Tutorial 9 (Wk10)
Bayesian Networks

This week's lecture covers Bayesian approaches for data analysis. One part of this is the concept of Bayesian Networks. In this tutorial we will look at the theory and in next week's lab we will implement some examples using Python. The solutions will be published on Wednesday.

Introduction to Bayesian Networks

Bayesian Networks (BNs), also known as *Belief Networks*, are probabilistic graphical models that represent relationships between random variables using a Directed Acyclic Graph (DAG). Each node in the graph corresponds to a random variable, and each directed edge represents a conditional dependency. The structure of the DAG allows Bayesian Networks to compactly represent complex joint probability distributions by capturing the conditional independence relationships between variables.

Conditional Independence

A key feature of Bayesian Networks is ***conditional independence***. This means that some variables become independent of others when specific conditions (or evidence) are known.

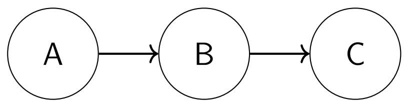
For example, consider three variables A, B, and C connected as $A \to B \to C$ in a Bayesian Network:

- The edge $A \to B$ implies that B is directly influenced by A.
- The edge $B \to C$ implies that C is directly influenced by B.

Conditional independence allows us to state that A and C are independent of each other given B. This is mathematically written as:

$$P(C \mid A, B) = P(C \mid B)$$

This means that once we know B, the probability of C no longer depends on A. This property simplifies computations and reduces the number of parameters needed to model the system.



Example

Imagine a weather prediction system.

- A: It is cloudy.
- B: It is raining.
- C: The ground is wet.

The relationship is modeled as $A \to B \to C$:
Clouds (A) cause rain (B). Rain (B) wets the ground (C).

Using conditional independence:

1. If we know it's raining ($B =$ true), the probability of the ground being wet (C) does not depend on whether it's cloudy (A).

2. This simplifies the computation of $P(A, B, C)$:

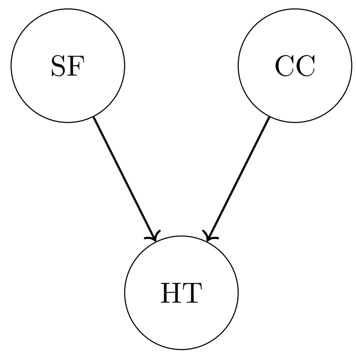
$$P(A, B, C) = P(A) \cdot P(B \mid A) \cdot P(C \mid B)$$

Example 2

Suppose we are building a system to diagnose diseases. A high temperature (HT) could be caused by:

- Seasonal Flu (SF).
- Contact with COVID-19 (CC).

Using Bayesian Networks, we model this as:



$$P(HT, SF, CC) = P(HT \mid SF, CC)P(SF)P(CC)$$

In this model:

- Seasonal flu (SF) and contact with COVID-19 (CC) are assumed to be independent.
- High temperature (HT) depends on SF and CC.

Conditional Probability Table (CPT)

CPTs quantify relationships in a Bayesian Network. The table lists probabilities of one variable given its parent(s). For example, for the node HT:

SF	CC	$P(HT = \text{true})$	$P(HT = \text{false})$
true	true	0.9	0.1
true	false	0.7	0.3
false	true	0.6	0.4
false	false	0.1	0.9

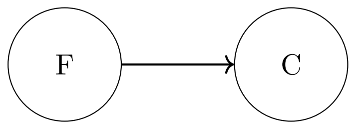
Conditional independence in this model implies that SF and CC are independent of each other but jointly influence HT. For example:

$$P(HT \mid SF, CC) = \text{value from CPT}$$

Exercises

Problem 1

Consider the following Bayesian network, where F represents "having the flu" and C represents "coughing."



CPTs:

$$P(F) = 0.1, \quad P(C \mid F = \text{true}) = 0.8, \quad P(C \mid F = \text{false}) = 0.3$$

Questions

1. Write the joint probability table for this network. (i.e., list probabilities of all combinations of variables).

   Answer: The joint probability $P(F, C)$ is computed using the chain rule of probability:

   $$P(F, C) = P(F) \cdot P(C \mid F)$$

   This means:
   Multiply the probability of F with the conditional probability of C given F.
   Using the provided values:

F	C	$P(F, C)$
true	true	$P(F = \text{true}) \cdot P(C = \text{true} \mid F = \text{true}) = 0.1 \cdot 0.8 = 0.08$
true	false	$P(F = \text{true}) \cdot P(C = \text{false} \mid F = \text{true}) = 0.1 \cdot 0.2 = 0.02$
false	true	$P(F = \text{false}) \cdot P(C = \text{true} \mid F = \text{false}) = 0.9 \cdot 0.3 = 0.27$
false	false	$P(F = \text{false}) \cdot P(C = \text{false} \mid F = \text{false}) = 0.9 \cdot 0.7 = 0.63$

2. What is $P(C = \text{true})$?

Answer: To compute $P(C = \text{true})$, we use the concept of **marginalisation**. **Marginalisation** is the process of summing over all possible values of another variable to compute the probability of a specific variable. Here, we marginalise over F:

$$P(C = \text{true}) = \sum_{F} P(C = \text{true}, F)$$

From the joint probability table:
$P(C = \text{true}, F = \text{true}) = 0.08$,
$P(C = \text{true}, F = \text{false}) = 0.27$.

Therefore:

$$P(C = \text{true}) = P(C = \text{true}, F = \text{true}) + P(C = \text{true}, F = \text{false})$$

Substituting values:

$$P(C = \text{true}) = 0.08 + 0.27 = 0.35$$

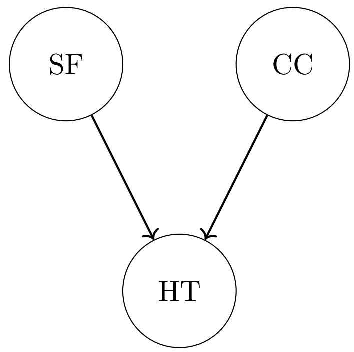
So this means there is a 35% chance of coughing, regardless of whether someone has the flu.

3. Which nodes are conditionally independent?

Answer: In the given Bayesian network:
F (Flu) is a parent of C (Coughing).
Conditional independence means that two variables are independent given the knowledge of a third variable. Here, C is directly influenced by F, so:
C and any other variable (if added to the network) would be conditionally independent given F.
For example, if there were a third variable A not directly connected to C, C and A would be independent given F. Since there are no other variables in the network, C is only dependent on F.

Problem 2

Consider a situation where Bob has a high temperature (HT), which could be caused by Seasonal Flu (SF) or Contact with COVID-19 (CC).



Questions

1. Write the joint probability distribution for this network.

   Answer: The joint probability distribution in a Bayesian Network is computed using the chain rule of probability. For this network:

   $$P(HT, SF, CC) = P(HT \mid SF, CC)P(SF)P(CC)$$

   Explanation:
   $P(SF)$ is the prior probability of having Seasonal Flu.
   $P(CC)$ is the prior probability of having Contact with COVID-19.
   $P(HT \mid SF, CC)$ is the conditional probability of Bob having a high temperature (HT) given that he has both Seasonal Flu (SF) and Contact with COVID-19 (CC).
   The joint probability combines these factors to model the full distribution of HT, SF, and CC.

2. Compute $P(HT = \text{true})$ given the following:

   $$P(SF = \text{true}) = 0.1, \quad P(CC = \text{true}) = 0.05, \quad P(HT = \text{true} \mid SF, CC) = 0.9$$

Answer: To compute $P(HT = \text{true})$, we use the **law of total probability**. The law states:

$$P(HT = \text{true}) = \sum_{SF,CC} P(HT = \text{true} \mid SF, CC) P(SF) P(CC)$$

$P(HT = \text{true})$ represents the probability that Bob has a high temperature, regardless of the specific causes (SF or CC).

We marginalise (sum) over all possible combinations of SF (true/false) and CC (true/false). This means we calculate the contribution of each scenario and add them up.
Each term in the summation is computed as:

$$P(HT = \text{true} \mid SF, CC) \cdot P(SF) \cdot P(CC)$$

Let's compute one term as an example:
If $SF = \text{true}$ and $CC = \text{true}$, the contribution is:

$$P(HT = \text{true} \mid SF = \text{true}, CC = \text{true}) \cdot P(SF = \text{true}) \cdot P(CC = \text{true})$$

$$= 0.9 \cdot 0.1 \cdot 0.05 = 0.0045$$

After computing for all combinations of SF and CC, sum them up to get the total probability:

$$P(HT = \text{true}) = 0.0045 + \dots$$

In this example, we assume only this one term contributes (for simplicity).

3. Which nodes are conditionally independent?

Answer: In this network, SF (Seasonal Flu) and CC (Contact with COVID-19) are **conditionally independent** given HT (High Temperature).
Explanation:
Conditional independence means that once we know HT (e.g., Bob has

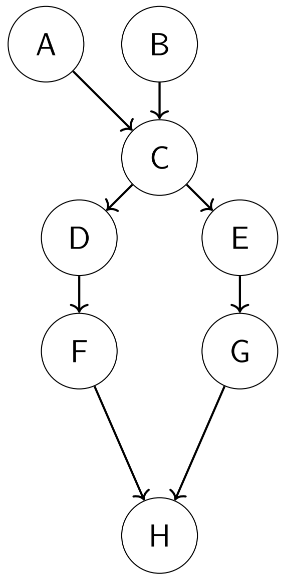
a high temperature), the knowledge of one variable (SF) does not provide additional information about the other variable (CC).
This is a common property in Bayesian Networks where multiple parents influence a single child node. The parents are conditionally independent given the child.

Mathematically:

$$P(SF, CC \mid HT) = P(SF \mid HT) \cdot P(CC \mid HT)$$

Problem 3

Consider the following Bayesian network:



Questions

1. Are D and E necessarily independent given evidence about both A and B?

   Answer: No. The path $D \to C \to E$ is not blocked.

   Independence in Bayesian networks depends on whether paths between variables are "blocked" by conditioning on other variables.
   D and E are not independent because they are both children of C, and there is a direct path ($D \to C \to E$).
   Even with evidence about A and B, the path through C remains open, as C is not conditioned on or blocked. Therefore, D and E are dependent.

2. Are A and C necessarily independent given evidence about D?

Answer: No. They are directly dependent. The path $A \to C$ is not blocked.

A is a parent of C in the network. This means there is a direct dependency between A and C.
Conditioning on D (a child of C) does not block the direct edge $A \to C$, so the two variables remain dependent.
Dependency is preserved because information about A directly influences C regardless of D.

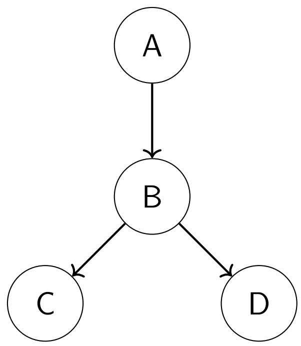
3. Are A and H necessarily independent given evidence about C?

Answer: Yes. All paths from A to H are blocked.
To check independence, we examine all possible paths from A to H in the network.
Path 1: $A \to C \to D \to F \to H$.
- This path is blocked because C is conditioned on, and it lies on the direct path.
Path 2: $A \to C \to E \to G \to H$.
- Similarly, this path is blocked by conditioning on C.
Because all paths from A to H are blocked by conditioning on C, A and H are independent given C.

Problem 4

Using the Directed Acyclic Graph (DAG) and the CPTs below, calculate the probability:

$$P(A = \textbf{true}, B = \textbf{true}, C = \textbf{true}, D = \textbf{true})$$



The graph shows the following relationships:

- A influences B,
- B influences both C and D.

The Bayesian Network is defined as:

$$P(A, B, C, D) = P(A) \cdot P(B \mid A) \cdot P(C \mid B) \cdot P(D \mid B)$$

Conditional Probability Tables (CPTs):

1. $P(A)$**: Probability of** A

A	$P(A)$
false	0.6
true	0.4

2. $P(B \mid A)$**: Probability of** B **given** A

A	B	$P(B \mid A)$
false	false	0.01
false	true	0.99
true	false	0.7
true	true	0.3

3. $P(C \mid B)$**: Probability of** C **given** B

B	C	$P(C \mid B)$
false	false	0.4
false	true	0.6
true	false	0.9
true	true	0.1

4. $P(D \mid B)$**: Probability of** D **given** B

B	D	$P(D \mid B)$
false	false	0.02
false	true	0.98
true	false	0.05
true	true	0.95

Answer: We aim to compute:

$$P(A = \text{true}, B = \text{true}, C = \text{true}, D = \text{true})$$

Using the formula:

$$P(A, B, C, D) = P(A) \cdot P(B \mid A) \cdot P(C \mid B) \cdot P(D \mid B)$$

Substituting values from the CPTs:

- $P(A = \text{true}) = 0.4$,
- $P(B = \text{true} \mid A = \text{true}) = 0.3$,
- $P(C = \text{true} \mid B = \text{true}) = 0.1$,
- $P(D = \text{true} \mid B = \text{true}) = 0.95$.

$$P(A = \text{true}, B = \text{true}, C = \text{true}, D = \text{true}) = 0.4 \cdot 0.3 \cdot 0.1 \cdot 0.95$$

Performing the calculation:

$$P(A, B, C, D) = 0.0114$$