

An Ecolab Perspective on the Bedau Evolutionary Statistics

Russell K. Standish

High Performance Computing Support Unit
University of New South Wales
Sydney, 2052
Australia
R.Standish@unsw.edu.au
<http://parallel.hpc.unsw.edu.au/rks>

Abstract

At Alife VI, Mark Bedau proposed some evolutionary statistics as a means of classifying different evolutionary systems. Ecolab, whilst not an artificial life system, is a model of an evolving ecology that has advantages of mathematical tractability and computational simplicity. The Bedau statistics are well defined for Ecolab, and this paper reports statistics measured for typical Ecolab runs, as a function of mutation rate. The behaviour ranges from class 1 (when mutation is switched off), through class 3 at intermediate mutation rates (corresponding to scale free dynamics) to class 2 at high mutation rates. The class 3/class 2 transition corresponds to an error threshold. Class 4 behaviour, which is typified by the Biosphere, is characterised by unbounded growth in diversity. It turns out that Ecolab is governed by an inverse relationship between diversity and connectivity, which also seems likely of the Biosphere. In Ecolab, the mutation operator is conservative with respect to connectivity, which explains the boundedness of diversity. The only way to get class 4 behaviour in Ecolab is to develop an evolutionary dynamics that reduces connectivity of time.

Introduction

At Alife VI, Mark Bedau proposed some evolutionary statistics (Bedau *et al.*, 1998) as a means of classifying different evolutionary systems. The intent here is to find a general scheme analogous to Wolfram's (1984) classification scheme of cellular automata. Three statistics are proposed:

Diversity (D): The number of species or components in the system

Mean Cumulative Evolutionary Activity (\bar{A}_{cum}): Activity of a species is defined as the population count of that species, the vector n in Ecolab terms. Evolutionary activity subtracts from this the neutral or nonadaptive part. This is achieved by running a *neutral shadow model*, that is identical with the original model, except that natural selection must be "turned off". Finally, this activity is accumulated over time of the species, and then averaged over all species.

New Evolutionary Activity (A_{new}): This corresponds to the number of new species crossing a threshold, divided by the diversity.

Bedau describes four classes of evolutionary behaviour, as in the following table:

| Class | $D(t)$ | $A_{cum}(t)$ | $A_{new}(t)$ | Description |
|-------|-----------|--------------|--------------|-----------------------|
| 1 | bounded | zero | zero | none |
| 2 | bounded | unbounded | none | unbounded, uncreative |
| 3 | bounded | bounded | positive | bounded, creative |
| 4 | unbounded | positive | positive | unbounded, creative |

Note that in Bedau *et al.* (1998), only 3 classes are mentioned — class 2 was added later in his presentation at Alife VI. Bedau has applied his statistics to a number of artificial life models, including Echo (Holland, 1995) and Tierra (Ray, 1991), none of which exhibit class 4 behaviour. By contrast, the same statistics applied to the fossil record (at least for the Phanerozoic — the period of time since the appearance of multicellular life in the Cambrian) — show a strong class 4 behaviour. Further, Bedau speculates that the global economy and internet traffic are also class 4, particularly as they show strong growth over a significant period of time. Since no artificial life systems to date appear to show class 4 behaviour, the gauntlet has been laid down to discover such a system to work out whether this classification difference is fundamental or not.

Ecolab (Standish, 1994), whilst not an artificial life system, is a model of an evolving ecology that has advantages of mathematical tractability and computational simplicity. It lies in between the extremely simplistic models of (for example) Bak and Sneppen (1993) or Newman (1997) and artificial life models of evolution such as *Tierra* or *Avida*. One of its key characteristics is that its dynamics are defined by the ecological interactions between the species, rather than ad hoc exogenous dynamics. The Bedau statistics are well defined for it, so it is interesting to see what class behaviour Ecolab exhibits. Furthermore, an Ecolab-like model is possible for all artificial life systems (valid in a continuum limit). For

Copyrighted Material

example, the equations of motion for Tierra are given in Standish(Standish, 1997).

Ecolab

The Ecolab model (as opposed to the Ecolab simulation system) is based on an evolving Lotka-Volterra ecology. The defining equation is given by:

$$\dot{n} = r * n + n * \beta n + \gamma * \nabla^2 n + \mu(r * n) \quad (1)$$

where n is the species density, r the effective reproduction rate (difference between the intrinsic birth and death rates in the absence of competition), β the matrix of interaction terms between species, γ the migration rate and μ the mutation operator. All of these quantities (apart from β , which is a matrix) are vectors of length n_{sp} , the number of species in the ecology. The operator $*$ denotes elementwise multiplication. The mutation operator returns a vector of dimensionality greater than n_{sp} , with the first n_{sp} elements set to zero — in effect expanding the dimensionality of the space, a key feature of this system. For a more detailed exposition of the various properties of the model, in particular, the precise form of the mutation operator, the reader is referred to the previous published papers, as well as the Ecolab Technical Report, which are all available from the Ecolab Web Site¹.

For the purposes of this paper, it is worthwhile expounding a little on the properties of the mutation operator. It models point mutations in particular (other mutation types, such as recombination are simply not modeled within Ecolab). Point mutations in genotype space, which satisfy Poisson statistics, give rise random mutations, with locality, in phenotype space. Since the only phenotypic properties of interest to the model are the parameters r , β and γ the parameters are mutated according to a normal or lognormal distribution (according as the parameters are reals or positive (or negative) respectively), using a sample from the Poisson distribution for the width. The two parameters governing mutation (width of the Poisson distribution, and the rate at which mutations are attempted) are related via a simple proportional factor (called the “species radius (or separation)”) that is kept constant throughout the simulations reported here. Each species has its own mutation rate — given as a vector μ .

Each of these phenotypic parameters are initialised from a uniform distribution. The relevant input parameters for a run are then maximum and minimum values for each of r , the diagonal of β , the offdiagonal of β , μ , γ and the species radius ρ . The complete system may be scaled in the time dimension, fixed by what value is chosen for the timestep. In this case, $\max_i r_i = 0.1$, so one timestep corresponds to about a 14th of the doubling time of the fastest reproducing

ecology. This is a compromise between continuity of the simulation and computational expense. The ratio $\frac{\max_i r_i}{\max_i \beta_{ii}}$ roughly corresponds to the carrying capacity of the ecology. This is chosen to be about 100 so that behaviour near the equilibrium is reasonably continuous rather than stochastic. The ratio of offdiagonal to diagonal terms relates to how negative definite β is. Since mutations tends to drive the matrix away from being negative definite (system stability), the maximum of the offdiagonal terms is chosen to make the initial system marginally unstable. The species radius $\rho = 0.1$ was chosen empirically to make new species *phenotypically* distinct from its parent species.

Having fixed the other parameters according to the above criteria, the remaining degrees of freedom are μ and γ . In this paper, we vary the maximum mutation rate in different simulations, but keep the distribution of migration rates fixed.

One other feature worth noting is that the mutation operator will also randomly add or drop connections between species, according to an exponential distribution. Thus, the mutation operator is in fact highly conservative — with the lognormally mutated parameters capped (in the case of μ and γ) or restricted by the requirements of boundedness (diagonal components of β)(Standish, 1998; Ecolab Technical Report).

Neutral Shadow Model

An important feature for improving the accuracy of the evolutionary statistics is the use of a *neutral shadow model*. This model should be as similar as possible to the original model, but with all selection turned off. In the case of Ecolab, this is accomplished by running a shadow population density vector n' , and when n is updated, the shadow vector is updated by a random permutation of the updates. Thus each shadow species behaves in the long run like an average species. Activity is also tracked at the same time, with the activity vector being updated by the difference between the population density and the shadow population density, provided that difference is positive.

The new activity statistic A_{new} is computed by summing the number of species that have crossed a threshold. In (Bedau *et al.*, 1998), this threshold is determined by plotting the activity distributions for both the original and the shadow model, and taking the cross-over point as the threshold. This turned out to be 50 individuals, rather than the arbitrary 10 individuals used in other Ecolab studies. In fact the two distributions are nearly equal over the range 10–50, but if an activity is above 50, then it is highly likely to be due to adaptive behaviour.

¹<http://parallel.hpc.unsw.edu.au/rks/ecolab.html>

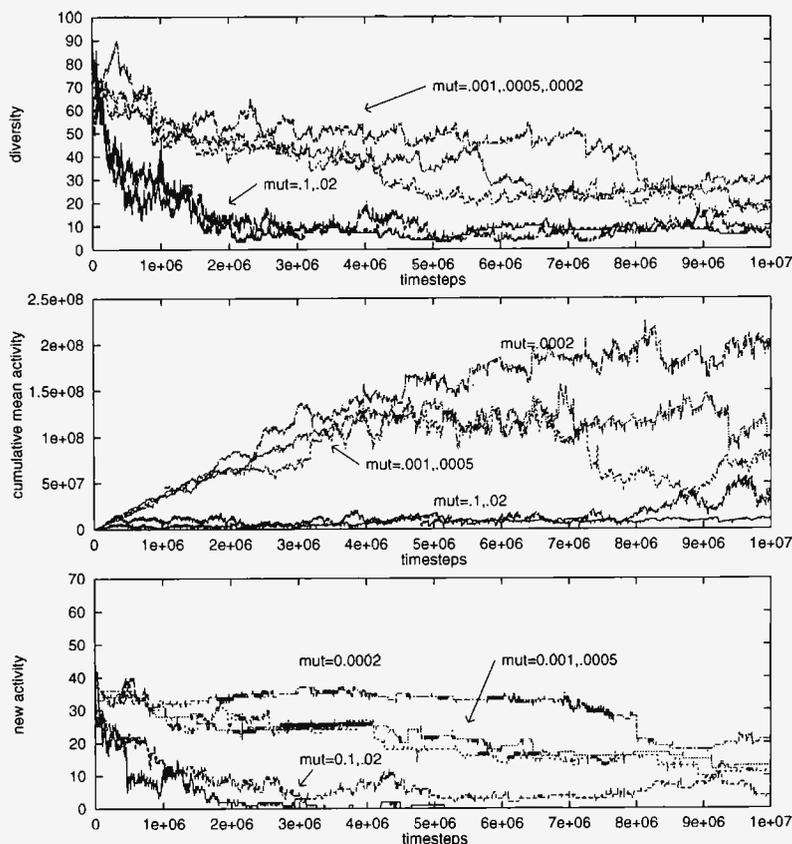


Figure 1: A typical run for panmictic Ecolab at varying mutation rates, showing the Bedau statistics: diversity, cumulative mean activity and new activity

Behaviour of Ecolab

Figure 1 shows the Bedau statistics for typical Ecolab runs (panmictic, or spatially independent case), as a function of mutation rate. When the mutation rate is too low, class two behaviour is seen. Diversity remains constant, and activity grows unbounded as the system rapidly sheds unviable organisms and tends to a stable ecology. Conversely, for high levels of mutation, class one behaviour is seen. There is a constant churn of organisms, that do not have any chance to generate activity. For intermediate levels of mutation, an interesting situation arises. Here, the number of mutant organisms that successfully invade the ecosystem roughly balances the number lost through extinction (Standish, 1998). Scale free behaviour is observed in a number of statistics, such as the distribution of species lifetime. These same 3 states of behaviour have been observed in Avida (Adami *et al.*, 1998).

The code used for this simulation is available from the Ecolab web site as version 3.3 of the software. The model including the neutral shadow model is defined in `shadow.cc`, and a sample experimental script

given as `bedau.tcl`. The only parameters varied are the spatial dimensions and `mutation(random,maxval)`.

The evolutionary statistics were also collected for a spatially dependent Ecolab, however due to some implementation difficulties, run lengths exceeding 1×10^6 timesteps have not been achieved prior to this paper's deadline. Broadly speaking, though, the same behaviour is seen as the panmictic case, although there is a period of diversity growth in the early period prior to settling on a higher level of diversity than the panmictic case.

This can be understood by considering two extremes of spatially dependent Ecolab models, namely zero migration and infinite migration. Infinite migration effectively corresponds to the panmictic case again, whereas zero migration corresponds to a number of cells, independent of each other, each running the panmictic model. So we would expect in the case of zero migration, the diversity (in the long run) should be proportional to the number of cells (or the total area). The in between case of finite nonzero migration should also show an increase in diversity with area, due to partial

independence of each cell, but the increase should be sublinear, as migration causes some species to be identified between cells. Island Biogeography (MacArthur and Wilson, 1967) theory postulates that the relationship is $D \propto A^{-s}$ for some coefficient s , which presumably must depend in some fashion on the migration rates, but is generally in the range 0.2–0.35 for most empirical studies.

May's Stability Criterion

May (1972) proposed that random Lotka-Volterra webs would be unstable if

$$n_{sp} < \frac{1}{s^2 C} \quad (2)$$

where C is the connectivity, defined as the proportion of nonzero elements in β , and s is the interaction strength, defined as the standard deviation of the offdiagonal terms of β , divided by the average of the diagonal terms. Cohen and Newman (1985) showed that May's criteria does not hold for Lotka-Volterra systems in general, only a smaller class related to the models May studied. However, the inverse relationship between species number and connectivity does appear to hold (Pimm, 1982; Cohen and Newman, 1988; Cohen *et al.*, 1990).

Stability is not a relevant property in Ecolab, as really the persistent state (which includes the stable state

as a special case) is the attractor. However, the inverse relationship between diversity and connectivity does hold (Standish, 1998), for spatially dependent as well as panmictic cases. Therefore, in order for diversity to show an increasing trend, a corresponding decreasing trend must occur in connectivity. This ought to be true of the biosphere also, given the universality of this relationship.

As mentioned in section , the mutation operator is highly conservative with respect to connectivity. It assumes that a new species inherits the same connections as its parent, with random additions or deletions according to a symmetric distribution (just as likely to gain a connection as lose one). This has the effect of preserving the connectivity over time. In order for connectivity to decrease, different dynamics would need to be proposed, for example assuming that the mutant species did not compete with its parent.

One possibility for the cause of this growth in diversity is the mass extinctions, that have occurred a handful of times throughout the Phanerozoic. However, the only reasonable way of modeling this is to remove a random proportion of species from the ecology at a particular time. This operation does not alter the connectivity, as the links lost is exactly balanced by the reduced diversity. When implemented within Ecolab, one gets the characteristic rebound in diversity after the extinction event, however, the rebound is back to about the same diversity level as existed prior to the extinction.

Another possibility that actually would work in the right way is related to the fact that the Phanerozoic era corresponds to the breakup of the Pangaea supercontinent — firstly into Gondwana and Laurasia, then into the six continents we know today. Assuming that there is almost no migration between the continents (thus 6 equal-sized continents would support 6 times the diversity of one continent that size) and that the species-area law within a continent has $D \propto A^3$, we would expect that a breakup of a single supercontinent into 6 equal sized pieces should produce $6^{1-3} = 3.5$ times the diversity of the original supercontinent. This factor accounts for a significant fraction of the diversity growth since the Permian.² (Benton, 1995)

Clearly this is a very rough “back of the envelope” calculation, but it is sufficient to show that continental breakup needs to be allowed for in determining if there is any intrinsic evolutionary processes driving diversity growth.

²In case anyone thinks that this result is an argument in favour of habitat fragmentation for promotion of diversity, this is a question of scale. Over short timescales habitat fragmentation is bad for diversity, as is any major environmental change. Only over evolutionary timescales will the diversity bounce back.

Acknowledgements

I wish to thank Mark Bedau for many illuminating discussions, and for assistance in developing the neutral shadow model for Ecolab. I also wish to thank the *New South Wales Centre for Parallel Computing* for computational resources required for this project.

References

- Chris Adami, Ryoichi Seki, and Robel Yirdaw. Critical exponents of species-size distribution in evolution. In Chris Adami, Richard Belew, Hiroaki Kitano, and Charles Taylor, editors, *Artificial Life VI*, pages 221–227, Cambridge, Mass., 1998. MIT Press.
- Per Bak and Kim Sneppen. Punctuated equilibrium and criticality in a simple model of evolution. *Phys. Rev. Lett.*, 71:4083, 1993.
- Mark A. Bedau, Emile Snyder, and Norman H. Packard. A classification of long-term evolutionary dynamics. In Chris Adami, Richard Belew, Hiroaki Kitano, and Charles Taylor, editors, *Artificial Life VI*, pages 228–237, Cambridge, Mass., 1998. MIT Press.
- M. J. Benton. Diversification and extinction in the history of life. *Science*, 268:52–58, 1995.
- J. E. Cohen, T. Luczac, C. M. Newman, and Z.-M. Zhou. Stochastic structure and nonlinear dynamics of food webs. *Proc. R. Soc. Lond. B*, 240:607–627, 1990.
- J. E. Cohen and C. M. Newman. When will a large complex system be stable? *J. Theo. Bio.*, 113:153–156, 1985.
- J. E. Cohen and C. M. Newman. Dynamic basis of food web organisation. *Ecology*, 1988.
- John H. Holland. *Hidden Order: How Adaption Builds Complexity*. Helix Books, 1995.
- R. H. MacArthur and E. O. Wilson. *The Theory of Island Biogeography*. Princeton UP, Princeton, 1967.
- R. M. May. Will a large complex system be stable. *Nature*, 238:413–414, 1972.
- M. E. J. Newman. A model of mass extinction. *J. Theo. Bio.*, 189:235–252, 1997.
- S. L. Pimm. *Food Webs*. Chapman and Hall, London, 1982.
- Tom Ray. An approach to the synthesis of life. In C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, editors, *Artificial Life II*, page 371. Addison-Wesley, New York, 1991.
- R. K. Standish. Embryology in Tierra: A study of a genotype to phenotype map. *Complexity International*, 4, 1997. <http://www.csu.edu.au/ci>.
- R. K. Standish. Statistics of certain models of evolution. *Phys. Rev. E*, 59:1545–1550, 1999.
- R. K. Standish. The role of innovation within economics. In W. Barnett, C. Chiarella, S. Keen, R. Marks, and H. Schnabl, editors, *Commerce, Complexity and Evolution*, volume 11 of *International Symposia in Economic Theory and Econometrics*, pages 61–79. Cam-

- bridge UP, 2000.
- Russell Standish. Cellular Ecolab. In Russell Standish, Bruce Henry, Simon Watt, Robert Marks, Robert Stocker, David Green, Steve Keen, and Terry Bossomaier, editors, *Complex Systems '98 — Complexity Between the Ecos: From Ecology to Economics*, page 80. Complexity Online, <http://life.csu.edu.au/complex>, 1998. also in *Complexity International*, **6** <http://www.csu.edu.au/ci>.
- Russell K. Standish. Ecolab documentation. Available at <http://parallel.acsu.unsw.edu.au/rks/ecolab.html>.
- Russell K. Standish. Population models with random embryologies as a paradigm for evolution. In *Complex Systems: Mechanism of Adaption*. IOS Press, Amsterdam, 1994. also *Complexity International*, **2**, <http://www.csu.edu.au/ci>.
- S. Wolfram. Cellular automata as models of complexity. *Nature*, 311:419–424, 1984.