# Bayesian Belief Network

# What Independencies does a Bayes Net Model?

- In order for a Bayesian network to model a probability distribution, the following must be true by definition:

  Each variable is conditionally independent of all its non-descendants in the graph given the value of all its parents.

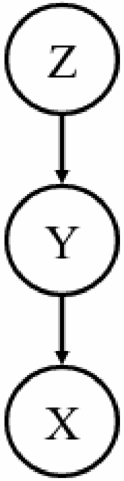
- This implies

$$P(X_1 \ldots X_n) = \prod_{i=1}^{n} P(X_i \mid parents(X_i))$$

- But what else does it imply?

# What Independencies does a Bayes Net Model?

- Example:




Given $Y$, does learning the value of $Z$ tell us nothing new about $X$?

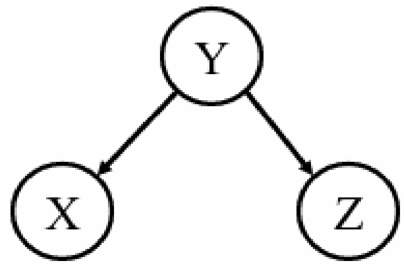
I.e., is $P(X|Y, Z)$ equal to $P(X | Y)$?

Yes. Since we know the value of all of $X$'s parents (namely, $Y$), and $Z$ is not a descendant of $X$, $X$ is conditionally independent of $Z$.

Also, since independence is symmetric, $P(Z|Y, X) = P(Z|Y)$.

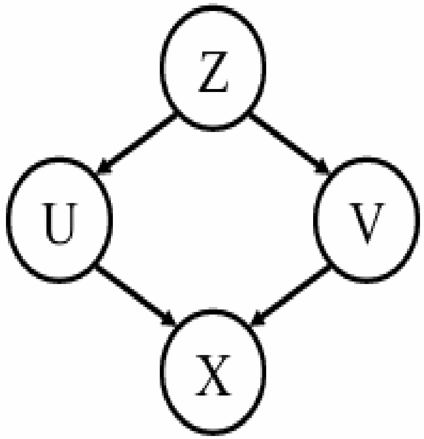
# What Independencies does a Bayes Net Model?

- Let $I<X,Y,Z>$ represent $X$ and $Z$ being conditionally independent given $Y$.



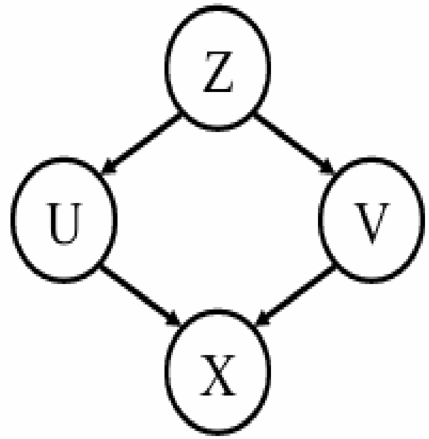

- $I<X,Y,Z>$? Yes, just as in previous example: All X's parents given, and Z is not a descendant.

# What Independencies does a Bayes Net Model?



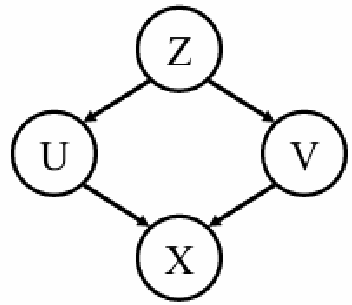

- $I\langle X, \{U\}, Z\rangle$?

# What Independencies does a Bayes Net Model?



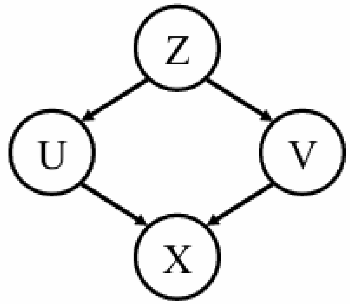

- $I\langle X, \{U\}, Z\rangle$? No.

## What Independencies does a Bayes Net Model?



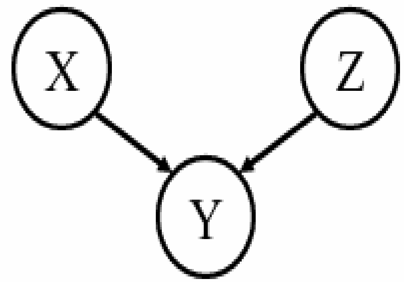

- $I\langle X,\{U,V\},Z\rangle$?

## What Independencies does a Bayes Net Model?




- $I\langle X, \{U, V\}, Z\rangle$? Yes.

# Things get a little more confusing




- X has no parents, so we're know all its parents' values trivially
- Z is not a descendant of X
- So, $I\langle X,\{\},Z\rangle$, even though there's a undirected path from $X$ to Z through an unknown variable $Y$.

# To determine if variables are independent

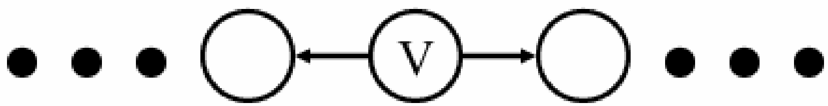
## *d-separation* to the rescue

- Fortunately, there is a relatively simple algorithm for determining whether two variables in a Bayesian network are conditionally independent: *d-separation*.

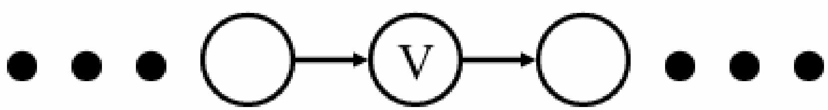
- Definition: $X$ and $Z$ are *d-separated* by a set of evidence variables $E$ iff every undirected path from $X$ to $Z$ is "blocked", where a path is "blocked" iff one or more of the following conditions is true: ...

# A path is "blocked" when...

- There exists a variable *V* on the path such that
  - it **is** in the evidence set *E*
  - the arcs putting *V* in the path are "tail-to-tail"



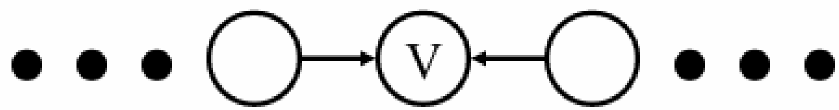

- Or, there exists a variable *V* on the path such that
  - it **is** in the evidence set *E*
  - the arcs putting *V* in the path are "tail-to-head"

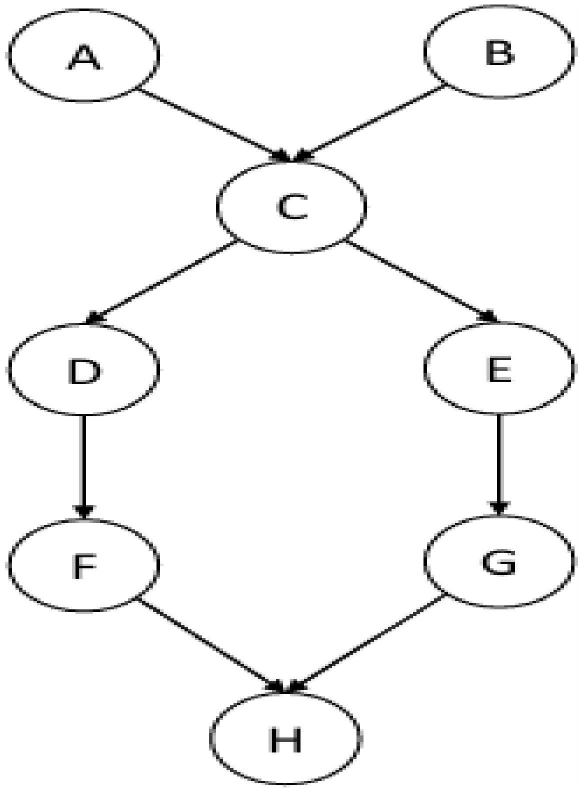



- Or, ...

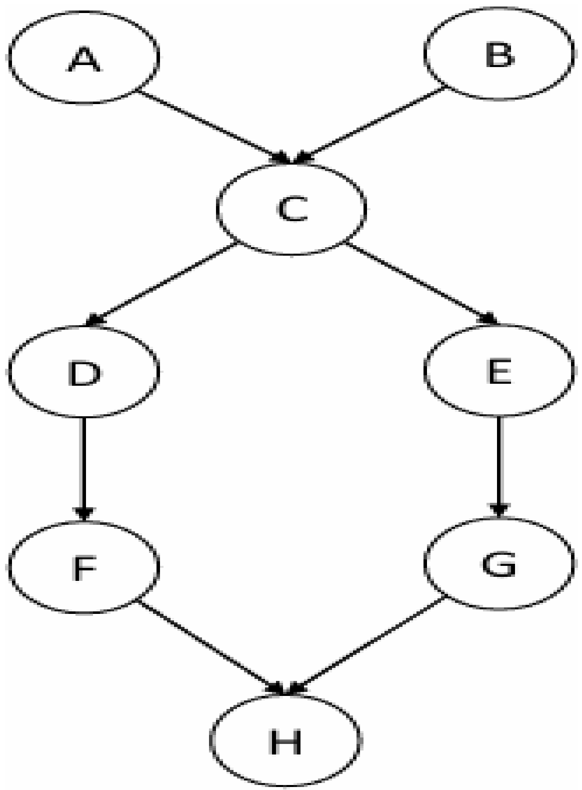
# A path is "blocked" when… (the funky case)

- … Or, there exists a variable $V$ on the path such that
  - it **is NOT** in the evidence set $E$
  - **neither are any of its descendants**
  - the arcs putting $V$ on the path are "head-to-head"



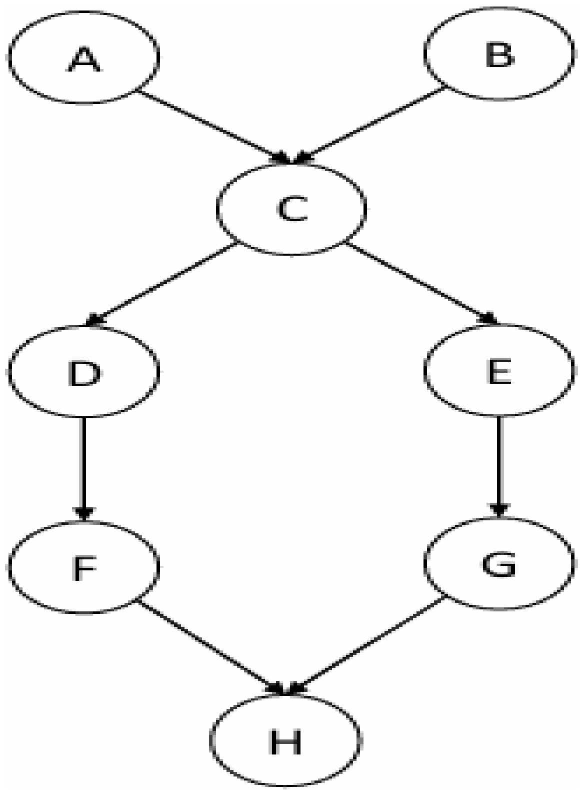



a) Are D and E necessarily independent given evidence about both A and B?

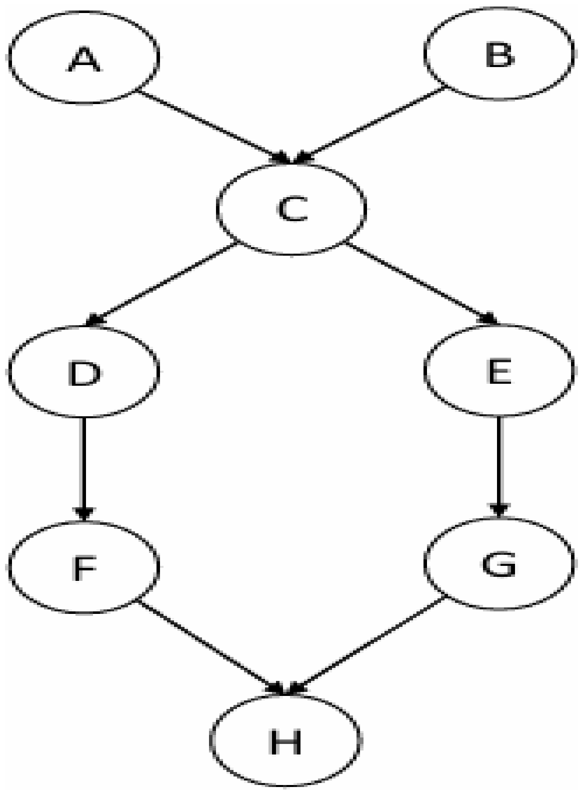



a) Are D and E necessarily independent given evidence about both A and B?

**Answer:**
No. The path D-C-E is not blocked.

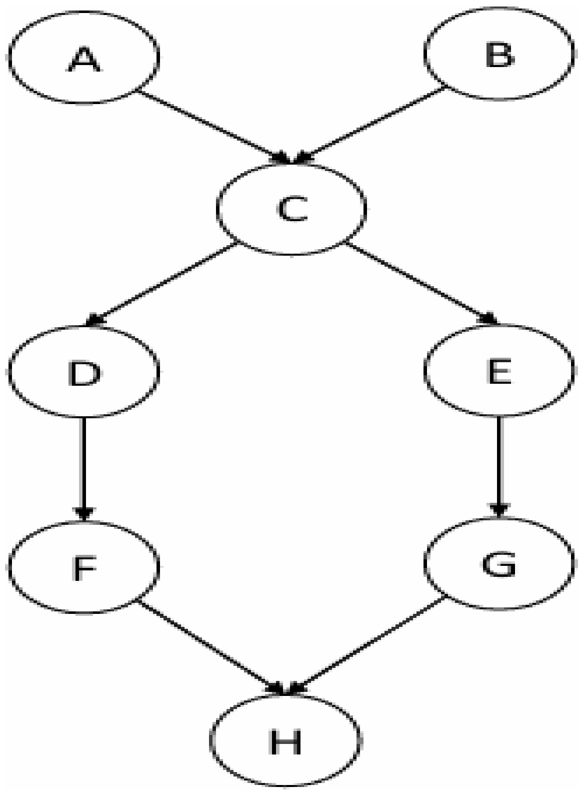



b) Are A and C necessarily independent given evidence about D?

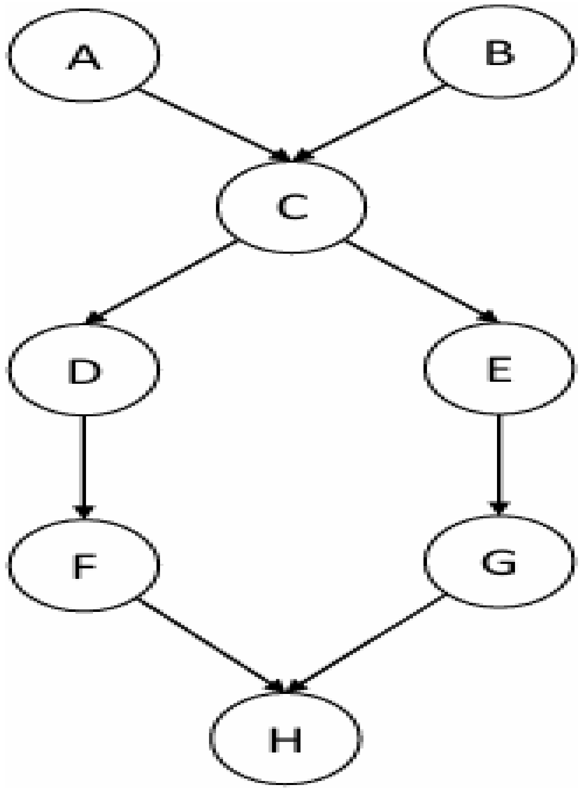



b) Are A and C necessarily independent given evidence about D?

**Answer:**
No. They are directly dependent. The path A-C is not blocked.




c) Are A and H necessarily independent given evidence about C?

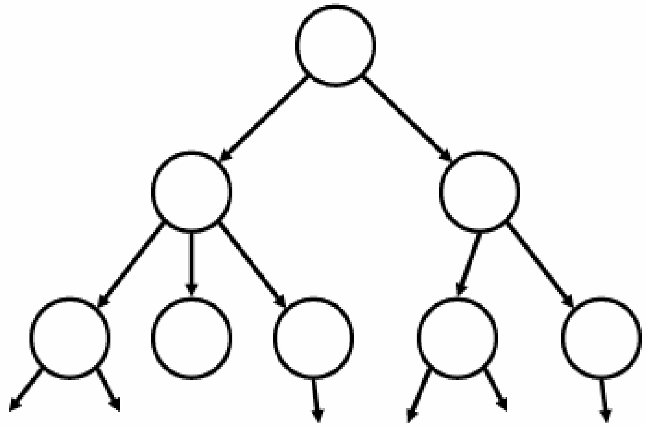



c) Are A and H necessarily independent given evidence about C?

**Answer:**
Yes. All paths from A to H are blocked.

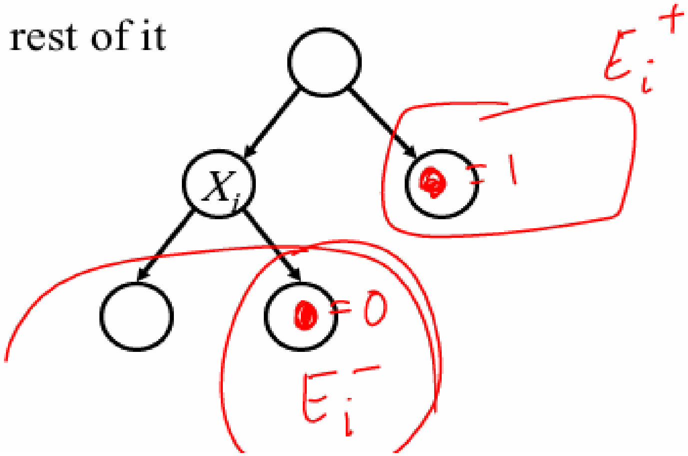
# Bayesian Network Inference

- Let's look a special class of networks: *trees / forests* in which each node has at most one parent.



# Bayesian Network Inference

- Suppose we want $P(X_i \mid E)$ where $E$ is some set of evidence variables.
- Let's split $E$ into two parts:
  - $E_i^-$ is the part consisting of assignments to variables in the subtree rooted at $X_i$
  - $E_i^+$ is the rest of it

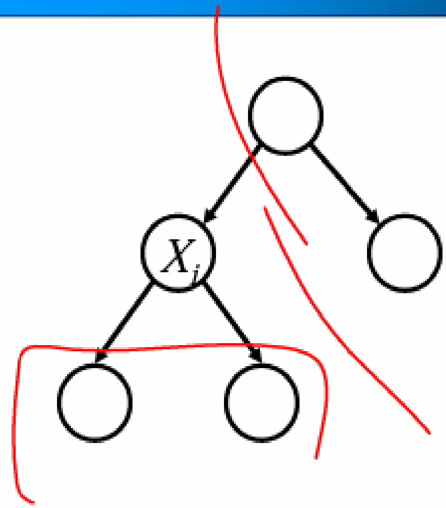


# Decomposing the probabilities, cont'd

$$P(X_i \mid E) = P(X_i \mid E_i^-, E_i^+)$$

$$= \frac{P(E_i^- \mid X_i, E_i^+)\, P(X_i \mid E_i^+)}{P(E_i^- \mid E_i^+)}$$

$$= \alpha\, P(E_i^- \mid X_i)\, P(X_i \mid E_i^+)$$



# Decomposing the probabilities, cont'd

$$P(X_i \mid E) = P(X_i \mid E_i^-, E_i^+)$$

$$= \frac{P(E_i^- \mid X, E_i^+) P(X \mid E_i^+)}{P(E_i^- \mid E_i^+)}$$