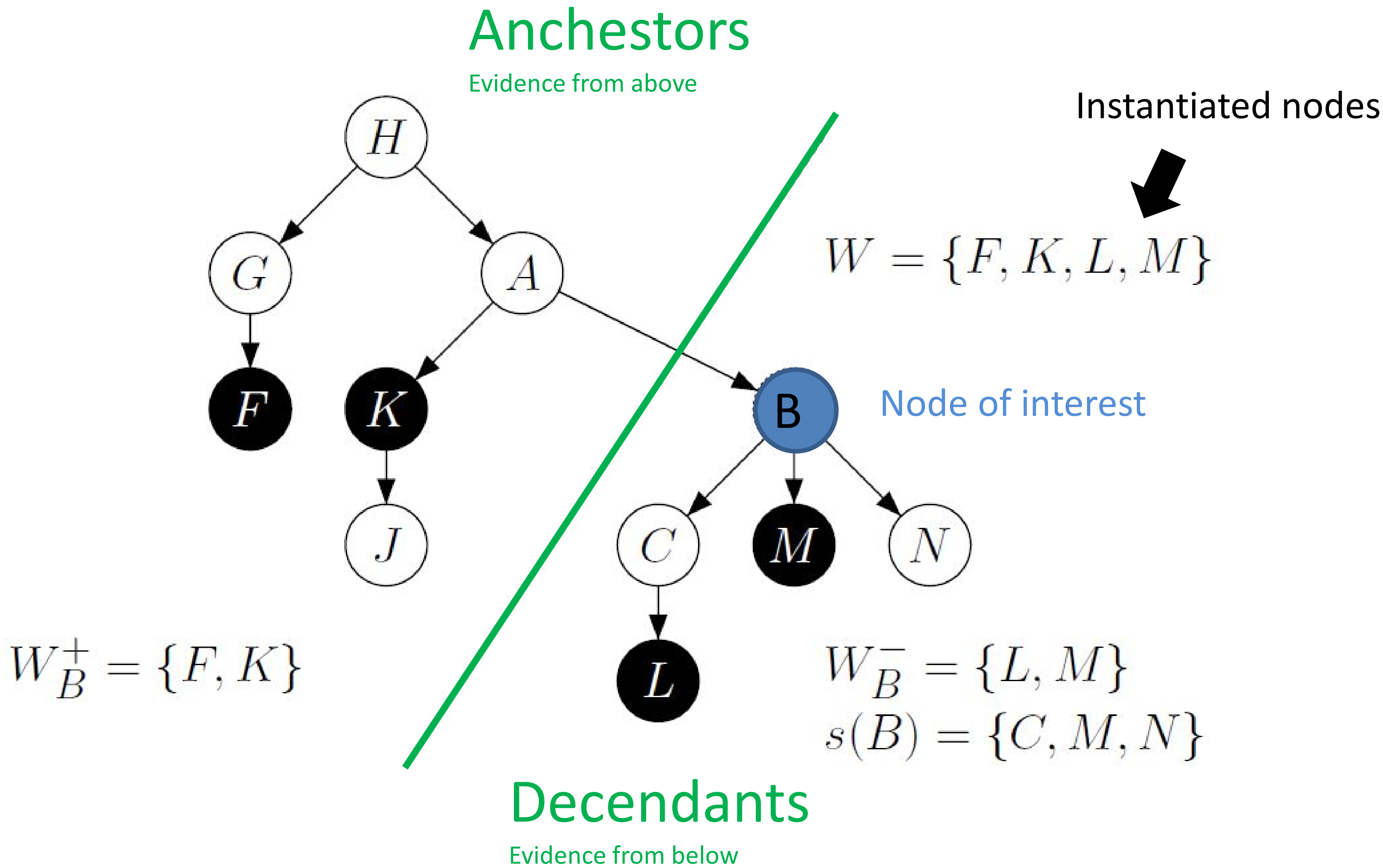
Evidence Propagation in Bayesian Trees

Example 1: Conditioning in a Bayesian Tree



The Formal Problem

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

Influence of the Evidence

$$P(B=b \mid W) = P(b \mid W_B^-, W_B^+) \quad \text{with } B \notin W$$

Def. Cond. Prob.

$$= \frac{P(W_B^-, W_B^+, B=b)}{P(W_B^-, W_B^+)}$$

Bayes Theorem

$$= \frac{P(W_B^-, W_B^+ \mid b)P(b)}{P(W_B^-, W_B^+)}$$

d-separation

$$= \frac{P(W_B^- \mid b)P(W_B^+ \mid b)P(b)}{P(W_B^-, W_B^+)}$$

Bayes Theorem

$$= \frac{P(W_B^+)}{P(W_B^-, W_B^+)} \times \underbrace{P(W_B^- \mid b)}_{\text{Evidence from "below"}} \times \underbrace{P(b \mid W_B^+)}_{\text{Evidence from "above"}}$$

This value depends only on B and W. It is the same for all values b of B. So it can be calculated by normalization using the evidences. This „normalization factor" is denoted by $\alpha_{B,W}$.

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:

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

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

π- and λ-Values

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

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

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

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

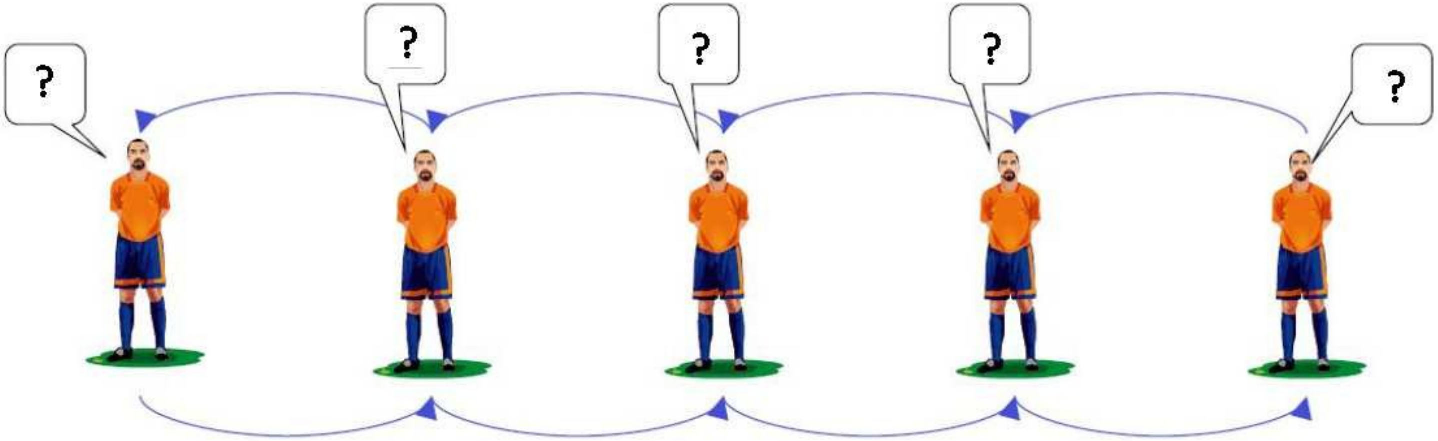
The values for λ and π are changing during the initialization and propagation phase. It is convenient to store the values as table in a node processor

	π	λ	P
b_1	0.25	0.8	0.229
b_2	0.75	0.9	0.771

Example 2 Belief Propagation via Message Passing

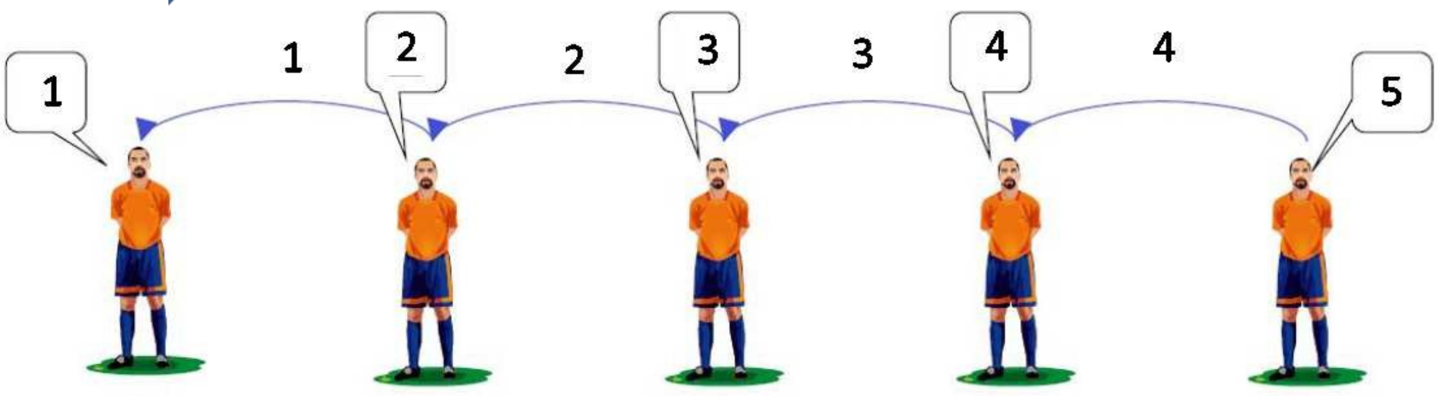
Global information can be shared locally by every entity.
New knowledge is distributed by message passing.

Example: How many people are we?



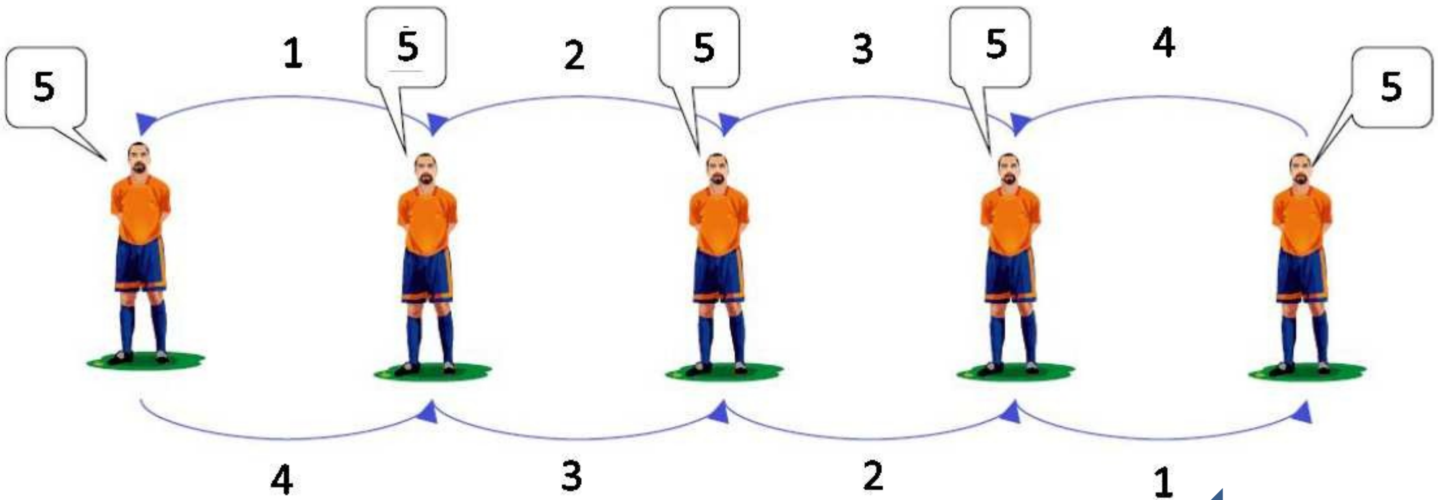
How many people are we?

Forward propagation



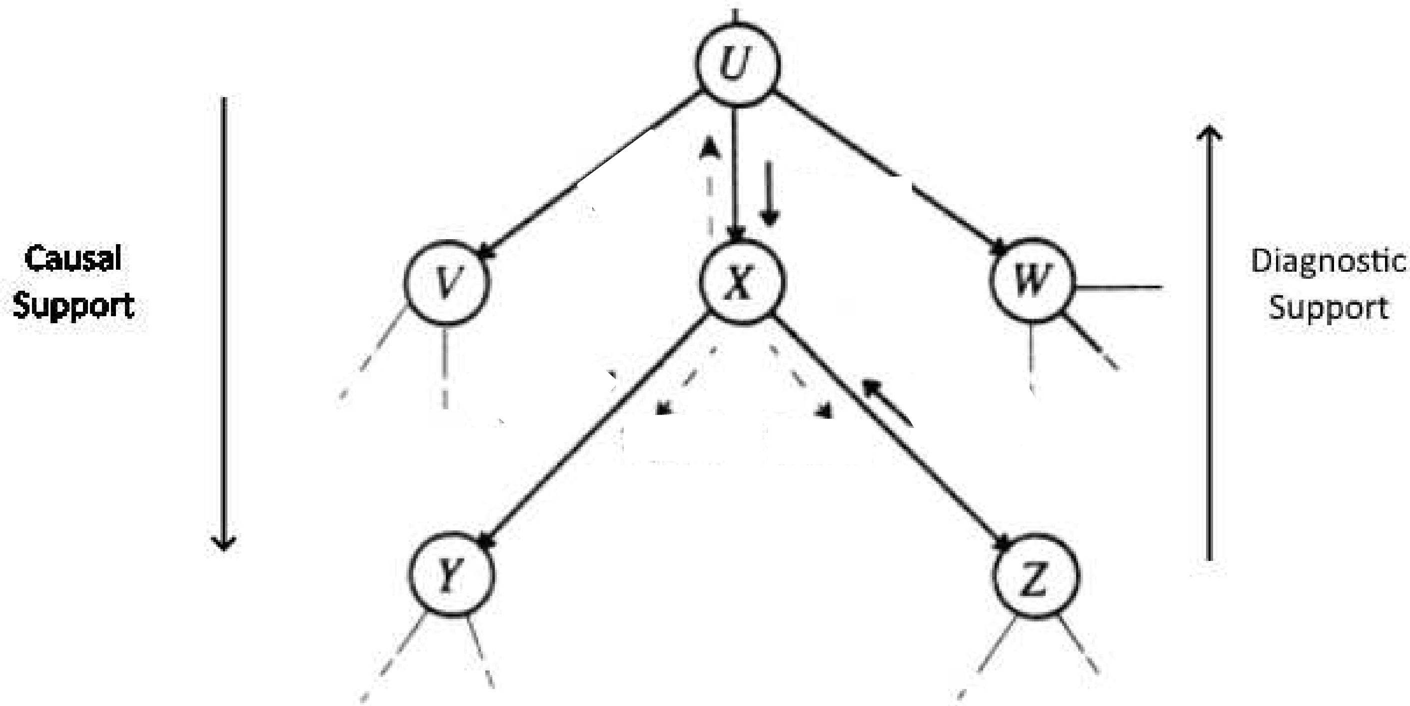
Example 2 Belief Propagation via Message Passing

How many people are we?



backward propagation

Belief Propagation in Trees via Message Passing



λ-message

Let $B \in V$ be an attribute and $C \in s(B)$ its direct children with the respective domains $\text{dom}(B) = \{ b_1, \ldots, b_i, \ldots, b_k\}$ and $\text{dom}(C) = \{c_1, \ldots, c_j, \ldots, c_m\}$.

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

The vector

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

is called λ-*message* from C to B.

λ-Message

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

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

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

π-Message

π-message

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

$$j = 1, \ldots, m :$$

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

The vector

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

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 $b \in \text{dom}(B)$ be a value of B's domain.

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

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

$$\pi_{A \to B}(a_j) = \pi(a_j) \cdot \prod_{C \in s(A) \setminus \{B\}} \lambda_{C \to A}(a_j)$$

$$= \tau_{B,W} \cdot \frac{P(a_j \mid W)}{\lambda_{B \to A}(a_j)}$$

Example 3 Belief Updating

Court Hearing

There are three suspects for murder $X = \{x1, x2, x3\}$

There is a holder of the weapon $Y = \{y1, y2, y3\}$

A fingerprint analysis Z of the weapon is performed Z

Bayes Network

$X \rightarrow Y \rightarrow Z$, $P(X,Y,Z) = P(X) \times P(Y|X) \times P(Z|Y)$

$P(X) = (0.8,0.1,0.1)$, $P(y|x) = 0.8$, if $x=y$, $P(y|x) = 0.1$, if $x \neq y$

The „Who is the murder?“-Belief changes over time in the light of new evidence.

Initial Belief

$\pi(X) = (0.8,0.1,0.1)$, $\lambda(X) = (1,1,1)$ (no diagnostic support), $P(X) = \alpha\pi(X)\lambda(X) = (0.8,0.1,0.1)$

$$\pi(Y) = (0.8,0.1,0.1) \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{pmatrix} = (0.66,0.17,0.17) , \lambda(Y) = (1,1,1) , P(Y) = (0.66,0.17,0.17)$$

For calculations it is convenient if the numbers in the vectors are not normalized, so we use again normalization factors. The final result is given in normalized form.

Example 3 Belief Updating

Updated belief due to the fingerprint analysis (evidential support)

Z passes diagnostic knowledge to Y in form of a new λ of Y: $\lambda(Y) = \alpha\,(0.8,0.6,0.5)$

$\pi(Y) = (\,0.66,0.17,0.17)$ does not change,

The new belief is P(Y)= $\alpha\pi(Y)\lambda(Y)=\alpha(08,0.6,0.5)(0.66,0.17,0.17) = (0.738,0,142,0.119)$

Y passes the new knowledge to X, where the conditional probabilities have to be used:

$$\lambda(X) = \beta \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{pmatrix} \begin{pmatrix} 0.8 \\ 0.6 \\ 0.5 \end{pmatrix} = \beta\,(0.75,0.61,0.54)$$

The new belief is P (X) = $\alpha(0.75,0.61,0.54)\ (0.8,0.1,0.1) = (0.840,0.085,0.075)$

Note that P is the updated belief, in general it is different from the prior probability P

Example 3 Belief Updating

Updated belief due to a rather strong alibi for suspect 1 (causal support)

The prior probability for X is revised to $\pi(X) = (0.28,0.36,0.36)$

$\lambda(X) = (0.75,0.61,0.54)$ is the same as before, $P(X) = (0.343,0.349,0.308)$

The revised prior is forwarded to Y

$$\pi(Y) = (0.28,0.36,0.36)\begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{pmatrix} = (0.30,0.35,0.35),\ \lambda(Y) = (0.75,0.61,0.54),$$

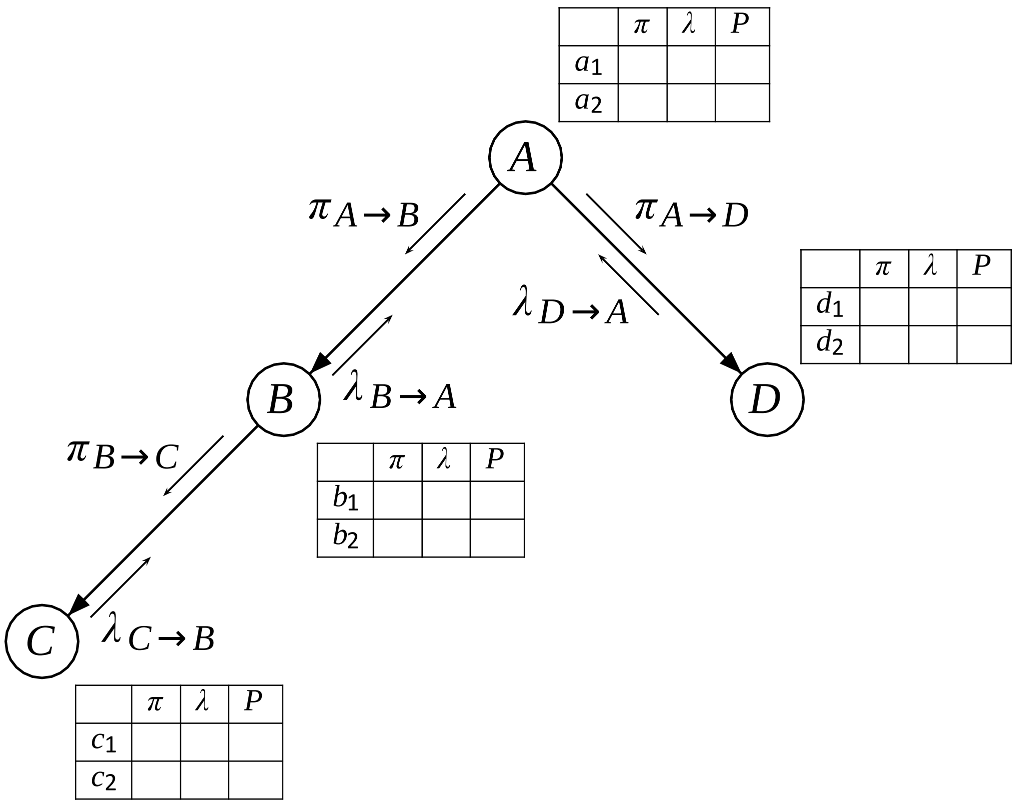
The new belief is $P(Y) = \alpha(0.3,0.35,0.35)\,(0.8,0.6,0.5) = (0.384,0.336,0.280)$

Thus suspect 2 becomes the strongest candidate for being the killer: $P(X=2) = 0.349$.

Note that suspect 1 is still more likely to be the owner of the fingerprint: $P(Y=1) = 0.384$

Example 4: Propagation in Belief Trees (1)

Belief Tree



Conditional Probabilities

$P(a_1) = 0.1$ $\quad P(b_1 \mid a_1) = 0.7$

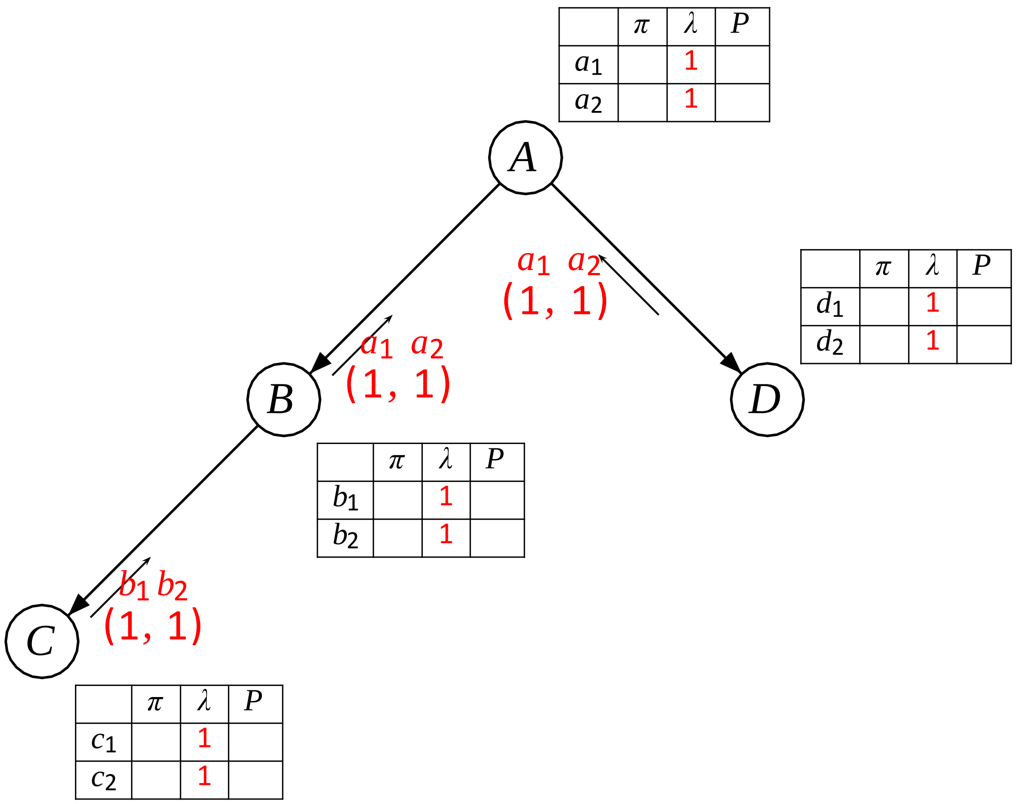
$P(b_1 \mid a_2) = 0.2$ $\quad P(d_1 \mid a_1) = 0.8$

$P(c_1 \mid b_1) = 0.4$

$P(d_1 \mid a_2) = 0.4$ $\quad P(c_1 \mid b_2) = 0.001$

Note that in the tables the P represents the „current“ belief, it changes.

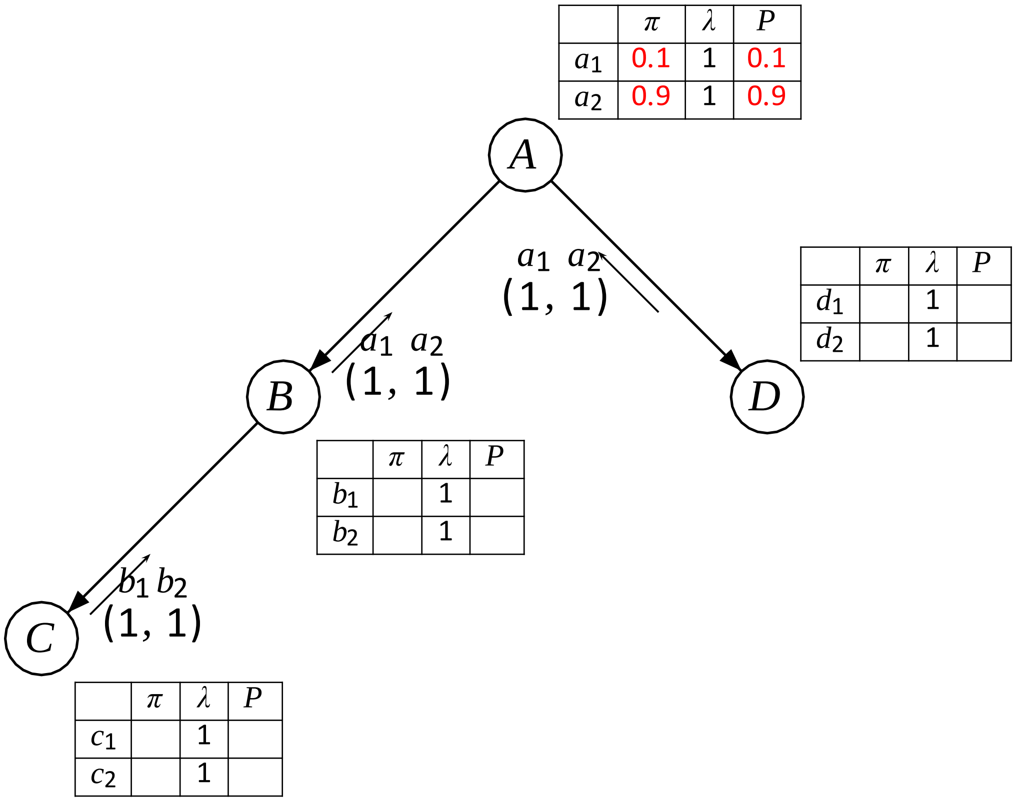
Belief Tree:



Initialization Phase:

Set all λ-messages and λ-values to 1.

Belief Tree:

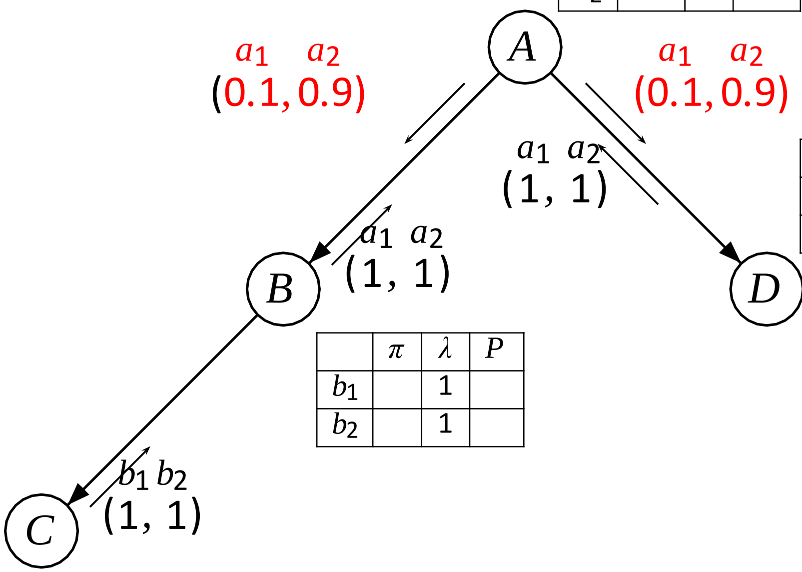


Initialization Phase:

Set all λ-messages and λ-values to 1.

$\pi(a_1) = P(a_1)$ and $\pi(a_2) = P(a_2)$

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	

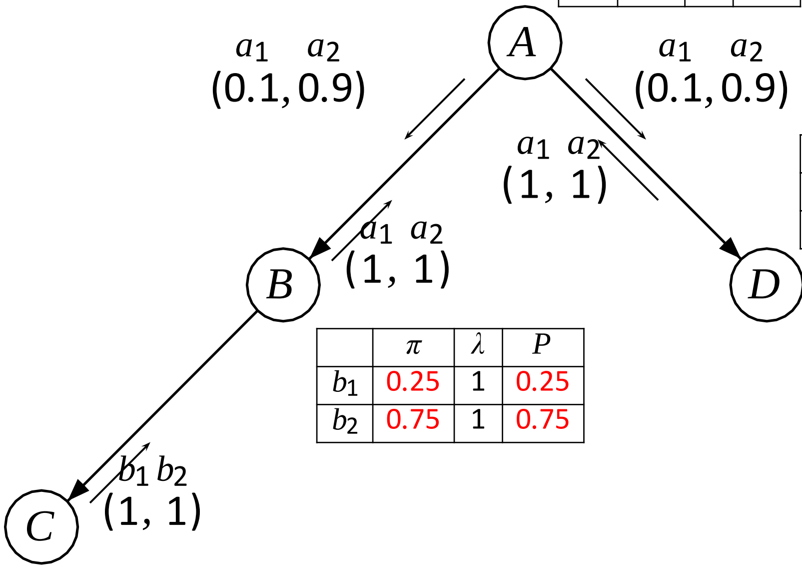
	π	λ	P
c_1		1	
c_2		1	

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.

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	

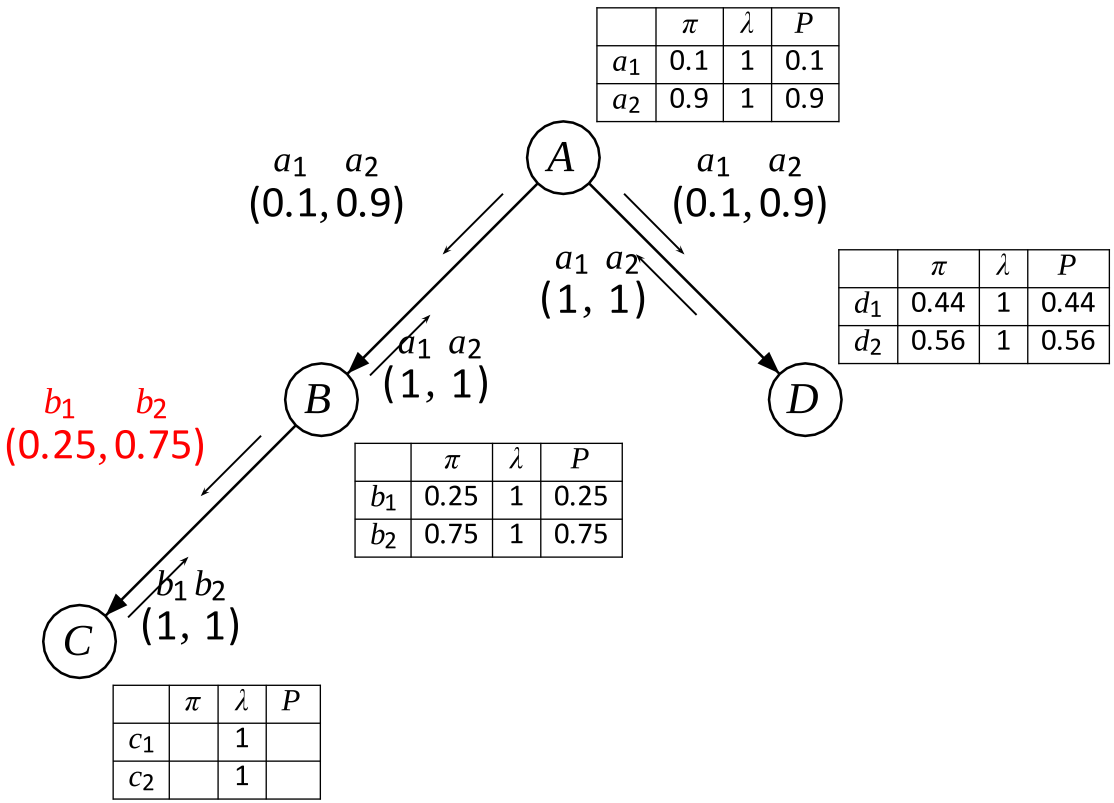
	π	λ	P
d_1	0.44	1	0.44
d_2	0.56	1	0.56

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.

Belief Tree:



Initialization Phase:

Set all λ-messages and λ-values to 1.

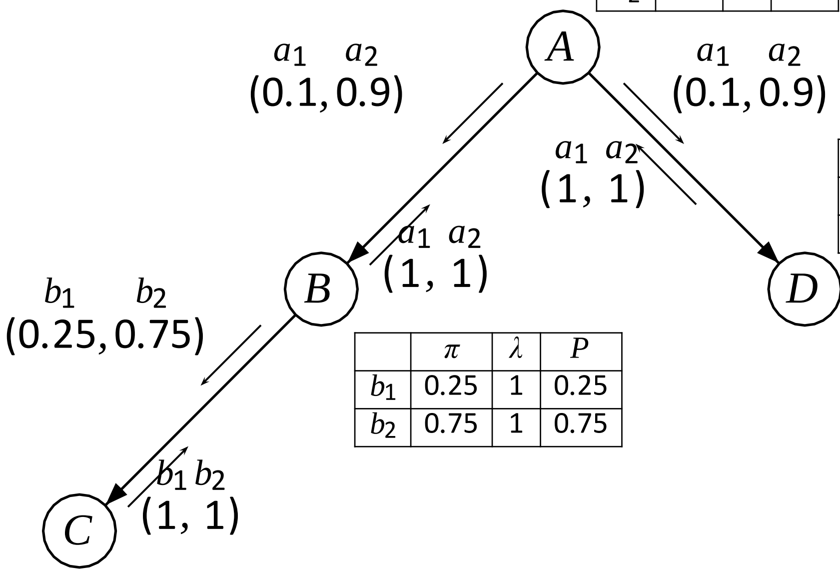
$\pi(a_1) = P(a_1)$ and $\pi(a_2) = P(a_2)$.

A sends π-messages to B and D.

B and D update their π-values.

B sends π-message to C.

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 $\pi(a_2) = P(a_2)$.

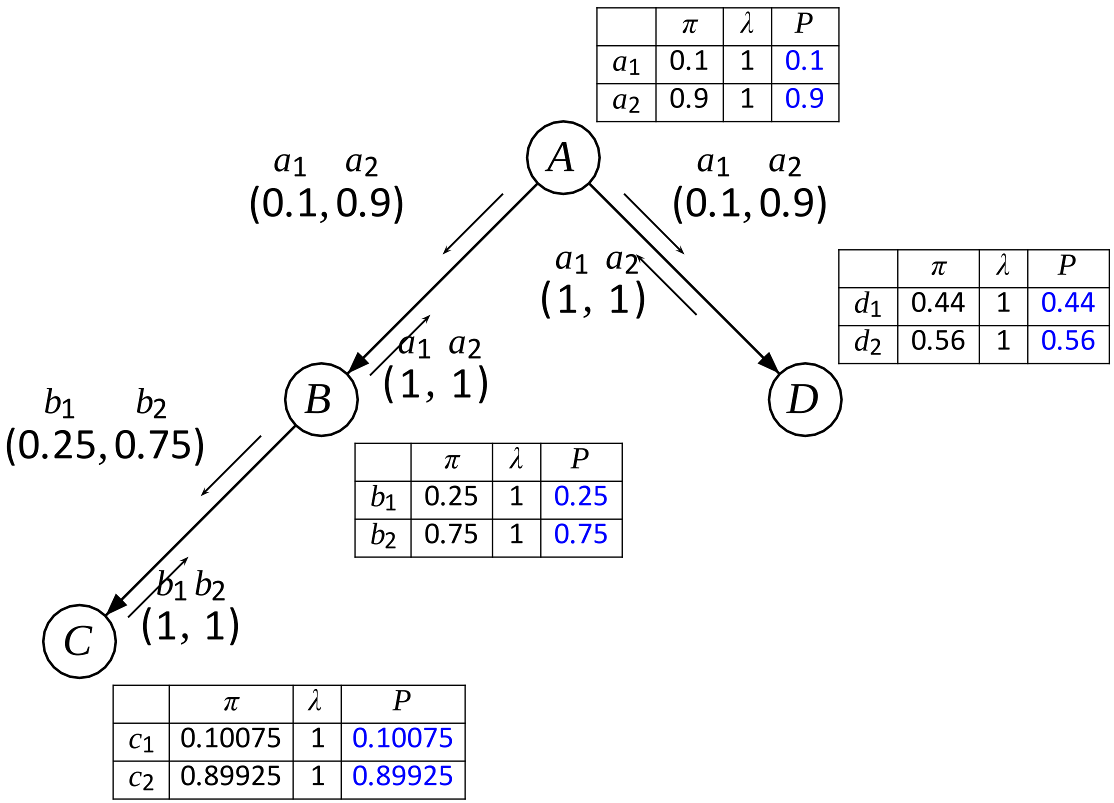
A sends π-messages to B and D.

B and D update their π-values.

B sends π-message to C.

C updates it π-value.

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

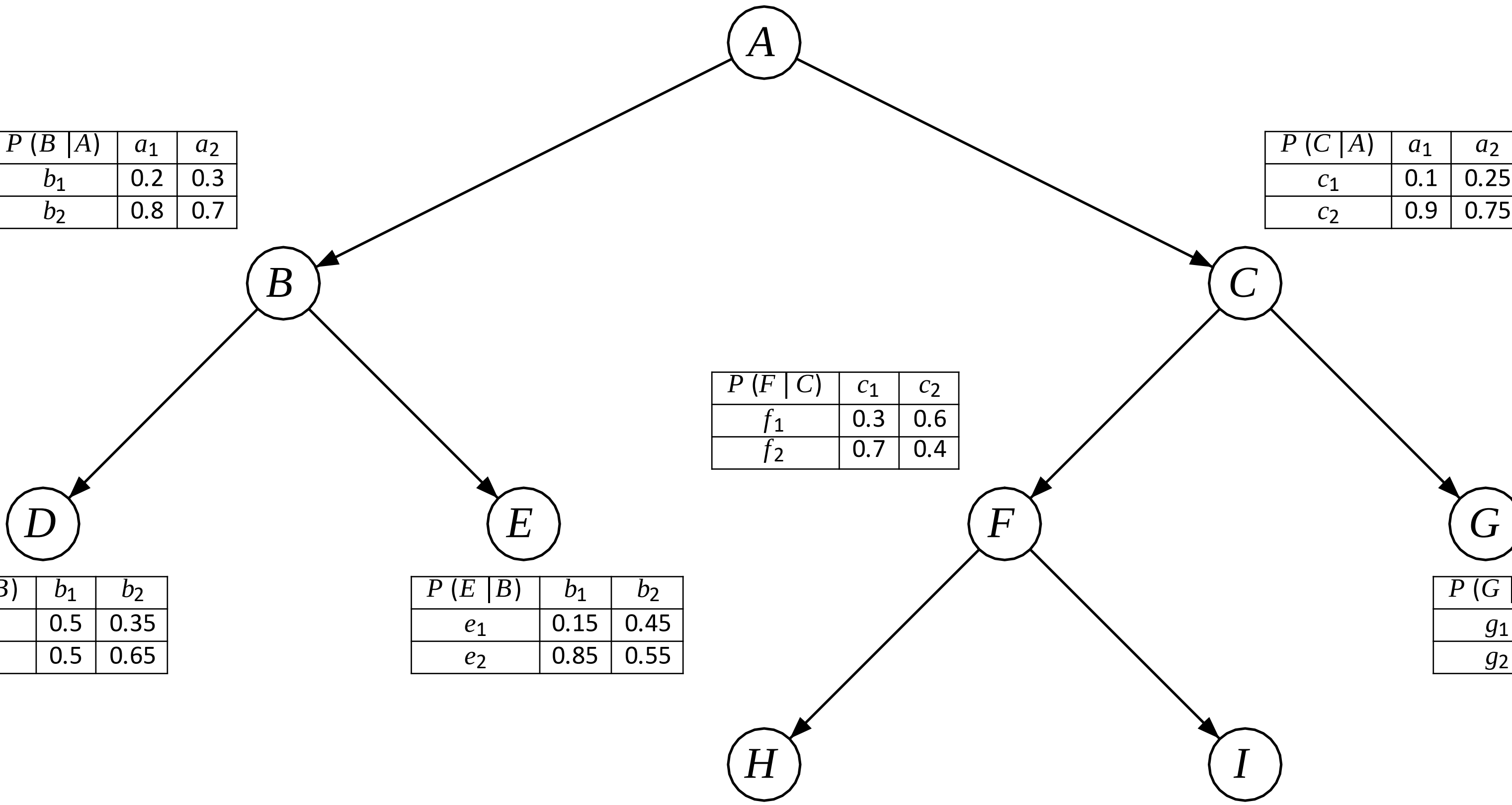
	π	λ	P
c_1	0.10075	1	0.10075
c_2	0.89925	1	0.89925

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.

Example 5: Larger Tree (1)



$P(A)$	$\emptyset$
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

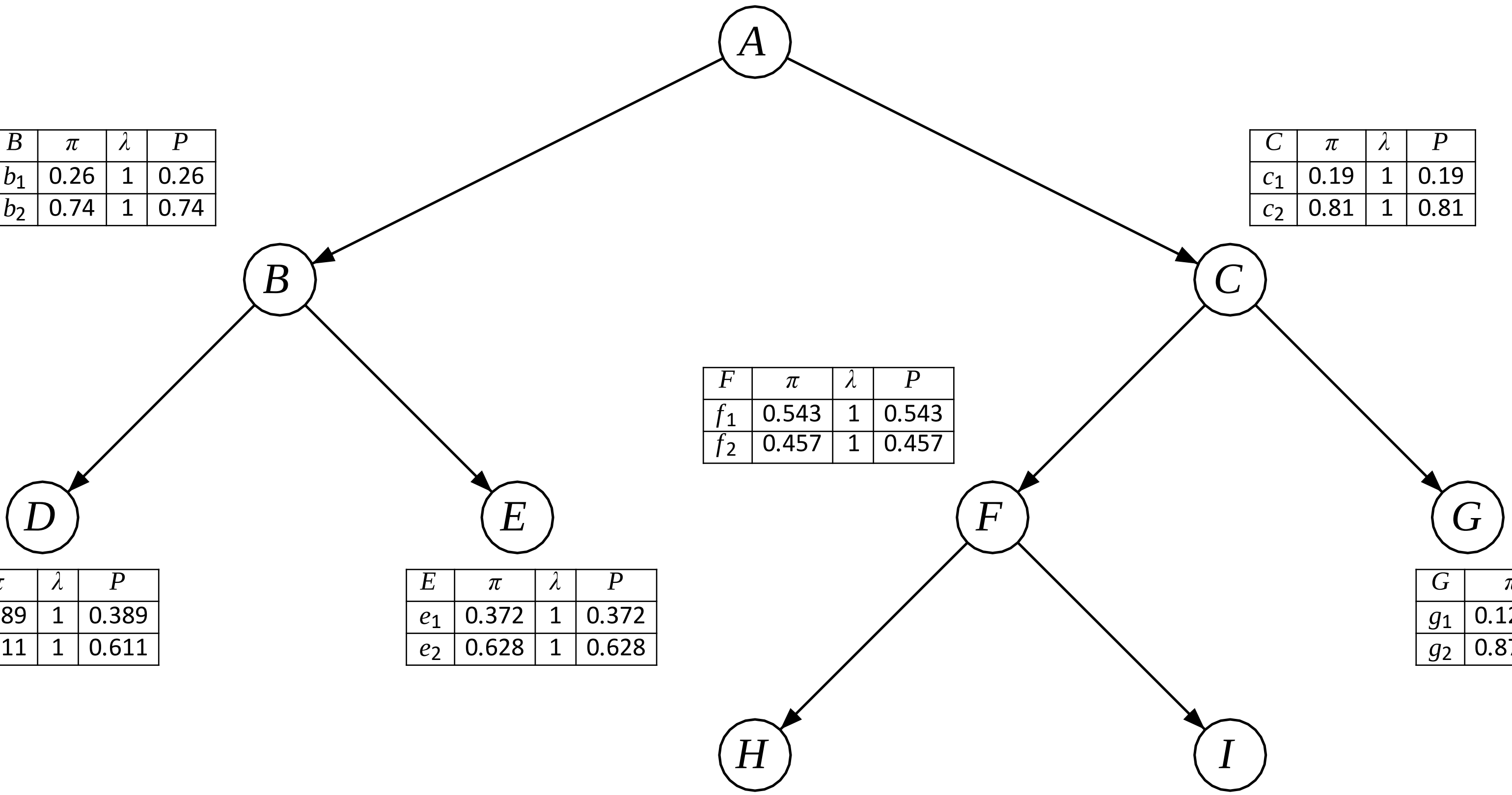
$P(E \mid B)$	b_1	b_2
e_1	0.15	0.45
e_2	0.85	0.55

$P(G \mid C)$	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 Tree (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

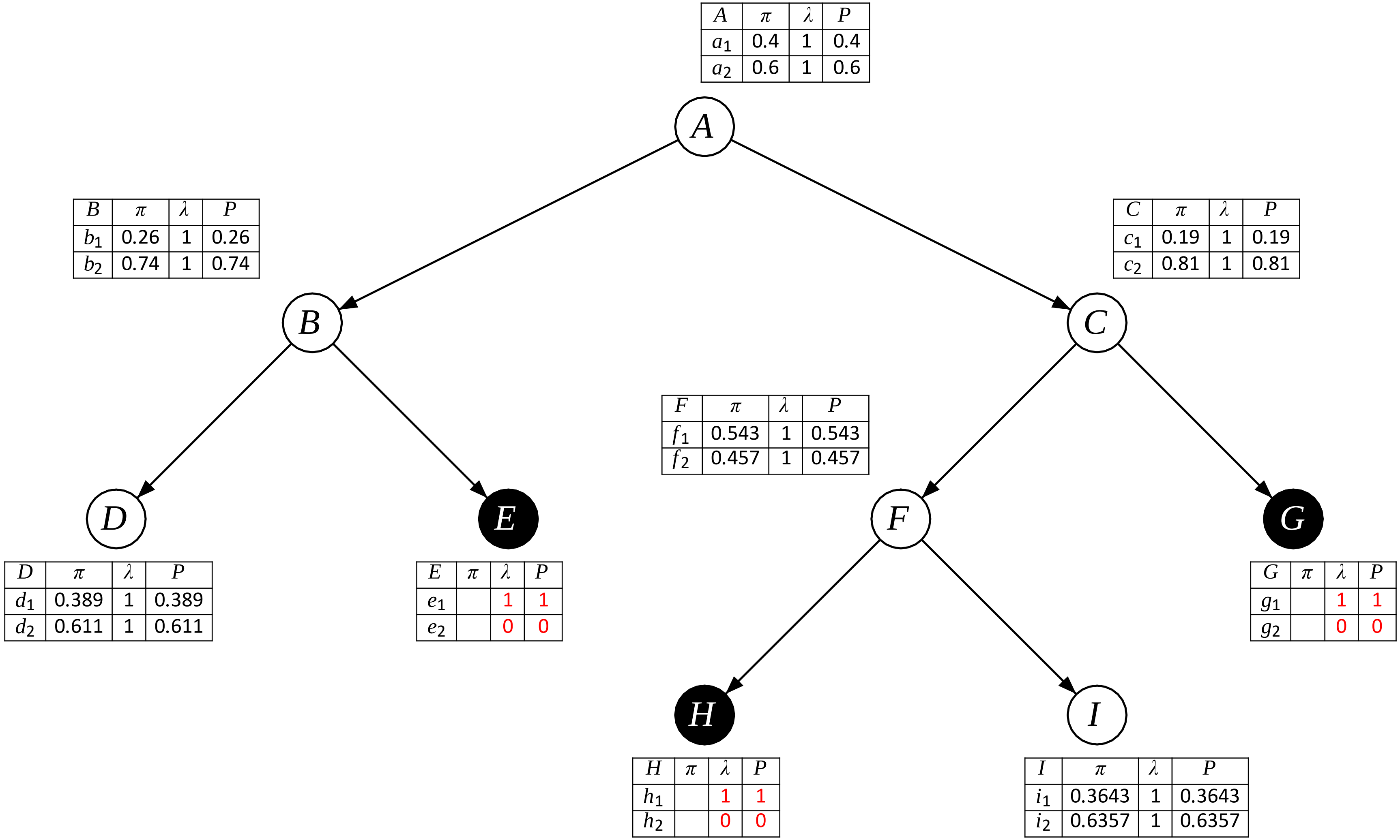
E	π	λ	P
e_1	0.372	1	0.372
e_2	0.628	1	0.628

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 Tree (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

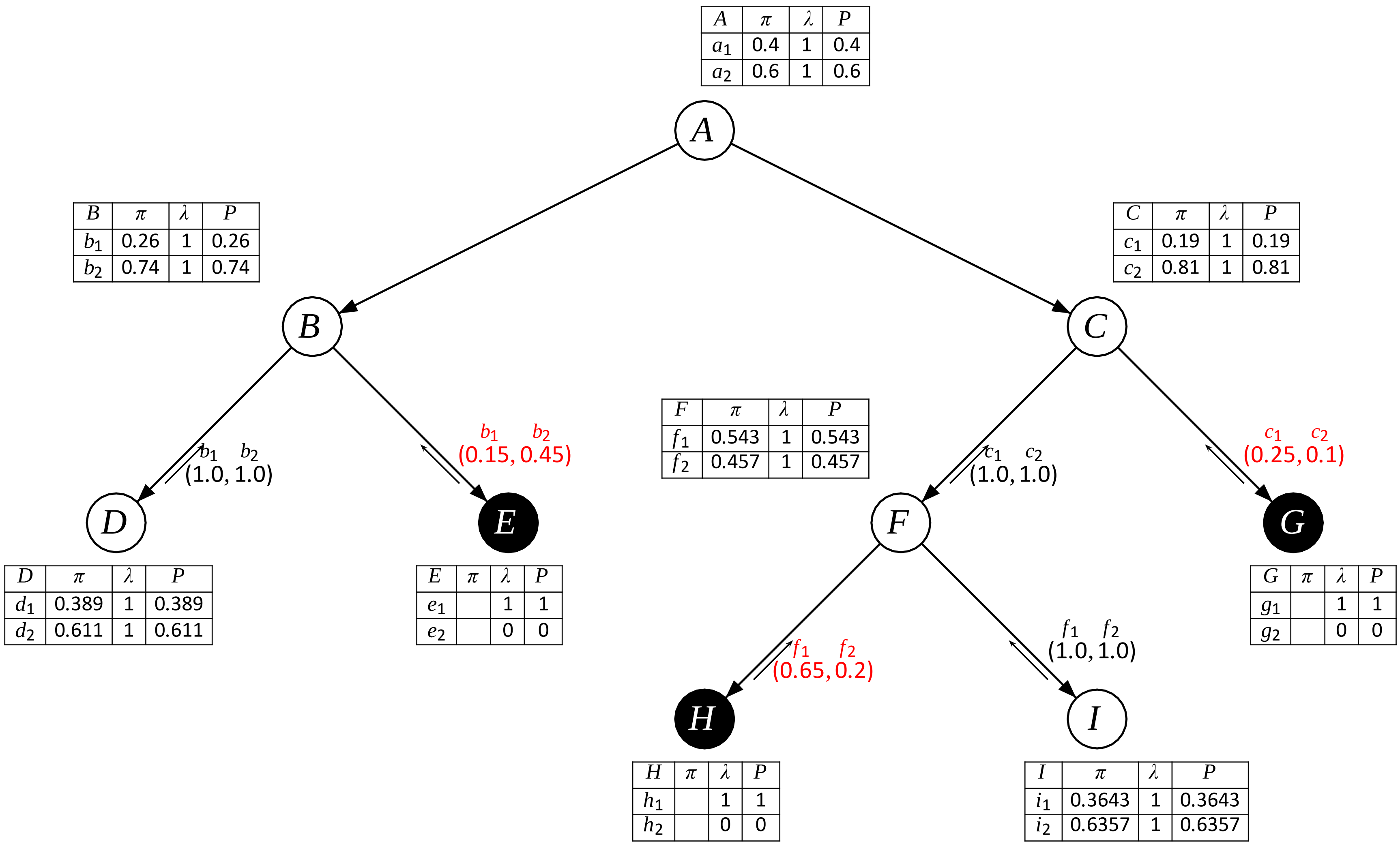
D	π	λ	P
d_1	0.389	1	0.389
d_2	0.611	1	0.611

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



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

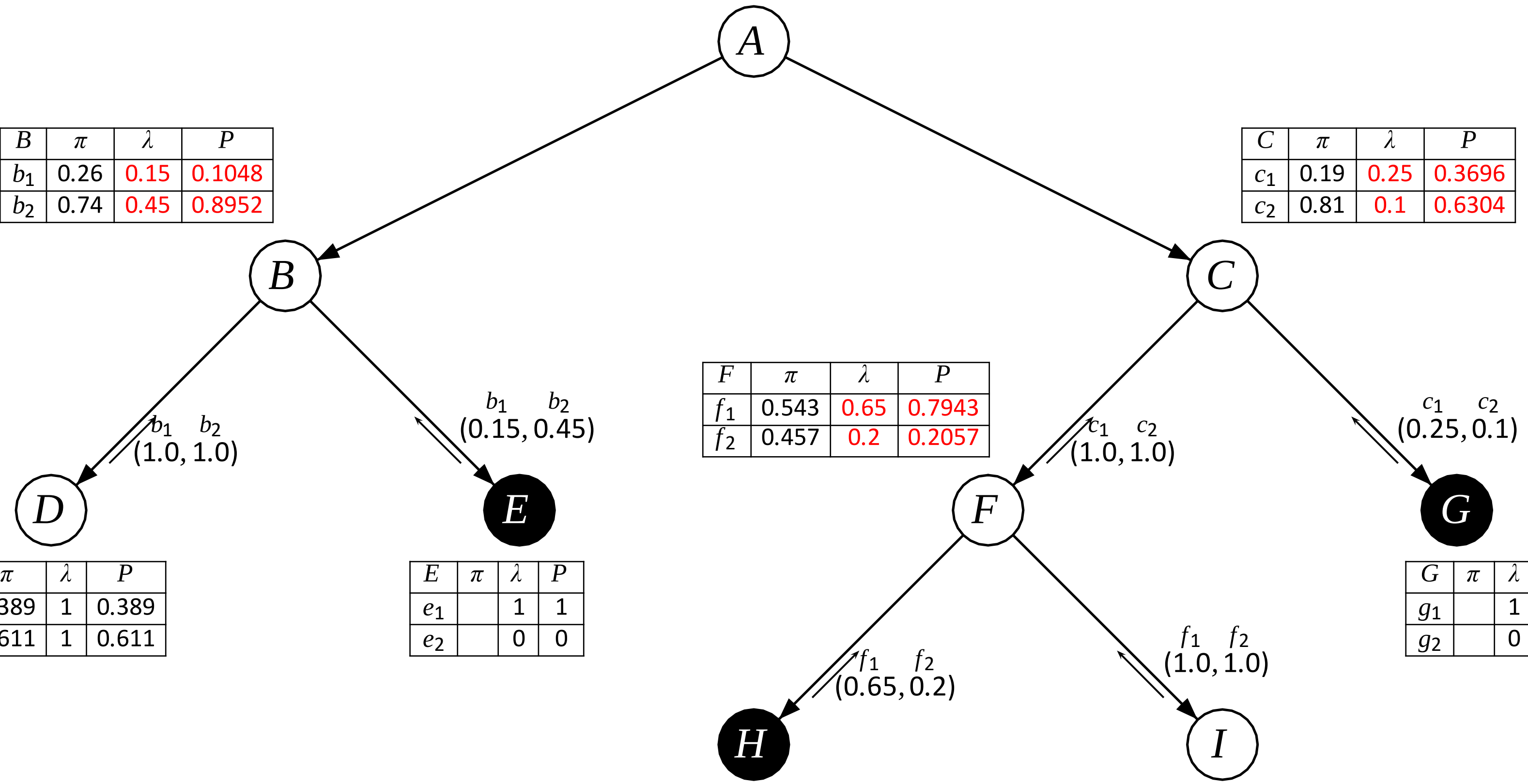
E	π	λ	P
e_1		1	1
e_2		0	0

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 Tree (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

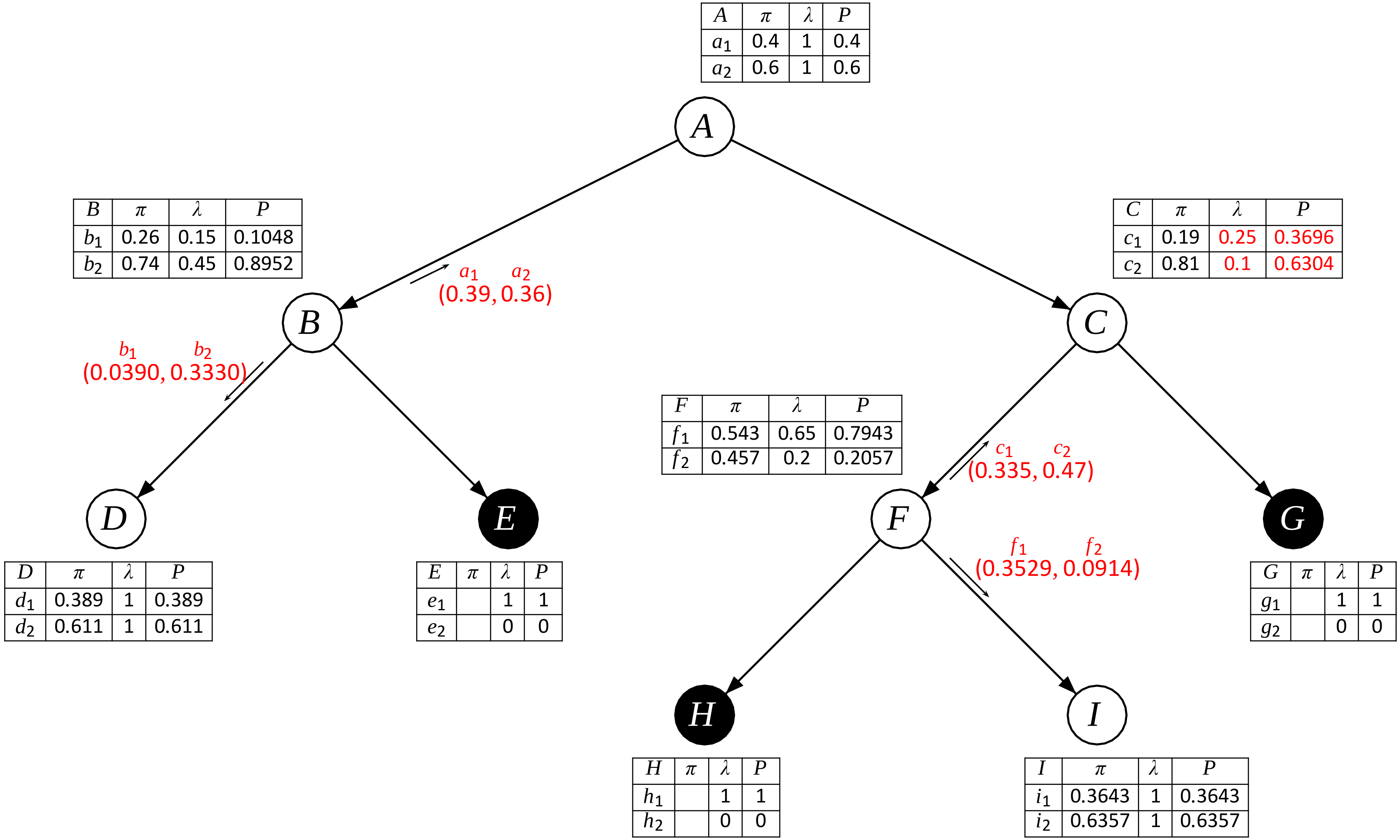
E	π	λ	P
e_1		1	1
e_2		0	0

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 Tree (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

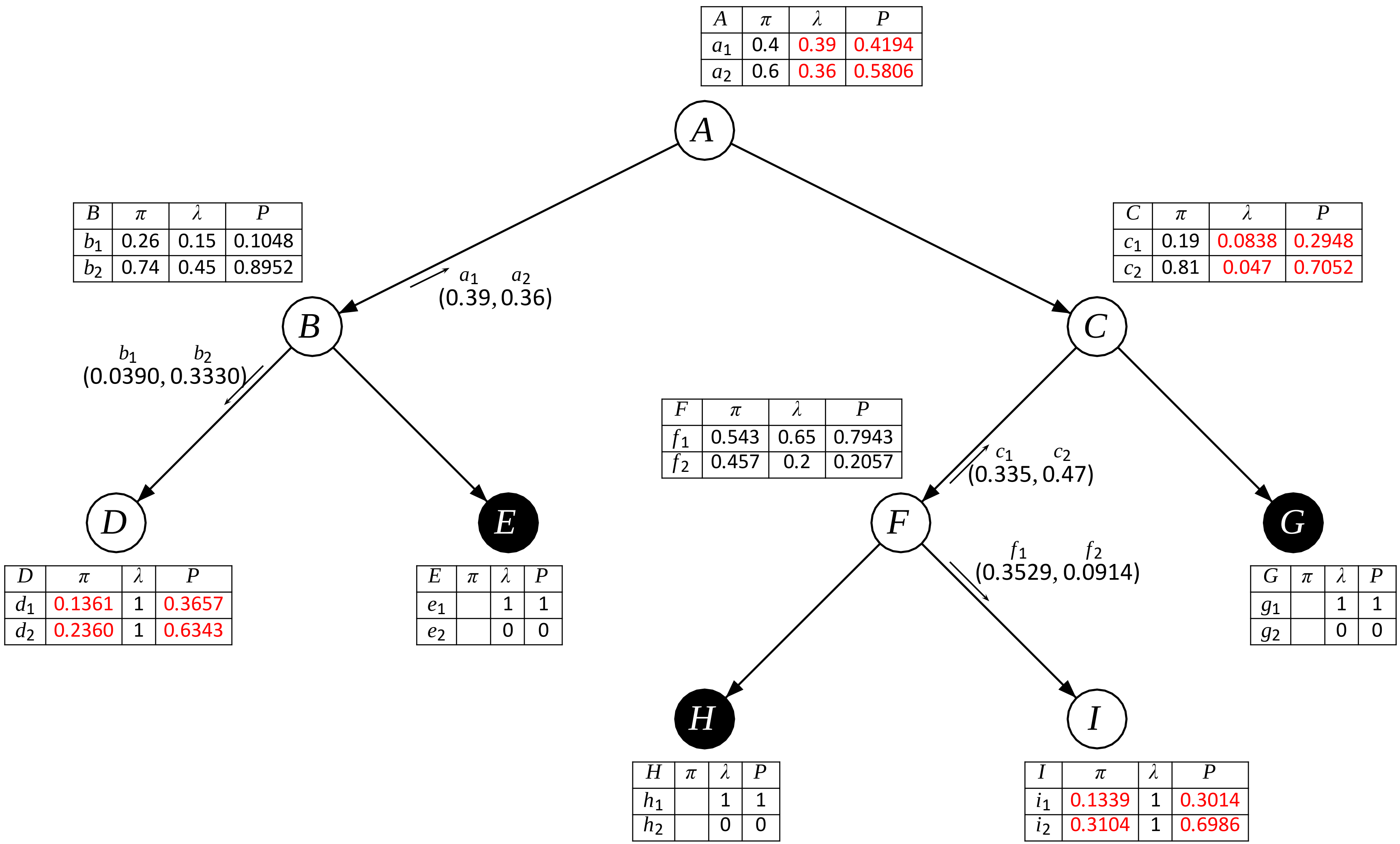
D	π	λ	P
d_1	0.389	1	0.389
d_2	0.611	1	0.611

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



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	0.1361	1	0.3657
d_2	0.2360	1	0.6343

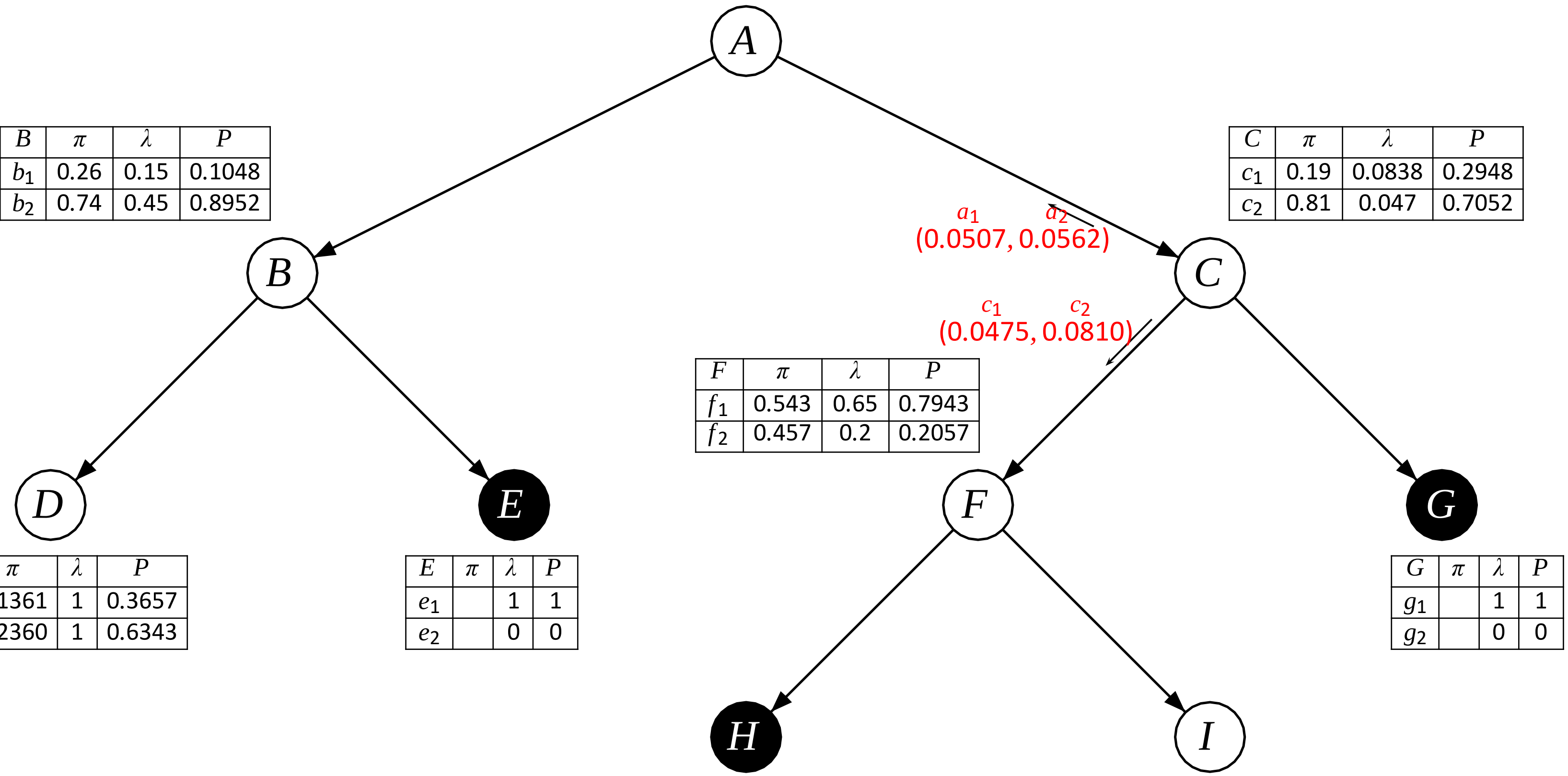
E	π	λ	P
e_1		1	1
e_2		0	0

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	0.1339	1	0.3014
i_2	0.3104	1	0.6986

Larger Tree (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	0.1361	1	0.3657
d_2	0.2360	1	0.6343

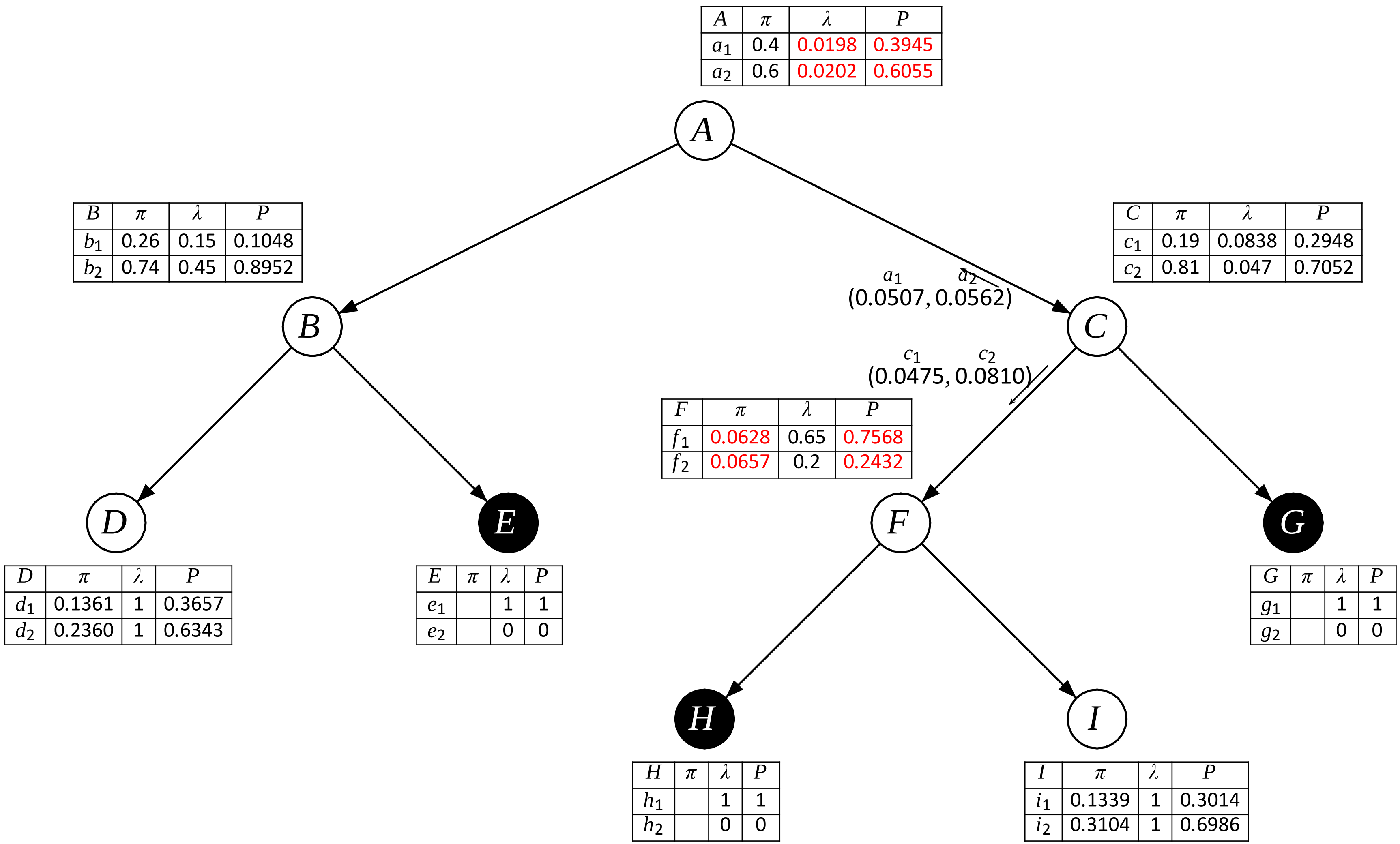
E	π	λ	P
e_1		1	1
e_2		0	0

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	0.1339	1	0.3014
i_2	0.3104	1	0.6986

Larger Tree (9): 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	0.0628	0.65	0.7568
f_2	0.0657	0.2	0.2432

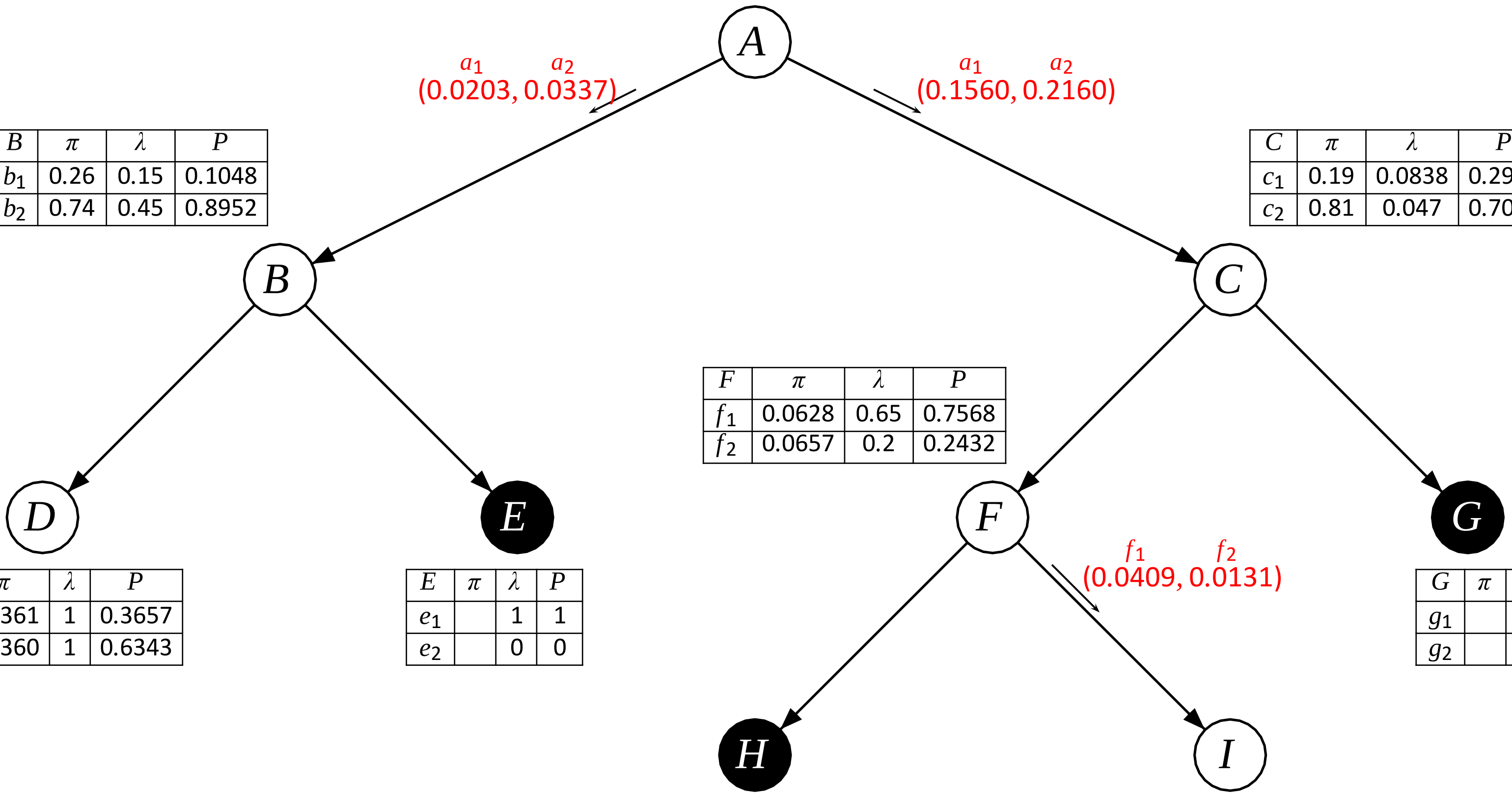
D	π	λ	P
d_1	0.1361	1	0.3657
d_2	0.2360	1	0.6343

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	0.1339	1	0.3014
i_2	0.3104	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	0.0628	0.65	0.7568
f_2	0.0657	0.2	0.2432

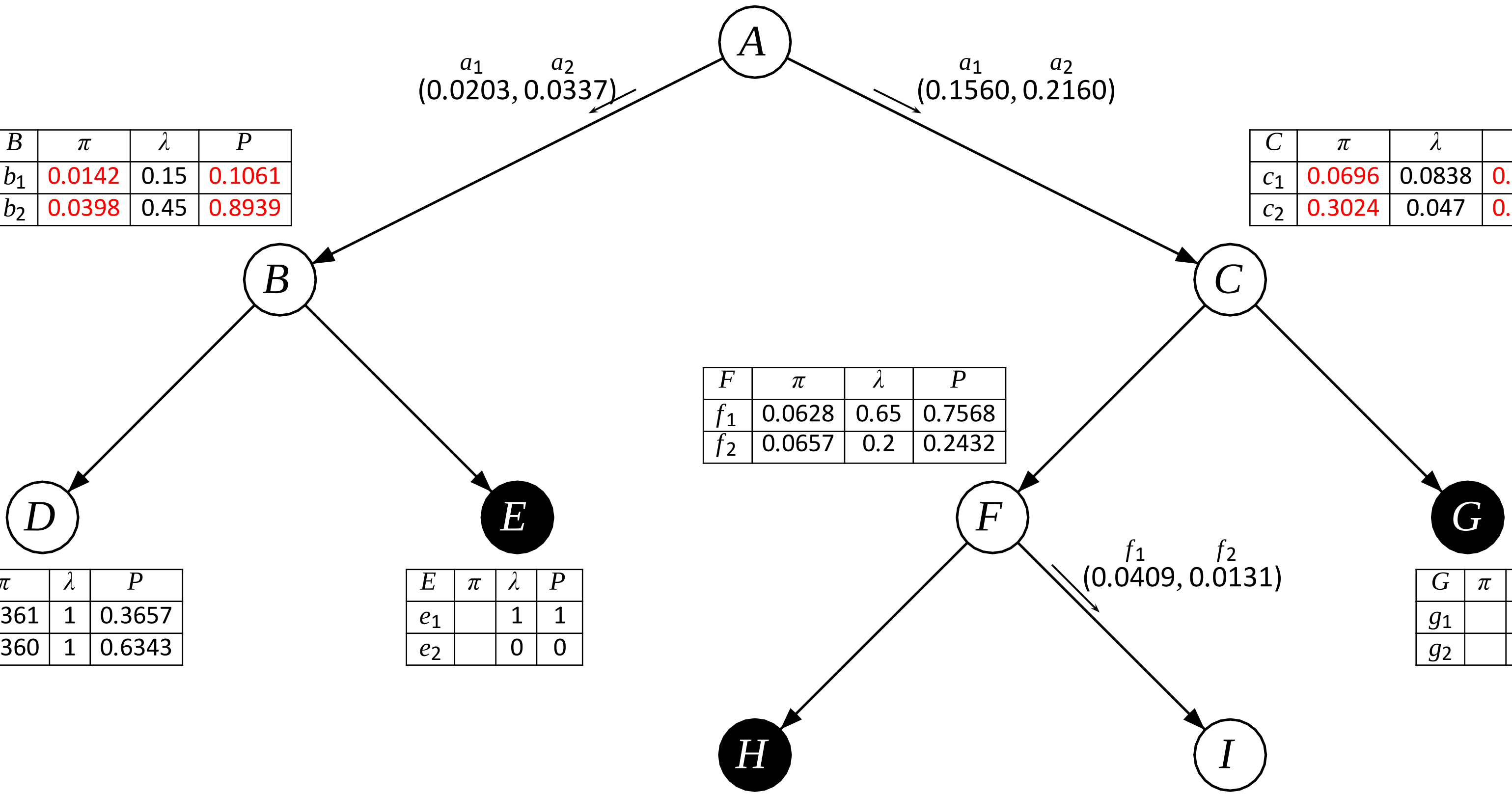
D	π	λ	P
d_1	0.1361	1	0.3657
d_2	0.2360	1	0.6343

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	0.1339	1	0.3014
i_2	0.3104	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	0.0142	0.15	0.1061
b_2	0.0398	0.45	0.8939

C	π	λ	P
c_1	0.0696	0.0838	0.2910
c_2	0.3024	0.047	0.7090

F	π	λ	P
f_1	0.0628	0.65	0.7568
f_2	0.0657	0.2	0.2432

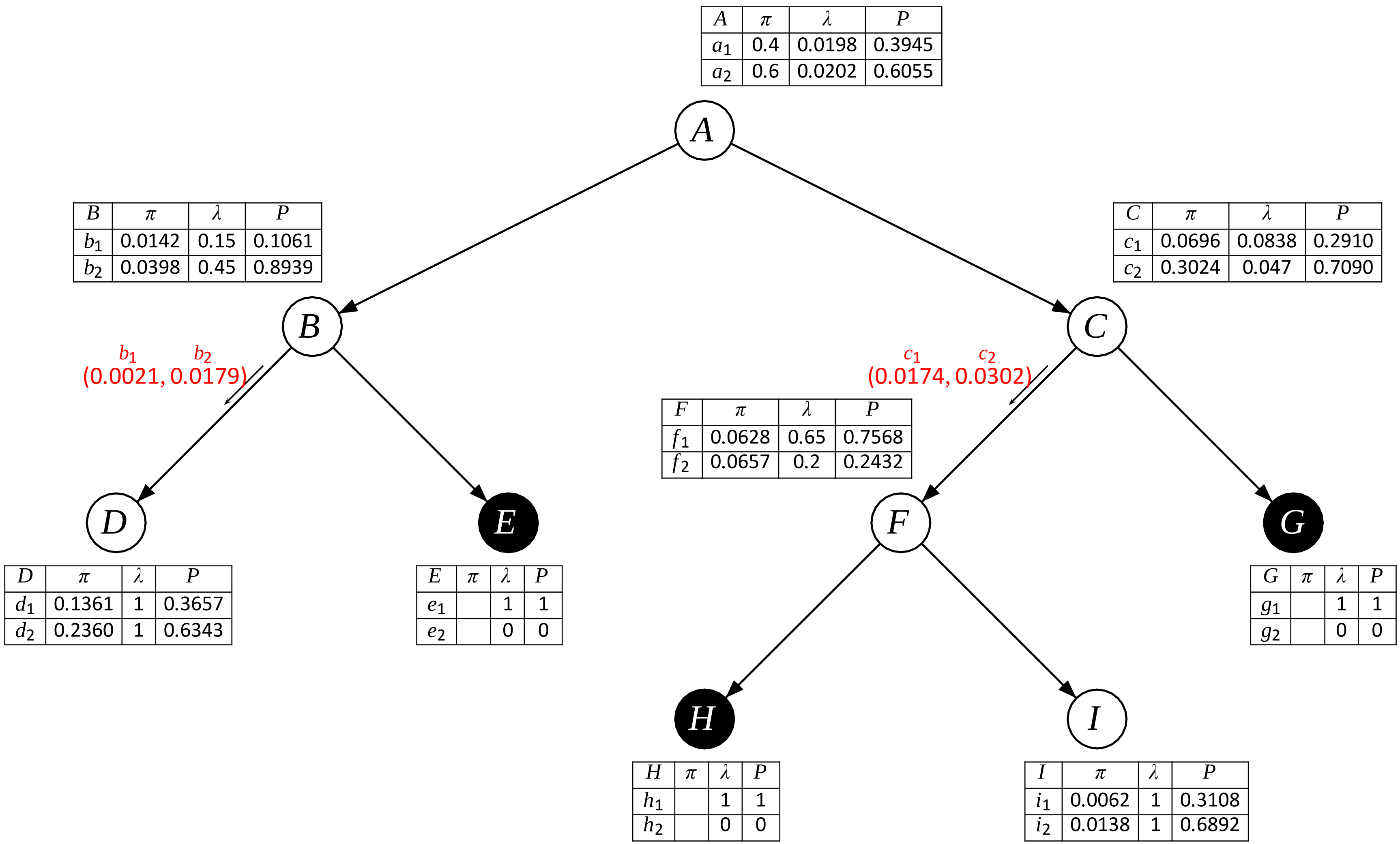
D	π	λ	P
d_1	0.1361	1	0.3657
d_2	0.2360	1	0.6343

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	0.0062	1	0.3108
i_2	0.0138	1	0.6892



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	0.0142	0.15	0.1061
b_2	0.0398	0.45	0.8939

C	π	λ	P
c_1	0.0696	0.0838	0.2910
c_2	0.3024	0.047	0.7090

F	π	λ	P
f_1	0.0628	0.65	0.7568
f_2	0.0657	0.2	0.2432

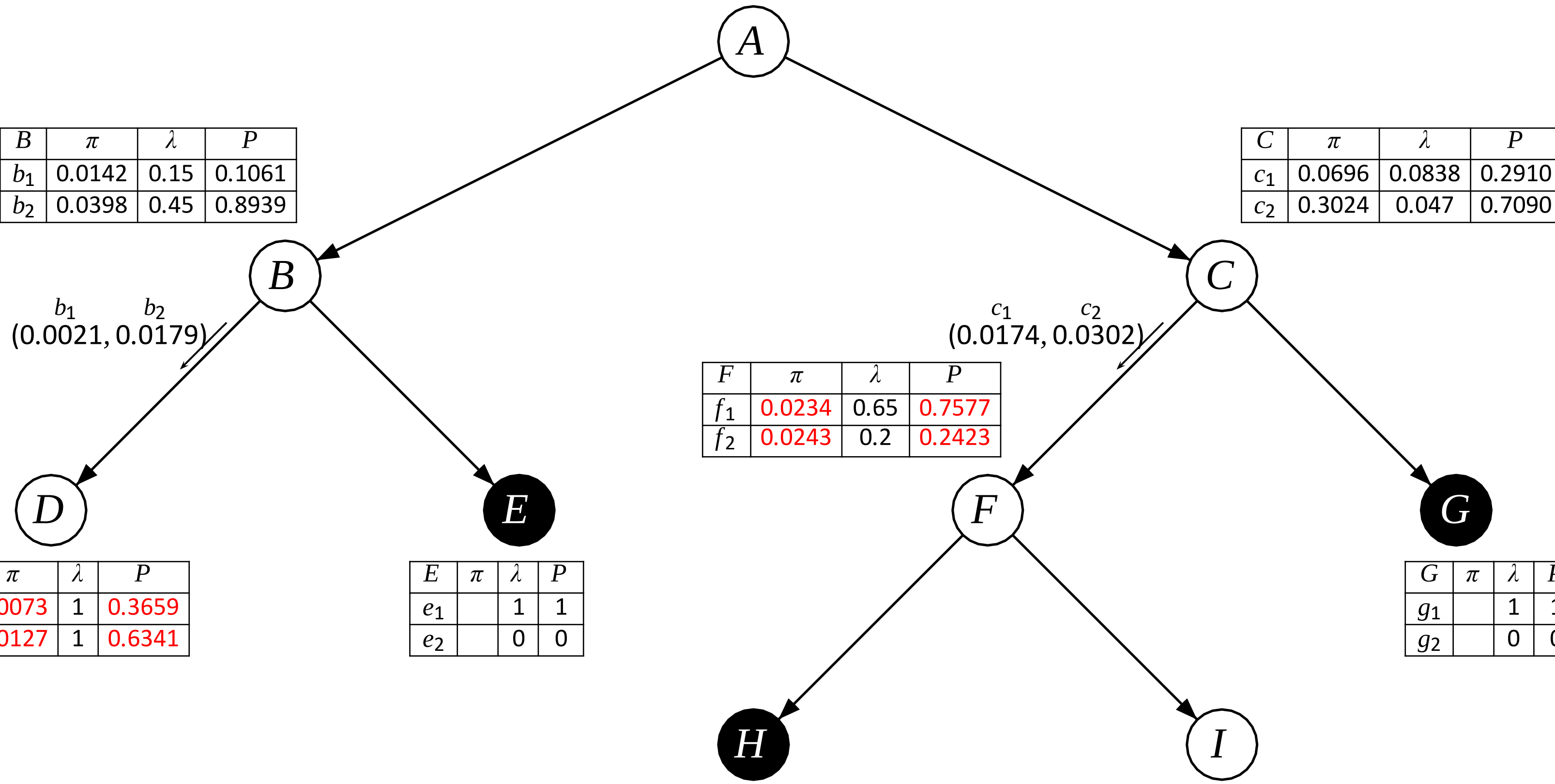
D	π	λ	P
d_1	0.1361	1	0.3657
d_2	0.2360	1	0.6343

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	0.0062	1	0.3108
i_2	0.0138	1	0.6892



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	0.0142	0.15	0.1061
b_2	0.0398	0.45	0.8939

C	π	λ	P
c_1	0.0696	0.0838	0.2910
c_2	0.3024	0.047	0.7090

F	π	λ	P
f_1	0.0234	0.65	0.7577
f_2	0.0243	0.2	0.2423

D	π	λ	P
d_1	0.0073	1	0.3659
d_2	0.0127	1	0.6341

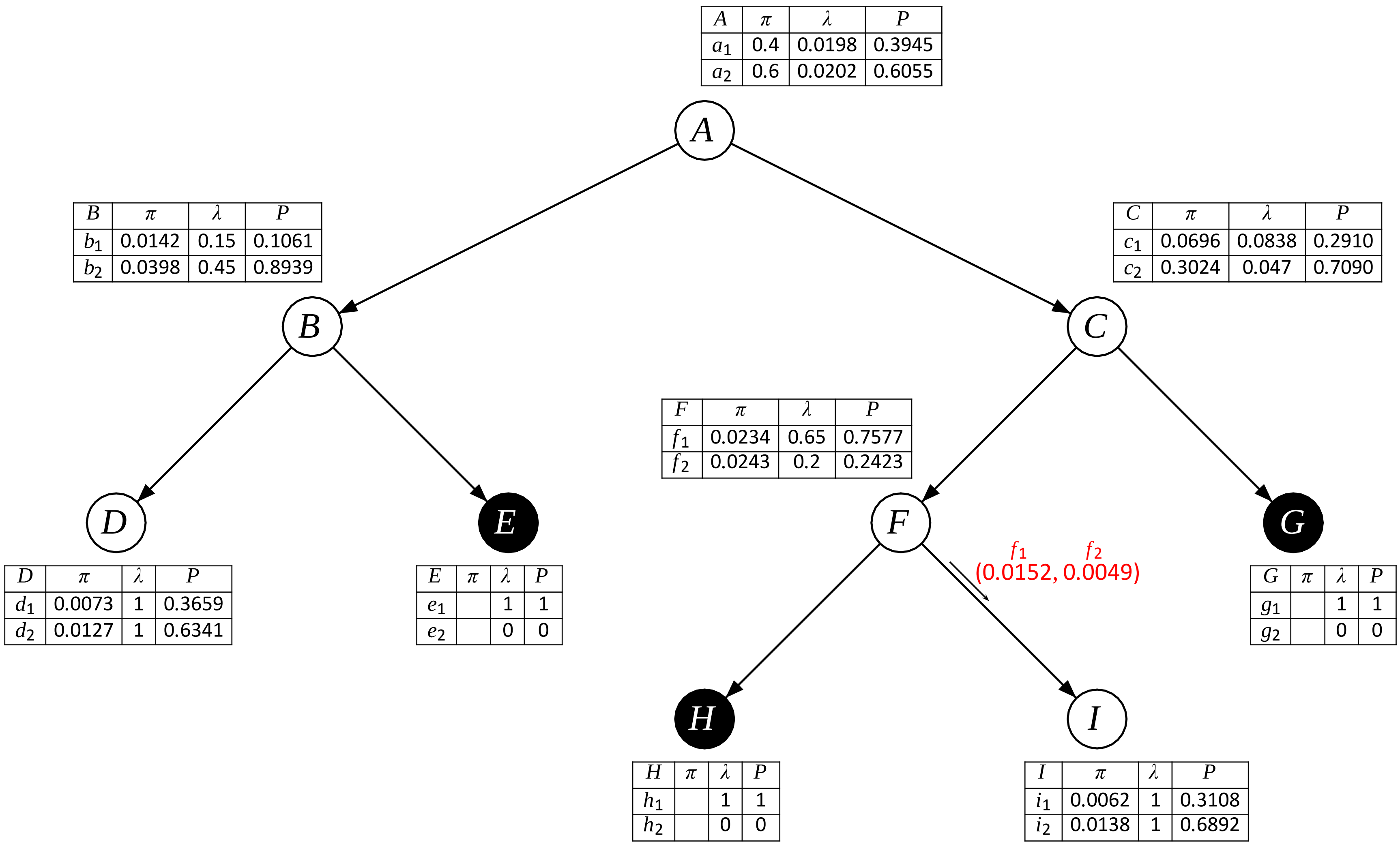
E	π	λ	P
e_1		1	1
e_2		0	0

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	0.0062	1	0.3108
i_2	0.0138	1	0.6892

Larger Tree (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	0.0142	0.15	0.1061
b_2	0.0398	0.45	0.8939

C	π	λ	P
c_1	0.0696	0.0838	0.2910
c_2	0.3024	0.047	0.7090

F	π	λ	P
f_1	0.0234	0.65	0.7577
f_2	0.0243	0.2	0.2423

D	π	λ	P
d_1	0.0073	1	0.3659
d_2	0.0127	1	0.6341

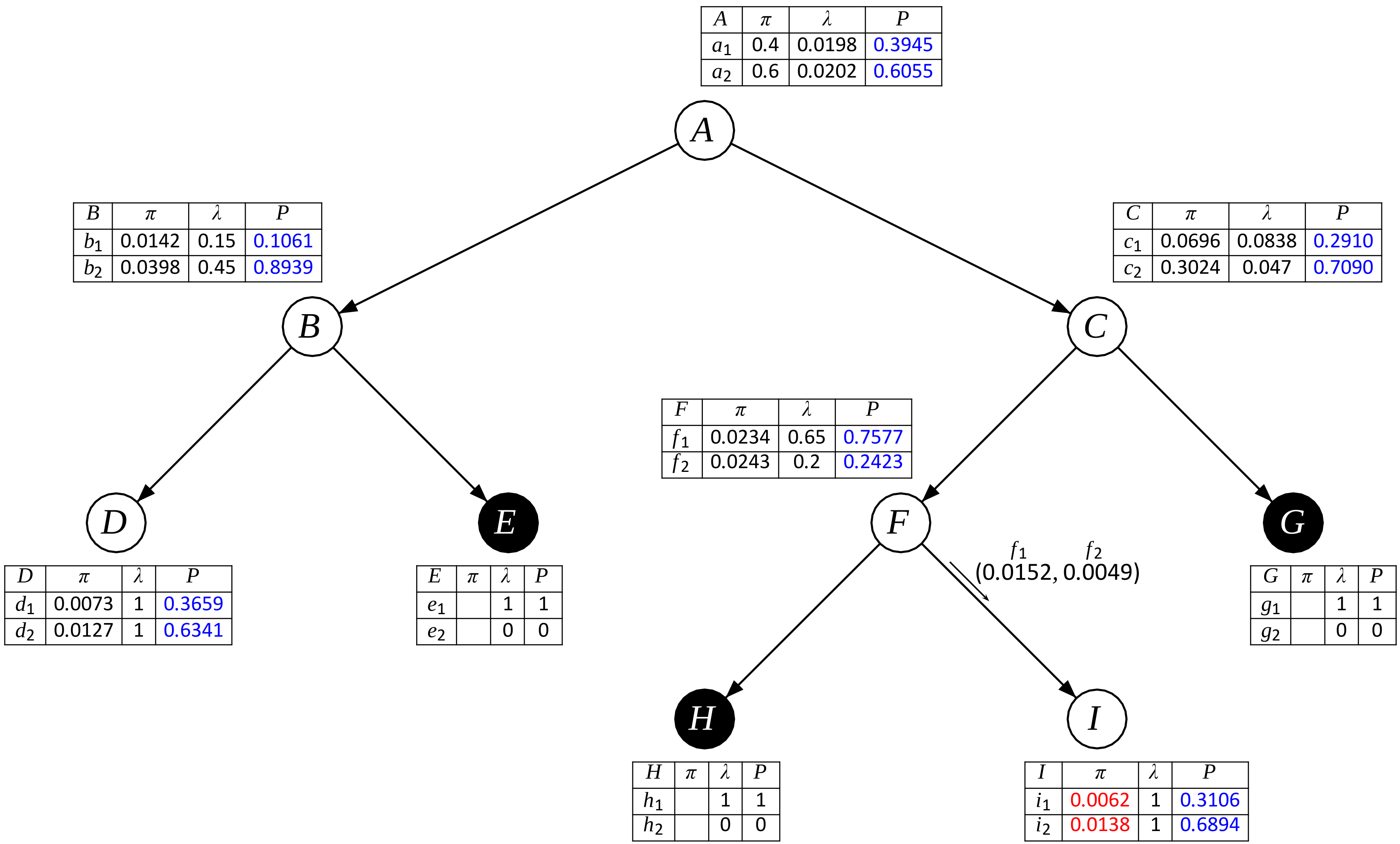
E	π	λ	P
e_1		1	1
e_2		0	0

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	0.0062	1	0.3108
i_2	0.0138	1	0.6892

Larger Tree (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	0.0142	0.15	0.1061
b_2	0.0398	0.45	0.8939

C	π	λ	P
c_1	0.0696	0.0838	0.2910
c_2	0.3024	0.047	0.7090

F	π	λ	P
f_1	0.0234	0.65	0.7577
f_2	0.0243	0.2	0.2423

D	π	λ	P
d_1	0.0073	1	0.3659
d_2	0.0127	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	0.0062	1	0.3106
i_2	0.0138	1	0.6894

Complexity of Belief Propagation in Trees

Properties of the updating scheme:

- efficient in storage and time
- m-ary tree with n states per node:
- every node stores $n^2 + mn + 2n$ real numbers and performes $2n^2 + mn + 2n$ multiplications per update

- The time complexity to obtain the posterior probability of all the variables in the tree is proportional to the diameter of the network (the number of arcs in the trajectory from the root to the most distant leaf).

The message passing mechanism can be directly extended to polytrees, as these are also singly connected networks. In this case, a node can have multiple parents, so the λ messages should be sent from a node to all its parents.

Example 6

answering probabilistic queries

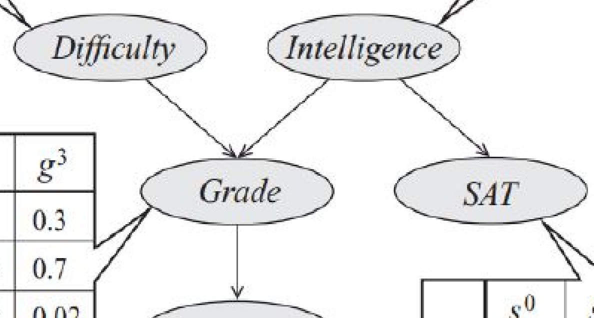
$$P(\mathbf{Y} = \mathbf{y} \mid \mathbf{E} = e) \quad ?$$

evidence

$$P(L = l^1 \mid S = s^1) = \frac{P(L=l^1, S=s^1)}{P(S=s^1)}$$

$$P(S = s^1) = \sum_{d,i,g,l} P(d, i, g, s, l)$$

an **inference** problem



d^0	d^1
0.6	0.4

i^0	i^1
0.7	0.3

	g^1	g^2	g^3
i^0,d^0	0.3	0.4	0.3
i^0,d^1	0.05	0.25	0.7
i^0,d^0	0.9	0.08	0.02
i^0,d^1	0.5	0.3	0.2

	s^0	s^1
i^0	0.95	0.05
i^1	0.2	0.8

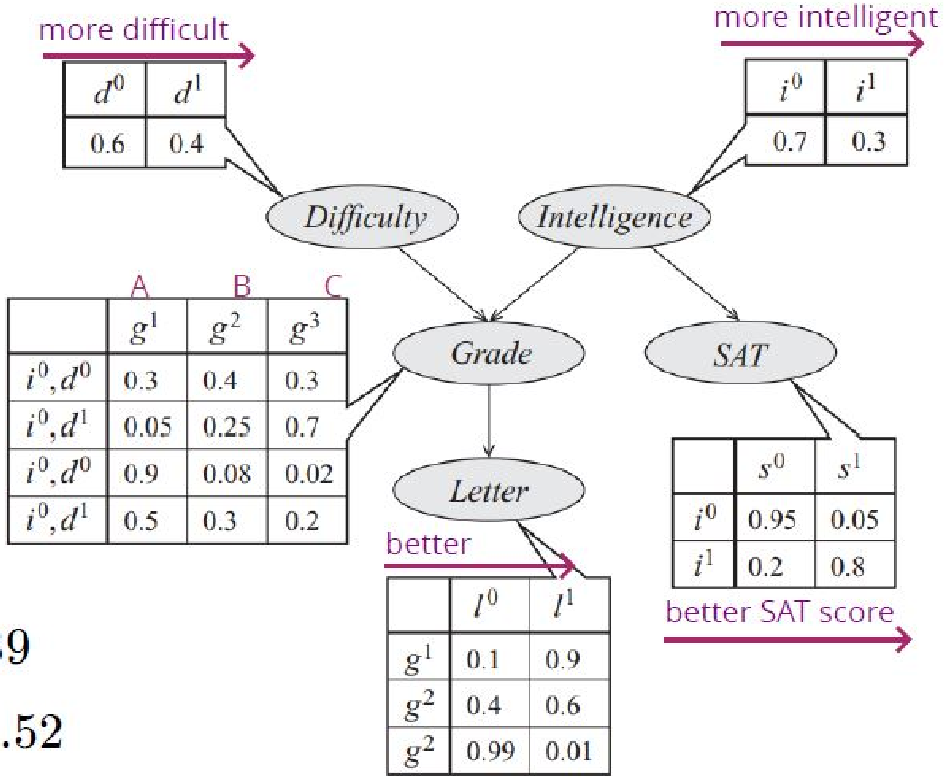
	l^0	l^1
g^1	0.1	0.9
g^2	0.4	0.6
g^2	0.99	0.01

causal reasoning (top-down)

- marginal prior
  - of getting a good letter

$$P(l^1) \approx .50$$

- marginal posterior
  - given low intelligence $P(l^1 \mid i^0) \approx .389$
  - ... and an easy exam $P(l^1 \mid i^0, d^0) \approx .52$



d^0	d^1
0.6	0.4

i^0	i^1
0.7	0.3

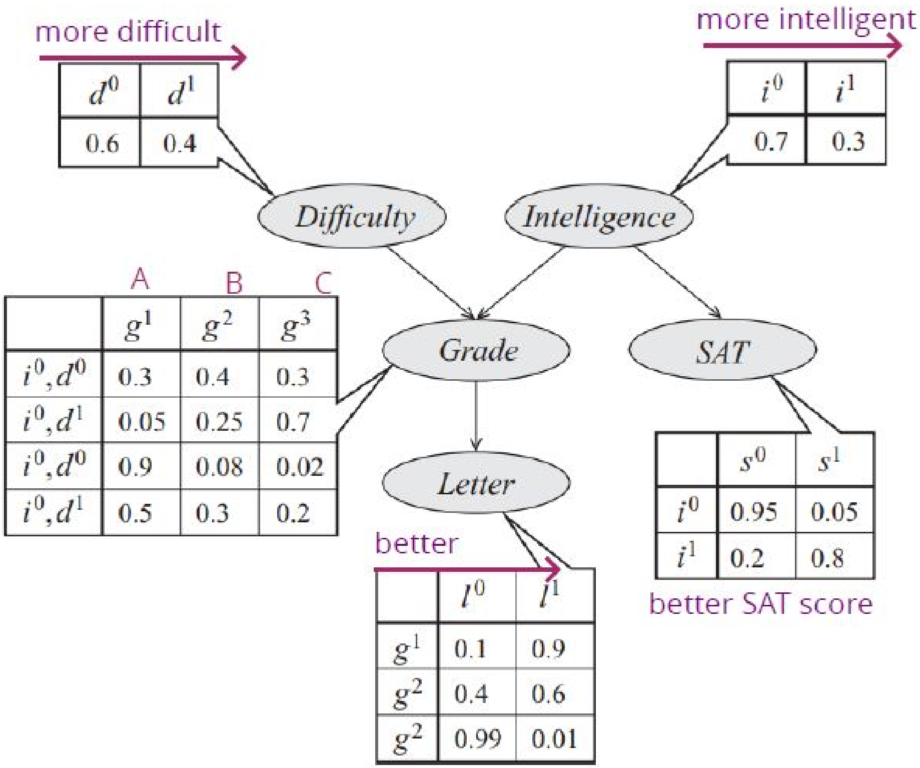
	g^1	g^2	g^3
i^0,d^0	0.3	0.4	0.3
i^0,d^1	0.05	0.25	0.7
i^0,d^0	0.9	0.08	0.02
i^0,d^1	0.5	0.3	0.2

	s^0	s^1
i^0	0.95	0.05
i^1	0.2	0.8

	l^0	l^1
g^1	0.1	0.9
g^2	0.4	0.6
g^2	0.99	0.01

evidential reasoning (bottom-up)

- (marginal) prior
  - of a high intelligence $P(i^1) \approx .30$
- (marginal) posterior
  - given a bad letter $P(i^1 \mid l^0) \approx .14$
  - ... and a bad grade $P(i^1 \mid l^0, g^3) \approx .08$



d^0	d^1
0.6	0.4

i^0	i^1
0.7	0.3

	g^1	g^2	g^3
i^0, d^0	0.3	0.4	0.3
i^0, d^1	0.05	0.25	0.7
i^0, d^0	0.9	0.08	0.02
i^0, d^1	0.5	0.3	0.2

	s^0	s^1
i^0	0.95	0.05
i^1	0.2	0.8

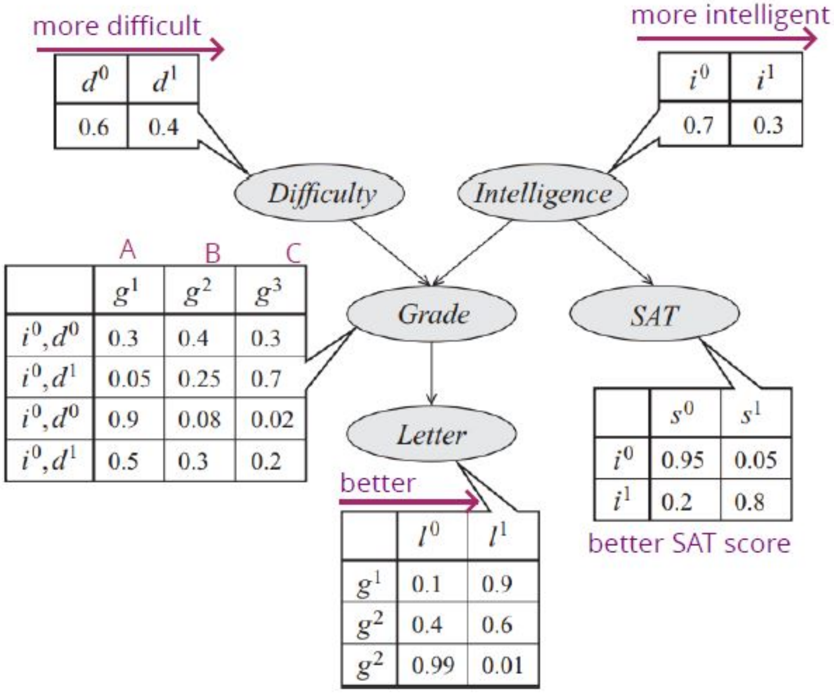
	l^0	l^1
g^1	0.1	0.9
g^2	0.4	0.6
g^2	0.99	0.01

Reasoning in a singly connected

Explaining

- prior
  - of a high intelligence $P(i^1) \approx .30$
- posterior
  - given a bad letter $P(i^1 \mid l^0) \approx .14$
  - ... and a bad grade $P(i^1 \mid l^0, g^3) \approx .08$
  - a difficult exam explains away the grade

$$P(i^1 \mid l^0, g^3, d^1) \approx .11$$



Problem: How to do the propagation in a general Bayesian Network, i.e. in a directed acyclic graph?