Inductive analogy by similarity and proximity

Kuipers, Theo A.F.

THEO A. F. KUIPERS

INDUCTIVE ANALOGY BY SIMILARITY AND PROXIMITY

1. INTRODUCTION

Modern inductive logic, conceived as a theory of inductive probabilities, is intended to be a useful instrument for learning from experience not only in objective contexts, where there is some underlying objective probability process, but also in other, non-objective contexts. However, for designing inductive systems, objective applications form the primary challenge, for they provide the objective, but unknown, goal which has to be approached in a rational way on the basis of increasing experimental evidence. Hence, in the following we will presuppose an underlying probability process, except when otherwise stated.

Initially there is assumed to be available only some general (true) knowledge about the underlying process, *the prior knowledge*. There also may be some *prior beliefs* (prejudices) which may or may not be true. A suitable inductive system has to start with such prior knowledge and beliefs. As soon as the experimental data may be considered to be literally symptomatic for the unknown objective parameters, the task of designing a system that learns from experience reduces to the task of letting it gradually converge to the data, at the expense of the prejudices. As is well known, Carnap's "continuum of inductive methods" is particularly prepared to approach a multinomial probability process, starting with equal probabilities for all outcomes, but no other prior beliefs are taken into account.

An important type of prior belief concerns considerations of analogy. Carnap distinguished in his posthumously published manuscript notes (Carnap, 1980, pp. 32—71) two fundamentally different kinds of analogical considerations. At the one hand, one outcome may be more similar to another one than a third one. This type of analogy is called *analogy by similarity*, where similarity can be assumed to be based on a distance function, which has to be available independently from the objective probabilities. The other type of analogy is called *analogy by proximity* and concerns differences due to the number of trials between different occurrences of the same outcome.

D. H. Helman (ed.), Analogical Reasoning, 299—313.

Carnap's exposition of analogy by proximity is only very sketchy (pp. 68–71), but also his treatment of analogy by similarity is rather restricted and unsatisfactory; in particular, it did not lead to full fledged inductive systems.

In this chapter we will first summarize, in sections 2 and 3, our construction of inductive systems where prior beliefs of analogy by similarity have a decreasing influence (Kuipers, 1984a, 1984b). These systems can be applied to multinomial processes, where a distance function is given, which may or may not be related (as a matter of objective, but unknown, fact) to differences between the corresponding objective probabilities.

Section 4 deals with one of the two plausible interpretations of "proximity" in "analogy by proximity", viz. proximity of past occurrences, called *proximity in the past.* It turns out to be possible to design systems with this influence, following to some extent (i.e. section 2) the strategy for analogy by similarity. However, it is not clear for what objective processes these systems would be useful, but there may be non-objective applications.

In section 5 Carnap's intended interpretation of "proximity" is considered, viz. proximity of future occurrences, called *proximity in the future*. In the light of the fact that systems having this kind of analogy seem, as far as objective applications are concerned, only useful if the objective process has, or at least may have, this kind of analogy as an objective feature, it is shown that a particular kind of Markov chain has this property and hence that this leads to the general question of how to approach a Markov chain. In section 6 this question is basically conceived as a matter of simultaneously approaching the component multinomial processes constituting a Markov chain.

Section 7 concludes with some evaluative remarks.

As in our earlier publications on inductive logic (see e.g. Kuipers, 1978 or 1980) we will use the mathematical way of presentation. Throughout we will assume a finite set K of mutually exclusive and together exhaustive outcomes $\{Q_1, \ldots, Q_{|K|}\}$, $2 \leqslant K < \infty$. It should be mentioned that "k" will only be used as an arbitrary index and may not be identified with the size of K, i.e. $|K|$. By e_n we indicate a sequence of n outcomes, hence $e_n \in K^n$, and e_nQ_i and e_ne_m are informal notations for $\langle e_n, Q_i \rangle$ and $\langle e_n, e_m \rangle$, respectively.

$\langle K, p \rangle$ indicates a (pure probability) *system*, where p is a probability pattern determined by the *special values* $p(Q_i/e_n)$ (i.e. the probability of Q_i given that e_n is the sequence of outcomes of the first n "experiments" or "trials") and the *product rule* $p(e_n Q_i) = p(e_n)p(Q_i/e_n)$. For simplicity, we will not only assume that the special values are, of course, nonnegative and add up to one, but even that they are positive. The expression $p(e_m/e_n)$ is defined as $p(e_n e_m/e_n)$.

$\langle K, q\rangle$, or simply q, will always indicate an objective system, i.e. a pure system characterizing the objective process mentioned in the context. $\langle K, p\rangle$ may refer to just a pure system, or specifically to an *inductive system*, i.e. a pure system designed to approach some objective system.

All inductive systems to be considered are designed with *initial symmetry*: $p(Q_i) \equiv 1/|K|$. Moreover, they can all be proven to be *closed*: $p(W^m/e_n)$ goes to 0 for proper subsets W of K and increasing m. This excludes positive values for proper universal hypotheses. In the concluding section both restrictions will be cancelled.

Finally, when a distance function on K is assumed, it is supposed to satisfy the standard conditions for a metric distance function: a real-valued non-negative function $d(Q_i, Q_j) = d_{ij}$, such that $d_{ij} = 0$ iff $i = j$, $d_{ij} = d_{ji}$, and $d_{ik} \leqslant d_{ij} + d_{jk}$. In fact, the last property, called triangularity, will not really be used in this chapter.

2. INDUCTIVE SYSTEMS WITH DECREASING VIRTUAL INFLUENCE

Suppose that the special values of a system can be written in the following way:

$$(1) \qquad p(Q_i/e_n) = \frac{n_i + \alpha_i(e_n) + \lambda/|K|}{n + \alpha(n) + \lambda}$$

such that

(2.1) $n_i(e_n) = n_i$ is the number of Q_i-occurrences in e_n, n_i/n is called the *empirical factor*

(2.2) $0 < \lambda < \infty$
$(\lambda/|K|)/\lambda = 1/|K|$ is called the *constant virtual factor*

(2.3) $0 \leqslant \alpha_i(e_n)$; $\sum_i \alpha_i(e_n) = \alpha(n)$

$\alpha_i(e_n)/\alpha(n)$ is called the *variable virtual factor*

(2.4) $\alpha(0) = 0$; $0 < \alpha(1) < 1$; $\beta(n) =_{df} \alpha(n) - \alpha(n-1)$ decreases from $\alpha(1)$ to 0.

It is easy to check that a special value can now be written as a weighted sum of the three mentioned factors, with nonnegative weights adding up to 1. Note that we get Carnap's continuum of inductive methods by putting (improperly, according to (2.4)): $\alpha(n) \equiv 0$.

It is also possible to show that systems satisfying (1) and (2) have the following properties: *instantial confirmation* ($p(Q_i/e_nQ_i) > p(Q_i/e_n)$), *instantial convergence* (s.v. converges to n_i/n); moreover, these systems are *closed.*

Finally, the weight of the empirical factor increases from 0 to 1, the weight of the constant virtual factor decreases from 1 to 0, and the weight of the variable virtual factor first increases from 0 to a certain maximum, and then decreases to 0.

It is important to note that these systems do not satisfy the principle of order indifference (or exchangeability) with respect to the future ($p(Q_iQ_j/e_n) = p(Q_jQ_i/e_n)$), let alone the general principle of order indifference ($p(e'_n) = p(e_n)$ for any order permutation e'_n of e_n). However, they satisfy the principle of order indifference with respect to the past, $p(Q_i/e'_n) = p(Q_i/e_n)$, if and only if $\alpha_i(e'_n) = \alpha_i(e_n)$, where e'_n indicates again an arbitrary permutation of e_n.

In the light of all mentioned properties, and despite the non-properties, it is plausible to call systems satisfying (1) and (2) (inductive) systems with decreasing (constant and variable) virtual influence, or *DV-systems* for short.

3. ANALOGY BY SIMILARITY

In Kuipers (1984a) we have argued, contrary to Carnap's conjecture, that the explication of the idea of analogy by similarity, where similarity is based on a distance function d_{ij}, should not primarily focus on the principle "if $d_{ij} < d_{ik}$ then $p(Q_i/e_nQ_j) > p(Q_i/e_nQ_k)$", but on:

PAS *Principle of analogy by similarity*
if $d_{ij} < d_{ik}$ then
$$p(Q_j/e_nQ_j) - p(Q_j/e_nQ_i) < p(Q_k/e_nQ_k) - p(Q_k/e_nQ_i)$$

The idea behind PAS is the following. It is at least plausible to have *self-similarity*: the probability of Q_j is lower when the last result is not Q_j, but some other Q_i. Formally:

$$0 < p(Q_j/e_nQ_j) - p(Q_j/e_nQ_i)$$

Hence, it is plausible to call this difference the Q_i-*substitution loss* for Q_j. Now PAS requires: if Q_i is closer to Q_j than to Q_k, then the Q_i-substitution loss for Q_j is smaller than for Q_k. It is easy to check that PAS includes self-similarity as a special case (substitute $i = j$). Note also that PAS makes, beside this special case, only sense if $|K|$ is at least 3!

Now it is not difficult to show that a DV-system satisfies PAS if there is an *analogy matrix* of constant fractions $a_j(i)$, satisfying

$$(3) \qquad 0 \leqslant a_j(i) \quad \sum_j a_j(i) = 1 \quad a_i(i) = 0$$

and governing the virtual analogy influence of a Q_i-occurrence for Q_j according to the informal *assignment rule*: if Q_i is closer to Q_j than to Q_k, then Q_j profits more from Q_i than Q_k, formally:

$$(4) \qquad \text{if } d_{ij} < d_{ik} \text{ then } a_j(i) > a_k(i) \quad (i \neq j)$$

and according to the following *rule of distribution*:

$$(5) \qquad \alpha_j(e_nQ_i) = \alpha_j(e_n) + a_j(i) \cdot \beta(n+1)$$

that is, the acquired analogy influence of Q_j increases due to an occurrence of Q_i with the relevant fraction of the total analogy to be distributed, viz. $\beta(n+1)$ ($=_{\text{df}}\ \alpha(n+1) - \alpha(n)$!).

It is easy to see that $\alpha_j(e_n)$ will depend on the order of e_n, and hence that the systems resulting from (1)–(5) do not satisfy order indifference with respect to the past (see section 2), hence they are called Ordered Virtual Analogy by Similarity systems (OVAS-systems).

In Kuipers (1984b) we have shown that we get Unordered VAS-systems (UVAS-systems) when (5) is replaced by (6)

$$(6) \qquad \alpha_j(e_n)/\alpha(n) = \sum_i a_j(i) \cdot (n_i/n)$$

Both systems satisfy Carnap's tentative principle, mentioned at the beginning of this section, as soon as the matrix is symmetric, that is, when $a_j(i) = a_i(j)$. However, this condition is in general not plausible. The neighbourhood of predicates as seen from Q_i (of course including Q_j) may look, as far as distances are concerned, quite different from that seen from Q_j (including Q_i). In that case it is plausible that Q_i and Q_j have to distribute their analogy influence differently, and hence $a_j(i) \neq a_i(j)$ may result.

The paradigmatic context for using a VAS-system certainly is a multinomial process, i.e. a sequence of independent trials with (of course unknown) positive objective probabilities q_i for all $Q_i \in K$ (adding up to 1), where there is some reason to assume initially that similarity will imply, as a matter of empirical fact, similar probabilities. In other words, decreasing distance is assumed to be correlated with decreasing difference between the corresponding objective probabilities. In formal terms:

$$\text{if } d_{ij} < d_{ik} \text{ then } |q_i - q_j| < |q_i - q_k|$$

Being a prejudice, it may turn out to be right or wrong. VAS-systems are such that the role of the prejudice is gradually replaced by genuine matters of empirical fact. To get this kind of effect, it is necessary to abandon at least order indifference with respect to the future, despite the fact that the order of outcomes is by definition in no way constrained in a multinomial context. To be precise: it is easy to prove that the only systems satisfying (1) and the principle $p(Q_iQ_j/e_n) = p(Q_jQ_i/e_n)$ are Carnap-systems (systems in which $\alpha(n) \equiv \alpha_i(e_n) \equiv 0$), which evidently lack the described effect.

In the rest of this chapter the order of outcomes of underlying objective processes will be assumed to be constrained, in some way or other, by the nature of these processes.

4. ANALOGY BY PROXIMITY IN THE PAST

There are at least two ways to explicate the intuitive idea of analogy by proximity. The closest one to Carnap's sketch of the idea, will be presented in the next section. Here we will concentrate on an interpretation of which the first part can be treated along the same lines as analogy by similarity.

The idea is that an outcome profits more from a recent self-occurrence than from a self-occurrence longer ago. Carnap just hinted upon this idea in 1950 (Carnap, 1950, pp. 63–64). Formally, the elementary form of the idea can be expressed as follows:

PAPP *Principle of analogy by proximity in the past*
$p(Q_i/e_nQ_jQ_i) > p(Q_i/e_nQ_iQ_j)$ for all $i, j, e_n, i \neq j$

and the general form reads:

PAPP^G $p(Q_i/e_nQ_je_mQ_i) > p(Q_i/e_nQ_ie_mQ_j)$ for all $i, j, e_n, e_m, i \neq j$

Of course, application of a system satisfying PAPP will only be considered when there is reason to assume that the underlying objective process is also governed by something like PAPP. Unfortunately, we do not know of objective processes having PAPP$^{(G)}$ in a natural way.

Fortunately, there may be non-objective applications. That is, PAPP, or something like it, may be motivated on the basis of increasing uncertainty of past reports or (and) past methods of observation.

For the construction of systems satisfying PAPP we start from inductive systems with decreasing virtual influence (DV-systems) as defined in section 2; hence the resulting systems will satisfy all properties mentioned there.

The distribution of the virtual influence will now of course be completely different from the method used in section 3 for analogy by similarity. In order to realize PAPP it is plausible to add to (1) and (2):

(7) $\alpha_i(e_nQ_i) = \gamma(n+1)\alpha_i(e_n) + \beta'(n+1)$
$\alpha_i(e_nQ_j) = \gamma(n+1)\alpha_i(e_n) \quad i \neq j$

(8) $0 < \gamma(n) < 1$ ($\gamma(n)$ may be constant!)

From (2), (7) and (8) it follows that $\beta'(n+1) = \alpha(n+1) - \gamma(n+1)\cdot\alpha(n)$ and that $0 < \beta(n+1) =_{df} \alpha(n+1) - \alpha(n) < \beta'(n+1) < 1$.

Now it is easy to check that (7) and (8) directly lead to the realization of the informal idea that a self-occurrence in the last trial is more profitable than the one before as soon as we add the condition that guarantees $\alpha_i(e_nQ_jQ_i) > \alpha_i(e_nQ_iQ_j)$:

(9) $\beta'(n+1) > \gamma(n+1)\beta'(n)$

This leads, via (1) and (2), to the satisfaction of PAPP. Due to the decreasing character of the virtual influence, guaranteed by (2.4), there will be for any n some $m_n \geqslant 0$ such that for all $m \leqslant m_n$,

$\alpha_i(e_nQ_je_mQ_i)$ is larger than $\alpha_i(e_nQ_ie_mQ_j)$, but that it becomes smaller for $m > m_n$. Hence, PAPP^G will be satisfied for all n (only) up to this m_n, which is a plausible kind of restriction if and only if the objective process may be thought to be restricted in a similar way.

The resulting systems, i.e. systems satisfying (1), (2), (7), (8) and (9), will be called VAPP-systems. Of course, there are many ways in which the parametric conditions can be satisfied. One obtains a relatively simple example as follows: $0 < \gamma(n) \equiv \gamma < 1$, $0 < \beta < 1$, $0 < x < 1$, such that $\beta(n+1) = x^n\beta$. Hence, $\alpha(n) = \beta(1-x^n)/(1-x)$, and also $\beta'(n+1)$ can be expressed in terms of β, γ and x. Condition (9) leads to a last restriction: $x > \gamma$ or $1 + x > 2\gamma$.

VAPP-systems are conceptually so construed, by (7) and (8), that the already acquired analogy due to previous self-occurrences devaluates after each trial, in exchange for new profit from an eventual self-occurrence at the last trial, which leads to the desired effect (PAPP). [As a matter of technical coincidence, due to the decreasing character of $\beta(n+1)$, the same effect occurs in OVAS-systems (section 3), apart from the fact that there arise differences for different outcomes, except when the fractions $a_j(i)$ $(i \neq j)$ are all equal, and hence equal to $1/(|K|-1)$.]

Note also that the improper OVAS-system with $a_i(i) = 1$, and hence all other fractions zero, has an effect opposite to PAPP, just because $\beta(n+1)$ decreases. This system corresponds technically with a potential VAPP-system satisfying (1), (2) and (7) such that $\gamma(n) \equiv 1$, and hence $\beta'(n+1) = \gamma(n+1)$. But it does of course not satisfy (9) (which would even reduce to *in*creasing $\beta(n+1)$).

We conclude this section with a general remark. It is clear that VAPP-systems satisfy PAPP such that the positive difference $p(Q_i/e_nQ_jQ_i) - p(Q_i/e_nQ_iQ_j)$ decreases with increasing n. Objective processes with this effect may of course exist. But objective processes where this difference remains constant, that is, at least globally conceived, will also exist and may be more interesting. However, VAPP-systems are not well equipped for such processes, because there is no mechanism included which guarantees the gradual replacement of the decreasing $\beta(n+1)$ by the "experimental expression" of the corresponding objective value.

5. ANALOGY BY PROXIMITY IN THE FUTURE

Formulated in our way of presentation, Carnap (1980, pp. 68—71) had basically the following in mind with "analogy by proximity": after an occurrence of Q_i the probability that Q_i occurs again after n trials is greater than that it re-occurs after $n + 1$ trials. Formally, in an obvious notation, the elementary version reads:

PAPF *Principle of analogy by proximity in the future*

$$p(Q_i/e_nQ_i) =_{df} p^{(1)}(Q_i/e_nQ_i) >$$

$$p^{(2)}(Q_i/e_nQ_i) =_{df} \sum_{Q_j \in K} p(Q_jQ_i/e_nQ_i) \quad \text{for all } i, e_n$$

and the general version:

PAPF^G $p^{(m)}(Q_i/e_nQ_i) > p^{(m+1)}(Q_i/e_nQ_i)$ for all i, e_n, m

where the following definition is presupposed:

$$p^{(m)}(Q_i/e_nQ_i) = \sum_{e_m \in K^m} p(e_mQ_i/e_nQ_i)$$

Note that when the system satisfies the general principle of (positive) instantial confirmation (relevance); $p^{(m)}(Q_i/e_nQ_i) > p(Q_i/e_n)$, then PAPF^G might well be called the principle of decreasing positive relevance.

Before we try to construct inductive systems satisfying PAPF$^{(G)}$ we will consider the question of what type of objective probability process satisfies it. There are of course many possibilities, hence the problem is to select one or more basic categories. Here we will restrict our attention to a suitable subclass of Markov chains.

As is well known (Feller, 1968, Ch. XV), in a Markov chain K is considered as the set of possible states of a concrete system, and there are initial probabilities a_i ($a_i \geqslant 0$, $\Sigma_i a_i = 1$) that the system starts in Q_i and there are fixed transition probabilities q_{ij} ($q_{ij} \geqslant 0$, $\Sigma_j q_{ij} = 1$) for the transition from Q_i to Q_j.

If there is a relevant distance function d_{ij} the following relations may hold (as a matter of fact or definition), for all i, j, k:

(10.1) $q_{ij} > q_{ik}$ iff $d_{ij} < d_{ik}$

(10.2) $q_{ji} > q_{ki}$ iff $d_{ij} < d_{ik}$

Note that (10) is guaranteed by the general condition

(11) $q_{ij} > q_{kl}$ iff $d_{ij} < d_{kl}$

Expressed in appealing terms, (10.1) implies, due to $d_{ii} = 0 < d_{ij}$ $(i \neq j)$, that all states are

selfish: $q_{ii} > q_{ij}$ $(i \neq j)$

whereas (10.2) implies that they are

unattractive: $q_{ii} > q_{ji}$ $(i \neq j)$

As suggested, it may well be that the distance function is definitorically based on the transition probabilities, but in a proper inductive context we do not know this link.

Higher order transition probabilities in a Markov chain are of course defined recursively such that the probability of reaching Q_j, starting from Q_i, in $m + 1$ steps, is just the sum of all probabilities of reaching an arbitrary Q_k in m steps immediately followed by Q_j. Formally: $q_{ij}^{(m+1)} = \Sigma_k q_{ik}^{(m)} q_{kj}$. It is easily shown that these probabilities satisfy

$$q_{ij}^{(n+m)} = \sum_k q_{ik}^{(n)} q_{kj}^{(m)}$$

For a Markov chain PAPF and PAPF^G respectively reduce to

(12) $q_{ii} = q_{ii}^{(1)} > q_{ii}^{(2)}$

(12)^G $q_{ii}^{(m)} > q_{ii}^{(m+1)}$ for all m

Markov chains satisfying (12) will be called (elementary) *divergent* and those satisfying even (12)^G generally divergent.

A natural question is what condition on (first order) transition probabilities guarantees (12)^G. Unfortunately, we only succeeded in (easily) proving

if $q_{ii} > q_{ji}$ for all $j \neq i$ then $q_{ii} > q_{ii}^{(m)}$ for all $m > 1$
if $q_{ii}^{(m)} > q_{ji}^{(m)}$ for all $j \neq i$ then $q_{ii}^{(m)} > q_{ii}^{(m+1)}$ for all m.

In words: if all states are unattractive then the chain is at least "globally" divergent; if all states are also "higher order" unattractive the chain is generally divergent.

In general, it is our impression that, although there is an enormously extented literature on Markov chains, properties such as $(12)^G$ have not been studied explicitly.

A divergent Markov chain is nevertheless a Markov chain, and hence the question of how to approach such a chain by an inductive system leads to the general question of approaching in a rational way a Markov chain.

6. HOW TO APPROACH A MARKOV CHAIN?

Let us assume to be confronted with a Markov chain with set of states K and positive, but further unknown, initial and transition probabilities. Now it is important to realize that a Markov chain is essentially a combination of $|K|$ multinomial process: for each i, q_{ij}, $j = 1, \ldots,$ $|K|$, forms a multinomial process. Assuming that they are mutually unrelated, it is plausible to decompose the evidence and to approach these processes independently by Carnap system.

Let $n_{ij}(e_n) = n_{ij}$ indicate the number of times that an occurrence of Q_i in e_n is followed by an occurrence of Q_j and let e_n^i indicate that the last member of e_n is Q_i. Of course we have $\Sigma_j n_{ij}(e_n^k) = n_i$ if $i \neq k$ and $\Sigma_j n_{ij}(e_n^i) = n_i - 1$.

The decomposition idea now leads straightforwardly to

$$(13) \qquad p(Q_j/e_n^i) = \frac{n_{ij} + \lambda/|K|}{n_i - 1 + \lambda}$$

Instantial confirmation is now of course only satisfied if we consider the multinomial processes separately. On the other hand, the fact that the component systems are closed implies immediately that the resulting compound system is also closed.

From (13) it follows immediately that $p(Q_i/e_n^i) > p(Q_j/e_n^i)$ $(i \neq j)$ if and only if $n_{ii} > n_{ij}$. The latter will on the average be the case if Q_i is selfish ($q_{ii} > q_{ij}$, for all j). On the other hand there is no simple condition on e_n that guarantees $p(Q_i/e_n^k Q_i) > p(Q_i/e_n^k Q_j)$ $(i \neq j, k)$, but it is nevertheless clear that this is on the average the case when Q_i is unattractive ($q_{ii} > q_{ji}$, for all j).

The crucial question is of course whether the resulting system satisfies $\text{PAPF}^{(G)}$ when the underlying Markov chain is divergent (i.e. satisfies it). Unfortunately, this is not uniformly the case, but again we may assume that e_n is on the average such that at least PAPF holds.

Systems defined by (13) will generally be called Markov (chain) Approaching systems (MA-systems).

What is the correct approach when we only know that the objective process is a proper Markov chain *or* a multinomial process? Note that a multinomial process is a degenerate Markov chain in the sense that all transition probabilities q_{ij} are independent of i, and hence $q_{ij} = a_j$. Of course, the foregoing decomposed procedure will in the limit lead to this degenerate case, i.e. $p(Q_j/e^i_n)$ will approach a_j, for both $n_{ij}/(n_i - 1)$ and n_j/n approach it.

There may, however, be good reasons to take both options more on an equal footing. One possibility seems at first sight:

$$p(Q_j/e^i_n) = \frac{n_j + n_{ij} + \lambda/|K|}{n + n_i - 1 + \lambda}$$

If the underlying process is in fact multinomial there is no problem, but if it is a Markov chain the "multinomial arrangements" will not disappear.

The following weighted version meets this problem:

$$p(Q_j/e^i_n) = \rho(n)\frac{n_{ij} + \lambda_{ma}/|K|}{n_i - 1 + \lambda_{ma}} + (1 - \rho(n))\frac{n_j + \lambda_{mu}/|K|}{n + \lambda_{mu}}$$

where $\rho(n)$ increases from $\rho(0) = 0$ to 1.

7. CONCLUDING REMARKS

We start with some remarks about the generality of the presented inductive systems. To begin with, it is easy to see that it is possible to change all systems such that they do not start with equal probability for all outcomes, but with outcome-dependent specific values, based on some "width-function" on the outcomes and the prior belief that this function is representative for the relevant objective probabilities.

The second plausible generalization concerns *open* systems. Open systems are based on a (non-trivial) *prior distribution*, on the so-called constituents, and *conditional systems* (see Kuipers, 1978 or 1980, for precise definitions). When the conditional systems are chosen with some kind of analogy influence, it will have effect of this kind on the resulting unconditional systems, although it may be difficult to characterize this effect in detail.

It should here be mentioned that open systems leave room for a kind of analogy by similarity which is excluded for closed systems. The choice of the prior distribution may be based on the distances between the outcomes of which a constituent claims that they are exemplified in the universe of the context. In Kuipers (1984a) we have studied this (other) type of analogy by similarity, there called "existential analogy" (as opposed to "instantial analogy", which was summarized in this chapter), in particular the effect, of plausible restrictions of this kind, on the posterior distribution.

Let us now come back to the results of the present paper. Although the treatment of analogy by proximity is certainly not yet as satisfactory as that of analogy by similarity, there seems to be some progress; first, the clear distinction between two interpretations of "proximity"; second, for proximity in the past, the distinction between objective processes having the feature in a stable way, or only in a decreasing way; third, a suitable inductive system to approach the latter kind of process; fourth, and finally, the reformulation of a restricted kind of proximity in the future as the important problem of approaching a Markov chain. For this last mentioned problem, a first solution has also been presented.

It is clear that many questions are still left open. However, to be honest, as far as purely objective applications are concerned, we have some doubts about whether the idea of analogy by proximity is after all very important as such. It may well be that the intuitive idea is always translatable as a, relatively arbitrary, special kind of the problem of approaching some well known type of objective process, such as a Markov chain. A so-called "Markov process" is another plausible and more sophisticated possibility. But, before attacking this last problem, it seems advisable to consider first the question whether there are better ways of approaching a Markov chain than the first attempt presented in this chapter.

To be sure, analogy by proximity may well be more important for non-objective applications, e.g. in the light of the decreasing value of past reports or methods of observation. However, in a purely non-objective context it is by definition impossible to formulate a convergence criterion as a condition of adequacy. Hence, the best thing to do in this connection may be to focus first on a mixed context, e.g. an objective multinomial context in which there is nevertheless decreasing certainty about past reports or methods of observation.

We conclude with some remarks about the relevance of inductive logic as presented here. As argued in Kuipers (1984c), the original motivation from philosophy of science for inductive logic is largely outdated. But it is becoming more and more clear that inductive systems are useful in statistics, and not only as sophisticated versions of Bayesian statistics (see e.g. Festa, 1989). Finally, it is not only evident that inductive systems can easily be put at work in a computer programme, the challenge to design learning programmes will more and more include the task of programming the learning of objective probability processes in a responsible way, as a special branch of approaching the truth. In other words, there is arising a new area of applications of, and questions for, inductive logic.

ACKNOWLEDGEMENTS

The author would like to thank Roberto Festa, David Helman, Ilkka Niiniluoto and the anonymous referee for their comments on the first version.

Department of Philosophy
University of Groningen
Westersingel 19
9718 CA Groningen
The Netherlands

REFERENCES

Carnap, R. (1950), *Logical Foundations of Probability*, Chicago: Univ. of Chicago Press, second impression 1963.

Carnap, R. (1980), 'A basic system of inductive logic, Part 2', in Jeffrey (ed.), pp. 7–155.

Feller, W. (1968), *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3 ed., New York: Wiley.

Festa, R. (1989), *Similarity, Truthlikeness and the Optimum Inductive Method*, in press.

Jeffrey, R. (ed.) (1980), *Studies in Inductive Logic and Probability*, Vol. II, University of California Press.

Kuipers, T. (1978), *Studies in Inductive Probability and Rational Expectation*, Synthese Library 123, Dordrecht: Reidel.

Kuipers, T. (1980), 'A survey of inductive systems', in Jeffrey (ed.), pp. 183–192.

Kuipers, T. (1984a), 'Two types of inductive analogy by similarity', *Erkenntnis* **21**, pp. 63–87.

Kuipers, T. (1984b), 'Inductive analogy in Carnapian spirit', *PSA* **1**, pp. 157–167.

Kuipers, T. (1984c), 'What remains of Carnap's program today?', spec. issue *Epistemologia*, Agazzi, E. and Costantini, D. (eds.), pp. 121–152.