EPSRC DTP PhD Research Project

Project Title: Uncertainty Quantification in Explainable Optimisation

Primary Supervisor details: Dr David Walker (D.J.Walker2@exeter.ac.uk; https://computerscience.exeter.ac.uk/people/profile/index.php?web_id=djw247).

Additional Supervisory team details: Dr Tinkle Chugh (T.Chugh@exeter.ac.uk; https://computerscience.exeter.ac.uk/people/profile/index.php?web_id=tc489).

Department: Computer Science

Location: Innovation 1, Streatham campus, Exeter

PhD Programme: PhD Computer Science

Project Description:

Explainable AI (XAI) has arisen as a tool for conveying understanding of how machine learning models work, however recent work has highlighted the potential for using XAI within the domain of nature-inspired optimisation [1]. Potential use cases for XAI within optimisation include explaining the process taken by an optimiser to generate solutions to a problem to the problem owner, and assisting the decision maker in their selection of a final solution to implement from a set of solutions generated by the optimiser by contrasting the alternatives with methods such as counterfactuals.

Incorporating XAI into metaheuristic optimisation lowers the barrier to understanding their behaviour and the solutions they generate, which in turn increases their potential for wider adoption and industrial impact.

A popular XAI tool within the ML community is LIME [2], which provides explanations of complex black-box models by fitting cheap to compute and interpretable linear models to mimic the global behaviour of the full model. By examining the behaviour of these linear models the user can gain an appreciation of how their data is being processed by the model to produce the prediction that it has resulted in. Recently, a similar approach was taken to explain the importance of individual variables comprising a solution to an optimisation problem, showing the potential of such an approach [3].

Uncertainty is inherent to metaheuristics – from the stochasticity arising from random initialisation of a search population to the random moves underpinning solution generation mechanisms such as crossover and mutation, as well as noise in the solution and objective

space combined with potentially multimodal search landscapes – the results that we seek to explain (solutions and algorithm processes) must be explained in the presence of noise. By quantifying this uncertainty we will provide a more robust explanation of the process and solutions.

This project will explore the potential of using Bayesian models to enable the interpretation of the evolutionary process while quantifying uncertainty around the data. The project will utilise optimisation problems from multiple domains, which will be agreed upon by the project team. Due to the research interests of the project team, one of these will relate to optimisation within the offshore renewable energy domain. An additional research challenge is evaluating the explanations generated. One way of doing this is to employ user experience studies to evaluate the usefulness of the explanations with human users, which necessitates the use of problems that are accessible to a wide range of participants. For this reason we expect to also consider problems such as timetabling, personnel scheduling and packing problems from operations research.

References

[1] J. Bacardit *et al.*, "The intersection of evolutionary computation and explainable AI", Proc. GECCO Workshop on ECXAI, 2022.

[2] M. Ribeiro *et al*., ""Why should I trust you?" Explaining the predictions of any classifier", Proc. KDD, 2016.

[3] M. Singh *et al.*, "Towards Explainable Metaheuristic: Mining Surrogate Fitness Models for Importance of Variables", Proc. GECCO Workshop on ECXAI, 2022.

Project specific enquiries: David Walker, Tinkle Chugh