Accelerating Convergence Towards the Optimal Pareto Front

Mohsen Davarynejad, Jafar Rezaei, Jos Vrancken, Jan van den Berg, Carlos A. Coello Coello

m.davarynejad@tudelft.nl
www.tudelft.nl/mdavarynejad

CEC’11, New Orleans
June 08, 2011
Brief Introduction to Evolutionary Algorithms,
Multi-objective Optimization,
Adaptive Fuzzy Fitness Granulation,
Numerical Results,
Conclusions.
EAs: Pros and cons

- Stochastic, global search.
- No requirement for derivative information.
- Need large number of fitness evaluations.
- Not well-suitable for on-line optimization.
Overview

Multi-objective Optimization

Brief Intro ...

Adaptive Fuzzy Fitness Granulation

Results

Summery and Contribution

Objective function

Decision vector $x$

(e.g. simulation model)

Optimization Algorithm:
only allowed to evaluate $f$

Objective vector $f(x)$
EMO: Pros and cons

- Stochastic, global search,
- No requirement for derivative information,
- Need large number of fitness evaluations,
- In the case of multiple objective optimization problems, the complexity increases with the number of objectives.
### Overview

**Brief Intro**

**Adaptive Fuzzy Fitness Granulation**

**Results**

**Summary and Contribution**

---

#### Multi-objective Optimization

**TU-Delft, TBM / CINVESTAV-IPN**

Accelerating Convergence Towards the Optimal Pareto Front

---

![Diagram](image)

- **Diversity**
  - How to maintain a diverse Pareto set approximation?

- **Convergence**
  - How to guide the population towards the Pareto set?
EMO: Pros and cons

- Stochastic, global search,
- No requirement for derivative information,
- Need large number of fitness evaluations,
- In the case of multiple objective optimization problems, the complexity increases with the number of objectives.

One of the solutions is fitness approximation.
Motivations for fitness approximation

- Fitness evaluation is highly time-consuming: to reduce computation time
- Fitness is noisy: to cancel out noise
- Search for robust solutions: to avoid additional expensive fitness evaluations
Fitness Approximation Methods

- **Problem approximation**
  - To replace experiments with simulations
  - To replace full simulations/models with reduced simulations/models

- **Ad hoc methods**
  - Fitness inheritance (from parents)
  - Fitness imitation (from brothers and sisters)

- **Data-driven functional approximation (Meta-Models)**
  - Polynomials (Response surface methodology)
  - Neural networks, e.g., multilayer perceptrons (MLPs), RBFN
  - Support vector machines (SVM)
Graduation and Granulation

- The basic concepts of graduation and granulation form the core of FL and are the principal distinguishing features of fuzzy logic.
- In fuzzy logic everything is or is allowed to be graduated, or equivalently, fuzzy.
- Furthermore, in fuzzy logic everything is or is allowed to be granulated, with a granule being a clump of attribute-values drawn together by indistinguishability, similarity, proximity or functionality.
- Graduated granulation, or equivalently fuzzy granulation, is a unique feature of fuzzy logic.
- Graduated granulation is inspired by the way in which humans deal with complexity and imprecision.\*
Similarity Measure

\[ \mu_{k,r} (x_{j,r}) = \exp \left( \frac{- (x_{j,r} - c_{k,r})^2}{(\sigma_k)^2} \right), \quad k = 1, 2, \ldots, l, \quad (1) \]

\[ \bar{\mu}_{j,k} = \frac{\sum_{r=1}^{m} \mu_{k,r} (x_{j,r})}{m} \quad (2) \]
Minimum similarity threshold

\[
f (X_j^i) = \begin{cases} 
  f (C_k) & \text{if } \max_{k \in \{1,2,...,l\}} \{\mu_{j,k}\} > \theta^i , \\
  f (X_j^i) & \text{otherwise.}
\end{cases}
\]
Controls the radius of influence of each granule

A distance measurement parameter that controls the degree of similarity between two individuals.

\[ \sigma_k = \sigma_{\min} \ast ((1 - gr_{\sigma}) + gr_{\sigma} \ast \text{rank}(k)) \]  \hspace{1cm} (4)
Controlling the granule pool length and protecting new pool members

\[ L_k = \begin{cases} 
L_k + 1 & \text{if } k = K , \\
L_k & \text{otherwise} ,
\end{cases} \]  

To ensure that new granules have a good chance to survive a number of steps.
- A pool with two parts with sizes \( \varepsilon N_G \) and \((1 - \varepsilon)N_G\).
- The first part is a FIFO queue. New granules are added to this part.
- Once it grows above \( \varepsilon N_G \), then the top of the queue is moved to the other part.
- Removal from the pool takes place only in the \((1 - \varepsilon)N_G\) part.
- This is to ensure that new granules have a good chance to survive a number of steps.
The approach is compared with the standard NSGA-II (Deb, etc. 2000).

The results reported are based on the following parameter values:

- Population size = 50.
- Crossover rate = 0.9. (SBX)
- Binary tournament selection.
- Mutation rate of $1/L$, $L =$ number of decision variables.
Numerical Results

- ZDT1-6 (2 objective), a standard set of MOPs,
- 10 design variables, [0, 1],
- Algorithm performance is measured in terms of:
  - Generational Distance ($GD$): Measures how far the given solutions are on the average from the true Pareto optimal front.
  - Hypervolume indicator $I_H$ : Measures the volume of the dominated portion of the objective space and which is enclosed by the reference set.
  - Set Coverage (SC): Measures the percentage of solutions in Pareto front covered by the other set.
AFFG-NSGA-II utilized parameter values and reference points used for calculating $I_H$.

<table>
<thead>
<tr>
<th>Problem</th>
<th>$\sigma_{\text{min}}$</th>
<th>$N_G$</th>
<th>Reference point</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZDT1</td>
<td>$2^{-4}$</td>
<td>100</td>
<td>[1.1, 3.5]</td>
</tr>
<tr>
<td>ZDT2</td>
<td>$2^{-5}$</td>
<td>100</td>
<td>[1.1, 5.0]</td>
</tr>
<tr>
<td>ZDT3</td>
<td>$2^{-5}$</td>
<td>100</td>
<td>[1.1, 6.0]</td>
</tr>
<tr>
<td>ZDT4</td>
<td>$2^{-6}$</td>
<td>100</td>
<td>[1.1, 140]</td>
</tr>
<tr>
<td>ZDT6</td>
<td>$2^{-5}$</td>
<td>100</td>
<td>[1.1, 9.0]</td>
</tr>
</tbody>
</table>
Numerical Results

- 30 distinct runs of each simulation.
- Wilcoxon rank-sum test.
- Each run is restricted to 1,000 fitness function evaluations.
Mean and standard deviation of $GD$ metrics.

<table>
<thead>
<tr>
<th>Problem</th>
<th>AFFG-NSGA-II mean, $\sigma$</th>
<th>NSGA-II mean, $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZDT1</td>
<td>0.010165, 0.005744</td>
<td>0.102095, 0.029859</td>
</tr>
<tr>
<td>ZDT2</td>
<td>0.018143, 0.008509</td>
<td>0.716683, 0.365823</td>
</tr>
<tr>
<td>ZDT3</td>
<td>0.098656, 0.022421</td>
<td>0.236176, 0.048486</td>
</tr>
<tr>
<td>ZDT4</td>
<td>11.160124, 4.239201</td>
<td>20.191547, 11.658247</td>
</tr>
<tr>
<td>ZDT6</td>
<td>0.553227, 0.090989</td>
<td>0.906796, 0.102573</td>
</tr>
</tbody>
</table>
### Numerical Results

Mean and standard deviation of $I_H$ metrics.

<table>
<thead>
<tr>
<th>Problem</th>
<th>AFFG-NSGA-II $\text{mean, } \sigma$</th>
<th>NSGA-II $\text{mean, } \sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZDT1</td>
<td>$3.408204, 0.052768$</td>
<td>$2.689226, 0.164173$</td>
</tr>
<tr>
<td>ZDT2</td>
<td>$4.524421, 0.110119$</td>
<td>$2.227951, 0.350130$</td>
</tr>
<tr>
<td>ZDT3</td>
<td>$6.106243, 0.198963$</td>
<td>$4.516725, 0.267211$</td>
</tr>
<tr>
<td>ZDT4</td>
<td>$130.929275, 10.401000$</td>
<td>$122.559419, 9.497713$</td>
</tr>
<tr>
<td>ZDT6</td>
<td>$4.817657, 0.327744$</td>
<td>$2.565694, 0.518935$</td>
</tr>
</tbody>
</table>
Numerical Results

Mean and standard deviation of SC metrics.

<table>
<thead>
<tr>
<th>Problem</th>
<th>AFFG-NSGA-II mean, $\sigma$</th>
<th>NSGA-II mean, $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZDT1</td>
<td>1.000000, 0.000000</td>
<td>0.000000, 0.000000</td>
</tr>
<tr>
<td>ZDT2</td>
<td>1.000000, 0.000000</td>
<td>0.000000, 0.000000</td>
</tr>
<tr>
<td>ZDT3</td>
<td>0.995745, 0.023307</td>
<td>0.003401, 0.018630</td>
</tr>
<tr>
<td>ZDT4</td>
<td>0.613805, 0.455574</td>
<td>0.324147, 0.427815</td>
</tr>
<tr>
<td>ZDT6</td>
<td>0.725224, 0.167394</td>
<td>0.113340, 0.123955</td>
</tr>
</tbody>
</table>
Numerical Results

ZDT1

TU-Delft, TBM / CINVESTAV-IPN

Accelerating Convergence Towards the Optimal Pareto Front
Numerical Results

ZDT1
Numerical Results

ZDT2

Accelerating Convergence Towards the Optimal Pareto Front
Numerical Results

ZDT2

TU-Delft, TBM / CINVESTAV-IPN
Accelerating Convergence Towards the Optimal Pareto Front
Numerical Results

ZDT3
Numerical Results

ZDT3

Accelerating Convergence Towards the Optimal Pareto Front
Numerical Results

ZDT4
Numerical Results

ZDT4
Numerical Results

ZDT6
Numerical Results

ZDT6
ZDT1 with a budget of 5,000 fitness function evaluations
Summery

- Fitness approximation Replaces an accurate, but costly function evaluation with approximate, but cheap function evaluations.
- Meta-modeling and other fitness approximation techniques have found a wide range of applications.
- Proper control of meta-models plays a critical role in the success of using meta-models.
Contribution

Why AFFG?

- Avoids initial training,
- Uses guided association to speed up search process,
- Gradually sets up an independent model of initial training data to compensate the lack of sufficient training data and to reach a model with sufficient approximation accuracy.
- Avoids the use of model in unrepresented design variable regions in the training set.
L3S Problem

- Lot Sizing with Supplier Selection,
  - Lot Sizing: The process of determining the size of the order quantities for each component in a product in each time period,
  - Supplier Selection: The process of identifying, evaluating and contracting with suppliers.

- Objectives,
  - Cost,
  - Reliability.
NDP Problem

- Network Design Problem,

- Objectives,
  - Efficiency,
  - Climate: Total emission CO2,
  - Noise: Weighted average sound power level.
Accelerating Convergence Towards the Optimal Pareto Front

Mohsen Davarynejad, Jafar Rezaei, Jos Vrancken, Jan van den Berg, Carlos A. Coello Coello

m.davarynejad@tudelft.nl
www.tudelft.nl/mdavarynejad

CEC’11, New Orleans
June 08, 2011