

# Propagation in Clique Trees

# Potential Representation

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

$$\psi_i : \prod_{X_j \in W_i} \text{dom}(X_j) \rightarrow \mathbb{R}$$

It is said that  $P(V)$  *factorizes* according  $(\{W_1, \dots, W_p\}, \{\psi_1, \dots, \psi_p\})$  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

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

$$\text{dom}(A) = \{a_1, a_2\}$$

$$\text{dom}(B) = \{b_1, b_2\}$$

$$\text{dom}(C) = \{c_1, c_2\}$$

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

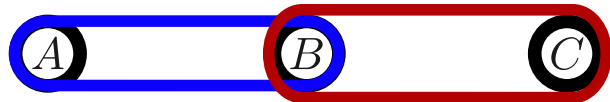
$$\psi_1 : \{a_1, a_2\} \times \{b_1, b_2\} \rightarrow \mathbb{R}$$

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

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

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

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



# Factorization of a Belief Network

Let  $(V, E, P)$  be a belief network and  $\{C_1, \dots, C_p\}$  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  $(\{C_1, \dots, C_p\}, \{\psi_1, \dots, \psi_p\})$  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  $\{C_1, \dots, C_p\}$  be a set of cliques w. r. t.  $V$ . The sets

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

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

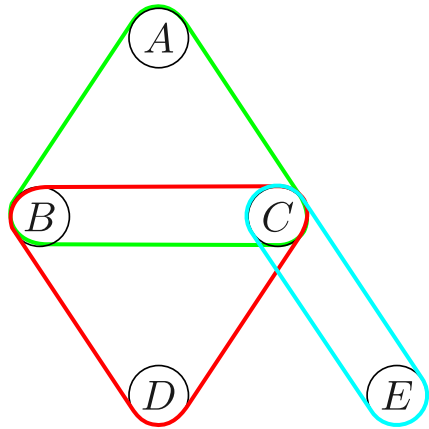
$$R_i = C_i \setminus S_i$$

# Decomposition w. r. t. a Join-Tree

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

$$R_i \perp\!\!\!\perp (C_1 \cup \dots \cup C_{i-1}) \setminus S_i \mid S_i \quad \text{for } i > 1$$

# Example



$$S_1 = \emptyset$$

$$R_1 = \{A, B, C\}$$

$$f(A) = C_1$$

$$S_2 = \{B, C\}$$

$$R_2 = \{D\}$$

$$f(B) = C_1$$

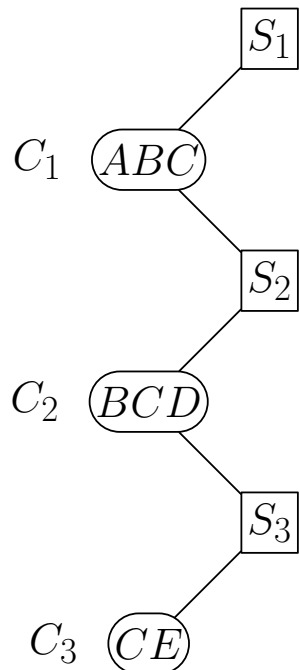
$$S_3 = \{C\}$$

$$R_3 = \{E\}$$

$$f(C) = C_1$$

$$f(D) = C_2$$

$$f(E) = C_3$$



$$\psi_1(C_1) = P(A) \cdot P(C | A) \cdot P(B|A)$$

$$\psi_2(C_2) = P(D | B, C)$$

$$\psi_3(C_3) = P(E | C)$$

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

# Propagation in Join Trees

## Main Idea

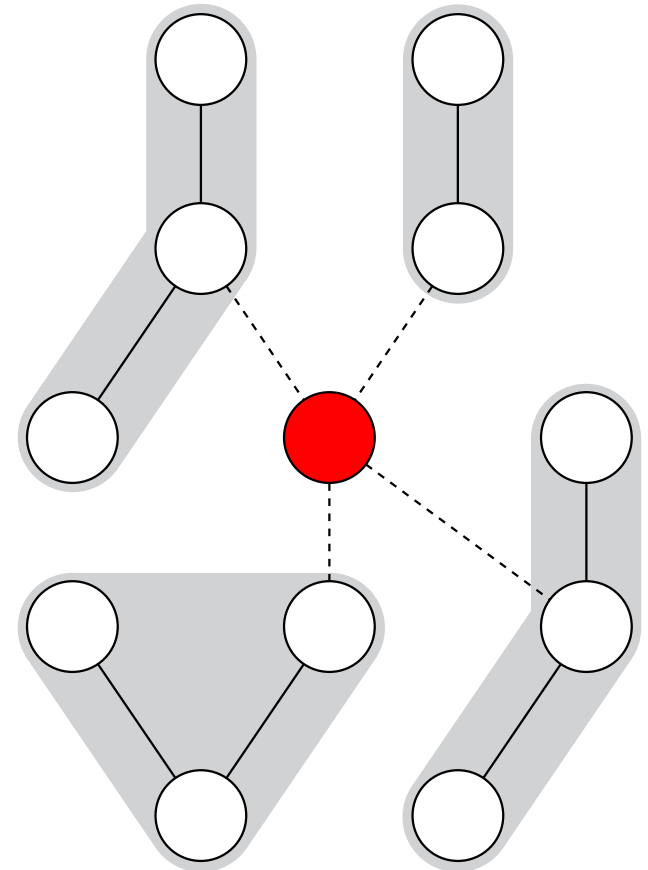
Incorporate evidence into the clique potentials.

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  $\psi_i$ .

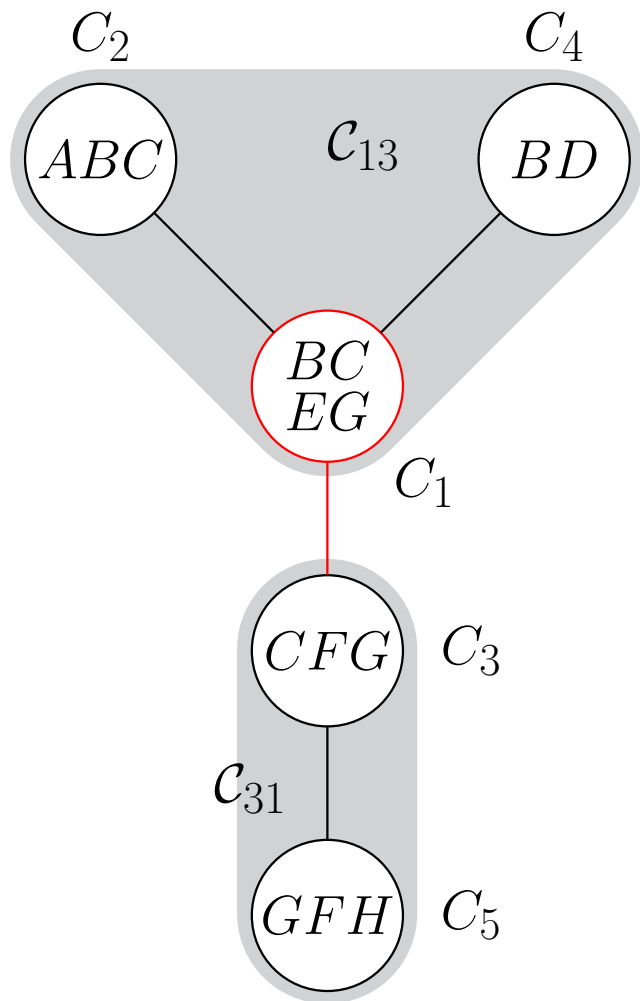
If for an attribute  $E$  some evidence  $e$  becomes known, we alter all potential functions of cliques containing  $E$  as follows:

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

All other potential functions are unchanged.



# Notation and Nomenclature



In general:

Clique  $C_i$  has  $q$  neighboring cliques  $B_1, \dots, B_q$ .

$C_{ij}$  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  $C_{ij}$ .

$V = X_{ij} \cup X_{ji}$  (complementary sets)

$S_{ij} = S_{ji} = C_i \cap C_j$  (not shown here)

$R_{ij} = X_{ij} \setminus S_{ij}$  (not shown here)

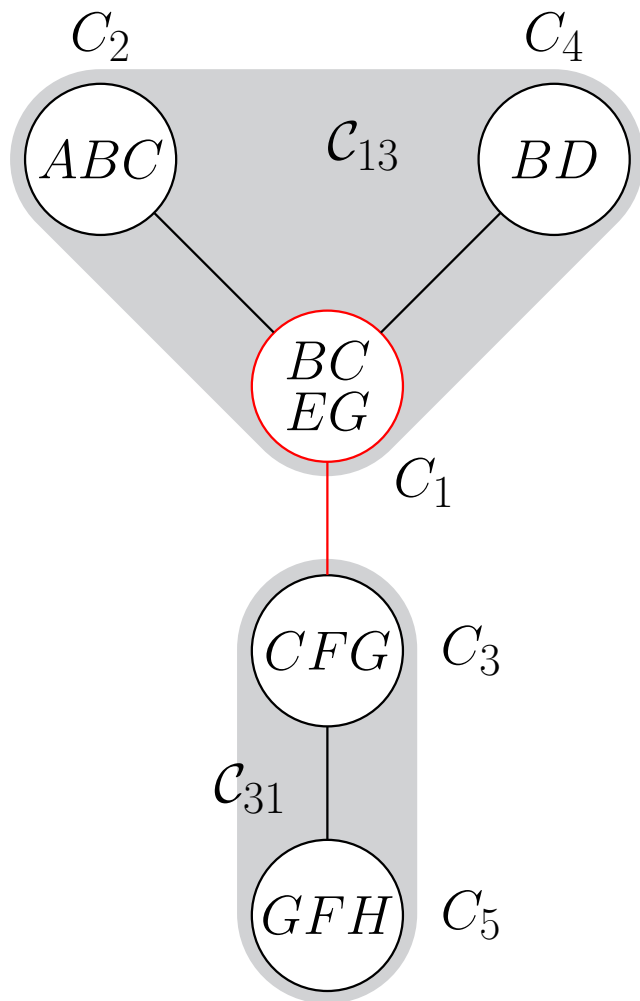
Here:

Neighbors of  $C_1$ :  $\{C_2, C_4, C_3\}$ ,  $C_{13} = \{C_1, C_2, C_4\}$

$X_{13} = \{A, B, C, D, E, G\}$ ,  $S_{13} = \{C, G\}$

$V = X_{13} \cup X_{31} = \{A, B, C, D, E, F, G, H\}$

$R_{13} = \{A, B, D, E\}$ ,  $R_{31} = \{F, H\}$



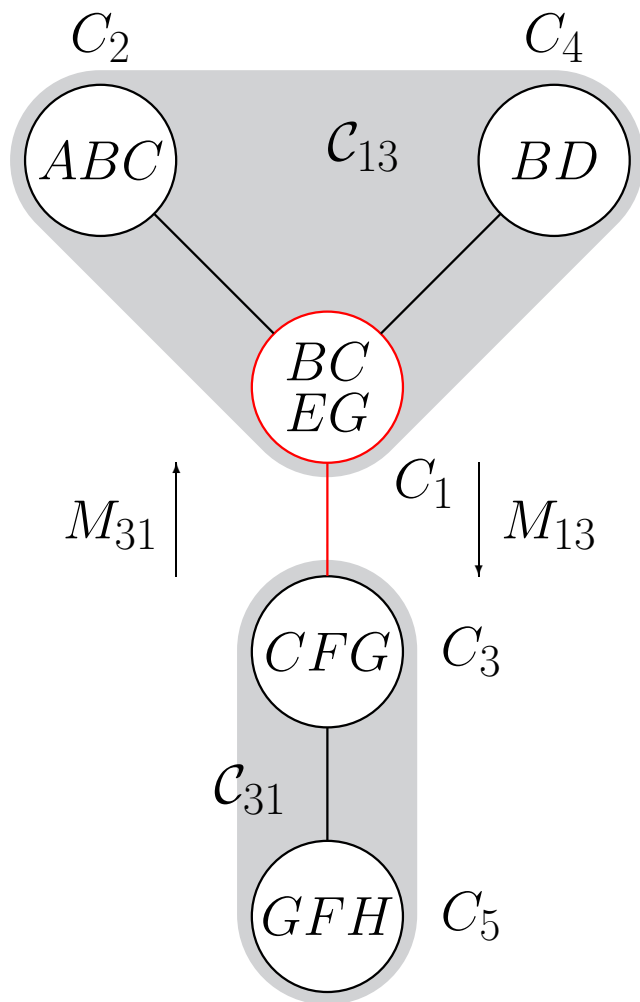
**Task:** Calculate  $P(s_{ij})$ :

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

$$\begin{aligned}
 V \setminus S_{13} &= (X_{13} \cup X_{31}) \setminus S_{13} \\
 &= R_{13} \cup R_{31}
 \end{aligned}$$

$$\begin{aligned}
 V \setminus \{C, G\} &= \{A, B, D, E\} \cup \{F, H\} \\
 &= \{A, B, D, E, F, H\}
 \end{aligned}$$

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  $P(s_{ij})$ :

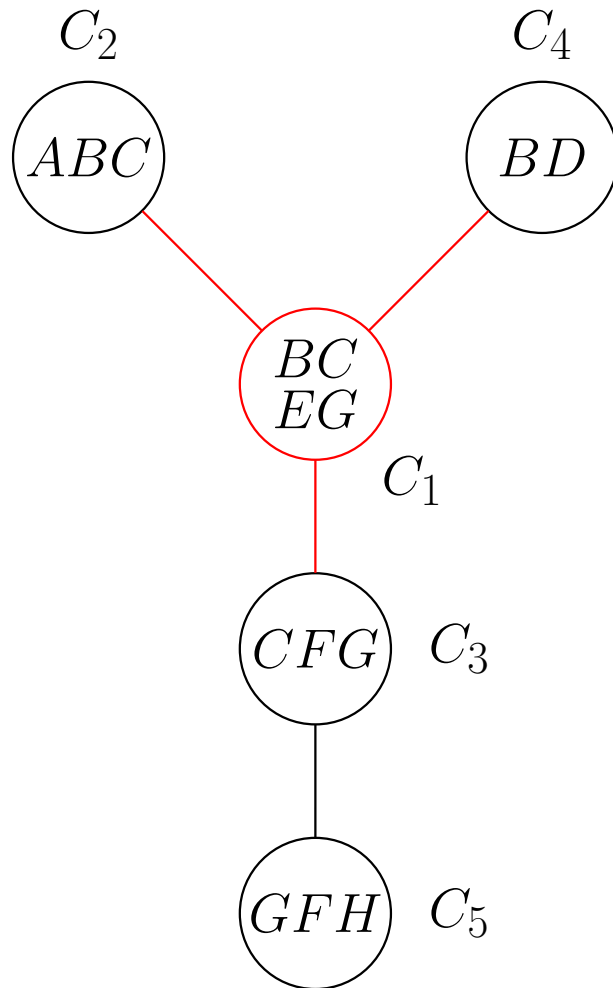
$$P(s_{ij}) = \sum_{v \setminus s_{ij}} \prod_{k=1}^m \psi_k(c_k)$$

$$\stackrel{\text{last slide}}{=} \sum_{r_{ij} \cup r_{ji}} \prod_{k=1}^m \psi_k(c_k)$$

$$\stackrel{\text{sum rule}}{=} \left( \sum_{r_{ij}} \prod_{c_k \in C_{ij}} \psi_k(c_k) \right) \cdot \left( \sum_{r_{ji}} \prod_{c_k \in C_{ji}} \psi_k(c_k) \right)$$

$$= M_{ij}(s_{ij}) \cdot M_{ji}(s_{ij})$$

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

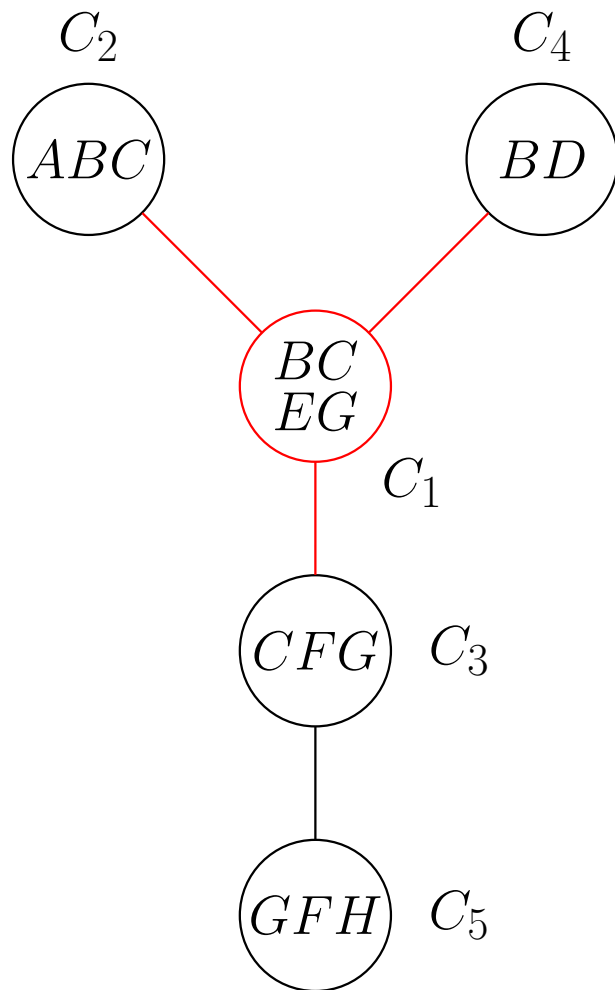


**Task:** Calculate  $P(c_i)$ :

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

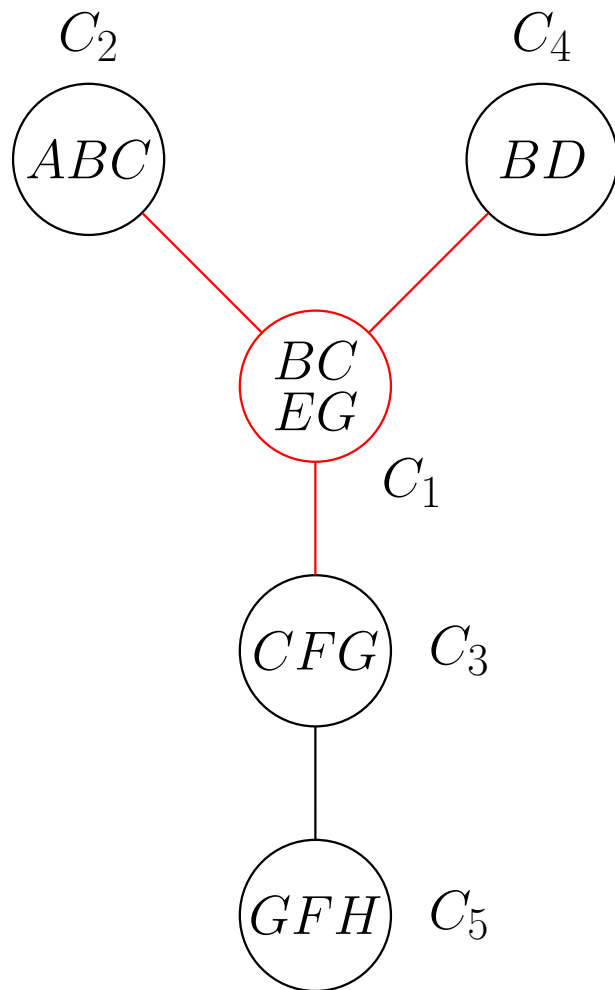
Example:

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



**Task:** Calculate  $P(c_i)$ :

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



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:

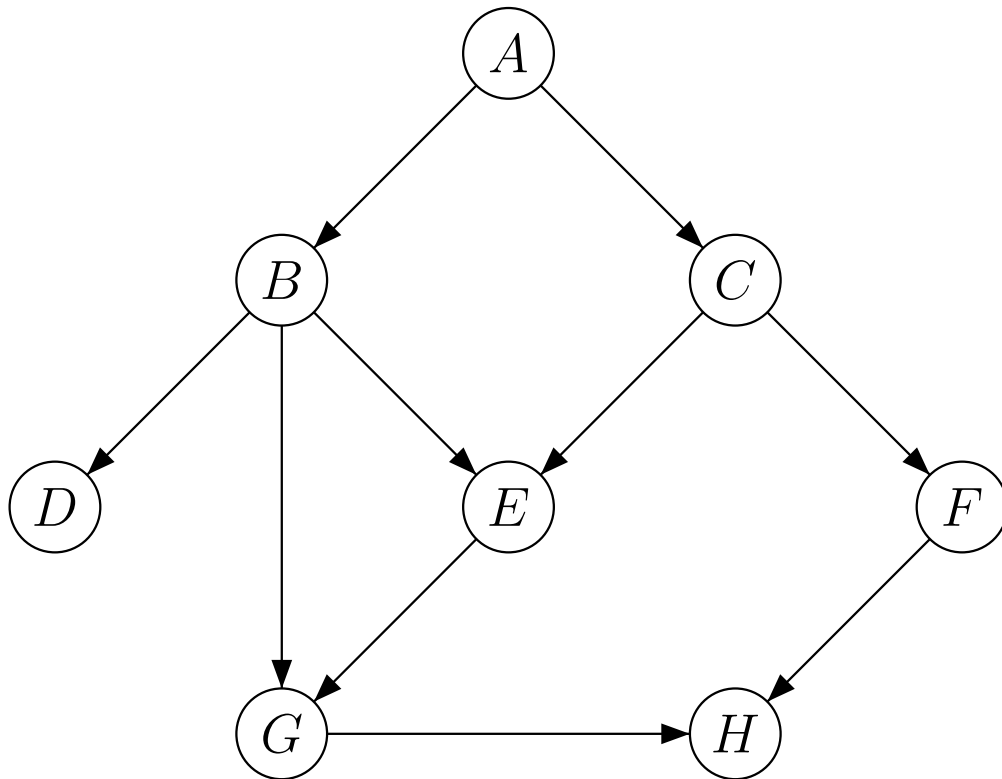
$$P(c_i) \propto \psi_i(c_i) \prod_{j=1}^q M_{ji}(s_{ij})$$

4. For every  $X \in V$  calculate the a-posteriori probability:

$$P(x_i | e) = \sum_{c_k \setminus x_i} P(c_k)$$

where  $C_k$  is the smallest clique containing  $X_i$ .

# Example: Putting it together



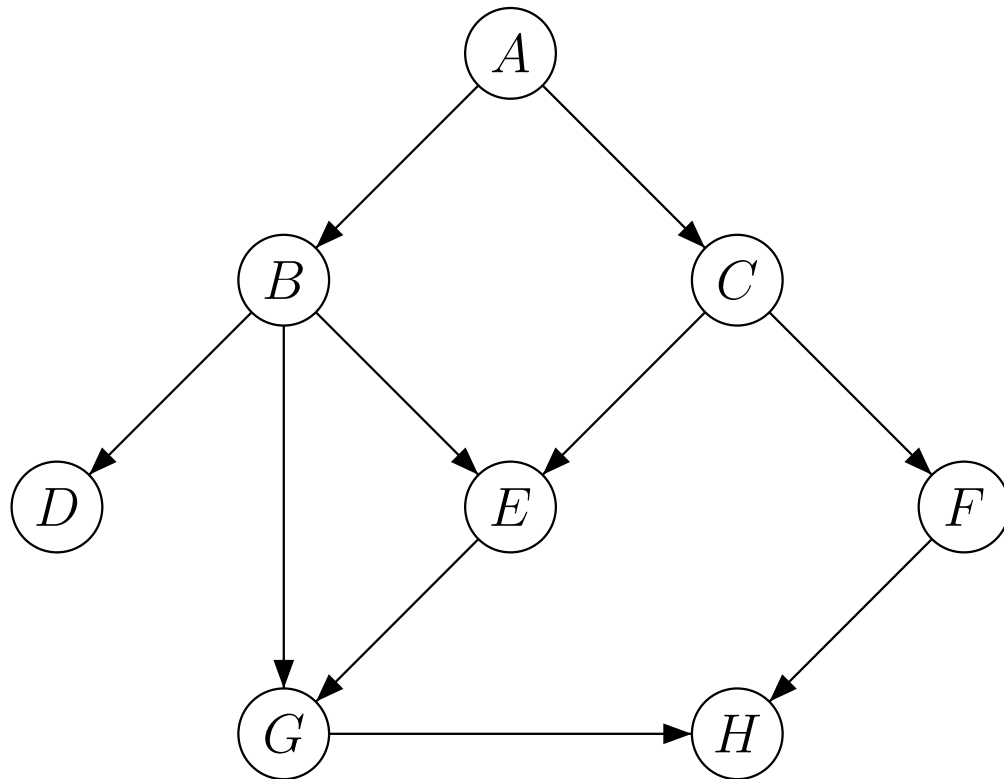
**Goals:** Find the marginal distributions and update them when evidence  $H = h_1$  becomes known.

**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.

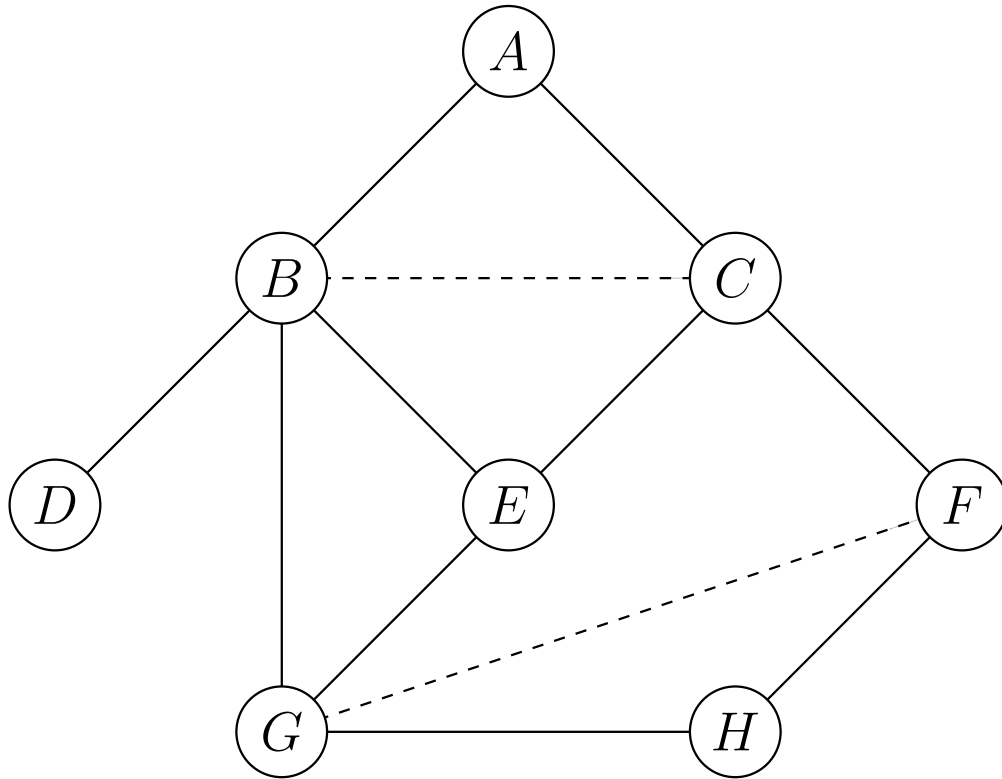


# Example: Step 1: Find a Join-Tree



**Join-Tree creation:**

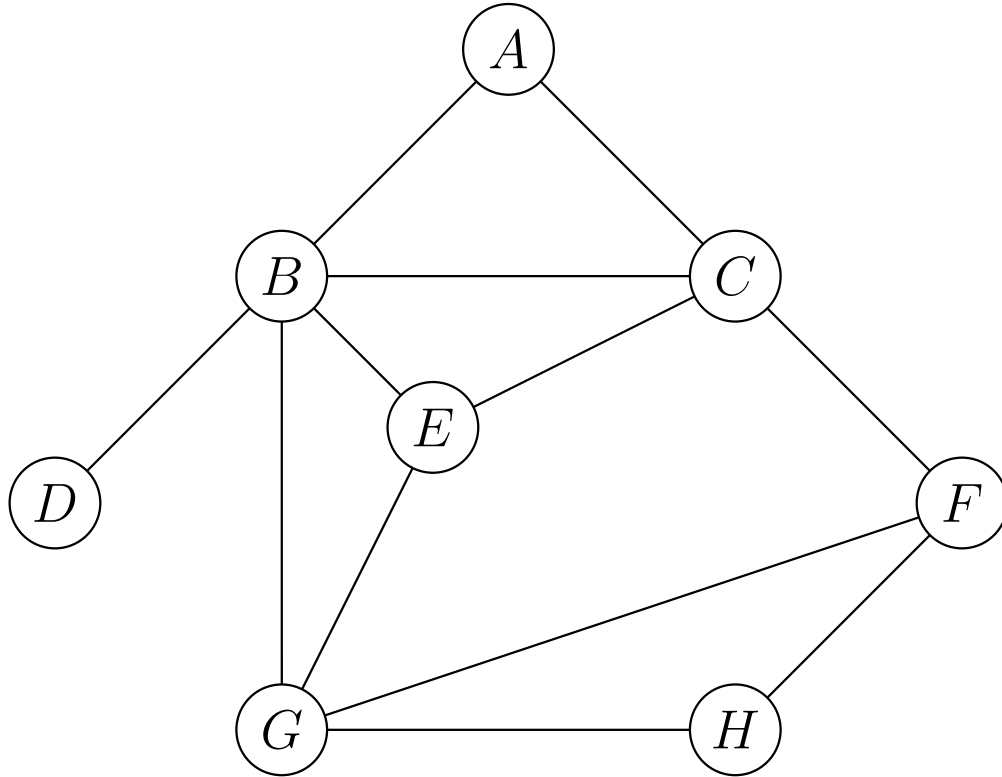
# Example: Step 1: Find a Join-Tree



## Join-Tree creation:

1. Moralize the graph.

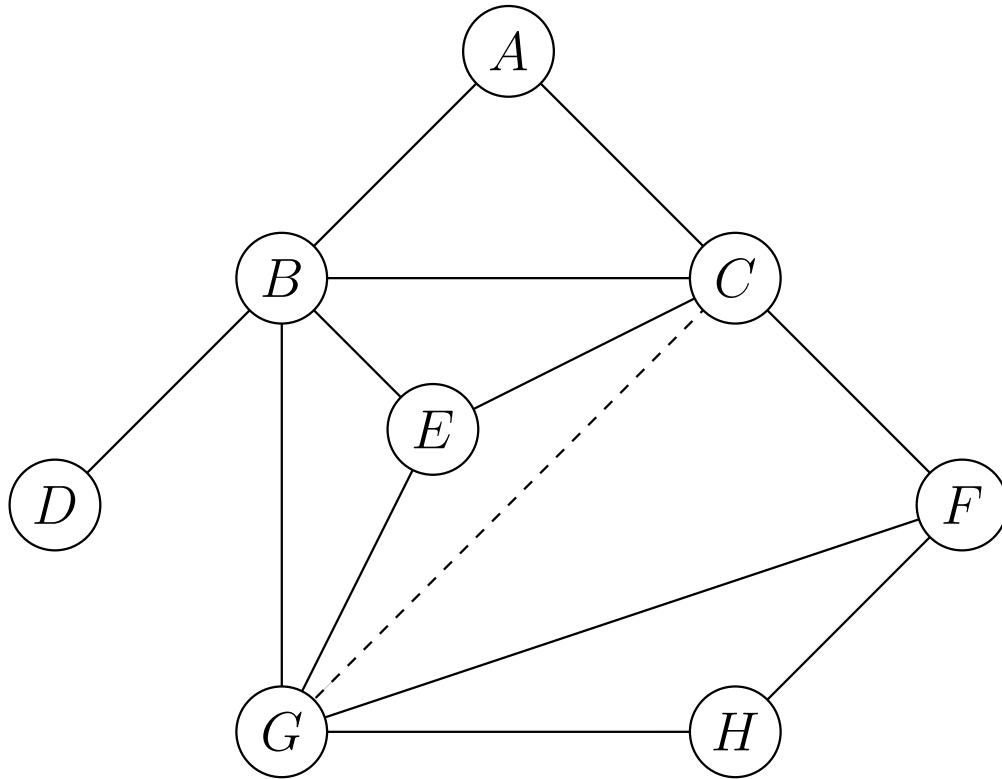
# Example: Step 1: Find a Join-Tree



## Join-Tree creation:

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

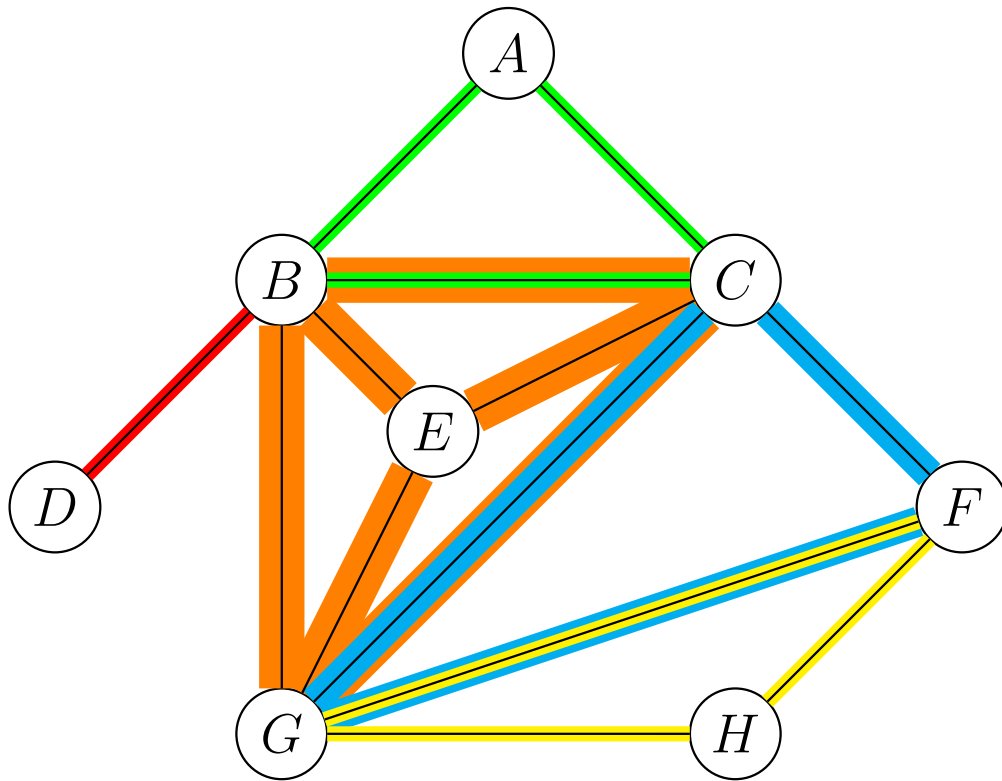
# Example: Step 1: Find a Join-Tree



## Join-Tree creation:

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

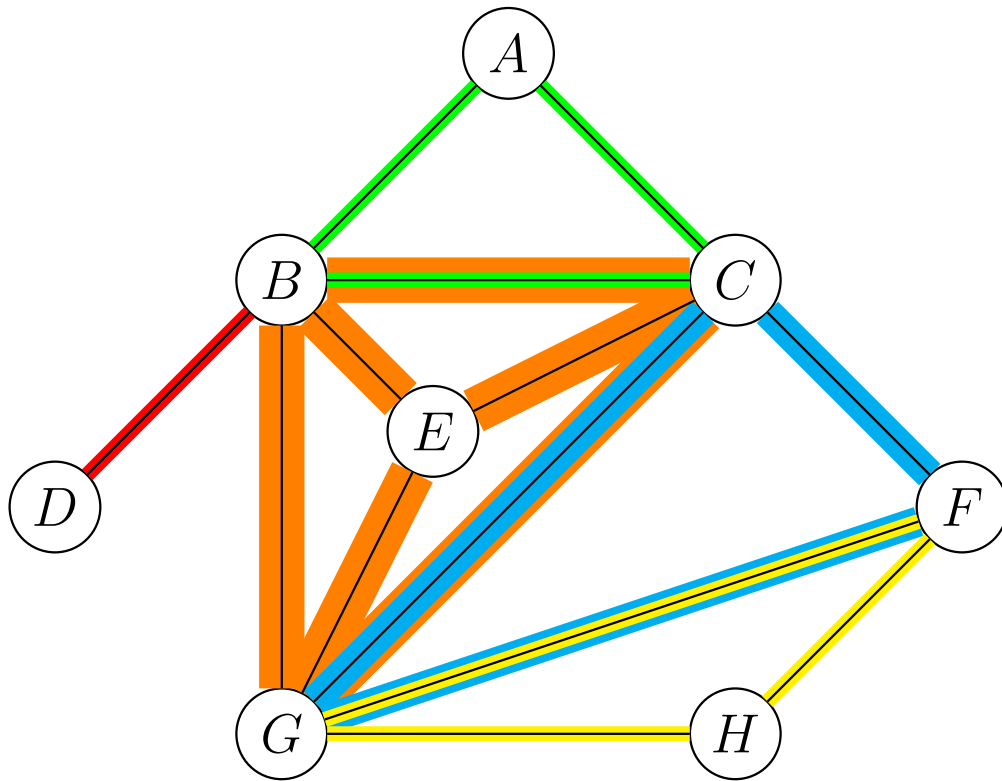
# Example: Step 1: Find a Join-Tree



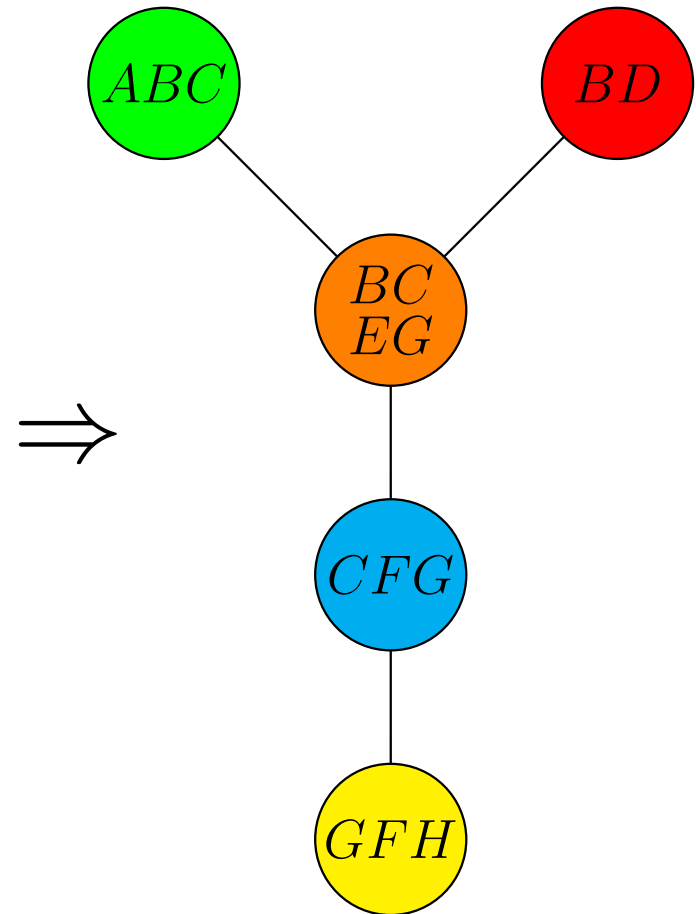
## Join-Tree creation:

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

# Example: Step 1: Find a Join-Tree

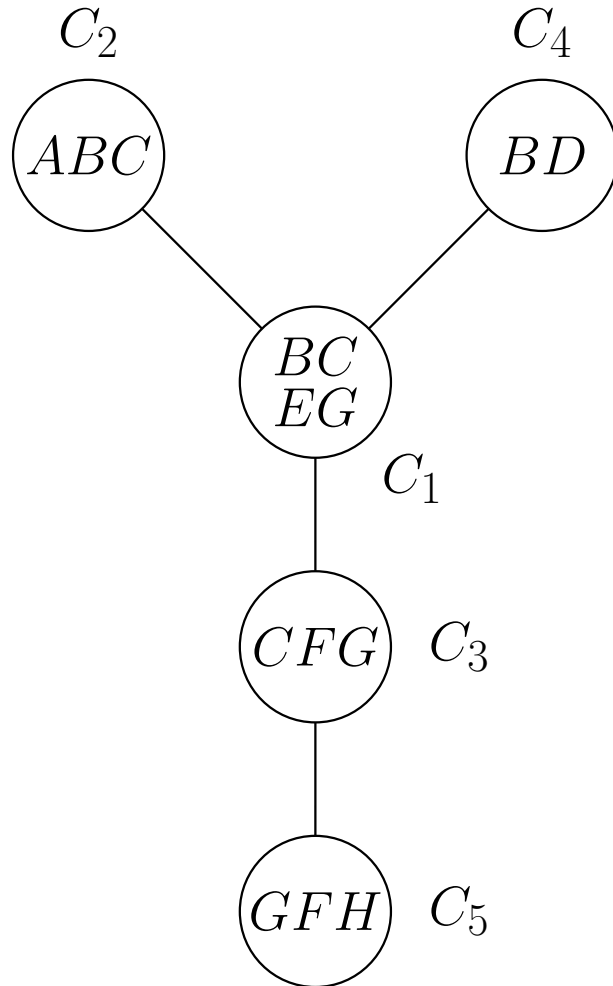


Example Bayesian network



One of the join trees

## Example: Step 2: Specify the Factor Potentials



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

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

**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  $C_i$ :

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

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:

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



## 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:

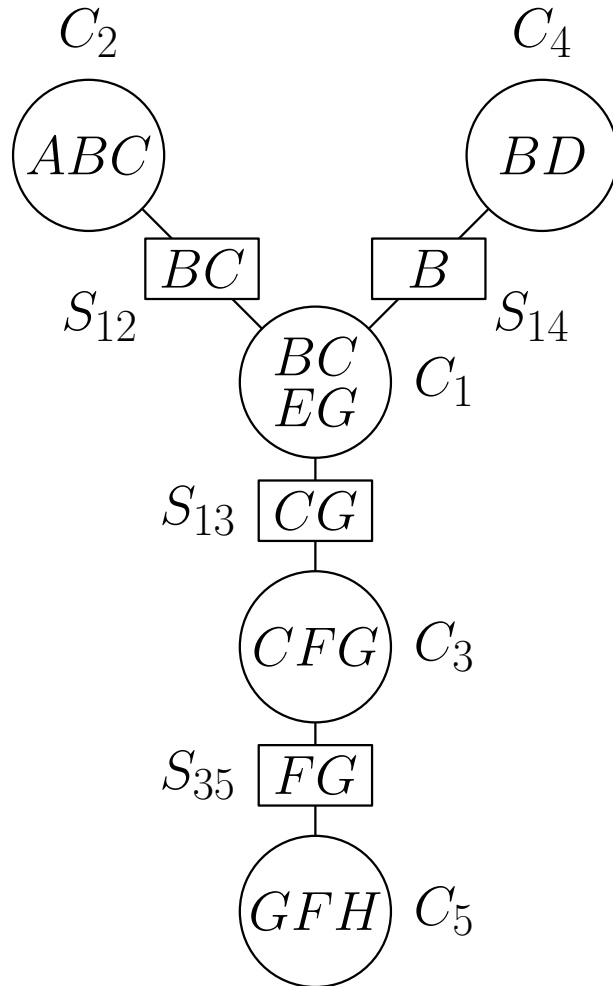
$$\begin{aligned}\Psi_1(b, c, e, g) &= P(e \mid b, c) \cdot P(g \mid e, b) \\ \Psi_2(a, b, c) &= P(b \mid a) \cdot P(c \mid a) \cdot P(a) \\ \Psi_3(c, f, g) &= P(f \mid c) \\ \Psi_4(b, d) &= P(d \mid b) \\ \Psi_5(g, f, h) &= P(h \mid g, f)\end{aligned}$$

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

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

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

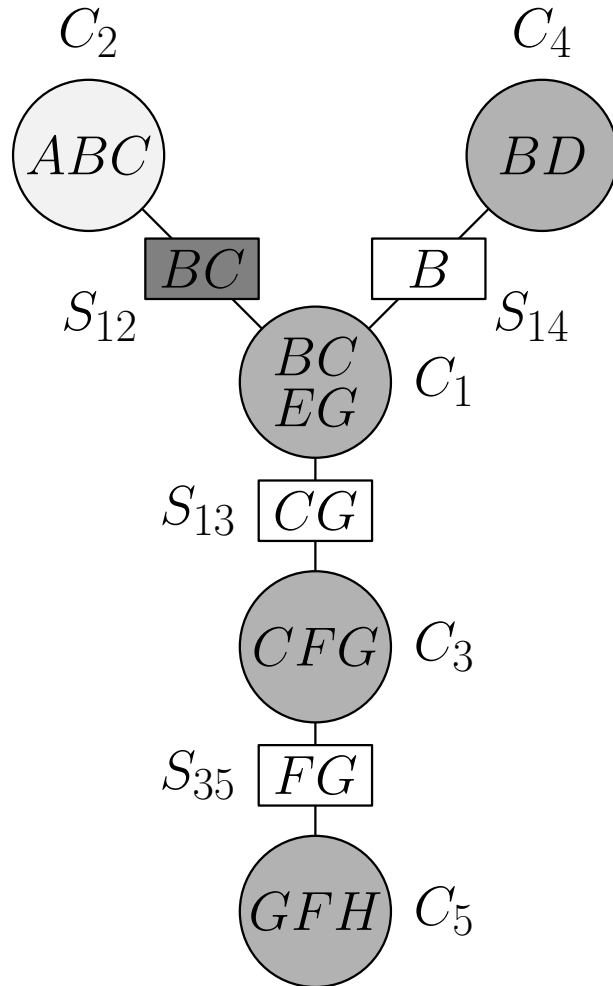
# Example: Closer Look on Option 2: Separation in a Join-Tree



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

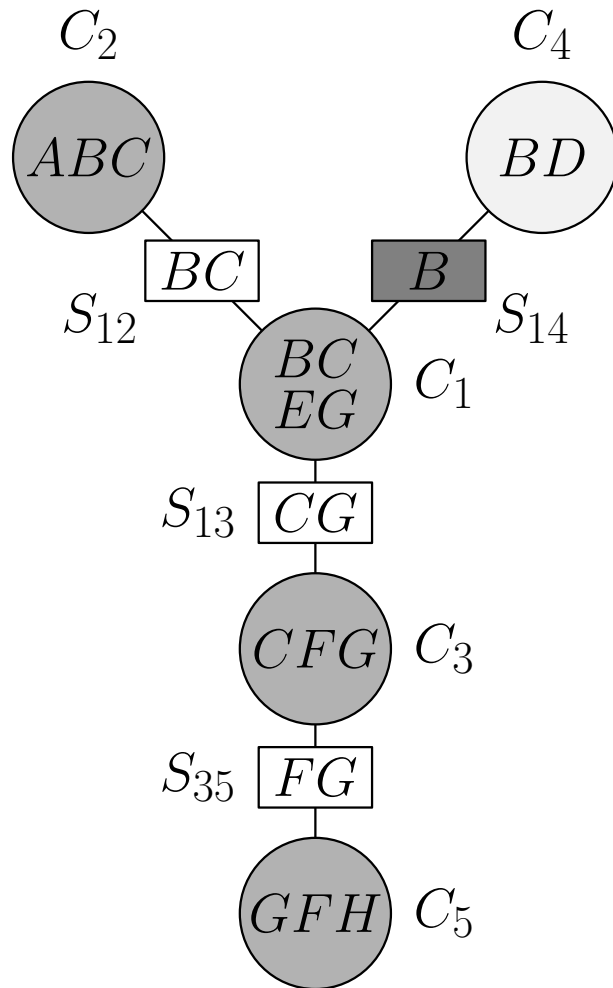


Encoded independence statements:

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



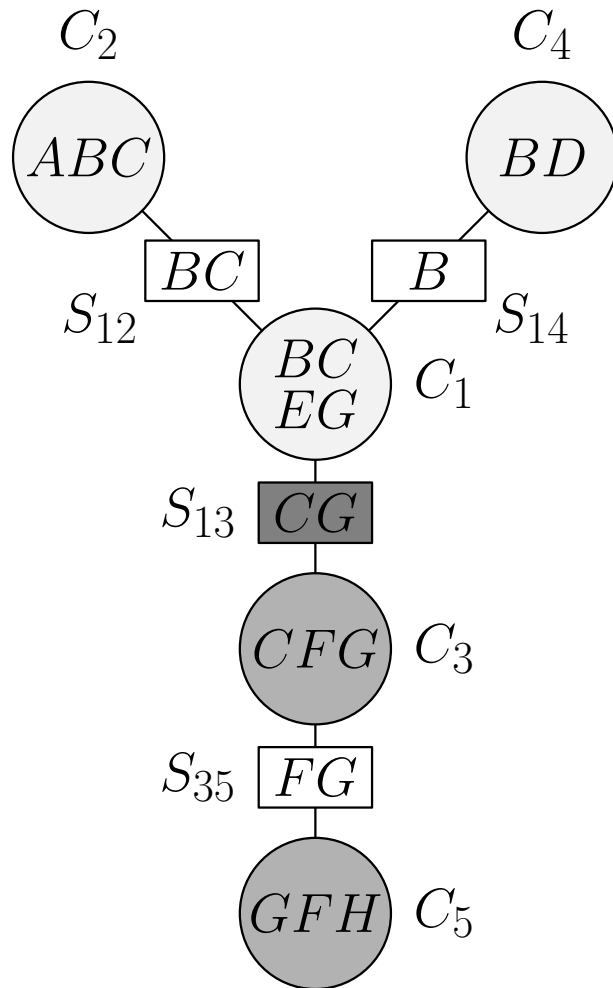
Encoded independence statements:

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

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

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

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



Encoded independence statements:

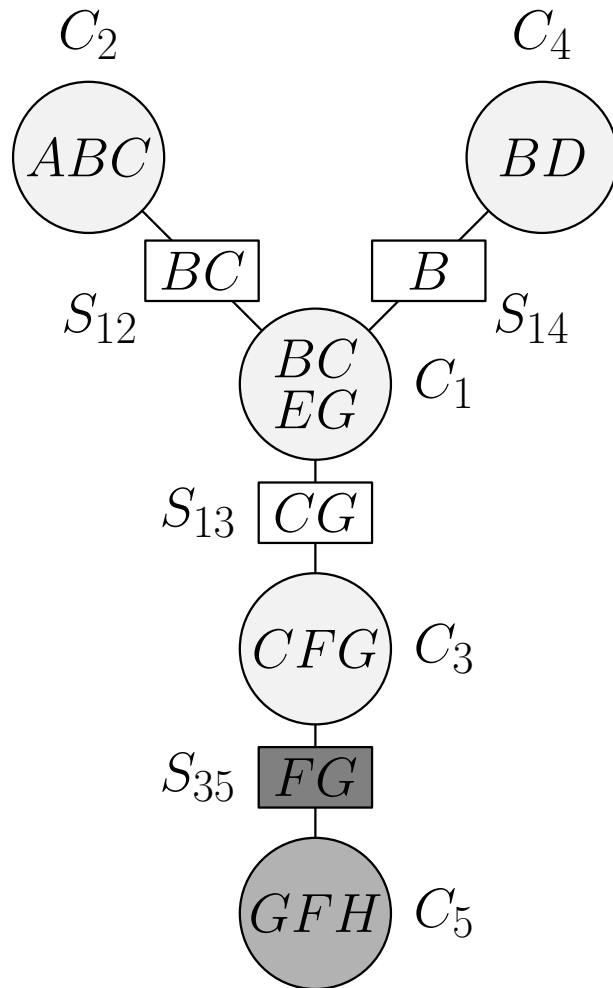
Given any separator, the variables in the cliques on one side become independent of the variables in the cliques on the other side.

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

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

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

# 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.

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

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

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

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

## Example: Closer Look on Option 2: Decomposition

The four separation statements translate into the following independence statements:

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

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

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

## Example: Closer Look on Option 2: Decomposition

The four separation statements translate into the following independence statements:

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

Exploiting the above independencies yields:

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



## Example: Closer Look on Option 2: Decomposition

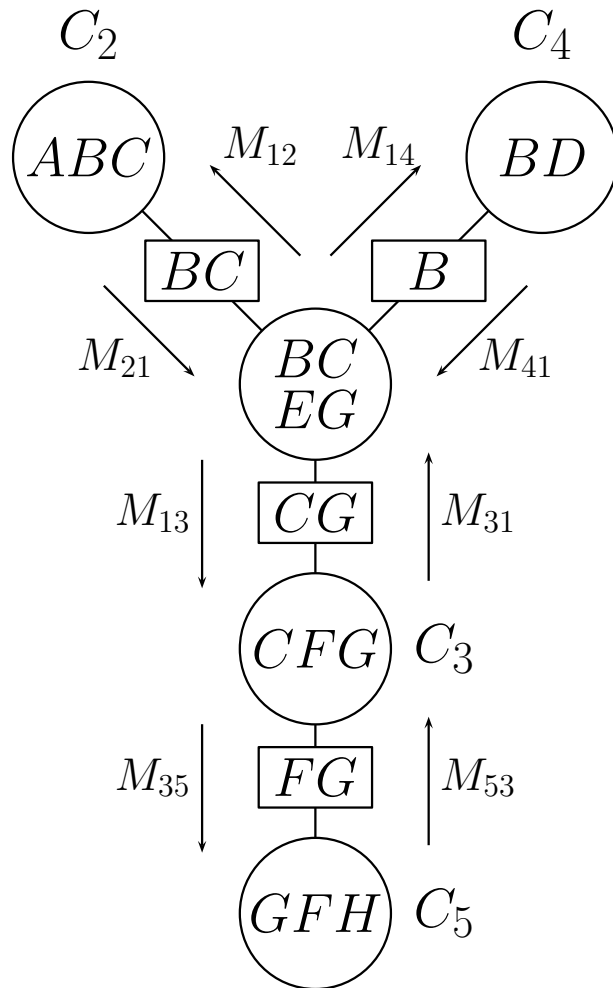
The four separation statements translate into the following independence statements:

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

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:

$$M_{41}(b) = \sum_d \Psi_4(b, d)$$

$$M_{53}(f, g) = \sum_h \Psi_5(f, g, h)$$

$$M_{21}(b, c) = \sum_a \Psi_2(a, b, c)$$

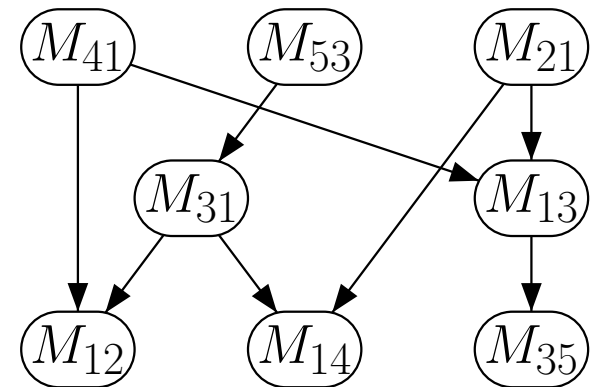
$$M_{31}(c, g) = \sum_f \Psi_3(c, f, g) M_{53}(f, g)$$

$$M_{13}(c, g) = \sum_{b,e} \Psi_1(b, c, e, g) M_{21}(b, c) M_{41}(b)$$

$$M_{12}(b, c) = \sum_{e,g} \Psi_2(b, c, e, g) M_{31}(c, g) M_{41}(b)$$

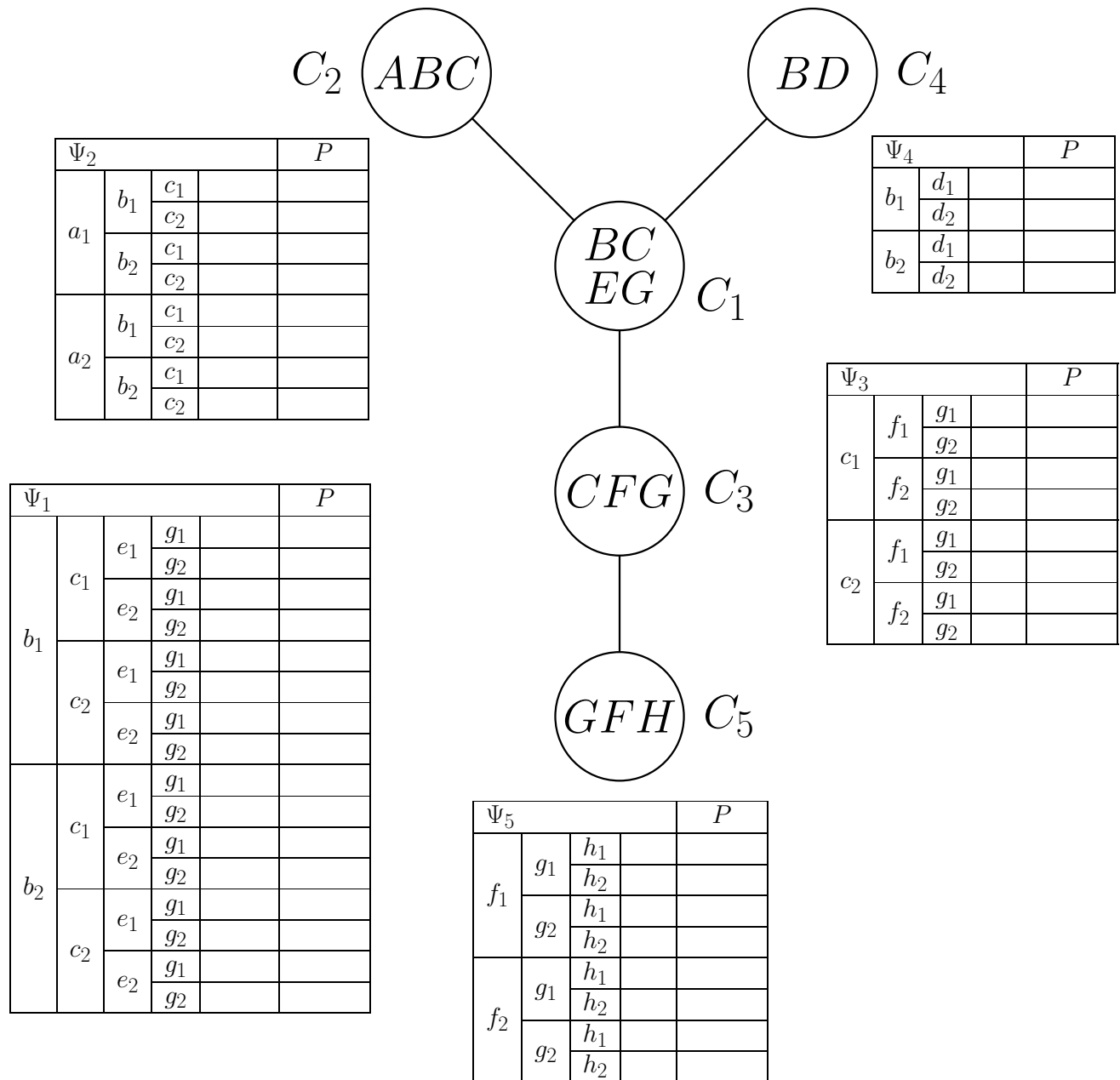
$$M_{14}(b) = \sum_{c,e,g} \Psi_1(b, c, e, g) M_{21}(b, c) M_{31}(c, g)$$

$$M_{35}(f, g) = \sum_c \Psi_3(c, f, g) M_{13}(c, g)$$

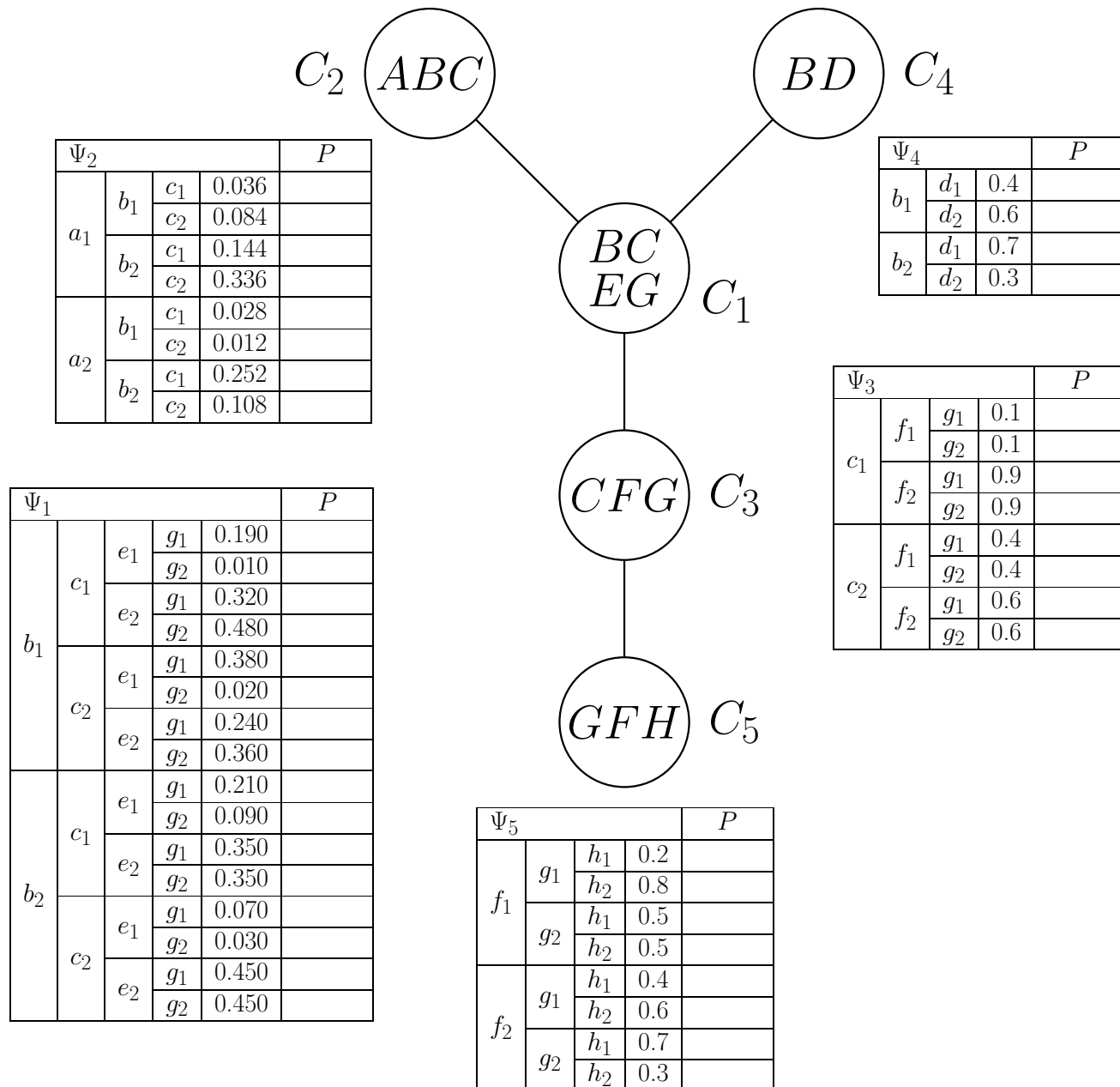


Arrows represent is-needed-for relations. Messages on the same level can be computed in any order. Messages are computed level-wise from top to bottom.

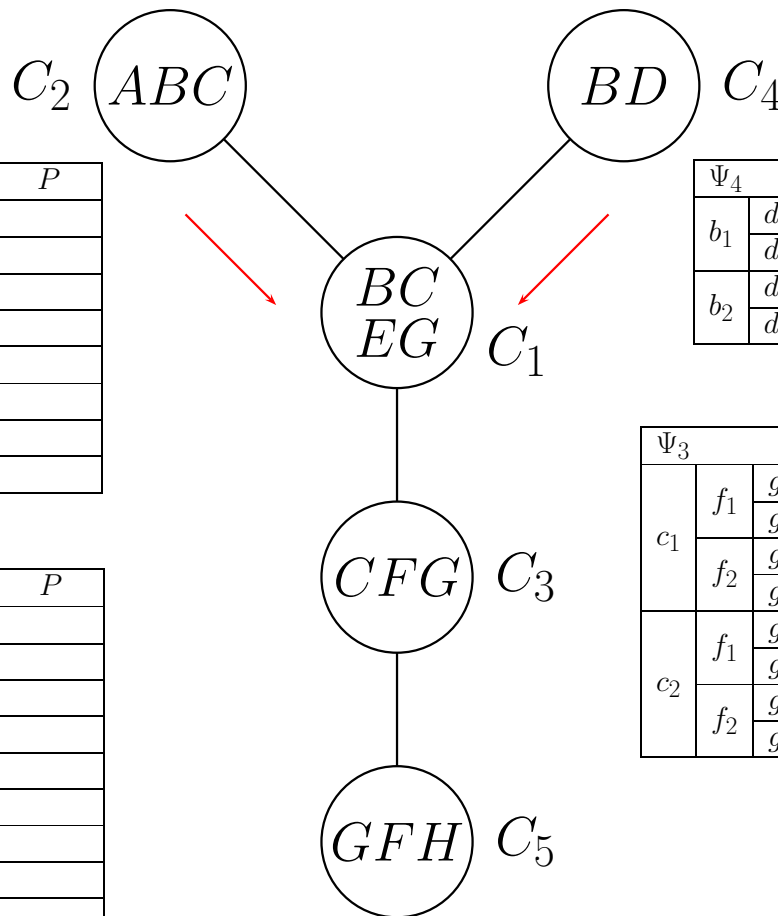
# Example: Step 3: Initialization (Potential Layouts)



# Example: Step 3: Initialization (Potential Values)



# Example: Step 3: Initialization (Sending Messages)



$\Psi_2$				$P$
$a_1$	$b_1$	$c_1$	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	

$\Psi_4$			$P$
$b_1$	$d_1$	0.4	
	$d_2$	0.6	
$b_2$	$d_1$	0.7	
	$d_2$	0.3	

$$M_{21} = \begin{pmatrix} b_{1,c_1} & b_{1,c_2} & b_{2,c_1} & b_{2,c_2} \\ 0.06, & 0.10, & 0.40, & 0.44 \end{pmatrix}$$

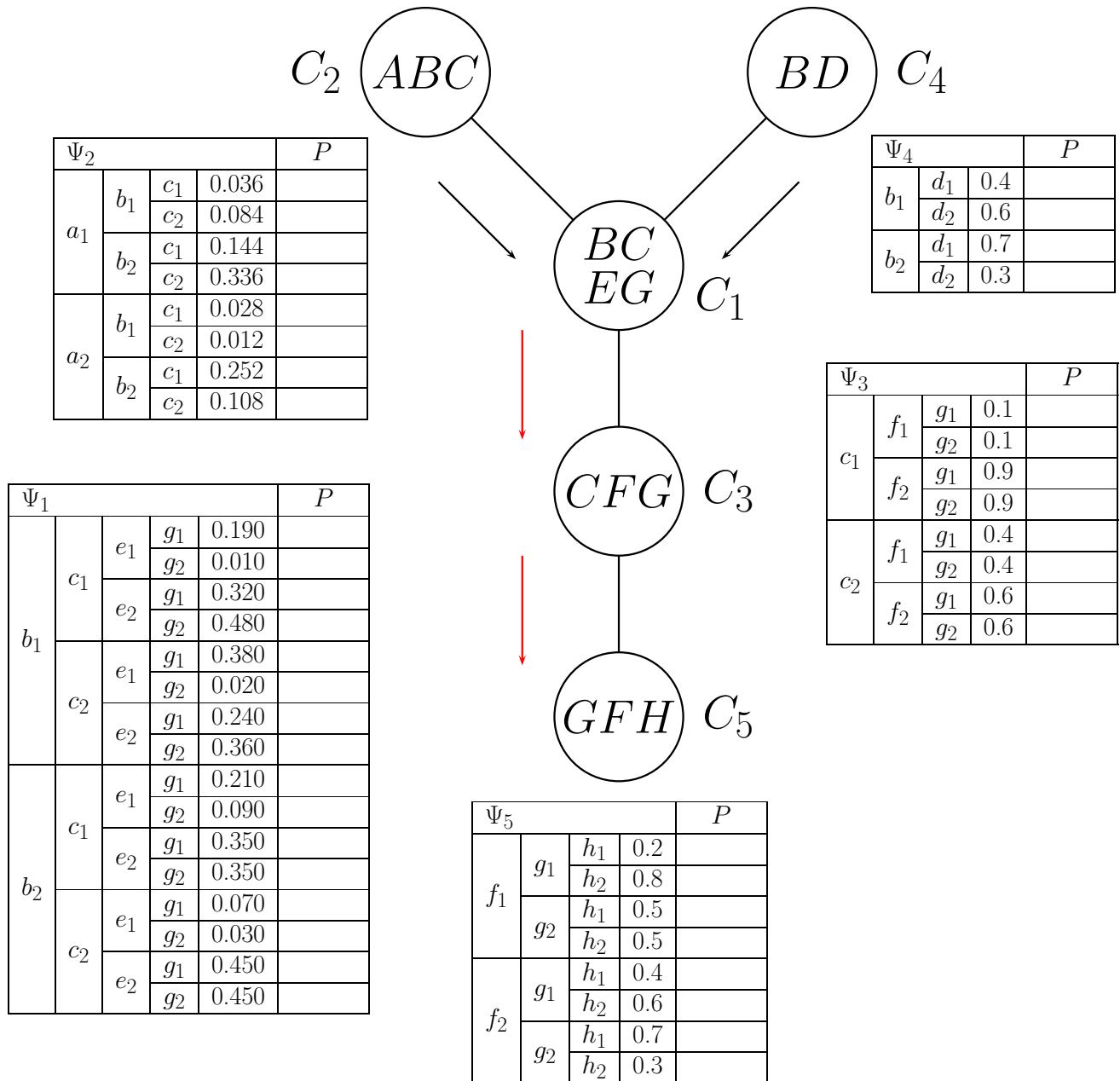
$$M_{41} = \begin{pmatrix} b_1 & b_2 \\ 1, & 1 \end{pmatrix}$$

$\Psi_1$				$P$
$b_1$	$c_1$	$e_1$	$g_1$	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		

$\Psi_3$				$P$
$c_1$	$f_1$	$g_1$	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	

$\Psi_5$				$P$
$f_1$	$g_1$	$h_1$	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)



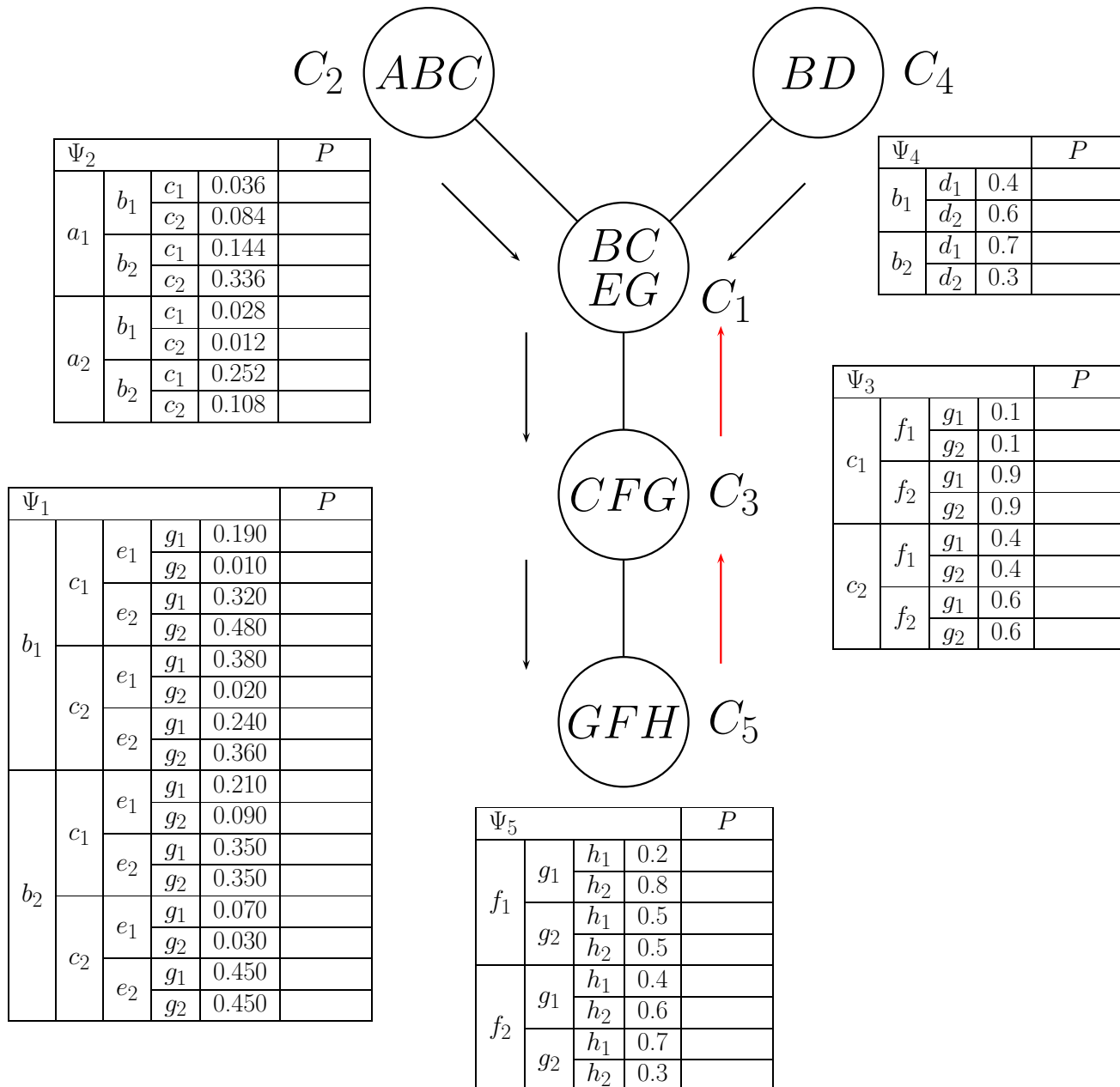
$$M_{21} = (b_{1,c_1} \ b_{1,c_2} \ b_{2,c_1} \ b_{2,c_2})$$

$$M_{41} = (b_1 \ b_2)$$

$$M_{13} = (c_{1,g_1} \ c_{1,g_2} \ c_{2,g_1} \ c_{2,g_2})$$

$$M_{35} = (f_{1,g_1} \ f_{1,g_2} \ f_{2,g_1} \ f_{2,g_2})$$

# Example: Step 3: Initialization (Sending Messages)



$$M_{21} = (b_{1,c_1} \ b_{1,c_2} \ b_{2,c_1} \ b_{2,c_2})$$

$$M_{41} = (b_1 \ b_2)$$

$$M_{13} = (c_{1,g_1} \ c_{1,g_2} \ c_{2,g_1} \ c_{2,g_2})$$

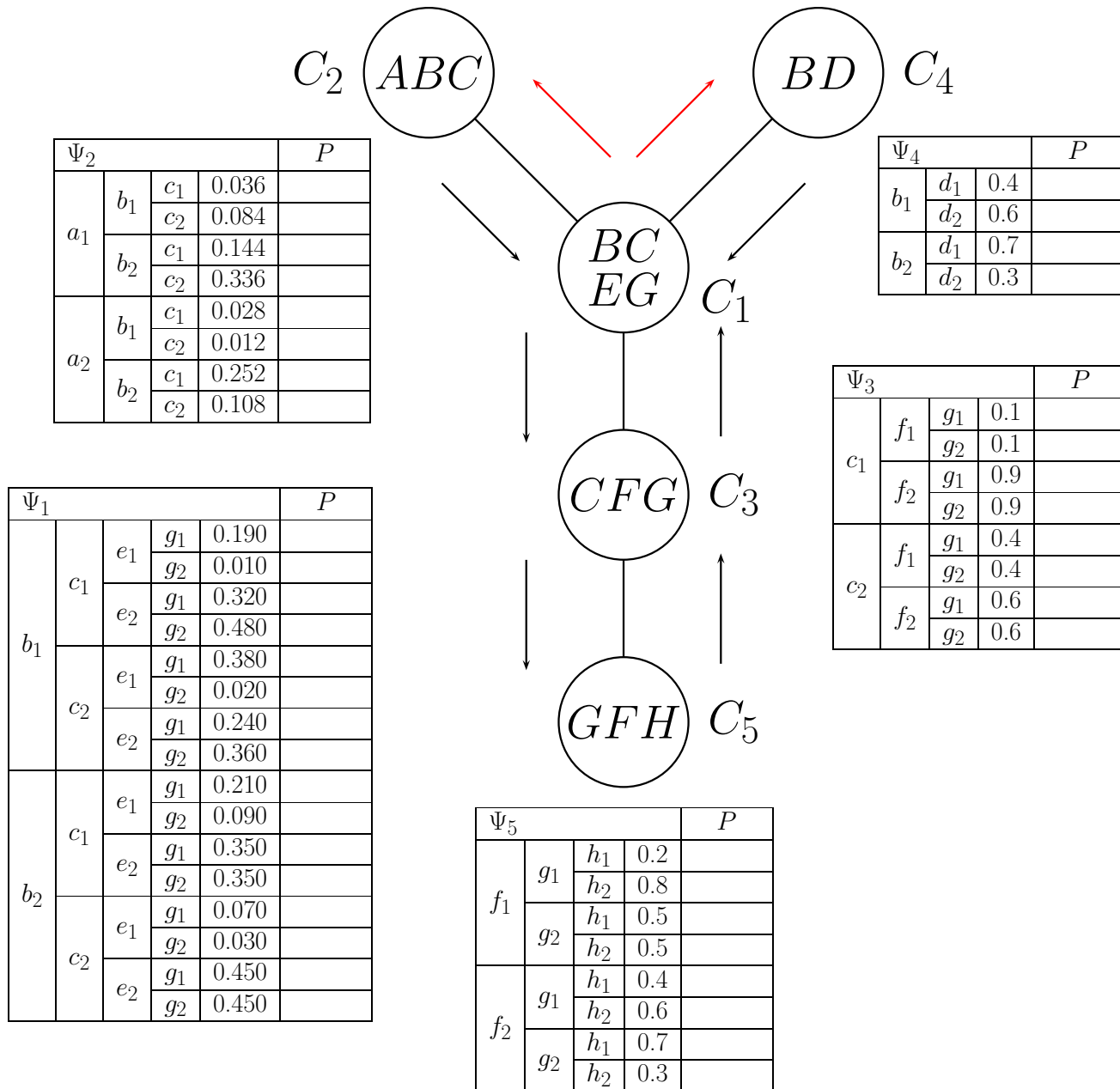
$$M_{35} = (f_{1,g_1} \ f_{1,g_2} \ f_{2,g_1} \ f_{2,g_2})$$

$$M_{53} = (1, 1, 1, 1)$$

$$M_{31} = (c_{1,g_1} \ c_{1,g_2} \ c_{2,g_1} \ c_{2,g_2})$$



# Example: Step 3: Initialization (Sending Messages)



$$M_{21} = (b_{1,c_1} \ b_{1,c_2} \ b_{2,c_1} \ b_{2,c_2})$$

$$= (0.06, 0.10, 0.40, 0.44)$$

$$M_{41} = (b_1 \ b_2)$$

$$= (1, 1)$$

$$M_{13} = (c_{1,g_1} \ c_{1,g_2} \ c_{2,g_1} \ c_{2,g_2})$$

$$= (0.254, 0.206, 0.290, 0.250)$$

$$M_{35} = (f_{1,g_1} \ f_{1,g_2} \ f_{2,g_1} \ f_{2,g_2})$$

$$= (0.14, 0.12, 0.40, 0.33)$$

$$M_{53} = (f_{1,g_1} \ f_{1,g_2} \ f_{2,g_1} \ f_{2,g_2})$$

$$= (1, 1, 1, 1)$$

$$M_{31} = (c_{1,g_1} \ c_{1,g_2} \ c_{2,g_1} \ c_{2,g_2})$$

$$= (1, 1, 1, 1)$$

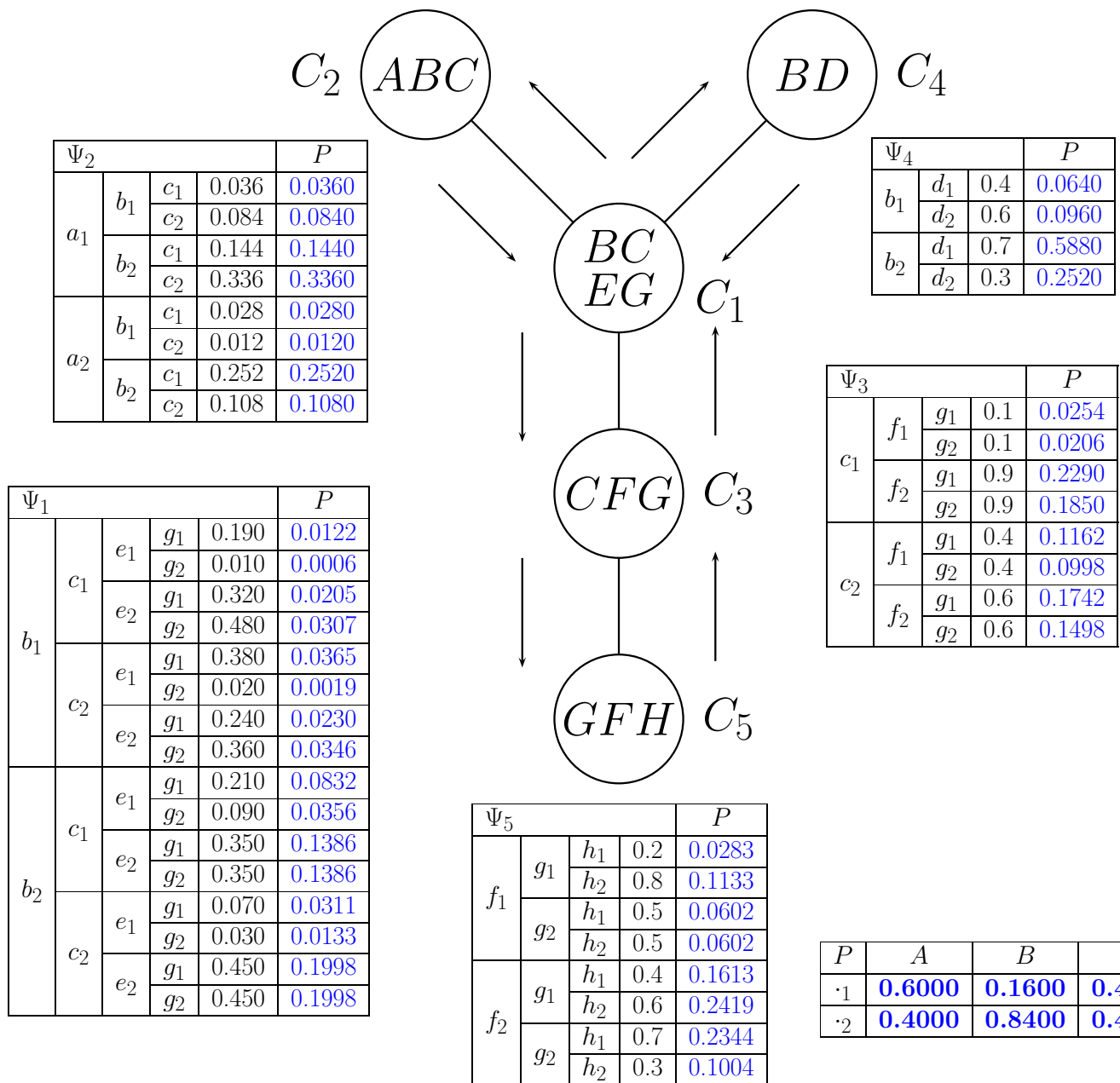
$$M_{12} = (b_{1,c_1} \ b_{1,c_2} \ b_{2,c_1} \ b_{2,c_2})$$

$$= (1, 1, 1, 1)$$

$$M_{14} = (b_1 \ b_2)$$

$$= (0.16, 0.84)$$

# Example: Step 3: Initialization Complete



$$M_{21} = \begin{pmatrix} b_{1,c_1} & b_{1,c_2} & b_{2,c_1} & b_{2,c_2} \\ 0.06, & 0.10, & 0.40, & 0.44 \end{pmatrix}$$

$$M_{41} = \begin{pmatrix} b_1 & b_2 \\ 1, & 1 \end{pmatrix}$$

$$M_{13} = \begin{pmatrix} c_{1,g_1} & c_{1,g_2} & c_{2,g_1} & c_{2,g_2} \\ 0.254, & 0.206, & 0.290, & 0.250 \end{pmatrix}$$

$$M_{35} = \begin{pmatrix} f_{1,g_1} & f_{1,g_2} & f_{2,g_1} & f_{2,g_2} \\ 0.14, & 0.12, & 0.40, & 0.33 \end{pmatrix}$$

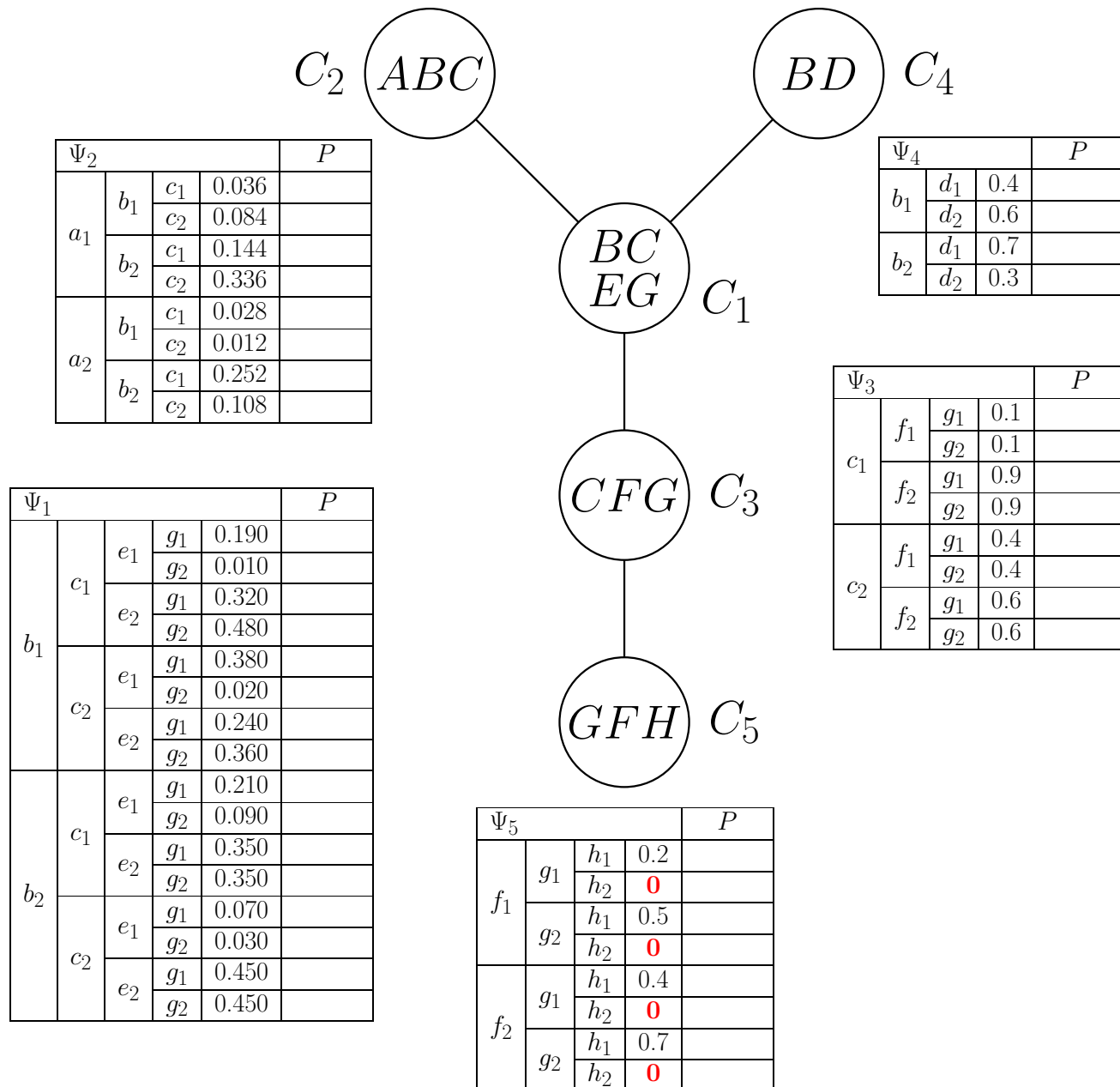
$$M_{53} = \begin{pmatrix} f_{1,g_1} & f_{1,g_2} & f_{2,g_1} & f_{2,g_2} \\ 1, & 1, & 1, & 1 \end{pmatrix}$$

$$M_{31} = \begin{pmatrix} c_{1,g_1} & c_{1,g_2} & c_{2,g_1} & c_{2,g_2} \\ 1, & 1, & 1, & 1 \end{pmatrix}$$

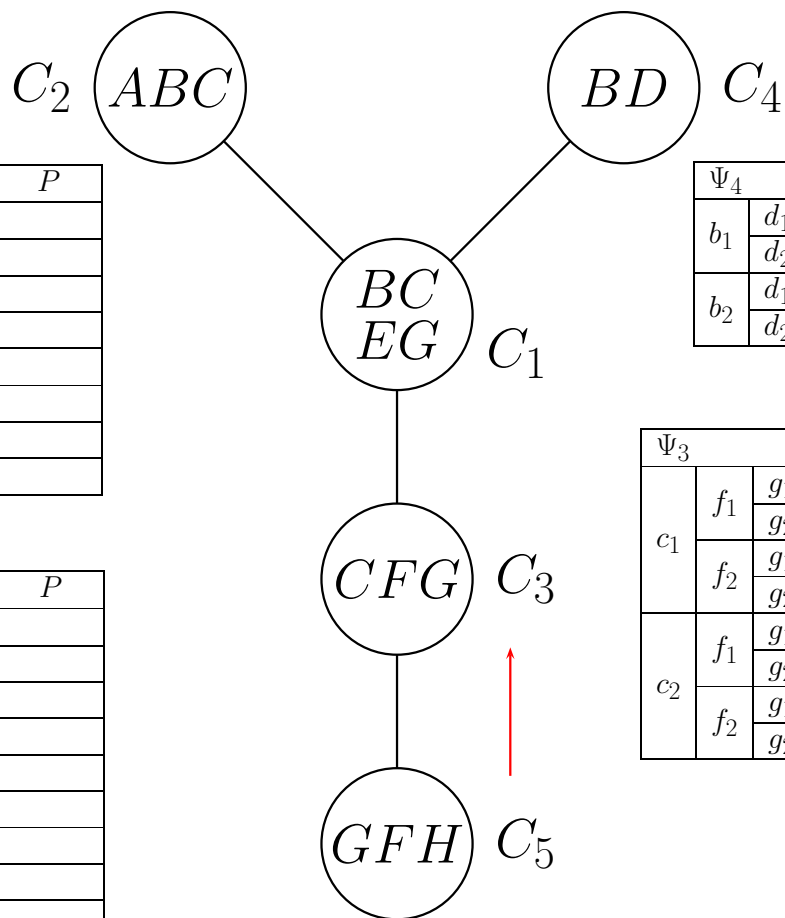
$$M_{12} = \begin{pmatrix} b_{1,c_1} & b_{1,c_2} & b_{2,c_1} & b_{2,c_2} \\ 1, & 1, & 1, & 1 \end{pmatrix}$$

$$M_{14} = \begin{pmatrix} b_1 & b_2 \\ 0.16, & 0.84 \end{pmatrix}$$

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



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



$$M_{53} = \begin{pmatrix} f_{1,g1} & f_{1,g2} & f_{2,g1} & f_{2,g2} \\ 0.2 & 0.5 & 0.4 & 0.7 \end{pmatrix}$$

$\Psi_2$				$P$
$a_1$	$b_1$	$c_1$	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	

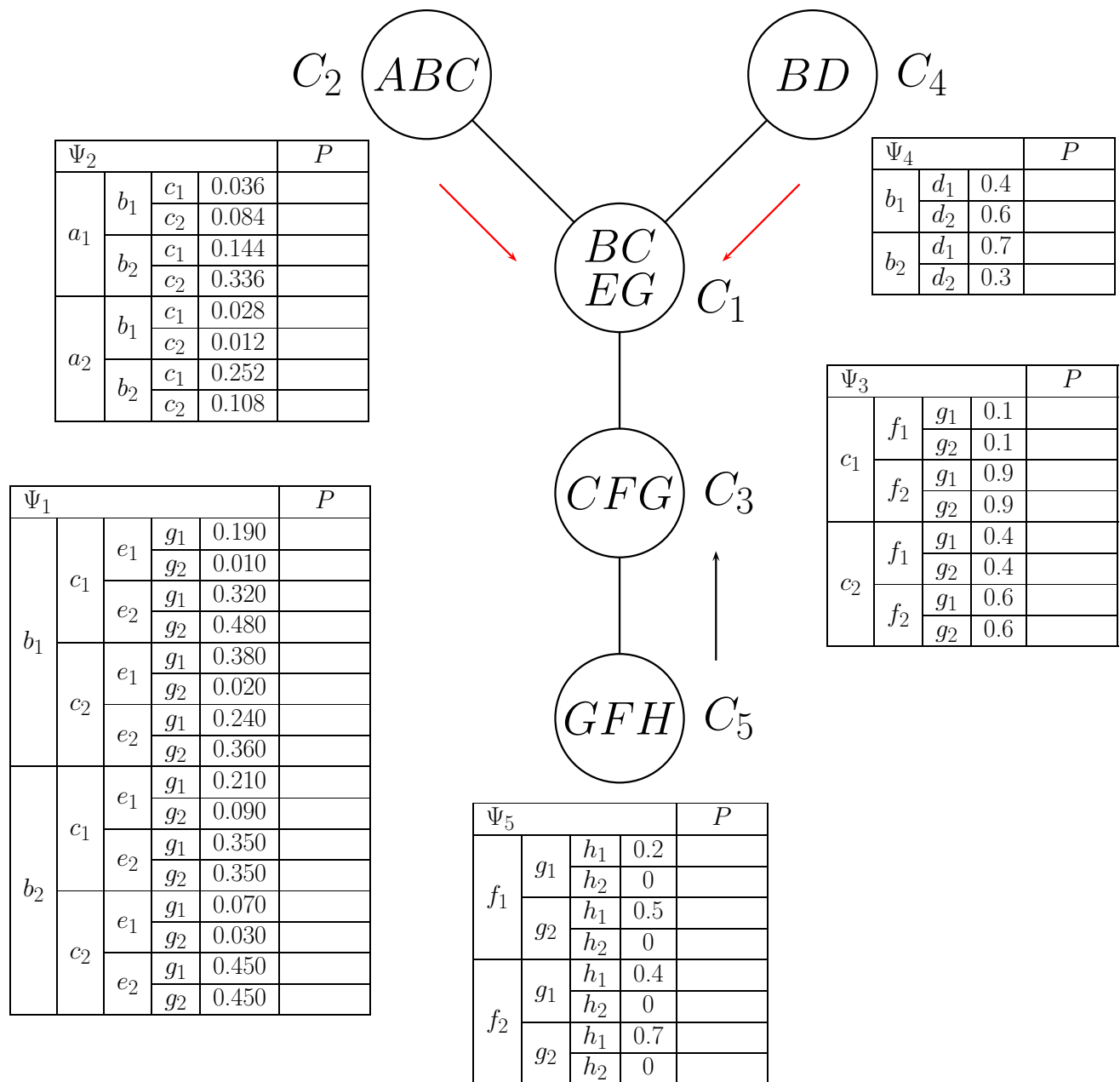
$\Psi_4$			$P$
$b_1$	$d_1$	0.4	
	$d_2$	0.6	
$b_2$	$d_1$	0.7	
	$d_2$	0.3	

$\Psi_1$				$P$
$b_1$	$c_1$	$e_1$	$g_1$	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

$\Psi_3$				$P$
$c_1$	$f_1$	$g_1$	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	

$\Psi_5$				$P$
$f_1$	$g_1$	$h_1$	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)

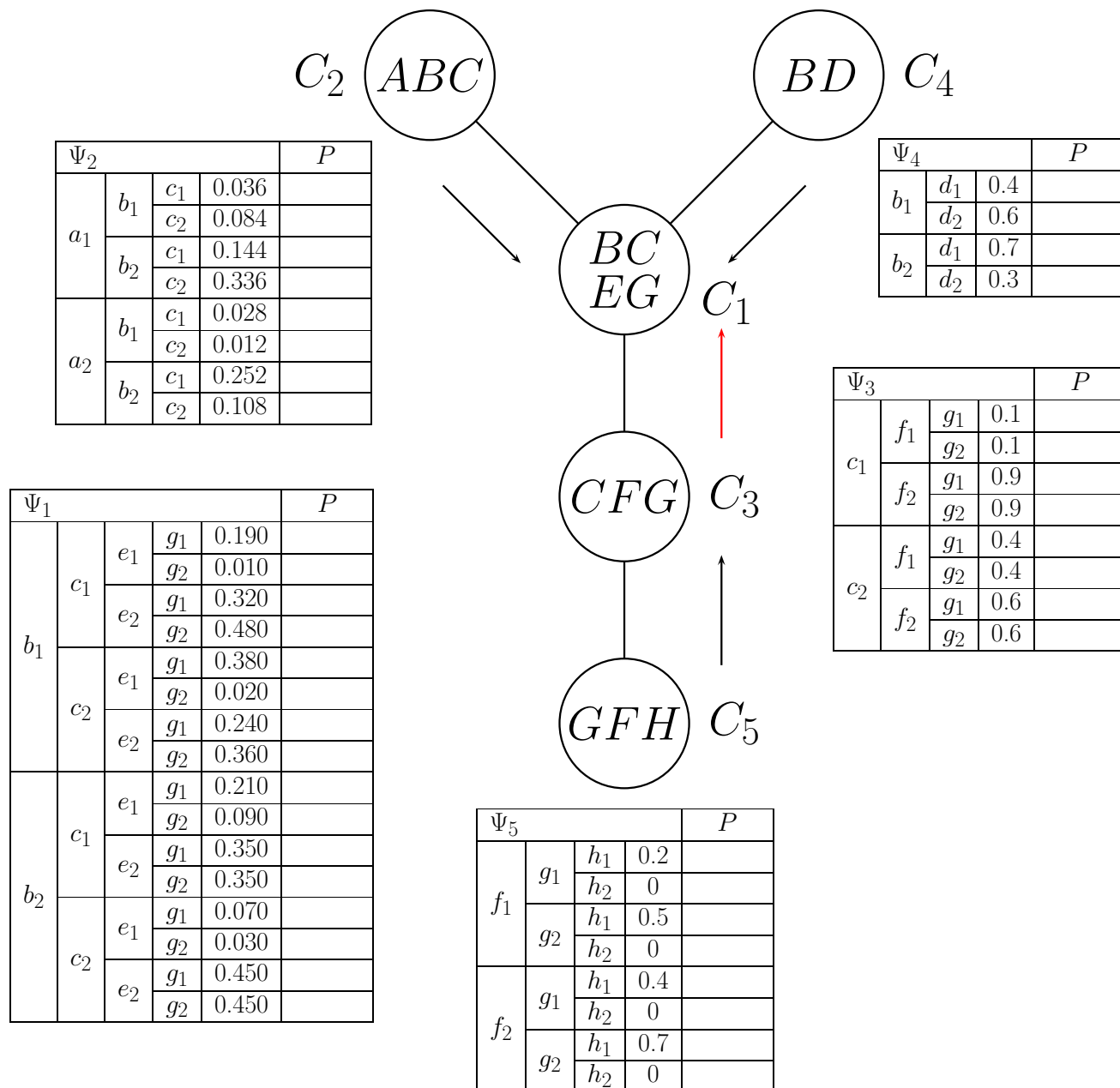


$$M_{53} = \begin{pmatrix} f_{1,g_1} & f_{1,g_2} & f_{2,g_1} & f_{2,g_2} \\ 0.2 & 0.5 & 0.4 & 0.7 \end{pmatrix}$$

$$M_{21} = \begin{pmatrix} b_{1,c_1} & b_{1,c_2} & b_{2,c_1} & b_{2,c_2} \\ 0.06 & 0.10 & 0.40 & 0.44 \end{pmatrix}$$

$$M_{41} = \begin{pmatrix} b_1 & b_2 \\ 1 & 1 \end{pmatrix}$$

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



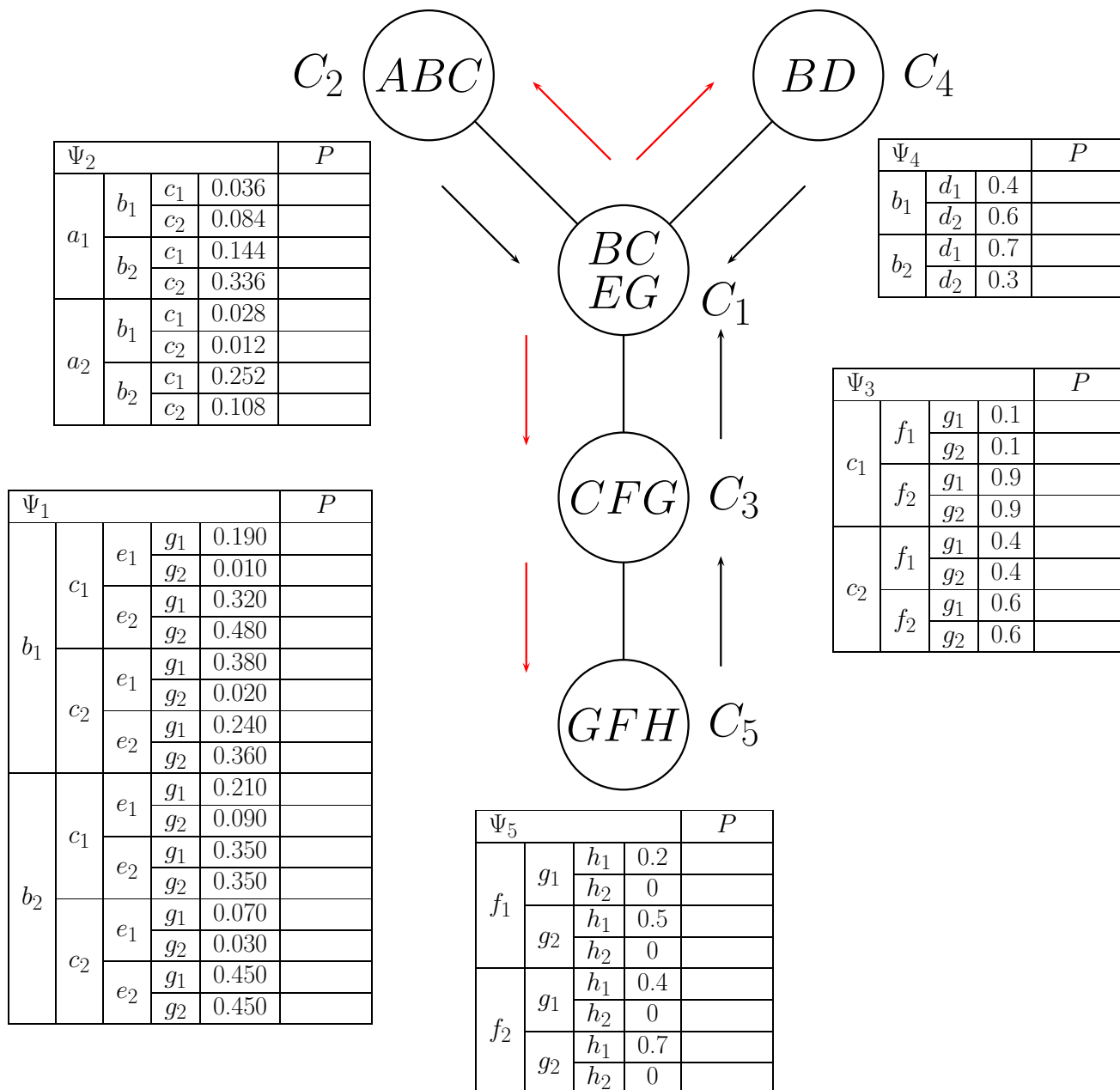
$$M_{53} = \begin{pmatrix} f_{1,g_1} & f_{1,g_2} & f_{2,g_1} & f_{2,g_2} \\ 0.2 & 0.5 & 0.4 & 0.7 \end{pmatrix}$$

$$M_{21} = \begin{pmatrix} b_{1,c_1} & b_{1,c_2} & b_{2,c_1} & b_{2,c_2} \\ 0.06 & 0.10 & 0.40 & 0.44 \end{pmatrix}$$

$$M_{41} = \begin{pmatrix} b_1 & b_2 \\ 1 & 1 \end{pmatrix}$$

$$M_{31} = \begin{pmatrix} c_{1,g_1} & c_{1,g_2} & c_{2,g_1} & c_{2,g_2} \\ 0.38 & 0.68 & 0.32 & 0.62 \end{pmatrix}$$

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



$$M_{53} = \begin{pmatrix} f_{1,g_1} & f_{1,g_2} & f_{2,g_1} & f_{2,g_2} \\ 0.2 & 0.5 & 0.4 & 0.7 \end{pmatrix}$$

$$M_{21} = \begin{pmatrix} b_{1,c_1} & b_{1,c_2} & b_{2,c_1} & b_{2,c_2} \\ 0.06 & 0.10 & 0.40 & 0.44 \end{pmatrix}$$

$$M_{41} = \begin{pmatrix} b_1 & b_2 \\ 1 & 1 \end{pmatrix}$$

$$M_{31} = \begin{pmatrix} c_{1,g_1} & c_{1,g_2} & c_{2,g_1} & c_{2,g_2} \\ 0.38 & 0.68 & 0.32 & 0.62 \end{pmatrix}$$

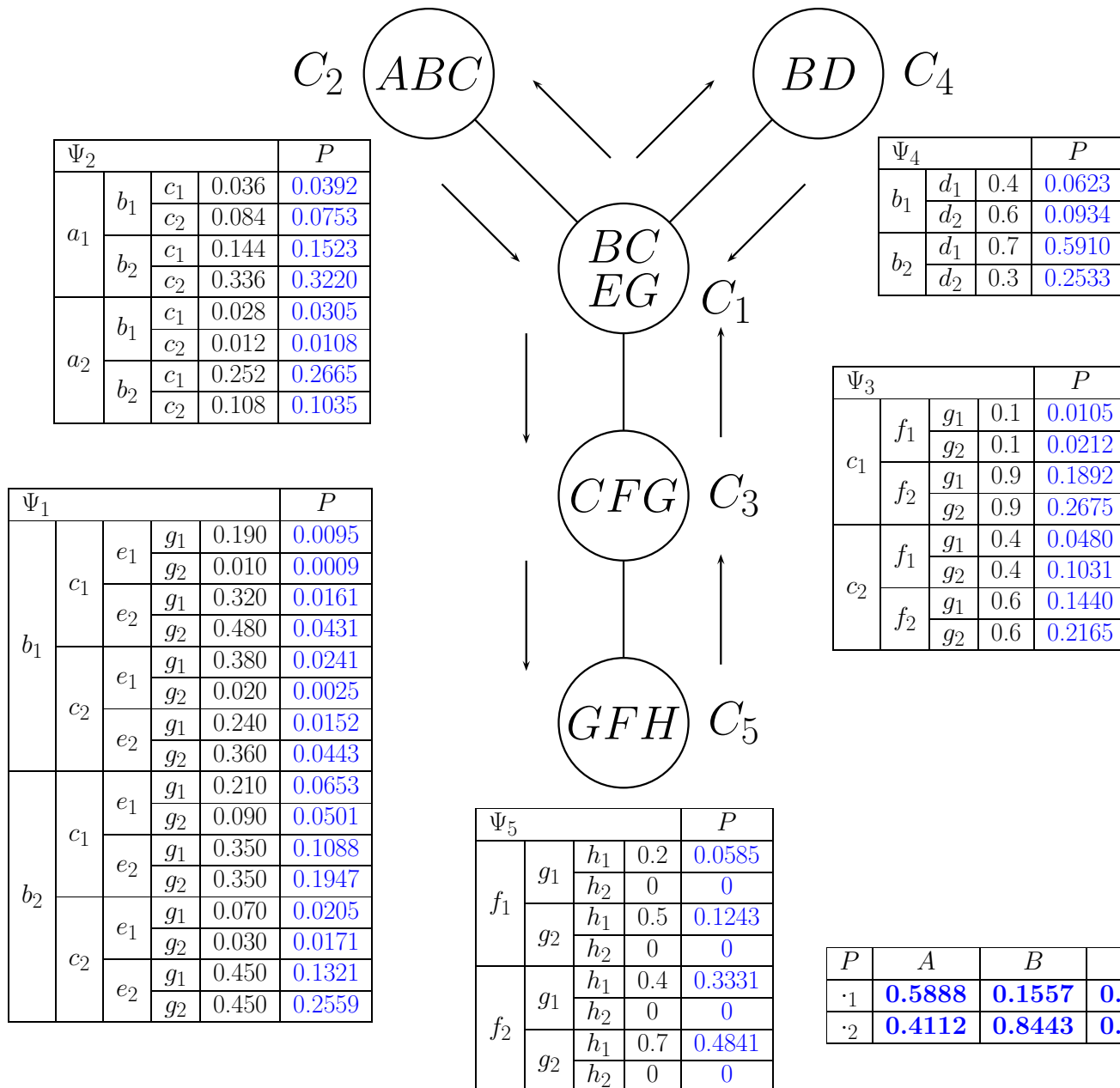
$$M_{12} = \begin{pmatrix} b_{1,c_1} & b_{1,c_2} & b_{2,c_1} & b_{2,c_2} \\ 0.527 & 0.434 & 0.512 & 0.464 \end{pmatrix}$$

$$M_{14} = \begin{pmatrix} b_1 & b_2 \\ 0.075 & 0.409 \end{pmatrix}$$

$$M_{13} = \begin{pmatrix} c_{1,g_1} & c_{1,g_2} & c_{2,g_1} & c_{2,g_2} \\ 0.254 & 0.206 & 0.290 & 0.250 \end{pmatrix}$$

$$M_{35} = \begin{pmatrix} f_{1,g_1} & f_{1,g_2} & f_{2,g_1} & f_{2,g_2} \\ 0.14 & 0.12 & 0.40 & 0.33 \end{pmatrix}$$

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



$$M_{53} = \begin{pmatrix} f_{1,g_1} & f_{1,g_2} & f_{2,g_1} & f_{2,g_2} \\ 0.2 & 0.5 & 0.4 & 0.7 \end{pmatrix}$$

$$M_{21} = \begin{pmatrix} b_{1,c_1} & b_{1,c_2} & b_{2,c_1} & b_{2,c_2} \\ 0.06 & 0.10 & 0.40 & 0.44 \end{pmatrix}$$

$$M_{41} = \begin{pmatrix} b_1 & b_2 \\ 1 & 1 \end{pmatrix}$$

$$M_{31} = \begin{pmatrix} c_{1,g_1} & c_{1,g_2} & c_{2,g_1} & c_{2,g_2} \\ 0.38 & 0.68 & 0.32 & 0.62 \end{pmatrix}$$

$$M_{12} = \begin{pmatrix} b_{1,c_1} & b_{1,c_2} & b_{2,c_1} & b_{2,c_2} \\ 0.527 & 0.434 & 0.512 & 0.464 \end{pmatrix}$$

$$M_{14} = \begin{pmatrix} b_1 & b_2 \\ 0.075 & 0.409 \end{pmatrix}$$

$$M_{13} = \begin{pmatrix} c_{1,g_1} & c_{1,g_2} & c_{2,g_1} & c_{2,g_2} \\ 0.254 & 0.206 & 0.290 & 0.250 \end{pmatrix}$$

$$M_{35} = \begin{pmatrix} f_{1,g_1} & f_{1,g_2} & f_{2,g_1} & f_{2,g_2} \\ 0.14 & 0.12 & 0.40 & 0.33 \end{pmatrix}$$

$P$	$A$	$B$	$C$	$D$	$E$	$F$	$G$	$H$
$\cdot_1$	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



# Summary

There are several exact inference methods such as variable elimination, clique tree propagation or recursive conditioning. These algorithms have complexity that is exponential with networks tree width. Exact inference is NP-hard.

In very large applications it is often useful to introduce topological structural constraints or restrictions on conditional probabilities, i.e. bounded variance algorithms.

There are also approximative inference methods.