

# Improving and still passing the ALife test: Component-normalised activity statistics classify evolution in Geb as unbounded

Alastair Channon

Department of Electronic and Computer Engineering  
 Anglesea Road, University of Portsmouth, PO1 3DJ. UK  
 alastair@channon.net <http://www.channon.net/alastair><sup>1</sup>

## Abstract

Bedau, Snyder & Packard's (1998) classification system for long-term evolutionary dynamics provides a test for open-ended evolution. Making this ALife test more rigorous, and passing it, are two of the most important open problems in the field. Previously (Channon, 2001) I presented the result that 'Geb', a system designed to verify and extend theories behind the generation of evolutionary emergent systems (Channon & Damper, 2000), has passed this test. However I also criticised the test, most significantly with regard to its normalisation method for artificial systems. This paper details a modified normalisation method, based on component activity normalisation, that overcomes these criticisms. It then presents the results of the revised test when applied to Geb, which indicate that this system does indeed exhibit open-ended evolution.

## Introduction

As the discipline of Artificial Life has developed, so has the need for quantifiable measures of success. This is especially true with regard to evolvability and open-ended evolution. Bedau et al. (Bedau & Packard, 1991; Bedau *et al.*, 1997; Bedau, Snyder & Packard, 1998; Bedau & Brown, 1999) have developed not only elegantly simple statistical measures for long-term evolution, but also a test for unbounded evolution. The test is so adaptable that it can be applied to any evolving system with an available record of its components' existence times, such as the biosphere's fossil record. Any artificial system can be tested, and those that have been include Tierra-like systems (Adami & Brown, 1994; Taylor & Hallam, 1998), Echo (Holland, 1975), Bugs (Packard, 1989), Lindgren's (1991) model of evolving strategies in the iterated prisoner's dilemma, Arthur's (1994) "Bar Problem"

<sup>1</sup>Source code and instructions for replicating the runs discussed in this paper are available from the author's home page.

and Ecolab (Standish, 2000). Previously only taxonomic families in the fossil record have passed the test.

Bedau et al.'s test is based on the following statistics (from Bedau *et al.*, 1997; Bedau, Snyder & Packard, 1998) which are calculated from the record of components' existence times.

Activity increment (by presence).

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

This is not the only activity increment that they have used, but it is the best for comparison across systems because it can be calculated for any system with a record of components' existence times.

Evolutionary Activity of a component.

$$a_i(t) = \begin{cases} \sum_{\tau=0}^t \Delta_i(\tau) & \text{if component } i \text{ exists at } t \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Diversity (number of components present).

$$D(t) = \#\{i : a_i(t) > 0\} \quad (3)$$

Total cumulative evolutionary activity (or just 'total activity').

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

Mean cumulative evolutionary activity (or just 'mean activity').

$$\bar{A}_{\text{cum}}(t) = \frac{A_{\text{cum}}(t)}{D(t)} \quad (5)$$

New evolutionary activity per component (or just 'new activity').

$$A_{\text{new}}(t) = \frac{1}{D(t)} \sum_{i:a_i(t) \in [a_0, a_1]} a_i(t) \quad (6)$$

CLASS	EVOLUTIONARY DYNAMICS	STATISTICAL SIGNATURE		
		$D$	$A_{\text{new}}$	$\bar{A}_{\text{cum}}$
1	none	bounded	zero	zero
2	bounded	bounded	positive	bounded
3a	unbounded ( $D$ )	unbounded	positive	bounded
3b	unbounded ( $\bar{A}_{\text{cum}}$ )	bounded	positive	unbounded
3c	unbounded ( $D$ & $\bar{A}_{\text{cum}}$ )	unbounded	positive	unbounded

Table 1: Classes of evolutionary dynamics and their statistical signatures, based on table 1 from Bedau, Snyder & Packard (1998)<sup>2</sup>. Rows 3b and 3c have been added to class 3 (see text).

For  $A_{\text{new}}$  to be a good measure of new activity, the range  $[a_0, a_1]$  should be chosen such that component activities within it can be considered both adaptively significant and not amongst the highest. For artificial systems, a ‘shadow’ is run, mirroring the real run in every detail except that whenever selection (artificial or natural) operates in the real system, random selection is employed in the shadow. The statistics from this shadow can then be used to determine  $a_0$  and levels of total and mean activity that can be considered adaptively significant.

After determining long-term trends in these statistics, the system being examined can be classified according to table 1. The hallmark of class 3 (unbounded evolutionary dynamics) is unbounded total cumulative evolutionary activity in combination with positive new evolutionary activity per component. Other possibilities exist with zero  $A_{\text{new}}$ , but these belong in class 1 (no evolutionary activity) because such cases have no significant new components. Table 1 in Bedau, Snyder & Packard (1998) only shows the first row (3a) for class 3, but footnote 1 in that paper acknowledges the other rows (3b and 3c). Note that table 1 includes all possibilities for positive  $A_{\text{new}}$ , because zero  $\bar{A}_{\text{cum}}$  implies zero  $A_{\text{new}}$ . So any system with unbounded evolutionary dynamics will belong to class 3 (one of 3a, 3b and 3c).

Previously (Channon, 2001) I presented the result that the artificial system ‘Geb’ (Channon & Damper, 1998a,b, 2000) has passed this test. However I also criticised the test, most significantly with regard to its normalisation method for arti-

cial systems. The test relies on normalisation (or validation) from a shadow that can drift away from core aspects of the real run that it is intended to shadow. For example, the components that exist in the real population at any one time (well into evolution) are almost certainly more densely clustered than those in the shadow. So the mutation of a real component is more likely to produce another high-activity component than the mutation of a shadow component. Once the real and shadow populations have been allowed to evolve, we are no longer comparing the real run with a true shadow. The longer the period since the shadow was initialised to match the real run, the less relevant the shadow is to the real run.

My other criticism of the test was in its use of mean activity when looking for unbounded activity growth, especially when classifying a system as belonging to class 3b. When diversity is bounded, retention (forever) of a single component results in unbounded mean activity. The test should not be so influenced by such components, and should rather look for trends in typical components. So it is median activity, not mean activity, that should be measured, and required to be unbounded for a system to be classified as within class 3b.

Before progressing to the two main contributions of this paper (the new normalisation method and its results when applied to Geb) it is first necessary to outline how the statistics have been implemented in Geb. These details apply equally to both the original and modified tests, which use the same statistics for the real run and only differ in their shadow and its use in normalising the real run’s statistics.

<sup>2</sup>Note that Bedau has since altered his class numbering scheme.

### Implementing the statistics in Geb

Geb is a virtual world containing autonomous organisms, each controlled by a neural network. Each neuron has a bit-string label, or ‘character’, which is used during development and for matching the neural outputs of one organism with basic behaviours (turning, killing, etc.) and with inputs of other organisms. An organism is born with a simple axiom network that results in reproduction. This develops through the application of a genetically determined Lindenmayer system (L-system) (Lindenmayer, 1968). Each L-system production rule has the following form:

$$\mathcal{P} \rightarrow \mathcal{S}_r, \mathcal{S}_n ; b_1, b_2, b_3, b_4, b_5, b_6$$

$\mathcal{P}$  Predecessor (first bits of node’s character)  
 $\mathcal{S}_r$  Successor 1: *replacement* node’s character  
 $\mathcal{S}_n$  Successor 2: *new* node’s character  
bits:  $b_1, b_2, b_3, b_4, b_5, b_6$  specify linkage details

The *successors* (1 and 2) are characters for the node(s) that replace the old node. If a successor has zero length then that node is not created.

An evolved genotype contains a large number of production rules (once decoded), but only the rules found to match neuron’s characters most closely are used during development. In this way, increasingly specific production rules can evolve, with regressive rules existing as fall-back options should a rule be damaged by crossover or mutation, and as material for further evolutionary exploration.

When a new organism is ‘born’, all possible production rules are decoded from its genotype. Then the developmental process is part-simulated in advance of it truly taking place, as a means of filtering out all the production rules that would never be used, either because they would never match any possible node’s character, or because more specific rules exist for each node that could develop. All rules that remain will be used if the organism lives long enough.

This makes the choice ‘production rules that survive the filtering process at birth’ a good choice for the component class. It turns out that there is a high degree of neutrality in successors, making them difficult to track. So a better choice for the component class, and the one I have used, is ‘predecessor

plus link details’ from production rules that survive the filtering process. See Channon (2001) for details. Having chosen this component class, there is a clear consequence for the possible classifications of evolutionary dynamics. Because the number of neurons that an organism can have is limited (for practical reasons), the number of production rules that can survive filtering is limited. And because the population size is small (a maximum of four hundred organisms), there is little room for more than a couple of species at a time. So diversity of these components will certainly be bounded, and we can rule out class 3a and 3c dynamics.

Geb’s shadow mirrors the real run in every detail except that selection is random. Whenever a real organism is killed, a randomly chosen shadow organism is also killed. Whenever a real organism is born (as the product of two real organisms), a new shadow organism is born as the product of two randomly chosen shadow organisms, using the same reproduction procedure with the same rates of crossover and mutation. For each initial real organism born with single-bit genotype ‘0’, an initial shadow organism is also born with single-bit genotype ‘0’.

Component snapshots were taken every one thousand timesteps. To put this in context, the example run reported lasted six million timesteps, during which time there were over five hundred and eighty million organism reproductions. In each timestep, every organism is updated. Because activity is intended as a measure of how much a component both is used (already covered above) and *persists*, I screen out (in each of the real and shadow populations) isolated occurrences: when a component occurs in the current snapshot but not the previous one.

In previously published work on Geb (before the implementation of evolutionary statistics), total extinction (population size dropping to one individual) was not mentioned because it had not been encountered. However, there was no mechanism in place to prevent it and, during the long trial runs undertaken when experimenting with evolutionary statistics, I encountered occasional runs in which total extinction occurred. So for the set of runs

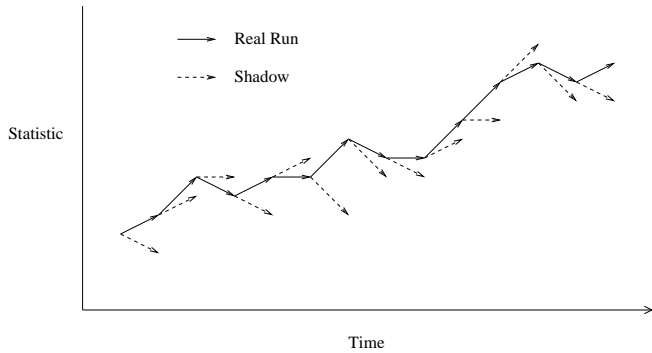


Figure 1: Illustration of the shadow-resetting method.

from which the example reported here is taken, I set a minimum number of organisms to twenty. The fact that total extinction is so rare despite the population size being so small (a maximum of four hundred organisms at any one time) indicates that there is no serious problem here. Once population sizes can feasibly be increased, the problem should in practice disappear rapidly.

### The shadow-resetting method

Because there is reason to doubt a method of normalising (or validating) evolutionary statistics that relies on a shadow that can drift away from core aspects of its real run, a new method needed to be developed that regularly resets the shadow (both components and activity history) to be identical to the real run. The rest of this paper details the development of such a method and reports on its results when applied to Geb.

The basic idea is that immediately after each snapshot (when an entry is made in the component existence record), the shadow run has its components reset to those of the real run. This allows us to compare inter-snapshot changes in activity in the real run with the changes we would expect from random selection, the result being an improved generic shadowing mechanism. When calculating evolutionary statistics (and indeed when recording component numbers), the shadow's history is taken to be that of the real run - see figure 1.

Most of the results below are from a typical run, drawn from a set of twenty carried out using this procedure. Atypical variations found within this set

are also reported and discussed. Figure 2 shows the raw real and shadow activity wave diagrams from the typical run. One obvious feature is that many of the waves keep increasing. As discussed in Channon (2001), this would also be true in a similar analysis of the biosphere's evolution. Genes that are beneficial to life tend to become basic for many species: humans have a significant proportion of genes in common with mice, flies and even plants.

Shadow waves follow the real waves, because the shadow is reset after each snapshot. The shadow loses components between snapshots far more frequently than the real run does. This is especially true of the lower-activity components, as we should expect. Adaptively significant production rules have many redundant copies on a typical genome such that should mutation break the rule at one point, it will still be decoded from elsewhere on the genome. Such components can survive even sustained periods of random selection. Yet even the highest activity components are frequently lost in the shadow and this provides verification that the snapshot interval (one thousand timesteps) is sufficient for comparing activity by presence. Employing a much higher reset frequency would require the use of an activity increment function which is sensitive to the number of occurrences of a component at any one time<sup>3</sup>, for example  $\Delta_i(t) = \text{fraction of organisms that have the } i^{\text{th}} \text{ component at time } t$ . However, because significantly more frequent snapshots are not feasible with the computational resources currently available, comparing activity by presence is still the best option.

Of course it would not make sense to calculate activity statistics based solely on the shadow's component existence record. Perhaps the most obvious course of action (and the one I took at first) is to calculate activity statistics for the shadow on the basis that at each timestep  $t$  we use the real run's component record for timesteps before  $t$ , and the shadow's component record at  $t$ . Figure 3 shows the resulting statistics from the shadow, alongside the real run's statistics. Activity (total, mean and median) is unbounded in the shadow only because the shadow has its components reset to those of the

<sup>3</sup>Thanks to Mark Bedau for bringing this to my attention.

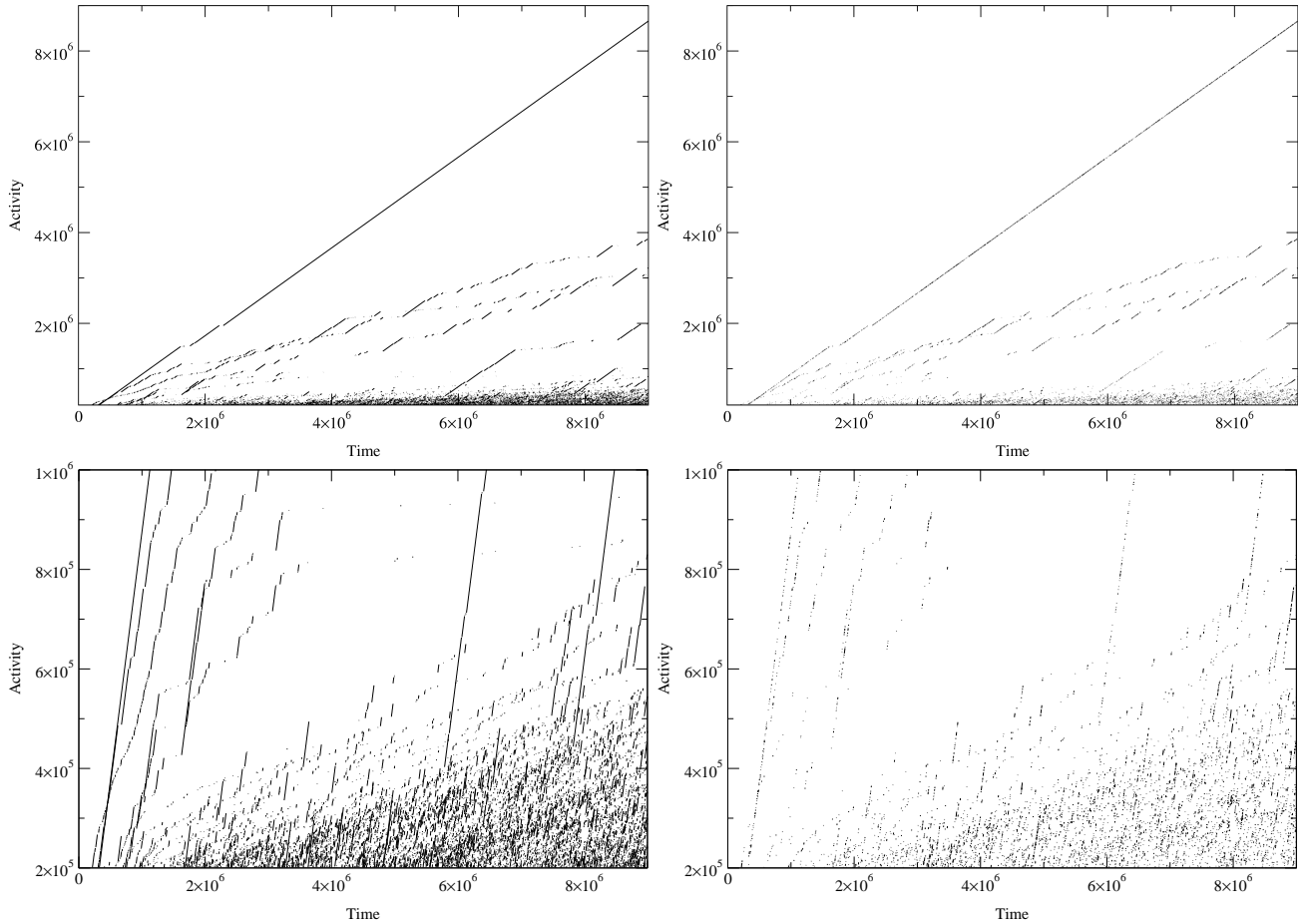


Figure 2: Activity wave diagrams for the real (left) and shadow (right) runs, with all horizontal (no-increase) lines removed. The lower diagrams show a magnified view of the activity range below 1 million.

real run after each snapshot. On average both total activity and diversity drop sharply in the shadow over each short (1000 timesteps) interval after it has been reset to the real run's state. The shadow's mean and median activity statistics show that (on average) it is the higher activity components that remain in the shadow, in agreement with the discussion (above) of the activity wave diagrams. Do not be confused by the fact that mean and median activity increase in the shadow over each inter-snapshot interval. This is due to the loss of lower activity components, not the result of any increase in component activity.

These results are encouraging, but they provide no route to normalising the real run's statistics in order to demonstrate a presence or lack of unbounded growth in, say, median activity. The method so far also provides no sound way of mea-

suring new activity. So naively calculating the shadow's activity statistics leads us to a dead-end. The idea of resetting the shadow run's state to match the real run's state just after each snapshot is a good one, but how can it be used to normalise the real run's statistics?

### Component activity normalisation

The solution is to normalise at the lower level of individual components' activities, rather than at the component-population level. This is done by subtracting the shadow's component activity increment from the real run's component activity increment, for each component. So when calculating activity by presence, a component's normalised activity is incremented if and only if it persists (and is used) in the real run but not in the shadow, and is decremented if and only if it persists (and is used)

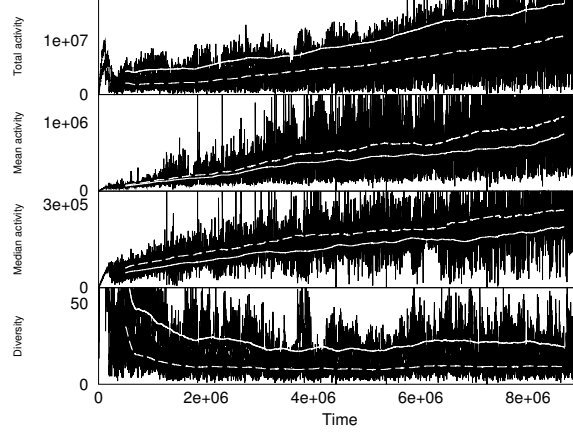


Figure 3: Total activity, mean activity, new activity and diversity from a typical Geb run and its regularly-reset shadow. Running averages are shown in white: solid for the real run, dashed for its shadow.

in the shadow but not in the real run (although activity will still be read as zero for that timestep). Here are the revised statistics:

Real run's component activity increment.

$$\Delta_i^R(t) = \begin{cases} 1 & \text{if component } i \text{ exists} \\ & \text{in the real run at } t \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

Shadow's component activity increment.

$$\Delta_i^S(t) = \begin{cases} 1 & \text{if component } i \text{ exists} \\ & \text{in the shadow at } t \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

Normalised component activity increment.

$$\Delta_i^N(t) = \Delta_i^R(t) - \Delta_i^S(t) \quad (9)$$

Normalised component activity.

$$a_i^N(t) = \begin{cases} \sum_{\tau=0}^t \Delta_i^N(\tau) & \text{if component } i \text{ exists} \\ & \text{in the real run at } t \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

Normalised diversity.

$$D^N(t) = \#\{i : a_i^N(t) > a_0^N\} \quad (11)$$

Note that this formula for  $D^N$  is only a suggestion for how diversity could be normalised when investigating systems with unbounded diversity: by

counting the number of components whose normalised activity has passed the threshold at which we consider them to be adaptively significant (see below). This method of normalising diversity is debatable. However, because Geb does not exhibit unbounded diversity I safely ignore that debate here, and do not calculate  $D^N$  for Geb. This is valid because no claim of unbounded diversity is being made, and because  $D^R$  (not  $D^N$ ) is the relevant value to use when calculating  $\bar{A}_{\text{cum}}^N$ ,  $\tilde{A}_{\text{cum}}^N$  and  $A_{\text{new}}^N$ , because  $D^R$  is the number of components that contribute to those statistics. Normalised total cumulative evolutionary activity.

$$A_{\text{cum}}^N(t) = \sum_{\substack{i: \text{component } i \text{ exists} \\ \text{in the real run at } t}} a_i^N(t) \quad (12)$$

Normalised mean cumulative evolutionary activity.

$$\bar{A}_{\text{cum}}^N(t) = \frac{A_{\text{cum}}^N(t)}{D^R(t)} \quad (13)$$

Normalised median cumulative evolutionary activity.

$$\tilde{A}_{\text{cum}}^N(t) = \text{Median}_{\substack{i: \text{component } i \text{ exists} \\ \text{in the real run at } t}} a_i^N(t) \quad (14)$$

Normalised new activity per component.

$$A_{\text{new}}^N(t) = \frac{1}{D^R(t)} \sum_{i: \text{component } i \text{ 'new'}} a_i^N(t) \quad (15)$$

See below for the details of calculating normalised new activity per component.

This is clearly the better approach, for it produces normalised component activities that measure how much each component's activity has increased above the increase that would have occurred had selection been random. So a component's normalised activity is a direct measure of the degree to which adaptive selection in the real run is causing the component to persist (and be used).

### Determining the normalised new-activity criteria

The final requirement, before these statistics can be used to classify evolutionary dynamics, is a method of determining when a component is newly adaptively significant: Clearly the method from Bedau, Snyder & Packard (1998) cannot be used with the revised shadowing mechanism. The method must provide a (normalised) activity level  $a_0^N$  at which a component can be considered adaptively significant, and a procedure for dropping a component from the list of new components. For the second of these concerns, a simple upper bound cannot be used, because normalised activity can both increase and decrease, so a component could potentially be considered 'new' forever. The simplest (and adequate) solution is to consider a component to be 'new' (newly adaptively significant) in the snapshot at which its activity reaches  $a_0^N$ , and never after that. So each component is considered new at most once. This leaves the issue of determining  $a_0^N$ .

If the presence or absence of a component confers no adaptive advantage or disadvantage, then the real and shadow systems are equivalent for this component. Further, which is used as the reset-to system (after each snapshot) makes no difference to the component's activity. So the (normalised) activity distribution for this class of components will be symmetric about the origin. Therefore, provided we can make the assumption that the most negative normalised activity encountered during a run is from such a component, we can negate this value to find a level at which normalised activity can be considered adaptively significant. Even if this assumption does not hold, negating the most negative activity encountered provides a value above which activity can be considered adaptively significant, even if this bound is higher than necessary.

This method can be expected to work well when activity is calculated by presence (as it is in Bedau and Packard's test and so also here), where changes in component activity ( $\Delta_i$ ) are small when compared with the activities of non-adaptive components. However, we should not expect it to provide a good bound when calculating activity by, for example, usage ( $\Delta_i = \# \text{components } i \text{ at } t$ ), where the most negative activities arise from neutral mutations of high usage components, some of which a shadow will encounter before its real run does.

### Results and Discussion

Figure 4 shows the normalised activity waves from the typical run, already discussed above, drawn from the full set of twenty runs. Notice that the activity values are significantly lower than before normalisation - see figure 2.

In each of the twenty runs, the lowest normalised activity encountered was greater than -30, with -10 being a more typical value. For simplicity I used a new-activity threshold ( $a_0^N$ ) of 30 on all runs. Figure 5 shows the resulting statistics for the typical run. Normalised median activity is unbounded (as are normalised total activity and normalised mean activity) and normalised new activity is positive. These results clearly fall into class 3b (according to this classification system): unbounded evolutionary activity.

These results are typical of the twenty runs carried out for this set of experiments. However, five of the runs encountered problems, causing their results to be atypical. Three of these effectively met total extinction. In the implementation section I noted that I imposed a minimum limit on the number of organisms, in an attempt to avoid total extinction. However, if population size hits this limit and does not increase rapidly, then many reproductions may occur with selection effectively random. This causes evolutionary activity to plummet as adaptive traits are lost. Once lost, this activity cannot be regained, except by the evolution of new adaptive components. These results are not a cause for concern, for the same reasons mentioned above: once population sizes can feasibly be increased, the problem should in practice disappear rapidly.

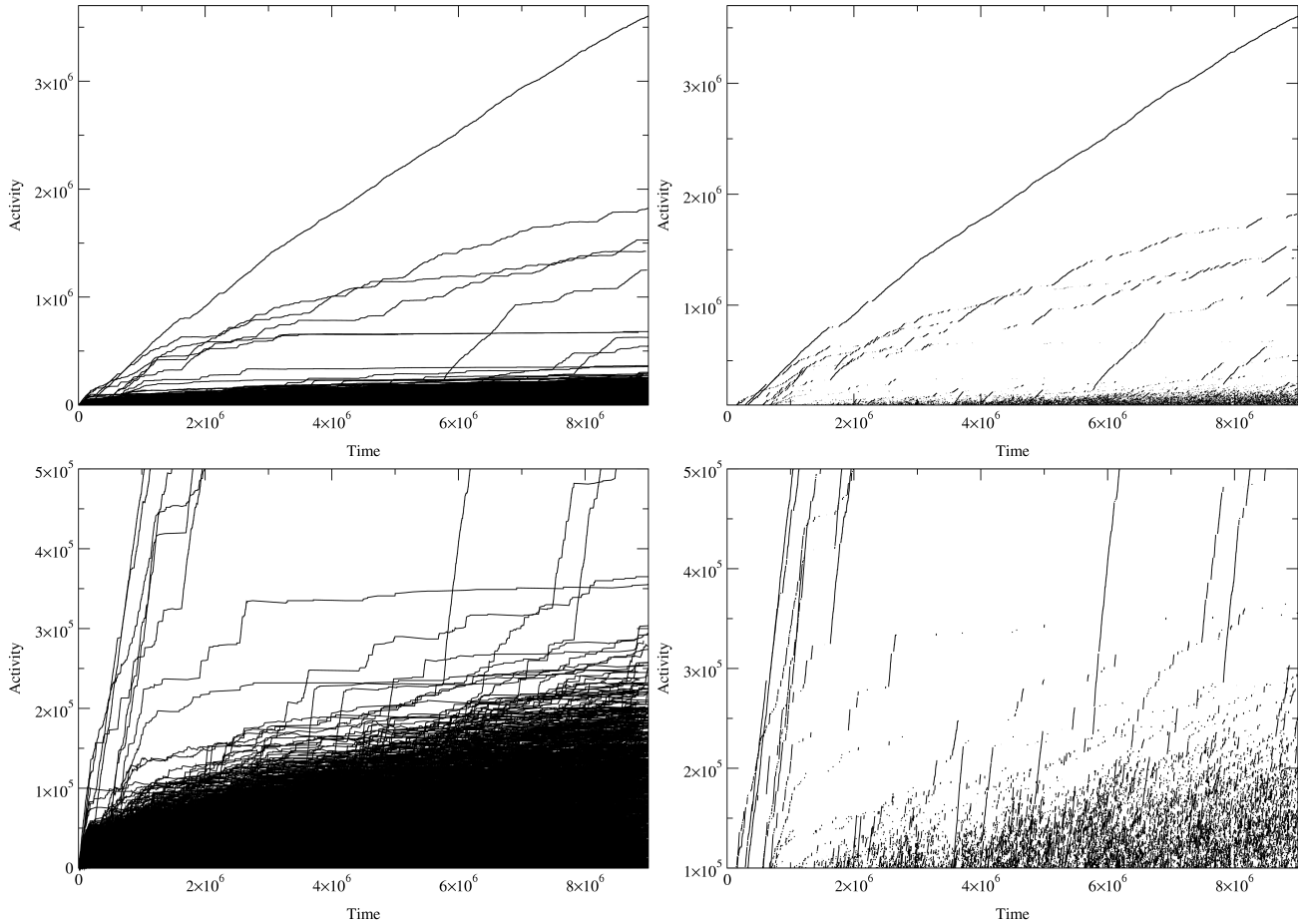


Figure 4: Normalised activity wave diagrams. Those on the right have had all horizontal (no-increase) lines removed. The lower diagrams show a magnified view of the activity range below 0.5 million.

Each of the other two atypical runs stagnated when the only existing species stopped reproducing. Of course this would ordinarily be a very poor strategy. It is easy to imagine how the bad gene (production rule) could have spread through a population of just one species as fit individuals reproduced with the new unfit ones, causing their children to pick up the dominant bad gene. However, one would not expect this to pose a threat to a different species. This is easily verified: introducing just a few organisms from any of the other evolved populations (from the other runs), causes the old organisms to be rapidly displaced by the newcomers. So this result is also not a cause for concern, for the same reason: it is due to the small population size, which cannot support more than one or two species at a time. Note also that both types of atypical run were also seen in Channon (2001).

It is easy to demonstrate that unbounded growth (or more accurately unbounded non-monotonic but *directed* increase) in activity, with positive new activity, is not a trivial consequence of unbounded genotype length. Consider the analogous system  $\text{Geb}^R$ , in which selection is random but all other details are as in Geb. Whenever a (randomly chosen) real organism is killed in  $\text{Geb}^R$ , a randomly chosen organism is also killed in its shadow. Whenever a real organism is born in  $\text{Geb}^R$  (as the product of two randomly chosen real organisms), a new shadow organism is born as the product of two randomly chosen shadow organisms. Of course running either real or shadow system from the same snapshot more than once would produce different results on each run, because of the stochastic nature of the systems. So normalised activity would unfold as a random walk, with ‘step’ probability



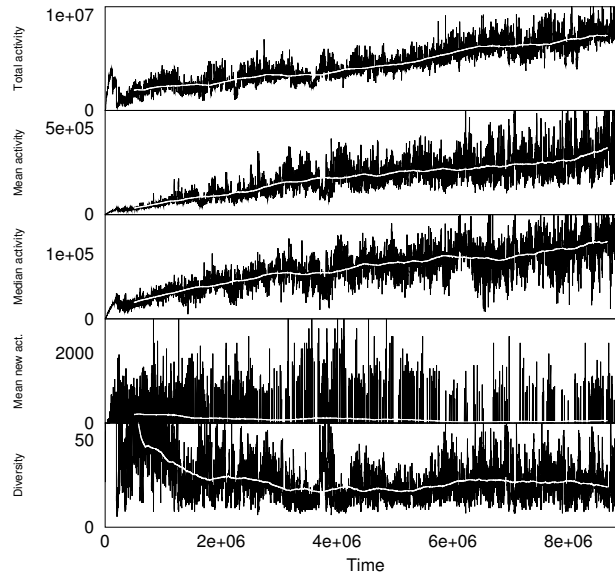


Figure 5: Normalised total activity, normalised mean activity, normalised median activity, normalised new activity, and real diversity. Running averages are shown in white.

distribution changing at each snapshot but always symmetric about zero.

Could unbounded growth in activity, with positive new activity, be a trivial consequence of unbounded genotype length in a biotic selection system? Certainly not, for new activity drops to zero in systems such as Tierra. Could unbounded growth in activity be a trivial consequence of unbounded genotype length in a biotic selection system that exhibits unending positive new activity? No, because the requirement remains that activity be retained, so that it can accumulate. For example, a (diversity-bounded) biotic selection system that continually generates new components only by mutation along (phenotypically) neutral networks would only be able to *use* a finite number of neutral variants at any one time. It would lose activity whenever a component is lost from (ceases being used in) the system.

### Conclusions

Both of my criticisms of the original test have been addressed. The revised shadowing method used here ensures that the normalisation of statistics is through a shadow that remains true to its real run, and median rather than (or rather as well as) mean activity has been used in the classification.

Geb has demonstrated class 3 behaviour: unbounded evolutionary activity. And this time we can have a greater degree of confidence in the results. However, this is a new variant of a previous test, and it is not beyond possibility that it could be improved upon. Certainty in these results can only come about through the application of the test to a range of evolutionary systems. That may take some time, since there are no other known artificial systems that even pass the original test. So for now we must be content with the conclusion that there is reason to believe that this system exhibits unbounded evolution.

While the caution of the previous paragraph is warranted, it is at least possible to say with certainty that these results qualitatively surpass those from previous artificial evolutionary systems. No previous biotic selection artificial evolutionary system has demonstrated unbounded evolutionary activity with positive new activity. As such these results provide validation of the theory behind Geb's design: a design constructed to satisfy the set of requirements for an evolutionary system within which increasingly complex advantageous behaviours can emerge, as uncovered by evaluating previous artificial systems.

## Acknowledgements

Thanks to Mark Bedau for discussion and many helpful comments on this work. Thanks also to the anonymous reviewers for constructive criticism.

## References

- Adami, C., and Brown, C. T. 1994. Evolutionary learning in the 2D artificial life system 'Avida'. In Brooks, R. A., and Maes, P., eds., *Proceedings of Artificial Life IV*, 377–381. Cambridge, MA: MIT Press.
- Arthur, W. B. 1994. Inductive reasoning and bounded rationality. *American Economic Review (Papers and Proceedings)* 84:406–411.
- Bedau, M. A., and Brown, C. T. 1999. Visualizing evolutionary activity of genotypes. *Artificial Life* 5(1):17–35.
- Bedau, M. A., and Packard, N. H. 1991. Measurement of evolutionary activity, teleology, and life. In Langton, C. G.; Taylor, C.; Farmer, J. D.; and Rasmussen, S., eds., *Proceedings of Artificial Life II*, 431–461. Redwood City, CA: Addison-Wesley.
- Bedau, M. A.; Snyder, E.; Brown, C. T.; and Packard, N. H. 1997. A comparison of evolutionary activity in artificial evolving systems and the biosphere. In Husbands, P., and Harvey, I., eds., *Proceedings of the Fourth European Conference on Artificial Life (ECAL97)*, Brighton, 125–134. Cambridge, MA: MIT Press.
- Bedau, M. A.; Snyder, E.; and Packard, N. H. 1998. A classification of long-term evolutionary dynamics. In Adami, C.; Belew, R.; Kitano, H.; and Taylor, C., eds., *Proceedings of Artificial Life VI, Los Angeles*, 228–237. Cambridge, MA: MIT Press.
- Channon, A. D., and Damper, R. I. 1998a. Evolving novel behaviors via natural selection. In Adami, C.; Belew, R.; Kitano, H.; and Taylor, C., eds., *Proceedings of Artificial Life VI, Los Angeles*, 384–388. Cambridge, MA: MIT Press.
- Channon, A. D., and Damper, R. I. 1998b. Perpetuating evolutionary emergence. In Pfeifer, R.; Blumberg, B.; Meyer, J.-A.; and Wilson, S., eds., *From Animals to Animats 5: Proceedings of the Fifth International Conference on Simulation of Adaptive Behavior (SAB98)*, Zurich, 534–539. Cambridge, MA: MIT Press.
- Channon, A. D., and Damper, R. I. 2000. Towards the evolutionary emergence of increasingly complex advantageous behaviours. *International Journal of Systems Science* 31(7):843–860.
- Channon, A. D. 2001. Passing the ALife test: Activity statistics classify evolution in Geb as unbounded. In Kelemen, J., and Sosik, P., eds., *Advances in Artificial Life: Proceedings of the Sixth European Conference on Artificial Life (ECAL2001)*, Prague, 417–426. Heidelberg: Springer-Verlag.
- Holland, J. H. 1975. *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI: University of Michigan Press. 2nd edition: MIT Press, 1992.
- Lindenmayer, A. 1968. Mathematical models for cellular interaction in development. *Journal of Theoretical Biology* 18:280–315. Parts I and II.
- Lindgren, K. 1991. Evolutionary phenomena in simple dynamics. In Langton, C. G.; Taylor, C.; Farmer, J. D.; and Rasmussen, S., eds., *Proceedings of Artificial Life II*, 295–312. Redwood City, CA: Addison-Wesley.
- Packard, N. H. 1989. Intrinsic adaptation in a simple model for evolution. In Langton, C. G., ed., *Santa Fe Institute Studies in the Sciences of Complexity, Vol VI: Proceedings of the Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems (Artificial Life I)*, 141–155. Redwood City, CA: Addison-Wesley.
- Standish, R. K. 2000. An ecolab perspective on the Bedau evolutionary statistics. In Bedau, M. A.; McCaskill, J. S.; Packard, N. H.; and Rasmussen, S., eds., *Proceedings of Artificial Life VII*, 238–242. Cambridge, MA: MIT Press.
- Taylor, T., and Hallam, J. 1998. Replaying the tape: An investigation into the role of contingency in evolution. In Adami, C.; Belew, R.; Kitano, H.; and Taylor, C., eds., *Proceedings of Artificial Life VI, Los Angeles*, 256–265. Cambridge, MA: MIT Press.