An Observation on Carnap's Continuum and Stochastic Independencies

J. B. Paris

School of Mathematics
The University of Manchester
Manchester M13 9PL

jeff.paris@manchester.ac.uk

Abstract

We characterize those identities and independencies which hold for all probability functions on a unary language satisfying the Principle of Atom Exchangeability. We then show that if this is strengthen to the requirement that Johnson's Sufficientness Principle holds, thus giving Carnap's Continuum of inductive methods for languages with at least two predicates, then new and somewhat inexplicable identities and independencies emerge, the latter even in the case of Carnap's Continuum for the language with just a single predicate.

Keywords: Carnap's Continuum, Atom Exchangeability, Stochastic Independence, Inductive Logic.

Introduction

To this day the *Continuum of Inductive Methods* described by Carnap in [1], [2], [3], [4] continues to be adapted and promoted as paradigm solutions to various problems within Inductive Logic. For example arithmetic combinations of these functions figure almost exclusively in recent attempts to provide probability functions exhibiting certain specific features of analogical influence, see [5], [11], [12], [15], [16].

There seem to be several good reasons for this focus. Firstly this Continuum has a widely acceptable justification in terms of its 'rationality': There is a putatively rational requirement, namely Johnson's Sufficientness Principle, that we can impose on an inductive method (i.e. probability function) which forces it to be precisely a member of Carnap's Continuum (see also Johnson's earlier derivation of this in [9]), at least when we assume that the language has more than one predicate. Secondly the Continuum has a simple form, making it easy to work with, whilst the parameter it involves has a clear interpretation which readily permits generalizations.

Carnap's original goal in his Inductive Logic programme was to develop an inductive method which could be applied to real world problems of induction, or more generally the assignment of probabilities based on some finite body of evidence, and which furthermore was *logical* in the sense that it's conclusions followed mechanically from the evidence via certain precisely formulated rules or principles. The arrival on the scene of Goodman's Grue Paradox, [6], [7], however highlighted an evident flaw in the *practicality* of the approach; that in real (as opposed to toy) examples there is usually so much available evidence that even if it could be suitably formulated in the language of the problem it would be completely infeasible to take it as one's premise set.

Whilst many philosophers have seen this as the end of the programme as a practical, rather than simply a theoretical, project, nevertheless apparently similar aspirations to Carnap's still seem to underlie papers such as those on analogical reasoning cited above. One explanation for this is that whilst *all* our available knowledge in an real world situation is just too much to handle nevertheless most of it should be redundant or irrelevant and possibly what really does matter *can* be simply formulated. This raises the question we shall consider in this paper, to what extent is this a reasonable assumption for the members of Carnap's Continuum, more precisely under what circumstances is a sentence θ stochastically independent of a sentence ϕ for all members of Carnap's Continuum?

Before that however we need to spend a little time introducing some standard notation. The experienced reader might therefore be advised to skip the next section, only referring back to it as necessary.

Notation

Let L be a predicate language with just q (unary) predicates, $P_1, P_2, \ldots, P_q$, constants a_1 for $i = 1, 2, 3, \ldots$ and no other relation, constant or function symbols. As usual the intention here is that these a_i exhaust the universe.

Let $\alpha_1(x), \alpha_2(x), \ldots, \alpha_{2^n}(x)$ denote the *atoms* of L, that is the 2^q formulae of L of the form

$$\pm P_1(x) \wedge \pm P_2(x) \wedge \ldots, \wedge \pm P_q(x).$$

So for example the atoms in the case $q = 2$ are $P_1(x) \wedge P_2(x)$, $P_1(x) \wedge \neg P_2(x)$, $\neg P_1(x) \wedge P_2(x)$, $\neg P_1(x) \wedge \neg P_2(x)$. Knowing which atom an a_i satisfies tells us exactly which of the $P_j(x)$ a_i does or does not satisfy, and hence tells us everything there is to know about a_i.

A *state description*, $\Theta(b_1, b_2, \ldots, b_m)$, for distinct choices $b_1, b_2, \ldots, b_m$ from the a_i, is a sentence of the form

$$\bigwedge_{i=1}^{m} \alpha_{j_i}(b_i), \tag{1}$$

and similarly tells us all there is to know about $b_1, b_2, \ldots, b_m$.

Notice that the state descriptions for $b_1, b_2, \ldots, b_m$ are disjoint and any quantifier free sentence $\phi(b_1, b_2, \ldots, b_m)$ of L is logically equivalent to a disjunction

$$\bigvee_{k=1}^{s} \Theta_k(b_1, b_2, \ldots, b_m)$$

of distinct state descriptions $\Theta_k(\vec{b})$ for $b_1, b_2, \ldots, b_m$. Hence if w is a probability function on L (for a definition see for example [8] or [14]) then

$$w(\phi(b_1, b_2, \ldots, b_m)) = \sum_{k=1}^{s} w(\Theta_k(b_1, b_2, \ldots, b_m)). \tag{2}$$

We say that w satisfies *Constant Exchangeability, Ex,* if $w(\phi(b_1, b_2, \ldots, b_m))$ depends only on $\phi(x_1, x_2, \ldots, x_m)$ and not on the (distinct) instantiating constants $b_1, b_2, \ldots, b_m$. By (2) it is already enough that this holds for state descriptions. Since all the probability functions we shall consider will satisfy Ex our results will apply for general $b_1, b_2, \ldots, b_m$ once proven for $a_1, a_2, \ldots, a_m$.

The *spectrum* of a state description $\Theta(b_1, \ldots, b_m)$ as in (1) is the multiset[1] $\overline{n} = \{\, n_1, n_2, \ldots, n_{2^q} \,\}$, where n_i is the number of times that the atom $\alpha_i(x)$ appears amongst the $\alpha_{j_1}(x), \alpha_{j_2}(x), \ldots, \alpha_{j_m}(x)$.

[1] Multisets are just like sets except that elements may be repeated.

We say that w satisfies *Atom Exchangeability, Ax,* if $w(\Theta(b_1, b_2, \ldots, b_m)$ depends only on the spectrum $\overline{n}$ of the state description $\Theta(b_1, b_2, \ldots, b_m)$. In this case we shall write $w(\overline{n})$ for $w(\Theta(\vec{b}))$.

Finally we say that w satisfies *Johnson's Sufficientness Principle, JSP,* if for a state description $\Theta(b_1, b_2, \ldots, b_m)$ as in (1), $w(\alpha_i(b_{m+1}) \,|\, \Theta(b_1, b_2, \ldots, b_m))$ depends only on m and n_i. It is well known that JSP implies Ax which in turn implies Ex.

As shown originally by Johnson, [9] (and independently later by Kemeny, see [4, section 19] and [10]) if the number of predicates, q, is at least 2 and the probability function w satisfies JSP then w is a member of *Carnap's Continuum of Inductive Methods*. That is, $w = c_\lambda$ for some $0 \leq \lambda \leq \infty$ where, with the above notation, c_λ is the probability function satisfying Ax such that

$$c_\lambda((\alpha_i(b_{m+1}) \wedge \Theta(b_1, b_2, \ldots, b_m)) \;=\; \frac{(n_i + \lambda/2^q)}{(m + \lambda)} \cdot c_\lambda(\Theta(b_1, b_2, \ldots, b_m)). \quad (3)$$

The cases $\lambda = 0, \infty$ here are rather exceptional and until further notice we shall restrict ourselves to $0 < \lambda < \infty$ when discussing the c_λ (though still referring to these as Carnap's Continuum.)

Stochastic Independence and Ax

Let w be a probability function on L satisfying Ax. Then from (2) for $\phi(a_1, \ldots, a_m)$ a sentence of L,

$$w(\phi(a_1, \ldots, a_m)) = \sum_{\overline{n}} f_\phi(\overline{n}) w(\overline{n}),$$

where the $\overline{n}$ range over the possible spectra $\{ n_1, n_2, \ldots, n_{2^q} \}$ with $\sum_{i=1}^{2^q} n_i = m$ and $f_\phi(\overline{n})$ is the number of state descriptions in (2) with spectrum $\overline{n}$.

Hence the stochastic independence[2] of $\theta(a_1, \ldots, a_m)$ and $\phi(a_1, \ldots, a_m)$ with respect to w, i.e.

$$w(\theta(a_1, \ldots, a_m) \wedge \phi(a_1, \ldots, a_m)) = w(\theta(a_1, \ldots, a_m)) \cdot w(\phi(a_1, \ldots, a_m))$$

[2]If $w(\phi) \neq 0$ we may replace this condition by the equivalent $w(\theta \,|\, \phi) = w(\theta)$ if in the context that seems more appropriate.

amounts to the identity,

$$\left(\sum_{\overline{n}} f_{\theta\wedge\phi}(\overline{n})w(\overline{n})\right) = \left(\sum_{\overline{n}} f_{\theta}(\overline{n})w(\overline{n})\right) \cdot \left(\sum_{\overline{n}} f_{\phi}(\overline{n})w(\overline{n})\right),$$

equivalently,

$$\left(\sum_{\overline{n}} f_{\theta\wedge\phi}(\overline{n})w(\overline{n})\right) \cdot \left(\sum_{\overline{n}} f_{\top}(\overline{n})w(\overline{n})\right) = \left(\sum_{\overline{n}} f_{\theta}(\overline{n})w(\overline{n})\right) \cdot \left(\sum_{\overline{n}} f_{\phi}(\overline{n})w(\overline{n})\right) \tag{4}$$

since

$$\sum_{\overline{n}} f_{\top}(\overline{n})w(\overline{n}) = 1.$$

We can now turn (briefly as it happens) to the main question we are interested in: When does (4) hold for c_λ with $0 < \lambda < \infty$? Given the aspirations outlined in the first section one may hope that we should certainly have independence when $\theta(\vec{a})$ and $\phi(\vec{a})$ are respectively logically equivalent to sentences $\theta'(a_1, a_2, \ldots, a_r)$ and $\phi'(a_{r+1}, a_{r+2}, \ldots, a_m)$ which have no predicates nor constants in common. However, as already pointed out in [8][3], this can fail, for example for $0 < \lambda < \infty$,

$$c_\lambda(P_2(a_3) \wedge P_2(a_4) \,|\, P_1(a_1) \wedge P_1(a_2)) > c_\lambda(P_2(a_3) \wedge P_2(a_4)).$$

Given this observation one might rashly be inclined to quite the opposite view, that the c_λ do not satisfy any independencies such as (4) except in the trivial cases when $\pm\theta \equiv \top$ or $\pm\phi \equiv \top$. This is not true however in at least two ways. Firstly the identity (4) with $w = c_\lambda$ is equivalent to a polynomial identity in the variable λ which will sometimes hold for a finite set of roots λ. We will dispense with such chance independencies by considering which independencies of the form (4) hold for *all* the c_λ (with $0 < \lambda < \infty$).

There is however a second way in which (4) can hold for all the c_λ. For if $q \geq 2$ then the $f(\overline{n})$ will have common divisors and simply by taking disjunctions of

[3]On the other hand there is an argument why this inequality is desirable: Namely the $P_1(a_1) \wedge P_1(a_2)$ provides evidence that the individuals a_i are similar and hence should support the view that a_3, a_4 are similar, in particular positively supporting $P_2(a_3) \wedge P_2(a_4)$.

state descriptions we can construct sentences $\theta(a_1, a_2, \ldots, a_m), \phi(a_1, a_2, \ldots, a_m)$ such that for some constant k, with $0 < k < 1$, and each $\overline{n}$

$$f_{\theta\wedge\phi}(\overline{n}) = kf_{\phi}(\overline{n}), \quad f_{\theta}(\overline{n}) = kf_{\top}(\overline{n}).$$

In this case the equality (4) will of course hold for *all* probability functions w satisfying Ax. For that reason we shall now temporarily break off from the main interest of this paper and consider the question of what identities and independencies must hold for all probability functions satisfying Ax.[4] In fact Theorem 3 below will show that the independencies described above are exactly those which hold for all probability functions w satisfying Ax. First though we will derive some special cases which are perhaps of independent interest.

Proposition 1. *If equation (4) holds for all probability functions w on L satisfying Ax then either*

$$\sum_{\overline{n}} f_{\theta}(\overline{n})w(\overline{n})$$

is constant for all probability functions w satisfying Ax or

$$\sum_{\overline{n}} f_{\phi}(\overline{n})w(\overline{n})$$

is constant for all probability functions w satisfying Ax.

Proof Suppose that (4) held for every probability function w satisfying Ax but that for two such functions w_1, w_2,

$$\sum_{\overline{n}} f_{\theta}(\overline{n})w_1(\overline{n}) \neq \sum_{\overline{n}} f_{\theta}(\overline{n})w_2(\overline{n}). \tag{5}$$

In this case θ cannot be logically equivalent to $\perp$ and by small perturbations of w_1, w_2 if necessary we may assume that neither side of (5) is zero.

[4]It is easy to see that if we weaken Ax to Ex then the only independencies satisfied by all probability functions satisfying Ex are those of the form $w(\theta \wedge \phi) = w(\theta) \cdot w(\phi)$ when $\pm\theta \equiv \top$ or $\pm\phi \equiv \top$.

The probability function $(w_1 + w_2)/2$ will also satisfy Ax and if according to w_i $(i = 1, 2)$ (4) gives $A_i = B_i C_i$ where

$$A_i = \sum_{\overline{n}} f_{\theta \wedge \phi}(\overline{n}) w_i(\overline{n}), \quad B_i = \sum_{\overline{n}} f_{\phi}(\overline{n}) w_i(\overline{n}), \quad C_i = \sum_{\overline{n}} f_{\theta}(\overline{n}) w_i(\overline{n})$$

then according to $(w_1 + w_2)/2$ (4) gives

$$2(A_1 + A_2) = (B_1 + B_2)(C_1 + C_2).$$

If $B_1 B_2 = 0$ this together with (5) gives $B_1 = B_2\ (= 0)$. Otherwise multiplying by $B_1 B_2$ and eliminating the C_1, C_2 gives

$$(A_1 B_2 - A_2 B_1)(B_1 - B_2) = 0$$

and again with (5) it follows that $B_1 = B_2$. In other words the second possibility in the proposition pertains for w_1 and w_2.

Now if the proposition failed there would be w_1, w_2, w_3, w_4 such that

$$\sum_{\overline{n}} f_{\theta}(\overline{n}) w_1(\overline{n}) \neq \sum_{\overline{n}} f_{\theta}(\overline{n}) w_2(\overline{n}),$$

$$\sum_{\overline{n}} f_{\phi}(\overline{n}) w_1(\overline{n}) = \sum_{\overline{n}} f_{\phi}(\overline{n}) w_2(\overline{n}),$$

$$\sum_{\overline{n}} f_{\theta}(\overline{n}) w_3(\overline{n}) = \sum_{\overline{n}} f_{\theta}(\overline{n}) w_4(\overline{n}),$$

$$\sum_{\overline{n}} f_{\phi}(\overline{n}) w_3(\overline{n}) \neq \sum_{\overline{n}} f_{\phi}(\overline{n}) w_4(\overline{n}),$$

with all such quantities non-zero. But by applying the same argument as above for $(w_1 + w_3)/2$ and $(w_2 + w_4)/2$ we see that this is not possible. The result follows. ■

Proposition 2. *With the notation of the previous proposition suppose that for all probability functions w satisfying Ax that*

$$\sum_{\overline{n}} f_{\theta}(\overline{n}) w(\overline{n}) = k \tag{6}$$

for some constant k. Then for each $\overline{n}$, $f_{\theta}(\overline{n}) = k f_{\top}(\overline{n})$.

Proof Given reals $s_1, s_2, \ldots, s_{2^q} \geq 0$ and not all zero let $w_{\vec{s}}$ be the probability function on L such that

$$w_{\vec{s}}(\overline{n}) = (2^q!)^{-1} \sum_{\sigma} s_{\sigma(1)}^{n_1} s_{\sigma(2)}^{n_2} \ldots s_{\sigma(2^q)}^{n_{2^q}} (s_1 + s_2 + \ldots + s_{2^q})^{-m}$$

where σ ranges over the permutations of $1, 2, \ldots, 2^q$. Then $w_{\vec{s}}$ satisfies Ax and (6) gives that

$$\sum_{\overline{n}} f_\theta(\overline{n}) (2^q!)^{-1} \sum_{\sigma} s_{\sigma(1)}^{n_1} s_{\sigma(2)}^{n_2} \ldots s_{\sigma(2^q)}^{n_{2^q}} = k(s_1 + s_2 + \ldots + s_{2^q})^m.$$

Since we can take each s_i to be algebraically independent this is only possible if the coefficients of $s_1^{n_1} s_2^{n_2} \ldots s_{2^q}^{n_{2^q}}$ on both sides agree, from which the result follows. ∎

Theorem 3. *Equation (4) holds for all probability functions w on L satisfying Ax if and only if $\pm\theta \equiv \top$ or $\pm\phi \equiv \top$ or for some constant k, $0 < k < 1$,*

$$f_{\theta \wedge \phi}(\overline{n}) = k f_\phi(\overline{n}), \quad f_\theta(\overline{n}) = k f_\top(\overline{n}) \qquad \text{for all } \overline{n} \tag{7}$$

or for some constant k, $0 < k < 1$,

$$f_{\theta \wedge \phi}(\overline{n}) = k f_\theta(\overline{n}), \quad f_\phi(\overline{n}) = k f_\top(\overline{n}) \qquad \text{for all } \overline{n}. \tag{8}$$

Proof Assume that θ, ϕ are neither tautologies nor contradictions. In the first case of Proposition 1 we may assume that

$$\sum_{\overline{n}} f_\theta(\overline{n}) w(\overline{n}) = k \tag{9}$$

for some constant k, with $0 < k < 1$, and for all probability functions w satisfying Ax. Hence from (4),

$$\sum_{\overline{n}} f_{\theta \wedge \phi}(\overline{n}) w(\overline{n}) = k \sum_{\overline{n}} f_\phi(\overline{n}) w(\overline{n}).$$

Now using $w_{\vec{s}}$ as in the proof of Proposition 2 we obtain that $f_{\theta \wedge \phi}(\overline{n}) = k f_\phi(\overline{n})$ for all $\overline{n}$ and since we already have $f_\theta(\overline{n}) = k f_\top(\overline{n})$ for all $\overline{n}$, again by Proposition 2 with (9), the result (7) follows.

In the second case

$$\sum_{\overline{n}} f_\phi(\overline{n}) w(\overline{n}) = k$$

for some constant k, with $0 < k < 1$ and (8) follows analogously.

The converse is of course immediate from our earlier observations. ■

We remark that by utilizing $w_{\vec{s}}$ with each of the s_i are algebraically independent we obtain probability functions which satisfy Ax and whose only independencies are those which all probability functions satisfying Ax must satisfy.

Stochastic Independence and JSP

We now return again to considering those non-trivial identities

$$c_\lambda(\theta(\vec{a}) \wedge \phi(\vec{a})) = c_\lambda(\theta(\vec{a})) \cdot c_\lambda(\phi(\vec{a}))$$

which hold for all $0 < \lambda < \infty$, where now 'non-trivial' means not holding for all probability functions satisfying Ax. It turns out that there are many such independencies but first we prove a negative result, recalling that m is the number of constants mentioned in $\theta(\vec{a})$ and $\phi(\vec{a})$:

Theorem 4. *For $m \leq 3$, equation (4) holds for all probability functions c_λ in Carnap's Continuum if and only if it holds for all probability functions w satisfying Ax.*

Proof It is enough to prove the result for $m = 3$. For (4) to hold for c_λ with $\theta(\vec{a}), \phi(\vec{a})$ non-contradictory requires that for $f_1 = f_{\theta\wedge\phi}(\{1,1,1\})$, $f_2 = f_{\theta\wedge\phi}(\{2,1\})$, $f_3 = f_{\theta\wedge\phi}(\{3\})$, $h_1 = f_\theta(\{1,1,1\})$, $g_1 = f_\phi(\{1,1,1\})$, $t_1 = f_\top(\{1,1,1\})$ etc.,

$$(f_1\mu^2 + f_2\mu(\mu+1) + f_3(\mu+1)(\mu+2))(t_1\mu^2 + t_2\mu(\mu+1) + t_3(\mu+1)(\mu+2)) = (h_1\mu^2 + h_2\mu(\mu+1) + h_3(\mu+1)(\mu+2))(g_1\mu^2 + g_2\mu(\mu+1) + g_3(\mu+1)(\mu+2)) \tag{10}$$

where $\mu = \lambda/2^q$ and none of these polynomials is identically zero. Let $f = f_1 + f_2 + f_3$ etc.. Clearly if (10) is to hold for all μ then $(f/h) = (g/t)$.

Factorizing the polynomial factors in (10) gives, say,

$$(f(\mu+\zeta_1)(\mu+\zeta_2))(t(\mu+\delta_1)(\mu+\delta_2)) =$$

$$(h(\mu+\gamma_1)(\mu+\gamma_2))(g(\mu+\beta_1)(\mu+\beta_2)). \qquad (11)$$

There are now various possibilities:

(a) $\{\zeta_1,\zeta_2\}=\{\gamma_1,\gamma_2\}$. In this case a similar phenomenon must hold for the β,δ and, since the polynomials μ^2, $\mu(\mu+1)$, $(\mu+1)(\mu+2)$ are linearly independent, this gives

$$f_{\theta\wedge\phi}(\overline{n})=(f/h)f_\theta(\overline{n}), \quad f_\phi(\overline{n})=(f/h)f_\top(\overline{n}),$$

and so this independency holds for all probability functions satisfying Ax.

(b) $\{\zeta_1,\zeta_2\}=\{\beta_1,\beta_2\}$. This case follows as in the previous case.

(c) Not cases (a) or (b). Notice that in this case the ζ_1,ζ_2,β_1 etc. must all be real since otherwise ζ_1,ζ_2 must be conjugates, etc. and one of the previous cases must have held. But since $t_1=2^q(2^q-1)(2^q-2)$, $t_2=3\cdot 2^q(2^q-1)$, $t_3=2^q$,

$$t(\mu+\delta_1)(\mu+\delta_2)=t\mu^2+(t_2+3t_3)\mu+2t_3=2^q((2^q\mu)^2+(2^q\mu)+2)$$

which has complex roots, so this case cannot occur and the required result follows.

■

However for $q\geq 2$ the situation changes once $m>3$.

Proposition 5. *For $m>3$ and $q\geq 2$ there are identities of the form (4) which hold for all probability functions c_λ in Carnap's Continuum but fail for some probability function w satisfying Ax.*

Proof Take $q=2$ and the usual atoms $\alpha_1(x)=P_1(x)\wedge P_2(x)$, $\alpha_2(x)=P_1(x)\wedge\neg P_2(x)$, $\alpha_3(x)=\neg P_1(x)\wedge P_2(x)$, $\alpha_4(x)=\neg P_1(x)\wedge\neg P_2(x)$. In this case one can check that

$$2c_\lambda(\alpha_1^2\alpha_2^2)=c_\lambda(\alpha_1^2\alpha_2\alpha_3)+c_\lambda(\alpha_1^3\alpha_2), \qquad (12)$$

where $\alpha_1^2\alpha_2^2$ is short for $\alpha_1(a_1)\wedge\alpha_1(a_2)\wedge\alpha_2(a_3)\wedge\alpha_2(a_4)$ etc..

Writing $\vec{a}$ for a_1,a_2,a_3,a_4 let

$$\varphi(\vec{a})=(P_2(a_2)\wedge\neg(\Theta_1(\vec{a})\vee\Theta_2(\vec{a})))\vee\Theta_3(\vec{a})\vee\Theta_4(\vec{a})$$

where $\Theta_1(\vec{a}), \Theta_2(\vec{a}), \Theta_3(\vec{a}), \Theta_4(\vec{a})$ are respectively the state descriptions

$$\alpha_1(a_1) \wedge \alpha_1(a_2) \wedge \alpha_1(a_3) \wedge \alpha_4(a_4),$$
$$\alpha_1(a_1) \wedge \alpha_1(a_2) \wedge \alpha_2(a_3) \wedge \alpha_4(a_4),$$
$$\alpha_1(a_1) \wedge \alpha_2(a_2) \wedge \alpha_2(a_3) \wedge \alpha_1(a_4),$$
$$\alpha_1(a_1) \wedge \alpha_2(a_2) \wedge \alpha_1(a_3) \wedge \alpha_2(a_4).$$

In this case $c_\lambda(P_1(a_1)) = 1/2 = c_\lambda(P_1(a_2))$ and by counting contributing state descriptions for $\vec{a}$ of a particular spectrum $\overline{n}$ we see that when $\overline{n} = \{2,2\}$, $\{2,1,1\}$, $\{3,1\}$, $\{4\}$, $\{1,1,1,1\}$, $P_1(a_1)$ has 18, 72, 24, 2, 12 such respectively, as does $P_2(a_2)$. If we were to take $\varphi(\vec{a}) = P_2(a_2)$ we would obtain corresponding figures of 9, 36, 12, 1, 6 for $P_1(a_1) \wedge \varphi(\vec{a})$. However if we just remove two state descriptions from $P_1(a_2)$ and add two extra ones as in the $\varphi(\vec{a})$ defined above the corresponding figures for $\varphi(\vec{a})$ and $P_1(a_1) \wedge \varphi(\vec{a})$ come out to be 20, 71, 23, 12, 2 and 11, 35, 11, 6, 1.

Using (12) it now follows that

$$c_\lambda(P_1(a_1) \wedge \varphi(\vec{a})) = c_\lambda(P_1(a_1)) \cdot c_\lambda(\varphi(\vec{a})).$$

However we can certainly find probability functions w satisfying Ax for which

$$w(P_1(a_1) \wedge \varphi(\vec{a})) \neq w(P_1(a_1)) \cdot w(\varphi(\vec{a})).$$

For example let w^δ be as in the Nix-Paris Continuum, see [13], and $0 < \delta < 1$. In this case for $\nu = (1+3\delta)/(1-\delta)$ and $C = 4^{-5}(1-\delta)^4$,

$$w^\delta(\{2,2\}) = 2C(\nu^2+1),\ \ w^\delta(\{2,1,1\}) = C(\nu+1)^2,\ \ w^\delta(\{3,1\}) = C(\nu^3+\nu+2)$$
$$w^\delta(\{4\}) = C(\nu^3+3),\ \ w^\delta(\{1,1,1,1\}) = 4C\nu,$$

and

$$w^\delta(\varphi(\vec{a}) \wedge P_1(a_1)) = C(11(2\nu^2+2)+35(\nu+1)^2+11(\nu^3+\nu+2)+6(4\nu)+(\nu^3+3))$$

which is not in general the same as

$$w^\delta(\varphi(\vec{a})) \cdot w^\delta(P_1(a_1)) = 2^{-1}C(18(2\nu^2+2)+72(\nu+1)^2+24(\nu^3+\nu+2)+12(4\nu)+2(\nu^3+3)). \blacksquare$$

We shall delay further discussion of this example until the next section. Right now we will consider the case when $q = 1$, that is when our language only has a single predicate, P say. In this case we can show Theorem 5 also for $m = 4$. When $m = 5$ we again return to the situation of Proposition 5, though the same method will not adapt. That is, when we only have a single predicate we cannot utilize a non-trivial identity of the form

$$c_\lambda(\psi(\vec{a})) = kc_\lambda(\eta(\vec{a}))$$

for some constant k to construct $\theta(\vec{a}), \phi(\vec{a})$ such that for all $0 < \lambda < \infty$

$$c_\lambda(\theta(\vec{a}) \wedge \phi(\vec{a})) = c_\lambda(\theta(\vec{a})) \cdot c_\lambda(\phi(\vec{a})),$$

whilst this identity fails for some probability function w satisfying Ax. In more detail:

Theorem 6. *Suppose that L has only a single predicate (i.e. $q = 1$) and*

$$c_\lambda(\psi(\vec{a})) = kc_\lambda(\eta(\vec{a})) \tag{13}$$

for all $0 < \lambda < \infty$. Then $f_\psi(\overline{n}) = kf_\eta(\overline{n})$ for all $\overline{n}$ and hence the identity (13) holds for all probability functions satisfying Ax.

Proof Let $x = \lambda/2$ and set

$$g_n(x) = \prod_{j=0}^{n-1}(x + j),$$

so $g_0(x) = 1$. Notice that

$$c_\lambda(\Theta(a_1, \dots, a_m)) = \frac{g_n(x)g_{m-n}(x)}{\prod_{j=0}^{m-1}(j + 2x)} = \frac{g_n(x)g_{m-n}(x)}{g_m(2x)}$$

for $\Theta(\vec{a})$ a state description having spectrum $\{n, m - n\}$.

For $0 \leq n \leq m$,

$$g_{[m/2]}(x) \mid g_n(x)g_{m-n}(x)$$

where as usual $[m/2]$ is the integer part of $m/2$. Let

$$q_n(x) = \frac{g_n(x)g_{m-n}(x)}{g_{[m/2]}(x)}.$$

for $n \leq m/2$. By considering the values of $q_0(x), q_1(x) \ldots, q_{[m/2]}$ at $0, -1, -2, \ldots,$ $-[m/2]+1$ it follows that these $q_n(x)$ are linearly independent.

Since for $q = 1$ the maximum length of a spectrum for this language is 2,

$$c_\lambda(\psi(\vec{a})) = g_{[m/2]}(x) \sum_{n \leq [m/2]} f_\psi(\{n, m-n\}) q_n(x),$$

$$c_\lambda(\eta(\vec{a})) = g_{[m/2]}(x) \sum_{n \leq [m/2]} f_\eta(\{n, m-n\}) q_n(x).$$

From (13) we must have

$$g_{[m/2]}(x) \sum_{n \leq [m/2]} f_\psi(\{n, m-n\}) q_n(x) = k g_{[m/2]}(x) \sum_{n \leq [m/2]} f_\eta(\{n, m-n\}) q_n(x)$$

so using the above linear independencies we must have

$$f_\psi(\{n, m-n\}) = k f_\eta(\{n, m-n\})$$

for $n \leq [m/2]$ and the result follows. ■

An alternative approach to that given in the proof of Proposition 5 does however give a non-trivial independency for $m = 5$ and $q = 1$ as we now show.

Proposition 7. *Suppose that L has only a single predicate (i.e. $q = 1$). Then there are $\theta(a_1, \ldots, a_5)$, $\phi(a_1, \ldots, a_5)$ such that for all $0 < \lambda < \infty$*

$$c_\lambda(\theta(\vec{a}) \wedge \phi(\vec{a})) = c_\lambda(\theta(\vec{a})) \cdot c_\lambda(\phi(\vec{a}))$$

but this fails for some probability function w satisfying Ax.

Proof Using the notation of the proof of the previous theorem, for $m = 5$

$$g_{[m/2]}(x) = x(x+1)(x+2),$$

$$\begin{array}{ll} g_0(x)g_5(x) = x(x+1)(x+2)(x+3)(x+4), & q_0(x) = (x+3)(x+4), \\ g_1(x)g_4(x) = x^2(x+1)(x+2)(x+3), & q_1(x) = x(x+3), \\ g_2(x)g_3(x) = x^2(x+1)^2(x+2), & q_2(x) = x(x+1), \end{array}$$

and

$$\begin{aligned}
2q_0(x) + 10q_1(x) + 20q_2(x) &= 8(2x+1)(2x+3), \\
q_0(x) + 4q_1(x) + 15q_2(x) &= 2(5x+6)(2x+1), \\
4q_1(x) + 12q_2(x) &= 8x(2x+3), \\
q_1(x) + 9q_2(x) &= 2x(5x+6).
\end{aligned}$$

Omitting mention of $\vec{a} = \langle a_1, a_2, a_3, a_4, a_5 \rangle$ let Θ_1, Θ_2 be the two state descriptions with spectrum $\{5\}$, let $\Phi_1, \ldots, \Phi_{10}$ be the 10 state descriptions with spectrum $\{1, 4\}$, let $\Psi_1, \ldots, \Psi_{20}$ be the 20 state descriptions with spectrum $\{2, 3\}$ and let

$$\begin{aligned}
\phi(a_1, \ldots, a_5) &= \Theta_1 \vee \bigvee_{i=1}^{4} \Phi_i \vee \bigvee_{i=1}^{15} \Psi_i, \\
\theta(a_1, \ldots, a_5) &= \bigvee_{i=4}^{7} \Phi_i \vee \bigvee_{i=7}^{18} \Psi_i, \\
\psi(a_1, \ldots, a_5) &= \Phi_4 \vee \bigvee_{i=7}^{15} \Psi_i.
\end{aligned}$$

Then $\phi(\vec{a}) \wedge \theta(\vec{a}) \equiv \psi(\vec{a})$ and

$$c_\lambda(\psi(\vec{a}) \,|\, \phi(\vec{a})) = \frac{q_1(x) + 9q_2(x)}{q_0(x) + 4q_1(x) + 15q_2(x)} = \frac{2x(5x+6)}{2(5x+6)(2x+1)},$$

$$c_\lambda(\theta(\vec{a}) \,|\, \top) = \frac{4q_1(x) + 12q_2(x)}{2q_0(x) + 10q_1(x) + 20q_2(x)} = \frac{8x(2x+3)}{8(2x+1)(2x+3)},$$

so

$$c_\lambda(\theta(\vec{a}) \wedge \phi(\vec{a})) = c_\lambda(\theta(\vec{a})) \cdot c_\lambda(\phi(\vec{a})).$$

However by Theorem 1 this identity is clearly not trivial in the sense of being satisfied by *all* probability functions w satisfying Ax. ■

Discussion

The previous sections have shown that there are in fact many 'mysterious' identities and independencies which hold for all the c_λ in Carnap's Continuum, since, when suitably phrased the results for $0 < \lambda < \infty$ can be extended to include also $\lambda = 0, \infty$ by a continuity argument. These are 'mysterious' in the sense that they do not hold for *all* probability functions satisfying Ax, if they did then by Theorems 2, 3, they would be easily explained and comprehended. Instead their derivation must require the stronger assumption of JSP rather than just Atom Exchangeability. However whilst the content of Johnson's Sufficientness Principle appears easy to grasp, and for the sake of argument accept, this seems not at all to be the case for these mysterious consequences. For example from the identity (12) and Atom Exchangeability we can obtain that

$$\begin{aligned} &c_\lambda((P_2(a_1) \leftrightarrow \neg P_2(a_2)) \wedge P_2(a_3) \wedge \neg P_2(a_4) \wedge \bigwedge_{i=1}^{4} P_1(a_i)) \\ &\qquad = c_\lambda(P_2(a_1) \wedge P_2(a_2) \wedge \neg P_1(a_4) \wedge \neg P_2(a_4) \wedge \bigwedge_{i=1}^{3} P_1(a_i)) \qquad (14) \end{aligned}$$

whilst from Proposition 5 we have that

$$c_\lambda(P_1(a_1) \,|\, \varphi(a_1, a_2, a_3, a_4)) = 1/2 = c_\lambda(P_1(a_1)),$$

where $\varphi(\vec{a})$ is (the somewhat incomprehensible)

$$\begin{aligned} &[P_2(a_2) \wedge \neg(P_1(a_3) \wedge \neg P_1(a_3) \wedge \neg P_2(a_4) \wedge \bigwedge_{i=1}^{2}(P_1(a_i) \wedge P_2(a_i))] \\ &\qquad \vee [P_2(a_1) \wedge \neg P_2(a_2) \wedge (P_2(a_3) \leftrightarrow \neg P_2(a_4)) \wedge \bigwedge_{i=1}^{4}(P_1(a_i)]. \end{aligned}$$

The immediate conclusion this leads us to then seems to be that there is much more that is hidden and mysterious in Johnson's Sufficientness Principle, and in turn Carnap's Continuum, than we might have expected. Whether or not one can give an enlightening explanation which will dispel the fog remains to be seen.

References

[1] Carnap, R., *Logical Foundations of Probability*, University of Chicago Press, Chicago, Routledge & Kegan Paul Ltd., 1950.

[2] Carnap, R., *The Continuum of Inductive Methods,* University of Chicago Press, 1952.

[3] Carnap, R., *Studies in Inductive Logic and Probability,* Volume I, eds. R.Carnap & R.C.Jefrey University of California Press, Berkeley and Los Angeles, 1971.

[4] Carnap, R., A Basic System of Inductive Logic, in *Studies in Inductive Logic and Probability*, Volume II, ed. R.C.Jeffrey, University of California Press, 1980, pp7-155.

[5] Festa, R., Analogy and Exchangeability in Predictive Inferences, *Erkenntnis,* 1996, **45**:229-252.

[6] Goodman, N., A Query on Confirmation, *Journal of Philosophy*, 1946, **43**:383-385.

[7] Goodman, N., On Infirmities in Confirmation-Theory, *Philosophy and Phenomenology Research*, 1947, **8**:149-151.

[8] Hill, M.J., Paris, J.B. & Wilmers, G.M., Some Observations on Induction in Predicate Probabilistic Reasoning, *Journal of Philosophical Logic*, 2002, **31**:43-75.

[9] Johnson, W.E., Probability: The Deductive and Inductive Problems, *Mind*, **41**:409-423, 1932.

[10] Kemeny, J.G., Carnap's Theory of Probability and Induction, in *The Philosophy of Rudolf Carnap*, ed. P.A.Schilpp, La Salle, Illinois, Open Court, 1964, 711-738.

[11] Maher, P., Probabilities for Multiple Properties: The Models of Hesse, Carnap and Kemeny, *Erkenntnis*, 2001, **55**:183-216.

[12] di Maio, M.C., Predictive Probability and Analogy by Similarity, *Erkenntnis*, 1995, **34**(3):369-394.

[13] Nix, C.J. & Paris, J.B., A Continuum of Inductive Methods arising from a Generalized Principle of Instantial Relevance, *Journal of Philosophical Logic,* 2006, **35**(1):83-115.

[14] Paris, J.B., *The Uncertain Reasoner's Companion,* Cambridge University Press, 1994.

[15] Romeijn, J.W., Analogical Predictions for Explicit Simmilarity, *Erkenntnis,* 2006, **64**(2):253-280.

[16] Skyrms, B., Analogy by Similarity in Hyper-Carnapian Inductive Logic, in *Philosophical Problems of the Internal and External Worlds,* eds. J.Earman. A.I.Janis, G.Massey & N.Rescher, University of Pittsburgh Press, 1993, pp273-282.