Reinforcement Learning

0. Class info

Mario Martin

CS-UPC

February 17, 2021

Contact information

Advanced Techniques in Computational Intelligence (previously second part of the URL course):

- Mario Martin
- Office 202, second floor Omega Building.
- email: mmartin@cs.upc.edu
- Web page for course: https://www.cs.upc.edu/~mmartin/ATCI-RL.html

In coronavirus times:

We cannot plan in advance because situation is changing so, **check** the *Racó* **regularly** for news about the course. During coronavirus times we will work as follows:

In coronavirus times:

We cannot plan in advance because situation is changing so, **check** the *Racó* **regularly** for news about the course. During coronavirus times we will work as follows:

- Each week (starting from February 15th) you will find, in the web page of the course, videos commenting the set of slides for that week.
- You are free to advance in your own pace with the slides, but remember the deadlines for evaluation works you will see later.
- Each Friday (starting from February 19th), on scheduled class time, I will open a *meet* session to answer questions and to have live feedback about the course. I will publish the link to the meet session on the *Racó*.
- As soon as we can return to *normal life* with live classes, we will go back to the standard way of teaching.

Syllabus

First part:

1. Basic concepts of Reinforcement Learning
2. Basic RL algorithms: Model based methods
   - Value iteration and Policy iteration
   - Asynchronous versions
3. Basic RL algorithms: Model free methods
   - Backups and Role of exploration
   - Monte Carlo, Q-learning and Sarsa
   - Temporal Differences: TD(λ)
4. Function approximation in RL
   - Non-parametric methods and linear methods
   - Deep Learning for RL
5. Policy gradient methods
   - Actor-Critic methods
   - REINFORCE and TRPO
   - DDPG, TD3, SAC

Syllabus

Second part:

6. Monte Carlo Tree Search and Alfa-Zero algorithms
   - Bandit problems
   - Monte Carlo Tree Search
   - Alpha Zero
7. Sample efficiency I: Model Based Reinforcement Learning (MBRL)
8. Sample efficiency II: Sparse Rewards problem
   - Exploration, Curricular Learning, Hierarchical Learning, Hindsight Experience Replay (HER)
9. Sample efficiency III: Multiple task and life-long learning
   - Transfer learning, Multi-task learning, Meta-learning
10. Multi-agent RL
    - RL in Cooperative, Competitive and Mixed cases problems

Resources

- Slides: On the web page of the curse.
- Books:
  - Sutton and Barto, *An Introduction to Reinforcement Learning*, 2nd Edition (2018). MIT Press. Available here.
  - Szepesvari. *Algorithms for Reinforcement Learning* (2010). Morgan and Claypool. Complete draft available here.
- Recommended courses with materials:
  - (Basic) David Silver's course Reinforcement Learning, 2015.
  - (Advanced) Sergey Levine's course CS 294 (Berkeley): Deep Reinforcement Learning, Fall 2018.
- Software: See Lab section in web page of the course.

Course Evaluation

The evaluation of the course will consist on three parts:

1. A test questionnaire about the topics of the course
2. Implementation of a domain and/or a reinforcement learning algorithm (f.i. in the OpenGym framework and python).
3. A research paper about the current state of the art of one topic related to RL.

Final grade will be resulting of this formula:

$$\text{Grade} = 0,20 * \text{Test} + 0,30 * \text{ Implementation} + 0,5 * \text{Paper}$$

Course Evaluation: Quiz

- Test will be about basics of RL that will include the topics covered in the first part of the course (see syllabus)
- I will provide a personalized test to each student that has to be filled and returned after a week.
- Test can be done at home. Quizs will be available to students and also delivered through the "Practicals" section of the *racó*.
- The test will done approximately during week from **5th to 12th of April**.

Course Evaluation: Implementation

- Work could consist in implementing an algorithm, an environment or a exploration technique explained in class
- You are free to choose the task and the algorithm
- The implementation must be in the python language and one of the platforms proposed.
- You will have to write short report about the implementation and results.
- The deadline for delivering this work is **May 7th 2021**.

Course Evaluation: Paper

- You will have to write brief paper about the current state of the art of the research in a advanced topic covered in the second part of the course.
  1. You will have to look for papers related to the topic
  2. You will have to collect relevant bibliography on the topic
  3. You will have to choose the more relevant papers
  4. You will have to summarize the problems described and to summarize the approaches presented on those papers
  5. You will have to situate the problem in the area and to describe the relation of the topic with other areas
- You will find a list of topics in the web page of the curse. Choose the topic and send me and e-mail. Each student should choose a different topic. Topics will be assigned using the first-to-choose-first-to-assign policy.
- You can work also on a topic not in the list *after my approval*.
- The deadline for delivering this report is **June 21th 2021**.

Lab classes

- Any `OS` is Ok.
- I will assume some knowledge of `python`. I will assume also you have `python 3.x` installed in your computer. Install Anaconda distribution if you don't have any installed.
- Don't worry. We will install the different platforms to play with different algorithms and environments. Instructions will be available also in the web page of the course.