Improving the Understanding of Neuromuscular Gait Control using Deep Reinforcement Learning

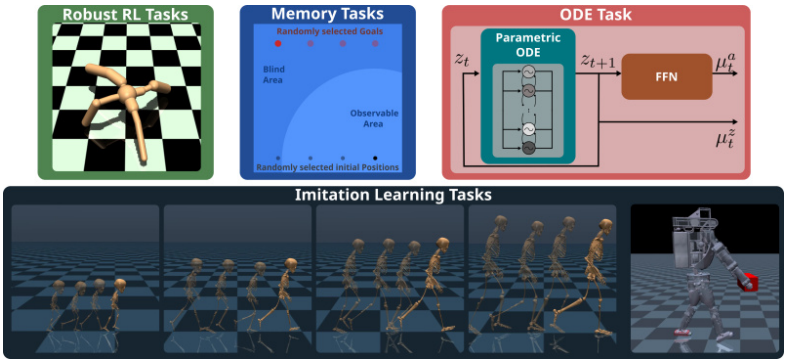

Figure 1: An overview of the tasks addressed in this project

Project Manager
Firas Al-Hafez

Researchers
Dr. Davide Tateo

Principal Investigator
Dr. Guoping Zhao

Project Term
2022 - 2023

Clusters
Lichtenberg Cluster Darmstadt

Software
PyTorch

Institute
Intelligent Autonomous Systems

University
Technische Universität Darmstadt

Introduction

Understanding the underlying reward mechanisms of human locomotion and transferring this knowledge to humanoid robots is a critical challenge in biomechanics and robotics. This report synthesizes findings from three pivotal papers that contribute to this field. These studies not only advance robotics by improving humanoid robot control but also enhance our understanding of biomechanics through the development of both torqueactuated and muscle-actuated human models.

Methods

1. LocoMuJoCo: The Imitation Learning Benchmark LocoMuJoCo introduces the first imitation learning benchmark for humanoid robots, essential for understanding human locomotion. It includes a diverse array of environments, such as quadrupeds, bipeds, and musculoskeletal human models. The benchmark incorporates real retargeted human motion capture data, providing various types of datasets that include perfect states, with or without actions. This comprehensive dataset serves as a foundational tool for evaluating and comparing different imitation learning algorithms. LocoMuJoCo supports both the Gymnasium and Mushroom-RL libraries, facilitating the implementation of reinforcement learning (RL) methods. The environments are designed to test the robustness and adaptability of algorithms under varying conditions.

2. LS-IQ: Imitation Learning with Implicit Rewards The study addresses the challenge of imitation learning from observation, where traditional methods like behavioral cloning fail due to the absence of action data. Adversarial methods, although commonly used, are often difficult to train and prone to overfitting. LS-IQ introduces an innovative approach by utilizing the relationship between reward and action value functions in

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

reinforcement learning. Instead of fitting an explicit reward model, implicit rewards are derived, providing a stable learning process. The paper includes a thorough analysis of why previous methods were unstable and how to mitigate these issues.

3. S2PG: Learning with Stateful Policies Stateful policies are crucial for managing partial observability or incorporating inductive biases into the policy. The study uses recurrent neural networks (RNNs) and central pattern generators (CPGs) to model these policies. S2PG presents a new stochastic stateful policy gradient estimator, designed to train these policies more efficiently. The approach addresses the limitations of backpropagation through time (BPTT), which is traditionally used but cumbersome for stateful policies. The method allows for training policies that adapt to unobserved and changing dynamics, significantly improving the adaptability of humanoid robots. This is achieved through parallelized environment simulations and advanced policy optimization techniques.

Results

1. LocoMuJoCo

- Developed and validated a comprehensive benchmark environment using diverse human motion capture data.
- Established baseline performance metrics for various imitation learning algorithms, demonstrating the benchmark's utility in advancing research.

2. LS-IQ

- Formulated an implicit reward framework and demonstrated its theoretical stability.
- Empirical validation showed improved performance and stability over existing adversarial methods on standard benchmarks.

3. S2PG

- Introduced a stochastic stateful policy gradient estimator and validated its effectiveness.
- Successfully trained humanoid locomotion policies that adapt to changing dynamics, showcasing significant improvements in adaptability and performance.

Discussion

These studies collectively advance our understanding of human locomotion and its application to biomechanics and humanoid robots. LocoMuJoCo sets a new standard for benchmarking imitation learning algorithms, providing a robust and diverse dataset that has gained widespread acceptance in the research community. This benchmark is essential for developing more accurate models of human motion, which can inform the design of prosthetics and rehabilitation strategies.

LS-IQ addresses critical challenges in imitation learning from observation by introducing a stable and effective implicit reward

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

framework, thereby overcoming the limitations of previous methods. This approach is particularly valuable for biomechanics, where understanding the nuanced rewards associated with human movement can lead to better predictive models of human behavior.

S2PG offers a novel approach to training stateful policies, enhancing the capability of humanoid robots to adapt to dynamic environments and partial observability. The insights gained from this research can be applied to develop more adaptive and responsive prosthetic limbs, improving the quality of life for individuals with mobility impairments.

Publications

Al-Hafez, F.; Tateo, D.; Arenz, O.; Zhao, G. and Peters, J. LS-IQ: Implicit reward regularization for inverse reinforcement learning. In Proceeding of the International Conference on Learning Representations, Kigali, Rwanda, May (2023)

Al-Hafez, F.; Zhao, G.; Peters, J.; Tateo, D. LocoMuJoCo: A comprehensive imitation learning benchmark for locomotion. In 6th Robot Learning Workshop, NeurIPS, New Orleans, Louisiana, United States, December (2023b)

Al-Hafez, F.; Zhao, G.; Peters, J.; Tateo, D. Time-Efficient Reinforcement Learning with Stochastic Stateful Policies. In Proceeding of the International Conference on Learning Representations, Vienna, Austria, May (2024)

Last Update: 2024-07-09 11:20