Failure of Pareto-based MOEAs : Does Non-dominated Really Mean Near to Optimal?

IKEDA Kokolo
Tokyo Institute of Technology
Nagatuta-cho 4259
Midori-ku, Yokohama city
226-8502, JAPAN
psyche@fe.dis.titech.ac.jp

KITA Hajime
National Intstitution
for Academic Degrees
Otsuka 3-29-1 Bunkyou-ku, Tokyo
112-0012, JAPAN
kita@niad.ac.jp

KOBAYASHI Shigenobu
Tokyo Institute of Technology
Nagatuta-cho 4259
Midori-ku, Yokohama city
226-8502, JAPAN
Kobayasi@dis.titech.ac.jp

Abstract-

Many multi-objective evolutionary algorithms (MOEAs) have been proposed over the years. Main part of the most successful algorithms such as PESA, or NSGA-II, are the Pareto based selection strategy that decide survivors using dominance among individuals. However, does the Pareto based selection strategy always succeed in finding the Pareto optimal solutions? This paper shows a very simple example that can cause serious troubles on the Pareto based MOEAs. In such an instance, various solutions, which are apart from the true Pareto-optimums, are left as hardly-dominated solutions. We define such solutions as dominance resistant solutions (DRSs), and show a class of problems which produces DRSs easily. To cope with this difficulty we propose α-domination strategy that relaxes the domination introducing weak trade-off among objectives. With α-domination strategy, the DRSs are effectively purged from the population.

1 Introduction

Multi-objective problems have been investigated in many communities with various motivations and approaches. In recent days, some modern *multi-objective evolutionary algorithms* (MOEAs) have shown excellent performance for test problems proposed in [Deb 1998]. Pareto Archived Evolution Strategy (PAES) [Knowles and Corne 1999], Strength Pareto Evolutionary Algorithm (SPEA)[Zitzler and Thiele 1999], Pareto Envelope-based Selection Algorithm (PESA)[Corne *et al.* 2000], and (Fast Elitist) Non-dominated Sorting Genetic Algorithm (NSGA-II)[Deb *et al.* 2000], are such sophisticated MOEAs.

These MOEAs are based on two common ideas, one is the non-dominated strategy and the other is the crowding based diversity maintenance. They consider non-dominated solutions as proper candidate solutions to be kept. If the population or the archive gets full, some of the non-dominated solutions are killed based on another criterion. That is, the solution in the most crowded area is considered to be useless. These ideas seems very reasonable. However, if there exist hardly-dominated but clearly worse solutions, and they are spreaded in the population, the Pareto-based strategies fail in purging such solutions. This paper shows a very simple example of such a case, and discusses how to treat this difficulty.

The organization of this paper is as follows. Section 2 overviews basic differences between Pareto-based algorithms and other MOEAs, and show a superiority of the Pareto-based algorithms. In Section 3, we define a class of solutions that cause serious trouble on the Pareto-based algorithms, and a simple example is shown, further the instance is generalized. In Section 4, we introduce an idea to cope with this difficulty. Finally Section 5 is the conclusion and future work.

2 Comparison of a Pareto-based and Non-Pareto-based MOEAs

2.1 Pareto-based and Non-Pareto-based MOEAs

The main idea of the Pareto-based MOEAs is that non-dominated solutions in a population have the advantage of being survivors. For usual problems, the Pareto-based MOEAs are considered to be superior to Non-Pareto-based MOEAs at least to the extent of searching Pareto optima.

In this paper, we employ a simple Pareto-based MOEA called PbEA for experiments. PbEA is inspired by PESA, but it is simplified only to have the population (archive) size N as a parameter so as to clarify the characteristics of modern Pareto-based MOEAs. No technique to speed up search is employed. The algorithm of the PbEA is as follows.

1. Initial population of N individuals are randomly created.

2. A pair of parents are selected randomly, and a candidate solution is generated applying the crossover or the mutation to the parents.

3. If the candidate dominates an individual in the population, replace it with the candidate, and go to Step 2.

4. If the candidate is dominated by an individual in the population, go to Step 2.

5. If an individual in the population is dominated by another, replace it with the candidate, and go to Step 2.

0-7803-6657-3/01/$10.00

6. (N+1 solutions are on the Pareto front.) Replace the most crowded solution with the candidate.

7. If the termination criterion holds, stop the algorithm, otherwise go to Step 2.

There also have been proposed various algorithms which don't employ the concept of domination. We call such algorithms Non-Pareto-based MOEAs. The Predator Prey Model (PPM) approach [Laumanns *et al.* 1999] is one of such MOEAs. In the PPM, there are N preys (solutions) and M predators. Each predator has its own preference over a single objective function, and kills the worst prey it can reach. Clearly, preys that are on the Pareto front may be killed by a predator, though its probability will be lower. Thus the PPM searchs the solutions of the multi-objective optimization problems.

2.2 Comparative Experiments

We show the behavior of the PbEA and two types of the PPM proposed in [Laumanns *et al.* 1999]. For fair comparison, the same configurations of the PPM and the test problem with [Laumanns *et al.* 1999] is used. The following test problem is used.
Test-1 :

$$\underset{x,y}{\text{minimize}} \quad g_1, g_2$$

$$g_1(x,y) = -10\exp(-0.2\sqrt{x^2+y^2}\,)$$
$$g_2(x,y) = |x|^{4/5} + |y|^{4/5} + 5\,(\sin^3 x + \sin^3 y)$$
$$-50 \le x \le 50, \qquad -50 \le y \le 50.$$

For the PbEA, the Gaussian mutation with the constant stepsize same with the PPM and no crossover is used. Population size $N = 100$, which is fewer than 900 in the PPM.

Figure 1 shows solutions attained by the MOEAs. The PPM with the constant stepsize (Fig.1(a)) keeps a number of good solutions, but it also holds many non-Pareto solutions. Even if we extend the maximum generation, this situation doesn't change. Figure 1(b) shows solutions of the PPM with the decreasing stepsize. Though solutions closer to the Pareto front are obtained, solutions are divided into several clusters, and if we extend generation they converge to almost same points.

Figure 1(c) shows the solutions by the PbEA. We can see two advantages of the PbEA over the PPM : 1) the PbEA holds very few non-Pareto solutions, and 2) the PbEA keeps wider Pareto front stably. As shown in this example, Pareto-based MOEAs perform well on usual problems.

3 Dominance Resistant Solutions and Failure of Pareto-based MOEAs

Researchers of MOEAs often consider that solutions sampled by MOEAs are classified into two types, i.e.,

1. solutions that were/will be dominated by other solutions before too long, and

2. solutions that are hardly dominated, and are (near to) the Pareto optima.

In other words, we are obsessed with the idea that 'hardly-dominated' means that they are good solutions.

However, as Kita *et al.* have pointed out [Kita *et al.* 1996], some solutions in a certain problem survive over very long generations though they are fairly inferior qualitatively, because solutions that can dominate these troubling solutions are scarcely found.

We define such solutions as Dominance Resistant Solutions (DRSs). That is,

- DRSs are extremely inferior to others in at least one objective, and therefore they are apart from Pareto optima.

- However, solutions that dominate a DRS are scarcely found.

Currently, the above definition of DRSs is not sufficiently clear in mathematical sense. It is because we want discuss DRSs from the viewpoint of practical usage of MOEAs.

If DRSs occupy a large area in the search space than that of the true Pareto optima, the Pareto-based algorithms will fail in finding the good approximation of the Pareto optima. That is, all non-dominated solutions are treated as survivor on these MOEAs, and DRSs spread out over the search space and remain long. Thus, solutions near to Pareto optima are never intensively searched.

3.1 A Simple Example of DRSs

DRSs and failure of the Pareto-based MOEAs are observed in a very simple 2-variable and 3-objective problems,
Test-2 :

$$\underset{x,y}{\text{minimize}} \quad f_1, f_2, f_3$$

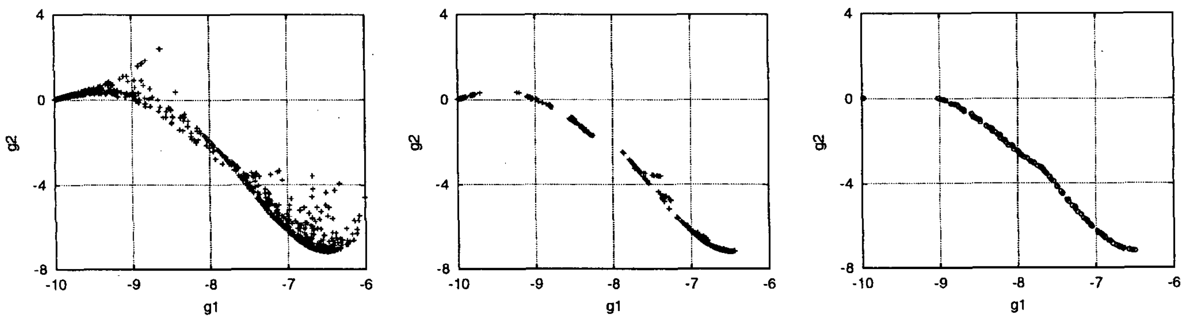
$$f_1(x,y) = x^2$$
$$f_2(x,y) = (x-20)^2$$
$$f_3(x,y) = y^2$$
$$-50 \le x \le 50, \qquad -50 \le y \le 50.$$

It is easily confirmed that the true Pareto-optima are the line segment :

$$(x,y) = (x,0), \ \ 0 \le x \le 20.$$

In Test-2, for example, $(x_A, y_A) = (8, 40)$, $(f_1, f_2, f_3) = (64, 144, 1600)$ is apart from Pareto-optimal.
However, $(x_B, y_B) = (12, 6)$, $(f_1, f_2, f_3) = (144, 64, 36)$, which is nearer to Pareto-optimal, never dominate (x_A, y_A).

In fact, solutions that dominate $(8, 40)$ are restricted to the points on the line segment $(x,y) = (8,y), -40 < y < 40$



(a)PPM with the constant stepsize (b)PPM with the decreasing stepsize (c) PbEA

Figure 1: Approximation of the Pareto set of Test-1 with the MOEAs.

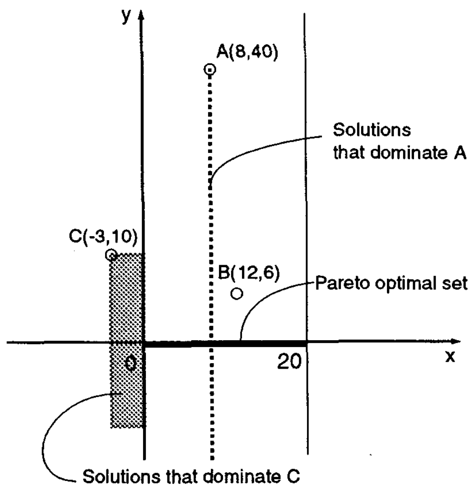


Figure 2: Variable space of Test-2. C is easily dominated, but A is hardly dominated by other solutions.

(Fig.2). In real-coded algorithms with the Gaussian mutation, such solutions on the line will never be found. In binary coded algorithms with the bit-flip mutation, such solutions can be found with unilateral alternation of a single variable. However, given the rotated problem of the Test-2, the binary coded algorithm also will fail.

Thus, for problem Test-2, solutions in the area $0 < x < 20,\ -50 \leq y \leq 50$ are similarly never dominated. This means that in Test-2, about 20% of the search space are DRSs.

3.2 Behavior of MOEAs on Test-2

So as to show seriousness of the failure of the Pareto-based MOEAs in such problems, we try to solve Test-2 with the PPM and the PbEA. All configurations are same as the previous experiments.

Figure 3(a) shows that, with PPM all solutions gathers near to the Pareto optimal variable values, $(x, y) = (x, 0),\ 0 < x < 20$, though some roughness and divisions are observed. By contrast, Fig.3(b) shows miserable behavior of the PbEA. As is predicted, solutions spread over the area $0 \leq x \leq 20,\ -50 \leq y \leq 50$. The PbEA fails in searching good solutions.

3.3 A Class of Problems which Produce DRSs Easily

In this section, we pick up a class of multi-objective optimization problems, and show the class is hard for Pareto-based algorithms because of DRSs.

Let P_1 be the K-objective optimization problem defined in the n-dimensional real space $X = R^n$, and let $f_1, ..., f_K$ be the objective functions to be minimized. Let X^* be the Pareto optimal set of P_1. We also consider another L-objective problem P_2 defined in the m-dimensional real space $Y = R^m$, and let $g_1, ..., g_L$ be the objective functions to be minimized, and Y^* be the Pareto optimal set of P_2.

Now, concatinating P_1 and P_2, consider the $(n+m)$-variables, $(K+L)$-objective problem P_3:

$$\underset{x_1,\dots,x_n,y_1,\dots,y_m}{\text{minimize}} \quad o_1, ..., o_{K+L}$$

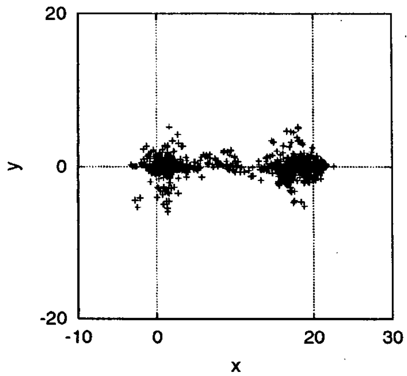
$$o_i := \begin{cases} f_i(x_1, ..., x_n) & \text{if } 1 \leq i \leq K, \\ g_{i-K}(y_1, ..., y_m) & \text{if } K+1 \leq i \leq K+L. \end{cases}$$

We call multi-objective optimization problem of this type *block-separable problem.* Test-2, the previous sample belongs to this class. Now let us denote $V = X \times Y$ and V^* be the search space and the Pareto optimal set of P_3 respectively.

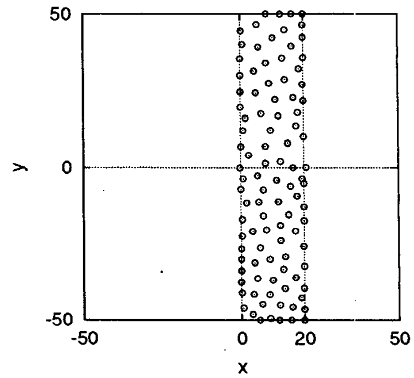
Prop. The Pareto optimal set of P_3 is $X^* \times Y^*$. (Proof is in Appendix)

Next we show a fact which causes DRSs seriously, that solutions in $X^* \times Y$ are hardly dominated by other solutions in P_3. Let $x^* \in X^*$, $y^- \in Y \setminus Y^*$, and define two sets $X^0 \subset X$ and $Y^+ \subset Y$ as follows.

$$X^0 := \{x \in X \mid \forall i,\ f_i(x) = f_i(x^*)\}$$



(a)PPM with the constant stepsize



(b) PbEA

Figure 3: Behaviors of the MOEAs for Test-2. Shown in variable(decision) space.

$$Y^+ := \{y \in Y \mid y \prec y^-\}$$

where relation '$\prec$' represents domination. That is, $a \prec b$ means a dominates b.
Theorem $(x, y) \prec (x^*, y^-) \Leftrightarrow (x, y) \in X^0 \times Y^+$.
Proof
$\Leftarrow$) trivial by definitions.
$\Rightarrow$) As $\forall i,\ o_i(x, y) \leq o_i(x^*, y^-),\ \ \forall i,\ f_i(x) \leq f_i(x^*)$.
On the other hand, $\forall i,\ f_i(x^*) \leq f_i(x)$ because $x^* \in X^*$.
In consequence, $\forall i,\ f_i(x^*) = f_i(x)$, and here $x \in X^0$.
Now the following equations are obtained,

$$1 \leq \forall i \leq K,\ o_i(x, y) = o_i(x^*, y^-) \quad (1)$$

$$1 \leq \forall i \leq K + L,\ o_i(x, y) \leq o_i(x^*, y^-) \quad (2)$$

$$1 \leq \exists i \leq K + L,\ o_i(x, y) < o_i(x^*, y^-). \quad (3)$$

From (1) and (3), we can conclude

$$K + 1 \leq \exists i \leq K + L,\ o_i(x, y) < o_i(x^*, y^-). \quad (4)$$

Consequently from (2) and (4), $y \prec y^-$ because $o_i(x, y) = g_{i-K}(y)$ for $K + 1 \leq i \leq K + L$. Then $y \in Y^+$, finally we get $(x, y) \in X^0 \times Y^+$.

Usually, the measure of X^0 is zero in X. Therefore the measure of $X^0 \times Y^+$ is also zero in V. So $(x^*, y) \in X^* \times Y$ are hardly dominated. Compared with $X^* \times Y$, the measure of $X^0 \times Y^+ \subset X^* \times Y$ is usually zero measured in $X^* \times Y$. On the other hand $X^* \times Y$ occupies wider area in V. This fact means that large portion of solutions of block-separable instances behave as DRSs (Figure 4), as observed in the previous example. As block-separable problems are easily considered and not so special, we believe the DRSs is serious problem for modern MOEAs.

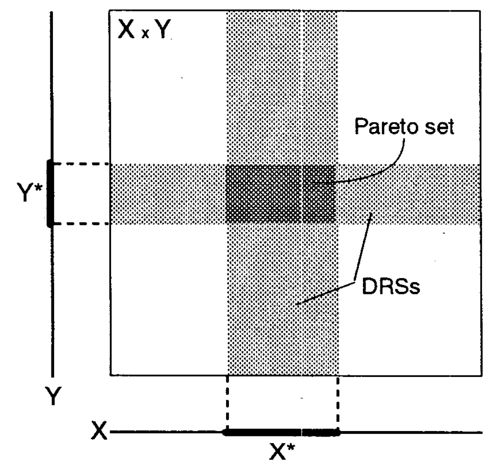


Figure 4: When $X^* \subset X$ is the Pareto set of P_1, and $Y^* \subset Y$ is of P_2, $X^* \times Y^*$ is the Pareto set of P_3. Solutions in $X^* \times (Y \setminus Y^*)$ and $(X \setminus X^*) \times Y^*$ are hardly dominated.

4 α-domination Strategy

4.1 Concept of α-domination

Difficulties of DRSs are raised from the strict definition of the domination. α-domination is a relaxation of the strict domination taking weak trade-off among objectives into consideration. Similar extensions has been discussed in, e.g. [Yu *et al.* 1974a] [Yu 1974b].

The fundamental idea of α-domination is setting upper/lower bounds of tradeoff rates between two objectives. While it is very difficult to assess exact tradeoff rate between two objectives, it will not so hard to assess lower(or upper) bound of substitution between two objectives, i.e. to set $\alpha_{ij} \geq 0$ such that *unit gain in f_j is at least worth of α_{ij}*

gain in f_i. For example, it is diffuicult to assess precise exchange rate of two currencies such as U.S. dollars and Euro, but we can say 1 dollar has at least the value of 0.1 Euro, and 1 Euro has at least the value of 0.1 dollar.

Let α_{ij} and α_{ij} be upper and lower bounds of the tradeoff rate between f_i and f_j respectively, i.e.

$$\alpha_{ji} \le \frac{\Delta f_i}{\Delta f_j} \le \frac{1}{\alpha_{ij}},$$

where Δf_i and Δf_j are equivalent change of f_i and f_j.

4.2 Definition, Corollary and Significance of α-domination

Consider K-objective minimization problem
$f_1(x), ..., f_n(x) \quad x \in X$. Let

$$g_i(x,y) := f_i(x) - f_i(y) + \sum_{j \ne i}^{1..K} \alpha_{ij}(f_j(x) - f_j(y)) \quad x, y \in X$$

$$(N.B. \; g_i(x,y) := \sum_{j=1}^{K} \alpha_{ij}(f_j(x) - f_j(y)) \text{ with } \alpha_{ii} = 1.)$$

Def.1. a solution x α-dominates a solution y (denoted by $x \overset{\alpha}{\prec} y$) $\Leftrightarrow \forall i \; g_i(x,y) \le 0$, and $\exists i \; g_i(x,y) < 0$.

When $\alpha_{ij} \equiv 0 \; \forall(i \ne j)$, definition of α-domination equals to that of domination. α-domination permits x to dominate y if x is slightly inferior to y in an objective but largely superior to y in some other objectives.

Def.2. $x \in X$ is α-Pareto optimal
$\Leftrightarrow \{y \in X \mid y \overset{\alpha}{\prec} x\} = \emptyset$.

About these new concepts, following proper corollaries are composed. (Proofs are in Appendix)

Cor.1. $x \prec y \Rightarrow x \overset{\alpha}{\prec} y$.

Cor.2. The α-Pareto optimal set $X^{\alpha} = \{ x \in X \mid x$ is α-Pareto optimal $\}$ is a subset of the Pareto optimal set X^*.

Replacing the strict domination used in Pareto-based MOEAs with α-domination, we will be able to search the α-Pareto optimal set instead of the Pareto optimal set.

α-domination will bring two benefits. Firstly, seriously worse solutions of DRSs can not survive any longer in α-Pareto based MOEAs. Additionally, we can gain fine-drawn candidates for decision making. Solutions that α-Pareto based MOEAs exclude, Pareto optimal but not α-Pareto optimal solutions, are extreme solutions which may not be selected by human as the final solution.

4.3 Numerical Experiments

To show α-domination can cope with DRS-problems, we carried out an experiment on Test-2 by the PbEA with α-domination. Parameters α_{ij} were set to const c in all $i \ne j$. $c = \frac{1}{3}, \frac{1}{9}, \frac{1}{100}$ are used.

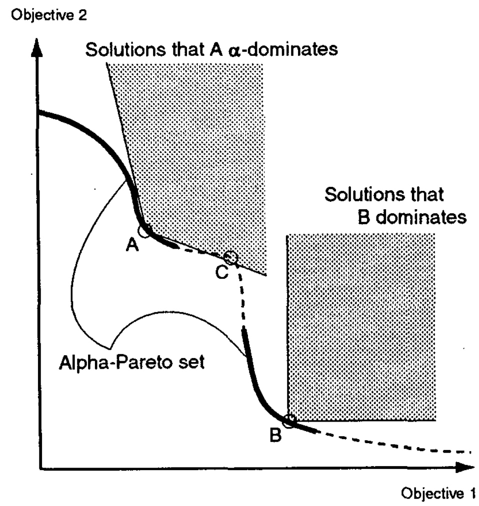


Figure 5: α-domination and strict domination in a minimization problem. C is Pareto optimal but is not α-Pareto optimal, because C is α-dominated by A.

In Fig.6, we can see the sound behavior of the PbEA with α-domination. We observe some tradeoff in performance of the MOEA with the parameter α. If we use larger α_{ij} (Fig.6(a)), solutions gathers closer to the Pareto front, but some region of the Pareto front are not covered. If we use the smaller α (Fig.6(c)), the wider Pareto front is found, however, solutions far from the Pareto front remains more because to dominate the DRSs gets difficult. Of course when c decreases to 0, search with the PbEA fails as shown at Fig.3(b).

5 Conclusion and Future work

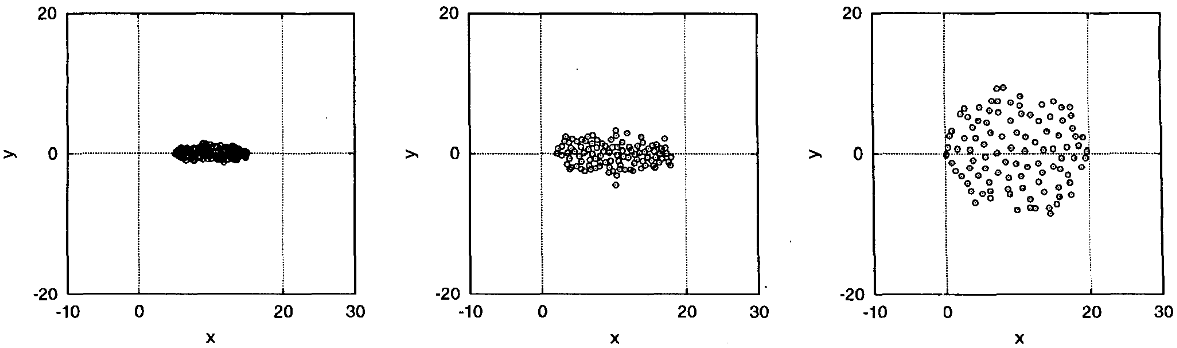
In this paper failure of modern Pareto-based multi-objective evolutionary algorithms (MOEAs) caused by dominance resistant solutions (DRSs) is shown. As an idea to cope with DRSs, α-domination strategy is proposed. The following two subjects should be studied as future work to show importance of the DRSs: 1) Do other/bigger classes of DRSs-problem exist? 2) Does DRSs really occurs and are serious in real world problems?

About these points, we consider that problems with many objectives may have DRSs easily, as simply it gets difficult for a solution to be improved simultaneously, and messy objectives likely cause block-separation discussed in Section 3.3.

Appendix: Proof of Proposition and Corollaries

Prop. The Pareto optimal set of P_3 is $X^* \times Y^*$.
Proof

Let $x^- \in X \setminus X^*$ and $y^0 \in Y$.
As $x^- \notin X^*$, $\exists x^+ \in X$ s.t. $x^+ \prec x^-$.
That is, $\forall i, \; f_i(x^+) \le f_i(x^-)$, and $\exists i, \; f_i(x^+) <$



(a) $\alpha = \frac{1}{3}$, $(x, 0)\ 5 \le x \le 15$ (b) $\alpha = \frac{1}{9}$, $(x, 0)\ 2 \le x \le 18$ (c) $\alpha = \frac{1}{100}$, $(x, 0)\ \frac{20}{101} \le x \le \frac{2000}{101}$

Figure 6: Behaviors of the PbEA with α-domination on Test-2 , shown in variable space. Captions show the α-Pareto optimal set of each α.

$f_i(x^-), 1 \le i \le K$.

Here $x^-, y^0 \notin V^*$, to be dominated by $x^+, y^0 \in V$ because

$$1 \le \forall i \le K,\ o_i(x^+, y^0) \le o_i(x^-, y^0)$$
$$1 \le \exists i \le K,\ o_i(x^+, y^0) < o_i(x^-, y^0)$$
$$K+1 \le \forall i \le K+L,\ o_i(x^+, y^0) = o_i(x^-, y^0).$$

Therefore, $((X \setminus X^*) \times Y) \cap V^* = \emptyset$. In the same way, we obtain $(X \times (Y \setminus Y^*)) \cap V^* = \emptyset$.

Next we show $(x^*, y^*) \in V^*$ where $x^* \in X^*,\ y^* \in Y^*$. Assuming $(x^*, y^*) \notin V^*$,

$$\exists (x^+, y^+) \in V,\ (x^+, y^+) \prec (x^*, y^*).$$

In this time, $1 \le \exists i \le K+L,\ o_i(x^+, y^+) < o_i(x^*, y^*)$. However, if $1 \le i \le K$, it means $f_i(x^+) < f_i(x^*)$, this contradicts to $x^* \in X^*$. Similarly if $K+1 \le i \le K+L$, this contradicts to $y^* \in Y^*$. Now the assumption is rejected.

By these arguments, $V^* = X^* \times Y^*$ is shown.

Cor.1. $x \prec y \Rightarrow x \overset{\alpha}{\prec} y$.

Proof

From definition of domination, $\forall i \quad f_i(x) - f_i(y) \le 0,\ \exists i\ f_i(x) - f_i(y) < 0$.

As $\forall i, j \quad \alpha_{ij} \le 0$ and $f_j(x) - f_j(y) \le 0$,

$$\forall i \quad \sum_{j \ne i}^{1..K} \alpha_{ij}(f_j(x) - f_j(y)) \le 0.$$

Therefore, $\forall i\ g_i(x, y) \le 0,\ \exists i\ g_i(x, y) < 0$ when $f_i(x) - f_i(y) < 0$. Thus $x \overset{\alpha}{\prec} y$.

Cor.2. The α-Pareto optimal set $X^\alpha = \{\ x \in X \mid x$ is α-Pareto optimal $\}$ is a subset of the Pareto optimal set X^*.

Proof

Let x be an α-Pareto optimal.

Then $\{\ y \in X \mid y \overset{\alpha}{\prec} x\} = \emptyset$.

If $w (\in X) \prec x,\ w \overset{\alpha}{\prec} x$ (*Cor.1.*).

So $\{\ w \in X \mid w \prec x\} \subset \{w \in X \mid w \overset{\alpha}{\prec} x\} = \emptyset$, i.e. x is a Pareto optimal.

Bibliography

[Deb 1998] K. Deb : Multi-Objective Genetic Algorithms : Problem Difficulties and Construction of Test Problems. *Technical Report CI-49/98, Department of Computer Science, University of Dortmund.*

[Deb *et al.* 2000] K. Deb, S. Arawal, A. Pratap, and T. Meyarivan : A Fast Elitist Non-dominated Sorting Genetic Algorithm for Multi-objective Optimization : NSGA-II, *Parallel Problem Solving from Nature(PPSN)6, pp. 849-858 (2000)*

[Corne *et al.* 2000] D. W. Corne, J. Knowles, M. J. Oates : The Pareto Envelope-Based Selection Algorithm for Multiobjective Optimization, *PPSN6*, pp. 839-848 (2000)

[Knowles and Corne 1999] J. Knowles, D. W. Corne : The Pareto Archived Evolution Strategy : A New Baseline Algorithm for Pareto Multiobjective Optimisation *Congress on Evolutionary Computation(CEC)* , pp. 98-105 (1999)

[Zitzler and Thiele 1999] E. Zitzler, L. Thiele : Multiobjective Evolutionary Algorithms : A Comparative Case Study and the Strength Pareto Approach, *IEEE Transactions on Evolutionary Computation,* 2(4), 257-272 (1999)

[Laumanns *et al.* 1999] M. Laumanns, G. Rudolph, and H.P. Schwefel : A Spatial Predator-Prey Approach to Multi-objective Optimization : A Preliminary Study, *PPSN5*, pp. 241-249 (1999)

[Kita *et al.* 1996] H. Kita, Y. Yabumoto, N. Mori and Y. Nishikawa: Multi-Objective Optimization by Means of the Thermodynamical Genetic Algorithm, *PPSN4*, pp. 504-512 (1996).

[Yu *et al.* 1974a] P.L.Yu et al. : Compromise Solutions, Domination Structures and Sulukvadze's Solution, *J. Opt. The. and App.*, 13-3 362/378 (1974).

[Yu 1974b] P.L.Yu: Cone Convexity, Cone Extreme Points and Nondominated Solutions in Decision Problems with Multiobjectives, *J. of Opt. Th. and App.*, 14-3, 319/376 (1974).