MA/SA/BA

Safe Quadruped Locomotion: Out-of-distribution Detection and Recovery

Background

Deep RL locomotion policies are strong in-distribution, yet under large pushes, slips, sensor glitches, or unseen terrains they can flip or thrash, risking damage to the robot itself and nearby people/objects. As legged robots move from labs into homes, offices, plants, and the outdoors, safety becomes non-negligible [1].

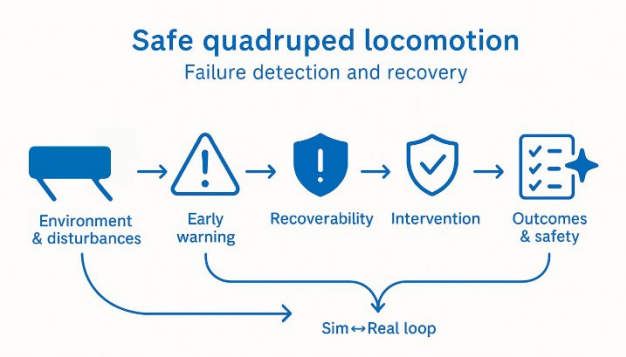
Building on this motivation, prior work explores two complementary angles: (i) training-time safe RL and constrained policy optimization (e.g., CPO/Lagrangian variants) that aim to satisfy constraints while learning; and (ii) deployment-time safety filters from control (e.g., HJ reachability, control barrier functions, predictive filtering), which minimally adjust commands to keep the system within a safe set. While theory and synthetic results are strong, applications to legged locomotion remain comparatively limited, leaving open questions that are central to real-world deployment: Can we detect out-of-distribution or failure-bound states with enough lead time to act? Where is the practical boundary between recoverable and unrecoverable states (a viability margin), and how can we estimate distance to it online? When should control strategy switch from the nominal controller to a safety/recovery mechanism, balancing incident reduction against performance?

Your Tasks

Through this project, you will gain hands-on expertise in state-of-the-art reinforcement learning algorithms and frameworks.

Your tasks include:

- Perform an in-depth literature review on safe reinforcement learning, constrained policy optimization and their application on quadruped locomotion tasks.
- Explore the potential solution space and develop at least one safety mechanism.
- Build the simulation evaluation framework.
- Investigate sim-to-real transfer to our **Unitree Go2** robot.



Conceptional illustration of safe quadruped locomotion

Requirements

- High motivation and interest in robot learning & control.
- Six months working time.
- Background in robotics, deep learning and reinforcement learning is an advantage.
- Programming skills (Python, PyTorch).

Supervisor: Prof. Alois Knoll.

Advisor: Yulong Xiao, yulong.xiao@tum.de

Lehrstuhl für Robotik, Künstliche Intelligenz und Echtzeitsysteme, TUM School of Computation, Information and Technology

[1] S. Ha, J. Lee, M. van de Panne, Z. Xie, W. Yu, and M. Khadiv, "Learning-based legged locomotion; state of the art and future perspectives," Jun. 03, 2024, arXiv: arXiv:2406.01152. Accessed: Oct. 02, 2024. [Online]. Available: http://arxiv.org/abs/2406.01152

*Unitree Go2 picture from https://www.unitree.com