Co-evolution improves the efficiency of preference learning methods when the Decision Maker's aspirations develop over time

Michał Tomczyk **Miłosz Kadziński**

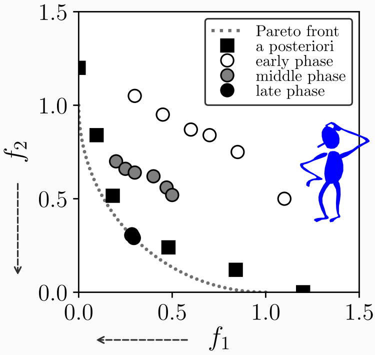
Institute of Computing Science
Poznan University of Technology, Poland

michal.tomczyk@cs.put.poznan.pl
www.cs.put.poznan.pl/mtomczyk/

Introduction

The work concerns preference-based evolutionary multiple-objective optimization. The involvement of the DM's aspirations in the search is beneficial twofold:

- The evolutionary speed can be improved due to narrowing the search space.
- Ultimately, solutions of much better DM-perceived quality can be constructed.



Introduction

The preference-based evolutionary hybrids are joint works of the fields of:

- Evolutionary Multiple-objective Optimization, and
- Multiple-Criteria Decision Analysis.

When developing new methods**, state-of-the-art postulates from both streams should be employed** to ensure the efficiency and validity of proposals, yielding a significant contribution to the literature in this way.

Evolutionary Multiple-objective Optimization (EMO)

Multiple-Criteria Decision Analysis (MCDA)

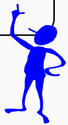
Introduction

Preference learning

<u>Preference learning:</u> Learning global model parameters of the assumed preference model using incomplete data, i.e., the DM's limited feedback. The learning process is often referred to as **preference disaggregation**.

It can be viewed as a subdiscipline of machine learning, but having much stronger assumptions:

1. The employed preference model should not be a black box but should have good explanatory properties. we
2. The methods should minimize the DM's cognitive effort.



Doumpos, M., Zopounidis, C. (2014). The Robustness Concern in Preference Disaggregation Approaches for Decision Aiding: An Overview. In: Rassias, T., Floudas, C., Butenko, S. (eds) Optimization in Science and Engineering. Springer, New York, NY. https://doi.org/10.1007/978-1-4939-0808-0_8

Example: Assume the DM's value system aligns well with a weighted Chebyshev function:

$$\min f(\boldsymbol{s}) = max\,\{w_1 s_1, w_2 s_2, \dots, w_M s_M\}$$

Hence, use **it as a preference model**. If you combine EMO with MCDA, **every component now has two perspectives**:

- **EMO**: the model can be used to assess solutions in the population.
- **MCDA**: The preference model is a mathematical formalism that should well reflect how the DM evaluates solutions.

The major model parameters are here objective weights.

Naive approach: ask the DM's for the weights. **Instead**, learn them via a preference disaggregation process.

Preference disaggregation: The DM's feedback is often represented in a natural, easy-to-comprehend form, e.g., pairwise comparisons of solutions (holistic, indirect judgments).

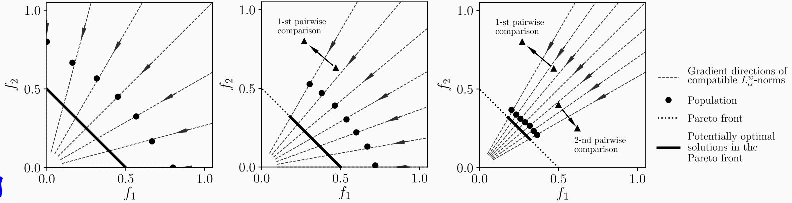
Example: Consider a pair $\boldsymbol{s}^a = [0.6, 0.4]$ and $\boldsymbol{s}^b = [0.4, 0.6]$ and assume that the stated that (s)he prefers $\boldsymbol{s}^a$ over $\boldsymbol{s}^b$, denoted by $\boldsymbol{s}^a \succ \boldsymbol{s}^b$. Preference disaggregation often refers to **a process of exploiting model parameters compatible with the DM's feedback:**

$$\boldsymbol{w} \text{ is compatible iff } \max\{w_1 s_1^a,\ w_2 s_2^a\} < \max\{w_1 s_1^a,\ w_2 s_2^a\}, \forall_{\boldsymbol{s}^a \succ \boldsymbol{s}^b}\ DM's\ statement$$

Naturally, there may be (infinitely) many such compatible model parameters. **There is a plethora of procedures in MCDA** for exploiting them and deriving recommendations. **Robust methods** take into account a whole space of compatible model parameters.

IEMO/D

- **Evolutionary framework:** IEMO/D is based on the MOEA/D algorithm.
- **Preference elicitation scheme:** Interactive, based on pairwise comparisons of solutions
- **Preference representation:** Represents the DM's by using a fine representation of the space of compatible L-norms.
- **Preference learning & Robustness concern:** After each preference elicitation, this set is incorporated into the decomposition-based framework as optimization goals (**potential optimality**).

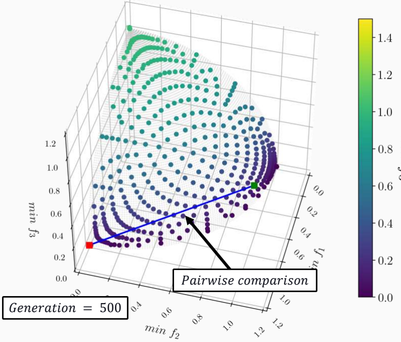


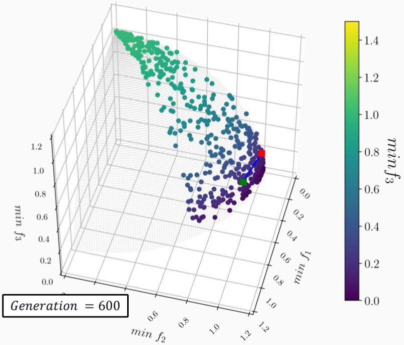
***IEMO/D:** M. K. Tomczyk and M. Kadziński. Decomposition-based interactive evolutionary algorithm for multiple objective optimization. IEEE Transactions on Evolutionary Computation 24, 2 (2020), 320–334*

Introduction
IEMO/D

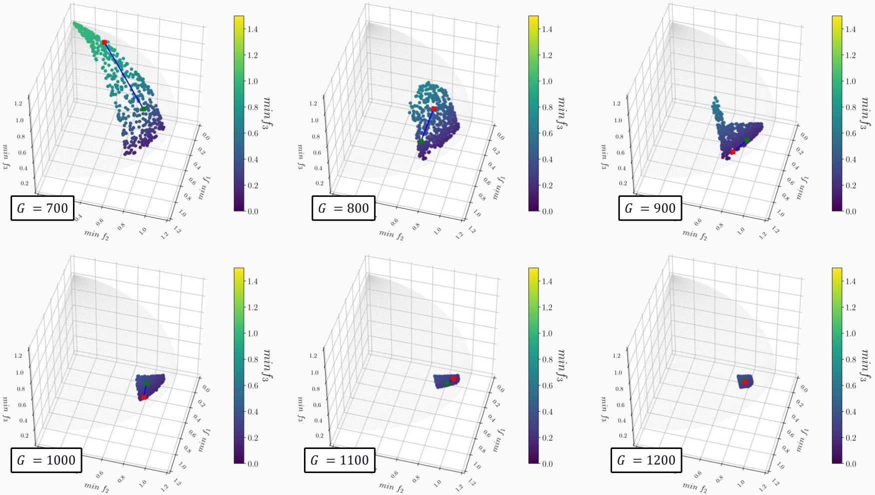
Example search:

- Assumed preference model: Weighted Chebyshev function
- Exhaustive search (surplus computational resources): $\textit{population size} = 496$ **for a more readable presentation**.
- **Pairs selection policy**: PWIT (maximizes the information-gain from the DM's feedback).
- DM's answers were simulated using a WCF with $\boldsymbol{w} = [0.3, 0.2, 0.5]$.





EMOSOR: *M. K. Tomczyk and M. Kadziński. EMOSOR: Evolutionary multiple objective optimization guided by interactive stochastic ordinal regression. Computers & Operations Research 108, 2019, 134–154*



$G = 700$
$G = 800$
$G = 900$
$G = 1000$
$G = 1100$
$G = 1200$
$min\ f_1$
$min\ f_2$
$min\ f_3$

Research question

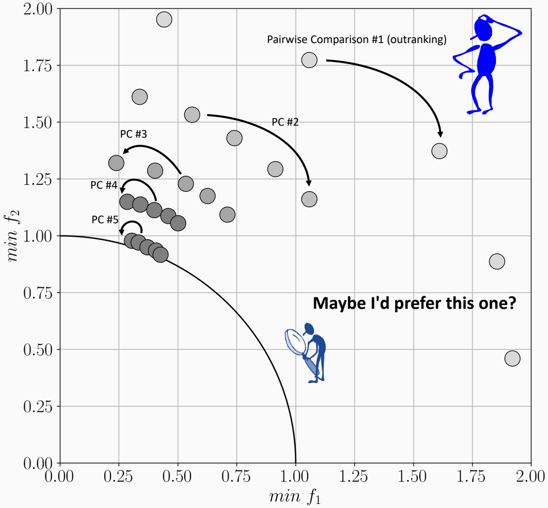
Incorporation of the preference-learning paradigm into evolutionary search raises many new essential research questions:

- How to properly select a preference model to simultaneously ensure a proper convergence and good relevance with the DM's aspirations?
- How to deal with potential inconsistency caused by (a) an improperly selected model or (b) DM's irrationality?
- When to trigger the interactions, and which solutions should be selected for the DM for the critical evaluation to maximize the information gained from their feedback?
- How to handle the DM's will to redefine their aspirations during the search?
- And many more...

Motivation

The DM's aspirations may not be well-established from the beginning. In an interactive process, (s)he may begin better refining them, though. Consequently, the DM may:

- change their answering policy,
- change their previous feedback.



Motivation

Consequences/goals:

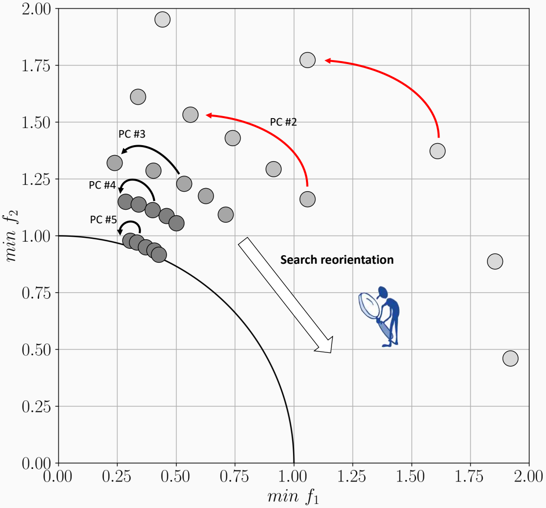
EMO: Re-orient the search

MCDA: Re-learn the DM's aspirations

Problems:

EMO: Shifting the population elsewhere may be computationally expensive.

MCDA: Now, the population consists of (probably) irrelevant solutions to the DM. The efficiency of the preference learning process depends on the solutions presented to the DM for evaluation. Presenting them irrelevant solutions for comparison is a bad idea.



The proposal

Proposal (extension to IEMO/D):

Divide (constant computational resources) the population into two species (co-evolution):

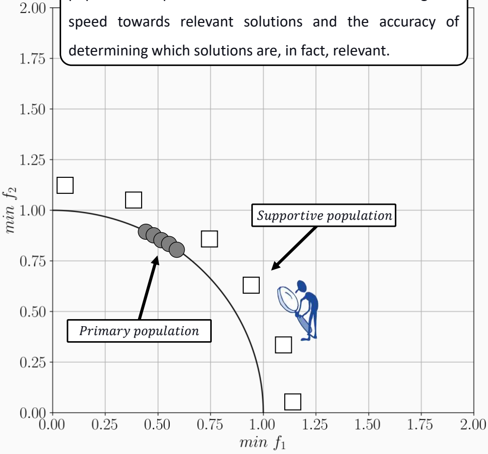
- Primary: follow DM's aspirations
- Supportive: approximate the PF.

Reason: maintain solutions representing a broader spectrum of trade-offs between objectives.

Benefits:

- **EMO:** Search re-orientation does not need to start "from scratch."
- **MCDA:** a better pair of solutions for comparison can be selected from the population.

The proportions between the primary and supportive populations represent a trade-off between the convergence speed towards relevant solutions and the accuracy of determining which solutions are, in fact, relevant.

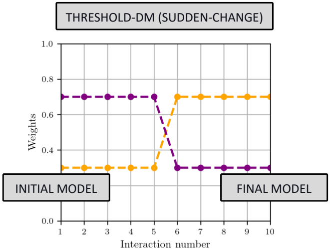


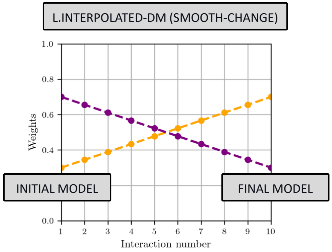
The artificial DM

In the experiments, the DM's answers were simulated using the WCF:

$$\min f(\boldsymbol{s}) = max\,\{w_1 s_1, w_2 s_2, \ldots, w_M s_M\}.$$

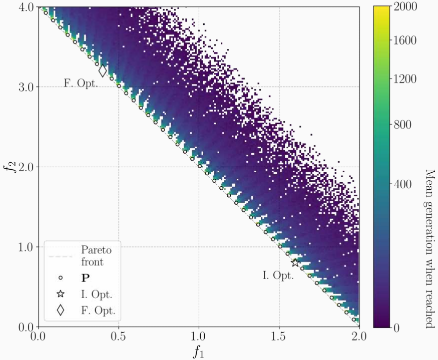
We assumed that the **DM's weights (WCF) could change dynamically to simulate the DM's preference system changing in time**. For this reason, we implemented **two models** of the dynamic artificial DM, whose weights are expressed as a function of the interaction number. **Anytime the DM's value system changes**, the already provided pairwise comparisons are revisited and, possibly, reversed. Further, the DM's preferences are disaggregated again.

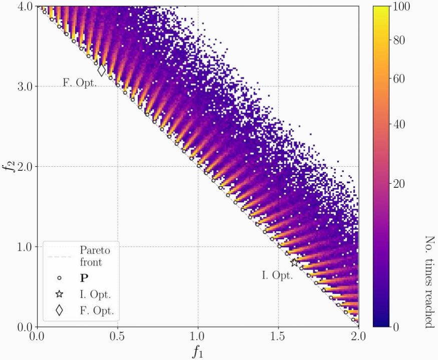




MOEA/D

- Results involve 100 independent runs of the method
- Discretization: 200 x 200.





***ViPEMO:** Miłosz Kadziński, Michał K. Tomczyk, Roman Słowiński, Preference-based cone contraction algorithms for interactive evolutionary multiple objective optimization, Swarm and Evolutionary Computation, Volume 52, 2020, 100602,*

Example results

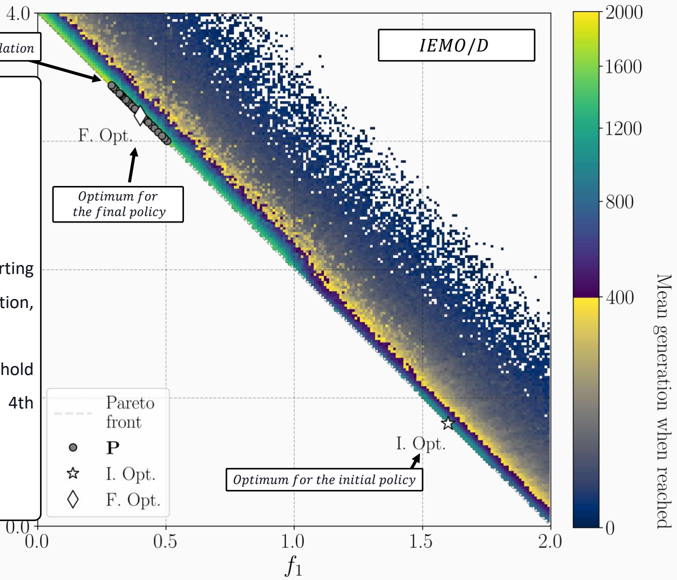
- DM's initial policy:

  $$\boldsymbol{w} = [0.2, 0.8]$$

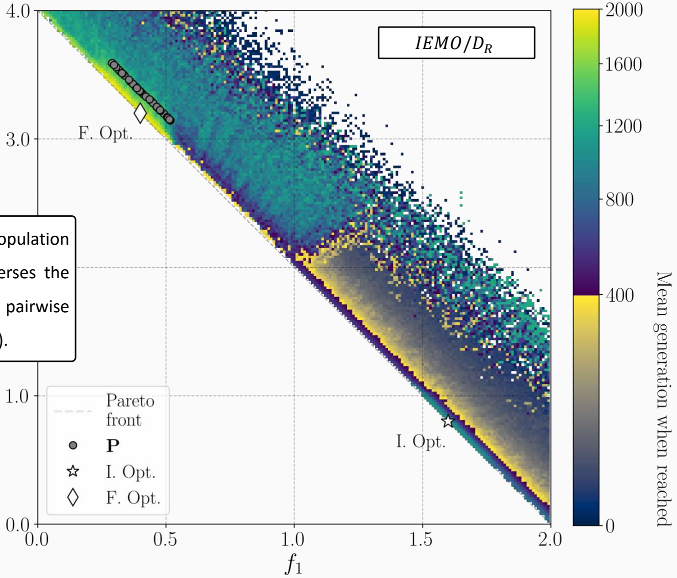
- DM's final policy:

  $$\boldsymbol{w} = [0.8, 0.2]$$

- 2000 generations
- Interactions: 6 (starting from the 400th generation, equally spaced)
- Changing policy: threshold (change just before 4th interaction)
- Population size: 45.
- WFG3

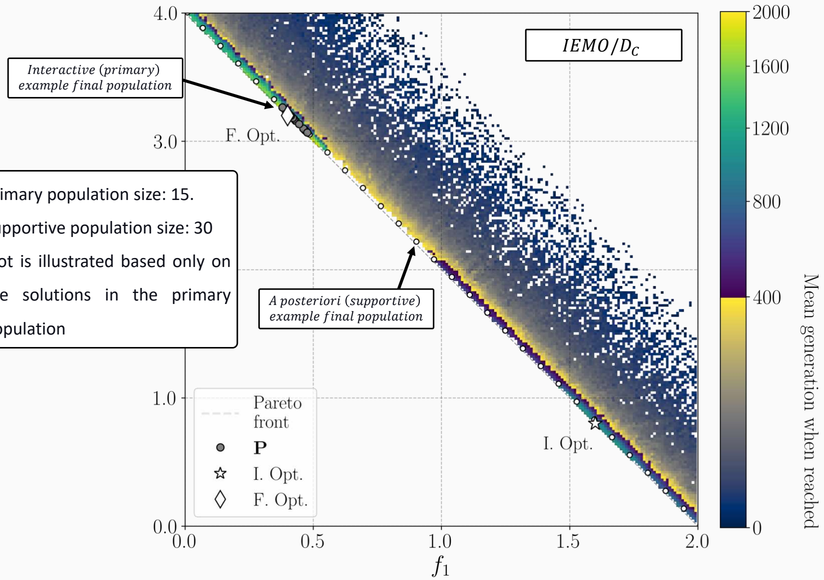


Example results



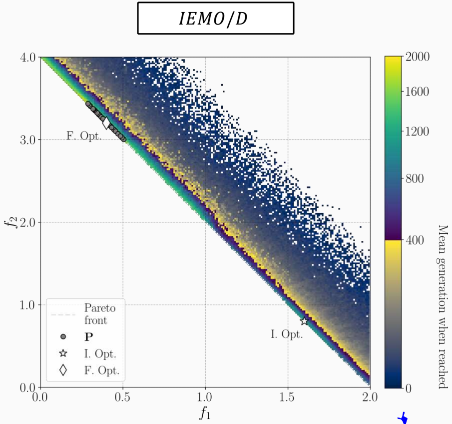
$IEMO/D_R$: **restart** the population every time the DM reverses the already provided pairwise comparison (**benchmark**).

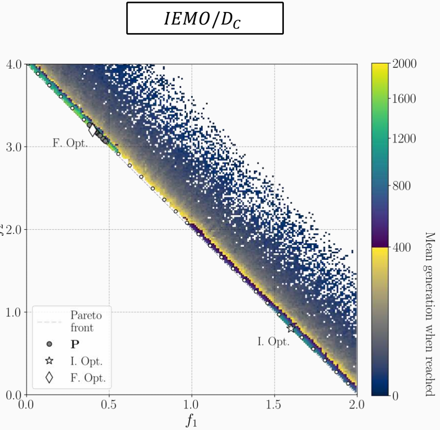
Example results



- Primary population size: 15.
- Supportive population size: 30
- Plot is illustrated based only on the solutions in the primary population

Example results





Example results

Extra: splitting the original population into primary and supportive:

- **Increases** efficiency of the re-learning and re-orienting processes.
- However, it **diminishes** the efficiency of the convergence.

Solution: dynamic resource reallocation (dynamic mode). The DM can label the pairwise comparisons as "uncertain" or "certain." If the answer is sure, the method can remove those scalar model instances that prove incompatible with the DM's answer and increase the primary population size by the same amount.

Example results

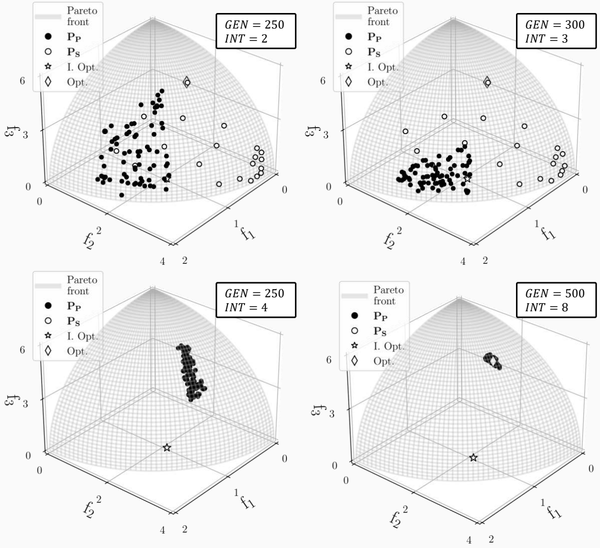
- DM's initial policy:

  $$\boldsymbol{w} = [0.2, 0.2, 0.6]$$

- DM's final policy:

  $$\boldsymbol{w} = [0.5, 0.3, 0.2]$$

- 500 generations
- Interactions: 8 (starting from the 120th generation, equally spaced)
- Changing policy: threshold (change just before 4th interaction)
- Population size: 96 (30 = primary; 66 = supportive).
- WFG4



Example results

Summary

- The work addresses the problem of potential DM's will to redefine their preferences during the interactive EMO search in the context of **preference learning**.
- The proposal: Interactive method that runs in the co-evolutionary mode. There are two populations: **primary** (responsible for converging in line with the DM's learned aspirations) and **supportive** (provides a broader spectrum of solutions).
- Extra: The method can run in the dynamic mode, allowing it to neglect the effect of diminishing convergence speed.
- Both concepts proved their usefulness during a series of extensive experiments.

Avenues for future research

- We will investigate the performance of the proposed method when coupled with different, potentially more flexible preference models.
- We will design other dynamic procedures for reallocating resources that would not rely on additional information provided by the DM. E.g., we could check if a supportive solutional already well approximates the PF and, if so, remove it and store it in the archive.
- We will examine the usability of the proposed method by involving real DMs.

Thank you for your attention!

michal.tomczyk@cs.put.poznan.pl