Given the Bayesian network for the Holmes scenario (included on the last page), we would like to answer the following question:

What is the probability that a burglary is happening given that Dr. Watson and Mrs. Gibbon both call? In other words, what is $P(B|W \wedge G)$?

A few terminologies:

- B is a query variable.
- W and G are evidence variables.
- All other variables are hidden/latent variables.

$$\begin{aligned} P(B|W \wedge G) &= \frac{P(B \wedge W \wedge G)}{P(W \wedge G)} \\ &= \frac{P(B \wedge W \wedge G)}{\sum_{b \in \{t,f\}} P(B = b \wedge W \wedge G)} \text{ by the sum rule} \\ &= \frac{P(B \wedge W \wedge G)}{P(B \wedge W \wedge G) + P(\neg B \wedge W \wedge G)} \end{aligned}$$

Computing this conditional probability requires computing two joint probabilities.

How do we compute $P(B \wedge W \wedge G)$?

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

which again uses the sum rule.

By the properties of the Bayesian network, we can rewrite the above expression as follows.

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

How many operations do we need to do?

- Sum up 8 terms: 7 additions.
- 5 multiplications per term for 8 terms: $5 * 8 = 40$ multiplications in total.
- 47 operations in total.

Can we do better?

We will perform the computations from right to left.

- Reorder the terms.
- Sum out variables one by one.
- Cache intermediate results.

If a term does not involve the variable being summed, we can pull that term out of that summation. Pull $P(B = t)$ out of all three summations.

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

R is involved in only one term. So let's sum out R first. After reordering, we will move each summation to the right until the summation is only for the terms involving the variable being summed out.

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

Let's sum out E before summing out A. Thus, we will order the terms and the computations as

A, E, and R, as follows.

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

Notice that $\sum_{r \in \{t,f\}} P(R = r | E = e) = 1$. We can remove the last summation right away. Why could we do this?

Rules:

- Any leaf node that is neither a query variable nor an evidence variable is irrelevant to the query.
- Every variable that is not an ancestor of a query variable or evidence variable is irrelevant to the query.

We need to evaluate the following quantity.

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

How many operations do we need to do?

- For each value of A, for the summation over E, 2 multiplications and 1 addition. $3 * 2 = 6$ operations.
- For the summation over A, 4 multiplications and 1 addition.
- Finally, 1 more multiplication.
- In total, 9 multiplications and 3 additions. (12 operations in total.)

The Holmes scenario

Mr. Holmes lives in a high crime area and therefore has installed a burglar alarm. He relies on his neighbors to phone him when they hear the alarm sound. Mr. Holmes has two neighbors, Dr. Watson and Mrs. Gibbon.

Unfortunately, his neighbors are not entirely reliable. Dr. Watson is known to be a tasteless practical joker and Mrs. Gibbon, while more reliable in general, has occasional drinking problems.

Mr. Holmes also knows from reading the instruction manual of his alarm system that the device is sensitive to earthquakes and can be triggered by one accidentally. He realizes that if an earthquake has occurred, it would surely be on the radio news.