City University of Hong Kong
Course Syllabus

offered by School of Data Science
with effect from Semester A 2019/20

Part I Course Overview

Course Title:	**Foundation of Reinforcement Learning**
Course Code:	**SDSC4001**
Course Duration:	**One Semester**
Credit Units:	**3**
Level:	**B4**
Proposed Area: *(for GE courses only)*	☐ Arts and Humanities ☐ Study of Societies, Social and Business Organisations ☐ Science and Technology
Medium of Instruction:	**English**
Medium of Assessment:	**English**
Prerequisites: *(Course Code and Title)*	**SDSC2002 Convex Optimization *or* MA3515 Introduction to Optimization** ***and*** **MA2506 Probability and Statistics**
Precursors: *(Course Code and Title)*	**Nil**
Equivalent Courses: *(Course Code and Title)*	**Nil**
Exclusive Courses: *(Course Code and Title)*	**Nil**

Part II Course Details

1. Abstract

(A 150-word description about the course)

This advanced elective course introduces the essential elements and mathematical foundations of the modern reinforcement learning: the optimal control theory, including dynamic programming and numerical techniques. It emphasizes both the fundamental theories in control theory and the numerical methods in context of reinforcement learning algorithms. It also equips students with computing algorithms and techniques for applications to some practical problems.

2. Course Intended Learning Outcomes (CILOs)

(CILOs state what the student is expected to be able to do at the end of the course according to a given standard of performance.)

No.	CILOs#	Weighting* (if applicable)	Discovery-enriched curriculum related learning outcomes (please tick where appropriate)		
			A1	*A2*	*A3*
1.	Explain clearly basic concepts in reinforcement learning, optimal control and dynamic programming.	10%	√		
2.	Understand the concept, theory and properties of Markov Decision Process and the fundamentals of optimal control and dynamic programming	25%	√	√	
3.	Explain and apply the methods and theories of Markov decision process and optimal control and dynamic programming to the reinforcement learning context.	25%	√	√	
4.	Explain algorithms of reinforcement learning in the context of data science and machine learning.	20%		√	√
5.	Apply reinforcement learning to formulating and solving real-life problems	20%		√	√
		100%			

** If weighting is assigned to CILOs, they should add up to 100%.*

 # Please specify the alignment of CILOs to the Gateway Education Programme Intended Learning outcomes (PILOs) in Section A of Annex.

A1: *Attitude*
Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.

A2: *Ability*
Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to self-life problems.

A3: *Accomplishments*
Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

3. Teaching and Learning Activities (TLAs)

(TLAs designed to facilitate students' achievement of the CILOs.)

TLA	Brief Description	CILO No.					Hours/week (if applicable)
		1	2	3	4	5	
Lecture	Learning through **teaching** is primarily based on lectures.	√	√	√	√	√	39 hours in total
Take-home assignments	Learning through **take-home assignments** helps students understand techniques of basic methods in as well as their applications in solving optimal control problems.	√	√	√	√		after-class
Online applications	Learning through **online examples for applications** helps students create and formulate mathematical models and apply to a range of practical problems in economics/science.				√		after-class

4. Assessment Tasks/Activities (ATs)

(ATs are designed to assess how well the students achieve the CILOs.)

Assessment Tasks/Activities	CILO No.					Weighting*	Remarks
	1	2	3	4	5		
Continuous Assessment: 40%							
Test	√	√		√		15-30%	Questions are designed for the part of the course to see how well the students have learned basic concepts of methods in optimal control and recognized their applications in solving problems.
Hand-in assignments	√	√	√	√		0-15%	These are skills based assessment to enable students to demonstrate the understanding of theories and the ability of applying dynamic programming methods in a diversity of problems.
Formative take-home assignments	√	√	√	√		0%	The assignments provide students chances to demonstrate their achievements on techniques of optimal control and dynamic programming learned in this course.
Examination: 60% (duration: 3 hours)	√	√	√	√	√	60%	Examination questions are designed to see how far students have achieved their intended learning outcomes. Questions will primarily be skills and understanding based to assess the student's versatility in basic methods of mathematical programming.
**The weightings should add up to 100%.*						100%	

For a student to pass the course, at least 30% of the maximum mark for the examination must be obtained.

5. Assessment Rubrics

(Grading of student achievements is based on student performance in assessment tasks/activities with the following rubrics.)

Assessment Task	Criterion	Excellent (A+, A, A-)	Good (B+, B, B-)	Fair (C+, C, C-)	Marginal (D)	Failure (F)
1. Test	Ability to understand the basic concepts of methods and recognize their applications in solving application problems	High	Significant	Moderate	Basic	Not even reaching marginal levels
2. Hand-in assignments	Ability to apply the techniques in a diversity of problems	High	Significant	Moderate	Basic	Not even reaching marginal levels
3. Examination	Ability to solve problems of reinforcement learning and Markov decision process with fundamental methods.	High	Significant	Moderate	Basic	Not even reaching marginal levels
4. Formative take-home assignments	Ability to demonstrate students' achievements on techniques learned in this course	High	Significant	Moderate	Basic	Not even reaching marginal levels

Part III Other Information (more details can be provided separately in the teaching plan)

1. Keyword Syllabus

(An indication of the key topics of the course.)

- Reinforcement learning: sequential decision model; state space; action space; policy, planning, rewards and transition probability; value function; model-free and model-based methods
- Markov Decision Process: discrete dynamic programming; discrete stochastic dynamics programming; optimal value function, open loop control, feedback control
- The principle of Optimality: Bellman equation
- Policy evaluation, greedy policy, policy improvement, value iteration and policy iteration
- Temporal-Difference Learning, Monte Carlo method, Q-learning

2. Reading List

2.1. Compulsory Readings

(Compulsory readings can include books, book chapters, or journal/magazine articles. There are also collections of e-books, e-journals available from the CityU Library.)

1.	Lecture note
2.	*Reinforcement Learning: An Introduction* by Richard S. Sutton and Andrew G. Barto, The MIT Press 2017.

2.2. Additional Readings

(Additional references for students to learn to expand their knowledge about the subject.)

1.	*Introduction to Stochastic dynamic programming* By Sheldon Ross, 1983.
2.	"Optimal Control Theory: An Introduction" (Dover Books on Electrical Engineering), by Donald E. Kirk. 2004.
3.	Deterministic and Stochastic Optimal Control by W. Fleming and R. Rishel. Springer. 1975.