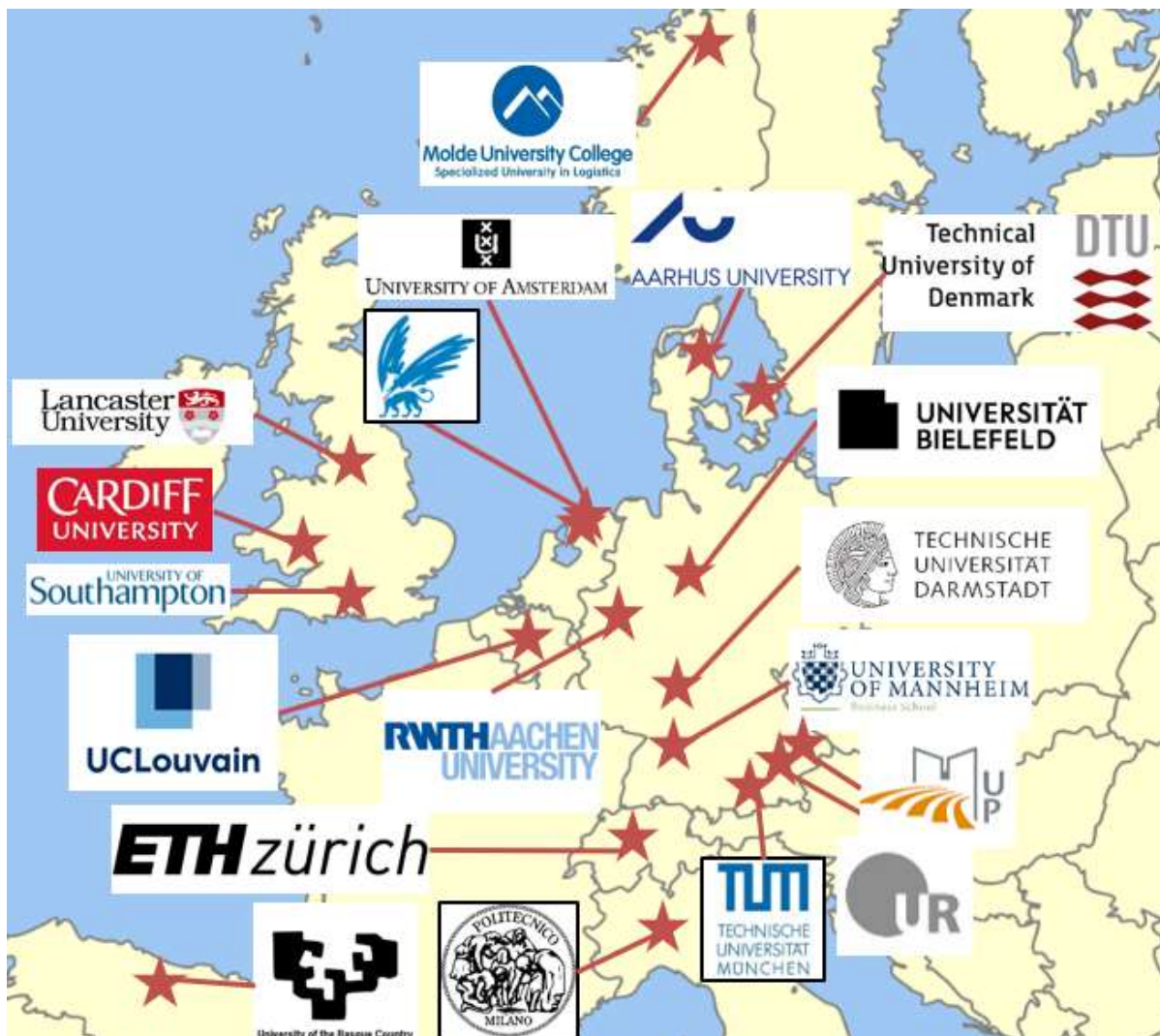


Final report EURO PhD School

“Reinforcement Learning Applied to Operations Research”

This EURO PhD school made students in stochastic modelling familiar with several Reinforcement Learning (RL) techniques. They applied these concepts to solve Operations Research problems through a hackathon-like workshop. 23 participants from 9 European countries joined the PhD school from 17th to 24th of July in Castle Gimborn, Germany.



Universities of the participants

The PhD school is organized by *Peter Jacko* (Lancaster University, UK), *Odysseas Kanavetas* (Leiden University, The Netherlands), *Ger Koole* (VU University Amsterdam, The Netherlands), and *Raik Stolletz* (University of Mannheim, Germany). The school is funded by the Association of European Operational Research Societies (EURO), the EURO working group on Stochastic Modelling (StochMod), and the German Operations Research Society (GOR e.V.).



Sponsor: EURO



Sponsor: EWG STOCHMOD



Sponsor: GOR

By learning the Reinforcement Learning techniques and applying them, participants obtained hands-on knowledge of Reinforcement Learning in the context of stochastic modeling problems. The school contained online and in-class Lectures from specialists, and time to work in groups on one of the application areas: inventory management, queueing systems, and pricing. After a brief introduction to Markovian Decision Processes, the participants had to find appropriate problems to solve with Temporal Difference Learning, Deep Reinforcement Learning, and Multi-Armed Bandits. After learning about different algorithms in tutorials, students implemented these and compared them with each other. Analyzed problems and results are presented and discussed with experts.

The tutorials were given by *Shipra Agrawal* (Columbia University, USA), *Pierre-Luc Bacon* (Université de Montréal, Canada), *Gerhard Neumann* (Karlsruhe Institute of Technology, Germany). *Vincent-Francois Lavet* (VU Amsterdam, The Netherlands), *Joren Gijsbrechts* (Universidade Católica Portuguesa, Portugal), *Emilie Kaufmann* (Université de Lille, France) and *Tor Lattimore* (DeepMind, UK), see the attached program or <http://stochmod.eu/EPS/Program/> for more details.

During the whole week the participants implemented the learned techniques, also extended the sampling codes that our experts prepared for them, and worked on their presentations. The final day there was an evaluation meeting where the students presented what they studied and what they learnt during the school.

EURO PhD School “Reinforcement Learning Applied to Operations Research”

When: 17-24 July 2022 (<http://stochmod.eu/EPS>)

Where: IBZ Schloss Gimborn, Marienheide, Germany (<https://www.ibz-gimborn.de/>)

Organized by EURO Working Group on Stochastic Modeling (<http://stochmod.eu>) and supported by EURO (<https://www.euro-online.org/>)

Overview

Goal: Make PhD students in stochastic modeling familiar with Reinforcement Learning techniques by DIY in a hackathon-like workshop.

Background: Markov decision processes are a familiar technique in stochastic modeling. Variants of many stochastic models of decision-making problems lend themselves for an analysis using Reinforcement Learning (In the operations research and control literature, Reinforcement Learning is often called Approximate Dynamic Programming). Examples are inventory and controlled queueing models with unknown arrivals distribution, or high-dimensional revenue management problems with unknown demand function. By learning the Reinforcement Learning techniques and applying them directly to simple problems we hope that participants will obtain hands-on knowledge of Reinforcement Learning in the context of stochastic modeling problems.

Attendees: PhD students, mainly from European countries (members of the EURO), are expected to have familiarized themselves with the basics of (finite) Markov decision processes, stochastic dynamic programming and Monte Carlo simulation before the start of the workshop.

Format: Tutorials and lectures by specialists, a mixture of in-person and online, with time to work in small groups on various operations research problems. Specialists will actively support the groups during implementation sessions.

Speakers and Organizers

<u>Shipra Agrawal</u> Columbia University USA	<u>Pierre-Luc Bacon</u> Université de Montréal & Mila Canada	<u>Vincent François-Lavet</u> VU Amsterdam Netherlands
<u>Joren Gijsbrechts</u> Universidade Católica Portuguesa Portugal	<u>Peter Jacko</u> Lancaster University & Berry Consultants UK	<u>Odysseas Kanavetas</u> Leiden University Netherlands
<u>Emilie Kaufmann</u> CNRS & Université de Lille France	<u>Ger Koole</u> VU Amsterdam Netherlands	<u>Tor Lattimore</u> DeepMind UK
<u>Gerhard Neumann</u> Karlsruhe Institute of Technology Germany	<u>Katharina Senz</u> University of Mannheim Germany	<u>Raik Stolletz</u> University of Mannheim Germany

Schedule

Sunday 17 July

16:15 Shuttle from Cologne main train station
17:15-18:15 Accommodation check-in
18:30-19:30: Dinner
19:30 Program starting
21:30 Program ending

Monday 18 July - Saturday 23 July

07:30-08:30: Breakfast
09:00 Program starting
10:30-11:00: Coffee & biscuits
13:00-14:00: Lunch
15:30-16:00 or 16:00-16:30: Coffee & pastries
18:30-19:30: Dinner
21:30 Program ending

Sunday 24 July

7:30-08:30 Breakfast
9:30-10:00 Accommodation check out
10:00 Program starting
11:30 Program ending
11:30-12:30: Lunch
12:45 Shuttle to Cologne central train station

Social Events

Monday 18 July

11:00-13:00 Walk

Thursday 21 July

16:30-21:30 Walk & dinner at Wiesengrund (www.gasthaus-wiesengrund.de/18.html) & bus

Saturday 23 July

19:30-21:30 Free time

Program

(Asterisk * means “online” attendance):

Sunday 17 July

- 19:30-20:30: Intro to RL and summer school program (Ger, Raik, Odysseas)
- 20:30-21:30: Discussion of OR problems tractable by RL (Ger, Odysseas)

Monday 18 July

- 09:00-10:30: 3 minutes presentation by each participant & Making groups (Ger, Odysseas)
- 11:00-13:00: Social event (Ger, Raik, Odysseas)
- 14:00-15:30: MDPs & OR/RL Problems - Tutorial (Shipra*)
- 16:00-18:00: MDPs & OR/RL Problems - Implementation in groups (Ger, Odysseas)
- 19:30-21:00: Value-based Methods I: Temporal Difference - Tutorial (Pierre-Luc*)

Tuesday 19 July

- 09:00-10:30: Value-based Methods I: Temporal Difference - Implementation in groups (Ger, Odysseas)
- 11:00-12:30: Value-based Methods II: Policy Gradient & Actor-Critic - Tutorial (Gerhard*)
- 14:00-16:00: Value-based Methods II: Policy Gradient & Actor-Critic - Implementation in groups (Ger, Odysseas)
- 16:30-18:00: Value-based Methods - Preparation of presentations
- 19:30-21:30: Value-based Methods - Group presentations (Ger, Odysseas)

Wednesday 20 July

- 09:00-10:30: Deep Reinforcement Learning I - Tutorial (Vincent)
- 11:00-12:30: Deep Reinforcement Learning I - Discussion of OR problems & implementation in groups (Joren*, Vincent, Odysseas)
- 14:00-15:30: Deep Reinforcement Learning II - Tutorial (Vincent)
- 16:00-18:00: Deep Reinforcement Learning II - Implementation in groups (Vincent, Odysseas)
- 19:30-21:30: Further group work

Thursday 21 July

- 09:00-10:30: Deep Reinforcement Learning III - Tutorial (Vincent)
- 11:00-12:00: Deep Reinforcement Learning III - Implementation in groups (Vincent, Odysseas)
- 12:00-12:30: Deep Reinforcement Learning - Preparation of presentations
- 14:00-16:00: Deep Reinforcement Learning - Group presentations (Vincent, Odysseas)
- 16:30-21:30: Social event & Dinner (Vincent, Odysseas, Emilie)

Friday 22 July

- 9:00-10:30: Multi-armed Bandits I: Stochastic - Tutorial (Emilie)
- 11:00-12:30: Multi-armed Bandits I: Stochastic - Discussion of OR problems & implementation in groups (Emilie, Odysseas)
- 14:00-15:30: Multi-armed Bandits II: Bayesian - Tutorial (Emilie)
- 16:00-18:00: Multi-armed Bandits II: Bayesian - Implementation in groups (Emilie, Odysseas)
- 19:30-21:30: Further group work

Saturday 23 July

- 9:00-10:30: Multi-armed Bandits III: Contextual - Tutorial (Tor)
- 11:00-12:30: Multi-armed Bandits III: Contextual - Implementation in groups (Tor, Odysseas)
- 14:00-15:00: Multi-armed Bandits - Advanced lecture (Tor) - not meant for implementation
- 15:00-16:00: Multi-armed Bandits - Preparation of presentations
- 16:30-18:30: Multi-armed Bandits - Group presentations (Tor, Odysseas)

Sunday 24 July

- 10:00-11:30: Evaluation & potential next steps (Tor, Odysseas)