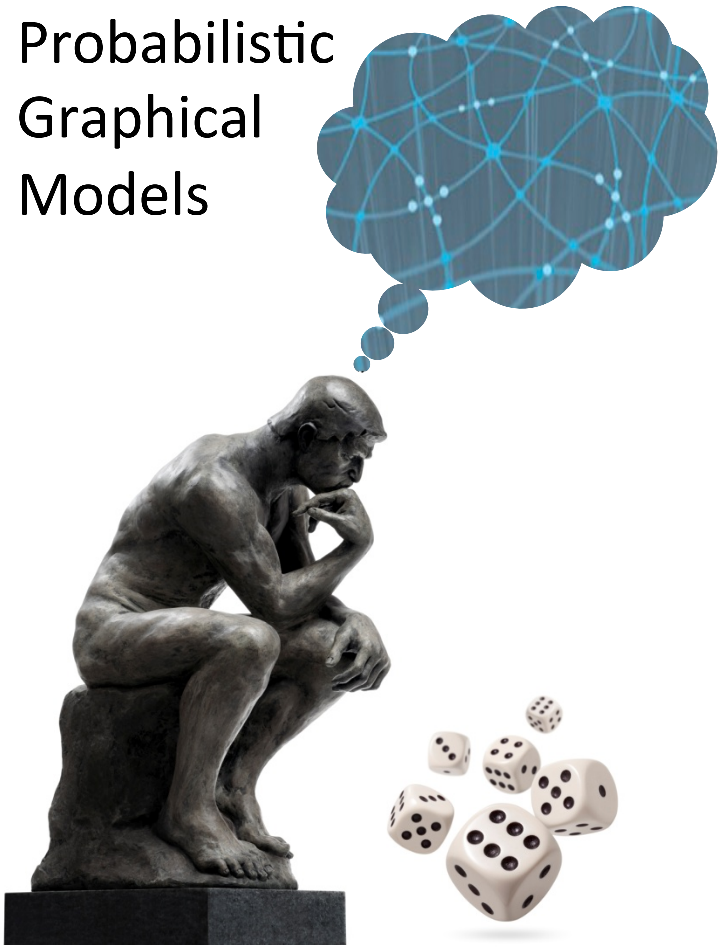
Probabilistic
Graphical
Models



Inference

Variable Elimination

Complexity Analysis

Eliminating Z

$$\psi_k(\boldsymbol{X}_k) = \prod_{i=1}^{m_k} \phi_i$$ factor product

$$\tau_k(\boldsymbol{X}_k - \{Z\}) = \sum_Z \psi_k(\boldsymbol{X}_k)$$ marginalization

Reminder: Factor Product

$$\psi_k(\boldsymbol{X}_k) = \prod_{i=1}^{m_k} \phi_i$$

each row $m_k - 1$ products

$N_k = |Val(X_k)|$

A B

a^1	b^1	0.5
a^1	b^2	0.8
a^2	b^1	0.1
a^2	b^2	0
a^3	b^1	0.3
a^3	b^2	0.9

B C

b^1	c^1	0.5
b^1	c^2	0.7
b^2	c^1	0.1
b^2	c^2	0.2

A, B, C

a^1	b^1	c^1	0.5·0.5 = 0.25
a^1	b^1	c^2	0.5·0.7 = 0.35
a^1	b^2	c^1	0.8·0.1 = 0.08
a^1	b^2	c^2	0.8·0.2 = 0.16
a^2	b^1	c^1	0.1·0.5 = 0.05
a^2	b^1	c^2	0.1·0.7 = 0.07
a^2	b^2	c^1	0·0.1 = 0
a^2	b^2	c^2	0·0.2 = 0
a^3	b^1	c^1	0.3·0.5 = 0.15
a^3	b^1	c^2	0.3·0.7 = 0.21
a^3	b^2	c^1	0.9·0.1 = 0.09
a^3	b^2	c^2	0.9·0.2 = 0.18

Cost: $(m_k-1)N_k$ multiplications

Daphne Koller

Reminder: Factor Marginalization

each number used exactly once

marg B

$$\tau_k(\boldsymbol{X}_k - \{Z\}) = \sum_Z \psi_k(\boldsymbol{X}_k)$$

$N_k = |Val(X_k)|$

Cost: $\sim N_k$ additions

a^1	b^1	c^1	0.25
a^1	b^1	c^2	0.35
a^1	b^2	c^1	0.08
a^1	b^2	c^2	0.16
a^2	b^1	c^1	0.05
a^2	b^1	c^2	0.07
a^2	b^2	c^1	0
a^2	b^2	c^2	0
a^3	b^1	c^1	0.15
a^3	b^1	c^2	0.21
a^3	b^2	c^1	0.09
a^3	b^2	c^2	0.18

a^1	c^1	0.33
a^1	c^2	0.51
a^2	c^1	0.05
a^2	c^2	0.07
a^3	c^1	0.24
a^3	c^2	0.39

Daphne Koller

Complexity of Variable Elimination

- Start with m factors
  - $m \leq n$ for Bayesian networks (one for every variable)
  - can be larger for Markov networks
- At each elimination step generate 1 factor
- At most n elimination steps
- Total number of factors: $m^* \leq m+n$

Daphne Koller

Complexity of Variable Elimination

- $N = \max(N_k)$ = size of the largest factor
- Product operations: $\sum_k (m_k-1)N_k \le N \underbrace{\sum_k (m_k-1)}_{\le m^*}$
  each factor multiply in at most once
  $N m^*$
- Sum operations: $\le \sum_k N_k \le N \cdot$ #elimination steps $\le N \cdot n$
- Total work is linear in N and m*

Daphne Koller

Complexity of Variable Elimination

- Total work is ~~linear in N and m~~ exponential blowup
- $N_k = |Val(\mathbf{X}_k)| = O(d^{r_k})$ where # variables in k^{th} factor
  - $d = \max(|Val(X_i)|)$ d values in their scope
  - $r_k = |\mathbf{X}_k|$ = cardinality of the scope of the k^{th} factor

Complexity Example

$$\tau_1(D) = \sum_{C} \phi_C(C)\phi_D(C,D)$$ 2

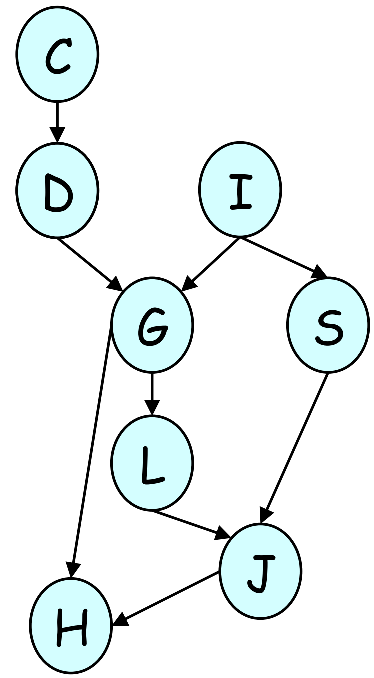
$$\tau_2(G,I) = \sum_{D} \phi_G(G,I,D)\tau_1(D)$$ 3

$$\tau_3(S,G) = \sum_{I} \phi_S(S,I)\phi_I(I)\tau_2(G,I)$$ 3

$$\tau_4(G,J) = \sum_{H} \phi_H(H,G,J)$$ 3

$$\tau_5(J,L,S) = \sum_{G} \phi_L(L,G)\tau_3(S,G)\tau_4(G,J)$$ 4

$$\tau_6(J) = \sum_{L,S} \phi_J(J,L,S)\tau_5(J,L,S)$$ 3



Complexity and Elimination Order

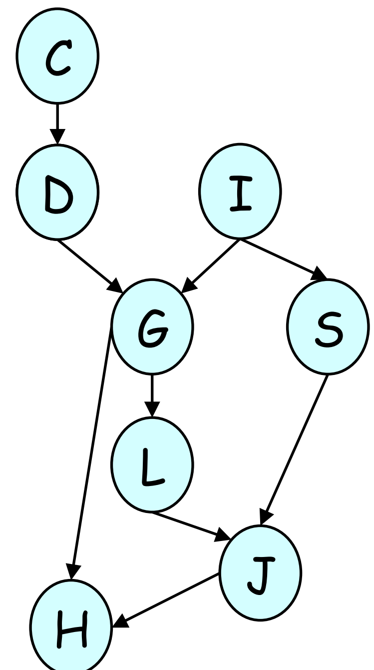
$$\sum_{L,S,G,H,I,D,C} \phi_J(J,L,S)\underline{\phi_L(L,G)}\phi_S(S,I)\underline{\phi_G(G,I,D)}\underline{\phi_H(H,G,J)}\phi_I(I)\phi_D(C,D)\phi_C(C)$$

- Eliminate: G

6

L, G, I, D, H, J

$$\sum_G \phi_L(L,G)\phi_G(G,I,D)\phi_H(H,G,J)$$

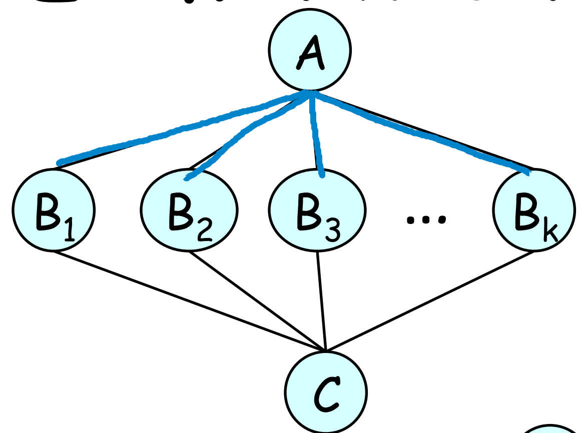


Complexity and Elimination Order

Eliminate A first:

$\{A, B_1, \ldots, B_k\}$

size of factor is exp in k

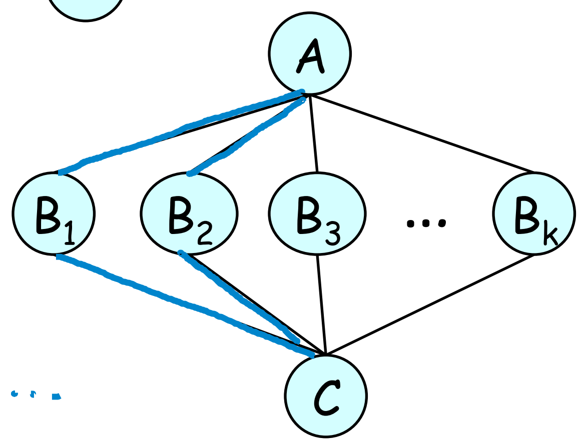


$\prod_i \tau_i(A,C)$

Eliminate B_i's first:

$\underbrace{\phi_{A_i}(A,B_i) \cdot \phi_{C_i}(C,B_i)}_{\text{scope } A,\, B_i,\, C} \Longrightarrow \tau_1(A,C)$

$\tau_2(A,C) \ \ldots$



Daphne Koller

Summary

- Complexity of variable elimination linear in
  - size of the model (# factors, # variables)
  - size of the largest factor generated
- Size of factor is exponential in its scope
- Complexity of algorithm depends heavily on elimination ordering