## **From Bucket-Elimination To Bucket Trees**



Definition: *T* is a bucket tree.

Theorem: *T* is an i-map of *G*.

- Variable-elimination can be viewed as message-passing (elimination) using a rooted bucket tree.
- Any variable (bucket) can be the root.

## **Generalization:Eliminate (sum over)** Variables Not in Separators



• Multiply all incoming messages, and  $P_i$ 's in the bucket and sum over  $B_1 \cap B_2$ .

• 
$$\boldsymbol{I}_{B_1}^{B_2}(s) = \sum_{B_1-s} (\Pi \boldsymbol{I}_i) \cdot (\Pi \boldsymbol{P}_i)$$

- Given a rooted bucket tree, *T*, every node can be the "root" of the variables-elimination computation.
- If B<sub>3</sub> is the root, bucket B<sub>2</sub> and then Bucket B<sub>1</sub> should be processed; π-messages sent from B<sub>2</sub> to B<sub>1</sub> and from B<sub>1</sub> To B<sub>3</sub>

#### **Bucket Propagation Algorithm**

- Input: A bucket tree  $B_1 \dots B_n$
- Output: For Each  $B_i$  and parent  $B_j$ , functions  $\lambda_i^j(S_{ij})$  and  $\pi_i^j(S_{ij})$  are exchanged.



Top Down:

• Let  $s \ \boldsymbol{l}_1 \dots \boldsymbol{l}_k$  messages from child nodes of  $B_i$ ,  $P_1 \dots P_r$  in  $B_i$  original functions.

• 
$$\boldsymbol{l}_{i}^{j}(S_{ij}) = \sum_{B_{i}-B_{j}} \prod_{i} \boldsymbol{l}_{i} \bullet \prod_{j} P_{j}$$

Bottom Up:

• Let  $\pi_{i_{j}}^{i}$  be received from  $B_{j}$ .

• 
$$\mathbf{p}_{i}^{k}(S_{ki}) = \sum_{B_{k}-B_{i}} (\prod_{j} P_{j}) \bullet \mathbf{p}_{j}^{i} \bullet \prod_{i \neq k} \mathbf{l}_{i}$$

- The belief of  $B_i$
- $P(B_i) = \prod_i P_j \bullet \prod_i \mathbf{l}_i \bullet \mathbf{p}_j^i$
- if x index Bucket i get Bel(x) by summing out Bel(x) =  $\alpha \sum_{S_{ij}} P(B_i)$

# **Propagation in a Bucket Tree**

Definitions:

- Let *G* be a Bayesian network, *d*, an ordering and  $B_1 \dots B_n$  the final bucket created processing along  $d = x_1 \dots x_n$ .
- Let  $B_i$  be the set of variables appearing in bucket *i* when it is processed.

Bucket Tree:

• A bucket tree has each  $B_i$  cluster as a node and there is an arc from  $B_i$  to  $B_j$  if the function created at  $B_i$  was placed in  $B_j$ 

Graph-Based Definition:

Let G<sub>d</sub> be the induced graph along d. Each variable x and it's earlier neighbors in a node, B<sub>x</sub>. There is an arc from B<sub>x</sub> to B<sub>y</sub> if y is the closest parent of x.

## **Upwards Messages On The Bucket Tree**



$$\Pi(A) = P(A)$$
  

$$\Pi_{B}^{P}(A, B) = P(B, A) \bullet \boldsymbol{I}_{C}^{B}(A, B)$$
  

$$\Pi_{B}^{C}(A, B) = P(B, A) \bullet \Pi(A) \bullet \boldsymbol{I}_{D}^{B}(A, B)$$
  

$$\Pi_{C}^{E}(B, C) = \sum_{A} P(C, A) \bullet \Pi_{B}^{C}(A, B)$$