Introductory Example: The Knapsack Problem

- Weight: 750g, Profit: 5
- Weight: 1500g, Profit: 8
- Weight: 300g, Profit: 7
- Weight: 1000g, Profit: 3

Single objective:
choose subset that
- maximizes overall profit
- w.r.t. a weight limit

Multiobjective:
choose subset that
- maximizes overall profit
- minimizes overall weight

The Search Space

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

Approaches: • profit more important than cost (ranking)

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

Approaches: • profit more important than cost (ranking)
  • weight must not exceed 2400g (constraint)
When to Make the Decision

Before Optimization:
- ranks objectives, defines constraints, …
- searches for one (green) solution

After Optimization:
- searches for a set of (green) solutions
- selects one solution considering constraints, etc.

Outline

1. Introduction: Why multiple objectives make a difference
2. Basic Principles: Terms one needs to know
3. Algorithm Design: Do it yourself
4. Performance Assessment: How to compare algorithms
5. Applications Domains: Where EMO is useful
6. Further Information: What else
Optimization Problem: Definition

A general optimization problem is given by a quadruple \((X, Z, f, rel)\) where

- \(X\) denotes the decision space containing the elements among which the best is sought; elements of \(X\) are called decision vectors or simply solutions;
- \(Z\) denotes the objective space, the space within which the decision vectors are evaluated and compared to each other; elements of \(Z\) are denoted as objective vectors;
- \(f\) represents a function \(f: X \rightarrow Z\) that assigns each decision vector a corresponding objective vector; \(f\) is usually neither injective nor surjective;
- \(rel\) is a binary relation over \(Z\), i.e., \(rel \subseteq Z \times Z\), which represents a partial order over \(Z\).

Objective Functions

- Usually, \(f\) consists of one or several functions \(f_1, \ldots, f_n\) that assign each solution a real number. Such a function \(f_i: X \rightarrow \mathbb{R}\) is called an objective function, and examples are cost, size, execution time, etc.
- In the case of a single objective function (\(n=1\)), the problem is denoted as a single-objective optimization problem; a multiobjective optimization problem involves several (\(n \geq 2\)) objective functions.

Comparing Objective Vectors

The pair \((Z, rel)\) forms a partially ordered set, i.e., for any two objective vectors \(a, b \in Z\) there can be four situations:

- \(a\) and \(b\) are equal: \(a \ rel \ b\) and \(b \ rel \ a\)
- \(a\) is better than \(b\): \(a \ rel \ b\) and not \((b \ rel \ a)\)
- \(a\) is worse than \(b\): not \((a \ rel \ b)\) and \(b \ rel \ a\)
- \(a\) and \(b\) are incomparable: neither \(a \ rel \ b\) nor \(b \ rel \ a\)

Example: \(Z = \mathbb{R}^2\), \((a_1, a_2) \ rel \ (b_1, b_2) :\iff a_1 \leq b_1 \ Land a_2 \leq b_2\)

Preference Structures

- The function \(f\) together with the partially ordered set \((Z, rel)\) defines a preference structure on the decision space \(X\) that reflects which solutions the decision maker/user prefers to other solutions:
  \(x_1 \ prefrel x_2 :\iff f(x_1) \ rel \ f(x_2)\)

  One says:
  - Two solutions \(x_1, x_2\) are equal \(\iff x_1 = x_2\);
  - A solution \(x_1\) is indifferent to a solution \(x_2\) \(\iff x_1 \ prefrel x_2\) and \(x_1 \neq x_2\);
  - A solution \(x_1\) is preferred to a solution \(x_2\) \(\iff x_1 \ prefrel x_2\);
  - A solution \(x_1\) is strictly preferred to a solution \(x_2\) \(\iff x_1 \ prefrel x_2\) and not \((x_2 \ prefrel x_1)\);
  - A solution \(x_1\) is incomparable to a solution \(x_2\) \(\iff\) neither \(x_1 \ prefrel x_2\) nor \(x_2 \ prefrel x_1\).
The Notion of Optimality

- A solution \( x \in X \) is called optimal with respect to a set \( S \subseteq X \) iff no solution \( x' \in S \) is strictly preferred to \( x \), i.e., for all \( x' \in S \): \( x' \text{ prefrel } x \Rightarrow x \text{ prefrel } x' \).
- In other words, \( f(x) \) is a minimal element of \( f(S) \) regarding the partially ordered set \((Z, \text{rel})\).

Illustration of Pareto Optimality

Maximize \((y_1, y_2, \ldots, y_k) = f(x_1, x_2, \ldots, x_n)\)

Pareto(-optimal) set = set of all Pareto-optimal solutions

Pareto Dominance

Assumption:
- \( n \) objective functions \( f_i : X \rightarrow \mathbb{R} \) where \( Z = \mathbb{R}^n \)
- all objectives are to be maximized

Usually considered relation: weak Pareto dominance
- optimization problem: \((X, \mathbb{R}^n, (f_1, \ldots, f_n), \leq)\)
- weak Pareto dominance:
  \[ x_1 \leq x_2 : \iff \forall 1 \leq i \leq n : f_i(x_1) \geq f_i(x_2) \]
- Pareto dominance: strict version of weak Pareto dominance
  \[ x_1 < x_2 : \iff x_1 \leq x_2 \land x_2 \not\leq x_1 \]

Decision and Objective Space

Pareto set
- non-optimal decision vector
- non-optimal objective vector

Pareto front
Pareto Set Approximations

Pareto set approximation (algorithm outcome) = set of incomparable solutions

- \( A \) is better than \( B \) = not worse in all objectives and sets not equal
- \( C \) dominates \( D \) = better in at least one objective
- \( A \) strictly dominates \( C \) = better in all objectives
- \( B \) is incomparable to \( C \) = neither set weakly better

Preference Information

Preference information (here) = any additional information that refines the dominance relation on approximation sets (partial order \( \rightarrow \) total order)

Example:

- optimization goal = maximize size \( S \) of dominated objective space

Note: every algorithm implicitly or explicitly makes assumptions about the decision maker’s preferences (limited memory, selection)

What Is the Optimization Goal?

- Find all Pareto-optimal solutions?
  - Impossible in continuous search spaces
  - How should the decision maker handle 10000 solutions?
- Find a representative subset of the Pareto set?
  - Many problems are NP-hard
  - What does representative actually mean?
- Find a good approximation of the Pareto set?
  - What is a good approximation?
  - How to formalize intuitive understanding:
    - close to the Pareto front
    - well distributed

Outline

1. Introduction: Why multiple objectives make a difference
2. Basic Principles: Terms one needs to know
3. Algorithm Design: Do it yourself
4. Performance Assessment: How to compare algorithms
5. Applications Domains: Where EMO is useful
6. Further Information: What else
**Design Choices**

- **representation**
- **fitness assignment**
- **mating selection**
- **parameters**
- **environmental selection**
- **variation operators**

---

**Ranking Solutions**

- **aggregation-based**
  - weighted sum
  - VEGA
- **criterion-based**
  - SPEA2

---

**Example: VEGA**

1. Select according to $f_1, f_2, f_3, \ldots, f_k$
2. Shuffle population $M$
3. $k$ separate selections
4. Mating pool $M'$

---

**Aggregation-Based Ranking**

- Multiple objectives $(y_1, y_2, \ldots, y_k)$
- Transformation
- Parameters
- Single objective $y$

---

**Example: weighting approach**

$$y = w_1 y_1 + \ldots + w_k y_k$$

**Note:** weights need to be varied during the run...
Example: Weighted Sum

Underlying concept:
- Convert all objectives except of one into constraints
- Adaptively vary constraints

Example: Multistart Constraint Method

Underlying concept:
- Convert all objectives except of one into constraints
- Adaptively vary constraints
**Example: Multistart Constraint Method (Cont’d)**

Extension to n objectives: ECEA [Laumanns et al. 2006]
- $f_1$ is the objective to optimize
- The boxes are defined by constraints on $f_2$ and $f_3$

Extension to n objectives: ECEA [Laumanns et al. 2006]
- $f_1$ is the objective to optimize
- The boxes are defined by constraints on $f_2$ and $f_3$

**Dominance-Based Ranking**

Types of information:
- **dominance rank** by how many individuals is an individual dominated?
- **dominance count** how many individuals does an individual dominate?
- **dominance depth** at which front is an individual located?

Examples:
- MOGA, NPGA: dominance rank
- NSGA/NSGA-II: dominance depth
- SPEA/SPEA2: dominance count + rank

**Refining Rankings**

Ranks | Pure Dominance Rank | Refined Ranking
--- | --- | ---
1 | 0 | 0.245
2 | 0 | 0.311
3 | 0 | 0.329
4 | 0 | 1
5 | 0 | 2
6 | 0 | 2
7 | 0 | 3
8 | 0 | no selection pressure within equivalence classes
9 | 0 | modified dominance relation

Examples:
- MOGA, NPGA: dominance rank
- NSGA/NSGA-II: dominance depth
- SPEA/SPEA2: dominance count + rank

**Example: MOGA and SPEA2**

**MOGA** [Fonseca, Fleming 1993]
- R (raw fitness) = #dominating solutions
- S (strength) = #dominated solutions

**SPEA2** [Zitzler et al. 2002]
- R (raw fitness) = $\sum$ strengths of dominators

**Refining Rankings**

Ranks | Pure Dominance Rank | Refined Ranking
--- | --- | ---
1 | 0 | 0.245
2 | 0 | 0.311
3 | 0 | 0.329
4 | 0 | 1
5 | 0 | 2
6 | 0 | 2
7 | 0 | 3
8 | 0 | no selection pressure within equivalence classes
9 | 0 | modified dominance relation
Methods Based On Euclidean Distance

Density estimation techniques:

- **Kernel**
  - MOGA density estimate = sum of $f$ values where $f$ is a function of the distance

- **Nearest neighbor**
  - SPEA2 density estimate = volume of the sphere defined by the nearest neighbor

- **Histogram**
  - PAES density estimate = number of solutions in the same box

Computation Effort Versus Accuracy

Two Nearest Neighbor Variants

- **Objective-Wise Euclidean Distance**
  - NSGA-II: faster, good for 2 objectives
  - SPEA2: slower, good for 3 objectives and more

[Deb et al. 2002] [Zitzler et al. 2002]

The Problem of Deterioration

Observation:
The use of Euclidean distance can lead to deterioration

Knapsack problem

[Laumanns et al. 2002]

Refinement of Dominance Relations

Integration of Goals, Priorities, Constraints:

[Fonseca, Fleming 1998]

A is preferable over B \iff

\begin{align*}
\min_i f_i(A) - f_i(B) \\
\text{subject to:} \\
aw_i \neq av_i \\
aw_i \leq gw_i \\
aw_i \leq gw_i - 1 \\
gw_i \geq gw_i - 1 - 1
\end{align*}

Continuous dominance “relations”: [Zitzler et al. 2003]

- $I_{e+}(A,B) = \min_i f_i(A) - f_i(B)$
- $I_{e+}(A,B) \geq 0$ and $I_{e+}(B,A) < 0 \iff A$ dominates $B$

(binary additive epsilon quality indicator)
**Example: IBEA**

**Question:** How to continuous dominance “relations” for fitness assignment? [Zitzler, Künzli 2004]

**Given:** function \( I \) (binary quality indicator) with

\[
A \text{ dominates } B \iff I(A, B) < I(B, A)
\]

**Idea:** measure for “loss in quality” if A is removed

Fitness: \[
F'(x^1) = \sum_{x^2 \in P \setminus \{x^1\}} I(\{x^2\}, \{x^1\})
\]

...corresponds to continuous extension of dominance rank
...blurs influence of dominating and dominated individuals

**Example: IBEA (Cont’d)**

**Fitness assignment:** \( O(n^2) \)

Fitness:

\[
F'(x^1) = \sum_{x^2 \in P \setminus \{x^1\}} -e^{-I(\{x^2\}, \{x^1\})/\kappa}
\]

- parameter \( \kappa \) is problem- and indicator-dependent
- no additional diversity preservation mechanism

**Mating selection:** \( O(n) \)

- binary tournament selection, fitness values constant

**Environmental selection:** \( O(n^2) \)

- iteratively remove individual with lowest fitness
- update fitness values of remaining individuals after each deletion

---

**Further Design Aspects**

- **Constraint handling:** How to integrate constraints into fitness assignment?
- **Archiving / environmental selection:** How to keep a good approximation?
- **Hybridization:** How to integrate, e.g., local search in a multiobjective EA?
- **Preference articulation:** How to focus the search on interesting regions?
- **Robustness and uncertainty:** How to account for variations in the objective function values?
- **Data structures:** How to support, e.g., fast dominance checks?

---

**Constraint Handling & Multiple Objectives**

<table>
<thead>
<tr>
<th>penalty functions</th>
<th>constraints as objectives</th>
<th>modified dominance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Add penalty term to fitness</td>
<td>Introduce additional objective(s)</td>
<td>extend to infeasible solutions</td>
</tr>
</tbody>
</table>

overall constraint violation

- [Michalewicz 1992]
- [Wright, Loosemore 2001]
- [Deb 2001]

constraints treated separately

- [Coello 2000]
- [Fonseca, Fleming 1998]
Archiving / Environmental Selection

**Variant 1:** without archive
- old population
- offspring
- new population

**Variant 2:** with archive
- old population
- offspring
- new population
- archive

Deterministic truncation
- archive = only nondominated solutions

Additional selection criteria:
- Density information / other preferences
- Time
- Chance

Outline

1. Introduction: Why multiple objectives make a difference
2. Basic Principles: Terms one needs to know
3. Algorithm Design: Do it yourself
4. Performance Assessment: How to compare algorithms
5. Applications Domains: Where EMO is useful
6. Further Information: What else

Once Upon a Time...

... multiobjective EAs were mainly compared visually:

ZDT6 benchmark problem: IBEA, SPEA2, NSGA-II

Performance Assessment: Approaches

1. Theoretically (by analysis): difficult
   - Limit behavior (unlimited run-time resources)
   - Running time analysis
2. Empirically (by simulation): standard

Problems: randomness, multiple objectives
Issues: quality measures, statistical testing, visualization, benchmark problems, parameter settings, …
Two Approaches for Empirical Studies

**Attainment function approach:**
- Applies statistical tests directly to the samples of approximation sets
- Gives detailed information about how and where performance differences occur

**Quality indicator approach:**
- First, reduces each approximation set to a single value of quality
- Applies statistical tests to the samples of quality values

<table>
<thead>
<tr>
<th>Indicator</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypervolume</td>
<td>6.3431</td>
<td>7.1924</td>
</tr>
<tr>
<td>ε-indicator</td>
<td>1.2060</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>

Empirical Attainment Functions

three runs of two multiobjective optimizers

Attainment Plots

50% attainment surface for IBEA, SPEA2, NSGA2 (ZDT6)

Attainment Function Analysis

- A Kolmogorov-Smirnov test examines the maximum difference between two cumulative distribution functions
- A KS-like test can be used to probe differences between the empirical attainment functions of a pair of optimizers, A and B
- The null hypothesis is that the attainment functions of A and B are identical
- The alternative hypothesis is that the distributions differ somewhere

[Fonseca et al. 2001]
Statistical Assessment

ZDT6
- IBEA – NSGA-II: significant difference (p=0)
- IBEA – SPEA2: significant difference (p=0)
- SPEA2 – NSGA-II: significant difference (p=0)

Knapsack
- IBEA – NSGA-II: no significant difference
- IBEA – SPEA2: no significant difference
- SPEA2 – NSGA-II: no significant difference

Quality Indicator Approach

Goal: compare two Pareto set approximations A and B

<table>
<thead>
<tr>
<th></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>0.3673</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 $C = \text{quality measure(s)} + \text{Boolean function}$

Example: $\varepsilon$-Quality Indicator

Two solutions:
$E(a,b) = \max_{1 \leq i \leq n} \min_{a \leq f_i(b)} \varepsilon \cdot f_i(a) \geq f_i(b)$

Two approximations:
$E(A,B) = \max_{b \in B} \min_{a \in A} E(a,b)$

Unary quality indicator: $I(A) = E(A,R)$ where R is a reference set

Power of Unary Quality Indicators

Important: compliance with dominance relations [Zitzler et al. 2003]
Statistical Assessment (Kruskal Test)

**ZDT6**
Epsilon

- IBEA \(\sim 0\)
- NSGA-II \(\sim 0\)
- SPEA2

- IBEA
- NSGA-II \(1\)
- SPEA2 \(1\)

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

**DTLZ2**
R

- IBEA \(\sim 0\)
- NSGA-II \(\sim 0\)
- SPEA2

- IBEA
- NSGA-II \(1\)
- SPEA2 \(1\)

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

Knapsack/Hypervolume: H0 = No significance of any differences

Performance Assessment Tools

- Reference set calculation
- Attainment function calculation
- Indicators
- Statistical testing procedures

http://www.tik.ee.ethz.ch/pisa
Outline

1. Introduction: Why multiple objectives make a difference
2. Basic Principles: Terms one needs to know
3. Algorithm Design: Do it yourself
4. Performance Assessment: How to compare algorithms
5. Applications Domains: Where EMO is useful
6. Further Information: What else

Design Space Exploration

Examples: computer design, biological experiment design, etc.

Example Applications

Architecture exploration:
- min. cost
- max. performance
- min. power consumption

Genetic marker selection:
- min. cost
- max. sensitivity

Application: Genetic Programming

Problem: Trees grow rapidly
- Premature convergence
- Overfitting of training data

Common approaches:
- Constraint (tree size limitation)
- Penalty term (parsimony pressure)
- Objective ranking (size post-optimization)
- Structure-based (ADF, etc.)

Multiobjective approach: Optimize both error and size

Keep and optimize small trees (potential building blocks)
Multiobjective approach (SPEA2) can find
- a correct solution with higher probability
- a correct solution slightly faster
- more compact (correct) solutions

than alternative approaches on even-parity problem.

[Bleuler et al. 2001]

1. Introduction: Why multiple objectives make a difference
2. Basic Principles: Terms you need to know
3. Algorithm Design: Do it yourself
4. Performance Assessment: Once upon a time
5. Applications Domains: Where EMO is useful
6. Further Information: What else

The Concept of PISA

Algorithms
- SPEA2
- NSGA-II
- PAES

Applications
- knapsack
- TSP
- network design

© Eckart Zitzler
ETH Zürich

GECCO 2006
Tutorial on EMO

The EMO Community

Links:
- EMO mailing list: http://w3.ualg.pt/lists/emo-list/
- EMO bibliography: http://www.lania.mx/~ccoello/EMOO/

Events:
- Conference on Evolutionary Multi-Criterion Optimization (EMO 2007 to be held in Japan)

Books:
- Multi-Objective Optimization using Evolutionary Algorithms
  Kalyanmoy Deb, Wiley, 2001

References