Collective Preferences in Evolutionary Multi-Objective Optimization: Techniques and Potential Contributions of Collective Intelligence

Daniel Cinalli
Universidade Federal Fluminense
Niterói (RJ) Brazil
drcinalli@gmail.com

Luis Martí
Pontifícia Universidade Católica do Rio de Janeiro
Rio de Janeiro (RJ), Brazil
lmarti@ele.pu-rio.br

Nayat Sanchez-Pi
Instituto de Lógica, Filosofia e Teoria da Ciência (ILTC)
Niterói (RJ) Brazil
nayat@iltc.br

Ana Cristina Bicharra Garcia
ADDLabs, Universidade Federal Fluminense
Niterói (RJ) Brazil
bicharra@ic.uff.br

ABSTRACT

This paper reviews suitable techniques of interactive and preference-based evolutionary multi-objective algorithms to achieve feasible solutions in Pareto-optimal front. We discuss about possible advantages of collective environments to aggregate consistent preferences in the optimization process. Decision maker can highlight the regions of Pareto frontier that are more relevant to him and focus the search only on those areas previously selected. In addition, interactive and cooperative genetic algorithms work on refining users’ preferences throughout the optimization process to improve the reference point or fitness function. Nevertheless, expressing preferences from a unique or small group of decision makers may raise unilateral choices issues and pour hints in terms of search parameter. Supported by a large group of human interaction, collective intelligence is suggested to enhance multi-objective results and explore a wider variety of answers.

Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous;
J.4 [Computer Applications]: Social and behavioral sciences;
G.1.6 [Mathematics of Computing]: Optimization;
L.6 [Science and Technology of Learning]: Society/Community

General Terms

Algorithms

Keywords

collective intelligence; preferences; reference points; evolutionary multi-objective optimization algorithms

1. INTRODUCTION

Multi-objective optimization problems (MOPs) are optimization problems where two or more objective functions must be simultaneously optimized [16]. As opposed to single-objective where just one function is used to mathematically describe the subject, MOP allows different and separated functions to model each objective of the problem. Consequently, the solution to a MOP is not a single point, but a set of points that represent different trade-offs of the objective functions. Therefore, its is frequently necessary to involve a higher-level decision making entity that chooses a solution from that set according to some a priori preferences.

Multi-objective optimization algorithms pursue two distinct goals of similar importance. The first one is to drive the convergence of solutions as close as possible to the optimal trade-off surface. This is a natural direction to assure near-optimality properties to any optimization process. The second goal tries to produce a uniform distribution of answers over the true optimal set. Diversity in the pool of final solutions guarantees good representation of alternatives to decision makers.

The MOP solution set can be determined by mathematical programming methods if the problem has certain characteristics (e. g., linearity or convexity of the objective functions). However, in the general case, the MOP is an NP-complete problem and, hence, deterministic search techniques are unsuitable. Metaheuristic approaches are a viable alternative to handle this issue in a stochastic fashion. Due to their population-based nature, evolutionary algorithms (EAs) can find multiple optimal solutions in one iteration run. Therefore, EAs have been applied to address this task in a reasonable time frame since the mid-eighties.

Multi-objective evolutionary algorithms (MOEAs) [8] can—or even must—take advantage of decision makers’ pref-
erences to drive the search through the objective space, focusing only on solutions of interest, instead of all possible solutions.

Nevertheless, expressing a preference from a unique decision maker in the optimization process may raise some concerns regarding unilateral choices and possibly pour hints in terms of search parameter. In this case, a global outcome built on the aggregation of vast and diverse users’ intelligence would be a helpful input parameter not only to reference the search, but also to explore a wider diversity of answers and enhance the results through multi-user interaction.

The present work reviews suitable techniques of preference-based multi-objective algorithms and interactive EAs to achieve feasible solutions in Pareto front. It also brings a discussion about possible advantages of collective intelligence environments to aggregate consistent preferences in the optimization process and overcome expected difficulties of the MOPs.

The remainder of this paper is organized as follows. Section 2 covers some definitions for MOPs and collective intelligence field. Section 3 outlines different techniques of preferences and reference points MOEAs. Possible contributions of collective intelligence to MOEAs are explored in Section 4. Finally, some conclusive remarks are put forward.

2. FOUNDATIONS

Basically, a MOP minimizes \( F(x) = \{ f_1(x), ..., f_k(x) \} \) subject to \( m \) constraints \( c_i(x) \leq 0, i = 1, ..., m \); where \( x \in \Omega \) is an \( n \)-dimensional decision variable. The solution to this problem can be expressed by relying on the Pareto dominance relationship. A \( x = \{ x_1, ..., x_k \} \) is said to dominate \( v = \{ v_1, ..., v_k \} \) (denoted as \( x \prec v \)) iff \( \forall i \in \{ 1, ..., k \}, \ x_i \leq v_i \) and \( \exists i \in \{ 1, ..., k \} \) such that \( x_i < v_i \). The Pareto-optimal set, \( P^* \), of solutions in multi-objective optimization for a given MOP is defined as \( P^* = \{ x \in \Omega, \not\exists x' \in \Omega \) such that \( F(x') < F(x) \). Elements of \( P^* \) are called non-dominated solutions. Similarly, the codomain of the set is known as the Pareto-optimal front [9, 8].

In other contexts, MOP’s solutions are called non-inferior, admissible or efficient solutions. The elements of this set are non-dominated with respect to all the others. Although elements of this solution set are equality optimal, it is required an interaction with a human decision maker to select the appropriated answer for the desired purpose.

MOEAs follow the common concepts of evolutionary algorithms. In general, an EA detains population of individual solutions subject to a set of operators and evaluated by one or many fitness functions. Operators are in charge of tasks ranging from reproduction of the individuals (crossover and mutation) to pushing selection and competition pressure on them [19]. The fitness function measures if certain solution satisfies the objective functions, it assigns a value to each available solution in the current population. The underlying idea of MOEAs is to identify the set of individuals non-dominated by the rest of the population in every generation. This process of comparison continues until, for example, neither non-dominated solutions comes out.

The quality of MOEAs’ trade-off fronts can be quantitatively evaluated by many different techniques. The hypervolume or S-metric [18, 34] is a quantitative measure of the space covered by the non-dominated points. It represents the union of hypercubes \( a_i \) defined by a non-dominated point \( a_i \) and a reference point \( x_{ref} \) as defined in (1).

\[
S(M) = \Lambda(\bigcup_{m \in M} a_i \in M) = \Lambda(\bigcup_{m \in M} \{ x | m < x < x_{ref} \}). \quad (1)
\]

A second measure widely used is the coverage of two sets [34]:

\[
C(X', X'') = \frac{|\{ a' \in X'; \exists a' \in X': a' \prec a'' \}|}{|X''|}, \quad (2)
\]

where \( X', X'' \subseteq X \) are two sets of decision vectors and function \( C \) maps the percentage of domination from one set to another in the interval \([0, 1]\). Although convex regions may be preferred by the hypervolume, it can be used independently to evaluate the performance of different multi-objective algorithms. The coverage of two sets technique has no restriction related to the shape of Pareto front, but it does not express how much better one set is over the other.

The field known as collective intelligence (COIN) [24, 27] has emerged after the development of social network technologies and interactive on-line systems in the beginning of 2000. COIN is a self-organized group intelligence arisen from participatory and collaboration actions of many individuals. Shared tasks or issues are handled by singular contributions in such a manner that their aggregation process outcome better results than each individual input. This phenomenon develops a sui-generis intelligence, bigger than its isolated pieces, and creates a final product of their combination.

A decision-making process over the Internet has to manage users’ interactions and get valuable knowledge concealed or dispersed in the group. This environment includes large and heterogeneous audiences that are mostly independent among each others and are not specialized in the subject. Therefore, the problem must be decomposed in tasks that sustain diversity and transient members’ attendances. Collective decision-making engages all the knowledge generated in the first stage of users’ interaction and aggregates singular inputs into a global perspective for a decision. Some techniques to reveal the group opinion or preferences are [29, 24]: voting, recommendation systems, judgement aggregation, prediction markets and rating scales. Others initiatives collect the best information available from the crowd: wikis, document ranking and deliberation maps. While the elicitation of ideas approach assembles answers or suggestions to make them converge as a consensus.

3. PREFERENCES IN MOEAS

There is a rich diversity of approaches to deal with MOPs. Many of them are aggregative methods and transform the problem from multi-objective to mono-objective. This conversion manages user-defined parameters (preferences) to bias the search during the optimization phase and guide it towards the final result. Some traditional techniques are dependent to the selection of objectives’ weights or boundaries. The weighted sum method [10] associates a weight to each objective function and takes its sum to revert the problem to a mono-objective equation: \( F(x) = \sum_{m=1}^{M} \lambda_m f_m(x) \). Lexicographic ordering method [17, 3] prevents the manipulation of \( a \) priori weights by setting priorities to the objective functions and arrange them in order of importance. The \( c \)-constraint approach selects one objective function to be optimized and restricts all the others within user-specified inequality constraints [16]:

\[
\min f_i(x); \text{subject to } f_k(x) \leq \varepsilon_k \text{ where } k = 1, \ldots, n; k \neq i.
\]
Distance-to-a-reference method [9] compares a vector \( f \in \mathbb{R}^k \) correspondent to an ideal or reference objective chosen by the decision maker against a solution candidate \( x \). The Goal Attainment technique [11] also uses a vector \( f \in \mathbb{R}^k \) of reference objective function plus a weighted vector \( w \) to minimize a scalar coefficient \( \Psi \) in such a way that: \( f_i(x) - w_i, \Psi \leq x_k \), where \( i = 1, \ldots, k \).

Many evolutionary frameworks have been applied to MOPs and broadly used to optimization. Some MOEAs like: SMS-EMOA [6]; MO-CMA-ES [20]; NSGA-II [13] or SPEA2[33], exploit a set of solutions in parallel and supply a starting point to new developments on the field. All of them look for the Pareto optimal set, \( P^* \), that ideally contains all non-dominated solutions uniformly distributed along the optimal frontier. Besides the search for a trade-off set, the decision maker still must identify a final answer from it and \( P^* \) might be extremely large or possibly infinite. It raises a pertinent question: when and why can preferences be included into the search process?

MOEAs techniques are classified by its articulation of preferences: \( a$-$priori \) approach performs decision before the searching process; interactive (progressive) method combines search and decision making; \( a$-$posteriori \) technique searches before making decision.

In this context, preferences are the expression of values and subjective impressions regarding the trade-offs points. It transforms qualitative feelings into quantitative values to restrict the objective space. It is largely sensitive to the non-technical, aesthetic and qualitative experiences of the decision makers. Preference is represented by a set of criteria \( I \), where \( i \in I \) corresponds to one preference information assigned to one attribute. The boundaries are usually set by the ideal (\( z^* \)), utopian (\( z^+ \)) and nadir (\( z^{nad} \)) points [30]. Therefore, a local preference can be expressed as a vector of weights over the objectives, a lexicographic sorting of objectives, a set of constrains, trade-off information or reference points for the search, among others representations.

The challenge is no longer just to obtain a diversity of answers in the entire high-dimensional Pareto front, but also achieve a consistent global preference for the problem and retrieve the expected solutions from it. In this sense, the more reliable is the preference vector, the better are the trade-off answers obtained. Reference points and interactive techniques are different strategies used to bring more relevant solutions to the decision makers.

Decision maker’s preferences can highlight the regions of Pareto frontier that are more significant to him/her. Instead of a whole computation of the front, reference point multi-objective algorithms get suggestions or hints to focus the operation only on those areas previously selected. The outcome is a preferred sub-set of Pareto front that enriches the analysis of the decision maker(s).

A reference point assigns desirable values to each one of the objective functions. The resulting objective vector represents the decision maker’s preference. Let \( X \) be the set of feasible solutions in the decision space. To each decision vector \( x \in \Omega \) is assigned one objective vector, \( z \in \mathbb{Z} \).

Given a reference point \( z^o \) for an \( n \)-objective optimization problem of minimizing \( F(x) = \{ f_1(x), \ldots, f_k(x) \} \), one of the families of achievement scalarizing functions can be stated as follows [21]:

\[
\sigma(z, z^0, \lambda, \rho) = \max_{i=1, \ldots, n} \{ \lambda_i(z_i - z_i^0) \} + \rho \sum_{i=1}^{n} \lambda_i(z_i - z_i^0), \tag{3}
\]

where \( \sigma \) is a mapping from \( Z \) onto \( \mathbb{R} \), \( z^0 = (z_1^0, \ldots, z_n^0) \) is a reference point vector, \( \lambda = (\lambda_1, \ldots, \lambda_n) \) is a scaling coefficients vector, and \( \rho \) is an arbitrary small positive number. Because of the scaling coefficients and reference points, many functions can be formed. The achievement problem can be built as follows:

\[
\min \sigma(z, z^0, \lambda, \rho) \tag{4}
\]

subject to \( x \in \Omega \)

Wierzbicki’s reference point approach [30, 31] concentrates the search of non-dominated solutions near the selected point and relies on the scalarization method. If the answers obtained do not satisfy the users, a new reference point is proposed and the process repeats.

Deb et al. [15] extend the number of reference points to cover more than one region in Pareto front and reach more significant solutions. They proposed a reference-point-based NSGA-II procedure (R-NSGA-II) to find a set of solutions in the neighbourhood of the corresponding Pareto optimal. Figure 1 shows five reference points and their respective points in Pareto frontier extracted from the well-known ZDT1 test suite [32].

Deb and Kumar [12] combined the NSGA-II algorithm with the light beam search to produce a set of solutions in the region of interest. The hybrid light beam MOEA modifies the crowding operator to incorporate preferences.

Rudolph et al. [26] developed an intersection between the reference point concept and the averaged Hausdorff distance (\( \Delta_P \)) as a selection criterion to measure the distance of the outcome set \( X \subset \Omega \) to the image of \( P^* \).

Said et al. [5] presented a new dominance relation that replaces the Pareto dominance and favours solutions near the reference point indicated. Named reference solution-based dominance (r-dominance), the r-NSGA-II develops a strict partial order among Pareto-equivalent solutions and gives a stronger selection pressure on the population.

The Interleaving Guidance method [7] guides the EA in both the decision and objective spaces to accelerate and specialize the convergence process. The decision space \( S \) is divided into multiples hyperspheres \( S = [s_0, s_1, \ldots, s_n] \). These
spheres use preferences information to converge the solutions towards areas with non-dominated points.

Facing a collective scenario where many users have to take part of the decision process, a few and limited work on reference point can be found. Pfeiffer and others [25] rank the solutions according their Euclidean distances to each reference point. Those points close to the reference points are assigned with the lowest crowding distance, which favours its selection by the algorithm.

The negotiation support system called W-NSS-GPA [4] takes the reference points and the decision makers’ hierarchy levels as weights to calculate the Social Reference Point (SRP, the aggregation of all preferences). However, W-NSS-GPA uses the average mean of reference points to reach the SRP and this strategy do not represent the diversity of a genuine collective preference (see next section).

Interactive genetic algorithms were successfully applied to get feedbacks of transitional results throughout the evolution process. MOEAs can handle intermediate non-dominated solutions to the decision maker and improve the search with a reference point or fitness function adjustments. Deb et al. [14] proposed an interactive MOEA approach termed PI-EMO. This technique changes progressively the value function after every few generations to lead the optimization on more preferred solutions. After t generations, the user analyses $\eta (\geq 2)$ well-sparsed non-dominated solutions and provides a complete or partial ranking of preference information about them. Thiele et al. [28] introduced a preference-based interactive algorithm (PBEA) that adapts the fitness evaluation with an achievement scalarizing function to guarantee an accurate approximation of the desired area in Pareto-optimal front. The algorithm IGAMII [2] applies fuzzy logic to simulate the human decision maker and relieve the constant interaction during the evolution.

4. INTEGRATING COLLECTIVE INTELLIGENCE

In general case, MOP is built on complex scenarios of high-dimension and large problem space. Even meta-heuristics approaches like MOEAs can have difficulties to solve the problems in a reasonable time and quality. On the other hand, human beings are used to multi-objective situations in their everyday lives. Those complex scenarios that are hard for computer might be easier or natural for ordinary people. So, why not to take advantage of people’s mind and improve the multi-objective algorithms to find better solutions?

COIN can be a special contribution to make MOEAs go beyond their reach. Human characteristics such as perception, strategy, weighting factors, among others subjectivities might be introduced into the algorithm to generate a better pool of answers and enhance the optimization process. The diversity of people and the synergy of actions inside collective environments can be used to understand the conflicting situation involving the objectives and give hints to a jump in its quality.

MOEA community neglected the scenario where multiples decision makers would express their preferences. Hence, very few algorithms consider more than one user to select the reference point for MOPs or to interact throughout the evolution process. Assuming a unique decision maker and a priori preferences in the optimization, the search faces the risk of mistaken guidelines. A chosen reference point might be too ambitious and miss the set of non-dominated objective vectors or too easily attainable and jeopardizes the search [22]. A collective reference point produced by the interaction and aggregation of opinions from many participants may provide a more accurate reference point than designed by only one expert (unilateral).

There are plenty of examples where COIN techniques overcome individuals’ results and legitimize the attempt to incorporate this method within MOEAs.

In Threadless a community of 500.000 people rates members t-shirt designs, InnoCentive site hosts companies’ problems and offers a cash prize to the one who presents the most preferred solution. Both initiatives harness collective ideas, elaborate global preferences to hit the target and outperform a design expert.

Affinova delivers a service to companies who want to improve their innovation and marketing rates in consumer packaged goods, retail, financial services and design. Its platform empowers teams to develop ideas, collect consumer feedback and predict the best execution plan for them. Danone, a global food company, used their services to launch the Activia product line in USA and the result beat the initial forecast by four times [1].

Another example is the puzzle game about protein folding: Foldit; it uses the human brain’s natural three-dimensional pattern matching to solve the problem of protein structure prediction. The highest scoring solutions are analysed by researchers and validated if applicable in real problems or not. Users in Foldit has already helped to decipher the crystal structure of the Mason-Pfizer monkey virus (M-PMV) retroviral protease [23].

5. A MOEA+COIN CASE STUDY

Considering a common problem of resource placement and assignment faced by many companies, industries and government: a domain with areas for resource extraction must address processing units in such a way that optimizes its operational costs and production of collected resources.

This general idea drives the allocation of available resources to operate in an economic way and, at the same time, prioritize the performance or production. The problem —to put it in simple terms— has to find a good solution for positioning the processing units according the resource area. It is formally represented as

$$\min \sum_{i=1}^{N} \sum_{j=1}^{M} \sigma_{ij} d_{ij} + \sum_{j=1}^{M} c_{j} \mu, \quad (5)$$

$$\max \sum_{i=1}^{N} \sum_{j=1}^{M} \sigma_{ij} v_{j}. \quad (6)$$

Let $\mu$ be the cost of one processing unit, $v$ the productive capacity of one processing unit linked to one resource area, $N$ a set of available positions to production units, $M$ a set of available positions to resource area and $D$ a distance matrix $(d_{ij})_{N \times M}$, where $n \in N$ and $m \in M$. The decision variables are the processing unit $c_{j} \in (j \in M)$ that assumes 1 if it is placed at position $j$ or 0 otherwise and $\sigma_{ij}$ that assumes 1 if there is a link between the resource area at position $i \in N$ and the processing unit at position $j \in M$. 
This MOP might increase the search complexity with many different constraints from real life and several interdependencies among the variables. Progressive articulation of preferences and collective intelligence can implement a dynamism not managed by a priori methods and enhance the efficiency of this MOP.

Hence, the problem described is a candidate for this experiment due to some reasons: a) the objectives and decision variables are meaningful to the group, the problem is intuitive and allows an interaction with the crowd’s cognition; b) incentive engines and gamification can be used to retain the users’ interest on the interaction during the optimization; c) the problem can be decomposed in small blocks to be presented to the participants; d) the feedbacks can be parallelized in synchrony with the evolution of individuals in a MOEA.

This paper presented a case study and hinted a hybrid approach that will be developed in future research: Interactive Collective Intelligence Evolutionary Algorithm (ICIEA). The algorithm will be applied to the resource placement problem explained before. So far, an experimental phase has started. The computational model was developed and the evolutionary operators (crossover, mutation and selection) were adjusted for the problem. Some preliminary results have been achieved without any collective interaction. Table 1 shows the best results after 60 independent executions per EA. Using the same scenario, every MOEA execution performed 300 generations for 200 individuals (population). The crossover and mutation rate were 0.30 and 0.10, respectively. Figure 2 reports the distribution of results comparing different multi-objective genetic algorithms. For this problem, NSGA-II returned better outcomes than SPEA2 and the values are more concentrated to the mean. Figure 3 demonstrates one solution from Pareto-optimal set composed of six resource areas and two processing units.

Once the right evolutionary operators are tuned to this MOP, the research plans the integration of COIN as the second stage of the experiment. Three new approaches will be implemented: voting with reinforced evaluation; collective local search; collective reference point. Voting with reinforced evaluation approach submits some individuals from population to the users and asks for their preferred ones. Those chosen will be selected for the next generation with an extra value, based on a multi-attribute utility function, in its fitness. Collective local search will open the individuals for users’ update, following the Find-Fix-Verify method. After the change, the new individuals can be directly inserted into the next generation or an aggregation can be implemented (clustering) to produce not only a new sample, but also a collective reference point to drive the search through the group’s preferences. At the end, the hypervolume and coverage measure will be used to compare the actual results without COIN with the future ones.

<table>
<thead>
<tr>
<th>Crossover</th>
<th>Mutation</th>
<th>Selection</th>
<th>S metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td>Uniform NSGA-II</td>
<td>Uniform SPEA2</td>
<td>0.6058</td>
</tr>
<tr>
<td>One Point</td>
<td>Uniform NSGA-II</td>
<td>Uniform SPEA2</td>
<td>0.6008</td>
</tr>
<tr>
<td>Two Point</td>
<td>Uniform NSGA-II</td>
<td>Uniform SPEA2</td>
<td>0.6032</td>
</tr>
<tr>
<td>Best P*</td>
<td>Uniform NSGA-II</td>
<td>Uniform SPEA2</td>
<td>0.5972</td>
</tr>
</tbody>
</table>

The combination of MOEA and COIN will bring human subjectivity and cognition into the optimization search. This benefit will not come from one decision maker only, but a group of people that is more intelligent when is working together.

6. FINAL REMARKS
In this paper we have discussed about preferences in multi-objective optimization evolutionary algorithms. MOEAs can take advantage of decision makers’ preferences to guide the search through relevant regions of Pareto-optimal front. Suitable implementation techniques of preference-based multi-objective algorithms and interactive EAs were pointed out as an alternative to handle the dynamism not expected by a priori methods. Particularly, we review some techniques of reference points.

However, a global outcome built on the aggregation of multiples users’ intelligence (COIN) would be a helpful input parameter not only to reference the search, but also to explore a wider diversity of answers and enhance the results through participants’ interaction.
Finally, a resource allocation case study problem was presented and future works outlined.

References


