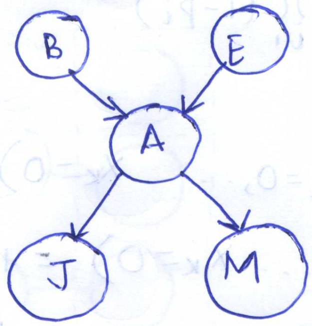
Review

B = burglary
E = earthquake
A = alarm
J = John calls
M = Mary calls



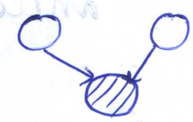
* Belief network (BN) = DAG + CPTs
* Conditional independence

$$P(X_1, X_2, \ldots X_n) = P(X_1)\,P(X_2/X_1) \ldots\ldots P(X_n \mid X_1, X_2, \ldots, X_{n-1})$$ product rule

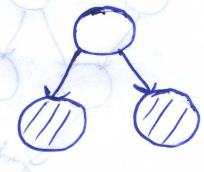
$$= \prod_i P(X_i / X_1, X_2, \ldots X_{i-1})$$

$$= \prod_i P(X_i / pa(X_i))$$ ← parents of X_i

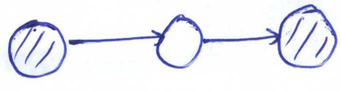
* Types of reasoning
  1) Competing ~~events~~ explanations of observed event



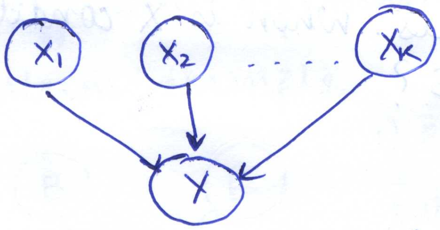
  2) Multiple events with common explanation



  3) Interfering events



* Representing CPTs



Suppose $X_i \in \{0, 1\}$
and $Y \in \{0, 1\}$
How to represent CPT $P(Y=1/X_1, X_2, \ldots X_K)$?

(i) lookup table

$O(2^k)$ can store arbitrary CPT

x_1	x_2	. . .	x_K	$P(Y=1/x_1, x_2, \ldots x_K)$
0	0		0	
1	0		0	
0	1		0	

(ii) "deterministic" node

"AND" $P(Y=1/x_1, x_2, \ldots x_K) = \prod_{i=1}^{k} x_i$

"OR" $P(Y=0/x_1, x_2, \ldots x_K) = \prod_{i=1}^{k} (1-x_i)$

(iii) noisy-OR node

Use k numbers $p_i \in [0,1]$ to parametrize $O(2^k)$ elements for $x_i \in \{0, 1\}$

$$P(Y=0/x_1, x_2, \ldots x_K) = \prod_{i=1}^{k} (1-p_i)^{x_i}$$

$$P(Y=1 / X_1, X_2, \ldots X_K) = 1 - \prod_{i=1}^{K} (1-p_i)^{X_i}$$

Why called "noisy-OR"?

Look at $P(Y=1 / X_1=0, X_2=0, \ldots X_K=0) = 1 - \prod_i \underbrace{(1-p_i)^0}_{1} = 0$

$P(Y=1 / X_1=1, X_2=0, X_3=0, \ldots X_K=0) = 1 - (1-p_1)^{(1)} \prod_{j>1}^{K} \underbrace{(1-p_j)^0}_{1}$

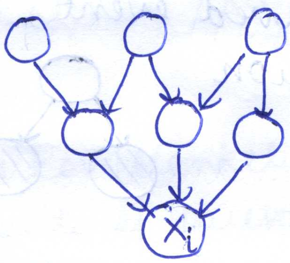
only $X_1 = 1$

$= p_1$

Intuition: $p_i \in [0,1]$ is the probability that $X_i = 1$ by itself triggers $Y_i = 1$

conditional independence

A node is conditionally independent of its non-parent ancestors given its parents.



$$P(X_i / pa(X_i)) = P(X_i / X_1, X_2, \ldots X_{i-1})$$

More generally -

Let X, Y and E refer to sets of nodes. When is X conditionally independent of Y given evidence E?

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

$P(Y|E,X) = P(Y|E)$?

$P(X,Y|E) = P(X|E)P(Y|E)$?

e.g. $X = \{X_i\}$

$E = \{pa(X_i)\}$

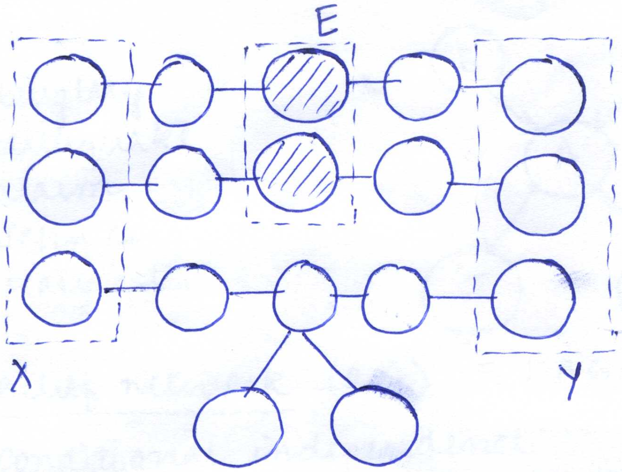
$Y = \{X_1, X_2, \ldots X_{i-1}\} - pa(X_i)$

* d-separation

"direction-dependent" separation

Relate conditional independence to graph-theoretic properties

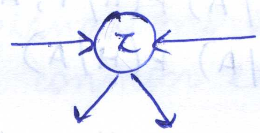
$P(X,Y|E) = P(X|E)P(Y|E)$ if and only if every undirected path (ignoring edge direction) from any node in X to any node in Y is "d-separated" by E.



<u>Definition</u> : a path π is <u>d-separated</u> if there exists a node $z \in \pi$ for which one of 3 conditions hold :

1) $z \in E$ with  is an intervening event

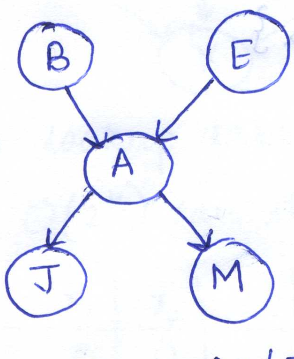
2) $z \in E$ with  is a common explanation

3) $z \notin E$, $desc(z) \notin E$,  no observed common effects

* Proof that d-separation $\Longleftrightarrow$ conditional independence is difficult, beyond course.

* Efficient algorithm exists for tests of d-separation

* Alarm example



$B \to X$
$A \to E$
$M \to Y$

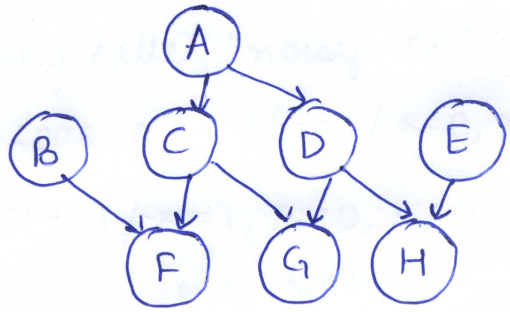
1) $P(B|A,M) \stackrel{?}{=} P(B|A)$
true, alarm is an intervening event

2) $P(J,M/A) \stackrel{?}{=} P(J/A)\,P(M/A)$
true, alarm is common explanation of J, M

3) $P(B/E) = P(B)$
true, condition III

4) $P(B,E/J) \stackrel{?}{=} P(B/J)\,P(E/J)$
false, explaining away
condition III is not true because $desc(A) \in E$.
$\swarrow J$

* Loopy BN ~~com~~ example



Statement	T/F
$P(D/H) \overset{?}{=} P(D/E,H)$	false
$P(F,H/A) \overset{?}{=} P(F\|A)\,P(H/A)$	true
$P(F,G,H/A) \overset{?}{=} P(F/A)\,P(G\|A)\,P(H/A)$	false

→ 2 paths to check :

$F \leftarrow C \leftarrow A \rightarrow D \rightarrow H$ II

$F \leftarrow C \rightarrow G \leftarrow D \rightarrow H$ condition III

<u>Note</u>: $P(F,G,H/A) = P(F/A)\,P(H/F,A)\,P(G|F,H,A)$ [Product rule]

$= P(F|A)\,P(H/A)\,P(G|F,H,A)$ [conditional independence]

<u>Def</u>: Markov blanket B_x of individual node X consists of parents and children of X and <u>parents of children</u> "spouses" of X (not including X)

<u>Theorem</u>: a node X is conditionally independent of all nodes outside B_x given nodes inside B_x.

$P(X/B_x, Y) = P(X|B_x)$ where $Y \notin \{X, B_x\}$