The Hong Kong University of Science and Technology

UG Course Syllabus (Fall 2025-26)

[Course Title] Reinforcement Learning

[Course Code] COMP4901Z

[No. of Credits] 3

[Any pre-/co-requisites]

Math: You should have familiar background in Linear Algebra (e.g., matrix inversion) and Probability (e.g., expectation, sampling).
Machine Learning: Basic machine learning knowledge (e.g., gradient backpropagation) and deep learning knowledge (e.g., neural network) as needed.
Programming: Python (necessary for assignment), PyTorch (better)

Name: [Instructor(s) Name] CHEN, Long

Email: [Your Email Address] longchen@ust.hk

Course Description

[Briefly describe the course content, key topics or themes, objectives, methods of instruction, e.g., lectures, discussions, projects].

Reinforcement learning (RL) is a computational learning approach where an agent tries to maximize the total amount of reward it receives while interacting with a complex and uncertain environment. It not only shows strong performance in lots of games (such as Go), but also becomes an essential technique in many today's real-world applications (such as LLM training, and embodied AI). This course aims to teach the fundamentals and the advanced topics of RL. The course content includes the introduction of basic RL elemnets (including MDP, dynamic programming, policy iteration), value-based approaches (DQN), policy-based approaches (policy gradient, actor critic), model-based RL, and RL techniques in today's computer vision or AI applications. To better enhance the understanding, we will also contain some Python/Pytorch implementations.

Syllabus

- Markov Decision Processes
- Dynamic Programming
- Monte Carlo and Temporal Difference Learning
- Q-Learning
- DQN and advanced techniques
- Policy Gradient
- Actor Critic
- Advanced Policy Gradient
- Continuous Controls
- Imitation Learning
- Model-based RL

Intended Learning Outcomes (ILOs)

By the end of this course, students should be able to:

1. Being familiar with different foundational RL concepts, and differentiate RL from other machine learning algorithms.
2. Formulate real-world problems in RL framework (i.e., Markov Decision Processes (MDPs)).
3. Understand and implement value-based methods (e.g., Q-learning, Deep Q-Networks (DQN)) and policy-based methods (e.g., Policy Gradients, Proximal Policy Optimization (PPO)) to solve discrete/continuous control problems
4. Understand advanced value-based and policy-based methods, and apply these techniques into designing more advanced algorithms.
5. Evaluate RL algorithms from different metrics.
6. Know the strengths and weaknesses of state-of-the-art RL algorithms, and their applications in real-world scenarios.

Assessment and Grading

This course will be assessed using criterion-referencing and grades will not be assigned using a curve. Detailed rubrics for each assignment are provided below, outlining the criteria used for evaluation.

Assessments:

[List specific assessed tasks, exams, quizzes, their weightage, and due dates; perhaps, add a summary table as below, to precede the details for each assessment.]

Assessment Task	Contribution to Overall Course grade (%)	Due date
In-class Quiz	20%	29/11/2025*
Assignment	20%	29/11/2025*
Midterm examination	20%	29/11/2025*
Final examination	40%	29/11/2025*

* Assessment marks for individual assessed tasks will be released within two weeks of the due date.

Mapping of Course ILOs to Assessment Tasks

[add to/delete table as appropriate]

Assessed Task	Mapped ILOs	Explanation
In-class Quiz	ILO1, ILO2, ILO3. ILO4, ILO5, ILO6	This task will cover all these aspects by designing different quizs.
Assignment	ILO1, ILO2, ILO3. ILO4	The coding assignment will help to students to implement and understand these new algorithms.
Midterm examination	ILO1, ILO2, ILO3. ILO4, ILO5, ILO6	The exam will cover all these aspects.
Final examination	ILO1, ILO2, ILO3. ILO4, ILO5, ILO6	The exam will cover all these aspects.

Grading Rubrics

[Detailed rubrics for each assignment will be provided. These rubrics clearly outline the criteria used for evaluation. Students can refer to these rubrics to understand how their work will be assessed.]

Final Grade Descriptors:

[As appropriate to the course and aligned with university standards]

Grades	Short Description	Elaboration on subject grading description
A	Excellent Performance	Demonstrates a comprehensive grasp of subject matter, expertise in problem-solving, and significant creativity in thinking. Exhibits a high capacity for scholarship and collaboration, going beyond core requirements to achieve learning goals.
B	Good Performance	Shows good knowledge and understanding of the main subject matter, competence in problem-solving, and the ability to analyze and evaluate issues. Displays high motivation to learn and the ability to work effectively with others.
C	Satisfactory Performance	Possesses adequate knowledge of core subject matter, competence in dealing with familiar problems, and some capacity for analysis and critical thinking. Shows persistence and effort to achieve broadly defined learning goals.
D	Marginal Pass	Has threshold knowledge of core subject matter, potential to achieve key professional skills, and the ability to make basic judgments. Benefits from the course and has the potential to develop in the discipline.
F	Fail	Demonstrates insufficient understanding of the subject matter and lacks the necessary problem-solving skills. Shows limited ability to think critically or analytically and exhibits minimal effort towards achieving learning goals. Does not meet the threshold requirements for professional practice or development in the discipline.

Course AI Policy

[State the course policy on the use of generative artificial intelligence tools to complete assessment tasks.]

The students cannot use AI tools in all these assignment tasks.

Communication and Feedback

Assessment marks for individual assessed tasks will be communicated via Canvas within two weeks of submission. Feedback on assignments will include [specific details, e.g., strengths, areas for improvement]. Students who have further questions about the feedback including marks should consult the instructor within five working days after the feedback is received.

Resubmission Policy

[If applicable, explain the policy for resubmitting work or reassessment opportunities, including conditions and deadlines.]

There is no resubmission opportunity unless with very strong justifications.

Required Texts and Materials

Reference book: Excel and Excel VBA Programming for Beginners – 3rd Edition for Office 2013

Richard S. Sutton. Reinforcement Learning: An Introduction. Second Edition.
Kevin P. Murphy. Reinforcement Learning: An Overview.

Academic Integrity

Students are expected to adhere to the university's academic integrity policy. Students are expected to uphold HKUST's Academic Honor Code and to maintain the highest standards of academic integrity. The University has zero tolerance of academic misconduct. Please refer to Academic Integrity | HKUST – Academic Registry for the University's definition of plagiarism and ways to avoid cheating and plagiarism.

Additional Resources

Online course content to be published in HKUST canvas when the semester begins.