GECCO 2018
Tutorial on
Evolutionary Multiobjective Optimization

Dimo Brockhoff
dimo.brockhoff@inria.fr

updated slides will be available at
http://www.cmap.polytechnique.fr/~dimo.brockhoff/
Multiobjective Optimization

Multiple objectives that have to be optimized simultaneously

\( \text{cost} \)

\( \text{performance} \)

better

incomparable

worse

\( \text{max} \)

\( \text{min} \)
**A Brief Introduction to Multiobjective Optimization**

**Observations:**
1. There is no single optimal solution, but
2. Some solutions (●) are better than others (○)

---

![Graph showing observations in multiobjective optimization](image.png)

- **performance**
- **cost**
- **better**
- **worse**
- **incomparable**

Observations:
- There is no single optimal solution, but
- Some solutions (●) are better than others (○).
A Brief Introduction to Multiobjective Optimization

\[ u \text{ weakly Pareto dominates } v \ (u \leq_{\text{par}} v) : \ \forall 1 \leq i \leq k : f_i(u) \leq f_i(v) \]

\[ u \text{ Pareto dominates } v \ (u <_{\text{par}} v) : \ u \leq_{\text{par}} v \land v \not\leq_{\text{par}} u \]
A Brief Introduction to Multiobjective Optimization

\[ u \text{ weakly Pareto dominates } v \quad (u \leq_{\text{par}} v) : \quad \forall 1 \leq i \leq k : f_i(u) \leq f_i(v) \]

\[ u \text{ Pareto dominates } v \quad (u <_{\text{par}} v) : \quad u \leq_{\text{par}} v \land v \nless_{\text{par}} u \]
A Brief Introduction to Multiobjective Optimization

- Pareto dominance
- Cone dominance
- $\varepsilon$-dominance

Diagram with performance on the y-axis and cost on the x-axis, showing points and shaded regions representing different dominance concepts.
**A Brief Introduction to Multiobjective Optimization**

**Pareto set:** set of all non-dominated solutions (decision space)

**Pareto front:** its image in the objective space

---

**currently non-dominated front** (approximation)
**Pareto set:** set of all non-dominated solutions (decision space)

**Pareto front:** its image in the objective space
A Brief Introduction to Multiobjective Optimization

- decision space
  - $x_1$, $x_2$, $x_3$
  - solution of Pareto-optimal set
  - non-optimal decision vector

- objective space
  - $f_1$, $f_2$
  - vector of Pareto-optimal front
  - non-optimal objective vector
A Brief Introduction to Multiobjective Optimization

ideal point: best values
nadir point: worst values

obtained for Pareto-optimal points
Multiobjective Optimization
combination of optimization of a set and a decision for a solution

Optimization vs. Decision Making

decision making
selecting a solution

optimization
finding the good solutions
Selecting a Solution: Examples

Possible ranking: performance more important than cost

Possible Approaches:

- ranking: performance more important than cost

Graph showing relationship between performance and cost with data points plotted along the axes.
Selecting a Solution: Examples

Possible Approaches:

1. ranking: performance more important than cost
2. constraints: cost must not exceed 2400

diagram showing the relationship between performance and cost with selected solutions.
When to Make the Decision

Before Optimization:

- rank objectives, define constraints,…
- search for one (good) solution
When to Make the Decision

Before Optimization:

rank objectives, define constraints,...

search for one (good) solution

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:

Before Optimization:
When to Make the Decision

**Before Optimization:**
- rank objectives, define constraints,…
- search for one (good) solution

**After Optimization:**
- search for a set of (good) solutions
- select one solution considering constraints, etc.
When to Make the Decision

Before Optimization:
- rank objectives, define constraints, ...
- search for one (good) solution

After Optimization:
- search for a set of (good) solutions
- select one solution considering constraints, etc.

Focus: learning about a problem
- trade-off surface
- interactions among criteria
- structural information
- also: interactive optimization
Two Communities...

- established field (beginning in 1950s/1960s)
- bi-annual conferences since 1975
- background in economics, math, management and social sciences
- focus on optimization and decision making

- quite young field (first papers in mid 1980s)
- bi-annual conference since 2001
- background in computer science, applied math and engineering
- focus on optimization algorithms
- MCDM track at EMO conference since 2009
- special sessions on EMO at the MCDM conference since 2008
- joint Dagstuhl seminars since 2004
One of the Main Differences

Blackbox optimization

\[ x \in X \xrightarrow{f} (f_1(x), \ldots, f_k(x)) \]

only mild assumptions

→ EMO therefore well-suited for real-world engineering problems

- non-linear
- noisy
- non-differentiable
- uncertain
- expensive
- (integrated simulations, real experiments)
- many objectives
- huge search spaces
- many constraints
The Other Main Difference

Evolutionary Multiobjective Optimization
- set-based algorithms
- therefore possible to approximate the Pareto front in one run
Multiobjectivization

Some problems are easier to solve in a multiobjective scenario

example: TSP
[Knowles et al. 2001]

\[ \pi \in S_n \rightarrow f(\pi) \]

\[ \pi \in S_n \rightarrow (f_1(\pi, a, b), f_2(\pi, a, b)) \]

Multiobjectivization
by addition of new “helper objectives” [Jensen 2004]
job-shop scheduling [Jensen 2004], frame structural design
[Greiner et al. 2007], VRP [Watanabe and Sakakibara 2007], ... 
by decomposition of the single objective 
TSP [Knowles et al. 2001], minimum spanning trees [Neumann and Wegener 2006], protein structure prediction [Handl et al. 2008a], ...
also backed up by theory e.g. [Brockhoff et al. 2009, Handl et al. 2008b]
related to constrained and multimodal single-objective optimization 
see also this recent overview: [Segura et al. 2013]
Often innovative design principles among solutions are found

Example:
Cantilever beam topology optimization

[Bandaru and Deb 2015]
Often innovative design principles among solutions are found.

Example:
Clutch brake design
[Deb and Srinivasan 2006]

<table>
<thead>
<tr>
<th>Solution</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$x_4$</th>
<th>$x_5$</th>
<th>$f_1$</th>
<th>$f_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min. $f_1$</td>
<td>70</td>
<td>90</td>
<td>1.5</td>
<td>1000</td>
<td>3</td>
<td>0.4704</td>
<td>11.7617</td>
</tr>
<tr>
<td>Min. $f_2$</td>
<td>80</td>
<td>110</td>
<td>1.5</td>
<td>1000</td>
<td>9</td>
<td>2.0948</td>
<td>3.3505</td>
</tr>
</tbody>
</table>
Often innovative design principles among solutions are found

**Innovization** [Deb and Srinivasan 2006]

- using machine learning techniques to find new and innovative design principles among solution sets
- learning from/about a multi-objective optimization problem

**Other examples:**
- SOM for supersonic wing design [Obayashi and Sasaki 2003]
- Biclustering for processor design and knapsack [Ulrich et al. 2007]
- Successful case studies in engineering (noise barrier design, polymer extrusion, friction stir welding) [Deb et al. 2014]
The History of EMO At A Glance

- 1985: First EMO algorithms
- 1990: Dominance ranking
- 1995: Elitist EMO algorithms
- 2000: Visual performance assessment
- 2005: Attainment functions
- 2010: Convergence proofs
- 2015: Preference articulation

1980s: Dominance ranking, Elitist EMO algorithms
2000s: Visual performance assessment, Attainment functions
2010s: Convergence proofs, Preference articulation

Key developments:
- 1985: First EMO algorithms
- 1990: Dominance ranking
- 1995: Elitist EMO algorithms
- 2000: Visual performance assessment
- 2005: Attainment functions
- 2010: Convergence proofs
- 2015: Preference articulation

Topics:
- Statistical performance assessment
- Convergence proofs
- Design of test problems
- Running time analyses
- MCDM+ EMO (interactive EMO)
- Many-objective problems
- Expensive EMO (Surrogates)
- Design of Performance indicators
- Indicator-based EMO
- Scalarmization-based EMO
- Indicator-based EMO
- Inno-vization
The History of EMO At A Glance

Overall: 11190 references by March 26, 2018

https://emoo.cs.cinvestav.mx/
The EMO Community

The EMO conference series:

- EMO 2001
  - Zurich, CH
- EMO 2003
  - Faro, PT
- EMO 2005
  - Guanajuato, MX
- EMO 2007
  - Matsushima, JP
- EMO 2009
  - Nantes, FR
- EMO 2011
  - Ouro Preto, BR
- EMO 2013
  - Sheffield, GB
- EMO 2015
  - Guimarães, PT
- EMO 2017
  - Münster, DE
- EMO 2019
  - East Lansing, MI, USA

Many further activities:
special sessions, special journal issues, workshops, tutorials, ...

© Dimo Brockhoff
EMO tutorial, GECCO’2018, Kyoto, Japan, July 2018
Overview

The Big Picture

Basic Algorithm Design Principles and Concepts

Performance Assessment and Benchmarking

Preference Articulation
Fitness Assignment: Principal Approaches

aggregation-based
problem decomposition
(multiple single-objective optimization problems)

solution-oriented
scaling-dependent

max
max
max
max

max
max
max
max

criterion-based
VEGA

changing goals

dominance-based
SPEA2, NSGA-II
“modern” EMOA

set-oriented
scaling-independent

© Dimo Brockhoff
A scalarizing function $s$ is a function $s : \mathbb{Z} \rightarrow \mathbb{R}$ that maps each objective vector $u = (u_1, \ldots, u_n) \in \mathbb{Z}$ to a real value $s(u) \in \mathbb{R}$.
Solution-Oriented Problem Transformations

Example 1: weighted sum approach

\[ y = w_1 y_1 + \ldots + w_k y_k \]

Disadvantage: not all Pareto-optimal solutions can be found if the front is not convex
Solution-Oriented Problem Transformations

Example 2: weighted Tchebycheff

\[ y = \max_i | \lambda_i (z_i - y_i) | \]

Several other scalarizing functions are known, see e.g. [Miettinen 1999]
General Scheme of Most Set-Oriented EMO

mating selection (stochastic)

population (archiv)

offspring

environmental selection (greedy heuristic)

fitness assignment
partitioning into dominance classes

rank refinement within dominance classes
... goes back to a proposal by David Goldberg in 1989.
... is based on pairwise comparisons of the individuals only.

- **dominance rank**: by how many individuals is an individual dominated? *MOGA, NPGA*
- **dominance count**: how many individuals does an individual dominate? *SPEA, SPEA2*
- **dominance depth**: at which front is an individual located? *NSGA, NSGA-II, most of the recently proposed algorithms*
Illustration of Dominance-Based Partitioning

- Dominance rank:
  - Points are ranked based on their dominance depth.
  - Each point is assigned a rank.

- Dominance depth:
  - The depth of a point is determined by its rank.
  - Points with lower ranks have higher dominance depth.

\[ f_1 \rightarrow \min \rightarrow f_2 \rightarrow \min \]

© Dimo Brockhoff
EMO tutorial, GECCO'2018, Kyoto, Japan, July 2018
36
Refinement of Dominance Rankings

**Goal:** rank incomparable solutions within a dominance class

1. Diversity information

   **Kernel method**
   - diversity = function of the distances
   - ![Kernel method diagram]

   **k-th nearest neighbor**
   - diversity = function of distance to k-th nearest neighbor
   - ![K-th nearest neighbor diagram]

   **Histogram method**
   - diversity = number of elements within box(es)
   - ![Histogram method diagram]

2. (Contribution to a) quality indicator
Example: NSGA-II Diversity Preservation

Crowding Distance (CD)

- sort solutions with regard to each objective
- assign CD maximum value to extremal objective vectors
- compute CD based on the distance to the neighbors in each objective

\[
CD(i) = \frac{d_1(i)}{f_{1,\text{max}} - f_{1,\text{min}}} + \cdots + \frac{d_m(i)}{f_{m,\text{max}} - f_{m,\text{min}}}
\]
Selection in SPEA2 and NSGA-II can result in *deteriorative cycles*

non-dominated solutions already found can be lost
Remark: Many-Objective Optimization

- high number of objectives
  - percentage of non-dominated solutions within a random sample quickly approaches 100%
  - optimization is mainly guided by diversity criterion
  - apply secondary criterion compliant with dominance relation
Latest Approach (SMS-EMOA, MO-CMA-ES, HypE, …)
use hypervolume indicator to guide the search: refines dominance

Main idea
Delete solutions with the smallest hypervolume contribution
\[ d(s) = I_H(P) - I_H(P / \{s\}) \]
iteratively

But:
- can also result in cycles if reference point is not constant [Judt et al. 2011]
- expensive to compute exactly [Bringmann and Friedrich 2009]
- less and less practically important [Guerreiro and Fonseca 2017]
Indicator-Based Selection

- Concept can be generalized to any quality indicator

A (unary) quality indicator $I$ is a function $I: \Psi = 2^X \mapsto \mathbb{R}$ that assigns a Pareto set approximation a real value.

Multiobjective Problem $\xrightarrow{\text{Indicator}}$ Single-objective Problem

- for example: R2-indicator [Brockhoff et al. 2012], [Trautmann et al. 2013], [Díaz-Manríquez et al. 2013]

- Generalizable also to contribution to larger sets
  HypE [Bader and Zitzler 2011]: Hypervolume sampling + contribution if more than 1 (random) solution deleted
The Optimization Goal in Indicator-Based EMO

When the goal is to maximize a unary indicator…

- we have a single-objective problem on sets
- but what is the optimum?
- important: population size $\mu$ plays a role!

Optimal $\mu$-Distribution:

A set of $\mu$ solutions that maximizes a certain unary indicator $I$ among all sets of $\mu$ solutions is called *optimal $\mu$-distribution* for $I$. [Auger et al. 2009a]

Optimal \( \mu \)-Distributions for the Hypervolume

Hypervolume indicator refines dominance relation
\[ \implies \] most results on optimal \( \mu \)-distributions for hypervolume

Optimal \( \mu \)-Distributions (example results)

[Auger et al. 2009a]:
- contain equally spaced points iff front is linear
- density of points \( \propto \sqrt{-f'(x)} \) with \( f' \) the slope of the front

[Friedrich et al. 2011]:
- optimal \( \mu \)-distributions for the hypervolume correspond to \( \varepsilon \)-approximations of the front

\[
\begin{align*}
\text{OPT} & \quad 1 + \frac{\log(\min\{A/a, B/b\})}{n} \\
\text{HYP} & \quad 1 + \frac{\sqrt{A/a} + \sqrt{B/b}}{n - 4} \\
\text{logHYP} & \quad 1 + \frac{\sqrt{\log(A/a) \log(B/b)}}{n - 2}
\end{align*}
\]

! (probably) does not hold for > 2 objectives
Open Questions:

- How do the optimal $\mu$-distributions look like for $>2$ objectives?
- How to compute certain indicators quickly in practice?
  - several recent improvements for the hypervolume indicator
    [Yildiz and Suri 2012], [Bringmann 2012], [Bringmann 2013]
    [Guerreiro and Fonseca 2017]
- How to do indicator-based subset selection quickly?
  - also here several recent improvements
    [Kuhn et al. 2014], [Bringmann et al. 2014], [Guerreiro et al. 2015]
- What is the best strategy for the subset selection?

Further open questions on indicator-based EMO available at
http://simco.gforge.inria.fr/doku.php?id=openproblems
**Decomposition-Based Selection: MOEA/D**

**MOEA/D:** Multiobjective Evolutionary Algorithm Based on Decomposition [Zhang and Li 2007]

**Ideas:**
- optimize N scalarizing functions in parallel
- use best solutions of neighbor subproblems for mating
- keep the best solution for each scalarizing function
- update neighbors
- use external archive for non-dominated solutions
- several variants and enhancements
Remark: Variation in EMO

- at first sight not different from single-objective optimization
- most research on selection mechanisms (until now)
- but: convergence to a set $\neq$ convergence to a point

Open Question:
- how to achieve fast convergence to a set?

Related work:
- set-based gradient of the HV [Emmerich et al. 2007]
- multiobjective CMA-ES [Igel et al. 2007, Voß et al. 2010, Krause et al. 2016]
- RM-MEDA [Zhang et al. 2008]
- set-based variation [Bader et al. 2009]
- set-based fitness landscapes [Verel et al. 2011]
- offline and online configuration based on libraries of variation operators [Bezerra et al. 2015, Hadka and Reed 2013]
Overview

The Big Picture

Basic Algorithm Design Principles and Concepts

Performance Assessment and Benchmarking

Preference Articulation
Once Upon a Time...

... multiobjective EAs were mainly compared visually:

ZDT6 benchmark problem: IBEA, SPEA2, NSGA-II
Two Approaches for Empirical Studies

Attainment function approach
- applies statistical tests directly to the approximation set
- detailed information about how and where performance differences occur

Quality indicator approach
- reduces each approximation set to a single quality value
- applies statistical tests to the quality values

<table>
<thead>
<tr>
<th>Indicator</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypervolume indicator</td>
<td>6.3431</td>
<td>7.1924</td>
</tr>
<tr>
<td>$\varepsilon$-indicator</td>
<td>1.2090</td>
<td>0.12722</td>
</tr>
<tr>
<td>$R_2$ indicator</td>
<td>0.2434</td>
<td>0.1643</td>
</tr>
<tr>
<td>$R_3$ indicator</td>
<td>0.6454</td>
<td>0.3475</td>
</tr>
</tbody>
</table>

see e.g. [Zitzler et al. 2003]
Empirical Attainment Functions

\[ f_2(x) \]

\[ f_1(x) \]

Run 1

© Manuel López-Ibáñez
[López-Ibáñez et al. 2010]
Empirical Attainment Functions

© Manuel López-Ibáñez
[López-Ibáñez et al. 2010]
Empirical Attainment Functions

© Manuel López-Ibáñez
[López-Ibáñez et al. 2010]
Empirical Attainment Functions

© Manuel López-Ibáñez
[López-Ibáñez et al. 2010]
Empirical Attainment Functions

© Manuel López-Ibáñez
[López-Ibáñez et al. 2010]
Empirical Attainment Functions

© Manuel López-Ibáñez
[López-Ibáñez et al. 2010]
Empirical Attainment Functions: Definition

The Empirical Attainment Function \( \alpha(z) \) "counts" how many solution sets \( \mathcal{X}_i \) attain or dominate a vector \( z \) at time \( T \):

\[
\alpha_T(z) = \frac{1}{N} \sum_{i=1}^{N} 1\{\mathcal{X}_i \preceq_T z\}
\]

with \( \preceq_T \) being the weak dominance relation between a solution set and an objective vector at time \( T \).

Note that \( \alpha_T(z) \) is the empirical cumulative distribution function of the achieved objective function distribution at time \( T \) in the single-objective case ("fixed budget scenario").
Empirical Attainment Functions in Practice

© Manuel López-Ibáñez
[López-Ibáñez et al. 2010]

latest implementation online at
http://eden.dei.uc.pt/~cmfonsec/software.html
R package: http://lopez-ibanez.eu/eaftools
see also [López-Ibáñez et al. 2010, Fonseca et al. 2011]
Plotting Average Runtimes

**Note:** success probability can be naturally replaced by the average runtime of an artificially restarted algorithm (aRT):

```
code available at http://github.com/numbbo/coco/
see also [Brockhoff et al. 2017]
```
Quality Indicator Approach

Idea:
- transfer multiobjective problem into a set problem
- define an objective function (“quality indicator”) on sets
- use the resulting total (pre-)order (on the quality values)

Question:
Can any total (pre-)order be used or are there any requirements concerning the resulting preference relation?
⇒ Underlying dominance relation should be reflected!
\[
A \preceq B \iff \forall y \in B \exists x \in A \ x \leq_{\text{par}} y
\]
Refinements and Weak Refinements

1. \( \preccurlyeq \text{ ref} \) **refines** a preference relation \( \preccurlyeq \) iff

\[ A \preccurlyeq B \land B \not\succ A \Rightarrow A \preccurlyeq B \land B \not\succ A \]  

(\text{better } \Rightarrow \text{ better})

\[ \Rightarrow \text{ fulfills requirement} \]

2. \( \preccurlyeq \text{ ref} \) **weakly refines** a preference relation \( \preccurlyeq \) iff

\[ A \preccurlyeq B \land B \not\succ A \Rightarrow A \preccurlyeq B \]  

(\text{better } \Rightarrow \text{ weakly better})

\[ \Rightarrow \text{ does not fulfill requirement, but } \preccurlyeq \text{ does not contradict} \]

! sought are total refinements…  

[Zitzler et al. 2010]
Example: Refinements Using Indicators

\[ A \preceq B :\iff I(A) \geq I(B) \]

\( I(A) = \) volume of the weakly dominated area in objective space

\[ A \preceq B :\iff I(A,B) \leq I(B,A) \]

\( I(A,B) = \) how much needs A to be moved to weakly dominate B

uniary hypervolume indicator

binary epsilon indicator

refinement

© Dimo Brockhoff
EMO tutorial, GECCO’2018, Kyoto, Japan, July 2018
Example: Weak Refinement / No Refinement

**Weak Refinement**

\[ A \preceq B : \iff I(A, R) \leq I(B, R) \]

\( I(A, R) \) = how much needs \( A \) to be moved to weakly dominate \( R \)

**No Refinement**

\[ A \preceq B : \iff I(A) \leq I(B) \]

\( I(A) \) = variance of pairwise distances

Unary epsilon indicator

Unary diversity indicator
Quality Indicator Approach

Goal: compare two Pareto set approximations A and B

<table>
<thead>
<tr>
<th>Quality Indicator Approach</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>hypervolume</td>
<td>432.34</td>
<td>420.13</td>
</tr>
<tr>
<td>distance</td>
<td>0.3308</td>
<td>0.4532</td>
</tr>
<tr>
<td>diversity</td>
<td>0.3637</td>
<td>0.3463</td>
</tr>
<tr>
<td>spread</td>
<td>0.3622</td>
<td>0.3601</td>
</tr>
<tr>
<td>cardinality</td>
<td>6</td>
<td>5</td>
</tr>
</tbody>
</table>

Comparison method $C = \text{quality measure(s)} + \text{Boolean function}$

A, B $\xrightarrow{\text{quality measure}} \mathbb{R}^n \xrightarrow{\text{reduction}} \mathbb{R}^n \xrightarrow{\text{Boolean function}} \text{interpretation}$

"A better"
Example: Box Plots

- **epsilon indicator**
  - IBEA, NSGA-II, SPEA2
  - DTLZ2
  - Knapsack
  - ZDT6

- **hypervolume**
  - IBEA, NSGA-II, SPEA2
  - DTLZ2
  - Knapsack
  - ZDT6

- **R indicator**
  - IBEA, NSGA-II, SPEA2
  - DTLZ2
  - Knapsack
  - ZDT6
## Statistical Assessment (Kruskal Test)

### ZDT6 Epsilon

<table>
<thead>
<tr>
<th></th>
<th>IBEA</th>
<th>NSGA2</th>
<th>SPEA2</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBEA</td>
<td>~0</td>
<td></td>
<td>~0</td>
</tr>
<tr>
<td>NSGA2</td>
<td>1</td>
<td>~0</td>
<td></td>
</tr>
<tr>
<td>SPEA2</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Overall p-value = 6.22079e-17. Null hypothesis rejected (alpha 0.05)

### DTLZ2 R

<table>
<thead>
<tr>
<th></th>
<th>IBEA</th>
<th>NSGA2</th>
<th>SPEA2</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBEA</td>
<td>~0</td>
<td></td>
<td>~0</td>
</tr>
<tr>
<td>NSGA2</td>
<td>1</td>
<td>~0</td>
<td></td>
</tr>
<tr>
<td>SPEA2</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Overall p-value = 7.86834e-17. Null hypothesis rejected (alpha 0.05)

**Knapsack/Hypervolume:** $H_0 = \text{No significance of any differences}
Automated Benchmarking

- State-of-the-art in single-objective optimization: Blackbox Optimization Benchmarking (BBOB) with COCO platform
  https://github.com/numbbo/coco
- Release of a bi-objective test suite at BBOB-2016 workshop
- Focus on target-based runlengths
  - gives (nearly) anytime, interpretable results
  - defines problem=(test function instance, single-objective goal e.g. min. indicator difference to reference set, target precision)
  - reports average runtimes (aRT) to reach target precision
- COCO provides data profiles, scaling plots, scatter plots, tables, statistical tests, etc. automatically
Exemplary BBOB-2016 Results

Data from 15 submitted algorithms
Exemplary BBOB-2016 Results
A Few Recommendations

- always display *everything* you have
- look at *single runs*
- do each experiment at least twice
  
  (= look at the *variance* of your results)
- as quality indicators, use hypervolume, R2, or epsilon indicator
- see also the tutorial by Nikolaus Hansen on this topic (not restricted to single-objective optimization!)
Overview

The Big Picture

Basic Algorithm Design Principles and Concepts

Performance Assessment and Benchmarking

Preference Articulation
Articulating User Preferences During Search

What we thought: EMO is preference-less

Search before decision making: Optimization is performed without any preference information given. The result of the search process is a set of (ideally Pareto-optimal) candidate solutions from which the final choice is made by the DM.

What we learnt: EMO just uses weaker preference information
Nevertheless...

- the more (known) preferences incorporated the better
- in particular if search space is large

[Branke and Deb 2004] [Branke 2008] [Bechikh et al. 2015]

1. **Refine/modify dominance relation, e.g.:**
   - using goals, priorities, constraints
     [Fonseca and Fleming 1998a,b]
   - using different types of dominance cones
     [Branke and Deb 2004]

2. **Use quality indicators, e.g.:**
   - based on reference points and directions [Deb and Sundar 2006, Deb and Kumar 2007]
   - based on the hypervolume indicator
     [Brockhoff et al. 2013] [Wagner and Trautmann 2010]
   - based on the R2 indicator [Trautmann et al. 2013]
Example: Weighted Hypervolume Indicator

\[ I_H^w(A) = \int \sum_i w(\bar{z}) d\bar{z} \]

[Brockhoff et al. 2013]
Weighted Hypervolume in Practice

[Auger et al. 2009b]
Example: Desirability Function (DF)-SMS-EMOA

[Wagner and Trautmann 2010]
DF-SMS-EMOA in Practice
Example: R2-EMOA

Concept
Integration of preferences by varying the scalarizing functions

Position of ideal point
Example: R2-EMOA

Concept
Integration of preferences by varying the scalarizing functions

Restriction of the weight space
Interactive Approaches

Successive Preference Articulation = Interactive EMO

- recent interest of both EMO and MCDM community
- important in practice

Examples

- first interactive EMO: [Tanino et al. 1993]
- good overview: [Jaszkiewicz and Branke 2008]
- more recent work: [Brockhoff et al. 2014] [Branke et al. 2014]

Issues/Open Questions

- realistic scenarios/ value functions
- evaluation of interactive algorithms [López-Ibáñez and Knowles 2015]
Conclusions: EMO as Interactive Decision Support

- **modeling**
  - problem
  - specification
  - optimization
  - preference articulation
- **solution**
  - adjustment
  - analysis
  - visualization

decision making
The EMO Community

Links:
- EMO mailing list: https://lists.dei.uc.pt/mailman/listinfo/emo-list
- MCDM mailing list: http://lists.jyu.fi/mailman/listinfo/mcdm-discussion
- EMO bibliography: http://www.lania.mx/~ccoello/EMOO/
- EMO conference series: http://www.dep.uminho.pt/EMO2015/

Books:
- *Multi-Objective Optimization using Evolutionary Algorithms*  
  Kalyanmoy Deb, Wiley, 2001
- and more…
Software

PISA

A Platform and Programming Language Independent Interface for Search Algorithms

Welcome to the jMetal Web Site

jMetal is ...

jMetal stands for Metaheuristic Algorithm in Java, and it is an object-oriented Java-based framework for multi-objective optimization with metaheuristics.

You can use it to ...

The object-oriented architecture of framework and the included features allow you to experiment with the provided classic state-of-the-art techniques, develop your own algorithms, solve your optimization problem integrate jMetal in other tools, etc.

Our motivation is ...

The motivation driving us is to provide a framework that can be easily extended and customized by researchers and practitioners.

Summary of features

MOEA Framework

A Framework for Innovation

The MOEA Framework is a free and open source Java library for developing and experimenting with multiobjective evolutionary algorithms (MOEAs) and other general-purpose multiobjective optimization algorithms. The MOEA Framework supports genetic algorithms, differential evolution, particle swarm optimization, genetic programming, grammatical evolution, and more. A number of algorithms are provided out-of-the-box, including NSGA-II, NSGA-III, e-MOEA, IDES, MOEA/D. In addition, the MOEA Framework provides the tools necessary to rapidly design, develop, execute and statistically test optimization algorithms.

Key Features

- Fast, reliable implementations of many state-of-the-art multiobjective evolutionary algorithms
- Extensible with custom algorithms, problems and operators
- Supports master-slave, island-model, and hybrid parallelization
- Modular design for constructing new optimization algorithms from existing components
- Permissive open source license
- Fully documented source code

Download from SourceForge

© Dimo Brockhoff
Software

github.com/numbbo/coco/
Perspectives

Challenging Open (Research) Directions

- from algorithms to toolkits
  - libraries of modules for each task (selection, variation, etc.)
  - problem-specific algorithm configuration/ parameter tuning
- benchmarking
  - comparison with classical approaches
  - design/selection of practically relevant problems
  - Algorithm/toolkit recommendations for practice
- integration of EMO and MCDM into one field
- interactive preference articulation and learning
- interactive problem design
- integration of problem-specific knowledge

Questions?
Additional Slides
Dimo Brockhoff

RandOpt team
Inria Saclay - Ile-de-France
CMAP UMR 7641 École Polytechnique CNRS
Route de Saclay
91128 Palaiseau
France

After obtaining his diploma in computer science (Dipl.-Inform.) from University of Dortmund, Germany in 2005, Dimo Brockhoff received his PhD (Dr. sc. ETH) from ETH Zurich, Switzerland in 2009. Between June 2009 and October 2011 he held postdoctoral research positions---first at Inria Saclay Ile-de-France in Orsay and then at Ecole Polytechnique in Palaiseau, both in France. Since November 2011, Dimo has been a permanent researcher at Inria: from 2011 till 2016 with the Inria Lille - Nord Europe research center and since October 2016 with the Saclay - Ile-de-France research center, co-located with CMAP, Ecole Polytechnique. His most recent research interests are focused on evolutionary multiobjective optimization (EMO) and other (single-objective) blackbox optimization techniques, in particular with respect to benchmarking, theoretical aspects, and expensive optimization.
References


References


References


References


References


References


References


I would like to thank in particular Eckart Zitzler and Tobias Wagner who contributed significantly to the content of these slides over the years.

Many thanks go also to Carlos Fonseca and Manuel López-Ibáñez for pointing out some mistakes during my PPSN 2016 presentation on the same topic that should be corrected in these slides.