Title of the course: Reinforcement Learning

1. Credit requirement:(L-T-P: 3-0-0, Credit: 3)

2. Please select the committee for Approval: PGPEC

3. Name of the Dept: CSE

4. Please Specify the Level of the Subject: UG Final Year/PG level

5. Whether the subject will be offered as compulsory or elective: Elective

6. Prerequisite(s) for the subject, if any

(Please give the subject numbers and names): 1. CS60010: Deep Learning

Course Objective

Sequential decision-making is one of the major topics in machine learning. From experience, the task is to decide the sequence of actions to perform in an uncertain environment in order to achieve some goals that may not necessarily seem beneficial in near future but are optimal for getting better long term reward. Reinforcement learning (RL) is a paradigm that proposes a formal framework to this problem. With its root in behavioral psychology, it aims to model the trial-and-error learning process where an artificial agent may learn by interacting with its environment. Over the past few years, RL has become increasingly popular due to its success in addressing long-standing and interesting sequential decision-making problems in variety of domains like machine learning, computer vision, natural language processing and robotics. Several of these achievements are due to combination of RL with deep learning techniques that allows dealing with problems in very high dimensional state-space. With such renewed interest in RL, a separate full-fledged course on Reinforcement Learning can be a fantastic addition. Its a great time for the younger force of our country to delve into the intricacies of reinforcement learning. It will also augment the existing Machine Learning and Deep Learning courses.

The aim of the course will be to familiarize the students with the basic concepts as well as with the state-of-the-art research literature in deep reinforcement learning. The first part will introduce the basic building blocks of classical reinforcement learning i.e., Markov Decision Processes (MDP), the Bellman state and optimality equations, value and policy iteration strategies etc. This part will be rigorous with Mathematical formulations, derivations and proofs of uniqueness and convergence guarantees of solutions. This section will be concluded by going from MDPs to model free Reinforcement learning with discussions on Monte-carlo tree search, Temporal Difference learning, Q-learning etc. The second part of the course will combine classical theories of RL and the advanced concepts of Deep Learning towards a treatise of Deep Reinforcement Learning. The topics that will be covered in this part of the course are algorithms and strategies developed in the last 3-4 years of worldwide research in this area (Deep Q learning, Double Q Learning etc.). This will pave the path to the myriads of policy optimization based RL algorithms. This includes stochastic policy gradients (REINFORCE), deterministic policy gradients (DDPG, Actor-critic etc.), natural policy gradients, TRPO, PPO etc. and finally evolutionary strategies. Policy optimization will form a major part of the course to make the students conversant with the latest developments in policy optimization strategies of RL. The course will end with a brief introduction to other advanced topics under the bigger umbrella of RL. This includes partially observable Markov decision processes (POMDP), learning without explicit reward function and multi-agent reinforcement learning.

At completion, the course will

1. enable the students to understand and apply basic RL methodologies (policy/value iteration, Monte-Carlo, TD, Q-Learning, policy gradient methodologies etc.) for sequential decision making or prediction problems.
2. enable the students to understand and critically appraise related literature on both classical works and on current frontiers of RL.
3. enable the students to implement existing RL algorithms using industry level RL tools in python.

Contents/Syllabus:

Topic No	Contents	Hours
1	Introduction/Course Logistics	1
2	Linear Algebra and Probability Review	1
3	MP/MRP/MDP: Definition and examples, State/Action Value Functions	2
	Bellman expectation/optimality equation, Derivation/Existence/Uniqueness	3
4	Model based planning (prediction and control) of MDPs, value and policy iterations and their convergence	3
5	Model free prediction (Monte-Carlo and TD)	2
6	Model free control (On-policy, Off-policy, SARSA, TD, Q-Learning)	3
7	Function approximation	2
8	Deep NN Architectures for RL	2
9	Deep Q Learning, Double Q learning, replay memory	2
10	Policy Gradients – Stochastic Policy Gradients (REINFORCE)	1
	Deterministic Policy Gradients (DDPG)	2
	Actor-Critic, A3C	2
	Natural Policy Gradients - Introduction	1
	TRPO, PPO	2
	SVG	2
11	Exploration and Exploitation, Bandits	2
12	Brief introduction to advanced topics in RL - POMDP	2
	Learning without explicit reward function (LFD,	2

	Learning from Feedback)	
	Multi Agent RL	2
	Guest Lecture: (Possibly on autonomus driving and RL)	1

Resources:

Books:

1. Sutton & Barto, Reinforcement Learning: An Introduction, MIT Press, 2nd Edition.

General References:

1. Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming
2. François-Lavet, Henderson, Islam, Bellemare and Pineau, An Introduction to Deep Reinforcement Learning, now
3. Szepesvari, Algorithms for Reinforcement Learning, Morgan & Claypool

In addition a number of recent papers will be discussed reference to which will be provided as the course progresses.

Evaluation plan:

Homework (10%), Mid-term (10%), End-term (40%), Course project (40%)

Related Subjects offered by the Institute:

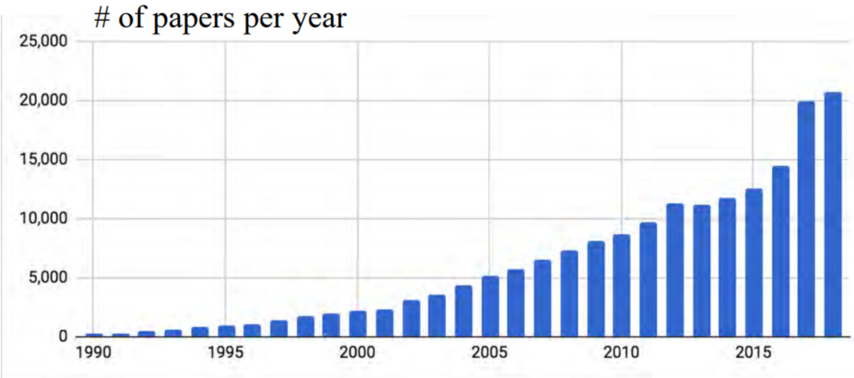
CS60045 Artificial Intelligence

a)Approximate percentage of overlap: 5%

b)Reasons for offering the new subject in spite of the overlap

The scope of the above subject is large and thus RL forms only a small part of it where RL ideas are briefly stated as an introductory topic among many other topics. The dedicated course on RL is designed after realizing the potential that being able to build and analyze applications on RL would be useful for the graduate and upper undergraduate students of our institute. Side by side the course would be of tremendous help for researcher scholars working in the area of machine learning, computer vision, natural Language Processing and Speech Signal Processing. The proposed course will go deep into the Mathematics of RL for both classical as well as new developments in the domain. Policy gradients is one of the major focus area of the course which is not covered in the above mentioned course. In addition, the discussions on Deep Reinforcement Learning would beautifully supplement the Deep Learning courses offered by the CSE Department and the Center for AI. Finally, the advanced topics (e.g., POMDP, Multi Agent RL) will help and motivate the younger force of our country to delve into the intricacies of Reinforcement Learning research which is growing at an unprecedented rate than ever (ref Fig. 1 below).

25+ years of RL papers



P. Henderson, R. Islam, P. Bachman, J. Pineau, D. Precup, D. Meger.
Deep Reinforcement Learning that Matters. AAAI 2017 *(+updates)*.

Fig 1: Number of papers using RL