

# Deep (Reinforcement) Learning

Columbia University E6885 section 001

Dr. Jon Krohn

jon@untapt.com

October 26th, 2018



# Outline

- 1 Dense Networks
- 2 Specialized Network Architectures
- 3 Deep Reinforcement Learning



# Outline

- 1 Dense Networks
- 2 Specialized Network Architectures
- 3 Deep Reinforcement Learning



# Outline

- 1 Dense Networks
- 2 Specialized Network Architectures
- 3 Deep Reinforcement Learning



Dense Nets

Specialized  
Nets

Deep RL

- 1 Dense Networks
- 2 Specialized Network Architectures
- 3 Deep Reinforcement Learning



# Essential Deep Learning Theory

whiteboarding





# Deep Neural Network

Interactive Jupyter Demo

[\[\*Deep Net in Keras notebook\*\]](#)





# Outline

- 1 Dense Networks
- 2 Specialized Network Architectures
- 3 Deep Reinforcement Learning



# Convolutional Neural Networks

Dense Nets

Specialized  
Nets

Deep RL

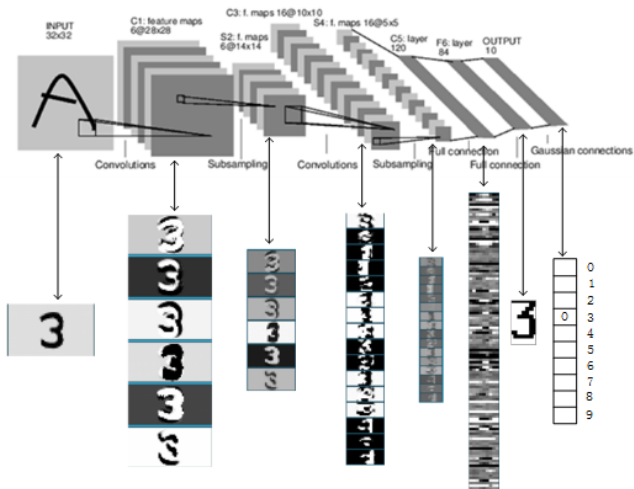
[\[Conv Layer Demo\]](#)

[\[Deep Viz\]](#)



# LeNet-5

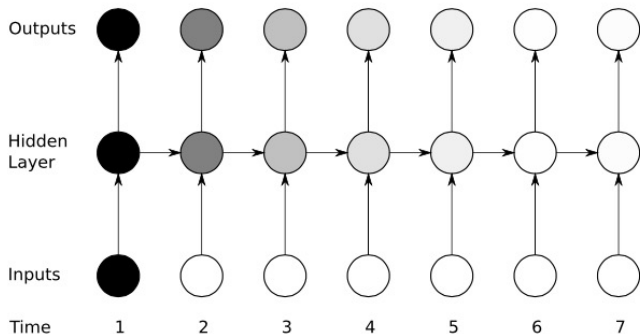
Static GitHub Demo



[\[Deep Net in Keras notebook\]](#)



# Recurrent Neural Networks



[\[Chris Olah blog post\]](#)

[\[Stacked Bidirectional LSTM notebook\]](#)

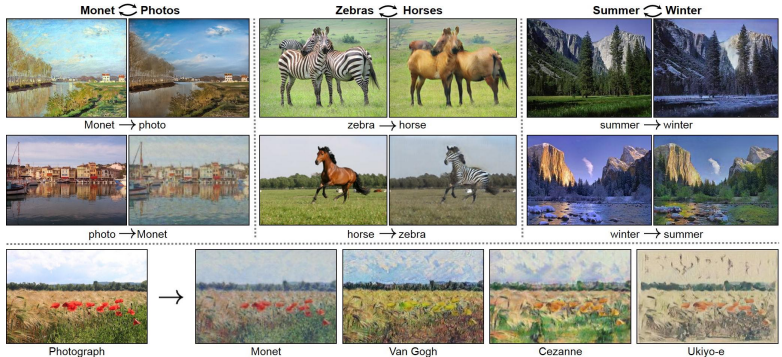


# Generative Adversarial Networks

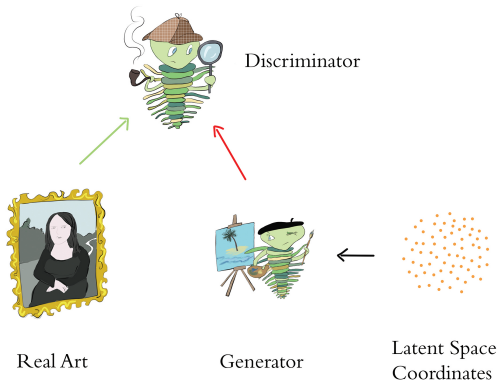
Dense Nets

Specialized Nets

Deep RL



# Generative Adversarial Networks



["Celebrity" Face Interpolation]

[Quick, Draw!]

[Generative Adversarial Network notebook]



# Outline

- 1 Dense Networks
- 2 Specialized Network Architectures
- 3 Deep Reinforcement Learning**



# AlphaGO

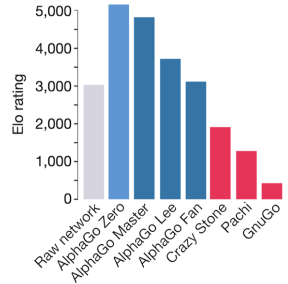
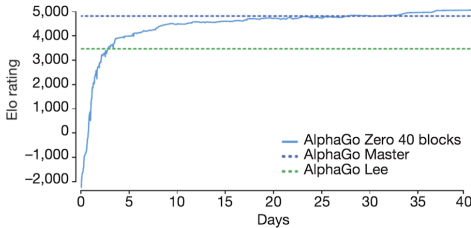
Silver et al. (2016)





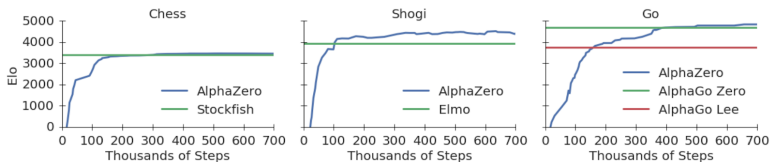
# AlphaGo Zero

Silver et al. (2017a)

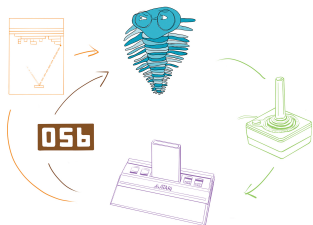
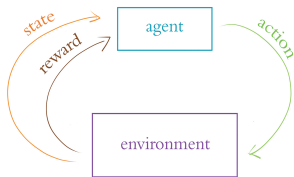


# AlphaZero

Silver et al. (2017b)

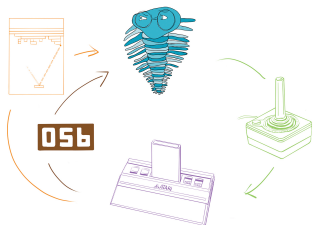
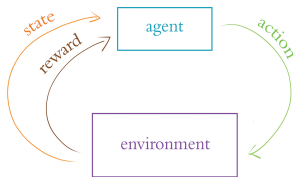


# Markov Decision Process



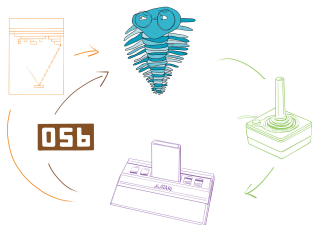
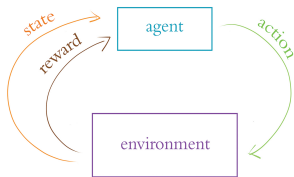
- $S$  : all possible states
- $A$  : " " actions
- $R$  : reward distribution given  $(s, a)$
- $P$  : transition prob. to  $S_{t+1}$  given  $(s, a)$
- $\gamma$  : discount factor

# Markov Decision Process



- $S$  : all possible states
- $A$  : " " actions
- $R$  : reward distribution given  $(s, a)$
- $P$  : transition prob. to  $S_{t+1}$  given  $(s, a)$
- $\gamma$  : discount factor

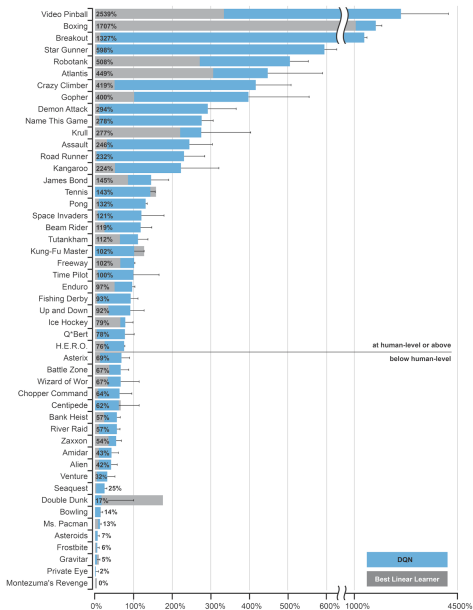
# Markov Decision Process



- $S$  : all possible states
- $A$  : " " actions
- $R$  : reward distribution given  $(s, a)$
- $P$  : transition prob. to  $S_{t+1}$  given  $(s, a)$
- $\gamma$  : discount factor

# Deep Q-Learning

Mnih et al. (2015)

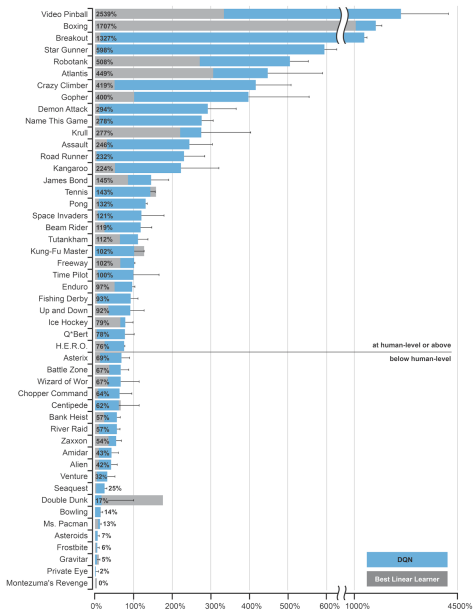


[Atari Games]



# Deep Q-Learning

Mnih et al. (2015)

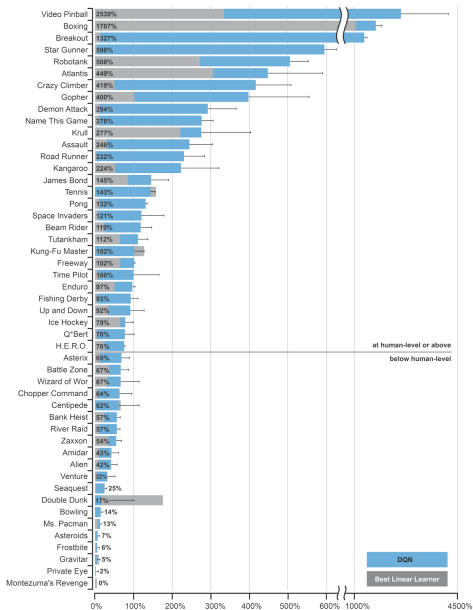


[Atari Games]



# Deep Q-Learning

Mnih et al. (2015)



[Atari Games]





## Deep Q-Learning

Interactive Jupyter Demo

$$S \xrightarrow{\pi} A$$

objective:

$$\text{find } \pi^* = \max_{\pi} \left( \sum_{t>0} \gamma^t r^t \right)$$

value function:  $V^{\pi}(s)$

Q-value function:  $Q^{\pi}(s, a)$

$$Q^*(s, a) \approx Q(s, a, \theta)$$

[\[Deep Q-Learning Network notebook\]](#)

[\[SLM-Lab\]](#)



# Additional Materials

- [\[Jon Krohn Resources\]](#)
- [\[\*Deep Learning Illustrated\* textbook\]](#)
- [\[\*Deep Learning\* at NYC Data Science Academy\]](#)



# Additional Materials

- [\[Jon Krohn Resources\]](#)
- [\[\*Deep Learning Illustrated\* textbook\]](#)
- [\[\*Deep Learning\* at NYC Data Science Academy\]](#)



# Additional Materials

- [\[Jon Krohn Resources\]](#)
- [\[\*Deep Learning Illustrated\* textbook\]](#)
- [\[\*Deep Learning\* at NYC Data Science Academy\]](#)



## Stay in Touch!

- **Twitter: @JonKrohnLearns**
- Medium: @jonkrohn
- [linkedin.com/in/jonkrohn](https://www.linkedin.com/in/jonkrohn) (mention Columbia class)



## Stay in Touch!

- Twitter: @JonKrohnLearns
- Medium: @jonkrohn
- [linkedin.com/in/jonkrohn](https://www.linkedin.com/in/jonkrohn) (mention Columbia class)



## Stay in Touch!

- Twitter: @JonKrohnLearns
- Medium: @jonkrohn
- [linkedin.com/in/jonkrohn](https://www.linkedin.com/in/jonkrohn) (mention Columbia class)

