

READING A BAYESIAN NETWORK

1) Conditional independence is ^{such} a fundamental concept because it tells you when a piece of knowledge is relevant.

For instance, knowing Rhonda's blood type is normally irrelevant to knowing Sam's (they are not related by blood), i.e. $P(s|r) = P(s)$.

But if I know their son Tim's blood type is, say, AB, then knowing Rhonda's blood type is suddenly relevant to Sam's (if Rhonda's is A, then Sam's cannot be A), i.e. $P(s|r, t) \neq P(s|t)$.

We represent the conditional independence of two variables X and Y given a set of variables Z as $X \perp\!\!\!\perp Y | Z$.

2) Conditional independence is tough to "see" in a distribution. In which of these is $A \perp\!\!\!\perp B$?

A	B	C	P_1	P_2
0	0	0	$1/32$	$1/32$
0	0	1	$3/32$	$3/32$
0	1	0	$6/32$	$6/32$
0	1	1	$6/32$	$6/32$
1	0	0	$3/32$	$3/32$
1	0	1	$1/32$	$3/32$
1	1	0	$4/32$	$2/32$
1	1	1	$8/32$	$8/32$

READING A BAYESIAN NETWORK

$$③ P_1(A=1) = \frac{3+1+4+8}{32} = \frac{1}{2}$$

$$P_1(A=1|B=0) = \frac{P_1(A=1, B=0)}{P_1(B=0)} = \frac{\frac{3+1}{32}}{\frac{1+3+3+1}{32}} = \frac{1}{2}$$

$$P_1(A=1|B=1) = \frac{P_1(A=1, B=1)}{P_1(B=1)} = \frac{\frac{4+8}{32}}{\frac{6+6+4+8}{32}} = \frac{1}{2}$$

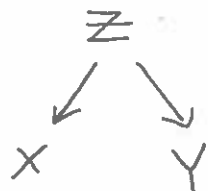
So $A \perp\!\!\!\perp B$ in P_1 .

$$P_2(A=1) = \frac{3+3+2+8}{32} = \frac{1}{2}$$

$$P_2(A=1|B=0) = \frac{P_2(A=1, B=0)}{P_2(B=0)} = \frac{\frac{3+3}{32}}{\frac{1+3+3+3}{32}} = \frac{3}{5}$$

So $A \not\perp\!\!\!\perp B$ in P_2 .

④ With a Bayesian network, conditional independence is much easier to "see". Let's first consider the ways in which a variable Z can link two other variables:



"fork"



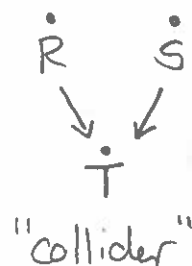
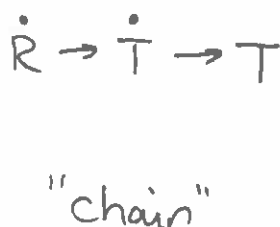
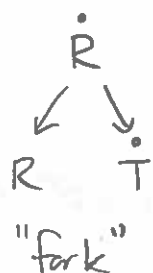
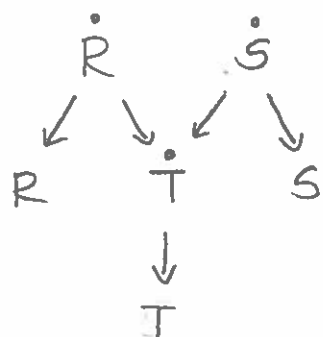
"chain"



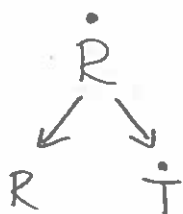
"collider"

READING A BAYESIAN NETWORK

5) Considering our blood type network, we can see examples of all of these:



6) What are the conditional independence relationships implied by these structures? Consider the fork:

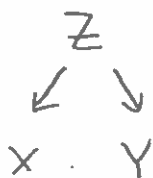


Intuitively, knowing Rhonda's blood type is relevant to knowing Tim's genotype (she's his mom). In other words, $R \not\perp T$.

However, if we know Rhonda's genotype already, then knowing Rhonda's blood type is now irrelevant to our opinion about Tim's genotype (it's superfluous). In other words, $R \perp T | R$.

READING A BAYESIAN NETWORK

7) We can establish this mathematically. Given a fork:



$$\begin{aligned} P(x|y,z) &= \frac{P(x,y,z)}{P(y,z)} = \frac{P(x,y,z)}{\sum_{x'} P(x',y,z)} \\ &= \frac{P(z)P(x|z)P(y|z)}{\sum_{x'} P(z)P(x'|z)P(y|z)} \quad \left[\text{by def'n of Bayes Net} \right] \\ &= \frac{P(z)P(x|z)P(y|z)}{P(z)P(y|z) \sum_{x'} P(x'|z)} \\ &= \frac{P(x|z)}{\sum_{x'} P(x'|z)} \\ &= P(x|z) \quad \left[\text{b/c } \sum_{x'} P(x'|z) = 1 \right] \end{aligned}$$

So $X \perp\!\!\!\perp Y | Z$.

8) Next, consider the chain:

$$\dot{R} \rightarrow \dot{T} \rightarrow T$$

Intuitively, knowing Rhonda's genotype is relevant to knowing Tim's blood type (she's his mom). In other words, $\dot{R} \not\perp\!\!\!\perp T$.

However, if we know Tim's genotype already, then information about Rhonda is now irrelevant to our opinion about Tim's blood type. In other words, $\dot{R} \perp\!\!\!\perp T | \dot{T}$.

READING A BAYESIAN NETWORK

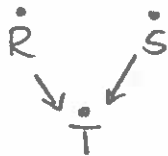
9) We can also prove this. Given a chain: $X \rightarrow Z \rightarrow Y$

$$\begin{aligned} P(x|y, z) &= \frac{P(x, y, z)}{\sum_{x'} P(x', y, z)} \\ &= \frac{P(x)P(z|x)P(y|z)}{\sum_{x'} P(x')P(z|x')P(y|z)} \\ &= \frac{P(x)P(z|x)P(y|z)}{P(y|z) \sum_{x'} P(x')P(z|x')} \\ &= \frac{P(x)P(z|x)}{P(z)} \\ &= \frac{P(x)}{P(z)} \cdot \frac{P(x|z)P(z)}{P(x)} \\ &= P(x|z) \end{aligned}$$

[by def'n of
Bayes Net]

So $X \perp\!\!\!\perp Y | Z$.

10) Colliders are a bit different:



Intuitively, knowing Rhonda's genotype is irrelevant to knowing Sam's genotype (they aren't blood relatives). In other words, $R \perp\!\!\!\perp S$.

However, if we know Tim's genotype, then information about Rhonda's can now be relevant to Sam's genotype (if Tim is AB, then knowing Rhonda is AO means Sam must have a B allele). $R \not\perp\!\!\!\perp S | T$.

READING A BAYESIAN NETWORK

① Given a collider $X \swarrow \searrow Y$
 $\quad \quad \quad \downarrow$
 $\quad \quad \quad Z$, we can show by example

that $X \not\perp Y | Z$ is possible:

X	Y	Z	
0	0	0	$\frac{1}{4}$
0	0	1	0
0	1	0	0
0	1	1	$\frac{1}{4}$
1	0	0	0
1	0	1	$\frac{1}{4}$
1	1	0	0
1	1	1	$\frac{1}{4}$

$$P(X=1|Z=1) = \frac{P(X=1, Z=1)}{P(Z=1)} = \frac{\frac{1}{4} + \frac{1}{4}}{\frac{1}{4} + \frac{1}{4} + \frac{1}{4}} = \frac{2}{3}$$

$$P(X=1|Y=0, Z=1) = \frac{P(X=1, Y=0, Z=1)}{P(Y=0, Z=1)} = \frac{\frac{1}{4}}{\frac{1}{4}} = 1$$

However:

$$P(X|Y) = \sum_z P(X, z|Y) = \frac{\sum_z P(X, Y, z)}{P(Y)}$$

$$= \frac{\sum_z P(X)P(Y)P(z|X, Y)}{P(Y)}$$

[from def'n
of Bayes Net]

$$= \frac{P(X)P(Y) \sum_z P(z|X, Y)}{P(Y)}$$

$$= P(X) \sum_z P(z|X, Y)$$

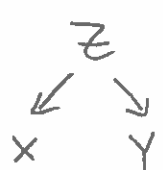
$$= P(X)$$

$$\left[\text{b/c } \sum_z P(z|X, Y) = 1 \right]$$


So $X \perp Y$.

READING A BAYESIAN NETWORK

⑫ In summary:

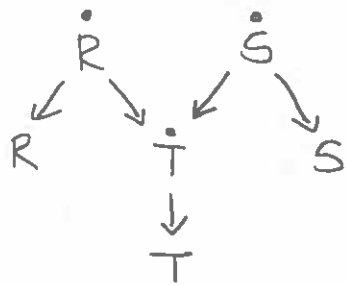
For a fork  or a chain $X \rightarrow Z \rightarrow Y$;

$X \not\perp\!\!\!\perp Y$ and $X \perp\!\!\!\perp Y | Z$

For a collider :

$X \perp\!\!\!\perp Y$ and $X \not\perp\!\!\!\perp Y | Z$

⑬ We can use these basic structures to determine the flow of relevance in a larger network:

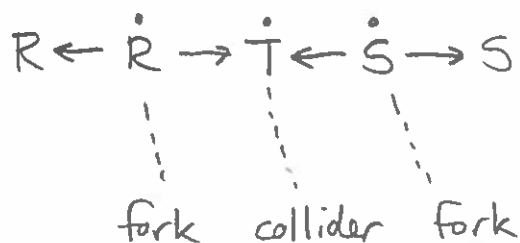


To determine the flow of relevance between R and S , we examine each path between them. There is only one in this case:

$R \leftarrow \dot{R} \rightarrow \dot{T} \leftarrow \dot{S} \rightarrow S$

READING A BAYESIAN NETWORK

- 14) Each intermediate node in the path is the center of a fork, chain, or collider:



- 15) Given a set Z of nodes, the path:

$$X \longleftrightarrow W_1 \longleftrightarrow \dots \longleftrightarrow W_k \longleftrightarrow Y$$

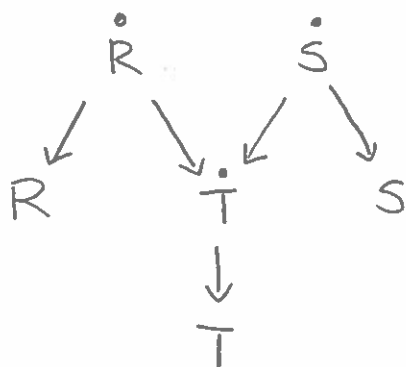
is blocked by Z if there exists W_i s.t.

- $W_i \in Z$ and W_i is the center of a fork or chain
- $W_i \notin Z$, no descendant of $W_i \in Z$, and W_i is the center of a collider.

If every path between X and Y is blocked by Z , we say that X and Y are d-separated by Z , and we write $X \perp Y \mid Z$.

READING A BAYESIAN NETWORK

16



$$R \perp S \mid \emptyset$$

(the collider at \dot{T} blocks the path)

$$R \not\perp S \mid \{\dot{T}\}$$

(T opens the collider at \dot{T})

$$R \perp S \mid \{\dot{R}, \dot{T}\}$$

(the fork at \dot{R} blocks the path)

$$R \not\perp T \mid \emptyset$$

(there is an unblocked path)

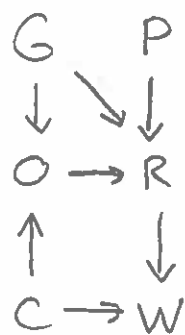
$$R \perp T \mid \{\dot{T}\}$$

(the chain at \dot{T} blocks the path)

17 Theorem: IF P is any distribution that factors according to Bayesian network G , then if $X \perp Y \mid Z$ in G , then $X \perp Y \mid Z$ in P .

READING A BAYESIAN NETWORK

18 Practice:



$G \perp P \mid \emptyset ?$

Yes.

$G \perp P \mid \{W\} ?$

No.

$G \perp W \mid \emptyset ?$

No.

$G \perp W \mid \{O\} ?$

No.

$G \perp W \mid \{R\} ?$

Yes.

$C \perp P \mid \emptyset ?$

Yes.

$C \not\perp P \mid \{W\} ?$

No.

$C \not\perp P \mid \{W, R\} ?$

No.