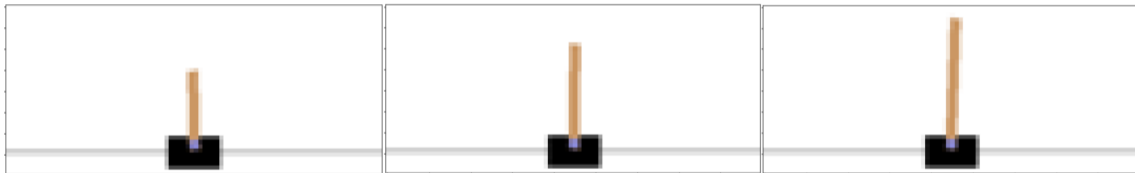


Self-Paced Context Evaluation for Contextual Reinforcement Learning

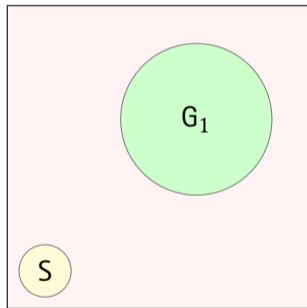
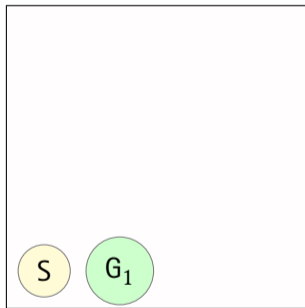
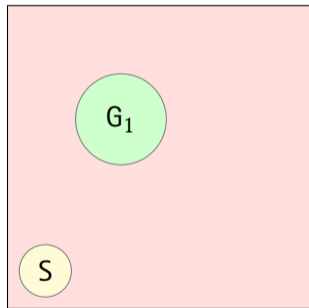
Theresa Eimer¹, André Biedenkapp², Frank Hutter^{2,3}, Marius Lindauer¹

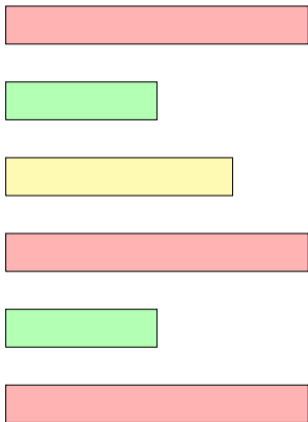
¹Leibniz Universität Hannover | ²Albert-Ludwigs Universität Freiburg | ³Bosch Center for Artificial Intelligence

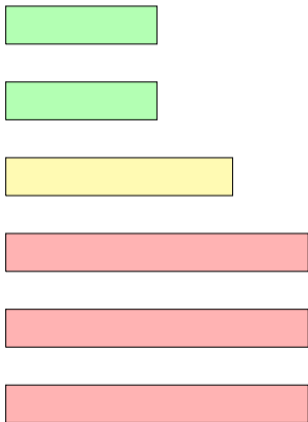
What is Context?



What is Context?







- ▶ Value function V can be used to generate instance curricula [Klink et al., 2020, Xie et al., 2020]
- ▶ Instance generator may not generally be available
- ▶ Alternative: order given contexts according to evaluation

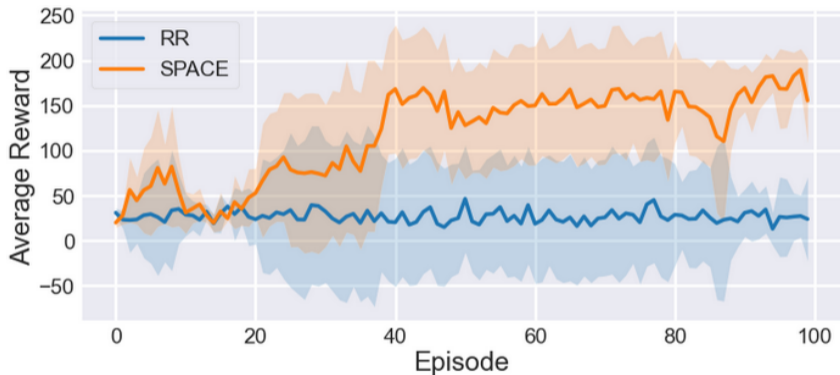


Figure: Performance of SPaCE & round robin on Contextual Cartpole.

- ▶ We use that $V(s_t, c_i)$ converges towards the maximum expected reward [Sutton and Barto, 1998]
- ▶ Using $t = 0$ gives us the expected reward per episode for each instance
- ▶ Performance improvement capacity (PIC): $d_t(i) = V_t(s_0, c_i) - V_{t-1}(s_0, c_i)$
- ▶ Evaluating PIC this way adds little overhead

Algorithm 1 SPaCE

```
1: size = 1
2: for i = 1 to num_iterations do
3:   if little change in V then
4:     ++size
5:   end if
6:    $d_t = \text{compute\_PIC}()$ 
7:    $\text{sort\_desc}(d_t)$ 
8:    $\text{instances} = \{i \in I \mid d_t(i) \in d_t[: \text{size}]\}$ 
9:   Train agent on instances
10: end for
```

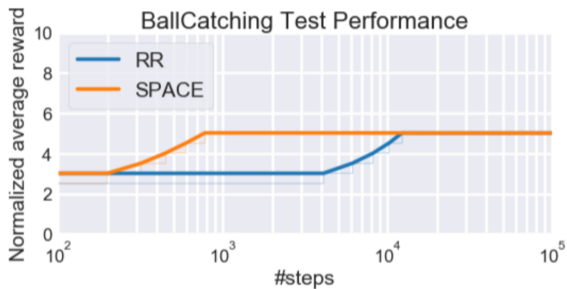



Figure: Train and test performance of SPaCE & round robin on AntGoal and Contextual BallCatching [Klink et al., 2020].


The value function can be a great tool in curriculum learning, even without any domain knowledge.


See our full paper for insights and further experiments on:

- ▶ More complex context representations
- ▶ The influence of the instance set size
- ▶ Convergence of SPaCE compared to Round Robin

Looking forward to seeing you at the poster!

 Klink, P., D'Eramo, C., Peters, J., and Pajarinen, J. (2020).
Self-paced deep reinforcement learning.
arXiv:2004.11812 [cs.LG].

 Sutton, R. S. and Barto, A. G. (1998).
Reinforcement Learning - an Introduction.
Adaptive computation and machine learning. MIT Press.

 Xie, Z., Ling, H. Y., Kim, N. H., and van de Panne, M. (2020).
ALLSTEPS: Curriculum-driven learning of stepping stone skills.
arXiv:2005.04323 [cs.GR].