Henry Kyburg

The Deductive Model: Does It Have Instances?

1. By the deductive model, I mean the deductive model of scientific inquiry. There are plenty of deductive systems around, of course: arithmetic, number theory, set theory, probability theory, even axiomatizations of scientific theories, for example in physics, in biology, and even in psychology. What I have in mind in its simplest form is the good old hypothetico-deductive model: formulate your hypothesis clearly; draw testable implications from it; put those implications to the test; if they fail the test, reject the hypothesis; and if they pass, the hypothesis is to that extent confirmed.

Everybody knows that things are a little more complicated than this, and numerous suggestions of alternatives have been made. Karl Popper (1959) suggests that we make sincere efforts to refute hypotheses having great content; but the mechanism of falsification was still deductive. Lakatos (1970) applied the deductive model only within a research program, and offers other criteria for evaluating research programs. Even Clark Glymour (1980), who rejects some aspects of the hypothetico-deductive method in his bootstrap technique, still holds to the general deductive model in many respects. What counts in the straightforward deductive model are deductive relations, given background assumptions, or given a theory T, or given a *really sincere* attempt to look for contrary evidence, or given a paradigm, or given a research program.

What is more curious is that the great modern inductivist Carnap, as well as others who have followed him in focussing on confirmation, also accepts the deductive model. Much of Carnap's work is concerned precisely with the attempt to elucidate the force of conforming instances: refuting on nonconforming instances never seemed problematic.

It is more curious yet that many philosophers and statisticians who have been concerned with statistical inference, and indeed the vast majority of those scientists who employ statistical inference as an everyday scientific tool, accept the deductive model. Significance testing, Neyman's rejection rules, fall right into the classical pattern. Loosely speaking, the idea is that you specify in advance some data such that *if* they are observed, you will reject the hypothesis being tested.

With respect to the issues I propose to focus on, the ordinary Bayesian view of scientific inference is just as wedded to the deductive model as those I have already mentioned. If a hypothesis H entails a statement E, and $\sim E$ is observed or otherwise comes to have probability 1, H must come to have probability 0. Furthermore, in the nonextreme cases the deductive rules of the probability calculus provide the connection between prior and posterior degrees of belief.

There are three aspects of the deductive model to which all the authors I have mentioned seem firmly committed.

First: There is a deductively closed and strictly consistent set of statements accepted by the scientist. This set may be empty (for a superextreme Bayesian); or it may include any statements that in a given context are regarded as "unproblematic"; or it may even include the very general sorts of statements that characterize a paradigm or a research program.

Second: There is a set of statements (axioms of theories, hypotheses) that are considered or contemplated. These are the statements that are up for test. They may be quite general, or they may be quite specific, but in any case they have empirical content, and they are to be confirmed, in some sense or other, by their instances or consequences, or they are to be refuted by counterinstances.

Third: There is a set of statements that is directly testable, statements that express the results of experimentation or observation directly and incorrigibly.

According to this general view, scientific procedure consists in putting statements of this third class to experimental verification or refutation; the result may be the deductive refutation of a statement belonging to the second class (Popper); or the increase in degree of confirmation of the hypothesis in the second class (Carnap); or the general rearrangement of degrees of belief assigned to items in the second class (subjectivist Bayesian); or the verification of a statement in the second class, in virtue of general assumptions embodied in the first class (Lakatos, Sneed); or the basis for a computation providing confirming instances of statements in the second class (Glymour); or it might be the grounds for the rejection of a null hypothesis (for the practicing statistician).

I shall argue that science doesn't work this way at all, not even an idealized version of science, and thus that there are no instances of the deductive model, and should be none. In particular, I shall deny each of the three assertions underlying the deductive model; the result will be one possible *inductive* model:

A: I shall assume that there is a set of statements accepted by the scientist, but I shall deny that the complete set of accepted statements is consistent or deductively closed.

B: I shall deny that scientific laws and hypotheses can be up for test, since I shall deny that they have empirical content at all. Since they have no empirical content, they neither can be "confirmed" by their instances or their consequences, nor can they even be refuted deductively by counterinstances.

C: I shall assume that there is no set of statements that unproblematically express the results of experimentation or observation. I shall indeed suppose that some such statements may be accepted on the basis of observation; but I shall not suppose that they are incorrigible, once accepted.

Nonetheless, I shall assert, what it has recently become fashionable to deny in some quarters, that scientific argument *is* argument, and that it conforms to canons of rationality. The bearing of evidence on the acceptability of laws, on the choice between theories, etc., can be reflected by formal logical relations.

2. First, some preliminaries. Like most proponents of the deductive model, I shall suppose we can represent scientific knowledge in a first-order language; I suppose that the language contains some axioms for set theory, and thus all the mathematics we need, including the measure theory needed for statistics.[1] In the course of scientific inquiry, this representation changes. There is change in the vocabulary; the term "phlogiston" is dropped; the term "quark" is introduced. There is also change in the meaning of terms that persevere through scientific change: the quantity *m* (interpreted "mass of") changes its meaning in the transition

from Newtonian to relativistic mechanics. This requires some explanation, a bit of which I'll provide later.

Note that there is nothing in the language itself that represents the distinction between observational and theoretical vocabulary.

A rational corpus will simply be a set of statements in the language, subject to a single constraint: that it contain all the consequences of any statement it contains. I include among the consequences those arrived at with the help of axioms of the language. We assume that these axioms are consistent; if we learned that they were inconsistent, we would change them to eliminate the inconsistency. It is already clear that no one's body of actual or dispositional beliefs can be a rational corpus—no one believes all the truths of set theory or is even inclined to believe them explicitly. Nevertheless it is a useful notion, for many people accept (say) Zermelo Frankel axioms for set theory, and thereby *commit* themselves to the consequences of those axioms. A rational corpus represents a set of statements to which one is committed regardless of one's doxastic state.

Note that neither deductive closure nor even consistency is built into the notion of a rational corpus. It is only the set of consequences of the axioms alone that is deductively closed. There is only one rational corpus that is explicitly inconsistent (the corpus containing all statements), but there are any number whose deductive closure is inconsistent. (For example, a corpus could contain the premises for the lottery paradox, but need not contain the conclusion: no ticket will win.)

Probability will be a function from pairs consisting of a rational corpus and a statement to subintervals of [0, 1]. It is objective and logical: i.e., given a rational corpus K and a statement S, the probability of S is completely determined to be the interval (p, q) regardless of what anyone knows or believes. Underlying any such probability is a statement of frequency (or, rarely, chance) that is a member of the corpus K. This statement of frequency may be empirical, or it may be a consequence of the axioms of our language.[2] (For example, the axioms of set theory, or the axioms characterizing "fair die.") The definition of probability involves no more than direct inference.

Rational corpora of individuals (or groups, or disciplines) come with indices reflecting criteria of acceptance. A sentence belongs to a corpus of level .9, for example, if its probability relative to a (specified) corpus of level higher that .9 is a subinterval of [.9, 1].[3]

There are two ways in which statements get into rational corpora, other than by being consequences of the axioms of the language: by inference and by observation. Both involve probability.

Acceptance by inference has already been characterized: S belongs to the corpus K of level r, if its probability relative to the specified corpus of level higher than r is greater than r. It is a theorem of probability (in my sense) that if p entails q, then the minimum probability of q is at least as great as that of p.

Acceptance by observation is more complicated. Borrowing some terminology from Mary Hesse (1974), let us take an observation *report* to be a sentence of our language we are motivated to accept by what happens to us. Not all such sentences can be true simultaneously; we know that some must be false, for the set of reports may well be inconsistent with the axioms of our language. From this we can infer that each kind of observation report is characterized by a certain long-run frequency of error. We accept an observation *report* (i.e., we construe it as an *observation sentence*) in the corpus of level r, just in case its probability of *not* being in error, relative to the specified metacorpus of level higher than r, is higher than r.[4]

Note that the *set* of observation sentences in the corpus of level r may still be inconsistent with our axioms. For example, we may have a large collection of observation reports of the form Rx, of which we know that less than 1 percent are erroneous. Still, if we have no way of picking out any particular ones that are in error, they may all be characterized by a probability of error of less than .01, and they may all thus end up in the rational corpus of level .99.

One more piece of machinery, and we'll turn to some examples. The predictive observational content of a rational corpus is the set of sentences in that corpus that get there by inference and not by observation, but that are of the sort that *could* get there by observation. Thus if Rx is a highly dependable form of observation report, and the sentence "Ra" appears in a rational corpus due to inference and not due to observation, then it is a member of the predictive observational content of that corpus. I assume that there is some natural way of constraining the predictive observational content of a rational corpus to be finite, so that we can compare the contents of two rational corpora expressed in different languages.[5]

3. Let us begin with measurement. We might suppose that our judgments of relative length were infallible and incorrigible. We would

have to suppose, then, that few of the conventional properties of relative length actually held. In particular, the transitivity of *being the same length as* would have to be regarded as false. But we would never have to reject an observation statement; and we could even include some predictive observation statements in our rational corpus: it might well be true of *almost* all triples of objects that if the first is longer than the second, and the second longer than the third, then the first is longer than the third.

But we might supposed instead that "longer than" satisfied the ordinary axioms, and that our judgments of relative length were sometimes erroneous. For example, it will follow from these axioms that *being the same length as* is transitive. Once we accept that as axiomatic, then there is no test that will refute the generalization. And that is important, because we have all observed prima facie refutations of it: the most obvious occur when we measure something a number of times and get a number of different values. In fact, under these circumstances (since we have a rather complete theory of error for this case) we do not accept any of the observation *reports* as observation *statements*, but rather use them, together with that theory of error, to infer a "theoretical" sentence (theoretical in scare quates) that we do add to our rational corpus: the length of x lies in the interval L plus or minus d. This is no doubt what an ordinary scientist means by an observation statement.

The result of this change from regarding the predicate *longer than* as strictly observational but without structure, to regarding it as having structure but being subject to observational error, is an enormous increase in the predictive observational content of our rational corpus. I cannot demonstrate this now, but it is at least intuitively plausible. That it is advantageous in this way does depend on facts about the world: on the fact that the errors of observation we introduce to allow ourselves to stick to the handy structure we attribute a priori to "longer than" are ordinarily rather small relative to our concerns with relative length. It also depends on the fact that there are ways of reducing those errors, but this is a complex fact and depends on far more in the way of scientific knowledge than the mere existence of relatively rigid bodies. And it depends importantly on the contribution to communication among people that standard scales of length make possible.[6]

Let us turn to something more empirical than the theory of lengths of rigid bodies. How about the law of thermal expansion: that the change in length of a object is proportional to its change in temperature. Well, of course that isn't really a law: it applies only to homogeneous substances of certain sorts, not undergoing changes of state, and then only within certain limits. So it is hedged around with provisos. And of course if we are to test it with a mercury thermometer we are presupposing that the law holds for mercury. What is going on here? Well, to begin with, we have a way of measuring length. We also have an indirect way of measuring temperature (possibly several ways) that may depend on the uniform expansion (or apparently uniform expansion) of mercury. That in itself might give us the idea of thermal expansion, as applied to metals in general, for example. This is reminiscent of Clark Glymour's idea of bootstrapping. There is no reason why a special case of the law shouldn't enter into the confirmation of the law in its generality. But even taking this for granted, do we now hold this law in a kind of epistemological limbo, and look for refuting instances? De we subject it to thorough test before we accept it?

How would we do it? We might take an iron bar and heat it up, carefully measuring it and its temperature before and after. This would allow us to compute a coefficient of thermal expansion of iron, but it is no *test* of the theory. Well then, we perform the experiment again, using a different bar and heating it a different amount. Does it expand the predicted amount? No. Does that refute the law? No. It gives us a different value for the coefficient of thermal expansion. Wait, though. We've forgotten the errors of measurement of length and temperature. These will be reflected in an error for the coefficient of thermal expansion, and we might demand that the second trial yield a number that is the same as the first, "within experimental error." Let us unpack this. From the distribution of errors of measurement of length and the distribution of errors of measurement of temperature, together with a bit of knowledge concerning the independence of these errors, we can derive a distribution of errors of (indirect) measurement of the coefficient of thermal expansion, *given* the truth of the law we are testing. The trouble is that these distributions are generally taken to be approximately normal, and thus to be unbounded: no conceivable set of measurements can conflict with the law. No set of numbers can fail to be an "instance" of the law "within experimental error."

Let us try again. We know a lot about the distribution of errors of measurement of length; rather less about the distribution of errors of measurement of temperature. Suppose they are both approximately normal. Given an observation report of a measurement, we can accept in a corpus of level r, the assertion that the measured length lies within certain limits. This is an *inductive* inference: it is a direct inference from the distribution of errors in general to a conclusion about the magnitude of an error in a particular case. The same is true for each of our four measurements of length and four measurements of temperature. This is still no help (although it does correspond to the way people sometimes talk), because what we require to achieve a possible counterinstance to the proposed law is the *conjunction* of these eight statements, and it is obvious that each of them can be acceptable at the level r while their conjunction is not. Very well, we can consider the joint distribution of the eight errors, and choose our bounds for each of the quantities in such a way that an appropriate conjunction appears in our corpus of level r; this conjunction might or might not refute the proposed law. But there are a lot of ways of choosing those bounds (we could take some quite narrowly and others quite broadly, or vice versa), some of which might lead to conflict with the proposed law, and some of which might not, and this all at the same level of practical certainty.

Let us try once more. This time we shall assume that the law is true, and that all of our measurements are to be understood as a computation of the coefficient C. From the truth of the law and the error distributions for the measurements of length and temperature, we can derive an error distribution for C. Our first set of measurements yields the result that the assertion that C lies in such and such an interval (I_1) is in our corpus of level r. We might suppose that our second set of measurements could yield the result that the assertion that C lies in I_2 is in our corpus of level r. Suppose that I_1 and I_2 do not overlap at all (or overlap very little); would this refute our law? or test it? This can't happen! Why not? Because in order for "$C \varepsilon I_2$" to have a probability of at least r, the (indirect) measurement of C must be a random member of the set of such measurements with respect to having a given error, and the previous measurement of C precludes this. Just as in sampling from a binomial population to estimate the binomial parameter p, the conditions of randomness require that we not make two inferences based on two parts of the sample, but consider the *whole* sample as the data for our inference, so in this case, the requirements of randomness demand that we base our inference concerning C on the *total* evidence available. Thus in fact what we have in our corpus of level r is neither "$C \varepsilon I_1$" nor "$C \varepsilon I_2$," but "$C \varepsilon I_c$," where I_c is the interval computed from the total evidence—i.e., all sets of observations.

Is there no way that the evidence can refute the law? Let us make one last try. Given that we know the distribution of errors in the measurements of lengths and temperatures, and given that the law is true, the distribution of errors of the indirect measurement of C will have a known distribution. Could a sample of measurements of C lead to the rejection of this distribution of errors? Not if the distribution is *known*. If the distribution is known, an awkward collection of sample values can only be evidence of bad luck. At a certain point, though, doesn't one reject the notion of bad luck, and thus the assumed distribution of errors of measurement of C? Sure. But notice that this assumed distribution depended on three things: the distribution of errors of measurement of length, and distribution of errors of measurement of temperature, and the law of thermal expansion. Let us simplify things by supposing that the evidence on which our knowledge of the distribution of errors of measurement of length is based is so extensive that we can take that statistical generalization for granted.

There are then two courses open to us. We may take our body of evidence as throwing new light on the distribution of errors of measurement of temperature, or we may suppose that it puts our law of thermal expansion in a bad light. How do we decide which? I propose that we decide in the same way as we decide whether or not to accept the transitivity of *being the same length as* or *being the same temperature as*. Thus it seems to me that the law should be regarded as a proposed *linguistic* reform: if we accept it, the data regarding the distribution of C are to be construed as data bearing on the distribution of errors in the measurement of temperature. And surely, if the distribution of observations of C is close to that expected, that is just what we would do. We would regard those observations as a bit more evidence regarding errors of measurement of temperature.

This means that in considering the law and its evidence, we are implicitly choosing between two languages: one that contains the law, and one that does not. To this choice we apply the criterion mentioned earlier: which language yields us a corpus of practical certainties with the greatest predictive observational content? If adopting the language containing the law yields the result that our measurements of temperature are subject to much more error than we thought, then the corpus of practical certainties using the language will have to forego observational predictions concerning temperature that the corpus based on the other language could retain. Note that it would have to forego observational predictions concerning temperature that had nothing to do with the law of thermal expansion, so that there is a holistic element in the criterion. But if adopting the language contianing the law led to the result that our measurements of temperature were only a little, or not at all, more inaccurate than we thought, the predictive observational content of our body of practical certainties would be significantly enlarged.

For simplicity, I have been speaking as if the law of thermal expansion were universal. It isn't, and most laws—at least, laws of this sort—have limited scope. The scope of the law is easy to build into this treatment. For a certain range of temperatures of a certain set of substances, distributions of values of their coefficients of thermal expansion are included as part of the data we have concerning the errors of measurement of temperature. For other temperatures and other substances, the law just doesn't apply. The full statement of the law in our language thus contains its limitations. Note that these limitations are *discovered* empirically: we do not confirm the law for each of a class of applications, but take it as universal, and discover its limitations.

Of course nowadays we know a lot more about the law of thermal expansion than is embodied in the special case I have been considering. This increasing sophistication is allowed for on the view of testing I am proposing. A simple illustration can be found in the gas law: At one point in time it is a feature of our scientific *language* that for relatively rarified gases, $PV = nRT$. Observations are to be construed as throwing light on the errors of measurement of pressure, volume, and temperature. (In fact, for high temperatures, this law does provide a means of measuring temperature by means of the gas thermometer.) At a later point in time it becomes a feature of our scientific *language* that for a wider range of instances, $(P + (a/v^2))\ (v - b) = RT$, where v is specific volume, and a and b are characteristic of the gas in question. This is van der Waals's equation, and it is clear that for small values of a and b, the ideal gas law is a reasonable approximation to the truth. Now of course what is taken as a feature of our scientific language is yet more complicated; the ideal gas law and van der Waals's equation provide handy computational approximations (instruments?) but are strictly false, even as applied to the best old instances.

5. What does this approach have to say in general about the confirmation, disconfirmation, degree of credibility, and testing of scientific laws and theories? In one sense, it settles the whole problem—it reduces it to epistemology. We don't confirm laws, we don't test them, we don't

disconfirm them, because they are taken to be a priori features of our language. But we do have criteria, expressed in terms of a certain sort of content of our bodies of knowledge, for preferring a language with one a priori feature to a language with a different a priori feature.

Nevertheless something goes on that seems like the process characterized by the hypothetico-deductive model, or like confirmation by instances, or like Glymour's bootstrapping, or like something involving simplicity. Something should be said by way to explanation of the apparent relevance of these notions.

When we have a great deal of knowledge about the distributions of errors of measurement (direct *or indirect*) of the quantities that appear in a proposed law, we don't have to think very hard when we obtain a set of measurements that are widely at variance with the law. We forget about the law. Why? Because it seriously undermines our presumed knowledge of error distributions, and thereby impoverishes rather than enriches the predictive content of our body of knowledge. To show in detail how this comes about is nontrivial. I have made some efforts along these lines in my article "The Justification of Deduction in Science" (Kyburg, forthcoming). I have also argued there that in a very weak sense the successful testing of a law (the establishment of approximately confirming instances) does enhance its "acceptability"—instance confirmation may well help, particularly when indirect measurement is involved, to reduce the dispersion of our errors of measurement.

Nevertheless the general progress of scientific inquiry is contrary to that suggested by the deductive model. As the accuracy of measurement improves in a given domain, the discrepancies between our laws and our experimental results become more pronounced. *Successful* laws are progressively *dis*confirmed.

There are some interesting connections between Glymour's notion of bootstrapping and the approach I have been exploring. One is the fact that on Glymour's reconstruction, part of a theory may be needed to compute a parameter in a hypothesis, or a theoretical quantity. This is illustrated in my case by the calculation of the coefficient C in the example I belabored. It is easy to see, however, that one can only *accept* in his rational corpus an inequality like $C_1 < C < C_u$. But that is as it should be, and it is exactly what is needed to yield new predictive content.

Another connection lies in the valuable notion of a *computation*. Glymour shows computations going only one way, from data to parameters. It is important to note that error already enters here: we must begin with observation reports, embodying error, and trace the distribution of error through the calculation to arrive at a distribution of error for the quantity or parameter in question. But in applications of science—even in the case of simple predictions—the computation goes both ways. We compute the (credible) values of theoretical quantities on the basis of observation reports, then pass through yet further computations of theoretical quantities (keeping track of error distributions as we go along), and then back down to observation *sentences* (e.g., the angle almost certainly lies between 130.15° and 130.25°), and thence, through errors of measurement again, to a prediction (e.g., the angle will be reported to be between 129.5° and 130.5°).[7] What strikes me as most important and valuable about the notion of a computation is that it allows us to calculate what I would like to call a *trajectory of error* from observation report to observation report, as well as the trajectory of error from observation report to observation sentence to parameter to theoretical quantity.

Let me turn to questions of generality and simplicity. One aspect of generality and simplicity has already been touched upon. We discover statistical limitations in the applications of laws that lead us to prefer a language in which the scope of a law is limited: in other words, we make our (simple) conventions as general as we can get away with making them. But there are other aspects. One concerns the reduction of errors and the elimination of "discrepancies." In addition to extending the useful range of applicability of a law or theory by improving our techniques of measurement—and the distribution of errors provides a direct criterion of "improvement"—we may have a theory according to which a certain quantity is *mainly* a function of something else. For example, the orbit of a planet is mainly due to forces acting between the planet and the sun, but also, strictly speaking, a function of the forces exerted by other celestial bodies. In such a case the theory itself may provide an account of the deviant probabilistic difference between prediction sentence and observation report. And at the same time, it can tell us how to go about taking account of these systematic errors. It can tell us (if only roughly) how to "subtract the effect of the sun" on the orbit of the moon, or in what ways it would be *legitimate* to falsify and idealize the data to be used in calculating theoretical parameters or quantities. But the theory can function in this way only if it is construed as an a priori feature of our language.

There is another way in which a theory can be used to reduce the component of error in predictive sentences. This may be relevant to the discussion of Ptolemaic and Copernican astronomy. Either theory, it has been pointed out, can be made to fit the phenomena reasonably well. To do so involves determining parameters, and this determination is subject to error. On the Ptolemaic theory we have just one set of parameters characterizing the motion of each planet relative to the earth. On the Copernican theory this relative motion is characterized by *two* sets of parameters: one characterizing the motion of the earth about the sun (call this set E) and one characterizing the motion of the ith planet about the sun (call this P_i). On the Ptolemaic theory, observation of a given planet helps *to reduce the uncertainty in the parameters characterizing its movement* about the earth, independently of observations of other planets. On the Copernican theory, there are more parameters to worry about (E as well as P_i), but observations on the ith planet help to reduce the errors in E—and hence in the apparent motions of other planets—as well as P_i. Generality can thus contribute to the control of error. And of course one's concerns are prospective as well as immediate: the prospect of better control of error in the future, particularly as regards systematic error, has a bearing on our choice of theoretical languages.

I suppose that in some sense or other I am a victim of East Coast Holism: Here I am not only saying that our body of scientific knowledge must be treated as a single ball of wax, but that scientific change involves the very language in which we do science. But it is one thing to solemnly intone "All in One," and something else again to provide enough analysis of the Great White Unity to allow one to tinker with one part at a time and to choose rationally between two alternative balls of wax that differ only in specifiable limited ways. It is the latter sort of thing that I think worth pursuing, for the sake of the improvement of the understanding. But I confess that I have no more than sketched the beginnings of an approach.

I have attempted to provide some prima facie plausibility for my three preposterous claims: A, that our body of scientific practical certainties need be neither consistent nor deductively closed; B, that scientific laws need not be construed as having empirical content, nor as being confirmed by their instances; and C, that we need not regard any observations at all as incorrigible. I should also remark, though, that the deductive model is to the inductive model as the ideal gas law to van der Waals's equation: When we have adopted van der Waals's equation, the ideal gas law has no instances, but it is a handy approximation.

The analogy yields a further implication: when we have yet more insight into science, the inductive model I have sketched will probably have no instances either, just as van der Waals's equation has no instances today. But I think the inductive model will even then serve as a calculational approximation over a wider range of situations than the deductive model does.

Notes

1. There has in recent years been a stir about replacing the consideration of formal languages by models. Although there are advantages in so doing—we needn't bother about unintended models of our languages, we can replace the notion of deducibility by the tighter notion of model-theoretic entailment—these advantages are not important for present purposes. And we shall have occasion to refer to certain actual sentences—observation *reports*—of the language of science anyway.

2. It is a set-theoretical truth that most subsets of a given set resemble it with respect to the proportion of objects having a given property. This is an assertion without empirical content. Yet it may nevertheless be useful, when, for example, a specific subset—a sample that we have drawn—is a *random member* of the set of n-membered subsets of a given set. Similarly, it is vacuous to say that a fair die, tossed fairly, has an equal chance of exhibiting each face; but this can nonetheless be a useful piece of information if we can be *practically certain* that a given setup involves a (practically) fair die being tossed (nearly) fairly.

3. We must *specify* the high-level corpus, for it may be that there are two corpora of level higher than .9 relative to which the statement has different probabilities. For more details about probability, see Kyburg (1974).

4. We speak here of *meta*corpus since the argument concerns frequencies of error among sentences.

5. In the case of individuals, for example, we might suppose a maximum lifespan of a hundred years, and a maximum number of kinds of discrimination expressible in a given vocabulary (say, 10,000), and a minimum time to make an observation (say, a tenth of a second), and conclude that the upper bound on the contents of a rational corpus of an individual is $100 \times 365 \times 24 \times 60 \times 60 \times 10 \times 10{,}000 = 3.1 \times 10^{14}$ items. Here is where we might build in a cost function for both computation and for observation. *Here* is where "logical omniscience" becomes constrained.

6. More detail will be found in Kyburg (1979).

7. The distinction, introduced by Mary Hesse (1974), between observation statements and observation reports, is a handy one. It is somewhat elaborated on in Kyburg (forthcoming).

References

Glymour, Clark. 1980. *Theory and Evidence*. Princeton: Princeton University Press.

Hesse, Mary. 1974. *The Structure of Scientific Inference*. Berkeley: University of California Press.

Kyburg, Henry E. Jr. 1974. *The Logical Foundations of Statistical Inference*. Dordrecht: Reidel.

Kyburg, Henry E. Jr. 1979. Direct Measurement. *American Philosophical Quarterly*. 16: 259-272.

Kyburg, Henry E. Jr. Forthcoming. The Justification of Deduction in Science.

Lakatos, Imre. 1970. Falsification and the Methodology of Scientific Research Programmes. In *Criticism and the Growth of Knowledge*, ed. Imre Lakatos and Alan Musgrave, Cambridge: Cambridge University Press, 1970.

Popper, Karl. 1959. *The Logic of Scientific Discovery*. London: Hutchinson and Co.