Pure Inductive Logic
Workshop Notes for ISLA 2014

Jeff Paris,
School of Mathematics,
University of Manchester,
Manchester M13 9PL, UK.

jeff.paris@manchester.ac.uk

Notes and slides available under Teaching Material at
www.maths.manchester.ac.uk/~jeff/

Context and Notation

L_q is the first order language with

- Constant symbols a_n, $n \in \mathbb{N}^+ = \{1, 2, 3, \dots\}$
- Predicate (i.e. unary relation) symbols $R_1, R_2, \dots, R_q$.

$SL_q, QFSL_q$ denote the sentences and quantifier free sentences of L_q.

Let M be a structure for L_q with universe the interpretations of the a_n (also denoted a_n).

> Question: Given an agent $\mathcal{A}$ inhabiting M and $\theta \in SL_q$ what probability $w(\theta)$ should $\mathcal{A}$ rationally, or logically, assign to θ?

Very Important Condition here: $\mathcal{A}$ knows nothing about M, s/he has no particular interpretation in mind for the constants and predicates.

More precisely:

> Question: Given an agent $\mathcal{A}$ inhabiting M, rationally or logically, what probability function w should $\mathcal{A}$ adopt?

Here $w : SL_q \to [0, 1]$ is a *probability function on* L_q if for all $\theta, \phi, \exists x\, \psi(x) \in SL_q$

(P1) $\models \theta \;\Rightarrow\; w(\theta) = 1$.

(P2) $\theta \models \neg\phi \;\Rightarrow\; w(\theta \vee \phi) = w(\theta) + w(\phi)$.

(P3) $w(\exists x\, \psi(x)) = \lim_{n\to\infty} w(\psi(a_1) \vee \psi(a_2) \vee \ldots \vee \psi(a_n))$.

Proposition 1 *Let w be a probability function on SL. Then for $\theta, \phi \in SL$,*

(a) $w(\neg\theta) = 1 - w(\theta)$.

(b) $\models \neg\theta \;\Rightarrow\; w(\theta) = 0$.

(c) $\theta \models \phi \;\Rightarrow\; w(\theta) \leq w(\phi)$.

(d) $\theta \equiv \phi \;\Rightarrow\; w(\theta) = w(\phi)$.

(e) $w(\theta \vee \phi) = w(\theta) + w(\phi) - w(\theta \wedge \phi)$.

For $\phi \in SL_q$ the corresponding conditional probability function $w(- \,|\, \phi)$ is a probability function such that for $\theta \in SL_q$,

$$w(\theta \,|\, \phi) \cdot w(\phi) = w(\theta \wedge \phi), \quad \text{i.e. } w(\theta \,|\, \phi) = \frac{w(\theta \wedge \phi)}{w(\phi)} \text{ if } w(\phi) > 0.$$

Specifying Probability Functions

Gaifman's Theorem 2 *Suppose that $w : QFSL_q \to [0,1]$ satisfies (P1) and (P2) for $\theta, \phi \in QFSL_q$. Then w has a unique extension to a probability function on L_q satisfying (P1),(P2),(P3) for any $\theta, \phi, \exists x\, \psi(x) \in SL_q$.*

Example

Let $\alpha_1, \ldots, \alpha_{2^q}$, the atoms of L_q, denote the 2^q formulae of the form

$$R_1^{\epsilon_1}(x) \wedge R_2^{\epsilon_2}(x) \wedge \ldots \wedge R_q^{\epsilon_n}(x)$$

where the $\epsilon_i \in \{0,1\}$ and $R^1 = R, R^0 = \neg R$.

Let

$$\vec{c} \in \mathbb{D}_{2^q} = \{\langle x_1, x_2, \ldots, x_{2^q}\rangle \,|\, x_i \geq 0, \sum_{i=1}^{2^q} x_i = 1\}$$

Define $w_{\vec{c}}$ on (instantiations) of atoms by

$$w_{\vec{c}}(\alpha_j(a_i)) = c_j, \quad j = 1, 2, \ldots, 2^q,$$

Extend $w_{\vec{c}}$ to *state descriptions*, that is conjunctions of atoms, by setting, for $b_1, b_2, \ldots, b_n$ distinct elements of $\{a_k \,|\, k \in \mathbb{N}^+\}$,

$$\begin{aligned} & w_{\vec{c}}(\alpha_{h_1}(b_1) \wedge \alpha_{h_2}(b_2) \wedge \ldots \wedge \alpha_{h_n}(b_n)) \\ &= w_{\vec{c}}(\alpha_{h_1}(b_1)) \times w_{\vec{c}}(\alpha_{h_2}(b_2)) \times \ldots \times w_{\vec{c}}(\alpha_{h_n}(b_n))) \\ &= c_{h_1} \times c_{h_2} \times \ldots \times c_{h_n} \\ &= \prod_{j=1}^{n} c_{h_j}. \end{aligned}$$

By the Disjunctive Normal Form Theorem, for $\theta(b_1, b_2, \ldots, b_n) \in QFSL_q$

$$\theta(b_1, b_2, \ldots, b_n) \equiv \bigvee_{k=1}^{r} \bigwedge_{i=1}^{n} \alpha_{h_{ik}}(b_i)$$

for some h_{ik}.

Set

$$\begin{aligned} w_{\vec{c}}(\theta(b_1,\ldots,b_n) &= w_{\vec{c}}\left(\bigvee_{k=1}^{r}\bigwedge_{i=1}^{n}\alpha_{h_{ik}}(b_i)\right) \\ &= \sum_{k=1}^{r} w_{\vec{c}}\left(\bigwedge_{i=1}^{n}\alpha_{h_{ik}}(b_i)\right) \\ &= \sum_{k=1}^{r}\prod_{i=1}^{n} w_{\vec{c}}(\alpha_{h_{ik}}(b_i)) \\ &= \sum_{k=1}^{r}\prod_{i=1}^{n} c_{h_{ik}}. \end{aligned}$$

By Gaifman's Theorem $w_{\vec{c}}$ extends to a probability function on L_q.

Rational Principles

Sources:

- Symmetry
- Irrelevance
- Relevance
- Analogy

The Constant Exchangeability Principle Ex

For $\theta(a_1, a_2, \ldots, a_n) \in SL_q$ and (distinct) $a_{i_1}, a_{i_2}, \ldots, a_{i_n}$

$$w(\theta(a_1, a_2, \ldots, a_n)) = w(\theta(a_{i_1}, a_{i_2}, \ldots, a_{i_n})).$$

The $w_{\vec{c}}$ satisfy Ex.

de Finetti's Representation Theorem 3 *A probability function w on the language L_q satisfies Ex just if it is a mixture of the $w_{\vec{c}}$.*

More precisely, just if

$$w = \int w_{\vec{x}}\, d\mu(\vec{x})$$

where μ is a normalized countably additive measure on the Borel subsets of

$$\mathbb{D}_{2^q} = \{\langle x_1, x_2, \ldots, x_{2^q}\rangle \,|\, 0 \leq x_1, x_2, \ldots, x_{2^q}, \sum_i x_i = 1\}.$$

Theorem 4 *Ex implies the:*

Principle of Instantial Relevance, PIR:
For $\theta(a_1, a_2, \ldots, a_n) \in SL_q$,

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

Proof Let the probability function w on L satisfy Ex.
Without loss of generality let $\alpha_i(x) = \alpha_1(x)$ and, for simplicity,

$$\theta(a_1, \ldots, a_n) \equiv \bigvee_{k=1}^{r} \bigwedge_{i=1}^{n} \alpha_{h_{ik}}(a_i).$$

Then for μ the de Finetti prior for w,

$$w(\theta(a_1, \ldots, a_n)) = \int_{\mathbb{D}_{2^q}} w_{\vec{x}}(\theta(a_1, \ldots, a_n))\, d\mu(\vec{x}) = \int_{\mathbb{D}_{2^q}} \sum_{k=1}^{r} \prod_{i=1}^{n} x_{h_{ik}}\, d\mu(\vec{x}) = A \text{ say},$$

$$w(\alpha_1(a_{n+1}) \wedge \theta(a_1, \ldots, a_n)) = \int_{\mathbb{D}_{2^q}} x_1 \sum_{k=1}^{r} \prod_{i=1}^{n} x_{h_{ik}}\, d\mu(\vec{x}),$$

$$w(\alpha_1(a_{n+2}) \wedge \alpha_1(a_{n+1}) \wedge \theta(a_1, \ldots, a_n)) = \int_{\mathbb{D}_{2^q}} x_1^2 \sum_{k=1}^{r} \prod_{i=1}^{n} x_{h_{ik}}\, d\mu(\vec{x})$$

and (1) amounts to

$$\left(\int_{\mathbb{D}_{2^q}} x_1 \sum_{k=1}^{r} \prod_{i=1}^{n} x_{h_{ik}}\, d\mu(\vec{x}) \right)^2 \leq \left(\int_{\mathbb{D}_{2^q}} \sum_{k=1}^{r} \prod_{i=1}^{n} x_{h_{ik}}\, d\mu(\vec{x}) \right) \cdot \left(\int_{\mathbb{D}_{2^q}} x_1^2 \sum_{k=1}^{r} \prod_{i=1}^{n} x_{h_{ik}}\, d\mu(\vec{x}) \right). \quad (2)$$

If $A = 0$ then this clearly holds, so assume that $A \neq 0$.

Then multiplying out

$$0 \leq \int_{\mathbb{D}_{2^q}} \left(x_1 A - \int_{\mathbb{D}_{2^q}} x_1 \sum_{k=1}^{r} \prod_{i=1}^{n} x_{h_{ik}}\, d\mu(\vec{x}) \right)^2 \sum_{k=1}^{r} \prod_{i=1}^{n} x_{h_{ik}}\, d\mu(\vec{x}). \quad (3)$$

and dividing by A^2 gives (2), and the result follows. ■

The Extended Principle of Instantial Relevance, EPIR

For $\theta(a_1, a_2, \ldots, a_n), \phi(a_{n+1}) \in SL_q$,

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

Principle of Predicate Exchangeability, Px

For $\phi(R_1, R_2, \ldots, R_m) \in SL_q$, where we explicitly display the predicate symbols occurring in ϕ, and (distinct) $1 \leq i_1, i_2, \ldots, i_m \leq q$,

$$w(\phi(R_1, R_2, \ldots, R_m)) = w(\phi(R_{i_1}, R_{i_2}, \ldots, R_{i_m})).$$

The $w_{\vec{c}}$ do not satisfy Px in general.

Unary Language Invariance, ULi

A probability function w on L_q satisfies Unary Language Invariance if there is a family of probability functions w^r, one on each language L_r for $r \in \mathbb{N}^+$, such that $w = w^q$, each member of this family satisfies Px and whenever $p \leq r$ then $w^r \restriction SL_p = w^p$.

Principles of Analogy

Counterpart Principle, CP

For any $\theta \in SL_q$, if $\theta' \in SL_q$ is obtained by replacing some of the predicate and constant symbols appearing in θ by (distinct) new ones not occurring in θ and $\psi \in SL_q$ only mentions constant and predicate symbols common to both θ and θ' then

$$w(\theta \,|\, \theta' \wedge \psi\,) \geq w(\theta \,|\, \psi).$$

CP is *analogical support by structural similarity*

Theorem 5 *Let the probability function w on L_q satisfy ULi. Then w satisfies the Counterpart Principle, CP.*

Proof We may assume that $w(\psi) > 0$.

Let the ULi family consist of w^r on L_r for $r \in \mathbb{N}^+$.

Then

$$w_\infty = \bigcup_{r=1}^{\infty} w_r$$

is a probability function on the infinite (unary) language $L_\infty = \{R_1, R_2, R_3, \ldots\}$ extending w and satisfying Ex and Px.

Let θ, θ', ψ be as in the statement of CP.

We may assume that all the constant and predicate symbols appearing in θ which are common to θ' are amongst $a_1, a_2, \ldots, a_n$, $R_1, R_2, \ldots, R_g$, and that the replacements are $a_{n+i} \mapsto a_{n+i+k}$ for $i = 1, \ldots, k$ and $R_{g+j} \mapsto R_{g+j+t}$ for $j = 1, \ldots, t$.

Suppressing these common constant and predicate symbols we can write

$$\theta = \theta(a_{n+1}, a_{n+2}, \ldots, a_{n+k}, R_{g+1}, R_{g+2}, \ldots, R_{g+t}),$$
$$\theta' = \theta(a_{n+k+1}, a_{n+k+2}, \ldots, a_{n+2k}, R_{g+t+1}, R_{g+t+2}, \ldots, R_{g+2t}).$$

Let

$$\theta_{i+1} = \theta(a_{n+ik+1}, a_{n+ik+2}, \ldots, a_{n+(i+1)k}, R_{g+it+1}, R_{g+it+2}, \ldots, R_{g+(i+1)t}) \in SL_\infty$$

so $\theta_1 = \theta$, $\theta_2 = \theta'$.

Define $\tau : QFSL_1 \to SL_\infty$ by

$$\tau(R_1(a_i)) = \theta_i, \quad \tau(\neg\phi) = \neg\tau(\phi), \quad \tau(\phi \wedge \eta) = \tau(\phi) \wedge \tau(\eta), \qquad \text{etc.}$$

Define $v : QFSL_1 \to [0,1]$ by

$$v(\phi) = w_\infty(\tau(\phi) \mid \psi).$$

Since w_∞ satisfies (P1-2) (on SL_∞) so does v (on $QFSL_1$).

Since w_∞ satisfies Ex + Px, for $\phi \in QFSL_1$, permuting the θ_i in $w(\tau(\phi) \mid \psi)$ will leave this value unchanged so permuting the a_i in ϕ will leave $v(\phi)$ unchanged. i.e. v satisfies Ex.

By Gaifman's Theorem v has an extension to a probability function on L_1 which still satisfies Ex.

Hence v satisfies PIR by Theorem 4, so

$$v(R_1(a_1) \mid R_1(a_2)) \geq v(R_1(a_1)).$$

But since $\tau(R_1(a_1)) = \theta$, $\tau(R_1(a_2)) = \theta'$ this gives

$$w_\infty(\theta \mid \theta' \wedge \psi) \geq w_\infty(\theta \mid \psi)$$

and

$$w(\theta \mid \theta' \wedge \psi) \geq w(\theta \mid \psi)$$

■

With the above notation we can also show that

$$w(\theta_{n+1} \mid \bigwedge_{i=1}^{m} \theta_i \wedge \bigwedge_{j=m+1}^{n} \neg\theta_j) \;\geq\; w(\theta_{n+1} \mid \bigwedge_{i=1}^{k} \theta_i \wedge \bigwedge_{j=k+1}^{n} \neg\theta_j)$$

whenever $m \geq k$.

Theorem 6 *Let the probability function w on L_q satisfy ULi and let*

$$\begin{aligned} \theta &= \theta(\vec{a_1}, \vec{a_2}, \vec{a_3}, \vec{R_1}, \vec{R_2}, \vec{R_3}) \\ \theta' &= \theta(\vec{a_1}, \vec{a_2}, \vec{a_4}, \vec{R_1}, \vec{R_2}, \vec{R_4}) \\ \theta'' &= \theta(\vec{a_1}, \vec{a_5}, \vec{a_6}, \vec{R_1}, \vec{R_5}, \vec{R_6}) \end{aligned}$$

and $\psi = \psi(\vec{a_1}, \vec{R_1})$ where the $\vec{a_i}, \vec{R_j}$ are all disjoint. Then

$$w(\theta \mid \theta' \wedge \psi) \geq w(\theta \mid \theta'' \wedge \psi).$$

A Failed Attempt

For atoms $\alpha_i(x) = \bigwedge_{n=1}^{q} R_n^{\epsilon_n}(x)$, $\alpha_j(x) = \bigwedge_{n=1}^{q} R_n^{\delta_n}(x)$, where the $\epsilon_n, \delta_n \in \{0,1\}$ and $R^1 = R, R^0 = \neg R$,

$$\begin{aligned} |\alpha_i - \alpha_j| &= \sum_{n=1}^{q} |\epsilon_n - \delta_n| \\ &= \text{the number of conjuncts } R_n \text{ on which } \alpha_i, \alpha_j \text{ differ.} \end{aligned}$$

Principle of Analogical Support by Distance:

If $\theta(a_1, \ldots, a_n) \in QFSL_q$ and

$$|\alpha_i - \alpha_j| < |\alpha_i - \alpha_k|$$

then

$$w(\alpha_i(a_{n+2}) \,|\, \alpha_j(a_{n+1}) \wedge \theta(a_1, \ldots, a_n)) \geq w(\alpha_i(a_{n+2}) \,|\, \alpha_k(a_{n+1}) \wedge \theta(a_1, \ldots, a_n)).$$

Unfortunately the only solutions, even for $q = 2$, are hardly 'rational'.