Annexe A: New/Revised Course Content in OBTL+ Format

Course Overview

The sections shown on this interface are based on the templates UG OBTL+ or PG OBTL+
If you are revising/duplicating an existing course and do not see the pre-filled contents you expect in the subsequent sections e.g. Course Aims, Intended Learning Outcomes etc. please refer to Data Transformation Status for more information.

Expected Implementation in Academic Year	AY2024-2025
Semester/Trimester/Others (specify approx. Start/End date)	Semester 1
Course Author * Faculty proposing/revising the course	Jeremie Houssineau
Course Author Email	jeremie.houssineau@ntu.edu.sg
Course Title	Reinforcement Learning
Course Code	MH4521
Academic Units	4
Contact Hours	57
Research Experience Components	Not Applicable

Course Requisites (if applicable)

Pre-requisites	MH2500 and MH3512
Co-requisites	
Pre-requisite to	
Mutually exclusive to	
Replacement course to	
Remarks (if any)	

Course Aims

This course will introduce the framework of reinforcement learning including some theoretical aspects and practical algorithms. The course will start with special cases such as multi-armed bandits before moving on to Markov decision processes and the corresponding planning and online reinforcement learning problems.

If you want to build a solid understanding of the principles and fundamental results backing reinforcement learning, as well as develop some intuition about the methodology and practical challenges of this approach, then this course is meant for you.

This course will equip you with a practical understanding of reinforcement learning, hence allowing you to apply this type of methods in machine-learning-related jobs. The theoretical insights gained in this course will also help you adapt to future developments in the field.

Course's Intended Learning Outcomes (ILOs)

Upon the successful completion of this course, you (student) would be able to:

ILO 1	Formulate suitable problem as Markov decision processes and apply relevant algorithms
ILO 2	Use Standard Proof Techniques And Adapt Them To New Situations
ILO 3	Evaluate effectiveness of a given learning method, both in theory and in practice
ILO 4	Implement standard reinforcement learning algorithms in Python

Course Content

- Multi-armed bandits: theory and practice
- Markov decision processes (MDPs)
- Fundamental Theorem of MDPs
- Temporal-Difference methods
- Policy gradient methods
- Model-based reinforcement learning

Reading and References (if applicable)

- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press, ISBN: 9780262039246.
- Lattimore, T., & Szepesvári, C. (2020). *Bandit algorithms*. Cambridge University Press, ISBN-10: 1108486827.
- Moerland, T. M., Broekens, J., Plaat, A., & Jonker, C. M. (2023). Model-based reinforcement learning: A survey. Foundations and Trends® in Machine Learning, 16(1), 1-118.
- Lapan, M. (2018). *Deep Reinforcement Learning Hands-On: Apply modern RL methods, with deep Q-networks, value iteration, policy gradients, TRPO, AlphaGo Zero and more*. Packt Publishing Ltd. ISBN-10: 1838826998

NOTE: The above readings comprise the foundational readings for the course and more up-to-date relevant readings will be provided when they are available.

Planned Schedule

Week or Session	Topics or Themes	ILO	Readings	Delivery Mode	Activities
1	Background material		Lattimore & Szepesvári 2020, Sections 2 and 3	In-person	Lectures
2	Multi-armed bandits: practice	1, 3	Sutton & Barto 2018, Section 2	In-person	Lectures and a tutorial
3	Multi-armed bandits: theory	2, 3	Lattimore & Szepesvári 2020, Section 4	In-person	Lectures and a tutorial
4	Markov decision processes: definition and practice	1	Sutton & Barto 2018, Section 3	In-person	Lectures, a tutorial and a computer lab
5	Markov decision processes: theoy	2	Lattimore & Szepesvári 2020, Section 38	In-person	Lectures and a tutorial
6	Temporal-Difference methods: on-policy vs. off policy	1, 4	Sutton & Barto 2018, Section 6	In-person	Lectures, a tutorial and a computer lab
7	Midterm test			In-person	
8	Policy gradient methods	1, 4	Sutton & Barto 2018, Section 13	In-person	Lectures and a tutorial
9	Policy gradient theorem	2, 3	Sutton & Barto 2018, Section 13	In-person	Lectures and a tutorial
10	Model-based reinforcement learning: principles	1	Moerland et al. 2023	In-person	Lectures, a tutorial and a computer lab
11	Model-based reinforcement learning: methods	1, 3	Moerland et al. 2023	In-person	Lectures and a tutorial

Week or Session	Topics or Themes	ILO	Readings	Delivery Mode	Activities
12	Advanced topics: function approximation and deep reinforcement learning	1, 4	Lapan 2018	In-person	Lectures and a tutorial
13	Advanced topics: Actor-critic methods	1, 4	Sutton & Barto 2018, Section 13	In-person	Lectures and a tutorial

Learning and Teaching Approach

Approach	How does this approach support you in achieving the learning outcomes?
Technology-enhanced learning: use of quizzes and polls in lecture to support learning	This helps students taking a more active role during lectures (anonymous, not assessed)
Optional individual project	Having a particular problem in mind helps establishing a context for algorithms and principles seen in lectures (not assessed)

Assessment Structure

Assessment Components (includes both continuous and summative assessment)

No.	Component	ILO	Related PLO or Accreditation	Weightage	Team/Individual	Rubrics	Level of Understanding
1	Continuous Assessment (CA): Assignment(4-question assignment to do at home)	1, 2, 3, 4	Not Applicable	10	Individual	Analytic	Extended Abstract
2	Continuous Assessment (CA): Test/Quiz(100-minute midterm test, short answer & computational questions)	1, 2, 3	Not Applicable	25	Individual	Analytic	Relational
3	Summative Assessment (EXAM): Final exam(2-hour exam, short answer & computational questions)	1, 2, 3	Not Applicable	60	Individual	Analytic	Relational
4	Continuous Assessment (CA): Class Participation(Participation in tutorials)	1, 2, 3	Not Applicable	5	Individual	Holistic	Extended Abstract

Description of Assessment Components (if applicable)

Additional details about assessment number:
1. Questions in assignments will vary in form, ranging from the derivation of additional results to the implementation and assessment of reinforcement learning algorithms in Python. Assignments will be focused on recently covered results, offering an opportunity to revisit the material studied during lectures and tutorials.
4. Students will be assigned tutorial questions to prepare in advance so that they can present their proposed solution to the rest of the tutorial group. This will be an occasion for students to learn from their peers.

Formative Feedback

You will receive formative feedback through written responses to your papers and verbal feedback through in-class discussion. You will receive summative group feedback on the exam following the conclusion of the module.

NTU Graduate Attributes/Competency Mapping

This course intends to develop the following graduate attributes and competencies (maximum 5 most relevant)

Attributes/Competency	Level
Problem Solving	Advanced
Sense Making	Intermediate

Course Policy

Policy (Academic Integrity)

Good academic work depends on honesty and ethical behaviour. The quality of your work as a student relies on adhering to the principles of academic integrity and to the NTU Honour Code, a set of values shared by the whole university community. Truth, Trust and Justice are at the core of NTU's shared values. As a student, it is important that you recognize your responsibilities in understanding and applying the principles of academic integrity in all the work you do at NTU. Not knowing what is involved in maintaining academic integrity does not excuse academic dishonesty. You need to actively equip yourself with strategies to avoid all forms of academic dishonesty, including plagiarism, academic fraud, collusion and cheating. If you are uncertain of the definitions of any of these terms, you should go to the academic integrity website for more information. On the use of technological tools (such as Generative AI tools), different courses / assignments have different intended learning outcomes. Students should refer to the specific assignment instructions on their use and requirements and/or consult your instructors on how you can use these tools to help your learning. Consult your instructor(s) if you need any clarification about the requirements of academic integrity in the course.

Policy (General)

You are expected to complete all assigned readings, activities, assignments, attend all classes punctually and complete all scheduled assignments by due dates. You are expected to take responsibility to follow up with assignments and course related announcements. You are expected to participate in all project critiques, class discussions and activities.

Policy (Absenteeism)

In-class activities make up a portion of your course grade. Absence from class without a valid reason will affect your participation grade. Valid reasons include falling sick supported by a medical certificate and participation in NTU's approved activities supported by an excuse letter from the relevant bodies. There will be no make-up opportunities for in-class activities.

Policy (Others, if applicable)

Diversity and inclusion policy

Integrating a diverse set of experiences is important for a more comprehensive understanding of science.

It is our goal to create an inclusive and collaborative learning environment that supports a diversity of perspectives and learning experiences, and that honours your identities; including ethnicity, gender, socioeconomic status, sexual orientation, religion or ability.

To help accomplish this:

If you are neuroatypical or neurodiverse, have dyslexia or ADHD (for example), or have a social anxiety disorder or social phobia;
If you feel like your performance in the class is being impacted by your experiences outside of class;
If something was said in class (by anyone, including the instructor) that made you feel uncomfortable;
Please speak to your teaching team, our school pastoral officer, or a peer or senior (either in-person or via email) about how we can help facilitate your learning experience.

As a participant in course discussions, you should also strive to honour the diversity of your classmates. You can do this by: using preferred pronouns and names; being respectful of others opinions and actively making sure all voices are being heard; and refraining from the use of derogatory or demeaning speech or actions.

All members of the class are expected to adhere to the NTU anti-harassment policy. if you witness something that goes against this or have any other concerns, please speak to your instructors or a faculty member.

Point-based marking (not rubric-based)