The Evolution Strategy. A Mathematical Model of Darwinian Evolution

Ingo Rechenberg

Technische Universität Berlin, Fachgebeit Bionik und Evolutionstechnik
Ackerstraße 71-76, D-1000 Berlin 65, Fed. Rep. of Germany

1. Introduction

Everybody is fascinated, realizing the technical performance of living systems. Thus we should remember that living beings are the result of a large-scale experiment on the Earth, called biological evolution. It may be worthy to reflect on this biological development. Over the past years it has become obvious that rules of biological evolution are the result of an evolution process itself. Suppose a population of organism with slightly modified hereditary rules compared to the existing norm. If these modifications will help the population to adapt faster to their particular environments, then this population will have a better chance to survive in future than a population with less effective hereditary rules. Therefore it should be assumed that evolution, during its action over more than a thousand million years, gave itself an optimal mode of operation. This hypothesis results to the following statements:

1. The imitation of rules of biological evolution should yield an excellent experimental method in engineering to design better technical apparatus.

2. The similarities between biological evolution and optimization give rise to the idea that common concepts in optimization theory may be used to describe Darwinian evolution.

2. Evolution Strategy as an Optimization Technique

It was in 1964 [1] when I started the first experiment to imitate the method of biological evolution in a laboratory of fluid mechanics. An aluminium plate flexible at five positions was mounted in an open wind tunnel (Fig. 1). The articulated plate can be altered stepwise. There are more than 345 million possible forms. The task is to find the shape with minimum drag. We all know that this is a flat plate, directed parallel to the air stream. But suppose we don't know that. Therefore the plate is set into a random starting configuration. To produce the random alternations of the five hinges (the mutations in biology) we used in our first experiment a mechanical apparatus (Fig. 2). Five balls, representing the five hinges of the plate, pass the pyramid of pins and land in the ground boxes. The box markings determine the alternations of angle.

Our experimental arrangement makes it possible to measure the fitness of the mutated shapes. The technical fitness is the drag of the plate, which has to become a minimum. At the beginning of the wind-tunnel experiment the plate was folded into a zig-zag shape of high drag. The experimental scheme to imitate rules of biological evolution will be discussed later in detail. The basic idea is to reject all mutations with increasing drag (decreasing fitness). But if a randomly generated form has a lower drag, then it becomes parental shape for the next mutations. Figure 3 shows the result of our first evolution-like experiment. The drag of the plate is plotted versus the number of generations. Below the diagramm the best of the plate after every ten generations is shown. We achieve the plain shape after 300 generations.

E. Frehland (ed.), *Synergetics — From Microscopic to Macroscopic Order*

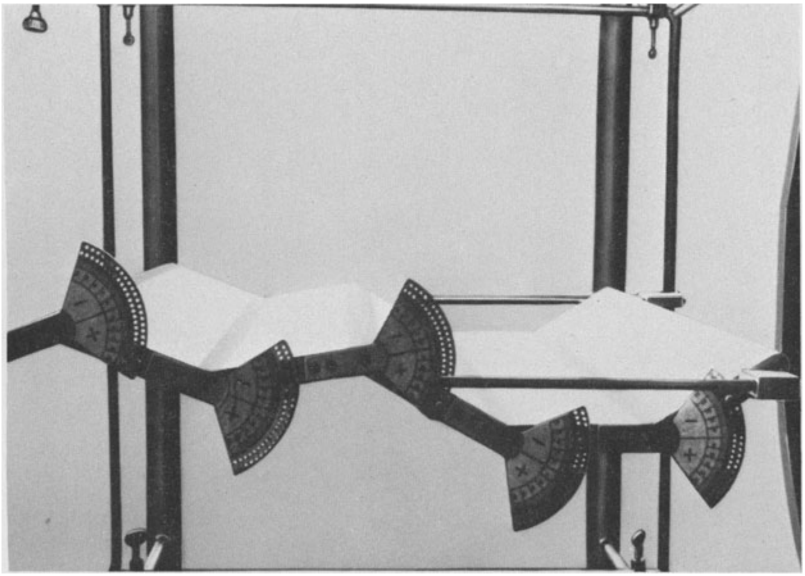
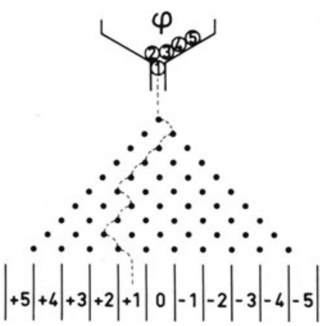


Fig. 1 The articulated plate for an experiment in the wind tunnel to imitate Darwinian evolution



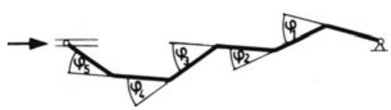


Fig. 2 Galton's pin board. Passing the pyramid the balls land in compartments. The labels indicate the alternations of angle

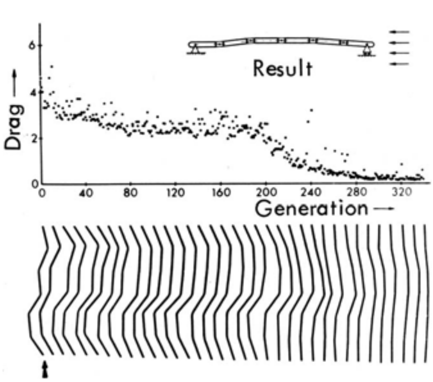


Fig. 3 The experimentum crucis. Starting the experiment the plate was set in a random configuration. Applying rules of biological evolution the plate develops to the form of minimum drag. The expected solution of a flat plate is found

For the next experiment we wished to change a boundary condition. Suppose we turn the wind tunnel. In the biological world this would be a change of the environment. Because it is too expensive to turn the complete wind tunnel, we turned the flat plate (Fig. 4). Now evolution goes on and the flat plate develops to an S-shaped curvature with minimum drag.

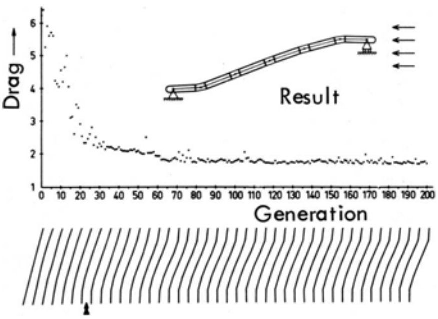


Fig. 4 Continuation of the experiment. The boundary conditions were changed by raising the leading edge of the plate. The new environmetal factor will cause the plate to evolve to an S-shaped curvature

Whilst the tests with the articulated plate are more of academic value, the following three experiments in the field of fluid dynamics are of practical interest [2]. The next task was to find the form of a right-angled pipe bend with minimum flow resistance. Figure 5 shows the experimental arrangement. Two flexible plastic hoses are held by adjustable bars in the deflection zone. The positions of the six bars of each pipe are the optimization variables. At the beginning of the experiment both pipe elbows were adjusted to a quarter of a circle. While one pipe bend was continuously varied according to the algorithm of evolution strategy, the second pipe remained unchanged as the reference system. Figure 6 shows the initial form and the optimum form one upon the other. The optimum pipe bend, having 10% less deflection losses, starts with a steadily increasing curvature (similar to Euler's spiral) and ends with a small reverse in curvature.

In another experiment made by SCHWEFEL [3] a two-phase supersonic flow nozzle was developed using evolution strategy. Heated water vapourises partly in the throat of a convergent-divergent nozzle. The expanding vapour then forms the propellent for the remaining liquid. It is impossible to calculate the shape of the nozzle for maximum thrust. For the experimental optimization the nozzle was made up of segments. A total of 400 segments with different conical borings were available. By continually

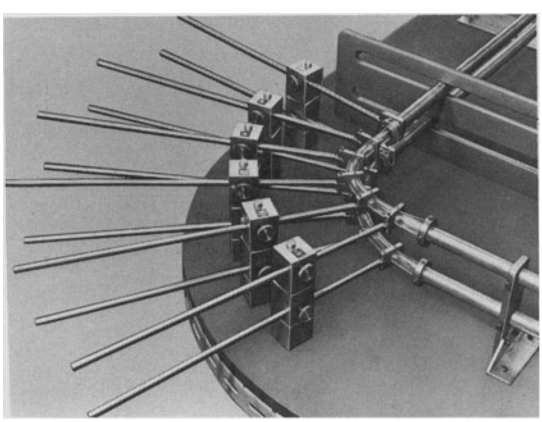


Fig. 5 Experimental arrangement for the evolution of a pipe elbow

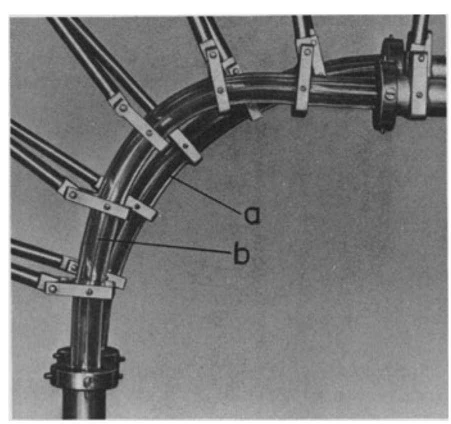


Fig. 6 Circular form (a) and optimum form (b) of the pipe elbow

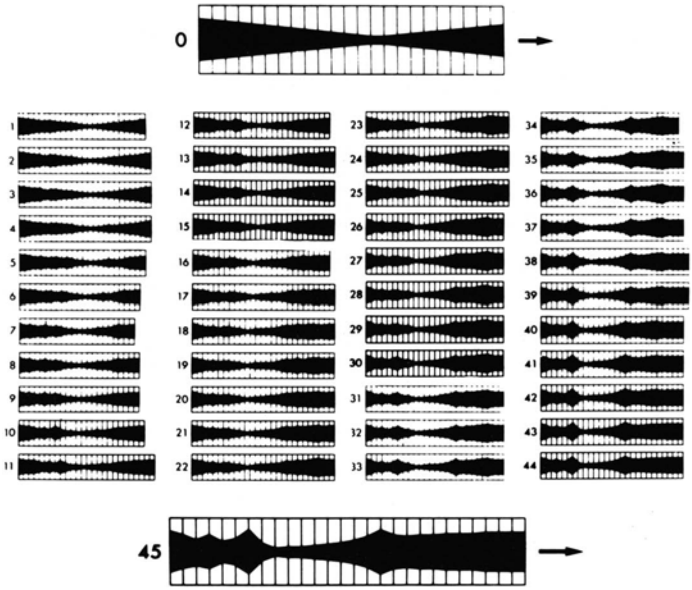


Fig. 7 Evolution of a hot-water flashing nozzle. A set of 400 segments, each conically bored, made it possible to build up different nozzle configurations

exchanging segments in accordance with the rules of evolution strategy, an optimum nozzle form was found which looks like a modern vase (Fig. 7). In this experiment the evolution strategy has invented something new . That is a chamber in the diverging part of the nozzle, where mixing takes place between liquid and vapour. The energetic efficiency of the initial conical nozzle, calculated on the basis of supersonic theory, was 55%. The optimum nozzle has an efficiency of nearly 80%.

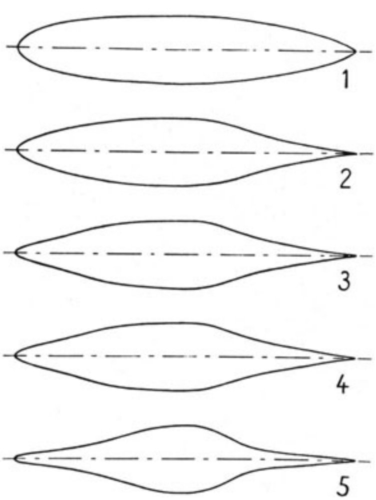


Fig. 8 Development of a body of revolution for minimum drag shown at intervals of 600 generations

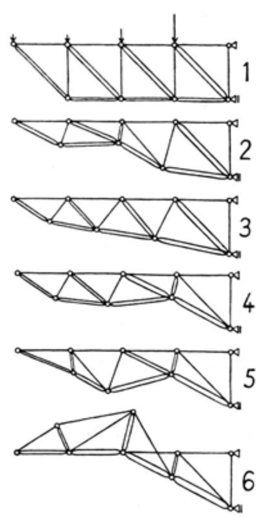


Fig. 9 Development of a lattice frame for minimum weight using evolution strategy

It is a classical problem of fluid dynamics to find the minimum drag profile of an axisymmetric body. Recently PINEBROOK solved this problem on a computer, combining boundary layer calculation techniques with the evolution strategy. Figure 8 gives an example from his work [4]. A body of revolution with a constant maximum diameter and a given length evolves to a dolphin-like form with minimum drag. The computer run required 2400 generations to fit the 21 intensities of the variational problem.

Finally an application of the evolution strategy is presented in the field of structural engineering. A lattice frame with given loads has to be constructed for minimum weight. The variables of the structure are the plain coordinates of the six joints. Figure 9 shows the development of the frame. The initial design as the result of a linear optimization procedure has a weight of 922 kg. The optimum solution, which looks like a crane jib, weights only 718 kg. This computer simulation was performed by HOEFLER [5].

3. Sequence of Operations for the Evolution Strategy

So far the evolution strategy has proved to be an excellent optimization procedure. Now the time has come to elucidate the exact algorithm. It is useful to introduce at this point a nomenclature for the different modes of the evolution strategy [6], [7]. This will be cone by the formal abbreviation:

$$(\mu \overset{+}{,} \lambda)\text{-ES} .$$

We shall call μ the number of parents of a generation and λ the number of descendants (the children of the parents). We start with the simplest imitation of evolution. This is a (1+1)-ES. A game of cards (data cards) demonstrates how this algorithm works (Fig. 10a). A data card represents a genotype in biology. By card re-

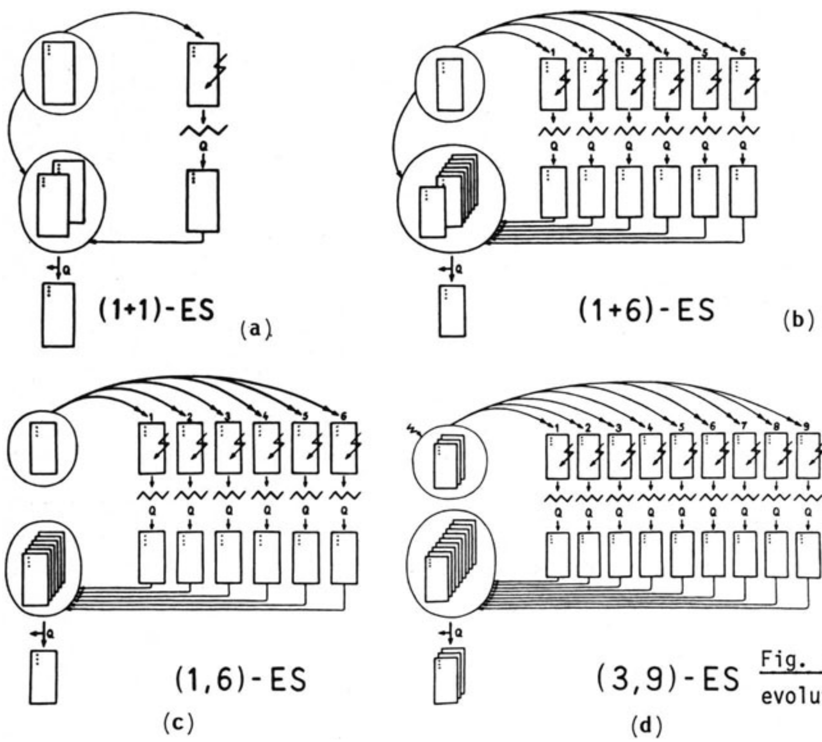


Fig. 10 Grades of evolution strategies

peating the parental information will be doubled. This is the DNA replication in biology. DNA replication errors are introducedin our game by using a random number generator. Then the randomly altered information of the offspring has to be realized. We get the phenotype, which is in our example a slightly changed form of the articulated plate. We measure the drag of the plate within the airstream of the windtunnel. The drag becomes the fitness of the offspring genotype. The parental card plus the offspring card are put into a ballot box. The best card holds the parental information for the next generation. This so-called two-membered evolution strategy has been applied in our first experiments.

A higher imitation level of biological evolution will be obtained with the algorithm of a (1+6)-ES. Now the parent will produce 6 offspring (Fig. 10b). Random alternations are introduced and the information is translated into the phenotypes. After the drag has been measured all 6 offspring cards plus the parental card are put into the ballot box. The best of them will become the parent for the next generation.

The next scheme takes into account that the parents have a limited lifespan (Fig. 10c). We set up a (1,6)-ES. The comma in this notation indicates that the parental card is not included in the selection. With the exception of this modification the operations are the same as in the preceding algorithm. This kind of evolution strategy is preferred in our mathematical treatment today.

Last we introduce a population of three parents. They produce in a random sequence 9 offspring. Because it is the comma version of the evolution strategy only the 9 offspring are put into the ballot box. The three best of them survive and become the parents for the next generation. This scheme will be named a (3,9)-ES.

What can we do with this formalism? Certainly we can apply the statistical operations to optimize an airfoil in a wind tunnel or even to find the solution of Rubik's cube. But we are far away from the formulation of an evolution equation. Actually I take a new course for the mathematical treatment of Darwinian evolution. The key idea is to interpret the method of biological evolution as a hill-climbing procedure. Hill climbing is a principle common to many optimization strategies. A hill-climbing strategy acts like gravity, forcing a ball to roll down the gradient of a hill, but it works in the opposite direction. The effect is that the vast space of possibilities is reduced to a narrow street, on which the optimum seeking takes place. Applying such a strategy you must make sure that a hill exists to climb up. This was the case for all our engineering experiments using the evolution strategy. I claim that there is no difference in biology. Piecewise smooth relationships between the fitness of an organism and the structure of the variables form a genetic landscape with hills to climb up.

To demonstrate the gradient climbing of a (1,10)-ES we look at an ordinary optimization problem in automobile engineering. The object is a carburettor with two adjusting screws. The contour lines in Fig. 11 represent screw settings of equal efficiency. To find the optimum setting we start with the parent at a random position (a). This parent will produce 10 children (b). Then the parent will die out (c). The offspring with the highest efficiency is declared to become the new parent (d). All other children die out (e). The new parent will produce the 10 children of the next generation (f). The setting point of the variables moves up the hill with a certain speed.

4. The Discovery of the Evolution Window

We will now direct our attention on the calculation of the rate of progress. To do this we must know the local form of the fitness function. A smooth fitness function may be described by a general quadratic equation with the variables $x_1, x_2, \ldots, x_n$:

$$(1) \qquad Q = Q_o + \sum_{k=1}^{n} a_k x_k - \sum_{i=1}^{n} \sum_{k=1}^{n} b_{ik} x_i x_k \ .$$

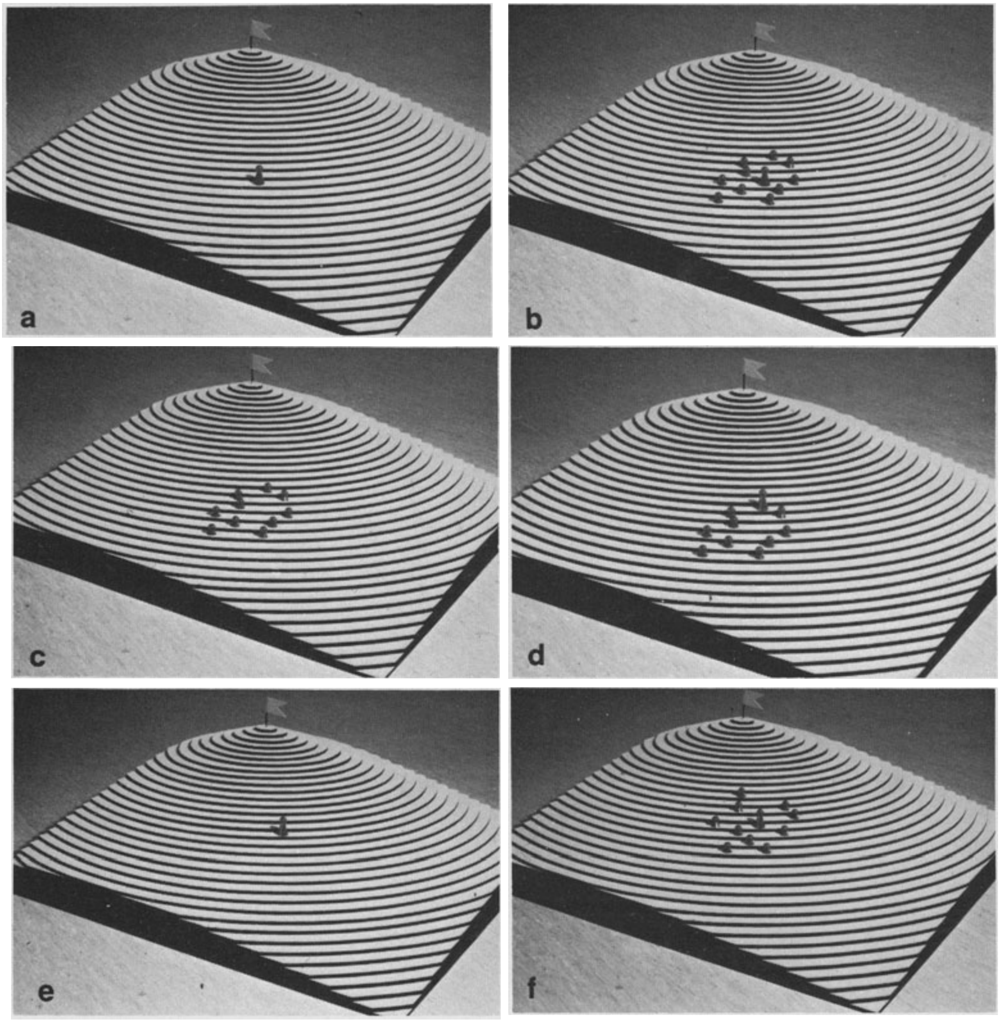


Fig. 11 The hill-climbing procedure of the multimembered evolution strategy

The variables are altered by random numbers satisfying a $(0,\sigma)$ normal distribution. Because the mutations are distributed spherically symmetric around the parental point, a rotation of the coordinate axes will be allowed. The transformation of equation (1) to the principal axes gives the result:

$$(2) \qquad Q = Q_0 + \sum_{k=1}^{n} c_k y_k - \sum_{k=1}^{n} d_k y_k^2 \quad .$$

For this general description of the local form of a fitness function I succeeded in calculating the rate of progress in the case of a (μ,λ)-ES [8]. Here we introduce a simplification. Assuming equal coefficients d_k=d for each variable - we speak of an isotropic non-linear local behaviour of the fitness function - the rate of progress is given by the formula:

(3) $$\varphi_{\mu,\lambda} = \frac{c_{\mu,\lambda}\,\delta}{\sqrt{n}} - \frac{\delta^2}{2r}$$

$c_{\mu,\lambda}$ = progress coefficient
δ = mutation step length
n = number of variables
$r = \sqrt{\sum c_k^2}/2d$.

All difficulties are concentrated in the evaluation of the progress coefficient. In the case of a $(1,\lambda)$-ES one has to solve the integral:

(4) $$c_{1,\lambda} = \sqrt{\frac{2}{\pi}}\,\frac{\lambda}{2^{\lambda-1}} \int_{z=-\infty}^{\infty} z\,e^{-z^2}\left[1+\mathrm{erf}(z)\right] dz \; .$$

Fortunately the progress coefficient will change only from 1,5 to 3,5 when λ is increased from 10 to 1000 [7].

Figure 12 shows the theoretical result in the form of a diagram . The climbing speed Φ (in a universal notation) is represented as a function of the mutation step size Δ (also in a universal notation). It is a very remarkable result that the evolution strategy works only within a small band of the mutation step size. I have named this band the "evolution window" (Evolutionsfenster). Out of this window no evolution occurs. This is a very exciting fact, because SCHWEFEL [6] has found in his doctor thesis that the main difference between the evolution strategy and other optimum seeking methods is the exponent by which the search effort will rise with increasing number of variables. This means that the result of the evolution window is of more general evidence. It is worthy to reflect on this fact for engineering optimization as well as for the development of economic and social systems. Right to the window you may locate the field of revolution with negative values of Φ . Left to the window you will have the region of conservatism with no progress. The logarithmic scale for Δ has been chosen, because this quantity (the mutation rate in biology) will change in a decimal power mode.

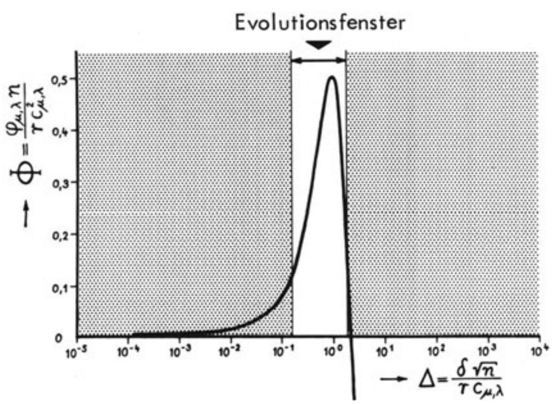


Fig. 12 The evolution window. Evolution strategy works only within a small band of the mutation step length

5. The Twofold Algorithm of the Evolution Strategy

The existence of the narrow evolution window gives rise to a new problem. How can the evolution strategy find the window in order to be effective? The answer is: the multimembered evolution strategy is a twofold optimization procedure. This very important feature of the evolution strategy is illustrated in Fig. 13 for the example of a (2,10)-ES. This time we have two parents on our fitness hill (a). Each parent will produce five offspring (b). But one of the two parents will do it with a large

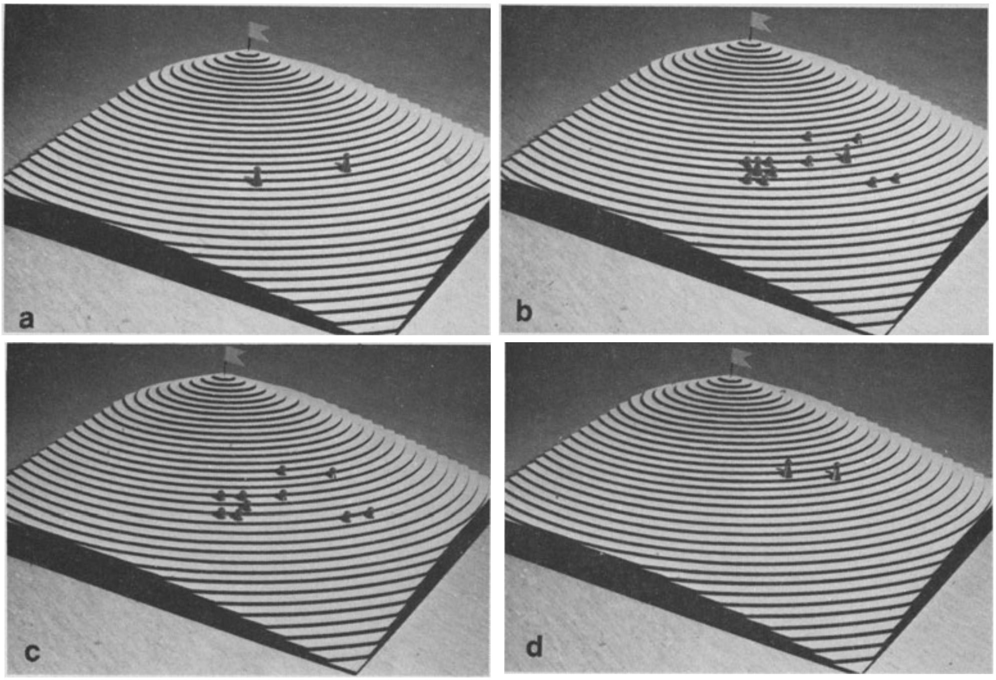


Fig. 13 Hill climbing and automatic adjustment of the mutation step length

step size, the other one with a small step size. In the language of biology: the two parents have different kinds of error correcting DNA polymerases. Next the parents will die out (c). The two offspring with the highest fitness are declared to become the parents of the next generation and all other children die out (d).

We come now to the main point: most probably these children will win having a step size to be located nearer to the centre of the evolution window. In biology, however, the mutation step size, or more precisely the mutation rate, is a hereditable character. That does mean that the step size of a surviving offspring has to be transferred to the corresponding parent of the next generation. Before creating a new set of children, however, we must not forget to mutate the step length of each parent. In this case the mutation step size adapts itself to the local topology of the fitness funtion to get maximum rate of progress.

We look at a similar situation in the life of man. Suppose you are an alpinist climbing a difficult mountain-side. You have picked out an appropriate climbing technique. You can't, however, decide if it is the best, because there is no standard of comparison. Next day you are climbing in a group. Each member of the group is using a slightly different climbing technique. After a short time it becomes evident which technique is the best and you may copy it.

The $(1,\lambda)$-ES, taking step length mutations into account, can be formalised as follows:

$$\delta_{N1}^{g} = \delta_{E}^{g} \cdot \xi_{1} \qquad \text{(offspring 1 of generation } g\text{)}$$
$$\underline{X}_{N1}^{g} = \underline{X}_{E}^{g} + \delta_{N1}^{g} \cdot \underline{Z}_{1}$$
$$\vdots \qquad\qquad \vdots$$
$$\delta_{N\lambda}^{g} = \delta_{E}^{g} \cdot \xi_{\lambda} \qquad \text{(offspring } \lambda \text{ of generation } g\text{)}$$
$$\underline{X}_{N\lambda}^{g} = \underline{X}_{E}^{g} + \delta_{N\lambda}^{g} \cdot \underline{Z}_{\lambda}$$

Let NB denote the best offspring: $Q(\underline{x}^g_{NB}) = Max\{Q(\underline{x}^g_{N1}), \cdots Q(\underline{x}^g_{N\lambda})\}$

$$\delta_E^{g+1} = \delta_{NB}^g \qquad \text{(parent of generation } g+1\text{)}$$
$$\underline{x}_E^{g+1} = \underline{x}_{NB}^g$$

In this algorithm the random vectors z_i have normally distributed components. In accordance to the logarithmic scale of the mutation step length in the evolution window, the deviates ξ_i are obtained from log-normally distributed numbers.

6. Gradient Climbing in Darwinian Evolution

What is the optimum mutation step length of the evolution strategy adapting itself to maximum rate of progress? Our hill-climbing theory gives the analytical expression:

$$\delta_{opt} = \frac{c_{\mu,\lambda} \cdot r}{\sqrt{n}} \quad . \tag{5}$$

At this point it has to be stated that the formulas (2),(3) and (5) are the result of an asymptotic theory ($n \to \infty$). The general solution becomes more complicated. However, we are just interested in a complex system having many degrees of freedom (variables). Now, if the number of variables is increasing, the optimum step length to get maximum rate of progress will decrease. If it would be possible to look into a multidimensional Euclidian space designed by the n variables of the problem, then one would observe that an evolution process looks like a one-dimensional diffusion process winding up the gradient path of the fitness function. The picture of an evolution process to find the optimum is not that of a concentrating cloud of points. It is an elongating chain of points following the gradient line of the fitness function (Fig. 14). The gradient path acts as a guiding thread from the starting point to an optimum of the functional landscape. Evolution strategy will not scatter in the vast space of possibilities.

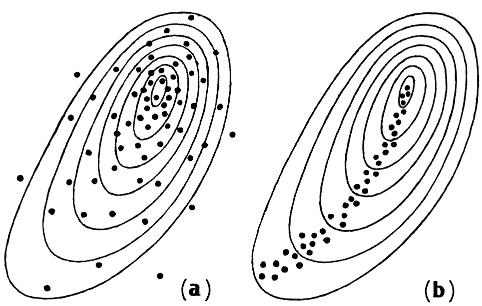


Fig. 14 Hill-climbing picture of the evolution strategy

a) False: shrinking cloud of random points

b) True: diffusion along the gradient line

Many people extrapolate the two-dimensional view of the problem into n dimensions. This turns out to be a fatal error. You can poke around in a two-dimensional manifold to find an optimum. Of course, applying pure random search you can jump from a local to a global optimum in two dimensions. However, you have no chance to repeat this in n dimensions. There is only one way to overcome the curse of dimensionality: to follow the gradient line. All optimum seeking methods act in this way. The evolution strategy, however, will do that with a minimum search effort.

References

1. I. Rechenberg: "Cybernetic Solution Path of an Experimental Problem", Royal Aircraft Establishment, Library Translation 1122, Farnborough (1965)
2. I. Rechenberg: Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution (Frommann-Holzboog, Stuttgart 1973)
3. J. Klockgether, H.-P. Schwefel: "Two-Phase Nozzle and Hollow Core Jet Experiments", Eleventh Symposium on Engineering Aspects of Magnetohydrodynamics. Pasadena, California (1970)
4. W. E. Pinebrook: "Drag Minimization on a Body of Revolution", PhD Thesis, Faculty of the Department of Mechanical Engineering of the University of Houston (1982)
5. A. Höfler, U. Leyssner, J. Wiedemann: "Optimization of the Layout of Trusses Combining Strategies Based on Michell's Theorem and on the Biological Principles of Evolution", AGARD Conf. Proc. 123, Mailand (1973)
6. H.-P. Schwefel: Numerical Optimization of Computer Models (Wiley, Chichester 1981)
7. I. Rechenberg: "Evolutionsstrategien", in Medizinische Informatik und Statistik, Vol. 8, Ed. B. Schneider and U. Ranft, Springer, Berlin 1978
8. I. Rechenberg: Lectures on the Evolution-Strategy, Technische Universität Berlin (unpublished)