LSINF2275 Project II
« Data mining and decision making »
Markov Decision Processes 2

Professor:	Marco Saerens marco.saerens@uclouvain.be
Address:	Université catholique de Louvain
	LSM & ICTEAM
	Place des Doyens 1
	B-1348 Louvain-la-Neuve
	Belgique
Telephone:	010 47.92.46.
Fax :	010 47.83.24.
Teaching assistants:	Bertrand Lebichot, Pierre Leleux

Objective

The objective of this project is to put into practice some algorithms for solving Markov decision processes and reinforcement learning discussed in the "data mining and decision making" lectures. This will be done through either the study of a new practical case or the study of theoretical extensions of the basic algorithms. This work is very open in that you will have the choice of the studied application or extension.

Assignment statement

As for the first project, you will be working in groups of 2 students exactly – the same groups as for the first project.

In this second assignment, you are asked to study an **interesting practical application** of Markov decision processes (MDP), reinforcement learning (RL), bandit problems (BP), or a **theoretical extension** of the basic algorithms together with an empirical validation. Good starting points are the two following books, together with their web sites:

- Sutton & Barto (2018) "Reinforcement learning: an introduction, 2nd ed.". MIT Press. Preprint available at http://incompleteideas.net/book/the-book-2nd.html. (chapter 16 describes some applications).
- Powell (2011) "Approximate dynamic programming, 2nd ed.". Wiley.
- Bertsekas & Tsitsiklis (1996) "Neuro-dynamic programming". Athena Scientific. (chapter 8 describes some applications).
- Powell & Ryzhov (2012) "Optimal learning". Wiley
- Szepesvari (2010) "Algorithms for Reinforcement Learning". Morgan & Claypool Publishers.

The topic is very open – more concretely, here are some potential ideas:

- Find (either in the literature, or create a new one) a nice practical application of MDP, RL or BP, for instance a game; describe it, implement it and validate it empirically.

- Tackle an ambitious practical application on OpenAI, https://openai.com/systems/ with MDP or RL.

- Investigate a theoretical extension of MDP, RL or BP. For instance, Monte-Carlo methods in RL, together with importance sampling. Describe briefly the theory behind the method and validate it on a practical.

- Investigate a theoretical extension of MDP, RL or BP. For instance, n-step SARSA methods in RL. Describe briefly the theory behind the method and validate it on a practical application.

- Investigate a theoretical extension of MDP, RL or BP. For instance, approximate RL based on neural networks. Describe briefly the theory behind the method and validate it on a practical application.

- Investigate a theoretical extension of MDP. For instance, an exploration strategy like, e.g., randomized shortest paths. Describe briefly the theory behind the method and validate it on a practical application.

- Investigate some potential solutions to deal with systems with a large (exponential) number of states in MDP, RL or BP and compare them empirically on a practical application.

- Investigate some potential solutions to deal with partially observable Markov decision processes (POMDP) and compare them empirically on a practical application.

The work will be based on scientific papers or books and should contain both a theoretical contribution (description of the method) as well as an empirical validation. The balance between the two components depends on your wish and expectation: the orientation of the work can be more theoretical or more practical. If, for instance, you decide to face a complex application, you can limit yourself to a description of the basic algorithms you used (MDP, RL or BP). Conversely, if you decide to study a theoretical extension, the practical application can just be the “Snakes and Ladders” game investigated in the first project.

The implementation should be in *Matlab, Octave, Python, Julia* or *R*.

Report

Please do not forget to mention your affiliation on the cover page, together with your name (SINF, INFO, MAP, STAT, BIR, DATS, etc).

You are asked to write a report (in English) of maximum 7 pages (without the code which has to be delivered in a different file), preferably in latex. This report will contain

- a description of your objectives and (theoretical + empirical) contributions.
- a synthesis of the theory needed to solve your problem.
- a description of the investigated empirical application.
- a short description of your implementation along with a presentation and a discussion of the results and comparisons. Please include graphics when relevant.
- bibliographical references.

This report must be uploaded on Sunday, May 20, 2018, before 23:55, together with the code (all files zipped together) on Moodle in the section « Assignments ». Do not forget to comment your code. This first project accounts for 3 points on 20. The weight of the second project will be 4 on 20. Thus, the final oral exam amounts to 13/20.

Good Work !