Integrated Project

Robot Learning
Today’s agenda!

- Sales Pitch of the IP
- Project Timeline
- Project Proposals
Sales Pitch
Among the most important questions ever: continue the research road to a Ph.D. (=Dr.)?

The personal and professional advantages are enormous!

An exciting life:

• follow your ideas & dreams...
• actively acquire knowledge and refine it…
• enjoy international conferences and visits with collaborators around the world…

• However, it ain’t for everybody!

• Your Master’s thesis will already decide on your chances!

• Do you wanna figure out whether there is a researcher in you?
A survey on the foundations of robust adversarial reinforcement learning

Janoesch Moos
Department of Computer Science
TU Darmstadt
Darmstadt, 64289
janoesch.moos@stud.tu-darmstadt.de

Key Benset
Department of Computer Science
TU Darmstadt
Darmstadt, 64289
key.benset@stud.tu-darmstadt.de

Abstract
Reinforcement learning algorithms are known to struggle with robustness and generalization to environment changes in terms of uncertainty and parameter perturbations. Different methods have been proposed to adapt approaches of robust perturbations. In this paper, we show how reinforcement learning can be adapted to reinforce learning. In this paper, we show how reinforcement learning can be adapted to reinforce learning. We discuss the transition from finite-horizon to infinite-horizon reinforcement learning and how it is possible to use reinforcement learning in situations with multiple agents and Markov games. Finally, we explain the adversarial reinforcement learning algorithm and how it can be adapted to multi-agent reinforcement learning. This approach has proven to increase the performance of reinforcement learning across different test scenarios and reduce the impact of parameter perturbations between training and test scenarios as well as simulation and real world.
Theoretical Investigations

Likelihood-free Inference in Reinforcement Learning: Relation to REPS and Sim-to-Real Applications

Yu Cui
Department of Computer Science
TU Dortmund
Hochschulstraße 10, 44227 Dortmund
yucui@cs.tu-dortmund.de

Marin Dusanić
Department of Computer Science
TU Dortmund
Hochschulstraße 10, 44227 Dortmund
marin.dusanic@tu-dortmund.de

Abstract

The difficulty of transferring optimized policies from simulation to reality (Sim2Real gap) forms the Achilles heel of reinforcement learning (RL) to real-world problems. Prior work addresses the Sim2Real gap by optimizing policy parameters. However, likelihood-free inference is intrinsically more challenging, as it seeks to infer a distribution over a set of parameters from observations. For every practical problem of likelihood-free inference, there are several approaches, such as Approximate Bayesian Computation (ABC) and Expectation-Maximization (EM). In this paper, we show connections between ABC and likelihood-free inference algorithms. By directly connecting between methods, we lay down a theoretical foundation for theoretical guarantees in ABC and theoretical results for likelihood-free inference algorithms. In particular, we demonstrate how ABC and likelihood-free inference algorithms are related. Additionally, we propose an algorithm that combines the benefits of both methods.

3 Connections between ABC, VI and ARB methods

When we choose the reward function of REPS to be a discrepancy $d$ similar to the one in ABC, we can observe similarities in what both algorithms are doing.

Both REPS and ABC sample from a prior, evaluate the quality of the samples, ABC keeps a collection of samples with discrepancy $d$ under a hard threshold, REPS weights the samples by their respective discrepancy. To discretize a continuous and therefore the ELBO can be evaluated directly. With $R(x) = \log p(x | \theta)$ and

$$p_{\theta}(x | \theta) \sim p_{\theta}(\theta) \exp \left( \frac{R(x)}{\eta} \right) \approx p_{\theta}(\theta) \frac{p_{\theta}(x | \theta) \beta}{\eta}$$

$\eta$ is determined by the KL-covariance of REPS and influences the weighting of prior and likelihood so that the step size is limited. In ABC, $p_{\theta}(x | \theta) \sim c$ is concave in $p_{\theta}(x | \theta)$ for $\eta \geq 1$, and convex for $\eta \leq 1$. The choice of $\eta$ affects the likelihood in favor of the likelihoods. Therefore, processing the observations is equivalent to processing all observations at once.

Sim-to-Real gap: ABC can substitute the log-likelihood $-d(x, \theta)$ with some discrepancy $d$ on observations $x$.

$$d(x, \theta) = \begin{cases} 0 & \text{if } x = x, \\ \infty & \text{else}. \end{cases}$$

For real observations $x$.
Implement Algorithms

Development and Implementation of an Approach for Multi-Object Tracking in the Context of Autonomous Driving

Tomáš Potoček
Intelligent Autonomous Systems
TU Darmstadt, Germany
tomas.potoccek.tu-darmstadt.de

Abstract

As the level of automation in vehicles increases, there is the need for a decision-making system that can operate autonomously in increasingly complex scenarios such as crowded urban or heavy traffic situations. Resolving the dynamics of moving objects in the environment in real-time is, therefore, a crucial component to enable autonomous driving vehicles. In this work, it is presented a unified framework of multi-object detection and tracking using 3D LIDAR, where detected clusters in the point cloud are used for tracking an unknown number of objects in the scene using a GM-PDO filter. The evaluation results using the KITTI dataset on ROS environment shows that the proposed framework for multi-object tracking can achieve promising real-time performance on complex urban scenarios.
Enhancing Intention Aware Movement Primitives

Abstract

In the future, robots will not only operate at a safe distance from humans but also have a closer social contact with them. To provide a comfortable human-robot cooperation, the robot is desirable to dynamically adapt its movements to not interrupt the workflow of its partner. The goal of this work is thereby to allow the robot to infer the workflow of its partner. For this, the robot is expected to be able to infer the partner's intention and predict its future movements. Based on a few early observations of human motion, probabilistic movement primitives (ProMPS) provide a probabilistic framework to model uncertainty and inherent correlation of human movement. In this work, we use a neural network ProMPS to predict human intentions in terms of most likely future trajectories based on the observed data. Compared to prior work that relied on linear motion models, the learned data can thereby improve the accuracy and stability of real-time motion prediction. In particular, we compare two different approaches of incorporating ProMPS to the prediction of trajectories with partial observation: first, a method which conditionally models the weight distribution of learned ProMPS to match the early observations, and second, an Expectation-Maximization based algorithm on observed data points. Experiments on trajectories show a better performance of the first approach in terms of prediction accuracy on recorded human motion data of 25 subjects compared to ground truth that can already be made after observing 30 percent of trajectory points.

Keywords: human-robot cooperation, trajectory prediction, intention-aware ProMPS

4 Experimental evaluation

In this section, we first describe the setup of experiments and the task performed by a human and the robot. In the experiments, we performed experiments with a human and the robot in the experiments of a user study conducted by Koen et al. [9]. After that, we present and discuss the results of different methods for phase-detection and trajectory prediction evaluated on the recorded data collected in the user study.

4.1 Experimental setup

Figure 2: (a) Illustration of the experimental setup where a human and the robot perform a pick-and-place task in a shared workspace. Adapted from [9]. (b) Comparison of estimated phases of a trajectory from student (blue), 3,3,2, and optimization method (red) 3,3,2.

In the experiment environment illustrated in Figure 2a, a human and the robot perform a pick-and-place task while sharing the workspace. The task of a human seated at A is to remove a block from B and deliver it to C. During the task, the human is allowed to use only one hand to carry the block, the robot is allowed to use one hand to carry the block, and the robot is not allowed to use the right hand to carry the block. On this hand, optical-tracking system. The task of the robot is to deliver the block from D to E. During the task, the robot is allowed to move freely to any location in the workspace without any restriction. The robot is required to perform the task without any restrictions. In the experiments, the robot is not required to perform the task without any restrictions.

The task is completed when the robot has moved to the correct location of the block, and the human have delivered the block to the correct location. The task is completed when the robot has moved to the correct location of the block, and the human have delivered the block to the correct location.
Write a Scientific Paper
Do a Mini-Conference!
What We Offer....

We offer you a glimpse how life as a researcher in robot learning is like

- Use the knowledge from the robot learning lecture right away
- Decide what problem you are interested in and implement it in our simulator
- Write a scientific paper with a team
- Have a mini-conference at the semester’s end
- Good projects can be continued as a Masters or Bachelors theses

- You are trying out how research life is like!
Background (technical) Knowledge

• Is very project-dependent

• But might help if you have:
  • Mathematics from the first semesters (calculus, statistics)
  • Programming (project dependent, usually C/C++, Python)
  • Computer science fundamentals (algorithms)

• Simultaneous or previous attendance of the Statistical Machine Learning and Robot Learning lectures is very helpful!

• Most important is that you have a wish to learn new topics!
The Timeline
The Timeline

1. Choose a project, email an IP coordinator and supervisors (Until 23.10.2022)

2. Topic Assignment (30.10.2022)

3. Work on the project…

4. Write up results into a paper (06.03.2023)

5. Peer review (15.03.2023)

6. Final submission (24.03.2022)

7. Presentation to the group (mini-conference) (28.03.2023)
The Projects
Memory-Free Continual Learning

**Supervisor:** Ahmed Hendawy, Carlo D’Eramo

**Motivation:**
- Can Deep Learning (DL) models *continually* learn?
- Once new task is encountered → DL models *forget* the learned tasks
- Absence of the previous tasks’ data → Parameter shift
- Why do not we replay the old data? → Expensive or Privacy

**Objective:**
- Alleviate the forgetting *without* saving old data in the raw format (e.g. images).
- Model the old data as a distribution [1], or a subspace [2].
- Supervised learning → Reinforcement Learning (RL) (optional)

**Requirements:**
- Good programming skills in Python
- Prior knowledge in DL and RL

Learn to play Tangram

Supervisor: Kay Hansel, Niklas Funk

Solving the game of Tangram is challenging:

• Combinatorial burden
• Requires high- and low-level reasoning
• Approach should generalize to different shapes

Goal:

• Exploit a combination of powerful graph-based representations (GNNs) & (multi-agent) RL for solving the task

Tactile Environment Interaction

Supervisor: Niklas Funk

Sense of touch is essential, and has the potential to severely improve robotic systems through enabling environment interaction!

Idea:
- Exploit vision-based tactile sensor [1,2] mounted on robot to interact with objects

Goal:
- Create classification pipeline that outputs whether object is static or moveable

Learning Bimanual Robotic Grasping

Supervisors: Julen Urain, Alap Kshirsagar

**Motivation:**
Bimanual grasps are required for manipulation of large, deformable, fragile objects

**Goals:**
Learn bimanual robotic grasps from dataset of dual-arm grasps [1]

**Requirements:**
Good programming in python,
Prior knowledge of RL/IL (optional)

Kinematically Constrained Humanlike Bimanual Robot Motion

Supervisors: Alap Kshirsagar, Vignesh Prasad

**Motivation and Objective:**
- While there have been advances towards imitation learning with movement primitives, for bimanual object carrying tasks, enforcing (Inverse-)Kinematic constraints for accurate behaviours is needed
- Objective: Explore methods for task space Inverse-Kinematic adaptation of learnt object carrying behaviours

**Project Goals:**
- Study literature on bimanual robot motion generation and task space adaptation of Learning-by-Demonstration approaches
- Implement and train models on an existing bimanual motion dataset for baseline behaviours
- Incorporate Inverse Kinematic task space adaptation when applying the learnt behaviours on a robot

**Requirements**
- Good programming skills in Python
- Basic Experience with AI/Machine Learning

Characterizing Fear-induced Adaptation of Balance by IRL

Supervisor: Alap Kshirsagar, Firas Al-Hafez

Motivation:
- Fear of falling has been found to correlate with increased sway in elderly adults [1]
- Q: How do computational goals underlying balance control change under fear?

Goal:
- Use Inverse Reinforcement Learning to infer human postural control goals

Requirements:
- Good Python programming, hands on experience with RL (preferred), experience with simulators (e.g., MuJoCo) (preferred)

Interactive Semi-Supervised Action Segmentation

Supervisors: Lisa Scherf, Vignesh Prasad, Felix Kaiser

Motivation and Objective:
- To transfer behaviours from humans to robots, understanding underlying actions in a task enables learning in a modular fashion and allows non-experts to teach robots
- Unsupervised Action Segmentation achieves this and needs low manual effort but fails to find ideal semantically relevant clusters [1, 2]
- Objective: Use human interactive input to improve the Action Segmentation in a semi-supervised manner.

Project Goals:
- Explore Literature on Un-/Semi- Supervised Action Segmentation and benchmark different approaches on an existing dataset.
- Develop a User Interface for interactive feedback.
- Improve the implemented algorithms iteratively with user feedback.
- Transfer learned demonstration to a robot.

Requirements
- Good programming skills in Python
- Basic Experience with AI/Machine Learning

References:
[3] https://telekinesis.ai
System identification and control for Telemax manipulator

**Supervisor:** Junning Huang, Davide Tateo

**Motivation:**
- System of Telemax is hard to identify: weird dynamics exists because of the joystick control.
- Control the manipulator is hard: the system is not a control-affine system, the dynamics is non-linear with respect to control input.

**Goal:**
- Grey-box model system identification for Telemax robot
- Learning to control for the Telemax manipulator with reinforcement learning

**Requirements**
- Good programming skills in Python
- Basic Experience with reinforcement learning, nonlinear system identification and control

[1] End-to-End Learning of Hybrid Inverse Dynamics Models for Precise and Compliant Impedance Control
Tactile Active Exploration of Object Shapes

Supervisor: Tim Schneider, Boris Belousov, Alap Kshirsagar

How do humans/robots utilize the sense of touch to understand object properties?

Idea:
- Start with a baseline on contour following [1]
- Exploit active inference with tactile data [2]

Goal:
- Generate explorative motions of the robot
- Compare to known strategies of humans


Homework (due end of day Sunday, 23.10.2022)

Write a short paragraph to answer the following questions:

1) Which project would you like to try and why?

2) Why do you think this project is important?

3) What helpful background do you have for the project and what makes you special for that project?

4) Your academic aspirations: 1 semester? 2 semesters? Future thesis?

The participants can only send two such proposals to our PhD students. Please specify your priority for the two projects.

If you already have a group, please send a joint email.

~Email niklas.funk@tu-darmstadt.de (cc) + supervisors with proposals~

After a meeting with the potential supervisor(s), topics will be assigned to students. Unfortunately, some students might not get topics!