ELSEVIER
Comput. Methods Appl. Mech. Engrg. 186 (2000) 125–140
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Case studies in evolutionary experimentation and computation

Ingo Rechenberg

FG Bionik und Evolutionstechnik, Technische Universität Berlin, Sekr. ACK 1, Ackerstrasse 71–76, D-13355 Berlin, Germany

Received 30 March 1999

Abstract

14 examples prove the potency of the evolution strategy. The exemplary compilation starts with the drag-minimization of a wing-like construction in a wind tunnel and ends with the creation of an extraordinary magic square on a computer. Prerequisite for a successful evolutionary optimization is the validity of strong (or at least piecemeal strong) causality for the quality function. The problem solution fails without any inherent order of the quality function. Euler's extension of Fermat's last theorem cannot be disproved with an evolutionary algorithm, because the design of a piecemeal causal pay-off function is not in sight. © 2000 Elsevier Science S.A. All rights reserved.

Keywords: Evolutionary algorithms (GA, ES, EP); Noisy fitness data; Design out fear; Optimization under noise; Convergence improvement techniques; Self-adaptation

1. Introduction

It fairly often happens that a genetic algorithm or an evolution strategy (ES) does not find an optimal solution really faster than a blind search. The reason is the lack of evolvability of the system. Disciples of the evolution strategy follow the argumentation of Darwin: evolvability of a system means that a series of small steps can be constructed, starting from a bad position, walking along moderate states, and ending at the best point. This implies the validity of regional strong causality for the quality function (small causes produce small effects).

The six examples of evolutionary experimentation (EE) and the eight examples of evolutionary computation should mainly demonstrate the evolvability of these problems. The applied ESs reach from a simple two membered-ES up to a nested multi-membered-ES. An example from the theory of numbers demonstrates that without an order à la Darwin evolutionary optimization fails.

Evolutionary experimentation means 'analog' computation in physical systems. EE was the challenge in the late 60th, when the ES got born. But it is rather expensive to change variables in the 'hardware'. With the incidence of the computers evolutionary computation triumphed. It remains unconsidered that for many engineering problems a sufficient mathematical model is not available. Expensive evolutionary experimentation is still the choice to solve these optimization problems.

In the following a short introduction into the nomenclature of the applied ESs is given. The simplest algorithm is written as (see [7])

$$(\mathbf{1} + 1) - \mathrm{ES}.$$

E-mail address: rechenberg@bionik.tu-berlin.de (I. Rechenberg).

0045-7825/00/$ - see front matter
PII: S0045-7825(99)00381-3

One parent (bold figure) produces one offspring. Parent plus offspring are put together and the best survives. Nearer to the model of biology is the (see [8])

$$(\mu + 1) - \text{ES}.$$

One of the μ parents is randomly chosen to produce a child. After that the most inferior must be removed from the $\mu + 1$ offspring. More biologically again is the (see [12,13])

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

A lottery box contains μ parents. We pick out one, produce an offspring and put the parent back into the box. We do this λ times. If the plus sign is written, the $\mu + \lambda$ individuals must be reduced to the μ best. But if the comma sign is valid, the parents die out and the μ best offspring get selected to become parents. Biological evolution works with sexual reproduction. An ES with recombination is written in the form (see [8])

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

According to the spelling: only a ρth of the parental information builds up an offspring. If $\rho = 2$ (normally in nature) a child gets half of the genes from the father and the other half from the mother. In this algorithm we pick out ρ parents, mix their variables randomly (dominant or intermediary) and put them back into the lottery box. We repeat this mixing for each offspring before mutating the variables. Finally, the nested ES is the most complex algorithm (see [9,10]).

$$[\mu'/\mu' \overset{+}{,} \lambda'(\mu/\mu \overset{+}{,} \lambda)^{\gamma}] - \text{ES}.$$

For this algorithm I prefer multi-recombination ($\rho = \mu$), which means that each parental population (after intermediary recombination) is represented by its ‘centroid’. The hierarchically topmost population (μ') is then represented by a ‘centroid of centroids’. From this position λ' founder populations get exposed. Each founder population plays the game of a $(\mu/\mu \overset{+}{,} \lambda) - \text{ES}$ during the period of γ generations. Only the nested ES can – with the help of founder populations – solve multimodal optimization problems. The prerequisite is the existence of a piecemeal strong causal order of the hills. And only the nested ES adapts strategic parameters (e. g. the mutation step-length) correctly. An approximate step-length adaptation rule is to produce λ offspring with various mutation step-lengths. Assume that the step-length of the winner gives higher rate of progress in the next generation. However, the one-generation step-length adaptation often fails. Evolution must run some time like an automobile-race to decide which strategic style (here the mutation step-length) is the best. This is what a nested ES does.

We have seen that the fraction bar in the ES-notation means a real algebraic operation. And the brackets too are more than markers. Under certain constraints the brackets may be manipulated algebraically. Then, however, we must clearly distinguish between parents and offspring. The number of parents is therefore indicated by bold figures.

2. Evolutionary experimentation

2.1. Evolution of a kink plate with minimum drag

ES got born when the following experiment was planned and executed. According to Fig. 1 six board-shaped strips were flexibly connected at their longitudinal edges. The joints could be individually adjusted and arrested in angular intervals of 2°, each joint allowing 51 steps. With their five joints, the interconnected strips could therefore assume 345 025 251 different configurations.

At the beginning of the experiment, the plate was folded into an irregular zigzag shape with the aim of developing a configuration having a minimum of drag. As it was known beforehand that a flat plate satisfies this requirement, the solution to the problem was self-evident.

The experiment, applying a $(\mathbf{1} + 1) - \text{ES}$, was performed in the summer of 1964 and was considered an ‘experimentum crucis’ [7]. The key questions at that time were whether, by using the basic

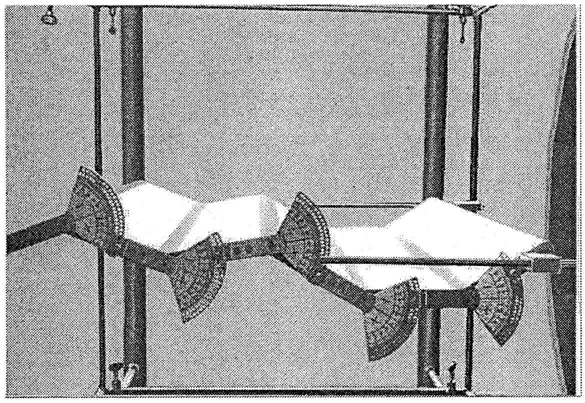

Fig. 1. Wind tunnel mounting for evolutionary experimentation with the zigzag plate.

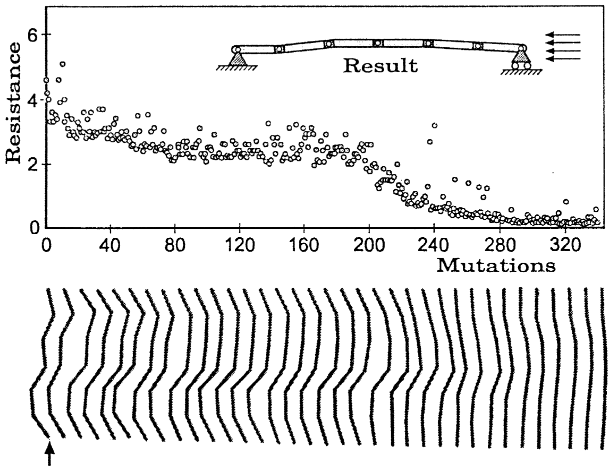

Fig. 2. The 'proof of Darwin' in a wind tunnel. Drag minimization of the kink plate by mutation and selection.

mechanisms of biological evolution (mutation and selection only), a flat configuration of the plate could be obtained and how much time was required for it. Skeptics anticipated that several million iterations would necessary. Fig. 2 depicts the magnitude of the drag of the plate versus the number of mutations.

The momentarily best configuration after every 10 mutations is shown below the diagram. Already after 300 generations the plate has assumed a form, which – in the technical sense – is planar. The drag increase due to the slight curvature of the plate was not measurable. It should be mentioned that the stagnation between generations 80 and 180 occurred only in the first experiment. Otherwise evolution was finished after 200 generations. Critics were not convinced of this primitive solution. Therefore the boundary condition (environment in biology) was changed: The wind tunnel being the environment of the kink plate changed its flow direction (Fig. 3).

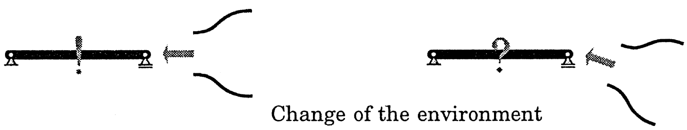

Fig. 3. Continuation of the 'experimentum crucis'.

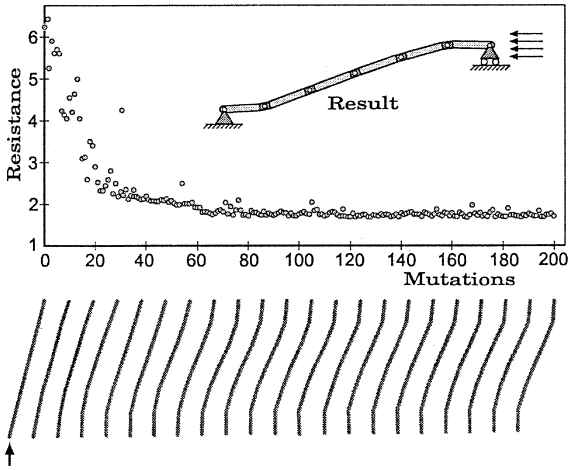

Fig. 4. Drag minimization of the kink plate when the environment changes.

In reality, however, not the wind tunnel but the plate was turned. Now evolution goes on. A plane sloping 14° against the airflow has a high resistance, because the flow separates from the upper side. The profile of minimum flow resistance turns out to be an S-shaped plate (Fig. 4). The solution does not have any technical application. Otherwise it would be a challenge for the discipline of computational fluid dynamics (CFD) to calculate this optimum shape.

2.2. Evolution of a pipe bend

More interesting in practice is the next optimization problem. The task was to find the form of a pipe bend with minimal flow losses. Fig. 5 shows two experimental setups to solve this problem. The left picture dates from the year 1965. A flexible hose is first guided in a straight pipe, then, in the 90° deflection, is held by 6 adjustable bars and finally, in the ‘smoothing section’, is again guided in a straight pipe. Two identical hoses (fed from the same kettle) permit to measure the difference between parent and offspring more exactly. What the ES has to do is minimizing the pressure difference between the entrance and the exit of the bending.

The right-hand picture shows a more modern version of the experiment. An evolution-strategic guided industrial robot adjusts the mutations now. The visible steel-wire-armoured hose contains in its interior a microcellular rubber hose. The total tube bend amounts to 180° now. The experimental setup was designed for a preliminary study in order to evolve an optimal inlet elbow for a well-known car company.

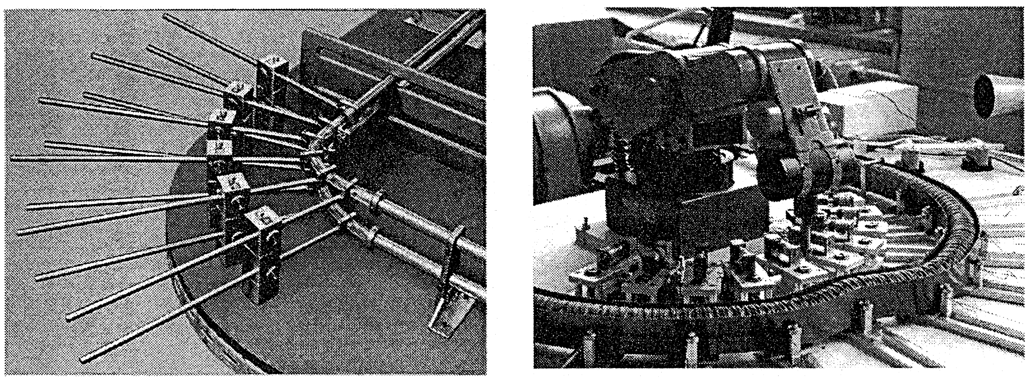

Fig. 5. Evolutionary experimentation by hand and with a robot. Six manually adjustable shafts (left) and 10 robot-controlled rope drives (right) determine the bending form of the pipe.

Both tubing are operated with pressurized air. The flow is fully turbulent. At the beginning of the experiment the pipe bends are adjusted to a quarter respectively to a half of a circle shape. Fig. 6 shows the results of the evolution-strategic experimentation. A fluid mechanics engineer would not have known the solutions before. We look to the 90° bend. While at the quarter circle bend the turn round starts with a sudden jump in curvature, there is a continuously increasing curvature for the optimum pipe bend. In the run out of the pipe bend a distinct little reversal of curvature occurs whose meaning is not yet clear. If the detour angle is 180° nearly the same details of the optimized curvature can be detected. In comparison to the initial circular configuration, both of the optimized configurations displayed a 2% reduction of the total losses. It is conventional to divide up the resistance of a pipe bend into an inevitable part of the thought stretched pipe bend and an additional contribution due to secondary flow in curvature. The optimal pipe bends then reduce the additional drag around about 10%. Because the length L of the corresponding pipe bend is very large compared to the diameter D of the pipe ($L/D = 31$), the optimal shapes might be used only for an oil pipeline. Again it would be a challenge for the discipline of computational fluid dynamics (CFD) to calculate these optimum shapes.

2.3. Evolution of a two-phase nozzle

For an almost spectacular evolution-strategic experiment a boiler with a steam power of 5 tons/h had to be operated at the Technical University of Berlin. The evolution strategists had been approached by industry with the request to optimize a two-phase jet nozzle. Superheated water was to be partially evaporated by the diminution of pressure in an appropriately tapered nozzle, the expanding steam being a propellant for the remaining fluid. The problem was to obtain a maximum velocity for the fluid steam mixture. The juxtaposition of fluid and steam leads to extremely complex flow fields in the nozzle, which defy the determination of its most advantageous shape by analysis.

Based on an idea of Hans-Paul Schwefel for an evolution-strategic experiment, the nozzle was represented by a series of segments [4,11]. A total of 330 compatible segments with conically shaped interior openings were prepared, allowing the assembly of potentially 10^{60} different nozzle configurations without discontinuities in their contours (Fig. 7).

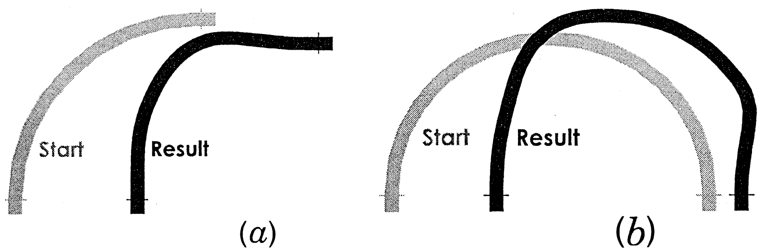

Fig. 6. Optimized shape of a 90°-pipe bend (a) and optimized shape of a 180°-pipe bend (b).

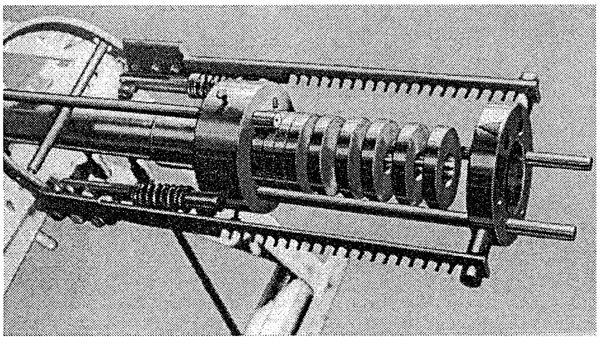

Fig. 7. Segmented hot steam nozzle. Mutability of the test object for evolutionary experimentation.

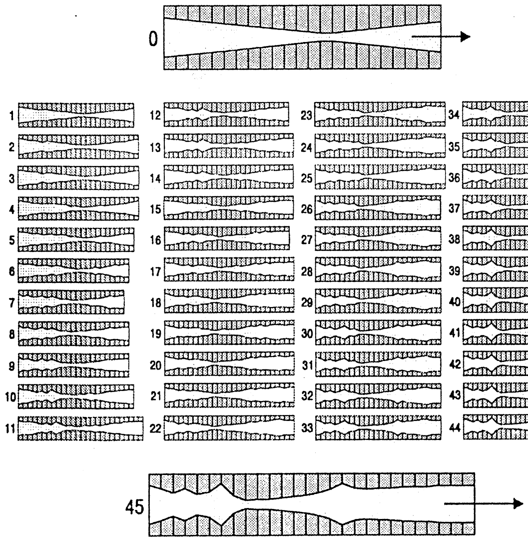
Fig. 8. Schwefel's evolutionary experiment with a two phase supersonic jet nozzle.

As initial parent configuration an analytically derived Laval-type nozzle with an especially long convergent inlet section was chosen. The continual exchange of segments by the experimental rules of a $(\mathbf{1}+1)$-ES produced a completely unexpected optimal configuration. Fig. 8 displays the initial, all successful intermediate and the final configuration of the nozzle.

The experiment was stopped after 400 generations. The efficiency raised from 55% (starting parent) to nearly 80% (optimum). The strangely shaped bulges of the optimized nozzle, of course, attracted the attention of fluid dynamicists. Further experiments with an optimized nozzle made of special glass provided insight into the function of the bulges. They promote the formation of rotational flow, which reduces the temperature gradient in the influx section and impedes the separation of liquid and steam phases in the exit section.

2.4. Evolution of a bird-like wing tip

Winglets are the 'dernier cri' in the aircraft construction. The wing ears shall reduce the margin vortex, which is induced by a flow round the wing ends due to the pressure difference between the upper and lower wing-side. The spread-wing of a bird is the invention of nature to reduce this vortex drag. In the classic bionics the engineer tries to imitate directly nature's solutions. It is however an illusion if he wants to copy the geometry of a bird-wing in the flight situation by means of a good photo. Therefore, a neo-bionicist starts with the fundamental idea he explored in biology and repeats the evolution under the special technical constraints. Fig. 9 shows such a repeated shortened evolution experiment in a wind tunnel.

The above experiment started with a flat configuration of the five wing-ears. Flexible strips of lead metal at the roots of the winglets made the system mutable. The angle of attack and the angle of staggering of each wing-ear could be altered manually by means of a special mechanical device. A computer program outputs the data of 12 mutated wing configurations, which are tested in the wind tunnel. The offspring numbers with the smallest drag to lift ratio are then transmitted to the computer. The (**4**/4, 12)-ES program declares these offspring to be the parents, mixes them and outputs the next generation.

The value drag/lift defines the glide slope of a bird or a sailplane. Later on, we compared in a control experiment the optimized bird-like wing with a non-slotted rectangular wing. From these wind tunnel data we got a diminution of the glide slope of 10.8% for the optimized wing. Because the wind tunnel model has a low aspect ratio (1 : 4), the wing tip vortex is rather strong. For a normal aircraft the expected improvement surely will be less than 5%. The presented experiment should be a pilot test to explore the evolvability of the multi-winglet system.

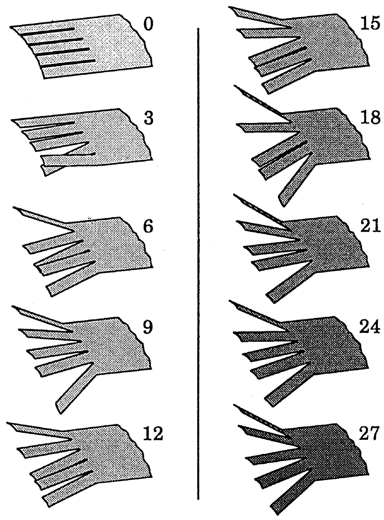

Fig. 9. Evolutionary experimentation on a bird-like wing. After 27 generations a spread wing has developed with minimum vortex drag.

3. Evolutionary computation

3.1. Evolution of an eye-lens

Skeptics of Darwin's theory of evolution cite the problem of the development of an eye. A 5% eye would not function and would be meaningless. Many precisely matched mutations would have to happen at the same time for an organ, which could function as an eye to develop. Darwin wrote:

> If it could be demonstrated that any complex organ existed, which could not possibly have been formed by numerous, successive, slight modifications, my theory would absolutely break down.

This is an anticipation of the principle of strong causality. The evolution of an eye does indeed proceed through many small steps:
1. Clustering of light sensitive cells.
2. Indentation of the light sensitive layer forming a recess.
3. Covering this recess with a jelly-like substance.
4. Shaping this substance into a light concentrator.

Today we can find archaic remains of the evolution of the eye in form of flat eyes, cup eyes, pit eyes and lens-shaped eyes. Fig. 10 shows a moldable lens of variable thicknesses d_k $(k = 1, 2, \ldots, n)$ as object of evolution.

The parallel rays of a light beam are diffracted to various degrees on passing through the glass body. Convergence of all rays at a point F is accomplished through adjustment of the thickness d_k at various

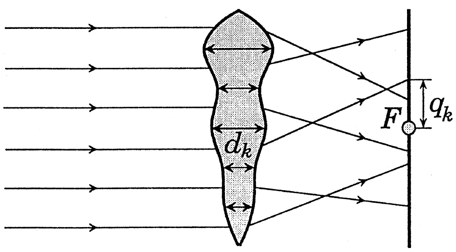

Fig. 10. Malleable glass body as object of evolution.

points (property of a convex lens). The deviation of this convergence can be expressed by the sum $\sum q_k^2$, the sum of the squares of the k individual deviations q_k. The optimization problem can be expressed by

$$\text{ray scattering} = \sum q_k^2 \rightarrow \text{min.}$$

The convex lens evolves as the ray scattering approaches the value zero. Mutations that reduce the scattering are positive with respect to selection. A life form with a primor-dial eye of better convergence can more successfully detect a danger than others. Evolutionary progress from a 1% eye to a 100% eye is possible notwithstanding the skepticism.

For the purpose of mathematical formulation the adjustable glass body with a diffraction index e is considered to be composed of prisms of height h. The base thicknesses d_k of the prisms are variables of the polygon lens. The rim thickness d_0 is preset. Light rays which impinge on the center (h / 2) of the prisms are taken as representative. The path of every light ray can be calculated according to the laws of geometrical optics (diffraction of thin prisms). The projections of the n rays on the image plane add up to the overall scattering. The quality function Q can be read directly from the drawing (Fig. 11):

$$Q = \sum_{k=1}^{n} \left[R - \frac{h}{2} - h(k-1) - \frac{b}{h}(\varepsilon - 1)(d_k - d_{k-1}) \right]^2 \rightarrow \text{min.}$$

It should be emphasized that the above equation applies only to very thin lenses. Much more complicated mathematics is needed in order to develop an achromatic super telephoto lens with the ES. For instance sin α and tan α can not be approximated by α and every ray must be exactly traced through the glass in order to find the exit point. Fig. 12 shows the development of a window pane to a magnifying glass by means of a (**1**, 10)-ES, whereby the lens was composed of 10 segments. The numbers in parentheses next to the number of generations indicate the amount of scattering.

3.2. Evolution of a video cable tree

Cable television comes to Hometown, but the cable ends at the Town Square. The first selectman charges the town surveyor to develop a plan for a least expensive fiber optic cable tree to all houses. Knowing that the shortest distance between two points is a straight line the surveyor designs a star network (Fig. 13(a)).

An evolution-strategist however believes that he can design a better solution and develops a variable net topology, consisting of nodes with three branches each. These nodes can be moved in a x–y coordinate system. Let be Δx_i and Δy_i the coordinate differences of the ends of the line i. Then the minimization problem can be formulated by

$$Q = \sum_{\text{all line elements } i} \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2} \rightarrow \text{min.}$$

At the start of the ES optimization all nodes overlap, corresponding thus to a star network. The minimum expense network developed with the (**1**, 10)-ES needs 40% less fiber cable than the start solution (Fig. 13(b)).

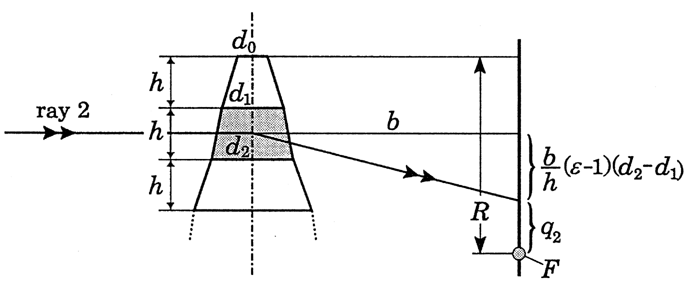

Fig. 11. Path of a light beam through a glass prism.

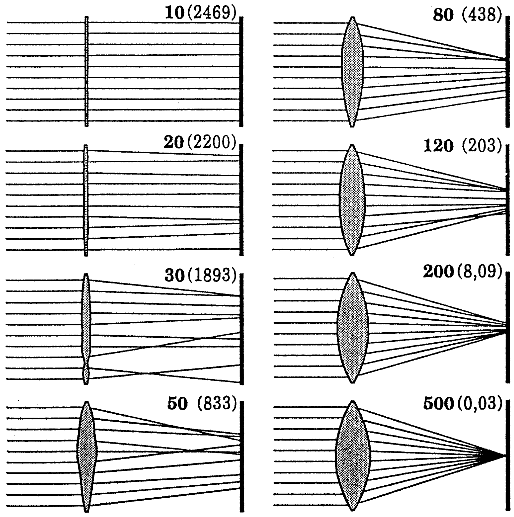

Fig. 12. From window pane to magnifying glass. The evolutionary computation needs 15 s on a Pentium II PC.

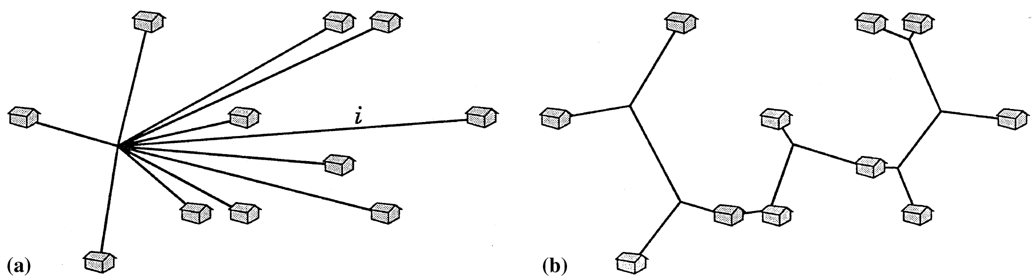

Fig. 13. (a) Videocable to the houses of Hometown. Evolutionary computation starts with a radial network. (b) Evolution-strategic minimum cable network of Hometown.

It should be pointed out that the form of the minimum solution depends on the basic structure of the network. The sequence by which the houses will be 'mathematically' connected cannot be changed during the optimization process. In order to find the global optimum the ES-optimization process must be repeated for other connecting topologies. Even nature has difficulties with optimizing a topology; e.g. the basic fish skeleton was not changed during the development of land animals. Evolution did not invent a new biological supporting structure but only modified the existing one. No one can convince me that the supporting biological structure of aquatic animals is the best solution for upright land animals.

3.3. ES design of a truss bridge

A truss structure is a load bearing structure consisting of rods flexibly connected at the joints. A student Holger Eggert based his program on a flat static truss bridge. The load distribution has been calculated according to the Ritter method. The objective is to design the cross sections of the round solid rods so that

- the rods in tension do not exceed the maximum permissible tensile stress, and
- the rods in compression are loaded just below the Euler buckling point.

Under such boundary conditions for the rods the total weight of the structure can be calculated. The goal is to find a solution with minimum weight. Similarly to the cabling network problem the joint positions are adjustable in the x–y coordinate system and the selected joint numbering constitutes an unchangeable connecting structure.

The objective of this computer experiment was the development of a minimum weight bridge. The two supporting joints determine the span of the bridge. The road (lower girder) is divided into 11 joints movable

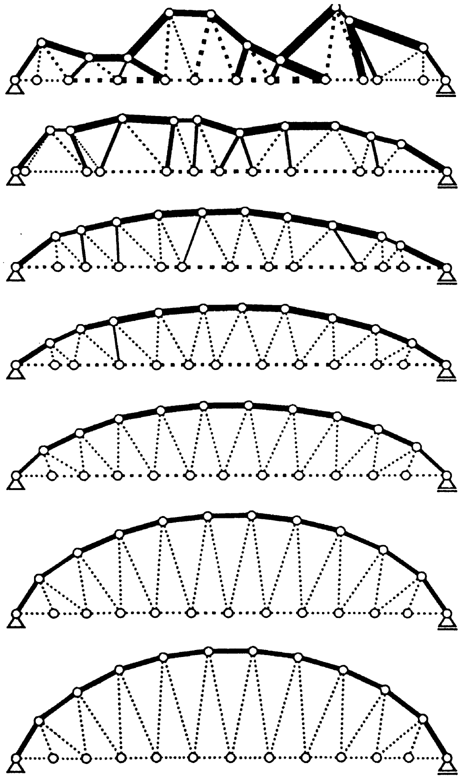

Fig. 14. Evolution of the form of an arched bridge with minimum weight. Solid lines are compression columns and dashed lines are tension bars.

only in the x direction. The 10 upper joints can be moved both in the x and the y direction. The problem is, therefore, defined by 31 variables. The load on the bridge is determined by the vehicles on the bridge and is defined by the civil engineer as equally distributed load. The own weight of the bridge is not taken into account separately. A (**10**/10, 100)-ES was used for the computer run as depicted in Fig. 14. The weight of the bridge construction is shown below the number of generations.

3.4. ES-optimization of a hull shape

What is the shape of a streamlined rotational body with least resistance? William E. Pinebrook [6] solved this problem with a classical $(\mathbf{1}+1)$-member ES with 1/5 success rule. The circular cross section and the length of the spindle are given. The pressure distribution is calculated in the first pass for every evolution-strategic generated variant. In the second pass the development of the boundary layer next to the wall is calculated. Pinebrook assumes realistically that, despite the stabilizing effect of the pressure drop, at 3% of the flow length turbulence will occur. The boundary layer computation yields the wall shear stresses. Summing these local friction forces over the entire surface gives the total friction force. In addition the pressure distribution is affected by the displacement thickness of the boundary layer. The pressure distribution of the potential flow, with a resultant zero force in flow direction, will change and an additional flow resistance will develop. The sum of these two drag components has to be minimized. Fig. 15 shows the development of the hull shape starting with the 'blimp-shape' and reaching the shape of a 'dolphin spindle'.

Some comments on the result: since a surface generates friction, one of the approaches to optimization consists of minimizing the body surface. Focusing only on this aspect, an intuitive solution can be found by pushing a rod through the center of a disk representing the body cross section. After dipping this construct into a soap solution the soap film will assume a surface of least area. However, a kink would develop at the 'midship-section'. The evolution-strategic optimization will smooth the surface. We want to think further:

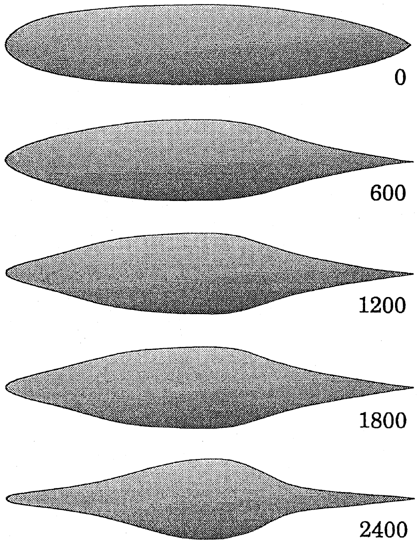

Fig. 15. 2400 generations for evolution of a hull spindle with least resistance. The circular cross section and the length of the spindle are given.

The diameter increasing in front of an ordinary body of revolution leads to a thin boundary layer through the acceleration of the flow. A thin boundary layer, however, produces large friction. Thus, would it be favorable to make the boundary layer thicker by a front-contour without flow acceleration (e.g. by a constant-diameter-column)? But even this would not work because thicker boundary layers distort the resistless pressure distribution of the potential flow so that a resultant pressure force remains. This small excursion into the details of flow theory shows how many complex compromises must be found for solving this optimizing problem. The novel spindle shape in Fig. 15 may represent such a compromise.

Recently Thorsten Lutz [5] presented the results of evolutionary computations of airship bodies with minimum drag for a given volume. For small Reynolds numbers (small body size) he obtained the well-known characteristics of a laminar flow body (Fig. 16(a)).

If the Reynolds number increases the optimal airship hull becomes more compact (Fig. 16(b)). The reason for this is that the flow has to be accelerated more in order to stabilize the laminar boundary layer in front of the body. For an even higher Reynolds number, valid for large airships, a very strong acceleration of the flow is needed to prevent premature transition of the laminar flow to turbulent flow. This leads to the pointed form looking quite unusual for an airship hull (Fig. 16(c)).

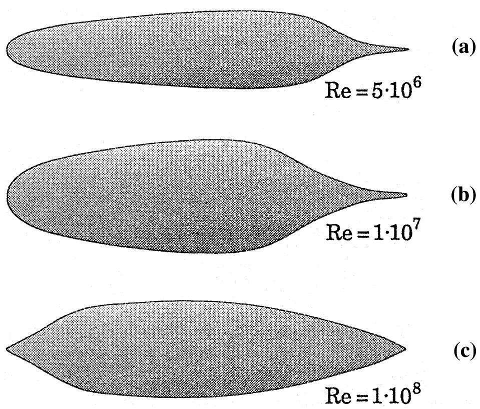

Fig. 16. Results of an evolutionary computation. Airship hulls with minimum drag for a given volume at different Reynolds numbers.

A typical ES-optimization run ends with large fluctuations of the resistance, because the optimization process always tries to increase the contour thickness just as gentle as it is needed for a stabilization of the laminar flow. This means, however, that a small mutation can shift the transition point abruptly to the front of the body, resulting in a sudden rise of the resistance. Finally it has to be mentioned that ES-optimization of the airship-hull needs a high sophisticated ES including covariance adaptation (see [2]).

3.5. Evolution of the trajectory of a stone throw

In 1744, the French scientist Pierre-Louis Moreau De Maupertius was convinced that he had discovered an all-embracing world plan, namely that nature proceeds with utmost thriftiness. Applied to a stone throw (of mass m) from a point a to a point b this principle can be expressed by

$$\int_a^b mv\,\mathrm{d}s \to \min .$$

It was Leibniz and Euler, who expressed similar thoughts before Maupertius. ES should be able to solve the problem of minimization, which nature does for every stone throw. Let the kinetic energy of the stone at the starting point be given by E_0. Then the conservation of energy in a gravitational field leads to the equation

$$\int_a^b m\sqrt{2(E_0 - mgy)/m}\,\mathrm{d}s \to \min .$$

For a numerical solution of the problem with ES polygon segments approximate the trajectory of the stone. The x and y values of the 10 polygon corners supply 20 variables. Including displacement of the x coordinates permits the generation of loops in the trajectory. However, uniform x increments of the benchmarks are rewarded. Fig. 17 shows how a trajectory, impossible according to physics, can be converted into the well-known parabola by a (**1**, 10)-ES.

One should remember that the history of physics records many extremum principles of the laws of nature:

1. The principle of Torricelli, according to which the center of gravity of a moving system assumes the lowest position in gravity field.
2. The principle of minimum energy of an elasto-static system in equilibrium.
3. The principle of Fermat according to which a ray of light always chooses the path of minimum time.
4. The principle of least coercion by Gauss, minimizing the deviation of the forced motion from the free motion.
5. The principle by Hertz according to which a motion chooses the smallest curvature (principle of the straightest path).

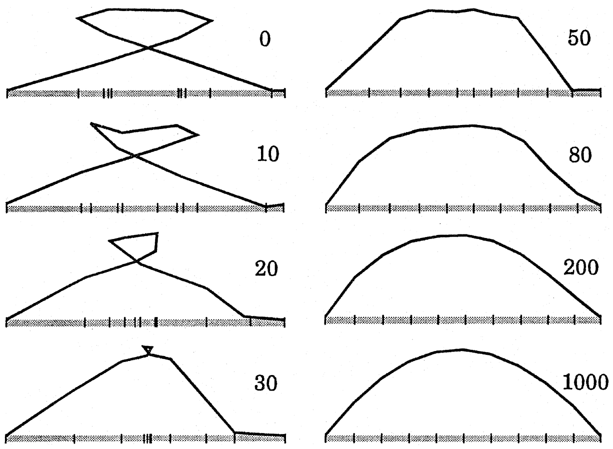

Fig. 17. Stone throw trajectory. ES-minimization of the action integral deforms the absurd loop trajectory to the well known parabola.

6. The principle of Hamilton according to which the time integral of the Lagrange function assumes a maximum, a minimum or a saddle point.

If one assumes that the speed and power of future computers will further increase one could speculate that many complex scientific problems would be solved by an evolution-strategic treatment of the corresponding extremum principle. The Hamilton principle can be applied to classical as well as quantum mechanical processes (chemical, thermodynamic and electrodynamic processes) and, therefore, offers a special advantage.

3.6. The largest small hexagon after Graham

What is the largest surface area, which a hexagon can assume if no two corners have a larger distance from each other as the unit distance? Ronald R. Graham [1] gave the answer. The optimum form reminds one of the silhouette of a diamond crystal. The surface area amounts to 0.674981, compared to the regular hexagon, which has an area of 0.64958.

It is a challenge for the evolution strategist to solve this problem for other polygons. The task is simple. The surface area of a polygon has to be maximized under the constraint that all corner distances are $\leqslant 1$. Fig. 18 shows the solutions for an octagon and a decagon, compared with Graham's solution for a hexagon on the left-hand side.

The evolutionary computer runs showed that for polygons with an odd number of sides the solution is always an equilateral polygon. For increasing number of sides of even polygons the solutions also approach the regular form.

3.7. Fermat's last theorem and Euler's supposition

The following function has to be minimized when x_1 to x_5 are natural numbers:

$$Q = \left|x_1^6 + x_2^6 + x_3^6 + x_4^6 + x_5^6 - x_6^6\right| \to \min.$$

Who solves this minimization problem would get much attention under the mathematicians. The exact solution of the above equation means that the sum of five natural numbers raised to the power of six results in a sixth natural number raised to the power of six.

$$x_1^6 + x_2^6 + x_3^6 + x_4^6 + x_5^6 = x_6^6, \quad x_1, x_2, \ldots, x_6 = \text{natural numbers}.$$

It was Pierre De Fermat, who stated that for instance the integer equation $x^6 + y^6 = z^6$ has no solution. Such form of an equation can be solved only for the power of two. (Andrew Wiles gave 1997 the definite proof for this supposition). Leonhard Euler formulated the extended assumption that even the equations below never exist.

$$\begin{aligned} &x_1^4 + x_2^4 + x_3^4 = x_4^4 \\ &x_1^5 + x_2^5 + x_3^5 + x_4^5 = x_5^5 \qquad\qquad x_i = \text{natural numbers}, \\ &x_1^6 + x_2^6 + x_3^6 + x_4^6 + x_5^6 = x_6^6 \\ &\ldots \end{aligned}$$

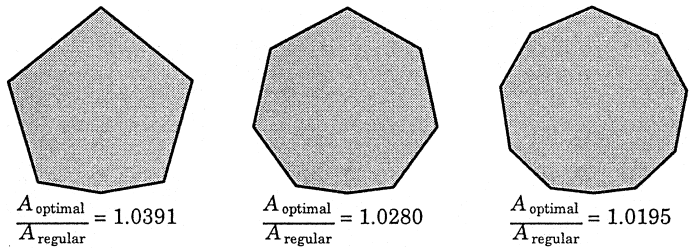

Fig. 18. The largest small hexagon, octagon and decagon.

But Euler has been mistaken. Frye showed 1988 by computer trials that the following sum of three fourth powers yields a forth fourth power:

$$95800^4 + 217519^4 + 414560^4 = 422481^4.$$

And 1966 a corresponding equation for the fifth power was discovered by Lander and Parkin:

$$27^5 + 84^5 + 110^5 + 133^5 = 144^5.$$

Is it possible to find a solution for the power of six through a minimization of the designed objective function? Is it feasible to use the ES for this task? A first exploration showed a difficulty. The deviation of Q from zero gets always small when the numbers x_i become small. Therefore the objective was normalized in the following form:

$$Q = \frac{|x_1^6 + x_2^6 + x_3^6 + x_4^6 + x_5^6 - x_6^6|}{x_1^6 + x_2^6 + x_3^6 + x_4^6 + x_5^6 + x_6^6} \rightarrow \min.$$

One realizes that for a given number x_6 a set of numbers x_1 to x_5 does exist, which fulfills the equation quite well. But the sum doesn't open exactly. One has the feeling that the problem behaves extremely multimodal. I spend two month on a PC to find a solution. Due to the expected multimodality of the objective a $[\mathbf{1}, 4(\mathbf{1},100)^{100}]$-ES was applied. But I failed. Below the best results are given for the power of six and seven:

$$67^6 + 124^6 + 456^6 + 884^6 + 1327^6 = (1346.00000000004163\ldots)^6,$$

$$20^7 + 38^7 + 83^7 + 83^7 + 488^7 + 881^7 = (883.000000000002135\ldots)^7.$$

It is a challenge for all disciples of evolutionary algorithms to disprove Euler's supposition with their ingenious ideas. According to Darwin, however, the problem could prove to be not evolvable. I could announce merrily that I have received three times the exact solution of Lander and Parkin applying a nested ES. But the search was restricted between 1 and 200 for each x_i. In this case the probability to hit the (known) solution is $5!/200^5 = 3.75 \times 10^{-10}$. The ES proved to be 10 times faster than a blind search. Suppose that for the power of six a solution exists with all $x_i \leqslant 1000$. Then probably $(1000^6/6!)/10 = 1.39 \times 10^{14}$ function evaluations are needed to find the solution. Therefore the example demonstrates that the ES is not able to solve exactly the general multimodal optimization problem. I do not believe that any other Evolutionary Algorithm will make this better. It is self-evident that an Evolutionary Algorithm cannot overcome the problem of 'NP hardness'. But it must be noticed that the ES is able to find a rather good solution.

3.8. Evolution of a magic square

The last problem seems to have the same complexity as the previous. It is also a problem of natural numbers. A magic square is a special structure of natural numbers, which are arranged in a matrix so that their sums horizontally, vertically and diagonally become equal to S. The sum S is called the magic sum. The 'Chinese square' is about 4000 years old. It is represented by a 3 by 3 square with all numbers from 1 to 9 resulting in the magic sum of $S = 15$. Albrecht Duerer depicted in his engraving 'Melancholy' a square of 4 by 4 cells with the magic sum $S = 34$ for the numbers 1 to 16. It is customary to expect a magic square to use consecutive natural numbers.

In order to develop a magic square with the help of the ES a quality function Q must be constructed. Concentrating on two different squares A and B a decision must be reached whether A is better than B or B better than A or A and B equally good. A function, which permits such a decision for a 3 by 3 square, is expressed by:

n_1	n_2	n_3
n_4	n_5	n_6
n_7	n_8	n_9

$$
\begin{aligned}
Q = {} & (n_1 + n_2 + n_3 - 15)^2 (n_4 + n_5 + n_6 - 15)^2 + (n_7 + n_8 + n_9 - 15)^2 \\
& + (n_1 + n_4 + n_7 - 15)^2 (n_2 + n_5 + n_8 - 15)^2 + (n_3 + n_6 + n_9 - 15)^2 \\
& + (n_1 + n_5 + n_9 - 15)^2 (n_3 + n_5 + n_7 - 15)^2 \rightarrow \text{min}.
\end{aligned}
$$

One must take care, that none of the numbers between 1 to 9 is lost by a mutation. Therefore the mutation rule reads:

- Pick randomly an integer from a first cell.
- Mutate this number within the restricted range.
- Find this new integer in a second cell.
- Exchange the integer numbers of these two cells.

For a small magic square the best mutation step size is the smallest variation ± 1, which guarantees strong causal behavior. However, Michael Herdy [3] generated 100 by 100 squares with every number from 1 to 10 000 appearing only once. To reach the solution in shortest time mutative step-size control must be used. It means that at the beginning more than one cell will be mutated (exchanged) and that the variation is larger than ± 1. Herdy introduces an additional feature in the square. He hides special numbers in the square (birth date, credit card ID or Duerer's magic square). Special algorithms are not known to generate magic squares for this task. A 98s Pentium PC requires about 2 h to evolve a 100 by 100 magic square under these conditions.

39	104	224	65	98	115	90	127	93	123	97	84	145	104	112	56	48	140	70	65
111	119	140	109	129	85	99	128	77	179	71	65	139	42	117	64	114	61	86	64
49	99	225	37	62	126	78	126	136	103	146	66	82	6	112	102	148	107	125	64
110	117	161	97	149	41	142	178	69	82	113	73	107	62	107	102	55	127	56	51
69	24	109	83	112	121	108	86	133	111	63	51	115	72	126	132	71	176	102	135
120	70	171	64	145	101	64	79	61	109	95	55	100	59	147	116	124	142	60	117
132	150	1	147	152	9	9	205	159	92	9	9	209	134	141	9	9	208	133	82
147	1	1	88	9	185	219	9	121	9	120	294	9	130	9	227	184	9	128	160
69	115	1	141	9	136	190	9	147	9	215	202	9	94	9	194	198	9	105	138
88	136	1	158	9	132	156	9	155	9	139	249	9	145	9	172	180	9	114	120
200	206	1	166	171	9	9	9	170	245	9	9	9	185	215	9	9	9	181	178
119	118	1	87	154	100	104	9	117	128	90	159	9	157	139	105	174	9	126	94
84	160	1	156	158	140	78	9	87	181	123	154	9	74	108	131	78	9	50	109
185	160	1	169	9	9	9	228	205	9	9	9	211	176	9	9	9	258	145	180
90	57	158	88	89	122	94	108	38	123	136	73	155	81	150	100	114	54	112	57
60	106	130	53	80	117	122	138	64	52	66	85	140	94	138	115	140	185	50	64
85	49	201	104	87	79	47	106	53	99	175	63	165	47	107	127	99	130	91	85
32	75	141	32	129	187	137	116	39	93	113	119	161	144	84	33	86	131	72	75
69	55	131	92	102	90	140	161	38	107	129	111	85	98	119	72	65	98	180	57
41	78	200	63	146	95	104	159	37	138	81	129	131	95	41	124	94	128	13	104

Fig. 19. Evolution of an extraordinary 20×20 magic square with the magic sum of 1999.

Suppose that we have to put more information in the square. The condition then must be dropped that every number appears only once. We get an improper magic square. Fig. 19 depicts such a square which shows the year 1999 in the center and which has the magic sum of 1999. For the turn of the millennium I have already completed the corresponding magic square. Both times a $[\mathbf{1}, 4(\mathbf{1}, 20)^{100}]$-ES has been used. The normal ES has been designed for real variables. The algorithm has not to be changed for integer optimization. One may simply replace 'n_i' by 'floor(x_i)' in the quality function, where x_i are real positive numbers. The ES is robust enough to climb up the stepped hill.

4. Conclusion

Disciples of evolutionary algorithms could criticize that the above examples (except the Euler-problem) are too simple. I am frequently asked whether a problem solved with an evolutionary algorithm could not be solved faster with a deterministic method. The information science distinguishes between tractable and not tractable problems. A problem which can be solved with an evolutionary algorithm belongs to the class of tractable problems. And the more intensively we deal with such a problem, the more skillfully we can make the strategy. Herdy [3] has shown that Rubiks cube having 6×6 elements on each side can be arranged with the ES. Nobody can buy such a cube today. But if a 6×6-cube would be for sale, a strategist would surely design a method to arrange the cube much faster than the ES will do.

We look now to the world of hard problems (NP-hard, NP-complete). There is no hope to solve the Hamiltonian path problem in polynomial expected time applying an evolutionary algorithm. No 'deterministic' strategy including the ES will do this. Certainly this is not a new statement. However, we may hope that an evolutionary algorithm, unable to solve the problem exactly, nevertheless finds a good solution. And we may further hope that this is done faster as if we work grimly to design the deterministic strategy which fits the special problem.

References

[1] R.L. Graham, The largest small hexagon, Journal of Combinatorial Theory, Series A 18 (1975) 165–170.
[2] N. Hansen, A. Ostermeier, Adapting arbitrary normal mutation distribution in evolution strategies: The covariance matrix adaptation, in: Proceedings of the 1996 IEEE International Conference on Evolutionary Computation, 1996.
[3] M. Herdy, Theory and application of the evolution strategy, Doctor-thesis, Technische Universität Berlin (in preparation), 1999.
[4] J. Klockgether, H.P. Schwefel, Two-phase nozzle and hollow-core jet experiments, in: Proceedings of the 11th Symposium on Engineering Aspects of MHD, Cal. Inst. Techn., 1970, pp. 141–148.
[5] T. Lutz et al., Shape Optimization of Axisymmetric Bodies in Incompressible Flow, Institut für Aerodynamik und Gasdynamik, Universität Stuttgart, 1996.
[6] W.E. Pinebrook, C. Dalton, Drag minimization on a body of revolution through evolution, Computer Methods in Applied Mechanics and Engineering (1983) 179–197.
[7] I. Rechenberg, Cybernetic solution path of an experimental problem. Royal Air-craft Establishment, Library Translation 1122, Farnborough, Reprint in: D.B. Fogel (Ed.), Evolutionary Computation, The Fossil Record, IEEE Press, Piscataway, NJ, 1965, pp. 301–309.
[8] I. Rechenberg, Evolutionsstrategie – Optimisierung technischer Systeme nach Prinzipien der biologischen Evolution, Frommann-Holzboog, Stuttgart, 1973.
[9] I. Rechenberg, Evolutionsstrategien. In: B. Schneider, U. Ranft (Eds.), Simulationsmethoden in der Medizin und Biologie, Springer, Berlin, Evolutionsstrategie '94, Frommann-Holzboog, Stuttgart, 1978.
[10] I. Rechenberg, Evolutionsstrategie 94, Frommann-Holzboog, Stuttgart, 1994.
[11] H.P. Schwefel, Experimentelle Optimierung einer Zweiphasendüse, Bericht 35 des AEG-Forschungsinstituts Berlin zum Projekt MHD-Staustrahlrohr, 1968.
[12] H.P. Schwefel, Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie, Birkhäuser, Basel, 1977.
[13] H.P. Schwefel, Numerical Optimization of Computer Models, Wiley, Chichester, 1995.