Global Optimization by means of Distributed Evolution Strategies

Günter Rudolph
Department of Computer Science
University of Dortmund
P.O. Box 500 500
D-4600 Dortmund 50

1 Introduction

Multi-membered evolution strategies, as introduced by Schwefel [1], are known to be robust optimum seeking procedures for a variety of parameter optimization problems. On account of their inherent parallelism multi-membered evolution strategies can be implemented easily on parallel computers. However, while there have been several studies on parallel versions of genetic algorithms [2][3], only little work has been done on parallel versions of evolution strategies.

The approach taken here to parallelize the evolution strategy can be considered as a coarse grained one: A large population is divided into several subpopulations - each of them 'living' on a different transputer. After one or more generations some individuals migrate to a neighboring subpopulation, or in other words: the distributed optimization processes exchange information during the search.

One field of application might be the search for a global optimum in high dimensional parameter spaces with nonlinear objectives - a problem which is *unsolvable in general* [4]. Even if the search space and the accuracy are bounded the general problem remains NP-hard. While those problems can be solved by means of complete enumeration in low dimensions this strategy must fail in high dimension due to the exponential increase of the necessary effort.

An analoguous argumentation about the solvability can be drawn to the general class of combinatorial problems: among them you can find the well-known *Travelling Salesman Problem (TSP)* which has been included into the testbed. To make the evolution strategy running on this kind of problems one has to map a feasible solution of the TSP to a so-called 'object variable vector' of the evolution strategy and vice versa. This will be explained in the next sections in more detail.

2 Distributed Evolution Strategies

Two new parameters have been added to the traditional evolution strategy. The parameter *migration period (MP)* represents the number of generations wherein no migration takes place. The number of the migrants is controlled by the parameter *migrants (Mig)*.

In addition to various parameter settings of these two parameters the recombination parameter of the traditional strategy has been varied, too. Using *intermediate recombination (I)* means, that the genotype/phenotype vector of the offspring is generated by evaluating the mean vector of the parents' vectors. *Discrete recombination (D)* can be regarded as a dynamic n-point crossover: the genome of the offspring is produced by choosing either the vector component of the first or the second parent with the same probability.

3 Continuous Problems

3.1 Convergence Speed

In order to determine the convergence speed of the distributed evolution strategy each variant at first was tested on a unimodal (strict convex) problem:

$$f(\underline{x}) = \sum_{i=1}^{30} x_i^2 \quad . \tag{1}$$

The starting point for each subpopulation was set at random to $x_i \approx \pm 10^{50}$ to obtain a large obversation period. From the data (fig. 1) we can assume that every variant has a *geometric* resp. *linear-R convergence* for the unimodal problem.

Definition :

Let (ε_k) be a nonnegative sequence with $\varepsilon_k \to 0$ for $k \to \infty$. If there is an index k_0, a number $C > 0$ and a number $r \in [0\,,\,1)$, so that

$$\varepsilon_k \leq C \cdot r^k \quad \forall k > k_0 \ , \tag{2}$$

then the sequence (ε_k) is said to be *geometrically convergent* or *linear-R convergent*. □

From the data we can derive

$$10^{-\frac{38}{800}k+50} \leq \varepsilon_k := ||\underline{x}_k - \underline{x}^*|| \leq 10^{-\frac{27}{800}k+50} \quad .$$

Hence, with $C = 10^{50}$ and $k_0 = 1$ we get

$$0.8964 \approx 10^{-\frac{38}{800}} \leq r \leq 10^{-\frac{27}{800}} \approx 0.9252 \quad ,$$

which fulfills (2). That means that the convergence speed does not suffer from migration obviously.

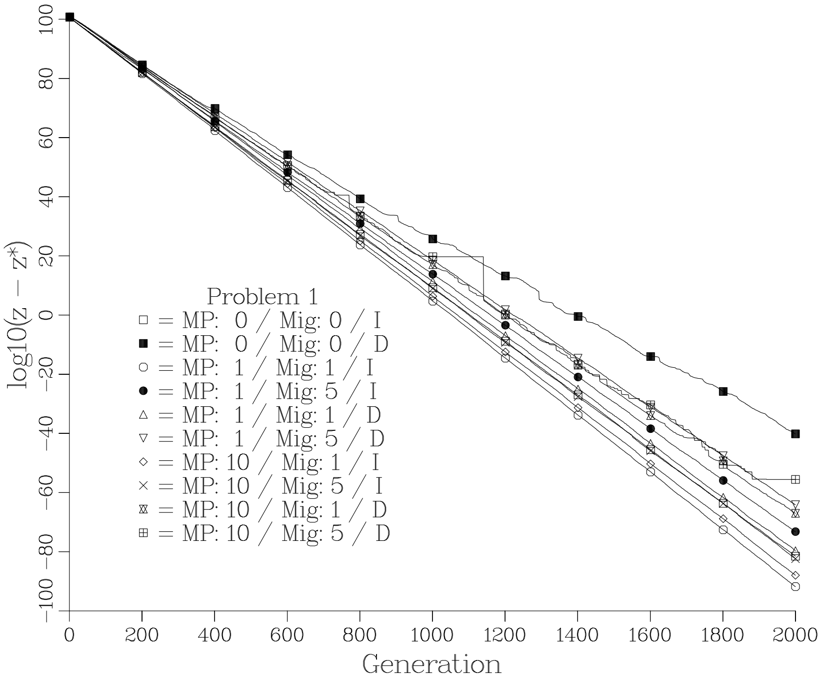

Figure 1: Best overall convergence for problem 1

3.2 Reliability

The reliability was tested for a generalized form of Rastrigin's problem [4] which can be considered as a perturbed version of problem (1):

$$f(\underline{x}) = \sum_{i=1}^{30} \left\{ x_i^2 + 50 \cdot [1 - \cos(2\,\pi\,x_i)] \right\} \quad . \tag{3}$$

The starting points were chosen at random in $S = \{\underline{x} \in \Re^{30} : |x_i| \leq 1000\}$, so that the function has 2001^{30} local minima in S. However, if the population is far from the global minimum the perturbations are dominated by the quadratic term. The closer the strategy approaches the global minimum the stronger the effects of the perturbations.

From figure 2 one can see that only those variants with discrete recombination *and* migration could find the global minimum: all intermediate variants and even the discrete variant *without* migration failed. But this might be explained by the symmetric arrangement of the local minima: Suppose, a subpopulation has converged to a local minimum (n = 5)

```
(0, 1, 0, 1, 1)'
```

and another to the local minimum

```
(1, 0, 0, 1, 0)'   .
```

By means of migration individuals from the subpopulations meet each other and making use of discrete recombination they perhaps produce an offspring with the following genome:

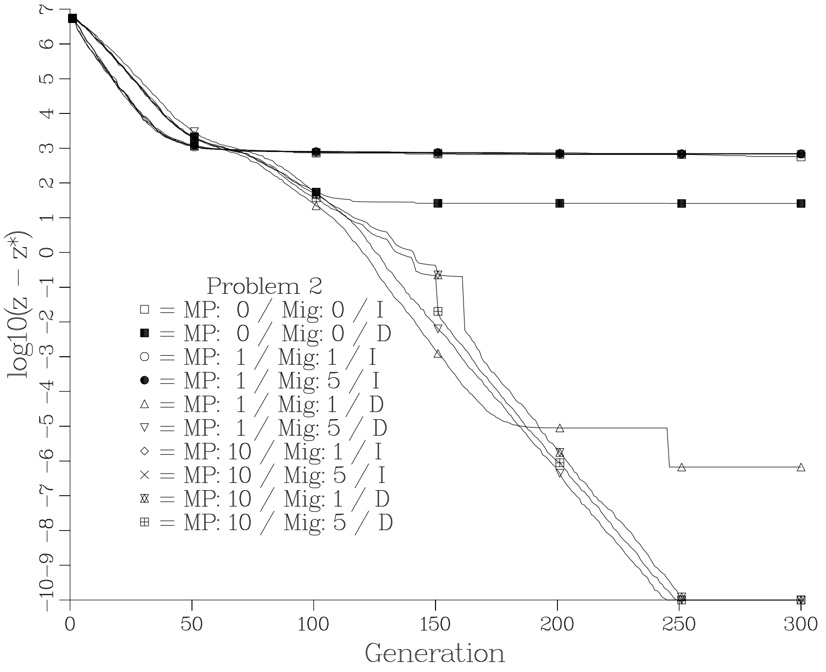

Figure 2: Best overall convergence for problem 2

```
(0, 0, 0, 1, 0)'  .
```

This offspring is much closer to the optimum than its parents and after some migrations in subsequent generations the global optimum can be found. If this is the right explanation, this effect will occur whenever the local optima are arranged regularly related to each vector component - regardless of the scaling.

On the other hand, if an offspring is produced by intermediate recombination the genome looks like

```
(0.5, 0.5, 0, 1, 0.5)'  .
```

Such offsprings are *always* worse than their parents so that the effects of migration will be eliminated by selection with high probabilty.

4 Combinatorial Problems

4.1 Genotype/Phenotype-Mapping

The genotype of an individual is represented by an object variable vector $\underline{x}$ with contiuous components and the strategy variable vector $\underline{\sigma}$. To get the phenotype resp. a feasible solution of the TSP the components of the object variable vector $\underline{x}$ are sorted, so that the corresponding sorted indices of the vector represent a feasible tour. The sorting was done by a variant of *heapsort* [5].

4.2 First results

The distributed algorithm with the above mapping was tested for a 100 cities problem. The best tour known so far has the length 21285. The best tour found by the distributed evolution strategy within 500 generations has the length 21381 and can be seen in fig. 3. Surely, it is not likely that an evolution strategy using this simple mapping will find the global minimum, but there might be some possibilities to invent a more sophisticated mapping. However, for a 30 cities problem the best known tour was found within 50 generations.

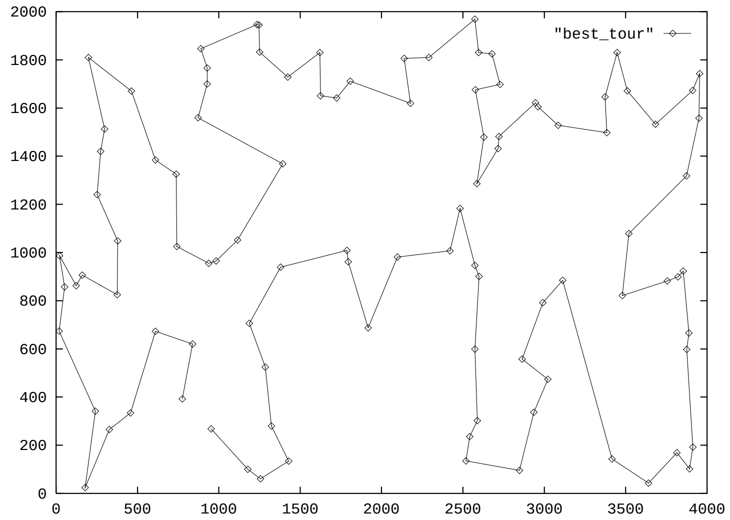

Figure 3: Best tour for the 100 cities problem

References

[1] Schwefel, H.-P. : *Numerical Optimization of Computer Models*, Wiley, Chichester 1981.

[2] Grefenstette, J.J. (ed.) : *Genetic algorithms and their applications*, Proc. of the Second Int'l. Conf. on Genetic Algorithms, July 28-31, 1987, Lawrence Earlbaum, Hillsdale 1987.

[3] Schaffer, J.D. (ed.) : *Genetic algorithms*, Proc. of the Third Int'l. Conf. on Genetic Algorithms, Morgan Kaufman, San Mateo 1989.

[4] Törn, A. and A. Žilinskas : *Global Optimization*, Lecture Notes in Computer Science Vol. 350, Springer, Berlin 1989.

[5] Press, W.H.; B.P. Flannery; S.A. Teukolsky and W.T. Vetterling : *Numerical Recipes*, Cambridge University Press, Cambridge 1986.