Proceedings of the 2007 IEEE Symposium on
Artificial Life (CI-ALife 2007)

Studying Long-term Evolution with Artificial Life

Antonio Lafusa
Department of Computing, Macquarie University
Sydney, NSW 2109, Australia

Abstract **- This paper describes an *artificial ecology system* populated by self-replicating machines, which is able to capture many of the qualitative and quantitative characteristics observed in real ecosystems, from the fractal nature of taxonomic and phylogenetic trees to the pattern of originations and extinctions in long-term evolution.**

The self-replicating machines are composed by elementary particles. The composition and behaviour of a self-replicating machine is determined by its genetic code, and new species arises by means of random mutations.

The resulting dynamics is characterised by the co-evolution of species competing for resources. The interplay between the species and environment plays a crucial role for self-regulation: extinctions and originations are mechanisms used by the ecosystem to maintain the dynamical equilibrium that is favourable to life itself.

I. INTRODUCTION

The study of long-term evolution, which considers the whole history of life, is mainly based upon the time series and statistics drawn from the fossil record. Many *trends*, or patterns have been discovered looking at the time series of extinctions and originations of species, which can be used to make hypotheses on the underlying principles of evolution. However, due to the intrinsic incompleteness of the fossil record and many biases involved in the process of collecting the data, many important hypotheses are still debated.

One of the most important open debates concerns the presence of fractals [1] in evolution, e.g. in the time series of origination and extinctions of species over hundreds of million of years [2, 3]. Consolidated signs of self-similarity have been discovered in long-term evolution by looking at the *taxonomy* of life. Taxonomy is a classification of organisms into species, genera, families, etc. Such a classification also represents evolutionary relationships between organisms, e.g. all the species belonging to the same family have the same common ancestor. Starting from the common ancestor of all the living organisms and tracing all the branches, which represents the way life differentiated itself, it is possible to build a taxonomic tree, namely the *tree of life*.

It turns out that the tree of life is fractal [4-6]: a small portion of the tree resembles the whole structure (it is scale-free). It means that the same underlying mechanisms of speciation, competition and selection, which have shaped the tree-of-life, have the same characteristics and outcomes for both small and extremely long time scales.

Another open debate in evolutionary ecology concerns mass extinctions [7]. From the fossil record, it is evident that the history of life has been characterised by events of mass extinction, which led to the majority of species disappearing in a very short period, alternated by periods of relative calm in the species turnaround. An example is the extinction of the dinosaurs, which occurred circa 65 millions year ago.

The classical theory explains the events of mass extinctions invoking external - or *abiotic* - causes (like meteorites, volcanisms, etc.). However, from the study of complex, self-organising systems, a new theory is emerging which explains mass extinctions as the consequences of the intrinsic dynamics of the co-evolving species, without any need for external causes. This position highlights the role of *biotic* (internal) causes over that of *abiotic* causes.

The need for a new theory and new models of macro-evolution, which explicitly incorporated the co-evolution of species (within a network of interactions), originated also from the fact that quantitative studies highlighting fractals were not easy to explain in the context of the classic theory. One of the more important attempts to model evolution within a physics framework is that of self-organized criticality (SOC), in which the system brings itself toward a *critical state,* being characterised by self-similar statistical distributions - namely *power-laws* - resulting in many measures [8-10]. In the SOC models the interaction between species is sufficient to explain those characteristics that the authors wanted to model, without any need for abiotic causes [11-17]. Further, palaeontologists criticized these models, especially their basic, overly simplistic assumptions [18]. However, even if SOC models are not widely accepted, there is a general agreement that the contribution of these models was very important in order to focus the attention on biotic interactions.

In general, it is not possible to perform experiments that can validate one theory: it is not feasible (it takes millions of years to observe evolution) or not ethically possible (like perturbing the equilibrium of an ecosystem to study its response).

The solution proposed in this paper is to use self-replicating machines (instead of real-organisms) to perform *experiments that are impossible in real ecosystems*. These simulations allow the observation of a relative long time period, and the collected data are extremely detailed. This new approach could be used to validate different positions in the current debates.

In this paper, an *artificial ecology system* is presented, which simulates the chemistry of the environment and all the organisms, composed by molecules (*artificial molecules)* fluctuating in a two-dimensional space.

The artificial molecules are modelled as elementary particles, each of them belonging to a certain substance, which can either be aggregated to form *structures* (or macro-molecules) or free in the environment. The system is designed so that the structures are chemical machines that are able to replicate themselves, aggregating other particles from the surrounding space.

The self-replicating structures (or self-replicating machines) interact with the environment, consuming substances for building copies of themselves and releasing waste substances, according to their *genotype*. The genotype is the information that determines and differentiates the behaviour of the different species, and it is contained inside the organism's structure. During the self-replication the genotype can be mutated with a small probability, originating new species.

It is important to note that every macroscopic property of the system emerges from the interaction of the elementary particles, according to a so-called *local rule* (i.e. the particles can only interact with their nearest neighbours).

In subsequent paragraphs, after the definition of the system, the main results obtained by observing the long-term evolution will be reported, which are found to be compatible with real data extracted from both the fossil record and taxonomy.

II. A WORLD POPULATED OF SELF-REPLICATING MACHINES

Since the first von Neumann's investigation [19] on the possibility of self-replicating machines, many models have been developed in order to better understand the origin and the general principles of life or to investigate the possibility of using self-replication as a revolutionary technology in space exploration [20] and molecular manufacturing systems [21]. In the multiplicity of studies, the underlying models are mostly *cellular automata* [22-27] or *artificial chemistry* models [28, 29]. Recently some extremely simple physical implementations of self-replicating robots (using robotic modules) have also appeared in literature [30, 31].

The system presented in this paper is a Particle Based System (PBS), in which the particles float in two-dimensional space in a random walk and can dynamically form *structures* composed of many linked particles. In this context, *a self-replicating machine* is a dynamic structure that is able to aggregate other particles and organise them toward a copy of itself.

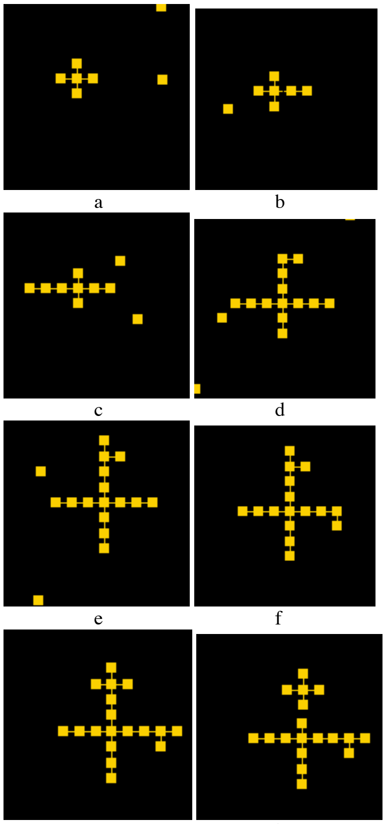
The particles exchange information through their links according to a local physics, with these interactions being the only form of control and organization. In particular, each particle has an internal state that is updated at regular intervals, according to the state of the linked particles. The maximum number of links for one particle is 4. The internal state of a particle is discrete, and we denote with k the number of possible states. Since a particle behaves like a *finite-state machine*, the result of interactions can be represented as a rule-table (similar to that used for cellular automata). At a given time t, the local configuration of a particle is given by its state and the state of the linked particles. The rule-table is a function that maps the local configuration at time t into the state of the particle at time $t+1$. At every time step, the particles are asynchronously updated.

The major characteristics of this system can be summarised as follows:

1. Each particle performs a random walk and is in one of the k possible states.
2. There are four wiring arms. Each wiring arm can be in one of two different states: *available* or *not available*.
3. A wiring (link) is spontaneously created between two particles if their wiring arms are both available and they are close enough to establish the relationship.
4. A wiring can be broken by one of the two particles.
5. A local rule determines the internal state of a particle, the availability of the wiring arms and the deletion of existing wirings (at time $t+1$), according to the internal state of the wired particles and the particle's own state (at time t). Therefore, it is a *finite-state machine*.
6. Each particle cuts an established wiring with a small probability P_D, which is a probability per wiring per time-step. This also causes a self-replicating machine to die because of instability, and all the particles to be liberated into the environment.

The definition of a rule completes the specification of the system and determines the behaviour and possible complex structures (see [32] for the particular rule-table that is used to achieve self-replicating structures). If we adopt this definition, the result is a world populated of self-replicating machines.

The series of figures 1a-h is a representation of a self-replicating machine that grows and produces copies. Each particle is represented by a little square.



a b

c d

e f

g h

Figure 1. *a)* A self-replicating machine in its minimal configuration of five modules. *b)* The structure starts to grow. *c-d-e-f)* The structure continues to aggregate other particles from the surrounding space. *g)* On the top arm, the formation of a child is completed. *h)* The newborn structure is detached and starts its own life.

The purpose of this system is to study co-evolving species in a complex environment: we need to differentiate the machines (i.e. obtaining different species) and we also need a number of different resources that can be produced and consumed by the machines in order to establish an interplay with the environment.

In order to accomplish these requirements, we modify the design of the system, adding a new property to the internal state of a particle: the *substance*. Therefore, one of the parameters of this system will be the number of substances (in the simulation we use values between 3 and 20). In each time-step of simulation, the concentration of a given substance in the environment is defined as the fraction of particles belonging to that substance.

The next step is to define the genetic code of a self-replicating machine, which can be mutated in order to produce new species with different behaviour. The genetic code will be stored in the internal state of the particle, and determines which substances can be bonded with to form a structure, and how these substances will be transformed. Moreover, a structure that aggregates a free particle from the environment transfers its genetic code to the aggregated particle. A mutation occurs in the genotype with a small probability P_M.

From figure 1, it is clear that a self-replicating machine generates other copies by means of a process of growth and division. If, during the growth, a mutation occurs, then the newborn structure detached from the parent belongs to a different species.

In figure 2, a snapshot of a simulation.

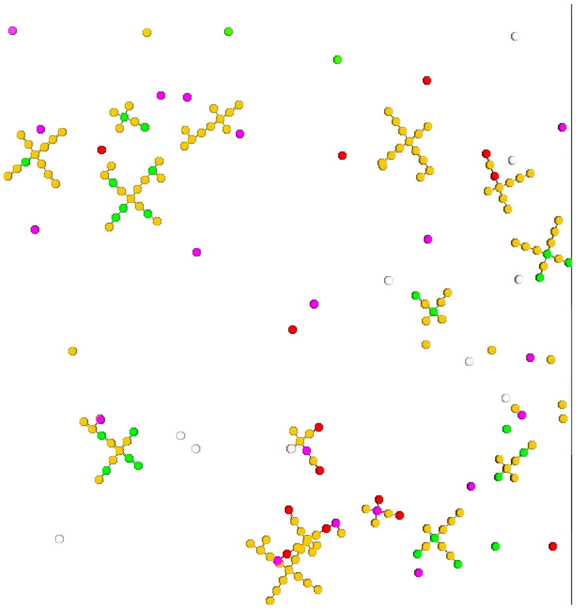


Figure 2. A snapshot of the simulation, showing a small portion of space with self-replicating structures and free particles. The colours represent different substances.

The following table summarises all the parameter of the system:

Parameter name	Description	Values used in the simulations
N	Number of particles in the system	800-10000
M	Number of substances	3-20
P_M	Probability of genotype mutation	0.01-0.25
P_D	Probability of wiring decay between two particles	0.0005-0.01

Table 1. Parameters of the system and values used in the experiments.

III. LONG-TERM EVOLUTION: EXPERIMENTAL RESULTS

In the previous section, the behaviour of the system has been defined at a microscopic scale, i.e. as a particle based system. In this section, we study the emergent long-term behaviour. The objective is to compare the experimental results of the simulations with data from fossil records and taxonomy, in order to validate our model of evolutionary ecology.

First, we need to introduce a particular mathematical property, which describe the scale invariance (i.e. the fact that we observe the same characteristics for every scale) of many phenomena in nature, namely the *power law* [33]. The mathematical form of a power law, between two quantities X and Y is:

$$Y = c\,X^{K}$$

where c is a constant of proportionality and K is the *exponent* of the power law.

If Y is the probability to observe the value X in a given phenomenon, then this phenomenon is said to be power-law distributed. For example, from the study of taxonomy, it is known that the number of species belonging to a given family is power-law distributed with an exponent between 2 and 3 [4-6].

If we plot the X-Y data, transforming the coordinates from linear to logarithmic, a power law is very easy to recognise because it appears as a *straight line*. The slope of the straight line on the X-Y plot is the exponent K.

Before describing the quantitative analysis, we will briefly give the definition of species and family used in this particular context: a *species* is a collection of machines having the same genotype; a *family* is a collection of species that have a common portion of the genotype, precisely the part of the genotype that determines the composition of a machine in terms of substances.

For each experiment, the following time series are studied:

a. *The number of self-replicating machines* present at a given time step (an example is shown in figure 3).
b. *Diversity:* defined as the number of species present at a given time step and number of families (see figure 3).
c. *Extinctions:* defined as the number of extinct species, accumulated over time windows of 25000 time-steps.
d. *Originations:* defined as the number of new species, accumulated over time windows of 25000 time-steps.
e. Concentrations of the substances in the environments.

The quantitative results obtained for the artificial ecology system are:

I. *Phylogenetic tree*: we keep track of the sub-species originated by each species, being able to reconstruct the history of the evolution as a tree, which turns out to be a fractal (or *scale-free*) tree. The degree (i.e. the number of direct descendants of a species) is power-law distributed with exponent –2.4. The same tree is also reconstructed for the families, resulting in the same degree distribution and the same power-law exponent (see figure 4).
II. *Species/family lifetimes*: the species lifetimes and the family lifetimes are power-law distributed with exponent –2 (see figure 5).
III. *Distribution of species-per-family*: the number of species per family is power-law distributed with exponent –2.4 (see figure 6).
IV. Power spectra analysis: the Power Spectral Density (PSD) of a time series is a useful transformation that has been recently applied to extinctions, originations and diversity of species in the fossil record. We applied the same method to the time series obtained from our model, and it turns out that the power spectra of *number of machines*, *originations*, *extinctions* and *diversity* follow a power law with exponent –1.76 (figures 7-10).

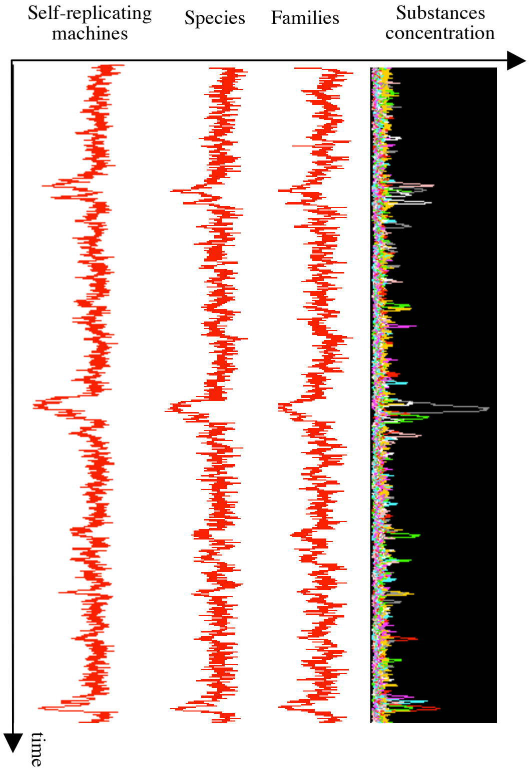
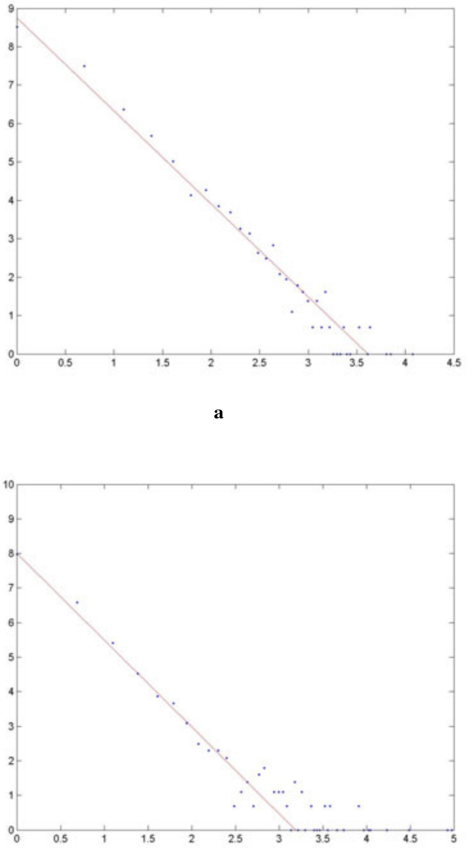


Figure 3. A time window of the simulation, which shows the number of machines, number of species, number of families and substances concentration over time (10^6 time-steps). It is easy to see that the system is characterised by period of calm alternated by bursts of extinctions and originations.



a

b

Figure 4. a) Degree distribution of the species tree in log-log scale (i.e. the number of direct descendants of a given species). b) Degree distribution of the families tree in log-log scale. The straight lines are linear fit to the data and the slope is –2.42 for (a) and –2.5 for (b).

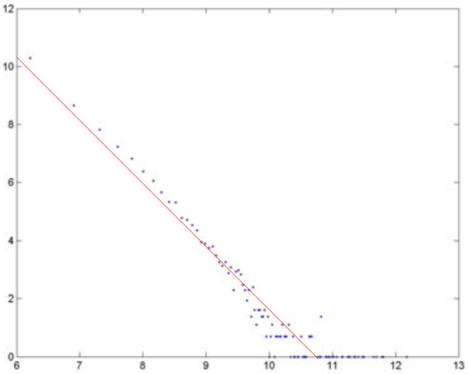


Figure 5. Lifetimes distribution of species in a log-log scale. The straight line corresponds to an exponent circa –2.

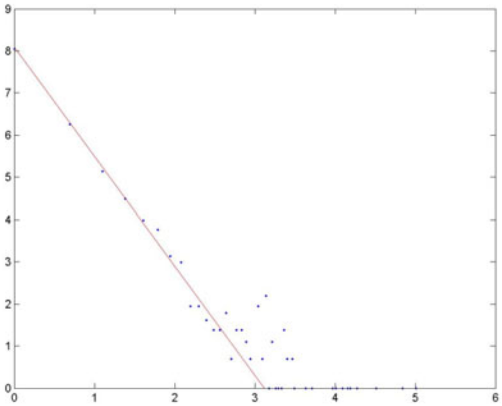


Figure 6. The distribution of the number of species per family in log-log scale. The straight line is a linear fit and the slope is –2.5.

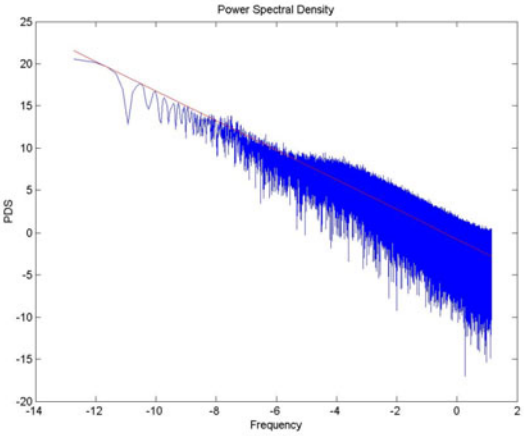


Figure 7. PSD of the number of machines time series in a log-log scale. The straight line is a linear fit with slope –1.76.

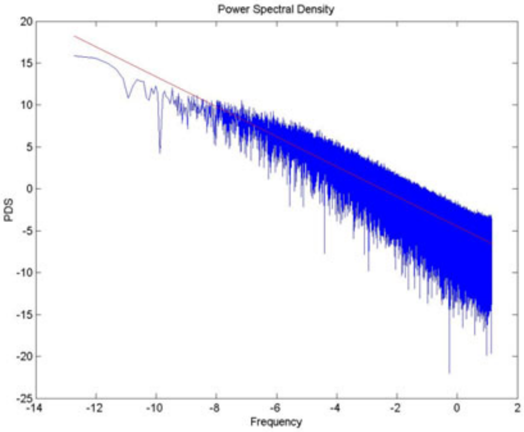


Figure 8. PSD of the number of species time series in a log-log scale. The straight line is a linear fit with slope –1.76.

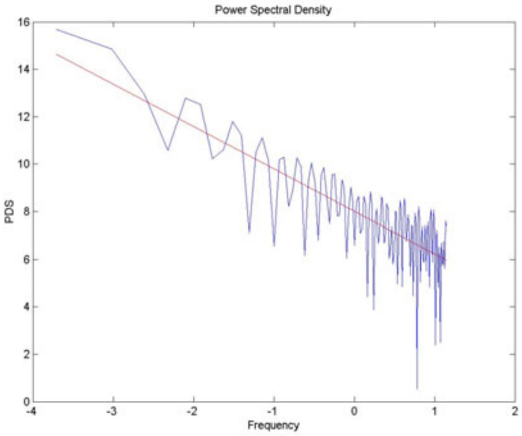


Figure 9. PSD of the number of species extinct, accumulated over time windows of 25000 timesteps. The straight line is a linear fit with slope –1.76.

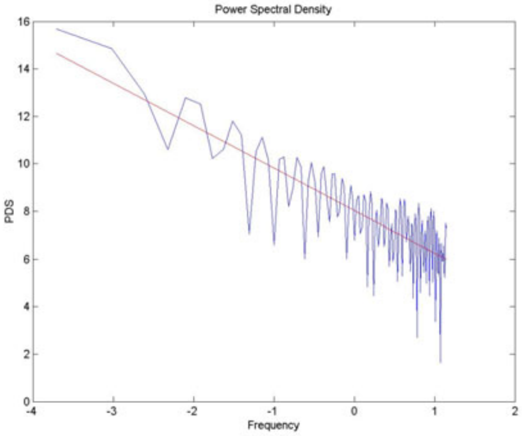


Figure 10. PSD of the number of new species originated, accumulated over time windows of 25000 timesteps. The straight line is a linear fit with slope –1.76.

IV. VALIDATION OF THE MODEL USING DATA FROM THE FOSSIL RECORD AND TAXONOMY

In order to validate the model, we summarised all the power-laws found in the experiments with self-replicating machines, compared with the data found in real ecosystems. From the following table, we can see that all the power-law exponents found in the artificial ecology system are in strong agreement with those found in the fossil record and taxonomy.

Power law exponents	Model	Known in literature
Lifetimes	2	Between 1.5 and 2 [17]
Taxonomy (species per families)	2.5	Between 1.84 and 2.41 [4, 5] Between 2 and 3 [6]
Phylogenetic tree	2.42 (species) 2.5 (families)	N/A
Extinctions power spectra	1.76	Between 1.7 and 2.14 [34]
Originations power spectra	1.76	N/A
Diversity power spectra	1.76	N/A

Table 2. Comparison of power-law exponents between the artificial ecosystem and fossils/taxonomic data. The (negative) sign of the exponents is omitted. Many of the cited results are obtained with different techniques and data sets. Both the exponents and the power-law form are still debated.

V. SELF-REGULATION AND MASS-EXTINCTIONS

One of the most interesting qualitative properties of the artificial ecology system is self-regulation. In particular when an excess of a given substance perturbs the environment, initially a phase of extinction occurs, which can reach the size of a mass extinction (most of the species involved). However, when the mass extinction reaches its intensity peak, many ecological niches are made available and the competition is very low. Consequently, a phase occurs of proliferation of new families and species, which consume all the excess of substances and bring the system in a new equilibrium. This new equilibrium is characterised by a balance between what is consumed and produced, so that the concentrations of all the substances are approximately constant.

The following is an example of self-regulation: If a new species X transforms a substance A into another substance B, and the population of this species grows very fast, it is likely to produce an excess of B in the environment, which may limit the growth of other species or bring other species to extinction. To recover the original equilibrium, the system normally produces two different behaviours: the first one is the extinction of X, after a period of rapid proliferation, because the species X depends on other species that get extinct in the modified environment; the second possible behaviour is the emergence of a new species Y that absorbs the excess of substance B, balancing the activity of the species X.

In general, examining all the experiments, the whole evolution in time is characterised by periods of calm, which are interrupted by bursts of extinctions and originations (see figure 3). In figure 11, one of these phases of instability is studied in detail, and the resulting time series are explained in terms of self-regulatory mechanism.

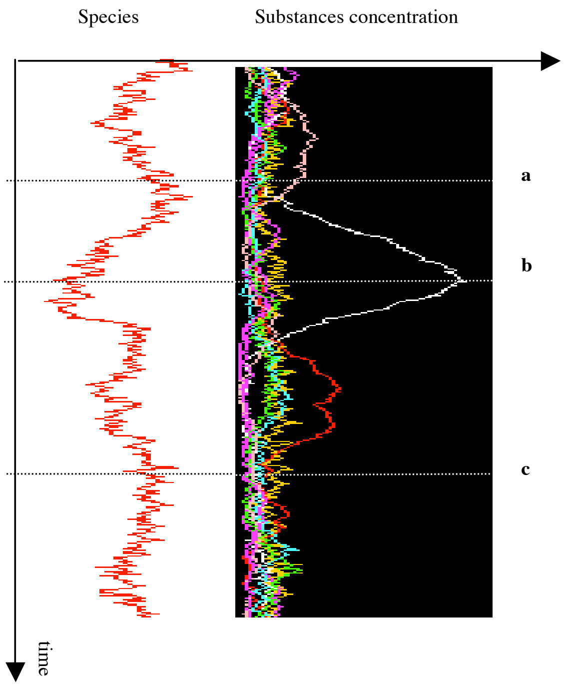


Figure 11. An example of self-regulatory mechanism: in (a), the species X emerges, which produces the white substance; from (a) to (b), the white substance is produced in excess, the equilibrium is perturbed and the system enters a phase of high instability, which corresponds to a mass-extinction. In (b) a new species Y emerges, which consumes the white substance, decreasing its concentration in the environment. Now Y is rapidly proliferating in the new conditions and originates other subspecies. One of the subspecies of Y transforms the white substance into the red substance, producing an excess of red substances that is not balanced by other species. Finally, in (c) a new species Z consumes the red substance in excess and the system enters a new phase of stability, characterised by a dynamical equilibrium between X, Y, Z and all the other species that survived to the previous mass extinction.

VI. CONCLUSION

The study of evolutionary ecology is limited by the fragmented and biased nature of the fossil record and, in general, it is not possible to validate one theory using just data sets that come from palaeontology. A new approach, proposed in this paper, is to simulate a world populated of self-replicating machines, interacting and co-evolving in their artificial environment. As a consequence, it is possible to perform experiments and observations that are impossible in real eco-system, being able to choose between different positions in the current debates and decide the direction of future investigations.

An important debate in the study of evolution is about mass extinctions. One position is to explain them with external (*abiotic*) causes, like meteorites; the other position is to explain mass extinction as a result of the intrinsic dynamics, which has internal (*biotic*) causes like a change in the equilibrium and relationships between different species.

Since the design of the artificial ecology system presented in this article is focused on the internal dynamics of co-evolving species, with no abiotic perturbation, the observation of mass extinctions is explainable without any need for abiotic causes.

Another debate is about the fractal nature of evolution. Even if there is general agreement about the presence of fractals in the evolutionary ecology patterns, some of the claims made on fossil records have been subsequently identified as artefacts [3], moreover the mechanisms that lead to fractals in evolution are not known and all the current studies remains hypotheses.

As a preliminary interpretation of these experimental results, the fractal nature of evolution seems to be a consequence of co-evolution and interplay with the environment. In particular, and accordingly with real data, the tree-of-life is scale-free and many other signs of self-similarity are found in the distribution of the species lifetimes and power spectra of all the time series extracted from these experiments.

A well known - and strongly debated - theory of ecology is *Gaia theory* [35], which acknowledges the fact that the Earth is a self-regulating system, which is similar to a living organism in the way it adjusts the internal parameters (e.g. temperature, concentration of CO_2, etc.) to safeguard the conditions for life. From a systemic point of view, it is a feedback chain between the environment and life. As pointed out in the previous section, this behaviour has been observed continuously during the experiments.

As a conclusive note, the high level of agreement between the model and real data indicates that this approach - studying the ecology of self-replicating machines - is useful in discriminating between different theories. Future developments can be focused on measuring other quantities that are suitable to be verified with real data, along with the observation of the mechanisms, often hidden in nature, which are responsible of those behaviours.

REFERENCES

[1] Mandelbrot, B. and Wheeler, J. A. The fractal geometry of nature. American Journal of Physics (1983).

[2] R.V. Sole, S.C. Manrubia, M. Benton and P. Bak, Self-similarity of extinction statistics in the fossil record, Nature, 388, 764-767 (1997).

[3] J.W. Kirchner and A. Weil, No fractals in fossil extinction records, Nature 395, 337-338 (1998).

[4] Burlando, B. The fractal dimension of taxonomic system. *Journal of Theoretical Biology*, 146, 99-114 (1990).

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

[6] Caretta-Cartozo C., Barthelemy, M, Garlaschelli D., Ricotta C. Caldarelli G. Scale- free properties of taxonomy data (unpublished).

[7] Raup, D. M. Extinction: Bad genes or bad luck? Oxford University Press (1993).

[8] P. Bak, C. Tang and K. Wiesenfeld. Self-organized criticality: an explanation for 1/f noise. Phys. Rev. Lett. 59,381–384 (1987).

[9] P. Bak, C. Tang and K. Wiesenfeld. Self-organized criticality. Phys. Rev. A 38,364–374 (1988).

[10] P. Bak. How nature works. Springer, New York (1996).

[11] P. Bak and K. Sneppen, Punctuated equilibria and criticality in a simple model of evolution, Phys. Rev. Lett. 71, 4083-4086 (1993).
[12] M. Paczuski, S. Maslov and P. Bak, Avalanche dynamics in evolution, growth and depinning models, Phys. Rev. E 53, 414-443 (1996).
[13] R.V. Sole and J. Bascompte, Are critical phenomena relevant to large scale evolution?, Proc. Roy. Soc. Lond. B 263, 161-168 (1996).
[14] R. V. Sole and S.C. Manrubia, Extinction and self-orgnized criticality in a model of large-scale evolution, Phys. Rev. E 54, R42-R45 (1996).
[15] R. V. Sole and S.C. Manrubia, Criticality and unpredictability in macroevolution, Phys. Rev. E 55, 4500-4507 (1997).
[16] S.C. Manrubia and M. Paczuski, A simple model of large scale organization in evolution, Int. J. Mod. Phys. C bf 9, 1025-1032 (1998).
[17] K. Sneppen, P. Bak, H. Flyvbjerg and M. H. Jensen. Evolution as a self-organized critical phenomenon. Proc. Natl. Acad. Sci. 92, 5209–5213 (1995).
[18] R. E. Plotnick and J. J. Sepkoski. A multiplicative multifractal model for originations and extinctions. Paleobiology, 27(1), 126-139 (2001).
[19] J. von Neumann. Theory of Self-Reproducing Automata. University of Illinois Press, Illinois. Edited and completed by A.W.Burks (1966).
[20] R. A. Freitas, W. A. Zachary. self-replicating, growing lunar factory. 5th Princeton/AIAA/SSI Conf. Space manufacturing. Princeton, N.J (1981).
[21] K.E. Drexler. Nanosystems: Molecular Machinery, Manufacturing, and Computation. New York: Wiley (1992).
[22] C. G. Langton. Self-reproduction in cellular automata. Physica D 10:.135-144. (1984).
[23] J. D, Lohn and J. A. Reggia. Discovery of self-replicating structures using a genetic algorithm, Proceedings of IEEE International Conference on Evolutionary Computation, 678-683 (1995).
[24] H. Sayama. Self-Replicating Worms that Increase Structural Complexity through Gene Transmission. In Bedau M. A., McCaskill J. S., Packard N. H., and Rasmussen S. eds. Artificial Life VII: Proceedings of the Seventh International Conference, 467-476. MIT Press (2000).
[25] H. Sayama. Self-protection and diversity in self-replicating cellular automata, *Artificial Life* 10: 83-98 (2004).
[26] E. Bilotta, A. Lafusa, P. Pantano. Life-like self-reproducer*s*, Complexity **9-1** (2003).
[27] E. Bilotta, A. Lafusa, P. Pantano. Is self-replication an embedded characteristic of artificial/living matter?, Artificial Life VIII*: Proceedings of the Eighth International Conference on the Simulation and Synthesis of Living Systems* (2002*).*
[28] T.J. Hutton. Evolvable self-replicating molecules in an artificial chemistry. *Artificial Life* 8(4): 341-356 (2002).
[29] A. Smith, P. Turney, and R. Ewashuk. Self-replicating machines in continuous space with virtual physics. *Artificial Life* 9:21-40 (2003).
[30] V. Zykov, E. Mytilinaios, B. Adams, and H. Lipson. Self-reproducing machines. *Nature* 435, 163–164 (2005).
[31] S. Griffith, D. Goldwater, J. M. Jacobson. Self-replication from random parts. *Nature* 437, 636 (2005).
[32] A. Lafusa. Self-replicating machines in a particle based system. PhD working paper, Macquarie University, Sydney (unpublished).
[33] M. E. J. Newman Power laws, Pareto distributions and Zipf's law, *Contemporary Physics* **46**, 323-351 (2005)
[34] M. E. J. Newman and G. J. Eble. Decline in extinction rates and self-similarity in the fossil record. SFI working paper 98-04-031 (1998). Available at www.santafe.edu.au/research/publications/workingpapers/98-04-031.ps
[35] J. Lovelock. Gaia: A new look at life on Earth. Oxford University Press (2000)..