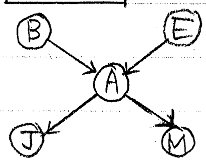
10/5 **Review**




B: burglary, E: earthquake, A: alarm,
J: John calls, M: Mary calls.

* Belief network BN = DAG + CPTs
* Conditional independence

  $P(X_i \mid X_1, X_2, \cdots, X_{i-1}) = P(X_i \mid pa(X_i))$
* Representing CPTs

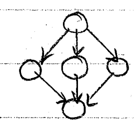
  lookup tables

  logical AND / OR

  noisy-OR

  sigmoid
* Conditional independence

  - A node $X_i$ is conditionally independent of its non-parent ancestors given its parents:



$P(X_i \mid pa(X_i)) = P(X_i \mid X_1, \cdots, X_{i-1})$

  - More generally:

  Let X, Y and E refer to sets of nodes in BN.

  When is X conditionally independent of Y given evidence E?

  When is $P(X \mid E, Y) = P(X \mid E)$?

  $P(X, Y \mid E) = P(X \mid E)\, P(Y \mid E)$?

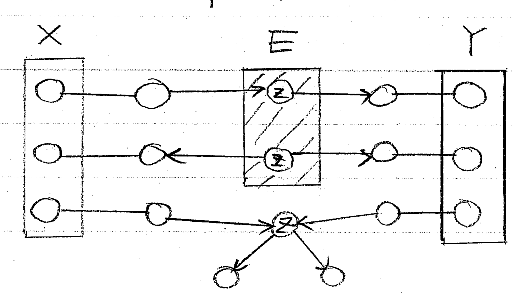
**d-separation**

"direction-dependent"

Relates conditional independence to graph-theoretic properties.

$P(X, Y \mid E) = P(X \mid E)\, P(Y \mid E)$ if and only if every undirected path from a node in X to any node in Y is "d-separated" by E.

<u>Def</u>: a path $\pi$ is d-separated if there exists a node $z \in \pi$ for which one of three conditions hold:




<u>Intuition</u>

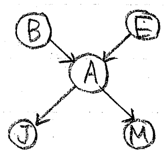
(I) $z \in E$ with $\rightarrow (z) \rightarrow$ is an "intervening" event in a causal chain.

(II) $z \in E$ with $\leftarrow z \rightarrow$ is a common cause / explanation.

(III) $z \notin E$ with $\rightarrow z \leftarrow$ and all descendants $(z) \notin E$.

↳ is an unobserved common effect.

Ex :




I) $P(B \mid A, M) = P(B \mid A)$?

true : A is an intermediate cause.

II) $P(J, M \mid A) = P(J \mid A)\, P(M \mid A)$?

true : A is a common cause or explanation.

III) $P(B, E) = P(B)\, P(E)$?

true : A is an unobserved common effect.

Ex.

$P(B, E \mid A) = P(B \mid A)\, P(E \mid A)$? Not true. Path $B \rightarrow A \leftarrow E$ fails all three conditions.

* Proof that d-sep $\Leftrightarrow$ conditional independence is hard.
* Algorithms exist for efficient tests of d-separation.

## Inference

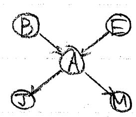
* Problem

$E$ = set of evidence nodes

$Q$ = set of query nodes

How to compute posterior distribution $P(Q \mid E)$?

* Types of inference




- diagnostic reasoning (from effects to causes) e.g., $P(B=1 \mid M=1)$
- causal reasoning (from causes to effects) e.g., $P(M=1 \mid B=1)$
- explaining away (about multiple causes) e.g., $P(B=1 \mid A=1, E=1)$
- mixed (causes and effects) e.g., $P(B=1, M=1 \mid A=1, J=1)$

- When can inference be done efficiently?

(i.e., polynomial time in size of DAG and CPTs)

Polytrees !!

## Polytrees

* A polytree is a graph without loops
  - a singly connected network (at most one undirected path between any two nodes)
* Goal : compute $P(X = x \mid E)$

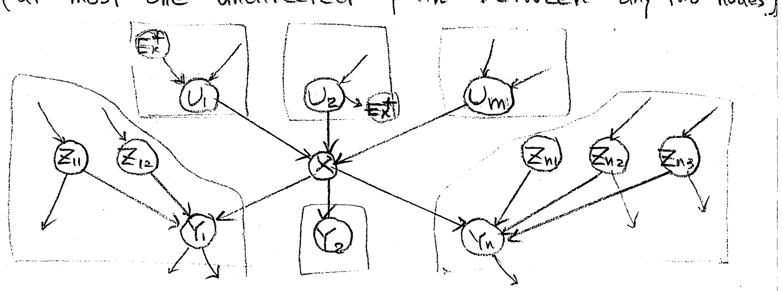
Node $X$

Evidence $E$

Parents $U_i$ $(i = 1, \cdots, m$ # parents$)$

Children $Y_j$ $(j = 1, \cdots, n$ # children$)$

Parents of children $Z_{jk}$ (excluding $X$)



* Note : Boxes don't overlap for polytrees — otherwise, there are loops.
* Types of evidence :

  $E_X^+$ = evidence "above" X connected thru parents.

  $E_X^-$ = evidence "below" X connected thru children.

  Assume $X \notin E$. (otherwise trivial).

  $E = E_X^+ \cup E_X^-$
* General strategy : recursion

$$P(X=x \mid E) = P(X=x \mid E_X^-, E_X^+)$$

$$= \frac{P(E_X^- \mid X=x, E_X^+)\, P(X=x \mid E_X^+)}{P(E_X^- \mid E_X^+)} \quad \text{generalized Bayes rule}$$

$$= \frac{P(E_X^- \mid X=x)\, P(X=x \mid E_X^+)}{P(E_X^- \mid E_X^+)} \quad \text{d-separation case I, } X = \text{intermediate event}$$

* Plan of attack
  - compute numerator up to constant factor. (independent of little "$x$").
  - compute denominator by $P(E_X^- \mid E_X^+) = \sum_x P(E_X^- \mid X=x)\, P(X=x \mid E_X^+)$ ← by normalization.
  - constant factors will divide out of ratio.
* Upstream recursion

$$P(X \mid E_X^+) = \sum_{\vec{u}} P(X, \vec{U}=\vec{u} \mid E_X^+) \quad \text{marginalization}$$

$$= \sum_{\vec{u}} P(X \mid \vec{U}=\vec{u}, E_X^+)\, P(\vec{U}=\vec{u} \mid E_X^+) \quad \text{product rule.}$$

$$= \sum_{\vec{u}} P(X \mid \vec{U}=\vec{u})\, P(\vec{U}=\vec{u} \mid E_X^+) \quad \text{d-separation I or II}$$

$$= \sum_{\vec{u}} P(X \mid \vec{U}=\vec{u}) \prod_{i=1}^{m} P(U_i=u_i \mid E_X^+) \quad \text{d-separation III } X = \text{unobserved common effect.}$$

[Notation $E_{U_i \setminus X}$ = evidence connected to $U_i$ except via path thru X.
(evidence in its own box)]

$$= \sum_{\vec{u}} \underbrace{P(X \mid \vec{U}=\vec{u})}_{\substack{\text{CPT}\\ \text{(given)}}} \prod_{i=1}^{m} \underbrace{P(U_i=u_i \mid E_{U_i \setminus X})}_{\text{recurse on parents.}} \quad \text{d-separation case III.}$$