

*An examination of the SEP Candidate Analogical
Inference Rule within Pure Inductive Logic*

Howarth, E. and Paris, J.B. and Vencovska, A.

2015

MIMS EPrint: **2015.74**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: <http://eprints.maths.manchester.ac.uk/>

And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

An examination of the SEP Candidate Analogical Inference Rule within Pure Inductive Logic*

E.Howarth[†], J.B.Paris and A.Vencovská[‡]

School of Mathematics
The University of Manchester
Manchester M13 9PL

lizhowarth@outlook.com

jeff.paris@manchester.ac.uk

alena.vencovska@manchester.ac.uk

August 6, 2015

Abstract

Within the framework of (Unary) Pure Inductive Logic we investigate four possible formulations of a probabilistic principle of analogy based on a template considered by Paul Bartha in the Stanford Encyclopedia of Philosophy [1] and give some characterizations of the probability functions which satisfy them. In addition we investigate an alternative interpretation of analogical support, also considered by Bartha, based not on the enhancement of probability but on the creation of possibility.

Key words: Analogy, Inductive Logic, Logical Probability, Rationality, Uncertain Reasoning.

*To appear in the Journal of Applied Logic.

[†]Supported by a UK Engineering and Physical Sciences Research Council Studentship.

[‡]Supported by a UK Engineering and Physical Sciences Research Council Research Grant.

Introduction

Paraphrasing his article in the Stanford Encyclopedia of Philosophy, SEP [1], Paul Bartha considers the following characterization of an individual analogical argument:

SOURCE (S)	TARGET (T)	
P	P^*	[positive analogy]
A	$\neg A^*$	[negative analogy]
$\neg B$	B^*	
Q	Q^*	(plausibly)

with an *analogical argument* being: *It is plausible that Q^* holds in the target domain because of certain known (or accepted) similarities with the source domain, despite certain known (or accepted) differences.*

In turn he examines a corresponding candidate analogical inference rule, CAIR for short:

Suppose S and T are the source and target domains. Suppose $P_1 \dots, P_n$ (with $n \geq 1$) represents the positive analogy, A_1, \dots, A_r and $\neg B_1, \dots, \neg B_s$ represent the (possibly vacuous) negative analogy, and Q represents the hypothetical analogy. In the absence of reasons for thinking otherwise, infer that Q^ holds in the target domain with degree of support $p > 0$, where p is an increasing function of n and a decreasing function of r and s .*

The primary intention of this paper is to formulate, as principles, mathematically more precise versions of CAIR within the framework of (*Unary*) *Pure Inductive Logic*, PIL for short, where ‘degree of support’ is identified with (subjective) probability, and to determine which probability functions satisfy these versions in the presence of certain other, widely accepted, symmetry requirements. We should point out that this differs somewhat from the ‘Applied Inductive Logic’ framework in which Bartha considers and dismisses CAIR, even as a non-starter. Following that we shall suggest and investigate within this formal framework an alternative interpretation of analogical support based on the creation of possibility, also considered by Bartha as what he terms ‘the modal conception’, see [1, Section 2.3].

This (*Unary*) *Pure Inductive Logic* framework is explained in, for example, [20], [21]. In short we work in a first order predicate language L_q with finitely many predicate, i.e. unary relation, symbols R_1, R_2, \dots, R_q , countably many constants

a_1, a_2, a_3, \dots , which are intended to name all the elements of the universe, and no function symbols nor equality. Let SL_q and $QFSL_q$ denote, respectively, the sentences and quantifier free sentences of this language.

A *probability function on L_q* is a function $w : SL_q \rightarrow [0, 1]$ such that for all $\theta, \phi, \exists x \psi(x) \in SL_q$:

- (P1) If $\models \theta$ then $w(\theta) = 1$
- (P2) If $\models \neg(\theta \wedge \phi)$ then $w(\theta \vee \phi) = w(\theta) + w(\phi)$
- (P3) $w(\exists x \psi(x)) = \lim_{m \rightarrow \infty} w(\bigvee_{i=1}^m \psi(a_i))$,

this last condition reflecting the intention that the constants a_i exhaust the universe.

The primary goal of PIL, as we would present it, is to investigate which such probability functions are *logical* or *rational* in the sense of corresponding to the subjective probabilities assigned by a rational agent *in the absence of any further knowledge or intended interpretation of the constant and predicate symbols*.

Whilst we have no precise definition of what we mean by ‘logical’ or ‘rational’ here, indeed such a clarification is essentially equivalent to the above goal, we do at least seem to have some intuitions about what constitutes being rational, or perhaps more usually what constitutes being irrational. For example in the circumstances of such zero knowledge it would seem to be irrational to treat any one constant differently from any other. Precisely then a rational probability function w should satisfy:

The Principle of Constant Exchangeability, Ex.

A *probability function w on SL_q satisfies Constant Exchangeability* if, for any permutation σ of $1, 2, \dots$ and $\theta(a_1, \dots, a_n) \in SL_q$,

$$w(\theta(a_{\sigma(1)}, \dots, a_{\sigma(n)})) = w(\theta(a_1, \dots, a_n)). \quad (1)$$

This principle is so widely assumed in this context that we shall henceforth take it, without further mention, *that all probability functions we discuss satisfy it*.

Similarly there would seem to be no rational reason to give two predicates different properties, nor even between a predicate and its negation. This leads to imposing two further requirements on a rational probability function to satisfy:

The Principle of Predicate Exchangeability, Px

If R_i, R_j are predicate symbols of L_q then for $\theta \in SL$, $w(\theta) = w(\theta')$ where θ' is the result of transposing R_i, R_j throughout θ .

The Strong Negation Principle, SN

For $\theta \in SL_q$, $w(\theta) = w(\theta')$ where θ' is the result of replacing each occurrence of the predicate symbol R in θ by $\neg R$.

In what follows we shall restrict our attention to probability functions w satisfying these three principles Ex, Px, SN.

There is a further principle which we will need subsequently and whose rationality may be argued for as follows. Suppose that on the basis of some considerations we have made the probability function w^{L_q} our rational choice of probability function on L_q and the probability function w^{L_r} our rational choice of probability function on L_r where $r \geq q$. Then since $SL_q \subseteq SL_r$ it would seem perverse if w^{L_r} did not agree with w^{L_q} on SL_q , since it would mean that what we considered a rationally justified value for the probability of $\theta \in SL_q$ depended on the presence or absence of relation symbols in the language which were not even mentioned in θ .

Given our earlier argument for the rationality of Px + SN this leads to the following ‘meta-rationality’ principle which it is desirable for a probability function w^{L_q} on L_q to satisfy, though unlike Ex, Px and SN we will not actually assume it as the default:

Unary Language Invariance with Strong Negation¹, ULi + SN

A probability function w on L_q satisfies Unary Language Invariance with SN if there is a family of probability functions w^{L_r} , one on each language L_r where $r \in \mathbb{N}^+ = \{1, 2, 3, \dots\}$, satisfying Ex + Px + SN and such that $w = w^{L_q}$ and whenever $r \leq s$ then w^{L_s} agrees with w^{L_r} on SL_r .

Whilst the rationality of observing symmetries and language invariance as expressed by the above principles seems to us hard to question, the rationality of arguments by analogy appears much less forceful. Nevertheless in the real world we often are somewhat influenced by analogies, for clear accounts of such within mathematics see [23], [24], and there have been several attempts, starting with Rudolf Carnap, to capture facets of analogy as a rational or logical principle within the framework of Inductive Logic, see for example [3], Carnap & Stegmüller [4] and later Festa [5], Hesse [9], [10], Maher [16], [17], di Maio [18], Romeijn [25], Skyrms

¹ULi alone, see [21], only requires each of the probability functions w^{L_r} to satisfy Ex + Px.

[26].

In each of the next four sections we will add to this list of ‘Principles of Analogy’ by proposing interpretations, or variants, of CAIR within the framework of PIL and, in Theorems 1, 2, 3, 4, investigating the probability functions that satisfy them (in the presence of our standing assumptions Ex, Px, SN). Subsequently we will broaden the remit by proposing a principle of analogy (Dolly’s Principle) based on the idea of analogy as a source of possibility (the modal conception as Bartha terms it) rather than increase of probability. Since mathematical results in these sections contain on occasions technicalities that some readers may wish to simply accept we shall now spend a little time introducing the terms that appear in their statements in order that they become directly accessible. A general overview of these results will then be given in the final section.

An *atom* of L_q is a formula of the form

$$R_1^{\epsilon_1} \wedge R_2^{\epsilon_2} \wedge \dots \wedge R_q^{\epsilon_q}$$

where $\epsilon_1, \epsilon_2, \dots, \epsilon_q \in \{0, 1\}$ and $R_i^1 = R_i, R_i^0 = \neg R_i$. So there are 2^q atoms for L_q , which we denote $\alpha_1(x), \alpha_2(x), \dots, \alpha_{2^q}(x)$, corresponding to the 2^q different choices for the \vec{c} . Notice that because we only have unary relation symbols in the language, knowing which atom a constant satisfies tells us all there is to know about that constant.

Similarly for (distinct) constants² b_1, b_2, \dots, b_n the *state description* that holds for them, that is the sentence

$$\bigwedge_{i=1}^n \alpha_{h_i}(b_i),$$

tells us all there is to know about these constants. By a theorem of Gaifman, see [6] or [21, Theorem 7.1], a probability function is uniquely determined by its values on state descriptions.

Let \mathbb{D}_{2^q} be the set of vectors

$$\left\{ \langle x_1, x_2, \dots, x_{2^q} \rangle \in \mathbb{R}^{2^q} \mid x_i \geq 0, \sum_i x_i = 1 \right\}.$$

For $\vec{c} \in \mathbb{D}_{2^q}$ the probability function $w_{\vec{c}}$ on L_q is defined by

$$w_{\vec{c}} \left(\bigwedge_{i=1}^n \alpha_{h_i}(b_i) \right) = \prod_{i=1}^n c_{h_i}.$$

²We may on occasions use b_1, b_2, b_3 etc. for constants from $\{a_1, a_2, a_3, \dots\}$, rather than $a_{i_1}, a_{i_2}, a_{i_3}$, etc., in order to avoid multiple subscripts.

In other words $w_{\mathcal{C}}$ treats the $\alpha_{h_i}(b_i)$ as stochastically independent with individual probabilities c_{h_i} , $i = 1, \dots, n$. This probability function satisfies Ex but not Px nor SN except under special circumstances. For future reference we recall (see [21, Chapter 8]) that the $w_{\mathcal{C}}$ are characterized by satisfying the

The Constant Irrelevance Principle

If $\theta, \phi \in QFSL$ have no constant symbols in common then

$$w(\theta \wedge \phi) = w(\theta) \cdot w(\phi).$$

We remark that the principle implies that $w(\theta \wedge \phi) = w(\theta) \cdot w(\phi)$ even when $\theta, \phi \in SL$ (not necessarily quantifier free), see [21, Chapter 8].

The functions $w_{\mathcal{C}}$ are fundamental since any probability function satisfying Ex can be expressed from them as an integral over \mathbb{D}_{2^q} :

De Finetti's Representation Theorem. *Let w be a probability function on SL_q satisfying Ex. Then there is a normalized and countably additive measure μ on the Borel subsets of \mathbb{D}_{2^q} such that*

$$\begin{aligned} w\left(\bigwedge_{i=1}^n \alpha_{h_i}(b_i)\right) &= \int_{\mathbb{D}_{2^q}} \prod_{j=1}^{2^q} x_j^{m_j} d\mu(\vec{x}) \\ &= \int_{\mathbb{D}_{2^q}} w_{\vec{x}}\left(\bigwedge_{i=1}^n \alpha_{h_i}(b_i)\right) d\mu(\vec{x}), \end{aligned} \quad (1)$$

where for $j = 1, 2, \dots, 2^q$, $m_j = |\{i \mid h_i = j\}|$, the number of times that j occurs amongst h_1, h_2, \dots, h_n .

de Finetti's Theorem finds numerous important, and slick, applications in PIL; for example Humburg's proof (see [14], or [21, Chapter 11]) of a result of Gaifman, [6], that we shall need later, that Ex implies

The (Extended) Principle of Instancial Relevance

For $\theta(a_1, a_2, \dots, a_n), \phi(a_1) \in SL$,

$$w(\phi(a_{n+2}) \mid \phi(a_{n+1}) \wedge \theta(a_1, a_2, \dots, a_n)) \geq w(\phi(a_{n+2}) \mid \theta(a_1, a_2, \dots, a_n)). \quad (2)$$

A second important family of probability functions on L_q are the $c_{\lambda}^{L_q}$, $0 \leq \lambda \leq \infty$, of Carnap's *Continuum of Inductive Methods* which, for $\lambda > 0$, are specified by

$$c_{\lambda}^{L_q}(\alpha_j(b_{n+1}) \mid \bigwedge_{i=1}^n \alpha_{h_i}(a_i)) = \frac{m_j + \lambda 2^{-q}}{n + \lambda}$$

where (again) $m_j = |\{i \mid h_i = j\}|$ and for $\lambda = 0$ by

$$c_0^{L_q} \left(\bigwedge_{i=1}^n \alpha_{h_i}(b_i) \right) = \begin{cases} 2^{-q} & \text{if } h_1 = h_2 = \dots = h_n, \\ 0 & \text{otherwise.} \end{cases}$$

These $c_\lambda^{L_q}$ satisfy Ex + Px + SN and even ULi with SN, the corresponding language invariant family being obtained by fixing the λ and letting the q range over \mathbb{N}^+ .

Given that convex sums of probability functions are again probability functions and that probability functions are determined by their values on state descriptions it is easy to check that

$$c_\infty^{L_q} = w_{\langle 2^{-q}, 2^{-q}, \dots, 2^{-q} \rangle}$$

$$c_0^{L_q} = 2^{-q} (w_{\langle 1, 0, 0, \dots, 0 \rangle} + w_{\langle 0, 1, 0, \dots, 0 \rangle} + \dots + w_{\langle 0, \dots, 0, 0, 1 \rangle}).$$

To simplify the notation in what follows we shall omit the superscript L_q in $c_\lambda^{L_q}$ when q is clear from the context.

Notice that any permutation of predicates, or transposition of $R_j, \neg R_j$, generates a permutation of the atoms α_i . We shall say that a permutation σ of atoms is *licensed by* Px + SN if it can be formed as a composition of such permutations. Notice that if w satisfies Px + SN then for such a σ ,

$$w \left(\bigwedge_{i=1}^n \alpha_{h_i}(a_{k_i}) \right) = w \left(\bigwedge_{i=1}^n \sigma(\alpha_{h_i})(a_{k_i}) \right). \quad (3)$$

Furthermore, since by Ex the left (and right) hand side is the same for any choice of distinct constants we shall, to simplify the notation, sometimes omit the instantiating constants and denote it simply as

$$w \left(\bigwedge_{i=1}^n \alpha_{h_i} \right)$$

or even

$$w(\alpha_1^{m_1} \alpha_2^{m_2} \dots \alpha_{2^q}^{m_{2^q}})$$

where m_j is the number of times that j appears in h_1, h_2, \dots, h_n .

For future reference note that *Atom Exchangeability* is the assertion that (3) holds for *any* permutation σ of the set of atoms, not just those licensed by Px + SN.

The General Analogy Principle

The first question we might feel obliged to address vis-a-vis CAIR is what exactly the forms of the Q, P, A, B are and what exactly is being treated analogously in the relationship between Q and Q^* etc. – what we shall refer to as the *carrier* of the analogy. The four versions we shall consider are really centred around possible answers to these questions within the framework of Unary³ PIL. In our first attempt at a formulation the Q, P are just quantifier free sentences and it is the constants which are the carriers:

The General Analogy Principle, GAP

For $\vec{a} = \langle a_3, a_4, \dots, a_k \rangle$ and $\psi(a_1, \vec{a}), \phi(a_1, \vec{a}) \in QFSL$,

$$w(\phi(a_2, \vec{a}) \mid \psi(a_1, \vec{a}) \wedge \psi(a_2, \vec{a}) \wedge \phi(a_1, \vec{a})) \geq w(\phi(a_2, \vec{a}) \mid \phi(a_1, \vec{a})). \quad (4)$$

In this principle then $\psi(a_1, \vec{a}) \wedge \psi(a_2, \vec{a})$ provides ‘evidence’ that a_1, a_2 are similar and hence should enhance (or at least not decrease) the probability that a_2 should again be similar to a_1 in satisfying $\phi(x, \vec{a})$ given that a_1 does.

A few comments are in order here. Firstly we shall identify (4) with

$$\begin{aligned} w(\phi(a_2, \vec{a}) \wedge \psi(a_1, \vec{a}) \wedge \psi(a_2, \vec{a}) \wedge \phi(a_1, \vec{a})) \cdot w(\phi(a_1, \vec{a})) \\ \geq w(\psi(a_1, \vec{a}) \wedge \psi(a_2, \vec{a}) \wedge \phi(a_1, \vec{a})) \cdot w(\phi(a_2, \vec{a}) \wedge \phi(a_1, \vec{a})), \end{aligned}$$

a convenience which satisfactorily allows us to dispense with the problem of conditioning on sentences with probability zero.⁴

Secondly, in this formulation we have taken as vacuous the negative analogies A_1, \dots, A_r and $\neg B_1, \dots, \neg B_s$. In particular then the monotonicity element of Bartha’s representation has been reduced to a single inequality.⁵ Thirdly, notice that by Ex the choice of constants a_1, a_2, \dots, a_k is not relevant since it implies the same principle for any distinct choice of constants. Finally, within this formulation we are restricting $\phi(a_1, \vec{a}), \psi(a_1, \vec{a})$ to be quantifier free, again for reasons which will shortly become clear.

³Several of our results apply also to Polyadic Inductive Logic, see [20], [21] for further details, but for simplicity we shall limit ourselves here to the purely unary.

⁴More generally we shall identify $(a/b) \geq (c/d)$ with $ad \geq bc$.

⁵Subsequent results will somewhat vindicate this decision.

As we now show GAP fails to satisfactorily capture our (presumably viable) intuitions about analogy.

Theorem 1. *Let w be a probability function on the unary language L_q satisfying $Px + SN$.⁶ Then*

(A) *If $q = 1$ then w satisfies GAP just if $w = c_0$.*

(B) *If $q \geq 2$ then w satisfies GAP just if $w = c_0$, even dropping the additional constants \vec{a} .*

Proof. (A): That c_0 (on L_1) satisfies GAP will be shown in part (B) below.

If $w = \nu c_\infty + (1 - \nu)c_0$ with $0 < \nu \leq 1$ then one can check that for $\phi(a_1, \vec{a}), \psi(a_1, \vec{a})$ being respectively

$$\begin{aligned} & ((R_1(a_1) \wedge (R_1(a_3) \vee R_1(a_4))) \vee (\neg R_1(a_1) \wedge (R_1(a_3) \vee \neg R_1(a_4)))) \wedge R_1(a_5) \wedge \neg R_1(a_6), \\ & ((R_1(a_1) \wedge (\neg R_1(a_3) \vee R_1(a_4))) \vee (\neg R_1(a_1) \wedge (\neg R_1(a_3) \vee \neg R_1(a_4)))) \wedge R_1(a_5) \wedge \neg R_1(a_6), \end{aligned}$$

we have

$$w(\phi(a_2, \vec{a}) \mid \psi(a_1, \vec{a}) \wedge \psi(a_2, \vec{a}) \wedge \phi(a_1, \vec{a})) = 2/3 < 5/6 = w(\phi(a_2, \vec{a}) \mid \phi(a_1, \vec{a})),$$

which provides the required counter-example.

So now suppose that w is not of the form $\nu c_\infty + (1 - \nu)c_0$ for any $0 \leq \nu \leq 1$. Then the de Finetti prior of w must have a support point $\langle c, 1 - c \rangle$ with $0 < c < 1/2$. (In other words every open set containing this point has non-zero measure.)

Let $\phi(a_1) = R_1(a_1)$. Then by the Extended Principle of Instantial Relevance, (2), and SN,

$$w(\phi(a_2) \mid \phi(a_1)) \geq w(R_1(a_1)) = 1/2. \quad (5)$$

Let

$$\psi(a_3, a_4, \dots, a_k) = R_1^{[mc]}(\neg R_1)^{[m(1-c)]}$$

where as usual $[mc]$ is the integer part of mc and $k = [mc] + [m(1 - c)] + 2$. Then

$$w(\phi(a_2) \mid \phi(a_1) \wedge \psi(\vec{a})) = \frac{w(R_1^{[mc+2]}(\neg R_1)^{[m(1-c)]})}{w(R_1^{[mc+1]}(\neg R_1)^{[m(1-c)]})} \approx c$$

for large m (see for example [21, Chapter 12]). Comparing with (5) we have the required counter-example.

⁶Recall the standing assumption that all probability functions considered satisfy Ex.

(B): Here we shall show the result even without the \vec{a} being present. We first need to introduce another probability function, ϖ , on L_q . For α_i an atom of L_q let α_i^c be that atom of L_q which disagrees with α_i on every $R_j(x)$, in other words, for $j = 1, 2, \dots, q$,

$$\alpha_i(x) \models R_j(x) \iff \alpha_i^c(x) \models \neg R_j(x).$$

Now let $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_{2^q-1}$ run through all vectors in \mathbb{D}_{2^q} which have zeros at all coordinates except for two, say the i th and j th, with $\alpha_i^c = \alpha_j$, and in those places the entry is $1/2$. Set

$$\varpi = 2^{1-q} \sum_{i=1}^{2^q-1} w_{\vec{e}_i}.$$

We now show that in the case of a unary language L with $q \geq 2$ predicates and a probability function w on L satisfying Ex, Px and SN and not of the form $\lambda\varpi + (1-\lambda)c_0$ for some $0 < \lambda \leq 1$ there are $\phi(a_1), \psi(a_1)$ for which (4) fails.⁷

To this end let $G \subset \{1, 2, \dots, 2^q\}$, $|G| = 2^{q-1}$ and let $x = w(\alpha_i\alpha_i)$ and $y_{ij} = w(\alpha_i\alpha_j)$. Notice that x is independent of i since for any atoms α_i, α_j there is a permutation σ of atoms licensed by SN such that $\sigma(\alpha_i) = \alpha_j$.

Since $|G| = 2^{q-1}$ and

$$\sum_{i \in G} \sum_{i \neq j \in G} y_{ij} = \sum_{\substack{i \neq j \\ i, j \in G}} y_{ij}$$

we can find $i \in G$, say $i = 1$ (so $1 \in G$), such that

$$2^{q-1} \sum_{1 \neq j \in G} y_{1j} \leq \sum_{\substack{i \neq j \\ i, j \in G}} y_{ij}. \quad (6)$$

Let

$$\begin{aligned} \phi(a_1) &= \bigvee_{i \in G} \alpha_i(a_1), \\ \psi(a_1) &= \alpha_1(a_1) \vee \bigvee_{i \notin G} \alpha_i(a_1). \end{aligned}$$

Notice that

$$w(\alpha_1(a_1)) = 2^{-q} = x + \sum_{1 \neq j \in G} y_{1j} + \sum_{j \notin G} y_{1j}. \quad (7)$$

⁷Since probability functions satisfying Px + SN on a language L continue to satisfy these principles when restricted to smaller languages it would actually suffice here to prove this part for $q = 2$. However, overall, that does not seem to be any simpler.

Then with the above abbreviations,

$$\begin{aligned}
w(\phi(a_1)) &= 2^{q-1}2^{-q}, \\
w(\phi(a_1) \wedge \phi(a_2)) &= 2^{q-1}x + \sum_{\substack{i,j \in G \\ i \neq j}} y_{ij}, \\
w(\psi(a_1) \wedge \psi(a_2) \wedge \phi(a_1)) &= x + \sum_{j \notin G} y_{1j}, \\
w(\psi(a_1) \wedge \psi(a_2) \wedge \phi(a_1) \wedge \phi(a_2)) &= x,
\end{aligned}$$

and the inequality (4) becomes

$$\frac{x}{x + \sum_{j \notin G} y_{1j}} \geq \frac{2^{q-1}x + \sum_{\substack{i,j \in G \\ i \neq j}} y_{ij}}{2^{q-1}2^{-q}}.$$

Multiplying out gives

$$\begin{aligned}
2^{q-1}2^{-q}x &\geq \left(x + \sum_{j \notin G} y_{1j} \right) \left(2^{q-1}x + \sum_{\substack{i,j \in G \\ i \neq j}} y_{ij} \right) \\
&= 2^{q-1}x^2 + x \left(\sum_{\substack{i,j \in G \\ i \neq j}} y_{ij} + 2^{q-1} \sum_{j \notin G} y_{1j} \right) + \sum_{j \notin G} y_{1j} \cdot \sum_{\substack{i,j \in G \\ i \neq j}} y_{ij} \\
&= 2^{q-1}x^2 + x \left(\sum_{\substack{i,j \in G \\ i \neq j}} y_{ij} + 2^{q-1}(2^{-q} - x - \sum_{1 \neq j \in G} y_{1j}) \right) \\
&\quad + \sum_{j \notin G} y_{1j} \cdot \sum_{\substack{i,j \in G \\ i \neq j}} y_{ij} \quad \text{by (7)}.
\end{aligned}$$

Cancelling out terms now gives

$$0 \geq x \left(\sum_{\substack{i,j \in G \\ i \neq j}} y_{ij} - 2^{q-1} \sum_{1 \neq j \in G} y_{1j} \right) + \sum_{j \notin G} y_{1j} \cdot \sum_{\substack{i,j \in G \\ i \neq j}} y_{ij}. \quad (8)$$

By (6) this right hand side is at least 0. We now show that for a suitable initial choice of G it must be strictly positive.

Since w is not of the form $\lambda\varpi + (1 - \lambda)c_0$ for some $0 < \lambda \leq 1$ there must be $i \neq j$ such that $w(\alpha_i\alpha_j) > 0$ and $\alpha_j \neq \alpha_i^c$, say α_i, α_j differ on r predicates where $1 \leq r < q$. Let $\alpha_k(x) \models R_1(x)$. Then there is a permutation σ licensed by SN such that $\sigma(\alpha_i) = \alpha_k$, $\sigma(\alpha_j)$ differs from α_k on r predicates⁸ and by Px + SN, $w(\alpha_k\sigma(\alpha_j)) = w(\alpha_i\alpha_j) > 0$. Suppose that $\sigma(\alpha_j) \models R_1(x)$. Then because $1 \leq r < q$ we can find a permutation τ licensed by SN + Px such that $\alpha_k = \tau(\alpha_k)$ and $\tau\sigma(\alpha_j) \models \neg R_1(x)$, and of course $w(\alpha_k\tau\sigma(\alpha_j)) > 0$. Similarly if $\sigma(\alpha_j) \models \neg R_1(x)$ we can find an atom α_s differing from α_k on r predicates such that $\alpha_s \models R_1(x)$ and $w(\alpha_k\alpha_s) > 0$.

In this case then we can take G to be the set of those atoms α_i such that $\alpha_i \models R_1(x)$ and obtain the required contradiction to (8). So GAP does not hold.

Turning now to the case where $w = \lambda\varpi + (1 - \lambda)c_0$ for some $0 < \lambda \leq 1$ and $q \geq 2$ let $\alpha_i, \alpha_j, \alpha_i^c, \alpha_j^c$ be distinct atoms and take

$$\phi(a_1) = \alpha_i(a_1) \vee \alpha_j(a_1) \vee \alpha_j^c(a_1), \quad \psi(a_1) = \alpha_i(a_1) \vee \alpha_i^c(a_1).$$

Then

$$\begin{aligned} w(\phi(a_1)) &= \lambda\varpi(\phi(a_1)) + (1 - \lambda)c_0(\phi(a_1)) = 3\lambda 2^{-q} + 3(1 - \lambda)2^{-q} \\ &= 3 \cdot 2^{-q}, \end{aligned}$$

$$\begin{aligned} w(\phi(a_1) \wedge \phi(a_2)) &= \lambda\varpi(\phi(a_1) \wedge \phi(a_2)) + (1 - \lambda)c_0(\phi(a_1) \wedge \phi(a_2)) \\ &= 5\lambda 2^{-q-1} + 3(1 - \lambda)2^{-q} \\ &= (3 - \lambda/2)2^{-q}, \end{aligned}$$

$$\begin{aligned} w(\psi(a_1) \wedge \psi(a_2) \wedge \phi(a_1)) &= \lambda\varpi(\psi(a_1) \wedge \psi(a_2) \wedge \phi(a_1)) \\ &\quad + (1 - \lambda)c_0(\psi(a_1) \wedge \psi(a_2) \wedge \phi(a_1)) \\ &= \lambda 2^{-q} + (1 - \lambda)2^{-q} = 2^{-q}, \end{aligned}$$

$$\begin{aligned} w(\psi(a_1) \wedge \psi(a_2) \wedge \phi(a_1) \wedge \phi(a_2)) &= \lambda\varpi(\psi(a_1) \wedge \psi(a_2) \wedge \phi(a_1) \wedge \phi(a_2)) \\ &\quad + (1 - \lambda)c_0(\psi(a_1) \wedge \psi(a_2) \wedge \phi(a_1) \wedge \phi(a_2)) \\ &= \lambda 2^{-q-1} + (1 - \lambda)2^{-q} = (1 - \lambda/2)2^{-q}, \end{aligned}$$

and the inequality (4) becomes

$$3(1 - \lambda/2) \geq 3 - \lambda/2,$$

⁸The permutations licensed by Px + SN are precisely those that preserve Hamming distance, see [11, Theorem 2].

which fails since $\lambda > 0$, and gives the required counter-example.

To complete case (B) we now show that GAP holds for c_0 on L_q for any q . Indeed we shall show that it holds even in the case where ϕ, ψ are sentences rather than just quantifier free. To this end notice (or see for example [8] where a similar result is derived) that for the *unary* language L_q , any sentence mentioning constants a_1, \dots, a_n is logically equivalent to a sentence in the form

$$\bigvee_{i=1}^s \left(\bigwedge_{j=1}^n \alpha_{k_{i,j}}(a_j) \wedge \bigwedge_{m=1}^{2^q} (\exists x \alpha_m(x))^{\epsilon_{i,m}} \right) \quad (9)$$

where the $\epsilon_{i,m} \in \{0, 1\}$ and as usual ψ^ϵ is ψ if $\epsilon = 1$ and $\neg\psi$ if $\epsilon = 0$.

Noticing that

$$c_0(\alpha_i \alpha_j) = \begin{cases} 2^{-q} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

we can see that for each of the disjuncts in (9) either

$$c_0(\alpha_{k_{i,1}}(a_1) \leftrightarrow (\bigwedge_{j=1}^n \alpha_{k_{i,j}}(a_j) \wedge \bigwedge_{m=1}^{2^q} (\exists x \alpha_m(x))^{\epsilon_{i,m}})) = 1$$

or

$$c_0(\perp \leftrightarrow (\bigwedge_{j=1}^n \alpha_{k_{i,j}}(a_j) \wedge \bigwedge_{m=1}^{2^q} (\exists x \alpha_m(x))^{\epsilon_{i,m}})) = 1.$$

From this it follows that there are $S, T \subseteq \{1, 2, \dots, 2^q\}$ such that

$$c_0(\phi(a_1, \vec{a}) \leftrightarrow \bigvee_{j \in S} \alpha_j(a_1)) = 1,$$

$$c_0(\psi(a_1, \vec{a}) \leftrightarrow \bigvee_{j \in T} \alpha_j(a_1)) = 1.$$

Hence

$$c_0(\phi(a_1, \vec{a})) = c_0(\phi(a_1, \vec{a}) \wedge \phi(a_2, \vec{a})) = 2^{-q}|S|$$

$$\begin{aligned} c_0(\psi(a_1, \vec{a}) \wedge \psi(a_2, \vec{a}) \wedge \phi(a_1, \vec{a})) &= c_0(\psi(a_1, \vec{a}) \wedge \psi(a_2, \vec{a}) \wedge \phi(a_1, \vec{a}) \wedge \phi(a_2, \vec{a})) \\ &= 2^{-q}|S \cap T| \end{aligned}$$

and the inequality (4) follows immediately. \square

It is perhaps worth remarking here that in the case of $q = 1$ the extra constants a_3, a_4, \dots, a_m employed in forming the counter-example to GAP cannot be dispensed with. In fact when $q = 1$ and the extra constants are absent GAP trivially holds for any w (and even continues to hold when ϕ, ψ may contain quantifiers provided w not of the form $\lambda c_0 + (1 - \lambda)w'$, with $0 < \lambda < 1$ and $w' \neq c_0$, see [19]).

The Equivalence Analogy Principle, EAP

GAP bears a superficial resemblance to the following analogy principle suggested by [2, Section 3] (though as far as we can tell Peirce viewed this as an *abduction* rather than an analogy principle):

The Equivalence Analogy Principle, EAP

For $\vec{a} = \langle a_3, a_4, \dots, a_k \rangle$ and $\psi(a_1, \vec{a}), \phi(a_1, \vec{a}) \in QFSL$,

$$w(\phi(a_1, \vec{a}) \leftrightarrow \phi(a_2, \vec{a}) \mid \psi(a_1, \vec{a}) \wedge \psi(a_2, \vec{a})) \geq w(\phi(a_1, \vec{a}) \leftrightarrow \phi(a_2, \vec{a})). \quad (10)$$

As with GAP it is the constants which are the carriers of the analogy and presumably, judging from their similarity, EAP's justification is based on the same intuitions, so one might have expected that they would again have the same solutions, or more aptly lack of solutions. This is indeed almost the case provided we allow the additional constants \vec{a} in EAP. However dropping these constants gives a somewhat different, but still very restricted, set of solutions, in contrast to any supposedly similar intuitions.

To start with we shall give a characterization of the probability functions satisfying EAP, in the presence of our customary additional default assumptions of Px + SN.

Theorem 2. *Let w be a probability function on the unary language L_q satisfying Px + SN. Then w satisfies EAP just if $w = c_0^{L_q}$.*

Proof. Suppose that $w \neq c_0^{L_q}$ satisfies EAP. Without loss of generality we may assume that $w(\alpha_1 \alpha_2) > 0$. Using de Finetti's representation theorem (and the fact that for $\vec{c} \in \mathbb{D}_{2^q}$ we have $w_{\vec{c}}(\alpha_1(\neg \alpha_1)) \leq \frac{1}{4}$), we can see that

$$w(\alpha_1) > w(\alpha_1^2(\neg \alpha_1)) \geq w(\alpha_1^2 \alpha_2^2) > 0. \quad (11)$$

Let

$$\psi(a_3) = \alpha_1(a_3), \quad \phi(a_1, a_3) = \alpha_1(a_1) \wedge \psi(a_3).$$

Notice that ψ does not actually depend on a_1 so $\psi(a_1, \vec{a}) = \psi(a_2, \vec{a}) = \psi(a_3)$.

Then

$$\begin{aligned}
w((\phi(a_1, a_3) \leftrightarrow \phi(a_2, a_3)) \wedge \psi(a_1, \vec{a}) \wedge \psi(a_2, \vec{a})) \\
&= w((\alpha_1(a_1) \leftrightarrow \alpha_1(a_2)) \wedge \alpha_1(a_3)) \\
&= w(\alpha_1(a_3)) - w(\neg(\alpha_1(a_1) \leftrightarrow \alpha_1(a_2)) \wedge \alpha_1(a_3)) \\
&= w(\alpha_1(a_3)) - 2w(\alpha_1(a_1) \wedge \neg\alpha_1(a_2) \wedge \alpha_1(a_3)).
\end{aligned}$$

Similarly

$$w(\phi(a_1, a_3) \leftrightarrow \phi(a_2, a_3)) = 1 - 2w(\alpha_1(a_1) \wedge \neg\alpha_1(a_2) \wedge \alpha_1(a_3)).$$

Thus for EAP to hold we must have

$$w(\alpha_1(a_3))(1 - 2w(\alpha_1(a_1) \wedge \neg\alpha_1(a_2) \wedge \alpha_1(a_3))) \leq w(\alpha_1(a_3)) - 2w(\alpha_1(a_1) \wedge \neg\alpha_1(a_2) \wedge \alpha_1(a_3)),$$

equivalently

$$w(\alpha_1(a_1) \wedge \neg\alpha_1(a_2) \wedge \alpha_1(a_3)) \leq w(\alpha_1(a_3)) \cdot w(\alpha_1(a_1) \wedge \neg\alpha_1(a_2) \wedge \alpha_1(a_3)),$$

which by (11) fails and gives the required contradiction. Notice that we only needed the single extra constant a_3 to derive this contradiction.

To complete the proof we need to show that $w = c_0^{L^q}$ satisfies EAP.

To this end let $\phi(a_1, \vec{a}) \in QFSL$ (where as usual $\vec{a} = \langle a_3, a_4, \dots, a_k \rangle$) and $1 \leq j \leq 2^q$. Since for this probability function w

$$w(\alpha_j(a_1) \rightarrow \alpha_j(a_m)) = 1,$$

for any m , if

$$\alpha_j(x) \wedge \bigwedge_{i=3}^k \alpha_j(a_i) \models \phi(x, \vec{a})$$

then

$$w(\alpha_j(a_1) \wedge \phi(a_1, \vec{a}) \wedge \phi(a_2, \vec{a})) = 2^{-q}, \quad w(\alpha_j(a_1) \wedge \neg\phi(a_1, \vec{a}) \wedge \neg\phi(a_2, \vec{a})) = 0,$$

while if

$$\alpha_j(x) \wedge \bigwedge_{i=3}^k \alpha_j(a_i) \models \neg\phi(x, \vec{a})$$

then

$$w(\alpha_j(a_1) \wedge \phi(a_1, \vec{a}) \wedge \phi(a_2, \vec{a})) = 0, \quad w(\alpha_j(a_1) \wedge \neg\phi(a_1, \vec{a}) \wedge \neg\phi(a_2, \vec{a})) = 2^{-q}.$$

In either case,

$$\begin{aligned} w(\alpha_j(a_1) \wedge (\phi(a_1, \vec{a}) \leftrightarrow \phi(a_2, \vec{a}))) \\ = w(\alpha_j(a_1) \wedge \phi(a_1, \vec{a}) \wedge \phi(a_2, \vec{a})) + w(\alpha_j(a_1) \wedge \neg\phi(a_1, \vec{a}) \wedge \neg\phi(a_2, \vec{a})) = 2^{-q} \end{aligned}$$

and summing over $1 \leq j \leq 2^q$ gives

$$w(\phi(a_1, \vec{a}) \leftrightarrow \phi(a_2, \vec{a})) = w(\phi(a_1, \vec{a}) \wedge \phi(a_2, \vec{a})) + w(\neg\phi(a_1, \vec{a}) \wedge \neg\phi(a_2, \vec{a})) = 1.$$

Since for $\theta, \xi \in SL_q$, $w(\theta \wedge \xi) = w(\xi)$ whenever $w(\theta) = 1$, it follows that (10) holds with equality. □

Note that the solution $c_0^{L^q}$ to (10) actually satisfies the stronger condition ULi + SN.

It is clear that in the proof of Theorem 2 the additional constants \vec{a} are playing an important role, and indeed that is the case. In particular, see [13], for $q \geq 2$ the probability functions $c_\lambda^{L^q}$ satisfy the weaker version of EAP without the additional constants \vec{a} , when, in fact exactly when,⁹

$$c_\lambda^{L^q}(\alpha_1 \alpha_2) \leq 2^{-q}(2^q - 1)^{-2}, \quad (12)$$

equivalently when

$$\lambda \leq \frac{2^q}{2^{2q} - 3 \cdot 2^q + 1}. \quad (13)$$

Consequently the only λ for which this principle can hold *for all* q is $\lambda = 0$.

Condition (13) is interesting because, to our knowledge, there are currently no other ‘rational principles’ considered in Inductive Logic which differentiate between the λ in the open range $(0, \infty)$. In this case the often preferred value for λ of 2^q (which corresponds to the uniform de Finetti prior for $c_\lambda^{L^q}$) lies above the bound given in (13) (for $q \geq 2$) so that with that somewhat popular choice this weaker version of EAP fails.

⁹There is a particular worst case here which occurs when $\phi(a_1) = \alpha_1(a_1)$ and $\psi(a_1) = \neg\alpha_2(a_1)$. Indeed this worst case and the bound given in (12) actually applies to *any* probability function satisfying *Atom Exchangeability* – for details of this and an investigation into variations on EAP and the underlying symmetry assumptions see [13].

Returning again to the superficial similarity between GAP and EAP one might initially have felt that

$$w(\phi(a_1, \vec{a}) \leftrightarrow \phi(a_2, \vec{a}) \mid \psi(a_1, \vec{a}) \wedge \psi(a_2, \vec{a})) \geq w(\phi(a_1, \vec{a}) \leftrightarrow \phi(a_1, \vec{a})), \quad (14)$$

$$w(\phi(a_2, \vec{a}) \mid \phi(a_1, \vec{a}) \wedge \psi(a_1, \vec{a}) \wedge \psi(a_2, \vec{a})) \geq w(\phi(a_2, \vec{a}) \mid \phi(a_1, \vec{a})), \quad (15)$$

were essentially expressing the same sentiment. The above discussion however show that (14) can hold whilst (15) fails. Conversely by taking $q = 2$, $\phi(a_1, \vec{a}) = \alpha_1(a_1)$, $\psi(a_1, \vec{a}) = \neg\alpha_2(a_1)$ and $\lambda > 4/5$ it can be checked that in this case (15) holds but (14) fails for $c_\lambda^{L_2}$.

We remark that although Theorems 1 and 2 are proved here for a unary language L it can be shown that when the additional constants \vec{a} are allowed they hold too, with the standard extension of c_0 (as $u^{(0,1,0,0,\dots),L}$, see [21, Chapter 29]), for not purely unary languages.¹⁰

The Constant Analogy Principle

The previous two attempts to capture even a part¹¹ of Bartha's representation of analogy can at best be said to tell us what is not possible in the presence of Px + SN. Perhaps there may be more probability functions satisfying these analogy principles if we dropped Px and/or SN, but given the obvious strong attraction of Px and SN on grounds of symmetry compared with the apparently hazy intuitions which begat GAP and EAP this would hardly seem a worthwhile investigation in the context.

¹⁰In more detail suppose that \mathcal{L} is a polyadic language and $w \neq c_0$ is a probability function on \mathcal{L} satisfying Px + SN. Then there must be some relation symbol P of \mathcal{L} and \vec{a}, \vec{b} such that $w(P(\vec{a}) \wedge \neg P(\vec{b})) > 0$. By considering $w(P(\vec{a}) \leftrightarrow \neg P(\vec{d}))$ where \vec{d} are new constants (not occurring in \vec{a} and \vec{b}) and using Ex we can see that we must have $w(P(\vec{c}_1) \wedge \neg P(\vec{c}_2)) > 0$ where the \vec{c}_1, \vec{c}_2 are disjoint. Now define the probability function v on L_1 by

$$v\left(\bigwedge_{i=1}^n R_1^{\epsilon_i}(a_i)\right) = w\left(\bigwedge_{i=1}^n P^{\epsilon_i}(\vec{c}_i)\right)$$

where \vec{c}_i are disjoint blocks of constants. Since w satisfies SN and Ex so does v . Let $\phi, \psi \in QFSL_1$ provide counter-examples required for Theorems 1 and 2 for v and let ϕ^*, ψ^* be the result of replacing each $R_1(a_i)$ in ϕ, ψ respectively by $P(\vec{c}_i)$. Then these provide the counter-examples required for Theorems 1 and 2 for w .

In the other direction to show that $c_0^{\mathcal{L}}$ is a solution notice that if \mathcal{L} has q relation symbols P_1, \dots, P_q then for $\theta \in QFSL$, $c_0^{\mathcal{L}}(\theta) = c_0^{L^q}(\theta')$ where θ' is the result of replacing each $P_i(a_{j_1}, a_{j_2}, \dots, a_{j_r})$ from \mathcal{L} by $R_i(a_{j_1})$.

¹¹Since negative analogies do not figure.

An alternative, which also seems closer to Bartha's intention, is to restrict the P, Q, A, B etc. to having the particularly simple form of just $R(a)$, i.e. a unary relation applied to a constant. This yields two further principles depending on whether we take the carrier of the analogy to be the constants or the relations.¹² In this section we take the analogy to be between the properties of two constants, the known positive analogies being instances where a predicate agreed on these two constants and a negative analogy when it disagreed. Precisely, for

$$\phi(x) = \bigwedge_{i=1}^n R_i^{\epsilon_i}(x), \quad \psi(x) = \bigwedge_{i=1}^n R_i^{\delta_i}(x),$$

we define the 'distance' between ϕ and ψ to be

$$[\phi - \psi] = \sum_{i=1}^n |\epsilon_i - \delta_i|$$

and propose:

The Constant Analogy Principle, CAP

For $\phi(x) = \bigwedge_{i=1}^n R_i^{\epsilon_i}(x)$ and $\psi(x) = \bigwedge_{i=1}^n R_i^{\delta_i}(x)$,

$$w(R_{n+1}(a_2) \mid R_{n+1}(a_1) \wedge \psi(a_2) \wedge \phi(a_1)) \tag{16}$$

is a decreasing (not necessarily strictly) function of $[\phi - \psi]$.

Given that in this principle no particular emphasis is being placed on the number of predicates R_1, R_2, \dots, R_n that we have at our disposal it seems natural to assume not just Px + SN but rather ULi + SN. In that case we have the following somewhat satisfying result.

Theorem 3. *Let the probability function w on L_q satisfy ULi + SN. Then w satisfies CAP.*

Proof. Since w is part of a ULi family w^{L_r} for $r \in \mathbb{N}^+$ we may take the union of all these probability functions to produce a probability function defined on sentences of the language $L = \bigcup_{r=1}^{\infty} L_r$ and extending w . To avoid introducing any additional notation, and because it will not cause any confusion, we will also use w to denote this probability function.

¹²It is also possible to consider a formalization in which it applies to sentences as carriers, see the *Counterpart Principle* of [12], [21, Chapter 22].

Let $\beta_1, \beta_2, \dots, \beta_{2n+1}$ denote the atoms of L_{n+1} , say

$$\begin{aligned}\beta_k(a_1) &= R_{n+1}(a_1) \wedge \phi(a_1), \\ \beta_j(a_2) &= R_{n+1}(a_2) \wedge \psi(a_2), \\ \beta_r(a_2) &= \neg R_{n+1}(a_2) \wedge \psi(a_2),\end{aligned}$$

where ϕ, ψ are as in the definition of CAP. Then

$$\begin{aligned}w(R_{n+1}(a_2) \mid R_{n+1}(a_1) \wedge \psi(a_2) \wedge \phi(a_1)) \\ = \frac{w(\beta_k(a_1) \wedge \beta_j(a_2))}{w(\beta_k(a_1) \wedge \beta_j(a_2)) + w(\beta_k(a_1) \wedge \beta_r(a_2))}.\end{aligned}\quad (17)$$

By using the permutations of atoms licensed by Px + SN we may suppose here that

$$\begin{aligned}\beta_k(a_1) &= \bigwedge_{i=1}^{n+1} R_i(a_1), \\ \beta_j(a_2) &= R_{n+1}(a_2) \wedge \bigwedge_{i=1}^m R_i(a_2) \wedge \bigwedge_{i=m+1}^n \neg R_i(a_2), \\ \beta_r(a_2) &= \neg R_{n+1}(a_2) \wedge \bigwedge_{i=1}^m R_i(a_2) \wedge \bigwedge_{i=m+1}^n \neg R_i(a_2),\end{aligned}$$

where $n - m = \lceil \phi - \psi \rceil$. From this it is clear that (16) is purely a function of $\lceil \phi - \psi \rceil$ (and n).

It only remains to show that it is a decreasing function of $\lceil \phi - \psi \rceil$. Assume for the moment that the

$$w(\beta_g(a_1) \wedge \beta_h(a_2))$$

are all non-zero. Then by dividing top and bottom by its numerator we see that (16) will be a decreasing function of $\lceil \phi - \psi \rceil$ just if

$$\frac{w(\beta_k(a_1) \wedge \beta_r(a_2))}{w(\beta_k(a_1) \wedge \beta_j(a_2))}\quad (18)$$

is an increasing function of $\lceil \phi - \psi \rceil$.

To this end define a function u on the state descriptions of the language L_1 (whose atoms are just $R_1, \neg R_1$) by

$$u(\neg R_1^s R_1^{t-s}) = 2^t w \left(\bigwedge_{i=1}^t R_i(a_1) \wedge \bigwedge_{i=1}^s \neg R_i(a_2) \wedge \bigwedge_{i=s+1}^t R_i(a_2) \right).$$

Notice that by Px the particular choice of distinct predicate symbols R_1, \dots, R_t here is irrelevant. Using the fact that w satisfies ULi + SN it can be checked that u is, or at least extends to, a probability function which satisfies Ex. With this definition the condition (18) becomes the requirement that

$$\frac{u((\neg R_1)^{m+1} R_1^{n-m})}{u((\neg R_1)^m R_1^{n+1-m})} \quad (19)$$

is an increasing function of m . This will follow once we show that

$$\frac{u((\neg R_1)^{m+1} R_1^{n-m})}{u((\neg R_1)^m R_1^{n+1-m})} \geq \frac{u((\neg R_1)^m R_1^{n-m+1})}{u((\neg R_1)^{m-1} R_1^{n+2-m})}. \quad (20)$$

Since u satisfies Ex we know by de Finetti's Theorem that for some countably additive measure μ on the Borel subsets of $[0, 1]$ that

$$u = \int w_{\langle x, 1-x \rangle} d\mu(x). \quad (21)$$

Using this and writing $h(x) = (1-x)^{m-1} x^{n-m}$, (20) becomes

$$\frac{\int (1-x)^2 h(x) d\mu(x)}{\int x(1-x) h(x) d\mu(x)} \geq \frac{\int x(1-x) h(x) d\mu(x)}{\int x^2 h(x) d\mu(x)}.$$

But multiplying out this reduces to

$$\int h(x) d\mu(x) \cdot \int x^2 h(x) d\mu(x) \geq \left(\int x h(x) d\mu(x) \right)^2$$

which holds by Hölder's Inequality.

Returning now to our earlier assumption that the $w(\beta_g(a_1) \wedge \beta_h(a_2))$ are all non-zero, if this fails then defining u as above we should have

$$u(R_1^k (\neg R_1)^j) = 0$$

for some j, k . However by considering again (21) this can only happen if

$$u = \nu w_{\langle 1, 0 \rangle} + (1-\nu) w_{\langle 0, 1 \rangle}$$

for some $0 \leq \nu \leq 1$, in which case the original requirement on (17) holds trivially. \square

Not all probability functions satisfying just Px + SN satisfy CAP. For example when $n + 1 = 3$ and we order the β_j as (in the obvious shorthand)

$$R_1^1 R_2^1 R_3^1, R_1^1 R_2^1 R_3^0, R_1^1 R_2^0 R_3^1, R_1^1 R_2^0 R_3^0,$$

$$R_1^0 R_2^1 R_3^1, R_1^0 R_2^1 R_3^0, R_1^0 R_2^0 R_3^1, R_1^0 R_2^0 R_3^0,$$

$b = 1/10$, $a = 1/5$ and w is

$$(12)^{-1} (w_{\langle a,b,b,a,b,b,b,b \rangle} + w_{\langle a,b,b,b,b,a,b,b \rangle} + w_{\langle a,b,b,b,b,b,a,b \rangle}$$

$$+ w_{\langle b,a,a,b,b,b,b,b \rangle} + w_{\langle b,a,b,b,a,b,b,b \rangle} + w_{\langle b,a,b,b,b,b,b,a \rangle}$$

$$+ w_{\langle b,b,a,b,a,b,b,b \rangle} + w_{\langle b,b,a,b,b,b,b,a \rangle} + w_{\langle b,b,b,b,a,b,b,a \rangle}$$

$$+ w_{\langle b,b,b,a,b,b,a,b \rangle} + w_{\langle b,b,b,a,b,a,b,b \rangle} + w_{\langle b,b,b,b,a,a,b,b \rangle})$$

then it can be checked that

$$\frac{w(\beta_1(a_1) \wedge \beta_8(a_2))}{w(\beta_1(a_1) \wedge \beta_4(a_2))} < \frac{w(\beta_1(a_1) \wedge \beta_7(a_2))}{w(\beta_1(a_1) \wedge \beta_3(a_2))},$$

contrary to the requirement on (18), despite w satisfying Px + SN (but not ULi + SN of course).

From Theorem 3 it follows that all the $c_\lambda^{L_q}$ satisfy CAP. Indeed it is quite straightforward to show, appealing to de Finetti's Theorem again, that *any* probability function satisfying Atom Exchangeability will satisfy CAP, whether or not it satisfies ULi + SN.

Satisfying as Theorem 3 may be however it does raise a slightly uncomfortable issue. Firstly the intuition behind it seems no different from that which initially prompted us to propose GAP. Given the failure of that principle can we really claim that Theorem 3 in some way justifies our intuition? Is it not more reasonable to conclude that this theorem is grounded not on 'analogy' but on some different basis? And indeed a study at the proof shows that the key step is an application of a provable version of the *Strong Principle of Instantial Relevance*, see [21, Chapter 21] or [22], in which case we could be said to be simply appealing to our intuitions about *relevance*¹³. Putting it another way then it could be said that the 'analogy' within CAP is really just reducible to 'relevance', raising for a moment the question whether, within the context of PIL, analogy is anything more than a special case of relevance.

¹³Or ultimately symmetry since Ex implies SPIR for L_1 .

The Predicate Analogy Principle

In contrast to the three previous sections we now consider an interpretation of Bartha's representation in which we take the predicates of the language to be the carriers of the analogy. That is we take the analogy to be between the properties of two predicates, the known positive analogies being instances where these predicates agreed on a constant and a negative analogy when they disagreed. Precisely, for

$$\phi = \bigwedge_{i=1}^n R_1^{\epsilon_i}(a_i), \quad \psi = \bigwedge_{i=1}^n R_2^{\delta_i}(a_i),$$

define the 'distance' between ϕ and ψ to be

$$[\phi - \psi] = \sum_{i=1}^n |\epsilon_i - \delta_i|$$

and set:

The Predicate Analogy Principle, PAP

For $\phi(\vec{a}) = \bigwedge_{i=1}^n R_1^{\epsilon_i}(a_i)$ and $\psi(\vec{a}) = \bigwedge_{i=1}^n R_2^{\delta_i}(a_i)$,

$$w(R_2(a_{n+1}) \mid R_1(a_{n+1}) \wedge \psi(\vec{a}) \wedge \phi(\vec{a})) \quad (22)$$

is a decreasing function of $[\phi - \psi]$ (for fixed n).

Notice that since only two predicate symbols appear in (22) it is natural to first study this principle when $q = 2$.¹⁴ Setting

$$\alpha_1(x) = R_1(x) \wedge R_2(x), \quad \alpha_2(x) = R_1(x) \wedge \neg R_2(x),$$

$$\alpha_3(x) = \neg R_1(x) \wedge R_2(x), \quad \alpha_4(x) = \neg R_1(x) \wedge \neg R_2(x).$$

this condition (22) is equivalent to¹⁵

$$\frac{w(\alpha_2(a_{n+1}) \wedge \phi(\vec{a}) \wedge \psi(\vec{a}))}{w(\alpha_1(a_{n+1}) \wedge \phi(\vec{a}) \wedge \psi(\vec{a}))},$$

being an increasing function of $[\phi - \psi]$, a fact that we shall use repeatedly in what follows.

¹⁴The characterization for $q > 2$ (with Px + SN) just requires the restriction to SL_2 to have the form we shall shortly be describing.

¹⁵Recall the convention introduced at footnote (4) concerning zero denominators.

One family of probability functions on L_2 satisfying Px + SN and PAP are the

$$u^{(b)} = 2^{-1}(w_{\langle b, 1/2-b, 1/2-b, b \rangle} + w_{\langle 1/2-b, b, b, 1/2-b \rangle})$$

where $0 \leq b \leq 1/2$. Clearly the $u^{(b)}$ satisfy Px + SN. To see that they also satisfy PAP notice that for ϕ, ψ as in the statement of PAP and $m = \lfloor \phi - \psi \rfloor$,

$$\frac{u^{(b)}(\alpha_2(a_{n+1}) \wedge \psi(\vec{a}) \wedge \phi(\vec{a}))}{u^{(b)}(\alpha_1(a_{n+1}) \wedge \psi(\vec{a}) \wedge \phi(\vec{a}))} = \frac{(1/2 - b)^{m+1} b^{n-m} + b^{m+1} (1/2 - b)^{n-m}}{(1/2 - b)^m b^{n-m+1} + b^m (1/2 - b)^{n-m+1}}$$

and this right hand side, when defined, is increasing in m (for fixed $n \geq m$).

A second family of probability functions on L_2 satisfying PAP + Px + SN, in this case rather trivially, are those of the form

$$v^{(d)} = 4^{-1}(w_{\langle d, 0, 0, 1-d \rangle} + w_{\langle 1-d, 0, 0, d \rangle} + w_{\langle 0, d, 1-d, 0 \rangle} + w_{\langle 0, 1-d, d, 0 \rangle})$$

where $0 \leq d \leq 1$. Trivially because any $\phi(\vec{a}) \wedge \psi(\vec{a})$ containing atoms both from $\{\alpha_1, \alpha_4\}$ and from $\{\alpha_2, \alpha_3\}$ gets probability zero.

In fact the probability functions which satisfy PAP + Px + SN are precisely those whose restriction to SL_2 is a convex mixture of probability functions from these two families. Precisely:

Theorem 4. *Let the probability function w on L_2 satisfy Px + SN. Then w satisfies PAP just if either w is a convex mixture of the $u^{(b)}$ or w is a convex mixture of the $v^{(d)}$ as above, equivalently, just if there is a countably additive measure μ on the Borel subsets of \mathbb{D}_4 such that for $\theta \in SL_2$,*

$$w(\theta) = \int_{\mathbb{D}_4} w_{\vec{x}}(\theta) d\mu(\vec{x})$$

and either $\mu(A) = 1$ where

$$A = \{\langle x_1, x_2, x_3, x_4 \rangle \in \mathbb{D}_4 \mid x_1 = x_4 \text{ and } x_2 = x_3\}.$$

or $\mu(B) = 1$ where

$$B = \{\langle x_1, x_2, x_3, x_4 \rangle \in \mathbb{D}_4 \mid x_2 = x_3 = 0 \text{ or } x_1 = x_4 = 0\}.$$

Proof. Suppose that w satisfies PAP + Px + SN. Let $\vec{b} = \langle b_1, b_2, b_3, b_4 \rangle$ be a support point of the de Finetti prior μ of w . We shall use [21, Lemma 12.1] which tells us that

$$\lim_{r \rightarrow \infty} \frac{\int_{\mathbb{D}_4} x_i \prod_{j=1}^4 x_j^{[rb_j]} d\mu(\vec{x})}{\int_{\mathbb{D}_4} \prod_{j=1}^4 x_j^{[rb_j]} d\mu(\vec{x})} = b_i, \quad (23)$$

where as usual $[rb_1]$ is the integer part of rb_1 etc.

Let

$$\begin{aligned}\phi_1 \wedge \psi_1 &= \alpha_1^{[rb_1]} \alpha_2^{[rb_2]} \alpha_3^{[rb_3]} \alpha_4^{[rb_4]}, \\ \phi_2 \wedge \psi_2 &= \alpha_1^{[rb_1]} \alpha_2^{[rb_3]} \alpha_3^{[rb_2]} \alpha_4^{[rb_4]},\end{aligned}$$

where the ϕ_i, ψ_i only mention the predicate R_1, R_2 respectively. Note that

$$w(\alpha_i \wedge \phi_1 \wedge \psi_1) = \int_{\mathbb{D}_4} x_i \prod_{j=1}^4 x_j^{[rb_j]} d\mu(\vec{x})$$

First assume that $b_1 \neq 0$. Using (23) we can make

$$\frac{w(\alpha_2 \wedge \phi_1 \wedge \psi_1)}{w(\alpha_1 \wedge \phi_1 \wedge \psi_1)}$$

arbitrarily close to

$$\frac{w_b(\alpha_2)}{w_b(\alpha_1)} = \frac{b_2}{b_1}$$

by picking r suitably large. Similarly, since by Px + SN, $\langle b_1, b_3, b_2, b_4 \rangle$ must also be a support point of μ , we can make

$$\frac{w(\alpha_2 \wedge \phi_2 \wedge \psi_2)}{w(\alpha_1 \wedge \phi_2 \wedge \psi_2)}$$

arbitrarily close to b_3/b_1 . But since

$$[\phi_1 - \psi_1] = [\phi_2 - \psi_2]$$

PAP gives that these two values $b_2/b_1, b_3/b_1$ must be equal. It follows that $b_2 = b_3$. If $b_2 = 0$ then we already have that $\langle b_1, b_2, b_3, b_4 \rangle \in B$. If $b_2 \neq 0$ a similar argument using the support points $\langle b_2, b_1, b_4, b_3 \rangle, \langle b_2, b_4, b_1, b_3 \rangle$ shows that $b_1 = b_4$. From this it follows that $\mu(A \cup B) = 1$ for A, B as above.

We claim that it must be the case that $\mu(A) = 1$ or $\mu(B) = 1$. For otherwise we can pick support points of μ , $\langle b, 1/2 - b, 1/2 - b, b \rangle \in A - B$ and, without loss of generality, $\langle d, 0, 0, 1 - d \rangle \in B - A$ (with $0 < b < 1/2$ and $d \neq 0, 1/2$). Then by PAP we must have equality between

$$\frac{\int_A x_1^k x_2 x_4^{n-k} d\mu(\vec{x}) + \int_{B-A} x_1^k x_2 x_4^{n-k} d\mu(\vec{x})}{\int_A x_1^{k+1} x_4^{n-k} d\mu(\vec{x}) + \int_{B-A} x_1^{k+1} x_4^{n-k} d\mu(\vec{x})}$$

and

$$\frac{\int_A x_1^j x_2 x_4^{n-j} d\mu(\vec{x}) + \int_{B-A} x_1^j x_2 x_4^{n-j} d\mu(\vec{x})}{\int_A x_1^{j+1} x_4^{n-j} d\mu(\vec{x}) + \int_{B-A} x_1^{j+1} x_4^{n-j} d\mu(\vec{x})}$$

for $k, j \leq n$. Since for any $\langle x_1, x_2, x_3, x_4 \rangle \in A$ we have $x_1 = x_4$ it must be the case that

$$\begin{aligned}\int_A x_1^k x_2 x_4^{n-k} d\mu(\vec{x}) &= \int_A x_1^j x_2 x_4^{n-j} d\mu(\vec{x}), \\ \int_A x_1^{k+1} x_4^{n-k} d\mu(\vec{x}) &= \int_A x_1^{j+1} x_4^{n-j} d\mu(\vec{x}),\end{aligned}$$

and by the existence of $\langle b, 1/2 - b, 1/2 - b, b \rangle$ these are non-zero. Furthermore,

$$\int_{B-A} x_1^k x_2 x_4^{n-k} d\mu(\vec{x}) = 0 = \int_{B-A} x_1^j x_2 x_4^{n-j} d\mu(\vec{x})$$

for $n > 0$, so it must be the case that

$$\int_{B-A} x_1^{k+1} x_4^{n-k} d\mu(\vec{x}) = \int_{B-A} x_1^{j+1} x_4^{n-j} d\mu(\vec{x}).$$

Hence

$$\frac{\int_{B-A} x_1^{[dm]+2} x_4^{[(1-d)m]} d\mu(\vec{x})}{\int_{B-A} x_1^{[dm]} x_4^{[(1-d)m]} d\mu(\vec{x})} = \frac{\int_{B-A} x_1^{[dm]+1} x_4^{[(1-d)m]+1} d\mu(\vec{x})}{\int_{B-A} x_1^{[dm]} x_4^{[(1-d)m]} d\mu(\vec{x})}. \quad (24)$$

Let

$$w' = (\mu(B - A))^{-1} \int_{B-A} w_{\vec{x}} d\mu(\vec{x}).$$

Using $\mu(A \cup B) = 1$ and $d \neq 0, 1/2$ we can see that $\langle d, 0, 0, 1 - d \rangle$ is a support point of w' . Hence it follows from [21, Lemma 12.1] that for large m the integrals (24) are close to $d^2, d(1 - d)$ respectively, which is impossible.

In the other direction let w satisfy Px + SN and $\mu(A) = 1$. Define a probability function v on the language L_1 with a single predicate symbol R_1 by

$$v \left(\bigwedge_{i=1}^m R_1^{\epsilon_i}(a_i) \right) = w \left(\bigwedge_{i=1}^m (\alpha_1(a_i) \vee \alpha_4(a_i))^{\epsilon_i} \right) = 2^m w \left(\bigwedge_{i=1}^m \alpha_{h_i}(a_i) \right)$$

for $h_i \in \{1, 4\}$ when $\epsilon_i = 1$, $h_i \in \{2, 3\}$ when $\epsilon_i = 0$. Then v satisfies Ex + SN and for $m = \lfloor \phi - \psi \rfloor$

$$\frac{w(\alpha_2 \wedge \phi \wedge \psi)}{w(\alpha_1 \wedge \phi \wedge \psi)} = \frac{v(R_1^{n-m}(\neg R_1)^{m+1})}{v(R_1^{n-m+1}(\neg R_1)^m)},$$

which we have already met as (19) and shown to be increasing in m under the assumption of Ex.

Finally in the case when $\mu(B) = 1$ it is straightforward to check that PAP holds, trivially in fact. □

Given Ax there are only two probability functions on L_2 satisfying this, c_0 and c_∞ . For suppose w satisfies PAP and Ax and as usual let

$$\begin{aligned}\alpha_1(x) &= R_1(x) \wedge R_2(x), & \alpha_2(x) &= R_1(x) \wedge \neg R_2(x), \\ \alpha_3(x) &= \neg R_1(x) \wedge R_2(x), & \alpha_4(x) &= \neg R_1(x) \wedge \neg R_2(x).\end{aligned}$$

Then by PAP

$$w(R_2(a_2) \mid R_1(a_2) \wedge \alpha_2(a_1)) = w(R_2(a_2) \mid R_1(a_2) \wedge \alpha_3(a_1)).$$

But that gives that either $w(\alpha_2(a_1) \wedge \alpha_1(a_2)) = 0$ or

$$w(\alpha_2(a_1) \wedge \alpha_2(a_2)) = w(\alpha_2(a_1) \wedge \alpha_3(a_2))$$

and with Ax the only possibilities here are c_0 and c_∞ .

Analogy as Possibility

In the previous four sections we have looked at formulations of analogical support as enhancement of probability. However as Bartha points out in [1] analogy can act to simply engender plausibility, or as we shall call it possibility. To give an example, the fact that the commonest bird in the United States in 1814 (the passenger pigeon) was extinct by 1914 may be used as an argument that ‘by analogy’, the monarch - arguably the currently commonest butterfly in the United States, may equally regrettably be extinct a century from now. For here it seems that the argument is aimed not so much at raising the probability as creating the possibility which we will take to mean producing a non-zero probability.

One explanation why we might see this as in any sense a worthwhile argument to make in a discussion on the future of the monarch is that viewed from a certain angle monarchs and passenger pigeons may be thought of as the same thing, at least as regards the features that are actually relevant here. Thus the realization that it had happened once argues that it could happen again.

The example might seem to correspond to the Extended Principle of Instantial Relevance, (2). Here however we shall propose an alternative formulation which is about creating possibility, and also more obviously captures this idea of ‘being thought of as the same thing as regards the relevant features’.

Dolly’s Principle, DP

For $\theta(a_1, a_2, \dots, a_m) \in SL$, if $\sigma : \{a_1, a_2, \dots, a_m\} \rightarrow \{a_1, a_2, \dots, a_m\}$ and $w(\theta(\sigma(a_1), \sigma(a_2), \dots, \sigma(a_m))) > 0$ then $w(\theta(a_1, a_2, \dots, a_m)) > 0$.

Notice that by repeated application (and Ex) it is enough that this principle holds for $\sigma(2) = 1, \sigma(i) = i$ for $i \neq 2$.

We shall now show that for the unary language L_q Ex already implies DP. First however we seem to need a (useful) lemma which applies even to a possibly polyadic language L .

Lemma 5. *Let $\theta(a_1, a_2, \dots, a_m) \equiv \phi(a_1, a_2, \dots, a_m)$. Then for $\sigma : \{a_1, a_2, \dots, a_m\} \rightarrow \{a_1, a_2, \dots, a_m\}$,*

$$\theta(\sigma(a_1), \sigma(a_2), \dots, \sigma(a_m)) \equiv \phi(\sigma(a_1), \sigma(a_2), \dots, \sigma(a_m)).$$

Proof. Let K be a structure for the language \mathcal{L} with the same relation symbols as L but only the constant symbols $\sigma(a_1), \sigma(a_2), \dots, \sigma(a_m)$ and suppose that

$$K \models \theta(\sigma(a_1), \sigma(a_2), \dots, \sigma(a_m)).$$

Then

$$K \models \theta(\sigma(a_1)^K, \sigma(a_2)^K, \dots, \sigma(a_m)^K),$$

where $\sigma(a_1)^K$ is the interpretation of the constant $\sigma(a_1)$ in K etc.. Clearly also

$$J \models \theta(\sigma(a_1)^K, \sigma(a_2)^K, \dots, \sigma(a_m)^K)$$

where J is a structure for L extending K in which $a_i^J = \sigma(a_i)^K$ for $i \leq m$. In other words

$$J \models \theta(a_1^J, \dots, a_m^J),$$

so

$$J \models \theta(a_1, \dots, a_m)$$

and by logical equivalence,

$$J \models \phi(a_1, \dots, a_m).$$

Reversing the above argument with ϕ in place of θ now gives

$$K \models \phi(\sigma(a_1), \dots, \sigma(a_m)),$$

as required. □

Notice that an immediate corollary of this result is that Super Regularity, i.e. that $w(\psi) > 0$ whenever $\psi \in SL$ is consistent, already implies DP (via showing that if $\theta(a_1, \dots, a_m)$ is inconsistent then so is $\theta(a_{\sigma(1)}, \dots, a_{\sigma(m)})$).

Theorem 6. *For the unary language L_q , Ex implies DP .*

Proof. Let $\theta(a_1, \dots, a_m) \in SL_q$. Then as at (9) θ is logically equivalent to a disjunction of sentences of the form

$$\bigwedge_{i=1}^m \alpha_{h_i}(a_i) \wedge \bigwedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\epsilon_j}.$$

If $w(\theta(\sigma(a_1), \sigma(a_2), \dots, \sigma(a_m))) > 0$ then by Lemma 5 this must also hold for the image under σ of this representation of θ so for at least one such disjunct we must have

$$w \left(\bigwedge_{i=1}^m \alpha_{h_i}(\sigma(a_i)) \wedge \bigwedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\epsilon_j} \right) > 0. \quad (25)$$

The only way this is possible is if $h_i = h_r$ whenever $\sigma(a_i) = \sigma(a_r)$, otherwise the sentence in (25) would be inconsistent so have probability 0. So dropping repeated conjuncts (25) can equivalently be written as

$$w \left(\bigwedge_{a_r \in Rg(\sigma)} \alpha_{g_r}(a_r) \wedge \bigwedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\epsilon_j} \right) > 0 \quad (26)$$

where $g_r = h_i$ for i such that $\sigma(a_i) = a_r$.

From (26), de Finetti's Theorem and the Constant Irrelevance Principle give

$$\int_{\mathbb{D}_{2^q}} \prod_{a_r \in Rg(\sigma)} x_{g_r} w_{\vec{x}} \left(\bigwedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\epsilon_j} \right) d\mu(\vec{x}) > 0. \quad (27)$$

So

$$\prod_{a_r \in Rg(\sigma)} x_{g_r} w_{\vec{x}} \left(\bigwedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\epsilon_j} \right)$$

must be non-zero for a non-null (with respect to μ) set of $\vec{x} \in \mathbb{D}_{2^q}$ and as a result the same must hold for

$$\prod_{a_r \in Rg(\sigma)} x_{g_r}^{s_r} w_{\vec{x}} \left(\bigwedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\epsilon_j} \right)$$

where s_r is the number of a_i mapped by σ to a_r . Hence

$$\int_{\mathbb{D}_{2^q}} \prod_{a_r \in Rg(\sigma)} x_{g_r}^{s_r} w_{\vec{x}} \left(\bigwedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\epsilon_j} \right) d\mu(\vec{x}) > 0. \quad (28)$$

But the left hand side of (28) is just what we get if we apply de Finetti's Theorem to this conjunct in the representation of $\theta(a_1, \dots, a_m)$, so the required result follows. \square

For unary languages then DP adds nothing new, it already follows from the standing assumption Ex. However this fact does not carry over to polyadic languages. For example if L is the language with a single binary relation symbol R and w is the obvious version of Carnap's m^\dagger (equivalently c_∞) on this language then

$$w(\forall x (R(a_1, x) \leftrightarrow R(a_2, x))) = 0$$

whilst

$$w(\forall x (R(a_1, x) \leftrightarrow R(a_1, x))) = 1.$$

Nevertheless there is still a wide class of polyadic probability functions which do satisfy DP, as we shall show in a forthcoming paper.

Conclusion

In short we have shown that in the presence of Ex + Px + SN the principles GAP, EAP and PAP place very severe demands on a probability function, and must now be considered dead ends, DP makes no demands at all whilst CAP is actually satisfied by a naturally attractive class of such probability functions, namely those that satisfy the somewhat stronger background condition of ULi + SN.

As far as Bartha's candidate representation is concerned then we can say that CAP seems to provide a viable formulation of it in the context of PIL whereas GAP, EAP and PAP do not. Still this raises the uncomfortable question of why they produce such different conclusions when they all appear to be based on similar intuitions about analogical support.

Given that CAP follows from ULi + SN it is an interesting question to ask from where in this background assumption the 'analogical support' originates. Inspecting the proof of Theorem 3 we see that the key inequality is (20) which derives from simply the assumption Ex via de Finetti's Theorem. This is exactly similar to the derivation from Ex of the Extended Principle of Instantial Relevance from which we might reasonably question whether 'analogy as enhancement of probability' is really anything more than 'relevance', an already quite well studied notion (see [21]).

Throughout this paper we have taken Ex + Px + SN, or ULi + SN, as our background assumptions. However the rather widespread acceptance of Johnson's Sufficiency Postulate (JSP) within Inductive Logic might on the contrary be

used as an argument for strengthening these to Ax , or $ULi + Ax$, since these are consequences of JSP. Doing so would still give CAP (and DP) as a consequence but would, for $q \geq 2$, restrict GAP, EAP and PAP down to the single probability function $c_0^{L^q}$ of Carnap's Continuum.

Combined with previous results in [11], [12] a pattern seems to be emerging with so called 'Analogy Principles', namely that they either hold, almost by chance, for some small family of otherwise (apparently) undistinguished probability functions, or they are actually consequences of some already established and acceptable principles such as $ULi + SN$ or Ax . In other words, to our knowledge we do not currently have any analogy principles which genuinely introduce new concepts without also reducing the field of 'rational probability functions' down to almost a triviality (and leading to the conclusion that such a version of 'reasoning by analogy' is both very powerful and very dangerous). Of course there are numerous further formulations and variations that one might base on CAIR and perhaps some of those might yet endorse it in the context of PIL. On the basis of what we have here however the picture of analogical support as presented in CAIR seems not to have materialized, an outcome in parallel with Bartha's own criticisms of CAIR within what we would term Applied Inductive Logic.

References

- [1] Bartha, P., Analogy and Analogical Reasoning, *The Stanford Encyclopedia of Philosophy (Fall 2013 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2013/entries/reasoning-analogy/>.
- [2] Burch, Robert, Charles Sanders Peirce, *The Stanford Encyclopedia of Philosophy (Winter 2014 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2014/entries/peirce/>.
- [3] 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.
- [4] Carnap, R. & Stegmüller, W., *Induktive Logik und Wahrscheinlichkeit*, Springer Verlag, Wien, 1959.
- [5] Festa, R., Analogy and Exchangeability in Predictive Inferences, *Erkenntnis*, 1996, **45**:229-252.
- [6] Gaifman, H., Concerning Measures on First Order Calculi, *Israel Journal of Mathematics*, 1964, **2**:1-18.

- [7] Gaifman, H., Applications of de Finetti's Theorem to Inductive Logic, in *Studies in Inductive Logic and Probability*, Volume I, Eds. R.Carnap & R.C.Jeffrey, University of California Press, 1971, pp235-251.
- [8] Grove, A.J., Halpern, J.Y. & Koller, D., Random Worlds and Maximum Entropy, *Journal of Artificial Intelligence Research*, 1994, **2**:33-88.
- [9] Hesse, M.B., *The Structure of Scientific Inference*, University of California Press, 1974.
- [10] Hesse, M.B., Analogy and Confirmation Theory, *Dialectica*, September 1963, **17**(2-3):284-292.
- [11] Hill, A. & Paris, J.B., An Analogy Principle in Inductive Logic, *Annals of Pure and Applied Logic*, 2013, **164**:1293-1321.
- [12] Hill, A. & Paris, J.B., The Counterpart Principle of Analogical Support by Similarity, *Erkenntnis*, Online First Article, October 2013.
- [13] Howarth, E., *New Rationality Principles in Pure Inductive Logic*, Doctoral Thesis, The University of Manchester, Manchester, UK, 2015. To become available at <http://www.maths.manchester.ac.uk/~jeff/>
- [14] Humburg, J., The Principle of Instantial Relevance, *Studies in Inductive Logic and Probability*, Eds. R.Carnap & R.C.Jeffrey, vol.I, University of California Press, Berkeley and Los Angeles, 1971, pp225-233.
- [15] Johnson, W.E., Probability: The Deductive and Inductive Problems, *Mind*, 1932, **41**:409-423.
- [16] Maher, P., Probabilities for Multiple Properties: The Models of Hesse, Carnap and Kemeny, *Erkenntnis*, 2001, **55**:183-216.
- [17] Maher, P., A Conception of Inductive Logic, *Philosophy of Science*, 2006, **73**:513-520.
- [18] di Maio, M.C., Predictive Probability and Analogy by Similarity, *Erkenntnis*, 1995, **43**(3):369-394.
- [19] Supplement to this paper, available at <http://www.maths.manchester.ac.uk/~jeff/papers/jp150303A1.pdf>
- [20] Paris, J.B., Pure Inductive Logic, *The Continuum Companion to Philosophical Logic*, L.Horsten & R.Pettigrew (eds.), Continuum International Publishing Group, London, 2011, pp428-449.

- [21] Paris, J.B. & Vencovská, A., *Pure Inductive Logic*, to appear in the Association of Symbolic Logic Series *Perspectives in Mathematical Logic*, Cambridge University Press.
- [22] Paris, J.B. & Waterhouse, P., Atom Exchangeability and Instantial Relevance, *Journal of Philosophical Logic*, 2009, **38**(3):313-332.
- [23] Polya, G., *Induction and Analogy in Mathematics, Volume I of Mathematics and Plausible Reasoning*, Geoffrey Cumberlege, Oxford University Press, 1954.
- [24] Polya, G., *Patterns of Plausible Inference, Volume II of Mathematics and Plausible Reasoning*, Geoffrey Cumberlege, Oxford University Press, 1954.
- [25] Romeijn, J.W., Analogical Predictions for Explicit Similarity, *Erkenntnis*, 2006, **64**(2):253-280.
- [26] 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.