Proposition de recherche doctorale

Interpretability and ethical issues in sequential decision problems

Mots clés :
- Directeur de thèse : Aurélie Beynier
- Co directeur de thèse : nicolas MAUDET
- Co-encadrant(s) :
- Unité de recherche : Laboratoire d'informatique de Paris 6
- Ecole doctorale : École Doctorale Informatique, Télécommunications, Électronique de Paris
- Domaine scientifique principal: Sciences et technologies de l'information et de la communication

Résumé du projet de recherche (Langue 1)
Markov Decision Processes (MDPs) and reinforcement learning (RL) are two very successful paradigms adopted in artificial intelligence for designing autonomous agents capable of dealing with sequential decision problems. While traditionally the "reward function", a basic component of both MDPs and RL, is assumed to be known from the start, it has been noted that in many situations it can be hard to fully specify it. In the last decade several works [3,4,5,6] proposed "preference-based" versions of MDPs and RL addressing the problem of dealing with incompletely specified reward function. The fact that the reward is not specified apriori but it is acquired opens a number of crucial issues that need to be tackled. The first one is that the autonomous agent should be able to explain why a policy is recommended; in particular the policy of an MDP maximizes expected utility and may be hard to understand for an human user [2]: producing a compact representation of a policy that can be used for communicating with the user is therefore a challenging problem. The second issue is that, in an open world, any behavior can be learnt apriori, and therefore we should investigate how to avoid "unethical" behaviors. In a recent foundational paper [1], Abel et al. envisions the prospect of using reinforcement to model an idealized ethical artificial agent. Among the several issues that this approach raises, and identified by Abel et al., we can cite the problem of teaching the agent, and of making policies interpretable. This PhD proposal follows these lines of research and investigates the prospect of using and explaining MDPs / RL policies respecting ethical requirements. We will investigate how an "ethical" utility function might be acquired from an expert using an interactive elicitation method. [1] D. Abel, J. MacGlashan, M.L. Littman: Reinforcement Learning as a Framework for Ethical Decision Making. AAAI Workshop 2016 [2] O. Zia Khan, P. Poupart, J. P. Black: Minimal Sufficient Explanations for Factored Markov Decision Processes. ICAPS 2009 [3] L. Torrey, M. E. Taylor: Teaching on a budget: agents advising agents in reinforcement learning. AAMAS 2013 [4] K. Regan, C. Boutilier: Regret-based Reward Elicitation for Markov Decision Processes. UAI 2009 [5] R. Akrour, M. Schoenauer, M. Sebag: APRIL: Active Preference Learning-Based Reinforcement Learning. ECML/PKDD 2012 [6] P. Weng, B. Zanuttini: Interactive Value Iteration for MDPs with unknown rewards