CS188: Artificial Intelligence.

Lecture Attendance Policy is suspended.

CS188: Artificial Intelligence.

Lecture Attendance Policy is suspended.
  Working out full ramifications.

CS188: Artificial Intelligence.

Lecture Attendance Policy is suspended.
  Working out full ramifications.

Written Homework 2 is out.

CS188: Artificial Intelligence.

Lecture Attendance Policy is suspended.
  Working out full ramifications.

Written Homework 2 is out.
  Get started!

Artificial Intelligence: Bayes' Nets

Representation. ✓

Artificial Intelligence: Bayes' Nets

Representation. ✓

Conditional Independences (Next.)

Artificial Intelligence: Bayes' Nets

Representation. ✓

Conditional Independences (Next.)

Probabilistic Inference

Artificial Intelligence: Bayes’ Nets

Representation. ✓

Conditional Independences (Next.)

Probabilistic Inference

Learning Bayes’ Nets from Data

Conditional Independence

X and Y are independent if

Conditional Independence

X and Y are independent if

$$\forall x, y : P(x, y) = P(x)P(y) \rightarrow X \perp\!\!\!\perp Y$$

Conditional Independence

X and Y are independent if

$$\forall x, y : P(x, y) = P(x)P(y) \rightarrow X \perp\!\!\!\perp Y$$

X and Y are conditionally independent given Z

Conditional Independence

X and Y are independent if

$$\forall x, y : P(x, y) = P(x)P(y) \rightarrow X \perp\!\!\!\perp Y$$

X and Y are conditionally independent given Z

$$\forall x, y, z : P(x, y|z) = P(x|z)P(y|z) \rightarrow X \perp\!\!\!\perp Y|Z$$

Conditional Independence

X and Y are independent if

$$\forall x, y : P(x, y) = P(x)P(y) \rightarrow X \perp\!\!\!\perp Y$$

X and Y are conditionally independent given Z

$$\forall x, y, z : P(x, y|z) = P(x|z)P(y|z) \rightarrow X \perp\!\!\!\perp Y|Z$$

Conditional Independence

X and Y are independent if

$$\forall x, y : P(x, y) = P(x)P(y) \rightarrow X \perp\!\!\!\perp Y$$

X and Y are conditionally independent given Z

$$\forall x, y, z : P(x, y|z) = P(x|z)P(y|z) \rightarrow X \perp\!\!\!\perp Y|Z$$

(Conditional) independence is a property of a distribution

Conditional Independence

X and Y are independent if

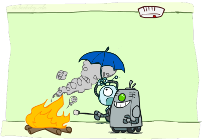
$$\forall x, y : P(x, y) = P(x)P(y) \rightarrow X \perp\!\!\!\perp Y$$

X and Y are conditionally independent given Z

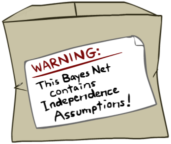
$$\forall x, y, z : P(x, y|z) = P(x|z)P(y|z) \rightarrow X \perp\!\!\!\perp Y|Z$$

(Conditional) independence is a property of a distribution

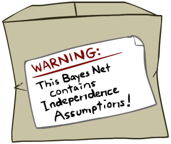
Example: Alarm $\perp\!\!\!\perp$ Fire $|$ Smoke



Bayes Nets: Assumptions

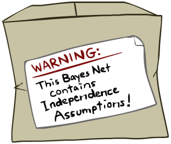


Bayes Nets: Assumptions



Assumptions we are required to make to define the Bayes net when given the graph:

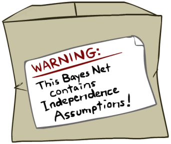
Bayes Nets: Assumptions



Assumptions we are required to make to define the Bayes net when given the graph:

$$P(x_i|x_1, x_2, \ldots, x_{i-1}) = P(x_i|parents(X_i))$$

Bayes Nets: Assumptions

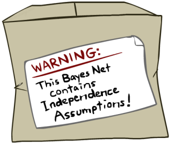


Assumptions we are required to make to define the Bayes net when given the graph:

$$P(x_i|x_1, x_2, \ldots, x_{i-1}) = P(x_i|parents(X_i))$$

Beyond above "chain rule for Bayes net" conditional independence assumptions

Bayes Nets: Assumptions



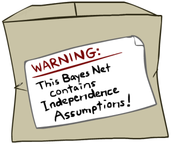
Assumptions we are required to make to define the Bayes net when given the graph:

$$P(x_i|x_1, x_2, \ldots, x_{i-1}) = P(x_i|parents(X_i))$$

Beyond above "chain rule for Bayes net" conditional independence assumptions

- Often additional conditional independences

Bayes Nets: Assumptions



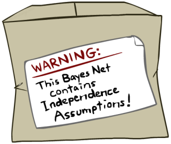
Assumptions we are required to make to define the Bayes net when given the graph:

$$P(x_i | x_1, x_2, \ldots, x_{i-1}) = P(x_i | parents(X_i))$$

Beyond above "chain rule for Bayes net" conditional independence assumptions

- Often additional conditional independences
- They can be read off the graph

Bayes Nets: Assumptions



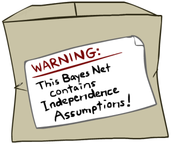
Assumptions we are required to make to define the Bayes net when given the graph:

$$P(x_i|x_1, x_2, \ldots, x_{i-1}) = P(x_i|parents(X_i))$$

Beyond above "chain rule for Bayes net" conditional independence assumptions

- Often additional conditional independences
- They can be read off the graph

Bayes Nets: Assumptions



Assumptions we are required to make to define the Bayes net when given the graph:

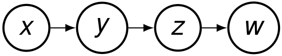
$$P(x_i|x_1, x_2, \ldots, x_{i-1}) = P(x_i|parents(X_i))$$

Beyond above "chain rule for Bayes net" conditional independence assumptions

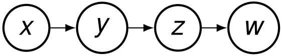
- Often additional conditional independences
- They can be read off the graph

Important for modeling: understand assumptions made when choosing a Bayes net graph

Example



Example



Conditional independence assumptions directly from simplifications in chain rule:

Example

$$x \rightarrow y \rightarrow z \rightarrow w$$

Conditional independence assumptions directly from simplifications in chain rule:

Additional implied conditional independence assumptions?

Independence in a BN

Independence in a BN

Important question about a BN:

Independence in a BN

Important question about a BN:

- Are two nodes independent given certain evidence?

Independence in a BN

Important question about a BN:

- Are two nodes independent given certain evidence?
- If yes, can prove using algebra (tedious in general)

Independence in a BN

Important question about a BN:

- Are two nodes independent given certain evidence?
- If yes, can prove using algebra (tedious in general)
- If no, can prove with a counter example

Independence in a BN

Important question about a BN:

- Are two nodes independent given certain evidence?
- If yes, can prove using algebra (tedious in general)
- If no, can prove with a counter example
- Example:

Independence in a BN

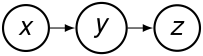
Important question about a BN:

- Are two nodes independent given certain evidence?
- If yes, can prove using algebra (tedious in general)
- If no, can prove with a counter example
- Example:

Independence in a BN

Important question about a BN:

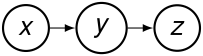
- Are two nodes independent given certain evidence?
- If yes, can prove using algebra (tedious in general)
- If no, can prove with a counter example
- Example:



Independence in a BN

Important question about a BN:

- Are two nodes independent given certain evidence?
- If yes, can prove using algebra (tedious in general)
- If no, can prove with a counter example
- Example:

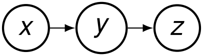


- Question: are X and Z necessarily independent?

Independence in a BN

Important question about a BN:

- Are two nodes independent given certain evidence?
- If yes, can prove using algebra (tedious in general)
- If no, can prove with a counter example
- Example:



- Question: are X and Z necessarily independent?
- Answer: no.

Independence in a BN

Important question about a BN:

- Are two nodes independent given certain evidence?
- If yes, can prove using algebra (tedious in general)
- If no, can prove with a counter example
- Example:

$$x \rightarrow y \rightarrow z$$

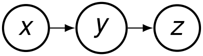
- Question: are X and Z necessarily independent?
- Answer: no.

  Example:

Independence in a BN

Important question about a BN:

- Are two nodes independent given certain evidence?
- If yes, can prove using algebra (tedious in general)
- If no, can prove with a counter example
- Example:

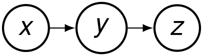


- Question: are X and Z necessarily independent?
- Answer: no.
  Example: low pressure causes rain, which causes traffic.

Independence in a BN

Important question about a BN:

- Are two nodes independent given certain evidence?
- If yes, can prove using algebra (tedious in general)
- If no, can prove with a counter example
- Example:

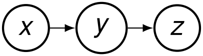


- Question: are X and Z necessarily independent?
- Answer: no.
  Example: low pressure causes rain, which causes traffic.
- X can influence Z, Z can influence X (via Y)

Independence in a BN

Important question about a BN:

- Are two nodes independent given certain evidence?
- If yes, can prove using algebra (tedious in general)
- If no, can prove with a counter example
- Example:



- Question: are X and Z necessarily independent?
- Answer: no.
  Example: low pressure causes rain, which causes traffic.
- X can influence Z, Z can influence X (via Y)
- Addendum: they could be independent: how?

Independence in a BN

Important question about a BN:

- Are two nodes independent given certain evidence?
- If yes, can prove using algebra (tedious in general)
- If no, can prove with a counter example
- Example:

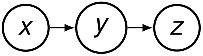
$$x \rightarrow y \rightarrow z$$

- Question: are X and Z necessarily independent?
- Answer: no.
  Example: low pressure causes rain, which causes traffic.
- X can influence Z, Z can influence X (via Y)
- Addendum: they could be independent: how?

Independence in a BN

Important question about a BN:

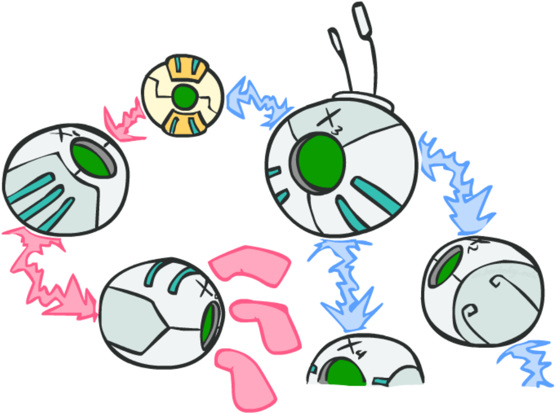
- Are two nodes independent given certain evidence?
- If yes, can prove using algebra (tedious in general)
- If no, can prove with a counter example
- Example:



- Question: are X and Z necessarily independent?
- Answer: no.
  Example: low pressure causes rain, which causes traffic.
- X can influence Z, Z can influence X (via Y)
- Addendum: they could be independent: how?

Same as Markov chain.

D-separation: Outline



D-separation: Outline

D-separation: Outline

Study independence properties for triples

D-separation: Outline

Study independence properties for triples

Analyze complex cases in terms of member triples

D-separation: Outline

Study independence properties for triples

Analyze complex cases in terms of member triples

D-separation: a condition / algorithm for answering such queries

Causal Chains

This configuration is a "causal chain"

Causal Chains

This configuration is a "causal chain"

X: Low pressure Y: Rain Z: Traffic

Causal Chains

This configuration is a "causal chain"

X: Low pressure Y: Rain Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

Causal Chains

This configuration is a "causal chain"

X: Low pressure Y: Rain Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

Causal Chains

This configuration is a "causal chain"

X: Low pressure   Y: Rain   Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z ? No!

Causal Chains

This configuration is a "causal chain"

X: Low pressure Y: Rain Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z ? No!

- One example set of CPTs for which X is not independent of Z is sufficient to show this independence is not guaranteed.

Causal Chains

This configuration is a "causal chain"

X: Low pressure Y: Rain Z: Traffic

$$P(x,y,z) = P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z ? No!

- One example set of CPTs for which X is not independent of Z is sufficient to show this independence is not guaranteed.
- Example:

Causal Chains

This configuration is a "causal chain"

X: Low pressure Y: Rain Z: Traffic

$$P(x,y,z) = P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z ? No!

- One example set of CPTs for which X is not independent of Z is sufficient to show this independence is not guaranteed.
- Example:

Causal Chains

This configuration is a "causal chain"

X: Low pressure Y: Rain Z: Traffic

$$P(x,y,z) = P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z ? No!

- One example set of CPTs for which X is not independent of Z is sufficient to show this independence is not guaranteed.
- Example:

Low pressure $\rightarrow$ rain $\rightarrow$ traffic,

Causal Chains

This configuration is a "causal chain"

X: Low pressure Y: Rain Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z ? No!

- One example set of CPTs for which X is not independent of Z is sufficient to show this independence is not guaranteed.
- Example:

  Low pressure $\rightarrow$ rain $\rightarrow$ traffic,

  High pressure $\rightarrow$ no rain $\rightarrow$ no traffic

Causal Chains

This configuration is a "causal chain"

X: Low pressure Y: Rain Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z ? No!

- One example set of CPTs for which X is not independent of Z is sufficient to show this independence is not guaranteed.
- Example:

  Low pressure $\rightarrow$ rain $\rightarrow$ traffic,

  High pressure $\rightarrow$ no rain $\rightarrow$ no traffic
- In numbers:

Causal Chains

This configuration is a "causal chain"

X: Low pressure Y: Rain Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z ? No!

- One example set of CPTs for which X is not independent of Z is sufficient to show this independence is not guaranteed.
- Example:

  Low pressure $\rightarrow$ rain $\rightarrow$ traffic,

  High pressure $\rightarrow$ no rain $\rightarrow$ no traffic

- In numbers:
- P(+y | +x) = 1, P(-y | - x) = 1,

Causal Chains

This configuration is a "causal chain"

X: Low pressure Y: Rain Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z ? No!

- One example set of CPTs for which X is not independent of Z is sufficient to show this independence is not guaranteed.
- Example:

Low pressure $\rightarrow$ rain $\rightarrow$ traffic,

High pressure $\rightarrow$ no rain $\rightarrow$ no traffic

- In numbers:
- P(+y | +x) = 1, P(-y | - x) = 1,
- P(+z | +y) = 1, P(-z | -y) = 1

Causal Chains

This configuration is a "causal chain"

X: Low pressure Y: Rain Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z ? No!

- One example set of CPTs for which X is not independent of Z is sufficient to show this independence is not guaranteed.
- Example:

  Low pressure $\rightarrow$ rain $\rightarrow$ traffic,

  High pressure $\rightarrow$ no rain $\rightarrow$ no traffic

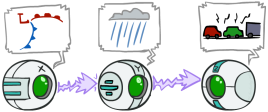
- In numbers:
- P(+y | +x) = 1, P(-y | - x) = 1,
- P(+z | +y) = 1, P(-z | -y) = 1

Casual Chain

This configuration is a “causal chain”

Casual Chain

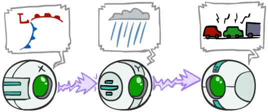
This configuration is a “causal chain”



X: Low pressure Y: Rain Z: Traffic

Casual Chain

This configuration is a “causal chain”

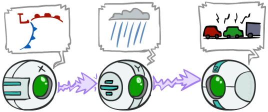


X: Low pressure Y: Rain Z: Traffic

$$P(x,y,z) = P(x)P(y|x)P(z|y)$$

Casual Chain

This configuration is a "causal chain"

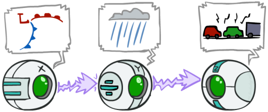


X: Low pressure Y: Rain Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

Casual Chain

This configuration is a “causal chain”



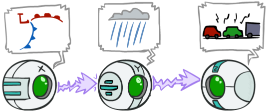
X: Low pressure Y: Rain Z: Traffic

$$P(x, y, z) =$$
$$P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z given Y?

Casual Chain

This configuration is a “causal chain”



X: Low pressure Y: Rain Z: Traffic

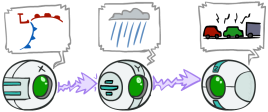
$$P(x, y, z) =$$
$$P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z given Y?

$$P(z|y, x) = \frac{P(x,y,z)}{P(x,y)}$$

Casual Chain

This configuration is a “causal chain”



X: Low pressure   Y: Rain   Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z given Y?

$$P(z|y, x) = \frac{P(x,y,z)}{P(x,y)} = \frac{P(x)P(y|x)P(z|y)}{P(x)P(y|x)}$$

Casual Chain

This configuration is a “causal chain”

X: Low pressure Y: Rain Z: Traffic

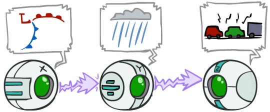
$$P(x,y,z) = P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z given Y?

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

Casual Chain

This configuration is a “causal chain”



X: Low pressure Y: Rain Z: Traffic

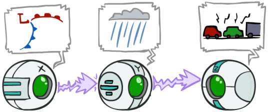
$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z given Y?

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

Casual Chain

This configuration is a “causal chain”



X: Low pressure Y: Rain Z: Traffic

$$P(x, y, z) =$$
$$P(x)P(y|x)P(z|y)$$

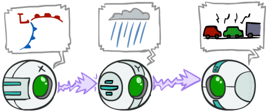
Guaranteed X independent of Z given Y?

$$P(z|y, x) = \frac{P(x,y,z)}{P(x,y)}$$
$$= \frac{P(x)P(y|x)P(z|y)}{P(x)P(y|x)}$$
$$= P(z|y)$$

Yes!!!!!

Casual Chain

This configuration is a “causal chain”



X: Low pressure Y: Rain Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z given Y?

$$P(z|y,x) = \frac{P(x,y,z)}{P(x,y)} = \frac{P(x)P(y|x)P(z|y)}{P(x)P(y|x)} = P(z|y)$$

Yes!!!!!

- Evidence along the chain “blocks” the influence

Casual Chain

This configuration is a “causal chain”

X: Low pressure Y: Rain Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

Guaranteed X independent of Z given Y?

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

Yes!!!!!

- Evidence along the chain “blocks” the influence

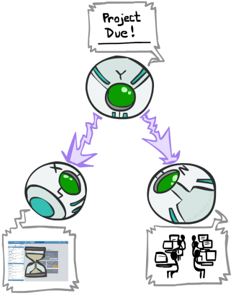
Common Cause

This configuration is a “common cause”

Common Cause

This configuration is a "common cause"

Y: Project due

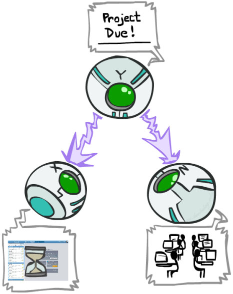


X: Forums busy   Z: Lab full

Common Cause

This configuration is a "common cause"

Y: Project due



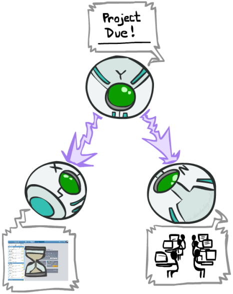
X: Forums busy   Z: Lab full

Guaranteed X independent of Z ? No!

Common Cause

This configuration is a “common cause”

Y: Project due



X: Forums busy   Z: Lab full

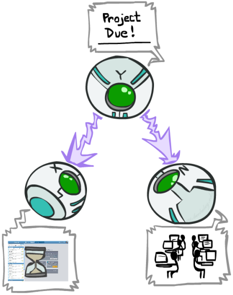
Guaranteed X independent of Z ? No!

- One example set of CPTs for which X is not independent of Z is sufficient to show this independence is not guaranteed.

Common Cause

This configuration is a “common cause”

Y: Project due



X: Forums busy Z: Lab full

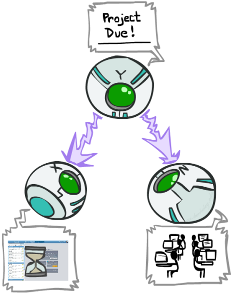
Guaranteed X independent of Z ? No!

- One example set of CPTs for which X is not independent of Z is sufficient to show this independence is not guaranteed.
- Example:

Common Cause

This configuration is a “common cause”

Y: Project due



X: Forums busy   Z: Lab full

Guaranteed X independent of Z ? No!

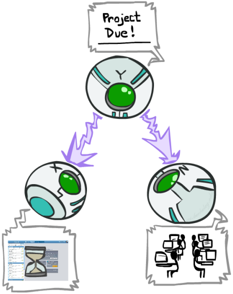
- One example set of CPTs for which X is not independent of Z is sufficient to show this independence is not guaranteed.
- Example:

  Project due $\leftrightarrow$ both forums busy and lab full
- In numbers:

Common Cause

This configuration is a "common cause"

Y: Project due



X: Forums busy   Z: Lab full

Guaranteed X independent of Z ? No!

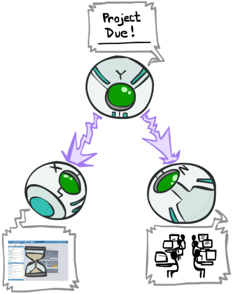
- One example set of CPTs for which X is not independent of Z is sufficient to show this independence is not guaranteed.
- Example:
  Project due $\leftrightarrow$ both forums busy and lab full
- In numbers:

$$P(+x \mid +y) = 1,\ P(-x \mid -y) = 1,$$

Common Cause

This configuration is a "common cause"

Y: Project due



X: Forums busy   Z: Lab full

Guaranteed X independent of Z ? No!

- One example set of CPTs for which X is not independent of Z is sufficient to show this independence is not guaranteed.
- Example:
  Project due $\leftrightarrow$ both forums busy and lab full
- In numbers:

$$P(+x \mid +y) = 1,\ P(-x \mid -y) = 1,$$

$$P(+z \mid +y) = 1,\ P(-z \mid -y) = 1$$

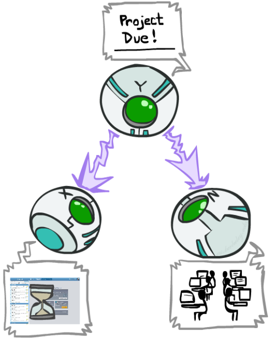
Common Cause

This configuration is a “common cause”

Common Cause

This configuration is a “common cause”

Y: Project due

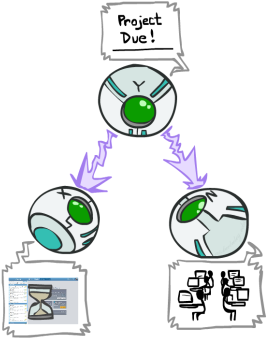


X: Forums busy Z: Lab full

Common Cause

This configuration is a “common cause”

Y: Project due



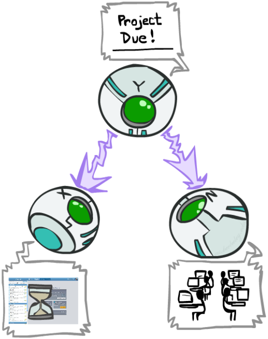
X: Forums busy

Z: Lab full

Common Cause

This configuration is a "common cause"

Y: Project due



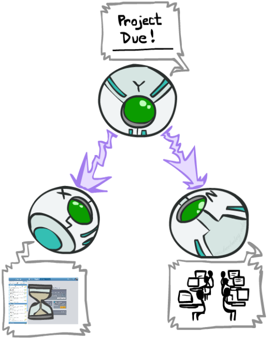
X: Forums busy Z: Lab full

Guaranteed X and Z independent given Y?

Common Cause

This configuration is a "common cause"

Y: Project due



X: Forums busy Z: Lab full

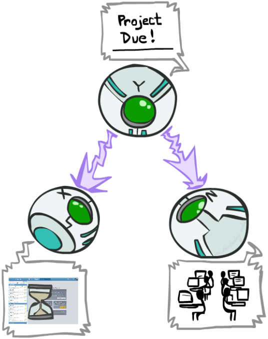
Guaranteed X and Z independent given Y?

$$P(z|x,y) = \frac{P(x,y,z)}{P(x,y)}$$

Common Cause

This configuration is a “common cause”

Y: Project due



X: Forums busy Z: Lab full

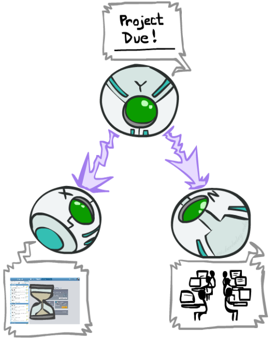
Guaranteed X and Z independent given Y?

$$P(z|x,y) = \frac{P(x,y,z)}{P(x,y)}$$
$$= \frac{P(y)P(x|y)P(z|y)}{P(y)P(x|y)}$$

Common Cause

This configuration is a "common cause"

Y: Project due



X: Forums busy Z: Lab full

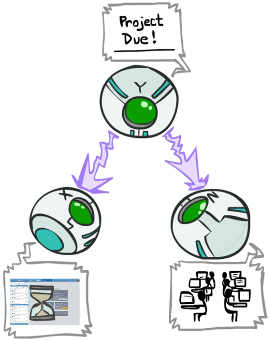
Guaranteed X and Z independent given Y?

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

Common Cause

This configuration is a "common cause"

Y: Project due



X: Forums busy   Z: Lab full

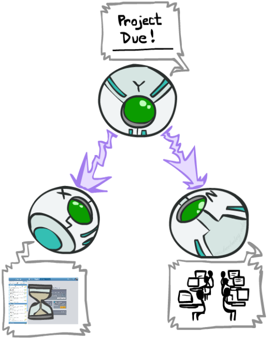
Guaranteed X and Z independent given Y?

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

Common Cause

This configuration is a "common cause"

Y: Project due



X: Forums busy   Z: Lab full

Guaranteed X and Z independent given Y?

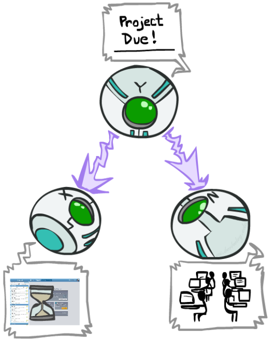
$$
\begin{aligned}
P(z|x,y) &= \frac{P(x,y,z)}{P(x,y)} \\
&= \frac{P(y)P(x|y)P(z|y)}{P(y)P(x|y)} \\
&= P(z|y)
\end{aligned}
$$

Yes!

Common Cause

This configuration is a "common cause"

Y: Project due



X: Forums busy Z: Lab full

Guaranteed X and Z independent given Y?

$$P(z|x,y) = \frac{P(x,y,z)}{P(x,y)}$$
$$= \frac{P(y)P(x|y)P(z|y)}{P(y)P(x|y)}$$
$$= P(z|y)$$

Yes!

Observing the cause blocks influence between effects.

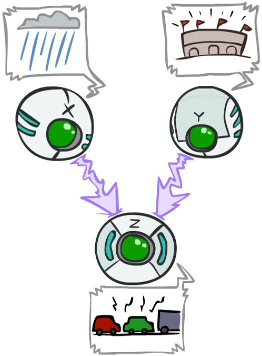
Common Effect

One last configuration: two causes of one effect (v-structures)

Common Effect

One last configuration: two causes of one effect (v-structures)

X: Raining Y:Ballgame

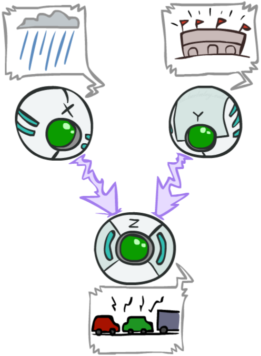


Z: Traffic

Common Effect

One last configuration: two causes of one effect (v-structures)

X: Raining Y:Ballgame



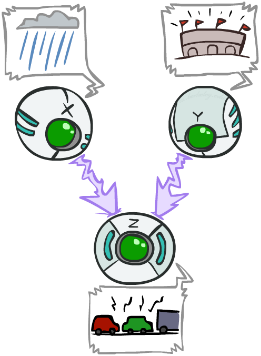
Are X and Y independent?

Z: Traffic

Common Effect

One last configuration: two causes of one effect (v-structures)

X: Raining Y:Ballgame



Z: Traffic

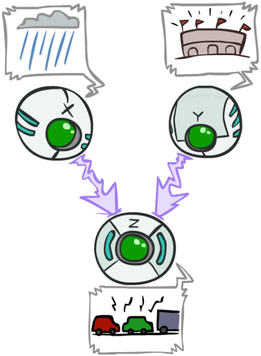
Are X and Y independent?

- Yes: the ballgame and the rain cause traffic, but they are not correlated

Common Effect

One last configuration: two causes of one effect (v-structures)

X: Raining Y:Ballgame



Z: Traffic

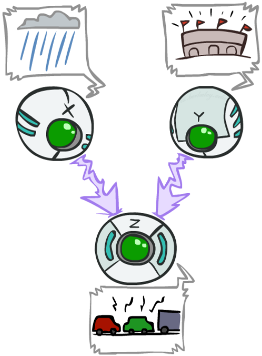
Are X and Y independent?

- Yes: the ballgame and the rain cause traffic, but they are not correlated
- Still need to prove they must be (try it!)

Common Effect

One last configuration: two causes of one effect (v-structures)

X: Raining Y:Ballgame



Z: Traffic

Are X and Y independent?

- Yes: the ballgame and the rain cause traffic, but they are not correlated
- Still need to prove they must be (try it!)

Common Effect

One last configuration: two causes of one effect (v-structures)

X: Raining Y:Ballgame

Z: Traffic

Are X and Y independent?

- Yes: the ballgame and the rain cause traffic, but they are not correlated
- Still need to prove they must be (try it!)

Are X and Y independent given Z?

Common Effect

One last configuration: two causes of one effect (v-structures)

X: Raining Y:Ballgame

Z: Traffic

Are X and Y independent?

- Yes: the ballgame and the rain cause traffic, but they are not correlated
- Still need to prove they must be (try it!)

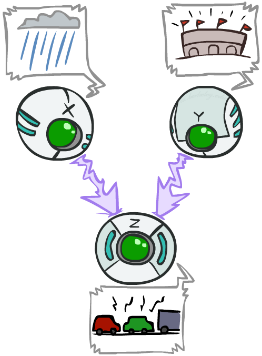
Are X and Y independent given Z?

- No: seeing traffic puts the rain and the ballgame in competition as explanation.

Common Effect

One last configuration: two causes of one effect (v-structures)

X: Raining   Y:Ballgame



Z: Traffic

Are X and Y independent?

- Yes: the ballgame and the rain cause traffic, but they are not correlated
- Still need to prove they must be (try it!)

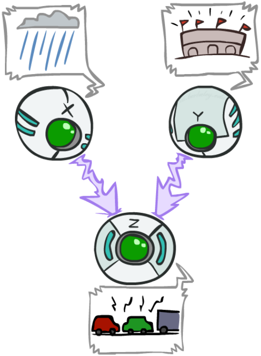
Are X and Y independent given Z?

- No: seeing traffic puts the rain and the ballgame in competition as explanation.

Common Effect

One last configuration: two causes of one effect (v-structures)

X: Raining Y:Ballgame



Z: Traffic

Are X and Y independent?

- Yes: the ballgame and the rain cause traffic, but they are not correlated
- Still need to prove they must be (try it!)

Are X and Y independent given Z?

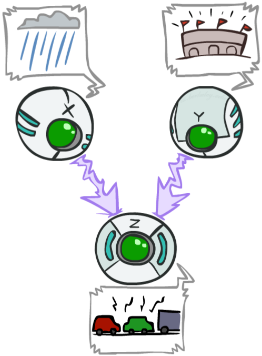
- No: seeing traffic puts the rain and the ballgame in competition as explanation.

This is backwards from the other cases

Common Effect

One last configuration: two causes of one effect (v-structures)

X: Raining Y:Ballgame



Z: Traffic

Are X and Y independent?

- Yes: the ballgame and the rain cause traffic, but they are not correlated
- Still need to prove they must be (try it!)

Are X and Y independent given Z?

- No: seeing traffic puts the rain and the ballgame in competition as explanation.

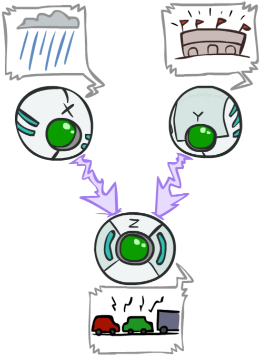
This is backwards from the other cases

- Observing an effect activates influence between possible causes.

Common Effect

One last configuration: two causes of one effect (v-structures)

X: Raining   Y:Ballgame



Z: Traffic

Are X and Y independent?

- Yes: the ballgame and the rain cause traffic, but they are not correlated
- Still need to prove they must be (try it!)

Are X and Y independent given Z?

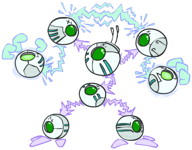
- No: seeing traffic puts the rain and the ballgame in competition as explanation.

This is backwards from the other cases

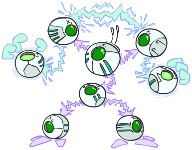
- Observing an effect activates influence between possible causes.

The General Case

The General Case

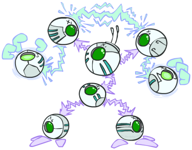


The General Case



General question: in a given BN, are two variables independent (given evidence)?

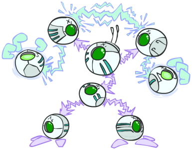
The General Case



General question: in a given BN, are two variables independent (given evidence)?

Solution: analyze the graph

The General Case

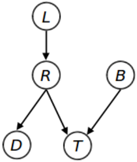


General question: in a given BN, are two variables independent (given evidence)?

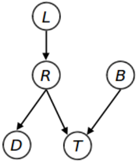
Solution: analyze the graph

Any complex example can be broken into repetitions of the three canonical cases

Reachability

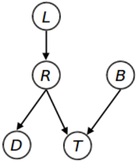


Reachability



Recipe: shade evidence nodes, look for paths in the resulting graph

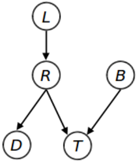
Reachability



Recipe: shade evidence nodes, look for paths in the resulting graph

Attempt 1: if two nodes are connected by an undirected path not blocked by a shaded node, they are conditionally independent

Reachability



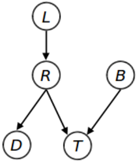
Recipe: shade evidence nodes, look for paths in the resulting graph

Attempt 1: if two nodes are connected by an undirected path not blocked by a shaded node, they are conditionally independent

Almost works, but not quite

- Where does it break?

Reachability

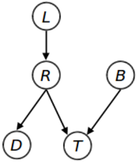


Recipe: shade evidence nodes, look for paths in the resulting graph

Attempt 1: if two nodes are connected by an undirected path not blocked by a shaded node, they are conditionally independent

Almost works, but not quite
- Where does it break?
- Answer: the v-structure at T doesn't count as a link in a path unless "active"

Reachability



Recipe: shade evidence nodes, look for paths in the resulting graph

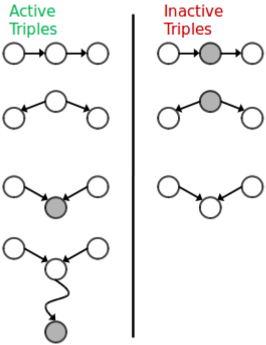
Attempt 1: if two nodes are connected by an undirected path not blocked by a shaded node, they are conditionally independent

Almost works, but not quite

- Where does it break?
- Answer: the v-structure at T doesn't count as a link in a path unless "active"

Active / Inactive Paths

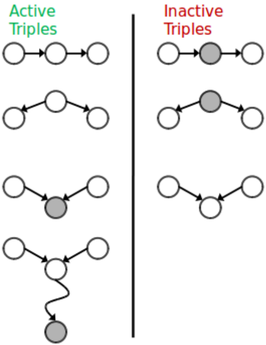
Question: Are X and Y conditionally independent given evidence variables Z?



Active / Inactive Paths

Question: Are X and Y conditionally independent given evidence variables Z?

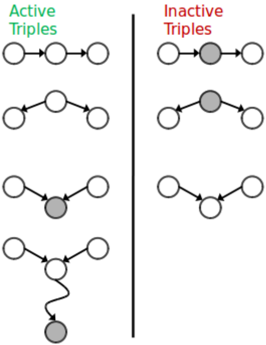
- Yes, if X and Y "d-separated" by Z



Active / Inactive Paths

Question: Are X and Y conditionally independent given evidence variables Z?

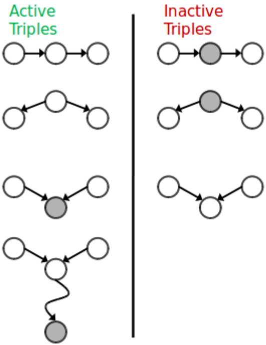
- Yes, if X and Y "d-separated" by Z
- Consider all (undirected) paths from X to Y



Active / Inactive Paths

Question: Are X and Y conditionally independent given evidence variables Z?

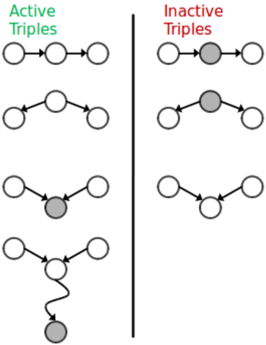
- Yes, if X and Y “d-separated” by Z
- Consider all (undirected) paths from X to Y
- No active paths = independence!



Active / Inactive Paths

Question: Are X and Y conditionally independent given evidence variables Z?

- Yes, if X and Y "d-separated" by Z
- Consider all (undirected) paths from X to Y
- No active paths = independence!

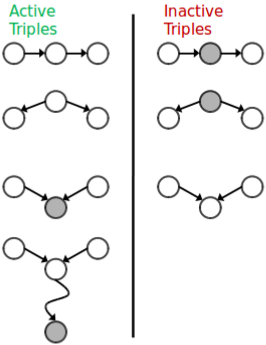


Active / Inactive Paths

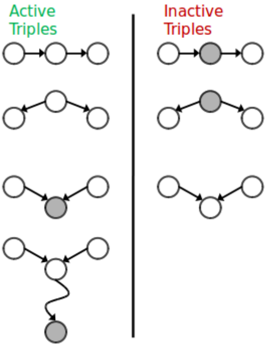
Question: Are X and Y conditionally independent given evidence variables Z?

- Yes, if X and Y “d-separated” by Z
- Consider all (undirected) paths from X to Y
- No active paths = independence!

A path is active if each triple is active:



Active / Inactive Paths



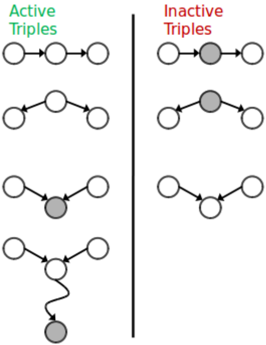
Question: Are X and Y conditionally independent given evidence variables Z?

- Yes, if X and Y “d-separated” by Z
- Consider all (undirected) paths from X to Y
- No active paths = independence!

A path is active if each triple is active:

- Causal chain:

Active / Inactive Paths



Question: Are X and Y conditionally independent given evidence variables Z?

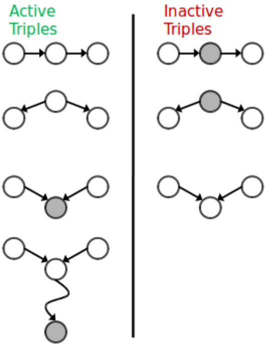
- Yes, if X and Y “d-separated” by Z
- Consider all (undirected) paths from X to Y
- No active paths = independence!

A path is active if each triple is active:

- Causal chain:

$$A \rightarrow B \rightarrow C$$

Active / Inactive Paths



Question: Are X and Y conditionally independent given evidence variables Z?

- Yes, if X and Y “d-separated” by Z
- Consider all (undirected) paths from X to Y
- No active paths = independence!

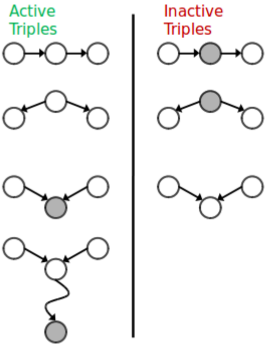
A path is active if each triple is active:

- Causal chain:

$$A \rightarrow B \rightarrow C$$

where B is unobserved (either direction)

Active / Inactive Paths



Question: Are X and Y conditionally independent given evidence variables Z?

- Yes, if X and Y “d-separated” by Z
- Consider all (undirected) paths from X to Y
- No active paths = independence!

A path is active if each triple is active:

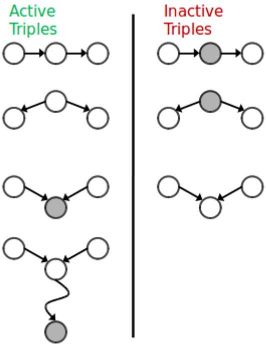
- Causal chain:

$$A \rightarrow B \rightarrow C$$

where B is unobserved (either direction)

- Common cause:

Active / Inactive Paths



Question: Are X and Y conditionally independent given evidence variables Z?

- Yes, if X and Y "d-separated" by Z
- Consider all (undirected) paths from X to Y
- No active paths = independence!

A path is active if each triple is active:

- Causal chain:

  $A \rightarrow B \rightarrow C$

  where B is unobserved (either direction)

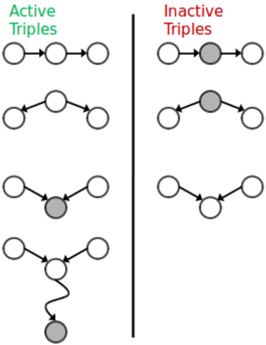
- Common cause:

  $A \leftarrow B \rightarrow C$ where B is unobserved

Active / Inactive Paths

Question: Are X and Y conditionally independent given evidence variables Z?

- Yes, if X and Y “d-separated” by Z
- Consider all (undirected) paths from X to Y
- No active paths = independence!



A path is active if each triple is active:

- Causal chain:

  $A \rightarrow B \rightarrow C$

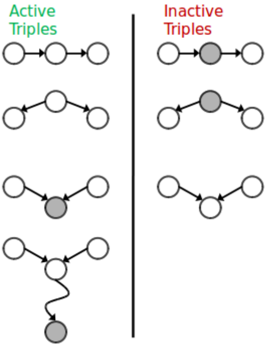
where B is unobserved (either direction)

- Common cause:

  $A \leftarrow B \rightarrow C$ where B is unobserved

- Common effect (aka v-structure)

Active / Inactive Paths



Question: Are X and Y conditionally independent given evidence variables Z?

- Yes, if X and Y "d-separated" by Z
- Consider all (undirected) paths from X to Y
- No active paths = independence!

A path is active if each triple is active:

- Causal chain:

  $A \rightarrow B \rightarrow C$

where B is unobserved (either direction)

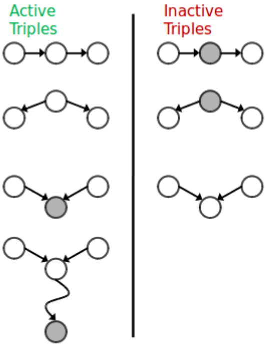
- Common cause:

  $A \leftarrow B \rightarrow C$ where B is unobserved

- Common effect (aka v-structure)

  $A \rightarrow B \leftarrow C$

Active / Inactive Paths



Question: Are X and Y conditionally independent given evidence variables Z?

- Yes, if X and Y "d-separated" by Z
- Consider all (undirected) paths from X to Y
- No active paths = independence!

A path is active if each triple is active:

- Causal chain:

  $A \to B \to C$
  where B is unobserved (either direction)

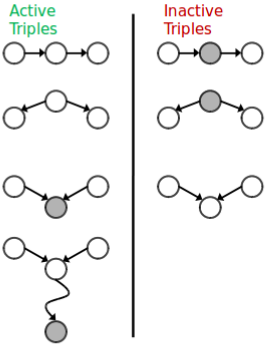
- Common cause:

  $A \leftarrow B \to C$ where B is unobserved

- Common effect (aka v-structure)

  $A \to B \leftarrow C$
  where B or one of its descendents is observed

Active / Inactive Paths



Question: Are X and Y conditionally independent given evidence variables Z?

- Yes, if X and Y "d-separated" by Z
- Consider all (undirected) paths from X to Y
- No active paths = independence!

A path is active if each triple is active:

- Causal chain:

  $A \rightarrow B \rightarrow C$

  where B is unobserved (either direction)

- Common cause:

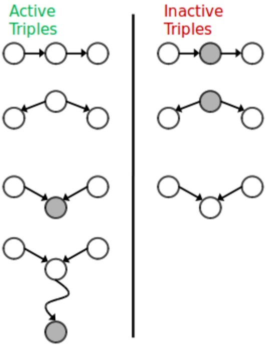
  $A \leftarrow B \rightarrow C$ where B is unobserved

- Common effect (aka v-structure)

  $A \rightarrow B \leftarrow C$

  where B or one of its descendents is observed

Active / Inactive Paths



Question: Are X and Y conditionally independent given evidence variables Z?

- Yes, if X and Y "d-separated" by Z
- Consider all (undirected) paths from X to Y
- No active paths = independence!

A path is active if each triple is active:

- Causal chain:

  $A \rightarrow B \rightarrow C$
  where B is unobserved (either direction)
- Common cause:

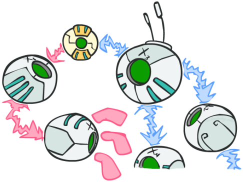
  $A \leftarrow B \rightarrow C$ where B is unobserved
- Common effect (aka v-structure)

  $A \rightarrow B \leftarrow C$
  where B or one of its descendents is observed

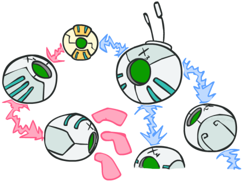
All it takes to block a path is a single inactive segment

D-separation

Query: $X_i \perp\!\!\!\perp X_j | \{X_{k_1}, \ldots, X_{k_n}\}$?



D-separation

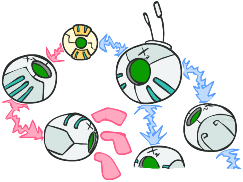


Query: $X_i \perp\!\!\!\perp X_j | \{X_{k_1}, \ldots, X_{k_n}\}$?

Check all (undirected!) paths between X_i and X_j

- If one or more active, then independence not guaranteed

D-separation

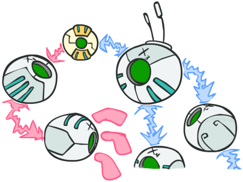


Query: $X_i \perp\!\!\!\perp X_j | \{X_{k_1}, \ldots, X_{k_n}\}$?

Check all (undirected!) paths between X_i and X_j

- If one or more active, then independence not guaranteed

D-separation



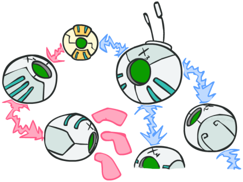
Query: $X_i \perp\!\!\!\perp X_j | \{X_{k_1}, \ldots, X_{k_n}\}$?

Check all (undirected!) paths between X_i and X_j

- If one or more active, then independence not guaranteed

$$X_i \not\perp\!\!\!\perp X_j | \{X_{k_1}, \ldots, X_{k_n}\}$$

D-separation



Query: $X_i \perp\!\!\!\perp X_j | \{X_{k_1}, \ldots, X_{k_n}\}$?

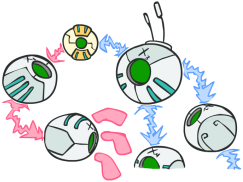
Check all (undirected!) paths between X_i and X_j

- If one or more active, then independence not guaranteed

$$X_i \not\perp\!\!\!\perp X_j | \{X_{k_1}, \ldots, X_{k_n}\}$$

- Otherwise (i.e. if all paths are inactive), then independence is guaranteed

D-separation



Query: $X_i \perp\!\!\!\perp X_j | \{X_{k_1}, \ldots, X_{k_n}\}$?

Check all (undirected!) paths between X_i and X_j

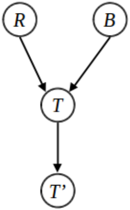
- If one or more active, then independence not guaranteed

$$X_i \not\perp\!\!\!\perp X_j | \{X_{k_1}, \ldots, X_{k_n}\}$$

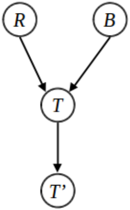
- Otherwise (i.e. if all paths are inactive), then independence is guaranteed

$$X_i \perp\!\!\!\perp X_j | \{X_{k_1}, \ldots, X_{k_n}\}$$

Example

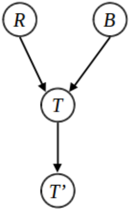


Example



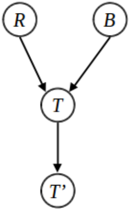
$R \perp\!\!\!\perp B$?

Example



$R \perp\!\!\!\perp B$? Yes.

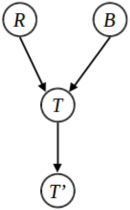
Example



$R \perp\!\!\!\perp B$? Yes.

$R \perp\!\!\!\perp B | T$?

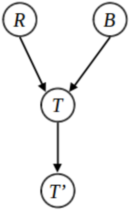
Example



$R \perp\!\!\!\perp B$? Yes.

$R \perp\!\!\!\perp B | T$? No.

Example

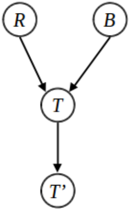


$R \perp\!\!\!\perp B$? Yes.

$R \perp\!\!\!\perp B | T$? No.

$R \perp\!\!\!\perp B | T'$?

Example

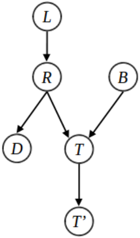


$R \perp\!\!\!\perp B$? Yes.

$R \perp\!\!\!\perp B | T$? No.

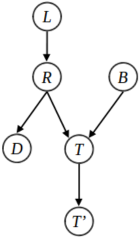
$R \perp\!\!\!\perp B | T'$? No

Example



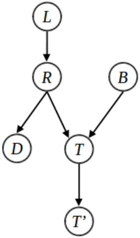
$L \perp\!\!\!\perp T' \mid T$?

Example



$L \perp\!\!\!\perp T' | T$? Yes

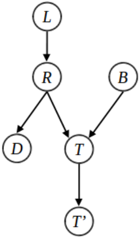
Example



$L \perp\!\!\!\perp T' | T$? Yes

$L \perp\!\!\!\perp B$?

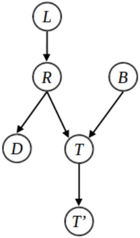
Example



$L \perp\!\!\!\perp T' \mid T$? Yes

$L \perp\!\!\!\perp B$? Yes

Example

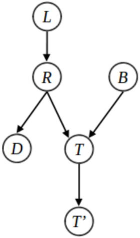


$L \perp\!\!\!\perp T' | T$? Yes

$L \perp\!\!\!\perp B$? Yes

$L \perp\!\!\!\perp B | T$?

Example

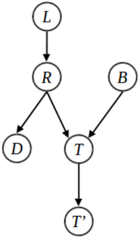


$L \perp\!\!\!\perp T' | T$? Yes

$L \perp\!\!\!\perp B$? Yes

$L \perp\!\!\!\perp B | T$? No

Example



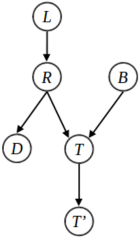
$L \perp\!\!\!\perp T' | T$? Yes

$L \perp\!\!\!\perp B$? Yes

$L \perp\!\!\!\perp B | T$? No

$L \perp\!\!\!\perp B | T'$?

Example



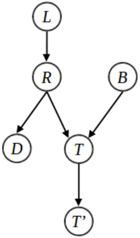
$L \perp\!\!\!\perp T' | T$? Yes

$L \perp\!\!\!\perp B$? Yes

$L \perp\!\!\!\perp B | T$? No

$L \perp\!\!\!\perp B | T'$? No

Example



$L \perp\!\!\!\perp T' | T$? Yes

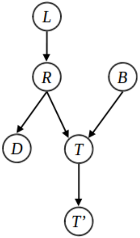
$L \perp\!\!\!\perp B$? Yes

$L \perp\!\!\!\perp B | T$? No

$L \perp\!\!\!\perp B | T'$? No

$L \perp\!\!\!\perp B | T, R$?

Example



$L \perp\!\!\!\perp T' | T$? Yes

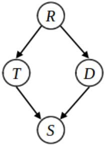
$L \perp\!\!\!\perp B$? Yes

$L \perp\!\!\!\perp B | T$? No

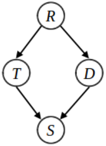
$L \perp\!\!\!\perp B | T'$? No

$L \perp\!\!\!\perp B | T, R$? Yes

Example



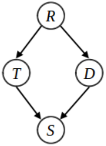
Example



Variables:

- R: Raining

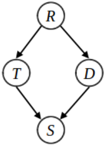
Example



Variables:

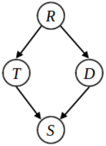
- R: Raining
- T: Traffic

Example



Variables:
- R: Raining
- T: Traffic
- D: Roof drips

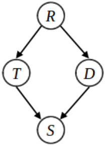
Example



Variables:

- R: Raining
- T: Traffic
- D: Roof drips
- S: I'm sad

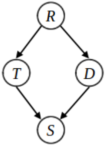
Example



Variables:

- R: Raining
- T: Traffic
- D: Roof drips
- S: I'm sad

Example

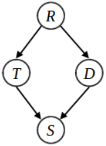


Variables:

- R: Raining
- T: Traffic
- D: Roof drips
- S: I’m sad

Questions:

Example



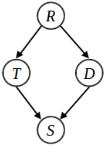
Variables:

- R: Raining
- T: Traffic
- D: Roof drips
- S: I'm sad

Questions:

$T \perp\!\!\!\perp D$?

Example



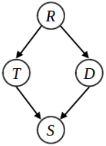
Variables:

- R: Raining
- T: Traffic
- D: Roof drips
- S: I'm sad

Questions:

$T \perp\!\!\!\perp D$? No..

Example



Variables:

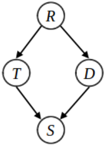
- R: Raining
- T: Traffic
- D: Roof drips
- S: I'm sad

Questions:

$T \perp\!\!\!\perp D$? No..

$T \perp\!\!\!\perp D | R$?

Example



Variables:

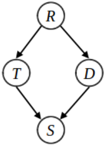
- R: Raining
- T: Traffic
- D: Roof drips
- S: I'm sad

Questions:

$T \perp\!\!\!\perp D$? No..

$T \perp\!\!\!\perp D | R$? Yes.

Example



Variables:

- R: Raining
- T: Traffic
- D: Roof drips
- S: I'm sad

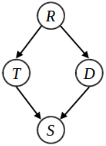
Questions:

$T \perp\!\!\!\perp D$? No..

$T \perp\!\!\!\perp D | R$? Yes.

$T \perp\!\!\!\perp D | R, S$?

Example



Variables:

- R: Raining
- T: Traffic
- D: Roof drips
- S: I'm sad

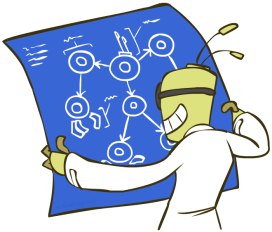
Questions:

$T \perp\!\!\!\perp D$? No..

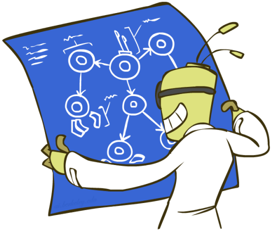
$T \perp\!\!\!\perp D | R$? Yes.

$T \perp\!\!\!\perp D | R, S$? No.

Structure Implications

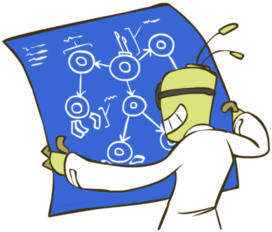


Structure Implications



Given a Bayes net structure, can run d-separation algorithm to build a complete list of conditional independences that are necessarily true of the form

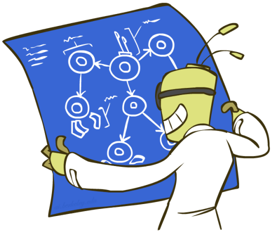
Structure Implications



Given a Bayes net structure, can run d-separation algorithm to build a complete list of conditional independences that are necessarily true of the form

$$X_i \perp\!\!\!\perp X_j | \{X_{k_1}, \ldots, X_{k_n}\}$$

Structure Implications

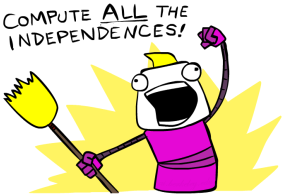


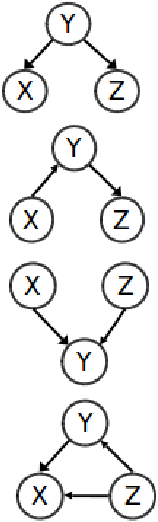
Given a Bayes net structure, can run d-separation algorithm to build a complete list of conditional independences that are necessarily true of the form

$$X_i \perp\!\!\!\perp X_j | \{X_{k_1}, \ldots, X_{k_n}\}$$

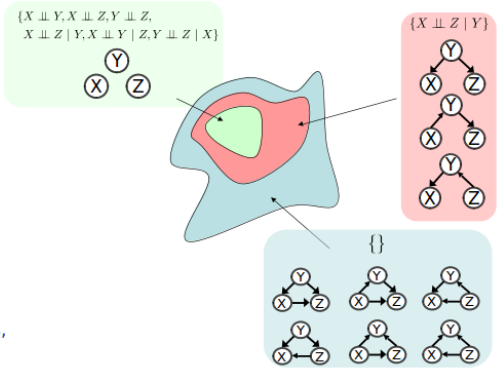
This list determines the set of probability distributions that can be represented

Computing All Independences





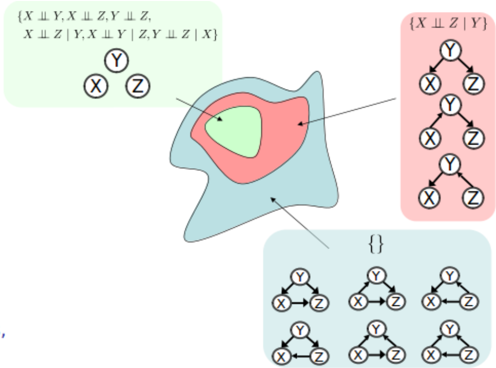
Topology Limits Distributions



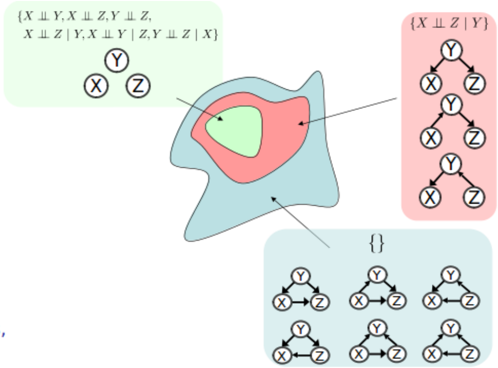
s,

Topology Limits Distributions

Given some graph topology G, only certain joint distributions can be encoded



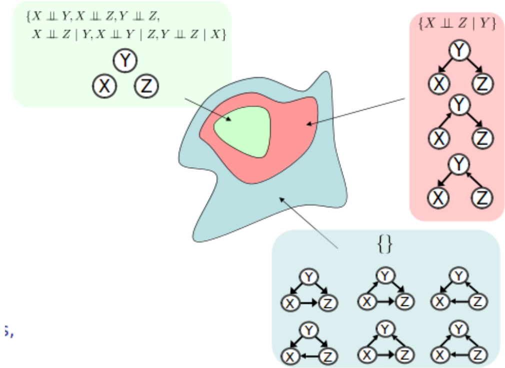
Topology Limits Distributions



Given some graph topology G, only certain joint distributions can be encoded

The graph structure guarantees certain (conditional) independences

Topology Limits Distributions

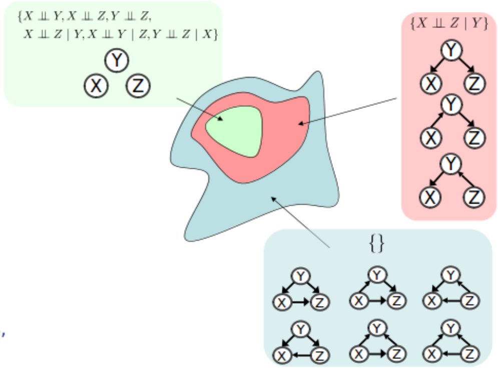


Given some graph topology G, only certain joint distributions can be encoded

The graph structure guarantees certain (conditional) independences

(There might be more independence)

Topology Limits Distributions



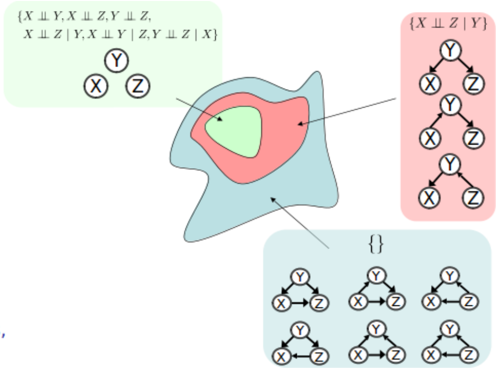
Given some graph topology G, only certain joint distributions can be encoded

The graph structure guarantees certain (conditional) independences

(There might be more independence)

Adding arcs increases the set of distributions, but has several costs

Topology Limits Distributions



Given some graph topology G, only certain joint distributions can be encoded

The graph structure guarantees certain (conditional) independences

(There might be more independence)

Adding arcs increases the set of distributions, but has several costs

Full conditioning can encode any distribution

Bayes Nets Representation Summary

Bayes Nets Representation Summary

Bayes nets compactly encode joint distributions

Bayes Nets Representation Summary

Bayes nets compactly encode joint distributions

Guaranteed independencies of distributions can be deduced from BN graph structure

Bayes Nets Representation Summary

Bayes nets compactly encode joint distributions

Guaranteed independencies of distributions can be deduced from BN graph structure

D-separation gives precise conditional independence guarantees from graph alone

Bayes Nets Representation Summary

Bayes nets compactly encode joint distributions

Guaranteed independencies of distributions can be deduced from BN graph structure

D-separation gives precise conditional independence guarantees from graph alone

A Bayes' net's joint distribution may have further (conditional) independence that is not detectable until you inspect its specific distribution

Bayes' Nets

Bayes' Nets

Representation

Bayes' Nets

Representation ✓

Bayes' Nets

Representation ✓

Conditional Independences

Bayes' Nets

Representation ✓

Conditional Independences ✓

Bayes' Nets

Representation ✓

Conditional Independences ✓

Probabilistic Inference

- Enumeration (exact, exponential complexity)

Bayes' Nets

Representation ✓

Conditional Independences ✓

Probabilistic Inference

- Enumeration (exact, exponential complexity)
- Variable elimination (exact, worst-case exponential complexity, often better)

Bayes' Nets

Representation ✓

Conditional Independences ✓

Probabilistic Inference

- Enumeration (exact, exponential complexity)
- Variable elimination (exact, worst-case exponential complexity, often better)
- Probabilistic inference is NP-complete

Bayes' Nets

Representation ✓

Conditional Independences ✓

Probabilistic Inference

- Enumeration (exact, exponential complexity)
- Variable elimination (exact, worst-case exponential complexity, often better)
- Probabilistic inference is NP-complete
- Sampling (approximate)

Bayes' Nets

Representation ✓

Conditional Independences ✓

Probabilistic Inference

- Enumeration (exact, exponential complexity)
- Variable elimination (exact, worst-case exponential complexity, often better)
- Probabilistic inference is NP-complete
- Sampling (approximate)

Learning Bayes' Nets from Data

Bayes' Nets

Bayes' Nets

Representation

Bayes' Nets

Representation ✓

Bayes' Nets

Representation ✓

Conditional Independences

Bayes' Nets

Representation ✓

Conditional Independences ✓

Bayes' Nets

Representation ✓

Conditional Independences ✓

Probabilistic Inference (Next).

- Enumeration (exact, exponential complexity)

Bayes' Nets

Representation ✓

Conditional Independences ✓

Probabilistic Inference (Next).

- Enumeration (exact, exponential complexity)
- Variable elimination (exact, worst-case exponential complexity, often better)

Bayes' Nets

Representation ✓

Conditional Independences ✓

Probabilistic Inference (Next).

- Enumeration (exact, exponential complexity)
- Variable elimination (exact, worst-case exponential complexity, often better)
- Probabilistic inference is NP-complete

Bayes' Nets

Representation ✓

Conditional Independences ✓

Probabilistic Inference (Next).

- Enumeration (exact, exponential complexity)
- Variable elimination (exact, worst-case exponential complexity, often better)
- Probabilistic inference is NP-complete
- Sampling (approximate)

Bayes' Nets

Representation ✓

Conditional Independences ✓

Probabilistic Inference (Next).

- Enumeration (exact, exponential complexity)
- Variable elimination (exact, worst-case exponential complexity, often better)
- Probabilistic inference is NP-complete
- Sampling (approximate)

Learning Bayes' Nets from Data

Examples:

- Posterior probability

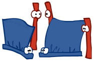
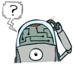
Examples:

- Posterior probability
- Most likely explanation:

Examples:

- Posterior probability
- Most likely explanation:

Examples:

- Posterior probability
- Most likely explanation:

Inference

Examples:

- Posterior probability
- Most likely explanation:

Inference

Inference: calculating some useful quantity from a joint probability distribution

Inference by Enumeration

General case:

- Evidence variables:

$E_1, \ldots, E_k = e_1, \ldots, e_k$

Inference by Enumeration

General case:

- Evidence variables: $E_1, \ldots, E_k = e_1, \ldots, e_k$
- Query* variable: Q

Inference by Enumeration

General case:

- Evidence variables:

$E_1, \ldots, E_k = e_1, \ldots, e_k$

- Query* variable: Q
- Hidden variables: $H_1, \ldots, H_r$

Inference by Enumeration

General case:

- Evidence variables: $E_1, \ldots, E_k = e_1, \ldots, e_k$
- Query* variable: Q
- Hidden variables: $H_1, \ldots, H_r$

Inference by Enumeration

General case:

- Evidence variables:

$E_1, \ldots, E_k = e_1, \ldots, e_k$

- Query* variable: Q
- Hidden variables: $H_1, \ldots, H_r$

All Variables:

Evidence $\cup$ $\{$Q$\}$ $\cup$ Hidden.

Inference by Enumeration

General case:

- Evidence variables:

$E_1, \ldots, E_k = e_1, \ldots, e_k$

- Query* variable: Q
- Hidden variables: $H_1, \ldots, H_r$

All Variables:

Evidence $\cup$ $\{$Q$\}$ $\cup$ Hidden.

We want: $Pr(Q|e_1, \ldots, e_k)$

Inference by Enumeration

General case:

- Evidence variables:

$E_1, \ldots, E_k = e_1, \ldots, e_k$

- Query* variable: Q
- Hidden variables: $H_1, \ldots, H_r$

All Variables:

Evidence $\cup$ $\{$Q$\}$ $\cup$ Hidden.

We want: $Pr(Q|e_1, \ldots, e_k)$

* Works fine with multiple query variables, too

Inference by Enumeration

General case:

- Evidence variables: $E_1, \ldots, E_k = e_1, \ldots, e_k$
- Query* variable: Q
- Hidden variables: $H_1, \ldots, H_r$

All Variables:

Evidence $\cup$ $\{$Q$\}$ $\cup$ Hidden.

We want: $Pr(Q|e_1, \ldots, e_k)$

* Works fine with multiple query variables, too

Step 1: Select the entries consistent with the evidence

Inference by Enumeration

General case:

- Evidence variables: $E_1, \ldots, E_k = e_1, \ldots, e_k$
- Query* variable: Q
- Hidden variables: $H_1, \ldots, H_r$

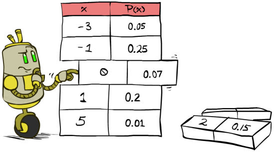
All Variables:

Evidence $\cup$ $\{$Q$\}$ $\cup$ Hidden.

We want: $Pr(Q|e_1, \ldots, e_k)$

* Works fine with multiple query variables, too

Step 1: Select the entries consistent with the evidence



Inference by Enumeration

General case:

- Evidence variables: $E_1, \ldots, E_k = e_1, \ldots, e_k$
- Query* variable: Q
- Hidden variables: $H_1, \ldots, H_r$

All Variables:

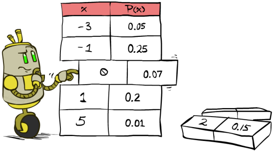
Evidence $\cup$ $\{$Q$\}$ $\cup$ Hidden.

We want: $Pr(Q|e_1, \ldots, e_k)$

* Works fine with multiple query variables, too

Step 1: Select the entries consistent with the evidence

Step 2: Sum out H to get joint of Query and evidence



Inference by Enumeration

General case:

- Evidence variables: $E_1, \ldots, E_k = e_1, \ldots, e_k$
- Query* variable: Q
- Hidden variables: $H_1, \ldots, H_r$

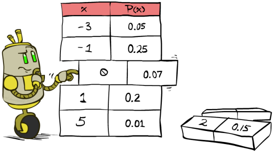
All Variables:

Evidence $\cup$ $\{Q\}$ $\cup$ Hidden.

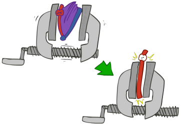
We want: $Pr(Q|e_1, \ldots, e_k)$

* Works fine with multiple query variables, too

Step 1: Select the entries consistent with the evidence



Step 2: Sum out H to get joint of Query and evidence



Inference by Enumeration

General case:

- Evidence variables: $E_1, \ldots, E_k = e_1, \ldots, e_k$
- Query* variable: Q
- Hidden variables: $H_1, \ldots, H_r$

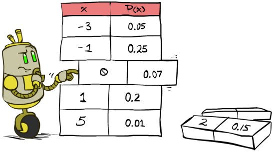
All Variables:

Evidence $\cup$ $\{$Q$\}$ $\cup$ Hidden.

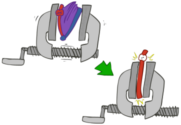
We want: $Pr(Q|e_1, \ldots, e_k)$

* Works fine with multiple query variables, too

Step 1: Select the entries consistent with the evidence



Step 2: Sum out H to get joint of Query and evidence



$P(Q, e_1, \ldots, e_k) =$
$\sum_{h_1, \ldots, h_r} P(Q, h_1, \ldots, h_r, e_1, \ldots, e_k)$

Step 3: Normalize.

Inference by Enumeration

General case:

- Evidence variables:

$E_1, \ldots, E_k = e_1, \ldots, e_k$

- Query* variable: Q
- Hidden variables: $H_1, \ldots, H_r$

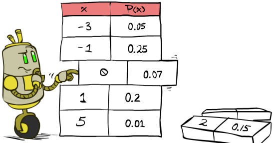
All Variables:

Evidence $\cup$ {Q} $\cup$ Hidden.

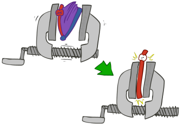
We want: $Pr(Q|e_1, \ldots, e_k)$

* Works fine with multiple query variables, too

Step 1: Select the entries consistent with the evidence



Step 2: Sum out H to get joint of Query and evidence



$P(Q, e_1, \ldots, e_k) = \sum_{h_1,\ldots,h_r} P(Q, h_1, \ldots, h_r, e_1, \ldots, e_k)$

Step 3: Normalize.

$Z = \sum_q P(Q = q, e_1, \ldots, e_k)$

Inference by Enumeration

General case:

- Evidence variables:

$E_1, \ldots, E_k = e_1, \ldots, e_k$

- Query* variable: Q
- Hidden variables: $H_1, \ldots, H_r$

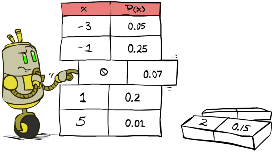
We want: $Pr(Q|e_1, \ldots, e_k)$

* Works fine with multiple query variables, too

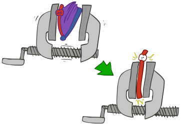
All Variables:

Evidence $\cup$ $\{$Q$\}$ $\cup$ Hidden.

Step 1: Select the entries consistent with the evidence



Step 2: Sum out H to get joint of Query and evidence



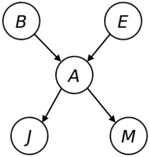
$P(Q, e_1, \ldots, e_k) = \sum_{h_1,\ldots,h_r} P(Q, h_1, \ldots, h_r, e_1, \ldots, e_k)$

Step 3: Normalize.

$Z = \sum_q P(Q = q, e_1, \ldots, e_k)$

$P(Q|e_1, \ldots, e_k) = \frac{1}{Z} P(Q, e_1, \ldots, e_k)$

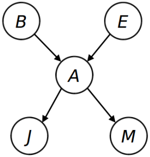
Inference by Enumeration in Bayes' Net



Given unlimited time, inference in BNs is easy.

$P(B|+j,+m) \propto P(B,+j,+m)$

Inference by Enumeration in Bayes' Net

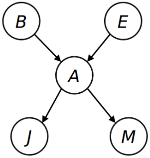


Given unlimited time, inference in BNs is easy.

$$P(B \mid +j, +m) \propto P(B, +j, +m)$$

$$= \sum_{e,a} P(B, e, a, +j, +m)$$

Inference by Enumeration in Bayes' Net



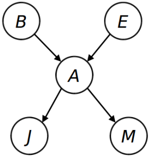
Given unlimited time, inference in BNs is easy.

$$P(B|+j,+m) \propto P(B,+j,+m)$$

$$= \sum_{e,a} P(B,e,a,+j,+m)$$

$$= \sum_{e,a} P(B)P(e)P(a|B,e)P(+j|a)P(+m|a)$$

Inference by Enumeration in Bayes' Net



Given unlimited time, inference in BNs is easy.

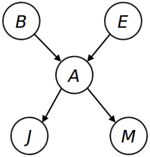
$$P(B|+j,+m) \propto P(B,+j,+m)$$

$$= \sum_{e,a} P(B,e,a,+j,+m)$$

$$= \sum_{e,a} P(B)P(e)P(a|B,e)P(+j|a)P(+m|a)$$

$$= P(B)P(+e)P(+a|B,+e)P(+j|+a)P(+m|+a)$$

Inference by Enumeration in Bayes' Net



Given unlimited time, inference in BNs is easy.

$$P(B|+j,+m) \propto P(B,+j,+m)$$

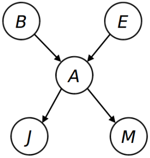
$$= \sum_{e,a} P(B,e,a,+j,+m)$$

$$= \sum_{e,a} P(B)P(e)P(a|B,e)P(+j|a)P(+m|a)$$

$$= P(B)P(+e)P(+a|B,+e)P(+j|+a)P(+m|+a)$$

$$+P(B)P(+e)P(-a|B,+e)P(+j|-a)P(+m|-a)+$$

Inference by Enumeration in Bayes' Net

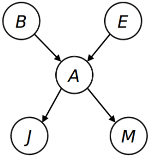


Given unlimited time, inference in BNs is easy.

$$
\begin{aligned}
P(B|+j,+m) &\propto P(B,+j,+m) \\
&= \textstyle\sum_{e,a} P(B,e,a,+j,+m) \\
&= \textstyle\sum_{e,a} P(B)P(e)P(a|B,e)P(+j|a)P(+m|a)
\end{aligned}
$$

$$
\begin{aligned}
&= P(B)P(+e)P(+a|B,+e)P(+j|+a)P(+m|+a) \\
&\quad + P(B)P(+e)P(-a|B,+e)P(+j|-a)P(+m|-a) + \\
&\quad P(B)P(-e)P(+a|B,-e)P(+j|+a)P(+m|+a)
\end{aligned}
$$

Inference by Enumeration in Bayes' Net



Given unlimited time, inference in BNs is easy.

$$P(B|+j,+m) \propto P(B,+j,+m)$$

$$= \sum_{e,a} P(B,e,a,+j,+m)$$

$$= \sum_{e,a} P(B)P(e)P(a|B,e)P(+j|a)P(+m|a)$$

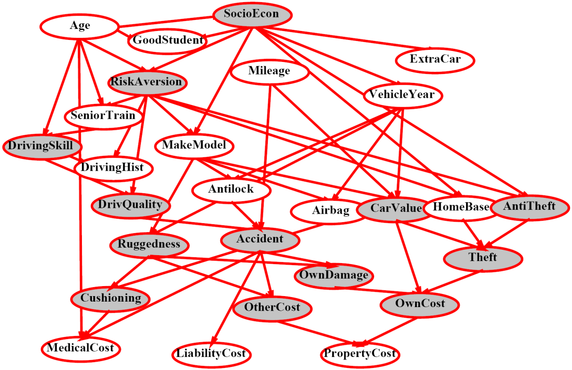
$$= P(B)P(+e)P(+a|B,+e)P(+j|+a)P(+m|+a)$$

$$+P(B)P(+e)P(-a|B,+e)P(+j|-a)P(+m|-a)+$$

$$P(B)P(-e)P(+a|B,-e)P(+j|+a)P(+m|+a)$$

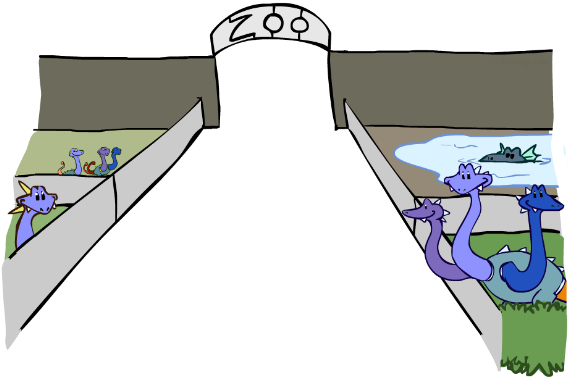
$$+P(B)P(-e)P(-a|B,-e)P(+j|-a)$$

Inference by Enumeration?

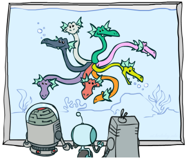


$$P(antilock|observedvariables) = ?$$

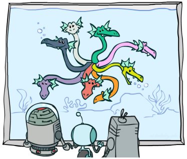
Factor Zoo



Factor Zoo I



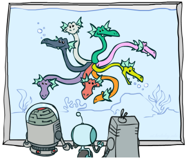
Factor Zoo I



Joint distribution: P(X,Y)

- Entries P(x,y) for all x, y

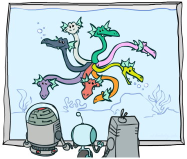
Factor Zoo I



Joint distribution: P(X,Y)

- Entries P(x,y) for all x, y
- Sums to?

Factor Zoo I

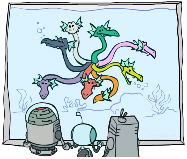


Joint distribution: P(X,Y)

- Entries P(x,y) for all x, y
- Sums to? 1

Factor Zoo I

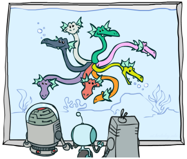
$P(T, W)$



Joint distribution: P(X,Y)

- Entries P(x,y) for all x, y
- Sums to? 1

Factor Zoo I



Joint distribution: P(X,Y)

- Entries P(x,y) for all x, y
- Sums to? 1

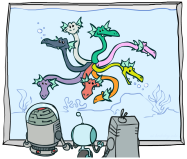
Selected joint: P(x,Y)

- Slice of joint distribution

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Factor Zoo I



Joint distribution: P(X,Y)

- Entries P(x,y) for all x, y
- Sums to? 1

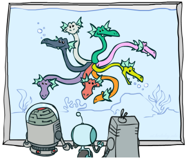
Selected joint: P(x,Y)

- Slice of joint distribution
- Entries P(x,y):

$P(T,W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Factor Zoo I



Joint distribution: P(X,Y)

- Entries P(x,y) for all x, y
- Sums to? 1

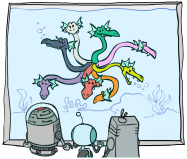
Selected joint: P(x,Y)

- Slice of joint distribution
- Entries P(x,y):
  fixed x, all y

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Factor Zoo I



Joint distribution: P(X,Y)

- Entries P(x,y) for all x, y
- Sums to? 1

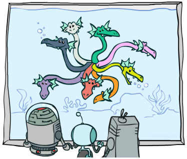
Selected joint: P(x,Y)

- Slice of joint distribution
- Entries P(x,y): fixed x, all y
- Table P(cold, W)?

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Factor Zoo I



Joint distribution: P(X,Y)

- Entries P(x,y) for all x, y
- Sums to? 1

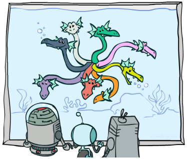
Selected joint: P(x,Y)

- Slice of joint distribution
- Entries P(x,y):
  fixed x, all y
- Table P(cold, W)?
- Sums to?

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Factor Zoo I



Joint distribution: P(X,Y)

- Entries P(x,y) for all x, y
- Sums to? 1

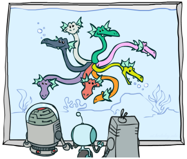
Selected joint: P(x,Y)

- Slice of joint distribution
- Entries P(x,y): fixed x, all y
- Table P(cold, W)?
- Sums to? P(x).

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Factor Zoo I



Joint distribution: P(X,Y)

- Entries P(x,y) for all x, y
- Sums to? 1

Selected joint: P(x,Y)

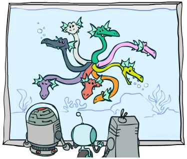
- Slice of joint distribution
- Entries P(x,y):
  fixed x, all y
- Table P(cold, W)?
- Sums to? P(x).

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$P(cold, W)$

Factor Zoo I



Joint distribution: P(X,Y)

- Entries P(x,y) for all x, y
- Sums to? 1

Selected joint: P(x,Y)

- Slice of joint distribution
- Entries P(x,y):
  fixed x, all y
- Table P(cold, W)?
- Sums to? P(x).

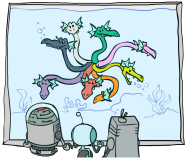
$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$P(cold, W)$

T	W	P
cold	sun	0.2
cold	rain	0.3

Factor Zoo I



Joint distribution: P(X,Y)

- Entries P(x,y) for all x, y
- Sums to? 1

Selected joint: P(x,Y)

- Slice of joint distribution
- Entries P(x,y):
  fixed x, all y
- Table P(cold, W)?
- Sums to? P(x).

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$P(cold, W)$

T	W	P
cold	sun	0.2
cold	rain	0.3

Number of capital letters (variables)= dimensionality of the table

Factor Zoo II

Single conditional: $P(Y|x)$

- Entries $P(y|x)$ for fixed x, all y

Factor Zoo II

Single conditional: $P(Y|x)$

- Entries $P(y|x)$ for fixed x, all y
- Sum of entries?

Factor Zoo II

Single conditional: $P(Y|x)$

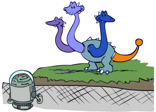
- Entries $P(y|x)$ for fixed x, all y
- Sum of entries? 1

Factor Zoo II

Single conditional: $P(Y|x)$

- Entries $P(y|x)$ for fixed x, all y
- Sum of entries? 1

T	W	P
cold	sun	0.4
cold	rain	0.6



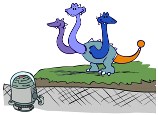
Family of conditionals:

Factor Zoo II

Single conditional: $P(Y|x)$

- Entries $P(y|x)$ for fixed x, all y
- Sum of entries? 1

T	W	P
cold	sun	0.4
cold	rain	0.6



Family of conditionals:

$P(X|Y)$

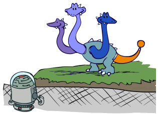
- Multiple conditionals

Factor Zoo II

Single conditional: $P(Y|x)$

- Entries $P(y|x)$ for fixed x, all y
- Sum of entries? 1

T	W	P
cold	sun	0.4
cold	rain	0.6



Family of conditionals:

$P(X|Y)$

- Multiple conditionals
- Entries $P(x|y)$ for all

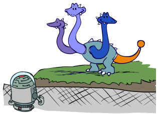
x, y

Factor Zoo II

Single conditional: $P(Y|x)$

- Entries $P(y|x)$ for fixed x, all y
- Sum of entries? 1

T	W	P
cold	sun	0.4
cold	rain	0.6



Family of conditionals:

$P(X|Y)$

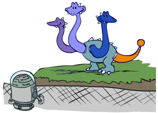
- Multiple conditionals
- Entries $P(x|y)$ for all x, y
- Sum?

Factor Zoo II

Single conditional: $P(Y|x)$

- Entries $P(y|x)$ for fixed x, all y
- Sum of entries? 1

T	W	P
cold	sun	0.4
cold	rain	0.6



Family of conditionals:

$P(X|Y)$

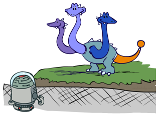
- Multiple conditionals
- Entries $P(x|y)$ for all x, y
- Sum? $|Y|$

Factor Zoo II

Single conditional: $P(Y|x)$

- Entries $P(y|x)$ for fixed x, all y
- Sum of entries? 1

T	W	P
cold	sun	0.4
cold	rain	0.6

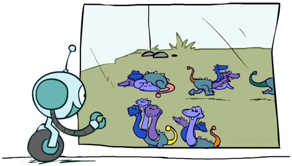


Family of conditionals:

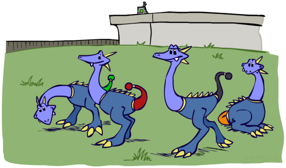
$P(X|Y)$

- Multiple conditionals
- Entries $P(x|y)$ for all x, y
- Sum? $|Y|$

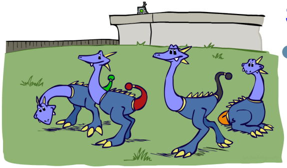
T	W	P
hot	sun	0.8
hot	rain	0.2
cold	sun	0.4
cold	rain	0.6



Factor Zoo III



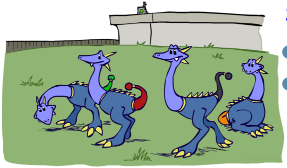
Factor Zoo III



Specified family: $P(y|X)$

- Entries $P(y|x)$ for fixed y, but for all x

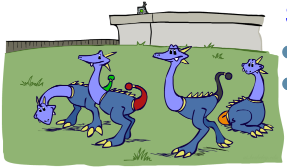
Factor Zoo III



Specified family: $P(y|X)$

- Entries $P(y|x)$ for fixed y, but for all x
- Sums to?

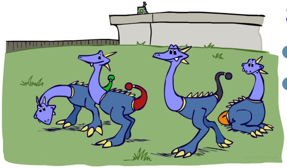
Factor Zoo III



Specified family: $P(y|X)$

- Entries $P(y|x)$ for fixed y, but for all x
- Sums to? who knows!

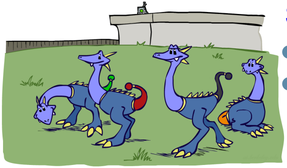
Factor Zoo III



Specified family: $P(y|X)$

- Entries $P(y|x)$ for fixed y, but for all x
- Sums to? who knows!

Factor Zoo III



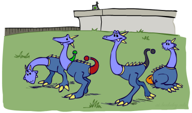
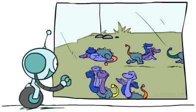
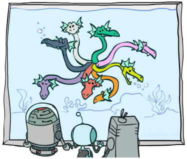
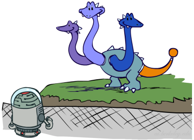
Specified family: $P(y|X)$

- Entries $P(y|x)$ for fixed y, but for all x
- Sums to? who knows!

$P(rain|T)$

T	W	P
hot	rain	0.2
cold	rain	0.6

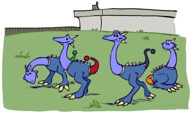
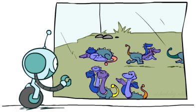
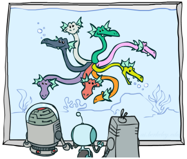
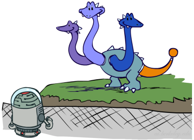
Factor Zoo Summary

In general, when we write $P(Y_1, \ldots, Y_N | X_1, \ldots, X_M)$

- It is a “factor,” a multi-dimensional array

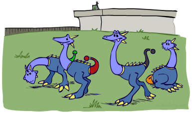
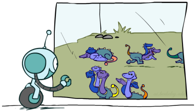
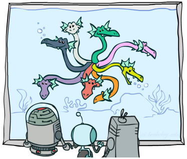
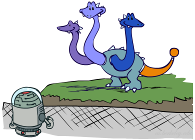
Factor Zoo Summary

In general, when we write $P(Y_1, \ldots, Y_N | X_1, \ldots, X_M)$

- It is a "factor," a multi-dimensional array
- Its values are $P(y_1, \ldots, y_N | x_1, \ldots, x_M)$

Factor Zoo Summary

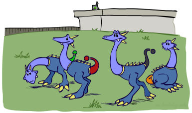
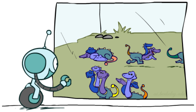
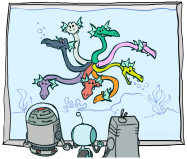
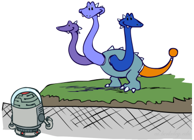





In general, when we write $P(Y_1, \ldots, Y_N | X_1, \ldots, X_M)$

- It is a "factor," a multi-dimensional array
- Its values are $P(y_1, \ldots, y_N | x_1, \ldots, x_M)$
- Any assigned (=lower-case) X or Y is a dimension missing

(selected) from the array

Factor Zoo Summary

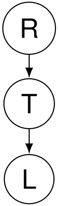
In general, when we write $P(Y_1, \ldots, Y_N | X_1, \ldots, X_M)$

- It is a "factor," a multi-dimensional array
- Its values are $P(y_1, \ldots, y_N | x_1, \ldots, x_M)$
- Any assigned (=lower-case) X or Y is a dimension missing (selected) from the array

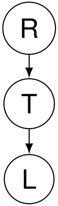
Example: Traffic Domain

$P(R)$

T	L
+r	0.1
-r	0.9



Example: Traffic Domain



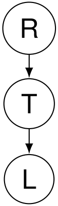
$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

Example: Traffic Domain



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

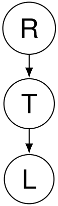
+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Random Variables

Example: Traffic Domain



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

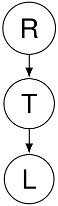
$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Random Variables

- R: Raining

Example: Traffic Domain



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

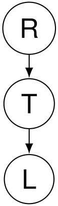
$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Random Variables

- R: Raining
- T: Traffic

Example: Traffic Domain



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

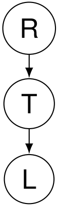
$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Random Variables

- R: Raining
- T: Traffic
- L: Late for class!

Example: Traffic Domain



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

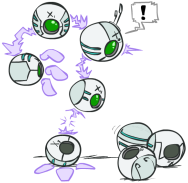
$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Random Variables

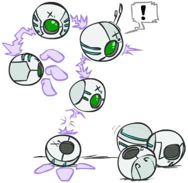
- R: Raining
- T: Traffic
- L: Late for class!

Inference by Enumeration: Procedural Outline



Inference by Enumeration: Procedural Outline

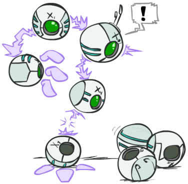
Track objects called factors.



Inference by Enumeration: Procedural Outline

Track objects called factors.

Initial factors are local CPTs (one per node)



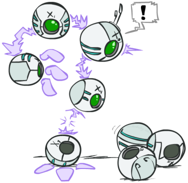
$P(R)$

T	L
+r	0.1
-r	0.9

Inference by Enumeration: Procedural Outline

Track objects called factors.

Initial factors are local CPTs (one per node)



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

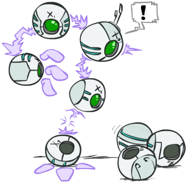
$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Inference by Enumeration: Procedural Outline

Track objects called factors.

Initial factors are local CPTs (one per node)



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

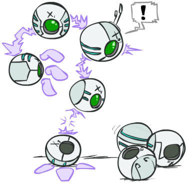
Any known values are selected.

E.g. if we know $L = +l$, the initial factors are

$P(R)$

T	L
+r	0.1
-r	0.9

Inference by Enumeration: Procedural Outline



Track objects called factors.

Initial factors are local CPTs (one per node)

$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Any known values are selected.

E.g. if we know $L = +l$, the initial factors are

$P(R)$

T	L
+r	0.1
-r	0.9

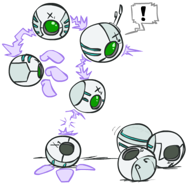
$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Inference by Enumeration: Procedural Outline



Track objects called factors.

Initial factors are local CPTs (one per node)

$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Any known values are selected.

E.g. if we know $L = +l$, the initial factors are

$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Procedure: Join all factors, then eliminate all hidden variables

Operation 1: Join Factors

Operation 1: Join Factors

First basic operation: joining factors

Operation 1: Join Factors

First basic operation: joining factors

Combining factors:

- Just like a database join

Operation 1: Join Factors

First basic operation: joining factors

Combining factors:

- Just like a database join
- Get all factors over the joining variable

Operation 1: Join Factors

First basic operation: joining factors

Combining factors:

- Just like a database join
- Get all factors over the joining variable
- Build a new factor over the union of the variables involved

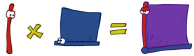
Operation 1: Join Factors

First basic operation: joining factors

Combining factors:

- Just like a database join
- Get all factors over the joining variable
- Build a new factor over the union of the variables involved

Operation 1: Join Factors

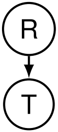


First basic operation: joining factors

Combining factors:

- Just like a database join
- Get all factors over the joining variable
- Build a new factor over the union of the variables involved

Example: Join on R



$P(R)$

T	L
+r	0.1
-r	0.9

$\times$

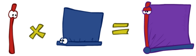
$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$\rightarrow$

R,T

Operation 1: Join Factors

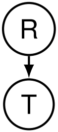


First basic operation: joining factors

Combining factors:

- Just like a database join
- Get all factors over the joining variable
- Build a new factor over the union of the variables involved

Example: Join on R



$P(R)$

T	L
+r	0.1
-r	0.9

$\times$

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$\rightarrow$

R,T

$P(R, T)$

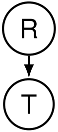
Operation 1: Join Factors

First basic operation: joining factors

Combining factors:

- Just like a database join
- Get all factors over the joining variable
- Build a new factor over the union of the variables involved

Example: Join on R



$P(R)$

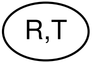
T	L
+r	0.1
-r	0.9

$\times$

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$\rightarrow$

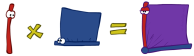


$P(R, T)$

+r	+t	0.08
+r	-t	0.02
-r	+t	0.09
-r	-t	0.81

- Computation for each entry: pointwise products

Operation 1: Join Factors

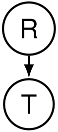


First basic operation: joining factors

Combining factors:

- Just like a database join
- Get all factors over the joining variable
- Build a new factor over the union of the variables involved

Example: Join on R



$P(R)$

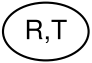
T	L
+r	0.1
-r	0.9

$\times$

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$\rightarrow$

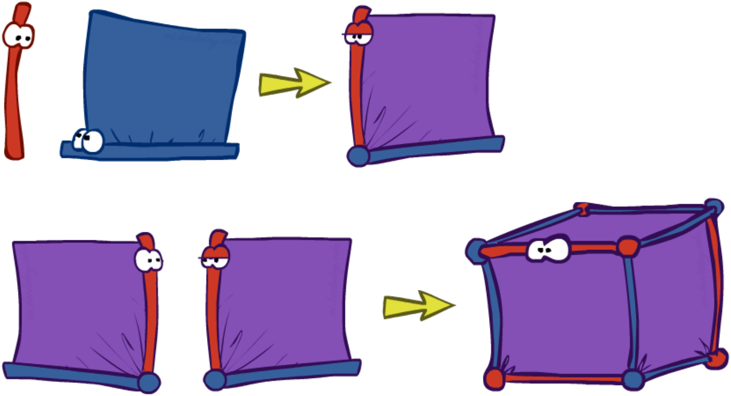


$P(R, T)$

+r	+t	0.08
+r	-t	0.02
-r	+t	0.09
-r	-t	0.81

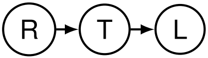
- Computation for each entry: pointwise products

Example: Multiple Joins

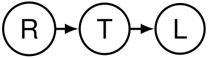


Example: Multiple Joins

Example: Multiple Joins



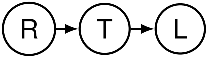
Example: Multiple Joins



$P(R)$

T	L
+r	0.1
-r	0.9

Example: Multiple Joins



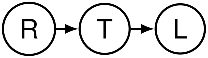
$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

Example: Multiple Joins



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

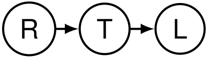
+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

Join R

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Example: Multiple Joins



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

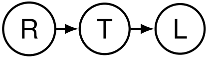
Join R

$\rightarrow$

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Example: Multiple Joins



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

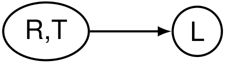
+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

Join R

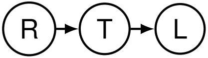
$\rightarrow$

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9



Example: Multiple Joins



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

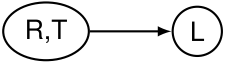
+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Join R

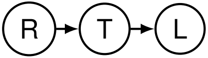
$\rightarrow$



$P(R, T)$

+r	+t	0.08
+r	-t	0.02
-r	+t	0.09
-r	-t	0.81

Example: Multiple Joins



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

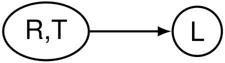
+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Join R

$\rightarrow$



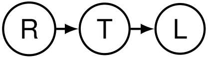
$P(R, T)$

+r	+t	0.08
+r	-t	0.02
-r	+t	0.09
-r	-t	0.81

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Example: Multiple Joins



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

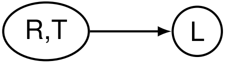
+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Join R

$\rightarrow$



$P(R,T)$

+r	+t	0.08
+r	-t	0.02
-r	+t	0.09
-r	-t	0.81

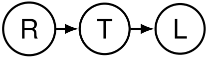
$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Join T.

$\rightarrow$

Example: Multiple Joins



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

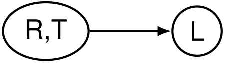
+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Join R

$\rightarrow$



$P(R,T)$

+r	+t	0.08
+r	-t	0.02
-r	+t	0.09
-r	-t	0.81

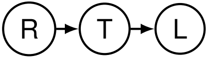
$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Join T.

$\rightarrow$

Example: Multiple Joins



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

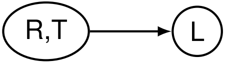
+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Join R

$\rightarrow$



$P(R,T)$

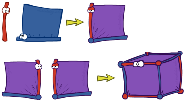
+r	+t	0.08
+r	-t	0.02
-r	+t	0.09
-r	-t	0.81

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Join T.

$\rightarrow$



Example: Multiple Joins

$R \rightarrow T \rightarrow L$

$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Join R $\rightarrow$

$R,T \rightarrow L$

$P(R, T)$

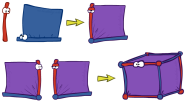
+r	+t	0.08
+r	-t	0.02
-r	+t	0.09
-r	-t	0.81

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

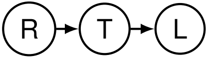
Join T. $\rightarrow$

R,T,L



$P(R, T, L)$

Example: Multiple Joins



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

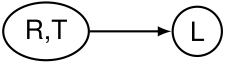
+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Join R

$\rightarrow$



$P(R,T)$

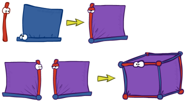
+r	+t	0.08
+r	-t	0.02
-r	+t	0.09
-r	-t	0.81

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Join T.

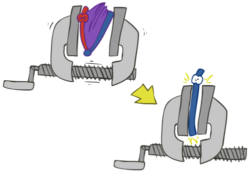
$\rightarrow$



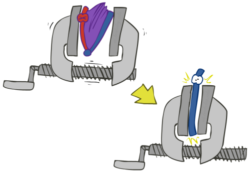
$P(R,T,L)$

+r	+t	+l	0.024
+r	+t	-l	0.056
+r	-t	+l	0.002
+r	-t	-l	0.018
-r	+t	+l	0.027
-r	+t	-l	0.063
-r	-t	+l	0.081
-r	-t	-l	0.729

Operation 2: Eliminate

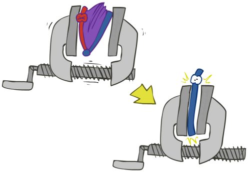


Operation 2: Eliminate



Second basic operation: marginalization

Operation 2: Eliminate

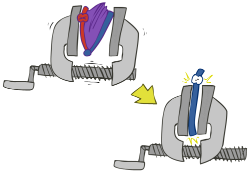


Second basic operation: marginalization

Take a factor and sum out a variable

- Shrinks a factor to a smaller one

Operation 2: Eliminate

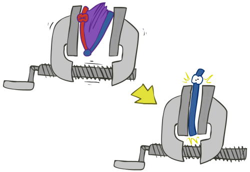


Second basic operation: marginalization

Take a factor and sum out a variable

- Shrinks a factor to a smaller one
- A projection operation

Operation 2: Eliminate

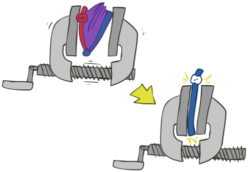


Second basic operation: marginalization

Take a factor and sum out a variable

- Shrinks a factor to a smaller one
- A projection operation

Operation 2: Eliminate



Second basic operation: marginalization

Take a factor and sum out a variable

- Shrinks a factor to a smaller one
- A projection operation

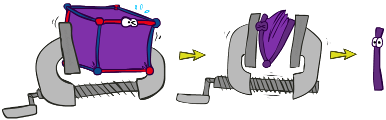
Example:

+r	+t	0.08
+r	-t	0.02
-r	+t	0.09
-r	-t	0.81

Sum R $\rightarrow$

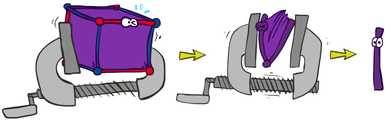
+t	0.17
-t	0.83

Multiple Elimination



R,T,L

Multiple Elimination



R,T,L

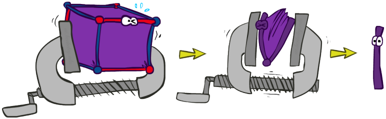
$P(R, T, L)$

+r	+t	+l	0.024
+r	+t	-l	0.056
+r	-t	+l	0.002
+r	-t	-l	0.018
-r	+t	+l	0.027
-r	+t	-l	0.063
-r	-t	+l	0.081
-r	-t	-l	0.729

Sum out R $\rightarrow$

T,L

Multiple Elimination



R,T,L

$P(R,T,L)$

+r	+t	+l	0.024
+r	+t	-l	0.056
+r	-t	+l	0.002
+r	-t	-l	0.018
-r	+t	+l	0.027
-r	+t	-l	0.063
-r	-t	+l	0.081
-r	-t	-l	0.729

Sum out R $\rightarrow$

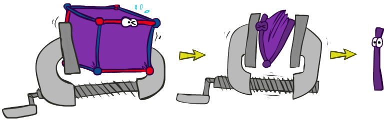
T,L

+t	+l	0.026
+t	-l	0.119
-t	+l	0.083
-t	-l	0.747

Sum out T $\rightarrow$

L

Multiple Elimination



R,T,L

$P(R,T,L)$

+r	+t	+l	0.024
+r	+t	-l	0.056
+r	-t	+l	0.002
+r	-t	-l	0.018
-r	+t	+l	0.027
-r	+t	-l	0.063
-r	-t	+l	0.081
-r	-t	-l	0.729

Sum out R $\rightarrow$

T,L

+t	+l	0.026
+t	-l	0.119
-t	+l	0.083
-t	-l	0.747

Sum out T $\rightarrow$

L

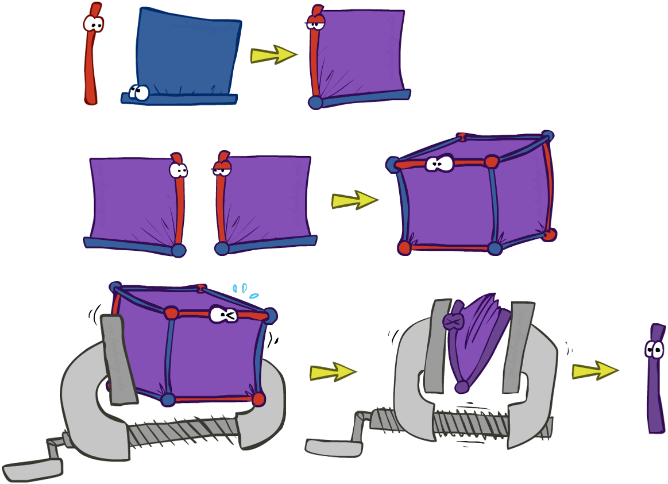
+l	0.134
-l	0.866

Thus Far: Multiple Join, Multiple Eliminate

Inference by Enumeration!

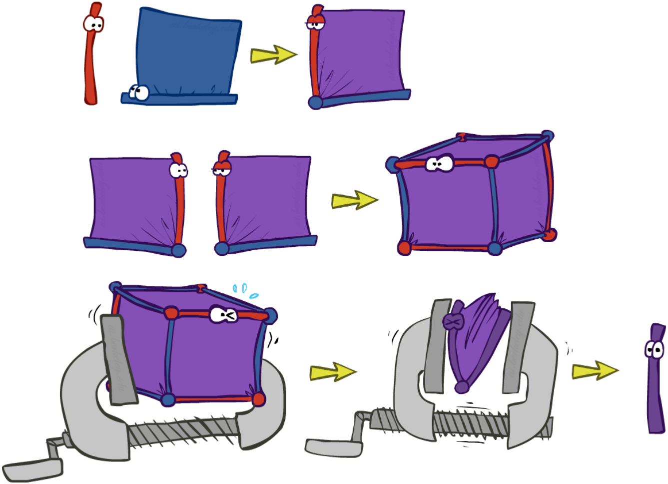
Thus Far: Multiple Join, Multiple Eliminate

Inference by Enumeration!



Thus Far: Multiple Join, Multiple Eliminate

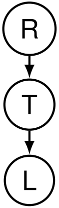
Inference by Enumeration!



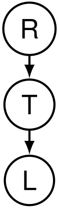
Example: Traffic Domain (revisited)

$P(R)$

T	L
+r	0.1
-r	0.9



Example: Traffic Domain (revisited)



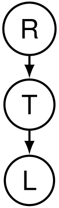
$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

Example: Traffic Domain (revisited)



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

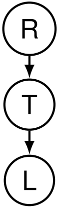
+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Random Variables

Example: Traffic Domain (revisited)



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

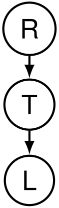
$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Random Variables

- R: Raining

Example: Traffic Domain (revisited)



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

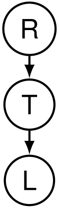
$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Random Variables

- R: Raining
- T: Traffic

Example: Traffic Domain (revisited)



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

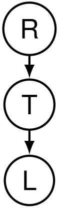
$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Random Variables

- R: Raining
- T: Traffic
- L: Late for class!

Example: Traffic Domain (revisited)



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

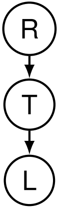
+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Random Variables

- R: Raining
- T: Traffic
- L: Late for class!

$P(L)$

Example: Traffic Domain (revisited)



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

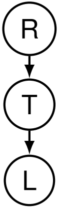
+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

$P(L) = \sum_{r,t} P(r,t,L)$

Random Variables

- R: Raining
- T: Traffic
- L: Late for class!

Example: Traffic Domain (revisited)



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

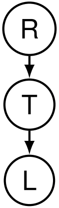
+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

Random Variables

- R: Raining
- T: Traffic
- L: Late for class!

$P(L) = \sum_{r,t} P(r,t,L) = \sum_{r,t} P(r)P(t|r)P(L|t).$

Example: Traffic Domain (revisited)



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

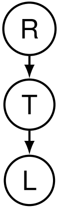
Random Variables

- R: Raining
- T: Traffic
- L: Late for class!

$$P(L) = \sum_{r,t} P(r,t,L) = \sum_{r,t} P(r)P(t|r)P(L|t).$$

or

Example: Traffic Domain (revisited)



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

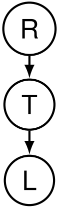
Random Variables

- R: Raining
- T: Traffic
- L: Late for class!

$P(L) = \sum_{r,t} P(r,t,L) = \sum_{r,t} P(r)P(t|r)P(L|t)$.

or $P(L) = \sum_t P(t)P(L|t)$ and $P(T) = \sum_r P(r)P(T|R)$.

Example: Traffic Domain (revisited)



$P(R)$

T	L
+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9

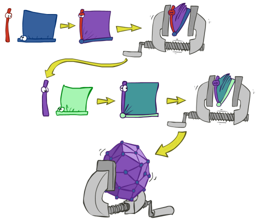
Random Variables

- R: Raining
- T: Traffic
- L: Late for class!

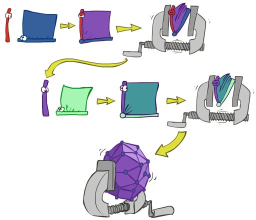
$P(L) = \sum_{r,t} P(r,t,L) = \sum_{r,t} P(r)P(t|r)P(L|t)$.

or $P(L) = \sum_t P(t)P(L|t)$ and $P(T) = \sum_r P(r)P(T|R)$.

Inference by Enumeration vs. Variable Elimination

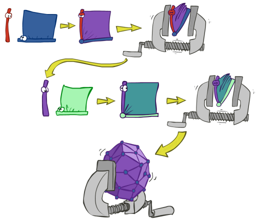


Inference by Enumeration vs. Variable Elimination



Why is inference by enumeration so slow?

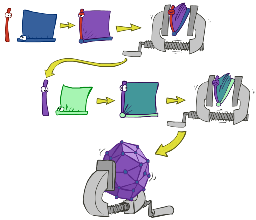
Inference by Enumeration vs. Variable Elimination



Why is inference by enumeration so slow?

- Join whole joint distribution before sum out the hidden variables

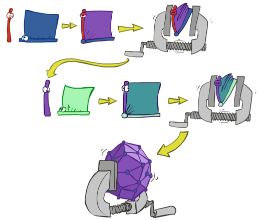
Inference by Enumeration vs. Variable Elimination



Why is inference by enumeration so slow?

- Join whole joint distribution before sum out the hidden variables

Inference by Enumeration vs. Variable Elimination

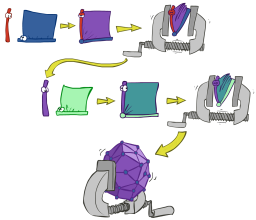


Why is inference by enumeration so slow?

- Join whole joint distribution before sum out the hidden variables

Idea: interleave joining and marginalizing!

Inference by Enumeration vs. Variable Elimination



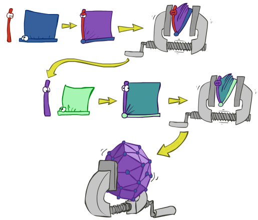
Why is inference by enumeration so slow?

- Join whole joint distribution before sum out the hidden variables

Idea: interleave joining and marginalizing!

- Called “Variable Elimination”

Inference by Enumeration vs. Variable Elimination



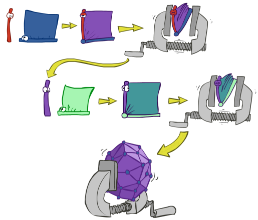
Why is inference by enumeration so slow?

- Join whole joint distribution before sum out the hidden variables

Idea: interleave joining and marginalizing!

- Called “Variable Elimination”
- Still NP-hard, but usually much faster than inference by enumeration

Inference by Enumeration vs. Variable Elimination



Why is inference by enumeration so slow?

- Join whole joint distribution before sum out the hidden variables

Idea: interleave joining and marginalizing!

- Called “Variable Elimination”
- Still NP-hard, but usually much faster than inference by enumeration