
CIS 471/571 (Fall 2020):
Introduction to Artificial Intelligence

Assignment Project Exam Help

Lecture 14: Bayes Nets – Independence

<https://tutorcs.com>

WeChat: cstutorcs

Thanh H. Nguyen

Source: <http://ai.berkeley.edu/home.html>



Announcement

- Homework 4: Bayes Nets and HMMs
 - Will be posted today (Nov 12, 2020)
 - Deadline: Nov 24, 2020

Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

Probability Recap

- Conditional probability $P(x|y) = \frac{P(x, y)}{P(y)}$
- Product rule $P(x, y) = P(x|y)P(y)$
- Chain rule $P(X_1, X_2, \dots, X_n) = P(X_1)P(X_2|X_1)P(X_3|X_1, X_2) \dots \prod_{i=1}^n P(X_i|X_1, \dots, X_{i-1})$
- X, Y independent if and only if: $\forall x, y : P(x, y) = P(x)P(y)$
- X and Y are conditionally independent given Z if and only if:

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



Bayes' Nets

- A Bayes' net is an efficient encoding of a probabilistic model of a domain

- Questions we can ask:

- Inference: given a fixed BN, what is $P(X \mid e)$?
- Representation: given a BN graph, what kinds of distributions can it encode?
- Modeling: what BN is most appropriate for a given domain?



Bayes' Net Semantics

- A directed, acyclic graph, one node per random variable
- A conditional probability table (CPT) for each node

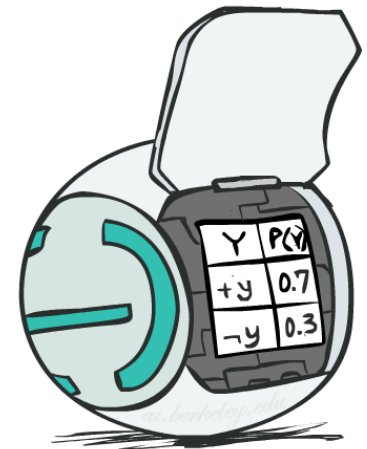
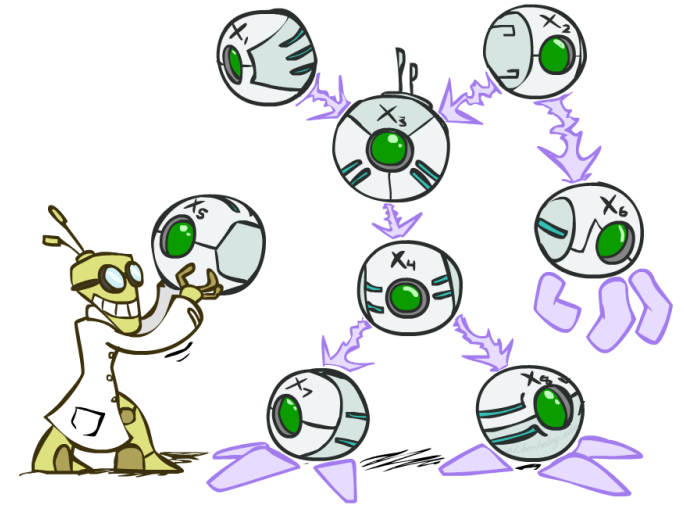
- A collection of distributions over X , one for each combination of parents' values

$$P(X|a_1 \dots a_n)$$

WeChat: cstutorcs

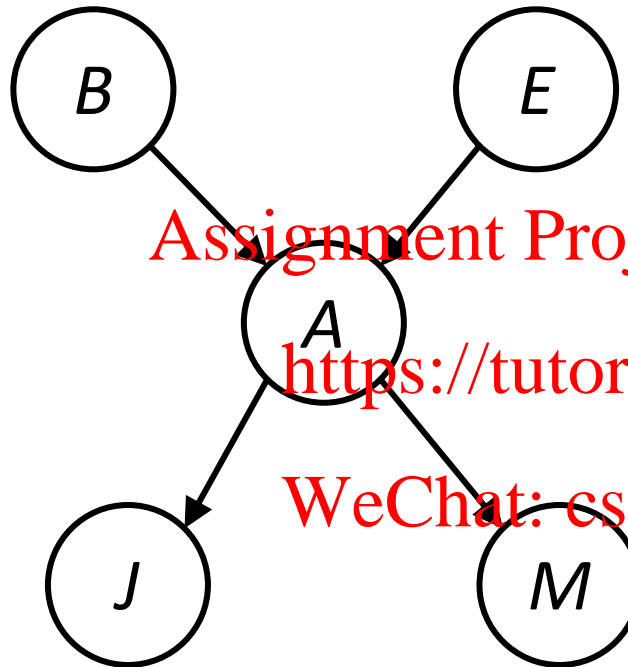
- Bayes' nets implicitly encode joint distributions
 - As a product of local conditional distributions
 - To see what probability a BN gives to a full assignment, multiply all the relevant conditionals together:

$$P(x_1, x_2, \dots, x_n) = \prod_{i=1}^n P(x_i | \text{parents}(X_i))$$



Example: Alarm Network

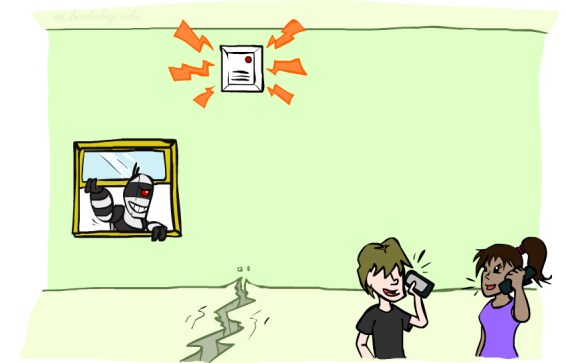
B	P(B)
+b	0.001
-b	0.999



E	P(E)
+e	0.002
-e	0.998

A	M	P(M A)
+a	+m	0.7
+a	-m	0.3
-a	+m	0.01
-a	-m	0.99

A	J	P(J A)
+a	+j	0.9
+a	-j	0.1
-a	+j	0.05
-a	-j	0.95



B	E	A	P(A B,E)
+b	+e	+a	0.95
+b	+e	-a	0.05
+b	-e	+a	0.94
+b	-e	-a	0.06
-b	+e	+a	0.29
-b	+e	-a	0.71
-b	-e	+a	0.001
-b	-e	-a	0.999

$$P(+b, -e, +a, -j, +m) =$$

Assignment Project Exam Help

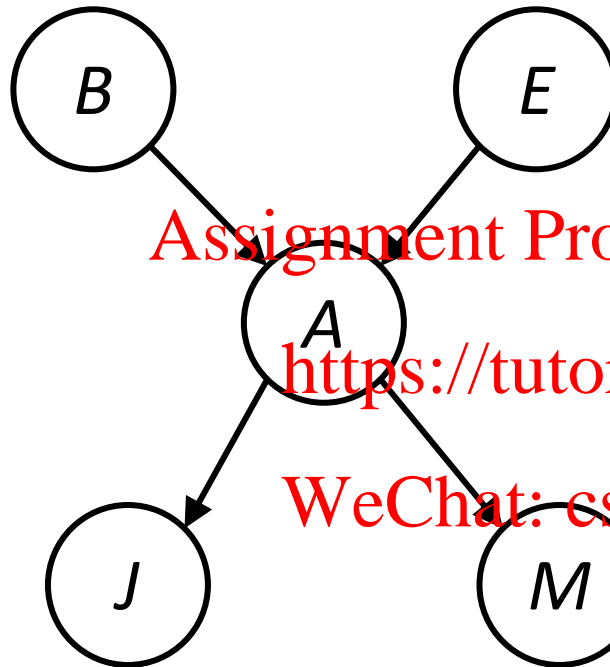
<https://tutorcs.com>

WeChat: cstutorcs



Example: Alarm Network

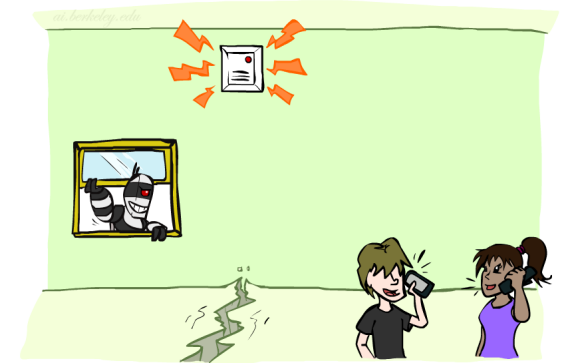
B	P(B)
+b	0.001
-b	0.999



E	P(E)
+e	0.002
-e	0.998

A	J	P(J A)
+a	+j	0.9
+a	-j	0.1
-a	+j	0.05
-a	-j	0.95

A	M	P(M A)
+a	+m	0.7
+a	-m	0.3
-a	+m	0.01
-a	-m	0.99



B	E	A	P(A B,E)
+b	+e	+a	0.95
+b	+e	-a	0.05
+b	-e	+a	0.94
+b	-e	-a	0.06
-b	+e	+a	0.29
-b	+e	-a	0.71
-b	-e	+a	0.001
-b	-e	-a	0.999

$$\begin{aligned}
 P(+b, -e, +a, -j, +m) &= \\
 P(+b)P(-e)P(+a|+b, -e)P(-j|+a)P(+m|+a) &= \\
 0.001 \times 0.998 \times 0.94 \times 0.1 \times 0.7 &=
 \end{aligned}$$



Size of a Bayes' Net

- How big is a joint distribution over N Boolean variables?

$$2^N$$

- Both give you the power to calculate

$$P(X_1, X_2, \dots, X_n)$$

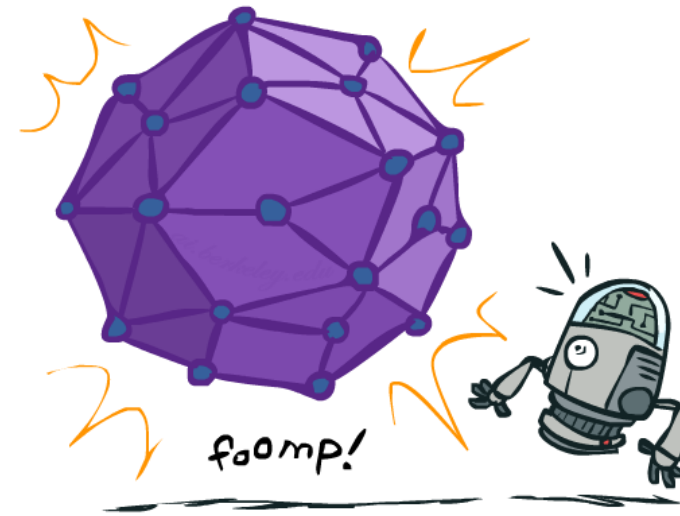
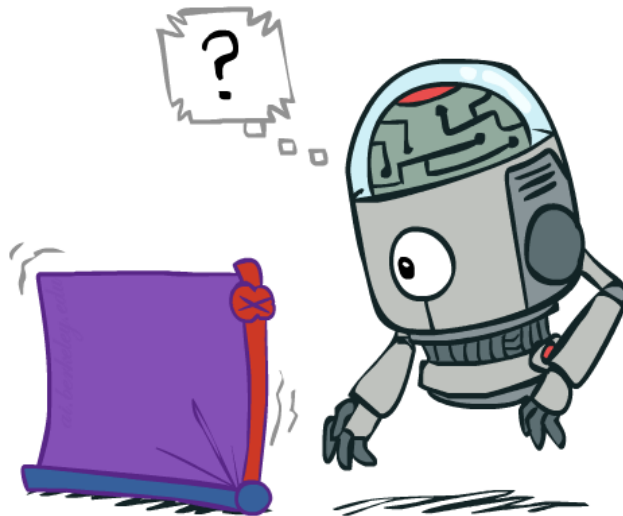
Assignment Projects Exam Help
Also, Huge space savings!

- How big is an N -node net if nodes have up to k parents?

$$O(N * 2^{k+1})$$

Also, easier to elicit local CPTs
<https://tutorcs.com>

WeChat: cstutorcs Also, faster to answer queries (coming)



Bayes' Nets

✓ Representation

- Conditional Independences
<https://tutorcs.com>
- Probabilistic Inference
WeChat: cstutorcs
- Learning Bayes' Nets from Data



Conditional Independence

- X and Y are **independent** if

$$\forall x, y \quad P(x, y) = P(x)P(y) \quad \text{---} \rightarrow \quad X \perp\!\!\!\perp Y$$

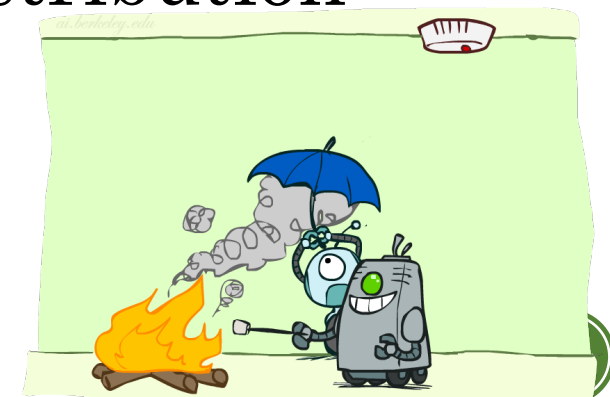
Assignment Project Exam Help

- X and Y are **conditionally independent** given Z

$$\forall x, y, z \quad P(x, y|z) = P(x|z)P(y|z) \quad \text{---} \rightarrow \quad X \perp\!\!\!\perp Y|Z$$

- (Conditional) independence is a property of a distribution

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



Bayes Nets: Assumptions

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

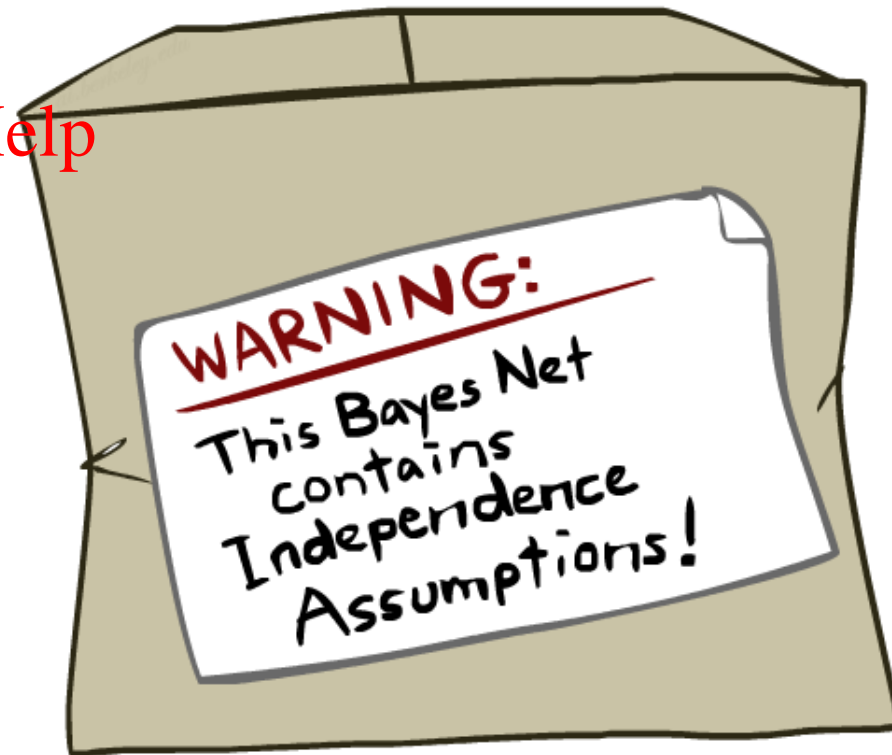
$$P(x_i | x_1 \cdots x_{i-1}) = P(x_i | \text{parents}(X_i))$$

Assignment Project Exam Help

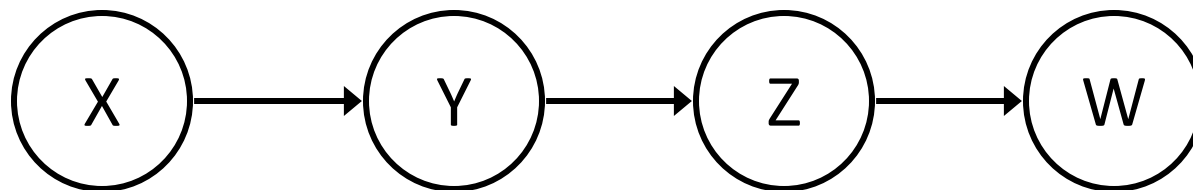
- Beyond above “chain rule \rightarrow Bayes net” conditional independence assumptions <https://tutorcs.com>

- Often additional conditional independence assumptions
- They can be read off the graph

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



Example



- Conditional independence assumptions directly from simplifications in chain rule:

<https://tutorcs.com>

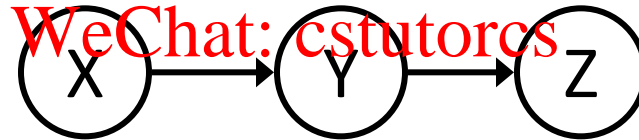
WeChat: cstutorcs

- Additional implied conditional independence assumptions?



Independence in a BN

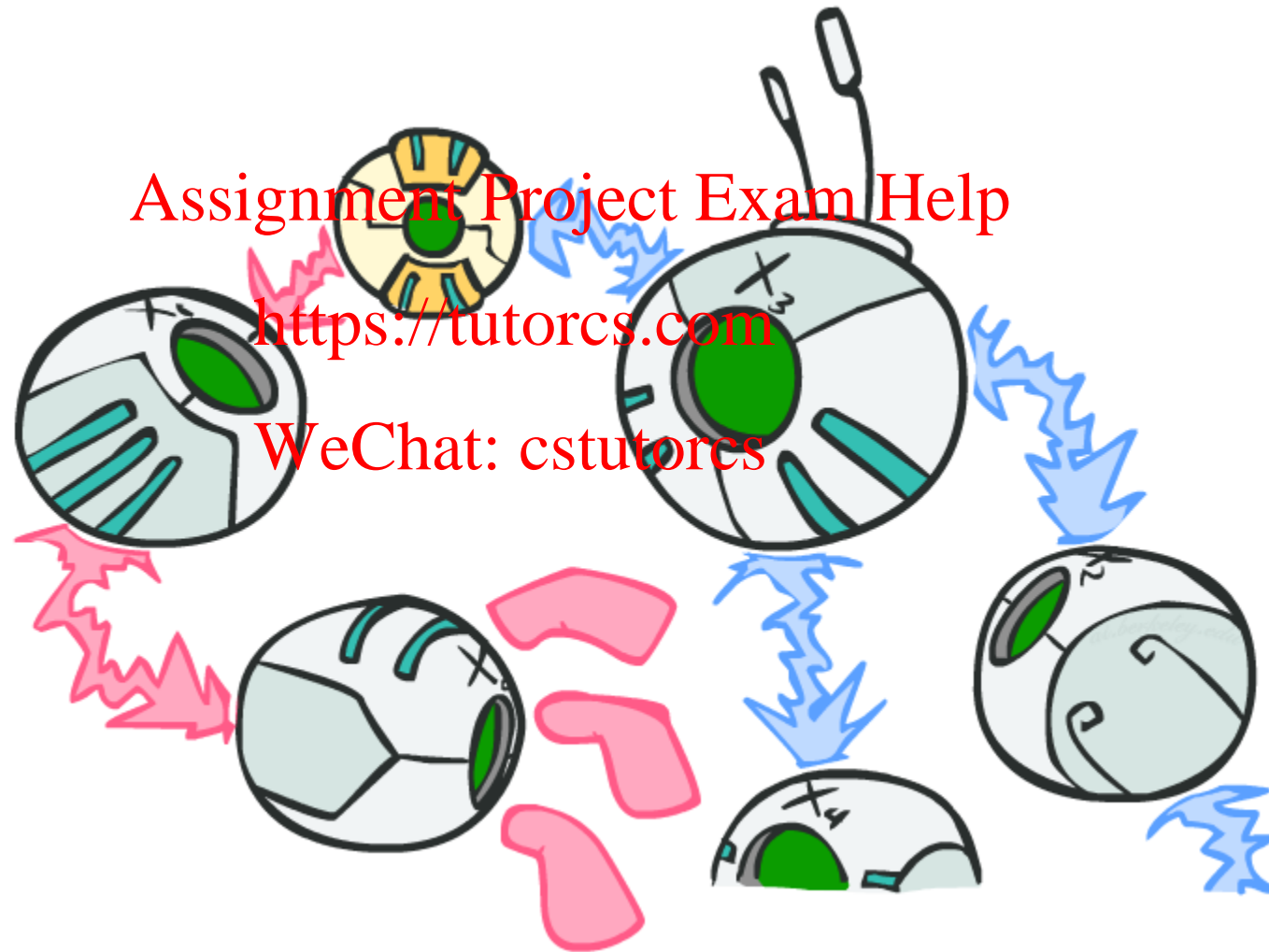
- Important question about a BN:
 - Are two nodes independent given certain evidence?
 - If yes, can prove using algebra (tedious in general)
 - If no, can prove with a counter example
 - Example: <https://tutorcs.com>



- 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?



D-separation: Outline



D-separation: Outline

- Study independence properties for triples

Assignment Project Exam Help

- Analyze complex cases in terms of member triples

<https://tutorcs.com>

WeChat: cstutorcs

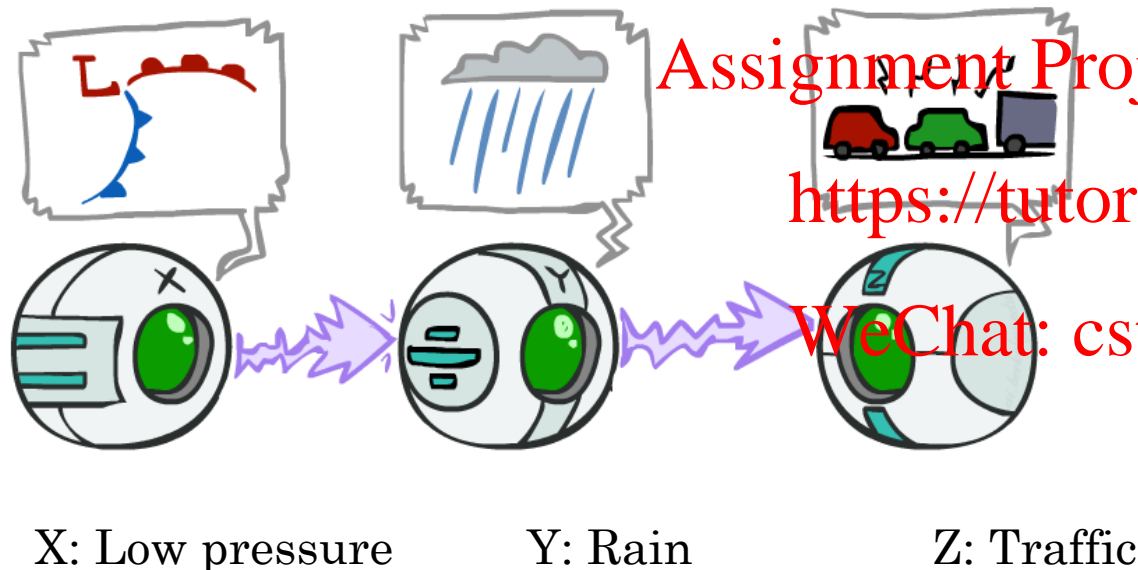
- D-separation: a condition / algorithm for answering such queries



Causal Chains

- This configuration is a “causal chain”

- Guaranteed X independent of Z ? *No!*



Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

- 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 causes rain causes traffic,
high pressure causes no rain causes no traffic

- In numbers:

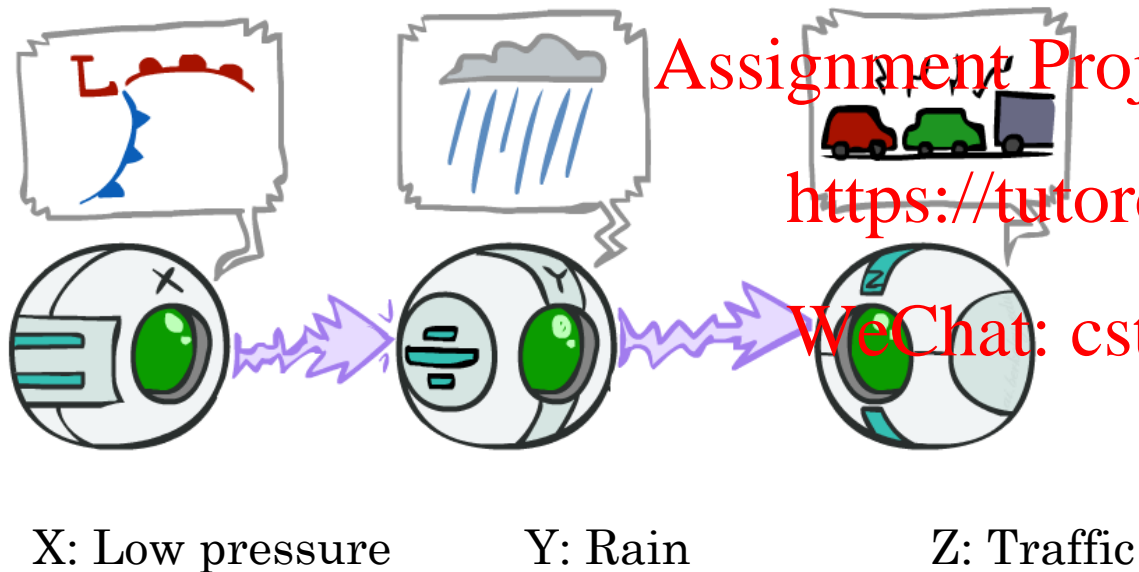
$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

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



Causal Chains

- This configuration is a “causal chain”
- Guaranteed X independent of Z given Y?



Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

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

Yes!

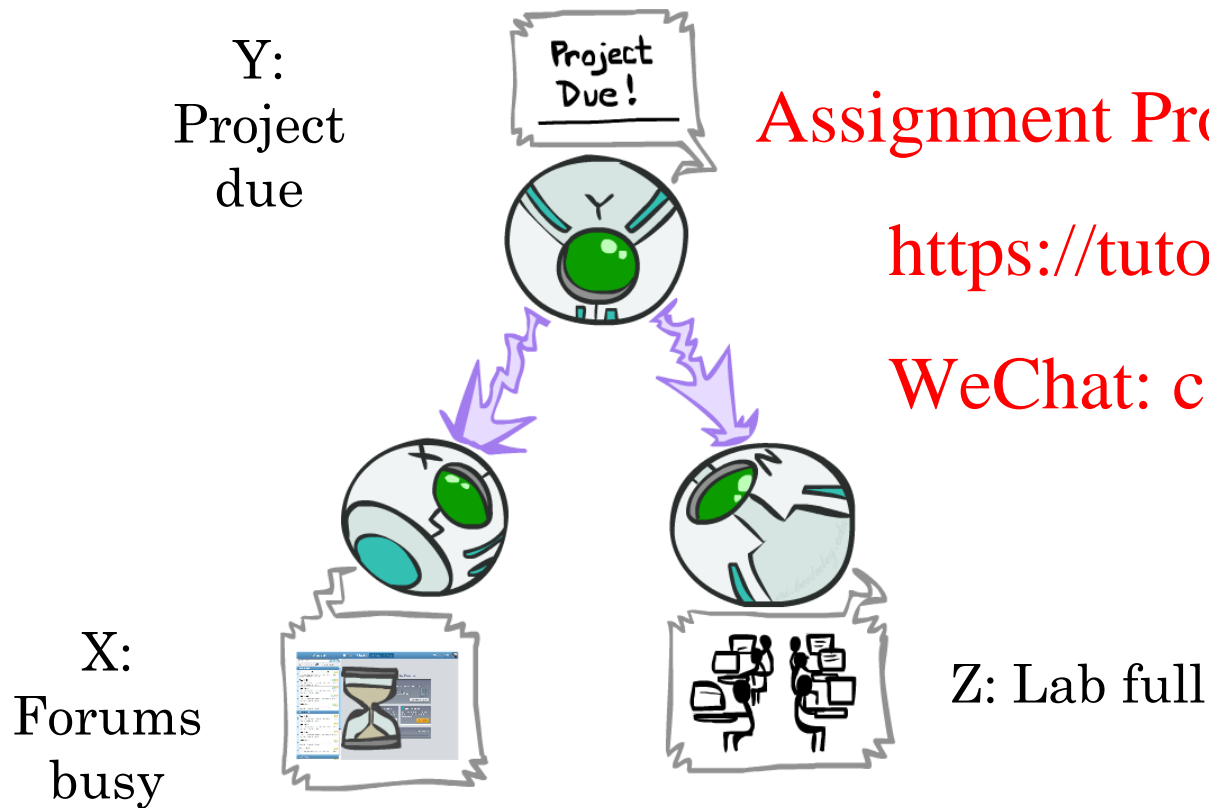
$$P(x,y,z) = P(x)P(y|x)P(z|y)$$

- Evidence along the chain “blocks” the influence



Common Cause

- This configuration is a “common cause”
- Guaranteed X independent of Z ? **No!**



Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

- 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 causes both forums busy and lab full

- In numbers:

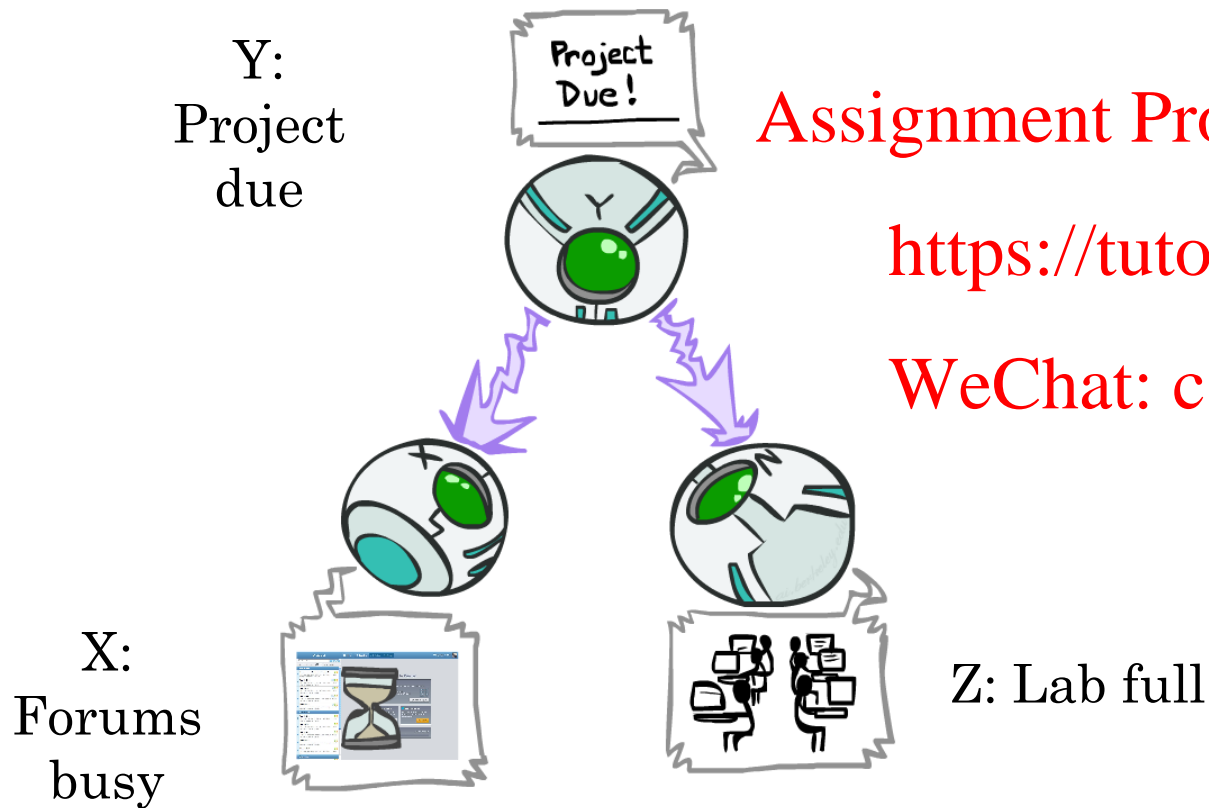
$$P(+x \mid +y) = 1, P(-x \mid -y) = 1, \\ P(+z \mid +y) = 1, P(-z \mid -y) = 1$$

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



Common Cause

- This configuration is a “common cause”
- Guaranteed X and Z independent given Y?



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

Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

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

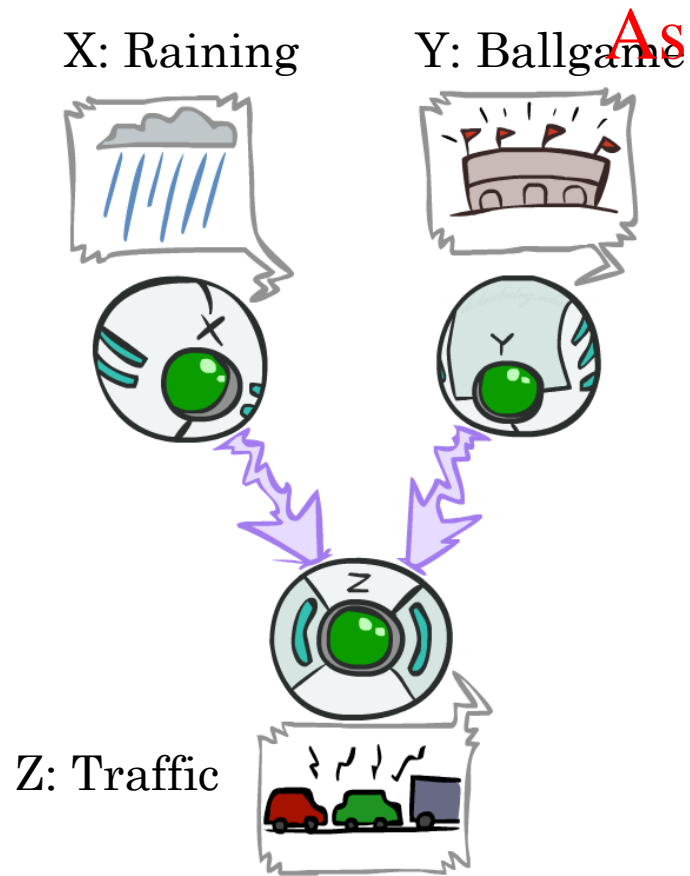
Yes!

- Observing the cause blocks influence between effects.



Common Effect

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



- 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!)

<https://tutorcs.com>

WeChat: cstutorcs

- 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

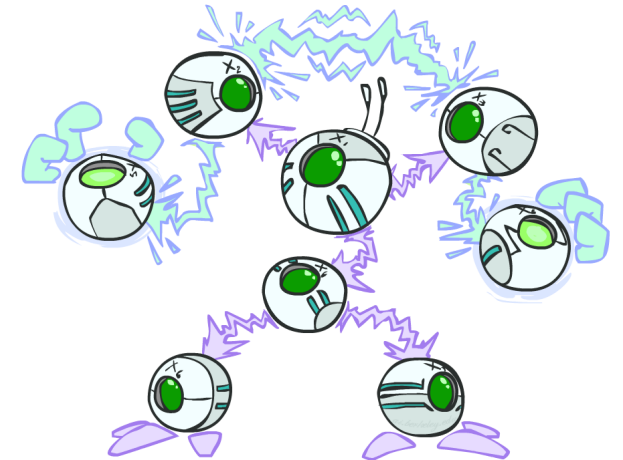
- General question: in a given BN, are two variables independent (given evidence)?

Assignment Project Exam Help

- Solution: analyze the graph <https://tutorcs.com>

WeChat: cstutorcs

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



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!

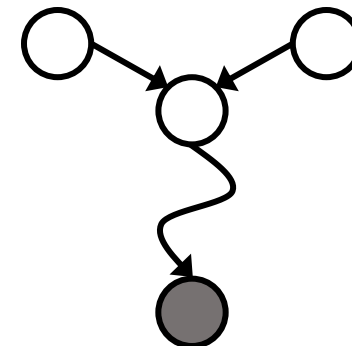
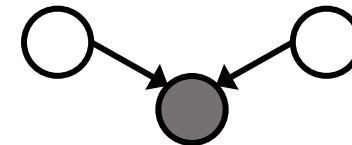
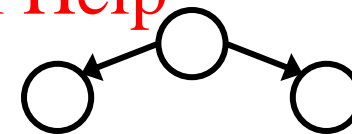
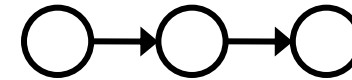
<https://tutorcs.com>

- A path is active if each triple is active.

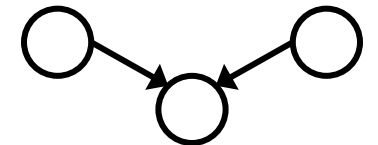
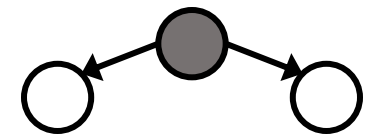
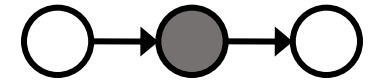
- Causal chain $A \rightarrow B \rightarrow C$ where B is unobserved (either direction)
- Common cause $A \leftarrow B \rightarrow C$ where B is unobserved
- Common effect (aka v-structure)
 $A \rightarrow B \leftarrow C$ where B or one of its descendants is observed

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

Active Triples



Inactive Triples

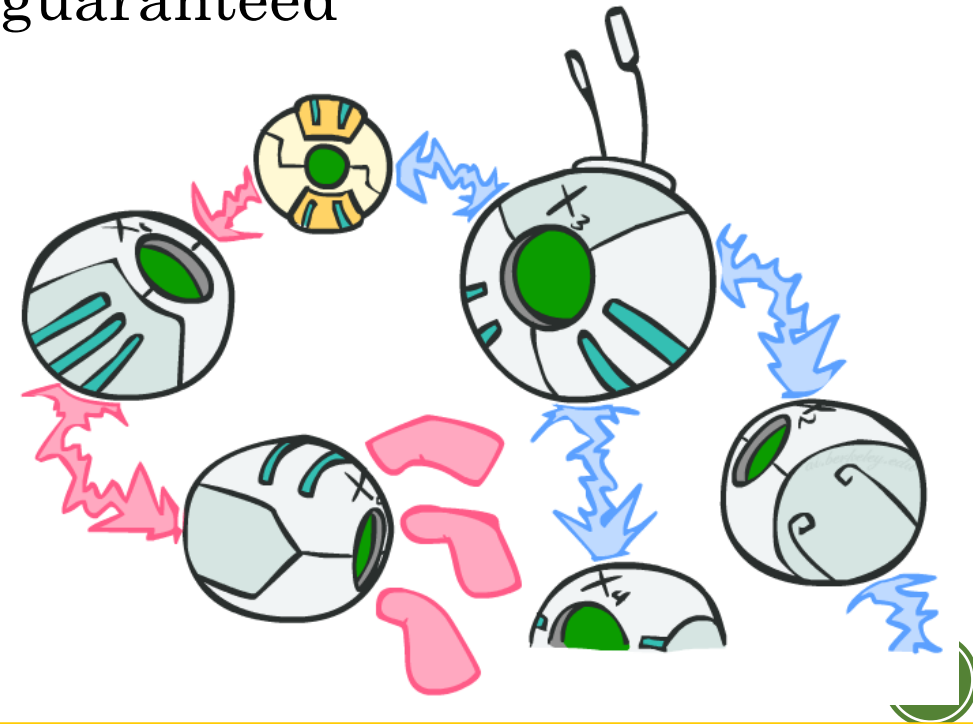


D-Separation

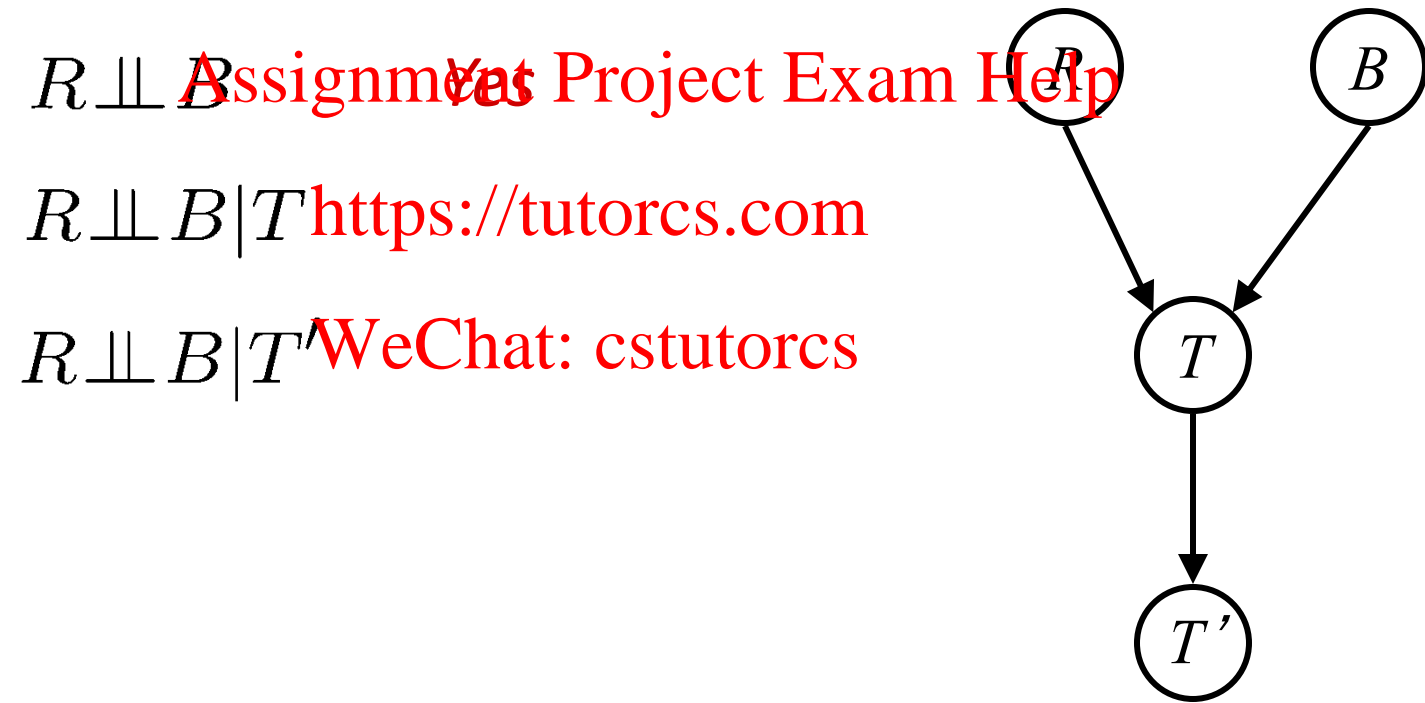
- Query: $X_i \perp\!\!\!\perp X_j | \{X_{k_1}, \dots, X_{k_n}\} ?$
- Check all (undirected) paths between X_i and X_j
 - If one or more active, then independence not guaranteed
- Otherwise (i.e. if all paths are inactive), then independence is guaranteed

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

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



Example



Example

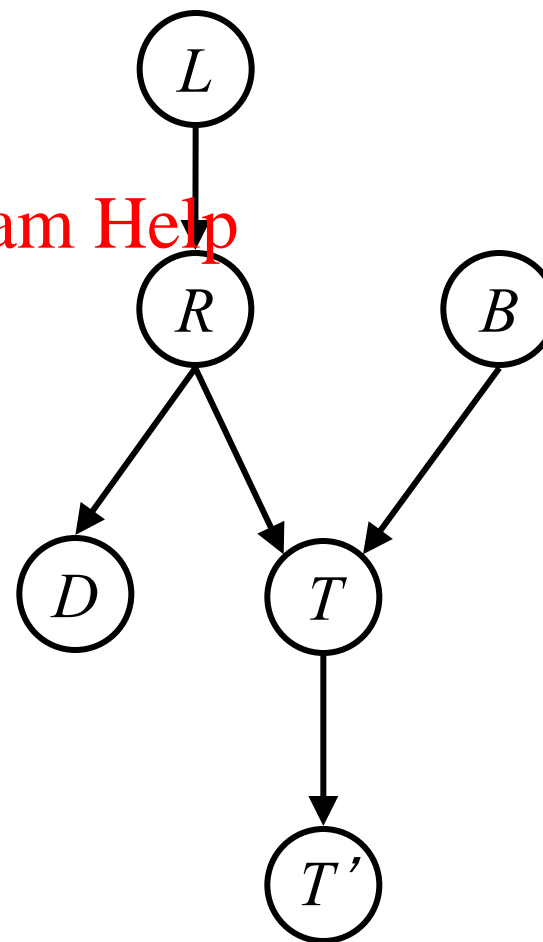
$L \perp\!\!\!\perp T' \mid T$ Yes

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

$L \perp\!\!\!\perp B \mid T$ Yes

$L \perp\!\!\!\perp B \mid T'$ No

$L \perp\!\!\!\perp B \mid T, R$ Yes



Example

- Variables:

- R: Raining

- T: Traffic

- D: Roof drips

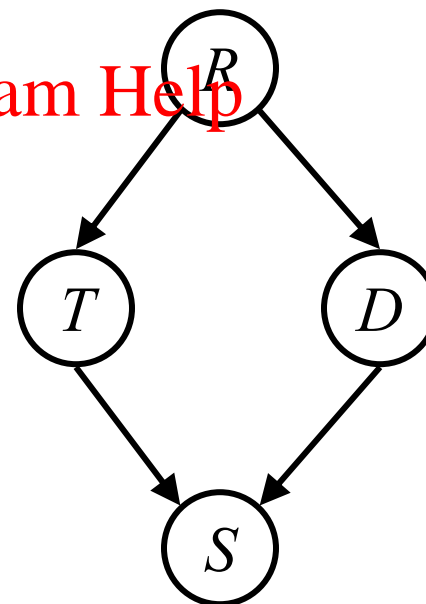
- S: I'm sad

- Questions:

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

$$T \perp\!\!\!\perp D | R \quad \text{Yes}$$

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

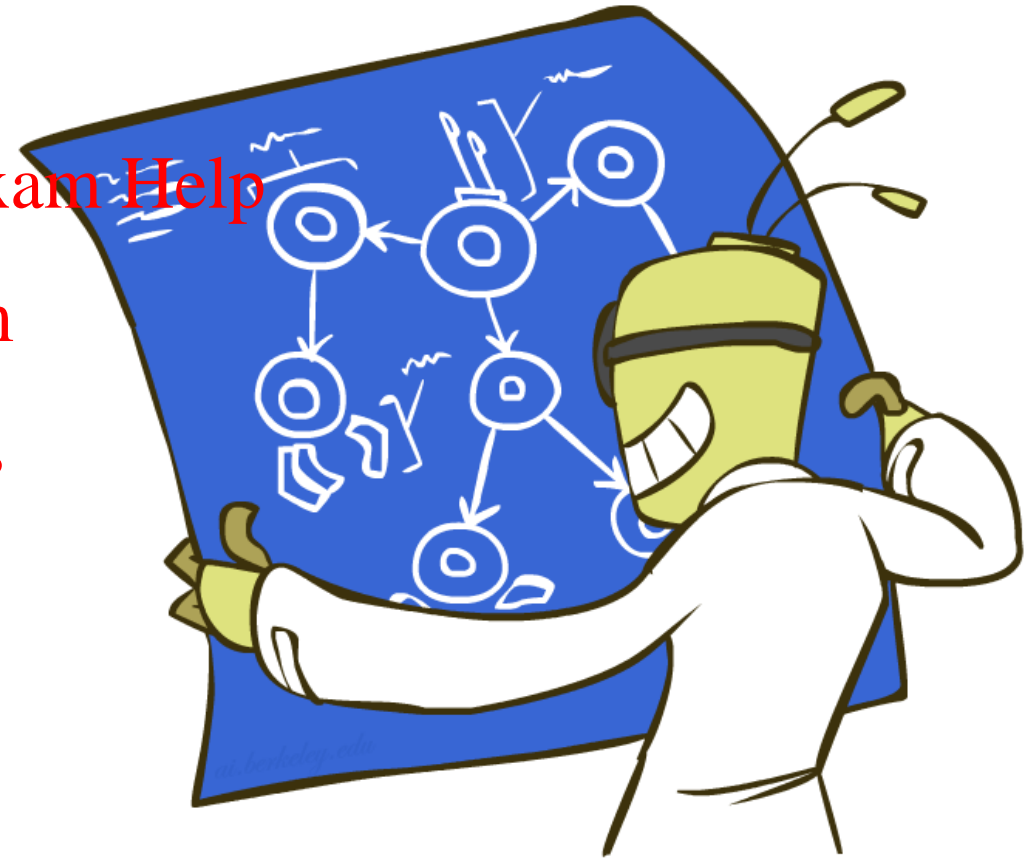


Structure Implications

- 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

$$X_i \perp\!\!\!\perp X_j \mid \{X_{k_1}, \dots, X_{k_n}\}$$

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



Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs



Computing All Independences

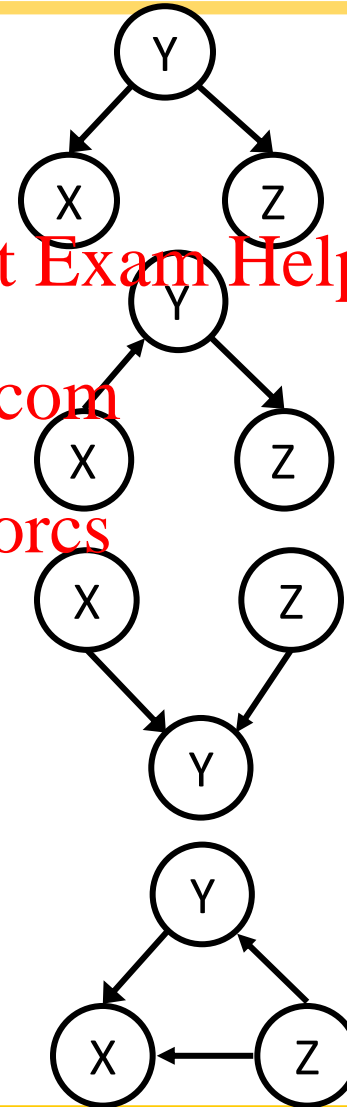
COMPUTE ALL THE
INDEPENDENCES!



Assignment Project Exam Help

<https://tutorcs.com>

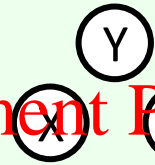
WeChat: cstutorcs



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

$\{X \perp\!\!\!\perp Y, X \perp\!\!\!\perp Z, Y \perp\!\!\!\perp Z,$
 $X \perp\!\!\!\perp Z \mid Y, X \perp\!\!\!\perp Y \mid Z, Y \perp\!\!\!\perp Z \mid X\}$

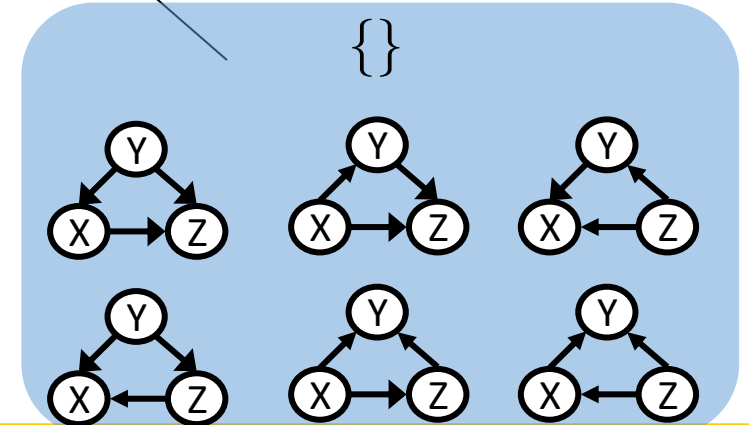
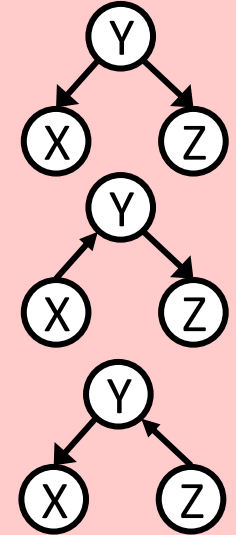


Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

$\{X \perp\!\!\!\perp Z \mid Y\}$



Bayes Nets Representation Summary

- Bayes nets compactly encode joint distributions
- Guaranteed independence of distributions can be deduced from BN graph structure
<https://tutorcs.com>
- 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

✓ 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

Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

