

Population fixed-points for functions of unitation

Jonathan E. Rowe
 Artificial Intelligence Group
 Dept. Computer & Information Science
 De Montfort University
 Milton Keynes MK7 6HP
 Great Britain
jrowe@dmu.ac.uk

Abstract

The dynamical systems model for the simple genetic algorithm due to Vose [12] can be simplified to the case of zero crossover, and to fitness functions that divide the search space into relatively few equivalence classes. This produces a low-dimensional system for which the fixed-point can be calculated; it is the leading eigenvector of the system. This technique, applied elsewhere [11] to Royal Road functions, is adapted in this paper to apply to functions of unitation. Infinite population fixed-points are calculated for some simple examples, including trap functions that have previously been analysed in terms of deception [2]. The effects of eigenvectors outside the population space are explored, and finite population behaviour is examined in this way. Two surprising theorems are demonstrated for infinite populations: that every population distribution is the fixed-point for some fitness function; and that fitness functions exist for which evolution goes “backwards” towards the global minimum. Theoretical results are backed up with experiments using real, finite population genetic algorithms.

1 Introduction

The dynamical systems model for the simple genetic algorithm (SGA) suggested earlier by Michael Vose [12] represents populations as vectors $p = (p_0, \dots, p_{s-1})$, where s is the size of the search space (that is, the number of different chromosomes). In this vector, p_j is the proportion of the population taken up by the chromosome corresponding to j . The set of

vectors:

$$\Lambda = \left\{ (x_0, \dots, x_{s-1}) : \sum_{j=0}^{s-1} x_j = 1, x_j \geq 0 \right\}$$

is known as the *simplex*. Clearly, all real populations correspond to points within the simplex. However, it is theoretically interesting to also consider vectors outside the simplex (with some $x_j < 0$), but whose components still sum to 1. In particular, it will be seen that fixed-points outside the simplex may affect the behaviour inside the simplex.

The action of the SGA on a population p is given by the heuristic operator

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

(where \circ denotes operator composition). Each component $(\mathcal{G}p)_j$ is the probability that chromosome j is sampled in the next generation. It is known that $\mathcal{G}p$ is the expected next generation, and that the actual transition rule from one generation to the next approaches $\mathcal{G}p$ as the population size tends to infinity. For this reason, this is sometimes referred to as the infinite population model.

\mathcal{F} gives the effects of proportional selection and is defined in terms of a diagonal matrix:

$$S_{k,k} = f(k)$$

where $f(k)$ is the fitness of chromosome k . The selection operator is then:

$$\mathcal{F}p = \frac{\mathcal{S}p}{\langle f \rangle(p)}$$

where $\langle f \rangle(p)$ is the average (mean) fitness of the population p given by

$$\langle f \rangle(p) = \sum_{j=0}^{s-1} p_j f(j)$$

Note that this definition is valid also for p outside the simplex.

The effects of one-point crossover and mutation (at a given rate) are given by \mathcal{M} (see [12] for details). For the case of zero crossover, this is simply a matrix M such that $M_{i,j}$ is the probability that chromosome j is mutated to chromosome i . Assuming a non-zero mutation rate and a non-zero fitness function, the following facts are known about this case:

1. Fixed-points of \mathcal{G} are eigenvectors of $G = MS$, once they have been normalised so that their components sum to one.
2. Eigenvalues of G give the average fitness of the corresponding fixed-point populations.
3. Exactly one normalised eigenvector of G is in the simplex Λ .
4. This eigenvector corresponds to the (simple) maximal eigenvalue.
5. The sequence $p, \mathcal{G}p, \mathcal{G}^2p, \dots$ converges for each $p \in \Lambda$ (that is, \mathcal{G} is *focused*).

The first two points follow from simple algebra. The other points come from the Perron-Frobenius theorem, which applies because G is positive (see for example [7]), and the fact that Λ is convex.

In practice, it may be difficult to calculate the eigensystem of G , not least because the matrices involved are extremely large. G is an $s \times s$ matrix where s is the number of chromosomes. Thus $s = 2^n$, for binary strings of length n . The situation can be made more tractable if one can define a set of equivalence classes on the set of chromosomes which is respected by the fitness function. This is the approach taken by van Nimwegen, Crutchfield and Mitchell [11] in analysing the dynamics of a simple genetic algorithm on the class of Royal Road fitness functions [4]. An overview of their results can be found in [10]. This approach is feasible because the number of distinguishable fitness values that can be taken on in a Royal Road function is very limited (see [12] for a further theoretical treatment of this idea).

The main focus of the study in [11] is the effect of finite populations (for certain mutation rates) in producing *metastable* states. During a typical run, the population will tend to linger in certain non-optimal states. This is due to the influence of fixed-points outside the simplex. If there is a fixed-point outside, but very close to the simplex, the force $|Gp - p|$ will be very small close to that fixed-point, since G is continuous. So a real population, inside the simplex, will have an expected next generation very close to itself, if it is near that fixed-point. For a finite population, there are only a limited number of viable points within the simplex, and so the population may stay in the vicinity of that point until a freak mutation sends it a suitably long distance away.¹

This paper applies the same technique to functions of unitation. These are functions of bitstrings that depend only on the number of ones in a string. The set of chromosomes can therefore be divided up into equivalence classes based on this unitation value. Once appropriate selection and mutation matrices have been defined, the fixed-points of some simple fitness functions will be studied. It will be seen that the interplay between fitness function and the mutation rate can be critical in determining where in the search space the fixed-point will lie.

2 Statistical dynamics for functions of unitation

Suppose we have a function of unitation:

$$u : \{0, 1\}^n \rightarrow \mathbb{R}$$

where $u(x)$ depends only on the number of ones in x and is always non-negative. That is, u is entirely defined by a function

$$f : \{0, \dots, n\} \rightarrow \mathbb{R}^+$$

By considering only the equivalence classes of chromosomes with equal numbers of ones, we can represent an infinite population by a vector

$$p = (p_0, p_1, \dots, p_n)$$

¹An alternative suggestion for finding metastable states for finite populations is to perform all eigensystem calculations to a precision of $1/m$, where m is the population size. Experiments with this approach have so far had only occasional success.

where p_k is the proportion of the population having exactly k ones. We now define the selection matrix S to be a diagonal matrix whose entries are

$$S_{k,k} = f(k)$$

The mutation matrix M , whose entries give the probabilities that a string with j ones becomes a string with i ones, is given by

$$M_{i,j} = \sum_{k=0}^{n-j} \sum_{l=0}^j \delta_{j+k-l,i} \binom{n-j}{k} \binom{j}{l} q^{k+l} (1-q)^{n-k-l}$$

where q is the mutation probability, and $\delta_{x,y}$ is the Kronecker delta function:

$$\delta_{x,y} = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases}$$

The matrix M is the same as the matrix T in [11].

We now have

$$\mathcal{G}p = \frac{MSp}{\langle f \rangle(p)}$$

where \mathcal{G} is the heuristic operator, adapted for our unitation equivalence classes. Given any population p , the expected next generation is $\mathcal{G}p$, and moreover this is the sampling distribution for the next generation. If v is a fixed-point of \mathcal{G} then

$$\frac{MSv}{\langle f \rangle(v)} = v$$

and so

$$Gv = MSv = \langle f \rangle(v)v$$

which means v is an eigenvector of $G = MS$ with its average fitness as the corresponding eigenvalue.

As outlined above, if there are no zero fitness values, the only (normalised) eigenvector in the simplex is the leading one, which demonstrates that there is a unique fixed-point for non-zero functions of unitation. If the fitness function can take on a zero value, then a population comprising copies of such a zero-valued string is also a fixed-point. However, any string with zero fitness will not get selected for the next generation under proportional selection, and so in practice we can ignore such cases.

We are now in a position to consider some example functions of unitation and calculate the eigenvectors of G . All calculations presented were computed using Mathematica version 2.2.3 running on Microsoft Windows 95. A number of experimental runs of a genetic algorithm are also presented. Unless otherwise stated, a population size of 500 is used (initialised at random), with experiments lasting for 200 generations, and averaged over 10 runs. Error bars of one standard deviation are shown. For these runs it should be noted that Stochastic Universal Sampling (SUS) [1] was used, rather than proportional selection. SUS ensures that, within the limits of the given population size, a population as near to that expected for proportional selection is generated. The above theory should still produce a good approximation of the fixed-points for this selection scheme. However, populations are

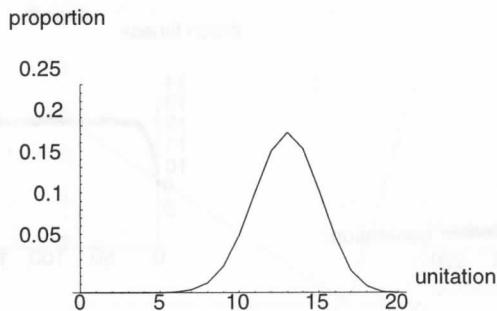


Figure 1: Fixed-point population distribution for onemax: $n = 20, q = 0.05$.

likely to exhibit much less variation from one generation to the next than with proportional selection. The advantage of this is that it should be seen more clearly when a fixed-point has been reached. A further advantage is that SUS is computationally more efficient than proportional selection, enabling more experiments to be conducted.

3 Simple examples of quasispecies

3.1 The onemax problem

The standard onemax problem is defined by $f(k) = k$ and so we have:

$$S_{k,k} = k$$

Figure 1 shows the fixed-point distribution calculated as the leading eigenvector of G , for the case $n = 20, q = 0.05$. The corresponding eigenvalue (mean fitness) is 12.94. It can be seen that this distribution is clustered some distance from the global optimum. Such a population is sometimes referred to as a *quasispecies* [3].

Figure 2 shows the results of 10 experimental runs. The Euclidean distance from the current population vector to the fixed-point (averaged over these runs) is shown for 200 generations, along with error bars of one standard deviation. In each run, the initial population was filled with strings selected with a uniform probability distribution. The average population fitness trajectory for the same 10 runs is also shown. It can be seen that the eigenvalue of 12.94 provides a good estimate of the average population fitness once the populations have converged.

3.2 Thompson's experiment

It has been suggested [9] that genetic algorithms may be good for finding *robust* solutions to problems, as they tend to settle on areas of the search space that are relatively flat with respect to mutation. This suggestion by Adrian Thompson was illustrated experimentally with a function of unitation on five bits defined by:

$$f(0) = 10$$

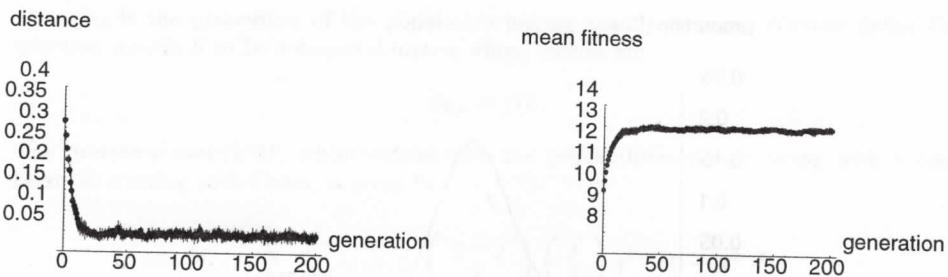


Figure 2: GA behaviour on the onemax problem $n = 20, q = 0.05$. Left: distance of population from fixed-point. Right: average population fitness.

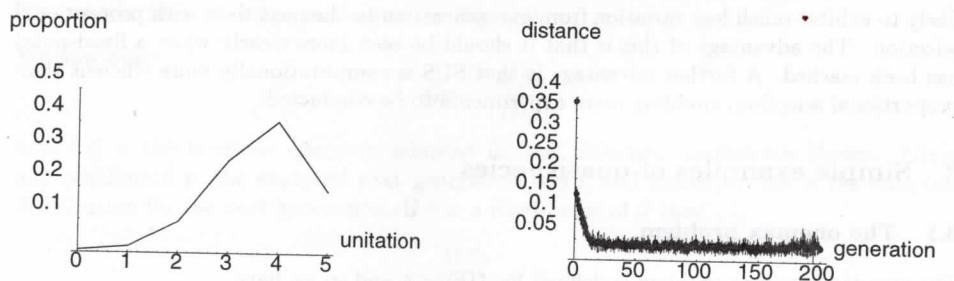


Figure 3: Left: fixed-point population for Thompson's function ($q = 0.2$). Right: distance of population from this fixed-point.

$$\begin{aligned} f(1) &= 0 \\ f(2) &= 0 \\ f(3) &= 0 \\ f(4) &= 5 \\ f(5) &= 9 \end{aligned}$$

Thompson demonstrated that a population starting entirely on the global optimum will move across to the suboptimum, which is broader (though lower) and therefore more robust. The infinite population fixed-point can be calculated using the eigenvector method and is shown in Figure 3. This provides an explanation for Thompson's experimental result, which indicated that the global optimum is unstable. A graph showing the distance from the current population to the fixed-point over 200 generations (averaged over 10 runs) is also shown. Again, the initial populations are filled with strings selected with a uniform probability distribution.

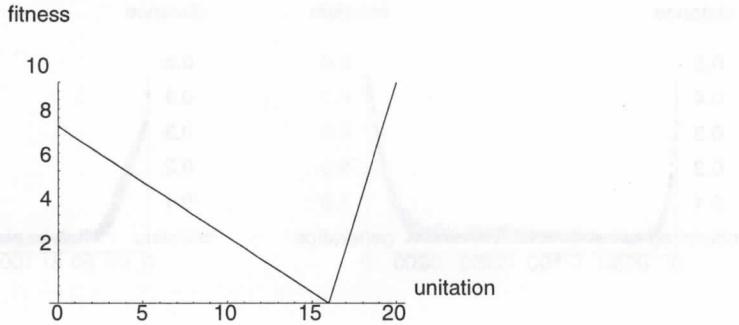


Figure 4: A typical trap function (not fully deceptive according to [2]).

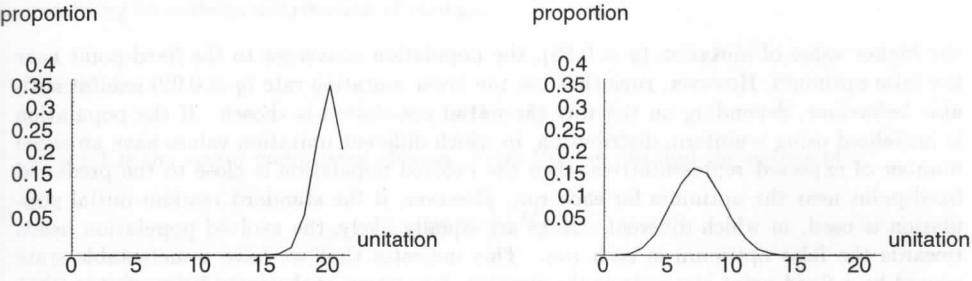


Figure 5: The fixed-point population for the trap function of Figure 4 is located near to the optimum for $q = 0.02$ (left) but is near the false optimum for $q = 0.05$ (right).

4 Trap functions and deception

Trap functions are piecewise linear functions of unitation, with a global optimum occurring with n ones and a local optimum at 0 ones (see Figure 4). The idea is that, under certain circumstances, a GA will be “misled” into the false optimum. Indeed, the infinite fixed-point population typically resides near one or other of the optima. If the mutation rate is small enough, then the population can sit on the steeper global optimum slope. For larger mutation rates, however, this slope is too steep and the fixed-point now resides near the local optimum (see figure 5).

These functions have been analysed in terms of the amount of deception they contain [2]. The notion of deception has already been questioned as a useful tool for analysing finite population behaviour [5, 8]. We now see that the effect of deception in a fitness function is dependent on the operators that are chosen. Deception analysis looks at a landscape independent of the operators. However, operators and landscape are interlinked: some mutation rates mean that the area surrounding the global optimum is unstable, even for functions that are not fully deceptive.

The graphs of distance from current population to fixed-point are shown in Figure 6. For

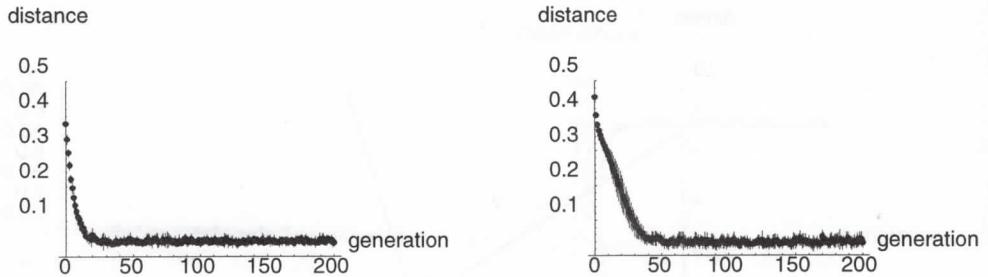


Figure 6: Distance from current population to fixed-point for the trap function with $q = 0.05$ (left) and $q = 0.02$ (right). Initial populations are drawn from uniform distribution of strings (left) and uniform distribution of unitation class (right).

the higher value of mutation ($q = 0.05$), the population converges to the fixed-point near the false optimum. However, runs that use the lower mutation rate ($q = 0.02$) exhibit variable behaviour, depending on the way the initial population is chosen. If the population is initialised using a uniform distribution, in which different unitation values have an equal number of expected representatives, then the evolved population is close to the predicted fixed-point near the optimum for each run. However, if the standard random initial population is used, in which different strings are equally likely, the evolved population heads towards the false optimum in each run. This indicates that we have a metastable state caused by a fixed-point just outside the simplex. Inspection of the second eigenvector (that is, with the second largest eigenvalue) indicates that it is extremely close to the simplex (it has one negative component of -3×10^{-14} after normalising). This eigenvector is illustrated in Figure 7. A graph of distance from current population to this (normalised) eigenvector is also shown, and confirms that this is indeed acting as an attractor.

5 Evolution can go backwards

It is usually taken for granted that an evolutionary process will tend to favour individuals with higher fitness values, and indeed this assumption is built into the mechanism of proportional selection. This describes the dynamic relation of one generation to the next. However, in the fixed-point limit, things can be very different. It is possible that the proportion of individuals in a population, within a given fitness class, be *inversely proportional* to that fitness. To achieve this, we require that the generational operator G have the following as an eigenvector (suitably scaled):

$$p_k = \frac{1}{f(k)}$$

$k = 0, \dots, n$. Using the definition of G gives:

$$MSp = \lambda p$$

for some λ . But the diagonal entries of S are simply the fitness values. That is:

$$S_{k,k} = f(k)$$

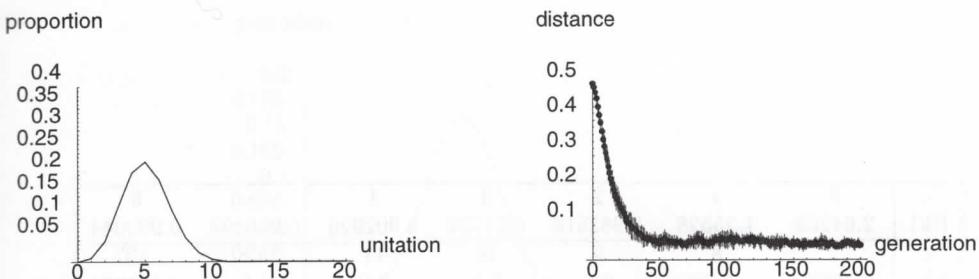


Figure 7: Left: the second normalised eigenvector for the trap function of Figure 4 is located very near to the simplex and so creates a metastable state in its vicinity ($q = 0.02$). Right: distance from current population to metastable state. The initial populations were chosen according to uniform distribution of strings.

Therefore:

$$M\mathbf{1} = \lambda p$$

where $\mathbf{1}$ is the vector comprising all ones. From this we can find the entries of p

$$p_k = \frac{1}{\lambda} (M\mathbf{1})_k$$

and so (choosing $\lambda = 1$):

$$f(k) = \frac{1}{(M\mathbf{1})_k} = \frac{1}{\sum_{l=0}^n M_{k,l}} \quad (1)$$

This function is tabulated in Table 1 for $n = 20$, $q = 0.05$, and illustrated in Figure 8, along with its fixed-point.

It should be noted that the reason this method works is that we are dealing with equivalence classes of individuals. Two individuals are equivalent if they share the same unitation value. Because of the way unitation is defined, these classes are of different sizes. It is therefore possible to assign different fitness values to these classes in such a way that the numbers of individuals in a class, times the allocated fitness value, equals 1. If the equivalence classes had equal sizes, then all the allocated values would be equal and there would be no surprises. This phenomenon, then, is entirely due to the way unitation breaks up the search space into classes of unequal size.

Experiments show that for small populations, the trajectory is towards a metastable state, given by the third eigenvector, illustrated in Figure 9, which is extremely close to the simplex. The second eigenvector is some distance from the simplex and is assumed to have no effect (although see the discussion in the next section). Since this metastable distribution is close to the initial distribution given the standard random generation method, experiments were run on this example using a uniform initial population (sampled so that each unitation value is equally likely). A population of 50 was used. Larger populations drawn in this way are very close to the true fixed-point and so the metastable state is avoided. The graph of distance from current population to the metastable state is shown in Figure 10. Notice that

k	0	1	2	3	4	5	6
$f(k)$	2.64269	1.25528	0.983518	0.91732	0.902926	0.900398	0.900044
k	7	8	9	10	11	12	13
$f(k)$	0.900004	0.9	0.9	0.9	0.9	0.9	0.900004
k	14	15	16	17	18	19	20
$f(k)$	0.900044	0.900398	0.902926	0.91732	0.983518	1.25528	2.64269

Table 1: A fitness function over 20 bits that has a fixed-point population in which the proportion of individuals within each fitness class is inversely proportional to that fitness. $n = 20, q = 0.05$.

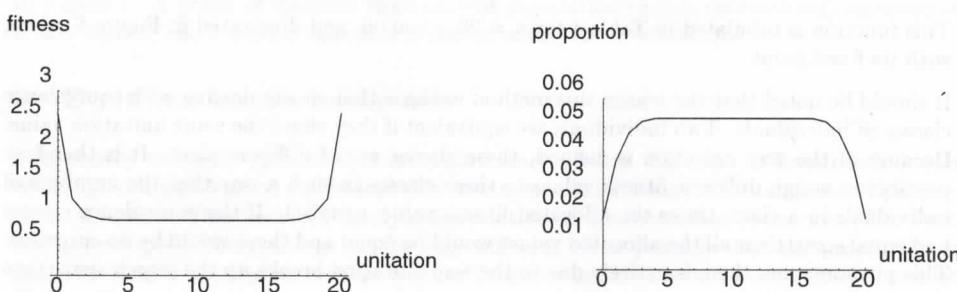
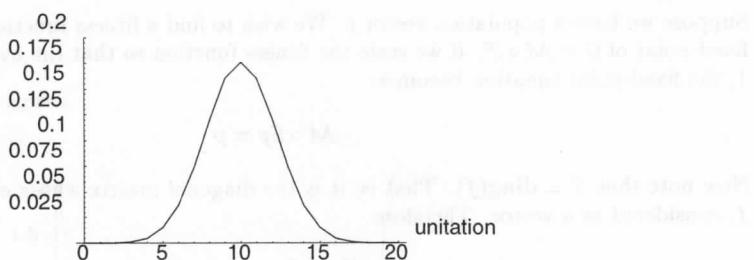
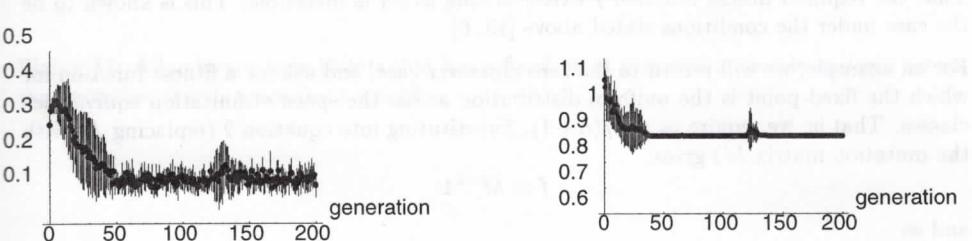


Figure 8: Left: graph of the fitness function tabulated in Table 1. Right: its fixed-point population ($n = 20, q = 0.05$).

proportion

Figure 9: The third eigenvector for the fitness function shown in Figure 8. $n = 20, q = 0.05$

distance

Figure 10: Left: distance from current population to metastable state for the fitness function shown in Figure 8. Right: The average population fitness over the same runs ($n = 20, q = 0.05$). The initial populations were chosen according to a uniform distribution of unitation classes. A population size of 50 was used.

this metastable state is centred around the global minimum, although the objective was to maximise the function. Plotting the average population fitness over time reveals that it is decreasing as the system evolves towards this minimum.²

6 Evolution can go anywhere

It is natural to wonder what can be said in general about the shape of infinite population fixed-points. Are the bell curves already seen typical? Are there any limits on the kinds of population that can form a fixed-point? How will the introduction of crossover affect the kinds of fixed-points that can be found? In fact, it turns out that *any* population distribution can be a fixed-point for some fitness function, even with non-zero, one-point crossover, given certain reasonable assumptions about crossover and mutation rates.

Theorem 1 *Given a population $p \in \Lambda$, there exists a fitness function such that p is a*

²The fluctuation at generation 125 is due to a random drift in one of the runs.

fixed-point of $\mathcal{G} = \mathcal{M} \circ \mathcal{F}$, provided $0 < q < 1/2$ and the crossover rate is less than 1.³

Suppose we have a population vector p . We wish to find a fitness function such that p is a fixed-point of $\mathcal{G} = \mathcal{M} \circ \mathcal{F}$. If we scale the fitness function so that the average fitness of p is 1, the fixed-point equation becomes:

$$\mathcal{M} \circ Sp = p$$

Now note that $S = \text{diag}(f)$. That is, it is the diagonal matrix whose entries are values of f , considered as a vector. Therefore

$$p = \mathcal{M} \circ \text{diag}(f)p = \mathcal{M} \circ \text{diag}(p)f$$

and so

$$f = \text{diag}(p)^{-1} \circ \mathcal{M}^{-1}p \quad (2)$$

Thus the required fitness function f exists so long as \mathcal{M} is invertible. This is known to be the case under the conditions stated above [13, 6].

For an example, we will return to the zero crossover case, and ask for a fitness function for which the fixed-point is the uniform distribution across the space of unitation equivalence classes. That is, we require $p_k = 1/(n+1)$. Substituting into equation 2 (replacing \mathcal{M} with the mutation matrix M) gives:

$$f = M^{-1}\mathbf{1}$$

and so

$$f(k) = \sum_{l=0}^n M_{k,l}^{-1}$$

A typical such landscape is illustrated in Figure 11, where the function is defined over 10 bits and $q = 0.08$. For populations of size 500, the behaviour of the genetic algorithm is quite interesting. The population seems to wander at random around the search space, never settling down. A typical run over 2000 generations is shown in Figure 12.

It would appear that the population is alternately visiting the fixed-point and a nearby metastable state. An inspection of the first three eigenvectors reveals that the second normalised eigenvector is a long way from the simplex (components of the order of 10^{10}), whereas the third is quite close. It might be thought that it is the third that is creating the metastable state. However, experiments reveal that this is not the case and, over many runs, populations do not approach this state. The answer is to reconsider the second eigenvector, in the light of the following two observations (which assume zero crossover):

Theorem 2 *If a population vector p lies on the line joining two normalised eigenvectors of G , then so does Gp . Moreover, Gp will be closer to the eigenvector with the greater eigenvalue.*

Consider two eigenvectors v and w . A point on the line connecting them may be expressed as

$$p = \lambda v + (1 - \lambda)w$$

³The generalisation of this theorem to include crossover is due to one of the reviewers.

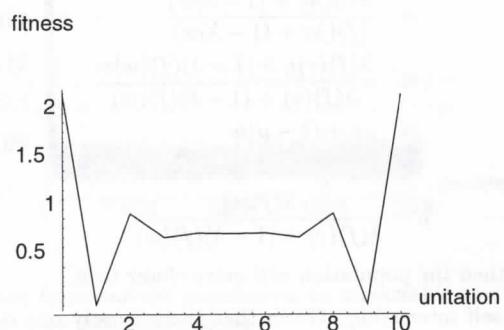


Figure 11: A landscape over 10 bits that has a fixed-point population that is uniform across the unitation search space for $q = 0.08$

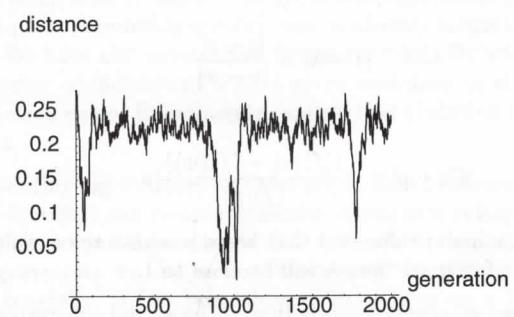


Figure 12: A typical trajectory (plotted as distance from the fixed-point) for the fitness function shown in Figure 11, over 2000 generations.

for some scalar λ . Then

$$\begin{aligned} \mathcal{G}p &= \frac{MSp}{\langle f \rangle(p)} \\ &= \frac{MS(\lambda v + (1 - \lambda)w)}{\langle f \rangle(\lambda v + (1 - \lambda)w)} \\ &= \frac{\lambda \langle f \rangle(v)v + (1 - \lambda)\langle f \rangle(w)w}{\lambda \langle f \rangle(v) + (1 - \lambda)\langle f \rangle(w)} \\ &= \mu v + (1 - \mu)w \end{aligned}$$

where

$$\mu = \frac{\lambda \langle f \rangle(v)}{\lambda \langle f \rangle(v) + (1 - \lambda)\langle f \rangle(w)}$$

If, say, $\langle f \rangle(v) > \langle f \rangle(w)$ then the population will move closer to v .

This observation is of itself interesting. Given that there is only one eigenvector actually in the simplex, it follows that the flow of p follows a straight line from each of the other eigenvectors to the fixed-point.

Theorem 3 *Let v be the fixed-point in the simplex and w another normalised eigenvector. Let $p \in \Lambda$ be on the line connecting v and w . Then there is a constant k such that*

$$|\mathcal{G}p - p| < k \frac{(\langle f \rangle(v) - \langle f \rangle(w))}{\langle f \rangle(p)} |p - v|$$

Again, set $p = \lambda v + (1 - \lambda)w$. Since $p \in \Lambda$ note that $\lambda > 0$. Then some algebra produces:

$$\mathcal{G}p - p = \lambda(1 - \lambda) \frac{(\langle f \rangle(v) - \langle f \rangle(w))}{\langle f \rangle(p)} (v - w)$$

Substituting

$$|1 - \lambda| = \frac{|p - v|}{|v - w|}$$

and taking norms gives:

$$|\mathcal{G}p - p| = \lambda \frac{(\langle f \rangle(v) - \langle f \rangle(w))}{\langle f \rangle(p)} |p - v|$$

Choosing k to be the maximum value of λ that keeps p within the simplex gives the result. Note that if v and w are far apart then λ will be close to 1.

The consequence of these two observations is that, if the second eigenvalue is extremely close to the leading eigenvalue then, even if the eigenvectors are far apart, the distance moved by a population within the simplex is small all along the line connecting the two eigenvectors. Returning to our example, the second eigenvector is indeed far from the simplex, but the eigenvalue, 0.9997, is close to that of the fixed-point (whose eigenvalue is 1). This indicates that there is not simply a metastable state but a metastable line. A little geometry allows us to calculate the shortest distance from a population to this line. We now plot this distance over 2000 generations (averaged over 10 runs). Figure 13 shows that populations indeed head towards that line and then stay more or less on it.

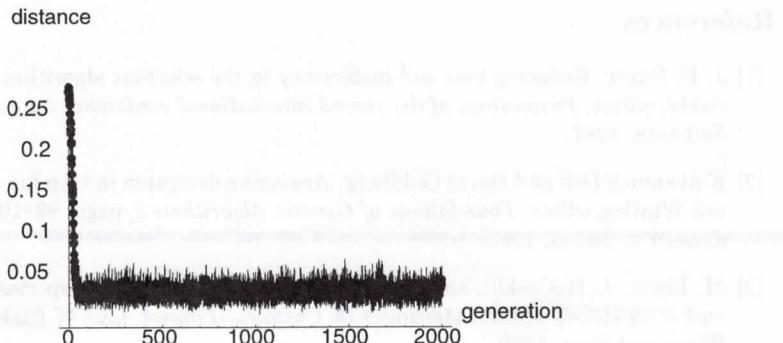


Figure 13: Distance from current population to metastable line for the fitness function shown in Figure 11, over 2000 generations.

7 Conclusion

The dynamical systems model of the GA [12] can be simplified in the case of zero crossover. By considering functions which take on a limited number of values, the dimensionality of the system can be reduced to the point where calculations are tractable. This method, used in [11] to study Royal Road functions, has been adapted to apply to functions of unitation. The infinite population fixed-points of these fitness functions, given a mutation rate, are readily calculated for problems of modest size.

It has been shown theoretically that any population distribution is potentially a fixed-point of some landscape (even with crossover). In particular, there exist non-flat fitness functions for which the fixed-point population is distributed uniformly across the entire search space of unitation classes. We have also constructed fitness functions for which the final population represents the number of individuals with a given unitation in inverse proportion to the fitness of that unitation value. Experiments confirm that evolution appears to go backwards on such landscapes.

The effect of non-leading eigenvectors on finite population behaviour has also been studied. It has been shown that they can create metastable states as previously predicted. It has also been shown that if the second eigenvalue is close to the first, then a line through the simplex connecting the eigenvectors will be metastable, even if the corresponding eigenvectors are far apart. More generally, it has been shown that if p is on a line connecting any two eigenvectors, then so is Gp (assuming zero crossover). The force, $|Gp - p|$ is proportional to the difference between the corresponding eigenvalues.

8 Acknowledgements

I would like to acknowledge the many useful comments of the reviewers of this paper, and helpful conversations with Michael Vose and Alden Wright.

References

- [1] J. E. Baker. Reducing bias and inefficiency in the selection algorithm. In J. J. Grefenstette, editor, *Proceedings of the second international conference on genetic algorithms*. Erlbaum, 1987.
- [2] Kalyanmoy Deb and David Goldberg. Analyzing deception in trap functions. In L. Darrell Whitley, editor, *Foundations of Genetic Algorithms 2*, pages 93–108. Morgan Kaufmann Publishers, 1993.
- [3] M. Eigen, J. McCaskill, and P. Schuster. The molecular quasi-species. In I. Prigogine and S. A. Rice, editors, *Advances in Chemical Physics LXXV*, pages 149–263. John Wiley and Sons, 1989.
- [4] Stephanie Forrest and Melanie Mitchell. Relative building-block fitness and the building block hypothesis. In L. Darrell Whitley, editor, *Foundations of Genetic Algorithms 2*, pages 109–126. Morgan Kaufmann Publishers, 1992.
- [5] John J. Grefenstette. Deception considered harmful. In L. Darrell Whitley, editor, *Foundations of Genetic Algorithms 2*, pages 75–91. Morgan Kaufmann Publishers, 1992.
- [6] Gary J. Koehler, Siddartha Bhattacharyya, and Michael D. Vose. General cardinality genetic algorithms. *Evolutionary Computation*, 5(4):439–459, 1997.
- [7] Erwin Kreysig. *Advanced engineering mathematics*. John Wiley & Sons, 1988.
- [8] Jon Rowe and Ian East. Deception and GA-hardness. In *Proceedings of the First Nordic Workshop on Genetic Algorithms*. University of Vaasa Press, 1995.
- [9] Adrian Thompson. Evolving fault tolerant systems. In *Proceedings of the First IEEE/IEEE International Conference on Genetic Algorithms in Engineering Systems*, pages 524–529. IEE, 1995.
- [10] Erik van Nimwegen, James P. Crutchfield, and Melanie Mitchell. Finite populations induce metastability in evolutionary search. *Physics Letters A*, 229(2):144–150, 1997.
- [11] Erik van Nimwegen, James P. Crutchfield, and Melanie Mitchell. Statistical dynamics of the Royal Road genetic algorithm. Technical Report 97-04-035, Santa Fe Institute, 1997. To appear in the Special Issue on Evolutionary Computation of the Journal of Theoretical Computer Science, A. E. Eiben and G. Rudolph, editors.
- [12] Michael D. Vose. Random heuristic search. Technical report, University of Tennessee, Knoxville, TN, 1998. To appear in the Special Issue on Evolutionary Computation of the Journal of Theoretical Computer Science, A. E. Eiben and G. Rudolph, editors.
- [13] Michael D. Vose and Alden H. Wright. The Walsh transform and the theory of the simple genetic algorithm. In S. Pal and P. Wang, editors, *Genetic Algorithms for Pattern Recognition*. CRC Press, 1996.