Propagation in Belief Networks

Objective

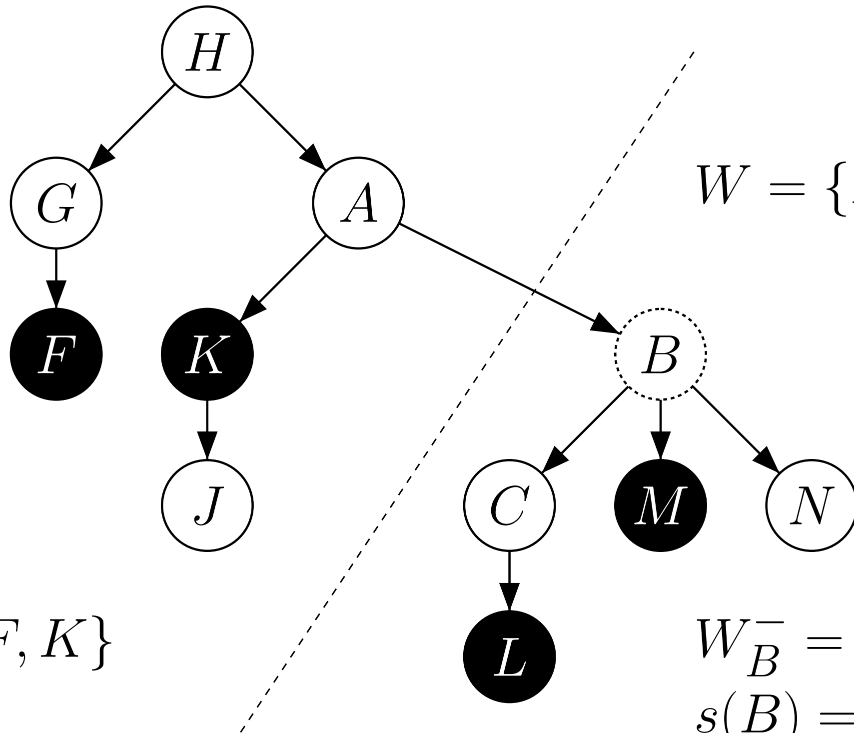
- **Given:** Belief network (V, E, P) with tree structure and $P(V) > 0$.
  Set $W \subseteq V$ of instantiated variables where
  a priori knowledge $W \neq \emptyset$ is allowed

- **Desired:** $P(B \mid W)$ for all $B \in V$

- **Notation:**

W_B^-	subset of those variables of W that belong to the subtree of (V, E) that has root B
W_B^+	$= W \backslash W_B^-$
$s(B)$	set of direct successors of B
Ω_B	domain of B
b^*	value that B is instantiated with

Example



$$W = \{F, K, L, M\}$$

$$W_B^+ = \{F, K\}$$

$$W_B^- = \{L, M\}$$

$$s(B) = \{C, M, N\}$$

Decomposition in the Tree

$$
\begin{aligned}
P(B = b \mid W) &= P(b \mid W_B^- \cup W_B^+) \qquad \text{with } B \notin W \\
&= \frac{P(W_B^- \cup W_B^+ \cup \{b\})}{P(W_B^- \cup W_B^+)} \\
&= \frac{P(W_B^- \cup W_B^+ \mid b) P(b)}{P(W_B^- \cup W_B^+)} \\
&= \frac{P(W_B^- \mid b) P(W_B^+ \mid b) P(b)}{P(W_B^- \cup W_B^+)} \\
&= \beta_{B,W} \quad \underbrace{P(W_B^- \mid b)}_{\text{Evidence from "below"}} \quad \underbrace{P(b \mid W_B^+)}_{\text{Evidence from "above"}}
\end{aligned}
$$

Since we ignore the constant $\beta_{B,W}$ for the derivations below, the following designations are used instead of $P(\cdot)$:

π-values and λ-values

Let $B \in V$ be a variable and $b \in \Omega_B$ a value of its domain. We define the π- and λ-values as follows:

$$\lambda(b) = \begin{cases} P(W_B^- \mid b) & \text{if } B \notin W \\ \quad 1 & \text{if } B \in W \wedge b^* = b \\ \quad 0 & \text{if } B \in W \wedge b^* \neq b \end{cases}$$

$$\pi(b) = P(b \mid W_B^+)$$

π- and λ-Values

$$\lambda(b) = \prod_{C \in s(B)} P(W_C^- \mid b) \qquad \text{if } B \in W$$

$$\lambda(b) = 1 \qquad \text{if } B \text{ leaf in } (V, E)$$

$$\pi(b) = P(b) \qquad \text{if } B \text{ root in } (V, E)$$

$$P(b \mid W) = \alpha_{B,W} \cdot \lambda(b) \cdot \pi(b)$$

λ-message

Let $B \in V$ be an attribute and $C \in s(B)$ its direct children with the respective domains $\mathrm{dom}(B) = \{B_1, \ldots, b_i, \ldots, b_k\}$ and $\mathrm{dom}(C) = \{c_1, \ldots, c_j, \ldots, c_m\}$.

$$\lambda_{C \to B}(b_i) \overset{\text{Def}}{=} \sum_{j=1}^{m} P(c_j \mid b_i) \cdot \lambda(c_j), \qquad i = 1, \ldots, k$$

The vector

$$\vec{\lambda}_{C \to B} \overset{\text{Def}}{=} \Big(\lambda_{C \to B}(b_i)\Big)_{i=1}^{k}$$

is called λ-*message* from C to B.

λ-Message

Let $B \in V$ an attribute an $b \in \mathrm{dom}(B)$ a value of its domain.
Then

$$\lambda(b) = \begin{cases} \rho_{B,W} \cdot \prod\limits_{C \in s(B)} \lambda_C(b) & \text{if } B \notin W \\ 1 & \text{if } B \in W \wedge b = b^* \\ 0 & \text{if } B \in W \wedge b \neq b^* \end{cases}$$

with $\rho_{B,W}$ being a positive constant.

π-message

Let $B \in V$ be a non-root node in (V, E) and $A \in V$ its parent with domain $\mathrm{dom}(A) = \{a_1, \ldots, a_j, \ldots, a_m\}$.

$$j = 1, \ldots, m:$$
$$\pi_{A \to B}(a_j) \stackrel{\text{Def}}{=} \begin{cases} \pi(a_j) \cdot \prod\limits_{C \in s(A) \setminus \{B\}} \lambda_C(a_j) & \text{if } A \notin W \\ 1 & \text{if } A \in W \wedge a = a^* \\ 0 & \text{if } A \in W \wedge a \neq a^* \end{cases}$$

The vector

$$\vec{\pi}_{A \to B} \stackrel{\text{Def}}{=} \Big(\pi_{A \to B}(a_j)\Big)_{j=1}^{m}$$

is called π-*message* from A to B.

π-Message

Let $B \in V$ be a non-root node in (V, E) and A the parent node of B.
Further let $b \in \mathrm{dom}(B)$ be a value of B's domain.

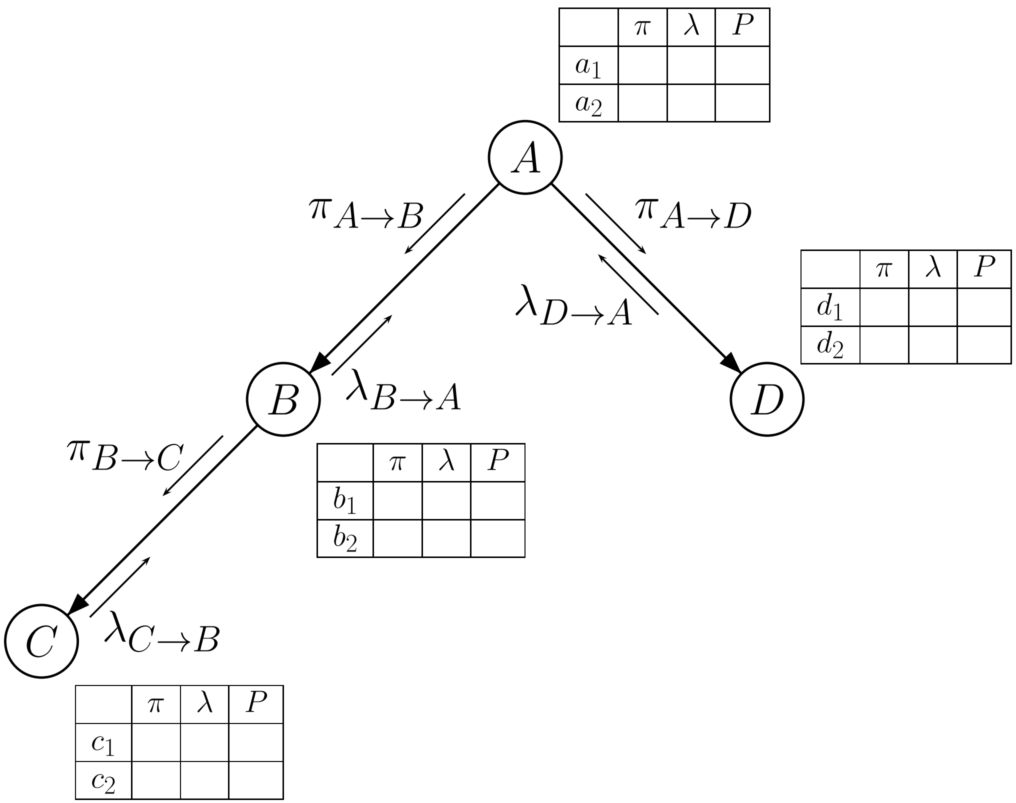
$$\pi(b) \;=\; \mu_{B,W} \;\cdot \sum_{a \in \mathrm{dom}(A)} P(b \mid a) \cdot \pi_{A \to B}(a)$$

Let $A \notin W$ a non-instantiated attribute and $P(V) > 0$.

$$\begin{aligned}
\pi_{A \to B}(a_j) \;&=\; \pi(a_j) \;\cdot \prod_{C \in s(A) \setminus \{B\}} \lambda_{C \to A}(a_j) \\
&=\; \tau_{B,W} \cdot \frac{P(a_j \mid W)}{\lambda_{B \to A}(a_j)}
\end{aligned}$$

Propagation in Belief Trees

Belief Tree:



Parameters:

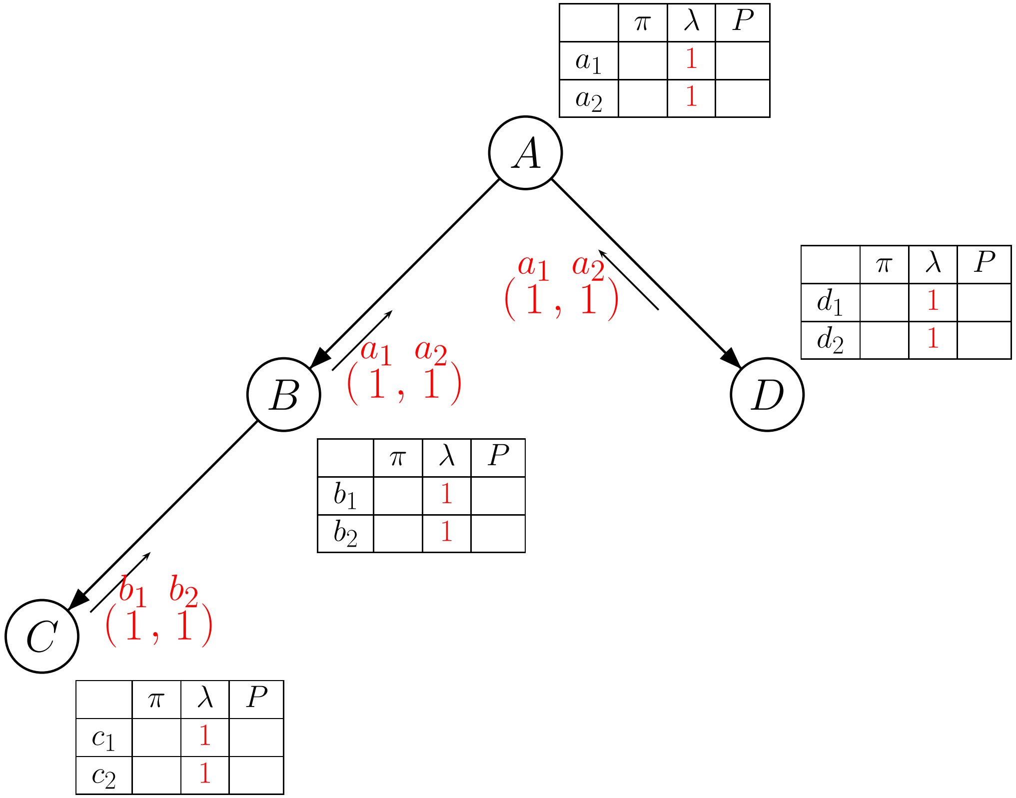
$$P(a_1) = 0.1 \quad P(b_1 \mid a_1) = 0.7$$
$$P(b_1 \mid a_2) = 0.2$$
$$P(d_1 \mid a_1) = 0.8 \quad P(c_1 \mid b_1) = 0.4$$
$$P(d_1 \mid a_2) = 0.4 \quad P(c_1 \mid b_2) = 0.001$$

Desired:

$$\forall X \in \{A, B, C, D\}: \; P(X \mid \emptyset) = ?$$

Propagation in Belief Trees (2)

Belief Tree:

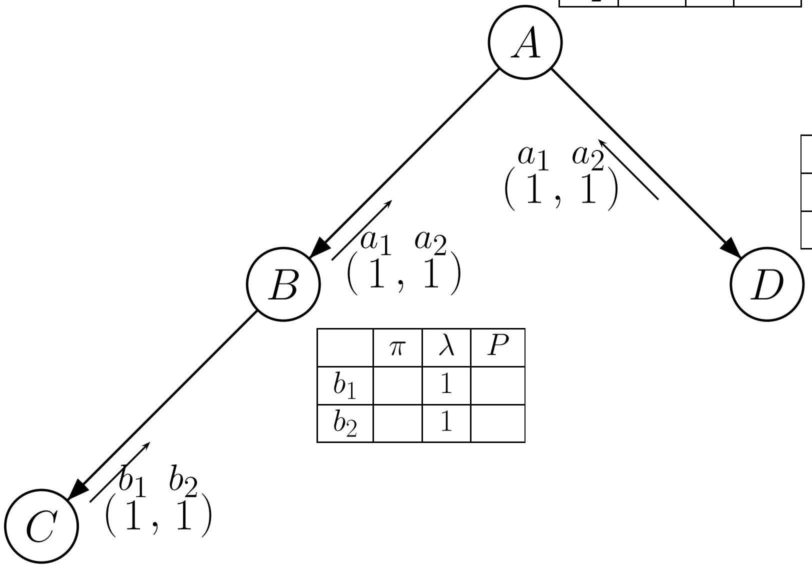


Initialization Phase:

- Set all λ-messages and λ-values to 1.

Propagation in Belief Trees (3)

Belief Tree:



	π	λ	P
a_1	0.1	1	0.1
a_2	0.9	1	0.9

	π	λ	P
b_1		1	
b_2		1	

	π	λ	P
c_1		1	
c_2		1	

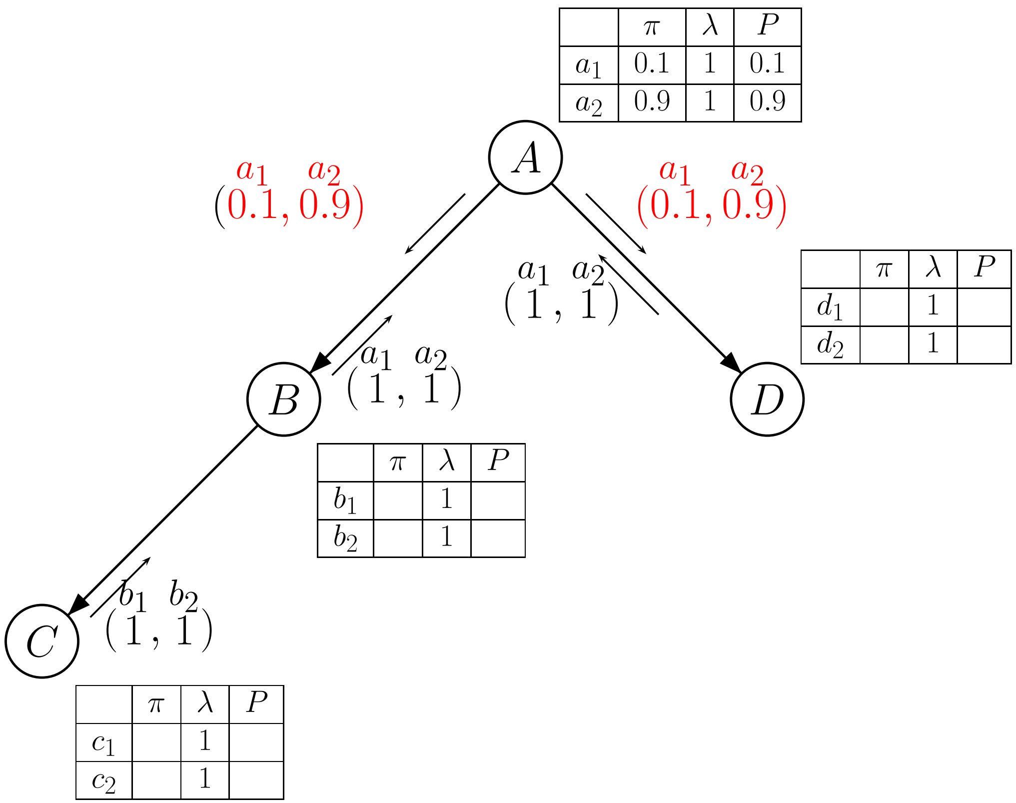
	π	λ	P
d_1		1	
d_2		1	

Initialization Phase:

- Set all λ-messages and λ-values to 1.
- $\pi(a_1) = P(a_1)$ and $\pi(a_2) = P(a_2)$

Propagation in Belief Trees (4)

Belief Tree:



	π	λ	P
a_1	0.1	1	0.1
a_2	0.9	1	0.9

	π	λ	P
d_1		1	
d_2		1	

	π	λ	P
b_1		1	
b_2		1	

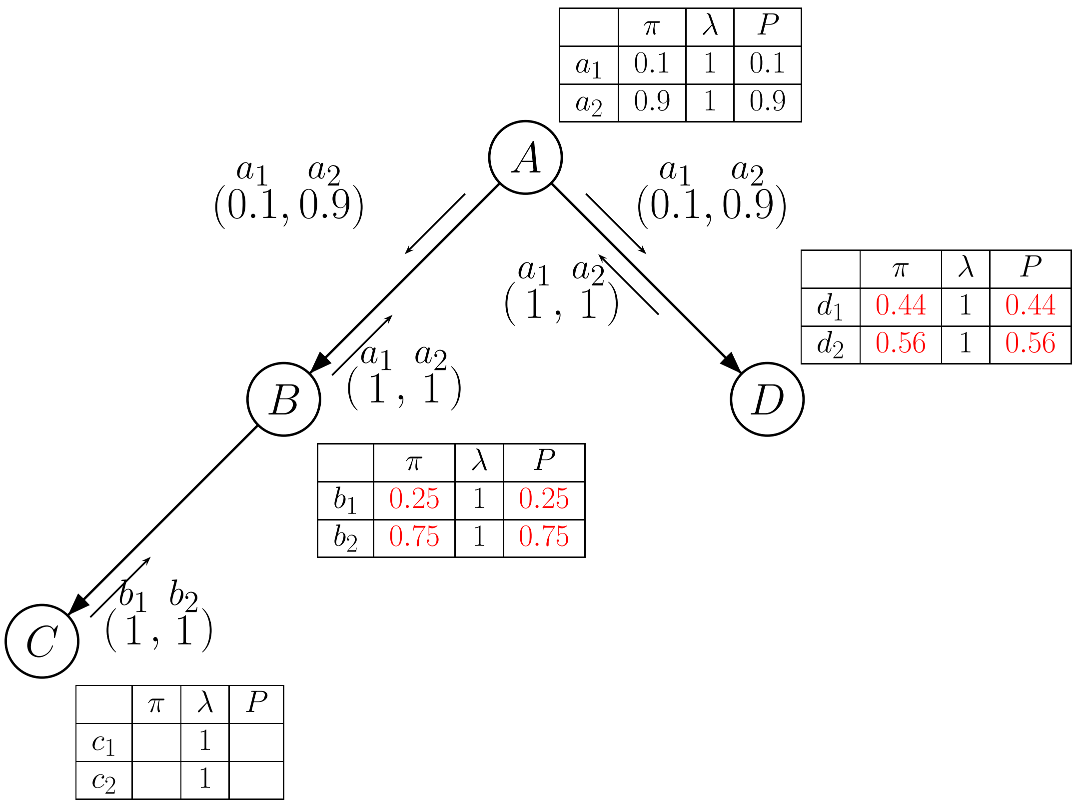
	π	λ	P
c_1		1	
c_2		1	

Initialization Phase:

- Set all λ-messages and λ-values to 1.
- $\pi(a_1) = P(a_1)$ and $\pi(a_2) = P(a_2)$.
- A sends π-messages to B and D.

Propagation in Belief Trees (5)

Belief Tree:



	π	λ	P
a_1	0.1	1	0.1
a_2	0.9	1	0.9

	π	λ	P
b_1	0.25	1	0.25
b_2	0.75	1	0.75

	π	λ	P
c_1		1	
c_2		1	

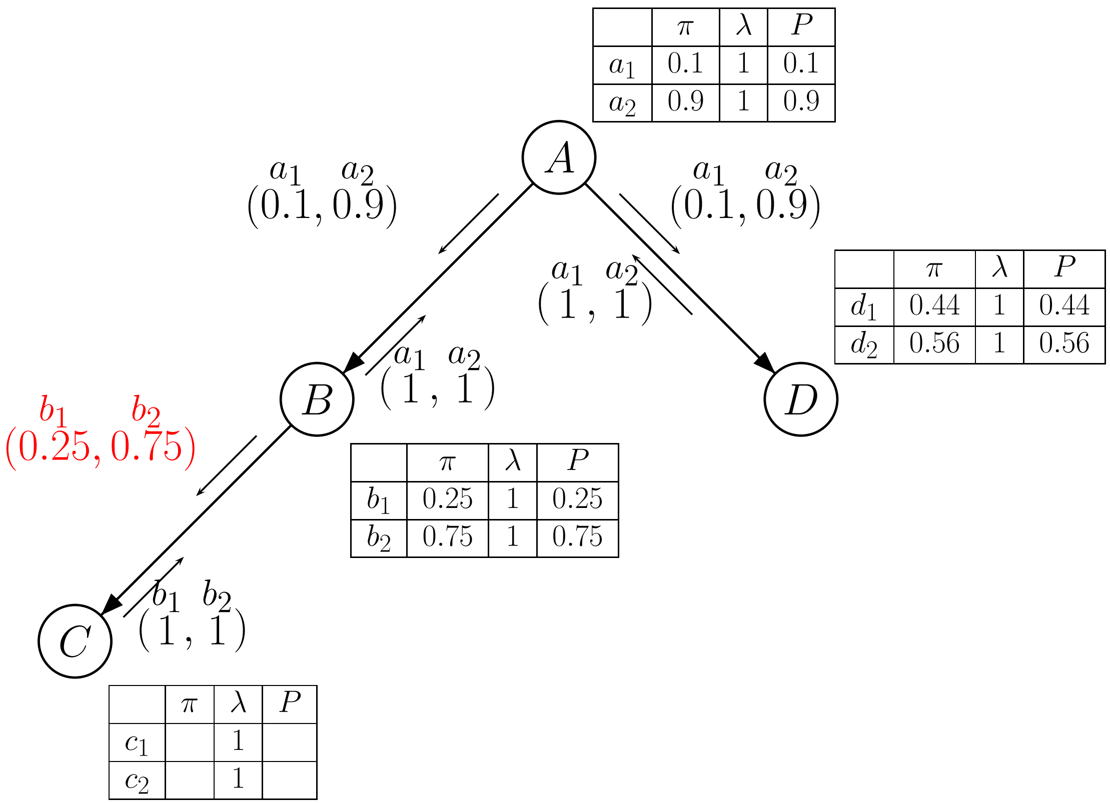
	π	λ	P
d_1	0.44	1	0.44
d_2	0.56	1	0.56

Initialization Phase:

- Set all λ-messages and λ-values to 1.
- $\pi(a_1) = P(a_1)$ and $\pi(a_2) = P(a_2)$.
- A sends π-messages to B and D.
- B and D update their π-values.

Propagation in Belief Trees (6)

Belief Tree:

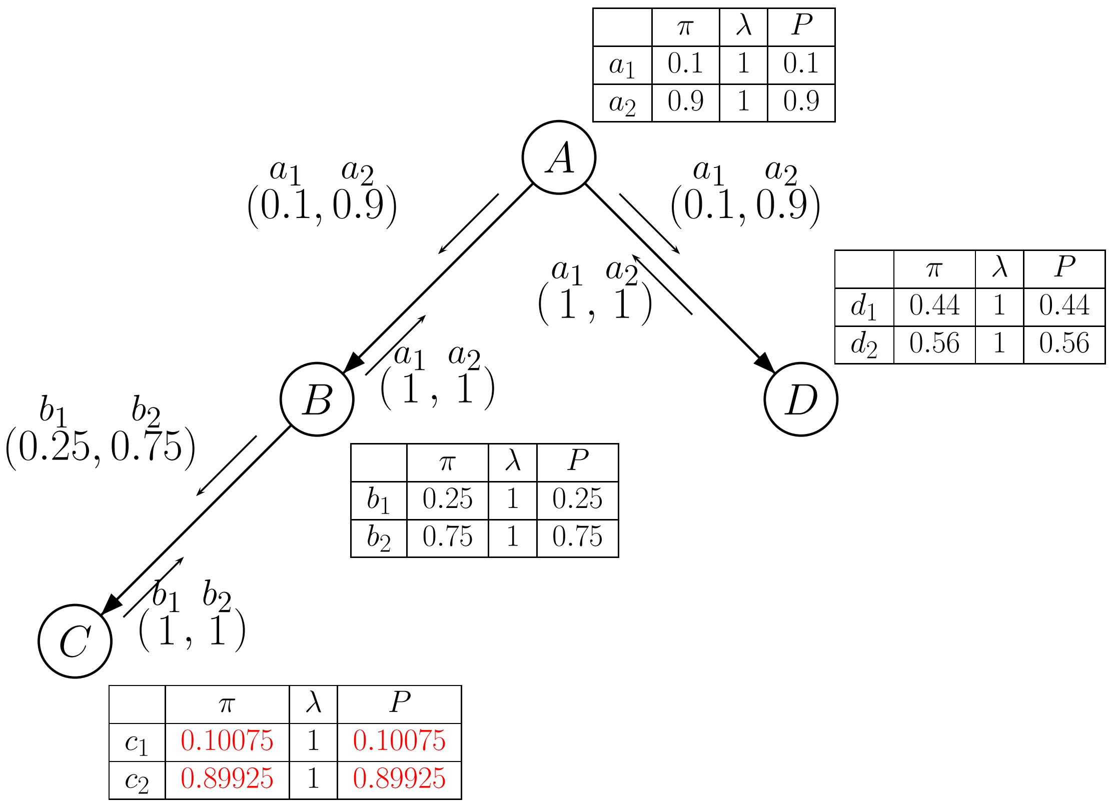


Initialization Phase:

- Set all λ-messages and λ-values to 1.
- $\pi(a_1) = P(a_1)$ and $\pi(a_2) = P(a_2)$.
- A sends π-messages to B and D.
- B and D update their π-values.
- B sends π-message to C.

Propagation in Belief Trees (7)

Belief Tree:



	π	λ	P
a_1	0.1	1	0.1
a_2	0.9	1	0.9

	π	λ	P
d_1	0.44	1	0.44
d_2	0.56	1	0.56

	π	λ	P
b_1	0.25	1	0.25
b_2	0.75	1	0.75

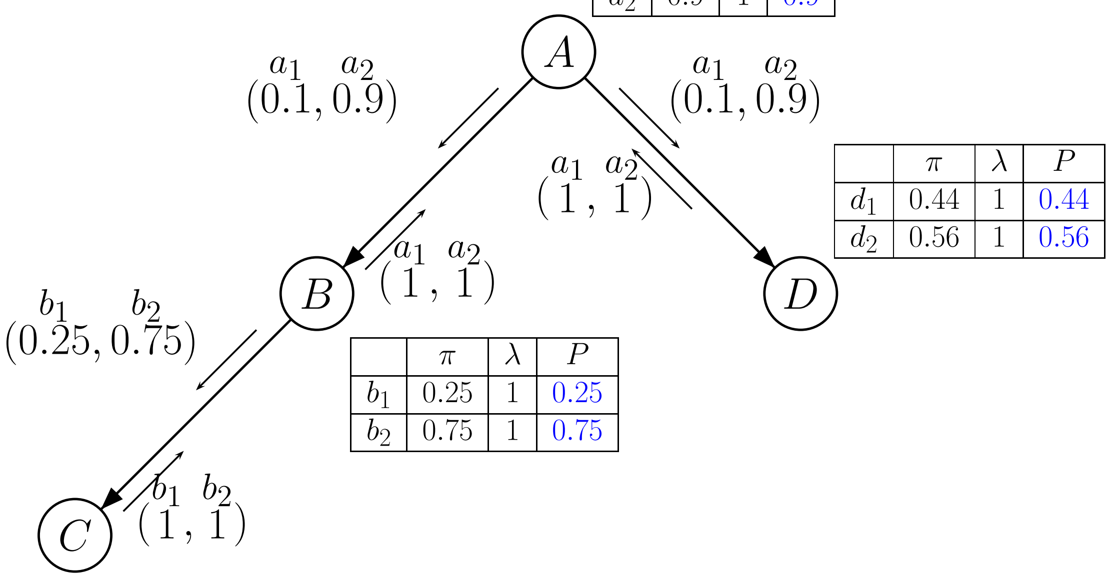
	π	λ	P
c_1	0.10075	1	0.10075
c_2	0.89925	1	0.89925

Initialization Phase:

- Set all λ-messages and λ-values to 1.
- $\pi(a_1) = P(a_1)$ and $\pi(a_2) = P(a_2)$.
- A sends π-messages to B and D.
- B and D update their π-values.
- B sends π-message to C.
- C updates it π-value.

Propagation in Belief Trees (8)

Belief Tree:



	π	λ	P
a_1	0.1	1	0.1
a_2	0.9	1	0.9

	π	λ	P
d_1	0.44	1	0.44
d_2	0.56	1	0.56

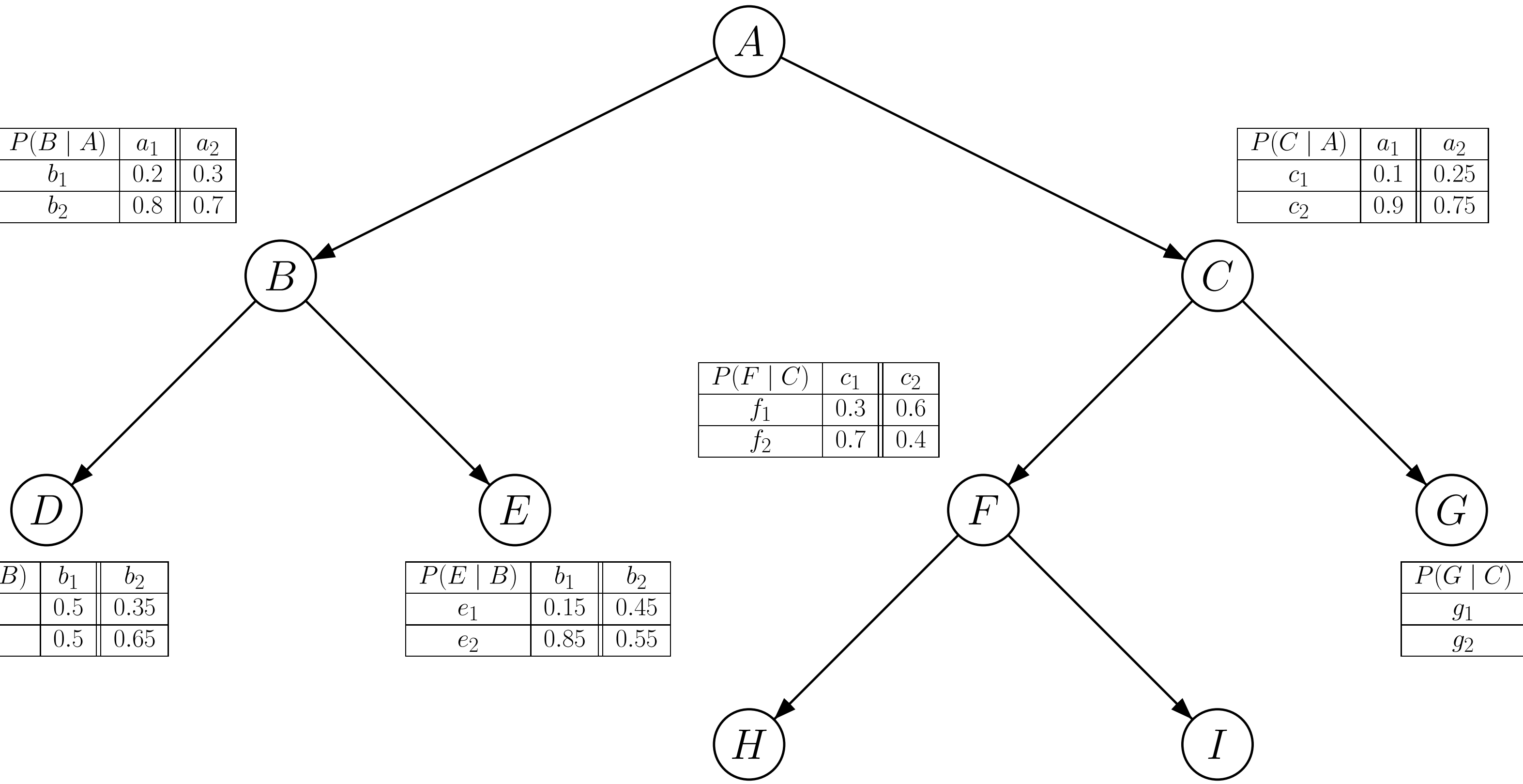
	π	λ	P
b_1	0.25	1	0.25
b_2	0.75	1	0.75

	π	λ	P
c_1	0.10075	1	0.10075
c_2	0.89925	1	0.89925

Initialization Phase:

- Set all λ-messages and λ-values to 1.
- $\pi(a_1) = P(a_1)$ and $\pi(a_2) = P(a_2)$.
- A sends π-messages to B and D.
- B and D update their π-values.
- B sends π-message to C.
- C updates it π-value.
- Initialization finished.

Larger Network (1): Parameters



$P(A)$	$\emptyset$
a_1	0.4
a_2	0.6

$P(B \mid A)$	a_1	a_2
b_1	0.2	0.3
b_2	0.8	0.7

$P(C \mid A)$	a_1	a_2
c_1	0.1	0.25
c_2	0.9	0.75

$P(F \mid C)$	c_1	c_2
f_1	0.3	0.6
f_2	0.7	0.4

$P(D \mid B)$	b_1	b_2
d_1	0.5	0.35
d_2	0.5	0.65

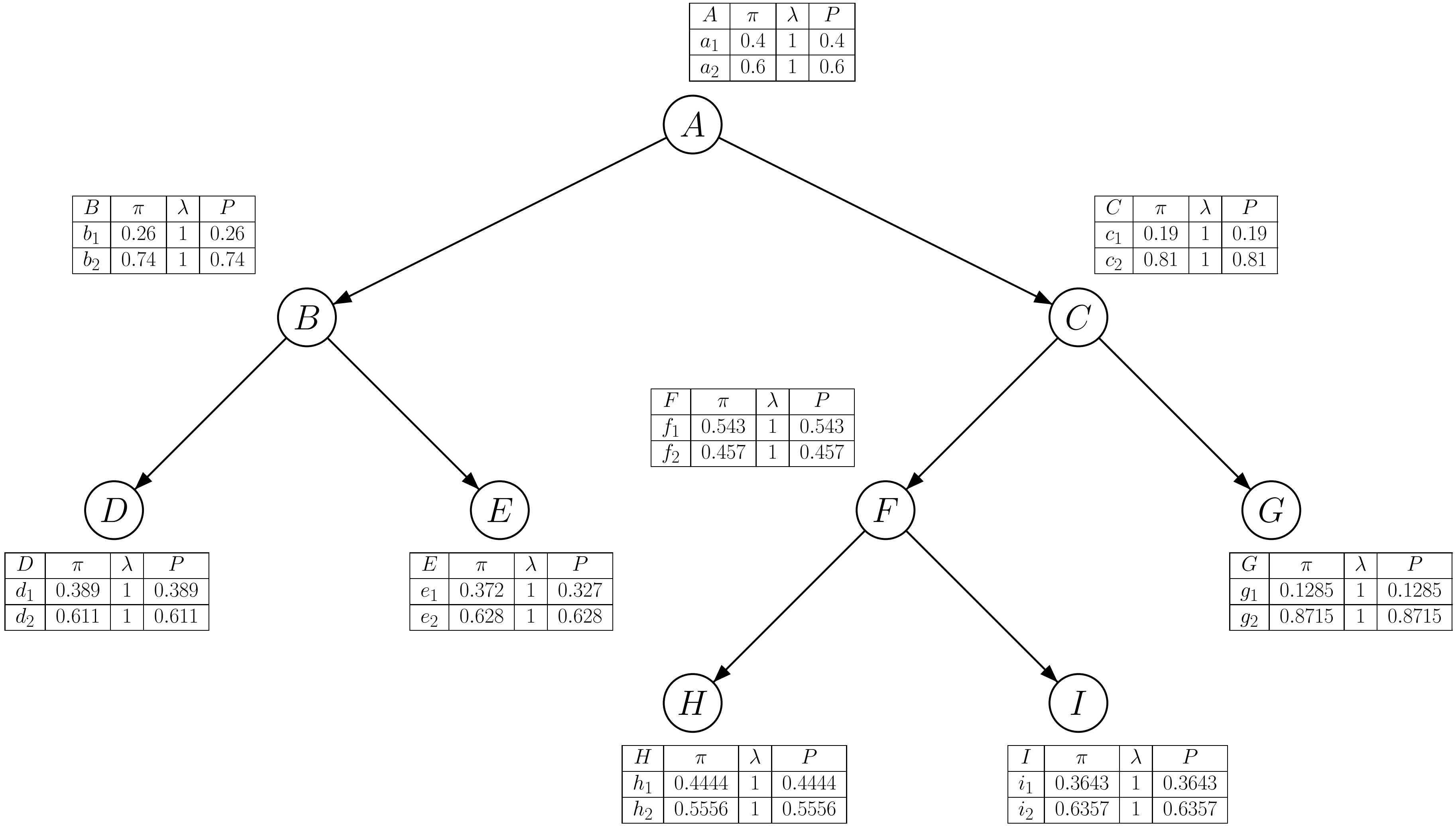
$P(E \mid B)$	b_1	b_2
e_1	0.15	0.45
e_2	0.85	0.55

$P(G \mid C)$	c_1	c_2
g_1	0.25	0.1
g_2	0.75	0.9

$P(H \mid F)$	f_1	f_2
h_1	0.65	0.2
h_2	0.35	0.8

$P(I \mid F)$	f_1	f_2
i_1	0.25	0.5
i_2	0.75	0.5

Larger Network (2): After Initialization



A	π	λ	P
a_1	0.4	1	0.4
a_2	0.6	1	0.6

B	π	λ	P
b_1	0.26	1	0.26
b_2	0.74	1	0.74

C	π	λ	P
c_1	0.19	1	0.19
c_2	0.81	1	0.81

F	π	λ	P
f_1	0.543	1	0.543
f_2	0.457	1	0.457

D	π	λ	P
d_1	0.389	1	0.389
d_2	0.611	1	0.611

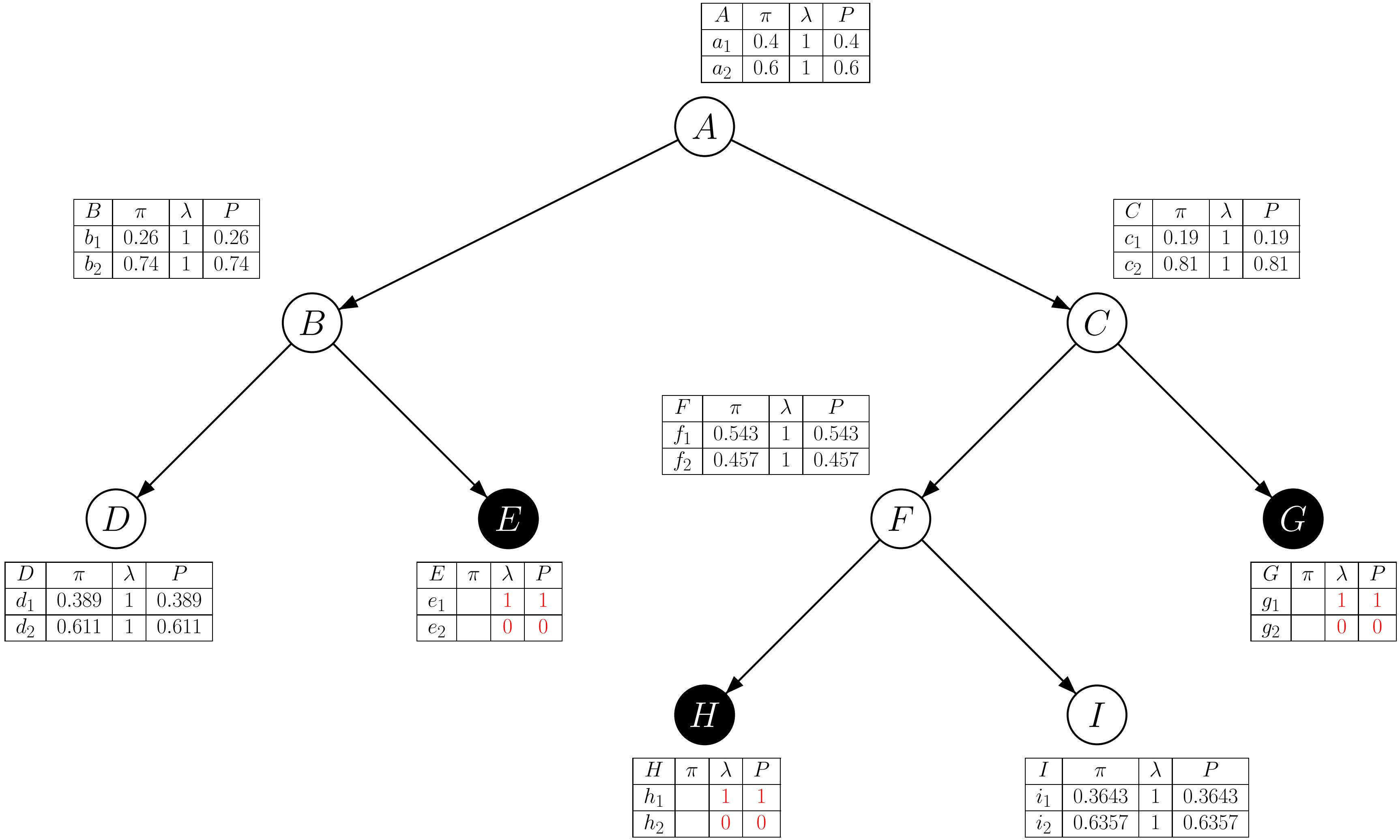
E	π	λ	P
e_1	0.372	1	0.327
e_2	0.628	1	0.628

G	π	λ	P
g_1	0.1285	1	0.1285
g_2	0.8715	1	0.8715

H	π	λ	P
h_1	0.4444	1	0.4444
h_2	0.5556	1	0.5556

I	π	λ	P
i_1	0.3643	1	0.3643
i_2	0.6357	1	0.6357

Larger Network (3): Set Evidence e_1, g_1, h_1



A	π	λ	P
a_1	0.4	1	0.4
a_2	0.6	1	0.6

B	π	λ	P
b_1	0.26	1	0.26
b_2	0.74	1	0.74

C	π	λ	P
c_1	0.19	1	0.19
c_2	0.81	1	0.81

F	π	λ	P
f_1	0.543	1	0.543
f_2	0.457	1	0.457

D	π	λ	P
d_1	0.389	1	0.389
d_2	0.611	1	0.611

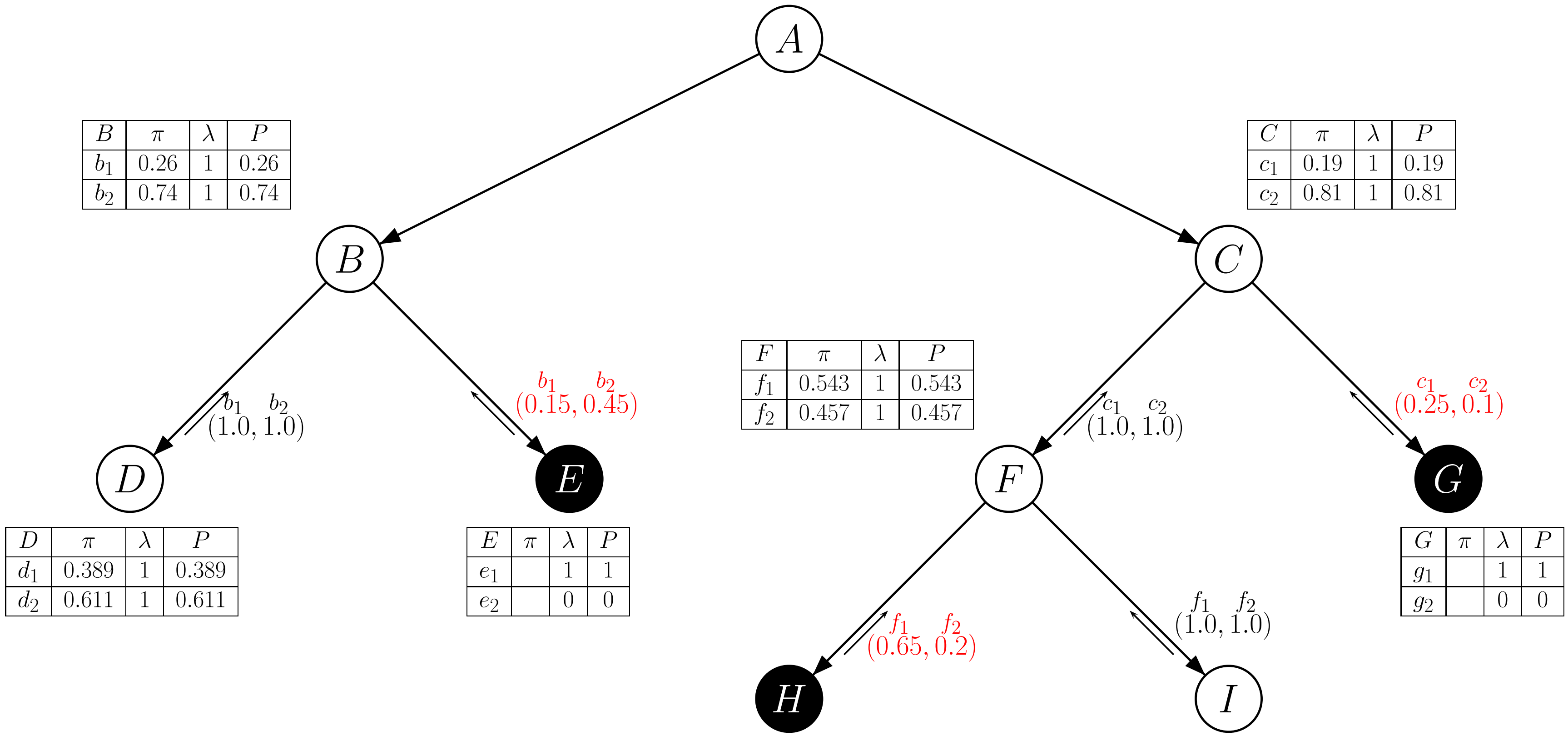
E	π	λ	P
e_1		1	1
e_2		0	0

G	π	λ	P
g_1		1	1
g_2		0	0

H	π	λ	P
h_1		1	1
h_2		0	0

I	π	λ	P
i_1	0.3643	1	0.3643
i_2	0.6357	1	0.6357

Larger Network (4): Propagate Evidence



A	π	λ	P
a_1	0.4	1	0.4
a_2	0.6	1	0.6

B	π	λ	P
b_1	0.26	1	0.26
b_2	0.74	1	0.74

C	π	λ	P
c_1	0.19	1	0.19
c_2	0.81	1	0.81

F	π	λ	P
f_1	0.543	1	0.543
f_2	0.457	1	0.457

D	π	λ	P
d_1	0.389	1	0.389
d_2	0.611	1	0.611

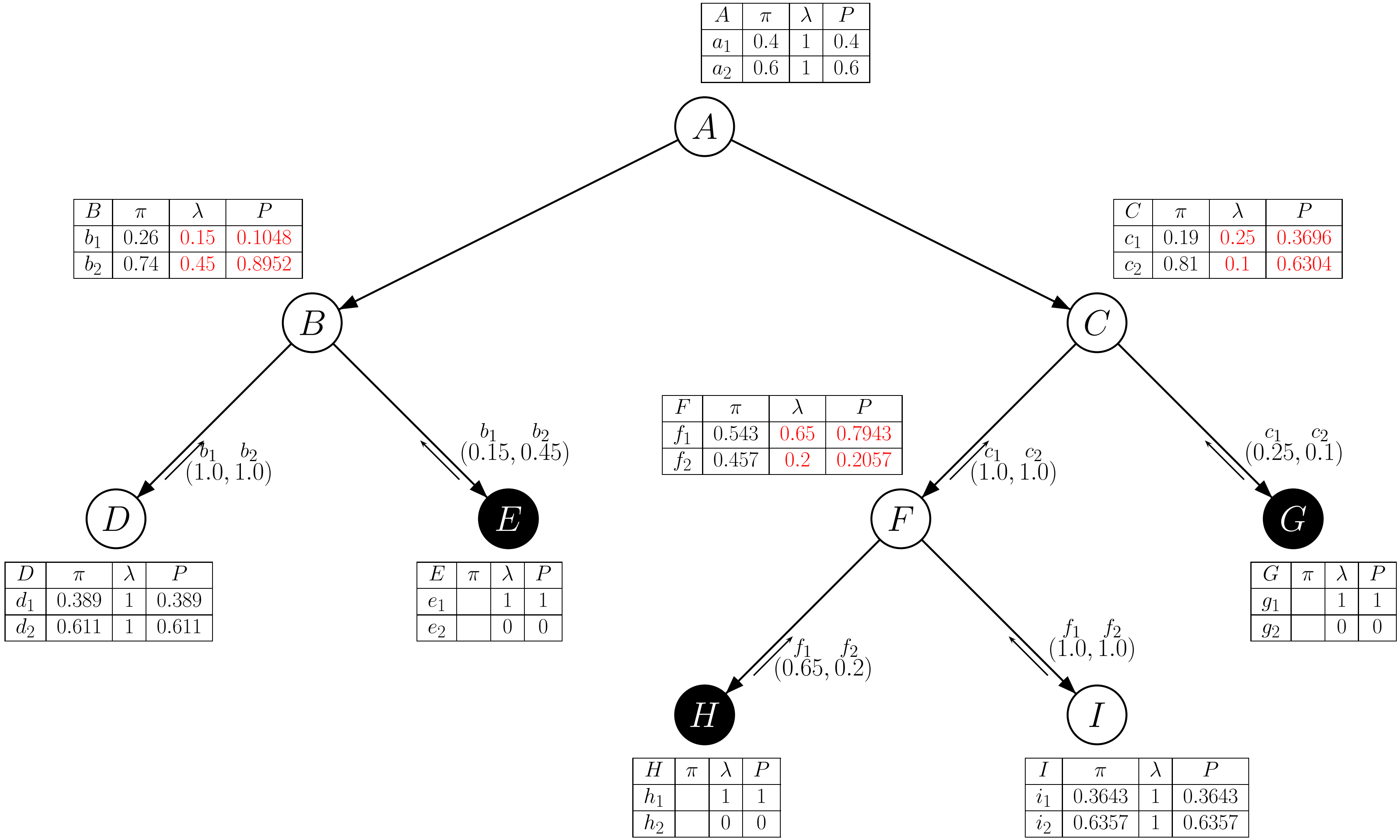
E	π	λ	P
e_1		1	1
e_2		0	0

G	π	λ	P
g_1		1	1
g_2		0	0

H	π	λ	P
h_1		1	1
h_2		0	0

I	π	λ	P
i_1	0.3643	1	0.3643
i_2	0.6357	1	0.6357

Larger Network (5): Propagate Evidence, cont.



A	π	λ	P
a_1	0.4	1	0.4
a_2	0.6	1	0.6

B	π	λ	P
b_1	0.26	0.15	0.1048
b_2	0.74	0.45	0.8952

C	π	λ	P
c_1	0.19	0.25	0.3696
c_2	0.81	0.1	0.6304

F	π	λ	P
f_1	0.543	0.65	0.7943
f_2	0.457	0.2	0.2057

D	π	λ	P
d_1	0.389	1	0.389
d_2	0.611	1	0.611

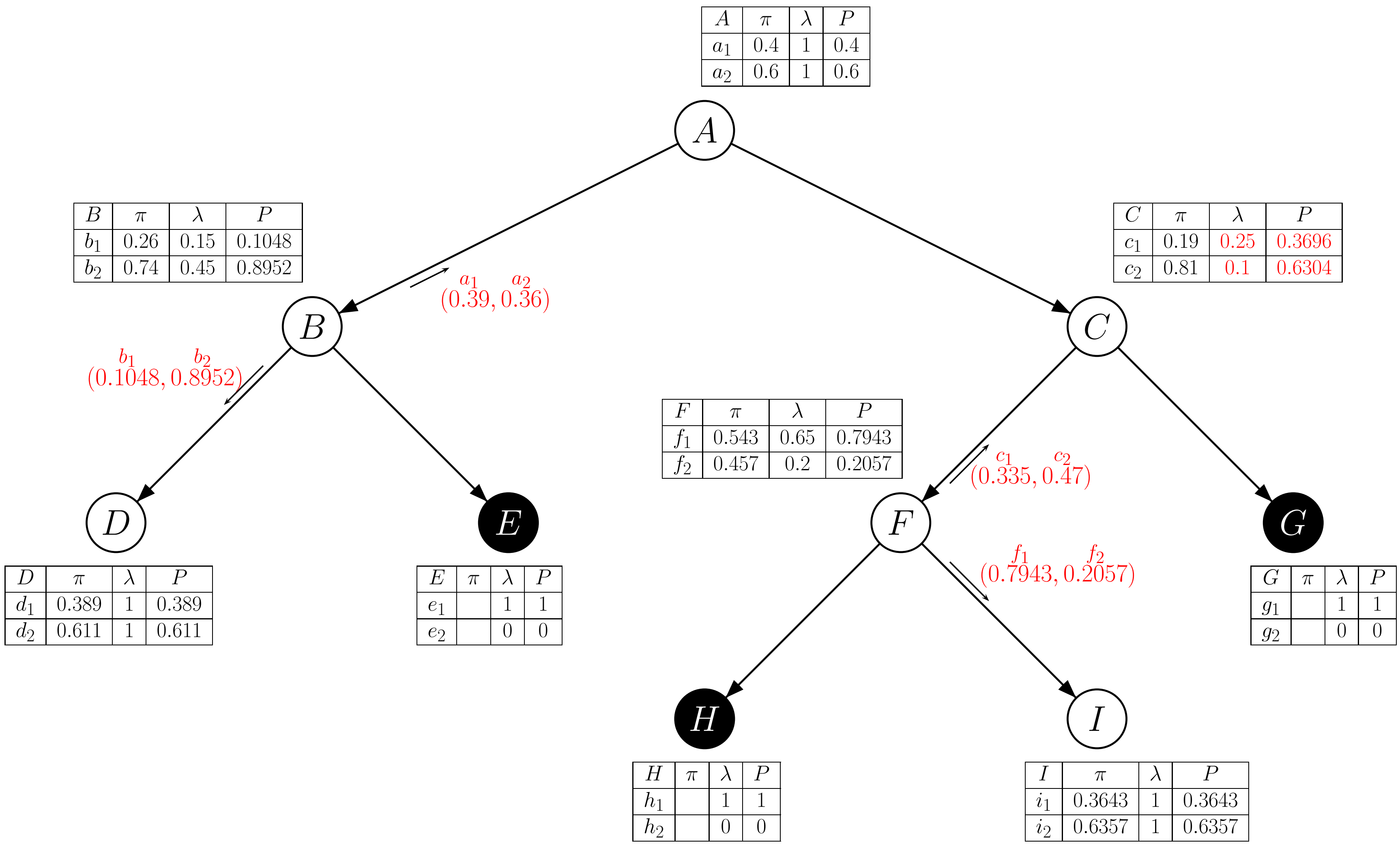
E	π	λ	P
e_1		1	1
e_2		0	0

G	π	λ	P
g_1		1	1
g_2		0	0

H	π	λ	P
h_1		1	1
h_2		0	0

I	π	λ	P
i_1	0.3643	1	0.3643
i_2	0.6357	1	0.6357

Larger Network (6): Propagate Evidence, cont.



A	π	λ	P
a_1	0.4	1	0.4
a_2	0.6	1	0.6

B	π	λ	P
b_1	0.26	0.15	0.1048
b_2	0.74	0.45	0.8952

C	π	λ	P
c_1	0.19	0.25	0.3696
c_2	0.81	0.1	0.6304

F	π	λ	P
f_1	0.543	0.65	0.7943
f_2	0.457	0.2	0.2057

D	π	λ	P
d_1	0.389	1	0.389
d_2	0.611	1	0.611

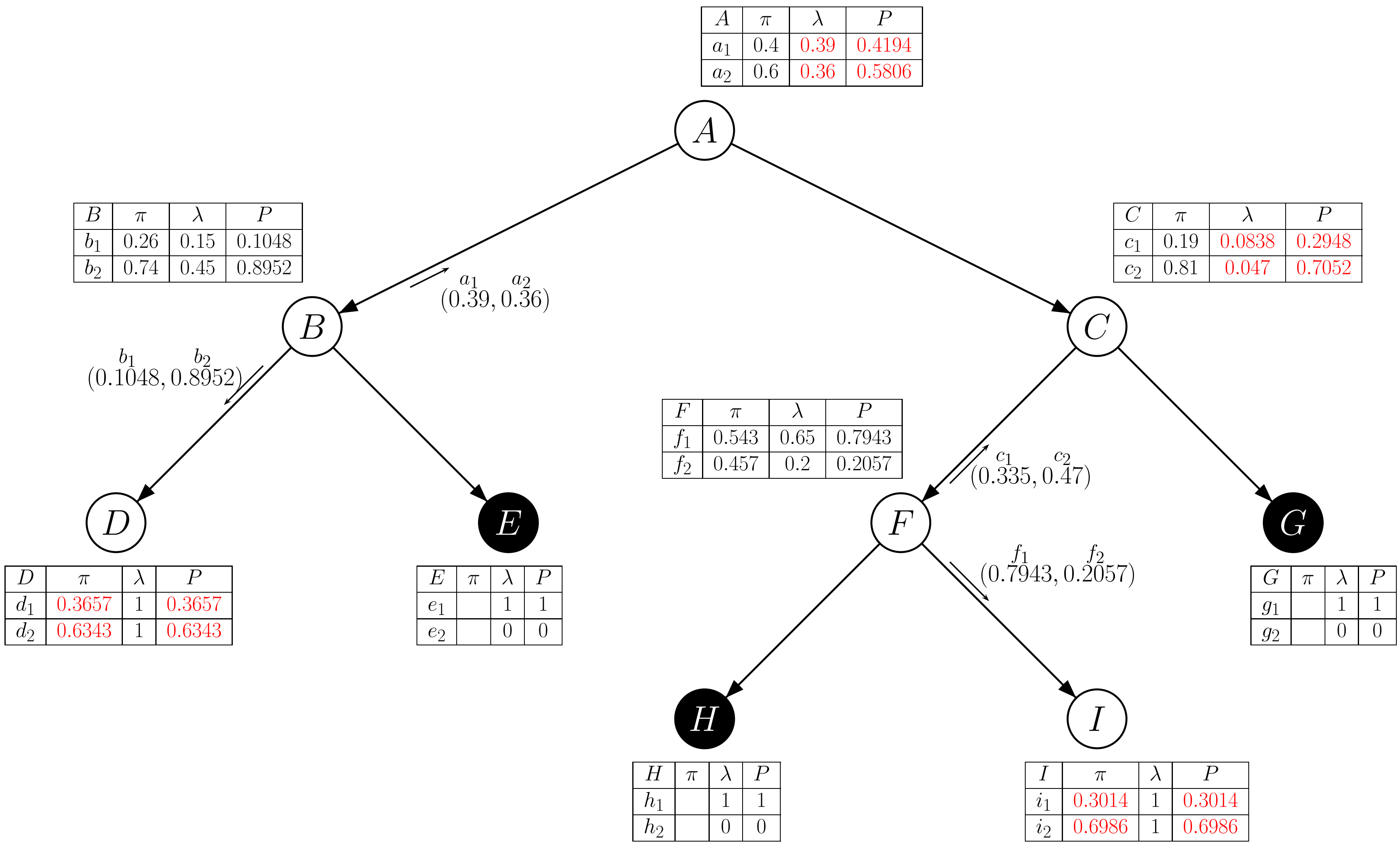
E	π	λ	P
e_1		1	1
e_2		0	0

G	π	λ	P
g_1		1	1
g_2		0	0

H	π	λ	P
h_1		1	1
h_2		0	0

I	π	λ	P
i_1	0.3643	1	0.3643
i_2	0.6357	1	0.6357

Larger Network (7): Propagate Evidence, cont.



A	π	λ	P
a_1	0.4	0.39	0.4194
a_2	0.6	0.36	0.5806

B	π	λ	P
b_1	0.26	0.15	0.1048
b_2	0.74	0.45	0.8952

C	π	λ	P
c_1	0.19	0.0838	0.2948
c_2	0.81	0.047	0.7052

F	π	λ	P
f_1	0.543	0.65	0.7943
f_2	0.457	0.2	0.2057

D	π	λ	P
d_1	0.3657	1	0.3657
d_2	0.6343	1	0.6343

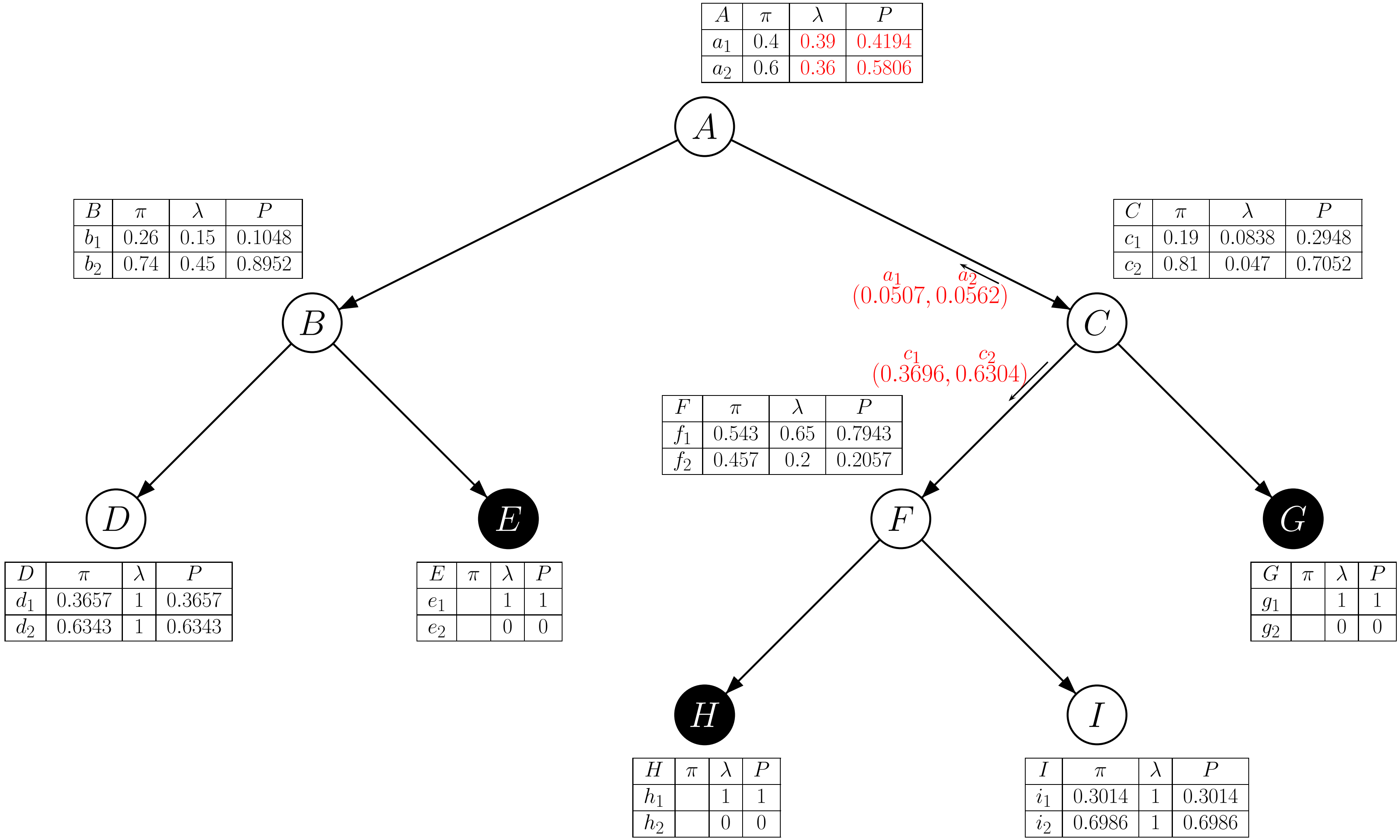
E	π	λ	P
e_1		1	1
e_2		0	0

G	π	λ	P
g_1		1	1
g_2		0	0

H	π	λ	P
h_1		1	1
h_2		0	0

I	π	λ	P
i_1	0.3014	1	0.3014
i_2	0.6986	1	0.6986

Larger Network (8): Propagate Evidence, cont.



A	π	λ	P
a_1	0.4	0.39	0.4194
a_2	0.6	0.36	0.5806

B	π	λ	P
b_1	0.26	0.15	0.1048
b_2	0.74	0.45	0.8952

C	π	λ	P
c_1	0.19	0.0838	0.2948
c_2	0.81	0.047	0.7052

F	π	λ	P
f_1	0.543	0.65	0.7943
f_2	0.457	0.2	0.2057

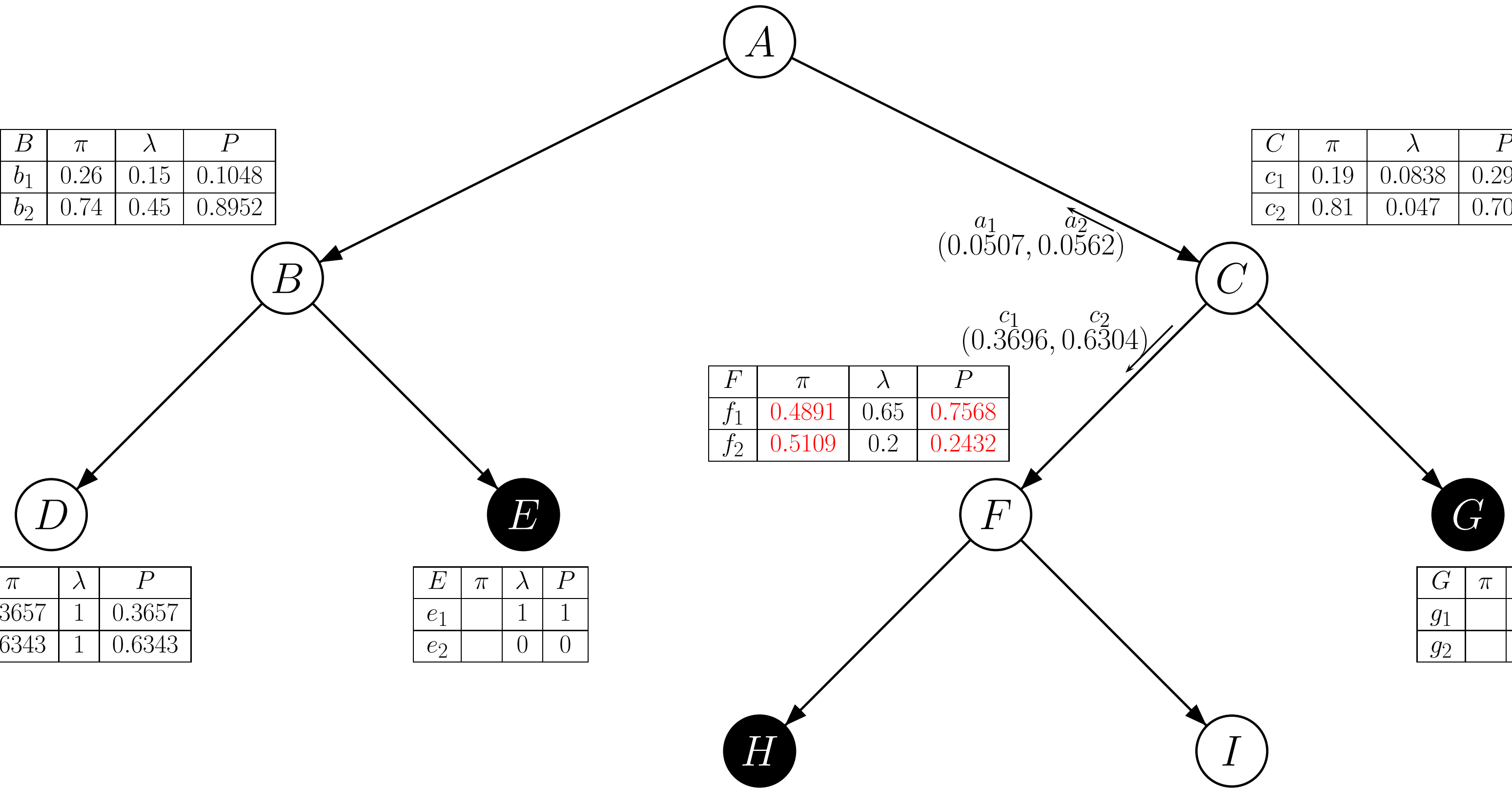
D	π	λ	P
d_1	0.3657	1	0.3657
d_2	0.6343	1	0.6343

E	π	λ	P
e_1		1	1
e_2		0	0

G	π	λ	P
g_1		1	1
g_2		0	0

H	π	λ	P
h_1		1	1
h_2		0	0

I	π	λ	P
i_1	0.3014	1	0.3014
i_2	0.6986	1	0.6986



A	π	λ	P
a_1	0.4	0.0198	0.3945
a_2	0.6	0.0202	0.6055

B	π	λ	P
b_1	0.26	0.15	0.1048
b_2	0.74	0.45	0.8952

C	π	λ	P
c_1	0.19	0.0838	0.2948
c_2	0.81	0.047	0.7052

F	π	λ	P
f_1	0.4891	0.65	0.7568
f_2	0.5109	0.2	0.2432

D	π	λ	P
d_1	0.3657	1	0.3657
d_2	0.6343	1	0.6343

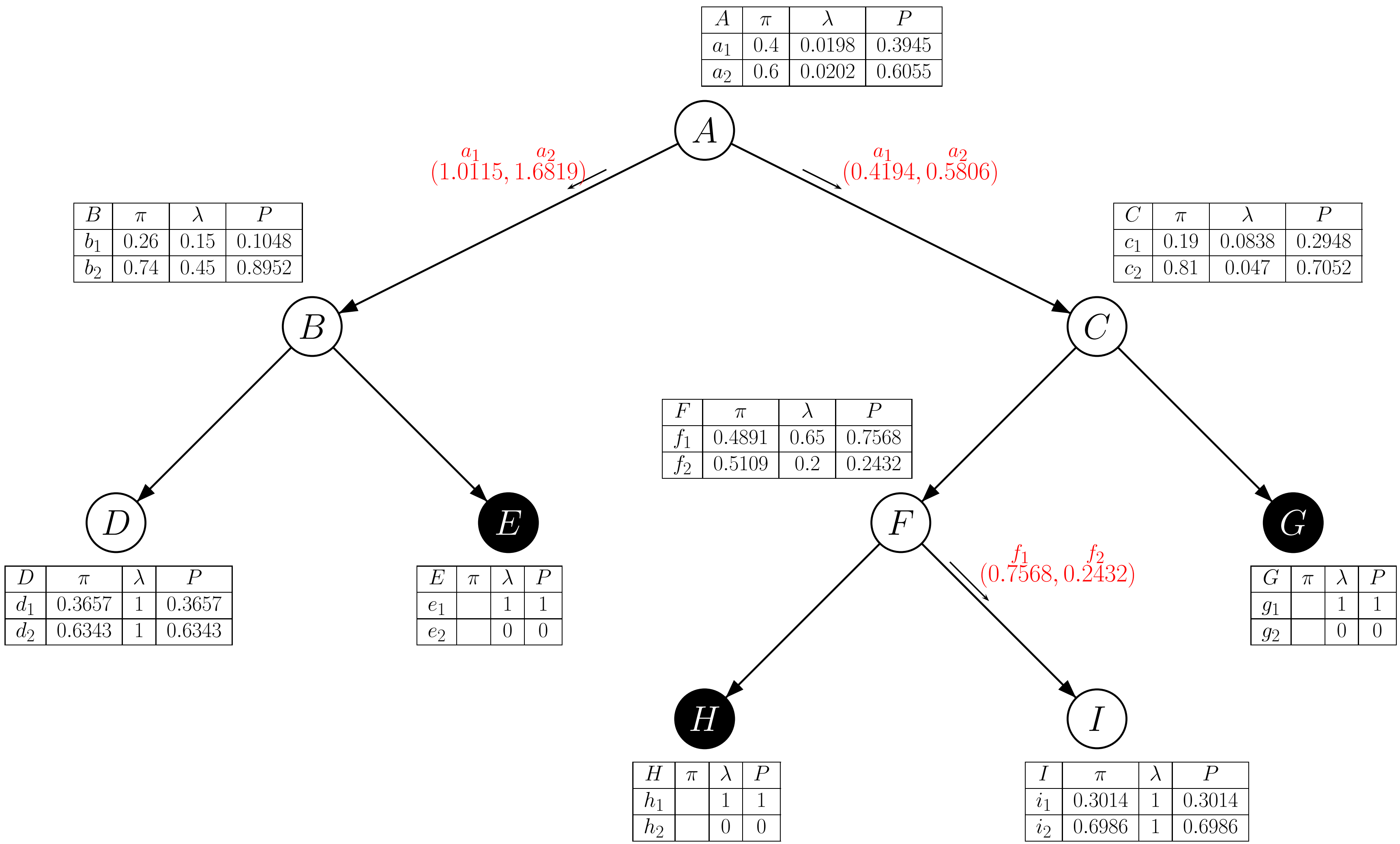
E	π	λ	P
e_1		1	1
e_2		0	0

G	π	λ	P
g_1		1	1
g_2		0	0

H	π	λ	P
h_1		1	1
h_2		0	0

I	π	λ	P
i_1	0.3014	1	0.3014
i_2	0.6986	1	0.6986

Larger Network (10): Propagate Evidence, cont.



A	π	λ	P
a_1	0.4	0.0198	0.3945
a_2	0.6	0.0202	0.6055

B	π	λ	P
b_1	0.26	0.15	0.1048
b_2	0.74	0.45	0.8952

C	π	λ	P
c_1	0.19	0.0838	0.2948
c_2	0.81	0.047	0.7052

F	π	λ	P
f_1	0.4891	0.65	0.7568
f_2	0.5109	0.2	0.2432

D	π	λ	P
d_1	0.3657	1	0.3657
d_2	0.6343	1	0.6343

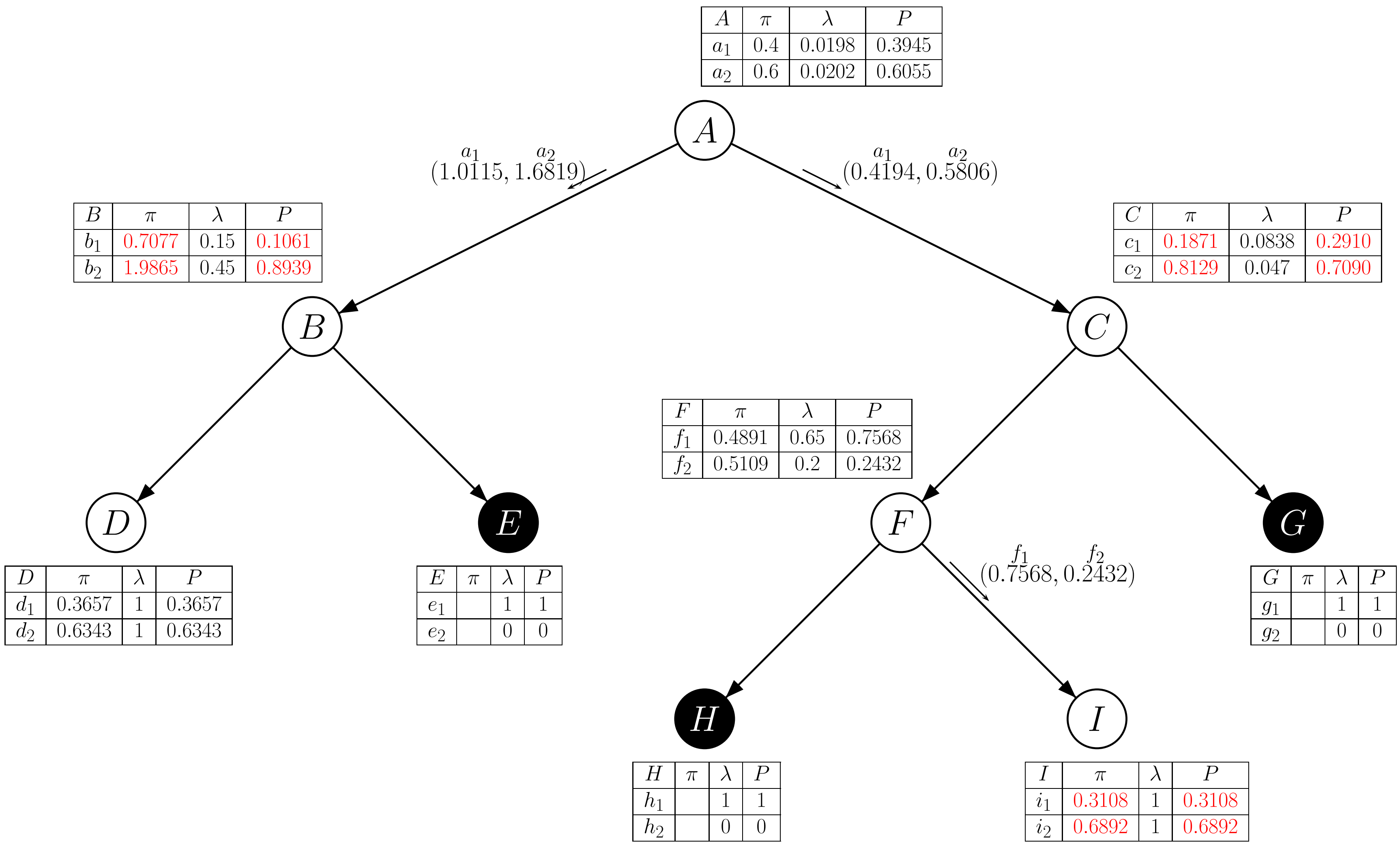
E	π	λ	P
e_1		1	1
e_2		0	0

G	π	λ	P
g_1		1	1
g_2		0	0

H	π	λ	P
h_1		1	1
h_2		0	0

I	π	λ	P
i_1	0.3014	1	0.3014
i_2	0.6986	1	0.6986

Larger Network (11): Propagate Evidence, cont.



A	π	λ	P
a_1	0.4	0.0198	0.3945
a_2	0.6	0.0202	0.6055

B	π	λ	P
b_1	0.7077	0.15	0.1061
b_2	1.9865	0.45	0.8939

C	π	λ	P
c_1	0.1871	0.0838	0.2910
c_2	0.8129	0.047	0.7090

F	π	λ	P
f_1	0.4891	0.65	0.7568
f_2	0.5109	0.2	0.2432

D	π	λ	P
d_1	0.3657	1	0.3657
d_2	0.6343	1	0.6343

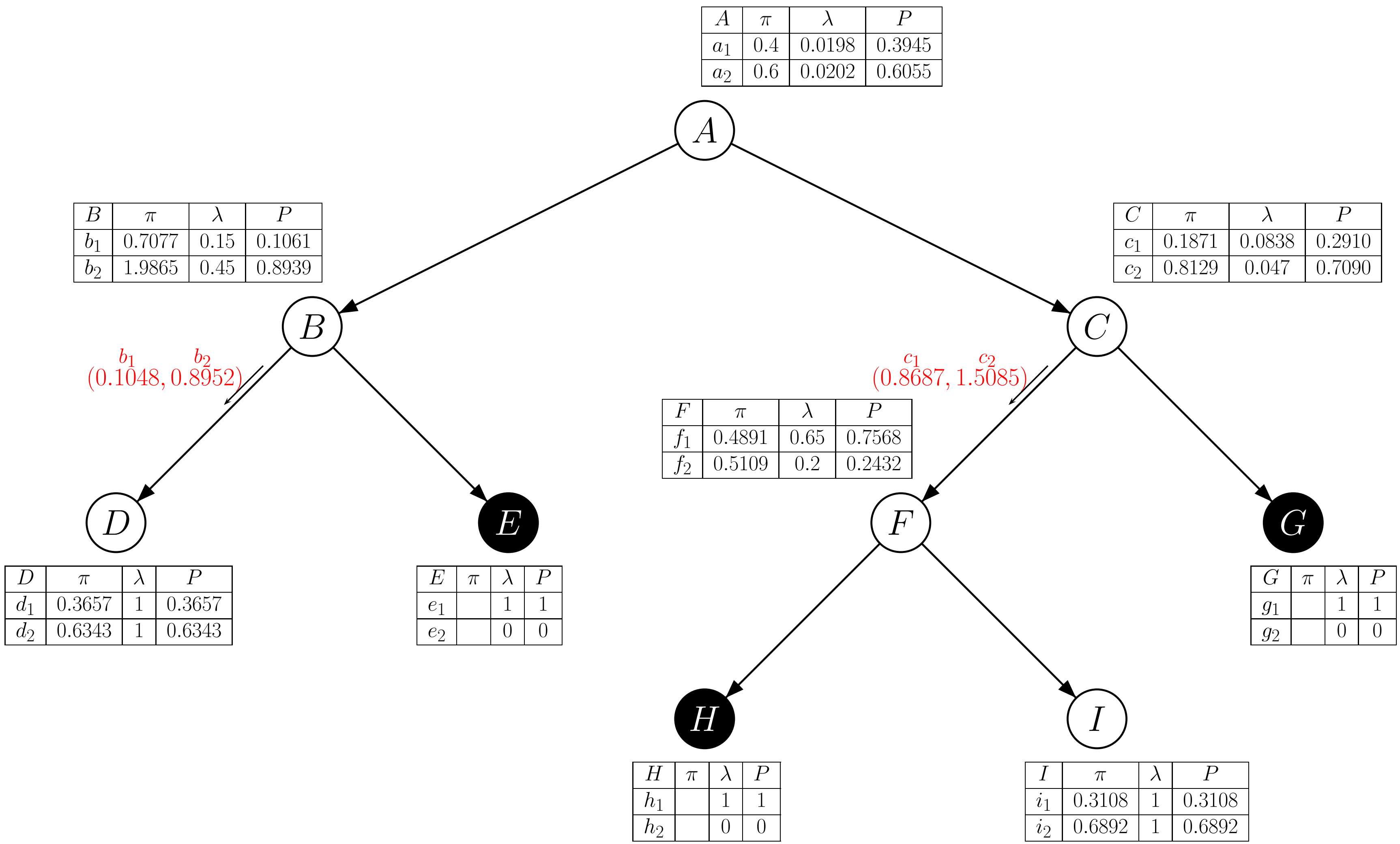
E	π	λ	P
e_1		1	1
e_2		0	0

G	π	λ	P
g_1		1	1
g_2		0	0

H	π	λ	P
h_1		1	1
h_2		0	0

I	π	λ	P
i_1	0.3108	1	0.3108
i_2	0.6892	1	0.6892

Larger Network (12): Propagate Evidence, cont.



A	π	λ	P
a_1	0.4	0.0198	0.3945
a_2	0.6	0.0202	0.6055

B	π	λ	P
b_1	0.7077	0.15	0.1061
b_2	1.9865	0.45	0.8939

C	π	λ	P
c_1	0.1871	0.0838	0.2910
c_2	0.8129	0.047	0.7090

F	π	λ	P
f_1	0.4891	0.65	0.7568
f_2	0.5109	0.2	0.2432

D	π	λ	P
d_1	0.3657	1	0.3657
d_2	0.6343	1	0.6343

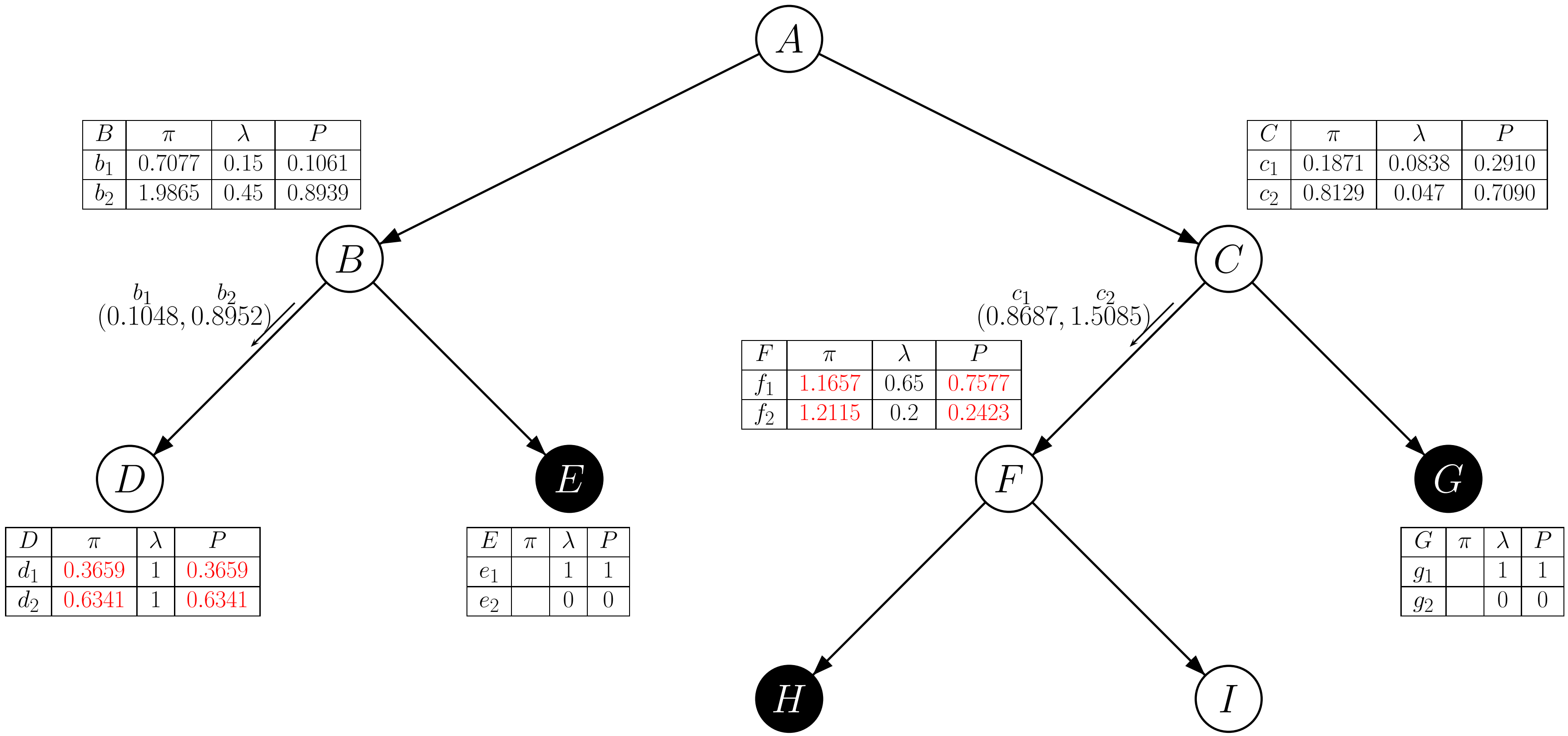
E	π	λ	P
e_1		1	1
e_2		0	0

G	π	λ	P
g_1		1	1
g_2		0	0

H	π	λ	P
h_1		1	1
h_2		0	0

I	π	λ	P
i_1	0.3108	1	0.3108
i_2	0.6892	1	0.6892

Larger Network (13): Propagate Evidence, cont.



A	π	λ	P
a_1	0.4	0.0198	0.3945
a_2	0.6	0.0202	0.6055

B	π	λ	P
b_1	0.7077	0.15	0.1061
b_2	1.9865	0.45	0.8939

C	π	λ	P
c_1	0.1871	0.0838	0.2910
c_2	0.8129	0.047	0.7090

F	π	λ	P
f_1	1.1657	0.65	0.7577
f_2	1.2115	0.2	0.2423

D	π	λ	P
d_1	0.3659	1	0.3659
d_2	0.6341	1	0.6341

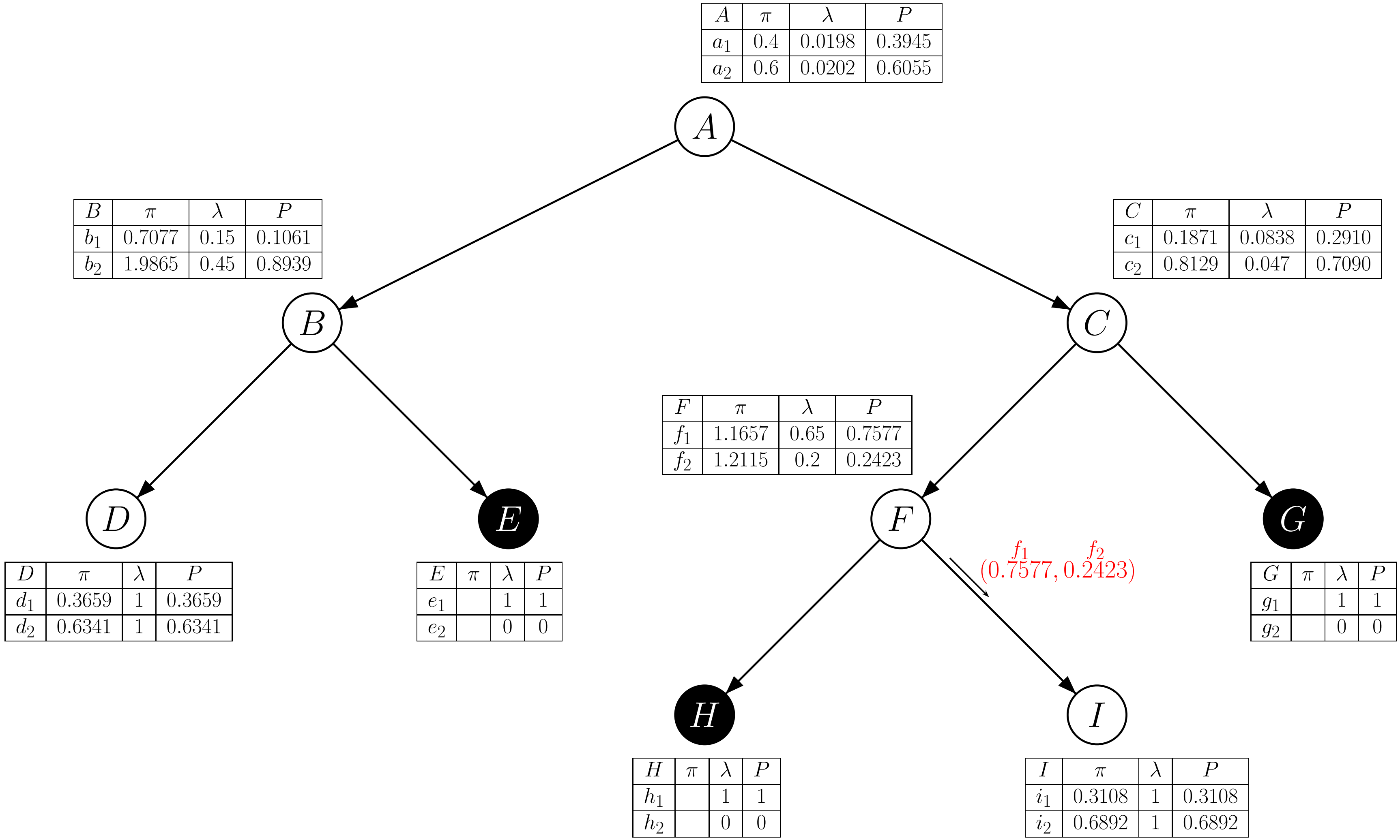
E	π	λ	P
e_1		1	1
e_2		0	0

G	π	λ	P
g_1		1	1
g_2		0	0

H	π	λ	P
h_1		1	1
h_2		0	0

I	π	λ	P
i_1	0.3108	1	0.3108
i_2	0.6892	1	0.6892

Larger Network (14): Propagate Evidence, cont.



A	π	λ	P
a_1	0.4	0.0198	0.3945
a_2	0.6	0.0202	0.6055

B	π	λ	P
b_1	0.7077	0.15	0.1061
b_2	1.9865	0.45	0.8939

C	π	λ	P
c_1	0.1871	0.0838	0.2910
c_2	0.8129	0.047	0.7090

F	π	λ	P
f_1	1.1657	0.65	0.7577
f_2	1.2115	0.2	0.2423

D	π	λ	P
d_1	0.3659	1	0.3659
d_2	0.6341	1	0.6341

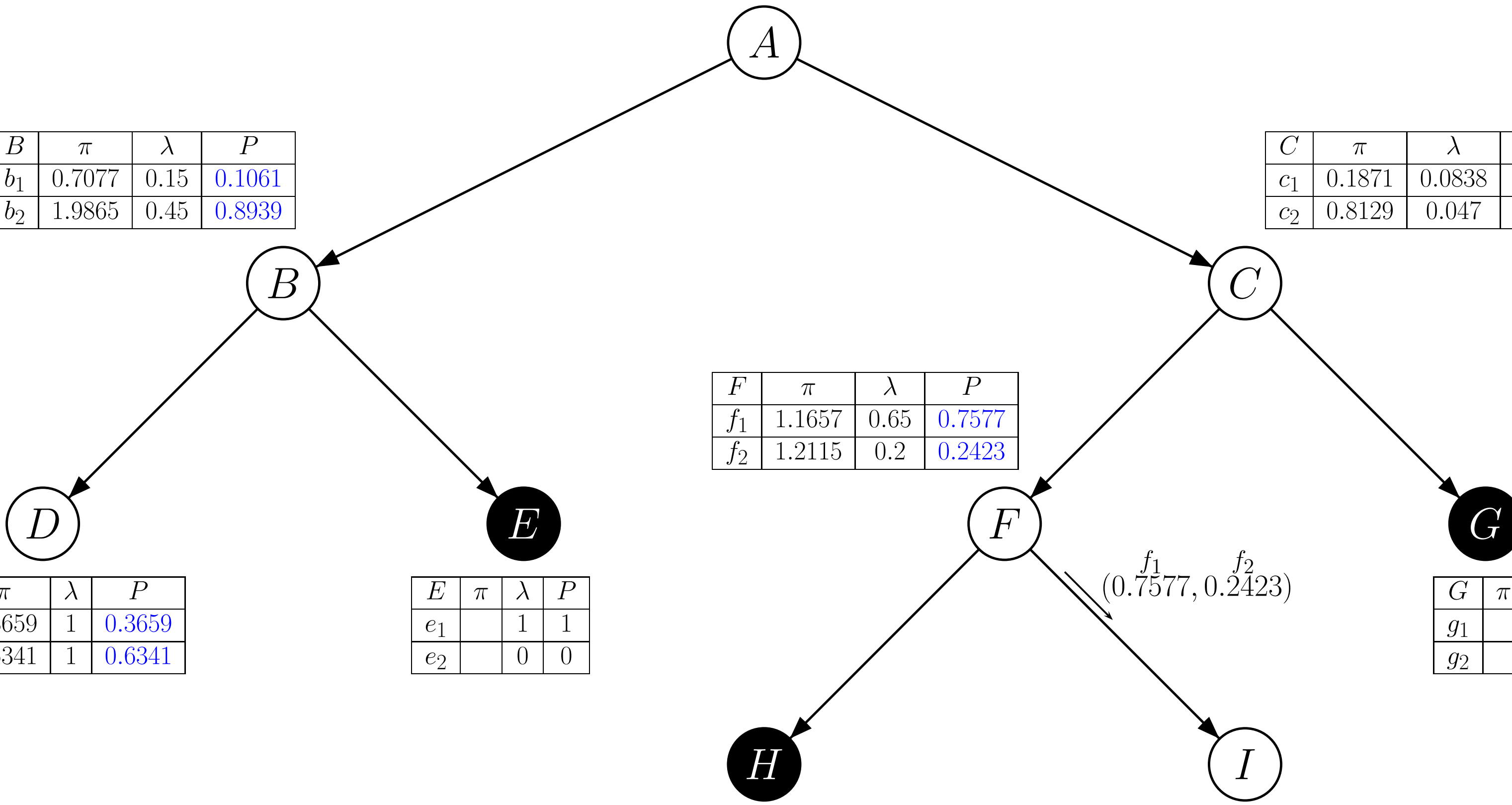
E	π	λ	P
e_1		1	1
e_2		0	0

G	π	λ	P
g_1		1	1
g_2		0	0

H	π	λ	P
h_1		1	1
h_2		0	0

I	π	λ	P
i_1	0.3108	1	0.3108
i_2	0.6892	1	0.6892

Larger Network (15): Finished



A	π	λ	P
a_1	0.4	0.0198	0.3945
a_2	0.6	0.0202	0.6055

B	π	λ	P
b_1	0.7077	0.15	0.1061
b_2	1.9865	0.45	0.8939

C	π	λ	P
c_1	0.1871	0.0838	0.2910
c_2	0.8129	0.047	0.7090

F	π	λ	P
f_1	1.1657	0.65	0.7577
f_2	1.2115	0.2	0.2423

D	π	λ	P
d_1	0.3659	1	0.3659
d_2	0.6341	1	0.6341

E	π	λ	P
e_1		1	1
e_2		0	0

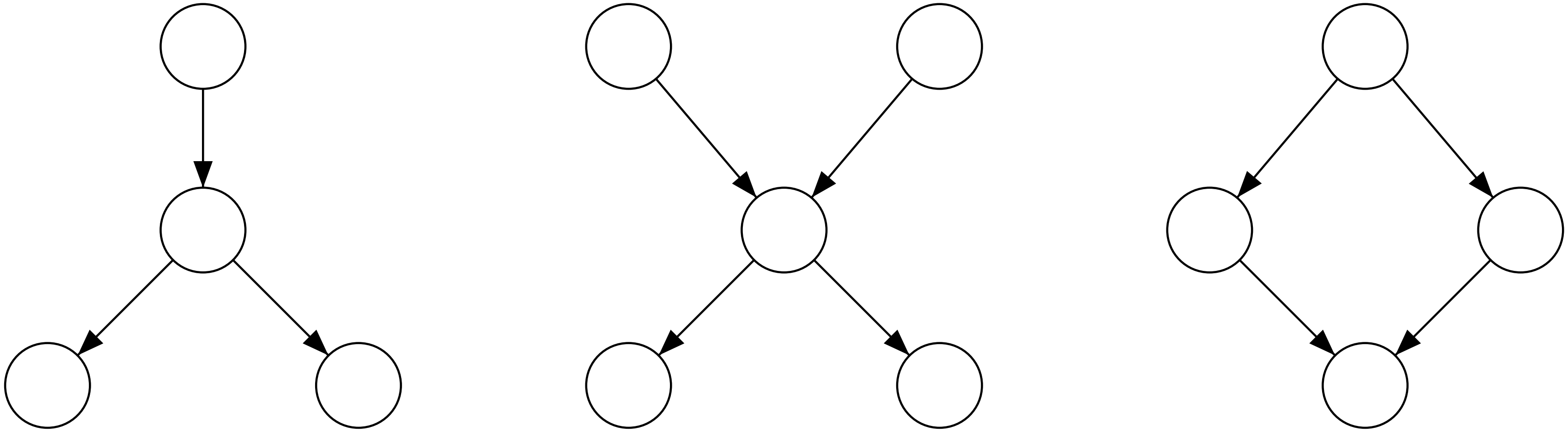
G	π	λ	P
g_1		1	1
g_2		0	0

H	π	λ	P
h_1		1	1
h_2		0	0

I	π	λ	P
i_1	0.3106	1	0.3106
i_2	0.6894	1	0.6894

Propagation in Clique Trees

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!

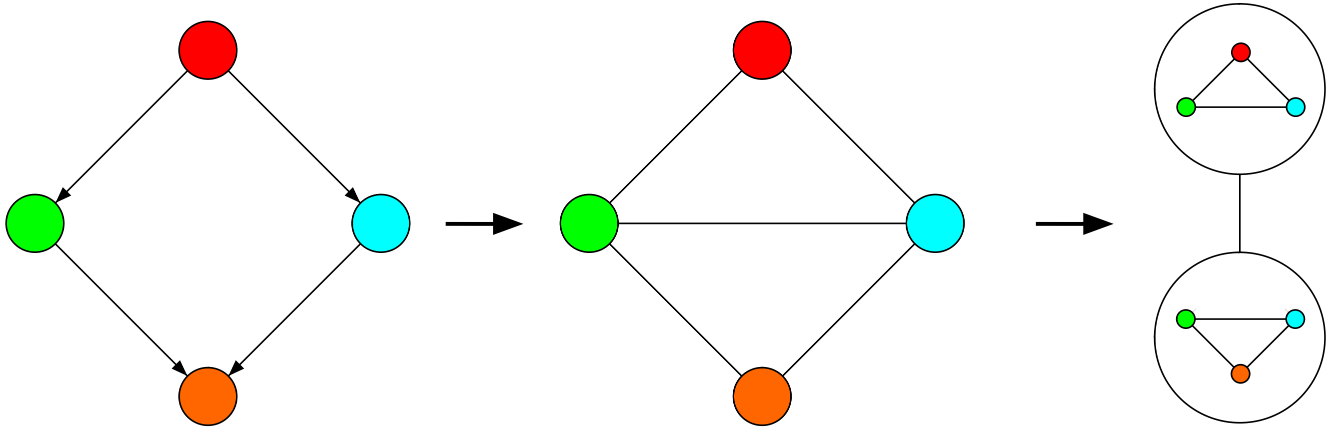
Idea

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.

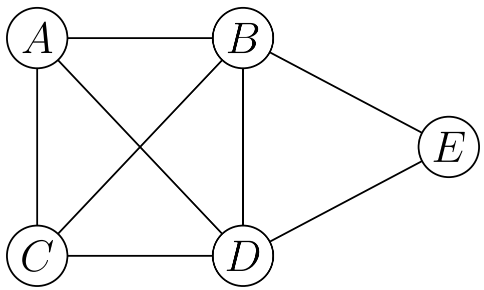
Complete Graph

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

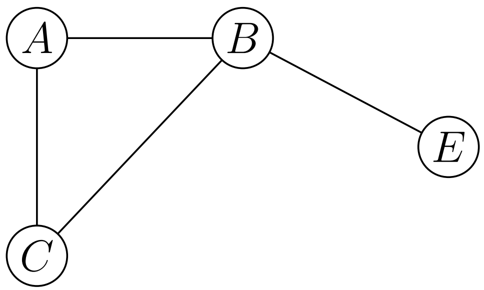
Induced Subgraph

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

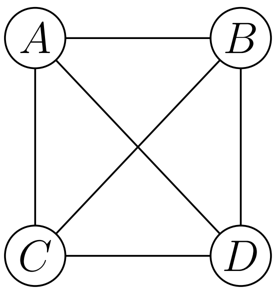
$$E_W = \{(u, v) \in E \mid u, v \in W\}.$$



Incomplete graph



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



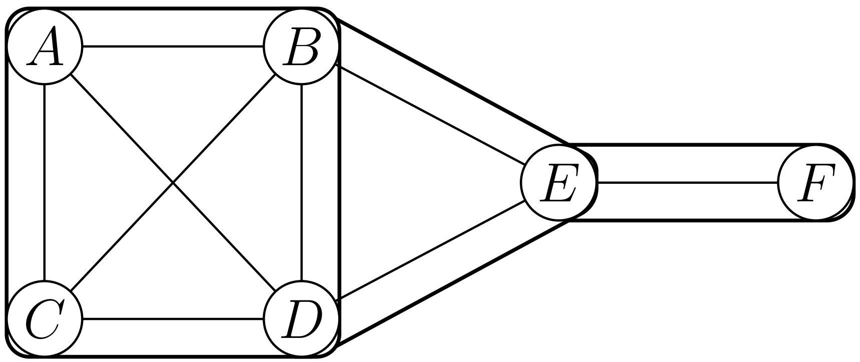
Complete (sub)graph

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

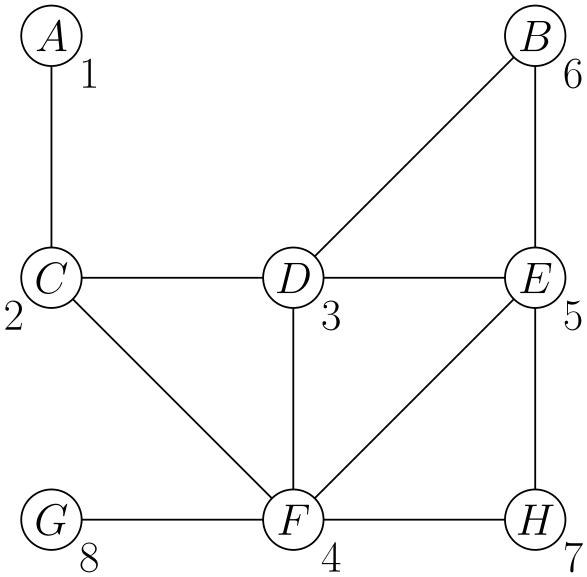
$$C_1 = \{A, B, C, D\}$$
$$C_2 = \{B, D, E\}$$
$$C_3 = \{E, F\}$$

Perfect Ordering

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

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

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



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

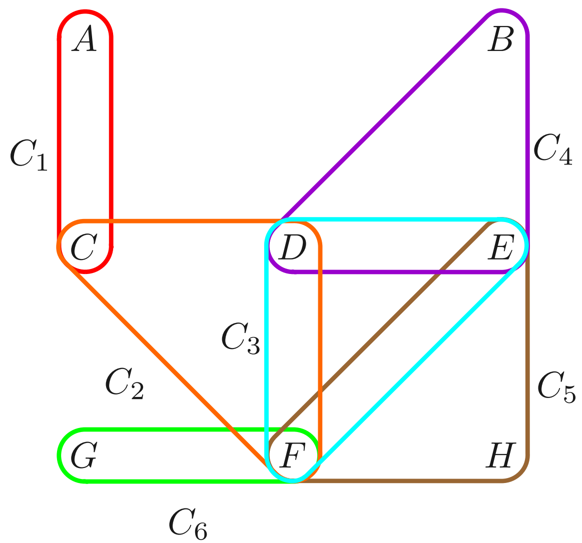
i	$\operatorname{adj}(v_i)$	$\operatorname{adj}(v_i) \cap \{v_1, \ldots, v_{i-1}\}$		
1	$\{C\}$	$\{C\} \cap \emptyset$	$= \emptyset$	complete
2	$\{A, D, F\}$	$\{A\} \cap \{A, D, F\}$	$= \{A\}$	complete
3	$\{C, B, E, F\}$	$\{A, C\} \cap \{C, B, E, F\}$	$= \{C\}$	complete
4	$\{G, C, D, E, H\}$	$\{A, C, D\} \cap \{G, C, D, E, H\}$	$= \{C, D\}$	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

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:

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

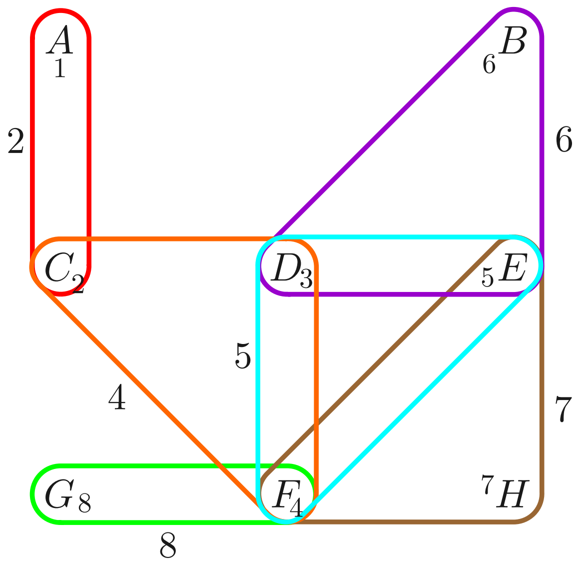


$\xi = \langle C_1,\ C_2,\ C_3,\ C_4,\ C_5,\ C_6 \rangle$

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

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.

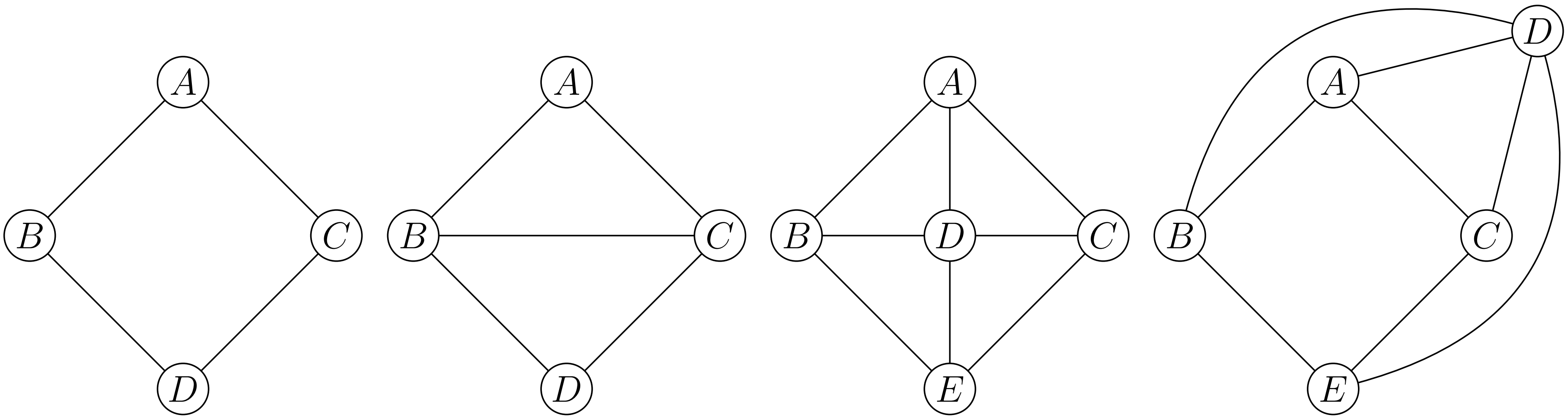


Clique	Rank		
$\{A, C\}$	$\max\{\alpha(A), \alpha(C)\}$	$= 2$	$\rightarrow C_1$
$\{C, D, F\}$	$\max\{\alpha(C), \alpha(D), \alpha(F)\}$	$= 4$	$\rightarrow C_2$
$\{D, E, F\}$	$\max\{\alpha(D), \alpha(E), \alpha(F)\}$	$= 5$	$\rightarrow C_3$
$\{B, D, E\}$	$\max\{\alpha(B), \alpha(D), \alpha(E)\}$	$= 6$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

How to get a perfect ordering?

Triangulated Graph

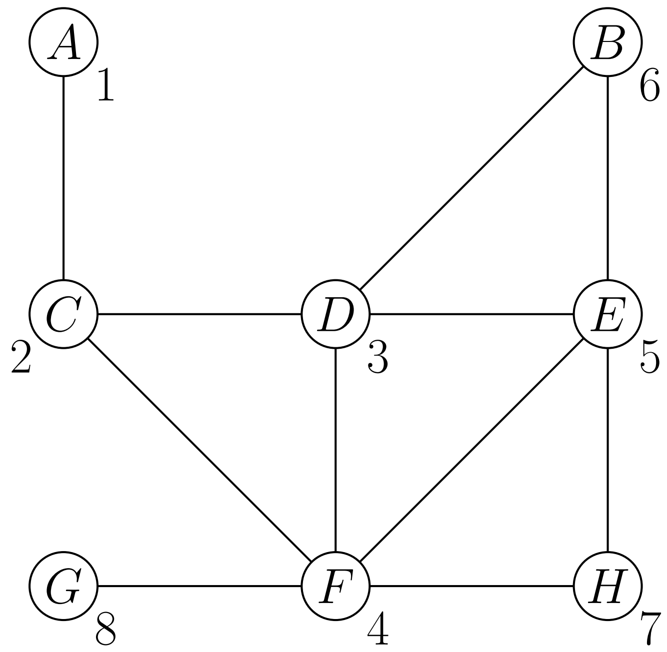
An undirected graph is called *triangulated* if every simple loop (i. e. path with identical start and end node but with any other node occurring at most once) of length greater 3 has a chord.



not triangulated | triangulated | not triangulated | no chord for $\langle A, B, E, C\rangle$

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

Triangulated Graphs (3)

An undirected graph is triangulated iff the ordering obtained by MCS is perfect.

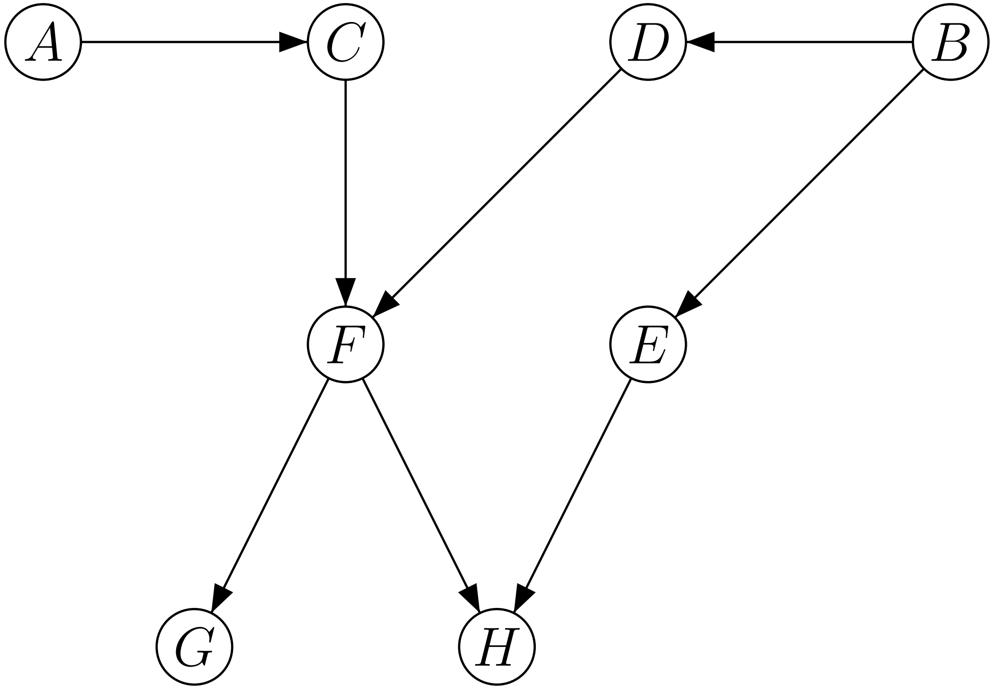
To check whether a graph is triangulated is efficient to implement. The optimization problem that is related to the triangulation task is NP-hard. However, there are good heuristics.

Moral Graph (Repetition)

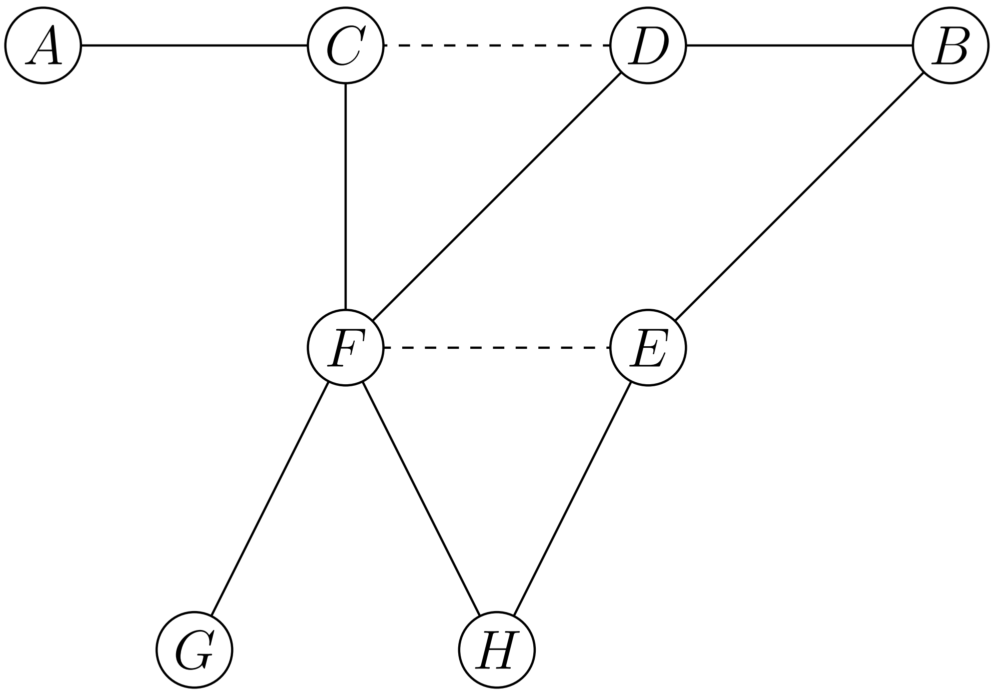
Let $G = (V, E)$ be a directed acyclic graph. If $u, w \in W$ are parents of $v \in V$ connect u and w with an (arbitrarily oriented) edge. After the removal of all edge directions the resulting graph $G_m = (V, E')$ is called the *moral graph* of G.

Join-Tree Construction (1)

Given directed graph.

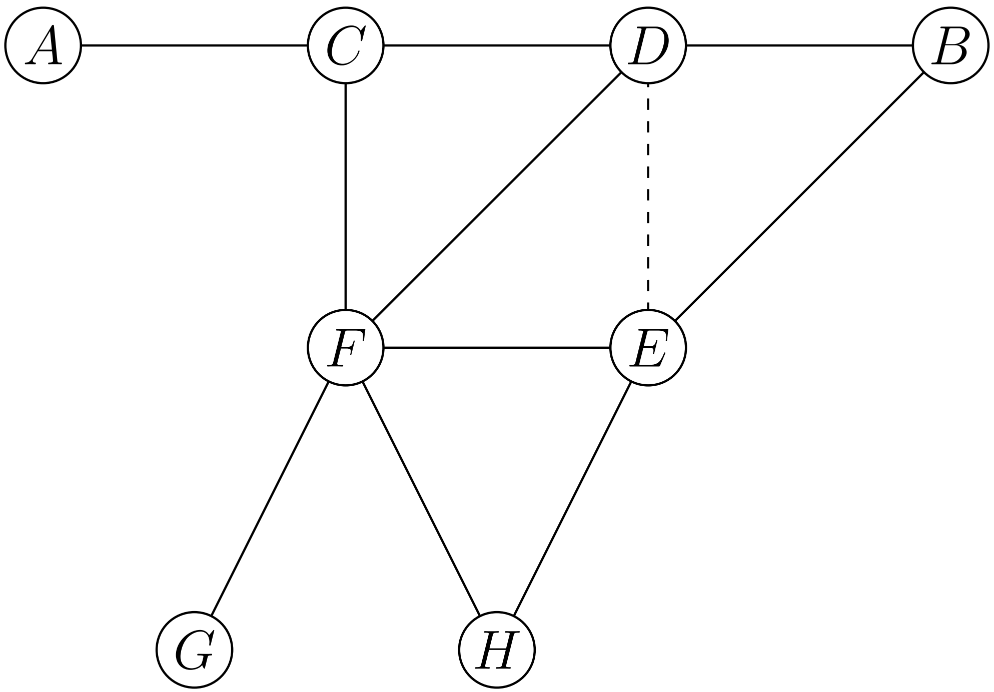


Join-Tree Construction (2)



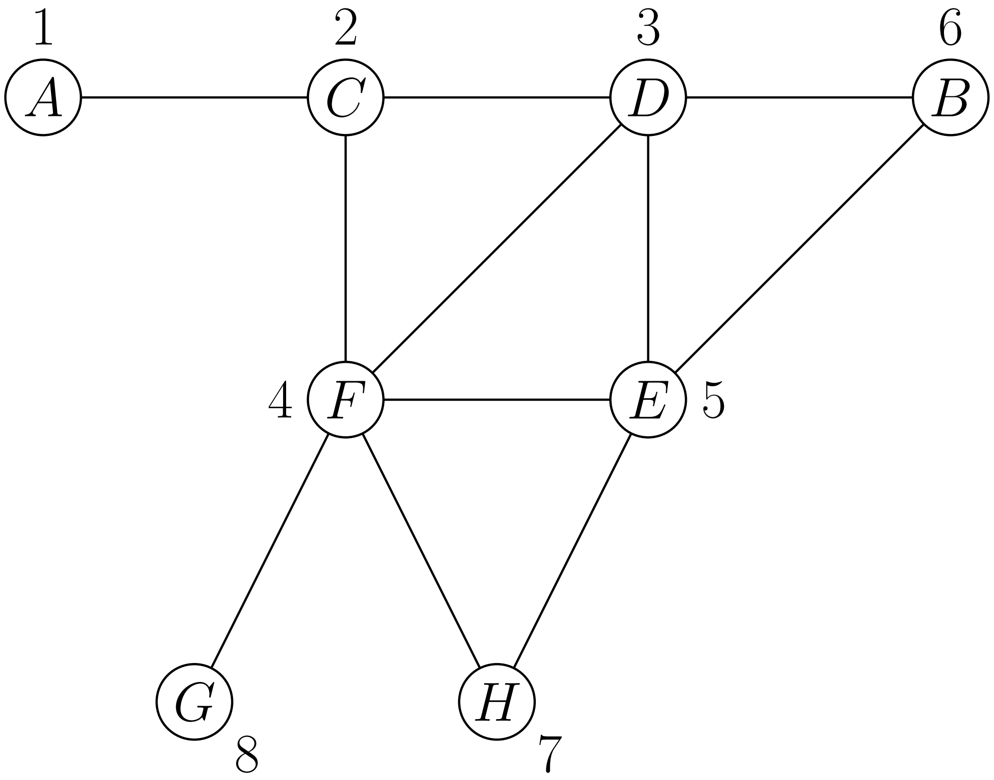
- Moral graph

Join-Tree Construction (3)



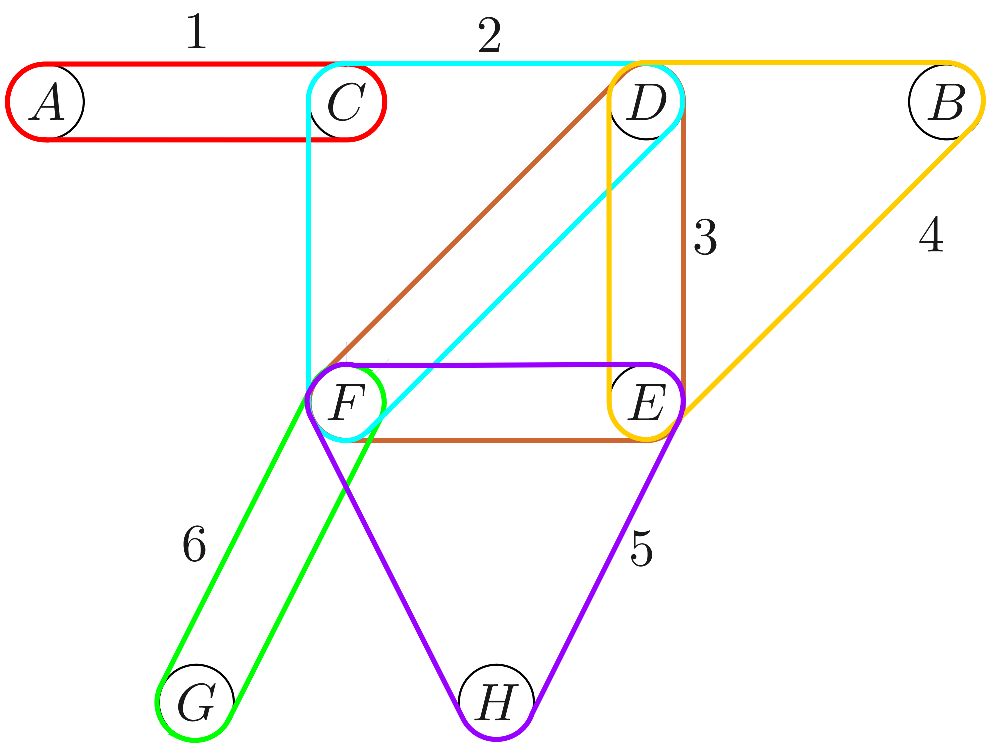
- Moral graph
- Triangulated graph

Join-Tree Construction (4)



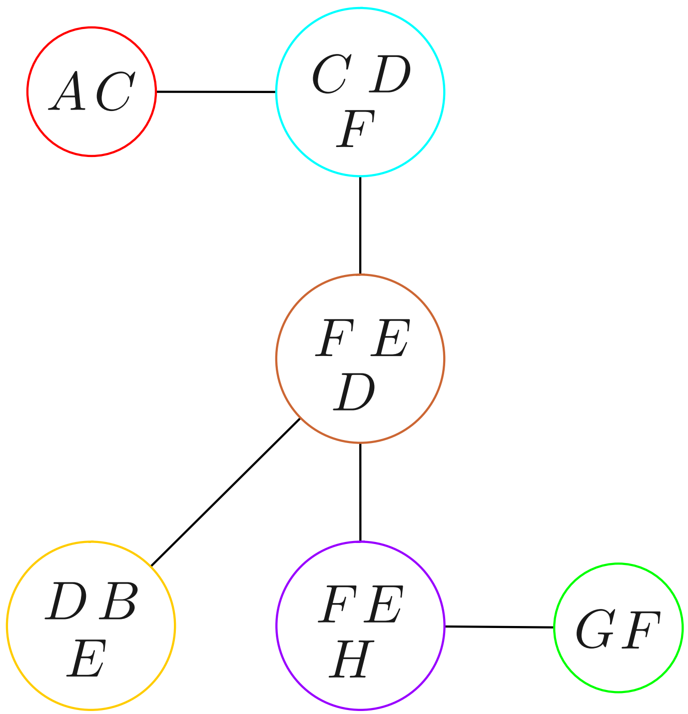
- Moral graph
- Triangulated graph
- MCS yields perfect ordering

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	$\mathrm{dom}(A) = \{a_1, a_2\}$	$\cdot_1 =$ existing
B	increased total serum calcium	$\mathrm{dom}(B) = \{b_1, b_2\}$	$\cdot_2 =$ notexisting
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 have to be specified!

Example: Qualitative Knowledge

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

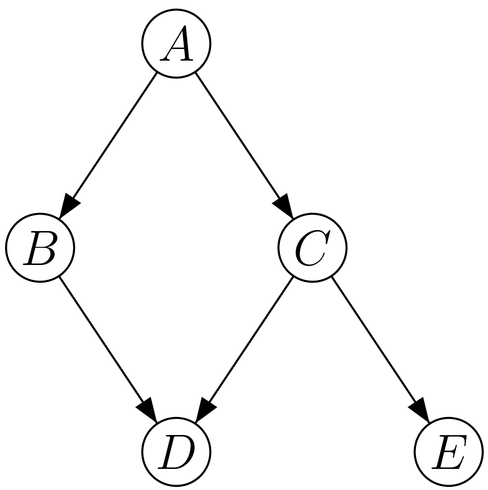
$$\left.\begin{aligned} P(d_1 \mid b_1, c_1) &= 0.8 \\ P(d_1 \mid b_1, c_2) &= 0.8 \\ P(d_1 \mid b_2, c_1) &= 0.8 \\ P(d_1 \mid b_2, c_2) &= 0.05 \end{aligned}\right\} \text{coma rare but common, if either cause is present}$$

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

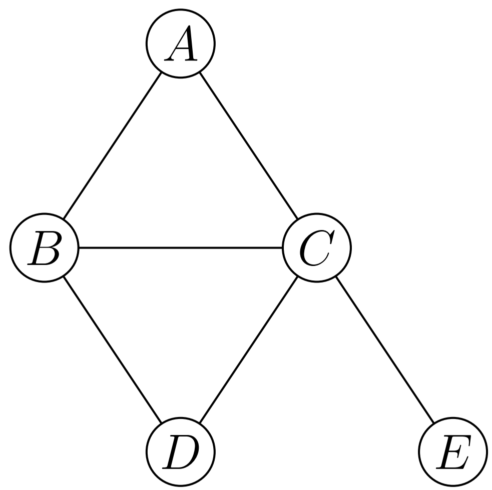
$$\left.\begin{aligned} P(c_1 \mid a_1) &= 0.2 \\ P(c_1 \mid a_2) &= 0.05 \end{aligned}\right\} \text{brain tumor rare, and uncommon consequence of metastases}$$

$$\left.P(a_1) = 0.2\right\} \text{incidence of metastatic cancer in relevant clinic}$$

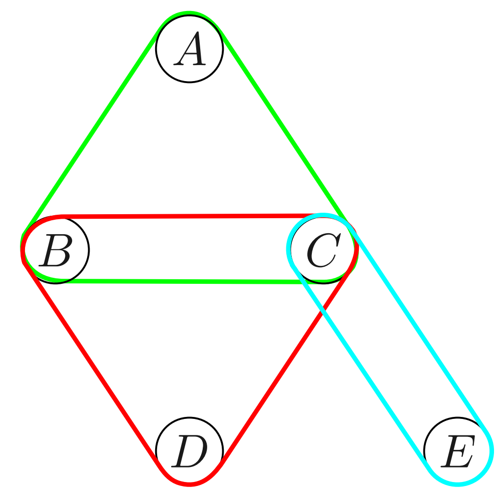
Example: Metastatic Cancer



Dependencies



Moralization/Triangulation



MCS, hyper graph

Clique tree with separator sets

Propagation on Cliques (3)

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
$\vdots$	$\vdots$
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
$\vdots$	$\vdots$
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

Potential representation:

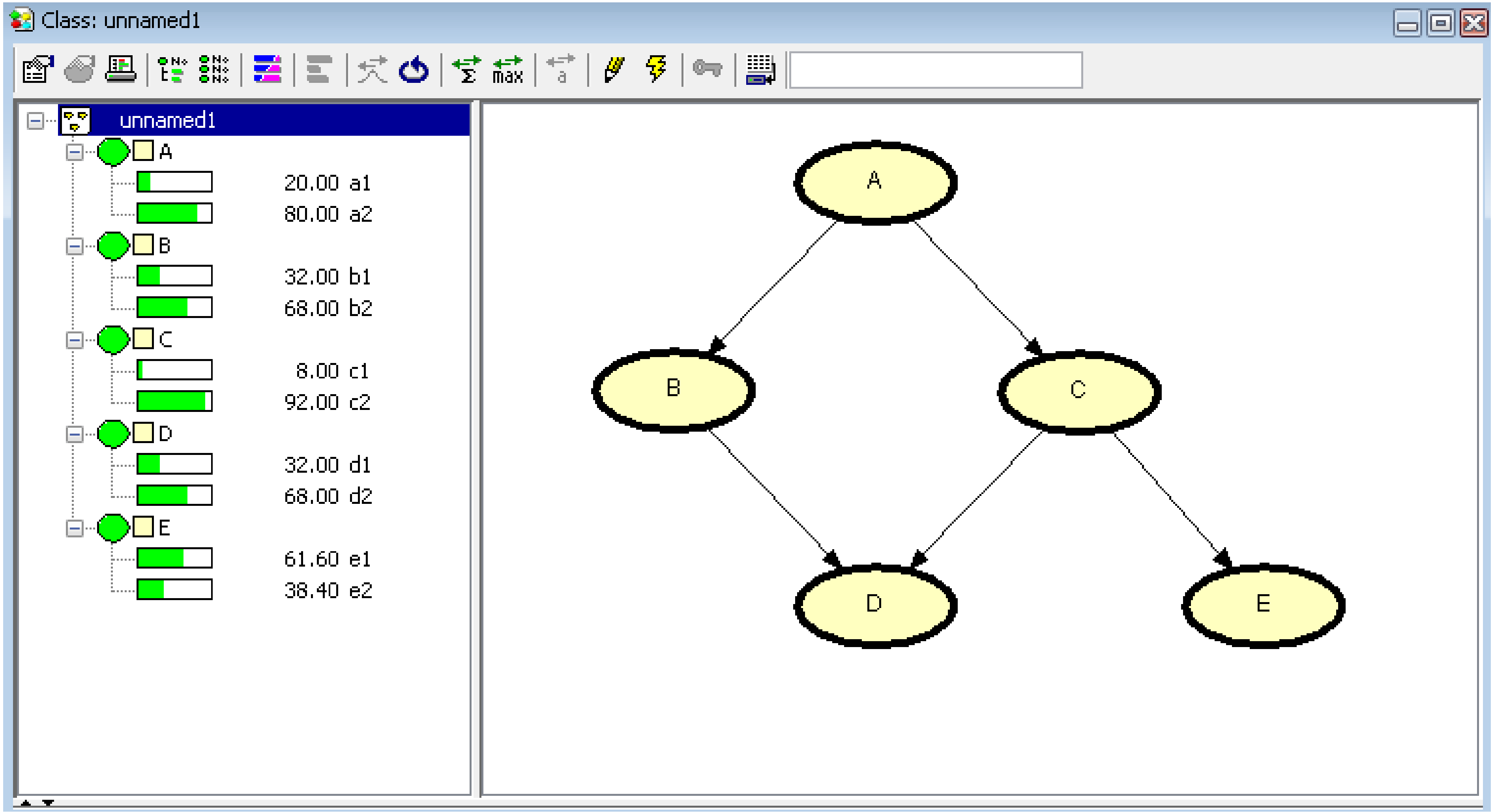
$$\begin{aligned} P(A,B,C,D,E,) &= P(A \mid \emptyset)P(B \mid A)P(C \mid A)P(D \mid BC)P(E \mid C) \\ &= \frac{P(A,B,C)P(B,C,D),P(C,E)}{P(BC)P(C)} \end{aligned}$$

Propagation on Cliques (4)

Propagation:

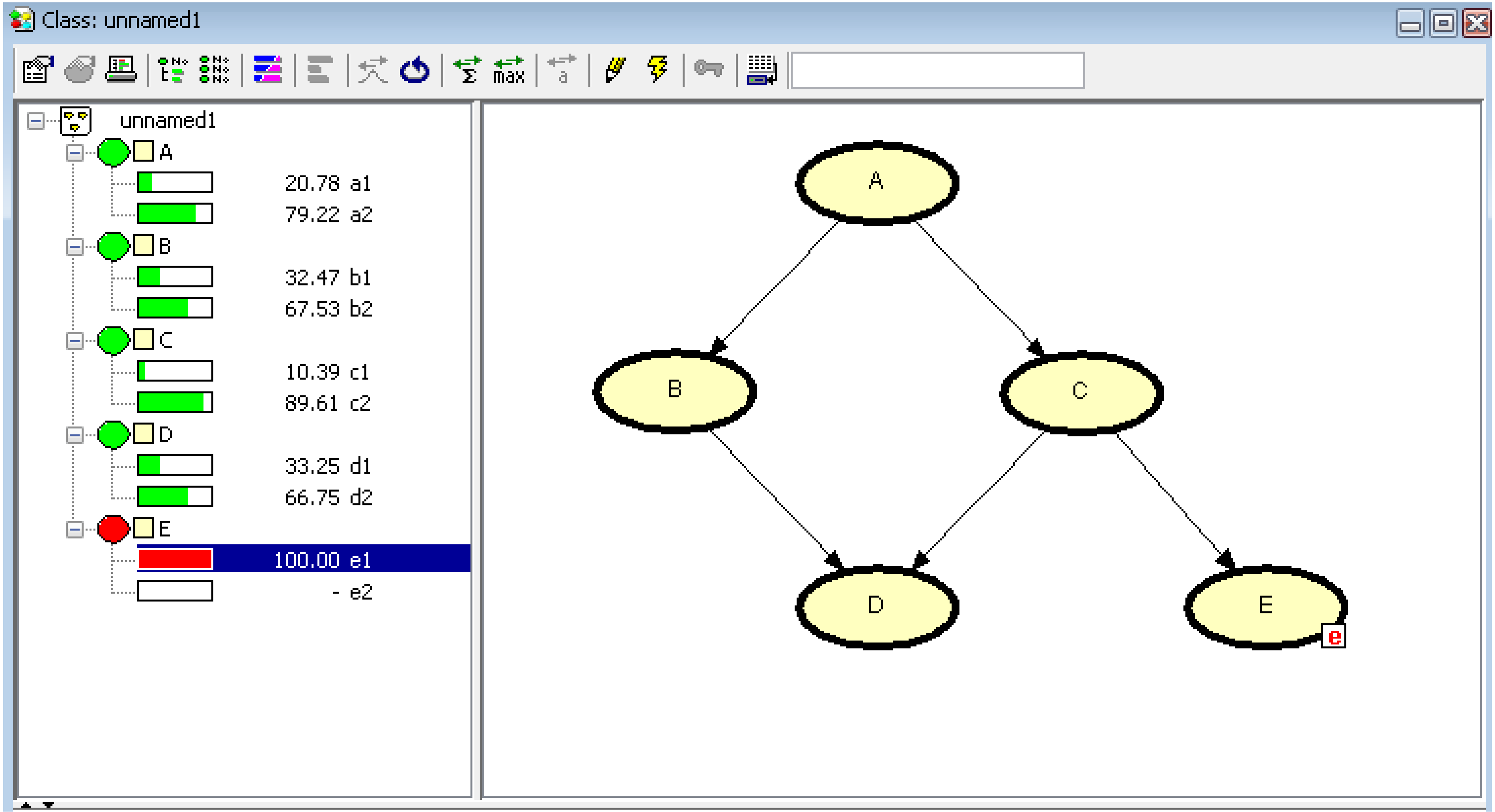
- $P(d_1) = 0.32$, evidence $E = e_1$, desired: $P^*(\ldots) = P(\cdot \mid \{e_1\})$

$$
\begin{array}{lll}
P^*(c) & = P(c \mid e_1) & \text{conditional marginal distribution} \\[2ex]
P^*(b, c, d) & = \dfrac{P(b, c, d)}{P(c)} P(c \mid e_1) & \text{multipl./division with separation prob.} \\[2ex]
P(b, c, d),\ P^*(b, c) & & \text{calculate marginal distributions} \\[2ex]
P^*(a, b, c) & = \dfrac{P(a, b, c)}{P(b, c)} P(b, c \mid e_1) & \text{multipl./division with separation prob.} \\[2ex]
P^*(d_1) & = P(d_1 \mid e_1) = 0.33 &
\end{array}
$$



Marginal distributions in the HUGIN tool.

Propagation on Cliques (6)



Conditional marginal distributions with evidence $E = e_1$

Potential Representation

Let $V = \{X_j\}$ be a set of random variables $X_j : \Omega \to \mathrm{dom}(X_j)$ 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

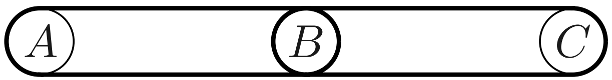
$$\psi_i : \bigtimes_{X_j \in W_i} \mathrm{dom}(X_j) \to \mathbb{R}$$

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.

Example



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

$$\begin{aligned}\mathrm{dom}(A) &= \{a_1, a_2\}\\ \mathrm{dom}(B) &= \{b_1, b_2\}\\ \mathrm{dom}(C) &= \{c_1, c_2\}\end{aligned}$$

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

$$\begin{aligned}\psi_1 &: \{a_1, a_2\} \times \{b_1, b_2\} \to \mathbb{R}\\ \psi_2 &: \{b_1, b_2\} \times \{c_1, c_2\} \to \mathbb{R}\end{aligned}$$

$$\begin{aligned}\psi_1(a, b) &= \tfrac{1}{4}\\ \psi_2(b, c) &= \tfrac{1}{2}\end{aligned}$$

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

Factorization of a Belief Network

Let (V, E, P) be an 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 the factor potentials are:

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

Separator Sets and Residual Sets

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

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

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

$$R_i = C_i \backslash S_i$$

Decomposition w. r. t. a Join-Tree

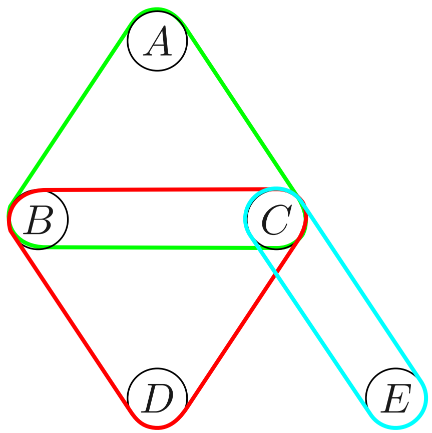
- Given a clique ordering $\{C_1, \ldots, C_p\}$ that satisfies the RIP, we can easily conclude the following separation statements:

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

- Hence, we can formulate the following factorization:

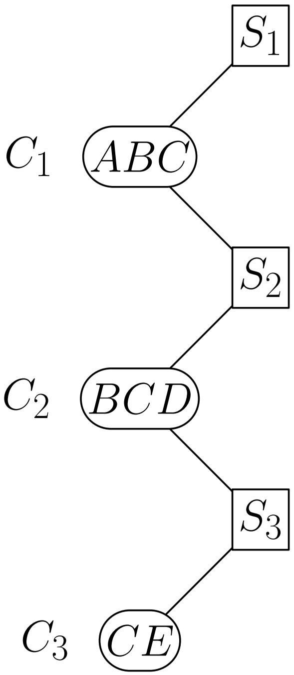
$$P(X_1, \ldots, X_n) = \prod_{i=1}^{p} P(R_i \mid S_i),$$

which also gives us a representation in terms of conditional probabilities (as for directed graphs before).



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

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

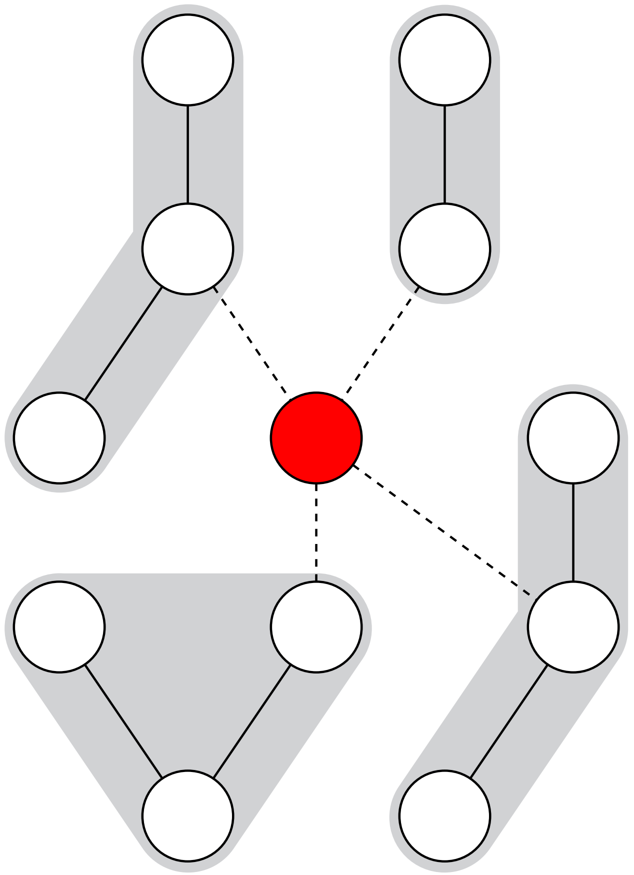


$$\psi_1(C_1) = P(A, B, C \mid \emptyset) = P(A) \cdot P(C \mid A) \cdot P(B|A)$$
$$\psi_2(C_2) = P(D \mid B, C)$$
$$\psi_3(C_3) = P(E \mid C)$$

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

Main Idea

- Incorporate evidence into the clique potentials.
- Since we are dealing with a tree structure, exploit the fact that a clique "separates" all its neighboring cliques (and their respective subtrees) from each other.
- Apply a message passing scheme to inform neighboring cliques about evidence.
- Since we do not have edge directions, we will only need one type of message.
- After having updated all cliques' potentials, we marginalize (and normalize) to get the probabilities of single attributes.



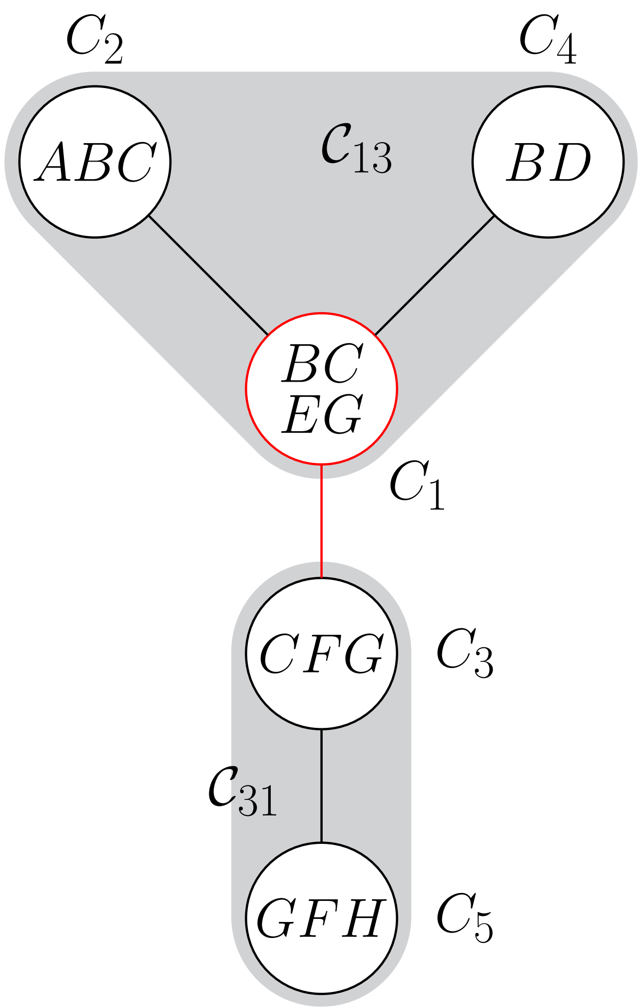
Incorporating Evidence

- Every clique C_i maintains a potential function ψ_i.
- If for an attribute E some evidence e becomes known, we alter all potential functions of cliques containing E as follows:

$$\psi_i^*(c_i) = \begin{cases} 0, & \text{if a value in } c_i \text{ is inconsistent with } e \\ \psi_i(c_i), & \text{otherwise} \end{cases}$$

- All other potential functions are unchanged.

Notation and Nomenclature

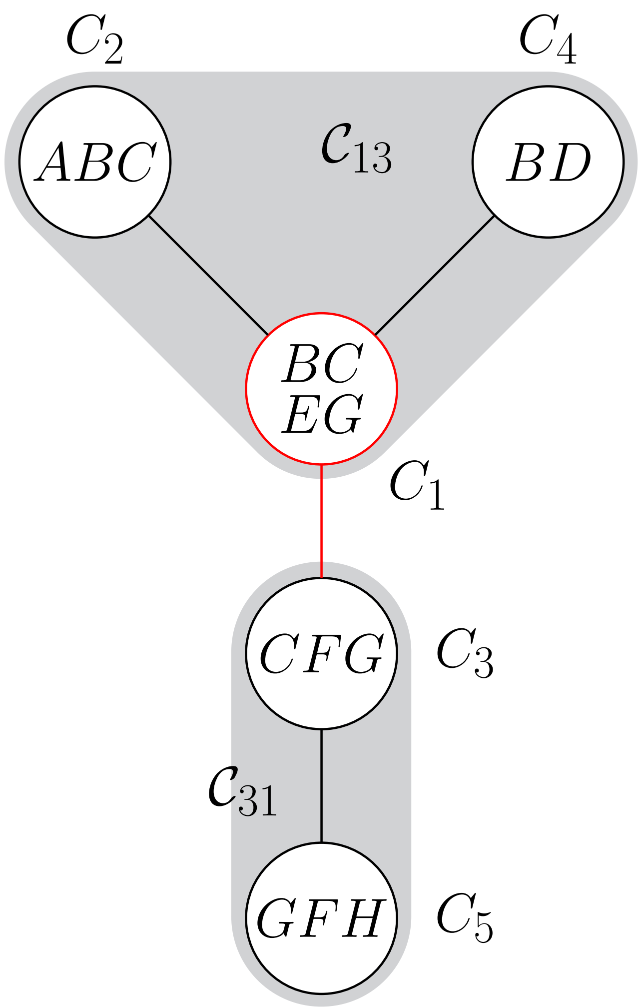


In general:

- Clique C_i has q neighboring cliques $B_1, \dots, B_q$.
- $\mathcal{C}_{ij}$ is the set of cliques in the subtree containing C_i after dropping the link to B_j.
- X_{ij} is the set of attributes in the cliques of $\mathcal{C}_{ij}$.
- $V = X_{ij} \cup X_{ji}$ (complementary sets)
- $S_{ij} = S_{ji} = C_i \cap C_j$ (not shown here)
- $R_{ij} = X_{ij} \setminus S_{ij}$ (not shown here)

Here:

- Neighbors of C_1: $\{C_2, C_4, C_3\}$, $\mathcal{C}_{13} = \{C_1, C_2, C_4\}$
- $X_{13} = \{A, B, C, D, E, G\}$, $S_{13} = \{C, G\}$
- $V = X_{13} \cup X_{31} = \{A, B, C, D, E, F, G, H\}$
- $R_{13} = \{A, B, D, E\}$, $R_{31} = \{F, H\}$

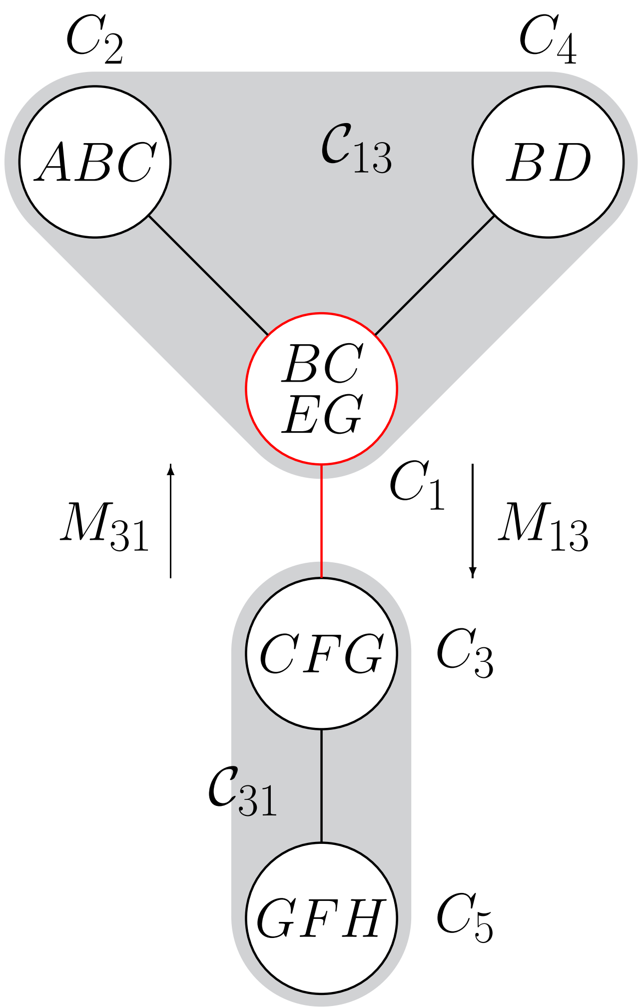


Task: Calculate $P(s_{ij})$:

$$\begin{aligned} V \setminus S_{ij} &= (X_{ij} \cup X_{ji}) \setminus S_{ij} \\ &= (X_{ij} \setminus S_{ij}) \ \cup \ (X_{ji} \setminus S_{ij}) \\ &= R_{ij} \ \cup \ R_{ji} \end{aligned}$$

$$\begin{aligned} V \setminus S_{13} &= (X_{13} \cup X_{31}) \setminus S_{13} \\ &= R_{13} \ \cup \ R_{31} \\ V \setminus \{C, G\} &= \{A, B, D, E\} \ \cup \ \{F, H\} \\ &= \{A, B, D, E, F, H\} \end{aligned}$$

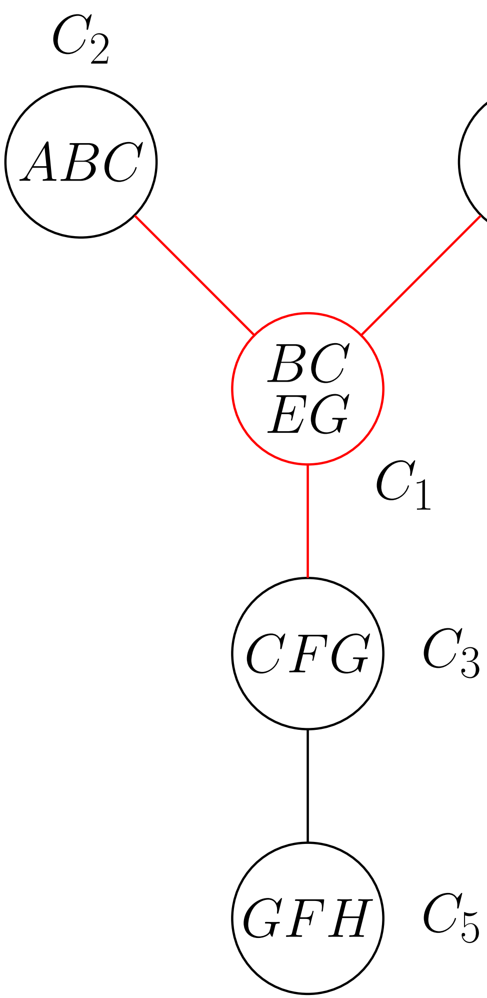
Note: R_{ij} is the set of attributes that are in C_i's subtree but not in B_j's. Therefore, R_{ij} and R_{ji} are always **disjoint**.



Task: Calculate $P(s_{ij})$:

$$
\begin{aligned}
P(s_{ij}) &= \sum_{v \setminus s_{ij}} \prod_{k=1}^{m} \psi_k(c_k) \\
&\overset{\text{last slide}}{=} \sum_{r_{ij} \cup r_{ji}} \prod_{k=1}^{m} \psi_k(c_k) \\
&\overset{\text{sum rule}}{=} \left(\sum_{r_{ij}} \prod_{c_k \in \mathcal{C}_{ij}} \psi_k(c_k) \right) \cdot \left(\sum_{r_{ji}} \prod_{c_k \in \mathcal{C}_{ji}} \psi_k(c_k) \right) \\
&= M_{ij}(s_{ij}) \cdot M_{ji}(s_{ij})
\end{aligned}
$$

M_{ij} is the message sent from C_i to neighbor B_j and vice versa.

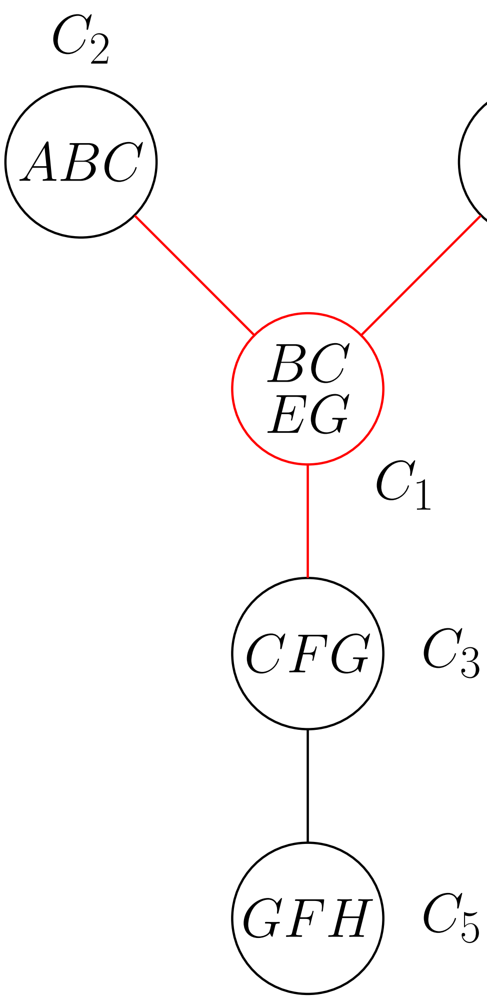


Task: Calculate $P(c_i)$:

$$
\begin{aligned}
V \setminus C_i &= \left(\bigcup_{k=1}^{q} X_{ki}\right) \setminus C_i \\
&= \bigcup_{k=1}^{q} \left(X_{ki} \setminus C_i\right) \\
&= \bigcup_{k=1}^{q} R_{ki}
\end{aligned}
$$

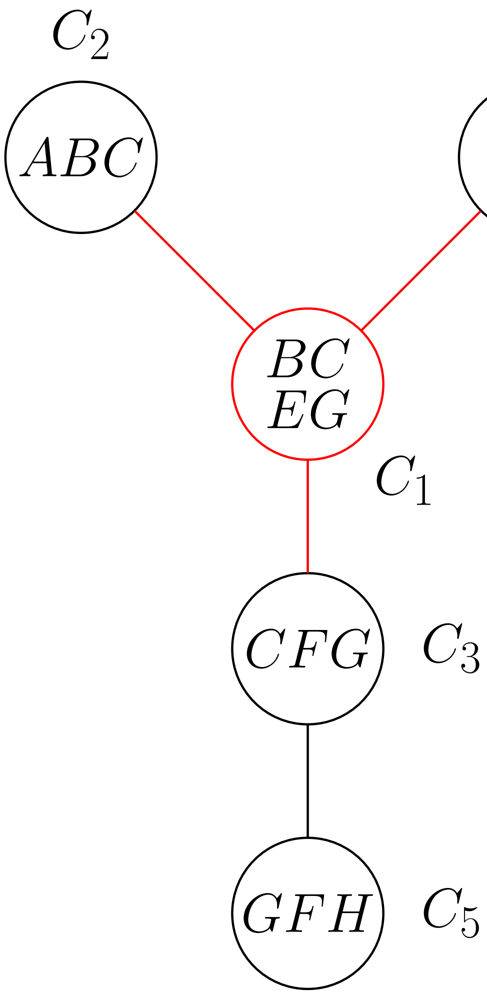
Example:

$$
\begin{aligned}
V \setminus C_1 &= R_{21} \cup R_{41} \cup R_{31} \\
\{A, D, F, H\} &= \{A\} \cup \{D\} \cup \{F, H\}
\end{aligned}
$$



Task: Calculate $P(c_i)$:

$$
\begin{aligned}
P(c_i) &= \underbrace{\sum_{v \backslash c_i}}_{\text{Marginalization}} \underbrace{\prod_{j=1}^{m} \psi_j(c_j)}_{\text{Decomposition}} \\
&= \psi_i(c_i) \sum_{v \backslash c_i} \prod_{i \neq j} \psi_j(c_j) \\
&= \psi_i(c_i) \sum_{r_{1i} \cup \cdots \cup r_{qi}} \prod_{i \neq j} \psi_j(c_j) \\
&= \psi_i(c_i) \underbrace{\left(\sum_{r_{1i}} \prod_{c_k \in \mathcal{C}_{1i}} \psi_k(c_k) \right)}_{M_{1i}(s_{ij})} \cdots \underbrace{\left(\sum_{r_{qi}} \prod_{c_k \in \mathcal{C}_{qi}} \psi_k(c_k) \right)}_{M_{qi}(s_{ij})} \\
&= \psi_i(c_i) \prod_{j=1}^{q} M_{ji}(s_{ij})
\end{aligned}
$$



Example: $P(c_1)$:

$$P(c_1) = \psi_1(c_1)M_{21}(s_{12})M_{41}(s_{14})M_{31}(s_{13})$$

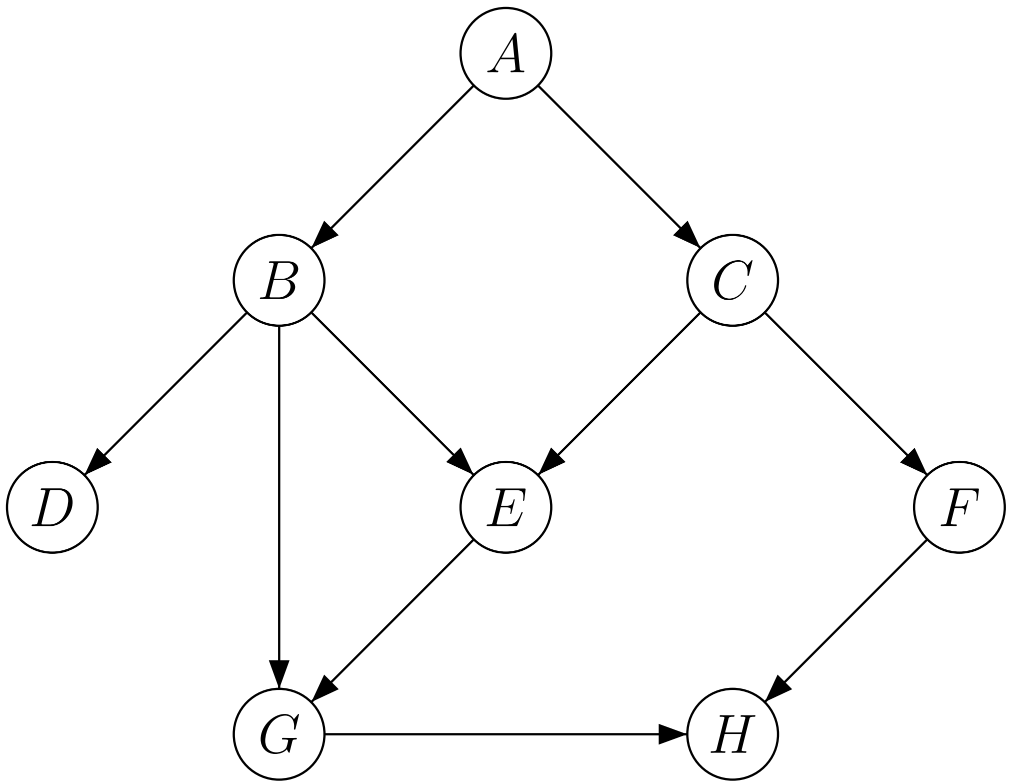
$M_{ij}(s_{ij})$ can be simplified further (without proof):

$$\begin{aligned} M_{ij}(s_{ij}) &= \sum_{r_{ij}} \prod_{c_k \in \mathcal{C}_{ij}} \psi_k(c_k) \\ &= \sum_{c_i \setminus s_{ij}} \psi_i(c_i) \prod_{k \neq j} M_{ki}(s_{ki}) \end{aligned}$$

Final Algorithm

- **Input:** Join tree $(\mathcal{C}, \Psi)$ over set of variables V and evidence $E = e$.
- **Output:** The a-posteriori probability $P(x_i \,|\, e)$ for every non-evidential X_i.
- **Initialization:** Incorporate evidence $E = e$ into potential functions.
- **Iterations:**
  1. For every clique C_i do: For every neighbor B_j of C_i do: If C_i has received all messages from the *other* neighbors, calculate and send $M_{ij}(s_{ij})$ to B_j.
  2. Repeat step 1 until no message is calculated.
  3. Calculate the joint probability distribution for every clique:
  $$P(c_i) \;\propto\; \psi_i(c_i) \prod_{j=1}^{q} M_{ji}(s_{ij})$$
  4. For every $X \in V$ calculate the a-posteriori probability:
  $$P(x_i \mid e) \;=\; \sum_{c_k \setminus x_i} P(c_k)$$
  where C_k is the smallest clique containing X_i.

Example: Putting it together

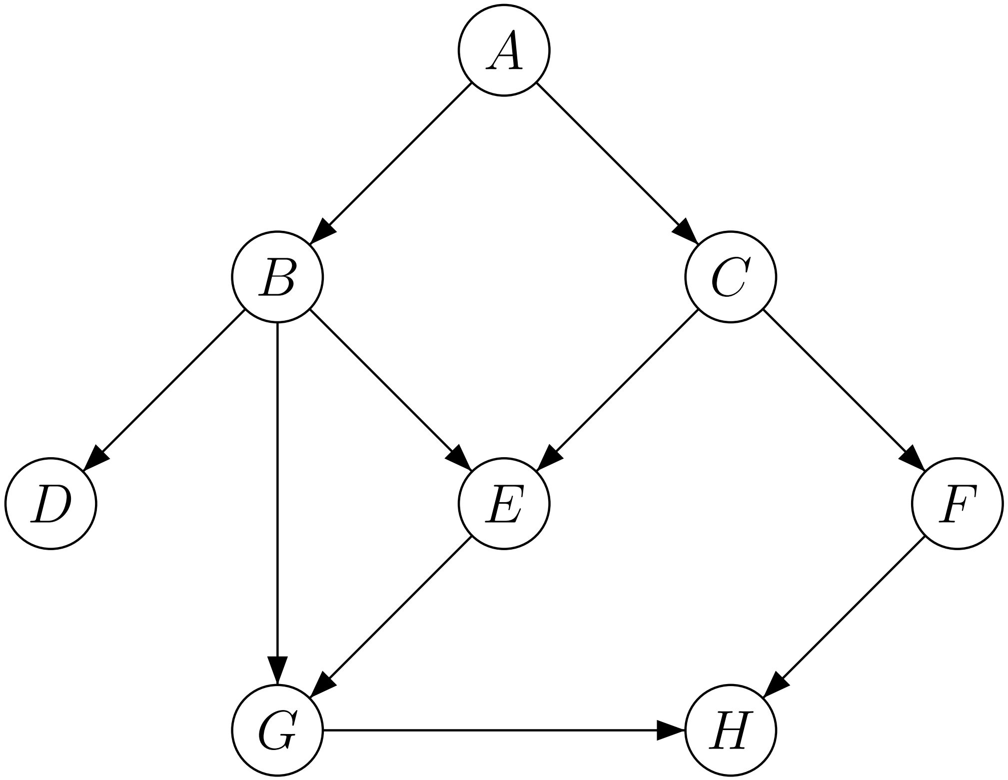


Goals: Find the marginal distributions and update them when evidence $H = h_1$ becomes known.

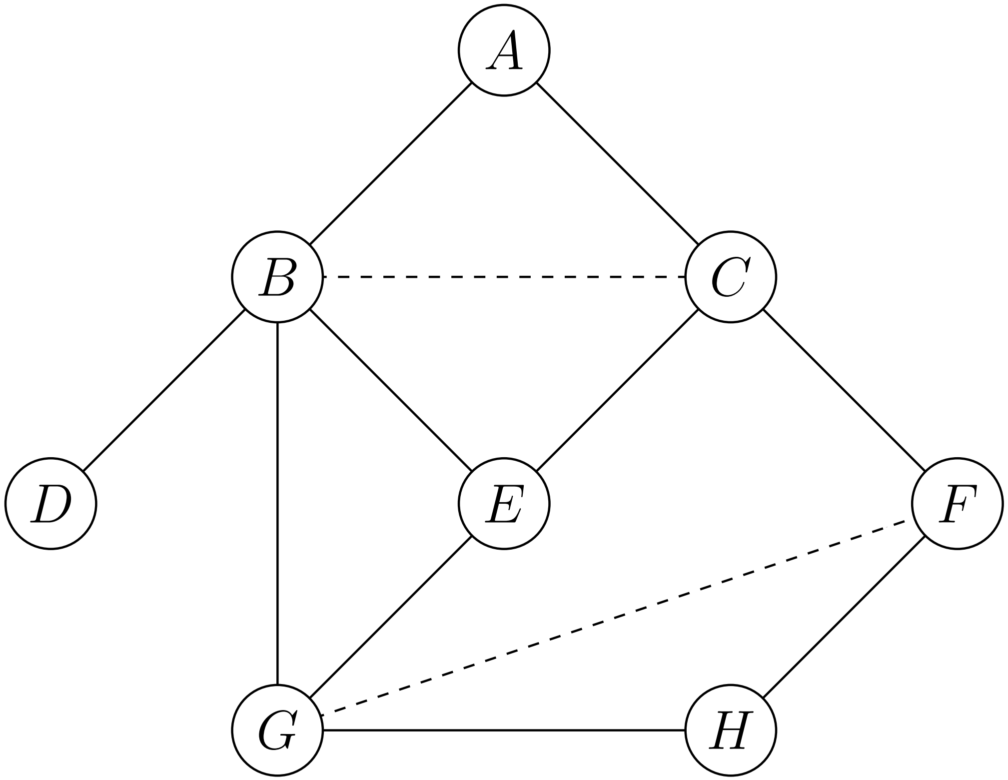
Steps:

1. Transform network into join-tree.
2. Specify factor potentials.
3. Propagate "zero" evidence to obtain the marginals before evidence is present.
4. Update factor potentials w. r. t. the evidence and do another propagation run.

Example: Step 1: Find a Join-Tree

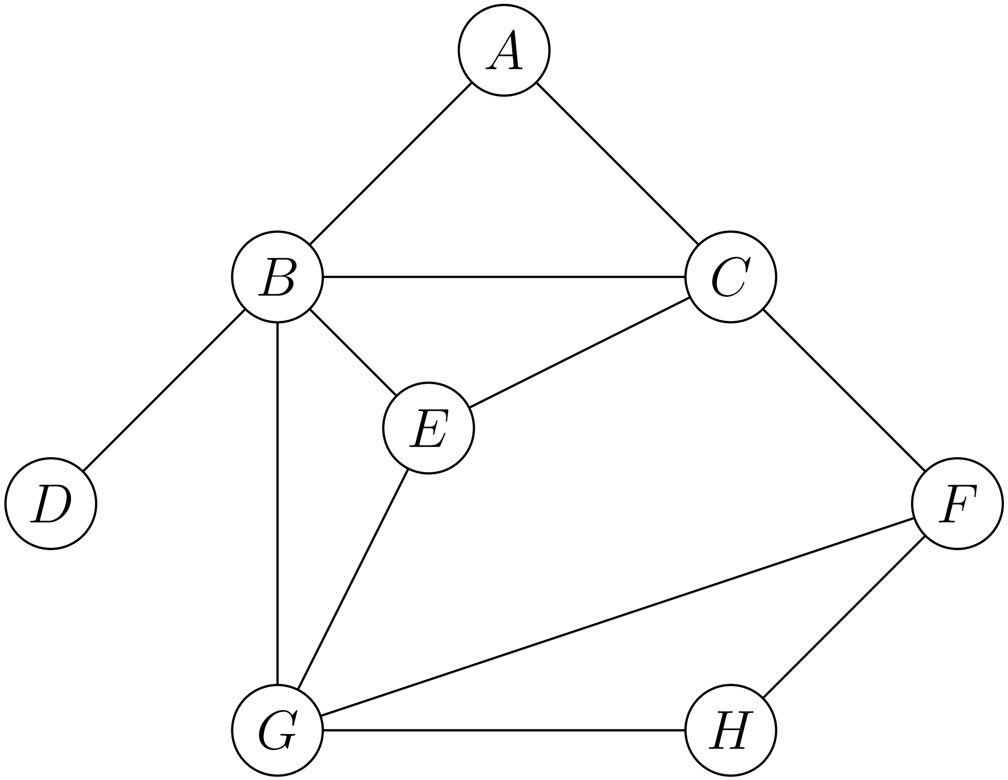


Join-Tree creation:



Join-Tree creation:

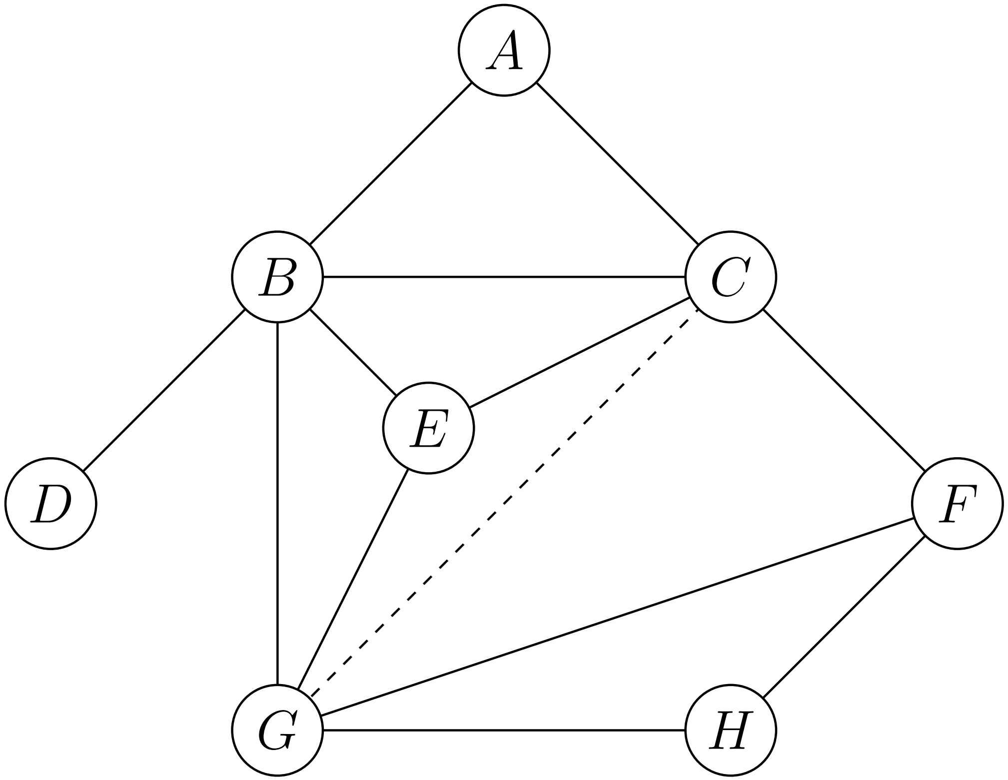
1. Moralize the graph.



Join-Tree creation:

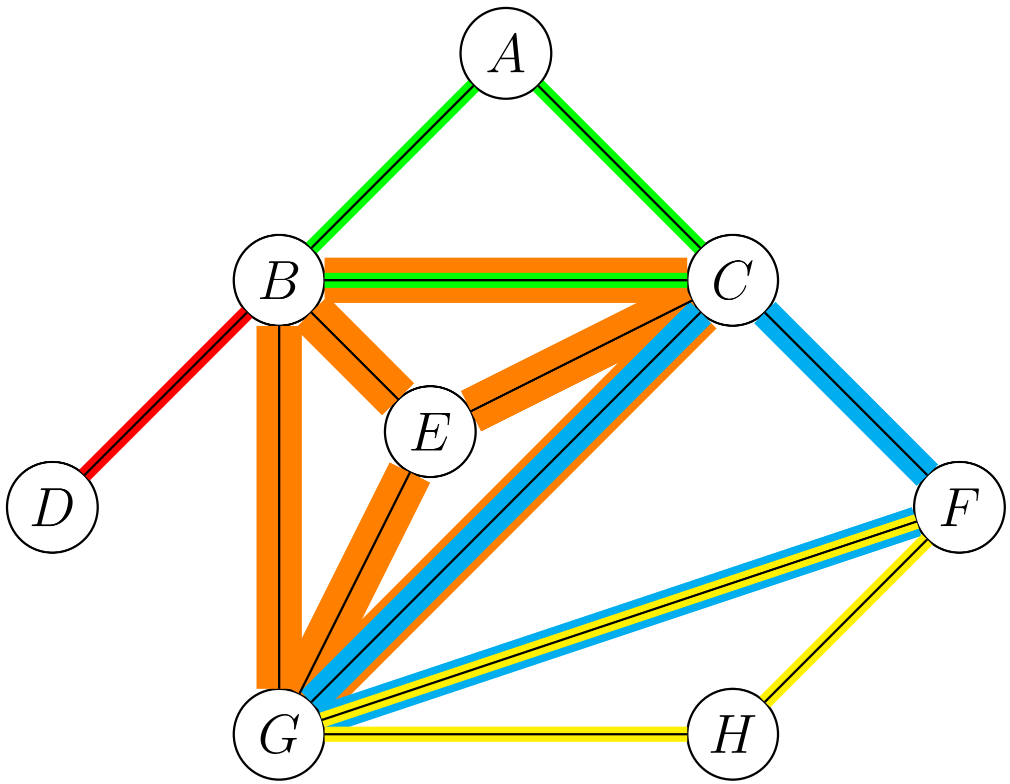
1. Moralize the graph.
2. Not yet triangulated.

Example: Step 1: Find a Join-Tree



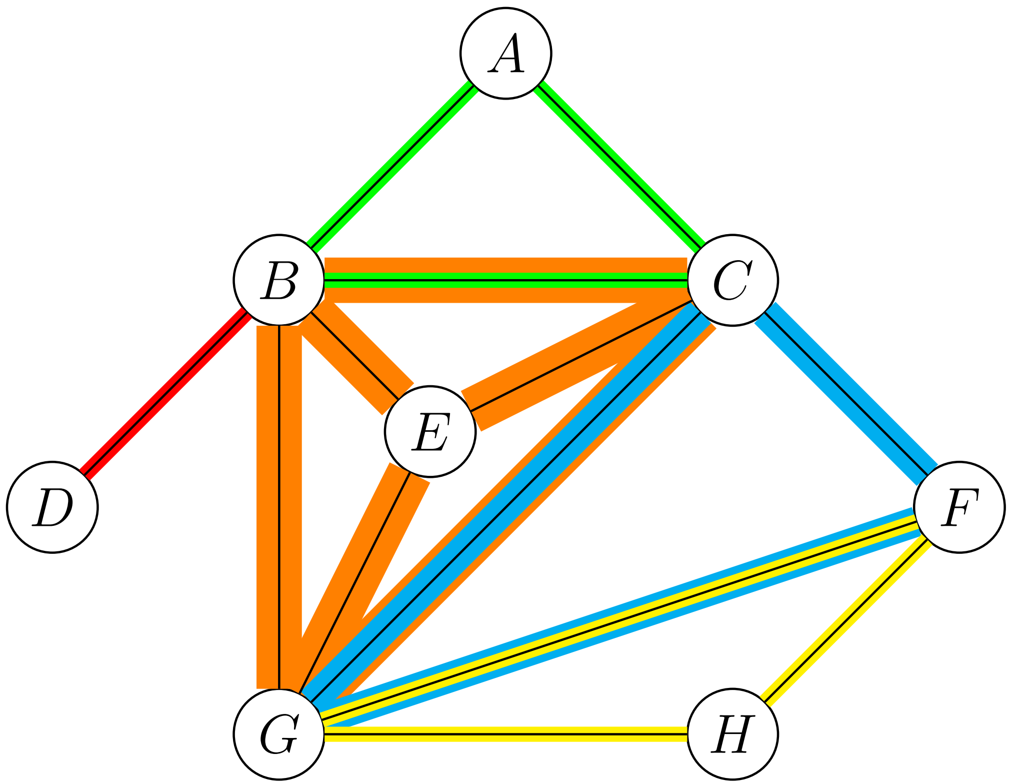
Join-Tree creation:

1. Moralize the graph.
2. Triangulate the graph.



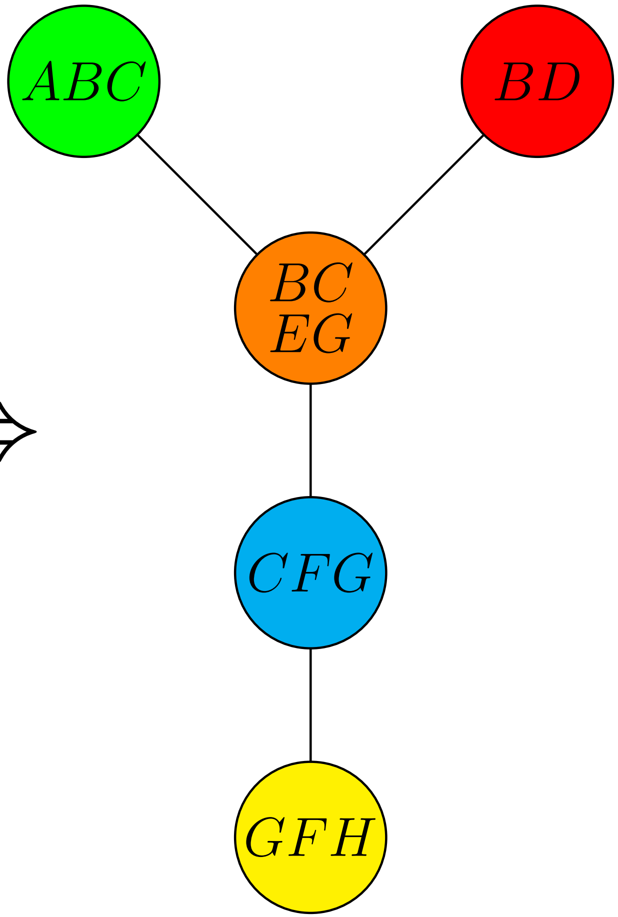
Join-Tree creation:

1. Moralize the graph.
2. Triangulate the graph.
3. Identify the maximal cliques.



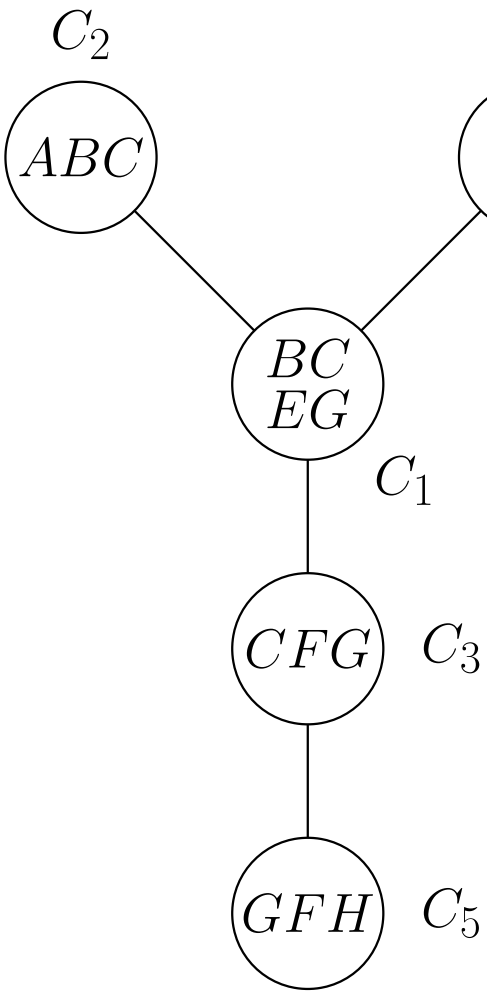
Example Bayesian network

$\Rightarrow$



One of the join trees

Example: Step 2: Specify the Factor Potentials



Decomposition of $P(A, B, C, D, E, F, G, H)$:

$$\begin{aligned} P(a, b, c, d, e, f, g, h) &= \prod_{i=1}^{5} \Psi_i(c_i) \\ &= \Psi_1(b, c, e, g) \cdot \Psi_2(a, b, c) \\ &\quad \cdot \Psi_3(c, f, g) \cdot \Psi_4(b, d) \\ &\qquad \cdot \Psi_5(g, f, h) \end{aligned}$$

Where to get the factor potentials from?

Example: Step 2: Specify the Factor Potentials

As long as the factor potentials multiply together as on the previous slide, we are free to choose them.

- **Option 1:** A factor potential of clique C_i is the product of all conditional probabilities of all node families properly contained in C_i:

$$\Psi_i(c_i) = 1 \cdot \prod_{\substack{\{X_i\} \cup Y_i \subseteq C_i \land \\ \text{parents}(X_i) = Y_i}} P(x_i \mid y_i)$$

  The 1 stresses that if no node family satisfies the product condition, we assign a constant 1 to the potential.

- **Option 2:** Choose potentials from the decomposition formula:

$$P(\bigcup_{i=1}^{n} C_i) = \frac{\prod_{i=1}^{n} P(C_i)}{\prod_{j=1}^{m} P(S_j)}$$

Example: Step 2: Specify the Factor Potentials

- **Option 1:** Factor potentials according to the conditional distributions of the node families of the underlying Bayesian network:

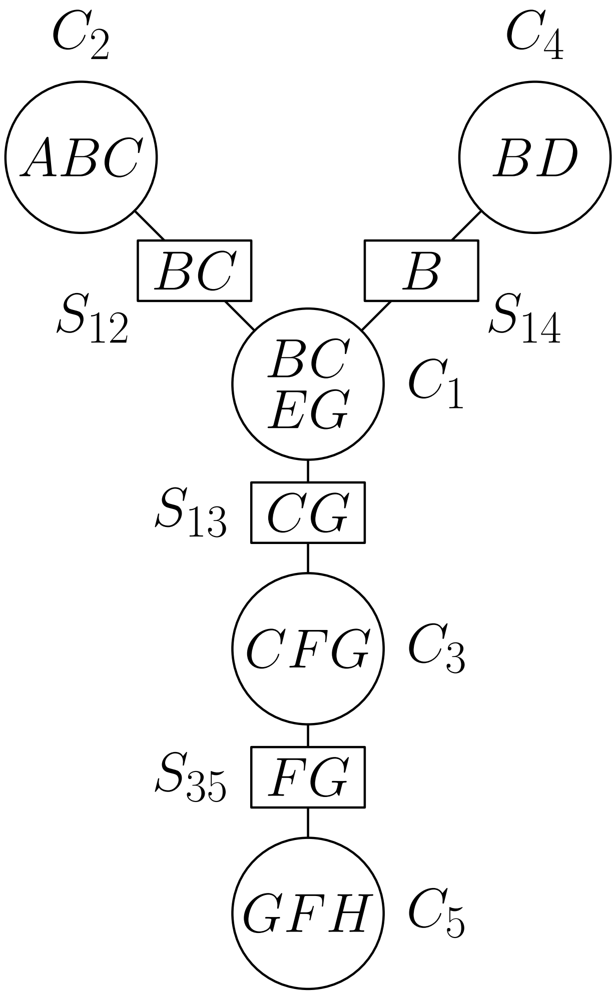
$$\begin{aligned}
\Psi_1(b,c,e,g) &= P(e \mid b,c) \cdot P(g \mid e,b) \\
\Psi_2(a,b,c) &= P(b \mid a) \cdot P(c \mid a) \cdot P(a) \\
\Psi_3(c,f,g) &= P(f \mid c) \\
\Psi_4(b,d) &= P(d \mid b) \\
\Psi_5(g,f,h) &= P(h \mid g,f)
\end{aligned}$$

(This assignment of factor potentials is used in this example.)

- **Option 2:** Factor potentials chosen from the join-tree decomposition:

$$\begin{aligned}
\Psi_1(b,c,e,g) &= P(b,e \mid c,g) \\
\Psi_2(a,b,c) &= P(a \mid b,c) \\
\Psi_3(c,f,g) &= P(c \mid f,g) \\
\Psi_4(b,d) &= P(d \mid b) \\
\Psi_5(g,f,h) &= P(h,g,f)
\end{aligned}$$

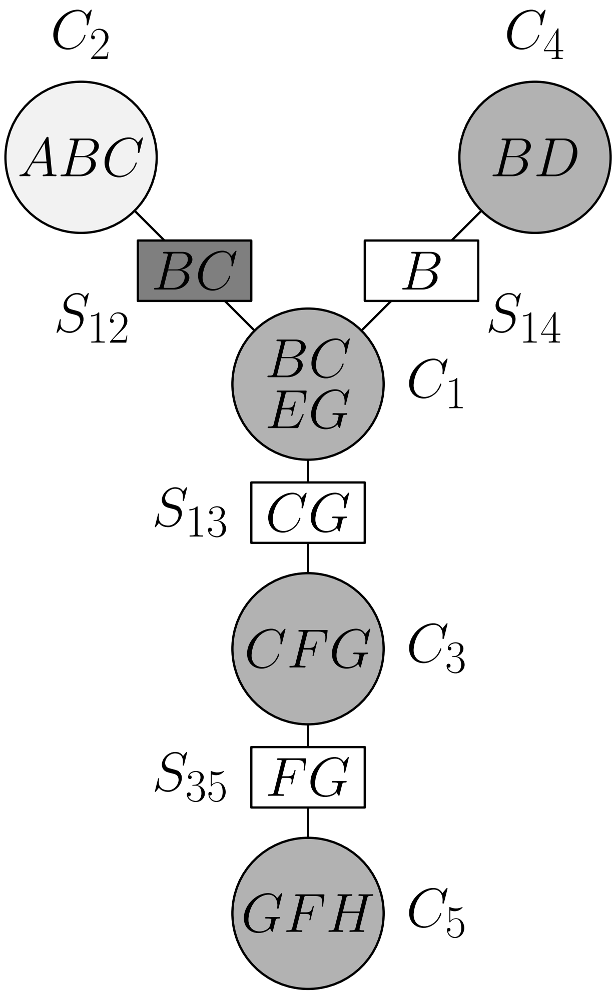
Example: Closer Look on Option 2: Separation in a Join-Tree



Encoded independence statements:

Given any separator, the variables in the cliques on one side become independent of the variables in the cliques on the other side.

Example: Closer Look on Option 2: Separation in a Join-Tree

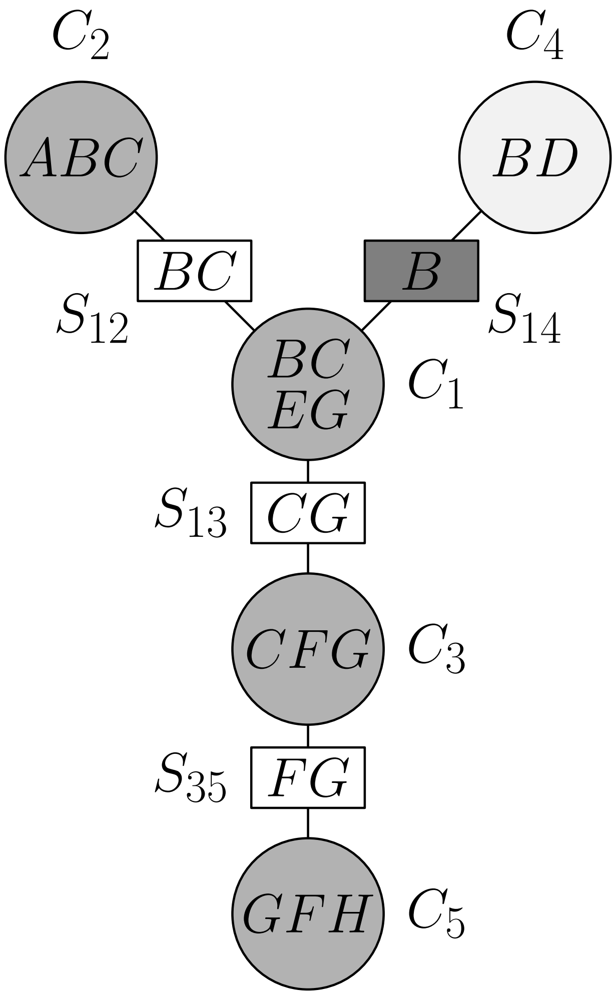


Encoded independence statements:

Given any separator, the variables in the cliques on one side become independent of the variables in the cliques on the other side.

$$A \perp\!\!\!\perp D, E, F, G, H \mid B, C$$

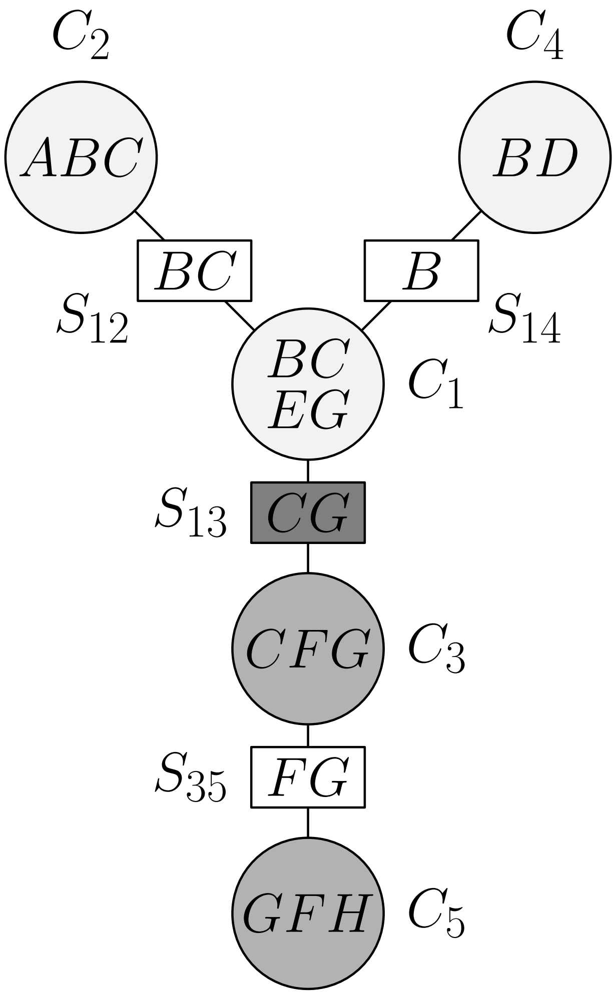
Example: Closer Look on Option 2: Separation in a Join-Tree



Encoded independence statements:

Given any separator, the variables in the cliques on one side become independent of the variables in the cliques on the other side.

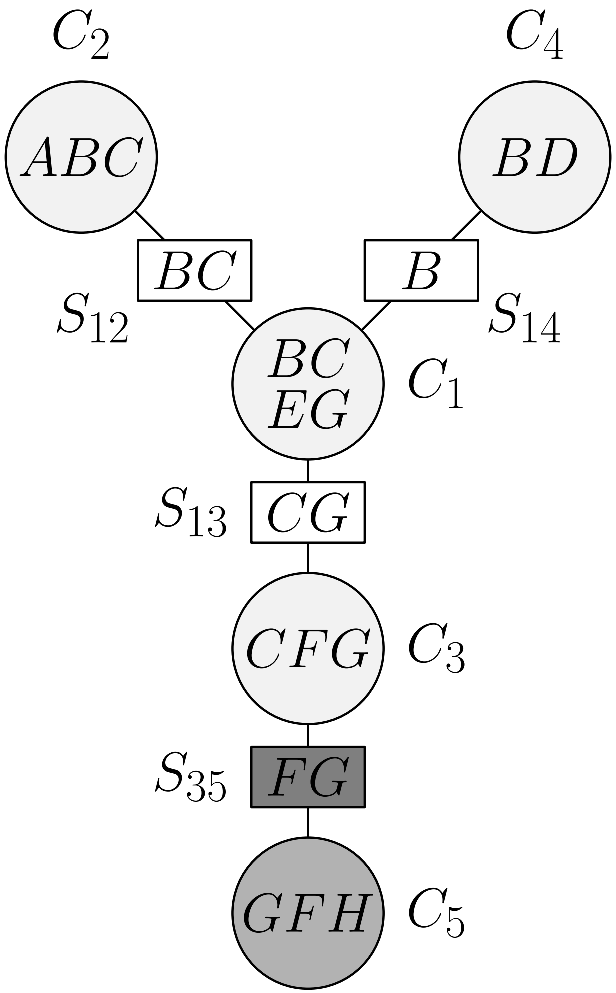
$$A \perp\!\!\!\perp D, E, F, G, H \mid B, C$$
$$D \perp\!\!\!\perp A, C, E, F, G, H \mid B$$



Encoded independence statements:

Given any separator, the variables in the cliques on one side become independent of the variables in the cliques on the other side.

$$A \perp\!\!\!\perp D, E, F, G, H \mid B, C$$
$$D \perp\!\!\!\perp A, C, E, F, G, H \mid B$$
$$A, B, E, D \perp\!\!\!\perp F, H \mid G, C$$



Encoded independence statements:

Given any separator, the variables in the cliques on one side become independent of the variables in the cliques on the other side.

$$A \perp\!\!\!\perp D, E, F, G, H \mid B, C$$
$$D \perp\!\!\!\perp A, C, E, F, G, H \mid B$$
$$A, B, E, D \perp\!\!\!\perp F, H \mid G, C$$
$$H \perp\!\!\!\perp A, B, C, D, E \mid F, G$$

Example: Closer Look on Option 2: Decomposition

The four separation statements translate into the following independence statements:

$$
\begin{array}{rlll}
A \perp\!\!\!\perp D, E, F, G, H \mid B, C & \Leftrightarrow P(A \mid B, C, D, E, F, G, H) & = P(A \mid B, C) \\
D \perp\!\!\!\perp A, C, E, F, G, H \mid B & \Rightarrow P(D \mid B, C, E, F, G, H) & = P(D \mid B) \\
A, B, E, D \perp\!\!\!\perp F, H \mid G, C & \Rightarrow P(B, E \mid G, C, F, H) & = P(B, E \mid G, C) \\
H \perp\!\!\!\perp A, B, C, D, E \mid F, G & \Rightarrow P(C \mid F, G, H) & = P(C \mid F, G)
\end{array}
$$

According to the chain rule we always have the following relation:

$$
\begin{aligned}
P(A, B, C, D, E, F, G, H) = {} & P(A \mid B, C, D, E, F, G, H) \cdot \\
& P(D \mid B, C, E, F, G, H) \cdot \\
& P(B, E \mid C, F, G, H) \cdot \\
& P(C \mid F, G, H) \cdot \\
& P(F, G, H)
\end{aligned}
$$

The four separation statements translate into the following independence statements:

$$
\begin{array}{rll}
A \perp\!\!\!\perp D, E, F, G, H \mid B, C & \Leftrightarrow P(A \mid B, C, D, E, F, G, H) & = P(A \mid B, C) \\
D \perp\!\!\!\perp A, C, E, F, G, H \mid B & \Rightarrow P(D \mid B, C, E, F, G, H) & = P(D \mid B) \\
A, B, E, D \perp\!\!\!\perp F, H \mid G, C & \Rightarrow P(B, E \mid G, C, F, H) & = P(B, E \mid G, C) \\
H \perp\!\!\!\perp A, B, C, D, E \mid F, G & \Rightarrow P(C \mid F, G, H) & = P(C \mid F, G)
\end{array}
$$

Exploiting the above independencies yields:

$$
\begin{aligned}
P(A, B, C, D, E, F, G, H) \;=\; & P(A \mid B, C) \cdot \\
& P(D \mid B) \cdot \\
& P(B, E \mid C, G) \cdot \\
& P(C \mid F, G) \cdot \\
& P(F, G, H)
\end{aligned}
$$

Example: Closer Look on Option 2: Decomposition

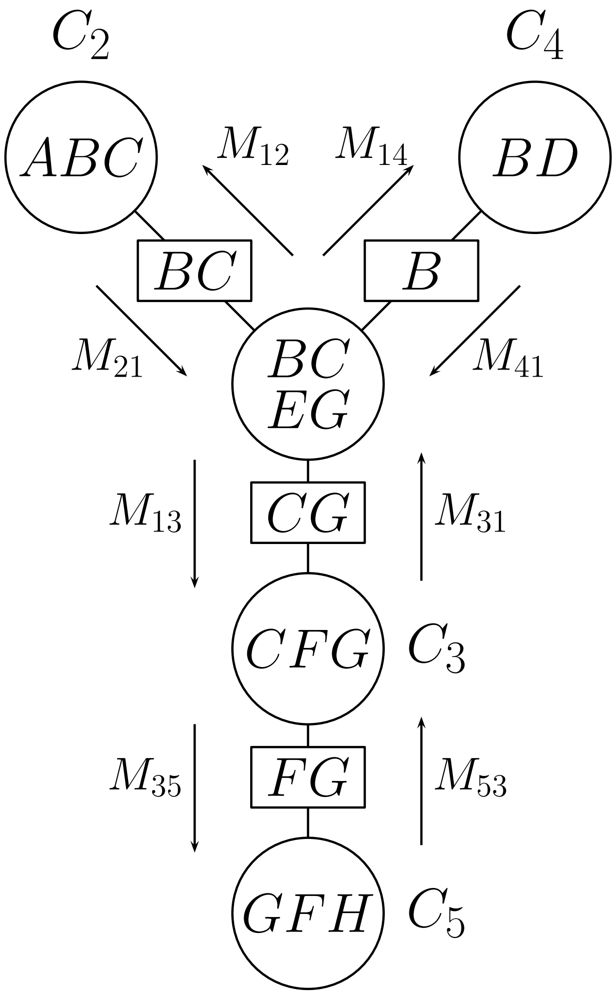
The four separation statements translate into the following independence statements:

$$\begin{array}{rlcl}
A \perp\!\!\!\perp D, E, F, G, H \mid B, C & \Leftrightarrow P(A \mid B, C, D, E, F, G, H) & = P(A \mid B, C) \\
D \perp\!\!\!\perp A, C, E, F, G, H \mid B & \Rightarrow P(D \mid B, C, E, F, G, H) & = P(D \mid B) \\
A, B, E, D \perp\!\!\!\perp F, H \mid G, C & \Rightarrow P(B, E \mid G, C, F, H) & = P(B, E \mid G, C) \\
H \perp\!\!\!\perp A, B, C, D, E \mid F, G & \Rightarrow P(C \mid F, G, H) & = P(C \mid F, G)
\end{array}$$

Getting rid of the conditions results in the final decomposition equation:

$$\begin{aligned}
P(A, B, C, D, E, F, G, H) &= P(A \mid B, C) P(D \mid B) P(B, E \mid C, G) P(C \mid F, G) P(F, G, H) \\
&= \frac{P(A, B, C) P(D, B) P(B, E, C, G) P(C, F, G) P(F, G, H)}{P(B, C) P(B) P(C, G) P(F, G)} \\
&= \frac{P(C_1) P(C_2) P(C_3) P(C_4) P(C_5)}{P(S_{12}) P(S_{14}) P(S_{13}) P(S_{35})}
\end{aligned}$$

Example: Step 3: Messages to be sent for Propagation



- According to the join-tree propagation algorithm, the probability distributions of all clique instantiations c_i is calculated as follows:

$$P(c_i) \;\propto\; \Psi_i(c_i) \prod_{j=1}^{q} M_{ji}(s_{ij})$$

- Spelt out for our example, we get:

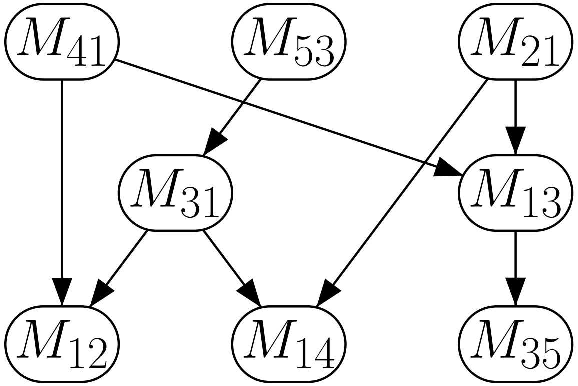
$$\begin{aligned}
P(c_1) &= P(b,c,e,g) &&= \Psi_1(b,c,e,g) && \cdot M_{21}(b,c) \cdot M_{31}(c,g) \cdot M_{41}(b) \\
P(c_2) &= P(a,b,c) &&\propto \Psi_2(a,b,c) && \cdot M_{12}(b,c) \\
P(c_3) &= P(c,f,g) &&\propto \Psi_3(c,f,g) && \cdot M_{13}(c,g) \cdot M_{53}(f,g) \\
P(c_4) &= P(b,d) &&\propto \Psi_4(b,d) && \cdot M_{14}(b) \\
P(c_5) &= P(f,g,h) &&\propto \Psi_5(f,g,h) && \cdot M_{35}(f,g)
\end{aligned}$$

- The $\propto$-symbol indicates that the right-hand side may not add up to one. In that case we just normalize.

Example: Step 3: Message Computation Order

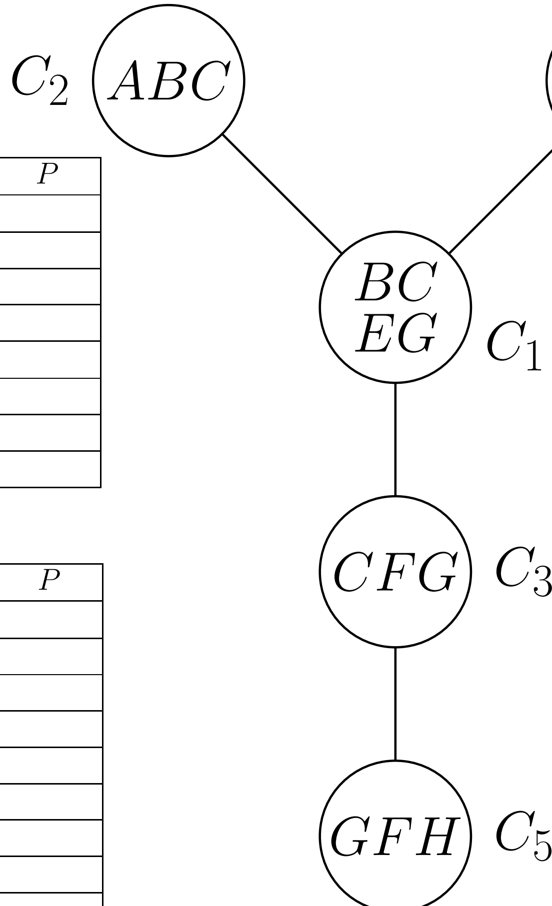
- The structure of the join-tree imposes a partial ordering according to which the messages need to be computed:

$$
\begin{aligned}
M_{41}(b) &= \sum_{d} \Psi_4(b, d) \\
M_{53}(f, g) &= \sum_{h} \Psi_5(f, g, h) \\
M_{21}(b, c) &= \sum_{a} \Psi_2(a, b, c) \\
M_{31}(c, g) &= \sum_{f} \Psi_3(c, f, g) M_{53}(f, g) \\
M_{13}(c, g) &= \sum_{b,e} \Psi_1(b, c, e, g) M_{21}(b, c) M_{41}(b) \\
M_{12}(b, c) &= \sum_{e,g} \Psi_2(b, c, e, g) M_{31}(c, g) M_{41}(b) \\
M_{14}(b) &= \sum_{c,e,g} \Psi_1(b, c, e, g) M_{21}(b, c) M_{31}(c, g) \\
M_{35}(f, g) &= \sum_{c} \Psi_3(c, f, g) M_{13}(c, g)
\end{aligned}
$$



Arrows represent is-needed-for relations. Messages on the same level can be computed in any order. Messages are computed level-wise from top to bottom.

Example: Step 3: Initialization (Potential Layouts)



Ψ_2				P
a_1	b_1	c_1		
		c_2		
	b_2	c_1		
		c_2		
a_2	b_1	c_1		
		c_2		
	b_2	c_1		
		c_2		

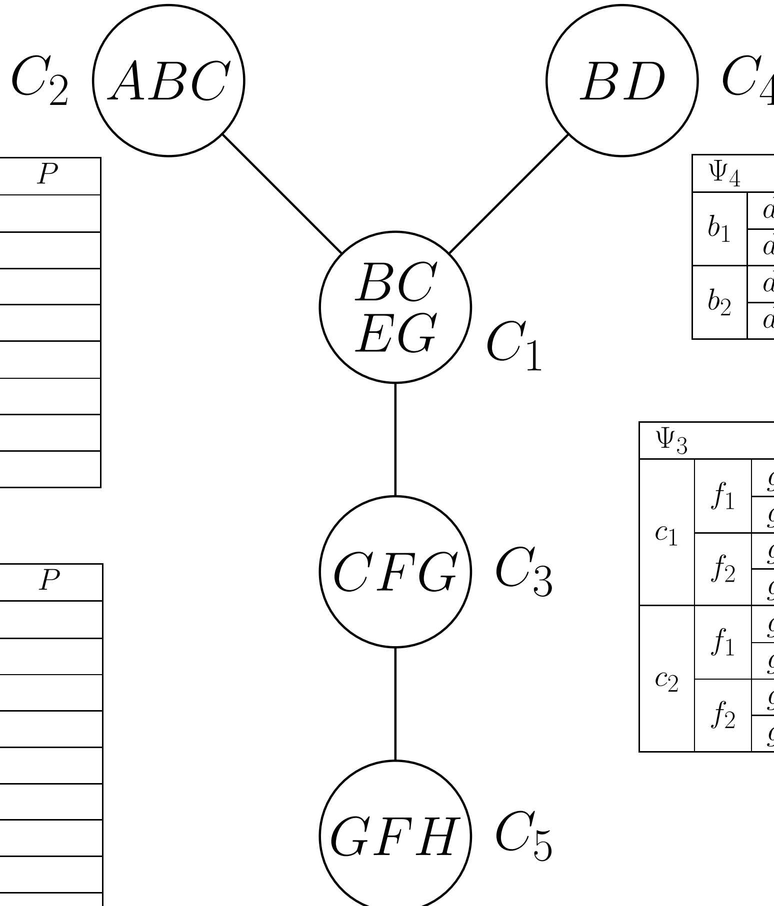
Ψ_4			P
b_1	d_1		
	d_2		
b_2	d_1		
	d_2		

Ψ_1					P
b_1	c_1	e_1	g_1		
			g_2		
		e_2	g_1		
			g_2		
	c_2	e_1	g_1		
			g_2		
		e_2	g_1		
			g_2		
b_2	c_1	e_1	g_1		
			g_2		
		e_2	g_1		
			g_2		
	c_2	e_1	g_1		
			g_2		
		e_2	g_1		
			g_2		

Ψ_3				P
c_1	f_1	g_1		
		g_2		
	f_2	g_1		
		g_2		
c_2	f_1	g_1		
		g_2		
	f_2	g_1		
		g_2		

Ψ_5				P
f_1	g_1	h_1		
		h_2		
	g_2	h_1		
		h_2		
f_2	g_1	h_1		
		h_2		
	g_2	h_1		
		h_2		

Example: Step 3: Initialization (Potential Values)



Ψ_2				P
a_1	b_1	c_1	0.036	
		c_2	0.084	
	b_2	c_1	0.144	
		c_2	0.336	
a_2	b_1	c_1	0.028	
		c_2	0.012	
	b_2	c_1	0.252	
		c_2	0.108	

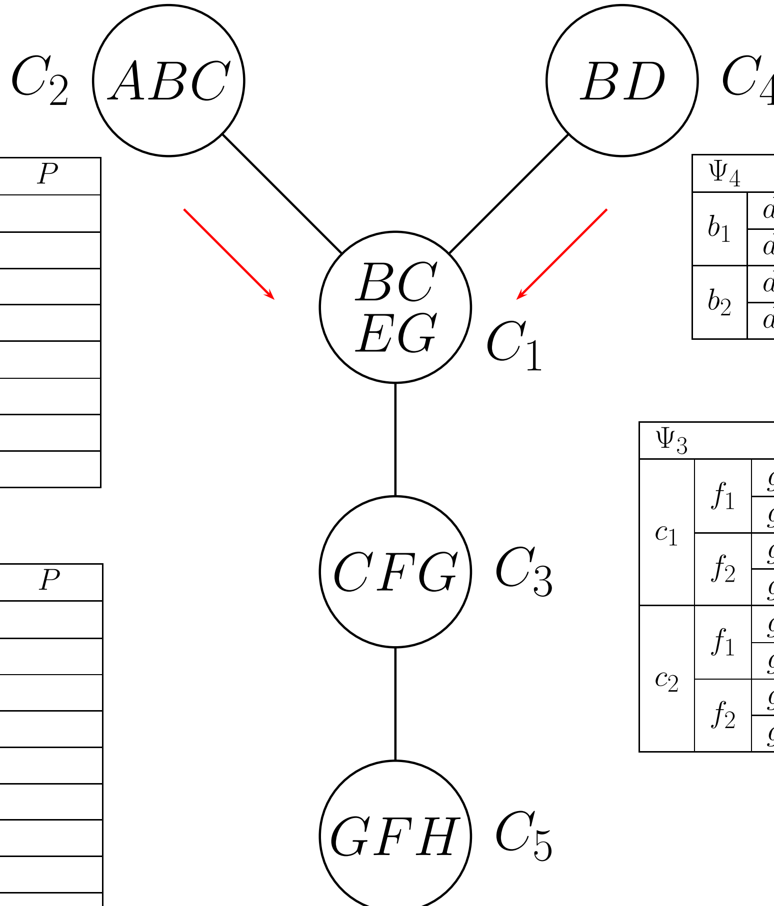
Ψ_4			P
b_1	d_1	0.4	
	d_2	0.6	
b_2	d_1	0.7	
	d_2	0.3	

Ψ_3				P
c_1	f_1	g_1	0.1	
		g_2	0.1	
	f_2	g_1	0.9	
		g_2	0.9	
c_2	f_1	g_1	0.4	
		g_2	0.4	
	f_2	g_1	0.6	
		g_2	0.6	

Ψ_1					P
b_1	c_1	e_1	g_1	0.190	
			g_2	0.010	
		e_2	g_1	0.320	
			g_2	0.480	
	c_2	e_1	g_1	0.380	
			g_2	0.020	
		e_2	g_1	0.240	
			g_2	0.360	
b_2	c_1	e_1	g_1	0.210	
			g_2	0.090	
		e_2	g_1	0.350	
			g_2	0.350	
	c_2	e_1	g_1	0.070	
			g_2	0.030	
		e_2	g_1	0.450	
			g_2	0.450	

Ψ_5				P
f_1	g_1	h_1	0.2	
		h_2	0.8	
	g_2	h_1	0.5	
		h_2	0.5	
f_2	g_1	h_1	0.4	
		h_2	0.6	
	g_2	h_1	0.7	
		h_2	0.3	

Example: Step 3: Initialization (Sending Messages)



$$M_{21} = \begin{array}{c} \begin{matrix} b_1,c_1 & b_1,c_2 & b_2,c_1 & b_2,c_2 \end{matrix} \\ \big(0.06, 0.10, 0.40, 0.44\big) \end{array}$$

$$M_{41} = \begin{array}{c} \begin{matrix} b_1 & b_2 \end{matrix} \\ \big(1, 1\big) \end{array}$$

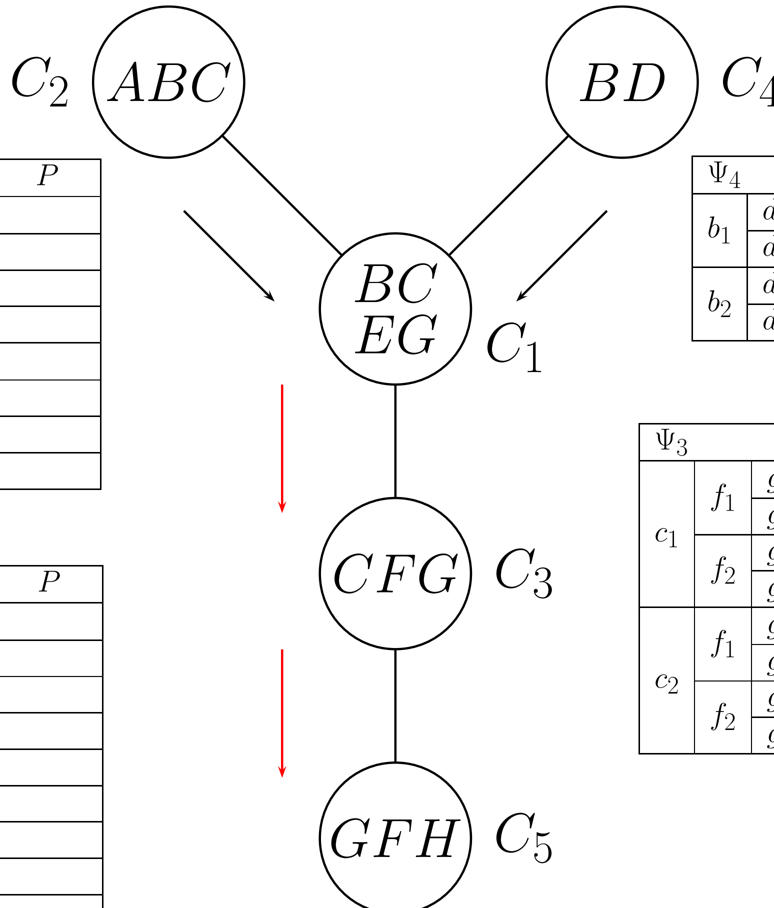
Ψ_2				P
a_1	b_1	c_1	0.036	
		c_2	0.084	
	b_2	c_1	0.144	
		c_2	0.336	
a_2	b_1	c_1	0.028	
		c_2	0.012	
	b_2	c_1	0.252	
		c_2	0.108	

Ψ_4			P
b_1	d_1	0.4	
	d_2	0.6	
b_2	d_1	0.7	
	d_2	0.3	

Ψ_3				P
c_1	f_1	g_1	0.1	
		g_2	0.1	
	f_2	g_1	0.9	
		g_2	0.9	
c_2	f_1	g_1	0.4	
		g_2	0.4	
	f_2	g_1	0.6	
		g_2	0.6	

Ψ_1					P
b_1	c_1	e_1	g_1	0.190	
			g_2	0.010	
		e_2	g_1	0.320	
			g_2	0.480	
	c_2	e_1	g_1	0.380	
			g_2	0.020	
		e_2	g_1	0.240	
			g_2	0.360	
b_2	c_1	e_1	g_1	0.210	
			g_2	0.090	
		e_2	g_1	0.350	
			g_2	0.350	
	c_2	e_1	g_1	0.070	
			g_2	0.030	
		e_2	g_1	0.450	
			g_2	0.450	

Ψ_5				P
f_1	g_1	h_1	0.2	
		h_2	0.8	
	g_2	h_1	0.5	
		h_2	0.5	
f_2	g_1	h_1	0.4	
		h_2	0.6	
	g_2	h_1	0.7	
		h_2	0.3	



Ψ_2				P
a_1	b_1	c_1	0.036	
		c_2	0.084	
	b_2	c_1	0.144	
		c_2	0.336	
a_2	b_1	c_1	0.028	
		c_2	0.012	
	b_2	c_1	0.252	
		c_2	0.108	

Ψ_4			P
b_1	d_1	0.4	
	d_2	0.6	
b_2	d_1	0.7	
	d_2	0.3	

Ψ_1					P
b_1	c_1	e_1	g_1	0.190	
			g_2	0.010	
		e_2	g_1	0.320	
			g_2	0.480	
	c_2	e_1	g_1	0.380	
			g_2	0.020	
		e_2	g_1	0.240	
			g_2	0.360	
b_2	c_1	e_1	g_1	0.210	
			g_2	0.090	
		e_2	g_1	0.350	
			g_2	0.350	
	c_2	e_1	g_1	0.070	
			g_2	0.030	
		e_2	g_1	0.450	
			g_2	0.450	

Ψ_3				P
c_1	f_1	g_1	0.1	
		g_2	0.1	
	f_2	g_1	0.9	
		g_2	0.9	
c_2	f_1	g_1	0.4	
		g_2	0.4	
	f_2	g_1	0.6	
		g_2	0.6	

Ψ_5				P
f_1	g_1	h_1	0.2	
		h_2	0.8	
	g_2	h_1	0.5	
		h_2	0.5	
f_2	g_1	h_1	0.4	
		h_2	0.6	
	g_2	h_1	0.7	
		h_2	0.3	

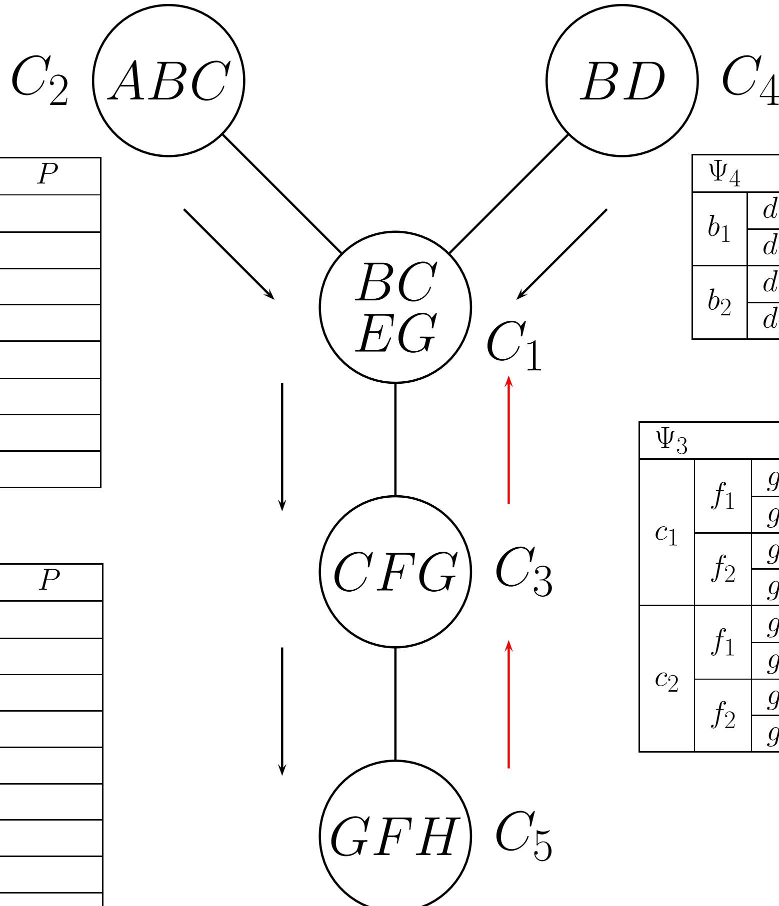
$$M_{21} = \begin{array}{cccc} b_1,c_1 & b_1,c_2 & b_2,c_1 & b_2,c_2 \\ \end{array} \left(0.06,\ 0.10,\ 0.40,\ 0.44\right)$$

$$M_{41} = \begin{array}{cc} b_1 & b_2 \end{array} \left(1,\ 1\right)$$

$$M_{13} = \begin{array}{cccc} c_1,g_1 & c_1,g_2 & c_2,g_1 & c_2,g_2 \end{array} \left(0.254,\ 0.206,\ 0.290,\ 0.250\right)$$

$$M_{35} = \begin{array}{cccc} f_1,g_1 & f_1,g_2 & f_2,g_1 & f_2,g_2 \end{array} \left(0.14,\ 0.12,\ 0.40,\ 0.33\right)$$

Example: Step 3: Initialization (Sending Messages)



Ψ_2				P
a_1	b_1	c_1	0.036	
		c_2	0.084	
	b_2	c_1	0.144	
		c_2	0.336	
a_2	b_1	c_1	0.028	
		c_2	0.012	
	b_2	c_1	0.252	
		c_2	0.108	

Ψ_4			P
b_1	d_1	0.4	
	d_2	0.6	
b_2	d_1	0.7	
	d_2	0.3	

Ψ_3				P
c_1	f_1	g_1	0.1	
		g_2	0.1	
	f_2	g_1	0.9	
		g_2	0.9	
c_2	f_1	g_1	0.4	
		g_2	0.4	
	f_2	g_1	0.6	
		g_2	0.6	

Ψ_1					P
b_1	c_1	e_1	g_1	0.190	
			g_2	0.010	
		e_2	g_1	0.320	
			g_2	0.480	
	c_2	e_1	g_1	0.380	
			g_2	0.020	
		e_2	g_1	0.240	
			g_2	0.360	
b_2	c_1	e_1	g_1	0.210	
			g_2	0.090	
		e_2	g_1	0.350	
			g_2	0.350	
	c_2	e_1	g_1	0.070	
			g_2	0.030	
		e_2	g_1	0.450	
			g_2	0.450	

Ψ_5				P
f_1	g_1	h_1	0.2	
		h_2	0.8	
	g_2	h_1	0.5	
		h_2	0.5	
f_2	g_1	h_1	0.4	
		h_2	0.6	
	g_2	h_1	0.7	
		h_2	0.3	

$$M_{21} = \left(\overset{b_1,c_1}{0.06}, \overset{b_1,c_2}{0.10}, \overset{b_2,c_1}{0.40}, \overset{b_2,c_2}{0.44}\right)$$

$$M_{41} = \left(\overset{b_1}{1}, \overset{b_2}{1}\right)$$

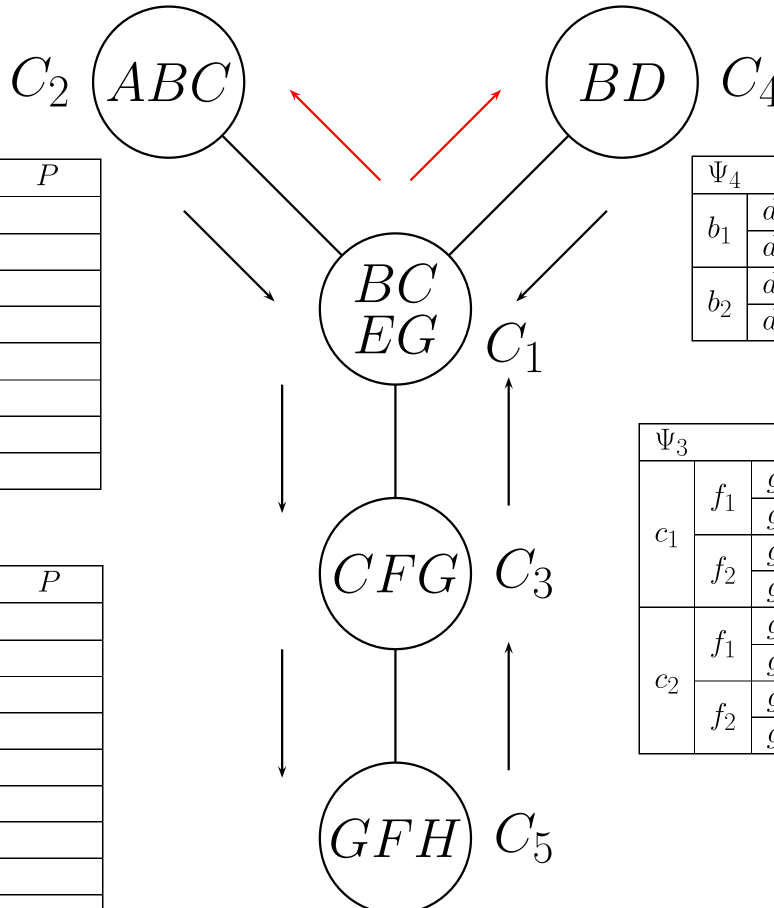
$$M_{13} = \left(\overset{c_1,g_1}{0.254}, \overset{c_1,g_2}{0.206}, \overset{c_2,g_1}{0.290}, \overset{c_2,g_2}{0.250}\right)$$

$$M_{35} = \left(\overset{f_1,g_1}{0.14}, \overset{f_1,g_2}{0.12}, \overset{f_2,g_1}{0.40}, \overset{f_2,g_2}{0.33}\right)$$

$$M_{53} = \left(\overset{f_1,g_1}{1}, \overset{f_1,g_2}{1}, \overset{f_2,g_1}{1}, \overset{f_2,g_2}{1}\right)$$

$$M_{31} = \left(\overset{c_1,g_1}{1}, \overset{c_1,g_2}{1}, \overset{c_2,g_1}{1}, \overset{c_2,g_2}{1}\right)$$

Example: Step 3: Initialization (Sending Messages)



Ψ_2				P
a_1	b_1	c_1	0.036	
		c_2	0.084	
	b_2	c_1	0.144	
		c_2	0.336	
a_2	b_1	c_1	0.028	
		c_2	0.012	
	b_2	c_1	0.252	
		c_2	0.108	

Ψ_4			P
b_1	d_1	0.4	
	d_2	0.6	
b_2	d_1	0.7	
	d_2	0.3	

Ψ_3				P
c_1	f_1	g_1	0.1	
		g_2	0.1	
	f_2	g_1	0.9	
		g_2	0.9	
c_2	f_1	g_1	0.4	
		g_2	0.4	
	f_2	g_1	0.6	
		g_2	0.6	

Ψ_1					P
b_1	c_1	e_1	g_1	0.190	
			g_2	0.010	
		e_2	g_1	0.320	
			g_2	0.480	
	c_2	e_1	g_1	0.380	
			g_2	0.020	
		e_2	g_1	0.240	
			g_2	0.360	
b_2	c_1	e_1	g_1	0.210	
			g_2	0.090	
		e_2	g_1	0.350	
			g_2	0.350	
	c_2	e_1	g_1	0.070	
			g_2	0.030	
		e_2	g_1	0.450	
			g_2	0.450	

Ψ_5				P
f_1	g_1	h_1	0.2	
		h_2	0.8	
	g_2	h_1	0.5	
		h_2	0.5	
f_2	g_1	h_1	0.4	
		h_2	0.6	
	g_2	h_1	0.7	
		h_2	0.3	

$$M_{21} = \left(\overset{b_1,c_1}{0.06}, \overset{b_1,c_2}{0.10}, \overset{b_2,c_1}{0.40}, \overset{b_2,c_2}{0.44}\right)$$

$$M_{41} = \left(\overset{b_1}{1}, \overset{b_2}{1}\right)$$

$$M_{13} = \left(\overset{c_1,g_1}{0.254}, \overset{c_1,g_2}{0.206}, \overset{c_2,g_1}{0.290}, \overset{c_2,g_2}{0.250}\right)$$

$$M_{35} = \left(\overset{f_1,g_1}{0.14}, \overset{f_1,g_2}{0.12}, \overset{f_2,g_1}{0.40}, \overset{f_2,g_2}{0.33}\right)$$

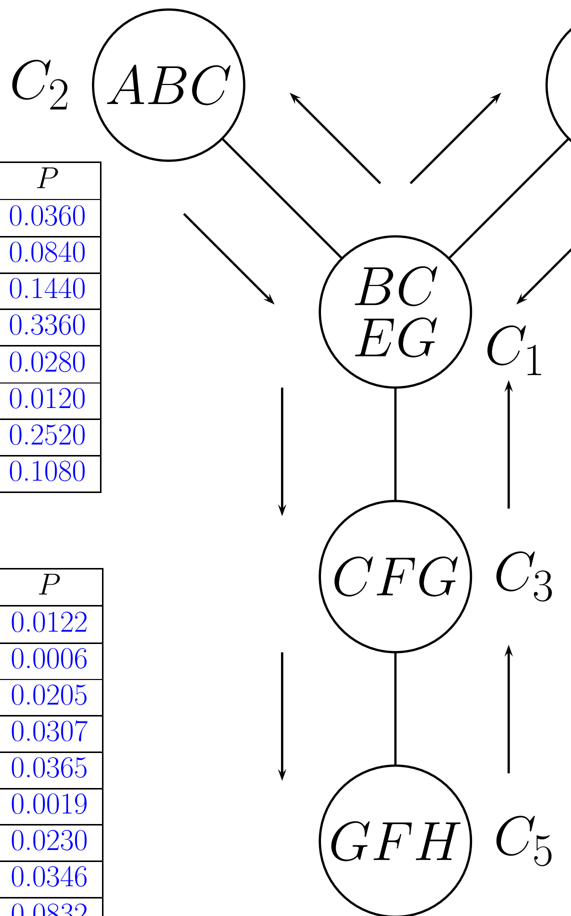
$$M_{53} = \left(\overset{f_1,g_1}{1}, \overset{f_1,g_2}{1}, \overset{f_2,g_1}{1}, \overset{f_2,g_2}{1}\right)$$

$$M_{31} = \left(\overset{c_1,g_1}{1}, \overset{c_1,g_2}{1}, \overset{c_2,g_1}{1}, \overset{c_2,g_2}{1}\right)$$

$$M_{12} = \left(\overset{b_1,c_1}{1}, \overset{b_1,c_2}{1}, \overset{b_2,c_1}{1}, \overset{b_2,c_2}{1}\right)$$

$$M_{14} = \left(\overset{b_1}{0.16}, \overset{b_2}{0.84}\right)$$

Example: Step 3: Initialization Complete



Ψ_2				P
a_1	b_1	c_1	0.036	0.0360
		c_2	0.084	0.0840
	b_2	c_1	0.144	0.1440
		c_2	0.336	0.3360
a_2	b_1	c_1	0.028	0.0280
		c_2	0.012	0.0120
	b_2	c_1	0.252	0.2520
		c_2	0.108	0.1080

Ψ_1					P
b_1	c_1	e_1	g_1	0.190	0.0122
			g_2	0.010	0.0006
		e_2	g_1	0.320	0.0205
			g_2	0.480	0.0307
	c_2	e_1	g_1	0.380	0.0365
			g_2	0.020	0.0019
		e_2	g_1	0.240	0.0230
			g_2	0.360	0.0346
b_2	c_1	e_1	g_1	0.210	0.0832
			g_2	0.090	0.0356
		e_2	g_1	0.350	0.1386
			g_2	0.350	0.1386
	c_2	e_1	g_1	0.070	0.0311
			g_2	0.030	0.0133
		e_2	g_1	0.450	0.1998
			g_2	0.450	0.1998

Ψ_5				P
f_1	g_1	h_1	0.2	0.0283
		h_2	0.8	0.1133
	g_2	h_1	0.5	0.0602
		h_2	0.5	0.0602
f_2	g_1	h_1	0.4	0.1613
		h_2	0.6	0.2419
	g_2	h_1	0.7	0.2344
		h_2	0.3	0.1004

Ψ_4			P
b_1	d_1	0.4	0.0640
	d_2	0.6	0.0960
b_2	d_1	0.7	0.5880
	d_2	0.3	0.2520

Ψ_3				P
c_1	f_1	g_1	0.1	0.0254
		g_2	0.1	0.0206
	f_2	g_1	0.9	0.2290
		g_2	0.9	0.1850
c_2	f_1	g_1	0.4	0.1162
		g_2	0.4	0.0998
	f_2	g_1	0.6	0.1742
		g_2	0.6	0.1498

$$M_{21} = \begin{array}{c} \begin{array}{cccc} b_1,c_1 & b_1,c_2 & b_2,c_1 & b_2,c_2 \end{array} \\ (0.06, 0.10, 0.40, 0.44) \end{array}$$

$$M_{41} = \begin{array}{c} \begin{array}{cc} b_1 & b_2 \end{array} \\ (1, 1) \end{array}$$

$$M_{13} = \begin{array}{c} \begin{array}{cccc} c_1,g_1 & c_1,g_2 & c_2,g_1 & c_2,g_2 \end{array} \\ (0.254, 0.206, 0.290, 0.250) \end{array}$$

$$M_{35} = \begin{array}{c} \begin{array}{cccc} f_1,g_1 & f_1,g_2 & f_2,g_1 & f_2,g_2 \end{array} \\ (0.14, 0.12, 0.40, 0.33) \end{array}$$

$$M_{53} = \begin{array}{c} \begin{array}{cccc} f_1,g_1 & f_1,g_2 & f_2,g_1 & f_2,g_2 \end{array} \\ (1, 1, 1, 1) \end{array}$$

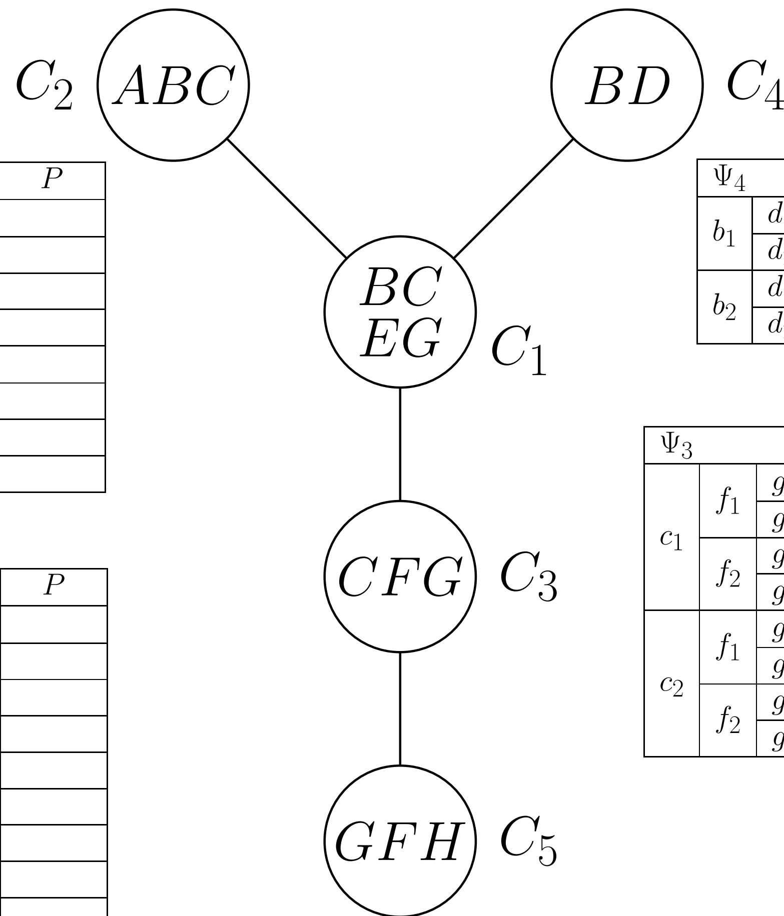
$$M_{31} = \begin{array}{c} \begin{array}{cccc} c_1,g_1 & c_1,g_2 & c_2,g_1 & c_2,g_2 \end{array} \\ (1, 1, 1, 1) \end{array}$$

$$M_{12} = \begin{array}{c} \begin{array}{cccc} b_1,c_1 & b_1,c_2 & b_2,c_1 & b_2,c_2 \end{array} \\ (1, 1, 1, 1) \end{array}$$

$$M_{14} = \begin{array}{c} \begin{array}{cc} b_1 & b_2 \end{array} \\ (0.16, 0.84) \end{array}$$

P	A	B	C	D	E	F	G	H
$\cdot_1$	**0.6000**	**0.1600**	**0.4600**	**0.6520**	**0.2144**	**0.2620**	**0.5448**	**0.4842**
$\cdot_2$	**0.4000**	**0.8400**	**0.4500**	**0.3480**	**0.7856**	**0.7380**	**0.4552**	**0.5158**

Example: Step 4: Evidence $H = h_1$ (Altering Potentials)



Ψ_2				P
a_1	b_1	c_1	0.036	
		c_2	0.084	
	b_2	c_1	0.144	
		c_2	0.336	
a_2	b_1	c_1	0.028	
		c_2	0.012	
	b_2	c_1	0.252	
		c_2	0.108	

Ψ_4			P
b_1	d_1	0.4	
	d_2	0.6	
b_2	d_1	0.7	
	d_2	0.3	

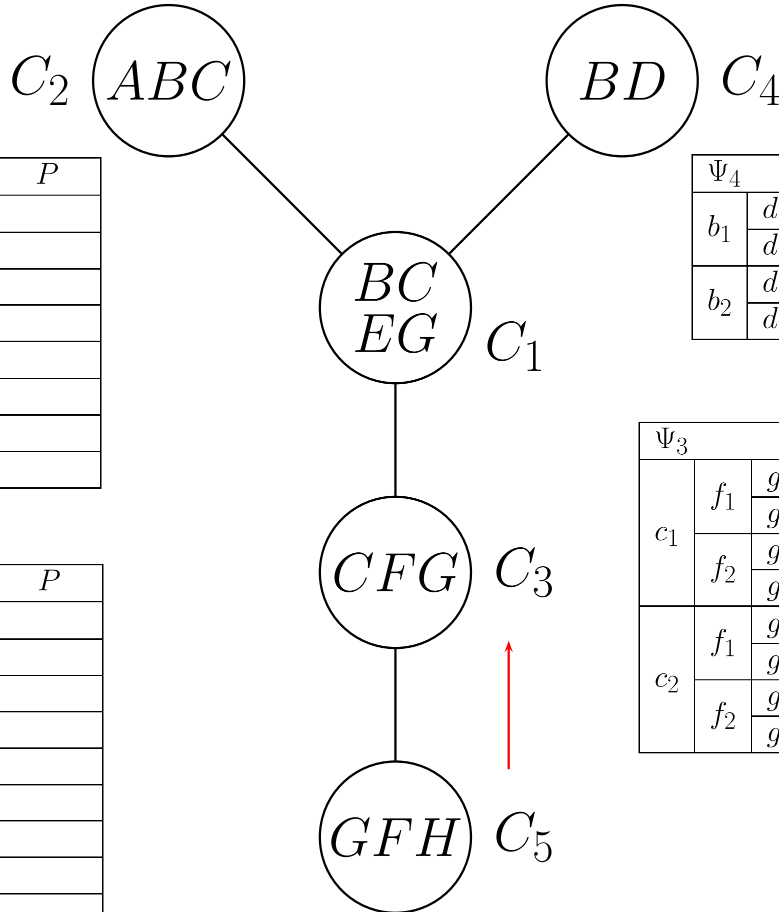
Ψ_3				P
c_1	f_1	g_1	0.1	
		g_2	0.1	
	f_2	g_1	0.9	
		g_2	0.9	
c_2	f_1	g_1	0.4	
		g_2	0.4	
	f_2	g_1	0.6	
		g_2	0.6	

Ψ_1					P
b_1	c_1	e_1	g_1	0.190	
			g_2	0.010	
		e_2	g_1	0.320	
			g_2	0.480	
	c_2	e_1	g_1	0.380	
			g_2	0.020	
		e_2	g_1	0.240	
			g_2	0.360	
b_2	c_1	e_1	g_1	0.210	
			g_2	0.090	
		e_2	g_1	0.350	
			g_2	0.350	
	c_2	e_1	g_1	0.070	
			g_2	0.030	
		e_2	g_1	0.450	
			g_2	0.450	

Ψ_5				P
f_1	g_1	h_1	0.2	
		h_2	**0**	
	g_2	h_1	0.5	
		h_2	**0**	
f_2	g_1	h_1	0.4	
		h_2	**0**	
	g_2	h_1	0.7	
		h_2	**0**	

Example: Step 4: Evidence $H = h_1$ (Sending Messages)

$$M_{53} = \begin{array}{cccc} f_1,g_1 & f_1,g_2 & f_2,g_1 & f_2,g_2 \end{array} \atop \left(\begin{array}{cccc} 0.2, & 0.5, & 0.4, & 0.7 \end{array} \right)$$



Ψ_2				P
a_1	b_1	c_1	0.036	
		c_2	0.084	
	b_2	c_1	0.144	
		c_2	0.336	
a_2	b_1	c_1	0.028	
		c_2	0.012	
	b_2	c_1	0.252	
		c_2	0.108	

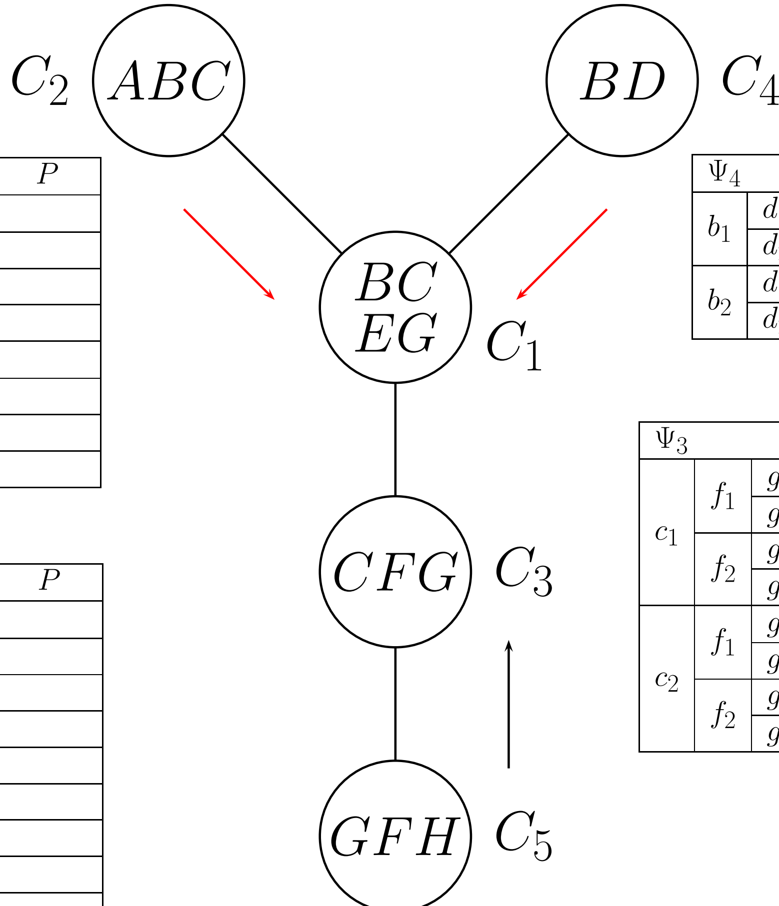
Ψ_4			P
b_1	d_1	0.4	
	d_2	0.6	
b_2	d_1	0.7	
	d_2	0.3	

Ψ_3				P
c_1	f_1	g_1	0.1	
		g_2	0.1	
	f_2	g_1	0.9	
		g_2	0.9	
c_2	f_1	g_1	0.4	
		g_2	0.4	
	f_2	g_1	0.6	
		g_2	0.6	

Ψ_1					P
b_1	c_1	e_1	g_1	0.190	
			g_2	0.010	
		e_2	g_1	0.320	
			g_2	0.480	
	c_2	e_1	g_1	0.380	
			g_2	0.020	
		e_2	g_1	0.240	
			g_2	0.360	
b_2	c_1	e_1	g_1	0.210	
			g_2	0.090	
		e_2	g_1	0.350	
			g_2	0.350	
	c_2	e_1	g_1	0.070	
			g_2	0.030	
		e_2	g_1	0.450	
			g_2	0.450	

Ψ_5				P
f_1	g_1	h_1	0.2	
		h_2	0	
	g_2	h_1	0.5	
		h_2	0	
f_2	g_1	h_1	0.4	
		h_2	0	
	g_2	h_1	0.7	
		h_2	0	

Example: Step 4: Evidence $H = h_1$ (Sending Messages)



Ψ_2				P
a_1	b_1	c_1	0.036	
		c_2	0.084	
	b_2	c_1	0.144	
		c_2	0.336	
a_2	b_1	c_1	0.028	
		c_2	0.012	
	b_2	c_1	0.252	
		c_2	0.108	

Ψ_4			P
b_1	d_1	0.4	
	d_2	0.6	
b_2	d_1	0.7	
	d_2	0.3	

$$M_{53} = \begin{array}{c} \quad f_1,g_1 \;\; f_1,g_2 \;\; f_2,g_1 \;\; f_2,g_2 \\ \left(0.2\;,\; 0.5\;,\; 0.4\;,\; 0.7 \right) \end{array}$$

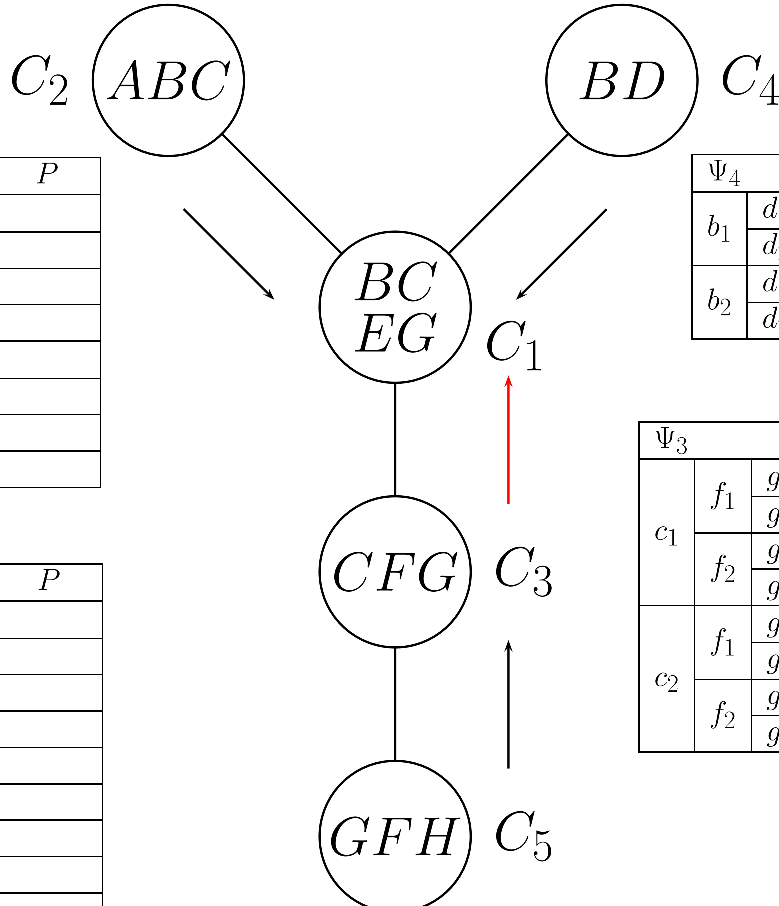
$$M_{21} = \begin{array}{c} \quad b_1,c_1 \;\; b_1,c_2 \;\; b_2,c_1 \;\; b_2,c_2 \\ \left(0.06, 0.10, 0.40, 0.44 \right) \end{array}$$

$$M_{41} = \begin{array}{c} \quad b_1 \;\; b_2 \\ \left(1\,,\, 1 \right) \end{array}$$

Ψ_1					P
b_1	c_1	e_1	g_1	0.190	
			g_2	0.010	
		e_2	g_1	0.320	
			g_2	0.480	
	c_2	e_1	g_1	0.380	
			g_2	0.020	
		e_2	g_1	0.240	
			g_2	0.360	
b_2	c_1	e_1	g_1	0.210	
			g_2	0.090	
		e_2	g_1	0.350	
			g_2	0.350	
	c_2	e_1	g_1	0.070	
			g_2	0.030	
		e_2	g_1	0.450	
			g_2	0.450	

Ψ_3				P
c_1	f_1	g_1	0.1	
		g_2	0.1	
	f_2	g_1	0.9	
		g_2	0.9	
c_2	f_1	g_1	0.4	
		g_2	0.4	
	f_2	g_1	0.6	
		g_2	0.6	

Ψ_5				P
f_1	g_1	h_1	0.2	
		h_2	0	
	g_2	h_1	0.5	
		h_2	0	
f_2	g_1	h_1	0.4	
		h_2	0	
	g_2	h_1	0.7	
		h_2	0	



Ψ_2				P
a_1	b_1	c_1	0.036	
		c_2	0.084	
	b_2	c_1	0.144	
		c_2	0.336	
a_2	b_1	c_1	0.028	
		c_2	0.012	
	b_2	c_1	0.252	
		c_2	0.108	

Ψ_4			P
b_1	d_1	0.4	
	d_2	0.6	
b_2	d_1	0.7	
	d_2	0.3	

$$M_{53} = \begin{array}{c} \quad f_1,g_1 \;\; f_1,g_2 \;\; f_2,g_1 \;\; f_2,g_2 \\ \left(0.2\,,\; 0.5\,,\; 0.4\,,\; 0.7\right) \end{array}$$

$$M_{21} = \begin{array}{c} \quad b_1,c_1 \;\; b_1,c_2 \;\; b_2,c_1 \;\; b_2,c_2 \\ \left(0.06, 0.10, 0.40, 0.44\right) \end{array}$$

$$M_{41} = \begin{array}{c} b_1 \;\; b_2 \\ \left(1,\; 1\right) \end{array}$$

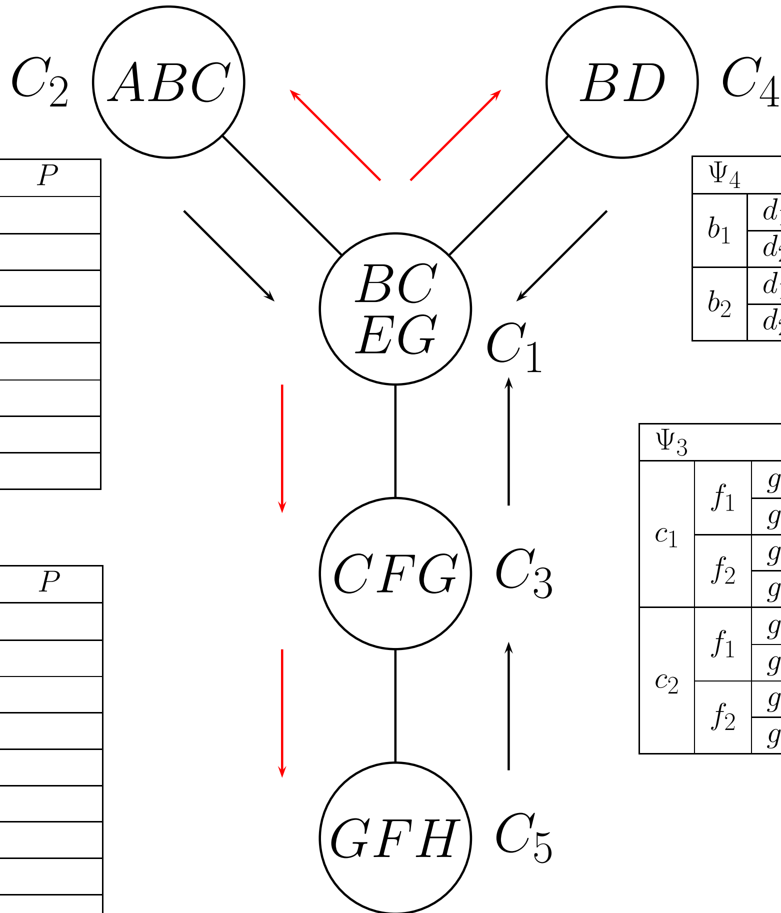
$$M_{31} = \begin{array}{c} \quad c_1,g_1 \;\; c_1,g_2 \;\; c_2,g_1 \;\; c_2,g_2 \\ \left(0.38, 0.68, 0.32, 0.62\right) \end{array}$$

Ψ_3				P
c_1	f_1	g_1	0.1	
		g_2	0.1	
	f_2	g_1	0.9	
		g_2	0.9	
c_2	f_1	g_1	0.4	
		g_2	0.4	
	f_2	g_1	0.6	
		g_2	0.6	

Ψ_1					P
b_1	c_1	e_1	g_1	0.190	
			g_2	0.010	
		e_2	g_1	0.320	
			g_2	0.480	
	c_2	e_1	g_1	0.380	
			g_2	0.020	
		e_2	g_1	0.240	
			g_2	0.360	
b_2	c_1	e_1	g_1	0.210	
			g_2	0.090	
		e_2	g_1	0.350	
			g_2	0.350	
	c_2	e_1	g_1	0.070	
			g_2	0.030	
		e_2	g_1	0.450	
			g_2	0.450	

Ψ_5				P
f_1	g_1	h_1	0.2	
		h_2	0	
	g_2	h_1	0.5	
		h_2	0	
f_2	g_1	h_1	0.4	
		h_2	0	
	g_2	h_1	0.7	
		h_2	0	

Example: Step 4: Evidence $H = h_1$ (Sending Messages)



Ψ_2				P
a_1	b_1	c_1	0.036	
		c_2	0.084	
	b_2	c_1	0.144	
		c_2	0.336	
a_2	b_1	c_1	0.028	
		c_2	0.012	
	b_2	c_1	0.252	
		c_2	0.108	

Ψ_4			P
b_1	d_1	0.4	
	d_2	0.6	
b_2	d_1	0.7	
	d_2	0.3	

Ψ_3				P
c_1	f_1	g_1	0.1	
		g_2	0.1	
	f_2	g_1	0.9	
		g_2	0.9	
c_2	f_1	g_1	0.4	
		g_2	0.4	
	f_2	g_1	0.6	
		g_2	0.6	

Ψ_1					P
b_1	c_1	e_1	g_1	0.190	
			g_2	0.010	
		e_2	g_1	0.320	
			g_2	0.480	
	c_2	e_1	g_1	0.380	
			g_2	0.020	
		e_2	g_1	0.240	
			g_2	0.360	
b_2	c_1	e_1	g_1	0.210	
			g_2	0.090	
		e_2	g_1	0.350	
			g_2	0.350	
	c_2	e_1	g_1	0.070	
			g_2	0.030	
		e_2	g_1	0.450	
			g_2	0.450	

Ψ_5				P
f_1	g_1	h_1	0.2	
		h_2	0	
	g_2	h_1	0.5	
		h_2	0	
f_2	g_1	h_1	0.4	
		h_2	0	
	g_2	h_1	0.7	
		h_2	0	

$$M_{53} = \begin{array}{cccc} f_1,g_1 & f_1,g_2 & f_2,g_1 & f_2,g_2 \\ \big(\,0.2\,, & 0.5\,, & 0.4\,, & 0.7\,\big) \end{array}$$

$$M_{21} = \begin{array}{cccc} b_1,c_1 & b_1,c_2 & b_2,c_1 & b_2,c_2 \\ \big(0.06, & 0.10, & 0.40, & 0.44\big) \end{array}$$

$$M_{41} = \begin{array}{cc} b_1 & b_2 \\ \big(1, & 1\big) \end{array}$$

$$M_{31} = \begin{array}{cccc} c_1,g_1 & c_1,g_2 & c_2,g_1 & c_2,g_2 \\ \big(0.38, & 0.68, & 0.32, & 0.62\big) \end{array}$$

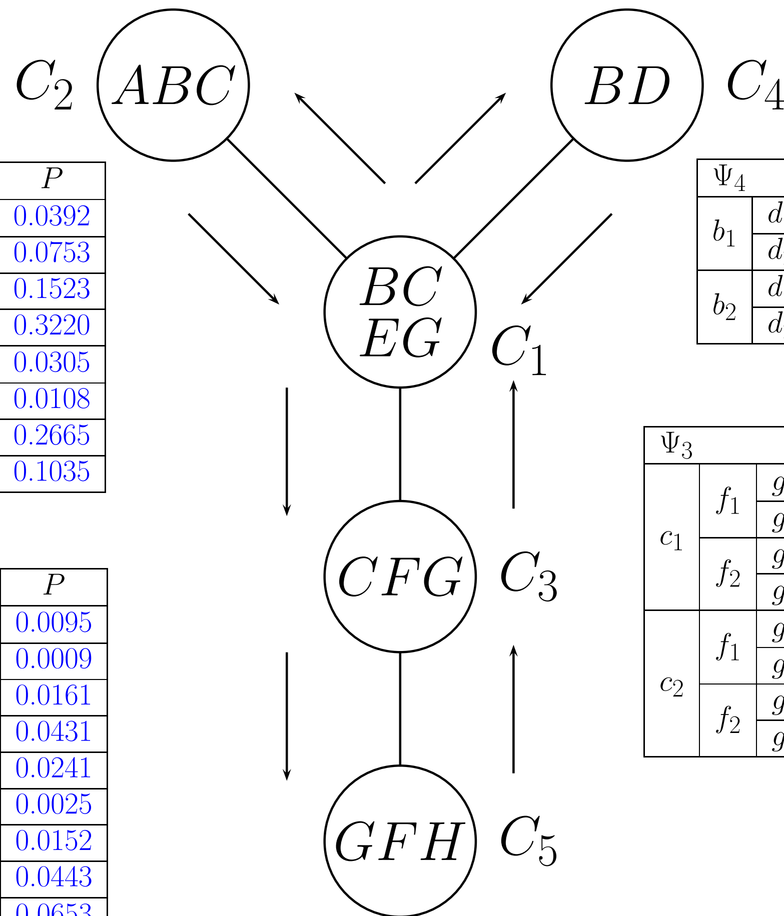
$$M_{12} = \begin{array}{cccc} b_1,c_1 & b_1,c_2 & b_2,c_1 & b_2,c_2 \\ \big(0.527, & 0.434, & 0.512, & 0.464\big) \end{array}$$

$$M_{14} = \begin{array}{cc} b_1 & b_2 \\ \big(0.075, & 0.409\big) \end{array}$$

$$M_{13} = \begin{array}{cccc} c_1,g_1 & c_1,g_2 & c_2,g_1 & c_2,g_2 \\ \big(0.254, & 0.206, & 0.290, & 0.250\big) \end{array}$$

$$M_{35} = \begin{array}{cccc} f_1,g_1 & f_1,g_2 & f_2,g_1 & f_2,g_2 \\ \big(0.14, & 0.12, & 0.40, & 0.33\big) \end{array}$$

Example: Step 4: Evidence $H = h_1$ Incorporated



Ψ_2				P
a_1	b_1	c_1	0.036	0.0392
		c_2	0.084	0.0753
	b_2	c_1	0.144	0.1523
		c_2	0.336	0.3220
a_2	b_1	c_1	0.028	0.0305
		c_2	0.012	0.0108
	b_2	c_1	0.252	0.2665
		c_2	0.108	0.1035

Ψ_4			P
b_1	d_1	0.4	0.0623
	d_2	0.6	0.0934
b_2	d_1	0.7	0.5910
	d_2	0.3	0.2533

Ψ_1					P
b_1	c_1	e_1	g_1	0.190	0.0095
			g_2	0.010	0.0009
		e_2	g_1	0.320	0.0161
			g_2	0.480	0.0431
	c_2	e_1	g_1	0.380	0.0241
			g_2	0.020	0.0025
		e_2	g_1	0.240	0.0152
			g_2	0.360	0.0443
b_2	c_1	e_1	g_1	0.210	0.0653
			g_2	0.090	0.0501
		e_2	g_1	0.350	0.1088
			g_2	0.350	0.1947
	c_2	e_1	g_1	0.070	0.0205
			g_2	0.030	0.0171
		e_2	g_1	0.450	0.1321
			g_2	0.450	0.2559

Ψ_3				P
c_1	f_1	g_1	0.1	0.0105
		g_2	0.1	0.0212
	f_2	g_1	0.9	0.1892
		g_2	0.9	0.2675
c_2	f_1	g_1	0.4	0.0480
		g_2	0.4	0.1031
	f_2	g_1	0.6	0.1440
		g_2	0.6	0.2165

Ψ_5				P
f_1	g_1	h_1	0.2	0.0585
		h_2	0	0
	g_2	h_1	0.5	0.1243
		h_2	0	0
f_2	g_1	h_1	0.4	0.3331
		h_2	0	0
	g_2	h_1	0.7	0.4841
		h_2	0	0

$$M_{53} = \begin{array}{c} f_1,g_1 \quad f_1,g_2 \quad f_2,g_1 \quad f_2,g_2 \\ (\,0.2\,,\ 0.5\,,\ 0.4\,,\ 0.7\,) \end{array}$$

$$M_{21} = \begin{array}{c} b_1,c_1 \quad b_1,c_2 \quad b_2,c_1 \quad b_2,c_2 \\ (0.06, 0.10, 0.40, 0.44) \end{array}$$

$$M_{41} = \begin{array}{c} b_1 \quad b_2 \\ (1\,,\ 1) \end{array}$$

$$M_{31} = \begin{array}{c} c_1,g_1 \quad c_1,g_2 \quad c_2,g_1 \quad c_2,g_2 \\ (0.38, 0.68, 0.32, 0.62) \end{array}$$

$$M_{12} = \begin{array}{c} b_1,c_1 \quad b_1,c_2 \quad b_2,c_1 \quad b_2,c_2 \\ (0.527, 0.434, 0.512, 0.464) \end{array}$$

$$M_{14} = \begin{array}{c} b_1 \quad b_2 \\ (0.075, 0.409) \end{array}$$

$$M_{13} = \begin{array}{c} c_1,g_1 \quad c_1,g_2 \quad c_2,g_1 \quad c_2,g_2 \\ (0.254, 0.206, 0.290, 0.250) \end{array}$$

$$M_{35} = \begin{array}{c} f_1,g_1 \quad f_1,g_2 \quad f_2,g_1 \quad f_2,g_2 \\ (0.14, 0.12, 0.40, 0.33) \end{array}$$

P	A	B	C	D	E	F	G	H
$\cdot_1$	**0.5888**	**0.1557**	**0.4884**	**0.6533**	**0.1899**	**0.1828**	**0.3916**	**1.0000**
$\cdot_2$	**0.4112**	**0.8443**	**0.5116**	**0.3467**	**0.8101**	**0.8172**	**0.6084**	**0.0000**

Building Graphical Models: Causal Modeling

Manual creation of a reasoning system based on a graphical model:

causal model of given domain

↓ **heuristics!**

conditional independence graph

↓ formally provable

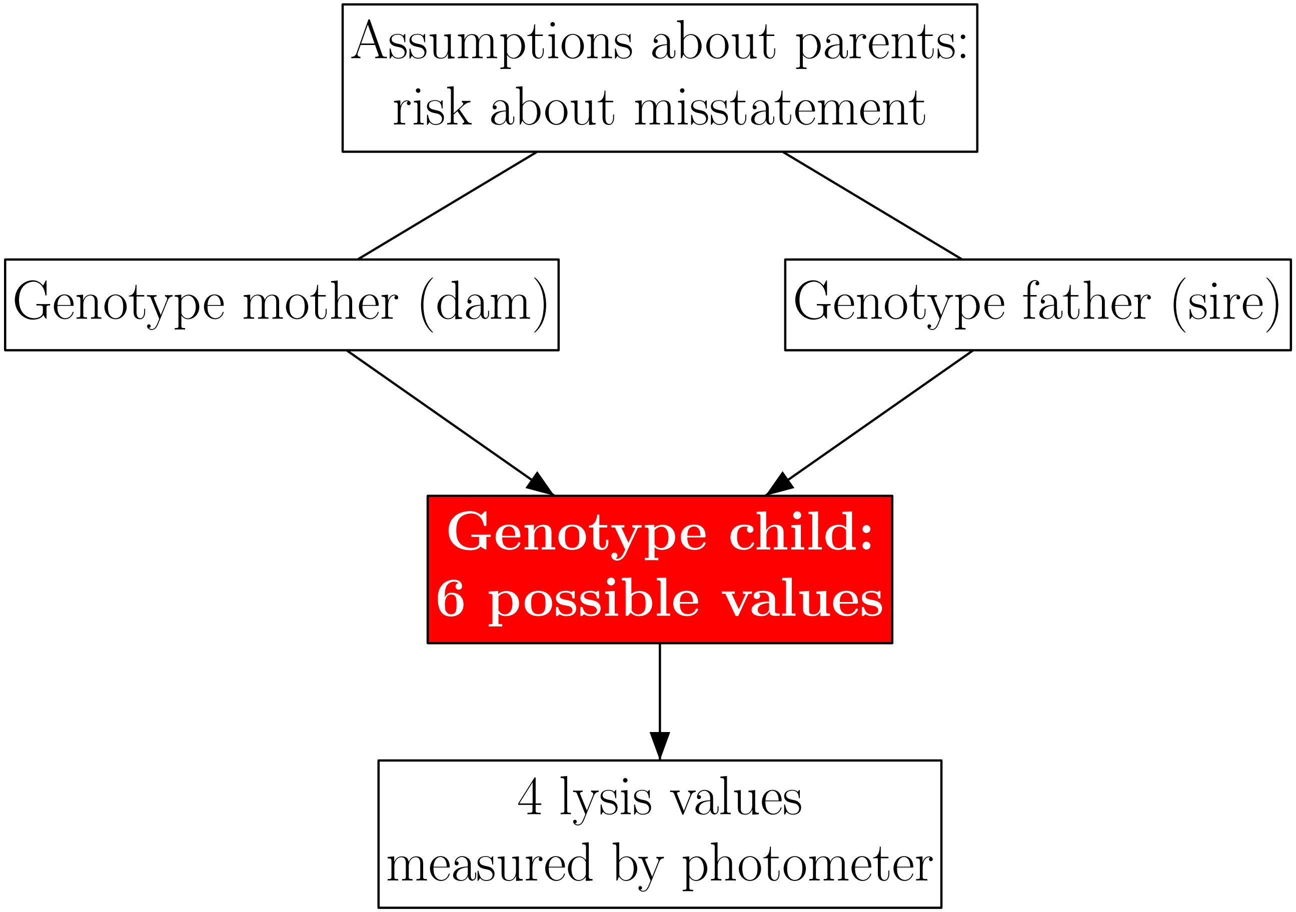
decomposition of the distribution

↓ formally provable

evidence propagation scheme

- Problem: strong assumptions about the statistical effects of causal relations.
- Nevertheless this approach often yields usable graphical models.

Example 1: Genotype Determination of Danish Jersey Cattle

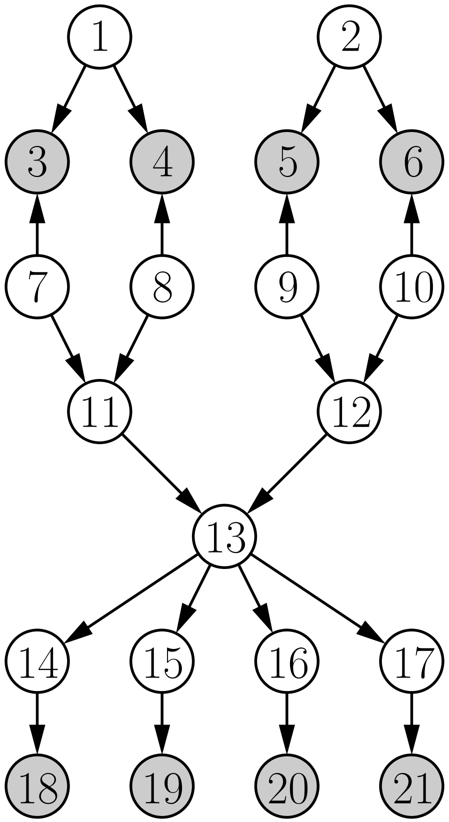


Reliability of databases

Inheritance rules

Blood group determination

Danish Jersey Cattle Blood Type Determination



21 attributes:
1 – dam correct?
2 – sire correct?
3 – stated dam ph.gr. 1
4 – stated dam ph.gr. 2
5 – stated sire ph.gr. 1
6 – stated sire ph.gr. 2
7 – true dam ph.gr. 1
8 – true dam ph.gr. 2
9 – true sire ph.gr. 1
10 – true sire ph.gr. 2
11 – offspring ph.gr. 1
12 – offspring ph.gr. 2
13 – offspring genotype
14 – factor 40
15 – factor 41
16 – factor 42
17 – factor 43
18 – lysis 40
19 – lysis 41
20 – lysis 42
21 – lysis 43

The grey nodes correspond to observable attributes.

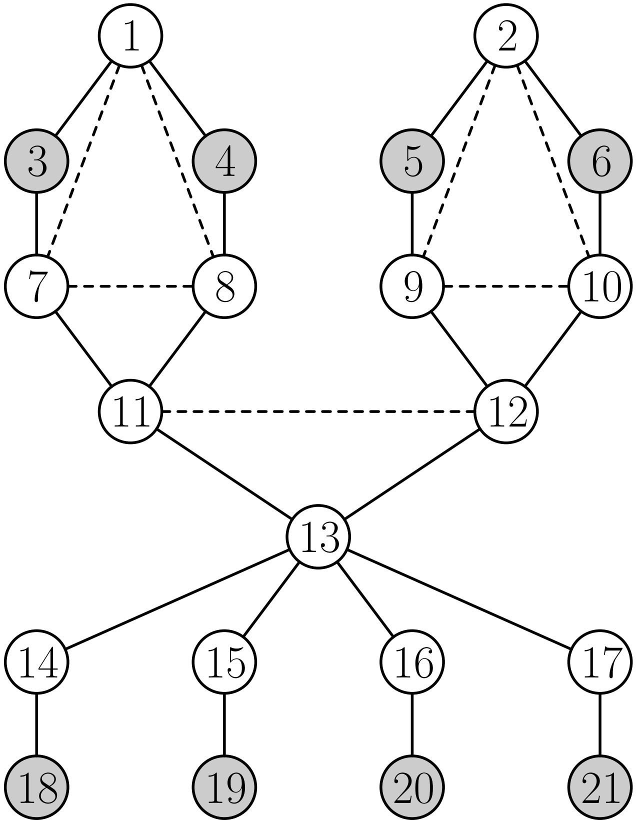
- This graph was specified by human domain experts, based on knowledge about (causal) dependences of the variables.

Example 1: Genotype Determination of Danish Jersey Cattle

- Full 21-dimensional domain has $2^6 \cdot 3^{10} \cdot 6 \cdot 8^4 = 92\,876\,046\,336$ possible states.
- Bayesian network requires only 306 conditional probabilities.
- Example of a conditional probability table (attributes 2, 9, and 5):

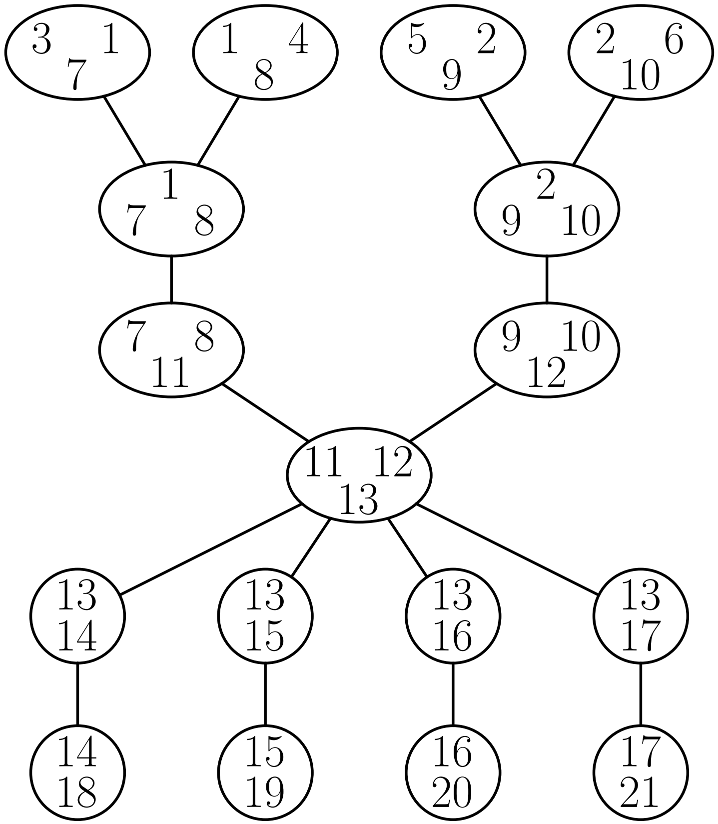
sire correct	true sire phenogroup 1	stated sire phenogroup 1 F1	V1	V2
yes	F1	1	0	0
yes	V1	0	1	0
yes	V2	0	0	1
no	F1	0.58	0.10	0.32
no	V1	0.58	0.10	0.32
no	V2	0.58	0.10	0.32

- The probabilities are acquired from human domain experts or estimated from historical data.



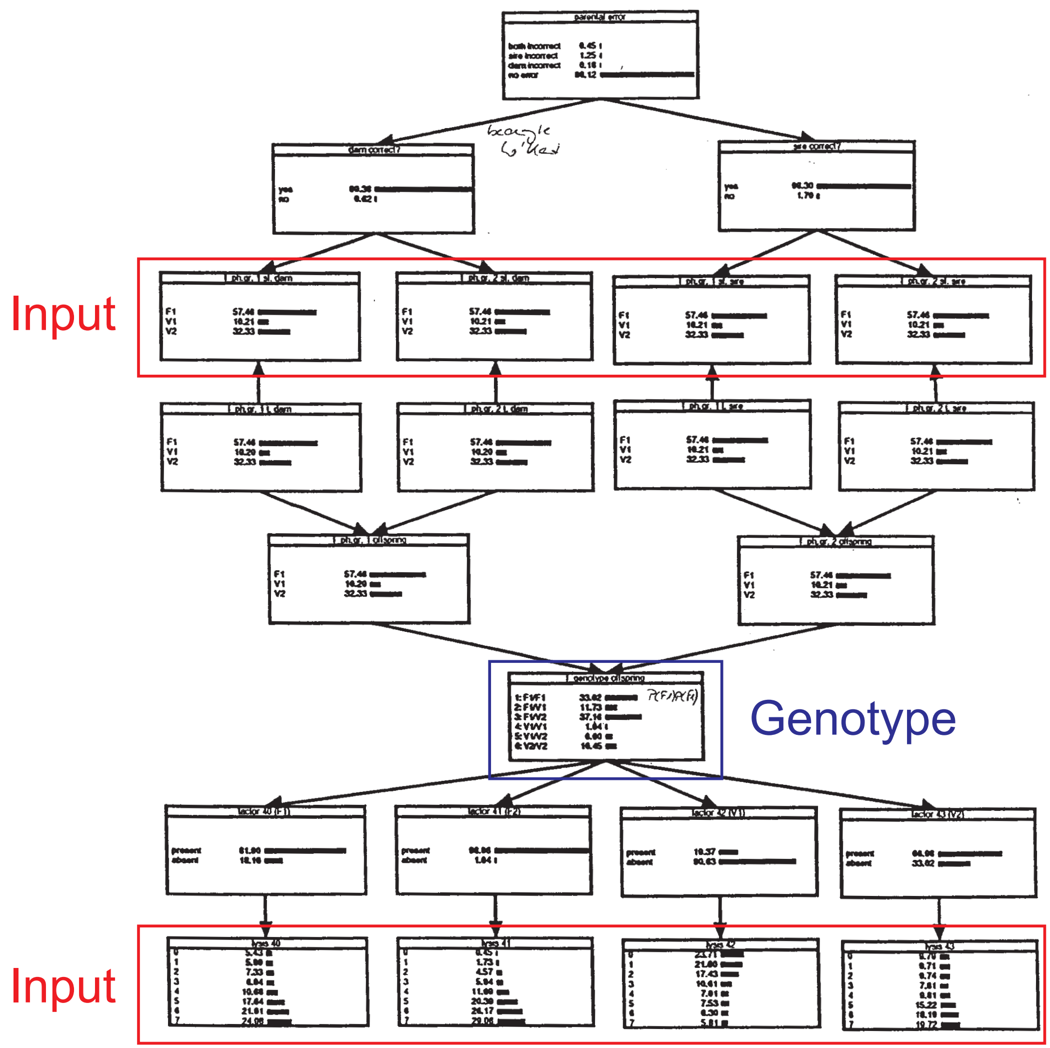
moral graph

(already triangulated)

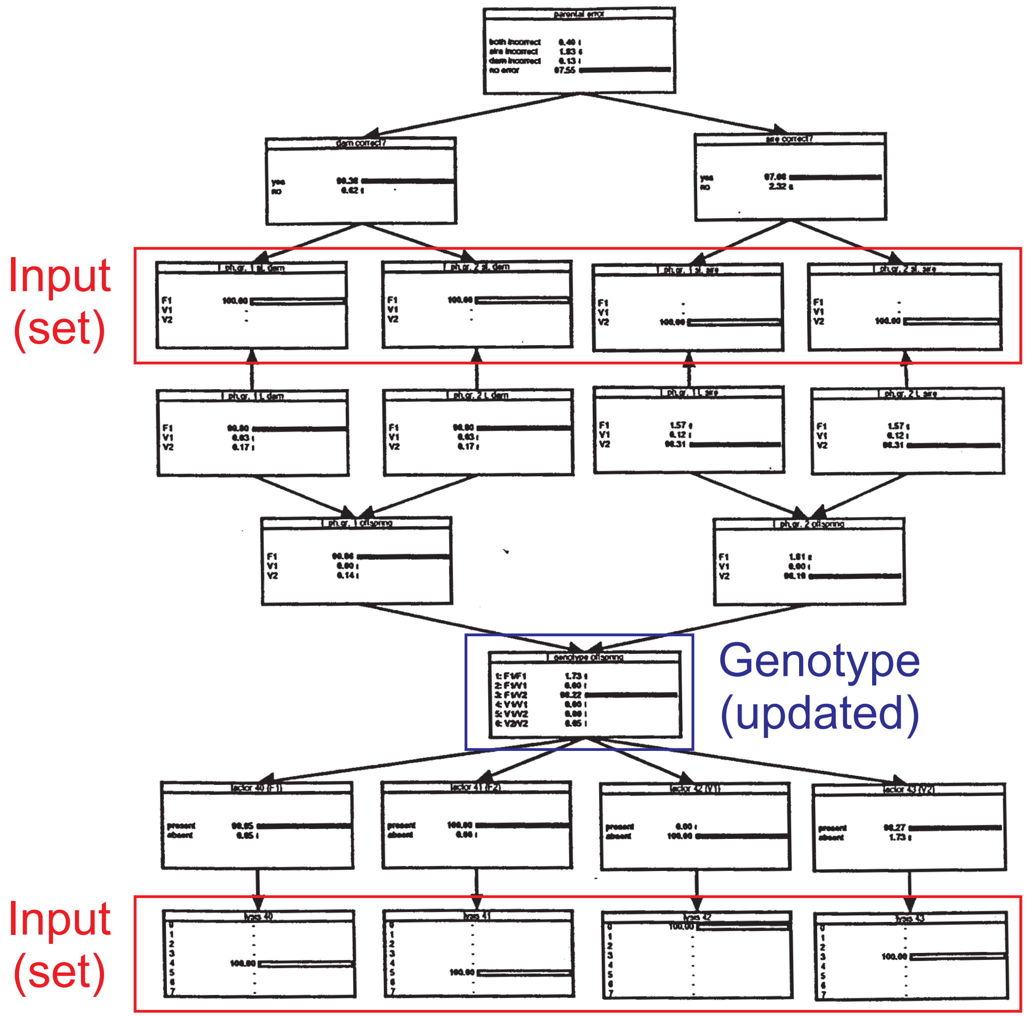


join tree

Marginal distributions before setting evidence:



Conditional distributions given evidence in the input variables:



Example 2: Item Planning at Volkswagen

Strategy of the VW Group

Marketing strategy	Vehicle specification by clients	Bestsellers defined by manufacturer
Complexity	Huge number of variants	Small number of variants

Vehicle specification

Equipment	fastback	2,8 l, 150 kW	Type Alpha	4	leather	...
Group	car body type	engine	radio	doors	seat cover	...

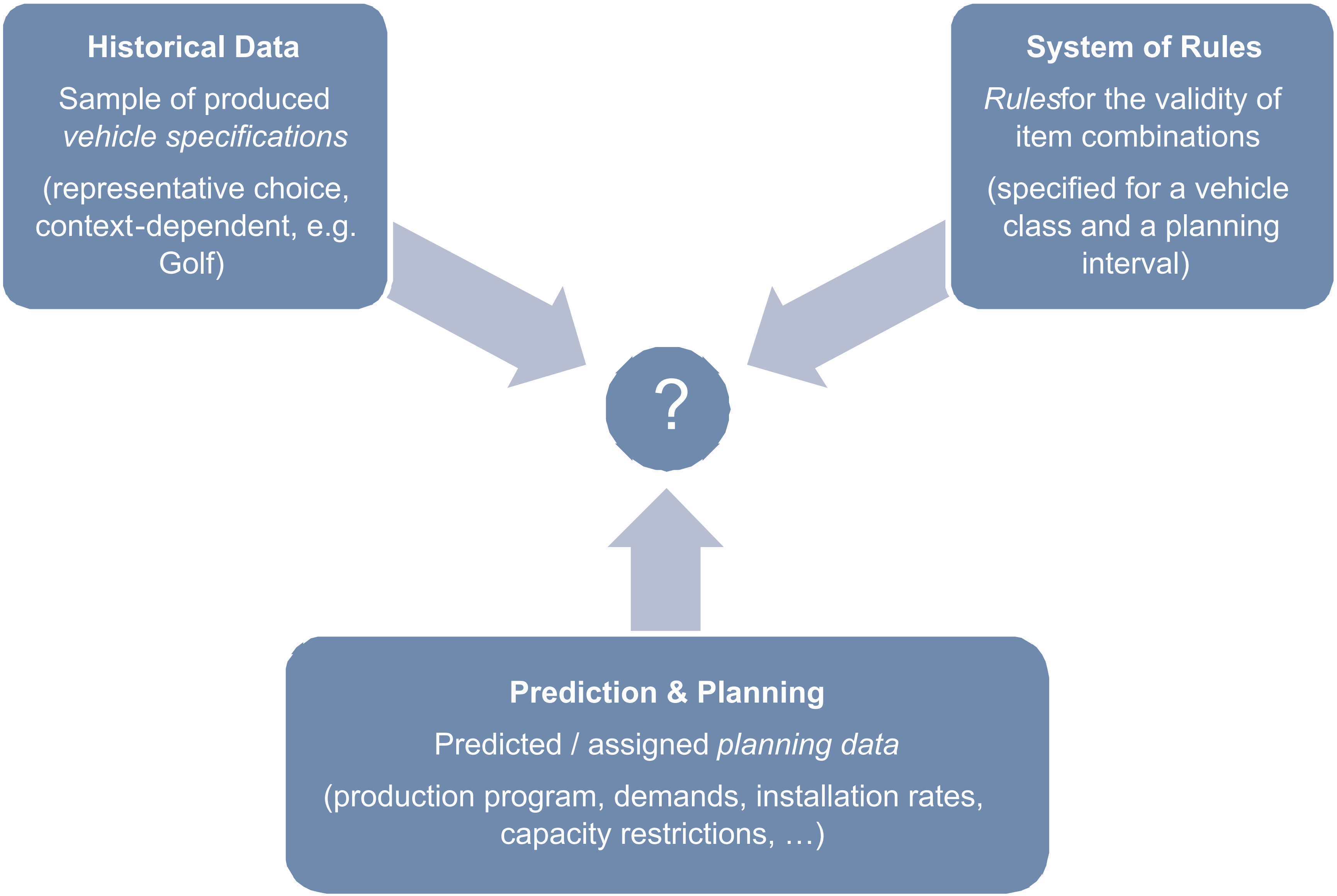
Example 2: Model “Golf”

- Approx. 200 equipment groups
- 2 to 50 items per group
- Therefore more than 2^{200} possible vehicle specifications
- Choice of valid specifications is constrained by a rule system (10000 technical rules, plus marketing and production rules)

Example of technical rules:

- **If** Engine=e_1 **then** Transmission=t_3
- **If** Engine=e_4 and Heating=h_2 **then** Generator $\in \{g_3, g_4, g_5\}$

Problem Representation



Complexity of the Planning Problem

Equipment table

	Engine	Transmission	Heating	Generator	$\cdots$
1	e_1	t_3	h_1	g_1	$\cdots$
2	e_2	t_4	h_3	g_5	$\cdots$
	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
100000	e_7	t_1	h_3	g_2	$\cdots$

Installation rates

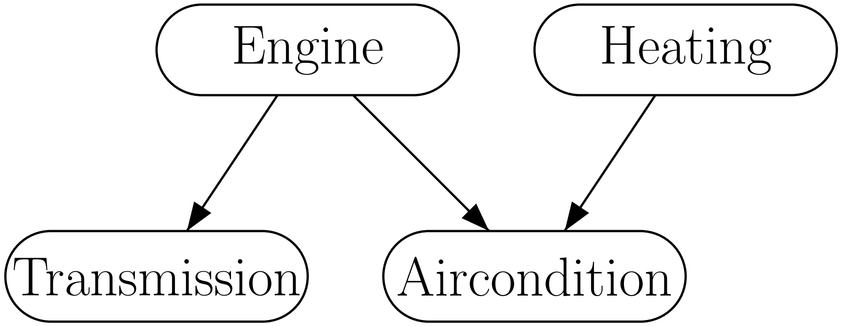
Engine	Transmission	Heating	Generator	$\cdots$	Rate
e_1	t_1	h_1	g_1	$\cdots$	0.0000012
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$

Result is a 200-dimensional, finite probability space

- $P(\text{Engine} = e_1, \text{Transmission} = t_3) = ?$
- $P(\text{Heating} = h_1 \mid \text{Generator} = g_3) = ?$

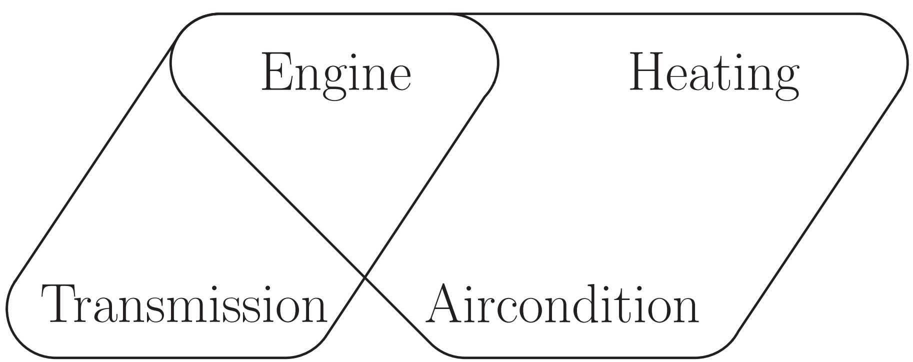
Problem of complexity!

Solution: Decomposition into Subspaces



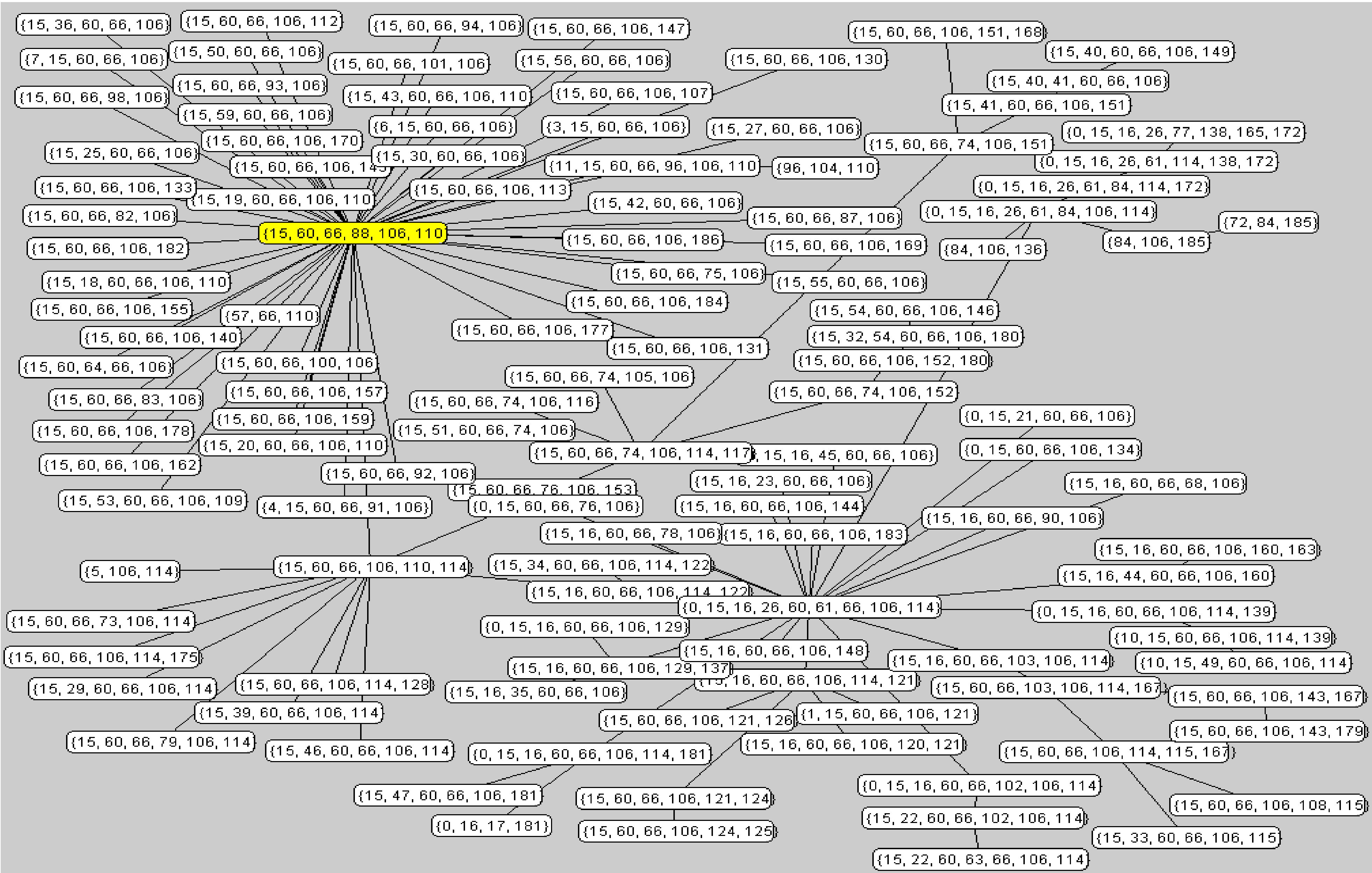
Bayesian Network

$$P(E,H,T,A) = P(A \mid E,H,T) \cdot P(T \mid E,H) \cdot P(E \mid H) \cdot P(H)$$
$$\stackrel{\text{here}}{=} P(A \mid E,H) \cdot P(T \mid E) \cdot P(E) \cdot P(H)$$



Hypergraph Decomposition

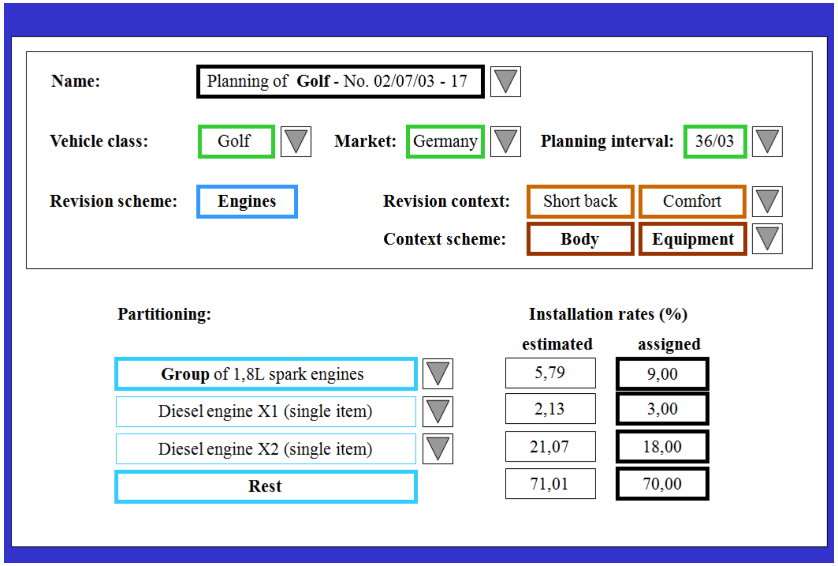
Clique Tree of the VW Bora



- **Application:**
  - **Compute item demand**
    Calculation of installation rates of equipment combinations
  - **Simulation**
    Analyze customer requirements (e. g. of persons having ordered a navigation system for a VW Polo)

- **Input:** Equipment combinations

- **Operation:** Compute
  - the conditional network distribution and
  - the probabilities of the specified equipment combinations.



Implementation and Deployment

- Projec leader: Jörg Gebhardt
- Client server system
- Server on 6–8 maschines
- Quadcore platform
- Terabyte hard drive
- Java, Linux, Oracle
- WebSphere application server
- Software used daily worldwide
- 15 developers
- 4000 Bayesian networks are currently used