Deep Reinforcement Learning under Uncertainty



Project Manager
Dr. Joni Pajarinen

Principal Investigator
Prof. Jan Peters (PhD)

Project Term
2018 - 2019

Project Areas
Computer Science

Institute
Intelligent Autonomous Systems

University
Technische Universität Darmstadt

Introduction

Traditionally robots have been used in factories in predefined structured environments. Recently, robots are employed more and more in unstructured partially unknown environments such as homes. In order to adapt to new environments and to perform tasks which require advanced skills, robots need to learn to cope with uncertainty and partial observability. To achieve this, robots need more efficient computational methods for learning directly from environmental stimuli. However, learning to perform advanced skills directly from observations is a challenging task. In this project, we investigated new techniques in deep reinforcement learning that allowed the robot to learn to perform tasks based on trial and error. In particular, we investigated how the robot can learn tasks which require long term planning and how to perform information gathering for accomplishing the assigned tasks. A computation cluster was needed because of the heavy computations in deep reinforcement learning.

Methods

The project will focus on method development in deep reinforcement learning under partial observability which is an unsolved crucial challenge. The methods optimize policies for agents such as robots or autonomous vehicles. A policy executes an action which modifies the (stochastic) world, makes an observation about the world, and receives a reward signal. The computational challenge comes from

1. simulating the world for each executed action, that is, generating samples, and

2. policy optimization based on the collected samples.

Results

We were able to develop a new policy search algorithm that can optimize policies under partial observability and yields improved performance w.r.t. the comparison methods.

Discussion

In the future, we will investigate improved memory representations, exploration, and how to guide policy optimization in a more efficient manner. The algorithms used will include gradient based policy search methods such as proximal policy optimization (PPO), Trust Region Policy Optimization (TRPO), Truncated Natural Policy Gradient, REINFORCE (VPG), and Reward-Weighted Regression (RWR). Moreover, we will also use methods such as deep Q-learning in discrete action tasks. We will also compare our results with well known gradient-free black box optimization methods such as the Cross Entropy Method (CEM) and Covariance Matrix Adaption Evolution Strategy (CMA-ES). We will develop new computational methods in sub-projects 1 - 4. In sub-project 1, we focus on the memory representation of autonomous agents. This requires evaluating methods with long-short term memory (LSTM) and related neural network techniques.

Publications

Pajarinen, J.; Thai, H. L.; Akrour, R.; Neumann, G.: Compatible Natural Gradient for Deep Reinforcement Learning, Mach Learn 108, 1443–1466 (2019) https://doi.org/10.1007/s10994-019-05807-0

Last Update: 2022-04-22 16:45