Research Internship

"Challenges of Sequential Decision Modelling in AgroEcology"

Odalric-ambrym Maillard
odalric.maillard@inria.fr
Scool, Inria Lille – Nord Europe

Vincent Levavasseur
vincent@verdeterreprod.fr
Ver-de-terre Production

Keywords: Reinforcement Learning, Markov Decision Processes, Dynamical Systems, Control Theory.
Supervision: The intern will be co-advised by Odalric-Ambrym Maillard from Inria team-project Scool and Vincent Levavasseur from Ver-De-Terre Production.
Place: This internship will be primarily held at the research center Inria Lille – Nord Europe, 10 avenue Halley, 59650 Villeneuve d'Ascq, France, in the Inria team-project Scool (previously known as SequeL). The intern will have the possibility to go on mission and interact with Ver-De-Terre Production.

Reinforcement Learning is one of three main Machine Learning paradigm (together with supervised and unsupervised learning). It focuses on formalizing and addressing Decision making in the context of a partially known dynamical system. The most common formalism used today in Reinforcement Learning is that of discrete time, rested Markov Decision Processes (MDP). This indeed enables to handle a variety of situations and to apply techniques such as value iteration of dynamic programming in order to compute an optimal decision policy. The pillars of MDP formalism are the notion of state, that enables to describe the instantaneous status of the system at some time, action that formalizes the interaction with the learner, as well as the reward and state transition functions that describe how good an action is considered and how the system evolves following an action.

The context of AgroEcology naturally involves Sequential Decision making. Indeed a farmer constantly interacts with a land, seen as a partially known dynamical system, by means of sequential observations and actions, and at least at a high level, the farmer has a global objective that is to grow plants. Hence, at first sight, it seems that observations can be a good indicator of the notion of states, actions of the farmers naturally map to actions in an MDP, the farmer's objective can lead to a reward function, and the evolution of the system maps to the transition function. Such a view is beneficial. However, the notion of MDP classically considered in the Reinforcement Learning literature also imposes stringent restrictions, that may cause a naive formalism to yield huge state-action spaces and miss some important parts of the farmer's decision making process. For instance, the farmer constantly plans under a variety of constraints, incorporating uncertain predictions, makes experiments, need to adjust plan depending on events, rather than fixed time steps, and has an objective that is a complicated mix of ecological, economical and human aspects.

In this internship, you will focus on the modeling of the farmer's decision process. You will challenge the MDP model, in its efficiency to describe the farmer's decision making process. A number of other model exists especially in the control literature (including e.g. event-based systems, renewal processes). You will first establish a literature review of the different models, showing their advantage and disadvantage with respect to standard MDPs, in particular in terms of their ability to accurately model farmer's decisions, and the efficiency in solving them. To assist in this task, you will also look at the way the farmer's decisions are considered in the agronomy literature on Decision Systems, and have the unique opportunity to discuss with farmers experts, via Ver-De-Terre production.

Beyond bibliography and its relevance to address the challenges of sequential decision making, you will illustrate the benefit of each formalism by coding them on some simplified synthetic environments. This would also be a good place to try some simple and sound ideas for a learning agent to make better decisions.

The successful candidate is expected to have a solid mathematical background in statistics and machine learning, especially in reinforcement learning, have good to excellent programming skills, have good formalism skills and to learn quickly. No prior agroecology knowledge is expected, but a taste for the topic is appreciated.

Here are a few bibliography pointers:

- Markov Decision Processes, Martin Putterman, 1994.
- A tutorial on event-based optimization: a new optimization framework, Xia, Jia, Cao, Discrete Event Dynamic Systems, 2014. https://link.springer.com/article/10.1007/s10626-013-0170-6
- Causal Reinforcement Learning tutorial, Bareinboim, ICML, 2019. https://crl.causalai.net
- Causal variables from reinforcement learning using generalized Bellman equations, Herlau, 2020. https://arxiv.org/pdf/2010.15745.pdf
- Numerical methods for simulation and optimization of piecewise deterministic Markov processes : application to reliability, B. de Saporta, F. Dufour, H. Zhang, Mathematics and statistics series Wiley-ISTE, 2015.
- Tightening Exploration in Upper Confidence Reinforcement Learning, Bourel, Maillard, Talebi, ICML 2020. https://arxiv.org/abs/2004.09656
- Improved exploration in factored average-reward MDPs, Talebi, Jonsson, Maillard https://arxiv.org/pdf/2009.04575.pdf
- Mathematics of Statistical Sequential Decision Making, Maillard, Habilitation thesis, 2019 https://hal.archives-ouvertes.fr/tel-02162189
- Agriculture softwares and platforms: https://smag.tech, https://www.agroptima.com/fr/cahier-de-culture, https://agritux.net, https://mesparcelles.fr/, https://ekylibre.com/
- The ontologies community of practice: A CGIAR initiative for big data in agrifood systems. Arnaud et al. 2020. Patterns, 1:100105, 13 p. https://doi.org/10.1016/j.patter.2020.100105

CENTRE DE RECHERCHE
LILLE – NORD EUROPE
Parc scientifique de la Haute-Borne
40 avenue Halley - Bât A - Park Plaza
59650 Villeneuve d'Ascq - France
Tél. : +33 (0)3 59 57 78 00
Fax : +33 (0)3 59 57 78 51

www.inria.fr