

# [RL22] Q&A Session on Policy-Based RL #1

---

09.06.2022

Christopher Mutschler

Let's play Kahoot!

**Kahoot!**

# Let's play Kahoot!

The image shows a screenshot of the Kahoot! website homepage. At the top, there is a navigation bar with the Kahoot! logo on the left and several menu items: News (with a red notification badge), School, Work, Home, Study, Academy, and AccessPass. On the right side of the navigation bar, there are buttons for 'Contact sales', 'Explore content', 'Play' (circled in red), 'Sign up', 'Log in', and a language selector 'EN'. Below the navigation bar, there are four main promotional cards. The first card on the left is purple and white, titled 'Make learning awesome!' and features a photo of two young women laughing. The second card is blue and white, titled 'Make your team superstar presenters' and features a photo of a group of people in a meeting. The third card is white and blue, titled 'NEW! Create a branded experience with Kahoot! themes' and features a photo of a Kahoot! game interface with a score of 55427 and names like Rik, Mal, Mille, and Martin. The fourth card is purple and white, titled 'Meet Kahoot! Kids!' and features a colorful illustration of a rocket and stars. Each card has a green call-to-action button.

**Make learning awesome!**  
Kahoot! delivers engaging learning to billions.  
[Sign up for free!](#)

**Make your team superstar presenters**  
Set your whole team up to deliver awesome presentations with Kahoot! 360 Spirit, our best plan from only \$16 per month.  
[Learn more >](#)  
[Buy now](#)

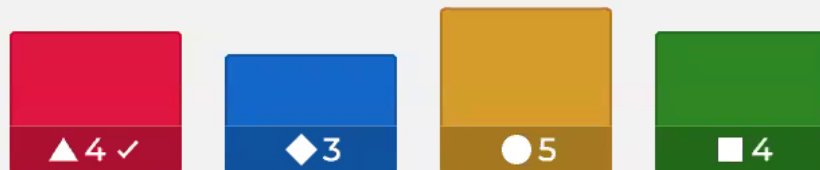
**NEW! Create a branded experience with Kahoot! themes**  
Boost audience engagement by customizing your kahoots for your work setting.  
[Choose Kahoot! 360 Pro Max](#)

**Meet Kahoot! Kids!**  
Spark your child's curiosity for learning with our new playful app experience.  
[Get started today](#)



# Deadly Triad: what is a major source of instability?

Next



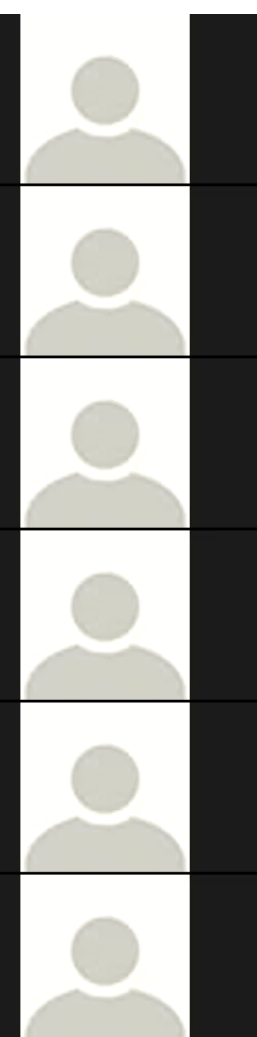
Show media

▲ Off-policy Learning ✓

◆ Dynamic Programming ✗

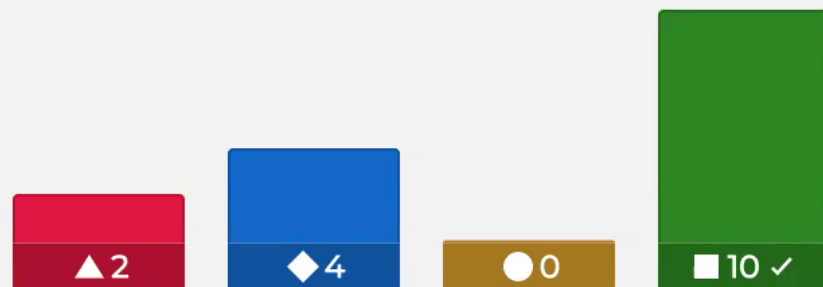
● On-Policy Learning ✗

■ Overfitting ✗



# Deadly Triad: what is a major source of instability?

Next



Show media

▲ Linear Features



◆ Unknown Dynamics



● Generalization

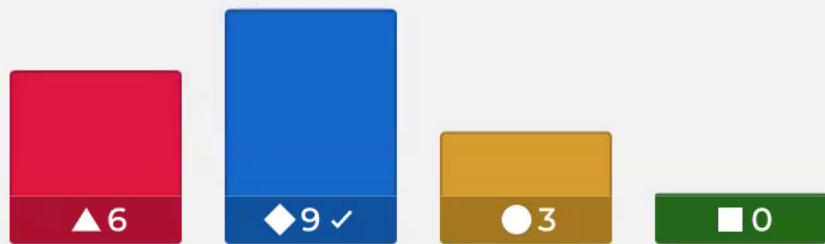


■ Function Approximation



# Deadly Triad: what is a major source of instability?

Next



Show media

▲ Stochastic Gradient Descent ✕

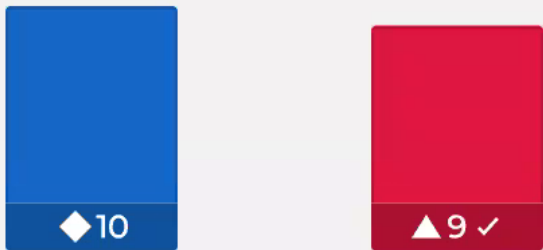
◆ Bootstrapping ✓

● I.I.D. ✕

■ Activation Functions ✕

# Policy Gradients efficiently make use of replay buffers

Next



Show media

◆ True ✕

▲ False ✓



# Policy Gradients can solve problems that cannot be solved with value-based approaches

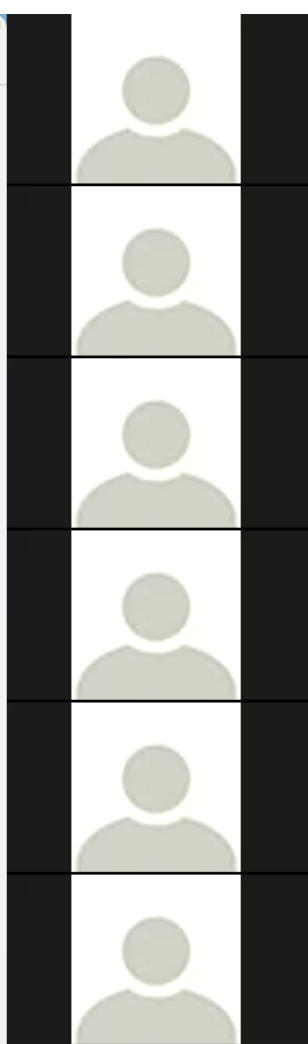
Next



Show media

◆ True ✓

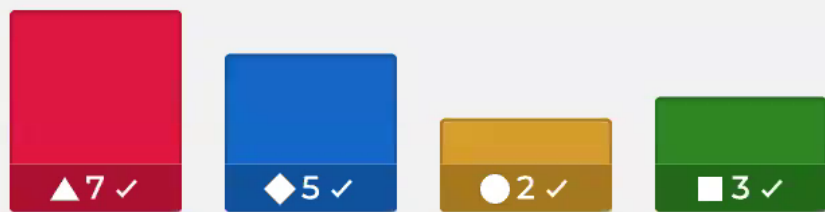
▲ False ✗





# What reduces variance of the policy gradient?

Next



Show media

▲ reward-to-go ✓

◆ state-value function ✓

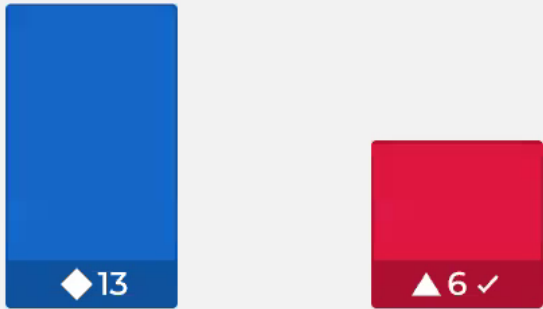
● state-action-value function ✓

■ n-step TD-error ✓



# The estimation of the policy gradient requires transition probabilities

Next



Show media

◆ True ✕

▲ False ✓

