Computational Evolution Modeling and Neutral Theory

A. SCOTT & S. BAHAR

Department of Physics & Astronomy and Center for Neurodynamics, University of Missouri - St Louis, St Louis, MO

Introduction

Optimization studies may be simplified by considering phenotypes of organisms as opposed to explicit genotypes in computational evolutionary models. By taking this approach, I explore competition among organisms with different mutability (maximum mutation size) over flat fitness landscapes, in which all organisms have identical fitness. This allows the results to be interpreted in the context of the neutral theory of evolution, in which there is no selection criterion imposed on the organisms, so mutations occur randomly and evolution occurs by "drift". **I show that the results of competition over many generations on various neutral fitness landscapes result in no clear optimal mutability in general.**

Model

Environment – The axes of a phenospace are two arbitrary phenotypes. An overlaid fitness landscape is set to a single value, so that there is **no** selection criterion, i.e., no explicit Darwinian natural selection. The phenospace and fitness landscape take the shape of a plane with 45 units along each axis.

Organisms – The phenospace is initially populated with 300 uniformly distributed organisms. In addition to being assigned a position within the phenospace, each organism is assigned a mutability, μ, selected with uniform probability from the range $0<\mu<7$.
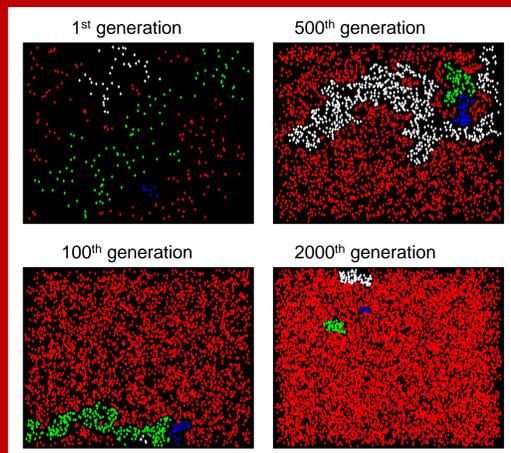

Figure 1: Example of organisms (red) in phenospace. The yellow, black, and blue dots represent three different clusters as determined by reproductive isolation (closed sets of mating organisms).

Mating – Mating between organisms is strictly with the nearest neighbor in phenospace (i.e., an assortative mating scheme). Each organism is selected as a reference organism (RO) once in each generation. The organism mates with its nearest neighbor and produces a number of offspring corresponding to the fitness of the landscape. The offspring are placed according to a uniform random distribution within a region determined by the μ value corresponding to the RO parent, as shown in Figure 2. The mutability μ thus represents the maximum mutation size available to the offspring. The mutability of the RO is assigned to each of its offspring. After all organisms have been an RO, they are removed, leaving only their offspring for the next generation. Organisms born outside of the bounds of the phenospace are removed.
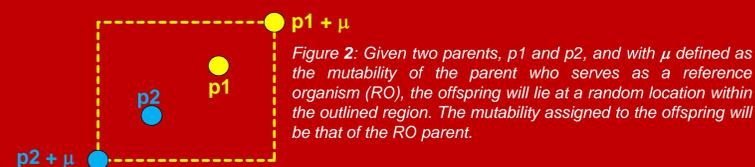

Figure 2: Given two parents, p1 and p2, and with μ defined as the mutability of the parent who serves as a reference organism (RO), the offspring will lie at a random location within the outlined region. The mutability assigned to the offspring will be that of the RO parent.

Overcrowding – After all the organisms have reproduced, any offspring at a distance less than 0.25 units from any other offspring are removed.

Random Removal – In addition to removal based on overcrowding, the system is further randomized by the removal of a random number of organisms (between 0 and 70% of the existing population). The percentage of organisms to be removed is randomly selected at each generation.

Competition Experiments – Each simulation was run until either (1) all values of μ had become extinct except one surviving value, or (2) the simulation had run for 2000 generations. All simulations were performed on PCs using a custom-written program in MATLAB. All random numbers were generated using the pseudo random number generator, Mersenne Twister.

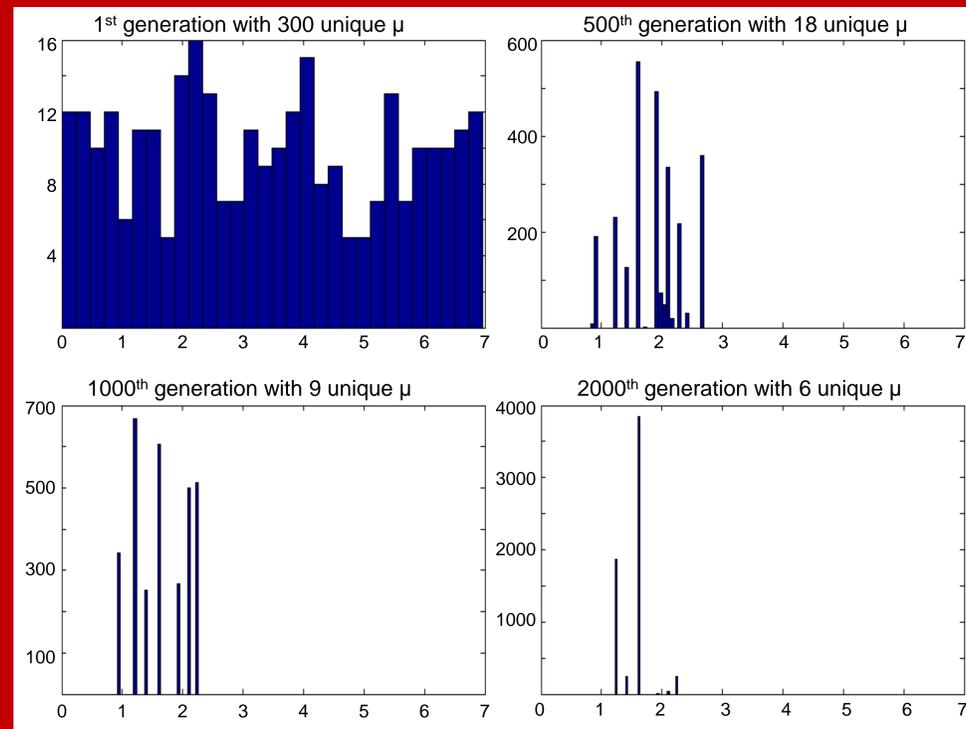For a more detailed description of the algorithms used, see [1].

Figure 3: Examples of population histograms, broken down into μ values. Data is from one simulation with fitness = 2.

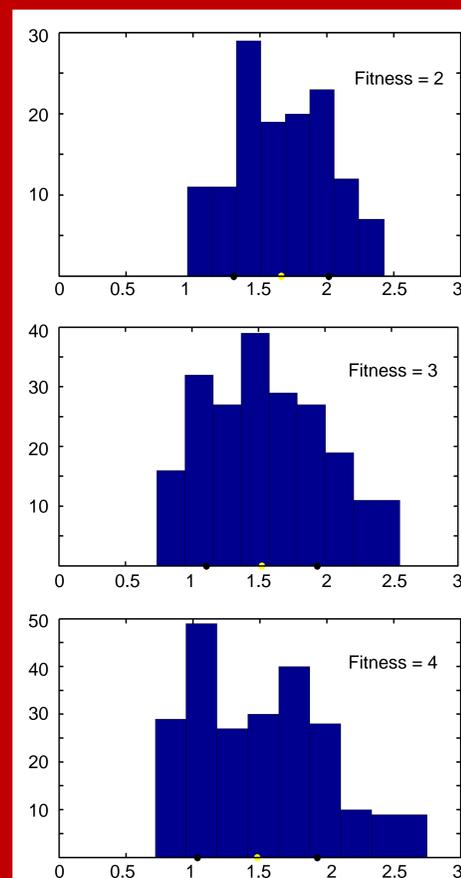| Table 1	Fitness = 2	Fitness = 3	Fitness = 4
Range of surviving μ	0.9570 : 2.4275	0.7364 : 2.5626	0.7178 : 2.7495
$<\mu>$ +/- σ_μ	1.6601 +/- 0.3539	1.5261 +/- 0.4168	1.4808 +/- 0.4474
μ of $<N>_{max}$	~1.55	~1.65	~1.95
$<N>_{max}$	461	625	948
$<N>_{max}$ % of $<N>$	14%	8%	11%

Figure 4: Frequency histograms of unique μ values (shown along the horizontal axis) in the final generations over all 20 simulations for each fitness. Yellow dot represents the average, $<\mu>$. Black dots represent 1 σ_μ from $<\mu>$. 132 μ's are represented for fitness = 2. 122 μ's are represented for fitness = 3. 222 μ's are represented for fitness = 4.
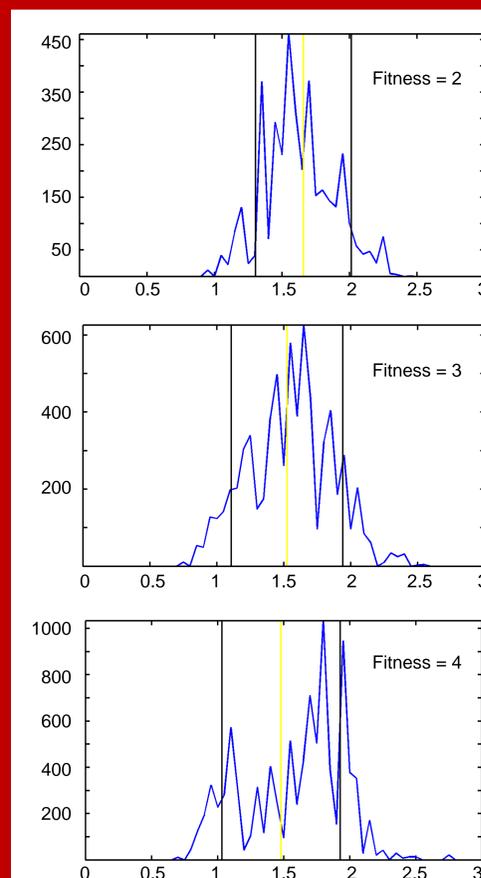

Figure 5: Final generation data for all 20 simulations for each fitness level. The blue curve indicates average population as a function of μ. For each corresponding fitness level, the yellow vertical line is the average μ occurrence from Figure 4, and the black vertical lines are a distance σ_μ from the mean from Figure 4.

Results

Results of the 60 simulations (20 simulations for each level of fitness) are illustrated in the Figures and summarized in Table 1. The mean and standard deviation of remaining μ, $<\mu>$ and σ_μ, are given for the final generation, averaged over all simulations. $<N>$ is the sum of all populations for each generation, averaged over 20 simulations, and $<N>_{max}$ is the maximum population averaged over the same 20 simulations, both of the final generation.

Discussion and Conclusions

Determining "optimality" in the context of neutral theory in this model may lead to different interpretations, since typical evolutionary models use fitness as an optimization criterion [2-3]. Differing characteristics may be optimized according to μ. If the frequency of a mutability value in the final population is used as a measure of optimization (Figure 4), then the μ value distribution for neutral fitness of 2 and 4 appears to be bimodal, indicating two optimal μ. In these cases, two local maximal values occur, with a separation of 2 σ_μ. This simple criterion of peak separation of 2 σ_μ is not satisfied for a neutral fitness landscape of 3, where there is less convincing evidence for possible bimodality. If the criteria for an optimal μ is instead defined by the greatest proportion of the population, then an optimal μ may be found for each level of fitness. However, the significance of the finding may not be sufficient to accept that μ value as optimal. Instead, there may be multiple μ's or, more simply, a range of μ's which are optimal. In either case, there is still no clear winner. The underlying reasons for this in the present model, as well as its possible broader implications and relation to other studies, remains to be investigated in future work.

Competition simulations including only two μ values assigned evenly among the initial population might reveal the strength of competition between various potentially optimal mutability values. Preliminary simulations allow μ values approximately determined by $<\mu>$ +/- σ_μ (which are very near the peaks in Figure 4) to "duel". Initial results (for fitness = 2) show that competition between the two populations of μ is very balanced (Figure 6). This contrasts with the "principle of competitive exclusion" which holds in simple two-dimensional nonlinear dynamics models of species interactions, such as the Lotka-Volterra model (System A), which states that for competition between two species for the same limited resource, one species drives the other to extinction [4]. A potential problem with this, however, is that in the Lotka-Volterra model, direct competition (equivalent overpopulation effect in the model presented here) affects species unequally, whereas overpopulation in the present model affects all organisms equally, since organisms can crowd out members of their own species as well as others.
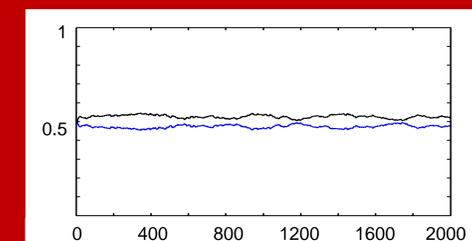

Figure 6: Competition between two μ values. The proportion of the population is shown as a function of generation. Data is averaged over 20 simulations of two competitors taken from $<\mu>$ +/- σ_μ according to Figure 4 with fitness = 2 .

(μ_{blue} = 1.31 & μ_{black} = 2.00).

As suggested by the results shown in Figure 6, it is possible that even in simulations of "dueling" μ populations there is still no winner. Such a result would imply that there are multiple optimal μ's or that a relatively wide range of μ's may be equally "optimal". If this is the case, a mathematical model in which the "principle of competitive exclusion" fails may shed light. A proposed nonlinear system which may describe competitions between two populations of μ may be represented in the equations below (System B). The populations of each μ are x and y, r and s are growth rates of each population, a and b are carrying capacities, and c is an overcrowding effect. Notice that these are modified logistic equations which include a coupling term that depends on the total population. With added noise, this system or a similar system may be studied for coexisting stable populations for each μ. Bimodality might then be explained by parameters in the mathematical system by relating them to constants within the computational model.

$$\frac{dx}{dt} = x(a-x-by)$$
$$\frac{dy}{dt} = y(c-dx-y)$$

System A: Lotka-Volterra model, with parameters a, b, c, and d. Variables x and y represent populations of competing species.

$$\frac{dx}{dt} = rx\left(1 - \frac{x}{a} - c(x+y)\right)$$
$$\frac{dy}{dt} = sy\left(1 - \frac{y}{b} - c(x+y)\right)$$

System B: Proposed model

References

1. Dees ND, Bahar S (2010) Mutation Size Optimizes Speciation in an Evolutionary Model. PLoS ONE 5(8): e11952. doi:10.1371/journal.pone.0011952.
2. Bedau MA, Packard NH (2003) Evolution of evolvability via adaptation of mutation rates. Biosystems 69: 143-169.
3. Clune J, Misevic D, Ofria C, Lenski RE, Elena SF, Sanjuán R (2008) Natural selection fails to optimize mutation rates for long term adaptation on rugged fitness landscapes. PLoS Comput Biol 4(9): e1000187.
4. Strogatz S, Nonlinear Dynamics and Chaos, Perseus Books Publishing, Cambridge, MA, 1994.