

The Research Group
Artificial Intelligence Laboratory

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

Mathieu Reymond

to obtain the degree of Doctor of Sciences

Title of the PhD thesis:

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

Promotor:

Prof. dr. Ann Nowé

Co-promotor:

Dr. Diederik M. Roijers

The defence will take place on
**Wednesday, August 30, 2023 at 14h in
auditorium I.0.01**

Members of the jury

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

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

Prof. dr. Sam Verboven (VUB)

Prof. dr. Peter Vamplew (Federation University
Australia)

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

Curriculum vitae

Mathieu Reymond graduated as MSc in Computer Science with a specialization in Artificial Intelligence at the VUB in 2017. He then joined the Artificial Intelligence Lab, VUB as a PhD student and teaching assistant.

His research focuses on decision-making problems with multiple objectives. He has published 9 peer-reviewed articles in international conferences and journals, 5 peer-reviewed articles in workshops, and 1 preprint.

His teaching duties allowed him to supervise 2 master theses and 1 bachelor thesis.

Abstract of the PhD research

Decision makers acting in real-world problems often have to take into account multiple objectives. When maximizing one objective comes at the cost of another, the objectives are in conflict, and the decision maker must find a compromise between them. The optimal trade-offs might differ on a case-by-case basis, as they depend on the preferences, or utility, of the decision maker.

In this dissertation, we investigate utility-based approaches, that take into account prior knowledge over the decision maker's utility, to learn his or her preferred trade-off. By making efficient usage of this prior knowledge, we can swiftly discard undesirable trade-offs, significantly narrowing down the search space for the optimal solution, and thus enhancing the speed at which we find the desired trade-off.

We analyze different scenarios, depending on the amount of prior knowledge available. First, we consider that the utility is known a priori, and propose a novel multi-objective reinforcement learning algorithm that optimizes directly on said utility. We show that, by explicitly considering multiple objectives, we learn the optimal trade-off in a more stable and efficient manner than using single-objective solvers. Second, we consider the interactive scenario, with only partial prior knowledge over the utility, but where we can query the decision maker to learn about its preferences. We show that we can optimize the timing of our queries during the learning process to maximally improve our chances of learning the preferred trade-off. Third, we assume no prior knowledge over the utility. Learning a single, but conditional solution on preferences allows us to reuse the samples learned for different trade-offs, thus improving the efficiency of the search. By generalizing our solution to all possible preferences, we can learn any possible trade-off, such that the decision maker can choose its preferred solution a posteriori.

Throughout our research, we primarily focus on sequential decision-making problems, where a solution is found after having taken a sequence of actions. We demonstrate our findings for unknown utility functions on a real-world use-case, epidemic outbreaks. By reducing the number of social contacts at key points in time, we can control the spread of the epidemic. We learn a variety of optimal strategies that balance between the hospitalizations due to the outbreak and the social contact reduction, which can support the decision maker in taking an informed decision, knowing the available alternatives.