Evolution of Evolvability via Adaptation of Mutation Rates

Mark A. Bedau

Reed College, 3213 SE Woodstock Blvd., Portland, OR 97202, USA

`mab@reed.edu`

Norman H. Packard

Prediction Company, 236 Montezuma St, Santa Fe NM 87501, USA

n@predict.com

Abstract

We examine a simple form of the evolution of evolvability—the evolution of mutation rates—in a simple model system. The system is composed of many agents moving, reproducing, and dying in a two-dimensional resource-limited world. We first examine various macroscopic quantities (three types of genetic diversity, a measure of population fitness, and a measure of evolutionary activity) as a function of fixed mutation rates. The results suggest that (i) mutation rate is a control parameter that governs a transition between two qualitatively different phases of evolution, an ordered phase characterized by punctuated equilibria of diversity, and a disordered phase of characterized by noisy fluctuations around an equilibrium diversity, and (ii) the ability of evolution to create adaptive structure is maximized when the mutation rate is just below the transition between these two phases of evolution. We hypothesize that this transition occurs when the demands for evolutionary memory and evolutionary novelty are typically balanced. We next allow the mutation rate itself to evolve, and we observe that evolving mutation rates adapt to values at this transition. Furthermore, the mutation rates adapt up (or down) as the evolutionary demands for novelty (or memory) increase, thus supporting the balance hypothesis.

1

1   Introduction

Evolvability—the capacity for evolution to create adaptations—has received considerable attention recently in both the biological and evolutionary computation communities. The practical benefit for evolutionary computation of understanding and enhancing evolution's adaptive power is obvious, but the theoretical implications of evolvability are also significant. There is general agreement that evolvability is crucial for understanding the orgin of complex adaptations [9, 10, 21, 34, 23] as well as the process of open-ended evolution [32, 26], a central open problem in artificial life [7, 8, 5]. In addition, Wagner and Altenberg [1, 34] argue that evolvability can unify and explain a host of seemingly related issues, including epistasis, genetic canalization, genetic modularity, developmental constraints, developmental and morphological integration, and biological versatility.

Studies of evolvability center on understanding how it originates [21, 1, 34] and how it evolves [33, 9, 10, 1, 34]. This paper treats the second issue. Evolvability has been criticized on the grounds that it involves group selection and it bestows a future rather than present benefit of individuals, but previous discussions have thoroughly defended evolvability's credentials [33, 9, 10, 1, 34, 23, 35, 11] so current research can focus on constructive tasks rather than polemics. Previous accomplishments include verbal and biological arguments for evolvability's feasibility and centrality [9, 10, 23, 35, 11] and on mathematical and computational demonstrations of the evolution of evolvability in simple situations such as infinite populations with a fixed fitness function. The present work takes the next step and considers the evolution of evolvability in a finite population adapting to a dynamic endogenous fitness function.

A system's evolvability depends on its ability to produce adaptive phenotypic variation, and this hinges on both the extent to which the system's phenotype space contains adaptive variation and the ability of evolutionary search to locate it. Two main factors control the effectiveness of the evolutionary search process: the way in which genetic operators traverse genotype space, and the way that genotypes are phenotypically expressed (the

2

genotype-phenotype mapping). For evolutionary search to explore a suitable variety of viable evolutionary pathways, genetic operators must generate sufficient "novelty." At the same time, since evolutionary adaptations are built through successive improvements, the evolutionary process must "remember" incremental improvements. Genetic operators like mutation rate simultaneously affect the demands for both evolutionary novelty and memory, so evolvability requires genetic operators to balance these competing demands successfully. (In the evolutionary computation literature this is known as the tension between "exploration" and "exploitation" [18].) Furthermore, the appropriate balance between evolutionary novelty and memory can continually change as evolution progresses, so evolvability requires genetic mechanisms to adjust the balance flexibly. Thus, understanding evolvability involves addressing the following questions:

- Where are evolutionary novelty and memory balanced?

- How does this balance shift during the course of evolution?

- Can the genetic mechanisms that structure evolution flexibly adapt to this shifting balance?

A simple way to address these questions is to let second-order evolution control the genetic mechanisms that structure first-order evolution—thus studying the evolution of evolvability. And the simplest genetic mechanism that structures the variety of evolutionary alternatives is the mutation rate.

Optimal and evolving mutation rates are discussed in both the biological and evolutionary computation literatures. A fundamental open question in evolutionary biology is whether mutation rates are as low as physically possible or whether they are at an optimal positive value [25]. The premise that mutations are generally harmful has been used to argue that natural selection of mutation rates can only go in one possible direction: toward zero (e.g., see [36]). On the other hand, mathematical analyses of certain one-locus, two-allele models with a separate modifier (mutation rate) locus have shown that evolution yields optimal, positive mutation rates under some sets of conditions (see, e.g., [17, 16, 19, 24]). But it remains unclear what, if anything, such results might reveal about the evolution of mutation rates in a finite population with many loci and many alleles per loci, and when the context of evolution continually changes and thus the fitness function unpredictably varies.

The discussion of evolving mutation rates in the evolutionary computation literature mainly focuses on the issue of automated control of evolution, in the context of efforts to minimize the time required to solve function optimization problems. Preliminary efforts were devoted to proving that automated control is feasible (e.g., [15, 2]), and current research in this area is proceeding along several fronts [12]. Although the potential technological value of this work is obvious, it is unclear how, if at all, it will illuminate the theoretical questions about the evolution of evolvability. In particular, all the above-cited work presumes that evolution is driven by a fixed and externally-specified fitness function, whereas the most interesting theoretical issues concern evolution with implicitly-specified fitness functions which continually change in the course of evolution [30]. These theoretical concerns have been addressed more directly in a study of evolving mutation rates in a host–parasite model [20], but explicit fitness functions still play some role in that work.

We aim to address some of the fundamental theoretical questions concerning the evolution of evolvability in the context of adaptation of mutation rates. Although simple, our model has rich evolutionary potential, with a finite population containing many loci and many alleles per loci, and an implicit fitness function that can continually change as the population evolves. After describing the model in section 2, we investigate how the model behaves as a function of different *fixed* mutation rates in section 3, by observing genetic diversity (section 3.1), fitness (section 3.2), and evolutionary activity (section 3.3). Those observations suggest that evolution's ability to create useful adaptive structure is maximized at transitional mutation rates that separate two phases of evolution characterized by genetic order and genetic disorder. We hypothesize (section 3.4) that these mutation rates are critical for the system's evolvability because they balance the competing demands for evolutionary novelty and memory. When we let the mutation rates evolve in section 4, we see that they evolve to the critical values at which evolvability is maximized, so we test the hypothesis that this adaptation of mutation rates is driven by the balance between novelty and memory by engineering circumstances that change the relative value of novelty and memory.

2   A Simple Model of Sensorimotor Evolution

Our evolutionary system is composed of many agents that could be called organisms, on analogy with biological systems. It has been used in various studies of the evolution of sensory-motor functionality [30, 6, 4, 3, 7, 31]. The system consists of many agents that exist together in an environment, in this case a toroidal lattice. The lattice has a real-valued field defined on it, $\mathcal{E}^t(x)$, which may be interpreted as an resource field. The resource field is driven by periodically adding resources (from an external source), $\mathcal{E}^{t+\Delta t}(x) = \mathcal{E}^t(x) + \Delta(x - \xi^t)$, where Δ is a function over the plane and ξ^t is a position in the plane. The agents are constantly gathering resources, extracting them at their location in the field and then expending them through their behavior. Thus, the agents function as the system's resource sinks, and the whole system is dissipative.

In analogy with biological systems, the dynamics of the population as a whole is comprised of all the birth-life-death cycles of the agents. Births occurs when agents accumulate enough resources to reproduce (see below), deaths occur when agents run out of resources, and the lives of agents consist of their interactions the environment. We label each agent with the index i, let I^t be the set of agents existing at t. Time is discrete. One unit of time t is marked by each agent interacting with the environment.

During its lifetime, each agent exchanges information with the environment by sensing and affecting the resource field in its local neighborhood. We assume that there is a discrete set of different possible sensory states, $s \in S = \{s_1, ..., s_{N_S}\}$. The agents exchange no information with each other directly (although this would be an easy generalization). Each agent (labeled with the index i) has certain information associated with it: (i) a current location, x_i^t, (ii) a current sensory state, s_i^t, (iii) a current reservoir of resources, E_i^t, and (iv) a sensory-motor map, φ_i, that yields a behavior given sensory information as input. In this model, an agent's behavior is a vector $\vec{b}$ denoting the agent's movement in the environment, $x_i^t \to x_i^{t+1} = x_i^t + \vec{b}_i^t$; in general, $\vec{b}_i^t = \varphi_i(s_i^t)$ (an exception is explained below). We assume that $\vec{b}_i^t$ is a member of a discrete set of different possible behaviors, $\vec{b}_i^t \in B = \{\vec{b}_1, ..., \vec{b}_{N_B}\}$. (Another easy generalization would be to include other kinds of possible behaviors.) Each behavior causes a change in the agent's supply of resources: an augmentation from extracting resources at the agent's new location, a

5

constant-sized reduction, and a reduction proportional to the distance moved, $E_i^t \rightarrow E_i^{t+1} = E_i^t + \alpha(\mathcal{E}^{t+1}(x_i^{t+1})) - \beta - \gamma(\vec{b}_i^t)$. Here, $\alpha(z)$ may be interpreted as the resource extraction function, β as the constant metabolic cost of surviving, and $\gamma(z)$ as a function for the variable metabolic cost of a specific behavior.

The sensory-motor map φ_i operating on a sensory input s_i^t has a particularly simple form because the sensory input is discrete. Since $s_i^t \in \{s_1, \ldots, s_{N_S}\}$, we may identify the function φ_i with its graph, a set of N_S behavior values, $\{\varphi_{is}\}$. Pursuing the biological analogy, we will consider the sensory-motor strategy elements $\{\varphi_{is}\}$ as i's *genome*, with N_S loci, and each particular element $\vec{b} = \varphi_{is}$ as the trait (or allele) at the sth locus of i's genome.

Reproduction occurs when an agent's resource supply exceeds a threshold, E_r. The parent splits its resources with its child, and the child inherits its parent's strategy elements $\{\varphi_{is}\}$, except for mutations. The genome of each agent i contains one special gene, μ_i—the rate at which i's strategy elements mutate when i reproduces, i.e., the probability that a strategy element of i's children is chosen (with equal probability) from the set of possible behaviors B. We also introduce a *meta-mutation* rate parameter, μ_μ—the probability that i's children's mutation-rate gene is chosen (with equal probability) from the interval $[\mu_i - \epsilon, \mu_i + \epsilon]$. The value of μ_μ is fixed during the course of a given simulation, and the value of μ_i is fixed during the course of i's lifetime (as are all of i's genes).

In order to investigate how natural selection affects the evolutionary dynamics of behavioral strategies and mutation rates, we introduce a *behavioral noise* parameter, B_0, defined as the probability that $\vec{b}_i^t$ is chosen at random from B rather than determined by $\varphi_i(s_i^t)$. If $B_0 = 1$, then agents survive and reproduce differentially, and children inherit their parents' strategy elements (except for mutations), but the inherited strategies $\{\varphi_{is}\}$ reflect only random genetic drift.

This model provides a simple setting for empirical study of the evolution of evolvability. Agents' immediate environments produce sensory states, which then trigger actions by means of the agents' sensory-motor maps, and these actions subsequently change the environment. In this way, the agents' sensory-motor maps influence the conditions of their own evolution. Since the agents' survival and reproduction depends directly on their ability to contin-

ually find resources in their environment, the implicit fitness functions in this model are constantly buffeted by the contingencies of natural selection and, thus, unpredictably change. This first-order evolution is structured by the sensory-motor maps actually compared and tested by natural selection. One especially simple mechanism that regulates the variety of maps available for evolutionary exploration is the mutation rate; the higher the mutation rates, the greater the variety. Thus, by allowing mutation-rate genes to evolve, we can study second-order evolution of evolvability.

Appendix A contains the values of all parameters for the empirical results reported here.

3   Phases of Evolution

We first studied the present model *without* evolving mutation rates. In this case, a global mutation parameter μ is applied to all agents, and reproduction copies a parent's strategy elements $\{\varphi_{is}\}$ to its child with probability μ of being chosen randomly from the set of possible strategy elements B. We measured several "macroscopic" quantities, where macroscopic means an average is performed over all agents or all sites in the world:

Diversity. Three different versions of diversity are measured, each of which is a different variance of genetic distributions in the population of agents.

Fitness. Implicitly, the goal of agents is to accumulate resources. Thus, the overall fitness of the population is inversely reflected by the amount of unconsumed resource in the world.

Evolutionary Activity. This quantity is derived from usage statistics, which reflect how well natural selection has tested the traits that persist in the population. If components are well-tested on an evolutionary time scale, evolutionary activity is positive.

These macrovariables are each defined in the following subsections. All of these quantities are measured as a function of mutation rate, which we set as a fixed parameter of the system for a given experimental run over which a measurement is made.

3.1 Diversity

We define three types of diversity for a population of agents. Recall that $\vec{b} = \varphi_{is}$ is the sth allele of the ith agent, which is a two dimensional vector representing the behavior of i when stimulated by s. $\|z\|$ is the two dimensional Euclidean norm, i.e., length, of the vector z.

Within-gene diversity. The variance of each allele value at a given locus, averaged over all loci:

$$
\begin{aligned}
W^t &= \langle \mathrm{Var}_i(\varphi_{is}) \rangle_s^t \\
&= \frac{1}{N_{I^t} N_S} \sum_{i \in I^t} \sum_{s \in S} \|\varphi_{is} - \langle \varphi_{is} \rangle_i^t\|^2,
\end{aligned}
\tag{1}
$$

where $\langle \varphi_{is} \rangle_i^t = 1/N_{I^t} \sum_{i \in I^t} \varphi_{is}$ is the average at t over agents of the value of the sth allele.

Between-gene diversity. The variance of the average allele values at each locus:

$$
\begin{aligned}
B^t &= \mathrm{Var}_s(\langle \varphi_{is} \rangle_i^t) \\
&= \frac{1}{N_S} \sum_{s \in S} \|\langle \varphi_{is} \rangle_i^t - \langle \varphi_{is} \rangle_{is}^t\|^2,
\end{aligned}
\tag{2}
$$

where $\langle \varphi_{is} \rangle_{is}^t = 1/(N_{I^t} N_S) \sum_{i \in I^t} \sum_{s \in S} \varphi_{is}$ is the average at t of all the alleles at all the loci of all the agents.

Total diversity. The variance of each allele value at each locus:

$$
\begin{aligned}
D^t &= \mathrm{Var}_{is}^t(\varphi_{is}) \\
&= \frac{1}{N_{I^t} N_S} \sum_{i \in I^t} \sum_{s \in S} \|\varphi_{is} - \langle \varphi_{is} \rangle_{is}^t\|^2.
\end{aligned}
\tag{3}
$$

An easy calculation shows that $D^t = W^t + B^t$.

The most notable aspect of the empirically observed diversity dynamics is that, as mutation rate is varied, there is a transition from punctuated equilibria dynamics to random-looking fluctuations. Typical diversity time series at different mutation rates are shown in Figure 1 without selection and

in Figure 2 with selection. We see the same pattern in both cases: At very low mutation rates diversity exhibits punctuated equilibrium dynamics, and at very high mutation rates diversity dynamics are random-looking fluctuations. Intermediate mutation rates show more complicated fluctuations. (Further details about these different diversity dynamics are found elsewhere [4, 3].) Furthermore, total diversity D is dominated by within-gene diversity W at low mutation rates, and by between-gene diversity at high mutation rates B, as the correlations among the diversity dynamics shows (Figure 3). To further resolve the relationship among the different diversity components, we define:

$$\frac{C^t = 4W^t B^t}{(D^t)^2} \tag{4}$$

The quantity C^t reflects the extent to which the total diversity D^t is not dominated by either of its components W^t or B^t. (The factor of 4 scales C^t between 0 and 1.) The transition between the two phases of evolution is evident in the time series of C^t (Figure 4). Examining how the time average $C = \langle C^t \rangle_t$ depends on mutation rate (Figure 5) shows that the transition between phases occurs when the mutation rate is in the range $10^{-3} \leq \mu \leq 10^{-2}$. The transition occurs both when selection is acting and when it is absent, and it seems sharper without selection.

3.2 Fitness

The fact that our model is resource-driven allows us to consider the overall fitness of the population in terms of its efficiency at extracting resources from their environment. A crude measure of this efficiency is the residual resource in the environment $\mathcal{E}^t_{resid}$, i.e., the resource that is not contained in the agents. Using the notation described in section 2,

$$\mathcal{E}^t_{resid} = \sum_x \mathcal{E}^t(x). \tag{5}$$

In Figure 6 we show the time average $\mathcal{E}_{resid} = \langle \mathcal{E}^t_{resid} \rangle_t$ of residual resource as a function of mutation rate. Note that the no selection case shows $\mathcal{E}_{resid}$ to be relatively constant; this is expected, as in this case no evolutionary learning is taking place. Note also that when selection is present population performance is maximized at mutation values around the transition, $10^{-3} \leq \mu \leq 10^{-2}$.

9

3.3 Evolutionary Activity

The computation of evolutionary activity is based on measuring the adaptive significance displayed by each trait (where, in this study, each strategy entry $\vec{b} = \varphi_{is}$ is considered a trait for the ith agent). We consider a trait's adaptive significance to be the positive contribution it has made to the ongoing existence of its host agent (further discussion of this suggestion occurs elsewhere [6, 7, 8, 31].) We judge such a contribution to be indicated by the degree to which the trait has been well tested by natural selection, with the trait persisting through repeated use and, in particular, accumulating more usage than would be expected *a priori*. Measuring a trait's adaptive significance, then, requires distinguishing its actual usage from its usage weighted by the extent to which it exceeds *a priori* expectations.

In order to measure "raw" usage of the various traits, we simply assign a usage variable u_{is}^t to each trait $\vec{b} = \varphi_{is}$ of each agent. The usage, u_{is}^t, is initially zero when an element enters the population through mutation (or at the very beginning of the simulation), and is incremented every time the element is actually used, i.e., when the agent receives the sensory input s and thus performs the behavior $\vec{b} = \varphi_{is}$:

$$u_{is}^{t+1} = u_{is}^t + \delta(s_i^{t+1} - s), \tag{6}$$

where $\delta(x)$ is the Dirac delta function, which is one if $x = 0$ and zero otherwise. We add the proviso that if $B_0 > 0$ and behavioral noise prevents $\varphi_i(s_i^{t+1})$ from determing $\vec{b}_i^{t+1}$, then the above rule for usage incrementing does not apply and $u_{is}^{t+1} = u_{is}^t$. Note that if $B_0 = 1$ then $u_{is}^t = 0$ for all i, s, and t.

We must expect some raw usage to be insignificant, however, i.e., to be usage that accrues by random triggering of a trait even though the trait lacks selective value. To eliminate the contribution of this insignificant usage, we modify the above algorithm and weight raw usage increments by a function W:

$$\tilde{u}_{is}^{t+1} = \tilde{u}_{is}^t + W\delta(s_i^{t+1} - s). \tag{7}$$

W can depend on a variety of factors, such as the state of the agent, the state of the trait being activated, and the mutation rate μ.

For the experiments reported here, W depends on the age A_{is} of the sth trait of the ith agent, and on the *drift duration*, t_μ. The drift duration

t_μ is computed by observing age distributions of traits for the no-adaptive-evolution case ($B_0 = 1$); these distributions have an exponential fall-off that depends on the mutation rate μ. The drift duration is defined as the age at which the frequency of observed ages goes to zero for the population sizes of the current experiment. Then we set $W = \theta(A_{is} - t_\mu)$, where the age A_{is} of the sth trait of the ith agent is defined as the number of times steps since φ_{is} was originally introduced by mutation at the sth locus in i's genetic lineage, and $\theta(x)$ is the Heaviside theta function, zero for $x < 0$ and one otherwise. Simpler weighting functions have also been used [6], and other schemes for screening off insignificant usage have been pursued in related work [7, 8, 31].

Evolutionary activity is simply the sum of the weighted usage:

$$A^t = \sum_{i \in I^t} \sum_{s \in S} \tilde{u}_{is}^t. \tag{8}$$

Figure 7 shows how the time average $A = \langle A^t \rangle_t$ of evolutionary activity depends on mutation rate. We see that the activity appears to follow the law $A \propto \mu^\alpha$, with $\alpha \approx -2.3 \pm 0.3$. Limitations of computational resources currently prevent us from exploring how activity varies at lower mutation rates (the drift duration becomes prohibitively long). Nevertheless, over the mutation spectrum that we can observe, we see that evolutionary activity steadily increases as the mutation rate approaches the transitional region $10^{-3} \leq \mu \leq 10^{-2}$.

3.4 Balancing Evolutionary Novelty and Memory

There are three main observations to be gleaned from our measurements of diversity, fitness, and evolutionary activity. First, diversity measurements indicate that there is a transition with distinctly different diversity dynamics for low mutation rates and high mutation rates. So far, because of the lack of measurement precision, it is not clear how sharp the transition is; there might even be a range of mutation values over which the transition takes place. Second, measurement of residual resource as a function of mutation rate shows a minimum that coincides with the transition region. Third, evolutionary activity increases as the mutation rate approaches the transition region from above. The basic picture painted by these results has two parts:

- The mutation rate is governs a transition between two qualitatively different phases of evolutionary dynamics. When the mutation rate

is significantly below the phase transition, the whole gene pool tends to remain frozen at a given strategy; the gene pool has a tremendous amount of redundancy and is highly ordered. When the mutation rate is significantly above the phase transition, the gene pool tends to be a continually changing plethora of randomly related strategies; it is highly disordered. The phase transition itself occurs in a characteristic mutation rate region.

- The ability of evolution to create useful adaptive structure is maximized at or just below the mutation rates that transition into the genetically disordered systems.

The upshot of these two effects is that evolutionary adaptation tends to be maximized when the gene pool is "at the edge of genetic disorder." These results are quite reminiscent of the error threshold from molecular evolution— the critical mutation rate at which evolutionary adaptations are destroyed more quickly than natural selection can produce them [13, 14]. It is still an open question to what extent our results might be a reflection of an error threshold, for our results are obtained in a significantly more complicated context than that in which the error threshold has been demonstrated, and our bifurcation of diversity dynamics occurs whether or not selection is present.

We suggest that these results can be explained by considering the competing demands for evolutionary "novelty" and "memory" (also known in the evolutionary computation literature as the tension between "exploration" and "exploitation"). Evolution's ability to build adaptive structure depends on the accessibility of viable evolutionary alternatives; the right range of accessible alternatives can make evolution easy, while too many or too few can make adaptation difficult or even impossible. In the present model, the population can evolve better sensory-motor strategies only if it can test enough novel strategies (its need for evolutionary "novelty"). At the same time, the populations sensory-motor strategies can adapt to a given environment only if strategies that prove beneficial can persist in the gene pool (its need for evolutionary "memory"). The mutation rate simultaneously affects both. The lower the mutation rate, the greater the number of genetic strategies remembered from parents; the higher the mutation rate, the greater the number of novel genetic strategies introduced with children. Too much mutation (not enough memory) will continually flood the population with

12

new random strategies; too little mutation (not enough novelty) will tend to lock the population at arbitrary strategies. Successful adaptive evolution requires a mutation rate suitably intermediate between these extremes, where the competing demands for memory and novelty are suitably balanced. Our hypothesis is that mutation rates around the transition optimally balance the competing evolutionary demands for novelty (high mutation) and memory (low mutation).[1] The balance between evolutionary novelty and memory that is critical for evolvability is evidently "at the edge of genetic disorder."

4 Evolution of Mutation Rates

The previous results show that evolvability in our model system is greatest when mutation rates are in the range $10^{-3} \leq \mu \leq 10^{-2}$. This naturally raises the question whether second-order meta-evolution can adapt the system so that its mutation rate is in this critical region. We pursued this question by extensively simulating evolving mutation rates, holding all model parameters constant except for systematically varying the meta-mutation rate, μ_μ, and switching adaptive evolution on and off ($B_0 = 0$ and $B_0 = 1$). In the initial population sensory-motor genes were assigned randomly and initial mutation-rate genes were either assigned randomly or all set to an arbitrary value. We then observed various "macroscopic" properties of the model, including the distribution of mutation rates,

$$M^t(\mu) = \sum_{i \in I^t} \delta(\mu_i - \mu). \tag{9}$$

Figure 8 shows that, as one could expect, when the agents' sensory-motor genes merely drift ($B_0 = 1$), the mutation distribution drifts up and down, with the width of the distribution and the rate of change being proportional to μ_μ. (Note that $M^t(\mu)$ does not spread out indefinitely because the population is continually truncated due to the random contingencies of finding resources.) On the other hand, when adaptive evolution happens ($B_0 = 0$), the mutation distribution eventually becomes clustered at low mutation rates

---

[1] If our transition between genetic order and disorder is a manifestation of the error threshold, then our results provide independent evidence to support the argument of Ochoa et al. [28] that the error threshold explains optimal mutation rates in evolutionary algorithms.

(Figure 9). As one can see, the width of the distribution and the rate at which the distribution falls are typically proportional to μ_μ, which is to be expected.

Figure 10 presents a range of data from a typical simulation with adaptation of sensory-motor and mutation-rate genes. When we resolve the low end of the mutation distribution in exponentially-smaller-sized bins, we can clearly see second-order evolution pushing the bulk of the mutation distribution into the bin (third from the top) collecting mutation rates in the range $10^{-3} < \mu_i \leq 10^{-2}$, which exactly corresponds to the previously observed transition. Population mean mutation rates corroborate this conclusion. These effects are robust across all kinds of initial conditions, even those which are so extreme as to effectively prevent adaptive evolution.

In addition, Figure 10 shows a marked increase in population fitness (i.e., decrease in residual resource) corresponding to the drop in mean mutation rate—clear evidence that the mutation rate drop is associated with increasing adaptation of sensory-motor maps. Note that this fitness increase is not evidently associated with a marked increase in population level; rather, the agents are evidently becoming much more efficient at extracting available resources from their environment.

These results further the case for the evolutionary significance of the transition. More important, we now have present clear evidence that second-order evolution can adapt mutation rates to the transition. In this sense, effective first-order evolution can emerge from second-order evolution, i.e., evolvability can evolve.

If we combine this conclusion with our earlier hypothesis that transitional mutation rates optimally balance the competing demands for novelty and memory, we would predict that the mutation distribution can be *raised* temporally by suddenly increasing the demands for evolutionary novelty. To test this hypothesis, we occasionally externally perturbed the model after first- and second-order evolution had substantially progressed, forcing evolutionary adaptation to start from scratch. For example, one kind of perturbation we employed was to randomize each agents' sensory-motor genes but leave their mutation-rate genes untouched.

These perturbation data confirm our prediction. As soon as the external perturbation shifts the balance toward the need for more evolutionary novelty, we observe the striking sequence of events illustrated in Figure 11: (a) the residual resource in the environment sharply rises, showing that the

14

population fitness has precipitously dropped;[2] (b) immediately after the fitness drop the mean mutation rate dramatically rises as the mutation rate distribution shifts upward; (c) by the time that the mean mutation rate has risen to its highest point the drop in residual resource level shows that the populations fitness has substantially improved, so the balance shifts toward the need for more evolutionary memory; (d) the fitness levels and mutation rates eventually return to their previous equilibrium levels. As Figure 11 illustrates, population level fluctuations play no evident role in these effects.

These results show the mutation rate distribution shifting up and down just as the balance hypothesis predicts. A change in the context for evolution can increase the need for rapid exploration of a wide variety of sensory-motor strategies and thus dramatically shift the balance toward the need for novelty. Then, subsequent sensory-motor evolution can reshape the context for evolution in such a way that the balance shifts back toward the need for memory.

The underlying regularity behind these results is that mutation rates evolve in such a way that evolutionary novelty and memory remain suitably balanced as required for maximal evolvability. The balance point is typically at "the edge of genetic disorder," but if the contingencies of evolution shift the point of balance elsewhere then the mutation rates adapt appropriately.[3]

5   Conclusion

Evolvability requires that the competing demands for evolutionary novelty and memory are suitably balanced. Our investigation of a simple evolutionary model with finite population and endogenous fitness function suggests that this balance point is located along the mutation spectrum at a transition separating two phases of evolution: genetic order and genetic disorder.

---

[2]Notice that the residual resource levels in this simulation are significantly lower than those in Figure 10. This is apparently due to the fact that the simulation shown in Figure 11 used a significantly smaller genome ($N_S = 32$ vs. $N_S = 1024$). Evidently, on these time scales at certain regions of parameter space, fitness can become significantly higher if the task of evolutionary adaptation is simplified through smaller genomes. The adaptation of mutation rates after perturbation happened for all size genomes studied.

[3]The similarity between this result and the famous—and controversial—claims about "adaptation to the edge of chaos" [29, 22, 27] is obvious, but it is an open question whether this similarity is more than superficial.

Furthermore, when we allow the mutation rates in this model to themselves adapt, the second-order evolution adapts the mutation rates as the demands for novelty and memory unpredictably shift. We thus have a concrete illustration of how evolution automatically and flexibly tunes the genetic mechanisms as the requirements for evolvability change.

Acknowledgements

References

[1] Altenberg, L. 1995. Genome growth and the evolution of the genotype-phenotype map. In: W. Banzhaf and F. H. Eeckman (Eds.), *Evolution and Biocomputation: Computational Models of Evolution*. Springer, Berlin, pp. 205–2159.

[2] Bäch, T. 1992. Self-adaptation in genetic algorithms. In: F. J. Varela and P. Bourgine (Eds.), *Towards a Practice of Autonomous Systems*. MIT Press, Cambridge, pp. 263–271.

[3] Bedau, M. A. 1995. Three illustrations of artificial life's working hypothesis. In: W. Banzhaf and F. H. Eeckman (Eds.), *Evolution and Biocomputation: Computational Models of Evolution*. Springer, Berlin, pp. 53–68.

[4] Bedau, M. A., and Bahm, A. 1994. Bifurcation structure in diversity dynamics. In R. Brooks and P. Maes, eds., *Artificial Life IV*. MIT Press, Cambridge, pp. 258–268.

[5] Bedau, M. A., McCaskill, J. S., Packard, N. H., Rasmussen, S., Adami, C. Green, D. G., Ikegami, T., Kaneko, K., Ray, T. S. 2000. Open problems in artificial life. *Artificial Life* 6, 363–376.

[6] Bedau, M. A. and Packard, N. H. 1992. Measurement of evolutionary activity, teleology, and life. In: C. G. Langton, C. E. Taylor, J. D. Farmer, and S. Rasmussen (Eds.), *Artificial Life II*, Addison-Wesley, Redwood City, CA, pp. 431–461.

[7] Bedau, M. A., Snyder, E., Brown, C. T., and Packard, N. H. 1997. A comparison of evolutionary activity in artificial evolving systems and the biosphere. In: P. Husbands and I. Harvey, (Eds.), *Fourth European Conference on Artificial Life*. MIT Press, Cambridge, pp. 125-134.

[8] Bedau, M. A., Snyder, E., and Packard, N. H. 1998. A classification of long-term evolutionary dynamics. In: C. Adami, R. Belew, H. Kitano, and C. Taylor, (Eds.), *Artificial Life VI*. MIT Press, Cambridge, pp. 228-237.

[9] Conrad, M. 1982. Natural selection and the evolution of neutralism. *BioSystems* 15, 83–85.

[10] Conrad, M. 1990. The geometry of evolution. *BioSystems* 24, 61–81.

[11] Dawkins, R. 1989. The evolution of evolvability. In: C. G. Langton (Ed.), *Artificial Life*. Addison-Wesley, Redwood City, pp. 201–220.

[12] Eiben, A. E., Hinterding, R., Michalewicz, Z. 1999. Parameter control in evolutionary algorithms. *IEEE Transactions on Evolutionary Computation* 3 (2), 124–141.

[13] Eigen, M. 1971. Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 58, 465–523.

[14] Eigen. M., McCaskill, J., and Schuster, P. 1979. Molecular quasi-species. *J. Phys. Chem.* 92, 6881-6891.

[15] Fogel, D. B., Fogel, L. J., and Atmar, J. W. 1991. Meta-evolutionary programming. In: R. R. Chen (Ed.), *Proceedings of the 25th Asilomar Conference on Signals, Systems and Computers*. Maple Press, San Jose, pp. 540–545.

[16] Gillespie, J. H. 1981. Evolution of the mutation rate at a heterotic locus. *Proceedings of the National Academy of Science, U.S.A.* 78, 2452–2454.

[17] Gillespie, J. H. 1991. Mutation modification in a random environment. *Evolution* 35, 468–476.

[18] Holland. J. H. 1975. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence.* University of Michigan Press, Ann Arbor. (Second Edition: MIT Press, 1992.)

[19] Holsinger, K. E., and Feldman, M. W. 1991. Modifiers of mutation rate: evolutionary optimum with complete selfing. *Proceedings of the National Academy of Science, U.S.A.* 80, 6732–6734.

[20] Kaneko, K., and Ikegami, T. 1992. Homeochaos: dynamic stability of a symbiotic network with population dynamics and evolving mutation rates. *Physica D* 56, 406–429.

[21] Kauffman, S. A. 1990. Requirements for evolvability in complex systems: orderly dynamics and frozen components. *Physica D* 42, 135–152.

[22] Kauffman, S. A. 1993. *The Origins of Order: Self-Organization and Selection in Evolution.* Oxford University Press, New York.

[23] Kirschner, M., and Gerhart, J. 1998. Evolvability. *Proc. Natl. Acad. Sci USA* 95, 8420–8427.

[24] Liberman, U., and Feldman, M. W. 1986. Modifiers of mutation rate: a general reduction principle. *Theoretical Population Biology* 30, 125–142.

[25] Maynard Smith, J. 1989. The limitations of evolutionary theory. In: J. Maynard Smith, *Did Darwin Get It Right?*. Chapman and Hall, New York, pp. 180–191.

[26] McMullin, B. 2000. John von Neumann and the evolutionary growth of complexity: looking backwards, looking forwards... *Artificial Life* 6, 347-361.

[27] Mitchell, M., Hraber, P. T., and Crutchfield, J. P. 1993. Revisiting the edge of chaos: evolving cellular automata to perform computations. *Complex Systems* 7, 89–130.

[28] Ochoa, G., Harvey, I., and Buxton, H. 1999. Error thresholds and their relation to optimal mutation rates. In: D. Floreano, J.-D. Nicoud, F. Mondada, (Eds.), *Advances in Artificial Life*. Springer-Verlag, Heidelberg, pp. 54–63.

[29] Packard, N. H. 1988. Adaptation toward the edge of chaos. In: Kelso, J. A. S., Mandell, A. J., and Shlesinger, M. F. (Eds.), *Dynamic Patterns in Complex Systems*. World Scientific, Singapore.

[30] Packard, N. H. 1989. Intrinsic adaptation in a simple model for evolution. In: C. G. Langton (Ed.), *Artificial Life*. Addison-Wesley, Redwood City, pp. 141–155.

[31] Rechtsteiner, A., and Bedau, M. A. 1999. A generic model for quantitative comparison of genotypic evolutionary activity. In: D. Floreano, J.-D. Nicoud, F. Mondada, (Eds.), *Advances in Artificial Life*. Springer-Verlag, Heidelberg, pp. 109-118.

[32] Taylor, T. 1999. On self-reproduction and evolvability. In: D. Floreano, J.-D. Nicoud, F. Mondada, (Eds.), *Advances in Artificial Life*. Springer-Verlag, Heidelberg, pp. 94–103.

[33] Wagner, G. P. 1981. Feedback selection and the evolution of modifiers. *Acta Biotheoretica* 30, 79–102.

[34] Wagner, G. P., and Altenberg, L. 1996. Complex adaptations and the evolution of evolvability. *Evolution* 50, 967–976.

[35] West-Eberhard, M. J. 1998. Evolution in the light of development and cell biology, and *vice versa*. *Proc. Natl. Acad. Sci. USA* 95, 8417–8419.

[36] Williams, G. C. 1966. *Adaptation and Natural Selection*. Princeton University Press, Princeton, N.J.

A   Parameter Values for the Simulations

The world has 128 sites on each side, and periodic boundary conditions. Individuals are updated sequentially, in an arbitrary order; no effort is made for parallel updating of all agents. Δ produces a pyramidal energy pile; $\Delta(\xi)$

19

has a peak of 255 at ξ and falls linearly to 0 at a distance of 8 sites from ξ, where ξ^t is chosen at random from the world sites. Each of the four nearest neighbors and the current site produces two bits of sensory information, yielding 10 bits total, or $N_S = 1024$. $\vec{b}$ is a displacement from zero to fifteen steps in one of the eight compass directions, so $N_B = 128$. $\alpha(\mathcal{E}(x))$ is a function returning all of $\mathcal{E}(x)$ up to a threshold $\mathcal{E}_m = 100$, $\beta = 20$, and $\gamma(\vec{b})$ is a function returning the magnitude of $\vec{b}$. $E_r = 500$. In the evolution of mutation experiment, the interval from which new mutation rates are chosen is truncated at 0 or 1 if $\mu - \epsilon < 0$ or $\mu + \epsilon > 1$; $\epsilon = .0025$.
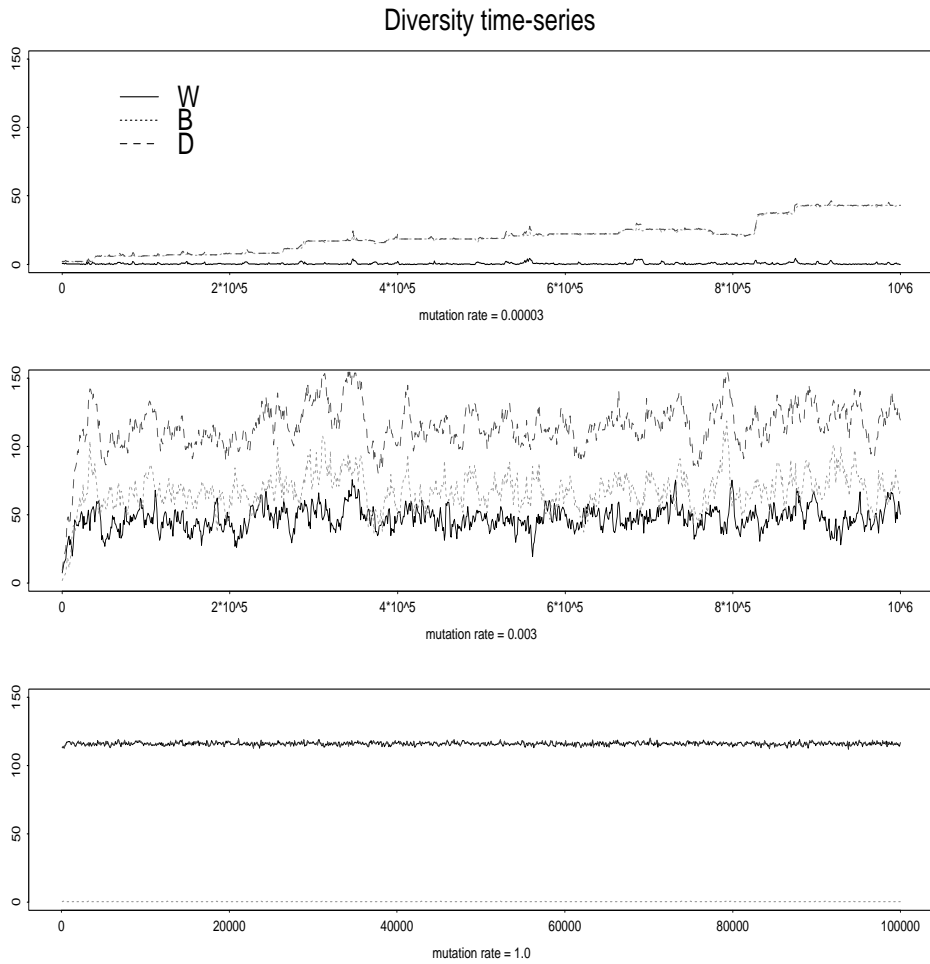
Figure 1: Time series of W^t, B^t, and D^t for three different mutation rates, illustrating a transition in diversity dynamics. Above: $\mu = 3 \times 10^{-5}$. Middle: $\mu = 3 \times 10^{-3}$. Below: $\mu = 1.0$ These measurements were made without adaptive evolution ($B_0 = 1$).
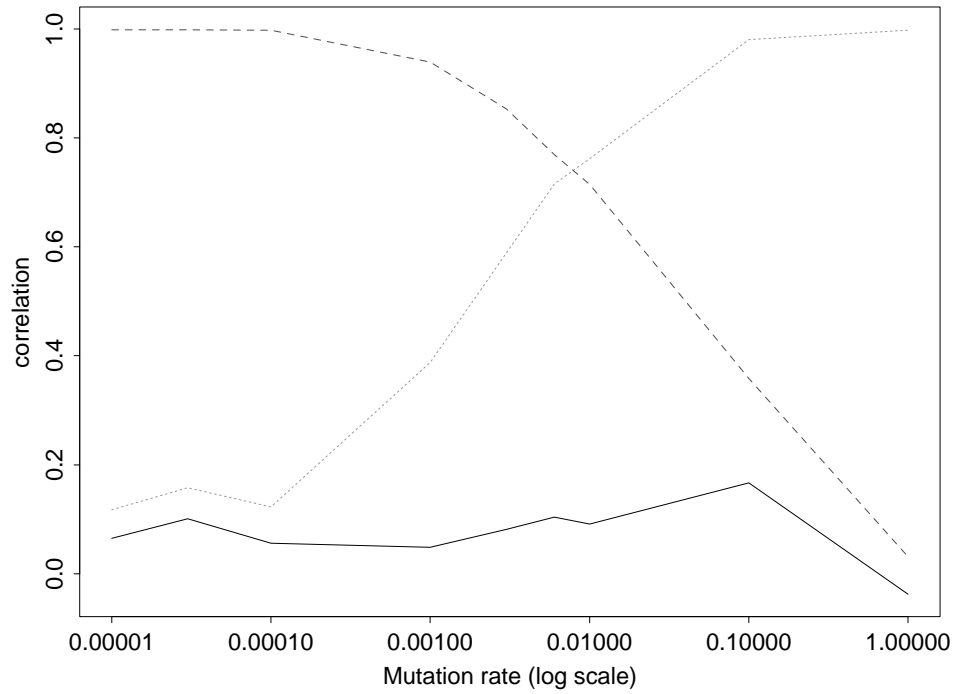
Diversity time-series (with selection)
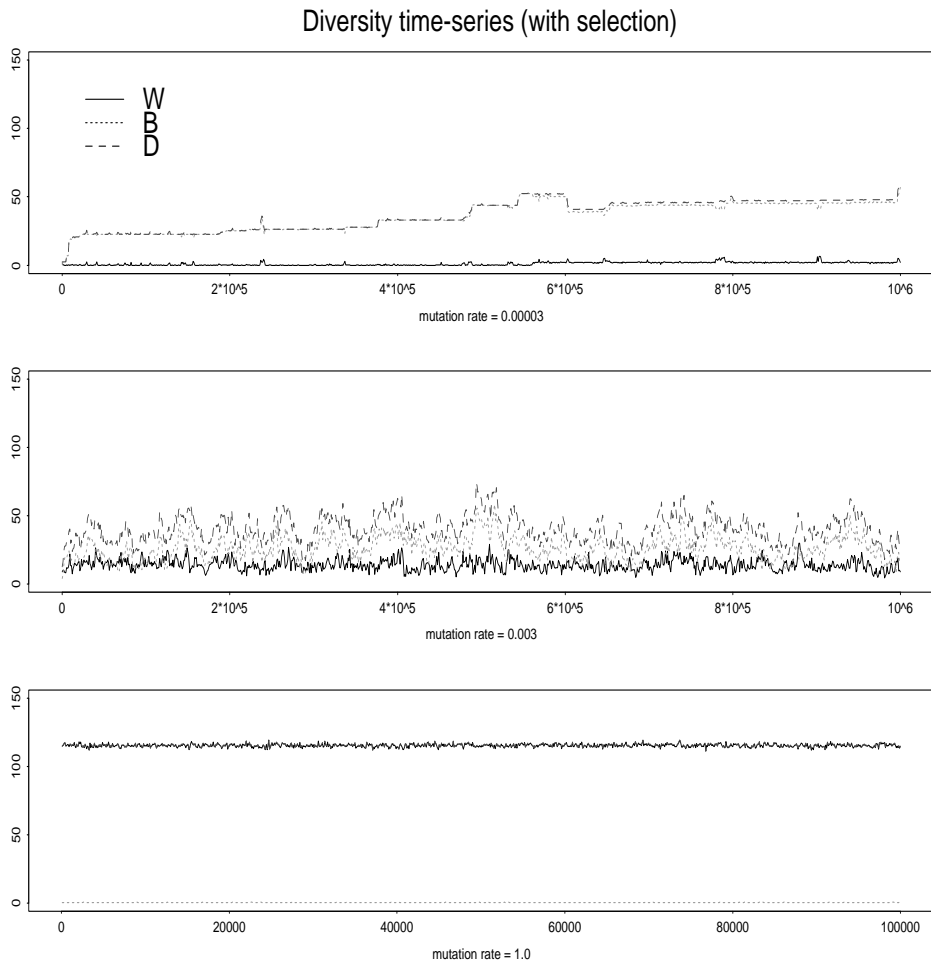
Figure 2: Time series of W^t, B^t, and D^t for three different mutation rates, as in the previous figure, except that adaptive evolution operates ($B_0 = 0$). The same transition in diversity dynamics is present. Above: $\mu = 3 \times 10^{-5}$. Middle: $\mu = 3 \times 10^{-3}$. Below: $\mu = 1.0$

22

Diversity correlations

Figure 3: The correlations among the diversity components: $\mathrm{Corr}(W,B)$ [solid line], $\mathrm{Corr}(W,D)$ [little dashes], $\mathrm{Corr}(B,D)$ [large dashes], where $W = \langle W^t \rangle_t$, $B = \langle B^t \rangle_t$, and $D = \langle D^t \rangle_t$. Note that at mutation rates below the transition D is well correlated with B, and at mutation rates above the transition D is well correlated with W. Adaptive evolution is absent in these data ($B_0 = 1$). Qualitatively similar results occur when adaptive evolution occurs ($B_0 = 0$).
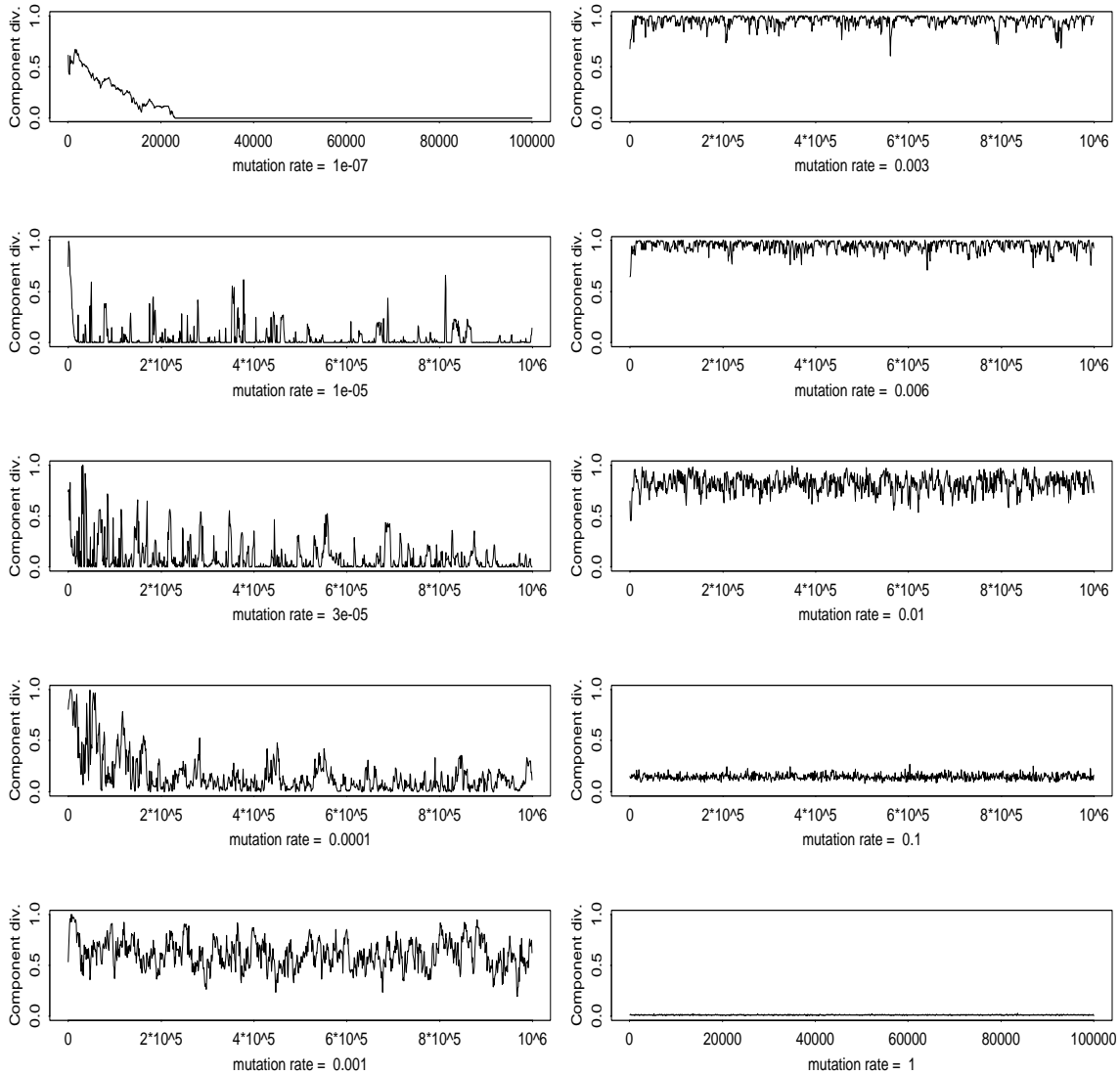
Figure 4: Time series of the component diversity, C^t, for a series of mutation rates when adaptive evolution is absent $(B_0 = 1)$. The transition is evident in the changes in these dynamics. Qualitatively similar results occur when adaptive evolution occurs $(B_0 = 0)$.

24
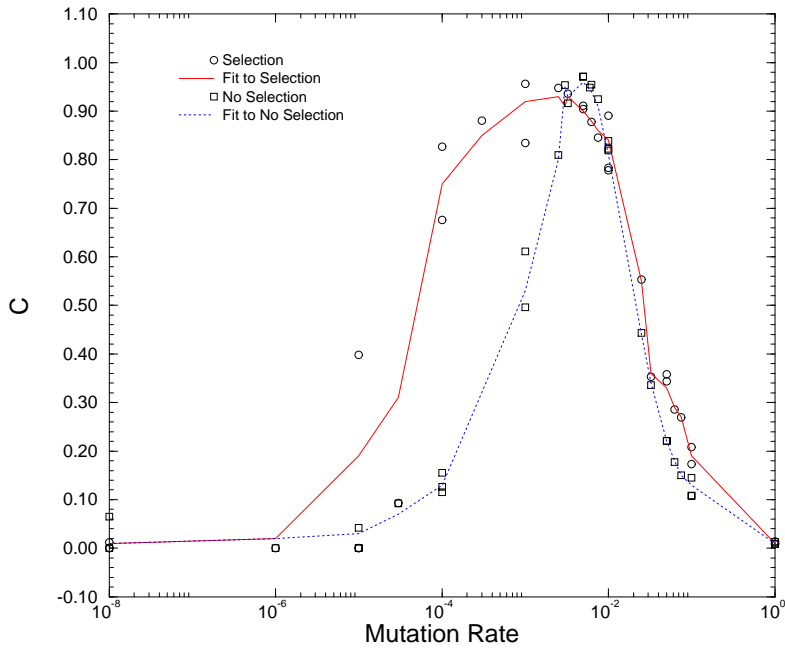
Figure 5: The extent to which total diversity is not dominated by one of its components, revealed by the time average $C = \langle C^t \rangle_t$, as a function of mutation rates, both with selection $(B_0 = 0)$ and without selection $(B_) = 1)$. A transition in shown around mutation rates in the range $10^{-3} \le \mu \le 10^{-2}$. The transition appears sharper without selection.
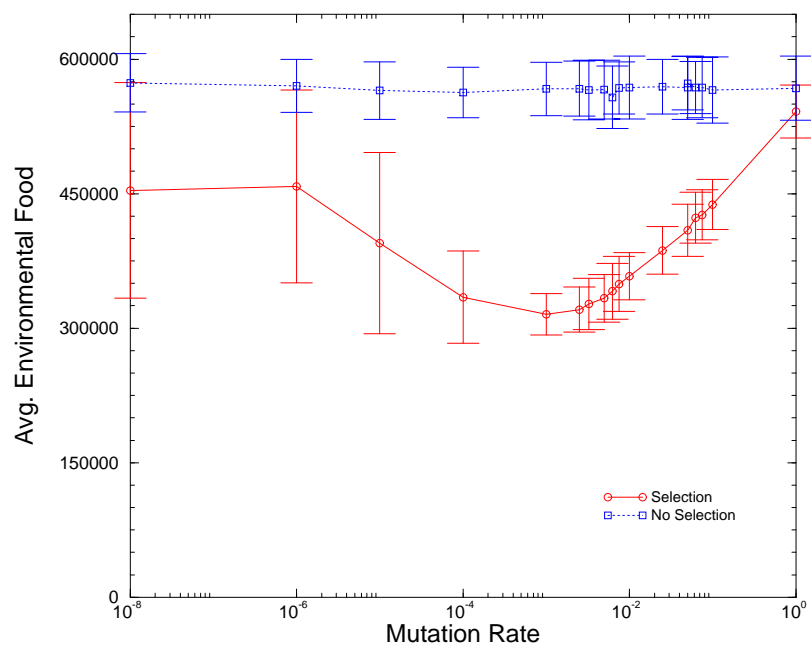
Figure 6: $\mathcal{E}_{resid}$, averaged over 10^4 time steps. Upper curve is without selection, lower curve is with selection. Maximal fitness occurs at mutation rates around the transition, $10^{-3} \leq \mu \leq 10^{-2}$.
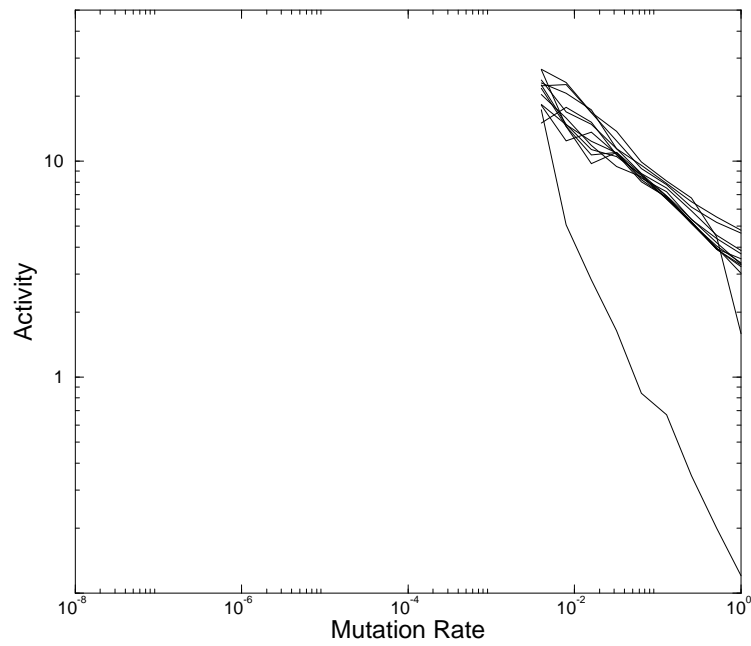
Figure 7: Evolutionary activity A as a function of mutation rate for several values of behavioral noise, $0 \leq B_0 \leq 1$. Note that evolutionary activity steadily increases as the mutation rate approaches the transitional region $10^{-3} \leq \mu \leq 10^{-2}$.
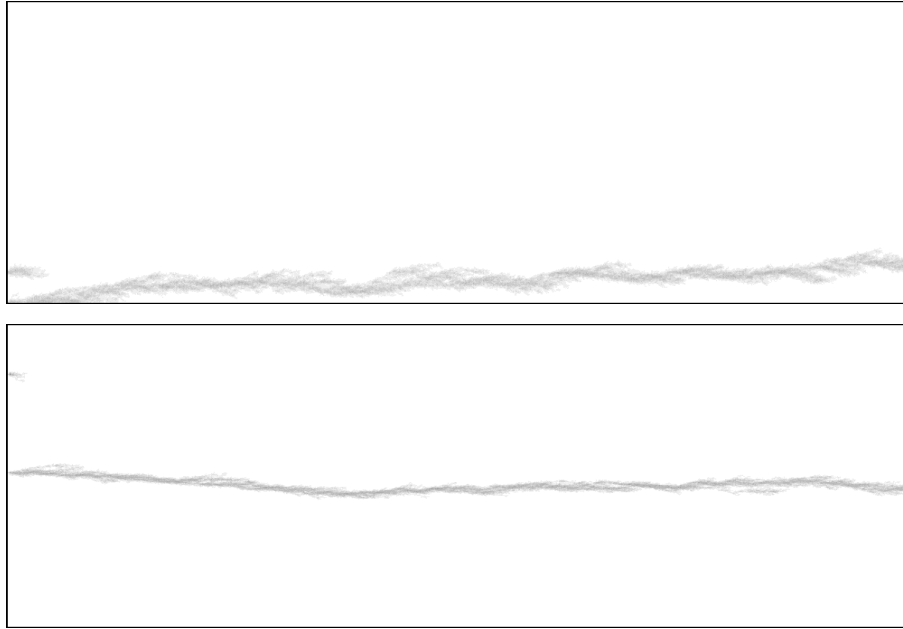
Figure 8: Dynamics of mutation rate distribution, $M^t(\mu)$, without adaptive evolution ($B_0 = 1$), at two different meta-mutation rates, for 10^5 time steps. Time increases to the right. The y-axis depicts 500 equal-sized bins, ranging from $1.0 \leq \mu_i < 0.998$ at the top to $0.002 \leq \mu_i \leq 0.0$ at the bottom. The gray-scale indicates a bin's frequency in the population (darker for greater frequency). Above: $\mu_\mu = 1.0$. Below: $\mu_\mu = 0.33$.

Figure 9: Dynamics of mutation rate distribution, $M^t(\mu)$, exactly as in the previous figure, but *with* adaptive evolution ($B_0 = 0$). Above: $\mu_\mu = 1.0$. Below: $\mu_\mu = 0.33$.
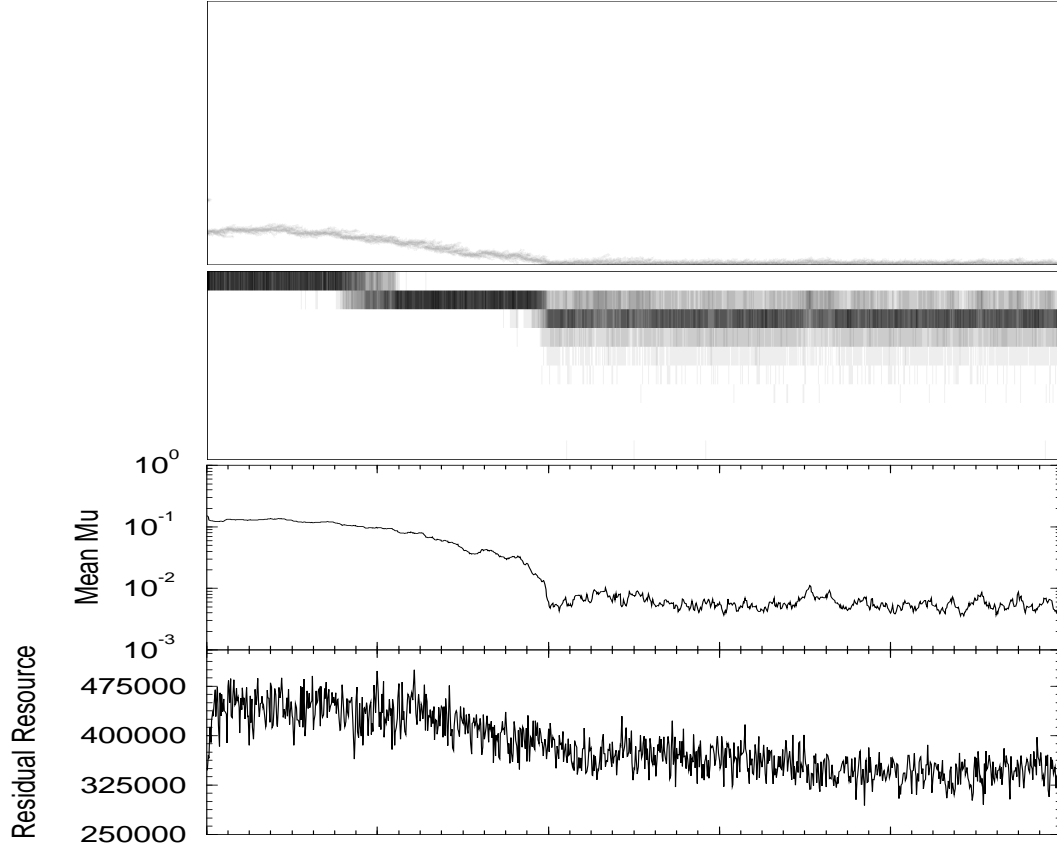
Figure 10: Time series of the mutation distribution $M^t(\mu)$ (with two forms of binning), population mean mutation rate $\langle \mu_i \rangle_{i \in I^t}$ (note the log-normal scale), and residual resource, from a typical simulation with adapting sensory-motor maps and mutation rates. 10^5 time steps are shown, with data sampled every 100 time steps; the mutation distribution of the initial population is flat across the mutation spectrum; $N_S = 1024$; $\mu_\mu = 0.66$. $M^t(\mu)$ above with 500 equal-sized bins, below with ten exponentially-smaller-sized bins (from top, $10^0 \leq \mu_i < 10^{-1}$, $10^{-1} \leq \mu_i < 10^{-2}$, etc.).
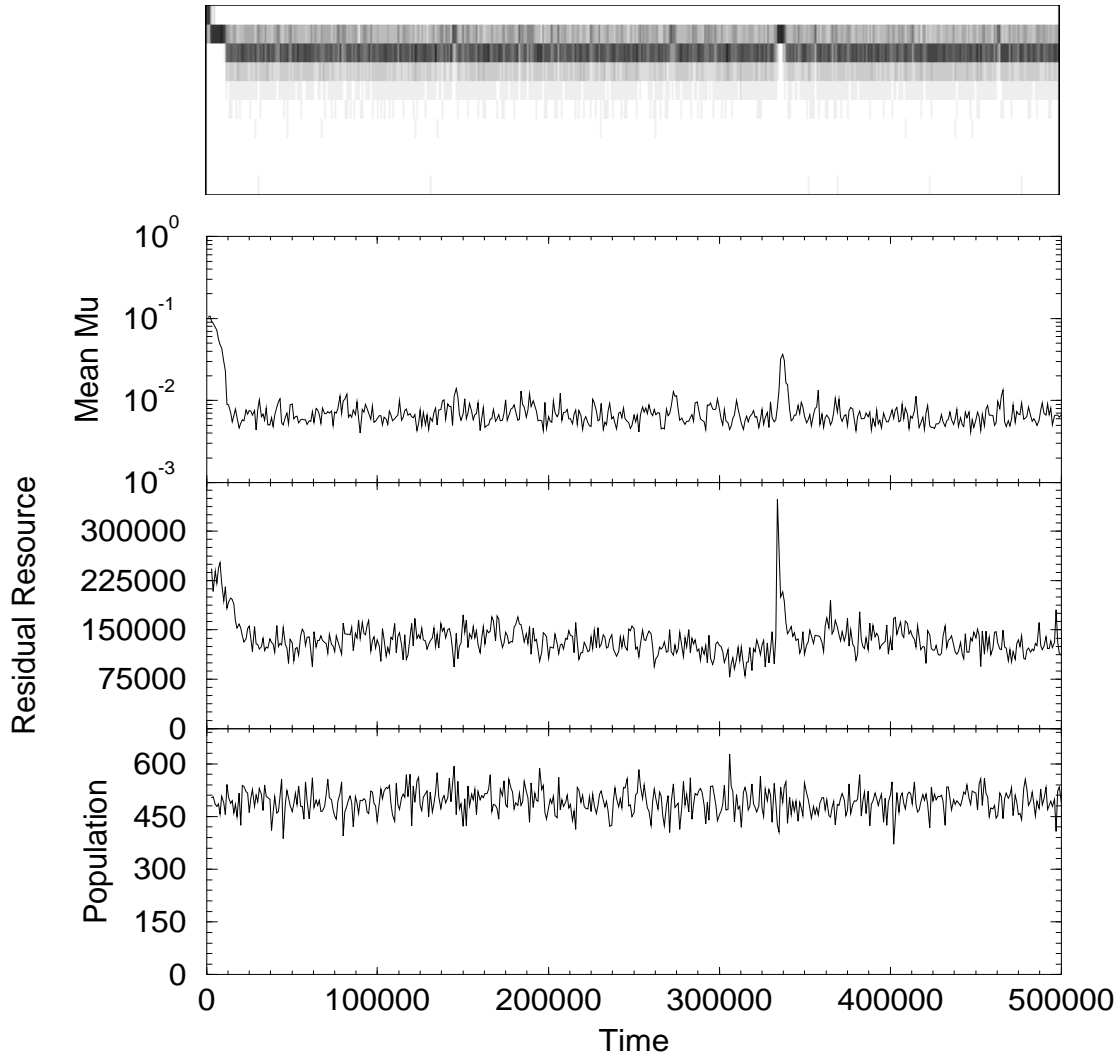
Figure 11: Time series of the mutation distribution $M^t(\mu)$ (with ten exponentially-smaller-sized bins, $10^0 \leq \mu_i \leq 10^{-1}$, $10^{-1} < \mu_i \leq 10^{-2}$, etc.), population mean mutation rate $\langle \mu_i \rangle_{i \in I^t}$ (note the log-normal scale), residual resource, and population level, from a typical simulation with adapting sensory-motor maps and mutation rates. 10^5 time steps are shown, with data sampled every 1000 time steps; the mutation distribution in the initial population is flat across the mutation spectrum; $N_S = 32$; $\mu_\mu = 0.66$. At time step 333334 all agents' sensory-motor genes—but not those governing mutation-rates—are randomized. 31