RUTGERS
Rutgers Business School
Newark and New Brunswick

REINFORCEMENT LEARNING 26:711:685
Spring 2020
Wednesday 3:00 PM - 5:50 PM

Michael N. Katehakis
Home page
100RR, Rm. 3031
Office Hours: after class, and by appointment

COURSE DESCRIPTION

This course covers topics in data driven sequential decision making. It first introduces the models and algorithms for multi-armed bandits and contextual bandits. Then it provides an advanced survey of basic theory for Markovian Decision Processes (MDPs), a basic model for the problem of optimal sequential decision making under uncertainty. The remainder of the course is devoted to Reinforcement Learning (RL) a paradigm for learning the parameters of an MDP from data, small state space MDPs and extensions for large state spaces.

The course aims to introduce the fundamental models, the main algorithms, performance guarantees, lower bounds, and applications in the industry.

COURSE MATERIALS

Recommended Texts:

- Dimitri P. Bertsekas Reinforcement Learning And Optimal Control, Athena Scientific, July 2019.
- Csaba Czepesvári, Algorithms for Reinforcement Learning
- Sheldon M. Ross: Introduction to Stochastic Dynamic Programming $\geq$ 7-th Edition, Elsevier

Other resources:

◇ *Other Resources:* I will provide additional at: reading materials during the course.
◇ Start with: Review of Probability.

PREREQUISITES

a) Proficiency in Python. Class assignments will be in Python (using numpy and Tensorflow).

b) College Calculus, Linear Algebra

c) Basic Probability and Statistics

LEARNING GOALS AND OBJECTIVES

The objective is to introduce you to the *models and techniques* of Reinforcement Learning. This will necessitate

some use of theory and computer manipulations of data. However, the goal is not that you will become experts on computer software, but rather to give you a perspective on how Reinforcement Learning models and techniques are used in practice. Hence, much of the material will be presented in a setting of practical situations. The focus is on ways of modeling and structural ideas about systems that evolve randomly in time.

Specific goals: There will be an introduction to the concepts and applications of those basic ideas that are considered to be most important for the practical analysis of Reinforcement Learning, listed below.

COURSE SCHEDULE[1]

Tentative Course Outline.

- ◇ Weeks 1-3: Basic concepts
  - Conditional Probabilities and Conditional Expectations.
  - Large Deviations.
  - Concentration Inequalities.
- ◇ Weeks 4-6: Mutli-Armed Bandits
  - PS (Thompson) Sampling.
  - Robbins early work.
  - Gittins indices.
  - UCB algorithms vs ϵ greedy and related algorithms.
  - Contextual bandits.
- ◇ Weeks 7-13: Reinforcement Learning.
  - Basic Markovian Decision Processes Models.
  - Optimality Criteria.
  - Value functions and Solution Methods.
  - Value iteration
  - Policy iteration
  - Linear programming
  - Asynchronous value iteration
  - Basic RL Models
  - The pitfalls of 'Certainty equivalence'.
  - Burnetas and Katehakis UCBs.
  - Optimistic Programming.
  - PS (Thompson) Sampling.
  - Q-learning
  - Deep reinforcement learning

[1] Tentative.

ACADEMIC INTEGRITY

Students are responsible for understanding the RU Academic Integrity Policy. Students must sign the RU Honor Pledge. See business.rutgers.edu/ai for more details.

ATTENDANCE AND PREPARATION POLICY

Grade performance is a demonstrated function of attendance, preparation and participation. You can get behind very easily by skipping classes, resulting in a poor understanding of the material, which will show up as a poor grade for the class. Any class sessions missed by the student are the student's responsibility to make up, not the instructor's. Late arrival that causes disruption, early departure that causes disruption, excessive conversation among students (a disruption in its own right), inappropriate use of electronic devices that cause disruptions, and other actions that disrupt the classroom are unacceptable.

If I am to be absent, my department chair or I will send you notice via email and Blackboard as far in advance as possible. If you are to be absent, report your absence in advance at https://sims.rutgers.edu/ssra/. If your absence is due to religious observance, a Rutgers-approved activity, illness, or family emergency/death and you seek makeup work, also send me an email with full details and supporting documentation within 3 days of your first absence.

If your absence is due to religious observance, a Rutgers-approved activity, illness, or family emergency/death and you seek makeup work, also send an email with full details and supporting documentation within 3 days of your absence. For weather emergencies, consult the campus home page. If the campus is open, class will be held. Expect me to arrive on time for each class session. I expect the same of you. Expect me to remain for the entirety of each class session. I expect the same of you. Expect me to prepare properly for each class session. I expect the same of you. Complete all background reading and assignments. You cannot learn if you are not prepared. The minimum expectation is that for each 3-hour class session, you have prepared by studying for at least twice as many hours. Expect me to participate fully in each class session. I expect the same of you. Stay focused and involved. You cannot learn if you are not paying attention.

TESTS, PROJECTS AND GRADING

A combination of lectures, homework, and coding assignments, students will become well versed in key ideas and techniques for RL. Assignments will include the basics of reinforcement learning in addition, students will do and present an RL final project. Your grade[2] will be based on a final Project, , homework assignments, and class participation, as follows:

- Final Project 50%
- Homework and class participation and conduct 50%

The following rules apply: If you have a disability that influences testing procedures, provide me an official letter from the Office of Disability Services at the start of the semester. No cell phones or other electronics are allowed in the testing room. You must show a valid Rutgers photo ID to enter the room and to turn in the exam. Alternate seating; do not sit next to another student or in your usual seat. Use the bathroom prior to the exam start; bathroom breaks, if essential, will be escorted. Your exam will not be accepted unless you sign the Honor Pledge.

CLASSROOM CONDUCT

The course will be largely taught using computer presentation. Class-related material (lecture notes, messages, etc.) will be

[2]Your final grade is not subject to negotiation. If you feel I have made an error, submit your written argument to me within one week of receiving your final grade. Clarify the precise error I made and provide all due supporting documentation. If I have made an error, I will gladly correct it. But I will adjust grades only if I have made an error.

posted on the class site. Additional visual material and demos may be shown in some classes. *Most of your work will take place outside the classroom, as you study, and apply the material to which you are introduced in class.*

Homework Assignments are designed to help you learn the material discussed in class. In addition doing a thorough job on the homework assignments is the best preparation for the quiz and the final examination. There are three types of assignments: read, prepare, and hand in.

- **Read:** When the assignment is to read some material, this reading is an important introduction to the topics to be discussed in class. I will make the assumption that you have done the reading before class and have understood much (but not necessarily all) of it. When the assignment is to read a problem, that problem will often be used in class to introduce new concepts.
- **Prepare:** Fully analyze the problem. Be ready to discuss it in class, with the numbers computed, etc. I will call on people, so please be ready.
- **Hand In:** The same as prepare, but you must turn in your analysis. All written assignments must be handed in at the beginning of class on the day they are due, and so you will probably want to make a copy of your assignment for reference during class. All written assignments will be graded. These assignments should be submitted in typed form using a word processor. Please write your name, RUID and email on all homework submitted. Team work on this homework is not allowed. Unless a documented reason is produced for unusual circumstances, late submissions will not be accepted more than a week late.

Your class participation will be evaluated subjectively, but will rely upon measures of punctuality, attendance, familiarity with the required readings, relevance and insight reflected in classroom questions, and commentary. Your class participation will be judged by what you add to the class environment, regardless of your technical background. Although several lectures will be didactic, we will rely heavily upon interactive discussion within the class. Students will be expected to be familiar with the readings, even though they might not understand all of the material in advance. In general, questions and comments are encouraged. Comments should be limited to the important aspects of earlier points made, and reflect knowledge of the readings. You may called on to answer questions about the homework or classroom discussion. Your classroom participation evaluation is based on the extent to which you contribute to the learning environment. However, correcting a mistake of the professor and asking what appear to be "dumb questions" about what is being covered are positive contributions. In the case of so-called "dumb questions," very often half of the class will have the same questions in mind and are relieved to have them asked.

Other requirements are: On-time arrival to classes, with uninterrupted attendance for the duration.
Maintenance of a professional atmosphere - use respectful comments and humor.
Refraining from distracting or disrespectful activities, e.g., avoiding side conversations.
Courtesy towards all participants in the classroom.

SUPPORT SERVICES

If you need accommodation for a disability, obtain a Letter of Accommodation from the Office of Disability Services. The Office of Disability Services at Rutgers, The State University of New Jersey, provides student-centered and student-inclusive programming in compliance with the Americans with Disabilities Act of 1990, the Americans with Disabilities Act Amendments of 2008, Section 504 of the Rehabilitation Act of 1973, Section 508 of the Rehabilitation Act of 1998, and the New Jersey Law Against Discrimination. https://ods.rutgers.edu

If you are a military veteran or are on active military duty, you can obtain support through the Office of Veteran and Military Programs and Services. http://veterans.rutgers.edu/

If you are in need of mental health services, please use our readily available services. [Select for inclusion in syllabus based on course location] [Rutgers University-Newark Counseling Center: http://counseling.newark.rutgers.edu/] [Rutgers Counseling and Psychological Services ? New Brunswick: http://rhscaps.rutgers.edu/]

If you are in need of physical health services, please use our readily available services. [Select for inclusion in syllabus based on course location] [Rutgers Health Services ? Newark: http://health.newark.rutgers.edu/] [Rutgers Health Services ? New Brunswick: http://health.rutgers.edu/]

If you are in need of legal services, please use our readily available services: http://rusls.rutgers.edu/