

Autonomous and Adaptive Systems

Introduction to the Course/Administrivia

Mirco Musolesi

mircomusolesi@acm.org

Instructor: Mirco Musolesi

- ▶ Email address: mirco.musolesi@unibo.it
- ▶ Personal website: <https://www.mircomusolesi.org>
- ▶ Office hours: please send email to the address above.

Teaching Assistant: Giorgio Franceschelli

- ▶ Email address: giorgio.franceschelli@unibo.it
- ▶ Personal website: <https://giorgiofranceschelli.github.io/>
- ▶ Office hours: please send me email to the address above.

Topics of the Module

- ▶ Introduction to the design of adaptive and autonomous systems: intelligent agents and intelligent machines, automatic vs autonomous decision-making.
- ▶ Introduction to Reinforcement Learning (RL): multi-armed bandits, Montecarlo methods, tabular methods, approximation function methods, and policy-based methods.
- ▶ Applications of RL to games, classic control theory problems and robotics.
- ▶ Introduction to algorithmic game theory for multi-agent learning systems: cooperation and coordination, social dilemmas, and Multi-Agent Reinforcement Learning.
- ▶ Bio-inspired adaptive systems.
- ▶ Intelligent machines that create: Generative Learning and AI creativity.
- ▶ The "brave new world": agentic AI systems, design of agents based on foundational models, training using Reinforcement Learning from human-feedback (RLHF) and Direct Preference Optimization (DPO).
- ▶ Open problems and the future: safety, value alignment, super-intelligence, controllability, and self-awareness.
- ▶ Ethical and philosophical implications of AI and autonomous systems.

Administrivia: Organisation of the Module

- ▶ We will have a variety of activities, including practical sessions, discussion of research papers with students' presentations, etc.
- ▶ Strong focus on the state-of-the-art (and what's next): we will read papers from leading AI/ML conferences (but also classics from the ML field), try software, etc.
- ▶ I will suggest papers to read before lectures, which will discuss together.
- ▶ The participation to these activities contributes to the final mark as discussed above.

Administrivia: Assessment

- ▶ The assessment will be based on an oral exam (90%) and class participation (10%).
- ▶ 6 oral exam sessions that will be announced on the course webpage and on the institutional website.
- ▶ The credit for the class participation is only valid during the current academic year. It cannot be transferred to the following one.

Administrivia: Assessment

- ▶ 6 exams per year:
 - ▶ June and July
 - ▶ September and October
 - ▶ January and February
- ▶ The exam will be structured as follows:
 - ▶ Discussion of a compulsory mini-project with presentation (max 3 slides + slide for the title) with a working demo if possible.
 - ▶ Questions about the topics covered during the module.

Administrivia: Assessment

- ▶ The title and type of mini-project does not need to be approved in advance.
- ▶ The project can be of two types (please note that this is different with respect to the previous academic years). Either:
 - ▶ A project based on an environment given to you by us (more on this environment later in the course); or
 - ▶ A project based on an environment designed by you.
- ▶ The project has to be about solving a problem in reinforcement learning (and/or generative learning) linked to the topics of the module.

Administrivia: Assessment

- ▶ An up to 6-page short report (paper-style) about the project has to be submitted in advance, in any case **before the exam registration deadline**.
- ▶ You need to submit the code together with the report.
- ▶ Please submit the report by email to the tutor of the course (giorgio.franceschelli@unibo.it).
- ▶ You should not submit zip files but a link to a repository (cloud service, GitHub, etc).

Administrivia: Project

- ▶ Indications about potential projects will be made during the course.
- ▶ Many material online: **reusing/repackaging code is not considered a valid project.**
- ▶ I will provide clear indications about what constitute an original contribution. This will become clear after covering the relevant topics in the module.

Administrivia: Plagiarism

- ▶ The report/code will be checked for plagiarism.
- ▶ You should not use any code that you find online.
 - ▶ In case you use any code (even a couple of lines), this should be stated.
 - ▶ The fact you re-used code will be taken into consideration in the final assessment of the project.
 - ▶ The core part of the project in any case must be original, you cannot just submit a modified version of a project found in GitHub for example.
- ▶ I will follow the procedures of the University in case of plagiarism cases.

Administrivia: Format of the Report

- ▶ The report has to be written in English.
- ▶ The report must be submitted using the NeurIPS LaTeX style that can be found at this address:

<https://media.neurips.cc/Conferences/NeurIPS2024/Styles.zip>