

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

Eugenio Tiziano Gabriele Bargiacchi

ter behaling van de graad van Doctor in de Wetenschappen

Titel van het proefschrift:

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

Promotoren:

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

De verdediging heeft plaats op

Maandag 15 januari 2024 om 10.30u in Auditorium E.O.10

Samenstelling van de jury

Prof. dr. Ann Dooms (VUB, voorzitter)
Prof. dr. Bart Bogaerts (VUB, secretaris)
Prof. dr. Bas Ketsman (VUB)
Prof. dr. Jilles Dibangoye (INSA Lyon, Frankrijk)
Prof. dr. Matthijs Spaan (TU Delft, Nederland)

Curriculum vitae

Eugenio Bargiacchi behaalde een B.Sc. in Computer Science aan de Università Statale di Milano in Italië (2011), waarbij hij zich richtte op informaticafundamenten en statistiek. Hij volgde dit met een M.Sc. in Kunstmatige Intelligentie aan de Vrije Universiteit Amsterdam in Nederland (2016). Hij begon zijn doctoraat in Kunstmatige Intelligentie aan het AI Lab aan de Vrije Universiteit Brussel in 2017, en kreeg een FWO-beurs om in 2019 aan zijn doctoraat te werken. Hij werkt nu in de privésector.

Abstract van het doctoraatsonderzoek

Multi-agent reinforcement learning (RL) biedt de mogelijkheid om autonoom te leren hoe meerdere actoren het beste kunnen opereren, waarbij de acties van een van hen de operaties van de anderen kunnen beïnvloeden. Het deelgebied van coöperatief multi-agent leren is bijzonder belangrijk, omdat het zich richt op die domeinen waar de actoren een gezamenlijke taak moeten uitvoeren, ongeacht hun individuele voorkeuren. Zo heeft coöperatieve multi-agent RL het potentieel om de efficiëntie van reële settings zoals verkeerscontrole, magazijnbeheer, controle van windmolenparken en vele andere aanzienlijk te verhogen. Een van de belangrijkste uitdagingen bij het leren in deze scenario's is hoe om te gaan met grote aantallen agenten, vooral zonder een enorme hoeveelheid gegevens nodig te hebben, die moeilijk en duur te verzamelen kunnen zijn. Dit komt omdat, naarmate het aantal te controleren agenten toeneemt, het aantal mogelijke gezamenlijke beleidslijnen exponentieel toeneemt, waardoor naïeve benaderingen onpraktisch of zelfs onuitvoerbaar worden.

In dit proefschrift pakken we specifiek de aspecten schaalbaarheid en steekproefefficiëntie in coöperatieve multi-agent RL aan. We benaderen deze problemen door gebruik te maken van modelgebaseerde methoden, die ons in staat stellen om gelijktijdig twee verschillende doelen te bereiken: het incorporeren van eerdere domeinkennis over een gegeven probleem in het leerproces, en zoveel mogelijk informatie halen uit elke interactie met de echte omgeving. In het bijzonder richten wij ons op domeinkennis in de vorm van coördinatiegrafieken, die localiteitsinformatie bevatten over de omgeving, d.w.z. over welke agenten direct of indirect met elkaar kunnen interageren. Deze grafieken maken veel efficiënter leren mogelijk, door het probleem in kleinere, eenvoudiger componenten te verdelen. Deze componenten kunnen dan weer worden samengebracht met behulp van gespecialiseerde optimalisatie-algoritmen die speciaal zijn ontworpen om te werken op gefactoriseerde domeinen.

Onze belangrijkste bijdragen bestaan uit verschillende nieuwe multi-agent bandit algoritmen voor zowel spijtminimalisatie als best-arm identificatie, alsook een nieuw multi-agent MDP algoritme dat efficiënt kan worden opgeschaald naar settings met honderden agenten. Daarnaast leveren wij theoretische bewijzen voor de bandit algoritmen die de geldigheid van onze aanpak aantonen, alsmede uitgebreide empirische tests voor al onze gepresenteerde methoden tegen een verscheidenheid van state-of-the-art benchmarks.