On Modeling Life

Chris Adami

W.K. Kellogg Radiation
  Laboratory 106-38
California Institute of
  Technology
Pasadena, CA 91125

Abstract We present a theoretical as well as experimental investigation of a population of self-replicating segments of code subject to random mutation and survival of the fittest. Under the assumption that such a system constitutes a minimal system with characteristics of life, we obtain a number of statements on the evolution of complexity and the trade-off between entropy and information.

I Introduction

What are the defining characteristics of life? For biologists and those studying living systems, this is the question that predates other questions and generates passionate arguments; for there are several ways how to answer this question. On the one hand, there is the philosophical approach. We would invent a "Turing Test of Life" in the tradition of the empiricists: "Anything that appears to be alive, is alive." On the other hand, we may draw up a list of characteristics and decide the question case by case, through a look-up table of sorts. To the scientist interested in the foundations of life, however, neither of these approaches is satisfactory. Rather, the scientific method applied to the problem of obtaining the defining characteristics of life would demand that we construct a *model* of life, one that displays the necessary attributes to pass the "Turing Test of Life," and then *strip it down to the most simple system that still displays these characteristics*. Anything that, if removed from this most simple system would result in a drastic change, removing the lifelikeness, must be considered as a defining characteristic of this system. This does not preclude the existence of other systems passing the same test, with different characteristics. It does, however, establish the model under consideration as a baseline, and its characteristics as universal in its class. It is after this step has been taken that we can further enrich the model to learn more about life.

Clearly, this kind of approach is impossible using natural systems, as all life currently extant shows considerably more sophistication than needed just to survive. Thus, the construction of the minimally living system must be artificial. This has recently become possible with the advent of Tom Ray's *Tierra* [13–15], and systems based on his idea. Briefly, Tierra-like, (or "auto-adaptive genetic" [1]) systems consist of a population of self-replicating strings of information coded in a suitable manner that "live," and coevolve in an environment subject to mutation. The information contained in the self-replicating strings is the information necessary to carry out the task of self-replication, as well as any information that might give a certain genotype an advantage over another. Typically, this is information about the population itself, and the environment it lives in. Such a system displays a wealth of complexity and "life-like" characteristics [1–3,13–15]; enough to suggest answers to a number of fascinating questions relating to life.

Our approach then will be this: We assume that such systems of self-replicating strings *do* constitute a "minimal living system," and if not minimal then close enough to

support the notion that the results obtained using the model are reflections of its "alive-ness," rather than artifacts. While the theoretical results are independent of a particular implementation, the experimental ones are not. It is beyond the scope of this article to examine how close the particular system is to the abstract minimal living system, yet our results suggest that it is a good approximation, and in many respects universal. In the following, we present theoretical arguments and experiments supporting them that suggest answers to a number of fascinating problems, including the emergence of complexity in living systems, and the nature of the evolutionary process.

Let us attempt to isolate the main characteristics of the minimal living system. Pri-marily, we find that the cornerstone of this system is self-replication. The statistical mechanics of the ensemble of strings with and without self-replication are so vastly different (see section 2) that we can determine with confidence that we have isolated a defining characteristic of living systems. While this in itself is not surprising, it is specifically the self-replication of the information contained in the genome that is nec-essary (and that in this simple system is tantamount to replication of the entity). The consequences of this fact alone are manifold. Considering the evidence for genetic similarity in tRNA from the time of the origin of life to today [8], this capability seems to allow for the conservation of information over infinite periods of time in a noisy environment.

The keys for the evolution of complexity in this system are: bit-wise mutation acting on the genome of members of the population, and survival of the fittest. The former is again an obvious choice for a necessary condition: Without mutation, a population of self-replicating strings shows no patterns, no evolution, and no complexity. Clearly, mutation works against the force that attempts to preserve information, yet at the same time mutation *creates* information, or rather transfers it from the environment into the genome of the strings, where again it will be preserved by self-replication. We find that "survival of the fittest" is synonymous with *survival of the most populous.* Thus, fitness in this system is directly related to replication rate. In the absence of a mechanism that enables one string to prevent the replication of another (kill-strategy), it is conceivable that "fitness = replication rate" is a universal feature of such low-level organisms.

Another mystery of living systems is associated with the evolution of complexity and its apparent contradiction with the second law of thermodynamics, the law of increasing entropy. While it is known that dissipative systems involving microscopic irreversibility can evolve from disorder to order, the precise mechanism is often unclear. In this model, the evolution of complexity appears to proceed by a trade-off between entropy and information. Information is transferred from the environment into the genome of the population via chance mutations that allow the mutated string to exploit the information for a higher replication rate. As a consequence, the (information-theoretic) entropy of the population of strings drops by a commensurate amount until offspring of the mutated string have reestablished disorder and the previous level of entropy is (roughly) restored. The time between such events is distributed according to a power law, and we thus observe periods of stasis interrupted by "avalanches of invention." Again, if we accept this system as a model for the most primitive and basic living system, the evidence suggests that evolution is Darwinian in nature even though not gradual, but rather punctuated: the picture promoted by Gould and Eldredge [9]. Below, we shall give more details and results of experiments that support the claims made here.

2 Statistical Mechanics of Self-Replicating Bit-Strings

Let us approach the model system from a theoretical point of view and write down a few of the basic equations governing the dynamics of self-replicating strings. Consider a system of N strings of ℓ instructions, where each instruction can take on p values

($p = 2$ for binary strings, $p = 4$ for DNA, and $p = 32$ for Ray's Tierra). Also, let there be N_g different genotypes (different combinations of instructions) extant in the population, $N_g \leq N$, and let n_i denote the number of strings of genotype i. Then,

$$N = \sum_i^{N_g} n_i. \tag{1}$$

Let us further assume that the total number of strings in the population is constant: $dN/dt = 0$ (this constraint can be relaxed trivially). Also, denote the replication rate (offspring per unit time) of genotype i by ε_i. Defining the average replication rate (average "fitness")

$$\bar{\varepsilon} = \sum_i^{N_g} \frac{n_i}{N} \varepsilon_i \equiv \sum_i^{N_g} \rho_i \varepsilon_i, \tag{2}$$

where ρ_i is the genotype *density* of species i, we can immediately write down the equation that describes the time-evolution of n_i in a "mean-field" approximation

$$
\begin{aligned}
n_i(t+1) - n_i(t) &= (\varepsilon_i - \langle\varepsilon\rangle - R\ell)\ n_i(t) + \frac{N}{N_g} R\ell \\
&\equiv \gamma_i n_i(t) + C
\end{aligned}
\tag{3}
$$

with obvious definitions for γ_i and C, the "source" term. Here, we introduced the mutation rate R in units [mutations $\cdot$ (site $\cdot$ time)$^{-1}$]. Thus, $R\ell$ is the probability that a string of length ℓ is hit by a mutation per unit time, that is, per instruction executed. The constant term on the right hand side of Equation 3 is modeling mutations acting on genotypes $j \neq i$ and ensures $\dot{N} = 0$. Equation 3 can be solved in the equilibrium limit, $d\bar{\varepsilon}/dt = 0$:

$$n_i(t) = \frac{N}{N_g} \left\{ \left(1 + \frac{R\ell}{\gamma_i}\right) e^{\gamma_i t} - \frac{R\ell}{\gamma_i} \right\} \tag{4}$$

using an equilibrium boundary condition, $n_i(0) = N/N_g = \bar{n}$. For $C \approx 0$ this is simply

$$n_i(t) = n_i(0) e^{\gamma_i t} \tag{5}$$

where $\gamma_i = \varepsilon_i - \bar{\varepsilon} - R\ell$ is the "growth factor." Clearly, in equilibrium $\gamma_i = 0$ only for the best genotype and $\gamma_i < 0$ for inferior ones.

In order to make contact with statistical mechanics, define the *inferiority*

$$E_i = \varepsilon_{\text{best}} - \varepsilon_i \tag{6}$$

where $\varepsilon_{\text{best}}$ is the highest replication rate in the population. Then, $E_i \geq 0$ and we can define the ground-state ("vacuum") of the population by

$$\langle 0|E|0\rangle = \bar{E} = 0. \tag{7}$$

Let us also define the "fine-grained" (Shannon) entropy

$$S = -\sum_i^{N_g} \frac{n_i}{N} \log\left(\frac{n_i}{N}\right) \tag{8}$$

with the property

$$\frac{\partial S}{\partial E} \simeq \frac{1}{R\ell}, \tag{9}$$

which suggests that the system of self-replicating strings is analogous to atoms with eigen-energies E_i in a heat bath of temperature $R\ell$, which initiates transitions $E_i \to E_j$ via mutations. In fact, detailed balance arguments lead to the equation determining the equilibrium distribution $n_i(E_i, \bar{E})$

$$\frac{1}{n_i} - \frac{1}{n_j} = \frac{1}{C}(E_i - E_j) \tag{10}$$

with the solution

$$n_i(E_i, \bar{E}) = \frac{C}{E_i - \bar{E} + R\ell}. \tag{11}$$

In equilibrium, we find $\bar{E} \to R\ell$, universally. However, this equilibrium is disrupted by spontaneous phase transitions triggered by mutations creating genotypes with $\varepsilon_i > \varepsilon_{\text{best}}$ with a "latent heat" $L = \Delta\varepsilon = \varepsilon_{\text{best}}^{\text{new}} - \varepsilon_{\text{best}}^{\text{old}}$. After such an event the entropy of the system defined by Equation 2 drops, and then rebounds. In Figure 1 we compare the entropy of a system of strings subject to mutation with or without replication in the *Avida* environment [3].[1] If we start both systems in the uniform state $\bar{E} = 0$ (which implies $S = 0$: all strings are identical), we see that the self-replicating population reaches a plateau with $S < \log(N)$ while the non-self-replicating system reaches the maximum value rather quickly. In addition, the population of self-replicating strings can lower the entropy spontaneously, a feat that is impossible for the non-replicating population which must strictly adhere to the second law of thermodynamics.

3   Entropy and Information

Let us investigate more closely the time-evolution of entropy in systems of self-replicating entities. From the point of view of statistical mechanics rather than information theory, entropy is just the logarithm of the number of states in the volume of "phase-space" occupied by the population. Genetic phase-space, or genotype-space, is spanned by the combinatorics to assemble the basic building blocks (or bits), and is finite for finite-length strings. The population forms a "cloud" in genotype-space that moves slowly according to the fitness-landscape. Each genotype moves along a trajectory determined by a gradient term (adaptive process) and a noise term (mutations). As a consequence, the time-evolution of the genotype of a string is determined by the usual Langevin equation. While genotype-space is very high-dimensional, the number of states occupied by the population is by definition N_g, and thus,

$$S = \log(N_g), \tag{12}$$

---

[1] We chose this system because the number of strings can be kept strictly constant in Avida.
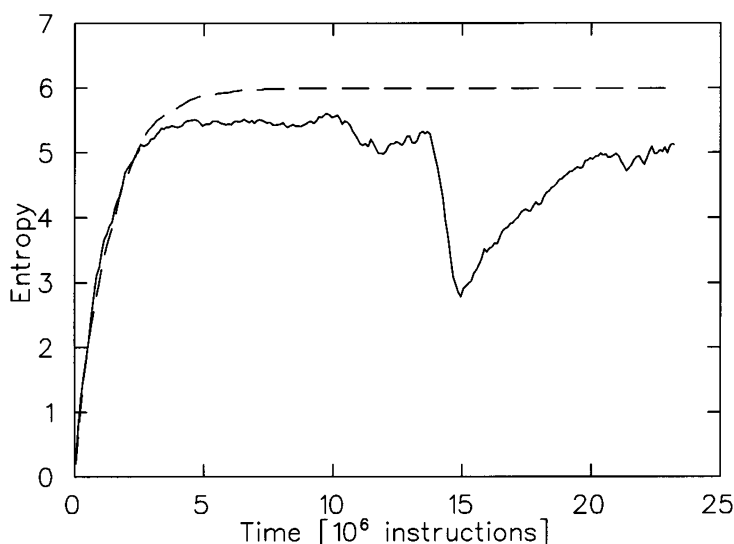
Figure I. Entropy of a population of 400 self-replicating (solid line) and nonreplicating (dashed line) strings started in the uniform state.

which is equivalent to the fine-grained entropy in the large N limit. The number of genotypes N_g is an equilibrium quantity, depending on mutation rate $R\ell$, total number of strings N, and most importantly, the average inferiority of the population. Clearly, only in infinite systems after infinite time is $\log(N_g) = \ell \log(p)$. For finite systems, the number of genotypes is determined by the rate of birth of new genotypes and the rate of extinction of existing ones. The rate of birth of new, viable genotypes is determined by the ratio of "cold" to "hot" spots in the genome. A cold spot on the string is an instruction that, if mutated, leads to either death or severe inferiority. Hot spots are instructions that are neutral under mutation and are thus responsible for the ε-degeneracy of genotypes. Thus, if there were no hot spots in the genome, there could not be any diversity, and in fact no evolution could take place, as was conjectured some time ago [11].

Let us denote the number of cold spots in a string of length ℓ by v_c. Then the birth rate of new genotypes is $R\ell N$, while the birth rate of *viable* genotypes is $(1 - v_c/\ell)R\ell N$. On the other hand, the extinction rate of existing (viable) genotypes depends (again in a mean-field approximation) on the average inferiority of the population, $\bar{E}$. Clearly, for zero inferiority there will be no extinctions, while extinctions are maximal during a phase transition, when the average inferiority is high. We thus find

$$\dot{N_g} = NR\ell \left(1 - \frac{v_c}{\ell}\right) - \bar{E}N_g \tag{13}$$

and thus in equilibrium

$$N_g = N\frac{R\ell}{\bar{E}}\left(1 - \frac{v_c}{\ell}\right). \tag{14}$$

We are now ready to examine the dropping of entropy observed through phase transitions. The decrease seen in Figure 1 is of course due to the discovery of information and a corresponding new $\varepsilon_{\text{best}}$. Through such a transition, $\bar{E} \to \bar{E} + \Delta\varepsilon$ and $\Delta S = \Delta\varepsilon/(R\ell)$.

However, with time $\bar{E}$ relaxes, and the entropy rebounds, but not necessarily to its previous equilibrium value. Consider, for instance, the situation where the information gained in the process requires more cold spots ν_c for storage than before. Then, a phase transition effectively freezes some spots that were originally hot spots, and thus reduces the effective length of the string. This is the situation in Figure 1; the entropy rebounds, but reaches a lower plateau than before, due to the incorporation of information in the genome. We are thus witnessing a curious trade-off between information and entropy. In fact, it has been conjectured [17] that the physical entropy is not given by just the fine-grained entropy, but rather is the sum of Shannon entropy $S(\rho)$ and algorithmic complexity $K(\rho)$, the latter being a measure of information stored in a string. In this scenario, complexity can emerge without violating the second law, simply by trading information for entropy while keeping the physical entropy constant.

4 Fractal Structure of Evolution

Earlier we described how macroscopically the system is in an equilibrium state only rarely disrupted by chance mutations that happen to improve the fitness of the population. Here, we would like to take a more quantitative look at this behavior, and work out the consequences for the structure of evolution.

More and more evidence is surfacing that certain quantities relating to living systems are distributed according to power laws, such as the frequency of extinction events of a certain size [12]. It has been known for a long time that there is a structure in the taxonomic system, since it was pointed out by Willis [16] that there is a very large number of taxa with only one or a few subtaxa, but only a few taxa with many subtaxa. Recently, there has been an extensive analysis of the frequency distribution of the number of taxa N that have n subtaxa [4, 5], with the result that $N \sim n^{-D}$ with $D \approx 2 \pm 0.5$. This fractal structure of evolution has been obtained from a painstaking analysis of fossil records from the Cambrian to the Tertiary (families within orders) as well as a selection of cataloged flora and fauna (subclades within clades, families within orders, and species within genera). Yet, an explanation for this structure could not be offered. Generally speaking, such a distribution can be expected if no scale determines the length of time that a certain species dominates a population, and the number of subfamilies is proportional to this period of time. This hypothesis can be tested using the minimal living system.

Recently [2], we suggested that self-replicating systems typically operate in a self-organized critical state [6, 7]. Drawing the analogy to the sandpile model, we could identify information as the agent that causes the self-organization, and which is transported via a dissipative process through the system. We also identified the growth-factor γ_i (see Equation 5) as the critical variable. In the equilibrium situation, the population is dominated by the current "best" genotype and its ε-degenerate offspring, all of which have $\gamma_i = 0$. An advantageous mutation however can disrupt this equilibrium by creating a new "best" genotype with $\gamma_i > 0$. Thus, the vacuum with $\bar{E} = 0$ is now a false one, and a phase transition must occur. This happens in the manner of the sandpile avalanches, with the new information transmitted through the system via the offspring. Gradually, all genotypes with a subcritical replication rate will become extinct and replaced: The system has returned to its critical state. A necessary condition for the identification of the self-organized critical state is a power spectrum of the fitness history (replication rate as a function of time) that is of the power-law type. A typical fitness history (normalized replication rate of the most abundant genotype in the population as a function of evolutionary time) obtained using the Tierra system is displayed in Figure 2.
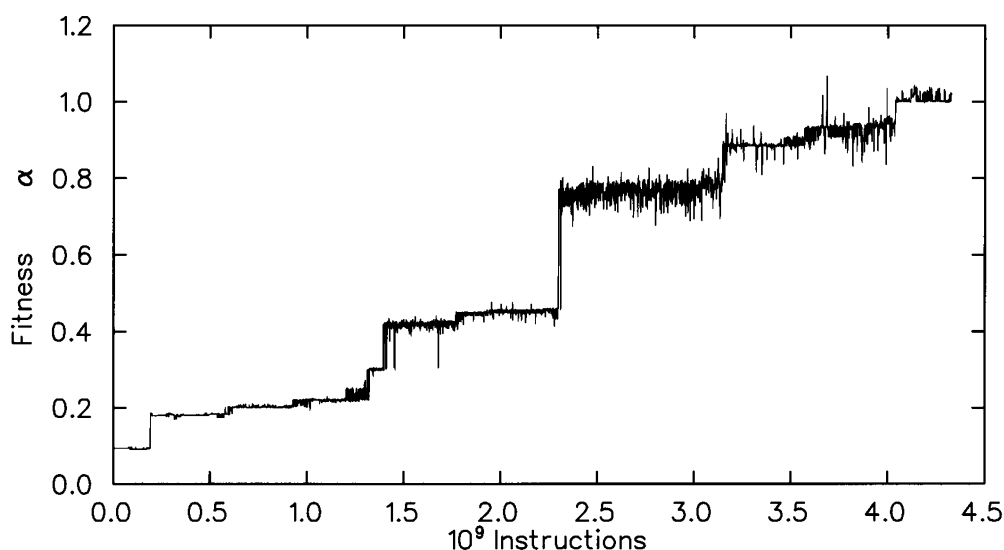
Figure 2. Fitness curve for a typical run. The (normalized) fitness α of the most successful (i.e., most populous) genotype is plotted as a function of time, measured every million instructions for a mutation rate $R = 0.65 \times 10^{-3}$.
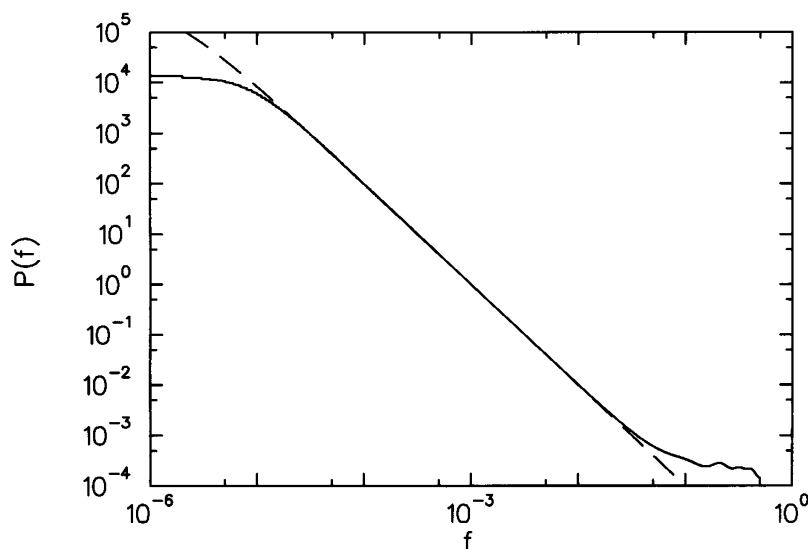

Figure 3. Power spectrum $P(f)$ of a typical fitness curve $\alpha(t)$ (Figure 2). The dashed line is a fit to $P(f) \sim f^{-\beta}$ with $\beta = 2.0 \pm 0.05$.

The associated power spectral density in Figure 3 shows a near perfect $1/f^2$ spectrum over 4 orders of magnitude. While a power-law spectral density is a necessary condition for the identification of a self-organized critical state, it is not sufficient. Indeed, a $1/f^2$ spectrum is typical for a simple random walk, and thus for the most basic dissipative process. Incidentally, the process of evolution can be described as a guided random walk in genotype space, albeit a *non-Markovian* one.

On the other hand, the distribution of waiting times between events of a certain size in a random walk model is exponential, rather than of the power law type. The latter
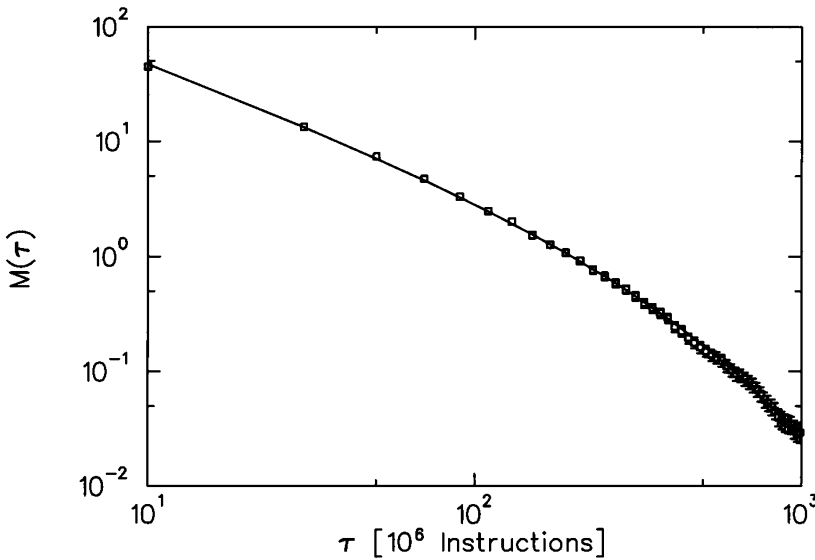
Figure 4. Integrated distribution $M(\tau)$ of times between phase transitions τ (length of epoch) measured in millions of executed instructions.

is expected if the time between such events is not controlled by any scale. The fundamental time scales in self-replicating systems of the sort discussed here are the gestation time (the time for a string to gestate one offspring) and the average time between two mutations hitting a single string (see [1]). Both of these time scales, however, are much smaller than the lifetime of single genotypes. Also, as discussed earlier in connection with the heat-bath analogy, the system can be understood as a superposition of a very large number of meta-stable states, with transitions between them induced by the mutations. In such a system, the time between avalanches or "events" is distributed according to a power law, as is the size and duration of the events [4, 5]. Using the Tierra system at a mutation rate of $R = 0.65 \times 10^{-3}$ (mutations per instruction executed), with a total number of 131,072 sites inhabited by typically 600–1,400 strings of length 60–150, we have measured the frequency distribution $N(\tau)$ of waiting times τ (lengths of "epochs") between phase transitions (which are signalled by a discontinuous jump in fitness as in Figure 2). Figure 4 shows the *integrated* distribution function

$$M(\tau) = \frac{1}{\tau} \int_{\tau}^{\infty} N(t) \, dt. \tag{15}$$

The integrated quantity is distributed with the same exponent as $N(t)$ (as can be shown by direct differentiation), but is more reliable due to improved statistics.

The data were obtained from 50 runs under identical conditions (save the random number seed) resulting in 512 waiting times, and fitted to a power law multiplied by a cutoff function, $N(\tau) \sim \tau^{-\alpha} \exp(-\tau/T)$.[2] The integrated distribution function can then be fitted by an incomplete gamma function

$$M(\tau) \sim \tau^{-1} \Gamma(1 - \alpha, \tau/T), \tag{16}$$

---

2 Because all runs were terminated after execution of between 500 million and 4 billion instructions, we cannot expect to measure the distribution of waiting times larger than 500 million instructions with satisfactory accuracy.

which yields $\alpha = 0.6 \pm 0.1$ and $T = 540 \pm 40$, where all times are in units of millions of executed instructions. We expect the exponent of the power law to be universal, while the cutoff of course is not.

There are a number of fascinating consequences of such a distribution. If the time between phase transitions, as well as the size of fitness increases during such transitions are distributed according to a power law, curves such as the one depicted in Figure 2 are fractal in nature, that is, they (ideally) appear similar *at all scales* ("Devil's Staircase"[3]). If this is the case, we can imagine that a more sophisticated model, one that can run many orders of magnitudes longer than the one we used, with unlimited fitness increases, would yield just such a curve, where each element of the curve when examined closely would have the same structure as the one in Figure 2. Then, we must admit that fitness increases of enormous magnitude can happen almost instantly, driven by microscopic point mutations *only*,[4] albeit extremely infrequently. Also, the model implies that at each such phase transition, a large number of genotypes goes extinct dynamically (see Equation 14) without external interference (this was also observed in [15]). Thus, Raup's [12] distribution of extinction events would equally follow.

On the other hand, it is tempting to believe that the longer a species dominates a population, the more subspecies will be generated, that is, $n \sim \tau^\delta$, with a model-dependent exponent δ. As a consequence, the frequency distribution of taxa with n subtaxa, $N(n)$, must also be given by a power law.

5 Conclusions

We proposed to model life using artificial living systems that are "minimal" in their characteristics. We determined self-replication of information, bit-wise mutation of the information (providing the noise), and an environment of information (providing the gradient of the fitness landscape) as necessary characters of such a model. The artificial life systems Tierra and Avida were used to determine universal characteristics of living systems, such as the evolution of complexity through transfer of information present in the environment into the genome, and its relation to the second law of thermodynamics. Furthermore, we were able to conjecture that living systems typically evolve toward a self-organized critical state and that this state determines fractal population structures. This also constitutes strong evidence for a punctuated equilibrium picture of evolution as favored by the fossil record.

References
1. Adami, C. (1994). Learning and complexity in genetic auto-adaptive systems. *Physica, D 80*, 154–170.

2. Adami, C. (1993). Self-organized criticality in living systems. KRL preprint MAP-167, Caltech.

3. Adami, C. & Brown, C. T. (1994). Evolutionary learning in the 2D artificial life system "Avida." In R. A. Brooks and P. Maes (Eds.), *Artificial life IV*: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems, (pp. 377–381). Cambridge, MA: MIT Press.

---

3 See, e.g., [10].

4 In all of the previous discussion we have ignored mutations of the crossover type, which effectively happen in systems such as Tierra, but which can be viewed as *consequences* of point mutations. Very little is as yet known about the impact of crossover mutations in the self-replicating systems described here, but we must assume that some of the phase transitions are, in fact, fitness improvements due to crossover mutations.

4. Burlando, B. (1990). The fractal dimension of taxonomic systems. *Journal of Theoretical Biology, 146*, 99–114.

5. Burlando, B. (1993). The fractal geometry of evolution. *Journal of Theoretical Biology, 163*, 161–172.

6. Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of $1/f$ noise. *Physical Review Letters, 59*, 381–384.

7. Bak, P., Tang, C., & Wiesenfeld, K. (1988). Self-organized criticality. *Physical Review, A38*, 364–374.

8. Eigen, M., Lindemann, B. F., Tietze, M., Winkler-Oswatitsch, R., Dress, A., and von Haeseler, A. (1989). How old is the genetic code? Statistical geometry of tRNA provides an answer. *Science, 244*, 673–679.

9. Gould, S. J. & Eldredge, N. (1993). Punctuated Equilibrium Comes of Age. *Nature, 366*, 223-227.

10. Mandelbrot, B. (1977). *The fractal geometry of nature.* New York: W. H. Freeman.

11. Maynard Smith, J. (1970). Natural selection and the concept of a protein space. *Nature, 225*, 563–564.

12. Raup, D. M. (1986). Biological extinction in earth history. *Science, 231*, 1528–1533.

13. Ray, T. S. (1992). An Approach to the synthesis of life. In C. G. Langton et al. (Eds.), *Artificial Life II*: Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems (Santa Fe Institute Studies in the Sciences of Complexity, Vol. 10), (pp. 371–408) Reading, MA: Addison-Wesley.

14. Ray, T. S. (1994). An evolutionary approach to synthetic biology: Zen and the art of Creating Life. *Artificial Life, 1*, 179–209.

15. Ray, T. S. (1994). Evolution, complexity, entropy, and artificial reality. *Physica, D 75*, 239–263.

16. Willis, J. C. (1922). *Age and area.* Cambridge, England: Cambridge University Press.

17. Zurek, W. H. (1990). Algorithmic information content, Church-Turing thesis, physical entropy, and Maxwell's demon. In W. H. Zurek (Ed.), *Proceedings of the 1988 Workshop on Complexity, Entropy, and the Physics of Information* (pp. 73–89) (Santa Fe Institute Studies in the Sciences of Complexity, Vol. 8). Reading, MA: Addison-Wesley.