Reinforcement Learning

Coursework Introduction

Michael Herrmann, David Abel

Based on slides by Stefano V. Albrecht

11 February 2025

THE UNIVERSITY of EDINBURGH
informatics

Lecture Outline

- Coursework Outline
- Marking Details
- Submission Instructions

- Getting Started
- Contact

Coursework overview

Material Covered

1. Dynamic Programming (20 Marks)
2. Tabular Reinforcement Learning (30 Marks)
3. Deep Reinforcement Learning for Discrete Actions (30 Marks)
4. Deep Reinforcement Learning for Continuous Actions (20 Marks)

Question 1: Dynamic Programming

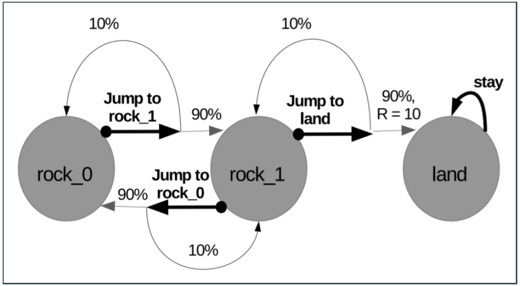

Figure 1: Example MDP for Exercise 1

- Implement functions for Value Iteration & Policy Iteration
- Create your own MDPs to debug your implementation
- Marked based on correctness of your implementation

Question 2: Tabular Reinforcement Learning

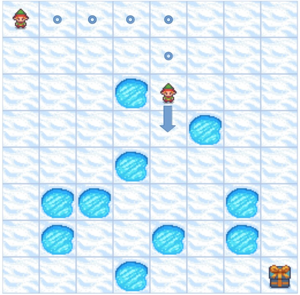

Figure 2a: FrozenLake Environment

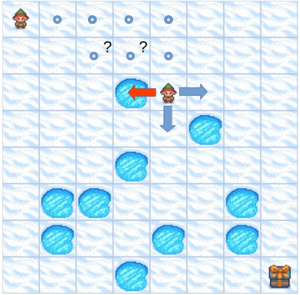

Figure 2b: Slippery FrozenLake

- Implement functions for $\mathcal{Q}$-Learning and Monte Carlo with ε-greedy policy
- Marked based on:
  - Correctness of Implementation
  - Hyperparameter tuning questions

Question 3: Deep Reinforcement Learning

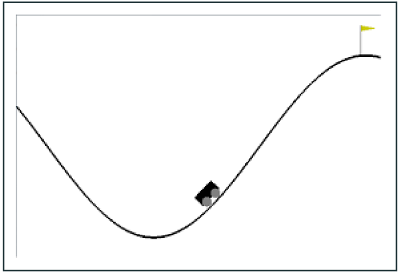

Figure 3: Rendering of the MountainCar environment

- Implement deep reinforcement learning algorithms
    - Value-based
    - Policy gradient
- Marked based on:
    - Correctness of Implementation
    - Hyperparameter tuning & scheduling
    - Loss understanding
- It is recommended to try first the CartPole problem (no marks)

Question 4: Deep Deterministic Policy Gradient

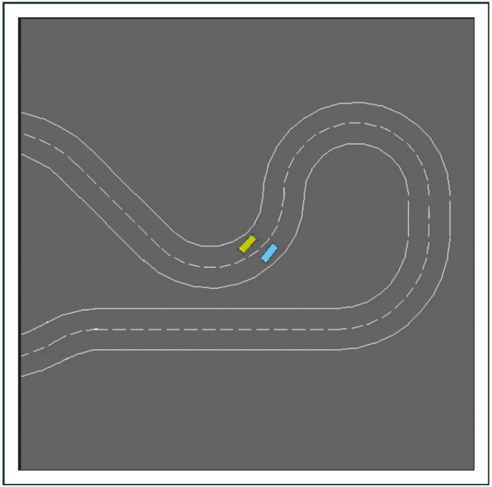

Figure 4: Rendering of the Racetrack

- Implement the Deep Deterministic Policy Gradient (DDPG) algorithm
- Stable hyperparameters for Racetrack are provided to check your algorithm's correctness
- Tune the deep networks' sizes
- Provide a trained model for DDPG in Racetrack
- Marked based on:
  - Correctness of Implementation
  - Mean returns achieved by your agent at the end of training

Bonus Question 5: Tricks and further work

- Same settings and environments as previous questions
- You are free to choose algorithmic details and tuning
- Marking based on:
  - Mean end of training returns achieved by your agent
  - A description of your strategy, approach incl. clarity and originality.
- Marking scale (bonus marks are not subject to negotiation)
  - 5 bonus marks: Reasonable effort
  - 10 bonus marks: Creative approach
  - 15 bonus marks: Impressive result
  - 20 bonus marks: Publishable result

Organisation

General remarks

- In case of any differences, the coursework description takes precedence over these slides.
- Please start reading the coursework description already now and send any comments to the course organiser.
- The coursework description may be improved over the first days after release.
- Use a virtual environment for your work on this coursework.

Marking Details

- Unit tests to evaluate correctness
- Read documentations for desired outputs of each function
- Evaluation of performance measured by average returns
- See coursework document for expected performance of each environment
- Write-up
- Fill the functions in `answer_sheet.py`

Submission Details

- Organize submission files as instructed in coursework document
- Submit as zip file through LEARN
- LEARN assignment submission guidelines:

  `https://blogs.ed.ac.uk/ilts/2019/09/27/assignment-hand-ins-for-learn-guidance-for-students/`

- Deadline: March 28th 2025, 12:00 Noon
- Submission will be possible from beginning of March.

Getting Started

Getting Started

- Read the coursework description (download from Learn)
- Read the code base documentation
- Understand the training script for each question
- Read the cited papers for each question

Codebase Setup

- Install Python3
- Set up virtual environment
- Download code base from Learn
- Install package dependencies

Contact Us

- Join a demo session per week (W6 - W9)
- Questions should be posted on Piazza
- Clarification will be posted by TA team
- Use tags in Piazza to organize posts for easy search
- Please do not discuss implementation details and solutions!

Demonstration Session

Labs will be held from W6 (after the break) to W9: Wednesdays, 4:10 pm and 5:10 pm and Thursdays, 10:00 am and 11:10 am.

- W6: Installation & setup troubleshooting
- W7: PyTorch Demonstration
- W8: Q&A with priority on Q1, Q2 and Q3
- W9: Q&A with priority on Q4 and Q5

Come with questions prepared! These are Q&A's, not lectures!

Short video introductions will be available for some of the labs.

Common Pitfalls!

- Extension deadlines are strict! Never submit late on an extension!
- Forks are public. Keep your code private and never share it.
- Start early!
- Question have different difficulty, focus on the easier ones first.
- Some subjects are still to be introduces in the lectures, be patient.
- Following suggestions of our RL tutorial lecture, keep track of performance and hyperparameters using plotting tools.
- Except for Q5, do not implement optimisations or tricks outside of the provided specifications (even if they improve performance)!

Any questions?