Master thesis:
Learning and decision making under uncertainty

Supervision:	Urtzi AYESTA
	Research Director CNRS and Ikerbasque Professor at UPV/EHU
	`http://www.irit.fr/~Urtzi.Ayesta`
	Maaike VERLOOP
	Associate Researcher CNRS
	`http://verloop.perso.enseeiht.fr/`

Location:	IRIT-ENSEEIHT, RMESS group, Toulouse, and Donosti/St. Sebastian
Starting date:	Any time before summer 2019.
Fellowship:	Monthly allowance will be provided according to the rules set by CNRS.
PhD perspectives:	After completion of the project, the student might be offered a PhD position.

Description of research area: Lots of artificial systems (either software agent, or hardware robot) obtain their data sequentially, along time. For instance, this data may consist of information posted on social media, clicks of customers on banners, or attachments on social networks. The theory of sequential decision making (aka reinforcement learning) deals with how to learn over time the behavior of the system and to make effective decisions in order to optimize certain performance criteria. Applications of sequential decision theory are very numerous in industry, and it is at the heart of the artificial intelligence technological revolution. In combination with deep neural networks, it plays a key role in obtaining supra-human intelligent algorithms in games like Go, Atari etc.

Objectives of the internship: The main focus of the internship is to investigate sequential decision making in the presence of a non-observable environment whose state changes over time. This happens in many application areas, including recommendation systems, scheduling etc. For example, think of Facebook or Twitter where user-dependent links (such as contacts, ads, etc) must be placed. The software can collect information on the behavior of the user and learns the profile of the user. Based on this, it adapts its decisions in order to maximize revenue. However, the user's behavior might strongly depend on its mood, or other variable that is non-observable by the system. Another example can be downloading data by mobile users. Here, a base station must decide at every time epoch to which mobile user to send data. The instantaneous channel quality might be unknown to the controller of the base station, but changes over time due to fading.

The objective of the internship will be to design and analyze an efficient algorithm to learn to make decisions in such settings. A main drawback of learning algorithms is that the computational complexity grows exponentially in the dimension of the problem. To overcome this problem, the main idea will be to combine techniques from sequential learning with those from optimal stochastic control (MDP). The latter permits to determine optimal structures, that will help when designing learning algorithms.

Keywords: Sequential learning, reinforcement learning, artificial intelligence, multi-armed bandits, Markov processes

Requirements: Candidates should have a background in (applied) mathematics, operations research, computer science or electrical engineering. Experience in stochastic modeling, stochastic optimization, queueing theory or sequential learning will be appreciated. In the course of the project the candidate is expected to perform some numerical experiments (Matlab, C++, Python ...).

Relevant bibliography:

- D.P. Bertsekas, Dynamic programming and optimal control II, Athena Scientific, 2005
- Richard S. Sutton and Andrew G. Barto, Reinforcement Learning: An Introduction, 2018

Application: Please send a detailed curriculum vitae, along with a brief cover letter motivating your interest, to urtzi.ayesta@irit.fr and verloop@irit.fr