Inferences in Bayesian Networks

Alice Gao

Lecture 12

Based on work by K. Leyton-Brown, K. Larson, and P. van Beek

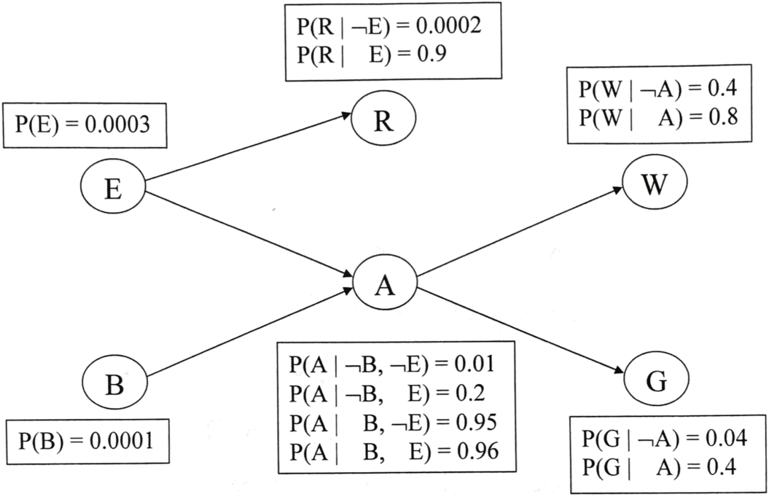
Outline

Learning Goals

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

- Define factors. Manipulate factors using operations restrict, sum out, multiply and normalize.
- Compute a prior or a posterior probability given a Bayesian network.
- Describe/trace/implement the variable elimination algorithm for calculating a prior or a posterior probability given a Bayesian network.

A Bayesian Network for the Holmes Scenario



Answering a Question

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

$$P(B|W \wedge G)$$

- Query variables: B
- Evidence variables: W and G
- Hidden variables: E and A (Let's drop Radio.)

Answering a Question

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

$$P(B|W \land G) = \frac{P(B \land W \land G)}{P(B \land W \land G) + P(\neg B \land W \land G)}$$

Answering this question require computing two joint probabilities. Let's compute the first one $P(B \land W \land G)$.

Number of operations

How many addition and multiplication operations do we need to perform to evaluate the following expression?

$$\sum_{e \in \{t,f\}} \sum_{a \in \{t,f\}} \sum_{r \in \{t,f\}} P(B=t)P(E=e)P(A=a|B=t \wedge E=e) \\ P(R=r|E=e)P(G=t|A=a)P(W=t|A=a)$$

(A) Less than 10

(B) 10-25

(C) 26-40

(D) 41-55

(E) More than 55

Number of operations

How many addition and multiplication operations do we need to perform to evaluate the following expression?

$$
\begin{aligned}
&P(B = t \wedge W = t \wedge G = t) \\
&= P(B = t) \sum_{a \in \{t,f\}} P(G = t | A = a) P(W = t | A = a) \\
&\qquad \sum_{e \in \{t,f\}} P(E = e) P(A = a | B = t \wedge E = e)
\end{aligned}
$$

(A) Less than 10

(B) 10-25

(C) 26-40

(D) 41-55

(E) More than 55

Factors

- A factor is a representation of a function from a tuple of random variables into a number.
- We will write factor f on variables $X_1, \ldots, X_j$ as $f(X_1, \ldots, X_j)$.
- A factor denotes a distribution over the given tuple of variables in some (unspecified) context
  - e.g., $P(X_1, X_2)$ is a factor $f(X_1, X_2)$
  - e.g., $P(X_1, X_2, X_3 = v_3)$ is a factor $f(X_1, X_2)$
  - e.g., $P(X_1, X_3 = v_3 | X_2)$ is a factor $f(X_1, X_2)$

Restrict a factor

- Our first operation: we can assign values to some variables of a factor.
  - $f(X_1 = v_1, X_2, \ldots, X_j)$, where $v_1 \in dom(X_1)$, is a factor on $X_2, \ldots, X_j$.
  - $f(X_1 = v_1, X_2 = v_2, \ldots, X_j = v_j)$ is a number that is the value of f when each X_i has value v_i.

Example: Restrict a factor

$r(X, Y, Z)$:

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

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

Example: Restrict a factor

What is $f_1(A = t, B)$?

f_1:

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

Sum out variables

Our second operation: we can sum out a variable, say X_1 with domain $\{v_1, \ldots, v_k\}$, from factor $f(X_1, \ldots, X_j)$, resulting in a factor on $X_2, \ldots, X_j$ defined by:

$$\begin{aligned} &(\sum_{X_1} f)(X_2, \ldots, X_j) \\ &= \quad f(X_1 = v_1, \ldots, X_j) + \cdots + f(X_1 = v_k, \ldots, X_j) \end{aligned}$$

Example: Sum out variables

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

Example: Sum out variables

What is $\sum_C f_3$?

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

Multiplying factors

- Our third operation: factors can be multiplied together.
- 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 $(f_1 \times f_2)(\overline{X}, \overline{Y}, \overline{Z})$ defined by:

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

- Note: it's defined on all $\overline{X}, \overline{Y}, \overline{Z}$ triples, obtained by multiplying together the appropriate pair of entries from f_1 and f_2.

Example: Multiplying factors

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

$f_1 \times f_2$:

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

Example: Multiplying factors

What is $f_1 \times f_2$?

f_1:

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

f_2:

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

Variable elimination algorithm

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

- Construct a factor for each conditional probability.
- Set the observed variables to their observed values.
- For each of the other variables $X_{s_i} \in \{X_{s_1}, \ldots, X_{s_k}\}$, sum out X_{s_i}
- Multiply the remaining factors.
- Normalize by dividing the resulting factor $f(X_q)$ by $\sum_{X_q} f(X_q)$.

Revisiting the Learning Goals

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

- Define factors. Manipulate factors using operations restrict, sum out, multiply and normalize.
- Compute a prior or a posterior probability given a Bayesian network.
- Describe/trace/implement the variable elimination algorithm for calculating a prior or a posterior probability given a Bayesian network.