SYLLABUS

Course Type		Elective		
Course Code		EE807		
Course Name	Korean	전기공학특강 <강화학습이론>		
Course Name	English	Mathematical Foundation of Reinforcement Learning		
Instructors		Song Chong **(정송)**		
Lecture: Exp.: Credit(Homework)		3:0:3(6)	Prerequisites	EE528
Mutually Recognized Course(Undergrad./Grad.) : (x)			Term	Spring
Class material	http://netsys.kaist.ac.kr/EE807/ee807.html		Classroom/ Time	E3-2 2203 Mon/Wed 14:30-16:00

TA	Name	Office Hours	Office	Tel	Email
	Jeongmin Bae	TBA	IT Center(N1)-912	042-350-5473	jmbae@netsys.kaist.ac.kr
	Sewoong Lee	TBA	IT Center(N1)-912	042-350-5473	swlee@netsys.kaist.ac.kr

Textbook	Neuro-Dynamic Programming, Dimitri P. Bertsekas and John Tsitsiklis, Athena Scientific, 1996 Dynamic Programming and Optimal Control, Vol. I, Dimitri P. Bertsekas, 4th Ed., Athena Scientific, 2017 Dynamic Programming and Optimal Control, Vol. II: Approximate Dynamic Programming, Dimitri P. Bertsekas, 4th Ed. Athena Scientific, 2012 Lecture Notes
Late Submission & Copy Penalty of Assignment	Students can use up to 24 hours of late submission with penalty 30% degradation of your original score. There is no late submission allowed after 24 hours. Also, there is no score on the copied version.
Grading Policy	Midterm (35%) Final (35%) Homework (20%) Attendance (10%) (Homework includes three programming assignments.)
Descriptions of Courses	The subject of this course is sequential decision making under uncertainty in a system whose evolution is influenced by decisions. The decision made at any given time depends on the state of the system and the objective is to select a decision making rule that optimizes a certain performance criterion. Such problems can be solved, in principle, using the classical methods of dynamic programming. In practice, however, the applicability of dynamic programming to many important problems is limited by the enormous size of the underlying state spaces. “Neuro-dynamic programming” or "Reinforcement Learning" which is the term used in the Artificial Intelligence literature, uses neural networks and other approximation architectures to overcome such bottlenecks to the applicability of dynamic programming. The methodology allows systems to learn about their behavior through simulation, and to improve their performance through iterative reinforcement. The focus of this course is to understand the mathematical foundations of this methodology in light of the convergence and degree of suboptimality of different algorithms.

•Schedule

Period	Topics	Remarks
Week 1	Markov Decision Processes and Dynamic Programming	
Week 2	Markov Decision Processes and Dynamic Programming	
Week 3	Markov Decision Processes and Dynamic Programming	
Week 4	Simulation-Based Methods	

Week 5	Simulation-Based Methods	
Week 6	Value Function Approximation	
Week 7	Value Function Approximation	
Week 8	Mid-term exam	
Week 9	Value Function Approximation	
Week 10	Value Function Approximation	
Week 11	Policy Search Methods	
Week 12	Policy Search Methods	
Week13	Policy Search Methods	
Week 14	Online Learning and Games	
Week 15	Online Learning and Games	
Week 16	Final exam	

* Laboratory topics should be given in Remark for courses with lab hours.

Name of Professor : Sign

Chairman : Sign