Probabilistic Methods in Artifical Intelligence

Hadar Tal

Hebrew University of Jerusalem, Israel

May 24, 2024

1 Bayesian Networks

Definition 1.1 (Bayesian Network)
*A **Bayesian Network** is a directed acyclic graph (DAG) $G = (V, E)$, where each node $X_i \in V$ represents a random variable, and each edge $X_i \to X_j$ represents a direct dependency between the random variables X_i and X_j. Each node X_i is associated with a conditional probability distribution $P(X_i|Pa(X_i))$, where $Pa(X_i)$ denotes the parents of X_i in the graph G.*

Theorem 1.1 (Factorization Theorem)
The joint probability distribution of a Bayesian Network can be written as:

$$P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P(X_i|Pa(X_i)) \tag{1.1}$$

Lemma 1.1 (Chain Rule)
The joint probability distribution of a set of random variables $X_1, X_2, \ldots, X_n$ can be written as:

$$P(X_1, X_2, \ldots, X_n) = P(X_1)P(X_2|X_1)P(X_3|X_1, X_2)\ldots P(X_n|X_1, X_2, \ldots, X_{n-1}) \tag{1.2}$$