Microsoft

Reinforcement Learning:
Course Overview

ROLAND FERNANDEZ

Researcher, MSR AI

Instructor, AI School

Outline ("Course Overview")

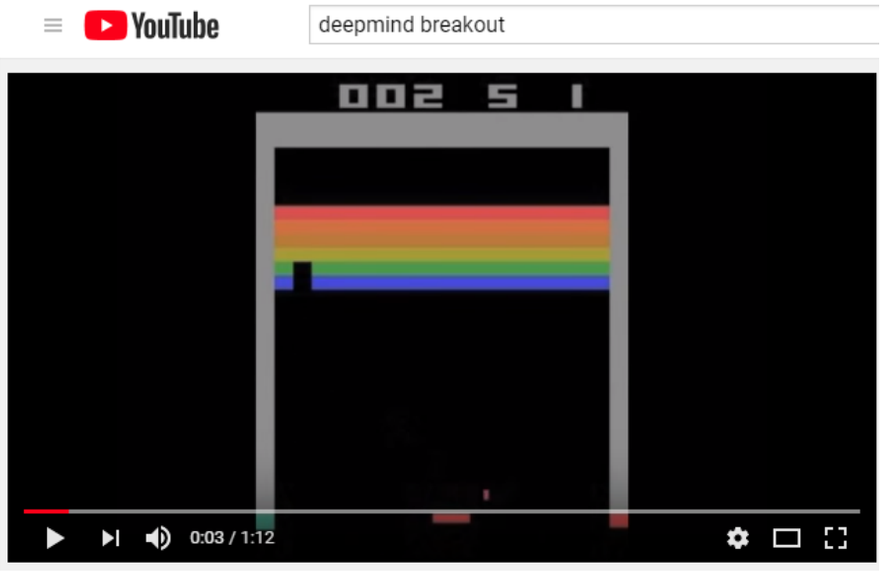
- Applications of RL
- What you will learn
- Modules Overview
- Labs Overview
- Books
- How to Install Lab Software

Outline

- Applications
- What you will learn?
- Modules Overview
- Labs Overview
- Books, Quizzes, Grading
- How to Install Lab Software

Applications

- **Playing Atari with Deep Reinforcement Learning** (Mnih, 2013)
  - Paper: https://www.cs.toronto.edu/~vmnih/docs/dqn.pdf
  - Video: https://www.youtube.com/watch?v=TmPfTpjtdgg



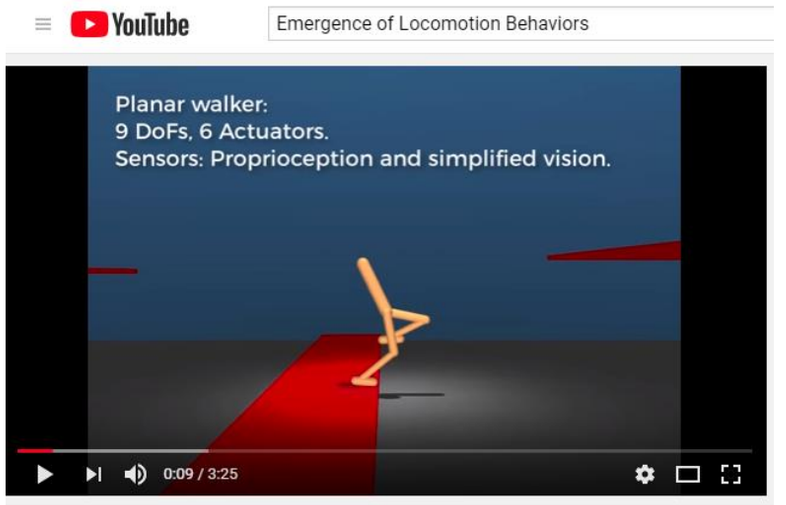
Applications

- **Hybrid Reward Architecture for Reinforcement Learning** (van Seijen, 2017)
  - Paper: https://arxiv.org/abs/1706.04208
  - Video: https://www.youtube.com/watch?v=zQyWMHFjewU

Divide and conquer: How Microsoft researchers used AI to master Ms. Pac-Man

Applications

- **Emergence of Locomotion Behaviors in Rich Environments** (Heess, 2017)
  - Paper: https://arxiv.org/pdf/1707.02286.pdf
  - Video: https://www.youtube.com/watch?v=hx_bgoTF7bs

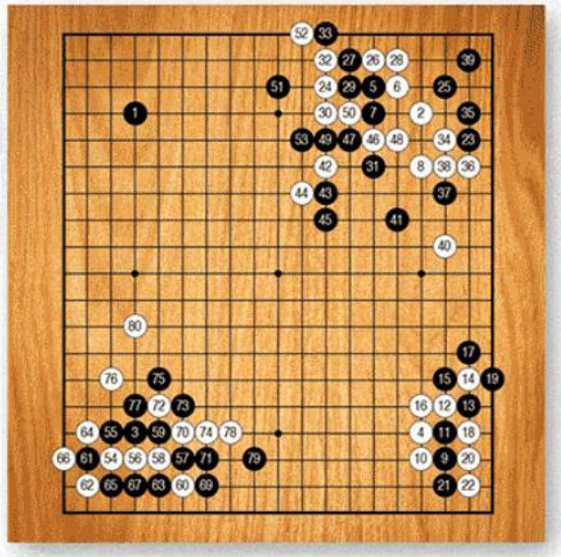


Applications

- AI-controlled Sailplane:
  - Paper: https://news.microsoft.com/features/science-mimics-nature-microsoft-researchers-test-ai-controlled-soaring-machine/
  - Video: https://www.youtube.com/watch?v=daINKmR1M-4

Applications

- **Mastering the game of Go without human knowledge** (Silver, 2017)
  - Paper: http://www.nature.com/nature/journal/v550/n7676/full/nature24270.html
  - Video: https://deepmind.com/blog/alphago-zero-learning-scratch/



(68) at (61)

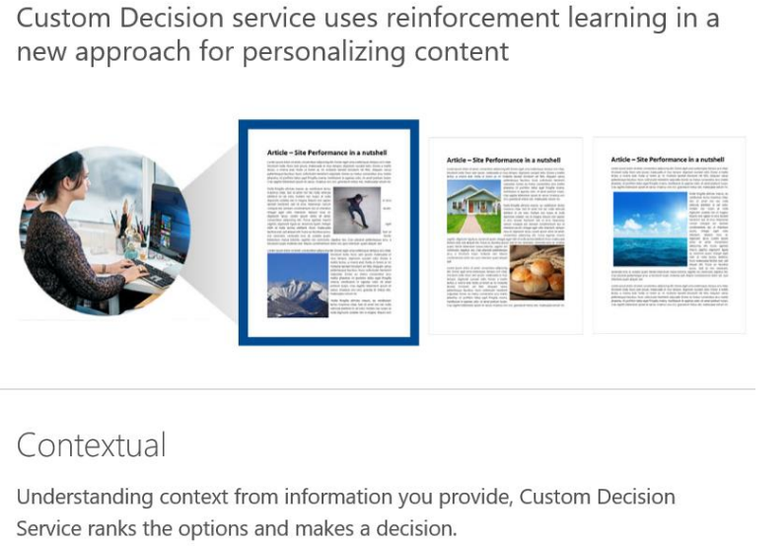
Captured Stones

70 hours

AlphaGo Zero plays at super-human level. The game is disciplined and involves multiple challenges across the board.

Applications

- Contextual Bandits at Microsoft
  - Paper: A Contextual-Bandit Approach to Personalized News Article Recommendation (Li, 2010)
  - Microsoft Azure Custom Decision Service: https://azure.microsoft.com/en-us/services/cognitive-services/custom-decision-service/



Outline

- Applications
- What you will learn
- Modules Overview
- Labs Overview
- Books
- How to Install Lab Software

What you will learn

- Intuition, formal notation, and how to implement core RL techniques:
  - Various Bandit algorithms
  - Q learning
  - DQN (Deep-Q Network)
  - REINFORCE
  - AC3 (Asynchronous Actor Critic)

Outline

- Applications
- What you will learn
- Modules Overview
- Labs Overview
- Books
- How to Install Lab Software

Modules Overview

- M0: Course Overview
- M1: Introduction to Reinforcement Learning
- M2: Bandits
- M3: The Reinforcement Learning Problem
- M4: Dynamic Programming
- M5: Temporal Difference Learning
- M6: Approximate Solution Methods
- M7: Policy Search Methods

Outline

- Applications
- What you will learn
- Modules Overview
- Labs Overview
- Books
- How to Install Lab Software

Labs Overview

- M0: (Overview) Install required lab software
- M1: (Introduction): Implement basic environment and agent
- M2: (Bandits) Tackle Bandit problem with various algorithms
- M3: (The RL Problem): Solve GridWorld with Optimal Value function
- M4: (DP) Solve Gridworld with Value & Policy Iteration
- M5: (TD) Learn Gridworld with Q-Learning and Sarsa
- M6: (Approximate Methods) Learn Cartpole with DQN
- M7: (Policy Search) Learn Cartpole with REINFORCE & A/C algorithms

Lab Software Used

- Anaconda3
- OpenAI Gym
- CNTK 2.2
- Chainer

Outline

- Applications
- What you will learn
- Modules Overview
- Labs Overview
- Books
- How to Install Lab Software

Books

- **Reinforcement Learning – An Introduction** (Sutton and Barto, 2017):
  - http://www.incompleteideas.net/sutton/book/the-book-2nd.html

Outline

- Applications
- What you will learn
- Modules Overview
- Labs Overview
- Books, Quizzes, Grading
- How to Install Lab Software

Summary for this Module (“Course Overview”)

- Applications
- What you will learn
- Modules Overview
- Labs Overview
- Books, Quizzes, Grading
- How to Install Lab Software