Course Overview and Introduction

40-957 Special Topics in Artificial Intelligence: Probabilistic Graphical Models
Sharif University of Technology

Soleymani
Spring 2015
Course info

- Instructor: Mahdieh Soleymani
  - Email: soleymani@sharif.edu

- Website: [http://ce.sharif.edu/courses/93-94/2/ce956-1](http://ce.sharif.edu/courses/93-94/2/ce956-1)

- Lectures: Sun-Tue (15-16:30), Room 402
- Office Hours: Sat-Monday 10:00-11:00

- Teacher assistants:
  - Hassan Hafez
  - Ahmad Khajeh nekhad
  - Mostafa Tavassolipour
Text book


- Some papers
Evaluation policy

- Mid-term: 25%
- Final: 35%
- Home works & course works: 40%
PGMs as a framework

- General-purpose framework for representing uncertain knowledge and learning and inference in uncertain conditions.
  - Exploiting structure in complex distributions to describe them compactly

- A graph-based representation as the basis of encoding a complex distribution
  - denotes conditional dependence structure between random vars
    - One view: Graph represents a set of independencies
    - Another view: Graph shows a skeleton for factorizing a joint distribution
  - allows declarative representation (with clear semantics) of the probabilistic knowledge
PGMs as a framework

- Intuitive & compact data structure for representation
- Efficient reasoning using general-purpose algorithms
- Sparse parameterization (enables us to elicit or learn from data)
History

- Wright 1921, 1934 and before

- Bayesian networks are independently developed by Spiegelhalter and Lauritzen in statistics and Pearl in computer science in the late 1980’s

- First applications (1990’s): expert systems and information retrieval
Reasoning under uncertainty

- Partial knowledge of the state of the world
  - Noisy or incomplete observations
  - We may not know or cover all the involved phenomena in our model
  - Partial knowledge can cause the world seems to be stochastic
  - Innate non-determinism

- To deal with partial knowledge and/or stochastic worlds we need reasoning under uncertainty
PGM: declarative representation

- Separation of knowledge and reasoning
- We need to specify our model for a specific application that represents our probabilistic knowledge
- There is a general suite of reasoning algorithms that can be used.
PGMs: some application areas

- Machine Learning and computational statistics
- Computer vision: e.g., segmenting and denoising images
- Robotics: e.g., robot localization and mapping
- Natural Language Processing
- Speech recognition
- Information Retrieval
- AI: game playing, planning
- Computational Biology
- Networks: decoding messages (sent over a noisy channel)
- Medical diagnosis and prognosis
- …
Graphical models: directed & undirected

- Two kinds of graphical models:
  - Directed: Bayesian Networks (BNs)
  - Undirected: Markov Random Fields (MRFs)

Causality relations

Correlation of variables
Graphical models: directed & undirected

[Pathfinder Project, 1992]
We will cover three aspects of the graphical models:

- Representation of probabilistic knowledge
- Inference algorithms on these models
- Using the data to acquire the distribution
Representation, inference, and learning

- **Representation**: When variables tends to interact directly with few other variables (local structure)

- **Inference**: answering queries using the model
  - algorithms for answering questions/queries according to the model and/or based given observation.

- **Learning**: of both the parameters and the structure of the graphical models
Representation

- Representing large multivariate distributions directly and exhaustively is hopeless:
  - The number of parameters is exponential in the number of random variables
  - Inference can be exponential in the number of variables

- PGM representation
  - Compact representation of the joint distribution
  - Transparent
    - We can combine expert knowledge and accumulated data to learn the model
  - Effective for inference and learning
Medical diagnosis example

- **Representation**

  ![Diagram showing the relationships between symptoms and diseases]

- **Inference:** Given symptoms, what disease is likely?
- **Eliciting or learning the required probabilities from the data**
Image denoising example
Plan in our course

- **Fundamentals of Graphical Models:**
  - Representation
    - Bayesian Network
    - Markov Random Fields
  - Exact inference
  - Basics of learning

- **Case studies: Popular graphical models**
  - Multivariate Gaussian Models
  - Temporal models

- **Approximate inference**
  - Variational methods
  - Monte Carlo algorithms

- **Advanced topics and latest developments**
Classical probabilistic methods

- Machine learning tasks with probabilistic approach:
  - Density estimation
  - Classification, regression
  - Clustering
  - Dimensionality reduction

- Many of the classical multivariate probabilistic methods are special cases of GMs

- Graphical model as a way to view all of them in a common formalism
Probability review

- Marginal probabilities
  \[ P(X) = \sum_y P(X, Y = y) \]

- Conditional probabilities
  \[ P(X|Y) = \frac{P(X, Y)}{P(Y)} \]

- Bayes rule:
  \[ P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)} \]

- Chain rule:
  \[ P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i|X_1, \ldots, X_{i-1}) \]