Bayesian Networks: Independence

Forest Agostinelli
University of South Carolina

Topics Covered in This Class

- **Part 1: Search**
  - Pathfinding
    - Uninformed search
    - Informed search
  - Adversarial search
  - Optimization
    - Local search
    - Constraint satisfaction
- **Part 2: Knowledge Representation and Reasoning**
  - Propositional logic
  - First-order logic
  - Prolog
- **Part 3: Knowledge Representation and Reasoning Under Uncertainty**
  - Probability
  - Bayesian networks
- **Part 4: Machine Learning**
  - Supervised learning
    - Inductive logic programming
    - Linear models
    - Deep neural networks
    - PyTorch
  - Reinforcement learning
    - Markov decision processes
    - Dynamic programming
    - Model-free RL
  - Unsupervised learning
    - Clustering
    - Autoencoders

Outline

- Chains
- Common cause
- Common effect
- D-Separation examples

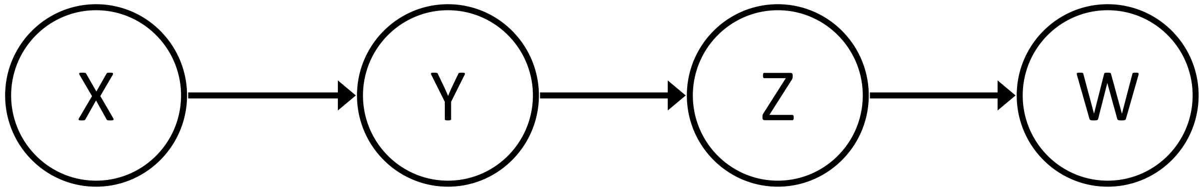
Bayesian Networks: Independence

- Bayesian Networks make the assumption that a variable is independent of its non-descendents given its parents
- This induces other conditional independence assumptions
- We can understand these assumptions by looking at the graph structure
- Understanding these conditional independencies will help us design our Bayesian Network

Notation

- X and Y are independent
  - $X \perp Y$
  - $P(X,Y) = P(X)P(Y)$
- X and Y are conditionally independent given Z
  - $X \perp Y|Z$
  - $P(X,Y|Z) = P(X|Z)P(Y|Z)$
- W is conditionally independent of X and Y given Z
  - $W \perp \{X,Y\}|Z$
  - $P(W,X|Z) = P(W|Z)P(X|Z)$
  - $P(W,Y|Z) = P(W|Z)P(Y|Z)$

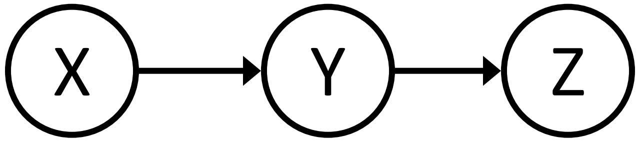
Independence Example



- Conditional independence assumptions directly from simplifications in chain rule:
  - $P(X,Y,Z,W) = P(X)P(Y|X)P(Z|X,Y)P(W|X,Y,Z)$
  - $P(X,Y,Z,W) = P(X)P(Y|X)P(Z|Y)P(W|Z)$
- Conditional independence assumptions
  - $Z \perp X|Y$
  - $W \perp \{X,Y\}|Z$
- Additional implied conditional independence assumptions?
  - $W \perp X|Y$

Independence in Bayesian Networks

- 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 of conditional probability tables (CPTs)
  - Example:



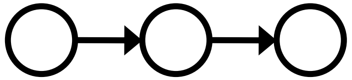
  - Question: are X and Z necessarily independent?
    - Answer: no. X can influence Z, Z can influence X (via Y)
    - Addendum: they *could* be independent: how?
      - We can write the conditional probability tables so that they actually are independent. However, they are not *guaranteed* to be independent

Examining Triples

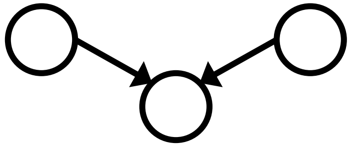
- We will later use D-separation to show whether any two variables X_i and X_j are guaranteed to be conditionally independent given $X_{k_1}, \dots, X_{k_n}$

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

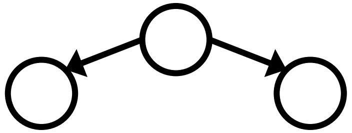
- This will involve examining triples of nodes that are connected via an undirected path (the network still must be a DAG)
- This can simplify our computation as we only have to know independence regarding all possible combinations of triples
  - **Chains**



  - **Common cause**



  - **Common effect**



Structure of Proofs

- X and Y are independent
  - **if and only if** $P(X,Y) = P(X)P(Y)$
  - **if and only if** $P(X|Y) = P(X)$
  - **if and only if** $P(Y|X) = P(Y)$
- X and Y are conditionally independent given Z
  - **if and only if** $P(X,Y|Z) = P(X|Z)P(Y|Z)$
  - **if and only if** $P(X|Y,Z) = P(X|Z)$
  - **if and only if** $P(Y|X,Z) = P(Y|Z)$
- Therefore, if we can show that any of these relationships hold, we can prove independence
- However, if can give a counterexample, meaning a set of conditional probability tables in which this does not hold, then we can not guarantee independence

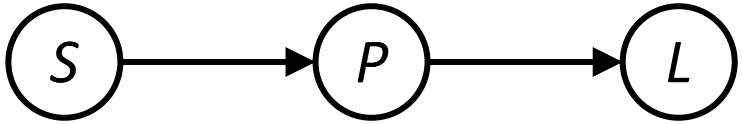
Chains: Not Given Middle

- L guaranteed to be independent of S?
- No! Let's give a counter example

P(S=+s)
0.5

S	P(P=+p|S)
+s	0.9
-s	0.1

P	P(L=+l|P)
+p	0.9
-p	0.1



S: Sun   P: Photosynthesis   L: Leaves green

$$P(S, P, L) = P(S)P(P|S)P(L|P)$$

Chains: Not Given Middle

- $P(+s,+l) = \sum_p P(+s,p,+l) = P(+s)\sum_p P(p|+s)P(+l|p)$
- $= 0.5 * (0.9 * 0.9 + 0.1 * 0.1) = 0.41$
- $P(+s) = \sum_{p,l} P(+s,p,l) = P(+s)\sum_p P(p|+s)\sum_l P(l|p)$
- $= 0.5 * [0.9 * (0.9 + 0.1) \ + 0.1 * (0.1 + 0.9)] = 0.5$
- $P(+l) = \sum_{s,p} P(+s,p,l) = \sum_p P(+l|p)\sum_s P(+s)P(p|+s)$
- $= 0.9 * (0.5 * 0.9 \ + 0.5 * 0.1) + 0.1 * (0.5 * 0.1 + 0.5 * 0.9) = 0.5$
- $P(+s,+l) \neq P(+s)P(+l)$
- $0.41 \neq 0.5 * 0.5$

P(S=+s)
0.5

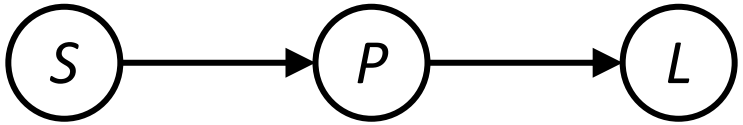
S	P(P=+p|S)
+s	0.9
-s	0.1

P	P(L=+l|P)
+p	0.9
-p	0.1

$P(S,P,L) = P(S)P(P|S)P(L|P)$

Chains: Given Middle

- L guaranteed to be independent of S given P?
- Yes. Let's prove it using the assumptions encoded in this BN.
- Proof
  - $P(L|S,P) = \frac{P(L,S,P)}{P(S,P)}$ //**product rule**
  - $= \frac{P(S)P(P|S)P(L|P)}{P(P|S)P(S)}$ //**assumptions** (numerator) **product rule** (denominator)
  - $P(L|P)$

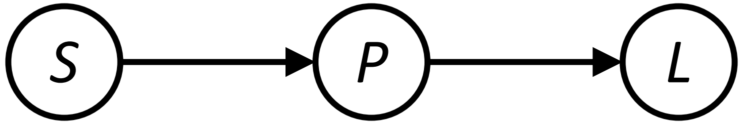


S: Sun P: Photosynthesis L: Leaves green

$$P(S,P,L) = P(S)P(P|S)P(L|P)$$

Chains: Given Middle

- Other ways to prove?
- $P(L,S|P) = \frac{P(L,S,P)}{P(P)}$
- $= \frac{P(S)P(P|S)P(L|P)}{P(P)}$
- $= \frac{P(S)P(S|P)P(P)P(L|P)}{P(S)P(P)}$
- $= P(S|P)P(L|P)$

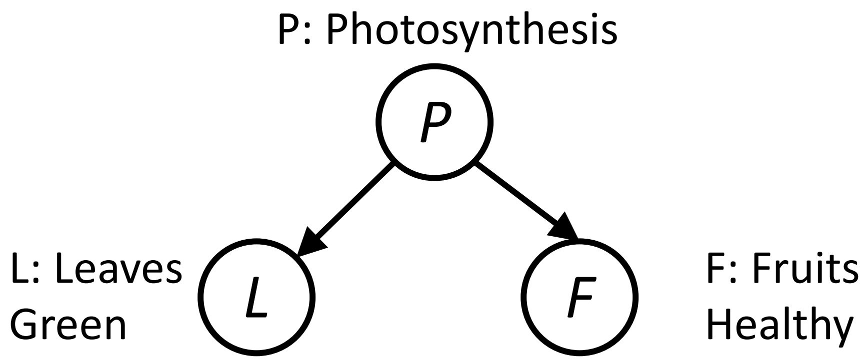


S: Sun   P: Photosynthesis   L: Leaves green

$P(S,P,L) = P(S)P(P|S)P(L|P)$

Common Cause: Not Given Cause

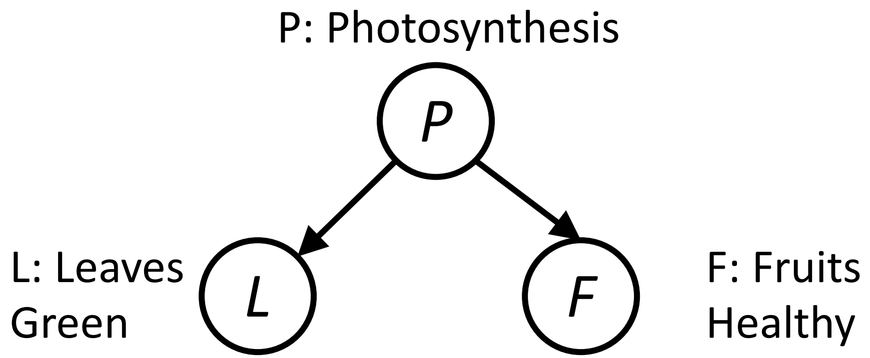
- Is L and F guaranteed to be independent?
- No. Let's prove it with a counter example
- Counter example
  - $P(L = T|P = T) = 0.9$
  - $P(L = T|P = F) = 0.1$
  - $P(F = T|P = T) = 0.9$
  - $P(F = T|P = F) = 0.1$



$$P(P, L, F) = P(P)P(L|P)P(F|P)$$

Common Cause: Given Cause

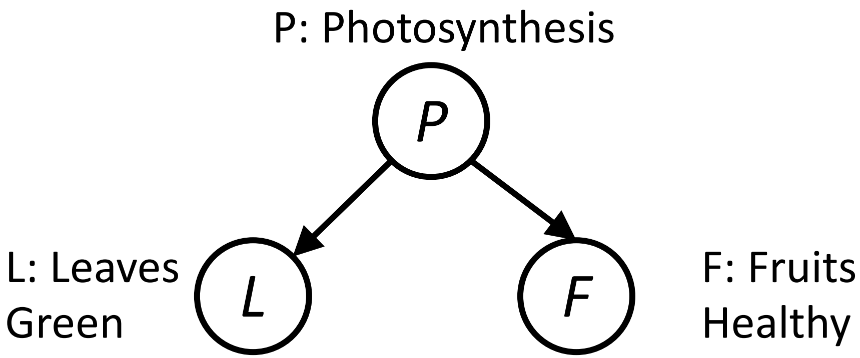
- Is L and F guaranteed to be independent given P?
- Yes. Let's prove it using the assumptions encoded in this BN.
- Proof
  - $P(L|F,P) = \frac{P(P,L,F)}{P(F,P)}$
  - $= \frac{P(P)P(L|P)P(F|P)}{P(F|P)P(P)}$
  - $= P(L|P)$



$$P(P,L,F) = P(P)P(L|P)P(F|P)$$

Common Cause: Given Cause

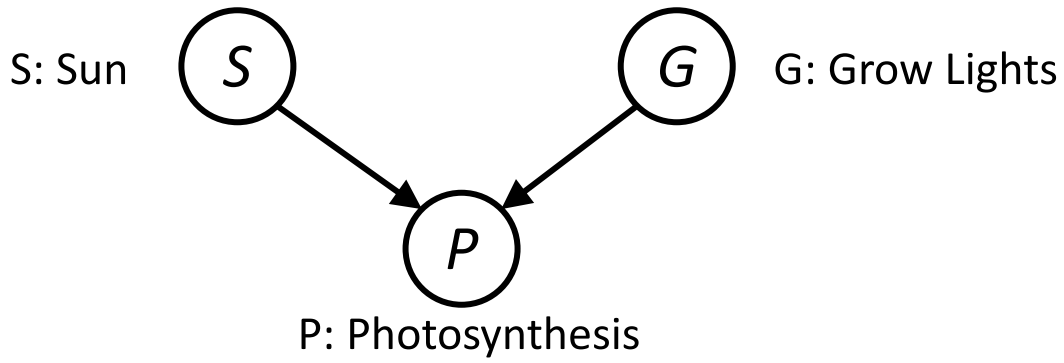
- Other ways to prove?
- Show $P(L,F|P) = P(L|P)P(F|P)$
- $P(L,F|P) = \frac{P(L,F,P)}{P(P)}$
- $= \frac{P(P)P(L|P)P(F|P)}{P(P)}$
- $P(L|P)P(F|P)$



$P(P,L,F) = P(P)P(L|P)P(F|P)$

Common Effect: Not Given Effect

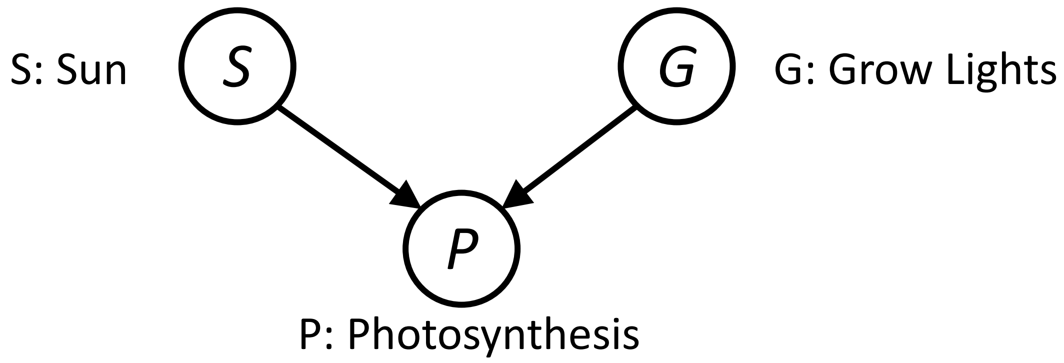
- S guaranteed to be independent of G?
- Yes. Let's prove it using our assumptions
- Proof
  - $P(s, g) = \sum_p P(s, g, p)$
  - $= \sum_p P(s)P(g)P(p|s, g)$
  - $= P(s)P(g) \sum_p P(p|s, g)$
  - $= P(s)P(g)$



$$P(S, G, P) = P(S)P(G)P(P|S, G)$$

Common Effect: Given Effect

- S guaranteed to be independent of G given P?
- No. Observing photosynthesis puts the two explanations in competition with one another
- Observing an effect activates influences between possible causes



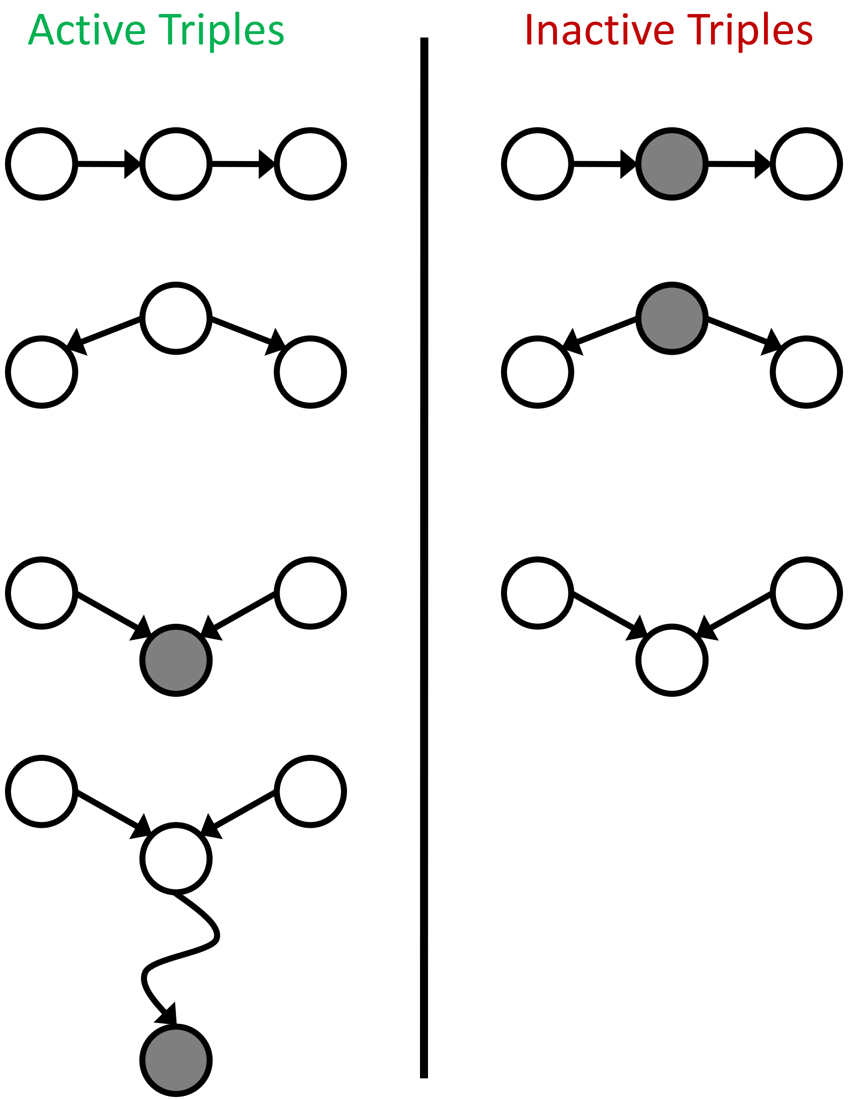
$$P(S,G,P) = P(S)P(G)P(P|S,G)$$

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 -> 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 descendants* is observed

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



D-Separation

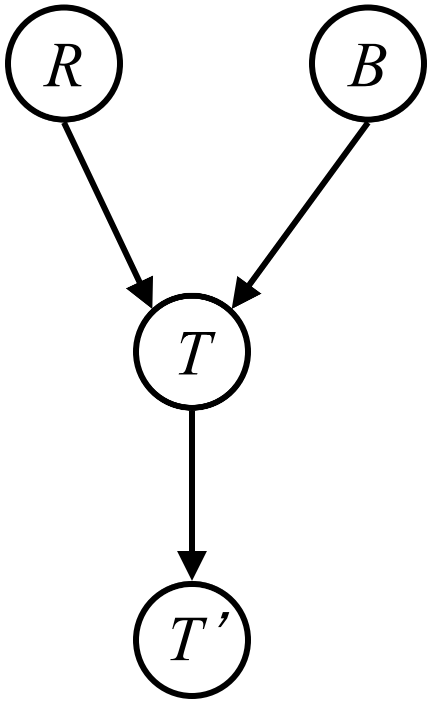
- Show whether any two variables X_i and X_j are guaranteed to be conditionally independent given $X_{k_1}, \ldots, X_{k_n}$

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

- Check all undirected paths from X_i to X_j
- If all paths are "inactive" then it is conditionally independent
  - A path is active if each consecutive triple in the path is active
  - If a single consecutive triple is inactive then the entire path is inactive
- If one or more paths is active, then independence is not guaranteed
- If all paths are inactive, then independence is guaranteed

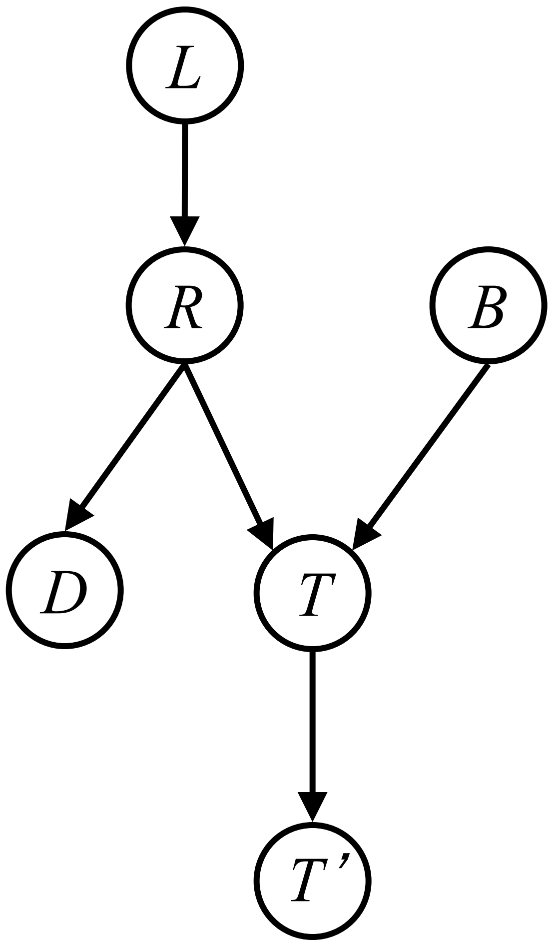
Example

- $R \perp B$
  - Yes
- $R \perp B|T$
  - No
- $R \perp B|T'$
  - No



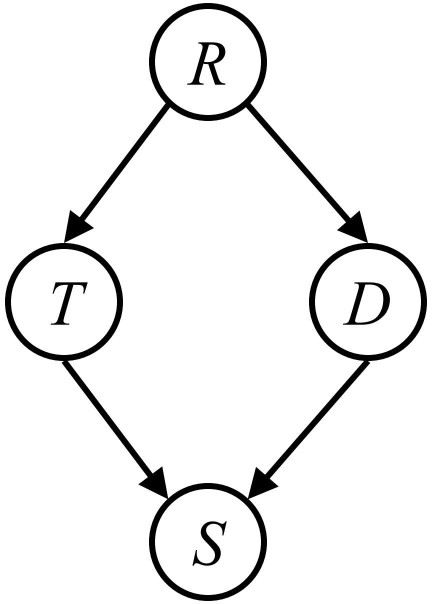
Example

- $L \perp T' | T$
  - Yes
- $L \perp B$
  - Yes
- $L \perp B | T$
  - No
- $L \perp B | T'$
  - No
- $L \perp B | T, R$
  - Yes



Example

- $T \perp D$
  - No
- $T \perp D | R$
  - Yes
- $T \perp D | R, S$
  - No



Markov Blanket

- Markov Blanket: The parents, children, and children's parents

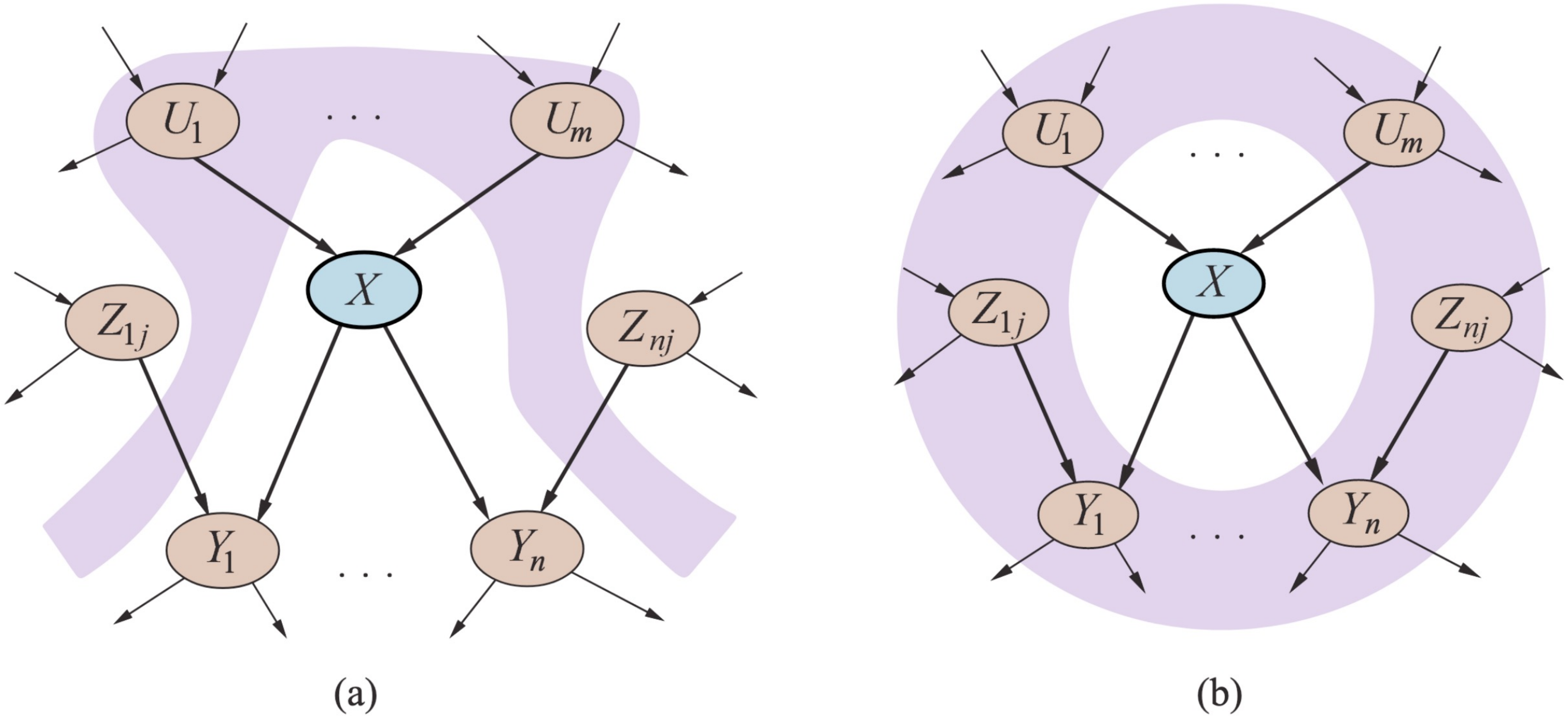
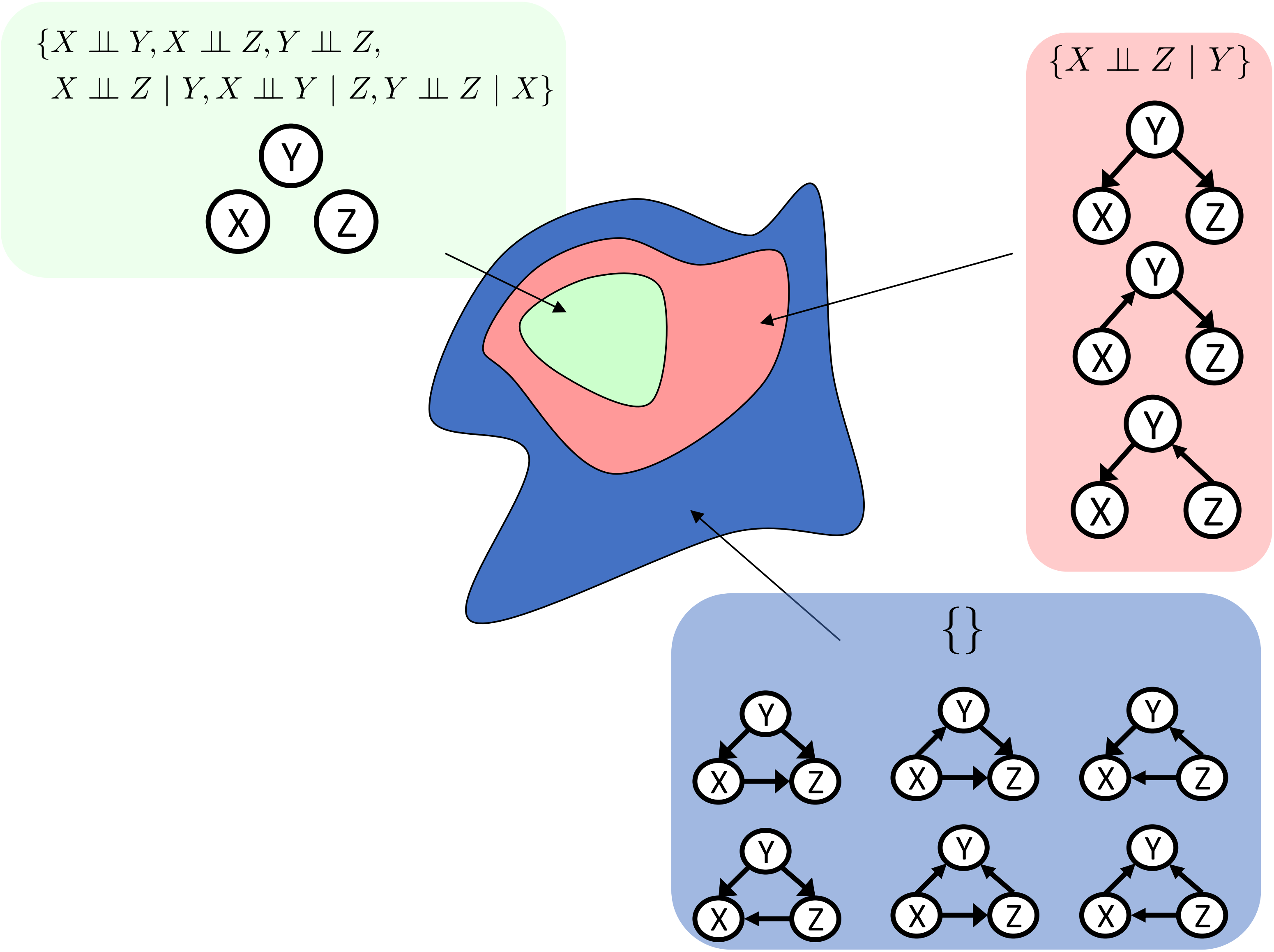


Figure 13.4 (a) A node X is conditionally independent of its non-descendants (e.g., the Z_{ij}s) given its parents (the U_is shown in the gray area). (b) A node X is conditionally independent of all other nodes in the network given its Markov blanket (the gray area).

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



Next Time

- Inference in Bayesian Networks