

Steady-State vs. Generational Genetic Algorithms : A Comparison of Time Complexity and Convergence Properties

David Noever
Subbiah Baskaran

SFI WORKING PAPER: 1992-07-032

SFI Working Papers contain accounts of scientific work of the author(s) and do not necessarily represent the views of the Santa Fe Institute. We accept papers intended for publication in peer-reviewed journals or proceedings volumes, but not papers that have already appeared in print. Except for papers by our external faculty, papers must be based on work done at SFI, inspired by an invited visit to or collaboration at SFI, or funded by an SFI grant.

©NOTICE: This working paper is included by permission of the contributing author(s) as a means to ensure timely distribution of the scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the author(s). It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author's copyright. These works may be reposted only with the explicit permission of the copyright holder.

www.santafe.edu



SANTA FE INSTITUTE

**Steady-state vs. generational genetic algorithms:
A comparison of time complexity and convergence properties**

David Noever* and Subbiah Baskaran⁺

The genetic algorithm (GA) represents a powerful class of search and optimization techniques developed in analogy to genetic laws and natural selection. Best solutions are allowed to evolve subject to some fitness criteria, while internally the mechanics are left largely as a black box. For steady-state GAs, efforts directed towards finding general recursion relations have failed, thus obscuring previous comparisons with the other preferred GA based on generational reproduction. For a binary GA using steady-state reproduction, new recursion relations are found in closed form for three trial cases: (1) constant average population fitness; (2) exponentially bounded variation in average fitness; and (3) constant jump size of average fitness between successive generations. In the latter case, the generational GA is found to be a subset of the steady-state for a jump size, $K=1$. In general the resulting equations provide the relationships of practical interest between estimated run time, problem size, and fitness ratio, along with defining a striking set of new parameters which together give a framework to quantify how the steady-state GA balances its elite selection against the need for diversity and mixing between individual alternatives. Two characteristics of the steady-state analysis are derived as a decay correlation time for the average population fitness (or half-life) and an entropy-like measure of fitness diversity and information exchange within a large population.

**Biophysics Branch, ES-76, National Aeronautics and Space Administration, George C. Marshall Space Flight Center, Huntsville, AL 35812 USA*

+Institute for Theoretical Chemistry, University of Vienna, Währingerstrasse 17, 1-1090, Vienna, Austria

Overview and Problem Statement

The genetic algorithm (GA) is a massively parallel, numerical search and optimization technique which acts in a somewhat analogous fashion to natural genetic laws. As developed in 1975 (Holland, 1975) and subsequently applied to a host of optimization problems in engineering, machine learning and physics (Davis, 1991; Goldberg, 1989), the method operates on partial (unoptimized) solutions and generates mutated or recombined copies of possible alternatives until a good solution emerges or evolves. At each generation step, the partial solutions can be mixed using a cross-over operation which exchanges segments of these partial solutions; this operation draws its significance from natural genetics and reproduction, where the chromosomes play the role of partial solutions or schemas which exchange information prior to subsequent reproduction and copying to the next generation. The internal mechanics of how the genetic algorithm arrives at its optimal solution remain a topic of great interest.

In general the mathematical workings of the GA's operation has remained somewhat shrouded: it may work, but how? This note explores one analytical approach to understand how such numerical searches can select and promote elite solutions at the expense of poor solutions, *while at the same time keeping a steady and robust pool of new alternatives*. The question is one of diversity and the concept of a population's disorder or variety is compared between two versions of the genetic algorithm: proportional selection both at each generation and at steady-state (Goldberg and Deb, 1991; Syswerda, 1991; Ankenbrandt, 1990).

To anticipate the outcome, closed form recursion relations are derived for example cases of the steady state method which allows birth and death of solutions. The time

evolution of the steady-state's average fitness is compared with generational selection for both time complexity and convergence properties (Shaffer, et al., 1989; Goldberg and Deb, 1991; Grefenstette, 1986). The analysis extends the inductive proofs of Ankenbrandt (1991) to include steady-state alternatives. Goldberg and Deb (1991) treated one version of the steady-state GA (Whitley's Genitor) and derived an integrated solution valid after many generations. (e.g. when a time derivative approximates the finite difference between generations). In agreement with Syswerda (1991), they generally remarked that other than by direct iteration, *a simplified solution for steady-state reproduction is non-trivial*. The first section of this note will consider various limited generalizations which lead towards the goal of closed analytic forms for the steady-state GA. Syswerda (1991) previously took up the problem of comparing steady-state and generational GAs and pointed out that while the two GA versions share their theoretical foundations in the schema theorem, in practice their reproductive behavior can differ markedly. The present aim is to understand further some of the hidden behavior which underlies the GA's remarkable success.

Recursion Relations for Proportionate Selection

General use of proportionate selection in the GA currently favors two reproduction strategies: either generational or steady-state. Syswerda (1991) has drawn the distinction clearly and identified generational reproduction as a method which replaces the entire population at each generation, while steady-state reproduction replaces only a few individuals at a given time. By numerically iterating their respective reproductive performance, he showed that the two methods can be made to look similar within the schema theorem (Holland, 1975), that is, to favor fit solution by increasing their

reproductive rate. A comparative summary of each method's capabilities and shortcomings appears in Table I.

Case I: Generational Reproduction

The Recurrence Relation

For discrete time steps between generations, $t=\{0,1,2,\dots\}$, let $P_{i,t}$ correspond to the proportion of alleles (or bit string values) set to the value one for a particular allele position i at generation t . Let $P_{i,0}$ represent P for the founder generation, $t=0$. For simplicity, all subsequent work will treat binary genetic algorithms which have alleles possessing either of two values, zero or one; the multivalued case is a trivial generalization at this level of proof. Let f_1 correspond to the organism's fitness (some survival probability) sampled with allele value one in a particular position j . Likewise, take f_0 to represent the fitness of all organisms sampled with allele value zero in position j . For any defined fitness ratio, $r=f_1/f_0$, the value will be considered time-independent and constant across generations. Ankenbrandt (1991) has discussed the alternatives to this assumption and shown that time-dependent fitness ratios can be placed within an analogous framework. More specifically, the minimum fitness ratio for the entire simulation appears to approximate the cross-generational fitness ratio; additionally the discrete fitness ratio can be examined directly prior to copying the population to the next generation.

Using an inductive proof and solving by inspection, Ankenbrandt (1991) has derived the solution for recursion in the generational model. Goldberg and Deb (1991) put forward the same solution in a compact form:

$$P_{i,t} = \frac{r_j^t P_{i,0}}{\sum_j r_j^t P_{j,0}} \quad (1)$$

where r is a polynomial of order equal to the generational time step, t . Since the inductive proof requires evaluation of only three terms, $P_{i=1,2,3}$, the derivation is repeated here using a slightly different formalism which will prove most valuable in subsequent comparisons with the steady-state version of reproduction.

Appendix A derives equation (1) in a particularly transparent notation. For a fitness ratio, $r=f_1/f_0$, two new terms can be introduced: the inverse proportion, $x_t=1/P_t$, and the fitness difference, $\Delta f=(f_1-f_0)/f_0$. Equation (1) can be rewritten simply as

$$x_{t+1} = \frac{1}{r} [\Delta f + x_t] \quad (2)$$

This more straightforward form can be iterated recursively and the closed form solution can be identified by inspection:

$$x_t = [(r-1) + x_{t-1}]/r = 1 - (1-x_0)/r^t \quad (3)$$

Substitution for x in (3) gives agreement with equation (1) and its time complexities derived previously (Ankensbrandt, 1991; Goldberg and Deb, 1991).

Time Complexity

Time complexity of a GA refers to the functional dependence of convergence time on

the population size and fitness. Two cases of time complexity can be considered: worst and average limits. The worst case complexity corresponds to the conditions ($P_f=1-1/n$; $P_o=1/n$), where $1/n$ is some small number corresponding the proportional representation of a single fit individual initially P_o and a near convergence on that fit individual which ends the simulation at P_f . For generational reproduction, the time complexity follows (Ankensbrandt, 1991) as:

$$t_c = \frac{\ln \left[\frac{P_f(1-P_o)}{P_o(1-P_f)} \right]}{\ln r} = \frac{\ln (n-1)^2}{\ln r} \quad (4)$$

In contrast, the average time complexity corresponds to the conditions ($P_f=1-1/n$; $P_o=0.5$), where $P_o=0.5$ is the proportion corresponding the random representation of a binary algorithm and again a near convergence on a fit individual which ends the simulation at P_f . The time complexity follows as:

$$t_c = \frac{\ln \left[\frac{P_f(1-P_o)}{P_o(1-P_f)} \right]}{\ln r} = \frac{\ln (n-1)}{\ln r} \quad (4)$$

A comparative plot of average and worst case time complexity is shown in Fig. 1 for various fitness ratios, r , and population sizes, n . In general, the worst case time complexity is twice the average case.

Case II: Steady State Reproduction

The steady-state recursion relation differs from generational reproduction owing to the birth and death of new individuals in each generation. The simple population balance for a single individual which is copied and one which is deleted is:

$$P_{t+1} = P_t \left[1 - \frac{1}{n} \right] + P_t \left[\frac{f_1}{S_t} \right] \quad (6)$$

where $S_t = P_t f_1 + (1 - P_t) f_0$ is the total average population fitness and n is the population size. Physically the first term on the right hand side represents individuals which live uncopied from one generation t to generation $t+1$, the second term represents the death of an unfit individual (required to keep population size constant), and the last term represents the copying of one individual with a reproductive rate, f_1/S_t . In general no simple closed form equation exists for this recursion relation; the failure is a direct consequence of the non-linearity which arises from the need to evaluate an updated version of the average population fitness, S_t , at each generation. This function S_t will be considered in subsequent section with care to search for approximations which permit a closed form comparison with the generational relations and time complexity. Here consideration is given to two possible techniques for evaluating the approximate behavior of the most fit proportion, P_t , in the steady-state GA. An integrated form is found valid for large populations, n , and many generations into the simulation. Secondly, a new model which assumes constant jump fraction, P_{t+1}/P_t , is proposed and found to possess interesting mathematical properties.

Example A: Differential equation for time evolution of the fittest proportion, $P(t)$

An Integrated Solution

In the limit of large population sizes, n , the approximation holds that $P_t(1 - 1/n) = P_t$. In this case, after many generations, the difference equation (6) can be rewritten simply as an ordinary differential equation in the fitness ratio, $r = f_1/f_0$ as:

$$\frac{dP}{dt} = \frac{r}{(r-1) + \frac{1}{P_t}} \quad (7)$$

Elementary integration gives the time dependence for the proportion of most fit individuals.

Theorem 1: In the limit of large population sizes, $n \gg 1$, and many generations, t large ($t \gg 1$), then the time evolution of the most fit proportion is

$$t = \frac{(r-1)}{r} P + \frac{1}{r} \ln P \quad (8)$$

and the corresponding convergence time is

$$t_c = \frac{(r-1)}{r} (P_f - P_o) + \frac{1}{r} \ln \frac{P_f}{P_o} \quad (9)$$

For early generations (small t) the integration overestimates the convergence rate, but becomes a better approximation for late generations.

Time Complexity

As was the case for the generational GA, two cases of time complexity can be considered: worst and average limits. The worst case complexity corresponds to the conditions ($P_f = 1 - 1/n$; $P_o = 1/n$), where $1/n$ is some small number corresponding the proportional representation of a single fit individual initially P_o and a near convergence on that fit individual which ends the simulation at P_f . The time complexity follows as:

$$t_c = \frac{(r-1)}{r} + \frac{\ln(n-1)}{r} \quad (10)$$

In contrast, the average time complexity corresponds to the conditions ($P_f = 1 - 1/n$; $P_o = 0.5$), where $P_o = 0.5$ is the proportion corresponding the random representation of a binary algorithm and again a near convergence on a fit individual which ends the simulation at P_f .

The time complexity follows as:

$$t_c = \frac{(r-1)}{r} \left(\frac{n-2}{2n} \right) + \frac{1}{r} \ln \left[\frac{2(n-1)}{n} \right] \quad (11)$$

A comparative plot of average and worst case time complexity is shown in Fig. 2 for various fitness ratios, r , and population sizes, n .

Example B: A constant jump size model: $K=1$ condition implies equivalence between generational and steady-state GA

While the integrated version of the steady-state GA is compact, its form remains somewhat unsatisfying for comparing with generational recursion relations. To arrive at the differential equation (7), the death process must be entirely neglected (valid for large populations, n). To address this uncertainty, this section proposes a well-defined problem with the constraint that the steady-state GA will increase the proportion of fit individuals by a constant jump fraction from one copy to the next. Syswerda (1991) first showed that with regards to the schema theorem, a random selection-and-deletion method best approximates the steady-state for comparison with generational reproduction. He found that in the random case, the steady-state iteration looks similar to generational iteration, except the steady-state is compressed over fewer total generation steps. The present case of a constant jump fraction is found to provide a useful alternative to random selection. The principal finding here is a new condition which establishes an equivalence between generational and steady-state GAs. *Generational reproduction appears as a subcase of the steady-state* for the particular condition that the jump fraction $K=1$. Subsequently, this condition will be called the $K=1$ matching condition for equivalence.

The Recursion Relation

Again introducing the notation for the inverse proportion, $x_t = 1/P_t$, and the fitness difference, $\Delta f = (f_1 - f_0)/f_0$. Equation (6) can be rewritten in Appendix A as

$$x_{t+1} = \frac{K}{r} [\Delta f + x_t] \quad (12)$$

for the case of a constant jump step for the proportion of most fit individuals between generations. This condition arises formally as

$$\frac{x_t - x_{t+1}}{x_t} = K \quad (13)$$

where K is defined as a non-negative constant for all increasing P_t .

By comparing the steady-state recursion (12) with the generational recursion (2), inspection reveals that for a constant jump fraction, generational reproduction is a subcase of steady-state reproduction for $K=1$. Thus, *the matching condition, $K=1$, implies dynamic equivalence.*

Theorem 2: The generational and steady-state GA show dynamic similarity subject to the assumption of a relatively large population size, $1/n \ll 1$, and $K=1$ for constant population jump size, $K = [P_{t+1} - P_t]/P_t$. The $K=1$ case is called the matching condition for equivalence.

In this way, the constant jump size can be thought to correspond physically to a linear stretch factor which maps the generational GA onto the broader steady-state behavior. This idea, while intuitively satisfying, also makes contact with Syswerda's (1991) arguments that the generational GA is identical to the steady-state GA with regards both to its characteristic convergence on a fit individual and the schema theorem. By brute force

iteration of the recursion relations, he succinctly showed that the steady-state GA corresponds to a compressed version of the generational GA. As derived analytically here, the measure of this compression is introduced as K .

Extending this idea, the compression factor, K , acts to pressure the steady-state population towards convergence. Large K values show earlier convergence and correspond physically to a greater jump size for the fittest proportion from one generation to the next. An additional practical feature of this formulation is that for certain classes of problems, such early convergence may prove disadvantageous. Spurious solutions may arise as local minima or random noise and thus prevent optimal selection. The problem has been well-examined in the GA literature and a number of ingenious ways around the difficulty have appeared. The present derivation raises this issue in another context. Consider the following conditions for $K < 1$ and steady-state reproduction applied to a set of prematurely convergent problems. Since the generational GA corresponds to $K = 1$ and frequently converges slower than the steady-state alternative, then equation (12) suggests that retuning the K compression factor to a lower value (or a value less than one) may provide a practical way to adjust convergence in a single stroke. Thus the matching condition $K = 1$ for equivalence not only poses a new criterion for comparing competing GA versions, but more importantly gives a tunable parameter for getting around the persistent problem of premature convergence and non-optimality.

The Recursion Relation

A direct test of the K values' significance to convergence depends on deriving the appropriate recursion relation. Inspection of (12) and (2) suggests that the steady-state

recursion should follow from the previously written relation for generational reproduction. However, close examination reveals that a fortuitous cancellation of terms in the generational case produces its simple form. The complete recursion relation for the corresponding steady-state case depends on recognizing that successive iteration give a geometric series in (K/r) which can be summed in closed form (for $K/r < 1$).

Theorem 3: For a constant jump size (K) model of generational reproduction, the K compression factor introduces the recursion relation:

$$x_{t+1} = \frac{K\Delta f}{r} \left[\sum_{\beta=0}^t \left(\frac{K}{r}\right)^\beta \right] + \left(\frac{K}{r}\right)^{t+1} x_0 \quad (14)$$

and for the specific condition, $(K/r) < 1$,

$$x_t = \frac{K\Delta f}{r-K} + \left(\frac{K}{r}\right)^t x_0 \quad (15)$$

The derivation is given in Appendix A and the summation is a geometric series in (K/r) .

For $(K/r) < 1$, this series sums to:

$$\sum_{\beta=0}^t \left(\frac{K}{r}\right)^\beta = \frac{1}{1 - \frac{K}{r}} = \frac{r}{r-K} \quad (16)$$

and the recursion relation (15) follows for constant jump size. The closed form (15) for the steady-state GA is the principal result of this section. Figure 3 compares the convergent behavior of the recursion relation for generational vs. steady-state and the $K=1$ matching condition .

Time Complexity

Elementary solution of (15) for steady-state reproduction gives the corresponding time complexity of the constant jump model.

Theorem 4: For a constant jump size of size K , generational reproduction yields an exact time complexity in the closed form

$$t_c = \frac{\ln\left(\frac{x_o}{x_t - \left(\frac{K\Delta f}{r-K}\right)}\right)}{\ln\left(\frac{r}{K}\right)} \quad (17)$$

Substituting $x_t = x_f = 1/P_f$ and $x_o = 1/P_o$, both worst and average time complexity can be studied. To make explicit the relation between time complexity both in the generational case and in the steady-state, constant-jump model, equation (17) can be modified by adding the negligibly small term $(K/r)^t$. This is of zero order within the assumptions $(K/r) < 1$ and many generations $t \gg 1$. This gives the particularly transparent form

$$x_t = \frac{K\Delta f}{r-K} + \left(\frac{K}{r}\right)^t (x_o - 1) \quad (18)$$

which equals the identical time complexity (4) as found in the generational case (for $K=1$)

$$t_c = \frac{\ln\left[\frac{P_f(1-P_o)}{P_o(1-P_f)}\right]}{\ln r} \quad (19)$$

and the general form (for $K > 1$):

$$t_c = \frac{\ln\left[\frac{(1-P_o)}{P_o} \frac{P_f(r-K)}{(r-K) - P_f K(r-1)}\right]}{\ln\left(\frac{r}{K}\right)} \quad (20)$$

As expected, the time complexity of the steady-state model equals the generational case for $K=1$.

As was the case for the generational GA, two cases of time complexity can be considered: worst and average limits. The worst case complexity corresponds to the

conditions ($P_f=1-1/n$; $P_O=1/n$), where $1/n$ is some small number corresponding the proportional representation of a single fit individual initially P_O and a near convergence on that fit individual which ends the simulation at P_f . The time complexity follows as:

$$t_c = \frac{\ln \left[\frac{(n-1)^2 (r-K)}{(r-K)n-(n-1)K(r-1)} \right]}{\ln \left[\frac{r}{K} \right]} \quad (21)$$

In contrast, the average time complexity corresponds to the conditions ($P_f=1-1/n$; $P_O=0.5$), where $P_O=0.5$ is the proportion corresponding the random representation of a binary algorithm and again a near convergence on a fit individual which ends the simulation at P_f . The time complexity follows as:

$$t_c = \frac{\ln \left[\frac{(n-1) (r-K)}{(r-K)n-(n-1)K(r-1)} \right]}{\ln \left[\frac{r}{K} \right]} \quad (22)$$

A comparative plot of average and worst case time complexity is shown in Fig. 4 for various fitness ratios, r , and population sizes, n .

Properties of Average Population Fitness

For a given recursion relation on the fit proportion, equivalent restrictions are implicitly put on the available iterative form for the average population fitness. This section considers the related question of: How does the average population fitness evolve with time given a recursive formula for the fit proportion? In three test cases, a closed form recursion can be found for the average population fitness, S_t and compared between cases and between generational and steady-state versions of the GA. Thus convergent behavior can arise

analytically given some steady state bounds; here, the overall population fitness is considered as a constant, an exponentially bounded function and the most general form written from the definition of the steady-state equation.

Steady State Bounds and Convergent Behavior

Test Case 1: Exponentially Bounded Average Population Fitness: Finding a Characteristic Time, λ

In this section, the question is posed: what is the time dependent nature of the average population fitness, S_t , during steady-state reproduction? What bounds can be placed on the fitness convergence of a successful algorithm? More specifically, one would like to define a characteristic time for the average fitness which captures its dynamic behavior. To do this, an exponential trial function is proposed and a decay time is extracted which depends only on the population size, n , and fitness ratio, $r=f_1/f_0$. In actual practice, the use of an exponential decay places relatively weak pressure on the population fitness.

For a given intergenerational time step, the average population fitness is constrained to be exponentially bounded in its approach to convergence. This condition corresponds to some limit on the rate on premature convergence. Mathematically this constraint can be written as:

$$[S_{t+1}-S_t] < S_t(e^{-\lambda t}-1) \quad (23)$$

for $\lambda t < 1$ or equally, $t < t_c$. Here the parameter λ is related to some constant decay time for the successive differences in the average population fitness. Thus for a bounded convergence, the equivalent recursion relation for the average population fitness, S_t , can be

written as:

$$S_{t+1}=S_t e^{-\lambda t} \quad (24)$$

which by induction has the closed form solution,

$$S_{t+1}=S_0 e^{-\lambda t^2} \quad (25).$$

Relation (25) is physically useful since its elegant and simple form allows one to solve directly for the characteristic time, λ , in a way that reflects the GA's initial parameters (P_0, S_0, n).

By substitution of (25) into the steady-state equation,

$$P_{t+1}=P_t \left[1 + \frac{f_1}{S_0 e^{-\lambda t^2}} - \frac{1}{n} \right] \quad (26)$$

For large populations, n , and many generations, the finite difference can be approximated by the differential and integrated formally to give:

$$P(t)= \exp \left[\frac{f_1}{S_0} \int e^{\lambda t^2} dt \right] \quad (27)$$

For non-optimal (low fitness) and decreasing populations in which $P_{t+1}-P_t < 0$, then the decay time, λ , is negative and the integrand in (27) is the Gaussian error function centered at zero time with variance $\sigma = (-2\lambda)^{-1}$.

In the more general case, however, the recursive formula (26) can be iterated directly and its closed form recovered from inspection. Rewriting (26) in the streamlined form suitable for recursion gives:

$$P_{t+1}=P_t [\beta + \alpha e^{\lambda t^2}] \quad (28)$$

In this case, $\beta = (1-1/n)$, $\alpha = f_1/S_0$, and $S_0 = f_1 P_0 + (1-P_0)f_0$. By repeated iteration and induction, the closed form version of (28) is

$$P_t = P_0 \prod_{m=0}^t (\beta + \alpha e^{\lambda m^2}) \quad (29)$$

To solve directly for the characteristic time, λ , the recursion relation (29) can be recast in the neighborhood of convergence ($t=t_c$), such that for the worst case time complexity (when $P_f/P_o=n-1$), then

$$\frac{P_f}{P_o} = n-1 = (\beta + \alpha e^{\lambda t_c^2}) \prod_{m=0}^{t_c-1} (\beta + \alpha e^{\lambda m^2}) \quad (30)$$

where the right hand side of (27) has been factored. Rearranging (30) gives

$$(\beta + \alpha e^{\lambda t_c^2}) = \frac{(n-1)}{\prod_{m=0}^{t_c-1} (\beta + \alpha e^{\lambda m^2})} \quad (31)$$

Since the generational product of the average fitness (the denominator of equation (31)) is very large, one can approximately rewrite (31) as

$$-\beta = \alpha e^{\lambda t_c^2} \quad \text{or equivalently, } \lambda = \frac{1}{t_c^2} \ln \frac{\beta}{\alpha} = \frac{1}{t_c^2} \ln \left[\frac{(n-1) S_o}{n f_1} \right] \quad (32)$$

Finally, by substituting for the definition of S_o and rewriting the bracketed logarithmic terms as $\ln[(n-1)\{1+(n-1)/r\}/n^2]$, then for large populations ($1/n \ll 1$) the characteristic time for convergent approach of the overall population fitness follows in terms of the fitness ratio and convergence time only. Hence,

Theorem 6: For an exponentially bounded approach of the average overall population fitness which satisfies the convergence condition,

$$[S_{t+1} - S_t] < S_t(e^{-\lambda t} - 1) \quad (33)$$

then in the limit of large population sizes, ($1/n \ll 1$), it follows that a characteristic time for average population fitness, λ , can be found equal to:

$$\lambda = \frac{1}{t_c^2} \ln \left(\frac{1}{r} \right) = \frac{1}{t_c^2} \ln \left(\frac{f_o}{f_1} \right) \quad (34)$$

Result (34) is the principal result of this section. The beautiful feature of its simple closed form is that it allows a GA optimization estimate for the convergent approach of the overall population fitness *analytically prior to a given numerical iteration*. That the characteristic time depends only on the convergence time and the fitness ratio makes this convergence guide of practical significance; in general, these parameters are known for a given population size and fitness assignment values.

To complete the discussion, one can use the characteristic time constant λ to define a half-life decay time for the difference in average population fitness, $S_{t_{1/2}}=2S_0$, or $S_t/S_0=1$. Direct substitution in the recursion relation (25) for $S_t=S_{t_{1/2}}$ gives the half-life for decay correlation in the average population fitness.

$$t_{1/2} = \sqrt{\frac{\ln(\frac{1}{2})}{\lambda}} = t_c \sqrt{\frac{\ln 2}{\ln r}} \quad (35)$$

This section concludes with:

Proposition 1: For an exponentially bounded approach of the average overall population fitness which satisfies the convergence condition,

$$[S_{t+1}-S_t] < S_t(e^{-\lambda t}-1) \quad (36)$$

then in the limit of large population sizes, ($1/n \ll 1$), it follows that a half-life for the decay correlation for differences in the average population fitness, $t_{1/2}$, can be defined which depends only on the fitness ratio, r , and the takeover time, t_c

$$t_{1/2} = t_c \sqrt{\frac{\ln 2}{\ln r}}$$

For a worst case takeover time using steady-state reproduction, the dependence of the general time t_c on population size, n , and fitness ratio, r , is given by equation (10). Thus for any given population size and fitness ratio, the half-life can be approximated as:

$$t_{1/2} = \left\{ \frac{(r-1)}{r} \left(\frac{n-2}{n} \right) + \frac{1}{r} \ln(n-1) \right\} \sqrt{\frac{\ln 2}{\ln r}} \quad (37)$$

Physically this half-life gives a characteristic number of generational time steps for the average population fitness to reach twice its initial value, $S_{t_{1/2}} = 2S_0$. This practically gives a concrete measure of correlation times for fitness which depends only on initial parameter selection (Figure 5).

Test Case 2: Constant Average Population Fitness: Defining an Appropriate Step-size, κ

To extend the analysis of the overall population fitness, the second test case considers a constant average fitness. This assumption is generally a poor approximation within the GA, since convergence relies on an ever-changing fitness for the population. However, the case is presented here to highlight a possible role for *defining a minimum population step size* between generations, in particular one which can give well-behaved convergence. One definition for the population fitness can be written as:

$$S_t = \frac{f_1}{\kappa} \quad (38)$$

This form is inferred from the previous test case with a constant jump size between most fit proportions, $P_{t+1}/P_t = f_1/S_t = \kappa$. Equation (38) states that the average fitness is constant and proportional to the fitness of one allele, f_1 . With this definition, the steady-state GA can be iterated to get the recursion relation in closed form

$$P_{t+1} = \left[1 - \frac{1}{n} + \kappa \right] P_t \quad (39)$$

and convergence time

$$t_c = \frac{\ln \left[\frac{P_f}{P_o} \right]}{\ln \left[1 + \frac{\kappa}{n} \right]} \quad (40)$$

For the standard cases of time complexity, both average and worst values can be found as before, but now as a function of (n, κ) . The worst and average cases for convergence time are:

$$t_c = \frac{n^2 \ln(n-1)}{\kappa} \quad (\text{worst case}); \quad t_c = \frac{n^2 \ln \left[\frac{2(n-1)}{n} \right]}{\kappa} \quad (\text{average case})$$

To give physical meaning to the value of κ , one can define a normalized step size for iteration as, $\epsilon = \kappa/n$, where ϵ equals the ratio between generations of the fit proportion P_{t+1}/P_t divided by the total population size. In this way, the parameter ϵ represents a scaled step size between generations of the overall population fitness.

To understand the dynamics of successive steps between generations and the observable consequences on this scaled step size, equation (40) can be rewritten as:

$$t_c = \frac{\ln \left[\frac{P_f}{P_o} \right]}{\ln \left[P_f + \frac{\epsilon}{n} \right]} \quad (41)$$

Equation (41) has the advantage that for small step sizes relative to the total number of individuals in the population ($\epsilon/n \ll 1$), then the upper bound on $P_f < 1$ makes the logarithmic approximation valid, $\ln[P_f + \epsilon/n] = \ln[1 + \epsilon/n] = \epsilon/n$ and

$$\epsilon = \left(\frac{P_f}{P_o} \right)^{\frac{n}{t_c}} \quad (42)$$

Therefore, for a worst case convergence ($P_f = 1 - n^{-1}$, $P_o = n^{-1}$), the scaled step size varies dynamically as $\epsilon = (n-1)^{n/t_c}$. In other words, for a particular population size n and a desired

number of generations to convergence, t_c , then the proportional increase in the most fit members should be at least

$$\kappa = \frac{P_{t+1}}{P_t} < n(n-1)^{\frac{n}{t_c}} \quad (43)$$

This relation allows the operator of a steady-state GA to guide the convergence by monitoring the intergenerational steps in the most fit population and comparing the ratio to the worst case scenario (43). Clearly, for longer convergence times, ($t_c \gg 1$), then $P_{t+1}/P_t = n$ and the proportional fitness evolves as time independent. Conversely for short convergence times ($n/t_c \gg 1$) and large populations, then $P_{t+1}/P_t = n^n/t_c$ and the proportional fitness evolves intimately as a power of the population size. The main result of this section is equation (43), namely the derivation of a lower bound on the intergenerational step size between the most fit proportion, P_{t+1}/P_t , along with the practical suggestion that *this criteria can guide on-line monitoring of steady-state convergence*. Equally, a physical meaning is attached to the requirement that population steps remain constant as a kind of scaled step size between generations.

Test Case 3: The General Recursion for Arbitrary Fitness: The Entropy Function for Diversity and Genomic Information

This section takes up the main questions explicitly: How does the genetic algorithm maintain diversity while selecting and favoring fit individuals? This prompts a related question which sheds some light on the issue of informational diversity in a population, namely how does the average population fitness evolve in time? For steady-state reproduction, an equivalent form to the recursion relation derived in (6) for the proportion

of fit individuals, P_t , can be written for S_t and solved (by integration) to give the time dynamics. The result of the following analysis will be a logarithmic, entropy-like definition for population diversity as a single fit individual emerges towards convergence.

From the definition of average overall population fitness (6), then

$$P_t = \frac{S_t - f_o}{f_1 - f_o}, \text{ or } P_{t+1} - P_t = \frac{S_{t+1} - S_t}{f_1 - f_o} \quad (44)$$

Likewise from the steady-state recursion relation for large populations,

$$P_{t+1} - P_t = \frac{f_1 P_t}{S_t} \quad (45)$$

Combining these equations (44) and (45) gives the recursion relation for average population fitness equal to:

$$S_{t+1} - S_t = f_1 \frac{f_o}{S_t} = \frac{\beta S_t + \alpha}{S_t} \quad (46)$$

where $\beta = f_1$ and $\alpha = -f_1 f_o$.

To ask for the time dependence of the average fitness requires the integrated form valid beyond the very early generations ($t \gg 1$), where $dS/dt = S_{t+1} - S_t$. The integral can be performed by parts and the time dependence extracted as follows.

Theorem 5: During steady-state reproduction, the average population fitness $S(t)$ evolves with time as the following form in the continuous limit for $\beta = f_1$ and $\alpha = -f_1 f_o$

$$t = \left(S - \frac{1}{\beta} \right) \ln (\beta S + \alpha) \quad (47)$$

Two interesting cases emerge. First consider a simplified version of (47) valid for small β values (small fitness, f_1) such that the approximation holds that $(S - \beta^{-1}) = -\beta^{-1}$ and

$$(\beta S + \alpha)^{\frac{-1}{\beta}} = e^t, \text{ or } t = \frac{-1}{\beta} \ln (\beta S + \alpha) = \frac{-1}{\beta} (\ln \alpha + \ln [\frac{\beta S}{\alpha} + 1]) \quad (48)$$

With the additional stipulation that $\beta S/\alpha \ll 1$, and $\ln(1+x)=x$ for small x , then a linear equation for the time evolution of the average population fitness emerges as:

$$S = \frac{1}{\beta} (\alpha t + \ln \alpha) \quad (49)$$

The slope is related to the fitness value assigned to allele, f_0 . This linear equation gives a convenient approximation which corresponds to a first-order approach to closed form solutions. The interest in this representation is that standard perturbation techniques can be built into the GA dynamics by including higher order polynomial expansions around this linear relation. In this way, the average fitness can be computed and compared between actual numerical iterations and analytical predictions. .

Secondly however consider the more striking version of (47) valid for large β values (large fitness, f_1) such that the approximation holds that $(S - \beta^{-1}) = S$ and

$$t = S \ln(S + \alpha) \quad (50)$$

For small α values corresponding to a low fitness value attached to the other allele ($f_1 \ll f_0$) compared to the population average, then (50) simplifies finally to:

$$t = S \ln S = \Psi(t, P) \quad (51)$$

The significant result of (51) is $\Psi(t, P)$. This entropy-like function is defined for the probability of finding an overall average population fitness in the particular state, S . The definition is motivated from the recognition that the product of a fitness probability function, S , and its logarithm, $\ln S$, is equivalent to the thermodynamic entropy of information about states of S . By analogy, a physical interpretation for the fitness evolution can be formulated in terms of this entropy as a measure of population diversity. Because of its linear dependence on time, the steady-state reproduction model is capable of maintaining considerable diversity late into a numerical simulation; equivalently, there exists a finite spread in the average fitness of the population (S0 even as particularly fit

individuals are selected across many generations. A more extended discussion of this point will follow in subsequent sections, but suffice for the present purposes to suppose that an entropy-like measure of fitness information arises directly from the steady-state GA's formalism. No *ad hoc* assumptions about its logarithmic form need be imposed. This result bears directly on related issues of putting a physical interpretation to the GA's delicate tradeoff between diversity and selection. It is worth noting the analogy between GA formalism and successful theories of information theory and entropy applied elsewhere in mathematic genetics and biology (see for example Gatlin, 1972).

Non-binary Generalization

For simplicity, this note has treated binary genetic algorithms, although the generalization to non-binary GAs can be developed in parallel by redefining the fitness ratio, $r=f_1/f^*$, where f^* is some composite fitness,

$$f^* = \frac{\sum_{j=1}^i f_j P_j}{\sum_{j=1}^i P_j} \quad (52)$$

The preceding analysis of the binary GA's time complexity can be written using this non-binary fitness ratio to give analogous results to the binary cases.

Uses and Extensions

The analysis applied here has generated characteristic times for convergence which give not only the time complexity (e.g order of magnitude dependence on population size) but

also includes various new measures of the population diversity and average fitness. In this way, the steady-state GA can yield a complementary framework for comparing its behavior to the analytically simpler generational GA. For the less tractable steady-state, this approach supplements the directly iterated solutions of Syswerda (1991) and the integrated solutions of Goldberg and Deb (1991). The closed-form, recursive solutions found here puts the steady-state GA on an equal footing with Ankenbrandt's proofs (1991) given for the generational GA and allows a fair and telling comparison. For select choices of parameters, this outline shows that the generational GA can arise as a subcase of the steady-state GA. Moreover, the same framework supports a derivation which naturally proposes a quantitative measure for the time history of population diversity within the steady-state model. An entropy-like term for the average fitness probabilities falls out of the analysis directly and with minimal assumptions..

It is interesting that the traditional thermodynamic notions of entropy as a logarithmic function of state probabilities can arise directly within a steady-state GA, particularly if the GA's formalism describes a life, death and birth process. The logarithm itself reflects the feature that in multiple state systems, entropy is an additive and not multiplicative variable. Similarly in the GA, average fitness probabilities for the population as a whole add sequentially over generations. It is intriguing to note that this characteristic logarithm appears effortlessly from the steady-state model with few *a priori* assumptions. Additional work should flesh out more explicit relations between the entropy function ($S \ln S$), population diversity and some average values for information mixing between generations. The philosophy of this note has been that a suitable context for such a link between disorder and genetic mixing can arise directly within the steady-state GA.

Conclusions and Summary

This note has compared and contrasted behavior between generational and steady-state versions of the commonly used genetic algorithms. Where possible, a recursive formalism has been derived analytically and in closed form. The time complexity for convergence to a most fit individual has been extracted as a function of population size. Additional dependencies on the fitness ratio can be computed for the specific cases of an exponentially bounded approach to convergence. The methods will find straightforward extension to other (non-exponential) functional approaches to optimality.

An entropic measure of average population fitness follows directly from the steady-state birth-and-death cycle. No *ad hoc* implementations or fits to an assumed logarithmic form need be applied. If this entropy equates physically with other probabilistic measures of population diversity, then one explanation for the GA's robustness may follow from analogous thermodynamic arguments about non-equilibrium disorder. The consequences of this novel behavior are ready for more comprehensive testing against various actual GA simulations. For some test cases, the GA is shown dynamically to vary its entropic measure linearly with time. In this context, the GA may draw on the rich tools applied to the modern study of self-organization and eventually open its black box to reveal more about how natural selection and reproduction can arrive both silently and mysteriously at such marvelous answers.

Acknowledgements

This material is based upon work supported in part by the Santa Fe Summer School on Complex Systems and the Santa Fe Institute. Support was also provided (for D.N.) by the Microgravity Sciences and Application Division, NASA and (for S.B.) by the Austrian fellowship agency, Oesterreicher Auslandsstudentendienst. S.B. gratefully thanks Prof. Murali Sheshardri and Prof. A. Nadarajan for their instruction, unceasing support and interest.

References

Davis, L. (Ed.) (1991). *Handbook of genetic algorithms*. New York: Van Nostrand Reinhold.

Ankenbrandt, C. (1991). An Extension to the Theory of Convergence and a Proof of the Time Complexity of Genetic Algorithms. Rawlins, G.J.E. (Ed.) (1991), *Foundations of genetic algorithms*. San Mateo, CA:Morgan Kaufmann Publishers.

Ankenbrandt, C. (1990). The Time Complexity of Genetic Algorithms and the Theory of Recombination Operators. Doctoral Dissertation, Department of Computer Science, Tulane University, New Orleans, LA.

Gatlin, L.L. (1972). *Information Theory and the Living System*. New York: Columbia University Press.

Grefenstette, J.J. (1986). Optimization of Control Parameters for Genetic Algorithms, IEEE Trans. on Systems, Man and Cybernetics, SMC-16, 1, p. 122-128.

Goldberg, D.E. and Deb, K. (1991) A Comparative Analysis of Secetion Schemes Used in Genetic Algorithms.Rawlins, G.J.E. (Ed.) (1991), *Foundations of genetic algorithms*. San Mateo, CA:Morgan Kaufmann Publishers.

Goldberg, D.E. (1989). *Genetic algorithms in search, optimization and machine learning*.

Reading, MA: Addison Wesley.

Holland, J.H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor, MI: University of Michigan Press.

Schaffer, J.D., Caruana, R.A., Eshelman, L.J., and Das, R. (1989). A Study of Control Parameters Affecting Online Performance of Genetic Algorithms for Function Optimization. *Proceedings of the Third International Conference on Genetic Algorithms*, San Mateo, CA: Morgan Kaufman. pp. 51-60.

Syswerda, G. (1991) A Study of Reproduction in Generational and Steady-State Genetic Algorithms. Rawlins, G.J.E. (Ed.) (1991), *Foundations of genetic algorithms*. San Mateo, CA: Morgan Kaufmann Publishers.

Appendix A: Inductive Proofs by Inspection for Recursion Relations of Binary Genetic Algorithms

(i) Generational case

Start with the recurrence relation given in equation (2) for P_{t+1} . The outline of an inductive proof develops the appropriate terms for P_1 through P_3 , then by inspection identifies the recurrence relation of the general form, $P_t = \phi(r, P_0)$ which is independent of all other P_t except P_0 . In the binary GA, generational reproduction implies that

$$P_{t+1} = \frac{f_1}{S_t} P_t \quad (A1)$$

where S_t defines the average population fitness, $S_t = f_1 P_t + (1 - P_t) f_0$. An iterated recursion relation is complicated by the appearance of P_t in the average population fitness, S_t . Ankenbrandt (1991) along with Goldberg and Deb (1991) have derived the elegantly and simple recursion relation

$$P_t = \frac{r^t P_0}{1 + P_0 (r^t - 1)} \quad (A2)$$

This appendix rederives (A2) in a particularly transparent notation which will prove advantageous for tackling the more difficult recursion relations required to solve for steady-state GAs. For a fitness ratio, $r = f_1/f_0$, two new terms can be introduced: the inverse proportion, $x_t = 1/P_t$, and the fitness difference, $\Delta f = (f_1 - f_0)/f_0$. Equation (A1) can be rewritten as

$$x_{t+1} = \frac{1}{r} [\Delta f + x_t] \quad (A3)$$

This more straightforward form can be iterated recursively and the closed form solution can be identified by inspection:

$$\begin{aligned}
 x_1 &= [(r-1)+x_0]/r = 1 - (1-x_0)/r \\
 x_2 &= [(r-1)+x_1]/r = 1 - (1-x_0)/r^2 \\
 x_3 &= [(r-1)+x_2]/r = 1 - (1-x_0)/r^3 \\
 x_t &= [(r-1)+x_{t-1}]/r = 1 - (1-x_0)/r^t
 \end{aligned} \tag{A4}$$

Substitution for x in (A4) gives agreement with equation (A2) and its time complexities derived previously.

(ii) Steady -state case: constant jump model and the $K=1$ condition for equivalence with generational GA

Start with the recurrence relation given in equation (6) for steady-state P_{t+1} .

$$P_{t+1} = P_t \left[1 - \frac{1}{n} \right] + P_t \left[\frac{f_1}{S_t} \right] \tag{A5}$$

The outline of an inductive proof develops the appropriate terms for P_1 through P_3 , then by inspection identifies the recurrence relation of the general form, $P_t = \phi(r, P_0)$ which is independent of all other P_t except P_0 . Again introducing the notation for the inverse proportion, $x_t = 1/P_t$, and the fitness difference, $\Delta f = (f_1 - f_0)/f_0$. Equation (A5) can be rewritten as

$$x_{t+1} = \frac{K}{r} [\Delta f + x_t] \tag{A6}$$

Appendix B: Note on Iterated Behavior of Average Population Fitness, S_t

Near convergence behavior: How does the fittest proportional fraction, (P_f/P_{t+1}) ,

depend on fitness, f , and population size, n ? Is it worth iterating anymore?

To explore the nature of the average population fitness, this appendix solves for S_t near convergence. The convergence condition given by Ankensbrandt (1991) equates $P_{t+1} - P_t = \gamma = 1/n < 1$. From the difference equation for the fit proportion, P_{t+1} , then the average population fitness can be solved for in a form that *depends only on the population size, n , and fitness ratio, f .*

Start with the recurrence relation given in equation (6) for P_{t+1} :

$$P_{t+1} = P_t \left[1 - \frac{1}{n} \right] + P_t \left[\frac{f_1}{S_t} \right] \quad (B1)$$

and the recursion relation for the average population fitness, $S_t = P_t f_1 + (1 - P_t) f_0$. Near convergence, $P_{t+1} = P_f = 1 - 1/n = (n-1)/n$, and therefore,

$$S_f = \frac{n(n-1)f_1}{2n-1} \quad (B2)$$

The fittest proportion evolves near convergence as

$$P_{t+1} = P_f \left[1 - \frac{1}{n} \right] + P_f \left[\frac{f_1}{S_t} \right] \quad (B3)$$

By substitution for the average population fitness, S , *the fit proportion evolves independent of assigned fitness values as:*

$$P_{t+1} = P_f \left[\frac{n^2}{n(n-1)} \right] \quad (B4)$$

which depends only on population size.

For times $t > t_c$, the time dependence for the development of this fit proportion in a convergent neighborhood can be solved for recursively by inspection as:

$$P_{t+1} = P_f \left\{ 1 - \frac{1}{n} \right\}^{t-t_c} \quad (B5)$$

Therefore near convergence, the most fit proportion of the population varies with population size as $(1-1/n)^{t-t_c}$. It is surprising that despite a changing overall fitness of the population, S_t , the behavior of the most fit individual evolves independent of any individual fitness values, f . In other words, less fit individuals may change their proportion dramatically near convergence (tending to zero frequency) and this behavior may appear within a changing average population fitness, but the dynamics of the fit fraction remains relatively unchanged. This follows directly from the definition of convergence itself. All fit solutions behave independent of fitness values, F , and approach convergence similarly beyond some fit proportion, P_f .

Table I. Steady-state vs Generational Reproduction: A Summary

Feature	Steady-state GA	Generational GA
Replacement step size at each time	few members	entire population
Selection criterion for copying	individual fitness values	individual fitness values
Selection criteria for deletion	random, least fit, exponential ranking, reverse fitness	no deletion
Relation to schema theorem	identical to generational	identical to steady-state
Population diversity	never fully convergent	less diversity
Ideal performance (Syswerda, 1991) (informal comparison)	time compressed	time expanded
New elite availability in subsequent steps	immediate availability	delayed availability
Application fields	classifier systems, Genitor, etc.	optima problems

Figure Captions

Figure 1. Time complexity behavior of generational reproduction. a) comparison between average and worst case convergence during generational reproduction; b) surface plot of worst case behavior as a function of population size, n , and fitness ratio, r .

Figure 2. Time complexity behavior of integrated form of steady-state reproduction. a) comparison between average and worst case convergence during steady-state reproduction; b) surface plot of worst case behavior as a function of population size, n , and fitness ratio, r ; c) comparison between average time complexity of steady-state vs. generational reproduction.

Figure 3. Time evolution of most fit proportion $x_t=1/P_t$ using steady-state reproduction with constant intergenerational jump size, K (e.g. compression factor). $K=1$ corresponds to generational reproduction. a) comparison of steady-state ($K>1$) to generational ($K=1$). b) surface plot of most fit proportion as a function of time and compression factor, K .

Figure 4. Time complexity of most fit proportion using steady-state reproduction with constant intergenerational jump size, K (e.g. compression factor). $K=1$ corresponds to generational reproduction. a) comparison of average and worst time complexity for steady-state ($K>1$) b) surface plot of time complexity as a function of population size n and compression factor, K .

Figure 5. Half-life behavior for average population fitness as a function of population size, n , and fitness ratio, r .

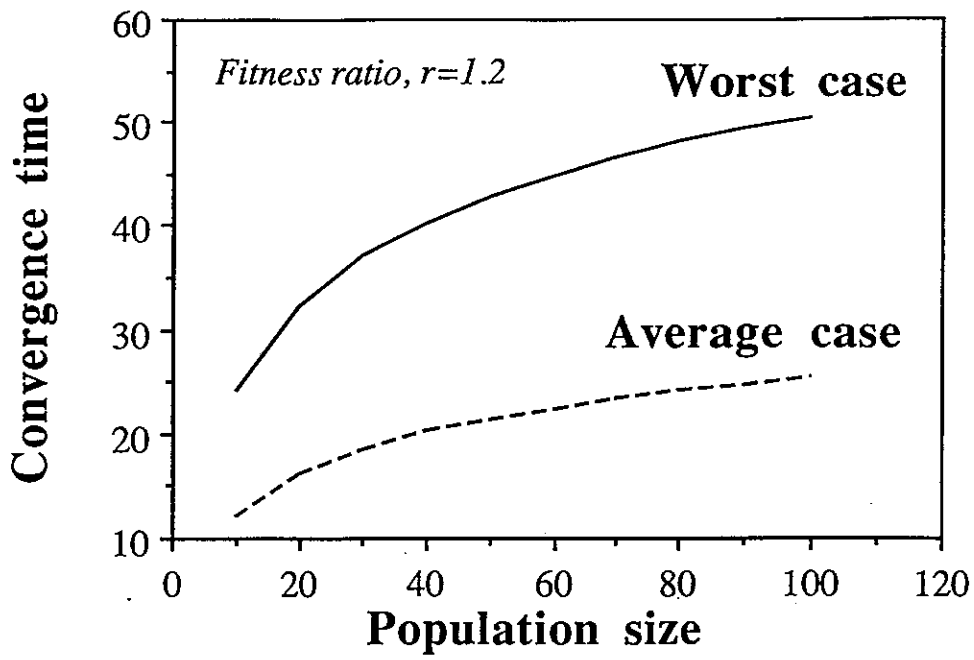


Figure 1a

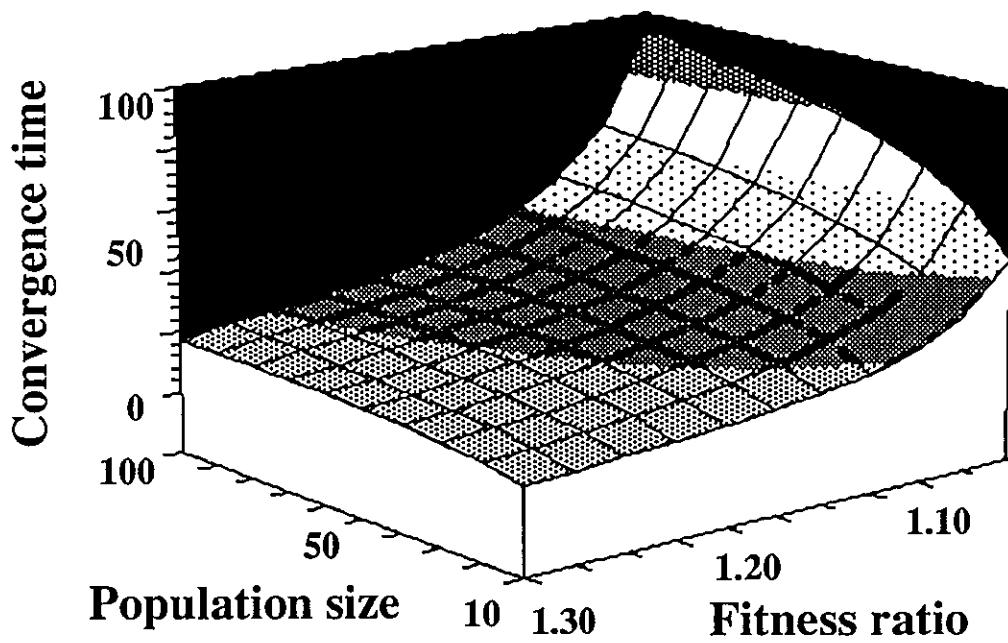


Figure 1b

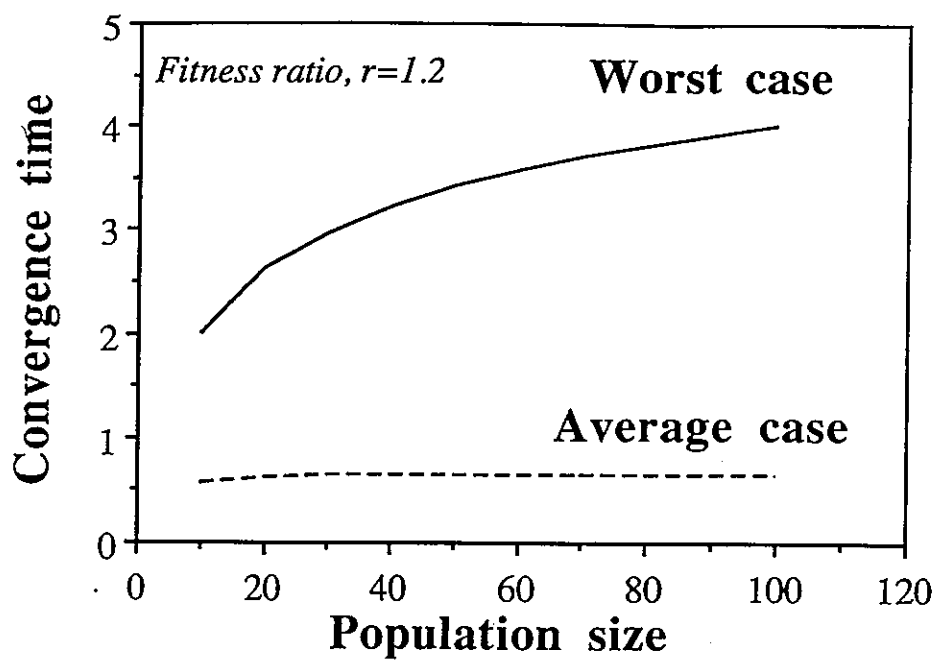


Figure 2a

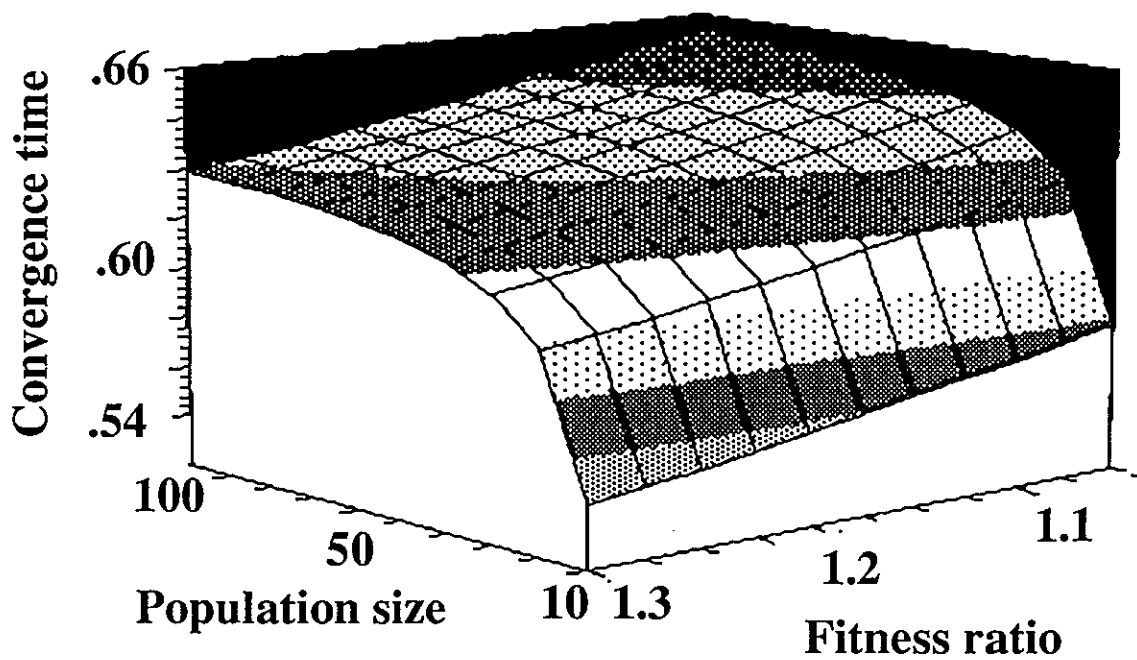


Figure 2b

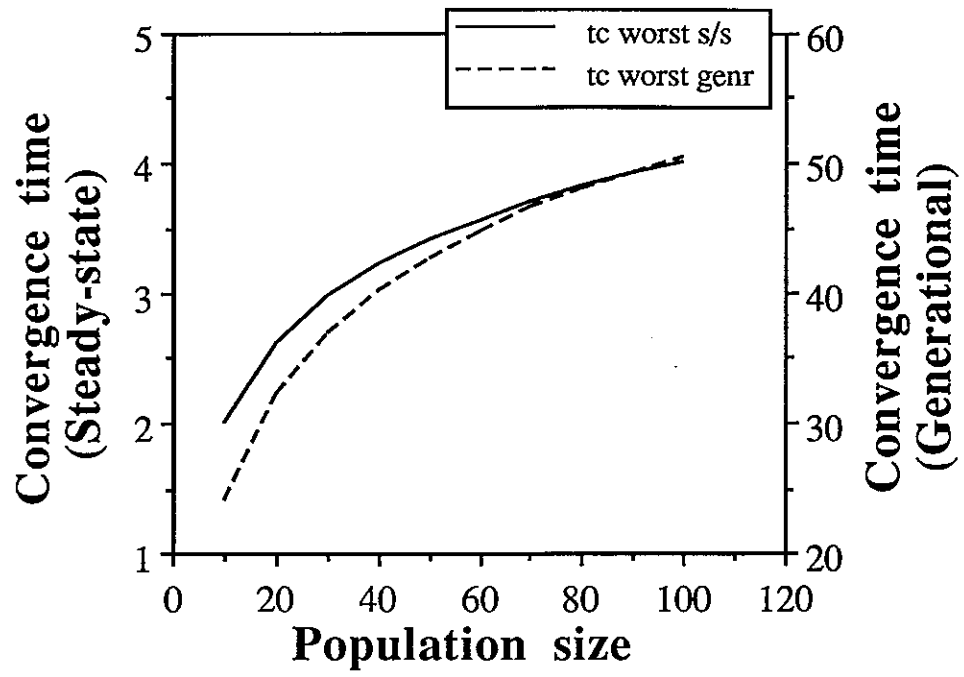


Figure 2c

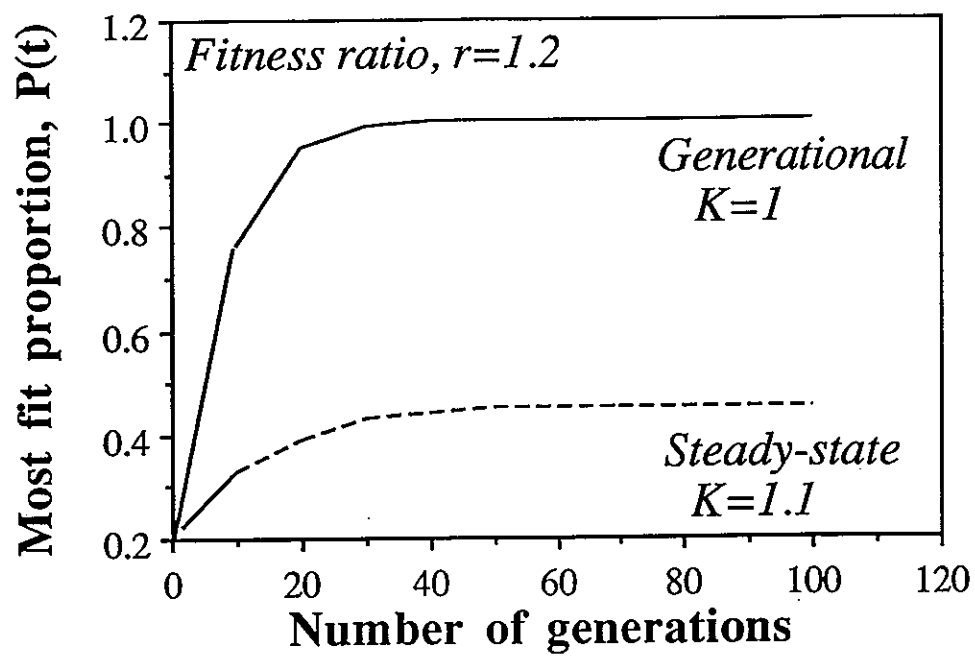


Figure 3a

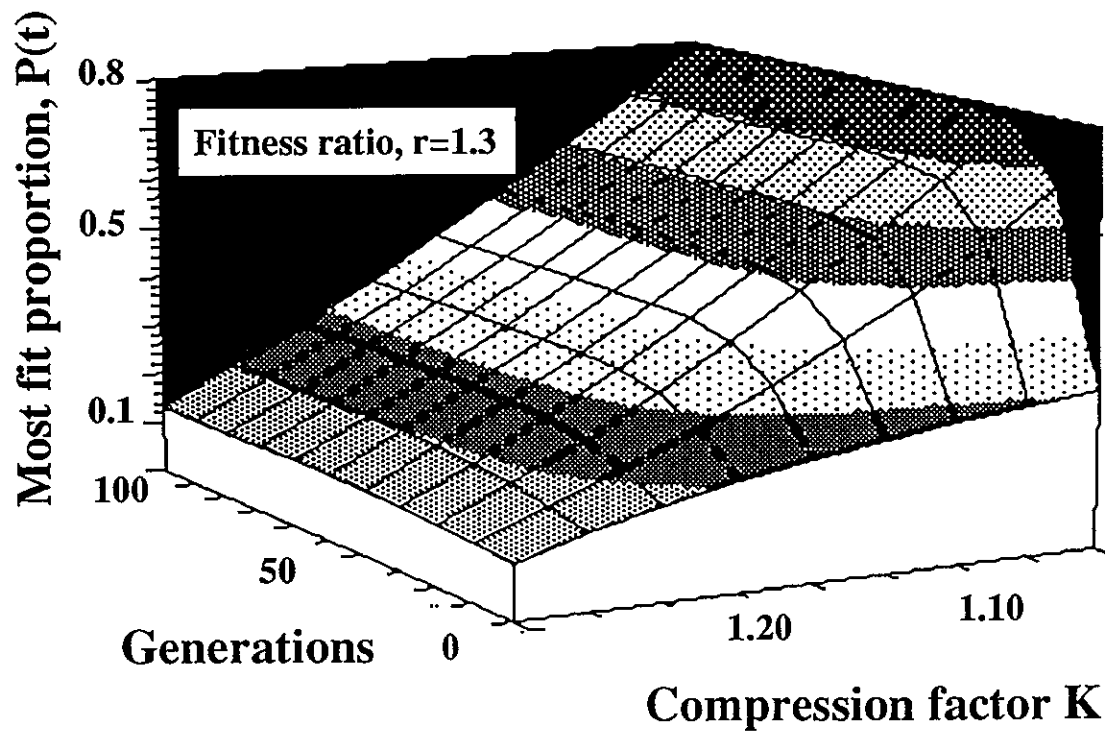


Figure 3b

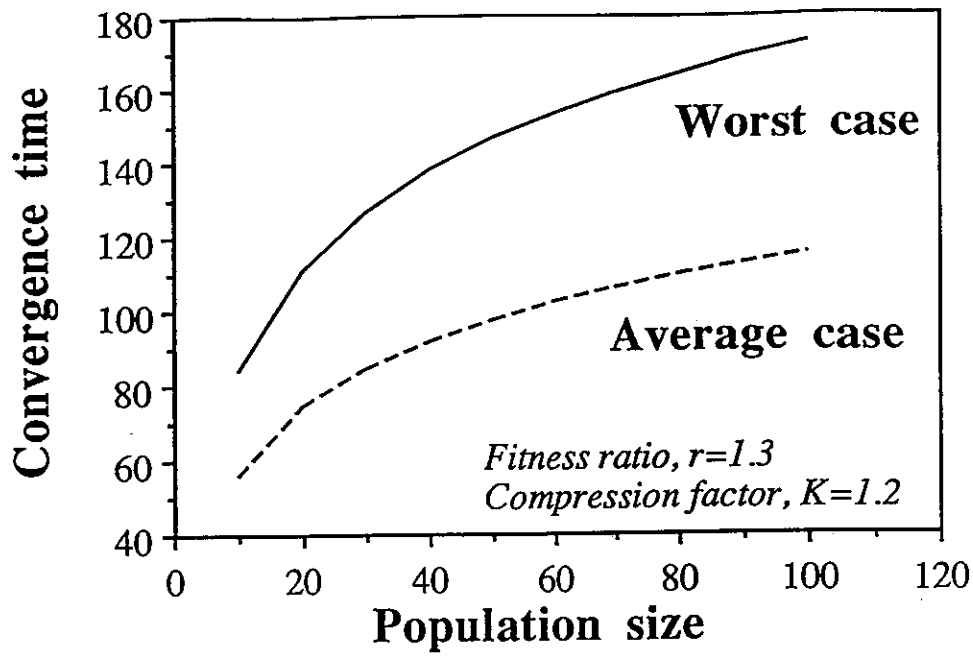


Figure 4a

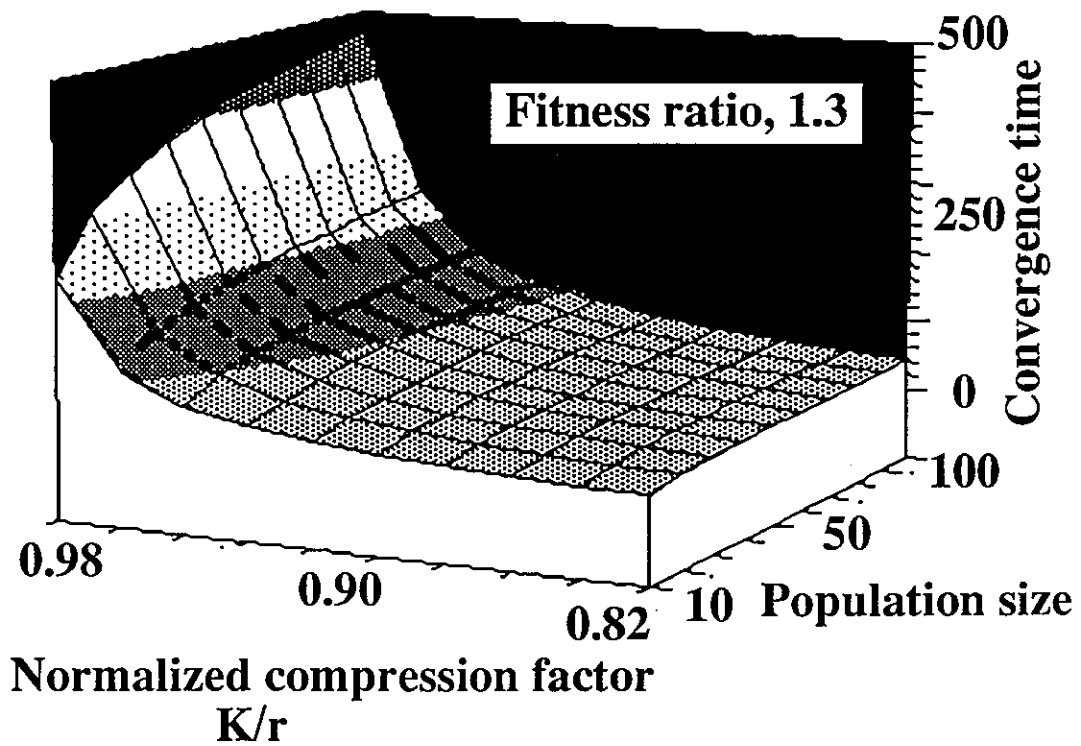


Figure 4b

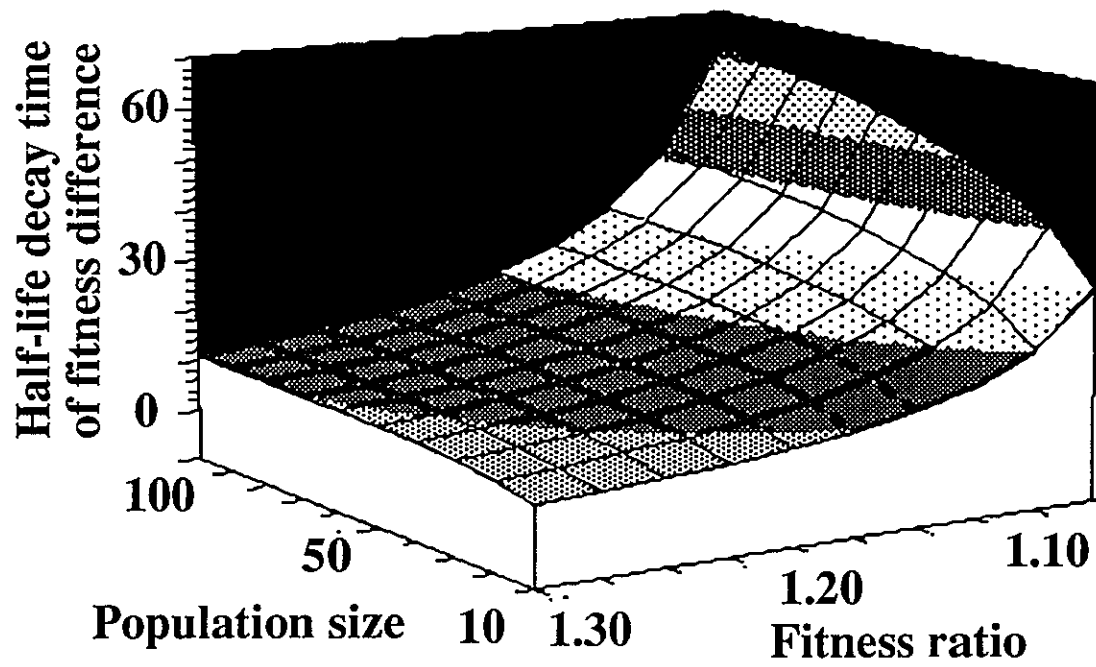


Figure 5

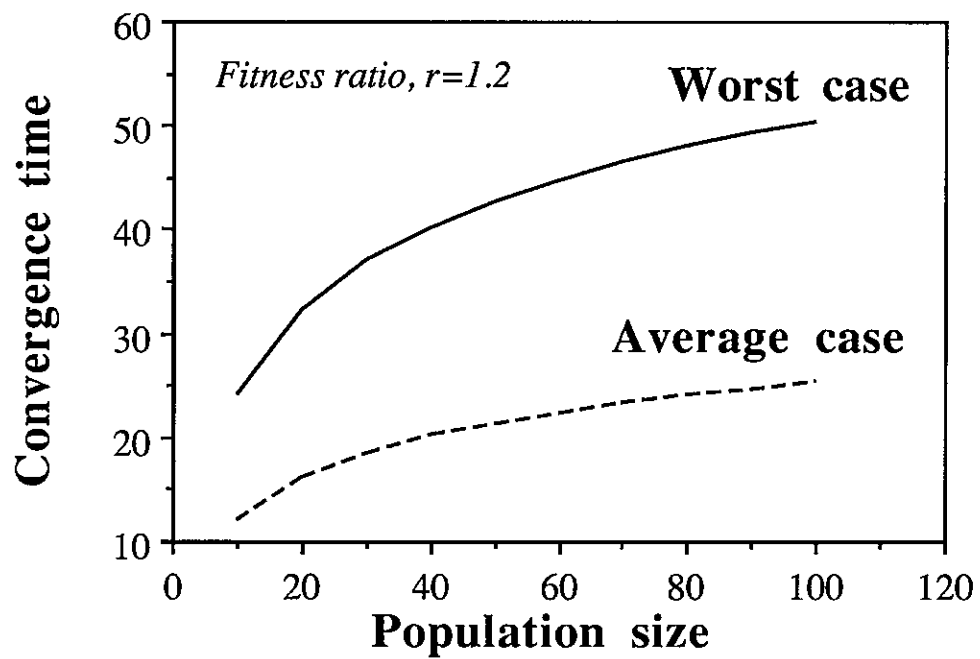


Figure 1a

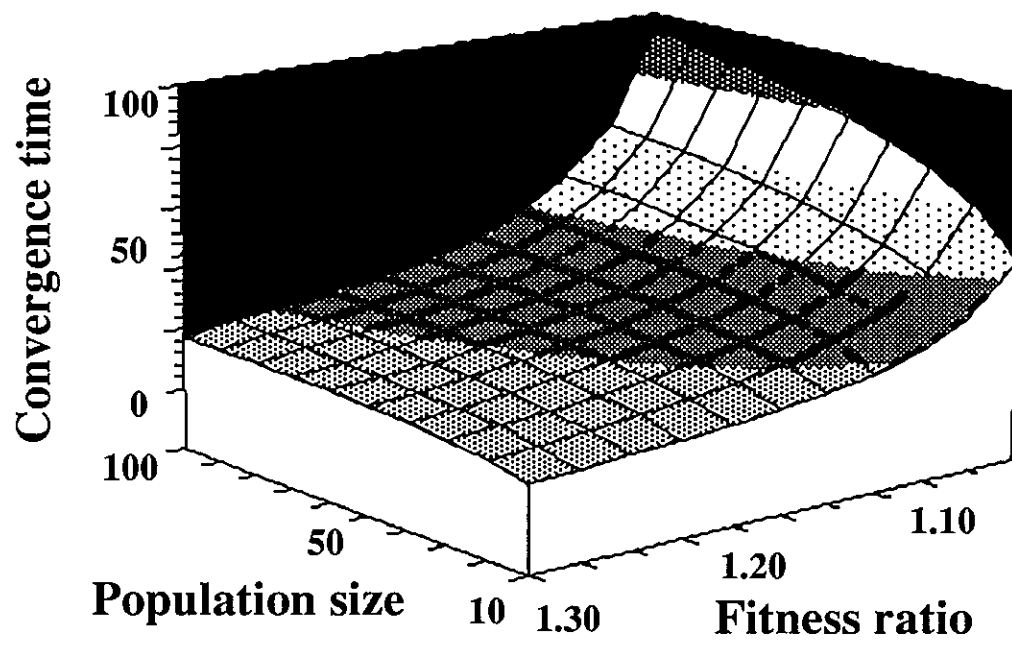


Figure 1b

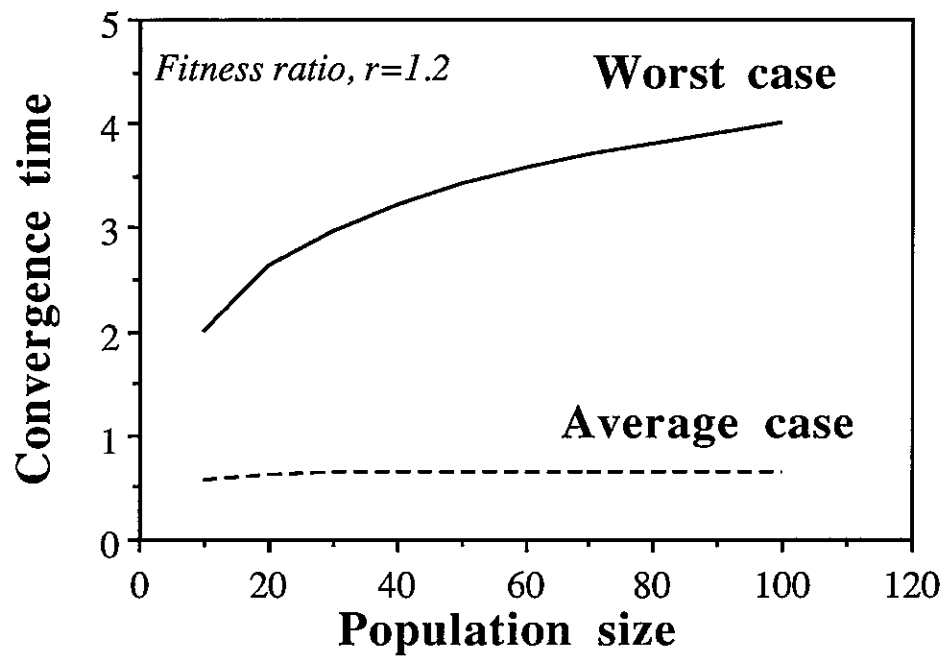


Figure 2a

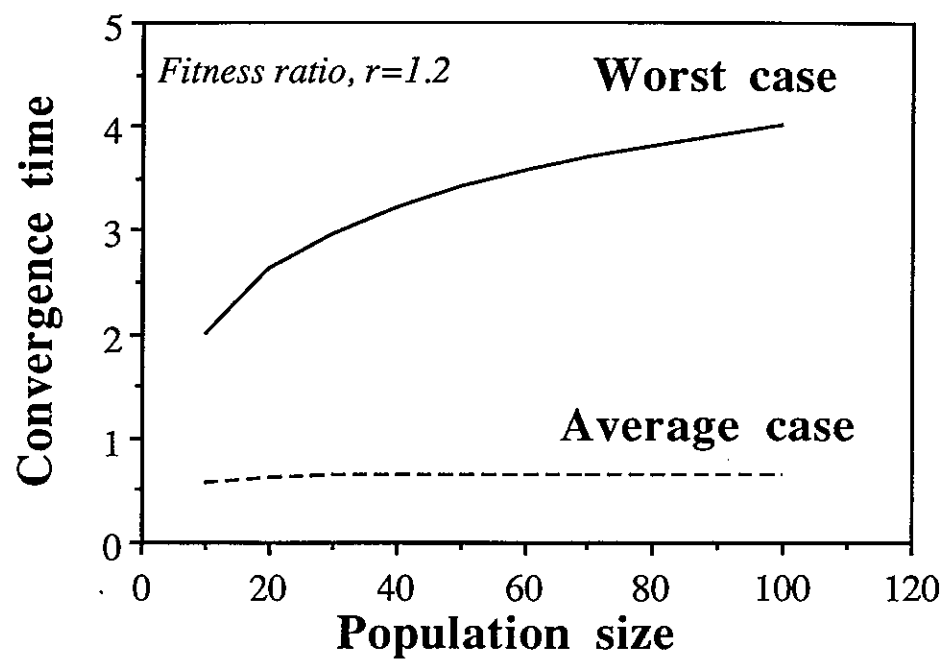


Figure 2a

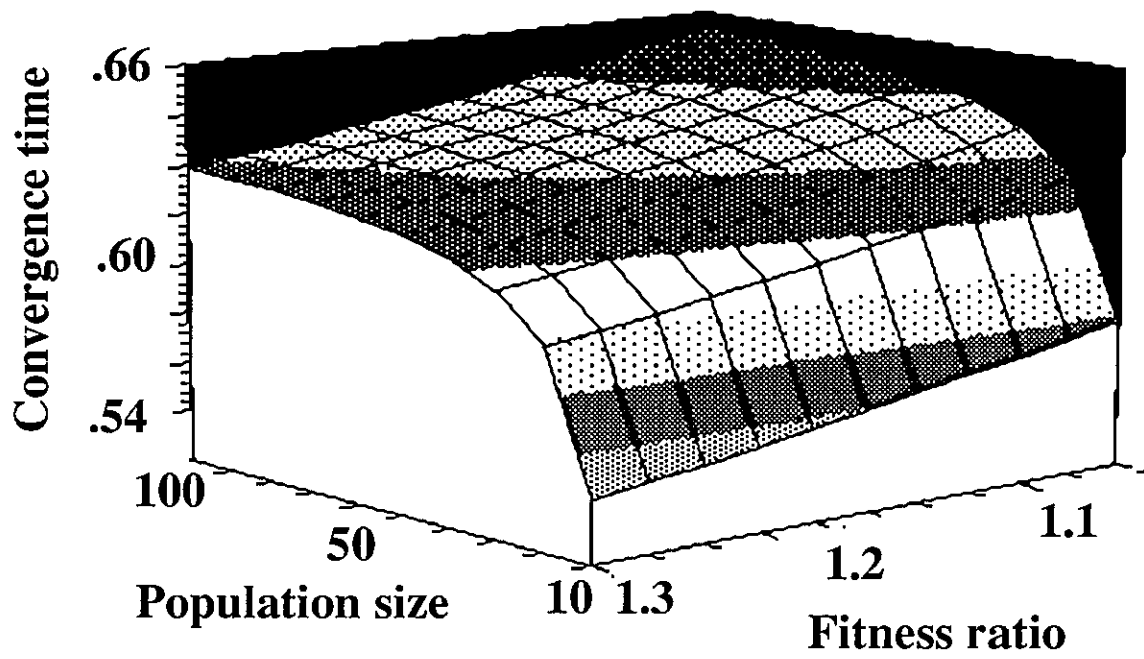


Figure 2b

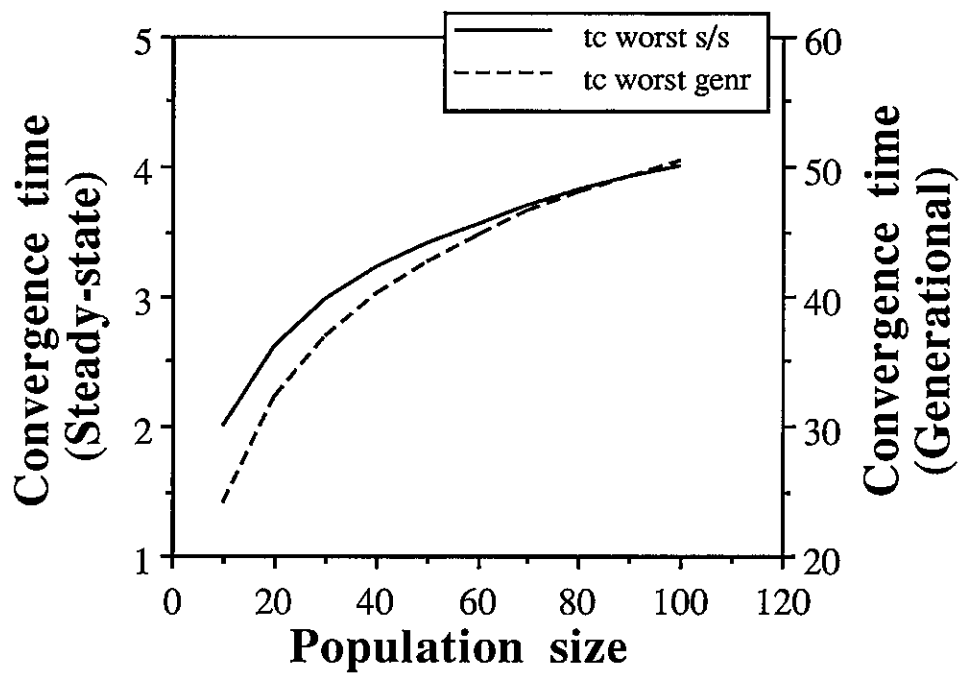


Figure 2c

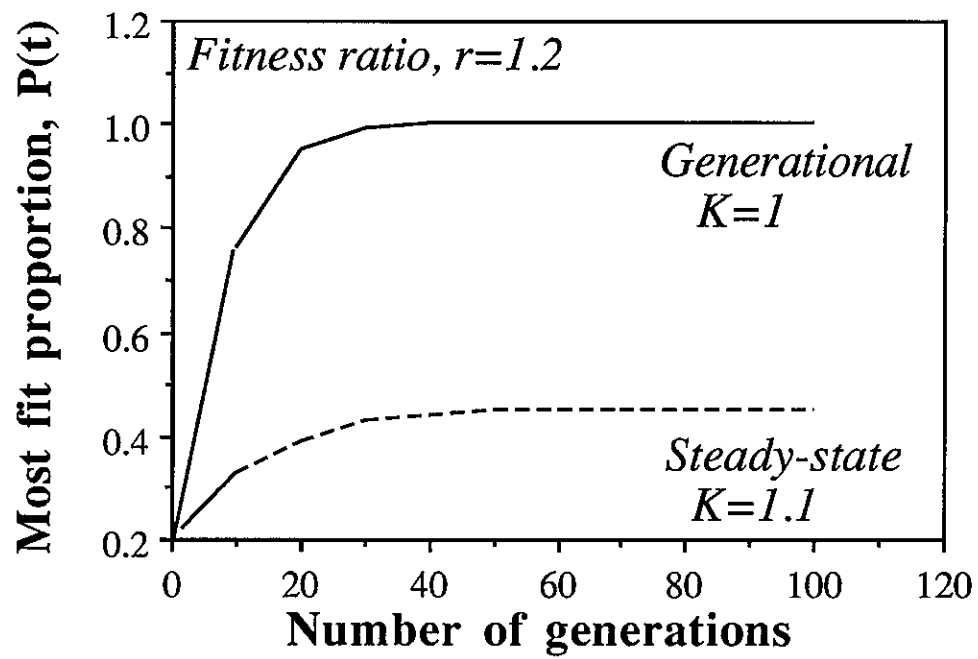


Figure 3a

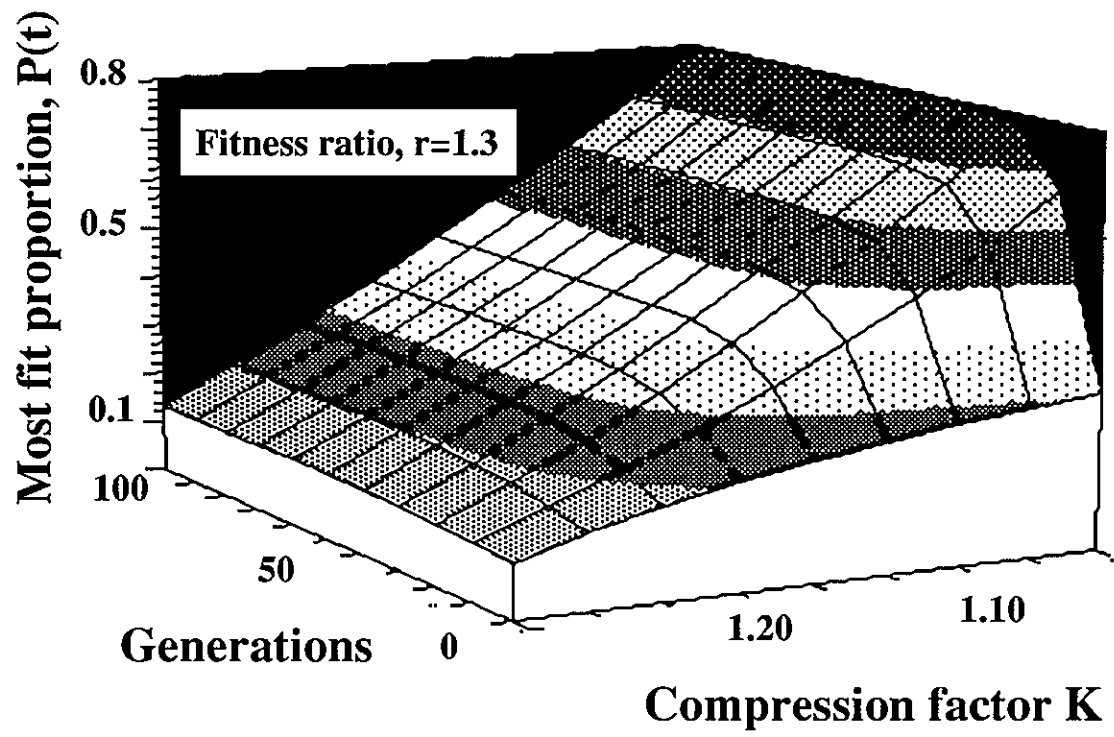


Figure 3b

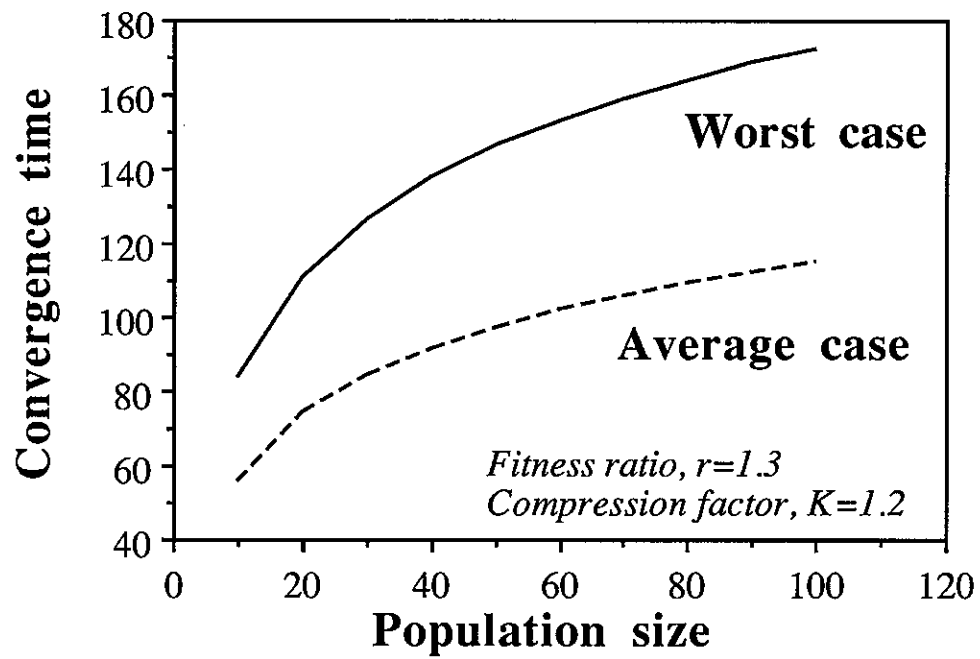


Figure 4a

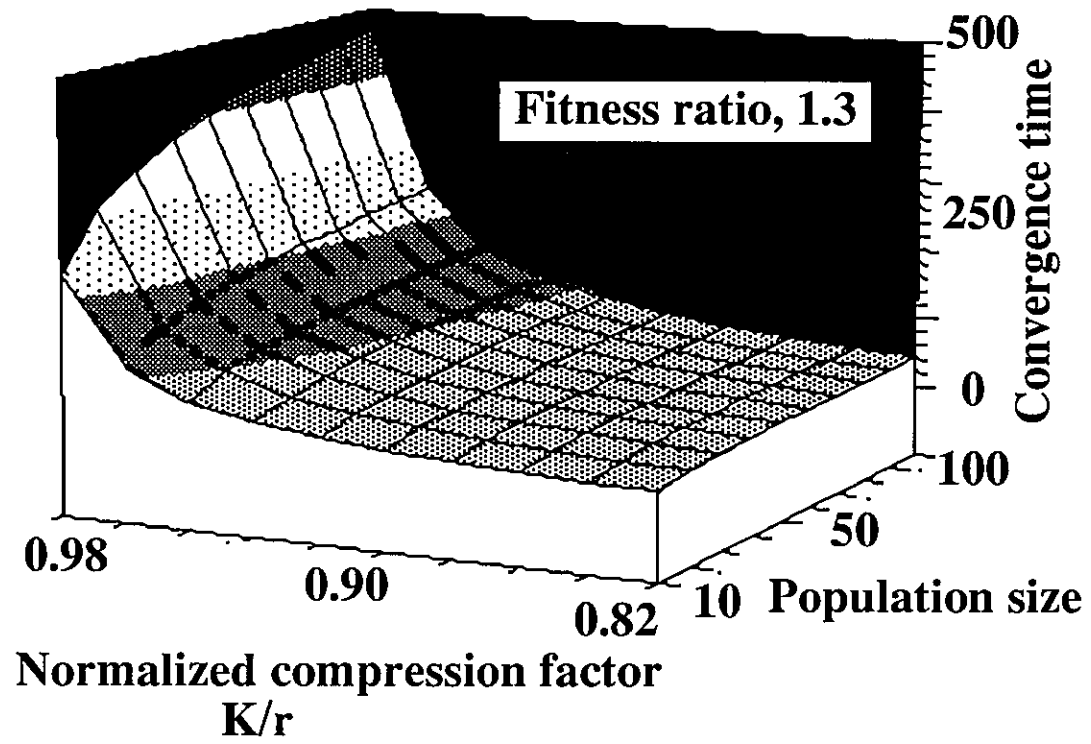


Figure 4b

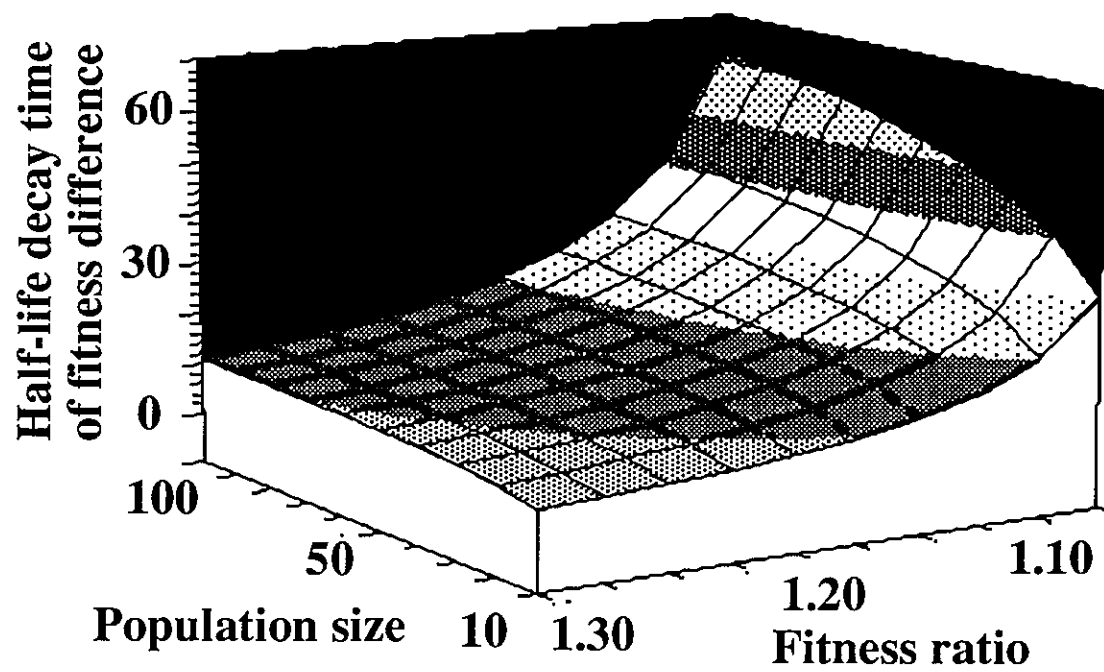


Figure 5