SINGULAR ANALOGY AND QUANTITATIVE INDUCTIVE LOGICS

John R. Welch

Theoria, 14 (1999), 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.

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 first approximation, one could say that an analogy is rationally acceptable only if its conclusion is more probable on the evidence than any rival conclusion based on the same evidence. The problem would then be to estimate these probabilities.

Quantitative inductive logics offer clear and intuitively satisfying results in many cases, but whatever their merits, critics complain that they misfire when applied to analogy. The genealogy of this claim seems to go back to Hesse (1963, Ch. 3; 1964; 1968) and Achinstein (1963). Consequent to these studies, the literature on analogy in quantitative inductive logics reveals a certain lacuna. Analogy as general inference has been studied systematically in Pietarinen (1972), for example, yet the type of singular inference captured in the traditional argument from analogy remains something of an untold story, receiving glancing attention at best.[1] There appear to be at least three reasons for this. The traditional argument from analogy is a relatively complex inference, first of all, molecular rather than atomic; hence it is unlikely to attract much attention from logicians working on inductive foundations. Secondly, the fact that analogical inference is not logically fundamental sometimes obscures its foundational role in cognition. Finally, since the initial successes of quantitative inductive logics were with singular inductions and the initial failures with inductive generalizations, the latter presented the immediate challenge. But now that the more recent of these logics have tenable policies on inductive generalization, at least one loose end concerning singular induction remains: singular analogy in its traditional form. To address it is the aim of this paper.

The inquiry proceeds as follows. Since how analogy fares in quantitative inductive logics cannot be reckoned without sorting out various forms of analogy, Section 2 proposes a typology of analogical inferences. Two of the resulting types are coextensive with the traditional argument from analogy, and Sections 3–7 detail their handling in the systems of Carnap (1952), Carnap and Stegmüller (1959), Carnap (1971, 1980), Hintikka (1966), and Hintikka and Niiniluoto (1976) in turn. These sections are partly expository, but they also extend and apply the original systems that have nothing to say about analogy (Carnap 1952, Hintikka and Niiniluoto 1976) or even positively exclude it (Hintikka 1966, p. 115). By the end of Section 7, an approach to singular analogy based on the work of Carnap (1952), Hintikka and Niiniluoto (1976), Kuipers (1978a, 1984), and Niiniluoto (1981) emerges. Section 8 ventures a few remarks about its philosophical interest.

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 <u>general analogies</u>, which include at least one quantified sentence, and <u>singular analogies</u>, 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 $\underline{a}$ and $\underline{b}$. It capitalizes instead on the similarity between $\underline{a}$'s property $\underline{FG}$ and $\underline{b}$'s inferred property $\underline{\overline{F}G}$. 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 analogies that fall neatly into neither category because they rely on both known similarity among individuals and similarity among properties. A simple example is

A3: $Fa \wedge Ga \wedge Ha$.

$Fb \wedge –Gb$.

So Hb.

Here the known similarity between individuals ($\underline{Fa}$ and $\underline{Fb}$) is balanced by the known dissimilarity ($\underline{Ga}$ and $–\underline{Gb}$); the balance is tipped by the property analogy between $\underline{FGH}$ for $\underline{a}$ and $\underline{F\overline{G}H}$ for $\underline{b}$. A fuzzy boundary is not necessarily fatal, however. The second and decisive consideration is that there is a higher vantage point from which A1, A2, and A3 can be viewed synoptically: the conclusions of all three arguments maximize similarity to the strongest properties represented by the evidence. That is, the evidence in both A1 and A2 includes the strongest property $\underline{FG}$, so we conclude $\underline{FG}$ rather than $\underline{F\bar{G}}$ in the first case and $\underline{\bar{F}G}$ rather than $\underline{\bar{F}\bar{G}}$ in the second. Likewise, because the evidence in A3 includes the strongest property $\underline{FGH}$ we infer $\underline{F\bar{G}H}$ instead of $\underline{F\bar{G}\bar{H}}$.

The key distinction for singular analogy is not then between individual and property analogy; it is, I suggest, between what I will call broad analogy and narrow analogy. Think of analogy as a bridge across a river; the analogy's premises are piers supporting a platform, the conclusion, over which one can pass to the other side. The piers can be broad in the sense that how the evidence is distributed for projected and unprojected properties is explicitly taken into account. A standard example is the roulette wheel of unknown bias. Given a small number of trials, the likelihood of a projected property—a five, say, on the next spin—may reasonably be judged higher if unprojected properties like neighboring numbers have been frequently successful (Jeffrey 1980, p. 3; Skyrms 1993, p. 274). But the piers can also be narrow in that the distribution of evidence among these unprojected properties is not taken into account. If 99 of 100 squirrels have been observed to be gray in 95 cases, russet in 3, and black in 1, the hypothesis that the remaining squirrel is gray would be based on a narrower range of evidence, one which slights the relative frequencies of the unprojected properties (russet and black) in favor of that of the projected property (gray).

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 $\underline{FG}$ but none that are $\underline{F\bar{G}}$ or $\underline{\bar{F}\bar{G}}$; nevertheless, the similarity relations among these properties would make it seem more probable that the next individual is $\underline{F\bar{G}}$ rather than $\underline{\bar{F}\bar{G}}$.[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 $\underline{P}_1$ 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, superficially similar instances will be found to be different in some respects. (1964, p. 320)

Achinstein's and Hesse's point should be conceded at once, I think: the usual analogy involves differences as well as similarities among analogues. Indeed, their point should even be strengthened: all analogy involves differences as well as similarities. For as long as we know enough to know that there are two objects involved, we know enough to know that some of their properties are different. When the objects are physical, for instance, their spatial properties must be different, and this difference brings others in its train.

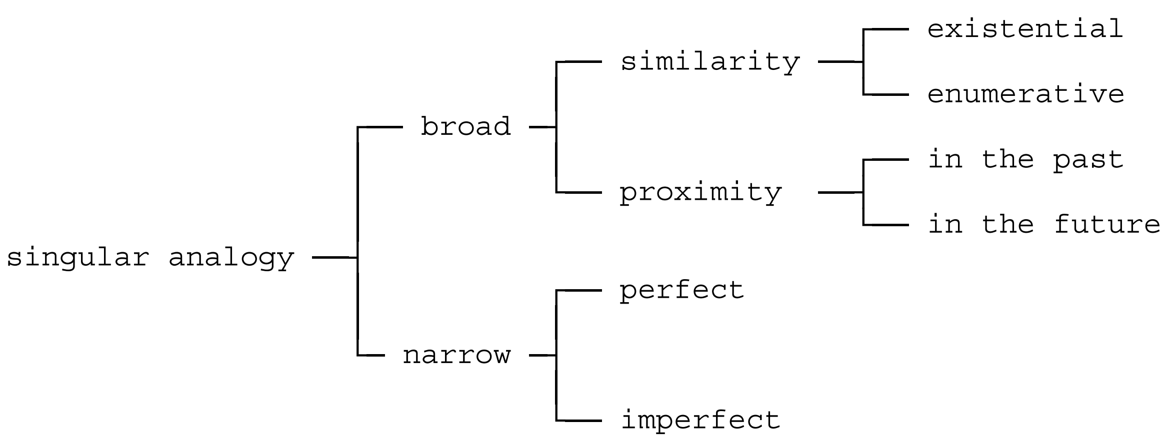
Nevertheless, this does not license the conclusion that the perfect analogy is "at best an idealization of the real situation." Why it does not can be gleaned from this nuanced passage from Mill.

> An argument from analogy, is an inference that what is true in a certain case, is true in a case known to be somewhat similar, but not known to be exactly parallel, that is, to be similar in all the material circumstances. An object has the property B: another object is not known to have that property, but resembles the first in a property A, not known to be connected with B; and the conclusion to which the analogy points, is that this object has the property B also. (1974, p. 794)

The argument Mill describes is our analogy A1. He does not claim, however, that there are no known differences between the two objects; he says only that they are "similar in all the material circumstances." Hence it is logically possible that the two objects are "somewhat similar," as he says, in that some of their properties are known to be different, but that all known relevant properties are the same.[4] This is not just a logical possibility, however; it is a recurrent feature of ordinary life. Think of the inference that the road under the next overpass will be slick because, five hundred meters back, the road under an overpass was slick. Or the hypothesis that a diskette manufactured by Erasem is defective because two others from the same box are.

So what we have are two kinds of narrow analogy. In one, some properties are different but all relevant properties are the same. In the other, all relevant properties are not the same. But the first is none other than perfect analogy, rightly viewed. The second is imperfect analogy. The distinction is not, as Achinstein and Hesse contend, that the imperfect has known differences and the perfect has none, for both have them.[5] It is that the known differences in the imperfect case are deemed relevant, while those in the 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:



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.

a) Contours of the λ-continuum

Carnap's inductive logics are built around a nucleus of basic themes. Among them is the segregation of statistical probability, whose statements are synthetic and belong to the object language, from logical probability, whose statements are analytic and proper to the metalanguage. A second is the view that it is logical probability, not statistical probability, which is "the basis of all inductive reasoning" (1963, p. 967). Finally, consistent with his approach to deductive logic from about 1935 on, Carnap treats inductive logic as a branch of semantics and accords a central role to the semantic concept of logical range—or model, as we would say today.[6] Whereas in deductive implication the models of the premises are entirely included in those of the conclusion, in inductive arguments only some of the premises' models are included in those of the conclusion. For example, if the probability of the conclusion given the premises is 2/3, the conclusion's models contain 2/3 of those of the premises.

Carnap sets out along these lines to explicate the concept of logical probability, limiting his λ-continuum of (1952) to first-order languages L with identity that have a finite number N of individual constants and a finite number k of logically independent primitive predicates. Within such languages, molecular predicates can be formed through Boolean operations on the primitive predicates, and of these molecular predicates, one group receives special attention: the class of Q-predicates. A Q-predicate is a strongest predicate formed by the conjunction of every primitive predicate, negated or not, of L. There are $2^k = K$ different Q-predicates, which are mutually exclusive and jointly exhaustive. For example, for L with k = 2 and primitive predicates 'F' and 'G', there are $2^2 = 4$ Q-predicates.

$Q_1(x)$	$F(x) \wedge G(x)$
$Q_2(x)$	$F(x) \wedge -G(x)$
$Q_3(x)$	$-F(x) \wedge G(x)$
$Q_4(x)$	$-F(x) \wedge -G(x)$

Together, they exhaust the kinds of individuals describable in the language.

Carnap's explication of logical probability is a definition of the degree of confirmation c(h,e) for any hypothesis h based on evidence e formulable within L. The guiding idea behind the λ-continuum is that c(h,e) should lie within an interval bounded by an empirical factor and a logical factor. The empirical factor is the ratio of favorable instances of a predicate in a sample 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:

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

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

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

(1) and (2) can be adapted for specific use in one of Carnap's two types of inductive methods. In Carnapian methods of the first kind, λ does not depend on K and is a constant. Then (1) takes the simpler form

(3) $$c(h_Q, e_Q) = \frac{n_Q + \frac{\lambda}{K}}{n + \lambda},$$

and (2) changes accordingly. But in methods of the second kind λ varies with K. Carnap's preferred method c* is the simplest such method, obtained by making λ(K) = K. In that case (1) reduces still further to

(4) $$c^*(h_Q, e_Q) = \frac{n_Q + 1}{n + K}.$$

The preceding definitions of degrees of confirmation are all for special situations in which a certain type of predicate is attributed to a single unobserved individual on the basis of a certain kind of sample. (1) has extremely far-reaching consequences, however. It is the characteristic function for the entire λ-continuum; once values characteristic of any inductive method in the continuum are assigned to it, then the values of c are fixed for any pair h,e of sentences where e is not logically false.

The transition from the special case represented by (1) to the general case is basically a three-step process. The first step is to use (1) in defining a measure function μ for state descriptions. A state description is a conjunction whose clauses consist of each of L's primitive predicates either affirmed or denied of each individual nameable in the language. Hence each such description reports a possible state of the world as completely as L permits. State descriptions can be stated equivalently in Q-form, that is, as a conjunction of N Q-sentences, each of which attributes a Q-predicate to one of the N individuals of L. Consider a state description t in Q-form such that 'Q_1' is instantiated by n_1 individuals, 'Q_2' by n_2, and so on to 'Q_K' by n_K. Each of these numbers n_Q is the Q-number of its associated predicate, and

$\Sigma \underline{n}_Q = \underline{N}$. Then relying on Carnap's product rule (1952, p. 12) and repeated applications of (1) (and writing 'λ' in place of 'λ(K)' for simplicity), we obtain

$$(5) \qquad \mu(t) = \frac{\prod_Q \left[\frac{\lambda}{K}\left(1+\frac{\lambda}{K}\right)\left(2+\frac{\lambda}{K}\right)\ldots\left(n_Q-1+\frac{\lambda}{K}\right)\right]}{\lambda\,(1+\lambda)\,(2+\lambda)\ldots(N-1+\lambda)}$$

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 μ(e) ≠ 0,

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

(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]

This description is repeated verbatim in Carnap (1950, p. 569).

Carnap is plainly speaking of perfect analogies to the exclusion of the imperfect variety. He makes short work of the subject, treating analogy as a special form of predictive inference: a conclusion about the composition of an unobserved sample drawn from the composition of an observed one. One result is a remarkably simple formula for perfect analogies in c*.[10] Let 'M_1' be the conjunction of primitive predicates, negated or not, known to be true of a, and let 'M_2' be the conjunction of primitive predicates known to be true of b. In addition, let a be understood to share all known properties of b, and to have other known properties as well. Hence the Q-predicates formed from 'M_1' are a proper subset of those formed from 'M_2', and the width of the former, w_1, is less than that of the latter, w_2. Now suppose we want the degree of confirmation of the analogical hypothesis that b has the a-properties it is not known to have—that b, in other words, is also M_1. According to (6), $c^*(M_1b, M_1a \wedge M_2b) = \mu^*(M_1a \wedge M_1b)/\mu^*(M_1a \wedge M_2b)$. The μ-values of this expression are obtained by repeated applications of (2) according to the product rule. Then since $\lambda(\underline{K}) = \underline{K}$ in $\underline{c}$*, we have

$$(7) \qquad c^*(M_1 b, M_1 a \wedge M_2 b) = \frac{w_1/K}{w_1/K}\,\frac{(1+w_1)/(1+K)}{(1+w_2)/(1+K)} = \frac{1+w_1}{1+w_2}.$$

(7) holds for what Carnap calls 'simple analogy', where the inference is from one individual to another, but multiple analogies from two or more individuals to another can be handled by replacing (7)'s empirical factor '1/1' with '2/2', '3/3', etc.

c) Assessment of narrow analogy in the λ-continuum

Carnap seems never to have generalized (7) for the whole λ-continuum, but it is quickly and usefully done. Where '$\underline{w}_2$' and '$\underline{w}_1$' have the meanings they have in (7),

$$(8) \qquad c(M_1 b, M_1 a \wedge M_2 b) = \frac{1+(w_1 \frac{\lambda(K)}{K})}{1+(w_2 \frac{\lambda(K)}{K})}.$$

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: $\underline{a}$ and $\underline{b}$ are $\underline{FG}$, $\underline{c}$ is $\underline{F\bar{G}}$, $\underline{d}$ is $\underline{\bar{F}G}$, and $\underline{e}$ is $\underline{F}$. In calculating the probability of the hypothesis that $\underline{e}$ is $\underline{G}$, the evidence concerning $\underline{d}$ should be excluded from the variant of (7) since it is incompatible with what is known about $\underline{e}$. Thus the empirical factor would be 2/3 with $\underline{FG}$, 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 $\underline{n}$ individuals $\underline{i}$ ($\underline{i}$ = $\underline{a}$,$\underline{b}$, ..., $\underline{y}$) have been examined and that $\underline{n}_1$ have $\underline{M}_1$, a property of width $\underline{w}_1$ attributed to the $\underline{n}$ + 1st individual $\underline{z}$ by the analogy's conclusion. Suppose also that $\underline{n}_2$ have $\underline{M}_2$, a property of width $\underline{w}_2 > \underline{w}_1$ that $\underline{z}$ is already known to have. Then where $\underline{E}_i$ is the conjunction of all the evidence about the individuals $\underline{i}$, the degree of confirmation in $\underline{c}$* of the analogical hypothesis that $\underline{z}$ is $\underline{M}_1$ can be calculated via (6) and (2) as:

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

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):

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

Though (10) holds all across the continuum, it has simpler special forms. For Carnapian methods of the first kind, 'λ($\underline{K}$)' reduces to 'λ' in both numerator and denominator, and for $\underline{c}$*, of course, (10) reduces to (9). The advantages of (10) are the advantages of (9) but magnified: easy yield of the same results obtainable with more labor from the characteristic function. The labor saved is usually considerable.

Critics of the λ-continuum like Achinstein and Hesse acknowledge its ability to deal with perfect analogies, but object to its handling of the imperfect type (Achinstein, 1963, p. 216; Hesse, 1964, p. 320). Achinstein, for example, considers the hypothesis that a sample conducts heat given that it is rhodium and that a sample of platinum and a sample of osmium conduct heat (1963, p. 217). Since rhodium, platinum, and osmium are all metals, they share some—but not all—primitive properties, thereby encouraging the imperfect analogical conclusion that the rhodium conducts heat because the platinum and osmium do. The degree of confirmation of this hypothesis given the evidence should be higher than the same hypothesis given only that the sample is rhodium. But in Carnap's continuum, as Achinstein shows, this is nowhere so (1963, pp. 220–21). Moreover, should nonmetals like oxygen and hydrogen be observed to conduct heat, that would provide the same support to the rhodium hypothesis as evidence of heat conduction by platinum and osmium; but the former have few primitive properties in common with rhodium and the latter have many. Hence the λ-system fails for imperfect analogies.

4. Carnap's and Stegmüller's λ-η Continuum

Carnap recounts that he modified his approach to inductive logic in 1951, soon after the writing but before the publication of (1952) (1963, pp. 75, 974 n45). The change was motivated by the realization that the λ-system is fully reliable only for predicates that belong to a single family, e.g., the family of colors. While the methods of (1952) can still be relied upon for approximations, the modified approach is to regard each primitive predicate as a member of a family composed of it and its negation. Three primitive predicates, therefore, become three families composed of two predicates apiece (1963, p. 974).

Carnap and Stegmüller exploit this modified approach in a new set of axioms for inductive logic in (1959, pp. 242–52). Because the axioms permit the derivation of the λ-continuum's characteristic function, perfect analogies fare here just as in the earlier continuum. But to improve the λ-continuum's results with imperfect analogies, the axioms also constrain possible definitions of a measure function μ different than the λ-continuum's (5) so that c(h,e) may be defined as in (6). Initially, two candidates for this measure function are considered.

To introduce the first, let $\mathcal{F}_1$ be a family of k_1 primitive predicates, and let $\mathcal{F}_2$ 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 $\mathcal{F}_{1,2}$ with $k_1k_2 = K$ members.

Say that each instantiated Q-predicate is true of $\underline{n}_{ij}$ individuals. Then the measure of a state description $\underline{t}$ is calculated as in the λ-system for a language with $\underline{k}_1\underline{k}_2$ Q-predicates. Thus the measure is given by (5) adapted to this special case:

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

(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 $\underline{k}$ primitive predicates with (5) as if it were a language with $\underline{k}$ 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 $\underline{L}$'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 $\underline{t}$ can be expressed as

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

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

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

$$(t_2) \quad FGw \wedge FGx \wedge \overline{F}\overline{G}y \wedge \overline{F}\overline{G}z$$

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

For $\underline{t}_1$ and $\underline{t}_2$ it should turn out that

$$c(F\overline{G}x \wedge \overline{F}Gy, FGw \wedge \overline{F}\overline{G}z) < c(FGx \wedge \overline{F}\overline{G}y, FGw \wedge \overline{F}\overline{G}z),$$

and we would therefore expect that (a) $\mu(\underline{t}_1) < \mu(\underline{t}_2)$. In addition, since for $\underline{t}_2$ and $\underline{t}_3$ we require that

$$c(\overline{F}\overline{G}y \wedge \overline{F}\overline{G}z, FGw \wedge FGx) < c(F\overline{G}y \wedge F\overline{G}z, FGw \wedge FGx),$$

we would expect that (b) $\mu(\underline{t}_2) < \mu(\underline{t}_3)$. But $\mu^{1,2}_{\lambda}$ meets (a) though not (b), for Q-measures treat all Q-predicates alike, ignoring the fact that the Q-predicates of $\underline{t}_3$ show analogy by similarity while those of $\underline{t}_2$ do not. $\mu^{1/2}_{\lambda}$, on the other hand, meets (b) though not (a); P-measures do pick up analogy by similarity, but overlook the difference between the disorderly world represented by $\underline{t}_1$ and the more orderly one represented by $\underline{t}_2$.

In an attempt to retain the advantages but excise the disadvantages of the first two measure functions, Carnap and Stegmüller propose another candidate. This third measure function is defined as a weighted mean of the first two by using an additional parameter, η, as the weight. Where $0 < \eta < 1$,

(13) $$\mu_{\lambda,\eta}(t) = [\eta \times \mu^{1/2}_{\lambda}(t)] + [(1-\eta) \times \mu^{1,2}_{\lambda}(t)].$$

η is what Kemeny calls the "analogy-constant" (1963, p. 733). The closer it is to 1, the more $\mu_{\lambda,\eta}$ approximates a P-measure and the more sensitive to imperfect analogy it becomes. As η approaches 0, however, $\mu_{\lambda,\eta}$ approximates a Q-measure and places more weight on the regularity with which Q-predicates are instantiated. But as long as its value stays within 0 and 1 noninclusive, $\mu_{\lambda,\eta}$ satisfies both (a) and (b).

Carnap and Stegmüller's definition of $\mu_{\lambda,\eta}$ is limited to just two families of predicates. But Hesse generalized both $\mu^{1,2}_{\lambda}$ and $\mu^{1/2}_{\lambda}$ for $\underline{n}$ predicate families and then defined an $\underline{n}$-family version of $\mu_{\lambda,\eta}$ (1964, p. 325):

(14) $$\mu_{\lambda,\eta}(t) = [\eta \times \mu^{1/2/\ldots/n}_{\lambda}(t)] + [(1-\eta) \times \mu^{1,2,\ldots,n}_{\lambda}(t)].$$

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*:

$$\eta^* = \frac{\lambda^*}{\lambda^*+1}. \tag{15}$$

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. 105–06).

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 views these extreme values as improper (1980, pp. 94–95). But the pivotal difference for our purposes is the restatement of the characteristic function (or representative function, as Carnap now calls it).[14] Assume a family of mutually exclusive primitive predicates and evidence e_P that one of these predicates, say 'P', is instantiated by n_P of n individuals. In addition, each primitive predicate has a γ-value ranging from 0 to 1 noninclusive that represents its relative width; the various γ-values sum to 1. Then the degree of confirmation of the hypothesis h_P that the next individual will instantiate 'P' is

(16) $$c(h_P, e_P) = \frac{n_P + \lambda\gamma_P}{n + \lambda}.$$

(16), the Basic System's representative function, amounts to a major expansion of the system of (1952).[15] The λ-continuum requires that the widths of the properties all be the same, but γ in the Basic System admits unequal as well as equal widths. As a result, two types of predicate families figure in the Basic System. λ-families satisfy the λ-condition that the degree of confirmation of hypotheses projected with their predicates depends only on n_P and n (1980, pp. 84, 87, 97). λ-γ- families, on the other hand, fulfill the λ-γ-condition, which consists of the λ-condition plus the additional stipulation that the various γ-values be equal (1980, pp. 87, 101). These latter families are therefore a special case of the former, and the λ-continuum (excluding the extreme systems where λ = 0 and ∞) is the corresponding subsystem of the Basic System.

How then does singular analogy fare in the Basic System? λ-γ-families obviously succeed no more nor less than the λ-continuum: perfect analogies, which can be dispatched by the short formula (10), do well; but imperfect analogies do not. λ-families could conceivably be treated along the lines of the λ-η continuum, but Carnap makes it clear that he is undecided about whether to keep his and Stegmüller's earlier solution or not (1980, p. 46). Moreover, though Carnap discusses broad analogy as well as narrow analogy for the first time in introducing the distinction between analogy by similarity and analogy by proximity (1980, pp. 40–41, 68–71), there is no doubt that broad analogy has no place in the Basic System; the λ-condition that both predicate families satisfy directly excludes it (Carnap 1980, p. 84; Jeffrey 1980, p. 3). For if degree of confirmation depends only on $\underline{n}_P$ and $\underline{n}$, it cannot depend on distance (similarity among predicates). Hence 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).

To introduce the α-λ continuum, we will need some Hintikkan terminology. Where $\underline{L}$ is a language with $\underline{k}$ logically independent primitive predicates (assumed, as in Hintikka's original papers, to be monadic), a constituent-predicate $\underline{Ct_i}(\underline{x})$ ($\underline{i}$ = 1, 2, ..., $\underline{K}$) is a complex predicate differing only in name from a Carnapian Q-predicate. There are $\underline{K}$ different Ct-predicates, which are of course mutually exclusive and jointly exhaustive. A constituent $\underline{C_w}$ is a closed sentence stating which Ct-predicates are instantiated and which are not. Each constituent has a width $\underline{w}$ equal to the number of Ct-predicates that it claims are instantiated. Thus constituents are equivalent to sentences of the following form:

$$(\exists x)\underline{Ct_1}(\underline{x}) \wedge (\exists x)\underline{Ct_2}(\underline{x}) \wedge \ldots \wedge (\exists x)\underline{Ct_w}(\underline{x}) \wedge (\forall x)(\underline{Ct_1}(\underline{x}) \vee \underline{Ct_2}(\underline{x}) \vee \ldots \vee \underline{Ct_w}(\underline{x}))\ .$$

There are $2^{K} - 1$ mutually exclusive constituents.[16] By contrast with Ct-predicates, which define possible kinds of individuals, constituents define possible kinds of worlds. They are the generalizations that form the core of Hintikka's systems, the primary reason that the continuum is so different from Carnap's methods.

b) Extending the α-λ continuum to narrow analogy

α-λ methods can be applied to singular analogy even though the evidence assumptions made in the original presentation of the continuum rule them out (1966, p. 115). The ensuing discussion concentrates on narrow analogies; the basic idea is to handle them as conditional probabilities. The analogy's conclusion is some hypothesis h, which is assumed to attribute a Ct-predicate to some object. The probability of h on the evidence e of the premises is therefore equal to $p(e \wedge h)/p(e)$. How, then, can the two required probabilities be calculated within α-λ? Let us take each in turn, beginning with the denominator.

i) $p(e)$ for narrow analogy

Given evidence e and a finite number of mutually exclusive and exhaustive hypotheses h_j ($j = 1, 2, ..., r$), Bayes' Theorem expresses $p(e)$ as follows:

(17)
$$\sum_{j=1}^{r} p(h_j)\, p(e|h_j)\,.$$

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

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

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 $\pi(0,z) =_{df} 1$,

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

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

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

When α = 0, all constituents receive equal prior probabilities. When α > 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 characteristic or representative function for the continuum.[19] Hintikka's representative function is conditional in that it depends on the assumption that a given constituent C_w compatible with the evidence is true.[20] Where e_Q is the evidence that n_Q of n individuals have the Ct-predicate 'Q', h_Q is the hypothesis that the next unexamined individual is Q, and w is the width of the constituent assumed true, the representative function is

(21) $$c(h_Q, e_Q) = \frac{n_Q + \frac{\lambda}{w}}{n + \lambda}.$$

Now suppose that n individuals have been observed and that n_1 have the Ct-predicate 'Q_1', n_2 have 'Q_2', and so on to n_c with 'Q_c'. Each Ct-predicate is instantiated by n_Q individuals, and

(22) $$\sum_{Q=1}^{c} n_Q = n.$$

Then $p(e|C_w)$ is equal to

(23) $$\frac{\prod_{Q=1}^{c} \pi(n_Q, \frac{\lambda}{w})}{\pi(n, \lambda)}.$$

Being able to determine $p(C_w)$ and $p(e|C_w)$, as we now are, would be sufficient to determine p(e), as we have just seen, provided that the evidence is Ct-homogeneous—provided, that is, that the evidence is simply that certain objects instantiate certain Ct-predicates. But that is precisely what we do not have in a narrow analogy. From the point of view of evidence, narrow analogies are hybrids. Some of their premises may relate objects to Ct-predicates, but other premises—the ones reporting partial observations—relate objects to primitive predicates. For example, our simple perfect analogy A1 attributes the Ct-predicate '$F(x) \wedge G(x)$' to a and the primitive predicate '$F(x)$' to b. We have already seen that (23) would suffice for $p(e|C_w)$ when e is like that of the first premise, which I shall call Ct-evidence. What remains, therefore, is to describe how to handle the conjunction of Ct-evidence with Pt-evidence, evidence about 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

(24) $$\underline{e}_{Q1} \vee \underline{e}_{Q2} \vee \ldots \vee \underline{e}_{Qn} ,$$

each clause of which attributes a Ct-predicate to an object. Conjoining (24) to the Ct-evidence $\underline{e}_C$ yields

(25) $$\underline{e}_C(\underline{e}_{Q1} \vee \underline{e}_{Q2} \vee \ldots \vee \underline{e}_{Qn}) ,$$

or, distributing,

(26) $$(\underline{e}_C \wedge \underline{e}_{Q1}) \vee (\underline{e}_C \wedge \underline{e}_{Q2}) \vee \ldots \vee (\underline{e}_C \wedge \underline{e}_{Qn}) .$$

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 $\underline{n}$ + 1st individual. A different Ct-predicate is attributed to this individual in each of (26)'s clauses; if the projected predicate is '$\underline{Q}_1$', the clause states that $\underline{n}_1 + 1$ of $\underline{n} + 1$ individuals, rather than $\underline{n}_1$ of $\underline{n}$, possess it, which makes the various $\underline{n}_q$ in (22) sum to $\underline{n} + 1$. With these differences in mind, let $\underline{x}$ be the number of predicate letters needed to transform the Pt-evidence into complete disjunctive normal form, and 2^x the number of disjuncts $\underline{d}$ of the evidence (26), each of which says that $\underline{c}_d$ Ct-predicates are instantiated. Then $\underline{p}(\underline{e}|\underline{C}_w)$ for Pt-evidence is

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

We are now ready to state p(e) for analogies. We assume that the same function that can relate λ to K for p(C_w) in (20) can also relate λ to w for p(e|C_w) in (27). Then p(e), which was expressed schematically as (18), can be fleshed out as

$$(28)\qquad \frac{\sum_{i=0}^{K-c}\left[\binom{K-c}{i}\pi\left(\alpha,\frac{(c+i)\cdot\lambda(K)}{K}\right)\sum_{d=1}^{2^x}\frac{\prod_{Q=1}^{c_d}\pi\left(n_Q,\frac{\lambda(c+i)}{c+i}\right)}{\pi(n+1,\lambda(c+i))}\right]}{\sum_{i=0}^{K}\binom{K}{i}\pi\left(\alpha,\frac{i\cdot\lambda(K)}{K}\right)}.$$

ii) p(e ∧ h) for narrow analogy

Having p(e) in hand, we need only p(e ∧ h) to be able to find the conditional probability of an analogical conclusion. p(e ∧ h) is like (28) except for a change to the expression for p(e|C_{c+i}). Since 'h' attributes a Ct-predicate to the n + 1st individual, it fills in the epistemic gaps left by the Pt-evidence; thus 'e ∧ h' is Ct-homogeneous and innocent of the complications inherent in e for analogies. As a result, p(e|C_{c+i}) is not based on (27) but on (23), which is modified so that n_1 + 1 of n + 1 individuals have the projected Ct-predicate 'Q_1'. p(e ∧ h) is then equal to

$$(29)\qquad \frac{\sum_{i=0}^{K-c}\left[\binom{K-c}{i}\pi\left(\alpha,\frac{(c+i)\cdot\lambda(K)}{K}\right)\frac{\prod_{Q=1}^{c}\pi\left(n_Q,\frac{\lambda(c+i)}{c+i}\right)}{\pi(n+1,\lambda(c+i))}\right]}{\sum_{i=0}^{K}\binom{K}{i}\pi\left(\alpha,\frac{i\cdot\lambda(K)}{K}\right)}.$$

The probability of an analogy's conclusion given its premises, p(h|e), can now be calculated as p(e ∧ h)/p(e).

c) Assessment of narrow analogy in the α-λ continuum

As in Carnap's λ-system, perfect analogies in the α-λ continuum have smooth sailing. The similarity is not just one of result, however, for Carnap'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 λ = λ(w) = w except in (19) and (20), where λ = λ(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 $\alpha \to \infty$.

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 λ(K) = K for prior probabilities of constituents in GCS, the P-components in numerator and denominator simplify to π(α, c + i). And since elsewhere in GCS λ = λ(w) = w, the α-λ representative function (21) reduces to

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

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

(31) $$c(h_M, e_M) = \frac{n_M + w_M}{n + w}.$$

Drawing all this together, and relying momentarily on (30) rather than (31), we can express the probability in GCS of an analogical hypothesis that a partially observed individual has a given Ct-predicate as $p(e \wedge h)/p(e)$:

(32) $$\frac{\sum_{i=0}^{K-c} \left[\binom{K-c}{i} \pi(\alpha, c+i) \frac{\prod_{Q=1}^{c} \pi(n_Q, 1)}{\pi(n+1, c+i)} \right]}{\sum_{i=0}^{K-c} \left[\binom{K-c}{i} \pi(\alpha, c+i) \sum_{d=1}^{2^x} \frac{\prod_{Q=1}^{c_d} \pi(n_Q, 1)}{\pi(n+1, c+i)} \right]}.$$

The special cases we are concerned with here are perfect analogies where both '$e \wedge h$' and 'e' agree on the number of Ct-predicates that are instantiated. Hence c is the same in the expressions for $p(e \wedge h)$ and $p(e)$, which suggests the term 'c-uniform' for these analogies. A1 of Section 2 is not c-uniform, for example, since '$e \wedge h$' recognizes one Ct-predicate while 'e' wavers between one and two. But if A1's evidence is augmented by the observation that another individual is $F\overline{G}$, the result is c-uniform; both '$e \wedge h$' and 'e' countenance two Ct-predicates.

One property of $\underline{c}$-uniform analogies is the identity of the N- and P-components in (32)'s numerator to those in the denominator. Consequently, if the nonidentical R-components for $\underline{p}(\underline{e} \wedge \underline{h})$ and for $\underline{p}(\underline{e})$ are represented as '$\underline{R}^{eh}$' and '$\underline{R}^{e}$' respectively, and the NPR components are given subscripts pegged to (32)'s $\underline{i}$ that indicate (indirectly) the width of their associated constituents, (32) has the structure

(33)
$$\frac{N_0 P_0 R^{eh}_{0} + N_1 P_1 R^{eh}_{1} + \ldots + N_{K\text{-}c} P_{K\text{-}c} R^{eh}_{K-c}}{N_0 P_0 R^{e}_{0} + N_1 P_1 R^{e}_{1} + \ldots + N_{K\text{-}c} P_{K\text{-}c} R^{e}_{K-c}} .$$

Now the last factor in the expansion of any R^{eh} component gives the probability on the evidence that the individual $\underline{z}$ of the analogy's conclusion has the Ct-predicate '$\underline{M}_1$'. The numerator of this factor takes the form $\underline{n}_1 + \underline{w}_1$, where $\underline{n}_1$ is the number of individuals known to have $\underline{M}_1$ and $\underline{w}_1$ is the predicate's width. Because this numerator is common to all $\underline{R}^{eh}$ components, it can be factored out of the expression for $\underline{p}(\underline{e} \wedge \underline{h})$. A parallel argument holds for the last factor in the expansion of any $\underline{R}^{e}$ component. The numerator of each such factor is $\underline{n}_2 + \underline{w}_2$, where $\underline{n}_2$ is the number of individuals known to have '$\underline{M}_2$', a disjunction of Ct-predicates that includes '$\underline{M}_1$', and $\underline{w}_2$ is its width. Because this numerator is common to all $\underline{R}^{e}$ components, it too can be factored out of the expression for $\underline{p}(\underline{e})$. Where '$\underline{R}^{eh-}$' and '$\underline{R}^{e-}$' are the R-components diminished by factoring, (33) is then

(34)
$$\frac{(n_1 + w_1)\ (N_0 P_0 R^{eh-}_{0} + N_1 P_1 R^{eh-}_{1} + \ldots + N_{K\text{-}c} P_{K\text{-}c} R^{eh-}_{K-c})}{(n_2 + w_2)\ (N_0 P_0 R^{e-}_{0} + N_1 P_1 R^{e-}_{1} + \ldots + N_{K\text{-}c} P_{K\text{-}c} R^{e-}_{K-c})} .$$

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 '$F\bar{G}$' 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:

$$\frac{\left(\frac{1}{2}\frac{0+1}{1+2}\frac{1+1}{2+2}\right) + \left(\frac{1}{3}\frac{0+1}{1+3}\frac{1+1}{2+3}\right) + \left(\frac{1}{4}\frac{0+1}{1+4}\frac{1+1}{2+4}\right)}{\left(\frac{1}{2}\frac{0+1}{1+2}\frac{2+2}{2+2}\right) + \left(\frac{1}{3}\frac{0+1}{1+3}\frac{2+2}{2+3}\right) + \left(\frac{1}{4}\frac{0+1}{1+4}\frac{2+2}{2+4}\right)}.$$

Factoring out the numerators in the last factor of each addend for both p(e $\wedge$ h) and p(e) gives

$$\frac{(1+1)\ \left[\left(\frac{1}{2}\frac{0+1}{1+2}\frac{1}{2+2}\right)+\left(\frac{1}{3}\frac{0+1}{1+3}\frac{1}{2+3}\right)+\left(\frac{1}{4}\frac{0+1}{1+4}\frac{1}{2+4}\right)\right]}{(2+2)\ \left[\left(\frac{1}{2}\frac{0+1}{1+2}\frac{1}{2+2}\right)+\left(\frac{1}{3}\frac{0+1}{1+3}\frac{1}{2+3}\right)+\left(\frac{1}{4}\frac{0+1}{1+4}\frac{1}{2+4}\right)\right]} = \frac{1+1}{2+2} = \frac{1}{2}.$$

The same result can be had from (9) in a fraction of the time.

Once the utility of (9) for c-uniform analogy has been established in GCS, it is tempting to stretch the point. If (9), which is native to c*, could be generalized as (10) for the entire λ-continuum, could it also be generalized for the α-λ continuum? Suppose that the evidence E_i concerning n individuals i (i = a,b, ..., y) is that n_1 possess a Ct-predicate 'M_1' of width w_1 and that n_2 have 'M_2', a disjunction of Ct-predicates of width w_2 that contains 'M_1'.

The $\underline{n}$ + 1st individual $\underline{z}$ also has $\underline{M}_2$. Then where $\underline{w}$ is the width of the constituent assumed true, the α-λ version of (9) for the analogical hypothesis that $\underline{z}$ is $\underline{M}_1$ would be

$$(35) \qquad p(M_1 z, E_i \wedge M_2 z) = \frac{n_1 + (w_1 \frac{\lambda(w)}{w})}{n_2 + (w_2)\frac{\lambda(w)}{w}} .$$

That (35) holds for c-uniform analogy all across the continuum can be easily shown by the factoring procedure rehearsed above for (9) in GCS. But the same proviso on logically incompatible evidence that applies to (9) and (10) applies also to (35). The result, as usual, is labor saved.

A common sort of c-uniform analogy merits mention apart. Besides c-uniformity, these analogies meet two further conditions: $\underline{c} = \underline{K}$ and $\lambda = \lambda(\underline{K})$. When $\underline{c} = \underline{K}$, all Ct-predicates are known to be instantiated; hence $\underline{w} = \underline{K}$ in (21), the α-λ continuum's representative function. Then provided $\lambda = \lambda(\underline{K})$, 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 $\underline{c}_0$

> ... the prior probabilities are assigned in such a way that $\underline{c}_0(\underline{Ct}_1(\underline{a}) \wedge \underline{Ct}_2(\underline{b}))$, where $\underline{Ct}_1$ _ $\underline{Ct}_2$, has the same value however similar or different $\underline{a}$ and $\underline{b}$ may be; that is, for example, if $\underline{Ct}_1$ is $\underline{P}_1\underline{P}_2\underline{P}_3, ..., \underline{P}_k$, it has the same value whether $\underline{Ct}_2$ is $\underline{P_1 P_2 P_3}, ..., \underline{\overline{P}_k}$, or $\underline{P_1 \overline{P}_2 \overline{P}_3}, ..., \underline{\overline{P}_k}$. 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 analogy, for example, as follows. Suppose that some of our primitive predicates are "metallic" in that they identify properties that metals have and nonmetals do not. Let these predicates form the family $\mathcal{F}_1$ and the remaining primitive predicates the family $\mathcal{F}_2$. Then the universe of discourse is partitioned twice: first according to the predicates of $\mathcal{F}_1$, then the resulting classes according to those of $\mathcal{F}_2$. The effect is asymmetry among the Ct-predicates: evidence that a property is present in metals like platinum and osmium counts more in hypotheses about rhodium than comparable evidence about nonmetals like oxygen and hydrogen. The fact that the primitive predicates for rhodium are more similar to those for metals than nonmetals is thereby registered. As Hintikka shows, this procedure is not limited to the case of two predicate families; it can also be achieved with a greater number of families and a corresponding increase in the number of partitions (1969, pp. 31–34).

This second solution relies on extra-systemic judgments about the ordering of primitive predicates, as Hintikka emphasizes. But unlike the first solution, it does not appear to be ad hoc. Hence not only does the α-λ continuum advance beyond Carnap's systems in its handling of inductive generalization; it brings imperfect analogy, which was never satisfactorily assimilated by Carnapian methods, under the tent. At this point, then, our resources for dealing with narrow analogy are not in bad shape—though, as we will see, they can be improved. But the problem of broad analogy has not been addressed at all.

7. Hintikka's and Niiniluoto's K-Dimensional System

As before, the discussion will proceed in three stages: a) contours of the K-dimensional system; b) applying the K-dimensional system to analogy; and c) assessment of analogy in the K-dimensional system.

a) Contours of the K-dimensional system

The glaring weakness of Carnap's inductive logics is their management of inductive generalization, a difficulty acknowledged by Carnap himself and apparently the root of his claim that the business of science is not to establish universal laws but to confirm instances of laws—primarily through singular inference.[25] Hence the program of Hintikka (1965b), (1965c), and (1966) to retrofit the Carnapian framework with a sustainable policy on 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 $\underline{n}_Q$ and $\underline{n}$ (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 $\underline{n}_Q$, $\underline{n}$, and $\underline{c}$ (the number of instantiated Ct-predicates). Hintikka and Niiniluoto express this second axiom by saying that, whereas the λ-function has the form $\underline{f}(\underline{n}_Q,\underline{n})$, KDS's function has the form $\underline{f}(\underline{n}_Q,\underline{n},\underline{c})$ (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 $\underline{K}$ free parameters, where $\underline{K}$ is, as before, the number of Ct-predicates specifiable in the language. The parameters are values for the representative function at $\underline{f}(0,\underline{c},\underline{c})$, where $\underline{c}$ = 1, 2, ..., $\underline{K}$–1, and for $\underline{f}(1,\underline{K}+1,\underline{K})$.

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 $\lambda(\underline{w})$ is proportional to $\underline{w}$ but without Hintikka's particular choice of the prior distribution $\underline{p}(\underline{C}_w)$ 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-systems is conditional in being relative to the truth of a constituent compatible with the evidence. Suppose evidence e_Q that n_Q of n individuals instantiate a Ct-predicate 'Q', and hypothesis h_Q that the next individual will also be Q. Then where w is the width of the constituent assumed true and ρ a real-valued parameter such that $0 < \rho \leq \infty$, the representative function can be formulated as

(36) $$c(h_Q, e_Q) = \frac{n_Q + \rho}{n + w\rho}.$$

Putting $\rho = \lambda/w$ and the prior probabilities of constituents in line with (19) or (20) generates the α-λ systems, including GCS (λ = K = w), the λ-continuum (K = w, α = ∞), and c* (λ = K = w, α = ∞). To this extent, Q-systems are familiar. But their parameters are specified differently than in KDS. Whereas the KDS parameters represent posterior probabilities concerning individuals, those of Q-systems range over prior probabilities of constituents. There are K–1 of these parameters $p(C_w)$, where w = 1, 2, ..., K–1.[28]

Generalizing the representative function for molecular predicates with $w \geq 1$ has already proved useful in the λ-continuum and in GCS, and it will prove so here as well. For ρ and w as in (36), evidence e_M that a molecular predicate 'M' of width w_M is instantiated by n_M of n individuals, and hypothesis h_M that the next individual will also be M, (36) can be amplified as

(37) $$c(h_M, e_M) = \frac{n_M + w_M\rho}{n + w\rho}.$$

(37) is the KDS analogue of the λ-continuum's (2) and GCS's (31).

b) Applying the K-dimensional system to analogy

Because of the equivalence between Q-systems and those of KDS, it is possible to study the latter, as Kuipers says, "in their 'Q-garb'" (1978a, p. 272). In particular, it is possible to show in a mathematically explicit way how to deal with narrow analogy in KDS. The strategy is fundamentally the same as it was in the α-λ continuum: the probability of the conclusion h given the evidence e of the premises is equal to p(e ∧ h)/p(e).

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

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

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

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

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

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

Finally, p(e ∧ h) is a parallel version of (29):

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

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 compatible with the evidence), P-components (the prior probabilities of these constituents), and R-components (based on the representative function).

Now suppose we take the special case of c-uniform analogy (analogies where both '$e \wedge h$' and 'e' concur on the number of instantiated Ct-predicates) that we explored previously within the α-λ continuum. Because c is the same in the expressions for both $p(e \wedge h)$ and $p(e)$, the N- and P-components of the numerator are identical to those of the denominator. But how are the R-components, obtained from iterated applications of (37), related to each other? We have already observed that R-components for analogies have the same number of factors: $n + 1$. In addition, the factoring sequence that led from (33) to (34) for GCS can be repeated here. For each R^{eh} component in $p(e \wedge h)$ concludes with a factor whose numerator is '$n_1 + w_1\rho$', where n_1 is the number of individuals instantiating 'M_1', the Ct-predicate projected by the analogy's conclusion, and w_1 is its width. And each R^e component in $p(e)$ has a final factor with a numerator of '$n_2 + w_2\rho$', where n_2 is the number of individuals instantiating 'M_2', a disjunction of Ct-predicates including 'M_1', and w_2 is its width. When both numerators are factored out, the result—structurally analogous to (34)—produces diminished R-components R^{eh-} in the numerator and R^{e-} in the denominator:

(42)
$$\frac{(n_1 + w_1\rho)\ (N_0 P_0 R_0^{eh-} + N_1 P_1 R_1^{eh-} + \ldots + N_{K-c} P_{K-c} R_{K-c}^{eh-})}{(n_2 + w_2\rho)\ (N_0 P_0 R_0^{e-} + N_1 P_1 R_1^{e-} + \ldots + N_{K-c} P_{K-c} R_{K-c}^{e-})}.$$

But since the numerator's diminished R-components are identical to the denominator's, they cancel along with the N- and P-components, leaving just

(43) $$\frac{(n_1 + w_1\rho)}{(n_2 + w_2\rho)} .$$

So for c-uniform analogies in KDS, one can take the express route via (43) instead of (41)/(40).[29] What (35) does for the α-λ continuum, (43) does for KDS. As always, however, any evidence incompatible with the analogy's conclusion must be excluded from (43)'s empirical factor.

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), Costantini (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

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

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

$$(45) \qquad r_{uv} = \frac{1}{1 + d_{uv}} \; .$$

If we have just two primitive predicates, '*F*' and '*G*', (45) gauges the degrees of resemblance between the Q-predicate 'FG' on the one hand and 'FG', '$F\overline{G}$', '$\overline{F}G$', and '$\overline{F}\overline{G}$' on the other to be 1, 1/2, 1/2, and 1/3 in turn. The application to (44) is straightforward: the analogy factors can be made proportional to the degrees of resemblance determined by (45). Specifically, (45) measures the degree of resemblance between the Q-predicates represented by the evidence and the Q-predicates being projected.

As a simple example, take our imperfect analogy A2. Relative to the minimal language where *K* = 4, λ-methods assign A2's conclusion a degree of confirmation of .5, thereby allotting to the rival conclusion that *b* is not *G* the same degree of confirmation. But that is to consider the less similar Q-predicate '$\overline{F}\overline{G}$' just as likely as the more similar '$\overline{F}G$'.

They are not equally likely, however, if we use (44) and (45) instead. (45) determines the degree of resemblance between the predicates in evidence ('$\underline{FG}$' and '$\underline{\bar{F}G} \vee \underline{\bar{F}\bar{G}}$') and the Q-predicates '$\underline{FG}$', '$\underline{\bar{F}G}$', and '$\underline{\bar{F}\bar{G}}$' to be 17/56, 15/56, and 13/56 respectively. Suppose we use these values as analogy factors in a version of (44) patterned after $\underline{c}^*$ ($\lambda(\underline{K}) = \underline{K}$). Then A2's conclusion receives a degree of confirmation of 8/15 (about .53), dropping that of the rival conclusion that $\underline{b}$ is not $\underline{G}$ to 7/15 (about .47). Comparably unequal results are obtained with other (44)-based methods.

So not only can (44) be relied upon for analogy by similarity; it affords smooth handling for imperfect analogy, and of course for perfect analogy as well. Compared to Carnap's and Stegmüller's technically adequate solution to the problem of imperfect analogy, (44) is both simpler and more general, and its analogy factors are grounded in an objective metric of similarity.

8. Conclusion

That quantitative inductive logics have matured to the point where they can handle both forms of narrow analogy as well as analogy by similarity is technically important, but it is not only that. For they provide us with a critical tool for assessing analogies that is mathematically explicit. If, 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. $\underline{e}_Q$ here is not $\underline{e}_Q$ in Carnap (1952) but $\underline{e}_i$. For the difference, see (1952, p. 12). Similarly, $\underline{h}_Q$ here is Carnap's $\underline{h}_i$. 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 <u>b</u> to <u>a</u> and <u>c</u> to <u>b</u> 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.

14. Both Carnap and Hintikka switched from the term 'characteristic function' to 'representative function' for describing their respective functions. The change was made "to avoid clashes with the standard mathematical and statistical terminology," as Hintikka explains in (1969, p. 44 n15).

15. (16) appeared in Stegmüller (1973, p. 496) prior to the publication of Part 2 of Carnap's Basic System.

16. The constituent that says that no Ct-predicate is instantiated is normally eliminated.

17. These methods are also applicable to finite universes. See Hintikka (1966, pp. 120–22).

18. It is worth noting that expressions like the π terms in (19) and (20) can be nonintegral. These expressions can be expressed as factorials, however, and the factorial function is routinely extended to nonintegral arguments with the gamma function. Thus (19), for example, can be generalized as

$$p(C_w) = \frac{\dfrac{\Gamma(\alpha + \frac{w\lambda}{K})}{\Gamma(\frac{w\lambda}{K})}}{\sum_{i=1}^{K} \binom{K}{i} \dfrac{\Gamma(\alpha + \frac{i\lambda}{K})}{\Gamma(\frac{i\lambda}{K})}} ,$$

and (20) in parallel fashion. Like rephrasings are at hand when needed for the α-λ formulas which follow.

19. As remarked in note 14, Hintikka switched from the term 'characteristic function' to 'representative function'. Even though he was still using the former in (1966), in accordance with his later usage I will henceforth use 'representative function'.

20. Hintikka's reasons for preferring his version of the function over Carnap's are presented in (1965c, pp. 28–30) and (1966, p. 119).

21. Details can be found in Hintikka (1965a).

22. GCS is described in Hintikka (1966, pp. 127–28). Hintikka introduced his combined system in (1965c) and generalized it in (1966).

23. Though Pietarinen focuses on inductive generalization, his approach is plainly adaptable to singular induction as well.

24. The emphasis is once again on general inference, but the adaptation to singular inference is straightforward.

25. An early statement is Carnap (1945, pp. 88, 90–93).

26. Kuipers speculates on the reasons for this new attempt in (1978a, p. 262).

27. Kuipers (1978b, Ch. 6) is also of interest on the relations among various Hintikkan systems.

28. Since the prior probabilities of all constituents sum to 1, $p(C_K)$ is fixed by the probabilities of the other constituents and is therefore not a parameter.

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, Harvard University Press, Cambridge, MA.

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

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

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

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 and R. Jeffrey (eds.), Studies in Inductive Logic and Probability, Vol. I, University of California Press, Berkeley, 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, University of California Press, Berkeley, pp. 7–155.

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

Costantini, 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, Kluwer, Dordrecht.

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

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

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, North-Holland, Amsterdam, pp. 220–223.

Hintikka, J.: 1965a, 'Distributive Normal Forms in First-Order Logic', in J. N. Crossley and M.A.E. Dummett (eds.), Formal Systems and Recursive Functions, North-Holland, Amsterdam, 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, North-Holland, Amsterdam, pp. 274–300.

Hintikka, J.: 1965c, 'On a Combined System of Inductive Logic', Acta Philosophica Fennica **18**, 21–30.

Hintikka, J.: 1966, 'A Two-Dimensional Continuum of Inductive Methods', in J. Hintikka and P. Suppes (eds.), Aspects of Inductive Logic, North-Holland, Amsterdam, pp. 113–132.

Hintikka, J.: 1968, 'Induction by Enumeration and Induction by Elimination' and 'Reply', in I. Lakatos (ed.), The Problem of Inductive Logic, North-Holland, Amsterdam, pp. 191–216, 223–231.

Hintikka, J.: 1969, 'Inductive Independence and the Paradoxes of Confirmation', in N. Rescher (ed.), Essays in Honor of Carl G. Hempel, D. Reidel, Dordrecht, pp. 24–46.

Hintikka, J.: 1970, 'On Semantic Information', in J. Hintikka and P. Suppes (eds.), Information and Inference, D. Reidel, Dordrecht, pp. 3–27.

Hintikka, J. and Niiniluoto, I.: 1976, 'An Axiomatic Foundation for the Logic of Inductive Generalization', in M. Przełęcki, K. Szaniawski, and R. Wójcicki (eds.), Formal

Methods in the Methodology of Empirical Sciences, D. Reidel, Dordrecht, and Ossolineum, Wrocław, pp. 57–81.

Jeffrey, R.: 1980, 'Introduction', in R. Jeffrey (ed.), Studies in Inductive Logic and Probability, Vol. II, University of California Press, Berkeley, pp. 1–6.

Kemeny, J.: 1963, 'Carnap's Theory of Probability and Induction', in Schilpp (1963), pp. 711–738.

Kuipers, T.: 1978a, 'On the Generalization of the Continuum of Inductive Methods to Universal Hypotheses', Synthese **37**, 260–72.

Kuipers, T.: 1978b, Studies in Inductive Probability and Rational Expectation, D. Reidel, Dordrecht.

Kuipers, T.: 1984, 'Two Types of Inductive Analogy by Similarity', Erkenntnis **21**, 63–87.

Kuipers, T.: 1988, 'Inductive Analogy by Similarity and Proximity', in Helman (1988), pp. 299–313.

Martin, J.J.: 1967, Bayesian Decision Problems and Markov Chains, Wiley, New York.

Mill, J.S.: 1974, A System of Logic, eds. J.M. Robson and R.F. McRae, University of Toronto and Routledge & Kegan Paul, Toronto (originally published: 1872).

Niiniluoto, I.: 1980, 'Analogy, Transitivity, and the Confirmation of Theories', in L.J. Cohen and M. Hesse (eds.), Applications of Inductive Logic, Oxford University Press, Oxford, pp. 218–234.

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

Niiniluoto, I.: 1988, 'Analogy and Similarity in Scientific Reasoning', in Helman (1988), pp. 271–298.

Pietarinen, J.: 1972, Lawlikeness, Analogy, and Inductive Logic, North-Holland, Amsterdam.

Schilpp, P., ed.: 1963, The Philosophy of Rudolf Carnap, Open Court, La Salle, IL and Cambridge University Press, London.

Skyrms, B.: 1991, 'Carnapian Inductive Logic for Markov Chains', Erkenntnis **35**, 439–460.

Skyrms, B.: 1993, 'Analogy by Similarity in Hyper-Carnapian Inductive Logic', in J. Earman, A. Janis, G. Massey, and N. Rescher (eds.), Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolf Grünbaum, University of Pittsburgh Press, Pittsburgh, pp. 273–282.

Spohn, W.: 1981, 'Analogy and Inductive Logic: A Note on Niiniluoto', Erkenntnis **16**, 35–52.

Stegmüller, W.: 1973, Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie, Band IV, Teil II, Springer-Verlag, Berlin.

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, Philosophical Papers, D. Reidel, Dordrecht, 1977, pp. 4–21.

Wittgenstein, L.: 1922, Tractatus Logico-Philosophicus, Routledge & Kegan Paul, London.