

Overview

Overview of this course

This two-week summer school introduces students to the theoretical foundations and algorithmic principles of reinforcement learning (RL). The course consists of **10 lectures**, each 2 hours long, and is targeted at advanced undergraduates or early graduate students.

Tentative Lecture Schedule

Lecture 1: What is Reinforcement Learning?

- Agent–environment interaction
- MDPs: states, actions, transitions, rewards
- Episodic vs. continuing tasks
- Policies and return definitions

Lecture 2: Value Functions and Optimality

- Bellman expectation equations
- Optimality and Bellman optimality
- Greedy policies

Lecture 3: Dynamic Programming

- Value iteration, policy iteration
- Generalized policy iteration
- Convergence

Lecture 4: Multi-Armed Bandits

- Regret minimization and exploration
- Epsilon-greedy, UCB, gradient bandits

Lecture 5: Monte Carlo Methods

- First-visit and every-visit prediction
- Monte Carlo control

Lecture 6: Temporal Difference Learning

- TD(0), SARSA, Q-learning
- Bootstrapping and online learning

Lecture 7: Function Approximation

- Curse of dimensionality
- Linear approximation and semi-gradient TD

Lecture 8: Policy Gradient Methods

- REINFORCE and the policy gradient theorem
- Variance reduction and baselines

Lecture 9: Actor–Critic & Deep RL Overview

- A3C-style learning and instability in Deep RL
- Replay buffers, target networks

Lecture 10: Frontiers and Applications

- Transfer, offline, and safe RL
- AlphaGo, DQN, RLHF