

INFO8003-1 Reinforcement Learning

Course Organisation

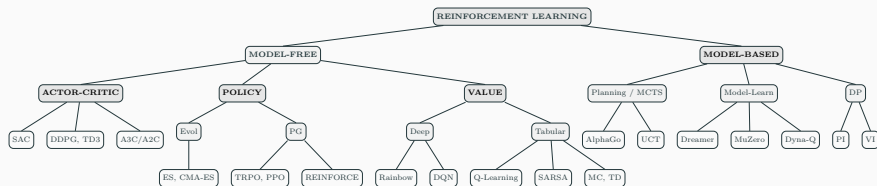
Contributors:

- Arthur Louette (arthur.louette@uliege.be) R134 B28
- Raphaël Fonteneau (Raphael.Fonteneau@uliege.be) 1/107 B28

February 3, 2026

- Course website.
- Both theoretical and practical sessions.
- Assistants:
 - Arthur Louette (arthur.louette@uliege.be) R134 B28;
 - Raphaël Fonteneau (Raphael.Fonteneau@uliege.be) 1/107 B28.
- Course on Tuesday afternoon.
- Permanence after the course at the assistants' office.

Course content



Legend: DP=Dynamic Programming, VI=Value Iteration, PI=Policy Iteration, PG=Policy Gradient, (S)AC=(Soft) Actor-Critic, ES=Evolution Strategies, Evol=Evolutionary, MC=Monte Carlo, TD=Temporal Difference, SARSA=State,Action,Reward,State,Action,DQN=Deep Q Network,TRPO=Trust-Region Policy Optimization,PPO=Proximal Policy Optimization, MCTS=Monte Carlo Tree Search

- Given by Arthur Louette & Julien Hansen.
- Goal:
 - Illustrating the concepts seen in the theoretical course;
 - Help for the project and overall understanding.
- Q/A after each course.
- After session homework:
 - A jupyter notebook to complete and submit.
- Correction of the previous notebook.

60% Exam:

- All the courses, practical sessions and questions about the project;
- Advice: have a look at the evaluations of previous years.

40% Projects and Notebooks:

- Notebooks to submit after the practical sessions;
- One project in groups of 2 students.