Probabilistic Graphical Models
Koller
The Bayesian Networks

Ahmad Asadi

March 12, 2016

Date Performed:	March 12, 2016
Instructor:	Dr. Nickabadi

1 Basics

Why to use independency Using random variable's independencies is effective in reducing the size of joint probability distribution over them.

Chain Rule The chain rule of conditional probabilities:

$$P(X_1, X_2, \cdots, X_n) = P(x_1) \cdot P(X_2|X_1) \cdot P(X_3|X_1, X_2) \cdots P(X_n|X_{n-1}, X_{n-2}, \cdots, X_1) \tag{1}$$

Factorization Splitting the overall joint probability distribution into several conditional probability distributions (CPDs) where multiplications of CPDs reconstructs joint probability distribution.

2 Naive/Idiot Bayes

The model includes a set of classes C, some number of features $X_1, \cdots, X_n$ and assumes that *features are conditionally independent given the instance's class*:

$$P(C, X_1, \cdots, X_n) = P(C) \cdot \Pi_{i=1}^{n} P(X_i|C) \tag{2}$$

Factors in this model:

a. A prior distribution $P(C)$

b. A set of CPDs $P(X_j|C)$

3 Bayesian Networks

DAG Representation The core element of representation in bayesian networks is a *directed acyclic graph (DAG)*, in which random variables are nodes and edges correspond to direct influence of random variables on each other.
This DAG can be viewed as:

a. A data structure that provides the skeleton for representing a joint distribution compactly in a factorized way.

b. A compact representation for a set of conditional independence assumptions about a distribution.

Local Probability Models A set of local probability models that represent the nature of the dependence of each variable on its parents, containing probability distributions of single random variables along with CPDs in model.

Reasoning Patterns A joint distribution P_B specifies the probability $P_B(Y = y|E = e)$ of any event y given any observations e. We condition the joint distribution on the event $E = e$ by eliminating the entries in the joint inconsistent with our observation e, and renormalizing the resulting entries to sum to 1; we compute the probability of the event y by summing the probabilities of all of the entries in the resulting posterior distribution that are consistent with y.

Causal Reasoning is a top-down flow of probability computations in DAG from causes to symptoms.

Evidential Reasoning is a bottom-up flow of probability computations in DAG from a symptom to its causes which yields the probability of each cause to be happened.

Intercausal Reasoning is a cross flow of probability computations in DAG from a cause to another cause passing collidors in the path.

Basic Independencies in Bayesian Networks