

# ADVANCED REINFORCEMENT LEARNING

JUNE 28–30, 2021 | Instructors: Pulkit Agrawal, Cathy Wu | [professional.mit.edu/arl](http://professional.mit.edu/arl)

*Note: All times are US Eastern Daylight Time. Schedule is subject to change.*

DAY 1: MONDAY, JUNE 28 (9:30am–7:30pm = 10 hours)	
9:30–11:00 AM	<b>Session 1: Intro, review of basic RL, overview, Why RL?</b> <ul style="list-style-type: none"> <li>• Statistical / ML perspective</li> <li>• Optimal control / Operational Research perspective</li> <li>• Economics perspective (discounting)</li> <li>• Online v/s Offline</li> <li>• On Policy v/s Off Policy</li> </ul>
11:00–11:30 AM	<b>Break</b>
11:30 AM–12:30 PM	<b>Session 2: Theoretical Results in RL</b>
12:30–1:30 PM	<b>Lunch</b>
1:30–2:00 PM	<b>Coffee Chat</b>
2:00–3:30 PM	<b>Session 3: Offline RL Theory &amp; Applications</b> <ul style="list-style-type: none"> <li>• Introduction to Offline RL</li> <li>• State-of-the art in Offline RL</li> <li>• Applications of Offline RL</li> </ul>
3:30–4:00 PM	<b>Break</b>
4:00–6:00 PM	<b>Hands-on implementation</b> <ul style="list-style-type: none"> <li>• Setup an environment into a format amenable to RL algorithms</li> <li>• Design reward functions to gain practical experience on reward hacking               <ul style="list-style-type: none"> <li>- Discussion on approaches to circumvent reward hacking</li> </ul> </li> <li>• Compare offline and online RL</li> </ul>
6:00–7:30 PM	<b>Reception</b>

# ADVANCED REINFORCEMENT LEARNING

JUNE 28–30, 2021 | Instructors: Pulkit Agrawal, Cathy Wu | [professional.mit.edu/arl](http://professional.mit.edu/arl)

*Note: All times are US Eastern Daylight Time. Schedule is subject to change.*

DAY 2: TUESDAY, JUNE 29 (9:30am–6:30pm = 9 hours)	
9:30–11:00 AM	<b>Session 1: Exploration v/s Exploitation</b> <ul style="list-style-type: none"> <li>• Exploration using Learning Progress, Prediction Error and State Visitation</li> <li>• Curriculum Learning</li> </ul>
11:00–11:30 AM	<b>Break</b>
11:30 AM–12:30 PM	<b>Session 2: Goal Based and Hierarchical RL</b>
12:30–1:00 PM	<b>Coffee Chat</b>
1:00–2:00 PM	<b>Lunch</b>
2:00–4:00 PM	<b>Session 3 (Interactive): Using Models and Demonstrations to Improve Sample Efficiency</b> <ul style="list-style-type: none"> <li>• Theory of improving sample efficiency</li> <li>• Implementing Models and Demonstrations</li> </ul>
4:00–4:30 PM	<b>Break</b>
4:30–5:15 PM	<b>Session 4: Deep Learning Architectures and RL</b> <ul style="list-style-type: none"> <li>• Memory Based RL</li> <li>• Transformers and Self-Attention</li> <li>• Episodic Control</li> </ul>
5:15–5:30 PM	<b>Break</b>
5:30–6:30 PM	<b>Session 5: State-of-the Art RL Algorithms</b> <ul style="list-style-type: none"> <li>• TD3</li> <li>• Soft Actor Critic (SAC)</li> <li>• RAINBOW</li> <li>• PPO</li> </ul>

# ADVANCED REINFORCEMENT LEARNING

JUNE 28–30, 2021 | Instructors: Pulkit Agrawal, Cathy Wu | [professional.mit.edu/arl](http://professional.mit.edu/arl)

*Note: All times are US Eastern Daylight Time. Schedule is subject to change.*

DAY 3: WEDNESDAY, JUNE 30 (9:30am–7:30pm = 10 hours)	
9:30–11:30 AM	<b>Session 1: Problem Clinic and Casting your problem into RL</b>
11:30 AM–12:30 PM	<b>Office Hours: 10 mins sign up slots</b>
12:30–1:30 PM	<b>Lunch</b>
1:30–3:00 PM	<b>Session 2: Applications to RL</b> <ul style="list-style-type: none"> <li>• Operations Research</li> <li>• Navigation</li> <li>• Manipulation</li> <li>• Urban Planning</li> </ul>
3:00–3:30 PM	<b>Break</b>
3:30–5:30 PM	<b>Session 3: Miscellaneous Topics</b> <ul style="list-style-type: none"> <li>• MCTS and its application to Alpha Go</li> <li>• Safety in RL</li> <li>• Are policy gradients true gradients?</li> <li>• Connection between RL and Evolutionary Algorithms</li> <li>• RL from the neuroscience view</li> </ul>
5:30–6:00 PM	<b>Break</b>
6:00–7:30 PM	<b>AMA session</b>