复旦大学大数据学院
School of Data Science, Fudan University
魏忠钰

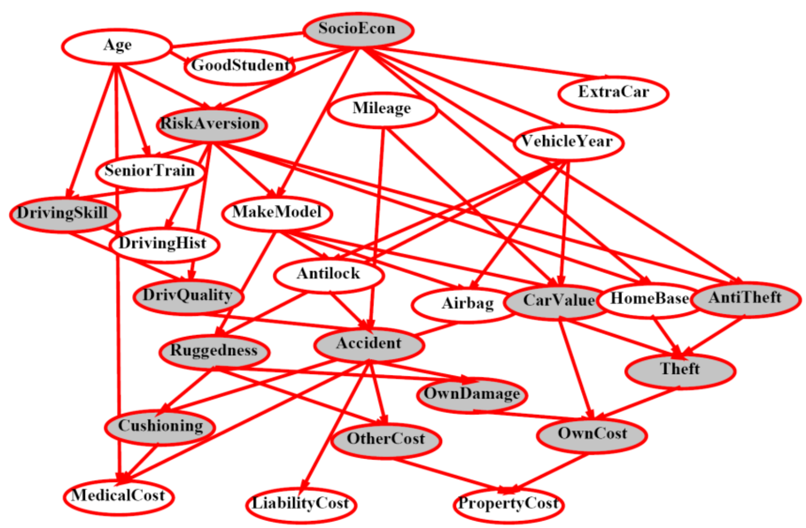
Bayes' Nets: Independence

May 23rd, 2018

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 | e)?
  - Representation: given a BN graph, what kinds of distributions can it encode?
  - Modeling: what BN is most appropriate for a given domain?



Conditional Independence

- X and Y are independent if

$$\forall x, y \;\; P(x, y) = P(x)P(y) \dashrightarrow \;\; X \perp\!\!\!\perp Y$$

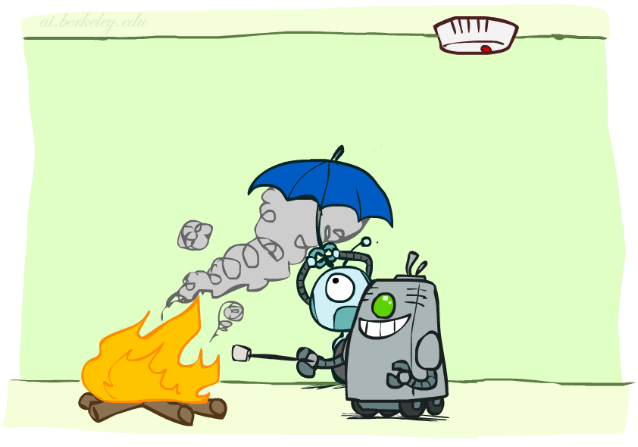
- X and Y are conditionally independent given Z

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

- (Conditional) independence is a property of a distribution

- Example:

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

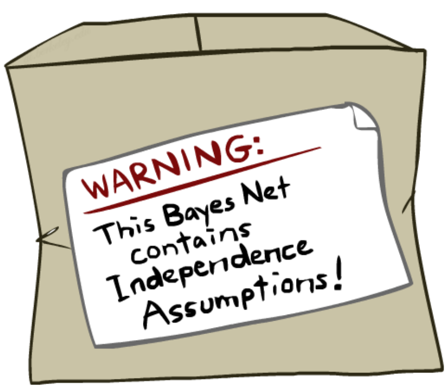


Bayes Nets: Assumptions

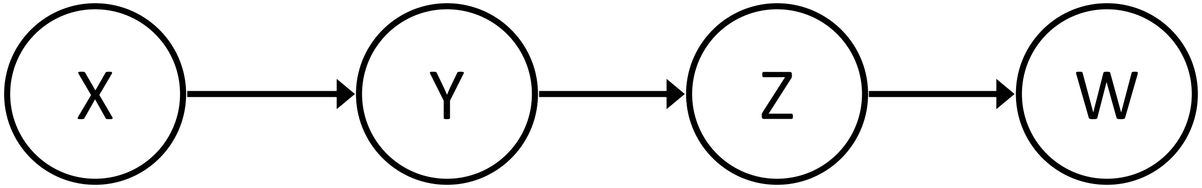
- Assumptions we are required to make to define the Bayes net when given the graph:

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

- Beyond above "chain rule → 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

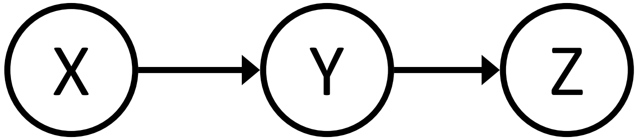


- Conditional independence assumptions directly from simplifications in chain rule:

- Additional implied conditional independence assumptions?

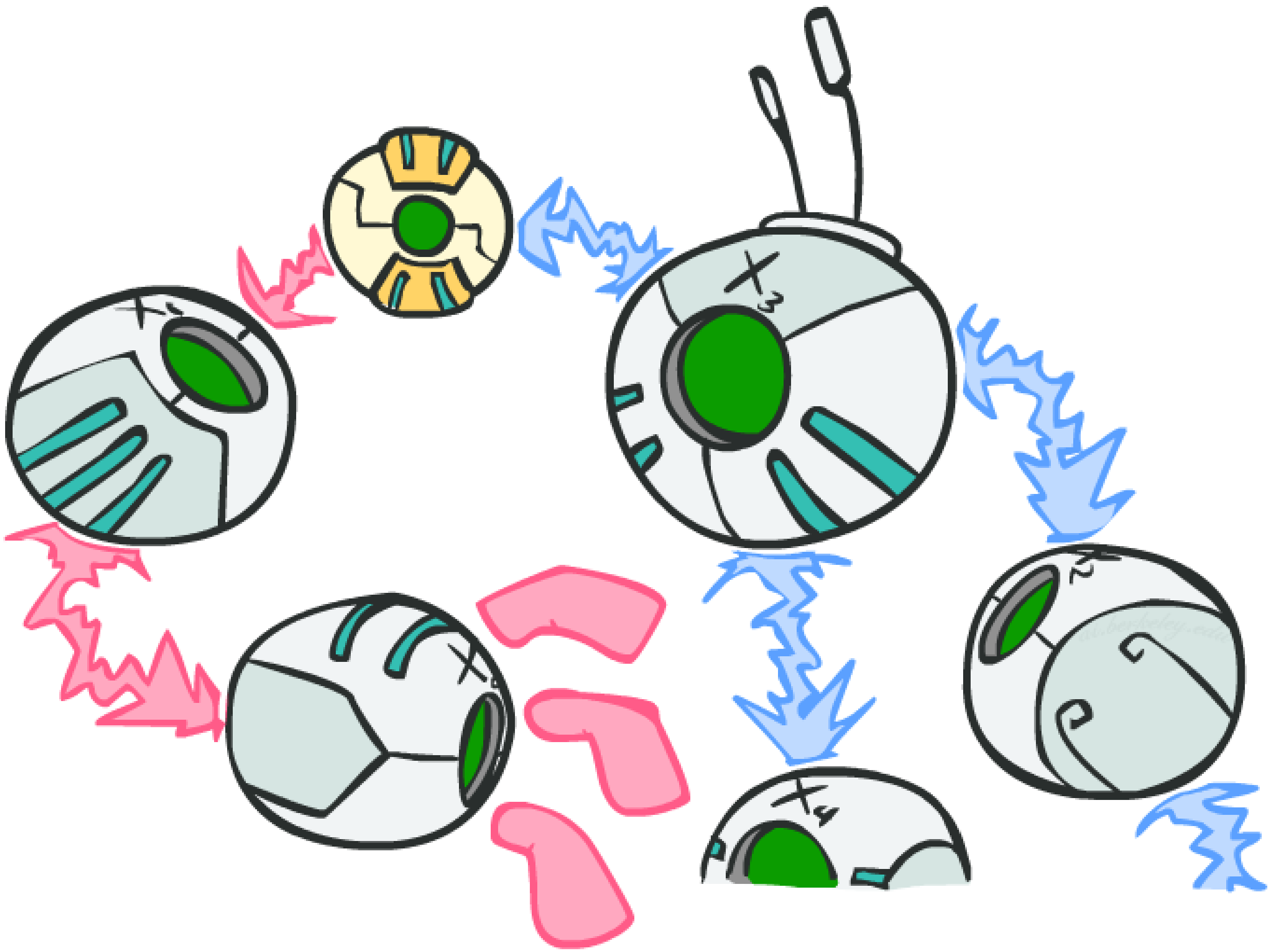
Independence in a BN

- 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:



  - 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)
    - They *could* be independent: how?

D-separation: Outline



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"

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 causes rain causes traffic, high pressure causes no rain causes 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 Given Y, X independent of Z ?

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

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

$$= P(z|y)$$

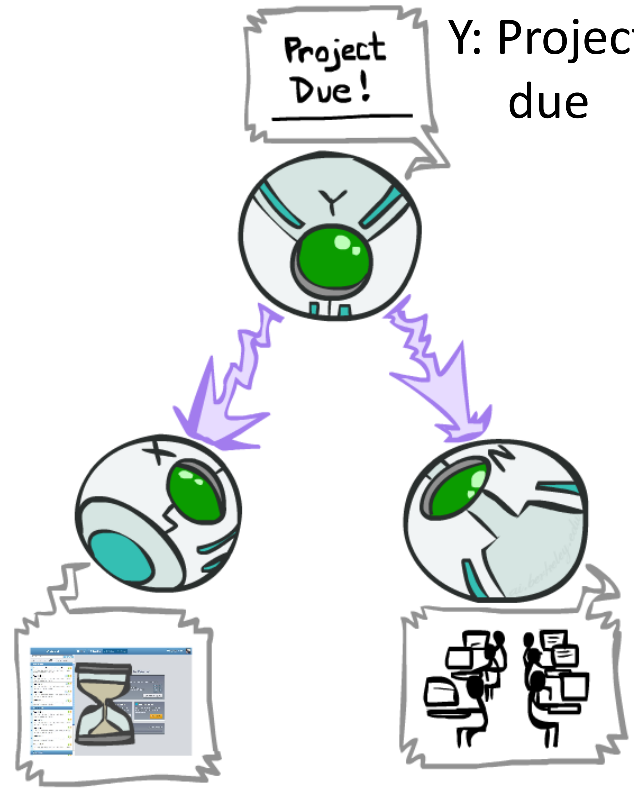
Yes!

- Evidence along the chain "blocks" the influence

Common Cause

- 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 causes both Email box busy and library full
    - In numbers

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



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

Common Cause

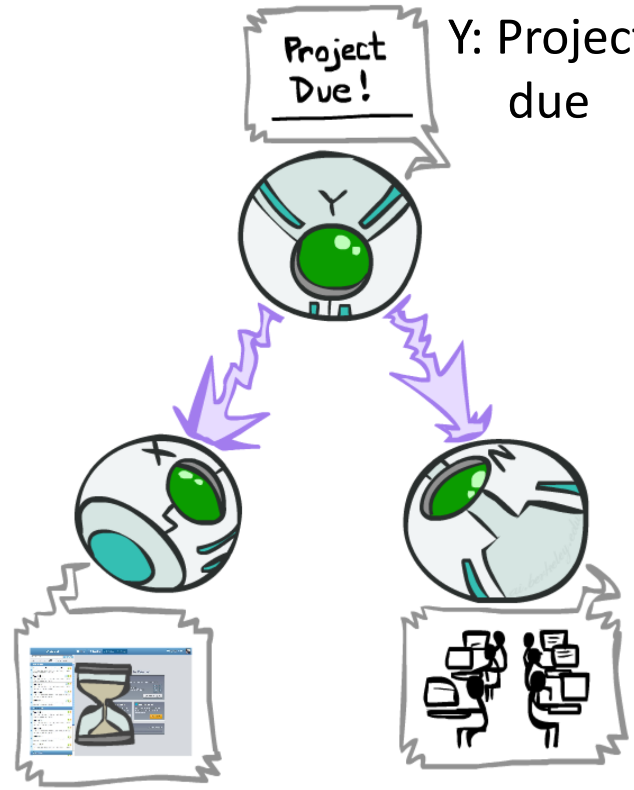
- Guaranteed X and Z independent given Y ?

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

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

$$= P(z|y)$$

Yes!



Y: Project due

X: Email Box busy

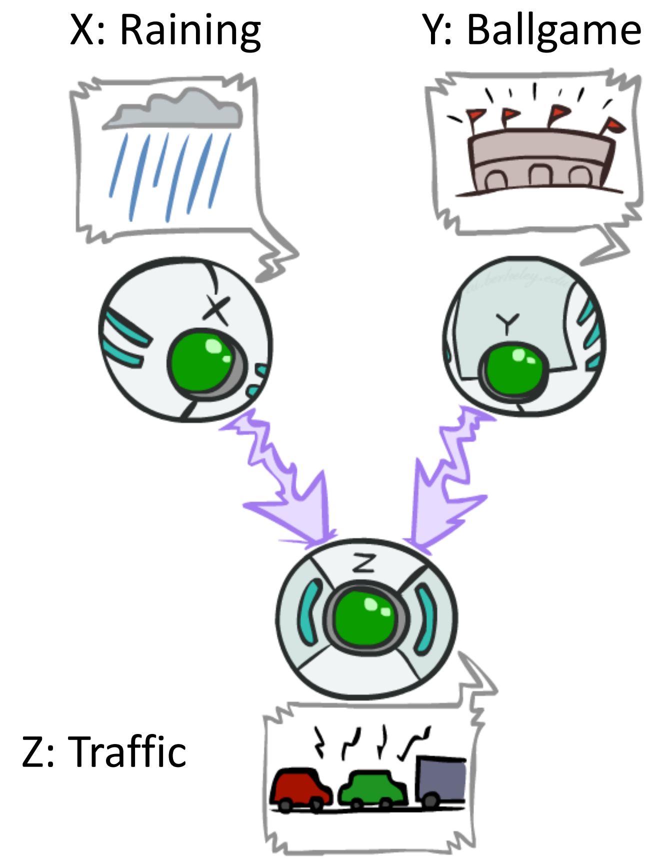
Z: Library full

Observing the cause blocks influence between effects.

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

Common Effect

- 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
- 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.

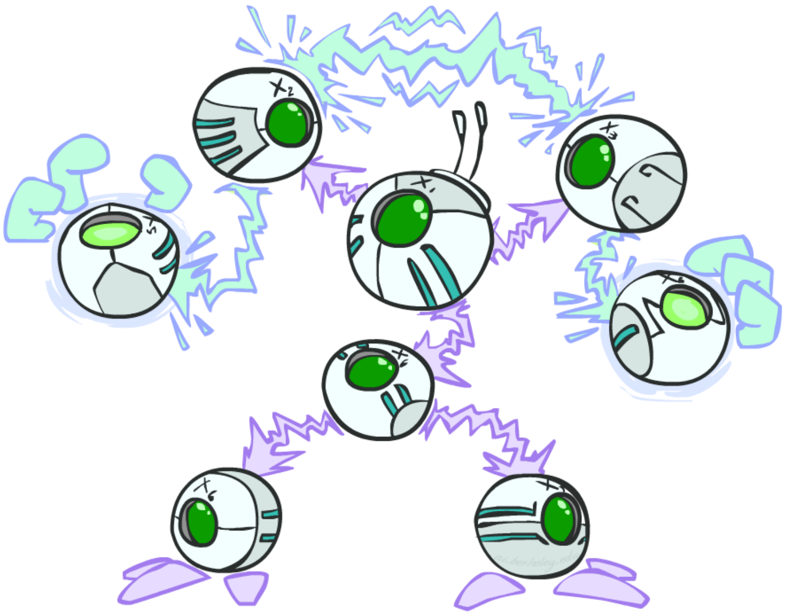


The General Case

The General Case

- General question: in a given BN, are two variables independent (given evidence)?
- Solution: analyze the graph
- Any complex example can be broken

  into repetitions of the three canonical cases

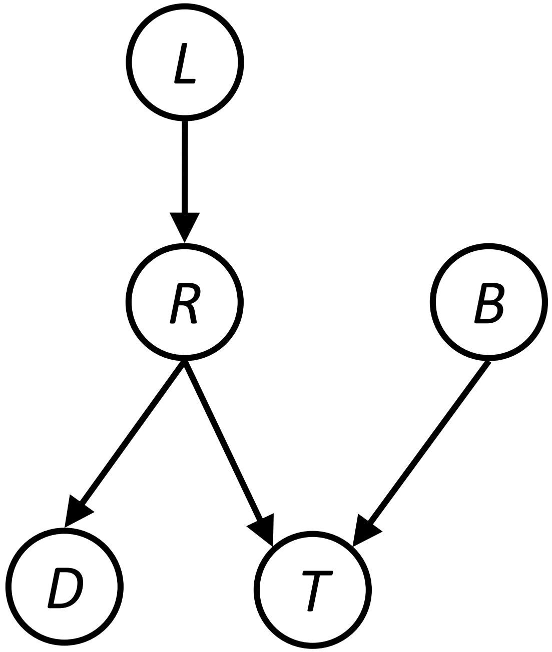


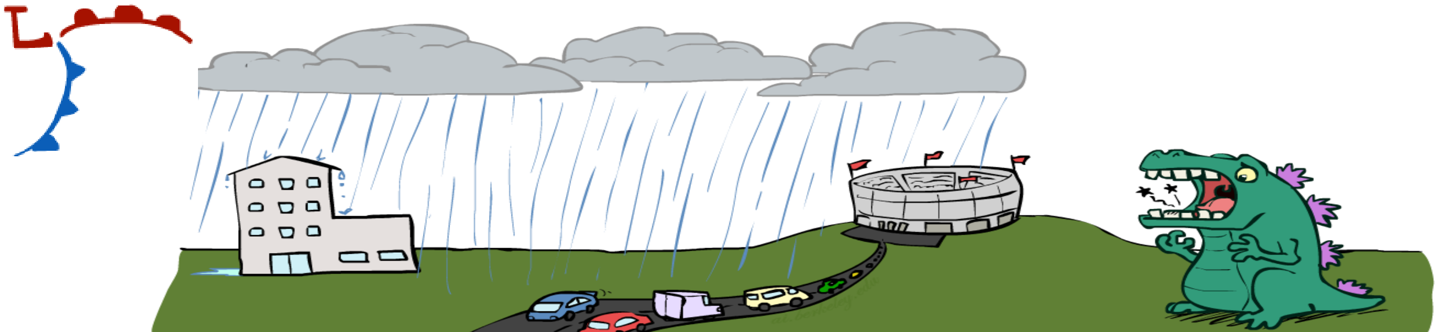
Reachability

- 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"



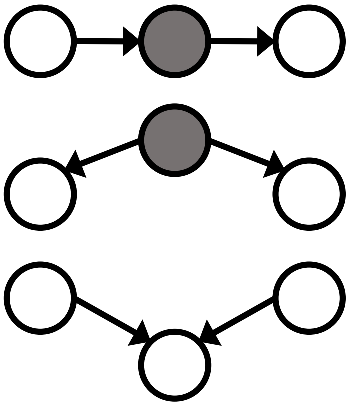


Active / Inactive Paths

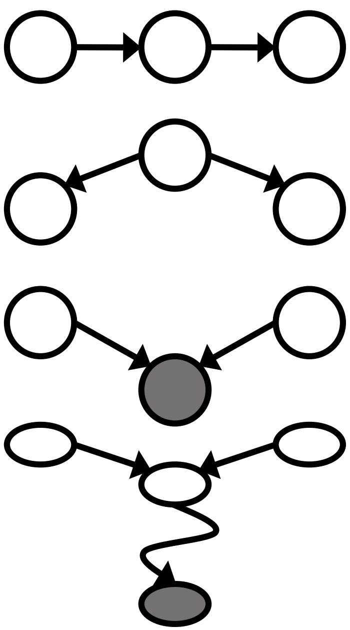
- Question: Are X and Y conditionally independent given evidence variables {Z}?
  - Yes, if X and Y "d-separated" by Z
  - Consider all (undirected) paths from X to Y
  - No active paths = independence!
- A path is active if every triple is active:
  - Causal chain A → B → C where B is unobserved (either direction)
  - Common cause A ← B → C where B is unobserved
  - Common effect (aka v-structure)

  A → B ← C where B *or one of its descendents* is observed
- All it takes to block a path is a single inactive segment

Inactive Triples



Active Triples



D-Separation

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

- Check all (undirected!) paths between X_i and X_j
  - If one or more active, then **independence not guaranteed**

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

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

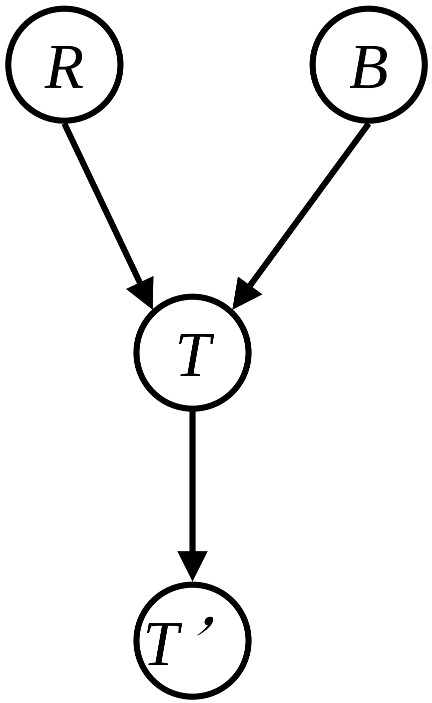
    $$X_i \perp\!\!\!\perp X_j | \{X_{k_1}, ..., X_{k_n}\}$$

Example

$R \perp\!\!\!\perp B$ *Yes*

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

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



Example

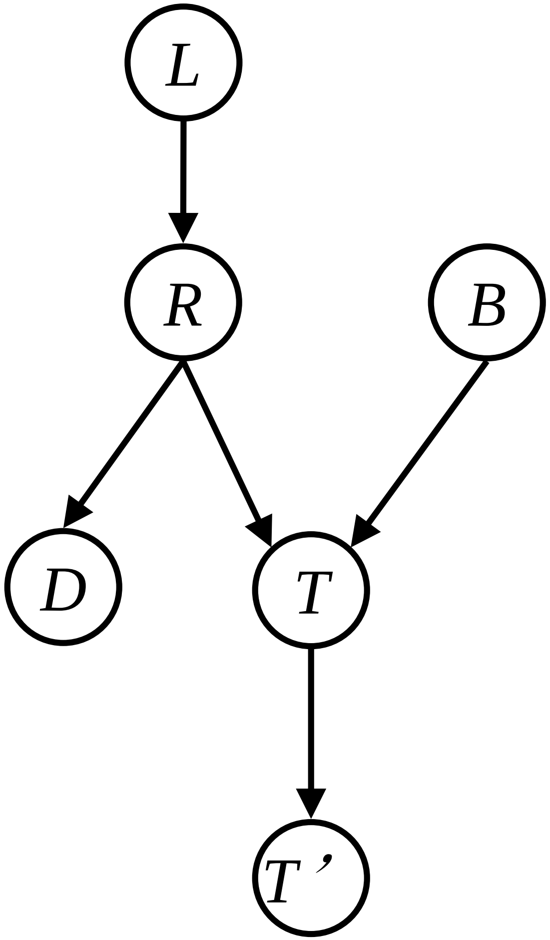
$L \perp\!\!\!\perp T' | T$ *Yes*

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

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

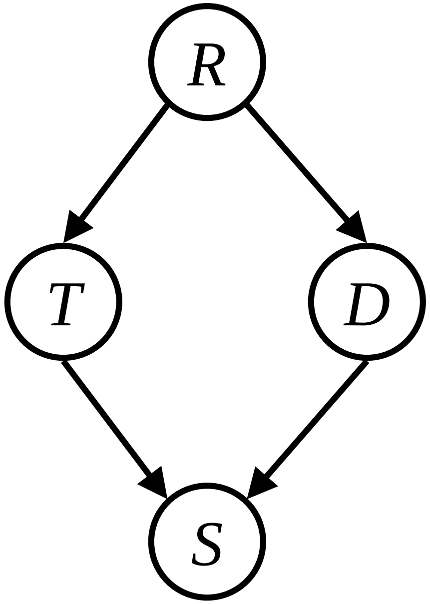
$L \perp\!\!\!\perp B | T'$

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



Example

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



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

$$T \perp\!\!\!\perp D | R$$ *Yes*

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

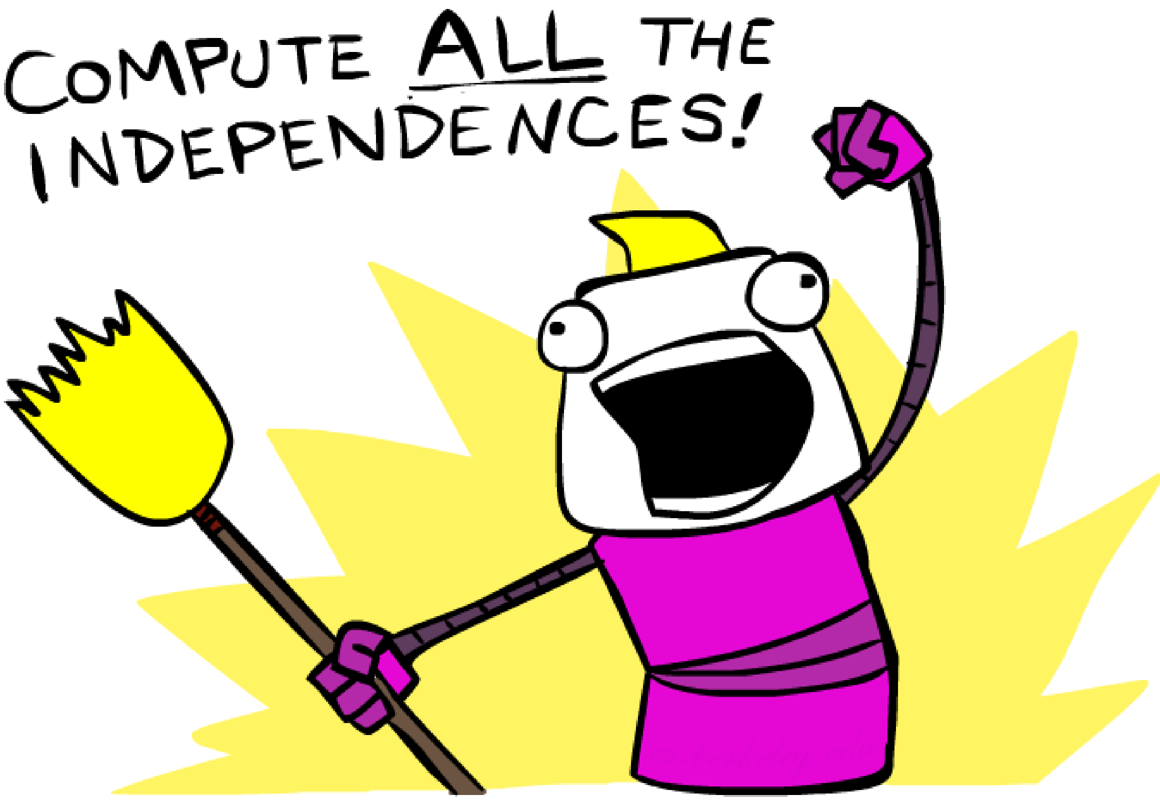
Structure Implications

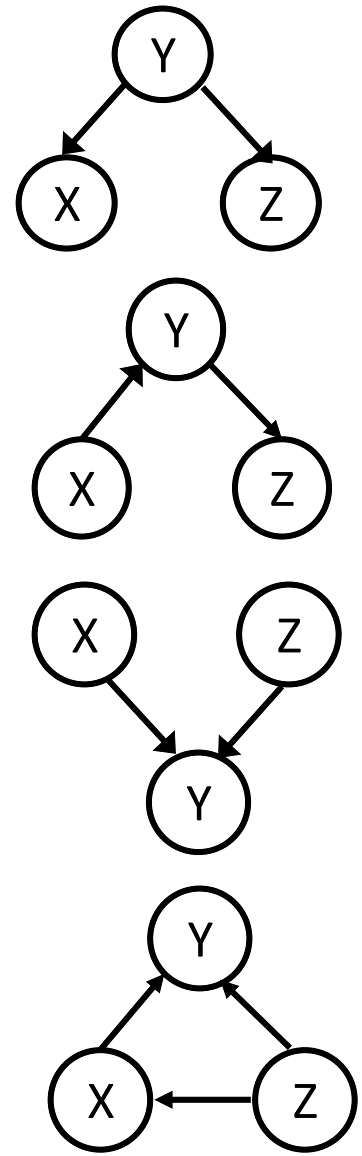
- Given a Bayes net structure, can run d-separation algorithm to build a complete list of conditional independences that are necessarily true of the form

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

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

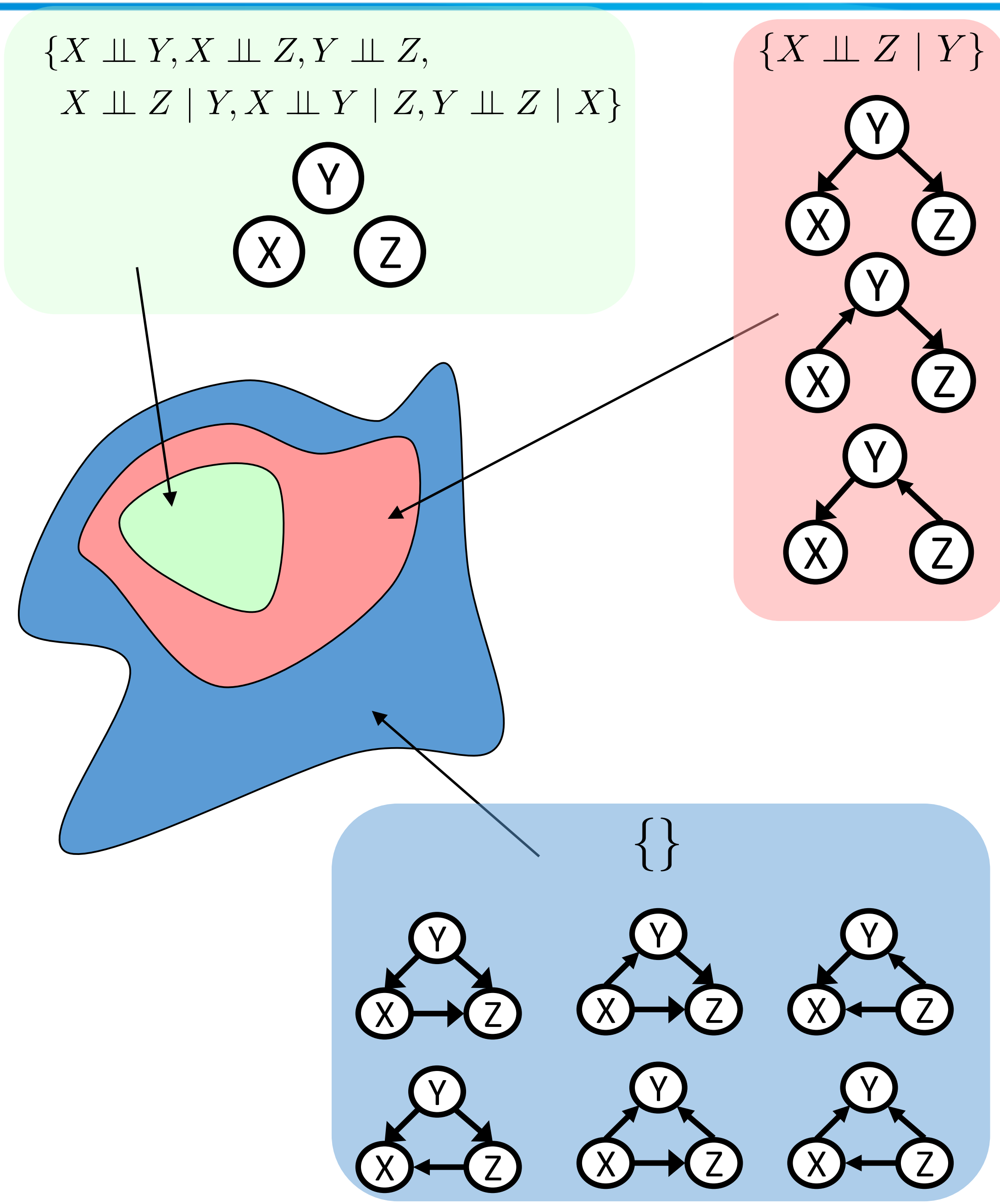
Computing All Independences





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)
- Adding arcs increases the set of distributions, but has several costs
- Full conditioning can encode any distribution



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

- 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