SCIENCES & BIOENGINEERING SCIENCES

The Research Group
Artificial Intelligence

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

Hélène Plisnier

to obtain the degree of Doctor of Sciences

Title of the PhD thesis:
Guiding the Exploration Strategy of a Reinforcement Learning Agent

Promotor:
Prof. dr. Ann Nowé (VUB)

The defense will take place on
Tuesday, April 18, 2023 at 16h in auditorium D.2.01

Members of the jury

Prof. dr. Elisa Gonzalez Boix (VUB, chair)
Dr. Roxana Rădulescu (VUB, secretary)
Prof. dr. Kris Steenhout (VUB)
Prof. dr. Aske Plaat (Leiden Institute of Advanced Computer Science)
Prof. dr. Philippe Preux (Inria)

Curriculum vitae

Hélène Plisnier (1995) obtained her BSc degree in Computer Science in 2015 and her MSc degree in Computer Science in 2017, both at the Université Libre de Bruxelles. She then joined the Artifical Intelligence Laboratory (VUB) as a PhD student. In 2019, she obtained a FWO strategic basic research grant. Hélène's research focuses on Reinforcement Learning (RL), specifically on how to improve the sample-efficiency of existing RL techniques. She invented the Actor-Advisor, and developed Shepherd, a web application making the use of RL methods more accessible to a non-expert audience. Her research resulted in 5 publications in international peer-reviewed conferences and workshops.

Abstract of the PhD research

Reinforcement Learning (RL) is a Machine Learning method mimicking the way humans learn to perform new tasks, specifically when it involves a great amount of trial and error. By nature, it is a progressive process that requires many interactions with the environment, and therefore time, before it can exhibit a satisfactory behavior. Consequently, an important challenge faced by current RL techniques is sample-efficiency. While most approaches to reduce the amount of samples needed focus on extracting a maximum amount of information out of each sample, I explore ways to improve the exploration strategy of the learner. Pushing the learning agent towards fruitful areas of the search space, and preventing it from wasting its time in undesirable areas, helps the agent reach a good policy faster and more efficiently. I present the Actor-Advisor, a general-purpose Policy Shaping method, allowing an external advisory policy to influence the actions selected by an RL agent. I extend my main contribution to a wide range of settings, such as discrete and continuous actions spaces, using on or off-policy Reinforcement Learning algorithms. I design the learning correction to let Policy Gradient-based methods benefit from off-policy external guidance, despite their strong on-policyness.

I evaluate the Actor-Advisor in two important RL sub-fields: learning from human intervention, and Transfer Learning. Although almost any source can be used as an advisor of an RL agent following the Actor-Advisor framework, the focus of this thesis is applying the Actor-Advisor to several novel Transfer Learning problems. Transfer Learning is resolutely related to sample-efficiency, since it aims at making the learning of new tasks faster by smartly reusing knowledge acquired in previous tasks. Finally, I introduce Self-Transfer, a learning trick inspired by Transfer Learning, in which an RL agent can easily improve its sample-efficiency by using an advisor pre-trained for a short while on the same task. I hope that my contributions will help promote the use of Reinforcement Learning methods in future real-life problems.