AI: REINFORCEMENT LEARNING

Grado en Computación e Inteligencia Artificial / Bachelor in Computer Science and Artificial Intelligence BCSAI SEP-2025
AIRL-CSAI.3.M.A

Area Computer Science
Number of sessions: 30
Academic year: 25-26
Degree course: THIRD
Number of credits: 6.0
Semester: 2º
Category: COMPULSORY
Language: English

Professor: **JOSÉ MANUEL REY GONZÁLEZ**

E-mail: jreyg@faculty.ie.edu

Consultant Director of numerous projects of strategic planning and business development, market studies and quality improvement, revenue growth and cost optimization, business processes reengineering, information systems design and implementation, for companies in virtually all industries, in both public and private sectors, with a systemic approach based on the identification of innovative competitive strategies, efficient organizational and process transformation and a firm leverage on the impulse of new technologies

Education

He is a Civil Engineer graduated with honors, special End Of Career Award, Bachelor of Economics and Business Administration, PhD student (unfinished) in Applied Quantitative Economics, and holds a CISA Title from the EDPAA (now ISACA) and a Master in Corporate Finance IESE, University of Navarra.

Professional Background

He became International Partner in Arthur Andersen in 1997, led the area of Innovation and Business Transformation, responsible for comprehensive quality management services, shared services centers, process reengineering and technological integration and the eBusiness area for strategic planning and development of interactive business and electronic commerce nationwide.

International Director of the World Excellence Center for Business Process Reengineering, member of the eBusiness leadership team in EMEIA and coordinator of this practice in the Mediterranean region.

He was a Managing Director in KPMG Consulting/BearingPoint, responsible for the Technology industry within the Telecommunications, Media and Content Sector and leader of the area of Process Innovation Solutions and Advanced Technology

He has served as Member of Board of Directors, Advisory boards or Management Consultant in matters of strategy and business development for several Spanish companies and Advisor of Venture Capital Companies and M&A Firms.

Arbitrator of the Madrid Arbitration Court, specialist in the Technology, Information and Communications sector and Developer of an advanced software platform for modelling business risks and economic projections, valuation of companies & portfolios and for building advanced risk & value information systems for Directors.

Public Activity and Teaching

He has been part of the Advisory Committees of eMobility and SIMO. He was a member of the Jury of the National eMobility Prize and participated in the Competitiveness Forums of Madrid and Castilla y León.

Instructor and speaker in numerous courses and seminars, national and international, of a general nature and related to his areas of specialization. Lecturer in various Business Management Master courses in subjects related to advanced economic and statistical analysis as well as organizer of the Master eBusiness of the Universidad Pontificia de Comillas, Speaker in the Executive Education Program of ESADE and in the Executive MBA of AEDE Business School in the subjects of "Information Systems", "Technological Environment" and "Systems Architecture".

He has published articles on business and technical content topics in newspapers such as Expansión, Cinco Días, Gaceta de los Negocios, Actualidad Económica, Dinero, etc., and directed or participated, as author or coordinator, in the publication of various books on issues of quality, organization and information technology (Las empresas del click / The Clickable Corporation [1999, Actualidad Económica], La Calidad en España/Quality in Spain [1995, Cinco Días]) and collaborated in the development of others (The organization in the information age: Learning, Innovation and Change [1995, IESE])

He was listed in the year 2000 by Actualidad Económica in the ranking of the 100 most relevant people on the Internet and eBusiness in Spain.

E-Mail: jreyg@faculty.ie.edu

Office Hours

Office hours will be on request. Please contact at:

jreyg@faculty.ie.edu

SUBJECT DESCRIPTION

Reinforcement Learning (RL) is a melting pot where excellent computational techniques of dynamic programming, heuristic search in problem spaces and the new techniques of functional representation of neural networks have converged with great success to build a highly scalable area of discipline that is allowing to efficiently attack a wide range of problems previously considered as practically unapproachable.

Throughout the course we will analyze each of these components in a structured way and we will see the types of approach combinations currently in use, from basic Q-Value models, to Deep QL and methods based on policy gradients, actor-critic, curiosity driven and other innovations in the dynamic definition of reward functions.

We will try to understand in which problem scenarios they are most advisable, their strengths and limitations and to anticipate some of the future trends that will progressively lead to advances and breakthroughs in this field.

LEARNING OBJECTIVES

The main objective of this course is to introduce students to the exciting field of RL and to set up a framework of knowledge that will allow them to make informed analysis of the opportunities and challenges for its successful application in Business.

To achieve this, we will focus on helping students grasp a sound understanding of RL techniques and their variations (model-free, model-based, online/offline RL, behavioural cloning, etc) and a good intuition of how to apply them to tasks of their own design.

It encompasses knowledge about:

- the fundamentals and theoretical principles of this discipline,
- the most relevant algorithms used in practice,
- the engineering process for the development and management of this type of projects and
- the understanding of the difficulties, obstacles and challenges to be faced in this path.

Therefore, it will combine both a theoretical and conceptual approach with a hands-on technical understanding of the different stages of these type of RL projects keeping always a clear perspective of the business issues and opportunities and the future evolution and innovations from the current state of the art.

To streamline the technical/computational requirements, we will concentrate on the implementations using Python, and a level of complexity ranging from basic to intermediate.

For some advanced examples we could also touch on some other development environments (such as LUA for robotics or Matlab for process-control) but just for showcase purposes. No programming activities will be required in these cases.

TEACHING METHODOLOGY

This program will try to put you in a good starting point for becoming a well-rounded RL practitioner, experiencing how to build RL models and apply them to problems in various fields.

We have tried to organize the different knowledge blocks integrating the conceptual and theoretical background first, then moving into practice/tutorials examples and then framing assignments of assorted levels of complexity.

We expect students to work 3 to 8 hours/week on average. This is a rough estimation of total hours the average student may take to complete all required coursework, including lecture and project time (Actual hours may vary depending on the individual student and each specific stage of the course)

The course will be both lecture and example-based, and will include group in-class discussions to promote learning and understanding of the course material in a variety of formats.

The course will have 6 main elements:

Lectures: We will explain the theoretical ideas, concepts and methods involved and will try to check on-time the correct understanding of the key elements. Questions and feedback will be the basic tool for this interactive dialog, in order to guarantee that the message gets across.

Examples/Tutorials/Cases: We will use profusely cases and examples related to the theory (and preparatory for the Assignments). Questions and what-if interactive analysis will be encouraged.

Discussions: as one critical skill for your future is how you present your work, we will encourage from time to time, for the relevant concepts in each section, some group discussions. A few of them will be announced in advance and will require some preparation from each student.

Assignments: practical exercises for experimenting with RL concepts, algorithms and techniques

Exams: There will just one formal test/exam, and a “test-exam” assignment will be included to get students used to answer questions about RL in different types of formats

Group work: For the final part of the course the student will have to work in a small group for preparing a final project presentation.

Assignments:

The best way to learn about a machine learning method is to program it yourself and experiment with it. So, the assignments will generally involve implementing machine learning algorithms, and experimentation to test your algorithms on some data. You will be asked to summarize your work, and analyze the results.

The implementations will be done in Python, creating Jupyter notebooks or pure Python programs (in some cases, numba or cython could be required for performance optimization), and the structure and review criteria will be reported for each specific job.

Though they are considered as individual work, the "individual assignments" will be realized (and delivered jointly) in small groups that will be defined by an special algorithm that will dynamically form the groups for each assignment ensuring that no two students will be assigned more than once together and trying to optimize and combine the skills of each one based on their previous results to maximize cross-learning, Therefore, collaboration on the assignments is only allowed (and enforced) within each group. Each student/group is responsible for his or her own work. Discussion of assignments and programs should be limited to clarification of the handout itself, and should not involve any sharing of pseudocode or code or simulation results.

The only exception will be for the Final Group Assignment, for which another specific algorithm will try to create optimized groups based on the preferences informed by each student.

The schedule of assignments is included in the syllabus and in general will correspond to the sessions planned as "with assignment" except for exceptions required for the best development of the content.

Assignments are due at the beginning of class/tutorial on the due date. Because they may be discussed in class that day, it is important that you have completed them by that day.

Optionally, some of the assignments could have the form of a “bake-off” (a competition between machine learning algorithms). We will give everyone some data for training a RL system, and you will try to develop the best method. We will then determine which system performs best on some unseen test data.

Exams:

There will just one formal test/exam. However, as an especial practice/assignment the students will have a series of test-exams to get them prepared for answering questions about RL (multiple-choice, true-false or open questions). These test-exams will be either automatically scored by the machine or in some cases will require a self-evaluation or self-assessment. So, they are a type of simulation of an open-book exam on all material covered up to that point in the lectures, tutorials, required readings, and assignments, but eventually there could be topics not covered in any of these that will require some online investigation or other type of research to find out (or understand) the possible answers. The final scoring of this “assignment” will be based on the number of test-exams done (freely decided by each student), their average results, and the degree of progressive improvement/learning shown.

[Most of] The questions for the final test/exam will be selected from the ones used for the practice

Attendance:
We expect students to attend all lectures, and all practice/tutorials. This is especially important because we will cover material in class that is not included in the reference readings. Also, the tutorials will not only be for review, practicing and answering questions, but new material and concepts will also be covered or explained.

IE University teaching method is defined by its collaborative, active, and applied nature. Students actively participate in the whole process to build their knowledge and sharpen their skills. Professor's main role is to lead and guide students to achieve the learning objectives of the course. This is done by engaging in a diverse range of teaching techniques and different types of learning activities such as the following:

Learning Activity	**Weighting**	**Estimated time a student should dedicate to prepare for and participate in**
Lectures	16.7 %	25.0 hours
Discussions	6.7 %	10.0 hours
Exercises in class, Asynchronous sessions, Field Work	53.3 %	80.0 hours
Group work	10.0 %	15.0 hours
Individual studying	13.3 %	20.0 hours
TOTAL	100.0 %	150.0 hours

AI POLICY

Generative artificial intelligence (GenAI) tools may be used in this course for research, ideation, generating an outline, proofreading, grammar check, coding, generating diagrams or knowledge graphs and/or image generation, with appropriate acknowledgement.
GenAI can also be used for the Assignments and the practice of preparation of the Exam, but it will not be allowed during the final Exam.
Such use is encorauged with the goal of developing an informed critical perspective on potential uses and generated outputs, so the student is responsible for any errors or omissions and the validation of the outputs of GenAI. A clear lack of attention to this task could be taken into account negatively in the process of correcting and grading the work.
AI is a tool, but one that you need to acknowledge using. Failure to do so is in violation of academic honesty policies. Acknowledging the use of AI will not impact your grade.
Suggested format to acknowledge the use of generative AI tools:
"*I acknowledge the use of [AI systems link] to [specify how you used generative AI]. The prompts used include [list of prompts]. The output of these prompts was used to [explain how you used the outputs in your work]*".
If you have chosen not to include any AI generated content in your assignment, the following disclosure is recommended:
"*No content generated by AI technologies has been used in this assignment*"

DISCLAIMER

The following description of the material covered is tentative.
An attempt will be made to cover all listed topics.

However; the pace in the classes will depend on the group performance and the recent advances in the field.

PROGRAM

SESSION 1 (LIVE IN-PERSON)

INTRODUCTION TO THE COURSE (SYLLABUS)

Procedural Learning and RL motivation. Overview of contents and analysis of the approaches and the relevant issues in this field.

SESSION 2 (LIVE IN-PERSON)

RL INTRODUCTION

Building blocks of current RL algorithms and approaches.

Different convergent routes to current RL (from MPC, Markov Chains, Genetic algorithms, etc...)

Book Chapters: Reinforcement Learning: An Introduction (Chapter 1, Section 1.7 Early history of Reinforcement Learning) (See Bibliography)

SESSION 3 (LIVE IN-PERSON)

RL BASIC CONCEPTS

Definition of intelligent agents and their types. Evaluative vs Instructive feedback. Associative vs Non-associative models.

Book Chapters: Reinforcement Learning: An Introduction (Chapter 1, Section 1.3 Elements of Reinforcement Learning) (See Bibliography)

SESSION 4 (LIVE IN-PERSON)

MULTI-ARMED BANDITS

Introduction to contextual bandits and its understanding as a basis for full RL

Interesting "low-hanging fruits" practical examples for its direct application in some type of business problems.

Book Chapters: Reinforcement Learning: An Introduction (Chapter 2, Tabular Solution Methods, Multi-armed bandits.) (See Bibliography)

SESSION 5 (LIVE IN-PERSON)

GRAPHS, SEARCH, MDP's

The understanding of the search space and its representation and the possible dimensions of heuristic approaches for search, prediction and control. Introduction to Markov Decision Processes.

Book Chapters: Reinforcement Learning: An Introduction (Chapter 3, Tabular Solution Methods, Finite Markov Decision Processes) (See Bibliography)

SESSION 6 (LIVE IN-PERSON)

REVIEW PRACTICE - MDP's

Review and apply concepts about MDP's

Practical cases of MDP's. Intuition and analytical solutions (for simple cases). Understanding of techniques for its representation (backup diagrams).

SESSION 7 (LIVE IN-PERSON)

DYNAMIC PROGRAMMING

The iterative value function and the Generalized Policy Improvement techniques.

Book Chapters: Reinforcement Learning: An Introduction (Chapter 4, Tabular Solution Methods, Dynamic Programming) (See Bibliography)

SESSION 8 (LIVE IN-PERSON)

MONTE CARLO METHODS

Explanation of Monte Carlo methods and the mathematical foundations behind the (extremely versatile in ML) Importance Sampling technique. The merge of these approaches with the Dynamic Programming approaches

Book Chapters: Reinforcement Learning: An Introduction (Chapter 5, Mote Carlo Methods, Importance Sampling) (See Bibliography)

SESSION 9 (LIVE IN-PERSON)

TEMPORAL DIFFERENCES [Assignment: ASSIGNMENT]

The integration of the previous different concepts in the current cornerstone of RL (for finite-MDP-like problems). Bellman Optimality Equation.

Book Chapters: Reinforcement Learning: An Introduction (Chapter 6, Temporal-Difference Learning and Chapter 7, n-step Bootstrapping) (See Bibliography)

SESSION 10 (LIVE IN-PERSON)

A FIRST PRACTICE - INTEGRATION OF CONCEPTS

Application of the principles studied to an initial problem. Comparison with the other possible solutions (harcoded, genetic algorithm, hamiltonian paths, etc). Live comparison and charting of all of them and discussion of pros and cons. Rationale behind the use of RL approaches.

SESSION 11 (LIVE IN-PERSON)

LEARNING AND PLANNING

A deeper comparative study of the characteristics and differences of model-based vs model-free methods.

Book Chapters: Reinforcement Learning: An Introduction (Chapter 8, Planning & Learning with tabular methods) (See Bibliography)

SESSION 12 (LIVE IN-PERSON)

EVOLUTIONARY COMPUTATION

Genetic algorithms and beyond, Non-dominated multi-objective optimization, the Pareto frontier and NEAT

Exploration of a sometimes considered independent field (genetic algorihtms and other non-derivative-based methods). Practical examples of their use in practical problems such as multi-objective optimization. Analysis of the non-dominated genetic algorithms (introduction to the concept of Pareto frontier). Exploration of Neuro-Evolution techniques and the current and future lines of research for connecting these types of solutions with the mainstream RL approaches.

SESSION 13 (LIVE IN-PERSON)

RL FRAMEWORKS

Overview of main RL Frameworks

Comparison of some currently used RL frameworks (from OpenAI gym, stable baselines, etc to Ray and RLib). Pros, cons and requisites for their use.

SESSION 14 (LIVE IN-PERSON)

INTERMEDIATE PRACTICE

Analysis and comparison of different algorithmic methods for the training of these types of models.

SESSION 15 (LIVE IN-PERSON)

APPROXIMATION METHODS

The key change in the strategy: the calculus of approximative functions (for value, q(s,a) or policy functions). The main advantages of this approach. Focus in one of the main (most relevant in practice) approaches: ANN function approximators.

Book Chapters: Reinforcement Learning: An Introduction (Chapter 9, On-policy prediction with approximation. Section 9.7 Nonlinear Function Approximation: Artificial Neural Networks. Chapter 10, Chapter 11 and Chapter 12 [Eligibility traces]) (See Bibliography)

SESSION 16 (LIVE IN-PERSON)

POLICY GRADIENTS

Theory behind PGM's and the "Policy Gradient Theorem"

Book Chapters: Reinforcement Learning: An Introduction (Chapter 13, Policy Gradient Methods) (See Bibliography)

SESSION 17 (LIVE IN-PERSON)

ACTOR-CRITIC [Assignment: ASSIGNMENT]

Definition and analysis of the actor-critic methods. The concept of "advantage". Main algorithms.

Book Chapters: Reinforcement Learning: An Introduction (Chapter 13, Section 13.5 Actor-critic methods) (See Bibliography)

SESSION 18 (LIVE IN-PERSON)

PRACTICE REVIEW

Dissection and discussion of sample solutions. Exploration of main difficulties.

SESSION 19 (LIVE IN-PERSON)

PROBLEM APPROACH FRAMEWORK

Exhaustive analysis of a classical problem (The Travelling Salesman Problem) under all possible lines of solutions (from Christofides to genetic algorithm, ant-colony optimization and simple or complex Reinforcement Learning) Discussion of pros and cons and initial understanding of "When, Why and How" of their possible uses.

SESSION 20 (LIVE IN-PERSON)

MODEL-BASED, OFF-LINE, INVERSE RL

Introduction to the novel field of IRL (learning an agent's objectives, values, or rewards by observing its behavior). Connection to related concepts of Behavioural Cloning, Imitation Learning and Teacher Agents.

SESSION 21 (LIVE IN-PERSON)

DEEP RL PRACTICE [Assignment: ASSIGNMENT]

Connecting the dots between RL and approximation (ANN) functions. Application to an image-based game learning problem. Comparison of results.

SESSION 22 (LIVE IN-PERSON)

SUMMARY RECAP [Assignment: GROUP ASSIGNMENT]

Recap session for all the concepts and algorithms introduced so far. Introductory extension to some new concepts (Rainbow algorithms, etc)

SESSION 23 (LIVE IN-PERSON)

RESEARCH PAPERS, INNOVATION TRENDS (A) [Assignment: EXAMINATION PRACTICE. BE PREPARED TO ANSWER QUESTIONS ABOUT RL]

Study and discussion of some of the most relevant papers published in the field.

Summary of new approaches (reward shaping, finite automata rewards, hierarchical RL...)

Book Chapters: Reinforcement Learning: An Introduction (Chapter 17 Frontiers) (See Bibliography)

SESSION 24 (LIVE IN-PERSON)

RESEARCH PAPERS, INNOVATION TRENDS (B)

Continuation of the study and discussion of some of the most relevant papers published in the field.

Analysis of obstacles and challenges in the field and in its practical application to business problems.

SESSION 25 (LIVE IN-PERSON)

RL APPLICATIONS

Exploration of business cases of practical application across the different industries.

SESSION 26 (LIVE IN-PERSON)

MULTI-AGENT RL

Incursion in the multi-agent reinforcement learning (MARL) sub-field. Possible dynamics for the behavior of multiple learning agents that coexist in a shared environment. Extension (or generalization) of the basic principles to these type of complex models.

SESSION 27 (LIVE IN-PERSON)

TOP CASES & COMPANIES

Company leaders in this field. Study of breakthrough examples (from AlphaGo, AlphaFold-2, Fusion Reactor Modeling, and beyond)

Book Chapters: Reinforcement Learning: An Introduction (Chapter 16 Applications and Case Studies) (See Bibliography)

SESSION 28 (LIVE IN-PERSON)

TEST-EXAM

Concept summary review examination. Not to be considered as a final exam. Its score will be added to the rest of the evaluation items. No minimum passing grade required.

SESSION 29 (LIVE IN-PERSON)

GROUP PROJECTS PRESENTATIONS (A)

Presentations, discussion and Q&A of group projects

SESSION 30 (LIVE IN-PERSON)

GROUP PROJECTS PRESENTATIONS (B), WRAP-UP

Presentations, discussion and Q&A of group projects

EVALUATION CRITERIA

We will behave as professionals. This means that I expect you to come to class prepared to discuss as if this was a meeting in your company. This is a small group, so we will take advantage of it doing a lot of direct interaction. Come prepared to class and be inquisitive.

As stated before, attendance to the classes is important and to help you with the right incentives approximately one third of your grade will come from this attendance, another one third will be obtained through your individual assignments and the final third will be gained through class participation (including group interactions and discussions) and from the final group work and presentation.

The following table summarizes the Ongoing grading (maximum) for each of the sessions and assignments:

SESSION	LECTURE	Participation	I.Assignment	G.Assignment	Test/Exam	TOTAL
RLI_01	INTRODUCTION TO THE COURSE (SYLLABUS)	4				4
RLI_02	RL INTRODUCTION	4				4
RLI_03	RL BASIC CONCEPTS	4				4
RLI_04	MULTI-ARMED BANDITS	4				4
RLI_05	GRAPHS, SEARCH, MDP's	4				4
RLI_06	REVIEW PRACTICE - MDP's	4				4
RLI_07	DYNAMIC PROGRAMMING	4				4
RLI_08	MONTE CARLO METHODS	4				4
RLI_09	TEMPORAL DIFFERENCES	4				4
RLI_10	A FIRST PRACTICE - INTEGRATION OF CONCEPTS	4				4
RLI_11	LEARNING AND PLANNING	4				4
RLI_12	EVOLUTIONARY COMPUTATION	4				4
RLI_13	RL FRAMEWORKS	4				4
RLI_14	INTERMEDIATE PRACTICE	4				4
RLI_15	APPROXIMATION METHODS	4				4
RLI_16	POLICY GRADIENTS	4				4
RLI_17	ACTOR-CRITIC	4	70			74
RLI_18	PRACTICE REVIEW	4				4
RLI_19	PROBLEM APPROACH FRAMEWORK	4				4
RLI_20	MODEL-BASED, OFF-LINE, INVERSE RL	4				4
RLI_21	DEEP RL PRACTICE	4	70			74
RLI_22	SUMMARY RECAP	4		60		64
RLI_23	RESEARCH PAPERS, INNOVATION TRENDS (A)	4	40			44
RLI_24	RESEARCH PAPERS, INNOVATION TRENDS (B)	4				4
RLI_25	RL APPLICATIONS	4				4
RLI_26	MULTI-AGENT RL	4				4
RLI_27	TOP CASES & COMPANIES	4				4
RLI_28	TEST-EXAM	4			180	184
RLI_29	GROUP PROJECTS PRESENTATIONS (A)	4		60		64
RLI_30	GROUP PROJECTS PRESENTATIONS (B), WRAP-UP	4				4
		120	180	120	180	600
		20%	30%	20%	30%	

criteria	**percentage**	**Learning Objectives**	**Comments**
Final Exam	30 %		

Individual Assignments	30 %		
Group Assignment	10 %		
Group Presentation	10 %		
Participation	20 %		

RE-SIT / RE-TAKE POLICY

Each student has four chances to pass any given course distributed over two consecutive academic years: ordinary call exams and extraordinary call exams (re-sits) in June/July.

Students who do not comply with the 80% attendance rule during the semester will fail both calls for this Academic Year (ordinary and extraordinary) and have to re-take the course (i.e., re-enroll) in the next Academic Year.

Evaluation criteria:

- Students failing the course in the ordinary call (during the semester) will have to re-sit the exam in June / July (except those not complying with the attendance rule, who will not have that opportunity and must directly re-enroll in the course on the next Academic Year).
- The extraordinary call exams in June / July (re-sits) require your physical presence at the campus you are enrolled in (Segovia or Madrid). There is no possibility to change the date, location or format of any exam, under any circumstances. Dates and location of the June / July re-sit exams will be posted in advance. Please take this into consideration when planning your summer.
- The June / July re-sit exam will consist of a comprehensive exam. Your final grade for the course will depend on the performance in this exam only; continuous evaluation over the semester will not be taken into consideration. Students will have to achieve the minimum passing grade of 5 and can obtain a maximum grade of 8.0 (out of 10.0) – i.e., “notable” in the in the re-sit exam.
- Retakers: Students who failed the subject on a previous Academic Year and are now re-enrolled as re-takers in a course will be needed to check the syllabus of the assigned professor, as well as contact the professor individually, regarding the specific evaluation criteria for them as retakers in the course during that semester (ordinary call of that Academic Year). The maximum grade that may be obtained in the retake exam (3rd call) is 10.0.

After ordinary and extraordinary call exams are graded by the professor, you will have a possibility to attend a review session for that exam and course grade. Please be available to attend the session in order to clarify any concerns you might have regarding your exam. Your professor will inform you about the time and place of the review session. Any grade appeals require that the student attended the review session prior to appealing.

- Students failing more than 18 ECTS credits in the academic year after the June-July re-sits will be asked to leave the Program. Please, make sure to prepare yourself well for the exams in order to pass your failed subjects.
- In case you decide to skip the opportunity to re-sit for an exam during the June / July extraordinary call, you will need to enroll in that course again for the next Academic Year as a re-taker and pay the corresponding extra cost. As you know, students have a total of four allowed calls to pass a given subject or course, in order to remain in the program.

BIBLIOGRAPHY

Compulsory

- Richard S. Sutton and Andrew G. Barto. (2018). *Reinforcement Learning: An Introduction.* 2nd edition. AbeBooks. ISBN 9780262039246 (Digital)

This book covers the ground essential to understanding much of the work out there published on RL. Could be hard going for the students without a relatively solid mathematical background (especially the Bellman equations and monte-carlo sections) but it is worth it. It is a must-read for anyone doing graduate research in reinforcement learning in order to get to grips with these matters. It talks about most important topics to get you started in the required direction.

BEHAVIOR RULES

Please, check the University's Code of Conduct here. The Program Director may provide further indications.

ATTENDANCE POLICY

Please, check the University's Attendance Policy here. The Program Director may provide further indications.

ETHICAL POLICY

Please, check the University's Ethics Code here. The Program Director may provide further indications.