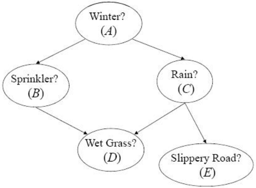
Logic and Bayesian Networks

Part 4: Variable Elimination

Jinbo Huang

Elimination



A	Θ_A
true	.6
false	.4

A	B	$\Theta_{B\|A}$
true	true	.2
true	false	.8
false	true	.75
false	false	.25

A	C	$\Theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.1
false	false	.9

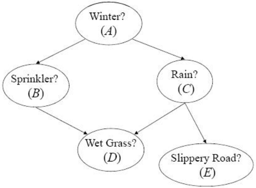
B	C	D	$\Theta_{D\|BC}$
true	true	true	.95
true	true	false	.05
true	false	true	.9
true	false	false	.1
false	true	true	.8
false	true	false	.2
false	false	true	0
false	false	false	1

C	E	$\Theta_{E\|C}$
true	true	.7
true	false	.3
false	true	0
false	false	1

$\Pr(D, E)$?

D	E	$\Pr(D, E)$
true	true	.30443
true	false	.39507
false	true	.05957
false	false	.24093

Elimination



A	Θ_A
true	.6
false	.4

A	B	$\Theta_{B\|A}$
true	true	.2
true	false	.8
false	true	.75
false	false	.25

A	C	$\Theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.1
false	false	.9

B	C	D	$\Theta_{D\|BC}$
true	true	true	.95
true	true	false	.05
true	false	true	.9
true	false	false	.1
false	true	true	.8
false	true	false	.2
false	false	true	0
false	false	false	1

C	E	$\Theta_{E\|C}$
true	true	.7
true	false	.3
false	true	0
false	false	1

$\Pr(D, E)$?

D	E	$\Pr(D, E)$
true	true	.30443
true	false	.39507
false	true	.05957
false	false	.24093

Sum out variables A, B, C from network

Elimination

A	B	C	D	E	Pr(.)
true	true	true	true	true	0.06384
true	true	true	true	false	0.02736
true	true	true	false	true	0.00336
true	true	true	false	false	0.00144
true	true	false	true	true	0.0
true	true	false	true	false	0.02160
true	true	false	false	true	0.0
true	true	false	false	false	0.00240
true	false	true	true	true	0.21504
true	false	true	true	false	0.09216
true	false	true	false	true	0.05376
true	false	true	false	false	0.02304
true	false	false	true	true	0.0
true	false	false	true	false	0.0
true	false	false	false	true	0.0
true	false	false	false	false	0.09600
false	true	true	true	true	0.01995
false	true	true	true	false	0.00855
false	true	true	false	true	0.00105
false	true	true	false	false	0.00045
false	true	false	true	true	0.0
false	true	false	true	false	0.24300
false	true	false	false	true	0.0
false	true	false	false	false	0.02700
false	false	true	true	true	0.00560
false	false	true	true	false	0.00240
false	false	true	false	true	0.00140
false	false	true	false	false	0.00060
false	false	false	true	true	0.0
false	false	false	true	false	0.0
false	false	false	false	true	0.0
false	false	false	false	false	0.0900

Summing out variables A

A	B	C	D	E	Pr(.)
true	true	true	true	true	0.06384
false	true	true	true	true	0.01995

B	C	D	E	Pr(.)
true	true	true	true	0.08379=0.06384+0.01995

Do it for all instantiations of B, C, D, E

Repeat to eliminate B, C

Elimination

A	B	C	D	E	Pr(.)
true	true	true	true	true	0.06384
true	true	true	true	false	0.02736
true	true	true	false	true	0.00336
true	true	true	false	false	0.00144
true	true	false	true	true	0.0
true	true	false	true	false	0.02160
true	true	false	false	true	0.0
true	true	false	false	false	0.00240
true	false	true	true	true	0.21504
true	false	true	true	false	0.09216
true	false	true	false	true	0.05376
true	false	true	false	false	0.02304
true	false	false	true	true	0.0
true	false	false	true	false	0.0
true	false	false	false	true	0.0
true	false	false	false	false	0.09600
false	true	true	true	true	0.01995
false	true	true	true	false	0.00855
false	true	true	false	true	0.00105
false	true	true	false	false	0.00045
false	true	false	true	true	0.0
false	true	false	true	false	0.24300
false	true	false	false	true	0.0
false	true	false	false	false	0.02700
false	false	true	true	true	0.00560
false	false	true	true	false	0.00240
false	false	true	false	true	0.00140
false	false	true	false	false	0.00060
false	false	false	true	true	0.0
false	false	false	true	false	0.0
false	false	false	false	true	0.0
false	false	false	false	false	0.0900

Summing out variables A

A	B	C	D	E	Pr(.)
true	true	true	true	true	0.06384
false	true	true	true	true	0.01995

B	C	D	E	Pr(.)
true	true	true	true	0.08379=0.06384+0.01995

Do it for all instantiations of B, C, D, E

Repeat to eliminate B, C

Exponential in number of variables

Elimination

A	B	C	D	E	Pr(.)
true	true	true	true	true	0.06384
true	true	true	true	false	0.02736
true	true	true	false	true	0.00336
true	true	true	false	false	0.00144
true	true	false	true	true	0.0
true	true	false	true	false	0.02160
true	true	false	false	true	0.0
true	true	false	false	false	0.00240
true	false	true	true	true	0.21504
true	false	true	true	false	0.09216
true	false	true	false	true	0.05376
true	false	true	false	false	0.02304
true	false	false	true	true	0.0
true	false	false	true	false	0.0
true	false	false	false	true	0.0
true	false	false	false	false	0.09600
false	true	true	true	true	0.01995
false	true	true	true	false	0.00855
false	true	true	false	true	0.00105
false	true	true	false	false	0.00045
false	true	false	true	true	0.0
false	true	false	true	false	0.24300
false	true	false	false	true	0.0
false	true	false	false	false	0.02700
false	false	true	true	true	0.00560
false	false	true	true	false	0.00240
false	false	true	false	true	0.00140
false	false	true	false	false	0.00060
false	false	false	true	true	0.0
false	false	false	true	false	0.0
false	false	false	false	true	0.0
false	false	false	false	false	0.0900

Summing out variables A

A	B	C	D	E	Pr(.)
true	true	true	true	true	0.06384
false	true	true	true	true	0.01995

B	C	D	E	Pr(.)
true	true	true	true	0.08379=0.06384+0.01995

Do it for all instantiations of B, C, D, E

Repeat to eliminate B, C

Exponential in number of variables

Solution: Elimination in *factored* form

Factors

B	C	D	f_1
true	true	true	.95
true	true	false	.05
true	false	true	.9
true	false	false	.1
false	true	true	.8
false	true	false	.2
false	false	true	0
false	false	false	1

D	E	f_2
true	true	0.448
true	false	0.192
false	true	0.112
false	false	0.248

Two factors: $f_1(b, c, d) = \Pr(d|b, c)$ and $f_2(d, e) = \Pr(d, e)$

- $f(x_1, \ldots, x_n)$: function from instantiation to number
- Can be joint or conditional probability
- *Trivial* factor: $n = 0$

Factors: Summing Out

- *Summing out* $Z \in \mathbf{X}$ from $f(\mathbf{X})$, where $\mathbf{Y} = \mathbf{X} \setminus \{Z\}$

$$(\sum_Z f)(\mathbf{y}) \stackrel{def}{=} \sum_z f(z, \mathbf{y})$$

- Commutative

$$\sum_Z \sum_W f = \sum_W \sum_Z f$$

- Summing out multiple variables $\sum_{\mathbf{Z}} f$: *marginalizing* variables $\mathbf{Z}$, *projecting* f on variables $\mathbf{Y}$ (other variables)
- Complexity $O(exp(w))$, where $w = |\mathbf{X}|$

Factors: Multiplication

- *Multiplying* $f_1(\mathbf{X})$ and $f_2(\mathbf{Y})$

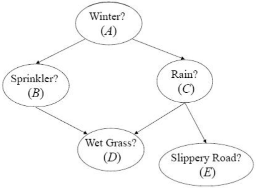
$$(f_1 f_2)(\mathbf{z}) \overset{def}{=} f_1(\mathbf{x}) f_2(\mathbf{y})$$

  where $\mathbf{Z} = \mathbf{X} \cup \mathbf{Y}$, $\mathbf{x} \sim \mathbf{z}$, $\mathbf{y} \sim \mathbf{z}$

- Commutative and associative

- Complexity $O(m \, exp(w))$ for m factors, where $w = |\mathbf{Z}|$

Prior Marginals by Elimination



A	Θ_A
true	.6
false	.4

A	B	$\Theta_{B\|A}$
true	true	.2
true	false	.8
false	true	.75
false	false	.25

A	C	$\Theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.1
false	false	.9

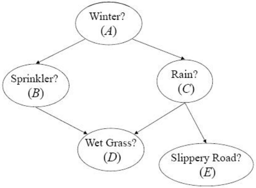
B	C	D	$\Theta_{D\|BC}$
true	true	true	.95
true	true	false	.05
true	false	true	.9
true	false	false	.1
false	true	true	.8
false	true	false	.2
false	false	true	0
false	false	false	1

C	E	$\Theta_{E\|C}$
true	true	.7
true	false	.3
false	true	0
false	false	1

Joint probability by chain rule

$$\Pr(a, b, c, d, e) = \theta_{e|c}\theta_{d|bc}\theta_{c|a}\theta_{b|a}\theta_a$$

Prior Marginals by Elimination



A	Θ_A
true	.6
false	.4

A	B	$\Theta_{B\|A}$
true	true	.2
true	false	.8
false	true	.75
false	false	.25

A	C	$\Theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.1
false	false	.9

B	C	D	$\Theta_{D\|BC}$
true	true	true	.95
true	true	false	.05
true	false	true	.9
true	false	false	.1
false	true	true	.8
false	true	false	.2
false	false	true	0
false	false	false	1

C	E	$\Theta_{E\|C}$
true	true	.7
true	false	.3
false	true	0
false	false	1

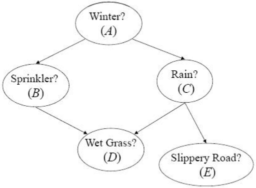
Joint probability by chain rule

$$\Pr(a, b, c, d, e) = \theta_{e|c}\theta_{d|bc}\theta_{c|a}\theta_{b|a}\theta_a$$

Joint probability as $\prod$ of factors

$$\Theta_{E|C}\Theta_{D|BC}\Theta_{D|A}\Theta_{B|A}\Theta_A$$

Prior Marginals by Elimination



A	Θ_A
true	.6
false	.4

A	B	$\Theta_{B\|A}$
true	true	.2
true	false	.8
false	true	.75
false	false	.25

A	C	$\Theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.1
false	false	.9

B	C	D	$\Theta_{D\|BC}$
true	true	true	.95
true	true	false	.05
true	false	true	.9
true	false	false	.1
false	true	true	.8
false	true	false	.2
false	false	true	0
false	false	false	1

C	E	$\Theta_{E\|C}$
true	true	.7
true	false	.3
false	true	0
false	false	1

Joint probability by chain rule

$$\Pr(a, b, c, d, e) = \theta_{e|c}\theta_{d|bc}\theta_{c|a}\theta_{b|a}\theta_a$$

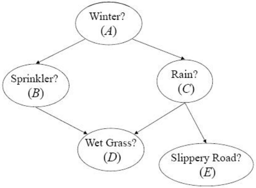
Joint probability as $\prod$ of factors

$$\Theta_{E|C}\Theta_{D|BC}\Theta_{D|A}\Theta_{B|A}\Theta_A$$

Marginals

$$\Pr(D, E) = \sum_{A,B,C} \Theta_{E|C}\Theta_{D|BC}\Theta_{D|A}\Theta_{B|A}\Theta_A$$

Prior Marginals by Elimination



A	Θ_A
true	.6
false	.4

A	B	$\Theta_{B\|A}$
true	true	.2
true	false	.8
false	true	.75
false	false	.25

A	C	$\Theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.1
false	false	.9

B	C	D	$\Theta_{D\|BC}$
true	true	true	.95
true	true	false	.05
true	false	true	.9
true	false	false	.1
false	true	true	.8
false	true	false	.2
false	false	true	0
false	false	false	1

C	E	$\Theta_{E\|C}$
true	true	.7
true	false	.3
false	true	0
false	false	1

Joint probability by chain rule

$$\Pr(a, b, c, d, e) = \theta_{e|c}\theta_{d|bc}\theta_{c|a}\theta_{b|a}\theta_a$$

Joint probability as $\prod$ of factors

$$\Theta_{E|C}\Theta_{D|BC}\Theta_{D|A}\Theta_{B|A}\Theta_A$$

Marginals

$$\Pr(D, E) = \sum_{A,B,C} \Theta_{E|C}\Theta_{D|BC}\Theta_{D|A}\Theta_{B|A}\Theta_A$$

Complexity still exponential in # of variables

Prior Marginals by Elimination: Early Summation

- Don't multiply all factors before summation
- Theorem: If X does not appear in f_1, then

$$\sum_X f_1 f_2 = f_1 \sum_X f_2$$

Prior Marginals by Elimination: Early Summation

- Don't multiply all factors before summation
- Theorem: If X does not appear in f_1, then

$$\sum_X f_1 f_2 = f_1 \sum_X f_2$$

- For example, if X appears only in f_n, then

$$\sum_X f_1 \dots f_n = f_1 \dots f_{n-1} \sum_X f_n$$

Prior Marginals by Elimination: Early Summation

- Don't multiply all factors before summation
- Theorem: If X does not appear in f_1, then

$$\sum_X f_1 f_2 = f_1 \sum_X f_2$$

- For example, if X appears only in f_n, then

$$\sum_X f_1 \ldots f_n = f_1 \ldots f_{n-1} \sum_X f_n$$

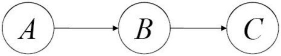
- Similarly, if X appears only in f_{n-1} and f_n, then

$$\sum_X f_1 \ldots f_n = f_1 \ldots f_{n-2} \sum_X f_{n-1} f_n$$

Prior Marginals by Elimination: Early Summation

- Multiply all factors that include X, sum out X from result
- Early summation reduces factor size, hence complexity of $\prod$

Prior Marginals by Elimination: Early Summation



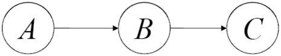
Compute $\Pr(C)$: eliminate A, then B

A	Θ_A
true	.6
false	.4

A	B	$\Theta_{B\|A}$
true	true	.9
true	false	.1
false	true	.2
false	false	.8

B	C	$\Theta_{C\|B}$
true	true	.3
true	false	.7
false	true	.5
false	false	.5

Prior Marginals by Elimination: Early Summation



A	Θ_A
true	.6
false	.4

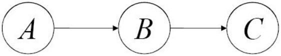
A	B	$\Theta_{B\|A}$
true	true	.9
true	false	.1
false	true	.2
false	false	.8

B	C	$\Theta_{C\|B}$
true	true	.3
true	false	.7
false	true	.5
false	false	.5

Compute $\Pr(C)$: eliminate A, then B

Two factors mention A: Θ_A, $\Theta_{B|A}$

Prior Marginals by Elimination: Early Summation



A	Θ_A
true	.6
false	.4

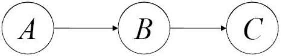
A	B	$\Theta_{B\|A}$
true	true	.9
true	false	.1
false	true	.2
false	false	.8

B	C	$\Theta_{C\|B}$
true	true	.3
true	false	.7
false	true	.5
false	false	.5

Multiply Θ_A and $\Theta_{B|A}$

A	B	$\Theta_A\Theta_{B\|A}$
true	true	.54
true	false	.06
false	true	.08
false	false	.32

Prior Marginals by Elimination: Early Summation



A	Θ_A
true	.6
false	.4

A	B	$\Theta_{B\|A}$
true	true	.9
true	false	.1
false	true	.2
false	false	.8

B	C	$\Theta_{C\|B}$
true	true	.3
true	false	.7
false	true	.5
false	false	.5

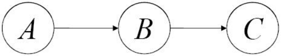
Multiply Θ_A and $\Theta_{B|A}$

A	B	$\Theta_A\Theta_{B\|A}$
true	true	.54
true	false	.06
false	true	.08
false	false	.32

Sum out A

B	$\sum_A \Theta_A\Theta_{B\|A}$
true	$.62 = .54 + .08$
false	$.38 = .06 + .32$

Prior Marginals by Elimination: Early Summation



A	Θ_A
true	.6
false	.4

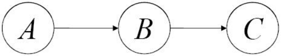
A	B	$\Theta_{B\|A}$
true	true	.9
true	false	.1
false	true	.2
false	false	.8

B	C	$\Theta_{C\|B}$
true	true	.3
true	false	.7
false	true	.5
false	false	.5

Two factors left, $\Theta_{C|B}$ & $\sum_A \Theta_A \Theta_{B|A}$, multiply

B	C	$\Theta_{C\|B} \sum_A \Theta_A \Theta_{B\|A}$
true	true	.186
true	false	.434
false	true	.190
false	false	.190

Prior Marginals by Elimination: Early Summation



A	Θ_A
true	.6
false	.4

A	B	$\Theta_{B\|A}$
true	true	.9
true	false	.1
false	true	.2
false	false	.8

B	C	$\Theta_{C\|B}$
true	true	.3
true	false	.7
false	true	.5
false	false	.5

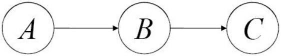
Two factors left, $\Theta_{C|B}$ & $\sum_A \Theta_A \Theta_{B|A}$, multiply

B	C	$\Theta_{C\|B} \sum_A \Theta_A \Theta_{B\|A}$
true	true	.186
true	false	.434
false	true	.190
false	false	.190

Sum out B

C	$\sum_B \Theta_{C\|B} \sum_A \Theta_A \Theta_{B\|A}$
true	.376
false	.624

Prior Marginals by Elimination: Early Summation



A	Θ_A
true	.6
false	.4

A	B	$\Theta_{B\mid A}$
true	true	.9
true	false	.1
false	true	.2
false	false	.8

B	C	$\Theta_{C\mid B}$
true	true	.3
true	false	.7
false	true	.5
false	false	.5

Biggest factor produced: 4 rows

Two factors left, $\Theta_{C|B}$ & $\sum_A \Theta_A \Theta_{B|A}$, multiply

B	C	$\Theta_{C\mid B} \sum_A \Theta_A \Theta_{B\mid A}$
true	true	.186
true	false	.434
false	true	.190
false	false	.190

Sum out B

C	$\sum_B \Theta_{C\mid B} \sum_A \Theta_A \Theta_{B\mid A}$
true	.376
false	.624

Prior Marginals by Elimination: Algorithm

Input: Bayesian network $\mathcal{N}$, variables $\mathbf{Q}$, order π on other variables

Output: prior marginal $\Pr(\mathbf{Q})$

$\mathcal{S} \leftarrow$ CPTs of network $\mathcal{N}$
for $i = 1$ to $|\pi|$ **do**
  $f \leftarrow \prod_k f_k$, where $f_k \in \mathcal{S}$ and mentions variable $\pi(i)$
  $f_i \leftarrow \sum_{\pi(i)} f$
  remove all f_k from $\mathcal{S}$, add f_i
return $\prod_{f \in \mathcal{S}} f$

Prior Marginals by Elimination: Algorithm

Input: Bayesian network $\mathcal{N}$, variables $\mathbf{Q}$, order π on other variables

Output: prior marginal $\Pr(\mathbf{Q})$

1: $\mathcal{S} \leftarrow$ CPTs of network $\mathcal{N}$
2: **for** $i = 1$ to $|\pi|$ **do**
3: $\quad f \leftarrow \prod_k f_k$, where $f_k \in \mathcal{S}$ and mentions variable $\pi(i)$
4: $\quad f_i \leftarrow \sum_{\pi(i)} f$
5: $\quad$ remove all f_k from $\mathcal{S}$, add f_i
6: **return** $\prod_{f \in \mathcal{S}} f$

Complexity (not counting line 6): $O(n\, exp(w))$, where w is $\#$ of variables of largest f_i, known as *width* of order π

Prior Marginals by Elimination: Algorithm

Input: Bayesian network $\mathcal{N}$, variables $\mathbf{Q}$, order π on other variables

Output: prior marginal $\Pr(\mathbf{Q})$

1: $\mathcal{S} \leftarrow$ CPTs of network $\mathcal{N}$
2: **for** $i = 1$ to $|\pi|$ **do**
3: $\quad f \leftarrow \prod_k f_k$, where $f_k \in \mathcal{S}$ and mentions variable $\pi(i)$
4: $\quad f_i \leftarrow \sum_{\pi(i)} f$
5: $\quad$ remove all f_k from $\mathcal{S}$, add f_i
6: **return** $\prod_{f \in \mathcal{S}} f$

How do we find all f_k on line 3 quickly (linear in $\#$ of such f_k)?

Prior Marginals by Elimination: Bucket Elimination

Bucket	Factors
E	$\Theta_{E\|C}$
B	$\Theta_{B\|A}$, $\Theta_{D\|BC}$
C	$\Theta_{C\|A}$
D	
A	Θ_A

Prior Marginals by Elimination: Bucket Elimination

Bucket	Factors
E	$\Theta_{E\|C}$
B	$\Theta_{B\|A}$, $\Theta_{D\|BC}$
C	$\Theta_{C\|A}$
D	
A	Θ_A

Bucket	Factors
E	
B	$\Theta_{B\|A}$, $\Theta_{D\|BC}$
C	$\Theta_{C\|A}$, $\sum_E \Theta_{E\|C}$
D	
A	Θ_A

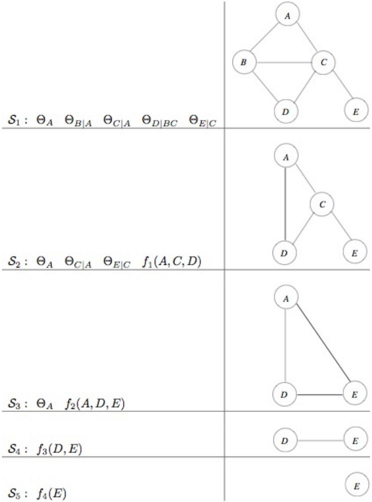
Width of Elimination Order

- Should prefer order with smaller width
- How to compute width, without actually running elimination?

Width of Elimination Order

- Should prefer order with smaller width
- How to compute width, without actually running elimination?
- Only care about size of factors, run abstract version of algorithm keeping track of factor sizes only

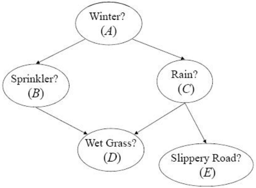
Width of Elimination Order



Computing Good Elimination Orders

- Finding optimal order is NP-hard
- Min-degree: eliminate variable with fewest neighbors
- Min-fill: eliminate variable leading to fewest fill-in edges

Posterior Marginals by Elimination



A	Θ_A
true	.6
false	.4

A	B	$\Theta_{B\|A}$
true	true	.2
true	false	.8
false	true	.75
false	false	.25

A	C	$\Theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.1
false	false	.9

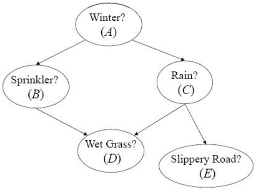
B	C	D	$\Theta_{D\|BC}$
true	true	true	.95
true	true	false	.05
true	false	true	.9
true	false	false	.1
false	true	true	.8
false	true	false	.2
false	false	true	0
false	false	false	1

C	E	$\Theta_{E\|C}$
true	true	.7
true	false	.3
false	true	0
false	false	1

$\Pr(D, E|\mathbf{e})$? $\mathbf{e} : A = \text{true}, B = \text{false}$

D	E	$\Pr(D, E\|\mathbf{e})$
true	true	.448
true	false	.192
false	true	.112
false	false	.248

Posterior Marginals by Elimination



A	Θ_A
true	.6
false	.4

A	B	$\Theta_{B\|A}$
true	true	.2
true	false	.8
false	true	.75
false	false	.25

A	C	$\Theta_{C\|A}$
true	true	.8
true	false	.2
false	true	.1
false	false	.9

B	C	D	$\Theta_{D\|BC}$
true	true	true	.95
true	true	false	.05
true	false	true	.9
true	false	false	.1
false	true	true	.8
false	true	false	.2
false	false	true	0
false	false	false	1

C	E	$\Theta_{E\|C}$
true	true	.7
true	false	.3
false	true	0
false	false	1

$\Pr(D, E|\mathbf{e})$? $\mathbf{e}: A = \text{true}, B = \text{false}$

D	E	$\Pr(D, E\|\mathbf{e})$
true	true	.448
true	false	.192
false	true	.112
false	false	.248

Compute *joint marginals* instead

D	E	$\Pr(D, E, \mathbf{e})$
true	true	.21504
true	false	.09216
false	true	.05376
false	false	.11904

Posterior Marginals by Elimination

- *Zero out* all rows of all factors inconsistent with $\mathbf{e}$
- Run elimination, result will be joint marginal $\Pr(\mathbf{Q}, \mathbf{e})$
- Add all entries to obtain $\Pr(\mathbf{e})$
- $\Pr(\mathbf{Q}|\mathbf{e}) = \frac{\Pr(\mathbf{Q},\mathbf{e})}{\Pr(\mathbf{e})}$

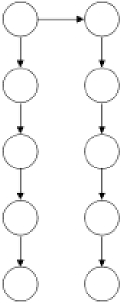
Posterior Marginals by Elimination

- *Zero out* all rows of all factors inconsistent with $\mathbf{e}$
- Run elimination, result will be joint marginal $\Pr(\mathbf{Q}, \mathbf{e})$
- Add all entries to obtain $\Pr(\mathbf{e})$
- $\Pr(\mathbf{Q}|\mathbf{e}) = \frac{\Pr(\mathbf{Q},\mathbf{e})}{\Pr(\mathbf{e})}$
- Run with $\mathbf{Q} = \emptyset$ for $\Pr(\mathbf{e})$

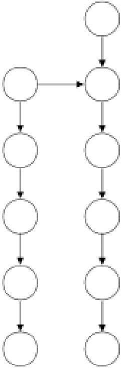
Network Structure and Complexity: Treewidth

- Complexity of elimination exp. in width of elimination order
- *Treewidth* is width of best elimination order for given network
- Quantifies how close the network resembles a tree

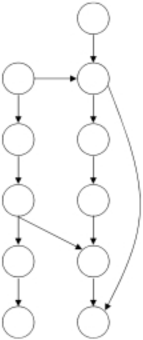
Network Structure and Complexity: Treewidth



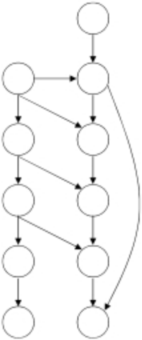
1



2



3



3

Network Structure and Complexity: Treewidth

- Trees have treewidth 1
- # of nodes has no genuine effect on treewidth
- # of parents per node has effect
  - Treewidth $\geq$ max # of parents per node
  - Equality holds for *polytrees*, or *singly-connected* networks
- Loops tend to increase treewidth
- # of loops has no genuine effect

Query Structure and Complexity: Network Pruning

- Consider computation of $\Pr(\mathbf{Q}, \mathbf{e})$ (includes prior marginals and probability of evidence as special cases)

Query Structure and Complexity: Network Pruning

- Consider computation of $\Pr(\mathbf{Q}, \mathbf{e})$ (includes prior marginals and probability of evidence as special cases)
- Pruning nodes: All leaves $\notin \mathbf{Q} \cup \mathbf{E}$, iteratively
  - Worst case: All leaves $\in \mathbf{Q} \cup \mathbf{E}$, no pruning
  - Best case: All $\mathbf{Q} \cup \mathbf{E}$ are roots, every node $\notin \mathbf{Q} \cup \mathbf{E}$ pruned

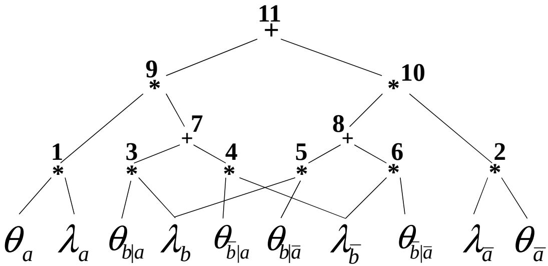
Query Structure and Complexity: Network Pruning

- Consider computation of $\Pr(\mathbf{Q}, \mathbf{e})$ (includes prior marginals and probability of evidence as special cases)
- Pruning nodes: All leaves $\notin \mathbf{Q} \cup \mathbf{E}$, iteratively
  - Worst case: All leaves $\in \mathbf{Q} \cup \mathbf{E}$, no pruning
  - Best case: All $\mathbf{Q} \cup \mathbf{E}$ are roots, every node $\notin \mathbf{Q} \cup \mathbf{E}$ pruned
- Pruning edges: For each edge $U \to X$, $U \in \mathbf{E}$
  - Remove edge, shrink CPT $\Theta_{X|\mathbf{U}}$ by removing rows inconsistent with $\mathbf{e}$ and removing column U

Query Structure and Complexity: Network Pruning

- Consider computation of $\Pr(\mathbf{Q}, \mathbf{e})$ (includes prior marginals and probability of evidence as special cases)
- Pruning nodes: All leaves $\notin \mathbf{Q} \cup \mathbf{E}$, iteratively
  - Worst case: All leaves $\in \mathbf{Q} \cup \mathbf{E}$, no pruning
  - Best case: All $\mathbf{Q} \cup \mathbf{E}$ are roots, every node $\notin \mathbf{Q} \cup \mathbf{E}$ pruned
- Pruning edges: For each edge $U \rightarrow X$, $U \in \mathbf{E}$
  - Remove edge, shrink CPT $\Theta_{X|\mathbf{U}}$ by removing rows inconsistent with $\mathbf{e}$ and removing column U
- *Effective treewidth* is treewidth of pruned network given query

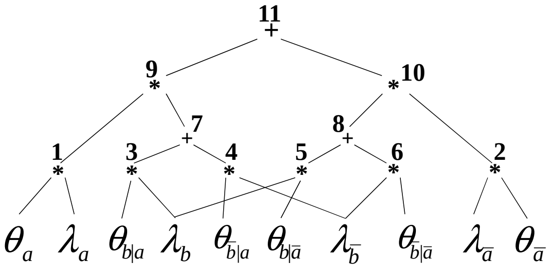
Arithmetic Circuits from Variable Elimination



A	B	$\Theta_{B\|A}$
true	true	$n_3 = \star(\lambda_b, \theta_{b\|a})$
true	false	$n_4 = \star(\lambda_{\bar{b}}, \theta_{\bar{b}\|a})$
false	true	$n_5 = \star(\lambda_b, \theta_{b\|\bar{a}})$
false	false	$n_6 = \star(\lambda_{\bar{b}}, \theta_{\bar{b}\|\bar{a}})$

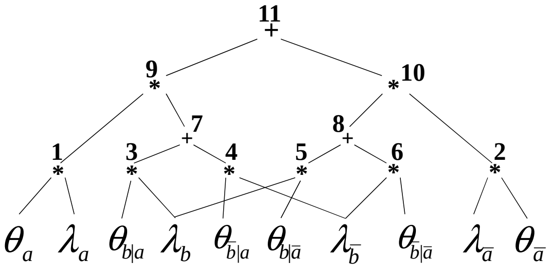
A	Θ_A
true	$n_1 = \star(\lambda_a, \theta_a)$
false	$n_2 = \star(\lambda_{\bar{a}}, \theta_{\bar{a}})$

Arithmetic Circuits from Variable Elimination



A	B	$\Theta_{B\|A}$
true	true	$n_3 = \star(\lambda_b, \theta_{b\|a})$
true	false	$n_4 = \star(\lambda_{\bar{b}}, \theta_{\bar{b}\|a})$
false	true	$n_5 = \star(\lambda_b, \theta_{b\|\bar{a}})$
false	false	$n_6 = \star(\lambda_{\bar{b}}, \theta_{\bar{b}\|\bar{a}})$

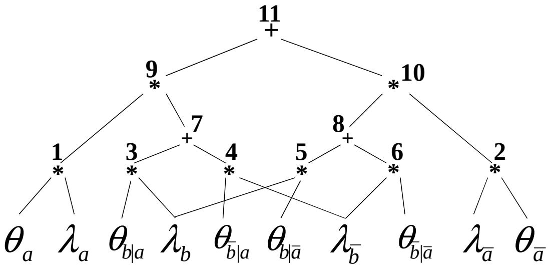
Arithmetic Circuits from Variable Elimination



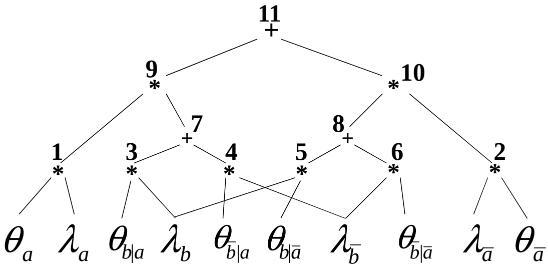
A	B	$\Theta_{B\|A}$
true	true	$n_3 = \star(\lambda_b, \theta_{b\|a})$
true	false	$n_4 = \star(\lambda_{\bar{b}}, \theta_{\bar{b}\|a})$
false	true	$n_5 = \star(\lambda_b, \theta_{b\|\bar{a}})$
false	false	$n_6 = \star(\lambda_{\bar{b}}, \theta_{\bar{b}\|\bar{a}})$

A	$\sum_B \Theta_{B\|A}$
true	$n_7 = +(n_3, n_4)$
false	$n_8 = +(n_5, n_6)$

Arithmetic Circuits from Variable Elimination



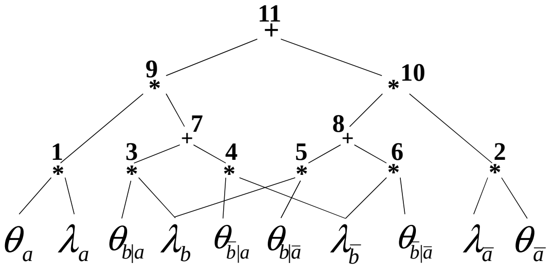
A	$\Theta_A \sum_B \Theta_{B\|A}$
true	$n_9 = \star(n_1, n_7)$
false	$n_{10} = \star(n_2, n_8)$



A	$\Theta_A \sum_B \Theta_{B\|A}$
true	$n_9 = \star(n_1, n_7)$
false	$n_{10} = \star(n_2, n_8)$

	$\sum_A \Theta_A \sum_B \Theta_{B\|A}$
	$n_{11} = +(n_9, n_{10})$

Arithmetic Circuits from Variable Elimination



Circuit size $O(n\, exp(w))$ as complexity of variable elimination

Variable Elimination vs. Compilation

Variable elimination

- $\Theta(n\, exp(w))$ in all cases
- A run of VE answers only one query
- Arithmetic circuit from VE useful for multiple queries, but still $\Theta(n\, exp(w))$

Variable Elimination vs. Compilation

Variable elimination

- $\Theta(n\,exp(w))$ in all cases
- A run of VE answers only one query
- Arithmetic circuit from VE useful for multiple queries, but still $\Theta(n\,exp(w))$

Compilation of logical encoding

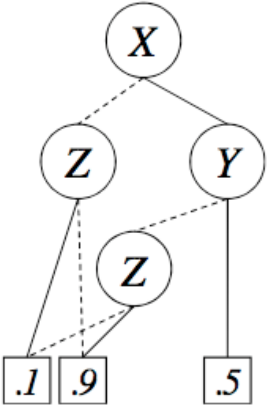
- Compilation $O(n\,exp(w))$ only in worst case, can be much faster
- Smaller arithmetic circuits, faster online query answering

Reviving Variable Elimination

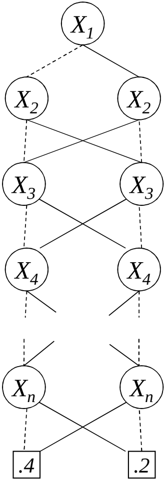
- Tables always have fixed size: exponential in # of variables
- Use non-tabular representations of factors to reduce size

Algebraic Decision Diagrams (ADDs)

X	Y	Z	$f(.)$
F	F	F	.9
F	F	T	.1
F	T	F	.9
F	T	T	.1
T	F	F	.1
T	F	T	.9
T	T	F	.5
T	T	T	.5

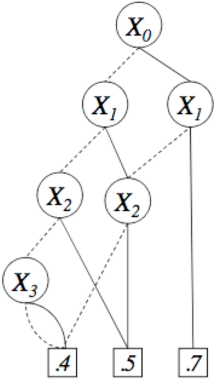


Compactness of ADDs

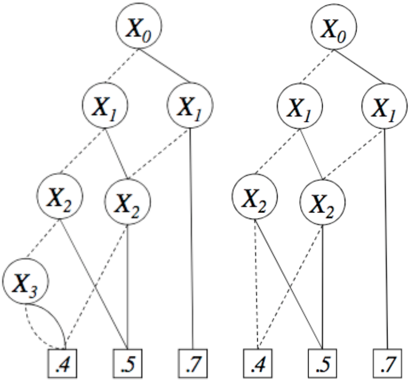


- $f(x_1, \ldots, x_n) = .2$ if odd $\#$ of x_i are true, .4 otherwise
- ADD size $O(n)$, tabular size $O(exp(n))$

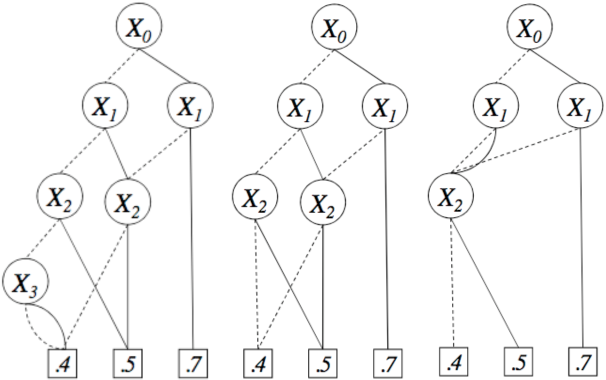
ADD Reduction



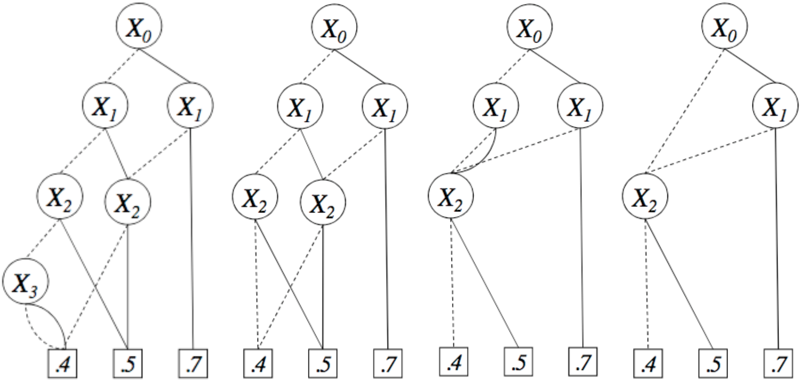
ADD Reduction



ADD Reduction



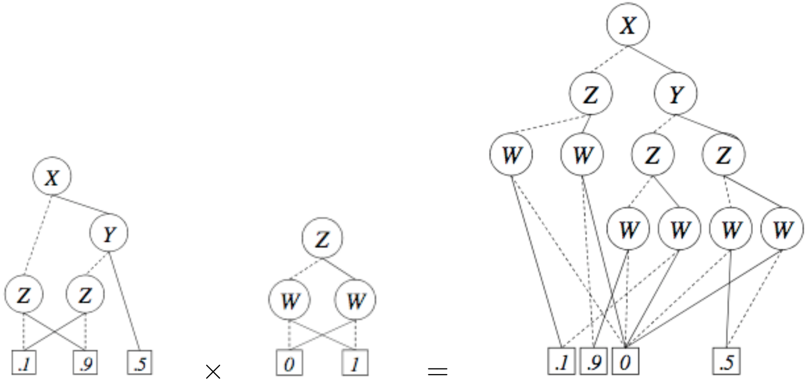
ADD Reduction



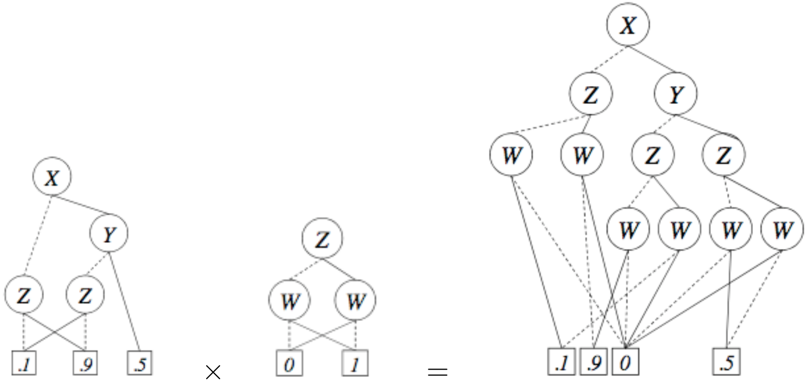
ADD Reduction

- Reduced ADDs are *canonical*: Unique for given variable order
- Size sensitive to variable order
- When used in elimination, reverse of elimination order tends to work well

ADD Operations: Apply

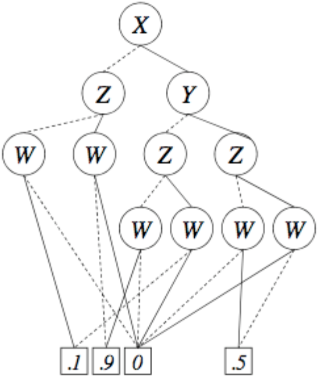


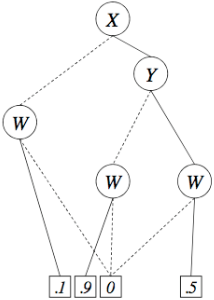
ADD Operations: Apply



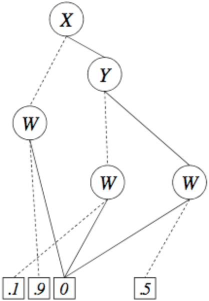
- Works with any binary operator: $+, -, \times, /$, etc
- Complexity $O(nm)$ (avoid redundant work with caching)

ADD Operations: Restrict



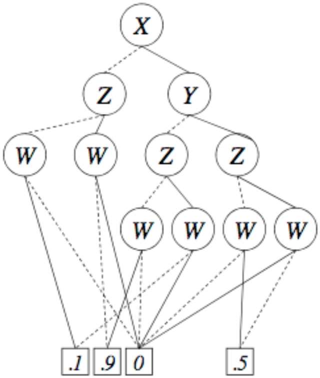


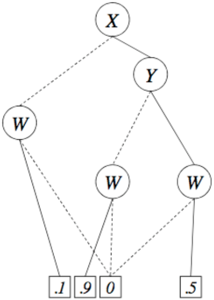
$\sum_Z f^{Z=F}$

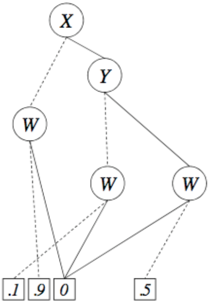


$\sum_Z f^{Z=T}$

ADD Operations: Restrict







$\sum_Z f^{Z=F}$ $\sum_Z f^{Z=T}$

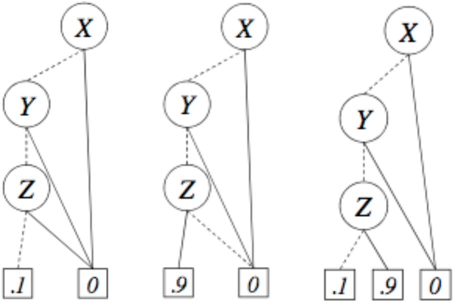
- Complexity $O(n)$, same for multiple variables

ADD Operations: Sum Out

$$\sum_X f = \sum_X f^{X=F} + \sum_X f^{X=T}$$

From Table to ADD

X	Y	Z	$f(.)$
F	F	F	.1
F	F	T	.9
F	T	F	.1
F	T	T	.9
T	F	F	.9
T	F	T	.1
T	T	F	.5
T	T	T	.5



- Each row to ADD, then add them up

Variable Elimination with ADDs

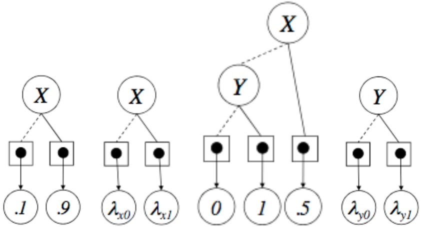
Use ADDs instead of tables

Multiplication and summing out by ADD operations

Arithmetic Circuits from Variable Elimination Revisited

X	Θ_X
x_0	.1
x_1	.9

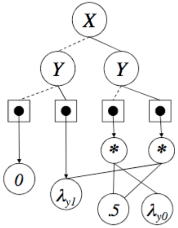
X	Y	$\Theta_{Y\|X}$
x_0	y_0	0
x_0	y_1	1
x_1	y_0	.5
x_1	y_1	.5

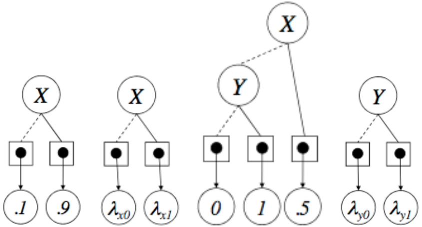


Arithmetic Circuits from Variable Elimination Revisited

X	Θ_X
x_0	.1
x_1	.9

X	Y	$\Theta_{Y\|X}$
x_0	y_0	0
x_0	y_1	1
x_1	y_0	.5
x_1	y_1	.5

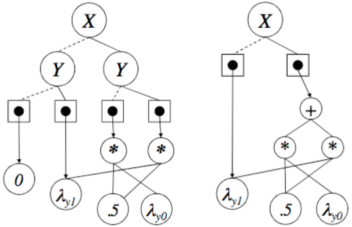


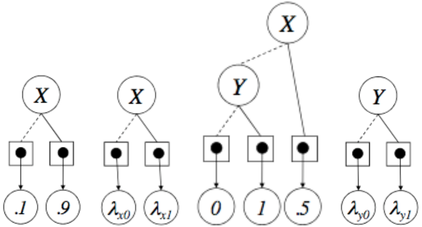


Arithmetic Circuits from Variable Elimination Revisited

X	Θ_X
x_0	.1
x_1	.9

X	Y	$\Theta_{Y\|X}$
x_0	y_0	0
x_0	y_1	1
x_1	y_0	.5
x_1	y_1	.5

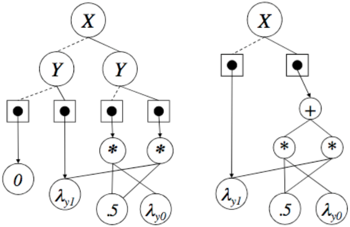


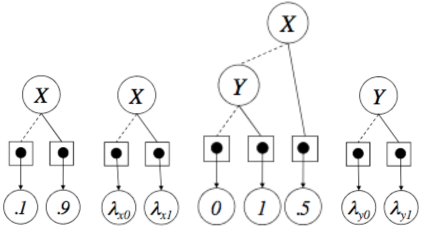


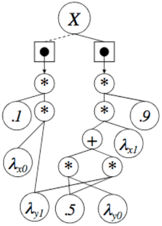
Arithmetic Circuits from Variable Elimination Revisited

X	Θ_X
x_0	.1
x_1	.9

X	Y	$\Theta_{Y\|X}$
x_0	y_0	0
x_0	y_1	1
x_1	y_0	.5
x_1	y_1	.5



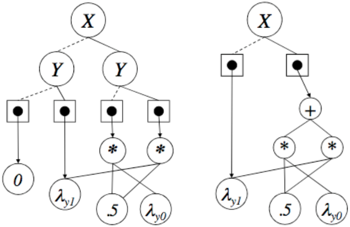


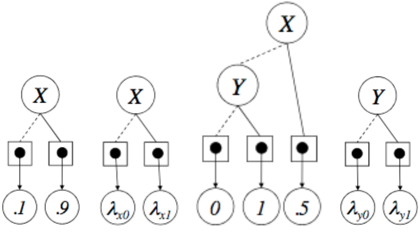


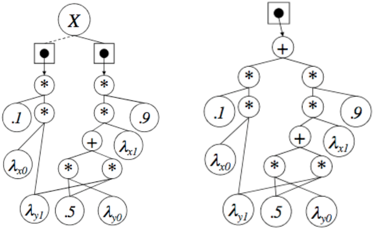
Arithmetic Circuits from Variable Elimination Revisited

X	Θ_X
x_0	.1
x_1	.9

X	Y	$\Theta_{Y\|X}$
x_0	y_0	0
x_0	y_1	1
x_1	y_0	.5
x_1	y_1	.5







Variable Elimination: Summary

- Factors, summing out, multiplication
- Prior marginals by elimination, bucket elimination
- Width of elimination order, min-degree, min-fill
- Posterior marginals and probability of evidence by zeroing out
- Treewidth and complexity
- Network pruning based on query
- Arithmetic circuits from variable elimination
- Variable elimination with ADDs