Multi-objective Estimation of Distribution Algorithms

Luis Martí

Laboratory of Adaptive Intelligence and Robotics
Dept. of Electrical Engineering
Pontifícia Universidade Católica do Rio de Janeiro

5ª Escola Luso-Brasileira de Computação Evolutiva
LNCC, Petrópolis, Fev. 18–21, 2014.
Research context

Context

- Real-world problems that involved the optimization of more than one objective at the same time.
- Actually, there were many objectives (> 200!).
- Evolutionary approaches to the problems.
Research context

Context

- Real-world problems that involved the optimization of more than one objective at the same time.
- Actually, there were many objectives (> 200!).
- Evolutionary approaches to the problems.

In this talk...

- Focused on optimization problems with many objectives.
- Using machine learning for improving search: EDAs.
- Model building in multi-objective EDAs was not correctly dealt with.
- Studied stopping criteria and convergence.
Evolutionary Multi-objective Optimization

At this point you should be aware that...

- Evolutionary algorithms are good tools for optimization.
- Multi-objective optimization is more than just an extension of single-objective optimization.
- Actually most—if not all—real problems are multi-objective.
Even choosing a fruit is a multi-objective problem!
Two-dimensional example

minimize $F(x) = \langle f_1(x), f_2(x) \rangle$, with $x \in D \subseteq \mathbb{R}^2$. 

Decision set ($D$)
minimize $F(x) = \langle f_1(x), f_2(x) \rangle$,  
with $x \in D \subseteq \mathbb{R}^2$. 

Decision set ($D$) 

Objective set ($O$)
Two-dimensional example

minimize $F(x) = \langle f_1(x), f_2(x) \rangle$, 
with $x \in D \subseteq \mathbb{R}^2$. 

Decision set ($D$)  

Objective set ($O$)  

Pareto-optimal front  

$x_2$  
$x_1$  

$f_2(x)$  

$f_1(x)$
Two-dimensional example

minimize $F(x) = \langle f_1(x), f_2(x) \rangle$, 
with $x \in \mathcal{D} \subseteq \mathbb{R}^2$. 

Decision set ($\mathcal{D}$)

Objective set ($\mathcal{O}$)

Pareto-optimal front
Two-dimensional example

minimize \( F(x) = \langle f_1(x), f_2(x) \rangle \),
with \( x \in D \subseteq \mathbb{R}^2 \).
Multi-objective optimization problem

\[
\text{minimize } F(x) = \langle f_1(x), \ldots, f_M(x) \rangle, \\
\text{with } x \in \mathcal{D}.
\]

- \mathcal{D}: feasible set — can be defined as constraints;
- \mathcal{O}: objective set;
- optimality — *Pareto dominance*;
- \mathcal{D}^*: Pareto-optimal set;
- \mathcal{O}^*: Pareto-optimal front, and;
- \mathcal{P}^*: optimizer solution.

A decision maker selects elements of \( \mathcal{P}^*_t \).
Why multi-objective optimization?

- Many problems can be formulated as an optimization problem.
- Frequently they are also intrinsically multi-objective.
- Most machine learning problems are multi-objective.
- It can also be used as a research and innovation tool.

Evolutionary approaches

- In the general case, MOPs are NP-hard.
- Heuristics are necessary.
- Multi-objective optimization evolutionary algorithms.
Many-objective problems

Problems with four or more objectives.

Many-objective problems

Problems with four or more objectives.

Challenges

- Scalability.

Many-objective problems

Problems with four or more objectives.

Challenges

- Scalability.
- Fitness assignment.

Many-objective problems

Problems with four or more objectives.

Challenges

- Scalability.
- Fitness assignment.
- Visualization.

Many-objective problems

Problems with four or more objectives.

Challenges

- Scalability.
- Fitness assignment.
- Visualization.
- Poor understanding of convergence and progress → stopping criteria.

Many-objective problems

Problems with four or more objectives.

Challenges

- Scalability.
- Fitness assignment.
- Visualization.
- Poor understanding of convergence and progress → stopping criteria.

Scalability

- Exponential relation between the number of objectives and the amount of required resources.¹
- Large populations are needed.

How to deal with many-objs

Approaches to many-objs

- “Better” fitness assignment;
- problem decomposition – MOEA/D;\(^2\)
- reduction of the number of objectives, and;
- better search “engines”.

Introducing learning in the search process might help!\(^3\)

...and this direction is (or was) unexplored!


A population of individuals;
Evolutionary algorithms

- A population of individuals;
- individuals are ranked using a fitness assignment function;
Evolutionary algorithms

- A population of individuals;
- individuals are ranked using a **fitness assignment function**;
- **evolution-inspired** operators are applied;
Evolutionary algorithms

- A population of individuals;
- individuals are ranked using a fitness assignment function;
- evolution-inspired operators are applied;
- fittest individuals have a more active role.
Estimation of distribution algorithms (EDAs)

From **population-based** to **model-based** optimization algorithms.

Evolutionary operators are substituted by...

...the construction of a model of the fittest individuals;

that model is sampled to produce new individuals.
Estimation of distribution algorithms (EDAs)

From **population-based** to **model-based** optimization algorithms.
Evolutionary operators are substituted by...
...the construction of a model of the fittest individuals;
that model is sampled to produce new individuals.

**Model-building process** — Bayesian networks, Mixtures, etc.
Main MOEDA approaches

Most MOEDAs are an extension of a single-objective EDA that uses a preexisting MOEA fitness assignment.

---


Main MOEDA approaches

Most MOEDAs are an extension of a single-objective EDA that uses a preexisting MOEA fitness assignment.

- Graphical algorithms:
  - \{m|h|mm|Mr\}-BOA;
  - Multi-objective mixture-based iterated density estimation algorithm (MIDEA);
    - Randomized leader algorithm;
    - $k$-means;
    - expectation-maximization.

- multi-objective CMA-ES;

- Regularity model-based multi-objective EDA (RM-MEDA);

- Multi-objective Parzen EDA (MOPED);

---


## Multi-objective EDA summary

<table>
<thead>
<tr>
<th>MOEDA</th>
<th>Domain</th>
<th>Fitness Assignment</th>
<th>Model Building</th>
</tr>
</thead>
<tbody>
<tr>
<td>mBOA</td>
<td>combinatorial</td>
<td>NSGA-II</td>
<td>Bayesian</td>
</tr>
<tr>
<td>mhBOA</td>
<td>combinatorial</td>
<td>NSGA-II</td>
<td>Bayesian</td>
</tr>
<tr>
<td>mmBOA</td>
<td>comb. + cont.</td>
<td>SPEA2</td>
<td>Bayesian</td>
</tr>
<tr>
<td>MrBOA</td>
<td>continuous</td>
<td>NSGA-II</td>
<td>Bayesian</td>
</tr>
<tr>
<td>Naïve MIDEA</td>
<td>continuous</td>
<td>Proprietary</td>
<td>Univariate dists.</td>
</tr>
<tr>
<td>MO-CMA-ES</td>
<td>continuous</td>
<td>Indicator</td>
<td>Cov. matrix</td>
</tr>
<tr>
<td>RM-MEDA</td>
<td>continuous</td>
<td>NSGA-II</td>
<td>Regularity prop.</td>
</tr>
<tr>
<td>MOPED</td>
<td>continuous</td>
<td>NSGA-II</td>
<td>Parzen estimator</td>
</tr>
</tbody>
</table>
Multi-objective EDAs (MOEDAs) in many-obj’s

It was expected that EDAs would make many-obj’s more tractable.

---

Multi-objective EDAs (MOEDAs) in many-obj’s

It was expected that EDAs would make many-obj’s more tractable.

- Scalability has not improved substantially by introducing MOEDAs.
- A preliminary experiment showed that non-robust approaches performed better than more sound ones.

---

Multi-objective EDAs (MOEDAs) in many-obj’s

It was expected that EDAs would make many-obj’s more tractable.

- Scalability has not improved substantially by introducing MOEDAs.
- A preliminary experiment showed that non-robust approaches performed better than more sound ones.\(^7\)

Why? — A hypothesis

- MOEDA approaches have mostly used off-the-shelf machine learning methods, but;
- model building is not a “typical” machine learning problem.

---

Understanding Model Building
Scalability: A salient issue for MOEDAs

Probably caused by undesirable properties of model-building algorithms.

1. Incorrect treatment of data outliers.
2. Loss of population diversity.
3. Excessive effort dedicated to the creation of the model.
4. The “jump-back” effect.
The outliers issue

\[ f_2(x) \]

\[ \leftarrow f_1(x) \]
The outliers issue

\[ f_2(x) \]

\[ \downarrow \]

model-building dataset \( (P'_t) \)

\[ \leftarrow f_1(x) \]
The outliers issue

$f_2(x)$

"outliers"

model

$\leftarrow f_1(x)$
Outliers in single-objective optimization
More formally: Set-wise vs. local error minimization

The “outlier” concept is not clearly defined nor understood.

- Unsupervised learning model-building data set: $\Psi = \{x_i\}$. 
More formally: Set-wise vs. local error minimization

The “outlier” concept is not clearly defined nor understood.

- Unsupervised learning model-building data set: $\Psi = \{x_i\}$.
- Machine learning model: $\mathcal{M}(x_i, \phi)$. 

There are many other forms of $E(\cdot)$ and $E_{\text{tot}}$. As outliers have little contribution to $E_{\text{tot}}$ they are disregarded! But, in our case, outliers are probably the most important elements of $\Psi$. 
More formally: Set-wise vs. local error minimization

The “outlier” concept is not clearly defined nor understood.

- Unsupervised learning model-building data set: $\Psi = \{x_i\}$.
- Machine learning model: $\mathcal{M}(x_i, \phi)$.
- Parameters $\phi$ are tuned minimizing the error over the learning data set:

$$E_{\text{tot}} = \sum_{x_i \in \Psi} E(\mathcal{M}(x_i, \phi)).$$

- There are many other forms of $E(\cdot)$ and $E_{\text{tot}}$. 
The “outlier” concept is not clearly defined nor understood.

- Unsupervised learning model-building data set: $\Psi = \{x_i\}$.
- Machine learning model: $M(x_i, \phi)$.
- Parameters $\phi$ are tuned minimizing the error over the learning data set:
  \[ E_{\text{tot}} = \sum_{x_i \in \Psi} E(M(x_i, \phi)) . \]
- There are many other forms of $E(\cdot)$ and $E_{\text{tot}}$.

As outliers have little contribution to $E_{\text{tot}}$ they are disregarded!
More formally: Set-wise vs. local error minimization

The “outlier” concept is not clearly defined nor understood.

- Unsupervised learning model-building data set: $\Psi = \{x_i\}$.
- Machine learning model: $\mathcal{M}(x_i, \phi)$.
- Parameters $\phi$ are tuned minimizing the error over the learning data set:
  \[ E_{\text{tot}} = \sum_{x_i \in \Psi} E(\mathcal{M}(x_i, \phi)) \, . \]
- There are many other forms of $E(\cdot)$ and $E_{\text{tot}}$.

As outliers have little contribution to $E_{\text{tot}}$ they are disregarded!.

But, in our case, outliers are probably the most important elements of $\Psi$. 
Diversity loss

This phenomenon has been documented in single and multi-objective EDAs.

Can be attributed to two main causes:
- biased selection processes, and;
- incorrect model building. — outliers again!

- Missing outliers leads to rapid homogenization.
- Some works “patch” current model-building algorithms:
  - Resampling;\(^8\)
  - “Mixed” model-building subset selection;\(^9\)
  - Injecting individuals from an EA being run in parallel.

---

“Overoptimal” models

EDAs use metaheuristics for determining the optimal model complexity.

- Structural risk minimization;
- Bayesian information criterion, and;
- many more.
- usually very computationally expensive...

A certain degree of “on-the-fly guessing” might provide an equally good solution at lower cost.
EDAs use metaheuristics for determining the optimal model complexity.

- Structural risk minimization;
- Bayesian information criterion, and;
- many more.
- usually very computationally expensive.

A certain degree of “on-the-fly guessing” might provide an equally good solution at lower cost.

We just need a “good model”, not the most compact one.
Bosman has reached similar conclusions regarding outliers and diversity loss:

- Adaptive Variance Scaling (AVS) and
- Standard–Deviation Ratio triggering (SDR).

---

The “Jump-Back” effect

Bosman has reached similar conclusions regarding outliers and diversity loss:

- Adaptive Variance Scaling (AVS) and
- Standard–Deviation Ratio triggering (SDR).

He has also noticed a cyclic behaviour in MOEDAs model building.

- Anticipated Mean Shift (AMS)
- Multi-objective Adapted Maximum-Likelihood Gaussian Model-miXture (MAMaLGaM-X).\textsuperscript{10}

This matter has not been properly addressed; new algorithms are required, or; we must find existing ones with the desired characteristics.
This matter has not been properly addressed; new algorithms are required, or; we must find existing ones with the desired characteristics.

Requirements

- Special learning laws sensible to outliers, and;
- self-organization for the on-the-fly determination of the optimal model complexity.
Competent model builders

- This matter has not been properly addressed;
- new algorithms are required, or;
- we must find existing ones with the desired characteristics.

Requirements

- Special learning laws sensible to outliers, and;
- self-organization for the on-the-fly determination of the optimal model complexity.

Set-wise error-based learning should be avoided.
Model-building growing neural gas (MB-GNG)\textsuperscript{11}

- One-layer network composed by a set of nodes;

Model-building growing neural gas (MB-GNG)\textsuperscript{11}

- One-layer network composed by a set of nodes;
- each node represents a local Gaussian density, and;

\vspace{1cm}

One-layer network composed by a set of nodes;
- each node represents a local Gaussian density, and;
- has an accumulated error that is used to determine where to inject more nodes;

---

Model-building growing neural gas (MB-GNG)\textsuperscript{11}

- One-layer network composed by a set of nodes;
- each node represents a local Gaussian density, and;
- has an accumulated error that is used to determine where to inject more nodes;
- three concurrent learning processes:
  - network adaptation;
  - node insertion, and;
  - node deletion.

One-layer network composed by a set of nodes; each node represents a local Gaussian density, and; has an accumulated error that is used to determine where to inject more nodes; three concurrent learning processes:

- network adaptation;
- node insertion, and;
- node deletion.

We added a cluster repulsion term for inducing a better spread of the nodes.

---

MB-GNG adaptation

\[ \Delta \mu_b = \epsilon_{\text{best}} (x - \mu_b) . \]
MB-GNG adaptation

\[ \Delta \mu_b = \epsilon_{\text{best}} (x - \mu_b). \]

\[ \Delta \mu_{\text{vic}} = \epsilon_{\text{vic}} (x - \mu_v) + \beta e \left( -\frac{d(\mu_v, \mu_b)}{\zeta} \right) \frac{\sum_{c_u \in V_b} d(\mu_u, \mu_b) (\mu_v - \mu_b)}{|V_b|}. \]
MB-GNG algorithm

A node, \( c_i := \langle \mu_i, \sigma_i, \xi_i, \mathcal{V}_i, v_{ij} \rangle; \ x \in \Psi \), model-building data set.

1: repeat
2: Randomly select \( x \in \Psi \).
3: Best-matching node, \( c_b \) and second best-matching node, \( c_{b'} \).
4: if \( c_{b'} \notin \mathcal{V}_b \) then
5: ▷ Make \( c_b \) and \( c_{b'} \) neighbors.
6: Edge age is set to \( \nu_{bb'} = 0 \) and error accumulator, \( \xi_b \), is updated.
7: Learning takes place in \( c_b \) and \( \forall c_{\text{vic}} \in \mathcal{V}_b \).
8: if \( t \) is node insertion iteration then
9: ▷ Node with largest accumulated error, \( c_e \), and the worst among its neighbors, \( c_{e'} \).
10: ▷ Insert new node between them.
11: Edge ages are increased — edges too old are removed — isolated nodes are removed.
12: until end condition met
MB-GNG: An example
MONEDA: Multi-objective neural EDA

Uses NSGA-II fitness assignment.

---

MONEDA: Multi-objective neural EDA

Uses NSGA-II fitness assignment.

MONEDA: Multi-objective neural EDA

Uses NSGA-II fitness assignment.

MONEDA: Multi-objective neural EDA

Uses NSGA-II fitness assignment.

\[ P_t \quad \xrightarrow{\text{best}} \quad \hat{P}_t \quad \xleftarrow{\text{worst}} \quad \text{MB-GNG} \]

\[ [\gamma \alpha n_{\text{pop}}] \]

---

MONEDA: Multi-objective neural EDA\textsuperscript{12}

Uses NSGA-II fitness assignment.

MONEDA: Multi-objective neural EDA

Uses NSGA-II fitness assignment.

MONEDA: Multi-objective neural EDA

Uses NSGA-II fitness assignment.

Validating MB-GNG

WFG4–WFG9

- Separability;
- uni-/multi-modality;
- parameter bias, and;
- deceptive locally optimal fronts.

Experiment design

- Each problem/dimension/algorithn repeated 30 times.
- Measured a set of performance indicators:
  - Hypervolume, additive epsilon, and $O^*$ coverage.
- Kruskal–Wallis for establishing the non-homogeneity of results.
- Pair-wise Conover–Inman procedure to determine winners.
- Performance index $P_{p,m}(A_i)$: perf. of algorithm $A_i$ wrt others.
Mean perf. index by problem

Mean perf. index by dim.
Assessing MONEDA: Test problems

Scaling on the number of objectives: $M = 3, 6, 9, 12.$
MONEDA and other algorithms — hypervolume

- Naïve MIDEA (nMID)
- MOPED (MOP)
- MONEDA (MON)
- MrBOA (MrB)
- RM-MEDA (RMM)
- NSGA-II (NSG)
- SPEA2 (SPE)

Mean perf. by problem

Mean perf. by dimension
Model-building CPU ops.

DTLZ3

DTLZ6

DTLZ7

WFG1

WFG2

WFG6
Measuring diversity loss

So far, we have demonstrated a better performance but what about diversity loss.

- The local correlation integral (LOCI) method\(^{13}\) is used to mark population outliers.
- It uses the multi-granularity deviation factor (MDEF) as outlier metric.
- MDEF is said to correctly deal with multi-density and multi-granularity.

We can measure how MDEF changes as algorithms progresses. Problems UF1–UF6.\(^{14}\)

---


Measuring diversity loss

UF1 Problem

HypE | MrBOA | Naive MIDEA

SDR | AVS | MIDEA | MAMALGAMx+ | MB-GNG

Graphs showing iteration vs. diversity loss for different algorithms.
Measuring diversity loss (II)

UF2 Problem

HypE

MrBOA

Naive MIDEA

SDR AVS MIDEA

MAMALGAMx +

MB-GNG
Measuring diversity loss (III)

UF3 Problem

<table>
<thead>
<tr>
<th>HypE</th>
<th>MrBOA</th>
<th>Naive MIDEA</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
<td><img src="image3.png" alt="Graph" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SDR AVS MIDEA</th>
<th>MAMALGAMx+</th>
<th>MB-GNG</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image4.png" alt="Graph" /></td>
<td><img src="image5.png" alt="Graph" /></td>
<td><img src="image6.png" alt="Graph" /></td>
</tr>
</tbody>
</table>
Measuring diversity loss (IV)

UF4 Problem

HypE  MrBOA  Naive MIDEA

SDR AVS MIDEA  MAMALGAMx+  MB-GNG
Measuring diversity loss (V)

UF5 Problem

<table>
<thead>
<tr>
<th>HypE</th>
<th>MrBOA</th>
<th>Naive MIDEA</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
<td><img src="image3.png" alt="Graph" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SDR AVS MIDEA</th>
<th>MAMALGAMx+</th>
<th>MB-GNG</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image4.png" alt="Graph" /></td>
<td><img src="image5.png" alt="Graph" /></td>
<td><img src="image6.png" alt="Graph" /></td>
</tr>
</tbody>
</table>
Measuring diversity loss (VI)

**UF6 Problem**

<table>
<thead>
<tr>
<th>HypE</th>
<th>MrBOA</th>
<th>Naive MIDEA</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
<td><img src="image3.png" alt="Graph" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SDR AVS MIDEA</th>
<th>MAMALGAMx+</th>
<th>MB-GNG</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image4.png" alt="Graph" /></td>
<td><img src="image5.png" alt="Graph" /></td>
<td><img src="image6.png" alt="Graph" /></td>
</tr>
</tbody>
</table>
Adaptive Resonance Theory (ART)

- ART\textsuperscript{15} is a theory of human cognition that has seen realization as neural networks.
- Adjusts categories in response to familiar inputs, and;
- creates new categories in response to inputs different enough from those previously seen.


Adaptive Resonance Theory (ART)

- ART\textsuperscript{15} is a theory of human cognition that has seen realization as neural networks.
- Adjusts categories in response to familiar inputs, and;
- creates new categories in response to inputs different enough from those previously seen.

Match-based learning


ART\textsuperscript{15} is a theory of human cognition that has seen realization as neural networks.

- Adjusts categories in response to familiar inputs, and;
- creates new categories in response to inputs different enough from those previously seen.

Match-based learning

- ART networks are not suitable for some classes of classical machine-learning applications\textsuperscript{16}.
- What is an inconvenience in that area is a feature in our case.


ART for model building

- **stability-plasticity dilemma**, or;
- when to learn and when to adapt.
- Adjust previously-learned categories in response to familiar inputs.
- Vigilance test allows to regulate the maximum tolerable difference.
- Create new “categories” in response to inputs different enough.
Model-building

- We used Gaussian ART\textsuperscript{17}, as it works in continuous domains.
- There are other ART networks suitable for combinatorial problems.


\textsuperscript{19}
**Model-building**

- We used Gaussian ART\(^\text{17}\), as it works in continuous domains.
- There are other ART networks suitable for combinatorial problems.

**Indicator-based selection**

- Monte Carlo approximation to hypervolume, as proposed for HypE\(^\text{18}\).
- Not very accurate but “low-cost” (computationally speaking).
- Suitable for fitness assignment in many-objective cases.

---


MARTEDA on WFG4–WFG9 — hypervolume

- MARTEDA
- MONEDA (Hyp)
- MONEDA (Pareto)
- naïve MIDEA
- MrBOA
- HypE
- SMS-EMOA
- NSGA-II

Mean perf. by problem

Mean perf. by dimension
Interesting Lines of Research
Interesting research topics

- Multi-objective estimation of distribution algorithms.
- Model inference, representation, etc.
- Multi-objective set-based optimization.
Interesting research topics

- Multi-objective estimation of distribution algorithms.
- Model inference, representation, etc.
- Multi-objective set-based optimization.

Other areas of interest

- Combining high-performance and super-computing paradigms with current evolutionary approaches.
- Links between convergence, stopping criteria and self-adaptation strategies and automatic parameter tuning methods.
- “Multi-objectivization” of machine learning algorithms.
Multi-objective EDAs

- Formalization and theoretical support of the results obtained so far.
- Impact of other learning paradigms like instance-based and match-based learning.
- Combination of information in decision and objective space.
- Model reuse across iterations.
MOEAs rely on individuals that represent candidate solutions.

Novel approach in which the search focus is moved from (independent) candidate solutions to sets of candidate solutions.

desired set properties can be expressed as a fitness function.

So far, there are sets of individuals but we could have better representations (fuzzy, Gaussian densities, etc.).

---

Set-based optimization areas

- Definition of search space and of neighbourhood relations in set-based multi-objective combinatorial optimization.
- Reformulate existing algorithms in terms of set-based multi-objective local search.
- New set-oriented local search approaches based on a quality indicators.
- Validation of the proposed algorithms on both academic, and real-world multi-objective optimization problems.
- The development of uncertainty-handling quality indicators.
- Extension of set-based local search approaches for taking the uncertainty arising in application areas into account.
Multi-objective problems are interesting from *theoretical* and *practical* reasons.

There is room inside EDAs for many interesting developments.

Set-based optimization is a promising novel area.

Machine learning methods benefit from a multi-objective approach.

Interested in practical applications.

lmarti@ele.puc-rio.br
http://www.giaa.inf.uc3m.es/miembros/lmarti – moving to PUC, eventually
HypE: An Algorithm for Fast Hypervolume-Based Many-Objective Optimization. 
TIK Report 286, Computer Engineering and Networks Laboratory (TIK), ETH Zurich.

The anticipated mean shift and cluster registration in mixture-based EDAs for multi-objective optimization.

Single objective = past, multiobjective = present, ?? = future.

An hybrid neural/genetic approach to continuous multi–objective optimization problems.
In Apolloni, B., Marino, M., and Tagliaferri, R., editors, Italian Workshop on Neural Neural Nets (WIRN), volume 2859 of Lecture Notes in Computer Science, pages 61–69. Springer.

Reidel, Boston.

Covariance matrix adaptation for multi-objective optimization.


MOEA/D: A multiobjective evolutionary algorithm based on decomposition.


Multiobjective optimization test instances for the CEC 2009 special session and competition.
Technical report, University of Essex, Colchester, UK and Nanyang Technological University, Singapore.

On set-based multiobjective optimization.