

Advanced Topics in Reinforcement Learning

Lecture 1: Course Introduction

Josiah Hanna

University of Wisconsin — Madison

About Me

- Assistant professor in CS since 2021.
- Research focus on AI and robotics with an emphasis on reinforcement learning.
- B.S. in CS and Math from the University of Kentucky.
- Ph.D. from the University of Texas at Austin.
- Post-doc at the University of Edinburgh.
- I enjoy: running, being outside, reading, and spending time with family (wife, kids (ages 1 and 3), 2 dogs).

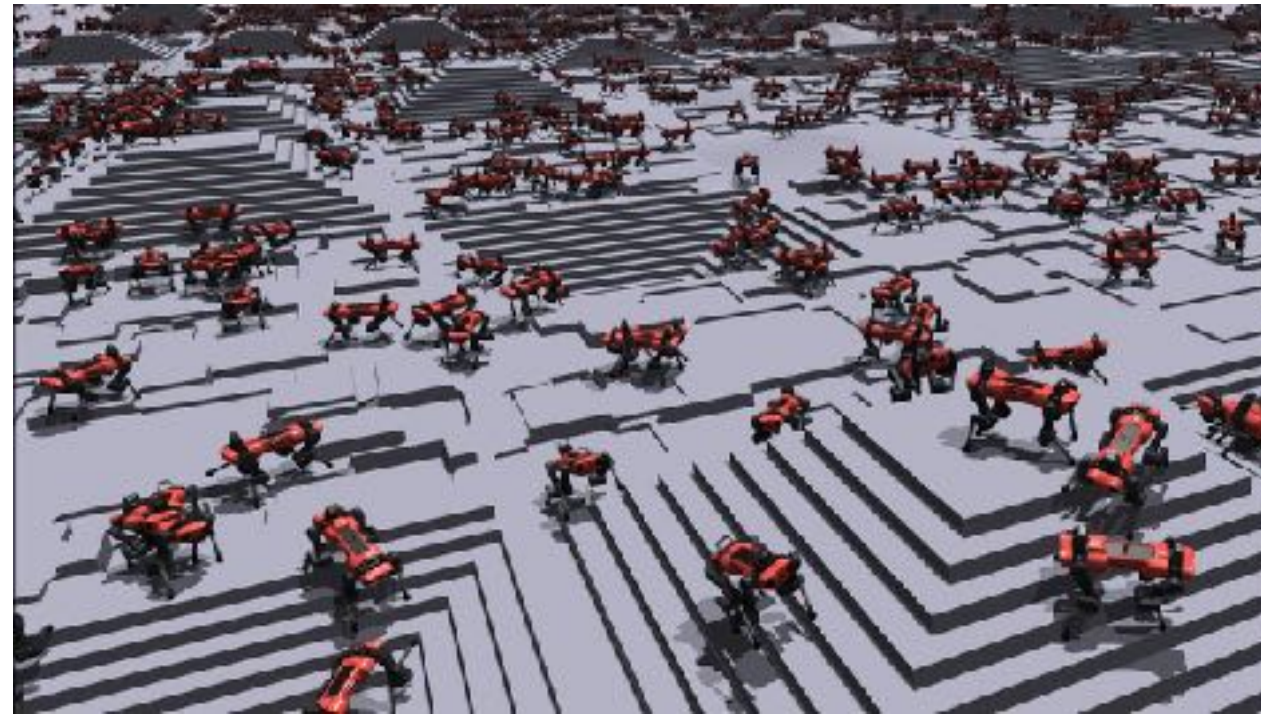
What is Reinforcement Learning?

- Type of machine learning that focuses on learning from evaluative feedback (rewards/punishments) and trial and error interaction.
- The learning agent takes actions, receives rewards, and over time learns to take actions that lead to the most reward.
- Think: training a dog to do tricks.



What Can RL Do?

- Play video games
- Play board games
- Control robots
- Recommend ads and web content
- Trade stocks
- Recommend medical treatments
- Control home thermostat systems
- Cooling of data centers
- Networking
- Databases
- Program Synthesis
- Post-train LLMs

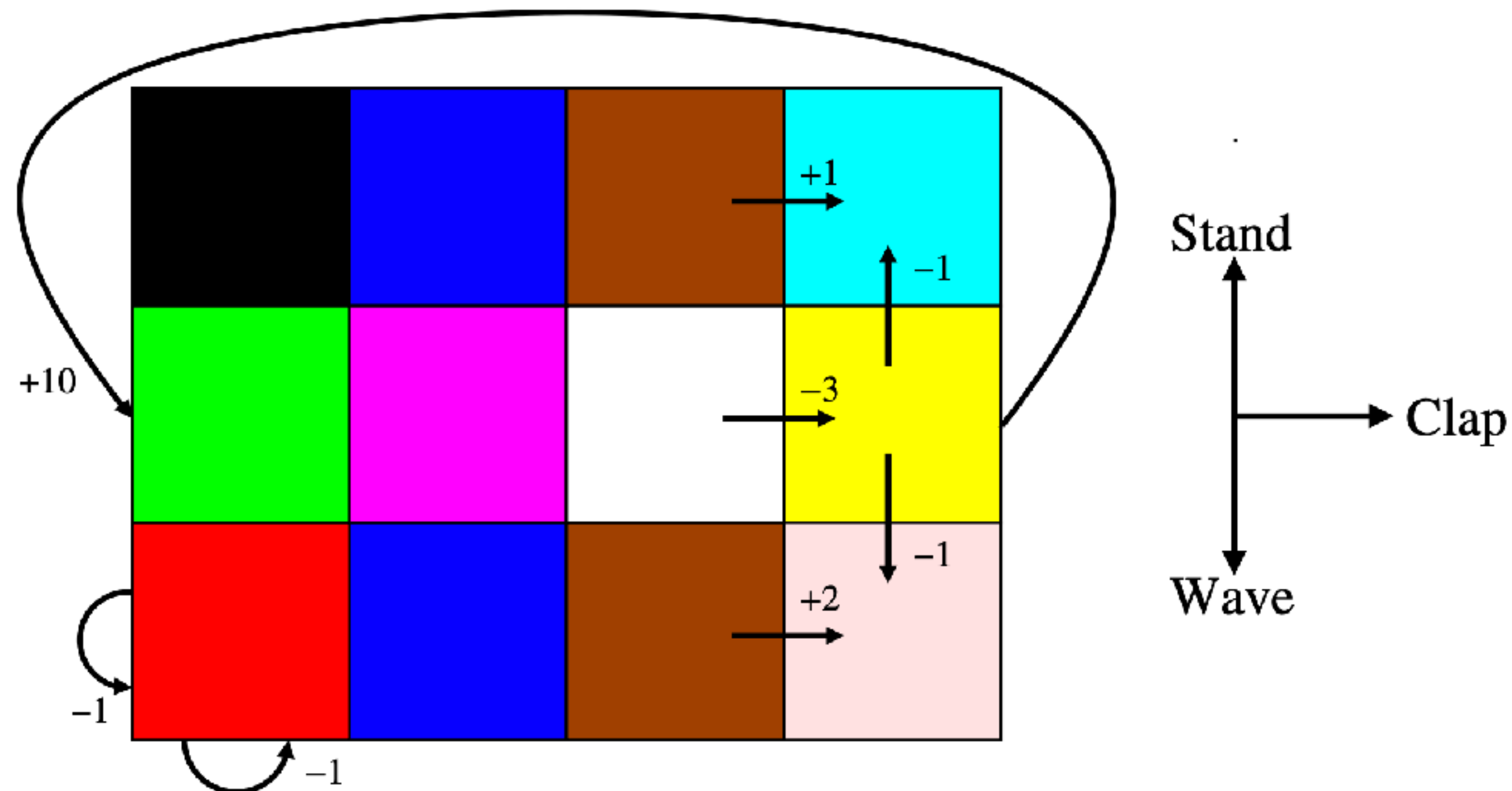


Be an RL Agent*

- You (as a class) are the learning agent.
- Three actions: stand, clap, or wave
- Observations: colors $\in \{\text{red, blue, orange, pink}\}$
- Rewards: depends on color you see and action you take.
- Goal: find the optimal policy.
 - Policy: mapping from colors to actions.
 - Optimal policy: policy that gives you the most reward.

Be an RL Agent

- How did you learn?
- What structure does the world have?

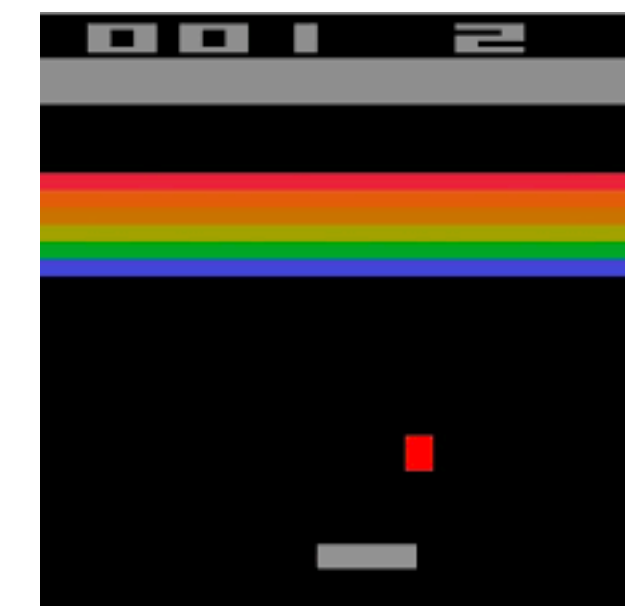


Reinforcement Learning Problems

- States: 3x4 grid
- Observations: colors
 - In this class (and the course textbook), states and observations will be treated the same.
- Actions: stand, clap, wave
- Rewards: +1, +2, -1, or +10
- State transitions dependent on action chosen.

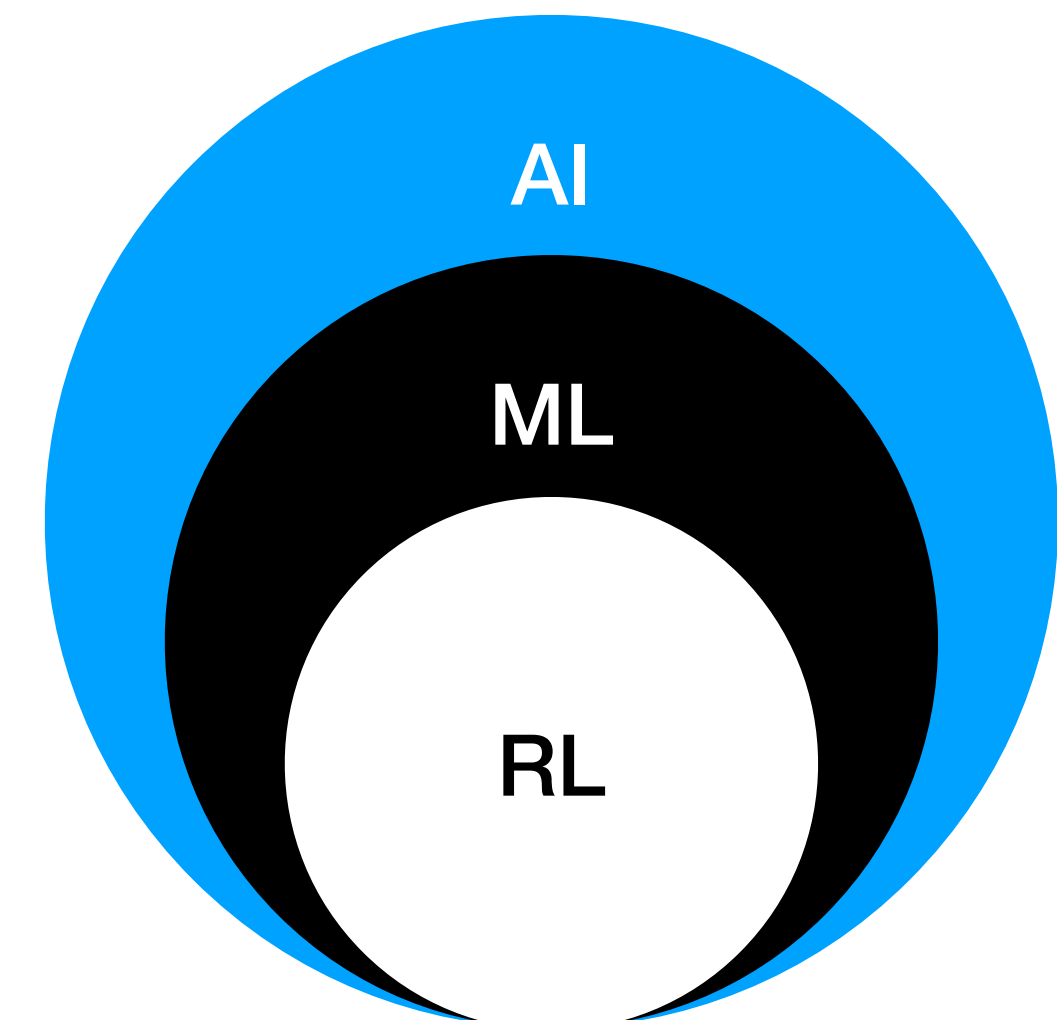
Example RL Problems

- What are the states? Actions? Rewards?
- Atari Breakout
- Home thermostat
- Stock trading



RL within Artificial Intelligence

- Supervised learning: learn from labelled examples.
 - Given a data set of $\{(X,Y)\}$, learn to map new instances of X to appropriate Y .
 - Ex: image classification, object detection, spam filtering.
- Unsupervised learning: discover structure in unlabelled data.
 - Ex: clustering, synthesizing images, language modeling
- Reinforcement learning: learn from evaluative feedback.
- Reinforcement learning also relates to AI planning.



Why is RL hard?

- Credit Assignment:
 - May take many actions before reward is received. Which ones were most important?
 - Example: you study 15 minutes a day all semester. The morning of the final exam, you eat a bowl of yogurt. You receive an A on the final. Was it the studying or the yogurt that led to the A?
 - Not trivial for people and animals either!
- Exploration vs. Exploitation
 - Should you keep trying actions that led to reward in the past or try new actions that might lead to even more reward?

Course Goals

After taking this course, you will be able to:

1. Explain fundamental RL concepts and apply fundamental RL algorithms.
2. Explain distinctions between advanced topics in RL research and the problems the research aims to address.
3. Complete an RL research project including implementation and experimental analysis of that implementation.

Related Courses

Some RL in context of
LLM post-training

Proof-based analysis of
RL, bandits, games.

Some RL in context of
robotics

This Course!

RL Fundamentals and
Advanced Topics

1-2 weeks on RL basics

10,000 Foot Preview

- RL Fundamentals (~ 2 months)
 - Tabular methods
 - RL with function approximation
 - Research papers in RL fundamentals
- Advanced Topics in RL Research (~ 1.5 month)

Schedule Overview

- See course webpage: https://pages.cs.wisc.edu/~jphanna/teaching/25fall_cs839/schedule.html

Syllabus

- Spend 10 minutes reading the course webpage.
- With 2-3 people sitting next to you, discuss potential confusion, questions, and concerns.
 - Prepare questions to ask after 10 minutes is up.
- https://pages.cs.wisc.edu/~jphanna/teaching/25fall_cs839/index.html

Pre-requisites

- Background survey: <https://forms.gle/hU6cM3gLxpNrdpq17>
- Probability and Statistics
 - Random variables, probability distributions, expectation, bias, variance, random sampling.
- Linear Algebra: dot-product, transpose, vector-matrix multiplication, matrix inverses.
- Calculus: basic differentiation with respect to scalar and vector variables.
- Machine Learning: neural networks, linear and logistic regression.
- Programming: one assignment in Python. Possibly advantageous for final project to know Python.

Class Periods

- **[Before class]** Required Weekly Readings
 - Most weeks will be from the course textbook.
 - **Submit reactions and questions by Monday at 12pm US central time.**
 - Submit on Gradescope by the deadline.
- Lecture and group activities
- Student paper presentations

Reading Responses

- Credit is based on evidence that the reading was completed.
- Responses and questions will be used to shape the week's lecture.
- Possible responses:
 - Questions
 - Critiques or suggestions for extensions.
 - What you want to learn about more.
 - Thoughts on what you find most important.

Student Presentations

- Goal: expose the entire class to a variety of RL research papers.
- One time during the semester, you will read and prepare a 10 minute presentation on an assigned paper.
- The number per class period may vary but in general there will be two speakers each class for the first part of the course.
- Sign-ups will be available soon.
- I have curated a list of papers to sign-up, however, feel free to email me with alternatives.

Attendance Policy

- Class attendance and participation is necessary for participation component of grade.
- Absences will be approved if an email is sent before class starts.
- Please remain at home if unwell!

Logistics

- Course Webpage: https://pages.cs.wisc.edu/~jphanna/teaching/25fall_cs839/index.html
- Piazza: <https://piazza.com/wisc/fall2025/cs839004/home>
- Canvas: <https://canvas.wisc.edu/courses/477251>
- Gradescope: <https://www.gradescope.com/courses/1109753>
 - Access code: X2RDNJ
- Presentation sign-ups: <https://docs.google.com/spreadsheets/d/1PMI8XO9IP84GW5jYFJi1qPo6E19ZKacw5nRKxY7YTu8/edit?usp=sharing>
- Office Hours: Tuesdays @ 11am-12pm (after lecture) or by appointment.

Action Items

- Join Piazza!
- Join Gradescope!
- Read Chapters 2 and 3 of course textbook (skim 2.5-2.9).
- Send a reading response by 12pm on Monday.
- Sign-up for a presentation (link on Piazza).