J Optim Theory Appl (2014) 162:845–855
DOI 10.1007/s10957-013-0498-y

Two-Sided Pareto Front Approximations

I. Kaliszewski · J. Miroforidis

Received: 16 February 2013 / Accepted: 25 November 2013 / Published online: 17 December 2013

Abstract A new approach to derive *Pareto front* approximations with evolutionary computations is proposed here.

At present, evolutionary multiobjective optimization algorithms derive a discrete approximation of the Pareto front (the set of objective maps of efficient solutions) by selecting feasible solutions such that their objective maps are close to the Pareto front. However, accuracy of such approximations is known only if the Pareto front is known, which makes their usefulness questionable.

Here we propose to exploit also elements outside feasible sets to derive pairs of such Pareto front approximations that for each approximation pair the corresponding Pareto front lies, in a certain sense, in-between. Accuracies of Pareto front approximations by such pairs can be measured and controlled with respect to distance between elements of a pair.

A rudimentary algorithm to derive pairs of Pareto front approximations is presented and the viability of the idea is verified on a limited number of test problems.

Keywords Multiobjective optimization · Evolutionary algorithms · Lower Pareto front approximations · Upper Pareto front approximations

I. Kaliszewski (✉) · J. Miroforidis
Systems Research Institute, Polish Academy of Sciences, Newelska 6, 01-447 Warsaw, Poland
e-mail: ignacy.kaliszewski@ibspan.waw.pl

J. Miroforidis
e-mail: janusz.miroforidis@ibspan.waw.pl

J. Miroforidis
Treeffect Co, Gdów, Poland

1 Introduction

Evolutionary multiobjective optimization (EMO) algorithms [1–3] derive finite approximations of Pareto fronts (i.e. sets of efficient solutions, as defined below). They can be regarded, conventionally, as *lower approximations* (we assume that all objectives are to be maximized), because all their elements are feasible solutions. Yet, with the exception of test problems, Pareto fronts are generally not known. Hence the exact accuracy of such approximations is not known either.

To rectify this, we propose to work with elements outside the feasible solution set (infeasible solutions), with the objective to provide *upper approximations* of Pareto fronts. A pair consisting of a lower and an upper approximation forms an *approximation* of the Pareto front, whose accuracy can be controlled by *distance* between the lower and the upper approximation. Thus, the approach proposed herein effectuates the idea of two-sided Pareto front approximations, which is as yet absent from the literature on EMO. Exploiting explicitly infeasible solutions to provide two-sided approximations of Pareto fronts offers a new turn in research in the field.

Our research has been motivated by the absence, to our best knowledge, of papers pertaining to EMO or multiobjective optimization, in which active use of infeasible elements would be harnessed to approximate the Pareto front. The only exception is perhaps paper [4]; however, in that work, infeasible solutions were not generated intentionally, as it is done in our work.

The outline of the paper is as follows. In Sect. 2, necessary definitions are provided; in particular; lower and upper *shells*, which yield specific lower and upper approximations of Pareto fronts, are defined. In Sect. 3, an approximation accuracy measure is set, and a relaxation of the definition of upper shell is proposed with the purpose to have a construct more suitable for computations than upper shell itself.

In Sect. 4, a rudimentary evolutionary algorithm for approximating Pareto fronts within given accuracy is presented. The algorithm has been run on five test problems taken from literature and the results are reported in Sect. 5. Directions for further research are proposed in Sect. 6, whereas Sect. 7 contains the concluding remarks.

2 Definitions and Notation

Multicriteria Optimization (MO) problem is formulated as

$$\begin{aligned} &\max f(x) \\ &x \in X_0 \subseteq \mathbb{R}^n, \end{aligned} \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}^k;\ f = (f_1, \ldots, f_k),\ f_i : \mathbb{R}^n \to \mathbb{R},\ i = 1, \ldots, k,\ k \geq 2$, are objective (criteria) functions; max denotes the operator of deriving all efficient elements (see the definition below). We assume that X_0 has an interior.

The *dominance relation* $\prec$ is defined on $\mathbb{R}^n$ as

$$x' \prec x \quad \Leftrightarrow \quad f(x') \ll f(x),$$

where $\ll$ denotes $f_i(x') \leq f_i(x),\ i = 1, \ldots, k,$ and $f_i(x') < f_i(x)$ for at least one i.

If $x \prec x'$ then x is *dominated* by x' and x' is *dominating* x.

An element x of X_0 is called *efficient* iff

$$\nexists x' \in X_0 \quad x \prec x'.$$

We denote the set of efficient elements by N and the set $f(N)$ (the *Pareto front*) by P, $P \subseteq f(X_0)$.

Lower shell is a finite nonempty set $S_L \subseteq X_0$, elements of which satisfy

$$\forall x \in S_L \ \nexists x' \in S_L \quad x \prec x' \tag{2}$$

(thus no element of S_L is dominated by another element of S_L).

We define the *nadir point* y^{nad} as

$$y_i^{\text{nad}} := \min_{x \in N} f_i(x), \quad i = 1, \ldots, k.$$

Upper shell is a finite nonempty set $S_U \subseteq \mathbb{R}^n \setminus X_0$, elements of which satisfy[1]

$$\forall x \in S_U \ \nexists x' \in S_U \quad x' \prec x, \tag{3}$$

$$\forall x \in S_U \ \nexists x' \in N \quad x \prec x', \tag{4}$$

$$\forall x \in S_U \quad y^{\text{nad}} \ll f(x). \tag{5}$$

3 Approximations of *P*

We aim at constructing numerically viable two-sided approximations of P.

To derive S_L for which $f(S_L)$ is "close" to P, any EMO algorithm can be used (cf. [1, 2, 6–9]).

Since the definition of an upper shell involves N, this construct is not a suitable approximation of N. A more suitable construct, referring to S_L instead of N, namely an *upper approximation* A_U, is obtained by replacing:
condition (3) by

$$\forall x \in A_U \ \nexists x' \in A_U \quad x' \prec x, \tag{6}$$

condition (4) by

$$\forall x \in A_U \ \nexists x' \in S_L \quad x \prec x', \tag{7}$$

condition (5) by

$$\forall x \in A_U \quad y^{\text{nad}}(S_L) \ll f(x), \tag{8}$$

where $y^{\text{nad}}(S_L)$ denotes an element of $\mathbb{R}^k$ such that

$$y_i^{\text{nad}}(S_L) := \min_{x \in S_L} f_i(x), \ i = 1, \ldots, k$$

($y^{\text{nad}}(S_L)$ varies with S_L).

[1] Throughout the paper we assume that an upper shell exists. It may exist, as shown in Sect. 5 for test problems, but it may not exist either (cf. [5]). An upper shell does not exist if for some $x' \in X_0$ there is no $x \in \mathbb{R}^n \setminus X_0$ such that $x' \prec x$.

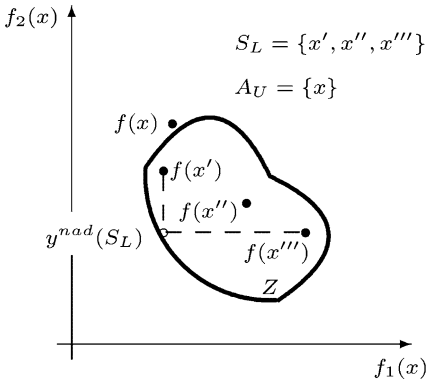

Fig. 1 An example when an element x, dominated by some element of N, belongs to A_U

By definition, an upper approximation A_U can contain elements which are dominated by some elements of N, as shown in Fig. 1, and certainly such elements are undesirable for the purpose. Condition (8) is meant to limit the domain for such elements. However, as S_L gets "closer" to N and $y^{\mathrm{nad}}(S_L)$ gets "closer" to y^{nad}, the chance for such elements being included in A_U decreases.

With S_L and A_U derived, the accuracy of the approximation of P by $f(S_L)$ and $f(A_U)$ can be measured as

$$\overline{acc}_P := \frac{1}{|S_L|} \sum_{x \in S_L} \min_{x' \in A_U} \left\| f(x) - f(x') \right\|,$$

or

$$acc_P := \max_{x \in S_L} \min_{x' \in A_U} \left\| f(x) - f(x') \right\|,$$

where $\| \cdot \|$ is a norm and $| \cdot |$ is cardinality of a set. In numerical experiments and applications, a form of normalization of $\overline{acc}_P$ and acc_P with respect to ranges of values of objective functions over e.g. S_L is advisable (cf. Sect. 5).

Those two indices measure only "closeness" of $f(S_L)$, $f(A_U)$. "Goodness" or "fairness" of the approximation of P by such constructs has to be ensured by standard EMO mechanisms.

In next section, we propose an algorithm for deriving two-sided approximations of P.

4 An Algorithm for Two-sided Approximations of P

The algorithm we propose below derives two-sided approximations of P, thus providing a way for approximation accuracy monitoring.

Let α_P denote the desired value of acc_P.

We limit the domain of searching in $\mathbb{R}^n \setminus X_0$ to some set

$$X_{\mathrm{DEC}} := \left\{ x \in \mathbb{R}^n \mid X_i^L \le x_i \le X_i^U, i = 1, \ldots, n \right\} \quad \text{such that} \quad X_0 \subseteq \mathrm{int}(X_{\mathrm{DEC}}).$$

By assumption, X_0 has an interior, hence elements of X_{DEC}, generated randomly, belong to X_0 with positive probability.

Algorithm EMO-APPROX

1. $j := 0,\ S_L^j := \emptyset,\ A_U^j := \emptyset$.
2. Select randomly η elements of X_0 and derive S_L^j.
3. Select randomly an element x of S_L^j and:
   3.1. derive an element $x' \in X_{\mathrm{DEC}}$ such that $x' \not\prec x$,
   3.2. if $x' \in X_0$ then update S_L^j and A_U^j with $S' = S_L^j \cup \{x'\}$, go to 3.4,
   3.3. update A_U^j with $A' = A_U^j \cup \{x'\}$,
   3.4. if $acc_P \leq \alpha_P$ or $j = j^{\max}$ then STOP,
   3.5. $j := j + 1$, go to 3.

In step 2, η is a parameter and derivation of S_L means that selected elements which do not satisfy condition (2) are to be removed.

In substep 3.1, to derive an element x' of the required properties, components of x are mutated until $x \in X_{\mathrm{DEC}}$ and $x' \not\prec x$ holds. Mutations can increase or decrease with probability 0.5 the value of a randomly selected component. The range of mutations decreases with the increasing j. If a mutation increases the ith component of x, then the value of this component after mutation is

$$x_i + \left(X_i^U - x_i\right) \times \left(1 - \mathrm{rnd}(0,1)^{2(1-\frac{j}{j^{\max}})}\right),$$

and if this mutation decreases the component, then the value of this component after mutation is

$$x_i - \left(x_i - X_i^L\right) \times \left(1 - \mathrm{rnd}(0,1)^{2(1-\frac{j}{j^{\max}})}\right).$$

Function rnd(0, 1) returns a random number from the range [0,1] with uniform probability. The presented method of mutation and the strategy of decreasing mutation range have been taken from the literature (cf. e.g. [6]).

In substep 3.2, the update of S_L^j means that elements of $S' = S_L^j \cup \{x'\}$ which do not satisfy condition (2) are to be removed from S', and only then $S_L^j := S'$. The update of A_U^j means that elements of A_U^j which do not satisfy condition (7) with respect to updated S_L^j are to be removed.

In substep 3.3, the update of A_U^j means that elements of $A' = A_U^j \cup \{x'\}$ which do not satisfy conditions (6), (7) and (8) are to be removed from A', and only then $A_U^j := A'$.

In substep 3.4, $j^{\max}$ is the maximal number of iterations in the algorithm.

There is no guarantee that the approximation accuracy monotonously improves by each iteration of EMO-APPROX (i.e. on $(i+1)$-th iteration acc_P takes a smaller value than on iteration i). The phenomenon is illustrated in Fig. 2. Indeed, suppose that $S_L = \{a, b\}$, $A_U = \{c, d\}$. Clearly, $acc_P^1 = \max\{\|f(a) - f(c)\|, \|f(b) - f(d)\|\}$ (the superscript indicates the iteration). Including e into S_L causes b to be eliminated from S_L (for e dominates b—condition (2)). Now we have $acc_P^2 = \max\{\|f(a) -$

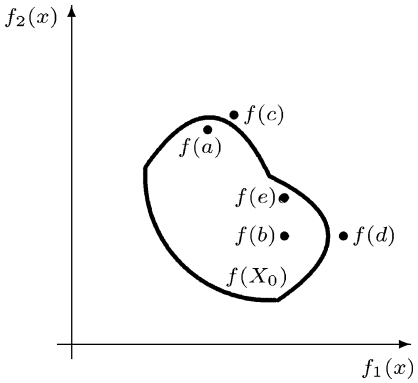

Fig. 2 Possible non-monotonous behavior of the algorithm

$f(c)\|, \|f(e) - f(d)\|\}$ and clearly $acc_P^2 \geq acc_P^1$, which means that the approximation accuracy has deteriorated. However, it can be expected that in successive iterations mutations of e or d can recover this local loss of accuracy.

As the algorithm is founded on genetic-type heuristics, no formal proof is offered that in general the algorithm is able to derive a two-sided approximation of P within a given accuracy. However, by means of two-sided approximations at least the behavior of such heuristics is put under control.

5 Numerical Experiments

We illustrate the behavior of EMO-APPROX on four test problems taken from [8], denoted DTLZ2a, DTLZ4a, DTLZ7a, Selri, and one taken from [10, 11], denoted Kita.

We normalized the accuracies $\overline{acc}_P$ and acc_P as follows:

$$\overline{acc}_P := \frac{1}{|S_L|} \sum_{x \in S_L} \min_{x' \in A_U} \left(\sum_{i=1}^{k} \left(\frac{f_i(x) - f_i(x')}{s_i^f} \right)^2 \right)^{\frac{1}{2}},$$

$$acc_P := \max_{x \in S_L} \min_{x' \in A_U} \left(\sum_{i=1}^{k} \left(\frac{f_i(x) - f_i(x')}{s_i^f} \right)^2 \right)^{\frac{1}{2}},$$

where $s_i^f := \max_{x \in S_L} f_i(x) - \min_{x \in S_L} f_i(x),\ i = 1, \ldots, k$ (the normalization factor varies with S_L).

We ran EMO-APPROX on the test problems with $j^{\max} = 9000$, $\eta = 100$ in each case, taking three shots of the algorithm behavior and the results it provided at $j = 3000$, $j = 6000$ and finally at $j = 9000$. Since it was not certain what values of the parameter α_P should be used, we set it to zero and we stopped the algorithm after the iteration count reached $j^{\max}$. In each case, X_{DEC} was assumed to be $[-0.2, 1.2] \times [-0.2, 1.2] \times \cdots \times [-0.2, 1.2]$ for all four DTLZ problems and $[-2.0, 9.0] \times [-2.0, 9.0]$ for the Kita problem.

Table 1 Test results—EMO-APPROX

Problem (n, m, b, k)	j	$\overline{acc}_P$	acc_P	$\#f$	$\|S_L\|$	$\|A_U\|$	$\|S_L\| + \|A_U\|$
DTLZ2a (8, 0, 8, 3)	3000	0.209	0.560	9054	752	93	845
	6000	0.206	0.574	18895	1643	157	1800
	9000	0.224	0.603	27796	2081	193	2274
DTLZ4a (8, 0, 8, 3)	3000	0.029	0.185	16369	248	180	428
	6000	0.017	0.110	30545	512	435	947
	9000	0.016	0.123	42436	805	633	1438
DTLZ7a (8, 0, 8, 3)	3000	0.030	0.259	5300	185	124	309
	6000	0.024	0.098	10638	348	209	557
	9000	0.019	0.158	15834	571	339	910
Selri (4, 0, 4, 3)	3000	0.048	0.229	3107	680	119	799
	6000	0.033	0.193	6116	1284	266	1550
	9000	0.026	0.219	9119	1926	377	2303
Kita (2, 3, 2, 2)	3000	0.024	0.905	4598	79	63	142
	6000	0.024	0.116	9082	117	86	203
	9000	0.017	0.076	13189	148	91	239

Table 1[2] shows the values of $\overline{acc}_P$ and acc_P for each problem and shot, where n, m, b and k are, respectively, the number of variables, the number of general constraints, the number of box constraints and the number of criteria; $\#f$ is the number of function f evaluations; $\|A_U\|$, $\|S_L\|$ and $\|A_U\| + \|S_L\|$ are the cardinality of, respectively, A_U, S_L and $A_U \cup S_L$.

Figures 3 and 4 present, respectively, the elements of S_L, A_U and $f(S_L)$, $f(A_U)$ for the Kita problem.

These test results constitute a rather limited base for drawing general conclusions. It is, nevertheless, possible to point to some regularities, which seem to be in line with the expected behavior of the algorithm.

Both accuracies improve monotonously for 3000, 6000 and 9000 iterations only in two instances (DTLZ7a and Kita), whereas for other problems improvements are not monotonous due to a phenomenon explained in previous section. In the remaining instances, there was no significant gain in increasing the number of iterations from 6000 to 9000, for at 9000 the accuracies are either worse or only slightly better.

In all instances, S_L has more elements then A_U. This is caused by the order in which the algorithm attempts to produce new elements of those sets—first in S_L and then in A_U. By interchanging this order, more balanced sets can be produced.

For the Kita, DTLZ2a and DTLZ4a problems, for which analytic forms of N are known, we have generated a number of elements of N (1852 elements for Kita, 2001 elements for DTLZ2a and DTLZ4a). In none of those problems is an element of A_U

[2]Pairs of sets S_L, A_U and $f(S_L), f(A_U)$ for all five problems at j equal to $3000, 6000, 9000$ can be obtained from the Authors on request.

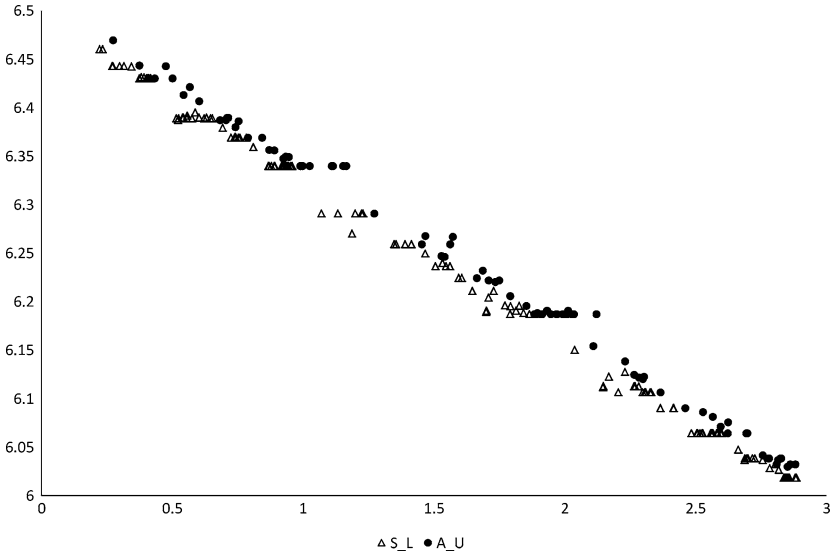

Fig. 3 Elements of S_L and A_U for the Kita problem

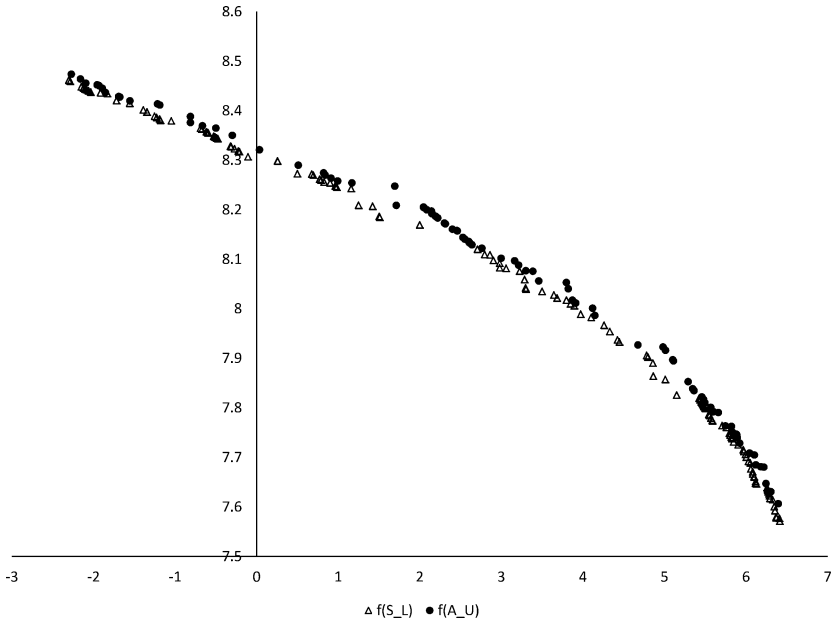

Fig. 4 Elements of $f(S_L)$ and $f(A_U)$ for the Kita problem

dominated by a generated element of N, which could be attributed to the strength of condition (8) (we have inspected only A_U derived in 9000 iterations).

Table 2 Test results—NSGA-II + EMO-APPROX

Problem (n, m, b, k)	Scaling	j	$\overline{acc}_P$	acc_P	# f	$\|S_L\|$	$\|A_U\|$	$\|S_L\| + \|A_U\|$
DTLZ2a (8, 0, 8, 3)	1	1000	0.223	0.624	3843	802	113	915
		2000	0.215	0.614	7592	1382	197	1579
		3000	0.207	0.608	11195	1898	247	2145
DTLZ4a (8, 0, 8, 3)	1	1000	0.044	0.191	4100	443	165	608
		2000	0.028	0.141	8201	577	247	824
		3000	0.023	0.119	12273	678	382	1060
DTLZ7a (8, 0, 8, 3)	0.01	1000	0.004	0.049	1811	555	315	870
		2000	0.003	0.078	3533	929	518	1447
		3000	0.002	0.078	5082	1314	690	2004
Selri (4, 0, 4, 3)	0.01	1000	0.036	0.223	1204	961	96	1057
		2000	0.023	0.123	2204	1673	166	1839
		3000	0.019	0.115	3204	2412	214	2626
Kita (2, 3, 2, 2)	0.001	1000	0.004	0.031	1552	471	295	766
		2000	0.003	0.023	2903	736	497	1233
		3000	0.002	0.020	4255	1041	634	1675

6 Further Directions

To demonstrate the potential in further fine tuning of the approach, we have coupled EMO-APPROX with the algorithm NSGA-II ([7]); the latter has proved itself a very effective in producing well distributed and accurate lower approximations (S_L) of P. In this experiment, NSGA-II has been instructed to derive, for each problem considered, a lower approximation of P with 200 elements in 200 generations ('generation' is NSGA-II parlance) to serve as a starting S_L for EMO-APPROX.

Next, EMO-APPROX has been run for each problem with the respective S_L. Because EMO-APPROX has been building successive S_L and A_U not from scratch, but from the results provided by NSGA-II, we have decreased $j^{\max}$ to 3000. And as we did not know how to account in EMO-APPROX for NSGA-II impact on mutating factor $(1 - \text{rnd}(0, 1)^{2(1-\frac{j}{j^{\max}})})$, for each problem we run the tandem NSGA-II + EMO–APPROX with four different scalings of the starting mutating factor values, namely with $1.0, 0.1, 0.01, 0.001$.

For each problem, at least one scaling produced lower or equal acc_P than algorithm EMO-APPROX run alone for 9000 iterations. From all such cases, we have selected those with the highest $\|S_L\| + \|A_U\|$. Results for runs selected in that manner are presented in Table 2.[3]

[3] Actually, for problem DTLZ2a, NSGA-II + EMO-APPROX in its best run produced slightly higher acc_P than EMO-APPROX, 0.608 versus 0.603, but considering the spread in the iteration count, such a difference can be regarded as nil.

Comparing Tables 1 and 2, it is evident that EMO-APPROX, producing two-sided approximations of the Pareto front, has benefited significantly from the preprocessing provided by NSGA-II—the tandem NSGA-II + EMO-APPROX has produced results comparable to, and in some cases distinctly better than those produced by EMO-APPROX alone, within three times less iterations. This strengthens our claim that, in the future, embedding the full range of EMO mechanisms into EMO-APPROX is worthwhile.

7 Conclusions

In this work, we limited ourselves to showing the viability of the idea to approximate Pareto fronts by pairs of lower and upper approximations. We also demonstrated how to improve interactively "closeness" between them. "Goodness" or "fairness" of approximations of P by such constructs constitutes the topic for further research. Another further topic is to provide means to derive pairs of S_L and S_U-like (A_U-like) constructs for cases where S_U does not exist. Some preliminary results pertaining to that issue have been already obtained and are reported in [5].

In the future experiments, to ensure even more uniform layouts of pairs $f(A_U)$, $f(S_L)$ along P, other genetic operators should be also exploited, whereas here we have confined ourselves only to the operator of mutation. We have done this deliberately to ensure clarity of the presentation and to demonstrate the viability of the concept of two-sided approximations of P.

The problem of providing accurate two-sided approximations with uniform layouts along Pareto fronts, being of interest in itself, has an immediate application in Multiple Criteria Decision Making, where a decision process can be enhanced if it is started with a not necessarily very accurate but uniform two-sided approximation of the Pareto set to roughly represent it. Next, in the course of the decision process, such approximations can be improved locally as directed by the decision maker's preferences [12, 13].

References

1. Deb, K.: Multi-Objective Optimization Using Evolutionary Algorithms. Wiley, Chichester (2001)
2. Coello Coello, C.A., Van Veldhuizen, D.A., Lamont, G.B.: Evolutionary Algorithms for Solving Multi-Objective Problems. Kluwer Academic, New York (2002)
3. Talbi, E.: Metaheuristics: From Design to Implementation. Wiley, New York (2009)
4. Legriel, J., Le Guernic, C., Cotton, S., Maler, O.: Approximating the Pareto front of multi-criteria optimization problems. In: International Conference on Tools and Algorithms for the Construction and Analysis of Systems, TACAS, pp. 69–83 (2010)
5. Kaliszewski, I., Miroforidis, J.: Real and virtual Pareto set upper approximations. In: Trzaskalik, T., Wachowicz, T. (eds.) Multiple Criteria Decision Making '11. The Publisher of University of Economics in Katowice, vol. 121 (2012)
6. Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs. Springer, Berlin (1996)

7. Deb, K., Pratap, A., Agarwal, S., Meyariva, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comput. **6**(2) (2002)
8. Knowles, J.: ParEGO: a hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. IEEE Trans. Evol. Comput. **10**, 50–66 (2005)
9. Hanne, T.: A multiobjective evolutionary algorithm for approximating the efficient set. Eur. J. Oper. Res. **176**, 1723–1734 (2007)
10. Kita, H., Yabumoto, Y., Mori, N., Nishikawa, Y.: Multi-objective optimization by means of the thermodynamical genetic algorithm. In: Voigt, H.-M., Ebeling, W., Rechenberg, I., Schwefel, H.-P. (eds.) Parallel Problem Solving from Nature–PPSN IV. Lecture Notes in Computer Science, pp. 504–512. Springer, Berlin (1996)
11. Kita test problem: http://delta.cs.cinvestav.mx/~ccoello/EMOO/testfuncs/, downloaded August 22, 2012
12. Kaliszewski, I., Miroforidis, J., Podkopayev, D.: Interactive multiple criteria decision making based on preference driven evolutionary multiobjective optimization with controllable accuracy—the case of ρ-efficiency. Systems Research Institute Report, RB/1/2011 (2011)
13. Kaliszewski, I., Miroforidis, J., Podkopayev, D.: Interactive multiple criteria decision making based on preference driven evolutionary multiobjective optimization with controllable accuracy. Eur. J. Oper. Res. **216**, 188–199 (2012)