

# Curriculum Reinforcement Learning

## SPaCE

Marius Lindauer



Winter Term 2021

- ▶ Most agents estimate the state-value function  $V$  in some way
- ▶ Common definition of  $V$ :

$$V(s) = \mathbb{E}_{\pi} \sum_t^T \gamma^t \cdot r_t$$

- ▶ Therefore  $V(s_0)$  estimates the total discounted reward for the whole episode
- ▶ Idea: use this information for curriculum generation

# Setting – what makes SPaCE special?

- ▶ Designed for deep contextual RL
- ▶ No prior knowledge about instance space required, e.g.:
  - ▶ size
  - ▶ difficulty
  - ▶ difficulty regions

# Algorithm Outline

- ▶ Hyperparameters: threshold  $\eta$ , increment size  $\kappa$
- ▶ Until desired number of steps is reached:
  - ▶ Choose the instances on which the evaluation has changed most (according to current instance set size)
  - ▶ Train on those instances
  - ▶ Evaluate if performance on the training set has changed by at least  $\eta$ . If not, increase instance set size by  $\kappa$

## Results - AntGoal

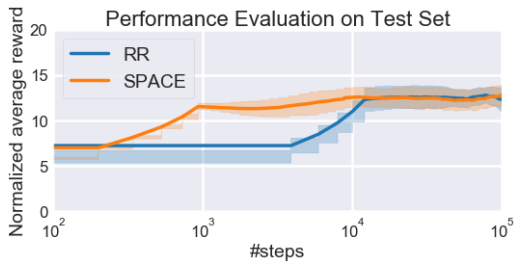


Figure: Comparing SPaCE and round robin on AntGoal with broken limbs.

## Results - PointMass

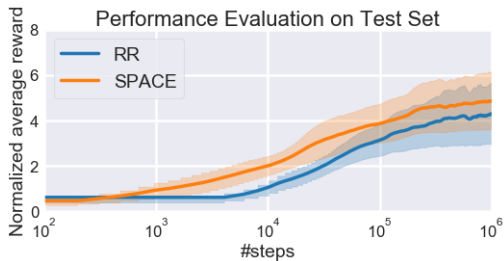


Figure: Comparing SPaCE and round robin on contextual PointMass.