Probabilistic Reasoning

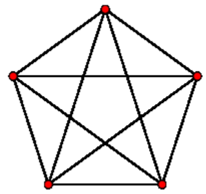
Unit # 11

How to Construct a JT

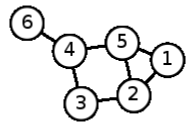
- Moralize the graph by adding links between parents of common children.
- Convert the given BN into an undirected graph by removing the arrowheads.
- Triangulate the graph.
  - If there is a cycle of length > 3 with no chord, add a chord between two non-adjacent vertices in the cycle.
- Order the nodes by maximum cardinality search.
  - Choose any node in the graph and label it 1.
  - For i=2 to n
    - Choose the node with the most labeled neighbors and label it i.
    - If any two labeled neighbors of i are not adjacent to each other
      - Add a link between those neighbors
      - Restart the algorithm.
- Find the cliques of the triangulated graph.
- Arrange as a junction tree.

Clique
(From Wikipedia)

- In graph theory, a clique in an undirected graph G = (V, E) is a subset of the vertex set C ⊆ V, such that for every two vertices in C, there exists an edge connecting the two.
- The size of a clique is the number of vertices it contains.
- Finding whether there is a clique of a given size in a graph (the clique problem) is NP-complete.

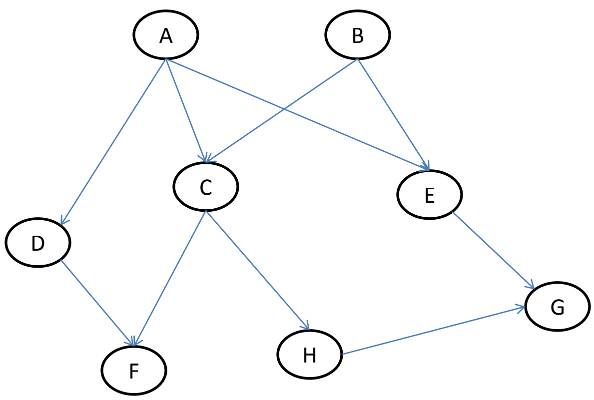


If a subgraph looks like this, the vertices in that subgraph form a clique of size 5.

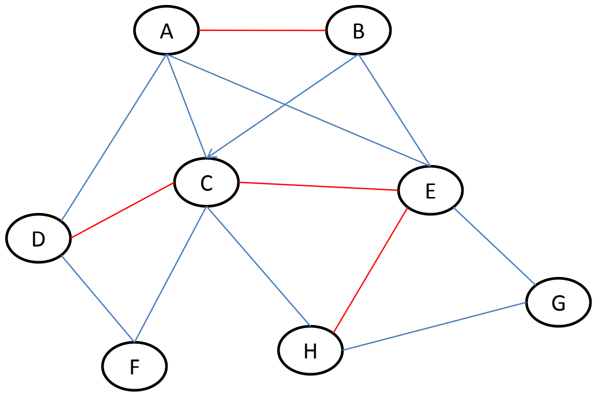


A graph with 6 vertices and 7 edges. On this graph vertices 1, 2, 5 is the only clique of size 3.

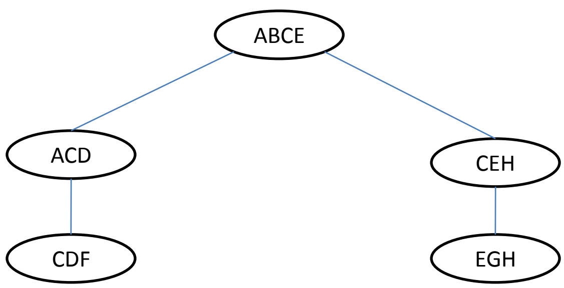
Junction Tree – Example I
(Source: Laskey)



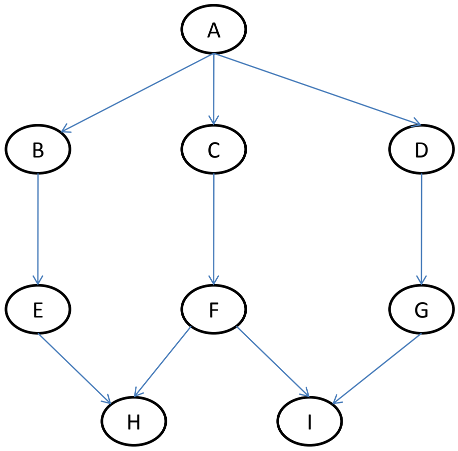
Junction Tree – Example I (Cont'd)



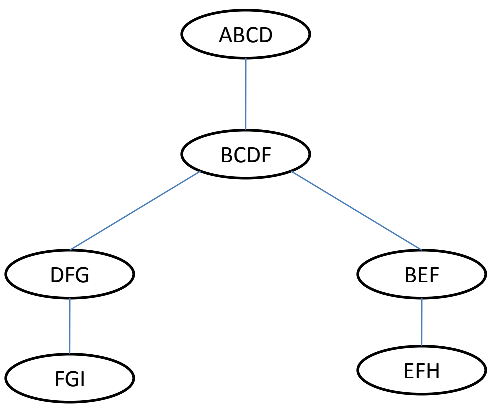
Junction Tree – Example I (Cont'd)



Junction Tree – Example II (Source: Korb & Nicholson)



Junction Tree – Example II (Cont'd)



Approximate Inference Mechanisms

- Simulation Based Schemes
  - Logic Sampling
  - Likelihood Weighting
- Approximate Algorithms
  - Loopy Belief Propagation

Logic Sampling (Laskey)

- Logic sampling takes random draws from the joint distribution of all variables in the network
- Each node keeps a running count of how many times each of its states is drawn
- To take one random draw (one observation on all variables):
  - Sample the nodes in order (so parents of a node are sampled before the node)
  - For root node(s), select state with probability P(state)
  - For non-root node(s) select state with probability P(state|sampled states of parents)
  - After all nodes have been sampled
    - Check whether sampled values of evidence values match observed states
    - If they don't match, throw away the draw
    - If they match, increment the count for sampled state of each non-evidence node
- Estimate P(Xt|Xe) by observed frequency among the retained samples

Algorithm
(From Jensen and Nielsen)

1. Let $(X_1, \ldots, X_n)$ be a topological ordering of the variables.
2. For $j = 1$ to N:
   a) For $i = 1$ to n:
      - Sample a state x_i for X_i using $P(X_i \mid \text{pa}(X_i) = \pi)$, where π is the configuration already sampled for $\text{pa}(X_i)$.

   b) If $\mathbf{x} = (x_1, \ldots, x_n)$ is consistent with $\mathbf{e}$, then

$$N(X_k = x_k) := N(X_k = x_k) + 1,$$

   where x_k is the state that was sampled for X_k.
3. Return:

$$P(X_k = x_k \mid \mathbf{e}) \approx \frac{N(X_k = x_k)}{\sum_{x \in \text{sp}(X_k)} N(X_k = x)}.$$