ORI 390R.16 - Markov Decision Processes

- **Time & Place:** Tue & Thurs 9:30-11am, ETC 5.132
- **Professor:** John J. Hasenbein
  - **Office:** ETC 5.128B
  - **Phone:** 471-3079
  - **Email:** *jhas@mail.utexas.edu* (This is the best way to contact me.)
  - **Office Hours:** Mondays, 10:30am-noon. You can also email me for an appointment.
- **Class Web Page:** All class materials will be posted on *Canvas*.
- **Required Text:** *Dynamic Programming and Optimal Control, Volume II* by Dimitri P. Bertsekas (4th Edition, Athena Scientific 2012). We will use one chapter from Volume I of this set, but that chapter is posted on Canvas.
- **Grading:** Problem sets will be assigned approximately every two weeks. There will be one mid-term exam and one final exam. Each exam will be worth 35% of your grade. Your homework average will comprise the other 30% of your grade.

  For the problem sets, you may discuss problems with your classmates and in fact are encouraged to do so. However, you should understand the solutions submitted by your group. A good rule of thumb is that you should be able to explain to me the solutions you have submitted.
- **Exams:** You are required to take the exams at the scheduled time. Make-up exams will not be given without a valid medical excuse. The first exam will be given in class on Tuesday, October 18th. The final exam will take place at the university scheduled date and time: Thursday, December 8th, 2-5pm.
- **Prerequisites:** Students should have knowledge of discrete-time and continuous-time Markov chains, and linear programming. Preferably, students should have taken ORI 390R.5 and ORI 391Q.5, or equivalent courses.
- **Email Communication**: For this class, email will be used as an official form of communication for notifying you of new homework assignments and other class updates. The University of Texas email policy can be found at `http://www.utexas.edu/cio/policies/university-electronic-mail-student-notification-policy`.
- **Students with disabilities:** The University of Texas at Austin provides upon request appropriate academic accommodations for qualified students with disabilities. For more information, contact the Office of the Dean of Students at 471-6259, 471-4641 TTY.
- **Course Evaluation:** Near the end of the course you will have an opportunity to anonymously evaluate the course and instructor using the standard College of Engineering evaluation form.

- **Class Web Site and Privacy:** For this class, web-based, password-protected class sites will be available via the *Canvas* system. The syllabus, handouts, assignments and other resources are types of information that may be available within this site. Site activities could include exchanging e-mail, engaging in class discussions and chats, and exchanging files. In addition, a class e-mail roster will be a component of the site. Students who do not want their names included in this electronic class roster must restrict their directory information in the Office of the Registrar, Main Building, Room 1. For information on restricting directory information see:
`http://registrar.utexas.edu/students/records/restrictmyinfo`.

Additional References

- *Dynamic Programming and Optimal Control, Volume I*, by Dimitri P. Bertsekas (Athena Scientific 2007).
- *Markov Decision Processes*, by Martin L. Puterman (Wiley 1994).
- *Introduction to Stochastic Dynamic Programming*, by Sheldon M. Ross (Academic Press 1983).
- *Stochastic Dynamic Programming and the Control of Queueing Systems*, by Linn I. Sennott (Wiley 1999).
- *Constrained Markov Decision Processes*, by Eitan Altman (Chapman & Hall 1999).
- *Introduction to Probability Models*, by Sheldon M. Ross (any edition, Academic Press).

Course Summary

The goal of this course is to introduce Markov decision processes (MDPs). MDPs are sometimes also referred to as controlled Markov chains and the underlying theory is called stochastic dynamic programming. MDPs have applications in many areas including revenue management (e.g., hotel, airline, and rental car pricing), control of queues, financial engineering, telecommunications, manufacturing, and economics.

The course will examine the models, theory, and applications of MDPs. MDPs may have countable or uncountable state and action spaces, and may be framed in discrete or continuous time. We focus mainly on countable state and action spaces, and the discrete-time setting. However, we also introduce the idea of uniformization to convert continuous-time problems into the discrete-time setting. In going through the topics below, I will introduce various examples in different application areas.

Course Outline

I. Introduction to MDPs

- MDP model basic components
- Applications and examples

II. Finite Horizon Models

- Elements of MDPs
- Dynamic programming algorithm
- The principle of optimality
- Examples: Stock options and the Secretary Problem
- Revenue management modeling

III. Infinite Horizon Discounted Problems

- Banach spaces and contraction mappings
- Convergence of the DP algorithm and Bellman's equation
- Value iteration and variants
- Policy iteration
- Linear programming methods

IV. Approximate Dynamic Programming

- The curses of dimensionality
- ADP via linear programming
- Lookahead and rollout methods
- Q-Learning

V. Infinite Horizon Average Cost Problems

- Counterexamples and thought experiments
- Finite space models
- Bellman's equations
- Blackwell optimality
- Multichain and unichain conditions
- Value iteration and policy iteration
- Linear programming solutions
- Infinite state space problems

VI. Continuous-time MDPs

- Uniformization
- Queueing Applications

VII. Partially Observable MDPs (time permitting)