# Inferences in Bayesian Networks
# Variable Elimination Algorithm

Alice Gao

Lecture 13

Readings: RN 14.4. PM 8.4.

# Outline

Learning Goals

Why Use the Variable Elimination Algorithm

The Variable Elimination Algorithm

Revisiting the Learning goals

# Learning Goals

By the end of the lecture, you should be able to

- Explain how we can perform probabilistic inference more efficiently using the variable elimination algorithm.
- Define factors. Manipulate factors using operations restrict, sum out, multiply and normalize.
- Describe/trace/implement the variable elimination algorithm for calculating a prior or a posterior probability given a Bayesian network.
- Explain how the elimination ordering affects the complexity of the variable elimination algorithm.
- Identify the variables that are irrelevant to a query.

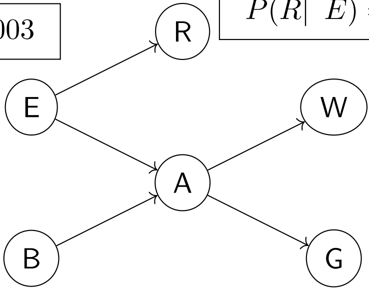
Learning Goals

Why Use the Variable Elimination Algorithm

The Variable Elimination Algorithm

Revisiting the Learning goals

# A Bayesian Network for the Holmes Scenario




$P(E) = 0.0003$

$P(B) = 0.0001$

$$P(R|\neg E) = 0.0002$$
$$P(R|\ E) = 0.9$$

$$P(W|\neg A) = 0.4$$
$$P(W|\ A) = 0.8$$

$$P(G|\neg A) = 0.04$$
$$P(G|\ A) = 0.4$$

$$P(A|\neg B \land \neg E) = 0.01$$
$$P(A|\neg B \land\ E) = 0.2$$
$$P(A|\ B \land \neg E) = 0.95$$
$$P(A|\ B \land\ E) = 0.96$$

# Answering a Question

*What is the probability that a burglary is happening given that Dr. Watson and Mrs. Gibbon both call?*

$$P(B|w \land g)$$

- Query variables: $B$
- Evidence variables: $W$ and $G$
- Hidden variables: $E$, $A$, and $R$.

# Answering the query using the joint distribution

$P(B|w \wedge g) =$

# Number of operations using the joint distribution

How many addition and multiplication operations do we need to calculate the expression below?

$$\sum_e \sum_a \sum_r P(B)P(e)P(a|B \wedge e)P(w|a)P(g|a)P(r|e)$$

(A) $\leq 10$

(B) 11-20

(C) 21-40

(D) 41-60

(E) $\geq 61$

# Answering the query using variable elimination algorithm

$$\sum_{e} \sum_{a} \sum_{r} P(B)P(e)P(a|B \wedge e)P(w|a)P(g|a)P(r|e)$$

# Number of operations via the variable elimination algorithm

How many addition and multiplication operations do we need to calculate the expression below?

$$P(B)\sum_{e}P(e)\sum_{a}P(a|B\wedge e)P(w|a)P(g|a)$$

(A) $\leq$ 10
(B) 11-20
(C) 21-40
(D) 41-60
(E) $\geq$ 61

Learning Goals

Why Use the Variable Elimination Algorithm

The Variable Elimination Algorithm

Revisiting the Learning goals

# Introducing the Variable Elimination Algorithm

- Performing probabilistic inference is challenging.
- Exact and approximate inferences.
- The variable elimination algorithm uses dynamic programming and exploits the conditional independence.

# Factors

- A function from some random variables to a number.
- $f(X_1, \ldots, X_j)$: a factor $f$ on variables $X_1, \ldots, X_j$.
- A factor can represent a joint or a conditional distribution. For example, $f(X_1, X_2)$ can represent $P(X_1 \wedge X_2)$, $P(X_1|X_2)$ or $P(X_1 \wedge X_3 = v_3|X_2)$.
- Define a factor for every conditional probability distribution in the Bayes net.

# Restrict a factor

**Restrict** a factor.

- Assign each evidence variable to its observed value.
- Restricting $f(X_1, X_2, \ldots, X_j)$ to $X_1 = v_1$, produces a new factor $f(X_1 = v_1, X_2, \ldots, X_j)$ on $X_2, \ldots, X_j$.
- $f(X_1 = v_1, X_2 = v_2, \ldots, X_j = v_j)$ is a number.

# Restrict a factor

$f_1(X, Y, Z)$:

| $X$ | $Y$ | $Z$ | 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 |

- What is $f_2(Y, Z) = f_1(x, Y, Z)$?
- What is $f_3(Y) = f_2(Y, \neg z)$?
- What is $f_4() = f_3(\neg y)$?

# Sum out a variable

**Sum out** a variable.

Summing out $X_1$ with domain $\{v_1, \ldots, v_k\}$ from factor $f(X_1, \ldots, X_j)$, produces a factor on $X_2, \ldots, X_j$ defined by:

$$(\sum_{X_1} f)(X_2, \ldots, X_j) = f(X_1 = v_1, \ldots, X_j) + \cdots + f(X_1 = v_k, \ldots, X_j)$$

# Sum out a variable

$f_1(X, Y, Z)$:

| $X$ | $Y$ | $Z$ | 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 |

What is $f_2(X, Z) = \sum_Y f_1(X, Y, Z)$?

# Multiplying factors

**Multiply** two factors together.

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

$$(f_1 \times f_2)(X, Y, Z) = f_1(X, Y) * f_2(Y, Z).$$

# Multiplying factors

$f_1$:

| $X$ | $Y$ | val |
|---|---|---|
| t | t | 0.1 |
| t | f | 0.9 |
| f | t | 0.2 |
| f | f | 0.8 |

$f_2$:

| $Y$ | $Z$ | val |
|---|---|---|
| t | t | 0.3 |
| t | f | 0.7 |
| f | t | 0.6 |
| f | f | 0.4 |

What is $f_1(X, Y) \times f_2(Y, Z)$?

# Normalize a factor

- Convert it to a probability distribution.
- Divide each value by the sum of all the values.

$f_1$:

| $Y$ | val |
|---|---|
| t | 0.2 |
| f | 0.6 |

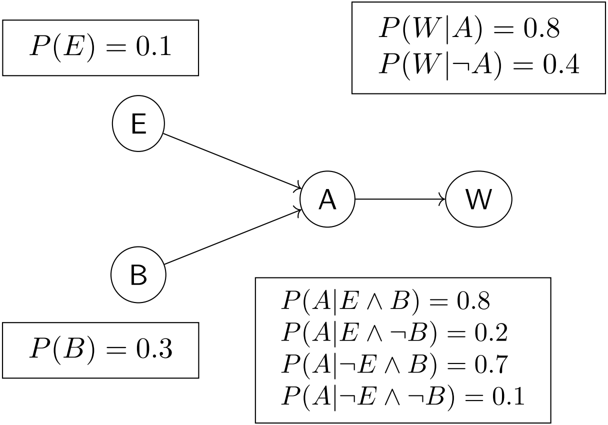
# Variable elimination algorithm

To compute $P(X_q | X_{o_1} = v_1 \land \ldots \land X_{o_j} = v_j)$:

- **Construct a factor** for each conditional probability distribution.
- **Restrict** the observed variables to their observed values.
- Eliminate each hidden variable $X_{h_j}$.
  - **Multiply** all the factors that contain $X_{h_j}$ to get new factor $g_j$.
  - **Sum out** the variable $X_{h_j}$ from the factor $g_j$.
- **Multiply** the remaining factors.
- **Normalize** the resulting factor.

# Example of VEA

Given a portion of the Holmes network below,
calculate $P(B|\neg A)$ using the variable elimination algorithm.



# Revisiting the Learning Goals

By the end of the lecture, you should be able to

- Explain how we can perform probabilistic inference more efficiently using the variable elimination algorithm.
- Define factors. Manipulate factors using operations restrict, sum out, multiply and normalize.
- Describe/trace/implement the variable elimination algorithm for calculating a prior or a posterior probability given a Bayesian network.
- Explain how the elimination ordering affects the complexity of the variable elimination algorithm.
- Identify the variables that are irrelevant to a query.