Exploration of population fixed-points versus mutation rates for functions of unitation

J Neal Richter[1], Alden Wright[2], John Paxton[1]

[1] Computer Science Department, Montana State University, 357 EPS,
Bozeman, Montana 59714
`{richter,paxton}@cs.montana.edu`
`http://www.cs.montana.edu`
[2] Department of Computer Science, Social Science Building Room 401
The University of Montana-Missoula, Missoula, MT 59812
`wright@cs.umt.edu`
`http://www.cs.umt.edu`

Abstract. Using a dynamic systems model for the Simple Genetic Algorithm due to Vose, we analyze the fixed point behavior of the model without crossover applied to functions of unitation. Unitation functions are simplified fitness functions that reduce the search space into a smaller number of equivalence classes. This reduction allows easier computation of fixed points. The effect of model fixed points that are outside of the population space, but near enough to influence behavior of the Simple GA will be examined.   This analysis will be applied to several fitness functions of unitation.

1   Introduction

The Vose infinite population model [1] of simple genetic algorithms is a dynamic systems model that represents populations as a vector of proportions. This vector has dimension s, where s is the size of the search space. Each entry in the vector is the proportion of members in the global population that are represented by a given chromosome in the search space. This representation allows utilization of techniques and theorems from the mathematical theory of dynamic systems to analyze the GA.

This paper is restricted to examining a class of fitness functions called 'functions of unitation'. These functions establish equivalence classes, allowing a reduction in the dimensionality of the corresponding Vose model. The analysis will also be restricted to mutation-selection GAs with no crossover, elitism or other advanced techniques.

Using the eigensystem of the mutation-only Vose model, fixed points of the mutation-selection GA can be found. Fixed points are population vectors such that applying the GA to them results in the same population vector. Fixed points are not the optimal points in the fitness landscape, they represent the expected long-run distribution of a population for a given GA with a large enough population. A fixed point may or may not contain a significant proportion of the global optima of the fitness landscape.

Fixed points will be calculated over a range of mutation rates for several fitness functions. Metastable states defined by the model will also be calculated. These states are fixed points of the model that exist outside the representable space of a real population. The metastable states near the population space can create metastable regions inside the space and have an effect on the GA under real populations [3].

This paper is largely an extension of Rowe [5] that further explores the effect of mutation rates on fixed points and metastable states. The intent of this exploration is to lay the groundwork for studying how adaptive mutation schemes can be understood and designed.

2 The Vose Dynamic Systems Model

The Vose infinite population model is largely the work of Michael Vose[1]. The intent of the model is to allow a mathematical analysis of the Simple Genetic Algorithm. For our purposes, we define the Simple Genetic Algorithm as a GA with proportionate selection, bitwise mutation and several standard crossover operators.

2.1 Background

The Vose model is a discrete dynamical system and is a kind of 'map'. Maps are discrete dynamic systems that translate an input to an output. An example is the well-known Logistic Map [2].

$$f(x) = \lambda x(1-x) \qquad \{\text{where } \lambda \in \Re \; ; \; \lambda > 0 \; ; x \in \Re \; \} \tag{1}$$

The Vose model translates the current population vector to the next population vector. Iterating the map simulates the trajectory of the GA (in the limit of infinite population size), where the current population vector becomes the input to the map, giving the next generation of the GA. This forms a sequence of population vectors p_1, p_2, p_3 ,..., p_k . This sequence is the trajectory of the GA model through the population space.

The Vose model is a deterministic dynamical system. For a sufficiently large population in a real GA, this model allows us to make fairly accurate predictions of the *expected* next population and the long term behavior of the population [3]. If the population is small, then the actual populations produced by the GA will have wide variability compared to the predicted population.

If the population is very large, then the actual populations produced by the GA should be close to the predicted model for some number of generations [1]. Thus the Vose model is called an *infinite population model.*

2.2 Theory

The following section summarizes Rowe [3] as it applies to selection-mutation models of GAs. Given a binary genome of length d, the search space of the GA is of size $s=2^d$.

To use the Vose model we represent the population as a vector of proportions of length s, $\boldsymbol{p} = (p_0,..., p_{s-1})$. Each p_i is the proportion of membership in the population by the binary string i.

For a 2-bit genome, a possible population vector is $\boldsymbol{p} = (0.1, 0.2, 0.5, 0.2)$. This could represent a population of 10 individuals, 1 copy of `00`, 2 copies of `01`, 5 copies of `10`, and 2 copies of `11`.

Note that the population vectors have the simplex property. Each component of the vector is in the range [0,1], and the sum of the components equals 1.

$$\Lambda = \left\{ (x_0,..., x_{s-1}): \sum_{i=0}^{s-1} x_i = 1, \quad x_j \in [0,1] \quad x \in \Re \right\} \tag{2}$$

Next the properties of the model for mutation-selection GAs are given. The mutation-selection infinite population model is created as follows. Operator G is defined by:

$$p_{t+1} = G(p(t)) = \frac{1}{\mu(p_t)} U\ S\ p_t \tag{3}$$

where $\mu(p_t)$ is the average fitness of the population p_t. The s x s mutation matrix U is composed of the probabilities that a chromosome string j will mutate into string i. Note that U is symmetric. The s x s selection matrix S is a diagonal matrix consisting of fitness values along the diagonal. Dividing by $\mu(p_t)$ implements proportionate selection.

$$\begin{aligned} U_{i,j} &= q^h (1-q)^{d-h} \quad h = \text{Hamming Distance } (i, j) \\ S_{k,k} &= f(x_k) \qquad\qquad k \in [0, s-1] \end{aligned} \tag{4}$$

From the theory we know these five properties of the US matrix [3].

1. US is a positive matrix: all entries are non-negative.
2. Fixed-points of the model are the normalized (so that all elements sum to 1) eigenvectors of US.
3. Only one normalized eigenvector is in the simplex (via Perron-Frobenius theorem [3][8]).
4. Eigenvalues of US are the average fitness of the population given by the corresponding eigenvector.
5. The largest eigenvalue corresponds with the lone eigenvector inside the simplex.

These properties allow the computation of the fixed points of the infinite population model for a given fitness function and mutation rate. By normalized vector we mean that all elements have been scaled such that they all sum to 1, creating a vector that obeys the simplex property. The lone normalized fixed point in the simplex is the global attractor of the dynamical system modeling the GA.

There can exist fixed points outside, but very near, the simplex. Such vectors do not obey the simplex property (they contain at least one negative number) and cannot represent a real population. Called 'metastable states' [3], they can exert an influence on the GA for real populations near these fixed points. They can form 'metastable regions' inside the simplex [3]. These regions are similar to the saddle points of elementary dynamical systems. Populations can escape from these regions, quick escape is likely only if the population size is small or the influence of the metastable state is weak (e.g. the metastable state is relatively far away from the simplex).

Note that fixed size populations form a subset of the simplex called a *lattice* [4]. Some simplex population vectors (an example is any vector with an element like 2/3) are not representable with finite populations. The finite population GA moves from lattice point to lattice point in the simplex. The smaller the population size, the sparser the lattice points in the simplex.

3 Functions of Unitation

Unitation functions are fitness functions where fitness is defined only by the number of ones in a chromosome $x:\{1,0\}^d$. All fitness values are non-negative:

$$u(x):\{0,1\}^d \rightarrow \Re^+ \tag{5}$$

An example function is

$$u(0) = 3 \qquad u(1) = 2 \qquad u(2) = 1 \qquad u(3) = 4$$

This definition allows us to reduce the dimensionality of the infinite proportionality population vector from 2^d x 1 to $(d+1)$ x 1. This vector is represented as $\boldsymbol{p} = \{p_0,...., p_d\}$, where p_k is the proportion of the population having exactly k ones. Note that this vector is of $d+1$ dimension as it must have entries for the all zeros case, the all ones case and all cases in between.

Using the above fitness function, a selection matrix S is defined as the $(d+1)$ x $(d+1)$ diagonal matrix $S_{k,k} = u(k)$. The mutation matrix U is defined as an $(d+1)$ x $(d+1)$ matrix with each entry computed using the following equation [4][5].

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

$$\text{where} \quad \delta_{j+k-l,i} = \begin{cases} 1 & if\ x = y \\ 0 & if\ x \neq y \end{cases} \tag{6}$$

$\delta_{x,y}$ is the Kronecker delta function and q is the mutation probability. This is the probability a given bit in the chromosome string mutates to its complement state.

3.1 Example Functions of Unitation

The three fitness functions pictured in Figure 1 are called NEEDLE, BINEEDLE and ONEMAX, and have been studied before by Rowe [5] and Wright [6]. Here d=20, α=20 and α=5 are used for NEEDLE and BINEEDLE.

$$\text{NEEDLE} = \begin{cases} 1+\alpha & \text{all ones string} \\ 1 & \text{otherwise} \end{cases}$$

$$\text{BINEEDLE} = \begin{cases} 1+\alpha & \text{all ones string} \\ 1 & \text{otherwise} \\ 1+\alpha & \text{all zeros string} \end{cases} \tag{7}$$

$$\text{ONEMAX} = \text{number of ones in string}$$
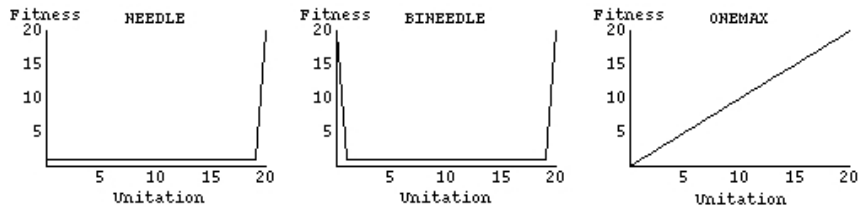
Figure 1. NEEDLE, BINEEDLE and ONEMAX fitness functions.

4   Fixed Points as a Function of Mutation Rate

Figure 2 shows two fixed points for the ONEMAX fitness function. These are the normalized leading eigenvector of G with mutation rate q=0.01 and q=0.05 both with d=20, providing 21 unitation classes. They show, for example, that for q=0.01 approximately 25% of the population should contain strings with 17 bits of value 1 after a sufficiently large number of generations of the mutation-selection GA have been computed and a very large population.
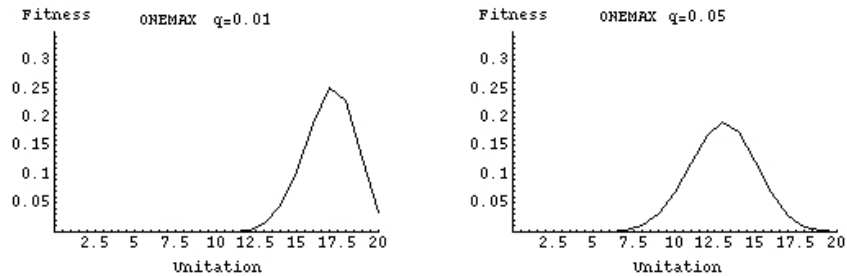

Figure 2. ONEMAX fixed point distributions for q=0.01 and q=0.05.

Next, a sequence of fixed points were computed for mutation rates starting at $q=0.01$, and the population distributions were plotted as a 3-dimensional surface. Figure 3 shows the ONEMAX fixed points plotted from $q=0.01$ to $q=0.20$. Note that at mutation rates near 0.01, the population contains a significant proportion of the $f(20)=20$ optimal mutation class. By $q=0.05$, there is near zero membership.
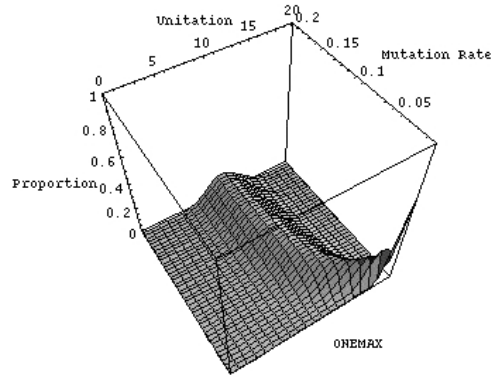

Figure 3. ONEMAX fixed point surface.

4.1 Fixed-point Surfaces for the NEEDLE and BINEEDLE Functions

Figures 4 and 5 contain the fixed point surfaces of NEEDLE and BINEEDLE for both $\alpha=5$ and $\alpha=20$. Figure 4 tells us that for low mutation rates, i.e. mutation rates below $q=0.025$ and $q=0.05$ respectively, NEEDLE has a significant proportion of the population at the maximum.
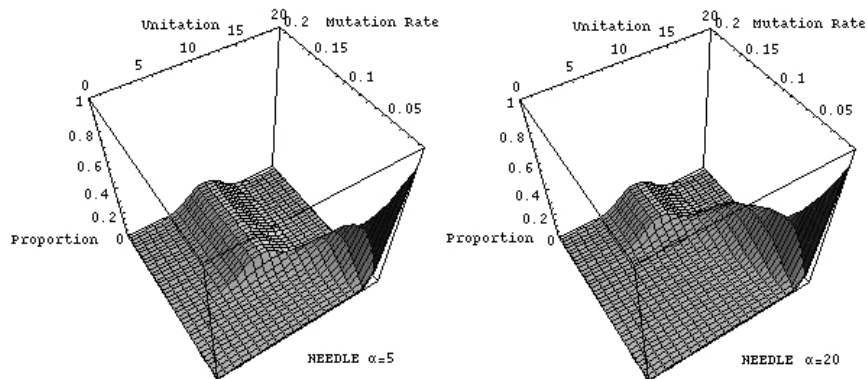

Figure 4. NEEDLE fixed point surfaces with $\alpha=5$ and $\alpha=20$.

In Figure 5 note the dramatic change in population distribution for BINEEDLE that occurs near $q=0.07$ (for $\alpha=5$). Above this mutation rate the population contains near

zero proportional membership by either global optimum. This tells us that mutation rates below this value are likely very important for good GA performance. For $\alpha=20$ this phase change occurs near $q=0.13$, indicating that this version of the fitness function is more tolerant of higher mutation rates. NEEDLE has similar properties.

Notice that for mutation rates greater than the critical q values for both functions, the population is centered around the unitation midpoint, a string of 10 ones and 10 zeros.
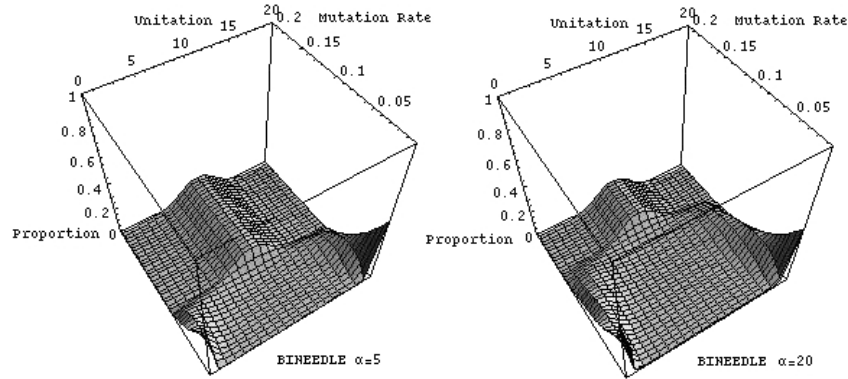

Figure 5. BINEEDLE fixed point surfaces with $\alpha=5$ and $\alpha.=20$.

The phase transitions for NEEDLE and BINEEDLE were studied in [6] and [9], and are shown by the Eigen model [10]. Ochoa and Harvey [9] restate the Eigen model for the GA community and show how the Eigen 'error thresholds' change under finite populations. Finite populations move the phase transition to lower mutation rates.

4.2 Fixed-point Surface for a Fully Deceptive Trap Function

Trap functions are piecewise linear functions that divide the search space into two Hamming space basins [7]. Each basin has an optimal point, one of which is the global optimum. In Deb and Goldberg [7], they set forth a set of conditions for calling a fitness function 'fully deceptive'. We have adopted a fully deceptive function from [7], and refer to it as DECTRAP.

$$f(x) = \begin{cases} 1 & \text{if } u(x) = d \\ 1 - \dfrac{1 + u(x)}{d} & \text{otherwise} \end{cases} \tag{8}$$

Figure 6 illustrates DECTRAP and its fixed-point surface. The trap function is very near a fitness of 1 for the all zeros string, and is fitness 1 for the all ones string. The all zeros basin takes up the majority of the function space.

The fixed-point surface has a drastic phase change at approximately $q=0.012$. Below this mutation rate a high proportion of the globally optimal string exists. Above

this mutation rate the fixed point contains nearly zero proportion of the global optimal. Notice again that as the mutation rate increases, the fixed point moves toward a population centered around the unitation midpoint.
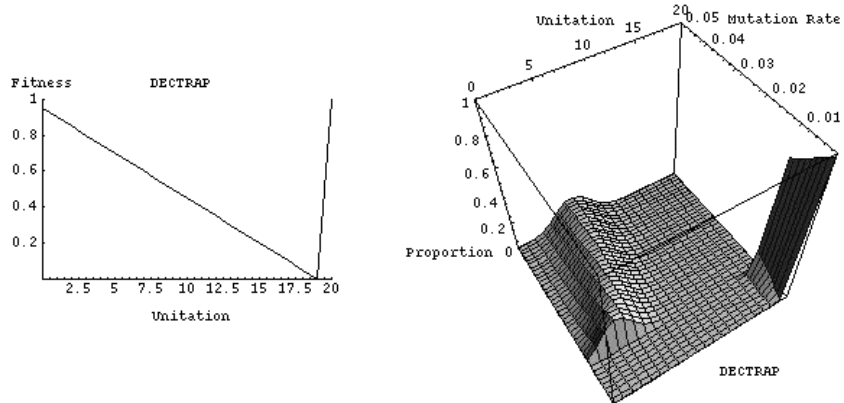

Figure 6. Fully deceptive trap functions DECTRAP and fixed point surface.

4.3 Fixed-point Surface for Functions with Two Traps

Figure 7 illustrates a trap function containing two traps, referred to as 2TRAP. The fixed point surface is very similar to the BINEEDLE surface with a critical phase change at q=0.04. Note that mutation rates below 0.02 are slightly superior since the fixed points are still centered much closer to the two local optima. As the mutation rate increases from 0.02 to 0.04, the population clusters move away from the maximums. The mutation rates are too high to maintain high membership.
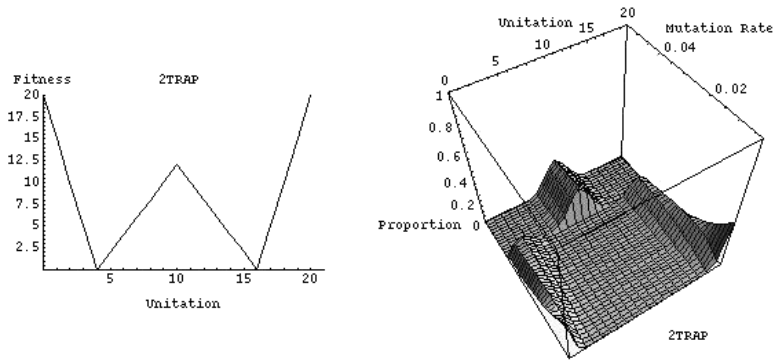

Figure 7. Double trap function 2TRAP and fixed point surface.

Figure 8 shows a deceptive double trap function, or DEC2TRAP. This function is modeled after the fully deceptive function given in [7]. The formula is given here:

$$f(x) = \begin{cases} 1 & \text{if } u(x) = d/2 \\ 1 - \dfrac{1+u(x)}{d/2} & \text{if } u(x) < d/2 \\ \dfrac{u(x) - d/2 - 1}{d/2} & \text{if } u(x) > d/2 \end{cases} \qquad (7)$$
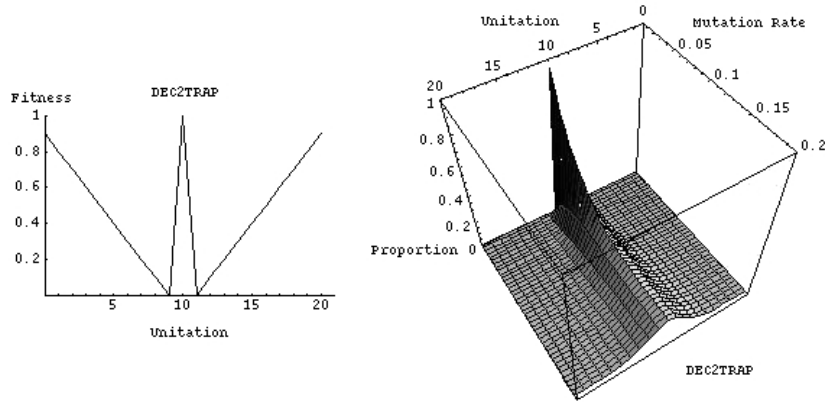

Figure 8. Deceptive double trap function DEC2TRAP and fixed point surface.

DEC2TRAP's fixed point landscape is very interesting in that it has virtually no membership of the $u(x)=0$ and $u(x)=20$ high fitness points. This is result is counter-intuitive. The regions on either side of the center optimal needle have smooth hills to climb that lead to the local maximums. A practioner might expect empirical GAs to retain membership in the local optimas, given that low mutation rates make it harder for a population to move outside the basin of either local maxima. Lack of elitism and other advanced features in the model partially explains the result, as well as the fact that after a high enough number of generations has passed, the Simple GA will converge totally into the basins of global maximums. Another reason is that there exist many strings in the $u(x)=10$ class, and only one string in the $u(x)=0$ and $u(x)=20$ classes.

5 Metastable States

Fixed points have a region around them generally called a 'basin of attraction' [2]. Loosely speaking, these are regions where the fixed point exerts influence. The Perron-Frobenius theorem [8] shows that the *US* matrix has only one fixed point inside the simplex. The other eigenvectors are called metastable states [3].

Rowe [3] defines 'metastable regions' as regions inside the simplex near metastable states close to the simplex. The continuity of Vose infinite population model implies

these regions exist [3]. The question of how these states move around the neighborhood of the simplex as the mutation rate changes is explored next.

For the functions previously described, the sum the negative components of the eigenvalues is computed. The lowest sum is potentially the closest metastable state to the simplex. Obviously this is not always the case, a true geometric distance will define exactly which of these vectors is closest to the simplex.

Computing a geometric distance involves finding the minimum distance from a point to an n-1 dimensional hypertetrahedron. For now, the first metric is used.

Figure 9 plots this distance metric for the $\alpha=5$ NEEDLE and BINEEDLE on a logarithmic scale. Also shown is the population proportion of the $u(x)=10$ string vs. mutation rate for comparison. Both functions have a similar metastable state movement and visually identical population proportion graphs. Note that the phase changes in all four graphs take place at approximately $q=0.075$ where the closest metastable state moves suddenly farther away. Figure 10 shows the identical graphs for $\alpha=20$. They show essentially the same effect as $\alpha=5$ except the phase changes take place at $q=0.13$.
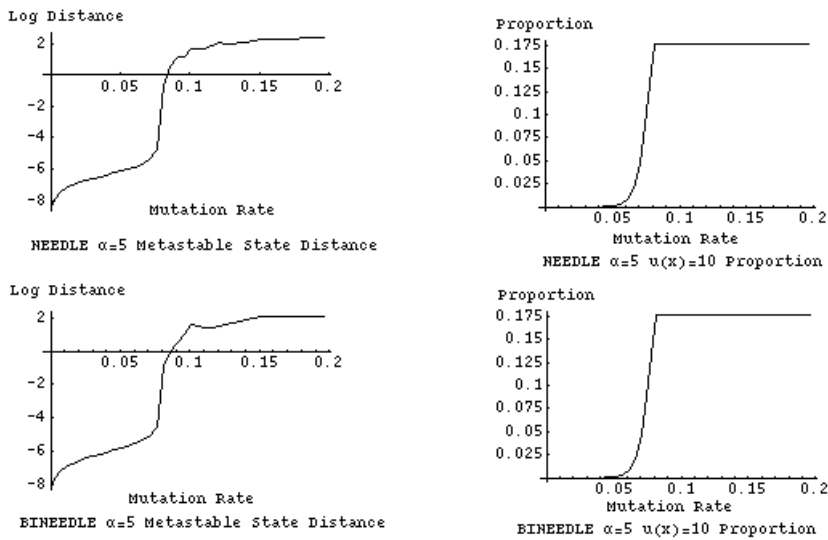

Figure 9. Metastable point distances and $u(x)=10$ proportions for $\alpha=5$ NEEDLE and BINEEDLE.

Figure 11 shows the same plots for the three trap functions ONEMAX. The plots for DECTRAP and DEC2TRAP show similar metastable state movement, with DEC2TRAP's nearest metastable state approximately twice as far away as DECTRAP for a given q. 2TRAP's metastable plot contains some interesting sharp edges that persisted for various settings from 30 to 40+ digits of precision.
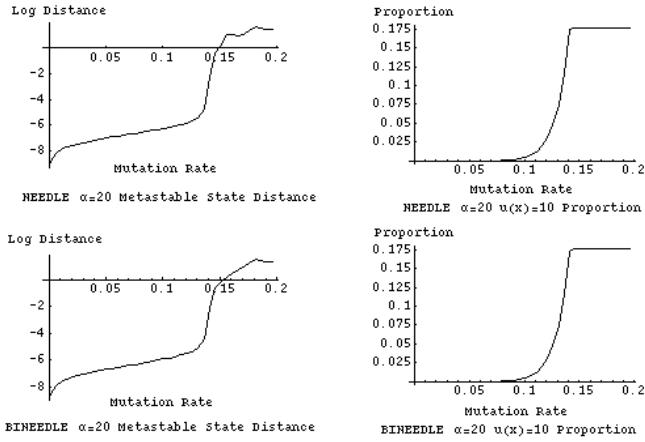
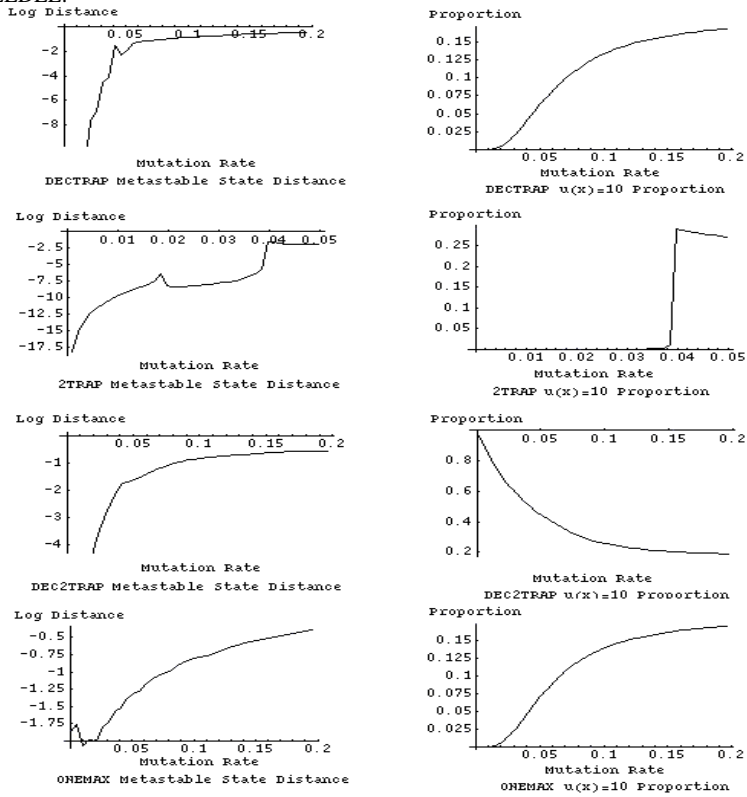**Figure 10.** Metastable state distances and $u(x)=10$ proportions for $\alpha=20$ NEEDLE and BINEEDLE.

Figure 12. Metastable point distances and $u(x)=10$ proportions for the trap functions and ONEMAX

6 Conclusions and Future Work

It is common to read in GA research papers that a specific mutation rate is used with no justification for a particular value. Hopefully this paper will prompt the readers to question the validity of their choice of mutation rate in the next GA project they undertake.

We have shown numerical results analyzing fixed-points and metastable state proximity for eight different fitness functions applied to functions of unitation. We chose to look at unitation functions since they provide a level of dimensional reduction to ease the tractability of theoretical analysis while still being interesting functions.

In the future we will investigate how the lessons learned here can influence the choice of adaptive mutation rate schemes and valid mutation rate ranges. We will also implement a geometric distance algorithm for finding the closest metastable states and apply this to other fitness functions.

We believe this type of analysis will help aid the understanding of simple mutation-only evolutionary algorithms like the (1+1) EA and related algorithms, as well as dynamic parameter schemes for these algorithms.

References

1. Vose, M.D. The Simple Genetic Algorithm: Foundations and Theory, MIT Press, Cambridge, MA. (1999)
2. Strogatz, S.H., Nonlinear Dynamics and Chaos. Addison-Wesley, Reading, MA. (1994)
3. Rowe, J. E. The Dynamical Systems Model of the Simple Genetic Algorithm, Theoretical Aspects of Evolutionary Computing, L. Kallel, B. Naudts, and A. Rogers, eds., Springer Verlag, New York, (2001) 31-57.
4. van Nimwegen E., James P. Crutchfield, J.P., Mitchell, M.: Finite Populations Induce Metastability in Evolutionary Search, Physics Letters A Vol. 229, Elsevier, New York (1997) 144-150.
5. Rowe, J.: Population fixed-points for functions of unitation. In Foundations of Genetic Algorithms. Vol. 5. Colin Reeves, and Wolfgang Banzhaf (eds.). Morgan Kaufmann, (1998) 69-84.
6. Wright A.H., Rowe J.E., Neil, J.R.: Analysis of the Simple Genetic Algorithm on the Single-peak and Double-peak Landscapes, Proc. of the 2002 Congress on Evolutionary Computation, Fogel et al. Editors, IEEE Press (1999), 214-219.
7. Deb, K., Goldberg, D.E.: Analyzing Deception in Trap Functions. In Foundations of Genetic Algorithms, Vol. 2. D Whitley, editor. Morgan Kaufmann, (1992): 93-108.
8. Brin, M., Stuck, G., Introduction to Dynamical Systems. Cambridge University Press, New York (2002).
9. Ochoa, G., Harvey, I. Recomcination and error thresholds in finite populations. In Foundations of Genetic Algorithms. Vol. 5. Colin Reeves, and Wolfgang Banzhaf (eds.). Morgan Kaufmann, (1998) 245-264.
10. Eigen, M., McCaskill, J., Schuster, P., Molecular quasi-species. Journal of Physical Chemistry, Vol 92, No 24, (1988) 6881-6891.