Evolvability in the Context of the Biosphere

Stephen Jones
387 Riley St, Surry Hills, NSW, 2010, Australia
sjones@culture.com.au

1: Introduction

Bedau *et al* [Bedau *et al,* 1997, Smith and Bedau, 1999] report an investigation into the quantification of adaptive behaviour as shown in various models of artificial evolution. They compare the accumulation of adaptive changes in the genotypes of several Artificial Life (AL) "organisms" with the cumulative adaptive behaviour of the same systems neutralised in such a way as to remove the evolutionarily adaptive aspects of the development of genotype diversity in the systems, replacing it with the results of an essentially random progress through activity space.

They also compare the evidence of adaptive evolution in the fossil record with a similarly derived neutral fossil record. In the artificial models the results show a series of increasing difference from the neutral level which then, at certain points, collapses back to almost the zero condition with a spike in the diversity statistics for that "moment". But they find a very interesting difference in the behaviour of the statistics where those statistics are derived from the actual data of fossil families in the biosphere. The difference is a continual increase in cumulative activity and diversity which appears to approach the exponential, and presumably reflects differences in the evolvability of the two types of system. Evolvability is here thought of as the capacity of a genotype to adaptively accommodate long-term changes in its context.

The question left hanging in the Bedau, *et al*, papers is the source of the difference in the evolvability of model versus natural systems. They make several suggestions which I shall look at, briefly, and then I shall offer further suggestions and some low-level formalisation in order to allow some development of model systems which may show similar characteristics.

In their discussion Bedau, *et al*, [Bedau *et al,* 1997, Smith and Bedau, 1999] canvass several differences between artificial models of evolution and the natural model.

1: In the Bugs and Evita models there are no "interesting interactions between organisms; no predator-prey connections, no cooperation, no communication, nothing." [Bedau *et al,* 1997]. In Tierra evolvability is largely determined by the functions of Tierran genetic operators (*eg*. mutation or insertion) [Ray, 1994]. In Echo although genotypic diversity increases it converges towards a simple trading ecology [Smith and Bedau, 1999]. Bedau *et al* conjecture that neither of these systems would show any greater propensity for evolvability than the Bugs or Evita systems [Bedau *et al,* 1997].

2: There may be an environment size issue. Obviously model systems do not have the vast environmental scale of the natural system, but they suggest that this will not change the essentially "qualitative character" of the statistics. Size of the system by itself should presumably only offer a constant multiplier factor to the numbers without affecting diversity and thus would only become relevant when the population of organisms outruns the available environment size.

3: Likewise making the available behaviours of the model systems more complex will not make the system behave more like the natural system. I suggest the real effect of complexity comes from interactions made available through the diversity of organisms and environmental niches not just the complexity of behaviour of any particular organism.

As Bedau, et al, suggest "the primary reason behind the biosphere's arrow of cumulative activity is that the dynamics of the biosphere constantly create new niches and new evolutionary possibilities through interactions between diverse organisms. This aspect of biological evolution dramatically amplifies both diversity and evolutionary activity..."[Bedau *et al,* 1997]

This paper concentrates on the explanation and amplification of this quote. Two reasons for this effect may be operating. First: the individual organisms in the system are active and autonomous, carrying out processes which emulate sensing and possibly communication and, assuming autonomy of the organisms, some level of intentionality [Jones, 1999]. Second: the environment itself will not be entirely neutral or even passive, but will undergo various degrees of impact from the organisms in it and will have varying degrees of impact on those organisms. That is, there are (potentially) multiple layers of complex relations between organism(s) and environment.

Thus the exponential growth of genetic adaptation in a biosphere as compared to the growth of genetic adaptation in current Artificial Life algorithms may be due to two aspects of the nature of these complex relationships between the organisms and their environments.

A: The nature of the interactions had with the environment. That is, the extent to which the organisms can have a changeable effect on the environment and the actual content of the environment at any particular epoch. These interactions include:

1: Interactions between organisms, *eg*. predator-prey.
2: Feed-forward effects of the organism's output onto the environment,
3: Feed-back effects of the environment on the organism,
4: Combinations of both feed-forward and feed-back as new feed from the environment, and
5: Cross-feed of the effects of other organisms' interactions with the environment.

There will be processes of delay in the impact of the combinations of feed-back from the organism and the, now different, feed-forward from the environment. The combinatorics may range from simple addition to major chemical interaction, to catastrophic loss of genotypes and/or environment. Cross-feed effects are generally embedded in the behavioural interaction between organisms and the consequents in the environment of aspects of this behavioural activity. They may be simply molecular or they may range across complexity-layers in the form of predator-prey relations or collaborative social behaviour.

B: The complexity of the systems of environment and organism and their interactions. There are several layers of complexity available to be involved.

1: Genotypic complexity within organisms through diversification of species over time,
2: Inter-organism and inter-species interaction through sensory, communicational and predator-prey interactions as a result of increased genotypic complexity,
3: Organism-environment interaction and consequences as a result of increased genotypic activity,
4: Complex intra-environmental interactions and their consequences onto organisms,
5: Appearance of "cultural" (essentially memetic rather than genetic) properties in interactions within and between genotypes.

A major difference between AL systems and "natural" systems is that the AL systems usually interact *within* only one layer whereas biosphere systems are able to interact across *several* layers at once. As genotypic diversity becomes more complex it will produce more complex interactions both between organisms and between organism and environment. For example, the appearance of new genotypes may promote the production of new resources into the environment as well as altering an organisms' capacity to handle different resources in the environment. The effect on the environment is to make it more complex, thus placing more diverse selection criteria on organisms and requiring more diverse adaptability. This in turn compounds the drive to genetic diversity.

Environments themselves are layered in complexity as well. These layers of environments may be seen as:

1: The passive environment which may have fixed food supply, the memory of the computer in which the system is running, the cell-culture medium in the biology lab (a system in which only sensory processes apply). These exist, largely, only in computer embodied simulations and instantiations, though this may also be the character of large scale monocultural agriculture.
2: An active environment in which multiple genotypes interact with the environment and directly change it, more like the ecological environment (a system in which sensory and communicative processes will be available to the organisms), and
3: The memetic cultural environment of the classes of information which are the vehicles of social interaction and in humans at least cultural/political interaction.

2: Organisms

What is it about an organism that can be directly affected by the environment such that its adaptive capacity is tested and through which the effectiveness of any adaptive changes to the genotype is discovered? Adaptive capacity would be the capacity of the organism to *use* the resources in its genes to handle day-to-day changes in its environment, whereas adaptive changes to the genotype allow whole species of organisms and their descendents to *have* the necessary resources in their genes, that they may be used. This is the difference between genetic change (eg. by mutation) and adaptation through learning. But at the same time it is clear that there is not a simplistic separation of the two processes in determining the survival of an organism with respect to the fitness criteria, but that they are intertwined. The use of an adaptive capacity may well be crucial in bringing the organism to reproduction.

An organism is essentially a package of some sort containing a string of program codes (in DNA or some other instruction set, either preset or evolved) embedded in the package within some sort of support framework, akin to the concept of a cell. This support framework will afford the necessary components and structures to allow the string to be operated upon so that instructions can be executed. In evolutionary systems there will be some way(s) in which the instructions string can be modified either by random mutation, some form of cross-over (*eg*. through sex) or by intervention from outside the system. Further it is necessary for the cell to be maintained and, again in evolutionary systems, for the cell to be reproduced [Ray, 1994].

The complexity of interactions available in the system will also contribute to genetic adaptation. The autonomy of an organism will bring internal feedback relations in which aspects of the internal system can emphasise the regulation of their local environment in intentional ways. When this spreads outside the organism-system you get social cultures-environments in which organisms communicate, sense and have intentionality. This then modulates genetic expression and genetic reproduction and may also alter mutation rates (by at least some sort of mechano-chemical impact on the DNA or the simulation of such). These more complex systems will diversify the nature of the selection criteria for any particular genotype within that environment.

The relations that an organism can have with its environment will have effects on the organism in two ways which can be characterised as changing (or adapting) the organism's or genotype's memory of how it adapted to any previous state of the environment. Within an organism there are two kinds of memory operating, the first is the genetic information in the DNA (or equivalent) which I call Darwinian memory and the second is the memory that each organism develops of its experience in the environment. I call this Lamarckian memory and it is a kind of learning

3: Environments

An environment is the container in which the organism (as genotype+package) operates. It may carry other contents such as food and (simulated) chemical products of the organisms' real or simulated metabolism. There will generally be a number of organisms of similar or different genotypes operating in the environment.

An environment will be either neutral (as in most, if not all, current AL systems) or it can be changeable through some class of interactions from the organism, or it can be active in the sense that the biosphere is. In a neutral environment, for example in the "Bugs" system, [Packard, 1989] there may be food or some other source of "energy" but there will be no direct effect from the organism onto that environment. The use of the food will be, as far as the environment is concerned, a passive act. In a changeable environment it is usually the case that the environment is changed from outside the total system organism+environment (O+E). This is the case, for example, in Tierra [Ray and Hart, 1998] because it is running inside otherwise normally employed computers doing their normal day-to-day tasks and only being Tierra when their owners are away from their desks. The environment changes according to the whims of the computer's owner. Tierra also changes the environment by changes in the population density of Tierran organisms in any particular computer, this will slow down the available machine-cycles-per-second-per-organism perhaps prompting some organisms to emigrate their daughters to another machine, but it does not directly effect the production of adaptive variations in the genotype. In a system in which the organisms generate products into the environment, which products may or may not be useful to other genotypes operating within the same environment, then the environment could be described as being "active" in the sense that we need here, though it is more a sense of being reactive. This is what happens within biological ecosystems.

There are essentially two types of interactions possible between the environment and the organisms/programs residing in that environment.

1: Outputs: an organism of sufficient complexity may produce by-products of "metabolic" processes as waste excreted into the environment. I don't know of any AL systems that do this but natural systems do it all the time. An organism may also secrete signals into the environment for purposes of probing the environment for useful information or for the development of communications with another organism. These by-products and signals may be behavioural as much as they are in any way, molecular processes. In fact molecular processes are a very low level layer of behavioural processes.

2: Inputs: Organisms of any autonomy will need food/fuel/energy resources which they will intake upon recognition, even if this only means that the food is in the current occupied space (*eg*. Bugs). If the organisms have any sensory input then they will input information of some sort about their environment, this may be information about food or about other organisms in the environment or any other representable environmental content. Organisms may also be "tricked" into inputting "poisons" by some kind of similarity to say "food" items. Again these inputs may be entirely behavioural as well as molecular.

Environments that are anything but neutral may be modified in various ways. Organisms which excrete wastes or secrete signals will inject this "matter" into the environment through some sort of "pore" or port in the boundary that distinguishes the organism from its environment. This material then exists in the environment changing the nature of the environment to some degree. If an organism is able to *use* the contents of the environment in some way then that organism will take up or input that material in through its boundary. Thus the content of the environment will be dynamic, changing over time as various cycles occur in the organisms operating within that environment. Further this environmental content being largely a product of the existence and processes of various organisms my be as much behavioural as molecular to the state that perhaps all inputs or outputs between organism and environment are layers and classes of behaviour.

A number of things can happen to significantly alter the state/content of the environment.

1: An organism may be massively productive of some sort of output and the output material then floods the environment.

2: An organism may produce outputs which are deleterious to other organisms in the environment, thus poisoning them.

3: An organism may need more input material than is available from the outputs of other organisms or the passive content of the environment and the organism may starve.

4: An organism may be too successful and reproduce too many children increasing the population pressure in the environment

5: In the case that there are multiple genotypes in the environment different behavioural-consequents in that environment will affect various genotypes in possibly very different ways

Usually changes in the environment will cause changes in an organism and may cause changes in a genotype. Also any change in the environment is going to mean a change in the fitness criteria selecting out which genotypes will continue to be appropriate to the current state of the environment thus affecting the survival of organisms in that environment. With any particular environmental change the survival rate of some organisms may be enhanced while for others it may diminish. It should also be recognized that in the biosphere the death of an organism means that its substance is left in the environment becoming resources for other organisms in that environment.

Some changes in the environment may well change the *expression* of certain genes/codes in the instruction string but this does not change the genotype as such. Changes in the genotype only happen in the interval of reproduction of an organism. They occur usually as mutations in the genotype, *some* of which *may* be positively adaptive to the new environmental conditions.

Changes in a genotype may also accumulate through the process of crossover in organisms using sexual reproduction. Where sexual reproduction is used the organisms will usually be complex enough to carry slight variations in their genes while still being of the same overall species/genotype.

4: Formalisation

To a large extent model evolutionary systems have concentrated on the genotypic aspect of adaptive change and have neglected the environmental contribution of interactions within and between organisms and between organisms and their environments, even when the environment consists solely of other organisms. Adaptive change of a genotype can only happen through alterations in the gene-string while within a particular organism adaptation can occur through learning and the variation of gene expression depending on environmental stimulus.

Feedback from the organism-as-a-system onto the gene through system-control of gene expression gives the organism a flexibility to handle immediate changes in the environment, some of which will change the reproduction rate and thus genetic adaptation. But it does not change the actual content of the genotype string, and only operates if the genotype already carries codes for better adapted proteins. Where alternative codes do not confer positive adaptive effects then the genotype fitness will degrade within that particular environment.

Here, then, I conceptualise aspects of an algorithm for an evolutionary system in which the environment is active in the sense that activity of the organisms in the environment leaves products/materials of varying kinds within the environment and that the organisms ingest materials found in the environment, thus effecting changes in the state of the environment. There may also be further chemical processes between materials in the environment further complexifying the situation.

As the environment changes, fitness criteria alter over longish sequences of reproductive cycles in the system O+E and so the genotype of an organism may only render it fit for the environment for short proportions of the system's total history. The hypothesis is that this variability in the environment should stimulate an increase in genetic diversity, and cumulative adaptive activity should approximate something more like the (possibly exponential) curve in fig.1 of Bedau [Bedau *et al,* 1997].

The active environment might be likened to a large-scale system in which feedback from the context onto the generators in the context will increase the tendency for those generators to change. Since changes can occur only in the reproductive interval, because of the types of change we are restricted to by definition of an evolutionary system, organism learning and the adaptive capacity of flexible gene transcription are initially ruled out. Where these survival adaptations become available the cumulative adaptive genetic variation will not be changed but survival periods for individual genotypes will be increased (and the rate of adaptive genetic variation may decrease). This may in fact restrict the rate of growth of the curve. A crucial exception arises where the age of reproduction is some significant proportion of the average lifetime of the organism (say approx 1/5th) and so the intra-organism adaptive flexibility through learning will increase the likelihood of reproduction thus also increasing opportunities for adaptive genotype mutation and crossover.

Organism-Environment interactions

I suppose a system of several genotypes (organisms) G1, G2, ...Gn in an environment E.

G1, ..., Gn will be using the resources of E for various survival tasks.

Since G1, ...,Gn are wholly within E (by definition) they use (*ie*. they ingest) resources R1(in), …, Rn(in) contained in E,

and they excrete/secrete substances P1(out), …, Pn(out) into E.

Under some versions of a possible system any of

P1(out), …, Pn(out) may = R1(in), …, Rn(in)

OR any of

P1(out), …, Pn(out) may = -(R1(in), …, Rn(in)).

The sign indicating whether the effect of ingestion of R1, …, Rn is positive (beneficial) or negative (deleterious) for G1,....,Gn

Consider the following possible scenarios

1: Simple gene expression

G1 --> P1

where --> reads “yields” or “produces”,

and P1 is a wholly intra-organism protein/product.

This system will have no effect on the environment

2: Complex gene expression

2a: G1 + Rx --> P1

where Rx is a resource from E required for the expression of the gene for P1,

and + reads "in the presence of".

pop(G1) or met(G1) then depends on conc(Rx) in E,

where pop is “population of”, met is “metabolic-rate of” and conc is “concentration of”.

If Rx is in short supply in E then pop(G1) will tend to decrease, and *vice versa*

If Rx is in plentiful supply in E then, *ceteris paribus*, pop(G1) will tend to increase

Positively adaptive mutations in G1’s descendents (G1desc) will then lead to increase in pop(G1desc).

2b: G1 + Rx --> P2

where P2 is secreted as a signaling molecule.

2c: G1 + Rx --> P1 & Pw

where Pw is an excreted waste product.

3: Multiple genotype processes

3a: G1 + R1 --> Pw1 AND G2 + R2 -->Pw2

where R1 = Pw2 AND R2 = Pw1

will produce a mutually beneficial relationship IFF pop(G1) = pop(G2) AND G1(prod(Pw1)) = G2(prod(Pw2))

where *eg*. G1(prod(Pw1)) is G1's rate of production of Pw1.

IF pop(G1) >> pop(G2) then E(conc(Pw1)) >> E(conc(Pw2))

and G2 will find much more R2 (= Pw1) available for use and there will be, *ceteris paribus*, a tendency for the relative populations of G1 and G2 to equilibrate through E “selecting” for G2

XOR

conversely for the situation where pop(G1) << pop(G2), E “selecting” for G1

XOR

IF G1(prod(Pw1)) >> G2(prod(Pw2)) then the lack of availability of Pw2 = R1 will be a rate-limiting condition on G1 + R1 --> Pw1 and again, *ceteris paribus*, there will be a tendency for the relative populations of G1 and G2 to equilibrate through E “selecting” for G2

XOR

conversely for the situation where G1(prod(Pw1)) << G2(prod(Pw2)),

E “selecting” for G1

3b: G1 + R1 --> Pw1 AND G2 + R2 -->Pw2

where R1 = -(Pw2) AND R2 = -(Pw1)

will produce a situation in which G1 or G2 may find itself being seriously selected against by E, because of, *eg*. G1 flooding E with what would be a “poison” for G2 or *vice versa*.

4: Another type of selection process in E against Gn manifests through the existence of multiple systems Gn(1, …, inf) [where inf reads “infinity”]. If the “behaviour”-output in E of Gn(1) is deleterious for Gn(/= 1) [where /= reads “other than”] the environment's selection pressure against Gn(/= 1) increases. In any system of Gn in E where Gn shows the number of Gn types in E > 1, then if these Gn types output any behaviour which is communicative and even intentional, this output may be regarded as behavioural and it implies that the system now contains a "social" culture. This behaviour can be either of negative, neutral or positive consequence on other Gn. For some Gn the particular value of behavioural consequence may range from -inf (deadly) … neutral … +inf(valuable) for each Gn, with any one Gn potentially having a totally different view of the sign-value of the behavioural consequence. This is the operation of a natural ecosystem because of the large number of Gn in the current E.

It is where the activity of one type of Gn in E is deleterious for some other Gm in E that E can be said to be selecting against that Gm. Under these conditions of E it is encumbent on the particular Gm being "selected" against to produce adaptive mutations which may then help its survival in E, for example by producing a mutation which changes the sign value of R from - to +. The conditions in 3a: tend to be self-regulative within the system E + (G1, ..., Gn) whereas the conditions in 3b: will force potentially catastrophic alterations in the fitness criteria for some particular Gn. 4: shows up as a more general version of 3: in that outputs now have a wider range of classes which include the predator-prey interaction through to teaching a child language. The kinds of outputs in 3: are simply a subset of the outputs available to 4: but they appear in lower complexity-level systems and are spawned as Darwinian memory with the system. Cultural memory as behavioural outputs is of course of the Lamarckian memory type

I have begun to model here an active environment which shows some of the characteristics of the natural biosphere of organism + environment. I consider it necessary to start to build computer programs from these scenarios to test the impact of a simulated active environment on the count of accumulated adaptive mutations mentioned in [Bedau *et al,* 1997].

References

Bedau, M.A., Snyder, E., Brown, C.T. and Packard, N.H. (1997) "A Comparison of Evolutionary Activity in Artificial Evolving Systems and in the Biosphere" in Husbands, P. and Harvey, I. (eds.) *Proc. 4th European Conf. on Artificial Life, ECAL97*. MIT Press/Bradford Books

Jones, S. (2000) "Sensing, Communication and Intentionality in Artificial Life" in Sugusaka, M. and Tanaka, H. (eds.) *Proc. 5th Int. Symposium on Artificial Life and Robotics*, Oita, Japan. pp26-29.

Packard, N.H. (1989). "Intrinsic Evolution in a simple model for evolution". In C.G.Langton, (ed.) *Artificial Life* (pp.141-155) Redwood City, Calif. Addison-Wesley.

Ray, T.S. and Hart, J. (1998) "Evolution of Differentiated Multi-threaded Digital Organisms" in: *Artificial Life VI Proceedings*, C.Adami, R.K.Belew, H.Kitano, and C.E.Taylor (eds.), pp295-304. The MIT Press, Cambridge. <http://www.hip.atr.co.jp/~ray/>

Ray, T.S. (1994) "An Evolutionary Approach to Synthetic Biology: Zen and the Art of Creating Life." *Artificial Life 1*: 179-209

Smith, R.M., and Bedau, M.A. (1999) “Is *Echo* a Complex Adaptive System” in *Evolutionary Computation* 7(1) pp45-68.