TA2-5 10:50

Proceedings of the 2000 IEEE
International Conference on Control Applications
Anchorage, Alaska, USA • September 25-27, 2000

Modified Evolution Strategies with a Diversity-Based Parent-Inclusion Scheme

Tao-Yuan Huang* and Yung-Yaw Chen
Institute of Electrical Engineering
National Taiwan University, Taipei, Taiwan, ROC
hubert@ipmc.ee.ntu.edu.tw

***Abstract*—There are two major variants of evolution strategies: the (μ, λ) evolution strategy and the ($\mu + \lambda$) evolution strategy. In a survey paper [1], it is suggested that preference of either strategy is most likely problem-dependent, and that both strategies are special cases of the general (μ, κ, λ) evolution strategy, where κ denotes the maximum life span of an individual. In this paper, a diversity-based fuzzy parent-inclusion scheme is devised to make compromises between both strategies and thus retain advantages from each. Simulation results provided do manifest the virtue of this modified evolution strategy.**

***Index terms*—Evolutionary Algorithms, Evolution Strategies, Diversity, Fuzzy Control**

I. INTRODUCTION

The optimization methods collectively called evolution strategies (ES) were originally developed by Rechenberg [2] and Schwefel [3] to solve difficult parameter optimization problems. The methodology belongs to the same family of biology-inspired numerical operations under the umbrella term *evolutionary computation*, as do genetic algorithms (GA) [4] and evolutionary programming (EP) [5].

Like all other evolutionary approaches, evolution strategies improve the population of possible answers by an iteration of the evolution cycle: reproduction and/or recombination, mutation, and selection. The flowchart of a typical cycle is provided in Figure 1.

For the sake of convenience in later discussions, the above procedure is hereby restated in popular nomenclature for evolution strategies. In each generation (iteration), there will be μ parents ($P(t)$) producing λ offspring ($P'(t)$) through the above mentioned evolution operators; then, μ individuals ($P(t+1)$) will be chosen from the population to be the parents in the next generation.

Although these population-based optimization algorithms make use of "only simple computations, such as additions, random number generations, and logical comparisons" and require only limited knowledge of the problems at hand [24], their performance of depends greatly on the interactive effects of the operation parameters on the optimization process.

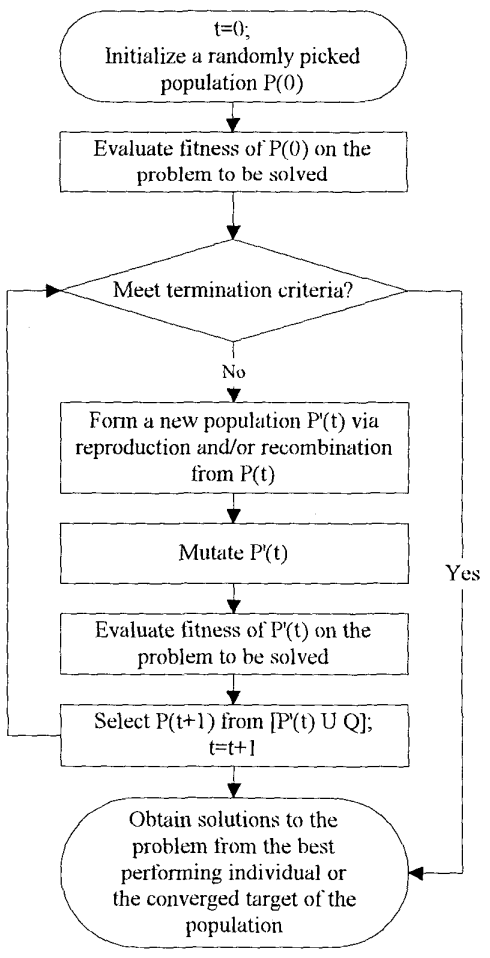


Fig. 1 Flowchart of a typical evolution cycle

In short, the evolution operators of recombination and mutation are to increase the diversity of the population in a favorable manner [25]. On the other hand, the operator of selection is responsible for convergence of solutions [24].

Indeed, diversity and convergence are the life of all population-based search methods. If an algorithm is able to maintain the *diversity of intermediate population* over an ever-shrinking region of search space, that is, monotonicity of convergence, it should be able to find an optimum point in a relatively short period. Nevertheless, it is hard to achieve this goal without "a certain optimality criterion for the combined process of selection, crossover, and mutation" [24].

In this paper, several diversity measures are employed as possible candidates for the aforementioned optimality criterion. These criterion candidates are then applied to modify popular implementations of evolution strategies. The results obtained therein not only manifest the virtue of the modified evolution strategy, but also vindicate diversity measures as effective optimality criteria for evolution operators.

II. Evolution Strategies

The evolution strategies used in this paper are as those described in [1]. The jth individual P_j of population in an n-dimension optimization problem is a vector comprised of *object variables*, $x_i \in \mathrm{R}$, $1 \le i \le n$, and the so-called *strategy parameters*, $\sigma_i \in \mathrm{R}_+$, $1 \le i \le n$. More formally, an individual in the population is expressed as $P_j = (X, \Sigma)$.

As apparently indicated in this formulation, the object variables part of individual vector represents a physical point in the search space. However, strategy parameters in the individual are responsible for the adaptation of mutation step sizes. Mutation is the major evolution operator in evolution strategies, with a relative recent incorporation of crossover operators as a result of hybridization between ES and GA.

A typical implementation of self-adapting mutation mechanism in evolution strategies is provided below for easy reference:

$$\sigma'_i = \sigma_i \cdot \exp[\tau' \cdot N(0,1) + \tau \cdot N_i(0,1)]$$
$$x'_i = x_i + \sigma'_i \cdot N_i(0,1)$$

with $\tau \propto (\sqrt{2n})^{-1}$ and $\tau \propto (\sqrt{2\sqrt{n}})^{-1}$, and

$N(0,1)$: normally distributed random variable with zero expectation and variance of one.

As mentioned before, even though the evolution cycle is characteristic of all evolutionary computation approaches, strategies stand out from the others by several features. Among these features are: (1) the use of real number to represent individuals in the search space, (2) the emphasis of mutation as a key evolutionary operator, (3) the mechanism of self-adaptation in mutation step sizes and inclination angles, and (4) deterministic selection based on ranking. Nevertheless, what has often been underestimated is the importance of different *parent-inclusion* schemes in evolution strategies as a search operator by itself.

There are two major variants of evolution strategies: the (μ, λ) evolution strategy (termed the *COMMA strategy* [6]) and the ($\mu + \lambda$) evolution strategy (termed the *PLUS strategy* [6]). In COMMA strategy, only the λ offspring enter the selection pool at the end of each generation. On the other hand, PLUS strategy performs selection over the union of μ parents and λ offspring.

Different studies have reported successful implementations with either strategy [7] [8]. However, in a survey paper [1], it is suggested that preference of either strategy is most likely problem-dependent, and that both strategies are special cases of the general (μ, κ, λ) evolution strategy, where κ denotes the maximum life span of an individual.

Since it is so hard to decide between the methods, compromises could be attempted by automatically switching between each strategy with a supervising mechanism. In this paper, a diversity-based fuzzy parent-inclusion scheme is devised to accommodate both PLUS and COMMA strategies.

This scheme employs a fuzzy logic controller, which use population diversity measures as inputs controls the parent-inclusion scheme. Simulation results of test problems are provided to manifest the virtues of the modified evolution strategies. It is concluded that the parent-inclusion scheme indeed retains advantages from PLUS and COMMA strategies and performs well to a variety of problems.

III. Exploration and Exploitation in Evolution Strategies

The behavior of most of evolution operators can be condensed into the two antipodal concepts of exploration and exploitation. For example, the procedure of mutation aims at exploring the search space, while recombination means to exploit useful information implicitly encoded in the current individuals so as to guide the direction of search. With exploration operators, new information can be introduced into the population to avoid premature convergence over local optima, whereas exploitation operators try to accelerate the rate of population convergence through the incorporation of learnt knowledge.

Although selection operators only seem to embody the biological idea of "survival of the fittest", it can still be endowed with the responsibility of search guidance. The distinction between the above-mentioned PLUS strategy and COMMA strategy is just one example. The constitution of the intermediate population, or the pool of genes under selection, differs in these two strategies. The PLUS strategy practices a *parent-inclusive* scheme, where the entire parent population joins in the selection pool. However, the COMMA strategy practices a *parent-exclusive* scheme, in which all individuals from the previous generation are excluded from selection.

Dissimilar search orientations are formulated in these two schemes. On the one hand, to select in the PLUS fashion is to exploit the knowledge of parents; on the other hand, selection in the COMMA style utilizes exploration results from the offspring only. Consequently, it is so observed in [9] that: "[the COMMA strategy] accept temporary deterioration that might help might help to leave the region of attraction of a local optimum and reach a better optimum, [while in PLUS strategy,] a monotonous course of evolution is guaranteed." Furthermore, the COMMA strategy is expected to "perform better on problems with an optimum moving over time, or on problems where the objective function is noisy" [7].

To sum up, the advantage of PLUS strategy is a guaranteed course of convergence to local/global optimum because of its insistence on elitism – keeping track of the best result ever found. In contrast, the explorative nature of the *non-elitist* COMMA strategy can either accelerate the convergence velocity rapidly or at least avoid noisy sub-optimal traps.

IV. Elitism, Diversity, and the Size of Mutation Step

Surfacing from the distinctions between PLUS strategy and COMMA strategy are the key notions of elitism and diversity. By elitism, convergence to global optimum can be proven as long as the possible range of mutation covers the region of optimization [10] [11] [12] [13]. By diversity, search operation can remain effective over a wide variety of problems, even though its convergence could only be attained under certain conditions [14]. To exemplify this point, a performance comparison of PLUS (PES) strategy, COMMA (CES) strategy, as well as simple GA on the sphere optimization problem:

$$P1: f_1(X) = \sum_i x_i^2, \qquad \text{where } X = \{x_i \mid i = 1..30\}$$

is conducted, along with results given in Figure 2.

The first noteworthy point in this example is the resemblance between the evolution tendencies of each algorithm with results obtained in [15], where CES, Evolutionary Programming, and simple GA are tested on the same problem. The remarkable similarity of PES evolution course with that of Evolutionary Programming is not a coincidence at all, since Evolutionary Programming equates a probabilistic variant of $(\mu + \mu)$ evolution strategy, disregarding minor differences in reproduction and self-adaptation mechanism [1] [9].

Another interesting observation with this example is the early saturation of GA. It can be shown that the phenomenon is related to mutation step size and that the convergent value is a function of population size, mutation step size, and the rate of mutation [22]. This premature convergence contrasts drastically with the superb performance of ES, which is accountable by the self-adaptation mechanism of strategy parameters, i.e. mutation step sizes.

A question may now come to the mind of the suspecting: why does PES perform poorly compared with CES in this simple problem? Again, this is related to the interaction of self-adapting strategy parameters with either parent-inclusion schemes, which is the subject of [22]. To put it succinctly, the rule of thumb will be: when the expected step-size is appropriate for the current problem, CES performs better. On the other hand, when the expected step-size is inappropriate, PES performs better. Finally, when the expected step-size or variance in fitness value with respect to a step-size is too small, PES and CES behave virtually the same.

From the above comments, it is clear that PES and CES sometimes do behave quite differently. This is why preference of either strategy is most likely

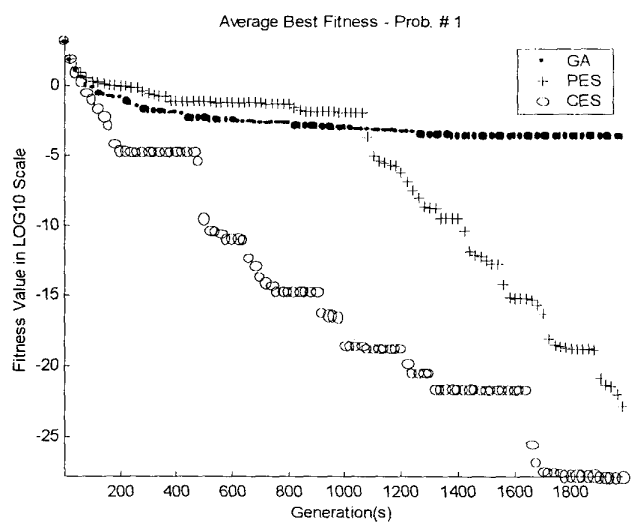


Fig. 2 Comparison of Performance of GA, PES, and CES on f_1

problem-dependent. Nevertheless, in order to gain benefits from the contradictory diversity-orientation and elitism, an adaptive parent-inclusion scheme must be used. That is, the new evolution strategy will pursue both convergence velocity and convergence reliability by changing the number of parents to be included across the generations and thus reach an ideal trade-off between these factors.

V. Evolution Strategy with an Adaptive Parent-Inclusion Scheme – A Diversity-Based Approach

No matter it is the case with dynamic or static optimization problems, it has been found that using non-constant parameters can improve the performance of evolutionary algorithms [16] [17] [23]. Consequently, the concept of *multi-strategy* approach has been proposed for on-line tuning of these parameters [18].

For the two variants of parent-inclusion schemes, with which the (μ, κ, λ) evolution strategy is a general form, the number of parents to be included in the selection pool is either all or none. In other words, $P(t+1)$ = SELECT $[P'(t) \cup Q]$, where Q can be either $P(t)$ or $\varnothing$, depending on which of the PLUS scheme or COMMA scheme is used.

To carry out the idea of an adaptive parent-inclusion scheme, a modified diversity-based selection is employed and the number of parents to be included varies between the extremities. Consequently, the algorithm now becomes: $P(t+1)$ = SELECT $[P'(t) \cup Q]$, where $Q = m$ randomly selected individuals from $P(t)$, and m = F (diversity measures of population).

Here, F (...) is a fuzzy logic controller (FLC), which feeds on diversity measures of population to give out the number of parents to be included in the selection pool. The underlying rationale is that the more of parents are to be included in the selection pool, convergence rate goes up and diversity goes down. Thus, the aims of the fuzzy logic controller F (...) are to modulate a dynamic balance between diversity and elitism by monitoring the diversity measures.

VI. Diversity Measures

Three sets of diversity measures are used, the first set is a modified set of indicators for fitness diversity and convergence rate as used in [18] [19],

$$F1: 1 - \frac{\text{Average Fitness (t)}}{\text{Best Fitness (t)}}, \text{ and}$$

$$F2: \frac{\text{abs(Best Fitness (t) - Best Fitness (t - 1))}}{\text{Best Fitness (t) - Average Fitness (t)}}.$$

It is denoted as *Fitness-diversity-based selection* (FDES). The second and third set of diversity measures employ average distribution radius and radius standard deviation and are denoted as *Spatial-diversity-based selection* (SDES), i.e.

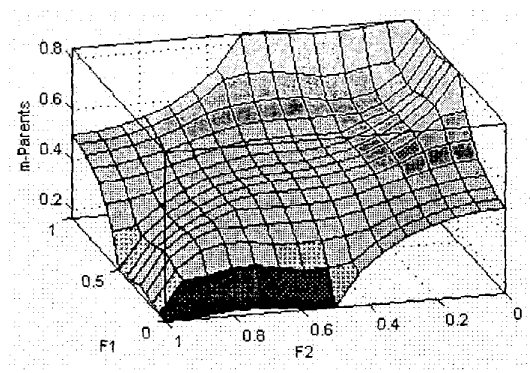
$S1$: Average Distribution Radius (cluster), and
$S2$: Radius STD (cluster).

The control surfaces of each of the FLC's used for parent inclusion in the parent-inclusion scheme are plotted in Figure 3 with normalized inputs.

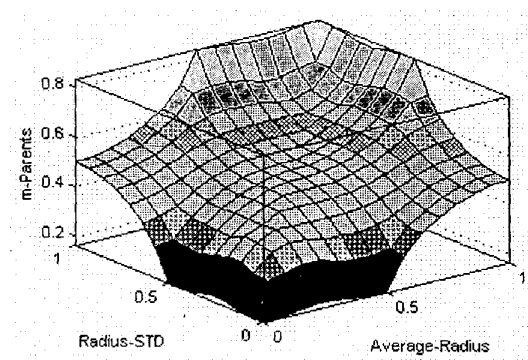
Despite their mutual use of the same formulae, the object clusters in each scheme differ. For the second set of indicators, the cluster in measure is the sub-population of offspring at each generation; however, for the third set, the cluster in measure is the sub-population of parents.

VII. Simulation Design

To compare the effectiveness of the above three modified



(a) FDES



(b) SDES

Fig. 3 Control Surface of FLC Used for Parent Inclusion Selction Scheme

evolution strategies to the original PLUS strategy and COMMA strategy, two optimization problems are used [15].

$$P1: f_1(X) = \sum_i x_i^2, \text{ where } X = \{x_i \mid i = 1..30\}$$

(Sphere function), and

$$P2: \begin{aligned} f_2(X) &= (-10)\exp((-0.05)\sqrt{\tfrac{1}{n}\sum_i x_i^2}) \\ &- \exp(\tfrac{1}{n}\sum_i 1.3\cos(0.5\pi x_i)) + 20 + e \end{aligned}$$

(Ackley function, n=30).

The sphere function is a continuous, uni-modal, convex function used to test the convergence velocity of the algorithms. The multi-modal Ackley function, however, has a number of randomly distributed extremes over the search space and a single global optimum. With this problem, convergence reliability is important. A 3-D graphical representation of the Ackley function is rendered in Figure 4.

To assess each approach fairly with their best performance, the following parameters are chosen or decided upon according to their population size, selection intensity, and previously reported performance after reviewing the numerical data from [15] [20] [21], and are listed in Table 1 for reference.

Simulation results are plotted in Figure 5 and Figure 6. As can be expected from intuition, the three modified evolution strategies indeed perform better with respect to either PLUS or COMMA evolution strategy. Moreover, the modified evolution strategies using spatial diversity measures stand out from the one using fitness diversity measures in both cases.

VIII. Conclusion

In this paper, modified evolution strategies with an adaptive fuzzy parent-inclusion scheme capitalizing population diversity information are introduced. The modified

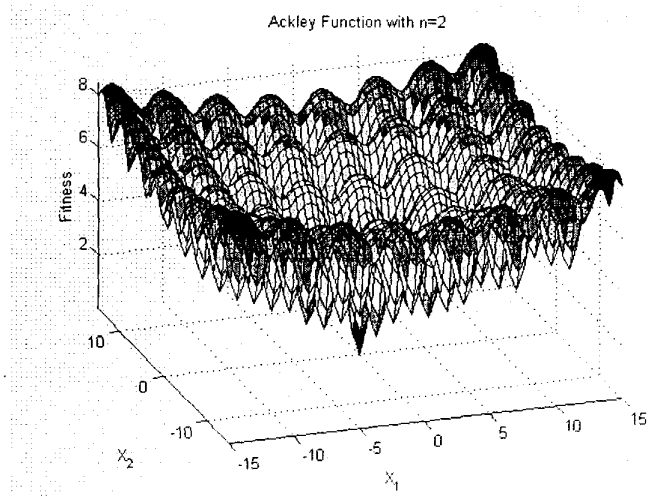


Fig. 4 3D Graphical Representation of Optimization Problem # 2

evolution strategies work because advantages of the original PLUS strategy and COMMA strategy are preserved and applied in accordance with population diversity.

The information is sent into a fuzzy logic controller to calculate the number of parents to be included in the selection pool. The purpose of this practice is to maintain a balance between diversity and elitism so that the strategy can adapt itself under all kinds of situations in a variety of problems. Simulation results have shown distinctive improvement on performance of these modified strategies with respect to the original PLUS and COMMA evolution strategies.

Moreover, the results show that spatial diversity measures are better than fitness-based diversity measures, even though additional computation is needed for calculation of spatial diversity.

Still another implication of this paper is the awareness of the importance of parent-inclusion scheme as a search guiding apparatus in evolutionary strategies. Finally, the applicability of this parent-inclusion scheme for various combinations of mutation setup should not be overlooked. This is because preference between PLUS strategy and COMMA strategy do change on a case-by-case basis, and

Table 1 Parameters of the Evolution Strategies Used in Simulation

Code Name	μ (Parents)	λ (Offspring)	Initial Strategy Parameter Value	Crossover		Selection Type
				Probability	Type	
PES	30	200	3	0.05	Discrete	PLUS selection
CES	30	200	3	0.05	Discrete	COMMA selection
FDES	30	200	3	0.05	Discrete	Fitness-diversity-based selection
SDES (O)	30	200	3	0.05	Discrete	Spatial-diversity-based selection (Offspring)
SDES (P)	30	200	3	0.05	Discrete	Spatial-diversity-based selection (Parent)

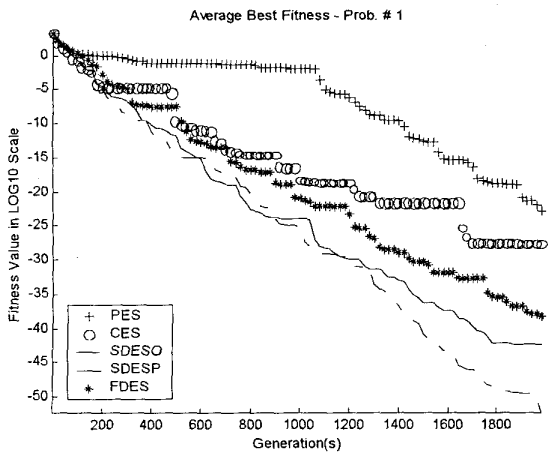


Fig. 5 Comparison of Performance of PES, CES, SDESO, SDESP, and FDES on f_1

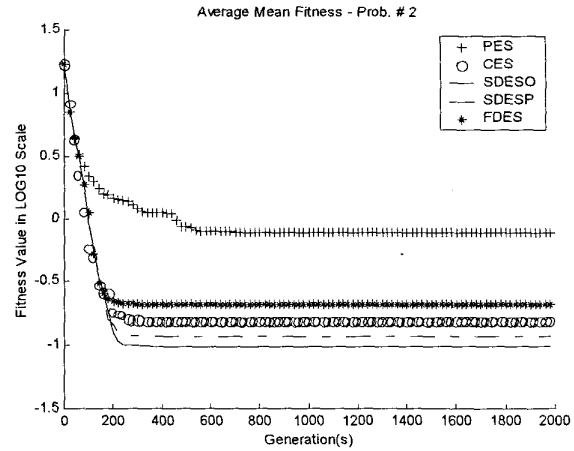


Fig. 6 Comparison of Performance of PES, CES, SDESO, SDESP, and FDES on f_2

sometimes both strategies just behave indistinguishably. The cause of this further complexity resides in the interaction between objective function landscape, mutation step-size, and mutation rate. Nevertheless, this interesting issue is above the scope of this paper and further research into the assertion is required.

IX. References

[1] T. Bäck, U. Hammel, and H.-P. Schwefel, "Evolutionary Computation: Comments on the History and Current State", IEEE Transactions on Evolutionary Computation, vol. 1, no. 1, pp. 3-17, Apr. 1997.

[2] I. Rechenberg, Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution. Stuttgart, Germany: Frommann-Holzboog, 1973.

[3] H.-P. Schwefel, Evolutionsstrategie und numerische Optimierung, Dissertation, Technische Universität Berlin, Germany, May 1975.

[4] J. H. Holland, "Outline for a Logical Theory of Adaptive Systems", J. Assoc. Comput. Mach., vol. 3, pp. 297-314, 1962.

[5] L. J. Fogel, "Autonomous Automata", Ind. Res., vol. 4, pp. 14-19, 1962.

[6] Jörg Heitkötter and David Beasley, eds., "The Hitch-Hiker's Guide to Evolutionary Computation: A list of Frequently Asked Questions (FAQ)", USENET: comp.ai.genetic. Available via anonymous FTP from rtfm.mit.edu/pub/usenet/news.answers/ai-faq/genetic/, 1999.

[7] Z. Michalewicz, Genetic Algorithms + Data Structures = Evolution Programs, 3rd Edition, Berlin: Springer, 1996.

[8] D. K. Gehlhaar and D. B. Fogel, "Tuning Evolutionary Programming for Conformationally Flexible Molecular Docking", Proc. 5th Annu. Conf. on Evolutionary Programming, Cambridge, MA: MIT Press, pp. 419–429, 1996.

[9] T. Bäck and H.-P. Schwefel, "Evolutionary Computation: An Overview", Proc. 3rd IEEE Conf. on Evolutionary Computation, Piscataway, NJ: IEEE Press, pp. 20-29, 1996.

[10] D. F. Fogel, Evolving Artificial Intelligence, Ph D. Thesis, University of California, San Diego, 1992.

[11] J. Born, Evolutionsstrategien zur Numerischen Lösung von Adaptationsaufgaben, Dissertation A, Humboldt-Universität, Berlin, 1978.

[12] F. J. Solis and R. J.-B. Wets, "Minimization by Random Search Techniques", Math. Operations Research, 6: 19-30, 1981.

[13] A. E. Eiben, E. H. L. Aarts, and K. M. Van Hee, "Global Convergence of Genetic Algorithms: A Markov Chain Analysis", Parallel Problem Solving from Nature, eds. H.-P. Schwefel and R. Männer, Berlin and Heidelberg: Springer, pp. 4-12, 1991.

[14] Günter Rudolph, "Convergence of Non-Elitist Strategies", Proc. 1st IEEE World Congress on Computational Intelligence, vol. 1, pp. 63-66, 1994.

[15] T. Bäck and H.-P. Schwefel, "An Overview of Evolutionary Algorithms for Parameter Optimization", Evolutionary Computation, vol. 1, no. 1, pp. 1-23, 1993.

[16] J. J. Grefenstette, "Optimization of Control Parameters for Genetic Algorithms", IEEE Trans. Systems, Man, and Cybernetics, vol. 16, no. 1, pp. 121-128, 1986.

[17] Laurent Dubois et al., "Improving Genetic Algorithms Using Fuzzy Logic", Systems Analysis, Modeling, and Simulation, vol. 18-19, pp. 241-244, 1995.

[18] Laurent Dubois and Toshio Fukuda, "On a Multi-Strategy Approach in Evolution Computation", Proceedings of IEEE International Conference on Evolutionary Computation, pp. 576-580, 1996.

[19] M. A. Lee and H. Takagi, "Dynamic Control of Genetic Algorithm Using Fuzzy Logic Techniques", Proc. ICGA'93, pp. 76-83, 1993.

[20] Tobias Blickle and Lothar Thiele, "A Comparison of Parent-inclusion schemes Used in Evolutionary Algorithms", Evolutionary Computation, vol. 4, no. 4, pp. 361-394, 1997.

[21] D. G. Mayer et al., "Survival of the Fittest – Genetic Algorithms versus Evolution Strategies in the Optimization of systems Models", Agricultural Systems, vol. 60, pp. 113-122, 1999.

[22] Tao-Yuan Huang and Yung-Yaw Chen, "Issues of Mutation Step-size in Evolution Strategies", to be published.

[23] Joáo Carlos Costa, Rui Tavares, and Agostinho Rosa, "An Experimental Study on Dynamic Random Variation of Population Size", 1999 IEEE International Conference on Systems, Man, and Cybernetics, 1999, vol. 1 , pp. 607-612, 1999.

[24] Xiaofeng Qi and Francesco Palmieri, "Theoretical Analysis of Evolutionary Algorithms With an Infinite Population Size in Continuous Space – Part I: Basic Properties of Selection and Mutation", IEEE Trans. Neural Networks, vol. 5, no. 1, pp. 102-119, January 1994.

[25] Xiaofeng Qi and Francesco Palmieri, "Theoretical Analysis of Evolutionary Algorithms With an Infinite Population Size in Continuous Space – Part II: Analysis of the Diversification Role of Crossover", IEEE Trans. Neural Networks, vol. 5, no. 1, pp. 120-129, January 1994.