Semester: VIII Subject: Advanced AI Academic Year:2024-2025

Module 1

Bayesian Network Example:

Ques. [MAY 2024 – 10 Marks]

A patient goes to the doctor for a medical condition, the doctor suspects three diseases as the cause of the condition. The three diseases are D1, D2, D3, which are marginally independent from each other. There are four symptoms S1, S2, S3, S4 which the doctor wants to check for presence in order to find the most probable cause of the condition. The symptoms are conditionally dependent to the three diseases as follows: S1 depends only on D1, S2 depends on D1 and D2. S3 is depends on D1 and D3, whereas S4 depends only on D3. Assume all random variables are Boolean, they are either 'true' or 'false'. i. Draw the Bayesian network for this problem.

ii. Write the expression for the joint probability distribution as a product of conditional probabilities.

iii. What is the number of independent parameters required to describe this joint distribution?

Solution:

GIVEN:

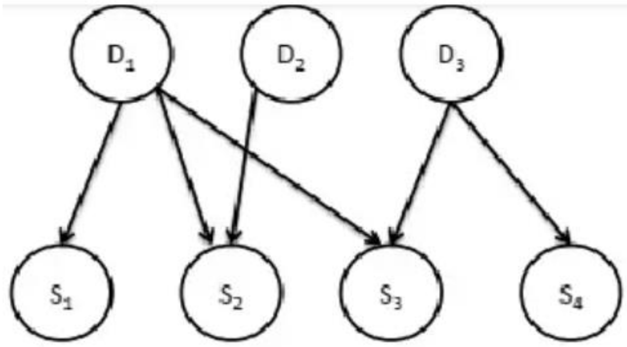
1. **Three diseases**: $D1, D2, D3$, which are **marginally independent**.
2. **Four symptoms**: $S1, S2, S3, S4$, which are conditionally dependent on diseases as follows:
   - $S1$: Depends on $D1$.
   - $S2$: Depends on $D1$ and $D2$.
   - $S3$: Depends on $D1$ and $D3$.
   - $S4$: Depends on $D3$.
3. All random variables are Boolean ($True/False$).

Semester: VIII Subject: Advanced AI Academic Year:2024-2025

1. Drawing the Bayesian Network

Nodes and Dependencies

- $D1, D2, D3$ are independent, so there are no edges between them.
- The dependencies between diseases and symptoms are:
  - $D1 \rightarrow S1, S2, S3$
  - $D2 \rightarrow S2$
  - $D3 \rightarrow S3, S4$



2. Joint Probability Distribution

The joint probability distribution can be expressed as a product of conditional probabilities based on the Bayesian network:

$$P(D1, D2, D3, S1, S2, S3, S4) = P(D1)P(D2)P(D3)P(S1|D1)P(S2|D1, D2)P(S3|D1, D3)P(S4|D3)$$

Here:

- $P(D1), P(D2), P(D3)$: Marginal probabilities of the diseases.
- $P(S1|D1), P(S2|D1, D2), P(S3|D1, D3), P(S4|D3)$: Conditional probabilities for symptoms given the diseases they depend on.

3. Calculating the Number of Independent Parameters

Step 1: Marginal probabilities for $D1, D2, D3$

Each Boolean variable requires **1 parameter** because the probability of the second state is determined by $1 - P(\text{True})$.

- $P(D1)$: 1 parameter.
- $P(D2)$: 1 parameter.
- $P(D3)$: 1 parameter.

Total for diseases: 3 parameters.

Step 2: Conditional probabilities for symptoms

Each conditional probability depends on its parent(s). We calculate the number of independent parameters required:

1. $P(S1|D1)$:
    - $D1$ has 2 states $(True, False)$.
    - For each state, $P(S1 = True)$ is independent.
    - **Parameters: 2.**
2. $P(S2|D1, D2)$:
    - $D1$ and $D2$ have $2 \times 2 = 4$ combinations of states.
    - For each combination, $P(S2 = True)$ is independent.
    - **Parameters: 4.**
3. $P(S3|D1, D3)$:
    - $D1$ and $D3$ have $2 \times 2 = 4$ combinations of states.
    - For each combination, $P(S3 = Tr$ is independent.
    - **Parameters: 4.**
4. $P(S4|D3)$:
    - $D3$ has 2 states $(True, False)$.
    - For each state, $P(S4 = True)$ is independent.
    - **Parameters: 2.**

Total for symptoms: $2 + 4 + 4 + 2 = 12$.

Step 3: Total Number of Parameters

Adding the parameters for diseases and symptoms:

$$\text{Total parameters} = \text{Parameters for diseases} + \text{Parameters for symptoms}$$

$$\text{Total parameters} = 3 + 12 = 15$$