# CMPUT 366: Intelligent Systems

Instructor: Adam White
University of Alberta
DeepMind Alberta

# Intelligent Systems

- To make persons, minds
- Introduction to the Science and Technology of Artificial Intelligence
    - touches on control theory, psychology, operations research, philosophy, and neuroscience
- A technical and conceptual foundation for understanding this large and complex set of issues

# Goals of Artificial Intelligence

- Scientific goal:
  - understand principles that make rational (intelligent) behavior possible, in natural or artificial systems.

- Engineering goal:
  - specify methods for design of useful, intelligent artifacts.

- Psychological goal:
  - understanding/modeling people
  - cognitive science

- Philosophical goal:
  - Understand what it means to be a person
  - Understand humanity's role in the universe

# The coming of artificial intelligence

- When people finally come to understand the principles of intelligence—what it is and how it works—well enough to design and create beings as intelligent as ourselves

- A fundamental goal for science, engineering, the humanities, …for all mankind

- It will change the way we work and play, our sense of self, life, and death, the goals we set for ourselves and for our societies

- It will lead to new beings and new ways of being, things inevitably *much <u>more powerful</u> than our current selves*

# AI is a great scientific prize

- cf. the discovery of DNA, the digital code of life, by Watson and Crick (1953)
- cf. Darwin’s discovery of evolution, how people are descendants of earlier forms of life (1860)
- cf. the splitting of the atom, by Hahn (1938)
    - leading to both atomic power and atomic bombs

# Discuss with your classmates

When will we understand the principles of intelligence well enough to create, using technology, artificial minds that rival our own in skill and generality?

Which of the following best represents *your* current views?

A. Never

B. Not during your lifetime

C. During your lifetime, but not before 2045

D. Before 2045

E. Before 2035

# Is human-level AI *possible*?

- If people are biological machines, then eventually we will reverse engineer them, and understand their workings
- Then, surely we can make improvements
  - with materials and technology not available to evolution
  - how could there not be something we can improve?
  - design can overcome local minima, make great strides, try things much faster than biology

# If AI is possible, then will it *eventually*, inevitably happen?

- No. Not if we destroy ourselves first
- If that doesn’t happen, then there will be strong, multi-incremental economic incentives pushing inexorably towards human and super-human AI
- It seems unlikely that they could be resisted
    - or successfully forbidden or controlled
    - there is too much value, too many independent actors

# Investment in AI is way up

- Google's prescient AI buying spree: Boston Dynamics, Nest, Deepmind Technologies, …
- New AI research labs at Facebook, Baidu, Allen Institute, Vicarious, Maluuba, DeepMind Alberta…
- Also enlarged corporate AI labs: Microsoft, Amazon, Adobe…
- Yahoo makes major investment in CMU machine learning department
- Many new AI startups getting venture capital
- New Canadian AI funding in Toronto, Montreal, and Edmonton
    - The Alberta Machine Intelligence Institute (AMII)

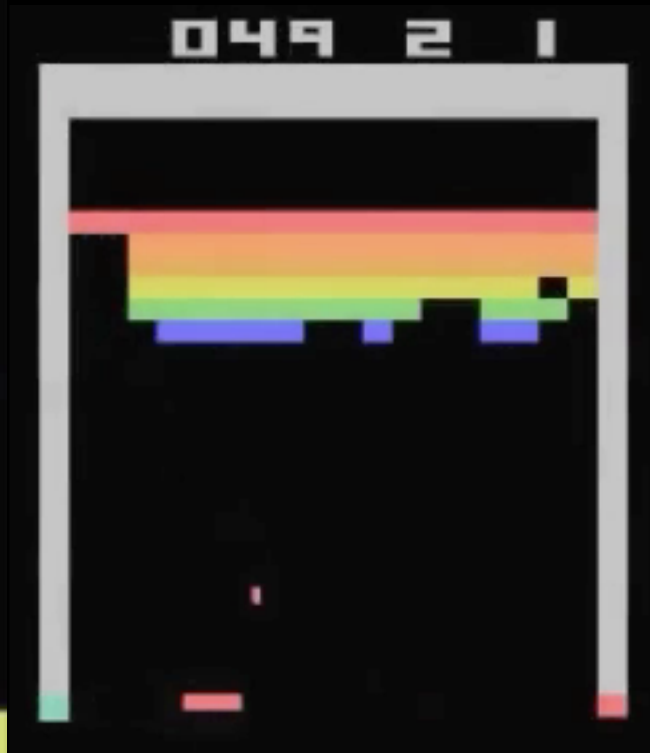
# Advances in AI abilities are coming faster;

## in the last 6 years:

- IBM's Watson beats the best human players of *Jeopardy!* (2011)
- Deep neural networks greatly improve the state of the art in speech recognition and computer vision (2012–)
- Google's self-driving car becomes a plausible reality (≈2013)
- Deepmind's DQN learns to play Atari games at the human level, from pixels, with no game-specific knowledge (≈2014, *Nature*)
- University of Alberta program solves Limit Poker (2015, Science),
  and then defeats professional players at No-limit Poker (2017, Science)
- Google Deepmind's AlphaGo defeats legendary Go player Lee Sedol (2016, Nature),
  and world champion Ke Jie (2017), vastly improving over all previous programs

# RL + Deep Learing Performance on Atari Games


Space Invaders

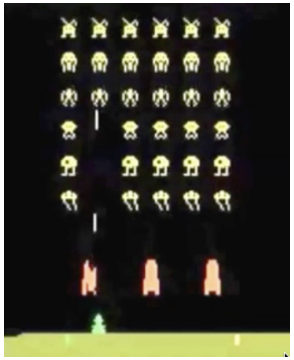



Breakout

Enduro

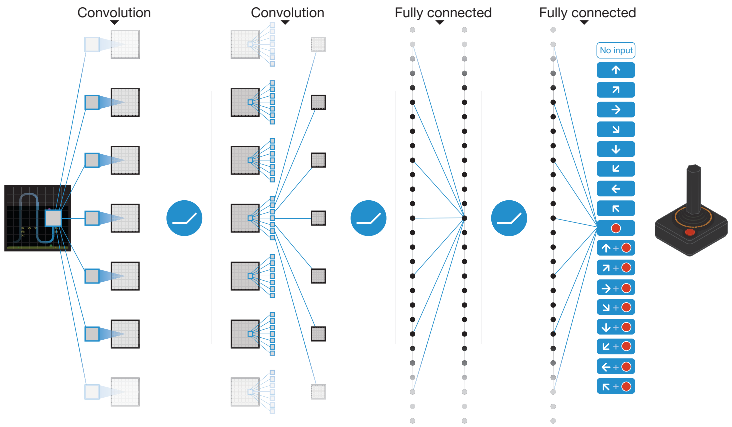
# RL + Deep Learning, applied to Classic Atari Games

Google Deepmind 2015, Bowling et al. 2012




- Learned to play 49 games for the Atari 2600 game console, without labels or human input, from self-play and the score alone

mapping raw screen pixels




to predictions of final score for each of 18 joystick actions

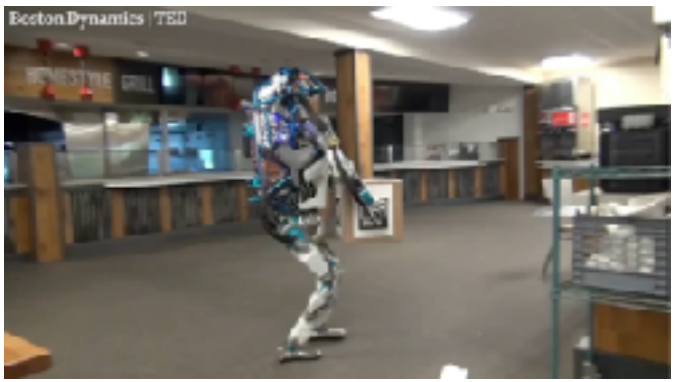
- Learned to play better than all previous algorithms and at human level for more than half the games

Same learning algorithm applied to all 49 games! w/o human tuning

# Cheap computation power drives progress in AI

- Deep learning algorithms are essentially the same as what was used in ‘80s
    - only now with larger computers (GPUs) and larger data sets
    - enabling today’s vastly improved speech recognition
- Similar impacts of computer power can be seen in recent years, and throughout AI’s history, in natural language processing, computer vision, and computer chess, Go, and other games

# Many fundamental research questions remain unresolved



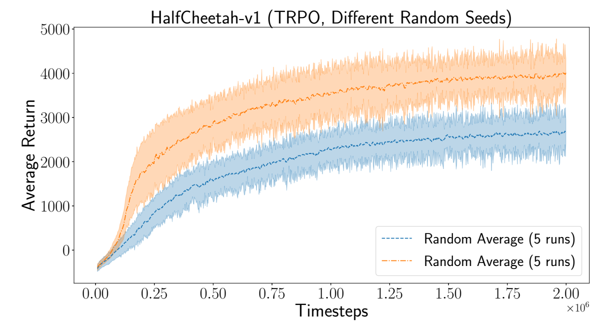




Figure 5: TRPO on HalfCheetah-v1 using the same hyperparameter configurations averaged over two sets of 5 different random seeds each. The average 2-sample $t$-test across entire training distribution resulted in $t = -9.0916$, $p = 0.0016$.

(Henderson et al, 2018)

| Game | ES | DQN w/ $\epsilon$-greedy | DQN w/ param noise |
|---|---|---|---|
| Alien | 994.0 | 1535.0 | **2070.0** |
| Amidar | 112.0 | 281.0 | **403.5** |
| BankHeist | 225.0 | 510.0 | **805.0** |
| BeamRider | 744.0 | **8184.0** | 7884.0 |
| Breakout | 9.5 | **406.0** | 390.5 |
| Enduro | 95.0 | 1094 | **1672.5** |
| Freeway | 31.0 | **32.0** | 31.5 |
| Frostbite | 370.0 | 250.0 | **1310.0** |
| Gravitar | **805.0** | 300.0 | 250.0 |
| MontezumaRevenge | 0.0 | 0.0 | **0.0** |
| Pitfall | 0.0 | -73.0 | -100.0 |
| Pong | **21.0** | **21.0** | 20.0 |
| PrivateEye | 100.0 | **133.0** | 100.0 |
| Qbert | 147.5 | **7625.0** | 7525.0 |
| Seaquest | 1390.0 | 8335.0 | **8920.0** |
| Solaris | **2090.0** | 720.0 | 400.0 |
| SpaceInvaders | 678.5 | 1000.0 | **1205.0** |
| Tutankham | 130.3 | 109.5 | **181.0** |
| Venture | **760.0** | 0 | 0 |
| WizardOfWor | **3480.0** | 2350.0 | 1850.0 |
| Zaxxon | 6380.0 | **8100.0** | 9050.0 |

(Plappert et al, 2017)

# AI is not like other sciences

- AI has Moore's law, an enabling technology racing alongside it, making the present special
- Moore's law is a slow fuse,
  leading to the greatest scientific and economic prize of all time
- So slow, so inevitable, yet so uncertain in timing
- The present is a special time for humanity, as we prepare for, wait for, and strive to create strong AI

# Algorithmic advances in Alberta

- World’s best computer games group for decades (see Bowling’s talk) including solving Poker
- Created the Atari games environment that our alumni, at Deepmind, used to show learning of human-level play
- Trained the AlphaGo team that beat the world Go champion
- World’s leading university in reinforcement learning algorithms, theory, and applications, including TD, MCTS
- ≈20 faculty members in AI

# Job opportunities in Alberta

- Huawei Edmonton Research lab
- Borealis AI
- Deepmind Alberta
- Several new labs and startups on the horizon

# Why are you here?

- What do you expect to learn?

For you, which of the following are essential abilities of an intelligent system that you would like to learn about (say in this course)?

The ability to:

A. sense and perceive the external world
B. choose actions that affect the world
C. use language and interact with other agents
D. predict the future
E. fool people into thinking that you are a person
F. have and achieve goals
G. reason symbolically, as in logic and mathematics
H. reason in advance about courses of action before picking the best
I. learn by trying things out and subsequently picking the best
J. have emotions, pleasure and pain
K. other?

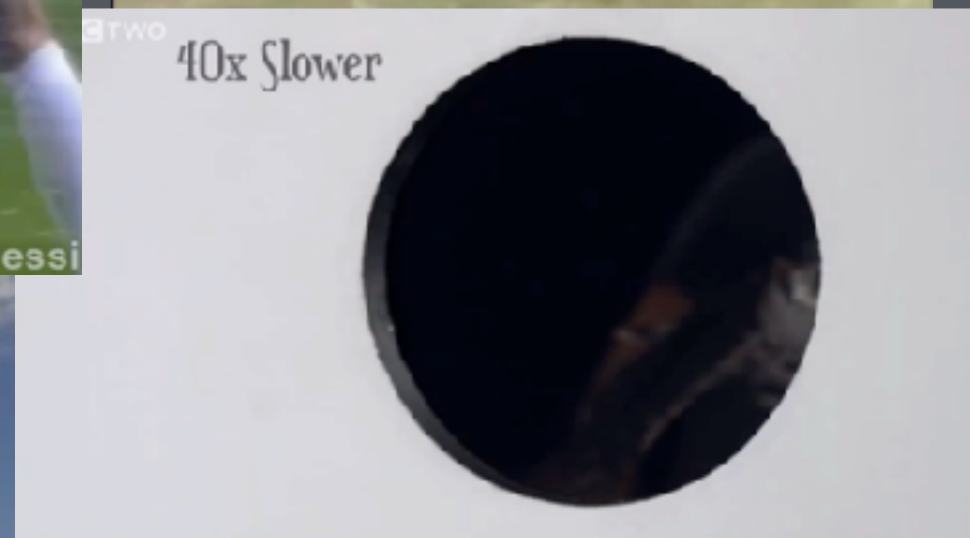
For you, which of the following are essential abilities of an intelligent system that you would like to learn about (say in this course)?

The ability to:

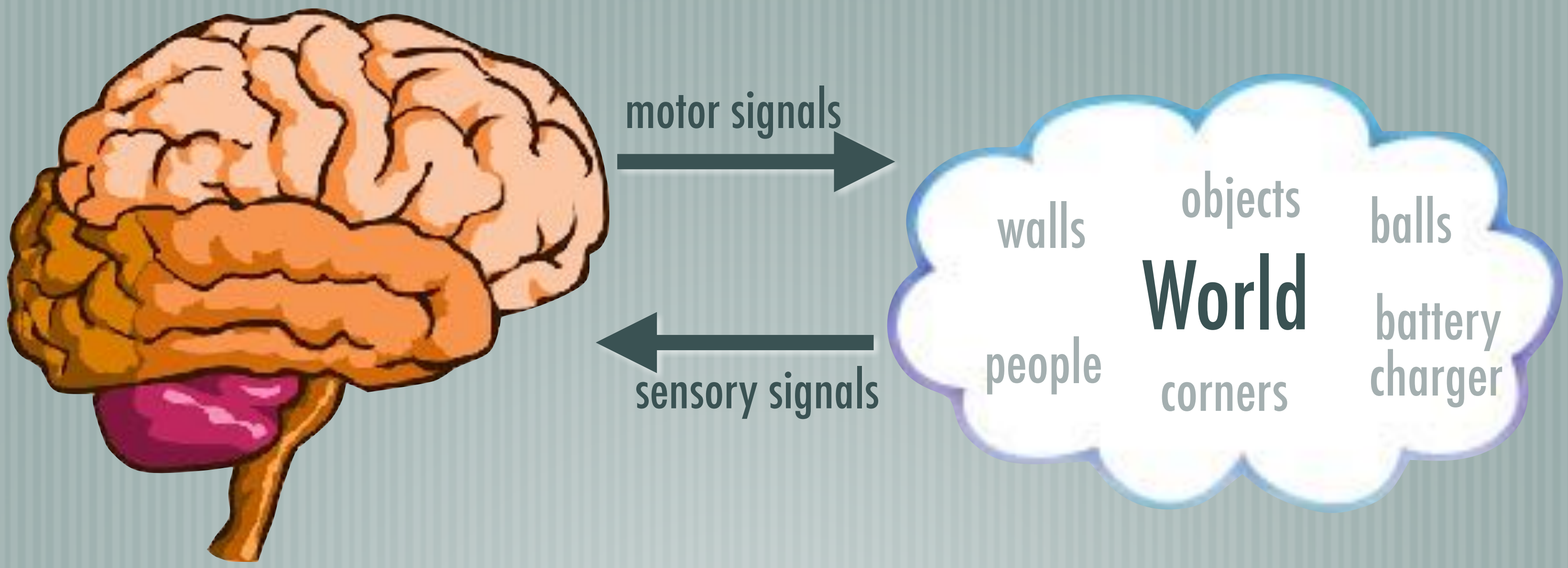
A. sense and perceive the external world
B. choose actions that affect the world
C. use language and interact with other agents
D. predict the future
E. fool people into thinking that you are a person
F. have and achieve goals
G. reason symbolically, as in logic and mathematics
H. reason in advance about courses of action before picking the best
I. learn by trying things out and subsequently picking the best
J. have emotions, pleasure and pain
K. other?

the mind's first responsibility is
real-time sensorimotor information processing

- Perception, action, & anticipation
  - as fast and reactive as possible

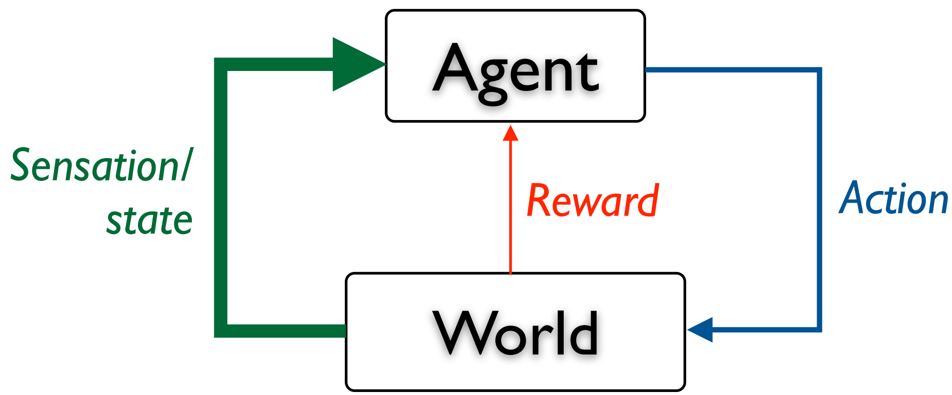



# Minds are sensori-motor information processors




the mind's job is to predict and control its sensory signals

# Reinforcement learning is *more autonomous learning*

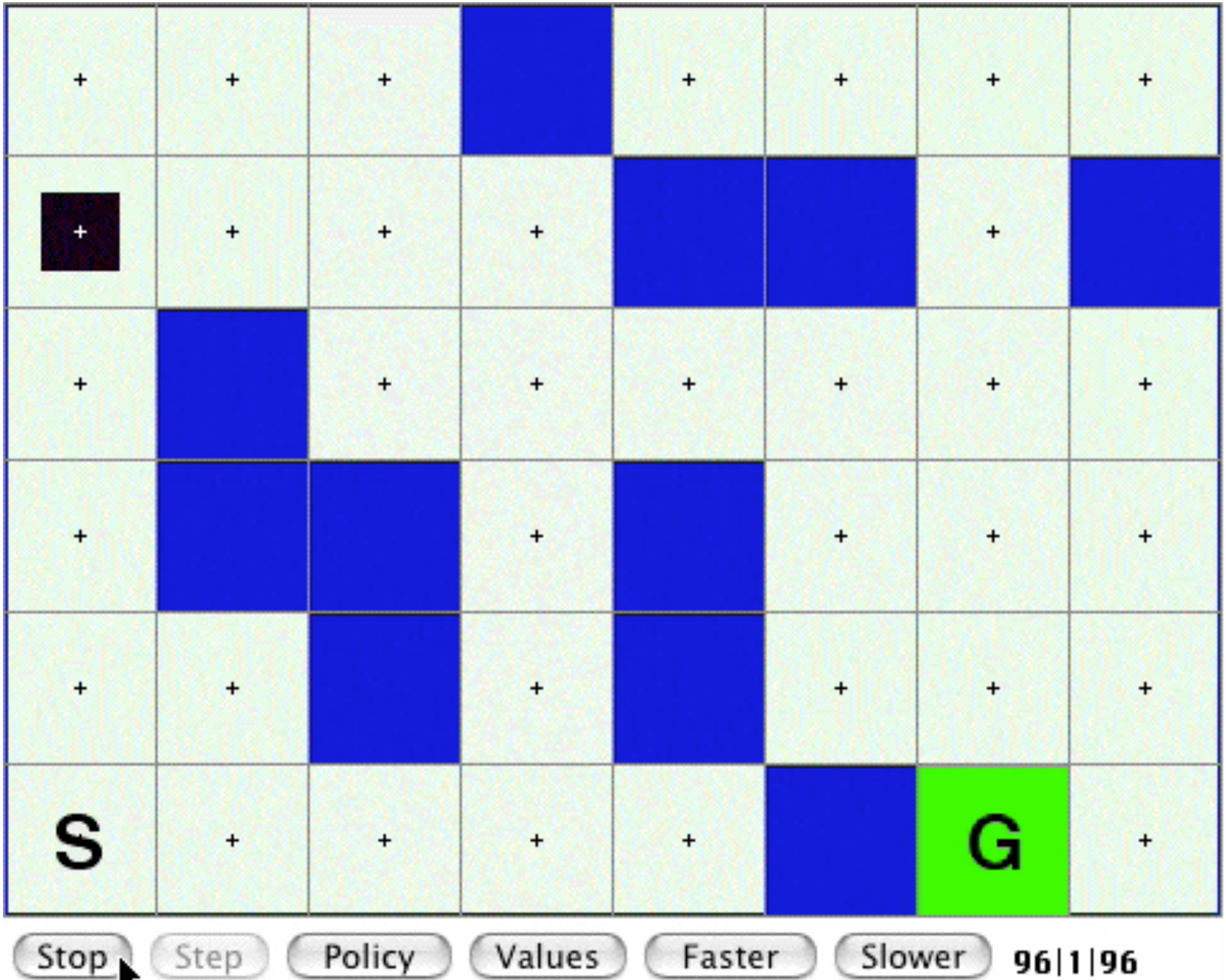



- Learning that requires less input from people
- AI that can learn for itself, during its normal operation

# Kinds of Reinforcement Learning

- *Model-free RL* — learning what to do by trying different things, remembering the best
- *Model-based RL* — learning how the world works, then computing what to do
- *Prediction learning* — learning what will happen next
- *Representation learning* — learning the features of state that generalize well
- *RL architectures* — putting it all together with massive computation

# Model-based RL: GridWorld Example



# Course Overview

- Main Topics:
  - Learning (by trial and error)
  - Planning (search, reason, thought, cognition)
  - Prediction (evaluation functions, knowledge)
  - Control (action selection, decision making)
- Recurring issues:
  - Demystifying the illusion of intelligence

# Order of Presentation

- Control: Bandits and Markov decision processes
- Stochastic planning (dynamic programming)
- Model-free reinforcement learning
- Planning with a learned model
- Learning with approximations

# High-level view

- Bandits and online learning (ch2):
  - **formalizing** a problem and discussing **solution methods**
  - A miniature version of the entire course
- Markov Decision Processes (ch3):
  - Our **formalization** of reinforcement learning and AI…no solution methods here
  - Students usually get impatient here

# High-level view (2)

- Classic MDP solution methods (ch's 4,5,6):
  - Dynamic programming (what if you knew how the world worked?)
  - Monte Carlo (what if you only learned from interaction)
  - Temporal difference learning (strengths of both)
- More advanced stuff:
  - Multi-step methods, planning

# High-level view (3)

- Everything up to and including chapter is tabular solution methods:
  - The foundation of modern RL
- In chapters 9, 10, 12 cover approximate solution methods:
  - Function approximation (including Neural Nets)
- The foundations established in chapter 3-8 will largely transfer to the function approximation case

# RL courses at UofA

- CMPUT 366: `Reinforcement learning I'(undergrad intro)
- CMPUT 609: Reinforcement Learning II (graduate, advanced topics)
  - Assumes background in RL (similar to 366)
- CMPUT 607: Reinforcement Learning in practice (applying RL algorithms and concepts to robots)
- CMPUT 659: RL, an optimization perspective

# Instruction Team

- Prof: Adam White
- TAs (grad students doing research in AI)
    - Andrew Jacobsen
    - Arash Pourzarabi
    - Han Wang
    - Kris De Asis
    - Niko Yasui
    - Yi Wan
    - Raksha Kumaraswamy
    - Paniz Behboudian

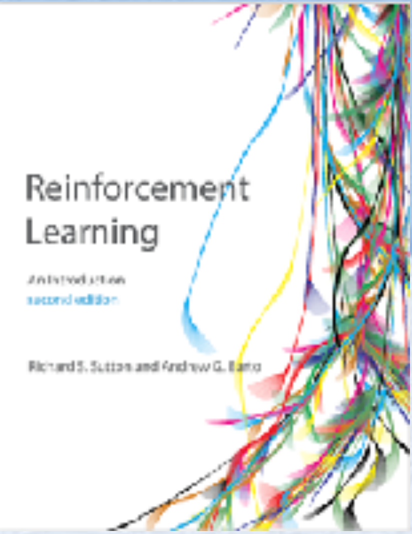
# CMPUT 366: Schedule of Classes and Assignments (subject to change)

| class num | date | lecture topic | Reading assignment (in advance) | Assignment due |
|---|---|---|---|---|
| **1** | Tue, Sep 4, 2018 | Artificial Intelligence; reasons for taking the course | Sutton & Barto Chapter 1 | |
| **2** | Thu, Sep 7, 2017 | Bandit problems | Sutton & Barto Chapter 2 (Section 2.7 optional) | |
| | **Mon, Sep 10, 2018** | Probability, Stats & Python tutorial | probabilities-expectations.pdf (in the dropbox) | |
| **3** | Tue, Sep 11, 2018 | Bandit problems plus RL examples | Sutton & Barto Chapter 2 (Section 2.7 optional) | |
| | **Wed, Sep 12, 2018** | Probability, Stats & Python tutorial | probabilities-expectations.pdf (in the dropbox) | |
| **4** | Thu, Sep 14, 2017 | Defining "Intelligent Systems" | Read the definition given for artificial intelligence in Wikipedia and in the Nilsson book on p13; google for and read "John McCarthy basic questions", | |
| | Mon, Sep 17, 2018 | A1 lab | | |
| **5** | Tue, Sep 18, 2018 | Markov decision problems | Sutton & Barto Chapter 3 thru 3.5 | A1 |
| | Wed, Sep 19, 2018 | RL-Glue tutorial | | |
| **6** | Thu, Sep 21, 2017 | Returns, value functions | Rest of Sutton & Barto Chapter 3 | |
| | Mon, Sep 24, 2018 | RL-Glue tutorial | | |
| **7** | Tue, Sep 25, 2018 | Bellman Equations | Sutton & Barto Summary of Notation, Sutton & Barto Section 4.1 | Thought question #1 |
| | Wed, Sep 26, 2018 | Lab | | |
| **8** | Thu, Sep 28, 2017 | Dynamic programming (planning) | Sutton & Barto Rest of Chapter 4 | |
| | Mon, Oct 1, 2018 | A2 lab | | |
| **9** | Tue, Oct 2, 2018 | Monte Carlo Learning | Sutton & Barto Chapter 5 thru 5.4 | A2 |
| | Wed, Oct 3, 2018 | A3 lab | | |
| **10** | Thu, Oct 5, 2017 | Off-policy Monte Carlo Learning | Sutton & Barto rest of Chapter 5 (except Sections 5.8, 5.9) | |
| | Mon, Oct 8, 2018 | **No lab (holiday)** | | |
| **11** | Tue, Oct 9, 2018 | Temporal-difference learning | Sutton & Barto Chapter 6 thru Section 6.3 | |
| | Wed, Oct 10, 2018 | A3 lab | | |
| **12** | Thu, Oct 12, 2017 | Temporal-difference learning | Sutton & Barto rest of Chapter 6 | |
| | Mon, Oct 15, 2018 | A3 lab | | |
| **13** | Tue, Oct 16, 2018 | Multi-step bootstrapping | Sutton & Barto Chapter 7 except Sections 7.4-6 | A3 |
| | Wed, Oct 17, 2018 | Midterm Review & practice questions | | |
| **14** | Thu, Oct 19, 2017 | Review | Sutton & Barto Chapters 2-7 | |
| | Mon, Oct 22, 2018 | Midterm Review & practice questions | | |

# Course Information

- Course eclass page
  - some official information, schedule
  - discussion list!
- Course Google Drive Folder (see eclass page for link)
  - assignments, slides, readings, test prep
- Lab is on Monday & Wed (ETLC E1 001), 5-7:50
  - a good place to do your assignments

# Textbooks




- Readings will be from web sources plus the following two textbooks (both of which are available as online electronically and open-access):
  - *Reinforcement Learning: An Introduction*, by R Sutton and A Barto, MIT Press. (bookdraft2018.pdf)
    - printed copies available at the bookstore—I hope!
  - *The Quest for AI*, by N Nilsson, Cambridge, 2010 (pdf)

# Evaluation

- Final Exam – 40%
- Midterm – 15%
- Assignment every two weeks (on tues.) – 40%
  - due in gradescope by due date at 11:59pm
  - bonus questions for extra credit
- 3 Thought questions – 5%
  - due in gradescope (no handwritten answers)
  - First one worth 1%, next two worth 2%

# Grades (366)

- Will be assigned on an *absolute scale* based on weighted % of points received:

| | |
|---|---|
| A+ | [90 -- 100]% |
| A | [85 -- 90]% |
| A- | [80 -- 85)% |
| B+ | [75 -- 80)% |
| B | [70 -- 75)% |
| B- | [65 -- 70)% |
| C+ | [60 -- 65)% |
| C | [55 -- 60)% |
| C- | [50 -- 55)% |
| D+ | [45 -- 50)% |
| D | [40 -- 45)% |
| F | [0 -- 40)% |

# Collaboration

- Working together to solve the problems is encouraged
- But you must write-up your answers individually
- You must acknowledge all the people you talked with in solving the problems

# What is Plagiarism

- Taking things from others and passing it off as your own work without credit

# Test time: are these ok?

- Writing down answers to assignments in a group?
- Getting a tutor to help write your code?
- Letting your friend look at your code or assignment question?
- Searching for and using assignment solutions from the internet?
- Not indicating on your assignment who you talked with?
- Discussing ideas without writing anything down?

# Policies on Integrity

- Cheating is reported to university whereupon it is out of our hands
- Possible consequences:
  - A mark of 0 for assignment
  - A mark of 0 for the course
  - A permanent note on student record
  - Suspension / Expulsion from university

# Labs

- Mondays 5-7:50, here and Wednesdays 5-7:50 in E1 001
- Next Mon & Wed: 1. Set up Python 3 & style expectations, 2. Review math (probabilities and expectations)
- In general:
  - get practice with problems like those on the assignments and exams
  - discuss the assignments and get questions answered
  - a great time and place to do the assignments!!

# Contacting us…

- Use the course discussion feature on eclass
  - Start a discussion
    - Read by prof and TAs
    - Remember: public!
- Meeting w/prof during office hours or by arrangement
- Go to the lab!

# Prerequisites

- Some comfort or interest in thinking abstractly and with mathematics
- Elementary statistics, probability theory
  - conditional expectations of random variables
  - there will be a lab session devoted to a tutorial review of basic probability
- Basic linear algebra: vectors, vector equations, gradients
- Programming skills (Python)
  - If Python is a problem, start working on it now

# for next time...

- Read Chapters 1 & 2 of Sutton & Barto text (online)
  - Read Chapter 2 fully
  - Use your judgement on Chapter 1

# Academic Integrity

- The University of Alberta is committed to the highest standards of academic integrity and honesty. Students are expected to be familiar with these standards regarding academic honesty and to uphold the policies of the University in this respect. Students are particularly urged to familiarize themselves with the provisions of the Code of Student Behavior (online at www.ualberta.ca/secretariat/appeals.htm) and avoid any behavior which could potentially result in suspicions of cheating, plagiarism, misrepresentation of facts and/or participation in an offence. Academic dishonesty is a serious offence and can result in suspension or expulsion from the University.

# AI Seminar !!!

- http://www.cs.ualberta.ca/~ai/cal/
- Friday noons, CSC 3-33, FREE PIZZA!
- Neat topics, great speakers
- For mailing list of announcements, google “mailman ualberta”, then sign up for ai-seminar