Analogy as Possible Identity within Pure Inductive Logic

J.B.Paris[*] and A.Vencovská[†]
School of Mathematics
The University of Manchester
Manchester M13 9PL

jeff.paris@manchester.ac.uk

alena.vencovska@manchester.ac.uk

April 18, 2017

Abstract

Within the framework of Pure Inductive Logic we investigate a regularity principle of Analogy based on the interpretation of analogy as possible identity and locate it with respect to the central principles of Constant and Spectrum Exchangeability.

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

Introduction

Suppose I come across the word 'gormful'. I know that the word 'gormless' means dull, stupid, and that the pair 'hopeless-hopeful' are opposites. I therefore think

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

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

it possible that 'gormful' means sharp, clever. Asked why I think that I might suggest it is *by analogy*. Pressed further and asked what I mean by analogous I might say that as far as I know the pairs ⟨hopeless, hopeful⟩ and ⟨gormless, gormful⟩ could in a certain sense be the same thing, namely both being pairs of opposites. For that reason I have some grounds for thinking that gormful means the opposite of gormless.

Notice that this argument by analogy is rather weak. It simply concludes that this meaning of gormful is *possible*, which we shall identify with having non-zero *subjective* probability. Apart from that it makes no assertion about the magnitude of this probability. Indeed there may very well be other arguments by analogy which conclude that gormful not having this meaning is also possible.

The purpose of this paper is to propose a principle based on this insight in the context of Pure Inductive Logic and to relate it to the currently central principles in the arena, most notably Constant and Spectrum Exchangeability.

The idea of analogical arguments as pointers to qualitative possibility, rather than an enhanced quantitative probability (see [16] for a short overview), is certainly not new and is discussed for example in Bartha's [1]. In particular in [7] we have already briefly considered a principle which is a further weakening of the one we shall study here.

Background

The setting for this paper is Pure Inductive Logic, PIL for short, which is explained in some detail in, for example, [14] or [15]. In short we work in a predicate language L with finitely many relation symbols $R_1, R_2, \ldots, R_q$ of arities $r_1, r_2, \ldots, r_q$ respectively, and constants $a_1, a_2, a_3, \ldots$ and no function symbols nor equality. We imagine an agent who lives in a first order structure M for L in which the interpretations of the constants a_i (also denoted a_i) constitute the universe. In the absence of any further information about M the agent intends to assign rational probabilities $w(\theta)$ to the sentences θ of L, the set of which we denoted by SL.

In other words the agent aims to select a probability function $w : SL \to [0, 1]$, i.e. satisfying

(i) If $\models \theta$ then $w(\theta) = 1$,

(ii) If $\theta \models \neg\phi$ then $w(\theta \vee \phi) = w(\theta) + w(\phi)$,

(iii) $w(\exists x\, \psi(x)) = \lim_{n\to\infty} w\left(\bigvee_{i=1}^{n} \psi(a_i)\right)$, together with various other perceived rationality requirements.

Precisely what such rationality requirements should be, and what they further entail, is the core concern of PIL. By now a number of such requirements, or more grandly 'principles', have been suggested on the grounds of being in some intuitive sense rational, despite the absence of any clear understanding of what that word means. PIL may be viewed as an attempt to furnish such an understanding.

One such rationality principle is Constant Exchangeability, based on the idea that in this uninterpreted situation the agent has no rational reason to treat the constants differently:

The Constant Exchangeability Principle, Ex.

If $\theta \in SL$ and the constant symbol a_j does not appear in θ then $w(\theta) = w(\theta')$ where θ' is the result of replacing each occurrence of a_i in θ by a_j.

Constant Exchangeability is widely accepted and in this paper we shall take it as implicit that all probability functions mentioned satisfy Ex.

Touching on a somewhat different facet of rationality is the notion of Regularity:

The Principle of Regularity, Reg

If $\theta \in QFSL$ is consistent then $w(\theta) > 0$.

Here $QFSL$ is the set of quantifier free sentences of L. We shall refer to a strenghtening of this principle, requiring $w(\theta) > 0$ for any $\theta \in SL$, as *Super Regularity*, SReg. The justification for Reg/SReg is that if $\theta \in QFSL/SL$ is possible, i.e. consistent, then it should be given some non-zero probability. It turns out that most of the popular purportedly rational probability functions in the area (in particular the c_λ^L for $0 < \lambda$ of Carnap's Continuum of Inductive Methods and their polyadic extensions, see [15]) naturally satisfy Reg. Despite apparently having the same underlying motivation however the same is not true of SReg; this is a principle which we rarely (if ever) meet without specifically forcing it to hold (see [15, Chapter 10]).

Another principle which will figure in this paper is Spectrum Exchangeability. To introduce it, we need some notation.

To avoid double suffices let $b_1, b_2, \ldots, b_m$ stand for some (distinct) constants[1] a_i of L. A *state description* for $b_1, b_2, \ldots, b_m$ is a quantifier free sentence of L of the

[1] Correctly we should refer to these as *constant symbols* but we imagine the reader will not object to our informality (and similarly for relations and relation symbols.)

form

$$\bigwedge_{j=1}^{q} \bigwedge_{c_1,c_2,\ldots,c_{r_j}} \pm R_j(c_1, c_2, \ldots, c_{r_j})$$

where the $c_1, c_2, \ldots, c_{r_j}$ range over not necessarily distinct elements from $\{b_1, b_2, \ldots, b_m\}$ and $\pm R_j$ stands for one of $R_j, \neg R_j$. So a state description for $b_1, b_2, \ldots, b_m$ specifies precisely which R_j hold for which r_j-tuples from $b_1, b_2, \ldots, b_m$ and which do not hold.

Given a state description $\Theta(b_1, b_2, \ldots, b_m)$ (upper case Greek letters will always denote state descriptions) we define an equivalence relation $\sim_\Theta$ on $\{b_1, b_2, \ldots, b_m\}$ by

$$b_i \sim_\Theta b_j \iff \Theta(\vec{b}) \wedge b_i = b_j \text{ is consistent when we add equality to } L.^2$$

In other words as far as $\Theta(\vec{b})$ is concerned b_j is *indistinguishable* from b_i.

The *spectrum* of $\Theta(\vec{b})$, denoted $\mathcal{S}(\Theta)$, is the multiset of sizes of the equivalence classes of $\sim_\Theta$.

The Principle of Spectrum Exchangeability, Sx

If $\Theta(b_1, b_2, \ldots, b_m), \Phi(b_1, b_2, \ldots, b_m)$ *are state descriptions with the same spectrum then* $w(\Theta(\vec{b})) = w(\Phi(\vec{b}))$.[3]

Spectrum Exchangeability is nowadays seen as a principle based on *irrelevance*. It says that as far as the probability assigned to a state description is concerned all that matters is how many distinguished constants there are and the multiplicities of those that are indistinguishable, exactly how they are distinguished is irrelevant.

Prima facie Sx is a strong principle and indeed that turns out to be the case in that it has many interesting and arguably desirable consequences, see [15, Chapter 27], which certainly makes it attractive. In fact there is a further argument for Sx, namely that if we restrict L to being unary, as was the case for the pioneers of Inductive Logic, Rudolf Carnap and W.E.Johnson, see for example [3], [4], [9], then Sx amounts to the *Principle of Atom Exchangeability* acceptable to both Carnap and Johnson. Interestingly Atom Exchangeability is usually thought of as a principle based on symmetry, as invariance of assigned probabilities under permutations of atoms, rather than irrelevance.[4]

---

[2] And of course the axioms of equality to $\models$.

[3] See any of [15, p193], [10], [11], for specific examples.

[4] Another natural way of extending Atom Exchangeability to the polyadic context is the Permutation Invariance Principle discussed e.g in [15, Chapters 39-41], which is entirely symmetry-based, see also [17].

For several of the proofs that follow we need to recall the definitions of the probability functions $u^{\bar{p},L}$, $v_t^{\bar{p},L}$ satisfying Sx, the associated notation, and the representation theorems for probability functions satisfying Sx:

We say that probability function w on SL is *homogeneous* if it satisfies Sx and for each $t \in \mathbb{N}^+ = \{1, 2, 3, \ldots\}$

$$\lim_{n\to\infty} w\left(\bigvee_{|\mathcal{S}(\Phi(a_1,\ldots,a_n))|=t} \Phi(a_1,\ldots,a_n)\right) = 0. \tag{1}$$

The disjunction is taken over all state descriptions of L for constants $a_1, \ldots, a_n$ with spectrum size t. Notice that since we are assuming Ex the $a_1, \ldots, a_n$ could be replaced here by any other (distinct) constants $b_1, \ldots, b_n$.

For a given state description $\Phi(b_1, \ldots, b_m)$ and a given vector $\vec{c} \in \mathbb{N}^m$, Φ is said to be *consistent with* $\vec{c}$ if for $1 \le s, t \le m$, $c_s = c_t \neq 0 \implies b_s \sim_\Phi b_t$. The set of all state descriptions for $\vec{b}$ which are consistent with $\vec{c}$ is denoted $\mathcal{C}(\vec{c}, \vec{b})$. For

$$\bar{p} \in \mathbb{B} = \{\langle p_0, p_1, p_2, \ldots\rangle \mid p_1 \ge p_2 \ge \ldots \ge 0,\ p_0 \ge 0\ \&\ \sum_i p_i = 1\},$$

the probability function $u^{\bar{p},L}$ is defined on a state description $\Phi(b_1, \ldots, b_m)$ by

$$u^{\bar{p},L}(\Phi(b_1,\ldots,b_m)) = \sum_{\substack{\langle c_1,\ldots,c_m\rangle \in \mathbb{N}^m \\ \Phi \in \mathcal{C}(\vec{c},\vec{b})}} |\mathcal{C}(\vec{c},\vec{b})|^{-1} \prod_{i=1}^{m} p_{c_i}. \tag{2}$$

The probability function $u^{\bar{p},L}$ extends uniquely to SL to satisfy Ex+Sx and furthermore it is homogeneous (see [11] or [15, chapter 29]) when

$$\bar{p} \in \mathbb{B}_\infty = \{\langle p_0, p_1, p_2, \ldots\rangle \in \mathbb{B} \mid p_0 > 0 \text{ or } p_i > 0 \text{ for all } i > 0\}.$$

Indeed the $u^{\bar{p},L}$ for $\bar{p} \in \mathbb{B}_\infty$ are the building blocks for all homogeneous probability functions:

Theorem 1. *Let w be a homogeneous probability function on SL. Then there is a measure*[5] *μ on the Borel subsets of $\mathbb{B}_\infty$ such that*

$$w = \int_{\mathbb{B}_\infty} u^{\bar{p},L}\, d\mu(\bar{p}).$$

Conversely given such a measure μ, w defined as above is a homogeneous probability function on SL.

[5] All measures will be taken to be countably additive and normalized.

For $t \in \mathbb{N}^+ = \{1, 2, 3, \ldots\}$, a probability function w on SL is t-heterogeneous if it satisfies Sx and

$$\lim_{n \to \infty} w \left(\bigvee_{|\mathcal{S}(\Phi(a_1,\ldots,a_n))|=t} \Phi(a_1, \ldots, a_n) \right) = 1. \tag{3}$$

Again, the disjunction is taken over all state descriptions of L for constants $a_1, \ldots, a_n$ with spectrum size t.

For $t \in \mathbb{N}^+$, let $\mathbb{N}_t = \{1, 2, \ldots, t\}$ and let

$$\bar{p} \in \mathbb{B}_t = \{\langle p_0, p_1, p_2, \ldots, \rangle \in \mathbb{B} \,|\, p_0 = 0 \ \& \ p_t > 0 = p_{t+1}\}.$$

The probability function $v_t^{\bar{p},L}$ is defined on state descriptions $\Phi(b_1, \ldots, b_m)$ in terms of vectors $\vec{c} \in (\mathbb{N}_t)^m$ and a function $\mathcal{G}(\vec{c}, \Phi)$. For a fixed $\vec{c}$, if Φ is not consistent with $\vec{c}$, i.e. if for some $1 \leq s, d \leq m$, $c_s = c_d$ but $b_s \not\sim_\Phi b_d$, then $\mathcal{G}(\vec{c}, \Phi)$ is zero. Otherwise let $c_{g_1}, \ldots, c_{g_r}$ be the first instances of each distinct entry in $\vec{c}$ and let Φ' be the state description for $b_{g_1}, \ldots, b_{g_r}$ entailed by Φ. Then $\mathcal{G}(\vec{c}, \Phi)$ takes the value

$$\frac{\text{Number of extensions of } \Phi' \text{ to a state description with spectrum } \mathbf{1}_t}{\text{Number of state descriptions with spectrum } \mathbf{1}_t}$$

where $\mathbf{1}_t$ is the spectrum of t 1's (that is, the spectrum of a state description for t constants in which all the constants are distinguishable from each other). Now set

$$v_t^{\bar{p},L}(\Phi(b_1, \ldots, b_m)) = \sum_{\langle c_1,\ldots,c_m \rangle \in (\mathbb{N}_t)^m} \mathcal{G}(\vec{c}, \Phi) \prod_{i=1}^{m} p_{c_i}. \tag{4}$$

As shown in, for example, [11] or [15, chapter 30], with this definition $v_t^{\bar{p},L}$ extends uniquely to a probability function satisfying Ex+Sx and furthermore is t-heterogeneous.

Moreover the $v_t^{\bar{p},L}$ for $\bar{p} \in \mathbb{B}_t$ are the building blocks for all t-heterogeneous probability functions:

Theorem 2. *Let w be a t-heterogeneous probability function on SL. Then there is a measure μ on the Borel subsets of $\mathbb{B}_t$ such that*

$$w = \int_{\mathbb{B}_t} v_t^{\bar{p},L} \, d\mu(\bar{p}).$$

Conversely given such a measure μ, w defined as above is a t-heterogeneous probability function on SL.

Probability functions satisfying Sx are a mixture of heterogeneous and homogeneous probability functions as the following theorem, see [8], [12], or [15, Theorem 30.2], explains.

The Ladder Theorem 3. *Any probability function w satisfying Sx can be expressed in the form*

$$w = \eta_0 w^{[0]} + \sum_{t=1}^{\infty} \eta_t w^{[t]}$$

where the $\eta_i \geq 0, \sum_i \eta_i = 1$, $w^{[0]}$ is homogeneous and $w^{[t]}$ is t-heterogeneous for $t > 0$.

Dolly's Constant Principle

The following principle, introduced in [7] (as simply Dolly's Principle, DP), for a probability function w on a language L is one attempt to capture the sort of possibility-by-analogy argument illustrated in the introduction to this paper:

Dolly's Constant Principle, DCP

Let $\theta \in SL$ and let θ^* be the result of replacing some constant symbols a_i (everywhere) in θ by $a_{\sigma(i)}$. Then if $w(\theta^*) > 0$, $w(\theta) > 0$.

Notice that unlike Ex the replacing constants may already appear in θ. So for example if

$$\theta = \forall x \, (R_1(a_1, a_2, x) \to R_2(x, a_1, a_3)))$$

we could have

$$\theta^* = \forall x \, (R_1(a_3, a_2, x) \to R_2(x, a_3, a_3))$$

by replacing a_1 by a_3.

We shall denote by QF-DCP the same principle but with θ restricted to being quantifier free.

DCP then is a regularity principle in that it says only that under certain conditions a probability must be non-zero without making any claims about magnitude. Notice that by repeated application DReg is equivalent to the principle where we replace just one constant.

The following result is proved in [7]:

Lemma 4. *If θ^* is consistent then so is θ.*

Consequently,

Corollary 5. *SReg implies DCP and Reg implies QF-DCP.*

The main result on DCP given in [7] is:

Theorem 6. *For L a unary language*[6] *Ex implies DP.*

For not purely unary languages QF-DP no longer follows from Ex, even when $q = 1$ and $r_1 = 2$. To see this let w be a probability function on the language L containing one binary predicate R such that

$$w(R(a_i, a_j)) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Then w satisfies Ex but not QF-DCP since $w(R(a_1, a_1)) = 1$ but $w(R(a_1, a_2)) = 0$.

Spectrum Exchangeability however does suffice for QF-DCP:

Theorem 7. *Sx implies QF-DCP.*

Proof. First suppose that w satisfies Sx and let $\theta(a_1, \ldots, a_{n+1}) \in QFSL$. If $w(\theta(a_1, \ldots, a_n, a_1) > 0$ then there must be some state description $\Theta(a_1, \ldots, a_{n+1})$ such that $\Theta(a_1, \ldots, a_{n+1}) \models \theta(a_1, \ldots, a_n)$ and $w(\Phi(a_1, a_2, \ldots, a_n)) > 0$ where

$$\Phi(a_1, a_2, \ldots, a_n) = \Theta(a_1, \ldots, a_n, a_1)).$$

But for $\Theta(a_1, \ldots, a_n, a_1)$ to even be consistent requires that $a_1 \sim_\Theta a_{n+1}$. Hence Φ is[7] also a state description and if the equivalence classes of $\sim_\Phi$ are $E_1, E_2, \ldots, E_r$ with $a_1 \in E_1$ then the equivalence classes of $\sim_\Theta$ are $\{a_{n+1}\} \cup E_1, E_2, \ldots, E_r$.

Now suppose that $\bar{p} \in \mathbb{B}$ and $u^{\bar{p},L}(\Phi) > 0$. Then from the definition of $u^{\bar{p},L}$ there must be some $\vec{c} = \langle c_1, c_2, \ldots, c_n \rangle \in (\mathbb{N}^+)^n$ compatible with Φ such that $\prod_{i=1}^n p_{c_i} > 0$. But then $\langle c_1, c_2, \ldots, c_n, c_1 \rangle$ must be compatible with Θ and $(\prod_{i=1}^n p_{c_i}) \cdot p_{c_1} > 0$ so $u^{\bar{p},L}(\Theta) > 0$.

Exactly similarly we can show that if $\bar{p} \in \mathbb{B}_t$ and $v_t^{\bar{p},L}(\Phi) > 0$ then $v_t^{\bar{p},L}(\Theta) > 0$. Finally since by Theorems 1, 2, 3, w must be a convex mixture of the $u^{\bar{p},L}$ and $v_t^{\bar{p},L}$ for $t \in \mathbb{N}^+$ we can conclude that we must also have $w(\Phi) > 0$, as required. □

We now turn to the case for general sentences rather than just quantifier free. First an observation. Let $\rho(z_1, z_2)$ be the formula of L

$$\bigwedge_{e=1}^{q} \bigwedge_{f=1}^{r_e} \forall x_1, \ldots, x_{f-1}, x_{f+1}, \ldots, x_{r_e}$$
$$(R_e(x_1, \ldots, x_{f-1}, z_1, x_{f+1}, \ldots, x_{r_e}) \leftrightarrow R_e(x_1, \ldots, x_{f-1}, z_2, x_{f+1}, \ldots, x_{r_e}))$$

[6] I.e. all the R_i are unary.

[7] Up to logical equivalence.

which expresses that z_1 and z_2 are permanently indistinguishable from each other. So $\rho(z_1, z_1)$ is a tautology.

Now suppose that $\theta(a_1, a_2, \ldots, a_n, a_{n+1}) \in SL$ was such that

$$w(\theta(a_1, a_2, \ldots, a_n, a_1)) > 0.$$

If we also have that

$$w(\theta(a_1, a_2, \ldots, a_n, a_1) \wedge \rho(a_1, a_{n+1})) > 0$$

then, since

$$\theta(a_1, a_2, \ldots, a_n, a_1) \wedge \rho(a_1, a_{n+1}) \models \theta(a_1, a_2, \ldots, a_n, a_{n+1}),$$

we would have the requirement of DCP in this case, i.e.

$$w(\theta(a_1, a_2, \ldots, a_n, a_{n+1})) > 0.$$

Conversely suppose that we have DCP. Then for any $\phi(a_1, a_2, \ldots, a_n) \in SL$ with $w(\phi) > 0$ we must also have that

$$w(\phi(a_1, a_2, \ldots, a_n) \wedge \rho(a_1, a_1)) = w(\phi(a_1, a_2, \ldots, a_n)) > 0$$

so by DCP

$$w(\phi(a_1, a_2, \ldots, a_n) \wedge \rho(a_1, a_{n+1})) > 0.$$

In summary then we have shown that DCP is equivalent to the assertion that:

Whenever $w(\phi(a_1, a_2, \ldots, a_n)) > 0$ *then* $w(\phi(a_1, a_2, \ldots, a_n) \wedge \rho(a_1, a_{n+1})) > 0.$

Now suppose that $v_t^{\bar{p},L}(\phi(a_1, a_2, \ldots, a_n)) > 0$. Since $v_t^{\bar{p},L}$ is t-heterogeneous

$$\lim_{m \to \infty} w\left(\bigvee_{|\mathcal{S}(\Theta(a_1,\ldots,a_m))|=t} \Theta(a_1, \ldots, a_m) \right) = 1. \tag{5}$$

Hence there must be some state description $\Theta(a_1, \ldots, a_m)$ with $m \geq n$ such that $|\mathcal{S}(\Theta(a_1, \ldots, a_m))| = t$ and

$$v_t^{\bar{p},L}(\phi(a_1, \ldots, a_n) \wedge \Theta(a_1, \ldots, a_m)) > 0. \tag{6}$$

Since $|\mathcal{S}(\Theta)| = t$ and $v_t^{\bar{p},L}$ is t-heterogeneous results in [6] (specifically Lemmas 4 and 8 of that paper) give that for any $\psi(a_1, \ldots, a_m) \in SL$ one of the following must hold:

$$v_t^{\bar{p},L}(\psi(a_1, \ldots, a_m) \wedge \Theta(a_1, \ldots, a_m)) = v_t^{\bar{p},L}(\Theta(a_1, \ldots, a_m)),$$

$$v_t^{\bar{p},L}(\neg\psi(a_1,\dots,a_m)\wedge\Theta(a_1,\dots,a_m)) = v_t^{\bar{p},L}(\Theta(a_1,\dots,a_m)).$$

Hence it must be that

$$v_t^{\bar{p},L}(\phi(a_1,\dots,a_n)\wedge\Theta(a_1,\dots,a_m)) = v_t^{\bar{p},L}(\Theta(a_1,\dots,a_m)). \tag{7}$$

Since $v_t^{\bar{p},L}(\Theta(a_1,\dots,a_m)) > 0$ we can find $\langle c_1, c_2, \dots, c_m\rangle \in \{1,2,\dots,t\}^m$ compatible with $\Theta(a_1,\dots,a_m)$ and such that $\mathcal{G}(\langle c_1,c_2,\dots,c_m\rangle,\Theta) > 0$. Let $\Theta^+(a_1,\dots,a_{m+1})$ be the state description extending Θ in which $a_{m+1}\sim_{\Theta^+} a_1$. Then $\langle c_1,c_2,\dots,c_m,c_1\rangle$ is compatible with Θ^+ and $\mathcal{G}(\langle c_1,c_2,\dots,c_m,c_1\rangle,\Theta^+) = \mathcal{G}(\langle c_1,c_2,\dots,c_m\rangle,\Theta)$, giving

$$0 < \mathcal{G}(\langle c_1,c_2,\dots,c_m,c_1\rangle,\Theta^+)\left(\prod_{i=1}^m p_{c_i}\right)\cdot p_{c_1} \le w(\Theta^+).$$

From the above we must have

$$v_t^{\bar{p},L}(\phi(a_1,\dots,a_n)\wedge\Theta^+(a_1,\dots,a_{m+1})) = v_t^{\bar{p},L}(\Theta^+(a_1,\dots,a_{m+1})). \tag{8}$$

We must similarly have that

$$v_t^{\bar{p},L}(\rho(a_1,a_{m+1})\wedge\Theta^+(a_1,\dots,a_{m+1})) = v_t^{\bar{p},L}(\Theta^+(a_1,\dots,a_{m+1})). \tag{9}$$

To see this let $1\le e\le q$, $1\le f\le r_e$ and $m+1, i_1, i_2, \dots, i_{f-1}, i_{f+1},\dots, i_{r_e}\le h$ and let $\Psi(a_1,a_2,\dots,a_h)$ be a state description extending $\Theta^+(a_1,\dots,a_{m+1})$ with $v_t^{\bar{p},L}(\Psi)>0$. Then since $v_t^{\bar{p},L}$ is t-heterogeneous $\mathcal{S}(\Psi)=t$ and we must still have $a_1\sim_\Psi a_{m+1}$. As above one of the following must hold

$$v_t^{\bar{p},L}(R_e(a_{i_1},\dots,a_{i_{f-1}},a_1,a_{i_f},\dots,a_{i_{r_e}})\wedge\Psi(a_1,\dots,a_h)) = v_t^{\bar{p},L}(\Psi(a_1,\dots,a_h)),$$

$$v_t^{\bar{p},L}(\neg R_e(a_{i_1},\dots,a_{i_{f-1}},a_1,a_{i_f},\dots,a_{i_{r_e}})\wedge\Psi(a_1,\dots,a_h)) = v_t^{\bar{p},L}(\Psi(a_1,\dots,a_h)).$$

Furthermore since $a_1\sim_{\Theta^+} a_{m+1}$ the corresponding one must hold with the designated a_1 replaced by a_{m+1}. Hence

$$v_t^{\bar{p},L}\big([R_e(a_{i_1},\dots,a_{i_{f-1}},a_1,a_{i_f},\dots,a_{i_{r_e}})\leftrightarrow R_e(a_{i_1},\dots,a_{i_{f-1}},a_{m+1},a_{i_f},\dots,a_{i_{r_e}})]$$
$$\wedge\ \Psi(a_1,\dots,a_h)\big) \;=\; v_t^{\bar{p},L}(\Psi(a_1,\dots,a_h)).$$

By summing over all such Ψ and then over all e, f and $a_{i_1},\dots,a_{i_{r_e}}$ (9) now follows.

Putting (8), (9) together now yields as required that

$$v_t^{\bar{p},L}(\phi(a_1,\dots,a_n)\wedge\rho(a_1,a_{m+1})>0$$

and of course by Ex we can replace a_{m+1} here by a_{n+1}.

The above considerations have shown that the $v_t^{\bar{p},L}$ satisfy DCP. Hence, using Theorem 2:

Theorem 8. *If w satisfying Sx is t-heterogeneous for some t, or more generally is a convex sum of heterogeneous probability functions, then w satisfies DCP.*

Corollary 9. *If $\bar{p} \in \mathbb{B}$ is such that $p_0 = 0 = p_m$ for some $m > 0$ then $u^{\bar{p},L}$ satisfies DCP.*

Proof. By Theorem 3 such a $u^{\bar{p},L}$ is a convex sum of heterogeneous probability functions. □

Theorem 10. *If $\bar{p} \in \mathbb{B}$ is such that $p_m > 0$ for all $m > 0$ then $u^{\bar{p},L}$ satisfies DCP.*

Proof. If L is purely unary then any probability function on SL is a convex sum of heterogeneous probability functions so the result follows from Theorem 8. So assume that R_1 is non-unary. Let $\bar{p}$ be as in the theorem. For $\vec{S} = S_1, S_2, \dots, S_h$ a partition of $\{1, 2, \dots, m\}$ let $\upsilon^{\vec{S}}(y_1, \dots, y_m)$ be

$$\bigwedge_{g=1}^{h} \bigwedge_{i,j \in S_g} \left(\rho(y_i, y_j) \wedge \bigwedge_{\substack{1 \le u \le m \\ u \notin S_g}} \neg\rho(y_i, y_u) \right), \tag{10}$$

and for $\Theta(a_1, \dots, a_m)$ a state description let

$$\Theta^{\vec{S}}(a_1, \dots, a_m) = \Theta(a_1, \dots, a_m) \wedge \upsilon^{\vec{S}}(a_1, \dots, a_m).$$

Then (as usual up to logical equivalence),

$$\bigvee_{\vec{S}} \Theta^{\vec{S}}(a_1, \dots, a_m) = \Theta(a_1, \dots, a_m)$$

where the disjunction is over $\vec{S}$ such that whenever i, j are in the same S_g then $a_i \sim_\Theta a_j$.

Furthermore in the case of such an $\vec{S}$, $u^{\bar{p},L}(\Theta^{\vec{S}}(a_1, \dots, a_m)) > 0$. To see this let $\Phi(a_1, \dots, a_t)$ be a state description extending $\Theta(a_1, \dots, a_m)$ with spectrum $\vec{T} = S_1, \dots, S_h, \{m+1\}, \{m+2\}, \dots, \{t\}$ (recall that L is not purely unary so this is possible for a suitably large choice of t). So

$$\Phi(a_1, \dots, a_t) \models \Theta(a_1, \dots, a_m) \tag{11}$$

and

$$\Phi(a_1, \dots, a_t) \models \bigwedge_{g=1}^{h} \bigwedge_{\substack{i \in S_g \\ u \notin S_g}} \neg\rho(a_i, a_u) \wedge \bigwedge_{m+1 \le i < j \le t} \neg\rho(a_i, a_j). \tag{12}$$

Now let $1 \leq c_1, \ldots, c_t \leq t$ be such that $\vec{c} = \langle c_1, \ldots, c_t \rangle$ is consistent with $\Phi(\vec{a})$, where $\vec{a} = \langle a_1, \ldots, a_t \rangle$, and $c_i = c_j$ whenever i, j are in the same S_g. Then since $p_n > 0$ for all $n > 0$,

$$u^{\bar{p},L}(\Theta(a_1, \ldots, a_m)) \geq u^{\bar{p},L}(\Phi(\vec{a})) \geq \mathcal{C}(\vec{c}, \vec{a})^{-1} \prod_{i=1}^{t} p_{c_i} > 0.$$

Furthermore for $k \geq t$, $\vec{b} = \langle a_1, a_2, \ldots, a_k \rangle$

$$\begin{aligned} u^{\bar{p},L}\left(\bigvee_{\Psi(\vec{b}) \models \Phi(\vec{a})} \Psi(\vec{b}) \right) &\geq \sum_{\Psi(\vec{b}) \models \Phi(\vec{a})} \sum_{\vec{d}} \mathcal{C}(\vec{d}, \vec{b})^{-1} \prod_{i=1}^{k} p_{d_i} \\ &= \mathcal{C}(\vec{c}, \vec{a})^{-1} \prod_{i=1}^{t} p_{c_i} \end{aligned}$$

where the $\vec{d}$ extend $\vec{c}$ and $\vec{d}$ is consistent with $\Psi(\vec{b})$. Whenever $c_i = c_j$ each such $\Psi(\vec{b})$ for which there is a consistent $\vec{d}$ must logically imply all instances of $\rho(a_i, a_j)$, i.e. of

$$\bigwedge_{e=1}^{q} \bigwedge_{f=1}^{r_e} \forall x_1, \ldots, x_{f-1}, x_{f+1}, \ldots, x_{r_e}$$
$$(R_e(x_1, \ldots, x_{f-1}, a_i, x_{f+1}, \ldots, x_{r_e}) \leftrightarrow R_e(x_1, \ldots, x_{f-1}, a_j, x_{f+1}, \ldots, x_{r_e}))$$

when the x_n come from $a_1, \ldots, a_k$. Hence if $\chi_k(a_1, \ldots, a_k)$ is the conjunction of all such instances for all $1 \leq i, j \leq t$ with $c_i = c_j$ then

$$\chi_{k+1}(a_1, \ldots, a_{k+1}) \models \chi_k(a_1, \ldots, a_k)$$

and

$$u^{\bar{p},L}(\Phi(\vec{a}) \wedge \chi_k(a_1, \ldots, a_k)) \geq \mathcal{C}(\vec{c}, \vec{a})^{-1} \prod_{i=1}^{t} p_{c_i} > 0.$$

Taking the limit of this left hand side as $k \to \infty$ gives that

$$u^{\bar{p},L}(\Phi(\vec{a}) \wedge \bigwedge_{g=1}^{h} \bigwedge_{i,j \in S_g} \rho(a_i, a_j)) \geq \mathcal{C}(\vec{c}, \vec{a})^{-1} \prod_{i=1}^{t} p_{c_i} > 0$$

and combining this with (11), (12) gives that

$$u^{\bar{p},L}(\Theta^{\vec{S}}(a_1, \ldots, a_m)) > 0.$$

Now suppose that $\theta(a_1,\dots,a_m)\in SL$ and $u^{\bar{p},L}(\theta(a_1,\dots,a_m))>0$. Then for some $\Theta(a_1,\dots,a_m)$, and hence $\Theta^{\vec{S}}(a_1,\dots,a_m)$, we must have

$$u^{\bar{p},L}(\theta(a_1,\dots,a_m))\wedge\Theta^{\vec{S}}(a_1,\dots,a_m))>0.$$

By the Lemma 15 and the proof of Lemma 16 in [6] one of

$$u^{\bar{p},L}(\theta(a_1,\dots,a_m))\wedge\Theta^{\vec{S}}(a_1,\dots,a_m))=u^{\bar{p},L}(\Theta^{\vec{S}}(a_1,\dots,a_m)), \tag{13}$$

$$u^{\bar{p},L}(\neg\theta(a_1,\dots,a_m))\wedge\Theta^{\vec{S}}(a_1,\dots,a_m))=u^{\bar{p},L}(\Theta^{\vec{S}}(a_1,\dots,a_m)),$$

must hold, so clearly it is the former. Without loss of generality let $1\in S_1$ and let $\Psi(a_1,\dots,a_{m+1})$ be the state description extending $\Theta(a_1,\dots,a_m)$ and such that $a_1\sim_\Psi a_{m+1}$. Then as above

$$\Psi^{\vec{U}}(a_1,\dots,a_{m+1})\models\Theta^{\vec{S}}(a_1,\dots,a_m),$$

$$w(\Psi^{\vec{U}}(a_1,\dots,a_{m+1}))>0,$$

where $\vec{U}=S_1\cup\{m+1\},S_2,S_3,\dots,S_h$. Hence with (13),

$$w(\theta(a_1,\dots,a_m)\wedge\Psi^{\vec{U}}(a_1,\dots,a_{m+1}))>0$$

and since

$$\Psi^{\vec{U}}(a_1,\dots,a_{m+1})\models\rho(a_1,a_{m+1})$$

the required result follows. □

However for L not purely unary when $p_0>0$ and $p_m>p_{m+1}=0$ for some m then $u^{\bar{p},L}$ fails DCP. A suitable counter-example here is the sentence $\chi(a_1,a_2,\dots,a_{2m},a_{2m+1})$ given by

$$\bigwedge_{i=1}^{m+1}\rho(a_{2i-1},a_{2i})\wedge\bigwedge_{1\le i<j\le m+1}\neg\rho(a_{2i-1},a_{2j-1})$$

where we have

(i) $u^{\bar{p},L}(\chi(a_1,a_2,\dots,a_{2m},a_{2m+1},a_{2m+1}))>0$, but

(ii) $u^{\bar{p},L}(\chi(a_1,a_2,\dots,a_{2m},a_{2m+1},a_{2m+2})=0$.

To show (i) let $c_{2i-1}=c_{2i}=i$ for $i=1,2,\dots,m$ and $c_{2m+1}=0$. Let $\Theta(a_1,a_2,\dots,a_{2m+1})$ be a state description with $a_{2i-1}\sim_\Theta a_{2i}$ for $i=1,2,\dots,m$ and $a_{2i-1}\nsim_\Theta a_{2j-1}$ for $1\le i<j\le m+1$. Then Θ is consistent with $\vec{c}$ and

$$\Theta(a_1,\dots,a_{2m+1})\models\bigwedge_{1\le i<j\le m+1}\neg\rho(a_{2i-1},a_{2j-1}). \tag{14}$$

Furthermore

$$u^{\bar{p},L}(\Theta(a_1,\dots,a_{2m+1}) \wedge \bigwedge_{i=1}^{m+1} \rho(a_{2i-1},a_{2i})) \geq \mathcal{C}(\vec{c},\vec{a})^{-1} \prod_{r=1}^{2m+1} p_{c_r} \ > \ 0. \tag{15}$$

To see this notice that if $\Phi(a_1,\dots,a_n)$ is a state description extending $\Theta(a_1,\dots,a_{2m+1})$ and $\vec{d} \in \{0,1,2,\dots,m\}^n$ extends $\vec{c}$ and is consistent with Φ then Φ logically implies the conjunction, $\xi_n(a_1,\dots,a_{2m+1})$ say, of instances of the

$$\bigwedge_{i=1}^{m} \bigwedge_{e=1}^{q} \bigwedge_{f=1}^{r_e} \forall x_1,\dots,x_{f-1},x_{f+1},\dots,x_{r_e}$$
$$(R_e(x_1,\dots,x_{f-1},a_{2i-1},x_{f+1},\dots,x_{r_e}) \leftrightarrow R_e(x_1,\dots,x_{f-1},a_{2i},x_{f+1},\dots,x_{r_e}))$$

when the x_j are given values from $a_1,a_2,\dots,a_n$.

Now of those state descriptions $\Phi(a_1,\dots,a_n)$ consistent with $\vec{d}$ extending $\vec{c}$ a fraction of at least $\mathcal{C}(\vec{c},\vec{a})^{-1}$ of them will extend $\Theta(a_1,\dots,a_{2m+1})$, so

$$u^{\bar{p},L}(\Theta(a_1,a_2,\dots,a_{2m+1}) \wedge \xi_n(a_1,\dots,a_n)) \ \geq$$
$$\geq \sum_{\vec{d}} \sum_{\substack{\phi(\vec{b}) \models \Theta(\vec{a}) \\ \Phi \in C(\vec{d})}} \mathcal{C}(\vec{d},\vec{b})^{-1} \prod_{r=1}^{n} p_{d_r} \ \geq \ \sum_{\vec{d}} \mathcal{C}(\vec{c},\vec{a})^{-1} \prod_{r=1}^{n} p_{d_r} \ = \ \mathcal{C}(\vec{c},\vec{a})^{-1} \prod_{r=1}^{2m+1} p_{c_r} \ > \ 0.$$

where $\vec{b} = \langle a_1,\dots,a_n \rangle$, the $\vec{d}$ range over elements of $\{0,1,2,\dots,m\}^n$ extending $\vec{c}$ and $C(\vec{d})$ is the set of state descriptions for $\vec{b}$ which are consistent with $\vec{d}$. Since this last term is independent of n, taking the limit as $n \to \infty$ yields (15).

Putting (14) and (15) together now gives (i)

Turning now to (ii) suppose on the contrary that $\Theta(a_1,\dots,a_{2m+2})$ was a state description such that

$$u^{\bar{p},L}(\chi(a_1,a_2,\dots,a_{2m},a_{2m+1}) \wedge \Theta(a_1,\dots,a_{2m+2})) > 0. \tag{16}$$

Then since the $\neg\rho(a_{2i-1},a_{2j-1})$ are (up to logical equivalence) Σ_1, by the generalization [15, Lemma 3.8] of condition (P3), there is a state description $\Phi(a_1,\dots,a_k)$ extending $\Theta(a_1,\dots,a_{2m+2})$ such that for $1 \leq i < j \leq m+1$

$$a_{2i-1} \not\sim_\Phi a_{2j-1}, \tag{17}$$

and so forcing

$$\Phi(a_1,\dots,a_k) \models \bigwedge_{1 \leq i < j \leq m+1} \neg\rho(a_{2i-1},a_{2j-1}),$$

and

$$u^{\bar{p},L}(\Phi(a_1,\ldots,a_k)\wedge\bigwedge_{i=1}^{m+1}\rho(a_{2i-1},a_{2i}))>0.$$

Now let n be large compared with k and let Y be the set of $\vec{c}\in\{0,1,\ldots,m\}^n$ such that the initial segment $\langle c_1,\ldots,c_k\rangle$ is consistent with Φ. From (17) it must already be the case that the non-zero values amongst the c_{2i-1} for $i=1,2,\ldots,m+1$ must all be different. So for some $1\leq i\leq m+1$, we must have $c_{2i-1}=0$. Let $\iota(\vec{c})$ pick out such an i.

Since n is large compared with k,

$$\sum_{\vec{d}}\prod_{r=1}^{n}p_{d_r}<3^{-1}u^{\bar{p},L}(\Phi\wedge\chi),\tag{18}$$

where the sum is over $\vec{d}\in Y$ which contain less than $k+np_0/2$ zeros.

Now suppose instead that $\vec{e}\in Y$ has at least $k+np_0/2$ zeros. Then the proportion of state descriptions Ψ for $\vec{b}=\langle a_1,\ldots,a_n\rangle$ extending Φ which are consistent with $\vec{e}$ and such that $a_{2\iota(\vec{e})-1}\sim_\Psi a_{2\iota(\vec{e})}$ will be less than some λ_n which tends to zero as $n\to\infty$. To see this consider constructing Ψ from Φ by extending to each of the constants $a_{k+1},a_{k+2},\ldots,a_n$ one step at a time and remaining consistency with $\vec{e}$. Every time we have $e_s=0$ there will be a free choice of extension and, because the language is not purely unary, at least half of these will destroy the indistinguishability of $a_{2\iota(\vec{e})-1}$ and $a_{2\iota(\vec{e})}$. In summary then we have that

$$\sum_{\Psi_{\vec{e}}}\mathcal{C}(\vec{e},\vec{b})^{-1}<\lambda_n\sum_{\Gamma_{\vec{e}}}\mathcal{C}(\vec{e},\vec{b})^{-1}$$

where the $\Gamma_{\vec{e}}$ range over state descriptions for $\vec{b}=\langle a_1,\ldots,a_n\rangle$ consistent with $\vec{e}$ and extending Φ and the $\Psi_{\vec{e}}$ range over the subset of these $\Gamma_{\vec{e}}$ for which $a_{2\iota(\vec{e})-1}\sim_{\Psi_{\vec{e}}} a_{2\iota(\vec{e})}$.

Hence for $\vec{d},\vec{e},\Psi,\Psi_{\vec{d}},\Psi_{\vec{e}},\Gamma_{\vec{e}}$ constrained as in the foregoing and $C(\Psi)$ the set of

$\vec{c} \in Y$ consistent with Ψ,

$$\begin{aligned}
u^{\bar{p},L}(\Phi \wedge \chi) &= \sum_{\Psi(\vec{b}) \models \Phi(\vec{a})} u^{\bar{p},L}(\Psi \wedge \chi) \;\leq \sum_{\substack{\Psi(\vec{b}) \models \Phi(\vec{a}) \\ \Psi(\vec{b}) \not\models \neg\chi}} u^{\bar{p},L}(\Psi) \\
&= \sum_{\substack{\Psi(\vec{b}) \models \Phi(\vec{a}) \\ \Psi(\vec{b}) \not\models \neg\chi}} \sum_{\vec{c} \in C(\Psi)} \mathcal{C}(\vec{c},\vec{b})^{-1} \prod_{r=1}^{n} p_{c_r} \\
&= \sum_{\vec{c} \in Y} \sum_{\substack{\Psi(\vec{b}) \models \Phi(\vec{a}), \vec{c} \in C(\Psi) \\ \Psi(\vec{b}) \not\models \neg\chi}} \mathcal{C}(\vec{c},\vec{b})^{-1} \prod_{r=1}^{n} p_{c_r} \\
&\leq \sum_{\vec{d}} \sum_{\Psi_{\vec{d}}} \mathcal{C}(\vec{d},\vec{b})^{-1} \prod_{r=1}^{n} p_{d_r} + \sum_{\vec{e}} \sum_{\substack{\Gamma_{\vec{e}} \\ \Gamma_{\vec{e}} \not\models \neg\chi}} \mathcal{C}(\vec{e},\vec{b})^{-1} \prod_{r=1}^{n} p_{e_r} \\
&\leq \sum_{\vec{d}} \prod_{r=1}^{n} p_{d_r} + \sum_{\vec{e}} \sum_{\Psi_{\vec{e}}} \mathcal{C}(\vec{e},\vec{b})^{-1} \prod_{r=1}^{n} p_{e_r}, \\
&\quad \text{since } \Psi \models \neg\chi \text{ if some } a_{2i-1} \not\sim_\Psi a_{2i} \text{ with } 1 \leq i \leq m+1, \\
&\leq 3^{-1} u^{\bar{p},L}(\Phi \wedge \chi) + \lambda_n \sum_{\Psi(\vec{b}) \models \Phi(\vec{a})} u^{\bar{p},L}(\Psi) \\
&\leq 3^{-1} u^{\bar{p},L}(\Phi \wedge \chi) + \lambda_n,
\end{aligned}$$

which provides the required contradiction.

Dolly's Relation Principle, DRP

In the previous section we gave a possible formulation of the informal analogy principle suggested in the introduction where constants were identified. An alternative is to consider treating relations of the same arity as identical, which in a similar fashion suggests the following formalization:

Dolly's Relation Principle, DRP

Let $\theta \in SL$ and let $\theta^\dagger$ be the result of replacing some relations symbols R_j (everywhere) in θ by $R_{\sigma(j)}$ of the same arity as R_j. Then if $w(\theta^\dagger) > 0$, $w(\theta) > 0$.

The replacing relation symbols may already appear in θ. So for example if

$$\theta = \forall x \, (R_1(a_1, a_2, x) \to R_2(x, a_1, a_3)))$$

we could have

$$\theta^\dagger = \forall x\, (R_2(a_1, a_2, x) \to R_2(x, a_1, a_3)).$$

We shall denote by QF-DRP the same principle but with θ restricted to being quantifier free.

Again we have:

Lemma 11. *If $\theta^\dagger$ is consistent then so is θ.*

Consequently,

Corollary 12. *SReg implies DRP and Reg implies QF-DRP.*

Unlike QF-DCP the principle QF-DRP does not follow simply from Ex even for purely unary languages.[8] To see this let L be the unary language with relation symbols R_1, R_2, R_3 and list the *atoms* of L,

$$R_1^{\epsilon_1}(x) \wedge R_2^{\epsilon_2}(x) \wedge R_3^{\epsilon_3}(x)$$

where the $\epsilon_i \in \{0,1\}$ and $R_j^1 = R_j, R_j^0 = \neg R_j$, as $\alpha_i(x)$, $i = 1, \ldots, 8$. Define the function κ on these atoms by

$$\kappa(R_1^{\epsilon_1}(x) \wedge R_2^{\epsilon_2}(x) \wedge R_3^{\epsilon_3}(x)) = (-1)^{\sum_i \epsilon_i}$$

and the probability function w for this language by[9]

$$w\left(\bigwedge_{i=1}^m \alpha_{h_i}(b_i)\right) = \begin{cases} 2^{-2m-1} & \text{if all the } \kappa(\alpha_{h_i}(x)) \text{ are equal,} \\ 0 & \text{otherwise.} \end{cases}$$

Then w satisfies Ex[10] but not QF-DRP since if $\theta(a_1, a_2)$ is

$$R_1(a_1) \wedge R_2(a_1) \wedge R_3(a_1) \wedge R_1(a_2) \wedge \neg R_2(a_2) \wedge R_3(a_2)$$

then $w(\theta(a_1, a_2)) = 0$ whilst if we form $\theta^\dagger(a_1, a_2)$ by replacing R_3 by R_1 in $\theta(a_1, a_2)$ then

$$\begin{aligned} w(\theta^\dagger(a_1, a_2)) &= w(R_1(a_1) \wedge R_2(a_1) \wedge R_1(a_2) \wedge \neg R_2(a_2)) \\ &\geq w(R_1(a_1) \wedge R_2(a_1) \wedge R_3(a_1) \wedge R_1(a_2) \wedge \neg R_2(a_2) \wedge \neg R_3(a_2)) \\ &= 2^{-7} \ > \ 0. \end{aligned}$$

Once we have Sx however QF-DRP does follow:

[8] In fact not even from Ex with the additional principles of Strong Negation and Predicate Exchangeability as the forthcoming example shows.

[9] By a Theorem of Gaifman [5] a probability function is determined by its values on state descriptions.

[10] And also the Principles of Predicate Exchangeability and Strong Negation.

Theorem 13. *Sx implies QF-DRP.*

Proof. Let $\theta(a_1, \ldots, a_n) \in QFSL$ and let $\theta^\dagger(a_1, \ldots, a_n)$ be the result of replacing each R_2 in θ by R_1 (which we assume has the same arity). Suppose that $w(\theta^\dagger) > 0$. Then there must be a state description $\Theta \models \theta$ such that $w(\Theta^\dagger) > 0$. This being the case $\Theta^\dagger$ must be consistent so in fact we must have that

$$\Theta \models R_1(\vec{b}) \leftrightarrow R_2(\vec{b})$$

for any $\vec{b}$ of the right length from $\{a_1, \ldots, a_n\}$. Since $w(\Theta^\dagger) > 0$ there must be a state description

$$\Psi(a_1, \ldots, a_n) \models \Theta^\dagger(a_1, \ldots, a_n)$$

with $w(\Psi) > 0$. Clearly if $a_i \sim_\Psi a_j$ then $a_i \sim_\Theta a_j$ so $|\mathcal{S}(\Psi)|$ must be at least that of $|\mathcal{S}(\Theta)|$. It follows that if $u^{\bar{p},L}(\Psi) > 0$ then $u^{\bar{p},L}(\Theta) > 0$ and similarly for the $v_t^{\bar{p},L}$. Consequently by the Representation Theorems 1, 2, 14 of [6] we see that if $w(\Psi) > 0$ then $w(\Theta) > 0$ and in turn $w(\theta) > 0$, completing the proof. □

Turning now to the full DRP notice that in a similar fashion to the equivalent version of DCP in the previous section we can prove that DRP is equivalent to:

> *Whenever $\theta \in SL$ does not mention the relation symbol R_j, $w(\theta) > 0$ and the relation symbol R_i has the same arity as R_j then $w(\theta \wedge \forall \vec{x}\,(R_i(\vec{x}) \leftrightarrow R_j(\vec{x}))) > 0$.*

Unfortunately Theorem 13 does not in general extend beyond $QFSL$. Trivially this is the case for homogeneous w by using Theorem 1 and the fact that for $\bar{p} \in \mathbb{B}_\infty$,

$$u^{\bar{p},L}(\forall \vec{x}\,(R_i(\vec{x}) \leftrightarrow R_j(\vec{x}))) = 0.$$

The situation is equally bad for t-heterogeneous probability functions when $t > 1$[11] and L is not purely unary.[12] To see this let R_1, R_2 be relation symbols of L of the same arity and let $L^\dagger = L - \{R_2\}$. Let $\Theta(a_1, \ldots, a_t)$ be a state description of L with $|\mathcal{S}(\Theta)| = t$ such that if $\Theta^\dagger(a_1, \ldots, a_t)$ is the restriction of Θ to $L^\dagger$ then $|\mathcal{S}(\Theta^\dagger)| = t - 1$. Then, see [6], there is a sentence $\zeta(a_1, \ldots, a_t)$ of L saying that $\Theta(a_1, \ldots, a_t)$ holds and any further $a_{t+1}, \ldots, a_m$ (for any m in fact) must look like copies of these $a_1, \ldots, a_t$ and $w(\zeta(\vec{a})) > 0$. Let $\zeta^\dagger(a_1, \ldots, a_t)$ be the corresponding sentence of $L^\dagger$ but with $\Theta^\dagger$ in place of Θ. Then it turns out, as expected, that

$$\zeta(\vec{a}) \models \zeta^\dagger(\vec{a})$$

[11] The only 1-heterogeneous probability function is (the Polyadic extension of) Carnap's c_0^L and in this case DRP is easily seen to hold.

[12] A similar result can be shown for L unary provided $1 < t \leq 2^{q-1}$.

so $w(\zeta^\dagger(\vec{a})) > 0$. However we must have that

$$w(\zeta^\dagger(\vec{a}) \wedge \forall \vec{x}\, (R_1(\vec{x}) \leftrightarrow R_2(\vec{x}))) = 0$$

since for any state description $\Psi(a_1, \dots, a_m)$ with $|\mathcal{S}(\Psi)| = t$ we must have

$$w(\zeta^\dagger(\vec{a}) \wedge \forall \vec{x}\, (R_1(\vec{x}) \leftrightarrow R_2(\vec{x})) \wedge \Psi(a_1, \dots, a_n)) = 0.$$

A Family of Functions Satisfying DCP and DRP.

The previous results provide families of probability functions that satisfy DCP/DRP but without actually providing a complete characterization. Indeed we do not know of any such characterization in terms of previously introduced properties.[13] What we now present (or recall) however is a rather malleable construction of probability functions satisfying these which may at least help guide the path to characterizations.

Let w be a probability function on SL, not necessarily even satisfying Ex. By a construction given by Gaifman (see [5] or [15, p190]) define $w^* : SL \to [0, 1]$ by:

$$w^*(\theta(a_1, a_2, \dots, a_m)) = \sum_{\tau} h_\tau w(\theta(a_{\tau(1)}, a_{\tau(2)}, \dots, a_{\tau(m)}))$$

where the τ range over maps from $\{1, 2, \dots, m\}$ to $\mathbb{N}^+$ and

$$h_\tau = \prod_{i=1}^{m} 2^{-\tau(i)}.$$

Then w^* is a probability function satisfying Ex. Indeed w^* satisfies DCP. To see this let $\theta(a_1, a_2, a_3, \dots, a_m) \in SL$ and suppose that

$$w^*(\theta(a_1, a_1, a_3, \dots, a_m)) = \sum_{\sigma} h_\sigma w(\theta(a_{\sigma(1)}, a_{\sigma(1)}, a_{\sigma(3)}, \dots, a_{\sigma(m)})) > 0,$$

where the σ have domain $\{1, 3, 4, \dots, m\}$. Then for some such σ

$$w(\theta(a_{\sigma(1)}, a_{\sigma(1)}, a_{\sigma(3)}, \dots, a_{\sigma(m)})) > 0$$

so

$$h_\tau w(\theta(a_{\tau(1)}, a_{\tau(2)}, a_{\tau(3)}, \dots, a_{\tau(m)})) > 0$$

[13] Of course it may be that no such characterization exists, that DCP/DRP represent genuinely new 'species'.

for that τ extending σ such that $\tau(2) = \sigma(1)$. Thus $w^*(\theta(a_1, a_2, a_3, \dots, a_m)) > 0$ as required for DCP.

The analogous construction with relation symbols (of the same arity) gives a construction of probability functions satisfying DRP (and by using one after the other yields probability functions satisfying both DCP and DRP).

As already remarked one of the values of these constructions is to furnish probability functions satisfying DCP/DRP and some additional properties.

Conclusion

In this paper we have proposed two compatible principles, DCP and DRP, which intend to capture an aspect of analogical support, support that is in terms of possibility. In this they join a number of other 'principle of analogical support' proposed by ourselves and others (see for example [2], [16]) , both formally and informally. Hopefully in time these various 'experiments' will provide a clearer picture and consensus of the meaning(s) and power of analogy.

Apart from introducing DCP and DRP the main part of this paper has involved attempting to relate these principles and their quantifier free versions to the (currently) fixed-star principles of Constant Exchangeability (Ex), Regularity (S/Reg), and Spectrum Exchangeability (Sx). Whilst some progress has been make in this direction exactly characterizing these principles in terms of existing landmark principles still seems some way off.

On a more optimist note however our results have shown that QF-DCP and QF-DRP possess some degree of *rationality* in that they follow from properties, for example Atom Exchangeability in the unary case, which have traditionally been quite widely regarded as rational principles.

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] Bartha, P., *By Parallel Reasoning,* Oxford University Press, 2009.

[3] Carnap, R., A Basic System of Inductive Logic, in *Studies in Inductive Logic and Probability, Volume I*, Eds. R.Carnap & R.C.Jeffrey, University of California Press, 1971, pp33-165.

[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] Gaifman, H., Concerning Measures on First Order Calculi, *Israel Journal of Mathematics*, 1964, **2**:1-18.

[6] Howarth, E. & Paris, J.B., The Theory of Spectrum Exchangeability, *Review of Symbolic Logic,* **8**(01):108-130, March 2014.

[7] Howarth, E., Paris, J.B. & Vencovská, A., An examination of the SEP Candidate Analogical Inference Rule within Pure Inductive Logic, *Journal of Applied Logic,* 2016, **14**:22-45.

[8] Landes, J., *The Principle of Spectrum Exchangeability within Inductive Logic*, Ph.D. Thesis, University of Manchester, 2009. Available at http://www.maths.manchester.ac.uk/∼jeff/

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

[10] Landes, J., Paris, J.B. & Vencovská, Some Aspects of Polyadic Inductive Logic,*Studia Logica*, 2008, **90**:3-16.

[11] Landes, J., Paris, J.B. & Vencovská, A., Representation Theorems for Probability Functions satisfying Spectrum Exchangeability in Inductive Logic, *International Journal of Approximate Reasoning*, 2009, **51**(1):35-55.

[12] Nix C.J. & Paris, J.B., A Note on Binary Inductive Logic, *Journal of Philosophical Logic*, 2007, **36**(6):735-771.

[13] Paris, J.B., *A Useful Lemma in Unary Predicate Logic*, available at http://www.maths.manchester.ac.uk/∼jeff/lecture-notes/useful.pdf

[14] Paris, J.B., *Pure Inductive Logic: Workshop Notes for Progic 2015,* available at http://www.maths.manchester.ac.uk/∼jeff/lecture-notes/Progic15.pdf

[15] Paris, J.B. & Vencovská, A., *Pure Inductive Logic*, in the Association of Symbolic Logic Series *Perspectives in Mathematical Logic*, Cambridge University Press, April 2015.

[16] Paris, J.B. & Vencovská, A., Combining Analogical Support in Pure Inductive Logic, submitted to Erkenntnis.

[17] Ronel, T. and Vencovská, A., *Invariance Principles in Polyadic Inductive Logic*, Logique at Analyse, 2014, **228**:541-561.