BAYESIAN NETWORK

Daegun Won

Outline

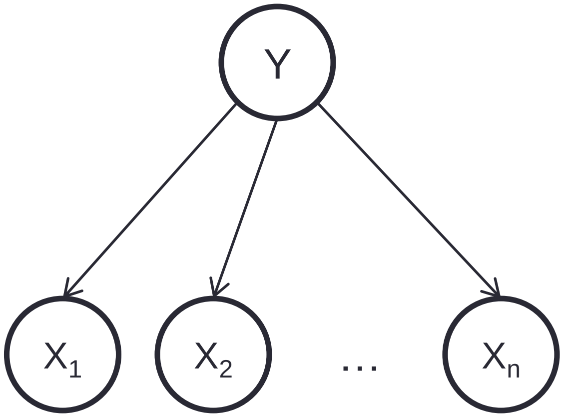
- Plate Notation
- d-separation
- Variable Elimination

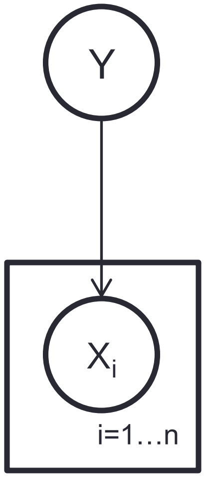
PLATE NOTATION

Plate notation

- Consider a Naïve Bayes classifier with “many” features
  - Bayes Net representation?
  - With 10 features?
  - With 100000 features?

Naïve Bayes simplified





Examples

- Linear/logistic regression?
- Gaussian model?
- Gaussian mixture model?

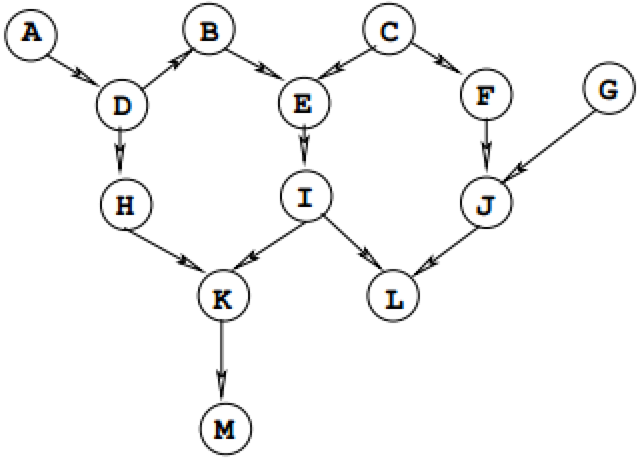
D-SEPARATION

d-separation

- Let **Z** be a set of observed variables.
- A trail $X_1 \leftrightarrow \ldots \leftrightarrow X_n$ is active given Z if
  - Whenever we have a collider $X_{i-1} \rightarrow X_i \leftarrow X_{i+1}$, then X_i or one of its descendants are in Z
  - No other node along the trail in Z
- **X** and **Y** are d-separated given **Z** if there is no active trail between any node X in **X** and Y in **Y** given **Z**

d-sep: examples

- d-sep (A,{C,F,G,J})?
- d-sep (A,G|M,B)?
- d-sep (A,M|E,K,L)?
- d-sep (I,M|E,K,L)?



Cond. Indep -> d-separation?

- Completeness
- Counterexample

	A = T	A = F
P(B = T|A)	.4	.4
P(B = F|A)	.6	.6

- P(B)? P(B|A)?

Okay then..

- Counterexample to completeness hypothesis
- Is d-separation really useful?
  - What's the point if d-separation can't catch all, or most of the conditional independencies?

Weaker completeness definition

- If $X \perp Y | Z$ in ***all*** distributions P that factorize over G, then X and Y are d-separated given Z
  - Contrapositive: If X and Y are not d-separated given Z in G, then X and Y are dependent in some distribution P that factorizes over G

- Strengthening the above..
  - For almost all distributions P that factorize over G, that is, for all distributions except for a set of measure zero in the space of CPD parameterizations, we have that I(P)=I(G)

VARIABLE ELIMINATION

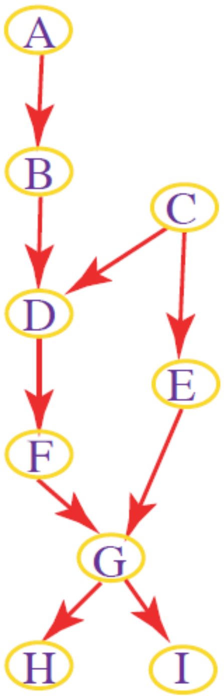
(Taken from www.cs.ubc.ca/~kevinlb/teaching/cs322 - 2006-7/Lectures/lect29.pdf)

VE: Summary

- $P(X_1, X_2, \dots, X_{n_X} | Z_1 = z_1, \dots, Z_{n_Z} = z_{n_Z})$
- Set the observed variables Z_k to their values
- For each of the irrelevant variables $Y_j \in \{Y_1, \dots, Y_{n_Y}\}$,
  - $\sum_{Y_j} f_1 \times \cdots \times f_n = (f_1 \times \cdots \times f_i)\left(\sum_{Y_j} f_{i+1} \times \cdots \times f_n\right)$
  - Let $f' = \left(\sum_{Y_j} f_{i+1} \times \cdots \times f_n\right)$
  - $\sum_{Y_j} f_1 \times \cdots \times f_i \times f'$

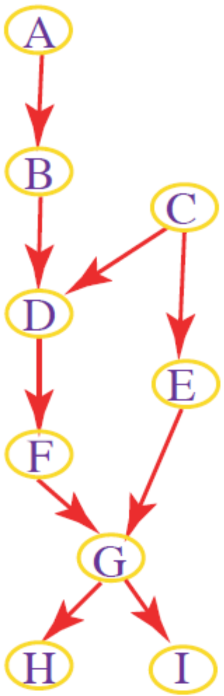
Variable Elimination

- Compute $P(G|H = h_1)$.
- Elimination order: A, C, E, H, I, B, D, F
- $P(G, H) = \sum_{A,B,C,D,E,F,I} P(A, B, C, D, E, F, G, H, I)$
- $= \sum_{A,B,C,D,E,F,I} P(A)P(B|A)P(C)P(D|B, C)P(E|C) P(F|D)P(G|F, E)P(H|G)P(I|G)$



Variable Elimination

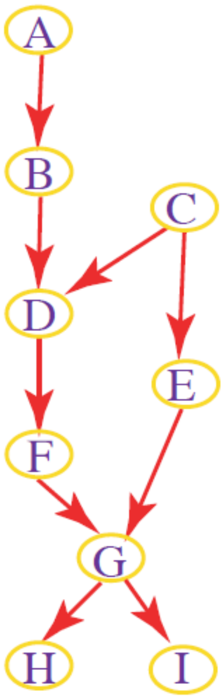
- Compute $P(G|H = h_1)$.
- Elimination order: A, C, E, H, I, B, D, F
- $P(G,H) = \sum_{A,B,C,D,E,F,I} P(A,B,C,D,E,F,G,H,I)$
- $= \sum_{A,B,C,D,E,F,I} \begin{matrix} P(A)P(B|A)P(C)P(D|B,C)P(E|C) \\ P(F|D)P(G|F,E)P(H|G)P(I|G) \end{matrix}$



- Eliminating A
  - $P(G,H) = \sum_{B,C,D,E,F,I} \begin{matrix} f_1(B)P(C)P(D|B,C)P(E|C) \\ P(F|D)P(G|F,E)P(H|G)P(I|G) \end{matrix}$
  - $f_1(B) = \sum_a P(A=a)P(B|A=a)$

Variable Elimination

- Compute $P(G|H = h_1)$.
- Elimination order: A, C, E, H, I, B, D, F
- $P(G,H) = \sum_{B,C,D,E,F,I} f_1(B)P(C)P(D|B,C)P(E|C)P(F|D)P(G|F,E)P(H|G)P(I|G)$

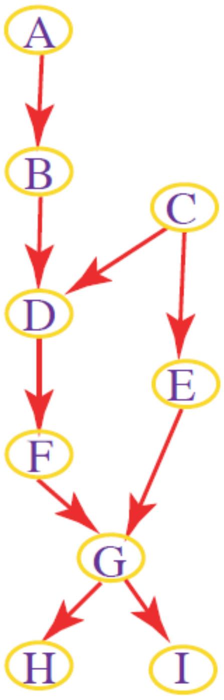


- Eliminating C
  - $P(G,H) = \sum_{B,D,E,F,I} f_1(B)f_2(B,D,E)P(F|D)P(G|F,E)P(H|G)P(I|G)$
  - $f_1(B) = \sum_a P(A = a)P(B|A = a)$
  - $f_2(B,D,E) = \sum_c P(C = c)P(D|B,C = c)P(E|C = c)$

Variable Elimination

- Compute $P(G|H = h_1)$.
- Elimination order: A, C, E, H, I, B, D, F
- $P(G, H) = \sum_{B,D,E,F,I} f_1(B) f_2(B, D, E) P(F|D) P(G|F, E) P(H|G) P(I|G)$

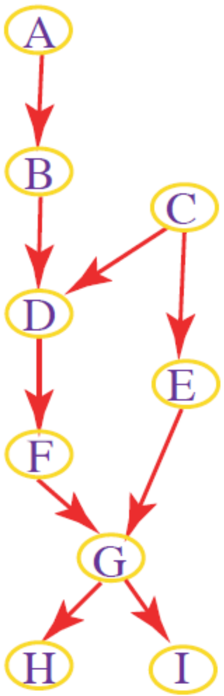
- Eliminating E
  - $P(G, H) = \sum_{B,D,F,I} f_1(B) f_3(B, D, F, G) P(F|D) P(H|G) P(I|G)$
  - $f_1(B) = \sum_a P(A = a) P(B|A = a)$
  - $f_2(B, D, E) = \sum_c P(C = c) P(D|B, C = c) P(E|C = c)$
  - $f_3(B, D, F, G) = \sum_e f_2(B, D, E = e) P(G|F, E = e)$



Variable Elimination

- Compute $P(G|H = h_1)$.
- Elimination order: A, C, E, H, I, B, D, F
- $P(G, H) = \sum_{B,D,F,I} f_1(B) f_2(B, D, E) P(F|D) P(G|F, E) P(H|G) P(I|G)$

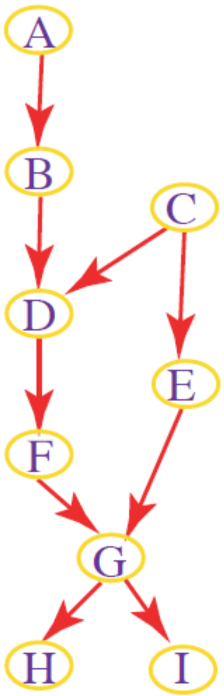
- Observe H
  - $P(G, H = h_1) = \sum_{B,D,F,I} f_1(B) f_3(B, D, F, G) P(F|D) f_4(G) P(I|G$
  - $f_1(B) = \sum_a P(A = a) P(B|A = a)$
  - $f_2(B, D, E) = \sum_c P(C = c) P(D|B, C = c) P(E|C = c)$
  - $f_3(B, D, F, G) = \sum_e f_2(B, D, E = e) P(G|F, E = e)$
  - $f_4(G) = P(H = h_1|G)$



Variable Elimination

- Compute $P(G|H = h_1)$.
- Elimination order: A, C, E, H, I, B, D, F
- $P(G, H = h_1) = \sum_{B,D,F,I} f_1(B) f_3(B, D, F, G) P(F|D) f_4(G) P(I|G)$

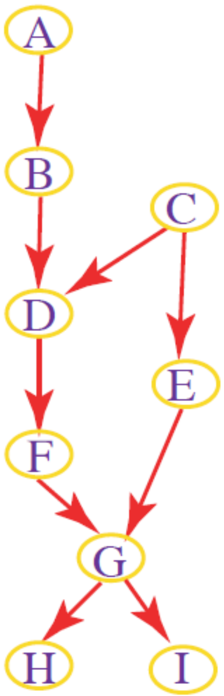
- Eliminating I
  - $P(G, H = h_1) = \sum_{B,D,F} f_1(B) f_3(B, D, F, G) P(F|D) f_4(G) f_5(G)$
  - $f_1(B) = \sum_a P(A = a) P(B|A = a)$
  - $f_2(B, D, E) = \sum_c P(C = c) P(D|B, C = c) P(E|C = c)$
  - $f_3(B, D, F, G) = \sum_e f_2(B, D, E = e) P(G|F, E = e)$
  - $f_4(G) = P(H = h_1|G)$, $f_5(G) = \sum_i P(I = i|G)$



Variable Elimination

- Compute $P(G|H = h_1)$.
- Elimination order: A, C, E, H, I, B, D, F
- $P(G, H = h_1) = \sum_{B,D,F} f_1(B) f_3(B, D, F, G) P(F|D) f_4(G) f_5(G)$

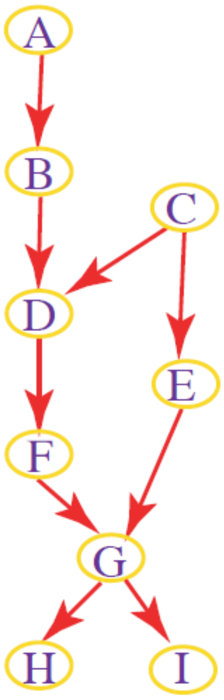
- Eliminating B
  - $P(G, H = h_1) = \sum_{D,F} f_6(D, F, G) P(F|D) f_4(G) f_5(G)$
  - $f_1(B) = \sum_a P(A = a) P(B|A = a)$
  - $f_2(B, D, E) = \sum_c P(C = c) P(D|B, C = c) P(E|C = c)$
  - $f_3(B, D, F, G) = \sum_e f_2(B, D, E = e) P(G|F, E = e)$
  - $f_4(G) = P(H = h_1|G)$, $f_5(G) = \sum_i P(I = i|G)$
  - $f_6(D, F, G) = \sum_b f_1(B = b) f_3(B = b, D, F, G)$



Variable Elimination

- Compute $P(G|H = h_1)$.
- Elimination order: A, C, E, H, I, B, D, F
- $P(G, H = h_1) = \sum_{D,F} f_6(D, F, G)P(F|D)f_4(G)f_5(G)$

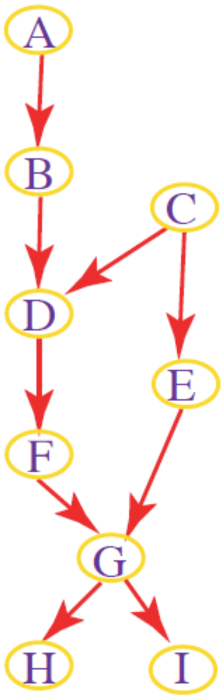
- Eliminating D
  - $P(G, H = h_1) = \sum_F f_7(F, G)f_4(G)f_5(G)$
  - $f_1(B) = \sum_a P(A = a)P(B|A = a)$
  - $f_2(B, D, E) = \sum_c P(C = c)P(D|B, C = c)P(E|C = c)$
  - $f_3(B, D, F, G) = \sum_e f_2(B, D, E = e)P(G|F, E = e)$
  - $f_4(G) = P(H = h_1|G)$, $f_5(G) = \sum_i P(I = i|G)$
  - $f_6(D, F, G) = \sum_b f_1(B = b)f_3(B = b, D, F, G)$



Variable Elimination

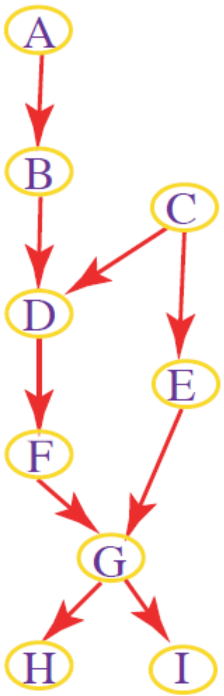
- Compute $P(G|H = h_1)$.
- Elimination order: A, C, E, H, I, B, D, F
- $P(G, H = h_1) = \sum_F f_7(F, G) f_4(G) f_5(G)$

- Eliminating F
  - $P(G, H = h_1) = f_8(G) f_4(G) f_5(G)$
  - $f_1(B) = \sum_a P(A = a) P(B|A = a)$
  - $f_2(B, D, E) = \sum_c P(C = c) P(D|B, C = c) P(E|C = c)$
  - $f_3(B, D, F, G) = \sum_e f_2(B, D, E = e) P(G|F, E = e)$
  - $f_4(G) = P(H = h_1|G)$, $f_5(G) = \sum_i P(I = i|G)$
  - $f_6(D, F, G) = \sum_b f_1(B = b) f_3(B = b, D, F, G)$
  - $f_8(G) = \sum_f f_7(F = f, G)$



Variable Elimination

- Compute $P(G|H = h_1)$.
- Elimination order: A, C, E, H, I, B, D, F
- $P(G, H = h_1) = f_8(G)f_4(G)f_5(G)$
  - $f_1(B) = \sum_a P(A = a)P(B|A = a)$
  - $f_2(B, D, E) = \sum_c P(C = c)P(D|B, C = c)P(E|C = c)$
  - $f_3(B, D, F, G) = \sum_e f_2(B, D, E = e)P(G|F, E = e)$
  - $f_4(G) = P(H = h_1|G)$, $f_5(G) = \sum_i P(I = i|G)$
  - $f_6(D, F, G) = \sum_b f_1(B = b)f_3(B = b, D, F, G)$
  - $f_8(G) = \sum_f f_7(F = f, G)$
- $P(G|H = h_1) = \frac{P(G,H=h_1)}{\sum_g P(G=g,H=h_1)}$



VE

- Exact inference
- Dynamic programming?
  - Memoizing intermediate results
- Other approximate inference methods
  - Loopy belief propagation
  - Sampling
  - Variational inference