

## A Classification of Long-Term Evolutionary Dynamics

Mark A. Bedau<sup>†</sup>, Emile Snyder<sup>†</sup>, Norman H. Packard<sup>‡</sup>

<sup>†</sup>Reed College, 3203 SE Woodstock Blvd., Portland OR 97202, USA, {mab, emile}@reed.edu

<sup>‡</sup>Prediction Company, 236 Montezuma St., Santa Fe NM 87501, USA, n@predict.com

Published in C. Adami, R. Belew, H. Kitano, and C. Taylor, eds.,  
*Artificial Life VI* (Cambridge: MIT Press), 1998, pp. 189–198.

### Abstract

We present empirical evidence that long-term evolutionary dynamics fall into three distinct classes, depending on whether adaptive evolutionary activity is absent (class 1), bounded (class 2), or unbounded (class 3). These classes are defined using three statistics: diversity, new evolutionary activity (Bedau & Packard 1992), and mean cumulative evolutionary activity (Bedau *et al.* 1997). The three classes partition all the long-term evolutionary dynamics observed in Holland’s Echo model (Holland 1992), in a random-selection adaptively-neutral “shadow” of Echo, and in the biosphere as reflected in the Phanerozoic fossil record. This classification provides quantitative evidence that Echo lacks the unbounded growth in adaptive evolutionary activity observed in the fossil record.

### Why Classify Evolutionary Dynamics?

We present and illustrate a classification of long-term evolutionary dynamics. Classifications of complex dynamical behavior are reasonably familiar, with Wolfram’s classification of cellular automata rules being one well-known example (Wolfram 1984), but there are few attempts to classify the dynamics specifically of adaptive evolution. Nevertheless, such a classification is at least implicitly presupposed by the debates in biology about such issues as the evolution of clay crystallites (Cairns-Smith 1982; 1985), the evolution of “memes” (Dawkins 1976), and the increasing complexity of life on Earth (Gould 1989; McShea 1996; Gould 1996). Likewise for claims in artificial life about systems exhibiting “open-ended evolution” or “perpetual novelty” or operating “far from equilibrium” (Lindgren 1992; Ray 1992; Holland 1992; 1995; Bedau *et al.* 1997). Indeed, the defining focus of the field of artificial life—simulating and synthesizing systems that behave essentially like living systems—implies such a classification. How can we tell whether artificial systems behave relevantly like real living systems without using at least an implicit classification of system behavior?

The classification question arises sharply only when we have many concrete instances to classify, so our inattention to the classification question was understandable

when we had a sample size of only one—the biosphere. But the advent of artificial life changes this. Scores of artificial evolving systems are now generating many thousands of instances of long-term evolutionary dynamics. So we now have ample empirical data to tackle the classification question rigorously.

On the basis of studying data from a variety of artificial life models and from the biosphere, we have concluded that long-term evolutionary dynamics fall into three different classes. Our procedure here is to define statistics characterizing evolutionary dynamics and then use them to define three classes of long-term evolutionary trends. We then illustrate these classes of evolutionary dynamics in three systems: Holland’s Echo model (Holland 1992; 1995), a random-selection model that shadows Echo’s dynamics, and the Phanerozoic biosphere as reflected in the fossil record. We choose these systems to illustrate the kinds of dynamics because (i) Echo, among artificial life models, is an especially promising candidate for exhibiting complex adaptive evolutionary dynamics, (ii) Echo’s random-selection shadow provides an adaptively-neutral null case which highlights adaptations in Echo, and (iii) the Phanerozoic fossil record presents our best evidence about long-term dynamics in natural evolving systems. We are in the process of classifying many other artificial and natural evolving systems.

### Evolutionary Activity Statistics

Our classification of evolutionary dynamics is based on statistics for quantifying adaptive evolutionary phenomena. These statistics have already been applied to various evolving systems in various ways for various purposes (Bedau & Packard 1992; Bedau 1995; Bedau *et al.* 1997; Bedau & Brown 1997). This section describes these statistics with maximal generality and then explains how they are applied here.

Our evolutionary activity statistics are computed from data obtained by observing an evolving system. In our view an evolving system consists of a population of components, all of which participate in a cycle of birth, life and death, with each component largely determined by inherited traits. (We use this “component” terminology

to maintain enough generality to cover a wide variety of entities, ranging from individual alleles to taxonomic families.) Birth, however, creates the possibility of innovations being introduced into the population. If the innovation is adaptive, it persists in the population with a beneficial effect on the survival potential of the components that have it. It persists not only in the component which first receives the innovation, but in all subsequent components that inherit the innovation, i.e., in an entire lineage. If the innovation is not adaptive, it either disappears or persists passively.

Our idea of evolutionary activity is to identify innovations that make a difference. Generally we consider an innovation to “make a difference” if it persists and continues to be used. Counters are attached to components for bookkeeping purposes, to update each component’s current activity as the component persists and is used. If the components are passed along during reproduction, the corresponding counters are inherited with the components, maintaining an increasing count for an entire lineage. Two large issues immediately arise:

1. What should be counted as a component, and what counts as the addition or subtraction of a component from the system? In most evolving systems components may be identified on a variety of levels. Previous work has studied components on the level of individual alleles (Bedau & Packard 1992; Bedau 1995) as well as genotypes (Bedau *et al.* 1997; Bedau & Brown 1997) and taxonomic families (Bedau *et al.* 1997).

Here we study entire genotypes and taxonomic families. The addition or subtraction of a given component consists of the origination or extinction of a given genotype or taxonomic family. It’s natural to choose genotypes and taxonomic families as components because adaptive evolution can be expected to affect the dynamics of those entities.

2. What should be a new component’s initial contribution to the evolutionary activity of the system and how should it change over time? To measure activity contributions we attach a counter to each component of the system,  $a_i(t)$ , where  $i$  labels the component and  $t$  labels time. These activity counters are purely observational devices. A component’s activity increases over time as follows,  $a_i(t) = \sum_{k \leq t} \Delta_i(k)$ , where  $\Delta_i(k)$  is the activity increment for component  $i$  at time  $k$ . Various activity incrementation functions  $\Delta_i(t)$  can be used, depending on the nature of the components and the purposes at hand.

Since genotypes and taxonomic families are components in the present context, it’s natural to measure a component’s contribution to the system’s evolutionary activity simply by its age. Everything else being equal, the more adaptive an innovative genotype or

taxonomic family continues to be, the longer it will persist in the system. So we choose an activity incrementation function that increases a component’s activity counter by one unit for each time step that it exists:

$$\Delta_i(t) = \begin{cases} 1 & \text{if component } i \text{ exists at } t \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Though there are ways to refine this simple counting method (Bedau & Packard 1992; Bedau 1995; Bedau *et al.* 1997), this version facilitates direct comparison with many other systems.

In some contexts activity statistics indicate a system’s *adaptive* evolutionary dynamics only after the activity increment  $\Delta_i(t)$  is normalized with respect to a “neutral” model devoid of adaptive dynamics (Bedau 1995; Bedau *et al.* 1997; Bedau & Brown 1997). Here we address this issue in two different ways. With respect to taxonomic families in the Phanerozoic biosphere, we consider this normalization to be accomplished *de facto* by the fossil record itself. In our view, the mere fact that a family appears in the fossil record is good evidence that its persistence reflects its adaptive significance. Significantly maladaptive taxonomic families would likely go extinct before leaving a trace in the fossil record. But measuring evolutionary activity in Echo data is another matter, because we know maladaptive genotypes contribute to Echo’s activity data. So, to screen off the activity of maladaptive Echo genotypes, we measure evolutionary activity in a “neutral shadow” of Echo. Then, by comparing the Echo and neutral shadow data we can tell how much (if any) of Echo’s evolutionary activity is due to the genotypes’ adaptive value. The details of this neutral screening are explained in subsequent sections.

Now, we can define various statistics based on the components in a system and their activity counters. Perhaps the simplest statistic—because it ignores activity information—is the system’s *diversity*,  $D(t)$ , which is simply the number of components present at time  $t$ ,

$$D(t) = \#\{i : a_i(t) > 0\} \quad , \quad (2)$$

where  $\#\{\cdot\}$  denotes set cardinality.

The values of the activity counters of each component in the system over all time can be collected in the *component activity distribution*,  $C(t, a)$ , as follows:

$$C(t, a) = \sum_i \delta(a - a_i(t)) \quad , \quad (3)$$

where  $\delta(a - a_i(t))$  is the Dirac delta function, equal to one if  $a = a_i(t)$  and zero otherwise. Thus,  $C(t, a)$  indicates the number of components with activity  $a$  at time  $t$ . (Normalizing the component activity distribution by the

diversity,  $\frac{C(t,a)}{D(t)}$ , gives the *fraction* of components in the population with activity  $a$  at time  $t$ .)

A measure of the continual adaptive success of the components in the system at a given time is provided by the *total cumulative evolutionary activity*,  $A_{\text{cum}}(t)$ , which simply sums the evolutionary activity of all the components at a given time:

$$A_{\text{cum}}(t) = \sum_i a_i(t) \quad (4)$$

$$\rightarrow \int_0^\infty aC(t,a) da . \quad (5)$$

(In practice, we compute activity statistics using the sum; the integral indicated is obtained in the limit when activity takes on a continuum of values.) As the integral shows, you can think about  $A_{\text{cum}}(t)$  as the mass in the component activity distribution weighted by its level of activity. So, the cumulative activity per component, or *mean cumulative evolutionary activity*,  $\bar{A}_{\text{cum}}(t)$ , is simply the cumulative evolutionary activity  $A_{\text{cum}}(t)$  divided by the diversity  $D(t)$ :

$$\bar{A}_{\text{cum}}(t) = \frac{A_{\text{cum}}(t)}{D(t)} . \quad (6)$$

We sometimes refer to mean cumulative evolutionary activity simply as “mean activity.”

Adaptive innovations correspond to new components flowing into the system and proving their adaptive value through their persistent activity. Let  $a_0$  and  $a_1$  define a strip through the component activity distribution function,  $C(t,a)$ , such that activity values  $a$  in the range  $a_0 \leq a \leq a_1$  are among the lowest activity values that can be interpreted as evidence that a component has positive adaptive significance. Then, one reflection of the rate of the evolution of adaptive innovations is the *new evolutionary activity*,  $A_{\text{new}}(t)$ , which sums the evolutionary activity per component with values between  $a_0$  and  $a_1$ :

$$A_{\text{new}}(t) = \frac{1}{D(t)} \sum_{i, a_0 \leq a_i(t) \leq a_1} a_i(t) \quad (7)$$

$$\rightarrow \frac{1}{D(t)} \int_{a_0}^{a_1} C(t,a) da . \quad (8)$$

We sometimes refer to new evolutionary activity per component just as “new activity.”

Since we view any appearance in the fossil record as evidence of a taxonomic family’s positive adaptive significance (recall above), we measure new activity in the fossil record in a strip right along the bottom of the component activity distribution. To screen off the low activity values which might reflect maladaptive genotypes in Echo, we use a “neutral shadow” of Echo to determine that activity level,  $a'$ , at which we can begin to have

confidence that a component’s activity reflects its positive adaptive value, and we let  $a_0$  and  $a_1$  define a small window surrounding  $a'$ .

There is more than one way to quantify diversity and evolutionary activity. For example, another useful measure of diversity is the Shannon entropy of the distribution of sizes of components in the system. In addition, the choice of what to count as a system’s components affects a system’s diversity as measured by  $D(t)$ . Likewise, the activity statistics are affected by choices about, among other things, what the system’s components are, how to define the component activity incrementation function,  $\Delta_i(t)$ , where to set  $a_0$  and  $a_1$ , how to define a “neutral” model, etc. Furthermore, there are other kinds of activity statistics besides those defined here (Bedau & Packard 1992; Bedau 1995; Bedau & Brown 1997). Our specific choices of diversity and evolutionary activity statistics here is motivated by the desire to directly compare the adaptive evolutionary dynamics in Echo and in the Phanerozoic biosphere.

## Classes of Evolutionary Dynamics

On the basis of observing evolutionary dynamics from a variety of artificial and natural evolving systems, we have concluded that there are three fundamentally different kinds of long-term evolutionary dynamics:

**Class 1.** No adaptive evolutionary activity: diversity  $D$  is bounded, new activity  $A_{\text{new}}$  is zero, and mean activity  $\bar{A}_{\text{cum}}$  is zero.

**Class 2.** Bounded adaptive evolutionary activity: diversity  $D$  is bounded, new activity  $A_{\text{new}}$  is positive, and mean activity  $\bar{A}_{\text{cum}}$  is bounded.

**Class 3.** Unbounded adaptive evolutionary activity: diversity  $D$  is unbounded, new activity  $A_{\text{new}}$  is positive, and mean activity  $\bar{A}_{\text{cum}}$  is bounded. Evolutionary activity is growing because  $D$  is unbounded,  $\bar{A}_{\text{cum}}$  is bounded, and total cumulative evolutionary activity,  $A_{\text{cum}}$ , is their product.<sup>1</sup>

(The Appendix precisely defines what we mean by a statistic being positive or bounded.) The three classes of evolutionary dynamics apply equally well to artificial and natural evolving systems. Although we sometimes lack sufficient evidence for an unambiguous classification, and although the available evidence sometimes is misleading, we have found that the evolutionary dynamics of any evolving system in which our statistics can be defined will eventually be seen to fall into one of these three classes.

The classification of a system’s evolutionary dynamics depends on certain decisions made when defining the

<sup>1</sup>Unbounded  $\bar{A}_{\text{cum}}$  and bounded (or unbounded)  $D$  would also yield unbounded  $A_{\text{cum}}$ , but we have never observed such dynamics.

statistics. In particular, diversity and activity statistics can be implemented only after the components of a system are identified and the activity incrementation function,  $\Delta_i(t)$ , is defined. Thus, a system could exhibit different classes of evolutionary dynamics at different levels of analysis (say, the genetic and the phenotypic levels).

The three classes of evolutionary dynamics are not logically exhaustive. Other classes of long-term evolutionary dynamics can be defined, such as a system showing bounded diversity, zero new activity, and unbounded mean activity, or a system showing bounded diversity, positive new activity, and unbounded mean activity. And, in fact, some evolving systems do *appear* to exhibit these two kinds of dynamics. However, when a system has evolved long enough to reveal its long-term evolutionary dynamics, and when its evolutionary activity data is appropriately normalized with a neutral model, we have always found its behavior to fall into one of the classes 1-3. (If further study were to reveal the need for additional classes of evolutionary dynamics, they should be definable with our statistics.)

In the first instance, our classification applies to the evolutionary dynamics in a given run of a given system. But if different runs of the same system at the same spot in parameter space all exhibit the same class of evolutionary dynamics, then the classification is a generic property of that system at that place in parameter space. Further, if the same class of evolutionary dynamics is exhibited by a system across a large area of parameter space, then the classification is even more generic. When adjacent regions in parameter space have different kinds of generic evolutionary dynamics, an important question is to identify and explain the line demarking these dynamics. Finally, a class of evolutionary dynamics might be shared as a generic feature across a large area of parameter space by a wide class of evolving systems, including both those found naturally and those constructed artificially.

## The Echo Model

John Holland created Echo in the attempt to produce a model that would illustrate the creation of complex structures by natural selection (Holland 1995). Echo's central explicit focus is to allow natural selection to shape the strategies by which a population of agents engage in various kinds of interactions. Detailed information about the Echo model is available elsewhere (Holland 1992; Jones & Forrest 1993; Holland 1994; Forrest & Jones 1994; Holland 1995; Hrabar, Jones, & Forrest 1997; SFI 1998).

An Echo world consists of a toroidal lattice of sites, each site having a resource fountain and a population of agents. (The Echo runs we describe here consist of worlds with only one site.) Different letters of the alphabet represent different types of resources available in

the world. A fixed amount of resources is distributed to each site at each time step, and unconsumed resources accumulate at a site up to a fixed ceiling.

An Echo agent consists of a "chromosome" that is composed of eleven sub-strings of the world's resources (letters of the alphabet) together with a reservoir storing excess resources. The sub-strings of the chromosome constitute an agent's *external tags* and *internal conditions* together with an uptake mask which specifies what resources the agent can take up from the environment. An agent's tags are external in the sense that other agents have access to them, while an agent's conditions are inaccessible to other agents. The tags and conditions are used to determine the outcome of the three types of interactions that Echo agents can engage in—combat, trade, and mating. Whether two agents interact and, if so, what type of interaction they have is determined by comparing the agents' tags and conditions. A string match of the appropriate tag and condition causes the interaction to take place. External tags and internal conditions allow complex (e.g. non-transitive) relationships to exist between the agents, and it is central to Echo's endogenous fitness function (a fitness function that is an emergent property of the environment and the other agents (Packard 1989)).

The combat interaction gives a good illustration of how tags and conditions are used. Two individuals engage in combat provided there is a prefix match between their combat conditions and the other individuals' offense tag. Each individual's payoff of the combat interaction is determined by a calculation based on the letters in the two individuals' offense and defense tags, and the winner of the combat is chosen probabilistically, based on the two individuals' relative payoffs. The losing agent gets a chance to flee, and otherwise is killed and loses its resources to the winner.

Trading and mating interactions use tags and conditions in a related way. Trading takes place if there is a prefix match between the trading condition of the first agent and the offense tag of the other agent. A trading interaction between two agents results in each agent transferring the excess of its trading resource (the amount of resources in the agent's reservoir over and above what it needs for reproduction) to the other agent. The mating interaction takes place if a bilateral match is found between the mating tags and conditions of two agents chosen to interact. The result of a successful mating interaction is more analogous to the types of genetic exchange seen in bacteria as opposed to sexual reproduction. The two participating agents exchange genetic material via crossover (at a random point in the chromosome) and replace their "parents" in the population.

Agents that have acquired enough resources in their reservoir to copy their chromosome reproduce asexually.

Asexual reproduction is subject to a probability,  $\mu$ , of a point mutation as well as probabilities of mutation by crossover and by insertion-deletion within the parent chromosome. As a part of asexual reproduction, parents give a fixed percentage of the resources remaining in their reservoir to their offspring. In addition to gathering resources from the environment, agents lose resources through a metabolic tax  $\tau$ , as well as by asexual reproduction, and they gain and lose resources by fighting and trading. The interaction probability,  $\iota$ , determines the probability that nearby agents will engage in the interactions that affect their resource levels. It is mutation together with the selection pressure due to competition for resources that drives the evolution of Echo’s population.

One time step in the Echo model consists of the following cycle of events: A proportion of the agents are selected to undergo interactions and the interactions take place. Resources at a site are distributed to those agents that can accept them. Agents are taxed probabilistically. Some agents are randomly killed and their resources returned to the environment. Agents that have not collected resources migrate to a randomly chosen neighboring site (in multi-site worlds). Finally, agents that have acquired sufficient resources reproduce asexually.

### Echo’s Neutral Shadow

In order to discern which features of Echo’s genotype statistics can be attributable to the genotypes’ adaptive significance, we created a “neutral shadow” of Echo. The crucial property of Echo’s neutral shadow is that its evolutionary dynamics are like Echo’s *except* that a genotype’s presence or concentration or longevity in the shadow population cannot be due to its adaptive significance.

Echo’s neutral shadow consists of a population of nominal “creatures” with nominal “genotypes.” A shadow “creature” has no spatial location and it cannot ingest resources or interact with other “creatures.” All it ever does is come into existence, perhaps reproduce (perhaps many times), and go out of existence; its only properties are its genotype and the times of its birth, reproductions (if any), and death.

Each Echo run has its own corresponding neutral shadow run. Changes in the Echo run sometimes cause corresponding changes in its neutral shadow, but changes in the neutral shadow never affect the run (hence the “shadow” terminology). The timing and number of birth and death events in the neutral shadow are directly copied from those in the normal Echo run, as is the neutral shadow’s mutation rate.

When some creature is born in the normal Echo run, it is time for a birth event in the shadow model, so a shadow parent chosen at random (with equal probability) from the shadow population reproduces. The new

shadow child inherits its parent’s genotype unless a mutation gives the child a new, unique genotype. When some creature dies in the normal Echo run, it is time for a death event in the shadow model and a “creature” is chosen at random (with equal probability) from the shadow population and killed. Thus, all selection in the neutral shadow is random.

The evolutionary dynamics in a neutral shadow is a neutral diffusion process in genotype space. Genotypes arise and go extinct, and their concentrations change over time, but the genotype dynamics are at best weakly linked to adaptation through the birth and death rates determined by adaptation in the normal Echo model. The birth, reproduction and death statistics that drive a neutral model “shadow” those of the Echo model, and those in the Echo model are (typically) affected by adaptation. Still, properties like the relative longevity and concentration of a genotype in the neutral shadow cannot be due to the genotype’s adaptive significance. All selection in the shadow model is random so no genotype has any adaptive significance. At the same time, by precisely mimicing the births, deaths, and mutation rate in a normal Echo run, the neutral shadow’s behavior helps us to determine which aspects of the behavior of the normal Echo run can be attributed to the adaptive significance of genotypes and which might reflect nothing more than the system’s underlying architecture or chance.

Figure 1 illustrates the difference between Echo and its neutral shadow. The Figure shows a “side view” of component activity distributions (from the Echo run and neutral shadow shown in Figure 4). These distributions have been collapsed (summed) along the temporal dimension and then divided by the total number of counts in both distributions. There is no guarantee that an Echo run and its neutral shadow will have the same number of genotypes. In fact, often the neutral shadow has more genotypes, since natural selection does not preferentially preserve those that are well adapted. By dividing the distributions by the total number of activity counts in both distributions, the value of each distribution at a given activity value  $a$  reflects the fraction of activity counts in each distribution that have activity  $a$ .

Note that, on average, the activity counters in the neutral shadow’s collapsed activity distribution are lower than those in Echo’s collapsed activity distribution; i.e., the distribution is squashed to the left. This is just how one would expect the neutral shadow’s random selection to affect a component activity distribution. By construction, individuals in Echo and its neutral shadow have the same birth, reproduction, and death rates, and their mutation rates are the same (indeed, all model parameters are identical). But while the selective force in the neutral shadow is entirely random, natural selection can pref-

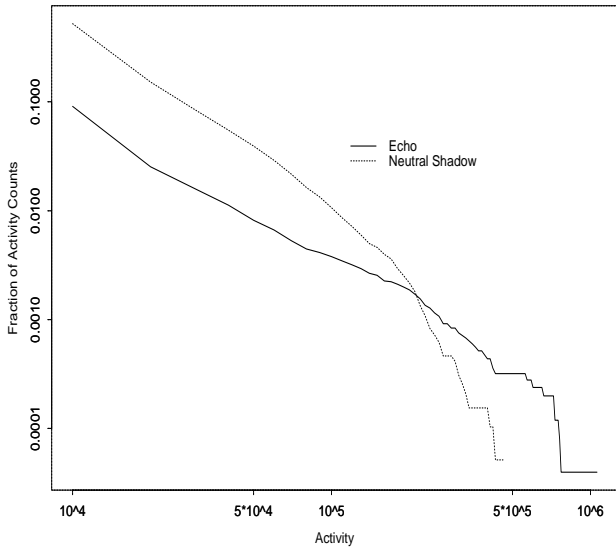


Figure 1: Log-log plot of the component activity distributions for the Echo run shown and neutral shadow shown in Figure 4, where these distributions have been collapsed (summed) along the temporal dimension. As one would expect, the neutral shadow’s genotypes show relatively lower activity. The point at which the distributions are equal is the activity value at which an activity count has the same probability of having occurred in the Echo and neutral shadow distributions. Since these distributions are equal at activity  $a' = 2.1 \times 10^5$ , we set  $a_0$  and  $a_1$  (used to calculate new activity,  $A_{\text{new}}$ ) slightly above and below this value, specifically,  $a_0 = 1.7 \times 10^5$  and  $a_1 = 2.5 \times 10^5$ .

entially cull poorly adapted genotypes and preserve well adapted genotypes in Echo. This squashes the low-activity end of Echo’s collapsed distribution and inflates its high-activity end. The difference between the two collapsed distributions quantifies *how much* natural selection affects the activity counts in Echo’s component activity distribution.

The point at which the two distributions have the same value (i.e., cross) reveals the activity value,  $a'$ , at which an activity count is equally likely to have been chosen from either distribution. Thus, to calculate new activity,  $A_{\text{new}}$ , we set  $a_0$  and  $a_1$  slightly above and below  $a'$ . (Recall our discussion of evolutionary activity statistics above.) Specifically, if we let  $a_{\text{max}}$  be the highest activity value at which either collapsed distribution is positive and if we let  $a'$  be the lowest value at which the two collapsed distributions cross, then we set  $a_0$  and  $a_1$  to be  $a' \pm (0.05 \times (a_{\text{max}} - a'))$ .

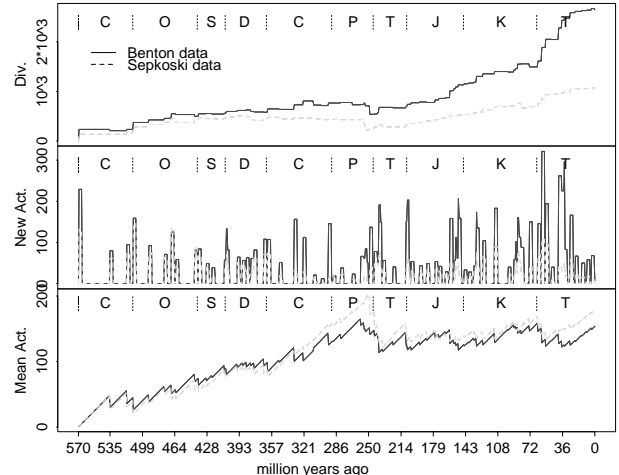


Figure 2: Diversity, new activity, and mean cumulative activity in the fossil data of Benton and Sepkoski. The labels at the top of each graph show the boundaries between the standard geological periods, thus: Cambrian, Ordovician, Silurian, Devonian, Carboniferous, Permian, Triassic, Jurassic, Cretaceous, Tertiary.

## The Fossil Data

We used two fossil data sets, each of which indicates the geological stages or epochs with the first and last appearance of taxonomic families. Benton’s data (Benton 1993) covers all families in all kingdoms found in the fossil record, for a total of 7111 families. Sepkoski’s data (Sepkoski 1992) indicates the fossil record for 3358 marine animal families. The duration of different stages and epochs varies widely, ranging over three orders of magnitude. In order to assign a uniform time scale to the fossil data, we used Harland’s time scale (Harland *et al.* 1990) to convert stages and epochs into time indications expressed in units of millions of years before the present.

We are most interested in classifying long-term trends among fossil species, but we study fossil families because much more complete data is available at this level of analysis (Valentine 1985; Sepkoski & Hulver 1985). Although fossil family data is certainly no precise predictor of fossil species data, there is evidence that species-level trends in the fossil record are reflected at the family level (see (Valentine 1985) and the references cited therein). Sepkoski and Hulver ((Sepkoski & Hulver 1985), p. 14) summarize the situation thus: ‘Although families do not display all of the detail of the fossil record, they should be sufficiently sensitive to show major evolutionary trends and patterns with characteristic time scales of fives to tens of millions of years.’ The trends we use to classify evolutionary dynamics occur in the fossil data on time scales at least that long.

## Results

We computed diversity, new activity, and mean cumulative activity in the Benton and Sepkoski fossil data sets (see Figure 2). We also computed these statistics from data produced by the Echo model and its neutral shadow at a variety of places in Echo’s parameter space (see Figures 3 and 4). Comparing the data from Echo and its neutral shadow allows us to normalize Echo’s evolutionary activity statistics.

The paucity of earlier fossil data lead us to restrict our attention to the Phanerozoic fossils, which start with the Cambrian explosion. The major extinction events, such as the mammoth one which ends the Permian period, and the famous K/T extinction which ended the age of the dinosaurs, are visible in the data. The overall trends in the fossil statistics are pretty unambiguous:  $D$  is unbounded,  $A_{\text{new}}$  is positive, and  $\bar{A}_{\text{cum}}$  is bounded. Thus the evolutionary dynamics of the Phanerozoic biosphere is in class 3.

We examined the evolutionary dynamics of Echo and its neutral shadow while varying three crucial parameters across their entire viable range. The per-locus mutation rate,  $\mu$ , ranged over  $0 \leq \mu \leq 1$ ; the interaction probability,  $\iota$ , ranged over  $0 \leq \iota \leq 1$ ; and the metabolic tax (metabolic tax),  $\tau$ , ranged over  $0 \leq \tau \leq 0.45$ . All other Echo parameters were held constant across all runs. (The CD available with this volume contains the Echo source code we used, including the parameter files, as well as evolutionary activity analysis software.)

To normalize the Echo data by a “neutral” model, we compare activity data from Echo and its neutral shadow. Not all of the activity generated by Echo reflects adaptive innovations. In fact, the neutral shadow’s activity shows how much “raw” activity accumulates in a non-adaptive analogue of Echo. So, we normalize Echo’s activity data by subtracting the neutral shadow’s new or cumulative activity from that of Echo. If the result is negligible or negative, then Echo’s normalized activity is nil. Since that level of raw activity has been observed in the neutral shadow, it is not evidence of the adaptive value of Echo’s components. On the other hand, if Echo’s level of raw activity is significantly higher than its neutral shadow, then we have good evidence that this residue—the normalized new or cumulative activity—indicates significant new and cumulative adaptive success of the system’s components.

After making sure that we were observing long-term trends and properly normalizing the activity data, we found that Echo’s evolutionary dynamics fell into either class 1 or class 2. Since long-term diversity dynamics were always bounded, class 3 dynamics never materialized.

If the mutation rate was very low (at or near zero), Echo and its neutral shadow show virtually identical evolutionary dynamics. Except for fleeting exceptions

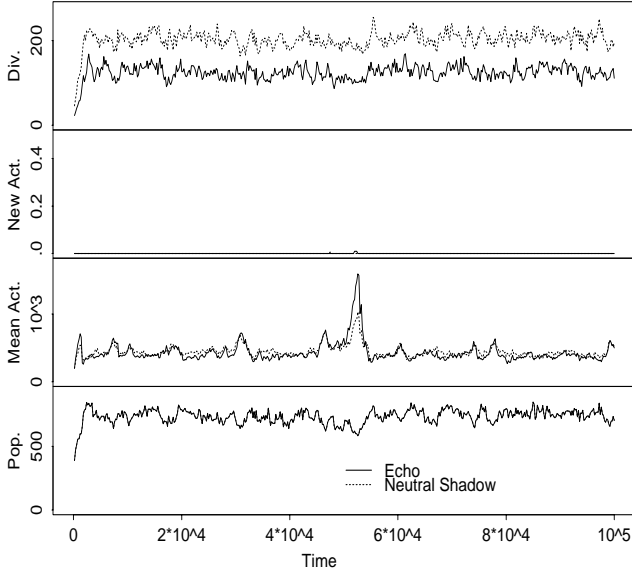


Figure 3: Typical diversity, new activity, and mean cumulative activity data from Echo and a neutral shadow when mutation rate  $\mu = 10^{-2}$ , interaction probability  $\iota = 0.05$ , and metabolic tax  $\tau = 0.01$ . Note that  $A_{\text{new}} = 0$  for both Echo and the neutral shadow and that  $\bar{A}_{\text{cum}}$  is not significantly higher in Echo than in the neutral shadow. (Here,  $a_0 = 6.4 \times 10^3$  and  $a_1 = 6.8 \times 10^3$ .)

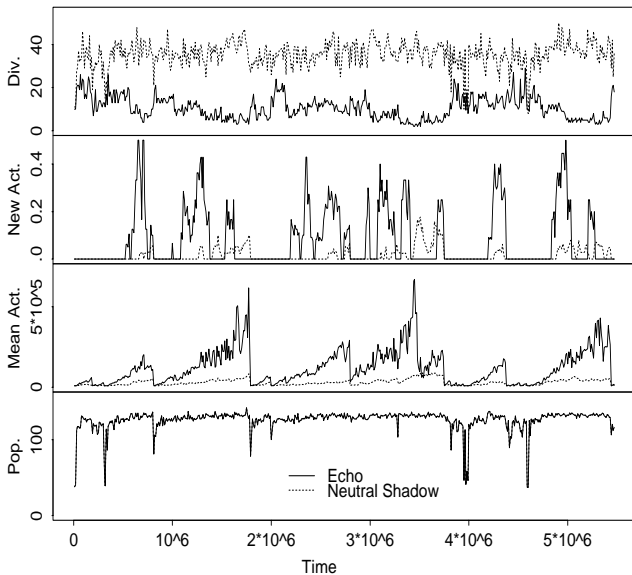


Figure 4: Typical diversity, new activity, and mean cumulative activity data from a Echo and a neutral shadow when mutation rate  $\mu = 10^{-2}$ , interaction probability  $\iota = 0.85$ , and metabolic tax  $\tau = 0.15$ . Note that  $A_{\text{new}}$  and  $\bar{A}_{\text{cum}}$  are significantly higher in Echo than in the neutral shadow. (Here,  $a_0 = 1.7 \times 10^5$  and  $a_1 = 2.5 \times 10^5$ .)

caused by a mutation, only one genotype exists at a time, so the indefinite trend is  $D(t) = 1$  and  $A_{\text{new}}(t) = 0$ . This causes the “raw” mean activity to increase with a slope of unity in both Echo and its neutral shadow, so the normalized mean activity is zero. Thus, the evolutionary dynamics of Echo when  $\mu \approx 0$  falls into class 1.

When the mutation rate is very high (at or near zero), the evolutionary dynamics of Echo and its neutral shadow are again virtually identical. A child’s genotype is virtually guaranteed to differ from that of its parent, so virtually every genotype has only one instance and  $D(t)$  remains very high. Furthermore, those genotypes that by chance have some adaptive significance have no chance to leave an imprint on the population, which means that the collapsed component activity distributions of Echo and the neutral shadow are virtually identical. Thus,  $a_0$  is set so high that  $A_{\text{new}}(t) \approx 0$  in both Echo and the neutral shadow. In addition, the two models have such similar “raw” mean activity dynamics that after normalization the consistent trend is  $\bar{A}_{\text{cum}}(t) \approx 0$ . Thus, the evolutionary dynamics of Echo when  $\mu \approx 1$  falls into class 1.

If the mutation is between these extremes, then the long-term evolutionary dynamics depend on other system parameters. Here we focus on two other parameters: interaction probability,  $\iota$ , and metabolic tax,  $\tau$ . Previous work has shown that these parameters are key determinants of evolutionary activity in Echo (Smith 1998). For example, when both  $\iota$  and  $\tau$  are very low, then even at moderate mutation rates Echo exhibits behavior reminiscent of what happens when the mutation rate is very high. For example, Figure 3 shows typical long-term statistical trends in Echo and its neutral model with  $\iota = 0.05$  and  $\tau = 0.01$ . The long-term trend is clearly that diversity is bounded and new activity is zero. Furthermore, since “raw” mean activity is about the same in Echo and the neutral shadow, normalized mean activity is approximately zero. This illustrates how, at very low interaction probability and metabolic tax, Echo has class 1 evolutionary dynamics regardless of the mutation rate.

On the other hand, at intermediate mutation rates, evolutionary activity in Echo increases significantly within a certain range of interaction probabilities and metabolic taxes, specifically, when  $0.5 \leq \iota \leq 1.0$  and  $0.15 \leq \tau \leq 0.4$  (Smith 1998). Some qualitative features of the evolutionary dynamics vary with the mutation rate, as one would expect, but those features of the statistics that determine long-term evolutionary dynamics remain the same. Figure 4 shows typical dynamics of the statistics from an Echo run and its neutral shadow within this region of parameter space. First, diversity is bounded (and significantly higher in the neutral shadow, as one would expect). Second, both new and mean activity are positive. Moreover, both are significantly higher

in Echo than in the neutral shadow, so Echo’s normalized new and mean activity are positive. Thus, Echo’s long-term evolutionary dynamics fall into class 2.

Finally, it is worth noting that, when normalized, the neutral shadows themselves have no new or cumulative evolutionary activity. Since data from a neutral shadow and *its* neutral shadow would look alike, subtracting one from the other would yield nothing. Thus, its normalized new and cumulative activity will be zero. In addition, since the qualitative shape of a neutral shadow’s diversity dynamic follows that of the Echo run which it shadows, and since all observed Echo runs show bounded diversity dynamics, so do all of Echo’s neutral shadows. For this reason, the long-term evolutionary dynamics of all observed neutral shadows of Echo falls into class 1.

Table 1 summarizes the three classes of evolutionary dynamics and the examples of each we have observed.

## Discussion

Our classification of long-term evolutionary dynamics in Echo, the neutral shadow, and the biosphere suggests three main conclusions:

*Conclusion 1:* New evolutionary activity measures the flow of adaptive innovations into an evolving system and mean cumulative evolutionary activity measures the continual adaptive success of such innovations. The primary evidence for this is the comparison between Echo and its neutral shadow and the effect of varying key Echo parameters (mutation rate, probability of interaction, metabolic tax) governing the process of adaptation. Further evidence supporting this conclusion comes from comparisons between other artificial evolving models and their neutral shadows (Bedau 1995; Bedau *et al.* 1997; Bedau & Brown 1997).

*Conclusion 2:* Comparison of the long-term evolutionary dynamics observed in Echo, its neutral shadow, and the Phanerozoic biosphere reveals these to be partitioned into three distinct classes: no adaptive evolutionary activity (class 1), bounded adaptive evolutionary activity (class 2), and unbounded adaptive evolutionary activity (class 3). All neutral shadow dynamics and some Echo dynamics fall into class 1, the rest of Echo dynamics fall into class 2, and only the biosphere dynamics fall into class 3.

*Conclusion 3:* If we accept conclusions 1 and 2, then Echo and the biosphere exhibit qualitatively different kinds of evolutionary dynamics. In particular, Echo lacks the unbounded growth in adaptive activity observed in the fossil record.

Classes 1-3 provide a classification of the evolutionary dynamics in artificial models and natural evolving systems. These classes have internal quantitative structure and they can be further subdivided, but we think that these three classes mark the most fundamental distinction among adaptive evolutionary dynamics. To be



CLASS	EVOLUTIONARY ACTIVITY	STATISTICAL SIGNATURE			EXAMPLES
		$D$	$A_{\text{new}}$	$\bar{A}_{\text{cum}}$	
1	none	bounded	zero	zero	Echo $_{\mu \approx 0}$ Echo $_{\mu \approx 1}$ Echo $_{\iota \approx 0, \tau \approx 0}$ all neutral shadows of Echo
2	bounded	bounded	positive	bounded	Echo $_{10^{-4} \leq \mu \leq 10^{-1}, 0.5 \leq \iota \leq 1.0, 0.15 \leq \tau \leq 0.4}$
3	unbounded	unbounded	positive	bounded	Phanerozoic biosphere

Table 1: Classes of evolutionary dynamics and their statistical signatures observed in Echo, Echo’s neutral shadow, and data from the fossil record. The Echo parameters varied in these examples are mutation rate,  $\mu$ , interaction probability,  $\iota$ , and metabolic tax,  $\tau$ .

sure, detecting these classes requires surmounting some practical problems. A system must be observed long enough for long-term trends to reveal themselves, and seeing a system’s specifically adaptive evolutionary activity might require normalization with a suitable “neutral” model. Nevertheless, the payoff of surmounting these obstacles is the ability to classify an evolving system by reference to an elusive and controversial (Gould & Lewontin 1979) but central property: the extent to which adaptations are being created by the process of evolution.

A weakness with the statistics we use to define classes 1-3 is the “emergence” problem: The statistics can be applied only after settling what a system’s components are and what counts as their activity, so the statistics would not directly reflect the evolutionary innovation of genuinely novel kinds of system components. The emergence problem does not arise when classifying the fossil data, because *post hoc* analysis has identified the relevant system components. Furthermore, with existing artificial life models, our understanding the system usually allows us to identify the relevant components confidently. Anyway, it’s unclear how serious the emergence problem will prove in practice. On the one hand, as discussed earlier, evolutionary activity statistics are always defined at a given level of analysis, and we should not expect to see the evolutionary activity at *all* levels with statistics defined at *one* level. On the other hand, we *would* often expect to see significant adaptive innovations echoed in activity statistics across many levels. For example, activity statistics defined at the level of individual cell types in the biosphere would show marked activity at the origination of multicellular life. So, activity statistics defined at one level will often indirectly indicate the emergence of higher levels of adaptive activity. The fact that we do not see this sort of signature in the Echo data indicates that higher levels of adaptive innovation are probably not occurring.

There are special problems and pitfalls inherent in using the fossil record to study long-term trends (Raup

1988). In particular, the “pull of the present” is a well-known sampling bias due to the fact that there are simply more recent fossils to study than older fossils. Future work will investigate the extent to which our classification of the evolutionary dynamics evident in the fossil record can be supported more rigorously.

Although we focus here only on Echo, its neutral shadow, and the Phanerozoic biosphere, our methodology and conclusions have quite broad import. Some natural evolving systems probably have class 2 dynamics. For example, space and time constraints might bound the adaptive activity of bacterial evolution in a chemostat. Other natural evolving systems probably show class 3 dynamics. Class 3 dynamics might even be detectable in systems like the global economy or internet traffic. We also suspect that no existing artificial evolving system has class 3 dynamics. In our opinion, creating such a system is among the very highest priorities of the field of artificial life. From one perspective, this is a negative result: Echo, and perhaps all other existing artificial evolutionary systems, apparently lack some important characteristic of the biosphere—whatever is responsible for its unbounded growth of adaptive activity. But at the same time this conclusion calls attention to the important constructive and creative challenge of devising an artificial model that succeeds where all others have failed. Here, again, classes 1-3 show their value, for they provide a feasible, objective, quantitative test of success.

**Acknowledgements.** Thanks to M. J. Benton and J. J. Sepkoski, Jr., for making their fossil data available. For help with Echo, thanks to Terry Jones, Simon Fraser, and Richard Smith. Thanks to Andreas Rechtsteiner and Richard Smith for on-going discussions about Echo and neutral models. Thanks for helpful comments to the anonymous reviewers for Artificial Life VI. Thanks also to the Santa Fe Institute for the hospitality, financial support, and computation resources which initiated and helped sustain this work.

## Appendix: Definitions

In this paper our operational definitions of what it is for a function  $f(t)$  to be unbounded or positive are as follows: The function  $f(t)$  is unbounded *iff*

$$\lim_{t \rightarrow \infty} \left( \frac{\sup(f(t))}{t} \right) > 0 , \quad (9)$$

where  $\sup(\cdot)$  is the supremum function. The function  $f(t)$  is positive *iff*

$$\lim_{t \rightarrow \infty} \left( \frac{\int_0^t f(t) dt}{t} \right) > 0 . \quad (10)$$

## References

- Bedau, M. A., and Brown, C. T. 1997. Visualizing evolutionary activity of genotypes. *Adaptive Behavior*. Submitted.
- Bedau, M. A., and Packard, N. H. 1992. Measurement of evolutionary activity, teleology, and life. In Langton, C.; Taylor, C.; Farmer, D.; and Rasmussen, S., eds., *Artificial Life II*. Addison-Wesley.
- Bedau, M. A.; Snyder, E.; Brown, C. T.; and Packard, N. 1997. A comparison of evolutionary activity in artificial evolving systems and in the biosphere. In Husbands, P., and Harvey, I., eds., *Proceedings of the Fourth European Conference on Artificial Life*. MIT Press/Bradford Books.
- Bedau, M. A. 1995. Three illustrations of artificial life's working hypothesis. In Banzhaf, W., and Eeckman, F., eds., *Evolution and Biocomputation—Computational Models of Evolution*. Springer.
- Benton, M. J., ed. 1993. *The Fossil Record 2*. London: Chapman and Hall.
- Cairns-Smith, A. G. 1982. *Genetic Takeover and the Mineral Origins of Life*. Cambridge: Cambridge University Press.
- Cairns-Smith, A. G. 1985. *Seven Clues to the Origin of Life*. Cambridge: Cambridge University Press.
- Dawkins, R. 1976. *The Selfish Gene*. New York: Oxford University Press.
- Forrest, S., and Jones, T. 1994. Modeling complex adaptive systems with echo. In Stonier, R. J., and Yu, X. H., eds., *Complex Systems: Mechanisms of Adaptation*. IOS Press.
- Gould, S. J., and Lewontin, R. C. 1979. The spandrels of san marco and the panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London Series B* 205:581–598.
- Gould, S. J. 1989. *Wonderful Life: The Burgess Shale and the Nature of History*. New York: Norton.
- Gould, S. J. 1996. *Full House: The Spread of Excellence from Plato to Darwin*. New York: Harmony Books.
- Harland, W. B.; Armstrong, R. L.; Cox, A. V.; Craig, L. E.; Smith, A. G.; and Smith, D. G. 1990. *A Geological Time Scale 1989*. Cambridge: Cambridge University Press.
- Holland, J. H. 1992. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. MIT Press/Bradford Books, 2 edition.
- Holland, J. H. 1994. Echoing emergence: objectives, rough definitions, and speculations for echo-class models. In Cowen, G. A.; Pines, D.; and Meltzer, D., eds., *Complexity: Metaphors, Models and Reality*. Addison-Wesley.
- Holland, J. H. 1995. *Hidden Order: How Adaptation Builds Complexity*. Helix Books.
- Hraber, P. T.; Jones, T.; and Forrest, S. 1997. The ecology of echo. In *Artificial Life 3*. 165–190.
- Jones, T., and Forrest, S. 1993. An introduction to sfi echo. Technical Report 93-12-074, Santa Fe Institute, Santa Fe NM.
- Lindgren, K. 1992. Evolutionary phenomena in simple dynamics. In Langton, C.; Taylor, C.; Farmer, D.; and Rasmussen, S., eds., *Artificial Life II*. Addison-Wesley.
- McShea, D. W. 1996. Metazoan complexity and evolution: is there a trend? *Evolution* 50:477–492.
- Packard, N. 1989. Intrinsic adaptation in a simple model for evolution. In Langton, C., ed., *Artificial Life*. Addison-Wesley.
- Raup, D. M. 1988. Testing the fossil record for evolutionary progress. In Nitecki, M. H., ed., *Evolutionary Progress*. Chicago: The University of Chicago Press.
- Ray, T. S. 1992. An approach to the synthesis of life. In Langton, C.; Taylor, C.; Farmer, D.; and Rasmussen, S., eds., *Artificial Life II*. Addison-Wesley.
- Sepkoski, J. J., and Hulver, M. L. 1985. An atlas of phanerozoic clade diversity diagrams. In Valentine, J. W., ed., *Phanerozoic Diversity Patterns: Profiles in Macroevolution*. Princeton: Princeton University Press. 11–39.
- Sepkoski, Jr., J. J. 1992. *A Compendium of Fossil Marine Animal Families*, volume 61. 2 edition. Milwaukee Public Museum Contributions in Biology and Geology.
- SFI. 1998. Echo home page. <http://www.santafe.edu/projects/echo>.
- Smith, R. 1998. Personal communication.
- Valentine, J. W. 1985. Diversity as data. In Valentine, J. W., ed., *Phanerozoic Diversity Patterns: Profiles in Macroevolution*. Princeton: Princeton University Press. 3–8.
- Wolfram, S. 1984. Cellular automata as models of complexity. *Nature* 311:419–424.