Master MVA: Reinforcement Learning **Lecture: -**

Organization of the Course

Lecturer: *Alessandro Lazaric* *http://researchers.lille.inria.fr/∼lazaric/Webpage/Teaching.html*

1 Objectives of the Course

- **Understand:** Mathematical foundation of the reinforcement learning paradigm
- **Use:** Overview of the state-of-the-art algorithms
- **Improve:** Discussion of the critical open questions in the field

2 Content

Lecture 1: Introduction to reinforcement learning

Lecture 2: Markov decision processes and dynamic programming

1. MDP, value functions, policies, Bellman operators
2. Value iteration
3. Policy iteration
4. Linear programming

Lecture 3: Reinforcement learning algorithms

1. Elements of stochastic approximation
2. TD(λ), Q-learning

Lecture 4: The exploration-exploitation dilemma: introduction to the multi-armed bandit framework

1. Stochastic bandit
2. Adversarial bandit
3. Extensions: from bandit to reinforcement learning

Lecture 5: Approximate dynamic programming

1. L_∞-norm analysis: Approximate value and policy iteration
2. Asymptotic analysis of LSTD, Bellman residual, LSPI, Fitted Q-iteration
3. Extensions to L_p-norm

Lecture 6: Finite-sample analysis of approximate dynamic programming

1. Elements of contraction of measures theory and statistical learning theory
2. Finite-sample analysis of LSTD and Bellman residual minimization

Lecture 7: Bonus topics

1. Transfer learning
2. The problem of state representation
3. Inverse reinforcement learning

3 Schedule

Date	*Topic*	*Classroom*
01/10	Intro/MDP	C013
08/10	Dynamic Programming	C013
15/10	RL Algorithms	C013
22/10	*TP* on DP and RL	(TBD)
29/10	Multi-arm Bandit (1)	C013
05/11	*TP* on Bandit	(TBD)
12/11	Multi-arm Bandit (2)	C013
19/11	*TP* on Bandit	(TBD)
26/11	Approximate DP	C013
03/12	Sample Complexity of ADP	C013
10/12	*TP* on ADP	(TBD)
17/12	Guest lectures + Internships	C013 (TBC)
14/01	Evaluation	C013 (TBC)

Lectures are from 11am to 1pm, TP should be from 11am to 1:15pm.

4 Evaluation

The evaluation will be based on a *review* of a set of papers on topics relevant to the course followed by an oral presentation. *Projects* on implementation or theoretical development are also available.

Towards the end of the course a set of possible *stages* will be presented, most of them with the possibility of continue with a PhD in the SequeL Team at INRIA-Lille.

5 Resources

Web page http://researchers.lille.inria.fr/∼lazaric/Webpage/Teaching.html

Main references

- *Apprentissage par renforcement*, former course by Rémi Munos.
  Available at `http://researchers.lille.inria.fr/~munos/master-mva/index.html`
- *Neuro Dynamic Programming*, Bertsekas et Tsitsiklis, 1996.

- *Introduction to Reinforcement Learning*, Sutton and Barto, 1998.
  Available at `http://webdocs.cs.ualberta.ca/ sutton/book/the-book.html`
- *Markov Decision Problems*, Puterman, 1994.
- *Processus Décisionnels de Markov en Intelligence Artificielle*, Sigaud et Buffet (eds.), 2004.
  Available at `http://researchers.lille.inria.fr/~munos/papers/files/bouquinPDMIA.pdf`
- *Algorithms for Reinforcement Learning*, Szepesvári, 2009.
  Available at: `http://www.ualberta.ca/ szepesva/RLBook.html`

Other courses

- *Reinforcement Learning in Practice* by Rich Sutton.
  `http://incompleteideas.net/rlai.cs.ualberta.ca/rlip/rlip.html`
- *Reinforcement Learning* by Benjamin Van Roy.
  `http://www.stanford.edu/class/msande338/`
- *Learning in Complex Systems* by Nahum Shimkin.
  `http://webee.technion.ac.il/people/shimkin/LCS10/LCS10index.html`

6 Contacts

Lecturer
Alessandro Lazaric
SequeL group, INRIA Lille - Nord Europe
alessandro.lazaric@inria.fr
`http://researchers.lille.inria.fr/~lazaric`

Assistant
Emilie Kaufmann
Equipe STA, Telecom ParisTech
emilie.kaufmann@telecom-paristech.fr
`http://perso.telecom-paristech.fr/~kaufmann/`

References