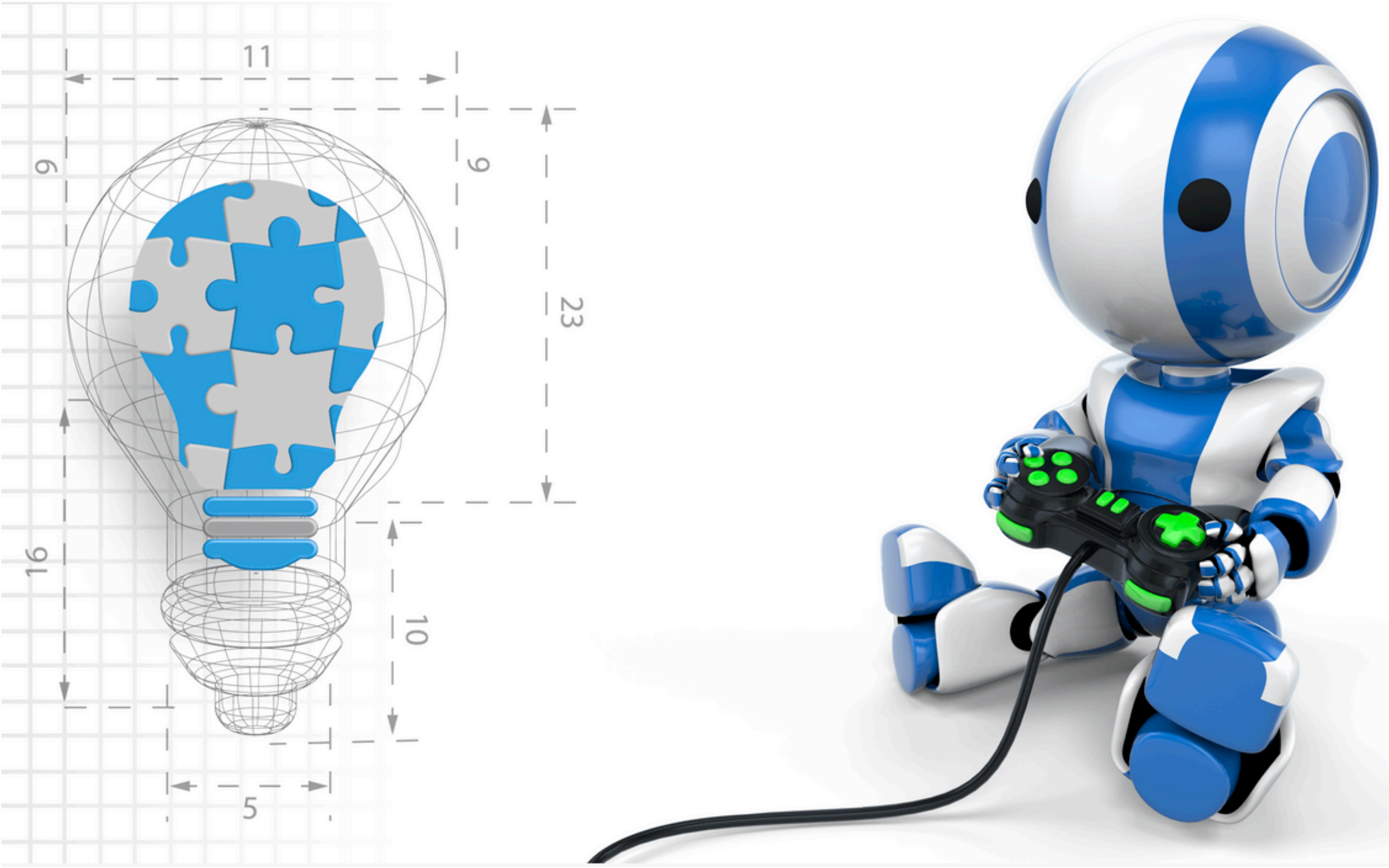

Towards Interpretable (and Robust) Reinforcement Learning Policies through Local Lipschitzness and Randomisation

Date of Meeting: 28. October 2020 via ZOOM
Begin: AM 10:00, **End:** AM 10:35
Participants: Dr.-Ing. Christopher Mutschler, Dr. Georgios Kontes, Lukas Schmidt, Hyeyoung Park
Project Partner: Dr. Georgios Kontes, Lukas Schmidt
Self-Learning Systems Group. Fraunhofer IIS, Nuremberg
Supervisor: Dr.-Ing. Christopher Mutschler, PD Dr. Fabian Scheipl
Student: Hyeyoung, Park 11365950 LMU

PROTOCOL SUMMARY

Project Overview

The outputs of neural networks applied in Reinforcement Learning (RL) are not only sensitive to small perturbations but also tricky to interpret. As long as this problem remains unsolved, RL may be underutilized in spite of the infinite possibilities of real-life application. On behalf of explainable RL, extraction of a policy and visualisation with the help of decision trees have been considered as a solution to boost interpretability of the policy. However, the size of a decision tree is often too large which hinders understanding the policy. This project will focus on visualising the effect of a local Lipschitzness constraint on the generated policy and creating a decision tree with a decent size.

Goals of this Project

1. Evaluating the performance of a policy with an added local Lipschitzness constraint to the loss function on a certain environment.
2. Visualising/comparing decision boundaries for unmodified and Lipschitz-constrained RL policies.
3. Extracting the Lipschitz-constrained policy and trying to visualise an understandable decision tree using a certain algorithm, e.g. VIPER algorithm.

Protocol Content

1. Assignment

- This Project is offered by Dr. Georgios Kontes, Lukas Schmidt from Self-Learning Systems Group. Fraunhofer IIS, Nuremberg
- The main investigator is Hyeyoung, Park (LMU) supervised by Dr.-Ing. Christopher Mutschler and Lukas Schmidt

2. Data / Status Quo

- Setting up RL Environment for Simulation:
    - a) CartPole is well-suited for visualisations and easily adaptable.
    - b) The findings can be extended in Mountain Car with varying physics properties (optional)
- Training a CartPole with various Policies.
- Creating a Git-repository named “Robust_RL”

3. Details

- Dimensionality Reduction: The velocity and position of the cart can be excluded in order to keep this experiment simple and readable. However, the angle of the pole should be included.
- Suggested option for agent framework:
- Batch-Constrained Deep Q-Learning (BCQ)
- Alternative: Spinning Up, rlpyt

4. Deliverables

- An adapted, two-variable CartPole environment that makes the visualisation of decision boundaries easy
- Visualisation and analysis of the decision boundaries for reinforcement learning agents trained with:
    a) Default training and Environment
    b) Domain Randomisation
    c) Adversarial Noise
- Visualisation and analysis of the decision boundaries for reinforcement learning agents trained with:
    a) Default environment and an added local Lipschitzness loss
    b) Domain Randomisation and an added local Lipschitzness loss
    c) Adversarial noise and an added local Lipschitzness loss
- Code with Documentation in repository
- Report that describes the methodology and the results

5. Schedule

Nov 13	Start
Dec 8	Adapted Environment + Visualisations
Jan 15	Modified Training Algorithm
Feb 22	Final Analysis and Writeup