CMPUT 366: Intelligent Systems and CMPUT 609: Reinforcement Learning & Artificial Intelligence

Instructors: Rich Sutton and 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

Intelligence is the ability to achieve goals

- “Intelligence is the most powerful phenomena in the universe” —Ray Kurzweil, c 2000
    - The phenomena is that there are systems in the universe that are well thought of as goal-seeking systems
- What is a goal-seeking system?
    - “Constant ends from variable means is the hallmark of mind” —William James, c 1890
    - a system that is better understood in terms of *outcomes* than in terms of *mechanisms*

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

- But it is also of significance beyond our species, beyond history

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

Milestones in the development of life on Earth

	year	Milestone	
The Age of Stars	14Bya	Big bang	
	4.5Bya	formation of the earth and solar system	
The Age of Replicators	3.7Bya	origin of life on earth (formation of first replicators)	Self-replicated things most prominent
		DNA and RNA	
	1.1Bya	sexual reproduction	
		multi-cellular organisms	
		nervous systems	
	1Mya	humans	
		culture	
	100Kya	language	
	10Kya	agriculture, metal tools	
	5Kya	written language	
The Age of Design	200ya	industrial revolution	Designed things most prominent
		technology	
	70ya	computers	
		nanotechnology	
	?	artificial intelligence	
		super-intelligence	
		…	

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

Socrative.com, Room 568225

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

Yes

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

Very probably, say 90%

When will human-level AI first be created?

- No one knows of course; we can make an educated guess about the probability distribution:
  - 25% chance by 2030
  - 50% chance by 2040
  - 10% chance never
- Certainly a significant chance within all of our expected lifetimes
  - We should take the possibility into account in our career plans

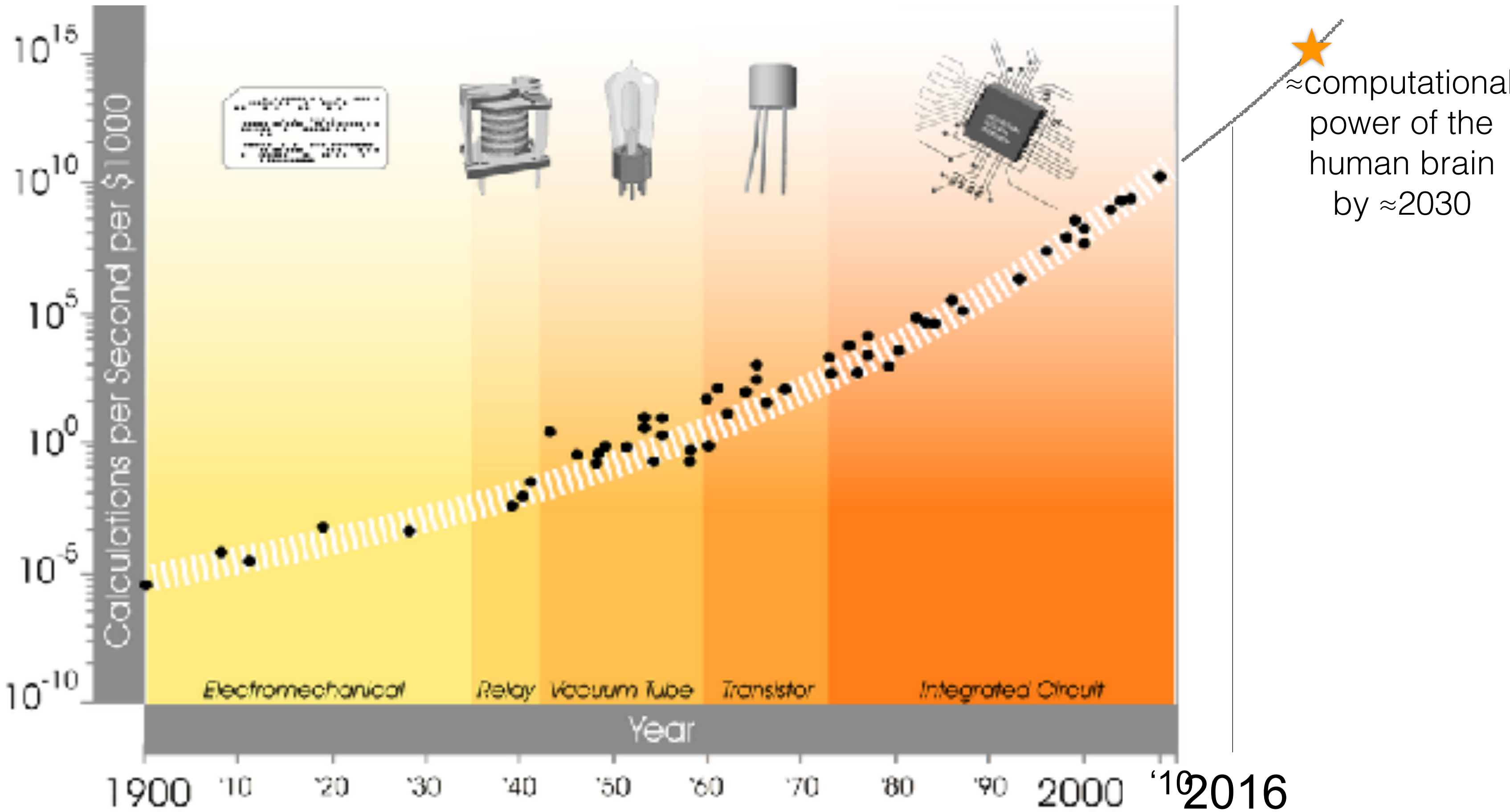
Investment in AI is way up

- 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

The 2nd industrial revolution

- The 1st industrial revolution was the *physical power* of machines substituting for that of people
- The 2nd industrial revolution is the *computational power* of machines substituting for that of people
  - Computation for perception, motor control, prediction, decision making, optimization, search
  - Until now, people have been our cheapest source of computation
  - But now our machines are starting to provide greater, cheaper computation

The computational revolution



Kurzweil AI

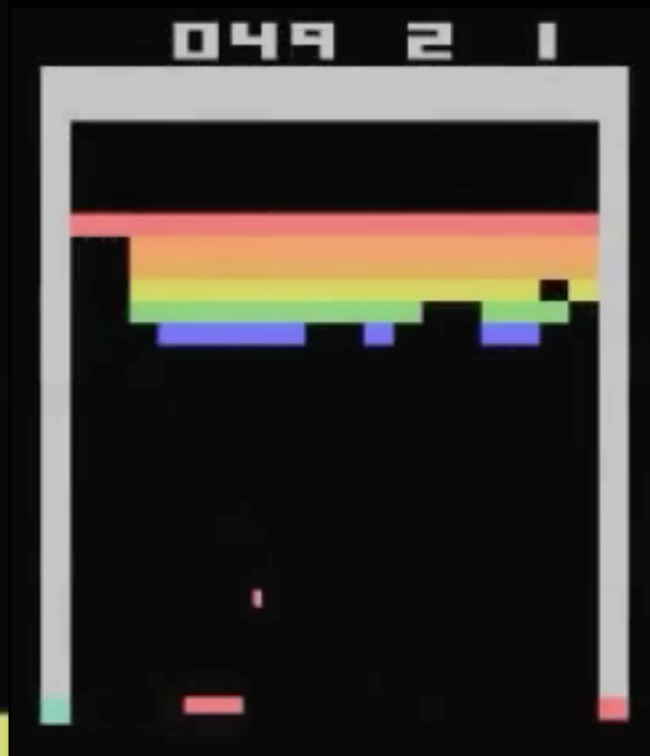
Advances in AI abilities are coming faster;

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

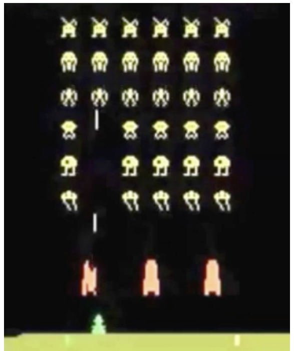


Breakout

Enduro

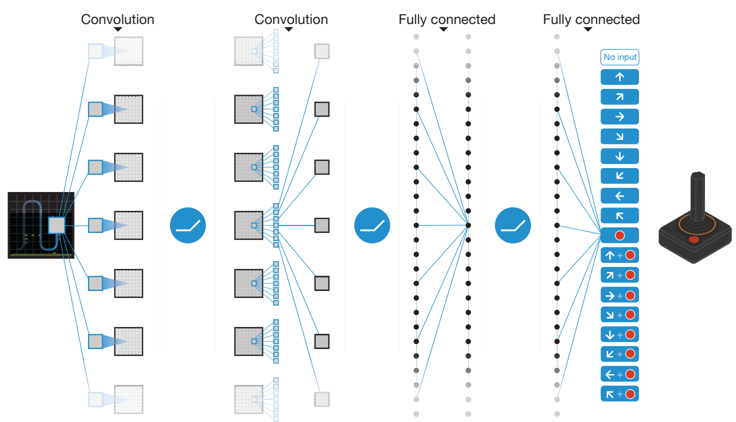
RL + Deep Learning, applied to Classic Atari Games

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

Advances in AI abilities are coming faster;
in the last 5 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's Cepheus solves Poker (2015, *Science*)
- Google Deepmind's AlphaGo defeats the world Go champion, vastly improving over all previous programs (2016)

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

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
- $\approx$20 faculty members in AI

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:

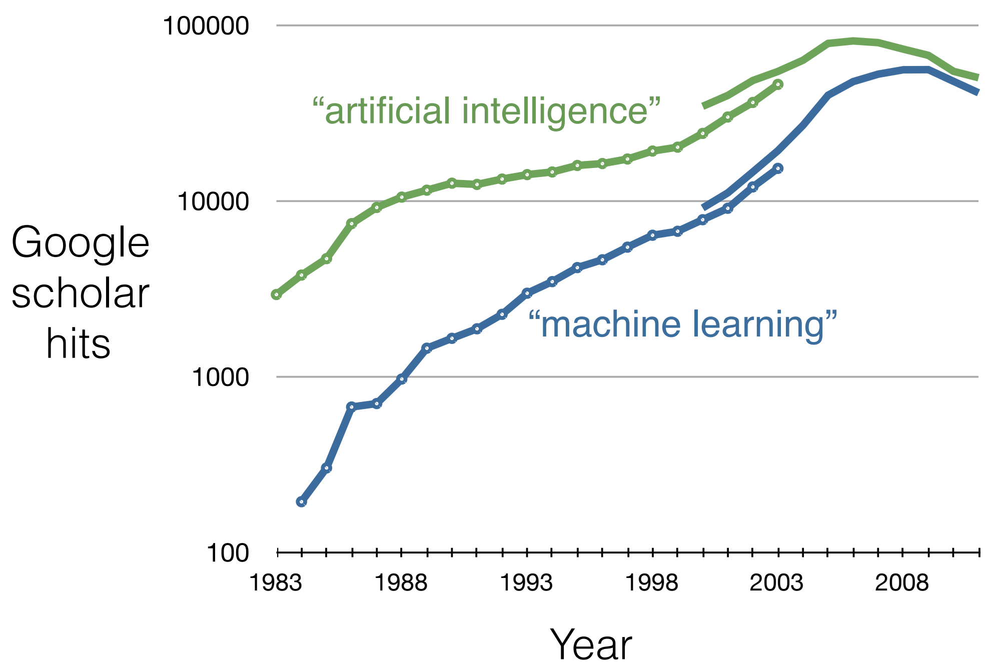
A. sense and perceive the external world
B. choose actions that affect the world
C. use language and interact with other agents
D. predict the future
E. fool people into thinking that you are a person
F. have and achieve goals
G. reason symbolically, as in logic and mathematics
H. reason in advance about 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?

Good Old-fashioned AI (GOFAI) and Modern Probabilistic AI

- AI was originally based more on deterministic symbolic logic, human intuition about thinking, and hand-crafted knowledge

- Over decades AI became more numeric, statistical, and based on data (learning)

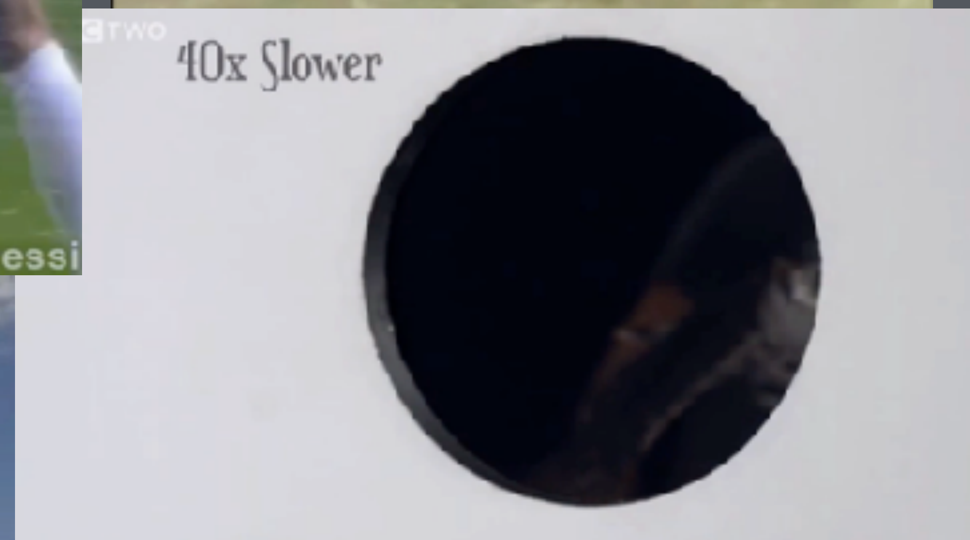
- And also much more integrated with engineering fields: statistics, decision theory, control theory, operations research, robotics computer science

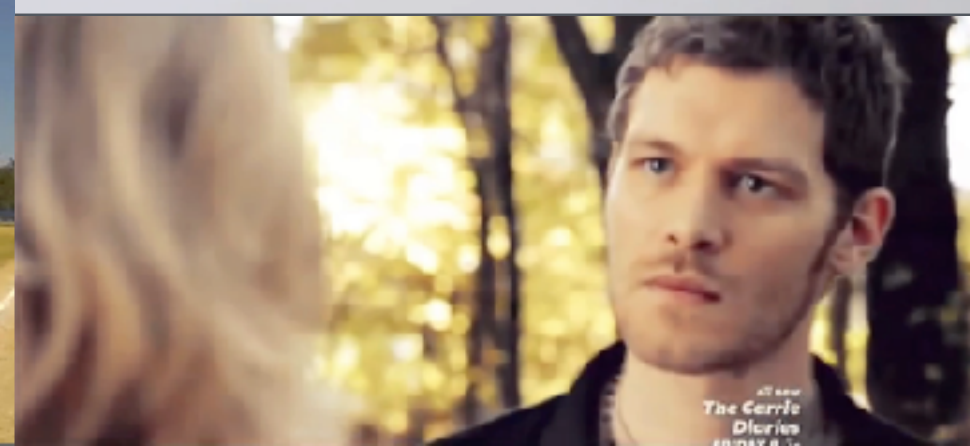


- Substantial convergence *and* divergence, with tensions and turf issues in both cases

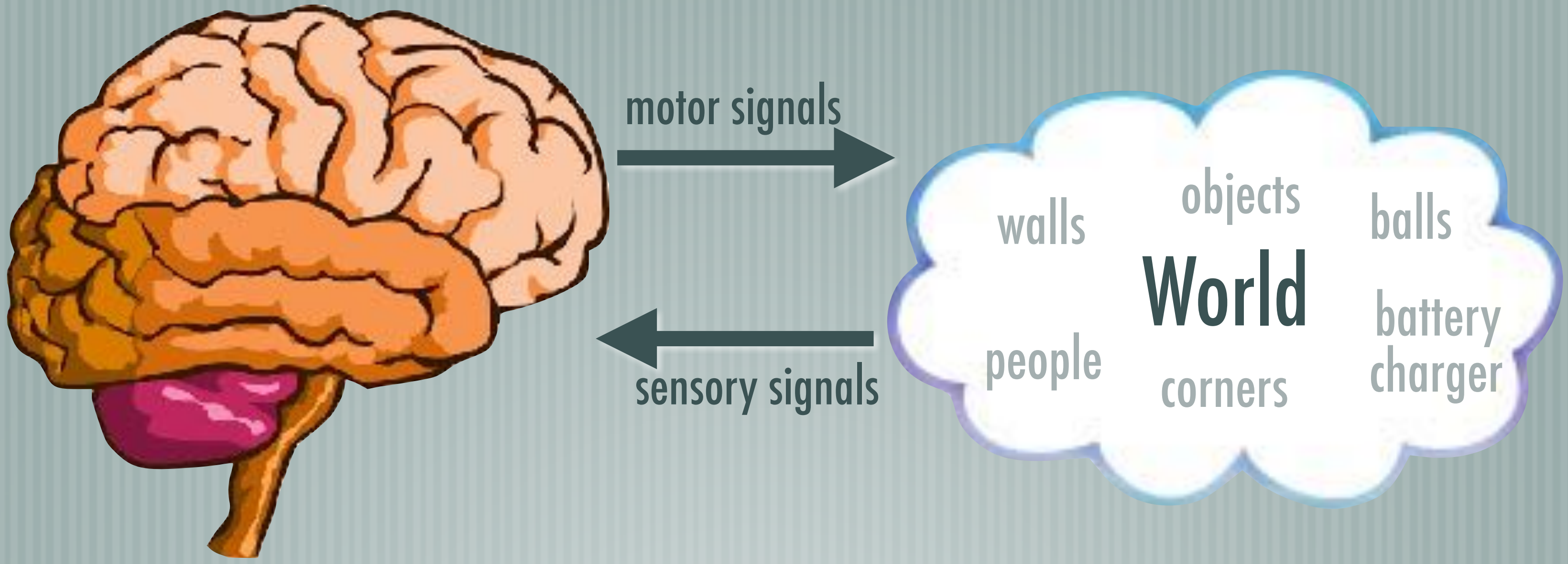
the mind's first responsibility is
real-time sensorimotor information processing

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



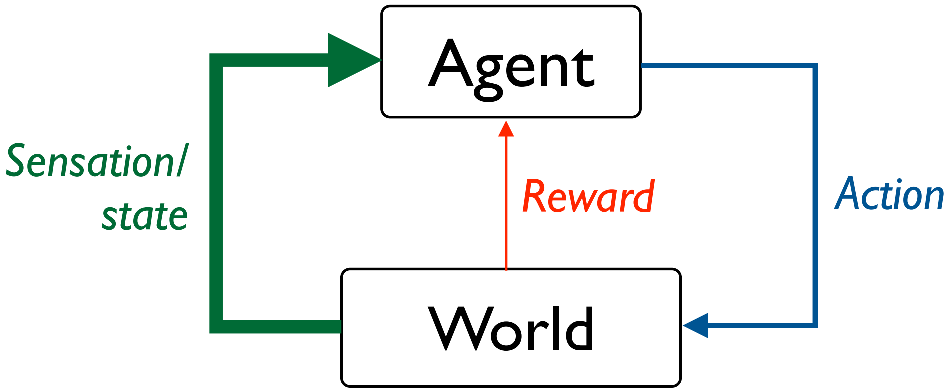


Minds are sensori-motor information processors



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

Reinforcement learning is *more autonomous learning*

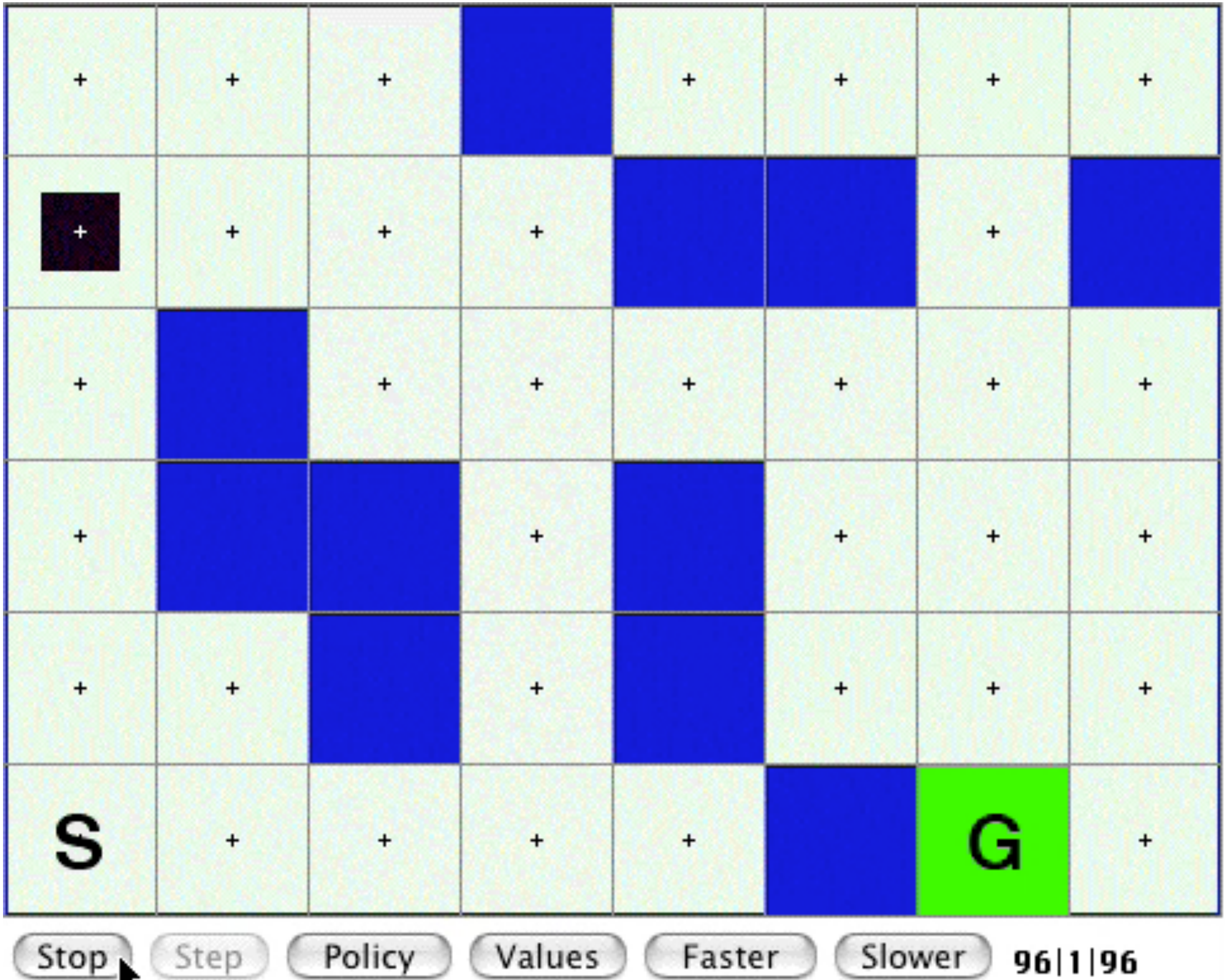


- Learning that requires less input from people
- AI that can learn for itself, during its normal operation

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
  - Purpose (goals, reward) vs Mechanism

Model-based RL: GridWorld Example



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

Instruction Team

- Profs: Adam White, Rich Sutton
- TAs (grad students doing research in AI)
  - Mohammad Ajallooeian
  - ...

CMPUT 366: Schedule of Classes and Assignments

class num	date	lecture topic	Reading assignment (in advance if possible)	Assignment due
1	Tue, Sep 5, 2017	The Magic of Artificial Intelligence; reasons for taking the course	Read section 1 of the Wikipedia entry for “the technological singularity”; see also Vinge2010 (http://www-rohan.sdsu.edu/faculty/vinge/misc/iaai10/) and Moravec1998 (http://www.transhumanist.com/volume1/moravec.htm)	
2	Thu, Sep 7, 2017	Bandit problems	Sutton & Barto Chapters 1 and 2 (Section 2.7 optional)	
	Mon, Sep 11, 2017	Probability, Python review/setup, textbook purchase	probabilities-expectations.pdf (in the dropbox)	
3	Tue, Sep 12, 2017	Bandit problems plus RL examples	Sutton & Barto Chapter 2 (Section 2.7 optional)	T1
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 18, 2017	A1 lab		
5	Tue, Sep 19, 2017	Markov decision problems	Sutton & Barto Chapter 3 first part TBD	A1, T2
6	Thu, Sep 21, 2017	Returns, value functions	Rest of Sutton & Barto Chapter 3	
	Mon, Sep 25, 2017	A2 lab		
7	Tue, Sep 26, 2017	Bellman Equations	Sutton & Barto Summary of Notation, Sutton & Barto Section 4.1	T3
8	Thu, Sep 28, 2017	Dynamic programming (planning)	Sutton & Barto Rest of Chapter 4	A2
	Mon, Oct 2, 2017	Tutoring lab		
9	Tue, Oct 3, 2017	Monte Carlo Learning	Sutton & Barto Chapter 5 thru 5.4	T5
10	Thu, Oct 5, 2017	Off-policy Monte Carlo Learning	Sutton & Barto rest of Chapter 5 (except Sections 5.8, 5.9)	
	Fri, Oct 6, 2017	Special Lab A3		
	Mon, Oct 9, 2017	No lab (holiday)		
11	Tue, Oct 10, 2017	Temporal-difference learning	Sutton & Barto Chapter 6 thru Section 6.3	A3, T6
12	Thu, Oct 12, 2017	Temporal-difference learning	Sutton & Barto rest of Chapter 6	
	Mon, Oct 16, 2017	A4 lab		
13	Tue, Oct 17, 2017	Multi-step bootstrapping	Sutton & Barto Chapter 7 except Sections 7.4-6	T7
14	Thu, Oct 19, 2017	Review	Sutton & Barto Chapters 2-7	A4
	Mon, Oct 23, 2017	Tutoring lab		
15	Tue, Oct 24, 2017	Midterm Exam	No new reading	
16	Thu, Oct 26, 2017	Models and planning	Sutton & Barto Chapter 8 thru Section 8.3	T8
	Mon, Oct 30, 2017	A5 lab		

Course Information

- Course Moodle page
  - some official information
  - discussion list!
- Course Google Drive Folder (see moodle page for link)
  - schedule, assignments, slides, projects
- Lab is on Monday, 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. (book366.pdf, book609.pdf)
        - we will use the in-progress, online 2nd edition
        - printed copies available at first lab— $20 exact
    - *The Quest for AI*, by N Nilsson, Cambridge, 2010 (pdf)

Evaluation

- Final Exam – 40%
- Midterm – 15%
- ≈1 Assignment every other week – 40%
  - due in gradescope by classtime
  - bonus questions required for 609, extra credit for 366
- Reading-Writing Exercise – 5%
  - due in gradescope by classtime for full credit

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
- You must completely understand and be able to justify your answers
- I reserve the right to check your understanding and possibly revise your marks

Labs

- Mondays 5-7:50, here
- This Monday: 1. Get books, 2. Set up Python 2.7 & style expectations, 3. 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
    - find a partner for the programming projects
    - a great time and place to do the assignments!!

Contacting us…

- Use the course discussion feature on the moodle page
  - Start a discussion
    - Read by prof and TAs
    - Remember: public!
- Meeting w/profs, TAs: at office hours or by arrangement

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
- Basic programming skills (Python)
  - If Python is a problem, choose a partner who is already comfortable with Python

for next time...

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

Policies on Integrity

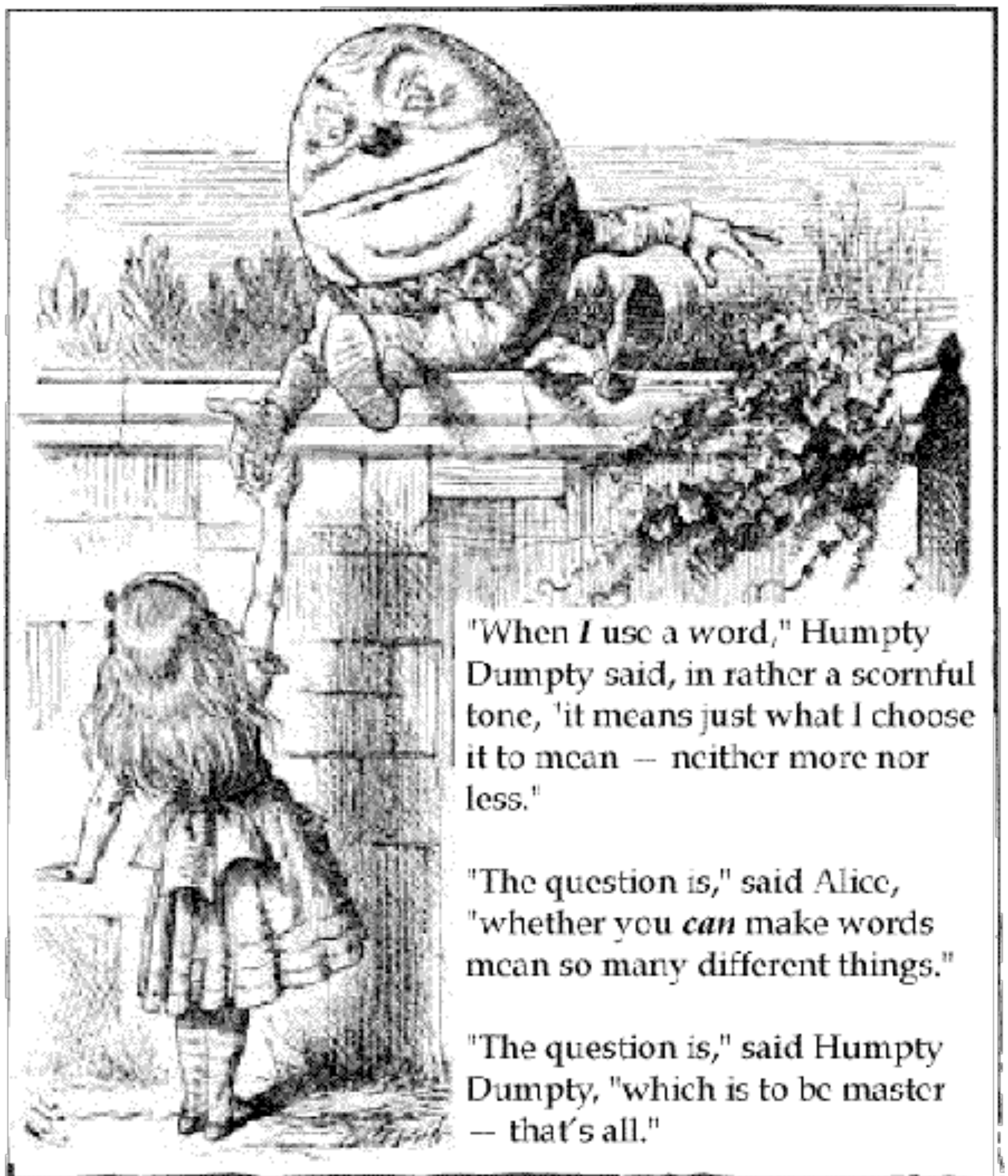
- Do not cheat on assignments:
  Discuss only general approaches to problem
- Do not take written notes on other's work
- Respect the lab environment. Do not:
  - Interfere with operation of computing system
  - Interfere with other's files
  - Change another's password
  - Copy another's program
  - etc.
- 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

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



From "Through the Looking Glass" by Lewis Carroll, 1871

A new view of the AI problem

and its implications for (and against) solution methods

- Minds are real-time information processors interacting with a firehose of data from a complex and arbitrary world
    - we must find *scalable* and *general* methods, to *learn* arbitrary stuff (no domain knowledge, no taking advantage of structure)
- We have immense computational resources, but it’s never enough; the complexity of the world is <u>always vastly greater</u>
    - we seek *computationally frugal* methods for finding *approximate* solutions (optimality is a distraction; relying on it is untenable)
- We have immense data, but not labeled examples
    - we must be able to learn from *unsupervised* interaction with the world, a.k.a. *self-labelling* (no human labels, not even from the web)


Boston Dynamics