High-Level Modeling, Imitation and Control for Locomotion, Autonomous Driving, and Resource Allocation

Project Manager
Dr. Oleg Arenz

Principal Investigator
Dr. Oleg Arenz

Project Term
2023 - 2025

Clusters
Lichtenberg II Cluster Darmstadt

Additional Software
MuJoCo, Isaac Sim, Simulink



Introduction

Reinforcement Learning (RL) and Imitation Learning (IL) enable intelligent systems to learn desired behaviors autonomously, typically by focusing on low-level actions. However, tackling complex real-world tasks requires a more human-like, high-level understanding, namely the ability to think abstractly and set future goals. This project is focused on developing new methods for RL and IL that allow agents to operate across these different levels of abstraction. The project is particularly concerned in learning abstract interfaces between the high-level goals and the low-level execution, often using latent representations that capture the essence of a skill or state. This approach enables us to solve challenging problems like teaching robots to walk with a human-like gait despite different physical forms, learning to imitate the specific driving style of a professional race car driver, and training smart policies to efficiently allocate resources for complex Extract-Transform-Load (ETL) data processing chains that require quick transferability across varying system details.

Methods

Our research centers around the encoding of desired robot motions using learnt abstract embeddings. This problem is central to developing smart agents capable of handling complex tasks, particularly in challenging physical domains like humanoid locomotion. To address this, we developed methods that use

Variational Autoencoders (VAEs) to encode complex behaviors, such as a walking gait, into concise, abstract representations. We are particularly interested in learning discrete representations suitable for modern transformer architectures. To that end, we created a novel VAE optimization that avoids typical gradient problems by borrowing techniques from policy search to effectively train a discrete encoder to select abstract codes. Complementary to this, we also explored using continuous VAE architectures to represent skills, where a high-level process, similar to Model Predictive Control (MPC), continuously optimizes the continuous latent code in real-time to dynamically steer the low-level execution. A separate, yet related, research direction focused on learning robust latent representations not for generating actions, but for modeling the system dynamics itself. We investigated using these representations for online system identification (SysID) of complex physical systems that are difficult to model exactly, using physics-informed grey-box models. Our method uses Deep Lagrangian Networks (DeLaN), which embed fundamental physics principles into the neural network structure. By learning a concise, low-dimensional latent dynamic representation, our model can quickly capture and adapt to residual, unmodeled complexities or shifts in parameters that occur during operation. This allows our robot to continuously and accurately adapt to the real-time changes of the system, enhancing prediction and control stability. When learning the low-level control with reinforcement learning for achieving high-level goals, we often also have low-level objectives, such as those given by demonstrations. For example in race driving, trading off the sparse RL targets (e.g., achieving the fastest lap time) and the coarse imitation learning objectives (e.g., matching the human driver's style) is a significant challenge. We propose a hybrid reinforcement/imitation learning method that adaptively trades of the conflicting objectives and applied it for evaluating different car setups. Finally, we tackled the challenge of transferable control in large-scale data systems. To enable the development of generalizable strategies, we built a high-fidelity simulator for Extract-Transform-Load (ETL) process chains. ETL chains, used for massive data migration, typically require manual and specific resource tuning. We used this simulator is to train a policy that can be quickly deployed on new or modified physical ETL systems with minimal retraining, demonstrating the power of abstract control in achieving rapid transferability.

Results

Our method for discrete variational autoencoding turned out to be effective not only for embedding motion trajectories, but also in other settings, for example, for encoding images of the ImageNet dataset. We also demonstrated on our continuous autoencoder that we can optimize the latent representation online to adapt locomotion trajectories of a quadruped robot for avoiding obstacles. We also evaluated our method for online latent system identification on a robot manipulator and demonstrated that our adaptive dynamics models can be used for precise model based tracking of trajectories, despite unexpected changes in the payload during execution. Our

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

method for hybrid reinforcement/imitation learning was able to outperform the lap time of a professional driver, while imitating their driving line. Backward tests show, that the simulation with our learned policy provide insights for setup tuning that are comparable with more expensive tests that require human drivers. For modelling ETL process chains, we developed a highly-parallelizable simulator and demonstrated that reinforcement learning policies can outperform heuristic approaches for dynamic resource allocation.

Discussion

This project successfully established the viability of abstract representations as a unifying concept for hierarchical control and modeling in complex domains. The robust performance achieved across motion encoding (using VAEs for both discrete and continuous planning) and online system identification (using latent DeLaN models) demonstrates that focusing on the right level of abstraction is fundamental to realizing high-performance, generalizable intelligent agents. The capability to successfully mediate conflicting high-level and low-level objectives (e.g. lap time vs driving line imitation) is crucial for learning control policies that realize high level objectives. Furthermore, the development of an abstract resource allocation policy for the ETL simulator, which uses state history to inform decisions, points toward its potential integration with online and adaptive system identification methods for robust real-world deployment. Our next steps will focus on developing generalized latent spaces capable of encoding skills transferable across different embodiments, and on rigorously validating the long-term integration of these abstract models and policies into full, robust, real-world control systems.

Last Update: 2026-02-03 09:54