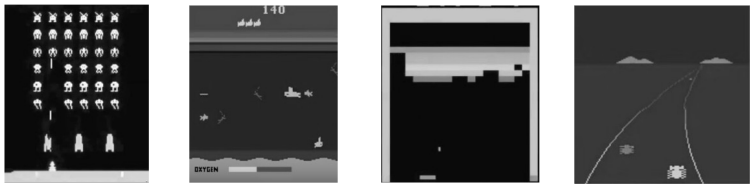
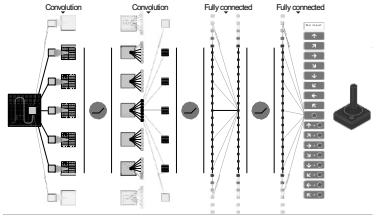
# RL + Deep Learning, applied to Classic Atari Games

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

47

47

## State of the Art

- Skydio: 'The Self-Flying Camera'

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

- Visual-inertial navigation and obstacle avoidance using multiple stereo camera pairs.
- Mobile NVidia processor onboard.

48

48

## State of the Art

- Self-Driving Car
  - Robots had to achieve extended missions in a mocked-up urban area, obeying traffic laws and avoiding other robots and cars.
  - Much more sophisticated sensor suites than in desert challenge (lasers, cameras, radars) to achieve all-around awareness.
  - Most car companies now have major autonomous driving projects. Other companies are developing 'autonomous taxi' services.

49

49

## State of the Art

- Waymo 360 degree Experience: A Fully Autonomous Driving Journey


https://www.youtube.com/watch?v=B8R148hFxPw&t=119s

- Waymo began as the Google self-driving car project in 2009.

50

50

## 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

51

51

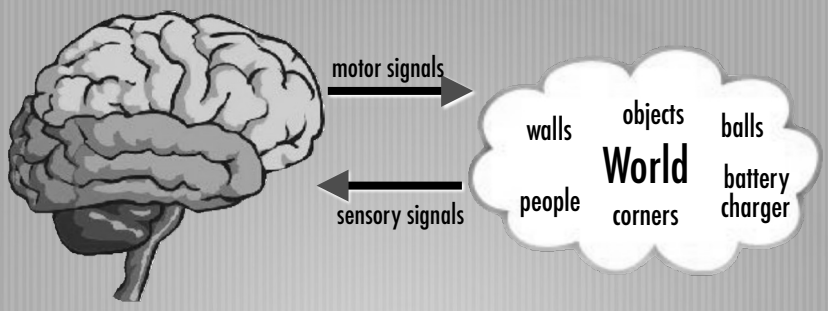
## AI is not like other sciences

- 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

52

52

# Minds are sensori-motor information processors

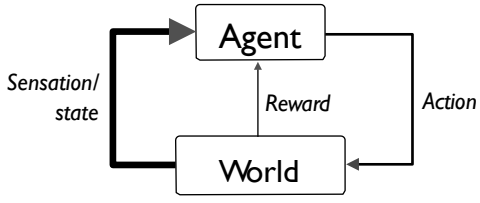



the mind's job is to predict and control its sensory signals

53

53

# Reinforcement learning is *more autonomous learning*




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

54

54

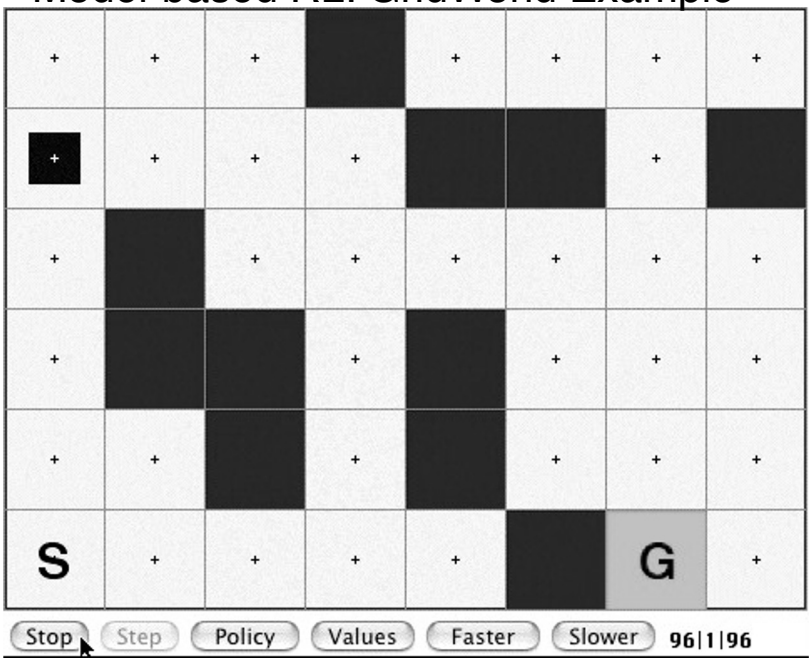
# Course Overview

- 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

55

55

## Model-based RL: GridWorld Example




56

56

# Order of Presentation

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

57

57

# Prerequisites

- Some comfort or interest in thinking abstractly and with mathematics
- Elementary statistics, probability theory
  - conditional expectations of random variables
- Basic linear algebra: vectors, vector equations, gradients
- Basic programming skills (Python)
  - Neural networks, backpropagation

58

58

Video time:
Boston Dynamic
https://www.youtube.com/watch?v=7Q3YW-3KCzU

Amazon warehouse robot
https://www.youtube.com/watch?v=Ox05Bks2Q3s

59

59


60

60


61

61