Fraunhofer IIS

Fraunhofer Institute for Integrated Circuits IIS

FAU FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

[RL22] Q&A Session on Model-based RL #2

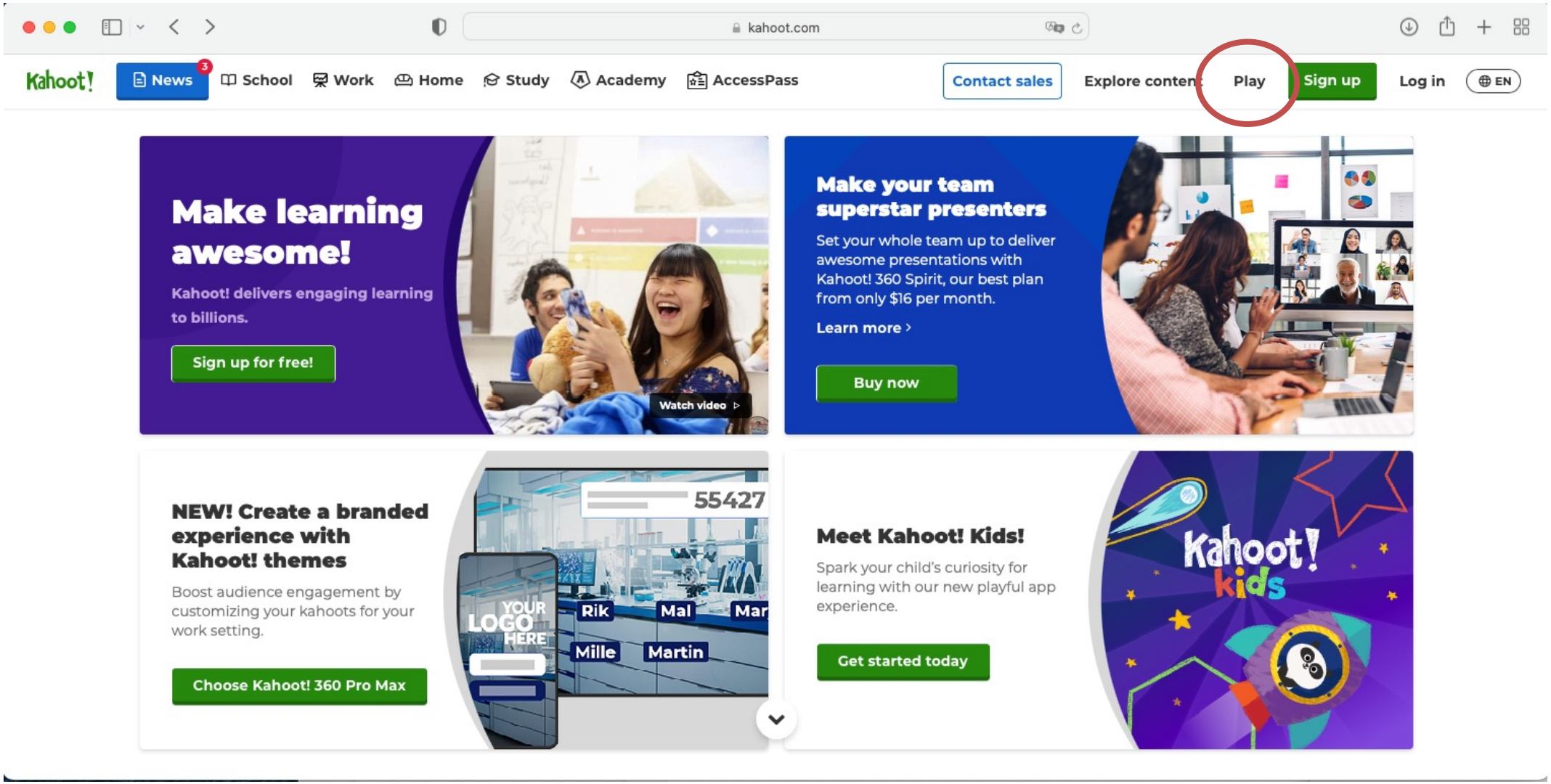
07.07.2022

Christopher Mutschler

Let‘s play Kahoot! again…

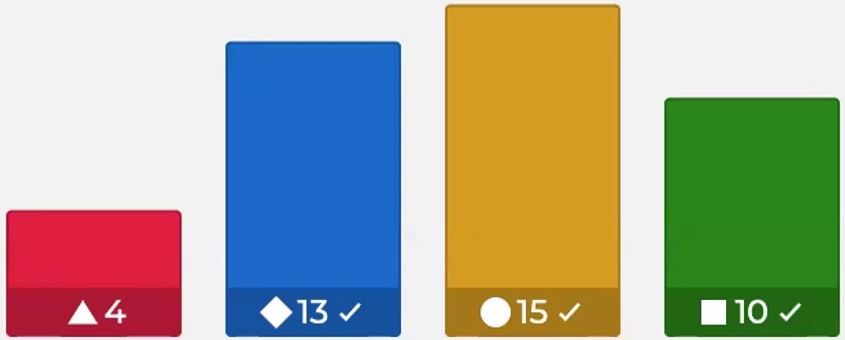
Kahoot!

Let‘s play Kahoot!



play.kahoot.it/v2/gameblock?quizId=a1826aa4-5dd0-4fbc-a5c2-6aa374d1abb3

Trajectory Optimization with Derivatives



Show media

- ▲ LSTMs help with the poor conditioning induced by the recursive gradients ✗
- ◆ Samples the initial action sequence randomly ✓
- ● Optimizes the action sequence via gradient descent ✓
- ■ Collocation makes downstream state changes less sensitive to early changes ✓

1/8

kahoot.it Game PIN: 2679590

Nicolas Kolbenschlag

Nicolas Kolbenschlag

Veronica Buckina

Veronica Buckina

FAU084414

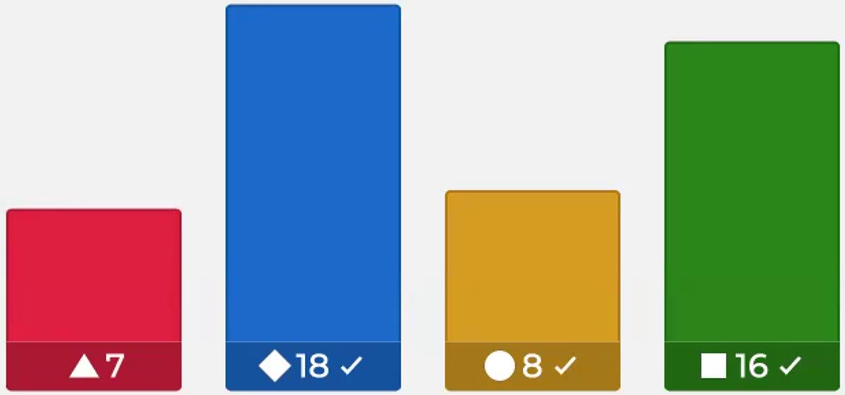
FAU084414

Takuma Nishimura

Takuma Nishimura

play.kahoot.it/v2/gameblock?quizId=a1826aa4-5dd0-4fbc-a5c2-6aa374d1abb3

Cross Entropy Maximization



Show media

- ▲ Uses gradients to optimize a number of action sequences ✗
- ◆ Fits a Gaussian to sample policy parameters at every iteration ✓
- ● Is able to find the global optimum ✓
- ■ Is easy to implement and embarrassingly easy to parallelize ✓

Nicolas Kolbenschlag

Nicolas Kolbenschlag

Veronica Buckina

Veronica Buckina

FAU084414

FAU084414

Takuma Nishimura

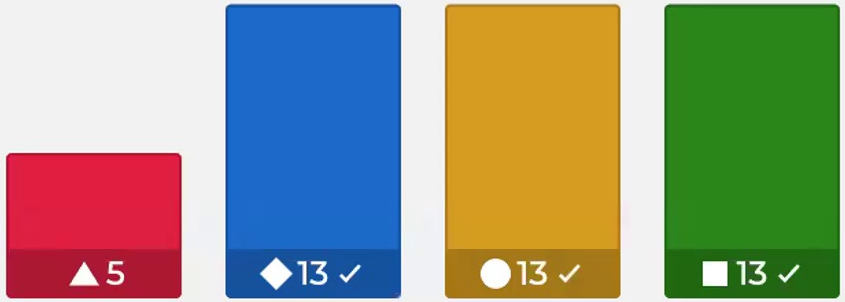
Takuma Nishimura

2/8

kahoot.it Game PIN: 2679590

play.kahoot.it/v2/gameblock?quizId=a1826aa4-5dd0-4fbc-a5c2-6aa374d1abb3

Uncertainty of models in RL



Show media

- ▲ Aleatoric uncertainty can be reduced close to 0 by training with more data ✗
- ◆ Episdemic uncertainty often leads to extrapolation errors ✓
- ● Episdemic uncertainty changes over the course of learning ✓
- ■ Aleatoric uncertainty captures the world's inherent stochasticity ✓

Nicolas Kolbenschlag

Nicolas Kolbenschlag

Veronica Buckina

Veronica Buckina

FAU084414

FAU084414

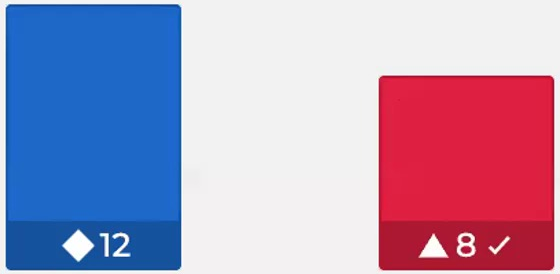
Takuma Nishimura

Takuma Nishimura

kahoot.it Game PIN: 2679590

play.kahoot.it/v2/gameblock?quizId=a1826aa4-5dd0-4fbc-a5c2-6aa374d1abb3

Model Predictive Control applies a sequence of actions at once (with a maximum length of the planning horizon)



Show media

True

False

kahoot.it Game PIN: 2679590

Nicolas Kolbenschlag

Nicolas Kolbenschlag

Veronica Buckina

Veronica Buckina

FAU084414

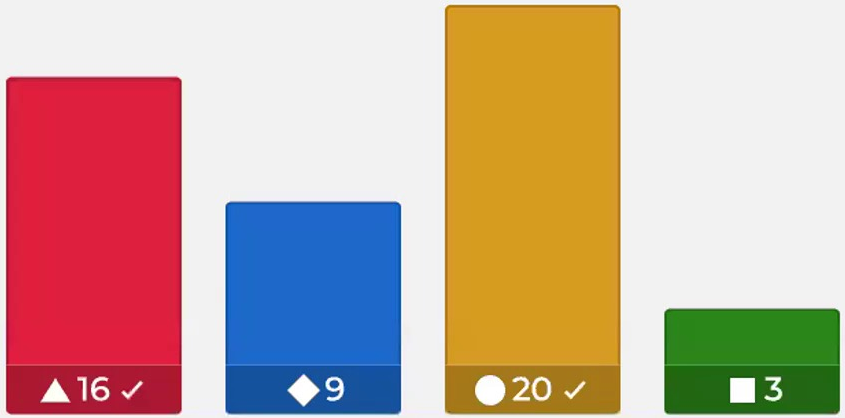
FAU084414

Takuma Nishimura

Takuma Nishimura

play.kahoot.it/v2/gameblock?quizId=a1826aa4-5dd0-4fbc-a5c2-6aa374d1abb3

Uncertainty Estimation: what models can we use?



Show media

- ▲ Model Ensembles ✓
- ◆ Recurrent Neural Networks ✗
- ● Gaussian Processes ✓
- ■ Support Vector Machines ✗

kahoot.it Game PIN: 2679590

Nicolas Kolbenschlag

Nicolas Kolbenschlag

Veronica Buckina

Veronica Buckina

FAU084414

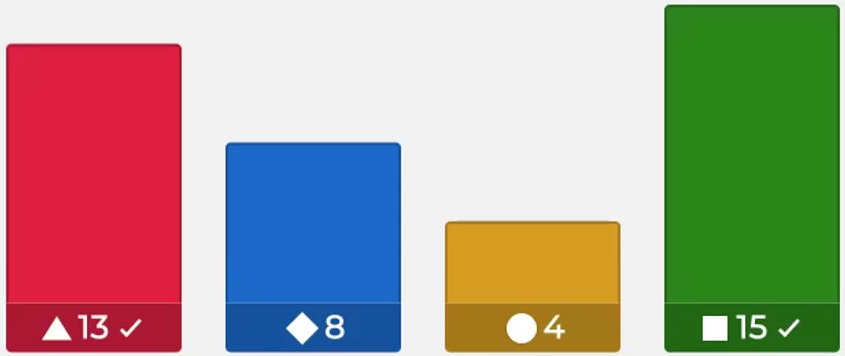
FAU084414

Takuma Nishimura

Takuma Nishimura

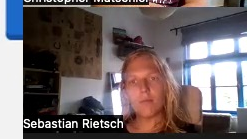
play.kahoot.it/v2/gameblock?quizId=a1826aa4-5dd0-4fbc-a5c2-6aa374d1abb3

Uncertainty Propagation: what methods can we use?



Show media

- ▲ Moment Matching ✔
- ◆ Stochastic Gradient Descent ✖
- ● Time-Difference Learning ✖
- ■ Trajectory Sampling ✔



Nicolas Kolbenschlag

Nicolas Kolbenschlag

Veronica Buckina

Veronica Buckina

FAU084414

FAU084414

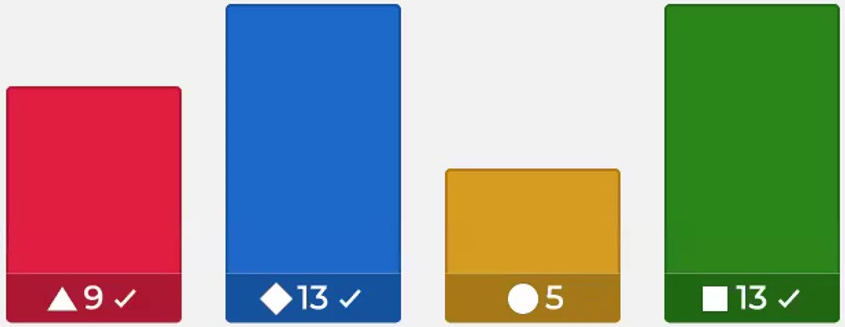
Takuma Nishimura

Takuma Nishimura

kahoot.it Game PIN: 2679590

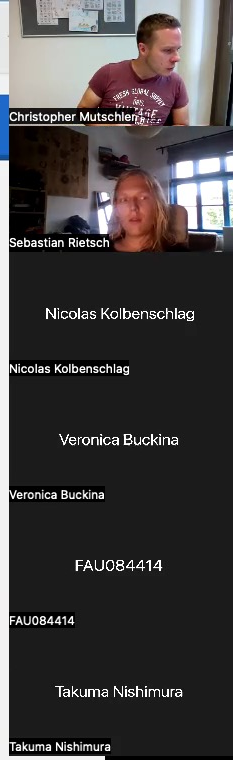
play.kahoot.it/v2/gameblock?quizId=a1826aa4-5dd0-4fbc-a5c2-6aa374d1abb3

PILCO



Show media

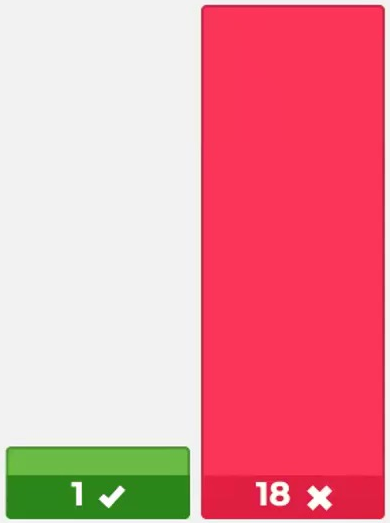
- ▲ Is robust to small model errors ✓
- ◆ Is very sample efficient ✓
- ● Does not require smooth dynamics ✗
- ■ Is computationally expensive w.r.t. the available training data ✓



kahoot.it Game PIN: 2679590

play.kahoot.it/v2/gameblock?quizId=a1826aa4-5dd0-4fbc-a5c2-6aa374d1abb3

MBRL: Sample Efficiency (top: requires many samples, bottom: requires little samples)



Show media

- Augmented Random Search ✓
- Policy-based RL ✓
- Value-based RL ✓
- Model-based RL ✓

Nicolas Kolbenschlag

Nicolas Kolbenschlag

Veronica Buckina

Veronica Buckina

FAU084414

FAU084414

Takuma Nishimura

Takuma Nishimura

kahoot.it Game PIN: 2679590