

Lecture plan for the second part of the course on Reinforcement Learning

Prof. Alfio Ferrara

Introduction

This lecture plan is for the second part of the course, lecturer Prof. Alfio Ferrara. The RL course will be taught in the period **January 15 - March 19, 2024**. Classes will be held in Via Celoria on **Monday** and **Tuesday** from **16:30 to 18:30**. The second part of the course this lecture plan refers to starts on **Mon, Feb 19, 2024** and end on **Tue, Mar 19, 2024**.

Class materials including slides and code example are distributed on the [GitHub repository](#). Other information and materials are also available on Prof. Cesa-Bianchi [website](#) and on the Ariel [website](#).

The course reference programming language is Python $\geq 3.9.x$.

Classes and topics

L1. Mon, Feb 12, 2024

Overview of implementation issues concerning tabular methods

- [coding] Implementation from scratch of MDP models
- [coding] Policy evaluation, Value iteration and Policy iteration algorithms
- [coding] Tutorial and introduction to [gymnasium](#)
- [coding] Example of policy evaluation, value iteration, and policy iteration using `gym`

L2. Tue, Feb 13, 2024

Implementation of estimation and control with tabular RL (using `gym`)

- [coding] Monte Carlo estimation and control
- [coding] Off-policy Monte Carlo control
- [coding] Temporal-Difference learning
- [coding] On-policy TD control (SARSA)

- [coding] Off-policy TD control (Q-Learning)

L3. Mon, Feb 26, 2024

Value Function Approximation (VFA)

- [theory] Introduction: Approximate solution methods
- [theory] Linear VFA
- [theory] Review of stochastic gradient descent (SGD)
- [theory] Model Free VFA for Policy Evaluation

L4. Tue, Feb 27, 2024

More on VFA and features engineering

- [theory] State features and features engineering
- [theory] Monte Carlo VFA for policy evaluation
- [coding] Examples of Monte Carlo first visit (using gym)
- [theory] TD learning for policy evaluation
- [coding] Examples of TD(0) methods for policy evaluation (using gym)
- [theory] Convergence of linear VFA in policy evaluation

L5. Mon, Mar 4, 2024

Control using value function approximation for state-action values

- [theory] Improve policy using VFA
- [theory] Theoretical setting with a ground-truth (similarities with supervised learning)
- [theory] Control without a ground-truth using different target values
- [theory] Monte Carlo control
- [theory] VFA On-policy control: SARSA
- [theory] VFA Off-policy control: Q-Learning

L6. Tue, Mar 5, 2024

Discussion on limits and differences between on-policy and off-policy control using VFA

- [theory] Instability and divergence of off-policy methods with VFA
- [theory] Baird's counterexample
- [coding] Implementation of Baird's counterexample using gym
- [coding] Comparison of On-policy (SARSA) and Off-policy (Q-Learning) with respect to divergence

L7. Mon, Mar 11, 2024

Introduction to Deep Q-Learning using neural networks as function approximations

- [theory] Introduction to DNN as function approximations
- [theory] Summary on Deep Neural Networks
- [theory] Summary of Convolutional Neural Networks (CNNs)
- [coding] Quick introduction to PyTorch

L8. Tue, Mar 12, 2024

Discussion and presentation of Deep Q-Learning from Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *nature*, 518(7540), 529-533.

- [theory] DQN architecture and methods to reduce the instability
- [theory] Experience replay
- [theory] Fixed Q targets
- [coding] DQN implementation using gym and PyTorch on Atari examples

L9. Mon, Mar 18, 2024

We discuss how, instead of estimating the value of actions in order to pick up the right one, we can learn a parametrized policy as a tool to select actions.

- [theory] Introduction to Policy Gradients methods
- [theory] Policy search
- [theory] The policy gradient theorem
- [theory] REINFORCE
- [theory] Actor-Critic methods

L10. Tue, Mar 19, 2024

Final project proposals presentation. Examples and discussion about modeling and implementing RL agents.

- [theory] Discussion on project proposals
- [coding] Hints and examples on coding best practices