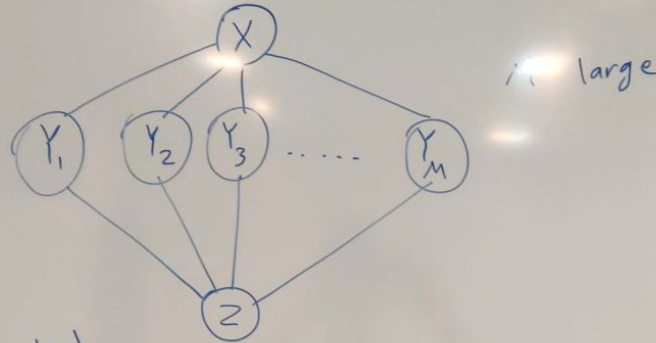


# Probabilistic Graphical Models

## Lectures 13

Message Passing - Junction Tree

# Elimination order



$$P(X, Z, Y_1, Y_2, \dots, Y_m) = \frac{1}{Z} \phi_1(X, Y_1) \phi_2(X, Y_2) \dots \phi_m(X, Y_m) \psi_1(Z, Y_1) \psi_2(Z, Y_2) \dots \psi_m(Z, Y_m)$$

$P(Z) = ?$  eliminate  $X, Y_1, \dots, Y_m$

first eliminate  $X$

$$P(Z, Y_1, Y_2, \dots, Y_m) = \frac{1}{Z} \prod_{i=1}^m \psi_i(Z, Y_i) \sum_X \prod_{i=1}^m \phi_i(X, Y_i)$$

$$= \frac{1}{Z} \prod_{i=1}^m \psi_i(Z, Y_i) \sum_X \delta(X, Y_1, Y_2, \dots, Y_m) \rightarrow \text{too large to compute}$$

# Elimination order



eliminate  $Y_i$ -s first

eliminate  $Y_1$

$$\begin{aligned} P(X, Z, Y_2, Y_3, \dots, Y_m) &= \frac{1}{2} \sum_{Y_1} \prod_{i=1}^m \psi_i(z, Y_i) \prod_{i=1}^m \phi_i(X, Y_i) \\ &= \frac{1}{2} \prod_{i=2}^m \psi_i(z, Y_i) \prod_{i=2}^m \phi_i(X, Y_i) \underbrace{\sum_{Y_1} \psi(z, Y_1) \phi(X, Y_1)}_{\tau_1(X, Z)} \\ &\quad \vdots \\ &\quad \underbrace{\sum_{Y_1} \delta(X, Z, Y_1)}_{\tau_1(X, Z)} \end{aligned}$$

eliminate  $X$

$$\begin{aligned} &\sum_X \tau_1(X, Z) \tau_2(X, Z) \dots \tau_m(X, Z) \\ &\sum_X \delta_{n+1}(X, Z) \end{aligned}$$

# Variable Elimination Limitations



K. N. Toosi  
University of Technology

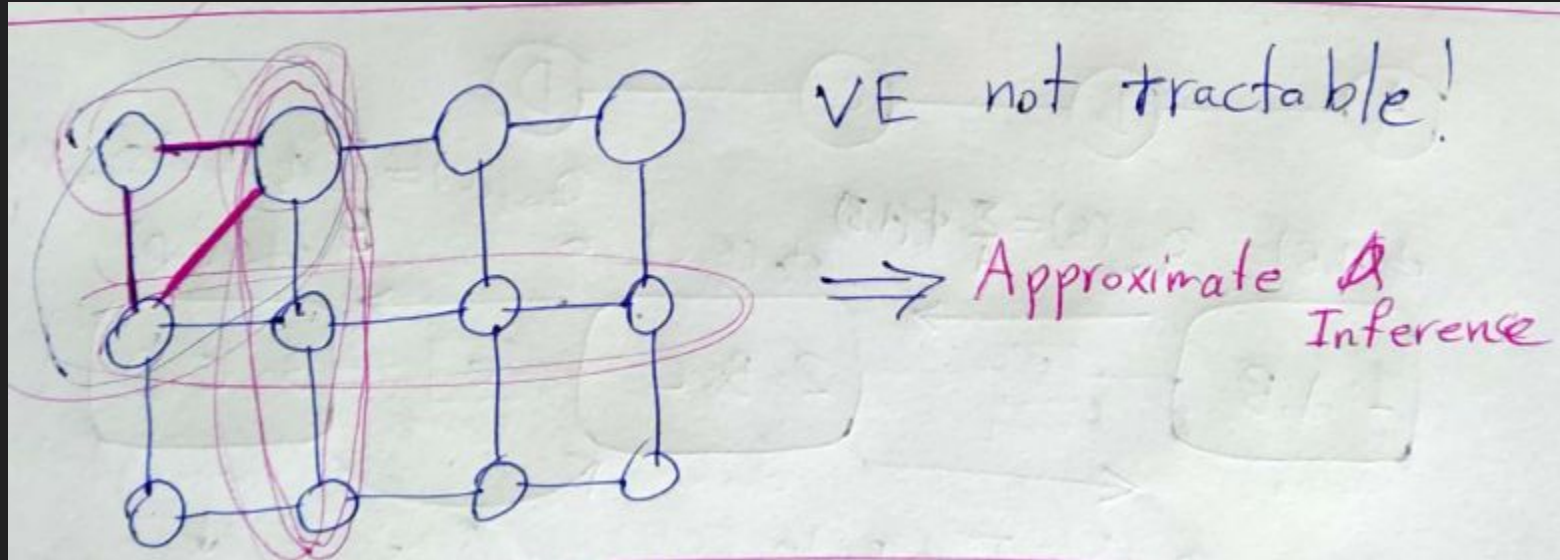
Fully Connected  
Bipartite Graph

$$P(X_1, \dots, X_m, Y_1, \dots, Y_n)$$
$$= \frac{1}{Z} \prod_{i=1}^m \prod_{j=1}^n \phi_{ij}(X_i, Y_j)$$

# Variable Elimination Limitations



K. N. Toosi  
University of Technology



# Inference



- Exact
  - Variable Elimination
  - Message Passing - Junction Tree
  - Graph-cuts
- Approximate
  - Message Passing - Loopy Belief Propagation
  - Graph-cut based
  - Sampling Based
  - Variational Inference

# Variable Elimination



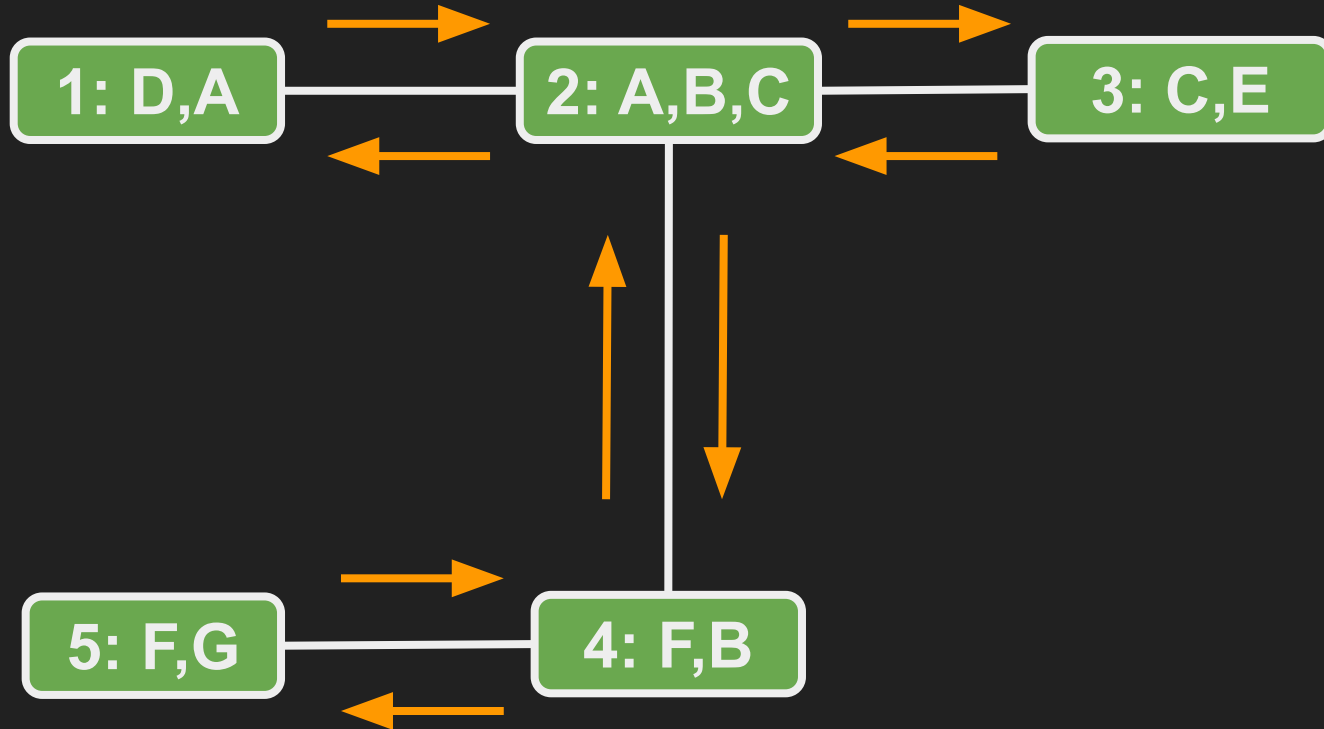
**K. N. Toosi**  
University of Technology

- Cannot compute multiple marginals using VE

# Message Passing

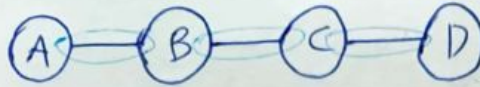


K. N. Toosi  
University of Technology





# Message Passing



17 (II)

$$P(A, B, C, D) = \frac{1}{2} \phi_1(A, B) \phi_2(B, C) \phi_3(C, D)$$

$$\tilde{P}(D) = \sum_C \sum_B \sum_A \phi_1(A, B) \phi_2(B, C) \phi_3(C, D)$$

Elim. A =  $\sum_C \sum_B \phi_2(B, C) \phi_3(C, D) \left[ \sum_A \phi_1(A, B) \right]$

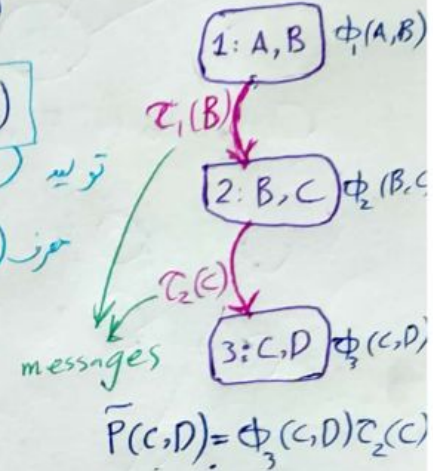
=  $\sum_C \sum_B \phi_2(B, C) \phi_3(C, D) \tau_1(B)$

*تولید* (Production) and *حذف* (Removal) arrows point to the  $\tau_1(B)$  term.

Elim. B =  $\sum_C \phi_3(C, D) \left[ \sum_B \phi_2(B, C) \tau_1(B) \right]$

Elim. C =  $\sum_C \phi_3(C, D) \tau_2(C)$

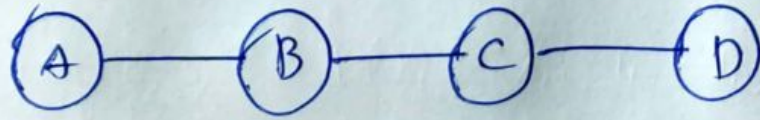
=  $\tau_3(D)$



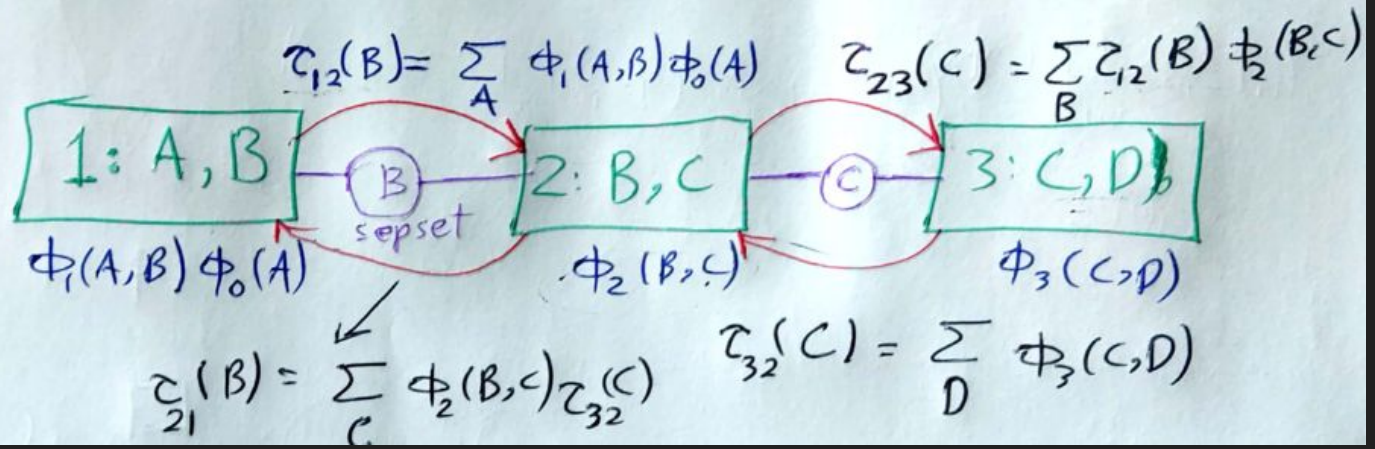
# Message Passing



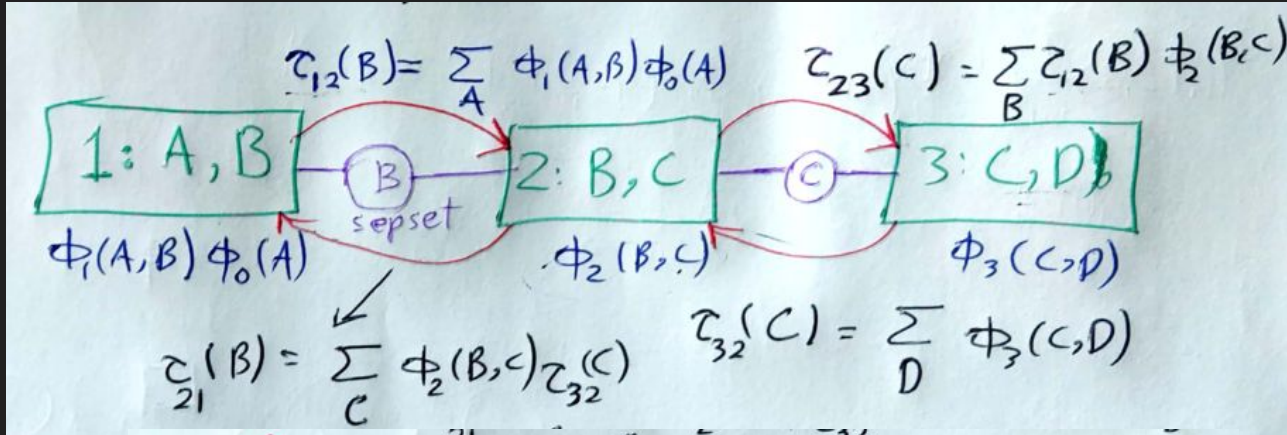
pgm 13 (II)



$$P(A, B, C, D) = \phi_1(A, B) \phi_2(B, C) \phi_3(C, D) \phi_0(A)$$



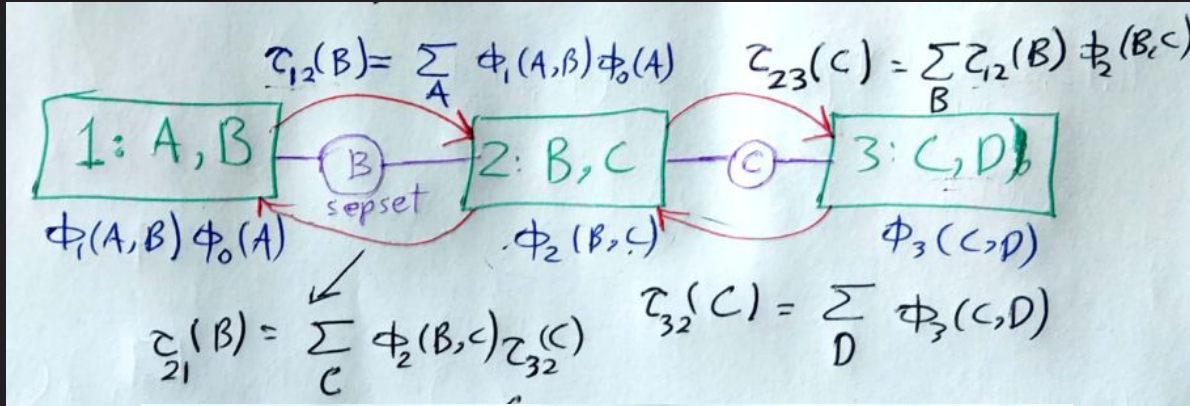
# Cluster Beliefs



belief

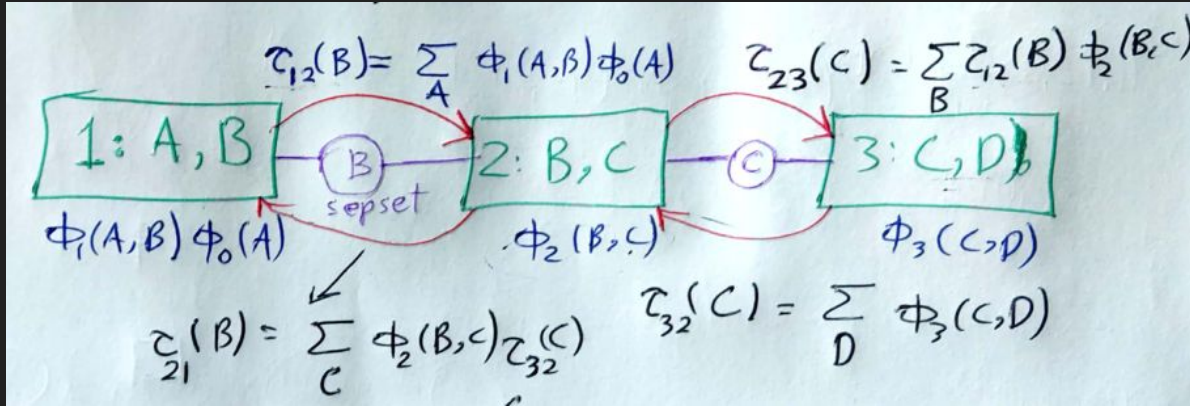
$$\begin{aligned}
 \beta(B, C) &= \phi_2(B, C) \tau_{12}(B) \tau_{32}(C) \\
 &= \phi_2(B, C) \left( \sum_A \phi_1(A, B) \phi_0(A) \right) \left( \sum_D \phi_3(C, D) \right) \\
 &= \sum_A \sum_D \phi_0(A) \phi_1(A, B) \phi_2(B, C) \phi_3(C, D) \\
 &= \sum_A \sum_D P(A, B, C, D) = P(B, C)
 \end{aligned}$$

# Cluster Beliefs



$$\begin{aligned}
 P(C, D) &= \phi_3(C, D) \tau_{23}(C) = \\
 &= \phi_3(C, D) \sum_B \tau_{12}(B) \phi_2(B, C) \\
 &= \phi_3(C, D) \sum_B \left[ \sum_A \phi_1(A, B) \phi_0(A) \right] \phi_2(B, C) \\
 &= \sum_A \sum_B \phi_3(C, D) \phi_1(A, B) \phi_0(A) \phi_2(B, C) \\
 &= \sum_A \sum_B P(A, B, C, D) = P(C, D)
 \end{aligned}$$

# Sepset Beliefs



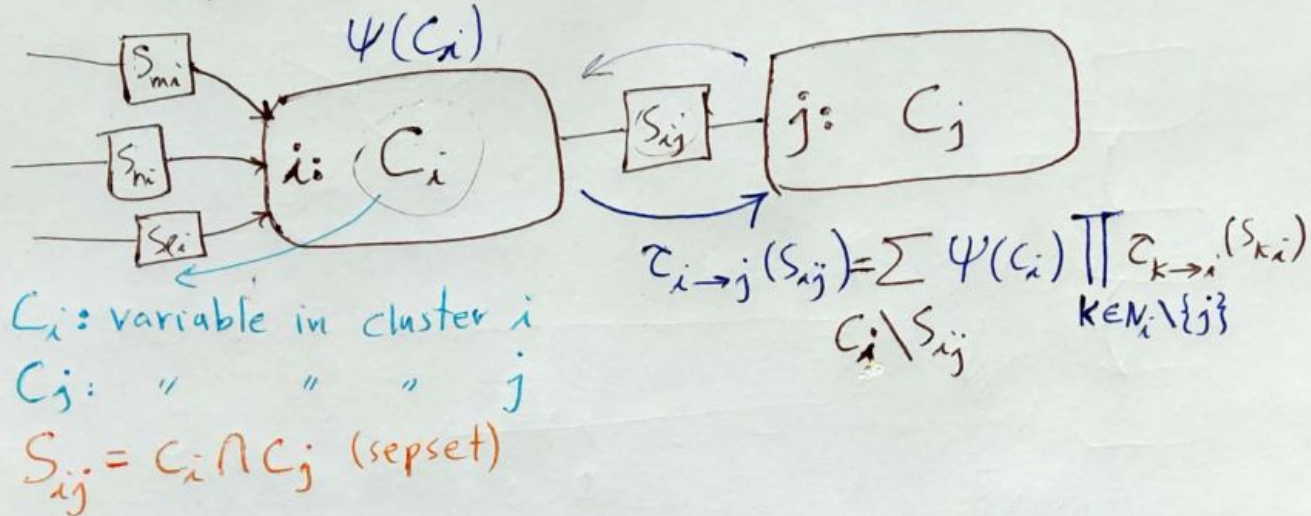
$$\begin{aligned}
 \beta(B) &= \tau_{12}(B) \tau_{21}(B) = \left[ \sum_A \phi_1(A, B) \phi_0(A) \right] \left[ \sum_C \phi_2(B, C) \tau_{32}(C) \right] \\
 &= \left[ \sum_A \phi_1(A, B) \phi_0(A) \right] \left[ \sum_C \phi_2(B, C) \right] \left[ \sum_D \phi_3(C, D) \right] \\
 &= \sum_A \sum_C \sum_D \phi_1(A, B) \phi_0(A) \phi_2(B, C) \phi_3(C, D) \\
 &= \sum_A \sum_C \sum_D P(A, B, C, D) = P(B)
 \end{aligned}$$

# Message Passing



- 1: A set of clusters
- 2: Assign Potentials to each cluster
- 3: Pass messages
- 4: Compute the Beliefs

17 III



# Message Passing



at the beginning  
we can only  
pass  $1 \rightarrow 2$   
 $6 \rightarrow 4$   
 $5 \rightarrow 4$   
 $3 \rightarrow 2$

cluster graph  
(Junction tree)

A message from cluster  $i$  to cluster  $j$  can only be sent if all incoming message to cluster  $i$  (except the one from cluster  $j$ ) are sent (are ready).

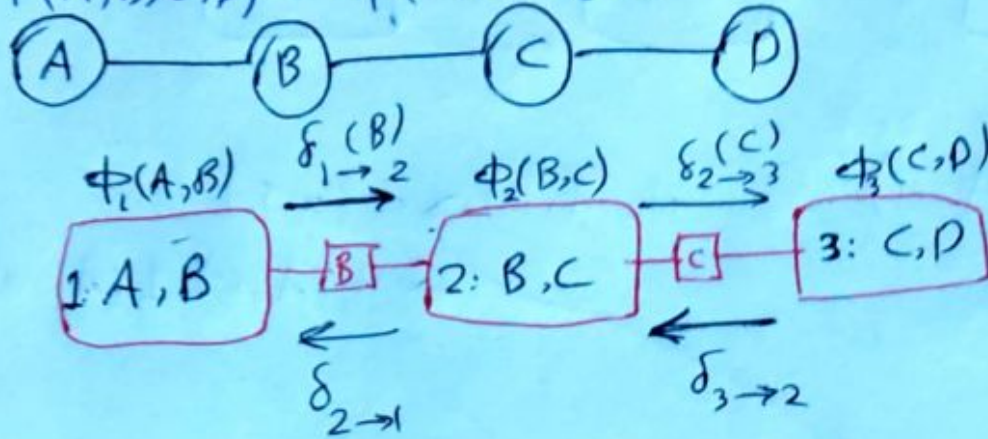
⇒ the cluster graph must be a tree (Junction-tree algorithm Exact Inference)  
start from the leaves of the tree.

# Junction Tree



Junction tree { tree graph  
nodes are clusters

$$P(A, B, C, D) = \phi_1(A, B) \phi_2(B, C) \phi_3(C, D)$$



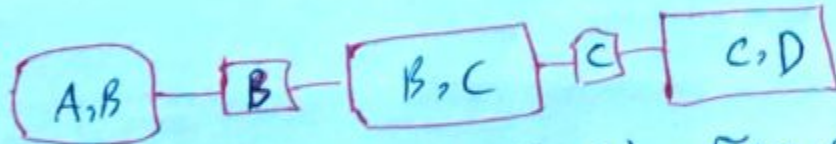
Belief propagation



# inference with evidence



What if we have evidence?  $P(X_t | X_e = x_e)$   
 case 1: if  $X_t, X_e$  are in the same cluster



$P(D | C=c) \Rightarrow ?$  compute  $\beta(C,D) = \tilde{P}(C,D) = \sum P(C,D)$

$$\tilde{P}(D, C=c) = \beta(D, c)$$

$$P(D | C=c) = \frac{\beta(D, c)}{\sum_{D'} \beta(D', c)} = \frac{\tilde{P}(D, c)}{\sum_D \tilde{P}(D, c)}$$

# inference with evidence



18 (II)

Diagram illustrating inference with evidence in a Hidden Markov Model (HMM) with three states:

- State 1:  $A, B$
- State 2:  $B, C$
- State 3:  $C, D$

Transitions and potentials:

- Transition 1:  $\phi_1(A, B)$  with evidence  $B$  and delta function  $\delta_{1 \rightarrow 2}(B)$
- Transition 2:  $\phi_2(B, C)$  with evidence  $C$  and delta function  $\delta_{2 \rightarrow 3}(C)$
- Transition 3:  $\phi_3(C, D)$  with evidence  $C$  and delta function  $\delta_{3 \rightarrow 2}(C)$

Joint probability calculation:

$$P(C, D | A=a) = \phi_1(A, B) \phi_3(C, D)$$

Revised expression with evidence:

$$\phi_1(A, B) \mathbb{1}(A=a)$$

Old delta function:

$$\delta_{1 \rightarrow 2}(B) = \sum_A \phi_1(A, B)$$

New delta function:

$$\delta_{1 \rightarrow 2}(B) = \phi_1(a, B)$$

Derivation:

$$= \sum_A \phi_1(A, B) \mathbb{1}(A=a) = \phi_1(a, B)$$

only need to recompute  $\delta_{1 \rightarrow 2}, \delta_{2 \rightarrow 3}$

Final expression for conditional probability:

$$P(B, C | A=a) = \sqrt{\delta_{1 \rightarrow 2}(B)}$$

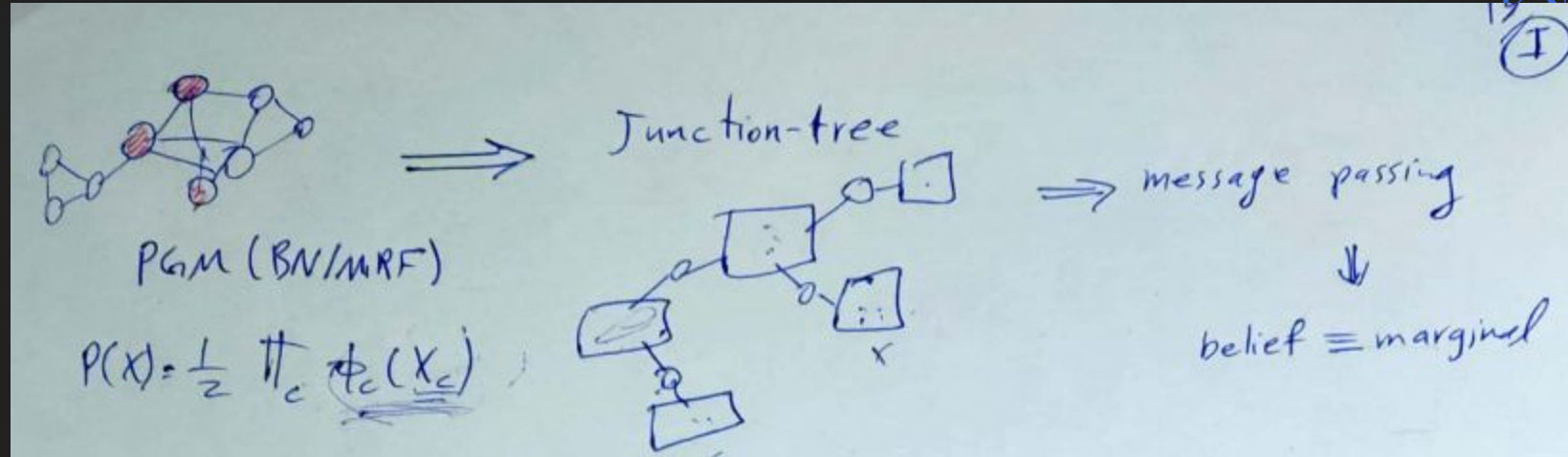
# Junction tree



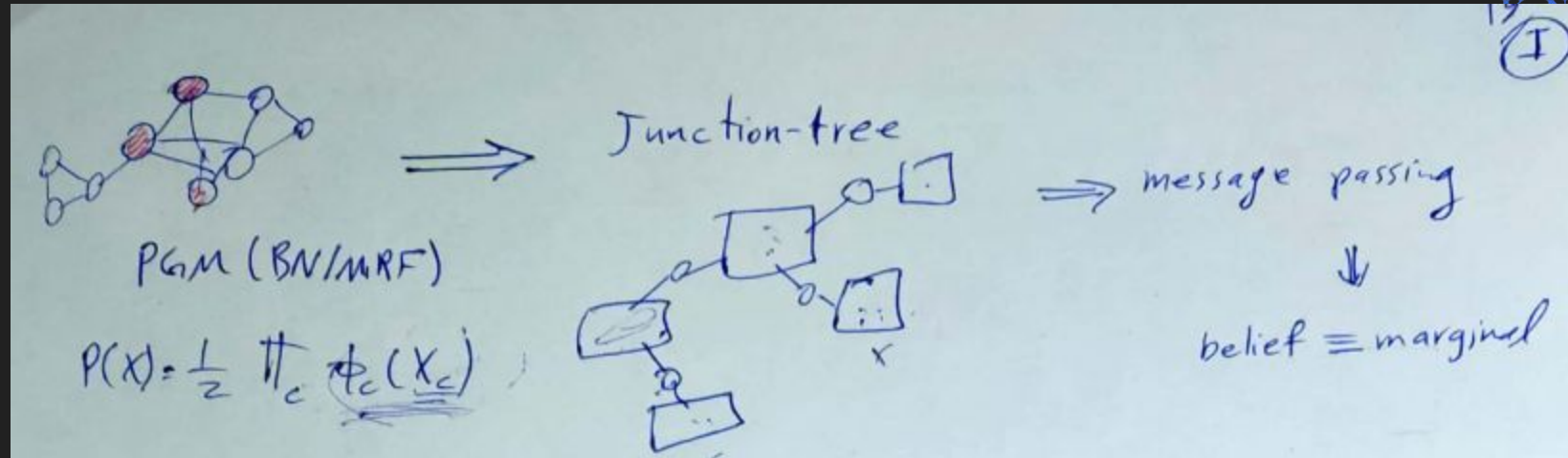
**K. N. Toosi**  
University of Technology

Message Passing is applicable only when the clusters are small  
(each cluster has few variables)

# How to build a cluster tree?



# How to build a cluster tree?



1. Using variable elimination

# Build cluster tree using VE



18 III

$P(C, D, I, G, S, L, J, H)$   
 $= P(C) P(D|C) P(I) P(G|D, I)$   
 $P(S|I) P(L|G) P(J|L, S)$   
 $P(H|G, J)$

$= \phi_1(C) \phi_2(D, C) \phi_3(I) \phi_4(G, D, I)$   
 $\phi_5(S, I) \phi_6(L, G) \phi_7(J, L, S)$   
 $\phi_8(H, G, J)$

Eliminate: C, D, I, H, G, S, L

# Build cluster tree using VE



K. N. Toosi  
University of Technology

18 III

$P(C, D, I, G, S, L, J, H)$   
 $= P(C) P(D|C) P(I) P(G|D, I)$   
 $P(S|I) P(L|G) P(J|L, S)$   
 $P(H|G, J)$

$= \phi_1(C) \phi_2(D, C) \phi_3(I) \phi_4(G, D, I)$   
 $\phi_5(S, I) \phi_6(L, G) \phi_7(J, L, S)$   
 $\phi_8(H, G, J)$

Eliminate: C, D, I, H, G, S, L

# How to build a cluster tree?



**K. N. Toosi**  
University of Technology

1. Using variable elimination
2. Determine valid cluster trees



# How to build a cluster tree?



Must possess two properties

## 1- Family Preservation

Each factor can be assigned to some cluster.

for each factor  $\phi_c(X_c)$  there must be a cluster  $C_i$  with variables  $X_{C_i}$  such that  $X_c \subseteq X_{C_i}$ .

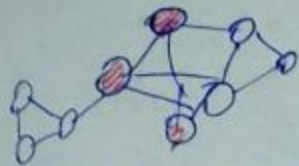
## 2- Running Intersection Property

for each variable  $X$  the subgraph containing of the cluster-tree containing  $X$  ~~is~~ is connected ( $\equiv$  also forms a tree).

# How to build a cluster tree?

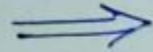


K. N. Toosi  
Technology

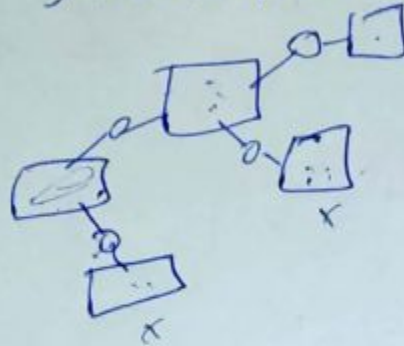


PGM (BN/MRF)

$$P(X) = \frac{1}{Z} \prod_c \phi_c(X_c)$$



Junction-tree



message passing



belief  $\equiv$  marginal

- 1- Family Preservation Property
- 2- Running Intersection Property (RIP)