

Variable Elimination

- Variable elimination is the dynamic programming variant of recursive conditioning.

Variable Elimination

- Variable elimination is the dynamic programming variant of recursive conditioning.
- Give a factorization, such as

$$P(D) = \sum_C P(D | C) \sum_B P(C | B) \sum_A P(A)P(B | A)$$

it does the innermost sums first, constructing representations of the intermediate **factors**:

- ▶ $\sum_A P(A)P(B | A)$ is a factor on B ; call it $f_1(B)$.

Variable Elimination

- Variable elimination is the dynamic programming variant of recursive conditioning.
- Give a factorization, such as

$$P(D) = \sum_C P(D | C) \sum_B P(C | B) \sum_A P(A)P(B | A)$$

it does the innermost sums first, constructing representations of the intermediate **factors**:

- ▶ $\sum_A P(A)P(B | A)$ is a factor on B ; call it $f_1(B)$.
- ▶ $\sum_B P(C | B)f_1(B)$ is a factor on C .

Variable Elimination

- Variable elimination is the dynamic programming variant of recursive conditioning.
- Give a factorization, such as

$$P(D) = \sum_C P(D | C) \sum_B P(C | B) \sum_A P(A)P(B | A)$$

it does the innermost sums first, constructing representations of the intermediate **factors**:

- ▶ $\sum_A P(A)P(B | A)$ is a factor on B ; call it $f_1(B)$.
- ▶ $\sum_B P(C | B)f_1(B)$ is a factor on C .
- Lecture covers:
 - ▶ Factors and factor arithmetic
 - ▶ Variable elimination algorithm

- A **factor** is a representation of a function from a tuple of random variables into a number.
- We write factor f on variables X_1, \dots, X_j as $f(X_1, \dots, X_j)$.

- A **factor** is a representation of a function from a tuple of random variables into a number.
- We write factor f on variables X_1, \dots, X_j as $f(X_1, \dots, X_j)$.
- You can assign some or all of the variables of a factor:
 - ▶ $f(X_1 = v_1, X_2, \dots, X_j)$, where $v_1 \in \text{domain}(X_1)$, is a factor on X_2, \dots, X_j .
 - ▶ $f(X_1 = v_1, X_2 = v_2, \dots, X_j = v_j)$ is a number that is the value of f when each X_i has value v_i .

The former is also written as $f(X_1, X_2, \dots, X_j)_{X_1 = v_1}$, etc.

Example factors

$r(X, Y, Z)$:

X	Y	Z	val
t	t	t	0.1
t	t	f	0.9
t	f	t	0.2
t	f	f	0.8
f	t	t	0.4
f	t	f	0.6
f	f	t	0.3
f	f	f	0.7

$r(X=t, Y, Z)$:

Y	Z	val
t	t	0.1
t	f	
f	t	
f	f	

Example factors

$r(X, Y, Z)$:

X	Y	Z	val
t	t	t	0.1
t	t	f	0.9
t	f	t	0.2
t	f	f	0.8
f	t	t	0.4
f	t	f	0.6
f	f	t	0.3
f	f	f	0.7

$r(X=t, Y, Z)$:

Y	Z	val
t	t	0.1
t	f	0.9
f	t	
f	f	

Example factors

$r(X, Y, Z)$:

X	Y	Z	val
t	t	t	0.1
t	t	f	0.9
t	f	t	0.2
t	f	f	0.8
f	t	t	0.4
f	t	f	0.6
f	f	t	0.3
f	f	f	0.7

$r(X=t, Y, Z)$:

Y	Z	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	

Example factors

$r(X, Y, Z)$:

X	Y	Z	val
t	t	t	0.1
t	t	f	0.9
t	f	t	0.2
t	f	f	0.8
f	t	t	0.4
f	t	f	0.6
f	f	t	0.3
f	f	f	0.7

$r(X=t, Y, Z)$:

Y	Z	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

Example factors

$r(X, Y, Z):$

X	Y	Z	val
t	t	t	0.1
t	t	f	0.9
t	f	t	0.2
t	f	f	0.8
f	t	t	0.4
f	t	f	0.6
f	f	t	0.3
f	f	f	0.7

$r(X=t, Y, Z):$

Y	Z	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

$r(X=t, Y, Z=f):$

Example factors

$r(X, Y, Z)$:

X	Y	Z	val
t	t	t	0.1
t	t	f	0.9
t	f	t	0.2
t	f	f	0.8
f	t	t	0.4
f	t	f	0.6
f	f	t	0.3
f	f	f	0.7

$r(X=t, Y, Z)$:

Y	Z	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

$r(X=t, Y, Z=f)$:

Y	val
t	
f	

Example factors

$r(X, Y, Z)$:

X	Y	Z	val
t	t	t	0.1
t	t	f	0.9
t	f	t	0.2
t	f	f	0.8
f	t	t	0.4
f	t	f	0.6
f	f	t	0.3
f	f	f	0.7

$r(X=t, Y, Z)$:

Y	Z	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

$r(X=t, Y, Z=f)$:

Y	val
t	0.9
f	

Example factors

$r(X, Y, Z)$:

X	Y	Z	val
t	t	t	0.1
t	t	f	0.9
t	f	t	0.2
t	f	f	0.8
f	t	t	0.4
f	t	f	0.6
f	f	t	0.3
f	f	f	0.7

$r(X=t, Y, Z)$:

Y	Z	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

$r(X=t, Y, Z=f)$:

Y	val
t	0.9
f	0.8

Example factors

$r(X, Y, Z)$:

X	Y	Z	val
t	t	t	0.1
t	t	f	0.9
t	f	t	0.2
t	f	f	0.8
f	t	t	0.4
f	t	f	0.6
f	f	t	0.3
f	f	f	0.7

$r(X=t, Y, Z)$:

Y	Z	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

$r(X=t, Y, Z=f)$:

Y	val
t	0.9
f	0.8

$r(X=t, Y=f, Z=f) =$

Example factors

$r(X, Y, Z)$:

X	Y	Z	val
t	t	t	0.1
t	t	f	0.9
t	f	t	0.2
t	f	f	0.8
f	t	t	0.4
f	t	f	0.6
f	f	t	0.3
f	f	f	0.7

$r(X=t, Y, Z)$:

Y	Z	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

$r(X=t, Y, Z=f)$:

Y	val
t	0.9
f	0.8

$r(X=t, Y=f, Z=f) = 0.8$

The **product** of factor $f_1(\overline{X}, \overline{Y})$ and $f_2(\overline{Y}, \overline{Z})$, where \overline{Y} are the variables in common, is the factor $(f_1 * f_2)(\overline{X}, \overline{Y}, \overline{Z})$ defined by:

$$(f_1 * f_2)(\overline{X}, \overline{Y}, \overline{Z}) = f_1(\overline{X}, \overline{Y})f_2(\overline{Y}, \overline{Z}).$$

Multiplying factors example

f_1 :

A	B	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

f_2 :

B	C	val
t	t	0.3
t	f	0.7
f	t	0.6
f	f	0.4

$f_1 * f_2$:

A	B	C	val
t	t	t	0.03
t	t	f	
t	f	t	
t	f	f	
f	t	t	
f	t	f	
f	f	t	
f	f	f	

Multiplying factors example

f_1 :

A	B	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

f_2 :

B	C	val
t	t	0.3
t	f	0.7
f	t	0.6
f	f	0.4

$f_1 * f_2$:

A	B	C	val
t	t	t	0.03
t	t	f	0.07
t	f	t	
t	f	f	
f	t	t	
f	t	f	
f	f	t	
f	f	f	

Multiplying factors example

f_1 :

A	B	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

f_2 :

B	C	val
t	t	0.3
t	f	0.7
f	t	0.6
f	f	0.4

$f_1 * f_2$:

A	B	C	val
t	t	t	0.03
t	t	f	0.07
t	f	t	0.54
t	f	f	
f	t	t	
f	t	f	
f	f	t	
f	f	f	

Multiplying factors example

f_1 :

A	B	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

f_2 :

B	C	val
t	t	0.3
t	f	0.7
f	t	0.6
f	f	0.4

$f_1 * f_2$:

A	B	C	val
t	t	t	0.03
t	t	f	0.07
t	f	t	0.54
t	f	f	0.36
f	t	t	
f	t	f	
f	f	t	
f	f	f	

Multiplying factors example

f_1 :

A	B	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

f_2 :

B	C	val
t	t	0.3
t	f	0.7
f	t	0.6
f	f	0.4

$f_1 * f_2$:

A	B	C	val
t	t	t	0.03
t	t	f	0.07
t	f	t	0.54
t	f	f	0.36
f	t	t	0.06
f	t	f	
f	f	t	
f	f	f	

Multiplying factors example

f_1 :

A	B	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

f_2 :

B	C	val
t	t	0.3
t	f	0.7
f	t	0.6
f	f	0.4

$f_1 * f_2$:

A	B	C	val
t	t	t	0.03
t	t	f	0.07
t	f	t	0.54
t	f	f	0.36
f	t	t	0.06
f	t	f	0.14
f	f	t	
f	f	f	

Multiplying factors example

f_1 :

A	B	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

f_2 :

B	C	val
t	t	0.3
t	f	0.7
f	t	0.6
f	f	0.4

$f_1 * f_2$:

A	B	C	val
t	t	t	0.03
t	t	f	0.07
t	f	t	0.54
t	f	f	0.36
f	t	t	0.06
f	t	f	0.14
f	f	t	0.48
f	f	f	

Multiplying factors example

f_1 :

A	B	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

f_2 :

B	C	val
t	t	0.3
t	f	0.7
f	t	0.6
f	f	0.4

$f_1 * f_2$:

A	B	C	val
t	t	t	0.03
t	t	f	0.07
t	f	t	0.54
t	f	f	0.36
f	t	t	0.06
f	t	f	0.14
f	f	t	0.48
f	f	f	0.32

Summing out variables

We can **sum out** a variable, say X_1 with domain $\{v_1, \dots, v_k\}$, from factor $f(X_1, \dots, X_j)$, resulting in a factor on X_2, \dots, X_j defined by:

$$\begin{aligned} & \left(\sum_{X_1} f \right) (X_2, \dots, X_j) \\ &= f(X_1 = v_1, \dots, X_j) + \dots + f(X_1 = v_k, \dots, X_j) \end{aligned}$$

Summing out a variable example

f_3 :

A	B	C	val
t	t	t	0.03
t	t	f	0.07
t	f	t	0.54
t	f	f	0.36
f	t	t	0.06
f	t	f	0.14
f	f	t	0.48
f	f	f	0.32

$\sum_B f_3$:

A	C	val
t	t	0.57
t	f	
f	t	
f	f	

Summing out a variable example

f_3 :

A	B	C	val
t	t	t	0.03
t	t	f	0.07
t	f	t	0.54
t	f	f	0.36
f	t	t	0.06
f	t	f	0.14
f	f	t	0.48
f	f	f	0.32

$\sum_B f_3$:

A	C	val
t	t	0.57
t	f	0.43
f	t	
f	f	

Summing out a variable example

f_3 :

A	B	C	val
t	t	t	0.03
t	t	f	0.07
t	f	t	0.54
t	f	f	0.36
f	t	t	0.06
f	t	f	0.14
f	f	t	0.48
f	f	f	0.32

$\sum_B f_3$:

A	C	val
t	t	0.57
t	f	0.43
f	t	0.54
f	f	

Summing out a variable example

f_3 :

A	B	C	val
t	t	t	0.03
t	t	f	0.07
t	f	t	0.54
t	f	f	0.36
f	t	t	0.06
f	t	f	0.14
f	f	t	0.48
f	f	f	0.32

$\sum_B f_3$:

A	C	val
t	t	0.57
t	f	0.43
f	t	0.54
f	f	0.46

Clicker Question

If $f(W, X, Y, Z)$ is a factor on variables $\{W, X, Y, Z\}$, then $f(W, X = 3, Y = \text{true}, Z)$ is a factor on

- A $\{W, X, Y, Z\}$
- B $\{X, Y\}$
- C $\{W, Z\}$
- D $\{\}$
- E none of the above

Clicker Question

If $f(W, X, Y, Z)$ is a factor on variables $\{W, X, Y, Z\}$, then $f(W = 17, X = 3, Y = \text{true}, Z = \text{false})$ is a factor on

- A $\{W, X, Y, Z\}$
- B $\{X, Y\}$
- C $\{W, Z\}$
- D $\{\}$
- E none of the above

Clicker Question

If f is a factor on $\{W, X, Y\}$ and

g is a factor on $\{W, U\}$

$(f * g)$ is a factor on

A $\{W, X, Y, U\}$

B $\{X, Y, U\}$

C $\{W\}$

D $\{f, g, W, X, Y, U\}$

E there is not enough information to tell

Clicker Question

If $f(W=3, X=4, Y=5) = 10$ and
 $g(W=3, U=12) = 15$
 $(f * g)(W=3, X=4, Y=5, U=12) =$

- A a factor on $\{W, X, Y, U\}$
- B 25
- C 150
- D none of the above
- E there is not enough information to tell

Exercise

Given factors:

s:

A	val
t	0.75
f	0.25

t:

A	B	val
t	t	0.6
t	f	0.4
f	t	0.2
f	f	0.8

o:

A	val
t	0.3
f	0.1

What are the following a function of?

i) $s * t$

A $\{A\}$

B $\{B\}$

C $\{A, B\}$

D $\{\}$

E none of the above

Exercise

Given factors:

s:

A	val
t	0.75
f	0.25

t:

A	B	val
t	t	0.6
t	f	0.4
f	t	0.2
f	f	0.8

o:

A	val
t	0.3
f	0.1

What are the following a function of?

i) $s * t$

A $\{A\}$

B $\{B\}$

C $\{A, B\}$

D $\{\}$

E none of the above

Exercise

Given factors:

s:

A	val
t	0.75
f	0.25

t:

A	B	val
t	t	0.6
t	f	0.4
f	t	0.2
f	f	0.8

o:

A	val
t	0.3
f	0.1

What are the following a function of?

i) $s * t$

A $\{A\}$

ii) $\sum_B(s * t)$

B $\{B\}$

C $\{A, B\}$

D $\{\}$

E none of the above

Exercise

Given factors:

s:

A	val
t	0.75
f	0.25

t:

A	B	val
t	t	0.6
t	f	0.4
f	t	0.2
f	f	0.8

o:

A	val
t	0.3
f	0.1

What are the following a function of?

- i) $s * t$
 - ii) $\sum_B (s * t)$
 - iii) $s * o$
- A {A}
 - B {B}
 - C {A, B}
 - D {}
 - E none of the above

Exercise

Given factors:

s:

A	val
t	0.75
f	0.25

t:

A	B	val
t	t	0.6
t	f	0.4
f	t	0.2
f	f	0.8

o:

A	val
t	0.3
f	0.1

What are the following a function of?

- i) $s * t$ A {A}
- ii) $\sum_B (s * t)$ B {B}
- iii) $s * o$ C {A, B}
- iv) $\sum_A s * t * o$ D {}
- E none of the above

Exercise

Given factors:

s:

A	val
t	0.75
f	0.25

t:

A	B	val
t	t	0.6
t	f	0.4
f	t	0.2
f	f	0.8

o:

A	val
t	0.3
f	0.1

What are the following a function of?

- i) $s * t$ A {A}
- ii) $\sum_B (s * t)$ B {B}
- iii) $s * o$ C {A, B}
- iv) $\sum_A s * t * o$ D {}
- v) $\sum_B (\sum_A s * t * o)$ E none of the above

- To compute the posterior probability of query Q given evidence $E = e$:

$$P(Q \mid E = e)$$

=

- To compute the posterior probability of query Q given evidence $E = e$:

$$\begin{aligned} P(Q \mid E = e) \\ &= \frac{P(Q, E = e)}{P(E = e)} \\ &= \end{aligned}$$

- To compute the posterior probability of query Q given evidence $E = e$:

$$\begin{aligned} P(Q \mid E = e) &= \frac{P(Q, E = e)}{P(E = e)} \\ &= \frac{P(Q, E = e)}{\sum_Q P(Q, E = e)}. \end{aligned}$$

- To compute the posterior probability of query Q given evidence $E = e$:

$$\begin{aligned} P(Q \mid E = e) &= \frac{P(Q, E = e)}{P(E = e)} \\ &= \frac{P(Q, E = e)}{\sum_Q P(Q, E = e)}. \end{aligned}$$

- So the computation reduces to the probability of $P(Q, E = e)$

- To compute the posterior probability of query Q given evidence $E = e$:

$$\begin{aligned} P(Q \mid E = e) &= \frac{P(Q, E = e)}{P(E = e)} \\ &= \frac{P(Q, E = e)}{\sum_Q P(Q, E = e)}. \end{aligned}$$

- So the computation reduces to the probability of $P(Q, E = e)$
- then normalize at the end.

Probability of a conjunction

- The variables of the belief network are X_1, \dots, X_n .
- The evidence is $Y_1 = v_1, \dots, Y_j = v_j$
- To compute $P(Q, Y_1 = v_1, \dots, Y_j = v_j)$:

Probability of a conjunction

- The variables of the belief network are X_1, \dots, X_n .
- The evidence is $Y_1 = v_1, \dots, Y_j = v_j$
- To compute $P(Q, Y_1 = v_1, \dots, Y_j = v_j)$:
we add the other variables,
 $Z_1, \dots, Z_k = \{X_1, \dots, X_n\} - \{Q\} - \{Y_1, \dots, Y_j\}$.
and sum them out.

Probability of a conjunction

- The variables of the belief network are X_1, \dots, X_n .
- The evidence is $Y_1 = v_1, \dots, Y_j = v_j$
- To compute $P(Q, Y_1 = v_1, \dots, Y_j = v_j)$:
we add the other variables,
 $Z_1, \dots, Z_k = \{X_1, \dots, X_n\} - \{Q\} - \{Y_1, \dots, Y_j\}$.
and sum them out.
- We order the Z_i into an **elimination ordering**.

$$P(Q, Y_1 = v_1, \dots, Y_j = v_j)$$

=

Probability of a conjunction

- The variables of the belief network are X_1, \dots, X_n .
- The evidence is $Y_1 = v_1, \dots, Y_j = v_j$
- To compute $P(Q, Y_1 = v_1, \dots, Y_j = v_j)$:
we add the other variables,
 $Z_1, \dots, Z_k = \{X_1, \dots, X_n\} - \{Q\} - \{Y_1, \dots, Y_j\}$.
and sum them out.
- We order the Z_i into an **elimination ordering**.

$$\begin{aligned} & P(Q, Y_1 = v_1, \dots, Y_j = v_j) \\ &= \sum_{Z_k} \cdots \sum_{Z_1} P(X_1, \dots, X_n)_{Y_1 = v_1, \dots, Y_j = v_j} \\ &= \end{aligned}$$

Probability of a conjunction

- The variables of the belief network are X_1, \dots, X_n .
- The evidence is $Y_1 = v_1, \dots, Y_j = v_j$
- To compute $P(Q, Y_1 = v_1, \dots, Y_j = v_j)$:
we add the other variables,
 $Z_1, \dots, Z_k = \{X_1, \dots, X_n\} - \{Q\} - \{Y_1, \dots, Y_j\}$.
and sum them out.
- We order the Z_i into an **elimination ordering**.

$$\begin{aligned} & P(Q, Y_1 = v_1, \dots, Y_j = v_j) \\ &= \sum_{Z_k} \cdots \sum_{Z_1} P(X_1, \dots, X_n)_{Y_1 = v_1, \dots, Y_j = v_j} \\ &= \sum_{Z_k} \cdots \sum_{Z_1} \prod_{i=1}^n P(X_i \mid \text{parents}(X_i))_{Y_1 = v_1, \dots, Y_j = v_j} \end{aligned}$$

Computing sums of products

Computation in belief networks reduces to computing the sums of products.

- How can we compute $ab + ac$ efficiently?

Computing sums of products

Computation in belief networks reduces to computing the sums of products.

- How can we compute $ab + ac$ efficiently?
- Distribute out a giving $a(b + c)$

Computation in belief networks reduces to computing the sums of products.

- How can we compute $ab + ac$ efficiently?
- Distribute out a giving $a(b + c)$
- How can we compute $\sum_{Z_1} \prod_{i=1}^n P(X_i \mid \text{parents}(X_i))$ efficiently?

Computation in belief networks reduces to computing the sums of products.

- How can we compute $ab + ac$ efficiently?
- Distribute out a giving $a(b + c)$
- How can we compute $\sum_{Z_1} \prod_{i=1}^n P(X_i | \text{parents}(X_i))$ efficiently?
- Distribute out those factors that don't involve Z_1 .

Variable elimination algorithm

To compute $P(Q \mid Y_1 = v_1 \wedge \dots \wedge Y_j = v_j)$:

- Construct a factor for each conditional probability.

Variable elimination algorithm

To compute $P(Q \mid Y_1 = v_1 \wedge \dots \wedge Y_j = v_j)$:

- Construct a factor for each conditional probability.
- Set the observed variables to their observed values.

Variable elimination algorithm

To compute $P(Q \mid Y_1 = v_1 \wedge \dots \wedge Y_j = v_j)$:

- Construct a factor for each conditional probability.
- Set the observed variables to their observed values.
- Sum out each of the non-observed non-query variables (the $\{Z_1, \dots, Z_k\}$) according to some elimination ordering.

Variable elimination algorithm

To compute $P(Q \mid Y_1 = v_1 \wedge \dots \wedge Y_j = v_j)$:

- Construct a factor for each conditional probability.
- Set the observed variables to their observed values.
- Sum out each of the non-observed non-query variables (the $\{Z_1, \dots, Z_k\}$) according to some elimination ordering.
- Multiply the remaining factors.
- Normalize by dividing the resulting factor $f(Q)$ by $\sum_Q f(Q)$.

Summing out a variable

To sum out a variable Z_j from a product f_1, \dots, f_k of factors:

- Partition the factors into
 - ▶ those that don't contain Z_j , say f_1, \dots, f_i ,
 - ▶ those that contain Z_j , say f_{i+1}, \dots, f_k

Summing out a variable

To sum out a variable Z_j from a product f_1, \dots, f_k of factors:

- Partition the factors into
 - ▶ those that don't contain Z_j , say f_1, \dots, f_i ,
 - ▶ those that contain Z_j , say f_{i+1}, \dots, f_k

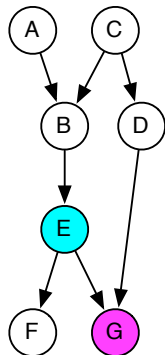
Then:

$$\sum_{Z_j} f_1 * \dots * f_k = f_1 * \dots * f_i * \left(\sum_{Z_j} f_{i+1} * \dots * f_k \right).$$

- Explicitly construct a representation of the rightmost factor. Replace the factors f_{i+1}, \dots, f_k by the new factor.

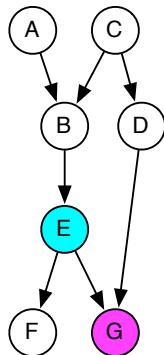
Example

$$P(E | g) =$$



Example

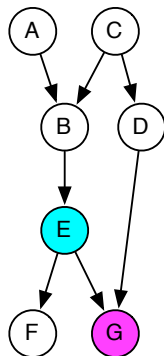
$$P(E | g) = \frac{P(E \wedge g)}{\sum_E P(E \wedge g)}$$



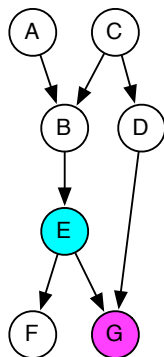
Example

$$P(E | g) = \frac{P(E \wedge g)}{\sum_E P(E \wedge g)}$$

$$P(E \wedge g) = \sum_F \sum_B \sum_C \sum_A \sum_D$$



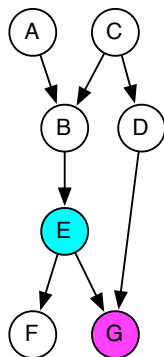
Example



$$P(E | g) = \frac{P(E \wedge g)}{\sum_E P(E \wedge g)}$$

$$\begin{aligned} P(E \wedge g) &= \sum_F \sum_B \sum_C \sum_A \sum_D P(A)P(B | AC) \\ &\quad P(C)P(D | C)P(E | B)P(F | E)P(g | ED) \\ &= \end{aligned}$$

Example



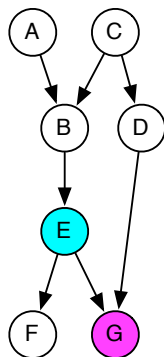
$$P(E | g) = \frac{P(E \wedge g)}{\sum_E P(E \wedge g)}$$

$$\begin{aligned} P(E \wedge g) &= \sum_F \sum_B \sum_C \sum_A \sum_D P(A)P(B | AC) \\ &\quad P(C)P(D | C)P(E | B)P(F | E)P(g | ED) \end{aligned}$$

=

$$\left(\sum_D P(D | C)P(g | ED) \right)$$

Example



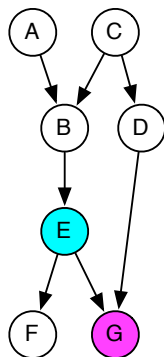
$$P(E | g) = \frac{P(E \wedge g)}{\sum_E P(E \wedge g)}$$

$$\begin{aligned} P(E \wedge g) &= \sum_F \sum_B \sum_C \sum_A \sum_D P(A)P(B | AC) \\ &\quad P(C)P(D | C)P(E | B)P(F | E)P(g | ED) \end{aligned}$$

=

$$\left(\sum_A P(A)P(B | AC) \right) \left(\sum_D P(D | C)P(g | ED) \right)$$

Example



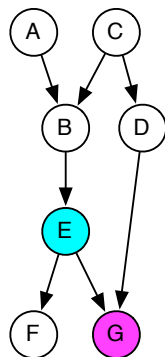
$$P(E | g) = \frac{P(E \wedge g)}{\sum_E P(E \wedge g)}$$

$$P(E \wedge g) = \sum_F \sum_B \sum_C \sum_A \sum_D P(A)P(B | AC) \\ P(C)P(D | C)P(E | B)P(F | E)P(g | ED)$$

=

$$\sum_C \left(P(C) \left(\sum_A P(A)P(B | AC) \right) \right. \\ \left. \left(\sum_D P(D | C)P(g | ED) \right) \right)$$

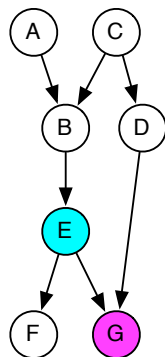
Example



$$P(E | g) = \frac{P(E \wedge g)}{\sum_E P(E \wedge g)}$$

$$\begin{aligned} P(E \wedge g) &= \sum_F \sum_B \sum_C \sum_A \sum_D P(A)P(B | AC) \\ &\quad P(C)P(D | C)P(E | B)P(F | E)P(g | ED) \\ &= \sum_B P(E | B) \sum_C \left(P(C) \left(\sum_A P(A)P(B | AC) \right) \right. \\ &\quad \left. \left(\sum_D P(D | C)P(g | ED) \right) \right) \end{aligned}$$

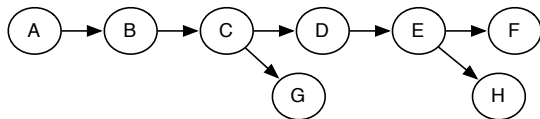
Example



$$P(E | g) = \frac{P(E \wedge g)}{\sum_E P(E \wedge g)}$$

$$\begin{aligned} P(E \wedge g) &= \sum_F \sum_B \sum_C \sum_A \sum_D P(A)P(B | AC) \\ &\quad P(C)P(D | C)P(E | B)P(F | E)P(g | ED) \\ &= \left(\sum_F P(F | E) \right) \\ &\quad \sum_B P(E | B) \sum_C \left(P(C) \left(\sum_A P(A)P(B | AC) \right) \right. \\ &\quad \left. \left(\sum_D P(D | C)P(g | ED) \right) \right) \end{aligned}$$

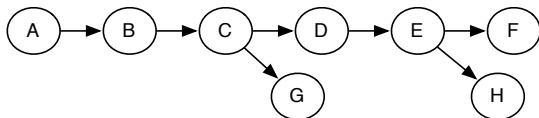
Variable Elimination example



Query: $P(G | f)$; elimination ordering: A, H, E, D, B, C

$$P(G | f) \propto$$

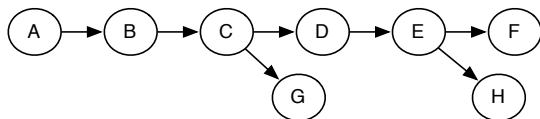
Variable Elimination example



Query: $P(G | f)$; elimination ordering: A, H, E, D, B, C

$$P(G | f) \propto \sum_C \sum_B \sum_D \sum_E \sum_H \sum_A P(A)P(B | A)P(C | B) \\ P(D | C)P(E | D)P(f | E)P(G | C)P(H | E)$$

Variable Elimination example



Query: $P(G | f)$; elimination ordering: A, H, E, D, B, C

$$P(G | f) \propto \sum_C \sum_B \sum_D \sum_E \sum_H \sum_A P(A)P(B | A)P(C | B) \\ P(D | C)P(E | D)P(f | E)P(G | C)P(H | E)$$

$$= \sum_C \left(\sum_B \left(\sum_A P(A)P(B | A) \right) P(C | B) \right) P(G | C) \\ \left(\sum_D P(D | C) \left(\sum_E P(E | D)P(f | E) \sum_H P(H | E) \right) \right)$$

Clicker Question

In variable elimination with factors:

$$f_0(W), f_1(W, X), f_2(X, Y), f_3(Y, Z)$$

If variable X is eliminated (summed out) first which factors are multiplied when summing X out:

- A none of them
- B f_1 and f_2
- C f_0 , f_1 and f_2
- D f_1 , f_2 and f_3
- E all of them

Clicker Question

In variable elimination with factors:

$$f_0(W), f_1(W, X), f_2(X, Y), f_3(Y, Z)$$

If variable Z is eliminated (summed out) first which factors are multiplied when summing Z out:

- A none of them
- B f_1 and f_2
- C f_0 , f_1 and f_2
- D f_1 , f_2 and f_3
- E all of them

Clicker Question

In variable elimination with factors:

$$f_0(W), f_1(W, X), f_2(X, Y), f_3(Y, Z)$$

If variable X is eliminated (summed out) first which factors remain after summing X out:

- A no factors remain
- B f_3 and $\sum_X f_0 * f_1 * f_2$
- C f_0, f_1, f_2, f_3 and $\sum_X f_1 * f_2$
- D f_0, f_3 and $\sum_X f_1 * f_2$
- E all of f_0, f_1, f_2, f_3

Pruning Irrelevant Variables (Belief networks)

Suppose you want to compute $P(X \mid e_1 \dots e_k)$:

- Prune any variables that have no observed or queried descendants.
- Connect the parents of any observed variable.
- Remove arc directions.
- Remove observed variables.
- Remove any variables not connected to X in the resulting (undirected) graph.

Clicker Question

Given evidence and a query, which variables can be pruned before running variable elimination

- A all of the variables that are not observed or queried
- B those variables after the query variable in the total ordering of variables that defines the belief network
- C all variables that are not observed or queried or are parents of queried variables
- D all variables that are not observed or queried and have no observed or queried descendents
- E none of the variables

Variable Elimination and Recursive Conditioning

- Variable elimination is the dynamic programming variant of recursive conditioning.
- Recursive Conditioning is the search variant of variable elimination.

Variable Elimination and Recursive Conditioning

- Variable elimination is the dynamic programming variant of recursive conditioning.
- Recursive Conditioning is the search variant of variable elimination.
- They do the same additions and multiplications.

Variable Elimination and Recursive Conditioning

- Variable elimination is the dynamic programming variant of recursive conditioning.
- Recursive Conditioning is the search variant of variable elimination.
- They do the same additions and multiplications.
- Space and time complexity $O(nd^t)$, for n variables, domain size d , and treewidth t .
 - treewidth is the number of variables in the smallest factor. It is a property of the graph and the elimination ordering.
- Recursive conditioning never modifies or creates factors; it only evaluates them.