Reinforcement Learning for Dynamic Stability in Robotic Systems

Alli Ajagbe, Noyonica Chatterjee, Surabhi Tannu

Objective:

The primary objective of our project is to develop and implement a robust reinforcement learning (RL) algorithm capable of ensuring dynamic stability in various robotic systems. Regardless of the disturbance or unbalancing scenario, our algorithm aims to enable robots to swiftly adapt and maintain their balance autonomously. This capability is crucial for a wide range of robotic applications, from robotic arms utilized in manufacturing processes to autonomous bots navigating unpredictable environments, where self-dependence for stability control is paramount.

Methodology:

Our methodology entails a comprehensive exploration of diverse RL algorithms, encompassing both model-free and model-based approaches, as well as on-policy and off-policy methods. Through extensive experimentation and analysis, we intend to identify the most effective algorithms for achieving dynamic stability in robotic systems. Furthermore, we plan to investigate hybrid models that integrate key components from multiple algorithms, aiming to leverage the strengths of each approach. To evaluate the performance of these algorithms, we have defined a set of metrics including convergence rate, training time, stability achievement, and robustness to novel scenarios. These metrics will be informed by thorough literature reviews and supported by relevant research findings in the field.

Timeline:

We broke down the project timeline down to 10 weeks-

- Week 1-2- For the first couple of weeks we plan to conduct a focused literature review on dynamic stabilization and reinforcement learning algorithms.
- Week 3-4 -Then we begin the development of the reinforcement learning models, targeting basic dynamic stabilization within the Gym Environment provided by OpenAI.
- Week 5-6 Begin development of the model targeting basic stabilization tasks with the environment. Start the preliminary testing to assess the models' performance and like find scopes of improvement
- Week 7-8 We plan to conduct in-depth testing and evaluation of the model comparing its performance against predefined metrics. Furthermore, we will work on optimization and seeing the robustness of these stabilization tasks.
- Week 8-10 We plan to test the model and create a possible simulation for the same, further conducting extensive testing and documenting all outcomes.

Plan of Action:

Our project plan will involve each member actively contributing to various aspects of the project. This entails parallel work on different RL algorithms, fostering a dynamic exchange of ideas and results during weekly team meetings. To streamline our development process and facilitate seamless code sharing and updates, we will maintain a dedicated GitHub repository, leveraging existing functionalities and repositories wherever possible. We also plan to meet with the professor and the TAs either weekly or biweekly, depending on the updates we would like to share regarding work done or doubts that need to be clarified.

Literature:

Some of the identified pieces of literature for our reinforcement learning use case are highlighted below:

- M. Posa, S. Kuindersma, and R. Tedrake, "*Optimization and Stabilization of Trajectories for Constrained Dynamical Systems*," in Proceedings of the International Conference on Robotics and Automation (ICRA), Stockholm, Sweden, 2016.
- Ao Xi, Chao Chen, "*Stability Control of a Biped Robot on a Dynamic Platform Based on Hybrid Reinforcement Learning*", Sensors 2020, 20(16), 4468.
- Lucian B, Tim B, et al, "Reinforcement Learning for Control: Performance, Stability, and Deep Approximators", Annual Reviews in Control, 2018.

Environment (If Any):

Since this is largely a software-based project, we intend to leverage OpenAI's Gym Environment for the CartPole balancing to estimate the performance of the algorithms we will be working with. This environment will provide us with a rough idea of how these algorithms will perform in comparison to others in real-life situations.