EEL 6935 **_Stochastic Control_** SPRING 2025

Control of systems subject to noise and uncertainty

Welcome to Control Systems & Reinforcement Learning

The course provides a range of approaches to modeling and control, with applications ranging from finance to power systems to medicine.

Markovian[1] models are given special attention because of their flexibility in modeling complex phenomena, and because this is the required setting in much of optimal control theory. The course will provide several approaches to the design of control laws, including dynamic programming theory and techniques to obtain control solutions with limited real time observations. Techniques from Lyapunov theory are introduced for stability and performance evaluation, and will provide some foundations for the construction and analysis of numerical methods for constructing approximately optimal control solutions.

In parallel with this theory is the development of model free methods for control design. Theory and practice of stochastic approximation will be presented as one essential ingredient in reinforcement learning (RL). All of the standard RL approaches will be surveyed, including SARSA, Q-learning, and actor-critic methods.

It is intended for graduate students who have some background in control and stochastic processes. Experience with *Matlab* or *Python* is also essential.

Why do we need noisy models? When you introduce the word "stochastic" to control, this just means that you are bringing in a larger range of tools for understanding how to control systems, and evaluate their performance. Name a tool from probability, and you have something useful for control synthesis. In particular, there is the question of information. This may mean the data available for control, or information about the system to be controlled. There may be variables of interest that are not directly observed, so we will want to estimate. Tools to be applied include nonlinear filtering and stochastic approximation (a foundation of reinforcement learning).

Time and Place: Tues 3:00-4:55 *&* Thur 4:05-4:55, Larsen 239

Office hours: Wednesdays in person 5104 Mala, 10:00-11:00am, and Mondays on Zoom, 1:00-2:00pm (*extended on request*). Graduate students Austin Cooper and Caio Lauand will at times provide time for additional office hours. We can be reached for questions by electronic mail via Canvas.

We cannot promise to respond to messages directed to @ece or @gmail.

Exams & homework: Homework problems will be assigned on a ∼bi-weekly basis (more frequent at the start), to be uploaded to canvas before lecture on the date due. They will be graded and returned the following week. *Late homework cannot be accepted.*

There will be two evening midterm exams in late February and late April (details to be announced soon), from 7:20 - 8:50 p.m. You will be allowed *one* sheet of notes ($8\frac{1}{2} \times 11$; both sides) in the first exam, and *two* in the second. Otherwise, the exams are closed-book and closed-notes.

Grading scheme: Homework[2] will count 10%, the midterm exams 70%, and the final project will count 20% towards the final grade in the course.

[1] A Markov chain is nothing more than a nonlinear state space model subject to noise.

[2] We encourage collaboration on homework!

Course Outline

I. Basics and previews   CS&RL Ch. 3 and 7

1) Controlled Markov models and MDPs.
2) Examples: both deterministic and stochastic
3) Temporal difference origins   [TD = Temporal difference]

II. Performance evaluation   Ch. 6

1) Probability review
2) Markov models and more examples
3) Lyapunov techniques for stability and performance
4) Monte-Carlo methods   **First Midterm Exam**

III. Numerical methods   Appendix B and handouts

1) Approximate dynamic programming.
2) Numerical techniques: policy and value iteration; LP formulation.
3) Partial information (belief state), with applications to
   Linear quadratic optimal control; Multi-armed bandits (with UCB heuristic).

IV. Adaptation and Learning   Ch. 8–10

1) Stochastic approximation: algorithms for learning
2) TD Learning
3) Q Learning   **Second Midterm Exam**
4) Actor-critic methods (time permitting)

   Review and thoughts for the future

References: The following are available free on-line (send your thanks to CUP[3]):

⊙ S. P. Meyn, *Control Systems and Reinforcement Learning* (CS&RL), 2022.
https://meyn.ece.ufl.edu/control-systems-and-reinforcement-learning

⊙ S. P. Meyn, *Control Techniques for Complex networks* (see last chapter).
https://meyn.ece.ufl.edu/publications/

⊙ S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability* (too advanced for this course).
https://meyn.ece.ufl.edu/publications/

The following are valuable background (send your thanks to Profs. Hajek and van Handel):

⊙ B. Hajek, *Exploration of Random Processes for Engineers*.
https://www.ifp.illinois.edu/~hajek/Papers/randomprocesses.html Review: $(\Omega, \mathcal{F}, \mathsf{P}) \star \mathsf{P}(A) \star \mathsf{E}[X \mid Y]$

⊙ R. Van Handel, *Lecture Notes on Hidden Markov Models*.
https://web.math.princeton.edu/~rvan/orf557/ (also in Files folder, `hmm080728.pdf`)

The following textbooks are of value, but not needed to follow the course:

⊙ D. Bertsekas and J. Tsitsiklis, *Neuro-Dynamic Programming*, 1996.

⊙ D. Bertsekas and S. Shreve, *Stochastic Optimal Control: The Discrete-Time Case*, 1978.
https://web.mit.edu/dimitrib/www/soc.html   Bertsekas has many other great resources

⊙ R. Sutton and A. Barto. *Reinforcement Learning: An Introduction.* MIT Press, 2018.   Available online

[3] Cambridge University Press