



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

## Mathieu Reymond

ter behaling van de graad van Doctor in de Wetenschappen

Titel van het proefschrift:

**On incorporating prior knowledge about the decision maker in multi-objective reinforcement learning**

Promotor:

**Prof. dr. Ann Nowé**

Co-promotor:

**Dr. Diederik M. Roijers**

De verdediging heeft plaats op

**Woensdag 30 augustus 2023 om 14u in aula  
I.0.01**

### Samenstelling van de jury

Prof. dr. Geraint A. Wiggins (VUB, voorzitter)

Prof. dr. Elisa Gonzalez Boix (VUB, secretaris)

Prof. dr. Sam Verboven (VUB)

Prof. dr. Peter Vamplew (Federation University  
Australia)

Prof. dr. Frederik Heintz (Linköping University,  
Sweden)

### Curriculum vitae

Mathieu Reymond studeerde in 2017 af als MSc in de Computerwetenschappen, met een specialisatie in Artificiële Intelligentie aan de VUB. Daarna startte hij als doctoraatstudent en onderwijsassistent in het Artificial Intelligence Lab, VUB.

Zijn onderzoek richt zich op besluitvormingsproblemen met meerdere doelstellingen. Hij heeft respectievelijk 9, 5 peer-reviewed artikelen gepubliceerd in internationale conferenties en journals, workshops en 1 preprint gepubliceerd.

Zijn onderwijstaken lieten hem toe tot het begeleiden van 2 master thesissen en 1 bachelorproef.

### Abstract van het doctoraatsonderzoek

Besluitvormers moeten vaak rekening houden met meerdere doelstellingen. Wanneer het maximaliseren van één doelstelling ten koste gaat van een andere, zijn de doelstellingen conflicterend, en moet de besluitvormer een compromisoplossing vinden. De optimale afwegingen kunnen per geval verschillen, omdat ze afhangen van de voorkeuren, of het nut, van de beslisser.

In dit proefschrift onderzoeken we nutsgebaseerde benaderingen, die rekening houden met voorkennis over het nut van de beslisser, om zo zijn of haar voorkeursafweging te leren kennen. Door efficiënt gebruik te maken van deze voorkennis kunnen we ongewenste afwegingen snel verwerpen, waardoor de zoekruimte naar de optimale oplossing aanzienlijk kleiner wordt en we dus sneller de gewenste afweging vinden.

Wij analyseren verschillende scenario's, afhankelijk van de beschikbare voorkennis. Ten eerste nemen we aan dat het nut a priori bekend is, en stellen we een nieuw multi-objectief reinforcement learning algoritme voor dat rechtstreeks op dat nut optimaliseert. We tonen aan dat, door expliciet rekening te houden met meerdere objectieven, we de optimale afweging leren op een stabielere en efficiëntere manier dan met single-objective algoritmen. Ten tweede beschouwen wij het interactieve scenario, met slechts gedeeltelijke voorkennis over het nut, maar waarbij wij met de beslisser kunnen interageren om zijn of haar voorkeuren te leren kennen. We kunnen hierbij de timing van onze vragen optimaliseren om de probabiliteit op het leren van de gewenste compromis te maximaliseren. Ten derde veronderstellen wij geen voorkennis over het nut. Het leren van een enkele, maar voorwaardelijke oplossing op voorkeuren stelt ons in staat de geleerde data te hergebruiken voor verschillende afwegingen, waardoor de efficiëntie van het zoekproces toeneemt. Door onze oplossing te generaliseren naar alle mogelijke voorkeuren, kunnen we elke mogelijke afweging leren, zodat de beslisser a posteriori zijn voorkeursoplossing kan kiezen.

In ons onderzoek richten wij ons voornamelijk op sequentiële besluitvormingsproblemen, waarbij een oplossing wordt gevonden na het nemen van een reeks handelingen. Wij demonstreren onze bevindingen voor onbekende nutsfuncties op een reëel praktijkgeval, namelijk epidemische uitbraken. Door het aantal sociale contacten op kritische momenten te beperken, kunnen we de verspreiding van de epidemie beheersen. Wij leren een verscheidenheid aan optimale strategieën die een evenwicht vinden tussen de ziekenhuisopnames ten gevolge van de uitbraak en de vermindering van het aantal sociale contacten, wat de besluitvormer kan helpen bij het nemen van een geïnformeerde beslissing, met kennis van de beschikbare alternatieven.