

Belief Networks

Introduction to Artificial Intelligence

CSE 151

Lecture 3

April 6, 2004

Last lecture

- Conditional Independence
- Bayesian (Belief) Networks: Syntax

This lecture

- Bayesian Networks
- Inference in Bayes nets

Reading

- Chapter 14.1-14.2

Using the full joint distribution

	Toothache=true	Toothache = false
Cavity=true	0.04	0.06
Cavity=false	0.01	0.89

What is the unconditional probability of having a Cavity?

$$\begin{aligned} P(\text{Cavity}) &= P(\text{Cavity} \wedge \text{Toothache}) + P(\text{Cavity} \wedge \sim \text{Toothache}) \\ &= 0.04 + 0.06 = 0.1 \end{aligned}$$

What is the probability of having either a cavity or a Toothache?

$$\begin{aligned} P(\text{Cavity} \vee \text{Toothache}) &= P(\text{Cavity}, \text{Toothache}) + P(\text{Cavity}, \sim \text{Toothache}) + P(\sim \text{Cavity}, \text{Toothache}) \\ &= 0.04 + 0.06 + 0.01 = 0.11 \end{aligned}$$

Marginalization

Given a condition distribution $P(X | Z)$, we can create the unconditional distribution $P(X)$ by marginalization:

$$P(X) = \sum_z P(X | Z=z) P(Z=z) = \sum_z P(X, Z=z)$$

In general, given a joint distribution over a set of variables, the distribution over any subset (called a **marginal distribution** for historical reasons) can be calculated by summing out the other variables.

Conditioning

Introducing a variable as an extra condition:

$$P(X|Y) = \sum_z P(X | Y, Z=z) P(Z=z | Y)$$

Intuition: often easier to assess each specific circumstance, e.g.,

$$\begin{aligned} &P(\text{RunOver} | \text{Cross}) \\ &= P(\text{RunOver} | \text{Cross}, \text{Light}=\text{green})P(\text{Light}=\text{green} | \text{Cross}) \\ &\quad + P(\text{RunOver} | \text{Cross}, \text{Light}=\text{yellow})P(\text{Light}=\text{yellow} | \text{Cross}) \\ &\quad + P(\text{RunOver} | \text{Cross}, \text{Light}=\text{red})P(\text{Light}=\text{red} | \text{Cross}) \end{aligned}$$

Absolute Independence

- Two random variables A and B are (absolutely) **independent** iff
$$P(A, B) = P(A)P(B)$$
- Using product rule for A & B independent, we can show:
$$P(A, B) = P(A | B)P(B) = P(A)P(B)$$
Therefore $P(A | B) = P(A)$
- If n Boolean variables are independent, the full joint is:
$$P(X_1, \dots, X_n) = \prod_i P(X_i)$$
Full joint is generally specified by $2^n - 1$ numbers, but when independent only n numbers are needed.
- Absolute independence is a very strong requirement, seldom met!!

Conditional Independence

- Some evidence may be irrelevant, allowing simplification, e.g.,

$$P(\text{Cavity} \mid \text{Toothache}, \text{CubsWin}) = P(\text{Cavity} \mid \text{Toothache}) = 0.8$$

- This property is known as **Conditional Independence** and can be expressed as:

$$P(X \mid Y, Z) = P(X \mid Z)$$

which says that X and Y conditionally independent given Z.

- If I have a cavity, the probability that the probe catches in it doesn't depend on whether I have a toothache:

1. $P(\text{Catch} \mid \text{Toothache}, \text{Cavity}) = P(\text{Catch} \mid \text{Cavity})$

i.e., Catch is conditionally independent of Toothache given Cavity

The same independence holds if I haven't got a cavity:

2. $P(\text{Catch} \mid \text{Toothache}, \sim\text{Cavity}) = P(\text{Catch} \mid \sim\text{Cavity})$

Using Conditional Independence

Full joint distribution can now be written as

$$\begin{aligned} &P(\text{Toothache}, \text{Catch}, \text{Cavity}) \\ &= P(\text{Toothache}, \text{Catch} \mid \text{Cavity}) P(\text{Cavity}) \\ &= P(\text{Toothache} \mid \text{Cavity}) P(\text{Catch} \mid \text{Cavity}) P(\text{Cavity}) \quad (\text{from 1.b}) \end{aligned}$$

Specified by: $2 + 2 + 1 = 5$ independent numbers

Compared to 7 for general joint

or 3 for unconditionally independent.

Belief (Bayes) networks

A simple, graphical notation for conditional independence assertions and hence for compact specification of full joint distributions.

Syntax:

1. a set of nodes, one per random variable
2. links mean parent “directly influences” child
3. a directed, acyclic graph
4. a conditional distribution (a table) for each node given its parents: $P(X_i \mid \text{Parents}(X_i))$

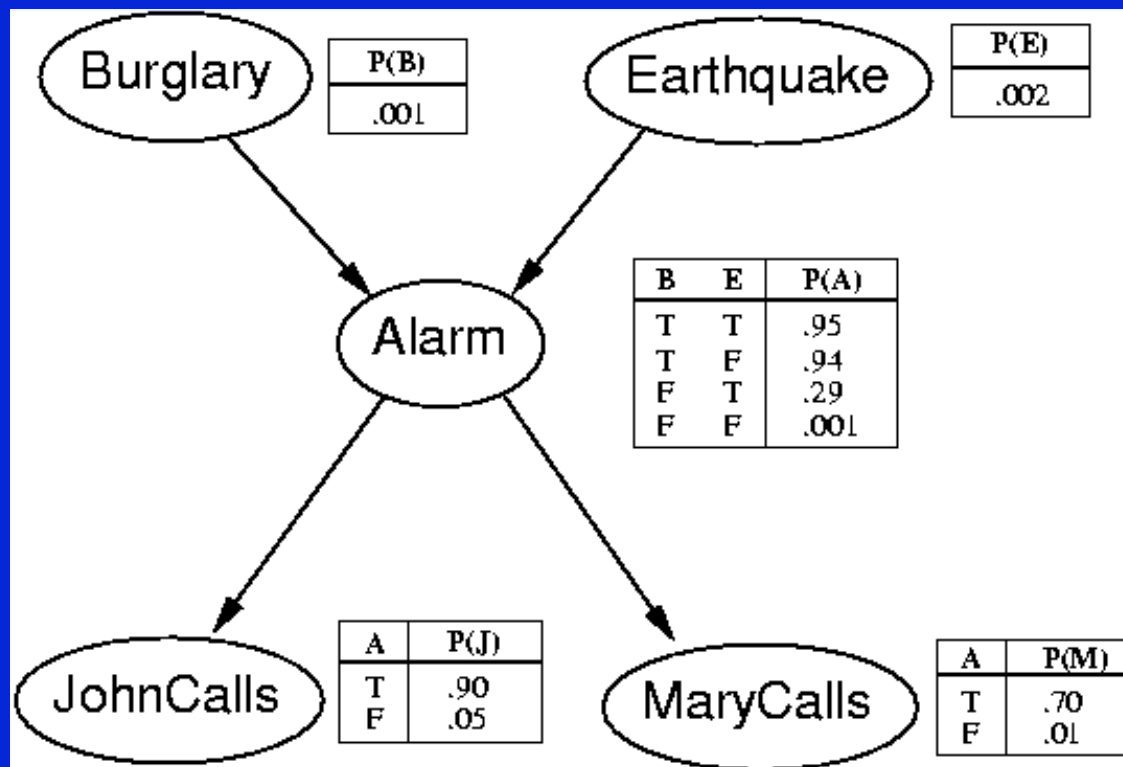
In the simplest case, conditional distribution represented as a conditional probability table (CPT)

Example

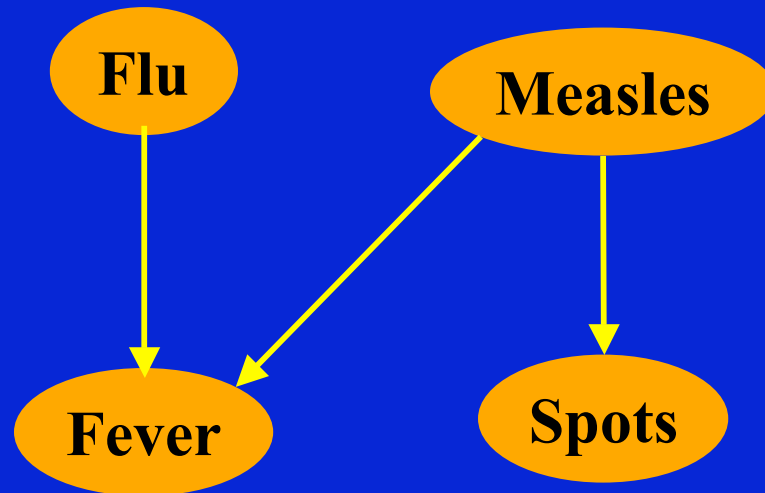
I'm at work, neighbor John calls to say my alarm is ringing, but neighbor Mary doesn't call. Sometimes it's set off by minor earthquakes. Is there a burglar?

Variables: Burglar, Earthquake, Alarm, JohnCalls, MaryCalls

Network topology reflects "causal" knowledge:



Example: Diseases & Symptoms



- Diseases have no parents, and hence have tables with prior probabilities.
- No links between diseases captures belief that diseases are independent of each other.
- No link between Spots & Flu indicates our belief that the flu doesn't cause spots.
- No links between symptoms captures our belief that these are independent of each other given the diseases.

Semantics

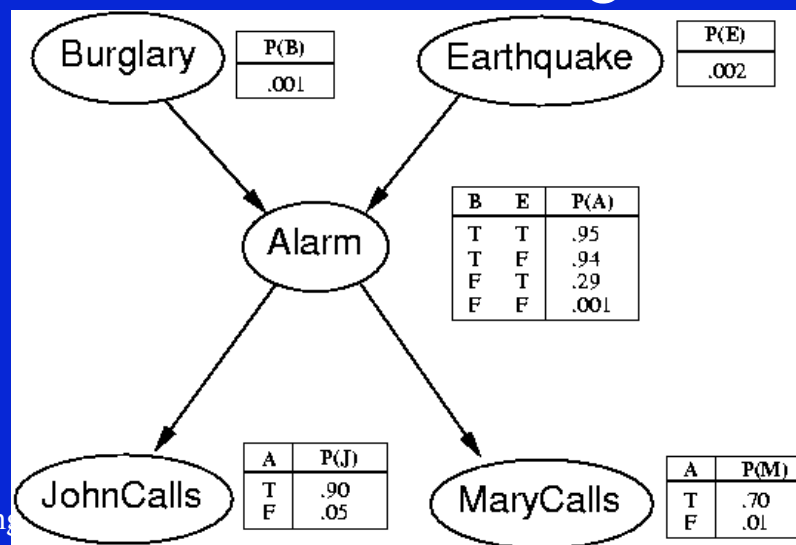
- “Global” semantics defines the full joint distribution as the product of the local conditional distributions:

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i \mid \text{Parents}(X_i))$$

e.g., $P(J \wedge M \wedge A \wedge \neg B \wedge \neg E)$

$$= P(\neg B)P(\neg E)P(A \mid \neg B \wedge \neg E)P(J \mid A)P(M \mid A)$$

- “Local” semantics: each node is conditionally independent of its nondescendants given its parents



Theorem:

Local semantics

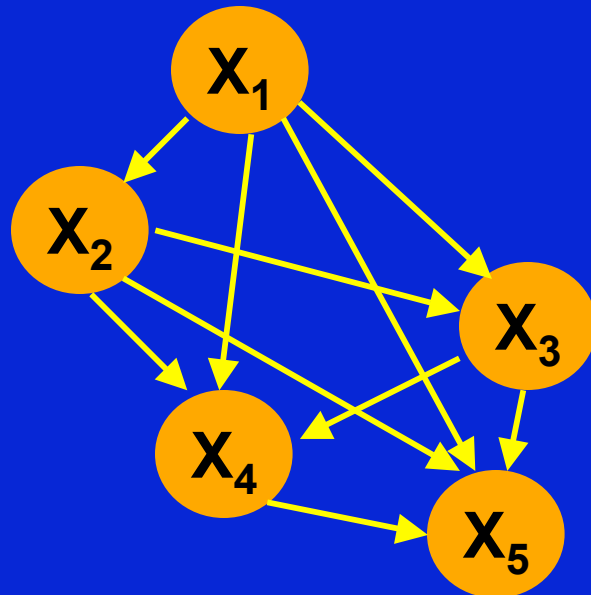
$\Leftrightarrow$ global semantics

Full Joint as fully connected Bayes Net

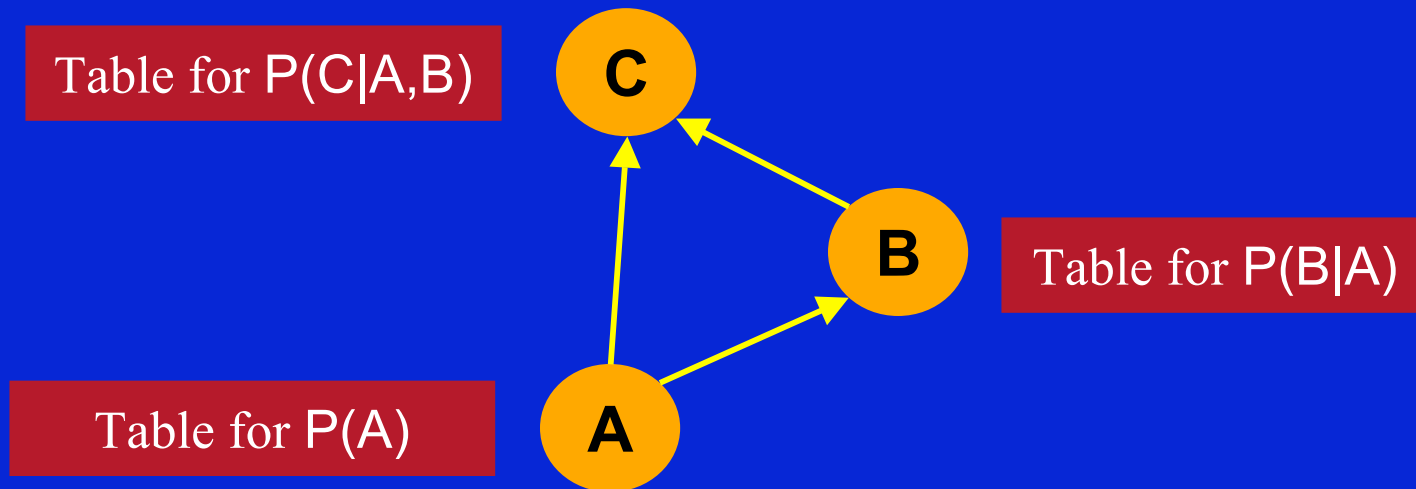
Chain rule is derived by successive application of product rule:

$$\begin{aligned} P(X_1, \dots, X_n) &= P(X_1, \dots, X_{n-1}) P(X_n | X_1, \dots, X_{n-1}) \\ &= P(X_1, \dots, X_{n-2}) P(X_{n-1} | X_1, \dots, X_{n-2}) P(X_n | X_1, \dots, X_{n-1}) \\ &= \prod_{i=1}^n P(X_i | X_1, \dots, X_{i-1}) \end{aligned}$$

What does this look like as a Bayes Net?

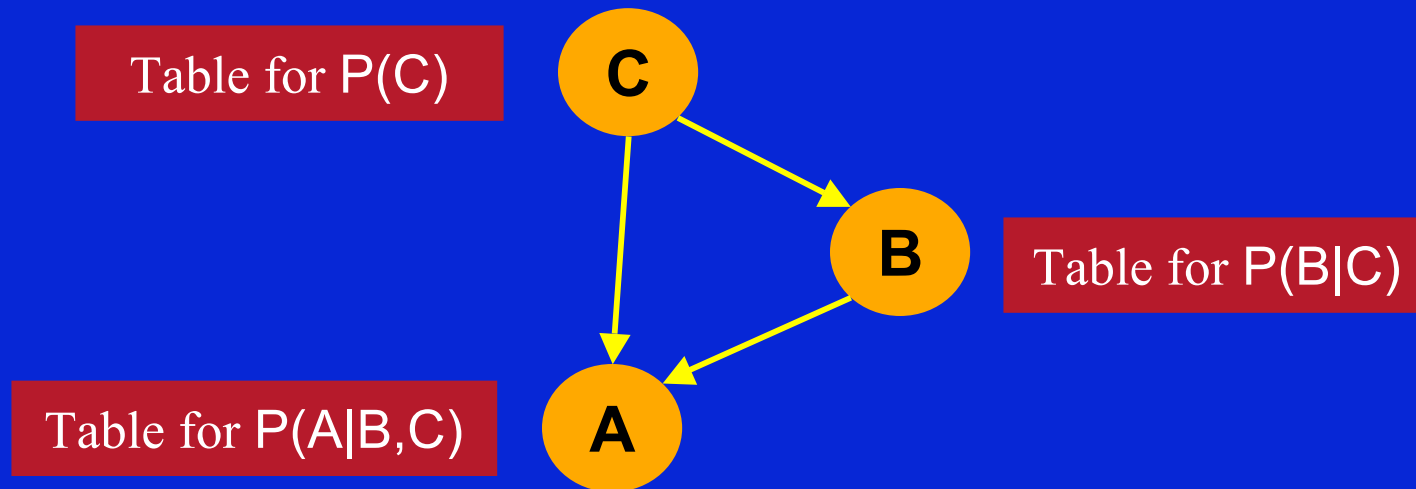


$$P(A,B,C)=P(C|A,B)P(B|A)P(A)$$



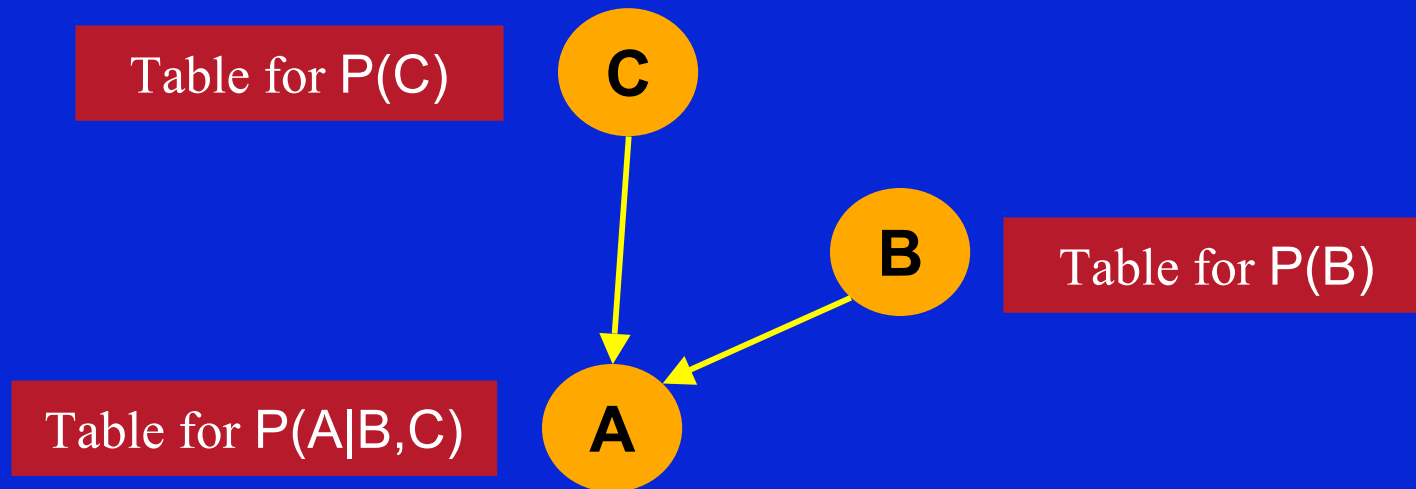
- This is as complicated a network as possible for three random variables
- It is not the only way to represent $P(A,B,C)$ as a Bayes Net.

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



- This is just as complicated a network as the previous network.
- Suppose B and C are independent of each other I.e. $P(B | C) = P(B)$. What does the Bayes net look like?

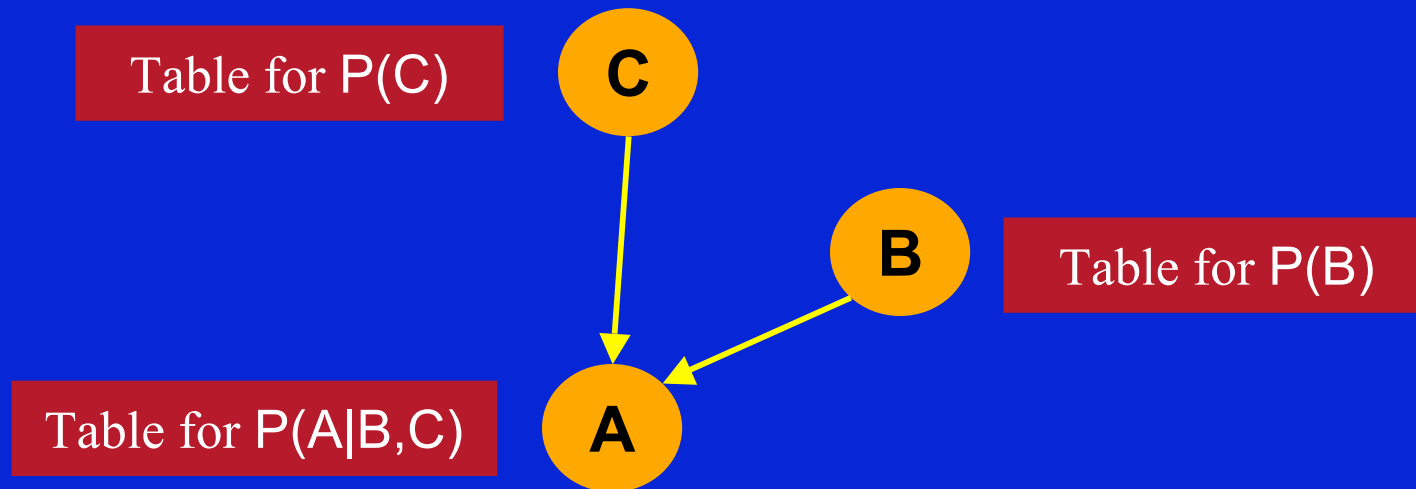
$$P(A,B,C)=P(A|B,C)P(B|C)P(C)$$



Suppose B and C are independent of each other i.e. $P(B | C) = P(B)$. What does the Bayes net look like?

Link between C and B goes away & B's table is simplified...The equation above simplifies also.

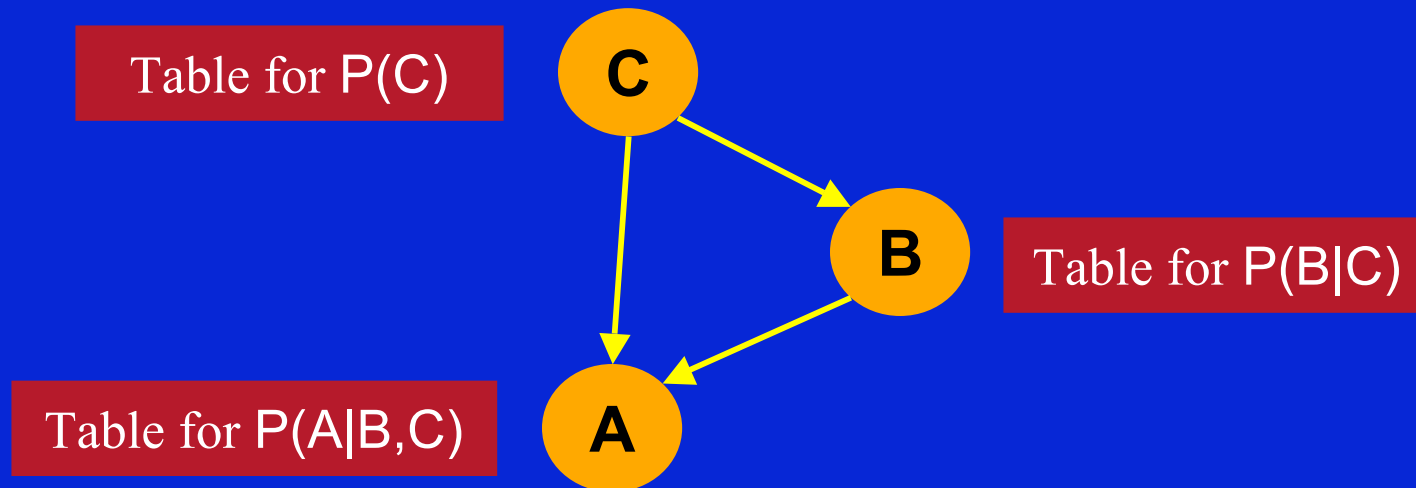
$$P(A,B,C)=P(A|B,C)P(B)P(C)$$



Suppose B and C are independent of each other i.e. $P(B | C) = P(B)$. What does the Bayes net look like?

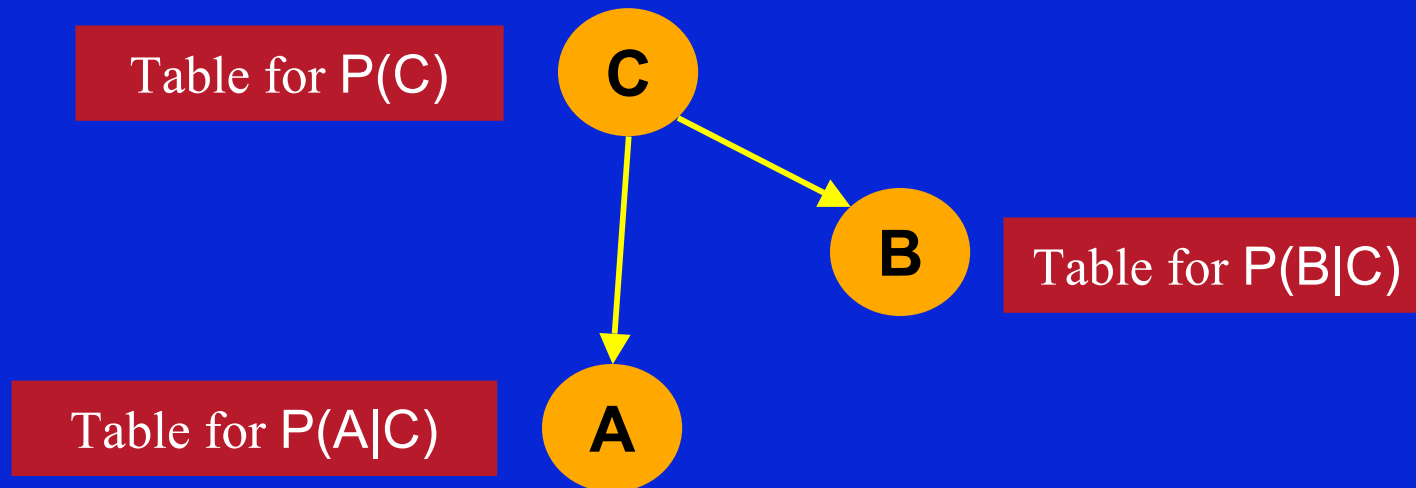
Link between C and B goes away & B's table is simplified, and the equation is simplified.

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



- Suppose A is independent of B given C, i.e. $P(A | B,C) = P(A|C)$. What does Bayes net look like?

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



- Suppose **A** is independent of **B** given **C**,
i.e. $P(A | B,C) = P(A|C)$. What does Bayes net
look like?

Link between **B & **A** disappears and **A**'s table is
simplified**

Constructing belief networks

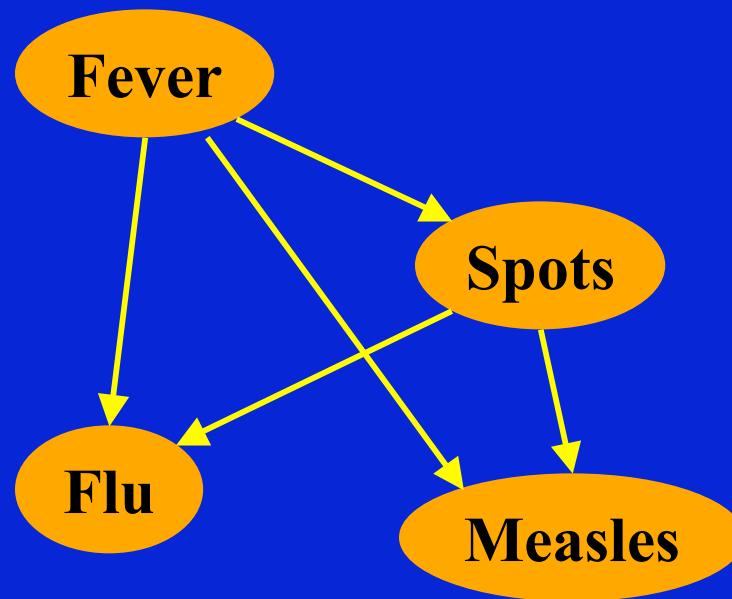
Choose an ordering of variables $X_1, \dots, X_n$

For $i = 1$ to n

- Add node X_i to the network
- Draw link from parents $X_1, \dots, X_{i-1}$ already in network satisfying conditional independence property, I.e. draw a link between X_k and X_i if
$$P(X_i | X_1, \dots, X_k, \dots, X_{i-1}) \neq P(X_i | X_1, \dots, X_{i-1})$$
- Create conditional probability table for node X_i

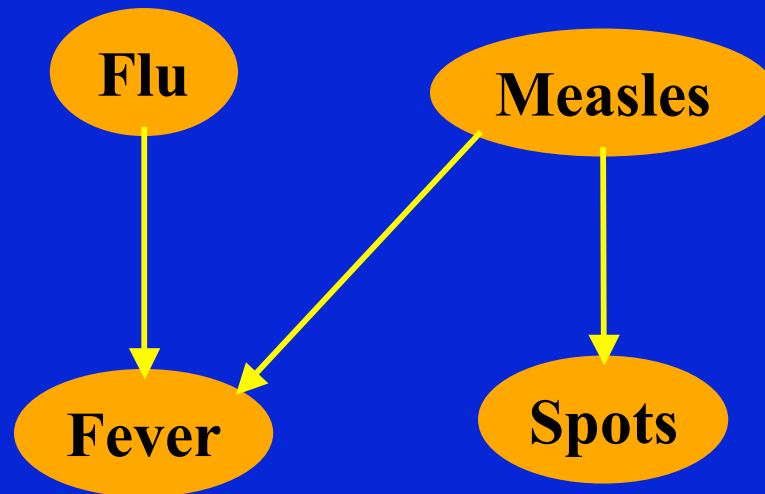
Note that there are many legal belief networks for a set of random variables, and the specific network depends upon the order.

An ordering: Fever, Spots, Flu, Measles



Not too intuitive

Another order: Flu, Measles, Fever, Spots



It is often better to start by
adding causes and then effects

Example

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

$P(J \mid M) = P(J)$? **No**

$P(A \mid J, M) = P(A \mid J)$? **No**

$P(A \mid J, M) = P(A)$? **No**

$P(B \mid A, J, M) = P(B \mid A)$? **Yes**

$P(B \mid A, J, M) = P(B)$? **No**

$P(E \mid B, A, J, M) = P(E \mid A)$? **No**

$P(E \mid B, A, J, M) = P(E \mid A, B)$? **Yes**

