

[RL22] Q&A Session on Exploration Strategies

14.07.2022

Christopher Mutschler

Let's play Kahoot! first...

Kahoot!

Let's play Kahoot!

The image shows a screenshot of the Kahoot! website homepage. The browser address bar shows 'kahoot.com'. The navigation menu includes 'Kahoot!', 'News', 'School', 'Work', 'Home', 'Study', 'Academy', 'AccessPass', 'Contact sales', 'Explore content', 'Play', 'Sign up', 'Log in', and 'EN'. The 'Play' button is circled in red. Below the navigation are four promotional cards:

- 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](#)

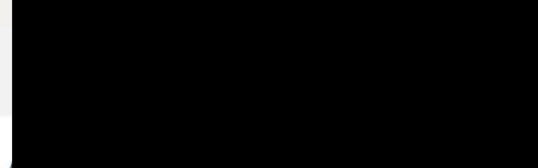
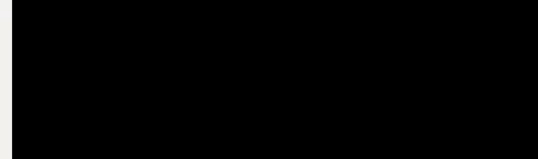
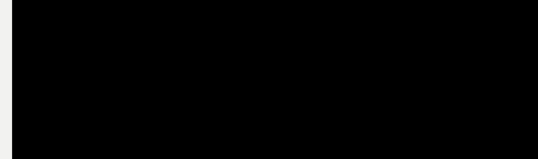
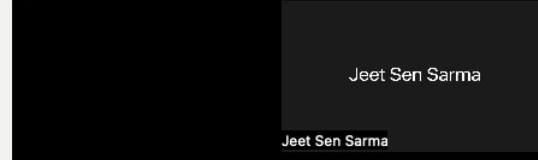
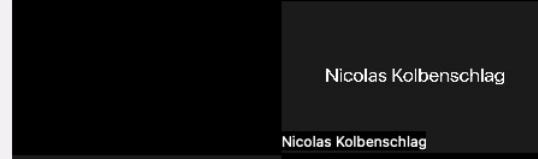
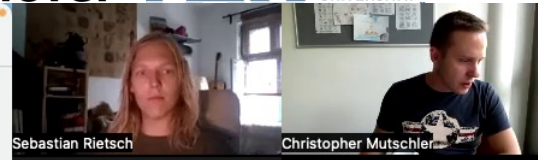
Course Logistics

Some remarks on the exam

- The exam will take place on **August, 10th** (8:15 – 9:45 ← this is still preliminary!)
- Preliminary room: Felix-Klein-Building (Cauerstraße 11), either H12 or H11

Some remarks on the evaluation

- Did you receive an invitation to an online evaluation??



Optimal Exploration: Sort by Theoretical Tractability (top: tractable, bottom: not tractable)

Next



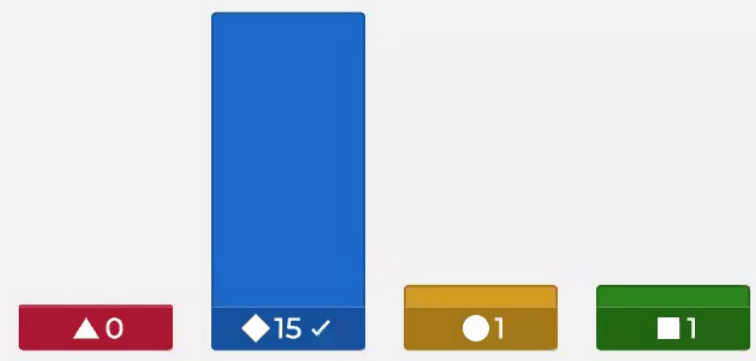
Show media

- ◆ Multi-armed Bandit ✓
- Contextual Bandit ✓
- Small and finite MDPs ✓
- ▲ Small MDPs with continuous state spaces ✓

Sebastian Rietsch
Christopher Mutschler
Pelin Genc
Pelin Genc
Nicolas Kolbenschlag
Nicolas Kolbenschlag
Abhinav
Abhinav
Jeet Sen Sarma
Jeet Sen Sarma

What helps us to quantify the exploration-exploitation trade-off?

Next



Show media

▲ Return ✕

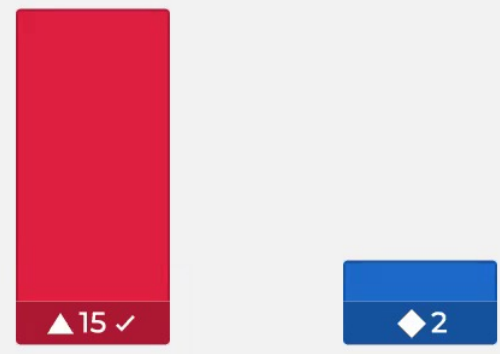
◆ Regret ✓

● Reward ✕

■ Cost ✕

What is true? "A good exploration algorithm..."

Next



Show media

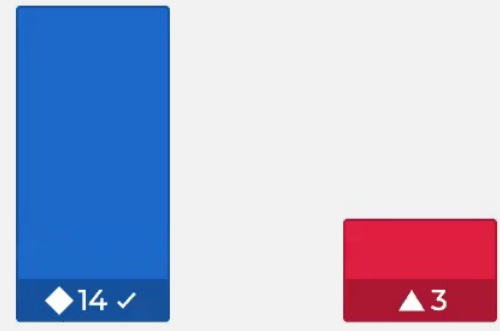
▲ ensures small visitation counts for large action regrets ✓

◆ yields high rewards for states with small visitation counts ✗

A grid of video conference windows. The top row shows two active windows for Sebastian Rietsch and Christopher Mutschler. Below them are several windows for other participants, some of which are dark, indicating they are muted or have their video off. The names visible are Pelin Genc, Nicolas Kolbenschlag, Abhinav, and Jeet Sen Sarma.

ϵ -greedy has a linear expected total regret

Next



Show media

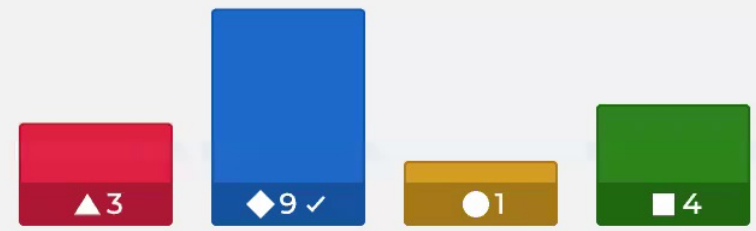
◆ True ✓

▲ False ✗

A grid of video conference participants. The top row shows two active video feeds: Sebastian Rietsch and Christopher Mutschler. Below them are several dark grey boxes with names: Pelin Genc, Pelin Genc, Nicolas Kolbenschlag, Nicolas Kolbenschlag, Abhinav, Abhinav, Jeet Sen Sarma, and Jeet Sen Sarma.

Upper Confidence Bounds - the expected total regret is (with $c = \sqrt{2}$)...

Next



Show media

▲ linear ✕

◆ logarithmic ✓

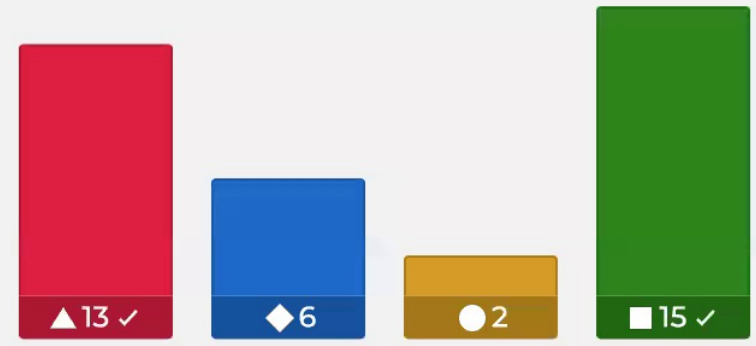
● exponential ✕

■ bounded ✕

Sebastian Rietsch
Christopher Mutschler
Pelin Genc
Nicolas Kolbenschlag
Abhinav
Jeet Sen Sarma

Bayesian UCB...

Next



Show media

▲ uses a Beta distribution for the posterior ✓

◆ uses a Gaussian process to estimate the rewards ✗

● uses Monte-Carlo dropout to robustly estimate rewards ✗

■ allows to add a prior on the reward distribution ✓

Sebastian Rietsch
Christopher Mutschler
Pelin Genc
Pelin Genc
Nicolas Kolbenschlag
Nicolas Kolbenschlag
Abhinav
Abhinav
Jeet Sen Sarma
Jeet Sen Sarma

Thompson Sampling samples an action based on the current prior, executes the best one, and updates the Beta distribution

Next



Show media

◆ True ✓

▲ False ✗

A grid of video conference windows. The top row shows two active windows for Sebastian Rietsch and Christopher Mutschler. Below them are several windows for other participants, some of which are muted or have their video off.

Participant Name	Status
Sebastian Rietsch	Active
Christopher Mutschler	Active
Pelin Genc	Muted
Nicolas Kolbenschlag	Muted
Abhinav	Muted
Jeet Sen Sarma	Muted