



**A T M E**  
College of Engineering



# MODULE 5

## Graphical Models



**A T M E**  
College of Engineering



## Introduction

Many real-world problems involve **uncertainty**.

Examples:

- Disease diagnosis
- Weather prediction
- Robot navigation
- Speech recognition
- Spam filtering



**A T M E**  
College of Engineering



# Bayesian Networks

A **Bayesian Network (BN)** is a **Directed Acyclic Graph (DAG)** in which:

- Nodes represent random variables.
- Directed edges represent conditional dependencies.
- Each node has an associated Conditional Probability Table (CPT).



**A T M E**  
College of Engineering



## Mathematical Definition

Let

$$X = \{X_1, X_2, \dots, X_n\}$$

be a set of random variables.

The joint probability distribution is

$$P(X) = \prod_{i=1}^n P(X_i \mid \text{Parents}(X_i))$$

This factorization is the key advantage of Bayesian Networks.



**A T M E**  
College of Engineering

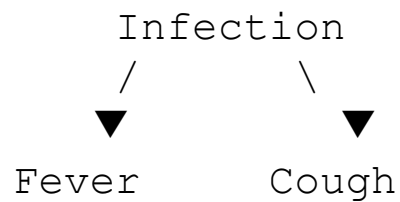


## Example

Disease Diagnosis

Variables:

- Fever
- Infection
- Cough



Interpretation

- Infection causes Fever.
- Infection causes Cough.
- Fever and Cough become conditionally independent once Infection is known.



**A T M E**  
College of Engineering



## Advantages

- Compact representation
- Handles uncertainty
- Supports probabilistic reasoning
- Efficient inference

## Applications

- Medical diagnosis
- Bioinformatics
- Fault diagnosis
- Recommendation systems
- Decision support systems



**A T M E**  
College of Engineering



# Approximate Inference

## Why Approximate Inference?

Exact inference in Bayesian Networks becomes computationally expensive when:

- Many variables exist
- Large CPTs
- Complex network structures

Exact inference is often NP-hard.

Approximate inference estimates probabilities instead of computing exact values.



**A T M E**  
College of Engineering



## **Types Sampling Methods**

- Monte Carlo Sampling
- Importance Sampling
- Likelihood Weighting
- Gibbs Sampling



**A T M E**  
College of Engineering



## Monte Carlo Sampling

Random samples are generated.

Probability is estimated as

$$P(A) \approx \frac{\text{Number of samples satisfying A}}{\text{Total samples}}$$
$$P(A) \approx \frac{\text{Number of samples satisfying A}}{\text{Total samples}}$$



**A T M E**  
College of Engineering



## Gibbs Sampling

One variable is updated at a time while keeping others fixed.

Algorithm

1. Initialize variables.
2. Update one variable.
3. Repeat.
4. Estimate probabilities.

Advantages

- Simple
- Good for large Bayesian Networks



**A T M E**  
College of Engineering



## Markov Random Fields (MRFs)

A **Markov Random Field** is an **undirected probabilistic graphical model**.

Unlike Bayesian Networks,

MRFs use **undirected edges**.



**A T M E**  
College of Engineering



## Applications

- Image segmentation
- Computer vision
- Protein structure prediction
- Object recognition

## Advantages

- Represents symmetric relationships
- Suitable for spatial data
- Flexible modeling

## Disadvantages

- Complex inference
- Partition function difficult to compute



**A T M E**  
College of Engineering



## **Hidden Markov Models (HMMs)**

**Many real-world problems involve sequences.**

**Examples**

**Speech recognition**

**Handwriting recognition**

**DNA analysis**

**Weather forecasting**

**Hidden Markov Models are designed for sequential data.**



**A T M E**  
College of Engineering



## Transition Probability

Probability of moving from one state to another.

$$A = P(S_j | S_i) \quad A = P(S_{_j} | S_{_i}) \quad A = P(S_j | S_i)$$

## Emission Probability

Probability of observation.

$$B = P(O | S) \quad B = P(O | S) \quad B = P(O | S)$$

## Initial Probability

Probability of starting state.

$$\pi = P(S_1) \quad \pi = P(S_1) \quad \pi = P(S_1)$$



**A T M E**  
College of Engineering



## **Three Fundamental Problems**

### **1. Evaluation**

Find probability of observation sequence.

Algorithm

Forward Algorithm.

### **2. Decoding**

Find hidden states.

Algorithm

Viterbi Algorithm.

### **3. Learning**

Estimate parameters.

Algorithm

Baum-Welch Algorithm.



**A T M E**  
College of Engineering



## Forward Algorithm

Computes

$$P(O \mid \lambda) P(O \mid \lambda)$$

where

O = observation sequence



**A T M E**  
College of Engineering



## Viterbi Algorithm

Finds

Most likely hidden state sequence.

Dynamic programming algorithm.



**A T M E**  
College of Engineering



## **Baum-Welch Algorithm**

Expectation-Maximization algorithm.

Learns

- Transition probabilities

Emission probabilities



**A T M E**  
College of Engineering



## Applications

- Speech recognition
- Machine translation
- POS tagging
- Gene prediction
- Gesture recognition
- Financial prediction

Robot localization



**A T M E**  
College of Engineering



## Advantages

- Excellent for sequence modeling
- Efficient algorithms
- Handles uncertainty

## Disadvantages

### First-order Markov assumption

- Limited long-range dependency modeling
- Computationally intensive for large state spaces