Bayesian network

Bayesian Network

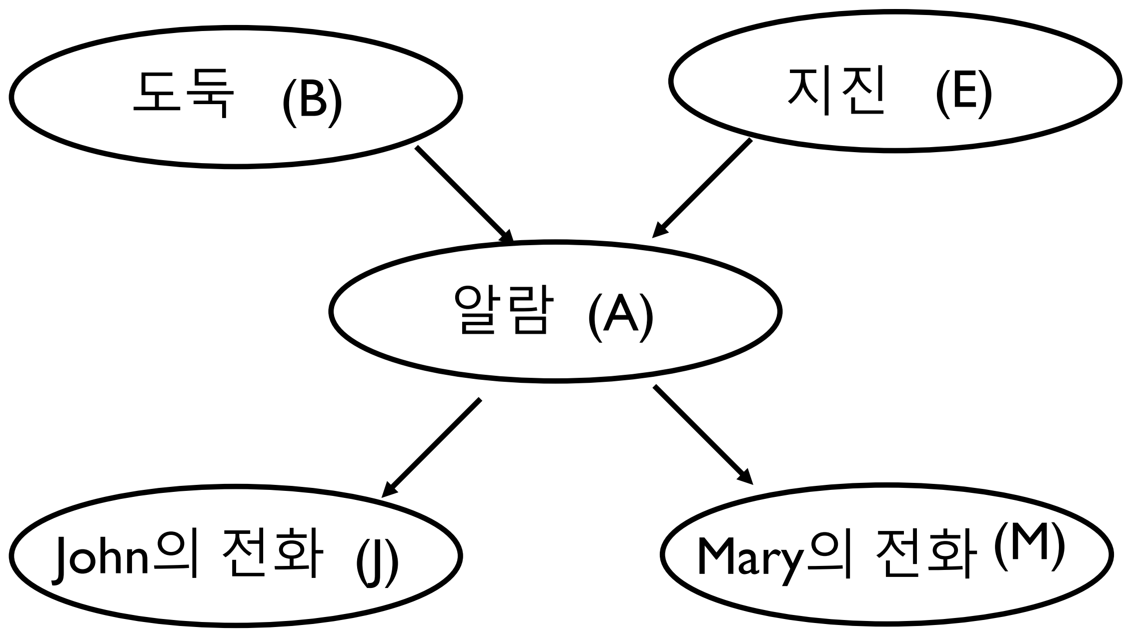
- A method for describing complex models (joint distributions) using conditional probability

- Global semantics defines the full joint distribution as the product of the local conditional distributions

$$P(X_1 \dots X_n) = \prod_i P(X_i | Pa(X_i))$$

- Local semantics means that each node is conditionally independent of its non-descendants given its parents

Bayesian network: an example



$$P(B, E, A, J, M) = P(B)\, P(E|B)\, P(A|B, E)\, P(J|B, E, A)\, P(M| B, E, A, J)$$

$$P(B, E, A, J, M) = P(B)\, P(E)\, P(A|B, E)\, P(J| A)\, P(M| A)$$

$E \perp\!\!\!\perp B$

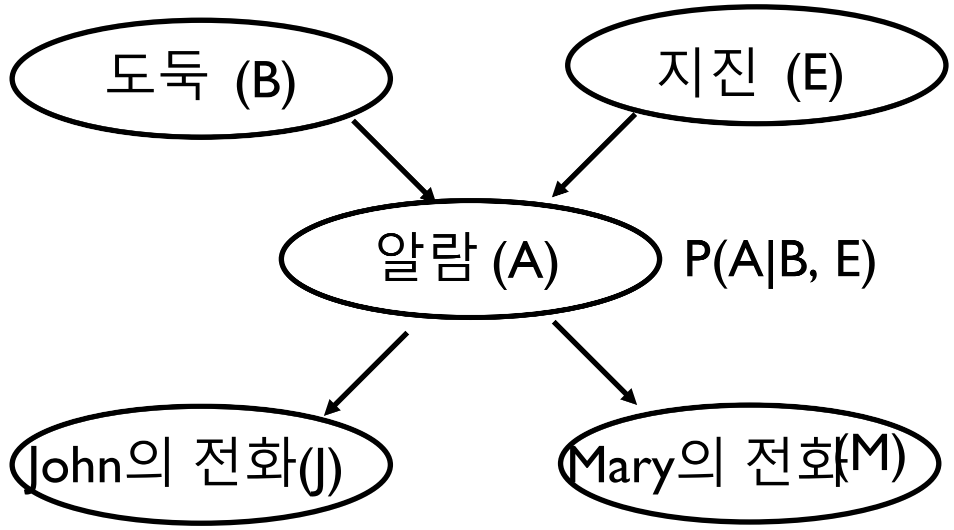
$J \perp\!\!\!\perp B, E \mid A$

$M \perp\!\!\!\perp B, E \mid A$

Bayesian network: an example

P(B) = 0.001

P(E) = 0.002



P(A|B, E)

B	E	P(A|B, E)
T	T	0.95
T	F	0.94
F	T	0.29
F	F	0.01

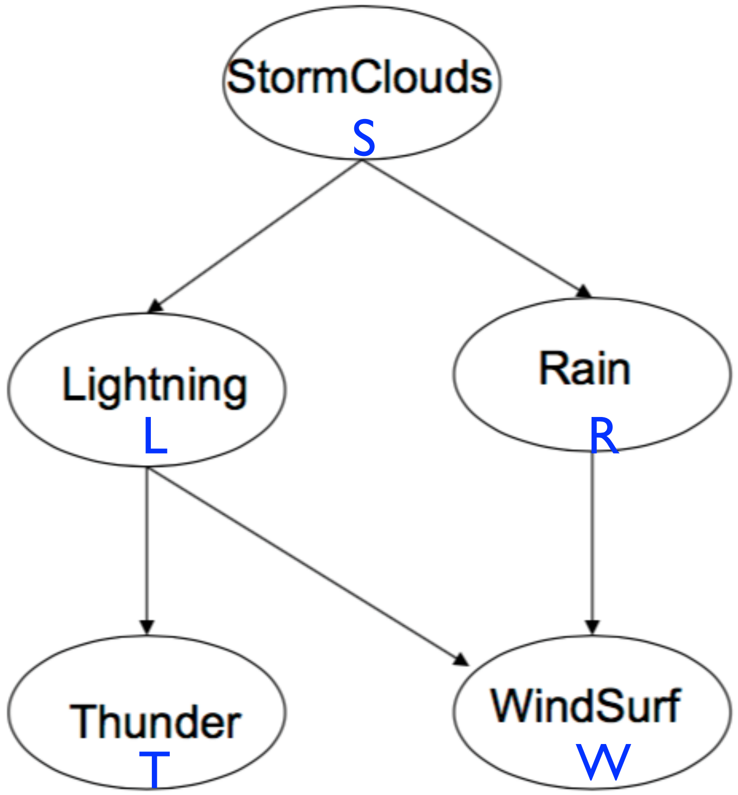
A	P(J|A)
T	0.90
F	0.05

A	P(M|A)
T	0.70
F	0.01

- What is the probability that the alarm has sounded, but neither burglary nor an earthquake has occurred, and John and Mary call?

$$P(a, \neg b, \neg e, j, m) = P(\neg b) \times P(\neg e) \times P(a \mid \neg b, \neg e) \times P(j \mid a) \times P(m \mid a)$$
$$= 0.999 \times 0.998 \times 0.01 \times 0.9 \times 0.7$$

Bayesian network: an example



- Chain rule of probability

  $P\ (S, L, R, T, W) = P(S)\ P(L \mid S)\ P(R \mid S, L)\ P(T \mid S, L, R)\ P\ (W \mid S, L, R, T)$

- Bayes net
  - $P\ (S, L, R, T, W) = P(S)\ P(L \mid S)\ P(R \mid S)\ P(T \mid L)\ P\ (W \mid L, R)$

Constructing Bayesian network

- What is the Bayes Network for X_1 and X_2 for two coins?

Constructing Bayesian network

- What is the Bayes Network for X_1 and X_2 for two coins?

$P(X_1)$

h	0.5
t	0.5

$P(X_2)$

h	0.5
t	0.5

$$P(X_1, X_2) = P(X_1)\,P(X_2)$$

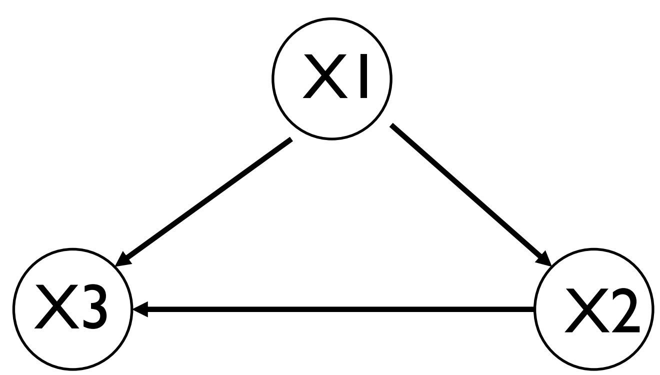
Constructing Bayesian network

- What is the Bayes Network for X1, …, X3 with no assumption of conditional independencies?

Constructing Bayesian network

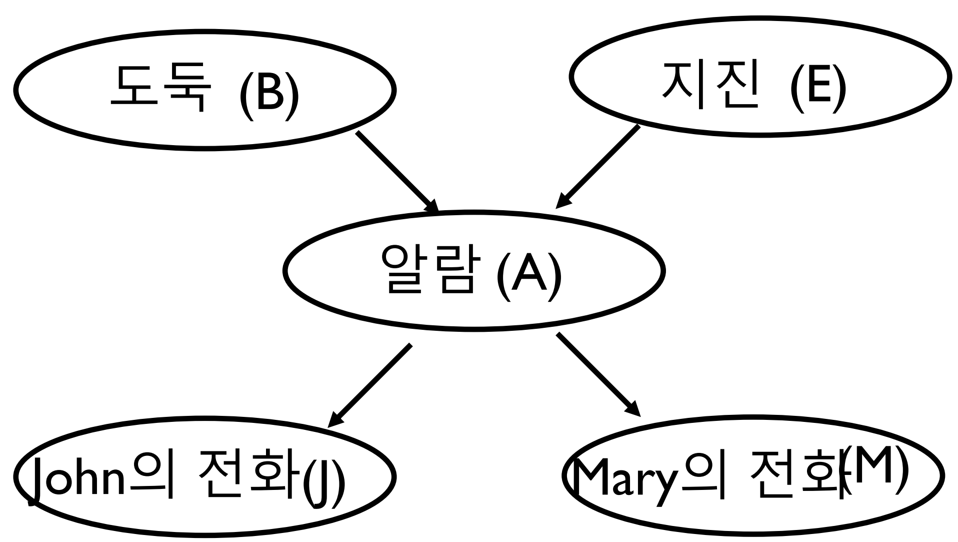
- What is the Bayes Network for X1, …, X3 with no assumption of conditional independencies?

$$P(X1, X2, X3) = P(X1)P(X2|X1)P(X3|X1, X2)$$



Constructing Bayesian network

- Each node is conditionally independent of its predecessors in the node ordering, given its parents
- Needs to choose parents for each node such that this property holds
- The parents of node X_i should contain all those nodes in $X_1, \ldots X_{i-1}$ that directly influence X_i
  - for example, $P(M \mid J, A, E, B) = P(M|A)$



Constructing Bayesian network

1. Choose an ordering of variables $X_1, \ldots, X_n$
2. For $i = 1$ to n

   add X_i to the network

   select parents from $X_1, \ldots, X_{i-1}$ such that

   $$\mathbf{P}(X_i | Parents(X_i)) = \mathbf{P}(X_i | X_1, \ldots, X_{i-1})$$

This choice of parents guarantees the global semantics:

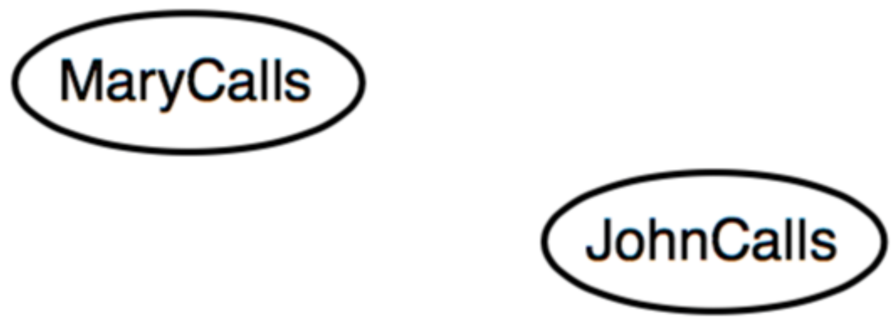
$$\begin{aligned} \mathbf{P}(X_1, \ldots, X_n) &= \prod_{i=1}^{n} \mathbf{P}(X_i | X_1, \ldots, X_{i-1}) && \text{(chain rule)} \\ &= \prod_{i=1}^{n} \mathbf{P}(X_i | Parents(X_i)) && \text{(by construction)} \end{aligned}$$

Constructing Bayesian network: an example

Suppose we choose the ordering M, J, A, B, E

Constructing Bayesian network: an example

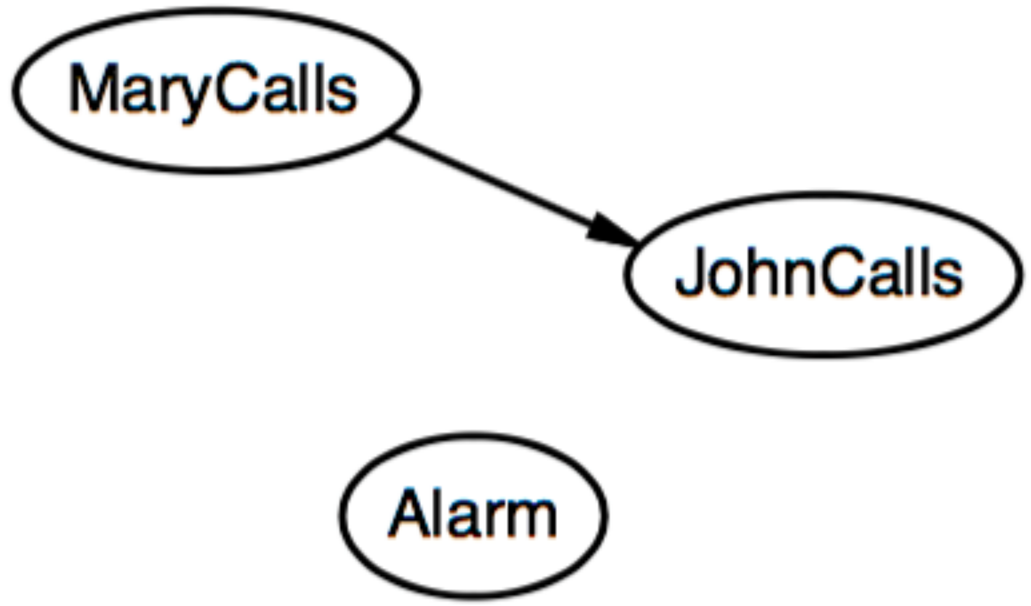
Suppose we choose the ordering M, J, A, B, E



$P(J|M) = P(J)$?

Constructing Bayesian network: an example

Suppose we choose the ordering M, J, A, B, E

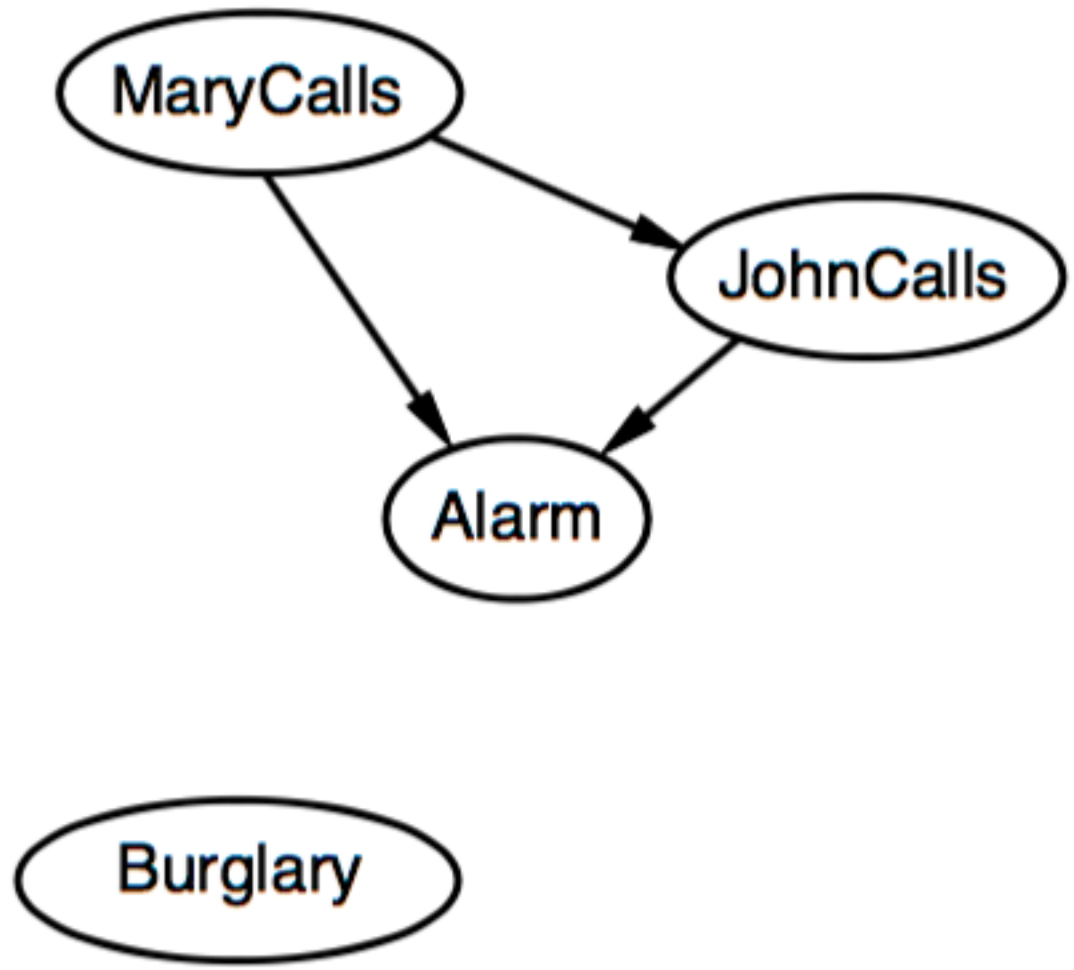


$P(J|M) = P(J)$? No (J is not independent of M)

$P(A|J,M) = P(A|J)$? $P(A|J,M) = P(A)$?

Constructing Bayesian network: an example

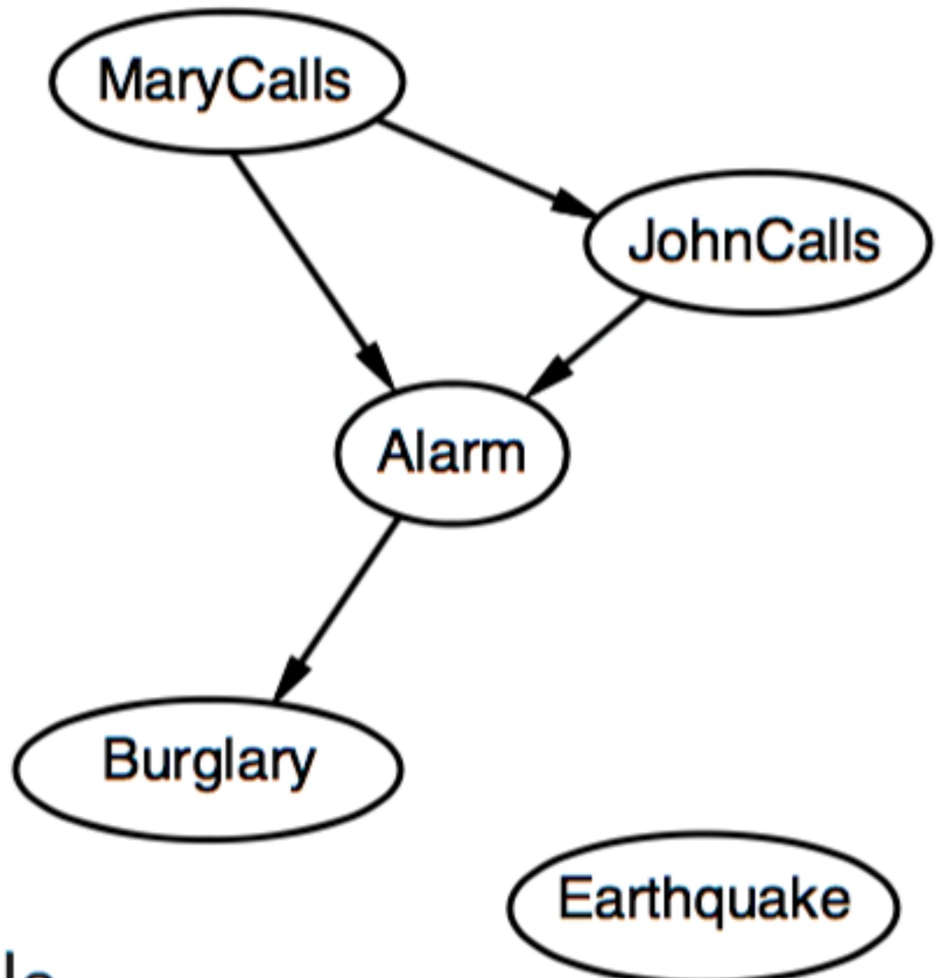
Suppose we choose the ordering M, J, A, B, E



$P(J|M) = P(J)$? No
$P(A|J, M) = P(A|J)$? $P(A|J, M) = P(A)$? No
$P(B|A, J, M) = P(B|A)$?
$P(B|A, J, M) = P(B)$?

Constructing Bayesian network: an example

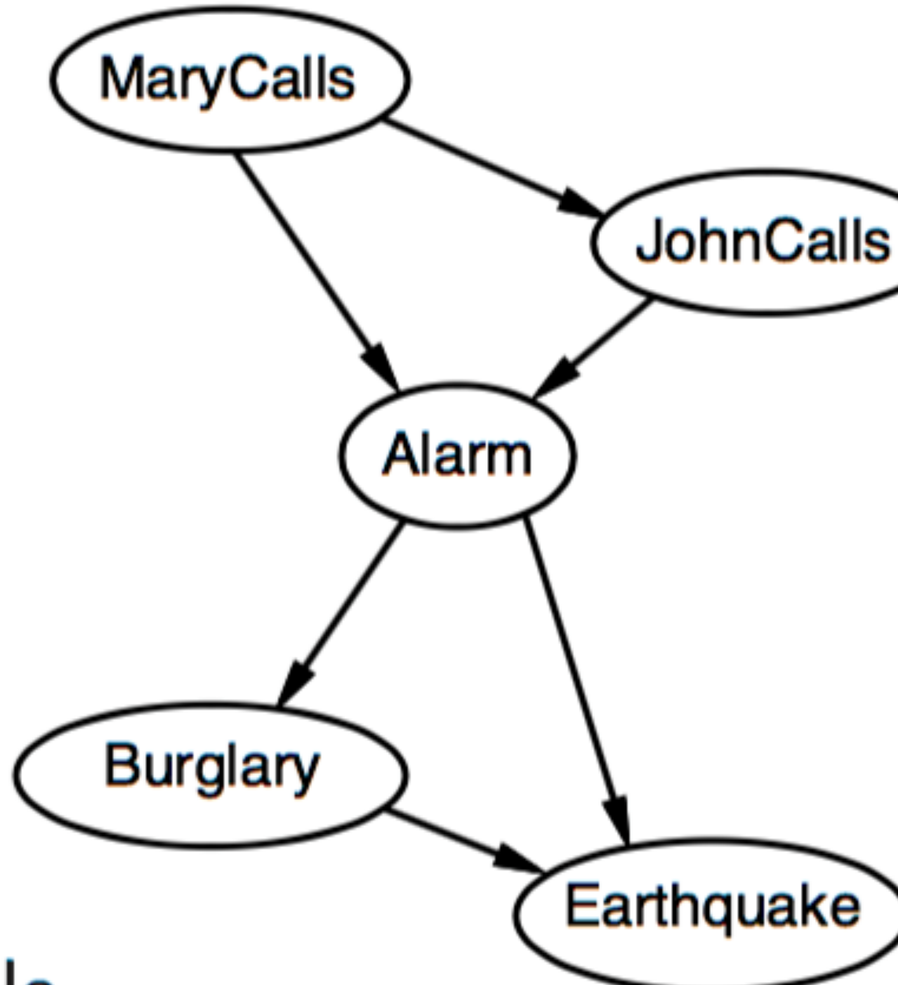
Suppose we choose the ordering M, J, A, B, E



$P(J|M) = P(J)$? No
$P(A|J,M) = P(A|J)$? $P(A|J,M) = P(A)$? No
$P(B|A,J,M) = P(B|A)$? Yes
$P(B|A,J,M) = P(B)$? No
$P(E|B,A,J,M) = P(E|A)$?
$P(E|B,A,J,M) = P(E|A,B)$?

Constructing Bayesian network: an example

Suppose we choose the ordering M, J, A, B, E



$P(J|M) = P(J)$? No

$P(A|J,M) = P(A|J)$? $P(A|J,M) = P(A)$? No

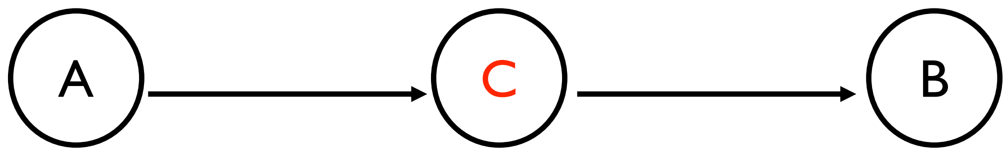
$P(B|A,J,M) = P(B|A)$? Yes

$P(B|A,J,M) = P(B)$? No

$P(E|B,A,J,M) = P(E|A)$? No

$P(E|B,A,J,M) = P(E|A,B)$? Yes

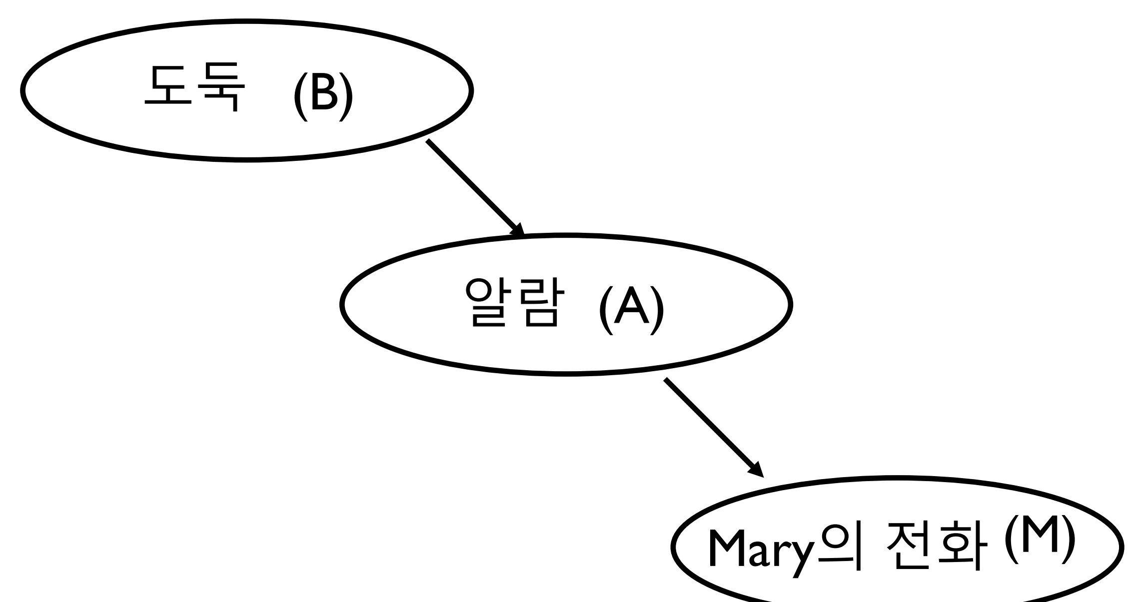
Bayesian network configuration: cascading



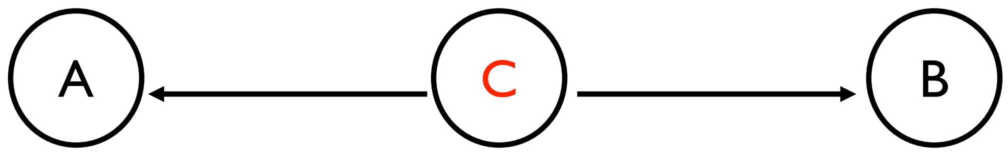
A is conditionally independent of B given C? $\rightarrow$ A ⫫ B | C?

$\rightarrow$ p(A,B|C) = p(A|C) p(B|C) ?

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



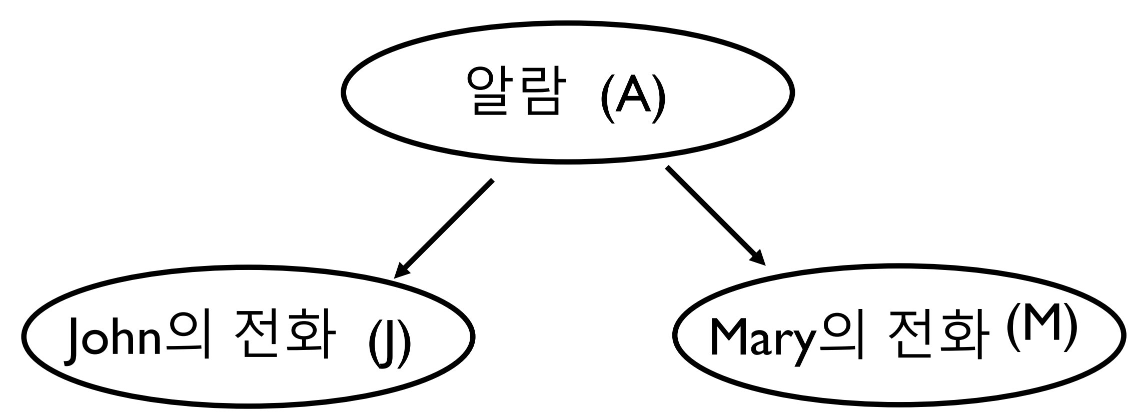
Bayesian network configuration : common parent



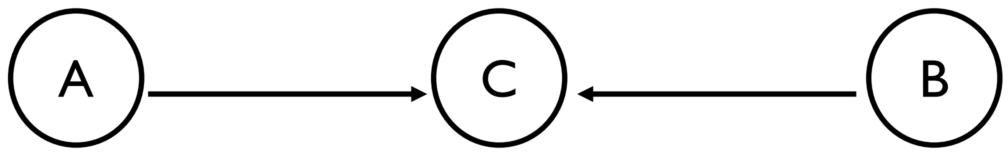
A is conditionally independent of B given C? → A ⫫ B | C?

→ p(A,B|C) = p(A|C) p(B|C) ?

$$P(A,B|C) = \frac{P(A,B,C)}{P(C)} = \frac{\cancel{P(C)}P(A|C)P(B|C)}{\cancel{P(C)}} = P(A|C)P(B|C)$$



Bayesian network configuration : common child



A is conditionally independent of B given C?

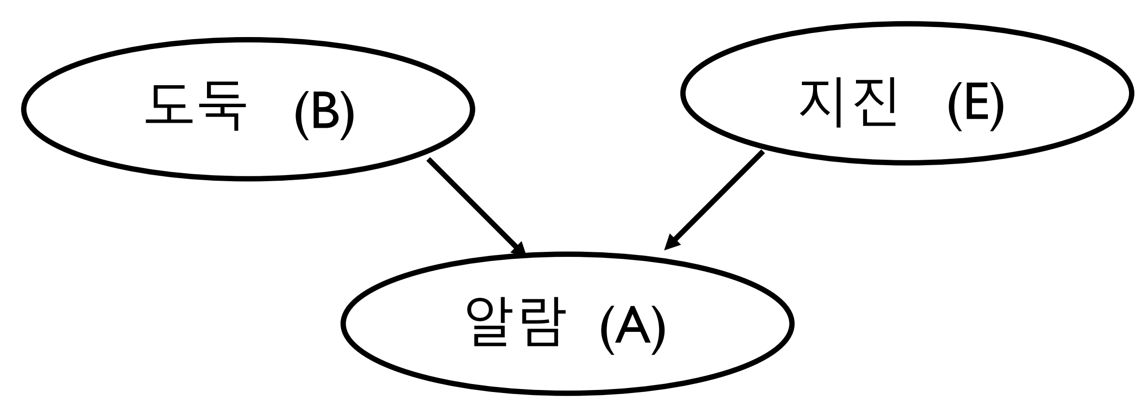
p(a,b|c) = p(a|c) p(b|c) ?

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

After marginalizing over C

$$P(A, B) = \sum_{C} P(A)P(B)P(C|A, B) = P(A)P(B)$$

Thus A and B are independent



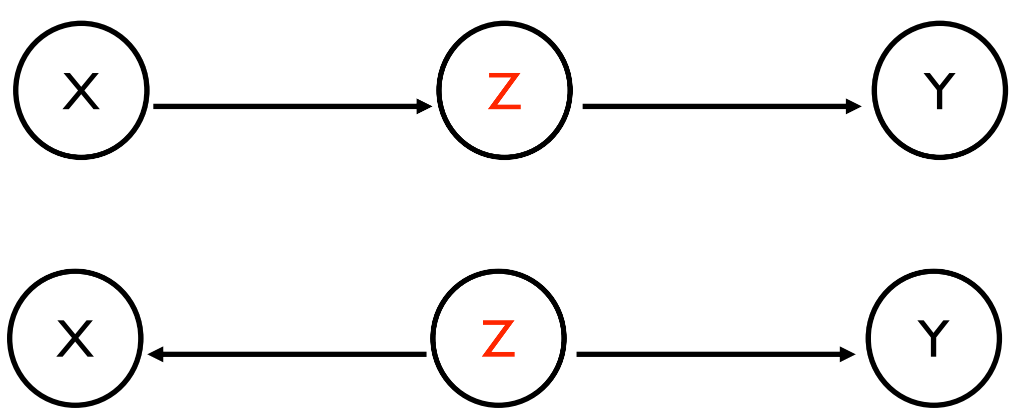
D-separated

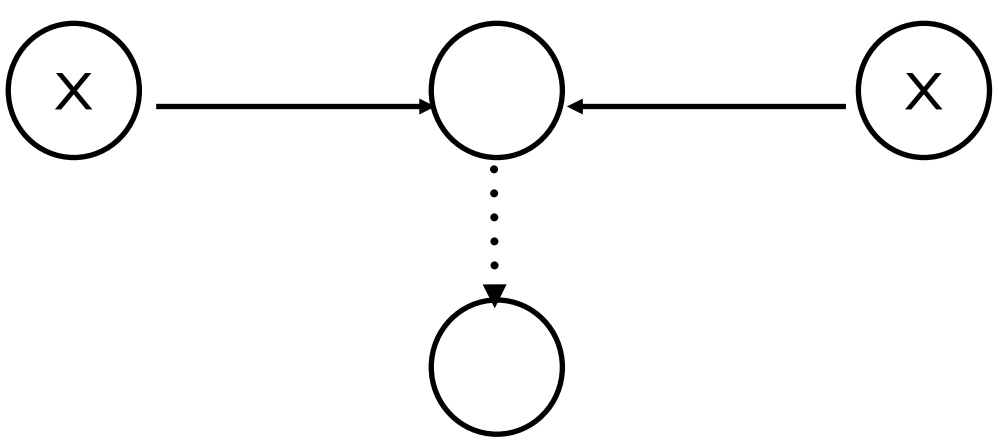
X and Y are D-separated by Z (conditionally independent given Z) iff every path from every variable in X to every variable in Y is blocked. (X, Y are independent)

A path from variable X to variable Y is blocked if it includes a node such that either

(a) the arrows on the path meet either "cascading" or "common parent", and the node is in the set Z

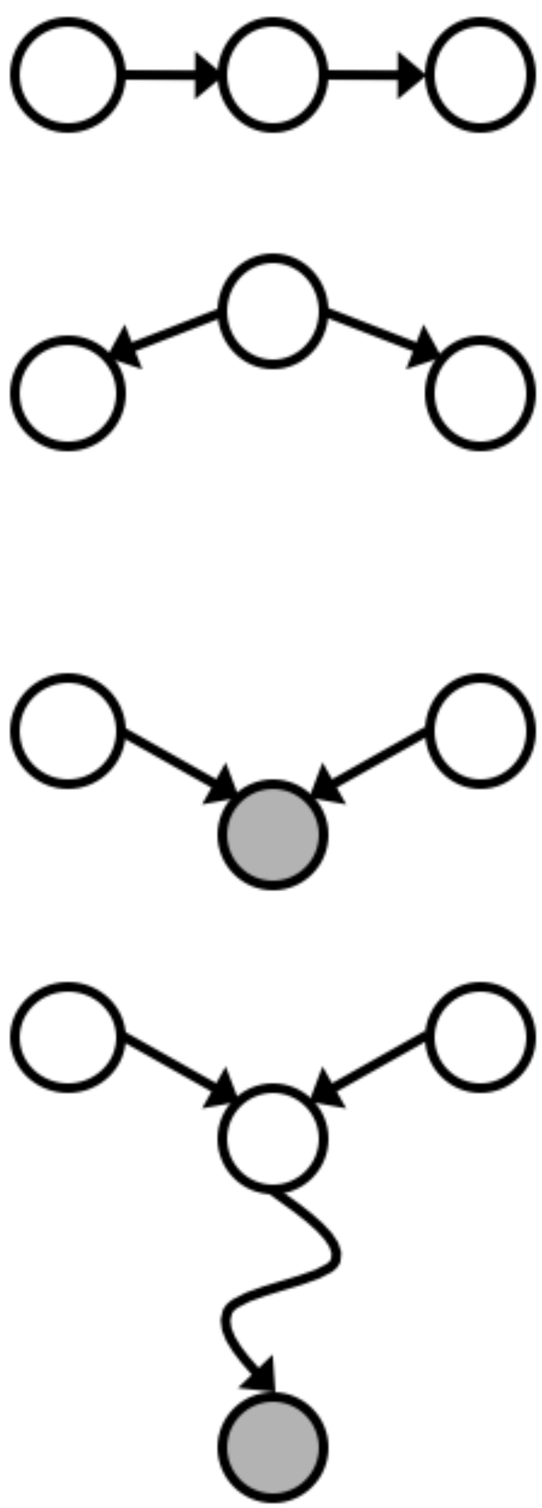
(b) the arrows meet "common child" at the node, and neither the node nor any of its descendants is in the set Z



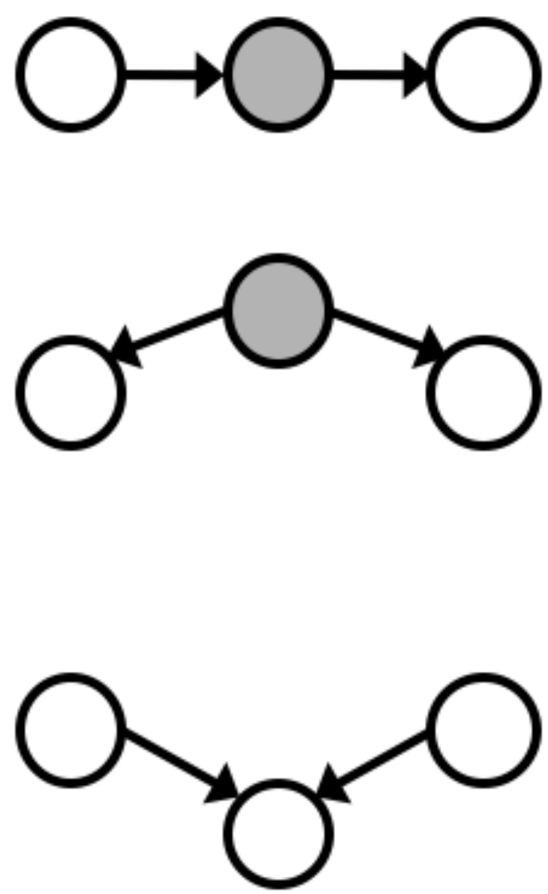


D-separated

not blocked



blocked



if all paths between A and B are blocked, independence is guaranteed

Markov blanket

Each node is conditionally independent of all others given its Markov blanket (parents, children, and children's parents)