Notre Dame Journal of Formal Logic
Volume 31, Number 3, Summer 1990

A Fully Logical Inductive Logic

JOHN NOLT*

Abstract Carnap and his successors have explored various *a priori* probability assignments to possible worlds (state descriptions) in an effort to generate plausible inductive probabilities. Such assignments typically incorporate an *a priori* bias in favor of more orderly worlds. This paper presents an alternative approach that abjures such *a priori* favoritism. Instead, inductive probabilities are derived from explicit assumptions about the structure of the actual world. It is shown that even very simple empirical assumptions (such as the hypothesis that there is a specific upper bound on the number of kinds of things) can yield plausible inductive probabilities for a wide range of inferences. The results for these simple assumptions are not, however, satisfactory in all cases; further work may produce better assumptions.

Any workable inductive logic must assume or presuppose that the world is (or is likely to be) fairly orderly. In the inductive logics developed by Carnap and his successors, for example, likely orderliness is presupposed by what amounts to an assignment of varying *a priori* weights to possible worlds (state descriptions) in proportion to their degree of order. This can be done in a variety of ways. In Carnap's systems, for example, one does it by choosing some finite value for a single parameter, λ.[1]

λ determines a relative weighting of two factors in the inductive projection of a property: its observed frequency and its logical width. Carnap calls these the *empirical* and the *logical* factors, respectively ([1], §7). Finite values of λ give some weight to the empirical factor, so that the assumption that a property has been found frequently typically increases the probability that it has further instances. But if $\lambda = \infty$ the empirical factor is disregarded, so that premises about one set of objects have no effect on the probability of conclusions about others.

*Carl Wagner provided valuable help at several points in the development of this paper. Two anonymous referees provided helpful criticisms of an earlier draft. Richard H. Hodson II wrote the PASCAL program that first enabled me to survey a wide range of results quickly.

Received July 11, 1988; revised January 29, 1989

The result is a *fully logical* inductive logic, in the sense that its inductive projection of properties depends on the logical factor alone. Empirical considerations can still influence inductive probabilities, but only if formulated as explicit assumptions of inductive inferences; they are not intrinsic to the logic.

Carnap rejected fully logical inductive logics on the grounds that accepting them "means refusing to give any regard to experience, to the results of observations, in making expectations or estimations" ([1], p. 38). But this opinion, which has since passed into conventional wisdom, is inaccurate.[2]

Granted, in a fully logical inductive logic premises about one set of objects do not affect the probability of a conclusion about others. But expectation can still be made to answer to experience if to the specific premises of an inductive inference we add a general assumption proclaiming the orderliness of the world. For with the addition of such an assumption (which says something about *all* objects), the premises do concern the objects mentioned in the conclusion and so can influence the conclusion's probability.

This idea is not new. The early literature on induction is peppered with attempts to formulate such an assumption—a so-called principle of the uniformity of nature. Many of these were misguided efforts to convert induction to deduction. But the sort of uniformity principle envisioned here would merely enable premises about one set of objects to increase the probability of a conclusion about others; it would not effect a wholesale transformation of induction into deduction.

The idea of formulating uniformity principles remains intriguing, despite its history of failure, because if successful it would provide a rigorous and general account of the sorts of order needed in *the actual world* to make sense of induction. Carnap and his successors constructed sophisticated weightings on *possible worlds* to achieve desired results. But these weightings do not represent anything real; at best they model the subjective predilections of idealized reasoners. While we cannot avoid adopting some *a priori* probability assignment, it may be interesting to see what could be accomplished by limiting consideration to the one that in effect presupposes no likelihood of uniformity at all, Carnap's $\lambda = \infty$. By this austere rejection of *a priori* favoritism toward order, we leave ourselves only one option: to introduce considerations of order exclusively through assumptions that say something about the actual world.

To illustrate this idea, I will describe a fully logical inductive logic for the monadic fragment of the predicate calculus that uses one reasonably interesting class of uniformity principles. While not the only uniformity principles worth considering, these are perhaps the simplest. They assert that the world contains no more than some finite number, k, of kinds of things. (The relevant notion of kind—which depends on a second parameter, π—will be explained shortly.) Lower values of k indicate greater uniformity.[3]

Thus, instead of a choice among values for λ, we have (supposing π to be fixed) a choice among values of k. The chief advantage over Carnapian methods and their generalizations is that since this choice determines an assertion *about the actual world*, it is falsifiable (indeed, finitely falsifiable); the choice of λ or other parameters of generalized Carnapian systems is not.

Induction can now be pursued conservatively by choosing a high value for k, or riskily by making k as low as possible, consistent with current knowledge.

In either case, we must increase the value of k if the current value is ever falsified by observation of more than k kinds. Conversely, if the accumulating data are very uniform, we may (in the Popperian spirit of venturing the strongest hypotheses) be emboldened to lower the value of k.

Of course, the construction of such a logic does nothing to solve Hume's problem. At best what it shows is that, contrary to conventional wisdom, a fully logical inductive logic can give plausible results—if used with the assumption that we live in the right kind of world. The Humean skeptic worries about the truth of that assumption. I will not worry much about it here (though I have a bit more to say below). Like Carnap's insistence on finite values of λ, it is an avatar of induction's inevitable leap of faith.

Echoing Carnap, we define a denumerable hierarchy of languages $\mathcal{L}_\pi^N$ for $N, \pi > 0$. $\mathcal{L}_\pi^N$ is the first-order language consisting of the names $\alpha_1, \alpha_2, \ldots, \alpha_N$, and the monadic predicates $P_1, P_2, \ldots, P_\pi$. Where s and t are variables or names and A is a sentence, '$A(s/t)$' designates the result of replacing each occurrence of t in A by s. Universally quantified formulas will be interpreted substitutionally and thus can be analyzed as (finite or infinite) conjunctions. By an *$N\pi$-state description* we mean a state description for the language $\mathcal{L}_\pi^N$.[4]

The *$N\pi$-range*, $r_{N\pi}(A)$, of a sentence A of $\mathcal{L}_\pi^N$ is the number of $N\pi$-state descriptions in which A is true. It is easily seen that if A is a sentence of $\mathcal{L}_\pi^N$, then for any $i > \pi$, $\mathrm{r}_{Ni}(A) = 2^{N(i-\pi)}\mathrm{r}_{N\pi}(A)$. It follows that the ratio $\mathrm{r}_{N\pi}(A)/2^{\pi N}$ remains constant for all values of π. (It is undefined if A is not in $\mathcal{L}_\pi^N$.) Thus for any sentence A of $\mathcal{L}_\pi^N$ we may set:

$$\mathrm{P}_N(A) = \mathrm{r}_{N\pi}(A)/2^{\pi N}.$$

We retain the relativization to N, since A itself may depend on N if A is universally quantified (i.e., is an N-membered conjunction). For any N, P_N satisfies the Kolmogorov axioms for the probability calculus for sentences of the languages $\mathcal{L}_\pi^N$. P_N is thus a measure of logical probability, relative to the size N of the possible universes of discourse.

The corresponding notion of conditional probability is defined in the usual way:

$$\mathrm{P}_N(C|A) = \mathrm{P}_N(C \,\&\, A)/\mathrm{P}_N(A), \text{ provided that } \mathrm{P}_N(A) \neq 0.$$

Notice that this could be equivalently expressed as:

$$\mathrm{P}_N(C|A) = \mathrm{r}_{N\pi}(C \,\&\, A)/\mathrm{r}_{N\pi}(A), \text{ provided that } \mathrm{r}_{N\pi}(A) \neq 0.$$

$\mathrm{P}_N(C|A)$ is to be regarded as a measure of partial entailment, the probability of a conclusion C, given assumption(s) S. In this paper I will not address the question of its relation, if any, to such notions as degree of rational belief, degree of confirmation, or rational betting quotient.

By a *π-predicate* (or, more simply, a "predicate") we shall mean a sentence of $\mathcal{L}_\pi^N$ that contains no names or quantifiers and has exactly one free variable. Where F is a predicate with free variable x and α is a name, the sentence $F(\alpha/x)$ is called the *ascription* of F to α. One important type of π-predicate is a π-kind. A *π-kind* is a conjunction $(F_1 x \,\&\ldots\&\, F_\pi x)$ where x is a variable and, for each i, F_i is either P_i or $\sim P_i$. π-kinds (or, more simply, "kinds") are what Carnap calls *Q*-predicates. For any π, the number of possible π-kinds is 2^π, which we hereafter write simply as K. Each π-predicate F is equivalent to a unique disjunction of π-kinds, the number of which is the *π-width* of F.

A *πnc-kind statement* ($1 \leq c \leq K$, $c \leq n$) is a conjunction of n ascriptions of π-kinds to n distinct names, in which a total of c kinds are ascribed to these names. Note that each $N\pi$-state description is a πNc-kind statement for some c. We often use the notation '$e_{\pi nc}$' to stand for some arbitrary πnc-kind statement. Whenever such an expression occurs within the scope of the operator 'P_N' we implicitly assume that the n names in $e_{\pi nc}$ are all in $\mathcal{L}_\pi^N$. (Otherwise the resulting expression is ill-defined.) Thus in such contexts we assume that $n \leq N$. Observe that:

(1) For any πnc-kind statement $e_{\pi nc}$, $\mathrm{P}_N(e_{\pi nc}) = 2^{-\pi n}$.

If Q is a kind with free variable x, then we say that Q *occurs in* a kind statement A if $Q(\alpha/x)$ is one of A's conjuncts, for some name α.

If A is a sentence of $\mathcal{L}_\pi^N$ containing no quantifiers and exactly n names, then A is equivalent to a disjunction of mutually exclusive kind statements of the form $e_{\pi nc}$. This disjunction, which we shall call A's *π-normal form*, is constructed as follows. List the $n\pi$ atomic sentences constructible from the n names in A and the π predicates in $\mathcal{L}_\pi^N$ in some standard order at the top of a bivalent truth table. Each horizontal line on the table then represents a unique πnc-kind statement in which an atomic sentence occurs negated or unnegated according to whether it is T or F on the line. A's π-normal form is just the disjunction (in order from the top to the bottom of the table) of each statement represented by a line on which A itself is T.

We now introduce the concept of a principle of the uniformity of nature. The *U*-principle* $\mathrm{U}^*_{N\pi k}$ for $\mathrm{N}, \pi > 0$ and $0 \leq k \leq K$, is the assertion that there are exactly k π-kinds. More specifically, it is the disjunction of all $\mathrm{N}\pi$-state descriptions that are πNk-kind statements. U*-principles have sometimes been called "consituent-structures" or "structural hypotheses" (see, e.g., Kuipers [9], p. 257). Notice that $\mathrm{U}^*_{N\pi k}$ is the null disjunction (a contradiction) for $k > N$ or $k = 0$.

The *U-principle* $\mathrm{U}_{N\pi k}$ is the assertion that there are at most k π-kinds. More precisely, $\mathrm{U}_{N\pi k} = \mathrm{U}^*_{N\pi 1} \vee \ldots \vee \mathrm{U}^*_{N\pi k}$ for $0 \leq k \leq K$. If $k = K$, $\mathrm{U}_{N\pi k}$ is tautologous, since each $N\pi$-state description is a disjunct of one of its disjuncts, and in each $M\pi$-world, $M \geq N$, an $N\pi$-state description is true. Moreover, for any sentence A of $\mathcal{L}_\pi^N$:

(2) $\mathrm{P}_N(\mathrm{U}_{N\pi k}|A) = \sum_{i=1}^{k} \mathrm{P}_N(\mathrm{U}^*_{N\pi i}|A)$

since for $k \neq j$, $\mathrm{U}^*_{N\pi k}$ and $\mathrm{U}^*_{N\pi j}$ are mutually exclusive.

A note on notation: we shall consider $\mathrm{U}_{N\pi k}$ or $\mathrm{U}^*_{N\pi k}$ only in the context of probabilities of the form P_N. Thus in expressions such as '$\mathrm{P}_N(\mathrm{U}_{N\pi k}|\mathrm{A})$' the double occurrence of 'N' is redundant; we shall henceforth abbreviate this to $\mathrm{P}_N(\mathrm{U}_{\pi k}|\mathrm{A})$', and likewise for other similarly redundant expressions.

As explained above, we regard $\mathrm{P}_N(C|A)$ as the inductive probability of conclusion C given assumption(s) A. However, as noted at the outset, our logic will be unable to "learn from experience" unless we supplement the premises with a uniformity principle. Thus we are interested mainly in probabilities of the form $\mathrm{P}_N(C|A \mathbin{\&} \mathrm{U}_{\pi k})$ – that is, $\mathrm{P}_N(C|A \mathbin{\&} \mathrm{U}_{N\pi k})$ – where $\mathrm{U}_{N\pi k}$ is some appropriate U-principle. It remains to say which U-principles are appropriate.

U-principles are determined by three parameters: π, k, and N. π specifies a degree of "coarseness" for kinds, while k limits the number of kinds of that coarseness. N is the number of objects in the world. Our interest will focus, not on specific values of N, but on $\lim_{N\to\infty} \mathrm{P}_N(C|A \ \& \ \mathrm{U}_{N\pi k})$. This may be regarded as an approximation of $\mathrm{P}_N(C|A \ \& \ \mathrm{U}_{N\pi k})$ for large finite values of N, and also as the probability of C, given both A and the assumption that at most k π-kinds are instantiated, for countably infinite worlds.

It is possible, though at the cost of uninteresting complication, to consider $\mathrm{P}_N(C|A \ \& \ \mathrm{U}_{N\pi k})$ where either C or A contains predicates or names not mentioned in $\mathrm{U}_{N\pi k}$. We shall not do so. Rather, we shall regard the choice of a U-principle as determining the choice of language, so that assuming $\mathrm{U}_{N\pi k}$ means adopting $\mathcal{L}_\pi^N$, to which any C and A we wish to consider must therefore belong. Conversely, given specific C and A, we may vary π and k in $\mathrm{P}_N(C|A \ \& \ \mathrm{U}_{N\pi k})$ freely, so long as π does not get so small as to exclude predicates of C and A. Thus $\mathrm{U}_{N\pi k}$ may (and typically will) contain more predicates than occur in C and A. Intuitively, these "extra" predicates represent (familiar or unfamiliar) properties whose presence or absence might be a factor in regularities involving the predicates of C and A.

Though we may vary the value of k freely, the interesting values are limited to a fairly specific range. Clearly we want $k < K$. For if $k = K$, $\mathrm{U}_{N\pi k}$ is tautologous and so adds nothing to the original assumptions. Moreover, on any interpretation of a language $\mathcal{L}_\pi^N$, if no stronger U-principle than $\mathrm{U}_{N\pi K}$ is true then there are no true nonlogical laws; every true universally quantified sentence is a tautology. (For every nontautologous universal generalization implies that at least one kind is uninstantiated.) Inductive reasoners certainly take it for granted that some nonlogical laws hold. Hence it does no great violence to their method to represent them as assuming that k takes some value less than K. (Justification of this assumption is something else again, but that is not our concern.)

At the other end of the scale, $\mathrm{U}_{N\pi k}$ is reasonable only if $k > 1$. For $\mathrm{U}_{N\pi 1}$ implies that for any π-predicate F and name α, if $F\alpha$ is true then so is $(x)Fx$. Not only is this unreasonable, for any interestingly interpreted language it is false. Indeed, in general $\mathrm{U}_{N\pi k}$ is plausible only if k is considerably greater than 1, since typically we know that the universe contains many kinds of things. Thus $\mathrm{U}_{N\pi k}$ is plausible only if $k < K$ and k is at least as large as the minimum number of π-kinds compatible with current knowledge.

Of course, inductive reasoners do not actually think in these terms, and it would certainly be gratuitous to suppose that they ever had some specific value of k in mind. They rely, rather, on a vague sense that the world is not too chaotic. U-principles are idealized representations of that vague sense.

Before examining the details of the logic just described, it may be useful to compare it with the axiomatizations of generalized Carnapian systems developed by Hintikka, Niiniluoto, and Kuipers (see [4], [6], and [9]–[12]). Direct comparison is not possible, since none of these writers employs U-principles as assumptions in the way outlined above. But indirect comparison can be achieved by treating a U-principle $\mathrm{U}_{N\pi k}$ not as an assumption, but as the *a priori* stipulation that $N\pi$-state descriptions containing more than k kinds have probability zero. Probabilities are then distributed evenly among the remaining $N\pi$-state descriptions. This yields a new probability operator of the form $\mathrm{P}_{N\pi k}$ such that

$P_{N\pi k}(C|A) = P_N(C|A \ \& \ U_{N\pi k})$, and one may inquire about the place of $P_{N\pi k}$ (or, more precisely, its value in the limit as $N \to \infty$) in the Hintikka–Niiniluoto–Kuipers systems.

On such a reconstrual $P_{N\pi k}(e_{\pi nc}) = 0$ for $c > k$; that is, it is regarded as impossible that more than k kinds should appear in a string of n ascriptions of kinds to individuals—because according to $P_{N\pi k}$ no more than k kinds can exist. This makes $P_{N\pi k}$ nonregular in Kuipers's sense.[5] The attention of Carnap and his successors has been focused primarily on regular systems; in particular, the nonregular systems defined by $P_{N\pi k}$ have not to my knowledge been singled out for extensive study. One reason for this inattention is obvious: nonregular systems entail certain finite *empirical* propositions. Carnapians, who wish to map the limits of plausible *a priori* probability distributions, generally eschew such empirical entanglements.

Things appear quite differently, however, if our aim is to explain as much of induction as possible in terms of tentative but categorical assumptions about the structure of the actual world. To that task we now return.

We shall consider the performance of the fully logical inductive logic described above in its application to four common kinds of inductive reasoning: simple induction, inductive generalization, singular predictive inference,[6] and analogical inference.

A *simple induction* is an inference from an assumption of the form $Fa_1 \ \& \dots \& \ Fa_n$ to a conclusion of the form Fa_{n+1} for some new name a_{n+1}. We must therefore characterize the function $P_N(Fa_{n+1}|Fa_1 \ \& \dots \& \ Fa_n \ \& \ U_{\pi k})$. The probability calculus allows us to re-express this as $P_N(Fa_1 \ \& \dots \& \ Fa_{n+1} |U_{\pi k})/P_N(Fa_1 \ \& \dots \& \ Fa_n|U_{\pi k})$; so it suffices to describe the function $P_N(Fa_1 \ \& \dots \& \ Fa_n|U_{\pi k})$ for any value of n.

Let w be the π-width of F. Then the π-normal form $E = \vee_i E_i$ of $Fa_1 \ \& \dots \& \ Fa_n$ is a disjunction of mutually exclusive πnc-kind statements for $1 \leq c \leq w$. So $P_N(Fa_1 \ \& \dots \& \ Fa_n|U_{\pi k}) = \sum_i P_N(E_i|U_{\pi k})$. Now we may think of the disjuncts of E simply as n-term sequences (with replacement) of the w kinds in F. The number of disjuncts containing exactly c kinds is equal to the number of n-term sequences (with replacement) of a w-membered set containing exactly c members of that set. Let us designate this quantity as $\text{seq}(c,n,w)$. Thus the number of disjuncts of the form $e_{\pi nc}$ in E is $\text{seq}(c,n,w)$. We shall see below (Corollary 1.1) that if A and B are both πnc-kind statements, then $P_N(A|U_{\pi k}) = P_N(B|U_{\pi k})$. Therefore $P_N(Fa_1 \ \& \dots \& \ Fa_n|U_{\pi k}) = \Sigma_{c=1}^{w} \text{seq}(c,n,w) \, P_N(e_{\pi nc}| U_{\pi k})$. But clearly $P_N(e_{\pi nc}|U_{\pi k}) = 0$ if $c > k$. So:

(3) $P_N(Fa_1 \ \& \dots \& \ Fa_n|U_{\pi k}) = \sum_{c=1}^{\min(k,w)} \text{seq}(c,n,w) \, P_N(e_{\pi nc}|U_{\pi k})$.

Hence:

(4) $P_N(Fa_{n+1}|Fa_1 \ \& \dots \& \ Fa_n \ \& \ U_{\pi k})$

$$= \frac{\sum_{c=1}^{\min(k,w)} \text{seq}(c,n+1,w) \, P_N(e_{\pi(n+1)c}|U_{\pi k})}{\sum_{c=1}^{\min(k,w)} \text{seq}(c,n,w) \, P_N(e_{\pi nc}|U_{\pi k})}.$$

Since $P_N(e_{\pi nc}|U_{\pi k}) = P_N(e_{\pi nc})P_N(U_{\pi k}|e_{\pi nc})/P_N(U_{\pi k})$, by (1) this may also be expressed as:

$$(5)\quad \frac{\sum_{c=1}^{\min(k,w)} \text{seq}(c, n+1, w)\ \text{P}_N(\text{U}_{\pi k} | e_{\pi(n+1)c})}{K \sum_{c=1}^{\min(k,w)} \text{seq}(c,n,w)\ \text{P}_N(\text{U}_{\pi k} | e_{\pi nc})}.$$

Our task is now reduced to describing $\text{seq}(c,n,w)$ and either $\text{P}_N(\text{U}_{\pi k} | e_{\pi nc})$ or $\text{P}_N(e_{\pi nc} | \text{U}_{\pi k})$. When no ambiguity threatens, we shall often omit mention of the parameter π. Thus, for example, '$\text{P}_N(\text{U}_{\pi k} | e_{\pi nc})$' becomes simply '$\text{P}_N(\text{U}_k | e_{nc})$'; and likewise for similar expressions.

The function seq is definable by induction on c and n.[7] For all nonnegative integers c, n, w:

(S1) $\text{seq}(c,n,w) = 1$, if $c = n = 0$
$= 0$, if $c = 0$ or $n = 0$ and $c \neq n$
$\text{seq}(c+1, n+1, w) = (w-c)\ \text{seq}(c,n,w) + (c+1)\text{seq}(c+1,n,w)$.

Some important consequences of this definition are stated below without proof:

(S2) If $c > n$ or $c > w$, then $\text{seq}(c,n,w) = 0$

(S3) $\text{seq}(c,n,w) = \binom{w}{c} \text{seq}(c,n,c)$

(S4) $\sum_{i=1}^{c} \text{seq}(i,n,c) = c^n$

(S5) For $c \leq n, w$, $\lim_{n\to\infty} \text{seq}(c,n,w)/\text{seq}(c+1,n,w) = 0$

(S6) For $c \leq n, w$, $\lim_{n\to\infty} \text{seq}(c,n,w) \Big/ \binom{w}{c} c^n = 1$

(S7) For $c \leq n, w$, $\lim_{n\to\infty} \text{seq}(c,n+1,w)/\text{seq}(c,n,w) = c$.

We now examine some fundamental characteristics of $\text{P}_N(\text{U}_k | e_{nc})$. For three special cases, the value of $\text{P}_N(\text{U}_k | e_{nc})$ is fairly obvious:

(6a) $\text{P}_N(\text{U}_k | e_{nc}) = 0$, if $k < c$
(6b) $= 1$, if $N - n \leq k - c$
(6c) $= 1$, if $k = K$.

If $k < c$ (6a), then given e_{nc} there must be more than k kinds; hence $\text{U}_{N\pi k}$ must be false. If $N - n \leq k - c$ (6b) then, given e_{nc}, even if the $N - n$ objects not mentioned in e_{nc} are all of different kinds, none of which occur in e_{nc}, the number of kinds in the world cannot exceed k; hence $\text{U}_{N\pi k}$ must be true. Finally, if $k = K$ (6c), $\text{U}_{N\pi k}$ is tautologous.

Clauses (6a) and (6b) serve as basis cases from which the following theorem provides a general way of computing $\text{P}_N(\text{U}_k | e_{nc})$ (substituting e_{nc} for A) by induction on N and k.

Theorem 1 *For $N, \pi, k > 0$ and any quantifier-free sentence A of $\mathfrak{L}_\pi^N$ such that $\text{r}_{(N+1)\pi}(A) \neq 0$: $\text{P}_{N+1}(\text{U}_{\pi k} | A) = [(K-k)\ \text{P}_N(\text{U}_{\pi(k-1)} | A) + k\ \text{P}_N(\text{U}_{\pi k} | A)]/K$.*

Proof: $\text{P}_{n+1}(\text{U}_{\pi k} | A) = \text{r}_{(N+1)\pi}(\text{U}_{\pi k}\ \&\ A)/\text{r}_{(N+1)\pi}(A)$. Since A is in $\mathfrak{L}_\pi^N$, any $(N+1)\pi$-state description contains at least one name, α_{N+1}, that does not occur in A. Moreover, each $(N+1)\pi$-state description has the form $(S\ \&\ Q\alpha_{N+1})$, where S is an $N\pi$-state description and $Q\alpha_{N+1}$ is the ascription of some π-kind Q to α_{N+1}. Now $(\text{U}_{\pi k}\ \&\ A)$ is true in exactly those $(N+1)\pi$-state descriptions $(S\ \&\ Q\alpha_{N+1})$ in which A is true and either (i) fewer than k kinds occur in S or (ii) exactly k kinds, one of which is Q, occur in S. Since A is quantifier-free and does not contain α_{N+1}, an $(N+1)\pi$-state description (S & $Q\alpha_{N+1}$) satisfies (i) iff ($U_{\pi(k-1)}$ & A) is true in S and Q is any π-kind. Since there are K π-kinds, the number of such $(N+1)\pi$-state descriptions is K $r_{N\pi}(U_{\pi(k-1)}$ & $A)$. And an $(N+1)\pi$-state description (S & $Q\alpha_{N+1}$) satisfies (ii) iff ($U^*_{\pi k}$ & A) is true in S and Q is one of the k kinds in S. Thus the number of such $(N+1)\pi$-state descriptions is $k\, r_{N\pi}(U^*_{\pi k}$ & $A)$. Moreover, $r_{(N+1)\pi}(A) = K\, r_{N\pi}(A)$, since A is true in an $(N+1)\pi$-state description S & $Q\alpha_{N+1}$ iff A is true in S and Q is one of the K π-kinds. Therefore:

$$P_{N+1}(U_{\pi k}|A) = [K\, r_{N\pi}(U_{\pi(k-1)} \,\&\, A) + k\, r_{N\pi}(U^*_{\pi k} \,\&\, A)]/K\, r_{N\pi}(A).$$

But by (2) and the definition of conditional probability this is:

$$[K\, P_N(U_{\pi(k-1)}|A) + k\, [P_N(U_{\pi k}|A) - P_N(U_{\pi(k-1)}|A)]]/K$$

which is the desired result.

Theorem 1, together with (S1), (5), and (6), provides an algorithm for computing $P_N(Fa_{n+1}|Fa_1$ & ... & Fa_n & $U_{\pi k})$. For the most part, the results are what one would expect. Where w is the π-width of F, if $w, k < K$ and $k > 1$, then $P_N(Fa_{n+1}|Fa_1$ & ... & Fa_n & $U_{\pi k})$ exceeds $P_N(Fa_{n+1})$ and becomes greater as m increases—with certain exceptions, noted below.[8] For $w = K$, Fa_{n+1} is tautologous, and so $P_N(Fa_{n+1}|Fa_1$ & ... & Fa_n & $U_{\pi k}) = 1$ for all values of n. For $k \geq K$, $U_{\pi k}$ is tautologous, so that $P_N(Fa_{n+1}|Fa_1$ & ... & Fa_n & $U_{\pi k}) = P_N(Fa_{n+1}) = w/K$. For $k = 1$, $P_N(Fa_{n+1}|Fa_1$ & ... & Fa_n & $U_{\pi k}) = 1$. We have already seen, however, that all reasonable values for k lie between 1 and K.

The only surprise is that as n increases and gets very close to N, $P_N(Fa_{n+1}|$ Fa_1 & ... & Fa_n & $U_{\pi k})$ may actually decrease. To see why, let us consider the specific case in which $\pi = 2$, $N = 3$, $w = 1$, and $k = 2$. Since $w = 1$, F is a kind; for example, F might be P_1x & P_2x. Now $P_3(F\alpha_3) = w/K = .25$. Calculation reveals that $P_3(F\alpha_3|F\alpha_1$ & $U_{2,2}) = .4$, which seems reasonable, but $P_3(F\alpha_3|F\alpha_1$ & $F\alpha_2$ & $U_{2,2})$ drops again to .25.

The drop occurs because we are considering only three-membered worlds. Given $F\alpha_1$ & $F\alpha_2$, two of the three objects must be of the same kind; hence there can be no more than two kinds. So $U_{2,2}$ is redundant: $P_3(F\alpha_3|F\alpha_1$ & $F\alpha_2$ & $U_{2,2}) = P_3(F\alpha_3|F\alpha_1$ & $F\alpha_2) = .25$. The drop, then, is a result of artificially restricting N to some finite value—in this case 3. We can in effect lift this restriction by considering P_N in the limit as $N \to \infty$. To do so, however, we need additional results, beginning with some corollaries to Theorem 1.

Corollary 1.1 *If A and B are both πnc-kind statements of $\mathcal{L}^N_\pi$, then for $k > 0$:*
(i) $P_N(U_{\pi k}|A) = P_N(U_{\pi k}|B)$
(ii) $P_N(A|U_{\pi k}) = P_N(B|U_{\pi k})$.

Proof: We establish (i) by induction on N. The basis case is $N = n$. Then by (6) if $k < c$, $P_N(U_{\pi k}|A) = P_N(U_{\pi k}|B) = 0$, and if $k \geq c$, $P_N(U_{\pi k}|A) = P_N(U_{\pi k}|B) = 1$. The inductive step is a straightforward application of the formula of Theorem 1. (ii) follows from (i), since by Bayes' Theorem

$P_N(A|U_{\pi k}) = P_N(A)P_N(U_{\pi k}|A)/P_N(U_{\pi k})$, which by (1) and (i) is $P_N(B)P_N(U_{\pi k}|B)/P_N(U_{\pi k}) = P_N(B|U_{\pi k})$.

Corollary 1.2 $P_{N+1}(U_k|e_{nc}) \leq P_N(U_k|e_{nc})$.

Proof: $P_N(U_k^*|e_{nc}) \geq 0$, so that $P_N(U_{k-1}|e_{nc}) \leq P_N(U_{k-1}|e_{nc}) + P_N(U_k^*|e_{nc}) = P_N(U_k|e_{nc})$ by (2). But then by Theorem 1, $P_{N+1}(U_k|e_{nc}) = [(K-k)P_N(U_{k-1}|e_{nc}) + k\,P_N(U_k|e_{nc})]/K \leq [(K-k)P_N(U_k|e_{nc}) + kP_N(U_k|e_{nc})]/K = P_N(U_k|e_{nc})$.

Corollary 1.3 *For* $k > 0$, $P_N(e_{1,1}|U_k) = P_N(e_{1,1}) = 1/K$.

Proof: By Bayes' Theorem, $P_N(e_{1,1}|U_k) = P_N(e_{1,1})P_N(U_k|e_{1,1})/P_N(U_k)$. We show by induction on N that $P_N(U_k) = P_N(U_k|e_{1,1})$, so the desired result follows by (1). The basis case is $N \leq k$. Clearly for $N \leq k$, $P_N(U_k) = P_N(U_k|e_{1,1}) = 1$, since the number of kinds had by objects cannot exceed the number of objects. Now suppose $N \geq k$. Let T be the tautology $P_1\alpha_1 \vee {\sim}P_1\alpha_1$. Then

$$\begin{aligned} P_{N+1}(U_k) &= P_{N+1}(U_k|T) \\ &= [(K-k)P_N(U_{(k-1)}|T) - kP_N(U_k|T)]/K && \text{(Theorem 1)} \\ &= [(K-k)P_N(U_{(k-1)}) - kP_N(U_k)]/K \\ &= [(K-k)P_N(U_{(k-1)}|e_{1,1}) - kP_N(U_k|e_{1,1})]/K \\ &&& \text{(Inductive hypothesis)} \\ &= P_{N+1}(U_k|e_{1,1}) && \text{(Theorem 1)} \end{aligned}$$

Theorem 2 *For any positive integer* x, $P_N(U_k|e_{nc}) = P_{N+x}(U_k|e_{(n+x)c})$.

Proof: By induction on N. The basis case is $N = n$, so that $N + x = n + x$. But then either $k < c$, in which case by (6a) both probabilities are 0, or $k \geq c$, whence by (6b) both are 1. The inductive step is a straightforward application of the formula of Theorem 1.

Corollary 2.1 *For any positive integer* x, $P_N(U_k|e_{nc}) \leq P_N(U_k|e_{(n+x)c})$.

Proof: Immediate by Corollary 1.2 and Theorem 2.

Theorem 3 *For* $c \leq k$, $P_N(U_k|e_{nc}) > 0$.

Proof: We can expand e_{nc} into an $N\pi$-state description S by conjoining to it (in a standard order) ascriptions of one of the kinds in e_{nc} to the $N - n$ names not in e_{nc}. Exactly c kinds occur in S, so that if $c \leq k$, $(U_{Nk}\ \&\ e_{nc})$ is true in S, and hence $r_{N\pi}(U_{Nk}\ \&\ e_{nc}) > 0$. But then by (1) and the definition of conditional probability, $P_N(U_k|e_{nc}) > 0$.

Theorem 4 *For* $c \leq k$ *and* $c < K$:

$$P_N(U_k|e_{(n+1)(c+1)}) = [KP_N(U_k|e_{nc}) - cP_N(U_k|e_{(n+1)c})]/(K-c).$$

Proof: Let α be any name of $\mathfrak{L}_N^\pi$ not in e_{nc}, and let $S_1, \ldots, S_c$ be the ascriptions of the c kinds in e_{nc} to α and $T_1, \ldots, T_{K-c}$ the ascriptions of the $K - c$ kinds not in e_{nc} to α. Clearly $(S_1 \vee \ldots \vee S_c) \vee (T_1 \vee \ldots \vee T_{K-c})$ is a tautology whose two disjuncts are mutually exclusive. Hence:

$$\mathrm{P}_N(S_1 \vee \ldots \vee S_c | e_{nc} \,\&\, \mathrm{U}_k) = 1 - \mathrm{P}_N(T_1 \vee \ldots \vee T_{K-c} | e_{nc} \,\&\, \mathrm{U}_k).$$

But now the disjuncts of $S_1 \vee \ldots \vee S_c$ and $T_1 \vee \ldots \vee T_{K-c}$ are also pairwise mutually exclusive, so this becomes:

$$\sum_{i=1}^{c} \mathrm{P}_N(S_i | e_{nc} \,\&\, \mathrm{U}_k) = 1 - \sum_{i=1}^{K-c} \mathrm{P}_N(T_i | e_{nc} \,\&\, \mathrm{U}_k)$$

which may be expanded by the probability calculus to:

$$\sum_{i=1}^{c} \frac{\mathrm{P}_N(S_i \,\&\, e_{nc})\mathrm{P}_N(\mathrm{U}_k | S_i \,\&\, e_{nc})}{\mathrm{P}_N(e_{nc})\mathrm{P}_N(\mathrm{U}_e | e_{nc})} = 1 - \sum_{i=1}^{K-c} \frac{\mathrm{P}_N(T_i \,\&\, e_{nc})\mathrm{P}_N(\mathrm{U}_k | T_i \,\&\, e_{nc})}{\mathrm{P}_N(e_{nc})\mathrm{P}_N(\mathrm{U}_k | e_{nc})}.$$

(By Theorem 3 the assumption that $c \leq k$ guarantees that both denominators are nonzero.) But for each i, $S_i \,\&\, e_{nc}$ is a $\pi(n+1)c$-kind statement and $T_i \,\&\, e_{nc}$ is a $\pi(n+1)(c+1)$-kind statement. So by (1) this becomes:

$$\sum_{i=1}^{c} \frac{\mathrm{P}_N(\mathrm{U}_k | S_i \,\&\, e_{nc})}{\mathrm{K}\mathrm{P}_N(\mathrm{U}_k | e_{nc})} = 1 - \sum_{i=1}^{K-c} \frac{\mathrm{P}_N(\mathrm{U}_k | T_i \,\&\, e_{nc})}{\mathrm{K}\mathrm{P}_N(\mathrm{U}_k | e_{nc})}$$

i.e.,

$$\sum_{i=1}^{c} \mathrm{P}_N(\mathrm{U}_k | S_i \,\&\, e_{nc}) = \mathrm{K}\mathrm{P}_N(\mathrm{U}_k | e_{nc}) - \sum_{i=1}^{K-c} \mathrm{P}_N(\mathrm{U}_k | T_i \,\&\, e_{nc}).$$

But by Corollary 1.1 this is just:

$$c\mathrm{P}_N(\mathrm{U}_k | e_{(n+1)c}) = \mathrm{K}\mathrm{P}_N(\mathrm{U}_k | e_{nc}) - (K-c)\mathrm{P}_N(\mathrm{U}_k | e_{(n+1)(c+1)}).$$

Corollary 4.1 *For $c \leq k$, $c < K$, and a nonnegative integer x:*

$$\mathrm{P}_N(\mathrm{U}_k | e_{n(c+x)}) \leq \mathrm{P}_N(\mathrm{U}_k | e_{nc}).$$

Proof:

$$\begin{aligned}\mathrm{P}_N(\mathrm{U}_k | e_{n(c+1)}) &= [K\mathrm{P}_N(\mathrm{U}_k | e_{(n-1)c}) - c\mathrm{P}_N(\mathrm{U}_k | e_{nc})]/(K-c) && \text{(Theorem 3)}\\ &\leq [K\mathrm{P}_N(\mathrm{U}_k | e_{nc}) - c\mathrm{P}_N(\mathrm{U}_k | e_{nc})]/(K-c) && \text{(Corollary 2.1)}\\ &= \mathrm{P}_N(\mathrm{U}_k | e_{nc}).\end{aligned}$$

Iterating this yields the desired result.

Theorem 5 *For $c \leq k \leq K$, $\lim_{N\to\infty}\mathrm{P}_N(\mathrm{U}_{k-1} | e_{nc})/\mathrm{P}_N(\mathrm{U}_k | e_{nc}) = 0$.*

Proof: Since $c \leq k$, $0 < \mathrm{P}_N(\mathrm{U}_k | e_{nc}) \leq 1$ by Theorem 3. Hence $[\mathrm{P}_N(\mathrm{U}_{k-1} | e_{nc})/\mathrm{P}_N(\mathrm{U}_k | e_{nc})] \leq \mathrm{P}_N(\mathrm{U}_{k-1} | e_{nc}) \leq \mathrm{P}_N(\mathrm{U}_{k-1} | e_{n1})$ (Corollary 4.1) $\leq \mathrm{P}_N(\mathrm{U}_{K-1} | e_{n1})$ by (2), since $k \leq K$. We show that $\lim_{N\to\infty}\mathrm{P}_N(\mathrm{U}_{K-1} | e_{n1}) = 0$. Since for all N, $\mathrm{P}_N(\mathrm{U}_{k-1} | e_{nc})/\mathrm{P}_N(\mathrm{U}_k | e_{nc}) \geq 0$, this proves the theorem. Now $\mathrm{P}_N(\mathrm{U}_{K-1} | e_{n1}) = \mathrm{r}_{N\pi}(\mathrm{U}_{K-1} \,\&\, e_{n1})/\mathrm{r}_{N\pi}(e_{n1})$. But $\mathrm{r}_{N\pi}(e_{n1}) = K^{(N-n)}$. Let $a_1, \ldots, a_n$ be the names in e_{n1}. $(\mathrm{U}_{K-1} \,\&\, e_{n1})$ is true in just those state descriptions in which the kind Q that occurs in e_{n1} is ascribed to $a_1, \ldots, a_n$ and in which the kinds ascribed to the remaining $N - n$ names belong to some $(K-1)$-membered subset of the K kinds of which Q itself is a member. There are $K - 1$ such subsets, but in general they are not disjoint. And for any such subset, there are $(K-1)^{N-n}$ ways of assigning its members to the $N-n$ names not in e_{n1}. So for all N, $r_{N\pi}(U_{K-1} \,\&\, e_{n1}) \leq (K-1)(K-1)^{N-n}$. But $\lim_{N\to\infty}(K-1)(K-1)^{N-n}/K^{N-n} = (K-1)\lim_{N\to\infty}(K-1)^{N-n}/K^{N-n} = 0$. Thus, since for all N, $P_N(U_{K-1}|e_{n1}) \geq 0$, $\lim_{N\to\infty}P_N(U_{K-1}|e_{n1}) = 0$.

Theorem 6 *For* $c \leq k$, $\lim_{N\to\infty}[P_N(e_{(n+1)c}|U_k)/P_N(e_{nc}|U_k)] = 1/k$.

Proof: Consider the reciprocal $P_N(e_{nc}|U_k)/P_N(e_{(n+1)c}|U_k)$. This has a definite nonzero value by Theorem 3. By the probability calculus this is: $[P_N(e_{nc})P_N(U_k|e_{nc})]/[P_N(e_{(n+1)c})P_N(U_k|e_{(n+1)c})]$, which by (1) is $KP_N(U_k|e_{nc})/P_N(U_k|e_{(n+1)c})$. But by Theorem 1 this is:

$$\frac{K[(K-k)P_{N-1}(U_{k-1}|e_{nc}) + kP_{N-1}(U_k|e_{nc})]}{KP_N(U_k|e_{(n+1)c})}$$

which by Theorem 2 is:

$$\frac{(K-k)P_N(U_{k-1}|e_{(n+1)c}) + kP_N(U_k|e_{(n+1)c})}{P_N(U_k|e_{(n+1)c})}.$$

But by Theorem 5, as $N\to\infty$ this tends to: $(K-k)\cdot 0 + k = k$. The desired limit is then the reciprocal $1/k$.

Theorem 7 *Let* $c \leq k$ *and* $c < K$. *Then*:

$$\lim_{N\to\infty}[P_N(e_{(n+1)(c+1)}|U_k)/P_N(e_{nc}|U_k)] = (k-c)/k(K-c).$$

Proof: $P_N(e_{(n+1)(c+1)}|U_k)/P_N(e_{nc}|U_k) = [P_N(e_{(n+1)(c+1)})P_N(U_k|e_{(n+1)(c+1)})]/[P_N(e_{nc})P_N(U_k|e_{nc})]$, which by (1) is $P_N(U_k|e_{(n+1)(c+1)})/KP_N(U_k|e_{nc})$. Now by Theorem 4 this is:

$$\frac{KP_N(U_k|e_{nc}) - cP_N(U_k|e_{(n+1)c})}{(K-c)KP_N(U_k|e_{nc})}$$

$$= \frac{1}{(K-c)} - \frac{cP_N(U_k|e_{(n+1)c})}{(K-c)KP_N(U_k|e_{nc})}$$

which in turn by the probability calculus and (1) is:

$$\frac{1}{(K-c)} - \frac{cKP_N(e_{(n+1)c}|U_k)}{(K-c)KP_N(e_{nc}|U_k)}.$$

But by Theorem 6, as $N\to\infty$ this tends to: $(k-c)/k(K-c)$.

Theorem 8 *For* $k > 0$, $\lim_{N\to\infty}P_N(e_{nc}|U_k) = k^{-n}\binom{k}{c}\Big/\binom{K}{c}$.

Proof: If $c > k$, $\binom{k}{c} = 0$, and so the result holds by (6a). Suppose, then, that $c \leq k$. Then by repeated applications of Theorem 6:

(i) $\lim_{N\to\infty}[P_N(e_{nc}|U_k)/P_N(e_{cc}|U_k)] = k^{c-n}$.

And by repeated applications of Theorem 7:

(ii) $\lim_{N\to\infty} [P_N(e_{cc}|U_k)/P_N(e_{1,1}|U_k)] = \Pi_{i=1}^{c-1}(k-i)/k(K-i)$.

But by Corollary 1.3, $P_N(e_{1,1}|U_k) = 1/K$, i.e., $(k-0)/k(K-0)$. Thus (ii) becomes:

(iii) $\lim_{N\to\infty} P_N(e_{cc}|U_k) = k^{-c}\Pi_{i=0}^{c-1}(k-i)/(K-i) = k^{-c}\binom{k}{c}\Big/\binom{K}{c}$.

Combining (iii) with (i) gives the desired result.

We now obtain a formula for simple induction as $N\to\infty$:

Theorem 9 *If F is a π-predicate of π-width w and k > 0:*

$$\lim_{N\to\infty} P_N(Fa_{n+1}|Fa_1 \,\&\ldots\&\, Fa_n \,\&\, U_{\pi k}) = \frac{\sum_{c=1}^{\min(k,w)} \text{seq}(c,n+1,w)\binom{k}{c}\Big/\binom{K}{c}}{k\sum_{c=1}^{\min(k,w)} \text{seq}(c,n,w)\binom{k}{c}\Big/\binom{K}{c}}.$$

Proof: Immediate from (4) and Theorem 8.

Notice that by (S3) this is:

$$(7)\quad \frac{\sum_{c=1}^{\min(k,w)} \text{seq}(c,n+1,c)\binom{w}{c}\binom{k}{c}\Big/\binom{K}{c}}{k\sum_{c=1}^{\min(k,w)} \text{seq}(c,n,c)\binom{w}{c}\binom{k}{c}\Big/\binom{K}{c}}.$$

A number of interesting results follow from this formula. For $k = K$, $U_{N\pi k}$ is tautologous, so that (7) reduces by (S3) and (S4) to $w/K = P_N(Fa_{n+1})$. Moreover, as k decreases from K the value of (7) increases. Hence for $k < K$, (7) always exceeds w/K, so that the premises of the simple induction do indeed enhance the probability of its conclusion. For $k = 1$, (7) takes the value 1, regardless of π and n. This is because the premises of the inference then assert both that the universe contains only one kind and that at least one object a_1 has F, so that the one extant kind must be in F; hence all objects must be F. This is counterintuitive; the inductive probability of the inference should vary directly with the sample size n. We have already seen, however, that the value $k = 1$ is extreme and antecedently implausible. For $1 < k < K$, the probability generally does vary directly with n.

There are two exceptions, namely the extreme values for w: $w = 1$ and $w = K$. For $w = K$, Fa_{n+1} is tautologous. Accordingly, (7) reduces by (S3) and (S4) to 1. For $w = 1$, (7) always takes the value $1/k$. Thus the limiting probability of simple induction as $N\to\infty$ does not vary with the sample size n if F is a predicate of width 1 (a kind). This peculiarlity is explained below, following Theorem 10. Where $0 < w$, $k < K$, however, the limiting probability does vary directly with n. The next theorem sets an upper bound on its variation.

Theorem 10 *For a π-predicate F of π-width w and $k > 0$:*

$$\lim_{n\to\infty} \lim_{N\to\infty} \mathrm{P}_N(Fa_{n+1}|Fa_1 \,\&\ldots\&\, Fa_n \,\&\, \mathrm{U}_{\pi k}) = \min(w/k,1).$$

Proof: By (S7), as $n \to \infty$ the quantity in Theorem 9 approaches:

$$\frac{\sum_{c=1}^{\min(k,w)} c \operatorname{seq}(c,n,w)\binom{k}{c}\Big/\binom{K}{c}}{k\sum_{c=1}^{\min(k,w)} \operatorname{seq}(c,n,w)\binom{k}{c}\Big/\binom{K}{c}}.$$

But by (S5) all the terms of both sums for $c < \min(k,w)$ become negligible relative to the term for which $c = \min(k,w)$ as $n \to \infty$. Hence as $n \to \infty$, this tends to $\min(k,w)/k = \min(w/k,1)$.

The explanation of this result is that as $N \to \infty$ the assumption $\mathrm{U}_{N\pi k}$ becomes, for the purposes of the inference, indistinguishable from $\mathrm{U}^*_{N\pi k}$, the assumption that there are exactly k kinds. This is so because as $N \to \infty$ the number of possible worlds containing fewer kinds becomes negligibly small relative to the number containing k. Now our logic regards all possible distributions of kinds among the N objects equally; and for large N, kinds are distributed in roughly equal numbers in nearly all the worlds containing k kinds. Thus the net effect of assuming $\mathrm{U}_{N\pi k}$ as N gets very large is to make it virtually certain that there are exactly k kinds, equally distributed among objects. The identity of these k kinds, however, is unknown. If the π-width of F is w, at most w of them can be in F; hence w/k is an upper bound on the probability of the inference.

Moreover, when $w = 1$ no more than one of the k extant kinds can be in F, so the observation of even one F implies that exactly one of the extant kinds is in F. Observation of additional F's therefore cannot increase the probability that more of the extant kinds are in F, as it does if $1 < w < K$. Given that kinds are equally distributed among the unexamined objects (which as $N \to \infty$ vastly outnumber the examined ones), the proportion of F's among the unexamined objects therefore tends to $1/k$ as $N \to \infty$, no matter how many F's we have observed. This is the explanation of the peculiarity noted above.

Notice, however, that for any language with a reasonably large number of predicates, we are hardly ever concerned with properties of width 1. Moreover, for any fixed finite value of N, $\mathrm{P}_N(Fa_{n+1}|Fa_1 \,\&\ldots\&\, Fa_n \,\&\, \mathrm{U}_{\pi k})$ generally increases as n increases toward N, even if $w = 1$ (though when n gets very close to N there is a final drop, as we saw earlier).

We turn next to *inductive generalization*, the inference from premises of the form $Fa_1 \,\&\ldots\&\, Fa_n$ to the conclusion $(x)Fx$, for some predicate F. From the probability calculus and the semantics for universal quantification, $\mathrm{P}_N((x)Fx|Fa_1 \,\&\ldots\&\, Fa_n \,\&\, \mathrm{U}_{\pi k}) = \mathrm{P}_N(Fa_1 \,\&\ldots\&\, Fa_N|\mathrm{U}_{\pi k})/\mathrm{P}_N(Fa_1 \,\&\ldots\&\, Fa_n|\mathrm{U}_{\pi k})$. Hence by (3), for a π-predicate F of π-width w:

$$(8)\quad \mathrm{P}_N((x)Fx|Fa_1 \,\&\ldots\&\, Fa_n \,\&\, \mathrm{U}_{\pi k}) = \frac{\mathrm{P}_N(Fa_1 \,\&\ldots\&\, Fa_N|\mathrm{U}_{\pi k})}{\sum_{c=1}^{\min(k,w)} \operatorname{seq}(c,n,w)\mathrm{P}_N(e_{nc}|\mathrm{U}_{\pi k})}.$$

Theorem 11 *For any π-predicate F of π-width w and $k > 0$:*

$$\lim_{N\to\infty} P_N((x)Fx|Fa_1 \& \ldots \& Fa_n \& U_{\pi k}) = \frac{\binom{w}{k}\Big/\binom{K}{k}}{\sum_{c=1}^{\min(k,w)} \text{seq}(c,n,w)k^{-n}\binom{k}{c}\Big/\binom{K}{c}}.$$

Proof: By Theorem 8, as $N \to \infty$ (8) approaches:

$$\frac{\lim_{N\to\infty} P_N(Fa_1 \& \ldots \& Fa_N|U_{\pi k})}{k^{-n}\sum_{c=1}^{\min(k,w)} \text{seq}(c,n,w)\binom{k}{c}\Big/\binom{K}{c}}.$$

But $P_N(Fa_1 \& \ldots \& Fa_N|U_{\pi k}) = P_N(Fa_1 \& \ldots \& Fa_N \& U_{\pi k})/P_N(U_{\pi k})$, which by (2) is $\sum_{c=1}^{k} P_N(Fa_1 \& \ldots \& Fa_N \& U^*_{\pi c})/\sum_{c=1}^{k} P_N(U^*_{\pi c})$. But $Fa_1 \& \ldots \& Fa_N \& U^*_{N\pi c}$ is true in just those $N\pi$-worlds containing exactly c kinds, all of which are in F. The number of such worlds is $\text{seq}(c,N,w)$. Thus $P_N(Fa_1 \& \ldots \& Fa_N|U_{\pi k}) =$

$$\frac{\sum_{c=1}^{k} \text{seq}(c,N,w)/K^N}{\sum_{c=1}^{k} \text{seq}(c,N,K)/K^N}.$$

Now by (S2) $\text{seq}(c,N,w) = 0$ if $c > w$. So by (S6) as $N \to \infty$ this approaches:

$$\frac{\sum_{c=1}^{\min(k,w)} \binom{w}{c}c^N}{\sum_{c=1}^{k} \binom{K}{c}c^N}.$$

Suppose $k \le w$. Then as $N \to \infty$, all other terms of the sums in the numerator and denominator become negligible relative to the term for which $c = k$. So this becomes $\binom{w}{k}\Big/\binom{K}{k}$ and so the theorem holds. On the other hand, suppose $k > w$. Then the largest term in the numerator is $\binom{w}{w}w^N = w^N$. But in that case, as $N \to \infty$, the whole expression goes to 0, in which case again the theorem holds, since its numerator is 0 whenever $k > w$.

Theorem 12 *Let F be a π-predicate of π-width w and $k > 0$. Then*:

$$\lim_{n\to\infty}\lim_{N\to\infty} P_N((x)Fx|Fa_1 \& \ldots \& Fa_n \& U_{\pi k}) = 1, \textit{ if } k \le w$$
$$= 0, \textit{ if } k > w.$$

Proof: For $k > w$ the desired result follows immediately from Theorem 11. Suppose, then, that $k \le w$. Now consider the denominator of the formula of Theorem 11. As $n \to \infty$, we know by (S5) that for $c > c'$, $\text{seq}(c', n, w)$ becomes negligible relative to $\text{seq}(c, n, w)$. Hence all other terms of the sum in the denominator become negligible relative to the case $c = k$. So as $n \to \infty$, the formula of Theorem 11 tends to:

$$\frac{\binom{w}{k} \Big/ \binom{K}{k}}{k^{-n}\text{seq}(k, n, w)\binom{k}{k} \Big/ \binom{K}{k}}.$$

But by (S6) this approaches:

$$\frac{\binom{w}{k} \Big/ \binom{K}{k}}{k^{-n}\binom{w}{k}k^{n}\binom{k}{k} \Big/ \binom{K}{k}} = 1.$$

This is a reasonable result. Provided that we adopt a sufficiently strong "inductive method" (i.e., set $k \leq w$), we can ensure that inductive generalization on F approaches 1 as the sample size m increases. Of course, $\text{U}_{N\pi k}$ may be false under the intended interpretation if we make k too small. But if $\text{U}_{N\pi k}$ is false for all $k \leq w$, then $(x)Fx$ is false too, since there exists an object of a kind not in F. In that case we certainly do not want the generalization to have probability 1, and indeed there is no true U-principle which makes it so.

We next investigate the *singular predictive inference.* This is an inference from the premise that exactly s of n specified objects have some predicate F to the conclusion that some other object has F. We must therefore describe the function $\text{P}_N(Fa_{n+1} | Fa_1 \;\&\ldots\&\; Fa_s \;\&\; {\sim}Fa_{s+1} \;\&\ldots\&\; {\sim}Fa_n \;\&\; \text{U}_{\pi k})$. This may be rewritten as:

$$(9)\quad \frac{\text{P}_N(Fa_{n+1} \;\&\; Fa_1 \;\&\ldots\&\; Fa_s \;\&\; {\sim}Fa_{s+1} \;\&\ldots\&\; {\sim}Fa_n | \text{U}_{\pi k})}{\text{P}_N(Fa_1 \;\&\ldots\&\; Fa_s \;\&\; {\sim}Fa_{s+1} \;\&\ldots\&\; {\sim}Fa_n | \text{U}_{\pi k})}.$$

Theorem 13 *Let F be a π-predicate of π-width w and $a_1, \ldots, a_{n+1}$ distinct names. Then for* $1 \leq s \leq n$ *and* $k > 0$:

$$\lim_{N\to\infty} \text{P}_N(Fa_{n+1} | Fa_1 \;\&\ldots\&\; Fa_s \;\&\; {\sim}Fa_{s+1} \;\&\ldots\&\; {\sim}Fa_n \;\&\; \text{U}_{\pi k})$$

$$= \frac{\sum_{c=1}^{k} \sum_{f=1}^{c} \text{seq}(f, s+1, w)\text{seq}(c-f, n-s, K-w)\binom{k}{c} \Big/ \binom{K}{c}}{k\sum_{c=1}^{k} \sum_{f=1}^{c} \text{seq}(f, s, w)\text{seq}(c-f, n-s, K-w)\binom{k}{c} \Big/ \binom{K}{c}}.$$

Proof: Let D_1 and D_2 be the π-normal forms of the conjunctions exhibited in the numerator and denominator of (9), respectively. D_1 is a disjunction of $\pi(n+1)c$-kind statements for various values of c. Each such kind statement consists of $s+1$ ascriptions of kinds in F and $(n+1) - (s+1)$ ascriptions of kinds in ${\sim}F$. In each, if the number of kinds in F is f, then the number of kinds in ${\sim}F$

is $c - f$. Therefore, since there are w kinds in F and $K - w$ kinds in $\sim F$, the number of $\pi(n + 1)c$-kind statements in D_1 containing exactly f kinds in F is $\text{seq}(f, s + 1, w)\text{seq}(c - f, (n + 1) - (s + 1), K - w)$. Similarly, the number of πnc-kind statements in D_2 containing exactly f kinds in F is $\text{seq}(f, s, w)\text{seq}(c - f, n - s, K - w)$. So by Corollary 1.1, (9) is:

$$\frac{\sum_{c=1}^{k} \sum_{f=1}^{c} \text{seq}(f, s + 1, w)\text{seq}(c - f,\ n - s,\ K - w)\mathrm{P}_N(e_{(n+1)c})|\mathrm{U}_k)}{\sum_{c=1}^{k} \sum_{f=1}^{c} \text{seq}(f, s, w)\text{seq}(c - f,\ n - s,\ K - w)\mathrm{P}_N(e_{nc})|\mathrm{U}_k)}.$$

Applying Theorem 8 gives the desired result.

Notice that Theorem 9 is just the special case of Theorem 13 in which $s = n$. For if $s = n$, then by (S1) $\text{seq}(c - f, n - s, K - w) = 0$, unless $f = c$, in which case $\text{seq}(c - f, n - s, K - w) = 1$; and $\text{seq}(f, s, w) = 0$ whenever $f > w$.

As with simple induction and inductive generalization, we wish to know how the inductive probability of the singular predictive inference changes as the sample size increases. But the important limiting value in this case is not the one in which $n \to \infty$ while other parameters stay the same; for in that case the proportion s/n of positive instances in the sample would continually decrease. Rather, we want n to increase while the proportion s/n remains constant. This occurs when both s and n are multiplied by some positive integer x; the larger the value of x, the greater the sample size. We want to know what happens as $x \to \infty$. The solution is given by Theorem 14.

Theorem 14 *Let F be a π-predicate of π-width w and $a_1, \ldots, a_{n+1}$ distinct names. Then for* $1 \leq s \leq n$ *and* $k > 0$:

$$\lim_{x\to\infty} \lim_{N\to\infty} \mathrm{P}_N(Fa_{xn+1}|Fa_1 \,\&\ldots\&\, Fa_{xs} \,\&\, \sim Fa_{xs+1} \,\&\ldots\&\, \sim Fa_{xn} \,\&\, \mathrm{U}_{\pi k}) = f/k,$$

for some real number f such that:

(i) *$f = w$ if $s/n \geq w/k$*
(ii) *$f = k + w - K$ if $s/n \leq (k + w - K)/k$*
(iii) *otherwise, $|f/k - s/n| < 1/k$ (i.e., f/k approximates s/n with an "error" of no more than $1/k$).*

Proof: Applying (S6) to the formula of Theorem 13 we see that as $n \to \infty$ this quantity tends to:

$$\frac{\sum_{c=1}^{k} \sum_{f=1}^{c} \binom{w}{f} f^{xs+1} \binom{K - w}{c - f} (c - f)^{x(n-s)} \binom{k}{c} \Big/ \binom{K}{c}}{k \sum_{c=1}^{k} \sum_{f=1}^{c} \binom{w}{f} f^{xs} \binom{K - w}{c - f} (c - f)^{x(n-s)} \binom{k}{c} \Big/ \binom{K}{c}}.$$

Now $(c - f)^{x(n-s)}$ attains its maximal value in both sums only when c is maximal, i.e., $c = k$. Thus as $x \to \infty$ all the terms of the outer sums in both the numerator and denominator become negligible relative to the term for which $c = k$, so that this approaches:

$$(10)\quad \frac{\sum_{f=1}^{k}\binom{w}{f}f^{xs+1}\binom{K-w}{k-f}(k-f)^{x(n-s)}\binom{k}{k}\Big/\binom{K}{k}}{k\sum_{f=1}^{k}\binom{w}{f}f^{xs}\binom{K-w}{k-f}(k-f)^{x(n-s)}\binom{k}{k}\Big/\binom{K}{k}}$$

$$= \frac{\sum_{f=1}^{k} f\binom{w}{f}\binom{K-w}{k-f}(f^s(k-f)^{n-s})^x}{k\sum_{f=1}^{k}\binom{w}{f}\binom{K-w}{k-f}(f^s(k-f)^{n-s})^x}.$$

Now once again, with these two sums as $x \to \infty$, all but some of the terms become negligible. Since $\binom{w}{f} = 0$ if $f \leq w$, and $\binom{K-w}{k-f} = 0$ if $k - f \geq K - w$, i.e., $f \geq w + k - K$, a term in either of the sums is nonzero only if $w + k - K \leq f \leq w$. Thus the nonnegligible terms are those for which f takes a positive integral value, from $w + k - K$ to $\min(k, w)$, that maximizes $f^s(k-f)^{n-s}$. For the moment, let us treat the latter expression as a real-valued function on f. The derivative of this function with respect to f is $(k-f)^{n-s-1}[f^s(n-s) + (k - f)sf^{s-1}]$, whose value is 0 iff $f = k$, $f = 0$ (an impossible case), or $fn - sk = 0$, i.e., $f/k = s/n$. For positive f, n, s, k and $f \leq k$ (the only possible cases), the derivative is positive if $f/k > s/n$ and negative if $f/k < s/n$. Thus for real f, $f^s(k - f)^{n-s}$ attains a unique maximum among the possible cases iff $f/k = s/n$ and it becomes smaller the more f/k differs from s/n.

But for nonzero terms of the sums, f can take only positive integral values from $w + k - K$ to $\min(k, w)$. Now suppose (Case (i)) that $s/n \geq w/k$. Then since $s/n \leq 1$, $\min(k, w) = w$. Thus f can attain no value greater than w, so that $f^s(k-f)^{n-s}$ is maximal for the possible integral values of f just when $f = w$. But in that case, as $x \to \infty$, (10) tends to w/k.

Similarly, suppose (Case (ii)) that $s/n \leq (k + w - K)/k$. Since $s/n > 0$, $k > K - w$; i.e., $k + w - K > 0$. Thus the minimal value attainable by f in the nonzero terms of the sums is $k + w - K$, so that $f^s(k-f)^{n-s}$ is maximal for the possible integral values of f just when $f = k + w - K$. So as $x \to \infty$, (10) tends to $(k + w - K)/k$.

Finally, suppose (Case (iii)) that $(k + w - K)/k < s/n < w/k$. Now there exists an integer $i < k$ such that $i/k < s/n < (i+1)/k$. We saw that the real-valued function $(f^s(k-f)^{n-s})$ has a maximum at $f = s/n$ and that its value decreases monotonically as f gets increasingly smaller or larger than s/n. Clearly, then, among integral values of f, either i or $(i + 1)$ will maximize $(f^s(k-f)^{n-s})$. There are three possible cases:

(a) $(i^s(k-i)^{n-s}) > ((i+1)^s(k-(i+1))^{n-s})$

(b) $(i^s(k-i)^{n-s}) < ((i+1)^s(k-(i+1))^{n-s})$

(c) $(i^s(k-i)^{n-s}) = ((i+1)^s(k-(i+1))^{n-s})$.

In cases (a) and (b), as $x \to \infty$ (10) obviously tends to i/k and $(i+1)/k$, respectively. In case (c), as $x \to \infty$ (10) tends to:

$$\frac{\left[i\binom{w}{i}\binom{K-w}{k-i}(i^s(k-i)^{n-s})^x\right] + \left[(i+1)\binom{w}{i+1}\binom{K-w}{k-(i+1)}((i+1)^s(k-(i+1))^{n-s})^x\right]}{k\left[\left[\binom{w}{i}\binom{K-w}{k-i}(i^s(k-i)^{n-s})^x\right] + \left[\binom{w}{i+1}\binom{K-w}{k-(i+1)}((i+1)^s(k-(i+1))^{n-s})^x\right]\right]}$$

$$= \frac{i\binom{w}{i}\binom{K-w}{k-i} + (i+1)\binom{w}{i+1}\binom{K-w}{k-(i+1)}}{k\left[\binom{w}{i}\binom{K-w}{k-i} + \binom{w}{i+1}\binom{K-w}{k-(i+1)}\right]}$$

$$= \frac{i\left[\binom{w}{i}\binom{K-w}{k-i} + \binom{w}{i+1}\binom{K-w}{k-(i+1)}\right] + \binom{w}{i+1}\binom{K-w}{k-(i+1)}}{k\left[\binom{w}{i}\binom{K-w}{k-i} + \binom{w}{i+1}\binom{K-w}{k-(i+1)}\right]}$$

$$= i/k + \frac{\binom{w}{i+1}\binom{K-w}{k-(i+1)}}{k\left[\binom{w}{i}\binom{K-w}{k-i} + \binom{w}{i+1}\binom{K-w}{k-(i+1)}\right]}$$

which lies between i/k and $(i+1)/k$.

This result holds because (as we saw with Theorem 10) for large N, the net effect of assuming $U_{N\pi k}$ is to make it virtually certain that there are exactly k kinds, equally distributed among the N objects. Yet, provided that $k < K$, equal distribution of exactly k extant kinds is compatible with very unequal distributions of wider properties. If we sample the universe for some π-predicate F of π-width $w > 1$, as our samples get larger (i.e., as $x \to \infty$) it becomes increasingly likely that the proportion of F's in the universe at large is as close as possible to the ratio of positive instances to observations, s/n. But given that kinds are virtually certain to be equally distributed and that some integral number of the k kinds are in F, this proportion must be of the form f/k, where f is an integer. The tendency will be for f to be such that f/k is as close as possible to s/n.

There are, however, two complications. The first is that if the k kinds are equally distributed, then the proportion of F's in the universe cannot exceed w/k (Case (i) in Theorem 14), since at most w of the k kinds are in F. The second is that for the same reason the proportion of F's in the universe cannot be less than $k - (K - w)$ (Case (ii)), since at most $K - w$ of the k kinds are in $\sim F$.

Suppose, for example, that $\pi = 4$, $k = 5$, and $w = 2$, but that with repeated sampling we find the proportion of F's in our sample to be running about .9.

Continued sampling will increase the probability that some new object has F from $w/K = 2/16 = .125$ (which is its probability in the absence of evidence) up toward $w/k = 2/5 = .4$, but not any higher.

These complications are ameliorated somewhat by the fact that in languages with reasonably large numbers of predicates we are almost always concerned with properties whose widths are neither very small nor very large. (Hence the widths of their negations are not very large or very small, either.) If all the properties that we wish to project inductively, together with their negations, have widths in some interval $[x,y]$, then by choosing $k \leq x$ we ensure that $w > k$ and also that $k \leq K - w$ (since $K - w$ is also in $[x,y]$), thus avoiding the limitations of Cases (i) and (ii).

Moreover, by making π large we can adopt correspondingly large values of k, and hence allow the inductive probabilities of singular predictive inferences to approach the sample proportion s/n to any desired degree of precision.

Thus even our very simple uniformity principles have a surprising robustness. While initially they might seem ill-adapted to the analysis of singular predictive inferences, in fact they can yield fairly reasonable results. They suffer, however, from one further anomaly; and it, unfortunately, is rather serious. As sample size increases (i.e., as x gets larger), the inductive probability as $N \to \infty$ of a singular predictive inference does not in all cases tend monotonically toward its limiting value f/k, as one would expect, but sometimes undergoes an initial oscillation. For example, let $\pi = 4$, $k = 8$, $w = 12$, $s = 3$, and $n = 4$, so that $s/n = 3/4 = .75$. Then where $x = 1$, we obtain by Theorem 13 the value .7408; for $x = 2$ (i.e., $s = 6$, $n = 8$) this increases to .7409, but for $x = 3$ it drops again to .7407 and continues to drop until $x = 6$, at which point it has the value .7402. Thereafter it steadily approaches the limiting value $f/k = 6/8 = .75$.

Again, this oddity can be minimized by making π large while keeping the proportion k/K constant. But it has undesirable effects that will not go away. As Carnap notes, the inductive probability of the singular predictive inference should always lie in the closed interval from s/n to w/K.[9] Yet in the example above, $s/n = w/K = .75$, while the inductive probabilities for finite n are all slightly lower than .75. The reason for this is most readily grasped by considering the case $x = 1$. Here we have observed four objects of which just three are F. This implies that at least one of the extant kinds is in $\sim F$, but it does not imply that at least three of the extant kinds are in F, since two or more of the observed F's may be of the same kind—a possibility which is increasingly likely the smaller the assumed value of k. Thus the sample shows only that at least one of the (presumably) k kinds is in $\sim F$ and at least one (and probably more) are in F. Since among the kinds we have not observed we expect the proportion in F to be $w/K = 3/4$, this suggests that the proportion of extant kinds in F may be a bit less than 3/4; hence the anomaly.

We turn finally to argument by analogy. Here at last our uniformity principles seem perfectly at home. Let F and G be π-properties such that for any name a, Fa implies Ga. By an *analogical inference* I mean an inference from a premise of the form Fa_1 & Ga_2 to the conclusion Fa_2. That is, we assume that a_2 is similar to a_1 in certain respects (namely G) and conclude that it is also similar in certain additional respects (namely those, if any, which differentiate F from G). We could, of course, consider analogical inferences involving more than two objects, but these are best regarded as a mixed mode, involving elements of both analogy and simple induction.

Let the π-widths of F and G be v and w respectively. Clearly $v \leq w$. Thus G is equivalent to a disjunction of w π-kinds, v of which are in F and $w - v$ of which are not in F. So G is equivalent to a sentence $F \vee D$, where D is a sentence of π-width $w - v$, and F and D are mutually exclusive. We wish to find $\mathrm{P}_N(Fa_2 | Fa_1 \,\&\, Ga_2 \,\&\, \mathrm{U}_{\pi k})$. This may be expressed as:

$$(11)\ \mathrm{P}_N(Fa_2 \,\&\, Fa_1 \,\&\, Ga_2 | \mathrm{U}_{\pi k})/\mathrm{P}_N(Fa_1 \,\&\, Ga_2 | \mathrm{U}_{\pi k})$$

$$= \mathrm{P}_N(Fa_1 \,\&\, Fa_2 | \mathrm{U}_{\pi k})/\mathrm{P}_N((Fa_1 \,\&\, Fa_2) \vee (Fa_1 \,\&\, Da_2) | \mathrm{U}_{\pi k})$$

$$= \mathrm{P}_N(Fa_1 \,\&\, Fa_2 | \mathrm{U}_{\pi k})/[\mathrm{P}_N(Fa_1 \,\&\, Fa_2 | \mathrm{U}_{\pi k}) + \mathrm{P}_N(Fa_1 \,\&\, Da_2 | \mathrm{U}_{\pi k})].$$

Theorem 15 *Let F and G be π-properties of π-widths v and w respectively, where $1 \leq v \leq w$, and a_1, a_2 be names in $\mathfrak{L}_\pi^N$. Then for $k > 0$:*

$$\mathrm{P}_N(Fa_2 | Fa_1 \,\&\, Ga_2 \,\&\, \mathrm{U}_{\pi k}) = \frac{\mathrm{P}_N(e_{2,1} | \mathrm{U}_k) + (v-1)\mathrm{P}_N(e_{2,2} | \mathrm{U}_k)}{\mathrm{P}_N(e_{2,1} | \mathrm{U}_k) + (w-1)\mathrm{P}_N(e_{2,2} | \mathrm{U}_k)}.$$

Proof: The number of disjuncts in the π-normal form of $Fa_1 \,\&\, Da_2$ is $v(w - v)$. Thus, applying (3) to (11) we obtain:

$$\frac{\mathrm{seq}(1,2,v)\mathrm{P}_N(e_{2,1} | \mathrm{U}_k) + \mathrm{seq}(2,2v)\mathrm{P}_N(e_{2,2} | \mathrm{U}_k)}{\mathrm{seq}(1,2,v)\mathrm{P}_N(e_{2,1} | \mathrm{U}_k) + \mathrm{seq}(2,2,v)\mathrm{P}_N(e_{2,2} | \mathrm{U}_k) + v(w-v)\mathrm{P}_N(e_{2,2} | \mathrm{U}_k)}.$$

By (S1) this is:

$$\frac{v\mathrm{P}_N(e_{2,1} | \mathrm{U}_k) + (v-1)v\mathrm{P}_N(e_{2,2} | \mathrm{U}_k)}{v\mathrm{P}_N(e_{2,1} | \mathrm{U}_k) + (v-1)v\mathrm{P}_N(e_{2,2} | \mathrm{U}_k) + v(w-v)\mathrm{P}_N(e_{2,2} | \mathrm{U}_k)}$$

which is the desired result.

Theorem 16 *Let F and G be π-properties of π-widths v and w respectively, where $1 \leq v \leq w$, and a_1, a_2 be any names. Then for $k > 0$:*

$$\lim_{N \to \infty} \mathrm{P}_N(Fa_2 | Fa_1 \,\&\, Ga_2 \,\&\, \mathrm{U}_{\pi k}) = \frac{(K-1) + (v-1)(k-1)}{(K-1) + (w-1)(k-1)}.$$

Proof: By Theorem 8, as $N \to \infty$ the formula of Theorem 15 tends to:

$$\frac{k/k^2K + (v-1)k(k-1)/(k^2K(K-1))}{k/k^2K + (w-1)k(k-1)/(k^2K(K-1))} = \frac{1 + (v-1)(k-1)/(K-1)}{1 + (w-1)(k-1)/(K-1)}.$$

The important consequences are apparent at once. For $k = K$, so that $\mathrm{U}_{\pi k}$ is tautologous, $\lim_{N\to\infty}\mathrm{P}_N(Fa_2 | Fa_1 \,\&\, Ga_2 \,\&\, \mathrm{U}_{\pi k}) = \mathrm{P}_N(Fa_2 | Ga_2) = v/w$. With decreasing k, $\lim_{N\to\infty}\mathrm{P}_N(Fa_2 | Fa_1 \,\&\, Ga_2 \,\&\, \mathrm{U}_{\pi k})$ increases monotonically, reaching 1 when $k = 1$. The probability of the inference for $k < K$ also varies directly with the degree of similarity assumed between a_1 and a_2, as measured by the ratio v/w. These consequences are plausible and straightforward.

I hope to have shown by these examples that fully logical inductive logics are not quite as inimical to induction as many have assumed. But I wish to suggest something broader: that our understanding of the relation between induction and order can be advanced by the general method illustrated in this paper. The method is to use fully logical inductive logics as standard contexts in which to compare the ramifications for induction of various assumptions about the uniformity of the world. We have examined one class of uniformity principles in the context of a very limited logic; other possibilities remain to be explored.

NOTES

1. Actually λ is a function, not a constant, but in most interesting cases it is a constant function, and hence Carnap treats it as a constant ([1], §11).
2. See, for example, [13], p. 208. Skyrms explains the conventional wisdom by saying that in fully logical inductive logics $P(A) = P(A \mid B)$; but in a footnote he restricts A and B to atomic sentences. Without the restriction, Skyrms's assertion is false; but even the qualified form is false if $A = B$. (This section is omitted from the third edition of the book.)
3. These uniformity principles are reminiscent of Keynes's Principle of Limited Variety ([8], Chapter XXII). But they are importantly different. Keynes's principle is the metaphysical assertion that (roughly) no object has infinitely many independent qualities. Our uniformity principles assert that the number of actually instantiated combinations of properties expressible in a given language is limited to some specific finite quantity. They are therefore both empirical and falsifiable. Moreover, though Keynes limited the application of his principle to analogical reasoning, it is shown below that uniformity principles based on kinds support a wide variety of inductive inferences.
4. For a detailed discussion of state descriptions, range, width, and related concepts, see [2], Part III.
5. Kuipers provides four axioms that encompass the systems studied in the Carnapian tradition. These vary slightly in each of [9], [10], and [11]. $P_{N\pi k}$ satisfies all these axioms, except for CA1 of [9], which would require $P_{N\pi k}(e_{\pi nc}) > 0$. In [10] and [11] the axiom corresponding to CA1 is loosened, permitting nonregular systems, i.e., those in which $e_{\pi nc}$ may have probability zero.
6. The term is Carnap's ([2], pp. 184–185, 567–569).
7. Equivalently, $\text{seq}(c, n, w)$ may be defined as $S_n^c w!/(w-c)! = S_n^c c! \binom{w}{c}$, where S_n^c is the number of partitions of n objects into c classes, i.e., a Stirling number of the second kind.
8. Proof of these facts is messy, and so is omitted here to save space. The general behavior of simple induction will be clarified when we consider what happens as $N \to \infty$, beginning with Theorem 9.
9. See [1], §8. In Carnap's notation, s/n is written as 's_M/s'.

REFERENCES

[1] Carnap, R., *The Continuum of Inductive Methods*, University of Chicago Press, Chicago, 1952.

[2] Carnap, R., *Logical Foundation of Probability*, Second edition, University of Chicago Press, Chicago, 1962.

[3] Hintikka, J., "Towards a theory of inductive generalization," pp. 274–288 in *Proceedings of the 1964 International Congress for Logic, Methodology and Philosophy of Science*, edited by Y. Bar-Hillel, North Holland, Amsterdam, 1965.

[4] Hintikka, J., "A two-dimensional continuum of inductive methods," pp. 113–132 in *Aspects of Inductive Logic*, edited by J. Hintikka and P. Suppes, North Holland, Amsterdam, 1966.

[5] Hintikka, J., "Induction by enumeration and induction by elimination," pp. 191–216 in *The Problem of Inductive Logic: Proceedings of the International Colloquium in the Philosophy of Science*, London, 1965, vol. 2, edited by I. Lakotos, North Holland, Amsterdam, 1968.

[6] Hintikka, J. and I. Niiniluoto, "An axiomatic foundation for the logic of inductive generalization," pp. 57–81 in *Formal Methods in the Methodology of Empirical Sciences*, edited by M. Przelecki, K. Szaniawski, and R. Wojcicki, Reidel, Dordrecht, 1976. (Reprinted in [7], pp. 157–181.)

[7] Jeffrey, R. C., *Studies in Inductive Logic and Probability*, vol. II, University of California Press, Berkeley, 1980.

[8] Keynes, J. M., *A Treatise on Probability*, MacMillan and Co., London, 1943.

[9] Kuipers, T. A. F., "On the generalization of the continuum of inductive methods to universal hypotheses," *Synthese*, vol. 37 (1978), pp. 255–284.

[10] Kuipers, T. A. F., *Studies in Inductive Probability and Rational Expectation*, Reidel, Dordrecht, 1978.

[11] Kuipers, T. A. F., "A survey of inductive systems," pp. 183–192 in [7].

[12] Niiniluoto, I., "On a *K*-dimensional system of inductive logic," *PSA* 1976, vol. 2, pp. 425–477.

[13] Skyrms, B., *Choice and Chance*, Second edition, Dickenson Publishing Company, Belmont, California, 1975.

Department of Philosophy
University of Tennessee
801 McClung Tower
Knoxville, Tennessee 37996-0480