This Technical Report describes the presentation given at
The IEEE Workshop on Genetic Algorithms, Simulated Annealing & Neural Nets
held in Glasgow, Scotland, May, 1990.

FORMALIZING GENETIC ALGORITHMS†

1 Introduction

Designed to search irregular, poorly understood spaces, GAs are general purpose algorithms developed by Holland (1975). Inspired by the example of population genetics, genetic search proceeds over a number of generations. The criteria of "survival of the fittest" provides evolutionary pressure for populations to develop increasingly fit individuals. Although there are many variants, the basic mechanism of a GA consists of:

1. Evaluation of individual fitness and formation of a gene pool.
2. Mutation and Recombination.

Individuals resulting from these operations form the members of the next generation, and the process is iterated until the system ceases to improve.

Fixed length binary strings are typically the members of the population. They are selected (with replacement) for the gene pool with probability proportional to their relative fitness, which is determined by the objective function. There, they are mutated and recombined by crossover. Mutation corresponds to flipping the bits of an individual with some small probability (the mutation rate). The simplest implementation of crossover selects two "parents" from the pool and, after choosing the same random position within each string, exchanges their tails. Crossover is typically performed with some probability (the crossover rate), and parents are otherwise cloned. This recombination cycle repeats, contributing one of the resulting "offspring" each time until the next generation is full.

While this description may suffice for successful application of the genetic paradigm, it is not particularly amenable to mathematical analysis. Our objective is to formalize a simple GA. We model GAs geometrically in sections 2 and 3 as dynamical systems in a high dimensional Euclidean space. In section 4, we develop basic structure of the model and establish preliminary results which demonstrate feasibility. The geometric object which simple genetic search explores is identified in section 5 and we consider simple examples illustrating some of its properties. In section 6 we indicate directions for future research.

†This research was supported by the National Science Foundation (IRI-8917545).

2 Preliminary Considerations

Let Ω be the set of all length ℓ binary strings, and let $N = 2^\ell$. Thinking of elements of Ω as binary numbers, we identify Ω with the interval of integers $[0, N-1]$. We also regard Ω as the product group

$$\mathcal{Z}_2 \times \ldots \times \mathcal{Z}_2$$

where $\mathcal{Z}_2$ denotes the additive group of integers modulo 2. The group operation $\oplus$ acts on integers in $[0, N-1]$ via these identifications, and we use $\otimes$ to represent componentwise multiplication.[1]

The t th generation of the genetic algorithm is modeled by a vector $s^t \in \mathcal{R}^N$, where the i th component of s^t is the probability that i is selected for the gene pool. Populations excluding members of Ω are modeled by vectors s^t having corresponding coordinates zero. Let $p^t \in \mathcal{R}^N$ be a vector with i th component equal to the proportion of i in the t th generation, and let $r_{i,j}(k)$ be the probability that k results from the reproductive process based on parents i and j.

Lemma 1 Let $\mathcal{E}$ denote expectation, then

$$\mathcal{E}\, p_k^{t+1} = \sum_{i,j} s_i^t\, s_j^t\, r_{i,j}(k)$$

Proof: The expected proportion of k in the next generation is computed by summing over all possible ways of producing k. If k results from reproduction based on parents i and j, then i is selected for reproduction with probability s_i^t, j is selected for reproduction with probability s_j^t, and k is the result of reproduction with probability $r_{i,j}(k)$. $\square$

Taking the limit as population size $\to \infty$, the law of large numbers gives $p_k^{t+1} \to \mathcal{E}\, p_k^{t+1}$. Thus Lemma 1 can be used to determine how the probability vector s^t changes from one generation to the next in a GA with infinite population. But first, we note an important property of $r_{i,j}(k)$:

Lemma 2 If reproduction is a combination of mutation and crossover, then

$$r_{i,j}(k \oplus l) = r_{i \oplus k, j \oplus k}(l)$$

Proof: Let $C(i,j)$ represent the possible results of crossing i and j. Note that $k \oplus l \in C(i,j)$ if and only if $k \in C(i \oplus l, j \oplus l)$. Let $X(i)$ represent the result of mutating i, for some fixed mutation. Note that $k \oplus l = X(i)$ if and only if $k = X(i \oplus l)$. Since reproduction is a combination of operations which commute with group translation, the result follows. $\square$

Let F be the nonnegative diagonal matrix with i th entry $f(i)$, where f is the objective function, and let M be the matrix with i,j th entry $m_{i,j} = r_{i,j}(0)$. Define permutations σ_j on $\mathcal{R}^N$ by

$$\sigma_j < y_0, \ldots, y_{N-1} >^T = < y_{j \oplus 0}, \ldots, y_{j \oplus (N-1)} >^T$$

[1] Hence, $\oplus$ is *exclusive-or* on integers and $\otimes$ is *logical-and*.

where vectors are regarded as columns, and T denotes transpose. Define operators $\mathcal{M}$, $\mathcal{F}$, and $\mathcal{G}$ on $\mathcal{R}^N$ by

$$\mathcal{M}(s) = <(\sigma_0\, s)^T M \sigma_0\, s, \ldots, (\sigma_{N-1}\, s)^T M \sigma_{N-1}\, s>^T$$

$$\mathcal{F}(s) = F s$$

$$\mathcal{G} = \mathcal{F} \circ \mathcal{M}$$

Theorem 1 Let $\sim$ represent the equivalence relation on $\mathcal{R}^N$ defined by $x \sim y$ if and only if $\exists \lambda > 0 \,.\, x = \lambda y$, then

$$\mathcal{E}\, s^{t+1} \sim \mathcal{G}(s^t)$$

Proof: We have

$$\begin{aligned}
\mathcal{E}\, p_k^{t+1} &= \sum_{i,j} s_i^t\, s_j^t\, r_{ij}(k) \\
&= \sum_{i,j} s_i^t\, s_j^t\, r_{i \oplus k, j \oplus k}(0) \\
&= \sum_{i \oplus k, j \oplus k} s_{i \oplus k}^t\, s_{j \oplus k}^t\, r_{ij}(0) \\
&= (\sigma_k\, s)^T M \sigma_k\, s
\end{aligned}$$

Since $s^{t+1} \sim F p^{t+1}$ (the probability of selection is proportional to relative fitness), the result follows. □

The (expected) behavior of a simple GA is therefore determined by two matrices; fitness information appropriate for selection is contained in F, while M encodes mixing information appropriate for recombination. Moreover, the relation

$$s^{t+1} \sim \mathcal{G}(s^t)$$

is an exact representation of the limiting behavior as population size $\to \infty$.

3 Formalization

The matrix M has many special properties, the most obvious of which are:

Theorem 2 The matrix M is nonnegative, symmetric, and for all i,j satisfies

$$1 = \sum_k m_{i \oplus k,\, j \oplus k}$$

Proof: M is nonnegative since its entries are probabilities, and is symmetric since $m_{i,j} = r_{i,j}(0)$ and the results of reproduction depend on the unordered set of parents.

Moreover,

$$1 = \sum_k r_{i,j}(k) = \sum_k r_{i\oplus k,\, j\oplus k}(0) = \sum_k m_{i\oplus k,\, j\oplus k}$$

□

The previous considerations lead us to the following formalization:

Definition 1 Given a nonnegative, injective objective function f defined on Ω, a *simple genetic search* is the pair of operators $(\mathcal{F}, \mathcal{M})$ except that M may be any matrix satisfying Theorem 2. An initial population is modeled by a point $s^0 \,\epsilon\, \mathcal{R}^N$, and the transition between generations is defined by $s^{t+1} \sim \mathcal{G}(s^t)$.

This formalization generalizes the mixing induced by mutation and crossover, and regards GAs with finite populations as approximations to the ideal of simple genetic search. Note that simple genetic search can be implemented without resorting to any population since $\mathcal{F}$ is simply matrix multiplication and $\mathcal{M}$ is represented by a collection of quadratic forms.

One natural geometric interpretation of simple genetic search is to regard $\mathcal{F}$ and $\mathcal{M}$ as maps from $\mathcal{S}$, the nonnegative points[2] of the unit sphere in $\mathcal{R}^N$, to $\mathcal{S}$ (since apart from the origin, each equivalence class of $\sim$ has a unique member of norm 1). An initial population then corresponds to a point on $\mathcal{S}$, the progression from one generation to the next is given by the iterations of $\mathcal{G}$, and convergence (of the GA) corresponds to a fixed point (of $\mathcal{G}$).

4 Basic Properties

Regarding $\mathcal{F}$ as a map on $\mathcal{S}$, its fixed points correspond to the eigenvectors of F, which are the unit basis vectors $u_0, \ldots, u_{N-1}$.[3]

Theorem 3 The basin of attraction of the fixed point u_j (of $\mathcal{F}$) is given by the intersection of $\mathcal{S}$ with the (solid) ellipsoid

$$\sum_i \left(s_i \frac{f(i)}{f(j)} \right)^2 < 1$$

Proof: Let $s \,\epsilon\, \mathcal{S}$. The cosine of the angle between s and u_j is given by the dot product $s \cdot u_j$, and the cosine of the angle between $\mathcal{F}(s)$ and u_j is given by $Fs/\|Fs\| \cdot u_j$. Hence, the angle between s and u_j is decreased by $\mathcal{F}$ when

$$s_j < \frac{s_j f(j)}{\|Fs\|}$$

which is equivalent to the statement of the theorem. □

Only the fixed points corresponding to the maximal value of the objective function f are in the interior of their basins of attraction. Hence *all other fixed points are unstable.* This follows from

[2] i.e., points with nonnegative coordinates.

[3] Here u_j differs from the zero vector only in that the j th component of u_j is 1.

the observation that when $f(j)$ is maximal, *no* point of $\mathcal{S}$ moves away from u_j since

$$\sum_i \left(s_i \frac{f(i)}{f(j)} \right)^2 \;\leq\; \sum_i s_i^2 \;=1$$

Regarding $\mathcal{M}$ as a map on $\mathcal{S}$, the set $\mathcal{M}_{fixed}$ of fixed points of $\mathcal{M}$ is more difficult to analyze; it can range from all of $\mathcal{S}$ to the single point $\upsilon \;=< \sqrt{\frac{1}{N}}, \ldots, \sqrt{\frac{1}{N}} >$. Moreover, intermediate behavior is possible; matrices corresponding to crossover can have surfaces of fixed points. In order to investigate $\mathcal{M}_{fixed}$ further, we need some properties of the differential $\mathcal{D}_{\mathcal{M}}(x)$ of $\mathcal{M}$ at x. We need to be careful, because the differential is changed by regarding $\mathcal{M}$ as a map from $\mathcal{S}$ to $\mathcal{S}$. We therefore interpret $\mathcal{M}$ strictly (i.e., as was originally defined) in what follows.

Lemma 3 Let the sum of the coordinates of $s \,\epsilon\, \mathcal{S}$ be denoted by $|\,s\,|$.

1. $\mathcal{M}(\alpha x) = \alpha^2 \mathcal{M}(x)$ for all $\alpha \,\epsilon\, \mathcal{R}$.
2. If $\mathcal{M}(x) \sim x$ then $\mathcal{M}(x) = |\,x\,|\, x$.
3. The i, j th component of $\mathcal{D}_{\mathcal{M}}(x)$ is $2 \sum_k m_{i \oplus j, k}\, x_{i \oplus k}$.
4. A maximal eigenvalue of $\mathcal{D}_{\mathcal{M}}(x)$ is $2\,|\,x\,|$.

Proof: Noting that each component of $\mathcal{M}(x)$ is a homogeneous polynomial of degree 2 in the coordinates of x establishes the first claim. Using Theorem 2,

$$|\,\mathcal{M}(x)\,| \;=\; \sum_k \sum_{i,j} m_{k \oplus i, k \oplus j}\, x_i x_j \;=\; \sum_{i,j} x_i x_j \sum_k m_{k \oplus i, k \oplus j} \;=\; |\,x\,|^2$$

Hence $\mathcal{M}(x) = \lambda x \implies |\,\mathcal{M}(x)\,| = \lambda\, |\,x\,| \implies \lambda = |\,x\,|$ which establishes the second claim. The calculation of $\mathcal{D}_{\mathcal{M}}$ follows from taking partial derivatives and using the symmetry of M to simplify the resulting Jacobian. The last claim follows from Perron – Frobenius theory,[4] since Theorem 2 implies that $\mathcal{D}_{\mathcal{M}}(x)$ is a nonnegative matrix with column sums equaling $2\,|\,x\,|$. □

We will need the following discrete analogue from Lyapunov's theory of stability: [5]

Lemma 4 Suppose that x is a fixed point of a map W and that the spectrum of the differential $\mathcal{D}_W(x)$ is contained in the open unit disk. Then x is asymptotically stable.

Note that for all k,

$$I - \upsilon\upsilon^T \;=\; \sigma_k^{-1}(I - \upsilon\upsilon^T)\,\sigma_k$$

[4]See H. Minc, *Nonnegative Matrices*, Wiley-Interscience, 1988.
[5]See G. R. Belitskii and Yu. I. Lyubich, *Matrix Norms and their Applications* (1988).

where the σ_k are regarded as permutation matrices.[6] Next define $M_* = (m^*_{i,j})$ to be the matrix determined by $m^*_{i,j} = m_{i \oplus j, i}$, and observe that

$$\mathcal{D}_{\mathcal{M}}(x) \;=\; 2 \sum_k \sigma_k^{-1} M_* \sigma_k \, x_k$$

Since the column sums of M_* are constant, as is also the case for $\mathcal{D}_{\mathcal{M}}(x)$, it follows from Perron – Frobenius theory that when M_* is positive, υ is the unique eigenvector for both M_*^T and $\mathcal{D}_{\mathcal{M}}(x)^T$. Moreover, since the corresponding eigenvalues are simple and maximal, this discussion leads us to a sufficient condition for a fixed point to be an attracter. Let

$$\Lambda \;=\; \{ x \,\epsilon\, \mathcal{R}^N \mid x \text{ is nonnegative and } |x| = 1 \}$$

Theorem 4 Let $x \,\epsilon\, \mathcal{M}_{fixed}$. If the matrix M_* is positive, then x is asymptotically stable whenever the second largest eigenvalue of M_* is less than $\frac{1}{2}$.

Proof: According to Lemma 4, it suffices to check the spectrum of the differential of $\mathcal{M}$. Since Λ is mapped into itself by $\mathcal{M}$, it suffices to consider the action of $\mathcal{M}$ restricted to Λ. The kernel of the projection $I - \upsilon\upsilon^T$ is normal to Λ, hence the spectral radius in question is $\rho = \rho(\mathcal{D}_{\mathcal{M}}(x)(I - \upsilon\upsilon^T))$. Because a matrix and its adjoint share the same norm and spectrum, the discussion following Lemma 4 shows

$$\begin{aligned} \rho \;&\le\; \|(I - \upsilon\upsilon^T)\, \mathcal{D}_{\mathcal{M}}(x)^T\| \\ &\le\; 2 \sum_k \|(I - \upsilon\upsilon^T)\, \sigma_k^{-1} M_*{}^T \sigma_k\| \, x_k \\ &=\; 2 \sum_k \|\sigma_k^{-1} (I - \upsilon\upsilon^T)\, M_*{}^T \sigma_k\| \, x_k \\ &=\; 2\, |x| \; \|(I - \upsilon\upsilon^T)\, M_*{}^T\| \end{aligned}$$

Since given any matrix, a Euclidean norm can be chosen to make its norm arbitrarily close to its spectral radius, the proof is completed by observing that premultiplication of $M_*{}^T$ by the projection $I - \upsilon\upsilon^T$ sends the maximal eigenvalue of $M_*{}^T$ to zero and otherwise leaves the spectrum alone. □

Although $\mathcal{M}_{fixed}$ can vary drastically, there is a group of symmetries which acts on it.

Theorem 5 For all j, and for every mixing matrix M, $\mathcal{M}(\sigma_j\, x) \;=\; \sigma_j \mathcal{M}(x)$. In particular, we have $\sigma_j \mathcal{M}_{fixed} \;=\; \mathcal{M}_{fixed}$, and $\upsilon \,\epsilon\, \mathcal{M}_{fixed}$.

Proof:

$$\begin{aligned} \sigma_j \mathcal{M}(x) \;&=\; \sigma_j < (\sigma_0\, x)^T M \sigma_0\, x, \ldots, (\sigma_{N-1}\, x)^T M \sigma_{N-1}\, x >^T \\ &=\; < (\sigma_{j+0}\, x)^T M \sigma_{j+0}\, x, \ldots, (\sigma_{j+N-1}\, x)^T M \sigma_{j+N-1}\, x >^T \end{aligned}$$

[6] Recall that $\upsilon = < \sqrt{\frac{1}{N}}, \ldots, \sqrt{\frac{1}{N}} >$, and, interpretating the σ_k as a permutation matrices, $\sigma_k = \sigma_k^{-1} = \sigma_k^T$.

$$
\begin{aligned}
&= \; <(\sigma_0\sigma_j\, x)^T M\sigma_0\sigma_j\, x, \ldots, (\sigma_{N-1}\sigma_j\, x)^T M\sigma_{N-1}\sigma_j\, x>^T \\
&= \; \mathcal{M}(\sigma_j\, x)
\end{aligned}
$$

Since $x = \mathcal{M}(x) \;\Rightarrow\; \sigma_j\, x = \sigma_j\mathcal{M}(x) = \mathcal{M}(\sigma_j\, x)$, it follows that $\sigma_j\mathcal{M}_{fixed} = \mathcal{M}_{fixed}$. Since $\sigma_j\mathcal{M}(\upsilon) = \mathcal{M}(\sigma_j\upsilon) = \mathcal{M}(\upsilon)$, it follows that $\mathcal{M}(\upsilon)$ is fixed by each σ_j and must therefore have equal components (i.e., $\mathcal{M}(\upsilon) \sim \upsilon$). □

Theorem 6 If the objective function f is positive, then a necessary condition for simple genetic search to converge to a population consisting of a single integer is that $m_{0,0} = 1$.

Proof: Since f is positive, the matrix F is nonsingular. Suppose that simple genetic search starting at some initial population s^0 will converge to a population consisting of a single integer, that is

$$\lim_{n\to\infty} \mathcal{G}^n(s^0) = u_j$$

Since $\mathcal{G}$ is the interleaving of $\mathcal{M}$ and $\mathcal{F}$, it follows that u_j must be a fixed point of $\mathcal{M}$: If not, then the continuity of $\mathcal{M}$ implies the existence an arbitrarily small neighborhood C of u_j such that $\mathcal{M}(C)$ is at least some fixed distance δ away from u_j. But the invertibility of F and the fact that u_j is an eigenvector imply that if C is sufficiently small, then $\mathcal{F}^{-1}(C)$ is within δ of u_j. Hence $\mathcal{G}(C)\cap C = \emptyset$ which contradicts convergence. Moreover, a direct calculation shows the k th component of $\mathcal{M}(u_j)$ is the k th component of u_j exactly when

$$m_{j\oplus k, j\oplus k} \;=\; \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise} \end{cases}$$

which, by Theorem 2, is equivalent to the condition that $m_{0,0} = 1$. □

We end this section with an illustrative example. Consider the case of one point crossover with mutation. If χ is the crossover rate and μ is the mutation rate, then a simple calculation shows $m_{i,j}$ is

$$\frac{(1-\mu)^\ell}{2}\left\{\eta^{|i|}\left(1-\chi+\chi\sum_{k=1}^{\ell-1}\eta^{-\Delta_{i,j,k}}\right)+\eta^{|j|}\left(1-\chi+\chi\sum_{k=1}^{\ell-1}\eta^{\Delta_{i,j,k}}\right)\right\}$$

where $\eta = \mu/(1-\mu)$, integers are to be regarded as bit vectors when occurring in $|\cdot|$, where division by zero at $\mu = 0$ and $\mu = 1$ is to be removed by continuity, and where

$$\Delta_{i,j,k} \;= |\,(2^k-1)\otimes i\,| - |\,(2^k-1)\otimes j\,|$$

Several computer runs calculating the spectrum of M_* support the following:

Conjecture 1 If $0 \le \mu \le 0.5$, then

- The second largest eigenvalue of M_* is $\frac{1}{2} - \mu$
- The third largest eigenvalue of M_* is $2\left(1 - \frac{\chi}{\ell-1}\right)\left(\frac{1}{2} - \mu\right)^2$

Applying Theorem 4, we would infer from this conjecture that every fixed point of $\mathcal{M}$ is an attracter when $0 < \mu < 0.5$. When $\mu = 0$, calculations indicate that the elements of $\mathcal{M}_{fixed}$ are not isolated but form a surface, which suggests the condition of Theorem 4 may be necessary and sufficient in this case.

The following conjecture of G. R. Belitskii and Yu. I. Lyubich applies to our example:

Conjecture 2 If $\max_{x \,\epsilon X} \rho(\mathcal{D}_W(x)) < 1$, where X is the compact and convex domain and codomain of W, then the fixed point of W is unique, and the sequence of iterates $W^k(x)$ converges to it for every choice of initial point x.

Applying Theorem 5, we would infer from this conjecture that v is the *unique* fixed point of $\mathcal{M}$ when $0 < \mu < 0.5$, and all of $\mathcal{S}$ is its basin of attraction. [7]

Finally, Theorem 6 confirms our intuition that a GA cannot converge to a population consisting of a single integer when $\mu > 0$. Moreover, population fitness cannot in general be monotonic since when $m_{0,0} \neq 1$ every population consisting of a single integer is unstable.

5 The GA-surface

Assume that the objective function f is nonzero so that $\mathcal{F}$ is invertible. Recall from Section 2 that a point $s^t \,\epsilon\, \Lambda$ (i.e., a probability vector), determines a population p^t by $s^t \sim \mathcal{F}(p^t)$ where s_i^t is regarded as the probability that i is selected for recombination. Moreover, the relationship between the t *th* and $t{+}1$ st generation is given by $p^{t+1} = \mathcal{M}(s^t)$. Because p_i^{t+1} describes the proportion of i in population $t{+}1$, it follows that $|\mathcal{F}(p^{t+1})|$ represents the average population fitness at generation $t{+}1$. For notational simplicity, we denote this population fitness by φ_{t+1}. Note that these remarks establish the relation $\varphi_{t+1} = |\mathcal{G}(s^t)|$.

Definition 2 The *GA-surface* corresponding to a simple genetic algorithm is the set

$$\{s^0 \prod_{i=1}^{\infty} \varphi_i^{-2^{-i}} \;|\; s^0 \,\epsilon\, \Lambda\}$$

The next theorem identifies the GA-surface as the object explored by simple genetic search.

Theorem 7 The GA-surface is differentiable and is mapped by $\mathcal{G}$ into itself.

Proof: Since φ_t is continuous as a function of the initial population s^0, and since Λ is compact, it follows that the infinite product defining the GA-surface converges uniformly. Moreover, the surface is differentiable since $\mathcal{G}$ is polynomial in every coordinate,

$$s^t = \frac{\mathcal{G}(s^{t-1})}{|\mathcal{G}(s^{t-1})|}$$

[7]Regarding $\mathcal{M}$ as a map on $\mathcal{S}$. Conjecture 2 would actually be applied with $X = \Lambda$ to obtain a unique fixed point on the simplex Λ.

and the denominator is uniformly bounded away from 0. From Lemma 3 we know that $\mathcal{G}(\alpha\, x) \;=\; \alpha^2 \mathcal{G}(x)$, hence

$$\mathcal{G}(s^0 \prod_{i=1}^{\infty} \varphi_i^{-2^{-i}}) \;=\; \mathcal{G}(s^0) \prod_{i=1}^{\infty} \varphi_i^{-2^{1-i}} \;=\; s^1 \varphi_1 \prod_{i=1}^{\infty} \varphi_i^{-2^{1-i}} \;=\; s^1 \prod_{i=1}^{\infty} \varphi_{i+1}^{-2^{-i}}$$

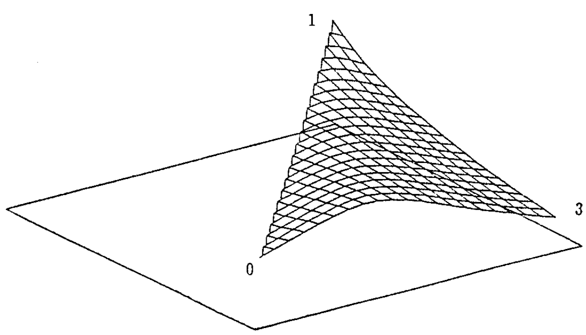
This finishes the proof, since the last expression is the point on the GA-surface corresponding to $s^1 \,\epsilon\, \Lambda$. □

Although population fitness cannot in general be monotonic, decreases are typically slight and transient. If we consider the L_1 norm[8] of the trajectory of successive generations through the GA-surface, we discover that it can be monotonic even when population fitness is not. Let $r_i = \varphi_i / \varphi_{i+1}$ be the ratio of population fitness for successive generations. The ratio of L_1 norms for the points on the GA-surface corresponding to s^1 and s^0 is

$$\prod_{i=1}^{\infty} \left(\frac{\varphi_i}{\varphi_{i+1}} \right)^{2^{-i}} \;=\; \sqrt{r_1 \sqrt{r_2 \sqrt{r_3 \sqrt{\cdots}}}}$$

Simple genetic search moves downhill with respect to the L_1 norm exactly when this nested sequence of square roots is less than 1. Consequently, population fitness may decrease between generations $(r_1, \ldots > 1)$ but if the decrease is slight and transient $(r_k, \ldots < 1)$ then the L_1 norm of the trajectory through the GA-surface can be monotonic. The GA-surface exaggerates the natural tendency of population fitness to be monotonic. We believe a simple GA approximates (with mild restrictions) hill climbing (*downhill*) on the GA-surface.

To illustrate, consider the minimal deceptive problem (Goldberg 1987). Let crossover and mutation be $\chi \;=\; 0.8$ and $\mu \;=\; 0.01$, and consider the objective function $f(0) \;=\; 4,\;\; f(1) \;=\; 3,$ $f(2) = 1,\; f(3) = 5$. This is a type-II problem analyzed by Goldberg.[9] The GA-surface for f is a curved three dimensional space in four dimensions. Intersecting with the coordinate hyperplane $u_2 = 0$ (the low fitness of 2 moves genetic search into this area), we obtain the following surface:

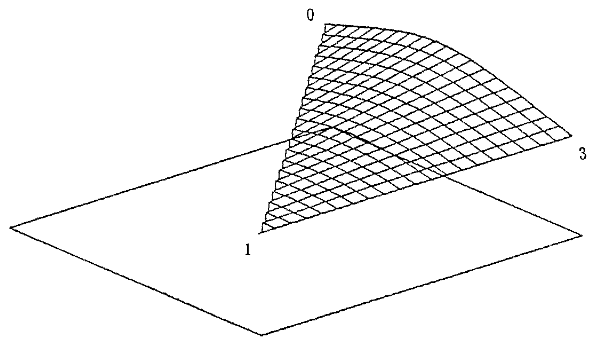


Each corner of this surface corresponds to a population consisting of the integer which labels it. The basin of attraction for each corner is approximately that region of the surface for which all downhill motion leads to it. Hence 0 and 3 are attracters, and the geometry of the GA-surface

[8] Note that $\|x\|_1 = |\, x\,|$ since all coordinates of points on the GA-surface are nonnegative.
[9] He assumed an infinite population (as in our model) but did not consider the effects of mutation.

makes the trajectory of most populations apparent.

If we consider the objective function $f(0) = 4,\ f(1) = 4.1,\ f(2) = .1,\ f(3) = 4.11$, we have a type-I problem. It is because the effects of mutation were not considered that Goldberg's results indicate GA convergence to the optimal. By intersecting the GA-surface corresponding to f with the coordinate hyperplane $u_2 = 0$ (as before, the low fitness of 2 moves genetic search into this area) we obtain:



The GA-surface is tilted slightly towards the forward corner indicating that 1 is the only attracter. Hence the GA should converge to a population near this suboptimal corner. In fact, the fixed point of $\mathcal{G}$ is $< .04, .77, .00, .18 >$. It is interesting to note that with $\mu = 0$, the GA-surface is tilted slightly towards the right corner, confirming Goldberg's result that without mutation, a type-I problem is not hard.

6 Future research

The mathematical elegance of the definitions, proofs, and conjectures associated with our model is indicative of fertility; our framework supports several promising research areas:

- integration of the actions of the transformations $\mathcal{F}$ and $\mathcal{M}$ into the behavior of $\mathcal{G}$,
- further development of the interaction between the geometry of the GA-surface and the trajectory of populations (under $\mathcal{G}$),
- incorporation of schemata analysis by developing its geometric analogue.

7 References

Belitskii, G. R. and Lyubich, Y. I. (1988), *Matrix Norms and their Applications*, Operator Theory: Advances and Applications Vol. 36, Birkhäuser.

Goldberg, D. E. (1987). Simple genetic algorithms and the minimal, deceptive problem. In L. Davis (ed.), *Genetic algorithms and simulated annealing* (pp 74 –88). London: Pitman.

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

H. Minc (1988), *Nonnegative Matrices.* Wiley-Interscience.