Author(s): Ehrgott, Matthias; Eichfelder, Gabriele; Küfer, Karl-Heinz; Lofi, Christoph; Miettinen, Kaisa; Paquete, Luís; Ruzika, Stefan; Sayın, Serpil; Steuer, Ralph E.; Stewart, Theodor J.; Stiglmayr, Michael; Vanderpooten, Daniel

Title: Personalization of Multicriteria Decision Support Systems

Year: 2018

Version: Accepted version (Final draft)

Please cite the original version:

Ehrgott, M., Eichfelder, G., Küfer, K.-H., Lofi, C., Miettinen, K., Paquete, L., Ruzika, S., Sayın, S., Steuer, R. E., Stewart, T. J., Stiglmayr, M., & Vanderpooten, D. (2018). Personalization of Multicriteria Decision Support Systems. In K. Klamroth, J. Knowles, G. Rudolph, & M. Wiecek (Eds.), Personalized Multiobjective Optimization : An Analytics Perspective (Dagstuhl Seminar 18031) (pp. 55-70). Dagstuhl Publishing. Dagstuhl Reports, 8. https://doi.org/10.4230/DagRep.8.1.33

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 Sayın, 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/disaggegating 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.

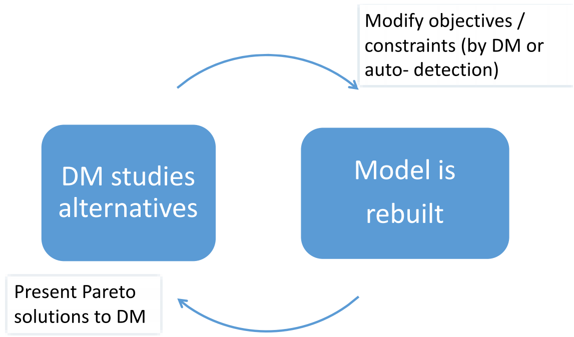

Figure 4 The framework for personalized decision support.

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 ϵ-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 $\mathcal{F} := \{f_i : \mathbb{R}^n \to \mathbb{R} \mid i = 1, \ldots, k\}$ 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 $h_1, \ldots h_m,\ g_1, \ldots, g_l : \mathbb{R}^k \to \mathbb{R}$ 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

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

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

▶ **Example 1.** Let $X = \mathbb{R}^n$ and $\mathcal{F} = \{f_1, f_2, f_3 \colon \mathbb{R}^n \to \mathbb{R}\}$. 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 $h_2(y_1, y_2, y_3) = w_2y_2 + w_3y_3$ we get

$$\min_{x\in X} \begin{pmatrix} f_1(x) \\ w_2f_2(x) + w_3f_3(x) \end{pmatrix}. \tag{$P_A(w_2,w_3)$}$$

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

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

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

$$\Delta_1 = f_1(\bar{x}), \qquad \Delta_2 = \Delta_3 = w_2f_2(\bar{x}) + w_3f_3(\bar{x}) + \delta$$

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

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

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

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 $\Delta_i \ge f_i(\bar{x})$, $i = 1, 2, 3$.
- 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 $\Delta_1 \ge f_1(\bar{x})$, $\delta \ge 0$ and

$$\Delta_2 = \Delta_3 = w_2f_2(\bar{x}) + w_3f_3(\bar{x}) + \delta\ .$$

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:

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

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 afromontane 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

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

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 $x_1, \ldots, x_m$ 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.

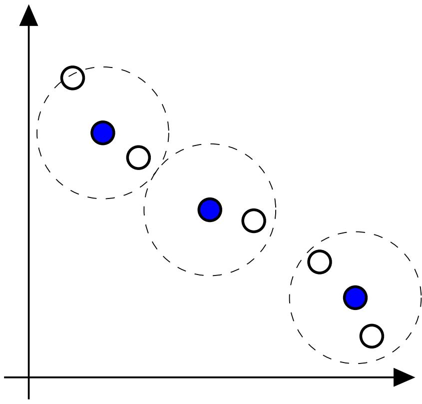

Figure 5 Illustration of a representation based on coverage.

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.,

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

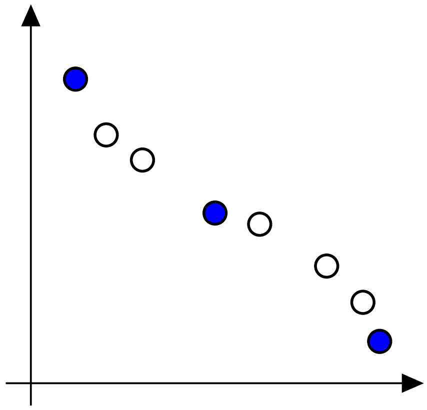

Figure 6 Illustration of representation based on uniformity.

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 ε'-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) $\min\{Cx : Ax \leqq b\}$ 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 $y^{AI}_k = \max\{y_k : y \in Y\}$ for $k = 1, \ldots, p$.
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 $v^k, k = 0, \ldots, p$ in $\mathbb{R}^p$

$$v^k_l = \begin{cases} y^{AI}_l, & l \neq k, \\ \phi + \hat{y}_k - e^T v^0 & l = k. \end{cases}$$

   - The convex hull S of $\{v^0, \ldots, v^p\}$ is a simplex containing Y.
   - The convex hull $\hat{S}$ of $\{v^1, \ldots, v^p\}$ 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 $\min\{t : q + te \in Y, t \geqq 0\}$ 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 representatiom.

▶ **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 $ds \leqq d(r^1, r^2) \leqq \sqrt{p}ds$. Hence, R is a ds-uniform representation of Y_N.*

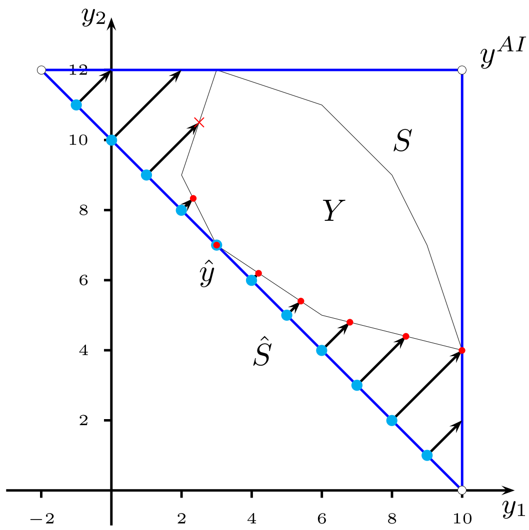

Figure 7 The revised boundary intersection method.

2. *Assume that the width* $w(S^j) \geqq ds$ *for the projection* S^j *of all maximal faces* Y^j *of* Y_N *on* $\hat{S}$*. Then* R *is a ds-uniform* $d_{\sqrt{p}ds}$*-representation of* Y_N*.*

In this section we outline how to adapt to the situation where Y_N is an explicitly given set of finitely many points. To this end, we now modify the RNBI method so that it becomes applicable to the case of $Y_N = Y = \{y^j : j \in J\}$ being an explicily given finite set. The main obstacle in doing this is that the sub-problem

$$\min\{t : q + te \in Y, t \geqq 0\}$$

that is solved for each reference pooint q will most often be infeasible. To avoid this situation, we replace Y in the sub-problem by $\hat{Y} = Y + \mathbb{R}^q$. Since $Y_N = \hat{Y}_N$, this has no effect on the Pareto optimal set, but the new sub-problem

$$\min\left\{t : q + te \in Y + \mathbb{R}^p_{\geqq}, t \geqq 0\right\}$$

is feasible. To solve it, we define $t(q) = \min_{j\in J} \max_{k\in\{1,\ldots,p\}}\{y^j_k - q_k\}$ and $r(q) = argmin_{j\in J} \max_{k\in\{1,\ldots,p\}}\{y^j_k - q_k\}$.

To compute, for reference point q, the intersecion of the ray $\{q + te : t \geqq 0\}$ with the cone $y^j + \mathbb{R}^q_{\geqq}$ dominated by y^j, the l_∞-distance $t(q)$ to the Pareto optimal point y^j is computed, and the closest point to q is chosen as a representative point $r(q)$. Then the representative set is $R = \{r(q) : q \in Q\}$.

There are a number of research questions related to this approach:

- Can quality guarantees in terms of uniformity and coverage error be proven?
- What is the cardinality of R given the cardinality of Q?

Representations based on clustering

A very simple, yet potentially effective idea for computing representations in the case of an explicitly given set of alternatives in the context of a geometric vision is based on *clustering*. The idea is to compute a certain number, say K clusters very quickly and retrieve information about the quality of the representation. In many real-life datasets, the density of points is 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 ε-*dominance* relation $\preceq_\varepsilon$ defined as follows:

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

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, $i + 1, \ldots, p$. 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 $y \preceq_\varepsilon y'$ (ε-*coverage*),
(ii) for any $y, y' \in K_{\varepsilon,\varepsilon'}$, not$(y \preceq_{\varepsilon'} y')$ and not$(y' \preceq_{\varepsilon'} y)$ (ε'-*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 ε to define the precision of the representation and then set ε' as large as possible. When it is possible to set $\varepsilon' = \varepsilon$, an $(\varepsilon, \varepsilon')$-kernel is called an ε-kernel.

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

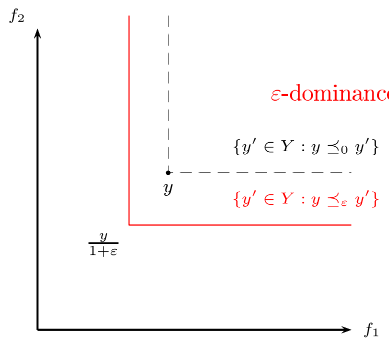

Figure 8 Dominance ($\preceq_0$) and ε-dominance ($\preceq_\varepsilon$) relations.

▶ **Theorem 4.**
- *If $p = 2$, an ε-kernel always exists (with $\varepsilon' = \varepsilon$).*
- *If $p \geq 3$, an $(\varepsilon, \varepsilon')$-kernel exists if and only if $\varepsilon' \leq \sqrt{1+\varepsilon} - 1$.*
- *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:

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

Note that this neighborhood is defined with a relation $\preceq_{\varepsilon'}$ which is used in the stability condition to define an $(\varepsilon, \varepsilon')$-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. Sayın. 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 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

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