INTELLIGENTE
AUTONOME
SYSTEME

Reinforcement Learning Lecture

Programming Challenge 1: Dynamic Programming

Boris Belousov and Hany Abdulsamad
TU Darmstadt, IAS

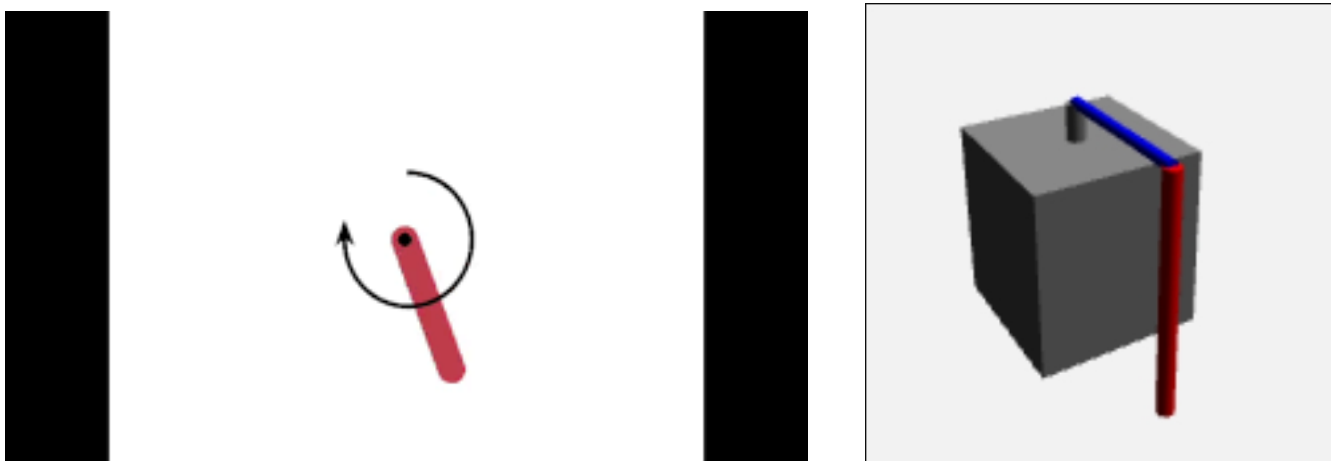
Winter Semester 2018/19

General Rules

- A template will be provided soon, but you can already start
- Submission due by 05.12.2018
- You are allowed to work in groups of 2-3 people
- One submission per group to be handed in over Moodle
- Grading criteria to be announced
- Bonus points to be announced

Challenge

- Solve the swing up task on: Pendulum-v0 and Qube-v0
- Algorithm should be able to preform a swing-up and stabilization of both systems for the infinite horizon case



Dynamics and Reward Learning

- Interaction with the system only through gym-env
  Initialize episode: $s = env.reset()$,
  Apply action: $s', r = env.step(a)$.
- Only the state and action spaces are known
- Choose a random exploration policy: $a = \mathcal{N}(\mu, \sigma)$
- Collect tuples of states, action and rewards: s', s, a, r

Dynamics and Reward Learning

- Solve regression to learn dynamics: $s_{t+1} = f(s_t, a_t)$
- Solve regression to learn reward: $r_t = g(s_t, a_t)$
- You can use Gaussian process, Neural networks, Random forest...
- Max. 10000/25000 samples for Pendulum-v0/Qube-v0 respec.
- Measure and plot your model accuracy for different number of samples

Dynamic Programming

- Choose proper/variable discretization of state-action space
- Use your dynamics and reward models to preform Dynamic Programming
- Implement both Value Iteration and Policy Iteration schemes
- Compute the cumulative reward over 100 episodes
- Compare results (plots and total reward) for different discretizations
- Plot best results of Value function and Policy for Pendulum-v0

Programming Environment

- Install miniconda https://conda.io/miniconda.html
- Create a new environment using provided YAML file https://conda.io/docs/user-guide/tasks/manage-environments.html#create-env-from-file
- Good code quality is absolutely important
- Submit python code and plots in pdfs in zip form to Moodle
- Add readme with your names and matriculation numbers