## **Junction tree propagation - BNDG 4-4.6**

Finn V. Jensen and Thomas D. Nielsen

### Message Passing in Join Trees

More sophisticated inference technique; used in most implemented Bayesian Network systems (e.g. Hugin).

**Overview:** 

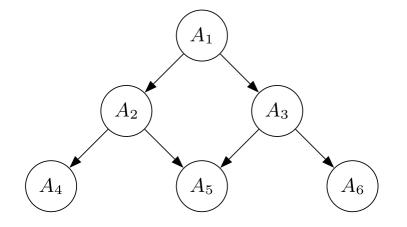
Given: Bayesian network for  $P(\mathbf{V})$ 

 $\Downarrow$  Preprocessing

### Construct a new internal representation for $P(\mathbf{V})$ called a *junction tree*

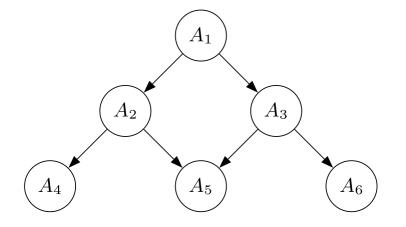
↓ Inference/Updating

Retrieve  $P(A \mid \mathbf{E} = \mathbf{e})$  for single  $A \in \mathbf{V}$ 



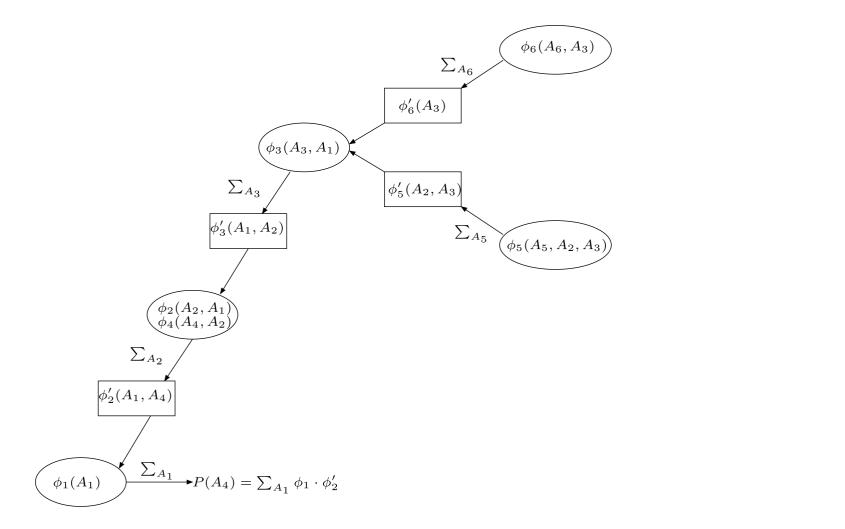
### $P(A_4)$

 $=\sum_{A_1}\sum_{A_2}\sum_{A_3}\sum_{A_5}\sum_{A_5}\sum_{A_6}\phi_1(A_1)\phi_2(A_2,A_1)\phi_3(A_3,A_1)\phi_4(A_4,A_2)\phi_5(A_5,A_2,A_3)\phi_6(A_6,A_3)$ 

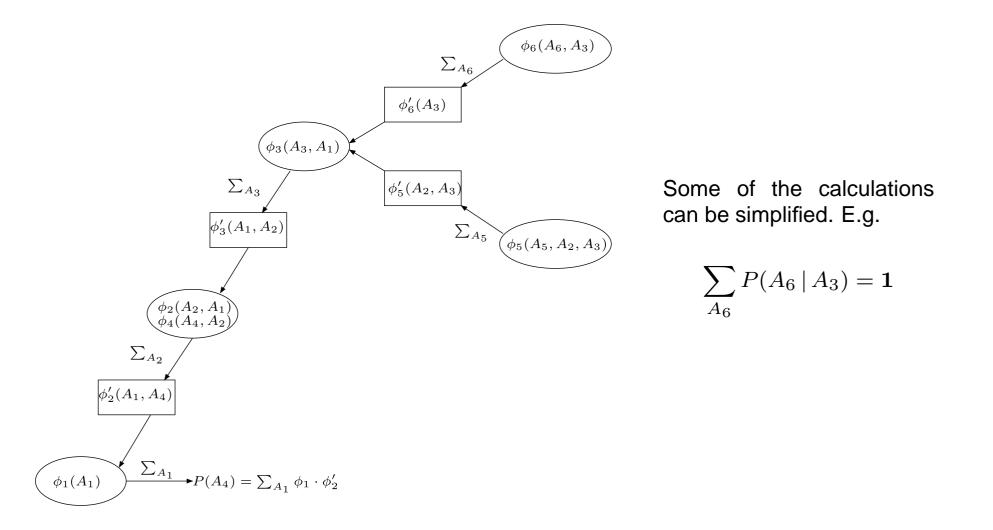


### $P(A_4)$

 $= \sum_{A_1} \sum_{A_2} \sum_{A_3} \sum_{A_5} \sum_{A_6} \phi_1(A_1) \phi_2(A_2, A_1) \phi_3(A_3, A_1) \phi_4(A_4, A_2) \phi_5(A_5, A_2, A_3) \phi_6(A_6, A_3)$  $= \sum_{A_1} \phi_1(A_1) \sum_{A_2} \phi_2(A_2, A_1) \phi_4(A_4, A_2) \sum_{A_3} \phi_3(A_3, A_1) \sum_{A_5} \phi_5(A_5, A_2, A_3) \sum_{A_6} \phi_6(A_6, A_3)$ 

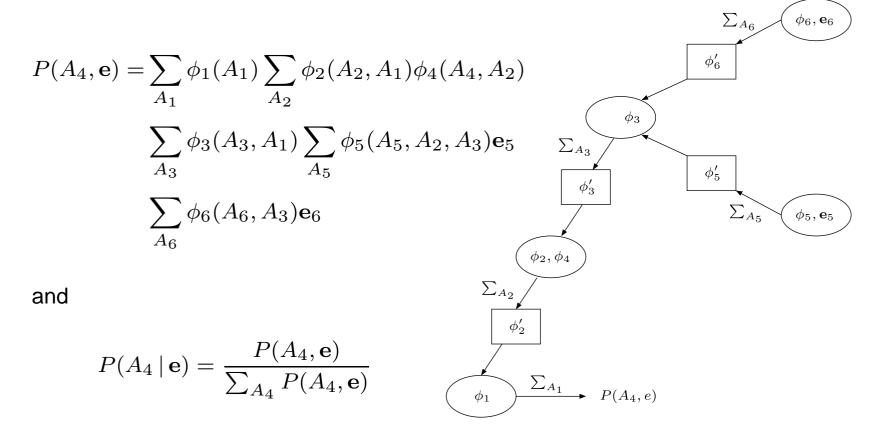


 $P(A_4) = \sum_{A_1} \phi_1(A_1) \sum_{A_2} \phi_2(A_2, A_1) \phi_4(A_4, A_2) \sum_{A_3} \phi_3(A_3, A_1) \sum_{A_5} \phi_5(A_5, A_2, A_3) \sum_{A_6} \phi_6(A_6, A_3)$ 

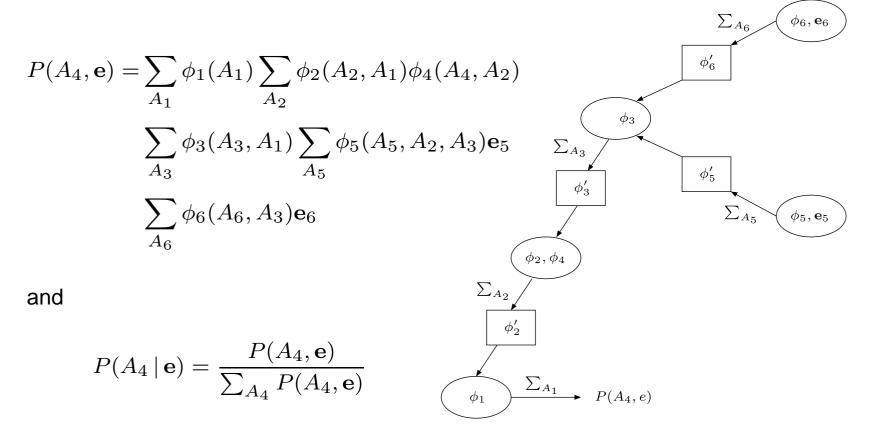


 $P(A_4) = \sum_{A_1} \phi_1(A_1) \sum_{A_2} \phi_2(A_2, A_1) \phi_4(A_4, A_2) \sum_{A_3} \phi_3(A_3, A_1) \sum_{A_5} \phi_5(A_5, A_2, A_3) \sum_{A_6} \phi_6(A_6, A_3)$ 

Assume evidence  $\mathbf{e} = (\mathbf{e}_5, \mathbf{e}_6)$ ; represented as 0/1 potentials. Then:

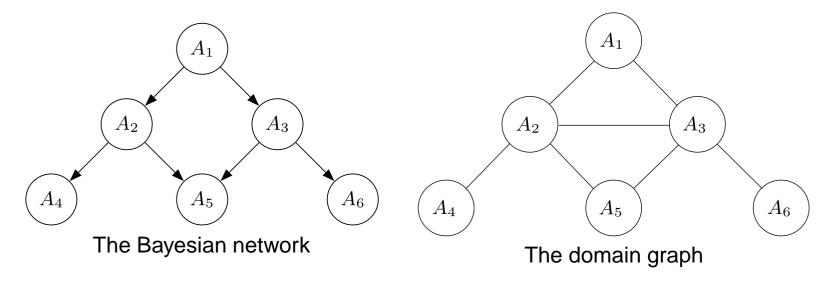


Assume evidence  $e = (e_5, e_6)$ ; represented as 0/1 potentials. Then:



The process is sufficiently general to handle all evidence scenarios!

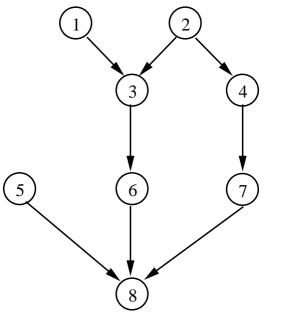
 $\Rightarrow$  We look for a general structure in which these calculations can be performed for all variables.



The link  $(A_2, A_3)$  is called a moral link.

#### **Moralization in general**

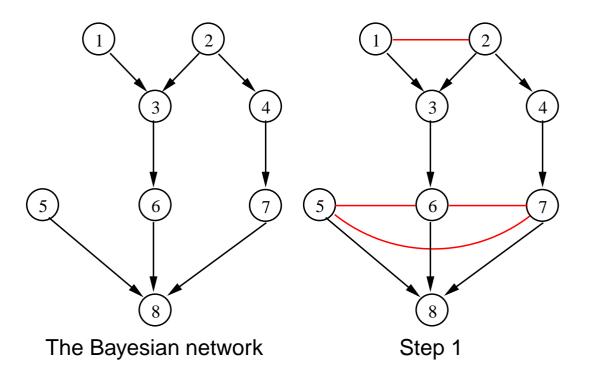
- For all nodes X: Connect pairwise all parents of X with undirected links.
- Replace all original directed links by undirected ones.



The Bayesian network

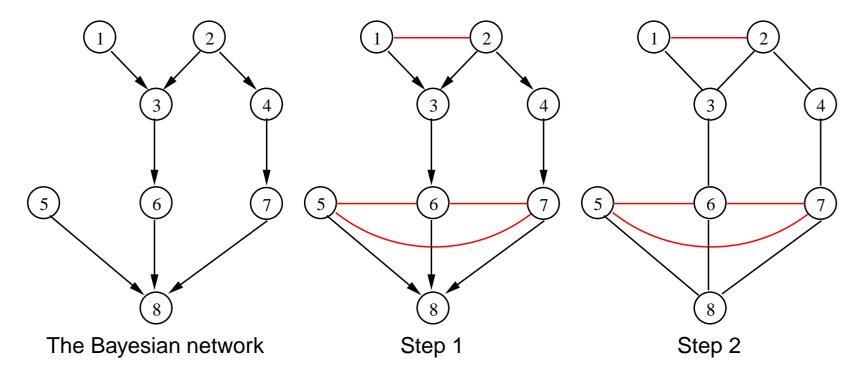
#### **Moralization in general**

- For all nodes X: Connect pairwise all parents of X with undirected links.
- Replace all original directed links by undirected ones.



#### **Moralization in general**

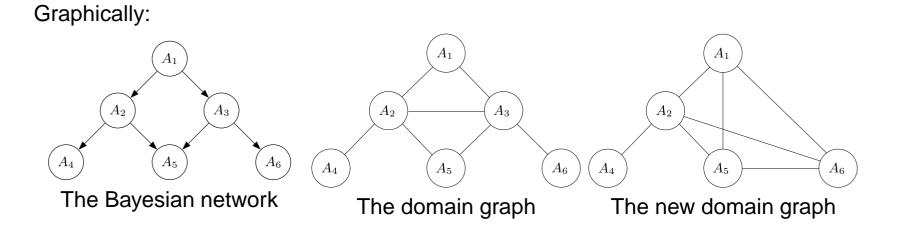
- For all nodes X: Connect pairwise all parents of X with undirected links.
- Replace all original directed links by undirected ones.



# **Eliminating a variable**

Suppose that when calculating  $P(A_4)$  we start off by eliminating  $A_3$ :

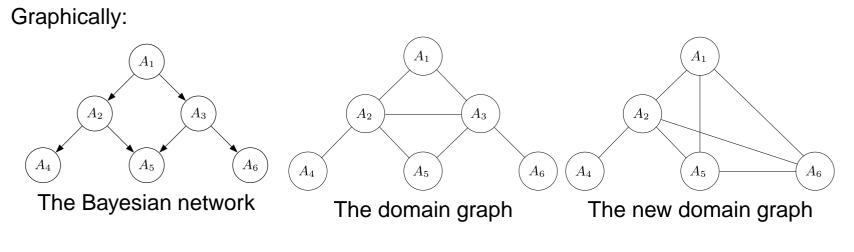
$$\phi'(A_1, A_2, A_5, A_6) = \sum_{A_3} \phi_3(A_3, A_1) \phi_4(A_4, A_2) \phi_5(A_5, A_2, A_3).$$



# **Eliminating a variable**

Suppose that when calculating  $P(A_4)$  we start off by eliminating  $A_3$ :

$$\phi'(A_1, A_2, A_5, A_6) = \sum_{A_3} \phi_3(A_3, A_1) \phi_4(A_4, A_2) \phi_5(A_5, A_2, A_3).$$



#### **Perfect elimination sequences**

If all variables can be eliminated without introducing fill-in edges, then the elimination sequence is called perfect. An example could be  $A_5, A_6, A_3, A_1, A_2, A_4$ .

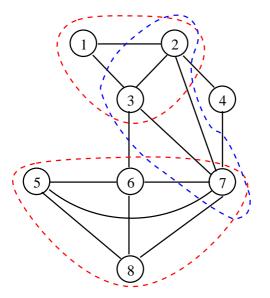
## **Perfect elimination sequences**

#### **Perfect elimination sequences**

All perfect elimination sequences produce the same domain set, namely the set of cliques of the domain graph.

#### Clique

A set of nodes is complete if all nodes are pairwise linked. A complete set is a clique if it is not a subset of another complete set.

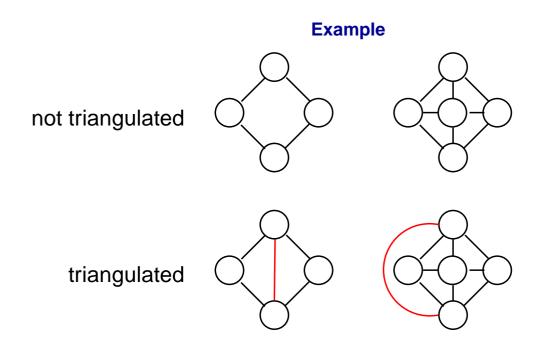


#### **Property**

Any perfect elimination sequence ending with A is optimal w.r.t. calculating P(A).

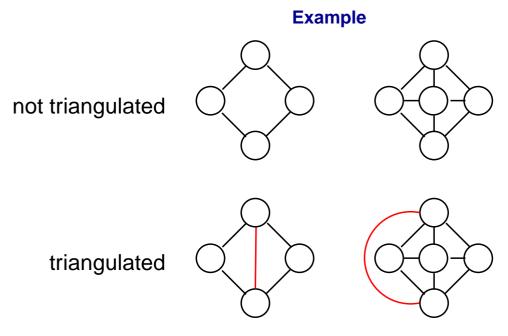
#### **Definition**

An undirected graph with a perfect elimination sequence is called a triangulated graph.



#### **Definition**

An undirected graph with a perfect elimination sequence is called a triangulated graph.

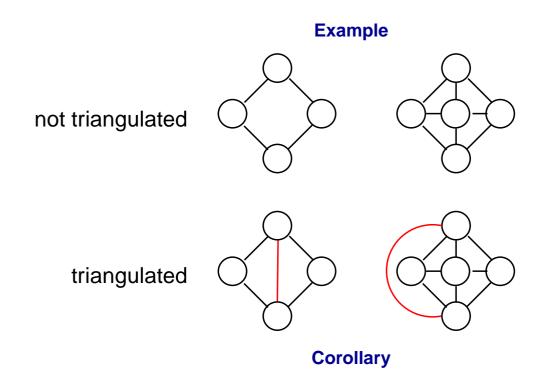


Corollary

A graph is triangulated iff all nodes can successively be eliminated without introducing fill-in edges.

#### **Definition**

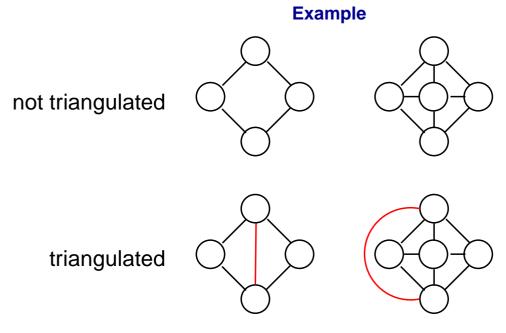
An undirected graph with a perfect elimination sequence is called a triangulated graph.



A graph is triangulated iff there does not exist a cycle of length  $\geq 4$  that is not cut by a cord.

#### **Definition**

An undirected graph with a perfect elimination sequence is called a triangulated graph.

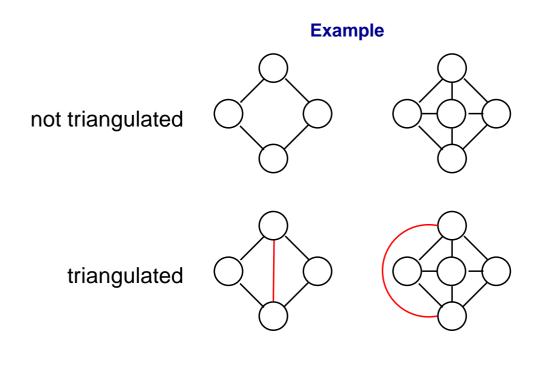


Corollary

In a triangulated graph, each variable A has a perfect elimination sequence ending with A.

#### **Definition**

An undirected graph with a perfect elimination sequence is called a triangulated graph.



Assumption

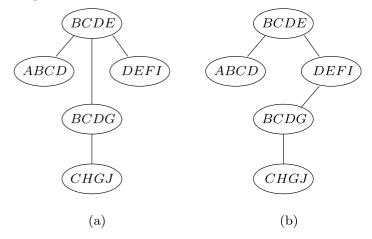
For now we shall assume that the domain graph is triangulated!

### Join Tree (Junction Tree)

Let  $\mathcal{G}$  be the set of cliques from an undirected graph, and let the cliques of  $\mathcal{G}$  be organized in a tree T. Then T is a join tree if

a variable V that is contained in two nodes C, C' also is contained in every node on the (unique) path connecting C and C'.

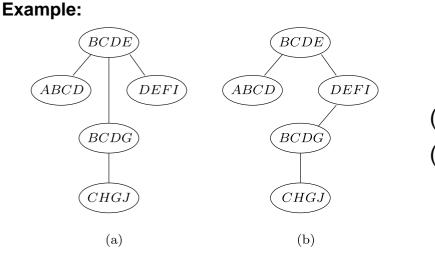
#### **Example:**

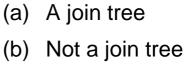


### Join Tree (Junction Tree)

Let  $\mathcal{G}$  be the set of cliques from an undirected graph, and let the cliques of  $\mathcal{G}$  be organized in a tree T. Then T is a join tree if

a variable V that is contained in two nodes C, C' also is contained in every node on the (unique) path connecting C and C'.

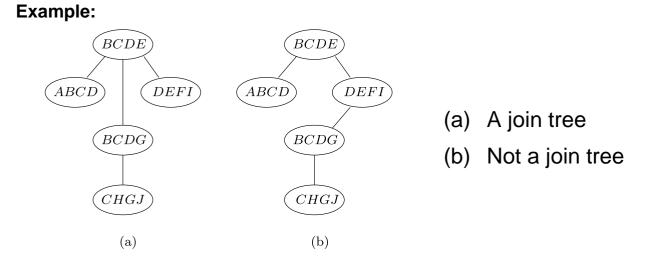




### Join Tree (Junction Tree)

Let  $\mathcal{G}$  be the set of cliques from an undirected graph, and let the cliques of  $\mathcal{G}$  be organized in a tree T. Then T is a join tree if

a variable V that is contained in two nodes C, C' also is contained in every node on the (unique) path connecting C and C'.



**Theorem:** An undirected graph G is triangulated if and only if the cliques of G can be organized in a join tree.

### From undirected graph to join graph

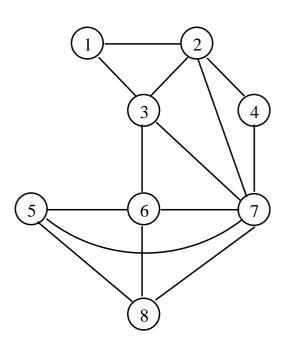
Define undirected graph  $(\mathcal{C}, \mathcal{E}^*)$ :

C: Set of cliques in triangulated moral graph

$$\mathcal{E}^*: \{ (\mathbf{C}_1, \mathbf{C}_2) \mid \mathbf{C}_1 \cap \mathbf{C}_2 \neq \emptyset \}$$

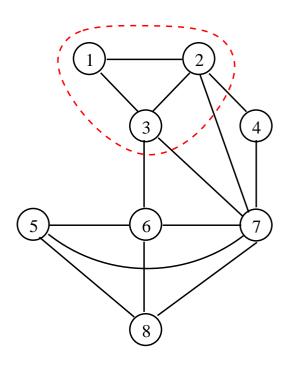
### From undirected graph to join graph

- $\mathcal{C}:\quad \text{Set of cliques in triangulated moral graph}$
- $\mathcal{E}^*: \quad \{(\mathbf{C}_1, \mathbf{C}_2) \mid \mathbf{C}_1 \cap \mathbf{C}_2 \neq \emptyset\}$



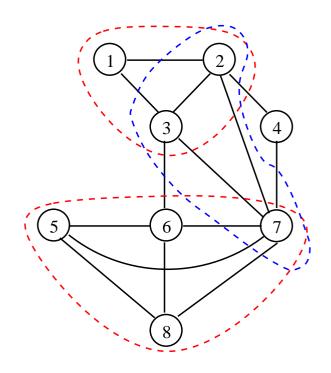
### From undirected graph to join graph

- $\mathcal{C}$ : Set of cliques in triangulated moral graph
- $\mathcal{E}^*: \{ (\mathbf{C}_1, \mathbf{C}_2) \mid \mathbf{C}_1 \cap \mathbf{C}_2 \neq \emptyset \}$



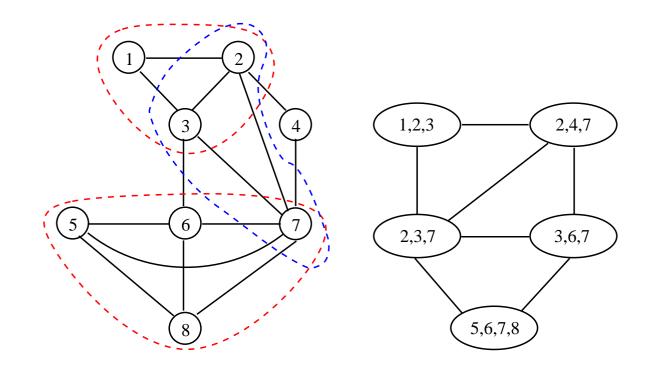
### From undirected graph to join graph

- C: Set of cliques in triangulated moral graph
- $\mathcal{E}^*: \{ (\mathbf{C}_1, \mathbf{C}_2) \mid \mathbf{C}_1 \cap \mathbf{C}_2 \neq \emptyset \}$



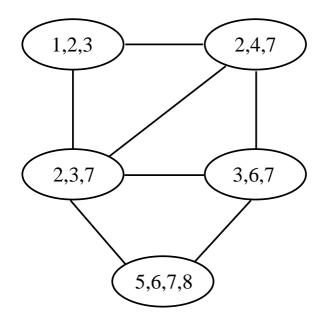
### From undirected graph to join graph

- $\mathcal{C}$ : Set of cliques in triangulated moral graph
- $\mathcal{E}^*: \{ (\mathbf{C}_1, \mathbf{C}_2) \mid \mathbf{C}_1 \cap \mathbf{C}_2 \neq \emptyset \}$

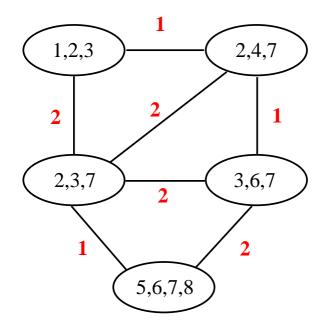


### From join graph to join tree

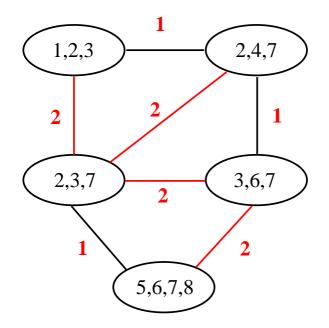
#### From join graph to join tree



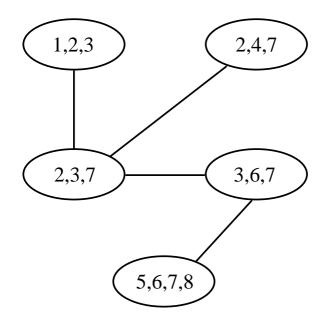
#### From join graph to join tree



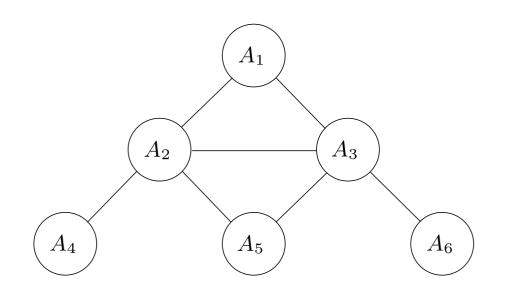
#### From join graph to join tree



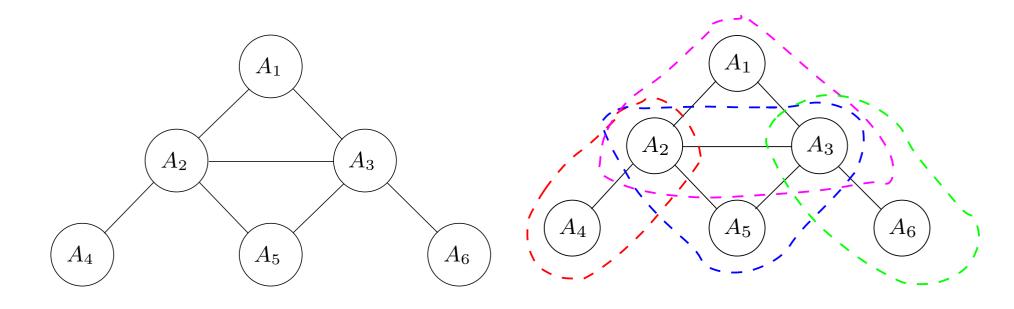
#### From join graph to join tree



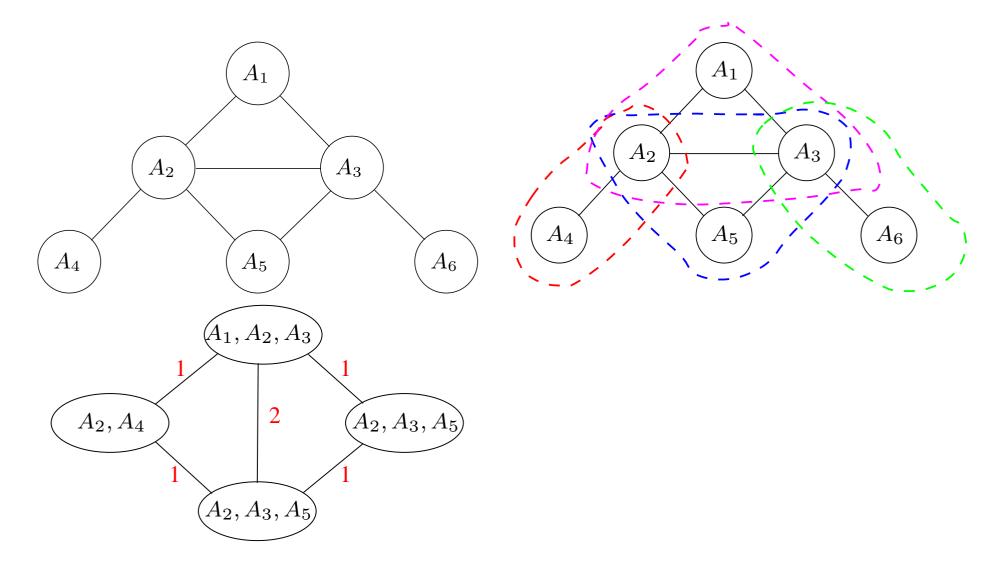




#### **Another example**

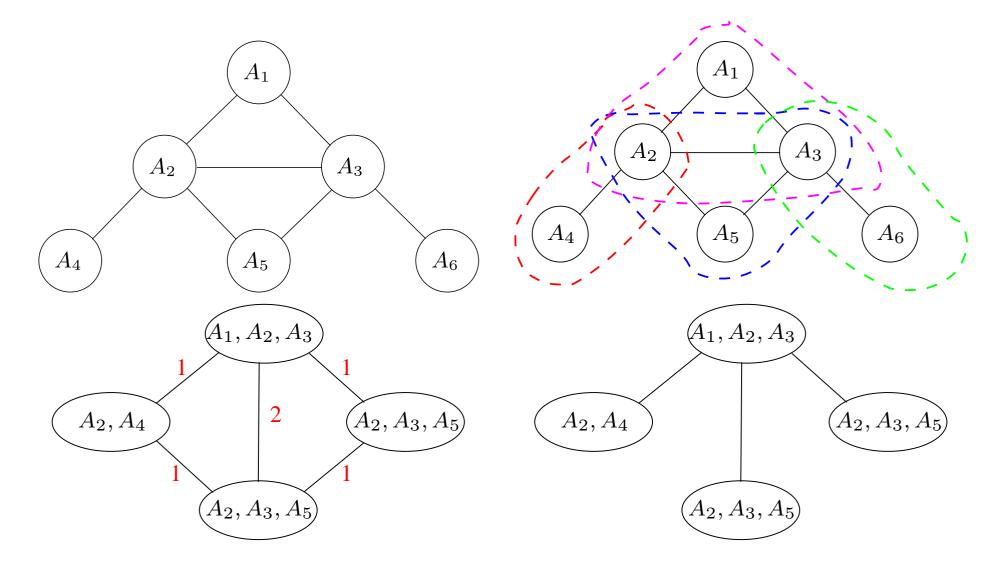


#### **Another example**



### Join tree construction

#### Another example



### **Junction tree construction**

#### Correctness

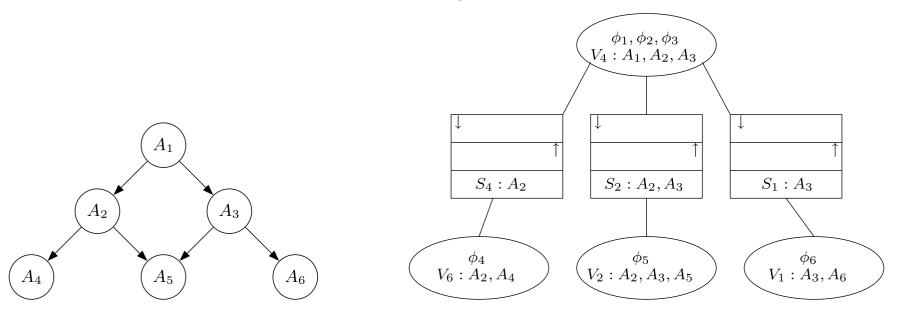
Let  $(\mathbf{V}, \rightarrow)$  be a triangulated graph,  $(\mathcal{C}, \mathcal{E}^*)$ . Every maximal spanning tree  $(\mathcal{C}, \mathcal{E})$  of  $(\mathcal{C}, \mathcal{E}^*)$  is a join tree for  $\mathbf{V}$  (Jensen, Jensen 1994).

If  $(\mathbf{V}, \rightarrow)$  is the triangulated moral graph of a Bayesian network, then there exists for every  $V \in \mathbf{V}$  a clique  $\mathbf{C} \in C$  with  $\{V\} \cup \operatorname{pa}(V) \subseteq \mathbf{C}$ .

#### Initialization

Let T be a join tree for the domain graph G with potentials  $\Phi$ . A junction tree for G consists of T with the additions:

- each potential  $\phi$  from  $\Phi$  is assigned to a clique containing dom ( $\phi$ ).
- each link has a separator attached.
- each separator contains two mailboxes, one for each direction.

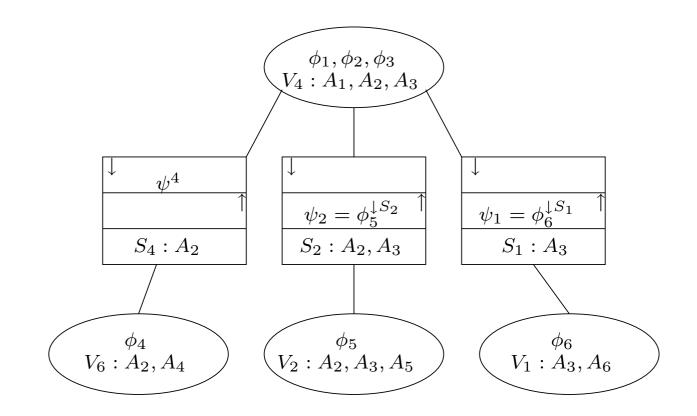


#### Example

#### **Propagation**

To calculate  $P(A_4)$  we find a clique ( $V_6$ ) containing  $A_4$  and send messages to that clique.

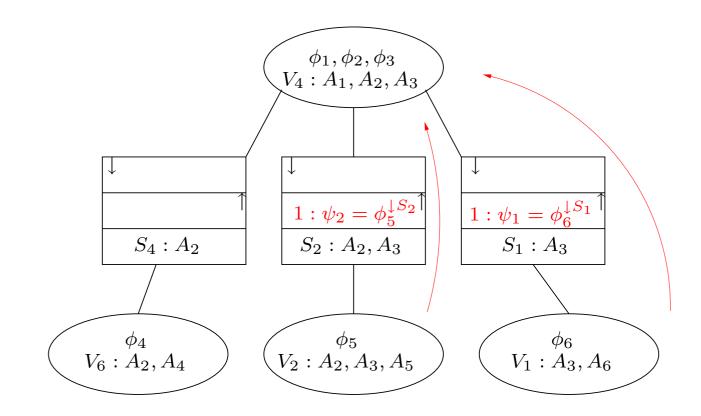
Example



#### **Propagation**

To calculate  $P(A_4)$  we find a clique ( $V_6$ ) containing  $A_4$  and send messages to that clique.

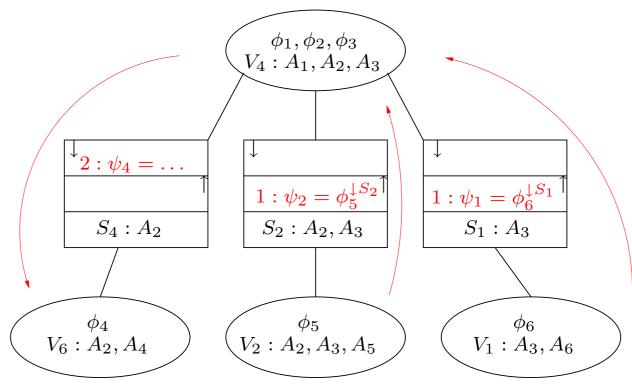
Example



#### **Propagation**

To calculate  $P(A_4)$  we find a clique ( $V_6$ ) containing  $A_4$  and send messages to that clique.

Example



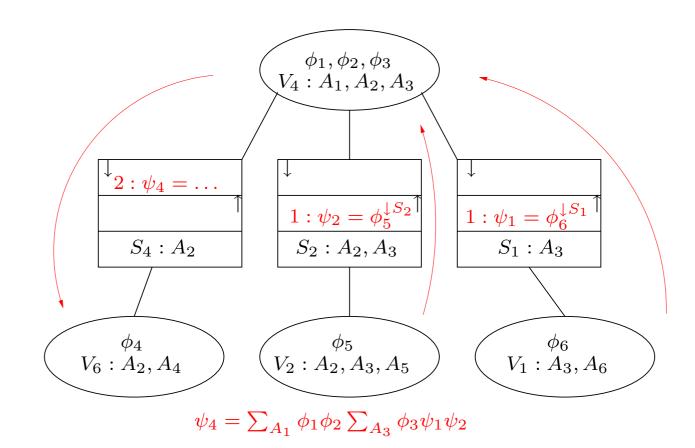
 $\psi_4 = \sum_{A_1} \phi_1 \phi_2 \sum_{A_3} \phi_3 \psi_1 \psi_2$ 

Now evidence has been collected to  $V_4$  and we get  $P(A_4) = \sum_{A_2} \phi_4 \psi_4$ .

#### **Propagation**

To calculate  $P(A_i)$  for any  $A_i$  we send messages away from  $V_6$ .

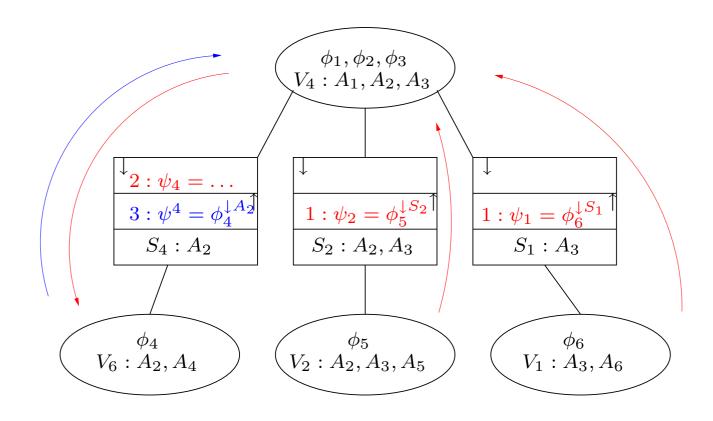
#### **Example continued**



#### **Propagation**

To calculate  $P(A_i)$  for any  $A_i$  we send messages away from  $V_6$ .

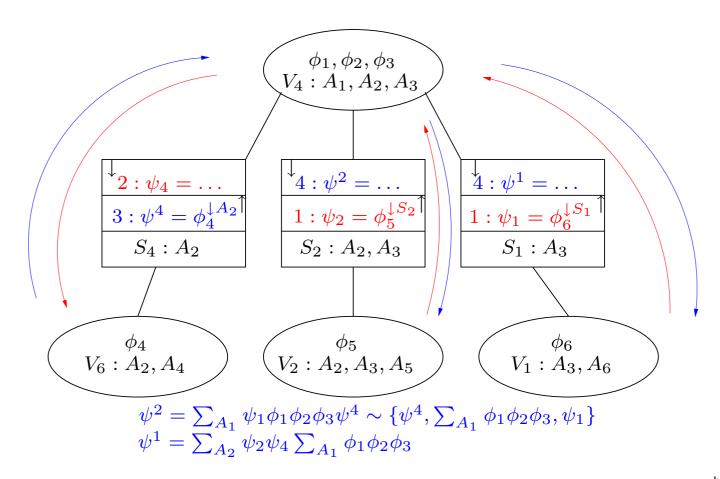
#### **Example continued**



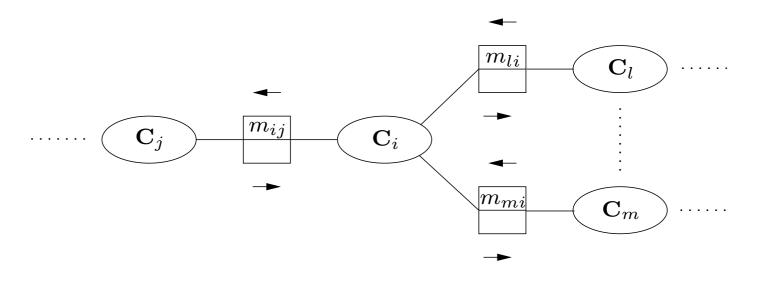
#### **Propagation**

To calculate  $P(A_i)$  for any  $A_i$  we send messages away from  $V_6$ .

#### **Example continued**

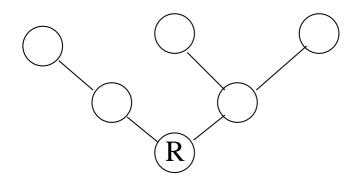


In general: Sending a message from  $\mathbf{C}_i$  to  $\mathbf{C}_j$ 





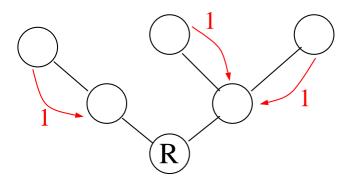
Messages passing in general



Collect and distribute messages

i) Collect messages to a preselected root R

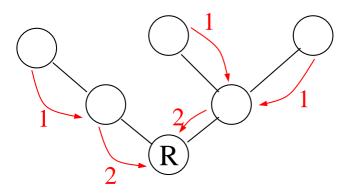
Messages passing in general



Collect and distribute messages

i) Collect messages to a preselected root R

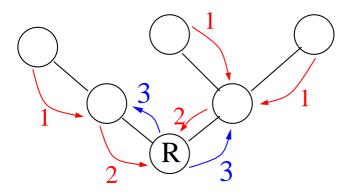
Messages passing in general



Collect and distribute messages

i) Collect messages to a preselected root R

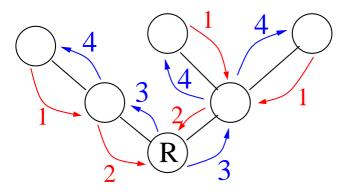
#### Messages passing in general



Collect and distribute messages

- i) Collect messages to a preselected root R
- ii) Distribute messages away from the root R

#### Messages passing in general



#### Collect and distribute messages

- i) Collect messages to a preselected root R
- ii) Distribute messages away from the root R

#### Theorem

Let the junction tree T represent the Bayesian network BN over the universe U and with evidence e. Assume that T is full.

1. Let *V* be a clique with set of potentials  $\Phi_V$ , and let  $S_1, \ldots, S_k$  be *V*'s neighboring separators and with *V*-directed messages  $\Psi_1, \ldots, \Psi_k$ . Then

$$P(V,e) = \prod \Phi_V \prod \Psi_1 \cdot \ldots \cdot \prod \Psi_k.$$

2. Let S be a separator with the sets  $\Psi_S$  and  $\Psi^S$  in the mailboxes. Then

$$P(S,e) = \prod \Psi_S \prod \Psi^S.$$

#### Theorem

Let the junction tree T represent the Bayesian network BN over the universe U and with evidence e. Assume that T is full.

1. Let *V* be a clique with set of potentials  $\Phi_V$ , and let  $S_1, \ldots, S_k$  be *V*'s neighboring separators and with *V*-directed messages  $\Psi_1, \ldots, \Psi_k$ . Then

$$P(V,e) = \prod \Phi_V \prod \Psi_1 \cdot \ldots \cdot \prod \Psi_k.$$

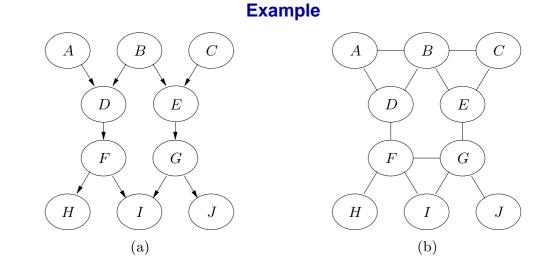
2. Let S be a separator with the sets  $\Psi_S$  and  $\Psi^S$  in the mailboxes. Then

$$P(S,e) = \prod \Psi_S \prod \Psi^S.$$

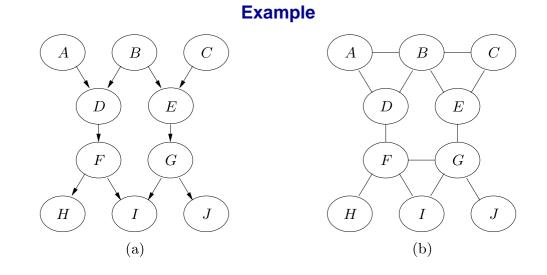
#### Evidence

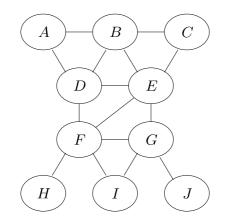
Evidence is inserted by adding corresponding 0/1-potentials to the appropriate cliques.

So far we have assumed that the domain graph is triangulated. If this is not the case, then we embed it in a triangulated graph.

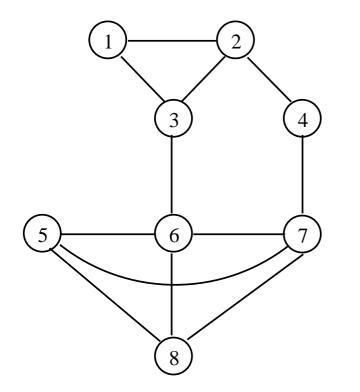


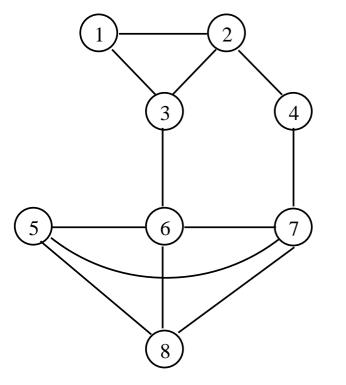
So far we have assumed that the domain graph is triangulated. If this is not the case, then we embed it in a triangulated graph.





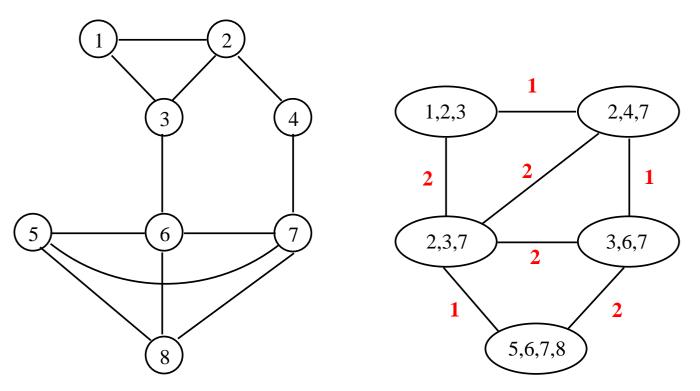
A triangulated graph extending (b)





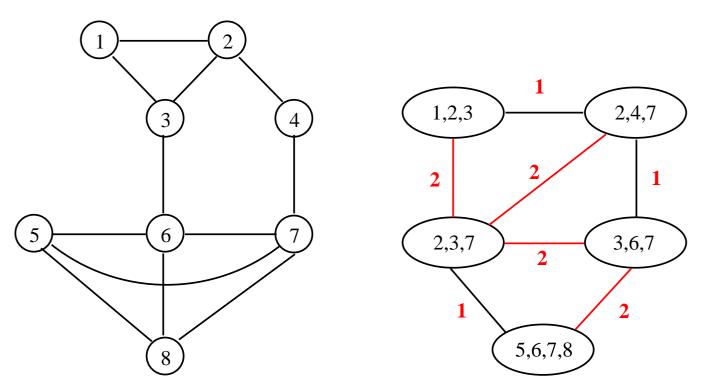
#### A heuristic

Repeatedly eliminate a simplicial node, and if this is not possible eliminate a node minimizing (fa(X) are noneliminated neighbors of X):



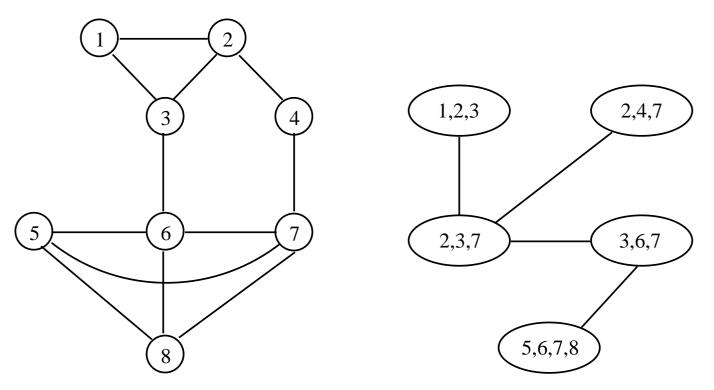
**A heuristic** 

Repeatedly eliminate a simplicial node, and if this is not possible eliminate a node minimizing (fa(X) are noneliminated neighbors of X):



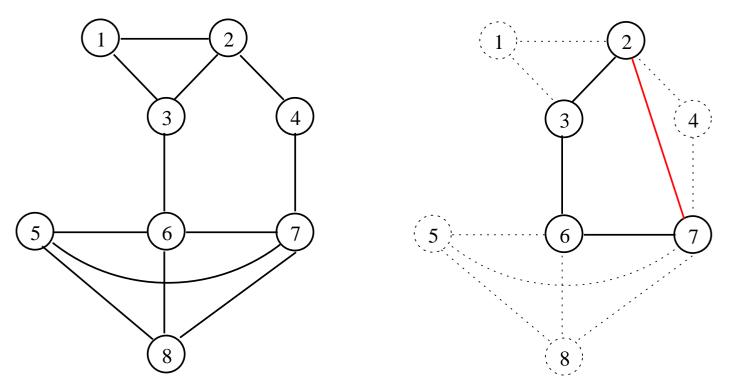
**A heuristic** 

Repeatedly eliminate a simplicial node, and if this is not possible eliminate a node minimizing (fa(X) are noneliminated neighbors of X):



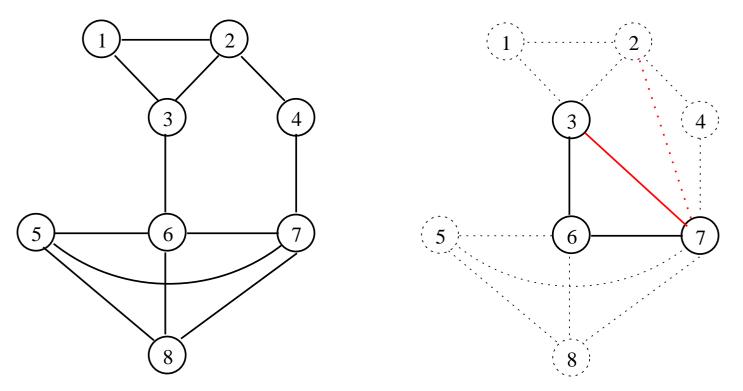
**A heuristic** 

Repeatedly eliminate a simplicial node, and if this is not possible eliminate a node minimizing (fa(X) are noneliminated neighbors of X):



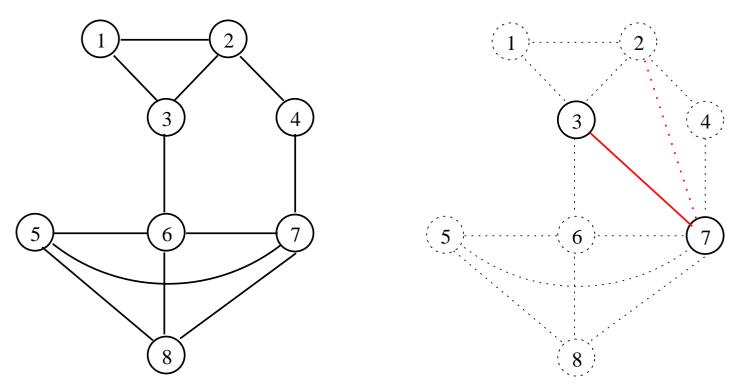
**A heuristic** 

Repeatedly eliminate a simplicial node, and if this is not possible eliminate a node minimizing (fa(X) are noneliminated neighbors of X):



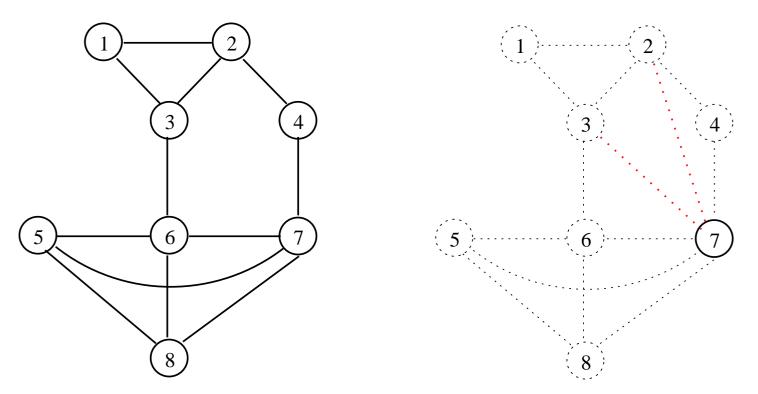
**A heuristic** 

Repeatedly eliminate a simplicial node, and if this is not possible eliminate a node minimizing (fa(X) are noneliminated neighbors of X):



**A heuristic** 

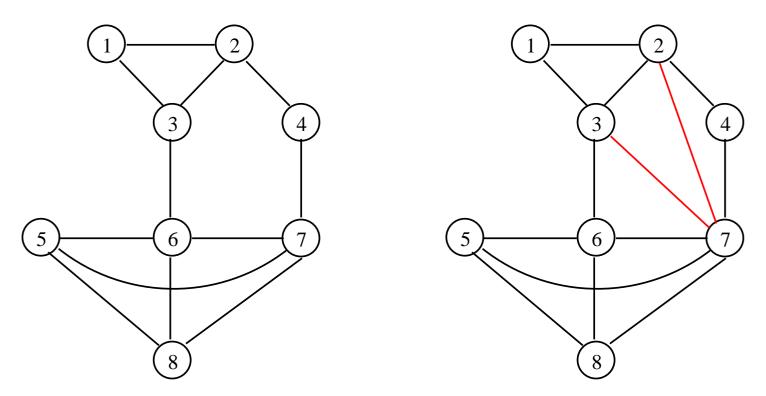
Repeatedly eliminate a simplicial node, and if this is not possible eliminate a node minimizing (fa(X) are noneliminated neighbors of X):



**A heuristic** 

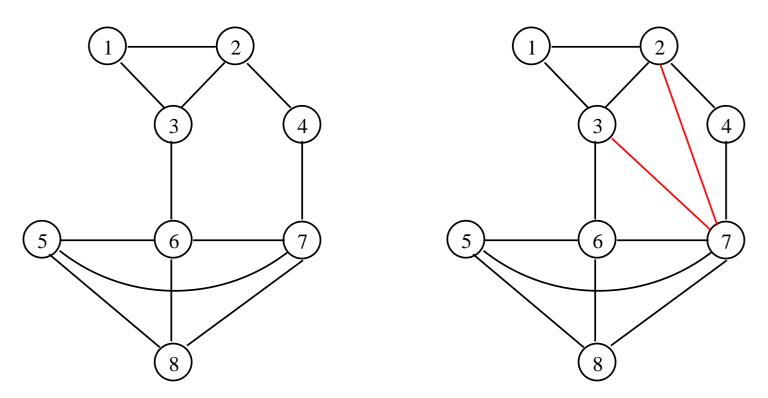
Repeatedly eliminate a simplicial node, and if this is not possible eliminate a node minimizing (fa(X) are noneliminated neighbors of X):

|fa(X)| (fill-in-size).



**A heuristic** 

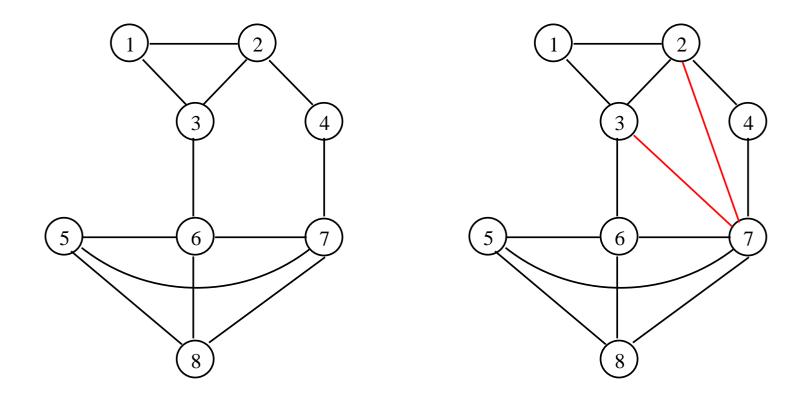
Repeatedly eliminate a simplicial node, and if this is not possible eliminate a node minimizing (fa(X) are noneliminated neighbors of X):



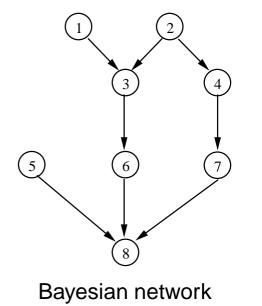
**A heuristic** 

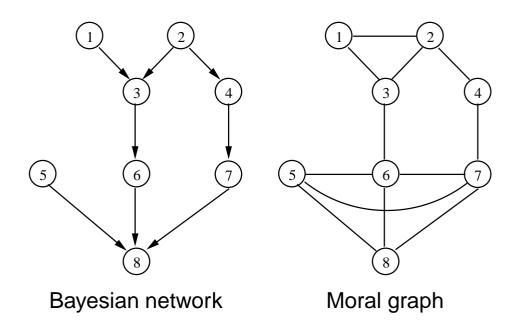
Repeatedly eliminate a simplicial node, and if this is not possible eliminate a node minimizing (fa(X) are noneliminated neighbors of X):

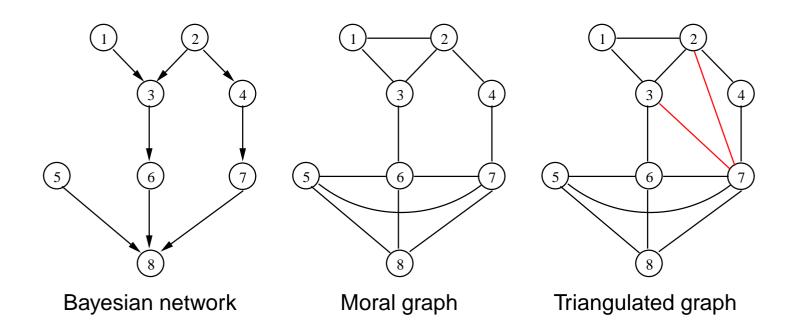
- $\blacktriangleright |fa(X)| (fill-in-size).$
- $\blacktriangleright | \operatorname{sp}(\operatorname{fa}(X)) | \text{ (clique-size).}$
- $\sum_{\{Y,Z\}:\{Y,Z\}\in nb(X)\wedge Z\not\in nb(Y)} |sp(\{Y,Z\})| \text{ (fill-in-weight)}$

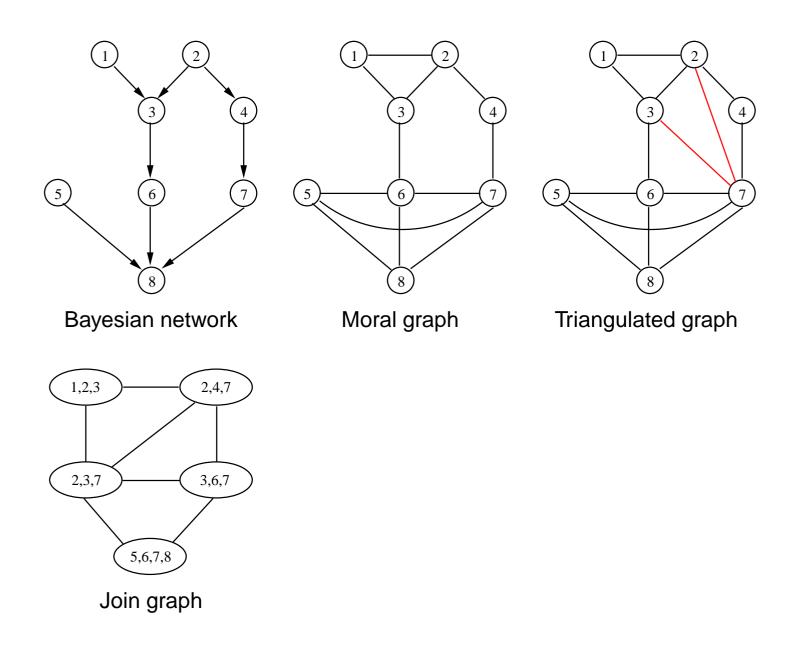


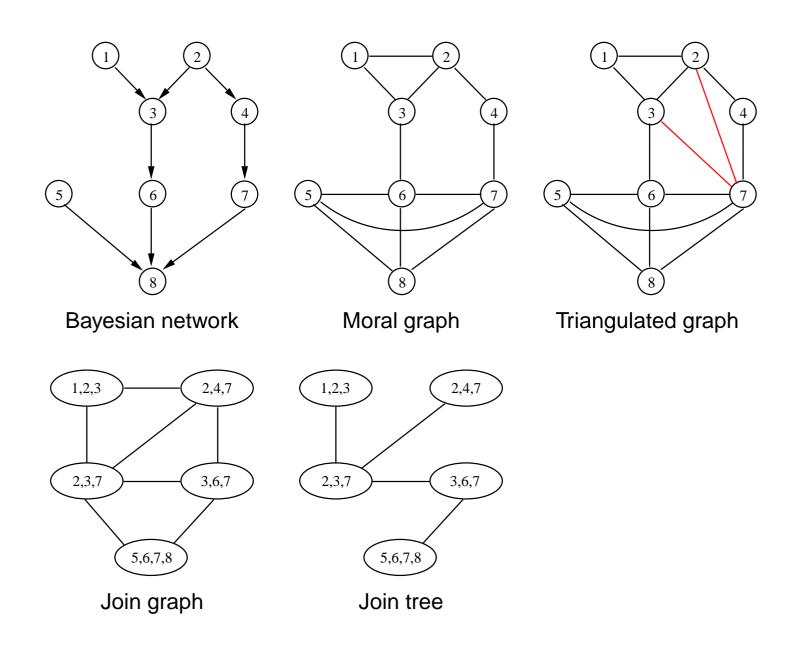
Usually there are many possible triangulations and finding the best one is NP-hard!

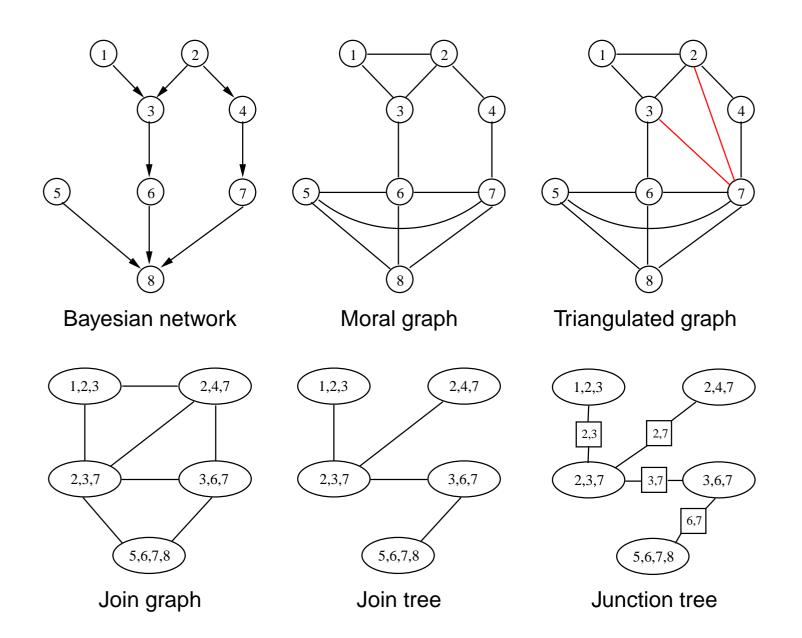












Summary

Given: Bayesian network BN for  $P(\mathbf{V})$ 

#### Summary

Given: Bayesian network BN for  $P(\mathbf{V})$ 

 $\Downarrow$  Moralize

Domain graph G for BN (also called the moral graph)

#### Summary

Given: Bayesian network BN for  $P(\mathbf{V})$ 

 $\Downarrow$  Moralize

Domain graph G for BN (also called the moral graph)

 $\Downarrow$  Triangulate

Triangulated graph for G

#### Summary

Given: Bayesian network BN for  $P(\mathbf{V})$ 

 $\Downarrow$  Moralize

Domain graph G for BN (also called the moral graph)

 $\Downarrow$  Triangulate

Triangulated graph for G

 $\Downarrow \mathsf{Find} \mathsf{ cliques}$ 

Join graph

#### Summary

Given: Bayesian network BN for  $P(\mathbf{V})$ 

 $\Downarrow$  Moralize

Domain graph G for BN (also called the moral graph)

 $\Downarrow$  Triangulate

Triangulated graph for G

 $\Downarrow \mathsf{Find} \mathsf{ cliques}$ 

Join graph

 $\Downarrow$  Find maximal spanning tree

Join tree

#### Summary

Given: Bayesian network BN for  $P(\mathbf{V})$ 

 $\Downarrow$  Moralize

Domain graph G for BN (also called the moral graph)

 $\Downarrow \text{Triangulate}$ 

Triangulated graph for G

 $\Downarrow \mathsf{Find} \mathsf{ cliques}$ 

Join graph

 $\Downarrow$  Find maximal spanning tree

Join tree

 $\Downarrow$ 

Perform belief updating by message passing.