A Note on the Empirical Evaluation of Intermediate Recombination

David B. Fogel
Natural Selection, Inc.
3333 N. Torrey Pines Ct., Suite 200
La Jolla, CA 92037
fogel@sunshine.ucsd.edu

Hans-Georg Beyer
University of Dortmund
Dept. of Computer Science
D-44221 Dortmund
Germany
beyer@ls11.informatik.uni-dortmund.de

Abstract
The effectiveness of intermediate recombination in evolution strategies is analyzed in light of the typical procedure of initializing trial solutions uniformly about the global optimum of benchmark functions. Analysis indicates that this procedure may predispose results in favor of intermediate recombination.
Keywords
Evolutionary algorithms, intermediate recombination, benchmark functions.

1. Introduction

Closed-form descriptions of the error convergence rates of evolutionary optimization algorithms have been difficult to attain. Indeed, there are no known error convergence rates for genetic algorithms, and the known convergence rates for evolution strategies and evolutionary programming are restricted, with few exceptions, to strongly convex functions (Bäck et al., 1993) or approximations of such functions (e.g., Beyer, 1995). The nonlinear and stochastic procedures of the algorithms make such analysis difficult or intractable. Many efforts have instead been directed toward empirical investigation: Convergence rates are assessed by executing multiple trials of an algorithm on a set of test functions and estimating the mean performance. In this manner, the efficiency of different search operators and selection procedures can be compared using statistical hypothesis testing.

One such search operator is termed *intermediate recombination*. This operator has many forms, but a routine implementation acts on two parents, say vectors x_S and x_T (using the notation of Bäck and Schwefel (1993)), and creates an offspring vector x' as follows:

$$x'_i = x_{S,i} + U \cdot (x_{T,i} - x_{S,i}) \tag{1}$$

where i is the ith component of the vectors and U is a uniform random variable in the range [0,1]. Under this operator, offspring are generated within a hyperbox of uniform density having vertices at the minimum and maximum values of each parent for each component. Intermediate recombination has been examined within research in evolution strategies[1] as a method to adjust the parameters that are used in evaluation of the error of the proposed solution (e.g., McDonnell & Waagen, 1995). (It has also been used with somewhat greater emphasis to adjust *strategy parameters* for self-adaptation (see Bäck & Schwefel, 1993).)

1 Bäck noted (personal communication) that the typical procedure within evolution strategies is to set $U = 1/2$ (see also Bäck and Schwefel, 1993) as a constant. This generates what is essentially a stochastic bisection search.

Evolution strategies often rely on a two-step variation process of recombination and mutation. For real-valued function optimization problems, this two-step procedure is frequently executed by using recombination in the generation of a predetermined number of offspring and then applying mutation in the form of multivariate zero-mean Gaussian perturbation to each offspring. The search procedure is typically begun by distributing trials uniformly over a range in each parameter of the objective function. The optimum solution is known a priori for most test functions (it is often set to be the origin) and the initial range for each parameter is typically symmetric about the optimum value of each component. Although this method of initialization has presumably been chosen so as not to bias the results, it will be proved that uniformly distributing trial solutions about the global extremum predisposes results in favor of intermediate recombination on the first generation.

2. Analysis

THEOREM: *If initial trial solutions are uniformly distributed symmetrically about the optimum solution, the use of intermediate recombination followed by independent zero-mean Gaussian perturbation generates offspring that are unbiased estimates of the optimum solution.*

PROOF: *Although the following analysis extends to multidimensional objective functions, for simplicity, consider a one-dimensional objective function $f(x)$, where x is a scalar value, and let the population size be restricted to two parents, X_1 and X_2. Let the minimum of $f(x)$ be $x = 0$. Under symmetric initialization, each parent takes on a value in accordance with a uniform random variable:*

$$X_1 \sim U(-L, L)$$
$$X_2 \sim U(-L, L)$$

where L is the upper limit and −L is the lower limit.

Let R be the intermediate recombination

$$R = X_1 + U(X_2 - X_1) \tag{2}$$

where $U \sim U(0, 1)$. Then R may be rewritten as

$$R = (1 - U)X_1 + UX_2 \tag{3}$$

X_1, X_2, *and U are statistically independent, with* $E(X_1) = E(X_2) = 0$, *and* $E(U) = 0.5$; *therefore,* $E(R) = 0$.

The variance of R can be determined by

$$\begin{aligned} V(R) &= E(R^2) - E(R)^2 && (4) \\ &= E(R^2) && (5) \\ &= E((1 - U)^2 X_1^2 + 2U(1 - U)X_1X_2 + U^2X_2^2) && (6) \\ &= E(X_1^2)E((1 - U)^2) + 2E(U(1 - U))E(X_1)E(X_2) + E(U^2)E(X_2^2) && (7) \end{aligned}$$

$$\begin{aligned} E(X_1^2) &= \int_{-L}^{L} x^2 f_{x_1}(x)\, dx && (8) \\ &= 1/(2L)[x^3/3|_{-L}^{L}] && (9) \\ &= L^2/3 && (10) \end{aligned}$$

Similarly $E(X_2^2) = L^2/3$ *and* $E(U^2) = 1/3$. *Therefore,*

$$E((1-U)^2) = E(1 - 2U + U^2) \quad (11)$$
$$= 1 - 2E(U) + E(U^2) \quad (12)$$
$$= 1/3 \quad (13)$$

Substituting,

$$V(R) = (L^2/3)(1/3) + (1/3)(L^2/3) \quad (14)$$
$$= (2/3)(L^2/3) \quad (15)$$
$$= (2/9)L^2 \quad (16)$$

Finally, adding a mutation following a zero-mean Gaussian random variable $N \sim N(0, \sigma^2)$ *to the result of the intermediate recombination yields an offspring*

$$X_{R+N} = R + N \quad (17)$$

with

$$E(X_{R+N}) = E(R) + E(N)$$
$$= 0 \quad (18)$$

Therefore, X_{R+N} *is an unbiased estimator of the global minimum of* $f(x)$ *and, moreover, due to the statistical independence of* R *and* N,

$$V(X_{R+N}) = V(R) + V(N) \quad (19)$$
$$= (2/9)L^2 + \sigma^2 \quad (20)$$

□

A similar theorem can be proved for the use of zero-mean Gaussian perturbations alone.

THEOREM: *If initial trial solutions are uniformly distributed symmetrically about the optimum solution, the use of zero-mean Gaussian perturbation generates offspring that are unbiased estimates of the optimum solution.*

PROOF: *A parent* $X_1 \sim U(-L, L)$ *has*

$$E(X_1) = 0 \quad (21)$$

and

$$V(X_1) = L^2/3 \quad (22)$$

The value of an offspring created solely through mutation is

$$X_N = X_1 + N \quad (23)$$

with expectation

$$E(X_N) = E(X_1) + E(N)$$
$$= 0 \quad (24)$$

Offspring created from the initial parents solely through mutation are thus also unbiased estimates of the minimum of $f(x)$ when the initialization is uniformly symmetric about the minimum and have variance

$$V(X_N) = V(X_1) + V(N) \quad (25)$$

$$= L^2/3 + \sigma^2 \quad (26)$$

□

Given two unbiased estimators, preference should always be given to the estimator with lower variance. The ratio of the variance obtained by not using intermediate recombination to the variance obtained by using intermediate recombination is

$$V(X_N)/V(X_{R+N}) = (L^2/3 + \sigma^2)/(2L^2/9 + \sigma^2) \quad (27)$$

$$= (3L^2 + 9\sigma^2)/(2L^2 + 9\sigma^2) \quad (28)$$

The ratio depends on the values of L and σ^2, but it will always be greater than 1 for positive L and σ^2; the variance using intermediate recombination will always be lower.[2] This follows solely from the initial distribution of parents being uniformly symmetric about the minimum of $f(x)$.

The preceding analysis applies only to the case of a one-dimensional objective function, but the results concerning the initial distribution of offspring extend to arbitrary multi-dimensional objective functions. However, the specific ratios of the variances (Equation 28) will vary for functions of higher dimensionality.

3. Conclusions

Intuitively, there are clearly cases in which intermediate recombination may be of help and also cases in which it may be detrimental. If the population bounds the global optimum, then repeatedly searching within a hyperbox defined by pairwise samples from the population will tend to act like a convex optimization procedure and this has the potential to converge quickly (Beyer, 1995). However, if no elements of the population bound the global optimum, intermediate recombination cannot discover it. Further, if the population resides in a locally optimal well, intermediate recombination may accelerate convergence to a false optimum and inhibit overall global convergence.

But more to the point, the conventional procedure of exclusively distributing initial populations uniformly about the global optimum should be abandoned. This procedure may unduly prejudice the results of empirical trials in favor of or against various search operators. In practice, the optimum of the objective function at hand is not known and it may not be possible to distribute initial trials symmetrically about the optimum. Indeed, it may not even be possible to bound the optimum with certainty. Evolutionary optimization algorithms should be tested on benchmark functions in various configurations that include initializing the population with large perturbations directed away from the optimum. Uniform initialization about the global optimum should be performed only as part of a broader spectrum of possible initializations and is itself only of limited interest.

2 A similar relationship can be proved for the special case of intermediate recombination when $U = 0.5$. Under such conditions, the expression indicated in Equation 28 would take on the form $(2L^2 + 6\sigma^2)/(L^2 + 6\sigma^2)$.

Acknowledgments

The authors thank the referees for their comments on and criticisms of this paper. D. B. Fogel acknowledges the assistance of T. Bäck, L. J. Fogel, J. R. McDonnell, A. V. Sebald, and D. E. Waagen in conducting reviews of an early draft.

References

Bäck, T. (1994). Personal communication, University of Dortmund, Germany.

Bäck, T., Rudolph, G., & Schwefel, H.-P. (1993). Evolutionary programming and evolution strategies: Similarities and differences. In Fogel, D. B. & Atmar, W. (Eds.), *Proceedings of the Second Annual Conference on Evolutionary Programming* (pp. 11–22). La Jolla, CA: Evolutionary Programming Society.

Bäck, T., & Schwefel, H.-P. (1993). An overview of evolutionary algorithms for parameter optimization. *Evolutionary Computation*, *1*(1), 1–24.

Beyer, H.-G. (1995). Toward a theory of evolution strategies: On the benefits of sex—the $(\mu/\mu, \lambda)$ theory. *Evolutionary Computation*, *3*(1), 81–111.

McDonnell, J. R., & Waagen, D. E. (1995). An empirical study of recombination in evolutionary search. In McDonnell, J. R., Reynolds, R. G., & Fogel, D. B. (Eds.), *Evolutionary Programming IV: The Proceedings of the Fourth Annual Conference on Evolutionary Programming* (pp. 465–478). Cambridge, MA: MIT Press.