On the generalization of the continuum of inductive methods to universal hypotheses

Kuipers, Theo A.F.

THEO A. F. KUIPERS*

ON THE GENERALIZATION OF THE CONTINUUM OF INDUCTIVE METHODS TO UNIVERSAL HYPOTHESES**

1. SUMMARY

Carnap's continuum of inductive methods (Carnap, 1952) has been considered, by himself and others, as a proof for the claim that the intuitive concept of rational degree of belief can be explicated, at least with respect to simple situations, in a satisfactory way. At the same time it has been considered as new evidence for the intuitive feeling that such an explication would only be possible for singular (or, individual) hypotheses but not for universal hypotheses. In particular, it was felt that it would not be possible to generalize Carnap's continuum in an acceptable way so that Carnap's continuum appears as an extreme special case.

In this paper it will be shown that this particular conjecture is false and that, consequently, the general conjecture is also false. The requirements for an acceptable generalization will be stated precisely and, in view of the literature on this subject, we have the strong conviction that these requirements will generally be admitted to be necessary and sufficient from the finitary (inductive) point of view.

The generalized continuum is not new, however. It is essentially contained in Hintikka's (1966) α-λ-system and it is essentially equivalent to the class of systems which have recently been introduced by Hintikka and Niiniluoto (1976). The main technical result of this article is the proof that the latter class of systems is equivalent to a particular subsystem of Hintikka's combined system. Hintikka and Niiniluoto could already conclude that it was possible to treat universal hypotheses in a fundamentally acceptable way. The equivalence theorem enables us to specify precisely why and in what sense we are justified to talk about *the* generalization of Carnap's continuum. Moreover it shows that this generalization is axiomatically as well as technically as simple as ever could be expected.

THEO A. F. KUIPERS

2. NOTATIONAL AND TERMINOLOGICAL CONVENTIONS

Carnap has presented his continuum of inductive methods completely in terms of the application he intended: the sentences of a monadic predicate language. But this continuum can also be described in purely mathematical terms, without reference to any particular application. The same holds for the systems that will be discussed in this paper. However, to make the intuitive understanding easier we shall use a terminology based on a very general type of application, including Carnap's favourite one, viz. the terminology used in experimental situations. This approach enables us moreover to simplify the symbolization at several points where misunderstandings are unlikely to arise.

Let there be described a repeatable experiment with K $(2 < K < \infty)$ elementary outcomes $Q_1, Q_2, \ldots, Q_K$ constituting the set T. Subsets of T will also be called outcomes. Performance of n experiments leads to an ordered sequence of elementary outcomes e_n, which is an element of the Cartesian product T^n and which may be used as evidence in relation to hypotheses concerning new experiments. In every particular context it will be clear whether, and in what way, the dummy expression for zero evidence, e_0, may be omitted or inserted.

Let $n_i(e_n)$, or simply n_i if e_n is fixed in the context, be the number of occurrences of Q_i in e_n. Let $c(e_n)$, or simply c, be the number of different Q_j's for which $n_i(e_n) > 0$; i.e. $c(e_n) = | \{Q_i/n_i(e_n) > 0\} |$. $H(e_n)$, or simply H, is the singular hypothesis that the next experiment (the $(n + 1)$th) will result in one of the elementary outcomes that do not occur in e_n or, for short, in a new elementary outcome. $H(e_n)$ therefore corresponds to the outcome $\{Q_i/n_i(e_n) = 0\}$. $\bar{H}(e_n)$ is the hypothesis that one of the elementary outcomes that have already occurred will occur. $\bar{H}(e_n)$ corresponds to the outcome $\{Q_i/n_i(e_n) > 0\}$. Of course we have that $n_i(e_0) = 0$, $c(e_0) = 0$ and $H(e_0)$ corresponds to T.

Let W be a non-empty subset of T. $C_W(n)$ is the (finite) hypothesis that the result of n experiments is such that all members of W have occurred (at least once) and no others. $C_W(n)$ corresponds to the set $\{e_n \in T^n/\forall_i [n_i(e_n) > 0 \leftrightarrow Q_i \in W]\}$. Note that $C_W(n)$ is empty iff $n < | W |$. C_W is the infinite hypothesis that in an infinite continuation of the experiments the elements of W will all occur and no others. C_W will be called a constitutional hypothesis of size $| W |$; it corresponds to the set $\cup_{k=|W|}^{\infty} C_W(k) WWW \ldots$ and

this is a subset of the infinite Cartesian product $WWW \ldots . S_w$ is the infinite hypothesis that precisely w elementary outcomes will occur in an infinite continuation of the experiments. It is called the structural hypothesis of size w and it corresponds to $\cup_{|W|=w} C_W$.

Later on it will be convenient to have a separate notation for an arbitrary constitutional hypothesis of size $w: C_w$. Let $C_w(e_n)$ indicate that C_w is compatible with e_n. Of course we have: $C_W(e_n)$ iff $e_n \in W^n$. Note that the number of C_w's compatible with e_n is equal to $\binom{K-c}{w-c}$ if $c(e_n) = c \leqslant w$, otherwise it is 0.

To simplify probability expressions we will use the following abbreviations[1]:

$$C_W e_n : C_W \cap e_n \, TTT \ldots ,$$

which is equal to

$$C_W \cap e_n \, WWW \ldots , \text{ because } C_W \subset WWW \ldots .$$

Note that $C_W e_n$ is non-empty iff C_W is compatible with e_n.

$$S_w e_n : \bigcup_{|W|=w} C_W e_n ,$$

which is equal to the same union restricted to those W for which $C_W(e_n)$ (and $|W| = w$).

Our concern will be restricted to regular consistent probability patterns with respect to $T, T^2, T^3, \ldots$: a real-valued function p on $T, T^2, T^3, \ldots$ and their power sets is such a pattern if for all $n > 0$

A1[2] $p(e_n) > 0$ for all $e_n \in T^n$

A2 $\sum_{e_n \in T^n} p(e_n) = 1$

A3 $\sum_{Q_i \in T} p(e_n Q_i) = p(e_n)$ for all $e_n \in T^n$ (consistency)

A4 $p(E_n) = \sum_{e_n \in E_n} p(e_n)$ for all $E_n \subset T^n$

The extension theorem of Kolmogorov guarantees that such a pattern has a unique extension to the (measurable subsets of the) infinite Cartesian product

TTT . . . . Of course this extension is such that

B1 $p(C_W) \geqslant 0$,   for all $W \subset T$, and consequently

$$p(S_w) = \sum_{|W| = w} p(C_W) \geqslant 0, \quad \text{for all } w = 1, 2, \ldots, K.$$

Moreover, from the fact that the constitutional hypotheses as well as the structural hypotheses are mutually non-overlapping and together exhaustive it follows that

B2 $\displaystyle\sum_{W \subset T} p(C_W) = 1$ and $\displaystyle\sum_{w=1}^{K} p(S_w) = 1$

The standard definition of conditional probability will be used: if A and B are subsets of the same set then $p(A/B) = p(A \cap B)/p(B)$, provided $p(B) \neq 0$. Probability expressions will be written as simply as possible. The following examples will suffice to illustrate the method:

$$p(Q_i/e_n) : p(e_n Q_i/e_n T) \quad (= p(e_n Q_i)/p(e_n))$$

$$p(Q_i Q_j/e_n) : p(e_n Q_i Q_j/e_n TT) \quad (= p(e_n Q_i Q_j)/p(e_n))$$

$$p(Q_i/e_n \bar{H}) : p(e_n(\bar{H} \cap Q_i)/e_n \bar{H}) \quad (= p(Q_i/e_n)/p(\bar{H}/e_n) \text{ if } n_i > 0)$$

All the foregoing expressions remain adequate if e_n is replaced by $C_w e_n$. Finally we shall use the abbreviations:

$$p(C_w/e_n) : p(C_w e_n/e_n TTT \ldots), \quad p(e_n/C_w) : p(C_w e_n/C_w).$$

The product rule, i.e. repeated application of the equality $p(e_n Q_i) = p(e_n) \cdot p(Q_i/e_n)$, shows that a consistent probability pattern is completely determined as soon as all 'special values' $p(Q_i/e_n)$ (including $p(Q_i/e_0) = p(Q_i)$) are specified. Note that A1 implies moreover that they have to be positive.

3. THE BALL-MODEL AND THE CONDITIONS OF ADEQUACY

Consider a ball of which every point on its surface is coloured by one of the colours $Q_1, Q_2, \ldots, Q_K$. The experiments are random throws and the (elementary) outcome of an experiment is the colour of the point of contact when the ball has come to rest. Let the objective probabilities be equal to the corresponding surface proportions. We assume also that if a colour occurs on

the surface then it has a positive objective probability. Let all this be the only information at the start of the experiments and let the outcomes of the consecutive experiments be the only new information we come to know.

Our aim is to construct a consistent probability pattern with respect to the outcomes which is based on 'rational' principles and satisfies certain minimum conditions of adequacy derived from the general requirement that we want 'to learn from experience'. In the ball-model the following four general conditions are both plausible and precise.

CA1   Positive instantial relevance:

$p(Q_i/e_nQ_i) > p(Q_i/e_n)$

CA2   Relative frequency convergence (Reichenbach-axiom):

If n_i/n approaches a limit, then $p(Q_i/e_n)$ has to approach the same limit.

CA3   Eliminative and enumerative relevance:

$p(C_c/e_nH) = 0; p(C_c/e_n\bar{H}) > p(C_c/e_n)$

(C_c indicates always the unique constitutional hypothesis of size c compatible with e_n)

CA4   Constitutional convergence:

If, after a finite number of experiments, c remains constant then $p(C_c/e_n)$ has to approach 1.

At this point it is difficult to formulate CA4 in a more precise way; below we will see how this condition can be satisfied in a perfectly clear way. Note that the first part of CA3 is satisfied in any consistent probability pattern as soon as $p(e_nH) > 0$, for C_ce_nH is empty. Observe, moreover, that the second part of CA3 as well as CA4 can only be satisfied if $p(C_c) > 0$.

Suppose that there occur on the ball precisely the colours belonging to the subset W of T. Then according to our assumptions the objective probability that in the long run precisely these colours will occur is 1. Hence it is acceptable in this application to interpret C_W as the hypothesis that precisely the colours of W occur on the ball.

Another application is the following urn-model. An urn contains at least K balls; each ball has one of the colours $Q_1, Q_2, \ldots, Q_K$ and the experiments are successive random drawings with replacement. The only problem with this model is that if it is known that the number of balls is finite then there seems

to be no possibility, in the patterns to be studied, for using the information that the objective probabilities are rational fractions.

The application intended by Carnap is essentially this urn-model but then with random drawings *without* replacement. More precisely, he assumed that the Q_i's constitute a family of mutually exclusive and jointly exhaustive predicates with respect to a randomly ordered countable universe. It will be clear that in this application, if the universe is (denumerably) infinite, C_W is equivalent to the *universal* hypothesis that all individuals of the universe exemplify only predicates belonging to W and that each of these predicates is actually exemplified. It is for this reason that we call the C_W's (and the S_w's) also universal hypotheses. If the universe contains only a finite number of N individuals, the described hypothesis corresponds to $C_W(N)$. In this paper, however, we shall only pay attention to the case that infinitely many experiments are in principle possible and also intended.

4. CARNAPIAN SYSTEMS AND THE REQUIREMENTS FOR AN ACCEPTABLE GENERALIZATION

The continuum of inductive methods (Carnap, 1952) is the set of consistent probability patterns for which there is a real number λ, $0 < \lambda < \infty$, such that

(1) $p(Q_i/e_n) = (n_i + \lambda/K)/(n + \lambda)$.

The parameter λ is determined as soon as one special value, for which $n_i \neq n/K$, has been specified, somewhere between $\min(n_i/n, 1/K)$ and $\max(n_i/n, 1/K)$.

Kemeny (1963) has shown that (1), and therefore the complete pattern, can be derived if the following material principles are added to the probability axioms:

POI   Principle of Order Indifference:
        $p(Q_iQ_j/e_n) = p(Q_jQ_i/e_n)$.
PRR[3] Principle of Restricted Relevance (or λ-principle):
        $p(Q_i/e_n) = f(n_i, n)$.

The proof is repeated in the appendix, together with other related proofs. In fact, these principles leave room for the extreme value $\lambda = \infty$ and, if $p(e_n) = 0$ would be allowed, also for the extreme value $\lambda = 0$. The pattern corres-

ponding to a particular finite positive parameter value will be called a Carnapian system.

It is easy to verify that any Carnapian system satisfies CA1 and CA2. It is also well-known, however, that a Carnapian system does not satisfy CA3 and CA4. This is due to the fact that $p(C_K) = 1$ (which will be proved later on). For this does not only imply that $p(C_w) = 0$ if $w \neq K$ (because of B2), but also that $p(C_K/e_n) = 1$ and $p(C_w/e_n) = 0$ for all e_n and $w \neq K$. It is now immediately seen that the condition of enumerative relevance is never satisfied and that the condition of constitutional convergence is only satisfied for $c = K$, but only in a trivial sense.

It has frequently been said that $p(C_K) = 1$ implies that a Carnapian system attaches the value 0 to all non-trivial universal hypotheses. But observe that C_K is in fact not a trivial hypothesis: it excludes the possibility that some elementary outcomes are in fact not realizable by the experiments.

Carnap and many others have held the opinion that it would not be possible to generalize the continuum in an acceptable way such that CA3 and CA4 become satisfied. And so Carnap drew the dramatic conclusion that it was not the task of pure science to pursue universal hypotheses and theories but rather to assign probabilities to finite hypotheses.

In our opinion the main requirements for a satisfactory generalization of Carnap's continuum are:

R1   It has to be based on 'rational' principles: there have to be good reasons for accepting them.

R2   The principles have to be finite: they have to impose general functional relations between probability values concerning finite numbers of experiments (as e.g. POI and PRR).

R3   Parameters have to be finite: their determination has to presuppose only considerations with respect to a finite number of experiments (as λ).

R4   It has to satisfy the conditions of adequacy CA1, . . . , CA4.

More than ten years ago Hintikka construed the so-called combined system (or, more generally, the α-λ-continuum) and he proved that this system satisfies the conditions of adequacy (see Hintikka, 1966). Though he had not presented the system explicitly in terms of principles and parameters, it

seemed perfectly clear that such a reconstruction of the system would bring out that R2 and in any case R3 were violated.

It is a plausible conjecture that the apparent violation of R2 and R3 by the α-λ system was one of the main reasons that Hintikka and Niiniluoto presented, in 1974, a new approach (see Hintikka and Niiniluoto, 1976). They proposed to replace PRR of the Carnapian systems by the, likewise finite, principle:

WPRR   Weak Principle of Restricted Relevance (or c-principle):
$$p(Q_i/e_n) = f_c(n_i, n).$$

They argued that, in the first place, WPRR is at least as defensible as PRR. They also showed that the resulting systems, here called P-systems, satisfy R3: their parameters are finite. Moreover they could not only prove that, under certain conditions, CA1 and CA2 are generally satisfied, but also that CA3 and CA4 are satisfied for the case $c = K - 1$. Finally they sketched a proof for the claim that CA3 and CA4 are generally satisfied. In sum, these new systems seemed to fulfill all the requirements R1, ..., R4.

Hintikka and Niiniluoto concluded that this new approach made it clear that it was, in principle, possible to give an axiomatic foundation for inductive strategies with respect to universal statements. However, the exact relation of the new systems to the Carnapian systems remained unclear, at least with regard to the admissible range of the new parameters. This fact was connected with an apparent general feature of the new systems: in sharp contrast to the Carnapian systems, the new systems seemed to be extraordinarily complicated. This feature made it hard to obtain much quantitative insight in the systems, which explains why the analysis of Hintikka and Niiniluoto was mainly restricted to qualitative considerations.

In this paper it will be shown that the class of P-systems is coextensive with the class of what we shall call Q-systems. These Q-systems are in fact those members of Hintikka's α-λ system in which $\lambda(w)$ is proportional to w but without Hintikka's particular choice of the prior distribution $p(C_w)$ in terms of α. The Q-systems contain the Carnapian systems as extreme cases in a straightforward way. They satisfy all four conditions of adequacy, and the equivalence theorem implies that they can be based on finite principles and finite parameters. The equivalence theorem of course also implies that Q-systems can be based on principles for which good reasons can be given.

However, in our opinion the defining principles for Q-systems are, apart from their infinite character, very reasonable. Finally, the mathematical 'machinery' of Q-systems is highly transparant; it is as simple as could reasonably be expected.

To justify the title of this article we confine ourselves, apart from proving all claims, to the remark that the weak principle of restricted relevance is obviously the slightest weakening of Carnap's principle of restricted relevance for which there are good reasons: the occurrence of a new elementary outcome falsifies an initially possible universal state of affairs. WPRR leaves room for the possbility to change our pattern in case of such events.

5. P-SYSTEMS

In this section we shall treat P-systems in a direct way as far as is necessary to prove the equivalence theorem. The content of this section is essentially contained in the paper by Hintikka and Niiniluoto, but the presentation is rather different.

Def. 1 A P_g-system is a consistent probability pattern with respect to $T, T^2, T^3, \ldots$ satisfying the principles

POI $\quad p(Q_i Q_j/e_n) = p(Q_j Q_i/e_n)$

WPRR $\quad p(Q_i/e_n) = f_c(n_i, n)$.

The following notational conventions will be very useful: since $p(H/e_n) = (K - c)f_c(0, n)$ we may replace $p(H/e_n)$ by $h(n, c)$ and $p(\bar{H}/e_n)$ $(= 1 - p(H/e_n) = 1 - h(n, c))$ by $g(n, c)$. Of course we have $h(0, 0) = p(H) = p(T) = 1$ and $h(n, K) = 0$ for $n \geqslant K$. Moreover the requirement that all $p(Q_i/e_n)$ have to be positive implies:

$$(2) \qquad 0 < h(n,c) < 1 \quad 0 < c \leqslant \min(K - 1, n).$$

On the basis of (2) it is easy to show that WPRR is equivalent to the combination of the three principles:

PR1 $\quad p(H/e_n) = h(n, c)(=_{df} 1 - g(n, c))$

PR2 $\quad p(Q_i/e_n H) = (p(Q_i/e_n)/p(H/e_n)) = 1/(K - c), n_i = 0$

PR3 $\quad p(Q_i/e_n \bar{H}) = (p(Q_i/e_n)/p(\bar{H}/e_n)) = k_c(n_i, n), n_i > 0.$

T1   In a P_g-system there is a real number ρ, $-1 < \rho \leqslant \infty$, such that

(3) $p(Q_i/e_n\bar{H}) = (n_i + \rho)/(n + c\rho)$ $(= k_c(n_i, n))$, $n_i > 0$.

The proof of this theorem, which is given in the appendix, is to a great extent similar to the proof of (1) from POI and PRR.

In this article we shall only study P-systems, which are, by definition, P_g-systems in which

(4) $0 < \rho \leqslant \infty$

Note that, because $g(n, K) = 1$ $(n \geqslant K)$, $k_K(n_i, n)$ and $f_K(n_i, n)$ correspond to $f(n, n_i)$ (of the Carnapian system) if ρ is replaced by λ/K.

T2   A P-system is, in addition to ρ, completely determined by the $(K - 1)$ (finite) parameters $h(c, c)$, $c = 1, 2, \ldots, K - 1$, or by $g(n, 1)$, $n = 1, 2, \ldots, K - 1$. (This does not imply that any choice of them in accordance with (2) is adequate. See section 8.)

Proof: From POI and WPRR follows: if $n_i > 0$ and $n_j = 0$ then $p(Q_iQ_j/e_n) = g(n, c) . k_c(n_i, n) . h(n + 1, c) . (1/(K - c)) = p(Q_jQ_i/e_n) = h(n, c)$. $(1/(K - c)) . g(n + 1, c + 1) . k_{c+1}(n_i, n + 1)$. Substitution of (3) and $g(n, c) = 1 - h(n, c)$ gives us the recursive relation, for $1 \leqslant c \leqslant \min(n, K - 1)$,

(5) $h(n + 1, c) = \dfrac{h(n, c)}{1 - h(n, c)} \cdot (1 - h(n + 1, c + 1)) \cdot \dfrac{n + c\rho}{n + 1 + (c + 1)\rho}$

That all $h(n, c)$ are now determined by the first set of parameters is easily seen by starting the calculation for $c = K - 1$ and $n = K - 1, K, K + 1, \ldots$ which is possible because $h(n, K) = 0$ $(n \geqslant K)$. That the second set is also prepared for this purpose is seen when (5) is rewritten as equation for $g(n + 1, c + 1)$ and the process is started for $c + 1 = 2$ and $n = 1$. The parameters are obviously finite. From PR1,2,3 and T1 it now follows that all special values are determined, and therefore the pattern,[4] Q.E.D.

The second set of parameters has only been given to show that the $h(c, c)$'s are not the only possible simple (finite) parameters. In what follows we shall however take these $h(c, c)$'s as parameters; but first let us introduce the Q-systems.

6. Q-SYSTEMS; ALL Q-SYSTEMS ARE P-SYSTEMS

The constitutional hypotheses are mutually non-overlapping and together exhaustive with respect to TTT .... This enables us to construct a consistent probability pattern by specifying (absolute) probability values for the C_W's and for each C_W a consistent probability pattern with respect to W, W^2, W^3, ... under the condition C_W. The absolute pattern is then obtained by conditionalization according to the rule of Bayes.

In the following definition we shall lay down restrictions on the conditional patterns in such a way that they are equal for constitutional hypotheses of the same size; for this reason we may replace the index 'C_w' by 'w'.

Def. 2 A Q-system is a consistent probability pattern with respect to T, T^2, T^3, ... satisfying the axioms:

Q1 $\quad q_{C_w}(Q_i/e_n) = q_w(Q_i/e_n) = f^w(n_i, n)^5, C_w(e_n Q_i)$

Q2 $\quad q_w(Q_i Q_j/e_n) = q_w(Q_j Q_i/e_n)$

Q3 $\quad q_w(Q_i/e_n \bar{H}) = q_v(Q_i/e_n \bar{H}), n_i > 0$

Q4 $\quad q(C_w) = q(C'_w)$.

According to Q4 constitutional hypotheses of the same size get, in a Q-system, the same value, therefore we have:

$$(6) \qquad q(S_w) = \binom{K}{w} q(C_w),$$

and, in combination with Q1, it follows that any reference to particular subsets W of T may be avoided.

Let us first consider the conditional patterns more in detail.

T3 $\quad$ For the conditional patterns of a Q-system there exists a unique real number ρ, $0 < \rho \leqslant \infty$, such that, except for $w = 2$ and $c = 1$,

$$(7) \qquad q_w(Q_i/e_n) = (n_i + \rho)/(n + w\rho).$$

The exception, which is equivalent to the case $w = 2$ and n_i is n or 0, can easily be restored by adding to Q1 in Def. 2 the simple condition:

Q1.1 $\quad f^2(1, 1) = f^2(1, 3)/(4f^2(1, 3) - 1)$,

or in terms of ρ : $f^2(1, 1) = (1 + \rho)/(1 + 2\rho)$.

The proof of T3 is given in the appendix. Note that the proofs of (1) and (7) are to a large extent similar, for Q1 and Q2 correspond to PRR and POI in Carnapian systems when w is replaced by K. Up to section 9 we shall assume

(8) $0 < \rho < \infty$

T3 says, in effect, that each particular conditional pattern is a Carnapian system. Therefore we now have the important theorem:

T4   For the conditional patterns of a Q-system holds

(9) $q_w(C_w) = 1; \quad q_w(C_v) = 0 \quad 1 \leqslant v < w$

Proof: (Added only for the sake of completeness.) Note first that the theorem is trivial for $w = 1$. Let W ($|W| = w > 1$) be a particular subset of T and let V be a proper non-empty subset of $W(|V| = v)$. Because $0 \leqslant q_w(C_v) \leqslant q_w(VVV\ldots)$ it is sufficient to prove that $q_w(VVV\ldots) = 0$. From (7) it follows that if $e_n \in V^n$, then $q_w(V/e_n) = (n + v\rho)/(n + w\rho)$. By the product rule we get

$$q_w(VVV\ldots) = \prod_{n=0}^{\infty} \frac{n + v\rho}{n + w\rho} = \prod_{n=0}^{\infty} \left(1 - \frac{(w - v)\rho}{n + w\rho} \right).$$

A well-known theorem says that the last product converges to a finite non-zero value if and only if the series

$$\sum_{n=0}^{\infty} \frac{(w - v)\rho}{n + w\rho}$$

converges. But this series is obviously comparable with

$$\sum_{n=1}^{\infty} 1/n,$$

which is well-known to be divergent. Because the factors in the original product are positive and monotone increasing to 1 this product has to be finite and non-negative; and hence it is 0, Q.E.D.

T4 leads us directly to

(10) $q_w(e_n) = q(e_n/C_w)$

and therefore to

(11) $q_w(Q_i/e_n) = q(Q_i/C_w e_n)$

Proof of (10): $q(e_n/C_w)$ is by definition equal to $q(C_w e_n)/q(C_w)$. By conditionalization $q(C_w e_n)$ becomes

$$\sum_{v=w}^{K} \binom{K-w}{v-w} q(C_v) q_v(C_w e_n).$$

Now

$$q_v(C_w e_n) = q_v(C_w) q_v(e_n/C_w).$$

The conclusion then follows by (9), Q.E.D.

From (10) and (11) it follows that we may write all conditionalizations in terms of $q(e_n/C_w)$ and $q(Q_i/C_w e_n)$, which also makes the formulas easier to read. Let us first reformulate (7):

(12) $q(Q_i/C_w e_n) = (n_i + \rho)/(n + w\rho).$

Direct consequences of (12) are:

(13) $q(H/C_w e_n) = \dfrac{(w-c)\rho}{n+w\rho}$ $q(\bar{H}/C_w e_n) = \dfrac{n+c\rho}{n+w\rho}$

(14) $n_i = 0 : q(Q_i/C_w e_n H) = \dfrac{1}{w-c}$ $n_i > 0 : q(Q_i/C_w e_n \bar{H}) = \dfrac{n_i + \rho}{n+c\rho}$

With the abbreviation

$$\eta(n, x) = x(x+1)(x+2) \ldots (x+n-1), \qquad \eta(0, x) = 1,$$

in which n is a positive integer and x a real number, we also obtain from (12), by the product rule,

(15) $q(e_n/C_w) = \prod_i \eta(n_i, \rho)/\eta(n, w\rho)$

Now it is a small step to:

T5   A Q-system is, apart from ρ, completely determined by the $K-1$ (infinite) parameters $q(C_w)$, $w = 1, 2, \ldots, K-1$; they have to be (nonnegative and) such that

$$\sum_{w=1}^{K-1} \binom{K}{w} q(C_w) < 1.$$

Proof: By conditionalization we obtain, using (15),

$$(16) \qquad q(e_n) = \sum_{w=c}^{K} \binom{K-c}{w-c} q(C_w) q(e_n/C_w)$$

$$= \prod_i \eta(n_i, \rho) \cdot \sum_{w=c}^{K} \binom{K-c}{w-c} q(C_w)/\eta(n, w\rho)$$

The theorem follows now directly from

$$\sum_{w=1}^{K} \binom{K}{w} q(C_w) = 1$$

and the requirement $q(e_n) > 0$ for all $c \leqslant \min(K, n)$, Q.E.D.

The special values $q(Q_i/e_n)$ can now directly be obtained from (16) by the equality $q(Q_i/e_n) = q(e_n Q_i)/q(e_n)$, but it is not worth while to write this out. Since $q(C_w/e_n) = q(C_w) \cdot q(e_n/C_w)/q(e_n)$ we obtain from (15) and (16) the following, important, result

$$(17) \qquad q(C_w/e_n) = [q(C_w)/\eta(n, w\rho)] \left/ \left[\sum_{v=c}^{K} \binom{K-c}{v-c} q(C_v)/\eta(n, v\rho) \right] \right.$$

Note that $q(C_w/e_n)$ depends only on w, n and c.

The following theorem is one of the main results of this paper:

T6   Q-systems are P-systems; and the parameter ρ in a Q-system corresponds to the parameter ρ in the P-formulation of that system.

Proof: POI is directly provable by conditionalization of $q(Q_i Q_j/e_n)$ and subsequent application of Q2. By conditionalization and substitution of (17) and (13) we get:

$$(18) \qquad q(H/e_n) = \sum_{w=c+1}^{K} \binom{K-c}{w-c} q(C_w/e_n) q(H/C_w e_n)$$

$$= \frac{\displaystyle\sum_{w=c+1}^{K} \binom{K-c}{w-c} \frac{q(C_w)}{\eta(n, w\rho)} \cdot \frac{(w-c)\rho}{n+w\rho}}{\displaystyle\sum_{v=c}^{K} \binom{K-c}{v-c} \frac{q(C_v)}{\eta(n, v\rho)}}$$

Therefore $q(H/e_n)$ depends only on n and c, which proves PR1. For the sake of completeness we specify also $q(\bar{H}/e_n)$; it can be obtained in the same way from (17) and (13), but of course also from $q(\bar{H}/e_n) = 1 - q(H/e_n)$ and (18):

$$(19) \qquad q(\bar{H}/e_n) = \frac{\displaystyle\sum_{w=c}^{K} \binom{K-c}{w-c} \frac{q(C_w)}{\eta(n, w\rho)} \cdot \frac{n+c\rho}{n+w\rho}}{\displaystyle\sum_{v=c}^{K} \binom{K-c}{v-c} \frac{q(C_v)}{\eta(n, v\rho)}}$$

Analogous to the way in which we obtained (18), by conditionalization from (14) and using that

$$\sum_{w=c}^{K} \binom{K-c}{w-c} q(C_w/e_n) = 1$$

we finally arrive at:

$$(20) \qquad q(Q_i/e_n H) = 1/(K-c), \qquad n_i = 0$$
$$(21) \qquad q(Q_i/e_n \bar{H}) = (n_i + \rho)/(n + c\rho), \qquad n_i > 0$$

PR2 is verified by (20). PR3 is obviously implied by (21). This completes the proof that a Q-system is a P-system. Comparison of (3) and (21) shows that the ρ of the Q-system corresponds to the ρ-parameter in its P-formulation. Of course it was for this fact that we used the same letter, Q.E.D.

7. ALL P-SYSTEMS ARE Q-SYSTEMS

This section will be devoted to the proof of the following theorem:

T7   All P-systems are Q-systems.

One way of proving such a theorem is of course to show that a P-system satisfies the Q-axioms. It turned out to give no essential problems to show that a P-system satisfies Q2, Q3 and Q4. Moreover we could prove that $p_w(Q_i/e_n)$ is a function of at most w, n_i, n and c. But we had to give up the attempt to show the final step leading up to Q1: that $p_w(Q_i/e_n)$ depends only on w, n, n_i and not on c. (But of course, if our claim is true, it must be possible to prove this last step, too.)

Fortunately it is possible to prove the theorem in a completely different

way. The main idea behind this proof is as follows. We start from a P-system (with ρ finite and positive) and try to construct a Q-system with the same special values. If we succeed, this is sufficient; if we do not succeed, we shall attempt to show that there is something wrong with the P-system, namely that it is not probabilistic.

From (3) and (21) we see that a necessary condition for a Q-system to be equivalent to a given P-system is that it has the same parameter-value ρ. This fact will be incorporated in what follows.

Let $Q_1, Q_2, Q_3, \ldots, Q_K$ be an arbitrary enumeration of all K elementary outcomes. Let e_c^c be the evidence $Q_1 Q_2 Q_3 \ldots Q_c$ $(1 \leqslant c \leqslant K)$. In a P-system holds, because of PR1, PR2 and the product rule:

$$(22) \quad p(e_c^c) = \frac{h(0,0)}{K} \cdot \frac{h(1,1)}{K-1} \cdot \frac{h(2,2)}{K-2} \cdots \frac{h(c-1,c-1)}{K-(c-1)}$$

$$= \frac{(K-c)!}{K!} \cdot \prod_{m=0}^{c-1} h(m,m).$$

Since we are trying to construct a Q-system in which, among other things, $p(e_c^c) = q(e_c^c)$, we define, on the basis of (16), the following set of K equations $E(c)$ $(c = 1, 2, \ldots, K)$ with K unknowns, $X(w)$ $(w = 1, 2, \ldots, K)$:

$$E(c) \quad \sum_{w=c}^{K} \binom{K-c}{w-c} X(w) \frac{\rho^c}{\eta(c, w\rho)} = \frac{(K-c)!}{K!} \cdot \prod_{m=0}^{c-1} h(m,m).$$

Note first that, because $h(0,0) = 1, E(1)$ can be transformed into

$$(23) \quad \sum_{w=1}^{K} \binom{K}{w} X(w) = 1.$$

Note further that $E(K)$ has only one unknown $(X(K))$ and that $E(c)$ $(c = 1, 2, \ldots, K-1)$ has one unknown more than $E(c+1)$, viz. $X(c)$. Hence the set of equations has a unique solution satisfying (23). Suppose now that this solution is non-negative, i.e. that

$$(24) \quad X(w) \geqslant 0 \qquad w = 1, 2, \ldots, K-1.$$

Since we have assumed $h(m, m)$ to be positive it follows that $X(K)$ is always positive.

From (23) and (24) we may conclude that the $X(w)$ $(w = 1, 2, \ldots, K-1)$

can be used as parameter values for the $q(C_w)$ in a Q-system. Now consider the Q-system determined by the parameters $\rho, q(C_w) = X(w)$. T6 tells us that this Q-system is also a P-system. From (16) and (22) it follows that $p(e_c^c) = q(e_c^c)$ and hence that $q(H/e_c^c) = h(c, c)$. T2 excludes that there are two different P-systems for which this holds: hence $p \equiv q$.

Now let us set aside our assumption that the solution is non-negative, and let us define generally:

$$(25) \qquad h_x(n, c) = \frac{\displaystyle\sum_{w=c+1}^{K} \binom{K - c}{w - c} \frac{X(w)}{\eta(n, w\rho)} \cdot \frac{(w - c)\rho}{n + w\rho}}{\displaystyle\sum_{v=c}^{K} \binom{K - c}{v - c} \frac{X(v)}{\eta(n, v\rho)}}$$

This definition was, of course, suggested by (18). Because $p(e_{c+1}^{c+1}) = (1/(K - c))\, h(c, c) \cdot p(e_c^c)$ it follows directly from the equations $E(c)$ and (25) that

$$(26) \qquad h_x(c, c) = h(c, c), \qquad c = 1, 2, \ldots, K - 1$$

Now consider the recursive relation (5). It can be checked that, if for some n and c, $h(n, c) = h_x(n, c)$ and also $h(n + 1, c + 1) = h_x(n + 1, c + 1)$ then $h(n + 1, c) = h_x(n + 1, c)$. In other words, *(25) is the explicit solution of (5)*, symbolically:

$$(27) \qquad h(n, c) = h_x(n, c), \qquad c = 1, 2, \ldots, \min(K, n).$$

(This general result is in fact not surprising for, under the restriction of a non-negative solution, it is an immediate consequence of the, already proved, fact that in that case the P-system is a Q-system.)

Suppose now that the solution of the equations $E(c)$ does not satisfy the non-negative condition (24). That is, let $X(w)$ be negative for some $w = 1, 2, \ldots, K - 1$. Let u be the largest index for which this holds. It follows from (25) and (27) that the numerator of $h(n, u)$ is positive for all $n \geqslant u$. Its denominator becomes negative as soon as

$$\sum_{v=u+1}^{K} \binom{K - u}{v - u} X(v) \frac{\eta(n, u\rho)}{\eta(n, v\rho)} < -X(u).$$

Because $-X(u) > 0$ this inequality holds when n is large enough, for

$\eta(n, u\rho)/\eta(n, v\rho)$ approaches 0 for $v > u$ (the proof of this limit-behaviour is essentially contained in the proof of T4), and therefore the whole left-hand sum approaches 0 by increasing n. We may conclude of course that, as soon as this happens, $h(n, u)$ is negative, and this is in conflict with (2). Therefore, our apparent P-system is not a probability pattern and, consequently, it is not a genuine P-system.

8. CARNAPIAN SYSTEMS ARE EXTREME SPECIAL CASES OF Q-SYSTEMS, AND Q-SYSTEMS SATISFY THE CONDITIONS OF ADEQUACY

The established equivalence between P- and Q-systems enables us to study P-systems in their 'Q-garb'. But we may of course also use symbolizations which were introduced for P-systems, such as $h(n, c)$, $g(n, c)$. In the context of a particular Q-system we shall call the Carnapian system with $\lambda = K\rho$ 'its (corresponding) C-system'. The following theorem will clarify the relation between a Q-system and the corresponding C-system.

T8   1   If $q(C_K) < 1$, then:

$$q(H/e_n) < (K - c)\rho/(n + K\rho) \qquad\qquad c \leqslant \min(n, K - 1)$$

$$q(\bar{H}/e_n) > (n + c\rho)/(n + K\rho) \qquad\qquad c \leqslant \min(n, K - 1)$$

$$q(Q_i/e_n) < \rho/(n + K\rho) \qquad n_i = 0 \qquad c \leqslant \min(n, K - 1)$$

$$q(Q_i/e_n) > (n_i + \rho)/(n + K\rho) \qquad n_i > 0 \qquad c \leqslant \min(n, K - 1)$$

2   If $q(C_K) = 1$, then the Q-system coincides with the C-system (i.e. all inequalities in 1. become equalities, including the case $c = K$)

Proof of 1: The first inequality follows directly from (18) and the fact that $(w - c)\rho/(n + w\rho) < (K - c)\rho/(n + K\rho)$ if $c \leqslant w < K$. The rest of the theorem gives trivial consequences of this inequality, (20) and (21).

Proof of 2: This follows directly from (18), (19), (20) and (21).

It might be thought that the requirement that the parameters $h(c, c)$ may not be larger than the corresponding C-values $((K - c)\rho/(c + K\rho))$, guarantees that they give rise to a (probabilistic) P-system. This is, however, not the case. The proof of T7 permits us to add to T2: *The admissible combinations of*

(positive) values for the parameters $h(c, c)$ are determined by the requirement that the equations $E(c)$ should lead to a non-negative solution. This requirement is easily seen to be stronger than the requirement that the parameters may not be larger than the corresponding C-values.

It is nevertheless possible to give a simple, sufficient, but not necessary, condition which guarantees that the equations have a positive solution $(X(w) > 0, w = 1, 2, \ldots, K)$, and therefore that the parameters give rise to a (probabilistic) P-system.

T9   The condition
$$0 < h(c, c) \leqslant \rho/(c + (c + 1)\rho) \quad c = 1, 2, \ldots, K - 1$$
guarantees a positive solution for the equations $E(c)$.

Proof: Note first that $X(K)$ is positive and smaller than 1. Because $p(e_{c+1}^{c+1}) = p(e_c^c)h(c, c)/(K - c)$ it is possible to derive from $E(c)$ and $E(c + 1)$ that

$$X(c) = \sum_{w=c+1}^{K} \binom{K - c}{w - c} X(w) \frac{\eta(c, c\rho)}{\eta(c, w\rho)} \left(\frac{(w - c)\rho}{(c + w\rho)h(c, c)} - 1 \right)$$

It is easy now to check that $X(c) > 0$ if the condition mentioned in the theorem is combined with the inductive hypothesis that $X(w) > 0$, $w = c + 1, \ldots, K$, Q.E.D.

Now we shall start to investigate the behaviour of a Q-system in the light of the conditions of adequacy that were introduced in section 3. At the same time we will derive some other important characteristics of Q-systems. In what follows we shall assume that the $q(C_w)$'s are all positive. It will be easy to check whether the inequality-sign ' $<$ ' has to be replaced by ' $\leqslant$ ' or by ' $=$ ' if this assumption is not (generally) satisfied. The proofs for the theorems will only be sketched.

We shall start with the condition of enumerative relevance (CA3, part two).

(28) $q(C_c/e_n \bar{H}) > q(C_c/e_n)$.

(Follows directly from (17).) An important consequence of (28) is

$$(29) \quad \sum_{w=c+1}^{K} \binom{K - c}{w - c} q(C_w/e_n \bar{H}) < \sum_{w=c+1}^{K} \binom{K - c}{w - c} q(C_w/e_n).$$

(The sum of the lefthand terms of (28) and (29) as well as the sum of the righthand terms must be 1.)

(30) $q(H/e_n\bar{H}) < q(H/e_n)$.

(From the first formulation of (18); use the fact that (13) implies $q(H/C_we_n\bar{H}) < q(H/C_we_n)$; finally, use (29).)

As counterpart of (30) we have:

(31) $q(\bar{H}/e_n\bar{H}) > q(\bar{H}/e_n)$.

From (21) we immediately obtain

(32) $q(Q_i/e_nQ_i\bar{H}) > q(Q_i/e_n\bar{H})$, $n_i > 0$.

Now we are in a position to verify CA1:

(33) $q(Q_i/e_nQ_i) > q(Q_i/e_n)$.

(For $n_i = 0$ directly from T8.1. For $n_i > 0$ it follows from (31) and (32).)

Now let us turn to the limit behaviour. The expression '$q(. ./e_n) \xrightarrow{(c)} L$' always indicates that $q(. ./e_n)$, conceived as real-valued function, has L as its limit if n goes to infinity (c remaining constant).

CA4 follows immediately from (17) and the fact that if $v > c$, then $\eta(n, c\rho)/\eta(n, v\rho) \to 0$, which was proved in the proof of T4. Hence we have

(34) $q(C_c/e_n) \xrightarrow{c} 1$; $q(C_w/e_n) \xrightarrow{c} 0, w > c$.

From (34) we easily get

(35) $q(\bar{H}/e_n) \xrightarrow{c} 1$; $q(H/e_n) \xrightarrow{c} 0$.

(Start from the conditional formulation of $q(H/e_n)$ in (18); use $q(H/C_we_n) \xrightarrow{c} 0$ if $w > c$, which is based on (13); finally, use the second part of (34).)

The Reichenbach-axiom (CA2) is based on the assumption that the Q-system is applied to experiments for which n_i/n (for all Q_i) goes to a certain limit, say q_i. It is well-known that it is problematic whether this assumption is mathematically acceptable, but the intuitive meaning is perfectly clear. The following results have the same shortcomings:

(36)   If $n_i/n \longrightarrow q_i > 0$, then $q(Q_i/e_n\bar{H}) \xrightarrow{c} q_i$.

         If $n_i/n \longrightarrow 0$ and $n_i > 0$, then $q(Q_i/e_n\bar{H}) \xrightarrow{c} 0$.

(Directly from (21).)

(37)   If $n_i/n \longrightarrow q_i$, then $q(Q_i/e_n) \xrightarrow{c} q_i$.

(For $q_i = 0$ and $n_i = 0$ this follows directly from (35). For the other cases: combine (35) and (36).)

Note that in the ball-model of section 3 we assumed that if the colour Q_i occurs on the ball, then $q_i > 0$. In this case the assumption of the Reichenbach-axiom implies that c will tend to a limit, so that we may replace (37) by:

(38)   If for all i either $n_i = q_i = 0$ or $n_i/n \to q_i > 0$, then
$q(Q_i/e_n) \to q_i$.

9. EXTREME CASES

In the preceding sections we restricted our attention to $2 < K < \infty$ and to finite positive values for ρ. In this section we shall make some claims and remarks about what happens if ρ or K takes an extreme value. The claims will not be proved for their proofs are very similar to the proofs in the preceding sections. The expression 'P-system' (or 'Q-system') will be used to refer to a system fullfilling all requirements for being a P-system except perhaps the condition that $p(e_n)$ has to be positive.

$\rho = \infty$.

Claims: — P-systems with $\rho = \infty$ are Q-systems with $\rho = \infty$, and vice versa.
          — Q-systems with $\rho = \infty$ and in which all $q(C_w)$ are positive satisfy CA1, CA3 and CA4 generally; however, they violate CA2.

Remarks: — We have separated this case only because the formulas get a different form.
          — The Carnapian system with $\rho = \infty$ violates all conditions of adequacy ($p(Q_i/e_n)$ is always $1/K$).

$K = 2$.

Claim: — All theorems about P- and Q-systems hold for $K = 2$ if we add the principle of linearity: in case of P-systems: if $c = 2$ and $n_i > 0$, then $p(Q_i/e_n\bar{H})$ is a linear function of n_i, in case of Q-systems: $q_2(Q_i/e_n)$ is a linear function of n_i.

Remarks: — It is well-known that the derivation of the Carnapian systems in case $K = 2$ requires also the related principle: $p(Q_i/e_n)$ is a linear function of n_i.

— The addition of the principle of linearity to the Q-axioms makes Q1.1 superfluous.

— Axiom Q3 does not imply any restriction in case $K = 2$; it is therefore superfluous.

$K = \infty$.

Claims: — If $q(S_w)$, $w = 1, 2, \ldots$ and ρ are taken as parameters in a Q-system such that $\Sigma_{w=1}^{\infty} q(S_w) \leqslant 1$, we get a completely acceptable pattern with respect to denumerably infinite many elementary outcomes. $\left(\text{Notice that the expression} \begin{pmatrix} K - c \\ w - c \end{pmatrix}\right.$ $q(C_w)$, which occurs in all conditionalizations, is equal to

$$\left.\begin{pmatrix} w \\ c \end{pmatrix} q(S_w) \middle/ \begin{pmatrix} K \\ c \end{pmatrix}\right).$$

— The corresponding P-systems can be obtained by taking $g(n, 1)$, $n = 1, 2, \ldots$ and ρ as parameters.

— The equivalence-theorem remains valid and these systems satisfy the conditions of adequacy in the same way as systems with finite K.

Remark: — In these systems it is much easier to give new names to the elementary outcomes when they occur for the first time. It is only in this reformulation that such systems satisfy $p(e_n) > 0$.

$\rho = 0$.

Claims: — 'Q-systems' with $\rho = 0$ are such that $q_w(Q_i/e_n) = n_i/n$, but they are inadequate because $q(H/e_n) = 0$.

— All 'Q-systems' with $\rho = 0$ give rise to the same pattern as the corresponding Carnapian extreme case: the so-called straight rule $p(Q_i/e_n) = n_i/n$.

Remark: — We do not know what 'P-systems' with $\rho = 0$ look like. We did not succeed in finding the explicit solution of (5) for this case; however, $h(n, c) = 0$ for all $c \leqslant \min(n, K)$, and therefore the

straight rule, is a solution. Our conjecture is that this solution is the only one for which $p(e_n)$ is never negative, but it might also be the case that there are several interesting solutions.

$-1 < \rho < 0$.

Claim: — 'Q-systems' with $\rho < 0$ are inadequate because they imply $q(H/e_n) < 0$.

Remark: — We do not know whether 'P-systems' with $-1 < \rho < 0$ are adequate. Apart from particular values of ρ in this interval, the equations $E(c)$ have a unique solution such that (25) remains the explicit solution of (5). But it is difficult to find out in what cases, if any, (25) leads to positive values for $p(H/e_n)$. If ρ takes certain rational values ($\rho = -v/w$ for some v, w such that $1 \leqslant v < w \leqslant K$), then not all equations are adequately defined. It is however our conjecture (in fact strong conviction) that all 'P-systems' with $-1 < \rho < 0$ are inadequate for the same reason as such 'Q-systems': there will be numbers n and c such that $p(H/e_n) < 0$.

10. CONCLUSION

The main conclusion of this article is of course that Carnap's continuum of inductive methods can be generalized in a completely acceptable way. The result is the class of Q-systems, with parameters

$$\rho(0 < \rho < \infty),$$

and

$$q(C_w), w = 1, 2, \ldots, K - 1 \left(\sum_{w=1}^{K-1} \binom{K}{w} q(C_w) < 1 \right).$$

We propose to call this class 'the stratified continuum of inductive methods', for obvious reasons. The equivalence-theorem tells us that this stratified continuum can be founded on 'rational' and finite principles and also that its members can in principle be characterized by finite parameters. The direct analysis of Q-systems shows that they behave in accordance with the conditions of adequacy for individual and universal hypotheses based on the intuitive notion of 'learning from experience'. To be precise, all members of

the stratified continuum satisfy the conditions for individual hypotheses as well as the condition of eliminative relevance for universal hypotheses unrestrictedly. They satisfy moreover the universal conditions of enumerative relevance and constitutional convergence for all C_w for which $q(C_w) > 0$. Finally, the stratified continuum contains the Carnapian continuum as extreme case: $q(C_K) = 1$.

The importance of the equivalence-theorem is of course primarily foundational. What has been shown is that choosing non-trivial initial values for the C_w's, which seems intuitively not acceptable from 'an inductive point of view', is not at all objectionable. Given a particular initial distribution (for the C_w's) we can calculate, by solving the equations $E(c)$ in the reverse way, essentially finite probability values that would give rise to the same pattern if they were taken as parameters in the P-formulation. In other words, the Q-system approach is completely acceptable from a finitary (inductive) point of view.

Fortunately, the Q-system approach is not only very attractive from a technical point of view, but it also seems intuitively more satisfactory to deliberate about the choice of the initial distribution, for apart from the apparent, but refuted, objection we shall, at least to our opinion, in general have more clear intuitions about the initial distribution than about the finite parameters in the P-formulation.

A main task for further research seems therefore to be the development of suggestions for initial distributions. In our opinion the choice has to be related to the particular type of application under consideration and to additional information — if present — with respect to the application, that is: to information which is not already built into the probabilistic framework. To give an example: we might know not only which elementary outcomes may occur but also that there will, with objective probability one, occur, at least so and so many elementary outcomes.

As to the application intended by Carnap, *viz.* a randomly ordered and denumerably infinite universe, Hintikka's (one-parametric) α-distribution (Hintikka, 1966) is very attractive as soon as there are reasons for letting the initial probability of C_w monotone increase with w. In Kuipers (1976), we have proposed a two-parametric distribution which leaves room for this and many other qualitative relations. By the appropriate choice of the two parameters it is then possible to realize a particular qualitative relation if

there are reasons for doing so on logical, statistical or metaphysical grounds.

Let us conclude this article by reformulating the distribution which has been proposed by Carnap in a discussion about Hintikka's α-distribution (Carnap, 1968). An initial distribution can of course not only be specified by the initial values for the constitutional hypotheses (or constituents, as Hintikka has called them) C_w but also by the initial values for the structural hypotheses (or constituent-structures as Carnap has called them) S_w, for we may obtain then the first values by $q(S_w) = \binom{K}{w} q(C_w)$. Carnap's proposal was essentially to apply the intuitive principle of indifference to the S_w's: give all of them the same initial value and therefore the value $1/K$. As is well-known the Carnapian system with $\lambda = K$, or, equivalently, $\rho = 1$, is such that all 'statistical descriptions' (i.e. for a given n, the class of series e_n with the same n_i's is one such description) get the same value. For $\rho = 1$ this remains true for all conditional patterns and all statistical descriptions compatible with the corresponding C_w. In our opinion the Q-system with the initial distribution which was proposed by Carnap ($q(S_w) = 1/K$) and with the value 1 for the parameter ρ is the most sophisticated way in which the classical principle of indifference can be applied in a truely inductive way.

University of Groningen,
The Netherlands

APPENDIX

This appendix contains a combined proof of T.1 and T.3. The proof of 4(1) (i.e. (1) of section 4) from POI and PRR is also included. We shall frequently apply the division-operation; that it is allowed is always essentially based on A1 of section 2.

Step 1

First, consider a conditional pattern q_w of a Q-system. From Q1 it follows that

(1) $q_w(H/e_n) = (w - c)f^w(0, n)$

(2) $q_w(Q_i/e_n H) = 1/(w - c),$ $n_i = 0$

(3) $q_w(Q_i/e_n \bar{H}) = f^w(n_i, n)/(1 - (w - c)f^w(0, n)),$ $n_i > 0.$

Hence q_w satisfies the three principles PR1, PR2 and PR3 of a P-system with $K = w$. Moreover Q2 corresponds to POI. Therefore: q_w is a P-system with $K = w$.

Step 2

Second, consider a P-system $(K > 2)$. From POI it follows, by the product rule, that

(4)   if $n_i > 0$, $n_j > 0$ (and $n_i + n_j \leqslant n - c + 2$ if $c > 2$ and $n_i + n_j = n$ if $c = 2$)

$$p(\bar{H}/e_n)p(Q_i/e_n\bar{H})p(\bar{H}/e_nQ_i)p(Q_j/e_nQ_i\bar{H}) =$$
$$p(\bar{H}/e_n)p(Q_j/e_n\bar{H})p(\bar{H}/e_nQ_j)p(Q_i/e_nQ_j\bar{H})$$

which may be transformed, on the basis of the PR-principles, into

(5) $k_c(n_i, n)k_c(n_j, n + 1) = k_c(n_j, n)k_c(n_i, n + 1)$.

From the probability axioms follows also:

(6) $\sum\limits_{i\,:\,n_i > 0} p(Q_i/e_n\bar{H}) = 1$.

From PR3 we get the special cases of (6):

(6.1) $k_c(n - c + 1, n) + (c - 1)k_c(1, n) = 1$

(6.2) $k_c(n - c + 1, n + 1) + k_c(2, n + 1) + (c - 2)k_c(1, n + 1) = 1$.

Let $c > 2$; substitution of $n_j = 1$ in (5) gives, for $1 \leqslant n_i \leqslant n - c + 1$,

(7) $k_c(n_i, n)k_c(1, n + 1) = k_c(1, n)k_c(n_i, n + 1)$.

Substitution of $n_i = n - c + 1$ resp. $n_i = 2$ leads to the special cases:

(7.1) $k_c(n - c + 1, n + 1) = \dfrac{k_c(n - c + 1, n)}{k_c(1, n)} \cdot k_c(1, n + 1)$

(7.2) $k_c(2, n + 1) = \dfrac{k_c(2, n)}{k_c(1, n)} \cdot k_c(1, n + 1)$.

Substitution of (7.1) and (7.2) in (6.2) gives:

(8) $k_c(1, n + 1) = 1 \left/ \left(\dfrac{k_c(n - c + 1, n)}{k_c(1, n)} + \dfrac{k_c(2, n)}{k_c(1, n)} + c - 2 \right) \right.$

and by substituting (6.1) in (8) we obtain:

(9) $k_c(1, n + 1) = k_c(1, n)/(1 + k_c(2, n) - k_c(1, n))$

With the following definition of λ_c

(10) $k_c(1, c + 1) \underset{df}{=} (1 + \lambda_c/c)/(c + 1 + \lambda_c)$

we are now in a position to prove:

(11)   for fixed $c > 2$ and $1 \leqslant n_i \leqslant n - c + 1$, it holds that

$k_c(n_i, n) = (n_i + \lambda_c/c)/(n + \lambda_c).$

Initial step: $n = c + 1$; therefore $n_i = 1$ or $= 2$; $k_c(1, c + 1)$ satisfies (11) by definition; that $k_c(2, c + 1)$ satisfies (11) now follows directly from (6.1).

Inductive step: suppose (11) holds for fixed $n \geqslant c + 1$, it then follows from (9) that it holds for $k_c(1, n + 1)$ and finally from (7) that it holds also for $k_c(n_i, n + 1)$, $1 < n_i \leqslant n - c + 1$. This completes the proof of (11) as far as $n \geqslant c + 1$.

Final step: that the claim is true for $n = c$, i.e. $k_c(1, c) = 1/c$, follows directly from (6) and PR3.

Step 3

Let $c \geqslant 2$; from POI, the product rule and the PR-principles it is easy to show first that $p(Q_iH/e_n) = p(HQ_i/e_n)$ for $n_i > 0$ and subsequently that

(12) $g(n, c)k_c(n_i, n)h(n + 1, c) = h(n, c)g(n + 1, c + 1)k_{c+1}(n_i, n + 1).$

Substitution of $c = 2$ in (12), using (11) for $c = 3$, leads to the conclusion that $k_2(n_i, n)$ $(1 \leqslant n_i \leqslant n - 1)$ is of the form $F(n) . (n_i + \lambda_3/3)/(n + 1 + \lambda_3)$. From (6) it follows that $k_2(n_i, n) + k_2(n - n_i, n) = 1$. This implies that $F(n) = (n + 1 + \lambda_3)/(n + 2/3 . \lambda_3)$, and therefore, with $\lambda_2 =_{df} 2/3 . \lambda_3$, we may conclude that (11) holds also for $c = 2$ and $1 \leqslant n_i \leqslant n - 1$. Note that (11) holds trivially for $c = 1$.

From (7) it follows that $k_c(n_i, n)/k_{c+1}(n_i, n + 1)$ may not depend on n_i. Hence λ_c/c has to be a constant for all $c = 2, 3, \ldots, K$, say ρ. Hence we have

now generally:

(12) $k_c(n_i, n) = (n_i + \rho)/(n + c\rho)$, $1 \leqslant n_i \leqslant n - c + 1$.

The necessary and sufficient condition which will guarantee that $k_c(n_i, n)$ is always positive is easily seen to be $-1 < \rho \leqslant \infty$. It is also easy to see that this condition ascertains that $k_c(n_i, n)$ is never larger than 1, and this completes the proof of T.1.

Step 4

In step 1 we argued that the conditional patterns q_w of Q-systems are P-systems with $w = K$. Hence we may interpret the proof of (12) as follows:

(13)   for each $w > 2$ there is a real number ρ_w, $-1 < \rho_w \leqslant \infty$ such that
$q_w(Q_i/e_n\bar{H}) = (n_i + \rho_w)/(n + c\rho_w), 1 \leqslant n_i \leqslant n - c + 1$.

From axiom Q3 it now follows that ρ_w is a constant, say ρ. Substitution of this result in (3) gives ($w > 2$)

(14) $f^w(n_i, n) = (1 - (w - c)f^w(0, n)) \cdot ((n_i + \rho)/(n + c\rho))$,

$1 \leqslant n_i \leqslant n - c + 1$.

Because $f^w(n_i, n)$ may not depend on c, it now follows, by comparing (14) for a fixed value of c ($2 \leqslant c < w$) with $c + 1$, that $f^w(0, n) = \rho/(n + w\rho)$, and therefore we have, for $w > 2$, that for all n_i

(15) $f^w(n_i, n) = (n_i + \rho)/(n + w\rho)$, $0 \leqslant n_i \leqslant n$

From Q3 and (13) it also follows for $1 \leqslant n_i \leqslant n - 1$, that $q_2(Q_i/e_n\bar{H})$ is equal to $(n_i + \rho)/(n + 2\rho)$. However, our argument from (14) to (15) cannot be applied here since we cannot compare two values for $c \geqslant 2$. Consider, therefore, the relation from Q1 and Q2:

(16) $f^2(0, n + 1) = f^2(0, n) \cdot f^2(n, n + 1)/f^2(n, n)$.

Since for $c = 2$, $q_2(Q_i/e_n\bar{H}) = q_2(Q_i/e_n)$, we have $f^2(n, n + 1) = (n + \rho)/(n + 1 + 2\rho)$. Suppose now that $f^2(0, n) = \rho/(n + 2\rho)$; then we have not only that $f^2(n, n) = (n + \rho)/(n + 2\rho)$ but from (16) we can also conclude by the relevant substitutions, that $f^2(0, n + 1) = \rho/(n + 1 + 2\rho)$. Combined with the special axiom Q1.1, which implies that $f^2(0, 1) = \rho/(1 + 2\rho)$, we obtain the result that $f^2(0, n) = \rho/(n + 2\rho)$, for all n. On the basis of (14) we now con-

clude that (15) holds also for $w = 2$. Note that (15) is trivial for $w = 1$. Finally, it is easily seen from (15) that the condition $0 < \rho \leqslant \infty$ is necessary and sufficient to guarantee that $f^w(n_i, n)$ is probabilistic; and this completes the proof of T.3.

Step 5

The principles POI and PRR correspond to Q1 and Q2. It is easy to check that the proof of (15) for a given $w > 2$ does not depend on the application of Q3 just prior to (14). Therefore, substitution of $\lambda = w\rho$ in (15) completes the proof of (1) of section 4 for $K = w > 2$.

NOTES

* The author wishes to thank Professor A. J. Stam for his stimulating suggestions during the research and Professor E. M. Barth, Doctor J. F. A. K. van Benthem and Professor J. J. A. Mooij for their comments on the first draft.

** At the end of the research for this article the author received from Professor Ilkka Niiniluoto his first draft of a paper entitled 'On a K-dimensional system of inductive logic' (i.e. on the system of P-systems in the present article). Some passages in that paper are closely, though only implicitly, related to the equivalence theorem, which is the central core of the present paper. The paper of Professor Niiniluoto will appear in the Proceedings of the 1976-PSA-meeting, Vol. 2.

[1] For simplicity we shall write 'e_n', even when the set containing only the n-tuple e_n, $\{e_n\}$, is intended.

[2] It is technically convenient to require regularity, that is to say, to exclude the possibility that $p(e_n)$ may be zero.

[3] This formulation of a principle or axiom has to be interpreted as: the probability value may only change if at least one of the arguments occurring at the right side changes.

[4] Note that, for all n and c, $h(n, c)$ can be calculated in a finite number of steps.

[5] This symbolization should not be misunderstood as the w-th power of $f(n_i, n)$; the index, w, indicates only a possible dependency on w.

Note 1 added in proof: An implicit assumption in the proof in Section 7 is that the solution of the equations $E(c)$ is such that the denominator in (25) is always non-zero. Suppose that this is not true. Let $c = c_0 < K$ be the largest c and, for this c, $n = n_0 + 1$ ($> c_0$) the smallest n for which the denominator of (25), i.e. of $h_X(n_0 + 1, c_0)$, is zero. It can be checked now that $h_X(n_0, c_0) = 1$ and that the proof of (27) remains valid for all n and c for which either $c > c_0$ or $c = c_0$ and $c_0 < n < n_0$. But this implies that $h(n_0, c_0) = 1$, which is in conflict with (2). Hence, for a genuine P-system the equations $E(c)$ are such that the denominator in (25) is always non-zero.

Note 2 added in proof: The class of Q-systems is of course a subclass of the class of systems that arises if Q3 is deleted in Def.2 of Section 6. In Ch.VI of our *Studies in Inductive Probability and Rational Expectation* (Synthese Library 123, Reidel, Dordrecht, forthcoming) this comprehensive class of inductive systems is studied extensively: axiomatic foundation, mutual relations, inductive properties, objective models and infinite extensions. In Ch.V the same is done with respect to a large class of Carnapian-like systems in which, however, $p(Q_i)$ need not be equal to $1/K$. In this book the logico-linguistic approaches to inductive logic have been replaced by a set-theoretic approach to rational expectation in contexts of theories and experiments and to suitable probability systems. In Ch.VII precise characterizations are given of the contexts in which Q-systems (and Carnapian systems) can be applied inductively, i.e. as rational expectation pattern.

BIBLIOGRAPHY

Carnap, Rudolf: 1952, *The Continuum of Inductive Methods*, The University of Chicago Press, Chicago.
Carnap, Rudolf: 1968, 'The Concept of Constituent-Structure', in I. Lakatos (ed.), *The Problem of Inductive Logic*, North-Holland, Amsterdam 1968, pp. 218–220.
Hintikka, Jaakko: 1966, 'A Two-Dimensional Continuum of Inductive Methods', in J. Hintikka and P. Suppes (eds.), *Aspects of Inductive Logic*, North-Holland, Amsterdam 1966, pp. 113–132.
Hintikka, Jaakko and Niiniluoto, Ilkka: 1976, 'Axiomatic Foundation for the Logic of Inductive Genealization', Przełęcki, *et al.*, 1976, pp. 57–81.
Kemeny, John: 1963, 'Carnap's Theory of Probability and Induction', in P. A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*, Open Court, LaSalle 1963, pp. 711–738.
Kuipers, Theo: 1976, 'A Two-Dimensional Continuum of A Priori Probability Distributions on Constituents', in Przełęcki, *et al.*, 1976, pp. 82–92.
Przełęcki, Marian, Szaniawski, Klemens, and Wójcicki, Ryszard (eds.): 1976, *Formal Methods in the Methodology of Empirical Sciences, Proceedings of the Conference for Formal Methods in the Methodology of Empirical Sciences, Warsaw, June 17–21, 1974*, Reidel, Dordrecht 1976.