

Online and Reinforcement Learning (OReL) – Spring 2025

Course Introduction

Mohammad Sadegh Talebi

m.shahi@di.ku.dk

Department of Computer Science



OReL 2025

- A course dealing with basic theory and algorithms of sequential decision making problems tackled via online learning and reinforcement learning.
- This is the 4th round, and we have 98 students (14% increase relative to 'OReL 2024').
- OReL is a restricted elective course in:
 - MSc Programme in Computer Science
 - MSc Programme in Statistics
 - MSc Programme in Mathematics-Economics



Course Team



Course Team: Instructors

**Christian Igel**

Professor at ML Section, DIKU
Head of AI Center

igel@di.ku.dk

(4 lectures)

**Yevgeny Seldin**

Professor and head of
ML Section, DIKU

seldin@di.ku.dk

(5 lectures)

**Sadegh Talebi**

Assistant Professor at
ML Section, DIKU

m.shahi@di.ku.dk

(7 lectures & course
management)



Course Team: Teaching Assistants



Hippolyte Bourel (x2)
PhD Candidate
at DIKU



Mathias Overby
MSc Student
at MATH



Tran Duy Vu
MSc Student
at DIKU



Course Plan



Weekly Plan

	Mon	Tue	Wed	Thu	Fri
9:15 – 10		TA (x3)		Lecture 1	
10 – 11		TA (x3)		Lecture 1	
11 – 12		TA (x3) + OTA		Lecture 1 + Q&A	
12 – 13:15					
13:15 – 14				Lecture 2	TA (x1)
14 – 15				Lecture 2	TA (x1)
15 – 16				Lecture 2 + Q&A	TA (x1)
					
20:59 (sharp)			HA DEADLINE		

↓
HOME
ASSIGNMENTS

- ‘TA’ denotes *physical* exercise sessions.
- ‘OTA’ denotes the *online* TA session (held over Zoom).
- You can attend any and as many sessions you like.
- Note: A slightly different plan in Week 6.
- For locations and potential updates, check “Where & When” on Absalon.



Tentative Lecture Plan

Week 1	9:15-12:00			13:15-16:00
Tuesday		Course Introduction	Overview of Online and Reinforcement Learning	-
Thursday	Theory of MDPs 1			Theory of MDPs 2

Week	Thursday, 9:15-12:00	Thursday, 13:15-16:00
2	Direct Policy Search	Policy Evaluation and Off-policy Evaluation
3	Off-policy Optimization and Q-Learning	Stochastic Bandits
4	Policy Gradient Methods	Stochastic Bandits (cont'd) + Adversarial Full Information
5	Adversarial Full Information (cont'd) + Adversarial Bandits	Deep RL I
6	Deep RL II	Theory of MDPs 3
7	Offline Evaluation of Bandits. Advanced Learning Settings (Contextual, Linear, ...)	PAC Exploration in MDPs
8	Regret Minimization in MDPs	Course Summary

Yevgeny Christian Sadegh



Subject to minor changes; check Absalon for the latest info.

Tentative Lecture Plan

Week 1	9:15-12:00		13:15-16:00
Tuesday	Course Introduction	Overview of Online and Reinforcement Learning	-
Thursday	Theory of MDPs 1 (RL)		Theory of MDPs 2 (RL)

Week	Thursday, 9:15-12:00	Thursday, 13:15-16:00
2	Direct Policy Search (RL)	Policy Evaluation and Off-policy Evaluation (RL)
3	Off-policy Optimization and Q-Learning (RL)	Stochastic Bandits (Online Learning)
4	Policy Gradient Methods (RL)	Stochastic Bandits (cont'd) + Adversarial Full Information (Online Learning)
5	Adversarial Full Information (cont'd) + Adversarial Bandits (Online Learning)	Deep RL I (RL)
6	Deep RL II (RL)	Theory of MDPs 3 (RL)
7	Offline Evaluation of Bandits. Advanced Learning Settings (Contextual, Linear, ...) (Online Learning)	PAC Exploration in MDPs (RL + Online Learning)
8	Regret Minimization in MDPs (RL + Online Learning)	Course Summary

Yevgeny Christian Sadegh



Subject to minor changes; check Absalon for the latest info.

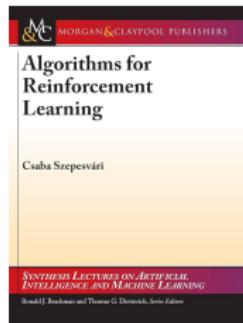
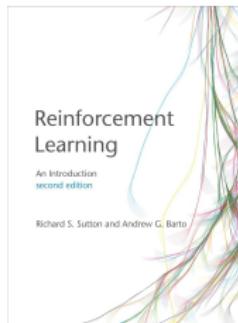
Course Material

- **Main material:** Lecture notes, slides, some papers, ... and blackboard

☰	Lecture Notes:
☰	🔗 [YS] Machine Learning -- The Science of Selection under Uncertainty
☰	🔗 [ST1] Notes on Theory of Discounted MDPs
☰	🔗 [ST2] Notes on Theory of Episodic MDPs
☰	🔗 [ST3] Notes on Theory of Average-Reward MDPs

- **Supplementary material (optional):**

- R. Sutton & A. Barto. [Reinforcement Learning: An Introduction](#).
- Cs. Szepesvári. [Algorithms for Reinforcement Learning](#).



Home Assignments



Home Assignments

- RELEASED ON WEDNESDAY
- Weekly home assignments
- Due dates: **Wednesdays at 20:59.**
 - **No resubmissions** (except for re-exam, if necessary)
 - Individual submission:
 - You are welcome to study together, but should write the solution **individually**.
 - Use of **LLMs** (e.g., ChatGPT, DeepSeek) is deemed **cheating**.
 - Final grade = the average grade of $n-1$ 'best' HAs
 - $n=8$ planned HAs
 - Submitting all HAs could help when your average is close to the boundaries.
- YOU CAN SUBMIT AS MANY TIMES AS YOU WANT BUT ONLY BEFORE THE DEADLINE (NOT SURE) (?)
- BUT THINGS SUCH AS "GRAMMALLY" CAN BE USED TO FIX THE GRAMMAR



Late Submissions



Late submissions will not be graded
... irrespective of the reason

Late submissions:

- Will not be graded ... irrespective of the reason
- But do submit, because we will consider (i) the number of submissions and (ii) their content when giving the final grade (i.e., it may help in borderline cases)

Sickness:

- Do NOT notify us about sickness/late submissions/etc.
- The final grade formula covers one potential emergency you may have during the course (e.g., sickness).
- If more than one emergency during the course, please, inform us at the end of the course and we will look into it.



Home Assignments: Feedback and Ref Solutions

- We do not hand out written reference solutions (only occasionally).
- But some reference solutions will be provided during **the Q&A hour** at the end of each lecture.
 - This hour is also meant for other questions related to course material.
 - You are also welcome to ask on Absalon (the **Discussions** forum).
- TAs provide feedback on your submissions:
 - *Short comments* regarding what was wrong when they take points.
 - They are not expected to provide written feedback on how to fix your mistakes.
 - You can ask them such things orally at a TA session (or a Q&A hour).
- In case of complaints:
 - Any questions regarding the feedback? Ask TAs first.
 - **We will not consider complaints below 10 points per assignment.** Contact the relevant TA first.



No lessons on the last week (should be only for TA and exams)
Passing grade for the exam should be 41-42% roughly