SCIENCES & BIOENGINEERING SCIENCES

The Research Group
Artificial Intelligence Lab

has the honor to invite you to the public defence of the PhD thesis of

Eugenio Tiziano Gabriele Bargiacchi

to obtain the degree of Doctor of Sciences

Title of the PhD thesis:

Controlling Large Scale Multi-Agent Environments with Model-Based Reinforcement Learning

Promotors:
Prof. dr. Ann Nowé (VUB)
Dr. Diederik M. Roijers (VUB)

The defence will take place on

Monday, January 15, 2024 at 10.30h in Auditorium E.0.10

Members of the jury

Prof. dr. Ann Dooms (VUB, chair)
Prof. dr. Bart Bogaerts (VUB, secretary)
Prof. dr. Bas Ketsman (VUB)
Prof. dr. Jilles Dibangoye (INSA Lyon, France)
Prof. dr. Matthijs Spaan (TU Delft, The Netherlands)

Curriculum vitae

Eugenio Bargiacchi obtained a B.Sc. in Computer Science at the Università Statale di Milano in Italy (2011), focusing on information technology fundamentals as well as statistics. He followed this with a M.Sc. in Artificial Intelligence at the Vrije Universiteit Amsterdam in The Netherlands (2016). He started his PhD in Artificial Intelligence at the AI Lab at the Vrije Universiteit Brussel in 2017, and obtained an FWO scholarship to work on his PhD in 2019. He now works in the private sector.

Abstract of the PhD research

Multi-agent reinforcement learning (RL) offers the opportunity to autonomously learn how to best operate multiple actors, where one's actions may affect the operations of the others. The subfield of cooperative multi-agent learning is especially important, as it focuses on those domains where the actors need to achieve a joint task, irrespectively of their individual preferences. Thus, cooperative multi-agent RL has the potential to significantly increase the efficiency of real world settings such as traffic control, warehouse management, wind farm control and many others. One of the main challenges when learning in these scenarios is how to handle large numbers of agents, especially without requiring vast amount of data, which may be hard and expensive to gather. This is because, as the number of agents to control increases, the number of possible joint policies increases exponentially, making naive approaches impractical if not outright infeasible.

In this dissertation, we specifically tackle the aspects of scalability and sample-efficiency in cooperative multi-agent RL. We approach these problems by leveraging model-based methods, which allow us to simultaneously achieve two distinct goals: to incorporate prior domain knowledge about a given problem into the learning process, and extract as much information as possible out of each interaction with the real environment. In particular, we focus on domain knowledge in the form of coordination graphs, which contain locality information about the environment, i.e. about which agents can directly or indirectly interact together. These graphs allow for much more efficient learning, by factorizing the problem into smaller, simpler components. These components can then be brought back together by using specialized optimization algorithms specifically designed to work on factored domains.

Our main contributions consist in several novel multi-agent bandit algorithms for both regret minimization and best-arm identification, as well as a novel multi-agent MDP algorithm that can efficiently scale to settings with hundreds of agents. We additionally provide theoretical proofs for the bandit algorithms proving the validity of our approach, as well as comprehensive empirical tests for all our presented methods against a variety of state-of-the-art benchmarks.