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.
Teaching assistants:	Sylvain Courtain, Pierre Leleux

Objective

The objective of this project is to put into practice some algorithms for solving Markov decision processes, reinforcement learning, or bandit problems 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 standard 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 3 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 on, for instance, the snakes and ladders game (already used in the first project). Good starting points are the following books or papers, 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

- Russell & Norvig (2010) "Artificial intelligence: a modern approach", 3d ed. Prentice-Hall.

- Szepesvari (2010) “Algorithms for reinforcement learning”. Morgan & Claypool Publishers.

- Phil Winder (2020) “Reinforcement learning: industrial applications of intelligent agents”. O’Reilly.

- Note also that several papers describing practical applications of Markov decision processes have been posted on Moodle.

Thus, the topic is **very open** – more concretely, here are some potential ideas:

- Find (either in the literature, or consider 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 application, which could be the snakes and ladders game.

    - n-step SARSA methods in RL. Describe briefly the theory behind the method and validate it on a practical application, which could be the snakes and ladders game.

    - Approximate RL based on simple linear regression or more complex neural networks. Describe briefly the theory behind the method and validate it on a practical application, which could be the snakes and ladders game.

- 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, which could be the snakes and ladders game.

- Investigate some potential solutions to deal with partially observable Markov decision processes (POMDP) and compare them empirically on a practical application, which could be the snakes and ladders game.

The work will be based on books or scientific papers and should contain a **theoretical contribution** (description of the used techniques) 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 practical 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 normally be in *Python* but, if needed, you could also use *Matlab, Octave, Julia* or *R* (if, for example, you find a nice environment already implemented in one of these languages).

Deliverables

The main deliverable of this project is a **video capsule**, of maximum 10 minutes, presenting your work (in French or English). For this video, you could simply use Teams or any video recording software such as, e.g., OBS studio.

You are also asked to write a **brief 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 brief synthesis of the theory needed to solve your problem.
- a description of the investigated empirical application (if relevant).
- a short description of your implementation along with a presentation / discussion of the results and comparisons. Please include graphics when relevant.
- bibliographical references.

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

The report and the code (zipped together) must be uploaded on Sunday, May 16, 2021, before 23:55 on Moodle in the section « Assignments ». Instructions for the delivering of the video capsule will be provided at the end of the semester (in May). This second project accounts for 5 points on 20, as the first one. Thus, the final oral exam amounts to 10/20.

Good work !