No Pain, No Gain: Landscapes, Learning Costs and Genetic Assimilation.

Giles Mayley
School of Cognitive and Computing Sciences,
University of Sussex,
Brighton, BN1 9QH, England.
Tel: +44 1273 606755 ext 3073.
email: gilesm@cogs.susx.ac.uk

CSRP 409
February 1996

Abstract

The evolution of a population can be guided by adaptive traits that are acquired by members of that population during their lifetime. This phenomenon, known as the Baldwin Effect, can

No Pain, No Gain: Landscapes, Learning Costs and Genetic Assimilation.*

Giles Mayley
School of Cognitive and Computing Sciences,
University of Sussex,
Brighton, BN1 9QH, England.
Tel: +44 1273 606755 ext 3073.
email: gilesm@cogs.susx.ac.uk

February 1996

1 Introduction

In 1896 J.M. Baldwin first identified a 'new factor' in evolutionary theory that has subsequently become known as the Baldwin effect (Baldwin, 1896). Baldwin was working at a time when there was still great debate between followers of the Darwinian and Lamarckian evolutionary frameworks (for historical review see Simpson (1952)). He was addressing the problems associated with fitting observable examples of apparent genetic specification of acquired characteristics into the Darwinian theory. That is, examples exist in nature in which traits that are first acquired by members of a population during their lifetimes, through adaptive processes, have become genetically fixed in subsequent generations[1]. The Lamarckian theory postulated a mechanism for the direct inheritance of characteristics acquired by a parent to its offspring through the transfer of information from phenotype to genotype. However, the Darwinian (non-Lamarckian) framework excludes such a mechanism and subsequent research into molecular biology has found no evidence for it. Baldwin's answer to this was to extend the idea of natural selection to include what he called *organic selection* which is a process whereby individuals could improve their chances of survival by adapting during their lifetime. That is, an individual that increased its selective advantage during its lifetime by learning a specific trait would pass on the learning ability to its offspring:

> 'The most plastic individuals will be preserved to do the advantageous things for which their variations show them to be most fit, and the next generation will show an emphasis of just this direction in its variations.' (Baldwin, 1896)

Baldwin then argues that these learnt traits can become genetically specified as evolution then

> 'Congenital variations, on the one hand, are kept alive and made effective by their use for adaptations in the life of the individual; and, on the other hand, adaptations become congenital by further progress and refinement of variation in the same lines of function as those which their acquisition by the individual called into play. But there is no need in either case to assume the Lamarkian *[sic]* factor.' (Baldwin, 1896)

At the time of Baldwin there was little known about the mechanics of evolution and so he was unable to establish the mechanisms through which acquired traits become genetically specified, resorting to vague terms such as 'refinement of variation'. By the 1950's, researchers were in a position to do so. Simpson (1952) suggested that a trait remained acquired until an advantageous mutation or set of mutations entered the gene pool that produced the same trait. Waddington (1953a) argued that this meant that there was no connection between the acquisition and subsequent *genetic assimilation*[2] of the trait, rendering the theory signifying very little. Instead, Waddington put forward his own ideas on developmental canalization (Waddington, 1942). He postulated that the development of environmentally driven adaptations is likely to become canalized by natural selection to produce the optimum response independent of the level of environmental stimulus received by a particular individual. Once canalized, the response can then be triggered by a variety of factors. That is, natural selection has provided itself with a means by which the desired genetic effect is likely to occur. Waddington provided laboratory support to his theories with examples of genetic assimilation occurring in populations of the fruit fly, *Drosophila melanogaster*, (Waddington, 1953b, 1956). However, as pointed out by Maynard Smith (1993), the large variability in the assimilated phenotypes indicated that full canalization had not taken place. Also, Waddington himself admits that in at least some of the experiments the changes were due to single mutations and that it was statistically possible that the assimilation was purely a chance effect with little connection between the environmental and genetic changes. Maynard Smith points out that Waddington's experiments also deviate from his theory in that the acquired characteristics were not adaptive to the stimulus that produced them. Maynard Smith's interpretation of the events seems the most plausible. The differences between individuals that are disguised by canalization are then exposed at times of environmental change, giving natural selection a variation on which to act. Individuals that are more able to adapt to their environment will then be selected for.

This paper attempts to investigate the conditions under which genetic assimilation takes place through the use of a series of computer models. First of all, a framework is given which describes the evolutionary circumstances under which the arguments, presented later on, are valid. The assumptions that have been made during this work are outlined in this section. It then goes on to outline the benefits that learning can give an evolving population but notes that there are also various costs to be paid by an individual for that ability to adapt during its lifetime. The evolutionary trade-off between these costs and benefits provides selection pressure for evolution to first exploit the benefits and then reduce the costs of lifetime adaptation such that acquired characteristics appear in the population and are then assimilated into the gene pool. Next, a property of the genotype to phenotype mapping, which has been named *neighbourhood correlation*, is described and the implications of this property on the combination of learning and evolution given. The cost of learning and neighbourhood correlation ideas are then applied to the Hinton and Nowlan (1987) model of learning and evolution. An series of new experiments is then described in which the ideas presented in this paper are more explicitly applicable.

[2] The term genetic assimilation was coined by Waddington and is the one that will be used throughout this paper to refer to the genetic specification of acquired characteristics.

individual's ability to learn is also under evaluation in the fitness trial as well as its performance in the given task after the learning process is complete. That is, the individual is under *continual assessment*. This is more analogous to biological systems. In this work, the presence of genetic assimilation in evolving populations using both assessment schemes is investigated through the use of a cost of learning function (section 3.2).

In most artificial systems that combine learning and evolution, an individual is initially generated from its genotype; it is then modified in some way by the learning process. Hopefully this will lead to an increase in its performance at the desired task. We thus have a distinction between its abilities at the beginning and end of its learning trial. If the unmodified individual has its fitness evaluated it will achieve a particular score. Let us call this the individual's *innate fitness*. If the individual is evaluated after it has learnt it will receive a different fitness. This is the score that the individual would be awarded in the posthumously assessed learning scheme above but it would receive rather less in the continually assessed scheme since it spent some time learning. This distinction between different scores is important to the arguments presented in section 3.2.

3 The Costs and Benefits of Learning

3.1 The Benefits of Learning

This section describes some of the benefits that learning can bestow on a population of evolving individuals.

The first is Baldwin's idea of organic selection in which he noted that evolution was almost able to 'predict' the direction in which to go. Adaptive members of the population are able to 'find' new advantageous behaviours that less plastic individuals are unable to perform. This means that these adaptive individuals gain the upper hand and are selected for. Thus evolution is guided by the actions of the individuals on which it operates. Secondly, Ackley and Littman (1991) state that the combination of learning and evolution increases the spatiotemporal bandwidth of the environment that a system can adaptively respond to. This means that a learning individual has the ability to cope with changes in the environment that are at a faster time scale than that on which evolution operates; and also to adapt to varied local spatial environmental differences for which evolution would have to specify a number of different responses. That is, it is more beneficial for evolution to provide a general purpose adaptive mechanism to cope with local variation than to provide several fixed behaviours to cover that variation.

Lastly, a learning mechanism may be able to provide an individual with behaviours that are simply very hard to evolve. For example, in humans, it is very difficult to imagine the English language being genetically specified at birth, even if the language were static and universal. It is better for evolution to provide an innate tendency to acquire a structured language and let learning sort out the details.

3.2 The Cost of Learning

The previous section outlined some of the benefits that learning can give an evolving individual. It is probably less obvious that there are also evolutionary costs to be paid by an individual for that ability to learn. Firstly, there are the increased *energy costs* that must be invested in an adaptive mechanism over a genetically fixed one. This is true in both natural and artificial systems. Assuming a learning mechanism is more complex than a fixed one, the ontogenetic process of development as well as the energy expended during the lifetime in the learning process itself will cost the individual energy that could be used in other pursuits. In the production of an artificial system, the constraints on its performance are largely economic. We can break this down into costs concerning the adequate provision of development time, CPU time, materials etc.

Once again, assuming that an adaptive architecture is more complex than a fixed one, the costs of programming and CPU times are factors to be taken into account when evaluating a system's performance. These costs are the artificial analogue of biological energy costs.

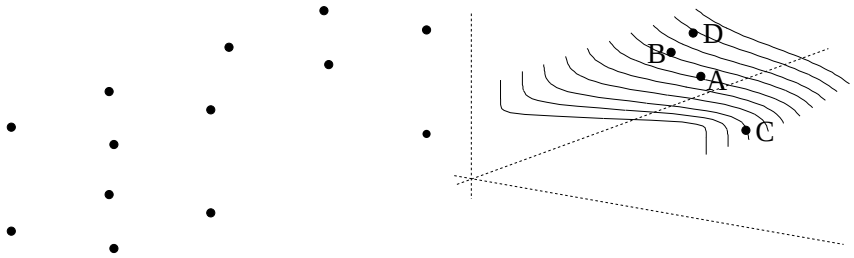
Secondly, learning means the individual increases its fitness in some way. This implies that the individual spends some period of its lifetime performing poorly. If we look upon the final fitness score that an individual achieves in an extended lifetime as an accumulation of the acquisition of small fitness increments, as is the case in of continually assessed learning, then during this period the individual is not adding as significantly as it might to its fitness, i.e. it is incurring a *time-wasting cost*. It may also be actually reducing its fitness by performing tasks incorrectly and accruing a 'cost of getting it wrong' or *incorrect behaviour cost*. This may have little significance, such as a bird eating a bad-tasting berry, but conversely may cause the individual irreparable damage, such as falling down a precipice (learning the hard way).

If the individual expends vast amounts of energy, takes a significant proportion of its lifetime or damages itself irrevocably whilst learning to perform a task to the pinnacle of perfection, it may be better, in evolutionary terms, to adopt a less perfect, genetically fixed solution to the problem. Since this behaviour is innate, a non-learning individual performs the given task right from the start of its lifetime and therefore incurs no time-wasting cost. It expends no energy adapting itself and its adequate performance ensures that, although it does not excel, it rarely does badly, avoiding incorrect behaviour costs.

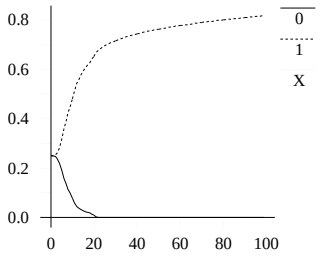
These costs are identical in biological and artificial systems as long as the individual is evaluated using the continual assessment scheme outlined in section 2. However, if the posthumous assessment learning scheme is used where the process of learning is not under evolutionary evaluation then these costs are excluded and the individuals only receive the benefits of learning.

3.3 Cost / Benefit Trade off

The inclusion of learning can both be beneficial and detrimental to the fitnesses of individuals in a population and therefore there is an evolutionary trade-off as to whether it is adopted. Since the assumption was made in section 2 that the existence of learning is under genetic control, then evolution might select against it if the costs are too high.


D
B
A
C

The experiment consists of a population of 1000 genotypes, 20 genes long. The genes can take one of a possible 3 alleles: **0**, **1** or **?**, distributed with initial relative frequencies 0.25, 0.25 and 0.5, respectively. When translated to the phenotype, the genes specify the states of 20 conceptual switches – a **0** corresponding to incorrectly set, a **1** to correctly set. The states of the switches defined by a **?** are determined through a sequence of 1000 learning trials. During each trial, all the **?**s are set to **0** or **1** with equal probability. If, on the ith trial, all the switches are set to **1** then the fitness of the genotype is calculated with the formula $F = 1 + 19(1000 - i)/1000$. Both a genotype specifying any **0**s (which thus has no chance of reaching the target phenotype of all **1**s) or one that has failed to set all its switches after 1000 learning trials will receive the minimum fitness of 1. One that consists of all **1**s, therefore avoiding any learning trials, is awarded the maximum fitness of 20. A standard G.A., applied without the learning trials, has no better than a random chance of finding the target solution of all **1**s since, amongst the 2^{20} possible genotypes (assuming only **0** and **1** alleles), there exists only a single spike of increased fitness. With the introduction of the trials, evolution is guided to the peak by the learning mechanism. That is, the genotypes that just happen to be near the peak attain an increased fitness score by climbing it during their lifetime.



be classified in the terms used in section 3.2 as a time-wasting cost, i.e. i only receives a fitness score once it has learnt J and is penalised by the number of trials it took to get there.

We can use the idea of learning costs and benefits to follow the varying selection pressures for and against learning that are present in an evolutionary sequence. This analysis is very similar to that presented by both Harvey (1993) and Belew (1989) but with the emphasis put on the cost of learning ideas presented here. In the early stages of evolution, most genotypes will contain at least one **0** and so no individuals will receive any selective advantage. This corresponds to the initial flat areas in the graphs in figure 2a. The number of **0**s in the population then falls as individuals that contain only **1**s and **?**s start to take over the population. Individuals with **0**s are strongly selected against in favour of those that can learn the correct phenotype. Those genotypes that contain only **1**s and **?**s receive a selective advantage that is proportional to the difference between the benefit they are receiving from learning and the cost they are incurring. At this stage in the evolutionary run the benefits heavily outweigh the costs. On average, a learning individual will have ten **?**s and ten **1**s. From figure 2b we can see that $B(I, J)$ and (I, J), for these proportions,

1 s as to

exists, with respect to the genetic operators, that links p and q. The use of crossover with no mutation in the HN model makes this point unclear and so a model is developed here that, it is hoped, encapsulates all the points and arguments presented in this paper.

6 Experimental Setup

This section describes a series of experiments that were conducted to look at the issues raised in this paper with regard to the circumstances under which genetic assimilation takes place. Evolutionary runs were conducted with and without learning on two different genotype to phenotype mappings, one of which has the property of neighbourhood correlation whilst the other has uncorrelated neighbourhoods. The learning experiments were conducted first without and then with a cost of learning.

6.1 Genoty e to Phenoty e Ma ings

Two genotype to phenotype mappings were used to show the different behaviours exhibited by the combination of learning and evolution on landscapes that have correlated and uncorrelated neighbourhoods. Both mappings take a genotype that is a binary string of length N and map it to a phenotype that is also a binary string of length N. Thus both genotypic spaces and phenotypic spaces are N-dimensional hypercubes. Learning and evolution are applied to maximise the number of **1**s in the phenotype.

In the first mapping, mapping 1, the phenotype is simply a copy of the genotype. This gives a correlated neighbourhood relationship with respect to the genetic operators (see section 6.2). Small changes in the phenotype correspond directly with small changes in the genotype.

The second mapping, mapping 2, has been designed to have maximally uncorrelated neighbourhoods. What is desired here is a mapping such that, for small changes in the phenotype, the corresponding genotype coding for the new phenotype must be as far away from the original genotype as possible. This was achieved through the use of the following mapping: for genotype of length N, where N is even, the phenotype is a copy of the genotype if there is an even number of **1**s in the genotype, or the inverse of the genotype if there is an odd number of **1**s in the genotype. This gives the relation that for a given phenotype, P, all the genotypes q_i that code for the phenotypes, Q_i, that are a Hamming distance of one away from P, are a Hamming distance of $(N-1)$ away from p, the genotype that codes for P. Note, the inverse of this is also true – small changes in genotype correspond to large changes in the phenotype. Also worth noting is that the relationship is not true for movements greater than one bit flip in either space. This led to a careful choice of genetic operator as described in section 6.2.

To get a better visualisation of mapping 2 we will look more closely at the case when $N = 4$, though the experimental results presented in section 7 use the equivalent mapping for $N = 20$.

$g \rightarrow ph$	00	01	11	10
00	0000	1110	0011	1101
01	1011	0101	1000	0110
11	1100	0010	1111	0001
10	0111	1001	0100	1010

Table 1: *Genotype to phenotype mapping 2 used to create an uncorrelated neighbourhood relationship. The row defines the first two bits of the genotype and the column the second two bits, giving the phenotype in the corresponding cell.*

Table

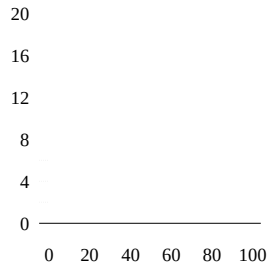
$ph \rightarrow g$	00	01	11	10
00	0000	1110		

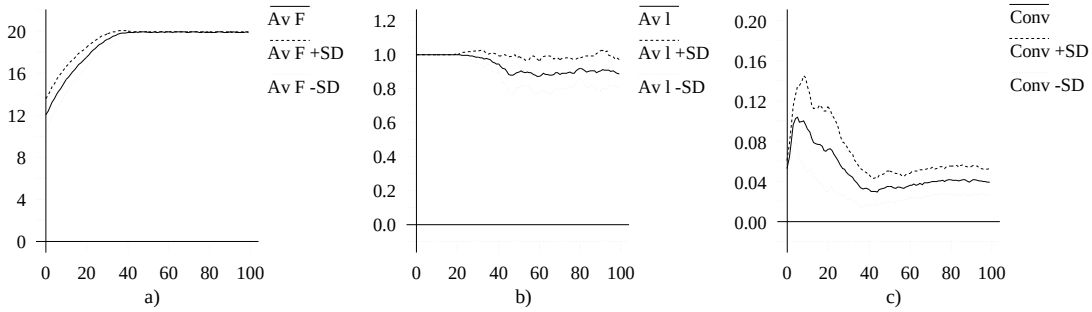
7 Results

The results presented in this section are the mean of twenty runs of the model in each of its six modes. Up to three variables are shown depending on their relevance to the discussion. The first, average fitness, is calculated as the mean of the fitnesses of all the individuals in the population each generation. These means are then averaged over the twenty runs and the standard deviations calculated. The second variable gives an indication as to the proportion of individuals in the population that used learning in achieving their fitness scores. The number of individuals that have been tagged as having used learning, as described in section 6.3, is divided by the number of individuals in the population each generation for each run. The graphs presented are then the mean of these proportions over the twenty runs with the standard deviations shown. The final variable is an indication of the convergence of the population. This is the average Hamming distance from each genotype to a consensus genotype each of whose loci are filled with the most common of the alleles over the population at that locus. The mean convergence over the twenty runs is shown with the standard deviation.

7.1 No learning

The first set of results are from evolutionary runs without learning, with fitness evaluated using equation 3. It can be seen that evolution alone was able to find the global optimum fairly easily on the mapping with neighbourhood correlation (figure 3a). Notably, the population 'unconverges' quite a lot before reconverging on the global optimum (figure 3b).




20
16
12
8
4
0
0 20 40 60 80 100
a)
Av F
Av F +SD
Av F -SD
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0 20 40 60 80 100
b)
Av l
Av l +SD
Av l -SD
0.20
0.16
0.12
0.08
0.04
0.00
0 20 40 60 80 100
c)
Conv
Conv +SD
Conv -SD

$$b = \frac{0.9b + a/10}{a + 0.905b} \tag{9}$$

Solving these for a and b, we get $a = 0.183$ and $b = 0.817$. So, to a first approximation, we would expect 0.183 of the population to have genotypes that consist of all **1**s as an artefact of the mutation operator and the

using equation 5. The average fitness of the population (figure 6a) reaches

8 Conclusion

The arguments and experiments presented in this paper discuss two criteria that are necessary for the genetic assimilation of acquired characteristics: the evolutionary cost of learning and the existence of a neighbourhood correlation relationship between phenotypic space and genotypic space.

For the genetic assimilation of a beneficial acquired characteristic, there must be an evolutionary cost to be paid for having that character adaptive over having it genetically specified. In systems where the learning is continually assessed, such as nature, there will always be the implicit costs that were discussed in section 3.2, i.e. energy costs, incorrect behaviour costs and time-wasting costs. Artificial evolution can also support systems where the learning is posthumously assessed, in which there are no implicit learning costs and, therefore, genetic assimilation does not take place. The cost of learning provides a selection pressure for evolution to genetically assimilate the acquired characteristic. For learnt characteristics to be genetically assimilated under posthumously assessed learning conditions, it would be necessary to impose an explicit cost of learning.

The second criterion for genetic assimilation is that the genotypic space and phenotypic space must have the property of neighbourhood correlation. In general, genotypic space and phenotypic spaces look very different. Assuming that for every phenotype there is a genotype that codes for it and that evolution and learning are working to maximize the same variable (i.e. fitness), each possible individual will occupy a point in both spaces. Neighbourhood correlation means that if two individuals have phenotypes that are close together in phenotype space then this implies that their genotypes are close together in genotype space. Genetic assimilation of a learned phenotypic trait takes place when there exists a path of increasing fitness between the genotype that encodes for the original innate trait and the genotype that encodes for the new trait in genotypic space. The path's existence depends on the genetic operators that evolution

Gray coded variables. Since Gray codes are widely used in genetic algorithms already, this poses little difficulty.

Returning to the debate talked about in section 1, it can be seen that the experiments presented here support Maynard Smith's assertion that environmental change uncovers differences in genotypes that can then bring about selection pressure for those individuals who are more able to adapt during their lifetime. That is, even if the whole population is able to learn the behaviour, those that are closer to that behaviour because of the variation brought about by environmental change, will incur less learning cost and therefore be selected for. Thus, the population moves towards the new behaviour and it becomes genetically specified. The unconverging and then reconverging of the population in figures 4c and 6c are a demonstration of this.

In conclusion, the cost of learning provides selection pressure for the genetic assimilation of acquired characteristics; neighbourhood correlation provides a path through which it can take place.