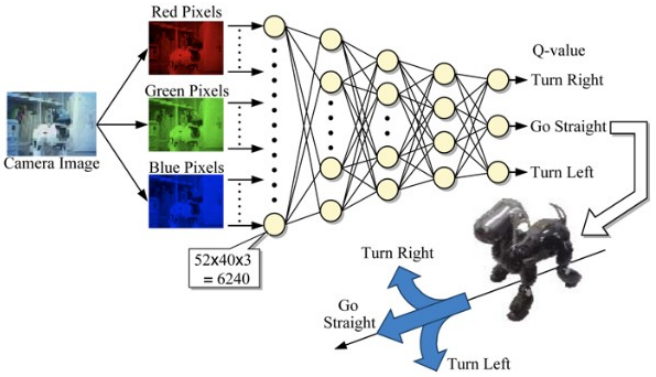
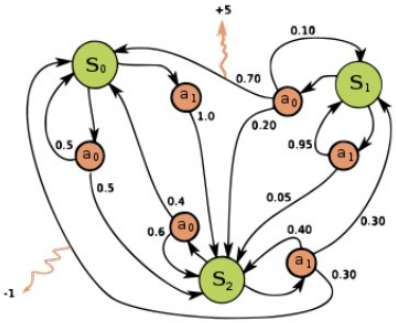
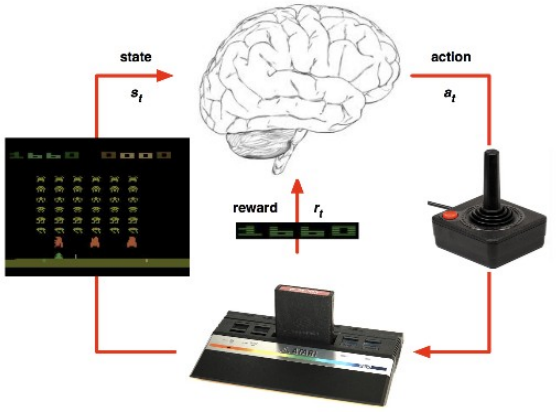
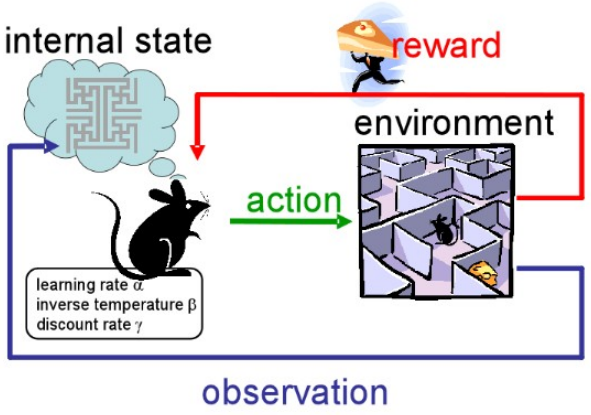
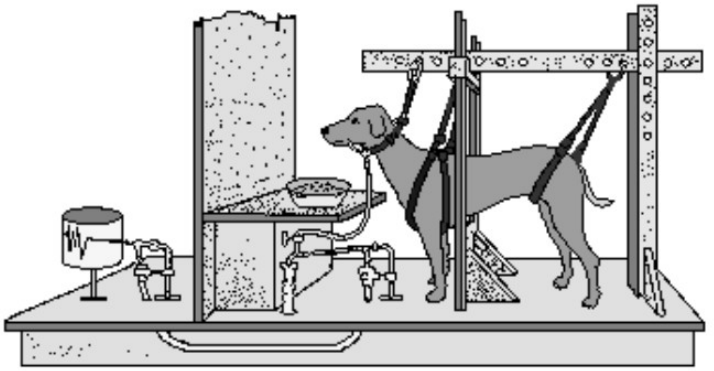
COMP 138: Reinforcement Learning











Instructor: Jivko Sinapov
Webpage: TBD

BE a reinforcement learner

- You, as a class, will act as the learning agent

BE a reinforcement learner

- You, as a class, will act as the learning agent
- **Actions:** wave, clap, or nod

BE a reinforcement learner

- You, as a class, will act as the learning agent
- **Actions:** wave, clap, or nod
- **Observations:** color, reward

BE a reinforcement learner

- You, as a class, will act as the learning agent
- **Actions:** wave, clap, or nod
- **Observations:** color, reward
- **Goal:** find an optimal *policy*

BE a reinforcement learner

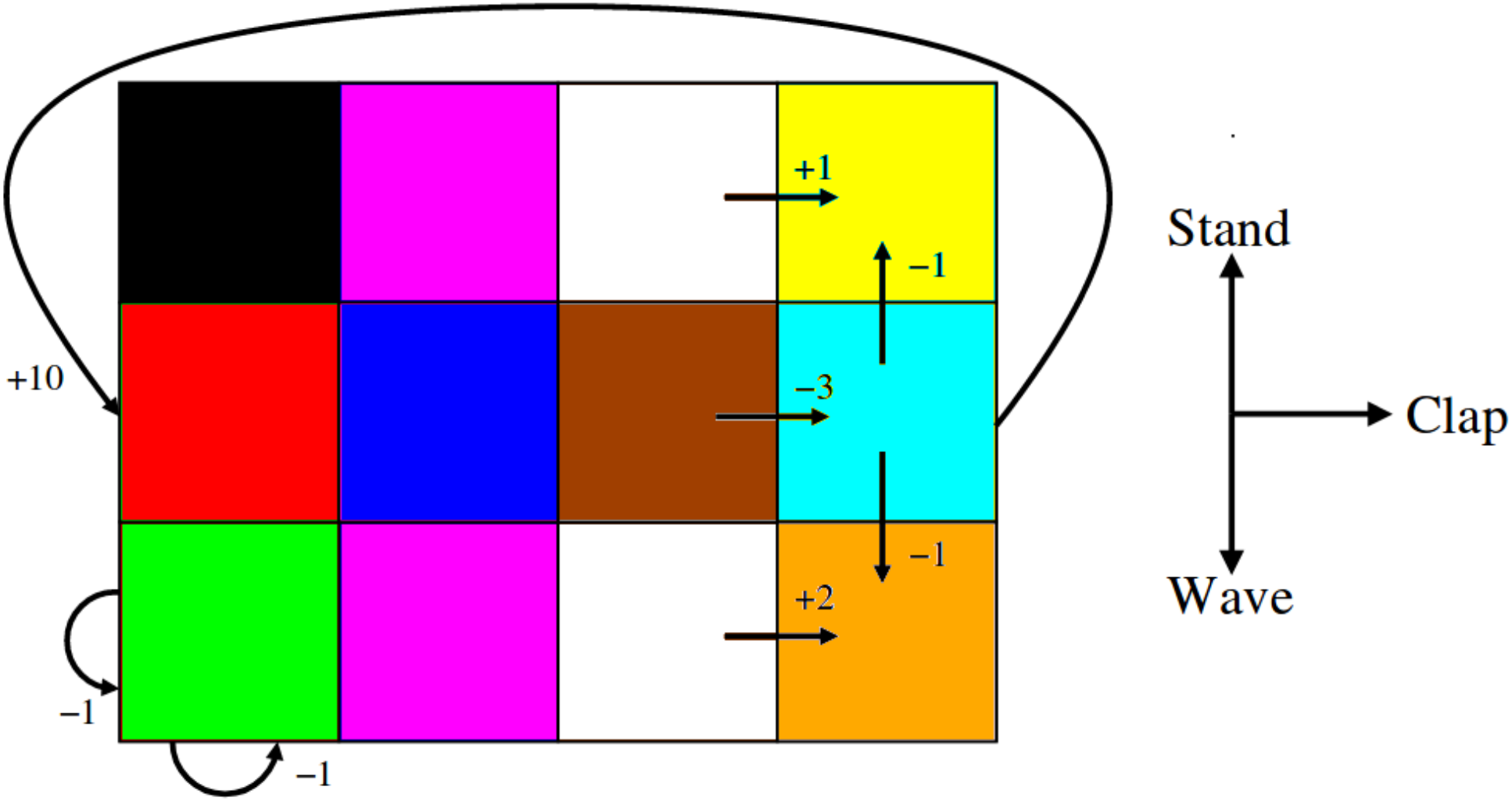
- You, as a class, will act as the learning agent
- **Actions:** wave, clap, or stand
- **Observations:** color, reward
- **Goal:** find an optimal *policy*
  - What is a policy? What makes a policy optimal?

How did you do it?

- What is your policy, and how is it represented?
- What does the world look like?

What actually happened...

What actually happened...



Note: yellow = universe and cyan = cat

Now, let’s formalize this

(board or writing projector)

About this course

- Reinforcement Learning **theory** & **practice**
- Theory at the start and practice towards end
- Syllabus = the course web page

Where does RL fall within the field of Artificial Intelligence?

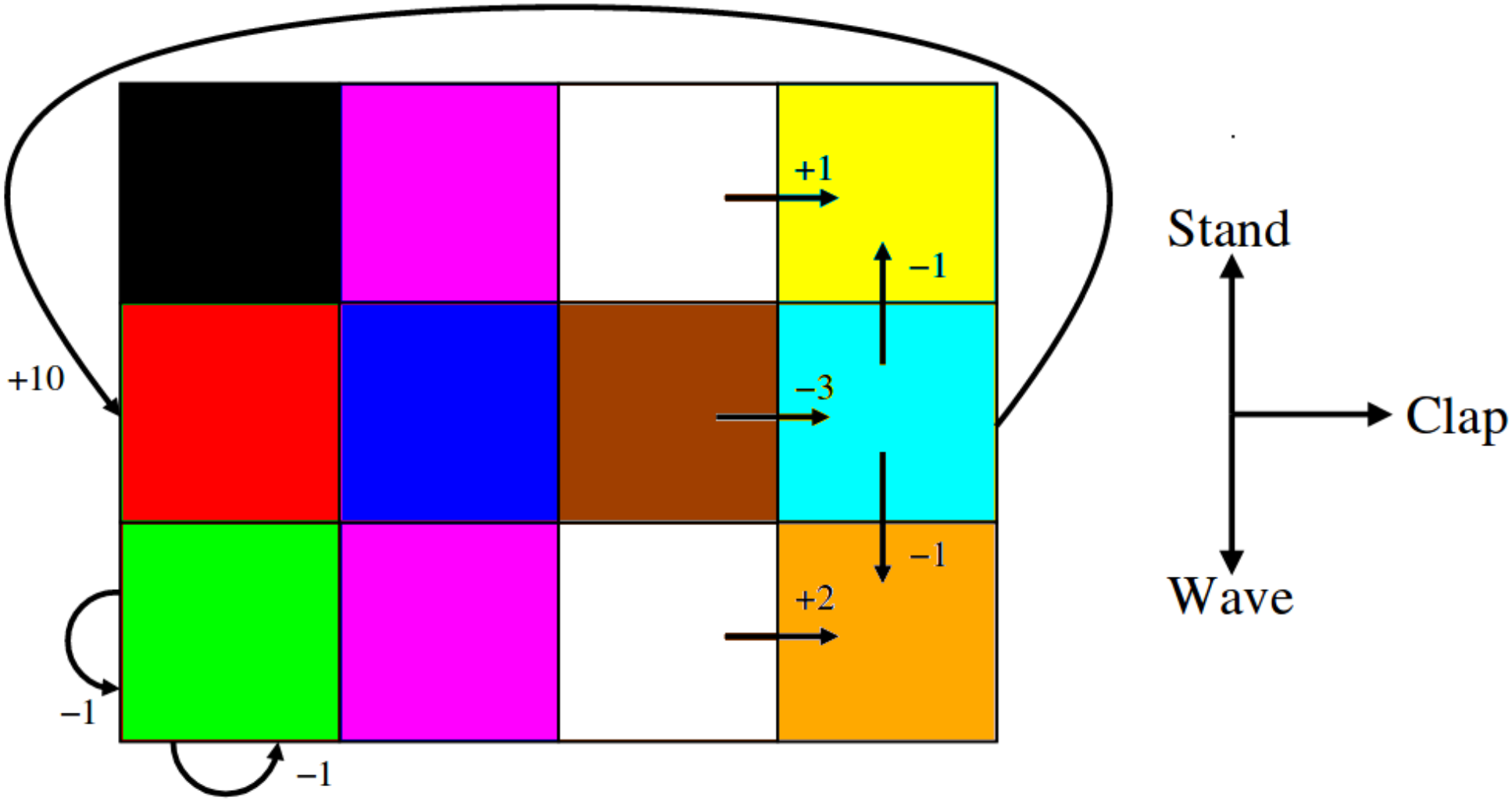
Where does RL fall within the field of Artificial Intelligence?

- AI → ML → RL

Where does RL fall within the field of Artificial Intelligence?

- AI → ML → RL
- Type of Machine Learning:
  - **Supervised**: learn from labeled examples
  - **Unsupervised**: learn from unlabeled examples
  - **Reinforcement**: learn through interaction

Reduced Formalism



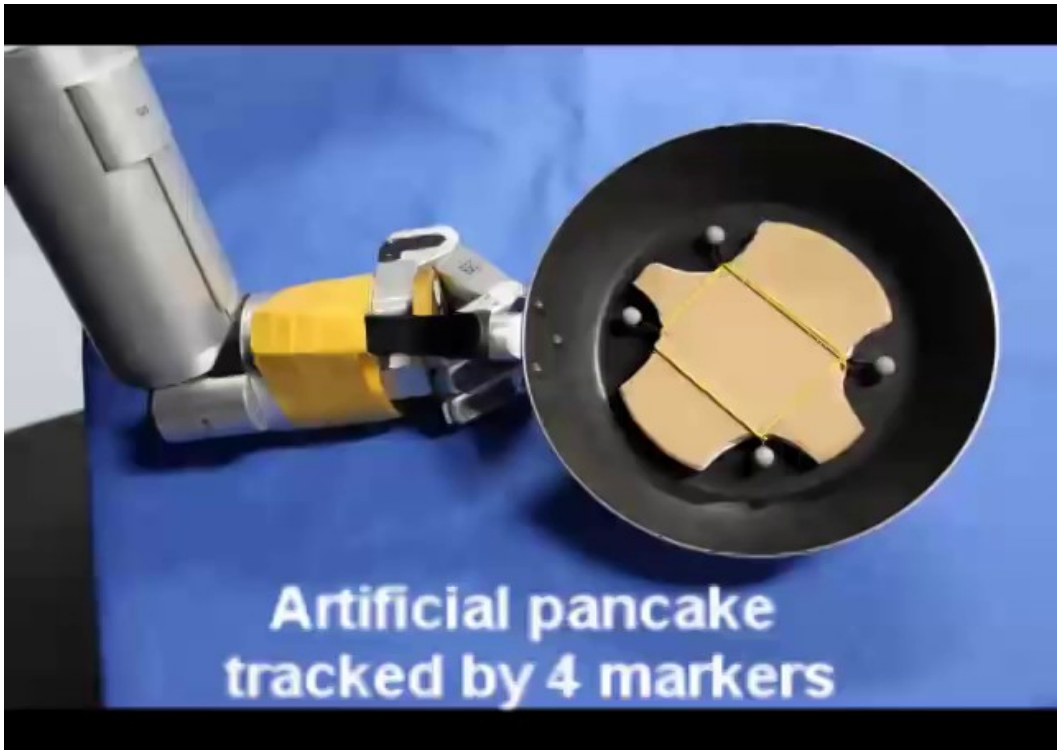
Reduced Formalism

(board or writing projector)

Take-home Message

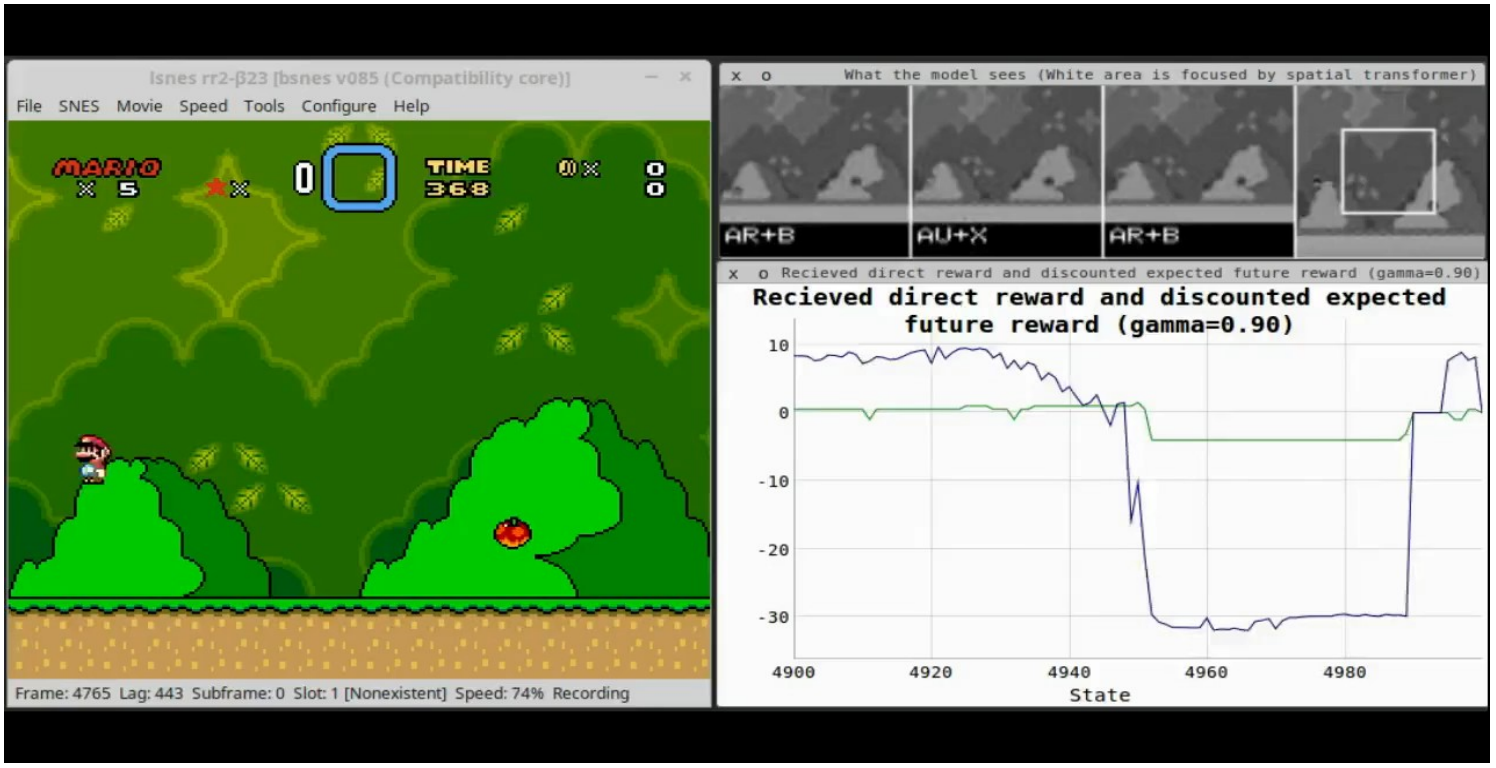
- Agent’s perspective: only the **policy** is under control
- State representation and reward function are given
- Focus on policy algorithms
- Appeal: program agents by just specifying goals
- Practice: need to pick state representation and reward function

Example Applications



https://www.youtube.com/watch?v=W_gxLKSsSIE

Example Applications



https://www.youtube.com/watch?v=L4KBBAwF_bE&t=96s

Reading Assignment

- Chapter 1 and 2 of Sutton and Barto
- Reading response on Canvas due 9/12 before class starts

Programming Assignments

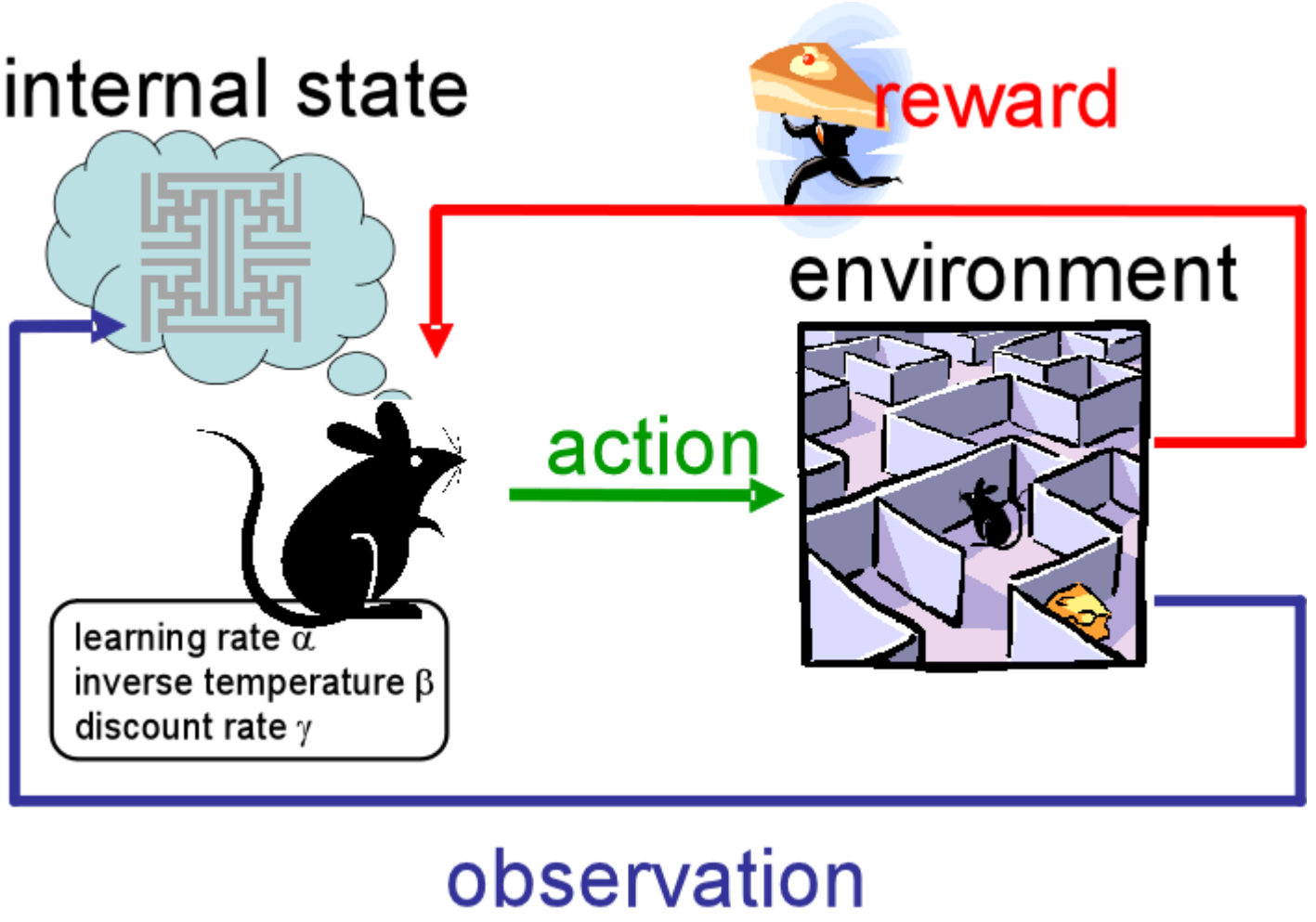
- Students are required to complete 4 minor programming assignments of their choosing
- Default options: programming exercises from Sutton and Barto (let’s look at some examples)

Discussion Moderation

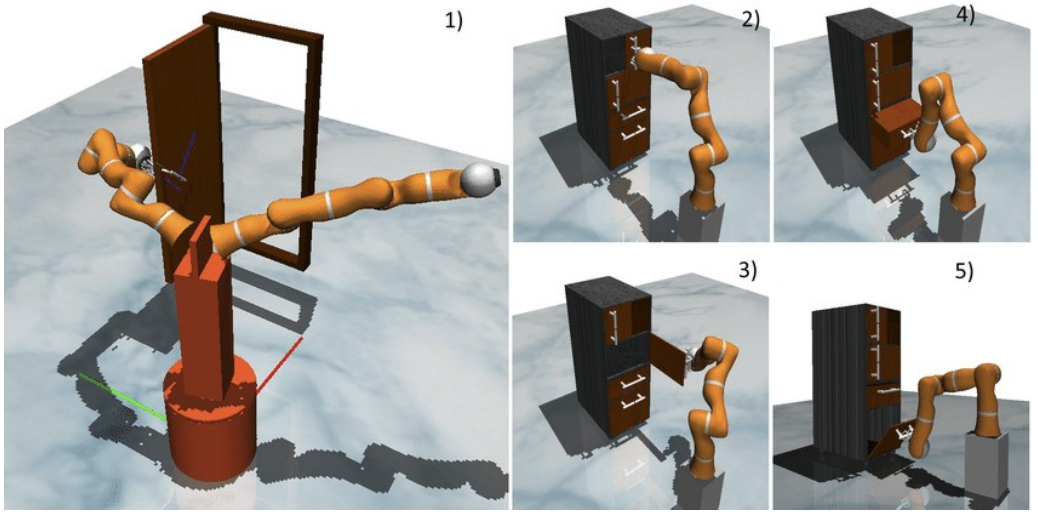
- Each graduate (MS or PhD) student will lead a reading discussion once during the semester
- Students can team up in a pair
- Sign up sheet will be posted to Canvas tonight
- Extra credit for anyone who volunteers for slots in the next week
- Presentation materials / notes or description of what will be discussed should be emailed to me 48 hours before the class

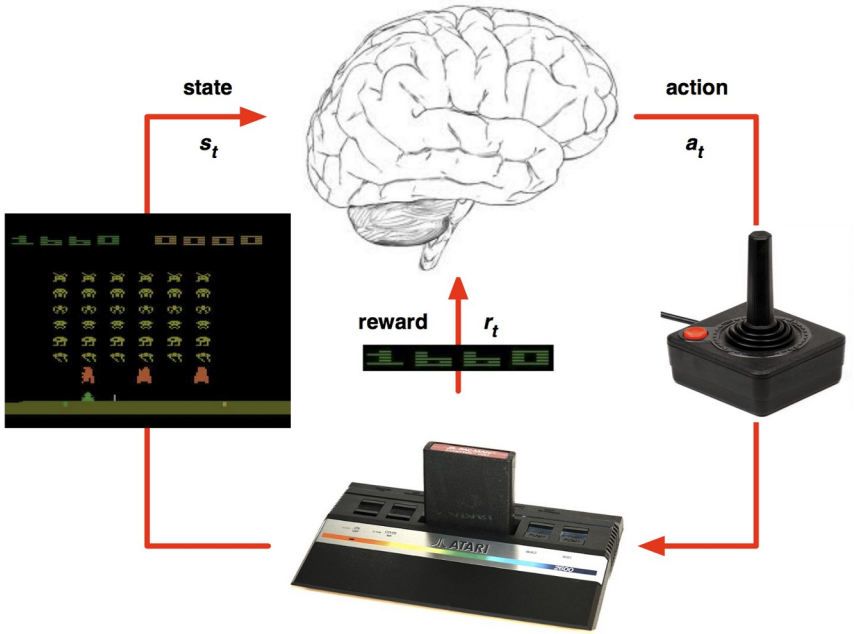
Next time...

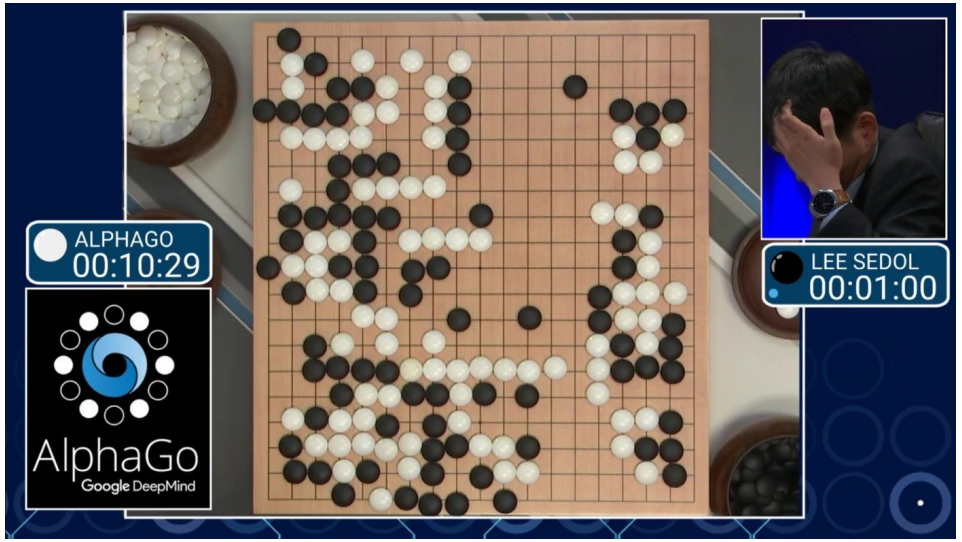
COMP 150: Reinforcement Learning

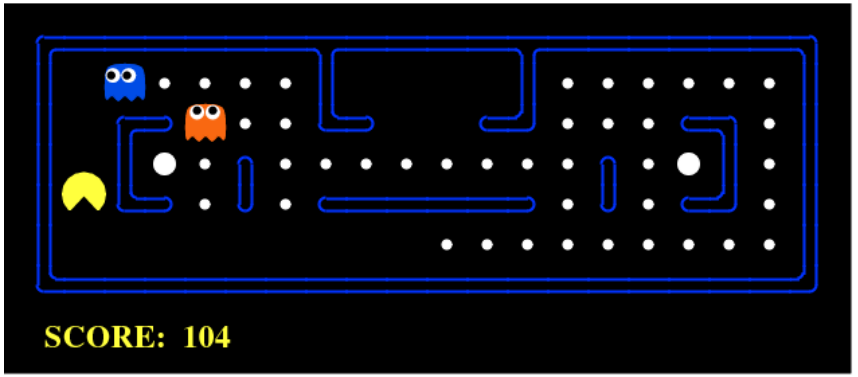


Domains and Applications

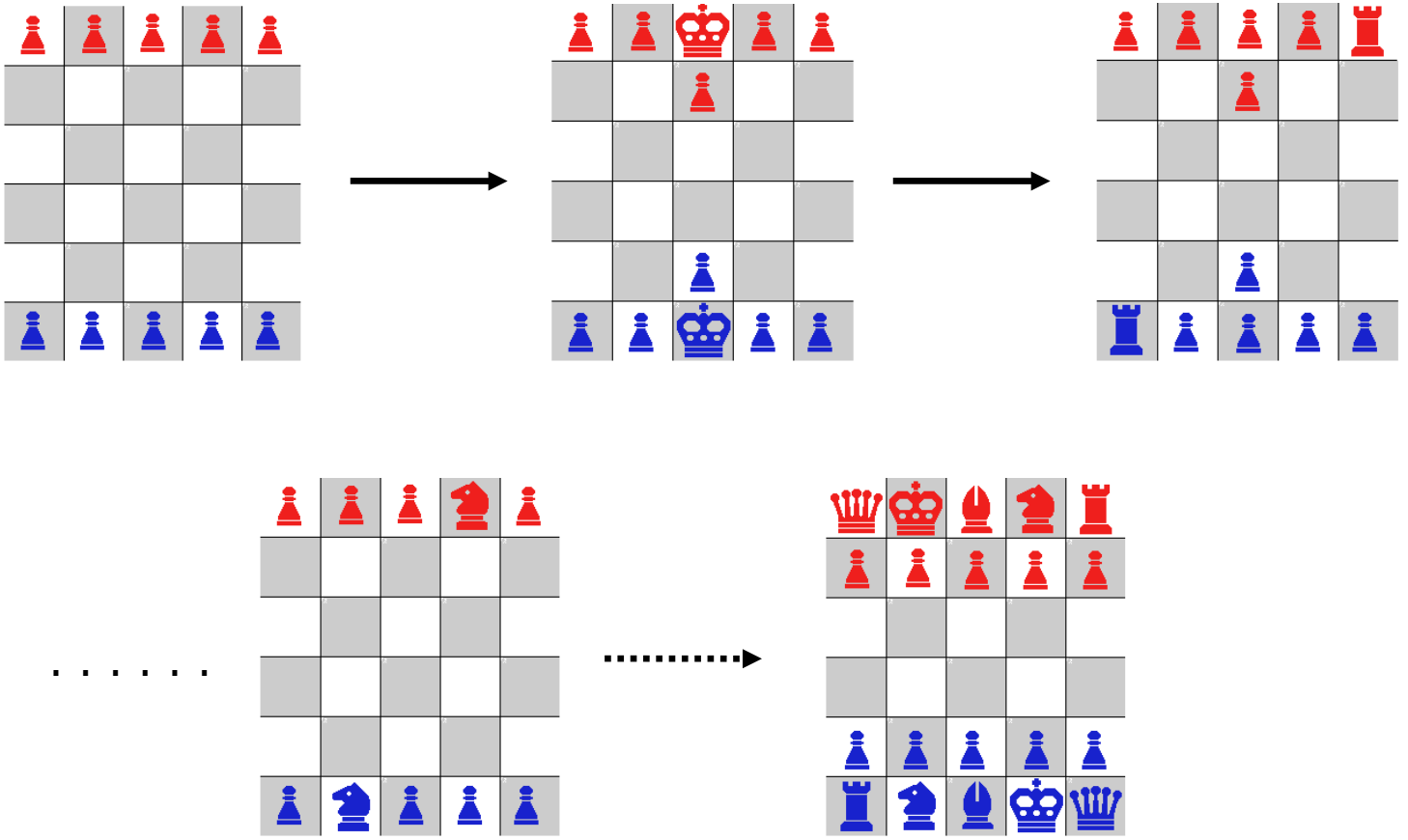








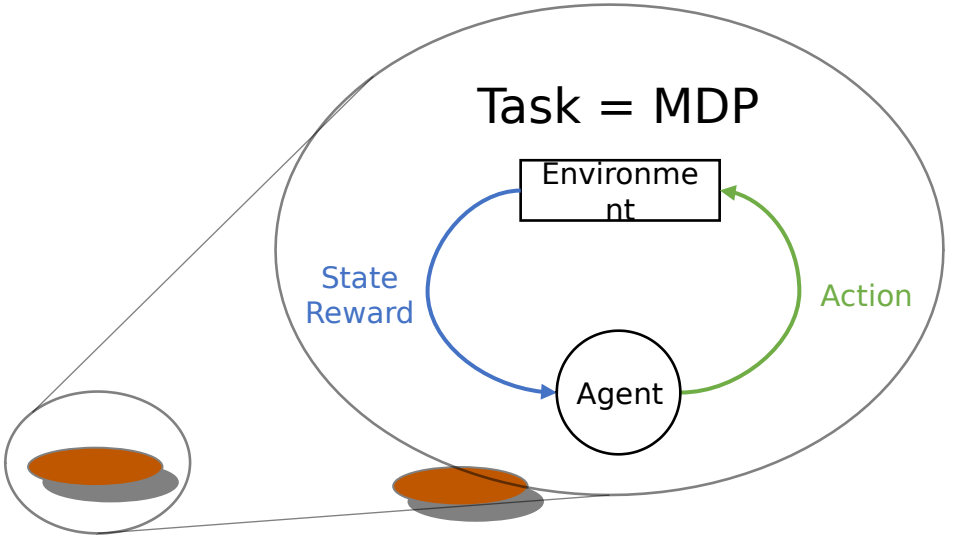
Curriculum Learning



Example QuickChess game variants

The Curriculum Learning Problem

Task Creation



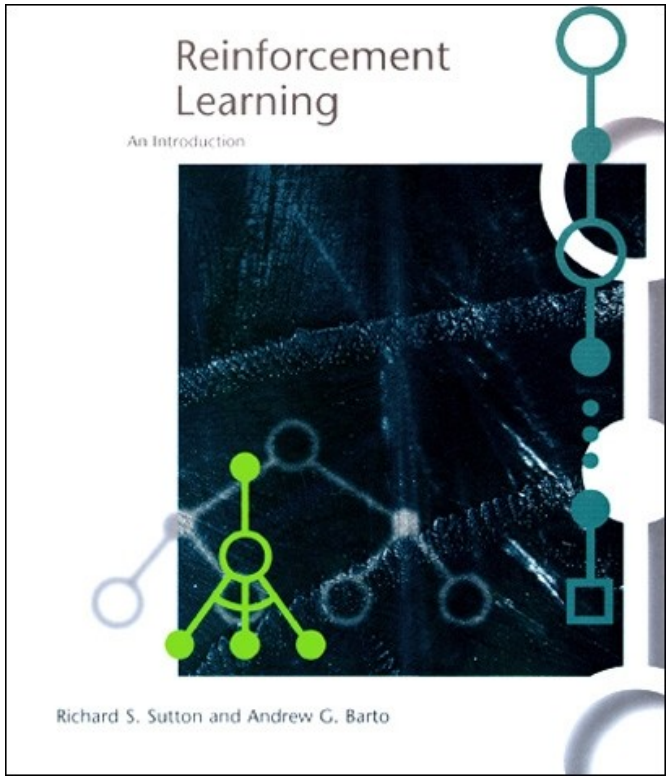
Target task

Sequencing

Transfer Learning

[Narverkar *et al* 2016]

Textbook



The authors have made the book available:
http://incompleteideas.net/book/bookdraft2017nov5.pdf

Course Organization

- Taught as a seminar: students take turns presenting the readings
- Will cover both theory and practice
- Final projects – you will complete a project in which you ask (and then answer) a relevant RL research question