BAYESIAN NETWORKS

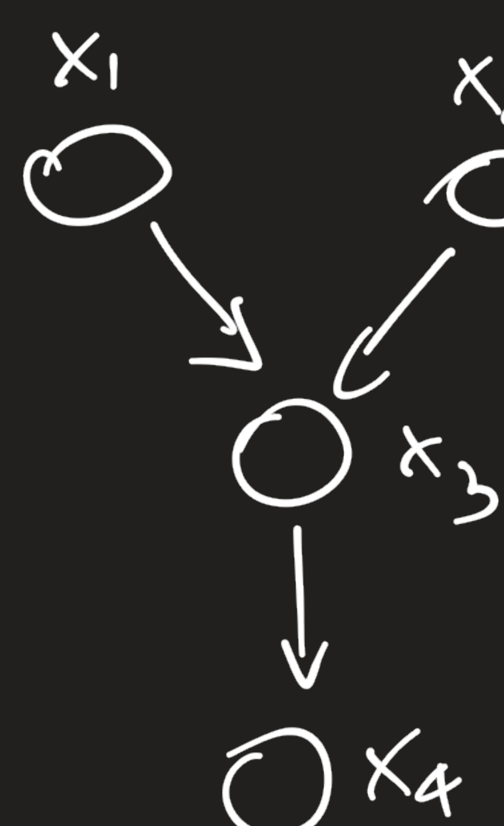
$$p(x_1,x_2,x_3,x_4) = p(x_4|x_3)\,p(x_3|x_1,x_2)\,p(x_2)\,p(x_1)$$

pa_k = parents of x_k $\quad pa_k \subseteq \{x_1, \dots, x_{k-1}\}$ are in the conditioning set of x_k

$$\Downarrow$$

$p(x_k \mid pa_k)$ is the factor for x_k

BAYESIAN NETWORK is DAG, $\{x_1, \dots, x_n\}$ VERTICES



(x_i, x_j), $i<j$ and $x_i \in pa_j$

CONVENTION

OBSERVED NODES SHADOWED

PLATED NODES

DETERMINISTIC QUANTITIES $\bullet^{\alpha}$ solid circles

$$p(x, z) = \underbrace{p(x \mid z)}_{\mathcal{N}(x \mid \mu(z), \sigma^2)} p(z)$$

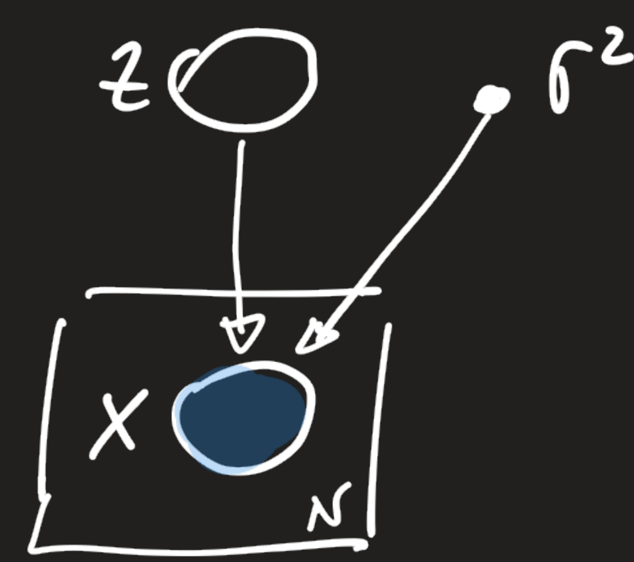
continuous $\mathcal{N}(\mu_z, \sigma^2)$

mixture component discrete $\{1 \ldots K\}$

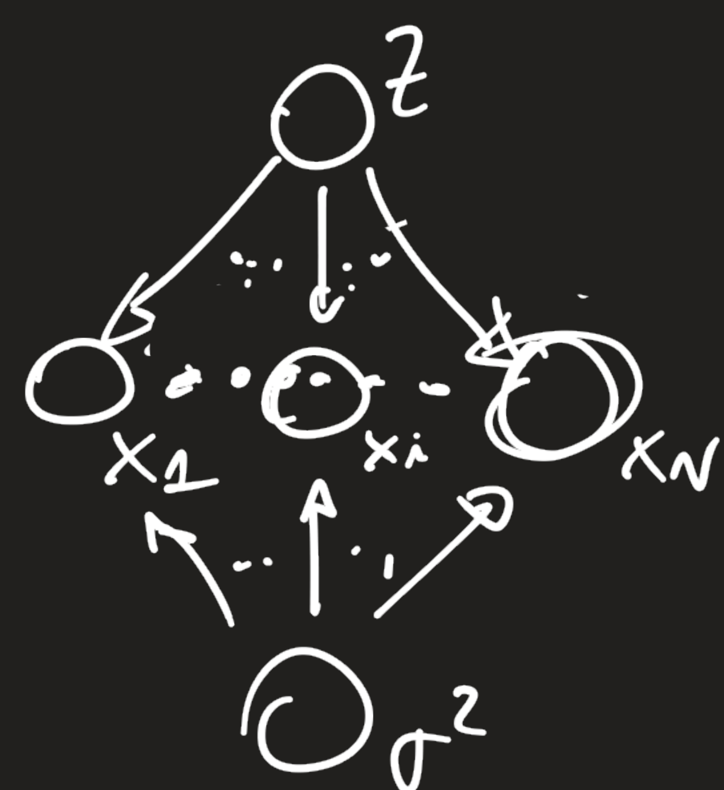
σ^2 constant

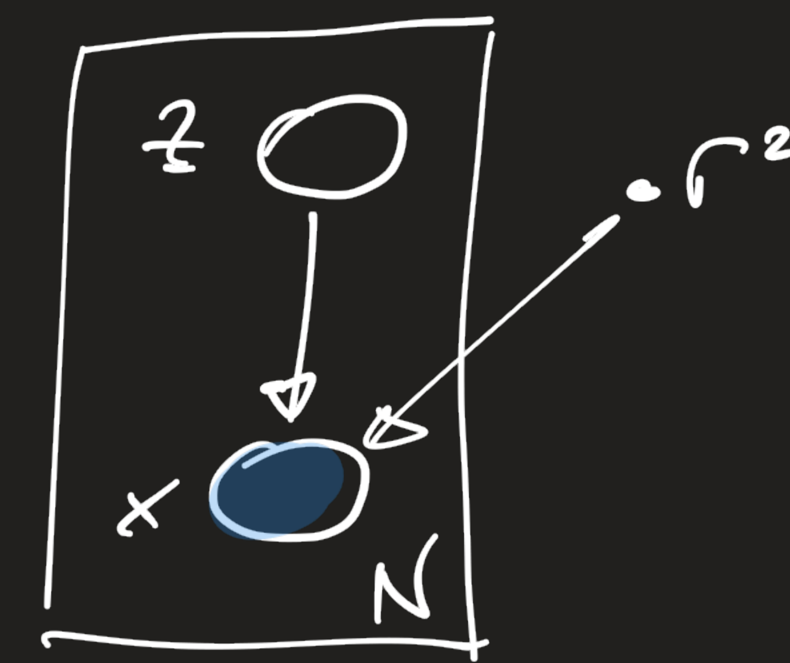
$x_1, \ldots, x_N = \underline{x}$

$$\underbrace{p(z \mid \underline{x})}_{p(z_1 \mid x_1)}$$



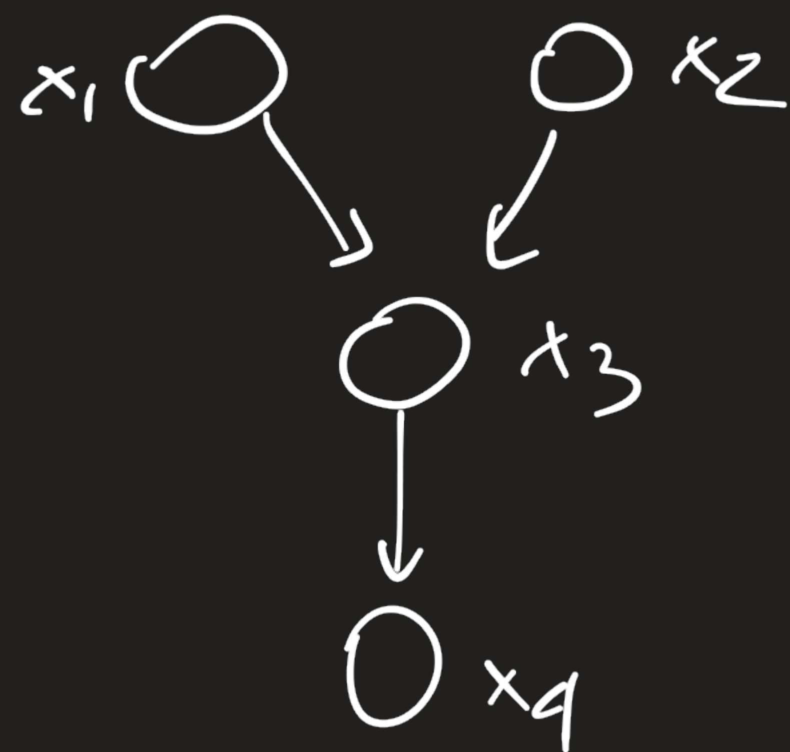
$\underline{x}$ come from the same z





ANCESTRAL SAMPLING

$$p(x_1, x_2, x_3, x_4) = p(x_4 \mid x_3)\, p(x_3 \mid x_1, x_2)\, p(x_2)\, p(x_1)$$

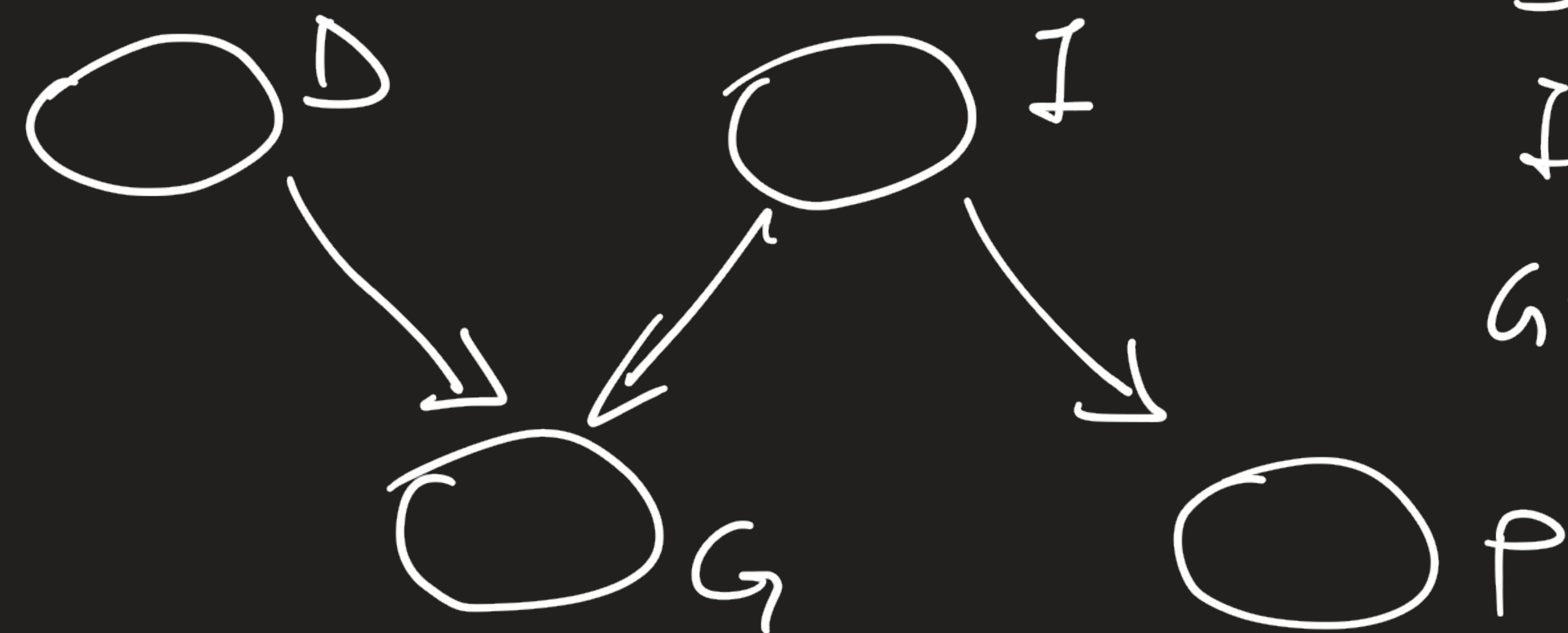


From top to bottom

from ancestrals to descendants

IF WE CAN SAMPLE FROM $\boxed{p(x \mid pa_x)}$

$P(D, I, P, G)$

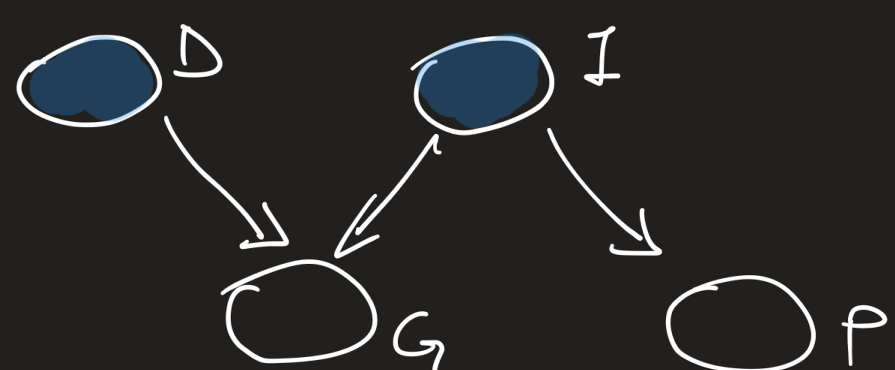


D: DIFFICULTY

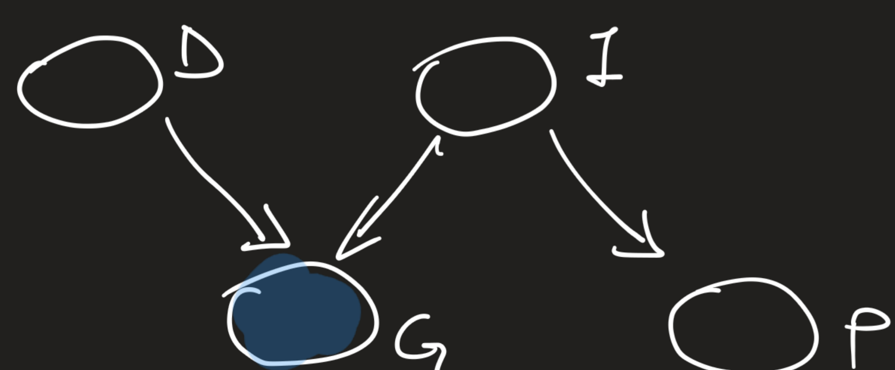
I: INTELLIGENCE

G: GRADE

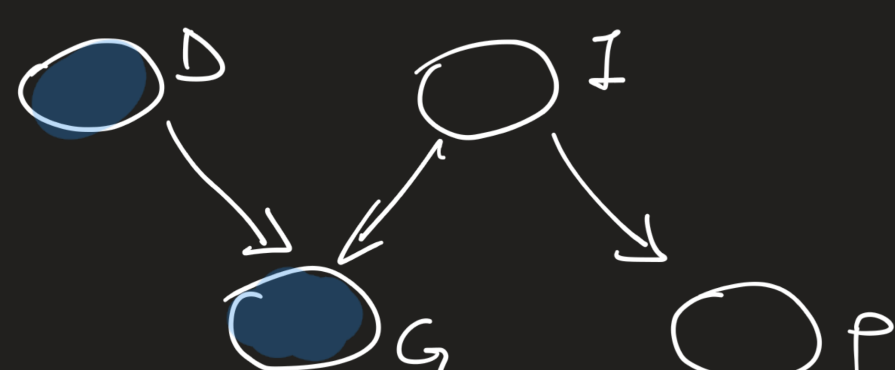
P: PASSED A HARD EXAM



↓ CAUSAL REASONING



↑ EVIDENTIAL REASONING



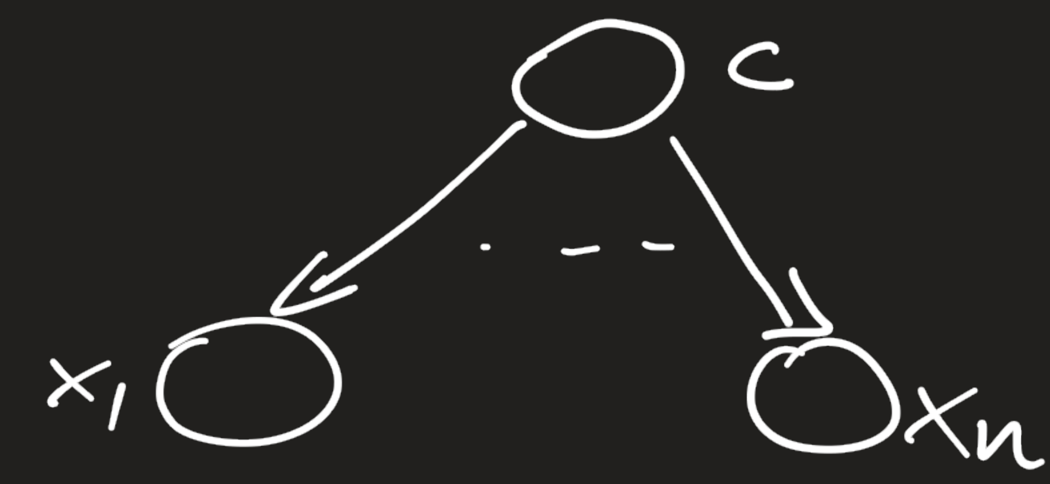
↓↑ INTERCAUSAL REASONING

NAIVE BAYES

$p(x_1, \dots, x_n \mid c)$ ~ generative model class conditional

{ FEATURES

$p(c \mid x_1, \dots, x_n)$



$$p(x_1, \dots, x_n \mid c) = \prod_{i=1}^{n} p(x_i \mid c)$$

TRAIN N.B. given c, fix x_i, consider observation only of feature x_i for points of class c, and fit a model of $\underline{p(x_i \mid c; \theta)}$

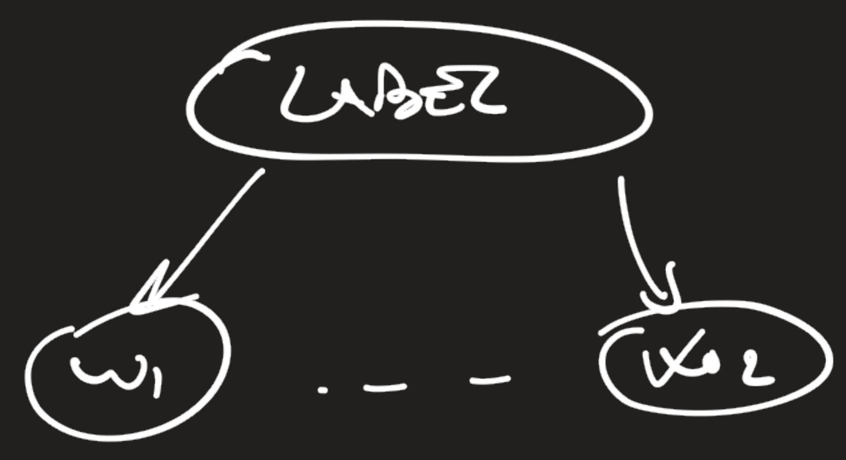
$p(c)$ estimated

$$p(c \mid x_1 \dots x_n) \propto p(c) \cdot p(x_1, \dots x_n \mid c) = p(c) \prod_i p(x_i \mid c)$$

$$\frac{p(c=0 \mid x_1 \dots x_n)}{p(c=1 \mid x_1 \dots x_n)}$$

GOOD FOR CLASSIFICATION ONLY

D_j: $\begin{cases} 0 \\ 1 & \text{if word } j \text{ is in DOC} \end{cases}$



w_i = # of occurrences of word j in doc.

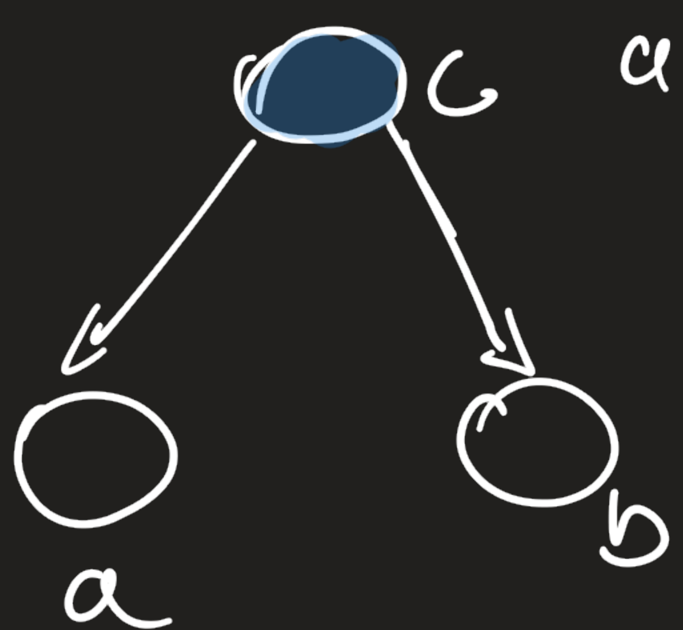
CONDITIONAL INDEPENDENCE IN BAYESIAN NETWORKS

def: a, b, c random variables.

a is cond. indep. of b given c $\quad a \perp\!\!\!\perp b \mid c$

iff

$$p(a \mid b, c) = p(a \mid c) \quad \text{OR} \quad p(a, b \mid c) = p(a \mid c)\, p(b \mid c)$$



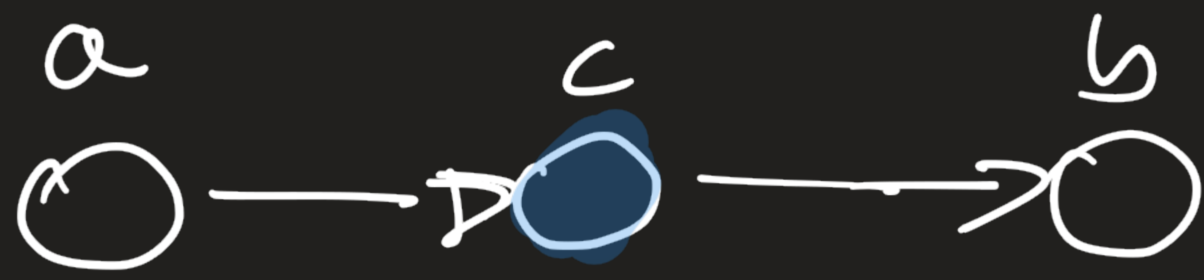
a, b TAIL TO TAIL with c

$$p(a,b) = \sum_c p(a \mid c)\, p(b \mid c)\, p(c) \neq p(a)\, p(b)$$

$$p(a,b \mid c) = \frac{p(a,b,c)}{p(c)} = \frac{p(a \mid c)\, p(b \mid c)\, \cancel{p(c)}}{\cancel{p(c)}} = p(a \mid c)\, p(b \mid c)$$

$a \not\!\perp\!\!\!\perp b$

$a \perp\!\!\!\perp b \mid c$



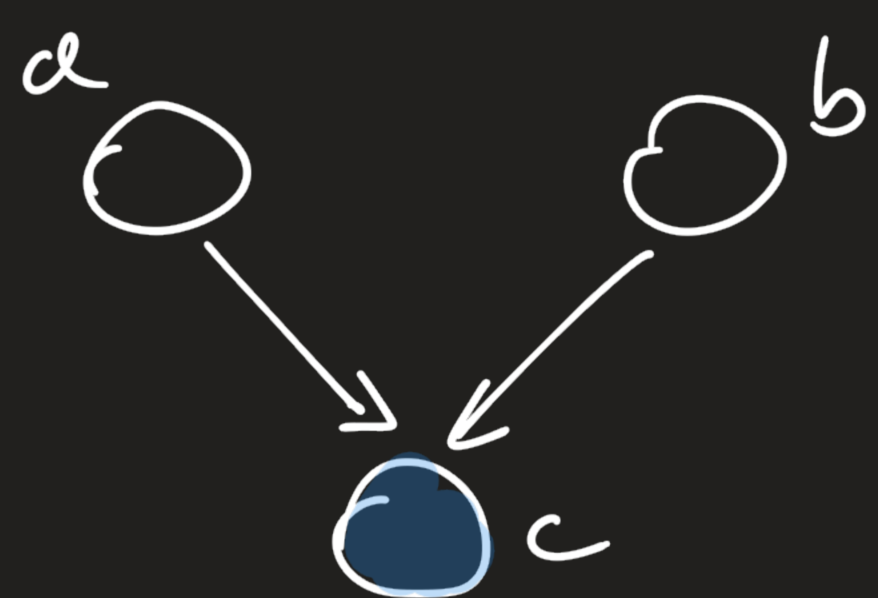
HEAD TO TAIL

$a \not\perp\!\!\!\perp b$

$$p(a,b) = \sum_c \underbrace{p(b|c)p(c|a)}\,p(a) = p(b|a)\,p(a)$$

$a \perp\!\!\!\perp b \mid c$

$$p(a,b|c) = p(b|c)\underbrace{\boxed{\frac{p(c|a)p(a)}{p(c)}}}_{p(a|c)} = p(b|c)\,p(a|c)$$



HEAD TO HEAD

$$p(a,b) = \sum_c p(a)p(b)p(c|a,b) = p(a)\,p(b)$$

$$p(a,b|c) = \frac{p(a)p(b)p(c|a,b)}{p(c)} \neq p(a|c)\,p(b|c)$$

$a \perp\!\!\!\perp b$

$a \not\perp\!\!\!\perp b \mid c \rightarrow$ also $a \not\perp\!\!\!\perp b \mid d$ - for d descendant of c

a, b, c , a path from a to b is **blocked** by c iff

- c is observed and the path is head to tail or tail to tail in c
- c is not observed, nor any descendant of c, and the path is head to head in c.

let A, B, C subsets, if all paths from a node in A to a node in B are blocked by a node in C,

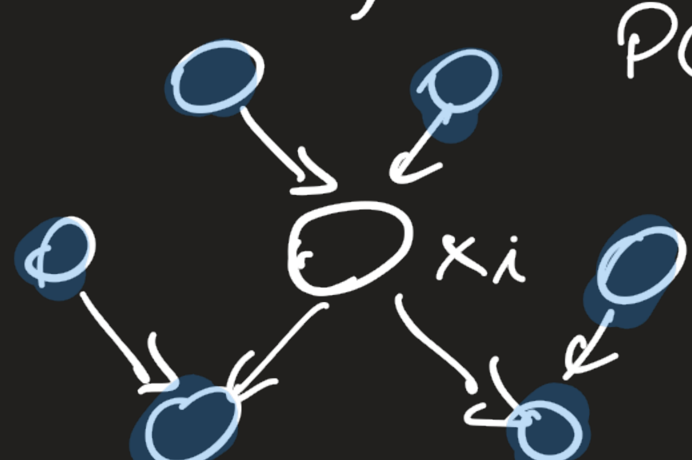
$$A \perp\!\!\!\perp B \mid C$$

MARKOV BLANKET

x_i, condition on $x_{\setminus i}$. Which nodes will remain in the conditioning set?

$$p(x_i \mid x_{\setminus i}) = \frac{p(x_1 \dots x_i \dots x_n)}{p(x_1 \dots x_{i-1} x_{i+1} \dots x_n)} = \frac{\prod_j p(x_j \mid pa_j)}{\sum_{x_i} \prod_j p(x_j \mid pa_j)}$$

only pa_i or those for which $x_i \in pa_j$



MARKOV BLANKET of x_i
- pa_i
- children
- co parents