PSYC 599 Spring 2023

Reinforcement Learning

Payam Piray

E-mail: piray@usc.edu
Office Hours: Wed. 2-4pm
Office: SGM 822

Class Hours: Mon. & Wed. 10-11:50
Class Room: THH 111

Course Description

Neuroscience and artificial intelligence (AI) have advanced rapidly in the past two decades. Both fields have enormously benefited from reinforcement learning (RL), the application of machine learning to sequential decision tasks with the goal of optimizing future return. In AI, RL is one of the three core artificial learning paradigms, along with supervised learning and unsupervised learning, and has led to several of the significant breakthroughs in the field in recent years (e.g., artificial RL systems defeating human champions of Go and chess). For neuroscience, RL provides a theoretical understanding of decision making and learning, which is needed to understand how to interpret, manipulate, and measure the neural mechanisms that support decision making in the brain. For instance, RL has facilitated breakthrough discoveries about the function of the neuromodulator dopamine in learning. RL also has applications in other disciplines that are concerned with sequential decision making, including economics and public policy.

Learning Objectives

This course is focused on core theoretical concepts in RL as well as its applications in both neuroscience and AI. Following a successful completion of the course, students are expected to have learned how to define a decision problem in the language of RL, use different algorithms to solve the problem, and be able to identify the pros and cons of each algorithm. Furthermore, students are expected to learn how RL is related to important topics in neuroscience, including the role of dopamine in learning, grid fields, and cognitive maps. Students are also expected to understand and work with deep RL solutions for large-scale problems. Students also get acquainted with advanced RL methods and their applications.

This course is intended for PhD students, and is cross-listed in psychology, neuroscience, electrical engineering, and computer science. Consequently, it is understood that students with different backgrounds will be differentially prepared for different parts of the course. An important learning component of this course will be the formal, mathematical analysis of the problems, which requires basic knowledge of probability and linear algebra. Furthermore, the course is interdisciplinary in terms of both topic and style. Students are expected to be interested in a deep understanding of RL and the frontiers of its applications in both AI and neuroscience.

- **Prerequisite(s):** Instructor permission
- **Recommended Preparation:** For Non-Engineering and Non-CS majors: Psych 625 or a similar course.

Course Notes

This course is a mix of lectures and paper discussions. The lectures will cover the fundamental theoretical concepts, while discussions — each led by a student — will focus on key applications in either AI or neuroscience. Required readings will be articles from the primary literature (usually one per week). The lectures may cover materials that are not covered in the readings, and students are responsible for the additional materials covered only in the lectures. Students are also expected to complete the assigned readings on time.

Although there is no official course textbook, there is one that covers roughly 70% of the lecture materials. There is no need to buy it tough; it is freely available online. Individual chapters of the textbook will be assigned for optional reading. Note that students must read the assigned primary literature papers, and they count toward the final grade.

- Sutton, R. S. and Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, Massachusetts, second edition [available at `http://incompleteideas.net/book/RLbook2020.pdf`]

Description and Assessment of Assignments

1. Reaction emails. Students are asked to write a brief note (one or two paragraphs) about their reactions to the week's reading assignments. These notes should be sent to the instructor prior to the discussion class (Sunday nights). All reaction emails should have the subject line "RL-course-reaction."
2. Paper discussion. Each student will present a series of assigned papers and will moderate the discussion that follows. These are similar to journal club meetings. Moderators should prepare slides, but they are encouraged not to work extensively on slides (or to understand every detail of the papers). It is crucial, however, that they have a fundamental understanding of the primary topics and concepts given in the assigned papers.
3. Paper presentation. Each student is required to present a recent theory paper related to the course, either from AI or neuroscience. A list of papers will be provided, and students are encouraged to choose a paper from the list (although it is possible to present a paper proposed by the student, subject to approval.) Paper presentations will be scheduled for the second half of the semester. The main goal of these presentations is for students to get experience presenting for an interdisciplinary audience.

4. Final project. In lieu of a final exam, there will be a final project. This will center around proposing an experiment to further test predictions of a theoretical paper (preferably the paper presented by the student) and simulating the experiment. This will include a final presentation (30 minutes), as well as a final report (4-6 pages.)

   The final report should include an introduction to the theoretical topic presented in the theory paper, the rationale for the proposed experiment, methods, results, and discussion. The code for the project should be submitted alongside the final report. It is NOT necessary that the suggested experiment be totally successful. It may indeed fail to some extent, if not entirely. In fact, even if their experiment fails, students may still get an A. However, students should go into further detail in their final report about why the experiment did not work and how to improve it in the future.

Grading Policy

- **10%** Active participation
- **15%** Reaction emails
- **15%** Paper discussion
- **25%** Paper presentation (midterm)
- **15%** Final presentation
- **20%** Final report

Assignment Submission Policy

Students are required to send their reaction emails at least one day before the corresponding discussion class. The final report is due on May 3.

Schedule and weekly learning goals

NOTE: The schedule is tentative and subject to change.

Week 1, Mon. 01/09. Introduction to RL

Week 1, Wed. 01/11. Dynamic programming

- Sutton and Barto, chapters 3-4 (optional)

Week 2, Mon. 01/16. Holiday

Week 2, Wed. 01/18. Model-free RL

- Sutton and Barto, chapters 5-6 (optional)

Week 3, Mon. 01/23. Discussion: RL in engineering vs RL in neuroscience

- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. W. H. Freeman and Company, San Francisco (CA), first edition (only chapter 1)

- Griffiths, T. L., Lieder, F., and Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, 7(2):217–229

Week 3, Wed. 01/25. Model-free RL in the brain

- Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599 (optinal)

- Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., and Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*, 16(7):966–973 (optinal)

- Daw, N. D. and Tobler, P. N. (2013). *Neuroeconomics: Chapter 15. Value Learning through Reinforcement: The Basics of Dopamine and Reinforcement Learning*. Academic Press, 2nd edition edition (optinal)

Week 4, Mon. 01/30. Discussion: Model-free vs Model-based in the brain

- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., and Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, 69(6):1204–1215

- Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, 8(12):1704–1711

- Keramati, M., Dezfouli, A., and Piray, P. (2011). Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS computational biology*, 7(5):e1002055 (optional)

Week 4, Wed. 02/01. Model-based RL

- Sutton and Barto, chapter 8 (optional)

Week 5, Mon. 02/06. Discussion: replay in the brain

- Foster, D. J. and Wilson, M. A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, 440(7084):680–683

- Pfeiffer, B. E. and Foster, D. J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, 497(7447):74–79

Week 5, Wed. 02/08. Deep RL

- Sutton and Barto, chapter 9 (optional)
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., and Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533 (optional)

Week 6, Mon. 02/13. Discussion: deep RL for games

- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., and Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484–489
- Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T., and Hassabis, D. (2017). Mastering the game of Go without human knowledge. *Nature*, 550(7676):354–359

Week 6, Wed. 02/15. Cognitive maps, grid fields and RL

- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55(4):189–208 (optional)
- Moser, E. I., Kropff, E., and Moser, M.-B. (2008). Place cells, grid cells, and the brain's spatial representation system. *Annual Review of Neuroscience*, 31:69–89 (optional)
- Stachenfeld, K. L., Botvinick, M. M., and Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, 20(11):1643–1653 (optional)

Week 7, Mon. 02/20. Holiday

Week 7, Wed. 02/22. Discussion: grid-like deep RL

- Banino, A., Barry, C., Uria, B., Blundell, C., Lillicrap, T., Mirowski, P., Pritzel, A., Chadwick, M. J., Degris, T., Modayil, J., Wayne, G., Soyer, H., Viola, F., Zhang, B., Goroshin, R., Rabinowitz, N., Pascanu, R., Beattie, C., Petersen, S., Sadik, A., Gaffney, S., King, H., Kavukcuoglu, K., Hassabis, D., Hadsell, R., and Kumaran, D. (2018). Vector-based navigation using grid-like representations in artificial agents. *Nature*, 557(7705):429–433

Week 8, Mon. 02/27. Paper presentation

Week 8, Wed. 03/01. Policy gradient methods

- Sutton and Barto, chapter 13 (optional)

Week 9, Mon. 03/06. Paper presentation

Week 9, Wed. 03/08. Paper presentation

Week 10, Mon. 03/13. Spring recess

Week 10, Wed. 03/15. Spring recess

Week 11, Mon. 03/20. Paper presentation

Week 11, Wed. 03/22. Exploration vs exploitation in RL

- Sutton and Barto, chapter 2 (otional)
- Gittins, J. C. (1979). Bandit Processes and Dynamic Allocation Indices. *Journal of the Royal Statistical Society. Series B (Methodological)*, 41(2):148–177 (optional)

Week 12, Mon. 03/27. Discussion: continuous control with deep RL

- Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., and Wierstra, D. (2019). Continuous control with deep reinforcement learning

Week 12, Wed. 03/29. Planning as Bayesian inference

- Todorov, E. (2009). Efficient computation of optimal actions. *Proceedings of the National Academy of Sciences of the United States of America*, 106(28):11478–11483 (optional)
- Levine, S. (2018). Reinforcement Learning and Control as Probabilistic Inference: Tutorial and Review. *arXiv:1805.00909 [cs, stat]* (optional)

Week 13, Mon. 04/03. Discussion: meta RL

- Finn, C., Abbeel, P., and Levine, S. (2017). Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks

Week 13, Wed. 04/05. Planning under uncertainty

- Kaelbling, L. P., Littman, M. L., and Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1):99–134 (optional)

Week 14, Mon. 04/10. Discussion: soft actor-critic

- Haarnoja, T., Zhou, A., Abbeel, P., and Levine, S. (2018). Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor

Week 14, Wed. 04/12. Final presentation

Week 15, Mon. 04/17. Final presentation

Week 15, Wed. 04/19. Final presentation

Week 16, Mon. 04/24. Final presentation

Week 16, Wed. 04/26. Review

Statement on Academic Conduct and Support Systems

Academic Conduct

Plagiarism — presenting someone else's ideas as your own, either verbatim or recast in your own words — is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in *SCampus* in Section 11, *Behavior Violating University Standards* `https://scampus.usc.edu/1100-behavior-violating-university-standards-and-appropriate-sanctions/`. Other forms of academic dishonesty are equally unacceptable. See additional information in *SCampus* and university policies on scientific misconduct, `http://policy.usc.edu/scientific-misconduct/`.

Discrimination, sexual assault, and harassment are not tolerated by the university. You are encouraged to report any incidents to the *Office of Equity and Diversity* `http://equity.usc.edu/` or to the *Department of Public Safety* `http://capsnet.usc.edu/department/department-public-safety/online-forms/contact-us`. This is important for the safety whole USC community. Another member of the university community — such as a friend, classmate, advisor, or faculty member — can help initiate the report, or can initiate the report on behalf of another person. *The Center for Women and Men* `http://www.usc.edu/student-affairs/cwm/` provides 24/7 confidential support, and the sexual assault resource center webpage `sarc@usc.edu` describes reporting options and other resources.

Support Systems

A number of USC's schools provide support for students who need help with scholarly writing. Check with your advisor or program staff to find out more. Students whose primary language is not English should check with the *American Language Institute* `http://dornsife.usc.edu/ali`, which sponsors courses and workshops specifically for international graduate students. *The Office of Disability Services and Programs* `http://sait.usc.edu/academicsupport/centerprograms/dsp/home_index.html` provides certification for students with disabilities and helps arrange the relevant accommodations. If an officially declared emergency makes travel to campus infeasible, *USC Emergency Information* `http://emergency.usc.edu/will` provide safety and other updates, including ways in which instruction will be continued by means of blackboard, teleconferencing, and other technology.

IMPORTANT: COVID-19 PROTOCOLS

Students must comply with all COVID-19 safety protocols outlined by federal, state, local, and university policies. These policies will likely evolve with the changing conditions of the COVID-19 pandemic and may include social distancing, the use of face coverings at all times, proof of vaccination, and regular COVID testing, among others. Depending on the policies outline by the above authorities, and the conditions of the class, the class might switch between meeting online and in person.