The new algorithms should be robust to environmental changes, both in terms of the observations that the system can make and the actions that the system can perform. Moreover, the algorithms should efficiently allocate the limited resources to minimal time, space, and sample complexity. In general, a successful sequential learning strategy should efficiently allocate the limited resources to exploitation (making the best decision based on our current, but possibly imperfect, knowledge) and to exploration (decisions that may appear sub-optimal but which may reduce the uncertainty and, as a result, could improve the relevance of future decisions).

This Ph.D. program is focused on sequential learning in structured and dynamic environments. The key aspect of this problem is that relatively little knowledge of the environment is available beforehand, and the learner (a virtual or physical agent) has to sequentially interact with the environment to learn its structure and then act optimally. This problem encompasses a wide range of applications depending on the definition of the structure of the environment, the sequential nature of the interaction between learner and environment, and the type of dynamics and evolution. In particular, this Ph.D. program will be driven by two application domains of scientific and societal importance: planning in games and in random environments. In both setups, the paradigm of reinforcement learning, in which the learner collects rewards to guide his/her actions, will be useful.

In many cases, the dynamic of the sequential interaction is particularly complex, since rewards keep changing over time and the set of actions (e.g., items that can be recommended in recommender systems) continuously evolve (products may be added or removed or change characteristics). Finally, both applications usually involve large spaces which ask for efficient learning algorithms with minimal time, space, and sample complexity. In general, a successful sequential learning strategy should efficiently allocate the limited resources to exploitation (making the best decision based on our current, but possibly imperfect, knowledge) and to exploration (decisions that may appear sub-optimal but which may reduce the uncertainty and, as a result, could improve the relevance of future decisions).

Many complex autonomous systems (e.g., electrical distribution networks, or smart grids) repeatedly select actions with the aim of achieving reinforcement. Reinforcement learning (RL) offers a powerful framework for acquiring adaptive behavior in this setting, associating a scalar reward with each action and learning from experience which action to select to maximize long-term reward. Although RL has produced impressive results recently (e.g., achieving human-level play in Atari games and beating the human world champion in the board game Go), most existing solutions only work under strong assumptions: the environment model is stationary, the objective is fixed, and trials end once the objective is met. The aim of this project is to advance the state of the art of fundamental research in lifelong RL by developing several novel RL algorithms that relax the above assumptions. The new algorithms should be robust to environmental changes, both in terms of the observations that the system can make and the actions that the system can perform. Moreover, the algorithms should be able to operate over long periods of time while achieving different objectives. The proposed algorithms will address three key problems related to lifelong RL: planning, exploration, and task decomposition. Planning is the problem of computing an action selection strategy given a (possibly partial) model of the task at hand. Exploration is the problem of selecting actions with the aim of mapping out the environment rather than achieving a particular objective. Task decomposition is the problem of defining different objectives and assigning a separate action selection strategy to each. The algorithms will be evaluated in two realistic scenarios: active network management for electrical distribution networks, and microgrid management. A test protocol will be
developed to evaluate each individual algorithm, as well as their combinations.

**Principales activités**

The PhD candidate will focus on one or more issues related to sequential learning in structured and evolving problems. The PhD candidate will first acquire expertise in different topics of machine learning such as online learning, multi-armed bandit, statistical learning theory, reinforcement learning, approximate dynamic programming, and algorithmic game theory. Then, the PhD candidate is expected to contribute to the advancement of the literature on this problem along many different lines: methodological (e.g., definition of general abstract models for a wide range of decision-making problems), theoretical (e.g., near optimality performance guarantees), and algorithmic (e.g., development of novel algorithms for specific decision-making problems). The research activity of the PhD candidate will be closely related to EU Chistera Delta project (http://www.chistera.eu/projects/delta). This will allow the PhD candidate to develop collaborations with other researchers participating in this research project and it may also allow him/her to spend part of his research activity at partner laboratories such as Montanuniversitat Leoben (Austria), Università Pompeu Fabra (Spain) or Université de Liège (Belgium). Possibility of internships in the industry research labs such as: Adobe Research in California or DeepMind in UK. The starting date of PhD program is flexible.

**Compétences**

The applicant must have a Master of Science in Computer Science, Statistics, Math, or related fields, possibly with the background in reinforcement learning, bandits, or optimization.

The working language in the lab is English, good written and oral communication skills are required.

**Avantages sociaux**

- Subsidised catering service
- Partially-reimbursed public transport
- Social security
- Paid leave
- Flexible working hours
- Sports facilities

More information about Lille:

http://www.lille3000.eu/portail/

http://www.lillemetropole.fr/mel.html

**Rémunération**

The gross monthly salary is 1982€ for 1st and 2nd year and 2085€ gross for the 3rd year.