

J. ALBERTO COFFA

HEMPEL'S AMBIGUITY¹

Explanation theory abandoned its pre-theoretical stage and became a respectable branch of philosophical inquiry when, in the late forties, Hempel began to develop his model of deductive nomological (D-N) explanation. In a sequence of now classic papers he succeeded in articulating an illuminating philosophical account of explanation which provided compelling evidence for the adequacy of the philosophical views captured by the D-N model.

In the early sixties Hempel turned his attention to the topic of inductive explanation. Until then, it had been generally believed that the inductive model had to be understood as a generalization – and a rather straightforward one at that – of the deductive model. Yet, already in the Hempel-Oppenheim paper a warning had been issued to the effect that such generalization raised “a variety of new problems”. Indeed, when finally, in a sequence of illuminating papers on inductive explanation, Hempel decided to face one of these problems, he felt forced to propose a theory of inductive explanation which differed drastically from pre-analytic consensus on the nature of such explanations.

Not the least of these departures was Hempel's implicit rejection of the claim that the deductive model is a limiting instance of the inductive model. Yet, much more than this was involved. We should like to argue that, in spite of misleading appearances of continuity, the philosophical understanding of explanations implicit in the model of inductive statistical (I-S) explanation which Hempel eventually produced is drastically different from, if not incompatible with, that which inspired his D-N model. One of the purposes of this paper is to draw attention to the nature and magnitude of the shift involved. Another is to explain why Hempel's views on inductive explanation ought not to be accepted.

The evolution of Hempel's thought was causally related to his analysis of a problem which, pending more illuminating designations, we will refer to as ‘Hempel's problem’. Due to it Hempel felt forced to propose an account of inductive explanation which contained a rather unexpected

feature, a peculiar form of relativization to knowledge that, according to Hempel, is unavoidably present in every adequate theory of inductive explanation.

The question whether such epistemic relativization as Hempel has introduced is avoidable or not is interesting in its own right. But it becomes pressing if one believes – as I do – that no characterization of inductive explanation incorporating that feature can be backed by a coherent and intelligible philosophy of explanation. If I am right, it is of more than passing interest to give a closer look at Hempel's problem. For unless one can find a way to avoid the conclusion Hempel drew from it, one may well have to accept that the concept of inductive explanation is as much of a conceptual delusion as some have claimed modal concepts to be.

My strategy in this paper will be the following. I will first attempt to locate the nature of Hempel's problem. This will prove to be unexpectedly difficult. Next, I will examine the epistemic relativization that Hempel felt forced to introduce, as well as his reasons for introducing it. Having done this, we will be in a position to examine the question of the unavoidability of Hempel's epistemic relativization. Up to this point the argument will be reasonably inter-paradigmatic. It will become less so when I finally turn to explain why an epistemically relativized theory of explanation à la Hempel cannot be taken to be a theory of explanation. In a way, my argument will boil down to the claim that I can't see how it could; but I will try to hide its unacademic form under the cloak of an analysis and rejection of those reasons that could conceivably be given in defense of such models. Here my remarks will be cautiously brief and somewhat cryptic. They will be even more so when, in the last paragraphs, I attempt to suggest the way in which an alternative theory of inductive explanation could be developed, based upon a solution of Hempel's problem. We turn, first, to Hempel's problem.

1. HEMPEL'S PROBLEM

There is a widespread tendency to view the simplest form of the D-N model (its 'basic form') as being that of

$$(I) \quad \frac{(x) (Fx \supset Gx) \quad Fa}{Ga} ,$$

where ' Ga ' is the explanandum, ' $(x) (Fx \supset Gx)$ ' is a nomic statement and ' Fa ' is a statement of initial (and boundary) conditions. When the premises of (I) are true, one says that (I) is a D-N explanation, or (for emphasis) a true D-N explanation; when they are well confirmed on the available evidence one says that (I) is a well confirmed D-N explanation, relative to the available evidence (or knowledge situation).

If one views the D-N model as in (I), attempts to generalize it into an inductive theory will naturally start fixing attention on what seem to be its only generalizable features: the kind of connection that the nomic premise asserts to obtain between the attributes it mentions, and the kind of connection obtaining between the premises and the conclusion of the argument. Both relationships can be generalized into probabilities, yet not into the same sort of probability. The deterministic connection asserted by the first premise can be weakened or generalized into a frequentist or statistical correlation, whereas the deductive link between premises and conclusion can be weakened or generalized into an inductive link. The deductive deterministic model can be thus generalized into an inductive statistical model.

Thus, it seems most natural to conclude that inductive explanations of basic form ought to be understood as follows, in what we will refer to as *the naive model of inductive explanation*. An I-S explanation (or true I-S explanation) of basic form will be an argument of the form

$$(II) \quad \frac{p(G, F) = r \quad Fa}{Ga}$$

together with the number $c(Ga, p(G, F)=r \& Fa)$ which, we assume, is also r ; where r is close to 1 and the premises of (II) are true, the first one being a law of nature.

Although probably no one has ever offered this precise characterization of inductive explanations, one may conjecture that something essentially like this has been lurking in people's minds when they talked about inductive explanations. Yet, as soon as Hempel turned his attention to the inductive theory of explanation, he noticed that the naive model was quite unacceptable.

Hempel observed that given that all of the premises of (I) are true, it is still possible that, for some property H , ' Ha ' is true and ' $p(-G, H)=s$ '

is a true nomic statement where s is close to 1. But then, not only (II) but also

$$(III) \quad \frac{p(-G, H)=s}{\begin{array}{l} Ha \\ \hline -Ga \end{array}}$$

would be an inductive argument with true premises, and a conclusion that is implied with high inductive probability by the premises. The existence of arguments like (II) and (III) is what Hempel called "the phenomenon of ambiguity".

Clearly, the phenomenon of ambiguity amounts to the fact that there can be naive inductive explanations of mutually inconsistent statements. The assumptions that most Texans are millionaires and that most philosophers are not millionaires are compatible with the assumption that Jones is a Texan philosopher. If all of these assumptions are true and if the statistical correlations are nomic, then we have two naive inductive explanations, one explaining that Jones is a millionaire and the other explaining that he is not.

Hempel felt that there was something very undesirable about this consequence; indeed, undesirable to the point of constituting a *reductio ad absurdum* of the naive model. Thus, according to Hempel, the phenomenon of ambiguity shows that there is something hopelessly wrong about that model; it shows that it incorporates a certain feature that immediately implies the worthlessness of the model. If we could identify this 'bad feature' precisely, we would have a well defined program and a problem that every explanation theorist should attempt to solve: to produce a definition of inductive explanation for which it can be shown that it does not share the bad feature in question.

Most of Hempel's explicit statements on the topic suggest that there is an obvious way of identifying this 'bad feature': it would consist in the fact that certain (otherwise plausible) characterizations of inductive explanation have instances with mutually inconsistent conclusions, or rather, in Hempelian terminology, that such definitions have inductive inconsistencies as instances (an *inductive inconsistency* being a pair of inductive arguments with mutually inconsistent conclusions which are implied with high inductive probability by their respective premises).

Let us say that a definition of inductive explanation *suffers from*

ambiguity when it has inductive inconsistencies as instances. (Note that since a definition of true inductive explanation must demand that its premises be true, such definitions may suffer from ambiguity only when the inductive inconsistencies that it has as instances have true premises. Hempel referred to this variety of ambiguity as *ontic ambiguity*. For definitions of inductive explanation demanding that the premises belong to a certain knowledge situation, the ambiguity in question was called by Hempel *epistemic ambiguity*.) The seemingly obvious construal indicated above is that the 'bad feature' is to suffer from ambiguity. Hempel's problem would then be that of finding a definition of inductive explanation that does not suffer from ambiguity.

This construal receives further support from the fact that throughout a series of papers in which Hempel has attempted to solve the problem raised by ambiguity, his efforts have had a consistently unique form: in all cases, a definition of inductive explanation was first offered, and then an attempt was made to prove that it did not suffer from ambiguity.

Nevertheless, a closer look at the structure of these attempts engenders a quite different impression concerning what Hempel's problem really is.

Notice first that if Hempel thought that his problem was that of producing a definition of inductive explanation not suffering from ambiguity, there would have been a most trivial solution which would apply in all imaginable cases, a solution which surely couldn't have escaped Hempel's attention. Given an arbitrary definition of inductive explanation, one could make it comply with the requirement to avoid ambiguity by adding to it a clause to the effect that the explanandum should be true (or, in an epistemic characterization, *known*). This would imply in the most straight-forward way that there are no (appropriately corrected) naive inductive explanations of mutually inconsistent statements. Furthermore the additional clause seems easily justifiable on traditional Hempelian standards according to which one can only explain what is the case (respectively, what is known to be the case). Thus, had Hempel thought that the problem uncovered by the phenomenon of ambiguity was the one described above, this problem would have deserved no attention whatsoever.

Yet, not only did Hempel think that the problem had no easy solution; in fact, for the case of true inductive explanations he thought that it

had no solution at all. And he claimed that this established the meaninglessness of the concept of true inductive explanation.

Anyone who interprets Hempel's problem as has been suggested above must also find it very difficult to comprehend why Hempel has set himself such narrow limits on what is to count as an allowable method to solve this problem. For, apart from a soon withdrawn appeal to the principle of total evidence in Hempel (1962), Hempel's way of dealing with his problem has always been to offer restrictions on the allowable reference classes in the nomic premises of inductive explanation. This was first done tentatively in Hempel (1962); a more precise requirement was offered in Hempel (1965), and Hempel claimed there that he could prove that the enforcement of this requirement guaranteed that his definition did not suffer from ambiguity. When Grandy showed that this was not the case, Hempel offered a considerably more complex requirement of the same kind in Hempel (1968) and a new proof that this new restriction on admissible reference classes guaranteed the desired effect. Yet, it turns out that the new proof also fails, in that it makes an unwarranted existential assumption.² The remarkable disproportion between the alleged aim of Hempel's seemingly unrealizable program and its limited means is underscored by the ironic fact that Hempel's definition of inductive explanation in Hempel (1968) requires that the explanandum belong to a knowledge situation; and this alone, under Hempel's demand that knowledge situations be consistent, immediately implies the result that Hempel has been at such great pains to prove on the basis of restrictions on admissible reference classes.

Are we to conclude that the Hempelian bird of happiness had always been in the philosopher's backyard? This seems unlikely. Rather, we would suggest, what seems called for is a revision of our understanding of what Hempel's problem is. Under our first construal, it appeared as a problem essentially concerned with the conclusions of inductive explanations, particularly, with the fact that they may contradict each other. But under this construal it proves impossible to make sense of Hempel's treatment of his problem. Hence, it seems necessary to provide a different account of what Hempel's problem is. The alternative that we would like to suggest is that Hempel is accurately, though maybe not clearly, perceiving a very real problem for inductive explanation, one that concerns not their conclusions but their premises; more precisely, the nomic

premise, and, within it, its reference class. We would like to suggest that when Hempel turned his attention to the theory of inductive explanation what he stumbled upon was the fact that the problem of defining a model of inductive explanation for single events was the other side of the coin of the single case problem. He stumbled, that is, upon the reference class problem.

There was in the theory of probability an old problem that had faded away, but not quite died. It was the problem of the reference class. This problem arises when, in the context of a frequency theory of probability, we attempt to answer the question 'what is the probability of a single event?', for example, 'what is the probability that this particular airplane will fall in its next flight?'. If one believes, as many frequentists did, that the only meaning probabilities can have is the frequency meaning, then in order to answer the question one must begin by identifying a reference class for the given event. But here is where the reference class problem arises. One and the same event can be associated with different reference classes, different to the extent that the probabilities of the event in such reference classes differ. This airplane belongs to the reference class of all airplanes and in that reference class the probability of its falling in its next flight is not high, but it also belongs to the reference class of airplanes whose wings will fall during the next flight, and as such its probability of not quite succeeding in reaching safely its destination is rather high. For the frequentist the question 'What is the probability of a single event' could make sense only if we could find a 'natural', 'appropriate' reference class for each event. But frequentists have traditionally held a principle of reference class democracy: in so far as the estimation of probabilities are concerned, all reference classes are created equal. Thus they would traditionally hold that it is strictly meaningless to assign a probability to a single event. The frequentist finds that, despite misleading appearances to the contrary, statements like 'the probability that this plane will fall on its next flight is high' are no more meaningful than guttural noises. Now, *if* one views statistical explanations as essentially concerned with placing the explanandum in an appropriate reference class, and *if* one countenances a frequentist interpretation for the nomic premise, then one of the main problems one will have to face while developing a theory of inductive explanation is the reference class problem. In effect, this is how Salmon saw the matter, since the essential

ingredient of his model is the claim that the 'appropriate' reference class in an explanation is the maximal homogenous reference class. We want to suggest that, although less obviously, this is also Hempel's understanding of what the main problem for a theory of inductive explanation is.

2. EPISTEMIC RELATIVIZATION

Let us leave for a moment the question as to what Hempel's problem is, and let me turn to consider briefly the conclusion Hempel felt forced to draw from it: the thesis of the epistemic relativity of inductive explanation. Since it is crucial to have a full understanding of what this thesis amounts to, I will devote a few paragraphs to the explanation of certain conceptual distinctions which will help us to grasp its force.

It is an obvious fact that the meaning of some expressions or concepts can be given without referring to knowledge, whereas that of others cannot. Let me call the latter epistemic and the former non-epistemic expressions. Examples of non-epistemic expressions are easy to find. 'Table', 'chair', 'electron', according to many, 'truth', would be typical instances. Examples of epistemic notions are also readily available. The best known instance may be that of the concept of confirmation. Although the syntactic form of expressions like 'hypothesis *h* is well-confirmed' may mislead us into believing that confirmation is a property of sentences, closer inspection reveals the fact that it is a relation between sentences and knowledge situations and that the concept of confirmation cannot be properly defined (that is, its meaning cannot be given) without a reference to sentences intended to describe a knowledge situation. Just as there are clear cut cases of each of these two kinds, there are concepts for which it is difficult to decide whether they are epistemic or not. Randomness is one such case. Some will argue that it describes possible properties of the world which may obtain or not obtain quite independently of the presence of knowledge. For these the concept is non-epistemic. Others argue that randomness means nothing if not relativized to knowledge, for the predication of randomness can only mean that the person who ascribes such a predicate lacks the (it is claimed, always potentially available) information that would allow him to uncover a certain order in the given system.

Having introduced the distinction between epistemic and non-

epistemic concepts, we go on to notice that there is a further interesting distinction to be drawn within the class of epistemic notions based upon the kind of role knowledge plays in them. On the one hand there are those epistemic notions in which knowledge enters essentially as an argument in a confirmation function, or, equivalently, as an ingredient in a statement of rational belief. And then there is the obscure and largely unintelligible remainder.

In the first group we find a significant example provided by Hempel's theory of deductive explanation. After introducing his non-epistemic notion of D-N explanation Hempel went on to say that he could define now the concept of a well confirmed D-N explanation, a well confirmed D-N explanation in a tacitly assumed knowledge situation *K* being, in effect, an argument which in knowledge situation *K* it is rational to believe is a D-N explanation, i.e., a true D-N explanation. In precisely the same fashion we could correctly and uninterestingly define the concepts of well-confirmed table, well-confirmed chair and well-confirmed electron, given that we started by having the concepts of table, chair and electron. Since we can only have reason to believe meaningful sentences, a confirmational epistemic predicate is an articulation of independently meaningful components.

Of course we can understand what a well-confirmed chair is because we began by understanding what a chair is. If 'x is a chair' had not had a meaning, we would not even have been able to make sense of the statement of rational belief made about it. Similarly, we can understand, if not appreciate, the notion of well-confirmed D-N explanation, because we were told first what kind of thing a D-N explanation is.³

Now we are in a position to state Hempel's thesis of the epistemic relativity of inductive explanation. As a consequence of his analysis of the phenomenon of ambiguity, Hempel concludes that the concept of inductive explanation, unlike its deductive counterpart, is epistemic; and he goes on to add that it is not epistemic in the sense in which well confirmed deductive explanations are. The concept of inductive explanation is a non-confirmational epistemic concept. Such is the thesis of the epistemic relativity of inductive explanation.

As Hempel is careful to point out, this means that there is no concept that stands to his epistemically relativized notion of inductive explanation as the concept of true D-N explanations stands to that of well confirmed

D-N explanation. According to the thesis of epistemic relativity there is no meaningful notion of true inductive explanation. Hence, we couldn't possibly have reasons to believe that anything is a true inductive explanation. Thus, it would be sheer confusion to see inductive explanations relative to K in Hempel's sense as those inductive arguments which in knowledge situation K is it rational to believe are inductive explanations.

It is then clear that, according to Hempel, there is a remarkable and surprising disanalogy between deductive and inductive explanations. When somebody asks us to give an account of deductive explanations, we can do so without referring to anybody's knowledge. If asked, for instance, what sort of thing would it be to explain deductively the present position of a planet, we would refer to descriptions of certain nomic and non-nomic facts but never to our or to anybody else's knowledge. This is a desirable feature in a non-psychologistic account of explanation. Yet, according to Hempel, when we ask what an inductive explanation of the same event would look like, there is no way in which an appropriate answer can be given without talking about knowledge; presumably, the knowledge available at the time of the explanation. Even more surprisingly, this reference to knowledge does not play the standard role that such references usually play, to wit, that of providing the epistemic platform for a judgment of rational belief. What role such reference plays is a question which deserves serious attention, since here we find the Achilles' heel of Hempel's whole construction.

Let me briefly survey what we have done so far. In effect, we have only done two things. First, we have inspected Hempel's problem, arguing that it isn't clear what precisely the problem is. We contended that in most of the relevant texts Hempel seems to imply that to show that a definition of explanation avoids the problem in question is to show something about the conclusions of such explanations, to wit, that they never contradict each other. But we also argued that some of Hempel's remarks, his blatantly ignoring a most trivial solution to this problem, and the otherwise unintelligible decision to circumscribe solutions to his problem to restrictions to the reference class, suggest that Hempel perceives that the real problem behind ambiguity is a problem having to do with the premises of explanations rather than with their conclusions, and that his problem is in fact, a new variant of the old reference class problem. After arguing for this claim we moved on to inspect the meaning

of a very peculiar conclusion that Hempel feels forced to draw from ambiguity. This is what we have just done. There is now a very natural gap to fill. We should explain what reasons Hempel has to argue that, given the problem of ambiguity, the thesis of the epistemic relativity of inductive explanation follows.

Here we find ourselves in an awkward position. For Hempel has said next to nothing explicitly on the connexion between ambiguity and epistemic relativity, and the little he has said does not carry much weight. Rather than engaging in a frustrating exercise of Hempelian exegesis, I will attempt to reconstruct conjecturally the train of thought that may have led Hempel to his remarkable conclusion.

3. THE ARGUMENT FOR EPISTEMIC RELATIVIZATION

We have seen that Hempel claims that there is no acceptable definition of inductive explanation which is not relativized to K . More precisely he seems to hold that not only is it the case that a definition of true I-S explanation may instantiate ontic ambiguity; he holds that it *must* instantiate it. And in the context of his assumption that a necessary condition for the acceptability of a concept of explanation is that it should not suffer from ambiguity, this implies the non-existence of non-relativized inductive explanations. Let us trace more carefully the form of Hempel's argument. The following assumptions seem to be involved.

Assumption 1. A definition of inductive explanation is inadmissible if it suffers from ambiguity.

Assumption 2. The only way to improve upon the naive definition of true I-S explanation is by introducing in the definition a new clause restricting those reference classes which are admissible in the nomic premise of the explanation.

Assumption 3. There is no clause as the one described in Assumption 2 such that (i) it makes no reference to knowledge, and (ii) when added to the definition of true I-S explanation it guarantees that the resulting definition does not suffer from ambiguity.

Before I try to inspect the sources of these assumptions, let me make a somewhat marginal remark. It seems clear that Hempel believes in the following two further assumptions:

Assumption 4. Inductive explanations relative to K ought to be defined

as inductive arguments of form (II) with premises in K (rather than in the class T of true sentences) which imply the conclusion with high inductive probability, and which verify a requirement delimiting the class of allowable reference classes in the nomic premise. Thus, in purely syntactic terms, a proposed definition of I-S explanation relative to K will differ from a proposed definition of true I-S explanation only in that the class of true sentences T plays in the latter the role that the class of known sentences plays in the former.

Finally, Hempel obviously believes that

Assumption 5. There is a definition of I-S explanation relative to K for which it can be proved that it does not suffer from ambiguity.

Now, the minor point I want to make is that these five assumptions are inconsistent with the claim that the class of true sentences is a possible knowledge situation, or, in other words, with the assumption that it is logically possible that someone could know precisely what is true. The contradiction is obvious since, if one admits that T is one of the classes over which the variable K ranges, assumption 3 affirms a universal of which assumption 5 denies an instance. For assumption 5 says that a certain definition of I-S explanation relative to K (where K is a free variable ranging over knowledge situations) does not suffer from ambiguity, in the sense of assumption 1. But when we replace the variable K by T in the alleged definition of I-S explanation relative to K , what we obtain, in view of assumptions 4 and 2, is a definition of true I-S explanation. Moreover, this definition of true I-S explanation does not suffer from (ontic) ambiguity, for there is, by assumption 5, a proof that, for all K (hence also for T), no instances of inductive inconsistencies obtain. Since, when K becomes T epistemic ambiguity becomes ontic ambiguity, the proof of the avoidance of epistemic ambiguity for some definition of I-S explanation relative to K is ipso facto the proof of the avoidance of ontic ambiguity for the corresponding definition of true I-S explanation.

One might retort that this would hardly affect Hempel since the class of true sentences is not a possible knowledge situation. This would mean that the conception of an omniscient being is logically incoherent and that it couldn't, it *logically* couldn't be the case that someone might happen to know all and only true statements, so that, as a matter of meaning, every knowledge situation should contain at least one falsehood

or lack one truth. This seems unlikely, but even if granted, it should be observed that a proof that a definition of I-S explanation relative to *K* does not suffer from epistemic ambiguity must use some assumption about *K* that is not true about *T*, for otherwise the same proof would establish for the associated concept of true inductive explanation that it avoids ontic ambiguity. Since none of the proofs of the alleged avoidance of epistemic ambiguity offered by Hempel so far make use of any features of *K* that are not true of *T*, had any of them been successful, it would have implied the falsehood of the epistemic relatively thesis.

This is the minor point I wanted to make. Let me now return to the main line of my argument. I was trying to understand the reasons which led Hempel to his claim that there are no true inductive explanations, and I said that the best sense I could make of such a claim was in terms of assumptions 1–3. Now, these assumptions are not implausible. The first one constitutes Hempel's 'official' view on what is wrong about ambiguity, and, no doubt, this claim has a strong intuitive appeal. The second assumption looks similarly plausible, but I have been arguing that its main contention concerning the limits imposed to particular solutions to ambiguity is based upon an understanding of Hempel's problem in conflict with that which inspires the first assumption. For, as I have pointed out, were one to take seriously the idea that all that is wrong with ambiguity is that some definitions of inductive explanation instantiate inconsistencies, then the restriction imposed by this assumption on possible solutions to that problem would be intolerably arbitrary due to the existence of the above-mentioned trivial solution.

Let us turn briefly to the crucial third assumption. It is here that, in spite of his explicit espousal of a propensity interpretation, Hempel's tacit appeal to a frequency conception of probability – with its reference class puzzlements – becomes apparent. Presumably, what could lead to this assumption is an argument like the following.

The source of ontic ambiguity lies in the fact that inductively inconsistent arguments with true premises fix attention upon what is only a partial aspect of the object of the explanation. The problem with our alternative explanations of Jones' financial condition is that one of them pays attention only to his being a Texan whereas the other pays attention only to his being a philosopher. Each alleged explanation is to be blamed on the grounds that it ignores a relevant aspect that the other

takes into account. This is the problem. In order to have a notion of true inductive explanation one must be able to solve it. But, Hempel says, one cannot solve it.

That we cannot solve it seems suggested by the fact that although the 'partial aspects' approach seems excellent to indicate which reference classes will not do, it seems unusable to decide which ones will do. Prima facie, it would seem that what needs to be enforced is a principle demanding that all relevant aspects of the explanandum object be taken into account. The frequentist will explicate this intuitive idea of relevance in terms of the long run convergence of the relative frequencies of the appropriate classes. When so understood, the principle seems to be as strong as to imply that no reference class is admissible, for every reference class seems condemned to be a partial relevant aspect of the explanandum object. The principle tells us that it is not enough to refer Jones to the class of Texans in order to explain his financial condition. But is it enough to refer him to the class of Texan philosophers? In all likelihood the demand to identify all relevant aspects implies that it is not. For it is most likely that there will be some other property of Jones', different from the one to be explained, which when conjoined to that of his being a Texan philosopher will determine a reference class in which the long run frequency of the outcome differs from the one it has in the original class. Unknown to everyone, Jones may have been born at the same time that a Chinese mandarin sneezed, and if the long run frequency of richness in the reference class of Texan philosophers born while a Chinese mandarin sneezes differs (as it well may) from that in the class of Texan philosophers, the demand to refer to all relevant aspects in an explanation rules out the class of Texan philosophers as an admissible reference class. Indeed, every reference class would be ruled out by some other reference class if a certain not unlikely assumption of the denseness of relevant reference classes is accepted: the assumption that given a reference class F and an attribute (explanandum) class G , there is a subclass of F (i.e., a class more specific than F) other than $F \& G$, a class to which the object of the explanation belongs, and in which the long run frequency of the explanandum property G is different from the one it has in F .

As far as I can see, it is something like the above considerations that may have led Hempel to act in agreement with assumption 3. And it is

this assumption, together with the rather natural assumption 2, which is to be held responsible for the thesis of epistemic relativity.

Only one more thing remains to be done before I can rest my case. I have argued that from a certain problem Hempel felt forced to draw the conclusion that the notion of inductive explanation is epistemically relativized. We have just seen what kinds of assumptions may have been involved in the argument leading to this conclusion. Now, I would like to explain why I find Hempel's conclusion worth avoiding. I will try to convince you that to accept Hempel's thesis of epistemic relativity amounts to accepting the claim that there are no inductive explanations, the concept of I-S explanation relative to *K* functioning as a placebo which can only calm the intellectual anxieties of the uncautious user. If I am right, anyone willing to hold that there are inductive explanations will have to begin by spotting a flaw in Hempel's argument. I will close this paper with a few remarks in which I will attempt to indicate in the barest outline how, by denying some of the assumptions on which Hempel's conclusion seems to rest, one can conceivably avoid epistemic relativity and introduce a satisfactory characterization of inductive explanation.

4. EPISTEMIC RELATIVITY REVISITED

Maybe the best way in which I can briefly convey my feelings about the oddity implicit in Hempel's theory of inductive explanation, is by noting, that in my view, Hempel's decision to develop a theory of I-S explanation relative to *K* after having argued that the notion of true inductive explanation makes no sense, seems comparable to that of a man who establishes conclusively that Hegel's philosophy is strict nonsense, and then proceeds to devote the rest of his life to produce the definitive edition of Hegel's writings. For I would like to suggest that the only purpose that could be served by the predicate 'being an inductive explanation relative to *K*' is that of identifying a class of inductive arguments, the respectability of which has been seriously undermined by Hempel's analysis of ambiguity.

Let me remind you the general form of Hempel's characterization of I-S explanation relative to *K*. As our fourth assumption indicates, Hempel views I-S explanations relative to a knowledge situation *K* as arguments of form (II) with premises in *K*, which imply the conclusion

with high inductive probability; furthermore, the probabilistic premise is nomic and its reference class complies with a certain requirement of maximal specificity. The heart of the definition is this last requirement which is supposed to work roughly in the following way. Pick a knowledge situation K ; now, a class to which the object of the explanation is known (in K) to belong will be an admissible reference class (relative to that knowledge situation) if it is the most specific class to which the object is known to belong (disregarding the explanandum property), or if it is a wider class but the probability of the explanandum property in all more specific classes to which the object of the explanation is known to belong, are known to be equal to that in the wider class.

Let me exemplify with our Texan philosopher, Jones. Suppose we know that he is a Texan and a philosopher, and suppose we know nothing other than the logical consequences of these facts. Then Hempel's principle says that we can explain his financial condition by referring him to the most specific class to which he is known to belong, that of Texan philosophers. And the principle adds that we could also refer him to the class of Texans, or to that of philosophers, provided that the probability of richness in the given class were equal to that in the class of Texan philosophers.

No doubt, you will have recognized in this description of the maximal specificity principle an epistemologized version of the demand to refer to all relevant aspects of the explanandum. By relativizing to knowledge that 'ontic' demand, the requirement becomes tractable, due to the fact that knowledge, or at any rate the instances of it that we are familiar with, is notable by how few traits of reality it captures. In effect, then, it is ignorance, rather than knowledge, that makes the maximal specificity principle look like a workable demand. It is because we don't know that Jones was born while that Chinese mandarin sneezed that we can confidently explain his financial condition by referring him to the class of Texan philosophers. Somewhat annoyingly, an increase in knowledge could leave us with no explanation at all. Even worse, if that denseness assumption of relevant classes which we seem to see behind Hempel's rejection of the demand for total relevant aspects were true, then as our knowledge increases, the principle becomes unusable. In the limit, God would find no inductive explanations relative to his knowledge situation: He knows too much.

Now, the question I would like to put to Hempel is the following. Take any I-S explanation relative to K , for some given K . It will be a sequence of formulas like (II). Assume, if you will, that the class K describes our knowledge situation. Now what is there about this inductive argument that makes it an explanation of its last formula? What reason could anyone have to say that it is an explanation of its conclusion?

It is not difficult to answer this question when we pose it, not for the inductive, but for the deductive case. If one asks, for example, what reason do we have to believe that a causal deductive explanation explains its explanandum, the answer is that its premises identify certain features of the world that are nomically responsible for the occurrence of the explanandum event.

Could we say, as in the deductive case, that I-S explanations relative to K explain because their premises somehow identify features of the world that are nomically responsible for the explanandum event? Certainly not. This is what we vaguely conceived to be possible while tacitly espousing the naive model, until Hempel shattered our illusions to pieces by focusing the reference class problem on the theory of explanation. Indeed, if there is no characterization of true inductive explanation, then it must be because there are no things which go on in the non-epistemic world of facts that can inductively explain the event. For if there were such non-epistemic going ons, their characterization would be a characterization of true inductive explanation. Thus, the possibility of a notion of true explanation, inductive or otherwise, is not just a desirable but ultimately dispensable feature of a model of explanation: it is the *sine qua non* of its realistic, non-psychologistic inspiration. It is because certain features of the world can be deterministically responsible for others that we can describe a concept of true deductive explanation by simply describing the form of such features. If there are features of the world which can be non-deterministically responsible for others, then we should be able to define a model of true inductive explanation. And, conversely, if we could define a model of true inductive explanation, there could be features of the world non-deterministically responsible for others. The thesis of epistemic relativity implies that, for Hempel, there are no such features. What, then, is the interest of I-S explanations relative to K ? Surely not, as we seen above, that in knowledge situation K we have reason to believe that they are inductive explanations. Then

what? We detect in Hempel's writings not even a hint as to what an answer to this question might be.

As it often happens when someone finds obviously absurd what someone else finds too obvious for defense, it is likely that we are approaching here a zone in which some of my deepest philosophical prejudices interfere with some of Hempel's deepest philosophical principles. Although I believe there is no way in which, consistently with his assumptions, Hempel can satisfy the urge to know why his I-S explanations relative to *K* are supposed to explain, there is a way in which he could overcome my criticism: by simply not taking it as such. He could agree, that is, that no answer can be given to my question, and he could then add that this is precisely the way it should be, for my question is, he might say, a pseudoquestion, the product of a serious philosophical confusion. I-S explanations relative to *K* are those arguments which, in all likelihood, educated people in knowledge situation *K* would take to be inductive explanations. Thus, one might rejoin that my question amounts to a challenge to the right of the scientific community to decide what things are explanations; a challenge raised from a suprascientific standpoint. But, the rejoinder would go on, there is no such standpoint since there are no standards to decide what is an explanation over and above the standards set by the scientific community.

Does it really make no sense to raise the question whether those arguments that the scientific community identifies as, believes to be or calls explanations, really are explanations? Does agreement at this (or any other) level imply truth? It is obvious that space (and other) limitations do not allow me to even start to deal seriously with this very deep question. But I would like to sketch very briefly the reasons I have to suspect it deserves a negative answer.

It is said that Lincoln once raised the question 'If we call a dog's tail a leg, how many legs does it have?', to which he immediately answered 'Four, because calling the tail a leg does not make it a leg'. If we call an inductive argument 'inductive explanation', does that make it an inductive explanation?

Some people tend to believe that when the calling is universal, or almost universal, for users of a language, this quantitative change somehow becomes qualitative and calling, then, does imply being. For them, what the ordinary man says will be philosophically important because what the ordinary man calls things is what everyone, or almost everyone, calls

them; hence (it is claimed) what they ought to be called in the given language; hence, what they are.

Other people tend to apply a more elitist version of the same argument: when the calling is universal, or almost universal in a community or specialists, then, once again, the quantitative change becomes qualitative and the fact that something is generally called an *X* in that community is sufficient ground to accept that this is what such things ought to be called; hence, what they are.

In so far as an argument can be detected for these two views, it would seem that they rely on a certain hypothesis concerning the way in which words receive their meaning. It is sometimes claimed that it would be an impossibility of some kind that there were no material objects because of the way in which the expression 'material object' used in that claim acquires its meaning; presumably, by ostension of objects that are supposed definitionally to qualify as entities to which the expression 'material object' is meant to apply (or in the presence of circumstances in which as a matter of definition, it is correct to apply the expression). Similarly, it is said, there must be things that we know, instances of knowledge, because of the way in which the word 'knowledge' is taught – which would be, at least partly, by the ostension of some paradigmatic cases of what is to count as knowledge.

One problem with this line of argument is that the meaning of expressions is not always given exclusively by ostension or by the application of the expression to actually given circumstances. At least sometimes, one must appeal to abstract conditions for the correct application of the term. We learn (and teach) the meaning of 'knowledge', for example, not only by learning (and teaching) to recognize alleged case of knowledge, but also by relating to that expression additional conceptual information, as, e.g., the information that a sentence is not a case of knowledge if it is not true (if a strong sense of 'knowledge' is meant) or well confirmed on the given evidence (if a weaker sense is intended).

Now, whenever alternative partial characterizations of the meaning of an expression are available, the question of their compatibility arises. If different instructions for the identification of instances of a concept are not compatible, it is not admissible to rely only on the verdicts of one, unless it has been made clear in the process of meaning ascription that such an instruction is to have overriding force.

It follows that the above-mentioned paradigm case argument can only be applied to expressions that receive their meaning exclusively or with overriding force via ostension. But for many philosophically important notions, meanings are assigned via the introduction of non-ostensive characterizations in the form of conditions to be satisfied by instances of the notion. Furthermore, it is often the case that such non-ostensional conditions tend to have overriding force, since we are taught to correct our judgments whenever they are not satisfied. We are taught to correct a knowledge claim, i.e., to say that what we called knowledge wasn't really such, when we find out that the statement claimed to be known was not true (in the strong sense) or was poorly confirmed on the available evidence (now in the weak sense). Hence, from the fact that we are taught (or teach) to call certain objects instances of knowledge, it does not follow that there is anything that we know.

Now, I would argue that the sense or senses of 'explanation' that are of interest in the philosophy of science do not receive their meaning exclusively or primarily by ostension but also by the specification of conditions that instances of such concepts ought to verify. Thus, from the fact that ordinary men or ordinary scientists are willing to call certain arguments 'explanations' in that sense, it will not follow that the things so-called happen to be explanations in that sense; nor even that there are any such explanations.

Thus, I remain convinced of the legitimacy of the question I have posed, and I conclude that the lack of an answer to it counts heavily against the significance of Hempel's theory of inductive explanation. I have just argued that Hempel's epistemic relativization is, if possible, worth avoiding. My search for the Hempelian assumptions that led to this thesis may now be seen as the search for the 'causes' of this intellectual malady. Maybe by removing some of them, we can also remove their 'effects'. In the few remaining paragraphs I would like to give a very rough idea of what I conceive to be an appropriate theory of inductive explanation.

5. SPECULATION

I side with Hempel's somewhat tacit belief that a good theory of inductive explanation ought to begin with an analysis of the reference class

problem. Thus, I accept the contention that, at present, the problem of defining inductive explanation is, essentially, the problem of identifying an appropriate requirement on admissible reference classes. I further agree with the demand (implicit in Hempel, explicit in Salmon⁴) that such requirements ought to be an explication of the demand to identify all *relevant* aspects of the explanandum. But I want to question Hempel's conclusion that this requirement is undefinable outside the limited framework provided by a human knowledge situation. Such conclusion seemed to rely upon the assumption that the only way to determine the relevance of a predicate is by determining its actual frequentist correlation with the explanandum predicate. But one might try to characterize the 'relevance' in those relevant aspects in a different way: not as statistical relevance but as nomic relevance, a predicate being nomically relevant to another when a law of nature determines that changes in the first one generate changes in the second one.

Perhaps I could make my plea for nomic relevance more appealing if I could explain why I feel that statistical relevance is neither necessary nor sufficient to determine whether a reference class is the 'appropriate' one.

Consider first the famous bellic episode described by Scriven in a somewhat different context: an atom bomb falls over a bridge and the bridge is destroyed; but it so happens that the cause of the destruction is not the explosion of the atom bomb for, a fraction of a second before the atom bomb explodes, 1000 kgs of dynamite are detonated on the main span of the bridge, causing its collapse. It would now seem that a deductive, indeed, causal explanation can be offered for the fact that the bridge is destroyed at a certain appropriately chosen time t , by appealing to the nomic premise that whenever an atom bomb explodes over a bridge at time t' (shortly before t), at time t the bridge is destroyed. In effect, we would be attempting to explain our explanandum deductively by referring the bridge to the class of bridges on which atom bombs fall a fraction of a second before their destruction. But it seems quite natural to contend that this alleged explanation is no explanation at all. The right explanation is provided by referring the bridge in question to the class of bridges on which 1000 kgs of dynamite are detonated. And this is so even though this reference class provides a statistically irrelevant partition for the property 'being destroyed a few seconds later' in the reference class defined in terms of the atom bomb explosion. Notice that the reason we

give for this choice is that the dynamite explosion was the nomicallly operant feature in the envisaged situation.

It is also possible to argue that statistical relevance is not sufficient in order that a reference class be preferred over another one. Thus, we should not care whether Jones was born while our Chinese mandarin sneezed because, even though it may well be that the class of persons born while a Chinese mandarin sneezed is statistically relevant to richness in the class of Texan philosophers, there is, in all likelihood, no natural law that correlates Texan philosophers born in such circumstances with the amount of money they possess. The only kind of property that could 'screen-off' that of being a Texan philosopher as inappropriate to explain Jones' financial condition is a property true of Jones (at the appropriate time) and such that a law of nature determines the nomic relevance of its conjunction with the screened-off property to Jones financial condition.

Thus, when nomicity and statistical relevance enter into conflict, it seems clear that nomicity always wins the day. They seldom enter into conflict, for statistical relevance is the evidence that we may have for the presence of nomicity. But a model of explanation, i.e., an account of what explanations are – rather than an account of what counts as evidence for their presence – should contain a reference to the explanatory features rather than to the symptoms of their presence. In this way, one may be led to speculate that to explain a single event is to refer the object of the explanation to its most specific property relative to the explanandum outcome. A property P will be the most specific property of an object relative to an outcome property Q whenever there is no other property R instantiated by the object of the explanation (during the appropriate time interval) such that the property $P \& R$ is nomicallly related to Q . There is some reason to suspect that the obscure notion of nomic relation can be clarified in a non *ad-hocish* way, consistently with the above speculation. But this is a subject for some other occasion.

Let me conclude with a summary of what I have attempted to do. I have first tried to explain the nature of what I take to be a most remarkable and unexpected development of Hempel's theory of explanation, his thesis of the epistemic relativity of inductive explanation. I have tried to understand Hempel's reasons for this claim, which seem to stem from his discovery of the problem of ambiguity. But, as we saw, Hempel's

identification of the nature of this problem was, itself, ambiguous. I have argued that the real difficulty is just the old reference class problem in a new guise. Then I contended that to accept Hempel's thesis is, in effect, to deny the existence of inductive explanations. If such consequence is to be avoided, some of Hempel's assumptions ought to be rejected. In agreement with my contention that Hempel's problem is the reference class problem, I held that one must revise his requirement of maximal specificity, reformulating it in ontic, rather than in epistemic terms. It is not obvious that this can be done; but it can be done if sense can be made of a certain appeal to nomicity related to that contemporary *inintelligibile* the propensity interpretation of probability. I can't possibly expect to have made this last claim plausible; but I would be satisfied if I had convinced you that there is a problem, maybe an interesting problem, where you thought there was none.

Dept. of History and Philosophy of Science, Indiana Univ., Bloomington, Ind., U.S.A.

NOTES

¹ This paper was read at the Lecture Series organized by the Center for the Philosophy of Science at the University of Pittsburgh. I gratefully acknowledge comments by Professors A. Grünbaum, W. C. Salmon, G. Massey, N. D. Belnap, Jr., and T. M. Simpson.

² It is assumed, without warrant, that for any *K* and any '*Gi*', there is a maximally specific predicate related to '*Gi*' in *K*. (It is possible to construct consistent and logically closed *K*'s for which this assumption is false.)

³ Prof. Gerald Massey has drawn my attention towards the apparent opacity of epistemically relativized predicates. He has pointed out that this raises serious doubts concerning the possibility of viewing them as expressing properties.

⁴ Particularly in Salmon (1970).

BIBLIOGRAPHY

- Hempel, Carl G., 'Deductive Nomological vs Statistical Explanation', in *Minnesota Studies in the Philosophy of Science* (ed. by H. Feigl and G. Maxwell), Univ. of Minnesota Press, Minneapolis, 1962, pp. 98-169.
- Hempel, Carl G., *Aspects of Scientific Explanation*, The Free Press, New York, 1965.
- Hempel, Carl G., 'Maximal Specificity and Lawlikeness in Probabilistic Explanation', *Philosophy of Science* 35 (1968), 116-134.
- Salmon, Wesley C., 'Statistical Explanation', in *Statistical Explanation and Statistical Relevance* (ed. by W. C. Salmon), Univ. of Pittsburgh Press, Pittsburgh, 1970, pp. 29-88.