Beyond the desired safety requirements that can be obtained by providing provably near-optimal methods, it remains a challenge to build the next generation of sequential learners, that can interact with complex systems in an autonomous way. Despite key advances in the understanding of sequential decision making under uncertainty over the last decade, specifically in multi-armed bandit theory and in reinforcement learning, much remains to be done in order to build the next generation of sequential learners, that can interact with complex and possibly evolving systems in an autonomous way. Beyond the desired safety requirements that can be obtained by providing provably near-optimal solutions, the exploration-exploitation trade-off is a fundamental challenge in sequential decision making. Should we follow a strategy that has been used a lot in the past and has empirically proven good until now (exploitation), or shall we explore a less known but potentially promising strategy (exploration)? A key difficulty to design a solution for these problems is that typically, when a decision is made, one only gets to see a noisy effect of that decision, and little about the effect of other alternatives. This gives rise to the study of the fundamental exploration-exploitation trade-off: Shall we follow a strategy that has been used a lot in the past and has empirically proven good until now (exploitation) or shall we explore a less known but potentially promising strategy (exploration)?

A solid mathematical background, specifically in mathematical statistics and/or optimization, information theory, and machine learning is required. A strong commitment for doing research, a taste for teamwork, for theory, for applications, and basic pedagogical skills are also required. Good writing and speaking abilities in English are essential.

Requirements:
- A Ph.D. in mathematics, statistics, computer science, or related fields.
- Experience in reinforcement learning or a related field.
- Strong research skills and a proven track record of publication.
- Excellent communication skills, both written and verbal.
- Experience in collaboration with industry and academia.

Benefits:
- Competitive salary and benefits package.
- Opportunities for professional development and career growth.
- Access to a dynamic and collaborative research environment.
- Opportunities for international travel and conferences.

About Inria
Inria, the French National Institute for computer science and applied mathematics, promotes "scientific excellence for technology transfer and society". Inria employs a staff of 2600, drawn from the world's most prestigious universities, to address the latest challenges in computer science and mathematics. The Institute operates eight Research Centres throughout France.

About the research centre or Inria department
Inria Lille - Nord Europe Research Centre was founded in 2008 and employs a staff of 360, including 300 scientists working in sixteen research teams. Recognised for its outstanding contribution the socio-economic development of the Nord - Pas-de-Calais Region, the Inria Lille - Nord Europe Research Centre undertakes research in the field of computer science in collaboration with a range of academic, institutional and industrial partners.

The strategy of the Centre is to develop an internationally renowned centre of excellence with a significant impact on the City of Lille and its surrounding area. It works to achieve this by pursuing a range of ambitious research projects in such fields of computer science as the intelligence of data and adaptive software systems. Building on the synergies between research and industry, Inria is a major contributor to skills and technology transfer in the field of computer science.

Context
The position is funded by the French ANR (Agence Nationale de la Recherche), with the context of the project: "BAnDits Against non-Stationarity and Structure" (BADASS). This is a four-year project, headed by Odalric-Ambry Maillard, in collaboration with Emilie Kaufmann and Richard Combes, that focuses on non-stationarity and structure in multi-armed bandit and reinforcement learning theory. It has specific funding to invite researchers, organize workshops, train students and generate a rich scientific activity around these questions. It also involves strong international connections with several research institutes abroad.

Understanding the dynamics of complex systems, and how to optimally act in them impacts all aspects of human societies where a careful management of natural, energetic, human and computational resources is required. To overcome the limitations of human capabilities to process large amounts of data, researchers from the field of machine learning and mathematical statistics for sequential decision making pursue the long-term goal of developing an optimal and automatic method that can, from partial observations and sequential interactions with a complex system, learn an optimal behavior.

A key difficulty to design a solution for these problems is that typically, when a decision is made, one only gets to see a noisy effect of that decision, and little about the effect of other alternatives. This gives rise to the study of the fundamental exploration-exploitation trade-off: Shall we follow a strategy that has been used a lot in the past and has empirically proven good until now (exploitation) or shall we explore a less known but potentially promising strategy (exploration)? Despite key advances in the understanding of sequential decision making under uncertainty over the last decade, specifically in multi-armed bandit theory and in reinforcement learning, much remain to be done in order to build the next generation of sequential learners, that can interact with complex and possibly evolving systems in an autonomous way.

Beyond the desired safety requirements that can be obtained by providing provably near-optimal solutions, the exploration-exploitation trade-off is a fundamental challenge in sequential decision making. Should we follow a strategy that has been used a lot in the past and has empirically proven good until now (exploitation), or shall we explore a less known but potentially promising strategy (exploration)? A key difficulty to design a solution for these problems is that typically, when a decision is made, one only gets to see a noisy effect of that decision, and little about the effect of other alternatives. This gives rise to the study of the fundamental exploration-exploitation trade-off: Shall we follow a strategy that has been used a lot in the past and has empirically proven good until now (exploitation) or shall we explore a less known but potentially promising strategy (exploration)? Despite key advances in the understanding of sequential decision making under uncertainty over the last decade, specifically in multi-armed bandit theory and in reinforcement learning, much remain to be done in order to build the next generation of sequential learners, that can interact with complex and possibly evolving systems in an autonomous way.

General Information
- Theme/Domain: Optimization, machine learning and statistical methods
- Town/city: Villeneuve d'Ascq
- Inria Center: CRI Lille - Nord Europe
- Starting date: 3/1/18
- Duration of contract: 1 year, 3 months
- Deadline to apply: 4/2/18

Contacts
- Inria Team: SEQUEL
- Recruiter: Maillard Odalric / odalric.maillard@inria.fr

The keys to success
A solid mathematical background, specifically in mathematical statistics and/or optimization, information theory, and machine learning are required. A strong commitment for doing research, a taste for teamwork, for theory, for applications, basic pedagogical skills are also required. Good writing and speaking abilities in English.

Required Diploma and experience: A Ph.D. in Machine Learning or Applied Mathematics, and an expertise in Reinforcement Learning (Bandits, Markov Decision Processes, etc.) and Mathematical statistics.

Conditions for application
- Defence Security: This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST). Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

Recruitment Policy: As part of its diversity policy, all Inria positions are accessible to people with disabilities.

Warning: you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.
performance guarantees, the two intertwined bottlenecks that must be studied and currently prevent the development of the field include the adaptation to the non-stationarity of the environment, and the often partially hidden structure of the environment generating the observation signal.

All travel expenses are covered within the limits of the scale in force.

Assignment

Assignments: The objective of this postdoctoral position, under the direct supervision of Odalric-Ambrym Maillard, is to study and develop novel decision-making strategies to handle the possibly unknown structure of the signals in the context of sequential prediction and reinforcement learning. Specific emphasis will be given to the notion of structure in BAnDits and to strategies designed for Markov Decision Processes with average gain criterion.

The proposed strategies will be developed with a generic application purpose and thus be given theoretically grounded performance guarantees under application-friendly assumptions.

For a better knowledge of the proposed research subject:
Please send an email to odalric.maillard@inria.fr, and visit the page of the team SequeL https://team.inria.fr/sequel/, or the Badass project https://project.inria.fr/badass/

Collaboration:
The successful candidate is expected to interact with the researchers in the sequel team and to strengthen his/her network of external collaborators.

Responsibilities:
The successful candidate will work along different work packages of the ANR project, especially the ones related to structure. He or she will also be encouraged to participate in the organization of a workshop or a tutorial as well.

Steering/Management:
The person recruited will be in charge of strengthening his or her publication record with novel and strong research articles, in order to get a good position in either academia or the industry.

Main activities

The successful candidate will be part of an ambitious research program, at the frontier of mathematical statistics, information theory and machine learning for sequential decision making focusing on reinforcement learning.

The standard activities of a Postdoctoral fellow include publishing research articles in top conferences or journals of the fields, communicating actively with other members of the team, strengthen his/her network of collaborators, and try novel and disruptive ideas. Meanwhile, the candidate will have the opportunity to enjoy the very beautiful, alive and welcoming city of Lille, as well as top working conditions for doing research in France.

Skills

Technical skills and level required: PhD in a field related to Machine Learning and/or Reinforcement Learning

Languages: fluent in English.

Benefits package

- Subsidised catering service
- Partially-reimbursed public transport

Remuneration

With a fixed term contract in the public administration, the salary will be 31836€ gross per year.