

## EXPLORATION OF STYLIZED FACTS IN THE ARTIFICIAL LIFE SYSTEM AVIDA

Shinta Koyano<sup>1</sup>, Lukáš Pichl<sup>2</sup>

**Abstract:** Population dynamics in the evolution, extinction, and re-evolution of various logic-function performing organisms is studied in the artificial life system, Avida. Following the work of Yedid (2009), we design an experiment involving two extinction regimes, pulse-extinction (corresponding to a random-kill event) and press-extinction (corresponding to a prolonged episode of rare resources). In addition, we study the effect of environmental topology (toroidal grid and clique graph). In the study of population dynamics, logarithmic returns are generally applied. The resulting distributions display a fat tail form of the power law: the more complex the logic function (in terms of NAND components), the broader the full width at half a maximum of the histogram. The power law exponents were in sound agreement with those of “real-life” populations and distributions. The distributions of evolutionary times, as well as post-extinction recovery periods, were very broad, and presumably had no standard deviations. Using 100 runs of 200,000 updates for each of the four cases (about 1 month of central processing unit time), we established the dynamics of the average population, with the effect of world topology.

**UDC Classification:** 004.8 **DOI:** <http://dx.doi.org/10.12955/cbup.v4.852>

**Keywords:** Artificial life, evolution, extinction, re-evolution, power law distribution.

### Introduction

The artificial life system, Avida (Lenski, Ofria, Pennock, & Adami, 2003), is a software platform in which digital organisms, namely computer programs from elements of a limited instruction set, self-replicate, evolve, and compete for computational time on a model hardware. The self-replication is a feature of the ancestor organism at the start of the computer simulation with the program execution resulting in copying of the ancestral organism, which splits into “parent” and “offspring” outputs. The digital organisms are capable of performing various logic functions, such as NOT, NAND, AND, OR\_N, OR, AND\_N, NOR, XOR, and EQU (Adami, Ofria, & Collier, 2000). Since the instruction set contains only NAND from the set of standard binary logical operators, the other functions are expressed in terms of its combinations. As the complexity of the logic function increases, organisms are rewarded with extra computational time, which allows them to make more offspring copies, and hence, effectively reproduce. The reproduction phase includes a genetic operator of mutation, by which the particular program instruction may be changed randomly with a small probability, thus driving evolution within the system. Although sexual reproduction has been introduced into Avida as an option (Covert III, Carlson-Stevermer, Derrberry, & Wilke, 2012), the default version used in this work did not include parental mating, and hence the cross-over operator was excluded in this study. Each digital organism is associated with a point on a space grid, where it lives. Organisms may be killed at random when their reproducing neighbors place their offspring at a particular location on the world grid. Two geometries are considered in this study: a 2D topology (60 × 60 toroidal grid) and a clique (graph of 3600 nodes, where each pair of nodes is connected by an edge link). Detailed explanation of the Avida software can be found in Lenski, Ofria, Pennock, & Adami, (2003). In this research study, we use Avida to examine the effect of world topology on the re-evolution of organisms, by performing various logic functions, especially the EQU function (includes five NANDs). We closely follow the work of Yedid, Ofria, & Lenski (2009) and Yedid, Stredwick, Ofria, & Agapow (2012), in which the extinction episode is introduced to the population in two modes: pulse extinction (random kill event) and press extinction (period of resource deprivation). The configuration files for the experiment, except for the effect of the world topology, are the same as those of Yedid et. al (2012). Avida has already been used in a number of diverse research studies, ranging from long-term adaptation studies or emergence of altruism (Clune, Misevic, Ofria, Lenski, Elena, & Sanjuan, 2008, Clune, Goldsby, Ofria, & Pennock, 2011), to the elucidation of the effects of the genetic operators (Nelson & Sanford, 2011; Goldsby & Cheng, 2008).

<sup>1</sup> Shinta Koyano, International Christian University, Mitaka, Tokyo, Japan

<sup>2</sup> Lukáš Pichl, International Christian University, Mitaka, Tokyo, Japan, [lukas@icu.ac.jp](mailto:lukas@icu.ac.jp)

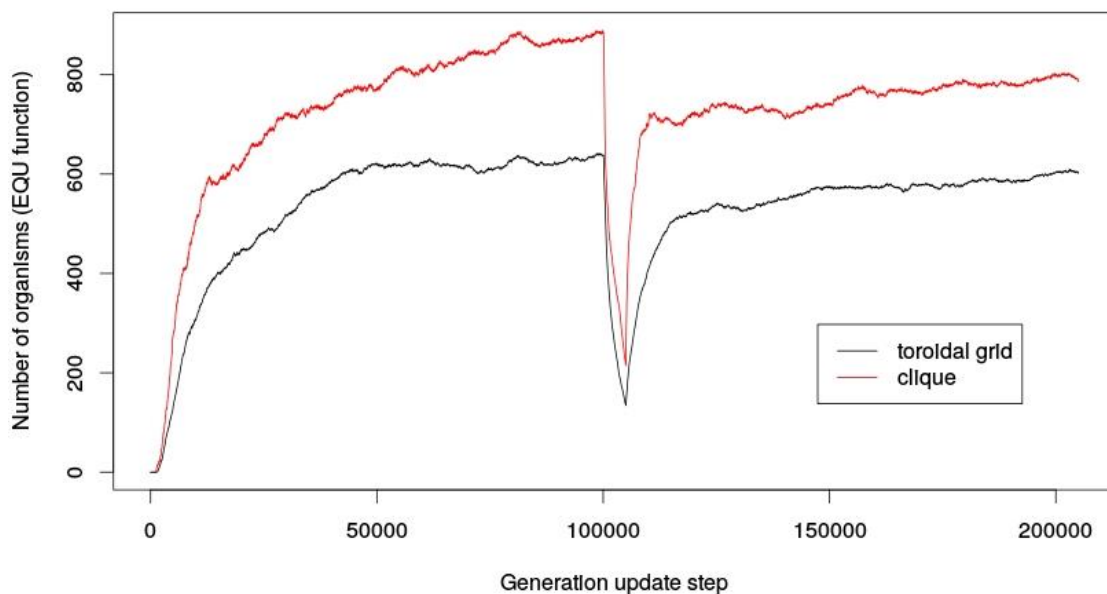
## Computational Experiment

All experiments were performed on a Dell PowerEdge T420 server, with 2 Intel Xeon E5-2407 2.2 GHz processors and 8 cores, running the GNU/Linux Fedora 23 operating system. Avida simulation studies were time consuming, since a single run process occupying one core lasted in the order of 3–4 hours for 100,000 updates period. The purpose of the present study was three-fold: (1) reproduce the work by Yedid et al. (2009, 2012); (2) study the effect of environmental geometry on the evolution or re-evolution of digital organisms (cf. Chow, Wilke, Ofria, Lenski, & Adami, 2004), namely the clique geometry and the toroidal grid geometry; and (3) analyze the distribution of the recovery times of computational functions.

Computer simulations were run with four basic configurations involving a pulse or press extinction regime, and a toroidal grid or clique geometry. Each configuration ran for a period required for 200,000 updates, involving 100,000 steps to evolve the complex features before extinction and 100,000 steps for the recovery. The evolution and recovery of the most complex function, EQU, was evident for an average of 100 runs in the press extinction regime (Figure 1) and the pulse extinction regime (Figure 2). The press extinction did not annihilate all digital organisms, but the re-evolution was much slower and could not be achieved within the computer simulation range for either of the two world topologies considered. However, as seen in Figure 2, a complete dip in the pulse extinction, and much faster recovery, especially for the clique geometry, showed the topology had greater inter-connection, as there was no notion of the local neighborhood, in contrast to the toroidal grid case. In addition, the clique geometry appeared a better mediator for growth of the EQU population, since the red curve was above the black curve throughout the simulation. The same reasoning applied to the clique population, i.e., it did not suffer from the limited size of the local neighborhood. This resulted in a more efficient spread of “fitter creatures” in the population.

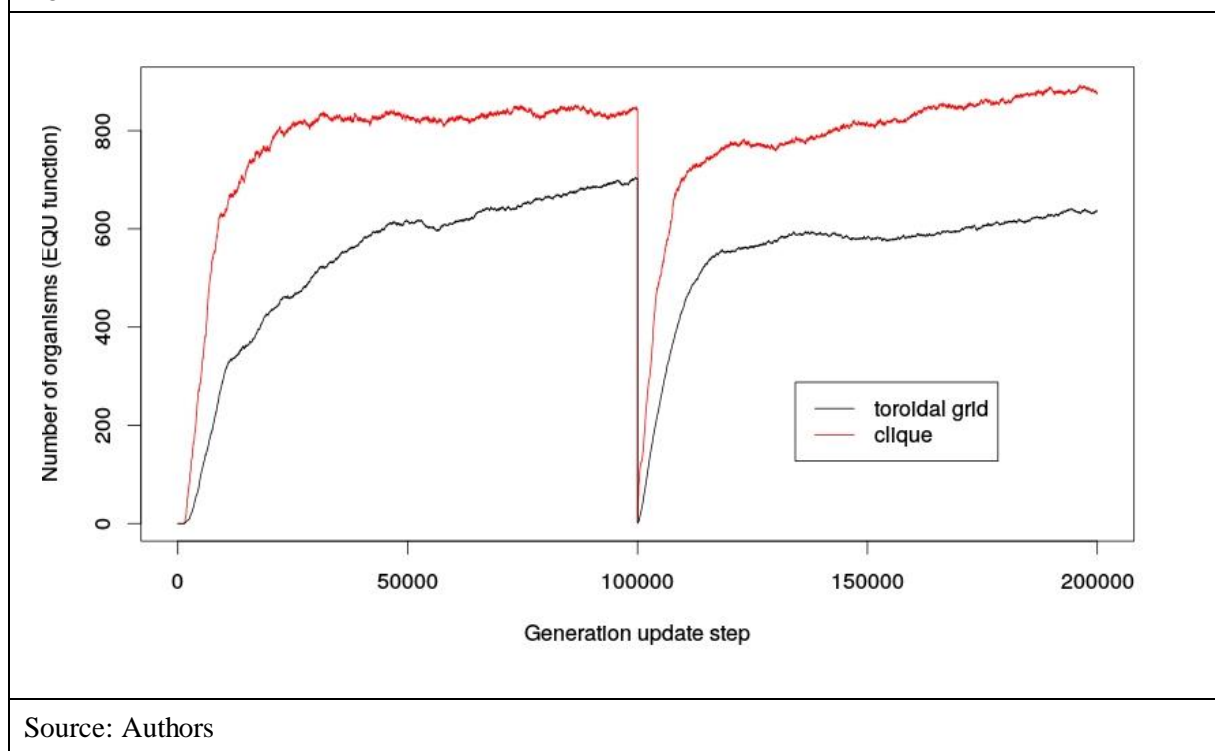
Initially, we planned to aggregate the individual recovery times for each type of digital organism, each extinction mode, and every geometry. However, the variation of the results corresponded to a power-law distribution without a standard deviation. The Results section provides examples (Tables 1 & 2).

Figure 1: Population dynamics for organisms performing EQU function in the press extinction regime



Source: Authors

Figure 2: Population dynamics for organisms performing EQU function in the pulse extinction regime



## Results and Discussions

Figures 1 and 2 show the main results of this paper. The following is a summary of the simulation results for the press extinction regime (both the toroidal grid and clique geometries). The results varied widely, with standard deviation estimates exceeding the mean recovery times for each of the nine logic functions computed by the digital organisms. The mean values of the recovery times were not well represented in the data, and were vastly different to the medians, with occasional outliers in the statistics. The results demonstrated the power-law, fat-tail probability density function, which decayed in the form of  $x^{-a}$ , where  $a$  lay in the interval between 2.5 and 3.0. Such a feature is characteristic of digital life environments, and this applied to both the complex feature evolution and the recovery times, which may not necessarily display a finite standard deviation. In other words, each of the simulation runs was relatively unique, and could develop or recover the digital organism populations at varying times from practically immediate to infinity.

The occasional appearance of zeros (Tables 1 and 2) shows that the press extinction regime was not always necessarily effective in suppressing the particular population. However, the symbol “NA” (meaning not available) indicated that in some scenarios recovery was not accomplished throughout the particular simulation run. We performed the Welch’s t-test on the means for each recovery distribution (results not shown) that identified geometry effect in some cases, but the geometry effect on the average populations in Figures 1 and 2 was more robust. To explore the stylized facts for various digital organisms, we computed the time-series of logarithmic returns on each population, i.e.,  $R_t = \log(P_t/P_{t-1})$ , where  $P_t$  was the population size at time  $t$ , and  $R$  represented the log return. The histograms of log returns showed values were stationary over time, unlike those of the population sizes. All distributions conformed to the power law, with exponents  $a$  limited in range from 2.3 to 3.0. The populations of more complex organisms also exhibited a broader range, as show in the bin width of the histogram in the log returns.

Table 1 presents a summary of the recovery times for each of the nine logic functions in the press extinction regime on toroidal grid. It shows the breadth of the artificial life scenarios. The evolution of the complex features, such as NAND, required a prior evolution of less complex functions, in order to commence. This evolutionary cascade nevertheless varied, and was different among each simulation.

Table 1: Recovery periods for logic functions in Avida (press extinction regime and toroidal geometry) The NA symbol means that until the end of simulation, i.e., after 200,000 updates, recovery had not taken place.

Run No.	Not	Nand	And	OrNot	Or	AndNot	Nor	Xor	Equals
1	25	1850	1650	1875	1025	3625	3750	1700	1750
2	1275	1375	2300	5050	3400	5675	4850	NA	NA
3	3100	14450	NA	95100	NA	7875	96825	2675	5750
4	NA	12525	6675	NA	NA	1775	2275	NA	1975
5	625	3650	2200	4250	21375	1375	2550	NA	NA
6	55875	2600	6375	26250	23200	22475	5900	22350	6950
7	14000	4350	0	3575	5325	6675	6425	NA	NA
8	1050	NA	NA	1575	1875	NA	1500	38075	51025
9	14200	1250	0	725	15775	375	70850	2325	NA
10	4650	8050	7525	4975	1350	4675	NA	NA	3400
11	NA	20400	16550	NA	NA	5025	3650	NA	NA
12	5175	4200	16750	1600	8550	85575	2675	1950	3875
13	9775	4950	4250	36375	12225	1775	3900	2050	3475
14	35125	2650	1325	1575	NA	48900	6875	4425	4400
15	48300	75	475	475	550	375	NA	NA	NA
16	10575	0	8075	41225	5125	8075	10250	NA	NA
17	875	10025	12100	37800	1375	6925	1475	NA	11925
18	3700	5050	3425	NA	19200	2225	4575	1925	4000
19	950	525	1600	1700	3800	23175	17625	750	2475
20	NA	0	0	1325	925	0	2700	NA	NA

NA: not available after 200,000 updates.

Source: Authors

Table 2: Recovery periods for logic functions in Avida (press extinction regime and clique geometry) The NA symbol means that until the end of simulation, i.e., after 200,000 updates, recovery had not taken place.

Run No.	Not	Nand	And	OrNot	Or	AndNot	Nor	Xor	Equals
1	675	7950	1125	575	500	550	67000	NA	350
2	NA	2100	0	225	2050	28600	0	0	0
3	500	6975	550	350	250	325	850	275	3475
4	1700	3225	5825	6400	1975	3450	1000	3175	3175
5	2025	0	650	71050	1425	225	5425	NA	NA
6	6100	1525	3225	1700	2650	10150	3475	1025	NA
7	25	20350	50	8725	100	65475	150	23950	18875
8	650	550	1675	NA	NA	1325	2350	NA	NA
9	NA	125	NA	1475	33075	3425	500	10450	3225
10	3400	3350	4225	27750	NA	2150	3775	NA	NA
11	95200	25	0	350	1200	1650	625	0	350
12	2150	32700	NA	1725	1400	NA	16475	750	1600
13	750	29275	NA	450	100	7250	350	7900	1450
14	31325	900	475	1075	775	50	1125	11025	1425
15	450	7575	5925	275	1150	NA	575	1075	1075
16	25	25	75	3175	75	3025	100	3400	19075
17	3750	150	80225	175	31325	NA	NA	1100	2475
18	50	25	50	900	25	50	91800	25	3325
19	425	1725	2775	NA	950	325	575	1525	1075
20	600	0	100	0	46925	0	50	350	1675

NA: not available after 200,000 updates

Source: Authors

Dips in the population curves in less complex functions were occasionally seen, often in conjunction with an emergence of a more complex trait. Table 2 presents an outline of the recovery times for each of the nine logic functions in the press extinction regime in the case of the clique approach. Again, the standard statistical techniques had limited use in this case because of large variances in the data. Thus, it was difficult to discern the topological effects by means of a statistical summary. The dynamics of average populations shown in Figures 1 and 2 offered a better picture on the effect of the world topology in Avida simulations.

## Conclusion

This work was devoted to studying the evolution of digital organisms in the artificial life research platform, Avida. It especially focused on re-evolution after an extinction period, which was created by selection pressure or a random pulse, and it followed the work of Yedid (2012). Using 100 runs for each geometry and every extinction regime, we found that an averaging process of individual-run recovery times had standard deviations that exceeded recovery times, with no upper boundary on the lengthiest recovery event. Instead of analyzing the statistical summaries of individual recovery periods, we discerned the difference in the topological effect from average recovery times. In the pulse extinction regime, the clique geometry produced faster recovery times for an average population than was found with the toroidal grid. For the press extinction regime, the recovery proceeded much slower than for the pulse extinction, and was not complete by the end of the simulation, i.e., after 200,000 updates. The analysis of the logarithmic returns of population sizes showed unexpected stylized features. It was found that the digital organisms of lower complexity (which were also in the lower position in the metabolic resource chain) produced a narrower width at half of the maximum in the histograms of log returns, than those of complex functions, such as the EQU (at the top of the metabolic hierarchy). In other words, the more complex organism, the higher the variance of the log returns. Furthermore, probability density functions of the log returns, in all cases, had fat tails, meaning they did not decay as fast as those displaying the normal distribution. They also exhibited a power law with inverse exponents in the range of 2.3–3.0. This stylized fact meant a relative abundance of extreme events (bubbles and crashes in population sizes) rather than the expectations related to the normal distribution. Finally, the value of the exponents was in accordance with that observed in living organisms, indicating that Avida is a realistic artificial life platform.

## References

- Adami, C., Ofria, C., & Collier, T. C. (2000) Evolution of biological complexity. *Proceedings of the National Academy of Sciences* 97, 4463-4468.
- Chow, S. S., Wilke, C. O., Ofria, C., Lenski, R. E., & Adami, C. (2004) Adaptive Radiation from Resource Competition in Digital Organisms, *Science* 305 (5680), 84-86.
- Clune, J., Misevic, D., Ofria, C., Lenski, R. E., Elena, S. F. & Sanjuan, R. (2008). Natural Selection Fails to Optimize Mutation Rates for Long-Term Adaptation on Rugged Fitness Landscapes, *PLoS Computational Biology* 4(9), e1000187.
- Clune, J., Goldsby, H. J., Ofria, C., & Pennock, R. T. (2011). Selective pressures for accurate altruism targeting: evidence from digital evolution for difficult-to-test aspects of inclusive fitness theory, *Proc Biol Sci.* 278 (1706), 666-674.
- Covert III, A. W., Carlson-Stevermer, J., Derrberry, D. Z., & Wilke, C. O. (2012). What does sex have to do with it: tracking the fate of deleterious mutations in sexual populations, *Artificial Life* 13,32-36.
- Goldsby, H. J. & Cheng, B. H. C. (2008). Avida-MDE: A Digital Evolution Approach to Generating Models of Adaptive Software Behavior, *GECCO'08: Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation 2008*, 1751-1758.
- Lenski, R. E., Ofria, C., Pennock, R. T., & Adami, C. (2003). The evolutionary origin of complex features. *Nature* 423 (6936), 139-144.
- Nelson, C. W. & Sanford, J. C. (2011). The effects of low-impact mutations in digital organisms. *Theoretical Biology and Medical Modelling* 8(9), 1-17.
- Yedid, G., Ofria, C. A., & Lenski, R. E. (2009). Selective press extinctions, but not random pulse extinctions, cause delayed ecological recovery in communities of digital organisms. *The American Naturalist* 173(4), E139-E154.
- Yedid, G., Stredwick, J., Ofria, C. A., & Agapow, P. M. (2012). A comparison of the effects of random and selective mass extinctions on erosion of evolutionary history in communities of digital organisms. *PloS one*, 7(5) e37233.