	UE-ML2: Sequential Decision Making	Semester 1
Contributes to	MICAS	

Coordinators:	Philippe CIBLAT, Telecom Paris Mireille SARKISS, Telecom SudParis	
Volume:	30h	**3 ects**
Hours:	Lectures: 18h, Exercises: 4.5h, Labs: 6h	
Assessment:	2 Assignments and 1 Final Exam	
Language:	English	

Objectives:
The course provides Markovian Decision Process Theory and shows how it can be applied in many settings. An introduction to Deep Reinforcement Learning (DRL) is also conducted.

Outcomes:
On completion of the course students should be able to:
- Identify a Markov Decision Process setting
- Check assumptions to exhibit optimal policy
- Code algorithms related to Markov Decision Process

Prerequisite
- Linear algebra
- Introduction to Probability and Statistics

Syllabus
- Finite Markov chain: transition matrix, recurrent states, transient states, steady-state, analysis through stochastic matrices, Graph representation : connected graph, Laplacian matrix, Hitting time
- Some examples about Sequential Decision-making process
- Markov Decision Process: the discounted case
  - Optimal offline policy, Bellman equation, value function, Q function
  - Value Iteration and Policy Iteration algorithms
- Markov Decision Process: the average case (optimal offline policy)
- Constrained Markov Decision Process: optimal offline policy
- Some suboptimal policies: Whittle's index with threshold policy
- Markov Decision Process without knowing transition probability
  - Value and value-cost Q function
  - Reinforcement learning : exploration/exploitation tradeoff
  - Epsilon-greedy, Boltzmann algorithms
  - Deep reinforcement learning: TD algorithm
- Application to Telecommunications (2 Labs)
  - Value Iteration
  - DRL

Bibliography:

- M. Puterman, "Markov Decision Processes: Discrete Stochastic Dynamic Programming", 1994.
- E. Altman, "Constrainted Markov Decision Processes", 1999.
- D. Bertsekas, "Dynamic Programming and Optimal Control", 1995.
- O. Hernandez-Lerma, "Adaptive Markov Control Processes", 1989.