Control Systems and Reinforcement Learning

EEL 6935 Section 1234

Class Periods: Days of week, period, and corresponding time of day

Location: **TBD**

Academic Term: Spring 2021

Instructor:
Sean Meyn
meyn@ece.ufl.edu 392-8934
Office Hours: TBD
Please contact through the Canvas website

Supervised Teaching Student:
Joel Mathias
Office Hours: TBD
Please contact through the Canvas website

Course Description

Reinforcement learning is a collection of tools for the design of decision and control algorithms. What makes RL different from traditional control is that the modelling step is avoided, and instead the control design is based on observations of the system to be controlled. Its origins are considered to be found in the 1980s program led by Richard Sutton, and the field made headlines in the public press more recently following the success of AlphaGo and other RL algorithms that beat grand masters at complex games like Go and Chess. Today it is hoped that RL will be an engine behind autonomous cars, as well as better decision making in fields ranging from medicine to finance. This course provides an introduction to RL through the lens of control theory. We will find that the DQN algorithm behind AlphaGo is related to classical control concepts going back to the 1960s. Given this intuition we will discover techniques to create new and potentially more reliable algorithms for decision and control.

Course Pre-Requisites / Co-Requisites

Required: EEL 5182 State Variable Methods in Linear Systems (or equivalent), and experience with Python or Matlab. Some exposure to probability at the undergraduate level is also required. The most important prerequisite is mathematical maturity.

Course Objectives

By the end of the semester, the student is expected to master these building blocks for RL design:

- Basics of dynamic programming for deterministic optimal control, and how dynamic programming equations motivate RL algorithms, including TD- and Q-learning.
- Basics of optimization in the context of machine learning: the role of convexity, and gradient descent techniques for the construction of algorithms.
- ODE methods to guide the creation of recursive algorithms, along with intuition regarding the solidarity between a practical recursive algorithm and its ODE approximation.

These building blocks will provide a rich set of tools for RL design that students will compare through numerical experiments using Python or Matlab.

Materials and Supply Fees

List if applicable

Required Textbooks and Software

Feedback Systems and Reinforcement Learning, Sean Meyn, 2020 (in preparation for publication, Cambridge University Press). A draft manuscript will be made available to students.

Recommended Materials

- R. Sutton and A. Barto. Reinforcement Learning: An Introduction. MIT Press. On- line edition at http://www.cs.ualberta.ca/~sutton/book/the-book.html, Cambridge, MA, 2nd edition, 2018.
- Karl J. Åström and Richard M. Murray. Feedback Systems: An Introduction for Scientists and Engineers, 2nd edition under construction: http://www.cds.caltech.edu/~murray/amwiki/index.php/Second_Edition and Murray's crash course from 2018, https://simons.berkeley.edu/talks/murray-control-1
- Csaba Szepesvari. Algorithms for Reinforcement Learning. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, 2010 https://sites.ualberta.ca/~szepesva/papers/RLAlgsInMDPs.pdf
- Theory of Reinforcement Learning: tutorials at the Simons Institute program, Aug. 19--Dec. 18, 2020, https://simons.berkeley.edu/programs/rl20

Course Schedule

Week 1: *You Have a Control Problem*. A survey of control philosophy from the point of view of both control and RL practitioners. Examples of models, state space models, and the meaning of "state" in different settings.
Week 2: Stability, performance, and control Lyapunov function design. Examples. HW1 due
Week 3: Optimal control and dynamic programming. Inverse dynamic programming.
A first glance at Q-learning. HW2 due
Week 4: The linear quadratic regulator and potential RL implementations. HW3 due
Week 5: ODE Methods for Algorithm Design.
What is an ODE? When is it stable? From ODE to algorithm. HW4 due
Week 6: Newton-Raphson flow. Applications to optimization.
One hour exam
Week 7: Stability of recursive algorithms and their ODE approximations. An introduction to stochastic approximation in a deterministic setting (quasi stochastic approximation, or QSA). Techniques to accelerate convergence, such as Ruppert-Polyak averaging (more advanced techniques may be introduced in assignments).
HW5 due
Week 8: Policy gradient methods for optimal control based on QSA techniques. HW6 due
Week 9: Survey of machine learning. Function approximation based on training data. The basics of kernel methods. The evolving theory of exploration.
Week 10: QSA and TD-learning, SARSA, and approximate policy iteration. HW7 due
Week 11: Watkins' Q-learning and its progeny
One hour exam
Week 12: Projected Bellman equations, and GQ learning
Week 13: Convex analytic approaches to optimal control.
LQR and semidefinite programming. HW8 due
Week 14: Convex Q-learning and DQN: variations and theory
Week 15: Future directions HW9 due

Project due during finals period

Attendance Policy, Class Expectations, and Make-Up Policy

Excused absences must be in compliance with university policies in the Graduate Catalog (https://catalog.ufl.edu/graduate/regulations/) and require appropriate documentation.

Evaluation of Grades

The homework assignments are designed to provide the students with an enhanced exposure to the theoretical concepts presented in lecture, and apply the theory to algorithm design. There will also be a significant component devoted to testing designs through computer simulation, using either Matlab or Python. The students are encouraged to work in groups, *but will **not** receive full credit unless they provide independent reports that describe their activities, and provide commentary on the numerical results obtained*.

The two midterms will be 90 minutes each: closed book, in-class, and comprehensive.

As for the final project, below is the format for my course on stochastic control. The instructions will be similar, but less demanding since the students will arrive with less mathematical maturity.

Instead of a final exam, each of you will pick out one research paper on stochastic control, and one paper that is referenced in your chosen paper. There are many excellent sources for papers: Some of my favorite work comes from the MIT school, Bertsekas, Tsitsiklis, and Van Roy (former student of Tsitsiklis, now at Stanford). You can find many references on my own website. *I want each of you to have unique papers.*

You will write a report on the two papers, using the following format.

(a) The report will be about 4 pages long, *not including any illustrations or computer plots you might want to include*, and it should be typed.

(b) The report will consist of three parts: A summary of the paper considered; a critique; a discussion of possible extensions of the results described in the paper; and a discussion on your supporting numerical experiments.

(c) The *summary* must be concise — consisting of approximately one page. It should be clear enough to allow a fellow student to understand the main ideas of the paper.

(d) The *critique* should compare the results of the paper to what has been discussed in class, and should indicate the merits/shortcomings of the paper. It should explain how the paper is related to the supporting reference that you chose.

(e) The *extensions* section should be treated like a proposal for a thesis. What further directions might be explored? What other methods could the author have considered? Are there related applications that might be examined?

(f) *Numerical experiments*. In the final part of the report, use a computer program such as Matlab to perform simulations to experiment with the issues raised in the paper. You will not receive credit for this part of the report unless you provide meaningful interpretations of your simulation results.

You will be free to collaborate, but you must work independently on the write-up. In particular, you must show creativity (your own creativity) in your interpretations of your results, and *you must draw connections between the work done in the project and our course material.*

Assignment	Total Points	Percentage of Final Grade
Homework Sets (9)	100 each	15%
Midterm exams (2)	100 each	60%
Review Paper / Project	100	25%
		100%

Grading Policy

Percent	Grade	Grade Points
90.0 - 100.0	A	4.00
87.0 - 89.9	A-	3.67
84.0 - 86.9	B+	3.33
81.0 – 83.9	B	3.00
78.0 - 80.9	B-	2.67
75.0 - 79.9	C+	2.33
72.0 – 74.9	C	2.00
69.0 - 71.9	C-	1.67
66.0 - 68.9	D+	1.33
63.0 - 65.9	D	1.00
60.0 - 62.9	D-	0.67
0 - 59.9	E	0.00

More information on UF grading policy may be found at: https://catalog.ufl.edu/graduate/regulations/

Students Requiring Accommodations
Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, https://www.dso.ufl.edu/drc) by providing appropriate documentation. Once registered, students will receive an accommodation letter which must be presented to the instructor when requesting accommodation. Students with disabilities should follow this procedure as early as possible in the semester.

Course Evaluation
Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at https://gatorevals.aa.ufl.edu/students/. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via https://ufl.bluera.com/ufl/. Summaries of course evaluation results are available to students at https://gatorevals.aa.ufl.edu/public-results/.

University Honesty Policy
UF students are bound by The Honor Pledge which states, "We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment." The Honor Code (https://sccr.dso.ufl.edu/policies/student-honor-code-student-conduct-code/) specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor or TAs in this class.

Commitment to a Safe and Inclusive Learning Environment
The Herbert Wertheim College of Engineering values broad diversity within our community and is committed to individual and group empowerment, inclusion, and the elimination of discrimination. It is expected that every person in this class will treat one another with dignity and respect regardless of gender, sexuality, disability, age, socioeconomic status, ethnicity, race, and culture.

If you feel like your performance in class is being impacted by discrimination or harassment of any kind, please contact your instructor or any of the following:
• Your academic advisor or Graduate Program Coordinator
• Robin Bielling, Director of Human Resources, 352-392-0903, rbielling@eng.ufl.edu

• Curtis Taylor, Associate Dean of Student Affairs, 352-392-2177, taylor@eng.ufl.edu
• Toshikazu Nishida, Associate Dean of Academic Affairs, 352-392-0943, nishida@eng.ufl.edu

Software Use

All faculty, staff, and students of the University are required and expected to obey the laws and legal agreements governing software use. Failure to do so can lead to monetary damages and/or criminal penalties for the individual violator. Because such violations are also against University policies and rules, disciplinary action will be taken as appropriate. We, the members of the University of Florida community, pledge to uphold ourselves and our peers to the highest standards of honesty and integrity.

Student Privacy

There are federal laws protecting your privacy with regards to grades earned in courses and on individual assignments. For more information, please see: https://registrar.ufl.edu/ferpa.html

Campus Resources:

Health and Wellness

U Matter, We Care:
Your well-being is important to the University of Florida. The U Matter, We Care initiative is committed to creating a culture of care on our campus by encouraging members of our community to look out for one another and to reach out for help if a member of our community is in need. If you or a friend is in distress, please contact umatter@ufl.edu so that the U Matter, We Care Team can reach out to the student in distress. A nighttime and weekend crisis counselor is available by phone at 352-392-1575. The U Matter, We Care Team can help connect students to the many other helping resources available including, but not limited to, Victim Advocates, Housing staff, and the Counseling and Wellness Center. Please remember that asking for help is a sign of strength. In case of emergency, call 9-1-1.

Counseling and Wellness Center: http://www.counseling.ufl.edu/cwc, and 392-1575; and the University Police Department: 392-1111 or 9-1-1 for emergencies.

Sexual Discrimination, Harassment, Assault, or Violence
If you or a friend has been subjected to sexual discrimination, sexual harassment, sexual assault, or violence contact the **Office of Title IX Compliance**, located at Yon Hall Room 427, 1908 Stadium Road, (352) 273-1094, title-ix@ufl.edu

Sexual Assault Recovery Services (SARS)
Student Health Care Center, 392-1161.

University Police Department at 392-1111 (or 9-1-1 for emergencies), or http://www.police.ufl.edu/.

Academic Resources

E-learning technical suppor*t*, 352-392-4357 (select option 2) or e-mail to Learning-support@ufl.edu. https://lss.at.ufl.edu/help.shtml.

Career Resource Center, Reitz Union, 392-1601. Career assistance and counseling. https://www.crc.ufl.edu/.

Library Support, http://cms.uflib.ufl.edu/ask. Various ways to receive assistance with respect to using the libraries or finding resources.

Teaching Center, Broward Hall, 392-2010 or 392-6420. General study skills and tutoring. https://teachingcenter.ufl.edu/.

Writing Studio, 302 Tigert Hall*,* 846-1138. Help brainstorming, formatting, and writing papers. https://writing.ufl.edu/writing-studio/.

Student Complaints Campus*:* https://care.dso.ufl.edu.

On-Line Students Complaints*:* http://www.distance.ufl.edu/student-complaint-process.