

# CMPUT 652: Reinforcement Learning with Robots

## Schedule

Class / week	Date	Topic	Material	Work
1 / 1	Sept 3	<a href="#">Introduction to AI &amp; RL with robots</a>		
2 / 1	Sept 5	<a href="#">The road to AI</a>		
3 / 2	Sept 10	<a href="#">Gradient descent &amp; averaging</a>	<a href="#">RL Ch 2</a> , <a href="#">ESL 2.4, 2.6</a>	
4 / 2	Sept 12	<a href="#">Function approximation with vector inputs</a>	RL Ch 2, <a href="#">Mahmood thesis Section 9.1</a>	*#(ESL 2.4, 2.6)
5 / 3	Sept 17	<a href="#">Bandits &amp; gradient descent</a>	RL Ch 2	*
6 / 3	Sept 19	<a href="#">Bandits, supervised learning and unification of objectives</a>	RL Ch 2, <a href="#">Schulman et al 2017</a>	*#(RL Ch 2)
7 / 4	Sept 24	<a href="#">Markov decision processes &amp; value functions</a>	RL Ch 3, <a href="#">Bertsekas (2012) Vol 2 Ch 1</a>	*
8 / 4	Sept 26	<a href="#">Dynamic programming</a>	RL Ch 4	*#(RL Ch 3, 4, due Sat), <b>Assignm. 1 release Sun</b>
9 / 5	Oct 1	<a href="#">Dynamic programming &amp; function approximation</a>	RL Ch 6, 9, 11.4, <a href="#">Olver notes 9</a>	*
10 / 5	Oct 3	<a href="#">Temporal difference methods &amp; off-policy learning</a>	RL Ch 9, 10, 11	*
11 / 6	Oct 8	<a href="#">Reflecting on off-policy learning and learning in real time</a>		*#(RL Ch 9.1-9.4, 9.8 11.4 & 13)
12 / 6	Oct 10	<a href="#">The big picture &amp; off-policy learning</a>	RL Ch 13, <a href="#">Mahmood thesis Ch 5</a> , <a href="#">Mahmood et al 2014 Neurips</a>	*, <b>Assignment 2 release,</b> <b>Assignment 1 due Oct 13</b>
13 / 7	Oct 15	Derivation of REINFORCE	RL Ch 13	*#( <a href="#">PPO</a> , <a href="#">DRL</a> )

14 / 7	Oct 17	Derivation of REINFORCE with gamma correction	RL Ch 13	*
15 / 8	Oct 22	Derivation of PPO from REINFORCE in nine steps	<a href="#">PPO</a>	*#( <a href="#">RL-setup</a> , <a href="#">10 steps</a> , <a href="#">unstated</a> )
16 / 8	Oct 24	PID controller, Move the motor & Read the packets ( <b>Lab</b> )		<b>Project abstract due</b>
17 / 9	Oct 29	<a href="#">Real-time learning tasks</a>	<a href="#">RL-setup</a>	<b>Assignm. 2 due</b>
18 / 9	Oct 31	Analyze observations ( <b>Lab</b> )	<a href="#">RL-setup</a>	<b>Project approval</b>
19 / 10	Nov 5	<a href="#">Discussion on proposal presentation and assignment 3</a>	<a href="#">RL-setup</a> , <a href="#">SenseAct</a>	#( <a href="#">SenseAct</a> , <a href="#">survey</a> )
20 / 10	Nov 7	Proposal presentations	<a href="#">SenseAct</a>	<b>Assignm. 3 release,</b> <b>Proposal presentation</b>
- / 11	Nov 12	Reading week	<a href="#">Baoding</a> , <a href="#">ROBEL</a>	
- / 11	Nov 14	Reading week	<a href="#">Async</a> , <a href="#">Walk</a> , <a href="#">TEXPLORE</a>	
21 / 12	Nov 19	Calculating lambda returns and advantage estimates	<a href="#">LfD</a>	# <a href="#">LfD</a>
22 / 12	Nov 21	Mind, mechanism & learning from demonstrations		
23 / 13	Nov 26	Online representation search	<a href="#">Mahmood thesis Ch 11 &amp; 12</a>	<b>Assignment 3 due</b>
24 / 13	Nov 28	Guest lecture: Patrick Pilarski		
25 / 14	Dec 3	General-purpose minds for robots		
26 / 14	Dec 5	Project presentations		<b>Project presentation</b>
	Dec 16			<b>Project report</b>

\* Class note due

# Reading summary due