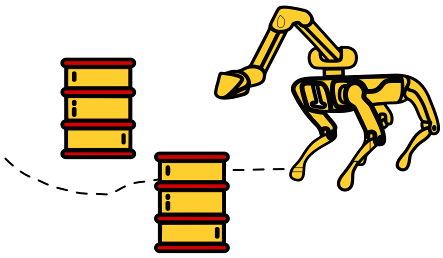
Online Safe Locomotion Learning in the Wild

Project Proposal for a Master Thesis

2024-05-03

Introduction



Reinforcement learning (RL) can potentially solve complex problems in a purely data-driven manner [1]. Still, the state-of-the-art in applying RL in robotics, relies heavily on high-fidelity simulators [2],[3]. While learning in simulation allows to circumvent sample complexity challenges that are common in model-free RL, even slight distribution shift ("sim-to-real gap") between simulation and the real system can cause these algorithms to easily fail. Recent advances in model-based reinforcement learning [4] have lead to superior sample efficiency, enabling online learning *without a simulator* [5]. Nonetheless, learning online cannot cause any damage and should adhere to safety requirements (for obvious reasons). The proposed project aims to demonstrate how existing safe model-based RL methods [6] can be used to solve the foregoing challenges.

Project Goals

The project aims to answer the following research questions:

- *How to model safe locomotion tasks for a real robotic system as a constrained RL problem?*
- *Can we use existing methods such as the one proposed by [6] to safely learn effective locomotion policies?*

Answering the above questions will encompass hands-on experience with a real robotic system (such as ANYmal) together with learning to implement and test cutting-edge RL methods. As RL on real hardware is not yet fully explored, we expect to unearth various challenges concerning the effectiveness of our methods in the online learning setting. Accordingly, an equally important goal of the project is to accurately identify these challenges and propose methodological improvements that can help address them.

A starting point would be to create a model of a typical locomotion task in Isaac Orbit [3] as a proof-of-concept. Following that, the second part of the project will be dedicated to extending the proof-of-concept to a real system.

Supervision

If you are a Master's student with

- basic knowledge in reinforcement learning, for instance, by taking *Probabilistic Artificial Intelligence* or *Foundations of Reinforcement Learning* courses;
- strong background in robotics and programming (`C++, ROS`),

please reach out to Yarden As or Chenhao Li. Feel free to share any previous materials, such as public code that you wrote, that could be relevant in demonstrating the above requirements.

References

[1] D. Silver *et al.*, “Mastering the game of Go with deep neural networks and tree search,” *Nature*, vol. 529, no. 7587, pp. 484–489, Jan. 2016, doi: 10.1038/nature16961.

[2] J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter, “Learning quadrupedal locomotion over challenging terrain,” *Science Robotics*, vol. 5, no. 47, p. eabc5986, 2020, doi: 10.1126/scirobotics.abc5986.

[3] M. Mittal *et al.*, “Orbit: A Unified Simulation Framework for Interactive Robot Learning Environments,” *IEEE Robotics and Automation Letters*, vol. 8, no. 6, pp. 3740–3747, Jun. 2023, doi: 10.1109/lra.2023.3270034.

[4] D. Hafner, T. Lillicrap, J. Ba, and M. Norouzi, “Dream to Control: Learning Behaviors by Latent Imagination.” 2020.

[5] P. Wu, A. Escontrela, D. Hafner, K. Goldberg, and P. Abbeel, “DayDreamer: World Models for Physical Robot Learning.” 2022.

[6] Y. As, I. Usmanova, S. Curi, and A. Krause, “Constrained Policy Optimization via Bayesian World Models.” 2022.