# Binary Pure Inductive logic

Alena Vencovská

$L$ - predicate language (without equality, no function symbols) with constant symbols $a_1, a_2, a_3 \ldots$ and finitely many predicate (relation) symbols. $SL$ is the set of sentences of $L$.

A *state description* for constants $a_1, \ldots, a_n$ is a sentence $\Theta(a_1, \ldots, a_n)$ that decides all atomic sentences involving these constants. Replacing constants by variables yields a *state formula*.

If $r$ is the maximum arity of the language then state formulae for $r$ variables are called *atoms*. If clear from context, 'atoms' can also refer to state descriptions for $r$ constants.

# Exchangeability

Exchangeability (possible formulation). A probability function $w : SL \to [0,1]$ satisfies *Exchangeability, Ex* if for any state description $\Theta(a_1, \ldots, a_n)$, $w(\Theta(a_1, \ldots, a_n))$ is invariant under permutations of constants.

In the unary case, Ex can be expressed in terms of signatures of state descriptions.

For $L = \{R_1, \dots, R_q\}$, all unary, atoms are:

$$\bigwedge_{i=1}^{q} \pm R_i(x),$$

where $\pm R$ is one of $R$, $\neg R$. We denote them $\alpha_1(x), \dots, \alpha_{2^q}(x)$.

State descriptions are sentences of the form

$$\alpha_{i_1}(a_1) \wedge \dots \wedge \alpha_{i_n}(a_n). \tag{1}$$

Ex says that the probability of (1) depends only on its signature $\langle n_1, \dots, n_{2^q} \rangle$ where $n_j$ is the number of times that $\alpha_j$ appears amongst the $\alpha_{i_1}, \dots, \alpha_{i_n}$.

We now restrict ourselves to $L = \{R\}$, $R$ binary.

A state description for $a_1, a_2, \ldots, a_n$ is a sentence of the form

$$\bigwedge_{i,j=1}^{n} \pm R(a_i, a_j).$$

We usually represent this by an $n \times n$ $\{0, 1\}$-matrix with 1 or 0 at the $i$th row, $j$th column depending on whether $R(a_i, a_j)$ or $\neg R(a_i, a_j)$ appears in the state description respectively.

# Example

$$\neg R(a_1,a_1) \wedge R(a_1,a_2) \wedge R(a_1,a_3) \wedge R(a_2,a_1) \wedge R(a_2,a_2)$$

$$\wedge R(a_2,a_3) \wedge \ R(a_3,a_1) \wedge \neg R(a_3,a_2) \wedge R(a_3,a_3)$$

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

Atoms are represented by $2 \times 2$ $\{0,1\}$-matrices (there are 16 of them).

# Example continued

State descriptions are composed of atoms. For example,

$$\Psi(a_1, a_2, a_3): \quad \gamma(a_1, a_2) \quad \wedge \quad \gamma(a_1, a_3) \quad \wedge \quad \eta(a_2, a_3)$$

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

The binary signature of a state description should in some sense be the vector of numbers of atoms that it is composed of.

It needs to be invariant under permutation of constants.

# Example continued

If $\Phi(a_1, a_2, a_3) = \Psi(a_2, a_1, a_3)$ then we have

$$\Phi(a_1, a_2, a_3) : \quad \bar{\gamma}(a_1, a_2) \quad \wedge \quad \eta(a_1, a_3) \quad \wedge \quad \gamma(a_2, a_3)$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

For signatures of state descriptions to be invariant under permutations of constants, ‘permuted’ atoms need to be counted together.

# Binary Signature, $L = \{R\}$

Ten classes of atoms, for example

$$\left\{ \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \right\} \quad \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \right\}$$

$$\left\{ \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\} \quad \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\} \quad \dots$$

We denote them $\Gamma_1, \Gamma_2, \dots \Gamma_{10}$.

For $L = \{R\}$, the signature of a state description $\Theta(a_1, \dots, a_n)$ is the vector $\langle n_1, \dots, n_{10} \rangle$ where $n_g$ is the number of pairs $\langle a_i, a_j \rangle$ with $i < j$ such that $\Theta$ restricted to $a_i, a_j$ is in $\Gamma_g$.

# Binary Exchangeability

Signature for any other $L$ with at most binary predicates is defined analogously.

*Binary Exchangeability, BEx.* Probability of a state description depends only on its signature.

BEx $\Rightarrow$ Ex but not conversely.

For $L = \{R\}$, with $\vec{Y} = \langle x_1, x_2; y_1, \ldots, y_{10} \rangle$, give a state description represented by a matrix the probability with which we obtain this matrix when placing 0 or 1 on the diagonal with probability $x_1$, $x_2$ respectively and then complete it according to probabilities $y_g$.

Hence we need $x_i, y_g \in [0, 1]$, $x_1 + x_2 = 1$ and for each ordered pair $\langle k, c \rangle$ with $k, c \in \{0, 1\}$ the sum of those $y_g$ for $g$ such that $\Gamma_g$ contains one or two atoms extending

$$\begin{pmatrix} k & \\ & c \end{pmatrix}$$

is 1 (the $y_g$ with 2 such atoms are counted twice).

# Explicit formula for the $w_{\vec{Y}}$

$\mathbb{D}_L$ is the set of $\vec{Y}$ as above. For $\vec{Y} \in \mathbb{D}_L$,

$$w_{\vec{Y}}(\Theta(a_1, \dots, a_n)) = x_1^{m_1} x_2^{m_2} \prod_{g=1}^{10} y_g^{n_g}$$

where $m_1, m_2$ is the number of times 0, 1 appears on the diagonal of the matrix representing $\Theta$ respectively and $\langle n_1, \dots, n_{10} \rangle$ is the signature of $\Theta$.

The $w_{\vec{Y}}$ extend to probability functions on $SL$.

For other languages with at most binary predicates the signature, $\mathbb{D}_L$ and the $w_{\vec{Y}}$ are defined analogously.

*Theorem.* Let $L$ contain at most binary predicates and let $w$ be a probability satisfying BEx. Then there exists a (normalised, $\sigma$-additive) measure $\mu$ on the Borel subsets of $\mathbb{D}_L$ such that for any $\theta \in SL$,

$$w(\theta) = \int_{\mathbb{D}_L} w_{\vec{Y}}(\theta) d\mu(\vec{Y}) \qquad (2)$$

Conversely, for a given measure $\mu$ on the Borel subsets of $\mathbb{D}_L$, the function defined by (2) is a probability function on SL satisfying BEx.

# Applications

This representation allows us to prove various results on conditioning for the BEx functions, similar to those leading to the Conditionalization theorem for the unary probability functions satisfying Ex. They apply to languages with binary predicates (any number) and possibly some unary predicates.

We say that $\vec{Y} \in \mathbb{D}_L$ is *extreme* if all $x_i$, $y_g$ are 0 or 1.

We say that $\vec{B} \in \mathbb{D}_L$ is a *support point* of $\mu$ if every open neighborhood of $\vec{B}$ has measure $> 0$. The set of support points of $\mu$ is its *support set*.

# Results

Let $w$ be a probability function satisfying BEx with de Finetti prior $\mu$ and let $\vec{B}_1, \vec{B}_2, \ldots, \vec{B}_k$ be distinct non-extreme support points of $\mu$. Let $\gamma_1, \gamma_2, \ldots, \gamma_k > 0$, $\sum_{j=1}^k \gamma_j = 1$ and let $v = \sum_{j=1}^k \gamma_j w_{\vec{B}_j}$. Then there exist quantifier-free sentences $\xi_n(a_1, \ldots, a_{t_n})$ such that for any quantifier-free sentence $\psi(a_1, \ldots, a_m)$,

$$\lim_{n\to\infty} w\left(\psi(a_{t_n+1}, \ldots, a_{t_n+m}) \,|\, \xi_n(a_1, \ldots, a_{t_n})\right) = v(\psi(a_1, \ldots, a_m)).$$

Let $w$ be a probability function on $SL$ satisfying BEx, with de Finetti prior $\mu$ and let $\epsilon > 0$. Then there are points $\vec{D}_1, \ldots, \vec{D}_k$ in the support of $\mu$, and $\gamma_j > 0$ with $\sum_j \gamma_j = 1$ such that for $v = \sum_j \gamma_j w_{\vec{D}_j}$,

$$|w(\psi(a_1, \ldots, a_m)) - v(\psi(a_1, \ldots, a_m))| < m^3 |SD(m)| \epsilon$$

for all quantifier-free sentences $\psi(a_1, \ldots, a_m)$.

($SD(m)$ is the number of state descriptions of $L$ for $m$ constants).

# Results

Let $w_1$ be probability a function satisfying BEx such that the support set of its de Finneti prior is $\mathbb{D}_L$. Then for any $w_2$ satisfying BEx and for a given $m$ and $\nu > 0$, there is a quantifier-free $\theta(a_1, \dots, a_n)$ such that for each quantifier-free $\psi(a_1, \dots, a_m)$,

$$|w_1(\psi(a_{n+1}, \dots, a_{n+m}) \,|\, \theta(a_1, \dots, a_n)) \\ - w_2(\psi(a_1, \dots, a_m))| < \nu.$$

Informally, within some prescribed accuracy and restricting ourselves to quantifier-free sentences involving only a limited number of constants, provided that the support set of the prior of $w_1$ is $\mathbb{D}_L$, we can obtain from it by conditioning any other $w_2$ satisfying BEx.