



Ubiquity Symposium

Evolutionary Computation and the Processes of Life

Life Lessons Taught By Simulated Evolution

by Hans-Paul Schwefel

Editor's Introduction

In this second article in the ACM Ubiquity symposium on evolutionary computation Hans-Paul Schwefel explores the effects of simulating evolutionary mechanisms.

Editor

Ubiquity Symposium

Evolutionary Computation and the Processes of Life

Life Lessons Taught By Simulated Evolution

by Hans-Paul Schwefel

Evolutionary algorithms were not only born as bio-inspired methods to improve or even optimize designs and processes simulated on digital computers, they should also serve as means to improve our understanding of nature's way to maintain and further develop living beings. I started into the field with a deep belief that better understood mechanisms of natural evolution would lead to improved optimization tools. Here is what I have already learnt during my search for understanding life by simulating some evolutionary principles.

According to evolutionary biologists, natural evolution works by means of five constitutive factors: heredity, change, surplus production, selection, and moribundity. Finite life span and thus death or forgetting is important as well as birth surplus. These two factors are not incorporated in all evolutionary algorithms (EA), but important as I have learned from my computer experiments. The three other factors are components of all EA, in different interpretations, however. Since the 1960s, three main approaches of simulating evolutionary factors have evolved, i.e., evolutionary programming [1, 2], genetic algorithms, [3, 4], and evolution strategies [5, 6, 7, 8]. All of these (and more) attempts to use paradigms gleaned from natural processes have been used for industrial improvement if not even optimization according to observations like those of Rosen [9], who found optimality principles in biology.

For a detailed description and history there exists a handbook of evolutionary computation, now one of the three sub-disciplines of computational intelligence—artificial neural networks, fuzzy logic, and evolutionary computation—edited by Bäck, Fogel, and Michalewicz [10].

Whether recombination or mutation is (more) important in hereditary change from one to the next generation has been controversial. The same is true for the reason why recombination can be beneficial. The reason why it can be harmful in multiobjective situations is obvious. To see this just imagine a task with only two variables and two objectives, where the parent individuals have already reached positions on or near the Pareto set (not front) being out of square or a curved piece of a line. Dominant recombination, where an offspring gets one character from one parent, the second one from the other, will always lead to an inferior solution. The same is true for intermediary recombination if the Pareto set is not a straight line.

Widely accepted is the view that mutation is an important factor in nature as well as in evolutionary optimization. The main question here is, how strong mutations should be. If changes are small, often nearly every second trial is successful—the success also being modest, however. Stronger mutations permit larger improvements, but more seldom. As Rechenberg has already shown, there is an “evolution window” within which the long-term mean success is (near) maximal [5]. The width of that window is about one decade, he found in two very different artificial situations (inclined ridge and sphere in high dimensional search spaces), and at its center the probability of success is about $1/5$. His formulae for maximal progress contain another variable, i.e., the distance to the peak (or bottom) of the hypersphere or the width of the corridor respectively or, more generally, the curvature radius of the landscape. Moreover, both the mutation strength and the progress rate are inversely proportional to the number of parameters that determine the individual performance or fitness. It is clear, therefore, that the mutation strength has not only to be adjusted once, but has to be adapted again and again during the search process. At the optimum it should vanish in order to keep the best position. If environmental changes occur mutations must eventually become stronger again, and after catastrophes evolution must perhaps start once more from scratch—at least in case of monoploidy. Polyploidy might help with information stored from previous situations.

How do living beings find the right mutation strength, i.e., mutation rate(s) for binary variables, standard deviation(s) of normally distributed changes for continuous variables or of geometrically distributed changes for integer/discrete variables?

Conventional optimization methods gather information in the vicinity of the current search position or from the history of the search process and use that information in order to place the next trials so that improvements become likely and as large as possible. For that purpose an

underlying model of the landscape is a prerequisite, seldom more sophisticated than linear (gradient type methods) or quadratic (Newton or quasi-Newton methods).

A knowledgeable experimenter would at least count trials, successes, and failures and adjust the variation strength according to Rechenberg's 1/5 success rule. It is striking that some commercial branches obey that rule without having ever heard of Rechenberg's theoretical results for the simplest, that is the two-membered or (1+1) evolution strategy with just one parent plus one descendant per generation or reproduction cycle, the better (in case of fitness equality the newer) of which “survives” and becomes parent of the next generation. In former times 80 percent of the disk records were recycled, because nobody could be attracted to buy them. It is said that only 20 percent of all catwalk presented fashion goods are ordered finally, and up to 85 percent of new foodstuffs disappear from the shelves of supermarkets within the first year. Such markets are groping in the dark like black-box optimizers, because consumer preferences are unpredictable. But the 1/5 (and even a 1/50) rule is no longer helpful in solving problems with active inequality constraints as well as such without constraints but with fitness level contours that form one or more sharp vertices. The sector open for possible improvements then is smaller than a half space; it may be tiny, indeed. Thus, one has to find a better means of adapting mutation strengths and other internal parameters of the search process, if it is to become reliable in more general situations.

Without an external observer, interpreter, and calculator, nature has established self-adaptation by adding control parameters into the genome, e.g., in the form of repair enzymes that are able to reduce (more or less) replication errors caused by too destructive premutations. Such enzymes are encoded in the so-called introns, and they underlie the same processes of change as the object variables or exons, which alone define the phenotype and thus determine the fitness. Within a population those individuals with proper introns should be more successful than others with inferior “internal models” of the environment. In principle, repair enzymes could prevent all mutations by repairing all pre-mutations, but they don't, because no more changes could occur and adaptation to environmental changes or evolution would no longer take place. Mutations are necessary for both exons and introns. This kind of two-level evolution works quite well if and only if some crucial conditions are fulfilled:

- sufficient population size,
- diversity of internal and object parameters,
- sufficient birth surplus,
- medium selection pressure, and

- finite life span.

Otherwise, the evolutionary process can run into a stalemate or conduct punctuated equilibria as Niles Eldredge and Stephen Jay Gould have termed the phenomenon [11]. Natural evolution is by no means as gradual as most encyclopaedias make their readers believe. There are many internal reasons for a bumpy instead of smooth way of development in addition to external ones as I am going to show in the following.

Let me explain all that by making use of the multimembered or (μ, λ) evolution strategy (ES) with μ parents producing λ offspring within one generation. The genome of each individual contains n object variables defining the fitness and at least one internal variable for the mutation strength or better up to n individual mutation strengths (one for each object variable (exon) or even additional introns for correlations between object variables. Since even simple linear correlations afford up to $O(n^2)$ introns it is no wonder that their number by far exceeds the number of exons. Now, in the case of real variables let the internal parameters be the standard deviation(s) or mean step sizes for mutating the real-valued object variables. The standard deviation(s) themselves are mutated by means of log-normal multipliers in order to avoid a bias toward larger or smaller mutations without necessity caused by natural selection.

A (10,100)-ES with one common standard deviation—now called (mean) step size for simplicity—works very well by both decreasing the step size when approaching the minimum (or maximum) and increasing it if appropriate, for example when the optimum is moving away. In this case of a so-called comma strategy, only the 10 best of the 100 offspring are used as parents for the next generation. The birth surplus is essential. A (10,10)-ES would generate a random walk in the whole search space. An elitist version like a (10+10)-EA, where the 10 best of parents plus offspring “survive,” may work with fixed or externally controlled mutation strength(s), but it is not suitable for the self-adaptation of internal parameters. Even an elitist (10+100)-ES version is slowed down sometimes by stagnation phases due to parents having achieved very good performance but with internal parameters that are not well suited for further progress. With a finite life span (here one generation or reproduction cycle only) and a small population this leads only to a short recession instead of a longer stagnation, and overall the search is not only more continuous but also quicker in the long term. Especially dynamic situations with moving optima—the usual situation in nature where any innovation of one species may change the fitness landscape for others like on a trampoline— require the

forgetting of interim solutions that have become defective. Programmed finite life span helps for quicker adaptation to changing environments and for any kind of progress. It may be seen as trick of nature if not even a defining feature of life. Already Wolfgang Goethe presumed lethality is an advantage for and no insufficiency of life.

If less than 10 out of 100 descendants survive, the learning of internal parameters is hampered, especially when up to n possibly different step sizes for the n object variables have to be adapted properly. With increased selection pressure it becomes profitable for the offspring to look for progress in lower-dimensional regions of the search space by sharply reducing some of the step sizes, because theoretical results tell us that progress rates are inversely proportional to the number of (effective) search dimensions, at least for a while until a relative optimum in subspace is reached. Then the vector of introns must be restructured to take up the search into the forgotten directions again. Such lightweight searchers can be avoided by lowering the selection pressure, so that the diversity of internal models of the searchers is maintained, sometimes including even necessary lateral thinkers. Such process needs collective intelligence to become successful. The intelligence is not concentrated in the overall best descendant as has been shown in [12].

Survival of the fittest, thus, is a bad advice and should not be taken literally as a recipe for the search of improvements. If really only one best offspring is taken as parent for creating all individuals of the next generation, which would be the case in an extreme (1,100)-ES, self-adaptation of the internal parameters would break down. Instead of Herbert Spencer's buzz phrase, Alfred Russel Wallace's characterization of natural selection as "extinction of the worst" should be used. Obviously lethality sells badly, even to Darwin himself, who finally took over Spencer's version in the third edition of his seminal work.

By the way, roulette wheel selection, as canonically used within the genetic algorithms frame, in which the number of mating events and thus descendants is proportional to the fitness of parents, is highly problematic. Moreover, it resembles mating selection rather than Darwinian natural selection. Laboratory animal researchers report the strongest individuals often burn out themselves in fights instead of engaging in mating and procreation. Natural selection can be seen as fight for survival, but rather against predators, parasites, and generally environmental conditions than against other individuals of their own species. This fight is sometimes modeled as tournament selection, e.g., in evolutionary programming, where the "individuals" are

thought of as species (therefore handled without recombination), but recently also in genetic algorithms applications, so that the buzz phrase “struggle for life” is taken literally as intra-species competition.

The considerations above contain only parts of the insight gained with multi-membered evolution strategies. Further extensions of the basic (μ, λ) -ES have been discussed —e.g., the somehow less “cruel” $(\mu, \kappa, \lambda, \rho)$ -ES with individuals having a maximum lifespan of $\kappa > 1$ reproduction cycles, where ρ denotes the recombination type [13,14]. It is much harder to be analyzed theoretically, however. Gene deletion and gene duplication for a variable number of variables, multi-cellularity in combination with somatic mutations during ontogeny of individuals for the case of binary variables, diploidy for various purposes, even two genders underlying different selection criteria for approaching narrow vertices formed by multiple active constraints, and above all predator-prey processes for multicriteria optimization have been found to be of help in such difficult decision making situations. The hope that one population spreads along a Pareto front may be futile. A multi-species approach with several local predators chasing the prey according to their own single criteria might turn out to be especially appropriate for asynchronous implementation on parallel computers.

Summary

There are several unexpected lessons I learned from simulating evolutionary mechanisms:

- Under unpredictable circumstances not even every second trial can be successful. The failure rate is at least as high as eighty percent, sometimes higher, for optimal progress. Some business areas have witnessed this fact.
- The evolutionary process can comprise stagnation and recession phases instead of being gradual. The fossil record is witness.
- Finite life span of individuals is a necessity as well as imperfect gene repair mechanisms for maintaining diversity and evolvability.
- Wisdom for advancement is not concentrated in the fittest individual, but distributed; natural intelligence is a collective phenomenon.
- Self-adaptation of internal parameters and rules requires large enough populations, birth surplus, medium selection pressure, and many more introns than exons.

In addition to these facts I learned their algorithmic reasons.

References

- [1] Fogel, L. J., Owens, A. J., Walsh, M. J. *Artificial Intelligence through Simulated Evolution*. Wiley, New York, 1966.
- [2] Fogel, D. B. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*. IEEE Press, New York, 1995.
- [3] Holland, J. H. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor. MI, 1975.
- [4] Goldberg, D. E. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, MA, 1989.
- [5] Rechenberg, I. *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. Frommann-Holzboog, Stuttgart, 1973.
- [6] Rechenberg, I. *Evolutionsstrategie '94*, Frommann-Holzboog, Stuttgart, 1994.
- [7] Schwefel, H.-P. *Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie*, Birkhäuser, Basel, 1977.
- [8] Schwefel, H.-P. *Evolution and Optimum Seeking*, Wiley, New York, 1995.
- [9] Rosen, R. *Optimality Principles in Biology*. Butterworths, London, 1967.
- [10] Bäck, Th., Fogel, D. B., and Michalewicz, Z. (eds.) *Handbook of Evolutionary Computation*. Institute of Physics Publ., Bristol and Oxford University Press, New York, 1997.
- [11] Gould, S. J., and Eldredge, N. Punctuated equilibria: The tempo and mode of evolution reconsidered. *Paleobiology* 3, (1977), 115-151.
- [12] Schwefel, H.-P. Collective intelligence in evolving systems, pp. 95–100. In *Ecodynamics, Contributions to Theoretical Ecology*. W. Wolff, C. J. Soeder, and F. Drepper (eds.), Springer, Berlin, 1988.
- [13] Rudolph, G., and Schwefel, H.-P. Simulated evolution under multiple criteria conditions revisited, pp. 249–261 in *Computational Intelligence – Research Frontiers*, J. M. Zurada, G. G. Yen, and J. Wang (eds.). Plenary and Invited Lectures at the IEEE World Congress on Computational Intelligence (WCCI 2008) at Hong Kong. Springer, Berlin, 2008.

[14] Schwefel, H.-P., and Rudolph, G. Contemporary evolution strategies. pp. 893–907. In *Advances in Artificial Life – Proc. Third European Conf. Artificial Life (ECAL'95)*, F. Morán, A. Moreno, J. J. Merelo, and P. Chacón (eds.). Springer, Berlin, 1995.

About the Authors

Hans-Paul Schwefel studied aero- and space-technology at the Technical University of Berlin (TUB). His diploma thesis of March 1965 was the first work on computer experiments with the German version of an evolutionary algorithm. After 10 years in industry and different fields of academia he received his Dr.-Ing. degree from TUB with a dissertation on *Evolutions strategie und numerische Optimierung*. He then acted as head of a computer aided systems analysis tools group at a large research center. From 1985 to 2006 he was the first holder of a chair for systems analysis at the faculty of computer science of the University of Dortmund. He was cofounder of the international conference series “Parallel Problem Solving from Nature” (PPSN), dean of the faculty, spokesman of the German collaborative research center on computational intelligence, cofounder and president of the Informatics Center Dortmund, and pro-rector for research and scientific staff advancement at the university. He has been member of the editorial boards of three journals and advisory board member of two book series in the field of evolutionary respective natural computation. His publication list comprises more than 170 entries. In 2002 he won an Evolutionary Computation Pioneer Award from the IEEE Neural Networks Society. He was elevated to Fellow of the IEEE in 2007. In the same year The University of Birmingham admitted him to the degree of Doctor of Science, *honoris causa*. In 2011 he received the IEEE Frank Rosenblatt Award sponsored by the IEEE Computational Intelligence Society.

DOI: 10.1145/2369296.2369297