A history of evolutionary computation

3 authors:

Kenneth De Jong
George Mason University

David B. Fogel
Natural Selection, Inc.

Hans-Paul Schwefel
Technische Universität Dortmund

A2.3 A history of evolutionary computation

Kenneth De Jong, David B Fogel and Hans-Paul Schwefel

Abstract

This section presents a brief but comprehensive summary of the history of evolutionary computation, starting with the ground-breaking work of the 1950s, tracing the rise and development of the three primary forms of evolutionary algorithms (evolution strategies, evolutionary programming and genetic algorithms), and concluding with the present time, in which a more unified view of the field is emerging.

A2.3.1 Introduction

No one will ever produce a completely accurate account of a set of past events since, as someone once pointed out, writing history is as difficult as forecasting. Thus we dare to begin our historical summary of evolutionary computation rather arbitrarily at a stage as recent as the mid-1950s.

At that time there was already evidence of the use of digital computer models to better understand the natural process of evolution. One of the first descriptions of the use of an evolutionary process for computer problem solving appeared in the articles by Friedberg (1958) and Friedberg *et al* (1959). This represented some of the early work in machine learning and described the use of an evolutionary algorithm for *automatic programming*, i.e. the task of finding a program that calculates a given input–output function. Other founders in the field remember a paper of Fraser (1957) that influenced their early work, and there may be many more such forerunners depending on whom one asks.

In the same time frame Bremermann presented some of the first attempts to apply simulated evolution to numerical optimization problems involving both linear and convex optimization as well as the solution of nonlinear simultaneous equations (Bremermann 1962). Bremermann also developed some of the early evolutionary algorithm (EA) theory, showing that the optimal mutation probability for linearly separable problems should have the value of $1/\ell$ in the case of ℓ bits encoding an individual (Bremermann *et al* 1965).

Also during this period Box developed his *evolutionary operation* (EVOP) ideas which involved an evolutionary technique for the design and analysis of (industrial) experiments (Box 1957, Box and Draper 1969). Box's ideas were never realized as a computer algorithm, although Spendley *et al* (1962) used them as the basis for their so-called *simplex design* method. It is interesting to note that the REVOP proposal (Satterthwaite 1959a, b) introducing randomness into the EVOP operations was rejected at that time.

As is the case with many ground-breaking efforts, these early studies were met with considerable skepticism. However, by the mid-1960s the bases for what we today identify as the three main forms of EA were clearly established. The roots of *evolutionary programming* (EP) were laid by Lawrence Fogel B1.4 in San Diego, California (Fogel *et al* 1966) and those of *genetic algorithms* (GAs) were developed at B1.2 the University of Michigan in Ann Arbor by Holland (1967). On the other side of the Atlantic Ocean, *evolution strategies* (ESs) were a joint development of a group of three students, Bienert, Rechenberg, and B1.3 Schwefel, in Berlin (Rechenberg 1965).

Over the next 25 years each of these branches developed quite independently of each other, resulting in unique parallel histories which are described in more detail in the following sections. However, in 1990 there was an organized effort to provide a forum for interaction among the various EA research communities. This took the form of an international workshop entitled *Parallel Problem Solving from Nature* at Dortmund (Schwefel and Männer 1991).

Since that event the interaction and cooperation among EA researchers from around the world has continued to grow. In the subsequent years special efforts were made by the organizers of *ICGA'91* (Belew and Booker 1991), *EP'92* (Fogel and Atmar 1992), and *PPSN'92* (Männer and Manderick 1992) to provide additional opportunities for interaction.

This increased interaction led to a consensus for the name of this new field, *evolutionary computation* (EC), and the establishment in 1993 of a journal by the same name published by MIT Press. The increasing interest in EC was further indicated by the *IEEE World Congress on Computational Intelligence (WCCI)* at Orlando, Florida, in June 1994 (Michalewicz *et al* 1994), in which one of the three simultaneous conferences was dedicated to EC along with conferences on neural networks and fuzzy systems. The dramatic growth of interest provided additional evidence for the need of an organized EC handbook (which you are now reading) to provide a more cohesive view of the field.

That brings us to the present in which the continued growth of the field is reflected by the many EC events and related activities each year, and its growing maturity reflected by the increasing number of books and articles about EC.

In order to keep this overview brief, we have deliberately suppressed many of the details of the historical developments within each of the three main EC streams. For the interested reader these details are presented in the following sections.

A2.3.2 Evolutionary programming

Evolutionary programming (EP) was devised by Lawrence J Fogel in 1960 while serving at the National Science Foundation (NSF). Fogel was on leave from Convair, tasked as special assistant to the associate director (research), Dr Richard Bolt, to study and write a report on investing in basic research. Artificial intelligence at the time was mainly concentrated around heuristics and the simulation of primitive neural networks. It was clear to Fogel that both these approaches were limited because they model humans rather than the essential process that produces creatures of increasing intellect: evolution. Fogel considered intelligence to be based on adapting behavior to meet goals in a range of environments. In turn, prediction was viewed as the key ingredient to intelligent behavior and this suggested a series of experiments on the use of simulated evolution of *finite-state machines* to forecast nonstationary time series with respect to C1.5
arbitrary criteria. These and other experiments were documented in a series of publications (Fogel 1962, 1964, Fogel *et al* 1965, 1966, and many others).

Intelligent behavior was viewed as requiring the composite ability to (i) predict one's environment, coupled with (ii) a translation of the predictions into a suitable response in light of the given goal. For the sake of generality, the environment was described as a sequence of symbols taken from a finite alphabet. The evolutionary problem was defined as evolving an algorithm (essentially a program) that would operate on the sequence of symbols thus far observed in such a manner so as to produce an output symbol that is likely to maximize the algorithm's performance in light of both the next symbol to appear in the environment and a well-defined payoff function. Finite-state machines provided a useful representation for the required behavior.

The proposal was as follows. A population of finite-state machines is exposed to the environment, that is, the sequence of symbols that have been observed up to the current time. For each parent machine, as each input symbol is offered to the machine, each output symbol is compared with the next input symbol. The worth of this prediction is then measured with respect to the payoff function (e.g. all–none, absolute error, squared error, or any other expression of the meaning of the symbols). After the last prediction is made, a function of the payoff for each symbol (e.g. average payoff per symbol) indicates the fitness of the machine.

Offspring machines are created by randomly mutating each parent machine. Each parent produces offspring (this was originally implemented as only a single offspring simply for convenience). There are five possible modes of random mutation that naturally result from the description of the machine: change an output symbol, change a state transition, add a state, delete a state, or change the initial state. The deletion of a state and change of the initial state are only allowed when the parent machine has more than one state. Mutations are chosen with respect to a probability distribution, which is typically uniform. The number of mutations per offspring is also chosen with respect to a probability distribution or may be fixed *a priori*. These offspring are then evaluated over the existing environment in the same manner as their parents. Other mutations, such as majority logic mating operating on three or more machines, were proposed by Fogel *et al* (1966) but not implemented.

The machines that provide the greatest payoff are retained to become parents of the next generation. (Typically, half the total machines were saved so that the parent population remained at a constant size.) This process is iterated until an actual prediction of the next symbol (as yet unexperienced) in the environment is required. The best machine generates this prediction, the new symbol is added to the experienced environment, and the process is repeated. Fogel (1964) (and Fogel *et al* (1966)) used 'nonregressive' evolution. To be retained, a machine had to rank in the best half of the population. Saving lesser-adapted machines was discussed as a possibility (Fogel *et al* 1966, p 21) but not incorporated.

This general procedure was successfully applied to problems in prediction, identification, and automatic control (Fogel *et al* 1964, 1966, Fogel 1968) and was extended to simulate coevolving populations by Fogel and Burgin (1969). Additional experiments evolving finite-state machines for sequence prediction, pattern recognition, and gaming can be found in the work of Lutter and Huntsinger (1969), Burgin (1969), Atmar (1976), Dearholt (1976), and Takeuchi (1980).

In the mid-1980s the general EP procedure was extended to alternative representations including ordered lists for the traveling salesman problem (Fogel and Fogel 1986), and real-valued vectors for continuous function optimization (Fogel and Fogel 1986). This led to other applications in route planning (Fogel 1988, Fogel and Fogel 1988), optimal subset selection (Fogel 1989), and training neural networks (Fogel *et al* 1990), as well as comparisons to other methods of *simulated evolution* (Fogel and Atmar 1990). F1.10
Methods for extending evolutionary search to a two-step process including evolution of the mutation variance were offered by Fogel *et al* (1991, 1992). Just as the proper choice of step sizes is a crucial part of every numerical process, including optimization, the internal adaptation of the mutation variance(s) is of utmost importance for the algorithm's efficiency. This process is called *self-adaptation* or autoadaptation C7.1
in the case of no explicit control mechanism, e.g. if the variances are part of the individuals' characteristics and underlie probabilistic variation in a similar way as do the ordinary decision variables.

In the early 1990s efforts were made to organize annual conferences on EP, these leading to the first conference in 1992 (Fogel and Atmar 1992). This conference offered a variety of optimization applications of EP in robotics (McDonnell *et al* 1992, Andersen *et al* 1992), path planning (Larsen and Herman 1992, Page *et al* 1992), *neural network design and training* (Sebald and Fogel 1992, Porto 1992, D1
McDonnell 1992), *automatic control* (Sebald *et al* 1992), and other fields. F1.3

First contacts were made between the EP and ES communities just before this conference, and the similar but independent paths that these two approaches had taken to simulating the process of evolution were clearly apparent. Members of the ES community have participated in all successive EP conferences (Bäck *et al* 1993, Sprave 1994, Bäck and Schütz 1995, Fogel *et al* 1996). There is less similarity between EP and GAs, as the latter emphasize simulating specific mechanisms that apply to natural genetic systems whereas EP emphasizes the behavioral, rather than genetic, relationships between parents and their offspring. Members of the GA and GP communities have, however, also been invited to participate in the annual conferences, making for truly interdisciplinary interaction (see e.g. Altenberg 1994, Land and Belew 1995, Koza and Andre 1996).

Since the early 1990s, efforts in EP have diversified in many directions. Applications in training neural networks have received considerable attention (see e.g. English 1994, Angeline *et al* 1994, McDonnell and Waagen 1994, Porto *et al* 1995), while relatively less attention has been devoted to evolving *fuzzy* D2
systems (Haffner and Sebald 1993, Kim and Jeon 1996). Image processing applications can be found in the articles by Bhattacharjya and Roysam (1994), Brotherton *et al* (1994), Rizki *et al* (1995), and others. Recent efforts to use EP in medicine have been offered by Fogel *et al* (1995) and Gehlhaar *et al* (1995). Efforts studying and comparing methods of self-adaptation can be found in the articles by Saravanan *et al* (1995), Angeline *et al* (1996), and others. Mathematical analyses of EP have been summarized by Fogel (1995).

To offer a summary, the initial efforts of L J Fogel indicate some of the early attempts to (i) use simulated evolution to perform prediction, (ii) include variable-length encodings, (iii) use representations that take the form of a sequence of instructions, (iv) incorporate a population of candidate solutions, and (v) coevolve evolutionary programs. Moreover, Fogel (1963, 1964) and Fogel *et al* (1966) offered the early recognition that natural evolution and the human endeavor of the scientific method are essentially similar processes, a notion recently echoed by Gell-Mann (1994). The initial prescriptions for operating on finite-state machines have been extended to arbitrary representations, mutation operators, and selection methods, and techniques for self-adapting the evolutionary search have been proposed and implemented.

The population size need not be kept constant and there can be a variable number of offspring per parent, much like the *($\mu + \lambda$) methods* offered in ESs. In contrast to these methods, selection is often made probabilistic in EP, giving lesser-scoring solutions some probability of surviving as parents into the next generation. In contrast to GAs, no effort is made in EP to support (some say maximize) *schema processing*, nor is the use of random variation constrained to emphasize specific mechanisms of genetic transfer, perhaps providing greater versatility to tackle specific problem domains that are unsuitable for genetic operators such as crossover. C2.4.4 B2.5

A2.3.3 Genetic algorithms

The first glimpses of the ideas underlying genetic algorithms (GAs) are found in Holland's papers in the early 1960s (see e.g. Holland 1962). In them Holland set out a broad and ambitious agenda for understanding the underlying principles of adaptive systems—systems that are capable of self-modification in response to their interactions with the environments in which they must function. Such a theory of adaptive systems should facilitate both the understanding of complex forms of adaptation as they appear in natural systems and our ability to design robust adaptive artifacts.

In Holland's view the key feature of robust natural adaptive systems was the successful use of competition and innovation to provide the ability to dynamically respond to unanticipated events and changing environments. Simple models of biological evolution were seen to capture these ideas nicely via notions of survival of the fittest and the continuous production of new offspring.

This theme of using evolutionary models both to understand natural adaptive systems and to design robust adaptive artifacts gave Holland's work a somewhat different focus than those of other contemporary groups that were exploring the use of evolutionary models in the design of efficient experimental optimization techniques (Rechenberg 1965) or for the evolution of intelligent agents (Fogel *et al* 1966), as reported in the previous section.

By the mid-1960s Holland's ideas began to take on various computational forms as reflected by the PhD students working with Holland. From the outset these systems had a distinct 'genetic' flavor to them in the sense that the objects to be evolved over time were represented internally as 'genomes' and the mechanisms of reproduction and inheritance were simple abstractions of familiar population genetics operators such as mutation, crossover, and inversion.

Bagley's thesis (Bagley 1967) involved tuning sets of weights used in the evaluation functions of game-playing programs, and represents some of the earliest experimental work in the use of diploid representations, the role of inversion, and selection mechanisms. By contrast Rosenberg's thesis (Rosenberg 1967) has a very distinct flavor of simulating the evolution of a simple biochemical system in which single-celled organisms capable of producing enzymes were represented in diploid fashion and were evolved over time to produce appropriate chemical concentrations. Of interest here is some of the earliest experimentation with adaptive crossover operators.

Cavicchio's thesis (Cavicchio 1970) focused on viewing these ideas as a form of adaptive search, and tested them experimentally on difficult search problems involving subroutine selection and pattern recognition. In his work we see some of the early studies on *elitist* forms of selection and ideas for adapting the rates of crossover and mutation. Hollstien's thesis (Hollstien 1971) took the first detailed look at alternate selection and mating schemes. Using a test suite of two-dimensional *fitness landscapes*, he experimented with a variety of breeding strategies drawn from techniques used by animal breeders. Also of interest here is Hollstien's use of binary string encodings of the genome and early observations about the virtues of Gray codings. C2.7.4 B2.7.4

In parallel with these experimental studies, Holland continued to work on a general theory of adaptive systems (Holland 1967). During this period he developed his now famous *schema analysis* of adaptive systems, relating it to the optimal allocation of trials using k-armed bandit models (Holland 1969). He used these ideas to develop a more theoretical analysis of his *reproductive plans* (simple GAs) (Holland 1971, 1973). Holland then pulled all of these ideas together in his pivotal book *Adaptation in Natural and Artificial Systems* (Holland 1975). B2.5

Of interest was the fact that many of the desirable properties of these algorithms being identified by Holland theoretically were frequently not observed experimentally. It was not difficult to identify the reasons for this. Hampered by a lack of computational resources and analysis tools, most of the early experimental studies involved a relatively small number of runs using small population sizes (generally less than 20). It became increasingly clear that many of the observed deviations from expected behavior could be traced to the well-known phenomenon in population genetics of *genetic drift*, the loss of genetic diversity due to the stochastic aspects of selection, reproduction, and the like in small populations.

By the early 1970s there was considerable interest in understanding better the behavior of implementable GAs. In particular, it was clear that choices of population size, representation issues, the choice of operators and operator rates all had significant effects of the observed behavior of GAs. Frantz's thesis (Frantz 1972) reflected this new focus by studying in detail the roles of crossover and inversion in populations of size 100. Of interest here is some of the earliest experimental work on multipoint crossover operators.

De Jong's thesis (De Jong 1975) broaded this line of study by analyzing both theoretically and experimentally the interacting effects of population size, crossover, and mutation on the behavior of a family of GAs being used to optimize a fixed test suite of functions. Out of this study came a strong sense that even these simple GAs had significant potential for solving difficult optimization problems.

The mid-1970s also represented a branching out of the family tree of GAs as other universities and research laboratories established research activities in this area. This happened slowly at first since initial attempts to spread the word about the progress being made in GAs were met with fairly negative perceptions from the artificial intelligence (AI) community as a result of early overhyped work in areas such as self-organizing systems and perceptrons.

Undaunted, groups from several universities including the University of Michigan, the University of Pittsburgh, and the University of Alberta organized an *Adaptive Systems Workshop* in the summer of 1976 in Ann Arbor, Michigan. About 20 people attended and agreed to meet again the following summer. This pattern repeated itself for several years, but by 1979 the organizers felt the need to broaden the scope and make things a little more formal. Holland, De Jong, and Sampson obtained NSF funding for *An Interdisciplinary Workshop in Adaptive Systems*, which was held at the University of Michigan in the summer of 1981 (Sampson 1981).

By this time there were several established research groups working on GAs. At the University of Michigan, Bethke, Goldberg, and Booker were continuing to develop GAs and explore Holland's *classifier* B1.5.2
systems as part of their PhD research (Bethke 1981, Booker 1982, Goldberg 1983). At the University of Pittsburgh, Smith and Wetzel were working with De Jong on various GA enhancements including the *Pitt approach* to rule learning (Smith 1980, Wetzel 1983). At the University of Alberta, Brindle continued to look at optimization applications of GAs under the direction of Sampson (Brindle 1981).

The continued growth of interest in GAs led to a series of discussions and plans to hold the first *International Conference on Genetic Algorithms (ICGA)* in Pittsburgh, Pennsylvania, in 1985. There were about 75 participants presenting and discussing a wide range of new developments in both the theory and application of GAs (Grefenstette 1985). The overwhelming success of this meeting resulted in agreement to continue *ICGA* as a biannual conference. Also agreed upon at *ICGA'85* was the initiation of a moderated electronic discussion group called *GA List*.

The field continued to grow and mature as reflected by the *ICGA* conference activities (Grefenstette 1987, Schaffer 1989) and the appearance of several books on the subject (Davis 1987, Goldberg 1989). Goldberg's book, in particular, served as a significant catalyst by presenting current GA theory and applications in a clear and precise form easily understood by a broad audience of scientists and engineers.

By 1989 the *ICGA* conference and other GA-related activities had grown to a point that some more formal mechanisms were needed. The result was the formation of the International Society for Genetic Algorithms (ISGA), an incorporated body whose purpose is to serve as a vehicle for conference funding and to help coordinate and facilitate GA-related activities. One of its first acts of business was to support a proposal to hold a theory workshop on the *Foundations of Genetic Algorithms (FOGA)* in Bloomington, Indiana (Rawlins 1991).

By this time nonstandard GAs were being developed to evolve complex, nonlinear variable-length structures such as rule sets, LISP code, and neural networks. One of the motivations for *FOGA* was the sense that the growth of GA-based applications had driven the field well beyond the capacity of existing theory to provide effective analyses and predictions.

Also in 1990, Schwefel hosted the first *PPSN* conference in Dortmund, which resulted in the first organized interaction between the ES and GA communities. This led to additional interaction at *ICGA'91* in San Diego which resulted in an informal agreement to hold *ICGA* and *PPSN* in alternating years, and a commitment to jointly initiate a journal for the field.

It was felt that in order for the journal to be successful, it must have broad scope and include other species of EA. Efforts were made to include the EP community as well (which began to organize its own conferences in 1992), and the new journal *Evolutionary Computation* was born with the inaugural issue in the spring of 1993.

The period from 1990 to the present has been characterized by tremendous growth and diversity of the GA community as reflected by the many conference activities (e.g. *ICGA* and *FOGA*), the emergence of new books on GAs, and a growing list of journal papers. New paradigms such as *messy GAs* (Goldberg *et al* 1991) and *genetic programming* (Koza 1992) were being developed. The interactions with other EC communities resulted in considerable crossbreeding of ideas and many new hybrid EAs. New GA applications continue to be developed, spanning a wide range of problem areas from engineering design problems to operations research problems to automatic programming.

C4.2.4
B1.5.1

A2.3.4 Evolution strategies

In 1964, three students of the Technical University of Berlin, Bienert, Rechenberg, and Schwefel, did not at all aim at devising a new kind of optimization procedure. During their studies of aerotechnology and space technology they met at an Institute of Fluid Mechanics and wanted to construct a kind of research robot that should perform series of experiments on a flexible slender three-dimensional body in a wind tunnel so as to minimize its drag. The method of minimization was planned to be either a one variable at a time or a discrete gradient technique, gleaned from classical numerics. Both strategies, performed manually, failed, however. They became stuck prematurely when used for a two-dimensional demonstration facility, a joint plate—its optimal shape being a flat plate—with which the students tried to demonstrate that it was possible to find the optimum automatically.

Only then did Rechenberg (1965) hit upon the idea to use dice for random decisions. This was the breakthrough—on 12 June 1964. The first version of an evolutionary strategy (ES), later called the $(1+1)$ ES, was born, with discrete, binomially distributed mutations centered at the ancestor's position, and just one parent and one descendant per generation. This ES was first tested on a mechanical calculating machine by Schwefel before it was used for the *experimentum crucis*, the joint plate. Even then, it took a while to overcome a merely locally optimal S shape and to converge towards the expected global optimum, the flat plate. Bienert (1967), the third of the three students, later actually constructed a kind of robot that could perform the actions and decisions automatically.

Using this simple *two-membered* ES, another student, Lichtfuß (1965), optimized the shape of a bent pipe, also experimentally. The result was rather unexpected, but nevertheless obviously better than all shapes proposed so far.

First computer experiments, on a Zuse Z23, as well as analytical investigations using binomially distributed integer mutations, had already been performed by Schwefel (1965). The main result was that such a strategy can become stuck prematurely, i.e. at 'solutions' that are not even locally optimal. Based on this experience the use of normally instead of binomially distributed mutations became standard in most of the later computer experiments with real-valued variables and in theoretical investigations into the method's efficiency, but not however in experimental optimization using ESs. In 1966 the little ES community was destroyed by dismissal from the Institute of Fluid Mechanics ('Cybernetics as such is no longer pursued at the institute!'). Not before 1970 was it found together again at the Institute of Measurement and Control of the Technical University of Berlin, sponsored by grants from the German Research Foundation (DFG). Due to the circumstances, the group missed publishing its ideas and results properly, especially in English.

In the meantime the often-cited two-phase nozzle optimization was performed at the Institute of Nuclear Technology of the Technical University of Berlin, then in an industrial surrounding, the AEG research laboratory (Schwefel 1968, Klockgether and Schwefel 1970), also at Berlin. For a hot-water flashing flow the shape of a three-dimensional convergent–divergent (thus supersonic) nozzle with maximum energy efficiency was sought. Though in this experimental optimization an exogenously controlled binomial-like distribution was used again, it was the first time that gene duplication and deletion were incorporated into an EA, especially in a $(1+1)$ ES, because the optimal length of the nozzle was not known in advance. As in case of the bent pipe this experimental strategy led to highly unexpected results, not easy to understand even afterwards, but definitely much better than available before.

First Rechenberg and later Schwefel analyzed and improved their ES. For the $(1+1)$ ES, Rechenberg, in his Dr.-Ing. thesis of 1971, developed, on the basis of two convex n-dimensional model functions, a convergence rate theory for $n \gg 1$ variables. Based on these results he formulated a $\frac{1}{5}$ success rule for adapting the standard deviation of mutation (Rechenberg 1973). The hope of arriving at an even better strategy by imitating organic evolution more closely led to the incorporation of the population principle and the introduction of recombination, which of course could not be embedded in the $(1+1)$ ES. A first *multimembered* ES, the $(\mu+1)$ ES—the notation was introduced later by Schwefel—was also designed by Rechenberg in his seminal work of 1973. Because of its inability to self-adapt the mutation step sizes (more accurately, standard deviations of the mutations), this strategy was never widely used.

Much more widespread became the $(\mu+\lambda)$ ES and (μ, λ) ES, both formulated by Schwefel in his Dr.-Ing. thesis of 1974–1975. It contains theoretical results such as a convergence rate theory for the $(1+\lambda)$ ES and the $(1, \lambda)$ ES ($\lambda > 1$), analogous to the theory introduced by Rechenberg for the $(1+1)$ ES (Schwefel 1977). The *multimembered* ($\mu > 1$) ESs arose from the otherwise ineffective incorporation of mutatable mutation parameters (variances and covariances of the Gaussian distributions used). Self-adaptation was achieved with the (μ, λ) ES first, not only with respect to the step sizes, but also with respect to correlation coefficients. The enhanced ES version with correlated mutations, described already in an internal report (Schwefel 1974), was published much later (Schwefel 1981) due to the fact that the author left Berlin in 1976. A more detailed empirical analysis of the on-line self-adaptation of the internal or strategy parameters was first published by Schwefel in 1987 (the tests themselves were secretly performed on one of the first small instruction multiple data (SIMD) parallel machines (CRAY1) at the Nuclear Research Centre (KFA) Jülich during the early 1980s with a first parallel version of the multimembered ES with correlated mutations). It was in this work that the notion of *self-adaptation by collective learning* first came up. The importance of recombination (for object as well as strategy parameters) and soft selection (or $\mu > 1$) was clearly demonstrated. Only recently has Beyer (1995a, b) delivered the theoretical background to that particularly important issue.

It may be worth mentioning that in the beginning there were strong objections against increasing λ as well as μ beyond one. The argument against $\lambda > 1$ was that the exploitation of the current knowledge was unnecessarily delayed, and the argument against $\mu > 1$ was that the survival of inferior members of the population would unnecessarily slow down the evolutionary progress. The hint that λ successors could be evaluated in parallel did not convince anybody since parallel computers were neither available nor expected in the near future. The two-membered ES and the very similar creeping random search method of Rastrigin (1965) were investigated thoroughly with respect to their convergence and convergence rates also by Matyas (1965) in Czechoslovakia, Born (1978) on the Eastern side of the Berlin wall (!), and Rappl (1984) in Munich.

Since this early work many new results have been produced by the ES community consisting of the group at Berlin (Rechenberg, since 1972) and that at Dortmund (Schwefel, since 1985). In particular, strategy variants concerning other than only real-valued parameter optimization, i.e. real-world problems, were invented. The first use of an ES for binary optimization using multicellular individuals was presented by Schwefel (1975). The idea of using several subpopulations and *niching mechanisms* for global C6.1 optimization was propagated by Schwefel in 1977; due to a lack of computing resources, however, it could not be tested thoroughly at that time. Rechenberg (1978) invented a notational scheme for such nested ESs.

Beside these nonstandard approaches there now exists a wide range of other ESs, e.g. several parallel concepts (Hoffmeister and Schwefel 1990, Lohmann 1991, Rudolph 1991, 1992, Sprave 1994, Rudolph and Sprave 1995), ESs for *multicriterion problems* (Kursawe 1991, 1992), for mixed-integer tasks F1.9 (Lohmann 1992, Rudolph 1994, Bäck and Schütz 1995), and even for problems with a variable-dimensional parameter space (Schütz and Sprave 1996), and variants concerning nonstandard step size and direction adaptation schemes (see e.g. Matyas 1967, Stewart *et al* 1967, Fürst *et al* 1968, Heydt 1970, Rappl 1984, Ostermeier *et al* 1994). Comparisons between ESs, GAs, and EP may be found in the articles by Bäck *et al* (1991, 1993). It was Bäck (1996) who introduced a common algorithmic scheme for all brands of current EAs.

Omitting all these other useful nonstandard ESs—a commented collection of literature concerning ES applications was made at the University of Dortmund (Bäck *et al* 1992)—the history of ESs is closed with a mention of three recent books by Rechenberg (1994), Schwefel (1995), and Bäck (1996) as well as three recent contributions that may be seen as written tutorials (Schwefel and Rudolph 1995, Bäck and Schwefel 1995, Schwefel and Bäck 1995), which on the one hand define the actual standard ES algorithms and on the other hand present some recent theoretical results.

References

Altenberg L 1994 Emergent phenomena in genetic programming *Proc. 3rd Annu. Conf. on Evolutionary Programming (San Diego, CA, 1994)* ed A V Sebald and L J Fogel (Singapore: World Scientific) pp 233–41

Andersen B, McDonnell J and Page W 1992 Configuration optimization of mobile manipulators with equality constraints using evolutionary programming *Proc. 1st Ann. Conf. on Evolutionary Programming (La Jolla, CA, 1992)* ed D B Fogel and W Atmar (La Jolla, CA: Evolutionary Programming Society) pp 71–9

Angeline P J, Fogel D B and Fogel L J 1996 A comparison of self-adaptation methods for finite state machines in a dynamic environment *Evolutionary Programming V—Proc. 5th Ann. Conf. on Evolutionary Programming (1996)* ed L J Fogel, P J Angeline and T Bäck (Cambridge, MA: MIT Press)

Angeline P J, Saunders G M and Pollack J B 1994 An evolutionary algorithm that constructs recurrent neural networks *IEEE Trans. Neural Networks* **NN-5** 54–65

Atmar J W 1976 *Speculation of the Evolution of Intelligence and Its Possible Realization in Machine Form* ScD Thesis, New Mexico State University

Bäck T 1996 *Evolutionary Algorithms in Theory and Practice* (New York: Oxford University Press)

Bäck T, Hoffmeister F and Schwefel H-P 1991 A survey of evolution strategies *Proc. 4th Int. Conf. on Genetic Algorithms (San Diego, CA, 1991)* ed R K Belew and L B Booker (San Mateo, CA: Morgan Kaufmann) pp 2–9

——1992 *Applications of Evolutionary Algorithms* Technical Report of the University of Dortmund Department of Computer Science Systems Analysis Research Group SYS-2/92

Bäck T, Rudolph G and Schwefel H-P 1993 Evolutionary programming and evolution strategies: similarities and differences *Proc. 2nd Ann. Conf. on Evolutionary Programming (San Diego, CA, 1993)* ed D B Fogel and W Atmar (La Jolla, CA: Evolutionary Programming Society) pp 11–22

Bäck T and Schütz M 1995 Evolution strategies for mixed-integer optimization of optical multilayer systems *Evolutionary Programming IV—Proc. 4th Ann. Conf on Evolutionary Programming (San Diego, CA, 1995)* ed J R McDonnell, R G Reynolds and D B Fogel (Cambridge, MA: MIT Press) pp 33–51

Bäck T and Schwefel H-P 1995 Evolution strategies I: variants and their computational implementation *Genetic Algorithms in Engineering and Computer Science, Proc. 1st Short Course EUROGEN-95* ed G Winter, J Périaux, M Galán and P Cuesta (New York: Wiley) pp 111–26

Bagley J D 1967 *The Behavior of Adaptive Systems which Employ Genetic and Correlation Algorithms* PhD Thesis, University of Michigan

Belew R K and Booker L B (eds) 1991 *Proc. 4th Int. Conf. on Genetic Algorithms (San Diego, CA, 1991)* (San Mateo, CA: Morgan Kaufmann)

Bethke A D 1981 *Genetic Algorithms as Function Optimizers* PhD Thesis, University of Michigan

Beyer H-G 1995a *How GAs do Not Work—Understanding GAs Without Schemata and Building Blocks* Technical Report of the University of Dortmund Department of Computer Science Systems Analysis Research Group SYS-2/95

——1995b Toward a theory of evolution strategies: on the benefit of sex—the $(\mu/\mu, \lambda)$-theory *Evolutionary Comput.* **3** 81–111

Bhattacharjya A K and Roysam B 1994 Joint solution of low-, intermediate- and high-level vision tasks by evolutionary optimization: application to computer vision at low SNR *IEEE Trans. Neural Networks* **NN-5** 83–95

Bienert P 1967 *Aufbau einer Optimierungsautomatik für drei Parameter* Dipl.-Ing. Thesis, Technical University of Berlin, Institute of Measurement and Control Technology

Booker L 1982 *Intelligent Behavior as an Adaptation to the Task Environment* PhD Thesis, University of Michigan

Born J 1978 *Evolutionsstrategien zur numerischen Lösung von Adaptationsaufgaben* PhD Thesis, Humboldt University at Berlin

Box G E P 1957 Evolutionary operation: a method for increasing industrial productivity *Appl. Stat.* **6** 81–101

Box G E P and Draper N P 1969 *Evolutionary Operation. A Method for Increasing Industrial Productivity* (New York: Wiley)

Bremermann H J 1962 Optimization through evolution and recombination *Self-Organizing Systems* ed M C Yovits *et al* (Washington, DC: Spartan)

Bremermann H J, Rogson M and Salaff S 1965 Search by evolution *Biophysics and Cybernetic Systems—Proc. 2nd Cybernetic Sciences Symp.* ed M Maxfield, A Callahan and L J Fogel (Washington, DC: Spartan) pp 157–67

Brindle A 1981 *Genetic Algorithms for Function Optimization* PhD Thesis, University of Alberta

Brotherton T W, Simpson P K, Fogel D B and Pollard T 1994 Classifier design using evolutionary programming *Proc. 3rd Ann. Conf. on Evolutionary Programming (San Diego, CA, 1994)* ed A V Sebald and L J Fogel (Singapore: World Scientific) pp 68–75

Burgin G H 1969 On playing two-person zero-sum games against nonminimax players *IEEE Trans. Syst. Sci. Cybernet.* **SSC-5** 369–70

Cavicchio D J 1970 *Adaptive Search Using Simulated Evolution* PhD Thesis, University of Michigan

Davis L 1987 *Genetic Algorithms and Simulated Annealing* (London: Pitman)

Dearholt D W 1976 Some experiments on generalization using evolving automata *Proc. 9th Int. Conf. on System Sciences (Honolulu, HI)* pp 131–3

De Jong K A 1975 *Analysis of Behavior of a Class of Genetic Adaptive Systems* PhD Thesis, University of Michigan
English T M 1994 Generalization in populations of recurrent neural networks *Proc. 3rd Ann. Conf. on Evolutionary Programming (San Diego, CA, 1994)* ed A V Sebald and L J Fogel (Singapore: World Scientific) pp 26–33
Fogel D B 1988 An evolutionary approach to the traveling salesman problem *Biol. Cybernet.* **60** 139–44
——1989 Evolutionary programming for voice feature analysis *Proc. 23rd Asilomar Conf. on Signals, Systems and Computers (Pacific Grove, CA)* pp 381–3
——1995 *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence* (New York: IEEE)
Fogel D B and Atmar J W 1990 Comparing genetic operators with Gaussian mutations in simulated evolutionary processing using linear systems *Biol. Cybernet.* **63** 111–4
——(eds) 1992 *Proc. 1st Ann. Conf. on Evolutionary Programming (La Jolla, CA, 1992)* (La Jolla, CA: Evolutionary Programming Society)
Fogel D B and Fogel L J 1988 Route optimization through evolutionary programming *Proc. 22nd Asilomar Conf. on Signals, Systems and Computers (Pacific Grove, CA)* pp 679–80
Fogel D B, Fogel L J and Atmar J W 1991 Meta-evolutionary programming *Proc. 25th Asilomar Conf. on Signals, Systems and Computers (Pacific Grove, CA)* ed R R Chen pp 540–5
Fogel D B, Fogel L J, Atmar J W and Fogel G B 1992 Hierarchic methods of evolutionary programming *Proc. 1st Ann. Conf. on Evolutionary Programming (La Jolla, CA, 1992)* ed D B Fogel and W Atmar (La Jolla, CA: Evolutionary Programming Society) pp 175–82
Fogel D B, Fogel L J and Porto V W 1990 Evolving neural networks *Biol. Cybernet.* **63** 487–93
Fogel D B, Wasson E C and Boughton E M 1995 Evolving neural networks for detecting breast cancer *Cancer Lett.* **96** 49–53
Fogel L J 1962 Autonomous automata *Industrial Res.* **4** 14–9
——1963 *Biotechnology: Concepts and Applications* (Englewood Cliffs, NJ: Prentice-Hall)
——1964 *On the Organization of Intellect* PhD Thesis, University of California at Los Angeles
——1968 Extending communication and control through simulated evolution *Bioengineering—an Engineering View Proc. Symp. on Engineering Significance of the Biological Sciences* ed G Bugliarello (San Francisco, CA: San Francisco Press) pp 286–304
Fogel L J, Angeline P J and Bäck T (eds) 1996 *Evolutionary Programming V—Proc. 5th Ann. Conf. on Evolutionary Programming (1996)* (Cambridge, MA: MIT Press)
Fogel L J and Burgin G H 1969 *Competitive Goal-seeking through Evolutionary Programming* Air Force Cambridge Research Laboratories Final Report Contract AF 19(628)-5927
Fogel L J and Fogel D B 1986 *Artificial Intelligence through Evolutionary Programming* US Army Research Institute Final Report Contract PO-9-X56-1102C-1
Fogel L J, Owens A J and Walsh M J 1964 On the evolution of artificial intelligence *Proc. 5th Natl Symp. on Human Factors in Electronics* (San Diego, CA: IEEE)
——1965 Artificial intelligence through a simulation of evolution *Biophysics and Cybernetic Systems* ed A Callahan, M Maxfield and L J Fogel (Washington, DC: Spartan) pp 131–56
——1966 *Artificial Intelligence through Simulated Evolution* (New York: Wiley)
Frantz D R 1972 *Non-linearities in Genetic Adaptive Search* PhD Thesis, University of Michigan
Fraser A S 1957 Simulation of genetic systems by automatic digital computers *Aust. J. Biol. Sci.* **10** 484–99
Friedberg R M 1958 A learning machine: part I *IBM J.* **2** 2–13
Friedberg R M, Dunham B and North J H 1959 A learning machine: part II *IBM J.* **3** 282–7
Fürst H, Müller P H and Nollau V 1968 Eine stochastische Methode zur Ermittlung der Maximalstelle einer Funktion von mehreren Veränderlichen mit experimentell ermittelbaren Funktionswerten und ihre Anwendung bei chemischen Prozessen *Chem.–Tech.* **20** 400–5
Gehlhaar *et al* 1995Gehlhaar D K *et al* 1995 Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming *Chem. Biol.* **2** 317–24
Gell-Mann M 1994 *The Quark and the Jaguar* (New York: Freeman)
Goldberg D E 1983 *Computer-Aided Gas Pipeline Operation using Genetic Algorithms and Rule Learning* PhD Thesis, University of Michigan
——1989 *Genetic Algorithms in Search, Optimization and Machine Learning* (Reading, MA: Addison-Wesley)
Goldberg D E, Deb K and Korb B 1991 Don't worry, be messy *Proc. 4th Int. Conf. on Genetic Algorithms (San Diego, CA, 1991)* ed R K Belew and L B Booker (San Mateo, CA: Morgan Kaufmann) pp 24–30
Grefenstette J J (ed) 1985 *Proc. 1st Int. Conf. on Genetic Algorithms and Their Applications (Pittsburgh, PA, 1985)* (Hillsdale, NJ: Erlbaum)
——1987 *Proc. 2nd Int. Conf. on Genetic Algorithms and Their Applications (Cambridge, MA, 1987)* (Hillsdale, NJ: Erlbaum)
Haffner S B and Sebald A V 1993 Computer-aided design of fuzzy HVAC controllers using evolutionary programming *Proc. 2nd Ann. Conf. on Evolutionary Programming (San Diego, CA, 1993)* ed D B Fogel and W Atmar (La Jolla, CA: Evolutionary Programming Society) pp 98–107
Heydt G T 1970 *Directed Random Search* PhD Thesis, Purdue University

Hoffmeister F and Schwefel H-P 1990 A taxonomy of parallel evolutionary algorithms *Parcella '90, Proc. 5th Int. Workshop on Parallel Processing by Cellular Automata and Arrays* vol 2, ed G Wolf, T Legendi and U Schendel (Berlin: Academic) pp 97–107

Holland J H 1962 Outline for a logical theory of adaptive systems *J. ACM* **9** 297–314

——1967 Nonlinear environments permitting efficient adaptation *Computer and Information Sciences II* (New York: Academic)

——1969 Adaptive plans optimal for payoff-only environments *Proc. 2nd Hawaii Int. Conf. on System Sciences* pp 917–20

——1971 Processing and processors for schemata *Associative information processing* ed E L Jacks (New York: Elsevier) pp 127–46

——1973 Genetic algorithms and the optimal allocation of trials *SIAM J. Comput.* **2** 88–105

——1975 *Adaptation in Natural and Artificial Systems* (Ann Arbor, MI: University of Michigan Press)

Hollstien R B 1971 *Artificial Genetic Adaptation in Computer Control Systems* PhD Thesis, University of Michigan

Kim J-H and Jeon J-Y 1996 Evolutionary programming-based high-precision controller design *Evolutionary Programming V—Proc. 5th Ann. Conf. on Evolutionary Programming (1996)* ed L J Fogel, P J Angeline and T Bäck (Cambridge, MA: MIT Press)

Klockgether J and Schwefel H-P 1970 Two-phase nozzle and hollow core jet experiments *Proc. 11th Symp. on Engineering Aspects of Magnetohydrodynamics* ed D G Elliott (Pasadena, CA: California Institute of Technology) pp 141–8

Koza J R 1992 *Genetic Programming* (Cambridge, MA: MIT Press)

Koza J R and Andre D 1996 Evolution of iteration in genetic programming *Evolutionary Programming V—Proc. 5th Ann. Conf. on Evolutionary Programming (1996)* ed L J Fogel, P J Angeline and T Bäck (Cambridge, MA: MIT Press)

Kursawe F 1991 A variant of evolution strategies for vector optimization *Parallel Problem Solving from Nature—Proc. 1st Workshop PPSN I (Lecture Notes in Computer Science 496) (Dortmund, 1991)* ed H-P Schwefel and R Männer (Berlin: Springer) pp 193–7

——1992 Naturanaloge Optimierverfahren—Neuere Entwicklungen in der Informatik *Studien zur Evolutorischen Ökonomik II (Schriften des Vereins für Socialpolitik 195 II)* ed U Witt (Berlin: Duncker and Humblot) pp 11–38

Land M and Belew R K 1995 Towards a self-replicating language for computation *Evolutionary Programming IV—Proc. 4th Ann. Conf on Evolutionary Programming (San Diego, CA, 1995)* ed J R McDonnell, R G Reynolds and D B Fogel (Cambridge, MA: MIT Press) pp 403–13

Larsen R W and Herman J S 1992 A comparison of evolutionary programming to neural networks and an application of evolutionary programming to a navy mission planning problem *Proc. 1st Ann. Conf. on Evolutionary Programming (La Jolla, CA, 1992)* ed D B Fogel and W Atmar (La Jolla, CA: Evolutionary Programming Society) pp 127–33

Lichtfuß H J 1965 *Evolution eines Rohrkrümmers* Dipl.-Ing. Thesis, Technical University of Berlin, Hermann Föttinger Institute for Hydrodynamics

Lohmann R 1991 Application of evolution strategy in parallel populations *Parallel Problem Solving from Nature—Proc. 1st Workshop PPSN I (Dortmund, 1991) (Lecture Notes in Computer Science 496)* ed H-P Schwefel and R Männer (Berlin: Springer) pp 198–208

——1992 Structure evolution and incomplete induction *Parallel Problem Solving from Nature 2 (Brussels, 1992)* ed R Männer and B Manderick (Amsterdam: Elsevier–North-Holland) pp 175–85

Lutter B E and Huntsinger R C 1969 Engineering applications of finite automata *Simulation* **13** 5–11

Männer R and Manderick B (eds) 1992 *Parallel Problem Solving from Nature 2 (Brussels, 1992)* (Amsterdam: Elsevier–North-Holland)

Matyas J 1965 Random optimization *Automation Remote Control* **26** 244–51

——1967 Das zufällige Optimierungsverfahren und seine Konvergenz *Proc. 5th Int. Analogue Computation Meeting (Lausanne, 1967)* **1** 540–4

McDonnell J R 1992 Training neural networks with weight constraints *Proc. 1st Ann. Conf. on Evolutionary Programming (La Jolla, CA, 1992)* ed D B Fogel and W Atmar (La Jolla, CA: Evolutionary Programming Society) pp 111–9

McDonnell J R, Andersen B D, Page W C and Pin F 1992 Mobile manipulator configuration optimization using evolutionary programming *Proc. 1st Ann. Conf. on Evolutionary Programming (La Jolla, CA, 1992)* ed D B Fogel and W Atmar (La Jolla, CA: Evolutionary Programming Society) pp 52–62

McDonnell J R and Waagen D 1994 Evolving recurrent perceptrons for time-series prediction *IEEE Trans. Neural Networks* **NN-5** 24–38

Michalewicz Z *et al* (eds) 1994 *Proc. 1st IEEE Conf. on Evolutionary Computation (Orlando, FL, 1994)* (Piscataway, NJ: IEEE)

Ostermeier A, Gawelczyk A and Hansen N 1994 Step-size adaptation based on non-local use of selection information *Parallel Problem Solving from Nature—PPSN III Int. Conf. on Evolutionary Computation (Jerusalem, 1994) (Lecture notes in Computer Science 866)* ed Y Davidor, H-P Schwefel and R Männer (Berlin: Springer) pp 189–98

Page W C, Andersen B D and McDonnell J R 1992 An evolutionary programming approach to multi-dimensional path planning *Proc. 1st Ann. Conf. on Evolutionary Programming (La Jolla, CA, 1992)* ed D B Fogel and W Atmar (La Jolla, CA: Evolutionary Programming Society) pp 63–70

Porto V W 1992 Alternative methods for training neural networks *Proc. 1st Ann. Conf. on Evolutionary Programming (La Jolla, CA, 1992)* ed D B Fogel and W Atmar (La Jolla, CA: Evolutionary Programming Society) pp 100–10

Porto V W, Fogel D B and Fogel L J 1995 Alternative neural network training methods *IEEE Expert* **10** 16–22

Rappl G 1984 *Konvergenzraten von Random-Search-Verfahren zur globalen Optimierung* PhD Thesis, Bundeswehr University

Rastrigin L A 1965 *Random Search in Optimization Problems for Multiparameter Systems (translated from the Russian original: Sluchainyi poisk v zadachakh optimisatsii mnogoarametricheskikh sistem, Zinatne, Riga)* Air Force System Command Foreign Technology Division FTD-HT-67-363

Rawlins G J E (ed) 1991 *Foundations of Genetic Algorithms* (San Mateo, CA: Morgan Kaufmann)

Rechenberg I 1965 *Cybernetic Solution Path of an Experimental Problem* Royal Aircraft Establishment Library Translation 1122

——1973 *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution* (Stuttgart: Frommann–Holzboog)

——1978 Evolutionsstrategien *Simulationsmethoden in der Medizin und Biologie* ed B Schneider and U Ranft (Berlin: Springer) pp 83–114

——1994 *Evolutionsstrategie ’94* (Stuttgart: Frommann–Holzboog)

Rizki M M, Tamburino L A and Zmuda M A 1995 Evolution of morphological recognition systems *Evolutionary Programming IV—Proc. 4th Ann. Conf on Evolutionary Programming (San Diego, CA, 1995)* ed J R McDonnell, R G Reynolds and D B Fogel (Cambridge, MA: MIT Press) pp 95–106

Rosenberg R 1967 *Simulation of Genetic Populations with Biochemical Properties* PhD Thesis, University of Michigan

Rudolph G 1991 Global optimization by means of distributed evolution strategies *Parallel Problem Solving from Nature—Proc. 1st Workshop PPSN I (Dortmund, 1991) (Lecture Notes in Computer Science 496)* ed H-P Schwefel and R Männer (Berlin: Springer) pp 209–13

——1992 Parallel approaches to stochastic global optimization *Parallel Computing: from Theory to Sound Practice, Proc. Eur. Workshop on Parallel Computing* ed W Joosen and E Milgrom (Amsterdam: IOS) pp 256–67

——1994 An evolutionary algorithm for integer programming *Parallel Problem Solving from Nature—PPSN III Int. Conf. on Evolutionary Computation (Jerusalem, 1994) (Lecture notes in Computer Science 866)* ed Y Davidor, H-P Schwefel and R Männer (Berlin: Springer) pp 139–48

Rudolph G and Sprave J 1995 A cellular genetic algorithm with self-adjusting acceptance threshold *Proc. 1st IEE/IEEE Int. Conf. on Genetic Algorithms in Engineering Systems: Innovations and Applications (GALESIA ’95) (Sheffield, 1995)* (London: IEE) pp 365–72

Sampson J R 1981 *A Synopsis of the Fifth Annual Ann Arbor Adaptive Systems Workshop* Department of Computing and Communication Science, Logic of Computers Group Technical Report University of Michigan

Saravanan N, Fogel D B and Nelson K M 1995 A comparison of methods for self-adaptation in evolutionary algorithms *BioSystems* **36** 157–66

Satterthwaite F E 1959a Random balance experimentation *Technometrics* **1** 111–37

——1959b *REVOP or Random Evolutionary Operation* Merrimack College Technical Report 10-10-59

Schaffer J D (ed) 1989 *Proc. 3rd Int. Conf. on Genetic Algorithms (Fairfax, WA, 1989)* (San Mateo, CA: Morgan Kaufmann)

Schütz M and Sprave J 1996 Application of parallel mixed-integer evolution strategies with mutation rate pooling *Evolutionary Programming V—Proc. 5th Ann. Conf. on Evolutionary Programming (1996)* ed L J Fogel, P J Angeline and T Bäck (Cambridge, MA: MIT Press)

Schwefel H-P 1965 *Kybernetische Evolution als Strategie der experimentellen Forschung in der Strömungstechnik* Dipl.-Ing. Thesis, Technical University of Berlin, Hermann Föttinger Institute for Hydrodynamics

——1968 *Experimentelle Optimierung einer Zweiphasendüse Teil I* AEG Research Institute Project MHD-Staustrahlrohr 11034/68 Technical Report 35

——1974 *Adaptive Mechanismen in der biologischen Evolution und ihr Einflußauf die Evolutionsgeschwindigkeit* Technical University of Berlin Working Group of Bionics and Evolution Techniques at the Institute for Measurement and Control Technology Technical Report Re 215/3

——1975 *Binäre Optimierung durch somatische Mutation* Working Group of Bionics and Evolution Techniques at the Institute of Measurement and Control Technology of the Technical University of Berlin and the Central Animal Laboratory of the Medical Highschool of Hannover Technical Report

——1977 *Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie (Interdisciplinary Systems Research 26)* (Basle: Birkhäuser)

——1981 *Numerical Optimization of Computer Models* (Chichester: Wiley)

——1987 Collective phenomena in evolutionary systems *Problems of Constancy and Change—the Complementarity of Systems Approaches to Complexity, Papers Presented at the 31st Ann. Meeting Int. Society Gen. Syst. Res.* vol 2, ed P Checkland and I Kiss (Budapest: International Society for General System Research) pp 1025–33

——1995 *Evolution and Optimum Seeking* (New York: Wiley)

Schwefel H-P and Bäck T 1995 Evolution strategies II: theoretical aspects *Genetic Algorithms in Engineering and Computer Science Proc. 1st Short Course EUROGEN-95* ed G Winter, J Périaux, M Galán and P Cuesta (New York: Wiley) pp 127–40

Schwefel H-P and Männer R (eds) 1991 *Parallel Problem Solving from Nature—Proc. 1st Workshop PPSN I (Dortmund, 1991) (Lecture Notes in Computer Science 496)* (Berlin: Springer)

Schwefel H-P and Rudolph G 1995 Contemporary evolution strategies *Advances in Artificial Life—Proc. 3rd Eur. Conf. on Artificial Life (ECAL'95) (Lecture Notes in Computer Science 929)* ed F Morán, A Moreno, J J Merelo and P Chacón (Berlin: Springer) pp 893–907

Sebald A V and Fogel D B 1992 Design of fault-tolerant neural networks for pattern classification *Proc. 1st Ann. Conf. on Evolutionary Programming (La Jolla, CA, 1992)* ed D B Fogel and W Atmar (La Jolla, CA: Evolutionary Programming Society) pp 90–9

Sebald A V, Schlenzig J and Fogel D B 1992 Minimax design of CMAC encoded neural controllers for systems with variable time delay *Proc. 1st Ann. Conf. on Evolutionary Programming (La Jolla, CA, 1992)* ed D B Fogel and W Atmar (La Jolla, CA: Evolutionary Programming Society) pp 120–6

Smith S F 1980 *A Learning System Based on Genetic Adaptive Algorithms* PhD Thesis, University of Pittsburgh

Spendley W, Hext G R and Himsworth F R 1962 Sequential application of simplex designs in optimisation and evolutionary operation *Technometrics* **4** 441–61

Sprave J 1994 Linear neighborhood evolution strategy *Proc. 3rd Ann. Conf. on Evolutionary Programming (San Diego, CA, 1994)* ed A V Sebald and L J Fogel (Singapore: World Scientific) pp 42–51

Stewart E C, Kavanaugh W P and Brocker D H 1967 Study of a global search algorithm for optimal control *Proc. 5th Int. Analogue Computation Meeting (Lausanne, 1967)* vol 1, pp 207–30

Takeuchi A 1980 Evolutionary automata—comparison of automaton behavior and Restle's learning model *Information Sci.* **20** 91–9

Wetzel A 1983 *Evaluation of the Effectiveness of Genetic Algorithms in Combinatorial Optimization* unpublished manuscript, University of Pittsburgh