VUB SCIENCES & BIOENGINEERING SCIENCES

The Research Group
Artificial Intelligence Lab

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

Raphaël Avalos Martinez de Escobar

to obtain the degree of Doctor of Sciences

Title of the PhD thesis:

Leveraging State Access in Partially Observable Sequential Decision-Making

Supervisor:
Prof. dr. Ann Nowé (VUB)
Co-supervisor:
Dr. Diederik M. Roijers (VUB, city of Amsterdam)

The defence will take place on
Friday, February 6, 2026 at 4.30 p.m.

VUB Etterbeek campus, Pleinlaan 2, Elsene, auditorium D.0.05

The defence can be followed through a live stream: https://avalos.fr/live_thesis

Members of the jury

Prof. dr. Coen De Roover (VUB, chair)
Prof. dr. Lynn Houthuys (VUB)
Prof. dr. Jan Lemeire (VUB)
Dr. Nicole Orzan (VUB)
Prof. dr. Karl Tuyls (IMEC, University of Liverpool, UK)
Prof. dr. Christopher Amato (Northeastern University, USA)

Curriculum vitae

Raphaël Avalos obtained his engineering degree in Computer and Data Science from Télécom ParisTech (2019) and his Master's degree in Applied Mathematics for AI from ENS Paris-Saclay (2019). In 2020, he joined the AI Lab at the VUB, where he was awarded an FWO Fellowship. His research has been published in leading international conferences and journals. During his PhD, Raphaël also served as a teaching assistant and was involved in the organisation of international workshops.

Abstract of the PhD research

Robots, autonomous systems, and multi-agent teams must often make decisions based on incomplete or noisy sensor information, a challenge known as partial observability. While most methods assume complete uncertainty, real systems can sometimes access accurate state information: either during training through simulation, or during deployment by activating costly high-precision sensors.

This thesis investigates how to strategically leverage such limited state access to improve decision-making under uncertainty.
We address this challenge across three distinct settings. In cooperative multi-agent reinforcement learning, we propose Local Advantage Networks (LAN), which stabilizes training of independent agents through a centralized critic conditioned on the full state that can be discarded during execution, achieving state-of-the-art performance on standard benchmarks.

For online planning with costly state queries, we develop AEMS-SR, a graph-based online planning algorithm designed to handle the complexity of costly state queries and which outperforms standard planners in test domains.

For model-based reinforcement learning, we introduce the Wasserstein Belief Updater (WBU), which learns latent belief updates and provides theoretical guarantees on belief quality by leveraging state access during training.

By treating state access as a valuable but limited resource and demonstrating when and how to exploit it, this thesis shows that agents can achieve more robust and effective decision-making in partially observable environments.