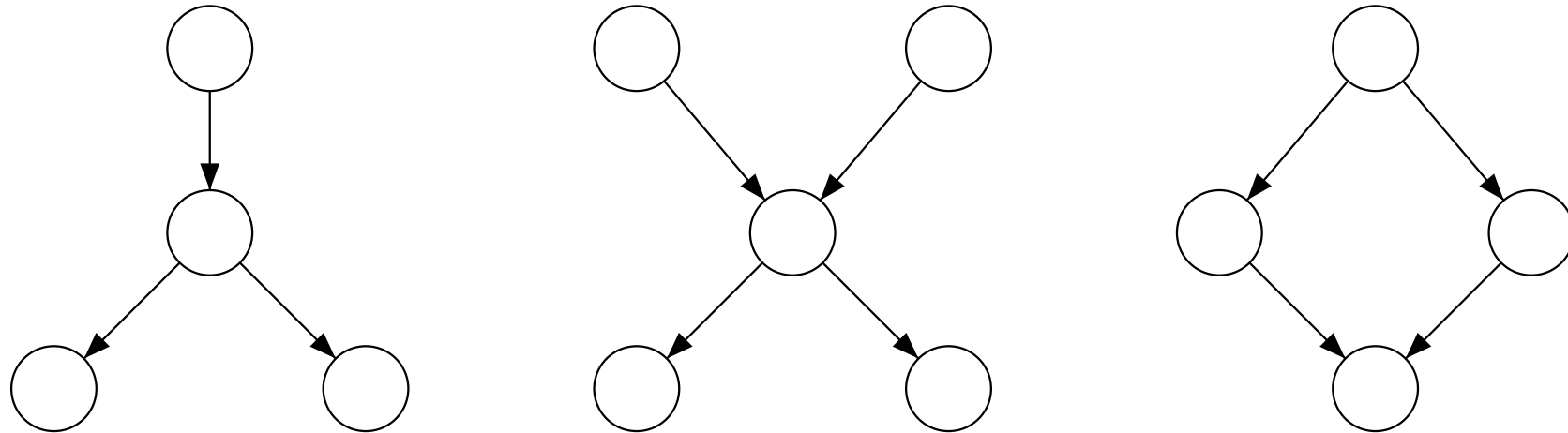


# Clique Tree Representations

# Problems



The propagation algorithm as presented can only deal with *trees*.

Can be extended to *polytrees* (i. e. singly connected graphs with multiple parents per node).

However, it cannot handle networks that contain loops!

## Main Objectives:

Transform the cyclic directed graph into a secondary structure without cycles.

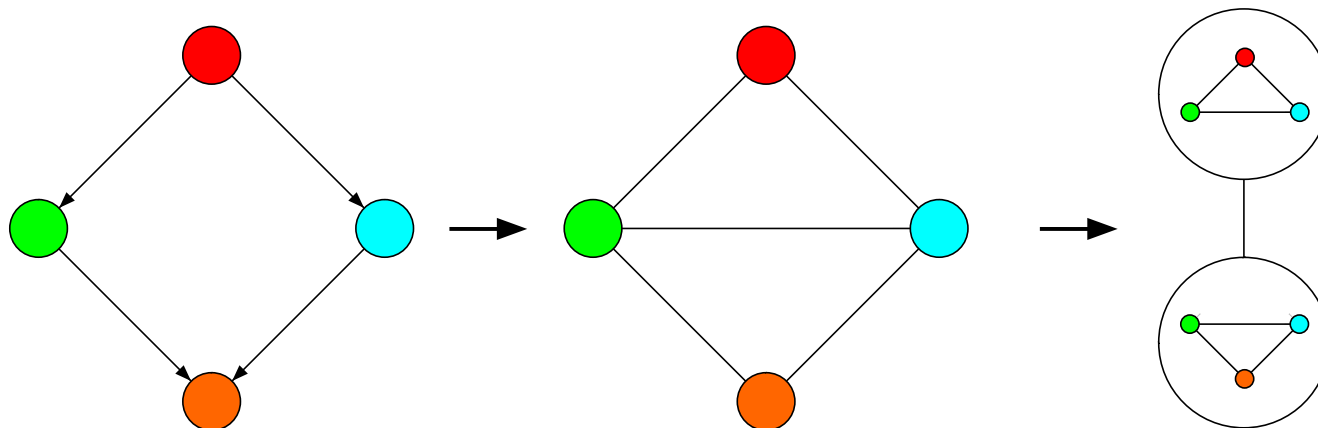
Find a decomposition of the underlying joint distribution.

## Task:

Combine nodes of the original (primary) graph structure.

These groups form the nodes of a secondary structure.

Find a transformation that yields tree structure.



## Idea (2)

### Secondary Structure:

We will generate an undirected graph mimicking (some of) the conditional independence statements of the cyclic directed graph.

Maximal cliques are identified and form the nodes of the secondary structure.

Specify a so-called potential function for every clique such that the product of all potentials yields the initial joint distribution.

In order to propagate evidence, create a **tree** from the clique nodes such that the following property is satisfied:

If two cliques have some attributes in common, then these attributes have to be contained in every clique of the path connecting the two cliques.  
(called the **running intersection property, RIP**)

### Justification:

Tree: Unique path of evidence propagation.

RIP: Update of an attribute reaches all cliques which contain it.

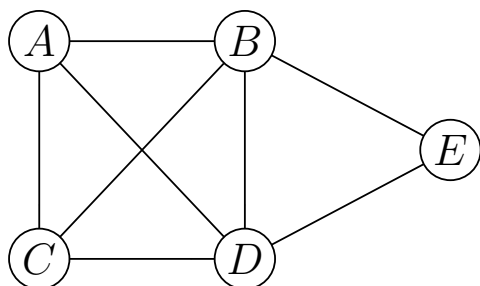
## Complete Graph

An undirected Graph  $G = (V, E)$  is called *complete*, if every pair of (distinct) nodes is connected by an edge.

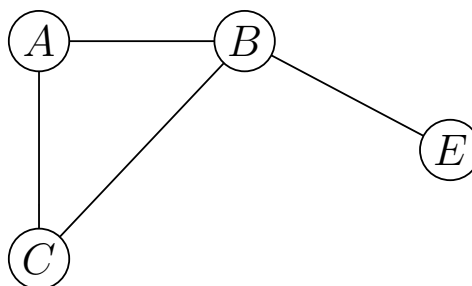
## Induced Subgraph

Let  $G = (V, E)$  be an undirected graph and  $W \subseteq V$  a selection of nodes. Then,  $G_W = (W, E_W)$  is called the *subgraph of  $G$  induced by  $W$*  with  $E_W$  being

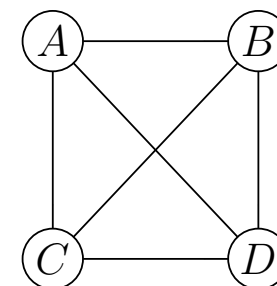
$$E_W = \{(u, v) \in E \mid u, v \in W\}.$$



Incomplete graph



Subgraph  $(W, E_W)$   
with  $W = \{A, B, C, E\}$



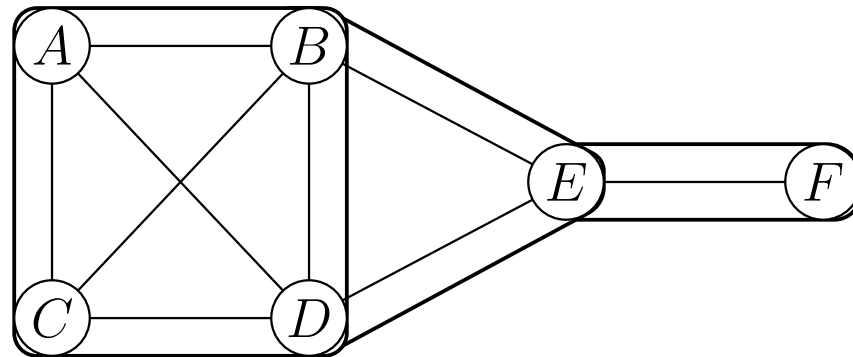
Complete (sub)graph

# Prerequisites (2)

## Complete Set, Clique

Let  $G = (V, E)$  be an undirected graph. A set  $W \subseteq V$  is called *complete* iff it induces a complete subgraph. It is further called a *clique*, iff  $W$  is maximal, i.e. it is not possible to add a node to  $W$  without violating the completeness condition.

- a)  $W$  is complete  $\Leftrightarrow W$  induces a complete subgraph
- b)  $W$  is a clique  $\Leftrightarrow W$  is complete and maximal



3 cliques

$$C_1 = \{A, B, C, D\}$$

$$C_2 = \{B, D, E\}$$

$$C_3 = \{E, F\}$$

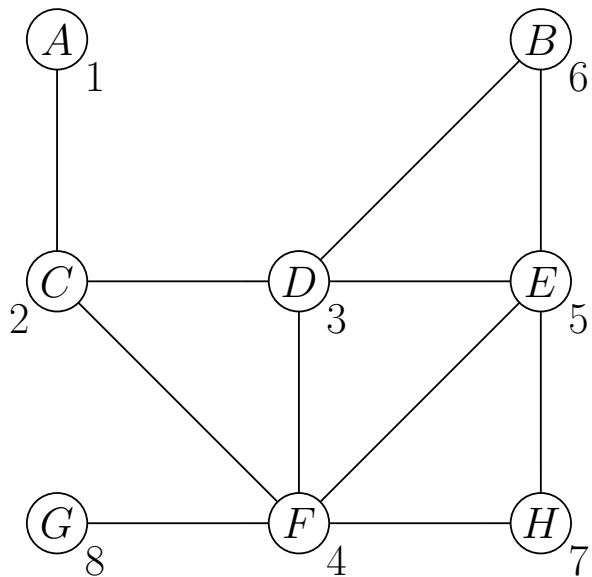
# Prerequisites (3)

## Perfect Ordering

Let  $G = (V, E)$  be an undirected graph with  $n$  nodes and  $\alpha = \langle v_1, \dots, v_n \rangle$  a total ordering on  $V$ . Then,  $\alpha$  is called *perfect*, if the following sets

$$\text{adj}(v_i) \cap \{v_1, \dots, v_{i-1}\} \quad i = 1, \dots, n$$

are complete, where  $\text{adj}(v_i) = \{w \mid (v_i, w) \in E\}$  returns the adjacent nodes of  $v_i$ .



$\alpha = \langle A, C, D, F, E, B, H, G \rangle$

$i$	$\text{adj}(v_i)$	$\{v_1, \dots, v_{i-1}\} \cap \text{adj}(v_i)$	
1	$\{C\}$	$\emptyset \cap \{C\}$	$= \emptyset$ complete
2	$\{A, D, F\}$	$\{A\} \cap \{A, D, F\}$	$= \{A\}$ complete
3	$\{C, B, E, F\}$	$\{A, C\} \cap \{C, B, E, F\}$	$= \{C\}$ complete
4	$\{G, C, D, E, H\}$	$\{A, C, D\} \cap \{G, C, D, E, H\}$	$= \{C, D\}$ complete
5	$\{B, D, F, H\}$	$\{A, C, D, F\} \cap \{B, D, F, H\}$	$= \{D, F\}$ complete
6	$\{D, E\}$	$\{A, C, D, F, E\} \cap \{D, E\}$	$= \{D, E\}$ complete
7	$\{F, E\}$	$\{A, C, D, F, E, B\} \cap \{F, E\}$	$= \{F, E\}$ complete
8	$\{F\}$	$\{A, C, D, F, E, B, H\} \cap \{F\}$	$= \{F\}$ complete

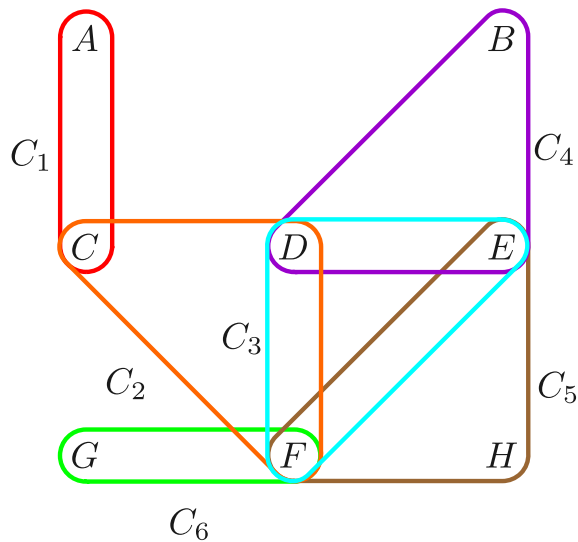
$\alpha$  is a perfect ordering

# Prerequisites (4)

## Running Intersection Property

Let  $G = (V, E)$  be an undirected graph with  $p$  cliques. An ordering of these cliques has the *running intersection property (RIP)*, if for every  $j > 1$  there exists an  $i < j$  such that:

$$C_j \cap (C_1 \cup \dots \cup C_{j-1}) \subseteq C_i$$



$$\xi = \langle C_1, C_2, C_3, C_4, C_5, C_6 \rangle$$

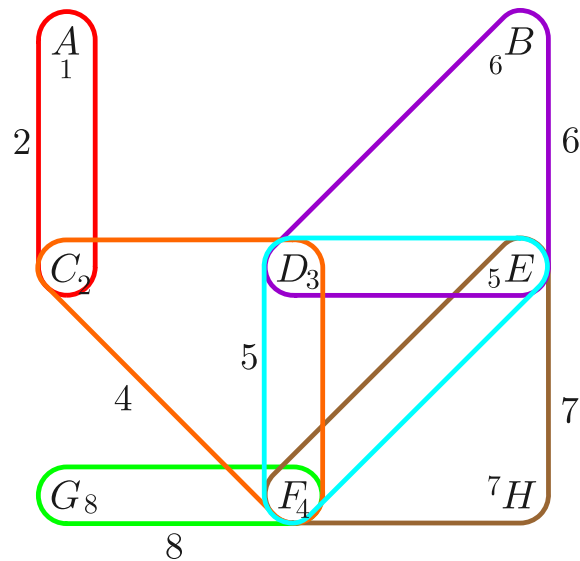
$j$			$i$
2	$C_2 \cap C_1$	$= \{C\}$	$\subseteq C_1$   1
3	$C_3 \cap (C_1 \cup C_2)$	$= \{D, F\}$	$\subseteq C_2$   2
4	$C_4 \cap (C_1 \cup C_2 \cup C_3)$	$= \{D, E\}$	$\subseteq C_3$   3
5	$C_5 \cap (C_1 \cup C_2 \cup C_3 \cup C_4)$	$= \{E, F\}$	$\subseteq C_3$   3
6	$C_6 \cap (C_1 \cup C_2 \cup C_3 \cup C_4 \cup C_5)$	$= \{F\}$	$\subseteq C_5$   5

$\xi$  has running intersection property



# Prerequisites (5)

If a node ordering  $\alpha$  of an undirected graph  $G = (V, E)$  is perfect and the cliques of  $G$  are ordered according to the highest rank (w. r. t.  $\alpha$ ) of the containing nodes, then this clique ordering has RIP.



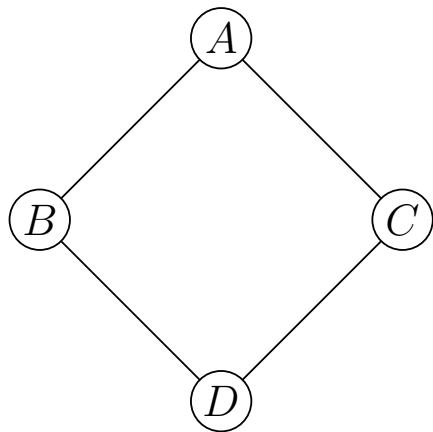
Clique	Rank
$\{A, C\}$	$\max\{\alpha(A), \alpha(C)\} = 2 \rightarrow C_1$
$\{C, D, F\}$	$\max\{\alpha(C), \alpha(D), \alpha(F)\} = 4 \rightarrow C_2$
$\{D, E, F\}$	$\max\{\alpha(D), \alpha(E), \alpha(F)\} = 5 \rightarrow C_3$
$\{B, D, E\}$	$\max\{\alpha(B), \alpha(D), \alpha(E)\} = 6 \rightarrow C_4$
$\{F, E, H\}$	$\max\{\alpha(F), \alpha(E), \alpha(H)\} = 7 \rightarrow C_5$
$\{F, G\}$	$\max\{\alpha(F), \alpha(G)\} = 8 \rightarrow C_6$

How to get a perfect ordering?

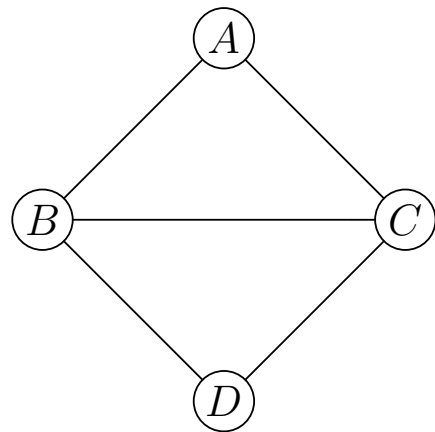
# Triangulated Graphs

## Triangulated Graph

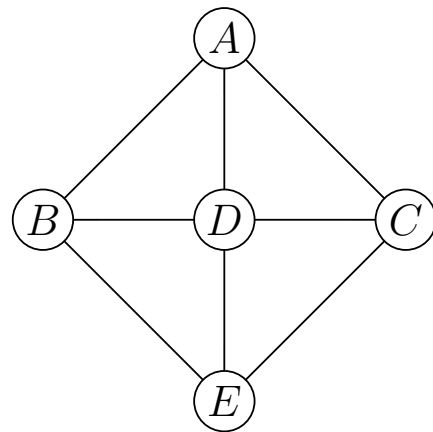
An undirected graph is called *triangulated* if every simple loop (i. e. path with identical start and end node but with any other node occurring at most once) of length greater 3 has a chord.



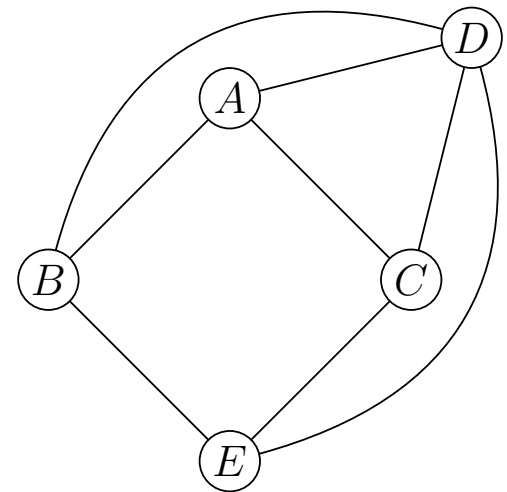
not triangulated



triangulated



not triangulated

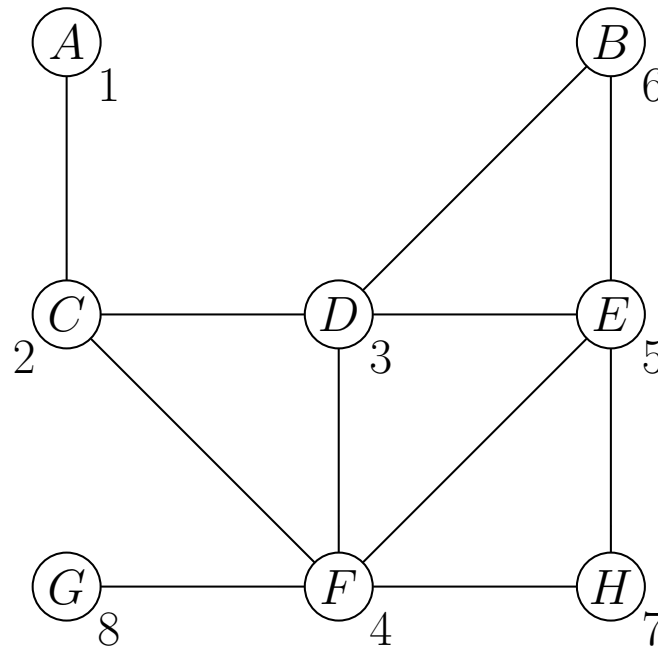


no chord for  $\langle A, B, E, C \rangle$

# Triangulated Graphs (2)

## Maximum Cardinality Search

Let  $G = (V, E)$  be an undirected graph. An ordering according *maximum cardinality search (MCS)* is obtained by first assigning 1 to an arbitrary node. If  $n$  numbers are assigned the node that is connected to most of the nodes already numbered gets assigned number  $n + 1$ .



3 can be assigned to  $D$  or  $F$

6 can be assigned to  $H$  or  $B$

# Triangulated Graphs (3)

If an undirected graph is triangulated, then the ordering obtained by MCS is perfect.

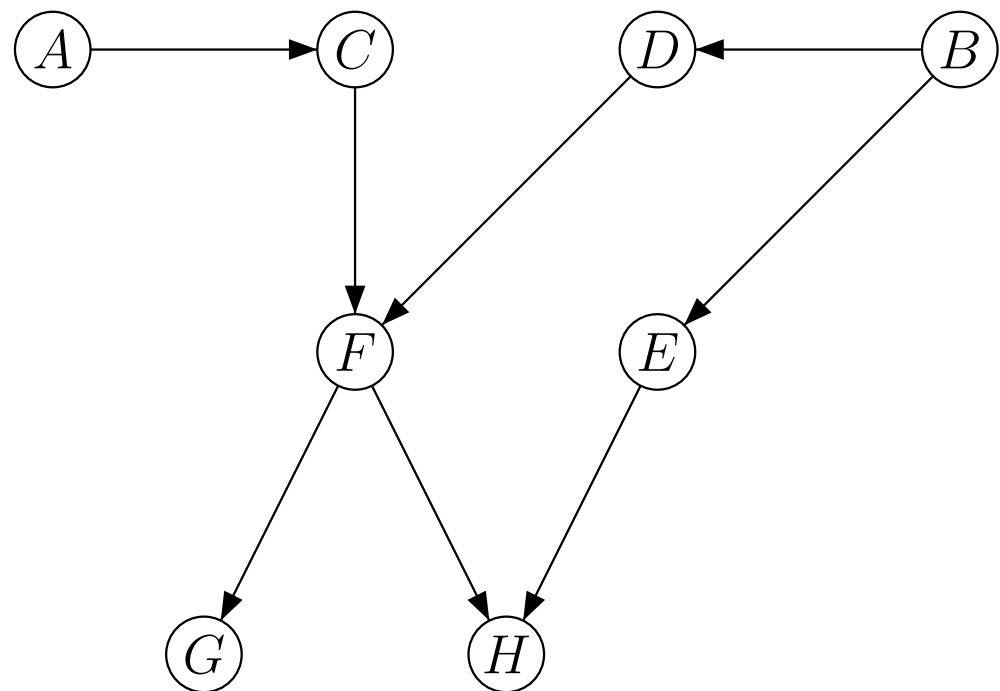
To check whether a graph is triangulated is efficient to implement. The optimization problem that is related to the triangulation task is NP-hard. However, there are good heuristics.

## **Moral Graph** (Repetition)

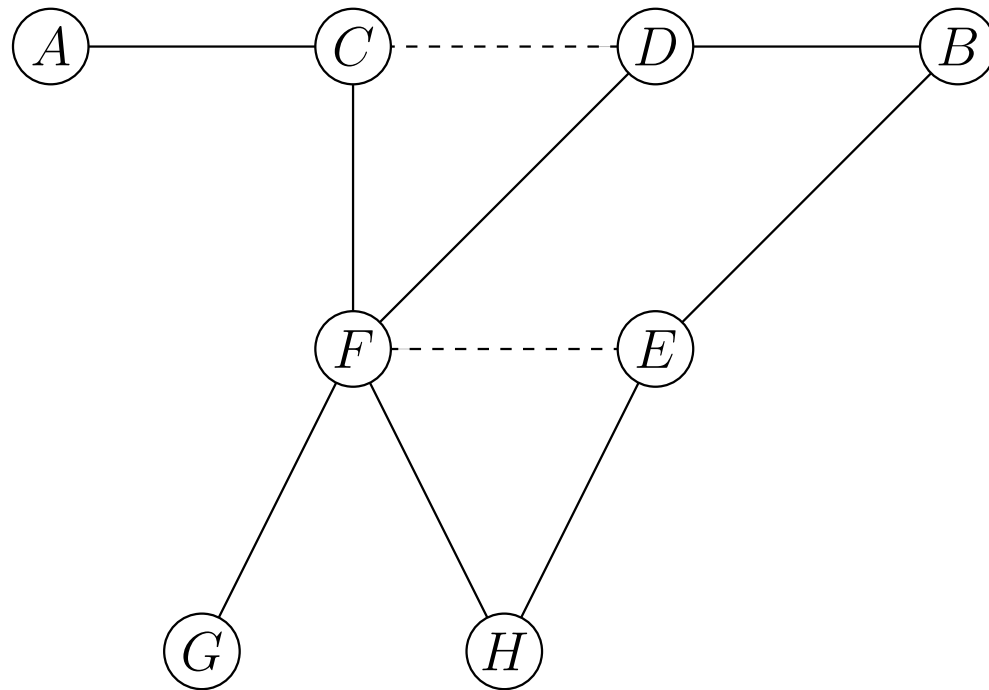
Let  $G = (V, E)$  be a directed acyclic graph. If  $u, w \in W$  are parents of  $v \in V$  connect  $u$  and  $w$  with an (arbitrarily oriented) edge. After the removal of all edge directions the resulting graph  $G_m = (V, E')$  is called the *moral graph* of  $G$ .

# Join-Tree Construction (1)

Given directed graph.

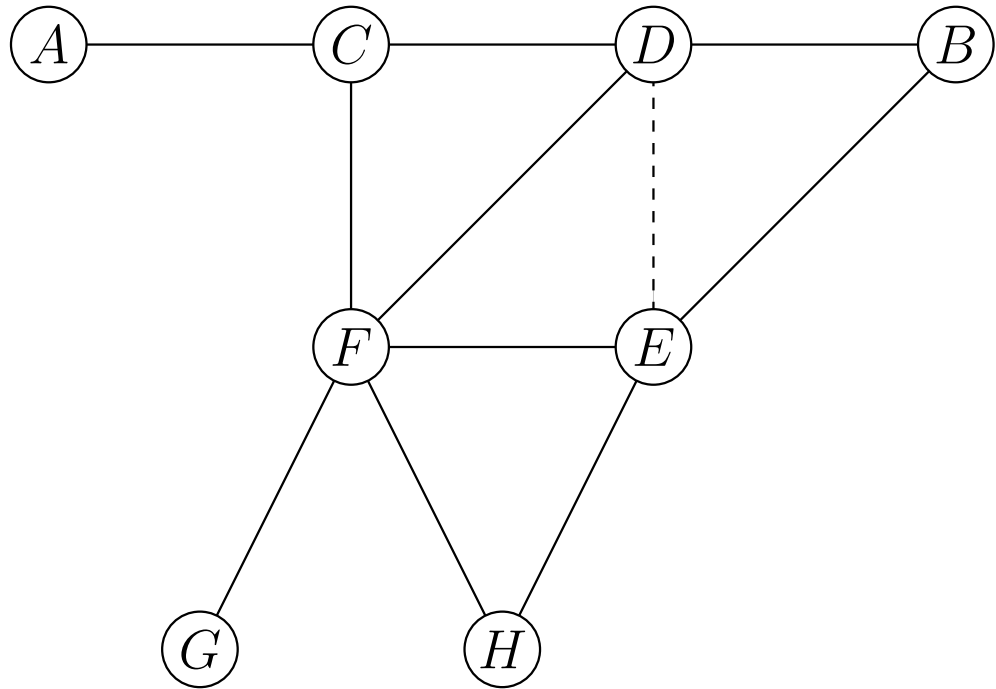


## Join-Tree Construction (2)



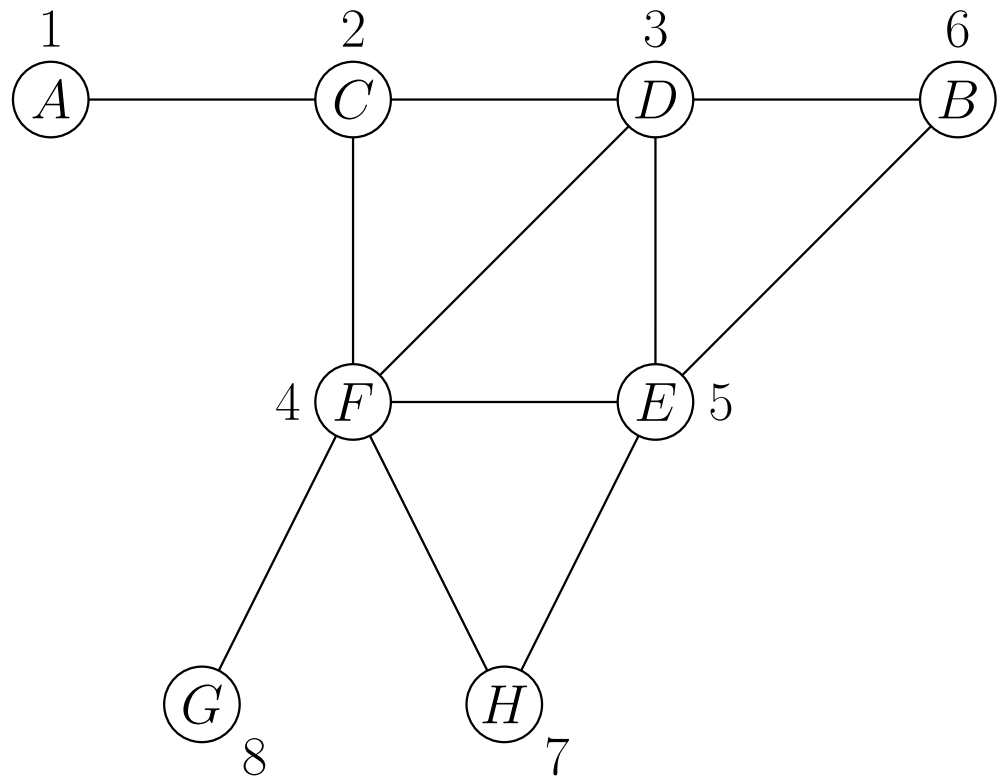
- Moral graph

# Join-Tree Construction (3)



- Moral graph
- Triangulated graph

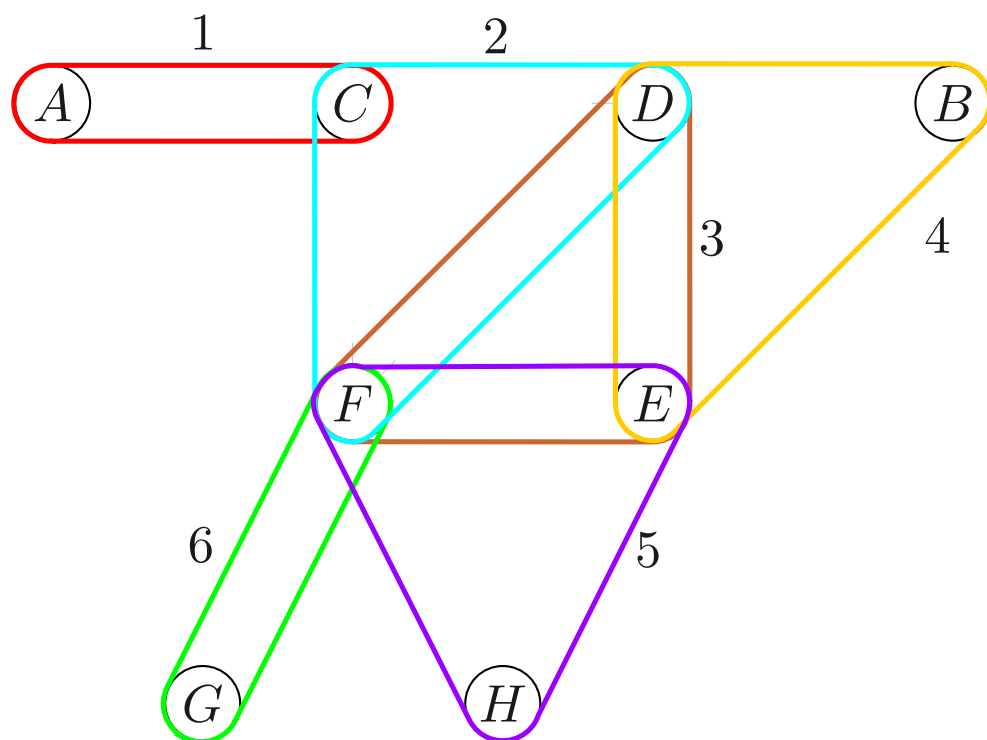
# Join-Tree Construction (4)



- Moral graph
- Triangulated graph
- MCS yields perfect ordering

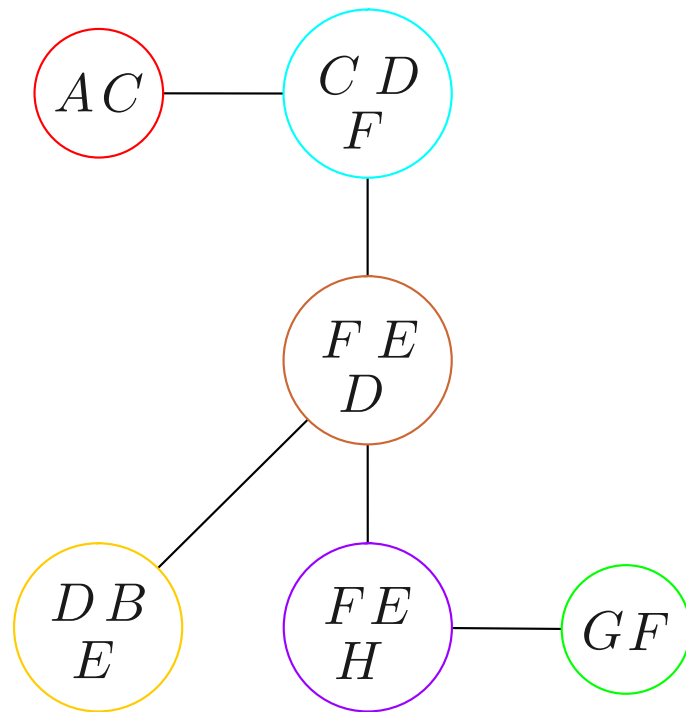


# Join-Tree Construction (5)



- Moral graph
- Triangulated graph
- MCS yields perfect ordering
- Clique order has RIP

# Join-Tree Construction (6)



- Moral graph
- Triangulated graph
- MCS yields perfect ordering
- Clique order has RIP
- Form a join-tree

Two cliques can be connected if they have a non-empty intersection. The generation of the tree follows the RIP. In case of a tie, connect cliques with the largest intersection. (e. g.  $DBE$ — $FED$  instead of  $DBE$ — $CFD$ ) Break remaining ties arbitrarily.

# Example: Expert Knowledge

## **Qualitative knowledge:**

Metastatic cancer is a possible cause of brain tumor, and is also an explanation for increased total serum calcium. In turn, either of these could explain a patient falling into a coma. Severe headache is also possibly associated with a brain tumor.

## **Special case:**

The patient has heavy headache.

## **Query:**

Will the patient fall into coma?

## Example: Choice of State Space

Attribute	Possible Values
$A$ metastatic cancer	$\text{dom}(A) = \{a_1, a_2\}$ $\cdot_1 = \text{existing}$
$B$ increased total serum calcium	$\text{dom}(B) = \{b_1, b_2\}$ $\cdot_2 = \text{notexisting}$
$C$ brain tumor	$\text{dom}(C) = \{c_1, c_2\}$
$D$ coma	$\text{dom}(D) = \{d_1, d_2\}$
$E$ severe headache	$\text{dom}(E) = \{e_1, e_2\}$

Exhaustive state space:

$$\Omega = \text{dom}(A) \times \text{dom}(B) \times \text{dom}(C) \times \text{dom}(D) \times \text{dom}(E)$$

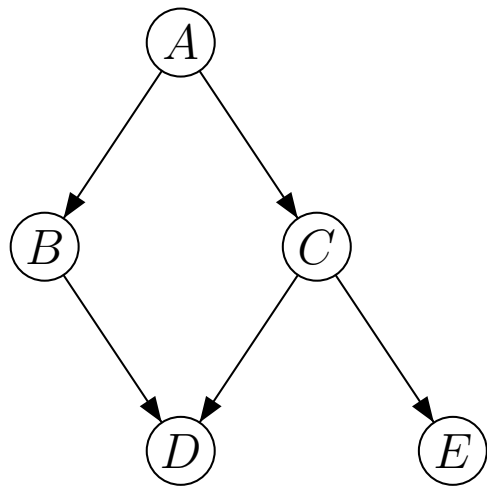
Marginal and conditional probabilities are of interest for the user!

# Example: Qualitative Knowledge

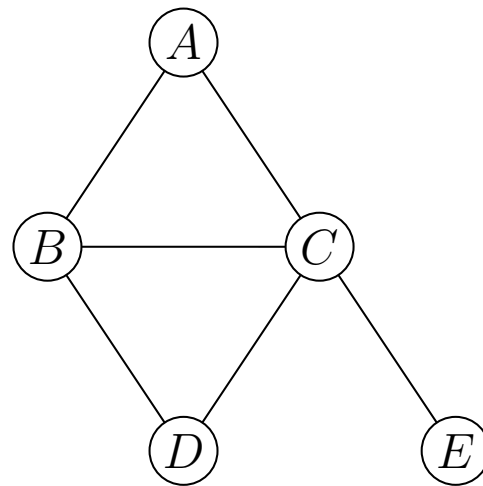
$$\begin{array}{l} P(e_1 | c_1) = 0.8 \\ P(e_1 | c_2) = 0.6 \end{array} \left. \vphantom{\begin{array}{l} P(e_1 | c_1) = 0.8 \\ P(e_1 | c_2) = 0.6 \end{array}} \right\} \text{headaches common, but more common if tumor present}$$
$$\begin{array}{l} P(d_1 | b_1, c_1) = 0.8 \\ P(d_1 | b_1, c_2) = 0.8 \\ P(d_1 | b_2, c_1) = 0.8 \\ P(d_1 | b_2, c_2) = 0.05 \end{array} \left. \vphantom{\begin{array}{l} P(d_1 | b_1, c_1) = 0.8 \\ P(d_1 | b_1, c_2) = 0.8 \\ P(d_1 | b_2, c_1) = 0.8 \\ P(d_1 | b_2, c_2) = 0.05 \end{array}} \right\} \text{coma rare but common, if either cause is present}$$
$$\begin{array}{l} P(b_1 | a_1) = 0.8 \\ P(b_1 | a_2) = 0.2 \end{array} \left. \vphantom{\begin{array}{l} P(b_1 | a_1) = 0.8 \\ P(b_1 | a_2) = 0.2 \end{array}} \right\} \begin{array}{l} \text{increased calcium uncommon,} \\ \text{but common consequence of metastases} \end{array}$$
$$\begin{array}{l} P(c_1 | a_1) = 0.2 \\ P(c_1 | a_2) = 0.05 \end{array} \left. \vphantom{\begin{array}{l} P(c_1 | a_1) = 0.2 \\ P(c_1 | a_2) = 0.05 \end{array}} \right\} \text{brain tumor rare, and uncommon consequence of metastases}$$
$$P(a_1) = 0.2 \quad \left. \vphantom{P(a_1) = 0.2} \right\} \text{incidence of metastatic cancer in relevant clinic}$$

# Example (1)

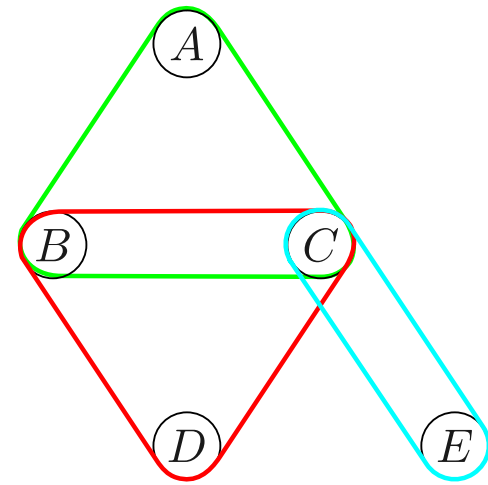
Example: Metastatic Cancer



Dependencies



Moralization/Triangulation



MCS, hyper graph

## Example (2)

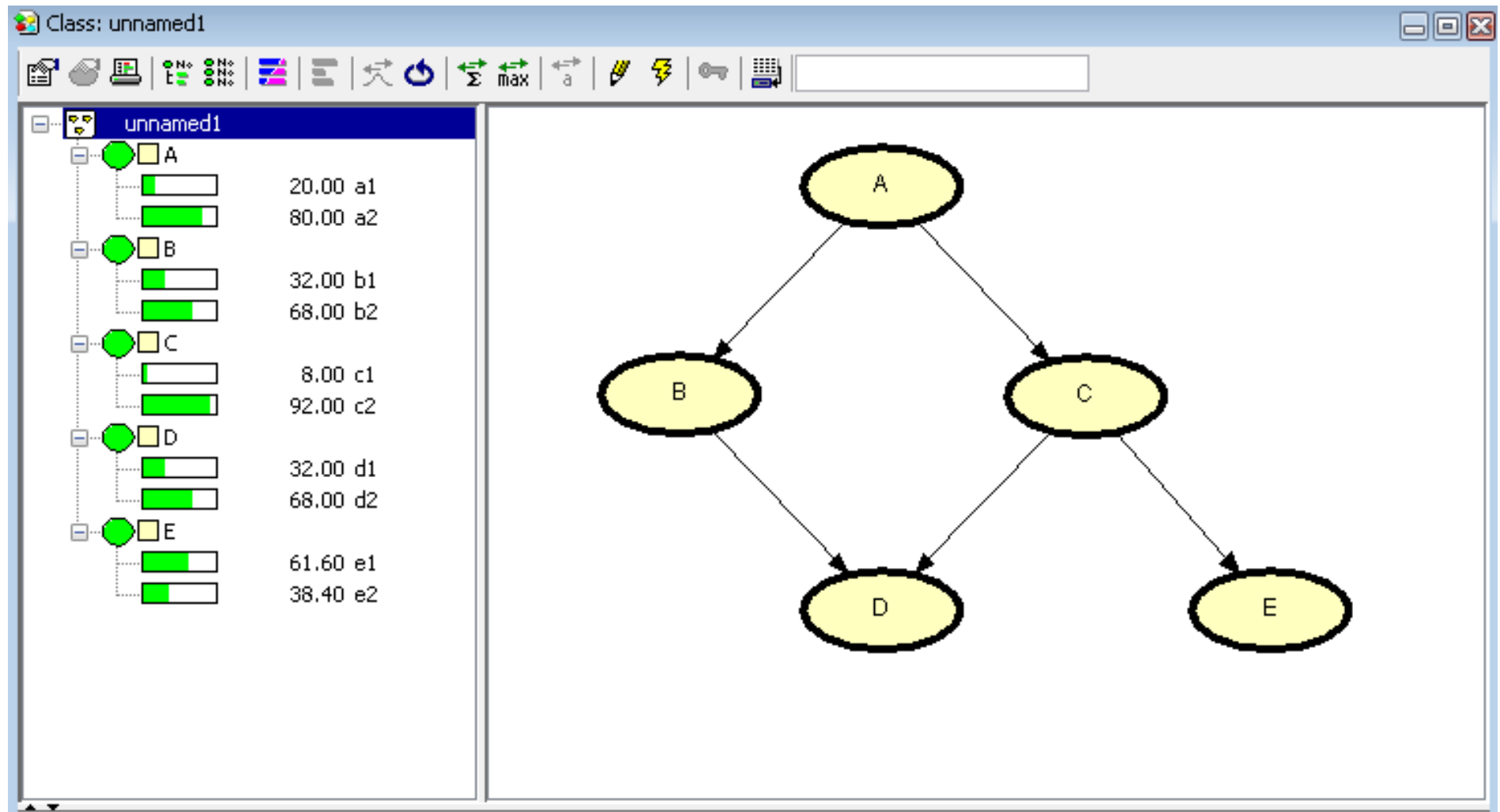
Quantitative knowledge:

$(a, b, c)$	$P(a, b, c)$	$(b, c, d)$	$P(b, c, d)$	$(c, e)$	$P(c, e)$
$a_1, b_1, c_1$	0.032	$b_1, c_1, d_1$	0.032	$c_1, e_1$	0.064
$a_2, b_1, c_1$	0.008	$b_2, c_1, d_1$	0.032	$c_2, e_1$	0.552
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$c_1, e_2$	0.016
$a_2, b_2, c_2$	0.608	$b_2, c_2, d_2$	0.608	$c_2, e_2$	0.368

Decomposition:

$$\begin{aligned} P(A, B, C, D, E) &= P(A)P(B | A)P(C | A)P(D | BC)P(E | C) \\ &= \frac{P(A, B)P(B, C, D)P(C, E)}{P(BC)P(C)} \end{aligned}$$

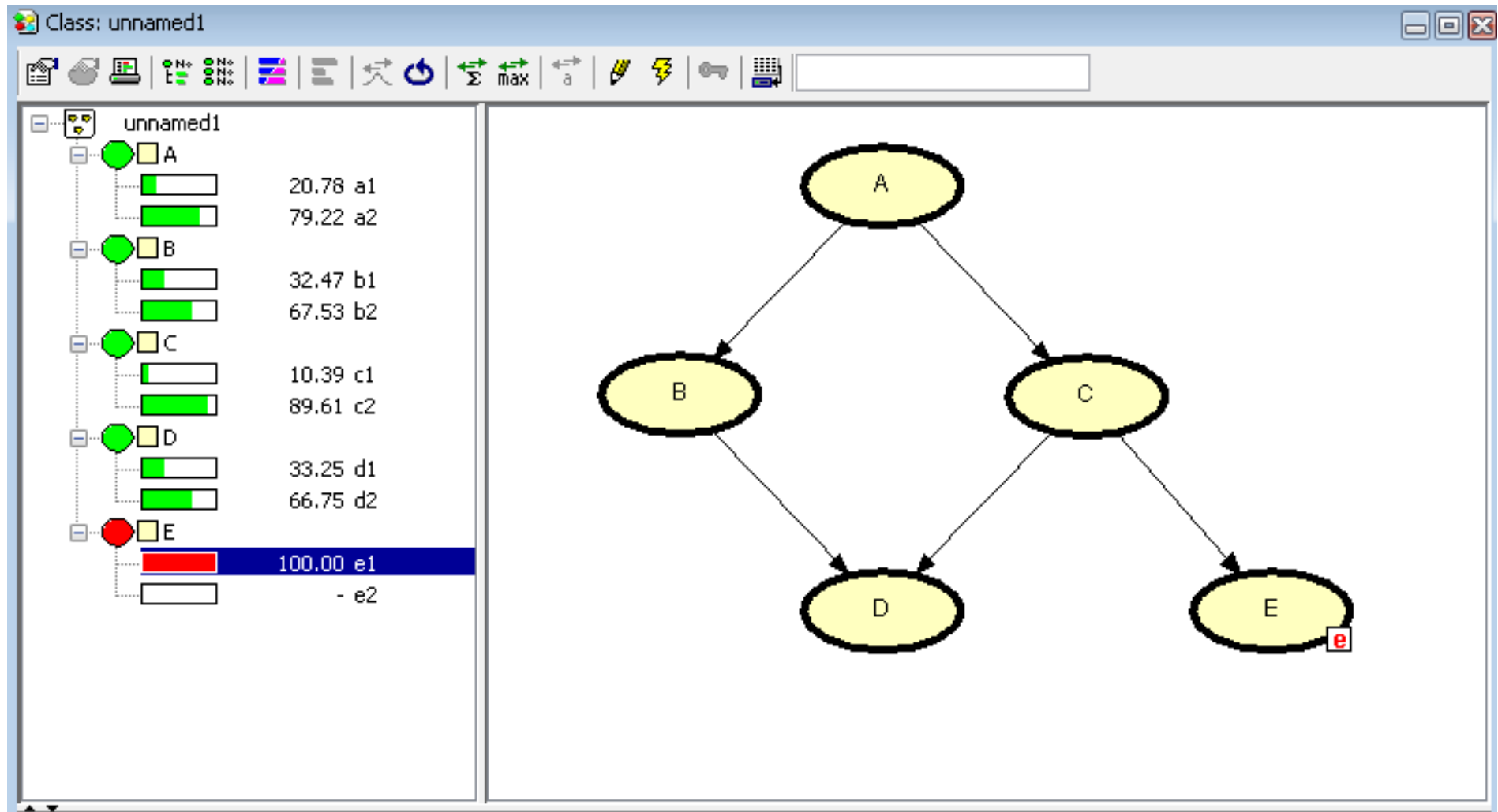
# Example (3)



Marginal distributions in the HUGIN tool.



# Example (4)



Conditional marginal distributions with evidence  $E = e_1$

# Loop-Breaking Techniques (1985-1990)

1. Conditioning (1985)
2. Stochastic simulation (1987)
3. Tree clustering (Spiegelhalter & Lauritzen 1986)
4. Node elimination (Shachter 1986)

Problems:

- Time exponential in tree-width (Dechter 1996)
- Autonomy is lost

# Applications of Bayesian Networks

- Medical Diagnosis
- Clinical Decision Support
- Complex Genetic Models
- Crime Risk Factors Analysis
- Spatial Dynamics in Geography
- Risk Management in Robotics
- Conservation of a threatened Bird
- Classification of Wines
- Student Modelling
- Sensor Validation
- An Information Retrieval System
- Reliability Analysis of Systems
- Terrorism Risk Management
- Credit-Rating of Companies
- Modelling of Mineral Potential
- Pavement and Bridge Management
- Complex Industrial Process Operation
- Probability of Default for Large Corporates
- Inference Problems in Forensic Science