Unearthing a Fossil from the History of Evolutionary Computation

David B. Fogel*

Natural Selection, Inc.

3333 N. Torrey Pines Ct., Suite 200

La Jolla, CA 92037

email: dfogel@natural-selection.com

Abstract. All of science relies on past experimentation and hypotheses. Unfortunately, the science of evolutionary computation is hampered by a general lack of awareness of many early efforts in the field. This paper offers a review of one such contribution from 1967 which employed self-adaptation, co-evolution, and assessed the utility of recombination in various settings. The conclusions, reconfirmed in recent literature, indicate that recombination (uniform or one-point crossover) is best applied in non-epistatic settings. Theoretical analysis supported the experimental findings and now raises questions concerning common applications of schema theory to describe the behavior of evolutionary algorithms.

Keywords: evolutionary algorithms, history of evolutionary computation

1. Introduction

All scientific endeavors rely on past experimentation and theory. To do science requires an iterative process of posing hypotheses, collecting data, analyzing the data in light of the hypotheses, and revising the hypotheses. Current theories should be framed in light of what is known to be true, what is known to be false, and what remains to be discovered. Evidence is required to support or refute theory. Failing to consider prior data, or choosing to ignore it because it does not fit within the scientist's "perfect" theory, degrades the integrity of the procedure, leading

*Address for correspondence: Natural Selection, Inc., 3333 N. Torrey Pines Ct., Suite 200, La Jolla, CA 92037

to a less than efficient search for appropriate explanations of studied phenomena — and yet this occurs all too frequently. Unfortunately, evolutionary computation is not immune to this "selective science." Indeed, it might serve as a prototypical example for students of history to examine years from now.

Evolutionary computation has a long and interesting history. Readers of popular science have been told that the field of study essentially began alone with genetic algorithms in the United States in the early 1960s (Levy, 1992, pp. 155-162; Waldrop, 1992, pp. 170-174; and others). Most practitioners of simulated evolution know otherwise. Many students of evolutionary computation have read of efforts in Germany to use evolution strategies to design physical devices in the mid-1960s (e.g., Rechenberg, 1965; and others), and of earlier evolutionary programming research in the United States to achieve artificial intelligence using variable-length representations (e.g., Fogel, 1962, 1964; and others). But these same students have been presented with considerable misinformation regarding these and other efforts. Certain key publications in evolutionary computation appear to distort or aggrandize different lines of investigation (e.g., Goldberg, 1989; Mitchell, 1996; and others).

For example, Goldberg (1989, pp. 104-106) unfortunately downplayed evolution strategies as being limited in their "schema processing," and dismissed evolutionary programming as incapable of solving other than very simple problems. It is difficult to explain this dismissal; there were many publications prior to 1989 investigating both methods theoretically and experimentally that indicated their practicality across diverse problems (e.g., Burgin, 1969; Rechenberg, 1973; Schwefel, 1981; Fogel, 1988; and very many others). Moreover, research in evolution strategies incorporated not only the cut and splice crossover operators common to genetic algorithms, but other more elaborate forms extending across multiple parents.[1]

If these sorts of inappropriate judgments were limited in scope merely to evolution strategies and evolutionary programming, this would present little problem to students of the history of evolutionary computation. The available literature is replete with examples of these techniques in multiple conferences and archive publications (e.g., Voigt et al., 1996; Angeline et al., 1997; Porto et al. 1998; and others), and efforts to downplay their importance in evolutionary computation are apparent and easily overcome. What is more difficult to overcome are the dismissals of key early work in evolutionary algorithms where the original authors are no longer active or alive, particularly when these efforts essentially duplicate and predate canonical genetic algorithms.

For example, Bremermann (1962), following Bremermann (1958) which concentrated on a mutation-based evolutionary algorithm, offered an optimization procedure with the following elements:

- Simulated chromosomes using binary or floating point representations
- A population of parents that give rise to offspring
- Recombination of pairs or more than two parents

---

[1] Fogel et al. (1966) also proposed a multiple parent majority logic operator within evolutionary programming acting on finite state machines. This was not put in practice, however, due to memory limitations on the size of the resulting offspring machine.

- An ability to handle problems with linear constraints

And this was not the extent of Hans Bremermann's contribution. Indeed the early period of evolutionary computation owes much to Bremermann and Rogson (1964), Bremermann et al. (1965, 1966), Bremermann (1967, 1968, 1970, 1973), and his students (e.g. Goguen and Goguen, 1967). Despite over 35 years of contributions up until his death in 1996, few who are active in evolutionary computation are aware of the magnitude of Bremermann's work and the degree to which it anticipated future developments in genetic algorithms. For example, Mitchell (1996, p. 3) described the introduction of a population-based algorithm using recombination, mutation, and selection in Holland (1975) as a "major innovation," yet Bremermann offered this more than a decade earlier.[2] Moreover, Goldberg (1989, p. 104) stated (unfortunately in error) that Bremermann merely proposed a recombination operator but did not present experimental results. In fact, Bremermann (1962) did indeed offer results of experiments that used both recombination and mutation to evolve solutions to linear systems of equations and linear programming problems; the results were even particularized by the machine time required and a recognition that the difficulty of the search problem in linear systems was dependent on the condition number of the matrix.

Bremermann is not alone in this oversight. Alex Fraser's undertakings have also been largely ignored by the evolutionary computation community, and yet Fraser (1957a) proposed an algorithm consisting of:

- Simulated chromosomes using a binary representation
- A population of parents that give rise to offspring
- Recombination of parents via a generalized n-point crossover operator
- Explicit construction of linkage groups of genes
- Selection based on maximizing, minimizing, or stabilizing against extreme values
- The possibility of varying the number of progeny per parent

Fraser continued to investigate artificial evolution to study properties of genetic adaptive systems throughout the late 1950s and 1960s (e.g., Fraser, 1957b, 1960a, 1960b, 1962; Fraser and Burnell, 1967a, 1967b; and others), culminating in an explicit recognition of evolution as a mechanism for generating purposive behavior (Fraser, 1968), and the textbook Fraser and Burnell (1970). Almost 20 years later, Goldberg (1989, p. 90) reported that "...there was no recognition in Fraser's writing that nature's search algorithm of choice might be useful in artificial settings. It remained for [John] Holland and his students to apply geneticlike operators to artificial problems in adaptation." But this characterization does not appear to do justice to the magnitude of the contribution and the understanding of genetic adaptive systems that Fraser offered. And there

---

[2]Mitchell (1996, p. 3) went further to suggest that Holland (1975) was the first attempt to put computational evolution on a firm theoretical footing, but earlier Bremermann (1958) and Bremermann et al. (1966) offered theory describing the optimum probability of mutation to maximize the rate of convergence in selected problems. Rechenberg (1973) also offered theoretical results on optimum settings for the standard deviation of zero-mean Gaussian mutation in evolution strategies, and there were several other prior efforts to describe computational evolution mathematically (e.g. Goguen and Goguen, 1967; Reed et al., 1967; and others).

were other experiments in simulated genetic adaptive systems that have received essentially no attention whatsoever (e.g., Barker, 1958; Crosby, 1963; and others).

Fortunately, many of these original contributions can be found in libraries and read in their original form. Further, some of the pioneers of these contributions are still available to recount their activities and detail the intended direction of their work. Fogel (1998) provides an edited set of selected readings that includes many of the above-mentioned examples, as well as Box (1957), Friedberg (1958), Friedman (1956), and others, based on the editor's communications with these pioneers and their colleagues collected over the past four years. This research suggests that the idea to simulate evolution on a computer for addressing engineering problems or studying general adaptive systems arose as many as 10 times, independently, over the course of two decades. Indeed, the field of evolutionary computation has no single "father," but rather a population of parents.

When considering the advances made between the early years of evolutionary computation and current lines of inquiry, several concepts and innovations might be considered. For example:

- *Self-adaptation*, where the evolutionary algorithm evolves the manner in which it searches for solutions online
- *Co-evolution*, where individual solutions are measured in light of their interactions with other solutions in the extant population, as opposed to a static evaluation function
- *Demes*, where a population is divided into multiple smaller groups allowing for niching to subproblems and migration of individuals between these groups
- *Emergence*, where a complex adaptive system generates intricate "emergent" behavior from simple locally interactive rules

and certainly other concepts could be offered. But in fact none of these are recent inventions: Self-adaptation appeared as early as Reed et al. (1967), co-evolution was offered in Barricelli (1963), demes were offered in Bossert (1967), and the identification of emergent patterns in an artificial life setting dates back to possibly the earliest efforts in evolutionary computation (Barricelli, 1954).

It is difficult to imagine what the attempts to simulate evolution on computers from the early 1950s must have involved. Even the "large" IBM computers that were used in evolutionary experiments in Friedberg (1958) and Fogel et al. (1965) were only about one-half the speed of an Apple II, with archaic programming by punch cards and no effective debugging tools. When it was first invented, evolutionary computation was an idea about three decades ahead of its time.

The history of evolutionary computation can be usefully applied to guide current investigations. Indeed, the field's direction might have been quite different if more consideration had been given to particular early efforts. Perhaps more attention could have been focused on using multiparent recombination (using more than two parents), following Kaufman (1967), Bremermann (1962), Bremermann et al. (1966), Fogel et al. (1966, p. 21), and others. The study of emergent properties and artificial life might have benefited from greater regard for the artificial ecosystems experiments of Conrad and Pattee (1970). Moreover, several early contributions

provided counterexamples to theoretical arguments and speculations offered mainly within the development of genetic algorithms (e.g., an unquestioned advantage to using crossover in evolutionary algorithms). One such paper was Reed et al. (1967), a truly remarkable contribution that will be reviewed in some detail here.

2. Simulation of Biological Evolution and Machine Learning

2.1. Overview

Jon Reed, Robert Toombs, and Nils Aall Barricelli published some of the earliest experiments in three distinct areas: co-evolution in gaming, self-adaptation, and comparing the effectiveness of different operators. Their paper: "Simulation of Biological Evolution and Machine Learning" appeared in the prominent *Journal of Theoretical Biology* (vol. 17, pp. 319-342, 1967). The framework included the evolution of probabilistic strategies for playing a simplified game of poker using parameters to affect the probabilities of crossover and mutation. The recombination operators included the now familiar one-point crossover, as well as uniform crossover, which was popularized 22 years later in Syswerda (1989). Across a series of different environments with continuous and discrete parameters, Reed et al. (1967) offered the first consistent evidence that recombination could be used to advantage on problems that were highly separable and were coded in a low cardinality alphabet. Their results also indicated that crossover (in the form of discrete recombination) provided no advantage in the cases of continuous parameters or strong interfering interactions between parameters (epistasis).

The poker game centered on hands comprising a single card, which could either be high or low. Three betting options were provided: pass, low bet, and high bet, these being of two, three, or seven units, respectively. The strategies for playing the game were composed of eight parameters, four of which defined probabilities for making various types of wagers, with the other four determining the mutation probabilities and the associated effective step size, as well as the potential for a strategy to crossover with another strategy. These self-adaptive (or so-called "strategy") parameters were also subject to random variation. In a particular setting for which an optimum strategy could be determined via the game theory of Von Neumann, the evolutionary simulation converged on nearly optimal plans, confirming the utility of the self-adaptive method.

Another basic contribution of the paper was the extension of evaluating solutions based on how well they compare with other members of the population, as opposed to a fixed fitness function that was provided *a priori*. This had been explored in an earlier paper by Barricelli (1963), as well as later in Fogel and Burgin (1969), but did not gain real attention again until Axelrod (1987). Other notable contributions of Reed et al. (1967) include the use of a population of as many as 10,000 individuals, which is large even by current standards, and a form of tournament selection to determine which strategies would become parents for the next generation.

2.2.   Detailed Description

More specifically, in the first experimental design, the representation for each player in a population was taken as eight positive numbers, wherein four parameters affected the individual's mutation and crossover properties and the other four defined the betting probabilities under different circumstances. Each strategy was evaluated in a series of 20 games against an opponent from the population. A game consisted of each player receiving one card. There were two types, high cards and low cards, with the probability of receiving a high card described as h. This was taken as $h = 0.5$ initially. Three different bets were allowed: pass, low bet, or high bet. These bets cost two, three, and seven units, respectively. If the two players bet differently, the player with the higher wager won that hand. If the players bet the same amount, the player with the higher card won; otherwise, the hand was a draw.

Four betting probabilities determined the players' strategies, namely the probabilities of (1) betting pass with a low hand (L_P), (2) betting pass with a high hand (H_P), (3) betting low with a low hand (L_L), and (4) betting low with a high hand (H_L). The other possible plays are functions of these parameters. Accompanying these parameters were four additional variables, a, b, c, d, which defined the manner in which the player would generate an offspring. The value $(1 - a)$ was the probability that one of the eight parameters would be replaced by a random number. The value $(1 - c)$ was a probability given to each of the four betting parameters to undergo an increment of size d. Finally, b served as a tag for whether or not the individual would undergo crossover. Depending on the experimental framework, conditions were considered where (1) only patterns with $b > 0.5$ were crossed, initially with uniform recombination, (2) crossover was obligatory regardless of the value of b, or (3) crossover was not allowed, again regardless of the value of b. The population consisted of 50 parents, which generated 50 offspring in each generation.

For $h = 0.5$, the optimum strategies could be determined from game theory (i.e., always bet high with a high card, low with a low card, and never pass), and Reed et al. (1967) remarked that in all three cases of handling crossover, the high hand betting probabilities were nearly optimized in fewer than 200 generations. The low hand betting probabilities were not as well optimized, but these were considered less important for the quality of play. "The quality of the game was fully competitive with average human players uninformed about game theory" (Reed et al., 1967). The fastest rate of optimization was observed when only patterns with $b > 0.5$ were recombined, but interestingly the rate of optimization when crossover was inhibited was greater than when it was mandatory.

The experimental framework was then modified to assess the emergent behavior when betting and variation parameters were taken from a small finite alphabet. The population size was increased to 500 players. Each player consisted of eight parameters encoded in 36 bits. These included: (1) a crossover parameter (3 bits), (2) the probability of betting high with a high card (5 bits), (3) the probability of betting low with a high card if no high bet is made (5 bits), (4) mutation probability (3 bits), (5) probability of betting high on a low card (5 bits), (6) probability of betting low with a low card if no high bet is made (5 bits), (7) mutation size

(5 bits), and 5 additional bits remaining for other assignment. The number of bits determined the degree of precision for each parameter. Although the implementation of mutation here was slightly different than as described above, self-adaptation was still used (the mutations could affect the mutation parameters), and crossover was adopted in a one-point cut and splice procedure (as was later popularized in genetic algorithms). The quality of each strategy was again determined using 20 games played against an opponent selected from the population, but selection immediately eliminated the loser in any pairwise competition and replaced it by a copy (subject to mutation and crossing) of the winner. Again, the three different rules for crossover (free, mandatory, prohibited) were compared. The results (Fig. 1) indicated that the most rapid optimization occurred when crossing over was prohibited.
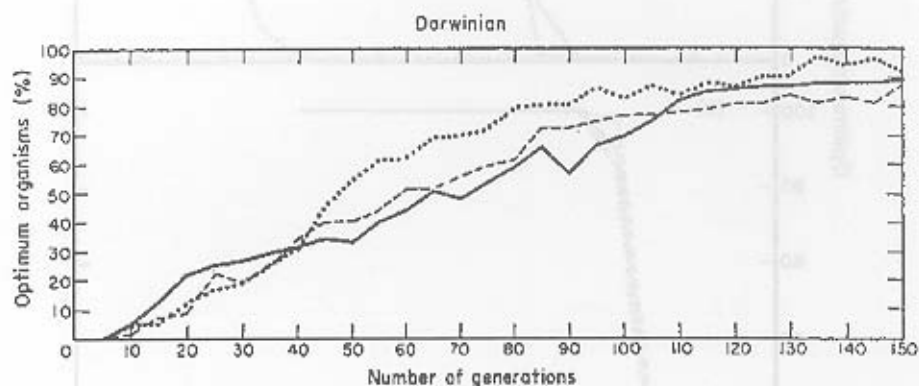

Figure 1.   The mean number of optimized betting strategies as a function of the number of generations averaged over three trials. The dotted line signifies the rate of optimization when crossover was prohibited, the dashed line indicates the performance under mandatory crossing, and solid line indicates the performance under "free" crossing (see text). The figure is from Reed et al. (1967). The title of "Darwinian" was given to describe the experiments with parameters varying over many settings controlled by possibly many "genes," as compared with a later "Mendelian" experiment in which parameters were controlled essentially by single bits that generated all-or-none probabilities of betting high, low, or pass with a high or low hand.

A separate experiment was also presented which was designed to eliminate the interaction among various parameters. Five parameters were considered, coded in binary. There were two frameworks for comparison: (1) in the "Darwinian" test, the sums of all five parameters for each of two players in a population were compared, with the larger sum being assigned the win, and (2) in the "Mendelian" test, only the sums of the leading bits for each parameter were compared. All players were initialized with zeroes in all bits. The same procedures as above were adopted for self-adaptation, mutation, and crossover, but the population size was increased to 10,000. In these two cases, crossing over was seen to dramatically improve performance on the Mendelian set up, but again provided no benefit in the Darwinian experiments (Fig. 2).
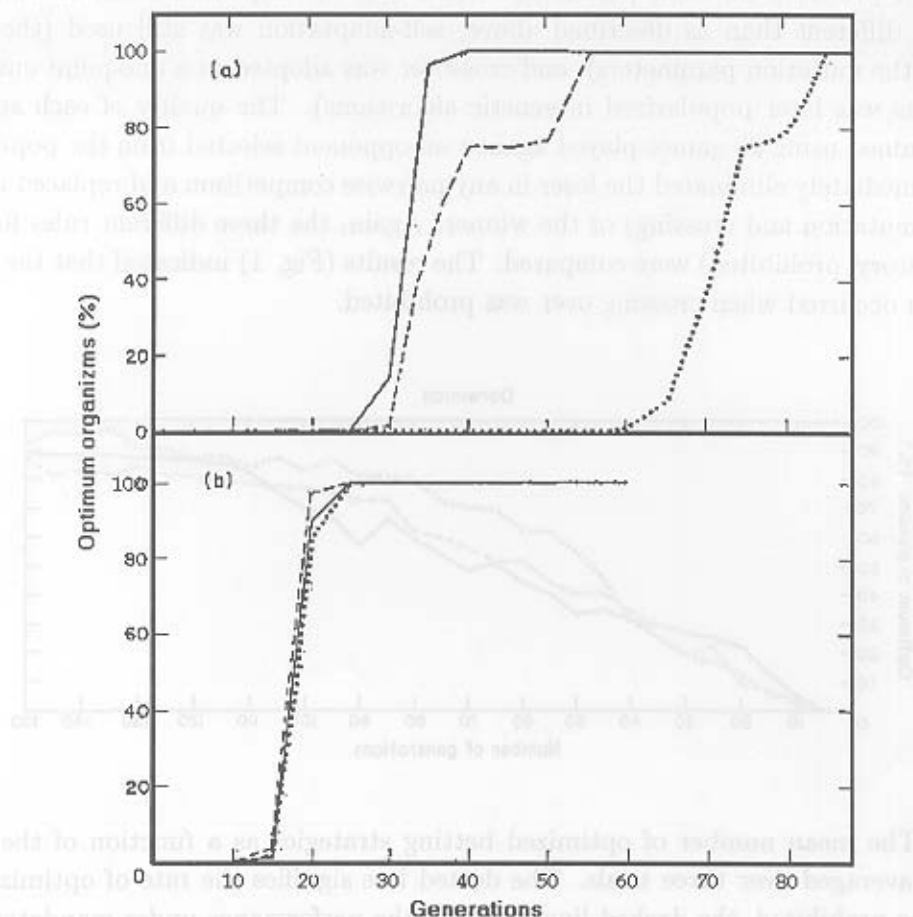
Figure 2. The mean number of optimized players in the (a) Mendelian and (b) Darwinian experiments designed to eliminate the interaction among various parameters. Each curve is the average of four trials. The dotted line corresponds to no crossover, the dashed line corresponds to mandatory crossover, and the solid line corresponds to "free" crossover (self-adapted). The results in the Mendelian case favored free and mandatory crossover; however, in the Darwinian case, there was essentially no observed difference between the methods. The figure is from Reed et al. (1967).

3. Discussion

Several sentences in Reed et al. (1967) convey the surprise of the result that crossing over might provide no benefit to the speed of adaptation (or perhaps even a negative effect). Separability (i.e., independent parameters) was in part identified as a key to expecting effective discrete

recombination methods. Unfortunately, this result did not gain wide notice. Only as recently as Salomon (1996) was it (re)demonstrated that the design of evolutionary algorithms relying on crossover and small mutation rates (i.e., canonical genetic algorithms) is best suited for problems posing independent parameters, and that the optimization performance of such techniques can be greatly reduced under coordinate rotations that induce parametric interdependencies. In retrospect, Reed et al. (1967) provided a preexisting counterexample to the speculation in Holland (1975) that a reproductive plan involving a high rate of one-point crossover and low rate of mutation operating on binary encodings would be generally able to exploit epistatic relationships among parameters.

Although Holland (1975) did not include experimental evidence to support the accompanying theory, the contemporary publication of De Jong (1975) did indicate the potential utility of genetic algorithms for function optimization. Unfortunately, the collection of benchmark functions that were studied have more recently been shown to be mostly separable, and therefore these results did not address the ability of a canonical genetic algorithm to handle epistatic effects, as found so often in real-world problems. Indeed, Davis (1991) offered a simple randomized hill-climbing procedure that served to essentially optimize each parameter in these functions individually, and this trivial algorithm outperformed the genetic algorithm across almost all cases examined, including those from De Jong (1975). Although De Jong (1975) served as a basis for all further genetic algorithm studies and applications into the late 1980s (Goldberg, 1989, p. 107), looking back, this emphasis was misplaced and accepted too easily. The available literature is now laden with examples indicating, in contrast with some strong claims of the universal importance of crossover in evolutionary algorithms (Goldberg and Lingle, 1985; Whitley and Hanson, 1989; and others), that there are cases both supporting and discouraging the use of recombination (Fogel, 1988; Fogel and Atmar, 1990; Schaffer and Eshelman, 1991; Bäck and Schwefel, 1993; Fogel and Atmar, 1994; Chellapilla, 1997, 1998; Angeline, 1997; and many others). Moreover, the "No Free Lunch" theorems of Wolpert and Macready (1997) essentially demonstrated that no single variation operator can have universal benefit across all problems, and Fogel and Ghozeil (1997a) indicated that there are fundamental equivalences between crossover and "mutation" operators (i.e., a two-parent crossover operator can be recast as a two-parent variation that does not employ the mechanism of crossover).

Reed et al. (1967) offered an interesting analysis of the issue of separability in two appendices. The first recapitulated Fisher's (1930) result of the speed of evolution, the competitive advantage that accrues to a species that interbreeds. This speed is defined as:

$$E = \sum_{r=1}^{n} S_r \frac{di_r}{dt}, \tag{1}$$

where t is the time, typically the number of generations, n is the number of positively selected mutations $(M_1, ..., M_n)$ expanding in the population, $S_1, ..., S_n$ are the selective values of the mutations, and i_r is the fraction of the gene population inheriting the mutation M_r. The mean speed of evolution for long time T as taken in relation to the time each individual mutation

requires to increase from minimal representation to covering the entire population is:

$$\overline{E}(T) = \sum_{r=1}^{n} \frac{1}{T} \int_{0}^{T} S_r \frac{di_r}{dt} dt \approx \sum_{r=1}^{n} \overline{S}_r(T) = n\overline{S}, \tag{2}$$

where $\overline{S}$ is the mean of the selective values of the n mutations for the time period T. $\overline{S}$ can be taken as a unit of the speed of evolution over T, leading to a statement of Fisher's Law:

"The maximum speed of evolution a species is capable of (or potential speed of
   evolution) is proportional to the number of genes. A species unable to interbreed is
   to be considered in this connection as a species with only one gene" (Reed et al.,
   1967).

Reed et al. (1967) concluded that achieving this maximum speed depends on having mutations expanding in different genes not interacting with each other epistatically such that they remove or revert the selective value of one of the mutations. This is the effect that was observed under primary experiments with the simplified game of poker.

The second appendix derived a similar result for p quantitative (phenotypic) characters $x_1, ..., x_p$ (such as height, weight, etc.) of individuals in a population. Consideration was given to the time variation of the mean values $\overline{x}_1, ..., \overline{x}_p$:

$$\frac{d\overline{x}_r}{dt} = \sum_{i=1}^{p} V^{ir} \frac{\partial s(\overline{x}_1, ..., \overline{x}_p)}{\partial \overline{x}_i}, \tag{3}$$

where V^{ij} is the covariance of x_i and x_j, and $s(.)$ is the selective value of the set of phenotypic characters. This led to a first approximation for the speed of evolution as:

$$E = \sum_{i,j=1}^{p} V^{ij} \frac{\partial s}{\partial \overline{x}_i} \frac{\partial s}{\partial \overline{x}_j}. \tag{4}$$

Reed et al. (1967) commented: "A main property of these formulas to which we wish to attract the reader's attention is that neither the number of genes nor the contribution of individual genes to the variances and covariances V^{ij} appear in the expressions of E and $d\overline{x}_r/dt$. This means that it makes no difference for the speed of evolution E nor for the time variation of the means $\overline{x}_r$ whether the quantitative characters $x_1, ..., x_p$ are controlled by several genes or by a single gene. In other words, the question whether the species does or does not have a breeding and crossover mechanism has no effect for the speed of quantitative ... evolution, at least in the first approximation range in which the above formulas are valid."

Readers who are familiar with the schema theorem of Holland (1975) will note a similarity between that theorem and the Fisherian relationships indicated above.[3] Discounting the disruptive effects of crossover and mutation, the theorem states that under proportional selection:

$$E \, P(H, t+1) = P(H, t) \frac{f(H, t)}{\overline{f}}, \tag{5}$$

[3]Holland (1997) noted that the schema theorem is an extension of Fisher's law expanded to handle the case of gene combinations.

where H is a particular schema (hyperplane), $P(H, t)$ is the proportion of the population at time t containing H, $f(H, t)$ is the mean fitness of all individuals in the population at time t that contain H, and $\overline{f}$ is the mean fitness of all individuals in the population. In other words, discounting the possibility of random variation on fitness (see Fogel and Ghozeil, 1997b), the expected proportion of the population to contain H in the next time step is proportional to the mean fitness of individuals that contain H relative to the mean fitness of the population. This sampling rate was deemed "nearly optimal" in light of mathematical analysis in Holland (1975) regarding the optimal allocation of trials to a k-armed bandit (each arm was analogous to a hyperplane) in light of minimizing expected losses.[4]

According to Eq. 5, schemata spread on average in each generation in proportion to realized mean fitness, but when Eq. 5 is iterated over successive generations, the mean realized fitnesses $f(H, t), f(H, t + 1), \ldots$, are not constant, nor even necessarily samples from a stationary distribution. Epistatic effects may serve to mitigate the propagation of different schemata. Thus no long-term predictions regarding the sampling rate of various schemata can be made with certainty *a priori*. Rather than being a "fundamental theorem of genetic algorithms" (cf. Goldberg, 1989, p. 33), the schema theorem can only reliably describe the statistical expectation of the propagation of schemata in a single generation based on their observed mean fitness (i.e., not when treated as random variables) and provides no general method for prescribing how to design a more efficient search algorithm. The application of the theorem breaks down under the possibility of epistatic effects, a result anticipated by Reed et al. (1967): "Evidently, Fisher's law was never meant to be applied to a case like this...."

4. Conclusions

Evolutionary computation could have benefited greatly from a more careful early study of the evidence and analysis reported in Reed et al. (1967). Not only was this one of the first systematic comparisons of various search operators, indicating primary data that could have had a significant effect on the emphasis on recombination in evolutionary algorithms, it was also profoundly forward looking in terms of incorporating sophisticated concepts of self-adaptation and competitive fitness evaluations. The simulations involved multiple trials with population sizes that were quite likely the largest applied up to the time of publication. If submitted for peer review today, the paper would likely be accepted for publication. Few contributions to the literature can withstand 30 years of time in this manner.

The student of evolutionary computation must acknowledge that many previous surveys of the field and its historical development have not served them well. This poses a significant challenge of reassessing the contribution of papers that have been misrepresented or misinterpreted, and uncovering other "fossils" that remain to be discovered. It also places the burden of the historian on the student of evolutionary computation, requiring them to assess earlier work in

---

[4]Macready and Wolpert (1996) asserted mathematically that this analysis is in error, and Rudolph (1997) provided a counterexample showing that the derivation in Holland (1975) does not yield minimal expected losses.

the context of the time it was offered, rather than current wisdom. This is particularly difficult when considering not only the obvious limited computing machinery, but also the unavailability of any method for easily communicating research results to the scientific community (e.g., internet or email). Despite these difficulties, there are now opportunities to rediscover and pursue research that has not received due attention. Reed et al. (1967) is merely one of many early explorations into evolutionary computation with the potential to have a profound impact on future investigations.

Acknowledgments

References

P.J. Angeline (1997) "Subtree crossover: Building block engine or macromutation?" Genetic Programming 1997: Proceedings of the 2nd Ann. Conf., J.R. Koza, K. Deb, M. Dorigo, D.B. Fogel, M. Garzon, H. Iba, R.L. Riolo (eds.), Morgan Kaufmann, San Francisco, CA, pp. 9-17.

P.J. Angeline, R.G. Reynolds, J.R. McDonnell, and R. Eberhart (eds.) (1997) Evolutionary Programming VI: Proceedings of the 6th Ann. Conf. on Evolutionary Programming, Springer-Verlag, Berlin.

R. Axelrod (1987) "The evolution of strategies in the iterated prisoner's dilemma," Genetic Algorithms and Simulated Annealing, L. Davis (ed.), Pitman, London, pp. 32-41.

T. Bäck and H.-P. Schwefel (1993) "An overview of evolutionary algorithms in parameter optimization," Evolutionary Computation, Vol. 1:1, pp. 1-24.

J.S.F. Barker (1958) "Simulation of genetic systems by automatic digital computers. III. Selection between alleles at an autosomal locus," Australian J. Biological Sciences, Vol. 11, pp. 603-612.

N.A. Barricelli (1954) "Esempi numerici di processi di evoluzione," Methodos, pp. 45-68.

N.A. Barricelli (1963) "Numerical testing of evolution theories. Part II: Preliminary tests of performance, symbiogenesis and terrestrial life," Acta Biotheoretica, Vol. 16, No. 3-4, pp. 99-126.

W. Bossert (1967) "Mathematical optimization: are there abstract limits on natural selection," Mathematical Challenges to the Neo-Darwinian Interpretation of Evolution, P.S. Moorhead and M.M. Kaplan (eds.), The Wistar Institute Press, Philadelphia, PA, pp. 35-46.

G.E.P. Box (1957) "Evolutionary operation: A method for increasing industrial productivity," Applied Statistics, Vol. 6:2, pp. 81-101.

H.J. Bremermann (1958) "The evolution of intelligence. The nervous system as a model of its environment," Technical Report, No. 1, Contract No. 477(17), Dept. of Mathematics, Univ. Washington, Seattle, July.

H.J. Bremermann (1962) "Optimization through evolution and recombination," Self-Organizing Systems-1962, M.C. Yovits, G.T. Jacobi, and G.D. Goldstein (eds.), Spartan Books, Washington D.C., pp. 93-106.

H.J. Bremermann (1967) "Quantitative aspects of goal-seeking self- organizing systems," Progress in Theoretical Biology, Vol. 1, Academic Press, NY, pp. 59-77.

H.J. Bremermann (1968) "Numerical optimization procedure derived from biological evolution processes," Cybernetic Problems in Bionics, H.L. Oestreicher and D.R. Moore (eds.), Bionics Symposium 1966, Dayton, OH, Gordon and Breach, NY, pp. 543- 561.

H.J. Bremermann (1970) "A method of unconstrained global optimization," Math. Biosciences, Vol. 9, pp. 1-15.

H.J. Bremermann (1973) "On the dynamics and trajectories of evolution processes," Biogenesis, Evolution, Homeostasis, A. Locker (ed.), Springer-Verlag, NY, pp. 29-37.

H.J. Bremermann and M. Rogson (1964) "An evolution-type search method for convex sets," ONR Tech. Report, Contracts 222(85) and 3656(58), Berkeley, CA, May.

H.J. Bremermann, M. Rogson, and S. Salaff (1965) "Search by evolution," Biophysics and Cybernetic Systems, M. Maxfield, A. Callahan, and L.J. Fogel (eds.), Spartan Books, Washington D.C., pp. 157-167.

H.J. Bremermann, M. Rogson, and S. Salaff (1966) "Global properties of evolution processes," Natural Automata and Useful Simulations, H.H. Pattee, E.A. Edlsack, L. Fein, and A.B. Callahan (eds.), Spartan Books, Washington D.C., pp. 3-41.

G.H. Burgin (1969) "On playing two-person zero-sum games against nonminimax players," IEEE Trans. Systems Science and Cybern., Vol. SSC-5, pp. 369-370.

K. Chellapilla (1997) "Evolving computer programs without subtree crossover," IEEE Trans. Evolutionary Computation, Vol. 1:3, pp. 209–216.

M. Conrad and H.H. Pattee (1970) "Evolution experiments with an artificial ecosystems," J. Theoret. Biol., Vol. 28, pp. 393-409.

J.L. Crosby (1963) "Evolution by computer," New Scientist, Vol. 17, pp. 415-417.

L. Davis (1991) "Bit-climbing, representational bias, and test suite design," Proc. of the 4th Intern. Conf. on Genetic Algorithms, R.K. Belew and L.B. Booker (eds.), Morgan Kaufmann, San Mateo, CA, pp. 18-23.

K.A. De Jong (1975) "An analysis of the behavior of a class of genetic adaptive systems," Doctoral dissertation, Univ. Michigan, Ann Arbor.

R.A. Fisher (1930) The Genetical Theory of Natural Selection, Clarendon Press, Oxford.

D.B. Fogel (1988) "An evolutionary approach to the traveling salesman problem," Biological Cybernetics, Vol. 6:2, pp. 139-144.

D.B. Fogel (ed.) (1998) Evolutionary Computation: The Fossil Record, IEEE Press, Piscataway, NJ, in press.

D.B. Fogel and J.W. Atmar (1990) "Comparing genetic operators with Gaussian mutations in simulated evolutionary processes using linear systems," Biological Cybernetics, Vol. 63:2, pp. 111-114.

D.B. Fogel and L.C. Stayton (1994) "On the effectiveness of crossover in simulated evolutionary optimization," BioSystems, Vol. 32:3, pp. 171-182.

D.B. Fogel and A. Ghozeil (1997a) "A note on representations and variation operators," IEEE Trans. Evolutionary Computation, Vol. 1:2, pp. 159-161.

D.B. Fogel and A. Ghozeil (1997b) Schema processing under proportional selection in the presence of random effects," IEEE Trans. Evolutionary Computation, Vol. 1:4, pp.290–293.

L.J. Fogel (1962) "Autonomous automata," Industrial Research, Vol. 4, pp. 14-19.

L.J. Fogel, "On the organization of intellect," Doctoral Dissertation, UCLA, 1964.

L.J. Fogel and G.H. Burgin (1969) "Competitive goal-seeking through evolutionary programming," Final Report under Contract No. AF 19(628)-5927, Air Force Cambridge Research Labs.

L.J. Fogel, A.J. Owens, and M.J. Walsh (1965) "Artificial intelligence through a simulation of evolution," Biophysics and Cybernetic Systems: Proc. of the 2nd Cybernetic Sciences Symposium, M. Maxfield, A. Callahan, and L.J. Fogel (eds.), Spartan Books, Washington D.C., pp. 131-155.

L.J. Fogel, A.J. Owens, and M.J. Walsh (1966) Artificial Intelligence through Simulated Evolution, John Wiley, NY.

A.S. Fraser (1957a) "Simulation of genetic systmes by automatic digital computers. I. Introduction," Australian J. Biological Sciences, Vol. 10, pp. 484-491.

A.S. Fraser (1957b) "Simulation of genetic systems by automatic digital computers. II. Effects of linkage of rates of advance under selection," Australian J. Biological Sciences, Vol. 10, pp. 492-499.

A.S. Fraser (1960a) "Simulation of genetic systems by automatic digiatl computers. 5-linkage, dominance and epistasis," Biometrical Genetics, O. Kempthorne (ed.), Pergamon Press, NY, pp. 70-83.

A.S. Fraser (1960b) "Simulation of genetic systems by automatic digital computers. VI. Epistasis," Australian J. Biological Cybernetics, Vol. 13:2, pp. 150-162.

A.S. Fraser (1960c) "Simulation of genetic systems by automatic digital computers. VII. Effects of reproductive rate, and intensity of selection, on genetic structure," Australian J. Biological Sciences, Vol. 13, pp. 344-350.

A.S. Fraser (1962) "Simulation of genetic systems," J. Theo. Biol., Vol. 2, pp. 329-346.

A.S. Fraser and D. Burnell (1967a) "Simulation of genetic systems. XI. Inversion polymorphism," Am. J. of Human Genetics, Vol. 19:3, pp. 270-287.

A.S. Fraser and D. Burnell (1967b) "Simulation of genetic systems. XII. Models of inverstion polymorphism," Genetics, Vol. 57, pp. 267-282.

A.S. Fraser (1968) "The evolution of purposive behavior," Purposive Systems, H. von Foerster, J.D. White, L.J. Peterson, and J.K. Russell (eds.), Spartan Books, Washington D.C., pp. 15-23.

A.S. Fraser and D. Burnell (1970) Computer Models in Genetics, McGraw-Hill, NY.

R.M. Friedberg (1958) "A learning machine: Part I," IBM J. Research and Development, Vol. 2:1, pp. 2-13.

G.F. Friedman (1956) "Selective feedback computers for engineering synthesis and nervous system analogy," Master's thesis, UCLA.

D.E. Goldberg (1989) Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, Reading, MA.

D.E. Goldberg and R. Lingle (1985) "Alleles, loci, and the traveling salesman problem," Proc. of an Intern. Conf. on Genetic Algorithms and Their Applications, J.J. Grefenstette (ed.), Lawrence Erlbaum, Hillsdale, NJ, pp. 154-159.

J.A. Goguen and N.H. Goguen (1967) "Contributions to the mathematical theory of evolutionary algorithms," ONR Tech. Report, Contracts 3656(08) and 222(85), Berkely, CA, Sept.

J.H. Holland (1975) Adaptation in Natural and Artificial Systems, Univ. Michigan Press, Ann Arbor.

J.H. Holland (1997) Personal communication, plenary lecture, 7th International Conference on Genetic Algorithms, East Lansing, MI.

H. Kaufman (1967) "An experimental investigation of process identification by competitive evolution," IEEE Trans. Systems Science and Cybernetics, Vol. SSC-3:1, pp. 11-16.

S. Levy (1992) Artificial Life: A Report from the Frontier where Computers Meet Biology, Vintage, NY.

W.G. Macready and D.H. Wolpert (1996) "On 2-armed Gaussian bandits and optimization," Tech. report, SFI-TR-96-03-009, Santa Fe Institute, Santa Fe, NM.

M. Mitchell (1996) An Introduction to Genetic Algorithms, MIT Press, Cambridge, MA.

V.W. Porto, N. Saravanan, D.E. Waagen, and A.E. Eiben (eds.) (1998) Evolutionary Programming VII: Proceedings of the 7th Ann. Conf. on Evolutionary Programming, Springer-Verlag, Berlin.

I. Rechenberg (1965) "Cybernetic solution path of an experimental problem," Royal Aircraft Establishment, Library Translation 1122.

I. Rechenberg (1973) Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution, Frommann-Holzboog, Stuttgart, Germany.

J. Reed, R. Toombs, and N.A. Barricelli (1967) "Simulation of biological evolution and machine learning: I. Selection of self-reproducing numeric patterns by data processing machines, effects of hereditary control, mutation type and crossing," J. Theo. Biol., Vol. 17, pp. 319-342.

G. Rudolph (1997) "Reflections on bandit problems and selection methods in uncertain environments," Proc. of the 7th Intern. Conf. on Genetic Algorithms, T. Bäck (ed.), Morgan Kaufmann, San Francisco, CA, pp. 166-173.

R. Salomon (1996) "Re-evaluating genetic algorithm performance under coordinate rotation of benchmark functions. A survey of some theoretical and practical aspects of genetic algorithms," BioSystems, Vol. 39, pp. 263-278.

J.D. Schaffer and L. Eshelman (1991) "On crossover as an evolutionarily viable strategy," Proc. of the Fourth International Conference on Genetic Algorithms, R.K. Belew and L.B. Booker (eds.), Morgan Kaufmann, San Mateo, CA, pp. 61-68.

H.-P. Schwefel (1981) Numerical Optimization of Computer Models, John Wiley, Chichester, U.K.

H.-M. Voigt, W. Ebeling, I. Rechenberg, H.-P. Schwefel (eds.) (1996) Parallel Problem Solving from Nature IV, Springer-Verlag, Berlin.

M.M. Waldrop (1992) Complexity: The Emerging Science at the Edge of Order and Chaos, Simon and Schuster, NY.

D. Whitley and T. Hanson (1989) "Optimizing neural networks using faster, more accurate genetic search," Proc. of the 3rd Int. Conf. on Genetic Algorithms, J.D. Schaffer (ed.), Morgan Kaufmann, San Mateo, CA, pp. 391-396.

D.H. Wolpert and W.G. Macready (1997) "No free lunch theorems for optimization," IEEE Trans. Evolutionary Computation, Vol. 1:1, pp. 67-82.