Course Overview

Mark Peot

1.0 Why take this course?

Graphical models are models for joint probability distributions.

Graphical models relate concepts of graphical independence (separation) to probabilistic independence. This independence implies that the underlying distribution factors, which is often a GOOD THING. If a joint distribution can be factored, then it *may* be possible to represent a distribution with less parameters.

We will see that if a distribution can factor, then less parameters are needed to capture the joint distribution and the distributive law of multiplication might be able to be used to simplify the computation of marginal and conditional distributions, often reducing the time complexity for computig a marginal from exponential in the number of variables to a low-order polynomial in the number of variables.

Some of you are taking this course to learn how to build probabilistic expert systems. In a probabilistic expert system, a joint distribution is used to model the relationship between uncertain, but important, variables (for example, boolean variables representing unknown diseases) and variable representing the outcome of tests or observations. The goal of the expert system is to compute the probability distribution over the unknown variables given values for some set of variables that we can observe. If a joint distribution can factor, then

- less parameters are needed to capture the joint distribution, reducing knowledge acquisition time, and
- inference algorithms can reduce the time required for inference from exponential in the number of variables to a low order polynomial, allowing you to develop faster expert systems with larger knowledge bases.

For you, this course will show how to build probabilistic expert systems from experts or from data, and will show you how to construct fast inference algorithms.

Some of you are taking this course in order to construct models for data. If we use a factored probability distribution, we need to use fewer parameters to represent a joint distribution. This reduces the amount of data required to learn a model and prevents overfitting. In this course, we will explore a number of issues related to model construction including parameter learning, structure learning, discrete and continuous models for data, and graph-based algorithms for clustering. We will also explore the relationship between graphical models and algorithms for time series data; demonstrating, for example, that algorithms such as the Kalman filter algorithm are really a special case of more general belief network algorithms.

2.0 Mechanics

2.1 Grading

The grade for the course will based on weekly homework (50%) and a course project (50%). Assigned homework will be due in class on the due date. Rules for the homework: work alone and cite any reference material that you use outside of the assigned texts.

Homework will be assigned only for the first half of the course, after which, a project will take precedence. Possibilities for projects include: probabilistic expert systems, algorithm development, theory development, decision support systems, learning software, extensions to theory, data analysis or literature surveys on a topic that interests you. The project culminates with a final report (8-15 pages) and

Homework:

HW 1. Assigned 1/15, Due 1/25.

HW 2. Assigned 1/25, Due 2/1.

HW 3. Assigned 2/1, Due 2/8.

HW 4. Assigned 2/8, Due 2/15.

HW 5. Assigned 2/15, Due 2/22.

HW 6. Assigned 2/22, Due 3/1.

Project:

Project proposal: 3/1.

Progress report I: 3/26.

Progress report II: 4/9.

In-class presentation: 4/19 - 4/23.

Final report: 4/23.

2.2 URL

The course web pages will be at http://www.stat.duke.edu/htdocs/courses/Spring99/sta294 for the first week or so of the semester. After this point, the page will move to http://www.stat.duke.edu/courses/Spring99/sta294. The web pages will contain links to all of the handouts; including lecture notes, homework and solutions.

2.3 Reading

Assigned Texts:
Castillo, Guiterrez, and Hadi; *Expert Systems and Probabilistic Network Models* [CGH]
Course Reader

Recommended Text:
Steffen Lauritzen; *Graphical Models*, [Lauritzen]

2.4 Contact Information

Phone: 660.6558

Email: peot@stat.duke.edu

Office: Levine Science Research Center, D105.

3.0 Calender

Week 1 (1/13 - 1/15): Introduction to graphs

Week 2: (1/20 - 1/22): No class 1/18 due to MLK. Introduction to independence, semi-graphoids, and graphoids. Intro to Markov properties of undirected graphs.

Week 3: (1/25 + 1/29): No class 1/27 due to a business trip. Markov properties of undirected, directed, and chain graphs.

Week 4: (2/1 - 2/5): Markov properties continued. Graph decomposition, clique trees, join trees, and triangulation.

Week 5: (2/8 - 2/12): The bucket elimination and join tree algorithms. Decomposition. The polytree and cutset conditioning algorithms for exact inference with discrete distributions. (2/8: STA 395: *Learning from what you don't observe.*)

Week 6: (2/15 - 2/19): Introduction to complexity and complexity results for exact algorithms. Probabilistic expert systems. Causal independence.

Week 7: (2/22 - 2/26): Model construction techniques. Case studies: Pathfinder, Microsoft Answer Wizard, and Microsoft Office Assistant.

Week 8: (3/1 - 3/5): The likelihood weighting and bounded variance and *AA* algorithms for approximate inference. Complexity results for approximate inference.

Week 9: (3/8 - 3/12): Normal distributions and graphical models.

Week 10: (3/15 - 3/19): Skiing.

Week 11: (3/22 - 3/26): Learning I: Algorithms for learning joint distributions and structure from complete data.

Week 12: (3/29 - 4/2): Learning II: EM, incomplete data and Autoclass.

Week 13: (4/5 - 4/9): Extensions for mixed distributions (gaussian and discrete). Strong decomposition.

Week 14: (4/12 - 4/16): Time series, approximation of time series.

Week 15: (4/19 - 4/23): Project presentations.