CS 583: Probabilistic Graphical Models

Topic: Summary

Mustafa Bilgic

http://www.cs.iit.edu/~mbilgic

https://twitter.com/bilgicm

Covered

- Chapter 1 – Introduction
- Chapter 2 – Foundations
- Representation
  - Chapter 3 – Bayesian network representation
  - Chapter 4 – Markov network representation
  - Chapter 5 – Local probabilistic models
  - Chapter 6 – Template-based models
  - Chapter 7 – Gaussian network models

Covered

- Inference
  - Chapter 9 – Variable elimination
  - Chapter 10 – Clique trees
  - Chapter 12 – Sampling
  - Chapter 13 – MAP inference

Covered

- Learning
  - Chapter 16 – Overview
  - Parameter estimation (Chapters 17, 19, 20)
    - Parameter estimation for Bayesian networks
    - Parameter estimation with incomplete data
    - Parameter estimation for Markov networks
  - Structure learning (Chapters 18 and 20)
    - Bayesian network structure learning
    - Markov network structure learning

Other Notable Topics

- Chapter 11 – Inference as optimization
- Chapter 15 – Inference in temporal models
- Chapter 21 – Causality
- Chapter 22 – Utilities and decisions
- Chapter 24 – Structured decision problems

Chapter 11 – Inference as optimization

- Approximate inference
- The underlying distribution is P
- Create a similar but "easier" distribution Q
- Answer queries using Q
- Notable approach: variational inference

Chapter 15 – Inference in temporal models

- Temporal models are template-based models
- One approach: "unroll" the network
  - Problem: can be fairly large, infinite even
- Specific structure and queries
- For example: tracking/filtering
  - $P\left(X^t \middle| o^{(1:t-1)}\right)$ i.e, our belief about current state given everything we have observed so far
- Notable algorithm: Viterbi

Chapter 21 – Causality

- We need to distinguish between correlation and causation
  - If X and Y are correlated, it could be
    - X causes Y
    - Y causes X
    - A hidden cause Z causes both X and Y
- Useful for interventions
- Useful for interpretability
- Notable approach: do-calculus
  - Differentiate between $P(Y \mid Z=z)$ versus $P(Y \mid do(Z=z))$

Chapter 22 – Utilities and decisions

- World states have
  - Uncertainty – we model this using probabilities
  - Utility – how much happy/unhappy we are in each of those world states
- We need to take some actions to maximize our expected utility
- For example:
  - A two-variable world: rain/not-rain, and have-umbrella/don't-have-umbrella
  - Utility: how much happy we are in one of those four world states
  - Action: take/don't-take umbrella

Chapter 24 – Structured decision problems

- Influence diagrams – Bayesian networks that represent
  - Uncertain nodes,
  - Action nodes, and
  - Utility nodes
- Inference: find the action that maximizes expected utility

Where do we go from here?

- In addition to what we discussed, learn at least
  - Inference and learning for template-based models
  - Causality
  - Variational inference
- Read and implement papers. For example:
  - Collective classification
    - http://www.cs.iit.edu/~ml/pdfs/sen-aimag08.pdf
  - Latent Dirichlet allocation
    - http://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf