University of Tehran
School of Electrical and Computer Engineering

<table>
<tr><td>Course:</td><td>Reinforcement Learning</td></tr>
<tr><td>Course type:</td><td>
<table>
<tr><td></td><td colspan="6">EE*</td><td colspan="4">CE*</td><td rowspan="2">Credit:

3</td></tr>
<tr><td></td><td>Com</td><td>E</td><td>P</td><td>B</td><td>Con</td><td>D</td><td>SW</td><td>HW</td><td>IT</td><td>MI</td></tr>
<tr><td>Required</td><td>☐</td><td>☐</td><td>☐</td><td>☐</td><td>☐</td><td>☐</td><td>☐</td><td>☐</td><td>☐</td><td>☒</td></tr>
<tr><td>Elective</td><td>☒</td><td>☐</td><td>☐</td><td>☒</td><td>☒</td><td>☐</td><td>☒</td><td>☒</td><td>☒</td><td>☐</td></tr>
</table>
</td></tr>
<tr><td>Level:</td><td>Undergraduate ☐
Graduate ☒</td></tr>
<tr><td>Co-requisite(s):</td><td>None.</td></tr>
<tr><td>Prerequisite(s):</td><td>Statistics and Probability (8101092)</td></tr>
<tr><td>Prerequisite by topic:</td><td>Programming mainly in Python</td></tr>
<tr><td>Textbook(s):</td><td>Main textbooks

[1] Sutton, Richard S., and Andrew G. Barto. Reinforcement learning: An introduction. The MIT Press, 2018
[2] Lattimore, Tor, and Szepesvári, Csaba, Bandit Algorithm. Cambridge Univ Press, 2020.
[3] Szepesvári, Csaba. "Algorithms for reinforcement learning." Morgan and Claypool, 2009.

For further reading

[3] Glimcher, P.W. and Fehr, E. and Camerer, C. and Poldrack, R.A , Neuroeconomics: Decision Making and the Brain, Elsevier, 2014.
[4] Olson M., Hergenhahn B.R. , An Introduction to the Theories of Learning (9th edition), Prentice-Hall, 2012.</td></tr>
<tr><td>Coordinator:</td><td>Majid Nili Ahmadabadi, Professor, School of ECE.</td></tr>
<tr><td>Goals:</td><td>Learning how the living agents interactively learn is of prime importance for creation of artificial systems; that adapt to the natural creatures and learn to improve the services they give to their users. Also, developing interactive learning methods to produce general AI is essential to create learning artificial systems for diversity of tasks and environments. The goal of this course is to provide the mathematical and computational basics needed to learn these two topics in a unified view, both at the individual and social levels. The focus is on reinforcement learning (RL) methods in discrete and continuous environments as well as on modeling learning behavior in individual and social contexts.</td></tr>
<tr><td>Outcome:</td><td>Upon successful completion of the course, students will be able to</td></tr>
</table>

	1. Do statistical analysis of learning and decision making behavior. 2. Formulate optimization and learning tasks into one reinforcement learning problem. 3. Get familiar with modeling real-world scenarios as a Multi-Armed Bandit (MAB) problem. 4. Understand basic properties of different MAB setting, and the behavior of corresponding learning algorithms. 5. Be able to analyze basic MAB problems via mathematical tools, and possibly extend them to other setups. 6. Develop and employ a diversity of reinforcement learning methods in: 6.1. discrete MDP environments 6.2. continuous MDP environments (mainly Deep RL) 6.3. discrete POMDP environments 7. Model human and animal learning and decision making behavior. 8. Analyze new reinforcement learning methods by mathematical and computational methods. 9. Improve the existing reinforcement learning methods.
Topics:	1- Introduction to human decision making methods and biases in individual and social context. 2- Reinforcement learning algorithms in single-step tasks. 3- Statistical analysis and modeling of learning behavior. 4- Stochastic Bandits with Finitely Many Arms. 5- Adversarial Bandits with Finitely Many Arms. 6- Markov models for discrete and continuous environments. 7- Dynamic programming methods in discrete MDP. 8- Reinforcement learning methods in discrete MDP. 9- Reinforcement learning methods in continuous MDP. 10- Deep reinforcement learning methods. 11- Introduction to social interactive learning.
Computer usage:	Implementing the projects using Python Software.
Assignments:	Six homework, including algorithm implementation and behavior modeling.
Projects:	Arbitrary topic related to the course.
Grading:	Assignments[1]: 30% Midterm exams: 20% Final exam: 30% Project[2] 20%
Further readings:	Some state of the art papers on learning in SMDP, POMDP, and continuous environments especially Deep RL.
Prepared by:	Majid Nili Ahmadabadi, Professor, School of ECE.
Date:	September, 10, 2020, revised September 20, 2021

*EE: Electrical Engineering		CE: Computer Engineering	
Com	Communications	SW	Software
E	Electronics	HW	Hardware
P	Power	IT	Information Technology
B	Bioelectronics	MI	Machine Intelligence and Robotics
Con	Control		

[1] Submitting assignments is mandatory.

[2] The project should target an application with economical or social value for the community.