Hessisches Kompetenzzentrum
für Hochleistungsrechnen

Active Exploration for Robotic Manipulation

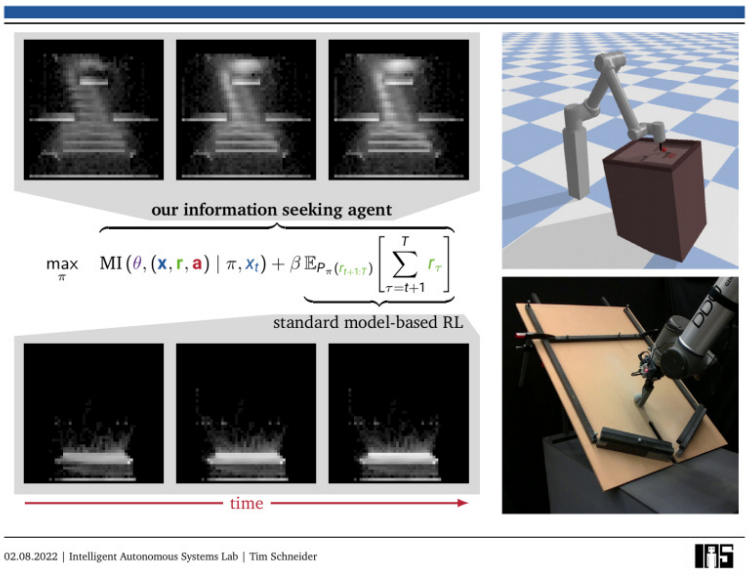

Figure 1: Overview of our method. The crucial difference to other model-based reinforcement learning methods is the addition of the mutual information (MI) term that encourages information seeking behavior.

Introduction

Robotic manipulation stands as a largely unsolved problem despite significant advances in robotics and machine learning in recent years. One of the key challenges in manipulation is the exploration of the complex contact dynamics of the environment, for which usually no accurate model is known apriori. In this work, we propose a model-based active exploration approach that enables efficient learning in sparse-reward robotic manipulation tasks. The proposed method estimates an information gain objective using an ensemble of probabilistic models and deploys model predictive control (MPC) to plan actions online that maximize the expected reward while also performing directed exploration. In this project, we evaluate our proposed algorithm in simulation on a challenging ball pushing task on tilted tables, where the target ball position is not known to the agent a-priori. Since our method relies heavily on GPU-accelerated deep neural network model learning, the use of appropriate hardware is imperative for its performance in these simulations. Our experiments serve as a fundamental application of active exploration in model-based reinforcement learning of complex robotic manipulation tasks.

Methods

In this work, we aim to solve sparse-reward Markov Decision Processes in a robotic manipulation setting. Our key contribution

Project Manager
Tim Schneider

Researchers
Prof. Dr. Georgia Chalvatzaki, Dr. Boris Belousov, Dr. Diego Romeres and Dr. Devesh K. Jha

Principal Investigator
Prof. Jan Peters (PhD)

Project Term
2021 - 2022

Clusters
Lichtenberg Cluster Darmstadt

Software
PyTorch

Additional Software
PyBullet

Institute
Intelligent Autonomous Systems

University
Technische Universität Darmstadt

Partners
MERL - MITSUBISHI ELECTRIC RESEARCH LABORATORIES

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

is a novel model-based reinforcement learning algorithm that approaches the problem of model learning from a Bayesian Optimal Experiment Design perspective. This algorithm uses data it gathers from the environment to learn an ensemble of neural network dynamics models. During roll-outs, these models are used to plan the agent's future behavior with a Model Predictive Control algorithm on the fly. However, instead of maximizing the reward signal directly, like many prior approaches do, we instead maximize the sum of the reward and an information gain term, that we estimate using the model ensemble. The advantage of this scheme is that the information gain term drives the agent to explore regions of the state space is hat not yet explored, while the reward term drives the agent to maximize its task performance. A central challenge in the implementation of this algorithm is that it has to be capable of running in real time to be applicable in the real world. Thus, we implemented the entire planning algorithm to run on a GPU, which allows us to compute new plans at a rate of approximately 5Hz. For our implementation, we require machines with NVIDIA CUDA capable GPUs, as they are present on the HHLR. The main use of the HHLR in this project was to conduct extensive experiments and hyperparameter search, which would not have been possible without a substantial amount of GPU resources.

Results

We evaluate our approach on three (two simulated and one real-world) challenging, sparse-reward manipulation tasks, in which the agent has to balance a ball on its finger to a goal position. What makes these tasks challenging is that the agent does not receive any information about its objective a-priori and instead has to explore the environment systemtically to find the goal. In our experiments, unlike the baselines, our method succeeds in finding the goal in all three tasks and proceeds to develop a robust strategy to solve them. It does so by exploring the environments systematically and therby solving these manipulation task without a dense extrinsic reward, but rather driven by its own curiosity.

Discussion

In this work, we developed an active exploration method that is capable of solving complex robotic manipulation tasks. Our main algorithmic contribution is the introduction of an information seeking strategy in model-based reinforcement learning, that balances between exploration of new states in the environment to improve the dynamics model and task performance. In our experiments we showed that our method can solve challenging exploration problems with sparse rewards in a robotic manipulation setting. Considering that none of the baselines were able to solve our tasks, we conclude that the information-seeking behavior of our agents is beneficial for solving such kind of problems Furthermore, our real-world experiments demonstrate the suitability of our method for learning complex manipulation tasks in the real world. In the future, we plan to incorporate tactile sensors into our setup. Tactile sensors would allow one to obtain more detailed feedback about the objects

being manipulated. Considering that humans deploy a variety of active haptic exploration strategies during manipulation, research on robotic active tactile exploration might bring us closer to human-level manipulation skills. One of the main limitations of our method that prevents it from being used on more dynamic tasks is that it is computationally comparably heavy and therefore limited to relatively slow tasks. Hence, an exciting future research direction is to tackle this issue with hierarchical controllers by combining our planning module with a learned low-level balancing controller running at a high frequency.

Publications

Schneider, T.; Belousov, B.; Chalvatzaki, G.; Romeres, D.; Jha, D. K.; Peters, J.: "Active Exploration for Robotic Manipulation", 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE (2022). https://doi.org/10.1109/IROS47612.2022.9982061

Tim Schneider, 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems, Active Exploration for Robotic Manipulation, Kyoto, Japan, 25.10.2022

Last Update: 2023-06-05 15:25