VE492 Final Exam Recitation

Ma Ziqiao

UM-SJTU Joint Institute

martin_maziqiao@sjtu.edu.cn

August 1, 2020

Bayesian Network

- Bayesian Networks
  A technique for describing complex joint distributions (models) using simple, local distributions (conditional probabilities).
- Notations
  - Nodes: variables with domains
    Can be assigned (observed) or unassigned (unobserved)
  - Arcs: interactions Encode conditional independence relationships.

Bayesian Network: Probabilities

- Chain rule (valid for all distributions):

$$P(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} P(x_i|x_1, \cdots, x_{i-1})$$

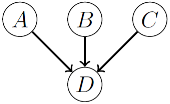
- Conditional independences:

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

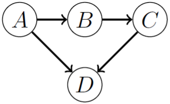
- $\Rightarrow P(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} P(x_i|\text{Parent}(X_i))$

Practice Problem

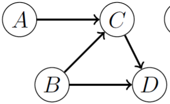
The full joint probability distribution $P(A, B, C, D)$ for four binary random variables is a 16-element array of numbers adding up to 1.0, so it has 15 degrees of freedom. A Bayes net makes explicit the conditional dependencies (and therefore also the conditional independence relations) among the random variables, which simplifies the information that must be provided to determine $P(A, B, C, D)$. For each of the Bayes' net diagrams below, express $P(A, B, C, D)$ as the product of conditional probabilities, and give the number of degrees of freedom required to specify the conditional probability tables.



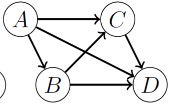
Graph 1.



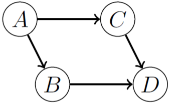
Graph 2.



Graph 3.



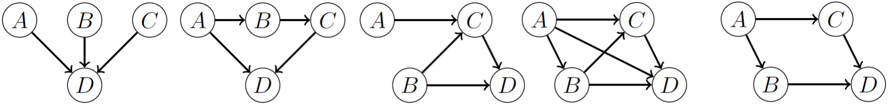
Graph 4.



Graph 5.

Practice Problem

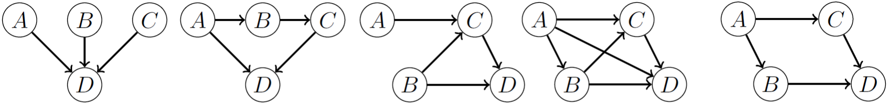
Provide $P(A, B, C, D)$ and the number of degrees of freedom for Graph 1-5.



Graph 1. Graph 2. Graph 3. Graph 4. Graph 5.

Practice Problem

Provide $P(A, B, C, D)$ and the number of degrees of freedom for Graph 1-5.



Graph 1. Graph 2. Graph 3. Graph 4. Graph 5.

Solution

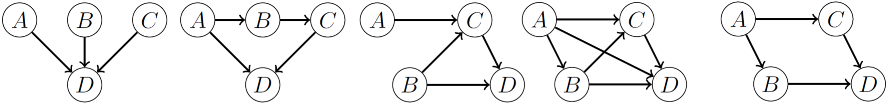
Graph 1:

$P(A, B, C, D) = P(D|A, B, C)P(A)P(B)P(C)$

Number of degrees of freedom $= = 2^3 + 1 + 1 + 1 = 11$

Practice Problem

Provide $P(A, B, C, D)$ and the number of degrees of freedom for Graph 1-5.



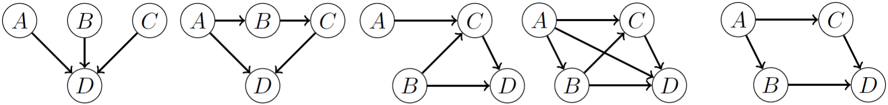
Graph 1. Graph 2. Graph 3. Graph 4. Graph 5.

Solution

Graph 2:
$P(A, B, C, D) = P(D|A, C)P(C|B)P(B|A)P(A)$
Number of degrees of freedom $= = 2^2 + 2 + 2 + 1 = 9$

Practice Problem

Provide $P(A, B, C, D)$ and the number of degrees of freedom for Graph 1-5.



Graph 1. Graph 2. Graph 3. Graph 4. Graph 5.

Solution

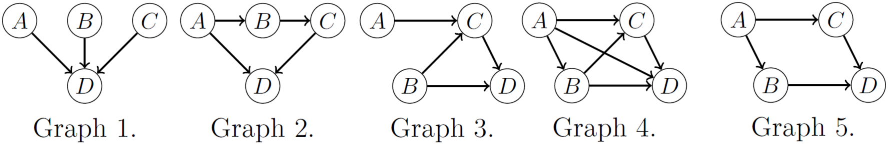
Graph 3:

$P(A, B, C, D) = P(D|B, C)P(C|A, B)P(B)P(A)$

Number of degrees of freedom $= = 2^2 + 2^2 + 1 + 1 = 10$

Practice Problem

Provide $P(A, B, C, D)$ and the number of degrees of freedom for Graph 1-5.



Solution

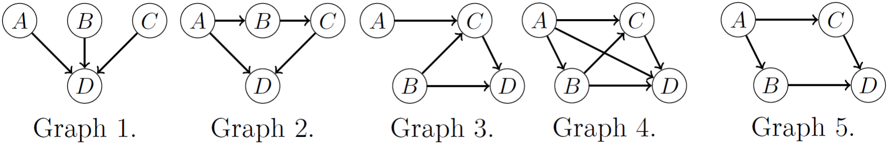
Graph 4:

$P(A, B, C, D) = P(D|A, B, C)P(C|A, B)P(B|A)P(A)$

Number of degrees of freedom $= = 2^3 + 2^2 + 2 + 1 = 15$

Practice Problem

Provide $P(A, B, C, D)$ and the number of degrees of freedom for Graph 1-5.



Solution

Graph 5:

$P(A, B, C, D) = P(D|B, C)P(C|A)P(B|A)P(A)$

Number of degrees of freedom $= = 2^2 + 2 + 2 + 1 = 9$

Bayesian Network: Conditional Independences

- Independence
  X and Y are independent ($X \perp\!\!\!\perp Y$) if $\forall x, y\ P(x, y) = P(x)P(y)$
- Conditional Independence
  X and Y are conditional independent given Z ($X \perp\!\!\!\perp Y|Z$) if $\forall x, y, z\ P(x, y|z) = P(x|z)P(y|z)$

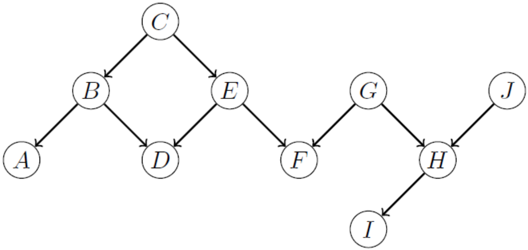
Bayesian Network: D-separation

A practical procedure of D-separation:

- Keep only ancestral graph;
- Connect nodes with common child;
- Make all edges undirect;
- Read off the properties.
  - Marginal independent if there is no path between nodes.
  - Conditional independent if all paths between nodes get across given nodes.

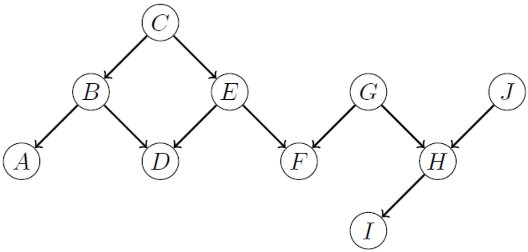
Practice Problem

Consider the following Bayes' net structure. Answer "yes" or "no" to the following questions regarding (conditional) independences and justify your answer.



Practice Problem

$$G \perp\!\!\!\perp I$$

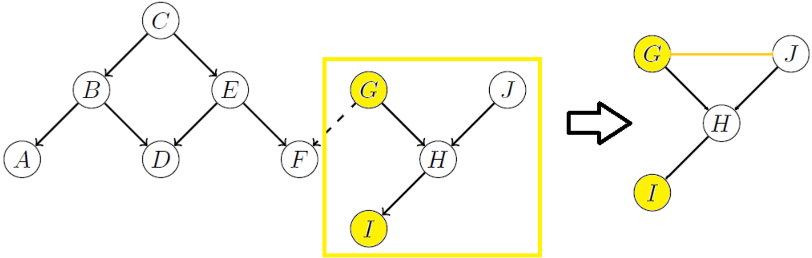


Practice Problem

$$G \perp\!\!\!\perp I$$

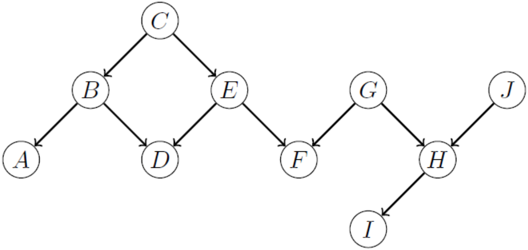
Solution

No, there is an active path (G, H, I).



Practice Problem

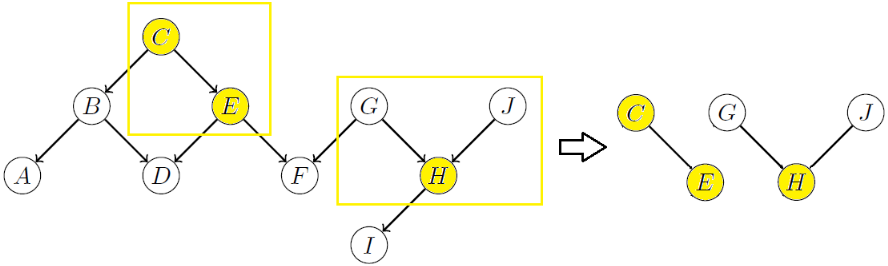
$$E \perp\!\!\!\perp H \mid C$$



$$E \perp\!\!\!\perp H | C$$

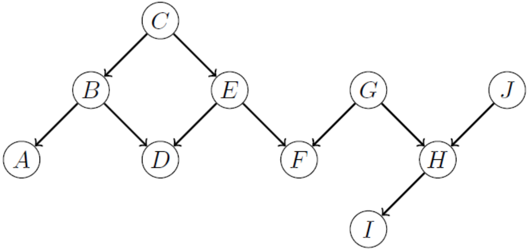
Solution

Yes.



Practice Problem

$$B \perp\!\!\!\perp H | A, C, D, E, F, G, J$$

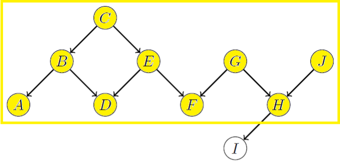


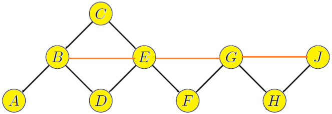
Practice Problem

$$B \perp\!\!\!\perp H | A, C, D, E, F, G, J$$

Solution

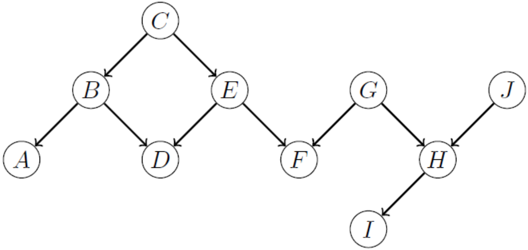
Yes.





Practice Problem

$$B \perp\!\!\!\perp F | A, C, D, G, H$$

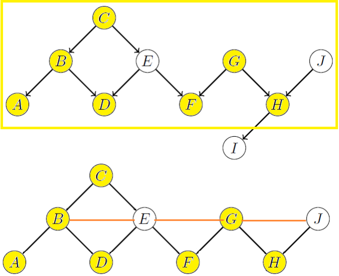


Practice Problem

$$B \perp\!\!\!\perp F | A, C, D, G, H$$

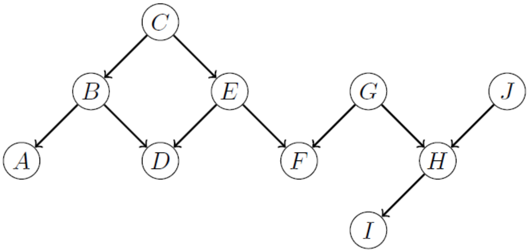
Solution

No.



Practice Problem

$$A, I \perp\!\!\!\perp C, G \mid B, H$$

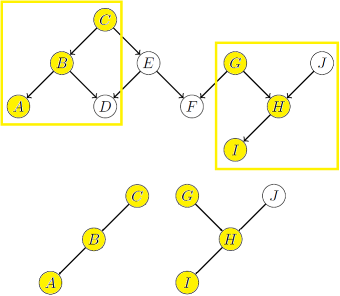


Practice Problem

$$A, I \perp\!\!\!\perp C, G \mid B, H$$

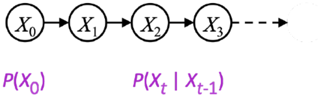
Solution

Yes.



Markov Problem

- State: Value of X at a given time
- Parameters:
  - Initial state probabilities
  - Transition probabilities (or dynamics) specify how state evolves over time
- Stationarity assumption: transition probabilities the same at all times $P(X'|X)$

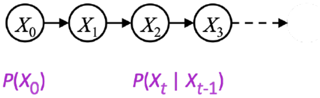


Markov Problem

- Markov assumption: future is independent of the past given the present
  - X_{t+1} is independent of $X_0, X_1, \cdots, X_t$
  - First-order Markov model
  - kth-order = dependencies on k earlier steps
- Joint distribution

$$P(X_0, X_1, \cdots, X_t) = P(X_0) \prod_{t=1}^{T} P(X_t | X_{t-1})$$

- Markov chain is a growable BN



Markov Problem

- Probability/stochastic vector π: A vector with non-negative entries that add up to one.
- Transition/stochastic Matrix T: A matrix whose (i, j) entry gives the probability that an element moves from the jth state to the ith state during the next step of the process.
  - If the Markov chain is time-homogeneous, then the transition matrix T is the same after each step.

$$\pi_{t+1} = T\pi_t$$

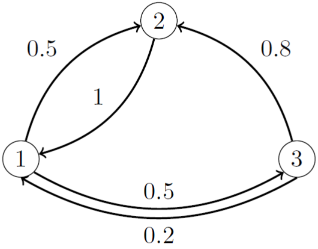
- The limiting distribution (if it exists) is called the stationary distribution π_∞ of the chain.

$$\pi_\infty = \pi_{\infty+1} = T\pi_\infty$$

Practice Problem

Consider the Markov chain.

- Assuming an uniform distribution over initial states, compute the probability of being in each state at time step 2.
- Compute its stationary distribution.



Practice Problem

Assuming an uniform distribution over initial states, compute the probability of being in each state at time step 2.

Solution

$$
\begin{aligned}
\pi_0 &= \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \\
T &= \begin{bmatrix} 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \\ 0.2 & 0.8 & 0 \end{bmatrix} \\
\pi_2 = T^2 \pi_0 &= \begin{bmatrix} 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \\ 0.2 & 0.8 & 0 \end{bmatrix}^2 \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}
\end{aligned}
$$

Practice Problem

Assuming an uniform distribution over initial states, compute the probability of being in each state at time step 2.

Solution

$$\pi_\infty = \begin{bmatrix} x & y & z \end{bmatrix}$$

$$T = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \\ 0.2 & 0.8 & 0 \end{bmatrix}$$

Practice Problem

Solution

$$
\begin{aligned}
\pi_\infty &= T\pi_\infty \\
\begin{bmatrix} x & y & z \end{bmatrix} &= \begin{bmatrix} 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \\ 0.2 & 0.8 & 0 \end{bmatrix} \begin{bmatrix} x & y & z \end{bmatrix}
\end{aligned}
$$

Hence,

$$
\begin{aligned}
x + y + z = 1 \\
x &= 0.5y + 0.5z \\
y &= x \\
z &= 0.2x + 0.8y \\
\Rightarrow x = y = z &= \frac{1}{3}
\end{aligned}
$$