future world (hypothetically extended to include an infinite sequence of tosses under the relevant circumstances) the limiting frequency could be 0 or 1, or anything in between, however "improbable" such a value may be – where, as we have seen, characterizing the appropriate sense of "improbable" here is a sticky problem for hypothetical limiting frequentists in connection with interpretation/idealization.

But perhaps (at least part of) the motivation for adopting what has come to be called a *propensity* view of probability stems from difficulties that confront even hypothetical relative frequency views in connection with probabilities of single events. Let x_1, x_2 , and x_3 be the three actual tossings of a fair coin under the celestial circumstances described earlier. And suppose that the actual relative frequency of tails in these three tosses is, in fact, $\frac{1}{3}$. We are interested in what the probability is that x_3 results in tails. Intuitively, let us assume for now, this single case probability is $\frac{1}{2}$. Let 'B' denote, in the model $M^* = \langle U^*, R^* \rangle$ corresponding to the actual world, the property of being an honest coin tossed honestly under the celestial circumstances described earlier, and 'A' the property of coming up tails. Should we identify the single case probability of x_3 's coming up tails with the hypothetical limiting frequency theory's construal of $P(A/B)$? Suppose that, on this construal, $P(A/B)$ were $\frac{1}{2}$, as we should expect (given, of course, some adequate solution to the problem of interpretation/idealization for the hypothetical limiting frequency theory). This would be some evidence in favor of the hypothetical limiting frequency interpretation's applicability to single cases, as well as its appropriate applicability to

sequences of events. But there are reasons why $P(A/B)$, under the hypothetical limiting frequency construal, may be equal to $\frac{1}{2}$ *other* than each element of a hypothetical infinite extension of $R^*(B)$ having, intuitively, a single case probability of $\frac{1}{2}$ of coming up tails. Suppose, for example, that during the first phase of any occurrence of the celestial configuration described – which lasts for half of the time of any such close opposition – all physically symmetrical coins are, in fact, *physically biased* roughly 2:1 for tails, while during the rest of the time of such a close opposition, all such coins are physically biased roughly 2:1 for heads, where the celestial configuration during these times is responsible, causally, for these biases. (Perhaps an example involving the tides would be more intuitive here.) Then $P(A/B)$, as a hypothetical relative frequency, should *still* be expected to be about $\frac{1}{2}$ – since about “half” (i.e., limiting relative frequency of $\frac{1}{2}$) of the elements of a hypothetical infinite extension of $R^*(B)$ may be expected to occur during the first phase, and the other “half” during the second phase, of close opposition – but, intuitively, if x_3 actually takes place during the second phase, the single case probability of x_3 ’s being a member of A should be roughly $\frac{1}{3}$.

Some ways to accommodate such cases as this readily suggest themselves, the basic idea behind all of them being that the reference class must be chosen properly. In the above case, for example, the appropriate class is not an extension of $R^*(B)$, but rather an extension of $R^*(B')$, the class of tosses of coins *during the second phase* of the celestial configuration described. But how shall the appropriate reference class be characterized in general?

One possibility is to say that the single case – x_3 in our example – should first be characterized uniquely by some set of permissible predicates – predicates permissible for the description of a reference class. Say that F_1, \dots, F_n are all permissible predicates and that x_3 is the only event in the course of the world’s actual history that satisfies each of these predicates. Then we might identify the probability of x_3 ’s being in A as the hypothetical limiting frequency of A in the class B' determined by F_1, \dots, F_n . Unlike the actual limiting frequency view, it doesn’t follow for hypothetical limiting frequencies that $P(A/B')$ is either 0 or 1. But this suggestion will not do, of course. And the reason is that F_1, \dots, F_n need not have anything to do, physically, with whether x_3 results in heads or tails, in order for them uniquely to pick out the event x_3 from all other events in the course of the world’s

history. If x_3 were the *only* actual event of a coin's being tossed under the celestial circumstances described above, then the F_i 's need only to describe that celestial configuration, so that, as we have seen, the hypothetical limiting frequency of A in B' may be $\frac{1}{2}$, even though the correct single case probability of x_3 's resulting in tails is, intuitively, $\frac{1}{3}$.

Two ways of accommodating this problem, consistent with a hypothetical limiting frequency conception of probability, suggest themselves. They are rough analogues for the hypothetical limiting frequency view of the two proposed solutions of the reference class problem for the actual limiting frequency interpretation, considered two sections back: (1) say that the appropriate class B' is the one determined by *every* permissible predicate actually satisfied by x_3 , and (2) say that the appropriate class B' is the broadest ontically homogeneous reference class for x_3 and A , in the sense of section 3, i.e., the broadest class such that $x_3 \in B'$ and for any $B'' \subseteq B'$, $P(A/B'') = P(A/B')$, where P , here, is still hypothetical limiting frequency. Thus, according to (1), an appropriate hypothetical infinite extension of $\{x_3\}$ will be a "narrowest" class, while, according to suggestion (2), such an extension will be a "broadest" class, not all elements of which satisfy every permissible predicate which x_3 satisfies.

Suggestion (1) could be criticized on the ground that causally irrelevant factors should not be included in the description of the reference class. But Eells and Sober (1983) argue that the *values* of hypothetical limiting frequencies will not be affected by the specification of causally irrelevant factors. But there is still a difficulty with suggestion (1), in connection with probabilistic *explanation*. Salmon's (1971) well-known counterexamples to Hempel's requirement of maximal specificity tell against the suggestion. I shall not rehearse these considerations here in detail. But the basic idea is that if we must specify *all* permissible predicates in the reference class description, then causally and explanatorily irrelevant factors will be specified as well as those that are relevant. But it seems that in explaining why an event exhibited some attribute, we should assign the event to a reference class determined by only the causally or otherwise explanatorily relevant factors, to avoid citing explanatorily irrelevant factors in the explanation of the character of the event.

As to suggestion (2), a simple example of Fetzer's (1981, p. 91; also 1977, pp. 199–200) – also in an explanatory context – tells quite

conclusively, it seems to me, against the idea. Suppose that Jones died of a brain tumor. Of course not everyone who has a brain tumor dies of it; assume, in fact, that brain tumors of the kind Jones had are irreducibly probabilistic causes of death. Say that $P(D/R \cap T)$, the hypothetical relative frequency of death among 60 year old human males – etc. – with a brain tumor, is r , in fact the *correct* “single case” value for Jones’s death. But suppose it is also true that $P(D/R \cap H)$, the hypothetical limiting frequency of death among 60 year old human males – etc. – *with a certain serious kind of heart disease*, is also equal to r . Then $R \cap T$ is not a *broadest* ontically homogeneous reference class for Jones and D . Any such reference class must have $(R \cap T) \cup (R \cap H)$, i.e., $R \cap (T \cup H)$, as a subset. Suppose that $R \cap (T \cup H)$ in fact is a *broadest* ontically homogeneous reference class for Jones and D . Under the suggestion that single case probabilities are hypothetical limiting frequencies in *broadest* ontically homogeneous reference classes, the event of Jones’s death cannot be probabilistically explained by assigning him to the class of 60 year old human males *with a brain tumor* (etc.) and citing the probability, r , of death in this class, but only by assigning him to the class of 60 year old human males (etc.) that *either* have a brain tumor *or* have that heart disease. The rationale behind taking the *broadest* homogeneous class was to *avoid* including causally *irrelevant* factors in explanation, but this formulation will *prohibit*, in cases such as this, specification of the *distinctively relevant* causally relevant factors.

Thus, suggestion (1) should be rejected because it *requires specification of causally irrelevant factors* in the description of the reference class, and suggestion (2) should be rejected because in some cases it will *prohibit specification of some factors that are distinctively causally relevant* for some single cases. These considerations, involving the conceptual adequacy of the hypothetical limiting frequency interpretation in connection with probabilities of single events, suggest an alternative conception of probability – a revision of the hypothetical limiting frequency view – on which the appropriate reference sequence is characterized not in terms of its subsequences and supersequences (and the relevant attribute and the single case in question), but rather in terms of the *operative causal conditions*, i.e., *the distinctively causally relevant factors*, themselves, where, after all, it was the failure of suggestions (1) and (2) to capture exactly these conditions that rendered them inadequate. Thus, Popper endorses the following alternative to

suggestions (1) and (2):

the frequency theorist is forced to introduce a modification of his theory – apparently a very slight one. He will now say that an admissible sequence of events (a reference sequence, a ‘collective’) must always be a sequence of repeated experiments. Or more generally, he will say that admissible sequences must be either virtual or actual sequences which are *characterized by a set of generating conditions* – by a set of conditions whose repeated realization produces the elements of the sequence. (1959, p. 34)

And Popper explains that, unlike frequency interpretations which take probability to be a property of sequences, somehow appropriately identified (or not), the propensity interpretation takes seriously the idea that an appropriate “sequence in its turn is defined by its set of *generating conditions*; and in such a way that probability may now be said to be *a property of the generating conditions*” (1959, p. 34). And Popper takes one more step. From the premise that actual and virtual frequencies depend on the experimental generating conditions, he concludes that “we have to visualize the conditions as endowed with a tendency, or disposition, or propensity, to produce sequences whose frequencies are equal to the probabilities; which is precisely what the propensity interpretation asserts” (1959, p. 35). As to probabilities of single events,

now we can say that the singular event a possesses a probability $p(a, b)$ owing to the fact that it is an event produced, or selected, in accordance with the generating conditions b , rather than owing to the fact that it is a member of a sequence b . (1959, p. 34)⁶

Thus, evidently, where B^* is an “experimental arrangement”, the propensity theory’s interpretation of ‘ $P(A/B^*) = r$ ’ is (roughly): B^* possesses a *universal* (or “*almost universal*”) *disposition* to produce, if repeated often, sequences B such that the limiting relative frequency of A ’s within B is r . (The reason for the qualification “or ‘almost universal’” is the same as that encountered in the previous section, as discussed below.) Thus, this “long run”⁷ propensity theory invokes just two concepts not present in the hypothetical limiting frequency theory investigated in the previous section: the idea of an *experimental arrangement* and the idea of a *certain kind of disposition of universal* (or “*almost universal*”) *strength* with which some experimental arrangements are endowed. Let us now consider the effect of introducing these two new ideas on the *conceptual adequacy* of the theory and on the possibility of satisfying the condition of *interpretation/idealization*.

As to their effect on conceptual adequacy, Fetzer has compared the

hypothetical limiting frequency theory with the propensity theory in relation to the way in which they may be invoked in accounting for certain frequency patterns that occur in the course of the actual world's history, in a passage worth quoting:

the difference between them is describable as follows: the dispositional interpretation provides a theoretical basis for accounting for these patterns in terms of the *system's initial conditions*, insofar as the occurrence of actual frequencies is explained by reference to the dispositional tendencies that generate them; while [hypothetical limiting] frequency interpretations, by contrast, yield an empirical basis for accounting for these patterns in terms of the *pattern's ultimate configuration*, since the occurrence of actual frequencies is explained by reference to the hypothetical frequencies which control them. Consequently, the kind of explanation provided by a dispositional interpretation for the occurrence of actual frequencies during the course of the world's history is broadly *mechanistic* in character, while the kind of explanation afforded by these frequency constructions for those same occurrences is broadly *teleological* in character. To the extent to which the progress of science has been identified with a transition from teleological to mechanistic explanations, therefore, there even appear to be suitable inductive grounds for preferring the dispositional to the frequency approach. (1981, pp. 77–78)

(Actually, Fetzer is here comparing the *single case* propensity view – which will be considered in the next section – with the hypothetical limiting frequency view, but these considerations apply equally, of course, in the comparison under examination here, as he later points out (p. 107).)

I agree that the “mechanistic” character of the long run propensity view constitutes a conceptual advantage for this view over the “teleological” hypothetical limiting frequency view, especially in connection with the problem of assigning probabilities to single events. For, as we have seen in the rejection of suggestions (1) and (2), above, it seems that an appropriate reference class for a single case probability cannot be characterized in terms just of (i) the membership of the single event in question in the class, (ii) the relevant attribute, and (iii) how hypothetical limiting frequencies change when the class is narrowed or broadened. The long run *propensity* view, on the other hand, specifies that a single case should be referred to the reference sequence of events *which are produced by the same experimental arrangement*, the intent of which is to hold constant just the controlling causal factors present in the single case in question. And this would surely seem to be an appropriate sequence, for what else could be relevant to the physical probability that a single event will exemplify a given attribute than the

physical circumstances under which the single event occurs? It seems clear that if a theory of probability that interprets probability in terms of sequences of events (together with other ideas) is to apply adequately to single events, then it must refer the single events to *such* reference classes, where the relevant physical circumstances under which the single event in question occurs are replicated in every element. So it seems that we should conclude, then, that *as far as conceptual adequacy goes, the long run propensity view is superior to the hypothetical limiting frequency theory.*

But what about satisfaction of the condition of interpretation/idealization, which is supposed to identify the conception in a precise manner and give the interpretation empirical (or cognitive, or, at least, "pre-probability-theoretical") significance? How are the two ideas of dispositions of universal (or "almost universal") strength and of experimental arrangements to be accommodated in a formal solution to the problem of interpretation/idealization?

I have two main points to make in connection with accommodating the idea of dispositions of universal or "almost universal" strength. First, it should be clear that the same sorts of considerations as those advanced in the previous section in connection with the hypothetical limiting frequency theory show that the long run propensity theory must also utilize some idea of "*almost* universal" strength of dispositions, rather than the idea of *strictly* universal strength, if the theory is to be able to accommodate the idea that trials in a sequence of events may be, intuitively, independent of each other. For if the trials are independent, then all sequences of results are "equipossible" and "equiprobable" (assuming that the relevant single case probabilities are supposed to be $\frac{1}{2}$). And even *without* independence and *without* the relevant single case probabilities all being equal to $\frac{1}{2}$, still it seems that *any* sequence of results – and hence *any* limiting relative frequency – should be granted to be *possible*, so that it would be incorrect to explicate probability in terms of a *strictly universal* disposition of an experimental arrangement to display its "characteristic" relative frequency. But we have already seen the serious difficulties involved in one way of trying to characterize the "almost universal" (or "almost every world") idea, and now it seems that these also confront the long run propensity theory in connection with the condition of interpretation/idealization.

Second, it seems that accommodating the dispositional idea (however the "almost universal" idea might be made clear) is closely connected

with accommodating the idea of the experimental arrangement. For, in the first place, it is experimental arrangements that are supposed to be endowed with the relevant dispositions, and, in the second place, not every "arrangement" can be said to possess a disposition of the relevant kind, as will be presently seen. Thus, it seems natural to try to characterize the relevant kind of disposition in terms of the appropriate kind of physical arrangement.

Suppose I have a coin tossing device which has a knob on it controlling a pointer which can be set at any position between 0 and 1, inclusively.⁸ If I set the pointer at position r , then the device will toss coins with a bias, intuitively, of $r:1-r$ in favor of tails, for all r between 0 and 1, inclusively. The internal mechanics of the device are not important. Now clearly, to say just that a coin is about to be tossed by this machine is not enough to specify an experimental arrangement, in a sense appropriate to the long run propensity interpretation of probability. Since such a specification of the arrangement does not include a specification of the setting of the control knob, the arrangement, so specified, does not possess an almost universal disposition to produce any particular limiting frequency of tails.

But now let us change the "initial conditions": the coin tossing device is put together with another device which rotates the pointer slowly back and forth at a constant speed from the 0-position to the 1-position to the 0-position, and so on. Now if the device is constructed in such a way that it tosses coins rapidly at constant short intervals, we can imagine that the combined device has an "almost universal" disposition to produce sequences of tosses with a "characteristic" limiting frequency of tails of $\frac{1}{2}$. But clearly again, even though the arrangement will "almost certainly" yield a "characteristic" limiting frequency, we have not specified an experimental arrangement in a sense appropriate for the long run propensity theory. For this theory is supposed to apply to single events, and, intuitively, when the pointer crosses the $\frac{2}{3}$ -position, the probability of the toss' landing tails is $\frac{2}{3}$, and not the "characteristic" limiting frequency of $\frac{1}{2}$ produced by the device.

Thus, suppose we include in the description of the arrangement a position of the pointer. Have we now succeeded in specifying an experimental arrangement in a sense appropriate for the long run propensity theory? It seems unlikely, even if this specification again yields a "characteristic" limiting frequency. For just as the position of the pointer clearly and overtly indicates a single case or short run bias,

there will no doubt also be *other* factors pertaining to conditions in and around the device which likewise introduce biases: the perhaps random movement of the air around the coin, the humidity, the tidal conditions, small earth tremors, and so on.

Note that for the combined device, the “characteristic” limiting frequency of tails depends on how the rate of rotation of the pointer varies as it sweeps across the dial: in the example above, this rate was assumed to be slow and constant. But if the pointer moved more slowly when it is in the interval $[0, \frac{1}{2}]$ than when it is in the interval $[\frac{1}{2}, 1]$, then the “characteristic” relative frequency of tails would be less than $\frac{1}{2}$. If some proposed specification of the experimental arrangement does not specify the position of the pointer, then we may say that the possible positions of the pointer are unspecified possible initial conditions, and that the “characteristic” limiting frequency of tails depends on both the experimental arrangement, as specified, and the “*distribution of initial conditions*” (as Sklar (1970) expresses the idea). In the example above, this distribution is determined by a function $v(r)$ = the absolute rate of speed of the pointer across the point r on the dial, this assumed to be small and constant for all sweeps across r .

So it seems that there are two options open to the long run propensity in connection with the nature of experimental arrangements: (i) specification of an experimental arrangement must include a specification of the distribution of unspecified initial conditions (as well as a specification of certain of the initial conditions), and (ii) all of the initial conditions must be held fixed. As to (i), three difficult problems arise. First, how is it to be decided *which* of the initial conditions are to be held fixed and which of them should be only partially specified by giving their distribution? Second, on what grounds should one distribution of the unfixed initial conditions to be preferred to another? Sklar has urged that “what this distribution would be [if the experiment were repeated often] is completely unconstrained by any *lawlike* features of the actual world whatsoever!” (1970, p. 363). On the other hand, Settle has reported private communication from Popper in which the latter conjectures that (in Settle’s words) “there is a law of nature, that unless they are constrained, initial conditions have a (‘natural’) propensity to scatter over the interval left open to them by the (constraining) experimental conditions” (Settle 1975, p. 391). Now if the initial conditions *did* have a propensity to scatter over some interval with some characteristic distribution as a matter of law, then perhaps we

would have an important improvement over the hypothetical limiting frequency theory: namely, a principle governing the extension of an actual sequence $R^*(B)$ to a sequence $R(B)$ in a B -maximal extension $M = \langle U, R \rangle$ of a lawlike future world. The principle would state that the initial conditions must be distributed over the elements of $R(B)$ according to how, as a matter of actual law, they must be distributed. But now we must ask whether this propensity to scatter over an interval with some "characteristic" distribution is a universal or an "almost universal" disposition, and if it turns out to be an "almost universal" disposition, then by now familiar problems emerge again. In any case, whether the disposition must be universal or only "almost universal" would seem to depend on a more precise formulation of the conjecture (e.g., are the different configurations of initial conditions *independent* of each other?), and perhaps on empirical investigation.

But, perhaps more importantly, the third difficulty with (i) pertains to the desire to make probability applicable to single events. If the exact configuration of initial conditions in a given trial (i.e., the actually obtaining values of the "hidden" variables) makes a physical difference with respect to the result of the trial, then it would seem inappropriate to leave any of the initial conditions unspecified.

This suggests consideration of alternative (ii). But the problem with suggestion (ii) is that it may very well be, as a matter of fact, a matter of physical law that in some cases, configurations of initial conditions cannot remain fixed from trial to trial. So if we have to consider a virtual sequence $R(B)$ in which all of the initial conditions remain unchanged from element to element, we may have to consider *nonlawlike* B -maximal extensions of lawlike future worlds. Also, if conditions such as being the n th element of $R(B)$ are to count as initial conditions, then there would also be logical difficulties with the idea of replicating an experiment, holding all of the initial conditions fixed. The single case propensity interpretation, to be considered in the next section, has, I think, a more plausible suggestion to offer along basically the same lines. So I shall postpone consideration of the "hold everything fixed" idea until then.

But perhaps it will be urged that we have been going about the explication of the two new concepts of the long run propensity interpretation in the wrong direction: instead of trying to characterize the ("almost") universal disposition idea in terms of the experimental arrangement, we should first try to characterize the disposition, and

then, in terms of this, characterize the relevant kind of experimental arrangement. Of course a particular, individual experiment cannot itself be repeated, literally speaking, so what is needed, of course, is a characterization of the relevant kind of *experiment type*. And according to this new approach, two particular experiments will be of the same type (of the appropriate kind of type) if they are *both endowed with the same universal or almost universal disposition* (of the appropriate kind). If this idea could be worked out satisfactorily – in connection with formal adequacy and interpretation/idealization – then, again, the long run propensity theory would have important advantages over the hypothetical limiting frequency theory as considered in the previous section. For then, the long run propensity theory would be in possession of a principle governing the extension of actual sequences $R^*(B)$ to infinite virtual, or hypothetical, sequences $R(B)$ in B -maximal extensions of lawful future worlds: we may say that an *admissible* such sequence for the purpose of assessing a probability $P(A/B)$ would be one whose every member was endowed with the very same disposition (of the appropriate kind) with which every member of $R^*(B)$ – which may consist of just one trial – is endowed. Thus, perhaps the characteristic structures of the long run propensity interpretation would be like those of the hypothetical limiting frequency interpretation except for having an additional component: say a function D from pairs $\langle M, x \rangle$ into properties of events, where, of course, we should not insist that $D(M, x) \in R$, where $M = \langle U, R \rangle$. Then, in the intended model, for any world M and event x , $D(M, x)$ is the universal or almost universal disposition of the relevant kind with which $M(x)$ is endowed, and an *admissible* extension of a sequence $R'(B)$ consisting, say, of just one event $M'(x)$ – where $M' = \langle U', R' \rangle$ is a lawful future world – would be a sequence $R(B)$ in a world $M = \langle U, R \rangle$, every member of which has the property $D(M, x) = D(M', x) = D(M^*, x)$. This, plausibly, might accommodate also the conceptual difficulty with the hypothetical limiting frequency view that, intuitively, lawful future worlds, and their B -maximal extensions, may have *statistical* laws differing from those that hold in the actual world, though, by the definition of lawful worlds, the *universal* laws that hold of them are the same as those that hold in the actual world.

Of course the above considerations only constitute a step in the direction of *conceptual* and *formal* adequacy of the long run propensity interpretation, leaving the problem of *interpretation/idealization* un-

touched. A solution to the latter problem is supposed to identify the intended model and associate independently understood concepts, objects or phenomena with the constituents of the intended model. Of course counterparts of the problems encountered earlier in connection with the hypothetical limiting frequency theory for interpretation/idealization remain (e.g., what does “almost every world” mean?), but a new part of the problem for the long run propensity view is to interpret the new function symbol ‘*D*’ – in other words, to explicate the relevant kind of “almost” universal disposition to produce sequences with a characteristic limiting frequency. From an antagonistic point of view, this property may be characterized as whatever it is that every member of hypothetical infinite sequence must have in order for the limiting frequency of the relevant attribute in that sequence to be appropriately transferable to any member of the sequence. From the other point of view, the postulation of the existence of this thing has been characterized as “a new physical hypothesis (or perhaps a metaphysical hypothesis) analogous to the hypothesis of Newtonian forces” (Popper 1959, p. 38). And, indeed, what seems to be lacking in the long run propensity interpretation is an explication of ‘*D*’ in terms of *old* concepts, objects or phenomena that are *already* understood *independently of probabilities or propensities*. Below I shall consider the question of whether propensity theories should be required to satisfy a condition of interpretation/idealization (for ‘*D*’), in light of the idea that such theories may be characterized as involving, after all, a *new* physical hypothesis to the effect that there are propensities, of some sort, which may be of a “new metaphysical category”. But, independently of the appropriateness of the condition, it seems appropriate to conclude that, while for reasons given several pages back, the long run propensity interpretation is *superior to the hypothetical limiting frequency theory as far as conceptual adequacy goes*, it is, for lack of an appropriate interpretation of ‘*D*’, *inferior with respect to interpretation/idealization*.

6.

There is a formulation of the single case propensity theory of probability that initially seems to have the advantage over the long run approach, in connection with interpretation/idealization, that the conception of probability offered can be explicated in terms understood

independently of the ideas of universal or "almost universal" dispositions – i.e., independently of the intended interpretation of the function symbol 'D' of the intended model of the long run account, as described at the end of the previous section. The single case theory which I shall consider below is, essentially, that of Fetzer and Nute (1979, 1980; see also Fetzer 1981, pp. 49–73).

The basic idea is that instead of looking at relative frequencies in sequences of *repetitions of experiments in single worlds* (and then perhaps at the relative frequency of worlds' exhibiting a given frequency), we look first at *sequences of lawful future worlds in which the single event in question takes place*, exhibiting or not exhibiting the relevant attribute. Thus, suppose we wish to give truth conditions for the statement ' $P(A/x) = r$ ', where x is the single event in question, say a toss of a coin, where A is the relevant attribute, say coming up tails, and where the statement means, intuitively, that the single case probability (propensity) of that toss' resulting in tails up is r . Let M_1, M_2, \dots be an infinite sequence of lawful future worlds, i.e., a sequence of worlds each of which obeys all the universal laws which the actual world M^* obeys and whose histories are the same as that of M^* at least up to the time of the event x . Then we may give truth conditions as follows:

$$\begin{array}{l} \text{'P(A/x) = r' is true in } M^* \text{ if and only if} \\ \lim_{k \rightarrow \infty} \frac{\# \{M_i : i \leq k \text{ and 'Ax' is true in } M_i\}}{k} = r. \end{array}$$

These truth conditions have the advantage over both the hypothetical limiting frequency interpretation and the long run propensity interpretation that they avoid the necessity of providing principles governing the extension of actual sequences of events to longer, ideally infinite, sequences of events that occur in some possible world. In particular, no recourse to the *D*-component of the long run propensity theory is necessary. Also, of course, there is no longer the problem of the reference class or experiment type, and, envisioning x as being, in every relevant world, numerically the very *same* event, we have, here, an interpretation that is truly applicable to the single case.

But there are difficulties with this approach that are similar to the problems encountered in connection with the hypothetical limiting frequency theory. Since there are, presumably, *nondenumerably* many lawful future worlds relevant to the probability statement in question (as argued in section 4), the problem arises of how to select from all of

these worlds an appropriate *denumerable* sequence, M_1, M_2, \dots , in terms of which the truth conditions for ' $P(A/x) = r$ ' should be given. Also, as long as Ax and $\bar{A}x$ are both physically possible, it would seem that there should be infinitely many lawful future Ax -worlds as well as infinitely many such $\bar{A}x$ -worlds, so that, for any value of r between 0 and 1, inclusive, there will be *some* sequence of worlds which, together with the suggested truth conditions, will yield the truth of ' $P(A/x) = r$ '. Thus again, Fetzer and Nute require only that "*almost every*" sequence of lawful future worlds satisfy the above truth conditions in order for ' $P(A/x) = r$ ' to be true, where this is made precise as follows. Let M^1, M^2, \dots be an infinite sequence of infinite sequences of lawful future worlds, where M^j_i is the i th member of the j th sequence. Then revised truth conditions are suggested as follows:

' $P(A/x) = r$ ' is true in M^* if and only if

$$\lim_{m \rightarrow \infty} \frac{\# \left\{ M^j : j \leq m \text{ and } \lim_{k \rightarrow \infty} \frac{\# \{ M^j_i : i \leq k \text{ and 'Ax' is true in } M^j_i \}}{k} = r \right\}}{m} = 1.$$

But the same sort of problem arises again. As noted above, if there are infinitely many Ax -worlds that are lawful and future, as well as infinitely many such $\bar{A}x$ -worlds, then there is, for any value of r , some sequence of worlds which, on the first suggested truth conditions, yields the truth of ' $P(A/x) = r$ '. But if there is even one such sequence, there are infinitely many such sequences: simply reorder the first n terms, for each finite n , to get infinitely many such sequences. Arrange these sequences into a sequence of them, in any order, and you get a sequence of sequences of lawful future worlds which, together with the second suggested truth conditions, yields the truth of ' $P(A/x) = r$ '. And this can be done for any value of r , as long as there are infinitely many lawful future Ax -worlds as well as infinitely many such $\bar{A}x$ -worlds. Thus, the problem is to specify principles for selecting an appropriate sequence of sequences of worlds from presumably nondenumerably many such sequences. We may take Fetzer and Nute's idea one step further and say that ' $P(A/x) = r$ ' is true if and only if "*almost every*" sequence of sequences of lawful future worlds satisfies the second suggestion, where *this* could be made precise in terms of an infinite sequence of infinite sequences of infinite sequences of lawful future worlds, but it is obvious that the same sort of problem would arise again. And so on.

Perhaps there may be another way of rigorously capturing the intuitive idea that the proportion of Ax -worlds in a random selection from all lawful future worlds will almost certainly be about r . Indeed, the difficulties with the above approach may suggest a measure theoretic approach, according to which ' $P(A/x) = r$ ' is supposed to be true if and only if $r = \sum_M \text{Probability}(M)$, where the summation is taken over lawful future Ax -worlds M , or, more generally, $P(A/x) = \int_{|Ax|} \mu(M) d\mu(M)$, where, again, the integral is over lawful future worlds in $|Ax|$. But now it is natural to ask where the probability function *Probability* comes from, or where the density function μ on worlds comes from.

Suppes (1973) has formulated a measure theoretical approach in which a probability function *Probability* can be inferred from an axiomatized quaternary relation $A/B \geq C/D$ (meaning that the propensity of A 's occurring given an occurrence of B is at least as great as the propensity of C 's occurring given an occurrence of D). But then the problem remains of giving a *physical interpretation* of the relation \geq , in the sense elaborated earlier in this paper. Note that this is different from the problem of *application* in particular physical contexts, whose solution may simply require an association of kinds of physical events with A , B , etc., and the assumption of additional axioms appropriate to the particular physical context. Suppes (1973) gives an example of this, relating to the phenomenon of radioactive decay.

Also, Giere (1976) formulates a measure theoretical kind of approach, in which finite, countable and continuous *possibility structures* are defined. These involve, basically, a set of possible worlds, a partition of this set, and (i) in the case of finite possibility structures, an equal measure over the possible worlds, (ii) in the case of countable possibility structures, an "equal measure" on the possible worlds obtained from a one-to-one correlation between the possible worlds and the interval $[0, 1]$ and a uniform density on $[0, 1]$, and (iii) in the case of continuous possibility structures, something more complicated. In each case, a probability space on the partition of worlds can be inferred, where this is used to give a formal definition of a "propensity function" on the set of final states of a stochastic system, where this set is an isomorph of the partition of worlds. In each case, the basic idea is, as Giere puts it, that "physical probabilities are a measure of the density of possibilities open to a system in a given initial state" (p. 338), where these possibilities correspond, in the formal theory, to the "equipossible worlds" (p.

338). But problems are: How to characterize possible worlds in such a way that they *are* “equipossible” (Bertrand paradoxes)? What does “equipossibility” mean? And how to come up with an appropriate “density of possibilities” (density function) in any given case? Now again, I do not think that all of these problems will be serious in all contexts of *application*, as it is clear from the examples that Giere gives that his formulation can be accommodated to many different kinds of stochastic phenomena, including radioactive decay and Bernoullian sequences. But this just displays the *formal adequacy* of the theory as applied to various kinds of phenomena, where the question under consideration is: How are the entities of the abstract possibility structures to be related to something independently understood in order that we may have an explication of propensity *via* interpretation/idealization?

In connection with his formulation of the propensity theory in the case of finite possibility structures, Giere himself states (where it is clear that these comments apply also to the countable and continuous cases),

The individual worlds and the uniform measure need have no direct physical correlates. That is the respect in which this semantics is merely a formal semantics. It would be an interesting physical hypothesis that underlying every stochastic process there exists a set of physically equipossible states. The above account of physical propensities is compatible with this hypothesis but does not require it. In any case, I doubt it is true.

Metaphysically speaking, then, what are physical propensities? They are weights over physically possible final states of stochastic trials – weights that generate a probability distribution over sets of states. The [uniform measure] function *u* provides only a formal and rather shallow analysis of this distribution. But is it not just the task of a propensity interpretation to explain what these weights are? No, because it cannot be done. We are faced with a new metaphysical category. (1976, p. 332)

And he suggests that what he has in mind may be the same point as Popper expressed by saying that his propensity theory involved a “*new physical hypothesis* (or perhaps metaphysical hypothesis) analogous to the hypothesis of Newtonian forces” (1959, p. 38).

This feature of single case propensity theories is also found in Fetzer’s theory, and he takes the idea a step further with a dispositional ontology according to which, for example, “*individual objects* are continuous sequences of instantiations of particular arrangements of [universal and statistical] dispositions” (1981, p. 41) and “*singular events* are continuous sequences of instantiations of particular arrangements of [universal and statistical] dispositions” (1981, p. 42). Here, dispositions are ontologically primitive, but

a dispositional predicate . . . may be informally defined as a set of ordered triples, each of which consists of a test trial description T^i , an outcome response description O^i , and a numerical strength specification r^k , i.e., $\{\langle T^1, O^1, r^1 \rangle, \langle T^2, O^2, r^2 \rangle, \dots\}$, where the number of members of these sets is determined by the variety of different trial tests and response outcomes that are ontological constituents of each specific disposition – a possibly infinite set. (1981, p. 37)

Thus it is clear that Fetzer, too, proposes the existence of propensities as “a new physical hypothesis”, where the single case propensity semantics with which this section began, involving infinite sequences of infinite sequences (and so on?) of trials under identical conditions are “required to display the complex character of propensity attributions” (private communication), but are not *definitive* of single case propensities. Whatever the detailed relation is between the propensities themselves and the semantics proposed, it is clear that for Fetzer as well, single case propensities are not *definable* in terms of independently understood concepts and phenomena.

That single case propensities are not definable in terms of independently understood concepts and phenomena (with the exception, in some cases perhaps, of a qualitative propensity relation) seems inevitable, if they belong to a “new metaphysical category”. Nevertheless, it might be hoped that empirical significance could be given to the concept of single case propensities by describing procedures whereby hypotheses involving the concept may be tested, hypotheses such as, “The single case propensity of *this* nucleus’ decaying before the end of 30 years is $\frac{1}{2}$ ”; and perhaps then, in terms of procedures for testing *such* hypotheses, we may be in a position to test the physical hypothesis that single case propensities obey some axiomatization of probability.

But the following problem naturally arises in connection with the idea of testing single case propensity statements: since, in any given single case x , the relevant attribute, say A , either occurs or fails to occur, the difference between, for example, $P(A/x) = 0.7$ and $P(A/x) = 0.8$ would seem to make no difference in experience. (See Reichenbach (1949, pp. 370–71) for another statement of this argument.) Despite the unrepeatability in principle of particular single events, however, a number of propensitists have suggested that single case propensity statements may be tested by reference to certain relative frequencies. Thus Fetzer says,

although a single case probability statement surely does pertain to singular events, it is somewhat misleading to suppose that the events involved only occur a single time. A

single case propensity statement, after all, pertains to every event of a certain kind, namely: to all those events characterized by a particular set of conditions. Such a statement asserts of all such events that they possess a specified propensity to generate a certain kind of outcome. Consequently, although the interpretation itself refers to a property of conditions possessed by single events, such statements may actually be tested by reference to the class of all such singular occurrences – which surely is not restricted to a single such event. (1971, p. 478)

But this would seem to raise the old problem of the reference class, or of the description of the experimental arrangement: *what is* the “certain kind” of event to which we should think of the single case in question as belonging, or *what is* the relevant “particular set of conditions”? Note that here, the reference class problem (or experiment type problem) arises in connection with the problem of *testing* probability statements, while for the relative frequency view and (as presented above) for the long run propensity view, this was a problem for the very explication of probability. And I would agree that a problem of this kind need not be as serious for a theory of probability if it arises *only* in connection with its theory of testing and *not* for the very explication of probability. But recall that we have decided to look for a kind of explication of the idea of single case propensities (for empirical interpretation) in empirical methods of testing propensity statements. So perhaps the problem is serious, as the following considerations may further suggest.

Now Fetzer has given an answer to the question posed above just after the quotation. He advances a requirement of maximal specificity according to which, roughly adapted to the context of this discussion, the “particular set of conditions” must include all “nomically relevant” factors, where a factor is nomically relevant, roughly, if its presence or absence in a given single case event would affect the single case propensity of that event’s having the relevant attribute (1981, pp. 50–51).⁹ Thus, it would seem that in order to test a single case propensity statement by looking at a relative frequency generated by an appropriate “particular set of conditions”, we must either already know or have good reason to believe or also test *other* single case propensity statements, namely, statements to the effect that (i) some particular set of conditions is present in every trial and (ii) that set of conditions is appropriate in the sense given above (all nomically relevant factors are held fixed in every trial) so that we may know that (iii) the single case propensity for the relevant attribute is the same in each trial. (Cf. Fetzer 1981, pp. 248–54.)

Giere has made the same point about testing single case propensity statements, in connection with tests of hypotheses about the half-lives of radioactive nuclei. He describes the standard procedure, which "assumes that each nucleus in the sample has the *same* half-life, whatever its value. Thus the test assumes the truth of some propensity statements, though not of course the truth of the hypothesis being tested" (1973, p. 478). But he goes on to argue that this feature of testing single case propensity hypotheses is not unique in science:

Consider the concept of an individual (as opposed to total) force in classical physics. Any attempt to determine the value of a particular force requires assumptions concerning other forces, e.g., that there are none operating or that their influence has been taken into account. Thus, if one regards the concept of an individual force as a legitimate empirical concept, one cannot dismiss single-case propensities solely on the ground that empirical tests of propensity hypotheses assume the truth of other propensity statements. (p. 479)

But it seems to me that the two cases are not parallel and that one can measure individual forces under significantly weaker assumptions than one can measure single case propensities. Consider the single case propensity hypothesis, "The probability that *this* nucleus will decay within 30 years is $\frac{1}{2}$." As Giere explains, the standard procedure for testing such a hypothesis is to obtain a large number of nuclei of the same kind and then count the numbers of them that decay within specified periods of time. The single case propensity hypothesis that such a test assumes to be true is that all the nuclei in the sample have the *same* half-life, whatever its value. How does one test a hypothesis concerning an individual force, say the force exerted by a certain spring when its length is two inches? One may first measure the total force present in the absence of the spring (say by observing the acceleration of some object) and then measure the total force present when the spring is introduced (say by observing the acceleration of an object when placed at the end of the spring in its two-inch configuration), and then calculate the difference between the two values. Presumably, the hypothesis concerning forces that Giere would say is assumed in such a test is that, when the spring is introduced, all of the *other* individual forces remain the same.

But note that the case of individual forces is quite different from that of single case propensities in that the individual forces can each be *separately measured* by *separate* tests which do not require the *same* assumptions in each case. In principle at least, each individual force can be eliminated and the total remaining force can be compared to the

original force to yield a measure of that force. In each case, of course, one assumes that all the remaining forces do not change, but since for each individual force, the remaining individual forces will constitute a *different* set of individual forces, the assumptions made in the different cases will not all be the same. Also, the law of addition of forces can in principle be tested by determining whether or not the sum of all the values determined for the individual forces add up to the value determined for the total force. (In practice, of course, this can only be done in the context of a background configuration of forces that cannot in practice be eliminated, but of course one can obtain additional confirmation for a hypothesis concerning an individual force by looking at the effect of introducing the relevant conditions in many different constellations of background forces.) In the case of the single case propensities of the individual radioactive nuclei decaying within 30 years, one cannot test each nucleus (of the same kind) separately: for each nucleus, the test of the relevant hypothesis is the very same test. Another difference is that in the case of individual forces, no assumption whatsoever must be made about the value of the individual force in question, whereas in the case of single case propensities, one must assume that the nucleus in question has the same propensity to decay within 30 years as do all the other nuclei in the sample.

Now it may be insisted that the analogy which Giere urges still holds, for in each case, it is still necessary to make *some* assumptions concerning the relevant kind of individual thing (propensity or force). But in the case of force, the method of testing itself makes it clear that *it is an individual kind of thing that is being measured: individual* forces, indirectly by measuring *different* pairs of *different* total forces and calculating differences. But in the case of testing single case propensity hypotheses, the method of testing does *not* give empirical significance to the idea that it is an *individual* or *single case* propensity – rather than, for example, a *long run* propensity – that is being measured. There is nothing in the general sketch of the method of *testing* single case propensity hypotheses under consideration which distinguishes the single case propensity interpretation from the long run propensity view.

Here is another aspect of this kind of difficulty in testing distinctively *single case* propensity hypotheses. That single case propensities exist is a new physical hypothesis. And until and unless this hypothesis is developed in more detail to the contrary, it would seem that two objects or sets of experimental conditions may differ in no respects whatsoever

except for a certain single case propensity with which they are endowed (and, of course, the physical consequences, e.g., pertaining to relative frequencies generated, of having the different single case propensities). On single case propensitist principles, it would seem conceivable, for example, that there is a certain "kind" of radioactive nucleus whose half-life has been tested extensively, where the *only* difference between individual nuclei of this kind is that they can actually have *different* ("single case") half-lives ("statistical hidden variables"), where the half-lives of these nuclei (which vary considerably) are distributed among such individual nuclei in such a way that testing random samples of them in the standard way has always yielded the same result as would be expected if they all had the *same* ("single case") half-life (which, say, scientists have been assuming): the different half-lives are distributed homogeneously among the nuclei of the relevant "kind". Then how to test the *single case* propensity hypotheses pertaining to the different nuclei? And how to test whether all the nuclei have the *same* half-life, or *different* half-lives so distributed among the nuclei that relative frequency tests give the same results as they would if all the nuclei had the same half-life? Now these considerations may only be valid in the absence of a well-developed theoretical background, here pertaining to how the structure of a nucleus (which may determine a *kind* of nucleus) is related to its half-life. But if Giere is correct in suggesting that "in the absence of a well-developed theoretical background, observed relative frequencies may provide the only evidence for propensity statements" (1973, p. 478), then examples such as this one strongly suggest that the single case and long run propensity conceptions of probability cannot be distinguished in terms of empirical methods of testing the relevant propensity hypotheses.

At the beginning of this section, we considered a formulation of the single case propensity theory of probability that initially seemed to be superior to the long run propensity theory with respect to interpretation/idealization. But problems for that formulation arose, problems of the same general kind as arose earlier for the hypothetical limiting frequency interpretation and which also confront the long run propensity theory. Recall that the long run propensity theory's solution to the problem of actual sequence extension was – in terms of the formulation I suggested in section 5 – to introduce a new two-place function *D*, which, in the intended model, has as its range of values a set of dispositional properties of a certain kind. While the introduction of *D*

rendered the long run propensity view superior to the hypothetical limiting frequency theory with respect to the condition of conceptual adequacy, it also rendered it inferior with respect to interpretation/idealization. Now it seems that the single case propensity view must do something very similar: it too must introduce a two-place function, say D^* , whose range, in the intended model, would be a set of dispositions, this time "single case statistical dispositions", rather than universal (or "almost universal") dispositions to produce statistical displays in the form of characteristic relative frequencies. Thus, temporarily leaving aside the question of conceptual adequacy and of the relevance of the condition of interpretation/idealization to propensity theories, it seems that the single case interpretation is *at least* as bad off as far as interpretation/idealization goes as is the long run view.

Just above, a third possible way of securing interpretation/idealization was considered – where the first two ways, of course, were *via* possible worlds and sequences of various kinds and *via* all this plus D , or D^* . The third was in terms of empirical procedures for testing hypotheses of the relevant kind, which, if successful, should help the theory to obtain at least empirical *interpretation*. But it seems that this approach cannot distinguish between the long run and single case dispositional approaches: both approaches would use the available finite relative frequencies in the very same way in tests of the relevant hypotheses. Roughly and intuitively speaking, the two theories *idealize* the relevant phenomena in different ways (one in terms of D and the other in terms of D^*), where the objects of *interpretation* of the two idealizations are the same: observable finite relative frequencies. Still leaving aside the question of conceptual adequacy and of the relevance of the condition of interpretation/idealization to propensity interpretations of probability, we can ask which of the two theories under consideration is better off in connection with the interpretation/idealization condition by asking: Which of the two idealizations is more fully interpreted in terms of the available relative frequencies? That is, we ask, intuitively: Which theory commits itself to the stronger concept (idealization), the concept a higher proportion of the content of which will therefore *lack* interpretation in terms of the common objects of interpretation? And it seems clear that it is the single case conception which is the stronger concept. The long run concept is, roughly, a two-component concept: the concept of universal (or "almost universal") dispositions *plus* the idea of limiting frequencies, where the

second component of the concept is quite well understood. The single case concept, however, is a kind of *one-component* “organic” union of the two components of the long run concept: the concept of a *partial or statistical disposition of a specific strength*, rather than a *universal* (or “*almost universal*” disposition to produce a statistical *display*. And in this case it seems that the whole (the concept of a *statistical disposition*) is greater, or stronger, than the sum of its parts (the concept of a *universal or “almost universal” disposition* plus the concept of a *display which is of a statistical character*, i.e., a sequence with a characteristic relative frequency). The single case propensity concept compresses the idea of a display of a statistical character into the concept of the disposition itself.

The question for interpretation/idealization is, here, how adequate the interpretation is relative to the concept, or intended idealization – that is, in this case, the extent to which observed relative frequencies capture the features or components of the idealization, or proposed concept. For the long run theory, what we always (or “almost always”) observe (or, at least what we can in principle observe) is the *direct manifestation* of the relevant disposition: the disposition is a disposition to produce sequences of events with a characteristic limiting frequency, and we can, in principle, observe a sequence with a characteristic frequency of the relevant attribute. What we don’t actually observe, of course, is the disposition itself. Thus, what we can in principle observe is the physical interpretation of one component of the two-component long run concept, and this is the direct manifestation of the relevant disposition. For the single case propensity view, on the other hand, what we actually can observe (relative frequencies in sequences) is *not* the direct manifestation of the relevant disposition, it seems. For the statistical disposition is supposed to operate directly on the single case, and *via its direct operation on single cases* it controls the observed relative frequencies in accordance with Bernoulli’s theorem. Thus, for the single case propensity interpretation, there are *two things which we do not observe when we observe relative frequencies: namely, the direct manifestation of the disposition, as well as, of course, the disposition itself*.¹⁰ Since (i) *observed relative frequencies* (that which it seems that both theories must use to secure, empirically, interpretation/idealization) are the *direct manifestations* of the relevant disposition on the long run view and also the *physical interpretation* of the other component of the two-component long run concept, and since

(ii) they are *not* the direct manifestations of the single case disposition, I conclude that, *in relation to the condition of interpretation/idealization, the long run propensity view fares better than the single case interpretation.*

Let us now turn to the comparison between the single case and long run propensity interpretations with respect to the condition of conceptual adequacy. We have already seen that both theories are, as Fetzer puts it, "broadly mechanistic" in character rather than "broadly teleological", like actual and hypothetical limiting frequency theories. But what of some of the other desiderata that should be brought to bear? Consider the problem of attributing probabilities to single cases, say the probability that event x will exemplify attribute A . Suppose that, as the long run propensitist requires, there really is a dispositional property $D(M^*, x)$; and suppose that, as the single case propensitist requires, there really is a single case propensity $D^*(M^*, x)$ for x itself to exhibit attribute A . (What these suppositions amount to, as far as a detailed *explication* of the concepts is concerned, is a problem for interpretation/idealization; here, we are interested only in the theoretical consequences of what is *intended* by advancing the concepts.) Now the possession of the property $D(M^*, x)$ by each member of a hypothetical infinite sequence of events in a lawlike future world is supposed to guarantee (or "almost guarantee") that the limiting relative frequency of A in the sequence is, say, r . But, as single case propensitists emphasize, what holds in the long run does not always matter in the single case (Hacking 1965, p. 50; see also Fetzer 1981, pp. 110–11).

Although Hacking's example (see the reference) pertains to *rational decision making*, the following considerations are intended to show that, as far as *physical probabilities* are concerned, it is also true that what matters in the single case need not matter in the long run (on the conceptions of single case and long run propensities under consideration). The possession of $D(M^*, x)$ by every member of a hypothetical infinite sequence of events in a lawful future world would seem to be compatible with (at least, is *not obviously incompatible* with) the members' possessing single case propensities for exhibiting A that differ from case to case, and differ from r . If both $D(M^*, x)$ and single case propensities exist, then all the possession of $D(M^*, x)$ by the members of a hypothetical infinite sequence has to guarantee (or "almost guarantee") in order to guarantee (or "almost guarantee") that the limiting relative frequency of A is r , is merely that the *average* of the single case propensities for A in the sequence is r , where this idea of

an “average” can be made precise in the obvious way. And indeed, on Popper’s conception, it *would* seem possible for possession of $D(M^*, x)$ by every member of a hypothetical infinite sequence in a lawful future world to guarantee *merely* that the *average* of the single case propensities is r , given his conjecture about configurations of initial conditions distributing themselves over the interval left open to them *as a matter of physical law*, and, hence, in lawlike future worlds (see section 5 above on this conjecture, and Settle 1975, p. 391). And then it would seem that even though *different* configurations of initial conditions would give rise to *different single case propensities*, still the *lawlike distribution* of the configurations of initial conditions over the interval left open to them would guarantee (or “almost guarantee”) the *characteristic limiting frequency*. Thus, what matters in the single case need not matter in the long run, assuming the truth of Popper’s conjecture. Until the concept of $D(M^*, x)$ is refined in such a way that it can be shown that possession of $D(M^*, x)$ by the relevant events cannot guarantee merely an average single case propensity for the relevant attribute among the single events in question – where the refinement does not make possession of $D(M^*, x)$ conceptually equivalent to the possession of a single case propensity – it would seem that as far as attributing probabilities to single events is concerned, the single case theory is conceptually superior to the long run interpretation.

Having considered the single case adequacy of the long run propensity approach, what now about the long run adequacy of the single case approach? According to Fetzer,

The most important benefit of the “single case” approach . . . is that it not only accounts for the meaning of *single case* probabilities but also solves the problem of *long run* probabilities; for, given the values of the relevant single case probabilities, calculations of long run probabilities for the various combinations of outcomes over various lengths of trials may be made on the basis of the mathematical principles [such as Bernoulli’s theorem] for statistical probabilities. Thus, the fundamental advantage of the single case interpretation is that it yields a construct which is theoretically significant for *both* the long run *and* the single case . . . (1981, p. 111)

Aside from whatever may be said in favor of the idea that some probabilistic phenomena are not “grounded from below” in terms of probabilistic laws on the level of individuals, but are rather “imposed from above” (see Hacking 1980, and Baird and Otte 1982 on this), this, of course, is correct; where, however, if some probabilistic phenomena actually were “imposed from above”, then perhaps in such cases a long

run approach would be more appropriate. But, aside from such worries, and given the general promise of Poisson's law of large numbers program of grounding probabilities from below, and given the theoretical difficulties of the long run conception in the single case, it seems appropriate to conclude that – though the long run propensity view may be superior to the single case theory in connection with interpretation/idealization – as far as conceptual adequacy is concerned, the single case propensity interpretation fares better than the long run theory.

7

I have argued that to the extent to which philosophical theories of objective probability have offered theoretically adequate *conceptions* of objective probability, in connection with such desiderata as causal and explanatory significance, applicability to single cases, etc., these theories have themselves failed to satisfy the *methodological* standard of interpretation/idealization, the requirement, roughly, that the conception offered be specified with the precision appropriate for a physical interpretation of an abstract formal calculus, and be fully interpreted in terms of concepts, objects or phenomena understood independently of probabilistic concepts. This *may* be grounds for scepticism about objective probability. On the other hand, perhaps we should take seriously the idea that propensity theories are, in part, proposals of a *new metaphysical or physical hypothesis* and that, therefore, we should not expect propensities to be explicable, in the way the condition of interpretation/idealization demands, in terms of *old*, or independently understood, concepts, objects or phenomena. Perhaps, in view of the idea that propensities are supposed to be entities of a “*new metaphysical category*”, it is inappropriate to foist the interpretation/idealization requirement on theories of propensity, since the requirement insists on explication of the proposed conception in terms of *old* ideas.

Indeed, in view of the foregoing discussion, it seems to me that the only way in which propensity theories can secure something like interpretation/idealization is *through* their conceptual adequacy, where the objects of interpretation, then, are such things as: theoretically adequate explanations of single events, and of physical regularities; causal laws; events and objects themselves; etc. (as is implicitly

suggested by Fetzer 1981, pp. 295–96). Thus, if these things can be *identified* (by which I do not mean fully understood) prior to an understanding of propensities, and if a propensity theory of probability can characterize a role that a certain concept (i.e., the propensity concept) plays in these things, it will thereby have established something like “bridge principles” connecting the theoretical concept of propensity with the independently identifiable things listed above, thereby also giving an implicit or partial definition of the theoretical concept. It seems to me that it can only be in terms of satisfaction of a weaker kind of condition of interpretation/idealization, formulated in the light of these ideas, that the propensity concept can be identified, where whether or not such a mode of identification would be entirely adequate is not entirely clear.

NOTES

* This paper was written, in part, under a grant from the Graduate School of the University of Wisconsin-Madison, which I gratefully acknowledge. I would also like to thank James H. Fetzer for many useful suggestions which improved an earlier draft, and Michael Byrd for useful discussions.

¹ On Jeffrey's (1965) theory, only a *family* of pairs of probability and desirability functions is determined by a coherent set of preference data, where neither function is uniquely determined. On other theories, the subjective probability function is determined uniquely, but the desirability function is not.

² Of course this does not imply that the set of values of a sequence-function is infinite. Also note that the idea of limiting frequency, defined below, applies in the case in which the reference class is finite: “Notice that a limit exists even when only a finite number of elements x_i belong to [the reference class B]; the value of the frequency for the last element is then regarded as the limit. This trivial case is included in the interpretation and does not create any difficulty in the fulfillment of . . . the . . . axioms” (Reichenbach 1949, p. 72).

³ Limiting frequencies aren't in general countably additive. See Van Fraassen (1979) on this and on the idea of limiting frequencies being defined on Boolean algebras.

⁴ For example, in Cartwright (1979) and Skyrms (1980); for discussion of these and other such theories, and further references, see Eells and Sober (1983).

⁵ My notation here differs from Kyburg's. Also, here, as in the sequel, the terms ‘ A ’, ‘ B ’, ‘ x ’, etc., are just that: terms of the relevant first order language. Sometimes, however, when it is clear what the relevant model M is, I shall use just ‘ A ’, ‘ B ’, ‘ x ’, etc., as names for what they denote *in* M ; at other times, I shall write ‘ $R(A)$ ’, ‘ $R(B)$ ’, etc., for the class or sequence which $M = \langle U, R \rangle$ assigns to ‘ A ’, ‘ B ’, etc., and $M(x)$ for what M assigns to an individual term x .

⁶ It seems that ‘ $p(a, b)$ ’ should not be read as ‘the probability of the singular event a

(*happening*)...’, but as ‘the probability of a certain event’s having the relevant attribute...’.

⁷ Some philosophers have said that there is an ambiguity in Popper’s writings in connection with whether his propensity theory is supposed to be a “long run” interpretation or a “single case” interpretation. In any case, in this section, I shall be considering the long run construal; in section 6, I consider a single case propensity approach.

⁸ I owe the idea of this machine to Harry Nieves, who invented it to make a somewhat different point.

⁹ Fetzer actually states the requirement in terms of *predicates* (rather than of “factors”) and of reference class *descriptions* (rather than of the classes themselves).

¹⁰ Of course one might say (as Fetzer has urged in private communication) that we actually do observe direct manifestations of single case propensities in each single event: namely, the occurrence or nonoccurrence of the relevant attribute. But, of course, such single case displays are inappropriate for the purpose at hand – namely, securing empirical interpretation for the *statistical* concept – since such single occurrences or nonoccurrences of the relevant attribute are, separately, completely uninformative in relation to the value of the single case propensity in question.

REFERENCES

- Baird, D. and R. E. Otte: 1982, ‘How to Commit the Gambler’s Fallacy and Get Away With It’, in P. D. Asquith and T. Nickles (eds.), *PSA 1982*, Philosophy of Science Association, East Lansing, Michigan, pp. 169–80.
- Benacerraf, P.: 1973, ‘Mathematical Truth’, *Journal of Philosophy* **70**, 661–79.
- Cartwright, N.: 1979, ‘Causal Laws and Effective Strategies’, *Nous* **13**, 419–37.
- Eells, E.: 1982, *Rational Decision and Causality*, Cambridge University Press, Cambridge, England and New York.
- Eells, E. and E. Sober: 1983, ‘Probabilistic Causality and the Question of Transitivity’, *Philosophy of Science* **50**, 35–57.
- Fetzer, J. H.: 1971, ‘Dispositional Probabilities’, in R. C. Buck and R. S. Cohen (eds.), *Boston Studies in the Philosophy of Science VIII (PSA 1970)*, Reidel, Dordrecht, pp. 473–82.
- Fetzer, J. H.: 1974, ‘Statistical Probabilities: Single Case Propensities vs. Long-Run Frequencies’, in W. Leinfellner and E. Köhler (eds.), *Developments in the Methodology of Social Science*, Reidel, Dordrecht, pp. 387–97.
- Fetzer, J. H.: 1977, ‘Reichenbach, Reference Classes, and Single Case “Probabilities”’, *Synthese* **34**, 185–217. Errata, *Synthese* **37** (1978), 113–14. Reprinted in Salmon (1979).
- Fetzer, J. H.: 1981, *Scientific Knowledge*, Reidel, Dordrecht.
- Fetzer, J. H. and D. E. Nute: 1979, ‘Syntax, Semantics, and Ontology: A Probabilistic Causal Calculus’, *Synthese* **40**, 453–95.
- Fetzer, J. H. and D. E. Nute: 1980, ‘A Probabilistic Causal Calculus: Conflicting

- Conceptions', *Synthese* **44**, 241–46. Errata, *Synthese* **48** (1981), 493.
- Giere, R. N.: 1973, 'Objective Single-Case Probabilities and the Foundations of Statistics', in P. Suppes, L. Henkin, A. Joja, and G. C. Moisil (eds.), *Logic, Methodology and Philosophy of Science IV*, North-Holland, Amsterdam, London, pp. 467–83.
- Giere, R. N.: 1976, 'A LaPlacean Formal Semantics for Single-Case Propensities', *Journal of Philosophical Logic* **5**, 320–53.
- Hacking, I.: 1965, *Logic of Statistical Inference*, Cambridge University Press, Cambridge, England and New York.
- Hacking, I.: 1980, 'Grounding Probabilities from Below', in P. D. Asquith and R. N. Giere (eds.), *PSA 1980*, Philosophy of Science Association, East Lansing, Michigan, pp. 110–16.
- Jeffrey, R. C.: 1965, *The Logic of Decision*, 2nd edn., 1983, University of Chicago Press, Chicago and London.
- Kyburg, H. E., Jr.: 1974, 'Propensities and Probabilities', *British Journal for the Philosophy of Science* **25**, 359–75.
- Kyburg, H. E., Jr.: 1978, 'Propensities and Probabilities', in R. Toumela (ed.), *Dispositions*, Reidel, Dordrecht, pp. 277–301. This is a slightly revised version of Kyburg (1974).
- Popper, K. R.: 1957, 'The Propensity Interpretation of the Calculus of Probability, and the Quantum Theory', in S. Körner (ed.), *Observation and Interpretation in the Philosophy of Physics*, Dover Publications, New York. pp. 65–70.
- Popper, K. R.: 1959, 'The Propensity Interpretation of Probability', *British Journal for the Philosophy of Science* **10**, 25–42.
- Reichenbach, H.: 1949, *The Theory of Probability*, University of California Press, Berkeley, Los Angeles.
- Russell, B.: 1948, *Human Knowledge: Its Scope and Limits*, Simon and Schuster, New York.
- Salmon, W. C.: 1967, *The Foundations of Scientific Inference*, University of Pittsburgh Press, Pittsburgh.
- Salmon, W. C.: 1971, *Statistical Explanation and Statistical Relevance*, University of Pittsburgh Press, Pittsburgh.
- Salmon, W. C.: 1979, *Hans Reichenbach: Logical Empiricist* (ed.), Reidel, Dordrecht.
- Settle, T.: 1975, 'Presuppositions of Propensity Theories of Probability', in G. Maxwell and R. M. Anderson, Jr (eds.), *Minnesota Studies in the Philosophy of Science VI: Induction, Probability and Confirmation*, University of Minnesota Press, Minneapolis, pp. 388–415.
- Sklar, L.: 1970, 'Is Probability a Dispositional Property?', *Journal of Philosophy* **67**, 355–66.
- Skyrms, B.: 1980, *Causal Necessity*, Yale University Press, New Haven, London.
- Suppes, P.: 1973, 'New Foundations of Objective Probability: Axioms for Propensities', in P. Suppe, L. Henkin, A. Joja, and G. C. Moisil (eds.), *Logic, Methodology and Philosophy of Science IV*, North-Holland, Amsterdam, London, pp. 515–29.
- Suppes, P.: 1974, 'Popper's Analysis of Probability in Quantum Mechanics', in P. A. Schilpp (ed.), *The Philosophy of Karl Popper*, Open Court, La Salle, Illinois, pp. 760–74.
- Van Fraassen, B. C.: 1979, 'Relative Frequencies', in Salmon (1979).

Von Mises, R.: 1957, *Probability, Statistics and Truth*, Macmillan, New York (2nd English edition).

Von Mises, R.: 1964, *Mathematical Theory of Probability and Statistics*, H. Geiringer (ed.), Academic Press, New York.

Dept. of Philosophy
University of Wisconsin-Madison
Madison, WI 53706
U.S.A.