Language Invariance and Spectrum Exchangeability in Inductive Logic

Landes, Juergen and Paris, Jeff and Vencovská, Alena

2008

MIMS EPrint: **2008.107**

ISSN 1749-9097

Language Invariance and Spectrum Exchangeability in Inductive Logic

Jürgen Landes*, Jeff Paris and Alena Vencovská†
School of Mathematics
University of Manchester
Manchester M13 9PL, UK

juergen_landes@yahoo.de
jeff.paris@manchester.ac.uk
alena.vencovska@manchester.ac.uk

August 15, 2008

Abstract

A sufficient condition is given for a probability function in Inductive Logic (with relations of all arities) satisfying spectrum exchangeability to additionally satisfy Language Invariance. This condition is shown to also be necessary in the case of homogeneous probability functions.

KEYWORDS: Uncertain reasoning, inductive logic, probability logic, spectrum exchangeability, language invariance.

Introduction

In common with recent developments in Inductive Logic, see for example [17] (and [1], [2], [3] for the classical approach), we shall work within a first order predicate language L with finitely many relation symbols, countably many constants $a_1, a_2, a_3, \ldots$ and no function symbols. The intention here is that these constants a_i exhaust the universe. Let $SL, QFSL$ respectively denote the sentences and quantifier free sentences of L.

We say that a function $w : SL \to [0, 1]$ is a *probability function* on L if it satisfies that for all $\theta, \phi, \exists x\, \psi(x) \in SL$:

(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) $w(\exists x\, \psi(x)) = \lim_{m\to\infty} w(\bigvee_{i=1}^{m} \psi(a_i))$.

Throughout w, possibly with with various annotations, will denote a probability function on L and, for the purposes of motivation, we shall be thinking of probabilities in the sense of de Finetti as subjective degrees of belief.

*Supported by a MATHLOGAPS Research Studentship.
†Supported by a UK Engineering and Physical Sciences Research Council (EPSRC) Research Assistantship.

By a theorem of Gaifman (see [5], where in fact the axioms (P1-3) were first formulated) any probability function defined on $QFSL$ (i.e. satisfying (P1) and (P2) for $\theta, \phi \in QFSL$) extends uniquely to a probability function on L. In this sense then we can largely limit our considerations to probability functions defined just on $QFSL$. Indeed, by the Disjunctive Normal Form Theorem it then follows that w is determined simply by its values on the *state descriptions*, that is sentences of the form

$$\bigwedge_{s=1}^{m} \bigwedge_{b_1,b_2,...,b_{r_s} \in \{a_{i_1},a_{i_2},...,a_{i_n}\}} \pm P_s(b_1, b_2, ..., b_{r_s})$$

where $P_1, P_2, \ldots, P_s$ are the relations of L with arities $r_1, r_2, \ldots, r_s$ respectively.

In Inductive logic we are basically interested in the choice of probability functions w on L when these are intended to represent the beliefs, i.e subjective probabilities, assigned by a *rational* or *logical* agent in the absence of any prior knowledge. The key restraint here is that this assignment should be *rational* or *logical* and it is customary to identify this with the requirement that w satisfies certain rational or common sense principles.

A number of such principles have been suggested, see for example [1], [2], [3], [8], [10], [13], [15], [16], [17], [19], the most basic of which asserts that w should not treat the constants a_i differently, equivalently:

The Constant Exchangeability Principle (Ex)

For $\theta, \theta' \in QFSL$, if θ' is obtained from θ by replacing, respectively, the (distinct) constant symbols $a_{i_1}, a_{i_2}, \ldots, a_{i_m}$ occurring in θ by the (distinct) constant symbols $a_{k_1}, a_{k_2}, \ldots, a_{k_m}$ then $w(\theta) = w(\theta')$.

We shall henceforth assume that all probability functions mentioned satisfy Ex. Notice that in this case the value of w on the state description as above does not depend on the choice of $i_1, i_2 \ldots, i_n$ and so we may without loss always take these to be $a_1, a_2, \ldots, a_n$.

There are two other such 'common sense' principles which we shall be concerned with in this paper. In order to explain the first of these we first need to introduce a little notation. Given a state description $\Theta(b_1, b_2, \ldots, b_n)$ where the b_i are distinct constants from L (i.e. choices of a_j) we say that b_i, b_j are *indistinguishable* mod Θ, written $b_i \sim_\Theta b_j$, if for all relations $P(x_1, x_2, \ldots, x_r)$ of L the sentence $P(b_{t_1}, b_{t_2}, \ldots, b_{t_r})$ appears positively as a conjunct in $\Theta(b_1, b_2, \ldots, b_n)$ if and only if $P(b_{s_1}, b_{s_2}, \ldots, b_{s_r})$ also appears positively as a conjunct in $\Theta(b_1, b_2, \ldots, b_n)$ where $\langle b_{s_1}, b_{s_2}, \ldots, b_{s_r}\rangle$ is the result of replacing any number of occurrences of b_i in $\langle b_{t_1}, b_{t_2}, \ldots, b_{t_r}\rangle$ by b_j or vice-versa. Clearly $\sim_\Theta$ is an equivalence relation.

Define the *spectrum* of Θ, denoted $\mathcal{S}(\Theta)$, to be the multiset of sizes of the (non-empty) equivalence classes with respect to $\sim_\Theta$.

The Spectrum Exchangeability Principle (Sx)

If $\Theta(b_1, b_2, \ldots, b_n), \Phi(c_1, c_2, \ldots, c_n)$ are state description and $\mathcal{S}(\Theta) = \mathcal{S}(\Phi)$ then $w(\Theta) = w(\Phi)$.

Clearly expressed in this form Sx implies Ex. In the early accounts of Inductive Logic, for example [1], [2], [3], [10], the language L was taken to be purely unary, that is the predicates of the language are just $P_1(x), P_2(x), \ldots, P_s(x)$ (but see [11]). In this case state descriptions have the simple form

$$\bigwedge_{i=1}^{n} \alpha_{h_i}(a_{t_i})$$

where the $\alpha_h(x)$, $h = 1, 2, \ldots, 2^s$ are the *atoms* of L, that is formulae of the form

$$\pm P_1(x) \wedge \pm P_2(x) \wedge \ldots \wedge \pm P_s(x),$$

and Sx reduces to *Atom Exchangeability*, Ax, asserting that

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

depends only on the multiset of $|\{i \,|\, h_i = j\}|$ for $j = 1, 2, \ldots, 2^s$.

The principle Ax was apparently happily accepted by Johnson and Carnap and the earlier investigators since it follows from Johnson's Sufficientness Principle[1] which they advocated.

The second principle which we shall be concerned with here is that of Language Invariance. The motivation behind this principle is that whilst we may at any one time be interested is some particular finite language L a rational choice of beliefs for that language should be capable of extension to a larger language. After all there is clearly no reason to suppose that there are only finitely many relations in existence and that L has already included all of them.

Language Invariance

The probability function w on L satisfies *Language Invariance*[2] if there exist a class of probability functions $w_{\mathcal{L}}$ for each finite predicate language $\mathcal{L}$ (as usual with constants a_i and no function symbols) such that whenever $\mathcal{L}'$ is a sublanguage of $\mathcal{L}$ then $w_{\mathcal{L}} \restriction S\mathcal{L}' = w_{\mathcal{L}'}$ and $w_L = w$.

In this case we shall describe the $w_{\mathcal{L}}$ as a *language invariant family* containing w.

In the next section we shall derive a sufficient condition for a probability function satisfying Spectrum Exchangeability to also satisfy Language Invariance.

A Sufficiency Condition for Language Invariance

Before stating and proving the main result of this paper we need to introduce a particular family of probability functions $u^{\overline{p}}_L$.

Let

$$\mathbb{B} = \{\langle x_0, x_1, x_2, ...\rangle \,|\, x_1 \geq x_2 \geq ... \geq 0,\, x_0 \geq 0,\, \sum_{i=0}^{\infty} x_i = 1\}$$

and endow $\mathbb{B}$ with the standard weak product topology inherited from $[0,1]^\infty$. Let

$$\overline{p} = \langle p_0, p_1, p_2, \ldots\rangle \in \mathbb{B}.$$

For a state description $\Theta(a_{i_1}, a_{i_2}, \ldots, a_{i_q})$ (from language L) and 'colors'

$$\vec{c} = \langle c_1, c_2, \ldots, c_q\rangle \in \{0, 1, 2, \ldots\}^q$$

(where 0 stands for the special color black) we define $j^{\overline{p}}(\Theta(a_{i_1}, a_{i_2}, \ldots, a_{i_q}), \vec{c})$ inductively as follows:

[1] See for example [19] or [20] for a formulation of this Principle in the notation of this paper.

[2] This differs from the earlier definition of Language Invariance given in [8] and [19] which was restricted to purely unary languages $\mathcal{L}$, $\mathcal{L}'$.

Set $j^{\overline{p}}(\top, \emptyset) = 1$. Suppose that at stage q we have defined the probability $j^{\overline{p}}(\Theta(a_{i_1}, a_{i_2}, \ldots, a_{i_q}), \vec{c})$. Pick color c_{q+1} from $0, 1, 2, \ldots$ according to the probabilities $p_0, p_1, p_2, \ldots$ and let

$$\vec{c}^{+} = \langle c_1, \ldots, , c_q, c_{q+1} \rangle.$$

If c_{q+1} is the same as an earlier color, c_j say, with $c_j \neq 0$ extend $\Theta(a_{i_1}, a_{i_2}, \ldots, a_q)$ to the unique state description $\Theta^{+}(a_{i_1}, a_{i_2}, \ldots, a_{i_q}, a_{i_{k+1}})$ for which $a_{i_j} \sim_{\Theta^+} a_{i_{q+1}}$. On the other hand if c_{p+1} is 0 or a previously unchosen color then randomly choose $\Theta^{+}(a_{i_1}, a_{i_2}, \ldots, a_{i_q}, a_{i_{q+1}})$ extending $\Theta(a_{i_1}, a_{i_2}, \ldots, a_{i_q})$ such that $\sim_\Theta$ and $\sim_{\Theta^+}$ agree on $\{a_{i_1}, a_{i_2}, \ldots, a_{i_q}\}^2$. Finally let $j^{\overline{p}}(\Theta^+, \vec{c}^{+})$ be $j^{\overline{p}}(\Theta, \vec{c})$ times the probability as described of then going from $\Theta, \vec{c}$ to $\Theta^+, \vec{c}^{+}$.

Having defined these $j^{\overline{p}}(\Theta, \vec{c})$ now set

$$u_L^{\overline{p}}(\Theta) = \sum_{\vec{c}} j^{\overline{p}}(\Theta, \vec{c}).$$

By a straightforward generalization of the result in [17] (where just two colors were considered) $u_L^{\overline{p}}$ satisfies Ex and Sx.

Theorem 1 *Let the probability function w on L satisfy Sx. Then for w to be language invariant it is sufficient that there is a countably additive measure μ on the Borel subsets of $\mathbb{B}$ such that for $\theta \in SL$,*

$$w(\theta) = \int_{\mathbb{B}} u_L^{\overline{p}}(\theta) d\mu. \tag{1}$$

Furthermore in this case if L contains at least one non-unary relation then the language invariant family containing w is unique.

We call μ as in this theorem the *de Finetti prior* of w.

Proof Suppose that (1) holds. Let $\mathcal{L}$ extend L and for $\phi \in S\mathcal{L}$ set

$$w_{\mathcal{L}}(\phi) = \int_{\mathbb{B}} u_{\mathcal{L}}^{\overline{p}}(\phi) d\mu, \tag{2}$$

in other words $w_{\mathcal{L}}$ has the same de Finetti prior as w, but the language has changed. We claim that $w_{\mathcal{L}} \restriction SL = w$. To show this it is enough to show that for a state description $\Theta(\vec{a})$ from language L, $w_{\mathcal{L}}(\Theta(\vec{a})) = w(\Theta(\vec{a}))$, and for this it is enough to show that

$$u_{\mathcal{L}}^{\overline{p}}(\Theta(\vec{a})) = u_L^{\overline{p}}(\Theta(\vec{a})). \tag{3}$$

Let $\Phi(\vec{a})$ be a state description for $\mathcal{L}$ extending $\Theta(\vec{a})$ (and with the same constants $\vec{a} = \langle a_1, a_2, \ldots, a_q \rangle$) and consider a summand $j_{\mathcal{L}}^{\overline{p}}(\Phi(\vec{a}), \vec{c})$ which yields $u_{\mathcal{L}}^{\overline{p}}(\Theta(\vec{a}))$ via

$$u_{\mathcal{L}}^{\overline{p}}(\Phi(\vec{a})) = \sum_{\vec{c}} j_{\mathcal{L}}^{\overline{p}}(\Phi(\vec{a}), \vec{c}).$$

This summand is formed by q choices of colors $c_1, c_2, \ldots, c_q$ and an increasing sequences of choices of state descriptions

$$\Phi_1(a_1), \Phi_2(a_1, a_2), \Phi_3(a_1, a_2, a_3), \ldots, \Phi_q(a_1, a_2, \ldots, a_q) = \Phi(\vec{a}).$$

Let $\Theta_k(a_1, a_2, \ldots, a_k)$ be the state description of L which $\Phi_k(a_1, a_2, \ldots, a_k)$ extends. Then

$$\Theta_q(a_1, a_2, \ldots, a_q) = \Theta(a_1, a_2, \ldots, a_q)$$

and for this same choice of colors $\vec{c}$ and Θ_k $j_L^{\overline{p}}(\Theta(\vec{a}), \vec{c})$ is a contributing summand to $u_L^{\overline{p}}(\Theta(\vec{a}))$. Furthermore the only difference between these two contributions is that at each choice of the kth state description $j_{\mathcal{L}}^{\overline{p}}(\Phi(\vec{a}),\vec{c})$ receives as a multiplicative factor one over the number of possible state descriptions in $\mathcal{L}$ extending $\Phi_{k-1}(a_1,a_2,\ldots,a_{k-1})$ whereas $j_L^{\overline{p}}(\Theta(\vec{a}),\vec{c})$ receives as a multiplicative factor one over the number of possible state descriptions in L extending $\Theta_{k-1}(a_1,a_2,\ldots,a_{k-1})$. Note that this depends on c_k, genuine choice being available only when c_k is either zero or not equal to any previous c_j. However, since otherwise in each case these factors depend only on the relations in L and $\mathcal{L}$ and not on the particular state descriptions $\Theta_{k-1}(a_1,\ldots,a_{k-1})$, $\Phi_{k-1}(a_1,\ldots,a_{k-1})$, overall

$$j_L^{\overline{p}}(\Theta(\vec{a}),\vec{c}) = M j_{\mathcal{L}}^{\overline{p}}(\Phi(\vec{a}),\vec{c})$$

where M is the number of possible choices (according to $\vec{c}$) of state descriptions of $\vec{a}$ in $\mathcal{L}$ extending $\Theta(\vec{a})$. But this means that

$$j_L^{\overline{p}}(\Theta(\vec{a}),\vec{c}) = \sum_{\Psi(\vec{a})} j_{\mathcal{L}}^{\overline{p}}(\Psi(\vec{a}),\vec{c})$$

where $\Psi(\vec{a})$ runs over the M many state descriptions of $\vec{a}$ in $\mathcal{L}$ admitted by $\vec{c}$ and extending $\Theta(\vec{a})$. Since

$$u_{\mathcal{L}}^{\overline{p}}(\Theta(\vec{a})) = \sum_{\Psi(\vec{a})}\sum_{\vec{c}} j_{\mathcal{L}}^{\overline{p}}(\Psi(\vec{a}),\vec{c})$$

where the sum is over all extensions $\Psi(\vec{a})$ in $\mathcal{L}$ of $\Theta(\vec{a})$, rearranging the summation on the right hand side yields

$$u_{\mathcal{L}}^{\overline{p}}(\Theta(\vec{a})) = u_L^{\overline{p}}(\Theta(\vec{a})),$$

as required.

Of course the required 'full' language invariant family for w can now be obtained by restricting/marginalizing these $w_{\mathcal{L}}$.

To show uniqueness suppose that L has some non-unary relation symbol and that there are two different language invariant families containing w, say w', w'' are the members of these families defined on $\mathcal{L} \supset L$ and they differ on some state description, $\Psi(a_1,a_2,\ldots,a_n)$ say.

We first define a well founded ordering on state descriptions $\Theta(a_1,a_2,\ldots,a_n)$ of $\mathcal{L}$ or L, for fixed n, by setting

$$\Theta(\vec{a}) \trianglelefteq \Phi(\vec{a}) \iff \sim_\Theta \text{ is a refinement of } \sim_\Phi .$$

We now show

$$w'(\Theta(\vec{a})) = w''(\Theta(\vec{a})) \tag{4}$$

by induction on this ordering. The least point in this ordering is when the equivalence classes of $\sim_\Theta$ are all singletons. In this case let $\Phi_L(a_1,a_2,\ldots,a_n)$ be a state description of L having this minimal spectrum. (This is where we need L to contain a non-unary relation symbol, to ensure that such a state description exists.) Then $\sim_{\Phi(\vec{a})}$ must again be this minimal spectrum for any state description $\Phi(a_1,a_2,\ldots,a_n)$ of $\mathcal{L}$ extending $\Phi_L(a_1,a_2,\ldots,a_n)$ and w' must take the same value on these by Sx. Hence, since

$$w(\Phi_L(a_1,a_2,\ldots,a_n)) = w'(\Phi_L(a_1,a_2,\ldots,a_n)) = \sum_{\Phi(\vec{a})} w'(\Phi(\vec{a})),$$

where the summation is over state descriptions $\Phi(\vec{a})$ extending $\Phi_L(\vec{a})$, we see that if M is the number of such $\Phi(\vec{a})$ then for any one of them

$$w'(\Phi(a_1,a_2,\ldots,a_n)) = M^{-1} w(\Phi_L(a_1,a_2,\ldots,a_n)).$$

Since this reasoning also applies to w'' (4) holds in this base case.

Now suppose that (4) holds for all $\Phi(\vec{a}) \lhd \Theta(\vec{a})$. Let $\Theta_L(\vec{a})$ be a state description of L having the same spectrum as $\Theta(\vec{a})$. Then again,

$$w(\Theta_L(\vec{a})) = \sum_{\Phi(\vec{a})} w'(\Phi(\vec{a})) \tag{5}$$

where the $\Phi(\vec{a})$ range over state descriptions in $\mathcal{L}$ extending $\Theta_L(\vec{a})$. Now all of these $\Phi(\vec{a})$ are less or equal $\Theta(\vec{a})$ in the ordering $\unlhd$, *and* $\Theta(\vec{a})$ does itself appear on the right hand side of this expression a non-zero number of times. Furthermore the identity (5) also holds with the probability function w'' in place of w', and by the inductive hypothesis these terms are the same except possibly for the argument $\Theta(\vec{a})$. But then of course they must also be the same in this case, as required to prove (4) and the theorem.

$\blacksquare$

An Application

For this section assume that our default language L has at least one non-unary relation. We first recall a classification[3] of probability functions w on L satisfying Sx.

Given a spectrum $\mathcal{S} = \{s_1, s_2, \dots, s_k\}$ let $|\mathcal{S}| = k$ and $\sum \mathcal{S} = \sum_{i=1}^k s_i$. For w on L satisfying Sx let $w(\mathcal{S}) = w(\Theta)$ for some/any state description Θ with spectrum $\mathcal{S}$. For $\Theta = \Theta(a_1, a_2, \dots, a_r)$ and $\mathcal{S}(\Theta) = \mathcal{S}$ (so $\sum \mathcal{S} = r$) let $\mathcal{N}^L(\mathcal{S}, \mathcal{T})$ be the number of state descriptions Φ extending Θ with $\mathcal{S}(\Phi) = \mathcal{T}$. By results in [16], [17], [13] this does not depend on the particular Θ with $\mathcal{S}(\Theta) = \mathcal{S}$ which is chosen.

We say that w is *homogeneous* if for all k

$$\lim_{r\to\infty} \sum_{|\mathcal{S}|=k, \sum \mathcal{S}=r} w(\mathcal{S}) = 0.$$

In other words the probability that all the a_i will fall in some fixed finite number of equivalence classes with respect to indistinguishability is zero.

We say that w is t*-heterogeneous* if

$$\lim_{r\to\infty} \sum_{|\mathcal{S}|=t, \sum \mathcal{S}=r} w(\mathcal{S}) = 1.$$

In other words the probability that all the a_i will fall in some t (non-empty) equivalence classes with respect to indistinguishability is 1.

The following theorem appears in [16], [17] for the case of a purely binary language and will appear in [13] for general not purely unary languages.

Theorem 2 *Let w satisfy Sx. Then there are probability functions $w^{[t]}$ satisfying Sx and constants $\eta_t \geq 0$ for $0 \leq t < \infty$ such that*

$$w = \sum_{i=0}^{\infty} \eta_i w^{[i]}, \qquad \sum_{i=0}^{\infty} \eta_i = 1,$$

$w^{[t]}$ is t-heterogeneous for $t > 0$ and $w^{[0]}$ is homogeneous. Furthermore the η_i are unique and so are the $w^{[i]}$ when $\eta_i \neq 0$.

[3]Given in [16], [17] for binary languages and more generally in the forthcoming [13].

The following result[4] will appear in the forthcoming paper [14].

Theorem 3 *Let w be a homogeneous probability function on L (not purely unary) satisfying Sx. Then there is a countably additive measure μ on the Borel subsets of $\mathbb{B}$ such that for $\theta \in SL$,*

$$w(\theta) = \int_{\mathbb{B}} u_L^{\overline{p}}(\theta) d\mu.$$

Using this result we have the following corollary to Theorem 1:

Corollary 4 *Let w be a homogeneous probability function on L (not purely unary) satisfying Sx. Then w satisfies Language Invariance.*

This is, to our way of thinking, a rather surprising result since it say in particular that just knowing a homogeneous w on a sublanguage consisting of a single non-unary relation is, provided we require Sx to be preserved, enough to determine it on all extensions of that language.

In contrast to Corollary 4 however:

Proposition 5 *Let $t > 1$ and let w be a t-heterogeneous probability function on L (not purely unary) satisfying Sx. Then w does not satisfy Language Invariance.*

Proof Suppose that w is a t-heterogeneous probability function on L and a member of some language invariant family. Let w' be a member of this family on $\mathcal{L} = L \cup \{P_1, P_2, \ldots, P_{t+1}\}$ where the P_i are new unary predicates.

Since L contains a non-unary relation we can find a state description $\Theta(a_1, a_2, \ldots, a_{t+1})$ for $\mathcal{L}$ whose restriction $\Theta_L(a_1, a_2, \ldots, a_{t+1})$ to L has a spectrum of length $t+1$. Hence if $w'(\Theta(a_1, a_2, \ldots, a_{t+1})) > 0$ then $w(\Theta_L(a_1, a_2, \ldots, a_{t+1})) > 0$, contradicting t-heterogeneity. So it must be the case that $w'(\Phi(\vec{a})) = 0$ whenever the spectrum of the state description $\Phi(\vec{a})$ has length greater than t (since for any such spectrum there is a state description with this spectrum which extends a state description on $t+1$ individuals of length $t+1$).

Now let $\Theta(a_1, a_2, \ldots, a_t)$ be a state description for $\mathcal{L}$ with spectrum of length t whose restriction $\Theta_L(a_1, a_2, \ldots, a_t)$ to L has a spectrum of length 1. Then if $\Phi(a_1, a_2, \ldots, a_{t+j})$ is a state description for $a_1, a_2, \ldots, a_{t+j}$ in $\mathcal{L}$ extending $\Theta(a_1, a_2, \ldots, a_t)$ and $w(\Phi(a_1, a_2, \ldots, a_{t+j})) \neq 0$ it must be the case that $\Phi(a_1, a_2, \ldots, a_{t+j})$ has spectrum of length t and in consequence the restriction $\Phi_L(a_1, a_2, \ldots, a_{t+j})$ to L must still have spectrum of length 1. Hence

$$\begin{aligned}
\sum_{|\mathcal{S}|=1, \sum \mathcal{S}=r} w(\mathcal{S}) &= \sum_{\Phi_L(a_1,\ldots,a_r) \text{ has spectrum of length } 1} w'(\Phi_L(\vec{a})) \\
&= \sum_{\Phi_L(a_1,\ldots,a_r) \text{ has spectrum of length } 1} w'(\Phi(\vec{a})) \\
&\geq \sum_{\Phi(a_1,\ldots,a_r) \text{ extends } \Theta(a_1,\ldots,a_t)} w'(\Phi(\vec{a})) \\
&= w'(\Theta_L(a_1, \ldots, a_t)) \\
&> 0,
\end{aligned}$$

so

$$\lim_{r\to\infty} \sum_{|\mathcal{S}|=1, \sum \mathcal{S}=r} w(\mathcal{S}) \geq w'(\Theta_L(a_1, \ldots, a_t)) > 0,$$

[4] A similar representation theorem can be proved for t-heterogeneous probability functions, see [18], but that will not be needed here.

contradicting t-heterogeneity. ■

Notice however that it is certainly possible to have mixtures of t-heterogeneous probability functions (for different t) which are language invariant. For example if we take $\overline{p} = \langle p_0, p_1, p_2, \dots \rangle \in \mathbb{B}$ with $p_0 = 0$ and $p_s = 0$ for $s > t$ then $u_L^{\overline{p}}$ is a convex combination of r-heterogeneous probability functions for $r \leq t$ and *is* language invariant by Theorem 1.

Proposition 5 does not hold if $t = 1$, the trivial probability function on L which gives probability 1 to all the a_i being indistinguishable (i.e. $1, 2, 3, \dots$ all being in the same equivalence class) provides, as L varies, the example of such a language invariant family.

We finally observe that the requirement in Proposition 5 that L contains a non-unary relation can be dropped if $t < 2^s$ where s is the number of unary relation symbols in L.

Conclusion

Since both Sx and Language Invariance are (we would claim) desirable principles in the context of assigning beliefs in the absence of any prior knowledge it is pleasing to have a sufficiency theorem for such probability functions in terms the particularly simple functions $u_L^{\overline{p}}$. This furthermore opens the possibility of deriving certain other properties of such functions by moving the onus of the task onto the much more malleable $u_L^{\overline{p}}$, examples of which will be given in the forthcoming [14].

References

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

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

[3] Carnap, R., A basic system of inductive logic, in *Studies in Inductive Logic and Probability*, Volume II, ed. R. C. Jeffrey, University of California Press, 1980, 7-155.

[4] De Finetti, B., La prevision: ses lois logiques, ses sources subjetive, *Annales de l'Institut Henri Poincaré*, 1937, **7**:1-68.

[5] Gaifman, H.: Concerning measures on first order calculi, *Israel journal of mathematics,* 1964, 2:1-18

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

[7] Goodman, N., On infirmities in confirmation-theory, *Philosophy and Phenomenology Research*, 1947, **8**:149-151.

[8] Hill, M.J., Paris, J.B. & Wilmers, G.M., Some observations on induction in predicate probabilistic reasoning, *Journal of Philosophical Logic*, 2002, **31**(1):43-75.

[9] Hoover, D.N., Relations on probability spaces and arrays of random variables. Preprint, Institute of Advanced Study, Princeton, 1979.

[10] Johnson, W.E., Probability: The deductive and inductive problems, *Mind*, 1932, **41**(164):409-423.

[11] Kemeny, J.G., Carnap's theory of probability and induction, in *The Philosophy of Rudolf Carnap*, ed. P.A.Schilpp, La Salle, Illinois, Open Court, 1963, 711-738.

[12] Krauss, P.H., Representation of symmetric probability models, *Journal of Symbolic Logic*, 1969, **34**(2):183-193.

[13] Landes, J., Doctorial Thesis, Manchester University, UK, to appear.

[14] Landes, J., Paris, J.B. & Vencovská, A., Representation Theorems for probability functions satisfying Spectrum Exchangeability in Inductive Logic. To be submitted to the *Journal of Symbolic Logic*.

[15] Nix, C.J. & Paris, J.B., A Continuum of inductive methods arising from a generalized principle of instantial relevance, *Journal of Philosophical Logic*, Online First Issue, DOI:10,1007/s 10992-005-9003x, ISSN 0022-3611 (Paper) 1573-0433, 2005.

[16] Nix, C.J., *Probabilistic Induction in the Predicate Calculus* Doctorial Thesis, Manchester University, Manchester, UK, 2005.
See `http://www.maths.man.ac.uk/~jeff/#students`

[17] Nix, C.J. & Paris, J.B., A note on binary inductive logic, to appear in the *Journal of Philosophical Logic*.

[18] Paris, J.B. & Vencovská, A., From unary to binary inductive logic, to appear in the *Proceedings of the Second Indian Conference on Logic and its Applications,* Mumbai, 2007.

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

[20] Vencovská, A., Binary Induction and Carnap's Continuum, Proceedings of the 7th Workshop on Uncertainty Processing (WUPES), Mikulov, 2006. See `http://mtr.utia.cas.cz/wupes06/articles/data/vencovska.pdf`