

| **Title:** Implementation of Bayesian networks |
| --- |

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Objective:** Understand inferences in uncertain scenarios

**Expected Outcome of Experiment:**

| **Course Outcome** | **After successful completion of the course students should be able to** |
| --- | --- |
| **CO 3** | Represent and formulate the knowledge to solve the problems using various reasoning techniques |

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Books/ Journals/ Websites referred:**

1. **“Artificial Intelligence: a Modern Approach” by Russel and Norving, Pearson education Publications**
2. **“Artificial Intelligence” By Rich and knight, Tata Mcgraw Hill Publications**
3. <https://machinelearningmastery.com/introduction-to-bayesian-belief-networks/>, last retried on April 02,2025
4. [https://towardsdatascience.com/introduction-to-bayesian-belief-networks-c012e3f59f 1b](https://towardsdatascience.com/introduction-to-bayesian-belief-networks-c012e3f59f1b) , last retried on April 02,2025

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Pre Lab/ Prior Concepts:** Conditional Probability theory, Probability theory

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Historical Profile: -** Uncertainty is an inherent challenge in artificial intelligence (AI), arising from incomplete, noisy, or ambiguous information. In real-world scenarios, AI systems must make decisions despite lacking full knowledge of the environment. Addressing uncertainty is crucial for building robust and reliable AI models that can reason, learn, and adapt effectively. Bayesian networks provide a powerful probabilistic framework to represent and manage uncertainty by modelling dependencies between variables. They enable AI systems to make informed predictions, update beliefs with new evidence, and handle complex decision-making under uncertainty.

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**New Concepts to be learned:** Uncertainty, reasoning with uncertain information, Bayesian network topology

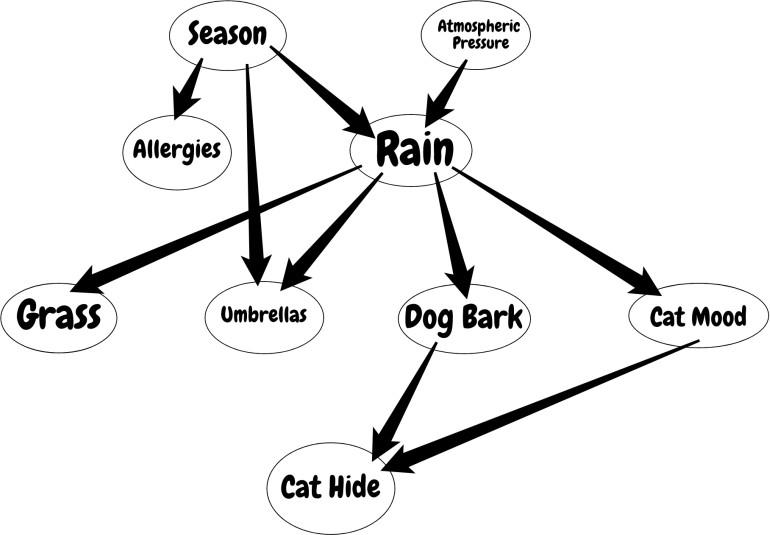
**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

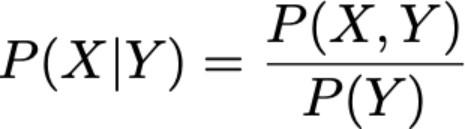
# Bayesian networks:

A Bayesian network (also known as a Bayes network, belief network, or decision network) is a probabilistic graphical model that represents a set of variables and their conditional dependencies via a directed acyclic graph (DAG). Bayesian networks are ideal for taking an event that occurred and predicting the likelihood that any one of several possible known causes was the contributing factor. For example, a Bayesian network could represent the probabilistic relationships between diseases and symptoms. Given symptoms, the network can be used to compute the probabilities of the presence of various diseases.

Efficient algorithms can perform inference and learning in Bayesian networks. Bayesian networks that model sequences of variables (e.g. speech signals or protein sequences) are called dynamic Bayesian networks. Generalizations of Bayesian networks that can represent and solve decision problems under uncertainty are called influence diagrams.

Bayesian Belief Network or Bayesian Network or Belief Network is a Probabilistic Graphical Model (PGM) that represents conditional dependencies between random variables through a Directed Acyclic Graph (DAG).



Bayesian Networks are applied in many fields. For example, disease diagnosis, optimized web search, spams filtering, gene regulatory networks, etc. And this list can be extended. The main objective of these networks is trying to understand the structure of causality relations. To clarify this, let’s consider a disease diagnosis problem. With given symptoms and their resulting disease, we construct our Belief Network and when a new patient comes, we can infer which disease or diseases may have the new patient by providing probabilities for each disease. Similarly, these causality relations can be constructed for other problems and inference techniques can be applied to interesting results.

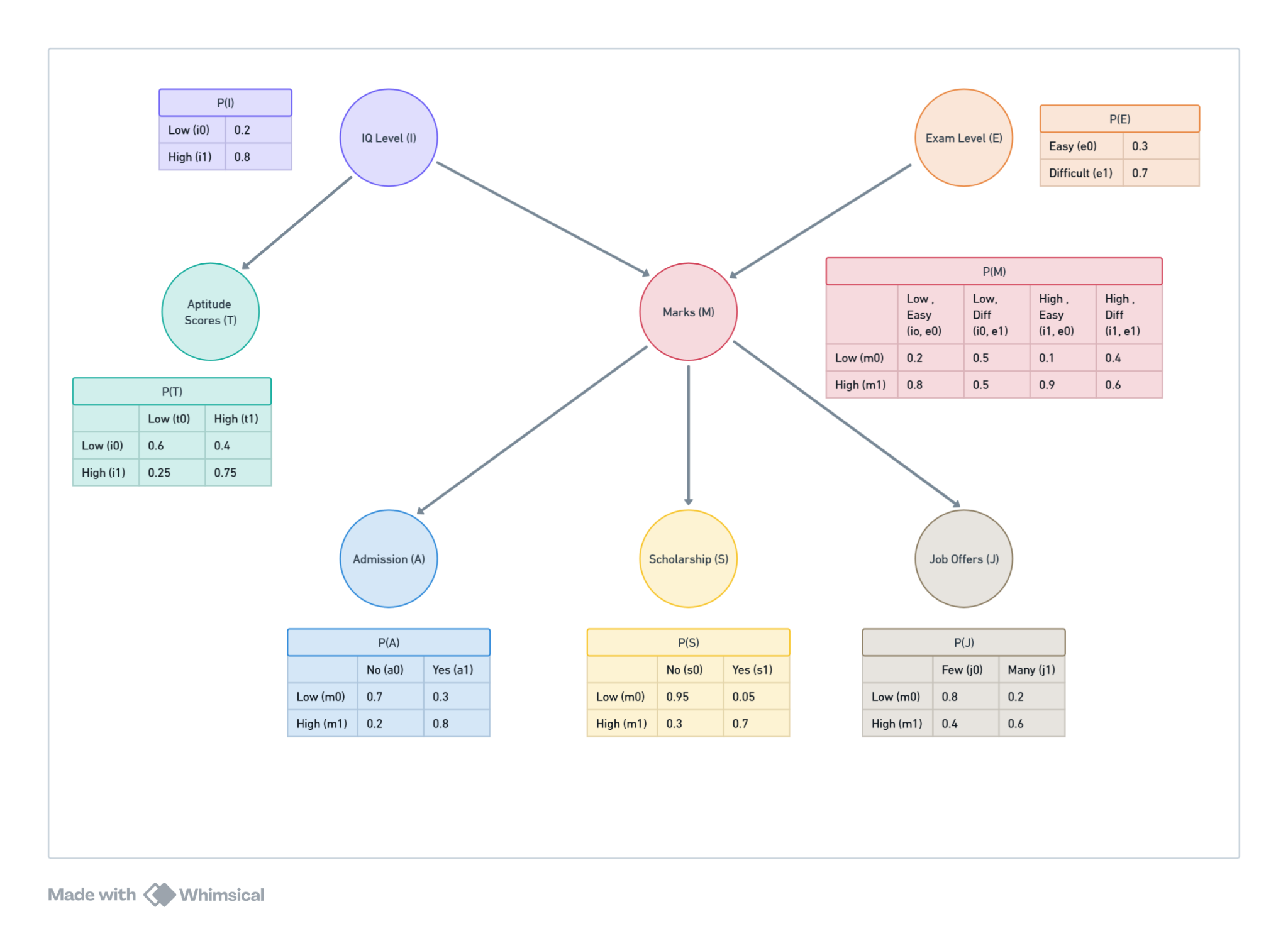
As you would understand from the formula, to be able to calculate the joint distribution we need to have conditional probabilities indicated by the network. But further that if we have the joint distribution, then we can start to ask interesting questions. For example, in the first example, we ask for the probability of “RAIN” if “SEASON” is “WINTER” and “DOG BARK” is “TRUE”.

**Work Description:** For the given problem, define the network, calculate the probabilities and query the system.

**Chosen Problem statement:**

This Bayesian Network (BN) models the relationships between various factors affecting a student's academic performance and subsequent career opportunities. The objective is to understand how factors such as **IQ Level, Exam Difficulty, and Aptitude Scores** influence **Marks**, which in turn affect **Scholarships, Admission, and Job Offers**.

This network enables probabilistic reasoning, allowing stakeholders (students, educators, and employers) to predict outcomes based on known variables and optimize educational strategies accordingly.

****

### Bayesian Network Structure:

* IQ → Marks ← Exam
* IQ → Aptitude Score
* Marks → Admission
* Marks → Scholarship
* Marks → Job Offer

### Conditional Probabilities Tables

#### 1. IQ Level (Prior Probability)

#### 

| **IQ Level (i)** | **P(i)** |
| --- | --- |
| i⁰ (Low) | 0.2 |
| i¹ (High) | 0.8 |

#### 

#### 

#### 

#### 

#### 

#### 

#### 2. Exam Level (Prior Probability)

#### 

| **Exam Level (e)** | **P(e)** |
| --- | --- |
| e⁰ (Easy) | 0.3 |
| e¹ (Hard) | 0.7 |

#### 

#### 

#### 

#### 

#### 

#### 

#### 3. Marks (Conditional on IQ and Exam Level)

#### 

| **IQ (i), Exam (e)** | **m⁰ (Low Marks)** | **m¹ (High Marks)** |
| --- | --- | --- |
| i⁰, e⁰ | 0.2 | 0.8 |
| i⁰, e¹ | 0.5 | 0.5 |
| i¹, e⁰ | 0.1 | 0.9 |
| i¹, e¹ | 0.4 | 0.6 |

#### 

#### 

#### 

#### 

#### 

#### 

#### 4. Admission (Conditional on Marks)

| **Marks (m)** | **a⁰ (No Admission)** | **a¹ (Admitted)** |
| --- | --- | --- |
| m⁰ | 0.7 | 0.3 |
| m¹ | 0.2 | 0.8 |

#### 

#### 

#### 

#### 

#### 

#### 5. Aptitude Score (Conditional on IQ)

#### 

| **IQ (i)** | **t⁰ (Low Score)** | **t¹ (High Score)** |
| --- | --- | --- |
| i⁰ | 0.6 | 0.4 |
| i¹ | 0.25 | 0.75 |

#### 

#### 

#### 

#### 

#### 

#### 6. Scholarship (Conditional on Marks)

#### 

| **Marks (m)** | **s⁰ (No Scholarship)** | **s¹ (Scholarship)** |
| --- | --- | --- |
| m⁰ | 0.95 | 0.05 |
| m¹ | 0.3 | 0.7 |

#### 

#### 

#### 

#### 

#### 

#### 

#### 7. Job Offer (Conditional on Marks)

#### 

| **Marks (m)** | **j⁰ (Few)** | **j¹ (Many)** |
| --- | --- | --- |
| m⁰ | 0.8 | 0.2 |
| m¹ | 0.4 | 0.6 |

**Implementation:**

*from* pgmpy.models *import* DiscreteBayesianNetwork

*from* pgmpy.factors.discrete *import* TabularCPD

*from* pgmpy.inference *import* VariableElimination

model = DiscreteBayesianNetwork([

('IQ', 'Marks'),

('Exam', 'Marks'),

('IQ', 'Aptitude'),

('Marks', 'Admission'),

('Marks', 'Scholarship'),

('Marks', 'JobOffer')

])

cpd\_iq = TabularCPD(*variable*='IQ', *variable\_card*=2, *values*=[[0.2], [0.8]])

cpd\_exam = TabularCPD(*variable*='Exam', *variable\_card*=2, *values*=[[0.3], [0.7]])

cpd\_marks = TabularCPD(*variable*='Marks', *variable\_card*=2,

*values*=[[0.2, 0.5, 0.1, 0.4],

[0.8, 0.5, 0.9, 0.6]],

*evidence*=['IQ', 'Exam'], *evidence\_card*=[2, 2])

cpd\_aptitude = TabularCPD(*variable*='Aptitude', *variable\_card*=2,

*values*=[[0.6, 0.25], [0.4, 0.75]],

*evidence*=['IQ'], *evidence\_card*=[2])

cpd\_admission = TabularCPD(*variable*='Admission', *variable\_card*=2,

*values*=[[0.7, 0.2], [0.3, 0.8]],

*evidence*=['Marks'], *evidence\_card*=[2])

cpd\_scholarship = TabularCPD(*variable*='Scholarship', *variable\_card*=2,

*values*=[[0.95, 0.3], [0.05, 0.7]],

*evidence*=['Marks'], *evidence\_card*=[2])

cpd\_job = TabularCPD(*variable*='JobOffer', *variable\_card*=2,

*values*=[[0.8, 0.4], [0.2, 0.6]],

*evidence*=['Marks'], *evidence\_card*=[2])

model.add\_cpds(cpd\_iq, cpd\_exam, cpd\_marks, cpd\_aptitude, cpd\_admission, cpd\_scholarship, cpd\_job)

*assert* model.check\_model()

inference = VariableElimination(model)

queries = [

('Scholarship', {'IQ': 1, 'Aptitude': 0}),

('JobOffer', {'IQ': 1, 'Scholarship': 1}),

('Exam', {'Admission': 1}),

('Aptitude', {'JobOffer': 1}),

('JobOffer', {'IQ': 0, 'Exam': 0})

]

*for* i, (var, evidence) *in* enumerate(queries, 1):

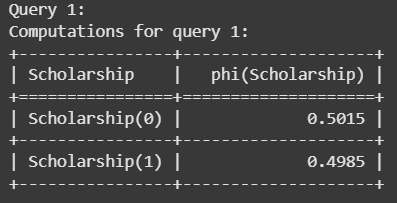
result = inference.query(*variables*=[var], *evidence*=evidence)

print(f"Query {i}:\nComputations for query {i}:\n{result}\n")

**Query 1:** What is the probability that a student with a high IQ but a low aptitude score still gets a scholarship?

(*This explores the interaction between IQ, Aptitude Score, and Scholarship while considering indirect effects through Marks.*)

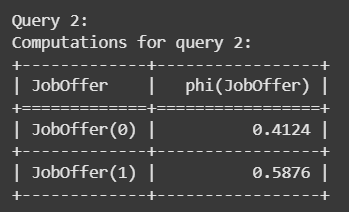
**Computations for query 1:**

****

**Query 2:** If a student has a high IQ and receives a scholarship, what is the probability that they will get many job offers?

(*This investigates how scholarship and IQ jointly influence job offers, factoring in indirect relationships through Marks and Admission.*)

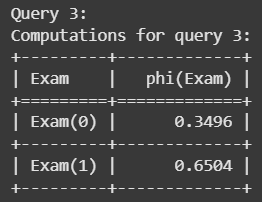
**Computations for query 2:**

****

**Query 3:** Given that a student gets admitted, what is the likelihood that they had taken a hard exam?

(*This reverses the dependency direction, using Admission as evidence to infer the difficulty level of the exam taken.*)

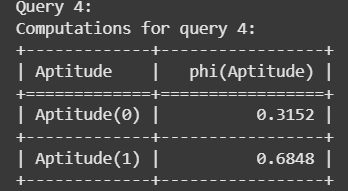
**Computations for query 3:**

****

**Query 4:** If a student receives many job offers, what is the probability that they had a high aptitude score?

(*This traces back from job success to aptitude, incorporating the intermediate influence of Marks and IQ.*)

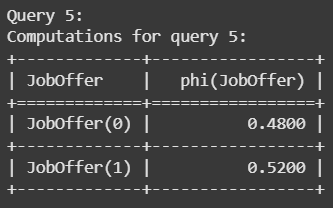
**Computations for query 4:**

****

**Query 5:** What is the probability of a student securing a job offer if they had a low IQ but took an easy exam?

(*This explores how a favorable exam condition can offset a disadvantageous IQ level in securing employment.*)

**Computations for query 5:**

****

**PostLab Questions:**

1. **Which of the following best describes a Bayesian Network?** a) A network of independent random variables  
 b) A graphical model representing conditional dependencies among variables  
 c) A deterministic rule-based AI model  
 d) A deep learning neural network

**Answer: b)**

2. **In a Bayesian Network, what do the edges between nodes represent?** a) Causal or probabilistic dependencies  
 b) Logical equivalence  
 c) Time-dependent transitions  
 d) Random connections

**Answer: a)**

**Descriptive Questions:**

**1.** **Explain the significance of Bayesian Networks in AI. How do they help in decision-making under uncertainty?**

Bayesian Networks (BNs) are probabilistic graphical models that represent relationships among variables using conditional dependencies. They help in decision-making under uncertainty by:

* **Modeling Complex Systems**: They capture dependencies between different factors, allowing for reasoning about cause and effect.
* **Handling Incomplete Data**: BNs allow predictions even when some variables are unknown by leveraging probabilistic inference.
* **Making Probabilistic Predictions**: They provide a structured way to compute the likelihood of outcomes given prior knowledge.
* **Enabling Decision Support Systems**: Many AI applications, including medical diagnosis and risk assessment, use BNs to guide decision-making based on observed data.

**2.**  **What are the main components of a Bayesian Network? Explain each briefly.**

A Bayesian Network consists of the following components:

* **Nodes (Variables):** Represent random variables in the system, such as IQ, Marks, or Job Offers.
* **Edges (Dependencies):** Directed edges between nodes represent conditional dependencies, indicating how one variable influences another.
* **Conditional Probability Tables (CPTs):** Define the probability of a variable given its parent nodes, quantifying the dependencies.
* **Inference Mechanism:** Algorithms like Variable Elimination or Belief Propagation allow for probabilistic reasoning given observed evidence.

**3.** **Suppose you have a Bayesian Network with three variables: Disease, Test Result, and Symptoms. Explain how you would use conditional probabilities to determine the likelihood that a patient has the disease given a positive test result.**

To determine the probability that a patient has the disease given a positive test result (*P(Disease | Test Result = Positive)*), we use **Bayes' Theorem**:

Steps:

1. **Define Prior Probability:** P(Disease)P(Disease)P(Disease) represents the general likelihood of having the disease.
2. **Use Conditional Probability:** P(TestResult∣Disease)P(Test Result | Disease)P(TestResult∣Disease) represents the probability of a positive test given that the patient has the disease.
3. **Calculate Evidence Probability:** P(TestResult)P(Test Result)P(TestResult) is computed using the law of total probability, considering both diseased and non-diseased cases.
4. **Compute Posterior Probability:** Apply Bayes' theorem to get the probability of having the disease given the test result.

This approach allows us to update beliefs based on observed data, a fundamental concept in Bayesian inference.

**Conclusion:** Bayesian Networks effectively model dependencies, enabling probabilistic reasoning for decision-making in education, admissions, scholarships, and career opportunities.