WOLFGANG SPOHN

ANALOGY AND INDUCTIVE LOGIC: A NOTE ON NIINILUOTO

The first part of this note contains some critical remarks about the forego-
ing paper of Niiniluoto (1981) (to which I shall always refer unless specified
otherwise). These remarks arrive at the conclusion that the systems pro-
posed by Niiniluoto, failing to satisfy symmetry, positive instantial relev-
ance, and the Reichenbach axiom, are more defective than seems accep-
table. At least the latter two properties may be regained for analogical
inference, as I try to show in the second part. The note closes with some
general and skeptical remarks to the effect that by trying to accommodate
analogy considerations, inductive logic regresses to the subjectivistic po-
sition.

1. CRITICISM

Apart from general doubts about inductive logic Carnap's λ-continuum,
though working well for enumerative induction, faced mainly two criti-
cisms: that it deals with generalizations in an inadequate way, i.e. that it is
unable to account for eliminative induction, and that it leaves no room for
considerations from analogy. The first difficulty is shown to be solvable in
a series of papers by Hintikka; and the K-dimensional system proposed by
Hintikka and Niiniluoto (1976) may be said to be *the* perfect generalization
of the λ-continuum with respect to eliminative induction. The second dif-
ficulty is less attended to, at least on a technical level, and in fact unsolved.
The aim of the foregoing paper of Niiniluoto is to overcome this problem
or at least to pave the way towards overcoming it; and his general strategy
to this effect is to modify the K-dimensional system so as to leave room for
analogy considerations without losing any or too many of the basic pro-
perties of the K-dimensional system. This results in two (families of) prob-
ability measures P_s and P_{es} presented in sections 5 and 6 of his paper. I
want to argue in the sequel that, after all, these modifications encroach on
the K-dimensional system too heavily to be acceptable:
(a) *About the system P_s:* As Niiniluoto himself states on p. 14, the prob-
ability measure P_s is identical with the K-dimensional system for exempli-

fied Q-predicates (i.e. the probabilities $f(n', n, c)$ remain unchanged for $n' > 0$ – I use his notation in this section) and thus enjoys the favourable properties of that system within in the area so confined. It does this, however, at the expense of allowing analogy considerations to affect *only* unexemplified Q-predicates. Because of this restrictedness Niiniluoto would, of course, not regard P_s as somethin to rest upon; he has introduced it preliminarily for the useful purpose of demonstrating in detail with a relatively simple system how considerations from analogy operate. Here I restate this restrictedness of P_s in order to point out prophylactically that the extent to which P_s preserves the properties of the K-dimensional system is somewhat deceptive. P_s does so only as far as it disregards analogy effects; and, we shall see, it ceases to do so as soon as it takes in these effects. (Niiniluoto's second system, P_{es}, is – so to speak – franker in this respect.)
(b) *Symmetry*: Niiniluoto himself says clearly on p. 20 and p. 27 that exchangeability or symmetry (with respect to individuals) does not hold for P_s and P_{es}. (For the reason mentioned in (a), and only for this reason, P_s satisfies what Niiniluoto calls a sort of restricted exchangeability on p. 20.) Niiniluoto seems to find this alright when saying that 'conditional inductive probabilities reflecting similarity influence from analogy may significantly depend upon the *order* in which individuals are observed in the new cells Q_i and Q_j' (p. 20). But in fact, this is only an explanation of why his formula (38) *does* not hold for P_s and $i, j \notin I_e$ and cannot count as a reason for which one *should* not wish symmetry to hold in light of analogy considerations. Observe, e.g., that the failing of symmetry also implies that in general

$$P(Q_i(a_{n+1}) \mid e_n^c) \neq P(Q_i(a_{n+2}) \mid e_n^c),$$

where either $P = P_s$ and $i \notin I_e$ or $P = P_{es}$, and that the *a priori* probabilities

$$P_s(e_n^c) \text{ and } P_{es}(e_n^c)$$

depend not only on the numbers n_i ($= 1, ..., K$), but also on the order in which the Q_i's get instantiated n_i times. In my view this is not acceptable; the symmetry condition, I think, loses not a tittle of its intuitive plausibility and of its philosophical justification (which I need not repeat here) merely by taking analogy influences into account. In particular, if the individual

constants $a_1, a_2, \ldots$ are supposed to be proper names for fixed individuals (about which nothing is known a priori), the symmetry condition still seems indispensible. There are the old cases in which symmetry is dispensible or even undesirable. The individual constants might be part of a coordinate language (in Carnap's sense); or the constant a_n might be interpreted as denoting the n-th individual to be observed whichever it may be; and so on. But all of this has nothing intrinsically to do with analogy.

(c) *Positive Instantial Relevance:* Contrary to Niiniluoto's claim on p. 15, the measure P_s violates the principle of positive instantial relevance. Because of (a) it can do so at most for the Q-predicates not yet exemplified; and this it actually does. Thus, it might happen for $j \notin I_e$ that

$$P_s(Q_j(a_{n+2})|\ e_n^c\ \wedge\ Q_j(a_{n+1})) \leq P_s(Q_j(a_{n+1})|\ e_n^c),$$

i.e. according to formula (28) on p. 15 that

$$f(1, n + 1, c + 1) \leq \frac{r(Q_j, e_n^c)}{r(e_n^c)} \cdot f(0, n, c).$$

Choose e.g. $K = 4$, the resemblance function r as in Niiniluoto's example 4 on p. 18, $e_1^1 = Q_1(a_1)$, and the Carnapian values for the parameters γ_c ($c = 1, 2, 3$). Then we have $f(1, 2, 2) = (1 + \lambda/4)/(2 + \lambda)$, $r(Q_2, e_1^1) = 1/2$, $r(e_1^1) = 4/9$, and $f(0, 1, 1) = (\lambda/4)/(1 + \lambda)$, and hence

$$f(1, 2, 2) \leq \frac{r(Q_2, e_1^1)}{r(e_1^1)} \cdot f(0, 1, 1), \text{ if, say, } \lambda \geq 24.$$

This observation carries over to Niiniluoto's measure P_{es}. In fact, P_{es} violates positive instantial relevance everywhere, i.e. not only for unexemplified Q-predicates. But again, I see no reason flowing from analogy considerations which in any way diminishes the desirability of positive instantial relevance.

(d) *The Reichenbach Axiom:* On p. 26 (after formula (47)) Niiniluoto makes some limit assertions about the measure P_{es}, which clearly show that P_{es} does not satisfy the Reichenbach axiom (that for increasing evidence e_n^c the conditional probabilities $P(Q_i(a_{n+1})|\ e_n^c)$ should converge to the limit of the relative frequencies n_i/n observed according to e_n^c – provided that this limit exists). It is also clear that this is not due to any special

features of Niiniluoto's example 6 discussed in that section. This violation of the basic principle of enumerative induction seems uncalled for by analogy considerations. Intuitively, I think, analogy influences would not be awarded the strength of deflecting inductive probabilities from the mark set in the Reichenbach axiom; rather, one would expect analogy influences to get weaker and weaker with increasing evidence and to fade entirely in the limit. This criticism, however, does not apply to the measure P_s, again for the reason mentioned in (a).

The balance is negative, I think. The systems P_s and P_{es} were so constructed as to handle the eliminative aspect of induction as satisfactorily as does the K-dimensional system. But the modifications of the K-dimensional system accounting for analogy considerations nullify at the same time its success on the other scores considered most essential so far.

2. ANOTHER PROPOSAL

Now, if we, still deferring a general discussion, assume a precise analogical inductive logic to be a reasonable project and if we have some conceptions about the desired functioning of analogy influences, what we then have is a technical problem: We are dealing with a very simple, but nevertheless not too restricted language with K monadic Q-predicates $Q_1, \ldots, Q_K$ and infinitely many individual constants $a_1, a_2, \ldots$, and we are looking for a probability measure P on the σ-field of propositions generated by this language which is
  (i) regular,
  (ii) symmetric (with respect to individuals),
  (iii) symmetric with respect to the Q-predicates,
  (iv) satisfies the principle of positive instantial relevance,
  (v) satisfies the Reichenbach axiom,
  (vi) gives positive probability to all contingent universal generalizations, and
  (vii) reflects analogy influences at least in a qualitatively satisfying way.
Other desiderata might be added, but this is the hard core, I think, and I address myself here only to this list. (i) and (iii) may be expected to present no problem in any case. Carnap's λ-continuum fulfills (i)-(v), but not (vi) and (vii); the K-dimensional system fulfills (i)-(vi), but not (vii); and the fact that the prima facie plausible attempts of Niiniluoto to tackle (vii) in a

systematic way were insufficient might make our problem appear too demanding. Indeed, it is demanding, and I cannot offer a solution to it. However, as far as I can see, the really tough job is to get symmetry or exchangeability. But (i) and (iii)-(vii) are less difficult to satisfy, as I shall show in the sequel.

The strategy now is not to modify some known system of inductive logic in the hope of keeping its favourable properties; this hope does not seem to be realizable. Rather, I shall dismiss symmetry outright, either for the strong, though unsatisfactory reason that it is too hard for me to get or for the better reason of having interpreted the individual constants in some of the ways mentioned earlier which make symmetry unwarranted. After that move we have great freedom in constructing suited probability measures. As our construction method I choose the one which has proved most effective in this field, namely, to specify a suitable characteristic function, i.e. to specify the probabilities

(1) $\qquad P(Q_i(a_{n+1})| e_n)$

for all n and all sample propositions e_n describing how the first n individuals $a_1, \ldots, a_n$ are distributed over $Q_1, \ldots, Q_K$. It is well known that *any* specification of these numbers such that $P(Q_i(a_{n+1})| e_n) > 0$ $(i = 1, \ldots, K)$ and $\sum_{i=1}^{K} P(Q_i(a_{n+1})| e_n) = 1$ uniquely determines a regular probability measure on the given σ-field. To restrict our somewhat excessive freedom, I make the common assumption that the probabilities (1) depend only on i and $\mathbf{n} = (n_1, \ldots, n_K)$, where n_j is the number of individuals belonging to Q_j according to e_n (and, hence, $\sum_{j=1}^{K} n_j = n$). Thus, we may write the probabilities (1) as

(2) $\qquad f_i(\mathbf{n})$.

Note that this assumption still falls short of exchangeability. We also had to satisfy the product or the quotient condition of Carnap, Jeffrey (1971), p. 153, in order to have full exchangeability.

In the following I shall not give the functions f_i directly. It will be more convenient to specify some suited functions g_i $(i = 1, \ldots, K)$ defined for all

K-tupels **n** of nonnegative integers and having arbitrary positive reals as values and to obtain the f_i from the g_i by normalization; i.e. having specified the g_i, we may first set

$$(3) \qquad g(\mathbf{n}) = \sum_{i=1}^{K} g_i(\mathbf{n})$$

and then define

$$(4) \qquad f_i(\mathbf{n}) = g_i(\mathbf{n})/g(\mathbf{n}).$$

So far, this is the notation I shall use. To get some idea how suitable functions g_i might look like, let me quickly review Carnap's λ-continuum and the K-dimensional system.

The λ-continuum is very simply characterized, namely by setting

$$(5) \qquad g_i(\mathbf{n}) = n_i + \lambda/K$$

for some $\lambda > 0$. Obviously, $g(\mathbf{n}) = n + \lambda$, and then the familiar expression for Carnapian characteristic functions ensues from (4). It is immediately clear that the measure so specified satisfies conditions (i) and (iii)–(v), and symmetry happily also holds, of course. As is well known, condition (vi), however, is not fulfilled. Since we want to satisfy it, let me once more briefly diagnose the mathematical source of trouble:

What we want is a measure P for which at least

$$(6) \qquad P(\wedge x \neg Q_i(x)) > 0 \text{ for all } i = 1, \ldots, K.$$

Abbreviating the proposition $\neg Q_i(a_1) \wedge \cdots \wedge \neg Q_i(a_n)$ by $A_{i,n}$ (and identifying $A_{i,0}$ with the tautology), we know that

$$(7) \qquad P(\wedge x \neg Q_i(x)) = \prod_{n=0}^{\infty} P(\neg Q_i(a_{n+1})| A_{i,n}) =$$

$$\prod_{n=0}^{\infty} [1 - P(Q_i(a_{n+1})| A_{i,n})].$$

Because of regularity, we have $0 < P(Q_i(a_{n+1})| A_{i,n}) < 1$ for all n; hence,

we may conclude further (cf. e.g. v. Mangoldt-Knopp (1962), p. 286) that the infinite product in (7) does not vanish if and only if

$$(8) \qquad \sum_{n=0}^{\infty} P(Q_i(a_{n+1})| A_{i,n}) < \infty.$$

Let us finally express (8) in terms of the characteristic function. To this end, define $\mathbf{n} = (n_1, \ldots, n_K)$ to be compatible with $A_{i,n}$ iff $\sum_{j=1}^{K} n_j = n$ and $n_i = 0$. Then (8) and therefore (6) holds if

$$(9) \qquad \sum_{n=0}^{\infty} \max f_i(\mathbf{n}) < \infty \text{ for all } i = 1, \ldots, K,$$

where the maximum is taken from all $\mathbf{n}$ compatible with $A_{i,n}$. (This is so, because $P(Q_i(a_{n+1})| A_{i,n}) \leq \max f_i(\mathbf{n})$.) (9) is somewhat stronger than (8), but in a desired way. For (9) implies not only (6), but also that all contingent universal generalizations, in fact all constituents, have positive probability. To indicate the idea for proving this, consider the universal proposition $U = \wedge x(\neg Q_i(x) \wedge \neg Q_j(x))$. Since by (9) we have

$$\sum_{n=0}^{\infty} \max f_i(\mathbf{n}) + \sum_{n=0}^{\infty} \max f_j(\mathbf{n}) < \infty,$$

a fortiori we have

$$\sum_{n=0}^{\infty} \max (f_i(\mathbf{n}) + f_j(\mathbf{n})) < \infty,$$

where the maximum is now taken from all $\mathbf{n}$ compatible with $A_{i,n}$ and $A_{j,n}$. This implies $P(U) > 0$ in the same way as (9) implies (6). Similarly, this also implies $P(U| B) > 0$, where B is the proposition that all the $K-2$ Q-predicates different from Q_i and Q_j are exemplified by the first $K-2$ individuals. Since $P(B) > 0$, this in turn implies that the constituent that all and only these Q-predicates are exemplified also has positive probability. And this argument generalizes to all constituents.

So, to get condition (vi), all we have to do is to choose the f_i or the g_i in such a way that (9) is satisfied. But by choosing the g_i as in (5) in the Carnapian way, we do not get (9), because then for all $\mathbf{n}$ with $n_i = 0$ $f_i(\mathbf{n}) =$

$(\lambda/K)/(n + \lambda)$, the infinite sum of which diverges. In other words: If we fix $g(\mathbf{n})$ to be $n + \lambda$, as seems to be convenient, then we cannot satisfy (9) by allowing any non-exemplified Q-predicate Q_i to keep its initial logical weight λ/K as g_i-value. And this is exactly the point where the K-dimensional system modifies the λ-continuum. In the K-dimensional system the non-exemplified Q-predicates get fined, so to speak, for not being exemplified; and the fine they have to pay increases with increasing evidence rapidly enough to satisfy (9); this means, in particular, that in the limit the fine eats up the whole initial logical weight. And what happens to the fine? It is distributed among the exemplified Q-predicates proportionally to their Carnapian g_i-values. Mathematically, this means that the K-dimensional system determines the functions g_i in this way:

$$(10) \qquad g_i(\mathbf{n}) = \begin{cases} \lambda/K \cdot (1 - \sigma_0(n, c)), & \text{if } n_i = 0, \\ (n_i + \lambda/K)(1 + \sigma_1(n, c)), & \text{if } n_i > 0, \end{cases}$$

where $\sigma_0(n, c)$ and $\sigma_1(n, c)$ are suitably chosen positive reals depending on the number n of individuals observed and on the number c of exemplified Q-predicates. Of course, we must have $(K - c) \cdot (\lambda/K) \cdot \sigma_0(n, c) = (n + c\lambda/K) \cdot \sigma_1(n, c)$ so that $g(\mathbf{n}) = n + \lambda$. Furthermore, $\sigma_0(n, c)$ is increasing with n and converges to 1 for $n \to \infty$. Again, the conditions (i) and (iii)–(v) are rather obviously satisfied, and Hintikka and Niiniluoto have chosen the 'fine' $\sigma_0(n, c)$ so ingeniously that symmetry, and many similarities to the λ-continuum, also hold. (In the notation of Niiniluoto (1977) we have $\sigma_0(n, c) = 1 - (K(n + \lambda))/\lambda(K - c)) \cdot [1 - \alpha(n, c)(n + c\lambda/K)]$.)

All this points to a way of constructing some functions g_i, or rather a considerable class of functions g_i, which are also able to deal with analogy influences. (In fact, even after specializing (1) to (2), we have a host of possibilities. But in this note I shall not engage in the philosophically still ill-founded enterprise of sounding all mathematical alternatives; one proposal will suffice.) The following features of construction seem to me to be suggested by the foregoing:

First, it seems to be a good move to fix $g(\mathbf{n})$ to be $n + \lambda$ and then to distribute the empirical mass n according to the absolute frequencies observed such that $g_i(\mathbf{n})$ equals n_i plus some portion from the logical or nonempirical mass λ still to be specified. This already guarantees us that the Reichenbach axiom is satisfied, and it also ensures positive instantial relev-

ance, provided that we avoid extreme fluctuations of the non-empirical mass. Moreover, it is clear that at the beginning the logical mass λ should be distributed evenly among the Q-predicates; this gives us symmetry with respect to the Q-predicates.

Hence, the second idea is to account for analogy influences by appropriately redistributing the non-empirical mass λ. As Niiniluoto has shown in detail, a redistribution as in (10) cannot afford this, since, e.g., all non-exemplified Q-predicates are treated there in the same way. Thus, the redistribution should make essential use of a distance function or similarity measure for the Q-predicates ruling the strength of analogy influences. This distance function may be chosen in some of the ways presented by Niiniluoto in section 4 of his paper. Here I only presuppose that for any i, j = 1, ..., K a number d_{ij} somehow measuring the distance or dissimilarity of Q_i to Q_j is given such that

(11) $d_{ij} = 0$ for $i = j$ and $d_{ij} > 0$ for $i \neq j$.

But there is a moral to be drawn from (10). At first sight it might seem that, if Q_i is very similar to Q_j, Q_i should profit by the observation of a Q_j-individual (i.e. the value of g_i should thereby increase). But then it becomes very difficult, if not impossible, to satisfy condition (vi) concerning universal generalizations. Thus, it seems advisable instead to proceed analogously to (10), i.e. to make the analogy influence of Q_j on Q_i always pay off negatively – the more the less similar Q_i is to Q_j. Intuitively, each Q_i now gets 'fined' for being dissimilar to the other Q_j's. (This does not at all have the consequence that there will be only negative analogy influences in the sense of Niiniluoto's paper.)

After these strategic remarks, the third step is to spell out the analogy influence numerically. Let us first consider the influence of Q_j on Q_i in isolation: It should depend on the distance d_{ji} of Q_j to Q_i and on the number n_j of observed Q_j-individuals; it should depend on nothing else (at least there is no plausible further candidate on which it might depend); and it should be bounded (this is suggested by the criticism (d) in section 1). If we fix for the moment the upper bound, i.e. the maximal analogy or, rather, disanalogy influence of Q_j on Q_i, to be 1, this means that this influence may be given by some function h which takes a positive real d and

a nonnegative integer m as arguments and which must have the following properties:

(12) $h(d, 0) = 0$,

(13) h is strictly increasing in d and in m, i.e. if $d = d'$ and $m < m'$ or $d < d'$ and $m = m'$, then $h(d, m) < h(d', m')$,

(14) $\lim\limits_{m \to \infty} h(d, m) = 1$ for each d,

(15) $\lim\limits_{d \to 0} h(d, m) = 0$ for each m,

(16) $\sum\limits_{m=1}^{\infty} \dfrac{1}{m} (1 - h(d, m)) < \infty$ for each d.

(16), which actually implies (14), already aims at getting (9) and need not worry us for the moment. A smooth function satisfying (12)–(16) would be $h(d, m) = 1 - \theta/(\theta + md)$ for some positive θ. But, in principle, we are free to choose any other function satisfying (12)–(16).

To summarize, the idea so far is that $g_i(\mathbf{n})$ equals the empirical weight n_i of Q_i plus its initial logical weight λ/K minus the 'fine', i.e. some combination of the disanalogy influences from all the other Q_j's. There remains the question: what combination? There is only one hard condition to satisfy: that the combination be such that in the limit it eats up the initial logical weight of unexemplified Q-predicates. For no very convincing reason, just because it does not look unplausible, I shall stipulate that the combination is a weighted average of the individual disanalogy influences, where the weights are to be Carnapian probabilities; thus the weight of the disanalogy influence of Q_j on Q_i is to be $(n_j + \mu/K)/(n + \mu)$ for some positive μ (which may very well be different from λ). This finally amounts to defining:

(17) $g_i(\mathbf{n}) = n_i + \dfrac{\lambda}{K} \cdot \left[1 - \sum\limits_{j \neq i} \dfrac{n_j + \mu/K}{n + \mu} \cdot h(d_{ji}, n_j) \right]$.

A last question is still open: What happens to the fines? There are three plausible possibilities which, however, are qualitatively equivalent (i.e.

with respect to our conditions (i) and (iii)–(vii)). The first is to credit to each Q_i all the fines for the other Q_j's which are due to it. The second is to distribute the total fine among the Q-predicates proportionally to their Carnapian g_i-values, as was done in (10). And the third is to distribute it proportionally to the g_i-values given by (17); with respect to the functions f_i, this is the same as not distributing it at all. I choose the last possibility and thus avoid changing or even complicating (17); but again, this is not a very strong reason. (Note that now $g(\mathbf{n}) = n + \lambda$ does not hold, after all; rather, we have $n < g(\mathbf{n}) \leq n + \lambda$.)

It is high time, I think, for an objection – the objection that the reasoning leading to (17) was, to say the least, very soft. It was indeed; there are many points which are debatable and for which some other alternative might be preferred – but unavoidably so at the present time. One can demonstrate mathematical possibilities, but the intuitive and philosophical discussion is not so advanced, to my knowledge, to provide good reasons for a rather definite choice among these possibilities; in fact, it seems questionable whether it can provide this at all. I shall come back to this point in the last section.

For the moment, I am content that my proposal (17) does satisfy the conditions (i) and (iii)–(vii): Regularity follows from the fact that $g_i(\mathbf{n}) > 0$ for all i and $\mathbf{n}$. Symmetry with respect to Q-predicates holds because $g_i(0, ..., 0) = \lambda/K$ for all i. Positive instantial relevance may be proved in this way: Let $\mathbf{n} = (n_1, ..., n_K)$ and $\mathbf{n}^i = (n_1, ..., n_{i-1}, n_i + 1, n_{i+1}, ..., n_K)$ and suppose that $g_i(\mathbf{n}) = n_i + (\lambda/K)(1 - x)$ and $g(\mathbf{n}) = n + \lambda(1 - y)$. Then we have $g_i(\mathbf{n}^i) = n_i + 1 + (\lambda/K) \cdot [1 - x \cdot (n + \mu)/(n + 1 + \mu)]$ and $g(\mathbf{n}^i) = n + 1 + \lambda \cdot [1 - y \cdot (n + \mu)/(n + 1 + \mu)]$ and hence $f_i(\mathbf{n}^i) > f_i(\mathbf{n})$. The Reichenbach axiom is an immediate consequence of the boundedness of $|g_i(\mathbf{n}) - n_i|$. That the inequality (9) and hence the condition (vi) hold, follows from (16). Finally, it is clear that some sort of analogy or disanalogy influence is systematically built in into (17). To make more vivid how it operates, let me give a numerical example which also indicates that the three adequacy conditions of Hesse (1964) (cf. Niiniluoto p. 10) are in fact satisfied by (17).

Let the language contain 4 logically independent primitive predicates M_1, M_2, M_3, and M_4. Thus there are $K = 16$ Q-predicates; set, in particular, $Q_1 = M_1 \wedge M_2 \wedge M_3 \wedge M_4$. $\bar{M}_i$ is to denote the negation of M_i, and expressions like $M_1 a \wedge M_2 a \wedge \bar{M}_3 a$ are abbreviated by $M_1 M_2 \bar{M}_3 a$.

Let the distance function among the Q-predicates be given by Niiniluoto's equation (21) (p.   ) so that $d_{ij} = 0, 1, 2, 3,$ or 4. Choose finally $\lambda = 32$, $\mu = 2$, and $h(d, m) = 1 - 12/(12 + md)$. Then we get the following numerical results (where $a = a_1$ and $b = a_2$):

(18)
$$P(Q_i b) = .0625 \text{ for all } i,$$
$$P(Q_1 b | Q_1 a) = .0957,$$
$$\text{if } d_{1i} = 1, \text{ then } P(Q_i b | Q_1 a) = .0620,$$
$$\text{if } d_{1i} = 2, \text{ then } P(Q_i b | Q_1 a) = .0604,$$
$$\text{if } d_{1i} = 3, \text{ then } P(Q_i b | Q_1 a) = .0590,$$
$$\text{if } d_{1i} = 4, \text{ then } P(Q_i b | Q_1 a) = .0578;$$

furthermore,

(19)
$$P(M_4 b) = .5,$$
$$P(M_4 b | M_1 M_2 M_3 M_4 a \wedge M_1 M_2 M_3 b) = .6070,$$
$$P(M_4 b | M_1 M_2 M_3 M_4 a \wedge M_1 M_2 \bar{M}_3 b) = .5064,$$
$$P(M_4 b | M_1 M_2 M_3 M_4 a \wedge M_1 \bar{M}_2 \bar{M}_3 b) = .5057,$$
$$P(M_4 b | M_1 M_2 M_3 M_4 a \wedge \bar{M}_1 \bar{M}_2 \bar{M}_3 b) = .5051,$$
$$P(M_4 b | M_1 M_2 M_4 a \wedge M_1 M_2 b) = .5631,$$
$$P(M_4 b | M_1 M_4 a \wedge M_1 \bar{M}_2 b) = .5061,$$
$$P(M_4 b | M_1 M_2 M_4 a \wedge \bar{M}_1 \bar{M}_2 b) = 5054,$$
$$P(M_4 b | M_1 M_4 a \wedge M_1 b) = .5367,$$
$$P(M_4 b | M_1 M_4 a \wedge \bar{M}_1 b) = .5058,$$

and finally, if M is any logical combination of M_1, M_2, and M_3,

(20)
$$P(M_4 b | M \bar{M}_4 a \wedge M b) = P(\bar{M}_4 b | M M_4 a \wedge M b)$$
$$= 1 - P(M_4 b | M M_4 a \wedge M b).$$

(19) exemplifies Hesse's conditions I and II, and with (20) we get her condition III.

These numbers are perhaps not very impressive. But this is due to a general feature of (17), namely that the impact of analogy considerations is rather weak according to (17), because it soon becomes dominated by the observed absolute frequencies. In fact, I have helped a bit by making λ large and μ small. But this is not so important now; what matters here, is

that this impact is qualitatively acceptable; and this seems to be the case and does not depend on the choice of λ and μ (it does depend, however, on the choice of the function h).

I have to add two remarks, a short and a longer one. First, let me say – though this is the sort of unredeemed claim one should not make in papers – that I suspect that the functions g_i as given by (17) (and also by (25) below) are symmetrizable, i.e. that it is possible to modify them in such a way that they keep their desired properties and, in addition, satisfy Carnap's quotient condition so that the resulting probability measure is symmetric. (But if so, then *not* because there would be something special about (17) and (25).) I am very unsure, however, whether the tremendous intricacy of verifying this is outweighed by its philosophical usefulness.

Secondly, while thinking about this note, I wondered why only general statements like 'there is no Q_i' etc. should have positive probability. Is it really reasonable to expect with certainty after having observed a Q_i-individual that there are other, in fact infinitely many Q_i-individuals? According to the K-dimensional system and also according to (17), yes. There we have

$$(21) \qquad P(\wedge x(x \neq a_1 \rightarrow \neg Q_i x) \mid Q_i a_1) = 0$$

– for the same reasons for which the λ-continuum failed to satisfy (6). Intuitively, however, I tend to deny this. Would you not agree that there is only one Beethoven (meaning that there was, is, and will be only one as great a genius as Beethoven, namely Beethoven himself)? Or to take a more scientific example: It may be presumptuous, but is it *unreasonable* to believe (at least with good chance) that there is only one place for consciousness in the universe, namely earth – even if we would still believe in an infinite universe? I think not, and therefore I finally want to point out that there is a mathematical remedy for (21):

The basic idea is similar to that leading to (17). The only difference is that now the amount by which Q_i gets fined (in comparison to (5)) has to eat up not only the initial logical weight λ/K in case Q_i remains unexemplified, but also the a posteriori weight $n_i + \lambda/K$ in case Q_i remains exemplified exactly n_i times. Thus, let φ be a real-valued two-place function such that

$$(22) \qquad 0 \leq \varphi(m, n) \leq 1$$

and such that for any non-decreasing sequence (m_n) of nonnegative integers with $m_n \leq n$ for all n:

(23)   if $\displaystyle\lim_{n \to \infty} \frac{m_n}{n} > 0$, then $\displaystyle\lim_{n \to \infty} \varphi(m_n, n) = 1$,

(24)   and if there is an m with $m_n \leq m$ for all n, then not only $\displaystyle\lim_{n \to \infty}$

$$\varphi(m_n, n) = 0, \text{ but even } \sum_{n=1}^{\infty} \frac{1}{n} \varphi(m_n, n) < \infty.$$

For example, $\varphi(m, n) = 1 - (1 - m/n)^m$ is such a function. Then we might modify (17) to

(25) $\displaystyle g_i(\mathbf{n}) = n_i \cdot \varphi(n_i, n) + \frac{\lambda}{K} \cdot \left[1 - \sum_{j \neq [} \frac{n_j + \mu/K}{n + \mu} \cdot h(d_{ji}, n_j) \right].$

These already very complicated characteristic functions not only satisfy the conditions (i) and (iii)–(vii), but also give positive probability to all contingent statements of the form 'there are exactly n_i Q_i-individuals and exactly n_j Q_j-individuals and...'. However, I am not going to argue this in all mathematical detail, since, again, the point was rather to demonstrate a mathematical possibility and not something of great philosophical or even practical relevance which deserves much scrutiny.

GENERAL REMARKS

As shown by the different attempts to construct an analogical inductive logic, there are various mathematical alternatives for accommodating analogy influences in *some* more or less general way which preserve other desirable properties of the confirmation function to a greater or lesser extent. But in what way *should* analogy influences be accommodated? So far, I think, there is no very helpful answer; there is little intuitive and philosophical guidance for choosing among the mathematical possibilities. To demonstrate this skepticism: Niiniluoto takes the three adequacy conditions of Hesse (1964) as an apparently unquestioned starting point for his own investigations. But in my view they are questionable, and the

following list of problems might make this clear (where I shall assume for the sake of specificity that we have to deal with k primitive monadic predicates $M_1, \ldots, M_k$):

(1) As Hesse presents her conditions and as Niiniluoto adopts them, they refer only to the beginning of the inductive process, i.e., only to the first two individuals to be observed. If so restrictedly interpreted, they leave us with no constraint at all for the effects of analogy influences after some evidence is obtained. One might reinterpret them so as to hold conditionally on any fixed evidence; but this does not really help, as we shall see.

(2) Hesse's conditions I and II are provided with an important and plausibly looking clause referring to the respective weights of the properties lending positive and negative analogy. But this clause is vague. What are these weights? Within the restricted interpretation, there is still a good answer to this (which is implicitly adopted by Niiniluoto): choose the weight of a property reciprocal to its width (in the technical sense). But this answer does not do for the generalized interpretation, which is, therefore, in need of concretization.

(3) One might perhaps wonder why this clause appears in condition I, but not in condition III. The following observation is relevant to this: The clause qualifying condition I is simply neglected by Niiniluoto's proposals as well as by mine. (In all these proposals we have, e.g., $P(M_k b) = 1/2$ and $P(M_k b \mid M_k a \wedge A) > 1/2$, where A is any evidence about analogies and/or disanalogies between a and b; cf. the equations (19) here and the equations (33) and (34) of Niiniluoto.) But there is an inherent reason for this neglection, namely the validity of (20) (or a generalized version thereof) within these proposals; this implies that $P(M_k b \mid M_k a \wedge A) > 1/2$ is equivalent to $1/2 > P(M_k b \mid \bar{M}_k a \wedge A)$ and hence to $P(M_k b \mid M_k a \wedge A) > P(M_k b \mid \bar{M}_k a \wedge A)$. In other words: These proposals satisfy the unqualified condition I, if and only if they satisfy the unqualified condition III. And this is true of any probability measure P for which $P(M_k b) = 1/2$ and (20) hold; and, intuitively, (20) should hold, since it looks very convincing that M_k and $\bar{M}_k$ are interchangeable in a situation as given in (20).

(4) So, should one omit the clause of condition I or add it to condition III? I have no clear intuitions about this. One might say: 'However small the similarities between a and b and however large their differences, a bit more similarity is still more likely than more differences.' Or one might

say: 'If the dissimilarities outweigh the similarities to too high a degree, then further dissimilarity in the remaining respects is more likely than similarity.' Metaphorically speaking: Does analogy influence act like gravitational force or rather like electromagnetic force (please, be so lenient to assume for the sake of the metaphor that equal charges attract and opposite charges repel one another)?

(5) There is a similar unclarity about condition II. Those who subscribe to it would say: 'If b is very similar to a, then b should be similar to a in the remaining respects, too; but if there is already a lot of dissimilarity between a and b, who cares about a bit more dissimilarity?' However, one might say as well: 'A little dissimilarity is always to be expected; so if b is already very similar to a, then the remaining respects are undecided. But total dissimilarity is very unlikely; if a and b are very dissimilar, they should be similar at least in the remaining respects.' Or to take another physical analogy: Does analogy influence act like gravitational force (which decreases with increasing distance) or like the restoring force of, say, a spring (which increases with increasing distance)? (Niiniluoto's proposals work according to the first alternative. My proposal is open in this respect; it depends on how the function h is chosen.)

To state (4), (5), and (1) in a nutshell: On which model exactly do we expect analogy influences to operate? And how exactly do they mix with observed frequencies? I fear that intuitions will vacillate and tend to various sides depending on the cases at hand and that no very definite and general principles concerning these questions will emerge from philosophical discussion.

These observations and the softness of the reasoning in section 2 bring me to my last point. I have the impression that by considering analogy influences, inductive logic changes its character in important respects. To speak somewhat vaguely (this will suffice): As always intended by Carnap and as expressed by its name, inductive logic is a rather aprioristic enterprise, hoping to arrive at a small or at least easily describable class of confirmation functions which are distinguished by some sort of objectivity or rationality. But now, with analogy influences taken into account, something like objectivity or just widely agreed reasonableness seems hard to reach, as the last paragraphs suggest. And the hope of getting a small class of confirmation functions now seems Utopian; the λ-continuum contained one free parameter, in the K-dimensional system there are K free para-

meters, and according to (17), which already embodies some arbitrary decisions, we can freely choose $K(K\text{-}1)/2$ distances, two parameters, λ and μ, and a whole function, h.

An aprioristic procedure seems questionable also on general grounds. Analogy, in its full sense only very meagerly captured in formal models as above, is a highly a-posteriori-matter; it is concerned with often rather vague considerations (or should we say: feelings?) of how theories in one empirical field might be carried over to another field; and passable intuitions about concrete analogies only evolve after a thorough-going examination of the subject matter at hand. This is not to say that Carnap's aprioristic program as characterized, e.g., by his transition from credence functions back to credibility and confirmation functions (cf. Carnap, Jeffrey (1971), pp. 13–27), would now become absurd in principle. It is only to say that it is scarcely feasible to generate all the complicatedly balanced analogy considerations from some initial distribution, that analogy is a long, perhaps too long way for rational reconstruction. And, in fact, all attempts to an analogical inductive logic were more or less openly determined not by the old Carnapian spirit, but by a very pragmatic attitude, content with looking for some nice desiderata and for ways to satisfy as many of them as possible.

All this indicates, I think, that, when considering analogy influences, we are thrown back to the subjectivistic position (within which mathematical investigations like Niiniluoto's may still be useful, but merely show what is and is not possible for a subjectivist). Whether this should be deplored or not – this belongs to another, large-scale battle-field.

Seminar für Philosophie, Logik and Wissenschaftstheorie
Universität München

BIBLIOGRAPHY

Carnap, R., Jeffrey, R. C., (eds.), *Studies in Inductive Logic and Probability*, vol. I, Berkeley: University of California Press, 1971.

Hesse, M., 'Analogy and Confirmation Theory', *Philosophy of Science 31* (1964), 319–327.

Hintikka, J., Niiniluoto, I., 'An Axiomatic Foundation for the Logic of Inductive Generalization', in: M. Przełecki, K. Szaniawski, R. Wójcicki (eds.), *Formal Methods in the Methodology of Empirical Sciences*, Dordrecht: D. Reidel, 1976, pp. 57–81.

v. Mangoldt, H., Knopp, K., *Einführung in die höhere Mathematik*, **2**. *Band, Stuttgart: Hirzel*, ¹²*1962*.

Niiniluoto, I., 'On a *K*-dimensional System of Inductive Logic', in: F. Suppe, P. D. Asquith (eds.), *PSA 1976*, vol. 2, East Lansing: Philosophy of Science Association, 1977, pp. 425–447.

Niiniluoto, I., 'Analogy and Inductive Logic', *Erkenntnis* **16** (1981) 1–34.