# IFORS2023: Comparing interactive multiobjective optimization methods with artificial decision-makers based on computational rationality

**Giovanni Misitano**[1] Bekir Afsar[1] Jussi Jokinen[1] Kaisa Miettinen[1]

[1]University of Jyväskylä (Finland), Faculty of Information Technology

July 10, 2023

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

# Giovanni Misitano

- I am Giovanni Misitano, a doctoral researcher from the Multiobjective Optimization Group at the University of Jyväskylä (Finland).
- The topic of my thesis is *Explainable Multiobjective Optimization*.
- President of the Finnish Operations Research Society (FORS)

# Overview

# Motivation I

- Preferences provided by a decision maker play a key role.
- Decision makers and preferences are subjective.
- Success measured in ability to support decision makers.

# Motivation II

Two important questions arise:

1. How can we compare interactive methods?
2. How can we support different decision makers?

We will combine ideas from cognitive science and reinforcement learning to begin address these questions.

For a general introduction, see[1].

## A Multiobjective Optimization Problem

$$\min_{\mathbf{x}\in S} \mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_k(\mathbf{x})), \tag{1}$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ is a *decision variable vector* consisting of $n$ decision variables; $S \subset \mathbb{R}^n$ is the *set of feasible decision variables*; and $f_i(\mathbf{x}) : S \to \mathbb{R} (i = 1, \dots, k)$ are conflicting *objective functions* to be minimized.

[1] Kaisa Miettinen. *Nonlinear multiobjective optimization*. Boston: Kluwer Academic Publishers, 1999.

# Multiobjective optimization II

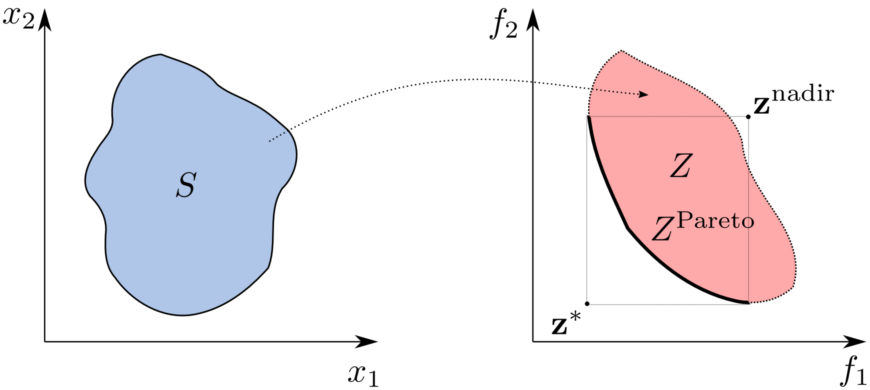
## Pareto optimality and the set of Pareto optimal solutions

If $\mathbf{x}\ $ is a Pareto optimal solution to (1), then there does not exist another solution $\mathbf{x}$ such that $f_i(\mathbf{x}) \quad f_i(\mathbf{x}\ )$ for *all* $i = 1; : : : ; k$, and $f_i(\mathbf{x}) < f_i(\mathbf{x}\ )$ for *some* $i = 1; : : : ; k$.

The set of all Pareto optimal solutions to (1) is known as the *set of Pareto optimal solutions*. The image of this set is $Z^{\text{Pareto}}$, and it is a subset of the image of the feasible set $Z$, i.e., $\mathbf{f} : S \; ! \; Z$ and $Z^{\text{Pareto}} \quad Z$.

# Multiobjective optimization III

$$\boxed{\min_{\mathbf{x}\in S} \mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}))}$$



# Interactive methods I

- Pareto optimal solutions cannot be fully compared.
- A decision maker (a domain expert) provides preferences.
- Multiobjective optimization methods utilize preferences to find the most preferred solution.

# Interactive methods II

- Interactive methods: preferences provided iteratively during optimization.
- Preference type example: *reference point* $\bar{\mathbf{z}} \in \mathbb{R}^k$ consisting of *aspiration levels*.
- Preferences change as the decision maker *constructs* (or *learns*) part of their preferences.

# Interactive methods III

# Arti cial decision makers

When decision makers are not available, how to study interactive methods?

Arti cial decision makers have been employed to compare interactive methods to some extent[2].

Do not capture human aspects, such as preference construction and cognitive limitatio such as memory.

How to model interaction between a decision maker and an interactive method?

[2] Bekir Afsar, Kaisa Miettinen, and Francisco Ruiz. \Assessing the performance of interactive multiobjective optimization methods: A survey". 54.4 (2021). doi : 10.1145/3448301 .

# Computational rationality I

Computational rationality: model human interaction by modeling its limitations.[3]

Find an optimal policy utilizing a reinforcement learning agent[4] that describes optimal interaction under the limitations assumed.

Interaction modeled as a partially observable Markov decision process.

This approach has been utilized successfully in modeling driving a car[5][6] and typing on a touchscreen keyboard[7], for example.

[3] Antti Oulasvirta, Jussi P. P. Jokinen, and Andrew Howes. "Computational Rationality as a Theory of Interaction". en. CHI Conference on Human Factors in Computing Systems. ACM, 2022, pp. 1–14. doi: 10.1145/3491102.3517739.

[4] Richard S Sutton and Andrew G Barto. Reinforcement learning: An introduction. MIT press, 2018.

[5] Jussi P. P. Jokinen, Tuomo Kujala, and Antti Oulasvirta. "Multitasking in Driving as Optimal Adaptation Under Uncertainty". Human Factors: The Journal of the Human Factors and Ergonomics Society 63.8 (2021), pp. 1324–1341. doi: 10.1177/0018720820927687.

[6] Jussi P. P. Jokinen and Tuomo Kujala. "Modelling Drivers' Adaptation to Assistance Systems". 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. ACM, 2021, pp. 12–19. doi: 10.1145/3409118.3475150.

[7] Jussi Jokinen, Aditya Acharya, Mohammad Uzair, Xinhui Jiang, and Antti Oulasvirta. "Touchscreen Typing As Optimal Supervisory Control". Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. ACM, 2021, pp. 1–14. doi: 10.1145/3411764.3445483.

# Computational rationality II

# Computational rationality III

- But when modeling interaction between a decision maker and an interactive method, what kind of limitations should we assume?
- For the purpose of this study, we have chosen two limiting factors: *preferences* and *memory*.

**?**

# Preferences I

- Many theories exist on how preferences can be modeled in a decision maker.
- In our work, we have assumed that preferences are *constructed* during the decision process (the interactive process)[8].
- Moreover, we have assumed that preferences are constructed from memory.

[8] Sarah Lichtenstein and Paul Slovic, eds. *The construction of preference*. Cambridge, England: Cambridge University Press, 2006.

- Decision makers prefer what is familiar to them.
- *Similarity*[9] between two objective vectors $\mathbf{z}^1$ and $\mathbf{z}^2$:

## Similarity

$$S(\mathbf{z}^1, \mathbf{z}^2) = S^{(1,2)} = \exp\left( -\lambda \sum_{j=1}^{k} w_j \left| z_j^1 - z_j^2 \right| \right), \tag{2}$$

where $\lambda > 0$ is the *discriminability* of the solutions being compared, and $0 \leq w_j \leq 1$ is the proportion of *attention* given to the $j^{\text{th}}$ objective value. We assume that $\sum_{j=1}^{k} w_j = 1$.

[9] Jana B. Jarecki and Jörg Rieskamp. "Comparing attribute-based and memory-based preferential choice". *DECISION* 49.1 (2022), pp. 65–90.

- Similarity (2) is used to compute the *quality* of an objective vector $\mathbf{z}$ by comparing it to memory:

**Quality based on similarity**

$$V(\mathbf{z}^i) = V^i = \frac{\sum_e S^{(i,e)} V^e}{\sum_e S^{(i,e)}}, \tag{3}$$

where $e$ represents the other objective vectors, e.g., in memory, the objective vector $\mathbf{z}^i$ is compared to, excluding the vector $i$ itself.

- Quality: how much the decision maker prefers an objective vector (the higher, the better).

# Memory I

- But what about memory?
- Assume a *retrograde model* memory: old information is replaced by new one as it presents itself[10,11].
- Memory is limited. Max $m$ objective vectors in memory at any time.

[10] M. Katkov, S. Romani, and M. Tsodyks. “Memory Retrieval from First Principles”. *Neuron* 94.5 (2017), pp. 1027–1032.

[11] Mikhail Katkov, Michelangelo Naim, Antonios Georgiou, and Misha Tsodyks. “Mathematical models of human memory”. *Journal of Mathematical Physics* 63.7 (2022), p. 073303.

# Memory II

## Simple memory model

Suppose $M$ is a set of objective vectors such that $|M| < m$, where $m$ is the size of our memory. Then, a simple memory model based on the principle of retrograde memory and similarity looks like:

$$
\begin{aligned}
&\text{if } |M| < m \text{ then } M \leftarrow M \cup \{\mathbf{z}^{\text{new}}\}, \\
&\text{else if } |M| = m \text{ then } M \leftarrow (M \setminus \{\mathbf{z}^{i}\}) \cup \{\mathbf{z}^{\text{new}}\}, \\
&\quad \text{where } i = \operatorname{argmax}_{j \in [1,m]} S(\mathbf{z}^{j}, \mathbf{z}^{\text{new}}) .
\end{aligned}
\tag{4}
$$

- When memory is not full, add new observation $\mathbf{z}^{\text{new}}$ to it.
- When memory is full, replace most similar object with $\mathbf{z}^{\text{new}}$.

# Memory III

- Humans tend to “clump” similar things together in their memory.
- We can assume certain objective vectors to always stay in memory ! steady part of preferences.
- Steady part can represent domain expertise.

# Current state of research

# What we are currently doing

- Reinforcement learning agent based on the preference (3) and memory models (4).
- Agent interacts with a reference point based interactive multiobjective optimization methods.
- Can work with other types of preferences.
- Experimenting with different reward functions.

# Our reinforcement learning agent I

# Our reinforcement learning agent II

# Our reinforcement learning agent III

# Our reinforcement learning agent IV

# Our reinforcement learning agent V

# Our reinforcement learning agent VI

# Our reinforcement learning agent VII

# Our reinforcement learning agent VIII

# Our reinforcement learning agent IX

Agent implemented in Python, gymnasium[12] and stable-baselines3[13].

Finding optimal policies: Proximal Policy Approximation[14].

DESDEO[15] to model multiobjective optimization problems and to access interactive methods.

Agent applied with the Reference Point Method[16].

---

[12] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. \Openai gym". arXiv preprint arXiv:1606.01540 (2016).

[13] Antonin Ra n, Ashley Hill, Adam Gleave, Anssi Kanervisto, Maximilian Ernestus, and Noah Dormann. \Stable-Baselines3: Reliable Reinforcement Learning Implementations". Journal of Machine Learning Research 22.268 (2021), pp. 1{8.

[14] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. \Proximal policy optimization algorithms". arXiv preprint arXiv:1707.06347 (2017).

[15] G. Misitano, B. S. Saini, B. Afsar, B. Shavazipour, and K. Miettinen. \DESDEO: The Modular and Open Source Framework for Interactive Multiobjective Optimization". IEEE Access 9 (2021), pp. 148277{148295.

[16] Andrzej P Wierzbicki. \A MATHEMATICAL BASIS FOR SATISFICING DECISION MAKING". Mathematical Modelling 3 (1982), p. 15.

# An optimal policy I

# An optimal policy II

- By studying found policies, we can gather valuable information on how di erent interactive methods can help decision makers.
- Can policies be used to compare interactive methods?
- Can policies be used as a tools to assist decision makers during an interactive proces

# An optimal policy III

Currently, the challenge is in finding a reward function that leads to a sensible policy. From such a policy, we can expect:

- If an agent has initially an objective vector(s) in its memory that exists, or is very near, a Pareto optimal objective vector, then the agent should converge in very few iterations.
- If an agent has initially an objective vector(s) in its memory that is not achievable and requires trading-off, we expect a longer deliberation before the agent decides to stop.
- Ideally, we could be able to identify a decision phase and a learning phase in an agent following an optimal policy[17].

[17] Jurgen Branke, Jürgen Branke, Kalyanmoy Deb, Kaisa Miettinen, and Roman Słowiński. Multiobjective optimization: Interactive and evolutionary approaches. Vol. 5252. Springer Science & Business Media, 2008.

# An optimal policy IV

But how good is an optimal policy?

We may only compare it to what we know from the literature when it comes to how decision makers may act during an interactive method.

We need further studies to compare the policies to how real decision makers act.

# Conclusions

# Conclusions

- New kind of arti cial decision maker based on computational rationality and reinforcement learning.
- Limited by memory and preferences.
- Optimal policies can be useful for comparing interactive methods and to support decis makers.
- Studies comparing our arti cial decision maker and real decision makers needed next.

# Let us network!

- My homepage: http://giovanni.misitano.xyz .
- Social media:
  - LinkedIn (linkedin.com/in/misitano),
  - Twitter (@misitano_g), and
  - GitHub (gialmisi ).
- DESDEO: https://desdeo.it.jyu.fi/
- The Multiobjective Optimization (research) Group: http://www.mit.jyu.fi/optgroup/

[1] Kaisa Miettinen. Nonlinear multiobjective optimization. Boston: Kluwer Academic Publishers, 1999.

[2] Bekir Afsar, Kaisa Miettinen, and Francisco Ruiz. \Assessing the performance of interactive multiobjective optimization methods: A survey". 54.4 (2021). doi : 10.1145/3448301.

[3] Antti Oulasvirta, Jussi P. P. Jokinen, and Andrew Howes. \Computational Rationality a a Theory of Interaction". en. CHI Conference on Human Factors in Computing System ACM, 2022, pp. 1{14. doi : 10.1145/3491102.3517739.

[4] Richard S Sutton and Andrew G Barto. Reinforcement learning: An introduction. MIT press, 2018.

# Bibliography II


[5] Jussi P. P. Jokinen, Tuomo Kujala, and Antti Oulasvirta. \Multitasking in Driving as Optimal Adaptation Under Uncertainty". Human Factors: The Journal of the Human Factors and Ergonomics Society 63.8 (2021), pp. 1324{1341. doi: 10.1177/0018720820927687.

[6] Jussi P. P. Jokinen and Tuomo Kujala. \Modelling Drivers' Adaptation to Assistance Systems". 13th International Conference on Automotive User Interfaces and Interactiv Vehicular Applications. ACM, 2021, pp. 12{19. doi: 10.1145/3409118.3475150.

[7] Jussi Jokinen, Aditya Acharya, Mohammad Uzair, Xinhui Jiang, and Antti Oulasvirta. \Touchscreen Typing As Optimal Supervisory Control". Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. ACM, 2021, pp. 1{14. doi: 10.1145/3411764.3445483.

[8] Sarah Lichtenstein and Paul Slovic, eds. The construction of preference. Cambridge, England: Cambridge University Press, 2006.

[9] Jana B. Jarecki and Jörg Rieskamp. "Comparing attribute-based and memory-based preferential choice". *DECISION* 49.1 (2022), pp. 65–90.

[10] M. Katkov, S. Romani, and M. Tsodyks. "Memory Retrieval from First Principles". *Neuron* 94.5 (2017), pp. 1027–1032.

[11] Mikhail Katkov, Michelangelo Naim, Antonios Georgiou, and Misha Tsodyks. "Mathematical models of human memory". *Journal of Mathematical Physics* 63.7 (2022), p. 073303.

[12] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. "Openai gym". *arXiv preprint arXiv:1606.01540* (2016).

[13] Antonin Raffin, Ashley Hill, Adam Gleave, Anssi Kanervisto, Maximilian Ernestus, and Noah Dormann. "Stable-Baselines3: Reliable Reinforcement Learning Implementations". *Journal of Machine Learning Research* 22.268 (2021), pp. 1–8.

# Bibliography IV


[14] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. "Proximal policy optimization algorithms". *arXiv preprint arXiv:1707.06347* (2017).

[15] G. Misitano, B. S. Saini, B. Afsar, B. Shavazipour, and K. Miettinen. "DESDEO: The Modular and Open Source Framework for Interactive Multiobjective Optimization". *IEEE Access* 9 (2021), pp. 148277–148295.

[16] Andrzej P Wierzbicki. "A MATHEMATICAL BASIS FOR SATISFICING DECISION MAKING". *Mathematical Modelling* 3 (1982), p. 15.

[17] Jurgen Branke, Jürgen Branke, Kalyanmoy Deb, Kaisa Miettinen, and Roman Słowiński. *Multiobjective optimization: Interactive and evolutionary approaches*. Vol. 5252. Springer Science & Business Media, 2008.