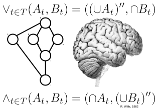
Bayesian networks: conditioning and d-separation

Bayesian Statistics and Machine Learning

Dominik Endres

November 25, 2015

Philipps Universität Marburg



Outline

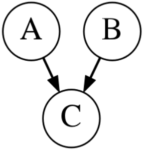
1. Conditioning: prediction and inference
    - Reminder: from Bayes net to factor graph
    - Conditioning with the sum-product algorithm
2. Determining independence: d-separation
    - The swine flu network
    - Determining dependency from the graph structure
    - Unobserved descendants
    - Unobserved ancestors
    - Observed ancestors
    - Observed intermediaries
    - Unobserved intermediaries
    - Observed descendants
    - Blocked and unblocked paths
    - The Markov blanket

Translating a Bayesian net into a factor graph

Bayesian network	factor graph
Node	Variable node
(Un)conditional distribution	Factor node
Origin node of directed edge	Edge from variable to factor
Target node of directed edge	Edge from factor to variable
Rand. Var. in unconditional marginal	Edge from factor to variable

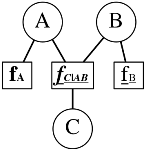
Example: $P(A, B, C) = P(A)P(B)P(C|A, B)$

Bay. net



Marginal distribution $P(A) =$
$= \sum_B \sum_C P(A)P(B)P(C|A, B)$

A factor graph



Marginal function $f(A) =$
$= \sum_A \sum_B f_A(A) f_B(B) f_{AB|C}(A, B, C)$

Summary: sum-product algorithm

Reminder:
$\mathbf{X} = X_1, \ldots, X_N$ are the arguments of $q(X_1, \ldots, X_N)$.
$\mathbf{s}_1, \ldots, \mathbf{s}_K \subseteq \mathbf{X}$, functions $f_1, \ldots, f_K$ depend only on the arguments in the corresponding $\mathbf{s}_i$.
$q(\mathbf{X}) = \prod_{i=1}^{K} f_i(\mathbf{s}_i)$

For *singly connected* graphs (undirected trees), we found:

- Message from factor t to variable Y:

$$\mu_{t\to Y}(Y) = \sum_{Z_1} \ldots \sum_{Z_M} t(Y, Z_1, \ldots, Z_M) \prod_{Z\in \mathrm{ne}_t \setminus Y} \mu_{Z\to t}(Z)$$

- Message from variable Y to factor t:

$$\mu_{Y\to t}(Y) = \prod_{l\in \mathrm{ne}_Y \setminus t} \mu_{l\to Y}(Y)$$

- Message from a factor leaf node f to variable Y: $\mu_{f\to Y}(Y) = f(Y)$
- Message from a variable leaf node Z to factor t: $\mu_{Z\to t}(Z) = 1$
- Marginal of Y: $f(Y) = \prod_{t\in \mathrm{ne}_Y} \mu_{t\to Y}(Y)$.
- Marginal of variable set $\mathbf{s}_i$: $f(\mathbf{s}_i) = f_i(\mathbf{s}_i) \prod_{Z\in \mathrm{ne}_{f_i}} \mu_{Z\to f_i}(Z)$.

Conditioning with the sum-product algorithm

Reminder:
$P(A) = \sum_B \sum_C P(A)P(B)P(C|A,B)$

Example: $P(A,B,C) = P(A)P(B)P(C|A,B)$

For prediction and inference, we want to evaluate conditionals, e.g. given $C = c$:

$$P(A|c) = \frac{P(A,c)}{P(c)}$$

The denominator can be computed with the sum-product algorithm. For the numerator, we need the (partial) marginal

$$P(A,c) = \sum_B P(A)P(B)P(c|A,B)$$

i.e. the sum over C has been dropped, because value C is now instantiated (set to value c).

Conditioning with the sum-product algorithm

Reminder:
$P(A) = \sum_B \sum_C P(A)P(B)P(C|A,B)$

Example: $P(A,B,C) = P(A)P(B)P(C|A,B)$

For prediction and inference, we want to evaluate conditionals, e.g. given $C = c$:

$$P(A|c) = \frac{P(A,c)}{P(c)}$$

The denominator can be computed with the sum-product algorithm. For the numerator, we need the (partial) marginal

$$P(A,c) = \sum_B P(A)P(B)P(c|A,B)$$

i.e. the sum over C has been dropped, because value C is now instantiated (set to value c).

Conditioning with the sum-product algorithm, contd.

In a general singly-connected graph, marginalizations happen in the factor nodes:

$$\mu_{t \to Y}(Y) = \sum_{Z_1} \dots \sum_{Z_M} t(Y, Z_1, \dots, Z_M) \prod_{Z \in \mathrm{ne}_t \setminus Y} \mu_{Z \to t}(Z)$$

$\Rightarrow$ to evaluate partial marginals:

- drop sums over the variables which are instantiated
- set these variables to their corresponding values.

For example, let $Y = a$ and $Z_M = b$. Then

$$\mu_{t \to Y}(Y = a) = \sum_{Z_1} \dots \sum_{Z_{M-1}} t(a, Z_1, \dots, Z_{M-1}, b) \prod_{Z \in \mathrm{ne}_t \setminus Y} \mu_{Z \to t}(Z)$$

and

$$\mu_{t \to Y}(Y \neq a) = 0.$$

Conditioning with the sum-product algorithm, contd.

In a general singly-connected graph, marginalizations happen in the factor nodes:

$$\mu_{t\to Y}(Y) = \sum_{Z_1} \dots \sum_{Z_M} t(Y, Z_1, \dots, Z_M) \prod_{Z \in \mathrm{ne}_t \setminus Y} \mu_{Z\to t}(Z)$$

$\Rightarrow$ to evaluate partial marginals:

- drop sums over the variables which are instantiated
- set these variables to their corresponding values.

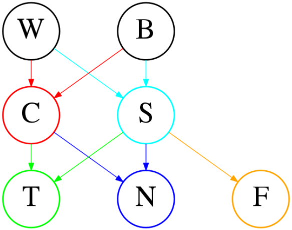
For example, let $Y = a$ and $Z_M = b$. Then

$$\mu_{t\to Y}(Y = a) = \sum_{Z_1} \dots \sum_{Z_{M-1}} t(a, Z_1, \dots, Z_{M-1}, b) \prod_{Z \in \mathrm{ne}_t \setminus Y} \mu_{Z\to t}(Z)$$

and

$$\mu_{t\to Y}(Y \neq a) = 0.$$

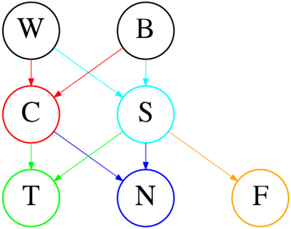
The swine flu network

The swine flu network	Variables
	W =workshop attendance
	B =bar visit
	C =common cold
	S =swine flu
	T =sore throat
	N =runny nose
	F =fever

Possibly relevant questions:

- If I had a cold, would I also get the swine flu? $P(S, C)$ (and $P(S|C)$).
- I've been to a workshop but not to a bar. If I had a cold, would I also get the swine flu? $P(S, C|w, \bar{b})$.
- Would I get a sore throat if I went to a bar tonight? $P(T, B)$.
- I see you've got a cold, but not the swine flu. Should I assume you have a sore throat if I learned that you went to a bar last night? $P(T, B|c, \bar{s})$.
- I see you have a cold. Did you go to a workshop or to a bar? $P(B, W|c)$.

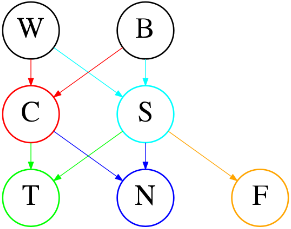
The swine flu network

The swine flu network	Variables
	W =workshop attendance B =bar visit C =common cold S =swine flu T =sore throat N =runny nose F =fever

Possibly relevant questions:

- If I had a cold, would I also get the swine flu? $P(S, C)$ (and $P(S|C)$).
- I've been to a workshop but not to a bar. If I had a cold, would I also get the swine flu? $P(S, C|w, \bar{b})$.
- Would I get a sore throat if I went to a bar tonight? $P(T, B)$.
- I see you've got a cold, but not the swine flu. Should I assume you have a sore throat if I learned that you went to a bar last night? $P(T, B|c, \bar{s})$.
- I see you have a cold. Did you go to a workshop or to a bar? $P(B, W|c)$.

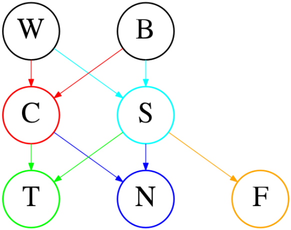
The swine flu network

The swine flu network	Variables
	W =workshop attendance B =bar visit C =common cold S =swine flu T =sore throat N =runny nose F =fever

Possibly relevant questions:

- If I had a cold, would I also get the swine flu? $P(S, C)$ (and $P(S|C)$).
- I've been to a workshop but not to a bar. If I had a cold, would I also get the swine flu? $P(S, C|w, \bar{b})$.
- Would I get a sore throat if I went to a bar tonight? $P(T, B)$.
- I see you've got a cold, but not the swine flu. Should I assume you have a sore throat if I learned that you went to a bar last night? $P(T, B|c, \bar{s})$.
- I see you have a cold. Did you go to a workshop or to a bar? $P(B, W|c)$.

The swine flu network

The swine flu network	Variables
	W =workshop attendance B =bar visit C =common cold S =swine flu T =sore throat N =runny nose F =fever

Possibly relevant questions:

- If I had a cold, would I also get the swine flu? $P(S, C)$ (and $P(S|C)$)
- I've been to a workshop but not to a bar. If I had a cold, would I also get the swine flu? $P(S, C|w, \bar{b})$.
- Would I get a sore throat if I went to a bar tonight? $P(T, B)$.
- I see you've got a cold, but not the swine flu. Should I assume you have a sore throat if I learned that you went to a bar last night? $P(T, B|c, \bar{s})$.
- I see you have a cold. Did you go to a workshop or to a bar? $P(B, W|c)$.

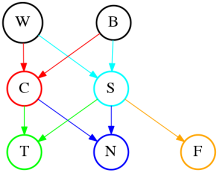
Determining dependency

The question:

If I had a cold, would I also get the swine flu? $P(S, C)$?

can be addressed in two ways:

- **quantitative**: translating the swine flu net into a (singly connected) factor graph and running the sum-product algorithm.
- **qualitative**: determining the dependency between S and C from the graph/factorization structure.



$$P(W, B, C, S, T, N, F) = P(W)P(B) \times P(C|W, B)P(S|W, B) \times P(T|C, S)P(N|C, S)P(F|S)$$

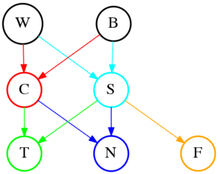
Determining dependency

The question:

If I had a cold, would I also get the swine flu? $P(S, C)$?

can be addressed in two ways:

- **quantitative**: translating the swine flu net into a (singly connected) factor graph and running the sum-product algorithm.
- **qualitative**: determining the dependency between S and C from the graph/factorization structure.



$$P(W, B, C, S, T, N, F) = P(W)P(B) \times P(C|W, B)P(S|W, B) \times P(T|C, S)P(N|C, S)P(F|S)$$

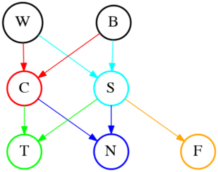
Determining dependency

The question:

If I had a cold, would I also get the swine flu? $P(S, C)$?

can be addressed in two ways:

- **quantitative:** translating the swine flu net into a (singly connected) factor graph and running the sum-product algorithm.
- **qualitative:** determining the dependency between S and C from the graph/factorization structure.



$$P(W, B, C, S, T, N, F) = P(W)P(B) \times P(C|W, B)P(S|W, B) \times P(T|C, S)P(N|C, S)P(F|S)$$

Unobserved descendants do not produce dependencies

Reminder:
$P(W, B, C, S, T, N, F) = P(W)P(B)P(C|W, B)P(S|W, B)P(T|C, S)P(N|C, S)P(F|S)$

To evaluate the functional form of $P(S, C)$, we marginalize all other variables:

$$P(S, C) = \sum_{\sim\{S,C\}} P(W, B, C, S, T, N, F)$$

Begin by marginalizing all descendants of S and C, i.e. T, N, F:

$$\begin{aligned} P(W, B, C, S) &= \sum_T \sum_N \sum_F P(W, B, C, S, T, N, F) \\ &= P(W)P(B)P(C|W, B)P(S|W, B) \\ &\quad \times \left[\underbrace{\sum_T P(T|C, S)}_{=1} \underbrace{\sum_N P(N|C, S)}_{=1} \underbrace{\sum_F P(F|S)}_{=1} \right] \end{aligned}$$

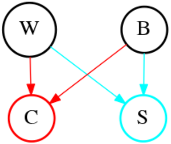
Unobserved descendants do not produce dependencies, contd.

Reminder:
$P(W,B,C,S,T,N,F) = P(W)P(B)P(C|W,B)P(S|W,B)P(T|C,S)P(N|C,S)P(F|S)$

The joint distribution

$$P(W,B,C,S) = P(W)P(B)P(C|W,B)P(S|W,B)$$

has the same dependency structure w.r.t. to S and C as the full joint distribution, since factors of unobserved descendants T,N,R sum to unity.



$$P(W,B,C,S) = \\ P(W)P(B) \\ \times P(C|W,B)P(C|W,B)$$

$\Rightarrow$ unobserved descendants do not change the dependency structrure.

Here: unobserved symptoms T,N,F tell us nothing about diseases C,S.

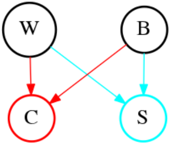
Unobserved descendants do not produce dependencies, contd.

Reminder:
$P(W, B, C, S, T, N, F) = P(W)P(B)P(C|W, B)P(S|W, B)P(T|C, S)P(N|C, S)P(F|S)$

The joint distribution

$$P(W, B, C, S) = P(W)P(B)P(C|W, B)P(S|W, B)$$

has the same dependency structure w.r.t. to S and C as the full joint distribution, since factors of unobserved descendants T, N, R sum to unity.



$$P(W, B, C, S) = P(W)P(B) \times P(C|W, B)P(C|W, B)$$

$\Rightarrow$ unobserved descendants do not change the dependency structrure.

Here: unobserved symptoms T, N, F tell us nothing about diseases C, S.

Unobserved ancestors produce dependencies

Reminder:
$P(W, B, C, S) = P(W)P(B)P(C|W, B)P(S|W, B)$

Compute marginals from $P(W, B, C, S)$:

$$P(C, S) = \sum_W \sum_B P(W)P(B)P(C|W, B)P(S|W, B)$$

$$P(C) = \sum_W \sum_B P(W)P(B)P(C|W, B) \underbrace{\sum_S P(S|W, B)}_{=1}$$

$$P(S) = \sum_W \sum_B P(W)P(B)P(S|W, B)$$

Thus, $P(C, S) \neq P(C)P(S)$.
$\Rightarrow$ unobserved ancestors (here W, B) produce dependencies.
$\Rightarrow$ C and S are marginally dependent: $C \not\perp\!\!\!\perp S$.
Here: unobserved "causes" W, B produce dependencies between diseases C, S

Unobserved ancestors produce dependencies

Reminder:
$P(W, B, C, S) = P(W)P(B)P(C|W, B)P(S|W, B)$

Compute marginals from $P(W, B, C, S)$:

$$P(C,S) = \sum_W \sum_B P(W)P(B)P(C|W,B)P(S|W,B)$$

$$P(C) = \sum_W \sum_B P(W)P(B)P(C|W,B) \underbrace{\sum_S P(S|W,B)}_{=1}$$

$$P(S) = \sum_W \sum_B P(W)P(B)P(S|W,B)$$

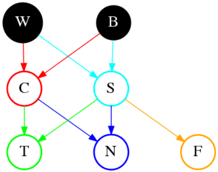
Thus, $P(C,S) \neq P(C)P(S)$.
$\Rightarrow$ unobserved ancestors (here W, B) produce dependencies.
$\Rightarrow$ C and S are marginally dependent: $C \not\perp\!\!\!\perp S$.
Here: unobserved “causes” W, B produce dependencies between diseases C, S

Observed ancestors do not produce dependencies

Next question: I've been to a workshop but not to a bar. If I had a cold, would I also get the swine flu? $P(C, S|w, \bar{b})$.

Answer by evaluating and marginalizing

$$P(w, \bar{b}, C, S) = \sum_{\sim\{S,C\}} P(w, \bar{b}, C, S, T, N, F)$$
$$= \sum_{T}\sum_{N}\sum_{F} P(w, \bar{b}, C, S, T, N, F)$$

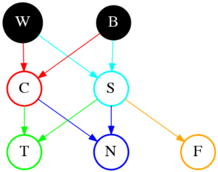


$$P(w, \bar{b}, C, S, T, N, F) = P(w)P(\bar{b}) \times P(C|w, \bar{b})P(S|w, \bar{b}) \times P(T|C, S)P(N|C, S)P(F|S)$$

Observed ancestors do not produce dependencies

Next question: I've been to a workshop but not to a bar. If I had a cold, would I also get the swine flu? $P(C, S|w, \bar{b})$.
Answer by evaluating and marginalizing

$$
\begin{aligned}
P(w, \bar{b}, C, S) &= \sum_{\sim\{S,C\}} P(w, \bar{b}, C, S, T, N, F) \\
&= \sum_T \sum_N \sum_F P(w, \bar{b}, C, S, T, N, F)
\end{aligned}
$$



$$
\begin{gathered}
P(w, \bar{b}, C, S, T, N, F) = \\
P(w)P(\bar{b}) \\
\times P(C|w, \bar{b}) P(S|w, \bar{b}) \\
\times P(T|C, S) P(N|C, S) P(F|S)
\end{gathered}
$$

Observed ancestors do not produce dependencies, contd.

Reminder:
$P(W,B,C,S,T,N,F) = P(W)P(B)P(C|W,B)P(S|W,B)P(T|C,S)P(N|C,S)P(F|S)$

We want to evaluate $P(S,C|w,\bar{b}) = \frac{P(w,\bar{b},S,C)}{P(w,\bar{b})}$ via

$$\begin{aligned}
P(w,\bar{b},S,C) &= \sum_T \sum_N \sum_F P(w,\bar{b},C,S,T,N,F) \\
&= P(w)P(\bar{b})P(C|w,\bar{b})P(S|w,\bar{b}) \\
&\quad \times \underbrace{\left[\sum_T P(T|C,S)\sum_N P(N|C,S)\sum_F P(F|S)\right]}_{=1} \\
P(w,\bar{b}) &= \sum_S \sum_C P(w,\bar{b},S,C) \\
&= P(w)P(\bar{b})\sum_C (C|w,\bar{b})\sum_S P(S|w,\bar{b}) = P(w)P(\bar{b})
\end{aligned}$$

Observed ancestors do not produce dependencies, contd.

Reminder:
$P(w, \bar{b}, S, C) = P(w)P(\bar{b})P(C|w, \bar{b})P(S|w, \bar{b})$
$P(w, \bar{b}) = P(w)P(\bar{b})$

Thus, we find

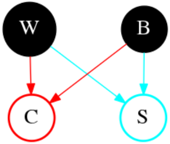
$$
\begin{aligned}
P(S, C|w, \bar{b}) &= \frac{P(w, \bar{b}, S, C)}{P(w, \bar{b})} \\
&= \frac{P(w)P(\bar{b})P(C|w, \bar{b})P(S|w, \bar{b})}{P(w)P(\bar{b})} \\
&= P(C|w, \bar{b})P(S|w, \bar{b})
\end{aligned}
$$

is a product of one factor depending on C and another depending on S.

$\Rightarrow$ C and S are independent given w and $\bar{b}$: $C \perp\!\!\!\perp S | w, \bar{b}$.

$\Rightarrow$ observed ancestors do not produces dependencies.

Here: observing "causes" W, B render diseases C, S independent.



$$
P(w, \bar{b}, C, S) = P(w)P(\bar{b}) \times P(C|w, \bar{b})P(C|w, \bar{b})
$$

Observed ancestors do not produce dependencies, contd.

Reminder:
$P(w, \bar{b}, S, C) = P(w)P(\bar{b})P(C|w, \bar{b})P(S|w, \bar{b})$
$P(w, \bar{b}) = P(w)P(\bar{b})$

Thus, we find

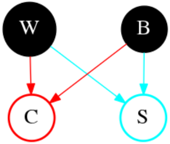
$$\begin{aligned} P(S, C|w, \bar{b}) &= \frac{P(w, \bar{b}, S, C)}{P(w, \bar{b})} \\ &= \frac{P(w)P(\bar{b})P(C|w, \bar{b})P(S|w, \bar{b})}{P(w)P(\bar{b})} \\ &= P(C|w, \bar{b})P(S|w, \bar{b}) \end{aligned}$$

is a product of one factor depending on C and another depending on S.

$\Rightarrow$ C and S are independent given w and $\bar{b}$: $C \perp\!\!\!\perp S | w, \bar{b}$.

$\Rightarrow$ observed ancestors do not produces dependencies.

Here: observing "causes" W, B render diseases C, S independent.



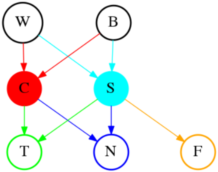
$$P(w, \bar{b}, C, S) = P(w)P(\bar{b}) \times P(C|w, \bar{b})P(C|w, \bar{b})$$

Observed intermediaries do not produce dependencies

Next question: I see you've got a cold, but not the swine flu. Should I assume you have a sore throat if I learned that you went to a bar last night? $P(T, B|c, \bar{s})$

Answer by evaluating and marginalizing

$$P(B, c, \bar{s}, T) = \sum_{\sim\{B,T\}} P(W, B, c, \bar{s}, T, N, F)$$
$$= \sum_{W} \sum_{N} \sum_{F} P(W, B, c, \bar{s}, T, N, F)$$



$$P(W, B, c, \bar{s}, T, N, F) = P(W)P(B) \times P(c|W, B)P(\bar{s}|W, B) \times P(T|c, \bar{s})P(N|c, \bar{s})P(F|\bar{s})$$

Observed intermediaries do not produce dependencies, contd.

Reminder:
$P(W,B,c,\bar{s},T,N,F) = P(W)P(B)P(c|W,B)P(\bar{s}|W,B)P(T|c,\bar{s})P(N|c,\bar{s})P(F|\bar{s})$

We want to evaluate $P(B,c,\bar{s},T)$ via

$$
\begin{aligned}
P(B,c,\bar{s},T) &= \sum_W \sum_N \sum_F P(W,B,c,\bar{s},T,N,F) \\
&= \sum_W P(W)P(B)P(c|W,B)P(\bar{s}|W,B)P(T|c,\bar{s}) \\
&\quad \times \underbrace{\left[\sum_N P(N|c,\bar{s}) \sum_F P(F|\bar{s})\right]}_{=1} \\
&= \sum_W P(W,B,c,\bar{s})P(T|c,\bar{s}) \\
&= P(B,c,\bar{s})P(T|c,\bar{s}) \\
&= P(B|c,\bar{s})P(T|c,\bar{s})P(c,\bar{s})
\end{aligned}
$$

Observed intermediaries do not produce dependencies, contd.

Reminder:
$P(B, c, \bar{s}, T) = P(B|c, \bar{s})P(T|c, \bar{s})P(c, \bar{s})$

Thus we find:

$$
\begin{aligned}
P(B, T|c, \bar{s}) &= \frac{P(B|c, \bar{s})P(T|c, \bar{s})P(c, \bar{s})}{P(c, \bar{s})} \\
&= P(B|c, \bar{s})P(T|c, \bar{s})
\end{aligned}
$$

is a product of one factor depending on B and another one depending on T.

$\Rightarrow B \perp\!\!\!\perp T | c, \bar{s}$.

$\Rightarrow$ observing intermediaries does not produces dependencies.

Here: observing the diseases C and S renders the symptom T independent of the "cause" of the disease B.

Observed intermediaries do not produce dependencies, contd.

Reminder:
$P(B, c, \bar{s}, T) = P(B|c, \bar{s})P(T|c, \bar{s})P(c, \bar{s})$

Thus we find:

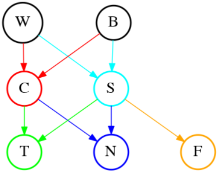
$$\begin{aligned} P(B, T|c, \bar{s}) &= \frac{P(B|c, \bar{s})P(T|c, \bar{s})P(c, \bar{s})}{P(c, \bar{s})} \\ &= P(B|c, \bar{s})P(T|c, \bar{s}) \end{aligned}$$

is a product of one factor depending on B and another one depending on T.
$\Rightarrow B \perp\!\!\!\perp T | c, \bar{s}$.
$\Rightarrow$ observing intermediaries does not produces dependencies.
Here: observing the diseases C and S renders the symptom T independent of the "cause" of the disease B.

Unobserved intermediaries produce dependencies

Next question: Would I get a sore throat if I went to a bar tonight? $P(T,B)$.
Answer: by evaluation of

$$\begin{aligned} P(T,B) &= \sum_{\sim\{T,B\}} P(W,B,C,S,T,N,F) \\ &= P(B) \sum_{\sim\{T,B\}} P(W,B,C,T,N,F|B) \end{aligned}$$



$$\begin{aligned} P(W,B,C,S,T,N,F) = \\ P(W)P(B) \\ \times P(C|W,B)P(S|W,B) \\ \times P(T|C,S)P(N|C,S)P(F|S) \end{aligned}$$

Unobserved intermediaries produce dependencies, contd.

Reminder:
$P(W, B, C, S, T, N, F) = P(W)P(B)P(C|W, B)P(S|W, B)P(T|C, S)P(N|C, S)P(F|S)$

We want to evaluate $P(T, B)$, $P(T)$ and $P(B)$ via

$$
\begin{aligned}
P(T, B) &= P(B) \sum_{\sim\{T,B\}} P(W, C, S, T, N, F|B) \\
&= P(B) \sum_{\sim\{T,B\}} P(W)P(C|W, B)P(S|W, B)P(T|C, S)P(N|C, S)P(F|S) \\
&= P(B) \sum_{W} \sum_{C} \sum_{S} P(W)P(C|W, B)P(S|W, B)P(T|C, S) \\
P(T) &= \sum_{B} P(T, B) \\
&= \sum_{B} P(B) \sum_{W} \sum_{C} \sum_{S} P(W)P(C|W, B)P(S|W, B)P(T|C, S)
\end{aligned}
$$

$\Rightarrow P(T, B) \neq P(B)P(T) \Rightarrow B \not\perp\!\!\!\perp T.$

Unobserved intermediaries produce dependencies, contd.

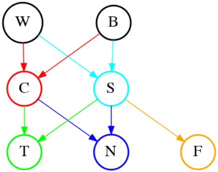
Reminder:
$P(T,B) = P(B)\sum_W \sum_C \sum_S P(W)P(C|W,B)P(S|W,B)P(T|C,S)$
$P(T) = \sum_B P(B)\sum_W \sum_C \sum_S P(W)P(C|W,B)P(S|W,B)P(T|C,S)$

If the intermediaries are unobserved, we find

$$P(T,B) \neq P(T)P(B) \Rightarrow B \not\perp\!\!\!\perp T$$

$\Rightarrow$ unobserved intermediaries produce depdendencies

Here: The symptom T is influenced by the "cause" of diseases B.



$$P(W,B,C,S,T,N,F) = P(W)P(B) \times P(C|W,B)P(S|W,B) \times P(T|C,S)P(N|C,S)P(F|S)$$

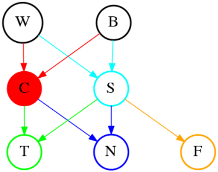
Observed descendants produce dependencies

Last question: I see you have a cold. Did you go to a bar or to a workshop? $P(B, W|c)$.

Answer by evaluation of

$$P(W, B, c) = \sum_{\sim\{W,B\}} P(W, B, c, S, T, N, F)$$

and marginalization to obtain $P(c)$.



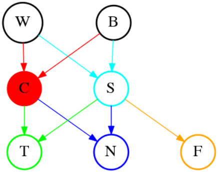
$$P(W, B, c, S, T, N, F) = P(W)P(B) \times P(c|W, B)P(S|W, B) \times P(T|c, S)P(N|c, S)P(F|S)$$

Observed descendants produce dependencies

Last question: I see you have a cold. Did you go to a bar or to a workshop? $P(B, W|c)$.
Answer by evaluation of

$$P(W, B, c) = \sum_{\sim\{W,B\}} P(W, B, c, S, T, N, F)$$

and marginalization to obtain $P(c)$.



$$P(W, B, c, S, T, N, F) = P(W)P(B) \times P(c|W, B)P(S|W, B) \times P(T|c, S)P(N|c, S)P(F|S)$$

Observed descendants produce dependencies, contd.

Reminder:
$P(W, B, C, S, T, N, F) = P(W)P(B)P(C|W, B)P(S|W, B)P(T|C, S)P(N|C, S)P(F|S)$

We want to evaluate $P(W, B|c)$ via

$$\begin{aligned}
P(W, B, c) &= \sum_{\sim\{W,B\}} P(W, B, c, S, T, N, F) \\
&= P(W)P(B)P(c|W, B) \underbrace{\sum_S P(S|W, B)}_{=1} \\
&\quad \times \underbrace{\left[\sum_T P(T|c, S) \sum_N P(N|c, S) \sum_F P(F|S)\right]}_{=1} \\
P(c) &= \sum_W \sum_B P(W)P(B)P(c|W, B) \\
\Rightarrow P(W, B|c) &= P(W)P(B)\frac{P(c|W, B)}{P(c)} \qquad (1)
\end{aligned}$$

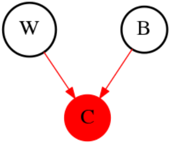
Observed descendants produce dependencies, contd.

Reminder:
$P(W, B, c) = P(W)P(B)P(c|W, B)$
$P(W, B|c) = P(W)P(B)\frac{P(c|W,B)}{P(c)}$

W and B are marginally independent: $W \perp\!\!\!\perp B$.
Observing C renders W and B dependent

$$P(W, B|c) = P(W)P(B)\frac{P(c|W, B)}{P(c)} \Rightarrow W \not\perp\!\!\!\perp B|c$$

$\Rightarrow$ observing common descendants produces a dependency between variables.
Here: the disease-"causes" W and B become dependent upon observing their effect C.



$$P(W, B, c) = P(W)P(B) \times P(c|W, B)$$

Observed descendants produce dependencies, contd.

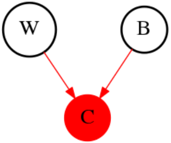
Reminder:
$P(W, B, c) = P(W)P(B)P(c|W, B)$
$P(W, B|c) = P(W)P(B)\frac{P(c|W,B)}{P(c)}$

W and B are marginally independent: $W \perp\!\!\!\perp B$.
Observing C renders W and B dependent

$$P(W, B|c) = P(W)P(B)\frac{P(c|W, B)}{P(c)} \Rightarrow W \not\perp\!\!\!\perp B|c$$

$\Rightarrow$ observing common descendants produces a dependency between variables.

Here: the disease-"causes" W and B become dependent upon observing their effect C.

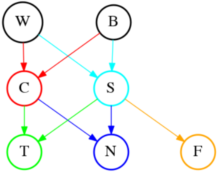


$P(W, B, c) =$
$P(W)P(B)$
$\times P(c|W, B)$

Determining independence graphically: d-separation

To check whether or when nodes in a Bayes net are independent, imagine information flowing from one node to another along the edges.

- The path connecting nodes does *not* have to follow the arrows.
- A path is *blocked* if dependency information cannot flow along it.
- If all paths between two nodes are blocked, then the nodes are independent.
- If at least one path is unblocked, then the nodes will be depdendent.



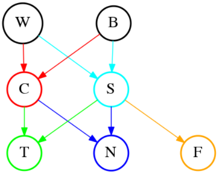
Three types of path components:

- head-to-tail, e.g. $B \to S \to N$
- head-to-head, e.g. $C \to N \to S$
- tail-to-tail, e.g. $C \to W \to S$

Determining independence graphically: d-separation

To check whether or when nodes in a Bayes net are independent, imagine information flowing from one node to another along the edges.

- The path connecting nodes does *not* have to follow the arrows.
- A path is *blocked* if dependency information cannot flow along it.
- If all paths between two nodes are blocked, then the nodes are independent.
- If at least one path is unblocked, then the nodes will be depdendent.



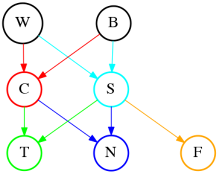
Three types of path components:

- head-to-tail, e.g. $B \to S \to N$
- head-to-head, e.g. $C \to N \to S$
- tail-to-tail, e.g. $C \to W \to S$

Determining independence graphically: d-separation

To check whether or when nodes in a Bayes net are independent, imagine information flowing from one node to another along the edges.

- The path connecting nodes does *not* have to follow the arrows.
- A path is *blocked* if dependency information cannot flow along it.
- If all paths between two nodes are blocked, then the nodes are independent.
- If at least one path is unblocked, then the nodes will be depdendent.



Three types of path components:

- head-to-tail, e.g. $B \rightarrow S \rightarrow N$
- head-to-head, e.g. $C \rightarrow N \rightarrow S$
- tail-to-tail, e.g. $C \rightarrow W \rightarrow S$

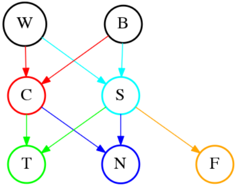
Blocking conditions: head-to-tail

We found:

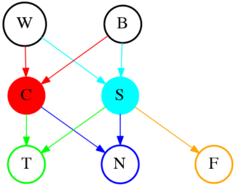
- Observed intermediaries do not produce dependencies.
- Unobserved intermediaries produce dependencies.

Conditions for **head-to-tail** path components:

- **blocked**: if the intermediate node is observed.
- **unblocked**: if the intermediate node is not observed.



B, T dependent



B, T independent

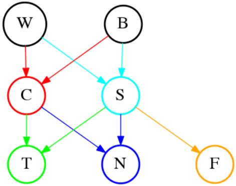
Blocking conditions: head-to-tail

We found:

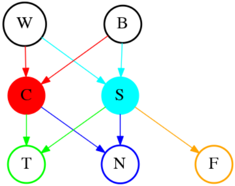
- Observed intermediaries do not produce dependencies.
- Unobserved intermediaries produce dependencies.

Conditions for **head-to-tail** path components:

- **blocked**: if the intermediate node is observed.
- **unblocked**: if the intermediate node is not observed.



B, T dependent



B, T independent

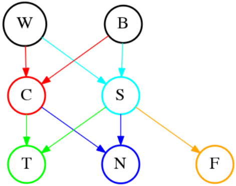
Blocking conditions: tail-to-tail

We found:

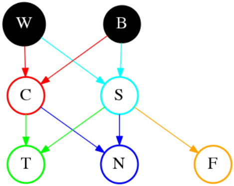
- Observed ancestors do not produce dependencies.
- Unobserved ancestors produce dependencies.

Conditions for **tail-to-tail** path components:

- **blocked**: if the ancestral node is observed.
- **unblocked**: if the ancestral node is not observed.



C, S dependent



C, S independent

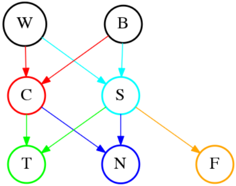
Blocking conditions: tail-to-tail

We found:

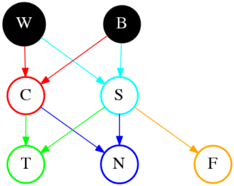
- Observed ancestors do not produce dependencies.
- Unobserved ancestors produce dependencies.

Conditions for **tail-to-tail** path components:

- **blocked**: if the ancestral node is observed.
- **unblocked**: if the ancestral node is not observed.



C, S dependent



C, S independent

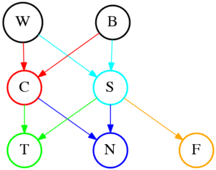
Blocking conditions: head-to-head

We found:

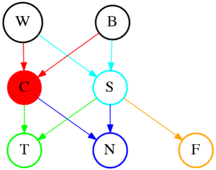
- Unobserved descendants do not produce dependencies.
- Observed descendants produce dependencies.

Conditions for **head-to-head** path components:

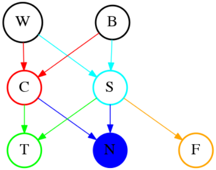
- **blocked**: if the descendant and all of its descendants are not observed
- **unblocked**: if the descendant or any of its descendants are observed.



W, B independent



W, B dependent



W, B dependent

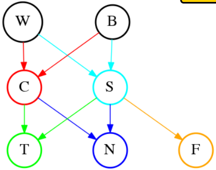
Blocking conditions: head-to-head

We found:

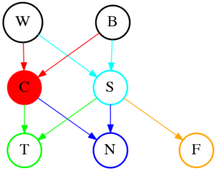
- Unobserved descendants do not produce dependencies.
- Observed descendants produce dependencies.

Conditions for **head-to-head** path components:

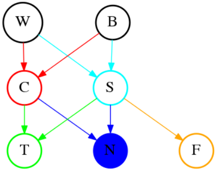
- **blocked**: if the descendant and all of its descendants are not observed
- **unblocked**: if the descendant or any of its descendants are observed.



W, B independent



W, B dependent

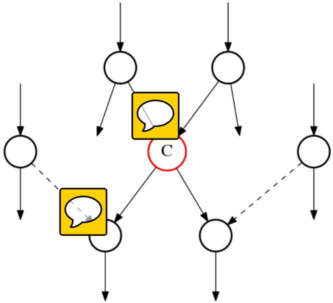


W, B dependent

Blocking all paths to a node

Question: Which nodes do we have to observe to render a given node C independent of all other nodes in the graph?

Answer: Block all paths connecting to C !

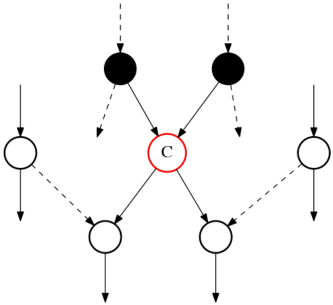


solid: open path, *dashed*: blocked path

Blocking all paths to a node

Question: Which nodes do we have to observe to render a given node C independent of all other nodes in the graph?

Answer: Block all paths connecting to C !



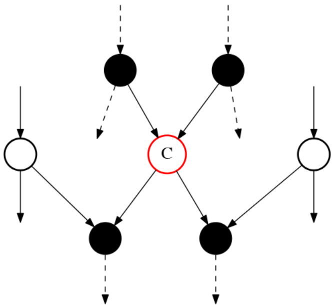
1. Block paths through the parents.

solid: open path, *dashed*: blocked path

Blocking all paths to a node

Question: Which nodes do we have to observe to render a given node C independent of all other nodes in the graph?

Answer: Block all paths connecting to C !



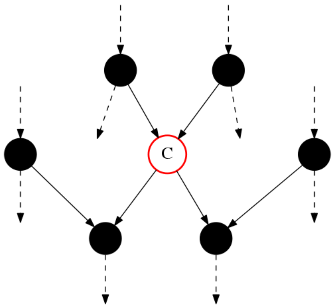
1. Block paths through the parents.
2. Block paths through children.

solid: open path, *dashed*: blocked path

Blocking all paths to a node

Question: Which nodes do we have to observe to render a given node C independent of all other nodes in the graph?

Answer: Block all paths connecting to C !



1. Block paths through the parents.
2. Block paths through children.
3. Block paths through siblings or co-parents.

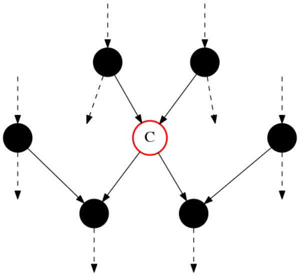
solid: open path, *dashed*: blocked path

The Markov blanket of a node

The set of all parents, children and co-parents is called the **Markov blanket** of a node.

Observing all nodes in the Markov blanket renders the node independent of any other node in the network.

This is e.g. important in Gibbs sampling.



solid: open path, *dashed*: blocked path