

De Onderzoeksgroep

Artificial Intelligence

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

Hélène Plisnier

ter behaling van de graad van Doctor in de Wetenschappen

Titel van het proefschrift:
Guiding the Exploration Strategy of a Reinforcement Learning Agent

Promotor:
Prof. dr. Ann Nowé (VUB)

De verdediging heeft plaats op
Dinsdag 18 april 2023 om 16u in aula D.2.01

Samenstelling van de jury

Prof. dr. Elisa Gonzalez Boix (VUB, voorzitter)
Dr. Roxana Rădulescu (VUB, secretaris)
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) behaalde haar bachelordiploma Computerwetenschappen in 2015 en haar masterdiploma Computerwetenschappen (specialisatie Artificiële Intelligentie en Embedded System Design) in 2017, beide aan de Université Libre de Bruxelles. Hierna vervoegde ze het Artificial Intelligence Laboratory (VUB) als doctoraatstudent. In 2019 behaalde ze een FWO strategisch basisonderzoeksmandaat. Hélène's onderzoek is gericht op Reinforcement Learning (RL), meer bepaald het verbeteren van de voorbeeldefficiëntie van bestaande RL algoritmen. Ze ontwikkelde de Actor-Advisor, een Policy Shaping methode voor algemene doeleinden, werkte op een verscheidenheid van Transfer Learning problemen, en ontwikkelde Shepherd, een web-toepassing dat het gebruik van RL algoritmen gemakkelijker maakt voor niet-experten. Haar onderzoek leidde tot 5 publicaties in internationale peer-reviewed conferenties en workshops.

Abstract van het doctoraatsonderzoek

Conditioning, of Reinforcement Learning (RL), is een Machine Learning methode dat de manier nabootst waarop mensen nieuwe taken leren uitvoeren, voornamelijk wanneer het gaat over vele keren vallen en opstaan. Van nature, is het een progressieve werkwijze dat veel interacties met de omgeving vereist, en daarom vergt het ook tijd, alvorens er een bevredigend gedrag vertoond kan worden. Bijgevolg, is voorbeeldefficiëntie een belangrijk probleem bij bestaande algoritmen. Terwijl de meeste technieken, die de hoeveelheid voorbeelden verminderen, focussen op het verkrijgen van zo veel mogelijk informatie, verken ik manieren om de exploratie strategie van de agent te verbeteren. De agent naar vruchtbare gebieden van de omgeving loodsen, en tijdsverspilling in ongewenste gebieden verwijderen, helpt de agent om sneller en efficiënter een goede policy leren. Ik introduceer de Actor-Advisor, een Policy Shaping methode voor algemene doeleinden, dat een extern adviserende policy toelaat om de acties die een RL agent selecteert te beïnvloeden. Ik voeg mijn voornaamste contributie toe aan een brede reeks van situaties, zowel discrete en continue actieruimtes, als on- en off-policy RL algoritmen. Ik ontwerp de leercorrectie om technieken gebaseerd op Policy Gradient te laten profiteren van off-policy externe begeleiding, ondanks dat zulke technieken sterk on-policy zijn.

Ik evalueer de Actor-Advisor in twee belangrijke RL subdomeinen: leren uit menselijke interventie, en Transfer Learning. Hoewel bijna onder welke bron als adviseur gebruikt kan worden, ligt de nadruk van dit proefschrift op het toepassen van de Actor-Advisor op verscheidene nieuwe Transfer Learning problemen. Transfer Learning is resoluut verwant met voorbeeldefficiëntie, aangezien het doel eveneens is om nieuwe taken sneller te leren, door kennis verworven in vorige taken slim te hergebruiken. Tenslotte, stel ik Self-Transfer voor, a leertruc geïnspireerd op Transfer Learning, waarin een RL agent zijn voorbeeldefficiëntie kan verbeteren door een adviseur te gebruiken die voor een korte periode voorgetraind op dezelfde taak. Ik hoop dat mijn contributies het gebruik van Reinforcement Learning methoden zullen helpen promoten op problemen in echte leven.