Integrated Project

Robot Learning

TECHNISCHE UNIVERSITÄT DARMSTADT

Today’s agenda!

- Sales Pitch of the IP
- Project Timeline
- Project Proposals

Sales Pitch

A Semester in the Life of a PhD!

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:

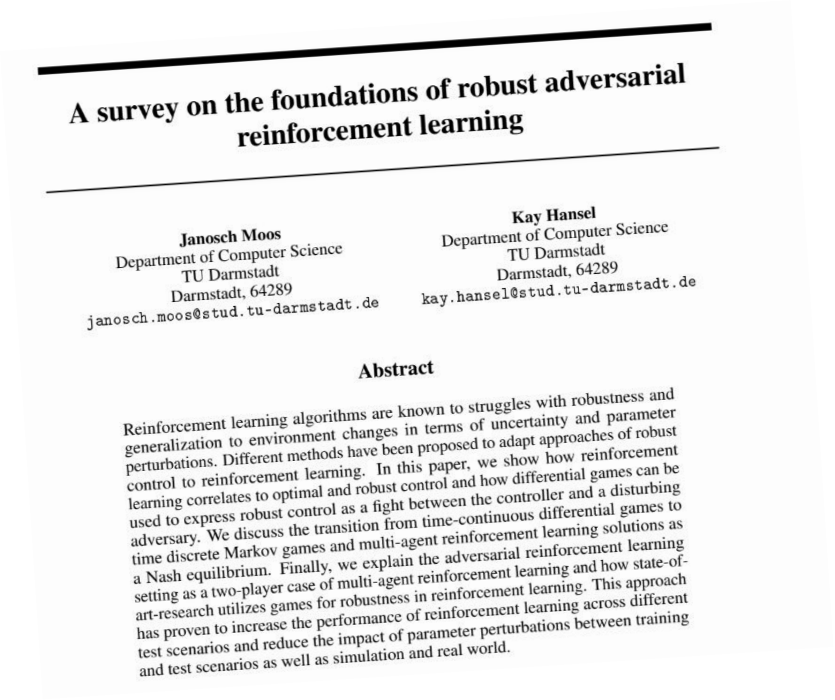
- follow *your* ideas & dreams...
- actively acquire knowledge and refine it…
- enjoy international conferences and visits with collaborators around the world…

- 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?

Literature Review



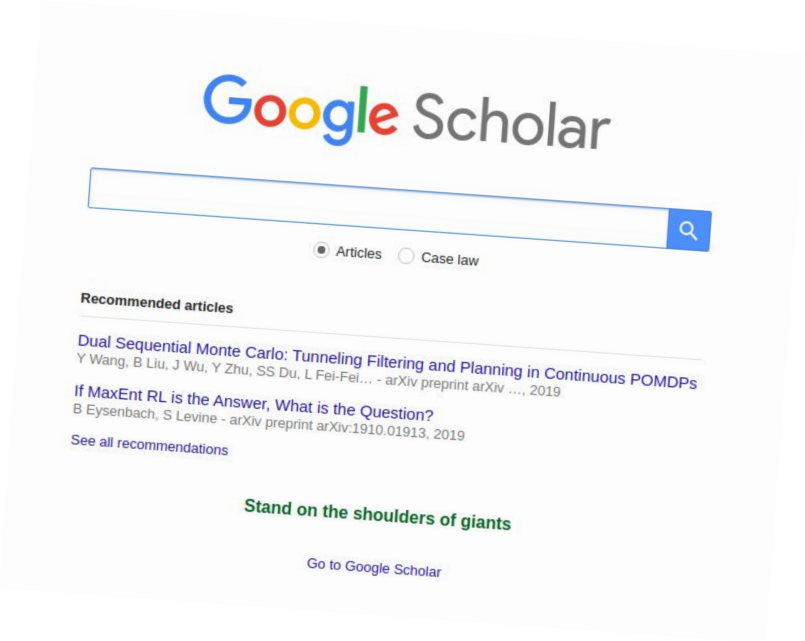
A survey on the foundations of robust adversarial reinforcement learning

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

Kay Hansel
Department of Computer Science
TU Darmstadt
Darmstadt, 64289
kay.hansel@stud.tu-darmstadt.de

Abstract

Reinforcement learning algorithms are known to struggles with robustness and generalization to environment changes in terms of uncertainty and parameter perturbations. Different methods have been proposed to adapt approaches of robust control to reinforcement learning. In this paper, we show how reinforcement learning correlates to optimal and robust control and how differential games can be used to express robust control as a fight between the controller and a disturbing adversary. We discuss the transition from time-continuous differential games to time discrete Markov games and multi-agent reinforcement learning solutions as a Nash equilibrium. Finally, we explain the adversarial reinforcement learning setting as a two-player case of multi-agent reinforcement learning and how state-of-art-research utilizes games for robustness in 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

Kai Cui *
Department of Computer Science
TU Darmstadt
Hochschulstraße 10, 64289 Darmstadt
kai.cui.de@gmail.com

Maximilian Hensel *
Department of Computer Science
TU Darmstadt
Hochschulstraße 10, 64289 Darmstadt
maxi.hensel@gmx.de

Abstract

The difficulty of transferring optimized policies from simulation to reality (Sim-to-Real gap) limits the applicability of reinforcement learning (RL) to real-world problems. Prior work addresses the Sim-to-Real gap by optimizing policy and simulation parameters alternatingly. Since the likelihood function of the simulator is intractable, prior work employed episodic relative entropy policy search (REPS) – a form of black box optimization – to implicitly find an approximate posterior over simulator parameters. The core problem of likelihood free inference arises in many other areas such as computational biology or population genetics and is usually solved by different approaches such as Approximate Bayesian Computation (ABC). In this paper, we draw connections between stochastic optimization algorithms such as REPS and approximate likelihood-free inference algorithms like ABC in order to better understand their relative strengths and weaknesses and hope to show the potential of these techniques outside their usual scope of application. Potential research directions including application of REPS to likelihood-free inference problems are sketched and discussed. Secondly, we propose an improvement over the state-of-the-art SimOpt algorithm by optimizing the simulator w.r.t. individual transitions rather than whole trajectories.

3 Connections between REPS, VI and ABC 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 and evaluate the quality of the samples. ABC keeps a collection of samples with discrepancy d under a hard threshold, resembling true posterior samples for sufficiently low thresholds. REPS weighs the samples by their respective discrepancy. To illustrate that REPS can generate real posterior samples, we focus on the ideal situation where $\log(p \mid \theta)$ is computable and therefore the ELBO can be evaluated directly. With $R(\theta) = \log p(x \mid \theta)$ and Equation (2) we obtain

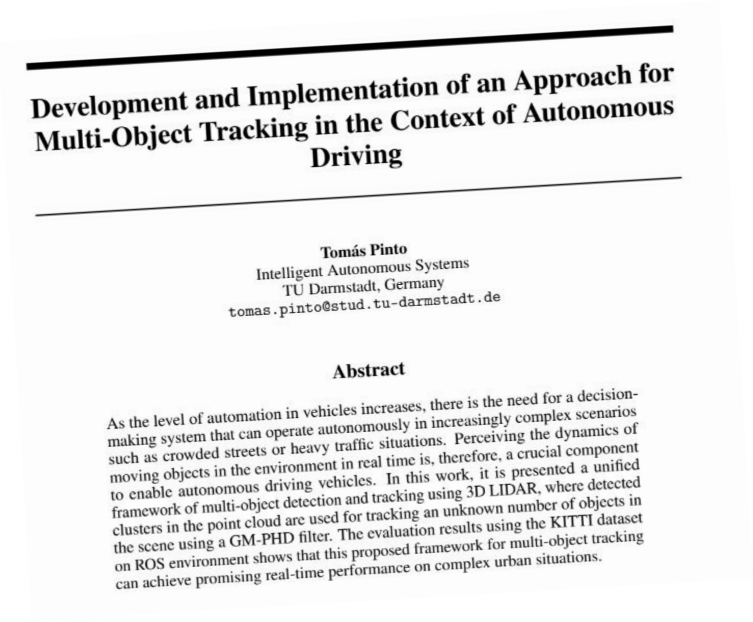
$$\underbrace{p_{n+1}(\theta \mid x)}_{\text{posterior}} \propto p_n(\theta) \exp\left(\frac{R(\theta)}{\eta}\right) = \underbrace{p_n(\theta)}_{\text{prior}} \underbrace{p_n(x \mid \theta)^{\frac{1}{\eta}}}_{\text{weighted likelihood}} .$$

η is determined by the KL-constraint of REPS and influences the weighting of prior and likelihood to limit the step size. In essence, $p_n(x \mid \theta)^{\frac{1}{\eta}}$ is concave in $p_n(x \mid \theta)$ for $\eta \geq 1$, and convex $\eta \leq 1$. The convexity skews the likelihoods in favor of likelier samples, while concavity equalizes the likelihoods of samples. If η was fixed to 1, we obtain true posterior samples. Note that if $\eta = 1$ sequentially processing the observations is equivalent to processing all observations at once.

Similar to ABC, we can substitute the log-likelihood $\log p(\hat{\mathbf{x}} \mid \mathbf{z})$ with some discrepancy on observations $\log p(\hat{\mathbf{x}} \mid \mathbf{z}) \propto -d(\hat{\mathbf{x}}, \mathbf{x})$, and therefore $R(\theta) = -d(\hat{\mathbf{x}}, \mathbf{x})$. For real observations $\hat{\mathbf{x}}$, under the limit of an extreme discrepancy function

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

Implement Algorithms



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

Tomás Pinto
Intelligent Autonomous Systems
TU Darmstadt, Germany
tomas.pinto@stud.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 streets or heavy traffic situations. Perceiving 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-PHD filter. The evaluation results using the KITTI dataset on ROS environment shows that this proposed framework for multi-object tracking can achieve promising real-time performance on complex urban situations.

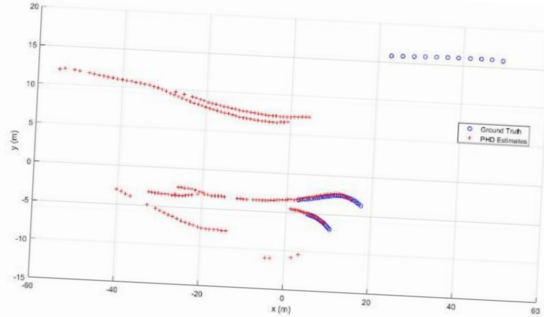


Figure 2: Tracking results on a KITTI dataset scene.

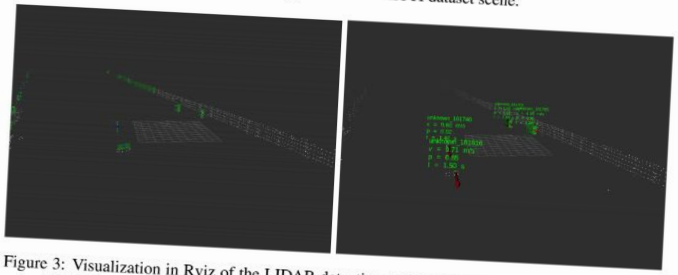


Figure 3: Visualization in Rviz of the LIDAR detection node output (Left) and the Tracker node output (Right).

Experimentation

Enhancing Intention Aware Movement Primitives

Artur Kruk, Yanhua Zhang
Department of Computer Science
Technology Universtiy of Darmstadt
Karolinenplatz 5, Darmstadt 64289
artur.kruk@stud.tu-darmstadt.de
yanhua.zhang@stud.tu-darmstadt.de

Abstract

In the future, robots will not only operate at a safe distance from humans but also have closer contact with them. To provide a comfortable human-robot cooperation, the robot is desirable to dynamically adapt its movements to not disrupt the workflow of its partner. The goal of this work is hereby to allow the robot to infer human's intention and predict its future movements based on a few early observations of human motion. Probabilistic movement primitives (ProMPs) provide a theoretical framework to model variability and inherent correlation of human motions. In this work, we are therefore learning ProMPs to predict human intentions in terms of most likely future trajectories, based on prior observed motion data. Compared to prior work that relied on linear motion models, the learned ProMPs can hereby improve the accuracy and stability of real-time motion predictions. In particular, we compare two different approaches of incorporating ProMPs for the prediction of trajectories with partial observations: First, a method which conditions the weight distribution of learned ProMPs to match the early observed data points and, second, an Expectation-Maximization based algorithm which allows for learning from trajectory with missing data and accounts for the spatial-temporal variability of the demonstrations. Experimental evaluations on recorded human motion data of 25 subjects show a better performance of the first method. A trajectory prediction close to ground truth 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 of a user study conducted by Koert et al. [9]. After that, we present and discuss the results of different methods for phase estimation and trajectory prediction evaluated on the recorded data collected in the user study.

4.1 Experimental setup

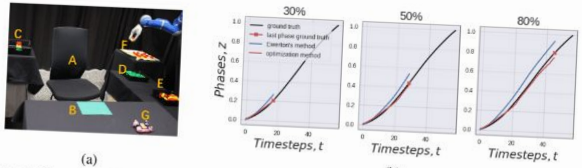


Figure 2: (a) Illustration of the experimental setup where a human and the robot perform a pick-and-place task in a shared workspace. Adapted from [9]. (b) Comparison of estimated phases of a trajectory from goal 2 to goal 1 after observing 30, 50 and 80 percent of trajectory points using Ewerton's method (blue) 3.3.1 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 their workspace. The task of a human seated at A is to assemble blocks, which need to be collected from D and E in a freely chosen order, at the work area B and deliver them to C. During the task, the human is allowed to use only one hand to collect the blocks. On this hand, the human wears a glove G with visual markers that are used to capture the hand movement via an optical-tracking system. The task of the robot is deliver blocks from F to E. During the task, the robot should adapt its movement according to the predicted movement of the human. However, since the focus of the project lies in predicting the movement of the human, we use the recorded data from the experiments where the robot is not involved in the task so that it does not influence the movements of the human. The task is completed when there are no more building blocks in D. Note that the hand of a human is moving between areas B, C, D, and E, so these areas represent motion goals.

Write a Scientific Paper


No. 12 CONCENTRIC STRUCTURES SURROUNDING LUNAR BASINS

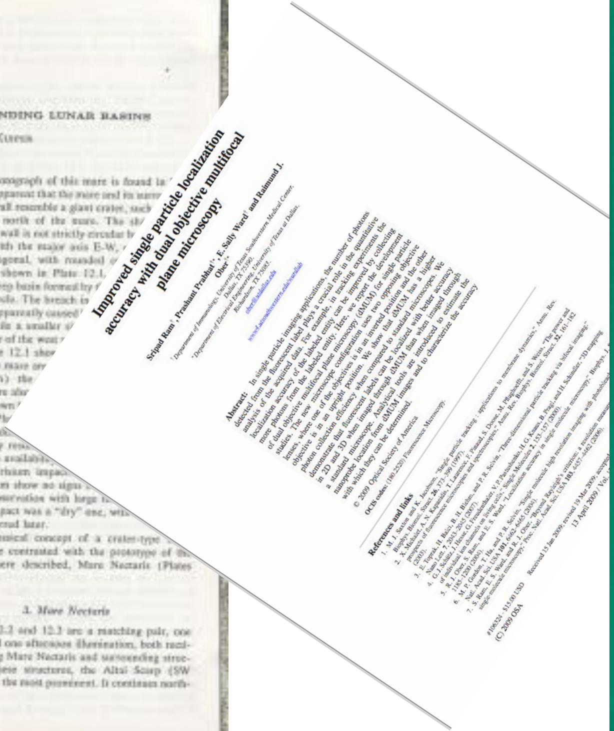
by W. K. Hartmann and G. P. Kuiper

June 20, 1962

1. Introduction

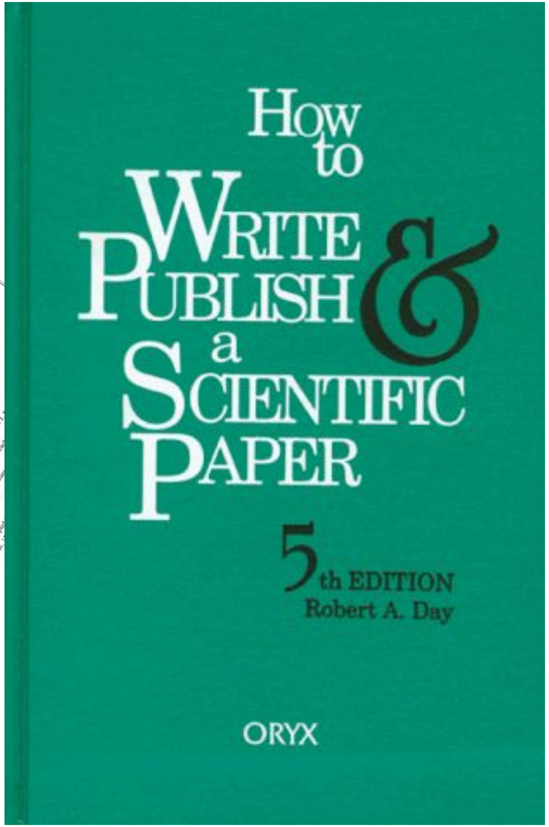
2. Well-Known Structural Features of Lunar Basins

3. Mare Nectaris


Improved single particle localization accuracy with dual objective multifocal plane microscopy

Sripad Ram, Prashant Prabhat, E. Sally Ward and Raimund J. Ober

References and links



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

TECHNISCHE
UNIVERSITÄT
DARMSTADT

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)

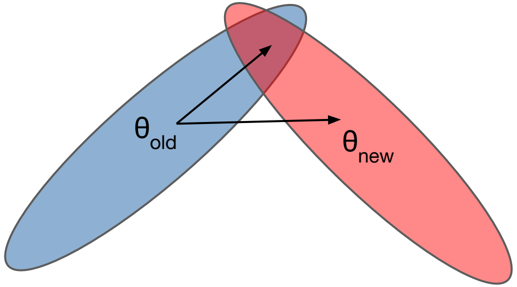
The Projects

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

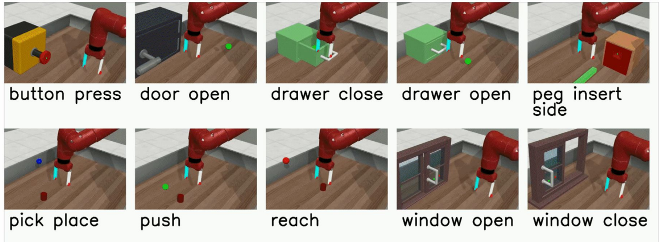


Objective:

- 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



[3]

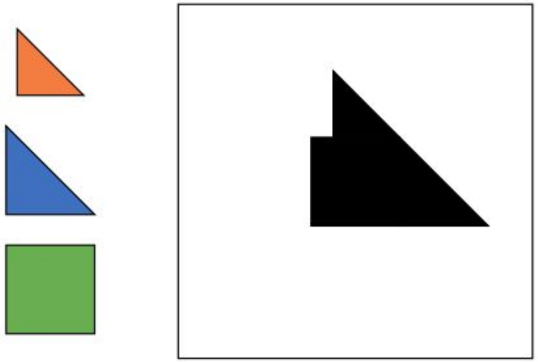
[1] Shin, Hanul, et al. "Continual learning with deep generative replay." *Advances in neural information processing systems* 30 (2017).
[2] Saha, Gobinda, Isha Garg, and Kaushik Roy. "Gradient projection memory for continual learning." *arXiv preprint arXiv:2103.09762* (2021).
[3] Yu, Tianhe, et al. "Meta-world: A benchmark and evaluation for multi-task and meta reinforcement learning." Conference on robot learning. PMLR, 2020.

Learn to play Tangram

Supervisor: Kay Hansel, Niklas Funk

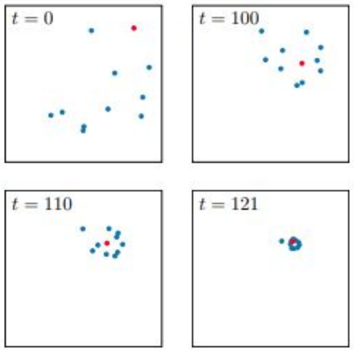
Solving the game of Tangram is challenging:

- 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



[1] Funk et al. Learn2Assemble with Structured Representations and Search for Robotic Architectural Construction. Conference of Robot Learning, 2021.
[2] Wenlong et al. One policy to control them all: Shared modular policies for agent-agnostic control. Proceedings of Machine Learning Research, 2020.

Tactile Environment Interaction

Supervisor: Niklas Funk

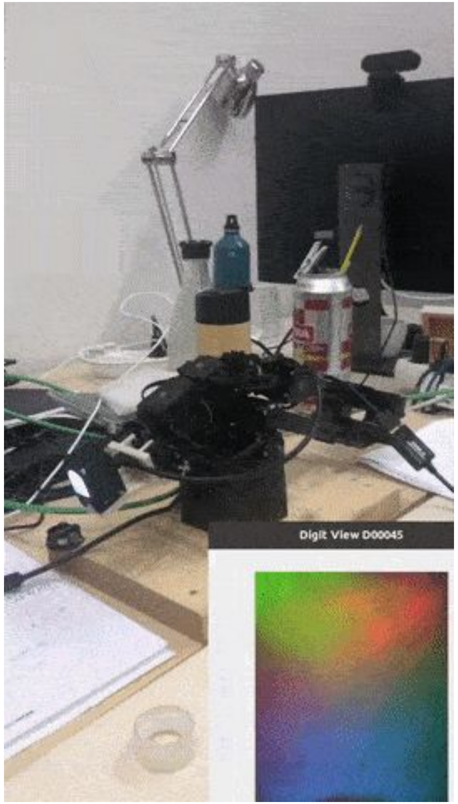
Sense of touch is essential, and has the potential to severely improve robotic systems through enabling environment interaction!

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



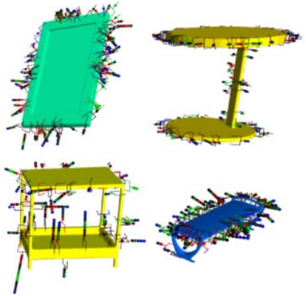
[1] Lambeta, Mike, et al. "DIGIT: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation."
[2] Belousov, Boris, et al. "Building a library of tactile skills based on fingervision."

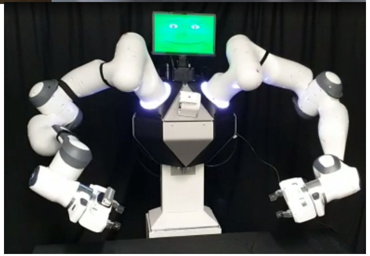
Learning Bimanual Robotic Grasping

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]



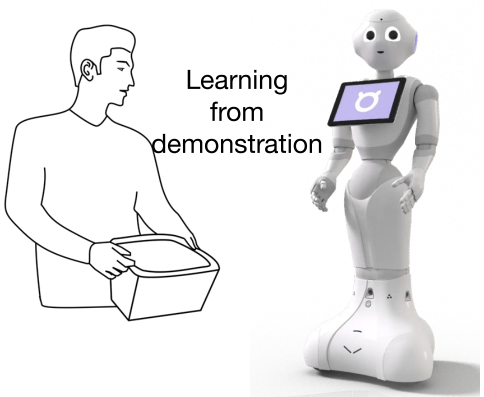


Requirements:
Good programming in python,
Prior knowledge of RL/IL (optional)

[1] Zhai, Guangyao, et al. "DA2: Dataset: Toward Dexterity-Aware Dual-Arm Grasping." IEEE Robotics and Automation Letters 7.4 (2022)

Kinematically Constrained Humanlike Bimanual Robot Motion

Supervisors: Alap Kshirsagar, Vignesh Prasad

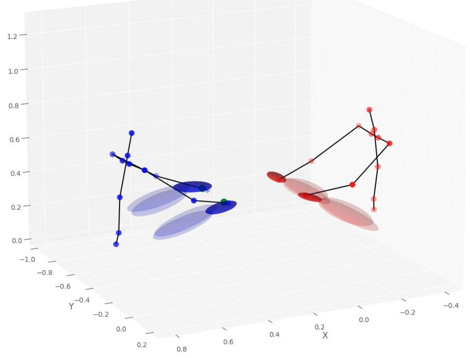


Motivation and Objective:

- 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

Gomez-Gonzalez et al. "Adaptation and robust learning of probabilistic movement primitives." IEEE Transactions on Robotics 36.2 (2020)
Campbell, J., and Amor, H. B. "Bayesian interaction primitives: A slam approach to human-robot interaction." Conference on Robot Learning (2017)

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)

[1] Davis, Justin R., et al. "The relationship between fear of falling and human postural control." Gait & posture 29.2 (2009): 275-279.

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:

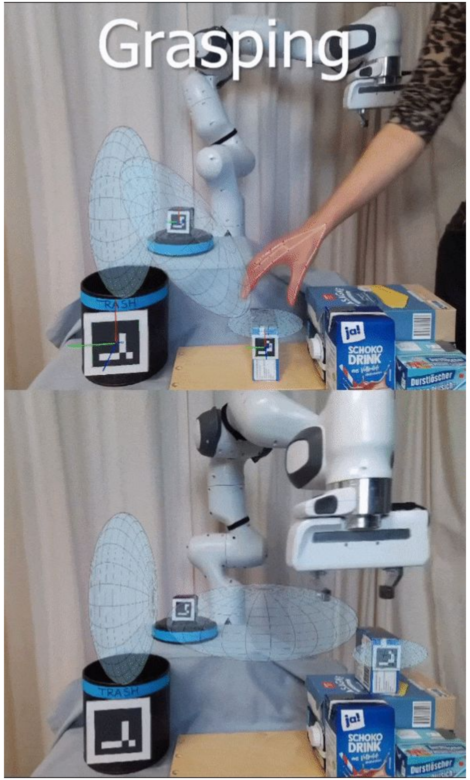
- Explore Literature on Un-/Semi- Supervised Action Segmentation and benchmark different approaches on an existing dataset.
- Develop a User Interface for interactive feedback.
- Improve the implemented algorithms iteratively with user feedback.
- Transfer learned demonstration to a robot.

Requirements

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

In cooperation with

[3]



[1] SSCAP: Self-supervised Co-occurrence Action Parsing for Unsupervised Temporal Action Segmentation, Wang et al, WACV 2022
[2] Nagano et al. "HVGH: unsupervised segmentation for high-dimensional time series using deep neural compression and statistical generative model." Frontiers in Robotics and AI (2019).
[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

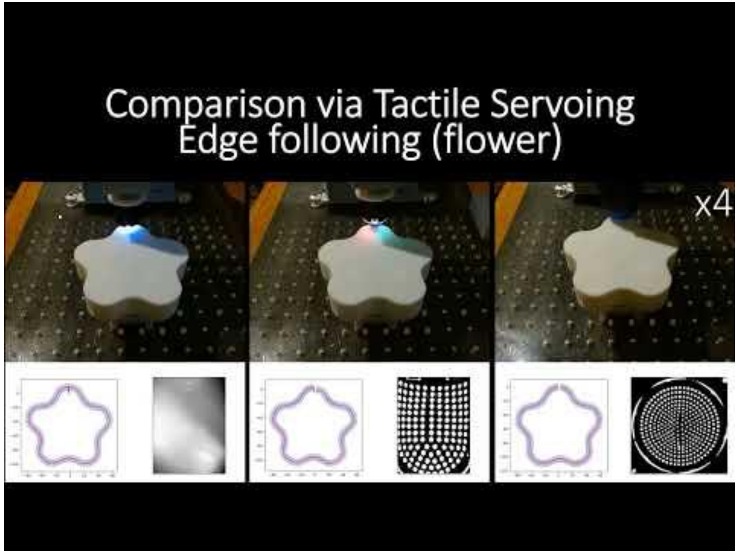
- Good programming skills in Python
- Basic Experience with reinforcement learning, nonlinear system identification and control

[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

[1] Lepora, N. F., Lin, Y., Money-Coomes, B., & Lloyd, J. (2022). Digitac: A digit-tactip hybrid tactile sensor for comparing low-cost high-resolution robot touch. IEEE Robotics and Automation Letters, 7(4), 9382-9388.
[2] Schneider, T., Belousov, B., Abdulsamad, H., & Peters, J. (2022). Active Inference for Robotic Manipulation. arXiv preprint arXiv:2206.10313.

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!