FUDAN SUMMER SCHOOL INTRODUCTION TO AI

Probabilistic Reasoning and Decision Making

Day 4 – Conditional Independence, d-separation

Conditional independence revisited

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

$$= \prod_{i=1}^{n} P(X_i|X_1, X_2, \dots, X_{i-1})$$

Suppose that in any given domain:

How to build a Bayes Net?

1. Choose random variables
2. Choose an ordering
3. While there are variables left:
    1. *Add node* X_i *to the BN*
    2. *Set parents of* X_i *to the minimal subset such that* X_i *is conditionally independent from non-parent nodes preceding i, given its parents.*
    3. *Define* $P(X_i|parent(X_i))$ *CPT*

How to build a Bayes Net?

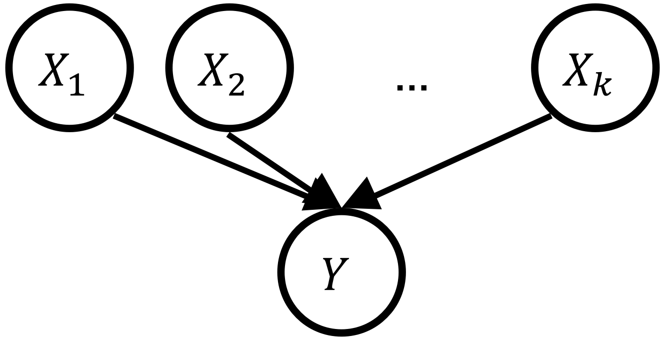
1. Choose random variables
2. Choose an ordering
3. While there are variables left:
    1. *Add node* X_i *to the BN*
    2. *Set parents of* X_i *to the minimal subset such that* X_i *is conditionally independent from non-parent nodes preceding i, given its parents.*
    3. *Define* $P(X_i|parent(X_i))$ *CPT*

Which is the better ordering to choose?

A. Root events first (i.e. causes), e.g. B, E, A, J, M
B. Child events first (i.e. consequences), e.g. M, J, A, E, B
C. It doesn't matter

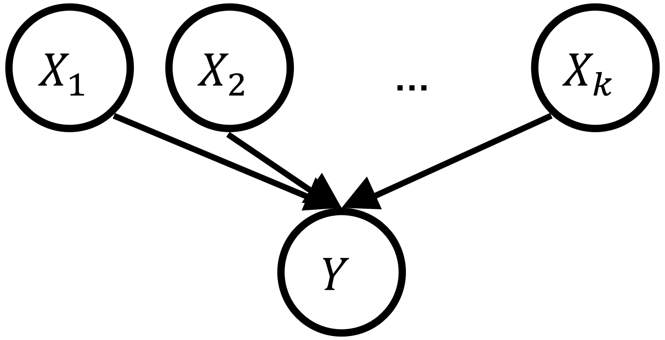
Representing CPTs

How do we represent the CPT?



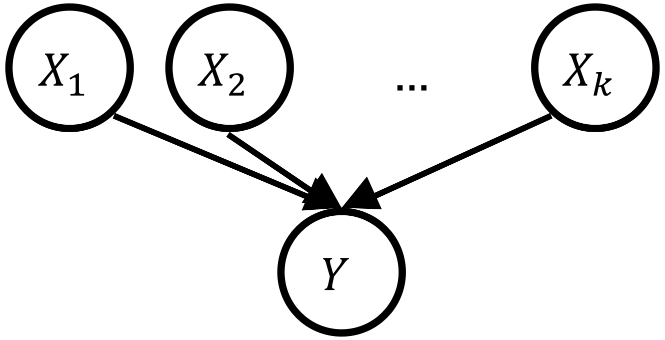
Representing CPTs

How do we represent the CPT?

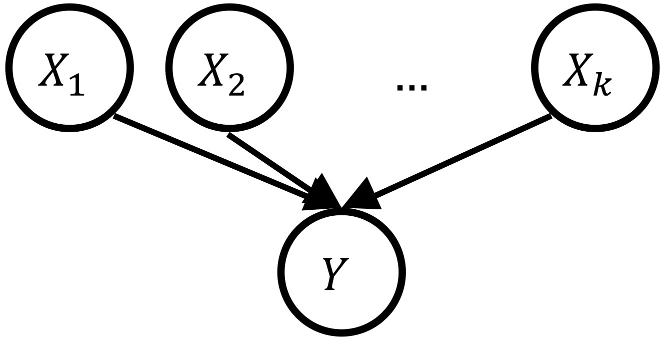


Representing CPTs

How do we represent the CPT?



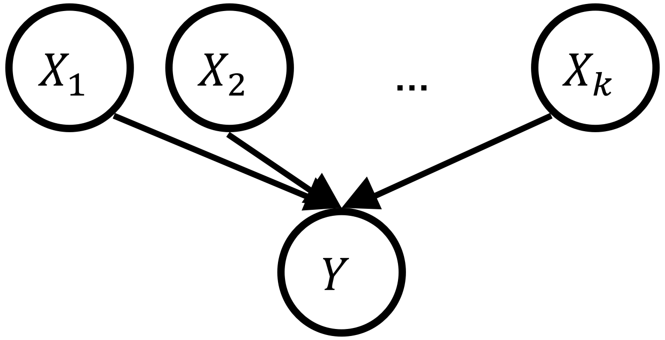
Noisy-OR



What is the value of $P(Y = 1 | X_1 = 0, X_2 = 0, \ldots, X_k = 0)$?

A. 0
B. p_i
C. 1
D. None of these
E. Not enough information given

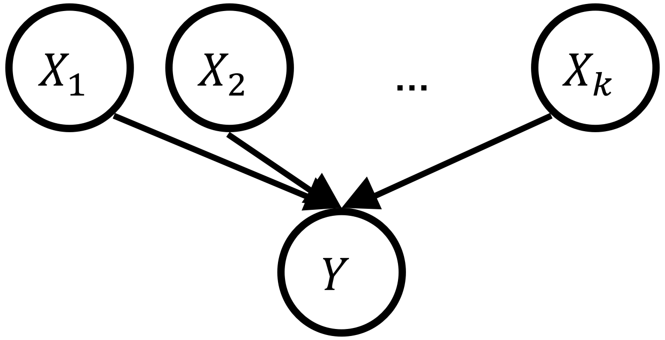
Noisy-OR



What is the value of $P(Y = 1 | X_1 = 0, X_2 = 0, \dots, X_i = 1, \dots, X_k = 0)$?

A. 0
B. p_i
C. 1
D. None of these
E. Not enough information given

Noisy-OR



Let $P(Y = 1 | X_1 = 0, \dots, X_i = 1, X_j = 1, \dots, X_k = 0) = \frac{1}{3}$

and $p_i = \frac{1}{5}$

What is p_j?

A. $^{1}/_{6}$

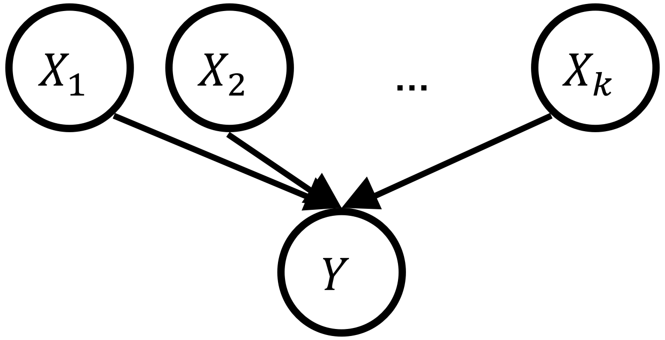
B. $^{2}/_{5}$

C. $^{1}/_{15}$

D. $^{2}/_{3}$

E. I have no idea

Noisy-OR



Let $P\left(Y=1|X_1=0,\dots,X_i=1,X_j=1,\dots,X_k=0\right)=\frac{1}{3}$

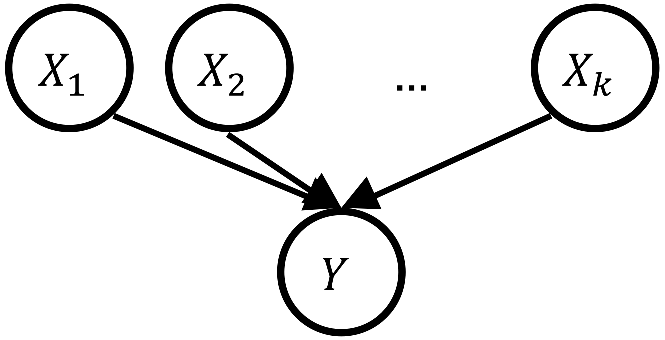
and $p_i=\frac{1}{5}$

What is p_j?

A. $^1/_6$
B. $^2/_5$
C. $^1/_{15}$
D. $^2/_3$
E. I have no idea

$$P(Y=1|X_1,\dots,X_k)=1-\prod_{i=1}^{k}(1-p_i)^{X_i}$$

Noisy-OR

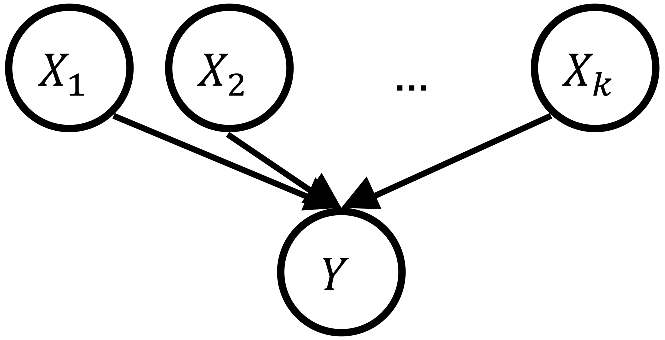


Is it true that P(Xi=1) = pi

A. Yes
B. NO

$$P(Y = 1|X_1, \dots, X_k) = 1 - \prod_{i=1}^{k}(1 - p_i)^{X_i}$$

Noisy-OR: Intuition

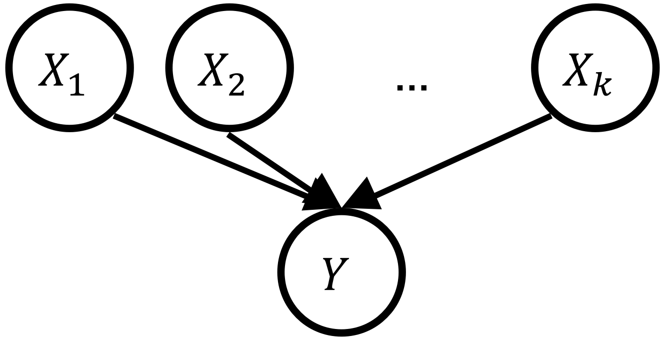


$$P(Y = 1 | X_1, \dots, X_k) = 1 - \prod_{i=1}^{k} (1 - p_i)^{X_i}$$

Let $P(Y = 1 | X_1 = 1, X_2 = 1, \dots, X_{i-1} = 1, X_i = 0, \dots, X_k = 0) = a$
and $P(Y = 1 | X_1 = 1, X_2 = 1, \dots, X_{i-1} = 1, X_i = 1, X_{i+1} = 0, \dots, X_k = 0) = b$
What is the relationship between a and b?

A. $a \leq b$
B. $a = b$
C. $a \geq b$
D. There is not enough information to determine the relationship

Noisy-OR: Intuition



$$P(Y = 1 | X_1, \ldots, X_k) = 1 - \prod_{i=1}^{k} (1 - p_i)^{X_i}$$

- When all parents are off, child is off with certainty (just like OR).
- Each parent that turns on increases (or at least does not decrease) the probability that the child will turn on.
- When all the parents are on, it still might not be certain that the child is on, but it is probably very likely.

Bayesian Networks: Network structure and conditional independence

- By our construction of a Bayes' net, we know that a node is conditionally independent of its non-parent ancestors given its parents:

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

$$= \prod_{i=1}^{n} P(X_i|X_1, X_2, \dots, X_{i-1})$$

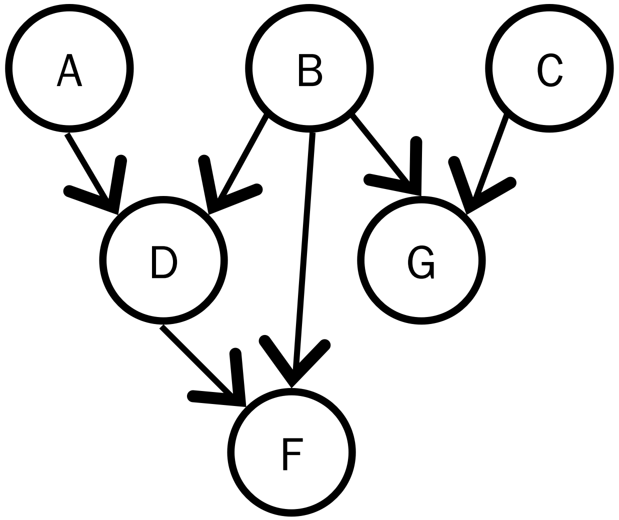
$$= \prod_{i=1}^{n} P(X_i|parents(X_i))$$

But what about more general assertions of independence?

Bayesian Networks: Network structure and conditional independence

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

Write three equalities using $P(X|Y,E)$, $P(Y|X,E)$, $P(X,Y|E)$, etc, that indicate this conditional independence.

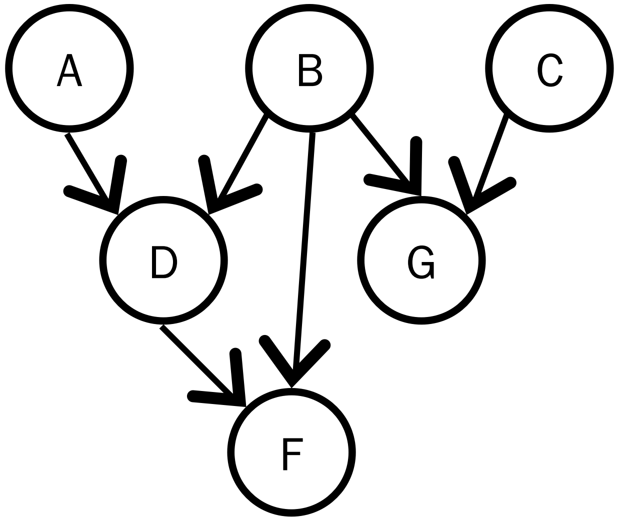


Bayesian Networks: Network structure and conditional independence

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

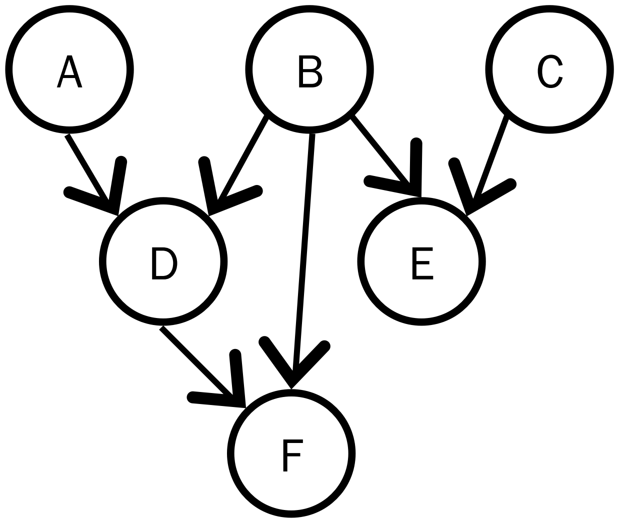
<u>Theorem</u>: $P(X, Y|E) = P(X|E)P(Y|E)$ if and only if every UNDIRECTED path from a node in X to a node in Y is "d-separated" ("blocked") by E.

(we will not prove this)



d-separation ("dependence" separation)

Definition: A path π (sequence of nodes connected by edges) is d-separated if there exists some node $z \in \pi$ for which one of the following three conditions hold:



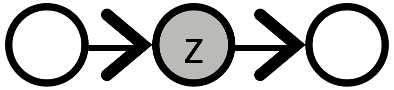
Conditional Independence

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

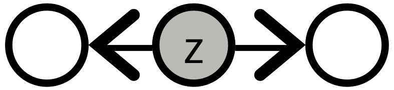
A. True
B. False

<u>Theorem</u>: $P(X,Y|E) = P(X|E)P(Y|E)$ if and only if **every** UNDIRECTED path from a node in X to a node in Y is "d-separated" ("blocked") by E.

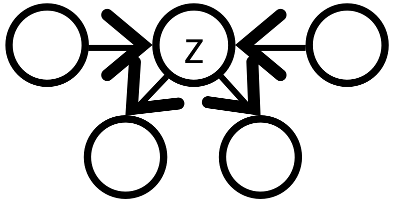
<u>Definition</u>: A path π (sequence of nodes connected by edges) is d-separated if there exists some node $z \in \pi$ for which **one** of the following three conditions hold:

1. 

$z \in E$ (observed event in causal chain)

2. 

$z \in E$ (observed common explanation)

3. 

$z \notin E, descendants(z) \notin E$ (no observed common effect)