ON THE APPLICATION OF INDUCTVE LOGIC

1. THE PROBLEM OF APPLICATION

Inductive logic[1] is here understood as a theory based on a definition of the logical concept of probability or degree of confirmation, as distinguished from the frequency concept of probability.[2] It seems justifiable to regard this theory as a kind of logic because, in spite of certain differences, it shows a striking analogy to deductive logic.[3] Its basic concept, the degree of confirmation, is in a certain sense a weak analogue of the concept of logical implication, the basic concept of deductive logic.

The classical theory of probability may be regarded as an attempt towards the construction of a system of inductive logic. However, modern criticism, especially that of Keynes,[4] has shown that the classical theory lacks a solid logical foundation and that some of its principles, if applied without restriction, even lead to contradictions. Modern systems in an incomplete axiomatic form without an explicit definition of the basic concept have been constructed by Keynes, Jeffreys,[5] and others. I have constructed an explicit definition of degree of confirmation (c*) and a theory based on this definition.[6] An alternative definition has been proposed by Helmer, Hempel, and Oppenheim.[7]

In the present paper I intend to discuss some problems which concern, not systematic questions *within* inductive logic, but rather the possibility and the conditions of the *application* of inductive logic to knowledge situations actually given or assumed. There are certain peculiar problems and difficulties involved in the application of *inductive* logic, different from those in the application of *deductive* logic. The chief difficulties which will be discussed here arise out of the following two circumstances. (1)

---

[1] See "On Inductive Logic," *Philosophy of Science*, XII (1945), pp. 72-97.

[2] See "The Two Concepts of Probability," *Philosophy and Phenomenological Research*, V (1945), pp. 513-532.

[3] See §2 of "Remarks on Induction and Truth," *ibid.,* VI (1946). pp. 590-602

[4] J. M. Keynes, *A Treatise on Probability* (1921).

[5] Harold Jeffreys, *Theory of Probability* (1939).

[6] A summary of the theory, stating the definition and some of the theorems, is given in the paper mentioned in footnote 1. The full theory will be developed in a book, *Probability and Induction*, which is in preparation.

[7] Olaf Helmer and Paul Oppenheim, "A Syntactical Definition of Probability and of Degree of Confirmation," *Journal of Symbolic Logic*, X (1945), pp. 25-60; Carl G. Hempel and P. Oppenheim, "A Definition of 'Degree of Confirmation'," *Philos. of Science,* XII (1945), pp. 98-115.

The systems of inductive logic at present available apply only to languages of a certain simple structure and hence to a simplified picture of the universe, while the practical application must be made to our actual, complex world. (2) In order to make it possible for us actually to calculate the value of the degree of confirmation of a given hypothesis with respect to given evidence, this evidence must be relatively simple; on the other hand the observational knowledge actually available to any adult person is enormously comprehensive and complicated.

I shall first explain the structure and interpretation of a language of the kind to which the definitions of degree of confirmation which have been constructed can be applied. Then the requirements for an application will be discussed, in particular the requirement that we must take as evidence the total knowledge available. Finally the problem of which properties are inductively projectible will be discussed. The discussions in this paper will be independent of the technical details of my definition of degree of confirmation; they apply to any definition applicable to languages of this kind.

2. STRUCTURE AND INTERPRETATION OF THE LANGUAGE

A language L of the kind to which my system of inductive logic applies is of a very simple structure, though not quite so limited as that to which deductive logic was restricted for more than two thousand years. It was beyond the latter chiefly in three respects:[5] (1) it allows for multiple quantification ("for every *x* there is a *y* . . ."); (2) it contains not only properties but also relations as primitive concepts; (3) it contains a sign of identity and hence can express cardinal numbers of properties. (In technical terms, L has the structure of lower functional logic with individual variables as the only variables.) The universe of discourse consists of individuals; their number is either finite or denumerably infinite. Every individual is designated by an individual constant. L contains primitive predicates designating properties of individuals and relations between individuals. The language L is to be interpreted as referring not to our actual world but to a simplified universe. The individuals are best regarded not as something like physical bodies but rather as positions (like space-time points in our actual world); hence an individual is not a complex or extended region but a single and indivisible entity.

An atomic sentence consists of a primitive predicate and one or more individual constants. A state-description is defined[9] as a conjunction whose components are some atomic sentences and the negations of all other

[8] The language to which Helmer's definition applies (see footnote 7) has a similar structure but contains only the first of the three features explained in the text.

[9] "On Inductive Logic," p. 73.

atomic sentences. The state-descriptions are meant as descriptions of the possible states of the whole universe. Therefore the following requirement must be fulfilled because, if it is violated, some state-descriptions become self-contradictory.

1. *Requirement of topical independence*. *a*. The *atomic sentences* must be logically independent of each other. (If, for instance, 'A' and 'B' were atomic sentences such that 'A' logically implied 'B', then any state-description containing 'A' and 'non-B' would be self-contradictory.) Consequently the individual constants and the primitive predicates must fulfill the following conditions.

b. *The individual constants* must designate different and entirely separate individuals. (If 'a' and 'b' designated the same individual, 'Pa and not Pb' would be impossible; if 'a' were a part of 'b' or had a part in common with 'b' and if 'H' designated 'hot' in the sense of 'being hot throughout', then 'Hb and not Ha' would be impossible.)

c. The *primitive predicates* must be logically independent of each other, (If '*x* is a raven' is understood as logically implying '*x* is black', then 'raven' and 'black' cannot both be primitive predicates, because 'b is a raven and b is not black' is impossible. 'Warm' and 'warmer' cannot both be primitive predicates, because 'a is warm and b is non-warm and b is warmer than a' is impossible.) (The above formulation applies to groups of two incompatible properties only, e g. P and non-P; for the sake of simplicity let us leave aside for the present discussion groups of more than two incompatible properties, for example a group of the following four properties: 'blue', 'red', 'yellow', 'neither blue nor red nor yellow'.)

The language L is here discussed for the purpose of inductive logic. The construction of a system of inductive logic for any given language presupposes that a system of deductive logic is available for this language. We are laying down requirements which the language L must fulfill so as to make possible the application of inductive logic. Now it is important to realize that some of these requirements are necessary even for deductive logic. This holds, for instance, for the requirement of logical independence. This requirement is, as we have seen, essential for the purpose of the state-descriptions as descriptions of logically possible states. Now the state-descriptions constitute the basis four deductive logic as well; for instance, we may define a logically true (analytic) sentence as one which holds in every state-description, and one sentence is said to imply logically a second sentence if the latter holds in every state-description in which the first holds. (If 'P' and 'Q' were primitive predicates and we were not assured whether or not they fulfilled the requirement of logical independence, then we should be unable to determine whether 'not Pa or Qa' is or is not analytic.)

We may imagine the primitive predicates as designating directly observable qualities or relations. (As examples we may think of something like 'blue', 'warm', 'darker', 'warmer', etc., abstracting from the fact that in our actual world these qualities and relations an observed only in extended regions.) Here a further requirement must be added.

2. *Requirement of simplicity*. The qualities and relations designated by the primitive predicates must not be analyzable into simpler components.

This requirement needs some explanation and discussion. First it means that, if a property is complex, that is, analyzable in terms of simpler properties, then it must not be chosen as primitive; it must rather be analyzed and then expressed by compound expressions. '*x* is a raven' is to be analyzed into '*x* is black and *x* is a bird, etc.' and hence to be expressed by a conjunction; and then 'bird' is to be analyzed further. The property '*x* occurs before or on V-E day and is red, or it occurs later and is non-red' (which will be discussed later) must not be taken as primitive even if there were a simple word for it in English or any other language; it must rather be analyzed an just stated and hence expressed by a disjunction of two conjunctions, some components of which are still compound.

It seems to me that the requirement of analyzing the concepts occurring holds even for deductive logic. One of the chief tasks of deductive logic is to determine whether a given sentence is logically true (analytic). For example, the sentence "all ravens are birds" is logically true, since its truth is based merely on the meanings of the terms occurring, independently of the contingency of facts. However, this deductive result can be established only with the help of the analysis of 'raven' mentioned above. For a deductive problem, for instance, for the question whether logical implication holds between two given sentences, it is necessary to carry the analysis of the terms involved at least to a set of terms which we recognize as logically independent. In practice this is often easy to see; but theoretically it involves quite serious problems. If the affirmative result that one sentence logically implies another or that a given sentence is logically true is established after some steps of analysis, then this result is definitive and the analysis need not go further. On the other hand, the situation with respect to a negative result, for instance, that two given sentences are logically independent or that a given sentence is factual (synthetic) is quite different. If at a certain stage in the analysis no logical dependence is found, the negative result may appear plausible and for practical purposes we may accept it as established. However, there is always the possibility, at least theoretically, that further analysis will reveal a logical dependence. Therefore, strictly speaking, the negative result is not definitively established until we drive the analysis down as far as possible. The properties or relations to which the analysis finally leads, those which cannot be analyzed further, we call *simple* properties or relations. The important fact is that even deductive logic requires, at least theoretically, an analysis into simple components.

It must be admitted that an exact explication of the concept of simplicity cannot easily be given. Nevertheless, the concept seems clear enough for many practical purposes because for many properties or relations it is quite clear that they are not simple but analyzable into simpler concepts; this holds, for example, for 'raven', 'house', 'milk', 'occurring before or on V-E day and being red, or occurring later and being non-red', and still more for 'brother', 'electrically charged', 'schizophrenic', and the like; the complexity of any of these latter concepts is obvious because a great number of observations is required for establishing the concept in a given instance. It cannot be denied that the question whether a given property is simple or not often involves serious difficulties. One difficulty consists in the fact that a property may be directly observable and yet analyzable, hence not simple. For example, gestalt-psychologists have pointed out that we recognize a thing as a dog immediately, that is, without first perceiving some separate items and then applying a process of reasoning resulting in the knowledge that the complex consisting of these items is a dog. Nevertheless we can analyze our recognition of a dog into simpler components, and we do so if we or somebody else has some doubt about it. Therefore the property of being a dog is not simple. On the other hand, a certain shade of blue is a simple property. We cannot analyze it into simpler components. (The spectral analysis of this blue into spectral colors as its components or the physical analysis of it in terms of electro-magnetic waves are, of course, not analyses in the present sense; they do not show the experience to be composite but rather establish, by way of induction, certain correlations between this color blue and other experiences.)

Now it seems to me that the situation in inductive logic with respect to the requirement of simplicity is fundamentally the same as in deductive logic but practically more difficult. I believe that for most of the problems in inductive logic it is necessary to carry the analysis to the end. It is true that this requirement of complete analysis involves great difficulties for the application of inductive logic to our actual world. It should, however, be kept in mind that these difficulties for inductive logic, although more complicated than those for deductive logic, are fundamentally not of a different kind. Both deductive and inductive logic, if their problems are to be soluble in an exact way, must be applied to a simplified universe. We may then presuppose that some simple, directly observable properties occurring in this universe have been taken as primitive for our language L.

The following requirement is perhaps necessary for inductive logic, though it is certainly not for deductive logic.

3. *Requirement of completeness*. The set of primitive predicates in the language L must be complete in the following sense. Every qualitative property or relation of the individuals, that is, every respect in which two positions in the universe may be found to differ by direct observation, must be expressible in L. This requirement is presupposed in my system of inductive logic; it is at present not yet quite dear whether it is essential for all forms of inductive logic.

On the basis of the primitive predicates, with the help of the other signs of L, other predicates may be introduced by definitions. All properties expressible by statement-forms in L (and hence designatable by predicates, primitive or defined) may then be classified into the following three kinds.

(1) *Purely qualitative properties*; they can be expressed without the use of individual constants, but not without primitive predicates. Examples: 'blue', 'non-blue', 'blue or non-warm'.

(2) *Purely positional properties*: they can be expressed without using primitive predicates. Examples: 'being the position a_{28}' (i.e. '$x = a_{28}$'), 'being neither a_{28} nor a_{30}' (i.e. '$x \neq a_{28}$ and $x \neq a_{30}$').

(3) *Mixed properties*; they do not belong to the kinds (1) or (2); hence every expression for them contains a primitive predicate and an individual constant. Examples: 'being red and not being a_{100}' (i.e. 'x is red and $x \neq a_{100}$).

It follows from these definitions that every full sentence of a purely qualitative property is factual (synthetic, neither logically true nor logically false) and every full sentence of a purely positional property is L-determinate (i.e., either logically true or logically false). (For this reason I have elsewhere called the purely positioned properties L-determinate properties.) (Note that 'x is red or x is not red' designates not a purely qualitative property but a purely positional property though a trivial one, because it is logically equivalent to '$x = x$'.)

3. THE PRINCIPLE OF TOTAL EVIDENCE

Let $c(h, e)$ be the degree of confirmation of the hypothesis h with respect to the evidence e. Let us suppose we have a definition of the function c and, based upon this definition, a theorem '$c\ (h, e) = q$', which states the value q of c for given h and e. A principle which seems generally recognized,[10] although not always obeyed, says that if we wish to apply such a theorem of the theory of probability to a given knowledge situation, then we have to take as evidence e the *total evidence* available to the person in question at the time in question, that is to say, his total knowledge of the

[10] Keynes, *op. cit*., p. 313, refers to "Bernoulli's maxim, that is reckoning a probability, we must take into account all the information which we have".

results of his observations.[11] An additional item of evidence *i* may be omitted only if it can be shown that this omission does not change the value of *c*, in other words, that c (*h*, *e* and *i*) = c (*h*, *e*). If this condition is fulfilled, *i* is said to be *irrelevant* for *h* with respect to *e*. In practice *i* is often omitted if its irrelevance is plausible though not actually proved. But if there is any doubt as to the irrelevance of *i*, it ought to be retained. And if, moreover, it is recognized that *i* is relevant, that is to say, that c(*h*, *e* and *i*) ≠ c(*h, e*), then it would obviously be wrong to omit *i*. Suppose, for example, that an observer *X* tells us that he has observed some inhabitants of Chicago with respect to the color of their hair and that he is interested in the degree of confirmation of the prediction *h* that a certain person who has not yet been observed and of whom nothing is known except that he is a Chicagoan, has red hair. *X* shows a list of 40 inhabitants of whom 20 are marked as red-haired and 20 as non-red-haired; *X* tells us that he has found these facts by observation; let *e* be this report. *X* tells us that according to his definition of *c* (in agreement with certain customary conceptions) the value of c in cases of this kind is equal to the relative frequency in the observed sample; hence c (*h*, *e*) = ½. He asks us whether he may apply this result in the sense that he may regard the probability of *h* for his situation at the present time as ½; and consequently, whether it would be reasonable for him to bet on *h* even odds. We ask him whether the list *e* concerning the 40 inhabitants represents all his observations of inhabitants. He replies, no, he had observed altogether 400 persons, but he was not very much interested in the other list *i* concerning the remaining 360 persons because all of them had been found not to have red hair. I believe that everybody would tell *X* that he committed a serious mistake by omitting an obviously relevant part of the evidence he had; and that he ought to take as the probability not c(*h*, *e*) but c(*h, e* and *i*) (thus not ½ but 1/20).

In a recent discussion note,[12] Nelson Goodman points out that there are properties which can be projected by induction from past observations to future events, and that there are other properties which are not thus projectible. He raises the question as to which properties are inductively projectible. And he tries to show with the help of two examples that both my definition of degree of confirmation (see footnote 1) and that by Helmer (see footnote 7) neglect this distinction and thereby yield implausible values in certain cases. The question of projectibility raised by Goodman is an important problem; I shall discuss it in the last section of this paper. At the present moment I want merely to point out that Goodman's reason-

[11] See "Remarks" (footnote 3), p. 594, the explanation of I6.

[12] Nelson Goodman. "A Query on Confirmation," *Journal of Philosophy*, XLIII (1946), pp. 383-385.

ing with respect to his two examples is not correct because he violates the principle of the total evidence in both of them. In the first example, ‘S*x*’ is meant as ‘*x* is a marble drawn before or on V-E day and is red, or is drawn later and in non-red’. The observational results include the fact that the marble a_1 was drawn ninety-eight days before V-E day and was red. Instead of this known fact, Goodman formulates the evidence concerning a_1 simply by ‘Sa_1’. This sentence, however, says merely that a_1 was drawn before or on V-E day and was red, or was drawn later and was non-red; it is true that this follows from the fact mentioned, but it is obviously lees than is known about a_1. In the second example, ninety-six tosses of balls have been observed; red and non-red balls occur in this sequence in a certain regular pattern. But then Goodman formulates the evidence simply by a conjunction of sentences which say which balls were red and which were non-red; thus the formulation omits the temporal regularity although it is obviously quite essential.

How can the temporal order of events be described? There are chiefly two alternative procedures. They will now be explained; the second seems to me more adequate. The first procedure consists in using a primitive predicate, say ‘R’ such that ‘R*xy*’ means that *x* is earlier than *y*; ‘*x* immediately precedes *y*’ can then be expressed by ‘R*xy* and there is no *z* such that R*xz* and R*zy*’. This procedure would presuppose that temporal relations are regarded as qualitative. This conception seems to me rather dubious; it would have the effect that, for instance, the asymmetry of R m merely contingent. The second procedure is possible only in a language which is somewhat stronger than the language L used in my system of inductive logic. It must be not a name-language, like L, but a coordinate-language;[13] that is to say, the forms of the individual expressions indicate their positional relations. In the simplest case, an accent ‘′’ is used; x' is the position immediately succeeding x. With the help of the accent, relations like earlier or later can be introduced by recursive definitions.[14] ‘a′′′ is later than a’ is here analytic (in analogy to the statement ‘the fourth of January, 1946, is later than the first of January, 1946’); the same holds for the general statements saying that the relation Earlier is asymmetric, irreflexive, and transitive. While my definition for c* applies only to a name-language, I have also constructed a tentative form of a similar definition for the degree of confirmation with respect to a coordinate-

[13] See *Logical Syntax of Language*, §3. For further discussions concerning the semantical character of coordinate-languages and the expression of positional properties and relations in them see Chapter II of my book *Meaning and Necessity* (1947).

[14] The individual may here be regarded as natural numbers, which function as coordinates for the positions. ‘*x* is later than *y*’ means here the same as ‘$x > y$’; for the definition of this relation on the basis of the accent see *Logical Syntax*, §20, D9.

language of the simple form just indicated. Thin language refers to a universe of discourse whose individuals (positions) exhibit a basic order of a one-dimensional, discrete structure; this order may, for example, be imagined as a temporal order. [This language permits a more adequate formulation of Goodman's property S; taking 'b' for 'V-E day', 'S*x*' is defined by '($x \leq$ b and x is red) or ($x >$ b and x is not red)'.]

As an instance of a case in which the temporal order of events is essential, let us consider again Goodman's second example. There are ninety-six individuals exhibiting the periodical pattern 'a_n is not red, and a_{n+1} is not red, and a_{n+2} is red'. Suppose that we describe this temporal order in the evidence *e* with the help of the second procedure just explained, that is, in a coordinate-language. Let *h* be the prediction 'a_{99} is red'. Then we find (on the basis of the tentative definition mentioned) that the degree of confirmation for *h* with respect to *e* is considerably higher than 1/3, in agreement with Goodman's requirement.

Goodman seems to be aware of the fact that his examples violate the principle of total evidence. However, he says[15] that "it would be fatal to accept" this principle; "if we shall have to express *all* the observed data in our statement of evidence, we shall have to include such particularized information—e. g., the unique date of each toss—that repetition in the future will be impossible." It is not clear what Goodman means by 'repetition' If by a repetition of an event *E* a later event in meant which has *all* the properties of the event *E* including its date, then any repetition is indeed impossible by definition. If, however, we understand, in agreement with customary usage, by a repetition of *E* a later event which is similar to *E* in certain properties not including temporal position, repetition is possible. If the event *E* described in the evidence *e* is complex, that is, consists of a large number of individuals, then a repetition will have only a small probability. (As an instance, take Goodman's example:[15] *e* describes an event *E* consisting of ninety-six consecutive tosses of balls exhibiting a wholly irregular distribution of colors, and *h* is the hypothesis that this distribution will be exactly repeated in the next ninety-six tosses. Since *h* (and likewise *e*) must be formulated not as one atomic sentence but as a conjunction of at least ninety-six atomic sentences, the value of $c(h, e)$, both in my theory and in that of Helmer-Hempel-Oppenheim, would be, not 1 or near to 1, as Goodman seems to think, but rather near to 0.)

4. IDEALIZATION AND APPLICATION

The interpretation of the language L indicated above refers not to the actual world in which we live but to a simplified universe. A system of inductive logic for the actual world, say, for the language of physics, in

[15] *Op. cit.*, in the next to the last paragraph.

which the space-time positions constitute a four-dimensional continuous order and the scale of values of the physical magnitudes a likewise continuous, would have to be much more complicated; the construction of such s system remains a task for the future.

Although our system applies to a simplified world, it is not useless as a basis for inductions concerning the actual world. Physics likewise uses certain simplified, idealized conceptions which would hold strictly only in a fictitious universe, for example those of frictionless movement, an absolutely rigid lever, a perfect pendulum, a mass point, an ideal gas, etc. These concepts are found to be useful, however, because the simple laws stated for these ideal cases hold approximately whenever the ideal conditions are approximately fulfilled. Similarly, there are actual situations which may be regarded as approximately representing the ideal conditions dealt with in the language L.

Suppose, for instance, that spherical bails of equal size are drawn from an urn; the surface of these balls is in general white, but some are marked with a red point, others not; some (without regard to whether they have a red point or not) have a blue point, others have not; and some have a yellow point, others not. A simple inspection does not reveal other differences between the balls. Then we may apply our system to the balls and their observed marks; we take as individuals the balls or rather the events of the appearance of the single balls, abstracting from the fact that the actual balls have distinguishable parts and that the very markings by which we distinguish them are parts of the balls. And we take the three kinds of markings as primitive properties as though they were the only qualitative properties of the balls, abstracting from the fact that a careful inspection of the actual balls would reveal many more properties in which they differ. Suppose we have drawn one hundred balls and found that forty of them had a certain property M, say, that of bearing a red point and a blue point. Suppose that this is all the knowledge we have concerning the balls and that we are interested in the probability of the hypothesis h that the next ball (if and when it appears) will have the property M. Then we shall take the observations of the hundred balls as our evidence e. This is again an idealization of the actual situation because in fact we have, of course, an enormous amount of knowledge concerning other things. We leave this other knowledge i aside because we regard it as plausible that it is not very relevant for h with respect to e, that is to say, that the value of $c(h, e)$, which we can calculate does not differ much from the value of $c(h, e \text{ and } i)$ which ought to be taken but would make the calculation too complicated. (Of course, we may be mistaken in this assumption; that is to say, a closer investigation might show that, in order to come to a sufficient approximation, certain other parts of the available knowledge must be included in the evidence; just as a physicist who assumes that the influence of the friction in a certain case is so small that he may neglect it may find by a closer analysis that its influence is considerable and therefore must be taken into account.) If the temporal order of the hundred ball drawings is known and seems to be relevant (for instance, if the sequence of the colors in their temporal order of appearance shows a high degree of regularity, as in Goodman's second example mentioned above), then we shall include in our evidence the description of this order. If the temporal order of the hundred drawings is not known (for instance, if we counted only the number of each kind without paying attention to the order), or if it is known but assumed to be not very relevant, then we shall take as evidence the conjunction of three hundred sentences each of which says of one of the hundred balls whether or not it has one of the three primitive properties. It will even be sufficient to take as evidence a conjunction of one hundred sentences each of which says of one of the hundred balls whether or not it is *M*. For certain rules of induction or definitions of degree of confirmation, among them the three to be mentioned below, it can be shown that the additional knowledge contained in the three hundred sentences is strictly irrelevant in this case.

Let us suppose that we have decided to take the latter conjunction of one hundred sentences concerning *M* and non-*M* as our evidence *e*. Then a system of inductive logic, although formulated for a simplified universe, may be applied to the actual knowledge situation just described. The application consists in calculating the value of the degree of confirmation *c* for the hypothesis *h* and the evidence *e* specified and taking this value as the probability sought.

5. EXAMPLES OF CALCULATION

We shall now apply several concepts of degree of confirmation to the evidence *e* and the hypothesis *h* in the example described in the preceding section. If we take the concept c* defined in my system, we obtain a value slightly less than 0.4, as we shall presently see. This result may then determine our willingness to bet on the prediction *h* that the next ball will be *M* at odds slightly less than two to three.

Let us see how the value of c* for this case is found. We apply here the theorem T of the singular predictive inference,[16] which says that

(T) $$c^*(h, e) = (s_1 + w_1) / (s + k).$$

s is the number of individuals in the observed sample, in our example 100. $s_1$is the number of individuals in this sample which are *M*, in our case

[16] "On Inductive Logic," §10, Theorem (l).

40. If L contains p primitive predicates, the number k of the strongest properties expressible in L is 2; in our example $p = 3$ and $k = 8$. w_1 is the logical width[17] of M; that is to say, M is a disjunction of w_1 of the k strongest properties; in our example, $w_1 = 2$. Thus we obtain c* = 42/108 = 0.389.

Let us use the theorem T just mentioned for the discussion of another problematic point in inductive logic. T shows that in the case of the singular predictive inference—and the same holds for many other cases—the value of c* for given sentences h and e depends also on the number p of primitive predicates in the language L, even if some of these predicates do not occur in the two sentences. With respect to the simplified universe, this fact does not involve any difficulty, since all properties occurring in this universe must be expressible in L. Therefore, if we can enlarge the language L by adding a new, independent primitive predicate, then L does not satisfy the requirement of completeness and hence cannot be taken as a basis for calculating c*. However, in any application to the actual world, as in the example discussed above, we usually disregard some of the properties actually occurring. Therefore here the dependence of c* on p deserves serious consideration. Thus Ernest Nagel[18] is right in raising the question whether this dependence does not make the definition of c* inadequate. However, his description of the dependence is not quite correct in two points. He says that, when new primitive predicates are added, (1) c* for a pair of given sentences "will in general diminish" and (2) c* can in this way "be reduced to as small a value as one pleases" In fact, however, c* will just as often increase as decrease and it can change is this way only within certain limits. The second point is essential; if Nagel's statement (2) were right, it would indeed constitute a serious objection. Let me explain the situation for the case of a singular predictive inference on the basis of the theorem T stated above, which is also quoted by Nagel. As explained in my paper,[19] T shows that c* is in this case always between w_1/k, the relative width of M, and s_1/s, the relative frequency of M in the observed sample. If the sample is small, c* is near to the former value; as the sample increases (assuming that the relative frequency in the sample remains the same) c* moves slowly towards the latter which it approaches as a limit. Now the addition of new primitive predicates has merely the effect that the movement of c* from the former to the latter value is slower; that is to say, in the enlarged language L′ a larger sample is required to bring c* as close to the latter value as in the

[17] *Op. cit.*, p. 84.
[18] E. Nagel, Review of "On Inductive Logic," *Journal of Symbolic Logic*, XI (1946) pp. 19-23; see point (3) on p. 22.
[19] *Op. cit.*, § 10.

original language L. If one primitive predicate is added, k is doubled but w_1 is likewise doubled, and hence the relative width remains unchanged. In our example, M is a conjunction of two primitive predicates; therefore the relative width of M is always ¼, independent of the number p of primitive predicates. For $p = 3$, we found above c* = 42/108 = 0.389. For $p = 4$, c* = 44/116 = 0.379; for $p = 10$, c* = 296/1124 = 0.263. Thus c* remains always above ¼, for any number of primitive predicates.

If we choose another rule of inductive or definition of degree of confirmation and apply it to the hypothesis h and evidence e of our example, we may find values which differ somewhat from the value of c*. For example, the definition proposed by Helmer (see footnote 7) and, in a certain sense, also Reichenbach's principle of induction[20] take a value equal to s_1/s, the relative frequency of M in the observed sample, thus in our case 0.4. Laplace's rule of succession takes $(s_1 + 1) / (s + 2)$; hence in our case 41/102 = 0.402.

Since there are several definitions of the degree of confirmation c, the question arises: on what basis shall we choose an adequate definition? How shall we judge the adequacy of a proposed definition? The simplest approach is the following. We imagine a knowledge situation and describe it in a sentence c, and further a hypothesis which we formulate by a sentence h. We chose e and h such that (1) they are simple enough to allow the application of the given definition of c, and (2) such that we have an intuitive impression of the value to which customary ways of inductive thinking would lead. Then we examine whether the value of c(h, e) calculated on the basis of the given definition is sufficiently in agreement with the intuitive value. Since the intuitive determination of a value is in general rather than vague, an approximate agreement will be regarded as sufficient. This is the case in the above example for all three calculated values, because customary inductive thinking takes the probability in cases of this kind with sufficiently large samples to be close to the relative frequency observed in the sample. If, however, the calculated value differ considerably from the intuitive one, we shall regard the definition as inadequate in the case in question. It will seldom occur that a proposed definition will generally yield inadequate values. (I have shown[21] that this is the case for Wittgenstein's definition.) More frequently we shall find that a proposed definition furnishes inadequate values only in certain special instances. In this case the definition need not be entirely abandoned; it may be that a suitable modification for it can be found. I believe that this is the case with a weaker form of Laplace's rule of succes-

[20] Hans Reichenbach, *Experience and Prediction* (1938), § 38.
[21] *Op. cit.*, §6.

sion (in its original strong form it leads to contradictions) and with Helmer's definition; this will be discussed at another place.

6. THE PROBLEM OF PROJECTIBILITY

Now we shall discuss Goodman's question as to what kinds of properties are inductively projectible. I think the latter term may be explained as follows. We call a property *W* inductively projectible if the following is always the case: the higher the relative frequency of *W* in an observed sample, the higher is, on this evidence, the probability that a non-observed individual has the property *W*. I assume that Goodman's question was meant in this sense. Then I propose this as a tentative answer: all purely qualitative properties (in the sense explained in §2) are inductively projectible; perhaps only these are; certainly the purely positional properties are not projectible, and I am inclined to believe that the mixed properties are not, but this requires further investigation. It seems that those authors who have formulated inductive rules without specifying the kind of properties to which they are applicable (for example, Laplace's rule of succession and Reichenbach's rule of induction) intended them to be applied to purely qualitative properties only. This was tacitly presupposed; a complete formulation should contain this or a similar restriction. The rules are, of course, meant to be applied not only to simple, directly observable properties but also to certain complex properties. Therefore we must be told to which of them we are permitted to apply them and to which not.

In my system of inductive logic the situation is somewhat different. Stipulations concerning the primitive terms of the language L like those explained earlier (§2) are sufficient. For this system, in distinction to the inductive rules mentioned, it is not necessary to give a classification of complex properties into projectible and nonprojectible properties and then to restrict the inductive procedure to projectible properties. The reason is that the system is based on an explicit definitions for the degree of confirmation c*.[22] By this definition, the value of c*(h, e) for any pair of sentences h and e in L of any form (except that a must not be self-contradictory) is uniquely determined. Now let *W* be a complex property of any kind expressible in L. (According to the requirement of simplicity, *W* is expressed in L not by a primitive predicate but by a compound expression.) We take as evidence each of a series of statements e_0, e_1, e_2, . . . , e_{10}, saying that of one thousand given individuals none, 100, 200, . . . , the entire thousand, are *W*; as h we take the statement that another individual is *W*. Then we apply the definition of c* to these statements. If the value of c*(h, e_n) increases with n, we shall say that *W* is projected by c* is this case.

[22] *Op. cit.*, §6.

If we find that W is projected by c* in all cases of this kind, we shall say that W is projectible by c*. In this way we can investigate the question which properties are projectible by c* and which are not. The answer to this question is determined by the definition of c*, although to prove a general result of this kind as a theorem based on the definition of c* may, of course, be a complicated task. The answer is determined although the definition of c*, in distinction to the complete formulation of the rules of induction mentioned earlier, is not restricted to projectible properties but applicable to all properties. But only for certain properties does the defintion always supply higher values for higher observed frequencies and thereby effect a projection for these properties. According to the presumption mentioned above, this holds for all purely qualitative properties and perhaps only for these. (For example, Goodman's property S is not purely qualitative but mixed because of the occurrence of 'V-E day', which is a name of a temporal position, hence an individual constant; this property S is certainly not projectible by c*.) Although the definition of c* is not restricted, many theorems concern a restricted class of properties. Thus most of the theorems stated in my article,[23] among them the theorem T mentioned above, refer only to factual, elementary properties,[24] that is, in our present terminology, purely qualitative properties. (Therefore these theorems cannot be applied, for instance, to the property S.)

If somebody were to criticize as axiom system of Euclidean geometry because it does not contain a rule specifying for which particular class of triangles the Pythagorean theorem holds, the author of the system might reply: no additional rule is required; the axiom system is complete; take it and discover yourself under what conditions the theorem holds. If anybody misses in my system of inductive logic a rule specifying the particular kind of properties for which inductive projection is permitted, the reply is: no additional rule is required; the definition of degree of confirmation is complete, and is sufficient to determine the kind of property for which projectibility holds.

RUDOLF CARNAP.

University of Chicago

[23] *Op. cit.*, §§9-14.
[24] *Op. cit.*, p. 84.