Guest editorial: special issue on evolutionary optimization, feature reduction and learning

Bing Xue¹ · Gang Chen¹

Published online: 12 August 2016
© Springer-Verlag Berlin Heidelberg 2016

Evolutionary techniques enjoy proven capability of delivering high-quality solutions to challenging problems in various scientific and technical fields. Cross-fertilization between evolutionary technique and machine learning approaches provide a new source of inspiration and has attracted great attention in recent years. Particularly evolutionary optimization, evolutionary feature reduction and evolutionary learning have been successfully applied to a wide range of real-world problems.

This special issue brings together some state-of-the-art works from a wide variety of topics concerning evolutionary optimization, feature reduction and learning, including multi-agent games, fitness landscape analysis, many-objective optimization, multi-objective feature selection, sentiment classification, clustering, web service composition, reinforcement learning and time series analysis. The ten papers included in this special issue originated from SEAL'14 (the 10th International Conference on Simulated Evolution and Learning) but have been substantially extended and revised from the conference version. The extended papers were again reviewed and revised rigorously in two rounds before acceptance.

The paper by Kriley et al. investigates the evolutionary dynamics of the well-known public goods game using a recently introduced modelling framework that encapsulates risk-sensitive assortment. The population of mobile agents in the game are divided into fix-sized interaction groups.

Each agent is bestowed with a risk-sensitive trait that guides local decision-making and migration. Through comprehensive computational simulation experiments, it is discovered in the paper that, when the size of each group increases, assortment levels tend to decrease and risk-averse agents tend to dominate the population. Meanwhile, this paper shows the existence of high variance in the proportion of cooperative agents both in groups and between different groups, suggesting that risk-seeking behaviour is an emergent property of mobility induced positive assortment.

The paper by Shirakawa and Nagao studies local landscape features since the neighbourhood structure has an important role in evolutionary algorithms. Through their study, Shirakawa and Nagao propose two new landscape features in vector forms, i.e. the bag of local landscape patterns (BoLLP) and the bag of evolvability (BoEvo). Both BoLLP and BoEvo are represented by the histograms of local information, involving both solutions and their neighbours. This paper also proposes to combine landscape features with different sample sizes so as to represent both fine-grained and coarse-grained landscape features. The proposed landscape features are further investigated on some well-known benchmark functions and the BBOB benchmark function set. The effectiveness of these features is further discussed based on the experimental results.

The paper by Sato focuses on improving the solution update method in the conventional MOEA/D algorithm for multi- and many-objective optimizations. In particular, a new chain-reaction solution update method has been successfully developed in the paper with the aim of improving the genetic diversity in the population by avoiding unnecessary solution duplications. Meanwhile, solutions will be updated in an orderly manner depending on the position of each offspring in the objective space. Experiments have been performed on the discrete knapsack and continuous WFG4 problems with

✉ Bing Xue
bing.xue@ecs.vuw.ac.nz

Gang Chen
aaron.chen@ecs.vuw.ac.nz

¹ Evolutionary Computation Research Group, School of Engineering and Computer Science, Victoria University of Wellington, Wellington, New Zealand

2–8 objectives. The results show that the proposed method can noticeably improve the search performance of MOEA/D.

The paper by Zhou et al. investigates binary particle swarm optimization (BPSO) for feature selection, particularly for sentiment classification tasks in text mining. Since PSO was originally proposed as a continuous optimization method, when the ideas were adopted to BPSO, there is a number of limitations. This paper discusses the major limitations of BPSO, such as unreasonable update formula of velocity and lack of evaluation on every single feature. Then a new BPSO algorithm is proposed, where a fitness proportionate selection mechanism is employed. The proposed BPSO is tested on sentiment classification-oriented feature selection problems. The experimental results show that the improved BPSO outperformed the original PSO to select a number of good features to increase the classification performance when using CART, NB and SVM on two sentiment classification data sets.

The paper by Gallagher investigates the use of clustering problems as a source of real-world and standard benchmark problems for the evaluation and comparison of black-box optimization algorithms. Based on three well-known data sets (i.e. Iris, Ruspini and German Towns), a set of 27 problem instances (from 4-D to 40-D) of varying difficulty is proposed. A set of baseline results on these problem instances are provided by using a number of clustering algorithms, including Covariance Matrix Adaptation–Evolution Strategy (CMA-ES) and several other standard algorithms. In addition, a web repository has also been created for this problem set to facilitate future use for algorithm evaluation and comparison. Overall, the results show that CMA-ES provides relatively good performance on many of these problems (with some variation of population size), often finding the global optimum and sometimes far outperforming the standard MATLAB solver (Nelder-Mead) and the widely used (non-black box) *k*-means clustering algorithm.

The paper by Sawczuk da Silva et al. investigates on web services composition to choose a set of web services from a large number of candidate services. This paper develops a new genetic programming (GP)-based method for web service composition, where three variations of GP methods are investigated using fitness penalizations, restricted population initialization and restricted genetic operators are used to enforce the correctness of solutions. The fitness function is to optimize the quality of service in each solution. A set of experiments have been carried out to compare the three GP variations with two PSO approaches. The results show that GP methods have a higher execution time for most data sets, but the results indicate that GP methods scale better than the PSO methods. In addition, there seems a trade-off between execution time and the quality of solutions when employing GP and PSO.

The paper by Yliniemi and Tumer studies the credit assignment issue in multi-agent multi-objective problems. Specifically, credit assignment through varied evaluations has been applied to two different policy selection paradigms to demonstrate its broad applicability. The reinforcement learning problem is first examined in the paper. The authors show that proper credit assignment can increase learning speed, enhance solution quality and retain effectiveness even when 20 % of agents act in their own interests. Subsequently, credit assignment is applied to NSGA-II. Experiment results indicate that NSGA-II considering credit assignment can dominate the best case performance of NSGA-II not considering credit assignment. The results strongly suggest that proper credit assignment is at least as important to performance as the choice of multi-objective algorithms.

The paper by Consoli et al. argues that, by using fitness measures of the offspring solutions alone, it is insufficient to determine the most suitable variation operators for use in order to improve the effectiveness of evolutionary algorithms. Driven by this understanding, an ensemble of four different online fitness landscape analysis techniques has been proposed in the paper to give a more accurate description of the current population. A credit assignment technique is further utilized to predict the reward of the most suitable variation operator based on two different reward measures. Experiment results in the paper show that the proposed approach can achieve solution quality comparable to state-of-the-art techniques. The results also reveal the beneficial effects of adaptive operator selection on the optimization ability.

The paper by Song et al. focuses on time series data analysis to automatically learning and discovering state patterns in time series data. A genetic programming (GP)-based method is proposed for learning from multichannel data streams based on body sensor readings, which is designed to not require prior knowledge of the relationships between channels and not require manually defined feature to be constructed, but using only raw input. The experiments have been conducted on both artificial and real-world multichannel time series data. The results show that when comparing with conventional classification methods which took both raw data and pre-defined feature data as input, the proposed GP method achieves significantly better results on raw inputs of synthetic single-channel state patterns, synthetic multi-channel state patterns and real-world multichannel state patterns. Further investigation on the evolved solutions analysis the reasons why the proposed method can work well.

The paper by Nguyen et al. focuses on feature selection for classification with the objectives of simultaneously reducing the dimensionality of the data and increasing the classification accuracy. This paper develops a new multi-objective PSO-based method to handle the two conflicting objectives, where a new archive updating local search method is pro-

posed to maintain a set of non-dominated solutions found during the evolutionary process, which are used as candidate leaders to guide the search of the swarm. Similarity-based interesting, removing and swapping operators are proposed to form the local search. The experiments on 12 benchmark data sets of varying difficulty show that in most cases, the proposed new method achieves better performance than three well-known evolutionary multi-objective algorithms and a current state-of-the-art PSO-based multi-objective feature selection approach.

We would like to express our sincere gratitude to the reviewers for their expertise and meticulous efforts in ensuring the quality of the papers for this special issue. Special thanks go to the following reviewers (in alphabetical order): Mario Andres Munoz Acosta, Wissam Albukhanajer, Harith Al-Sahaf, Slim Bechikh, Valerio Capraro, Rohitash Chan-

dra, Gang Chen, Shi Cheng, Richard Dazeley, Sam Devlin, Grant Dick, Zhun Fan, Wenglong Fu, Cheng-Lung Huang, Muhammad Iqbal, Jing Liu, Yuanning Liu, Olivi Matei, Yi Mei, Kourosh Neshatian, Peter Rockett, Tatsuya Sasaki, Andy Song, Chuan-Kang Ting, Peter Vamplew, Ling Wang, Huayang Xie, Bing Xue, Weijun Zeng, Zexuan Zhu, Juan Zou. Last but not least, we would like to thank the Managing Editor Professor Aniello Castiglione for his support and assistance during the editing process of this special issue.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.