THE 14th EUROPEAN WORKSHOP ON REINFORCEMENT LEARNING

École nationale supérieure d’arts et métiers (ENSAM)
8 Boulevard Louis XIV, 59800 Lille, France
October 1-3, 2018
2018 ORGANIZATION SUPPORT

EWRL would like to especially thank INRIA Lille - Nord Europe for the help in the organization of the EWRL 2018, investing time and resources.

2018 SPONSORS

EWRL gratefully acknowledge the generosity of those individuals and organizations who have provided financial support for the 14th European Workshop on Reinforcement Learning. Their financial support enables us to sponsor student travel and participation, general support to host the conference, and volunteers who assist during EWRL.

GOLD SPONSORS

DeepMind, criteo AI Lab, facebook, Microsoft

SILVER SPONSORS

amazon, Google
ORGANIZING COMMITTEE

GENERAL CHAIRS

Matteo Pirotta
Inria Lille

Ronan Fruit
Inria Lille

Florian Strub
Inria Lille

Mathieu Seurin
Inria Lille

ORGANIZING MEMBERS

Philippe Preux
Inria Lille

Jérémie Mary
Criteo Research

Olivier Pietquin
Google Brain

Program Committee

EWRL is thankful to all the people serving in the program committee for the time spent for reviewing the numerous submissions.

Marc Abeille  Anders Jonsson  Julien Perolat
Riad Akour  Emilie Kaufmann  Pierre Perrault
Oleg Arenz  Johannes Kirschner  Olivier Pietquin
Lilian Besson  Akshay Krishnamurthy  Joelle Pineau
Roberto Calandra  Alessandro Lazaric  Bilal Piot
Daniele Calandriello  Odalric-Abrym Maillard  Matteo Pirotta
Christian Daniel  Timothy Mann  Philippe Preux
Christos Dimitrakakis  Jérémie Mary  Marcello Restelli
Layla Elasri  Remi Munos  Mathieu Seurin
Ronan Fruit  Gergely Neu  Florian Strub
Victor Gabillon  Ann Nowe  Sadegh Talebi
Pratik Gajane  Laurent Orseau  Herke van Hoof
Matthieu Geist  Ronald Ortner  Claire Vernade
Mohammad Gheslaghi Azar  Ian Osband
Anna Harutyunyan  Simone Parisi
Maximilian Hütttenrauch  Vianney Perchet
## PROGRAM OVERVIEW

### Monday October 1 (tutorials)

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00-10:00</td>
<td>Check-in and welcoming coffee</td>
</tr>
<tr>
<td>9:45-10:00</td>
<td>Opening Remarks</td>
</tr>
<tr>
<td>10:00-12:00</td>
<td><strong>Csaba Szepesvári and Tor Lattimore</strong>&lt;br&gt;Advanced Topics In Exploration: The Role Of Randomization For Exploration In Bandits And Reinforcement Learning</td>
</tr>
<tr>
<td>12:00-14:00</td>
<td>Lunch break (on your own)</td>
</tr>
<tr>
<td>14:00-16:00</td>
<td><strong>Matteo Hessel</strong>&lt;br&gt;Deep Reinforcement Learning</td>
</tr>
<tr>
<td>16:00-16:30</td>
<td>Coffee break</td>
</tr>
<tr>
<td>16:30-18:30</td>
<td><strong>Andreas Krause and Felix Berkenkamp</strong>&lt;br&gt;Towards Safe Reinforcement Learning</td>
</tr>
</tbody>
</table>

### Tuesday October 2

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00-9:00</td>
<td>Check-in</td>
</tr>
<tr>
<td>9:00-9:45</td>
<td>Richard Sutton&lt;br&gt;Bootstrapping to Understandings</td>
</tr>
<tr>
<td>9:45-10:05</td>
<td>Contributed talk&lt;br&gt;<em>Anderson Acceleration for Reinforcement Learning</em></td>
</tr>
<tr>
<td>10:05-10:55</td>
<td>Poster session 1</td>
</tr>
<tr>
<td>10:55-11:40</td>
<td>Tze Leung Lai&lt;br&gt;<em>Multi-armed Bandits with Covariates: Theory and Applications</em></td>
</tr>
<tr>
<td>11:40-12:00</td>
<td>Contributed talk</td>
</tr>
<tr>
<td></td>
<td><em>When Gaussian Processes Meet Combinatorial Bandits: GCB</em></td>
</tr>
<tr>
<td>12:00-14:00</td>
<td>Lunch break (on your own)</td>
</tr>
</tbody>
</table>

### Wednesday October 3

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00-9:20</td>
<td>Contributed talk&lt;br&gt;Learning good policies from suboptimal demonstrations</td>
</tr>
<tr>
<td>9:20-10:05</td>
<td>Joelle Pineau&lt;br&gt;<em>Reproducibility, Reusability, and Robustness in Deep Reinforcement Learning</em></td>
</tr>
<tr>
<td>10:05-10:55</td>
<td>Poster session 3</td>
</tr>
<tr>
<td>10:55-11:40</td>
<td>Rémi Munos&lt;br&gt;<em>Off-Policy Deep Reinforcement Learning</em></td>
</tr>
<tr>
<td>11:40-12:00</td>
<td>Contributed talk&lt;br&gt;<em>Directed Policy Gradient for Safe Reinforcement Learning with Human Advice</em></td>
</tr>
</tbody>
</table>

### Tuesday October 2

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:00-14:20</td>
<td>Contributed talk&lt;br&gt;<em>Fighting Boredom in Recommender Systems with Linear Reinforcement Learning</em></td>
</tr>
<tr>
<td>14:20-15:05</td>
<td>Nicolò Cesa-Bianchi&lt;br&gt;<em>Trade-Offs in Adversarial Bandits</em></td>
</tr>
<tr>
<td>15:05-15:25</td>
<td>Contributed talk&lt;br&gt;<em>A0C: Alpha Zero in Continuous Action Space</em></td>
</tr>
<tr>
<td>15:25-16:10</td>
<td>Martin Riedmiller&lt;br&gt;<em>Machines that Learn from Scratch</em></td>
</tr>
<tr>
<td>16:10-17:10</td>
<td>Poster session 2</td>
</tr>
<tr>
<td>17:10-17:55</td>
<td>Gergely Neu&lt;br&gt;<em>A Unified View Of Entropy-Regularized Markov Decision Processes</em></td>
</tr>
<tr>
<td>17:55-18:15</td>
<td>Contributed talk&lt;br&gt;<em>Constraint-Space Projection Direct Policy Search</em></td>
</tr>
<tr>
<td>19:30-</td>
<td>Social Dinner</td>
</tr>
</tbody>
</table>

### Wednesday October 3

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:00-14:20</td>
<td>Contributed talk&lt;br&gt;Towards learning to best respond when losing control</td>
</tr>
<tr>
<td>14:20-15:05</td>
<td>Karl Tuyls&lt;br&gt;<em>New Horizons in Multi-Agent Learning</em></td>
</tr>
<tr>
<td>15:05-15:25</td>
<td>Contributed talk&lt;br&gt;<em>Transferring Value Functions via Variational Methods</em></td>
</tr>
<tr>
<td>15:25-16:10</td>
<td>Katja Hofmann&lt;br&gt;<em>Directions and Challenges in Multi-Task Reinforcement Learning</em></td>
</tr>
<tr>
<td>16:10-17:10</td>
<td>Poster session 4</td>
</tr>
<tr>
<td>17:10-17:55</td>
<td>Audrey Durand&lt;br&gt;<em>Real-world bandit applications: Bridging the gap between theory and practice</em></td>
</tr>
<tr>
<td>17:55-18:15</td>
<td>Contributed talk&lt;br&gt;<em>Counting to Explore and Generalize in Text-based Games</em></td>
</tr>
<tr>
<td>18:15-18:25</td>
<td>Closing Remarks</td>
</tr>
</tbody>
</table>
9:00 – 10:00 Check-in and welcoming coffee

9:50 – 10:00 Opening remarks

10:00 – 12:00 Tutorial 1

**ADVANCED TOPICS IN EXPLORATION: THE ROLE OF RANDOMIZATION FOR EXPLORATION IN BANDITS AND REINFORCEMENT LEARNING**
Csaba Szepesvári and Tor Lattimore

*Abstract.* Randomization appears as an ingredient in many algorithms for bandits and reinforcement learning. Sometimes randomization is unavoidable, like in adversarial bandit models. In other cases its role is to reduce the burden of computation, for example in combinatorial semibandits and stochastic linear bandits. In many cases designing a natural randomized algorithm based on Thompson sampling is more straightforward than the confidence-set construction usually required by other approaches. And these algorithms can work very well in practice. In this tutorial we plan to discuss the pros and cons of randomization for exploration in bandit problems and beyond.

12:00 – 14:00 Lunch break (on your own)

14:00 – 16:00 Tutorial 2

**DEEP REINFORCEMENT LEARNING**
Matteo Hessel

*Abstract.* The combination of deep learning with reinforcement learning (RL) has been the driving force behind many recent successes in the application of RL to increasingly complex domains. Extending reinforcement learning ideas and principles to work in combination with non linear function approximation can be non trivial, and introduces its own set of challenges. In this tutorial we will highlight some of these issues, describe how they are overcome in practice, and give an overview of the current state of the art in Deep RL methods.

16:00 – 16:30 Coffee break

16:30 – 18:30 Tutorial 3

**TOWARDS SAFE REINFORCEMENT LEARNING**
Andreas Krause and Felix Berkenkamp

*Abstract.* Reinforcement learning has seen stunning empirical breakthroughs. At its heart is the challenge of trading exploration – collecting data for learning better models – and exploitation – using the estimate to make decisions. In many applications, exploration is a potentially dangerous proposition, as it requires experimenting with actions that have unknown consequences. Hence, most prior work has confined exploration to simulated environments. In this tutorial, we will review work towards rigorously
reasoning about safety in reinforcement learning. Starting from a discussion of what safety in reinforcement learning means, we review several model-free and model-based approaches that consider additional safety constraints during learning. We conclude the tutorial with a summary of recent research in our group.
9:00 – 9:45 Invited talk

**BOOTSTRAPPING TO UNDERSTANDING**
Richard Sutton

9:45 – 10:05 Contributed talk

**ANDERSON ACCELERATION FOR REINFORCEMENT LEARNING**
Matthieu Geist and Bruno Scherrer.

10:05 – 10:55 Poster session 1 (with Coffee break)

10:55 – 11:40 Invited talk

**MULTI-ARMED BANDITS WITH COVARIATES: THEORY AND APPLICATIONS**
Tze Leung Lai

Abstract. In the past five years, multi-armed bandits with covariates, also called “contextual bandits” in machine learning, have become an active area of research in data science, stochastic optimization, and statistical modeling because of their applications to the development of personalized strategies in translational medicine and in recommender systems for web-based marketing and electronic business. After a brief review of classical (context-free) bandit theory, we describe a corresponding theory, covering both parametric and nonparametric approaches, for contextual bandits and illustrate their applications to personalized strategies.

11:40 – 12:00 Contributed talk

**WHEN GAUSSIAN PROCESSES MEET COMBINATORIAL BANDITS: GCB**
Guglielmo Maria Accabi, Alessandro Nuara, Francesco Trovò, Nicola Gatti and Marcello Restelli

12:00 – 14:00 Lunch break (on your own)

14:00 – 14:20 Contributed talk

**FIGHTING BOREDOM IN RECOMMENDER SYSTEMS WITH LINEAR REINFORCEMENT LEARNING**
Romain Warlop, Alessandro Lazaric and Jérémie Mary

14:20 – 15:05 Invited talk
TRADE-OFFS IN ADVERSARIAL BANDITS
Nicolò Cesa-Bianchi

Abstract. Adversarial bandits are a crisp mathematical framework for investigating sequential decision problems in the presence of partial and delayed feedbacks. In many applications, the algorithm has some control over the amount of feedback obtained after each round. However, these more informative feedbacks may come at a cost, forcing the algorithm to strike a balance. In this talk we focus on two examples of such trade-offs: one where the extra feedback is scattered over time, and one where some actions are more informative than others. In both cases, we show how the adversarial bandit setting allows dealing with these trade-offs in a rigorous and elegant way.

15:05 – 15:25 Contributed talk

A0C: ALPHA ZERO IN CONTINUOUS ACTION SPACE
Thomas Moerland, Joost Broekens, Aske Plaat and Catholijn Jonker

15:25 – 16:10 Invited talk

MACHINES THAT LEARN FROM SCRATCH
Martin Riedmiller

Abstract. Being able to autonomously learn from scratch is a key ability of intelligent systems - and has always been a central focus of our research. A particular challenge in real world robotic scenarios are methods that are at the same time highly data-efficient and robust, since data-collection on real robots is time intensive and often expensive. I will discuss two main research areas that are crucial for progress towards this goal: improved off-policy learning and better exploration. I will give examples of simulated and real robots that can learn increasingly complex tasks from scratch.

16:10 – 17:10 Poster session 2 (with Coffee break ☕️)

17:10 – 17:55 Invited talk

A UNIFIED VIEW OF ENTROPY-REGULARIZED MARKOV DECISION PROCESSES
Gergely Neu

Abstract. Entropy regularization, while a standard technique in the online learning toolbox, has only been recently discovered by the reinforcement learning community: In recent years, numerous new reinforcement learning algorithms have been derived using this principle, largely independently of each other. So far, a general framework for these algorithms has remained elusive. In this work, we propose such a general framework for entropy-regularized average-reward reinforcement learning in Markov decision processes (MDPs). Our approach is based on extending the linear-programming formulation of policy optimization in MDPs to accommodate convex regularization functions. Our key result is showing that using the conditional entropy of the joint state-action distributions as regularization yields a dual optimization problem closely resembling the Bellman optimality equations. This result enables us to formalize a number of
state-of-the-art entropy-regularized reinforcement learning algorithms as approximate variants of Mirror Descent or Dual Averaging, and thus to argue about the convergence properties of these methods. In particular, we show that the exact version of the TRPO algorithm of Schulman et al. (2015) actually converges to the optimal policy, while the entropy-regularized policy gradient methods of Mnih et al. (2016) may fail to converge to a fixed point.

17:55 – 18:15 Contributed talk

**CONSTRAINT-SPACE PROJECTION DIRECT POLICY SEARCH**
Riad Akrour, Jan Peters and Gerhard Neumann

**Posters of Day 1**
- Romain Warlop, Alessandro Lazaric and Jérémie Mary. Fighting Boredom in Recommender Systems with Linear Reinforcement Learning
- Riad Akrour, Jan Peters and Gerhard Neumann. Constraint-Space Projection Direct Policy Search
- Audrey Durand, Andrei Lupu and Doina Precup. Leveraging Observational Learning for Exploration in Bandits
- Thomas Moerland, Joost Broekens, Aske Plaat and Catholijn Jonker. A0C: Alpha Zero in Continuous Action Space
- Guglielmo Maria Accabi, Alessandro Nuara, Francesco Trovò, Nicola Gatti and Marcello Restelli. When Gaussian Processes Meet Combinatorial Bandits: GCB
- Matthieu Geist and Bruno Scherrer. Anderson Acceleration for Reinforcement Learning
- Tom Zahavy, Avinatan Hasidim, Haim Kaplan and Yishay Mansour. Planning in Hierarchical Reinforcement Learning: Guarantees for Using Local Policies
- Takuya Okano and Itsuki Noda. A Parameter Investigation of the ε-greedy Exploration Ratio Adaptation Method in Multi-agent Reinforcement Learning
- Boris Belousov and Jan Peters. Mean squared advantage minimization as a consequence of entropic policy improvement regularization
- Lukas Kemmer, Henrik von Kleist, Diego María De Grimaudet De Rochebouët, Nikolaos Tziortzios and Jesse Read. Reinforcement learning for supply chain optimization
- Nikolaos Tziortzios, Christos Dimitrakakis and Michalis Vazirgiannis. Randomised Bayesian Least-Squares Policy Iteration
- Peter Henderson, Joshua Romoff and Joelle Pineau. Where Did My Optimum Go?: An Empirical Analysis of Gradient Descent Optimization in Policy Gradient Methods
- Alberto Maria Metelli, Mirco Mutti and Marcello Restelli. Configurable Markov Decision Processes
- Mohammed Abdullah, Moez Draief and Aldo Pacchiano. Reinforcement Learning with Wasserstein Distance Regularisation, with Applications to Multipolicy Learning
- Reda Alami. Thompson Sampling for the non-Stationary Corrupt Multi-Armed Bandit
- Romain Laroche and Paul Trichelair. Safe Policy Improvement with Baseline Bootstrapping
- Peter Auer, Pratik Gajane and Ronald Ortner. Adaptively Tracking the Best Arm with an Unknown Number of Distribution Changes
- Ioannis Boukas, Damien Ernst, Anthony Papavasiliou and Bertrand Cornelusse. Intra-day Bidding Strategies for Storage Devices Using Deep Reinforcement Learning
- Ian Kash and Katja Hofmann. Combining No-regret and Q-learning
9:00 – 9:20 Contributed talk

LEARNING GOOD POLICIES FROM SUBOPTIMAL DEMONSTRATIONS
Yuxiang Li, Katja Hofmann and Ian Kash

9:20 – 10:05 Invited talk

REPRODUCIBILITY, REUSABILITY, AND ROBUSTNESS IN DEEP REINFORCEMENT LEARNING
Joelle Pineau

Abstract. In recent years, significant progress has been made in solving challenging problems across various domains using deep reinforcement learning. However reproducing results for state-of-the-art deep RL methods is seldom straightforward. High variance of some methods can make learning particularly difficult when environments or rewards are strongly stochastic. Furthermore, results can be brittle to even minor perturbations in the domain or experimental procedure. In this talk, I will discuss challenges that arise in experimental techniques and reporting procedures in deep RL, and will suggest methods and guidelines to make future results more reproducible, reusable and robust.

10:05 – 10:55 Poster session 3 (with Coffee break 🍵)

10:55 – 11:40 Invited talk

OFF-POLICY DEEP REINFORCEMENT LEARNING
Rémi Munos

Abstract. I will motivate the use of off-policy multi-steps learning in deep Reinforcement Learning agents, then mention theoretical guarantees in tabular environments, and finally report experiments in DMLab30.

11:40 – 12:00 Contributed talk

DIRECTED POLICY GRADIENT FOR SAFE REINFORCEMENT LEARNING WITH HUMAN ADVICE
Hélène Plisnier, Denis Steckelmacher, Tim Brys, Diederik Roijers and Ann Nowé

12:00 – 14:00 Lunch break (on your own) 🍽️

14:00 – 14:20 Contributed talk

TOWARDS LEARNING TO BEST RESPOND WHEN LOSING CONTROL
Richard Klima, Daan Bloembergen, Michael Kaisers and Karl Tuyls
NEW HORIZONS IN MULTI-AGENT LEARNING
Karl Tuyls

Abstract. Recent advances in deep learning and game theory have re-energized interest in multi-agent learning, both from an algorithmic and analytical point of view. Multi-agent learning is largely defined at the boundaries of reinforcement learning, game theory and multi-agent systems. Many real-world scenarios can be modelled as multi-agent systems, in which multiple autonomous decision makers interact in a single environment. The complex and dynamic nature of such interactions prevents hand-crafting solutions for all possible scenarios, hence learning becomes crucial. In this talk I will first hark back to the early days of multi-agent learning and discuss some of its foundational roots and show how the field has evolved into its current state of affairs. I will then discuss some of the current trends and advances in this historical context and illustrate some of these recent (algorithmic) results using ideas from evolutionary and empirical game theory, and discuss some of the major future challenges.

TRANSFERRING VALUE FUNCTIONS VIA VARIATIONAL METHODS
Andrea Tirinzoni, Rafael Rodriguez and Marcello Restelli

DIRECTIONS AND CHALLENGES IN MULTI-TASK REINFORCEMENT LEARNING
Katja Hofmann

Abstract. Multi-task reinforcement learning (RL) aims to develop approaches that learn to perform well across a range of related tasks, instead of specializing to a single task. This has high potential for real world applications, where sharing data across tasks could dramatically improve data efficiency to make RL approaches economically viable. In this talk I present two contributions. First, I present a novel Latent Variable Gaussian Process approach and demonstrate its ability to learn and leverage task structure for few-shot generalization in control tasks. Second, I introduce an RL challenge, MARLO, that we are currently organizing to encourage more research in multi-task (and multi-agent) RL. In the talk, I will focus on the challenges associated with defining tasks for evaluating generalization behaviour in RL, as a starting point for discussing open challenges.

REAL-WORLD BANDIT APPLICATIONS: BRIDGING THE GAP BETWEEN THEORY AND PRACTICE
Audrey Durand
Abstract. The multi-armed bandit setting is a well-known environment of choice for studying the exploration-exploitation trade-off. In order to provide powerful and robust algorithms, many efforts have been dedicated to obtain theoretical guarantees on the achievable performance under various bandit settings. However, although the original bandit setting takes roots in real-world applications, many assumptions required to obtain these guarantees do not hold in practice. One could then question the applicability of bandit strategies. Using first an application of contextual bandits to adaptive clinical trials, we will highlight the potential of bandit approaches under failure of assumptions. We will then provide an example of how to bridge the gap between theory and practice using an application of kernelized bandits to online tuning of imaging parameters for microscopy devices.

17:55 – 18:15 Contributed talk

COUNTING TO EXPLORE AND GENERALIZE IN TEXT-BASED GAMES
Xingdi Yuan, Marc-Alexandre Côté, Alessandro Sordoni, Matthew Hausknecht and Adam Trischler

Posters of Day 2

- Yuxiang Li, Katja Hofmann and Ian Kash. Learning good policies from suboptimal demonstrations
- Andrea Tirinzoni, Rafael Rodriguez and Marcello Restelli. Transferring Value Functions via Variational Methods
- Matteo Papini, Andrea Battistello and Marcello Restelli. Safely Exploring Policy Gradient
- Sudeep Raja Putta. Exponential Weights on the Hypercube in Polynomial Time
- Hélène Plinsier, Denis Steckelmacher, Tim Brys, Diederik Roijers and Ann Nowé. Directed Policy Gradient for Safe Reinforcement Learning with Human Advice
- Yao Liu and Emma Brunskill. When Simple Exploration is Sample Efficient: Identifying Sufficient Conditions for Random Exploration to Yield PAC RL Algorithms
- Xuedong Shang, Emilie Kaufmann and Michal Valko. Adaptive black-box optimization got easier: HCT needs only local smoothness
- Xingdi Yuan, Marc-Alexandre Côté, Alessandro Sordoni, Matthew Hausknecht and Adam Trischler. Counting to Explore and Generalize in Text-based Games
- Richard Klima, Daan Bloembergen, Michael Kaisers and Karl Tuyls. Towards learning to best respond when losing control
- M. Sadegh Talebi and Odalric-Ambrym Maillard. KL-UCRL Revisited: Variance-Aware Regret Bounds
- Eli Friedman and Fred Fontaine. Generalizing Across Multi-Objective Reward Functions in Deep Reinforcement Learning
- Zhaohan Guo and Emma Brunskill. Sample Efficient Learning with Feature Selection for Factored MDPs
- Nicolas Carrara, Olivier Pietquin, Romain Laroché, Tanguy Urvoys and Jean-Léon Bouraoui. A Fitted-Q Algorithm for Budgeted MDPs
- Thomas Moerland, Joost Broekens and Catholijn Jonker. The Potential of the Return
Distribution for Exploration in Reinforcement Learning
■ Ciara Pike-Burke and Steffen Grunewalder. Recovering Bandits
■ Howard Huang. An Empirical Study of Least-Squares Algorithms in Reinforcement Learning
■ Denis Steckelmacher, Hélène Plisnier, Diederik M. Roijers and Ann Nowé. Stable, Practical and On-line Bootstrapped Conservative Policy Iteration
■ Martin Gottwald, Mingpan Guo and Hao Shen. Neural Value Function Approximation in Continuous State Reinforcement Learning Problems
■ Kimia Nadjahi, Romain Laroche and Rémi Tachet des Combes. Soft Safe Policy Improvement with Baseline Bootstrapping
■ Simone Parisi, Voot Tangkaratt, Jan Peters and Mohammad Khan. TD-Regularized Actor-Critic Methods
■ Tom Zahavy, Matan Harouch, Nadav Merlis, Daniel J. Mankowitz and Shie Mannor. Learn What Not to Learn: Action Elimination with Deep Reinforcement Learning
■ Ronnie Stafford and John Shawe-Taylor. ACCME: Actively Compressed Conditional Mean Embeddings for Model-Based Reinforcement Learning
GALA DINNER

Couvent des Minimes
17 Quai du Wault, 59800 Lille, France

starting at 19:30
How to reach the restaurant:

- Bus 18 (approx. 20/30 minutes)
  Walk to the bus stop “Liberte” and take the bus toward “Anatole France”
  Get out at stop “Nationale”
  Walk to the restaurant

- Bus L1 (approx. 20/30 minutes)
  Walk to the bus stop “Liberte” and take the bus toward “Agrippin”
  Get out at stop “Champ De Mars”
  Walk to the restaurant
• Metro (approx. 20 minutes)
  Walk to the metro station “Grand Palais”
  Take Metro M2 (RED) toward “C.H. Dron”
  Get out at “Gare Lille Flandres” (2 stops)
  Take the Metro M1 (YELLOW) toward “CHU-Eurasante”
  Get out at “Rihour” (1 stop)
  Walk to the restaurant (approx. 10 minutes)
• Walk to the restaurant (approx. 20/30 minutes)