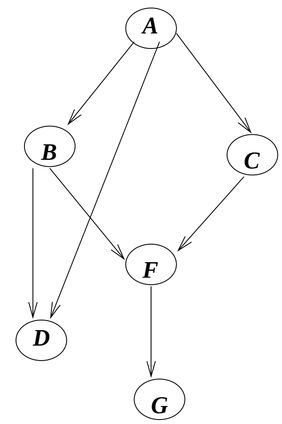
Bucket Elimination Example



$$P(\,A = \alpha\,) \quad =$$

$$= \sum_{c,b,f,d,g} P(\,A = \alpha,\;\; c, b, f, d, g\,)$$

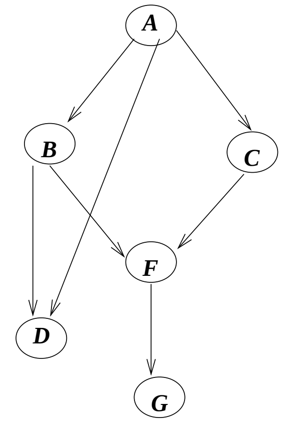
$$= \sum_{c,b,f,d,g} P(\,\alpha\,)\, P(\,c\,|\,\alpha\,)\, P(\,b\,|\,\alpha\,)\, P(\,f\,|\,b, c\,)\, P(\,d\,|\,\alpha, b\,)\, P(\,g\,|\,f\,)$$

$$= P(\,\alpha\,) \sum_c P(\,c\,|\,\alpha\,) \sum_b P(\,b\,|\,\alpha\,) \sum_f P(\,f\,|\,b, c\,) \times \sum_d P(\,d\,|\,\alpha, b\,) \sum_g P(\,g\,|\,f\,)$$

$$= \lambda_A(\,\alpha\,) \sum_c \lambda_{C|A}(\,c, \alpha\,) \sum_b \lambda_{B|A}(\,b, \alpha\,) \sum_f \lambda_{F|B,C}(\,f, b, c\,) \times \sum_d \lambda_{D|A,B}(\,d, \alpha, b\,) \sum_g \lambda_{G|F}(\,g, f\,)$$

Note: Each $P(\,X\,|\,Y\,)$ is from CP-table.

Using Buckets



$$P(\, A = \alpha \,) \quad =$$

$$\lambda_A(\,\alpha\,) \sum_c \lambda_{C|A}(\,c, \alpha\,) \sum_b \lambda_{B|A}(\,b, \alpha\,) \sum_f \lambda_{F|B,C}(\,f, b, c\,)$$
$$\times \sum_d \lambda_{D|A,B}(\,d, \alpha, b\,) \sum_g \lambda_{G|F}(\,g, f\,)$$

Buckets:

$\beta[A]$	$\beta[C]$	$\beta[B]$	$\beta[F]$	$\beta[D]$	$\beta[G]$
$\lambda_A(\,\alpha\,)$	$\lambda_{C|A}(\,c, \alpha\,)$	$\lambda_{B|A}(\,b, \alpha\,)$	$\lambda_{F|B,C}(\,f, b, c\,)$	$\lambda_{D|A,B}(\,d, \alpha, b\,)$	$\lambda_{G|F}(\,g, f\,)$

Plan: Sum out values of each bucket, 1-by-1,
creating new functions
Remove function from that bucket
Add new function to "earlier" bucket
. . .

At each time:

$$P(\, A = \alpha \,) \; = \; \sum_{c,b,f,d,g} \beta[A] \bowtie \beta[C] \bowtie \beta[B] \bowtie \beta[F] \bowtie \beta[D] \bowtie \beta[G]$$

Bucket-Elimination Algorithm

Order the variables $\langle X_1, \ldots, X_n \rangle$
Place each CPtable-function $\lambda_{X_i|X_{j1},\ldots,X_{jn}}(\ldots)$
into bucket $X_k = \text{argmax}_\ell\{X_i, X_{j1}, \ldots, X_{jn}\}$
Go through buckets in reverse order
$\langle \beta[G], \beta[D], \beta[F], \beta[B], \beta[C], \beta[A] \rangle$
For bucket X, with $\beta[X] = \{\lambda_{X,1}, \ldots, \lambda_{X,k}\}$:
Let $\lambda(\vec{Y}) = \sum_x \lambda_{X,1} \bowtie \cdots \bowtie \lambda_{X,k}$
Drop $\lambda(\vec{Y})$ into $\beta[Y_1]$
(Y_1 has highest index in ordering)

Buckets:

$\beta[A]$	$\beta[C]$	$\beta[B]$	$\beta[F]$	$\beta[D]$	$\beta[G]$
$\lambda_A(\alpha)$	$\lambda_{C\|A}(c, \alpha)$	$\lambda_{B\|A}(b, \alpha)$	$\lambda_{F\|B,C}(f, b, c)$	$\lambda_{D\|A,B}(d, \alpha, b)$	$\lambda_{G\|F}(g, f)$

% Given

g	f	$\lambda_{G\|F}(g, f)$
+	+	0.1
+	−	0.9
−	+	0.7
−	−	0.3

define

f	$\lambda_{\tilde{G}\|F}(f)$
+	0.1 + 0.7 = 0.8
−	0.8 + 0.3 = 1.1

% As $\lambda_{\tilde{G}|F}(f)$ depends on F, and not D . . .

$\lambda_A(\alpha)$	$\lambda_{C\|A}(c, \alpha)$	$\lambda_{B\|A}(b, \alpha)$	$\lambda_{F\|B,C}(f, b, c)$ $\lambda_{\tilde{G}\|F}(f)$	$\lambda_{D\|A,B}(d, \alpha, b)$	−

Operation#1: Summing Out

Buckets:

$\beta[A]$	$\beta[C]$	$\beta[B]$	$\beta[F]$	$\beta[D]$	$\beta[G]$
$\lambda_A(\alpha)$	$\lambda_{C\|A}(c,\alpha)$	$\lambda_{B\|A}(b,\alpha)$	$\lambda_{F\|B,C}(f,b,c)$	$\lambda_{D\|A,B}(d,\alpha,b)$	$\lambda_{G\|F}(g,f)$

% Given $\lambda_{G|F}(g,f) = \left\{ \begin{array}{cc|c} g & f & \lambda_{G|F}(g,f) \\ \hline + & + & 0.1 \\ + & - & 0.7 \\ - & + & 0.9 \\ - & - & 0.3 \end{array} \right\}$

% define $\lambda_{\tilde{G}|F}(f) = \left\{ \begin{array}{c|c} f & \lambda_{\tilde{G}|F}(f) \\ \hline + & 0.1 + 0.9 \;=\; 1.0 \\ - & 0.7 + 0.3 \;=\; 1.0 \end{array} \right\}$

% As $\lambda_{\tilde{G}|F}(f)$ depends on F, and not D . . .

$\lambda_A(\alpha)$	$\lambda_{C\|A}(c,\alpha)$	$\lambda_{B\|A}(b,\alpha)$	$\lambda_{F\|B,C}(f,b,c)$ $\boxed{\lambda_{\tilde{G}\|F}(f)}$	$\lambda_{D\|A,B}(d,\alpha,b)$	–

Trace, I

$\beta[A]$	$\beta[C]$	$\beta[B]$	$\beta[F]$	$\beta[D]$	$\beta[G]$
$\lambda_A(\alpha)$	$\lambda_{C\|A}(c,\alpha)$	$\lambda_{B\|A}(b,\alpha)$	$\lambda_{F\|B,C}(f,b,c)$	$\lambda_{D\|A,B}(d,\alpha,b)$	$\lambda_{G\|F}(g,f)$

% Let $\lambda_{\tilde{G}|F}(f) = \sum_g \lambda_{G|F}(g,f)$

$\lambda_A(\alpha)$	$\lambda_{C\|A}(c,\alpha)$	$\lambda_{B\|A}(b,\alpha)$	$\lambda_{F\|B,C}(f,b,c)$ $\lambda_{\tilde{G}\|F}(f)$	$\lambda_{D\|A,B}(d,\alpha,b)$	–

% Let $\lambda_{\tilde{D}|A,B}(\alpha,b) = \sum_d \lambda_{D|A,B}(d,\alpha,b)$

$\lambda_A(\alpha)$	$\lambda_{C\|A}(c,\alpha)$	$\lambda_{B\|A}(b,\alpha)$ $\lambda_{\tilde{D}\|A,B}(\alpha,b)$	$\lambda_{F\|B,C}(f,b,c)$ $\lambda_{\tilde{G}\|F}(f)$	–	–

Trace, II

$\beta[A]$	$\beta[C]$	$\beta[B]$	$\beta[F]$	$\beta[D$
$\lambda_A(\alpha)$	$\lambda_{C\|A}(c,\alpha)$	$\lambda_{B\|A}(b,\alpha)$ $\lambda_{\tilde{D}\|A,B}(\alpha,b)$	$\lambda_{F\|B,C}(f,b,c)$ $\lambda_{\tilde{G}\|F}(f)$	$-$

% Let $\lambda_{\tilde{F}:B,C}(b,c) = \sum_f \lambda_{F|B,C}(f,b,c) \times \lambda_{\tilde{G}|F}(f)$

$\lambda_A(\alpha)$	$\lambda_{C\|A}(c,\alpha)$	$\lambda_{B\|A}(b,\alpha)$ $\lambda_{\tilde{D}\|A,B}(\alpha,b)$ $\boxed{\lambda_{\tilde{F}:B,C}(b,c)}$	$-$	$-$

% Let $\lambda_{\tilde{B}:A,C}(\alpha,c) = \sum_b \lambda_{B|A}(b,\alpha) \times \lambda_{\tilde{D}|A,B}(\alpha,b) \times \lambda_{\tilde{F}:B,C}(b,c)$

$\lambda_A(\alpha)$	$\lambda_{C\|A}(c,\alpha)$ $\boxed{\lambda_{\tilde{B}:A,C}(\alpha,c)}$	$-$	$-$	$-$

% Let $\lambda_{\tilde{C}:A}(\alpha) = \sum_c \lambda_{C|A}(c,\alpha) \times \lambda_{\tilde{B}|A,C}(\alpha,c)$

$\lambda_A(\alpha)$ $\boxed{\lambda_{\tilde{C}:A}(\alpha)}$	$-$	$-$	$-$	$-$

Operation #2: Multiplying

Function: table wrt some NAMED variables

Given: $f(A=a, B=b)$, $g(A=a, C=c)$,
"Pointwise Product" $h = f \bowtie g$:

$$h(A=a, B=b, C=c) \quad = \quad f(A=a, B=b) \times g(A=a, C=c)$$

a	b	$f(A=a, B=b)$
t	t	0.30
t	f	0.70
f	t	0.91
f	f	0.09

a	c	$g(A=a, C=c)$
t	t	0.22
t	f	0.78
f	t	0.99
f	f	0.01

a	b	c	$h(A=a, B=b, C=c)$ $f(A=a, B=b) \times g(A=a, C=c)$
t	t	t	0.30 × 0.22
t	t	f	0.30 × 0.78
t	f	t	0.70 × 0.22
t	f	f	0.70 × 0.78
f	t	t	0.91 × 0.99
f	t	f	0.91 × 0.01
f	f	t	0.09 × 0.99
f	f	f	0.09 × 0.01

Bucket Elimination: Notes

- What to do with final buckets?

$\beta[A]$	$\beta[C]$	$\beta[B]$	$\beta[F]$	$\beta[D]$	$\beta[G]$
$\lambda_A(\alpha)$ $\lambda_{\tilde{C}:A}(\alpha)$	–	–	–	–	–

So $P(A=0) = \lambda_A(0) \times \lambda_{\tilde{C}:A}(0)$

$P(A=1) = \lambda_A(1) \times \lambda_{\tilde{C}:A}(1)$

- Can deal with > 1 variable:

To compute: $P(A=\alpha, C=\gamma)\ldots$

$\beta[A]$	$\beta[C]$	$\beta[B]$	$\beta[F]$	$\beta[D]$	$\beta[G]$
$\lambda_A(\alpha)$	$\lambda_{C\|A}(c,\alpha)$ $\lambda_{\tilde{B}:A,C}(\alpha,c)$	–	–	–	–

$$P\begin{pmatrix} A=\alpha \\ C=\gamma \end{pmatrix} = \lambda_A(\alpha)\cdot\lambda_{C|A}(c,\alpha)\cdot\lambda_{\tilde{B}:A,C}(\alpha,c)$$

- If dealing with CONDITIONAL ...

BucketElim Notes, II

- To compute

$$P(\,A = \alpha \,|\, C = 1\,) \;=\; \frac{P(\,A{=}\alpha,\ C{=}1\,)}{P(\,C{=}1\,)}$$

$$= \quad \frac{P(\,A = \alpha,\ C = 1\,)}{P(\,A = 0,\ C = 1\,) \;+\; P(\,A = 1,\ C = 1\,)}$$

$\Rightarrow$ Just need to compute $\boxed{P(\,\underline{A = \alpha},\ C = 1\,)}$ for all χ

- Here. . . can know $C = 1$.
  Can incorporate this. . .

Operation #3: Setting Value

View each CPtable as a function:

$$\lambda_{G|F}(g, f) = \left\{ \begin{array}{cc|c} f & g & \lambda_{G|F}(g, f) \\ \hline + & + & 0.1 \\ + & - & 0.9 \\ - & + & 0.7 \\ - & - & 0.3 \end{array} \right\}$$

as $\lambda_{G|F}: \{\mathtt{t}, \mathtt{f}\} \times \{\mathtt{t}, \mathtt{f}\} \mapsto [0, 1]$

SubDomain: When $G = \mathtt{t}$, get

$$\lambda_{\mathtt{t}G|F}(f) = \lambda_{G|F}(\mathtt{t}G, f)$$

$$\lambda_{\mathtt{t}G|F}(f) = \left\{ \begin{array}{c|c} f & \lambda_{\mathtt{t}G|F}(f) \\ \hline + & 0.1 \\ - & 0.7 \end{array} \right\}$$

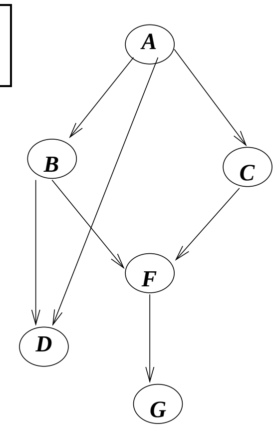
Similarly...

$$\lambda_{G|\mathtt{f}F}(g): \{\mathtt{t}, \mathtt{f}\} \mapsto [0, 1]$$

by

$$\lambda_{G|\mathtt{f}F}(\mathtt{t}G) = 0.70$$
$$\lambda_{G|\mathtt{f}F}(\mathtt{f}G) = 0.30$$

BucketElim: Efficiency Tricks



- If know values for some variables:

  For $P(\,\underline{A=\alpha},\ C=1, G=0\,)$

$\beta[A]$	$\beta[C]$	$\beta[B]$	$\beta[F]$	$\beta[D]$	$\beta[G]$
$\lambda_A(\,\alpha\,)$ $\lambda_{+C\mid A}(\,\alpha\,)$	$-$	$\lambda_{B\mid A}(\,b,\alpha\,)$	$\lambda_{F\mid B,C}(\,f,b,c\,)$ $\lambda_{-G\mid F}(\,f\,)$	$\lambda_{D\mid A,B}(\,d,\alpha,b\,)$	$-$

- Ignore bucket if d-separated (below evidence)

$$\begin{aligned}\lambda_{\tilde{D}\mid A,B}(\,\alpha,b\,) &= \textstyle\sum_d \lambda_{D\mid A,B}(\,d,\alpha,b\,)\\ &= \textstyle\sum_d P(\,D=d \mid A=\alpha, B=b\,) \ \equiv\ 1(\alpha,b)\end{aligned}$$

- Time/Space Efficiency:

  depends on size of largest bucket

  Eg: if every bucket ever has ≤ 3 variables,

  $O(D^3)$

  where $D = |Dom(X_i)|$ is size of domain

  $\Rightarrow$ Want to minimize

  "high-water-mark" size of largest bucket