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 14 August 2017 @ 1.30pm**

Room: **T121, T Building, Mt Helen Campus**
(Or Visimeet ID: 1718236)

Title: **"Interactive Multi-objective Reinforcement Learning"**

Abstract:

In multi-objective reinforcement learning (MORL), much attention is paid to generating optimal *solution sets* for unknown utility functions of users, based on the stochastic reward vectors only. In *online* MORL on the other hand, the agent will often be able to elicit preferences from the user, enabling it to learn about the utility function of its user directly. In this talk, we focus on online MORL with user interaction employing the multi-objective multi-armed bandit (MOMAB) setting.

We use Bayesian learning algorithms to learn about the environment and the user simultaneously. Specifically, we present two algorithms from our recent work: *Utility-MAP UCB (umap-UCB)* and *Interactive Thompson Sampling (ITS)*, and show empirically that the performance of these algorithms in terms of regret closely approximates the regret of UCB and regular Thompson sampling provided with the ground truth utility function of the user from the start.

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.