Visualization in Multiobjective Optimization

Bogdan Filipič and Tea Tušar
Tutorial at GECCO ’20

Department of Intelligent Systems
Jožef Stefan Institute
Ljubljana, Slovenia

Instructors

Bogdan Filipič is a senior researcher and head of Computational Intelligence Group at the Department of Intelligent Systems of the Jožef Stefan Institute, Ljubljana, Slovenia, and associate professor of Computer Science at the Jožef Stefan International Postgraduate School. His research interests are in artificial intelligence, stochastic optimization, and evolutionary computation.

Tea Tušar is a research fellow at the Department of Intelligent Systems of the Jožef Stefan Institute in Ljubljana, Slovenia. Her research interests include evolutionary algorithms for singleobjective and multiobjective optimization with emphasis on visualizing and benchmarking their results and applying them to real-world problems.

Final version

These slides as well as all the approximation sets used in this tutorial are available at http://dis.ijs.si/tea/research.htm

Contents

Introduction
A taxonomy of visualization methods
Visualizing approximation sets
Visualizing single approximation sets
Visualizing repeated approximation sets
Visualizing multiobjective landscapes
Summary
References
Introduction

Multiobjective optimization problem
Minimize

\[ f: X \rightarrow F \]

\[ f: (x_1, \ldots, x_n) \mapsto (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)) \]

- \( X \) is an \( n \)-dimensional decision space (or search space)
- \( F \subseteq \mathbb{R}^m \) is an \( m \)-dimensional objective space (\( m \geq 2 \))

Conflicting objectives \( \rightarrow \) a set of optimal solutions
- Pareto set in the decision space
- Pareto front in the objective space

Visualization in multiobjective optimization

- Solution sets in the decision or objective space (or both)
- Multiobjective landscapes—objective values in the decision space

Visualization of solution sets useful for:
- Analysis of solutions and solution sets
- Decision support in interactive optimization
- Analysis of algorithm performance

Visualization of multiobjective landscapes useful for:
- Revealing problem properties and difficulties
- Identifying basins of attraction of local optima

Visualization of solution sets in the decision space
- Problem-specific
- If \( X \subseteq \mathbb{R}^m \), any method for visualizing multidimensional solutions can be used
- Not the focus of this tutorial

Visualization of solution sets in the objective space
- Interested in sets of mutually nondominated solutions called approximation sets
- Different from ordinary multidimensional solution sets

Visualization of multiobjective landscapes
- Important for problem understanding, but few approaches exist
Challenges of visualizing solution sets in the objective space

- High dimension and large number of solutions
- Limitations of computing and displaying technologies
- Cognitive limitations

Visualization can be hard even in 2-D

Stochastic optimization algorithms

- Single run → single approximation set
- Multiple runs → multiple approximation sets

The Empirical Attainment Function (EAF) [23] or the Average Runtime Attainment Function (aRTA) [10] can be used in such cases

This tutorial does not cover

- Visualization of a few solutions for decision making purposes (see [41])
- Visualization of solution sets in the decision space
- General multidimensional visualization methods not previously used on approximation sets

This tutorial covers

- Visualization of solution sets in the objective space
  - Single approximation sets [2]
  - Repeated approximation sets [3, 10]
- Visualization of multiobjective landscapes

A taxonomy of visualization methods
A taxonomy of visualization methods [1]

Visualizing approximation sets

Visualizing single approximation sets

Methodology

Evaluating and comparing visualization methods

- No established methodology for evaluating or comparing visualization methods
- Propose benchmark approximation sets (analog to benchmark problems in multiobjective optimization)
- Visualize the sets using different methods
- Observe which set properties are distinguishable after visualization
- Only applicable to methods showing individual solutions or individual solution properties
Benchmark approximation sets

Three different sets that can be instantiated in any dimension
- Spherical with a clustered distribution of solutions (more at the corners and less at the center)
- Linear with a uniform distribution of solutions
- Knee-shaped with an even distribution of solutions

Size of each set
- 2-D: 50 solutions
- 3-D: 500 solutions
- 4-D: 500 solutions

An additional set with redundant objectives
- Adapted from [21]
- 12 objectives
- Can be instantiated for any number of $10^n$ solutions (here 100)

Desired properties of visualization methods

Demonstration on the 4-D spherical, linear and knee-shaped sets
- Preservation of the
  - Dominance relation between solutions
  - Front shape
  - Objective range
  - Distribution of solutions
- Robustness
- Handling of large sets
- Simultaneous visualization of multiple sets
- Scalability in number of objectives
- Simplicity

Demonstration on the 12-D approximation set
- Showing relations between objectives
Visualizing single approximation sets

- Scatter Plot Matrix
- Bubble Chart
- Parallel Coordinates
- Radar Chart
- Chord Diagram
- Heat Map
- Interactive Decision Map
- Narseh Method

Showing original objective values
- Scatter Plot Matrix
- Bubble Chart
- Parallel Coordinates
- Radar Chart
- Chord Diagram
- Heat Map
- Interactive Decision Map
- Narseh Method

Showing transformed objective values
- Scatter Plot Matrix
- Bubble Chart
- Parallel Coordinates
- Radar Chart
- Chord Diagram
- Heat Map
- Interactive Decision Map
- Narseh Method

Scatter plot matrix

Most often
- Scatter plot in a 2-D space
- Matrix of all possible combinations of objectives
  \[ m \text{ objectives} \rightarrow \frac{m(m-1)}{2} \text{ different combinations} \]

Alternatively
- Scatter plot in a 3-D space
  \[ m \text{ objectives} \rightarrow \frac{m(m-1)(m-2)}{6} \text{ different combinations} \]
Parallel coordinates

- \( m \) objectives \( \rightarrow m \) parallel axes
- Solution represented as a polyline with vertices on the axes
- Position of each vertex corresponds to that objective value
- No loss of information

Parallel coordinates

Spherical

Linear

Knee-shaped

Interactive decision maps

The Edgeworth-Pareto hull (EPH) of an approximation set \( A \) contains all points in the objective space that are weakly dominated by any solution in \( A \).

Interactive decision maps

- Visualize the surface of the EPH, not the actual approximation set
- Plot a number of axis-aligned sampling surfaces of the EPH
- Color used to denote third objective
- Fixed value of the forth objective
Interactive decision maps

Visualizing single approximation sets

Radial coordinate visualization

Also called RadViz

- Inspired from physics
- Objectives treated as anchors, equally spaced around the circumference of a unit circle
- Solutions attached to anchors with 'springs'
- Spring stiffness proportional to the objective value
- Solution placed where the spring forces are in equilibrium

Radial coordinate visualization

- Spherical
- Linear
- Knee-shaped
Prosections

- Visualize only part of the objective space
- Dimensionality reduction by projection of solutions in a section
- Need to choose prosection plane, angle and section width

Level diagrams

- $m$ objectives $\rightarrow m$ diagrams
- Plot solutions with objective $f_i$ on the $x$ axis and distance to the ideal point on the $y$ axis
Projections

Spherical and Linear

Hyper-space diagonal counting

- Inspired by Cantor’s proof that shows \(|\mathbb{N}| = |\mathbb{N}^2| = |\mathbb{N}^3| \ldots\)
- Discretize each objective (choose a number of bins)
- In the 4-D case
  - Enumerate the bins for objectives \(f_1\) and \(f_4\)
  - Enumerate the bins for objectives \(f_2\) and \(f_3\)
  - Plot the number of solutions in each pair of bins

Visualizing single approximation sets

Set properties (visualizing solutions dependently from each other)

Not optimization based

Optimization based

Hyper-space diagonal counting

| preservation of the |
|---|---|---|---|---|---|---|---|
| dominance relation | front shape | objective range | distribution of solutions | robustness | handling of large sets | simultaneous visualization | scalability | simplicity |
| ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ |

| preservation of the |
|---|---|---|---|---|---|---|---|
| dominance relation | front shape | objective range | distribution of solutions | robustness | handling of large sets | simultaneous visualization | scalability | simplicity |
| ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✓ |
Visualizing single approximation sets

![Diagram](image)

**Sammon mapping**

- A non-linear mapping
- Aims to preserve distances between solutions
  - $d_{ij}$ distance between solutions $x_i$ and $x_j$ in the objective space
  - $d_{ij}^*$ distance between solutions $x_i$ and $x_j$ in the visualized space
- Stress function to be minimized
  $$S = \sum_i \sum_{j\geq i} (d_{ij}^* - d_{ij})^2$$
- Minimization by gradient descent or other (iterative) methods

**Distance- and dominance-based mappings**

Both mappings

- Use nondominated sorting to split solutions to fronts
- Project solutions onto the circumference of circles (with circle radius proportional to front number)

**Distance-based mapping**

- Tries to preserve closeness of solutions
- Two solutions are very close if their relations to other solutions are mostly equal

**Dominance-based mapping**

- Aims at preserving dominance relations among solutions
- All $x \prec y$ can be shown correctly
- Tries to minimize cases where $x \not\prec y$ is not shown correctly
Distance- and dominance-based mappings

Distance-based mapping Dominance-based mapping

0 0.5 1 1.5 2 0 0.5 1 1.5 2
Linear Spherical Linear Spherical

Table:

<table>
<thead>
<tr>
<th>Dominance relation</th>
<th>Preservation of the front shape</th>
<th>Objective range</th>
<th>Handling of large sets</th>
<th>Simultaneous visualization</th>
<th>Scalability</th>
<th>Simplicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Self-organizing maps

- Self-organizing maps (SOMs) are neural networks
- Nearby solutions are mapped to nearby neurons in the SOM
- A SOM can be visualized using the unified distance matrix
- Distance between adjacent neurons is denoted with color
  - Similar neurons → light color
  - Different neurons (cluster boundaries) → dark color
Aggregation trees

- Binary trees that show relationships between objectives
- Iterative clustering of objectives based on their harmony
- Computation of different types of conflict
- Percentages quantify the conflict between objectives
- Colors used to show type of conflict
  - global conflict (black)
  - local conflict on 'good' values (red)
  - local conflict on 'bad' values (blue)
- Can be used to sort objectives in other representations (parallel coordinates, radial charts, heat maps)

Visualizing approximation sets

Visualizing repeated approximation sets
Visualizing repeated approximation sets

Showing performance at a time with the Empirical Attainment Function (EAF) [23]

Empirical attainment function

Goal-attainment
- Approximation set A
- A point in the objective space $\mathbf{z}$ is attained by $A$ when $\mathbf{z}$ is weakly dominated by at least one solution from $A$

Empirical attainment function

EAF values [23]
- Algorithm $A$, approximation sets $A_1, A_2, \ldots, A_r$
- EAF of $\mathbf{z}$ is the frequency of attaining $\mathbf{z}$ by $A_1, A_2, \ldots, A_r$
- Summary (or $k\%$-) attainment surfaces

Empirical attainment function

Differences in EAF values [36]
- Algorithm $A$, approximation sets $A_1, A_2, \ldots, A_r$
- Algorithm $B$, approximation sets $B_1, B_2, \ldots, B_r$
- Visualize differences between EAF values

Empirical attainment function

- Visualization with line plots and heat maps
Visualization of 3-D EAF

Need to compute and visualize a large number (over 10,000) of cuboids

**Exact case**
- EAF values: Slicing [3], Visualization of facets [13, 24]
- EAF differences: Slicing, Maximum intensity projection [56, 3]

**Approximated case**
- EAF values: Grid-based sampling [30], Slicing, Direct volume rendering [14, 3]
- EAF differences: Slicing, Maximum intensity projection, Direct volume rendering

Benchmark approximation sets

Two groups of spherical approximation sets
- 5 spherical approximation sets with a clustered distribution of solutions (different radii, 100 solutions in each)
- 5 spherical approximation sets with a uniform distribution of solutions (different radii, 100 solutions in each)

Exact 3-D EAF values and differences

**Slicing**
- Visualize cuboids intersecting the slicing plane
- Need to choose coordinate and angle

Clustered spherical

Uniform spherical

Exact 3-D EAF values and differences

**Slicing**

Clustered

Uniform

Difference

\[ \varphi = 5^\circ \]

Clustered

Uniform

Difference

\[ \varphi = 45^\circ \]
Approximated attainment surfaces

Grid-based sampling
Repeat for all $f_i f_j, i < j$ (i.e. $f_1 f_2, f_1 f_3$ and $f_2 f_3$):
- Construct a $k \times k$ grid on the plane $f_i f_j$
- Compute intersections between the attainment surface and the axis-aligned lines on the grid

Clusters

Median attainment surfaces

Visualizing repeated approximation sets

Showing performance over time with the Average Runtime Attainment Function (aRTA) [10]

Average Runtime Attainment Function

aRTA value
- Algorithm $\mathcal{A}$ run $r$ times
- All solutions that are nondominated at creation are recorded
- aRTA($z$) is the average number of evaluations needed to attain $z$

aRTA ratio
- Algorithms $\mathcal{A}$ and $\mathcal{B}$
- Visualize ratio between aRTA($z$) values for $\mathcal{A}$ and $\mathcal{B}$

Benchmark approximation sets

Two groups of sets mimicking convergence to a spherical front
- 5 sets mimicking logarithmic convergence to a spherical front with a clustered distribution (100 solutions each)
- 5 sets mimicking linear convergence to a spherical front with a uniform distribution (100 solutions each)
Average Runtime Attainment Function

Grid-based sampling

Clustered with logarithmic convergence

Uniform with linear convergence

Ratio

Visualizing multiobjective landscapes

Visualizing problem landscapes

- Single objective: visualize objective values in the decision space
- Multiple objectives: ?

Benchmark problems

The **bbob-biobj** test suite [11]

- Each bi-objective function constructed as the combination of two single-objective **bbob** functions
- Problems scalable in the number of decision variables
- Known single-objective optima, but not the Pareto set (or front)
- Included in the COCO platform

(https://github.com/numbbo/coco)
Benchmark problems with 2-D and 5-D decision spaces

Three bbob-biobj benchmark problems

- Double sphere problem \( F_1 = (f_1, f_1) \) in 2-D and 5-D, instance 1
- Sphere-Gallagher problem \( F_{10} = (f_1, f_{21}) \) in 2-D and 5-D, instance 1
- Double Gallagher problem \( F_{55} = (f_{21}, f_{21}) \) in 2-D and 5-D, instance 1

*Gallagher = Gallagher’s Gaussian 101-me Peaks Function

Visualizing multiobjective landscapes

Level sets

- Curves connecting points with the same value
- Orange = first objective, blue = second objective
- Demonstration on the 2-D benchmark problems

Line walks

- Equidistant sampling of the decision space along a line
- The line does not have to be parallel to an axis
- Not constrained by the decision space dimension
- Two display options
  - Show resulting values for each objective separately
  - Show resulting values in the objective space
- Demonstration on the 5-D benchmark problems
Line walks

Double sphere problem in 5-D

Decision space
- reference set (2017 of 2247057 points)
- cuts through single optima
- cut through both optima
- two random directions
- random cut in plane through optima

Objective space
- reference set (1999 of 2247057 points)
- cuts through single optima
- cut through both optima
- two random directions
- random cut in plane through optima

Sphere-Gallagher problem in 5-D

Decision space
- reference set (838 of 2377335 points)
- cuts through single optima
- cut through both optima
- two random directions
- random cut in plane through optima

Objective space
- reference set (2002 of 2377335 points)
- cuts through single optima
- cut through both optima
- two random directions
- random cut in plane through optima

Visualizing multiobjective landscapes

Double Gallagher problem in 5-D

Decision space
- reference set (363 of 928478 points)
- cuts through single optima
- cut through both optima
- two random directions
- random cut in plane through optima

Objective space
- reference set (11860 of 928478 points)
- cuts through single optima
- cut through both optima
- two random directions
- random cut in plane through optima

Showing transformed objective values
- Decision space approximated with a grid of points
- Show value using color (contours or the third dimension)
- Suitable only for 2-D decision spaces
Visualizing dominance ranks

- Discretized decision space (1000 × 1000 grid)
- Rank = number of grid points that dominate the current point
- All nondominated points have a rank of zero
- Visualize normalized ranks in logarithmic scale

Visualizing local dominance

- Discretized decision space (1000 × 1000 grid)
- Moore neighborhood = eight surrounding points
- Compute three different kinds of regions
  - Green Locally dominance-neutral regions
    - Points that are mutually nondominated with all their neighbors
    - Not equal to local Pareto sets
  - Pink Basins of attraction
  - White Boundary regions
- Can take a long time to compute
Visualizing cumulative gradients

- Discretized decision space (1000 \times 1000 grid)
- Compute the bi-objective gradient for all grid points
  \[ \mathbf{v} = \frac{\mathbf{v}_1}{||\mathbf{v}_1||} + \frac{\mathbf{v}_2}{||\mathbf{v}_2||} \]
- From a grid point, follow the path in the direction of the bi-objective gradient
- Sum all bi-objective gradient values along the path
- Visualize cumulative gradients in logarithmic scale

Global vs. local information

- Sphere-Gallagher problem

Visualizing multiobjective landscapes

How to handle such visualization when \( n > 2 \)?

Level sets, dominance ranks, local dominance and cumulative gradients
- Require cuts through the decision space (cf. slicing)
- Challenging to compute and interpret these methods in \( n \)-D

Line walks
- A useful alternative for high-dimensional decision spaces
- The presented information is very limited
**Summary**

- Visualization in multiobjective optimization useful for various purposes
- Customized methods are needed to address the peculiarities of approximation set visualization as well as multiobjective landscape visualization
- New visualization methods should first be analyzed using approximation sets and problems with known properties
- Visualization methods should also be evaluated with user studies (never done in multiobjective optimization and seldom in evolutionary computation [39])
Acknowledgement

The authors acknowledge the financial support from the Slovenian Research Agency (Research core funding No. P2-0209).

References


References vii


References viii


References ix


References x


Visualizing mutually nondominating solution sets in many-objective 
optimization. 

[54] D. J. Walker. 
Visualising multi-objective populations with treemaps. 

Identifying good algorithm parameters in evolutionary multi- and many-objective optimisation: A visualisation approach. 

Three-dimensional display in nuclear medicine. 

Study on effect of MOGA with interactive island model using visualization. 