SINGULAR ANALOGY AND QUANTITATIVE INDUCTIVE LOGICS

John R. WELCH*

Manuscript received: 1998.4.21.
Final version: 1999.3.9.

* Saint Louis University, Avenida del Valle 34 y 28, 28003 Madrid. E-mail: welchj@spmail.slu.edu

BIBLID [0495-4548 (1999) 14: 35; p. 207-247]

ABSTRACT: The paper explores the handling of singular analogy in quantitative inductive logics. It concentrates on two analogical patterns coextensive with the traditional argument from analogy: perfect and imperfect analogy. Each is examined within Carnap's λ-continuum, Carnap's and Stegmüller's λ-η continuum, Carnap's Basic System, Hintikka's α-λ continuum, and Hintikka's and Niiniluoto's *K*-dimensional system. It is argued that these logics handle perfect analogies with ease, and that imperfect analogies, while unmanageable in some logics, are quite manageable in others. The paper concludes with a modification of the *K*-dimensional system that synthesizes independent proposals by Kuipers and Niiniluoto.

Keywords: inductive, inductive logic, probability, analogy, perfect analogy, imperfect analogy.

CONTENTS

1. Introduction

When is it rational to be persuaded by an argument from analogy? One consideration would have to be logical form. Since arguments from analogy are not deductively valid, it would seem natural to require that they satisfy an inductive criterion. But what would the criterion stipulate? As a

Novel use is made of several formulas, generalizations of an idea of Carnap, for rapid calculation of the probability of certain analogies. The formulas emerge in the assessments of Carnap (1952) in Section 3c; of Hintikka (1966) in Section 6c; and of Hintikka and Niiniluoto (1976) in Section 7c.

2. *Types of Analogy*

Distinctions are especially wanted when it comes to analogy for, as J.S. Mill observes, "There is no word (...) which is used more loosely, or in a greater variety of senses, than Analogy" (1974, p. 554). If we begin with the obvious divide between *general analogies*, which include at least one quantified sentence, and *singular analogies*, which have no such sentences, we can focus on the latter, subdividing as necessary in a kind of Porphyrian tree. Locating the critical joint among singular analogies requires some attention to the root concept of similarity. Even a cursory review of the literature on analogy reveals that the relata of the similarity relation are not all of the same type. What is called analogy in some places features similarity among individuals, but other analogies are based on similarity among properties. An example of the former is the traditional argument from analogy:

A1: $Fa \wedge Ga$.
Fb.
So Gb.

Contrast A1 with the following argument, discussed in Pietarinen (1972, pp. 68-69) and elsewhere:

A2: $Fa \wedge Ga$.
$-Fb$.
So Gb.

The striking thing about A2 is that its premises show no similarity between the individuals *a* and *b*. It capitalizes instead on the similarity between *a*'s property *FG* and *b*'s inferred property $\overline{F}G$. One might then be inclined to posit two types of singular analogy: individual analogy for arguments like A1 and property analogy for those like A2. This would be premature, I believe, for two reasons. First of all, there are singular analo-

Developing the concept of broad analogy requires Carnap's distinction between *analogy by similarity* and *analogy by proximity* (1980, pp. 40-41, 68-71). As an example of the former, suppose that a small sample discloses individuals that are *FG* but none that are $F\overline{G}$ or $\overline{FG}$; nevertheless, the similarity relations among these properties would make it seem more probable that the next individual is $F\overline{G}$ rather than $\overline{FG}$.[2] Analogy by proximity occurs when the order of observation affects probability.[3] Suppose that a certain individual is known to instantiate a certain predicate; if order makes a difference, the probability that the next (most proximate) individual also has the predicate is assumed to be greater than the probability that the twentieth individual, say, has it. Both analogy by similarity and analogy by proximity have subspecies; the former is divided into *existential* and *enumerative* types in Niiniluoto (1988), and the latter branches into *proximity in the past* and *proximity in the future* in Kuipers (1988).

Like broad analogy, narrow analogy comes in more than one form. I propose to revive and reshape a distinction that appeared early on in the debate over quantitative inductive logic. Though this distinction was being drawn by both Hesse (1963, p. 121; 1964, pp. 320, 326) and Achinstein (1963, p. 216) at about the same time, the terminology I shall adopt is due to the latter. A *perfect analogy*, in Achinstein's sense, "attributes to an individual all of the properties which the observed individual is known to have" (1963, p. 216). Our A1 is an instance. An *imperfect analogy*, on the other hand, attributes to an individual only some of the properties of the observed individual, as in A2 and A3.

Though the imperfect-perfect distinction will be handy here, recasting it somewhat is necessary. To see why, notice Achinstein's claim that

> (...) the usual case of analogy, if not indeed what is meant by a case of analogy, is one in which an individual *b* mentioned in the evidence has many, *but not all*, of the properties of the individual *c* mentioned in the hypothesis. (1963, p. 216)

Likewise Hesse, criticizing Carnap on perfect analogy, asserts that

> (...) this type of argument is not what has been traditionally understood by argument from analogy, since analogical inference has always supposed *differences* as well as similarities between the two analogues. (...) That is to say the assumption, made in Carnap's type of inference, that the evidence ascribes to the individuals only the same property P_1 in both cases, and that there are not initially known to be any differences between them, is at best an idealization of the real situation. It will generally be the case that, if the total evidence is taken into account, superfi-

perfect case are not. Hence there is nothing unrealistic or idealized about perfect analogy; it too takes differences into account.

The preceding typology of singular analogy can be summarized as follows:

- singular analogy
  - broad
    - similarity
      - existential
      - enumerative
    - proximity
      - in the past
      - in the future
  - narrow
    - perfect
    - imperfect

Some of these kinds of singular analogy have been more conspicuous than others, and they will continue to be in this paper. Carnap thought that the similarity influences registered via analogy by similarity and analogy by proximity "have only secondary significance" (1980, pp. 41, 66, 70). In addition, some have expressed doubts "about whether the idea of analogy by proximity is after all very important as such" (Kuipers 1988, p. 311). However this may be, analogy by proximity is far removed from our present concerns and will be discussed no further here. Moreover, the other type of broad analogy, analogy by similarity, appears rather late in the literature on quantitative inductive logics. It seems to have come into focus only when it was noticed that certain inductive logics are more successful with some narrow analogies than with others. Pursuing analogy by similarity was a way of trying to fix that. Historically, then, narrow analogy was first; it includes the arguments from analogy discussed by Mill (1974, pp. 554-561). We will respect this priority here, concentrating on narrow analogy and treating broad analogy by similarity only as needed.

3. Carnap's λ-Continuum

Since the λ-continuum is a point of departure for later inductive systems, the following discussion develops its salient features. There are three subsections: a) contours of the λ-continuum; b) applying the λ-continuum to narrow analogy; and c) assessment of narrow analogy in the λ-continuum.

ple to the sample's total number of individuals. For molecular predicates, this ratio can be expressed as n_M / n; for Q-predicates, as n_Q / n. The logical factor is relative width, the coverage (so to speak) of the instantiated predicate relative to the total number K of L's Q-predicates. Because molecular predicates can always be analyzed into some number w of Q-predicates, their width is w and their relative width is w/K. Q-predicates, which are a special case. have width of 1 and relative width of $1/K$.

The values of the empirical and logical factors thus establish an interval somewhere between 0 and 1 inclusive, and $c(h,e)$ is to be located between or on the interval's endpoints. Exactly where is determined by identifying $c(h,e)$ in certain key situations with the value of a mathematical function: the weighted mean of the empirical and logical factors. If the weight of the empirical factor is conventionally set to n, the total number of individuals in the sample, then the value of the function is governed by the weight of the logical factor -a particularly simple form of the mean. Carnap calls this logical weight 'λ'. λ's value, which can be any real number from 0 to ∞ inclusive, is equal in weight to the observation of the same number of individuals. Its different values demarcate different inductive methods within the eponymous continuum.

Suppose now that observation of a determinate number of individuals yields evidence e_Q, which states no more than whether each observed individual has or does not have a Q-predicate 'Q'.[7] Suppose also that a hypothesis h_Q attributes 'Q' to an unobserved individual on the basis of e_Q. If it is further assumed that λ can vary with K but not with n_Q and n, then the desired degree of confirmation is given by the following expression:

$$c(h_Q, e_Q) = \frac{n_Q + \dfrac{\lambda(K)}{K}}{n + \lambda(K)} \quad . \qquad (1)$$

More generally, let the evidence e_M say only which individuals of a sample instantiate a molecular predicate 'M' and which do not, and let a hypothesis h_M ascribe 'M' to an unobserved individual given e_M. Then

$$c(h_M, e_M) = \frac{n_M + \dfrac{w\lambda(K)}{K}}{n + \lambda(K)} \quad . \qquad (2)$$

$$\mu(t) = \frac{\prod_{Q} [\frac{\lambda}{K} (1 + \frac{\lambda}{K}) (2 + \frac{\lambda}{K}) \dots (n_Q - 1 + \frac{\lambda}{K})]}{\lambda (1 + \lambda) (2 + \lambda) \dots (N - 1 + \lambda)} \quad (5)$$

for the measure of a state description in *L*.[8]

The second step is to extend this measure function so as to admit not just state descriptions but any sentence of *L* as argument. This is easily accomplished, however, since any sentence *s* of *L* that is not a state description provides less than the full description of the world supplied by a state description. It is therefore equivalent to a disjunction of more than one state description or, if it is logically false, to the negation of a disjunction of all state descriptions (Carnap, 1950, pp. 289-90; 1952, pp. 11, 18). Moreover, since state descriptions are mutually exclusive, Carnap's addition axiom (1952, p. 12) stipulates that the μ-value of *s* is equal to the sum of the μ-values of its component state descriptions. (5), therefore, provides the μ-value of any sentence *s* of *L* that is not logically false. If *s* is logically false, its μ-value is of course 0.

The final step is the general definition of $c(h,e)$ in terms of μ-values. For any sentences h,e where $\mu(e) \neq 0$,

$$c(h, e) = \frac{\mu (e \wedge h)}{\mu (e)}. \quad (6)$$

(6) can be viewed as an instantiation of the classical definition of conditional probability.

b) Applying the λ-continuum to narrow analogy

Once $c(h,e)$ is fully defined, it can be turned to specifics like analogy. Some extrapolation from earlier works is unavoidable, however, since Carnap does not treat the topic in (1952). In announcing what was to become the λ-continuum's central method, c^*, in (1945), he describes the inference by analogy as follows:

> The evidence known to us is the fact that individuals *a* and *b* agree in certain properties and, in addition, that *a* has a further property; thereupon we consider the hypothesis that *b* too has this property. (...) The hypothesis *h* says that *b* has not only the properties ascribed to it in the evidence but also the one (or several) ascribed in the evidence to *a* only, in other words, that *b* has all known properties of *a*. (...) (1945, p. 87)[9]

Multiple perfect analogies can be treated by adjusting the empirical factor as in (7).

Nor does Carnap observe that (7) can also be adapted to cases where the evidence is mixed in the sense that more than one Q-predicate is known to be instantiated. Care must be taken, however, so that any part of the evidence concerning predicates logically impossible for the partially known individual of the conclusion to instantiate is excluded from the empirical part of the formula.[11] Consider, for example, the following evidence: *a* and *b* are FG, *c* is $F\bar{G}$, *d* is $\bar{F}G$, and e is F. In calculating the probability of the hypothesis that *e* is G, the evidence concerning *d* should be excluded from the variant of (7) since it is incompatible with what is known about *e*. Thus the empirical factor would be 2/3 with FG, not 2/4, as can be verified with the characteristic function. Adhering to this proviso on evidence, then, (7) can be stated more generally. Suppose that n individuals i ($i = a, b, \ldots, y$) have been examined and that n_1 have M_1, a property of width w_1 attributed to the $n + 1$st individual z by the analogy's conclusion. Suppose also that n_2 have M_2, a property of width $w_2 > w_1$ that z is already known to have. Then where E_i is the conjunction of all the evidence about the individuals i, the degree of confirmation in c^* of the analogical hypothesis that z is M_1 can be calculated via (6) and (2) as:

$$c^* (M_1 z, E_i \wedge M_2 z) = \frac{(n_1 + w_1) \,/\, (n + K)}{(n_2 + w_2) \,/\, (n + K)} = \frac{n_1 + w_1}{n_2 + w_2}\,. \tag{9}$$

Just as (7) was generalized as (8) for the entire λ-continuum, (9) can be likewise expanded. Appealing once again to (6) and (2) yields the λ-continuum's version of (9):

$$c (M_1 z, E_i \wedge M_2 z) = \frac{n_1 + (w_1 \dfrac{\lambda (K)}{K})}{n_2 + (w_2 \dfrac{\lambda (K)}{K})}\,. \tag{10}$$

Though (10) holds all across the continuum, it has simpler special forms. For Carnapian methods of the first kind, '$\lambda(K)$' reduces to 'λ' in both numerator and denominator, and for c^*, of course, (10) reduces to (9). The advantages of (10) are the advantages of (9) but magnified: easy yield of

To introduce the first, let $\mathfrak{F}_1$ be a family of k_1 primitive predicates, and let $\mathfrak{F}_2$ be a family with k_2 such predicates. The Q-predicates Q_{ij} ($i = 1, 2,..., k_1$; $j = 1, 2,..., k_2$) formed from the predicates of these families comprise a pseudo-family $\mathfrak{F}_{1,2}$ with $k_1 k_2 = K$ members. Say that each instantiated Q-predicate is true of n_{ij} individuals. Then the measure of a state description t is calculated as in the λ-system for a language with $k_1 k_2$ Q-predicates. Thus the measure is given by (5) adapted to this special case:

$$\mu_\lambda^{1,2}(t) = \frac{\prod_{i=1}^{k_1}\prod_{j=1}^{k_2}[\frac{\lambda}{K}(1+\frac{\lambda}{K})(2+\frac{\lambda}{K})\ldots(n_{ij}-1+\frac{\lambda}{K})]}{\lambda(1+\lambda)(2+\lambda)\ldots(N-1+\lambda)} \quad . \qquad (11)$$

(11) provides a Q-measure, as I shall say, for it is keyed to state descriptions couched in terms of Q-predicates.

The second candidate is a P-measure, based upon state descriptions structured by primitive predicates. Here the idea is first to calculate the measure of each family of k primitive predicates with (5) as if it were a language with k Q-predicates in the λ-system, and then to take the product of the measures for each family. Where the measure of the distribution of L's individuals relative to the first family is μ_λ^1, and the measure of the same individuals' distribution relative to the second family is μ_λ^2, the measure of a state description t can be expressed as

$$\mu_\lambda^{1/2}(t) = \mu_\lambda^1 \times \mu_\lambda^2 \ . \qquad (12)$$

How well do these measure functions deal with imperfect analogies? Let us see. Consider the three state descriptions:[12]

(t_1) $FGw \wedge \overline{FG}x \wedge \overline{F}Gy \wedge \overline{FG}z$

(t_2) $FGw \wedge FGx \wedge \overline{FG}y \wedge \overline{FG}z$

(t_3) $FGw \wedge FGx \wedge F\overline{G}y \wedge F\overline{G}z$

For t_1 and t_2 it should turn out that

Moreover, she has shown that this generalized measure function satisfies certain general conditions (1964, pp. 322-23), and that this is sufficient to ensure reasonable confirmation values for imperfect analogies like Achinstein's rhodium example (1964, pp. 325-26).

Nevertheless, Hesse has also pointed out that this technique for imperfect analogies seems strangely imperfect. She objects to the "somewhat arbitrary and *ad hoc*" nature of Carnap's and Stegmüller's η-solution (1964, p. 325), and I concur. Although the λ-η system is a technically acceptable patch of the λ-continuum, returning satisfactory values for imperfect as well as perfect analogies, it provides no guidance on the choice of a value for η. Why one would shift η towards 0 or towards 1 is apparently to be decided on the spur of the moment.

5. Carnap's Basic System

The tenor of Carnap's posthumously published Basic System (1971, 1980) is caught by his conjecture that it is sufficient to base inductive logic on two magnitudes: the width and distance of properties (1980, p. 29).[13] Width is the logical width that figures so prominently in the λ-continuum. Distance, which Carnap conceives by "analogy to the dependence of a physical effect of one body on another upon the distance between the bodies," is similarity among properties (1980, p. 48). Each magnitude receives a parameter: γ for width and η for distance. In addition, λ is carried over from (1952) to represent logical weight. If a definite value has been chosen for λ, say λ^*, it can be used to determine the value of η, η^*:

$$\mu^* = \frac{\lambda^*}{\lambda^* + 1} . \qquad (15)$$

In the latter sections of the Basic System, in fact, λ displaces η and thereafter functions as the "main parameter" (1980, p. 93).

The result is a system which, without abandoning the λ-continuum, exhibits a number of critical differences. The axiomatic base has been simplified, first of all (1980, pp. 105-6). Moreover, both the λ-continuum's methods of the first kind and the second kind, including c^*, are regarded as inadequate *general* rules. The reason is that both assign the same λ-value to all predicate families of the same size, but Carnap now prefers to key λ to distance rather than size (1980, pp. 115-119). In addition, whereas the earlier system permitted λ to take values of 0 and ∞, the Basic System the Basic System imposes η-equality: all pairs of distinct predicates within a family are treated as equally similar (1980, p. 57). Despite Carnap's initial successes with perfect analogy, then, his bequest to inductive logic included analogy as a largely unsolved problem.

6. Hintikka's α-λ Continuum

Like the discussion of Carnap (1952), this section will be divided into three subsections: a) contours of the α-λ continuum; b) extending the α-λ continuum to narrow analogy; and c) assessment of narrow analogy in the α-λ continuum.

a) Contours of the α-λ continuum

Since the α-λ continuum of Hintikka (1966) has Carnap's λ-system as a special case, Hintikka can be said to take up the project of quantitative inductive logic where Carnap left it in (1963). Even so, there are major differences of approach. Carnap defends a logical interpretation of probability, for example, whereas Hintikka is more Bayesian, maintaining that there is no way to determine the values of inductive parameters like λ on strictly logical grounds (1969, pp. 38-40; 1970, pp. 23-25). Carnap's λ-continuum assigns zero probability to all generalizations in an infinite domain -a result most have found unacceptable, and which is absent from Hintikka's systems. Carnap keys on singular inductive inference (1952, p. 13), but Hintikka argues that, to prevent overdependence on domain, the focus should be on inductive generalization instead (1965b, p. 279).

Consequently, even though the focus of this study is analogy as singular inductive inference, inductive generalizations cannot be avoided in Hintikkan systems, for the values they assign to singular inductions are determined in part through values for inductive generalizations of a special sort. As its name indicates, the α-λ continuum is structured by two special parameters. Together, they influence both singular and general induction, but the parameter λ is like Carnap's λ, acting first and foremost on singular induction, whereas α's immediate effects are on inductive generalization. Although λ can be infinite and α cannot, α is comparable to λ in that it represents a priori considerations. Hintikka views it as an index of caution: the more irregularity we expect in the universe, and hence the slower we are to jump to lawlike conclusions, the higher α will be (1970, p. 21).

In the special case of Hintikka's continuum, $p(e)$ is therefore:

$$\sum_{i=0}^{K-c} \binom{K-c}{i} \; p\,(C_{c+1})\; p\,(e \mid C_{c+1}) \;, \tag{18}$$

where c is the number of Ct-predicates known to be instantiated, and the range of values for i permits the representation of the alternative constituents compatible with the evidence. Thus what we need are values for $p(C_w)$ and $p(e|C_w)$. In seeing how they are determined, we will follow Hintikka in assuming an infinite universe.[17]

Initially, let us consider $p(C_w)$. The methods of the α-λ continuum fix the prior probability of an arbitrary constituent of width w as follows. Where $\pi(a,z) =_{df} z \cdot (z+1) \cdot \ldots \cdot (z+a-1)$ if $a = 1, 2, 3,\ldots$ and $\pi(0,z) =_{df} 1$,

$$p\,(C_w) = \frac{\pi\,(\alpha, \dfrac{w\lambda}{K})}{\displaystyle\sum_{i=0}^{K} \binom{K}{i}\; \pi\,(\alpha, \dfrac{i\,\lambda}{K})} \,. \tag{19}$$

If λ is a constant, (19) anchors a Hintikkan version of Carnap's methods of the first kind. But if, as in Carnap's methods of the second kind, λ is a function $\lambda(K)$ of K, then (19) becomes

$$p\,(C_w) = \frac{\pi\,(\alpha, \dfrac{w \cdot \lambda(K)}{K})}{\displaystyle\sum_{i=0}^{K} \binom{K}{i}\; \pi\,(\alpha, \dfrac{i \cdot \lambda(K)}{K})} \,. \tag{20}$$

When $\alpha = 0$, all constituents receive equal prior probabilities. When $\alpha > 0$, however, constituents of different widths are given unequal prior probabilities, though the $\binom{K}{w}$ constituents having the same width all receive the same a priori weight.[18]

The second component of $p(e)$ is $p(e|C_w)$, the conditional probability of the evidence given a constituent. These probabilities are based upon the instantiation of primitive predicates provided by premises like the second.

The fusion of Ct- and Pt-evidence within α-λ can be accomplished by first restating the Pt-evidence in complete disjunctive normal form.[21] The result is a disjunction with the structure

$$e_{Q1} \vee e_{Q2} \vee \ldots \vee e_{Qn}, \tag{24}$$

each clause of which attributes a Ct-predicate to an object. Conjoining (24) to the Ct-evidence e_C yields

$$e_C(e_{Q1} \vee e_{Q2} \vee \ldots \vee e_{Qn}), \tag{25}$$

or, distributing,

$$(e_C \wedge e_{Q1}) \vee (e_C \wedge e_{Q2}) \vee \ldots \vee (e_C \wedge e_{Qn}). \tag{26}$$

Since (26) is Ct-homogeneous, (23) can now be applied. There are two differences compared to its application to Ct-evidence alone, however. The major difference is that it has to be applied more than once: once for each clause of (26). As these clauses are mutually exclusive, the probability of (26) is simply the summation of these repeated applications of (23). The minor difference is an adjustment for the partially observed $n + 1$st individual. A different Ct-predicate is attributed to this individual in each of (26)'s clauses; if the projected predicate is 'Q_1', the clause states that $n_1 + 1$ of $n + 1$ individuals, rather than n_1 of n, possess it, which makes the various n_q in (22) sum to $n + 1$. With these differences in mind, let x be the number of predicate letters needed to transform the Pt-evidence into complete disjunctive normal form, and 2^x the number of disjuncts d of the evidence (26), each of which says that c_d Ct-predicates are instantiated. Then $p(e|C_w)$ for Pt-evidence is

$$\sum_{d=1}^{2^x} \frac{\prod_{Q=1}^{c_d} \pi\left(n_Q, \frac{\lambda_{,}}{w}\right)}{\pi(n+1, \lambda)} \tag{27}$$

nap's system is a special case of Hintikka's, as remarked. Moreover, and more important for our purposes, there is a close connection between Carnap's c^* and Hintikka's generalized combined system (GCS), which is the point along the α-λ continuum where $\lambda = \lambda(w) = w$ except in (19) and (20), where $\lambda = \lambda(K) = K$ for prior probabilities of constituents.[22] The connection is that, as α grows without bound, the values obtained for singular inference in GCS approach those of c^* (1966, p. 128). For perfect analogies, this means that (9), the shortcut for perfect analogies in c^*, provides limit values for perfect analogies in GCS as $\alpha \rightarrow \infty$.

Indeed, there is an extensive subclass of cases for which (9) gives the exact value in GCS regardless of the value of α. To see what they are, let us first avail ourselves of some obvious simplifications. We have just seen that the probability of singular analogical hypotheses is determined in the α-λ continuum by the quotient (29)/(28). But the denominators of (29) and (28) cancel immediately. What remains both above and below the line is a summation over a product of the three factors within brackets: an N-component, as I will call it, for the number of constituents of a given width compatible with the evidence; a P-component for the prior probabilities of these constituents; and an R-component based on the appropriate representative function. In short, both numerator and denominator consist of a summation over products NPR. Now since $\lambda(K) = K$ for prior probabilities of constituents in GCS, the P-components in numerator and denominator simplify to $\pi(\alpha, c + i)$. And since elsewhere in GCS $\lambda = \lambda(w) = w$, the α-λ representative function (21) reduces to

$$c\,(h_Q, e_Q) = \frac{n_Q + 1}{n + w}\ , \qquad (30)$$

which permits simplification of the R-components. Finally, it will facilitate matters below to generalize (30) so that it applies not only to molecular predicates with $w = 1$ (Ct-predicates), but also to molecular predicates with $w \geq 1$ such as those that figure in the evidence for analogies. So just as (2) is (1) generalized in c^*, the following expression is (30) generalized in GCS. Let evidence e_M state that n_M of n individuals have a molecular predicate *'M'*, and let h_M be the hypothesis that the next individual will also be *M*. The width of *'M'* is w_M, which is distinct from the width w of the constituent that conditions the representative function as in (21). (30) then becomes

Now the last factor in the expansion of any R^{eh} component gives the probability on the evidence that the individual z of the analogy's conclusion has the Ct-predicate 'M_1'. The numerator of this factor takes the form $n_1 + w_1$, where n_1 is the number of individuals known to have M_1 and w_1 is the predicate's width. Because this numerator is common to all R^{eh} components, it can be factored out of the expression for $p(e \wedge h)$. A parallel argument holds for the last factor in the expansion of any R^e component. The numerator of each such factor is $n_2 + w_2$, where n_2 is the number of individuals known to have 'M_2', a disjunction of Ct-predicates that includes 'M_1', and w_2 is its width. Because this numerator is common to all R^e components, it too can be factored out of the expression for $p(e)$. Where 'R^{eh-}' and 'R^{e-}' are the R-components diminished by factoring, (33) is then

$$\frac{(n_1 + w_1) \quad N_0 P_0 R_0^{eh-} + N_1 P_1 R_1^{eh-} + \ldots + N_{K-c} P_{K-c} R_{K-c}^{eh-}}{(n_2 + w_2) \quad N_0 P_0 R_0^{e-} + N_1 P_1 R_1^{e-} + \ldots + N_{K-c} P_{K-c} R_{K-c}^{e-}} . \qquad (34)$$

In these cases, however, the N, P, and diminished R-components of the numerator are identical to those of the denominator. The result is wholesale canceling of the righthand parentheses; all that remains are the parentheses on the left. So here all the apparent complications of (29)/(28) boil down to a simple application of (9). But the condition on the statement of evidence given above with (9) must be respected as always.

As a quick example, suppose that the Ct-predicates 'FG' and '$F\overline{G}$' are instantiated respectively by a and b, and that a partially known individual c instantiates 'F'. We want the probability of the analogical hypothesis that c is also G. Here both '$e \wedge h$' and 'e' agree on the number of Ct-predicates that are instantiated: $c = 2$. Hence $K - c = 2$ in both numerator and denominator of (32), and the N- and P-components are plainly identical. All that really needs to be shown is that the diminished R-components of the numerator are identical to those of the denominator. The number of factors in the R-components of both $p(e \wedge h)$ and $p(e)$ for analogy is always $n + 1$; both expressions concern the same series of individuals, though from different points of view. The expanded R-components that follow are obtained through repeated applications of (31), and associated with constituents of widths 2, 3, and 4:

A common sort of c-uniform analogy merits mention apart. Besides c-uniformity, these analogies meet two further conditions: $c = K$ and $\lambda = \lambda(K)$. When $c = K$, all Ct-predicates are known to be instantiated; hence $w = K$ in (21), the α-λ continuum's representative function. Then provided $\lambda = \lambda(K)$, the rest of (21) reduces trivially to (1), the comparable λ-continuum function. Here, then, α-λ-systems collapse into λ-systems, and (35) is equivalent to (10).

The α-λ continuum is also like Carnap's earlier system in its handling of imperfect analogy. The λ-continuum does not deal adequately with imperfect analogy, as we have seen, nor does Hintikka's successor system. Hesse points out that the difficulty is the same in both systems: the symmetry of Q-predicates, in Carnap's case, or Ct-predicates, in Hintikka's. This symmetry ensures that for a confirmation function c_0

> (...) the prior probabilities are assigned in such a way that $c_0(Ct_1(a) \wedge Ct_2(b))$, where $Ct_1 \not\equiv Ct_2$, has the same value however similar or different a and b may be; that is, for example, if Ct_1 is $P_1P_2P_3,\ldots, P_k$, it has the same value whether Ct_2 is $P_1 P_2 P_3,\ldots, \overline{P}_k$, or $P_1 \overline{P}_2 \overline{P}_3,\ldots, \overline{P}_k$. Now although Hintikka's system is not formalized in this paper, it is clear that his *confirmation* functions, like Carnap's, are symmetrical with respect to *Ct*-predicates, and it therefore follows that these confirmation functions do not satisfy the analogy criterion. (1968, pp. 221-22)

There are at least two strategies for adjusting the α-λ system so that it can cope with imperfect as well as perfect analogies. One, based upon Carnap's and Stegmüller's approach in the λ-η continuum, is to define a measure function parallel to (13) above that uses an analogy constant like η to mediate between P- and Q-measures. This procedure has been illustrated by Pietarinen in (1972, pp. 91-94).[23] However, we have already noted Hesse's complaint that the Carnap-Stegmüller solution is *ad hoc*, and Niiniluoto makes the same charge against Pietarinen's extension of it to the α-λ system (1981, p. 2).

The other strategy, proposed by Hintikka, is a variant of the α-λ continuum in which the primitive predicates are ordered; the Ct-predicates then turn out to be asymmetrical (1968, p. 228). Hintikka works out this proposal in detail, suggesting three different ways of ordering the primitive predicates (1969, pp. 28-33). Though his main concern is to show that these methods can solve Hempel's and Goodman's paradoxes of confirmation, the extension to analogy has been carried out by Pietarinen (1972, pp. 94-99).[24] This ordering strategy would be applied to Achinstein's rhodium inductive generalization. Hintikka and Niiniluoto add a fourth installment, an axiomatic *K*-dimensional system (hereafter KDS) in (1976).[26]

The axiomatic base of KDS is slender. Like Carnap, first of all, Hintikka and Niiniluoto require a probability distribution that is symmetric (de Finetti's exchangeability) and satisfies the probability calculus. But unlike Carnap, whose characteristic or representative function depends on the sample only for n_Q and n (the number of observed individuals with a given Q-predicate and the total number of observed individuals), KDS's representative function relies on the sample for n_Q, n, and c (the number of instantiated Ct-predicates). Hintikka and Niiniluoto express this second axiom by saying that, whereas the λ-function has the form $f(n_Q,n)$, KDS's function has the form $f(n_Q,n,c)$ (1976, pp. 58-59). The additional argument ensures that the simplest constituent compatible with the evidence receives the highest confirmation in the long run (1976, pp. 60, 73).

In addition to these axioms, KDS includes K free parameters, where K is, as before, the number of Ct-predicates specifiable in the language. The parameters are values for the representative function at $f(0,c,c)$, where $c = 1, 2,..., K$-1, and for $f(1,K+1,K)$.

The parameters and axioms together determine a range of inductive systems. The range of KDS is not coextensive with that of the α-λ continuum, but the two do overlap considerably; GCS, for example, belongs to both α-λ and KDS (1976, pp. 59-60). Kuipers has shown that the systems of KDS "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" (1978a, p. 262).

Hintikka and Niiniluoto make it clear that their results are intended to be primarily qualitative (1976, pp. 60, 73). Commenting on this, Kuipers observes that the systems of KDS "seemed to be extraordinarily complicated," and that "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" (1978a, p. 262). Kuipers proves, however, that the systems of KDS, which he calls 'P-systems', are equivalent to a class of systems he calls 'Q-systems', and that "the mathematical 'machinery' of Q-systems is highly transparent; it is as simple as could reasonably be expected" (1978a, p. 263).[27]

Let us briefly examine these Q-systems, therefore, before turning to analogy in KDS. Kuipers presents them axiomatically (1978a, p. 265), but for our purposes it suffices to note a few salient features. Like (21), the analogous function for the α-λ continuum, the representative function for Q-

The basic ingredients for the requisite $p(e)$ are as in the α-λ continuum: $p(C_w)$ and $p(e|C_w)$. But $p(C_w)$ in a Q-system is a freely chosen parameter; though it can be set according to (19) or (20), it need not be. $p(e|C_w)$ for Ct-evidence is a variant of (23) obtained by replacing the α-λ representative function with (36), the corresponding Q-function. The result is

$$\frac{\prod_{Q=1}^{c} \pi\,(n_Q, \rho)}{\pi\,(n,\ w\,\rho)}\ . \tag{38}$$

$p(e|C_w)$ for the Pt-evidence of analogy is therefore

$$\sum_{d=1}^{2^x} \frac{\prod_{Q=1}^{c_d} \pi\,(n_Q, \rho)}{\pi\,(n+1,\ w\,\rho)}\ . \tag{39}$$

instead of (27). Reflecting these changes, $p(e)$ is then a scaled-down version of (28):

$$\sum_{i=0}^{K-c} [(\binom{K-c}{i})\ p\,(C_{c+1})\ \sum_{d=1}^{2^x} \frac{\prod_{Q=1}^{c_d} \pi\,(n_Q, \rho)}{\pi\,(n+1,\ (c+i)\,\rho)}\ ]\ . \tag{40}$$

Finally, $p(e \wedge h)$ is a parallel version of (29):

$$\sum_{i=0}^{K-c} [(\binom{K-c}{i})\ p\,(C_{c+1})\ \frac{\prod_{Q=1}^{c} \pi\,(n_Q, \rho)}{\pi\,(n+1,\ (c+i)\,\rho)}\ ]\ . \tag{41}$$

c) Assessment of analogy in the K-dimensional system

To track the behavior of perfect analogy in KDS, we recall that the probability of a partially observed individual having a given Ct-predicate is $p(e \wedge h)/p(e)$ expressed as in (41)/(40). This quotient, like (32), has the structure of (33): both numerator and denominator are summations over products of N-components (the number of constituents of a given width

KDS is successful with perfect analogy, as Niiniluoto has shown, but it does not return acceptable values for the imperfect variety (1981, pp. 7-10). Attempting to remedy that has led a number of thinkers to explore the kind of broad analogy Carnap called analogy by similarity. Among them are Niiniluoto (1980, 1981, 1988), Spohn (1981), Constantini (1983), Kuipers (1984), Skyrms (1993), and Festa (1997). Space constraints preclude a survey of this literature, but I will briefly describe Kuipers' approach in (1984), which Niiniluoto has explicitly endorsed (1988, p. 287).

Kuipers observes that (1), Carnap's characteristic function for the λ-system, can be looked at as an application of the straight rule to n_Q real empirical instances of a certain Q-predicate and $\lambda(K)/K$ virtual logical instances of the same predicate (1984, p. 69). Why not then treat analogy by analogy with these virtual logical instances? Why not add virtual analogical instances to (1) so that similarities among predicates are factored in? That is, let the number of virtual analogical instances of a specific Q-predicate on the evidence e be $\eta_Q(e) \geq 0$. Each Q-predicate will have its own $\eta_Q(e)$, which together add up to $\eta(n)$. Then (1) could be given an analogy factor $\eta_Q(e)/\eta(n)$ to go along with its empirical factor n_Q/n and its logical factor $1/K$. That is, (1) would become

$$c(h_Q, e_Q) = \frac{n_Q + \dfrac{\lambda(K)}{K} + \eta_Q(e)}{n + \lambda(K) + \eta(n)} . \qquad (44)$$

Like the empirical factors and logical factors, the various analogy factors sum to 1. (44) would hold only for the part of KDS coextensive with the λ-continuum, but Niiniluoto speculates on extending the procedure to the rest of KDS in (1988, pp. 289-292).

This is an attractive proposal, intuitive and clear, but how would the analogy factors be chosen? Intuitively, the idea is to make them proportional to the relative similarities of the Q-predicates. Techniques for measuring these similarities have been proposed by Niiniluoto (1981, pp. 12-14), Kuipers (1984, pp. 67, 73-74), and again by Niiniluoto (1988, pp. 279-80). Suppose we take the first of these proposals as an illustration. Let d_{uv} be the number of primitive predicates not shared by the Q-predicates 'Q_u' and 'Q_v'. Then the Q-predicates' degree of resemblance r can be expressed as suggested in Section 1, an analogy is rationally acceptable only if its conclusion is more probable on the evidence than any rival conclusion based on the same evidence, then these logics make the formal criterion of greater probability operational. What is more, if the argument from analogy has the epistemically foundational role I believe it can be shown to have, then logics such as these assume critical importance at the very roots of knowledge. An argument classifying something as a certain kind has not only the obvious constraint on true premises; it has, in addition, a usable check on its form.[30]

Notes

1 One of the few exceptions is Niiniluoto (1988).

2 Recent work on analogy by similarity includes Skyrms (1993) and Festa (1997).

3 Skyrms (1991) develops proposals by Kuipers (1988) and Martin (1967) in showing how to handle finite Markov chains as one kind of analogy by proximity.

4 Relevance is the line between evidence and knowledge or, put another way, evidence is relevant knowledge. How we make judgments of relevance is a psychological question, and how we ought to make them is a logical question. But that we make them is not a question at all; it is a fact. These issues, which are complex indeed, cannot be pursued further here.

5 It may be said in Achinstein's and Hesse's defense that Carnap's description of perfect analogy was probably the root of their error.

6 Carnap was converted to a semantic view of logic by Tarski. He relates in (1963, pp. 60-67, 71-72) that the concept of range came from Wittgenstein (1922) and Waismann (1930-31). He comments repeatedly on the centrality of range in both deductive and inductive logic; in (1942, pp. 96-97), for example, and in (1945, pp. 73-75).

7 e_Q here is not e_Q in Carnap (1952) but e_i. For the difference, see (1952, p. 12). Similarly, h_Q here is Carnap's h_i. The changes have been made in the interests of a more suggestive notation..

8 For the details of the derivation, see Carnap (1952, pp. 16-18, 30-31).

9 I have changed Carnap's individuals *b* to *a* and *c* to *b* in order to mesh with usage elsewhere in this paper.

10 The formula survives in Carnap's later work in (1950, p. 569) and Carnap and Stegmüller (1959, p. 227).

11 This does not violate Carnap's requirement of total evidence. See Hempel (1965, pp. 64-65).

12 Here I follow Hesse's presentation in (1964, p. 324).

13 Carnap talks of widths and distances of regions in an attribute space, but for our purposes these refinements are inessential.

29 That (43) is the natural extension of (9) in KDS was pointed out to me by Theo Kuipers (personal communication).

30 For help in locating hard-to-find sources, I am indebted to different people in different ways: Rick L. Chaney and Julie Arata Heringer of Saint Louis University's Madrid campus; Ron Crown of Saint Louis University's Frost campus; and Blanca Bengoechea and Ana María Jiménez of the Instituto de Filosofía at the Consejo Superior de Investigaciones Científicas in Madrid. In addition, the comments of two anonymous referees for *Theoria* aided considerably in the final revision of this paper, but any faults it may retain are my own.

BIBLIOGRAPHY

Achinstein, P.: 1963, 'Variety and Analogy in Confirmation Theory', *Philosophy of Science* 30, 207-221.

Carnap, R.: 1942, *Introduction to Semantics*, Cambridge, MA, Harvard University Press.

Carnap, R.: 1945, 'On Inductive Logic', *Philosophy of Science* 12, 72-97.

Carnap, R.: 1950, *Logical Foundations of Probability*, Chicago, University of Chicago Press.

Carnap, R.: 1952, *The Continuum of Inductive Methods*, Chicago, University of Chicago Press.

Carnap, R.: 1963, 'Probability and Inductive Logic', 'My Basic Conceptions of Probability and Induction', and 'An Axiom System for Inductive Logic', in Schilpp (1963), pp. 71-77, 966-973, 973-979.

Carnap, R.: 1971, 'A Basic System of Inductive Logic, Part 1', in R. Carnap, R. Jeffrey (eds.): *Studies in Inductive Logic and Probability*, Vol. I, Berkeley, University of California Press, pp. 33-165.

Carnap, R.: 1980, 'A Basic System of Inductive Logic, Part 2', in R. Jeffrey (ed.): *Studies in Inductive Logic and Probability*, Vol. II, Berkeley, University of California Press, pp. 7-155.

Carnap, R., Stegmüller, W.: 1959, *Induktive Logik und Wahrscheinlichkeit*, Wien, Springer.

Constantini, D.: 1983, 'Analogy by Similarity', *Erkenntnis* 20, 103-114.

Festa, R.: 1997, 'Analogy and Exchangeability in Predictive Inferences', *Erkenntnis* 45, 229-252.

Helman, D. (ed.): 1988, *Analogical Reasoning*, Dordrecht, Kluwer.

Hempel, C.: 1965, *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York, Free Press.

Hesse, M.: 1963, *Models and Analogies in Science*, London, Sheed & Ward.

Hesse, M.: 1964, 'Analogy and Confirmation Theory', *Philosophy of Science* 31, 319-327.

Hesse, M.: 1968, 'Enumeration, Elimination, and Analogy', in I. Lakatos (ed.): *The Problem of Inductive Logic*, Amsterdam, North-Holland, pp. 220-223.

Hintikka, J.: 1965a, 'Distributive Normal Forms in First-Order Logic', in J.N. Crossley, M.A.E. Dummett (eds.): *Formal Systems and Recursive Functions*, Amsterdam, North-Holland, pp. 48-91.

Hintikka, J.: 1965b, 'Towards a Theory of Inductive Generalization', in Y. Bar-Hillel (ed.): *Logic, Methodology and Philosophy of Science: Proceedings of the 1964 International Congress*, Amsterdam, North-Holland, pp. 274-300.

Waismann, F.: 1930-31, "Logische Analyse des Wahrscheinlichkeitsbegriffs," *Erkenntnis* 1, 228-248. English trans. 'A Logical Analysis of the Concept of Probability,' in F. Waismann: 1977, *Philosophical Papers*, Dordrecht, D. Reidel, pp. 4-21.
Wittgenstein, L.: 1922, *Tractatus Logico-Philosophicus*, London, Routledge & Kegan Paul.

John R. Welch has taught at the University of Massachusetts and Wellesley College; he is currently professor of philosophy at Saint Louis University (Madrid). Prior publications include works on corporate agency, Quine, the practical syllogism, Llull, science and ethics, Renaissance philosophy, the pragmatics and semantics of analogy, and rhetoric. His present research is on the application of inductive logics to problems of ethical classification. He edits the Value Inquiry Book Series entitled "Philosophy in Spain."