Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science

6.438 ALGORITHMS FOR INFERENCE

Information Sheet
Fall 2012

Lecturer
Tommi Jaakkola (tommi@mit.edu), 32-G470

Teaching Assistants
Sue Zheng (zhengs@mit.edu)
Joong Bum Rhim (jbrhim@mit.edu)

Administrative Assistant
Teresa Cataldo (cataldo@csail.mit.edu), 32-G475B

Lectures: Tuesdays and Thursdays, 9:30am–11am, Room 32-124

Recitations: Fridays, 10-11am or 11am-noon, Room 36-372

Office hours: Mondays, 1-2:30pm, 32-G4th floor lounge (Joong)
Tuesdays, 11:30-12:30, 32-G470 (Tommi)
Wednesdays, 3-4:30pm, 32-G4th floor lounge (Sue)

Welcome to 6.438 !

This is a graduate-level introduction to the principles of statistical inference with probabilistic models defined using graphical representations. As such, it is designed as a core graduate subject for students in the relevant subfields of both Area I and Area II. The material in this course constitutes a common foundation for work in machine learning, signal processing, artificial intelligence, computer vision, control, and communication. Moreover, it is the companion course to 6.437 (Inference and Information) which is offered each Spring; 6.437 and 6.438 may be taken in either order.

It is worth stressing that 6.438 is an *introductory* graduate subject: it is not an advanced graduate subject for students who have already have a mastery of statistical inference algorithms, yet want to understand such material at an even more sophisticated level. That said, the structure of this subject will be somewhat different than other such introductions to the topic.

Ultimately, the subject is about teaching you contemporary approaches to—and perspectives on—problems of statistical inference. The development of the material that forms the basis for this subject has historically been very much driven by applications. However, our focus in the course will not be on these applications—which form the basis for entire courses of their own—but rather on the common problem solving frameworks that they share. Nevertheless, we will cite various relevant applications as we develop the material and sometimes extract simplified examples from these contexts.

6.438 continues to mature, but is still a relatively new one in our curriculum, and as such we continue to experiment with content and pedagogy. This term will be no exception. While such evolution is an important part of course development, it also means that the subject will continue to be rougher and more raw than one with a longer history. As a result, this offering of 6.438 will appeal most to students who, as part of taking the subject, relish the role of beta-tester and are eager to contribute to the subject's evolution through their participation. Moreover, because the content is evolving, taking 6.438 this Fall will also require certain maturity in working with a variety of reading materials, not all perfectly matched to the subject as currently taught.

Prerequisites

The official prerequisites are 6.041 or 6.436, and 18.06, or their equivalents. Ultimately, what we require is *fluency* with both basic quantitative probabilistic analysis and linear algebra, together with some subsequent solid exposure to the engineering application of both. When in doubt, students whose undergraduate degrees are not from MIT should consult the staff to determine if they have had subjects that are effectively equivalent to the official prerequisites.

Reading

There is no existing text that perfectly matches the current content of 6.438 and the style in which we teach it. However, the notes

Michael I. Jordan, *Introduction to Probabilistic Graphical Models,*

chapters of which are posted on the 6.438 web site, have proven to be a popular complement to many of the lectures, containing useful additional details, perspectives, and insights. *Please understand that these notes have been made available for your personal use in this subject.* In getting permission to use these notes in 6.438, we have been asked to *not share or otherwise distribute any portions of these notes* to others outside the subject *in any form*, so please respect this request.

In addition, to fill in gaps or add additional insights, we will occasionally produce supplementary notes during the term. Any such notes will necessarily be spare, rough, and contain bugs, but will hopefully be a useful resource to you.

Finally, you will also find sections of the following books to be useful and more in-depth auxiliary references for parts of the term.

C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques,* MIT Press, 2009.

D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms,* Cambridge University Press, UK, 2003. (also available on-line at http://www.inference.phy.cam.ac.uk/mackay/itila/book.html)

We suggest you initially hold off purchasing any books while you are gauging their usefulness to you.

If you are interested in further reading, either to strengthen your background, reinforce some of the concepts from lecture, or to probe some topics in more detail, take a look at the additional references on the course web site. In particular, you'll find several papers containing a variety of useful insights, which are worth the effort to work through.

Problem Sets

There will be approximately 10 problem sets. Problem sets will be due in lecture. Problem sets must be handed in at the end of the class in which they are due. Problem set solutions will be available soon after the end of the due date's lecture.

While you should do all the assigned problems, only a small subset (chosen randomly or otherwise) will actually be graded. Don't be misled by the relatively few points assigned to homework grades in the final grade calculation. While the grade you get on your homework is only a minor component of your final grade, working through (and, yes, often struggling with at length!) the homework is a crucial part of the learning process and will invariably have a major impact on your understanding of the material and thus your final grade. Some of the problem sets will involve a Matlab component, to help you explore different aspects of the material. Also, problems labeled "practice" are additional problems for you to work through as time permits and if you think they would be helpful to you. They are never graded, but solutions will be provided.

In undertaking the problem sets, moderate collaboration in the form of joint problem solving with one or two classmates is permitted provided your writeup is your own.

Exams

There will be two exams in the subject, scheduled in the evening, most likely in the same room as lectures (32-124). The locations will be announced later. The midterm will be on Tuesday, October 23, and the final exam on December 11. Both exams will be designed to require about 1.5 hours of effort, but we'll use the three hour format to minimize the effects of time pressure. The exams will both be *closed book*. You will be allowed to bring *two* 8.5×11-inch sheet of notes (both sides) to the

Midterm, and *four* 8.5×11-inch sheets of notes to the Final.

Course Grade

The grade in the course is based upon our best assessment of your understanding of the material during the semester. Roughly, the weights used in grade assignment will be:

Midterm	30%
Final	40%
Homework	15% (completion) + 15% (grading)

The additional benefit from completing all the problem sets is that if you do better on the Final than the Midterm, then the Midterm will not count towards the grade. Instead, the Final exam will determine 70% of the grade. Note that you have this benefit only if you have done all the problem sets.

In addition, as always, other factors such as contributions to the lecture discussion and other interactions can make a difference in the final grade.

Course Web Site and Email

We will make announcements via email, and we will post various information and handouts on the course web site.

You should first make sure that you have an active Athena account (by visiting `http://web.mit.edu/accounts/` if necessary) as well as a personal certificate (by visiting `http://ist.mit.edu/services/certificates` if necessary). If you have problems or if you are not a regular MIT student, please contact any of the TAs for assistance.

The course web site is

http://web.mit.edu/6.438

You will need to have a valid certificate *and* be on the official course list to access the web site. If you have pre-registered for 6.438, this should already be set up; just double-check that you can access the web site (try to download a handout, for example).

If you have any questions during the term, you can reach us by sending email to

6.438-staff@mit.edu

Tentative Syllabus and Schedule

A tentative list of topics and a tentative schedule is as follows. We stress that this is preliminary and subject to change as the term evolves.

Homework				Lectures	
Out	**Due**	**Day**	**Date**	**#**	**Lecture Material**
		R	9/6	1	Introduction, overview, preliminaries
1		T	9/11	2	Directed probabilistic graphical models
		R	9/13	3	Undirected graphs
2	1	T	9/18	4	Factor graphs; generating and converting graphs
		R	9/20	5	Perfect maps, chordal graphs, Markov chains, trees
3	2	T	9/25	6	Gaussian graphical models
		R	9/27	7	Inference on graphs: elimination algorithm
4	3	T	10/2	8	Inference on trees: sum-product algorithm
		R	10/4	9	Example: forward-backward algorithm
		F	10/5		ADD DATE
		T	10/9		no class
		R	10/11	10	Sum-product algorithm with factor graphs
5	4	T	10/16	11	MAP estimation and min-sum algorithm
		R	10/18	12	Inference with Gaussian graphical models
		T	10/23	Q	Midterm Exam (through L11 and PS4)
		R	10/25	13	Example: Kalman filtering and smoothing
6	5	T	10/30	14	Junction tree algorithm
		R	11/1	15	Loopy belief propagation
7	6	T	11/6	16	Variational methods
		R	11/8	17	Approximate MAP by graph partitioning
8	7	T	11/13	18	Linear programming relaxations
		R	11/15	19	Sampling by Markov chain Monte Carlo
9	8	T	11/20	20	Parameter estimation in directed graphs
		W	11/21		DROP DATE
		R	11/22		No Class (Thanksgiving Vacation)
		T	11/27	21	Learning structure in directed graphs
		R	11/29	22	Modeling from partial observations
10	9	T	12/4	23	Learning undirected graphs
		R	12/6	24	Learning exponential family models
		F	12/7		LAST ASSIGNMENT DUE DATE
		M	12/11	Q	Final Exam (through L24 and PS9)