CSC384
Introduction to Artificial Intelligence: Uncertainty

November 11, 2014

Bayesian Networks

Terminology

Some common terms with intuitive definitions:

- **Parents**: $par(X_i)$
- **Children**
- **Descendents**
- **Ancestors**
- **Family**: A node and its parents
- **Evidence**: assignments for some variables

Bayesian Networks

Conditional independence

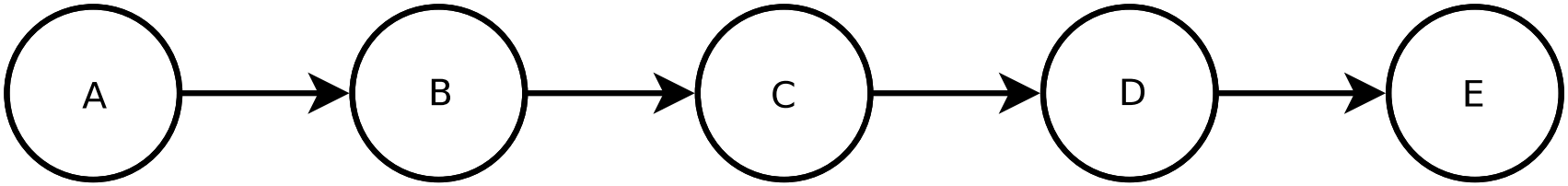
Variables are independent of their non-descendants given the values of their parents.

This leads to the following decomposition with the chain rule:

$$P(X_1, ..., X_n) = P(X_n|par(X_n)) * P(X_{n-1}|par(X_{n-1})) \\ *... * P(X_1|par(X_1))$$

Bayesian Networks

Conditional independence



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

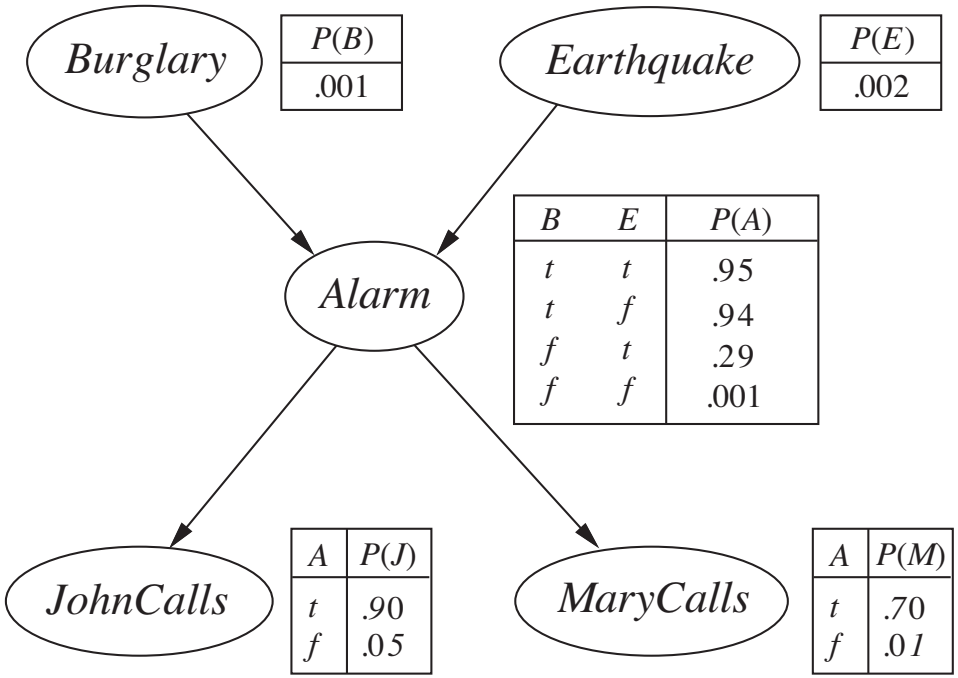
$$P(A,B,C,D,E) = P(E|D) * P(D|C) * P(C|B) * P(B|A) * P(A)$$

$$P(C) = \sum_{b_i \in Domain(B)} P(C|b_i)P(b_i)$$

$$= \sum_{b_i \in Domain(B)} P(C|b_i) \sum_{a_j \in Domain(A)} P(b_i|a_j)$$

Bayesian Networks

Conditional Independence

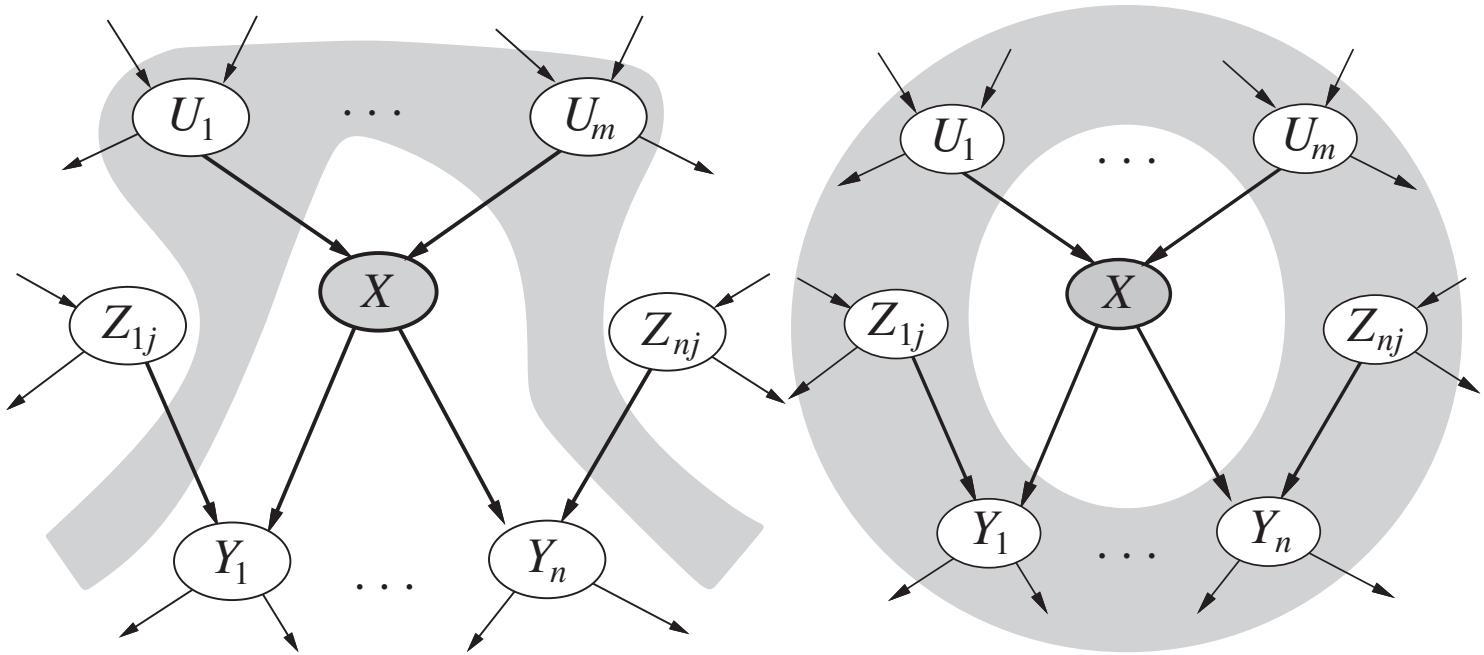


Are Earthquake and Burglary independent?

Are Earthquake and Burglary independent given alarm? Given JohnCalls?

Bayesian Networks

Conditional Independence



Left: A node X is conditionaly independent of its descendents given its parents (the gray area).

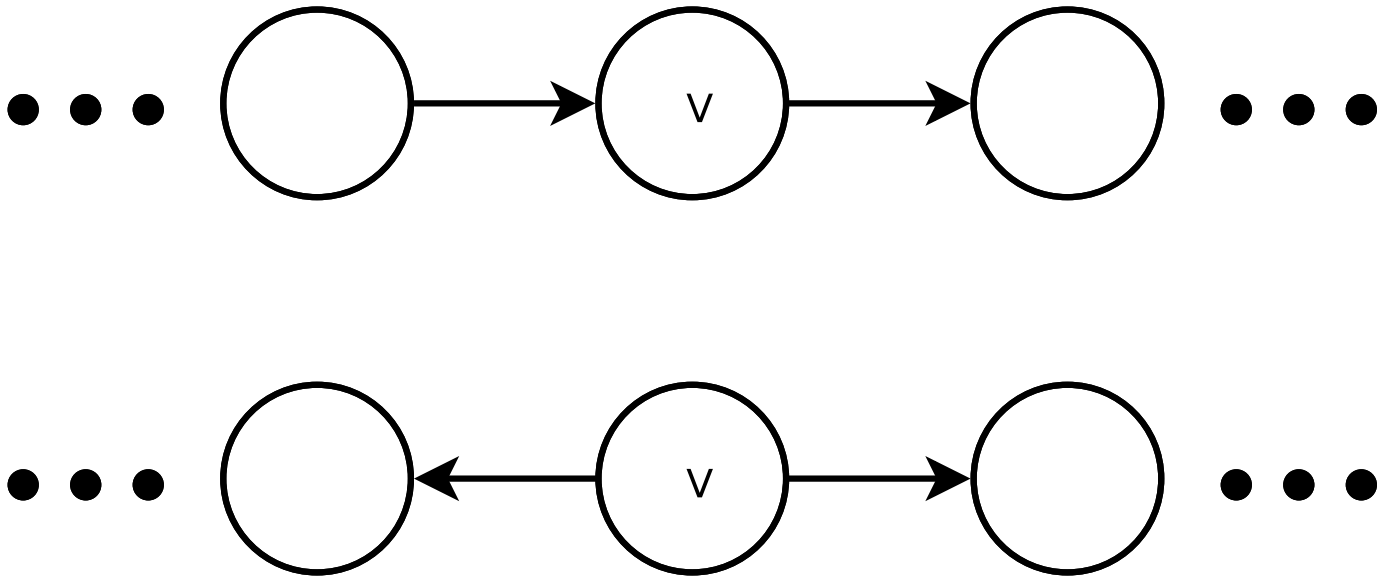
Right: A node X is conditionaly independent of all other nodes in the network given its Markov blanket (the gray area).

Bayesian Networks

Conditional Independence

D-Seperation (implies conditional independence):
Two variables are D-seperated if every undirected path between them contains a node with evidence (the conditional term) such that either:

- The arcs on the node are "tail to head"
- The arcs on the node are "tail to tail"



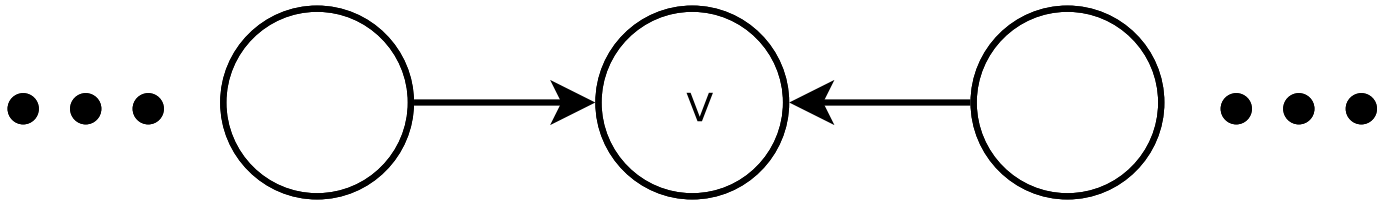
Or a node with no evidence satisifies an alternate condition

Bayesian Networks

Conditional Independence

The alternate conditions for satisfying a path in D-Seperation:

- The node is not in the evidence set
- No *descendent* of the node is in the evidence set
- the arcs on the node are "head to head"



Bayesian Networks

Constructing Bayes Nets

- Take an ordering of the variables
- Consider the chain rule:

$$P(X_1, .., X_n) = P(X_n|X_{n-1}...X_1) * ... * P(X_2|X_1) * P(X_1)$$

- For each X_i iteratively remove conditionally independent elements from its conditioning set $\{X_1, ..., X_{i-1}\}$ given the remaining variables. Repeat this until no more eliminations can occur
- What remains forms a Bayesian network

Bayesian Networks

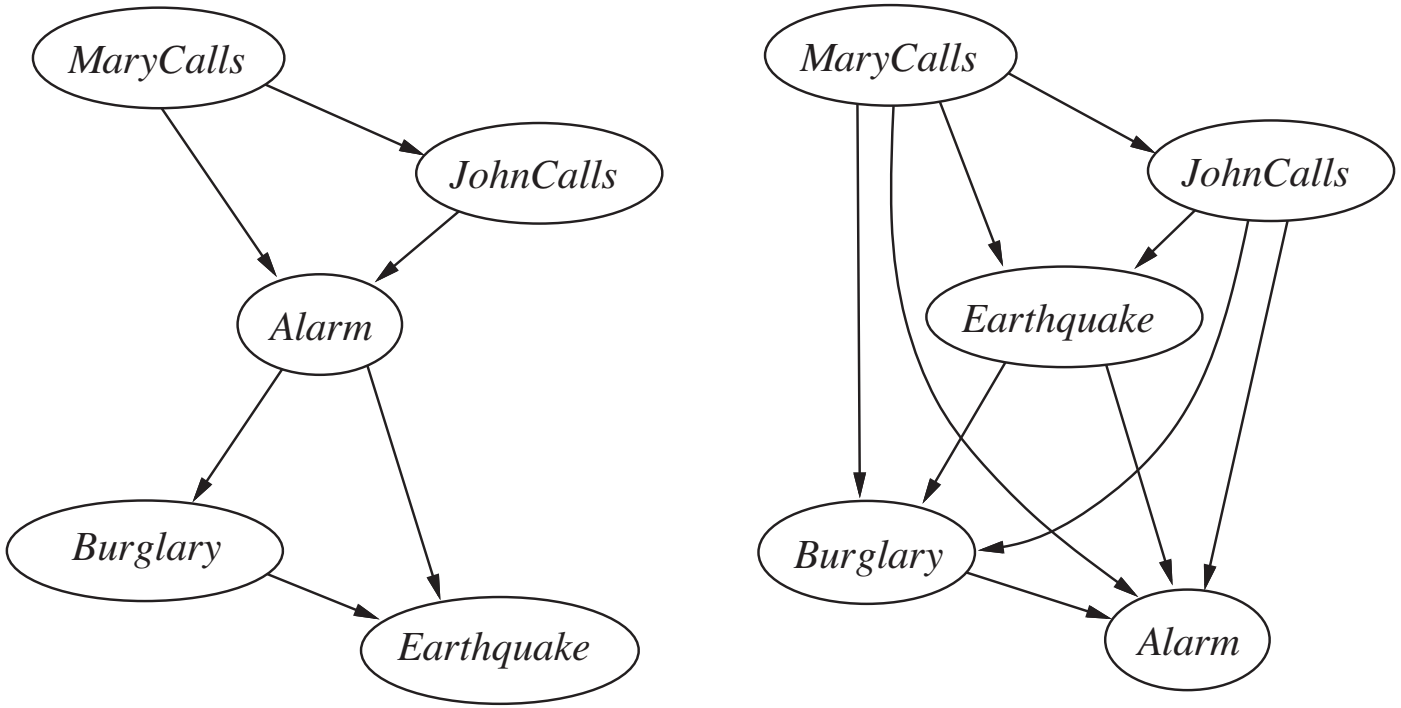
Constructing Bayes Nets

Forming the Bayes net:

- Specify numeric values for each conditional entry $P(X_i|par(X_i))$
- Note that the size of the tables for conditional entries is exponential in the number of parents!
- We gain massive savings in space compared to a table which contains all information by minimizing dependencies

Bayesian Networks

Constructing Bayes Nets



Possible results when given a poor ordering of variables