# Bayesian Network: Semantic and Factorization

A Bayesian Network (BN) is a probabilistic graphical model that represents a set of variables and their conditional dependencies using a Directed Acyclic Graph (DAG). It provides a compact representation of joint probability distributions (JPD) by factorizing them into local conditional probability distributions.

---

## 1. Semantic Interpretation of a Bayesian Network

The semantics of a Bayesian Network are based on conditional independence and causality:

1. Graphical Representation:  
 - Nodes represent random variables.  
 - Directed edges represent conditional dependencies.  
 - Absence of an edge means conditional independence between variables.

2. Conditional Independence Assumption:  
 - Each variable is conditionally independent of its non-descendants, given its parents.  
 - This assumption allows for efficient computation of probabilities.

3. Markov Property:  
 - Each node X is conditionally independent of its non-descendants given its parents:

P(X | All other nodes) = P(X | Parents(X))

### Example

Consider a Bayesian Network with three nodes:

Rain (R) → Wet Grass (W) ← Sprinkler (S)

- Conditional dependencies:  
 - Wet Grass (W) depends on both Rain (R) and Sprinkler (S).

Factorized probability:

P(R, S, W) = P(R) P(S) P(W | R, S)

- Conditional independence:  
 R ⊥ S (Rain and Sprinkler are independent unless otherwise specified.)

---

## 2. Factorization of a Bayesian Network

Factorization refers to the decomposition of the joint probability distribution (JPD) into smaller conditional probabilities using the chain rule.

### General Factorization Rule

Given a Bayesian Network with n nodes X1, X2, ..., Xn, the joint probability distribution is factorized as:

P(X1, X2, ..., Xn) = ∏ P(Xi | Parents(Xi)) (for i = 1 to n)

### Example

Consider a Bayesian Network with five variables:

A → B, A → C, B → D, C → D, D → E

The joint probability distribution is:

P(A, B, C, D, E) = P(A) P(B | A) P(C | A) P(D | B, C) P(E | D)

Each term represents a local conditional probability, making the computation of probabilities more efficient.

---

### Advantages of Factorization in Bayesian Networks

1. Computational Efficiency: Reduces the complexity of working with high-dimensional probability distributions.  
2. Modular Representation: Each variable is defined locally with its parents, making modeling easier.  
3. Efficient Inference: Uses message passing, belief propagation, or variable elimination for probability estimation.

---

## Summary

- Bayesian Networks represent variables and dependencies using DAGs.  
- The semantics rely on conditional independence and the Markov property.  
- Factorization breaks down the joint probability distribution (JPD) into smaller conditional probabilities for efficient computation.

This document provides a structured approach to understanding Bayesian Networks.