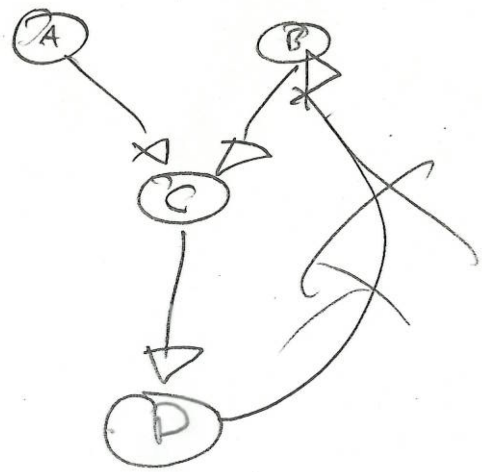
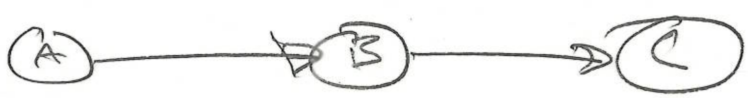
Bayes Nets:

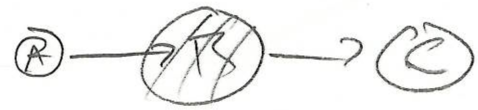


Directed,
no cycles

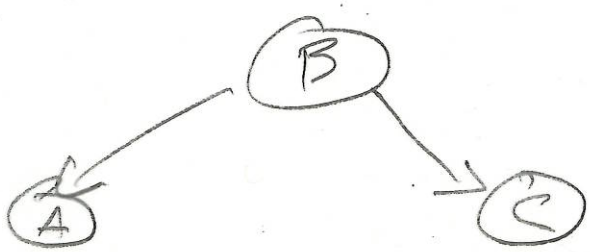
$$PDF = \prod_{nodes} P(\text{node} \mid \text{parents})$$

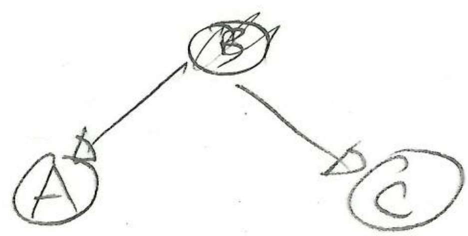
$A \perp C$



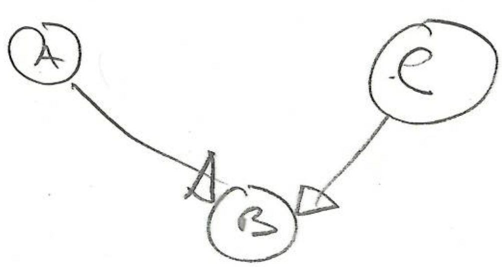


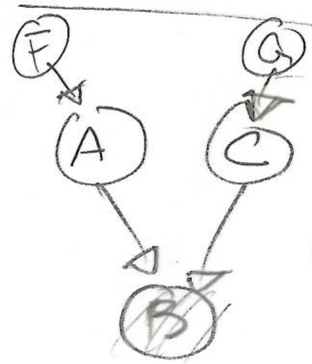
$A \not\perp C$





$A \perp C$





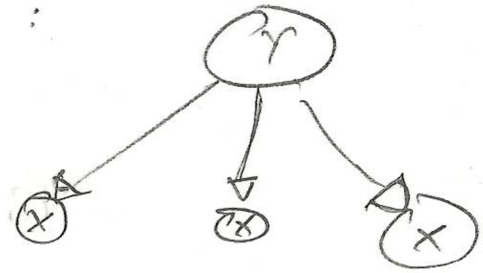
$A \not\perp C$

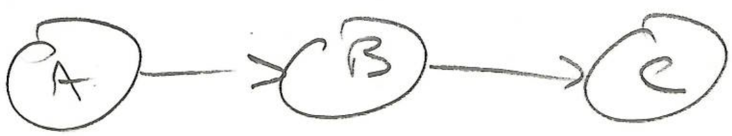
D-separation on arbitrary graph

Two nodes are independent if there is no valid path between them (all paths are d-separated)

Two nodes are dependent if <u>any</u> path between them is not d-separated

Naive Bayes:





$P(C)$

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

$$P(C) = P(A,B,C) + P(\sim A,B,C) + P(\sim A,\sim B,C) + P(A,\sim B,C)$$

Conditional Distributions

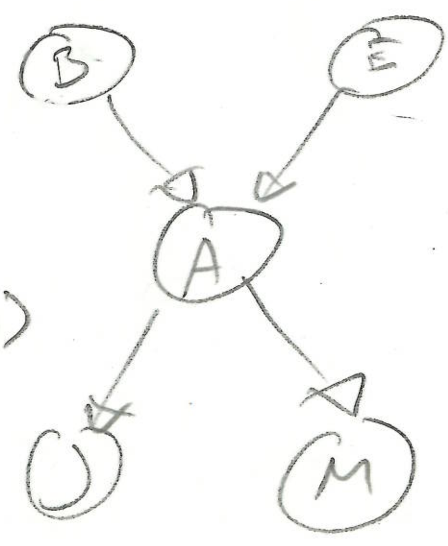
$A \rightarrow B \rightarrow C$ $\qquad P(A \mid C)$

$$P(A \mid C) = \frac{P(A\ C)}{P(C \mid A)P(A) + P(C \mid \sim A)P(\sim A)}$$

$$= \frac{P(A, C)}{P(A, C) + P(\sim A, C)}$$

A slightly better way ....
variable elimination

$$P(B, J, M) = ?$$



$$= \sum_a \sum_e P(B)P(e)P(a \mid B\ e) \, P(J \mid a)P(M \mid a)$$

$$= P(B) \sum_e \left[P(e) \sum_a \left[P(a \mid B\ e)\, P(J \mid a) P(M \mid a) \right] \right]$$

Rewrite probability statements as functions...

$$f_A(a, B, e) = P(a \mid B e)$$

$$f_J(a) = P(J \mid a)$$

$$f_M(a) = P(M \mid a)$$

$$P(B, J, M) = P(B) \sum_e P(e) \sum_a f_J(a) f_M(a) f_A(a, B, e)$$

$$f_{A,J,M}(B, e) = \sum_a f_J(a) f_M(a) f_A(a, B, e)$$

$$= P(B) \sum_e P(e) f_{AJM}(B, e)$$

$$f_{E,A,J,M}(B) = \sum_e P(e) f_{AJM}(B, e)$$

encapsulate the marginalizations within functions...

$$P(B, J, M) = P(B) f_{EAJM}(B)$$

BAM! variables eliminated!

expand this back out....

same function call

$$\left(f_j(a) + \dots \right) + \left(f_j(a) \dots \right)$$

$$\left(\underbrace{f_{AKCDE}(j)} + \dots\right) + \left(\underbrace{f_{AKCDE}(j)}\right)$$

$\sim 2^5$ calculations

$\sim 2^5$ calculations that I don't have to do

except.... I still had to do those 2^5 calculations in the first place

What to do when variable elimination fails?

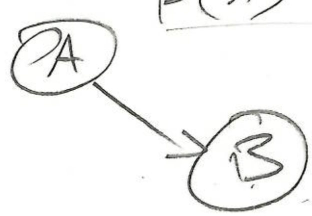
$\boxed{\text{Sample!}}$

If we have all conditional probability parameters...

Given some values (observations) from the graph, we can sample the rest

We should always have "free variables," nodes w/o parents = easy to sample

$$P(A)\begin{cases} a=1 & 0.3 \\ a=0 & 0.7 \end{cases}$$



$P(B|A)$

naive sampling

- start w/ free nodes
- sample rest

$P(C=1)$? $\frac{\# C=1}{\text{total samples}} = \frac{\sum I(C=1)}{\text{total samples}}$

$P(A=1) \equiv 0.000001$

naive sampling will always pick $A=0$:(

⟶ Weighted sampling

- Pick some <u>arbitrary</u> observations
- Calcute weight of assignment

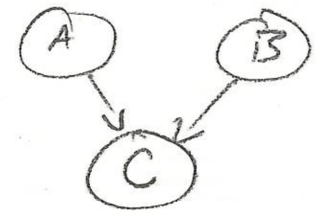
for example, $w = P(A=0)\,P(B=1)$

- sample the rest

Instead of counting, increment

by the weight

$$\hat{P}(C) = \frac{\sum w_i\, I(C=1)}{\sum w_i}$$



Set $A=1$, $B=0$

$w = P(A=1)\,P(B=0)$

- sample C a bunch

$$P(C=1 \mid A=1, B=0) = \frac{\sum w \; I(C=1)}{\sum_i w_i}$$

MCMC - Markov Chain Monte Carlo

- random walk
- iterate through nodes, resampling based on other values

$P(A)$? $P(B|A)$? $P(C|B)$

$$A \longrightarrow B \longrightarrow C$$

wanted $P(C)$? or $P(C|A)$?

- Need to estimate $P(B|A)$
- Get some observations
- E.M.
- Profit

→ More on this with HMMs later