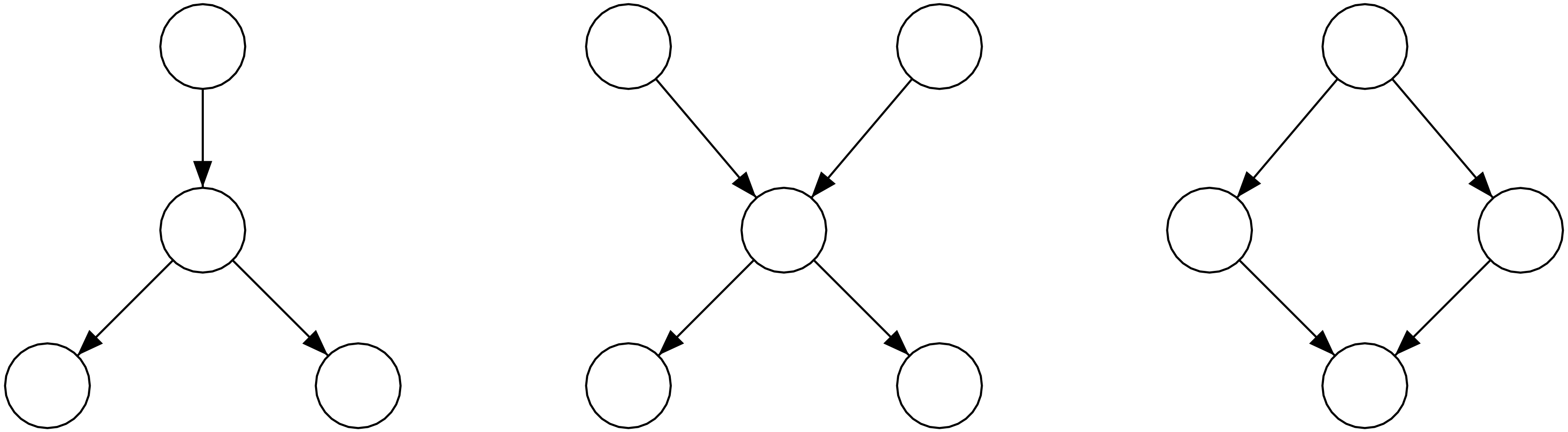
Clique Tree Representation

Problem: Loops



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

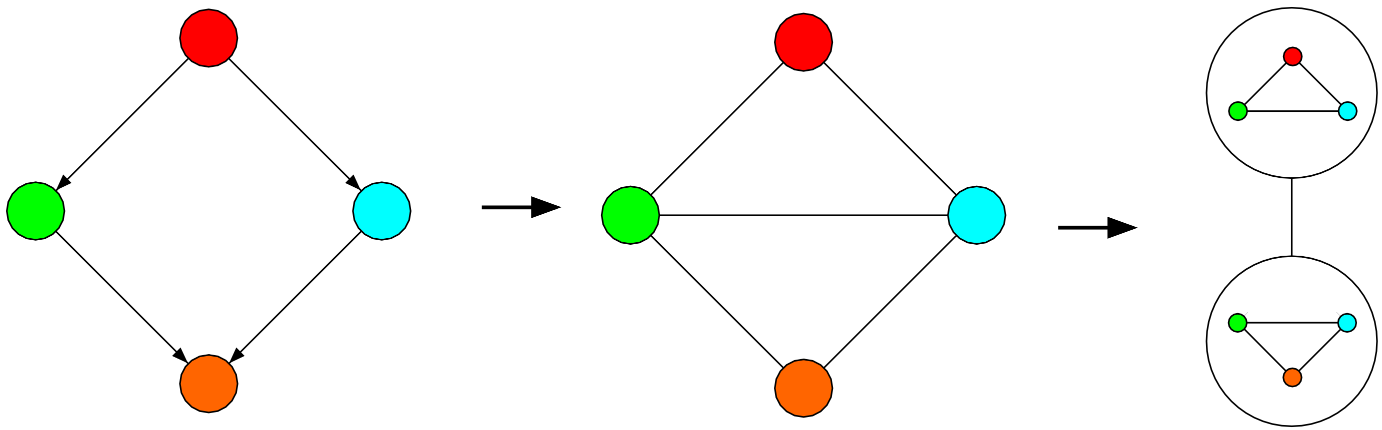
Can be extended to *polytrees* (i. e. singly connected graphs with multiple parents per node).

However, it cannot handle networks that contain loops!

Plan for solution

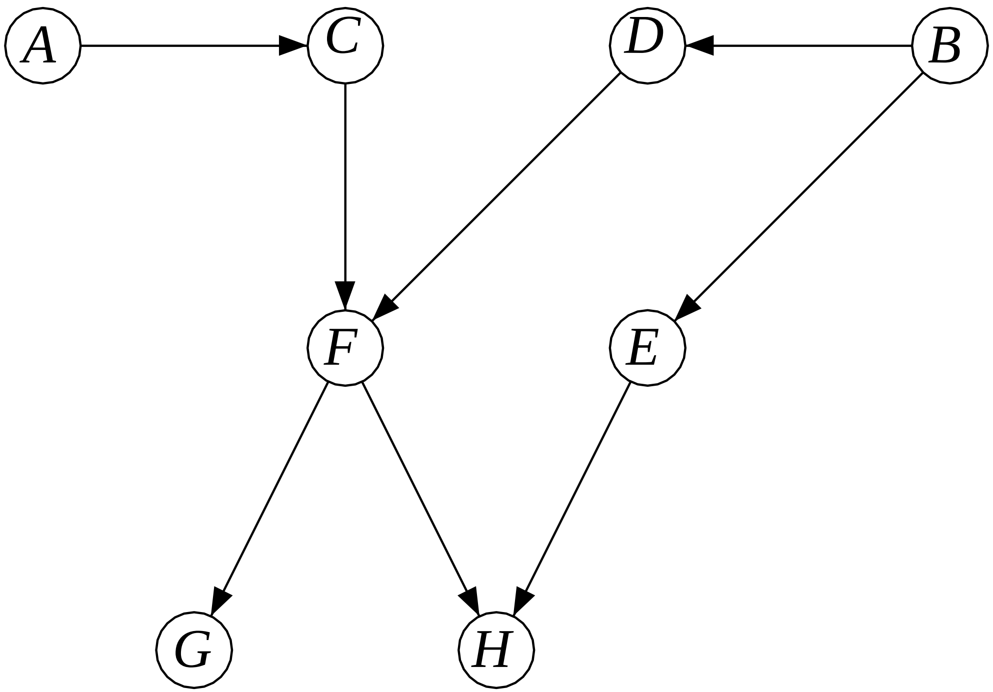
Transform the acyclic directed graph into a secondary structure with tree structure.

Find a decomposition of the underlying joint distribution.



Example: Join-Tree Construction

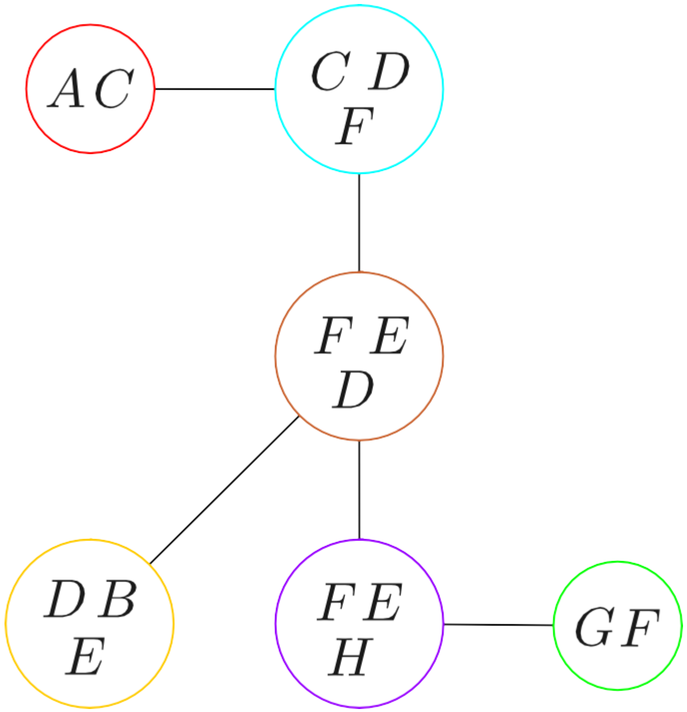
Given directed graph.



The result is a decomposition, represented in form of a join-tree

Transformation Algorithm

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



Plan for solution

In more detail

- **Generation of an undirected graph** mimicking (some of) the conditional independence statements of the cyclic directed graph.
- **Identification of maximal cliques of the undirected graph**
- **Creation of a clique tree** such that the **running intersection property** (RIP) is satisfied.
- **Factorization** with Potential Fuctions

Justification

Probability distribution: Decomposition using the clique tree

Tree: Unique path of evidence propagation

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

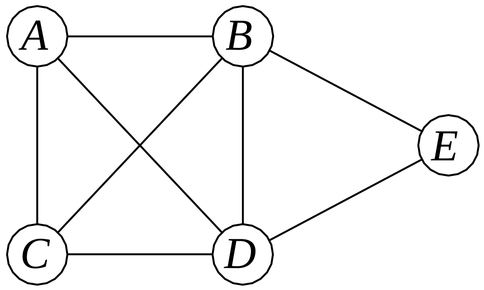
Potential functions: Efficient algorithms

Complete Graph

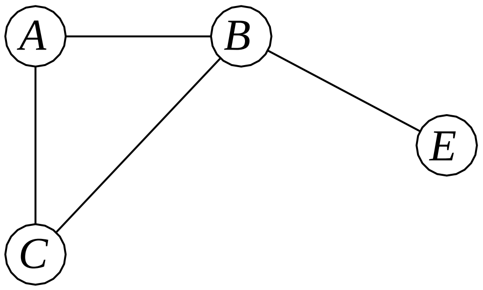
An undirected Graph $G = (V, E)$ is called *complete*, if every pair of (distinct) nodes is connected by an edge.

Induced Subgraph

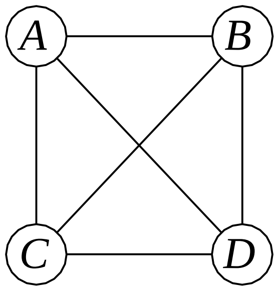
Let $G = (V, E)$ be an undirected graph and $W \subseteq V$ a selection of nodes. Then, $G_W = (W, E_W)$ is called the *subgraph of* G *induced by* W with
$E_W = \{(u, v) \in E \mid u, v \in W\}$.



Incomplete graph



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



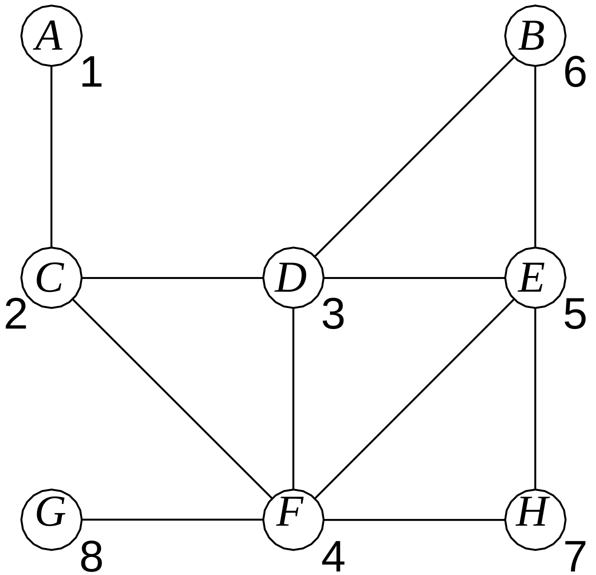
Complete (sub)graph

Perfect Ordering

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

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

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



$\alpha = (A, C, D, F, E, B, H, G)$

i	adj(v_i)	$\{v_1, \ldots, v_{i-1}\} \cap \text{adj}(v_i)$		
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

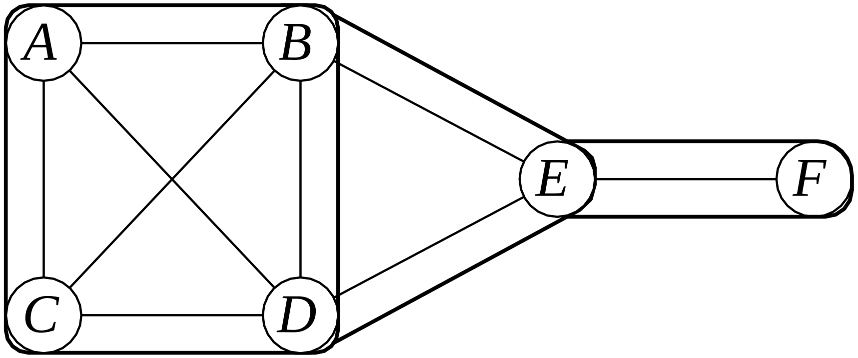
α is a perfect ordering

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



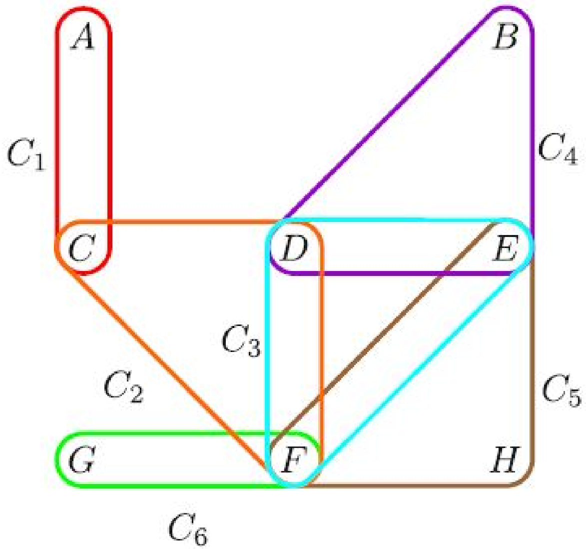
3 cliques

$C_1 = \{A, B, C, D\}$
$C_2 = \{B, D, E\}$
$C_3 = \{E, F\}$

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:

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

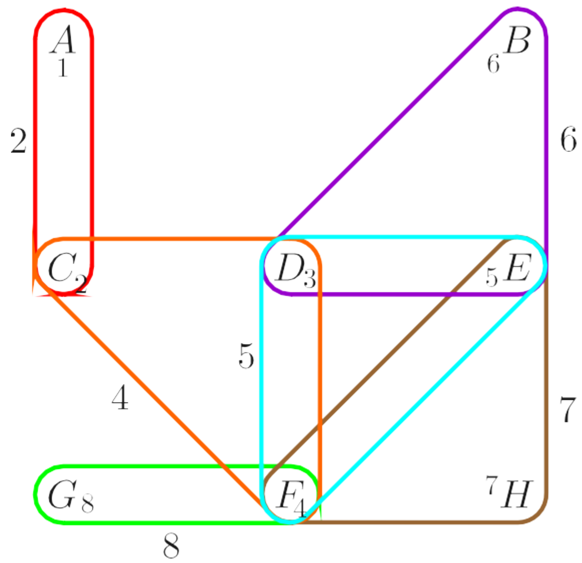


$\xi = \langle C_1,\ C_2,\ C_3,\ C_4,\ C_5,\ C_6 \rangle$

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

ξ has running intersection property

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

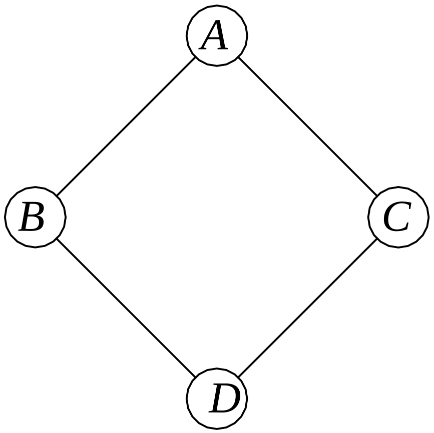


Clique	Rank		
$\{A, C\}$	$\max\{\alpha(A), \alpha(C)\}$	$= 2$	$\rightarrow C_1$
$\{C, D, F\}$	$\max\{\alpha(C), \alpha(D), \alpha(F)\}$	$= 4$	$\rightarrow C_2$
$\{D, E, F\}$	$\max\{\alpha(D), \alpha(E), \alpha(F)\}$	$= 5$	$\rightarrow C_3$
$\{B, D, E\}$	$\max\{\alpha(B), \alpha(D), \alpha(E)\}$	$= 6$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

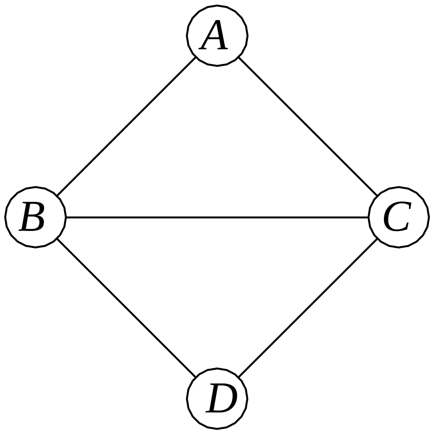
How to get a perfect ordering?

Triangulated Graph

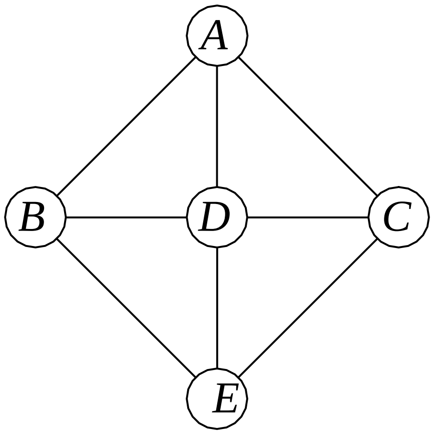
An undirected graph is called *triangulated* if every simple loop (i. e. path with identical start and end node but with any other node occurring at most once) of length greater 3 has a chord.



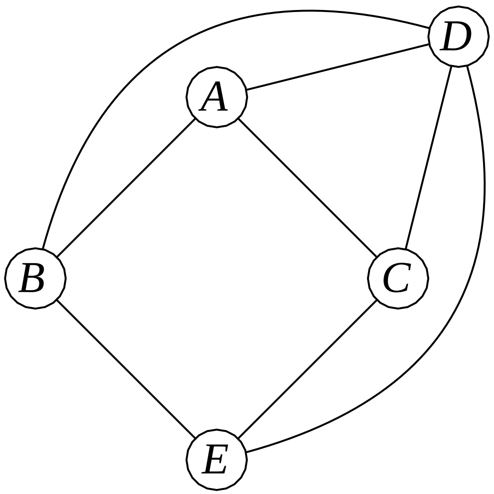
not triangulated



triangulated



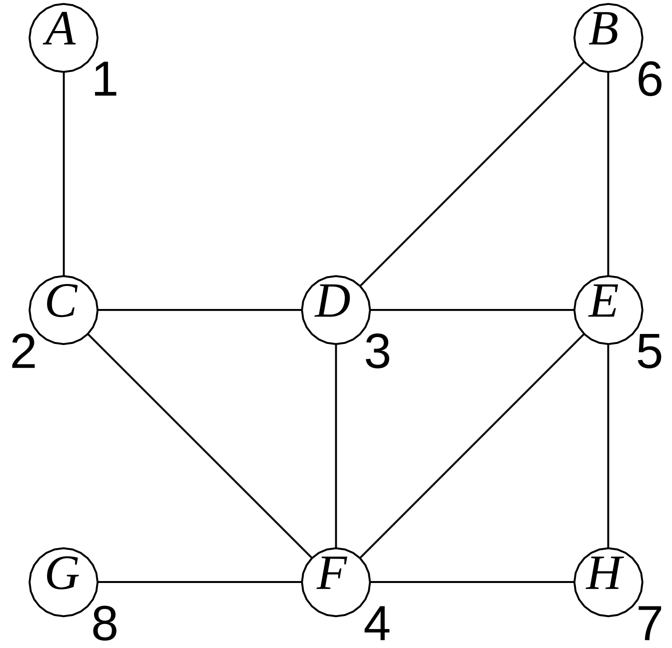
not triangulated



no chord for (A, B, E, C)

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 arbitray 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

Triangulation

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

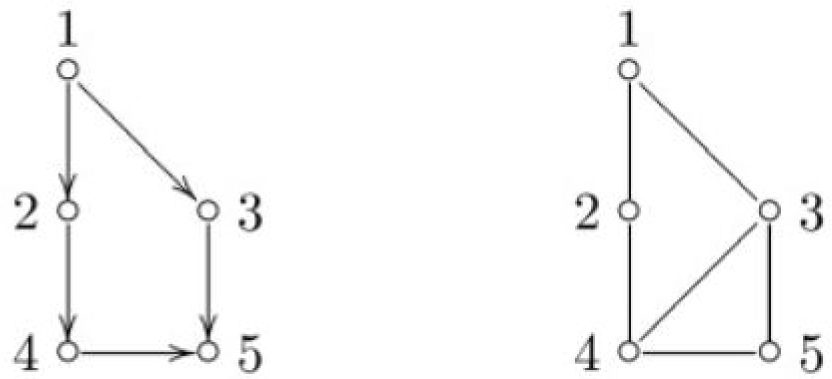
To check whether a graph is triangulated is efficient to implement.

How to find a „good“ triangulation?

The corresponding optimization problem („best“ triangulation, minimal number of additional edges) is NP-hard. However, there are heuristics for suboptimal but „good" solutions.

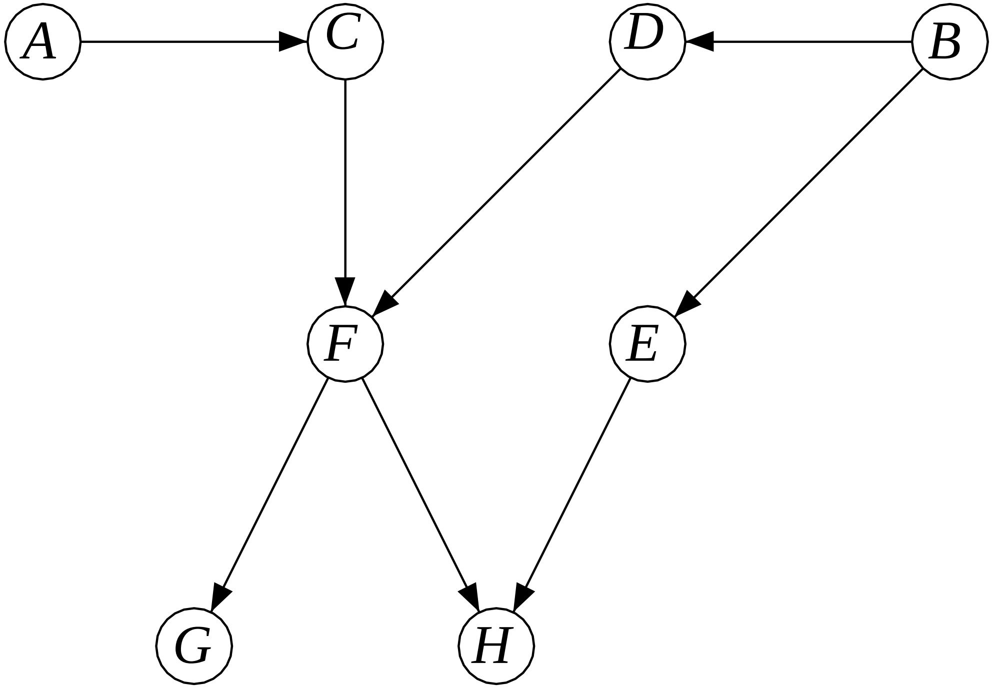
Moral Graph

Let $G = (V, E)$ be a directed acyclic graph. If $u, w \in W$ are parents of $v \in V$, then 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.

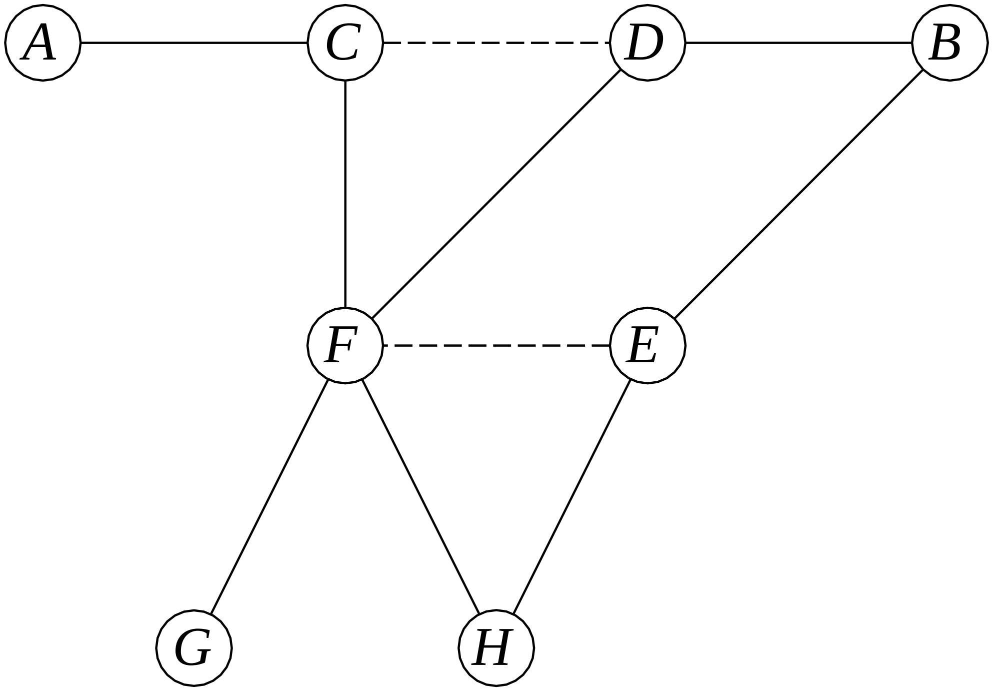


Example: Join-Tree Construction (1)

Given directed graph.

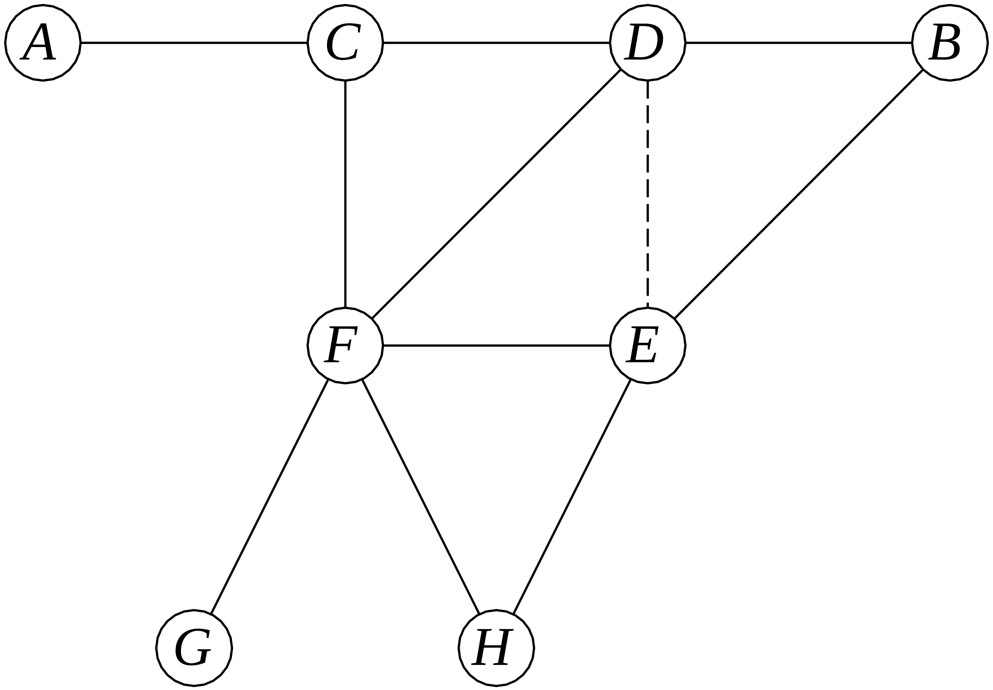


Join-Tree Construction (2)



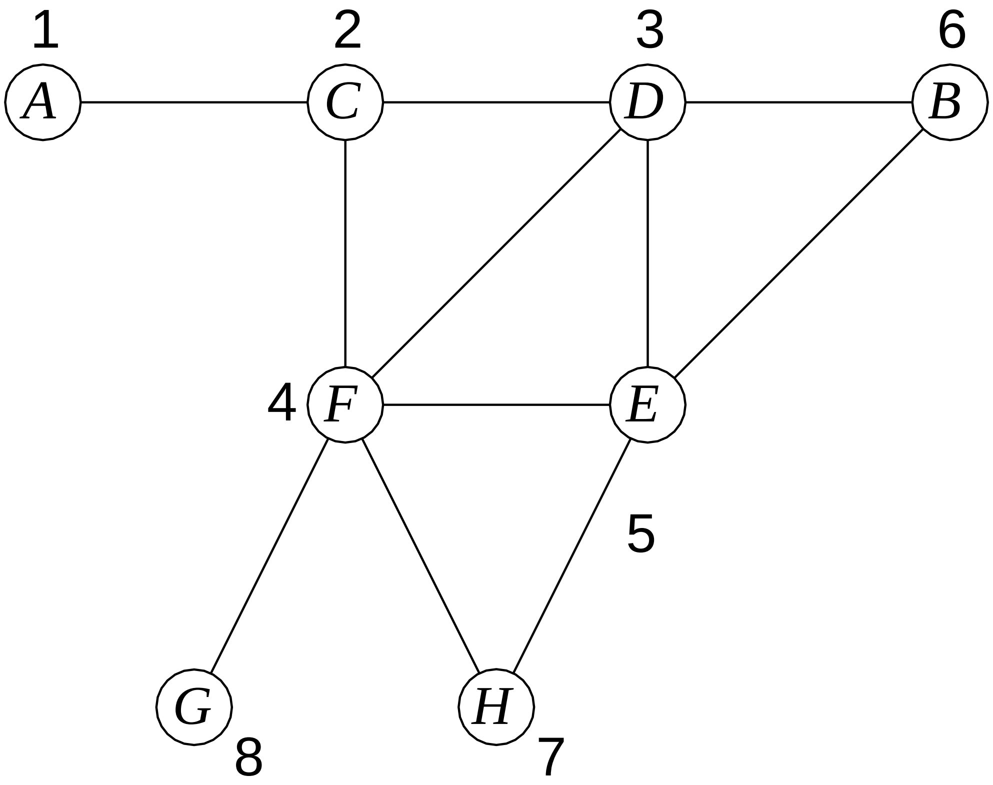
- Moral graph

Join-Tree Construction (3)

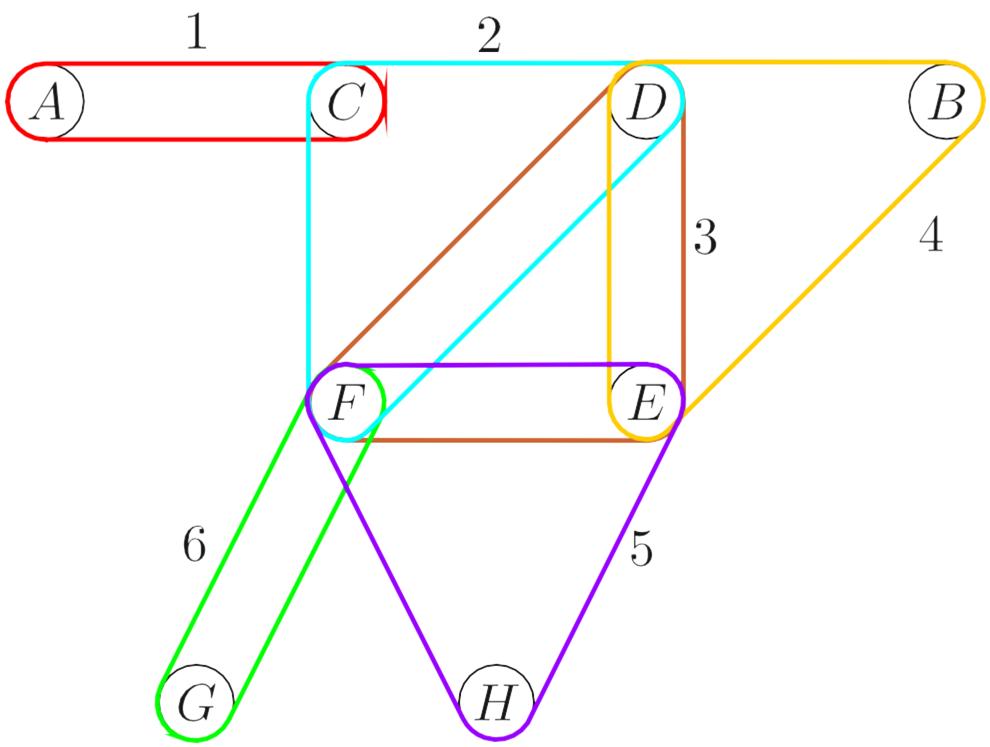


- Moral graph
- Triangulated graph

Join-Tree Construction (4)

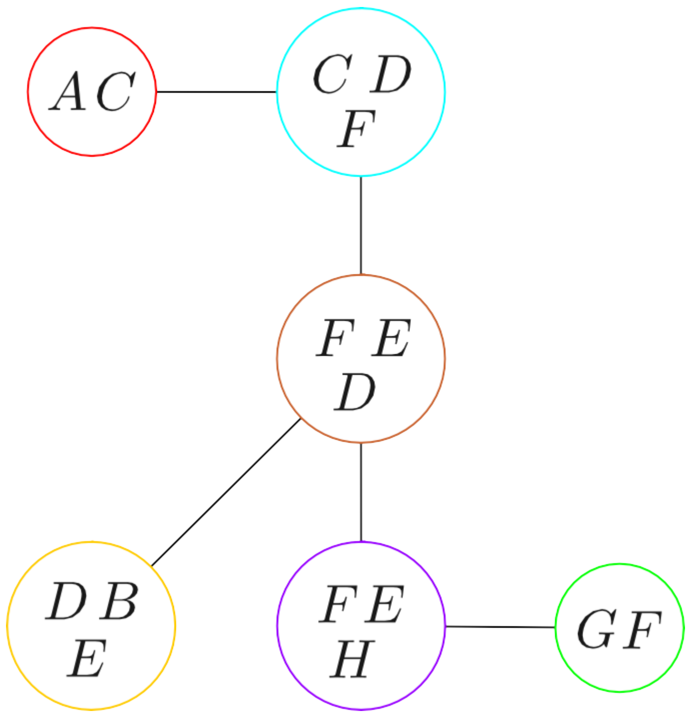


- Moral graph
- Triangulated graph
- MCS yields perfect ordering



- 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 ex- planation 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 suffers from heavy headache.

Query

Will the patient fall into coma?

Example: Choice of State Space

Attribute		Possible Values	
A	metastatic cancer	$\mathrm{dom}(A) = \{a_1, a_2\}$	$\cdot_1 =$ existing
B	increased total serum calcium	$\mathrm{dom}(B) = \{b_1, b_2\}$	$\cdot_2 =$ not existing
C	brain tumor	$\mathrm{dom}(C) = \{c_1, c_2\}$	
D	coma	$\mathrm{dom}(D) = \{d_1, d_2\}$	
E	severe headache	$\mathrm{dom}(E) = \{e_1, e_2\}$	

Exhaustive state space:

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

Marginal and conditional probabilities are of interest for the user!

Example 3: Choice of the conditional probabilities

$$\left.\begin{aligned} P(e_1 \mid c_1) &= 0.8 \\ P(e_1 \mid c_2) &= 0.6 \end{aligned}\right\}$$ headaches common, but more common if tumor present

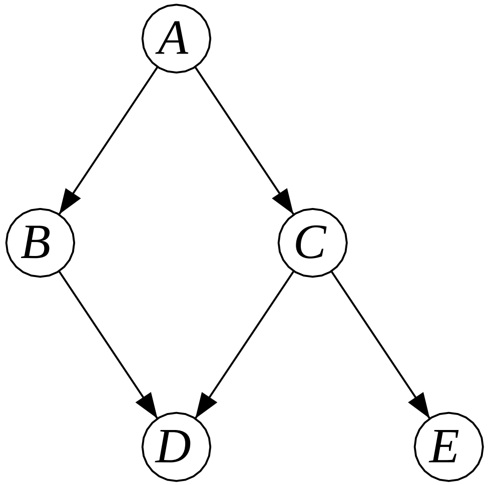
$$\left.\begin{aligned} P(d_1 \mid b_1, c_1) &= 0.8 \\ P(d_1 \mid b_1, c_2) &= 0.8 \\ P(d_1 \mid b_2, c_1) &= 0.8 \\ P(d_1 \mid b_2, c_2) &= 0.05 \end{aligned}\right\}$$ coma rare but common, if either cause is present

$$\left.\begin{aligned} P(b_1 \mid a_1) &= 0.8 \\ P(b_1 \mid a_2) &= 0.2 \end{aligned}\right\}$$ increased calcium uncommon, but common consequence of metastases

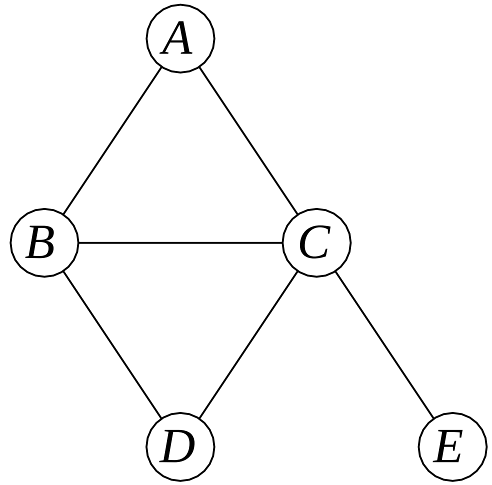
$$\left.\begin{aligned} P(c_1 \mid a_1) &= 0.2 \\ P(c_1 \mid a_2) &= 0.05 \end{aligned}\right\}$$ brain tumor rare, and uncommon consequence of metastases

$$P(a_1) = 0.2 \quad \}$$ incidence of metastatic cancer in relevant clinic

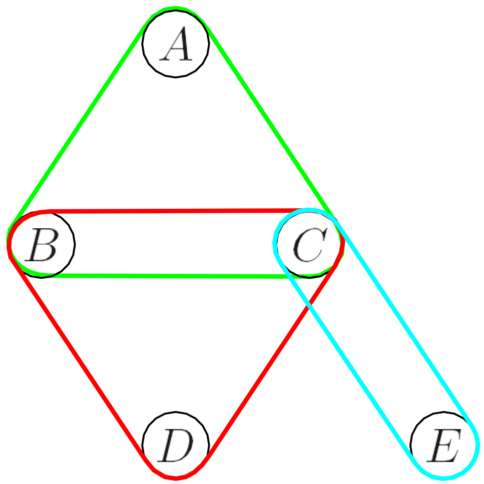
Example: Metastatic Cancer



Dependencies



Moralization/Triangulation



MCS, hyper graph

Example (2)

Quantitative knowledge:

(a, b, c)	$P(a, b, c)$
a_1, b_1, c_1	0.032
a_2, b_1, c_1	0.008
.	.
a_2, b_2, c_2	0.608

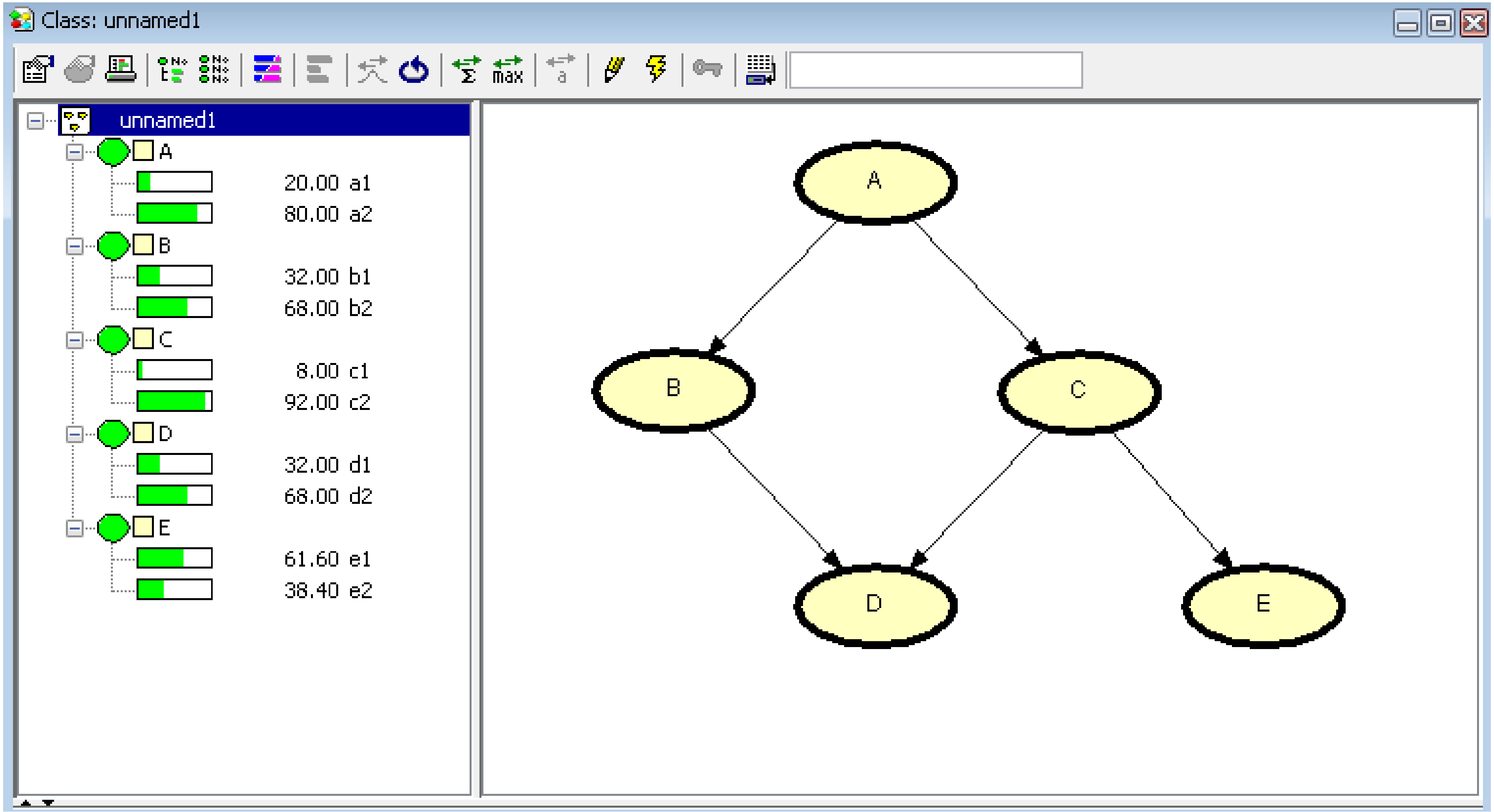
(b, c, d)	$P(b, c, d)$
b_1, c_1, d_1	0.032
b_2, c_1, d_1	0.032
.	.
b_2, c_2, d_2	0.608

(c, e)	$P(c, e)$
c_1, e_1	0.064
c_2, e_1	0.552
c_1, e_2	0.016
c_2, e_2	0.368

Decomposition:

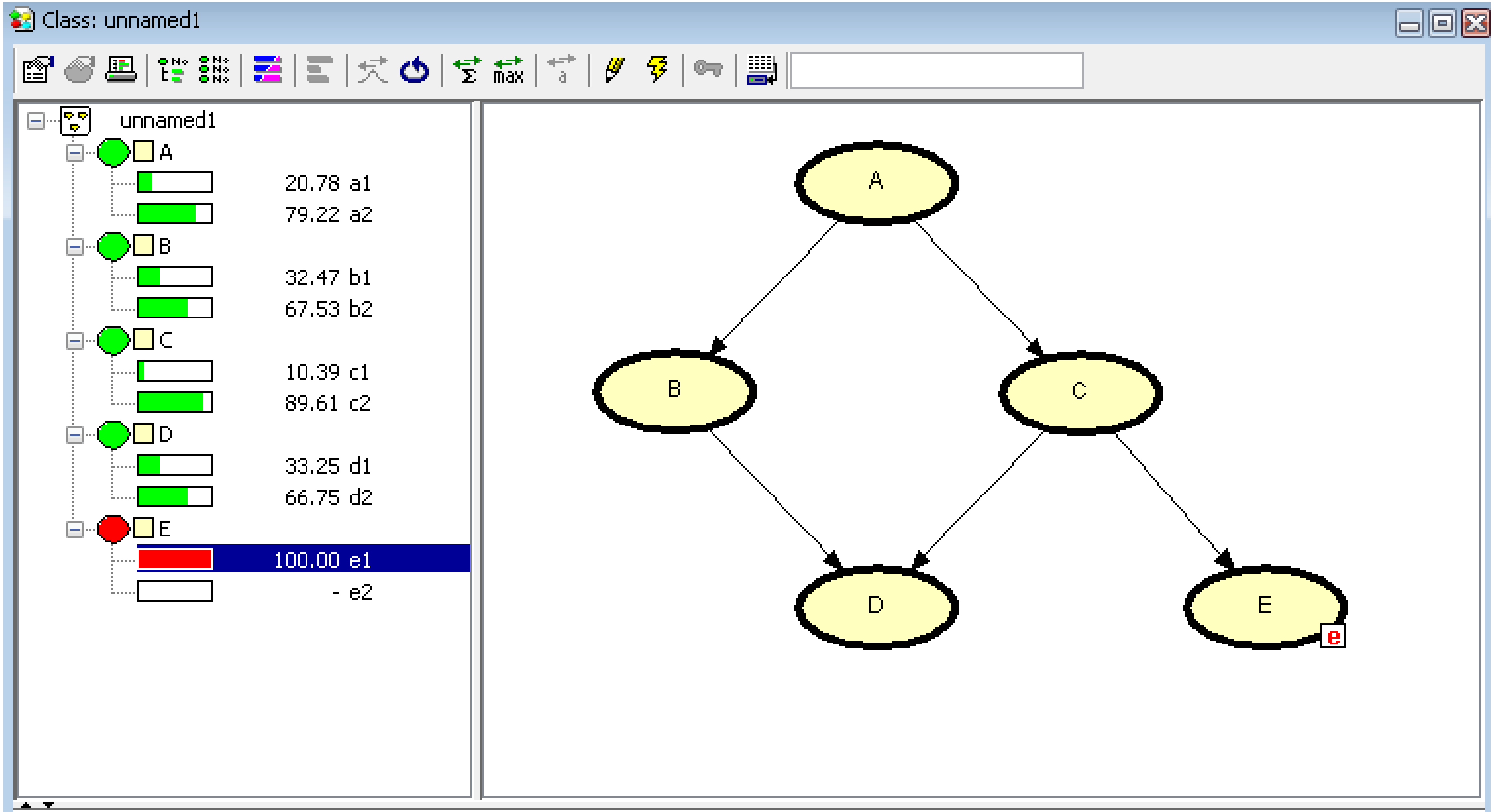
$$\begin{aligned} P(A, B, C, D, E) &= P(A)P(B \mid A)P(C \mid A)P(D \mid BC)P(E \mid C) \\ &= \frac{P(A, B, C)P(B, C, D)P(C, E)}{P(BC)P(C)} \end{aligned}$$

Example (3)



Marginal distributions in the HUGIN tool.

Example (4)



Conditional marginal distributions with evidence $E = e_1$

Potential Representation

Let $V = \{X_j\}$ be a set of random variables $X_j : \Omega \to \text{dom}(X_j)$ and P the joint distribution over V. Further, let

$$\{W_i \mid W_i \subseteq V, 1 \leq i \leq p\}$$

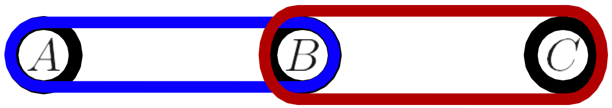
a family of subsets of V with associated functions

$$\psi_i : \bigtimes_{X_j \in W_i} \text{dom}(X_j) \to \mathbb{R}$$

It is said that $P(V)$ *factorizes* according $\big(\{W_1, \ldots, W_p\}, \{\psi_1, \ldots, \psi_p\}\big)$ if $P(V)$ can be written as:

$$P(v) = k \cdot \prod_{i=1}^{p} \psi_i(w_i)$$

where $k \in \mathbb{R}$, w_i is a realization of W_i that meets the values of v.



$V = \{A, B, C\}, \; W_1 = \{A, B\}, \; W_2 = \{B, C\}$

$\text{dom}(A) = \{a_1, a_2\}$
$\text{dom}(B) = \{b_1, b_2\}$
$\text{dom}(C) = \{c_1, c_2\}$

$P(a, b, c) = \frac{1}{8}$

$\psi_1 : \{a_1, a_2\} \times \{b_1, b_2\} \rightarrow \mathbb{R}$
$\psi_2 : \{b_1, b_2\} \times \{c_1, c_2\} \rightarrow \mathbb{R}$

$\psi_1(a, b) = \frac{1}{4}$
$\psi_2(b, c) = \frac{1}{2}$

$(\{W_1, W_2\}, \{\psi_1, \psi_2\})$ is a representation of P

Factorization of a Belief Network

Let (V, E, P) be a belief network and $\{C_1, \ldots, C_p\}$ the cliques of the join tree. For every node $v \in V$ choose a clique C such that v and all of its parents are contained in C, i. e. $\{v\} \cup c(v) \subseteq C$. The chosen clique is designated as $f(v)$.

To arrive at a factorization $(\{C_1, \ldots, C_p\}, \{\psi_1, \ldots, \psi_p\})$ of P, we define

$$\psi_i(C_i) = \prod_{v: f(v)=C_i} P(v \mid c(v))$$

In the Markov random field literature the clique functions are generally referred to as potential functions.

Separator Sets and Residual Sets

Let $\{C_1, \ldots, C_p\}$ be a set of cliques w. r. t. V. The sets

$$S_i = C_i \cap (C_1 \cup \cdots \cup C_{i-1}), \quad i = 2, \ldots, p, \qquad S_1 = \emptyset$$

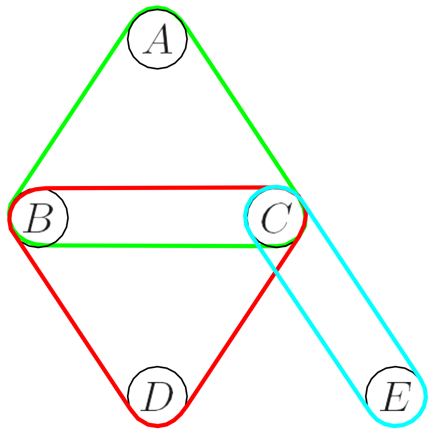
are called *separator sets* with their corresponding *residual sets*

$$R_i = C_i \backslash S_i$$

Given a clique ordering $\{C_1, \ldots, C_p\}$ that satisfies the running intersection property (RIP), we can conclude the following separation statements:

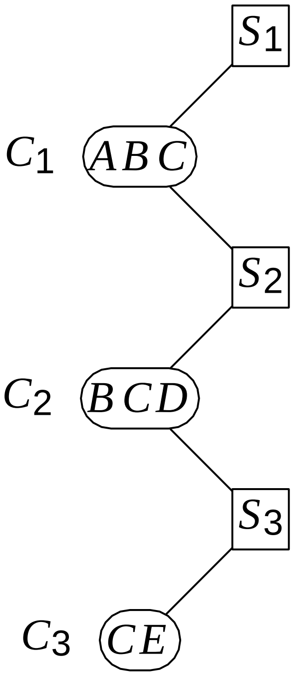
$$R_i \perp\!\!\!\perp C_1 \cup \cdots \cup C_{i-1})\backslash S_i \mid S_i \qquad \text{for } i > 1$$

Example 2



$$S_1 = \emptyset \qquad R_1 = \{A, B, C\}$$
$$S_2 = \{B, C\} \qquad R_2 = \{D\}$$
$$S_3 = \{C\} \qquad R_3 = \{E\}$$

$$f(A) = C_1$$
$$f(B) = C_1$$
$$f(C) = C_1$$
$$f(D) = C_2$$
$$f(E) = C_3$$



$$\psi_1(C_1) = P(A) \cdot P(C \mid A) \cdot P(B|A)$$
$$\psi_2(C_2) = P(D \mid B, C)$$
$$\psi_3(C_3) = P(E \mid C)$$

Propagation is accomplished by sending messages across the cliques in the tree. The emerging potentials are maintained by each clique.

A Few 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