CSC 696H: Topics in Reinforcement Learning Theory

TuTh 3:30pm-4:45pm, C E Chavez Bldg, Rm 305

Description of Course

Reinforcement learning (RL) has achieved great empirical success over the past few decades, and has been used in many fields such as robotics, healthcare, game playing, etc. This course will study RL from a theoretical perspective: when and how can we design RL algorithms with provable guarantees? Specifically, we will look at recent theoretical advances in several representative RL problems, such as RL with a generative model; exploration in RL; RL with function approximation; policy optimization in RL; offline RL. In the first half of this course, students will learn the necessary mathematical tools (such as Markov Decision Processes, concentration inequalities, optimization tools) for the design and analysis of RL algorithms. In the second half of this course, each registered student will present a recent paper on RL theory.

The goal of this graduate seminar course is to learn more about research in the general field of artificial intelligence. In this course, we will read and review research papers on artificial intelligence. We will also learn how to do research in computer science by reading, evaluating, presenting, and conducting a research project in artificial intelligence. Specific topics to be determined by current literature and faculty and student interest.

Course Prerequisites

Students must have strong familiarity with:

- Linear Algebra: linear space, basis, dimensions, linear transformations, matrices, eigenvalues and eigenvectors, positive definiteness of a matrix, matrix decompositions such as the SVD

- Multivariate Calculus: total derivative, gradient, linearity of the derivatives, (second-order) Taylor's expansion

- Basic probability theory: elementary events, definitions of probability, discrete and continuous random variables, distribution laws, (conditional) expectation, (conditional) independence, law of large numbers, central limit theorems

- Basic programming: fluency in at least one programming language (e.g. Matlab, Julia, Python, C, C++), using loops, lists, sorting, traversal in trees.

Other relevant background knowledge (e.g., reinforcement learning, learning theory, algorithm design and analysis, stochastic processes, optimization) is preferred but not required.

Instructor and Contact Information

Chicheng Zhang, GS 720, chichengz@cs.arizona.edu, https://zcc1307.github.io/

Course Modality

This class is scheduled to be taught in the in-person modality.

Course Format and Teaching Methods

In-person lectures (video-recorded), individual written and programming assignments, scribe note taking assignments, in-class student presentations

Obtaining Help

- **Advising:** If you have questions about your academic progress this semester, or your chosen degree program, consider contacting your graduate program coordinator and faculty advisor. Your program coordinator, faculty advisor, and the Graduate Center can guide you toward university resources to help you succeed.**Computer Science students** are encouraged to email gradadvising@cs.arizona.edu for advising related questions.
- **Life challenges:** If you are experiencing unexpected barriers to your success in your courses, please note the Dean of Students Office is a central support resource for all students and may be helpful. The Dean of Students Office can be reached at 520-621-2057 or DOS-deanofstudents@email.arizona.edu.
- **Physical and mental-health challenges**: If you are facing physical or mental health challenges this semester, please note that Campus Health provides quality medical and mental health care. For medical appointments, call (520)-621-9202. For After Hours care, call (520) 570-7898. For the Counseling & Psych Services (CAPS) 24/7 hotline, call (520) 621-3334.

Class Recordings

The lectures will be video-recorded in D2L; if in-class attendees do not wish to be identified by name, please notify the instructor.

For lecture recordings, which are used at the discretion of the instructor, students must access content in D2L only. Students may not modify content or re-use content for any purpose other than personal educational reasons. All recordings are subject to government and university regulations. Therefore, students accessing unauthorized recordings or using them in a manner inconsistent with UArizona values and educational policies (Code of Academic Integrity and the Student Code of Conduct) are also subject to civil action.

Course Objectives

This course will cover the following topics in reinforcement learning theory:

- Basic concepts in Markov decision processes: finite horizon episodic vs. infinite horizon with discounting; stationary vs. nonstationary policy, Bellman equations, occupancy measure
- Planning in Markov decision processes: policy iteration, value iteration, linear programming, and their computational complexities.
- RL with a generative model: sample complexity results
- Exploration in multi-armed and linear bandits: UCB, LinUCB
- Exploration in tabular RL: UCB-VI and extensions
- Exploration in RL with function approximation: LSVI-UCB, OLIVE
- (if time permits) Policy optimization: policy gradient, natural policy gradient and their convergence properties
- (if time permits) Offline RL: fitted Q-iteration

Successful students should be able to use the analytical tools covered in this course to understand

the rationale behind existing RL algorithms and develop new ones. Because of the theoretical nature of this course, students are expected to dedicate a significant amount of time on understanding mathematical concepts and skills outside the classroom.

Expected Learning Outcomes

Students will be able to:

- recognize RL problems in real-world situations, with clear definition of state, action, and rewards
- know what RL problems are more suitable to be modeled as finite horizon episodic setting, versus infinite horizon with discounting setting
- compute value functions, the optimal (action) value function, and the optimal policy given an MDP
- understand the convergence guarantees of policy iteration and value iteration
- understand the sample complexity analysis of RL with a generative model
- recognize the challenges of exploration in RL
- understand the "optimism in the face of uncertainty" principle in RL and recognize algorithms designed based on this principle
- understand the key analysis techniques of optimism-based RL algorithms
- understand the Bellman rank complexity measure and its bounds in tabular MDP and linear MDP settings

Absence and Class Participation Policy

The UA's policy concerning Class Attendance, Participation, and Administrative Drops is available at http://catalog.arizona.edu/policy/class-attendance-participation-and-administrative-drop

Students are encouraged to see the Graduate Program Coordinator (GPC) if they have concerns after the drop period (when a W will not appear on the transcript). The GPC will provide options and alternatives as appropriate for individual student situations.

The UA policy regarding absences for any sincerely held religious belief, observance or practice will be accommodated where reasonable: http://policy.arizona.edu/human-resources/religious-accommodation-policy.

Absences pre-approved by the UA Dean of Students (or dean's designee) will be honored. See https://deanofstudents.arizona.edu/absences

Participating in the course and attending lectures and other course events are vital to the learning process. As such, attendance is required at all lectures and discussion section meetings. Absences may affect a student's final course grade. If you anticipate being absent, are unexpectedly absent, or are unable to participate in class online activities, please contact me as soon as possible. To request a disability-related accommodation to this attendance policy, please contact the Disability Resource Center at (520) 621-3268 or drc-info@email.arizona.edu. If you are experiencing unexpected barriers to your success in your courses, the Dean of Students Office is a central support resource for all students and may be helpful. The Dean of Students Office is located in the Robert L. Nugent Building, room 100, or call 520-621-7057.

Illness

- If you feel sick, or may have been in contact with someone who is infectious, stay home. Except for seeking medical care, avoid contact with others and do not travel.

- Notify your instructor(s) if you will be missing a course meeting or an assignment deadline. Please understand that this does not guarantee an extension of the deadline.
- If you must miss the equivalent of more than one week of class, you should contact the Dean of Students Office DOS-deanofstudents@email.arizona.edu to share documentation about the challenges you are facing.

Makeup Policy for Students Who Register Late

If you register late for this class, contact me as soon as you do. You will be expected to submit all missed assignments within a week of your registration. It is your responsibility to catch up to the class content.

Course Communications

We will use D2L for communications and discussion.

Required Texts and Materials

Most of the lecture materials will be based on the book draft *Reinforcement Learning: Theory and Algorithms*, by Alekh Agarwal, Nan Jiang, Sham Kakade, and Wen Sun, available online.

See also excellent notes from other RL theory courses, e.g.,

- Chi Jin, ELE524: Foundations of Reinforcement Learning
- Nan Jiang, CS598: Statistical Reinforcement Learning
- Shipra Agrawal, IEOR 8100: Reinforcement Learning

Scheduled Topics/Activities

Week 1: Basic concepts in Markov decision processes: finite horizon episodic vs. infinite horizon with discounting; stationary vs. nonstationary policy, Bellman equations, occupancy measure

Week 2: Planning in Markov decision processes: policy iteration, value iteration, linear programming, and their computational complexities.

Week 3: RL with a generative model: sample complexity results; Homework 1 out

Week 4: Exploration in multi-armed and linear bandits: UCB, LinUCB

Week 5: Exploration in tabular RL: UCB-VI and extensions;

Week 6: Exploration in RL with linear and nonlinear function approximation: LSVI-UCB, OLIVE; Homework 2 out

Week 7: Policy optimization: policy gradient, natural policy gradient and their convergence properties

Week 8: Offline RL: fitted Q-iteration

Week 9-15: student presentations

Final Examination or Project

There will be no final exams for this course. Each student will select a paper, either among the provided list of papers, or upon instructor approval, and present it for an hour. The presentation

must include a clear exposition of the problem being addressed, the solution the paper proposes, a comparison of the solution with similar studies, key technical details and proofs, and possible extensions and open problems. Before the presentation, the student is required to schedule a meeting with the instructor to discuss their presentation materials (slides, etc). Throughout the course, the students are highly encouraged to schedule meetings with the instructor about their choice of paper for presentation, their reading progress, etc.

To receive full credit for class participation, students must attend and participate in the discussion of all classes. Students should contact the instructor regarding absences for make-up.

Grading Scale and Policies

The instructing staff will grade your assignments, project, and the final exam on a scale from 0 to 100, with the following weights:
- Class participation: 10%
- Scribe note taking: 10%
- Assignments: 30%
- Paper presentation: 50%

The final grade in the course is determined by the better of a per-class grading curve and overall performance:
- 90% or better: A;
- 80% or better: B;
- 70% or better: C;
- 60% or better: D;
- below 60%: E.

Every homework is due in 2 weeks, and will be returned to students before the next homework is due. Graded homework will be returned before the next homework is due. Exams will be returned within two weeks. Grading delays beyond promised return-by dates will be announced as soon as possible with an explanation for the delay.

As a rule for assignments, each late day for homework assignment submission will result in a deduction of 10% of the grade of the corresponding assignment (e.g., if a student submits their homework solution 6 days after the due date, and it gets a score of 12 (out of 15 points), the submission will receive (100%-60%) * 12 = 4.8 points.

You may petition the professor in writing for an exception if you feel you have a compelling reason for turning work in late.

Incomplete (I) or Withdrawal (W):

Requests for incomplete (I) or withdrawal (W) must be made in accordance with University policies, which are available at http://catalog.arizona.edu/policy/grades-and-grading-system#incomplete and http://catalog.arizona.edu/policy/grades-and-grading-system#Withdrawal respectively.

Dispute of Grade Policy: If you wish to dispute your grade for an assignment, you have two weeks after the grade has been turned in. In addition, even if you only dispute one portion of the grading for that unit, I reserve the right to revisit the entire unit (assignment or project).

Department of Computer Science Code of Conduct

The Department of Computer Science is committed to providing and maintaining a supportive educational environment for all. We strive to be welcoming and inclusive, respect privacy and confidentiality, behave respectfully and courteously, and practice intellectual honesty. Disruptive behaviors (such as physical or emotional harassment, dismissive attitudes, and abuse of department resources) will not be tolerated. The complete Code of Conduct is available on our department web site. We expect that you will adhere to this code, as well as the UA Student Code of Conduct, while you are a member of this class.

Classroom Behavior Policy

To foster a positive learning environment, students and instructors have a shared responsibility. We want a safe, welcoming, and inclusive environment where all of us feel comfortable with each other and where we can challenge ourselves to succeed. To that end, our focus is on the tasks at hand and not on extraneous activities (e.g., texting, chatting, reading a newspaper, making phone calls, web surfing, etc.).

Students are asked to refrain from disruptive conversations with people sitting around them during lecture. Students observed engaging in disruptive activity will be asked to cease this behavior. Those who continue to disrupt the class will be asked to leave lecture or discussion and may be reported to the Dean of Students.

Some learning styles are best served by using personal electronics, such as laptops and iPads. These devices can be distracting to other learners. Therefore, students who prefer to use electronic devices for note-taking during lectures should use one side of the classroom.

Threatening Behavior Policy

The UA Threatening Behavior by Students Policy prohibits threats of physical harm to any member of the University community, including to oneself. See http://policy.arizona.edu/education-and-student-affairs/threatening-behavior-students.

Notification of Objectionable Materials

This course will contain material of a mature nature, which may include explicit language, depictions of nudity, sexual situations, and/or violence. The instructor will provide advance notice when such materials will be used. Students are not automatically excused from interacting with such materials, but they are encouraged to speak with the instructor to voice concerns and to provide feedback.

Accessibility and Accommodations

At the University of Arizona, we strive to make learning experiences as accessible as possible. If you anticipate or experience barriers based on disability or pregnancy, please contact the Disability Resource Center (520-621-3268, https://drc.arizona.edu/) to establish reasonable accommodations.

Code of Academic Integrity

Students are encouraged to share intellectual views and discuss freely the principles and applications of course materials. However, graded work/exercises must be the product of independent effort unless otherwise instructed. Students are expected to adhere to the UA Code of Academic Integrity as described in the UA General Catalog. See http://deanofstudents.arizona.edu/academic-integrity/students/academic-integrity.

Uploading material from this course to a website other than D2L (or the class piazza) is strictly prohibited and will be considered a violation of the course policy and a violation of the code of academic integrity. Obtaining material associated with this course (or previous offerings of this course) on a site other than D2L (or the class piazza), such as Chegg, Course Hero, etc. or accessing these sites during a quiz or exam is a violation of the code of academic integrity. Any student determined to have uploaded or accessed material in an unauthorized manner will be

reported to the Dean of Students for a Code of Academic Integrity violation, with a recommended sanction of a failing grade in the course.

The University Libraries have some excellent tips for avoiding plagiarism, available at http://www.library.arizona.edu/help/tutorials/plagiarism/index.html.

Selling class notes and/or other course materials to other students or to a third party for resale is not permitted without the instructor's express written consent. Violations to this and other course rules are subject to the Code of Academic Integrity and may result in course sanctions. Additionally, students who use D2L or UA email to sell or buy these copyrighted materials are subject to Code of Conduct Violations for misuse of student e-mail addresses. This conduct may also constitute copyright infringement.

Nondiscrimination and Anti-harassment Policy

Required : The University of Arizona is committed to creating and maintaining an environment free of discrimination. In support of this commitment, the University prohibits discrimination, including harassment and retaliation, based on a protected classification, including race, color, religion, sex, national origin, age, disability, veteran status, sexual orientation, gender identity, or genetic information. For more information, including how to report a concern, please see http://policy.arizona.edu/human-resources/nondiscrimination-and-anti-harassment-policy

Our classroom is a place where everyone is encouraged to express well-formed opinions and their reasons for those opinions. We also want to create a tolerant and open environment where such opinions can be expressed without resorting to bullying or discrimination of others.

Additional Resources for Students

UA Academic policies and procedures are available at http://catalog.arizona.edu/policies
Visit the UArizona COVID-19 page for regular updates.

Campus Health
http://www.health.arizona.edu/
Campus Health provides quality medical and mental health care services through virtual and in-person care. Voluntary, free, and convenient COVID-19 testing is available for students on Main Campus. COVID-19 vaccine is available for all students at Campus Health.
Phone: 520-621-9202

Counseling and Psych Services (CAPS)
https://health.arizona.edu/counseling-psych-services
CAPS provides mental health care, including short-term counseling services.
Phone: 520-621-3334

The Dean of Students Office's Student Assistance Program
http://deanofstudents.arizona.edu/student-assistance/students/student-assistance
Student Assistance helps students manage crises, life traumas, and other barriers that impede success. The staff addresses the needs of students who experience issues related to social adjustment, academic challenges, psychological health, physical health, victimization, and relationship issues, through a variety of interventions, referrals, and follow up services.
Email: DOS-deanofstudents@email.arizona.edu
Phone: 520-621-7057

Survivor Advocacy Program
https://survivoradvocacy.arizona.edu/
The Survivor Advocacy Program provides confidential support and advocacy services to student survivors of sexual and gender-based violence. The Program can also advise students about relevant non-UA resources available within the local community for support.
Email: survivoradvocacy@email.arizona.edu

Phone: 520-621-5767

Campus Pantry

Any student who has difficulty affording groceries or accessing sufficient food to eat every day, or who lacks a safe and stable place to live and believes this may affect their performance in the course, is urged to contact the Dean of Students for support. In addition, the University of Arizona Campus Pantry is open for students to receive supplemental groceries at no cost. Please see their website at: *campuspantry.arizona.edu* for open times.

Furthermore, please notify me if you are comfortable in doing so. This will enable me to provide any resources that I may possess.

Preferred Gender Pronoun

This course affirms people of all gender expressions and gender identities. If you prefer to be called a different name than what is on the class roster, please let me know. Feel free to correct the instructor on your preferred gender pronoun. If you have any questions or concerns, please do not hesitate to contact me directly in class or via email (instructor email). If you wish to change your preferred name or pronoun in the UAccess system, please use the following guidelines:

Preferred name: University of Arizona students may choose to identify themselves within the University community using a preferred first name that differs from their official/legal name. A student's preferred name will appear instead of the person's official/legal first name in select University-related systems and documents, provided that the name is not being used for the purpose of misrepresentation. Students are able to update their preferred names in UAccess.

Pronouns: Students may designate pronouns they use to identify themselves. Instructors and staff are encouraged to use pronouns for people that they use for themselves as a sign of respect and inclusion. Students are able to update and edit their pronouns in UAccess.

More information on updating your preferred name and pronouns is available on the Office of the Registrar site at https://www.registrar.arizona.edu/.

Safety on Campus and in the Classroom

Familiarize yourself with the Evacuation and Active Shooter plans specific to Cesar E. Chavez Building. Also watch the video available at https://ua-saem-aiss.narrasys.com/#/story/university-of-arizona-cert/active-shooter

Confidentiality of Student Records

http://www.registrar.arizona.edu/personal-information/family-educational-rights-and-privacy-act-1974-ferpa?topic=ferpa

Land Acknowledgement Statement

We respectfully acknowledge the University of Arizona is on the land and territories of Indigenous peoples. Today, Arizona is home to 22 federally recognized tribes, with Tucson being home to the O'odham and the Yaqui. Committed to diversity and inclusion, the University strives to build sustainable relationships with sovereign Native Nations and Indigenous communities through

education offerings, partnerships, and community service.

Subject to Change Statement

Information contained in the course syllabus, other than the grade and absence policy, may be subject to change with advance notice, as deemed appropriate by the instructor.

Classroom attendance

- If you feel sick, or may have been in contact with someone who is infectious, stay home. Except for seeking medical care, avoid contact with others and do not travel.
- Notify your instructor(s) if you will be missing a course meeting or an assignment deadline.
- Non-attendance for any reason does **not** guarantee an automatic extension of due date or rescheduling of examinations/assessments.
- Please communicate and coordinate any request directly with your instructor.
- If you must miss the equivalent of more than one week of class, you should contact the Dean of Students Office DOS-deanofstudents@email.arizona.edu to share documentation about the challenges you are facing.
- Voluntary, free, and convenient COVID-19 testing is available for students on Main Campus.
- If you test positive for COVID-19 and you are participating in on-campus activities, you must report your results to Campus Health. To learn more about the process for reporting a positive test, visit the Case Notification Protocol.
- COVID-19 vaccine is available for all students at Campus Health.
- Visit the UArizona COVID-19 page for regular updates.

Statement on compliance with COVID-19 mitigation guidelines

As we enter the Fall semester, your and my health and safety remain the university's highest priority. To protect the health of everyone in this class, students are required to follow the university guidelines on COVID-19 mitigation. Please visit www.covid19.arizona.edu.