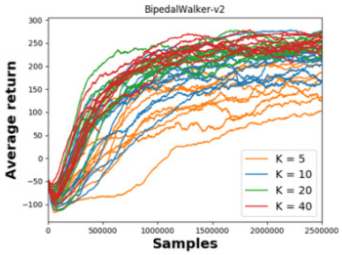
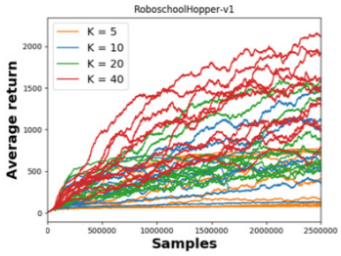
Hessisches Kompetenzzentrum
für Hochleistungsrechnen

Interpretable Reinforcement Learning





Introduction

Deep Reinforcement Learning can solve difficult high dimensional tasks, by exploiting the expressive representation power of Deep Neural Networks. However, to deploy Deep RL in the real world, the learned policy must be safe and trustable. However, these properties can be hardly achieved by a Deep Neural Network. The objective of this project is to investigate RL approaches that can achieve performances comparable to the Deep Learning ones, using an interpretable policy instead of a black-box Deep Neural Network. We based our policy representation on state prototypes, and we proceed in clustering the state space assigning a membership value for every state w.r.t. each cluster. Optimizing policies that are both expressive and interpretable is a major challenge, particularly because, to obtain these properties, we need to relax policy requirements (such as differentiability) or exploit hierarchical structures: in both fields, the Reinforcement Learning literature still lacks on well-established solutions, particularly in the context of robotic agents. The use of the HRZ cluster was crucial due to the need of running extensive comparisons with standard deep reinforcement learning, and the high computational demand of non- differentiable optimization.

Methods

In this project, we used standard differentiable optimization techniques (back-propagation and neural network optimizers) and stochastic optimization techniques, in particular, randomized search under heuristics.

Results

We were able to derive an interpretable reinforcement learning algorithm that worked comparably well w.r.tw. standard Deep Reinforcement Learning approaches. Most of the outcome of the research is still in the scientific review phase.

Project Manager
Dr. Davide Tateo

Researchers
Benedikt Volker

Principal Investigator
Dr. Davide Tateo

Project Term
2019 - 2020

Clusters
Lichtenberg Cluster Darmstadt

Software
PyTorch

University
Technische Universität Darmstadt

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN
TECHNISCHE UNIVERSITÄT DARMSTADT
UNIKASSEL VERSITÄT
Philipps Universität Marburg

Discussion

While the developed techniques look promising, we are still far from a completely interpretable reinforcement learning approach, and more work should be done in the area to reach the final objective. Our work shows that to have interpretable policies, it's often needed to trade off the simplicity of the algorithm and increase computational resources.

Figures

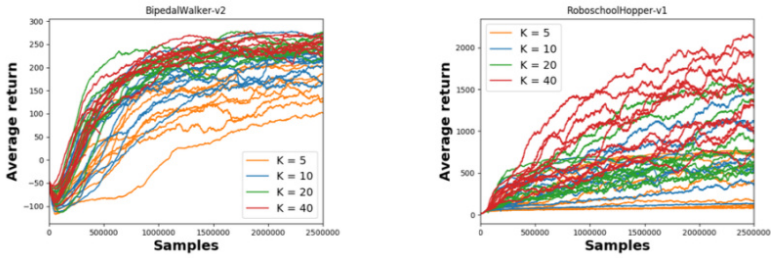


Figure 1: Performances of the proposed algorithm with different number of state clusters

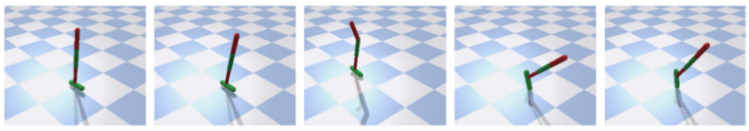


Figure 2: Interpretable clusters found by the algorithm in the Hopper environment

Publications

Akrour, R.; Tateo, D.; Peters, J.: Towards Reinforcement Learning of Human Readable Policies in ECML/PKDD Workshop on Deep Continuous-Discrete Machine Learning, (2019) https://www.ias.informatik.tu-darmstadt.de/uploads/Team/RiadAkrour/decodeml2019.pdf

Last Update: 2022-04-17 15:22