

De Onderzoeksgroep
Artificial Intelligence Lab

nodigt U graag uit op de openbare verdediging van het proefschrift van

Raphaël Avalos Martinez de Escobar

ter behaling van de graad van Doctor in de wetenschappen

Titel van het proefschrift:

**Leveraging State Access in
Partially Observable Sequential Decision-Making**

Promotor:

Prof. dr. Ann Nowé (VUB)

Co-promotor:

**Dr. Diederik M. Roijers (VUB, gemeente
Amsterdam)**

De verdediging heeft plaats op

vrijdag 6 februari 2026 om 16u30

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

De verdediging is ook te volgen via een livestream:
https://avalos.fr/live_thesis

Samenstelling van de jury

Prof. dr. Coen De Roover (VUB, voorzitter)

Prof. dr. Lynn Houthuys (VUB)

Prof. dr. Jan Lemeire (VUB)

Dr. Nicole Orzan (VUB)

Prof. dr. Karl Tuyls (IMEC, University of Liverpool,
VK)

Prof. dr. Christopher Amato (Northeastern
University, USA)

Curriculum vitae

Raphaël Avalos behaalde zijn ingenieursdiploma in Computer and Data Science aan Télécom ParisTech (2019) en zijn Master in Applied Mathematics for AI aan ENS Paris-Saclay (2019). In 2020 vervoegde hij het AI Lab VUB waar hij een FWO Fellowship behaalde. Zijn onderzoek werd publiceerd in toonaangevende internationale conferenties en tijdschriften. Tijdens zijn doctoraat was Raphaël ook assistent en betrokken bij de organisatie van internationale workshops.

Abstract van het doctoraatsonderzoek

Veel beslissingsproblemen in robotica, autonome systemen en multi-agentcoördinatie gaan gepaard met partiële observeerbaarheid, waarbij agenten moeten handelen op basis van onvolledige informatie over de omgeving. Deze uitdaging vloeit vaak voort uit de fysieke beperkingen van sensoren.. Terwijl de meeste methodes veronderstellen dat de toestand nooit beschikbaar is, kan deze in de praktijk soms worden voorspeld: tijdens training, bv via een simulator of door het activeren van hulpsensoren. Dergelijke toegang maakt is een waardevolle maar schaarse hulpbron.

Dit proefschrift onderzoekt hoe strategisch gebruik van deze hulpbron reinforcement learning (RL) en plannen kan verbeteren. In RL tonen we aan hoe toegang tot toestand-informatie tijdens training interne representaties kan sturen, het leerproces kan stabiliseren en de prestaties bij uitvoer zonder toestandinformatie kan verbeteren. In planning formaliseren en lossen we beslissingsproblemen op waarbij de agent de toestand tijdens uitvoering kan opvragen, waarbij winst van de informatie wordt afgewogen tegen de extra kost.

Ten eerste introduceren LAN voor coöperatief multi-agent RL, dat de training van onafhankelijke agenten stabiliseert via een gecentraliseerde baseline zonder gezamenlijke actie-waardefactorisatie te vereisen, en toch state-of-the-art prestaties behaalt. Ten tweede AEMS-SR, een grafgebaseerd online planningsalgoritme voor POMDP's, dat Anytime Error Minimization Search uitbreidt met heboud van ϵ -optimaliteits-garanties en dat beter presteert dan SoTa in testdomeinen. Ten derde ontwikkelen we de Wasserstein Belief Updater (WBU) voor model-based RL, die toestandinformatie tijdens training benut om latente belief-updates te leren. WBU biedt theoretische garanties voor de kwaliteit van beliefs via bisimulatieafstanden en behaalt sterke empirische prestaties in partieel observeerbare omgevingen.