CMPSCI 691T: Hierarchical Probabilistic Models for AI

Sridhar Mahadevan
mahadeva@cs.umass.edu

University of Massachusetts
Abstract

This course will be offered in Spring 2002. It meets every M/W from 9:05-10:20 in Room 140. The following slides describe the the course material in more depth.
Course Motivation

- 1990-current: probabilistic computing become a cornerstone of many areas of artificial intelligence, including machine learning, natural language, planning, robotics, and vision.

- Statistical process models (e.g., HMMs, MDPs) have grown enormously in their importance.

- Many real-world applications: information retrieval, networking, perception, robotics, vision, smart rooms, and wearable computing.
Course Motivation

Most applications have used "flat" models which scale badly to large problems.

This course will focus on recent advances on multiscale decomposable models which can lead to larger-scale applications.

We will cover theory, algorithms for learning (estimation) and control, and the main application areas of multi-scale models.
Prerequisites

One of the following courses:
- CMPSCI 683 (AI)
- CMPSCI 687 (RL)
- CMPSCI 689 (ML)

Also, you should have
- Undergraduate knowledge of probability and statistics
- Domain expertise in one of the major application areas
Recommended Reading

- **Statistical Language Learning** by Eugene Charniak, MIT Press, 1993
- **Statistical Methods for Speech Recognition** by Frederick Jelinek, MIT Press, 1997
- **Applied Probability Models with Optimization Applications** by Sheldon Ross, Dover, 1970
- **Markov Decision Processes** by Martin Puterman, Wiley, 1994
- **The Elements of Statistical Learning** by Hastie, Tibshirani, and Friedman, Springer, 2001
Course Grading

Student presentations (of project and selected papers) will be an integral part of the course. Depending on the number of students enrolled, the projects can be group-based. There will be NO exams or formal homework. Students can use the Bayes Network Toolbox in MATLAB for the projects.

<table>
<thead>
<tr>
<th>Section</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Assignments</td>
<td>20%</td>
</tr>
<tr>
<td>Class Presentations</td>
<td>20%</td>
</tr>
<tr>
<td>(Group) Project</td>
<td>60%</td>
</tr>
</tbody>
</table>
Course Topics

- Model dimensions: observability, controllability, and modularity
- Motivation: the need for multiscale representations
- Application Domains: decision-theoretic planning, face and gesture recognition, information retrieval, optimization, robotics, routing, smart rooms, wearable devices
- Overview of parameter estimation techniques: EM, maximum likelihood etc.
Hidden Markov Models

- Algorithms for learning HMMs (Baum-Welch, cross-entropy)
- Building classifiers using HMMs
- Input-output HMMs, observable operator models, and other variants
Hierarchical Hidden Markov Models

- Learning and using HHMMs.
- Model reuse and sharing
- Scaling issues
Semi-Markov Decision Processes

- Sequential decision-making over behaviors
- Hierarchical reinforcement learning
- Options, HAMs, and MAXQ formalisms
Partially Observable MDPs

- Memory structures: belief states, histories, tests
- Variable memory models, incl. suffix automata
- Predictive state representation models
- Learning POMDPs and suffix trees
- Hierarchical POMDPs
Compact Representations

- Dynamic Bayes nets
- Factorial HMMs
- Factored Markov decision processes
- Probabilistic inference (e.g. dynamic programming) over compact models