Bayesian Networks: a Combined Tuning Heuristic

Janneke H. Bolt

2016

Universiteit Utrecht

Outline

Bayesian network tuning

Sensitivity functions

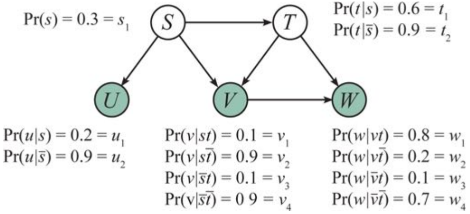
Tuning heuristics

Distance measures

Experiments

Bayesian network tuning

Adaptation of one or more parameters of a Bayesian network to enforce output to meet some constraint.



For example: the current value of $\Pr(st|uvw)$ is 0.045 but it should be 0.25. Tuning by adapting multiple parameters may be preferred.

Sensitivity functions

A sensitivity function describes an output probability of a network as a function of one or more of its parameters. The general form of a sensitivity function is:

$$\Pr(\mathbf{w}|\mathbf{u})(\mathbf{z}) = \frac{\sum_{\mathbf{z}_k \in \mathcal{P}(\mathbf{z})} \left(c_k \cdot \prod_{z_i \in \mathbf{z}_k} z_i \right)}{\sum_{\mathbf{z}_k \in \mathcal{P}(\mathbf{z})} \left(d_k \cdot \prod_{z_i \in \mathbf{z}_k} z_i \right)}$$

A two-way sensitivity function in the parameters z_1 and z_2 for example, has the following general form:

$$\Pr(\mathbf{w}|\mathbf{u})(z_1, z_2) = \frac{c_0 + c_1 \cdot z_1 + c_2 \cdot z_2 + c_3 \cdot z_1 \cdot z_2}{d_0 + d_1 \cdot z_1 + d_2 \cdot z_2 + d_3 \cdot z_1 \cdot z_2}$$

Tuning heuristics

When a network is tuned by changing multiple parameters a choice has to be made with respect to the relative amount of parameter change.

- Locally optimal
- Balanced
- Combined

Tuning heuristics: locally optimal

Consider a set parameters $\mathbf{z} = \{x, y_1, \ldots, y_{n-1}\}$. In a locally optimal tuning scheme, with respect to $\Pr(\mathbf{w} | \mathbf{u})$, all parameters $y_i \in \mathbf{z}$ are tied to x by:

$$y_i = \frac{sy_i^o}{sx^o} \cdot (x - x^o) + y_i^o$$

where sx^o and sy_i^o are the values of the partial derivatives of $\Pr(\mathbf{w} | \mathbf{u})(\mathbf{z})$ with respect to x and y_i at $\mathbf{z}^o$, respectively.

Advantage: locally optimal.

Disadvantage: the tuning range may be reduced.

Tuning heuristics: balanced

Consider a set parameters $\mathbf{z} = \{x, y_1, \ldots, y_{n-1}\}$. In a balanced tuning scheme, with respect to $\Pr(\mathbf{w}|\mathbf{u})$, all parameters $y_i \in \mathbf{z}$ are tied to x by

$$y_i = \begin{cases} \dfrac{x \cdot (x^o - 1) \cdot y_i^o}{x^o \cdot (x - 1 + y_i^o) - x \cdot y_i^o} & \text{if } y_i \updownarrow x \\ \dfrac{(x-1) \cdot x^o \cdot y_i^o}{-x + x \cdot x^o + x \cdot y_i^o - x^o \cdot y_i^o} & \text{if } y_i \not\updownarrow x \end{cases}$$

If $y_i \updownarrow x$, y_i and x simultaneously increase/decrease and if $y_i \not\updownarrow x$, an increase of x results in an decrease of y_i and vice versa. Moreover, x and y_i vary simultaneously in the entire interval $\langle 0, 1 \rangle$.

Advantage: in general a larger tuning range than the locally optimal heuristic.

Disadvantage: changes are locally not optimal.

Tuning heuristics: combined

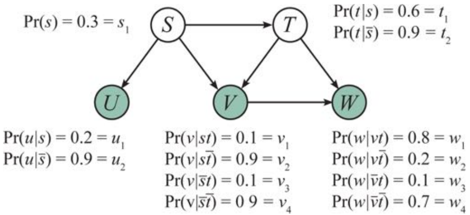
Consider a set parameters $\mathbf{z} = \{x, y_1, \ldots, y_{n-1}\}$. In a combined tuning scheme, with respect to $\Pr(\mathbf{w}|\mathbf{u})$, all parameters $y_i \in \mathbf{z}$ are tied to x by:

$$y_i = \begin{cases} \dfrac{x \cdot (sx^o \cdot y_i^o \cdot (y_i^o - 1) + sy_i^o \cdot (1 - x^o)) + sx^o \cdot y_i^o \cdot (1 - y_i^o) - sy_i^o \cdot x^o \cdot (1 - x^o)}{x \cdot (sx^o \cdot (y_i^o - 1) + sy_i^o \cdot (1 - x^o)) + sx^o \cdot (1 - y_i^o) - sy_i^o \cdot x^o \cdot (1 - x^o)} \\ \dfrac{x \cdot sx^o \cdot {y_i}^{o2}}{x \cdot (sx^o \cdot y_i^o - sy_i^o \cdot x^o) + sy_i^o \cdot x^{o2}} \\ \dfrac{x \cdot sx^o \cdot {y_i}^{o2} - sx^o \cdot {y_i}^{o2}}{x \cdot (sy_i^o \cdot (1 - x^o) + sx^o \cdot y_i^o) + sy_i^o \cdot x^o \cdot (x^o - 1) - sx^o \cdot y_i^o} \\ \dfrac{x \cdot (sx^o \cdot y_i^o \cdot (y_i^o - 1) - sy_i^o \cdot x^o) + sy_i^o \cdot x^{o2}}{x \cdot (sx^o \cdot (y_i^o - 1) - sy_i^o \cdot x^o) + sy_i^o \cdot x^{o2}} \end{cases}$$

where $sx^o, sy_i^o, \ldots, sy_{n-1}^o$ are the values of the partial derivatives of $\Pr(\mathbf{w}|\mathbf{u})(\mathbf{z})$ with respect to respectively $x, y_1, \ldots, y_{n-1}$ at $\mathbf{z}^o$.

This heuristic is locally optimal, yet, it covers the same tuning range as the balanced heuristic.

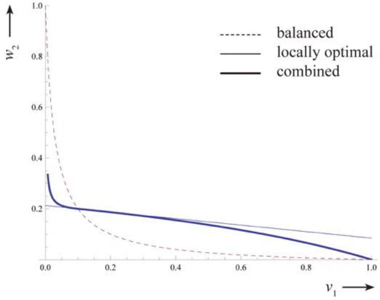
Example

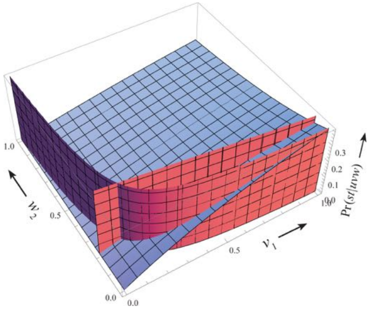


$$\Pr(st \,|\, uvw)(v_1, w_2) = \frac{2.88 \cdot v_1}{2.88 \cdot v_1 + 7.83 \cdot w_2 + 4.54}$$

Example

$$\Pr(st \,|\, uvw)(v_1, w_2) = \frac{2.88 \cdot v_1}{2.88 \cdot v_1 + 7.83 \cdot w_2 + 4.54}$$



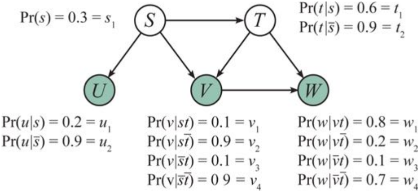


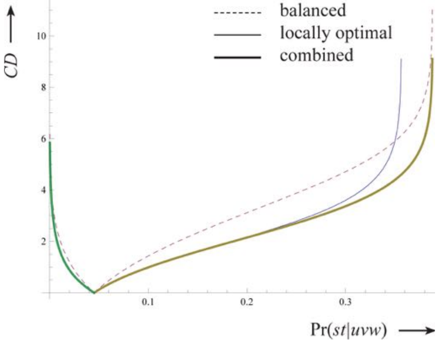
Distance measures

$$\mathrm{KL} = \sum_{\mathbf{v} \in \mathbf{V}} \mathsf{Pr}^{o}(\mathbf{v}) \cdot \ln \frac{\mathsf{Pr}^{o}(\mathbf{v})}{\mathsf{Pr}(\mathbf{v})}$$

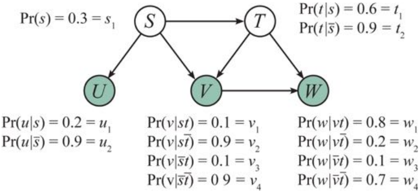
$$\mathrm{CD} = \ln \max_{\mathbf{v} \in \mathbf{V}} \frac{\mathsf{Pr}(\mathbf{v})}{\mathsf{Pr}^{o}(\mathbf{v})} - \ln \min_{\mathbf{v} \in \mathbf{V}} \frac{\mathsf{Pr}(\mathbf{v})}{\mathsf{Pr}^{o}(\mathbf{v})}$$

Example


$\Pr(s) = 0.3 = s_1$
S
T
$\Pr(t|s) = 0.6 = t_1$
$\Pr(t|\bar{s}) = 0.9 = t_2$
U
V
W
$\Pr(u|s) = 0.2 = u_1$
$\Pr(u|\bar{s}) = 0.9 = u_2$
$\Pr(v|st) = 0.1 = v_1$
$\Pr(v|s\bar{t}) = 0.9 = v_2$
$\Pr(v|\bar{s}t) = 0.1 = v_3$
$\Pr(v|\bar{s}\bar{t}) = 0\ 9 = v_4$
$\Pr(w|vt) = 0.8 = w_1$
$\Pr(w|v\bar{t}) = 0.2 = w_2$
$\Pr(w|\bar{v}t) = 0.1 = w_3$
$\Pr(w|\bar{v}\bar{t}) = 0.7 = w_4$


balanced
locally optimal
combined
CD
10
8
6
4
2
0.1
0.2
0.3
$\Pr(st|uvw)$

Experiments



$$\begin{aligned} \text{set1} &= \{v_1, u_2, w_2\} \\ \text{set2} &= \{v_1, v_2, v_3, v_4\} \\ \text{set3} &= \{v_1, v_2, u_1, u_2, w_1, w_2\} \end{aligned}$$

	'uniform'			'biased'		
	set 1	set 2	set 3	set 1	set 2	set 3
$\ln(OR) = -2$	31	35	13	40	43	10
$\ln(OR) = -1$	19	14	6	30	34	5
$\ln(OR) = -0.5$	9	7	3	21	22	3
$\ln(OR) = 0.5$	9	9	6	22	35	9
$\ln(OR) = 1$	24	32	13	32	45	18
$\ln(OR) = 2$	50	68	35	53	58	35

Experiments

	CD-distance, 'uniform parameterisation'								
	L vs B			L vs C			B vs C		
	set 1	set 2	set 3	set 1	set 2	set 3	set 1	set 2	set 3
$\ln(OR) = -2$	L	L	L	C	C	L	C	C	C
$\ln(OR) = -1$	L	L	L	C	C	L	C	C	C
$\ln(OR) = -0.5$	L	L	L	C	C	L	C	C	C
$\ln(OR) = 0.5$	L	B	L	C	C	C	C	C	C
$\ln(OR) = 1$	L	B	L	C	C	C	C	C	C
$\ln(OR) = 2$	B	B	B	C	C	C	C	B	C

	CD-distance, 'biased parameterisation'								
	L vs B			L vs C			B vs C		
	set 1	set 2	set 3	set 1	set 2	set 3	set 1	set 2	set 3
$\ln(OR) = -2$	B	B	L	C	C	L	C	C	C
$\ln(OR) = -1$	L	L	L	C	C	L	C	C	C
$\ln(OR) = -0.5$	L	L	L	C	C	L	C	C	C
$\ln(OR) = 0.5$	L	B	L	C	C	C	C	C	C
$\ln(OR) = 1$	L	B	L	C	C	C	C	C	C
$\ln(OR) = 2$	B	B	L	C	C	C	C	C	C

Questions?