The Convergence of a Multi-objective Evolutionary Algorithm Based on Grids

Yuren Zhou[1] and Jun He[2]

[1] School of Computer Science and Engineering,
South China University of Technology,
Guangzhou 510640, China
zhouyuren@hotmail.com
[2] School of Computer Science, The University of Birmingham,
Birmingham B15 2TT, UK

Abstract. Evolutionary algorithms are especially suited for multi-objective optimization problems. Many evolutionary algorithms have been successfully applied to various multi-objective optimization problems. However, theoretical studies on multi-objective evolutionary algorithms are relatively scarce. This paper analyzes the convergence properties of a simple pragmatic (μ+1)-MOEA. The convergence of MOEAs is defined and the general convergence conditions are studied. Under these conditions, it is proven that the proposed (μ+1)-MOEA converges almost surely to the Pareto-optimal front.

1 Introduction

Evolutionary algorithms (EAs), which adopt a population-based search, are especially suited for multi-objective optimization problems with several conflicting objectives [1-2]. EAs can search multiple objectives simultaneously and always keep the better solutions to next generation. Multi-Objective Evolutionary Algorithms (MOEAs) have been studied for more than ten years. It is generally recognized that Schaffer [3] was the first researcher to use EAs to handle vector optimization problems. Today various MOEAs, e.g., NSGA [4], SPEA [5], PAES [6] and NSGA-II [7], have been proposed and applied in many practical fields. Compared with a great amount of theoretical study on single-objective evolutionary algorithms [8-11], rigorous analysis of MOEAs is still in its infant phase, and attracts little attention from researchers [12-16]. Up to today only a few theoretical results on MOEAs have been obtained. Rudolph and Agapie [13] analyzed and proved MOEAs' convergence using Markov chain. Hanne [14] proposed a convergence theorem of function MOEAs with probability 1 under strict condition of "efficiency preserving", a requirement that some current pragmatic MOEAs do not meet. Laumanns [15] established a MOEA model which have both properties of converging to the Pareto-optimal front and maintaining a spread among obtained solutions, but he failed to rigorously define the convergence of MOEAs and prove that the MOEA model converges to the Pareto optimal sets. Recently Laumanns [16] presented the running time analysis of multi-objective EAs on pseudo-Boolean model problems.

This paper aims to discuss convergence of function MOEAs. We will introduce the rigorous definition of strong and weak convergences of MOEAs and discuss general

L. Wang, K. Chen, and Y.S. Ong (Eds.): ICNC 2005, LNCS 3611, pp. 1015 – 1024, 2005.

conditions that guarantee the convergence of MOEAs. Under these conditions, we show that the proposed MOEA converges almost surely to the Pareto optimal set.

The remainder of this paper is arranged as follows: Section 2 describes the $(\mu+1)$ MOEA and introduces some basic definitions and terms; Section 3 analyzes the proposed MOEA's convergence; Section 4 concludes the paper.

2 Definitions and Algorithm Description

Without loss of generality, consider following multi-objective optimization problem with n decision variables and m objectives:

(MOP) Maximize $\mathbf{y} = \mathbf{f}(\mathbf{x}) = (f_1(x_1,...,x_n), \ldots, f_m(x_1,...,x_n))$ (1)

Subject to $g_i(\mathbf{x}) \le 0$, i=1,…,q.

Where $\mathbf{x}=(x_1, \ldots, x_n) \in X \subset R^n$, $\mathbf{y}=(y_1,\ldots,y_m) \in Y \subset R^m$, $\mathbf{x}$ is the decision (parameter) vector, X is the decision space, $\mathbf{y}$ is the objective vector and Y is the objective space, g($\mathbf{x}$) is the constraint condition. This paper deals with non-constraint problems only.

2.1 Dominance Relation

Different from fully ordered scalar search spaces, multidimensional search spaces are only partially ordered, i.e., two different solutions are related to each other in two possible ways: either one dominates the other or none of them is dominated. Firstly let's introduce two basic definitions used in MOEAs: dominance relation and Pareto set.

Definition 1: Let $\mathbf{f}, \mathbf{g} \in R^m$, vector $\mathbf{f}$ is said to dominate vector $\mathbf{g}$ (written as $\mathbf{f} \succ \mathbf{g}$) if and only if

1) $\forall\, i \in \{1,...,m\}: f_i \ge g_i$;

2) $\exists\, j \in \{1,...,m\}: f_i > g_i$

if $\mathbf{f} \succ \mathbf{g}$, it also means that $\mathbf{g}$ is dominated by $\mathbf{f}$, denoted as $\mathbf{g} \prec \mathbf{f}$.

Definition 2: Let $F \subset R^m$ be a vector set, the set of vectors in F that are not dominated by any vector in F is called Pareto set of F, denoted as P(F), i.e., P(F) : = $\{\mathbf{g} \in F \mid \neg\exists \mathbf{f} \in F : \mathbf{f} \succ \mathbf{g}\}$.

Based on the above notation, now we define Pareto-optimal front, Pareto-optimal solution and Pareto-optimal set as follows.

Definition 3: Let $R_\mathbf{f}$ be the range of function $\mathbf{f}$ in MOP (1), the Pareto set of $R_\mathbf{f}$ is called Pareto-optimal front. That is P($R_\mathbf{f}$)=$\{\mathbf{y} \in R_\mathbf{f} \mid \neg\exists \mathbf{y}' \in R_\mathbf{f} : \mathbf{y}' \succ \mathbf{y}\}$.

Definition 4: Let P($R_\mathbf{f}$) be the Pareto-optimal front of MOP (1), the image source of P($R_\mathbf{f}$) under mapping $\mathbf{f}$ is said to be Pareto-optimal set, denoted as P($\mathbf{f}(\mathbf{x})$), i.e. P($\mathbf{f}(\mathbf{x})$) := $\{\mathbf{x} \in X \mid \neg\exists \mathbf{x}' \in X : \mathbf{f}(\mathbf{x}') \succ \mathbf{f}(\mathbf{x})\}$.

The vector in P(**f**(**x**)) is called Pareto-optimal solution.

The concept of ε-neighborhood, which is useful in discussing convergence of MOEA, is defined as follows.

Definition 5: Let **f**=$(f_1, f_2, \ldots, f_m) \in R^m$, ε> 0, the ε-neighborhood $N_\varepsilon(\mathbf{f})$ of **f** is defined as follows:

$$N_\varepsilon(\mathbf{f}) := \{ \mathbf{y} \in R^m \mid \mathbf{y}=(y_1, y_2, \ldots, y_m),\ y_i \in (f_i - \varepsilon, f_i + \varepsilon),\ i=1,\ldots,m\}$$

In fact, $N_\varepsilon(\mathbf{f})$ is an m–dimension hyper-box centered on **f** in R^m.

Let $F \subset R^m$, the union of the vectors' ε-neighborhood in F is said to be F's ε-neighborhood, denoted as $N_\varepsilon(F)$, i.e.,

$$N_\varepsilon(F) = \bigcup_{\mathbf{f} \in F} N_\varepsilon(\mathbf{f}).$$

2.2 (μ +1)-MOEA

Since there are more than one objectives to be optimized simultaneously in multi-objective optimization, the solution is no longer a single optimal point, rather a whole set of possible solutions with equivalent quality, i.e., Pareto-optimal set. This determines that the task faced by MOEAs is also two-objective:

To guide the search towards the Pareto-optimal set;

To maintain a diverse population in order to achieve a well distributed trade-off front.

Researchers have developed several MOEAs to implement the above tasks. One of them is the MOEA based on grids, developed mainly by Knowles [6] and Laumanns [15]. Based on ε -dominance concept (or grid), Deb [17] proposed a steady-state MOEA that had a good compromise in terms of convergence near to Pareto-optimal front, diversity of solutions and computational time. The basic idea of this MOEA is to divide the search space into a number of grids (or hyper-boxes) and to maintain the diversity by ensuring that a hyper-box can be occupied by only one solution. Many MOEAs (including the algorithm developed in [17]) make use of two co-evolving populations: an EA population and an archive population. This kind of MOEA is more difficult to analyze and we don't discuss it in this paper.

In this paper we only discuss a simple (μ +1) MOEA for MOP(1) based on grids introduced in [15][6]. This (μ +1) MOEA is composed of algorithm 1-3 where the details of Algorithm 1, Algorithm 3 and the notations can be found in [15]. It is similar to (μ +1) evolution strategy [18] and uses only one population.

The (μ +1) MOEA is described as follows:

Algorithm 1: Iterative search algorithm
t := 0
$A^{(0)} := \varnothing$
while terminate ($A^{(t)}$,t) = false **do**
  t := t+1

$\mathbf{f}^{(t)}$:= generate($A^{(t-1)}$)
$A^{(t)}$:= update($A^{(t-1)}, \mathbf{f}^{(t)}$)
end while
Output: $A^{(t)}$

Algorithm 1 is the general framework of iterative searching algorithm. The integer t denotes the evolution generation. The set $A^{(t)}$ is the population of objective space at generation t with a dynamic size. The vector $\mathbf{f}^{(t)}$ in objective space is the new individual yielded by using "generate" function. "Update" function is used to produce the next population from current population and new individual.

Algorithm 2: "generate" function
Input: $A^{(t-1)} = \{ \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_\mu \}$
$(A^{(t-1)})^{-1} := \{ \mathbf{f}^{-1}(\mathbf{a}_1), \mathbf{f}^{-1}(\mathbf{a}_2), \ldots, \mathbf{f}^{-1}(\mathbf{a}_\mu) \}$
$:= \{ \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_\mu \}$
$\mathbf{x_i}$:= Random $\{ \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_\mu \}$
$\mathbf{x}'$:= mutate $\mathbf{x_i}$
$\mathbf{f}^{(t)} := \mathbf{f}(\mathbf{x}')$
Output: $\mathbf{f}^{(t)}$

The purpose of algorithm 2 is to generate a new individual. The process works in a simple way: a individual is randomly selected from a population and then variation is conducted to generate a new individual.

Algorithm 3: "update" function
Input: A, **f**
$D := \{ \mathbf{f}' \in A |\ \text{box}(\mathbf{f}) \succ \text{box}(\mathbf{f}') \}$
if $D \neq \emptyset$ **then**
$A' := A \bigcup \{\mathbf{f}\} \backslash D$
else if $\exists\, \mathbf{f}' : (\text{box}(\mathbf{f}') = \text{box}(\mathbf{f}) \wedge \mathbf{f} \succ \mathbf{f}')$ **then**
$A' := A \bigcup \{\mathbf{f}\} \backslash \{ \mathbf{f}' \}$
else if $\neg \exists\, \mathbf{f}' : \text{box}(\mathbf{f}') = \text{box}(\mathbf{f}) \vee \text{box}(\mathbf{f}') \succ \text{box}(\mathbf{f})$ **then**
$A' := A \bigcup \{\mathbf{f}\}$
else
$A' := A$
end if
Output: A'

Algorithm 3 is the "update" function. Let the objective space Y be $\{(f_1, f_2, \ldots, f_m) \mid a_i \leq f_i \leq b_i \ (i=1,\ldots,m)\}$, $\delta_1, \delta_2, \ldots, \delta_m$ be previously given positive real vector (the smaller the number δ_i (i=1,...,m), the higher precision of the algorithm). Firstly, the objective space is divided into a number of m-dimension hyper-boxes, each having δ_i size in the i-th objective. The number of hyper-boxes in objective space Y is less than or equal to $\prod_{i=1}^{m}([\frac{b_i - a_i}{\delta_i}]+1)$. The box dominance relation can be easily generalized from the vector dominance relation. The algorithm allows at most one solution to be present in each hyper-box and always maintains a set of non-dominated boxes. Thus it maintains the diversity in the population and forces the population to converge to Pareto-optimal front. It is an elitist approach and its population size μ changes dynamically.

3 Convergence Analysis of (μ +1) MOEA

Unlike the single objective optimization problem where its optimal value is a real number, Pareto-optimal front in MOPs is a set of points in R^m. It is necessary to give the rigorous definition of convergence of MOEAs. Recall the definition of convergence for random variable sequence { X_n , $n \geq 1$}, there are definitions such as convergence almost surely, convergence in probability and convergence in mean [19].

Definition 6: Let {X, X_n (n=1, 2, …) } be random variables on a probability space (Ω, F, P), the random variable sequence X_n is said to converge almost surely to random variable X, if

$$P\{\lim_{n\to\infty} X_n = X\} = 1;$$

converge in probability to X, if

$$\lim_{n\to\infty} P\{| X_n - X |\leq \in\} = 1, \ \forall \varepsilon >0;$$

converge in mean to X, if

$$\lim_{n\to\infty} P\{| X_n - X |\} = 0.$$

Both convergence almost surely and convergence in mean implies convergence in probability whereas the converse is wrong in general.

Lemma 1: The following statements are equivalent

Random variable sequence X_n converge almost surely to random variable X;

$$\forall \varepsilon > 0, \ \lim_{m\to\infty} P\{| X_n - X | < \varepsilon \text{, for all } n \geq m\} = 1;$$

$$\forall \varepsilon > 0, \ \lim_{m\to\infty} P\{| X_n - X | \geq \varepsilon \text{, for some } n \geq m \} = 0;$$

$\forall\varepsilon > 0$, $P\{|X_n - X| \geq \varepsilon$ appears infinite times$\} = 0$;

$\forall\varepsilon > 0$, $P\{\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}(|X_n - X| \geq \varepsilon)\} = 0$.

Proof: see [19].

With the equivalent definitions in Lemma 1 we can define the convergence of (μ+1)-MOEA as follows.

Definition 7: Let $A^{(n)}$ (n=1,2,…) be the sequence of populations generated by (μ+1)-MOEA. Objective space is divided into m-dimension hyper-boxes. Let B_i (i=1,…,s) denote the hyper-boxes that contain the Pareto front of MOP(1), and F_i (i=1,…,s) the Pareto-optimal front in B_i. F_i's ε-neighborhood is denoted as $N_\varepsilon(F_i)$ (see Definition 5).

The (μ+1)-MOEA is said to converge almost surely to the front of MOP(1) if
$\forall\varepsilon > 0$, $\lim_{m\to\infty} P\{A^{(n)} \cap N_\varepsilon(F_i) \neq \varnothing$ for all $n \geq m\} = 1$, $1 \leq i \leq s$;

The (μ+1)-MOEA is said to converge in probability to the front of MOP(1) if
$\forall\varepsilon > 0$, $\lim_{n\to\infty} P\{A^{(n)} \cap N_\varepsilon(F_i) \neq \varnothing\} = 1$, $1 \leq i \leq s$.

Similar to Lemma 1, we have the equivalent definition as follows.

Lemma 2: The following statements are equivalent
The (μ+1)-MOEA converge almost surely to the front of MOP(1);
$\forall\varepsilon > 0$, $\lim_{m\to\infty} P\{A^{(n)} \cap N_\varepsilon(F_i) = \varnothing$ for some $n \geq m\} = 0$ $(1 \leq i \leq s)$;
$\forall\varepsilon > 0$, $P\{A^{(n)} \cap N_\varepsilon(F_i) = \varnothing$ appears infinite times$\} = 0$ $(1 \leq i \leq s)$;
$\forall\varepsilon > 0$, $P\{\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}(A^{(n)} \cap N_\varepsilon(F_i) = \varnothing)\} = 0$ $(1 \leq i \leq s)$.

Lemma 3: (Borel-Cantelli) Let $\{X_n$, n=1,2,…$\}$ be event sequence, then

$$\sum_{n=1}^{\infty} P(X_n) < \infty \Rightarrow P\{\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty} X_n\} = 0.$$

Proof: See [19].

Based on above definitions and lemmas, we give the main MOEA convergence theorems as follows.

Theorem 1: Let $A^{(n)}$ (n=1,2,…) be the sequence of populations generated by (μ+1)-MOEA. $N_\varepsilon(F_i)$ (i=1,…,s) is the ε-neighborhood as specified in definition 7. Let

$$\alpha_n^i := P\{A^{(n+1)} \cap N_\varepsilon(F_i) = \varnothing \mid A^{(n)} \cap N_\varepsilon(F_i) \neq \varnothing\},\ n \geq 1,\ i=1,\ldots,s;$$

$\beta_n^i := P\{ A^{(n+1)} \cap N_\varepsilon(F_i) = \varnothing \mid A^{(n)} \cap N_\varepsilon(F_i) = \varnothing \}$, $n \geq 1$, i=1,…,s;
$\gamma_n^i := \beta_1 \times \beta_2 \ldots \times \beta_n$, $n \geq 1$, i=1,…,s.
Then

(1) If $\forall \varepsilon > 0$, $\lim_{n\to\infty} \gamma_n^i = 0$(i=1, …, s), then ($\mu$ +1)-MOEA converges in probability to the Pareto-optimal front;

(2) If $\sum_{n=1}^{\infty} \gamma_n^i < \infty$ (i=1, …, s), then (μ +1)-MOEA converges almost surely to Pareto-optimal front.

Proof: Firstly note that $\alpha_n^i = 0$ ($n \geq 1$, i=1, …, s) because (μ +1)-MOEA uses an elitist preserving approach.

(1) Given $i \in \{1,2,\ldots s\}$, let
$P^{(i)}(n) = P\{ A^{(n)} \cap N_\varepsilon(F_i) = \varnothing \}$.
According to Bayesian formula, we obtain

$$\begin{aligned} P^{(i)}(n+1) &= P\{ A^{(n+1)} \cap N_\varepsilon(F_i) = \varnothing \} \\ &= P\{ A^{(n+1)} \cap N_\varepsilon(F_i) = \varnothing \mid A^{(n)} \cap N_\varepsilon(F_i) \neq \varnothing \} P\{ A^{(n)} \cap N_\varepsilon(F_i) \neq \varnothing \} \\ &\quad + P\{ A^{(n+1)} \cap N_\varepsilon(F_i) = \varnothing \mid A^{(n)} \cap N_\varepsilon(F_i) = \varnothing \} P\{ A^{(n)} \cap N_\varepsilon(F_i) = \varnothing \} \\ &= \alpha_n^i P\{ A^{(n)} \cap N_\varepsilon(F_i) \neq \varnothing \} + \beta_n^i P(n) \\ &= \beta_n^i \ldots \beta_1^i P(1) \\ &= \gamma_n^i P(1) \end{aligned}$$

Because $\lim_{n\to\infty} \gamma_n^i = 0$, we get $\lim_{n\to\infty} P^{(i)}(n+1) = 0$.
Hence, according to Definition 6(2), (μ +1)-MOEA converges in probability to Pareto-optimal front.

(2) From $P^{(i)}(n+1) = \gamma_n^i P(1)$, $\sum_{n=1}^{\infty} \gamma_n^i < \infty$ and Lemma 3, we know that the (μ +1)-MOEA converges almost surely to Pareto-optimal front. □

Theorem 2: Let the (μ +1)-MOEA be used to solve Problem MOP(1). Assume that the decision space X in MOP(1) is a compact set in R^n, the objective function **f**(**x**) is continuous on X, and the variation in Algorithm 2 of the (μ +1)-MOEA is a Gaussian variation. Then (μ +1)-MOEA converges almost surely to the Pareto-optimal front of MOP(1) .

Proof: Given $i \in \{1, 2, \ldots, s\}$, $\forall \varepsilon > 0$, according to Theorem 1(2), it is sufficient to show that there exists a constant $c \in (0,1)$, $\beta_n^i \leq c$ (n=1, 2, …).

The Gaussian variation of the (μ +1)-MOEA is denoted as $\mathbf{x}'$:=**x**+**Z**, where **Z** ~ $N(0,\sigma^2 I_n)$ is a normally distributed random vector and I_n denotes the n-dimension unit matrix.

Let $\mathbf{y}_0$ be a point on Pareto front in B_i, where $\mathbf{y}_0$ = **f**($\mathbf{x}_0$), $\mathbf{x}_0$ =($x_0^1, x_0^2, \ldots x_0^n$)$\in$ X.

Because **f**(**x**) is continuous on X, there exists a positive r>0 such that when **x** satisfies $\|\mathbf{x}-\mathbf{x}_0\|_\infty \le r$ (here $\|\ \|_\infty$ is the maximum norm), then $\|\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{x}_0)\|_\infty \le \varepsilon$, therefore **f**(**x**) $\in N_\varepsilon(F_i)$.

Let $D_{\mathbf{x}_0,r}$:= { **x**$\in$ X | $\|\mathbf{x}-\mathbf{x}_0\|_\infty \le r$}, for **x** =($x^1, x^2, \ldots x^n$)$\in$ X, we have

$$P\{\mathbf{x}+\mathbf{Z}\in D_{\mathbf{x}_0,r}\} = \prod_{k=1}^{n}\int_{x_0^k-x^k-r}^{x_0^k-x^k+r}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{u^2}{2\sigma^2}}du.$$

Let $P_1(\mathbf{x},\mathbf{x}_0)$= P{**x**+**Z** $\in D_{\mathbf{x}_0,r}$ }, then 0< $P_1(\mathbf{x}, \mathbf{x}_0)$<1.

Because $P_1(\mathbf{x},\mathbf{x}_0)$ is continuous on compact set X, there exists $x_1', \mathbf{x}_0' \in$ X, such that

$P_1(\mathbf{x},\mathbf{x}_0) \ge P_1(\mathbf{x}_1',\mathbf{x}_0') = \min_{\mathbf{x},x_0\in X} P_1(\mathbf{x},\mathbf{x}_0)$, and 0< $P_1(\mathbf{x}_1',\mathbf{x}_0')$<1.

Therefore

$P\{A^{(n+1)} \cap N_\varepsilon(F_i) \ne \varnothing \mid A^{(n)} \cap N_\varepsilon(F_i) = \varnothing\} \ge P_1(\mathbf{x}_1',\mathbf{x}_0')$

$\beta_n^i = P\{A^{(n+1)} \cap N_\varepsilon(F_i) = \varnothing \mid A^{(n)} \cap N_\varepsilon(F_i) = \varnothing\}$

$\le 1 - P_1(\mathbf{x}_1',\mathbf{x}_0')$ = c, c$\in$ (0,1). □

From the proof of Theorem 2, we know that if the Gaussian variation is replaced by Cauchy variation [20], Theorem 2 is still valid.

Furthermore, we have the following more general convergence theorem for the (μ +1)-MOEAs:

Theorem 3: Except for the variation, let all other assumptions remain the same as Theorem 2. Let the random variation vector be **Z**=($z_1, z_2, \ldots, z_n$), where z_i (i = 1,2,…n) are the independently identically distributed random variables whose density function is $\varphi(x)$ .If $\varphi(x)$ satisfies:

(1) $\varphi(x)$ is continuous on R;

(2) $\forall a, b \in R$, a<b, $\int_a^b \varphi(x)dx$ >0.

Then the (μ +1)-MOEA converges almost surely to Pareto-optimal front.

Proof: Similar to that of Theorem 2.

4 Conclusions and Future Work

Compared with single-objective evolutionary algorithms, the design and analysis of MOEAs are much more complicated. This paper has investigated the convergence properties of a simple pragmatic (μ +1)-MOEA based on grids [6,15]. We have established the conditions that guarantee the convergence of the algorithm, and proved that the (μ +1)-MOEA using either Gaussian variation or Cauchy variation is convergent. In more general, the proposed MOEA is proved to be convergent under the assumption that the variation parameter in the algorithm remains constant.

However, the convergent conditions presented in the paper do not hold for the algorithms with self-adaptation variation, which are widely applied in evolution strategy to improve convergence speed. This is one of our future works. Like the analysis in single-objective evolutionary algorithms [21,22], the limit behavior, time complexity, and dynamical behavior of MOEAs are also important topics in our future research.

References

1. C. A. C Coello, D. A. Van Veldhuizen, and G. B. Lamont: Evolutionary algorithms for solving multi-objctive problems. Norwell, MA: Kluwer (2002)
2. K. Deb: Multi-Objective Optimization Using Evolutionary Algorithms. Chichester, U.K.: Wiley (2001)
3. J. D. Schaffer: Multiple objective optimization with vector evaluated genetic algorithms. In: Grefenstte J J (ed.): Proceeding of an international conference on genetic algorithms and their applications. Pittsburgh PA: Morgan Kaufmann Publishers(1985)93-100
4. N. Srinivas, K. Deb: Multi-objective function optimization use non-dominated genetic algorithms. Evolutionary Computation, Vol.2(3): (1994)221-248
5. E. Zitzler, L. Thiele: Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. IEEE Transaction on Evolutionary Computation, Vol.3(4): (1999)257-271
6. J. D. Knowles, D. W. Corne: Approximating the nondominated front using the Pareto archived evolution strategy. Evolutionary Computation. Vol.8(2), (2000)149-172
7. K. Deb, S. Agrawal, A. Pratap, T. Meyarivan: A fast and elitist multi-objective genetic algorithms: NSGA- Ⅱ . IEEE Transaction on Evolutionary Computation. Vol.6(2), (2002)182-197
8. H.-G.Beyer, H.-P.Schwefel, and I. Wegener: How to analysis evolutionary algorithms. Theoretical Computation Science, Vol.287, (2002)101-130
9. G.Rudolph: Convergence Properties of evolutionary algorithms. Kovac, Hamburg(1997)
10. A. Bienveüe, O. Francois: Global convergence for evolution strategies in spherical problems: some simple proofs and difficulties. Theoretical Computer Science, Vol.306 (2003)269-289
11. C. H. Guo, H. W. Tang: Global convergence properties of evolution strategies. Mathematica Numerica Sinica. Vol.23(1): (2001)106-110
12. G. Rudolph: On a multi-objective evolutionary algorithm and its convergence to the Pareto set. In: Proc. Fifth IEEE Conf. Evolutionary Computation, Anchorage AK (1998)511-516

13. G. Rudolph, A. Agapie: Convergence properties of some multi-objective evolutionary algorithms. In: Proceedings of the 2000 Congress on Evolutionary Computation (CEC2000), Piscataway NJ (2000)1010-1016
14. T. Hanne: On the convergence of multiobjective evolutionary algorithms. European Journal of Operational Research, Vol.117(3): (1999)553-564
15. M. Laumanns, L. Thiele, K. Deb, E. Zitzler: Combining Convergence and diversity in evolutionary multi-objective optimization. Evolutionary Computation. Vol.10(3): (2002)263-282
16. M. Laumanns, L. Thiele, and E. Zitzler: Running time analysis of multiobjcetive evolutionary algorithms on pseudo-boolean functions. IEEE Transaction on Evolutionary Computation, Vol.8, No.2, (2004)170-182
17. Deb K, Mohan M, Mishra S: A fast multi-objective evolutionary algorithm for finding well-spread Pareto-optimal solutions. Indian Institute of Technology Kanpur, KanGal Report: Number 2003002(2003)
18. H. P. Schwefel: Evolution and optimum seeking. John Wiley & Sons, New York(1995)
19. K. L. Chung: A course in probability. Academic Press, New York (1974)
20. X. Yao, Y. Liu, G. Lin: Evolutionary programming made faster. IEEE Transaction on Evolutionary Computation. Vol.3(2): (1999)82-102
21. J. He and X. Yao: From an individual to a population: An analysis of the first hitting time of population-based evolutionary algorithms. IEEE Transactions on Evolutionary Computation, Vol.6(5): (2002)495-511
22. J. He and X. Yao: Towards an analytic framework for analysing the computation time of evolutionary algorithms. Artificial Intelligence, Vol.145(1-2): (2003)59-97