Ordering and Visualisation of Many-objective Populations

David Walker, Richard Everson and Jonathan Fieldsend

University of Exeter

19 July 2010
Introduction

**Ordering Many-objective populations**

- League tables are used to rank performance
- Often multiple *key performance indicators* measure a single item
- Information must be drawn from all of these KPIs to come up with a single measure of performance
- We construct a league table with multiple KPIs applying a single-objective ranking method to a many-objective data set — *the Power Index*

**Visualising Many-objective populations**

- Ordering methods can be used to support visualisation
- We present
  - A new dominance/graph theory-based method
  - A method to improve the comprehension of *heatmaps*
League Table Construction

An example — university league tables

- The Times *Good University Guide* 2009
- 8 KPIs
- Weighted sum approach
- Choosing weights can be difficult

Many-objective Optimisation Concepts

- In MOEAs, dominance approaches replaced weighted sum methods
- The job of a university is to maximise their performance over all KPIs simultaneously
- Let $y = (y_1, \ldots, y_k)$, for $k$ KPIs.
Pareto Sorting

Use dominance to locate the Pareto shells

Provides a partial ordering
A Graphical Visualisation

- Each university is represented by a node.
- An edge between two universities means that the origin university dominates the destination university.
- Universities are shown in their Pareto shells.
- For clarity, edges between universities are only shown from one shell to the next.
A Graphical Visualisation

- Each university is represented by a node.
- An edge between two universities means that the origin university dominates the destination university.
- Universities are shown in their Pareto shells.
- For clarity, edges between universities are only shown from one shell to the next.
Constructing the Graph

Probability of Dominance

\[ p(y_i \succ y_j) = \alpha + \beta \]

- \( \alpha \) — the fraction of objectives on which \( y_i \) is better than \( y_j \).
- \( \beta \) — half of the fraction of objectives on which \( y_i \) is equal to \( y_j \).

The probability that \( y_i \) would beat \( y_j \) in a tournament on a single randomly selected objective

Adjacency matrix \( W \)

\( W_{ij} = p(y_i \succ y_j) \)

Walker et al.  
Ordering and Visualisation of Many-objective Populations  
19 July 2010  
6 / 17
Consider the population as a graph, and rank the individuals according to the amount of outgoing traffic at each node.

Outflow $\sigma_i^{\text{out}}$

$$\sigma_i^{\text{out}} = \sum_j W_{ij}$$

- $y_i$ is ranked higher than $y_j$ if $\sigma_{y_i}^{\text{out}} > \sigma_{y_j}^{\text{out}}$
- $W_{ij}$ represents a transition from node $i$ to node $j$.

Equivalent to the Average Rank
**Power Index**

**Look for the most influential individual**

1 – Oxford  
2 – Imperial  
3 – Cambridge  
4 – UCL  
5 – LSE  
6 – Warwick  
7 – St Andrews  
8 – Durham  
9 – King’s  
10 – Bristol  
11 – York  
12 – N’ham  
13 – Bath  
14 – S’ton  
15 – Edinburgh

**Rank the individuals that dominate the most powerful individuals highest**
The power index is the limit of the sequence $u^t = Wu^{t-1}$

- $u^0 = 1$

- $u^1 = Wu^0$ (equivalent to the average rank and outflow, $u^1 = \sigma_{i}^{\text{out}}$)

- $u^2 = Wu^1$ (the average rank, with some proportion of the average rank of dominated universities)

- $u = \lim_{t \to \infty} \frac{u^t}{\sum_i u^t_i}$
Blue universities have the most power and tend to define the next shell in terms of the universities they dominate.

Red universities have the least power.
We have demonstrated a method based on graph theory and dominance for visualising many-objectives.

We cast dominance in a probabilistic framework.

The power index has been applied to many-objective populations for the first time.

We have improved the information which can be understood from the dominance graph by colouring by the power index.
Seriation of Heatmaps

- Heatmaps are a useful way of visualising a many-objective population
- They can be unclear because of the arbitrary ordering of individuals and objectives

We can order the population so that like individuals are close together in the permutation.

The population can be seriated to order individuals or objectives
Seriation of Objectives

Similarity between objectives:

\[ A_{mn} = 1 - \frac{1}{N} \sum_{i=1}^{N} (r_i^m - r_i^n)^2 \]

Distance summation:

For a permutation \( \pi_n \) of the \( n \)th objective, minimise

\[ g(\pi) = \sum_{m,n} A_{mn}(\pi_m - \pi_n)^2 \]
Seriation of Objectives

Relax $\pi$ to a continuous variable $z$ and minimise

$$h(z) = \sum_{m,n} A_{mn}(z_m - z_n)^2$$

subject to $\sum_n z_n = 0$ and $\sum_n z_n^2 = 1$

The solution is the Fiedler vector, the smallest non-zero eigenvector, of the graph Laplacian $L$ of $A$

$$L = D - A$$

where $D$ is a diagonal matrix, $D_{ii} = \sigma_{i\text{out}}$
Seriation of Objectives

\[ g(\pi) = 2.65 \times 10^9 \]

\[ g(\pi) = 2.23 \times 10^4 \]
Analysis of performance by key performance indicators can be considered in a multi-objective context.

Dominance coupled with graph theory provides an intuitive 2D representation.

A probabilistic interpretation of dominance allows us to represent the population as a graph.

The power index, which ranks a population based on the quality of individuals dominated, was applied to a multi-objective population ranking problem.

Heatmap clarity can be improved by seriating.