

Visualisation and Ordering of Many-objective Populations

David J. Walker, *Student Member, IEEE*, Richard M. Everson, *Member, IEEE*
and Jonathan E. Fieldsend, *Member, IEEE*

Abstract—We introduce novel methods of visualising and ordering multi- and many-objective populations. We compare individuals by the probability that one will beat another in a tournament on a randomly selected objective. This defines a weighted directed graph representing the population.

We introduce a novel graphical representation of the many objective population based on Pareto shells. We examine leagues, Pareto shells, preference ordering, average rank, outflow, the stationary distribution and the power index for ordering the population finding that the average rank is equivalent to outflow and that these together with the power index are generally superior.

Finally, we show how to seriate objectives to enhance the interpretability of heatmap visualisations.

I. INTRODUCTION

Evolutionary techniques have proven successful in solving multi-objective optimisation problems by locating approximations to the Pareto front. Significant effort is currently directed towards the move from multi-objective optimisation problems, problems with as many as 4 or 5 conflicting objectives, to *many-objective* optimisation, problems with 6 or more objectives. However, with a Pareto front of perhaps many solutions on hand, it is useful to be able to distinguish between solutions, to be able to visualise them and to order them to aid in selection of one or two. A common way of deciding on a solution is to present the decision maker with a graphical visualisation of the Pareto set, so that they can select one of the solutions.

When dealing with a bi-objective problem, it is a simple matter to plot the solutions on two-dimensional axes. If there is a knee in the Pareto front, a point beyond which favouring either objective will cause significant degradation in the other, then that is often the region of the set from which the solution is chosen. However, straightforward visualisation of solutions by plotting the objectives and identification of a knee region is more difficult with three objectives and out of the question with four or more.

In this paper we present a novel, easily comprehensible method of visualising the relations between many-objective solutions and apply ordering techniques to indicate which solutions may be preferred. The methods we present are also applicable to the visualisation and ordering of general many-objective datasets which are important in two additional areas: (a) selection of individuals for mutation and crossover in evolutionary algorithms, [1], [2] and (b), the focus of this

paper, analysing the performance of, for example, hospitals, schools or universities which are placed into league tables according to their performance on a number of key performance indicators (KPIs). Here we illustrate the visualisation and ordering techniques on the 2009 Times Good University Guide data, (GUG09) which scores the performance of UK universities in 2008 [3].

The remainder of this paper is organised as follows. The rest of this section presents some background material about dominance and Pareto optimality, before some common techniques for visualisation are discussed in section II. Section III presents a more detailed description of some ordering techniques, whilst Section IV demonstrates an improvement to a popular visualisation technique. Section V concludes.

A. Many-objective Populations and Dominance

Multi- and many-objective optimisation often employs the dominance relation to differentiate between putative solutions. Dominance allows us to quantify whether or not an individual, represented as a vector \mathbf{x} of M objectives is better or worse than another than \mathbf{x}' . \mathbf{x} is said to *dominate* \mathbf{x}' if it is no worse than \mathbf{x}' on any objective, and better on at least one. More formally, if the objectives are to be minimised:

$$\mathbf{x} \prec \mathbf{x}' \iff \forall i(x_i \leq x'_i) \wedge \exists i(x_i < x'_i). \quad (1)$$

A pair of individuals where neither dominates the other are said to be *mutually non-dominating* with respect to each other, and an individual which is not dominated by any other individual in the population is *non-dominated*.

B. Many-objective League Table Data

Throughout this paper, we illustrate the problem of visualising and ordering a many-objective data set: the Good University Guide (GUG09) [3], which summarises the performance of 113 UK universities in 2008. Universities are assessed according to the following 8 criteria: student satisfaction, based on their score in the National Student Survey (NSS); research quality; student–staff ratio; services and facilities spend; entry standards; completion; good honours; and graduate prospects.

The focus of the GUG is to construct a league table ranking the universities. This is achieved by normalising each KPI to have mean zero and unit variance, from which a score is arrived at by forming a weighted sum of the standardised KPIs. Universities are then ordered according to their score.

Here we regard this as a many-objective optimisation problem, where a university's aim is to achieve a good score on each of the criteria and we examine a number of ordering techniques which avoid the well-known problems of how to

The authors are with the College of Engineering, Mathematics and Physical Sciences, University of Exeter, EX4 4QF, UK. (email: {D.J.Walker, R.M.Everson, J.E.Fieldsend}@exeter.ac.uk). David Walker was supported by an EPSRC CASE Studentship with Motorola during the preparation of this paper.

combine KPIs measured in different units and choose *a priori* weightings.

II. VISUALISATION

Visualisation of a many-objective population can be done in two ways; the population can be reduced to a dimensionality such that a multi-objective visualisation technique (such as a scatter plot in 2 or 3 dimensions) can be used, or the entire space can be visualised directly using a technique such as a heatmap.

A. Visualisation by Dimension Reduction

A common approach to dimension reduction is to reduce an M -dimensional space by finding its *principal components*. Principal component analysis (PCA) finds a new lower-dimensional set of coordinates (the principal components) so that projection onto the principal components captures the maximum variance among all linear projections. The principal components are easily found as the eigenvectors of the covariance matrix of the population.

PCA is a linear method and two nonlinear techniques which have been used to find a low dimensional representation of data are the Generative Topographical Mapping [4] and Self Organising Maps (SOMs) [5]; see [6] for more on their use in multi-objective visualisation.

Closely related to the Generative Topographic Mapping and SOMs is Valdés and Barton’s use of a neural network to map the high-dimensional data to two or three dimensions while minimising a similarity metric, such as the Sammon error which attempts to preserve distances between data in the original and low-dimensional spaces [7]. Valdés and Barton then view the low-dimensional representation using virtual reality, which can be tailored to meet the needs of a decision maker.

However, the difficulty with all of these dimension reduction techniques is that the projection to lower dimensions inevitably destroys the dominance relationships between individuals. In fact it is easy to construct examples where the projection into even one fewer dimensions is bound to alter the dominance relationships between projected points, because if dimension reduction made no difference, one of the objectives would be superfluous.

B. Many-objective Visualisation

A straightforward way of visualising many-objectives is to draw a *parallel coordinate plot* [8] in which each individual is represented as a line of objective values. Whilst being scalable to any number of objectives, the resulting plots can be cluttered as each individual overlays another.

Pryke et al. [9] propose the use of a *heatmap* for visualising many-objective Pareto sets. In a heatmap plot numerical values corresponding to each objective and solution are shown by their ‘heat’ or colour. In order for the colours in the heat map to be meaningful, the values displayed in it must be on the same scale. Nazemi et al. [10] achieve this by linearly scaling all objectives to the range (0, 1).

Advantages of employing a heatmap for visualising a Pareto set include its scalability—heatmaps scale easily to any number of objectives—and heatmaps allow visualisation of both objective and decision space simultaneously. They are also straightforward to produce, the only computational work being the normalisation.

The main disadvantage can be their clarity. A heatmap for a Pareto front, in which the solutions are arbitrarily ordered, will likely appear to be a random collection of colours from which it can be difficult to extract any meaning. In section IV we show how solutions may be ordered to place similar ones together, thus enhancing their interpretability.

III. VISUALISING AND ORDERING MANY-OBJECTIVE POPULATIONS

In order to visualise and order a multi-objective population X it can be helpful to regard individuals $\mathbf{x}_i \in X$ in the population as nodes of a graph, with edges describing dominance relations between individuals. For example, $\mathbf{x}_i \prec \mathbf{x}_j$ might imply the presence of a directed edge from \mathbf{x}_i to \mathbf{x}_j . In general, the graph $G = (X, E, \mathbf{W})$ is defined by the set of nodes X , the set of directed edges $e_{ij} \in E$ and a set of non-negative weights corresponding to the edges, which we write as an adjacency matrix \mathbf{W} ; $W_{ij} > 0$ iff there exists an edge e_{ij} from \mathbf{x}_i to \mathbf{x}_j .

Suppose that of M objectives the individual \mathbf{x}_i is better than individual \mathbf{x}_j for d of them, then we weight the edges between them as $W_{ij} = d/M$ and $W_{ji} = 1 - d/M$. This may be interpreted in terms of a tournament between the individuals, in which an objective is selected at random and the dominating individual *on that objective* wins; the weights are the probabilities that each individual will win such a tournament. Clearly if $\mathbf{x}_i \prec \mathbf{x}_j$ then \mathbf{x}_i will win every tournament with probability 1. This generalises the idea of the *favour* relation proposed by Drechsler et al. [11] in which the winning individual in a tournament between two individuals is the one which is better on more objectives. We define $W_{ii} = 0$ for all i , so that $W_{ij} + W_{ji} = 1$ for all $i \neq j$. Moon and Pullman [12] call matrices with this property *generalised tournament matrices* (GTMs). Figure 1 shows the adjacency matrix of the GUG09 data.

A. Leagues

Moon and Pullman show that a GTM may be partitioned into *leagues* [12]; any player in one league will certainly beat (or dominate) any player in an inferior league. A GTM may be permuted into a normal form so that blocks comprising all ones lie in the top right hand corner. They define a *decomposable matrix* \mathbf{W} as one that can be permuted by a permutation \mathbf{Q} so that

$$\mathbf{Q}^T \mathbf{W} \mathbf{Q} = \begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \quad (2)$$

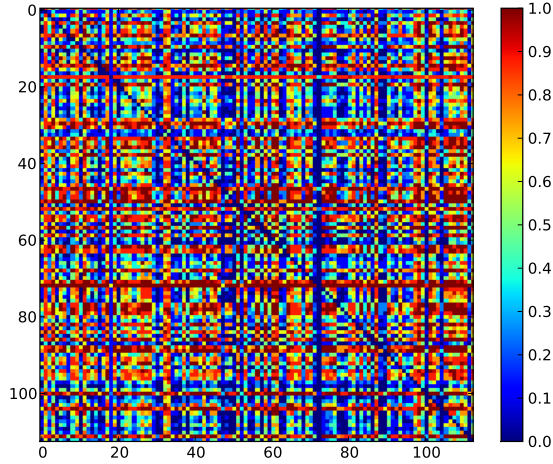


Fig. 1: The adjacency matrix \mathbf{W} of the GUG09 data. Universities are in alphabetical order and the colour of W_{ij} indicates the proportion of objectives for which university i is better than university j .

where \mathbf{A} and \mathbf{B} are square matrices. Then \mathbf{W} is a GTM if there exists a permutation \mathbf{Q} such that

$$\mathbf{Q}^T \mathbf{W} \mathbf{Q} = \begin{bmatrix} \mathbf{W}_1 & \mathbf{1} & \dots & \mathbf{1} \\ \mathbf{0} & \mathbf{W}_2 & & \vdots \\ \vdots & & \ddots & \mathbf{1} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{W}_l \end{bmatrix} \quad (3)$$

where each \mathbf{W}_i is a separate league, and every player in \mathbf{W}_i will beat every player in \mathbf{W}_j for $j > i$. As Moon and Pullman show, a GTM may be put into normal form by permuting the rows and columns so that the row sums $\sum_j W_{ij}$ are in decreasing order, after which the blocks can be read off. The GUG09 data comprises a single league; that is, no university (or group of universities) dominates every other university.

Partitioning the population into leagues does, however, not provide a total ordering of individuals. A player in one league can be said to be better than one in a lower league, but it does not differentiate between members of the same league.

B. Pareto Shells

A finer gradation of individuals within a league is provided by sorting individuals into *Pareto shells* of mutually non-dominating individuals, using the non-dominated sorting procedure that is used in some evolutionary optimisation algorithms [13].

The aim of performing a many-objective optimisation is to locate an estimate of the *Pareto front*, the set of globally optimal trade-offs for the given problem. Each individual in the Pareto front is not dominated, so that members of the Pareto front are mutually non-dominating with respect to each other. More formally, the Pareto front is defined as:

$$\mathcal{P}_0 = \{\mathbf{x} \in X \mid \neg \exists \mathbf{y} (\mathbf{y} \in X \wedge \mathbf{y} \prec \mathbf{x})\}. \quad (4)$$

Here we call \mathcal{P}_0 the zeroth *Pareto shell*. The first Pareto shell is then defined as the set of individuals that are not dominated

when members of \mathcal{P}_0 are removed from the population. Successive shells are defined in the same manner until the entire population has been exhausted; the j th Pareto shell is:

$$\mathcal{P}_j = \{\mathbf{x} \in X'_j \mid \neg \exists \mathbf{y} (\mathbf{y} \in X' \wedge \mathbf{y} \prec \mathbf{x})\} \quad (5)$$

where $X'_j = X \setminus \bigcup_{k=0}^{j-1} \mathcal{P}_k$.

In contrast to leagues, in which *any* individual in a superior league will always beat (dominate) *any* player in an inferior league, all that can be said about shells is that for each individual, \mathbf{x} in an inferior shell \mathcal{P}_j there is, in the next superior shell, *an* individual $\mathbf{x}' \in \mathcal{P}_{j-1}$ such that $\mathbf{x}' \prec \mathbf{x}$.

Pareto shells refine the ordering given by leagues and, because a single shell cannot include individuals from more than one league, each league may be partitioned into shells.

Figure 2 illustrates a new method of displaying the dominance relations in a population. We arrange nodes (individuals) in columns according to their Pareto shell and draw directed edges e_{ij} to indicate that $\mathbf{x}_i \prec \mathbf{x}_j$ for individuals in successive shells, $\mathbf{x}_i \in \mathcal{P}_k$ and $\mathbf{x}_j \in \mathcal{P}_{k+1}$. For clarity we show only the dominance relationships between successive shells. Nonetheless, this diagrammatic representation conveys information about powerful individuals within shells. Thus in \mathcal{P}_0 , Oxford dominates every university in \mathcal{P}_1 ; while in shell \mathcal{P}_1 , King's and in \mathcal{P}_3 , Leeds and Heriot-Watt are powerful individuals, responsible for defining the next most inferior shell.

The Pareto shells furnish a method for imputing missing objectives in a population. In the GUG09 data, a number of universities did not return one of the 8 KPIs, the National Student Survey score. If x_{im} , m th objective of individual \mathbf{x}_i is missing and suppose that the objectives are to be minimised, then we impute the missing value according to the following procedure:

- 1) Assign a very poor value to the missing value:
 $x_{im} := \max_i x_{im}$.
- 2) Sort the population into Pareto shells.
- 3) If $\mathbf{x}_i \in \mathcal{P}_j$, then assign $x_{im} := \max_{\mathbf{x}_k \in \mathcal{P}_j} x_{km}$.

This conservative imputation assigns to the missing value the worst value for the missing objective among the members of the best Pareto shell achieved without the benefit of the imputed value. Thus an individual with missing values cannot be promoted to a shell higher than justified by the data which is present.

Pareto shells, although providing finer discrimination than partitioning into leagues, still provide only a rough banding of individuals; we therefore consider other ordering methods.

C. Preference Ordering

Preference Ordering [14] introduces the *k-preference order* which is based on the idea of finding when an individual remains non-dominated as objectives are removed. An individual is considered efficient of order k if it is not dominated by any other solution considering all the $\binom{M}{k}$ objectives, so that individuals may be ordered according to smallest order at which they are efficient, a small k inferring greater fitness. Since members of inferior shells may not

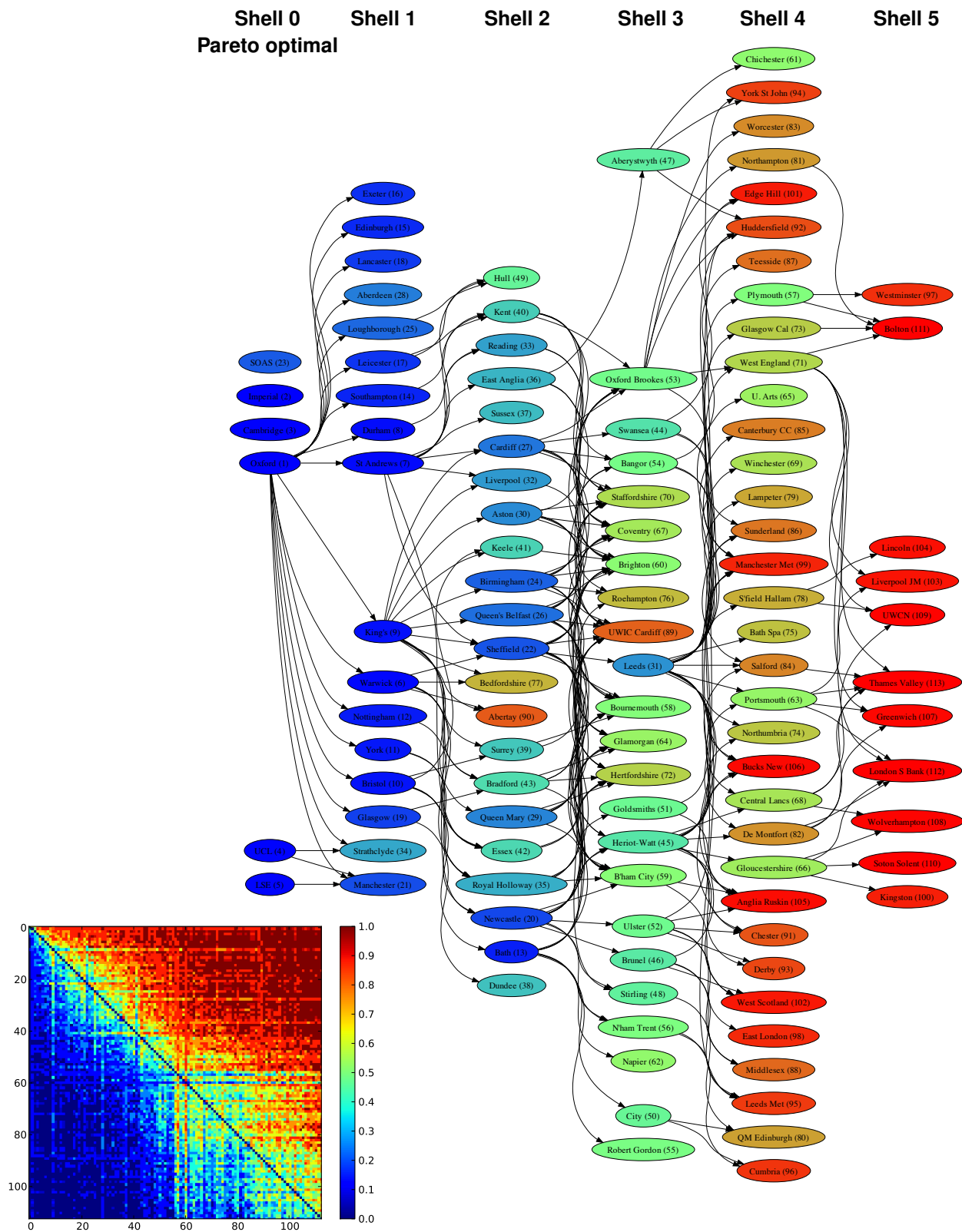


Fig. 2: The GUG09 data ordered by Pareto shells. Colours denote the power index ranking, which is described in section III-H (numerical ranks shown in brackets). Directed edges indicate that $\mathbf{x} \prec \mathbf{x}'$ for universities $\mathbf{x} \in \mathcal{P}_j$ and $\mathbf{x}' \in \mathcal{P}_{j+1}$. For clarity, only those dominance relationships between a shell and the succeeding one are shown. Vertical position within shells is immaterial and is determined by the graph layout algorithm. The insert shows the GUG09 data adjacency matrix ordered by the *average* shell, $\bar{S}(\mathbf{x})$; section III-F.

k	University	ρ	R_i	k	University	ρ	R_i
4	Warwick	0	0	8	Bristol	5	5
6	St Andrews	1	1	8	Manchester	6	6
6	King's	3	3	8	Southampton	7	7
7	Durham	2	2	8	Glasgow	8	8
7	York	4	4	8	Loughborough	9	9
7	Exeter	11	11	8	Leicester	10	10
7	Lancaster	13	13	8	Nottingham	12	15
7	Edinburgh	15	12	8	Aberdeen	14	14
				8	Strathclyde	16	16

TABLE I: Preference order for universities in \mathcal{P}_1 . Columns headed k indicate the efficiency of the corresponding university; columns headed ρ and R_i show the university's rank according to the power index (section III-H) and average rank (section III-D) respectively.

be non-dominated with any combination of the objectives the preference order tends not to discriminate well between individuals in inferior shells. Di Pierro et al. [14] compensate for this by considering each shell at a time, so that the preference order refines the partial order given by shells. Table I shows the efficiency of universities in \mathcal{P}_1 (without regard for universities in other shells). As the table shows, Warwick is clearly the most efficient university in this group, however, the preference order does not discriminate well between less powerful universities.

D. Average Rank

The particular scales of the separate objectives is immaterial in determining whether one individual is better or worse than another, so it is natural to rank individuals based on each objective. Consider the graph G_m , which describes the m^{th} objective with an associated adjacency matrix \mathbf{W}^m , a GTM with 0 and 1 entries only. Then the rank R_i^m of an individual i on objective m is

$$R_i^m = N - 1 - \sum_j W_{ij}^m. \quad (6)$$

High performing individuals are assigned numerically low ranks, with 0 being the best rank, while the worst individual has a rank of $N - 1$ because it is bettered by all others.

Ranks for separate objectives can then be averaged in order to produce an overall rank for each individual [15]. The average rank of individual i over all objectives is:

$$R_i = \frac{1}{M} \sum_{m=1}^M R_i^m. \quad (7)$$

An extension to the average rank is to weight the objectives, however since it is unclear how to choose the weights *a priori*, we treat all objectives with equal weight.

Corne and Knowles found that the best performing method of selecting individuals to be mutated at the next generation of a many-objective search was a variant of the average rank, which took into account only those individuals in the current Pareto set [1].

We note however, that the average rank cannot be used to order a two-objective non-dominated set because the sum of the ranks for any individual is $N - 1$.

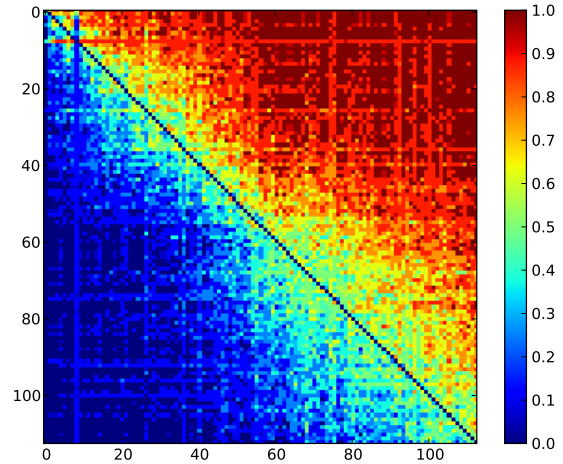


Fig. 3: The GUG09 data adjacency matrix ordered by the outflow σ^{out} or average rank.

E. Outflow

The outflow ranking method [16] orders nodes according to the outflow at that node. The outflow simply generalises the out-degree, and is the sum of the weights of leaving a node; in terms of the adjacency matrix,

$$\sigma_i^{\text{out}} = \sum_j W_{ij}. \quad (8)$$

Consequently node \mathbf{x}_i is ranked at least as high as node \mathbf{x}_j iff $\sigma_i^{\text{out}} \geq \sigma_j^{\text{out}}$. As van den Brink and Gilles note, outflow ranking generalises ranking by out-degree and ranking of tournaments by the Copeland score [16]. They also provide five axioms (anonymity, outflow monotonicity, positive responsiveness and strong order preservation) which characterise outflow ranking.

When there are several objectives each with adjacency matrix \mathbf{W}^m , then the outflow rank is equivalent to the average rank. This is because the overall adjacency matrix is an average of the \mathbf{W}^m :

$$\mathbf{W} = \frac{1}{M} \sum_{m=1}^M \mathbf{W}^m. \quad (9)$$

If $\mathbf{1}$ is the vector of 1s, the vector of outflows may be written:

$$\boldsymbol{\sigma}^{\text{out}} = \mathbf{W}\mathbf{1} = \frac{1}{M} \sum_{m=1}^M \mathbf{W}^m \mathbf{1} = (N - 1)\mathbf{1} - \frac{1}{M} \sum_{m=1}^M \mathbf{R}^m \quad (10)$$

where \mathbf{R}^m is the vector of ranks for the m^{th} objective (6). In addition, [1] observe that another ordering technique, *winning score*, is also equivalent to these two techniques.

The adjacency matrix for the GUG09 data, ordered by the average rank, is shown in Figure 3. In general universities with a average high probability of dominating others (those rows coloured largely red) have been grouped together towards to top right hand corner, while those which have a high average probability of being dominated are grouped towards the bottom left hand corner. An ideal permutation

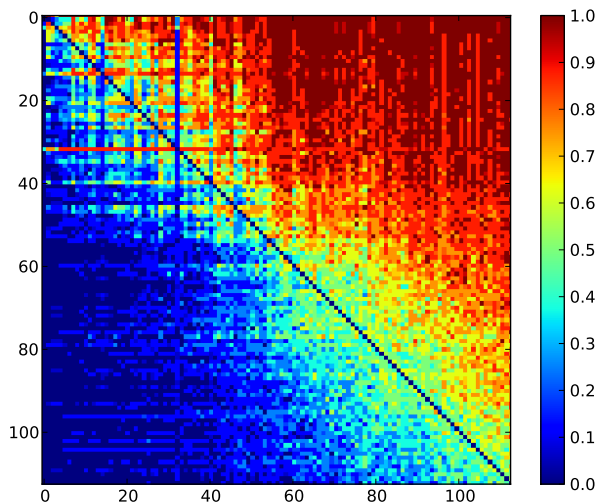


Fig. 4: The GUG09 data adjacency matrix ordered by the stationary distribution.

would group the red colour in the top right hand corner, with blue in the bottom left hand corner.

The outflow or average rank has many attractive features: it is easy to compute, easy to understand, and, as described previously, is the same measure used by Moon and Pullman to partition a graph into leagues [12]. Nonetheless it is a *local* method that uses only partial information about the graph (the sum of the weights leaving a node). In addition, as van den Brink and Gilles [16] mention, ranking by outflow is equivalent to ranking countries by total exports, rather than, for example, the trade balances which would be the result of ranking by the difference between outflow and inflow.

F. Average shell

The preference order ranks individuals by how robust they are as objectives are removed, but as noted above tends not to be very discriminative. A related idea is to enquire in which Pareto shell an individual resides as objectives are removed. Since with M objectives there are $\binom{M}{k}$ combinations of k objectives taken from the original M , we average over these combinations and denote the result by $\bar{S}_k(\mathbf{x})$. Note that $\bar{S}_M(\mathbf{x})$ is just the Pareto shell of the individual using all objectives whereas, providing that objective values are all distinct, $\bar{S}_1(\mathbf{x})$ is the average rank of \mathbf{x} . An individual may then be assessed by the average shell, weighted by the dimension of the reduced objective space:

$$\bar{S}(\mathbf{x}) = \frac{2}{M(M+1)} \sum_{k=1}^M k \bar{S}_k(\mathbf{x}) \quad (11)$$

The insert to figure 2 shows the GUG09 data adjacency matrix ordered by the $\bar{S}(\mathbf{x})$. Although the ordering groups generally similar individuals together, the outflow and power index rankings (which we describe below) appear preferable.

G. Stationary distribution

It is straightforward to define a random walk on the graph describing the multi-objective population. Consider a random

walker who at each time step jumps from the current node to one of the node's neighbours with the probability of a transition from node i to node j of $p_{ij} = W_{ij} / \sum_k W_{ik}$. The transition probabilities are therefore summarised in the matrix

$$\mathbf{P} = \mathbf{D}^{-1} \mathbf{W} \quad (12)$$

where \mathbf{D} is the diagonal matrix of row sums of \mathbf{W} which may be recognised as $\boldsymbol{\sigma}^{\text{out}}$, the vector of outflows (10). If \mathbf{p}_{t-1} is the (row) vector describing the probability of finding the walker at each node at time $t-1$, then the probability distribution at t is $\mathbf{p}_t = \mathbf{p}_{t-1} \mathbf{P}$. A graph is strongly connected when, for every two nodes \mathbf{x}_i and \mathbf{x}_j , there is a sequence of directed edges connecting \mathbf{x}_i to \mathbf{x}_j . This is true for nondecomposable GTMs in the Moon and Pullman sense [12] (i.e., GTMs representing a single league). The Perron-Frobenius theorem [17] implies that the transition probability matrix of a strongly connected graph has an eigenvalue 1 and unique left eigenvector $\boldsymbol{\pi}$ with $\pi_i > 0$ for all i which defines the stationary distribution of a random walk on G ; that is: $\boldsymbol{\pi} = \mathbf{P} \boldsymbol{\pi}$. The stationary probability of finding a walker at a node has been used in ranking applications, notably in the PageRank algorithm [18] in which frequently visited web pages are ranked highly. For ordering multi-objective populations, directed edges point *away* from powerful individuals so that powerful individuals are expected to be those with a low stationary probability.

The GUG09 data adjacency matrix ordered by the stationary distribution is shown in figure 4. However, although the broad ordering is as expected, there are some prominent anomalies. Oxford is ranked at 0, but King's is ranked at 1 although it is in \mathcal{P}_1 , because it dominates many universities in \mathcal{P}_2 so that a random walker arriving at King's will easily be transported away from it. Conversely, Cambridge in \mathcal{P}_0 is ranked at 32. This is due to it not returning an NSS score so that the NSS value used is the worst value for \mathcal{P}_0 . Cambridge is thus on the 'edge' of \mathcal{P}_0 and does not dominate many others which means that probability mass arriving at Cambridge is not easily transported away from it, leading to a relatively high probability in the stationary distribution and consequent poor rank.

H. Power Index

The *long path method* can be traced back to Wei [19] and Kendall [20]; it was called the *power index* by Berge [21] and has been used to rank single objective scores (e.g., [22]). The intuition behind the power index lies with the idea that the individuals in the population should be ranked not by the number of individuals that they dominate, but rather by the quality of those individuals. It is a global method and is defined in terms of the right eigenvector of \mathbf{W} . This is reached by considering the sequence:

$$\mathbf{u}^t = \mathbf{W} \mathbf{u}^{t-1} \quad (13)$$

for $t = 1, 2, \dots$, starting with $\mathbf{u}^0 = \mathbf{1}$. This means that \mathbf{u}^1 is the vector of average ranks or outflows, $\boldsymbol{\sigma}^{\text{out}}$; \mathbf{u}^2 is the

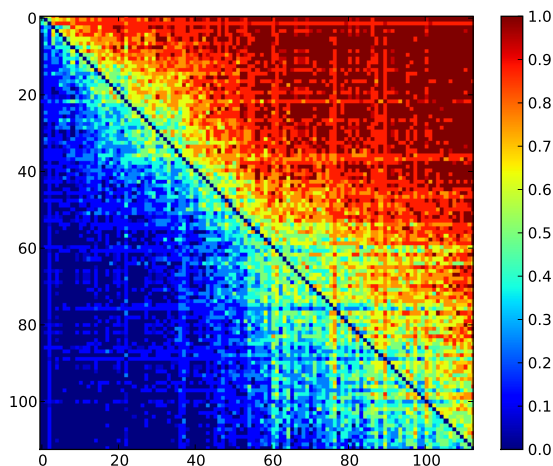


Fig. 5: The GUG09 data adjacency matrix ordered by the power index.

vector that assigns to \mathbf{x}_i the scores of all other individuals in proportion to the weight W_{ij} , that is $u_i^2 = \sum_j W_{ij} u_j^1$. The limit of this procedure is

$$\mathbf{u} = \lim_{t \rightarrow \infty} \frac{\mathbf{u}^t}{\sum_i u_i^t}. \quad (14)$$

This limit exists when G is strongly connected, and is the eigenvector of \mathbf{W} corresponding to the largest eigenvector; that is \mathbf{u} solves

$$\mathbf{W}\mathbf{u} = \lambda\mathbf{u}. \quad (15)$$

The Perron-Frobenius [17] theorem assures that \mathbf{W} has a unique positive eigenvector. The population is then ordered by the values of \mathbf{u} ; if u_i is the i th largest element of \mathbf{u} then \mathbf{x}_j is assigned rank i .

Figure 5 shows the adjacency matrix for the GUG09 data ordered by power index and nodes in Figure 2 are coloured according to their power index rank. Although the ordering given by power index and the average rank or outflow is broadly similar, figures 3, 5 and 6 show that they are indeed different because the power index of \mathbf{x} accounts for the power of the universities which, on average, it would beat when two objectives are selected at random. Note, however, that neither the outflow nor the power index orders respect the Pareto shell ordering, in that an $\mathbf{x} \in \mathcal{P}_j$ may be ranked better than $\mathbf{x}' \in \mathcal{P}_{j-1}$ and vice versa. Clear examples of this shown in Figure 2 are SOAS in \mathcal{P}_0 which is ranked at 23 by the power index, together with Bedfordshire and Abertay, both in \mathcal{P}_2 , but ranked at 77 and 90 respectively largely, as the figure shows, because they do not dominate any universities in \mathcal{P}_3 . We may infer that SOAS, Bedfordshire and Abertay are situated on the periphery of their shells because they do not dominate others. Conversely, those universities that dominate many others are ranked better, for example, Oxford (0), King's (9), Bath (13), Sheffield (22) and Leeds (31) all define the following shell.

Clearly the power index provides a good deal of additional information on the structure of a multi-objective population

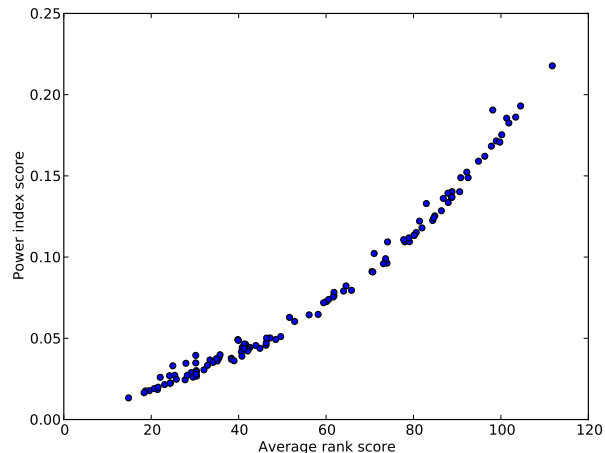


Fig. 6: Power index score versus outflow score for the GUG09 data. High values indicate best performance in both cases.

when it is comprised of a single or many Pareto shells. We note however, that like the average rank, it does not differentiate between members of a two-objective non-dominated set because in that case $W_{ij} = 1/2 \quad \forall i \neq j$ and the power index of each member is $1/\sqrt{N}$.

IV. SERIATION OF HEATMAPS

The left panel of figure 7 shows a heat map visualisation [9] of the 113 universities and 8 objectives of the GUG09 data. Here the colour scale shows the rank R_i^m of each university on each objective (cf. eq. (6)), which allows objectives measured on disparate scales and in disparate units to be straightforwardly compared. Although the heatmap conveys a good deal of information about the multi-objective population it can be unduly confused because the orders in which individuals are drawn on one axis and the order in which objectives are drawn on the other axis are arbitrary. The average rank or power index can usefully be used to order the individuals and in this section we *seriate* [23] the objectives to group similar objectives together.

We quantify the similarity between objectives m and n as

$$A_{mn} = 1 - \frac{1}{N^2} \sum_{i=1}^N (R_i^m - R_i^n)^2 \quad (16)$$

so that competing objectives which on average rank individuals differently have small A_{mn} , whereas objectives that rank individuals similarly have A_{mn} close to 1. Then if π_n is a permutation of the n th objective, we seek to minimise:

$$g(\pi) = \sum_{m,n} A_{mn} (\pi_m - \pi_n)^2. \quad (17)$$

In general this is NP-hard because the permutation is discrete. A way of obtaining an approximation is to relax the permutation π_n to a continuous variable z_n and instead minimise

$$h(\mathbf{z}) = \sum_{m,n} A_{mn} (z_m - z_n)^2 \quad (18)$$

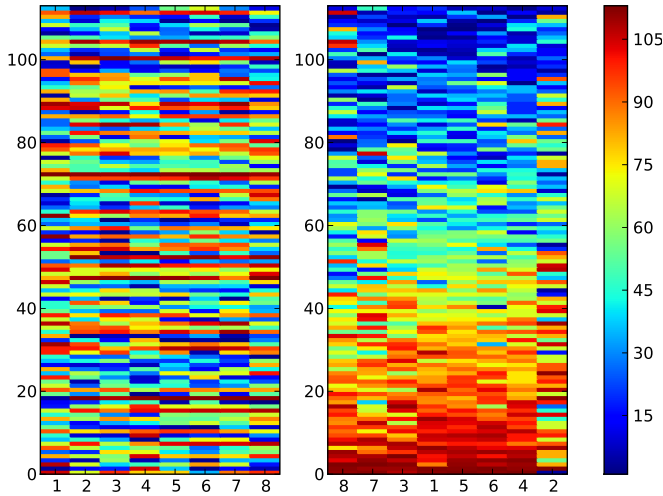


Fig. 7: *Left*: Heatmap representation of the GUG09 data. Objectives correspond to columns and universities, in alphabetical order, correspond to rows. *Right*: Heatmap after seriation to place like individuals and objectives close to each other.

subject to the constraints that (a) $\sum_n z_n = 0$, which takes care of the problem that adding a constant to all z_n does not change the order; and (b) $\sum_n z_n^2 = 1$, which rules out the trivial solution that all $z_n = 0$. With a solution to (18) on hand, the objectives are arranged in increasing order of the minimising \mathbf{z} . The solution to the constrained problem (18) may be written as the minimisation of $\mathbf{z}^T \mathbf{L} \mathbf{z}$ where \mathbf{L} is the *graph Laplacian* [23] of \mathbf{A} which is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$, with \mathbf{D} the diagonal matrix whose diagonal elements are $D_{nn} = \sum_m A_{nm}$. Then the solution to the constrained minimisation problem is the Fiedler vector of \mathbf{L} , namely the vector associated with the smallest non-zero eigenvalue of \mathbf{L} .

Figure 7 shows the GUG09 data with the objectives seriated to place similar objectives together. In addition we swapped the roles of objectives (columns) and individuals (rows) to seriate the heatmap a second time to place similar *individuals* together. The value of $g(\pi)$ for the unordered heatmap was 2.65×10^9 which dropped to 2.23×10^4 after the dual seriation. The seriated heatmap is more readily interpreted than the unordered map. By placing similar objectives together the extent to which objectives compete or are correlated becomes apparent; in the figure the 2nd objective (research quality) and the 7th objective (proportion of ‘good’ honours degrees) are seen to be less well correlated with other objectives.

V. CONCLUSION

The introduction of the probability of one individual beating another in a tournament on a randomly selected objective provides a novel visualisation method, drawing on tools from graph theory and generalised tournaments, with which we can view a many-objective population. We have examined a range of methods for ordering multi-objective populations. Among these, the average rank (which we

showed to be equivalent to the outflow rank) and the power index appear to be superior. The novel visualisation of a many-objective population by Pareto shells is enhanced by ranking information from the power index and we showed how to seriate objectives visualisation in a heatmap.

Although we have used these measures for ranking individuals for visualisation and decision making, we anticipate that these same ranking methods can be used for selection in evolutionary optimisation schemes.

REFERENCES

- [1] D. Corne and J. Knowles, “Techniques for Highly Multiobjective Optimisation: Some Nondominated Points are Better than Others,” in *Genetic and Evolutionary Computation Conference*, London, UK, 2007, pp. 773–780.
- [2] A. López Jaimes and C. A. Coello Coello, “Study of preference relations in many-objective optimization,” in *11th Annual conference on genetic and evolutionary computation*, 2009, pp. 611–618.
- [3] J. O’Leary, *The Times Good University Guide 2009*. Times Books, 2009.
- [4] C. Bishop, M. Svensén, and C. Williams, “GTM: The generative topographic mapping,” *Neural Computation*, pp. 215 – 235, 1998.
- [5] T. Kohonen, *Self-organising Maps*. Springer, 1995.
- [6] J. E. Fieldsend and R. M. Everson, “Visualisation of multi-class ROC surfaces,” in *Proceedings of the ICML 2005 workshop on ROC Analysis in Machine Learning*, 2005, pp. 49–56.
- [7] J. J. Valdés and A. J. Barton, “Visualizing high dimensional objective spaces for multi-objective optimization: A virtual reality approach,” in *IEEE Congress on Evolutionary Computation*, 2007, pp. 4199–4206.
- [8] C. M. Fonseca and P. J. Fleming, “Genetic algorithms for multi-objective optimization: Formulation, discussion and generalization,” in *Proceedings of the Fifth International Conference on Genetic Algorithms*. Morgan Kaufman, 1993, pp. 416–423.
- [9] A. Pryke, S. Mostaghim, and A. Nazemi, “Heatmap visualization of population based multi objective algorithms,” in *EMO*, 2006, pp. 361–375.
- [10] A. Nazemi, A. H. Chan, and X. Yao, “Selecting representative parameters of rainfall-runoff models using multi-objective calibration results and a fuzzy clustering algorithm,” in *BHS 10th National Hydrology Symposium, Exeter*, 2008.
- [11] N. Drechsler, R. Drechsler, and B. Becker, “Multi-objective optimisation based on relation favour,” *Lecture Notes in Computer Science*, vol. 1993, pp. 154–166, 2001.
- [12] J. Moon and N. Pullman, “On generalized tournament matrices,” *SIAM Review*, vol. 12, no. 3, pp. 384–399, 1970.
- [13] N. Srinivas and K. Deb, “Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms,” *Evolutionary Computation*, vol. 2, no. 3, pp. 221–248, Fall 1994.
- [14] F. di Piero, S. T. Khu, and D. Savic, “An investigation on preference ordering ranking scheme in multiobjective evolutionary optimization,” *IEEE Trans. on Evolutionary Comp.*, vol. 11, no. 1, pp. 17–45, 2007.
- [15] P. J. Bentley and J. P. Wakefield, “Finding acceptable solutions in the Pareto-optimal range using multiobjective genetic algorithms,” in *Soft Computing in Engineering Design and Manufacturing*, P. Chawdhry, R. Roy, and R. Pant, Eds. Springer-Verlag, 1998, pp. 231–240.
- [16] R. van den Brink and R. P. Gilles, “The outflow ranking method for weighted directed graphs,” *European Journal of Operational Research*, vol. 193, pp. 484–491, 2009.
- [17] R. A. Horn and C. R. Johnson, *Matrix Analysis*. CUP, 1990.
- [18] L. Page, S. Brin, R. Motwani, and T. Winograd, “The PageRank Citation Ranking: Bringing Order to the Web,” Stanford Digital Library Technologies Project, Stanford University, Tech. Rep., 1998.
- [19] T. Wei, “The algebraic foundation of ranking theory,” Ph.D. dissertation, Cambridge University, Cambridge, UK, 1952.
- [20] M. Kendall, “Further contributions to the theory of paired comparisons,” *Biometrics*, vol. 11, pp. 43–62, 1955.
- [21] C. Berge, *The theory of graphs*. London: Methuen, 1962.
- [22] J. P. Keener, “The Perron-Frobenius theorem and the ranking of football teams,” *SIAM Rev.*, vol. 35, no. 1, pp. 80–93, 1993.
- [23] J. E. Atkins, E. G. Boman, and B. Hendrikson, “A spectral algorithm for seriation and the consecutive ones problem,” *SIAM Journal on Computing*, vol. 28, no. 1, pp. 297–310, 1998.