

# Personalized Multiobjective Optimization: An Analytics Perspective

Edited by

Kathrin Klamroth<sup>1</sup>, Joshua D. Knowles<sup>2</sup>, Günter Rudolph<sup>3</sup>, and Margaret M. Wiecek<sup>4</sup>

1 Universität Wuppertal, DE, [klamroth@math.uni-wuppertal.de](mailto:klamroth@math.uni-wuppertal.de)

2 University of Birmingham, GB, [j.knowles@cs.bham.ac.uk](mailto:j.knowles@cs.bham.ac.uk)

3 Technische Universität Dortmund, DE, [guenter.rudolph@tu-dortmund.de](mailto:guenter.rudolph@tu-dortmund.de)

4 Clemson University, US, [wmalgor@clemson.edu](mailto:wmalgor@clemson.edu)

---

## Abstract

The Dagstuhl Seminar 18031 Personalization in Multiobjective Optimization: An Analytics Perspective carried on a series of five previous Dagstuhl Seminars (04461, 06501, 09041, 12041 and 15031) that were focused on Multiobjective Optimization. The continuing goal of this series is to strengthen the links between the Evolutionary Multiobjective Optimization (EMO) and the Multiple Criteria Decision Making (MCDM) communities, two of the largest communities concerned with multiobjective optimization today. Personalization in Multiobjective Optimization, the topic of this seminar, was motivated by the scientific challenges generated by personalization, mass customization, and mass data, and thus crosslinks application challenges with research domains integrating all aspects of EMO and MCDM. The outcome of the seminar was a new perspective on the opportunities as well as the research requirements for multiobjective optimization in the thriving fields of data analytics and personalization. Several multi-disciplinary research projects and new collaborations were initiated during the seminar, further interlacing the two communities of EMO and MCDM.

**Seminar** January 14–19, 2018 – <http://www.dagstuhl.de/18031>

**2012 ACM Subject Classification** Mathematics of computing → Mathematical optimization, Applied computing → Operations research, Computing methodologies → Artificial intelligence

**Keywords and phrases** multiple criteria decision making, evolutionary multiobjective optimization

**Digital Object Identifier** 10.4230/DagRep.8.1.33

**Edited in cooperation with** Richard Allmendinger

## 1 Executive Summary

*Kathrin Klamroth*

*Joshua D. Knowles*

*Günter Rudolph*

*Margaret M. Wiecek*

**License** © Creative Commons BY 3.0 Unported license

© Kathrin Klamroth, Joshua D. Knowles, Günter Rudolph, and Margaret M. Wiecek

The topic of the seminar, Personalization in Multiobjective Optimization, was motivated by ongoing changes in many areas of human activity. In particular, personalization, mass



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 3.0 Unported license

Personalized Multiobjective Optimization: An Analytics Perspective, *Dagstuhl Reports*, Vol. 8, Issue 01, pp. 33–99

Editors: Kathrin Klamroth, Joshua D. Knowles, Günter Rudolph, and Margaret M. Wiecek



Dagstuhl Reports

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

■ **Table 1** Working groups (WGs) crosslinking application challenges (rows) with research domains (columns). WG 1: Preference uncertainty quantification; WG 2: Personalization and customization of decision support; WG 3: Invariant rule extraction; WG 4: Complex networks and MCDA; WG 5: Metamodelling for interactive optimization.

	Modelling	Preferences	Algorithms
Platform design and product lines	WG3, WG5	WG1, WG3	WG3
Responsive and online personalization	WG2, WG5	WG1, WG2	WG2
Complex networks of decision makers	WG4, WG5	WG1, WG4	WG4

customization, and mass data have become essential in current business and engineering operations creating new challenges for academic and research communities. In the seminar, the EMO and MCDM communities, including junior and senior academic researchers as well as industry representatives, took an effort to jointly address the ongoing changes in the real-world with multiobjective optimization.

The purpose of multiobjective optimization is to develop methods that can solve problems having a number of (conflicting) optimization criteria and constraints, providing a multitude of solution alternatives, rather than pursuing only one “optimal” solution. In this aim the field has been highly successful: its methods have a track record of improving decision making across a broad swath of applications, indeed wherever there are conflicting goals or objectives. Yet, multiobjective optimization has so far focused almost exclusively on serving a single “decision maker”, providing solutions merely as potential (not actual) alternatives. In order to fulfill the demanding aims of mass-customization, product/service variation and personalization we see today in areas such as engineering, planning, operations, investment, media and Web services, and healthcare, new and innovative approaches are needed. This seminar took the first steps towards this goal by bringing together leading specialists in EMO and MCDM.

Personalization in multiobjective optimization as the main theme of the seminar has focused around three **application challenges** which are highly characteristic for real-world decision making and represent different ways that personalization is needed or delivered in an optimization setting. These were (i) Platform design and product lines, (ii) Responsive and online personalization, and (iii) Complex networks of decision makers. These three application challenges were crosslinked with three **research domains** that constitute the methodological core of multiobjective optimization and have been the foundation for the discussions at the previous Dagstuhl seminars. These were (1) Model building, (2) Preference modelling, and (3) Algorithm design and efficiency.

During the seminar, we formed five multi-disciplinary working groups (WGs) to implement the crosslinking between these application challenges and research domains, see Table 1. Each working group was focused on an application challenge (a row in Table 1; WGs 2, 3 and 4) or a research domain (a column in Table 1; WGs 1 and 5), all taking specific perspectives on the respective topics.

The program was updated on a daily basis to maintain flexibility in balancing time slots for talks, discussions, and working groups. The working groups were established on the first day in an open and highly interactive discussion. The program included several opportunities to report back from the working groups in order to establish further links and allow for adaptations and feedback. Some of the working groups split into subgroups and rejoined later in order to focus more strongly on different aspects of the topics considered. Abstracts of the talks and extended abstracts of the working groups can be found in subsequent chapters

of this report. Further notable events during the week included: (i) a hike on Wednesday afternoon with some sunshine (despite the quite terrible weather during the rest of the week), (ii) an announcements session allowing us to share details of upcoming events in our research community, and (iii) a wine and cheese party made possible by the support of the ITWM Kaiserslautern, represented by Karl-Heinz Küfer.

### Outcomes

Fourteen topical presentations were complemented by discussions in five working groups, covering the main themes of the seminar. The outcomes of each of the working groups can be seen in the sequel. Extended versions of their findings will be submitted to a Special Issue on “Personalization in Multiobjective Optimization: An Analytics Perspective” of the *Journal of Multicriteria Decision Analysis*, edited by Theo Stewart, that is guest edited by the organizers of this seminar. The submission deadline is July 31, 2018, and several working groups plan to submit extended versions of their reports to this special issue.

The seminar was highly productive, very lively and full of discussions, and has thus further strengthened the interaction between the EMO and MCDM communities. We expect that the seminar will initiate a new research domain interrelating multiobjective optimization and personalization, as it similarly has happened after the previous seminars in this series.

### Acknowledgments

A huge thank you to the Dagstuhl office and its very helpful and patient staff; many thanks to the organizers of the previous seminars in the series for the initiative and continuing advice; and many thanks to all the participants, who contributed in so many different ways to make this week a success. In the appendix, we also give special thanks to Joshua Knowles as he steps down from the organizer role.

## 2 Table of Contents

### Executive Summary

*Kathrin Klamroth, Joshua D. Knowles, Günter Rudolph, and Margaret M. Wiecek* 33

### Overview of Talks

Industrial applications of multicriteria decision support systems <i>Karl Heinz Küfer</i> . . . . .	38
Culturally tailored multicriteria product design using crowdsourcing <i>Georges Fadel</i> . . . . .	38
Metamodeling approaches for multiobjective optimization <i>Kalyanmoy Deb</i> . . . . .	39
Representations: Do they have potential for customer choice? <i>Serpil Sayın</i> . . . . .	39
Modelling complex networks of decision makers: An analytical sociology perspective <i>Robin Purshouse</i> . . . . .	40
Data-driven automatic design of multi-objective optimizers <i>Manuel López-Ibáñez</i> . . . . .	40
Maximizing the probability of consensus in group decision making <i>Michael Emmerich</i> . . . . .	41
Decision analytics with multiobjective optimization and a case in inventory management <i>Kaisa Miettinen</i> . . . . .	42
Actively learning a mapping for personalisation <i>Jürgen Branke</i> . . . . .	42
The NEMO framework for EMO: Learning value functions from pairwise comparisons <i>Roman Słowiński</i> . . . . .	43
Uncertainty quantification on Pareto fronts <i>Mickaël Binois</i> . . . . .	43
Innovization: Unveiling invariant rules from non-dominated solutions for knowledge discovery and faster convergence <i>Abhinav Gaur</i> . . . . .	43
Compressed data structures for bi-objective 0,1-knapsack problems <i>José Rui Figueira</i> . . . . .	44
Recent algorithmic progress in multiobjective (combinatorial) optimization <i>Andrzej Jaskiewicz</i> . . . . .	44

### Working Groups (WGs)

Multi-criteria decision making under performance and preference uncertainty (WG1) <i>Mickaël Binois, Jürgen Branke, Alexander Engau, Carlos M. Fonseca, Salvatore Greco, Miłosz Kadziński, Kathrin Klamroth, Sanaz Mostaghim, Patrick Reed, and Roman Słowiński</i> . . . . .	45
--	----



Personalization of multicriteria decision support systems (WG2) <i>Matthias Ehrgott, Gabriele Eichfelder, Karl-Heinz Küfer, Christoph Lofi, Kaisa Miettinen, Luís Paquete, Stefan Ruzika, Serpil Sayın, Ralph E. Steuer, Theodor J. Stewart, Michael Stiglmayr, and Daniel Vanderpooten</i> . . . . .	55
Usable knowledge extraction in multi-objective optimization: An analytics and “innovization” perspective (WG3) <i>Carlos A. Coello Coello, Kerstin Dächert, Kalyanmoy Deb, José Rui Figueira, Abhinav Gaur, Andrzej Jaszkiewicz, Günter Rudolph, Lothar Thiele, and Margaret M. Wiecek</i> . . . . .	70
Complex networks and MCDM (WG4) <i>Richard Allmendinger, Michael Emmerich, Georges Fadel, Jussi Hakanen, Johannes Jahn, Boris Naujoks, Robin Purshouse, and Pradyumn Shukla</i> . . . . .	76
Meta-modeling for (interactive) multi-objective optimization (WG5) <i>Dimo Brockhoff, Roberto Calandra, Manuel López-Ibáñez, Frank Neumann, Selvakumar Ulaganathan</i> . . . . .	85
<b>Topics of Interest for Participants for the Next Dagstuhl Seminar</b> . . . . .	95
<b>Changes in the Seminar Organization Body</b> . . . . .	95
<b>Seminar Schedule</b> . . . . .	95
<b>Participants</b> . . . . .	99

### 3 Overview of Talks

#### 3.1 Industrial applications of multicriteria decision support systems

*Karl Heinz Küfer (Fraunhofer ITWM – Kaiserslautern, DE)*

**License** © Creative Commons BY 3.0 Unported license  
© Karl Heinz Küfer

Most decisions in life are compromises: several objectives, most often arising from the four families cost, quality, time or environmental impact, have to be balanced. Decision making is rarely straight-forward because one cannot have best possible values for all of these goals simultaneously as they are at least partially in conflict. Many decision makers are reluctant to introduce decision support tools that directly show what the possible freedom of choice or inherent restrictions of the problems are. They often do not want to defend personal preferences or biases in decision rounds, which would become obvious by showing options and limitations in a transparent way. Others are in sorrows concerning the profile of or even their jobs. The talk will demonstrate and discuss examples of decision support tools in medical therapy planning, chemical process engineering and in the layout of renewable energy facilities, all of them in industrial practice for five or more years. Special attention is paid to the reception of such concepts in the companies and their impact if successfully implemented.

#### 3.2 Culturally tailored multicriteria product design using crowdsourcing

*Georges Fadel (Clemson University – Clemson, US)*

**License** © Creative Commons BY 3.0 Unported license  
© Georges Fadel

**Joint work of** Georges Fadel, Ivan Mata, Mo Chen, Paolo Guarneri, Manh Tien Nguyen

**Main reference** Ivan Mata, Georges Fadel, Anthony Garland, Winfried Zanker: “Affordance based interactive genetic algorithm (ABIGA)”, *Design Science*, Vol. 4, E5, 2018.

**URL** <https://doi.org/10.1017/dsj.2017.30>

The presentation describes an approach to involve crowds of users in the evolution of the design of a product by having them provide feedback to a tailored interactive multi-objective archive based micro- genetic algorithm. Affordances are defined as perceived opportunities for action, for instance, a ladder affords elevating the user and a glass affords containing a liquid. The users grade perceived affordances of a product and these are the criteria that the GA uses to evolve the shape of a product. The algorithm has multiple archives that store culturally biased solutions and use them in the evolution of solutions. After a number of generations, the designer can extract from the stored data which physical parameters affect specific affordances in the view of the users. The users will eventually be able to suggest additional affordances, and the designer would have to accept or not to add such a criterion to the system, and have possibly the designs evolve differently. A set of non-dominated solutions is then available to the designer to choose from. The system can be used by an individual to personalize a solution, or by a crowd to evolve the solution towards a more satisficing solution to the group.

### 3.3 Metamodeling approaches for multiobjective optimization

*Kalyanmoy Deb (Michigan State University – East Lansing, US)*

**License** © Creative Commons BY 3.0 Unported license  
© Kalyanmoy Deb

In multiobjective optimization, every objective function must be approximated with a suitable metamodel, particularly when a solution evaluation is computationally expensive. One straightforward approach is to model every objective function separately, but a number of other approaches are possible and may be more effective. In this talk, we proposed a taxonomy of different metamodeling frameworks and presented our recent results of each framework on multiple test problems. This research is motivated by practice and opens up a number of avenues for new research and application. Some of the methods highlighted are: (i) Specific metamodeling approaches (Kriging, RBF, or others) and their choice for every objective and constraint function, (ii) possible switching methods from one framework to another with iterations, (iii) possible other selection methods for metamodeling based on EMO methodologies, and (iv) possible use of trust region methods along with metamodeling approaches. Results on an industrial design problem was presented.

### 3.4 Representations: Do they have potential for customer choice?

*Serpil Sayın (Koc University – Istanbul, TR)*

**License** © Creative Commons BY 3.0 Unported license  
© Serpil Sayın

**Joint work of** Serpil Sayın, Gokhan Kirlik

**Main reference** Gokhan Kirlik, Serpil Sayın: “Bilevel programming for generating discrete representations in multiobjective optimization”, *Math. Program.*, Vol. 169(2), pp. 585–604, 2018.

**URL** <http://dx.doi.org/10.1007/s10107-017-1149-0>

Representations are subsets of nondominated sets that are expected to serve in the capacity of the original set. Finding representations makes most sense when the latter set is computationally difficult to obtain or practically difficult to explore. In recent years, there have been a number of studies that focused on delivering representations for multiobjective optimization problems. Some of these studies propose measures of quality to assess how well a representation or an approximation mimics the original set. These studies are mostly set in environments where finding the entire nondominated set is computationally challenging. Therefore they have not been discussed from the perspective of representing sets when all alternatives are explicitly available.

One problem in online retailing is presenting the items in a category to a potential customer. In most cases, the category contains a large selection of items. The user usually has a number of ways to customize the way she explores the category. For instance, filters may help limit values of interest for some relevant criteria. There may be choices offered to sort the items with respect to price, popularity, etc. I would like to ask the question if it is possible to design a new way of presenting a category to a customer based on what we know about representing nondominated sets. This would call for casting a customer's product choice problem as a multiple criteria one and delivering alternative mechanisms of navigating the category.

This discussion relates to the application challenge responsive and online personalization as well as representations in research domain.

### 3.5 Modelling complex networks of decision makers: An analytical sociology perspective

*Robin Purshouse (University of Sheffield – Sheffield, GB)*

**License** © Creative Commons BY 3.0 Unported license

© Robin Purshouse

**Joint work of** Robin C. Purshouse, Shaul Salomon, Gideon Avigad, Peter J. Fleming  
**Main reference** Shaul Salomon, Robin C. Purshouse, Gideon Avigad, Peter J. Fleming: “An Evolutionary Approach to Active Robust Multiobjective Optimisation”, in Proc. of the Evolutionary Multi-Criterion Optimization - 8th International Conference, EMO 2015, Guimarães, Portugal, March 29 -April 1, 2015. Proceedings, Part II, Lecture Notes in Computer Science, Vol. 9019, pp. 141–155, Springer, 2015.

**URL** [https://doi.org/10.1007/978-3-319-15892-1\\_10](https://doi.org/10.1007/978-3-319-15892-1_10)

Designers and planners who provide solutions for mass-markets and communities wish to understand how individuals in those markets and communities make choices about how they use, customise or reject those solutions. For example, a powertrain designer with fleet-level emissions and durability objectives wants to understand the different ways in which owners might operate a plug-in hybrid vehicle; a government planner with community-level health and revenue objectives wants to understand how citizens might choose to exploit a subsidised recreational facility. Whilst formulation of the higher-level multi-objective decision problem facing designers and planners has been addressed many times by researchers, far less attention has been paid to the, typically repeated, lower-level multi-objective decision problem faced by users, or to the interaction between these levels. Decisions at the lower-level are embedded within a complex socio-technical context, in which interactions between individuals can play a key role in how decisions are made and changed over time. This talk will introduce the framework of analytical sociology, pioneered by the Swedish sociologist Peter Hedström, as a means of modelling mass-customisation decision problems. Analytical sociology is a theory-based approach in which individual behaviours are driven by specified causal mechanisms. The talk will describe the three types of mechanism captured by the framework – situational, individual action, and transformational – and highlight the potential role of the designer and planner in shaping the decisions of heterogeneous individuals in mass-markets and communities.

### 3.6 Data-driven automatic design of multi-objective optimizers

*Manuel López-Ibáñez (University of Manchester – Manchester, GB)*

**License** © Creative Commons BY 3.0 Unported license

© Manuel López-Ibáñez

Recent work is increasingly showing that, given a library of good algorithmic components, automatically designed algorithms consistently outperform human-designed ones, even for thoroughly researched benchmark problems [1, 2, 3, 4]. The benefits of automated algorithm design rapidly increase for more complex and less studied problems, where the intuitions of human experts often fail. The transition from an expert-driven human-intensive design methodology to a data-driven CPU-intensive one also leads to the production of large amounts of data about the performance of algorithmic components. Despite some initial work in single-objective optimization and machine learning [5], it is still an open question how to use and analyze this data to gain insights about algorithmic components applied to multi-objective problems. Moreover, the transition to an automated design methodology

raises questions about performance metrics, the identification of equivalent and alternative algorithmic components, and the role of the decision-maker; questions that are particularly relevant in a multi-objective context.

## References

- 1 H. H. Hoos. Programming by optimization. *Communications of the ACM*, 55(2):70–80, 2012.
- 2 A. R. KhudaBukhsh, L. Xu, H. H. Hoos, and K. Leyton-Brown. SATenstein: Automatically Building Local Search SAT Solvers from Components. *Artificial Intelligence*, 232:20–42, 2016.
- 3 M. López-Ibáñez and T. Stützle. The Automatic Design of Multi-Objective Ant Colony Optimization Algorithms. *IEEE Transactions on Evolutionary Computation*, 16(6):861–875, 2012.
- 4 L. C. T. Bezerra, M. López-Ibáñez, and T. Stützle. Automatic Component-Wise Design of Multi-Objective Evolutionary Algorithms. *IEEE Transactions on Evolutionary Computation*, 20(3):403–417, 2016.
- 5 F. Hutter, H. H. Hoos, and K. Leyton-Brown. An Efficient Approach for Assessing Hyperparameter Importance. In *Proceedings of the 31th International Conference on Machine Learning*, volume 32, pages 754–762, 2014.

## 3.7 Maximizing the probability of consensus in group decision making

*Michael Emmerich (Leiden University – Leiden, NL)*


License © Creative Commons BY 3.0 Unported license  
© Michael Emmerich

Joint work of Michael Emmerich, Andre Deutz, Iryna Yevseyeva

Consider the scenario of selecting a portfolio of  $k$  alternative solutions from a set of  $n \gg k$  solutions. A moderator presents  $k$  solutions to a board of decision makers. The goal is to maximize the probability that the decision makers achieve consensus about at least one solution in the portfolio. In advance, decision makers formulated desirability functions for the objectives of concern – ranging from 0 (not acceptable) to 1 (fully satisfactory). Moreover, correlations between objectives may be formulated using a dependence graph. The analysis shows that the computation of the probability of consensus is related to specific integrals over the dominated space, which reduces to the hypervolume indicator after coordinate transformation in case of independent objectives. The problem of veto by overdemanding decision makers is discussed we propose a possible remedy by replacing the probability by higher momenta of the joint acceptance probability distribution.

### 3.8 Decision analytics with multiobjective optimization and a case in inventory management

*Kaisa Miettinen (University of Jyväskylä – Jyväskylä, FI)*

**License**  Creative Commons BY 3.0 Unported license

© Kaisa Miettinen

**Joint work of** Kaisa Miettinen, Juha Sipilä, Risto Heikkinen, Vesa Ojalehto

**URL** <http://www.jyu.fi/demo>

Thanks to digitalization, we have access to various types of data and must decide how to make the most of the data. We can use descriptive or predictive analytics but to make recommendations and informed decisions based on the data, we need prescriptive or decision analytics. If the problems contain multiple conflicting objectives, multiobjective optimization are to be applied.


We introduce the new thematic research area at the University of Jyväskylä called Decision Analytics utilizing Causal Models and Multiobjective Optimization (DEMO). The objective of DEMO is to develop elements of a seamless chain from data to decision support.

Lot sizing is an example of a data-driven optimization problem. It is important in production planning and inventory management, where a decision maker needs support, in particular, when the demand is stochastic. We consider the lot sizing problem of a Finnish production company and formulate four conflicting objectives. We solve it with two interactive multiobjective optimization methods. In interactive methods, a decision maker directs the search for the best balance between the conflicting objectives by providing preference information. In this way, (s)he can learn about what kind of solutions are available for the problem and also learn about the feasibility of one's preferences.

In the case considered, the decision maker found it useful to switch the method during the solution process. The results of this data-driven interactive multiobjective optimization approach are encouraging and demonstrate the practical value of decision analytics.

### 3.9 Actively learning a mapping for personalisation

*Jürgen Branke (University of Warwick – Warwick, GB)*

**License**  Creative Commons BY 3.0 Unported license

© Jürgen Branke

This talk tackles the problem of efficiently collecting data to learn a classifier, or mapping, from each user to the best personalisation, where users are described by continuous features and there is a finite set of personalisation options to choose from. An example would be online advertisements, where we want to learn the best possible advertisement and advertisement format for each user. We propose a fully sequential information collection policy based on Bayesian statistics and Gaussian Process models. In each step, they myopically allocate to the user the advertisement that promises the highest value of information collected.

### 3.10 The NEMO framework for EMO: Learning value functions from pairwise comparisons

*Roman Słowiński (Poznan University of Technology – Poznan, PL)*

**License** © Creative Commons BY 3.0 Unported license

© Roman Słowiński

**Joint work of** Roman Słowiński, Jürgen Branke, Salvatore Corrente, Salvatore Greco

Some years ago, we have proposed the NEMO framework to enhance multi-objective evolutionary algorithms by pairwise preference elicitation during the optimisation, allowing the algorithm to converge more quickly to the most relevant region of the Pareto front. The framework is based on Robust Ordinal Regression. Over the years, several variations have been developed, with different user preference models (linear, additive, Choquet-integral value functions) and different ways of integrating this information into evolutionary algorithms (as a surrogate fitness function, or by enriching the dominance relation). This presentation will provide an overview of the developments in this area.

### 3.11 Uncertainty quantification on Pareto fronts

*Mickaël Binois (University of Chicago – Chicago, US)*

**License** © Creative Commons BY 3.0 Unported license

© Mickaël Binois

In this short presentation, we review methods to approximate Pareto fronts in the case of expensive, possibly noisy, blackbox objective functions. We concentrate on methods involving Gaussian processes, which provide uncertainty quantification on the estimated Pareto front. Variations on the modeling include re-interpolation and nugget estimation, while the uncertainty is estimated from sampling, random closed sets or bootstrap.

### 3.12 Innovization: Unveiling invariant rules from non-dominated solutions for knowledge discovery and faster convergence

*Abhinav Gaur (Michigan State University – East Lansing, US)*

**License** © Creative Commons BY 3.0 Unported license

© Abhinav Gaur

Multi-objective optimization (MOO) problems lend themselves to not one but a set of optimal solutions also called Pareto-optimal (PO) solutions. Such PO solutions carry information on patterns that make these solutions concurrently optimal for multiple objectives/Customer preferences. Discovering such patterns from the PO solutions is called ‘Innovization’ or innovation through optimization. Some of the uses of carrying out an Innovization exercise are discovering principles that makes certain solutions PO for a MOO problem, automatically discovering optimization heuristics for a problem and, expediting black box MOO algorithms. In the context of “Personalized MOO”, the concepts in Innovization, Higher Level Innovization, Lower Level Innovization, Temporal Innovization have direct applications. For example, temporal Innovization can help us discover principles that govern how preferences of a class of customer have changed over time. Lower level innovization can help us discover preferences

of customers whose preferences lie at part of the PO front. Higher level Innovization can help us discover principles that govern the customer preferences as certain problem parameters are changed, and so on. Hence, the Innovization idea seems to be very relevant to the problem of studying “Personalized Multi Objective Optimization”.

### 3.13 Compressed data structures for bi-objective 0,1-knapsack problems

*José Rui Figueira (IST – Lisbon, PT)*

**License**  Creative Commons BY 3.0 Unported license  
© José Rui Figueira

Solving multi-objective combinatorial optimization problems to optimality is a computationally expensive task. The development of implicit enumeration approaches that efficiently explore certain properties of these problems has been the main focus of recent research. This article proposes algorithmic techniques that extend and empirically improve the memory usage of a dynamic programming algorithm for computing the set of efficient solutions both in the objective space and in the decision space for the bi-objective knapsack problem. An in-depth experimental analysis provides further information about the performance of these techniques with respect to the tradeoff between CPU time and memory usage.

### 3.14 Recent algorithmic progress in multiobjective (combinatorial) optimization

*Andrzej Jaszkiewicz (Poznan University of Technology – Poznan, PL)*

**License**  Creative Commons BY 3.0 Unported license  
© Andrzej Jaszkiewicz

Despite of many years of research in the area of multiobjective evolutionary algorithms and more generally multiobjective metaheuristics many real-life multiobjective problems, in particular combinatorial problems, constitute a serious challenge for existing methods. Recently an important progress has been made in the algorithmic toolbox of multiobjective optimization. Some of the new algorithms are focused on the combinatorial optimization, but many are more generally applicable. Some of the recently proposed or improved algorithms are:

- ND-Tree data structure and algorithm for the dynamic non-dominance problem [1]. ND-Tree allows for very efficient update of even large Pareto archives. It allows multiobjective evolutionary algorithms and other metaheuristics to store large sets of potentially Pareto-optimal solutions without loss of efficiency. ND-Tree can also be applied to efficiently solve the non-dominated sorting problem often used in evolutionary algorithms.
- Many-objective Pareto Local Search (MPLS) [2]. Pareto Local Search proved to be a very effective tool in the case of the bi-objective combinatorial optimization and it was used in a number of the state-of-the-art algorithms for problems of this kind. On the other hand, the standard Pareto Local Search algorithm becomes very inefficient for problems with more than two objectives. Many-Objective Pareto Local Search algorithm uses three new mechanisms to preserve the effectiveness of PLS in many-objective case.



The new mechanisms are: the efficient update of large Pareto archives with ND-Tree data structure, a new mechanism for the selection of the promising solutions for the neighborhood exploration, and a partial exploration of the neighborhoods.

- New efficient algorithms for calculating the exact hypervolume of the space dominated by a set of  $d$ -dimensional points. This value is often used as the quality indicator in the multiobjective evolutionary algorithms and other metaheuristics and the efficiency of calculating this indicator is of crucial importance especially in the case of large sets or many dimensional objective spaces. Recently significant improvements have been obtained in algorithms for calculating this indicator [3, 4, 5, 6, 7, 8]. They allow not only to speed-up computational experiments but also to use hypervolume within multiobjective algorithms, e.g. to guide the search or to define stopping conditions.

## References

- 1 A. Jaskiewicz and T. Lust. ND-Tree-based update: a Fast Algorithm for the Dynamic Non-Dominance Problem. *ArXiv e-prints*, arXiv:1603.04798, 2016.
- 2 A. Jaskiewicz. Many-Objective Pareto Local Search. *ArXiv e-prints*, arXiv: 1707.07899, 2017.
- 3 A. Jaskiewicz. Improved quick hypervolume algorithm. *Computers & Operations Research*, 90:72–83, 2018.
- 4 L. M. S. Russo and A. P. Francisco. Quick hypervolume. *IEEE Transactions on Evolutionary Computation*, 18(4):481–502, 2014.
- 5 L. M. S. Russo and A. P. Francisco. Extending quick hypervolume. *Journal of Heuristics*, 22(3):245–271, 2016.
- 6 R. Lacour, K. Klamroth, and C. M. Fonseca. A box decomposition algorithm to compute the hypervolume indicator. *Computers & Operations Research*, 79:347–360, 2017.
- 7 L. While, L. Bradstreet, and L. Barone. A fast way of calculating exact hypervolumes. *IEEE Transactions on Evolutionary Computation*, 16(1):86–95, 2012.
- 8 W. Cox and L. While. Improving the IWFG Algorithm For Calculating Incremental Hypervolume. *IEEE Congress on Evolutionary Computation*, pages 39690–3976, 2016.

## 4 Working Groups (WGs)

### 4.1 Multi-criteria decision making under performance and preference uncertainty (WG1)

*Mickaël Binois, Jürgen Branke, Alexander Engau, Carlos M. Fonseca, Salvatore Greco, Miłosz Kadziński, Kathrin Klamroth, Sanaz Mostaghim, Patrick Reed, and Roman Słowiński*

License © Creative Commons BY 3.0 Unported license

© Mickaël Binois, Jürgen Branke, Alexander Engau, Carlos M. Fonseca, Salvatore Greco, Miłosz Kadziński, Kathrin Klamroth, Sanaz Mostaghim, Patrick Reed, and Roman Słowiński

**Abstract.** We propose a novel methodology for interactive multi-objective optimization taking into account imprecision, ill-determination and uncertainty referring to both, the technical aspects determining evaluations of solutions by objective functions and the subjective aspects related to the preferences of the decision maker. With this aim, we consider a probability distribution on the space of the objective functions and a probability distribution on the space of the utility functions representing preferences of the decision maker. On the basis of these two probability distributions, without loss of generality supposed to be independent, one can compute a multi-criteria expected utility with a corresponding standard

deviation, that permit to assess a quality of each proposed solution. One can also compute an average multi-criteria expected utility and a related standard deviation for a set of solutions, which permit to assess a quality of a population of solutions. This feature can be useful in evolutionary multi-objective optimization algorithms to compare populations of solutions in successive iterations.

#### 4.1.1 Introduction

This paper summarizes the work of the *Preference Uncertainty Quantification* working group at the Dagstuhl seminar 18031 “Personalized Multi-objective Programming: An Analytics Perspective” that took place in Schloss Dagstuhl – Leibniz Center for Informatics - on January 14–19, 2018.

#### 4.1.2 Uncertainties

When dealing with multi-objective optimization problems, the decision makers (DMs), and the analysts helping them to solve these problems, are confronted in their reasoning with some uncertainties that are inherent to two kinds of “imperfect” information (see [2] and [3]):

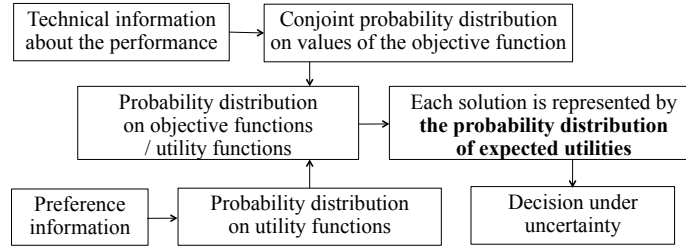
1. Information about the preferences of DMs is always partial and ill-defined. Even more, complete preferences do not exist a priori in DMs’ mind, because they evolve in the decision aiding process in interaction with an analyst. The preferences are formed in a constructive learning process in which DMs get a conviction that the most preferred solution has been reached for a given problem statement.
2. Information about consequences of considered solutions usually depend on hardly measurable or random variables. This makes that, in general, the evaluation of solutions with respect to different criteria is imprecise or uncertain.

Therefore, there is a need to take into account these two sources of uncertainty in an interactive multi-objective optimization process. A first consideration of this problem, but taking into account only uncertainty related to utility functions, has been proposed in [4].

#### 4.1.3 Problem formulation and basic notation

The multi-objective optimization process presented in this paper is formally represented as a multi-objective programming problem under performance and preference uncertainty as follows. Let  $X \subset R^n$  be an  $n$ -dimensional set of feasible decisions (or solutions, designs, alternatives, etc.) Let  $f: R^n \rightarrow R^m$  be an  $m$ -dimensional vector, called objective function, that maps each decision  $x \in X$  to a corresponding consequence or performance vector  $y = f(x)$ . To model performance uncertainty, we assume that each objective function  $f = (f_1, f_2, \dots, f_m)$  is a random element of some (for now: a priori) given set  $\mathcal{F}$  of cardinality  $k$ , i.e.,  $\mathcal{F} = \{f^1, f^2, \dots, f^k\}$  with random outputs  $y^i = (y_1^i, y_2^i, \dots, y_m^i)$  for each  $i \in \{1, 2, \dots, k\}$ . In other words, for each  $i \in \{1, 2, \dots, k\}$ , the vector function  $f^i = (f_1^i, f_2^i, \dots, f_m^i)$  is one realization of the random objective function  $f$ .

Moreover, under the additional assumption that this uncertainty is stochastic in nature, we can assign or estimate a stochastic probability vector  $p = (p_1, p_2, \dots, p_k)$  with  $\sum_{i=1}^k p_i = 1$  and with the interpretation that  $\Pr[f = f^i] = \Pr[y = y^i] = p_i$  for each  $i \in \{1, 2, \dots, k\}$ . In this way, we have defined a discrete probability distribution on the space of values taken by the objective function. Obviously, one can consider a generic probability distribution, not necessarily a discrete one. For a scheme of this setting, see the conceptual relationship between *technical information about the performance* and *conjoint probability distribution on values of the objective function* in Figure 1 on the top.



■ **Figure 1** Main idea underlying the proposed methodology.

Similarly, we can describe the uncertainty about preferences of the DM, considering a utility function  $u : R^m \rightarrow R$ , such that  $y \mapsto z = u(y)$ . Again,  $u$  is considered to be an element of a set  $\mathcal{U} = \{u^1, \dots, u^\ell\}$ , interpreted as a set of possible realizations of an uncertain utility function. Each utility function  $u^j \in \mathcal{U}$  has a probability  $\Pr[u = u^j] = q_j$ ,  $j = 1, \dots, \ell$ . This is marked in Figure 1 as *preference information* and *probability distribution of utility function*.

### A simple example

Consider a simple example, with  $n = 2$  and  $X = [0, 1]^2$ , so that the decision input to the objective functions is a vector  $x = (x_1, x_2)$  composed of two decision variables.

**Performance uncertainty.** Let us measure the performance of  $x$  in two dimensions, i.e.,  $m = 2$ , so that  $f : R^2 \rightarrow R^2$  with  $f = (f_1, f_2)$  for each objective realization. Moreover, consider  $k = 3$  uncertain realizations of the objective function, denoted by  $\mathcal{F} = \{f^1, f^2, f^3\}$ , with probabilities  $p = (p_1, p_2, p_3) = (0.5, 0.2, 0.3)$ , and taking the following form:

$$\begin{aligned} f^1(x) &:= (f_1^1(x_1, x_2), f_2^1(x_1, x_2)) = (x_1, x_2) \\ f^2(x) &:= (f_1^2(x_1, x_2), f_2^2(x_1, x_2)) = (\sqrt{x_1}, \sqrt[3]{x_2}) \\ f^3(x) &:= (f_1^3(x_1, x_2), f_2^3(x_1, x_2)) = (x_1^2, x_2^3). \end{aligned}$$

**Note:** Alternatively, supposing that the values taken by the objective function in each realization depend on the value taken on a basic reference realization (for example the mean value in case of an estimation through a Bayesian process) one can define the performance set  $Y := \{f(x) : x \in X\} \subset R^m$  and then use a transformation  $\phi^h : R^m \rightarrow R^m$  for each possible realization  $h = 1, \dots, k$ , so that for each  $y = f(x) \in Y$  we can also write  $\phi^h(y) = \phi^h(y_1, y_2)$  or  $\phi^h(f_1(x), f_2(x)) = (f_1^h(x), f_2^h(x))$ . For instance, in the considered example, we can take as a basic reference realization  $f^1(x) = f^1(x_1, x_2) = (f_1^1(x_1, x_2), f_2^1(x_1, x_2)) = (y_1, y_2) = (x_1, x_2)$ , and for each realization  $h = 1, 2, 3$ , suppose:

- $\phi^1(y) = \phi^1(y_1, y_2) = (y_1, y_2)$   
or  $\phi^1(f(x)) = \phi^1(f_1(x), f_2(x)) = (f_1^1(x), f_2^1(x)) = (f_1(x), f_2(x))$ ,
- $\phi^2(y) = \phi^2(y_1, y_2) = (\sqrt{y_1}, \sqrt[3]{y_2})$   
or  $\phi^2(f(x)) = \phi^2(f_1(x), f_2(x)) = (f_1^2(x), f_2^2(x)) = (\sqrt{f_1(x)}, \sqrt[3]{f_2(x)})$ ,
- $\phi^3(y) = \phi^3(y_1, y_2) = ((y_1)^2, (y_2)^3)$   
or  $\phi^3(f(x)) = \phi^3(f_1(x), f_2(x)) = (f_1^3(x), f_2^3(x)) = ((f_1(x))^2, (f_2(x))^3)$ .

**Preference uncertainty.** Suppose that we have a probability distribution on a set of  $\ell = 4$  utility functions describing the preference information as follows:

$$\begin{aligned} q_1 = 0.4 & : u^1(y) = 0.3y_1 + 0.7y_2, \\ q_2 = 0.3 & : u^2(y) = 0.5y_1 + 0.5y_2, \\ q_3 = 0.2 & : u^3(y) = 0.8y_1 + 0.2y_2, \\ q_4 = 0.1 & : u^4(y) = 0.9y_1 + 0.1y_2, \end{aligned}$$

where  $q_1, q_2, q_3, q_4$  are probabilities of realization of these utility functions.

**Expected utility and variance of a single solution.** In the following, we assume that the probability distributions of performance information and utility functions are independent from each other. Therefore, the joint probability distribution on the product space  $\mathcal{F} \times \mathcal{U}$  assigns to each pair  $(f^i, u^j)$  the probability  $\pi_{ij} = p_i \cdot q_j$  shown in the following matrix:

$$\Pi = \begin{bmatrix} \pi_{11} & \pi_{21} & \pi_{31} \\ \pi_{12} & \pi_{22} & \pi_{32} \\ \pi_{13} & \pi_{23} & \pi_{33} \\ \pi_{14} & \pi_{24} & \pi_{34} \end{bmatrix}^T = \begin{bmatrix} 0.20 & 0.08 & 0.12 \\ 0.15 & 0.06 & 0.09 \\ 0.10 & 0.04 & 0.06 \\ 0.05 & 0.02 & 0.03 \end{bmatrix}^T$$

For each decision  $x$  and each realization of its performance  $f^i$  in  $\mathcal{F}$ , one can compute the utility value  $u^j(f^i(x))$  that can be presented in the form of a matrix  $\mathbf{U}(x)$  with elements  $u^j(f^i(x))$  for  $i$  and  $j$ .

$$\mathbf{U}(x) = \begin{bmatrix} u^1(f^1(x)) & u^2(f^1(x)) & u^3(f^1(x)) & u^4(f^1(x)) \\ u^1(f^2(x)) & u^2(f^2(x)) & u^3(f^2(x)) & u^4(f^2(x)) \\ u^1(f^3(x)) & u^2(f^3(x)) & u^3(f^3(x)) & u^4(f^3(x)) \end{bmatrix}$$

Assuming that  $x = (0.5, 0.7)$ , one can compute the entries of matrix  $\mathbf{U}(x)$ , getting:

$$\mathbf{U}(0.5, 0.7) = \begin{bmatrix} 0.6400 & 0.6000 & 0.5400 & 0.5200 \\ 0.8337 & 0.7975 & 0.7433 & 0.7252 \\ 0.3151 & 0.2965 & 0.2686 & 0.2593 \end{bmatrix}$$

In order to compute the expected utility value  $E(u(f(x)))$  of decision  $x$ , we first need to compute the matrix:

$$\mathbf{V}(x) = \mathbf{U}(x) \times \Pi = [u^j(f^i(x)) \cdot \pi_{i,j}]_{\substack{i=1,\dots,k \\ j=1,\dots,\ell}}.$$

In our example, we get:

$$\mathbf{V}(0.5, 0.7) = \begin{bmatrix} 0.1280 & 0.0900 & 0.0540 & 0.0260 \\ 0.0667 & 0.0479 & 0.0297 & 0.0145 \\ 0.0378 & 0.0267 & 0.0161 & 0.0078 \end{bmatrix}$$

Then, the expected utility value  $E(u(f(x)))$  is obtained as:

$$E(u(f(x))) = \sum_{i=1}^k \sum_{j=1}^{\ell} u^j(f^i(x)) \cdot \pi_{ij}. \quad (1)$$

In our example, for  $x = (0.5, 0.7)$ , the expected utility value is  $E(u(f(0.5, 0.7))) = 0.5452$ . The variance is given by:

$$\sigma^2(u(f(x))) = \sum_{i=1}^k \sum_{j=1}^{\ell} (u^j(f^i(x)) - E(u(f(x))))^2 \cdot \pi_{ij}, \quad (2)$$

which, in our example, gives  $\sigma^2(u(f(x))) = 0.0339$ .

In general, the DM will try to maximize the expected value  $E(u(f(x)))$  and to minimize the variance of the selected solution  $\sigma^2(u(f(x)))$ . This principle can be applied in different procedures to select a solution  $x$  from a set of feasible solutions  $X \in R^n$ , such as:

- select a solution  $x \in X$  with the maximum expected utility value  $E(u(f(x)))$  provided that its variance  $\sigma^2(u(f(x)))$  is not greater than a given threshold  $\sigma^{2*}$ ;
- select a solution  $x \in X$  with the minimum variance  $\sigma^2(u(f(x)))$  provided that its expected utility value is not smaller than a given threshold  $E^*$ ;
- select a solution  $x \in X$  maximizing a scoring function  $S(E(u(f(x))), \sigma^2(u(f(x))))$  being not decreasing with respect to the expected utility value  $E(u(f(x)))$  and not increasing with respect to the variance  $\sigma^2(u(f(x)))$ , as it is the case of

$$S(E(u(f(x))), \sigma^2(u(f(x)))) = E(u(f(x))) - \lambda \cdot \sigma^2(u(f(x)))$$

where  $\lambda \geq 0$  is a coefficient representing a DM's aversion to risk.

Let us apply the above procedures to a set of feasible solutions  $X = \{x^1, x^2, x^3, x^4\}$ , where

- $x^1 = (0.5, 0.7)$ ,
- $x^2 = (0.8, 0.4)$ ,
- $x^3 = (0.4, 0.8)$ ,
- $x^4 = (0.9, 0.2)$ .

Let us observe that solution  $x^1$  is the same as solution  $x$  considered in the above simple example. Computing the expected utility value and the variance for each solution from  $X$  we get

- $E(u(f(x^1))) = 0.5452$ ,  $\sigma^2(u(f(x^1))) = 0.0339$ ,
- $E(u(f(x^2))) = 0.5768$ ,  $\sigma^2(u(f(x^2))) = 0.0350$ ,
- $E(u(f(x^3))) = 0.5496$ ,  $\sigma^2(u(f(x^3))) = 0.0323$ ,
- $E(u(f(x^4))) = 0.5643$ ,  $\sigma^2(u(f(x^4))) = 0.0377$ .

Consequently:

- if the DM wants to select a solution  $x \in X$  with the maximum expected utility value  $E(u(f(x)))$  provided that its variance  $\sigma^2(u(f(x)))$  is not greater than the threshold  $(\sigma^*)^2 = 0.0340$ , then solution  $x^3$  is selected;
- if the DM wants to select a solution  $x \in X$  with the minimum variance  $\sigma^2(u(f(x)))$  provided that its expected utility value is not smaller than the threshold  $E^* = 0.55$ , then solution  $x^2$  is selected;
- if the DM wants to select a solution  $x \in X$  maximizing a scoring function

$$S(E(u(f(x))), \sigma^2(u(f(x)))) = E(u(f(x))) - 2 \cdot \sigma^2(u(f(x))),$$

then we get

- $S(E(u(f(x^1))), \sigma^2(u(f(x^1)))) = 0.4773$ ,
  - $S(E(u(f(x^2))), \sigma^2(u(f(x^2)))) = 0.5068$ ,
  - $S(E(u(f(x^3))), \sigma^2(u(f(x^3)))) = 0.4850$ ,
  - $S(E(u(f(x^4))), \sigma^2(u(f(x^4)))) = 0.4890$ ,
- so that solution  $x^2$  is selected.

Another problem that can be considered in this context is the following. Suppose the DM wants to select one solution from  $X \subseteq R^n$ , which would maximize the expected utility value  $E(u(f(x)))$  and minimize the variance  $\sigma^2(u(f(x)))$ , taking into account a number of

■ **Table 2** A representation of Pareto-optimal solutions.

$x_1$	$x_2$	Expected value	Variance
0.388	0.862	0.580	0.030
0.351	0.899	0.585	0.031
0.314	0.936	0.592	0.031
0.302	0.948	0.594	0.032
0.292	0.958	0.597	0.032
0.284	0.966	0.598	0.033
0.276	0.974	0.600	0.033
0.270	0.980	0.602	0.034
0.263	0.987	0.603	0.035
0.258	0.992	0.605	0.035
0.252	0.998	0.606	0.036
0.250	1.000	0.607	0.036

constraints concerning decision variables  $h_s(x) \leq 0$ ,  $s = 1, \dots, S$ . Formally, this problem can be formulated as follows:

$$\text{maximize: } E(u(f(x)))$$

$$\text{minimize: } \sigma^2(u(f(x)))$$

subject to the constraints

$$x \in X, \tag{3}$$

$$h_s(x) \leq 0, \quad s = 1, \dots, S. \tag{4}$$

Obviously, in general, it is not possible to get an optimum value of  $E(u(f(x)))$  and  $\sigma^2(u(f(x)))$  for the same feasible  $x$ . Instead, one gets a set of Pareto-optimal solutions  $x$ , i.e., all solutions  $x \in X$  satisfying  $h_s(x) \leq 0$ ,  $s = 1, \dots, S$ , for which there does not exist any other solution  $\bar{x} \in X$  satisfying  $h_s(\bar{x}) \leq 0$ ,  $s = 1, \dots, S$ , having not worse expected utility value  $E(u(f(\bar{x})))$  and not worse variance  $\sigma^2(u(f(\bar{x})))$ , with at least one of the two being better, that is

$$E(u(f(\bar{x}))) > E(u(f(x))), \tag{5}$$

$$\sigma^2(u(f(\bar{x}))) \leq \sigma^2(u(f(x))) \tag{6}$$

or

$$E(u(f(\bar{x}))) \geq E(u(f(x))), \tag{7}$$

$$\sigma^2(u(f(\bar{x}))) < \sigma^2(u(f(x))). \tag{8}$$

Coming back to our example, we have  $X = [0, 1]^2$ , and let us consider the constraint  $h(x) = x_1 + x_2 - 1.25 \leq 0$ . Taking into account the set of objective functions  $\mathcal{F}$  and the set of utility function  $\mathcal{U}$  with respective probability distributions  $p$  and  $q$ , generating the conjoint probability distribution  $\Pi$  on  $\mathcal{F} \times \mathcal{U}$  introduced above, we can get a set of representative Pareto-optimal solutions presented in Table 2.

**Expected utility value and variance of a set of solutions.** Suppose we have a set of solutions  $X = \{x^1, \dots, x^r, \dots, x^t\} \subseteq R^n$ . In this case, it is possible to compute the expected utility value and the variance of this population of solutions, as follows:

$$E(u(f(X))) = \sum_{r=1}^t \sum_{i=1}^k \sum_{j=1}^{\ell} u_j(f^i(x^r)) \cdot \pi_{ij} \quad (9)$$

$$\sigma^2(u(f(X))) = \sum_{r=1}^t \sum_{i=1}^k \sum_{j=1}^{\ell} (u_j(f^i(x^r)) - E(u(f(X))))^2 \cdot \pi_{ij} \quad (10)$$

The expected utility value  $E(u(f(X)))$  and the variance  $\sigma^2(u(f(X)))$  can be computed using expected utility values and variances of particular solutions in the population, as well as covariances between these solutions:

$$E(u(f(X))) = \sum_{r=1}^t E(u(f(x^r))) \quad (11)$$

$$\sigma^2(u(f(X))) = \sum_{r=1}^t \sigma^2(u(f(x^r))) + 2 \sum_{r < s} \sigma(u(f(x^r)), u(f(x^s))) \quad (12)$$

where  $\sigma(u(f(x^r)), u(f(x^s)))$ ,  $r, s = 1, \dots, t$ ,  $r < s$ , is the covariance between  $u(f(x^r))$  and  $u(f(x^s))$ , that can be computed as follows:

$$\begin{aligned} \sigma(u(f(x^r)), u(f(x^s))) &= \sum_{i=1}^k \sum_{j=1}^{\ell} (u_j(f^i(x^r)) - E(u(f(x^r)))) \\ &\quad \cdot (u_j(f^i(x^s)) - E(u(f(x^s)))). \end{aligned} \quad (13)$$

The concepts of the expected utility value and the variance of a set of solution can be applied in multi-objective optimization algorithms with a different aim, for example:

- find a subset of solutions  $Y \subset X$  of a given cardinality  $q$  having the maximum expected utility value  $E(u(f(Y)))$ , provided that its variance  $\sigma^2(u(f(Y)))$  is not greater than a given threshold  $\bar{\sigma}^2$ ; the subset  $Y$  can be found by solving the following 0 – 1 quadratic programming problem:

$$\text{maximize: } \sum_{r=1}^t y_r E(u(f(x^r)))$$

subject to the constraints

$$\sum_{r=1}^t y_r \sigma^2(u(f(x^r))) + 2 \sum_{r=1}^{t-1} \sum_{s=r+1}^t y_r y_s \sigma(u(f(x^r)), u(f(x^s))) \leq \bar{\sigma}^2, \quad (14)$$

$$\sum_{r=1}^t y_r = q, \quad (15)$$

$$y_r \in \{0, 1\}, \quad r = 1, \dots, t; \quad (16)$$

the optimal subset  $Y$  will be composed of  $q$  solutions  $x^r \in X$  with  $y_r = 1$ ;

- find a subset of solutions  $Y \subset X$  of a given cardinality  $q$  having the minimum variance  $\sigma^2(u(f(Y)))$ , provided that the its expected value  $E(u(f(Y)))$  is not smaller than a given threshold  $\bar{E}$ ; the subset  $Y$  can be found by solving the following 0 – 1 quadratic programming problem:

$$\text{minimize: } \sum_{r=1}^t y_r \sigma^2(u(f(x^r))) + 2 \sum_{r=1}^{t-1} \sum_{s=r+1}^t y_r y_s \sigma(u(f(x^r)), u(f(x^s)))$$

subject to the constraints

$$\sum_{r=1}^t y_r E(u(f(x^r))) \geq \bar{E}, \quad (17)$$

$$\sum_{r=1}^t y_r = q, \quad (18)$$

$$y_r \in \{0, 1\}, r = 1, \dots, t; \quad (19)$$

again, the optimal subset  $Y$  will be composed of  $q$  solutions  $x^r \in X$  with  $y_r = 1$ .

Coming back to our example, let us consider again the solutions from the set  $X = \{x^1, x^2, x^3, x^4\}$ , and let us compute the covariances  $\sigma(u(f(x^r)), u(f(x^s)))$ , obtaining the following variance-covariance matrix  $\Sigma(X) = [\sigma(u(f(x^r)), u(f(x^s)))]$ , where  $\sigma(u(f(x^r)), u(f(x^r))) = \sigma^2(u(f(x^r)))$ :

$$\Sigma(X) = \begin{bmatrix} 0.0339 & 0.0258 & 0.0318 & 0.0157 \\ 0.0258 & 0.0350 & 0.0182 & 0.0334 \\ 0.0318 & 0.0182 & 0.0323 & 0.0064 \\ 0.0157 & 0.0334 & 0.0064 & 0.0377 \end{bmatrix}$$

Let us suppose that the DM wants to select a subset of solutions  $Y \subset X$  with cardinality  $q = 3$ , having the maximum expected utility value  $E(u(f(Y)))$ . Solving the 0-1 quadratic programming problem presented above, and without considering any constraint on the variance  $\sigma^2(u(f(Y)))$ , we get that the DM has to select the subset  $Y_1 = \{x^2, x^3, x^4\}$  with expected utility value  $E(u(f(Y_1))) = 1.6907$  and variance  $\sigma^2(u(f(Y_1))) = 0.2211$ .

If, in turn, the DM would like to select a subset of solutions  $Y \subset X$  with cardinality  $q = 3$ , having the minimum variance  $\sigma^2(u(f(Y)))$ , then, by solving the corresponding 0-1 quadratic programming problem presented above, and without considering any constraint on the expected value  $E(u(f(Y)))$ , the DM would get the subset  $Y_2 = \{x^1, x^3, x^4\}$  with expected utility value  $E(u(f(Y_2))) = 1.6591$  and variance  $\sigma^2(u(f(Y_2))) = 0.2118$ .

Suppose now that the DM would like to select a subset of solutions  $Y \subset X$  with cardinality  $q = 2$ , having the maximum expected utility value  $E(u(f(Y)))$  but under the condition that the variance  $\sigma^2(u(f(Y)))$  is not greater than 0.215. In this case, solving the corresponding 0-1 quadratic programming problem, the DM would get the subset  $Y_3 = \{x^3, x^4\}$  with expected utility value  $E(u(f(Y_3))) = 1.1139$  and variance  $\sigma^2(u(f(Y_3))) = 0.0828$ .

Finally, suppose that the DM would like to select a subset of solutions  $Y \subset X$  with cardinality  $q = 2$ , having the minimum variance  $\sigma^2(u(f(Y)))$  but under the condition that the expected utility value  $E(u(f(Y)))$  is not smaller than 1.1. In this case, the DM would get again the subset  $Y_4 = \{x^3, x^4\}$ .

The above two problems of selecting a subset of solutions of a given cardinality maximizing the expected utility value with a constraint on the variance, or minimizing the variance with a constraint on the expected value, can be interpreted as a discrete version of the Markowitz portfolio selection problem in the context of multi-objective optimization. It is sensible to consider also the classic continuous Markowitz portfolio selection problem which consists in searching for a vector

$$\mathbf{y} = [y_1, \dots, y_t], \quad y_r \geq 0, \quad r = 1, \dots, t, \quad \sum_{r=1}^t y_r = 1,$$

that maximizes the expected utility value

$$E(u(f(\mathbf{y}))) = \sum_{r=1}^t y_r E(u(f(x^r)))$$



subject to the constraint that the variance  $\sigma^2(u(f(\mathbf{y})))$  is not greater than a given threshold  $\bar{\sigma}^2$ , that is

$$\sigma^2(u(f(\mathbf{y}))) = \sum_{r=1}^t y_r \sigma^2(u(f(x^r))) + 2 \sum_{r=1}^{t-1} \sum_{s=r+1}^t y_r y_s \sigma(u(f(x^r)), u(f(x^s))) \leq \bar{\sigma}^2.$$

The classic Markowitz portfolio selection problem can also be formulated as minimization of the variance  $\sigma^2(u(f(\mathbf{y})))$  under the constraint that the expected utility value  $E(u(f(\mathbf{y})))$  is not smaller than a given threshold  $\bar{E}$ .

Coming back to our example, let us suppose that the DM wants to compute the vector  $\mathbf{y} = [y_1, \dots, y_4]$  having the maximum expected utility value  $E(u(f(\mathbf{y})))$  but under the condition that the variance  $\sigma^2(u(f(\mathbf{y})))$  is not greater than 0.025. In this case, the optimal vector is

$$\mathbf{y}^1 = [0 \quad 0.4223 \quad 0.3487 \quad 0.2289],$$

with its corresponding expected utility value  $E(u(f(\mathbf{y}^1))) = 0.5644$  and variance  $\sigma^2(u(f(\mathbf{y}^1))) = 0.025$ .

Instead, if we suppose that the DM wants to compute a vector  $\mathbf{y} = [y_1, \dots, y_4]$  having the minimum variance  $\sigma^2(u(f(\mathbf{y})))$  but under the condition that the expected utility value  $E(u(f(\mathbf{y})))$  is not smaller than 0.56, then the optimal vector is

$$\mathbf{y}^2 = [0 \quad 0.1599 \quad 0.4285 \quad 0.2118]$$

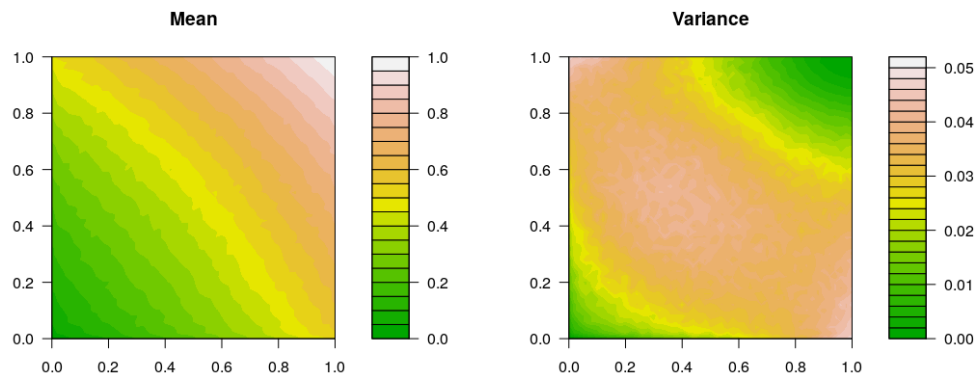
with its corresponding expected utility value  $E(u(f(\mathbf{y}^2))) = 0.56$  and variance  $\sigma^2(u(f(\mathbf{y}^2))) = 0.0224$ .

Let us finally remark, that the value of  $y_r, r = 1, \dots, t$ , can be interpreted as a score assigned by a fitness function to the corresponding solution  $x^r$  in an evolutionary optimization algorithm, such that the greater the value of  $y_r$  the more probably  $x^r$  should be selected to generate a new solution.

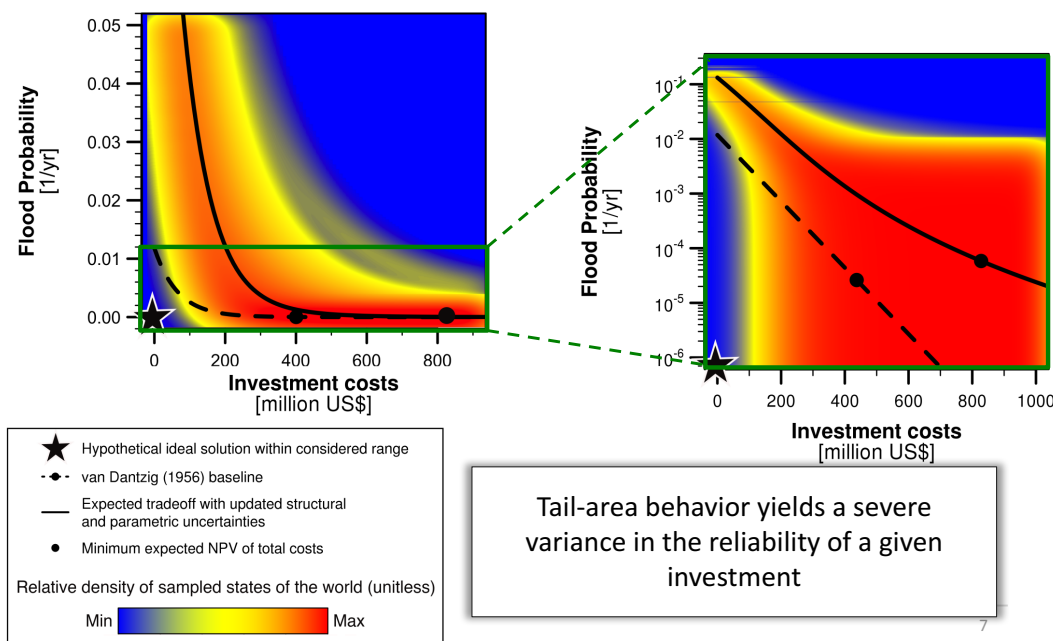
**Heat map visualization of averages and variances.** For a visualization of the situation that is described above consider Figure 2. For any two-dimensional input/decision/design/output variable  $x = (x_1, x_2)$  in the domain  $[0, 1] \times [0, 1]$ , we can compute the mean and variance of the  $l \cdot k$  (here,  $3 \cdot 4 = 12$ ) entries of the resulting matrix  $U(x)$  or  $U(y)$ . Then, the figure on its left and right side shows the thus computed mean values and variances for variables  $x$  or  $y$  of a discretized grid on  $[0, 1] \times [0, 1]$ .

#### 4.1.4 Application to sea-level rise and storm surge projections

This section describes a real-world application regarding the deep uncertainties in sea-level rise and storm surge projections. This example represents a probabilistic generalization of the classical Van Dantzig decision analytical application where the decision is to choose the level of increase in dike height to reduce flood risk [1]. The two objectives are probabilistic as a function of uncertainties in sea level rise due to climate change and local effects of the geophysics of storm surge (i.e., two different but interdependent geophysical models). Figure 3 illustrates the original deterministic Van Dantzig baseline, the mean trade-off between flood risk and investment, as well as the relative locations of the minimum net present values for investment. The challenge as emphasized in the log scale zoomed view is the mean Pareto front would not provide a DM an understanding of the severe variance



■ **Figure 2** Heat map visualization of averages (on the left) and variances (on the right).



■ **Figure 3** Real-world application about uncertainty in technical information (adapted from [1]).

in the potential outcomes for a given investment. For example, working with the mean trade-off an investment 800 Million US Dollars intended to provide a 1 in 10,000 year level of flood protection has a significant residual probability of dramatically less protection (severe damages and potential loss of life). This probabilistic Pareto space context poses a challenge to decision making, particularly given the potential uncertainties in preferences or risk aversion for the residual risks. It then motivates the question of understanding the potential joint probabilistic outcome of uncertain Pareto performance and uncertain DM's preferences.

#### 4.1.5 Open questions

In this report, we proposed a novel approach for interactive multi-objective optimization taking into account uncertainty referring to both the evaluations of solutions by objective functions as well as the preferences of the decision maker. We envisage the following directions for future research.

Firstly, we aim at developing methods for elicitation of probability distributions on objective performances and on utility functions. Secondly, we will propose some procedures for robustness analysis that would quantify the stability of results (utilities, ranks, and pairwise relations) obtained in view of uncertain performances and preferences. Thirdly, when aiming to select a set of feasible options, we will account for the interactions between different solutions. Fourthly, we will integrate the proposed methods with evolutionary multi-objective optimization algorithms with the aim of evaluating and selecting a population of solutions. Fifthly, we plan to adapt the introduced approach to a group decision setting, possibly differentiating between two groups of decision makers being responsible for, respectively, setting the goals and compromising these goals based on different utilities. Finally, we will apply the proposed methodology to real-world problems with highly uncertain information about the solutions' performances and decision makers' preferences.

#### References

- 1 P. C. Oddo, B. S. Lee, G. G. Garner, V. Srikrishnan, P. M. Reed, C. E. Forest, and K. Keller. Deep uncertainties in sea-level rise and storm surge projections: Implications for coastal flood risk management. *Risk Analysis*, page (in press).
- 2 B. Roy. Main sources of inaccurate determination, uncertainty and imprecision in decision models. *Math. Comput. Modelling*, 12:1245–54, 1989.
- 3 B. Roy. Do not confuse. *Int. J. Multicriteria Decision Making*, 6:112–17, 2016.
- 4 I. Yevseyeva, A. P. Guerreiro, T. T. M. Emmerich, and C. M. Fonseca. A portfolio optimization approach to selection in multiobjective evolutionary algorithms. In *International Conference on Parallel Problem Solving from Nature*, pages 672–681. Springer, 2014.

## 4.2 Personalization of multicriteria decision support systems (WG2)

*Matthias Ehrgott, Gabriele Eichfelder, Karl-Heinz Küfer, Christoph Lofi, Kaisa Miettinen, Luís Paquete, Stefan Ruzika, Serpil Sayin, Ralph E. Steuer, Theodor J. Stewart, Michael Stiglmayr, and Daniel Vanderpooten*

**License** © Creative Commons BY 3.0 Unported license

© Matthias Ehrgott, Gabriele Eichfelder, Karl-Heinz Küfer, Christoph Lofi, Kaisa Miettinen, Luís Paquete, Stefan Ruzika, Serpil Sayin, Ralph E. Steuer, Theodor J. Stewart, Michael Stiglmayr, and Daniel Vanderpooten

**Abstract.** In this report, personalization is approached from a learning perspective. We propose a framework for a decision support system to help a decision maker who faces the problem of identifying a most preferred from among a set of alternatives. Our framework encompasses the idea that the objectives and the constraints of the model may not be clear at the beginning and are likely to evolve throughout the decision process. Our proposal deviates from the vast literature on interactive methods by allowing the model to evolve in a very flexible way. We illustrate the need of personalized decision support systems with some applications. We also discuss ways to present solutions to a decision maker in a qualitative manner as this is an important part of the iterative learning and solution process.

#### 4.2.1 Introduction

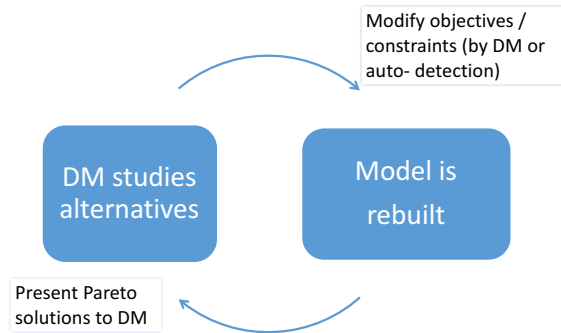
We approach personalization from a learning perspective and propose a framework for a decision support system to help a decision maker (DM) who faces the problem of identifying a most preferred solution from among a set of alternatives. Our framework is general in the sense that it allows for a continuous and discrete expression of alternatives. The alternatives may be explicitly available or may be defined implicitly via some functions (objectives and constraints). Our framework encompasses the idea that the objectives and the constraints of the model may not be clear at the beginning and are likely to evolve throughout the decision process. Thus, the process by which the DM modifies his/her perception of preferences through restructuring of the hierarchical decision model must be facilitated. This can be achieved, for instance, by adding/subtracting objectives, aggregating/disaggregating objectives, modifying constraints, converting constraints into objectives and vice versa while retaining insights gained from earlier phases of the analysis. Figure 4 illustrates the iterative decision making process. Some Pareto optimal solutions of an initial model are studied by a DM. These solutions reveal some findings about the problem to the DM or help him/her discover one's preferences. These are taken into account in a revised model and some carefully revised new Pareto optimal solutions are presented to the DM on the next round, and so forth. The process continues until the DM identifies a most preferred solution.

Our proposal deviates from the vast literature on interactive methods by allowing the model to evolve in its degree of flexibility. As the objectives and constraints of the model are modified, the Pareto optimal set shifts and changes. We have seen studies in the literature in which the solution method is switched depending on the phase of the solution process, i.e., the search. However, in these studies the model usually stays the same. Here, we understand personalization as enabling the model to evolve.

The influence of adding and subtracting objective functions to a multiobjective optimization problem has been considered in [9]. Furthermore, the relative importance of objectives is discussed and a definition of weights is given in [19, 20] (where weights are called coefficients) for objective functions as well as for groups of objective functions. This approach results in a convex combination of functions similar to linear combinations as discussed in [3]. There, strategies are discussed that reduce the size of the solution set of the multiobjective optimization problem for instance by combining several objectives linearly, i.e. by summing them up, before employing tools to solve the resulting multiobjective optimization problem. Using partial preference models, where weights are partially defined, is also a way of focusing on reduced solution sets of interest [13].

The need of iterating to find an appropriate model of a real-world problem to be solved is demonstrated in [2, 26] with cases in optimal shape design of an air intake channel and a two-stage separation process, respectively. In the latter case, an interactive multiobjective optimization method helped in validating and improving the model and only after that kind of iterating, the actual interactive solution process was conducted.

The idea of constraint optimization using multiobjective optimization models, i.e. the idea to transform constraints to objectives, as well as the other way around, is studied in [15]. Furthermore, e.g., in [16], it is demonstrated that converting a problem with one objective and four very demanding constraints can be solved by optimizing constraint violations besides the original objective, i.e., a problem with five objectives. Hence, in the literature, the relation between constrained and multiple objectives as well as between aggregated and disaggregated multiobjective optimization problems is already studied at least in parts, while several such models have so far not been used in an iterative manner on varying levels for steering a decision-making process.



■ **Figure 4** The framework for personalized decision support.

In [12], an unconstrained bi-objective discrete optimization problem is studied with the goal of finding representations that adhere to a given quality with respect to the  $\epsilon$ -indicator measure. The suggested approach is related to the Nemhauser-Ullman algorithm that has been proposed for the traditional knapsack problem which has one objective function and one constraint. The work of [23] brings this idea closer to the discussion in this report by formulating a bi-dimensional knapsack problem where one of the constraints is a soft constraint. The authors model the soft constraint as an objective function, thereby ending up with a uni-dimensional knapsack problem with two objectives. As such, they propose to compute representative solutions for the transformed problem so as to portray the trade-off between the objective function of the original problem and satisfaction or violation of its soft constraint.

In [11, Section 2] and [5, Section 3], a detailed review on the literature on modeling the relative importance of objectives is provided. In these references, as well as in [6], partial orderings, other than the natural orderings via (non)polyhedral cones, are examined for their impact on optimal (in that case, efficient) solution sets of multiobjective optimization problems. These examinations might help in understanding the relationship between (dis)aggregated multiobjective optimization problems.

### Problem formulation

To give a mathematical formulation of the problem of adding/subtracting and (dis)aggregating objectives, we make the following assumptions:

- Let a nonempty subset  $X \subseteq \mathbb{R}^n$  be given which describes the set of alternatives. For instance, the set  $X$  might be determined by some hard constraints given by laws of nature, which cannot be weakened and, thus, cannot be transformed to objective functions.
- Let  $\mathcal{F} := \{f_i: \mathbb{R}^n \rightarrow \mathbb{R} \mid i = 1, \dots, k\}$  be a finite collection of functions which are potentially of interest for particular models. Then, for particular model instances, some of these functions can appear in the formulation of the objective functions or in the constraints.
- Let  $h_1, \dots, h_m, g_1, \dots, g_l: \mathbb{R}^k \rightarrow \mathbb{R}$  be arbitrary functions describing which of the functions  $f \in \mathcal{F}$  are aggregated or chosen for the formulation of the individual objective functions or constraints of the particular model. Thereby,  $m \in \mathbb{N}$  and  $l \in \mathbb{N}$  also depend on the particular model instance.

Under these assumptions, a particular model instance can be expressed as

$$\min_{x \in S} h_1(f_1(x), \dots, f_k(x)), \dots, h_m(f_1(x), \dots, f_k(x)), \quad (\text{PMI})$$

where  $S := \{x \in X \mid g_i(f_1(x), \dots, f_k(x)) \leq 0, i = 1, \dots, l\}$ .

► **Example 1.** Let  $X = \mathbb{R}^n$  and  $\mathcal{F} = \{f_1, f_2, f_3: \mathbb{R}^n \rightarrow \mathbb{R}\}$ . For the aggregation functions  $h_j$  we only take linear combinations and selections into account. Thus, let weights  $w_2, w_3 > 0$  be given. With  $h_1(y_1, y_2, y_3) = y_1$  and  $h_2(y_1, y_2, y_3) = w_2 y_2 + w_3 y_3$  we get

$$\min_{x \in X} \begin{pmatrix} f_1(x) \\ w_2 f_2(x) + w_3 f_3(x) \end{pmatrix}. \quad (P_A(w_2, w_3))$$

The corresponding disaggregated multiobjective optimization problem with functions  $h_1(y_1, y_2, y_3) = y_1$ ,  $h_2(y_1, y_2, y_3) = y_2$ , and  $h_3(y_1, y_2, y_3) = y_3$  is

$$\min_{x \in X} \begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{pmatrix}. \quad (P_D)$$

When disaggregating the problem  $(P_A(w_2, w_3))$  one might be interested in keeping the properties of an already found Pareto optimal solution  $\bar{x} \in X$  of the bi-objective problem  $(P_A(w_2, w_3))$ . For instance, it might be the aim to keep the achieved level for the value  $f_1(\bar{x})$  while being willing to explore nearby values for  $f_2$  and  $f_3$ . With  $g_j(y_1, y_2, y_3) = y_j - \Delta_j$  for  $j = 1, 2, 3$  and with

$$\Delta_1 = f_1(\bar{x}), \quad \Delta_2 = \Delta_3 = w_2 f_2(\bar{x}) + w_3 f_3(\bar{x}) + \delta$$

for some scalar  $\delta \geq 0$ , also the following problem might be of interest.

$$\begin{aligned} \min \quad & \begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{pmatrix} \\ \text{s.t.} \quad & f_1(x) \leq \Delta_1, \\ & f_2(x) \leq \Delta_2, \\ & f_3(x) \leq \Delta_3, \\ & x \in X, \end{aligned} \quad (P_C(\Delta))$$

where we can write  $S = \{x \in X \mid f_i(x) \leq \Delta_i, i = 1, 2, 3\}$ .

The following relations are, for instance, obvious:

- If a point  $\bar{x} \in X$  is Pareto optimal for  $(P_D)$ , then  $\bar{x}$  is also Pareto optimal for  $(P_C(\Delta))$  for any  $\Delta \in \mathbb{R}^3$  with  $\Delta_i \geq f_i(\bar{x})$ ,  $i = 1, 2, 3$ .
- If a point  $\bar{x} \in X$  is Pareto optimal for  $(P_C(\Delta))$  for any  $\Delta \in \mathbb{R}^3$ , then  $\bar{x}$  is also Pareto optimal for  $(P_D)$ .
- If a point  $\bar{x} \in X$  is Pareto optimal for  $(P_A(w_2, w_3))$  for some weights  $w_2, w_3 > 0$ , then  $\bar{x}$  is also Pareto optimal for  $(P_D)$ .

Interesting questions are also, for instance, under which assumptions a Pareto optimal point  $\bar{x}$  of  $(P_A(w_2, w_3))$  for some weights  $w_2, w_3 > 0$  is at least feasible for  $(P_C(\Delta))$  for  $\Delta_1 \geq f_1(\bar{x})$ ,  $\delta \geq 0$  and

$$\Delta_2 = \Delta_3 = w_2 f_2(\bar{x}) + w_3 f_3(\bar{x}) + \delta.$$

### 4.2.2 Applications

Next we illustrate the need of personalized decision support systems with some applications.

#### Radiotherapy

In radiotherapy, the set of alternatives consists of applicable treatment plans  $x$ . We assume that the alternatives are judged by the DM solely based on the properties of the resulting dose distribution. At the highest level, the properties of the dose distribution predict the likelihood of treatment success or failure as well as the likelihood of specific complications and side effects related to the organs at risk. To represent and compute the dose distribution, the patient image is divided into (up to millions of) equal-sized voxels. The dose distribution  $D(x)$  is then the vector of all voxel dose values. For the sake of simplicity, each voxel either belongs to a target, to a specific organ at risk, or to normal tissue. For each target, there is a prescribed dose  $d^{presc}$  that is deemed adequate to kill all tumor cells. For evaluating a given dose distribution, a large collection of objective functions has been established in the radiotherapy community. Most of these objective functions in some way measure the average under- or overdose over all voxels belonging to a specific structure (target or organ at risk). However, other (“lower-level”) aspects of the dose distribution also play a role, such as smallish localized areas of too high dose far away from the target (“hot spots”). This is where aggregation and disaggregation come into play.

**Aggregation and disaggregation in radiotherapy.** The dose values in the individual voxels form a natural basis of lowest level and highest detail when assessing the dose distribution. The following implications can be assumed to hold for any DM’s utility function:

- For target voxels  $i$ , as long as the dose values are below the prescribed dose,  $d_i(x) < d_i(x')$  and all else equal, this implies that  $x$  is a worse treatment plan than  $x'$ .
- For target voxels  $i$ , as long as the dose values are above the prescribed dose,  $d_i(x) < d_i(x')$  and all else equal, this implies that  $x$  is a better treatment plan than  $x'$ .
- For risk and normal tissue voxels  $j$ ,  $d_j(x) < d_j(x')$  and all else equal, this implies that  $x$  is a better treatment plan than  $x'$ .

Fundamental (“atomic” or “lowest-level”) objective functions  $\mathcal{F}$  can be chosen as representations of these relations:

- For target voxels  $i$ :  $f_i^{UD}(x) = \max\{0, d^{presc} - d_i(x)\}$ .
- For target voxels  $i$ :  $f_i^{OD}(x) = \max\{0, d_i(x) - d^{presc}\}$ .
- For risk and normal tissue voxels  $j$ :  $f_j(x) = d_j(x)$ .

A decision process based on  $\mathcal{F}$  is infeasible. Given two unrelated dose distributions, a comparison may well exceed the mental capacity of a DM. Even if a trajectory is provided where in each comparison only a few fundamental functions differ, the search space would be too large and any search too unstructured for efficient decision making. Thus, “higher-level” functions are introduced that aggregate all fundamental functions of voxels of the same structure, for example, the squared organ at risk dose:

$$f_{risk}(x) = \sum_{j \in risk} (d_j(x))^2. \quad (20)$$

The aggregation simplifies the problem by treating every voxel within the structure as equal, disregarding position and spatial relationship to other voxels. Also, it handles the trade-off within the voxels of the same structure automatically, depending on the exact formulation of the aggregation (which can be chosen by the DM).

On the other hand, the aggregated function can cloud lower-level aspects of the DM's utility function. For example, the DM may be happy with the overall amount of dose for the organ at risk, but there is a certain region inside the organ at risk that still gets too much dose. One option would be to choose a different aggregated function, maybe using a higher coefficient in order to penalize higher doses more and force a different trade-off of voxel doses inside the organ at risk.

However, the discontent may be attributed more to the specific location and the spatial accumulation of higher dosed voxels, rather than the values themselves. In this case, the assumptions made when aggregating the fundamental functions – namely that all voxels are equally independent of location and spatial relationship to other voxels – breaks down. In this case, lower-level functions may need to be (re-)introduced in the variable model, i.e. the model must be disaggregated.

### Land use planning

Land use planning involves the allocation of facilities to specific locations or activities to specific areas within a region of land. In most non-trivial contexts, land-use planning involves many criteria, some at least of which will involve partially qualitative considerations such as social impacts of displacements, destruction of old burial sites and effects of biodiversity reduction. Typically also, conflict is generated between multiple stakeholders that needs some resolution before any decision can be implemented.

Two examples of land use planning problems with which one of the authors has been associated are the following. The first related to replacement of indigenous afro-montane grasslands on the eastern escarpment areas of South Africa by exotic commercial forestries [27]. The prime decision variables related to proportions of the region allocated to forestry, with subsidiary considerations including water supply to rural communities for subsistence and agriculture, and preservation of biodiversity in the region. The second example arose from restoration of land for nature conservation with associated partitioning of land into intensive and extensive agriculture, as well as other development activities, in the Netherlands [4]. The prime decision variables were binary, i.e. selection of activity for each designated parcel of land.

Land use planning provides a challenging context within which to seek personalization of decision support. Different stakeholders will have different perspectives on the same problem, which need to be provided for. As different groups work together and negotiate, problem structures and preference perceptions evolve dynamically, and this too needs to be captured in the decision support system.

Some dynamic issues which arose in these examples included the following:

- A need to incorporate policy (not entirely hard) constraints into the forestry development problem, that for any chosen proportion of area to forestry, the precise locations of the plantations were to be subject to environmental impact vetoes;
- The original decision support models for selection of land parcel activities focused on assessing the value of allocating each activity to each parcel as primary objectives. But deeper reflection led to a realization that system management requires the definition of further system-related criteria concerned with coherency of activities which are non-additively related to decision variables.



- In the water resources component of the South African forestry land allocation, one initially identified criterion was interests of rural village communities. But problems encountered while attempting to evaluate decision alternatives according to this criterion led to a realization that there were two relevant sub-criteria, that could be seen as “female” (close access to clean water) and “male” (availability of piecework on commercial farms).

Any decision support system must be able to cope with such often unexpected developments in the problem structure as regards both the decision space and the set of criteria.

#### 4.2.3 Research questions

In the following, we discuss some of the main research questions that need to be addressed in a personalized iterative decision making process as described in previous sections.

#### Aggregating/disaggregating functions as objectives and constraints

We start again by motivating our research questions with an example. Let us consider a problem where a DM wants to minimize cost  $f(x)$  and maximize quality  $g(x)$  of a product to be purchased:

$$\min f(x), \max g(x).$$

The quality may consist of two separate components:  $g(x) = w_1 g_1(x) + w_2 g_2(x)$ , cf. Example 1.

Let us suppose that a solution  $\bar{x}$  is identified by the DM after a first depiction of the Pareto front (in the objective space) of this problem. Now, the question is to find new solutions, not too far away from  $\bar{x}$ , of a possible disaggregated problem. Then one might solve the problem

$$\min f(x), \max g_1(x), \max g_2(x)$$

or

$$\begin{aligned} & \min f(x) \\ & \text{s.t.} \\ & g_1(x) \leq g(\bar{x}) + \Delta_1, \\ & g_2(x) \leq g(\bar{x}) + \Delta_2. \end{aligned}$$

Open research questions include:

- Are the relationships between reformulations of the problem stronger if  $g_1$  and  $g_2$  are somehow correlated? Does the strength of the relationship depend on  $\bar{x}$ ? From a practical point of view, is the non-correlated case of even more interest?
- How can a recommendation for an initial aggregation be made in order to start the decision-making process? How can objectives be added or removed? There can be settings when the model is blank (unknown) or very well-known. In the first case, the model is to be built by adding, in the other, by removing.
- An expressed constraint may be found to be irrelevant after learning that the range is too narrow to be relevant. The question is how to model this automatically.
- An objective can be converted into a constraint to eliminate unwanted alternatives or to save levels with specific objectives. The question is how to structure such approaches and what are the relations between the solutions found.

### Navigation

To form a good base for the selection step of a solution  $\bar{x}$  in an iterative process, a good presentation and a way to navigate between possible solutions is required. We state some known approaches as well as some open questions in the following.

**Navigation in a continuous space of alternatives.** For a continuous multiobjective optimization problem, a real-time navigation capability for the DM such as the following two-step process can be offered:

1. Optimizing a set of representative solutions  $x_1, \dots, x_m$  in an offline pre-computation, with objective function vectors  $F_i = F(x_i)$ . Explicitly or implicitly, the representative pairs  $(x_i, F_i)$  must have a neighborhood relationship defined, allowing neighboring solutions to be linearly interpolated. This means that for a subset  $I$  of mutually neighboring points, and for coefficients  $\lambda_i \geq 0$  with  $\sum_{i \in I} \lambda_i = 1$ 
  - any interpolated solution  $x = \sum_{i \in I} \lambda_i x_i$  is feasible,
  - for any interpolated point  $x$ , the objective function values  $F(x)$  differ from the Pareto optimal achievable values only by an acceptable error (“approximation quality”),
  - $F(\sum_{i \in I} \lambda_i x_i) \approx \sum_{i \in I} \lambda_i F_i$  in order for the navigation mechanisms of the item above to work (“triangulation of Pareto front approximation”).
2. Searching the space of interpolated solutions in real-time. This can be done by solving linear optimization problems in the interpolation coefficients.

For convex problems, this is understood (see, “sandwiching” [24] for the calculation of the representative solutions, and real-time navigation in [7, 17, 18]), but maybe not published well enough yet. In the convex case, many of the ingredients mentioned above come for free (neighborhood from calculating the convex hull, feasibility of interpolated solutions) or coincide (second and third bullet points as a consequence of sandwiching). For general nonconvex problems, this is not the case. One way of connecting objective and decision spaces for nonconvex problems has been proposed in [10]. Research questions include:

- Formalizing the approach, maybe embedding the convex case as a special case, in order to make it more known and understood in the community.
- Properties of nonconvex problems to facilitate this approach.
- Development, improvement, and description of algorithms for the calculation of representative solutions and for real-time navigation especially for the nonconvex case.

**Navigation in a discrete space of alternatives.** In a discrete case, the DM wants to find the preferred solution out of a finite but typically large set of alternatives. Such a decision problem can also be handled by real-time navigation mechanisms. However, interpolation is not possible. Thus, when traversing a set of alternatives, the direction and size of each navigational step cannot be controlled very well. Research questions include:

- How can the wishes of a DM be stated and interpreted in the context of discrete navigation?
- Should the DM follow a trajectory by jumping from alternative to alternative? If yes, how should the next alternative be chosen? Can this choice be defined by a particular distance measure or neighborhood relationship in the space of alternatives?
- Or should the navigation mechanism focus more on eliminating alternatives?

#### 4.2.4 Toward personalizing representations

Personalization is very much related to learning. In different domains, there may be different aspects of learning. Expert DMs may have a good understanding of the structure of a decision problem but they may still need to learn about the nature of the problem instance

(e.g. in radiotherapy) and gain insight in the conflicting nature of the objectives and feasible solutions as well as the feasibility of their preferences. Novice DMs may need to discover their objectives, constraints and solutions.

Throughout this section, we assume that some model (as a result of processes described in the sections before) is given together with an explicit list of  $n$  alternatives (e.g. items/products). The properties of these alternatives are described by criteria (defined on measurable scales). This is for example the case in online sales or consulting systems where customers are supported in choosing some product meeting their individual demands.

In many such practical applications, the set of Pareto optimal solutions exceeds a manageable cardinality. In order to analyze or visualize the set of alternatives and, thus, to assist the process of making a final decision, the DM requires a concise representation of the Pareto optimal set to obtain a quick overview. A good representation can still communicate the nature of the set while hiding options which are not informative. In the following, we investigate the influence of personalization on representations, adaptation of quality measures incorporating personal preferences and algorithms to compute a personalized representation in the context of explicitly given alternatives.

An idea to incorporate personalization in the computation of a representative subset is based on two functionalities which can be in principle applied in an arbitrary order during a decision making process:

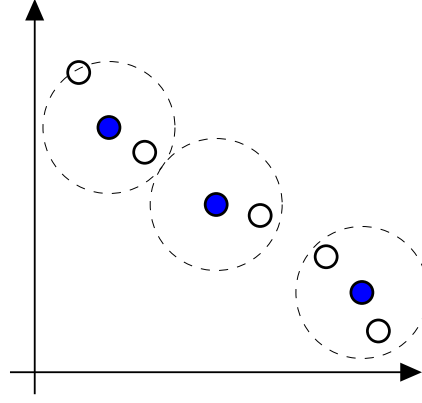
1. The computation of a good and concise representation for a given region of interest.
2. The determination of the set (or a representation) of neighbors wrt. to a selected point.

During the search for a finally preferred solution, a DM may iteratively make use of these two functionalities: A good representation for the problem/model at hand may be computed and analyzed, the model may be changed and the first functionality may be invoked again, or, eventually, a DM may be interested in the neighborhood of some selected point to be informed about similar alternatives. Before presenting some specific algorithmic ideas, we discuss these two functionalities in more detail first.

Concerning functionality 1, a crucial point relates to the notion of “goodness” of a representation, i.e. the quality of a representation. Certainly, one goal is to determine a representation  $R$  of the set of Pareto optimal points (also known as nondominated points)  $Y_N \in \mathbb{R}^p$  which is tractable for the DM and can be efficiently computed. We rely on the classical quality measures for discrete representations suggested and discussed in [8, 22], namely coverage, uniformity and cardinality which can be roughly characterized as follow.

- *Coverage*: any point in  $Y_N$  is represented or *covered* by at least one point in  $R$ .
- *Uniformity*, also called *spacing*: any two points in  $R$  are sufficiently *spaced*, avoiding redundancies.
- *Cardinality* refers to the cardinality  $|R|$  of the representation  $R$ . Since each representative point has to be computed with a certain effort, the cardinality should be small.

The concepts coverage and *spacing* can be implemented in a variety of ways. In principle, one can distinguish between a *geometric vision* based on *distances* and a *preference-oriented vision* using some *preference relation*. In a geometric vision, distances between points in  $Y_N$  and points in  $R$  are used to evaluate coverage. Likewise, uniformity is evaluated by calculating pairwise distances between points in  $R$ . Alternatively, a preference-oriented vision is based on a preference relation  $\preceq$ . For two points  $y$  and  $y'$ , one can then say that  $y$  *covers*  $y'$  if  $y \preceq y'$  which implies a notion of coverage. Analogously,  $y$  and  $y'$  are sufficiently *spaced* if not  $(y \preceq y')$  and not  $(y' \preceq y)$  which then defines the notion of uniformity/spacing.



■ **Figure 5** Illustration of a representation based on coverage.

#### 4.2.5 Algorithmic approaches for computing personalized representations

Based on the discussion in the previous section, several methods existing in the literature are proposed, which can be adapted, to meet the two functionalities mentioned. The first three methods are geometric-based approaches, while the fourth one is a preference-based approach. These efforts may be understood as a first attempt of computing personalized representations.

##### A geometric-based approach

In a geometric vision, coverage measures the quality of the representative subset by considering the distance of the unchosen elements to their closest elements in the subset. Formally, the coverage of a subset  $R \subseteq Y_N$  is computed as

$$I_C(R, Y_N) = \max_{y \in Y_N} \min_{y' \in R} \|y - y'\|.$$

The coverage representation problem consists of finding a subset of cardinality  $k$  that has the smallest coverage value, i.e.,

$$\min_{\substack{R \subseteq Y_N \\ |R|=k}} I_C(R, Y_N).$$

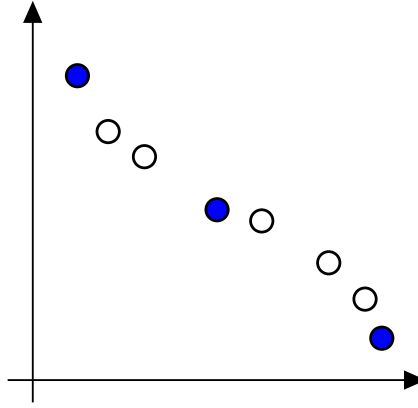
This problem is known as the  $k$ -center problem [14]. In the particular case of two objectives, it can be solved in a polynomial amount of time [28].

Similarly, in a geometric vision, uniformity measures how far apart the  $k$  chosen elements of the set  $R \subseteq Y_N$  are from each other. It is computed as the minimum distance between a pair of distinct elements as

$$I_U(R) = \min_{\substack{y, y' \in R \\ y \neq y'}} \|y - y'\|.$$

The goal of the uniformity representation problem is to find a subset  $R$ , with a given cardinality  $k$ , from a set  $Y_N$  that maximizes  $I_U(R)$ , i.e.,

$$\max_{\substack{R \subseteq Y_N \\ |R|=k}} I_U(R).$$



■ **Figure 6** Illustration of representation based on uniformity.

Note that this problem corresponds to a particular case of the  $k$ -dispersion problem in facility-location [21]. Also for the particular case of two objectives, this problem can be solvable in a polynomial amount of time [28].

Note, that functionality 2 suggests itself in a geometric vision: The neighborhood for the second functionality is an  $\varepsilon'$ -neighborhood of a selected point  $\bar{y}$ :

$$y \in Y_N : \|y - \bar{y}\| \leq \varepsilon'.$$

### The revised boundary intersection method

The revised boundary intersection (RNBI) method computes a discrete representation of the Pareto optimal set of a multiobjective linear optimization problem (MOLP)  $\min\{Cx : Ax \leq b\}$  with a bounded feasible set. It provides guarantees on both the uniformity and the coverage error of the representation, see [25]. The following is a description of the algorithm.

1. Input: MOLP data  $A, b, C$  and  $ds > 0$ .
2. Find  $y^{AI}$  defined by  $y_k^{AI} = \max\{y_k : y \in Y\}$  for  $k = 1, \dots, p$ .
3. Find a Pareto optimal point  $\hat{y}$  by solving the linear problem  $\phi := \min\{e^T y : y \in Y\}$ .
4. Compute  $p + 1$  points  $v^k, k = 0, \dots, p$  in  $\mathbb{R}^p$ 

$$v_l^k = \begin{cases} y_l^{AI}, & l \neq k, \\ \phi + \hat{y}_k - e^T v^0, & l = k. \end{cases}$$
  - The convex hull  $S$  of  $\{v^0, \dots, v^p\}$  is a simplex containing  $Y$ .
  - The convex hull  $\hat{S}$  of  $\{v^1, \dots, v^p\}$  is a hyperplane with normal  $e$  supporting  $Y$  in  $\hat{y}$ .
5. Compute equally spaced reference points  $q^i$  with a distance  $ds$  on  $\hat{S}$ .
6. For each reference point  $q$  solve the linear problem  $\min\{t : q + te \in Y, t \geq 0\}$  and eliminate dominated points from the resulting set  $R$ .
7. Output: Representation  $R$ .

The steps of the algorithm are illustrated in Figure 7.

Theorem 2 provides the quality guarantee for the method in terms of uniformity and coverage error of the generated representation.

► **Theorem 2.** *Let  $R$  be the representation of  $Y_N$  obtained with the RNBI method.*

1. *Let  $q^1, q^2$  be two reference points with  $d(q^1, q^2) = ds$  that yield Pareto optimal representative points  $r^1, r^2$ . Then  $ds \leq d(r^1, r^2) \leq \sqrt{p}ds$ . Hence,  $R$  is a  $ds$ -uniform representation of  $Y_N$ .*



not uniform, but has high-density clusters representing a certain "type" of outcome (e.g. products which are similar). To provide a quick overview of available Pareto optimal points, each of these types/clusters should be represented with one representative point. Clusters can be of different sizes, but would still be represented by a single point.

Such a clustering algorithm for realizing functionality 1 can be formulated as follows:

**Algorithm:** MSF-Clustering

Input:  $n$  items with their objective function values;  $K \in \mathbb{N}$

Output: A representation  $R$  with  $|R| = K$

1. Compute the pairwise distances between the items (wrt. their objective function values).
2. Sort these distances by increasing length.
3. Use Kruskal's algorithm to compute a Minimum Spanning Forest consisting of  $K$  trees (=cluster).
4. For each tree: Compute median/center item as the representative point of the cluster.
5. Return all representative points.

This clustering algorithm can be implemented in a running time of  $O(n^2 + n^2 \cdot \log_2 n^2 + n^2 \cdot \log_2^* n^2 + n^2) = O(n^2 \log n^2)$  and, thus, finds a representation in polynomial time. In case a DM then updates upper bounds on the values of the objectives (this is an operation which is likely to happen), a re-sorting can be implemented in  $O(n^2)$  which results in an  $O(n^2 \cdot \log_2^* n^2)$  algorithm for updating the representation.

Note that functionality 2, i.e. "display solutions close to some chosen representative point" can be very easily realized: all points in a cluster are displayed. Further research directions may clarify the quality of representations (wrt. uniformity and coverage) obtained with such a clustering algorithm.

### A preference-based approach

The first important question is the choice of the preference relation  $\preceq$  to be used to compute the representation  $R$ . Relation  $\preceq$  must be richer than the Pareto dominance relation in order to ensure conciseness of the representation. In cases where no a priori preference information is available, a natural candidate relation is the  $\varepsilon$ -dominance relation  $\preceq_\varepsilon$  defined as follows:

$$y \preceq_\varepsilon y' \text{ iff } y_i \leq (1 + \varepsilon)y'_i \quad i = 1, \dots, p,$$

where  $\varepsilon > 0$  can be interpreted as a tolerance/indifference threshold. Note that we can use different thresholds  $\varepsilon_i > 0$  for each criterion  $f_i$ ,  $i = 1, \dots, p$ . We can also use additive thresholds instead of multiplicative thresholds. The relation  $\preceq_\varepsilon$  enriches the standard Pareto dominance relation as illustrated in Figure 8.

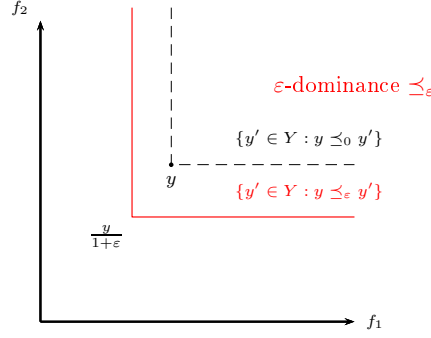
In order to implement functionality 1, which aims at producing a concise representation of a region of interest, we use the concept of an  $(\varepsilon, \varepsilon')$ -kernel, introduced in [1].

► **Definition 3.** Given  $\varepsilon, \varepsilon' > 0$ , an  $(\varepsilon, \varepsilon')$ -kernel is a set of points  $K_{\varepsilon, \varepsilon'} \subset Y$  satisfying:

- (i) for any  $y' \in Y_N$  there exists  $y \in K_{\varepsilon, \varepsilon'}$  such that  $y \preceq_\varepsilon y'$  ( $\varepsilon$ -coverage),
- (ii) for any  $y, y' \in K_{\varepsilon, \varepsilon'}$ , not( $y \preceq_{\varepsilon'} y'$ ) and not( $y' \preceq_{\varepsilon'} y$ ) ( $\varepsilon'$ -stability).

In order to guarantee the existence of an  $(\varepsilon, \varepsilon')$ -kernel, we must have  $\varepsilon' \leq \varepsilon$ . Considering that condition (i) prevails over condition (ii) in the definition of a good representation, we must first define a threshold  $\varepsilon$  to define the precision of the representation and then set  $\varepsilon'$  as large as possible. When it is possible to set  $\varepsilon' = \varepsilon$ , an  $(\varepsilon, \varepsilon')$ -kernel is called an  $\varepsilon$ -kernel.

Some important results, established in [1], are gathered in the following theorem.



■ **Figure 8** Dominance ( $\preceq_0$ ) and  $\epsilon$ -dominance ( $\preceq_\epsilon$ ) relations.

- **Theorem 4.** ■ *If  $p = 2$ , an  $\epsilon$ -kernel always exists (with  $\epsilon' = \epsilon$ ).*  
 ■ *If  $p \geq 3$ , an  $(\epsilon, \epsilon')$ -kernel exists if and only if  $\epsilon' \leq \sqrt{1 + \epsilon} - 1$ .*  
 ■ *If  $Y$  is defined explicitly, these concepts can be computed in a linear time.*

We show now how to implement functionality 2, which aims at producing alternatives similar to a (not necessarily) feasible reference point. Let  $\bar{y}$  be the reference point. The neighborhood of  $\bar{y}$  is:

$$\mathcal{N}(\bar{y}) = \{y \in Y_N : y \preceq_{\epsilon'} \bar{y} \text{ and } \bar{y} \preceq_{\epsilon'} y\}.$$

Note that this neighborhood is defined with a relation  $\preceq_{\epsilon'}$  which is used in the stability condition to define an  $(\epsilon, \epsilon')$ -kernel. It is indeed consistent to use this relation which was used to impose that two elements in  $R$  should not be too similar. This concept is clearly computable in a linear time.

#### 4.2.6 Conclusions

This report summarizes our findings on the topic of personalization of multicriteria decision support systems. With growing computational power, ever enlarging data storage capabilities, then increasing availability of large data sets and the success of multiobjective optimization methods, decision-making processes tend to ask more and more in the way of personalized aspects to make better, faster and more confident decisions. This is especially true on complex, professional applications which require sophisticated models and solution algorithms (e.g. radiotherapy treatment or landuse planning). In addition, everyday applications (such as online evaluations of products or sales for customers) with explicitly given sets of alternatives are subject to multiple criteria and a personalized perspective as well. This report identifies two central aspects which can be concisely described as “personalization in model building” for complex situations and “iterative computation of personalized representation systems” for explicitly given points. Initial ideas are presented with respect to both aspects. Yet, as the demand for personalization grows, more sophisticated concepts are still to be developed.

#### References

- 1 C. Bazgan, F. Jamain, and D. Vanderpooten. Discrete representation of the non-dominated set for multi-objective optimization problems using kernels. *European Journal of Operational Research*, 260(3):814–827, 2017.



- 2 T. Chugh, K. Sindhya, K. Miettinen, Y. Jin, T. Kratky, and P. Makkonen. Surrogate-assisted evolutionary multiobjective shape optimization of an air intake ventilation system. In *Proceedings of the 2017 IEEE Congress on Evolutionary Computation*, pages 1541–1548. IEEE, 2017.
- 3 S. Dempe, G. Eichfelder, and J. Fliege. On the effects of combining objectives in multi-objective optimization. *Mathematical Methods of Operations Research*, 82(1):1–18, 2015.
- 4 T. Eikelboom, R. Janssen, and T.J. Stewart. A spatial optimization algorithm for geodesign. *Landscape and Urban Planning*, 144:10–21, 2015.
- 5 A. Engau and M.M. Wiecek. Cone characterizations of approximate solutions in real vector optimization. *Journal of Optimization Theory and Applications*, 134(3):499–513, 2007.
- 6 A. Engau and M.M. Wiecek. Introducing nonpolyhedral cones to multiobjective programming. In V. Barichard, X. Gandibleux, and V. T’Kindt, editors, *Multiobjective Programming and Goal Programming, Proceedings*, pages 35–45. Springer, 2009.
- 7 P. Eskelinen, K. Miettinen, K. Klamroth, and J. Hakanen. Pareto Navigator for interactive nonlinear multiobjective optimization. *OR Spectrum*, 23:211–227, 2010.
- 8 S. Faulkenberg and M.M. Wiecek. On the quality of discrete representations in multiple objective programming. *Optimization and Engineering*, 11(3):423–440, 2010.
- 9 J. Fliege. The effects of adding objectives to an optimisation problem on the solution set. *Operations Research Letters*, 35(6):782–790, 2007.
- 10 M. Hartikainen and A. Lovison. PAINT-SiCon: Constructing consistent parametric representations of Pareto sets in nonconvex multiobjective optimization. *Journal of Global Optimization*, 62(2):243–261, 2015.
- 11 B.J. Hunt, M.M. Wiecek, and C.S. Hughes. Relative importance of criteria in multiobjective programming: A cone-based approach. *European Journal of Operational Research*, 207(2):936–945, 2010.
- 12 A.D. Jesus, L. Paquete, and J. Figueira. Finding representations for an unconstrained bi-objective combinatorial optimization problem. *Optimization Letters*, (12):321–334, 2018.
- 13 S. Kaddani, D. Vanderpooten, J.M. Vanpeperstraete, and H. Aissi. Weighted sum model with partial preference information: application to multi-objective optimization. *European Journal of Operational Research*, 260(2):665–679, 2017.
- 14 O. Kariv and S. L. Hakimi. An algorithmic approach to network location problems. I: The p-centers. *SIAM Journal on Applied Mathematics*, 37(3):513–538, 1979.
- 15 K. Klamroth and J. Tind. Constrained optimization using multiple objective programming. *Journal of Global Optimization*, 37(3):325–355, 2007.
- 16 K. Miettinen, M.M. Makela, and T. Mannikko. Optimal control of continuous casting by nondifferentiable multiobjective optimization. *Computational Optimization and Applications*, 11:177–194, 1998.
- 17 M. Monz. Pareto navigation - interactive multiobjective optimisation and its application in radiotherapy planning. *Technical University of Kaiserslautern, Department of Mathematics*, 2006.
- 18 M. Monz, K.-H. Küfer, T. Bortfeld, and C. Thieke. Pareto navigation - algorithmic foundation of interactive multi-criteria imrt planning. *Physics in Medicine and Biology*, 53:985–998, 2008.
- 19 V.D. Noghin. Relative importance of criteria: a quantitative approach. *Journal of Multi-Criteria Decision Analysis*, 6(6):355–363, 1997.
- 20 V.D. Noghin. What is the relative importance of criteria and how to use it in MCDM. In M. Koksalan and S. Zionts, editors, *Multiple Criteria Decision Making in the New Millennium, Proceedings*, pages 59–68. Springer, 2001.

- 21 S.S. Ravi, D. . Rosenkrantz, and G.K. Tayi. Facility dispersion problems: Heuristics and special cases. In F. Dehne, J.-R. Sack, and N. Santoro, editors, *Algorithms and Data Structures*, pages 355–366, Berlin, Heidelberg, 1991. Springer.
- 22 S. Sayin. Measuring the quality of discrete representations of efficient sets in multiple objective mathematical programming. *Mathematical Programming*, 87(3):543–560, 2000.
- 23 B. Schulze, L. Paquete, K. Klamroth, and J. Figueira. Bi-dimensional knapsack problems with one soft constraint. *Computers & Operations Research*, 78:15–26, 2017.
- 24 J.I. Serna. Approximating the nondominated set of  $\mathbb{R}^+$  convex bodies. *Technical University of Kaiserslautern, Department of Mathematics*, 2008.
- 25 L. Shao and M. Ehrgott. Discrete representation of non-dominated sets in multi-objective linear programming. *European Journal of Operational Research*, 255:687–698, 2016.
- 26 K. Sindhya, V. Ojalehto, J. Savolainen, H. Niemisto, J. Hakanen, and K. Miettinen. Coupling dynamic simulation and interactive multiobjective optimization for complex problems: An APROS-NIMBUS case study. *Expert Systems with Applications*, 41(5):2546–2558, 2014.
- 27 T.J. Stewart and A. Joubert. Conflicts between conservation goals and land use for exotic forest plantations in South Africa. In E. Beinat and P. Nijkamp, editors, *Multicriteria Analysis for Land-Use Management*, pages 17–31. Springer, 1998.
- 28 D. Vaz, L. Paquete, Fonseca C.M., K. Klamroth, and M. Stiglmayr. Representation of the non-dominated set in biobjective discrete optimization. *Computers & Operations Research*, 63:172–186, 2015.

### 4.3 Usable knowledge extraction in multi-objective optimization: An analytics and “innovization” perspective (WG3)

*Carlos A. Coello Coello, Kerstin Dächert, Kalyanmoy Deb, José Rui Figueira, Abhinav Gaur, Andrzej Jaszkiewicz, Günter Rudolph, Lothar Thiele, and Margaret M. Wiecek*

**License** © Creative Commons BY 3.0 Unported license

© Carlos A. Coello Coello, Kerstin Dächert, Kalyanmoy Deb, José Rui Figueira, Abhinav Gaur, Andrzej Jaszkiewicz, Günter Rudolph, Lothar Thiele, and Margaret M. Wiecek

**Abstract.** Knowledge extraction aims at detecting similarities and patterns hidden in the Pareto-optimal solutions arising from the outcome of a multi-objective optimization problem. The patterns may emerge from generic relationships of several variables or objective functions. Knowledge extraction is expected to bring out valuable information about a problem and is termed as a task of “innovization” elsewhere. While certain automated innovization methods have been proposed, in this report, we attempt to formalize the overall computational task from a machine learning and data analytics point of view. The results can be used to improve modeling and understand interdependencies among different objectives.

#### 4.3.1 Introduction

The topic was proposed by one of the participants (Deb) who has introduced the original idea, has been working on this topic for nearly two decades, and has called it “innovization” (innovation through optimization) [5, 6] of (technical) models, which leads to new designs, hence, true innovations.

The basic innovization idea has been used towards automated innovization methods, for example, in [2, 1, 3, 4]. The concept has been applied in practice, see, for example, [10, 13, 16]. Innovization methods have also been implemented by different other visualization or machine learning methods [14, 17, 18, 15, 7]. Since we do not aim at only reformulating the concept

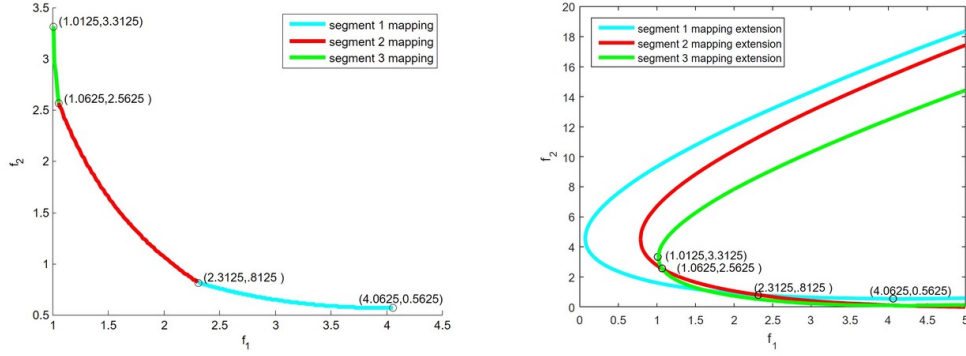
of innovization but also contributing new ideas, we use the more general term “knowledge extraction”. Due to the recent hype in machine learning and data analytics, this topic is of high interest. Moreover, it fits very well to the topic of this Dagstuhl seminar.

### 4.3.2 Problem statement

The main idea of knowledge extraction is as follows. Assume we have already modeled a problem with at least two conflicting objectives, that is, we have formulated a multi-objective optimization problem with certain variables, objective and constraint functions. Some of the variables might be restricted to take only discrete values which turns the problem into a mixed-integer multi-objective optimization problem. Furthermore, motivated from mechanical examples, the model contains certain parameters which are specified by the end users but might change in response to the new knowledge offered by the knowledge extraction procedure.

During the Dagstuhl seminar, we have decided to work on the following topics.

1. General Framework: Given two sets,  $P$  (target set) and  $Q$  (non-target set), from a problem,
  - RQ1:** What features of the problem (described by variables ( $x$ ), objectives ( $f$ ), inequality constraints ( $g$ ) and equality constraints ( $h$ ), or any other basis functions ( $b$ )) are present in  $P$  (but not in  $Q$ )?
  - RQ2:** How to represent features?
  - RQ3:** How to find rules (knowledge) in a computationally efficient way?
  - RQ4:** In what ways can we utilize the “knowledge”?
2. For RQ1: We shall show some examples to clarify the description of “feature”. Features to be considered will be of the type “if condition, then decision”. The outcome for such a feature when applied to a problem vector  $(x, f, g, h, b)$  is **true or false**. We shall consider problems having (i) continuous, (ii) mixed-integer, and (iii) combinatorial variables. We will refer to this as Task 1 in the following.
3. For RQ2: As Task 2, we shall identify subsets of variables defining features and show some examples. The following methods can be used:
  - User-supplied
  - ANOVA, Statistics
  - AIC, Entropy
  - Forward Selection, Backward Selection, and
  - Rough Sets.
4. For RQ3: We shall develop feature (knowledge) extraction procedures (referred to as Task 3) to find hidden features in problem instances and data using the following methods:
  - Genetic Programming (GP) to find general (free-form) features
  - Two-level Decision Tree/Forest Approach to find decision trees, and
  - Other generic or specific methods to find problem-specific structures.
5. For RQ4: We shall utilize the developed features (knowledge) to facilitate the following tasks (we refer to this as Task 4):
  - Knowledge elicitation to users in terms of (i) product platform scaling, (ii) putting focus on key concepts, (iii) using observed knowledge to build theory about knowledge, and (iv) provide leadership.
  - Online utilization of knowledge to improve convergence properties of optimization algorithms.



■ **Figure 9** The Pareto-optimal set (left) and curves determining its three nonlinear segments (right) for the BOQP (21).

- Knowledge accumulation to modify/trust the original problem in establishing the fact that (i) some variables may be redundant, (ii) some objectives may be redundant, and (iii) some constraints may be redundant.

### 4.3.3 Examples

#### Continuous optimization

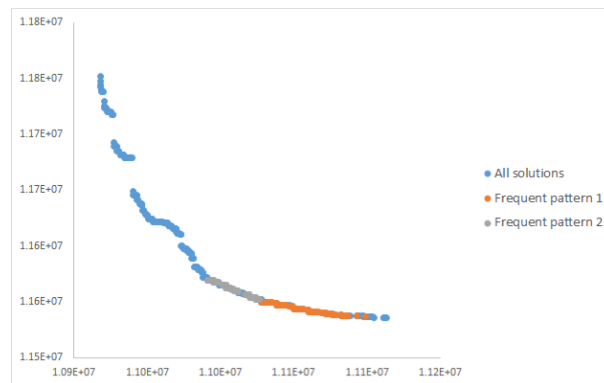
As a simple illustrative example consider the biobjective quadratic problem (BOQP) in which two quadratic objective functions are minimized on a feasible set determined by three linear constraints:

$$\begin{aligned}
 \min_{x_1, x_2} \quad & [f_1(x_1, x_2) = (x_1 - 2)^2 + (x_2 - 2)^2, \quad f_2(x_1, x_2) = x_1^2 + (x_2 - 3)^2] \\
 \text{s.t.} \quad & g_1(x_1, x_2) = x_1 + x_2 - 2.75 \leq 0 \\
 & g_2(x_1, x_2) = 2x_1 + x_2 - 3.75 \leq 0 \\
 & g_3(x_1, x_2) = x_2 \leq 2.25
 \end{aligned} \tag{21}$$

The Pareto-optimal set in the objective space  $(f_1, f_2)$  of this BOQP is shown in the left part of Figure 9. This set is composed of three curves whose equations, due to the simple structure of the objective functions and the feasible set, can be derived analytically. The three curves are depicted in the right part of this figure and have the following equations

$$\begin{aligned}
 \text{segment 1: } f_2 &= f_1 - \sqrt{16f_1 - 1} + 4.5 & \text{for } 2.3125 \leq f_1 \leq 4.0625 \\
 \text{segment 2: } f_2 &= f_1 - \sqrt{18f_1 - 14.0625} & \text{for } 1.0625 \leq f_1 \leq 2.3125 \\
 \text{segment 3: } f_2 &= f_1 - \sqrt{12.8f_1 - 12.96} + 2.3 & \text{for } 1.0125 \leq f_1 \leq 1.0625
 \end{aligned}$$

Recalling that the Pareto-optimal set is the image of the efficient set in the decision space  $(x_1, x_2)$ , we observe that in this particular example each of the three curves is the image of the efficient solutions that are located along (part of) the active constraint  $g_i(x_1, x_2) = 0, i = 1, 2, 3$ . In effect, we obtain the rules for the decision variables  $x_1$  and  $x_2$  in the form  $g_i(x_1, x_2) = 0$  and the answer to the research questions RQ1 and RQ2. Since in this very simple example the rules uniquely determine the Pareto-optimal set, the knowledge extraction is complete and the obtained knowledge is ultimate.



■ **Figure 10** Solutions in the objective space corresponding to two different frequent patterns.

### Combinatorial optimization

As another example we use a bi-objective traveling salesperson problem (TSP). We use real-life data provided by the company Emapa<sup>1</sup>. The data contains travel times and distances between each pair of 500 points located in Poland. The travel times and distances were estimated using data about the real-life road network. The goal is to find a Hamiltonian cycle in this graph taking into account two objectives: total travel time and total distance. The two objectives are obviously highly correlated but also partially in conflict. For example, a route utilizing highways may be longer but faster than a route using secondary roads.

To solve this problem we used a Two Phase Pareto Local Search (TPPSL) algorithm [12]. In the first phase we used the Lin-Kernighan heuristic [11] to generate an initial set of potentially Pareto-optimal solutions that are passed over to the second phase in which Pareto Local Search was run. As a result, 469 potentially Pareto-optimal solutions were found. Each solution is characterized by a set of 500 edges forming a Hamiltonian cycle. The first interesting observation is that these sets are highly similar. There are only 679 distinct edges that appear at least once in this set of solutions out of 124750 possible edges. Furthermore, 245 edges appear in all of the solutions. This set of common edges could be interpreted as a frequent pattern [9] with support (i.e., the number of matching solutions) equal to the number of all potentially Pareto-optimal solutions. In other words, each solution contains only 255 (out of 500) volatile edges, i.e., edges that do not belong to all of the Pareto-optimal solutions.

Furthermore, we can search for other interesting frequent patterns with lower support. For example, Figure 10 shows solutions supporting two different patterns presented in the objective space. The two patterns were selected such that they are supported by at least 100 solutions each, they contain many edges, and the sets of supporting solutions are disjoint. The first pattern is supported by 111 solutions and contains 346 edges. The second pattern is supported by 100 solutions and contains 367 edges. As can be seen from Figure 10, the solutions supporting the two patterns are located in different regions of the objective space. This is an interesting observation since the values of the objectives were not taken into account while selecting the two patterns. The two patterns could be understood as characterizations of two regions of the set of potentially Pareto-optimal solutions.

<sup>1</sup> <http://emapa.pl/>

#### 4.3.4 Usefulness of knowledge extraction

The a posteriori analysis of the results often reveals interesting information on the problem at hand. Unfortunately, this analysis is computationally demanding, in general. Therefore, it can be critically asked what this approach serves for when the optimization process has already been terminated. In the above mentioned combinatorial example, knowledge extraction might be used to reduce the problem size by neglecting edges which never or seldom belong to Pareto optimal solutions. This might have a considerable effect on computational time while basically maintaining the quality of the solutions obtained.

Another important and useful application is in the context of online algorithms. In many applications, the same optimization problem has to be solved over and over again with only slightly different parameter values, e.g., in the context of rolling planning in energy or water networks. In these cases, knowledge extraction might help in solving subsequent optimization problems much quicker and, thus, improving solution quality tremendously when only a very short computational time for optimization is available. Moreover, if time is restricted, a true multi-objective analysis offering different Pareto-optimal solutions is typically not possible. Hence, it is of urgent interest to learn an appropriate setting of ‘multi-objective’ parameters quickly. Also this task can be handled well by knowledge extraction methods.

We shall work on developing methodologies for each of the above topics and plan to write a journal quality paper.

#### 4.3.5 Conclusions

Pareto-optimal or near-Pareto-optimal solutions of multi-objective optimization problems often possess specific properties that can be, for example, seen from certain patterns in the variable values. Since general optimality conditions (like, for example, multi-objective variants of the KKT conditions) are often difficult to apply for practical and complex problems (for example, due to rather restrictive assumptions on the problem structure and due to the need of finding derivatives and a reliable solution of nonlinear equations and inequalities), the “innovization” procedure proposed by Deb [5] is a viable strategy. It is a two-step procedure in which first a set of preferable trade-offs and near-Pareto-optimal solutions are found by an EMO algorithm or a generative MCDM approach. In the second step, the optimized solutions are analyzed to decipher invariant features describing the variables, objectives, and constraint values that exist in the data. The usefulness of the basic innovization approach has been demonstrated by Deb and his collaborators over the past 15 years and certain efforts to automate the second step using machine learning procedures have also been proposed [2, 1, 8].

In this report, we have attempted to formalize a systematic procedure for the innovization task and to extend the basic concept to various computational, theoretical, and application domains.

#### Acknowledgments

The authors acknowledge the Dagstuhl Committee for providing the excellent facility and environment that are conducive for fruitful discussions.


#### References

- 1 S. Bandaru and K. Deb. Automated innovization for simultaneous discovery of multiple rules in bi-objective problems. In *Proceedings of Sixth International Conference on Evolutionary Multi-Criterion Optimization (EMO-2011)*, pages 1–15. Heidelberg: Springer, 2011.

- 2 S. Bandaru and K. Deb. Towards automating the discovery of certain innovative design principles through a clustering based optimization technique. *Engineering optimization*, 43(9):911–941, 2011.
- 3 S. Bandaru, A. H. C. Ng, and K. Deb. Data mining methods for knowledge discovery in multi-objective optimization: Part a – survey. *Expert Systems With Applications*, 70:139–159, 2017.
- 4 S. Bandaru, A. H. C. Ng, and K. Deb. Data mining methods for knowledge discovery in multi-objective optimization: Part b – new developments and applications. *Expert Systems With Applications*, 70:119–138, 2017.
- 5 K. Deb. Unveiling innovative design principles by means of multiple conflicting objectives. *Engineering Optimization*, 35(5):445–470, 2003.
- 6 K. Deb and A. Srinivasan. Innovization: Innovating design principles through optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2006)*, pages 1629–1636, New York: ACM, 2006.
- 7 Catarina Dudas, Amos HC Ng, Leif Pehrsson, and Henrik Boström. Integration of data mining and multi-objective optimisation for decision support in production systems development. *International Journal of Computer Integrated Manufacturing*, (ahead-of-print):1–16, 2013.
- 8 A. Gaur and K. Deb. Effect of size and order of variables in rules for multi-objective repair-based innovization procedure. In *Proceedings of Congress on Evolutionary Computation (CEC-2017) Conference*. Piscataway, NJ: IEEE Press, 2017.
- 9 Jiawei Han, Hong Cheng, Dong Xin, Xifeng Yan. Data Mining and Knowledge Discovery, August 2007, Volume 15, Issue 1, pp 55–86.
- 10 Nozomi Hitomi, Hyunseung Bang, and Daniel Selva. Extracting and applying knowledge with adaptive knowledge-driven optimization to architect an earth observing satellite system. In *AIAA Information Systems-AIAA Infotech@ Aerospace*, page 0794. 2017.
- 11 S. Lin and B.W. Kernighan. An effective heuristic algorithm for the traveling-salesman problem, *Operations Research*, 21, 1973, 498–516.
- 12 Thibaut Lust and Jacques Teghem. Two-phase Pareto local search for the biobjective traveling salesman problem, “*Journal of Heuristics*”, 2010, Jun, 16, 3, 475–510.
- 13 Shigeru Obayashi, Shinkyu Jeong, and Kazuhisa Chiba. Multi-objective design exploration for aerodynamic configurations. *AIAA Paper*, 4666:2005, 2005.
- 14 Shigeru Obayashi and Daisuke Sasaki. Visualization and data mining of pareto solutions using self-organizing map. In *Evolutionary multi-criterion optimization*, pages 796–809. Springer, 2003.
- 15 Akira Oyama, Taku Nonomura, and Kozo Fujii. Data mining of pareto-optimal transonic airfoil shapes using proper orthogonal decomposition. *Journal of Aircraft*, 47(5):1756–1762, 2010.
- 16 Kazuyuki Sugimura, Shigeru Obayashi, and Shinkyu Jeong. Multi-objective design exploration of a centrifugal impeller accompanied with a vaned diffuser. In *ASME/JSME 2007 5th Joint Fluids Engineering Conference*, pages 939–946. American Society of Mechanical Engineers, 2007.
- 17 H Taboada and D Coit. Data mining techniques to facilitate the analysis of the pareto-optimal set for multiple objective problems. In *Proceedings of the 2006 Industrial Engineering Research Conference (CD-ROM)*, 2006.
- 18 Tamara Ulrich, Dimo Brockhoff, and Eckart Zitzler. Pattern identification in pareto-set approximations. In *Proceedings of the 10th annual conference on Genetic and evolutionary computation*, pages 737–744. ACM, 2008.

## 4.4 Complex networks and MCDM (WG4)

*Richard Allmendinger, Michael Emmerich, Georges Fadel, Jussi Hakanen, Johannes Jahn, Boris Naujoks, Robin Purshouse, and Pradyumn Shukla*

License  Creative Commons BY 3.0 Unported license

© Richard Allmendinger, Michael Emmerich, Georges Fadel, Jussi Hakanen, Johannes Jahn, Boris Naujoks, Robin Purshouse, and Pradyumn Shukla

### 4.4.1 Introduction

This report introduces a novel, generic, multi-layered network model of large-scale multi-criteria decision making (MCDM), with a focus on the design and optimization of complex products and platforms. The report provides some examples of network structures in MCDM applications and develops two use-cases for the multi-layered model.

The starting point for this model was to view MCDM from the perspective of complex networks. The study of complex networks (CS) is a topic that recently received considerable attention across various disciplines [2]. Typical questions investigated in CS are:

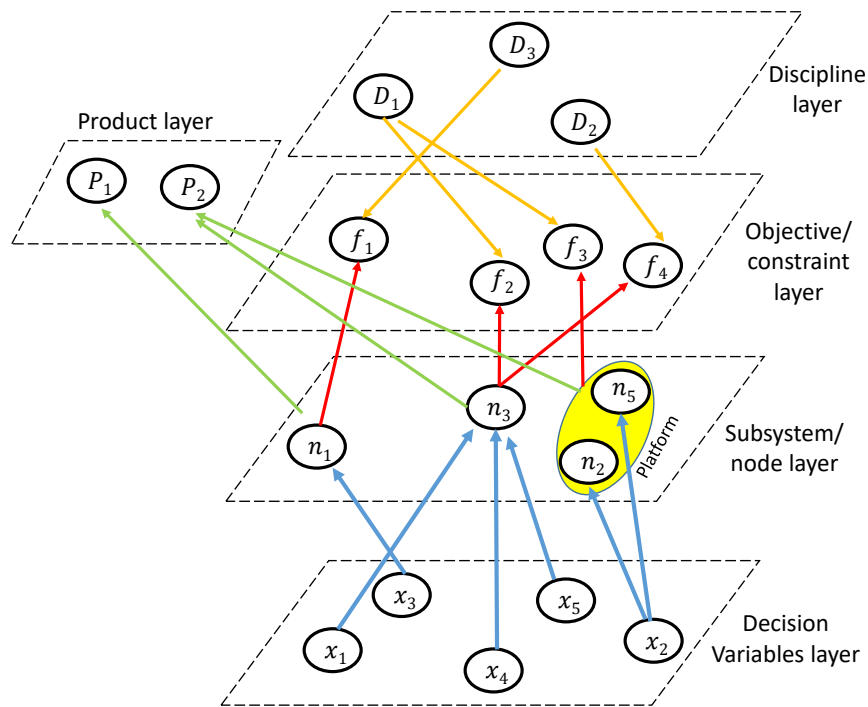
- modeling/formalization and visualization of networks;
- dynamics of networks – both in terms of states and in terms of structure;
- microscopic (e.g., node degrees, node centrality) and macroscopic (e.g., moments of the degree distribution, sparsity, modularity, community structure) properties of networks and how they influence each other;
- algorithms on complex networks.

### 4.4.2 Related work

In several publications, the modeling of design and optimization processes in terms of networks has been addressed. Here we provide only a snapshot of the current state of the art in this domain.

- Martins and Lambe [7] view networks by means of a matrix approach. Their focus is on the coupling between disciplines via shared design variables. The coupling matrix can be exploited by gradient based techniques via the chain rule and leads to efficient methods with sparse matrices. From a networks perspective, the matrices can be interpreted as adjacency matrices and therefore a translation into a network model is possible. However, it can be argued that the approach has a too strong focus on computational models to capture the entirety of a production environment, with aspects such as platforming, discipline specific decision making and multi-objectivity within subdisciplines. This is why we aim for a network model with a broader scope and emphasizing on linkage aspects, albeit less focused on quantitative aspects.
- Maulana et al. [6] introduce network models to model the relationships between objective functions in many objective optimization. Positive links indicate complementary objectives, negatively weighted links indicate conflicting objectives, whereas the non-existence of links signals that objectives can be optimized independently (for instance, because they depend on disjoint variable sets). They propose a method to derive these conflict graphs empirically from a correlation matrix and use the networks to decompose the problem, detect communities of objectives, and the relationship between these communities. Despite its usefulness in structuring many-objective optimization problems, the model by Maulana et al. is limited in scope. It represents only a single layer of the multidisciplinary problem – the objectives layer – and falls short in terms of modeling and integrating relationships between design variables, subsystems, and disciplines (or decision makers).





■ **Figure 11** Layered graph of a multidisciplinary optimization problem.

- Braha et al. [1] aims for models of a design department of a large enterprise representing interactions between designers and engineers as links. The model is not very detailed in terms of node and link semantics, but due to the large size of the data sets some interesting conclusions can be drawn about the general structure of the network, such as, scale-free degree distributions and small world properties.
- Ríos-Zapata et al. [9] consider complex decision networks in the context of traceability within product design processes. The work is preliminary in nature, but is interesting in its use of *traceability trees* to represent the multi-level relationships that connect high-level requirements to detailed design realisations. It is possible to reformulate this approach in MCDM terminology, where *Properties* are criteria, objectives or constraints ( $\mathbf{f}$ ), *Characteristics* are decision variables ( $\mathbf{x}$ ), *External Conditions* are parameters ( $\mathbf{p}$ ) and *Relations* are the models (simulations or expert opinions) that map  $\mathbf{x}$  to  $\mathbf{f}$ . The authors demonstrate the approach on a simple design problem (a portable cooler), highlighting the impact of detailed design choices on the properties of the product.
- Klamroth et al. [5] introduce the concept of *interwoven systems* for multi-objective optimization, in which design, optimization and decision making activities take place within the context of interacting sub-systems. Each sub-system can be viewed as a node within the global problem, with edges that represent shared variables and dependencies. The paper is particularly notable in developing Pareto optimality definitions for such interwoven systems.

#### 4.4.3 Towards a formalisation of complex networks for MCDM

Taking inspiration from the approach of Ríos-Zapata et al. [9], a multiobjective optimization problem in a multidisciplinary product design setting can be viewed as a layered graph, consisting of layers. The layered graph is visualized in Figure 11. Each layer has its own specific type of nodes.

- $L_1$  Elementary Variables: The nodes of this layer are the decision variables. One might also include the environmental variables which cannot be controlled.
- $L_2$  Subsystems: This layer consists of the subsystems of the production process. Subsystems have often their individual modeling and simulation approaches that are then combined in the multidisciplinary optimization.
- $L_3$  Objectives and Constraints: This layer consists of objectives and constraints. Some constraints and objectives are formulated across different subsystems. An example is the total mass of a car, to which different subsystem designs contribute, such as engine and chassis, but others not, such as navigation software.
- $L_4$  Disciplines: Disciplines are concerned with different aspects of the design. For instance, in car design, one might think of aerodynamics, car electronics, product marketing, and engine design. Typically, in a product design process, disciplines are represented by different teams with their own specific responsibilities. They are concerned with specific objectives and constraints. For instance, the aerodynamics of a car might be of concern for the marketing and for the environmental efficiency of a car.
- $L_5$  Products: the products are introduced into the model, in order to model platforming strategies. A platform is a hyperedge of subsystems that can be produced in a combined way and enter in this way into product. As a platform might include more than two nodes, hyperedges (subsets of nodes) are considered as a model.

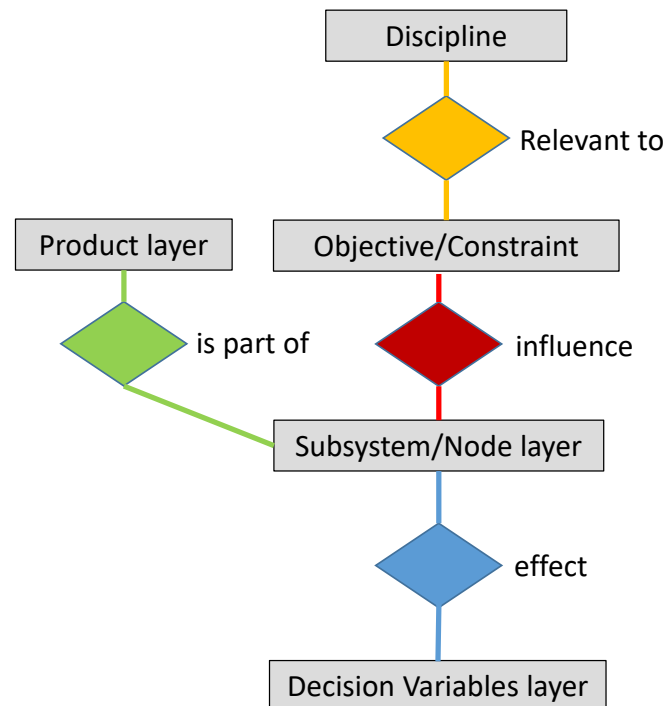
Relationships occur between nodes of different layers. Aiming for not modeling relationships that can be deduced by means of transitive closure, we model only relationships of the following types:

- $E_{12}$  Variable, Subsystem relationships: Subsystems can be viewed as functions that map decision variables to outputs, that are then used to compute objective and constraint function values. Formally,  $E_{12} \subseteq L_1 \times L_2$ .
- $E_{23}$  Subsystems, Objectives relationships: The behavior and properties of subsystems contribute to some of the objectives and constraints. Formally,  $E_{23} \subseteq L_2 \times L_3$ .
- $E_{34}$  Objectives, Disciplines relationships: Disciplines take into consideration certain objectives and constraints, and it is possible that objectives and constraints are shared among multiple disciplines. Formally,  $E_{34} \subseteq L_3 \times L_4$ .
- $E_{25}$  Subsystem/Platform, Products relationships: Products consist of subsets of subsystems, that might be grouped to subsets (platforms). Formally,  $E_{25} \subseteq \wp(L_2) \times L_5$ . Here  $\wp(L_2)$  denotes the set of potential platforms (subsets of subsystems with cardinality bigger than 1) and subsystems, represented by the singletons. Non-overlap in terms of subsystems applies.

All relationships are many-to-many relationships. There is total participation of each node set in the relationship sets. This is visualized in Figure 12. An overview of the components of a complex network for MCDM can also be found in Table 3.

#### 4.4.4 Examples of complex networks

The above framework can be used to represent a range of applications comprising interconnected components, which may be product parts and parameters, decision makers, and/or objectives. Examples of applications that would fit the framework include the design of a



■ **Figure 12** Entity Relationship Diagram.

product, planning and constructing of a facility, and determining the optimal location of facilities (e.g. location of a school). The next two sections look in more detail on how the framework can be mapped to two particular applications: designing a car (Section 4.4.4) and planning and constructing a power station (Section 4.4.4).

### Example I: Product design

Being able to model and facilitate the complex process of designing a sophisticated product was one of the main motivations of this work. The reason that this task is not straightforward is that a complex product, such as a car, consists of a large number of interconnected components, such as an engine, the car body, suspension, electrical supply system, etc. as illustrated in Figure 13. These components are developed by different teams often independently of each other and with more or less conflicting goals in mind.

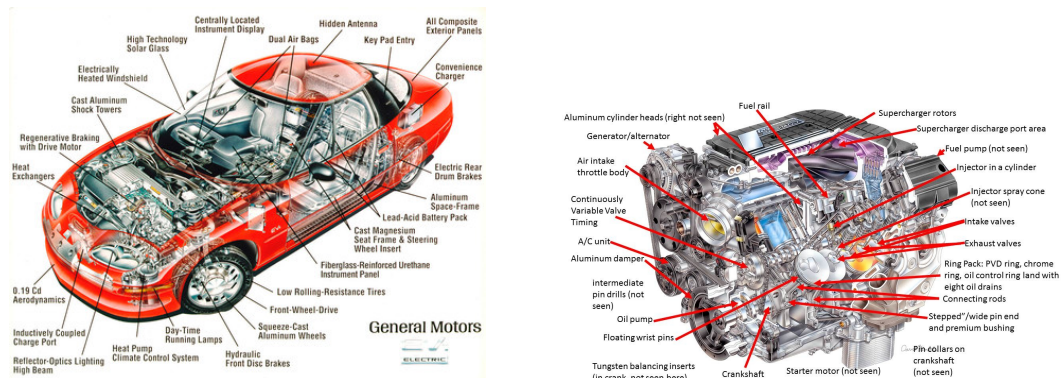
The availability of a structured framework to support the design of a complex product, such as a car, will make the design process more efficient and cheaper as well as provide a tool to visualize to the entire design team the various design components and their relationships. Ultimately, the framework will facilitate decision making in an environment that exists of many decision makers and different (conflicting) design goals (objectives).

What follows is a layer by layer mapping of the framework to the process of designing a car (a less formal mapping is carried out in the next example).

**Decision variables layer:** This layer comprises controllable parameters that have a direct influence on the shape, size and operation of every single component of a product. In the car design example, this may include appearance parameters, such as the dimensions of

Nodes	Edges	Layer
Disciplines or Discipline Decision Makers	Objectives and Constraints relevant to the discipline/decision making	Discipline
Objectives and Constraints	Nodes and Subsystems that influence objectives	Objective/Constraint
Subsystems and Nodes, can be grouped to platforms	Variables that effect the subsystems	Subsystem/Node
Elementary variables	controllable/observable	Decision Variable

■ **Table 3** Overview of components, i.e. nodes, edges, and layers, of a complex network for MCDM.



■ **Figure 13** Car parts (left plot, source: [www.pinterest.co.uk](http://www.pinterest.co.uk)) and engine parts (right plot, source: [www.anatomybody101.org](http://www.anatomybody101.org)).

a component and the location of a component within the overall product, and operation parameters, such as mass, energy, power, temperature, etc required to run a particular component.

**Subsystem / node layer:** The decision variable layer feeds into the subsystem layer in the sense that specifying the setting of each of the decision variables will define the appearance and working of a subsystem or component, such as the engine, battery, suspension, chassis, and car body. Each component has objectives and constraints (e.g. related to the power and noise of engine, and weight of chassis), which need to be accounted for when setting the decision variables. Typically, there is a decision maker for each subsystem (component) aiming to get the component at hand as optimal as possible.

The combination of several components can make up a platform. For instance, in the context of cars, the combination of engine and suspension type may define a platform. The characteristics of a platform, such as size and components involved, are monitored and decided by a decision maker, who is typically different from the ones governing the components involved in a platform.

**Disciplines layer:** This layer sits above the subsystem layer because it addresses multiple components to best satisfy a joint objective and meet certain constraints. Examples of disciplines in car design include the acoustics of a car (noise), structures, dynamics, aerodynamics, and heat transfer. It is common that each discipline has a decision maker associated with it.



■ **Figure 14** Combined cycle power plant at Düsseldorf, Germany (source: [www.siemens.com/press](http://www.siemens.com/press)).

**Product layer:** The product layer links multiple components and platforms to form the overall physical product, e.g. an actual car.

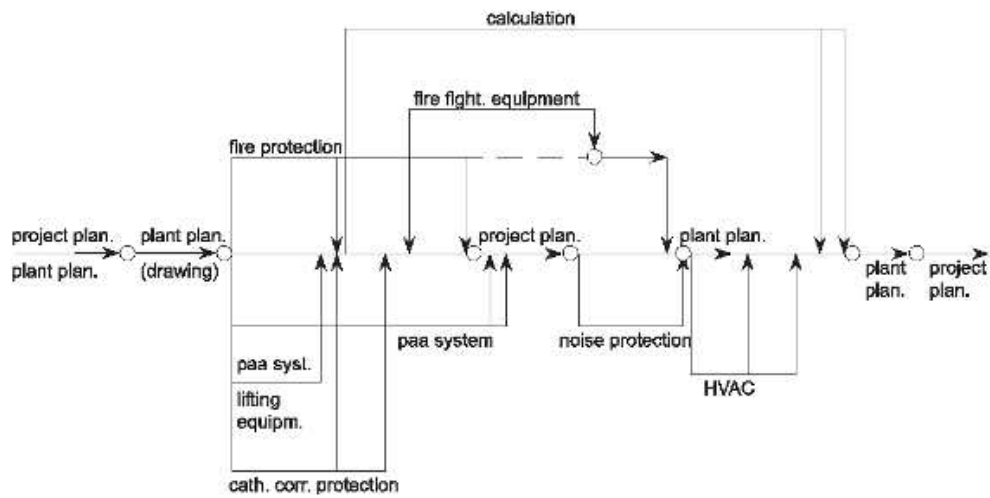
**Objective / constraint layer** All objectives and constraints considered in the subsystem layer and the discipline layer are mapped onto the objective / constraint layer. In general, there are several (conflicting) objectives in that layer including obvious ones, such as costs, but also several other objectives one needs to account for prior to rolling out a product, such as manufacturability, environmental impact, sustainability, product robustness, and customer satisfaction. These objectives are typically posed by the chief engineer.

### Example II: Power station planning and construction

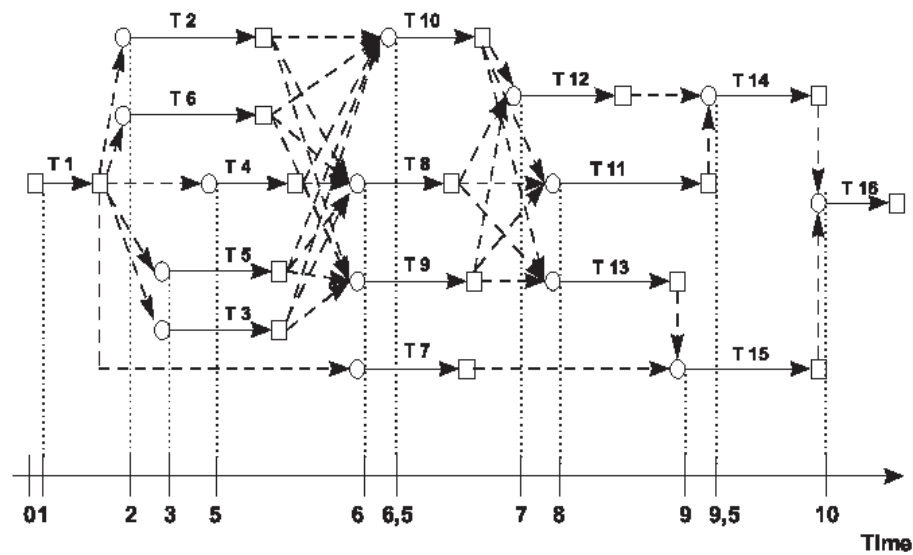
Planning and construction of power stations is a very difficult task, which leads to a very complex network with various complicated subsystems. Each power station is a personalized product (*product layer*) because it is designed to the specific requests of every customer. Various technologies (*subsystem layer*) are possible – the so-called combined heat and power plant is the most cost-efficient way to produce power and heat (e.g., compare Figure 14).

In this subsection we present and discuss results obtained by Hirschmann [3, 4]. There are several stages of the resulting engineering process including first planning, tender compiling, assembling and integration, putting into operation and service. Since this large problem is a discrete-continuous multiobjective optimization problem, the *variables layer* consists of real variables (e.g. duration times of subprocesses), integer variables (e.g. number of staff members) and attributes (e.g. describing the quality of tools). The *objective layer* considers five objectives: Project costs, fixed costs and duration have to be minimized, and flexibility and the effective use of the resources are to be maximized. Figure 15 illustrates the *discipline layer* together with the cooperation of the fields. Besides the classical disciplines such as electrical engineering and mechanical engineering, there are additional disciplines, such as fire and noise protection, among others.

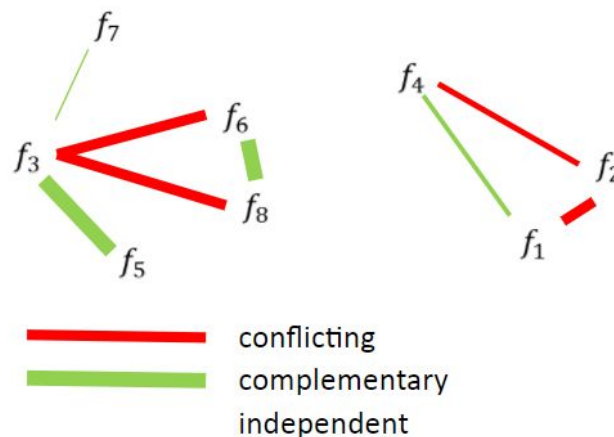
The resulting optimization problem leads to an optimal engineering process illustrated in Figure 16. This optimal process also considers time frames and is given in a simplified form. Several tasks are done in parallel, but there are also common nodes as a result of these subprocesses.



■ **Figure 15** Discipline layer of the construction of a power station.



■ **Figure 16** Optimal engineering process of the construction of a power station.



■ **Figure 17** An example of empirical correlations between objectives.

#### 4.4.5 Use cases for decision making in complex networks

Based on the above framework of complex networks, decision making is happening in different layers having decision makers with roles as described in examples of Section 4.4.4. Most of these decision situations are multi-objective by nature and the objectives in a lower level are typically a subset or a part of the objectives considered in higher levels which makes decision making in this setting complex. Next we will present some possible use cases for supporting multiple criteria decision making in complex networks.

##### Identifying conflict and redundancy

One use case is to use empirical correlations at objective layer to identify relationships between objectives, i.e. conflicting, harmonious and independent objectives [8]. When proceeding downwards, one can identify candidates for platforms that minimize potential conflict(s). In other words, what is the set of subsystems that can be used as a platform common to different products that minimizes potential conflicts. To evaluate this, new metric(s) are needed. On the other hand, when moving upwards, decision hotspots can be identified. That means identifying decision makers / disciplines with conflicts of interest related to the objectives considered requiring communication and negotiation in order to find consensus. Finally, within the objective layer, empirical correlations can be used to find and remove redundant objectives. An example of empirical correlations is shown in Figure 17 where green color denotes positive correlation while red color indicates conflicts.

##### Case-based reasoning for product design programmes

A further potential use case for a complex MCDM network is the ability to identify likely sources and degrees of conflict within product and platform design programmes. If existing product design programme exemplars can be captured using layered graphs, then network statistics can be used to quantify the features of these processes. For existing and past programmes, experiential design expertise is often available on the presence of conflict within the programme. This combined evidence could be used to develop case-based reasoning for new product design programmes, indicating the likely levels of conflict that will be experienced in the design of the product. This intelligence could be used by organisations in resource planning and management for forthcoming design programmes.



#### 4.4.6 Discussion and future research ideas

We originally started our work on complex networks and MCDM collecting research questions that came up thinking about the topic. Here are some of the research questions we discussed:

1. What examples of complex MCDA networks exist?
2. Can we simplify these to tractable examples?
  - What is minimum representation of multi-objective decision problem?
3. How do we represent these using formal languages?
4. How do we incorporate platform design issues?
5. How can we characterise the networks?
6. How do we analyse, design, optimise (on) these networks?
7. How do we introduce platforms in the networks?
8. How do we support decision making/consensus building on the networks?
9. What questions do we want to ask the network:
  - Who/what are the critical components wrt consensus finding?
  - Can we define useful metrics?
 These might be uniqueness, computability, resilience, conflicts (levels and causes) etc.

During our work on the topic, we have been able to find some answers to these questions. For example, Section 4.4.4 provides two examples of complex MCDM networks as an answer to question 1. In addition, our attempt to define use cases can be seen as an answer to question 2, however, not considering the minimisation aspect raised in the subquestion. Finally, question 3 was the starting point for Section 4.4.3 on formalisation. Answering the remaining question is part of future work.


#### References

- 1 Dan Braha and Yaneer Bar-Yam. The statistical mechanics of complex product development: Empirical and analytical results. *Management Science*, 53(7):1127–1145, 2007.
- 2 Stefano Boccaletti, Vito Latora, Yamir Moreno, Martin Chavez, and D-U Hwang. Complex networks: Structure and dynamics. *Physics reports*, 424(4-5):175–308, 2006.
- 3 Claudia Hirschmann. Optimierung von Engineering-Prozessen. *PhD thesis, University of Erlangen-Nürnberg*, 1998.
- 4 Claudia Hirschmann. Optimization of engineering-processes. *Poster*, 1998.
- 5 K. Klamroth, S. Mostaghim, B. Naujoks, S. Poles, R. Purshouse, G. Rudolph, S. Ruzika, S. Sayin, M.M. Wiecek, and X. Yao. Multiobjective optimization for interwoven systems. *Journal of Multi-Criteria Decision Analysis*, 24(1–2):71–81, 2017.
- 6 Asep Maulana, Zhongzhou Jiang, Jing Liu, Thomas Bäck, and Michael TM Emmerich. Reducing complexity in many objective optimization using community detection. In *Evolutionary Computation (CEC), 2015 IEEE Congress on*, pages 3140–3147. IEEE, 2015.
- 7 Joaquim RRA Martins and Andrew B Lambe. Multidisciplinary design optimization: a survey of architectures. *AIAA journal*, 51(9):2049–2075, 2013.
- 8 Robin C. Purshouse and Peter J. Fleming. Conflict, harmony, and independence: Relationships in evolutionary multi-criterion optimisation. *Lecture Notes in Computer Science*, 2632:16–30, 2003.
- 9 David Ríos-Zapata, Jérôme Pailhes, and Ricardo Mejía-Gutiérrez. Multi-layer graph theory utilisation for improving traceability and knowledge management in early design stages. In *Procedia CIRP 60*, pages 308–313, 2017.



## 4.5 Meta-modeling for (interactive) multi-objective optimization (WG5)

*Dimo Brockhoff, Roberto Calandra, Manuel López-Ibáñez, Frank Neumann, Selvakumar Ulaganathan*

License  Creative Commons BY 3.0 Unported license

© Dimo Brockhoff, Roberto Calandra, Manuel López-Ibáñez, Frank Neumann, Selvakumar Ulaganathan

### 4.5.1 Introduction

An important factor in evaluating multi-objective optimization (MOO) algorithms is data efficiency. For many real-world optimization problems, the number of evaluations of the objective functions that can be performed is limited due to cost, time or system constraints. Therefore, it is paramount for future MOO algorithms to be as data efficient as possible.

One approach for improving data efficiency is the use of meta-modeling (also called in different communities Kriging or Bayesian optimization). The underlying idea behind meta-modeling approaches is that explicitly building a model from the data collected during the optimization makes possible to use this data to efficiently reason about the next set of variables to evaluate. Moreover, the use of appropriate probabilistic meta-models makes the optimization more resilient to the stochasticity of the objectives.

In the context of meta-modeling in MOO, there are several open questions that apply also to an interactive context. One fundamental question, and the one that we discuss in this report is: What meta-model should be learned? Or akin: At which level of abstraction should we create the meta-models? In the literature, we can find meta-models at different levels of abstractions: (1) meta-models of the the multiple objective functions by means of a model per function [4, 9], (2) meta-model of the value of a scalarizing function [7, 2] that is defined in terms of some weights that need to be varied at runtime to approximate the whole Pareto front or (3) meta-models that predict some quality metric used by the optimization algorithm, for example, the Pareto ranking of solutions [8]. For detailed citations and discussion of related work, we refer to the recent review by Horn et al. [6]. However, despite these works, it is unclear which choice is preferable under different circumstances. Moreover, among the optimization algorithms that employ the third option, we did not find any work that directly model Pareto compliant quality metrics such as the hypervolume or epsilon metrics [14].

Many recent multiobjective optimizers employ the hypervolume indicator to measure the quality of the current solution set due to its advantageous theoretical properties [14, 11]—among them also some prominent model-based algorithms [4, 9]. Although previous works in the literature have shown that it is possible to compute the expected improvement of the hypervolume contribution directly from the meta-models of the objective functions [5], we did not find any work that has attempted to model directly the hypervolume contribution of a point in the decision space.

In this work, we discuss several advantages of modeling the hypervolume contribution, provide several alternative approaches for doing so, and present preliminary numerical results. Moreover, our idea of directly modeling the hypervolume contribution can be extended to other Pareto-compliant quality metrics [14], such as the quality metric guiding IBEA [12], which is based on the binary  $\epsilon$ -metric.

Another important question is which is the best meta-modeling technique to use for each case. However, we hypothesize that, in the context of (stochastic) MOO, the actual technique used has probably less impact than what is actually modeled. Thus, in this report, we focus on Gaussian processes (GP) [10] and we leave the use of other meta-models for future work.

#### 4.5.2 Bayesian multi-objective optimization (BMO)

Let us assume the following algorithmic framework applied to the classical Bayesian optimization scenario. We have a decision space that is a subset of  $\mathbb{R}^n$ , where  $n$  is the number of decision variables and a vector of  $M$  (expensive) objective functions  $\vec{f}(\vec{x}) = \{f_1(\vec{x}), \dots, f_M(\vec{x})\}$ , where  $f_i: \mathbb{R}^n \rightarrow \mathbb{R}$  ( $1 \leq i \leq M$ ), that are, without loss of generality, to be minimized. The optimization goal is to approximate, as well as possible, the Pareto-optimal set, that is the set of solutions  $X^*$  that are not dominated by any other feasible solution, that is,  $\vec{x}^* \in X^*$  iff  $\nexists \vec{x} \in \mathbb{R}^n$  such that  $\vec{f}(\vec{x}) \preceq \vec{f}(\vec{x}^*) \wedge \vec{f}(\vec{x}) \neq \vec{f}(\vec{x}^*)$ , where  $\preceq$  is the weak Pareto dominance relation.

A possible Bayesian optimization algorithm for solving the above problem is shown in Algorithm 1. In this algorithm, we assume that there is a method for generating an initial set of solutions available. The algorithm evaluates a single point  $\vec{x}_t$  per iteration  $t$  on the true vector of objective functions  $\vec{f}(\vec{x}_t)$ . Then, it builds a surrogate model  $\mathcal{M}$  based on the set of solutions evaluated up to the last iteration  $t$ ,  $X_t = \{\vec{x}_1, \dots, \vec{x}_t\}$  and their true objective function values  $Z_t = \{\vec{f}(\vec{x}_1), \dots, \vec{f}(\vec{x}_t)\}$ . How to build the model or what is the function (or functions) predicted by the model are left unspecified. As mentioned in the introduction, these may be the individual objective functions ( $f_i$ ), the value of some weighted aggregation (scalarization) of the objective functions, or some quality metric applied to the image of the solution set  $X_t$ . The model is then exploited at each iteration to suggest the next single solution  $\vec{x}_{t+1}$  to be evaluated on the true objective functions  $\vec{f}$ . Again, how the model is exploited depends on the particular implementation of this algorithmic model.

---

**Algorithm 1** Template for Bayesian Multiobjective Optimization (BMO)

---

- 1: Initially, a set of  $\mu$  solutions  $X_\mu = \{x_1, \dots, x_\mu\} \in \mathbb{R}^n$  is generated by means of random sampling, Latin Hypercube Design or some other method
  - 2: Compute  $Z_\mu$ , the image of  $X_\mu$  by evaluating the vector of true objective functions  $\vec{f}(x_i) \in \mathbb{R}^M$  for each  $x_i \in X_\mu$
  - 3: Set the iteration counter  $t$  to  $\mu$  (the number of so-far evaluated solutions)
  - 4: **repeat**
  - 5:   Build a model  $\mathcal{M}$  based on  $X_t$  and  $Z_t$
  - 6:   Use  $\mathcal{M}$  to suggest a new point  $\vec{x}_{t+1}$  based on an acquisition function (e.g., expected improvement)
  - 7:   Evaluate the true  $\vec{f}(\vec{x}_{t+1})$  and set  $X_{t+1} = X_t \cup x_{t+1}$  and  $Z_{t+1} = Z_t \cup \vec{f}(\vec{x}_{t+1})$
  - 8: **until** happy or running out of time
- 

#### 4.5.3 A surrogate model for the HV contribution

The goal of several highly effective multi-objective optimization algorithms is to maximize the hypervolume of the set of solutions found. The hypervolume of a solution set  $X = \{\vec{x}_1, \dots, \vec{x}_t\}$  ( $\vec{x}_i \in \mathbb{R}^n$ ,  $\forall i = 1, \dots, t$ ), given a reference point  $\vec{r} \in \mathbb{R}^M$  is the hypervolume of the objective space dominated by the solution set  $X$  and bounded above by the reference point:

$$HV(X) = \int \mathbf{1}_{\{\vec{z} \in \mathbb{R}^M \mid \exists \vec{x} \in X: \vec{f}(\vec{x}) \preceq \vec{z} \preceq \vec{r}\}}(\vec{z}) d\vec{z} \quad (22)$$

where  $\preceq$  is the weak Pareto dominance relation and  $\mathbf{1}_A(a)$  the indicator function, giving one if and only if  $a \in A$ . The Pareto-optimal set has the largest hypervolume of all feasible sets.

A way to guide the selection (or removal) of solutions during optimization is to select (or discard) solutions with the highest (resp. lowest) hypervolume contribution to the current

solution set, where the hypervolume contribution ( $HVC$ ) of a solution  $\vec{x}$  to a solution set  $X$  is the increment in hypervolume after the addition of  $\vec{x}$  to  $X$ , that is,  $HVC(\vec{x}, X) = HV(X \cup \vec{x}) - HV(X)$ . If  $\vec{x}$  is dominated by any solution in  $X$ , then  $HVC(\vec{x}, X) = 0$ .

In the context of Bayesian optimization, a previous work [5] has shown that it is possible to model the true objective functions and compute the hypervolume contribution of a solution directly from this model. Our proposal here is to model directly the hypervolume contribution (or some function related to this contribution) of a point in decision space, relative to the archive  $X_t$  of already evaluated solutions (and their image  $Z_t$ ), without building any model of the actual objective functions. One motivation for modeling directly the hypervolume contribution is that we would model a single “function” instead of  $M$  objective functions. Another motivation is that we conjecture that the landscape of the hypervolume contribution is likely to be more regular and easier to navigate and model than the combined landscape of the true objective functions.

To motivate this conjecture, we show in Fig. 18, for the simple problem of optimizing two Sphere functions with two decision variables, the hypervolume contribution of each point of the decision space with respect to a solution set of five solutions (marked with  $\times$ ). The center (hence, optimal) solution of each Sphere function is marked with a red and blue point, respectively. As shown by the plot, the hypervolume contribution is a multi-modal function but looks globally well-behaved with locally quadratic shapes and in the specific case of the five given solutions, a single global optimum with a large basin of attraction. Although this is not enough to prove our conjecture, specially when we move to higher dimensions and more complex problems, it does show that the landscape of the hypervolume contribution is not necessarily more complex than the combined landscape of the actual objective functions being optimized.

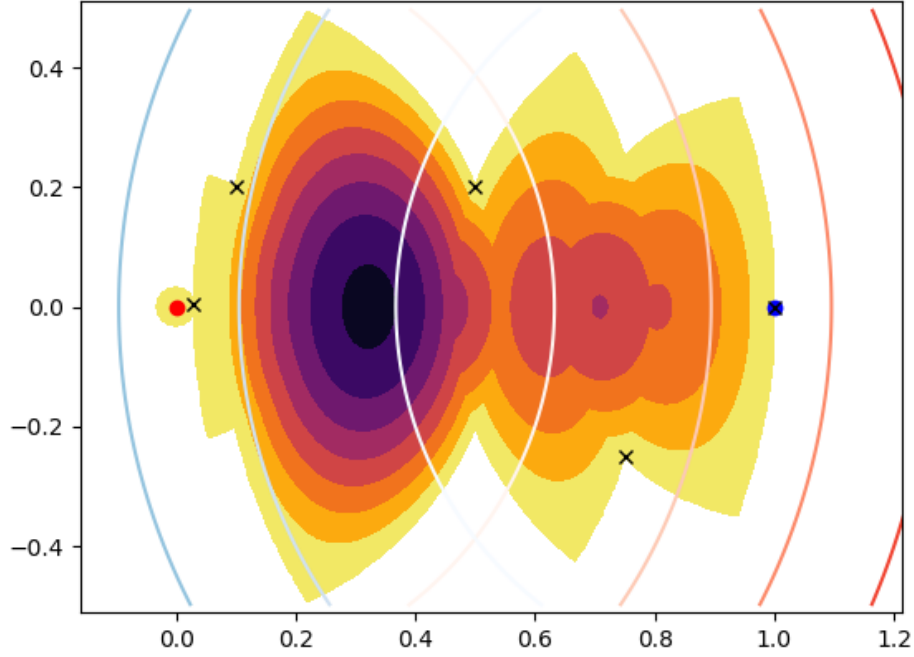
In order to build a model that predicts the hypervolume contribution of each solution in the decision space, we need to find out a way to build such a model using the information contained in our current solution set  $X_t$  and its image  $Z_t$ . We cannot simply use the hypervolume contribution of each point in  $X_t$  with respect to itself, since all solutions would have zero value. We discuss several possibilities in the next subsections.

### Method 1: Use information only from dominated solutions

A first approach is to keep the assumption that the hypervolume contribution of each point nondominated with respect to the current set  $X_t$  is zero, but assign a negative value to those points from  $X_t$  that are dominated. We have devised up to three different ways of doing the latter, which are illustrated in Fig. 19:

- (a) A first variant assigns  $HV(\{\vec{x}\}) - HV(ND_t)$  to each dominated point  $\vec{x} \in X_t$  where  $ND_t$  is the set of non-dominated points in  $X_t$ . The main advantage of this method is its simplicity. However, this variant is not smooth around zero when  $\vec{x}$  gets closer to the non-dominated set.
- (b) A second variant assigns  $HVC_{(-)}(\vec{x}, ND_t)$  to each dominated point  $\vec{x}$ , where  $HVC_{(-)}$  denotes the contribution of  $\vec{x}$  over the non-dominated set  $ND_t$  if we maximize instead of minimizing the objective functions and using the ideal of  $ND_t$  as the reference point for computing the hypervolume. We call this function, *negative hypervolume contribution*.
- (c) Another possibility is to use a distance metric from  $\vec{x}$  to  $ND_t$ .
- (d) Another simple strategy (not shown in Fig. 19) is to assign the negative dominance rank (from non-dominated sorting) to each dominated point.

We expect that the metrics above would be able to approximate the hypervolume contribution of solutions dominating the current solution set by exploiting the inherent



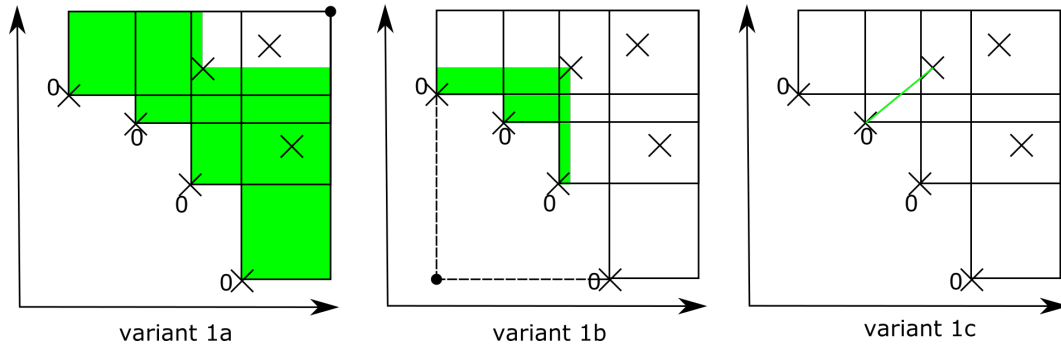
■ **Figure 18** Value of the hypervolume contribution of each point of the decision space of  $\mathbb{R}^2$  with respect to the solution set denoted by the symbols  $\times$ , when minimizing two Sphere functions whose optimal solutions in  $(0,0)$  and  $(1,0)$  respectively correspond to the blue and red points. Darker colors indicate larger values of the hypervolume contribution.

symmetries of meta-models such as Gaussian processes. However, the fact that no distinction is made for the nondominated solutions in our solution set may hinder the prediction power of such a model (and waste useful information). Thus, we propose next a way to assign a value to such nondominated solutions.

### Method 2: Use information from all solutions evaluated

The idea underlying our second proposed approach is to distinguish between points in the non-dominated set  $ND_t$  by assigning different values to each of them (instead of zero like in our first method above) in order to give even more information to the model. In particular, given a nondominated solution  $\vec{x} \in ND_t$ , we assign it its actual hypervolume contribution to the set  $X_t$  as  $HV(X_t) - HV(X_t \setminus \vec{x})$ . This should result in a model with higher values around solutions that are isolated in the objective space with the goal to force the Bayesian optimizer to suggest new solutions that are more likely to dominate a larger part of the objective space.

In the case of dominated points, we can use any of the variants discussed for method 1 above (see Fig. 19), leading to variants 2a, 2b, 2c, and 2d. Additionally, we could simply assign a value of zero for such points and only use the information provided by the nondominated solutions.



■ **Figure 19** Three ways of assigning an *HVC*-related value to the leftmost of three dominated solutions.

### Preliminary experiments

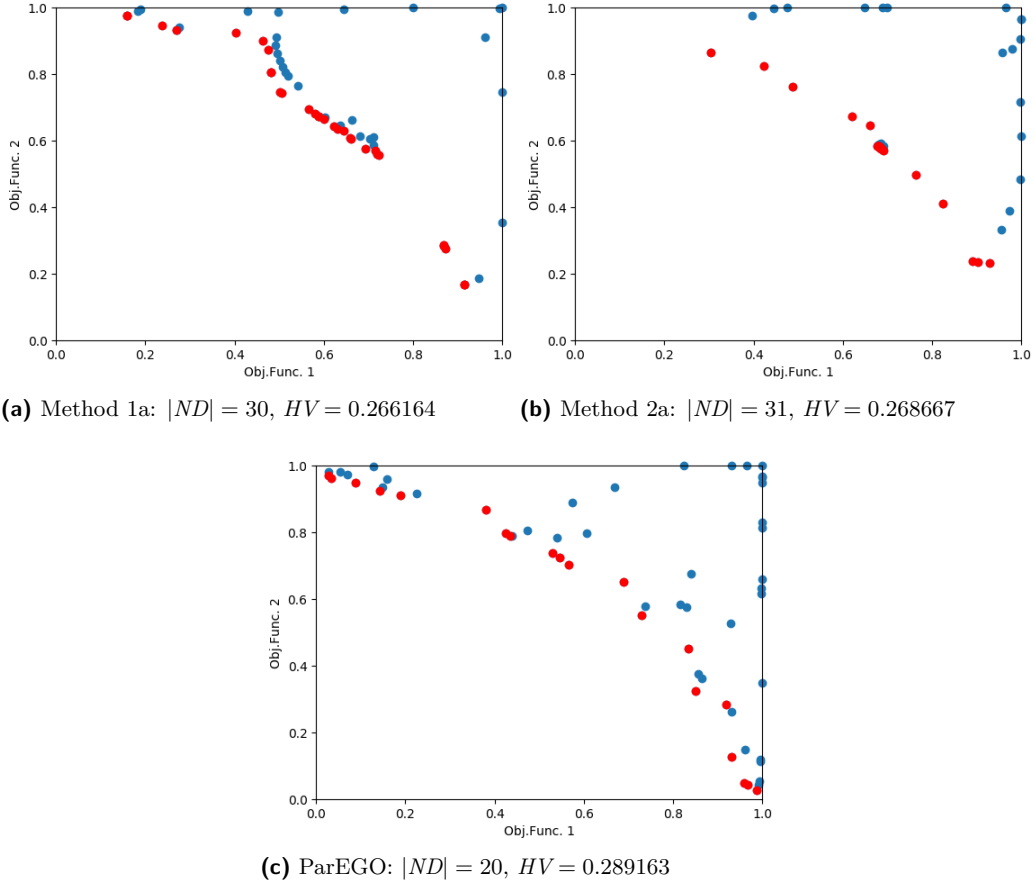
We carried out a few preliminary experiments to see whether the proposals above are able to guide optimization. In particular we analyze methods 1a and 2a. In method 1a, each dominated point  $\vec{x} \in X_t \setminus ND_t$  is assigned the value  $HV(\{\vec{x}\}) - HV(ND_t)$   $\vec{x} \in X_t$ , where  $ND_t$  is the set of non-dominated points in  $X_t$ , while points in  $ND_t$  have a value of zero. In method 2a, dominated points have the same value as in method 1a, but each nondominated solution  $\vec{x} \in ND_t$  is assigned its actual hypervolume contribution to the set  $X_t$  as  $HV(X_t) - HV(X_t \setminus \vec{x})$ .

We prototyped and integrated these two methods in the Algorithm 1 using the Opto framework [1]. Subsequently, we executed the algorithms for a maximum of 60 evaluations of the true objective function vector. We compare the results with the well-known ParEGO [7]. Figure 20 show the objective vectors of the final solution set produced by each approach on the bi-objective Double Sphere problem with dimension  $n = 2$ . Nondominated solutions are shown in red, while dominated solutions are shown in blue. The caption below each plot indicates the size and the hypervolume of the nondominated set produced by each approach. Although ParEGO produces the best results, it is encouraging that the first two runs of our proposed approaches produce reasonable results. In particular, Method 2a produces slightly better hypervolume but seems to have trouble generating solutions in the extremes of the Pareto frontier and it generates solutions that are too close to each other.

When looking at the solution space (Fig. 21), we can clearly see that the solutions produced by ParEGO are well-distributed along the Pareto set (green line), whereas the solutions produced by methods 1a and 2a are clustered in a smaller region. This suggests that the meta-model predicting the hypervolume contribution is not able to find extreme solutions and keeps predicting a high hypervolume contribution in that small region.

We also apply the three approaches to the more challenging ZDT1 problem [13] and results are shown in Fig. 22. To our surprise, our two methods are able to obtain slightly higher hypervolume values than ParEGO, although only method 2a shows an even distribution of solutions along the Pareto frontier, whereas method 1a produces solutions clustered in two small regions.

Nevertheless, a single run on each of two problems only provides some support to our initial conjecture that it is possible to guide optimization by directly modeling the hypervolume contribution without modeling the actual objective functions. However, a proper experimental analysis would be necessary to reach any definitive conclusions.



■ **Figure 20** Solutions, show in the objective space, produced by each approach when optimizing the Double Sphere problem after a maximum of 60 solution evaluations. Red dots indicate nondominated solutions, while blue dots are dominated ones.

#### 4.5.4 A surrogate model based on binary $\epsilon$ -metric

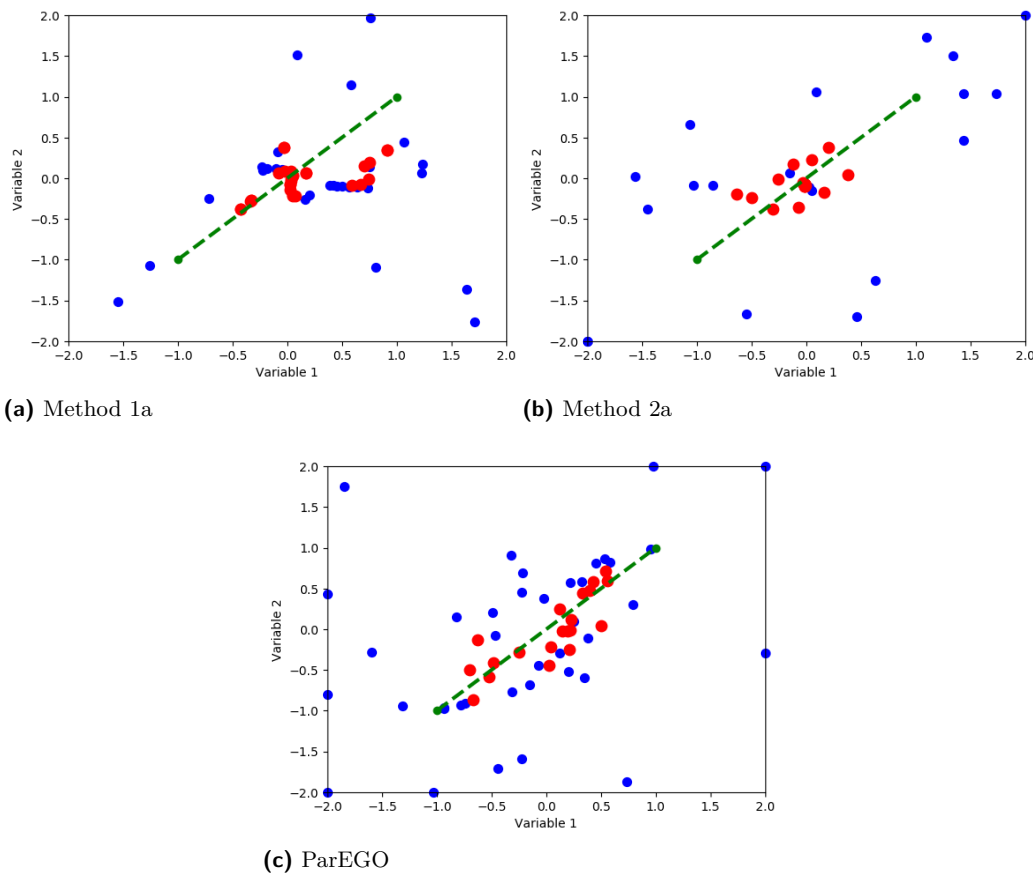
As shown above, trying to directly predict the hypervolume contribution requires the definition of alternative, but related metrics to assign a value to each point of our solution set  $X_t$ , since the actual  $HVC$  value of those points would be zero. Instead of considering the hypervolume contribution, a different approach to multi-objective model-based optimization was discussed in our working group: The direct usage of the fitness function in IBEA [12] as the objective function. This fitness function is defined as

$$F(\vec{x}_1) = \sum_{\vec{x}_2 \in X_t \setminus \{\vec{x}_1\}} -e^{-I(\vec{x}_2, \vec{x}_1)/\kappa} \quad (23)$$

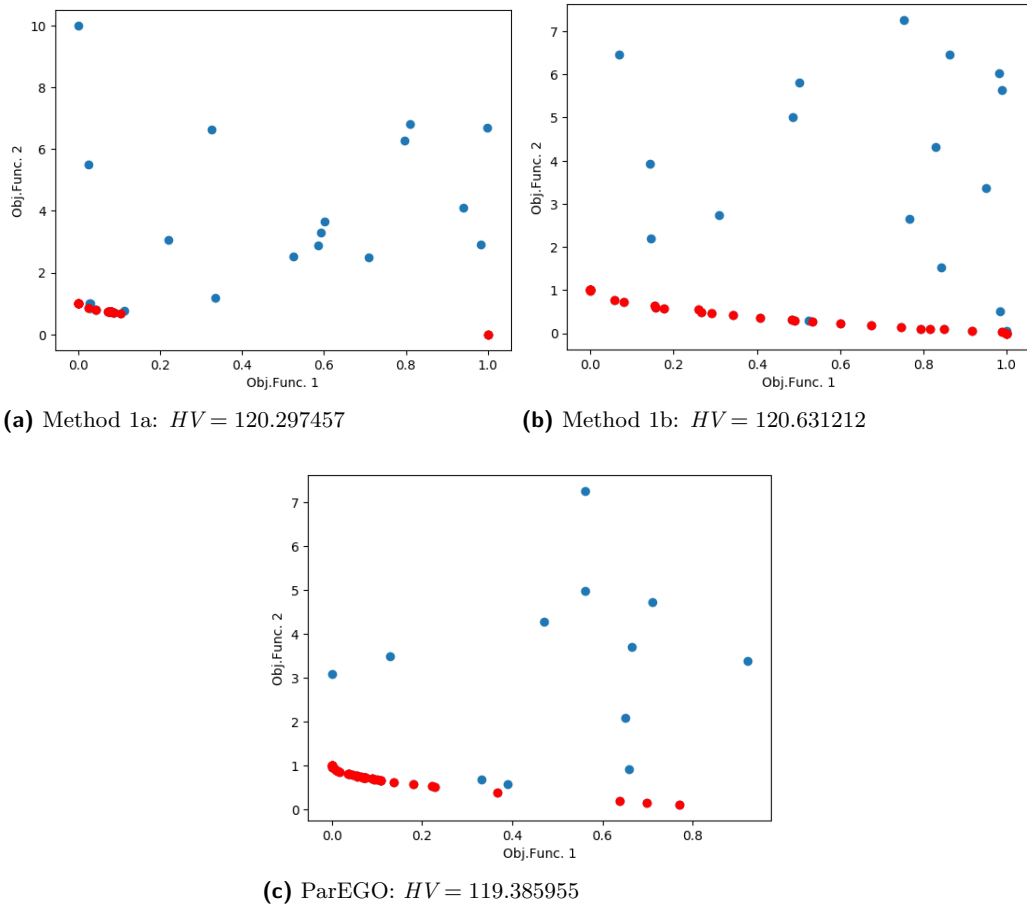
where  $\kappa$  is a normalization parameter and the metric  $I()$  above may be, for example, the additive binary  $\epsilon$ -metric:

$$I_{\epsilon+}(\vec{x}_2, \vec{x}_1) = \max_{i=1, \dots, M} f_i(\vec{x}_2) - f_i(\vec{x}_1) \quad (24)$$

The benefit of the above fitness metric is that it naturally assigns a value to every point in our solution set, and those values will usually be different, except for specific solution

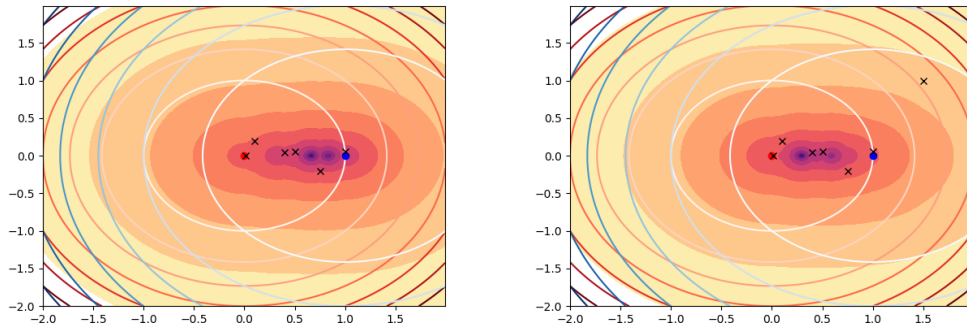


**Figure 21** Solutions, shown in the decision space, produced by each approach when optimizing the Double Sphere problem after a maximum of 60 solution evaluations. Red dots indicate nondominated solutions, while blue dots are dominated ones. The green dashed line corresponds to the optimal Pareto set.



■ **Figure 22** Solutions, show in the objective space, produced by each approach for the ZDT1 problem after 60 solution evaluations. Red dots indicate nondominated solutions, while blue dots are dominated ones.





■ **Figure 23** Value of IBEA's fitness function  $F()$  (using additive binary  $\epsilon$ -indicator) at each point of the decision space of  $\mathbb{R}^2$  with respect to the solution set denoted by the symbols  $\times$ , when minimizing two Sphere functions, whose optimal solutions in  $(0, 0)$  and  $(1, 0)$  respectively correspond to the blue and red points. Darker colors indicate larger values of  $F()$ . The right plot shows the landscape of  $F()$  after adding one (dominated) point to the solution set of the left plot.

sets that are unlikely to arise in real-world problems. In addition, IBEA has been shown to perform consistently well (when properly tuned for the scenario at hand) in a large number of scenarios, often outperforming more recent and popular multi-objective evolutionary algorithms [3]. Thus, this fitness function is likely to produce a similarly well-performing Bayesian optimizer.

A quick numerical experiment, however, showed that the IBEA fitness function has the disadvantage of fundamentally changing its landscape after adding dominated points to the solution set into the history. Figure 23 shows the landscape of function  $F(\vec{x})$  (Eq. 23) on a bi-dimensional decision space when optimizing two Sphere functions, with darker colors corresponding to higher values of  $F()$ . The optimal solutions of each Sphere function are shown as a red and a blue point, respectively, and contour lines denote the function value of each Sphere function. The current solution set  $X_t$  is denoted by  $\times$ . The left plot shows the landscape of  $F()$  with respect to five solutions in  $X_t$ . The right plot shows the landscape after adding an additional (dominated) solution to  $X_t$  at the top right. The difference in colors between the two plots shows that the landscape of the fitness function  $F()$  has changed after adding this point, in particular, the peaks of the function have shifted towards the red point.

#### 4.5.5 Conclusions

Many optimization problems have objective functions that are expensive to evaluate. In this case, meta-modeling allows predicting where to look for next solutions to be evaluated. The insights of newly evaluated solutions are then taken into account to update or refine the model for the problem at hand. Existing approaches either model the individual objective functions or a weighted aggregation thereof, however, we are not aware of any attempts at modeling directly the quality metrics that guide several multi-objective evolutionary algorithms.

In this report, we have discussed several ways to directly model the hypervolume contribution. Preliminary results on two problems suggest that this approach can guide a Bayesian multi-objective optimizer based on Gaussian Processes (GP), however, we also identified that the solutions generated have a low diversity and appear clustered in small regions of the decision and objective spaces. In addition, we also proposed how to model the fitness

function of IBEA, which is based on the binary  $\epsilon$ -indicator. This  $\epsilon$ -based fitness seems, in principle, easier to model directly than the hypervolume contribution, being able to directly provide a value for every point evaluated by the algorithm.

Further work is necessary to determine the advantages and disadvantages of the variants proposed here and empirically analyze their performance on multiple problems.

## References

- 1 Opto – a python framework for optimization. <https://github.com/robertocalandra/opto>
- 2 Aghamohammadi, N.R., Salomon, S., Yan, Y., Purshouse, R.C.: On the effect of scalarising norm choice in a ParEGO implementation. In: Trautmann, H., Rudolph, G., Klamroth, K., Schütze, O., Wiecek, M.M., Jin, Y., Grimme, C. (eds.) *Evolutionary Multi-criterion Optimization, EMO 2017*, pp. 1–15. *Lecture Notes in Computer Science*, Springer, Cham (2017)
- 3 Bezerra, L.C.T., López-Ibáñez, M., Stützle, T.: A large-scale experimental evaluation of high-performing multi- and many-objective evolutionary algorithms. *Evolutionary Computation* (2017), to appear
- 4 Emmerich, M.: Single-and multi-objective evolutionary design optimization assisted by gaussian random field metamodells. Dissertation, LS11, FB Informatik, Universität Dortmund, Germany (2005)
- 5 Emmerich, M.T.M., Yang, K., Deutz, A., Wang, H., Fonseca, C.M.: A multicriteria generalization of bayesian global optimization. In: Pardalos, P.M., Zhigljavsky, A., Žilinskas, J. (eds.) *Advances in Stochastic and Deterministic Global Optimization*, pp. 229–242. Springer International Publishing, Cham (2016)
- 6 Horn, D., Wagner, T., Biermann, D., Weihs, C., Bischl, B.: Model-based multi-objective optimization: taxonomy, multi-point proposal, toolbox and benchmark. In: *Evolutionary Multi-Criterion Optimization (EMO 2015)*. pp. 64–78. Springer (2015)
- 7 Knowles, J.D.: ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Transactions on Evolutionary Computation* 10(1), 50–66 (2006)
- 8 Loshchilov, I., Schoenauer, M., Sebag, M.: A mono surrogate for multiobjective optimization. In: *Genetic and Evolutionary Computation Conference (GECCO 2010)*. pp. 471–478. ACM (2010)
- 9 Ponweiser, W., Wagner, T., Biermann, D., Vincze, M.: Multiobjective optimization on a limited budget of evaluations using model-assisted  $\mathcal{S}$ -metric selection. In: *International Conference on Parallel Problem Solving from Nature*. pp. 784–794. Springer (2008)
- 10 Rasmussen, C.E., Williams, C.K.I.: *Gaussian Processes for Machine Learning*. The MIT Press (2006)
- 11 Zitzler, E., Brockhoff, D., Thiele, L.: The hypervolume indicator revisited: On the design of Pareto-compliant indicators via weighted integration. In: Obayashi, S., et al. (eds.) *Evolutionary Multi-criterion Optimization, EMO 2007, Lecture Notes in Computer Science*, vol. 4403, pp. 862–876. Springer, Heidelberg, Germany (2007)
- 12 Zitzler, E., Künzli, S.: Indicator-based selection in multiobjective search. In: Yao, X., et al. (eds.) *Proceedings of PPSN-VIII, Eighth International Conference on Parallel Problem Solving from Nature, Lecture Notes in Computer Science*, vol. 3242, pp. 832–842. Springer, Heidelberg, Germany (2004)
- 13 Zitzler, E., Thiele, L., Deb, K.: Comparison of multiobjective evolutionary algorithms: Empirical results. *Evolutionary Computation* 8(2), 173–195 (2000)
- 14 Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., Grunert da Fonseca, V.: Performance assessment of multiobjective optimizers: an analysis and review. *IEEE Transactions on Evolutionary Computation* 7(2), 117–132 (2003)

## 5 Topics of Interest for Participants for the Next Dagstuhl Seminar

During the summary session on Friday, all participants had an extensive discussion on the future challenges related to EMO and MCDM. This has led to a plethora of suggestions for future seminar topics continuing the series. Photographs of topics of interest for participants for the next Dagstuhl seminar on EMO & MCDM are shown in Figure 24. The suggestions will be used by the organizers towards the proposal for a continuation of the series.

## 6 Changes in the Seminar Organization Body

### Joshua Knowles steps down as co-organizer

On behalf of all the participants of the seminar, KK, GR and MW would like to extend our warm thank you to Joshua Knowles for his contributions to this Dagstuhl seminar series on Multiobjective Optimization as he steps down from the role of co-organizer, which he has held for three terms of office. To our large regret, Joshua could not be in Dagstuhl during the seminar week. Nevertheless, he has played a leading role in shaping the topic, sharpening the research questions and setting us all up on an exciting journey to personalization. We are very thankful for his advice and activities in the preparation of this and the previous seminars. Thank you, Joshua!

### Welcome to Carlos Fonseca

We are very pleased that our esteemed colleague Carlos Fonseca has agreed to serve as co-organizer for future editions of this Dagstuhl seminar series on Multiobjective Optimization.

## 7 Seminar Schedule

### Monday, January 15, 2018

#### 08:45 – 10:30: Welcome Session

- Welcome and Introduction
- Short presentation of all participants (2 minutes each!)
- Introduction to the topic of the seminar

#### Coffee Break

#### 11:00 – 12:00: Application Challenges

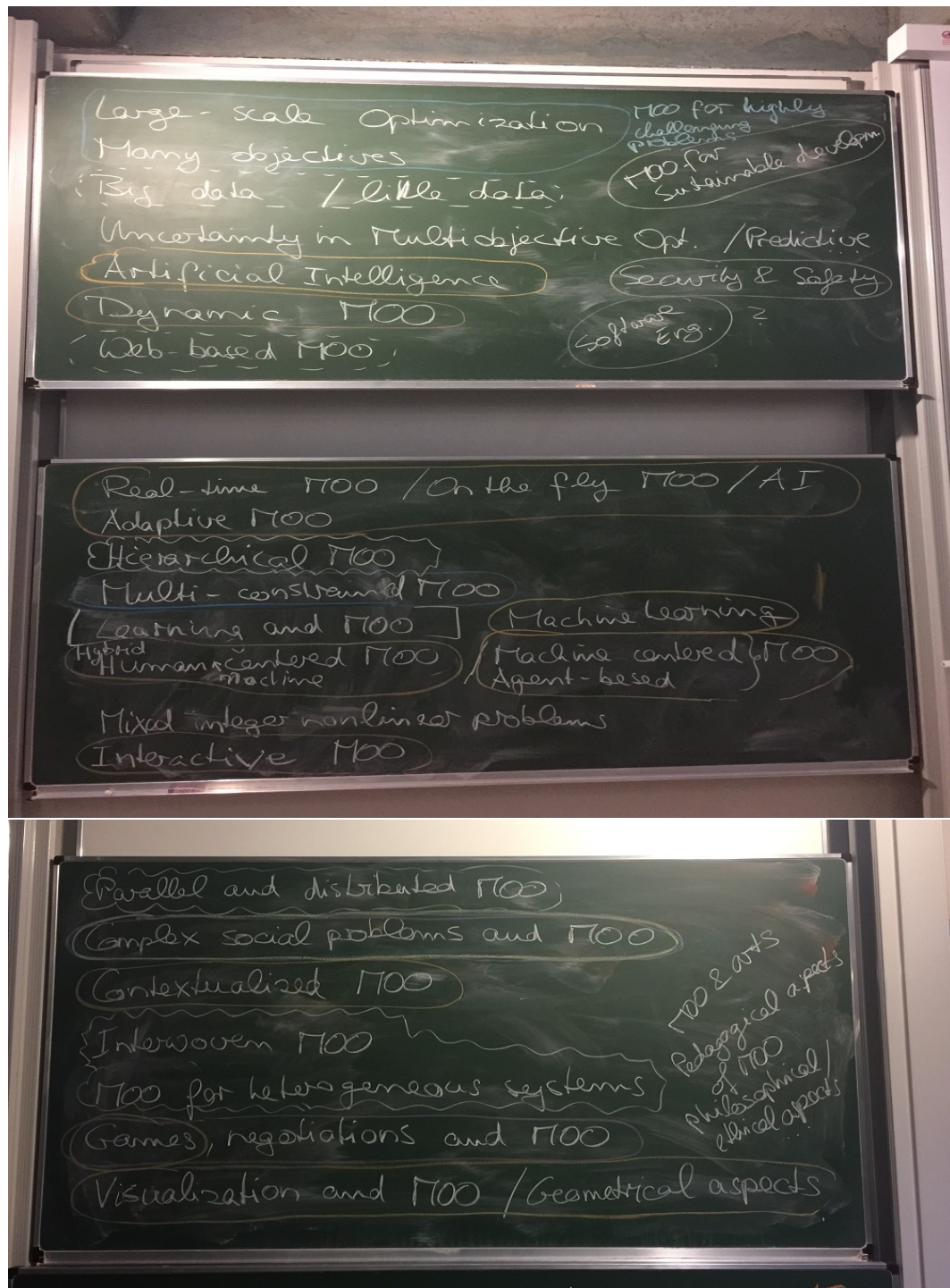
- Karl Heinz Küfer: Industrial Applications of Multicriteria Decision Support Systems
- Georges Fadel: Culturally Tailored Multicriteria Product Design using Crowdsourcing

#### Lunch

#### 13:30 – 14:30: Personalization in Model Building, Approximation, and Representation

- Kalyanmoy Deb: Metamodeling Approaches for Multiobjective Optimization
- Serpil Sayin: Representations: Do they have Potential for Customer Choice?

#### Coffee Break



■ **Figure 24** Topics of interest for participants for the next Dagstuhl seminar.

**15:00 – 15:30: Personalization and Preference Modelling**

- Robin Purshouse: Modelling Complex Networks of Decision Makers: An Analytical Sociology Perspective

**15:30 – 16:00: Personalization in Algorithm Design and Efficiency**

- Manuel López-Ibáñez: Data-Driven Automatic Design of Multi-Objective Optimizers

**Break****16:15 – 18:00: Group Discussion about Hot Topics and Working Groups****Tuesday, January 16, 2018****09:00 – 10:00: Decision Analytics and Consensus** Chair: Salvatore Greco

- Michael Emmerich: Maximizing the Probability of Consensus in Group Decision Making
- Kaisa Miettinen: Decision Analytics with Multiobjective Optimization and a Case in Inventory Management

**Coffee Break****10:30 – 12:00: Working Groups****Lunch****13:30 – 14:30: Personalization and Learning** Chair: Jussi Hakanen

- Jürgen Branke: Active Learning for Mapping Advertisements to Customers
- Roman Slowinski: The NEMO framework for EMO: Learning value functions from pairwise comparisons

**Coffee Break****15:00 – 17:00: Working Groups****17:00 – 18:00: Reports from Working Groups**

- 6 minutes / 3 slides per working group
- General discussion and working group adaptations

**Wednesday, January 17, 2018****09:00 – 10:00: Metamodelling and Knowledge Extraction** Chair: Carlos Fonseca

- Mickaël Binois: Uncertainty Quantification on Pareto Fronts
- Abhinav Gaur: Unveiling Invariant Rules from Non-Dominated Solutions for Knowledge Discovery and Faster Convergence

**10:00: Announcements****Coffee Break****10:30 – 12:00: Working Groups****Lunch****14:00: Group Foto (Outside)****14:05 – 16:00: Hiking Trip**

**16:30 – 18:00: Reports from Working Groups**

- 15 minutes / 5 slides per working group

**Thursday, January 18, 2018****9:00 – 10:00: Data Structures** Chair: Christoph Lofi

- José Rui Figueira: Compressed Data Structures for Bi-Objective  $\{0,1\}$ -Knapsack Problems
- Andrzej Jaskiewicz: Recent Algorithmic Progress in Multiobjective (Combinatorial) Optimization

**Coffee Break****10:30 – 12:00: Working Groups****Lunch****13:30 – 15:30: Working Groups****Coffee Break****16:00 – 17:00: Working Groups****17:00 – 18:00: Continuing the Dagstuhl Seminar Series****20:00: Wine & Cheese Party** (Music Room)**Friday, January 19, 2018****9:00 – 11:00: Presentation of Working Group Results****Coffee Break****11:30 – 12:00: Summary, Feedback, and Next Steps****Lunch & Goodbye**



## Participants

- Richard Allmendinger  
University of Manchester, GB
- Mickaël Binois  
Argonne National Laboratory –  
Lemont, US
- Jürgen Branke  
University of Warwick, GB
- Dimo Brockhoff  
INRIA Saclay – Palaiseau, FR
- Roberto Calandra  
University of California –  
Berkeley, US
- Carlos A. Coello Coello  
CINVESTAV – Mexico, MX
- Kerstin Dächert  
Universität Wuppertal, DE
- Kalyanmoy Deb  
Michigan State University, US
- Matthias Ehrgott  
Lancaster University  
Management School, GB
- Gabriele Eichfelder  
TU Ilmenau, DE
- Michael Emmerich  
Leiden University, NL
- Alexander Engau  
Lancaster University  
Management School, GB
- Georges Fadel  
Clemson University –  
Clemson, US
- José Rui Figueira  
IST – Lisbon, PT
- Carlos M. Fonseca  
University of Coimbra, PT
- Abhinav Gaur  
Michigan State University – East  
Lansing, US
- Salvatore Greco  
University of Catania, IT
- Jussi Hakanen  
University of Jyväskylä, FI
- Johannes Jahn  
Universität Erlangen-Nürnberg,  
DE
- Andrzej Jaszkievicz  
Poznan University of Technology,  
PL
- Milosz Kadzinski  
Poznan University of Technology,  
PL
- Kathrin Klamroth  
Universität Wuppertal, DE
- Karl Heinz Küfer  
Fraunhofer ITWM –  
Kaiserslautern, DE
- Christoph Lof  
TU Delft, NL
- Manuel López-Ibáñez  
University of Manchester, GB
- Kaisa Miettinen  
University of Jyväskylä, FI
- Sanaz Mostaghim  
Universität Magdeburg, DE
- Boris Naujoks  
TH Köln, DE
- Frank Neumann  
University of Adelaide, AU
- Luís Paquete  
University of Coimbra, PT
- Robin Purshouse  
University of Sheffield, GB
- Patrick M. Reed  
Cornell University, US
- Günter Rudolph  
TU Dortmund, DE
- Stefan Ruzika  
TU Kaiserslautern, DE
- Serpil Sayin  
Koc University – Istanbul, TR
- Pradyumn Kumar Shukla  
KIT – Karlsruher Institut für  
Technologie, DE
- Roman Slowinski  
Poznan University of Technology,  
PL
- Ralph E. Steuer  
University of Georgia, US
- Theodor J. Stewart  
University of Cape Town, ZA
- Michael Stiglmayr  
Universität Wuppertal, DE
- Lothar Thiele  
ETH Zürich, CH
- Selvakumar Ulaganathan  
Noesis Solutions – Leuven, BE
- Daniel Vanderpooten  
University Paris-Dauphine, FR
- Margaret M. Wiecek  
Clemson University, US

