Degree Project Management Tool (/)

UG4/MInf Projects - 24/25 (/ug4)

Learning Optimal Actions Using Differentiable Simulation

Principal Goal: Leverage differentiable simulation in RL

This project has been proposed by Michael Mistry (/ug4/person/1356) (mmistry@exseed.ed.ac.uk (mailto:mmistry@exseed.ed.ac.uk)). You have been approved **but not selected** for this project.

Model-free Reinforcement Learning (RL) has had an enormous impact in domains where reward functions are well-defined and actions are discrete (see *MuZero*) [1]. This success has not been fully transferred over to continuous domains such as robotics. We have seen teacher-student policies learn and generalise different behaviours for quadruped robots but this has been at the cost of very long and inefficient training runs. Optimisation-based control is model-based and uses derivative information, allowing it to generate clean intricate behaviours [2]. However, the methods need to be more scalable, without needing to solve problems every few milliseconds. Recent approaches have shown that we can leverage derivative information and differentiable simulation to learn value/policy functions much more efficiently than model-free RL [3, 4, 5]. These results however are limited and do not explore the tradeoffs. In this project, we aim to leverage differentiable simulators [6, 7] and any gradient-based methods to benchmark against widely used RL algorithms on complex nonlinear tasks.

1	https://www.nature.com/articles/s41586-020-03051-4.pdf (https://www.nature.com/articles/s41586-020-03051-4.pdf)
2	https://arxiv.org/pdf/2312.08961.pdf (https://arxiv.org/pdf/2312.08961.pdf)
3	https://arxiv.org/abs/2305.15244 (https://arxiv.org/abs/2305.15244)
4	http://proceedings.mlr.press/v144/ainsworth21a/ainsworth21a.pdf (http://proceedings.mlr.press/v144/ainsworth21a/ainsworth21a.pdf)
5	https://arxiv.org/abs/2204.07137 (https://arxiv.org/abs/2204.07137)
6	https://github.com/NVlabs/DiffRL/tree/main/dflex (https://github.com/NVlabs/DiffRL/tree/main/dflex)
7	https://ieeexplore.ieee.org/document/6386109 (https://ieeexplore.ieee.org/document/6386109)

Completion Criteria

Implementations of value or policy learning algorithms using a differentiable simulator. Test and compare results on physical (e.g. robotics) tasks in simulation.

Additional Information

Difficulty
3 - Hard

Popularity
5 students are interested in this project.

Desirable Skills
robotics, machine learning, simulation

Essential Skills
reinforcement learning

Unregister Interest | Back to the dashboard (/ug4)