Bayesian Networks: the Parental Synergy

Janneke H. Bolt

September 18th 2008

Universiteit Utrecht

Outline

- Definition
- Context
- Concluding remarks

The indicator function δ

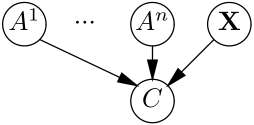
The indicator function δ compares a value assignment to a (set of) node(s) with some specific value assignment to the(se) node(s). It results in 1 if the number of differences is even and in -1 if the number of differences is odd.

The indicator function δ: example

Consider the ternary variables A and B. Given the value assignment $A = a_2$ and $B = b_2$ is found that:

$$
\begin{aligned}
\delta(a_1 b_1 \mid a_2 b_2) &= +1 \\
\delta(a_1 b_2 \mid a_2 b_2) &= -1 \\
\delta(a_1 b_3 \mid a_2 b_2) &= +1 \\
\delta(a_2 b_1 \mid a_2 b_2) &= -1 \\
\delta(a_2 b_2 \mid a_2 b_2) &= +1 \\
\delta(a_2 b_3 \mid a_2 b_2) &= -1 \\
\delta(a_3 b_1 \mid a_2 b_2) &= +1 \\
\delta(a_3 b_2 \mid a_2 b_2) &= -1 \\
\delta(a_3 b_3 \mid a_2 b_2) &= +1
\end{aligned}
$$

The parental synergy: definition

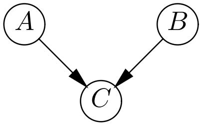


The parental synergy of **a** with respect to c_i given $\mathbf{X} = \mathbf{x}$, denoted as $Y^{\star}_{\mathbf{x}}(\mathbf{a}, c_i)$, is

$$Y^{\star}_{\mathbf{x}}(\mathbf{a}, c_i) = \sum_{\mathbf{A}} \delta(\mathbf{A} \mid \mathbf{a}) \cdot \Pr(c_i \mid \mathbf{A}\mathbf{x})$$

The parental synergy: example

AB	$\Pr(c \mid AB)$	$\delta(AB \mid a_2 b_2)$
$a_1 b_1$	0.7	$+1$
$a_1 b_2$	0.2	-1
$a_1 b_3$	0.4	$+1$
$a_2 b_1$	0.2	-1
$a_2 b_2$	1.0	$+1$
$a_2 b_3$	0.1	-1
$a_3 b_1$	0.3	$+1$
$a_3 b_2$	0.8	-1
$a_3 b_3$	0.9	$+1$



$$Y^{\star}(a_2 b_2, c) = \sum_{AB} \delta(AB \mid a_2 b_2) \cdot \Pr(c \mid AB) =$$

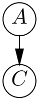
$$0.7 - 0.2 + 0.4 - 0.2 + 1.0 - 0.1 + 0.3 - 0.8 + 0.9 = 2.0$$

Interpretation (binary nodes)

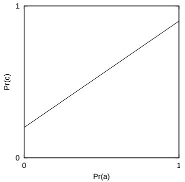
$Y^{\star}(c) = 0$

Interpretation (binary nodes)

$$Y^{\star}(c) = 0$$



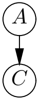
$$Y^{\star}(a, c) = \Pr(c \mid a) - \Pr(c \mid \bar{a})$$



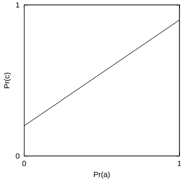
$\Pr(c \mid \bar{a}) = 0.2$; $\Pr(c \mid a) = 0.9$

Interpretation (binary nodes)

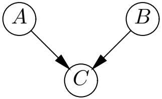
$Y^{\star}(c) = 0$



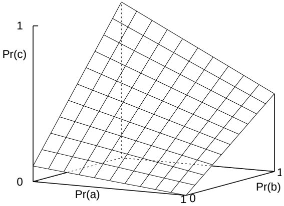
$Y^{\star}(a, c) =$
$\Pr(c \mid a) - \Pr(c \mid \bar{a})$



$\Pr(c \mid \bar{a}) = 0.2$; $\Pr(c \mid a) = 0.9$

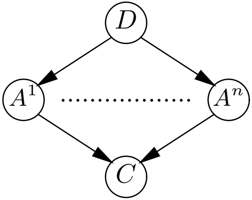


$Y^{\star}(ab, c) =$
$\Pr(c \mid ab) - \Pr(c \mid a\bar{b}) -$
$(\Pr(c \mid \bar{a}b) - \Pr(c \mid \bar{a}\bar{b}))$



$\Pr(c \mid ab) = 0.5$; $\Pr(c \mid a\bar{b}) = 0$;
$\Pr(c \mid \bar{a}b) = 1$; $\Pr(c \mid \bar{a}\bar{b}) = 0.1$

The prior convergence error



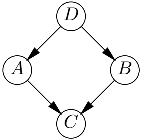
$$\Pr(c_i) = \sum_{\mathbf{A}} \Pr(c_i \mid \mathbf{A}) \cdot \Pr(\mathbf{A})$$

$$\widetilde{\Pr}(c_i) = \sum_{\mathbf{A}} \Pr(c_i \mid \mathbf{A}) \cdot \Pr(A^1) \cdot \ldots \cdot \Pr(A^n)$$

The prior convergence error equals

$$\Pr(c_i) - \widetilde{\Pr}(c_i)$$

The prior convergence error; binary nodes; two parent nodes



$$\Pr(c_i) - \widetilde{\Pr}(c_i) = l \cdot m \cdot n \cdot w$$

where

$$\begin{aligned} l &= \Pr(d) - \Pr(d)^2 \\ m &= \Pr(a \mid d) - \Pr(a \mid \bar{d}) \\ n &= \Pr(b \mid d) - \Pr(b \mid \bar{d}) \\ w &= \Pr(c_i \mid ab) - \Pr(c_i \mid a\bar{b}) - \Pr(c_i \mid \bar{a}b) + \Pr(c_i \mid \bar{a}\bar{b}) \end{aligned}$$

Note that

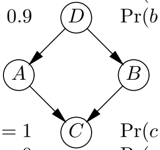
$$w = Y^{\star}(ab, c_i)$$

The prior convergence error; binary nodes; two parent nodes; illustration

$\Pr(d) = 0.5$

$\Pr(a \mid d) = 0.5$
$\Pr(a \mid \bar{d}) = 0.9$

$\Pr(b \mid d) = 0.1$
$\Pr(b \mid \bar{d}) = 0.9$



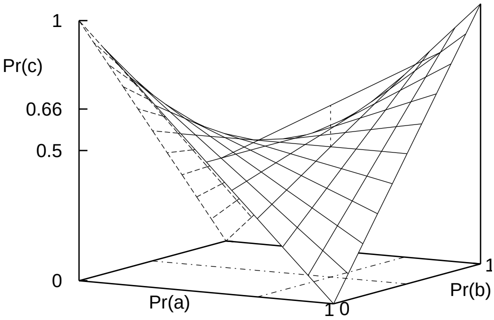
$\Pr(c \mid ab) = 1$
$\Pr(c \mid a\bar{b}) = 0$

$\Pr(c \mid \bar{a}b) = 0$
$\Pr(c \mid \bar{a}\bar{b}) = 1$

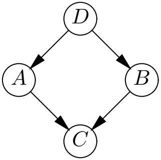
$$\Pr(c) - \widetilde{\Pr}(c) = l \cdot m \cdot n \cdot w$$

where

$$\begin{aligned} l &= \Pr(d) - \Pr(d)^2 \\ m &= \Pr(a \mid d) - \Pr(a \mid \bar{d}) \\ n &= \Pr(b \mid d) - \Pr(b \mid \bar{d}) \\ w &= \Pr(c \mid ab) - \Pr(c \mid a\bar{b}) \\ &\quad - \Pr(c \mid \bar{a}b) + \Pr(c \mid \bar{a}\bar{b}) \end{aligned}$$



The prior convergence error; binary nodes; two parent nodes; an alternative expression



$$\Pr(c_i) - \widetilde{\Pr}(c_i) = \big(\Pr(ab) - \Pr(a) \cdot \Pr(b)\big) \cdot w$$

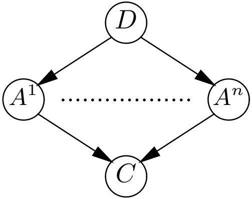
where

$$w = \Pr(c_i \mid ab) - \Pr(c_i \mid a\bar{b}) - \Pr(c_i \mid \bar{a}b) + \Pr(c_i \mid \bar{a}\bar{b})$$

Note that

$$w = Y^{\star}(ab, c_i)$$

The prior convergence error; binary nodes; an arbitrary number of parent nodes

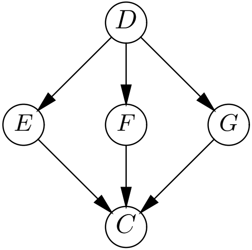


$$\Pr(c_i) - \widetilde{\Pr}(c_i) = \sum_{\mathbf{m}} \left(\Pr(a^x \ldots a^y) - \Pr(a^x) \cdot \ldots \cdot \Pr(a^y) \right) \cdot Y^{\star}_{\bar{a}^1 \ldots \bar{a}^n \setminus \bar{\mathbf{a}}^{\mathbf{m}}}(\mathbf{a}^{\mathbf{m}}, c_i)$$

where

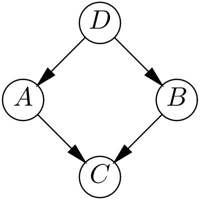
$$\mathbf{m} \in \mathcal{P}(\{1, \ldots, n\}) = \{x, \ldots, y\}$$
$$\mathbf{a}^{\mathbf{m}} = a^x \ldots a^y$$
$$\bar{\mathbf{a}}^{\mathbf{m}} = \bar{a}^x \ldots \bar{a}^y$$

The prior convergence error; binary nodes; an arbitrary number of parent nodes; example



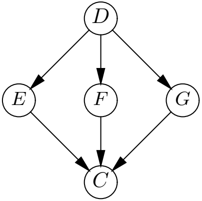
$$
\begin{aligned}
\Pr(c) - \widetilde{\Pr}(c) \quad = \quad & (\Pr(efg) - \Pr(e)\cdot\Pr(f)\cdot\Pr(g))\cdot Y^{\star}(efg, c) + \\
& (\Pr(ef) - \Pr(e)\cdot\Pr(f))\cdot Y^{\star}_{\bar{g}}(ef, c) + \\
& (\Pr(eg) - \Pr(e)\cdot\Pr(g))\cdot Y^{\star}_{\bar{f}}(eg, c) + \\
& (\Pr(fg) - \Pr(f)\cdot\Pr(g))\cdot Y^{\star}_{\bar{e}}(fg, c)
\end{aligned}
$$

The prior convergence error; multiple-valued nodes; two parent nodes



$$\Pr(c_i) - \widetilde{\Pr}(c_i) = \sum_{AB} \big(\Pr(AB) - \Pr(A) \cdot \Pr(B) \big) \cdot Y^{\star}(AB, c_i)/4$$

The prior convergence error; multiple-valued nodes; arbitrary number of parent nodes



$$\Pr(c_i) - \widetilde{\Pr}(c_i) =$$

$$\sum_{\mathbf{m}} \Big[\sum_{\mathbf{A^m}} \Big((\Pr(A^x \ldots A^y) - \Pr(A^x) \cdot \ldots \cdot \Pr(A^y)) \cdot \sum_{A^1, \ldots, A^n \setminus \mathbf{A^m}} Y^{\star}_{A^1, \ldots, A^n \setminus \mathbf{A^m}}(\mathbf{A^m}, c_i) \Big) \Big] / 2^n$$

where

$$\mathbf{m} \in \mathcal{P}(\{1, \ldots, n\}) = \{x, \ldots, y\}$$
$$\mathbf{A^m} = A^x \ldots A^y$$

To conclude

The parental synergy is an important factor in the expression of the prior convergence error. It determines the impact of the degree of dependency between the parent nodes on the size of the convergence error.

Questions?