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.

Two questions to consider

- Given a Bayesian network, is it a correct and good representation of the domain?
- How do we construct a Bayesian network that is a correct and good representation of the domain?

A Bayes network is a **correct** representation of the domain iff

- it makes the correct independence assumptions.

Among all the **correct** Bayes network representations,
a Bayesian network is a **good** representation of the domain iff

- the number of required probabilities is relatively small, and
- the probabilities required are natural to specify.

How do we construct a correct Bayesian network for a domain?

1. Determine the set of variables that are required to model the domain.
2. Order the variables, $\{X_1, ..., X_n\}$.
3. For $i = 1$ to n, do the following
   (a) Choose a minimum set of parents from X_1, ..., X_{i-1} such that $P(X_i|Parents(X_i)) = P(X_i|X_{i-1}, \ldots, X_1)$ is satisfied.
   (b) Create a link from each parent of X_i to X_i.
   (c) Write down the conditional probability table $P(X_i|Parents(X_i))$.

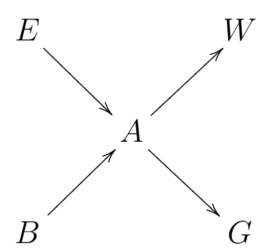
Construct a correct Bayesian network using the following ordering. (Let's drop Radio.)

$$B, E, A, W, G$$

Add each node to the Bayes network.

1. If E has no parent, we are claiming that B and E are independent, which is true.
2. If A has no parent, then we are claiming that A is independent of both E and B, which is false (A is directly affected by both B and E). If A has E has its only parent, then we are claiming that given E, A and B are independent, which is false (Given E, A is still directly affected by B). By the same reasoning, A cannot have B as its only parent. Thus, A must have both B and E as its parents.
3. Whether Watson calls directly depends on Alarm, and does not directly depend on Earthquake or Burglary. In other words, we are assuming that Watson does not directly observe Earthquake nor Burglary.

   No parent for W does not work because W is directly affected by A. If W has E, B, or both E and B as parents, then we are claiming that W is independent of A given E and B, which is false. If W has A has its only parents, then we are claiming that W is independent of B and E given A, which is true.
4. Using similar reasoning as W, G is independent of B and E given A. Thus, we can choose A as the only parent of G.



Some observations:

- A separates B and E from W and G. Given A, the two sides of the network are independent.
- A casual model: links are from causes to effects/consequences.
- This requires 10 probabilities in total. $1 + 1 + 4 + 2 + 2 = 10$.

Construct a correct Bayesian network using the following ordering. (Let's drop Radio.)

$$W, G, A, B, E$$

Compared to the previous ordering, we have switched the places of the W, G pair and the B, E pair.

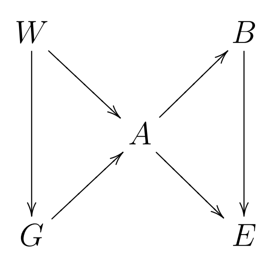
Add each node to the Bayes network.

1. If G has no parent, then we are claiming that G is independent of W, which is not the case. (W and G are dependent and they affect each other through A.) Therefore, G must have W as its parent.

2. If A has no parent, then we are claiming that A is independent of W and G, which is not the case (A directly influences both W and G.) If A has W as its only parent, then we are claiming that given W, A and G are independent, which is not the case (Given W, A still directly influences G). By the same reasoning, A cannot have G as its only parent. Thus, the only remaining possibility is for A to have both W and G as its parents.

3. Similar to the previous network, A separates B and E from W and G. Given A, the two sides of the network are independent. Given A, B is independent of W and G. Thus, we can have A as the only parent of B.

4. We know that given A, E is independent of W and G. Thus, we should choose parents of E amongst A and B. If B is the only parent of E, then we are claiming that W, G, and A are independent of E given B, which is false. If If A is the only parent of E, then we are claiming that W, G, and B are independent of E given A, which is false. W and G are independent of E given A. However, B is not independent of E given A. Both Burglary and Earthquake are independent causes of Alarm. Given Alarm, Burglary and Earthquake are dependent on each other. Thus, the only remaining possibility is to have both A and B as parents of E.



This requires 13 probabilities in total. $1 + 2 + 4 + 2 + 4 = 13$.

This is a diagnostic model. The links are from effects to causes. Choosing this ordering resulted in additional dependencies between independent causes and unrelated effects.

- Watson and Gibbon are independent given Alarm. If we consider Alarm after Watson and Gibbon, then Watson and Gibbon are dependent. We cannot take advantage of the independence property.
- Burglary and Earthquake are independent causes of Alarm. If we consider Burglary and Earthquake after Alarm, then they become dependent (Given Alarm, Burglary and Earthquake affect each other.) Again, we cannot take advantage of the fact that Burglary and Earthquake are independent.

Construct a correct Bayesian network using the following ordering.

$$W, G, E, B, A$$

Compared to the previous ordering, Alarm is after B and E rather than before B and E.

Add each node to the Bayes network.

1. Gibbon is affected by Watson.
2. Earthquake is affected by both Watson and Gibbon.
3. Burglary is affected by Watson, Gibbon, and Earthquake.
4. Alarm is affected by Watson, Gibbon, Earthquake and Burglary.

You will get a fully connected network. This will require specifying 31 probabilities.

For the previous example, we chose Alarm in the middle. By doing so, we can take advantage of the fact that Alarm separates W and G from B and E. In this example, we chose Alarm last. So we no longer have this separation.