Probabilistic Reasoning – Chapter 14.

Dhruv Rajan

October 21, 2017

Independence and *conditional independence* relationships are very powerful for modeling our world. **Bayesian Networks** introduce a systematic way of representing these relationships.

1 Representing Knowledge in an Uncertain Domain

Full joint probability distributions can answer any question about the domain, however, they become intractable as the number of variables grows. **Bayesian Networks** represent the same dependencies among variables, and give a concise specification of *any* full joint probability distribution.

A Bayesian Network is a DAG where each node is anntated with probability information, under the following specification:

1. A set of random variables makes up the nodes of the network. Variables may be discrete or continuous

2. A set of directed links or arrows connects pairs of nodes. If there is an arrow from node X to node Y, X is said to be a parent of Y.

3. Each Node X_i has a conditional probability distribution $P(X_i|Parents(X_i))$ which quantifies the effect of the parents on the node

4. The graph has no cycles (it is a DAG)

2 The Semantics of Bayesian Networks

2.1 Bayes' nets as Representations of a Full Joint Distribution

A generic query to a full joint distribution (FJD) is the probability of a conjunction of assignments to each variable: $P(X_1 = x_1 \wedge \cdots \wedge X_n = x_n)$. The following formula can be used.

$$P(x_1, \ldots x_n) = \prod_{i=1}^{n} P(x_i | parents(X_i)) \tag{1}$$

So the Bayesian network can answer any query, simply by summing the relevant joint CPDs for each node.

2.2 Constructing Bayesian Networks

The **chain rule** can be derived from the above formula, using the product rule to reduce the terms.

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i | x_{i-1}, \ldots, x_n) \tag{2}$$

Comparing this with (2), we see that if a Bayesian network is a correct encoding of the domain, then every node is conditionally indedpendent of all its predecessors, given its parents.

2.3 Compactness and Node Ordering

Bayesian Networks owe their compactness to the fact that they are locally structured. In most networks each node has some maximum number of parents k, which reduces the total amount of info needed for representation.

3 Efficient Representation of Conditional Distributions

canonical distributions can be specified completely. For example, a **deterministic node** has its value specified exactly by the values of its parents, with no uncertainty (Like figaro's Apply).