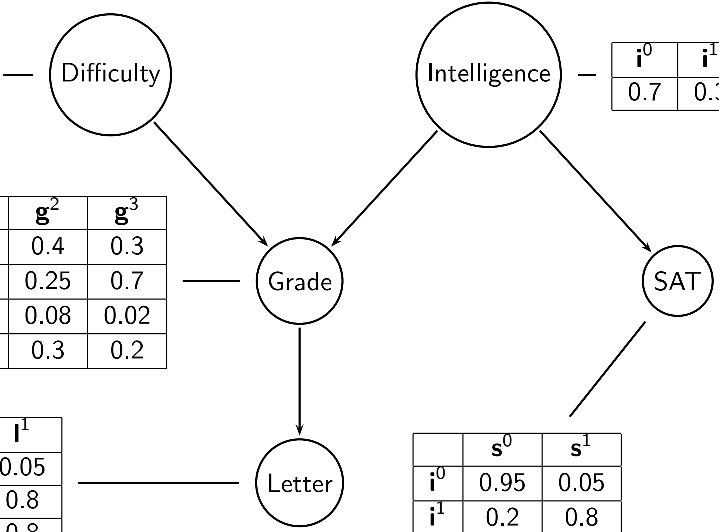
The max-min-hill-climbing algorithm

Michael Bauer

M.Sc. Comp. Science

May 3, 2014



d^0	d^1
0.6	0.4

i^0	i^1
0.7	0.3

	g^1	g^2	g^3
i^0, d^0	0.3	0.4	0.3
i^0, d^1	0.05	0.25	0.7
i^1, d^0	0.9	0.08	0.02
i^1, d^1	0.5	0.3	0.2

	l^0	l^1
g^1	0.95	0.05
g^2	0.2	0.8
g^3	0.2	0.8

	s^0	s^1
i^0	0.95	0.05
i^1	0.2	0.8

Reminder

Definition (Independence)

Let A, B denote random variables. Then A and B are independent iff

$$P(A \cap B) = P(A) * P(B). \tag{1}$$

Definition (Conditional Probability)

Let A, B denote random variables and $P(B) > 0$. The probability of A given B is defined as:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{2}$$

Definition (Conditional Independence)

Two variables X and Y are conditionally independent given $\mathbf{Z}$ w.r.t a probability distribution P, denoted as $Ind_p(X; Y|\mathbf{Z})$, if $\forall x, y, \mathbf{z}$, where $P(\mathbf{Z} = \mathbf{z}) > 0$,

$$P(X, Y|\mathbf{Z}) = P(X|\mathbf{Z}) * P(Y|\mathbf{Z}), \tag{3}$$

where $P(X, Y|\mathbf{Z}) = P(X \cap Y|\mathbf{Z})$.

It is equivalent

$$Ind(X; T|\mathbf{Z}) \Longleftrightarrow (Assoc(X; T|\mathbf{Z}) = 0), \tag{4}$$

where $Assoc(X; T|\mathbf{Z})$ is the strength association (dependency) of X and T given $\mathbf{Z}$.

Definition (Conditional Independence)

Two variables X and Y are conditionally independent given $\mathbf{Z}$ w.r.t a probability distribution P, denoted as $Ind_p(X;Y|\mathbf{Z})$, if $\forall x, y, \mathbf{z}$, where $P(\mathbf{Z} = \mathbf{z}) > 0$,

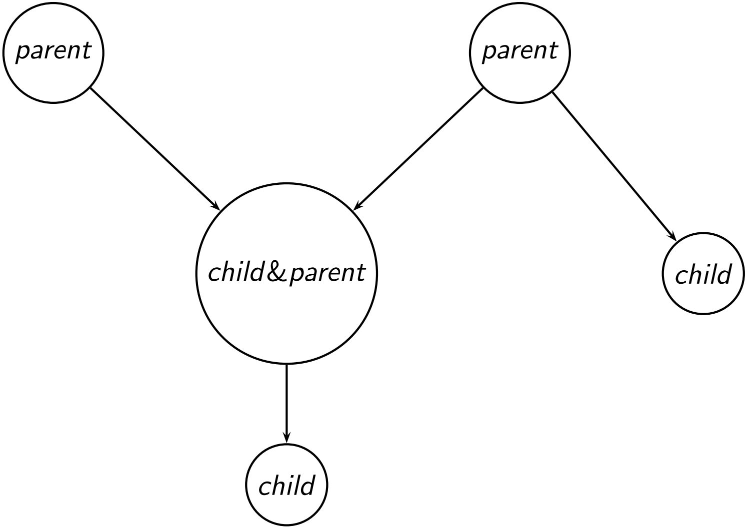
$$P(X, Y|\mathbf{Z}) = P(X|\mathbf{Z}) * P(Y|\mathbf{Z}), \tag{3}$$

where $P(X, Y|\mathbf{Z}) = P(X \cap Y|\mathbf{Z})$.
It is equivalent

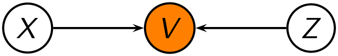
$$Ind(X; T|\mathbf{Z}) \iff (Assoc(X; T|\mathbf{Z}) = 0), \tag{4}$$

where $Assoc(X; T|\mathbf{Z})$ is the strength association (dependency) of X and T given $\mathbf{Z}$.

Directed Acyclic Graphs & Bayesian Networks



Collider

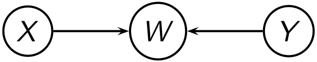


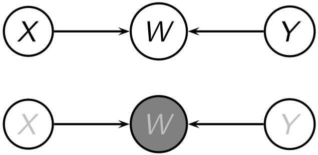
Blocked paths

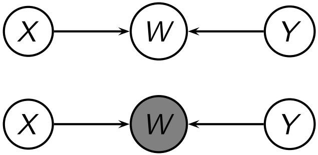
Define four sets

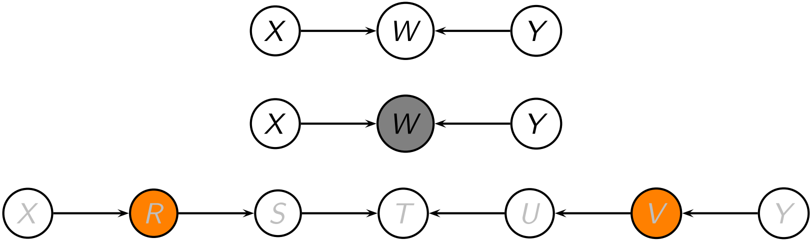
Let Z_1, Z_2, Z_3 and Z_4 denote sets with:

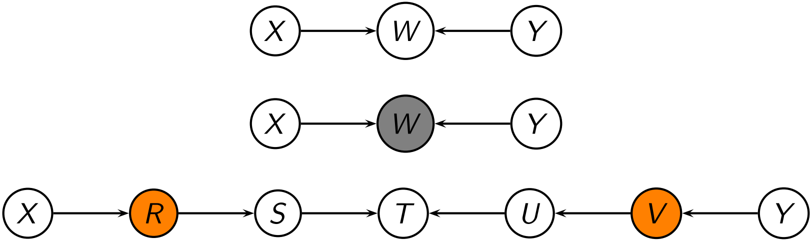
$$\begin{aligned} Z_1 &= \{\emptyset\} \\ Z_2 &= \{W\} \\ Z_3 &= \{R, V\} \\ Z_4 &= \{R, P\} \end{aligned}$$

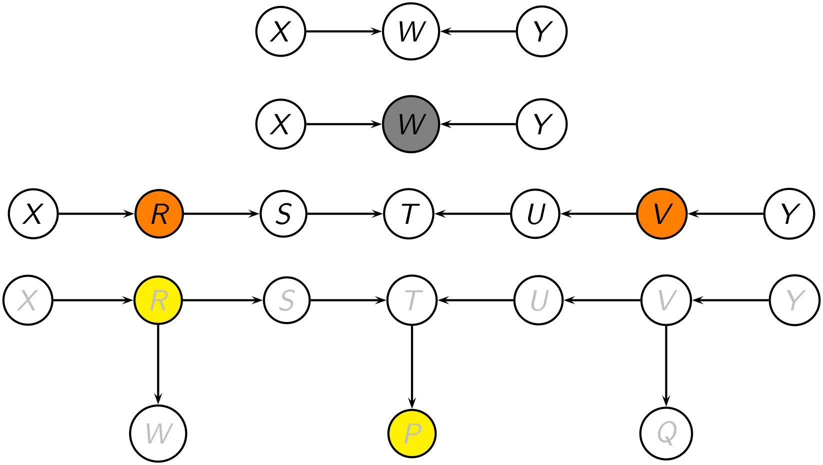

X
W
Y

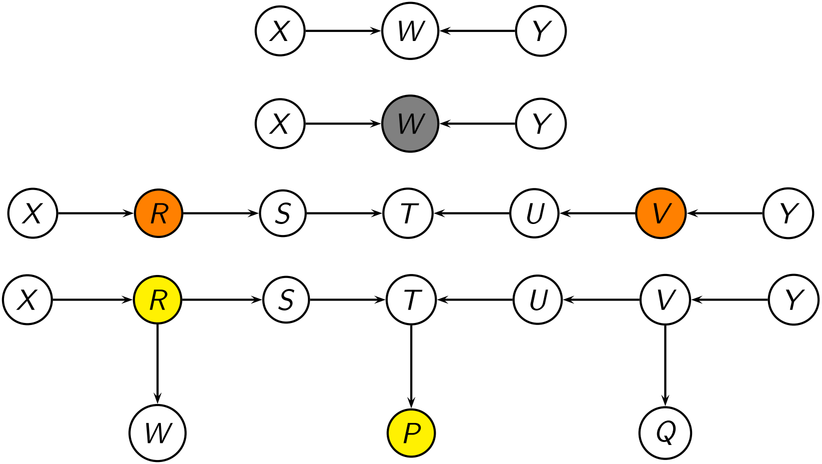

X
W
Y
X
W
Y


X
W
Y
X
W
Y


X
W
Y
X
W
Y
X
R
S
T
U
V
Y


X
W
Y
X
W
Y
X
R
S
T
U
V
Y

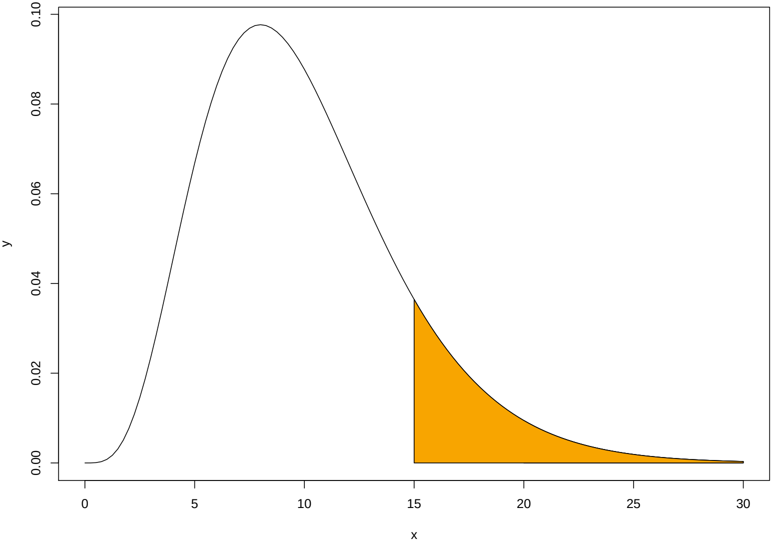

X
W
Y
X
W
Y
X
R
S
T
U
V
Y
X
R
S
T
U
V
Y
W
P
Q

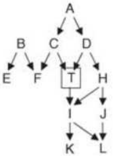

X
W
Y
X
W
Y
X
R
S
T
U
V
Y
X
R
S
T
U
V
Y
W
P
Q

Definition

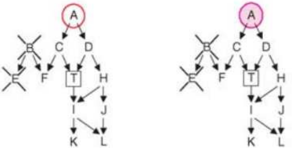
We calculate the G^2 value under the nullhypothesis of the conditional independence of $Ind_P(X_i, X_j|\mathbf{X}_k)$ holding. Then, the G^2 value is defined as:

$$G^2 := 2 * \sum_{a,b,c} S_{ijk}^{ab\mathbf{c}} * ln\left(\frac{S_{ijk}^{ab\mathbf{c}} * S_k^{\mathbf{c}}}{S_{ik}^{a\mathbf{c}} * S_{jk}^{b\mathbf{c}}}\right) \tag{5}$$


0.00
0.02
0.04
0.06
0.08
0.10
y
0
5
10
15
20
25
30
x



(a) Generating Network

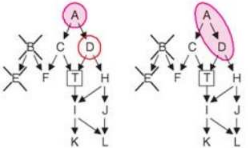


(b) (c)

Iteration 1

(b) A is selected. B & E are removed since $Ind(T; B|\{\})$ and $Ind(T; E|\{\})$.

(c) **CPC** $= \{A\}$

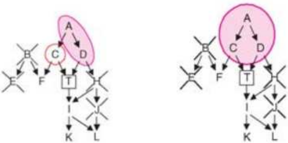


(d) (e)

Iteration 2

(d) D is selected.

(e) **CPC** $= \{A, D\}$

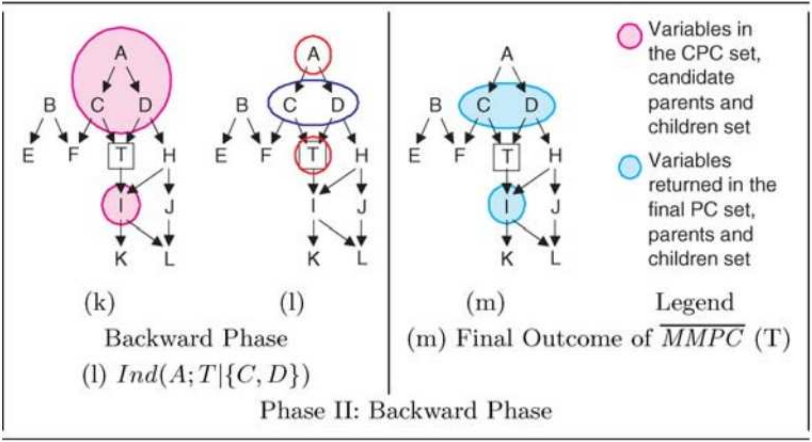


(f) (g)

Iteration 3

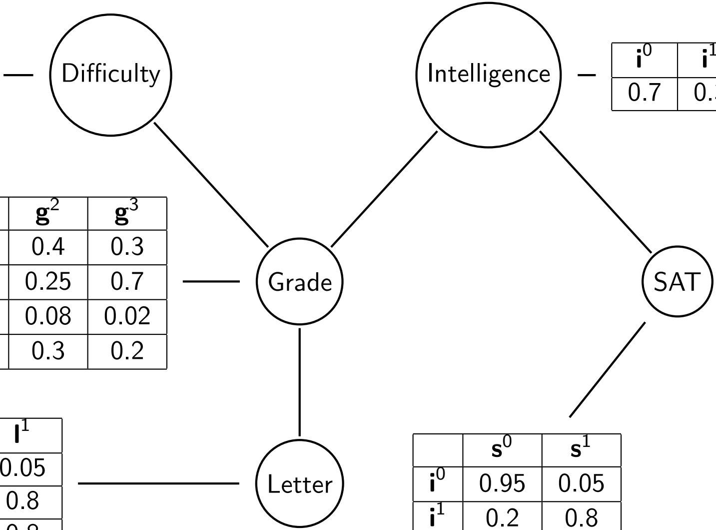
(f) C is selected. H & J are removed since $Ind(T; H|\{D\})$ and $Ind(T; J|\{D\})$.

(g) **CPC** $= \{A, C, D\}$


A
B
C
D
E
F
T
H
I
J
K
L
Variables in the CPC set, candidate parents and children set
Variables returned in the final PC set, parents and children set
(k)
(l)
(m)
Legend
Backward Phase
(l) $Ind(A;T|\{C,D\})$
(m) Final Outcome of $\overline{MMPC}$ (T)
Phase II: Backward Phase

d^0	d^1
0.6	0.4

i^0	i^1
0.7	0.3



	g^1	g^2	g^3
i^0,d^0	0.3	0.4	0.3
i^0,d^1	0.05	0.25	0.7
i^1,d^0	0.9	0.08	0.02
i^1,d^1	0.5	0.3	0.2

	l^0	l^1
g^1	0.95	0.05
g^2	0.2	0.8
g^3	0.2	0.8

	s^0	s^1
i^0	0.95	0.05
i^1	0.2	0.8

Thanks for your attention!