The max-min-hill-climbing algorithm

Michael Bauer

M.Sc. Comp. Science

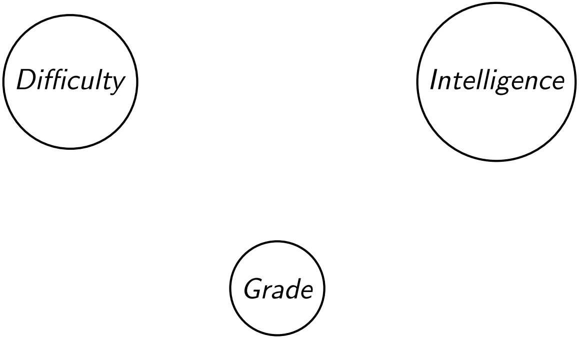
May 4, 2014

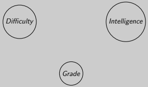
The max-min-hill-climbing algorithm

Michael Bauer
M.Sc. Comp. Science

May 4, 2014

- the max-min part is a constrained-based part
- hill-climbing is a greedy (local) search


Difficulty
Intelligence
Grade



- What is the algorithm about?
- Why do we need it?
- Where do I start from?
- having a matrix (dataframe in R) with observed date which follow probabilistic distributions
- talk about picture

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 $P(\mathbf{Z} = \mathbf{z}) > 0$,

$$P(X, Y|\mathbf{Z}) = P(X|\mathbf{Z}) * P(Y|\mathbf{Z}), \tag{1}$$

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

It is equivalent

$$Ind(X; Y|\mathbf{Z}) \Longleftrightarrow (Assoc(X; Y|\mathbf{Z}) = 0), \tag{2}$$

where $Assoc(X; Y|\mathbf{Z})$ is the strength association (dependency) of X and Y 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 $P(\mathbf{Z} = \mathbf{z}) > 0$,

$$P(X, Y|\mathbf{Z}) = P(X|\mathbf{Z}) * P(Y|\mathbf{Z}), \tag{1}$$

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

$$Ind(X; Y|\mathbf{Z}) \iff (Assoc(X; Y|\mathbf{Z}) = 0), \tag{2}$$

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

Definition (Conditional Independence)

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

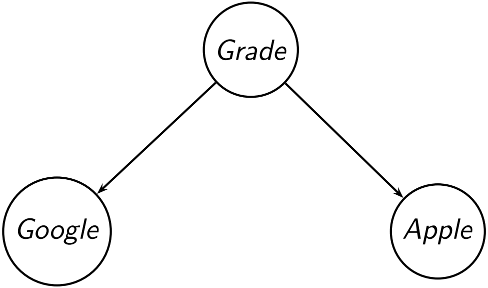
$$P(X, Y|\mathbf{Z}) = P(X|\mathbf{Z}) * P(Y|\mathbf{Z}), \tag{1}$$

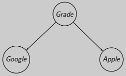
where $P(X, Y|\mathbf{Z}) = P(X \cap Y|\mathbf{Z})$.
It is equivalent

$$Ind(X; Y|\mathbf{Z}) \Longleftrightarrow (Assoc(X; Y|\mathbf{Z}) = 0), \tag{2}$$

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

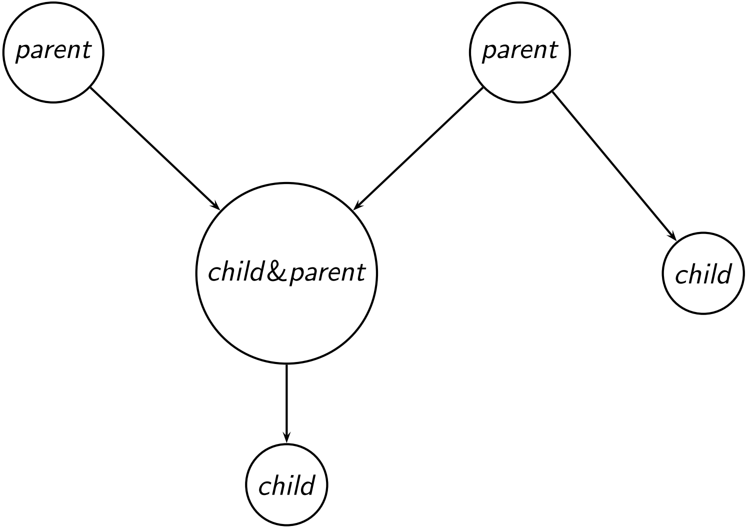
- <u>Independence:</u> coin toss; is a useful concept but not common. Most of the time two events are independent given an additional event
- <u>conditional probability:</u> pick a student -¿ intelligence -¿ grade
- <u>conditional independence:</u>
  - say admitted to Google, admitted to Apple
  - in most reasonable distributions these two events are not independent
  - assume they base their decision only on the grade
  - we assume that (pick one) student has grade A
  - the grade holds the relevant information
  - Google is conditional independent of Apple given grade A
- <u>Assoc:</u> Do not forget about Assoc -¿ Min and Max


Grade
Google
Apple



- DAG
- vertice - random variable
- edge - follow probability distributions

Bayesian Networks



Blocked paths & d-seperation

Define three sets

Let A_1, A_2 and A_3 denote sets with:

$$\begin{aligned} A_1 &= \{W\} \\ A_2 &= \{R, V\} \\ A_3 &= \{R, P\} \end{aligned}$$

Blocked paths & d-seperation

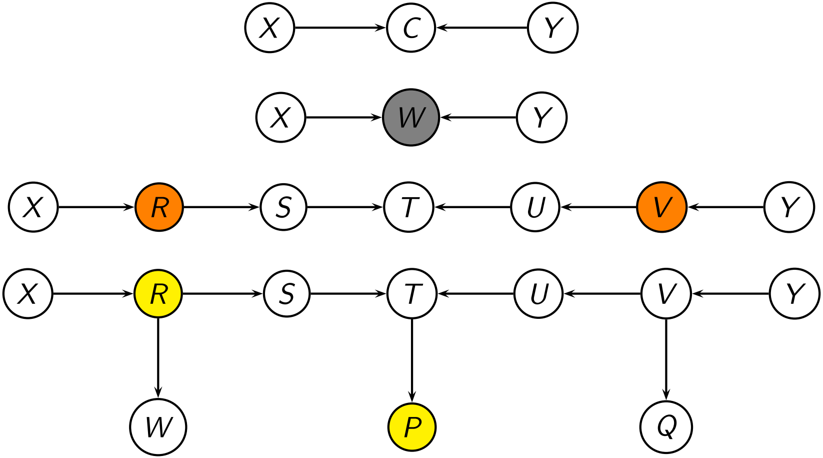
Define three sets

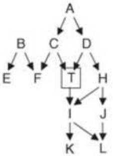
Let A_1, A_2 and A_3 denote sets with:

$$\begin{aligned} A_1 &= \{W\} \\ A_2 &= \{R, V\} \\ A_3 &= \{R, P\} \end{aligned}$$

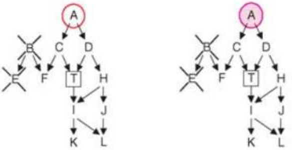
A path from X to Y is blocked by a set of nodes A if there is a node P on the path holding one of the following two conditions:

- P is not a collider and P is not in A or
- P is a collider and neither P or one of its descendants are in A.


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



(a) Generating Network

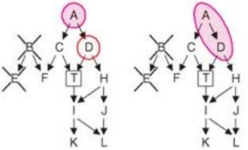


(b) (c)

Iteration 1

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

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

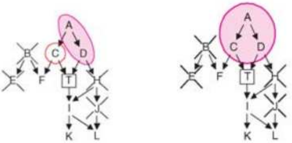


(d) (e)

Iteration 2

(d) D is selected.

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

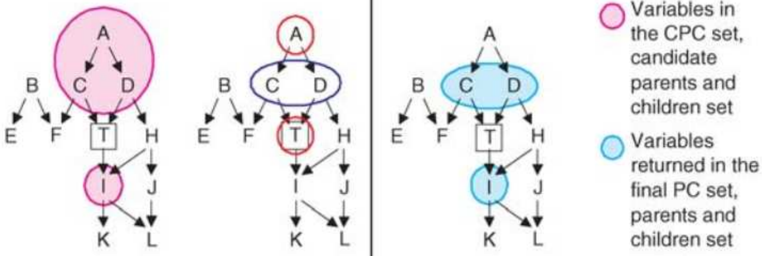


(f) (g)

Iteration 3

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

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



(k) (l) (m) Legend

Backward Phase

(l) $Ind(A; T|\{C, D\})$

(m) Final Outcome of $\overline{MMPC}$ (T)

Phase II: Backward Phase

Thanks for your attention!