COMP538: Introduction to Bayesian Networks

Introduction to Course

Nevin L. Zhang
lzhang@cse.ust.hk

Department of Computer Science and Engineering
Hong Kong University of Science and Technology

Fall 2008

Probabilistic Modeling/Reasoning

- Describe problem domain with random variables.
- Represent knowledge about the domain as relationships among the random variables.
  - Can be learned from data.
- Inference: After observing values of some variables, make inference about other variables of interest.

Examples:

- Medical diagnosis: symptoms and diseases can be modeled as random variables.
- Business decision making: e.g. predicting customers' future behavior based past purchasing history, important for direct marketing.
- Bio-informatics: genotypes and gene expression levels can be modeled as random variables
- Computer Science: Vision, speech understanding, sensor networks, localization, software engineering, etc

Joint Probability Distribution

$P(X_1, X_2, \ldots, X_n)$

- Conceptually, the simplest way to represent relationship among random variables.
- Facilitates all types of inferences between variables.
- **Big drawback**
  - Complexity exponential in n.
- A major reason for probability theory not playing an significant role in AI before 1980's.
- Alternatives explored instead:
  - Non-monotonic logic,
  - Uncertainty factors,
  - Belief function,
  - Fuzzy logic,
  - ...etc

Modularity

- The breakthrough came in early 1980s (Pearl 1986, 1988, Howard & Matheson 1984)
- In a joint probability distribution, every variable is, in theory, directly related to all other variables.
- Pearl and others realized:
    - It is often reasonable to make the assumption that each variable is directly related to only a few other variables.
    - This leads to **modularity**: Allowing decomposing a complex model into small manageable pieces.
    - Giving rise to **Bayesian networks**

Bayesian Networks

- Networks of random variables
  - Nodes represent random variables.
  - Arrows (links) represent dependence.
- Result of a marriage between probability theory and graph theory.
- Conditional independence implicitly represented:
  - Absence of links implies indpendence.
  - A random variable is directly related to only a few neighboring variables.
  - It is independent of all other variables given the neighboring variables.
- Also implicitly represents factorization of joint distribution.
- Facilitate the application of probability theory to many problems in AI, Applied Mathematics, Statistics, and Engineering that **are complex** and **involve uncertainty**.

Bayesian networks and other probabilistic models

Common framework for existing models

- Naive Bayes model, latent class model,mixture models,hidden Markov models, Markov chains, etc
- Facilitate easy share of progresses.

Provides a framework for new models:

- Dynamic Bayesian networks, Latent tree models, conditional Markov random field, etc

Use of Bayesian networks

- Proposed as a framework to build expert systems.
- Increasingly used a tool for data analysis

- Intuitive and easy to understand
    - Widely used as a communication tool before researchers.

Course Objectives

- Provide a solid training in the basic theory and methods of BNs.
  - Achieved through lectures by the instructor. 75% of class time.
- Convey an overview of the field.
  - Achieved through presentations by students. 25% of class time.