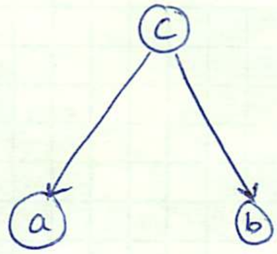
① Lecture 4 – Independence and d-separation

Cond indep



$p(a|c)p(b|c)p(c)$

① How do we rigorously show if $a \perp\!\!\!\perp b$ (if no other variables are observed?)

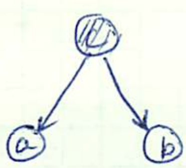
Need $p(a,b) = p(a)p(b)$. How do we get this?

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

$$= \cdots \quad \text{no way in general to factorize?}$$

$\Rightarrow a \not\perp\!\!\!\perp b \mid \phi$ ↳ Intuition knowing c can influence both a & b.

② How do things change if we have observation c?



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

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

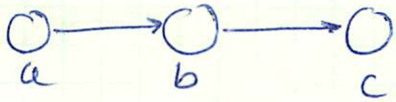
$$= \frac{p(a|c)p(b|c)p(c)}{p(c)}$$

$$= p(a|c)p(b|c) \quad \blacksquare$$

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

(2) Lecture 4 - | Indep & d-sep

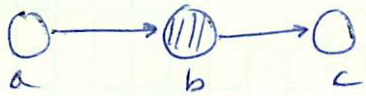
(3) A different case



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

$$p(a,c) = \sum_b p(a)\,p(b|a)\,p(c|b) \Rightarrow a \not\perp\!\!\!\perp c \mid \emptyset$$

$p(a,c|b)$



(use this in class:

)

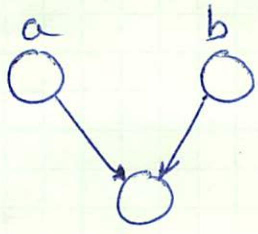
$$= \frac{p(a,b,c)}{p(b)} = \frac{\cancel{\sum_b p(a)\,p(b|a)\,p(c|b)}}{} = \frac{p(a)\,p(b|a)\,p(c|b)}{p(b)}$$

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

$$= p(a|b) \cdot p(c|b) \quad \blacksquare$$

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

(4) Finally

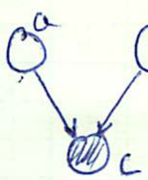


$$p(a,b,c) = p(a)\,p(b)\,p(c|a,b)$$

$$p(a,b) \overset{?}{=} p(a)\,p(b)$$

$$p(a,b) = \sum_c p(a)\,p(b)\,p(c|a,b) = p(a)\,p(b) \underbrace{\sum_c p(c|a,b)}_{=1} \quad \blacksquare$$

$\Rightarrow a \perp\!\!\!\perp b \mid \emptyset$



$a \perp\!\!\!\perp b | c$?

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

Does not factorize to $p(a|c)\,p(b|c)$

$\Rightarrow a \not\perp\!\!\!\perp b \mid c$

③ Lecture 4 – ~~[illegible]~~ Indep & d-sep

Toward
D-Separation: (directed separation)

① 

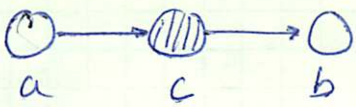
~~[illegible]~~

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

Consider path from a to b (ignoring arrows)

Node c is $\boxed{\text{"tail to tail"}}$ w.r.t. this path

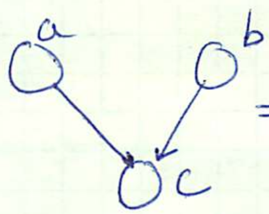
1. The presence of the path causes the nodes to be dependent
2. However, when we know c (i.e. condition on c) we "block" the path and a and b are ~~[illegible]~~ indep.

② 

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

Node c is $\boxed{\text{"head to tail"}}$

1. path a to b makes them dependent
2. given c, blocks the path, a & b are independent.

③ 

$\Rightarrow a \perp\!\!\!\perp b \mid \emptyset$

1. When c is unobserved $a \perp\!\!\!\perp b$
2. given c, $a \not\!\perp\!\!\!\perp b$

- node c is "head to head" w.r.t. a & b.
- ~~[illegible]~~ more generally: ~~[illegible]~~
  a head-to-head node unblocks a path if the node or any of its descendants is observed.

 Lecture 4 - indep & d-sep

D-separation (generalization of previous examples)

Consider directed graph: A, B, & C are arbitrary non-intersecting sets of nodes.

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

Consider: all possible paths: any node in A to any node in B

$S_i \in A$ to $S_j \in B$

Any such path is **blocked** if it includes a node s.t.: either

1) $S_k \in C$ and arrows on path meet

$\rightarrow \underset{S_k}{\bigcirc} \rightarrow$ or $\leftarrow \underset{S_k}{\bigcirc} \rightarrow$ (head-to-head or tail-to-tail)

or

2) $\rightarrow \underset{S_k}{\bigcirc} \leftarrow$ and $S_k \notin C$ and $\text{decendants}(S_k) \notin C$

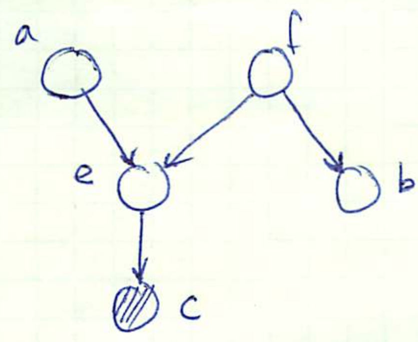
neither the node or any of its decendants are in C.

If all paths are blocked,

A is d-separated from B by C

and joint distr. over all vars satisfies

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



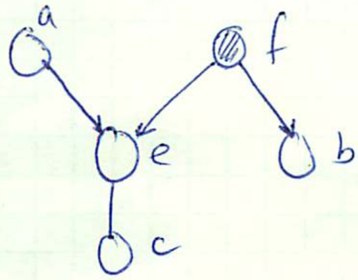
Consider path $a \rightarrow b$

f doesn't blocked $\notin C$ and it's a tail node

e doesn't block: head and $c \in C$ descendant.

$\Rightarrow a \not\perp\!\!\!\perp b \mid c$

Lecture4 Indep & d-sep



path $a \to b$ blocked by f.

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

Markov Blanket

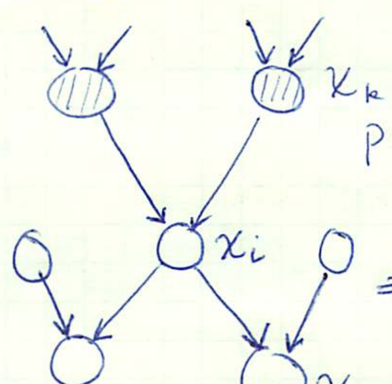
Consider
$$p(\underline{x}_i | \underline{x}_{\{j \neq i\}}) = \frac{p(x_1 \dots x_N)}{\int p(x_1 \dots x_N)\, dx_i}$$

$$= \frac{\prod_k p(x_k | pa(x_k))}{\int \prod_k p(x_k | pa(x_k))\, dx_i}$$

any factor that doesnt depend on x_i can be taken out of integral and will cancel with numerator.

What remains?

1)



$p(x_k | pa(x_k))$ does not include x_i

$\Rightarrow$ all parents of x_i cancel.

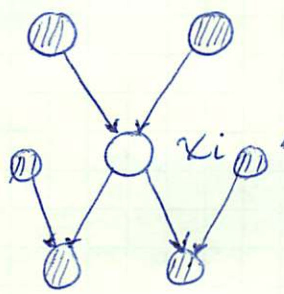
$p(x_l | pa(x_l))$ includes x_i and $pa(x_l)$

$\boxed{Q}$ What is the minimal set of nodes that isolates x_i from the rest of the graph? In other words for $p(x_i | C)$ how many nodes do we have to add to C before $p(x_i|C)$ is <u>independent</u> of the rest of the graph? $x_i \perp\!\!\!\perp B | C$ where B = every node but x_i & C.

1) need parents of x_i

2) what else? just children? $\longrightarrow$

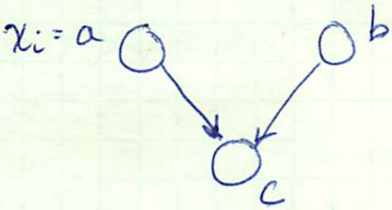
⑥ L4 - Indep & d-sep



Also need to "co-parents" of x_i.
The parents of x_i's children.

$$mb(i) = ch(i) \cup pa(i) \cup copa(i)$$

Consider this case again:

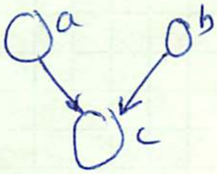


Recall $a \perp\!\!\!\perp b \mid \emptyset$

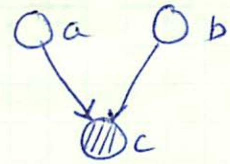
but $a \not\perp\!\!\!\perp b \mid c$

So: conditioning a c,

<u>induces</u> a dependence between a & b



$\Rightarrow a \perp\!\!\!\perp b \mid \emptyset$



$\Rightarrow a \not\perp\!\!\!\perp b \mid c \Rightarrow$

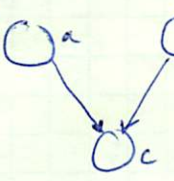
This is the phenomenon of "explaining away."

Example: a & b are two binary coins, c is the sum.

Q: Can you specify the network?

0	0.5
1	0.5

0	0.5
1	0.5



0	0.25
1	0.5
2	0.25

When we dont know sum, a & b are indep. i.e.

$$a \perp\!\!\!\perp b \mid \emptyset$$

but when we are given the sum c they are not:

$$a \not\perp\!\!\!\perp b \mid c.$$

7 L4 - Indep & d-sep

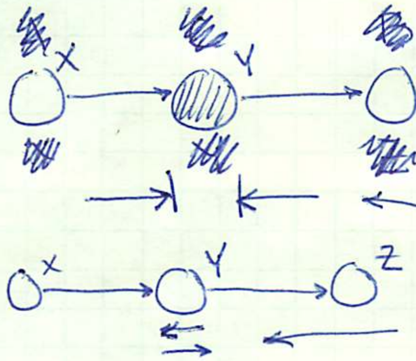
Summary of d-separation (Bayes ball algorithm)

They can bounce in both directions

place balls in A let them bounce around. If they can't reach B then

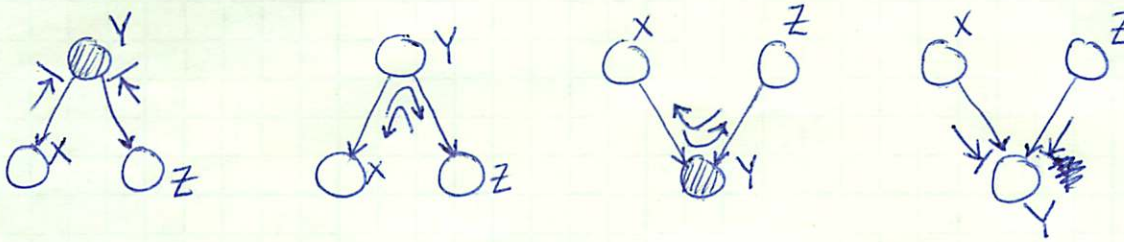
$A \perp\!\!\!\perp B \mid C$

or A is d-separated from B given C.

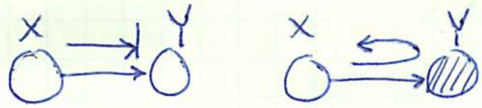


balls are blocked

balls



Boundary cases.



Why is this important?

- conditional independence relations are fundamental to what the graph represents
- understand and characterizing them plays an important role in:
  - simplifying the structure of the model
  - computations needed for inference & learning
- d-sep is a way of reading these relations directly from the model without any analytic manipulations.