COMPSCI 687: Reinforcement Learning
Syllabus

Fall 2022

Instructor:	Bruno Castro da Silva	**Time:**	Tue & Thur from 4:00 to 5:15 pm
Email:	bsilva@cs.umass.edu	**Place:**	Hasbrouck Lab Add, Room 124

1 About the course

This course will provide an introduction to, and comprehensive overview of, reinforcement learning (RL). Reinforcement learning is a branch of artificial intelligence focused on learning to make decisions—based on the interactions of an agent with its environment—in order to efficiently solve a problem. Reinforcement learning algorithms repeatedly answer the question "What should be done next?". They learn to solve tasks/problems via trial-and-error, even when there is no supervisor telling the algorithm what the correct decision would have been in a few sample situations or contexts.

RL algorithms have been successfully deployed in a wide range of real-life problems. Examples include:

- Robotics applications (*What activity should a robot perform next to more rapidly clean a kitchen?*);
- Achieving super-human performance in complex video games (*What action should the game character execute next?*);
- Performing package delivery by teams of drones (*When and where should a drone drop a package?*);
- Creating personalized recommendations (*Which advertisement, song, or movie should an intelligent system show or recommended to a particular user?*);
- Medical applications (*How much insulin should be injected next? What drug should be given next?*);
- Environmental applications (*Which countermeasure for an invasive species should be deployed next?*);
- Dialogue systems (*What sentence should be spoken next to keep a user engaged with the system?*);
- Helping better understand how our brains work (*Where and how in the brain are estimates of expected reward represented and updated, given an animal's experiences?*);
- Implementing brain-machine interfaces (*How should a neural decoder map brain signals to commands sent, e.g, to a robotic arm?*).

Applications such as these are bound to change the way we interact with artificial agents, making reinforcement learning one of the most promising areas of machine learning.

Broad topics covered in this course will include: Markov decision processes, reinforcement learning algorithms (model-free, batch/online, value function-based, actor-critics, policy gradient methods, etc.), and representations for reinforcement learning. Special topics may include ensuring the safety of reinforcement learning algorithms, hierarchical reinforcement learning, model-based algorithms, theoretical reinforcement learning, multi-agent reinforcement learning, and connections to animal learning.

In this course, each voice in the classroom has something of value to contribute. Please take care to respect the different experiences, beliefs, and values expressed by students and staff involved in this course. My colleagues and I support UMass' commitment to diversity, and welcome individuals regardless of age, background, citizenship, disability, sex, gender, gender identity, sexual orientation, education, ethnicity, family status, geographical origin, language, military experience, political views, race, religion, socioeconomic status, and work experience.

2 Website

This course's notes and syllabus will be hosted here. Homework assignments and other material will be posted on Moodle. Lectures will be recorded; recordings (along with .pdf slides) will also be on Moodle.

3 Class

Classes will be held on Tuesdays and Thursdays from 4:00pm–5:15pm in Hasbrouck Lab Add, Room 124.

4 Book

Parts of the course will be roughly based on the *second* edition of Sutton and Barto's book, *Reinforcement Learning: An Introduction*. It can be found on Amazon here. It is also available for free online here. Although the book is a fantastic introduction to the topic (and I encourage purchasing a copy if you plan to study reinforcement learning), owning the book is not a requirement.

5 Required background

We assume students have appropriate mathematical background in probability, statistics, multivariate calculus, linear algebra, and programming. The following references can provide a useful review:

- Probability Theory, by Maleki and Do.
- Linear Algebra and Matrix Calculus, by Kolter and Do.
- An Introduction to Statistical Learning, by James, Witten, Hastie and Tibshirani.
- Optimization: Any calculus textbook.
- Machine Learning: a Probabilistic Perspective, by Kevin Patrick Murphy.

6 Grading

Your grade will have three components:

1. **Homework Assignments** (50%): There will be frequent homework assignments, both written and programming. All assignments will have an equal weight.
2. **Midterm exam** (30%).
3. **Project** (20%): As reinforcement learning transitions from an academic curiosity to practical tools that you may use in your professional lives, it is critical that we study how to implement, fine-tune, and deploy these algorithms in practice. Further details will be available when the project is assigned, after the most relevant course material has been covered.

A cumulative grade in $[90\% - 100\%]$ will be an A- or A, $[75\%, 90\%)$ will be a B-, B, or B+, $[65\%, 75\%)$ will be a C-, C, or C+, and $[55\% - 65\%)$ will be a D or D+. Course grades will be curved only in students' favor (that is, these thresholds may be lowered, but a grade of 90% will not be lower than an A-). **Some extra credit opportunities may be given. Your grade may be reduced by any amount at the instructor's discretion due to inappropriate behavior, such as academic dishonesty.**

6.1 Late Policy

- Deadlines in this course are **strict**. A submission one minute after the deadline will receive zero credit. You are strongly encouraged to submit hours before any deadline.
- Having said that, to allow some flexibility to complete assignments (*homeworks*) given other constraints, you have a total of *five* free late days that you can choose to use when submitting a homework. You will be charged one late day for handing in an assignment within 24 hours after it is due, two late days for handing in an assignment within 48 hours after it is due, etc. Your assignment is considered late if either the written or code portions are submitted late. The late homework clock stops when both the written and code portions are submitted. After you have used up your late days, late homework will not count for credit except in special circumstances (e.g., illness documented by a doctor's note).
- All exams must be taken at the scheduled time unless *(1)* there is a documented conflict and arrangements have been made with the instructor before the exam; or *(2)* you have a medical emergency and you bring proof of such to the instructor before final grades for the given exam are computed. In any other case (unless those covered by the University's Academic Regulations), missing an exam will result in a grade of "F" for that exam.

7 Office Hours

The teaching assistants (TAs) this semester will be Aline Weber (alineweber@cs.umass.edu) and Dhawal Gupta (dgupta@cs.umass.edu).

Office hours will be held according to the following schedule (starting on 09/12), except *(1)* on holidays; *(2)* if the UMass official schedule follows a different day of the week; or *(3)* when noted otherwise.

	Monday	**Tuesday**	**Wednesday**	**Thursday**
8:30am–11:30pm				Dhawal
9am–1pm	Aline	Aline		
1:00pm–4:00pm				Dhawal
4:00pm–6:00pm			Dhawal	
5:15pm–6:30pm		Bruno		Bruno

Locations:

- **Monday, Tuesday, Wednesday, and Thursday office hours:**
  - TA office hours: LGRT T220.
  - Prof. da Silva's office hours will be held in immediately after each lecture. He will remain in the classroom for 1 hour and 15 minutes, post-lecture, helping students as needed.

8 SAT/Fail

At some time near the end of the semester (likely around the last day of class), you will be given the option to take the class SAT/Fail rather than for a letter grade. If you plan to take the course SAT/Fail, keep an eye out for an email (or a message on Moodle) from me around the end of the semester with instructions for requesting SAT/Fail. If you elect SAT/Fail, you will earn a SAT grade if your letter grade would have been a C or higher, and you will receive an F if your letter grade would have been lower. The above conditions *do not* hold for students with an academic honesty violation. In these cases, the requests described in this section are disallowed and/or un-approved.

9 Disability Services

The University of Massachusetts is committed to providing an equal educational opportunity for all students. If you have a documented physical, psychological, or learning disability on file with Disability Services, you may be eligible for academic accommodations to help you succeed in this course. If you would like to register with Disability Services, please visit their website or their office (161 Whitmore Administration Building; phone (413) 545–0892). Finally, if you have a documented disability that requires an accommodation, please notify me within the first two weeks of the semester so that we can make appropriate arrangements.

10 Cheating

- Cheating will not be tolerated. Assignments may include instructions about what forms of collaboration are allowed, if/when relevant.
- Copying answers or code from external sources (books, web pages, etc.), from other students, or from solutions to assignments from previous years is *always* considered cheating. To emphasize: no detectable copying is acceptable, even, e.g., copying a single sentence from an outside source. Sharing your code or solutions with other students is also considered cheating.
- The College of Information and Computer Sciences explicitly forbids any redistribution (including publicly available posting on an internet site) of any CICS course materials (including student solutions to course assignments, projects, exams, etc.) without the express written consent of the instructor of the course from which the materials come. Violations of this policy will be deemed instances of "facilitating dishonesty" (since a student making use of such materials would be guilty of plagiarism) and therefore may result in charges under the Academic Honesty Policy.
- Since students are expected to be familiar with this policy and the commonly accepted standards of academic integrity, ignorance of such standards is not normally sufficient evidence of lack of intent.
- **All instances of cheating will be reported to the university's Academic Honesty Board**, and will result in an F for the course.

11 COVID-19 and Face Covering Policy

Students and course staff in COMPSCI 687 are expected to do their part to slow the spread of COVID-19 and minimize the risk of illness for all community members. It is important to remember that community members may be vulnerable or live with vulnerable individuals.

Masks are welcome on campus, and we encourage everyone to respect the choices that individuals make about their own masking. Masking is **strongly** encouraged during the **first few weeks** of the Fall semester.

For general information, please refer to the official UMass FAQ regarding COVID-19 and masking.

Students and course staff in COMPSCI 687 must comply with all university policies regarding COVID-19.

- **If you test positive for COVID-19**, *you need to self-isolate*: don't come to class! Individuals who test positive for COVID-19 are required to isolate for a minimum of *five days* before returning to class. They should then continue to wear a mask for an additional *five days*.
- **If you were exposed to COVID-19**, you need to wear a high-grade mask (such as KN95, KF94, or N95 for *ten days*). You also need to get tested at least *five full days* after your exposure.

If you are in any doubt, please do not come to class—you can request an excused absence from any course meeting by sending the instructor an email; you will not be penalized for any missed quizzes, discussion exercises, etc.