

This invitation is extended to all Staff and HDR Students

RESEARCH SEMINAR

Centre for Informatics and Applied Optimisation

Speaker: Dr Diederik Roijers, Vrije
Universiteit Brussel, Belgium



Date: Thursday 17 August 2017 @ 11.30am

Room: T121, T Building, Mt Helen Campus
(Or Visimeet ID: 1607818)

Title: “Policy / Value Reuse in Incremental Multi-objective planning and learning”

Abstract:

Most decision problems have more than one objective, and are thus most naturally expressed using a vector-valued reward signal. While it might seem tempting to define one scalar reward function for these problems anyway by combining the different objectives in some way a priori, doing so can be error-prone, and will force human decision-makers to define (hypothetical) preferences before being able to look at the actually available alternatives. Instead, we can apply multi-objective planning or learning to obtain a set of possibly optimal policies and policy values per objective, and ask humans to define their preferences between these actual alternatives. This typically leads to a more well-informed decision, and is therefore highly preferable.

In this talk, we start from a brief overview of multi-objective decision making and how different assumptions about the utility function of the user lead to different solution sets. Thereafter, we zoom in on so-called outer loop solution methods, in which a multi-objective decision problem is solved as a sequence of single-objective problems. In contrast with so-called inner loop methods, outer loop methods can be applied across different multi-objective problem classes, provided that a suitable single-objective solution method exists. We propose the outer loop optimistic linear support (OLS) algorithm, which produces a convex coverage set, i.e., the optimal solution set when the utility function of the user is known to be linear (or policies can be stochastic). OLS can be applied to a variety of problems and applications, such as multi-objective coordination graphs, multi-objective MDPs and multi-objective POMDPs. We show how OLS can a) produce the optimal CCS, b) produce an epsilon-approximate CCS (in much less time) and c) can be sped-up by reusing policy and value information from previous iterations of the algorithm.

Biography:

Diederik M. Roijers did his masters in Computing Science at Utrecht University, before obtaining his PhD in Artificial Intelligence under the supervision of Shimon Whiteson and Frans A. Oliehoek at the University of Amsterdam in 2016, after which he joined the University of Oxford as a postdoctoral research assistant. He was awarded a Postdoctoral Fellowship Grant from FWO (Scientific Research Fund - Flanders) and has been working as an FWO Postdoctoral Fellow at the Vrije Universiteit Brussel since October 2016.

Diederik's research focuses on creating intelligent autonomous systems that assist humans in solving complex problems with multiple objectives. He studies the whole process of multi-objective decision making, including decision theoretic planning and learning, which enables agents to reason about the environments in which there are multiple reward signals; preference elicitation and machine learning to discover and model the preferences of human decision makers; and game theory and negotiation when there are multiple agents and/or decision makers involved. The ultimate goal of this process is to assist human decision makers in making well-informed decisions in complex real-world decision problems.

Deiderik is visiting Federation University for research collaboration with CIAO researcher A/Prof Peter Vamplew.

The seminar is easily available to staff and students at other FedUni campuses via the Visimeet system from their desk tops. If you would like to join the seminar please contact Evan Dekker, e.dekker@federation.edu.au who will organise an invitation to join the seminar.