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 speed the evolutionary process as traits that are initially acquired become genetically specified in later generations. This paper presents conditions under which this genetic assimilation *can take place. As well as the benefits that lifetime adaptation can give a population, there may be a cost to be paid for that adaptive ability. It is the evolutionary trade-off between these costs and benefits that provides the selection pressure for acquired traits to become genetically specified. It is also noted that genotypic space, in which evolution operates, and phenotypic space, on which adaptive processes (such as learning) operate, are, in general, of a fundamentally different nature. For an acquired characteristic to become genetically specified, then these spaces must have the property of* neighbourhood correlation *which means that a small distance between two individuals in phenotypic space implies that there is a small distance between the same two individuals in genotypic space.*

KEYWORDS: Baldwin Effect, Genetic assimilation, Learning cost, Neighbourhood correlation.

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 comes to work on a population of learning individuals:

 Submitted to the special edition of Evolutionary Computation, Evolution, Learning, and Instinct: 100 Years of the Baldwin Effect.

[1] The classic example of this is the existence of callosities on the rumps of ostrich foetuses (Waddington, 1942), (Maynard Smith, 1993)

> '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.

2 Evolutionary Framework

This section describes the evolutionary framework in which this work sits. It is necessary since there are several ways of combining learning with evolution in artificial systems that differ in details that are significant to the arguments presented. In this paper, no distinction is made between developmental adaptations, acquired characteristics and lifetime learning. Since it deals with the artificial and abstract, all these terms come under the blanket term lifetime adaptation and will be used interchangeably, with an arbitrary preference for the term 'learning'.

We start by stating a few assumptions:

1. This paper assumes that the individuals in the population are learning the same task on which evolution is operating. That is, learning operates on the same fitness measure that is used to provide selection pressure for the evolutionary process. In artificial systems this seems a sensible thing to do since the desired product of the evolutionary process is an individual (or population of individuals) that performs a specific task. It may be that the learning task is a subset of the entire behaviour of an individual, which is more often the case in biological evolution. However, since it contributes to the overall fitness of an individual, then the arguments presented here are still valid.

2. The assumption is also made that for every phenotype, there is a genotype that codes for it. That is, if an individual is produced from a genotype which is then modified by learning, there exists a genotype that codes for that modified phenotype. This assumption is important for the arguments presented in section 3.3 but has nothing to do with a Lamarckian inheritance.

3. Also, it is assumed that the existence of learning or adaptive traits in the population is under genetic control; or, alternatively, there could exist the possibility of the learning mechanism making no changes to the phenotype if the desired behaviour is being exhibited. The reason for this is to provide a mechanism by which the learned traits can become assimilated, either by being selected against or through redundancy. In artificial systems this could be done by simply reducing the learning rates of a neural network to zero under genetic control or through the provision of a learning scheme that does not change the phenotype once the desired behaviour is reached, e.g. an error-minimisation based scheme.

Most of the artificial evolution literature that deals with the Baldwin effect is concerned with the increase in performance that the combination of learning and evolution provides, e.g. (Hinton & Nowlan, 1987), (Whitley, Scott Gordon, & Mathias, 1994). It is usual, therefore, to start an evolutionary run from a population of individuals randomly distributed across genotype space as is the norm in conventional genetic algorithm practice. Here, however, it is the mechanisms of genetic assimilation that are under investigation. A simulation of environmental change is required and so the population starts off converged around a random point in genotype space. Evolution then acts to move the population as a whole through genotype space towards an optimum.

One way that researchers have combined learning with evolution is to evolve the parameters of a supervised learning scheme. The fitness of each individual is evaluated *after* the learning process by applying a set of test data and scoring the individual on how well it performs. An example of this involves the evolution of back-propagation neural networks as in (Chalmers, 1990). This scheme has the property that it is the result of the learning process that is assessed for fitness. That is, its fitness is awarded after the individual's 'learning lifetime', or *posthumously*. In contrast, another method is unsupervised learning. The individual learns and adapts during its lifetime and is scored continually throughout this process. Its final fitness score is then the accumulation of how well it performed the desired task from the moment it was 'born' until the end of its lifetime trial. There is no separation of the individual's lifetime into training and testing phases. The

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.

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. If we assume that the benefits outweigh the costs, these pressures can eventually lead to genetic assimilation. A population is pictured as a cloud of individuals converged around a point in genotype space. At times of evolutionary change, learning individuals that have the ability to reach an area of increased fitness would have the selective advantage over others in the population who could not. Thus the benefits of learning are exploited as these learning individuals take over the population. Once the population is full of individuals that can learn, then the selective pressure turns to reducing the cost of learning whilst maintaining the improved behaviour. The advantage then switches to those individuals who are genetically closer to the improved behaviour since they incur less learning cost in exhibiting that behaviour, subject to conditions that will be discussed later. Thus we would expect there to be a high degree of learning at times of environmental change which would then be reduced by the switch in selection pressure as evolution progresses. The learning is replaced by genetic specification, i.e. there is genetic assimilation.

Putting this more formally, we can say that a genotype, i, has an innate behaviour, I, (and associated innate fitness, $f(I)$) that is then modified by learning to behaviour J, then we can say that the fitness of i, $F(i)$, after learning is:

$$F(i) = f(I) + B(I, J) - C(I, J) \tag{1}$$

where $B(I, J)$ is the benefit received by i in changing its behaviour from I to J, and $C(I, J)$ is

the total learning cost incurred by this change. Thus, for learning to be useful then $B(I, J)$ must be positive and greater than $C(I, J)$, otherwise innate behaviour I would be selected for.

If there exists a genotype, j, that can express behaviour J innately (which we are assuming there is (section 2)) then $B(I, J)$ can be expressed as $f(J) - f(I)$, the difference between the two innate fitnesses, and equation 1 becomes:

$$F(i) = f(J) - C(I, J) \tag{2}$$

Consider a population converged around i whose members can all learn behaviour J. Genetic assimilation takes place when the population moves from around i to around j. For this to happen, the population must be able to move along a path between i and j in genotype space for which the associated phenotypes have an increasing fitness score as i gets closer to j. Assuming that as the population approaches j, less learning of the behaviour J takes place, then the increasing fitness scores will be due to the reduction in the cost of learning, i.e. the $C(I, J)$ term in equation 2. Also, the movement through genotype space is due to selection pressure for those individuals that are closer to j and so incur less cost of learning. Note, the path between i and j is not necessarily a straight one or one made up of consecutive points but must exist with respect to the genetic operators being used by evolution. This point will be expanded upon in section 4.

4 Neighbourhood Correlation

Many researchers talk about the idea that the combination of learning and evolution allows for a localised search of genotype space. In most experiments that have been presented this is in fact the case (Hinton & Nowlan, 1987). However, this idea does not generalize to all evolutionary circumstances. It is true that when an individual learns it does search a subset of possible individuals. However, whether or not this subset corresponds to genotypes that are localised around that individual is totally dependant on the genotype to phenotype mapping and the learning rule.

The point that needs to be kept in mind is that the spaces on which learning and evolution operate are completely different. For example, in the case of using a genetic algorithm to evolve a learning neural network, using a binary genotype of length N, the genetic operators work on an N-dimensional hypercube whereas the learning rule operates on a surface in a continuous space of different dimensionality. When learning and evolution are trying to maximise the same quantity (i.e. fitness) then the values in the two spaces are the same for each individual but the relative positions of the individuals in those spaces may differ considerably.

For example, in figure 1 genotypes a, b, c and d code for phenotypes A, B, C and D respectively. Let's consider the relationships between a (and A) and the other three individuals. a is close to b and A is close to B; a is close to c but A is some distance from C; a is some distance from d but A is close to D. We can say that there is a correlation between the distances in the two spaces between a and b but not between a and the others.

This brings us to an important idea. If all the points near a point, I, in phenotype space have corresponding points in genotype space that are near i and vice versa, then we can say the spaces have the property of *neighbourhood correlation* or have *correlated neighbourhoods*. If all the points that are near I in phenotype space have corresponding genotypes that are some distance from i and vice versa, then the spaces have *uncorrelated neighbourhoods*. The implications of this idea tie in with what was said in section 3.3 about the movement of populations through genotype space during genetic assimilation. If A is capable of learning B then genetic assimilation of B is able to occur since a is close to b and there exists a path of increasing fitness between them in both spaces. However, A could learn D in which case genetic assimilation is unlikely to occur. This is because the path that the genotypes have to take, to get to d, goes through c (for the sake

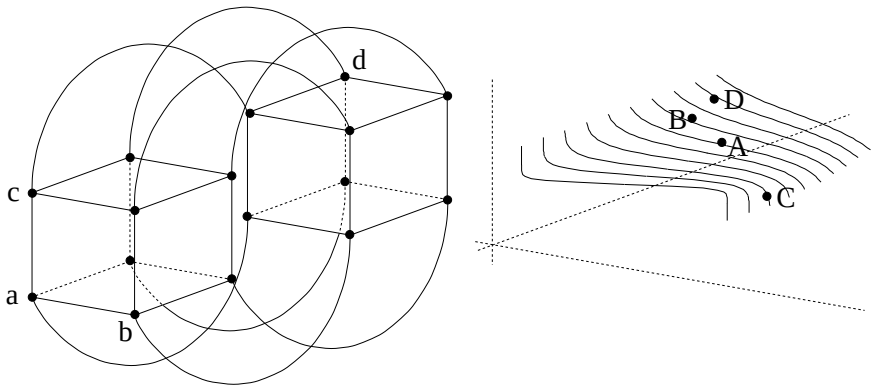

Figure 1: *A four dimensional hypercube and a surface in three dimensional space representing genotypic and phenotype space, respectively.*

of the argument). Since C is a less fit phenotype than A (e.g. by virtue of incurring a greater learning cost), A will be selected for in preference to C. A may still be modified to D by learning if it is evolutionarily advantageous to do so but D will not become genetically assimilated. This implies that if an individual's phenotype is modified to one corresponding to a genotype that is a large distance away in genotype space then learning can be detrimental to evolution. Thus, for the Baldwin Effect to work and genetic assimilation to take place, the genotypic and phenotypic spaces need to have the property of neighbourhood correlation.

In the above discussion, the idea of distance in the two spaces has been used fairly loosely. In genotype space the distance between two individuals depends on the genetic operators. For example, if the genetic operator used on the space in figure 1 consists of one and only one bit flip per breeding operation, then a is closer to b than c is. If, however, there were exactly two bit flips each breeding operation then c is close to b whilst a is an infinite distance away.

The crossover and mutation operators that are usually used in genetic algorithms allow greater freedom of movement around the genotype space than this. It is possible to get any string from any other. However, because this is very improbable, there is still some sense of genotypes being close together or far apart with respect to these operators.

Similarly, distance in phenotype space is defined in terms of the way the learning rule operates on the phenotypic variables. There may be areas that, although they have similar values to their variables, are not reachable from each other by the use of a particular learning rule.

It was mentioned in section 3.3 that a path of phenotypes with increasing fitness should exist between two points in genotype space for genetic assimilation to take place. The reality is looser than this statement implies since we are talking about the movement of a population under the crossover and mutation operators. It is the centre of gravity of the population that will move along an increasingly fit path rather than any specific individual.

5 Application of the Cost of Learning Theory to the Hinton and Nowlan Model

In this section the ideas discussed above are applied to the model of learning and evolution presented by Hinton and Nowlan (1987) who first showed the effectiveness of the combination of learning and evolution in a computational framework. Their work will be referred to as HN.

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.

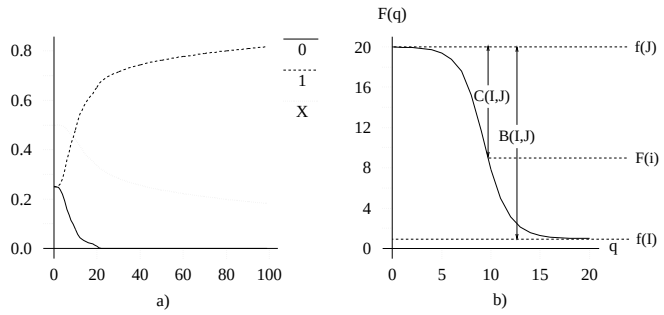

Figure 2: *a) Graph showing the relative frequencies of the* **0**, **1** *and* **?** *alleles from a replication of Hinton and Nowlan's experiment (average of 30 runs). b) Reproduction of Harvey's expected fitness graph. $F(q)$ is the expected fitness of a genotype with q* **?***s.*

Belew (1989) analyzes HN in terms of four distinct evolutionary phases. During his phase 2 he hints at the idea of learning costs:

> 'Phase 2 continues the same basic trend, but begins to also apply pressure against guessed alleles (**?**) in favor of genetically specified correct ones (**1**). In terms of the fitness function an individual is always better off not having to guess, and so there is a constant selective pressure against **?**s.'

If we look at Harvey's graph of expected fitness (Harvey, 1993), reproduced in figure 2b, then the learning costs and benefits become more explicit. In the notation of equations 1 and 2, a genotype consisting of all **1**s that is therefore at the global optimum is referred to as j and the single fitness peak, behaviour J. Its overall fitness, $F(j)$, and its innate fitness $f(J)$ are both equal to 20. An individual, i, whose genotype is anything other than j will receive an innate fitness, $f(i)$, of 1. Individual i might be able to reach J during its lifetime by appropriately setting its **?**s, and therefore receives the benefit of learning, $B(I, J)$, but will receive a rather smaller fitness score than $f(j)$. This difference in fitness scores can be considered the cost attributed to the process of learning and is shown as $C(I, J)$ in figure 2b. The shape of the graph was determined implicitly by HN's fitness function and their use of fitness proportional selection. The cost of learning can

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 $C(I, J)$, for these proportions, are roughly nineteen and eleven, giving a learning individual eight times as many offspring as a non-learning one. The proportion of **?**s in the population then starts to fall. An individual that has only nine **?**s will still receive a benefit of nineteen but will only pay a cost of about eight whilst still maintaining the optimal behaviour. The advantage then switches to those individuals who have more fixed (correct) genes, (**1**s) since they incur less learning cost and learning is selected against.

However, the proportion of **?**s in the population never falls to zero; there is only partial genetic assimilation. An explanation for this was put forward by Harvey (1993). He explains that the 'puzzle of the persistent question marks' is a result of genetic drift. That is, the selection pressure to remove the last **?**s, represented by the slope of the graph when q is close to zero, is not great enough to overcome the forces of genetic drift. Putting this in terms of the cost of learning, we can say that the number of **?**s in the population does not fall to zero since the cost of learning the last few is so small that there is not enough selective pressure to remove them against the forces of genetic drift. That is, genetic drift is the mechanism by which the **?**s are kept in the population and its effects are great enough to overcome the selection pressure created by the costs of having them there.

It is worth looking at which how well the HN model fits into the evolutionary framework and assumptions that have been made in section 2 so as to gauge the validity of the above analysis. The model is both learning and evolving to maximise the number of set switches (**1**s) in the phenotype and, therefore, evolution and learning are working on the same fitness measure. There is a genotype that encodes for every phenotype and, through the control of **?**s in the genotype, there is a genetic mechanism for selecting for and against learning. The evolutionary runs are started from a random population and not from a converged one. This is because HN were trying to show the utility of learning in an evolutionary context and for their purposes a random population was sufficient. At first, it appears that the learning scheme is posthumously assessed – an individual is awarded a fitness score at the end of its learning trials on which evolution acts. However, through their use of a fitness function that is a function of the number of trials taken by an individual to find the global optimum, the learning scheme is implicitly continually assessed. An individual is penalised for the proportion of time it spends not at the global optimum.

We now look at the neighbourhood relationships, as defined in section 4, that the HN model possesses. If we take a phenotype, P, and change it slightly to phenotype, Q, by flipping the setting of one of its switches, then there will always be a genotype, q, that codes for Q, that is only one allele change away from the genotype, p, that codes for P. That is, if we have two close phenotypes we can say that there exists two close genotypes that code for them. Note, the existence of **?** alleles means that there may be many qs that code for Q or that p and q may be the same, but the important point is that there is at least one q that is close to p and that a path

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 Genotype to Phenotype Mappings

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 1 shows mapping 2 in a Karnaugh map (Horowitz & Hill, 1989), which gives a good indication of the neighbourhood relationships. The concatenation of the row and column headings gives the four bit genotype and the phenotype it codes for can be read directly from the corresponding cell. Karnaugh maps have the property that if you move a distance of one cell in any direction (N, S, E, W), then the key that addresses the cells, in this case the genotype, only changes by one bit. Thus, we can see that if a single bit is flipped in genotype space then there is a movement of three bits (i.e. $N - 1$) in phenotypic space. That is, all the neighbours of a particular genotype, p, in genotypic space code for phenotypes that are a Hamming distance of three away from P.

6.2 Genetic Operator

A genetic operator was required that maintained the integrity of the neighbourhood relationships in the genotype to phenotype mappings described above in section 6.1. To this end, asexual reproduction was used with a single-point mutation operator. That is, an individual is selected for breeding and copied to the child. One randomly chosen bit may be flipped in the genotype of the child with a given probability. It is only possible, therefore, to have children that are either genetically identical to their parent or differing by only one bit. This preserves the neighbourhood relationships of the genotype to phenotype mappings.

6.3 Learning

As with the genetic operator, a learning algorithm was required that preserved the neighbourhood relations of the genotype to phenotype mappings. A one bit steepest ascent algorithm was chosen. The innate fitness of the phenotype is calculated. The fitness of all the phenotypes that are a Hamming distance of one away from the innate phenotype are then calculated and the highest score achieved by either these learnt phenotypes or the innate phenotype is take as the fitness score for that individual. If a learnt phenotype achieves a higher score than the innate one then the individual is logged as having used learning. If the individual cannot better its score through learning then it is logged as having not used learning. Note, since the goal is to maximise the number of **1**s in the phenotype and each individual is only allowed to flip one phenotypic bit at a time, then it is only possible for an individual to improve its score by one during learning.

We can look at the effect this movement around phenotypic space has on the genotype needed to code for a modified phenotype. Under mapping 1, a phenotype, such as `0111`, that can learn the target string, `1111`, through a one bit hill-climb, only has to have its genotype changed from `0111` to `1111` for the target to become assimilated. Table 2 shows what effect movement around phenotypic space has on the genotype needed to code for a modified phenotype under mapping 2.

If we have the phenotype, `0111`, that is able to learn the target string, `1111`, through a one bit hill-climb, it can be seen that the genotype has to change from `1000` to `1111` for that learnt behaviour to become assimilated. Under the one bit mutation operator this constitutes a large movement in genotype space for a small movement in phenotype space. The neighbourhoods under this mapping are maximally uncorrelated.

The cost of learning is simulated by subtracting a fixed penalty from the fitnesses of those individuals who have been logged as learning. This differentiates between schemes where learning is posthumously or continually assessed, as defined in section 2.

6.4 Experiment Overview

This section puts the above descriptions of the various components together to give an overall view of the model. The model is designed to operate in six different modes corresponding to combinations of evolution with or without learning, on mappings with and without neighbourhood

$ph \rightarrow g$	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 2: *Phenotype to genotype mapping 2 showing how movements around phenotypic space due to learning correspond to changes in the genotypes that code for the changed phenotype. The row defines the first two bits of the genotype and the column the second two bits, giving the phenotype in the corresponding cell. Note, this is the same as the genotype to phenotype mapping shown in table 1 except that movements are now in* phenotypic *space showing the corresponding movements in* genotypic *space.*

correlation and with or without a cost for learning (if learning is switched on). The results that are presented in section 7 are from experiments in which the genotype length, N, is set to 20. The genotype/phenotype mappings are calculated in exactly the same way as for $N = 4$ as described in section 6.1.

A converged population of size 100 is generated by creating a random genotype of 20 bits long and copying it to all the individuals. All the bits in all the genotypes are then mutated with a fixed low probability of 0.05. This gives a population converged around a random point in genotype space with an average Hamming distance from that point of 1. This was done to simulate the situation where a population has converged on one trait and then, due to environmental change, it has become favourable to exhibit a different trait. The genotypes are coded into phenotypes using one of the mappings described in section 6.1 and the fitnesses of the phenotypes evaluated. Depending on which mode of operation the model is in, the fitness of a phenotype is evaluated in different ways. When operating evolution only, the fitness of an individual is simply the number of **1**s in its phenotype, i.e.

$$F(i) = N - H(I, T) \tag{3}$$

where $F(i)$ is the fitness awarded to individual i, N is the genotype length and $H(I, T)$ is the Hamming distance from i's phenotype, I, to the target phenotype, T, which consists of all **1**s.

When the population is able to learn, the fitness of an individual is calculated using:

$$F(i) = N - H(I', T) \tag{4}$$

where I' is the phenotype that is closest to the target out of I and all the other phenotypes, I_j, that can be reached by I through the learning process described in section 6.3.

When there is a cost to learning, the fitness of an individual is calculated using:

$$F(i) = N - H(I', T) - c \tag{5}$$

where c is the cost of learning, set at 0 if $H(I', T) = H(I, T)$, 0.5 otherwise, i.e. the individual only gets penalised for learning if it needs to use it to increase its fitness.

Once the fitnesses of all the individuals in the population have been evaluated, they are bred asexually to form the next generation using simple linear rank based selection and the one point mutation operator described in section 6.2. The genotypes are picked to be mutated with a probability of 0.1 such that evolution operates on a slower time scale than learning. This genetic algorithm was run for 100 generations and various results were obtained that are presented in section 7.

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).

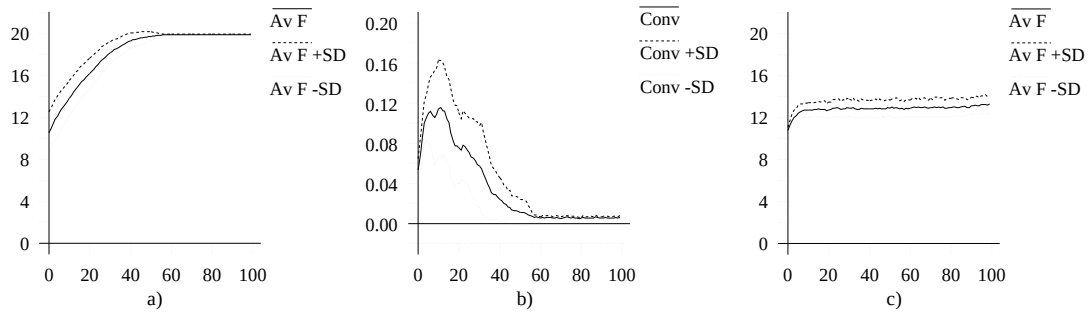

Figure 3: *Evolution without learning. a) Average fitness with correlated neighbourhoods. b) Convergence with correlated neighbourhoods. c) Average fitness with uncorrelated neighbourhoods.*

However, evolution has difficulty on the mapping with uncorrelated neighbourhood relations (figure 3c). This is to be expected since the task is hard with respect to the genetic operator.

7.2 Cost-Free Learning

Figure 4 shows the effect of the introduction of cost-free learning into a population evolving on landscapes with neighbourhood correlation. This set of results corresponds to the situation where learning is posthumously assessed, i.e. fitness is evaluated using equation 4 such that there is no learning cost. Comparing figure 4a with figure 3a, the first point to notice is that the population reaches the global optimum faster (i.e. in less generations) than evolution alone on the same landscape. Learning is guiding evolution to the peak through the Baldwin Effect. The second point to notice is that the population is less converged than in figure 3b. This is because not only genotypes that have a **1** at every locus receive the maximum fitness score but also those that are

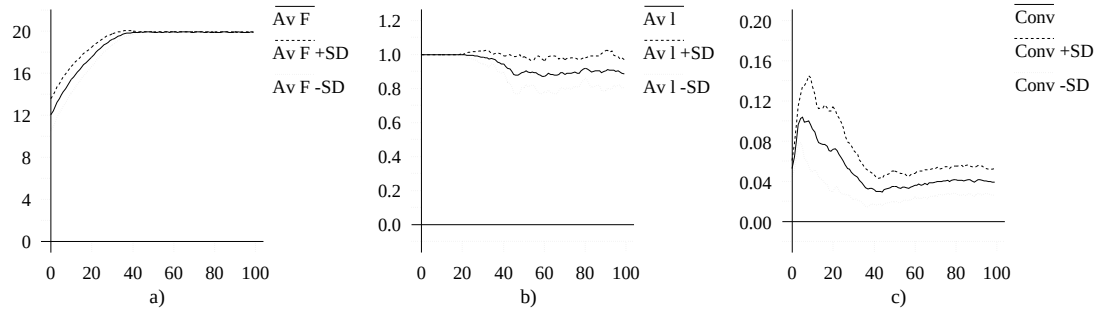

Figure 4: *Evolution and learning with correlated neighbourhoods but no cost of learning. a) Average fitness, b) proportion of individuals using learning, c) convergence.*

one bit flip away. This is because the latter are able to achieve the global optimum in phenotype space by learning. In figure 3a, b, evolution is working towards an optimum in genotype space that consists of only one possible genotype, whereas, the introduction of learning has expanded this optimum to one plus the number of genotypes that are a Hamming distance of one away in genotype space. Thus evolution can support a less converged population since there is no selection pressure between having N **1**s and $(N-1)$ **1**s. In the case of $N = 20$, there are now 21 genotypes that can achieve the highest fitness score. A good account of the expansion of optima that is caused by the interaction of learning and evolution was presented by Whitley et al. (1994).

The third point to note is that the proportion of learning individuals falls slightly once the population has converged around the optimum (figure 4b). This to be expected since some proportion of the genotypes that obtain the optimal fitness consist of all **1**s and therefore do not learn. At first sight it seems that the proportion of learning individuals falls too low. That is, there is only one possible genotype that can achieve maximum score without learning whilst there are twenty possible genotypes that can learn the optimal phenotype. One might, therefore, expect the graph in figure 4b to fall to around 0.95. However, this does not take into account those individuals that are mutated from a Hamming distance of one away from all **1**s to a Hamming distance of two away. To a first approximation, the proportion of the population that we would expect to have genotypes consisting of all **1**s can be calculated thus:

Let A, B and C be the sets of genotypes that are a Hamming distance of 0, 1 and $\geq$2 away from the genotype consisting of all **1**s, respectively. We make the assumptions that the population is converged around A and B, and, due to the linear rank selection, a negligible number of individuals from C get selected for breeding. The expected proportion of genotypes selected from A and B is a and b, respectively. After mutation, (with $N = 20$ and mutation rate $= 0.1$ (section 6.4)), these proportions will have changed to:

$$a' = a - \frac{a}{10} + \frac{b}{10} \cdot \frac{1}{20} \tag{6}$$

$$b' = b + \frac{a}{10} - \frac{b}{10} \cdot \frac{1}{20} - \frac{b}{10} \cdot \frac{19}{20} \tag{7}$$

Now, since we only select from A and B, whose members all get the same score, at steady state we can say that the expected proportions are $a'/(a'+b')$ and $b'/(a'+b')$, i.e.:

$$a = \frac{0.9a + b/200}{a + 0.905b} \tag{8}$$

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

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 selection process which corresponds to the experimental results. There is no genetic assimilation.

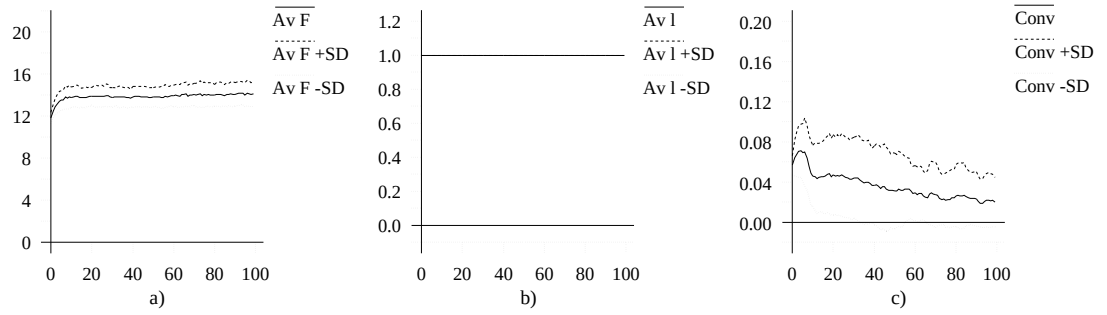

Figure 5: *Evolution and learning with uncorrelated neighbourhoods but no cost of learning. a) Average fitness, b) proportion of individuals using learning, c) convergence.*

A population of learning individuals with no cost of learning, evolving using the mapping with uncorrelated neighbourhoods, only does slightly better than evolution alone on the same landscapes. The improvement in average fitness in figure 5a over figure 3c is due to all the individuals in the population achieving a phenotype that is a Hamming distance of one closer to the target that their genotypes, as indicated by figure 5b. However, the learning behaviour is unable to guide evolution to areas of increased fitness due to the genotype and phenotype spaces having uncorrelated neighbourhoods.

7.3 Learning with Cost

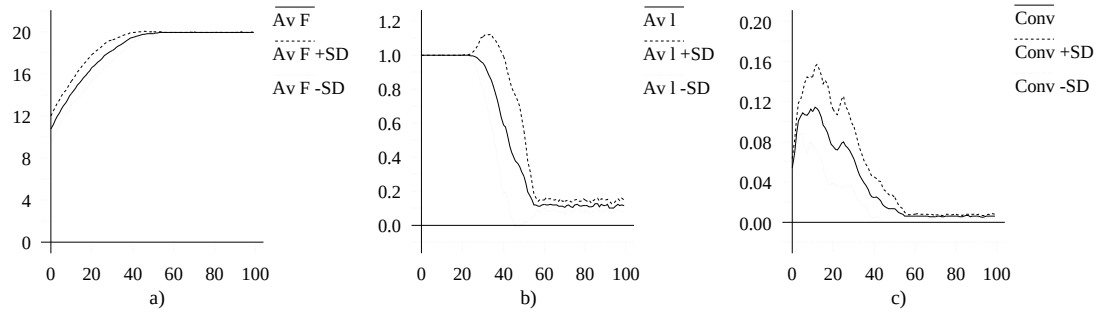

Figure 6: *Evolution and learning with correlated neighbourhoods and a cost for learning. a) Average fitness, b) proportion of individuals using learning, c) convergence.*

Figure 6 shows the case when evolution and learning are operating in spaces with correlated neighbourhoods where learning is penalised with a cost. This situation corresponds to the circumstance when learning is continually assessed and the fitness of each individual is evaluated

using equation 5. The average fitness of the population (figure 6a) reaches the optimal value faster than evolution alone but slower than when there is no cost of learning. We can see why by looking at figure 6b. This shows the proportion of learning individuals in the population. In the early stages of evolution, the population is full of learning individuals who are reaping the benefits of learning but are also paying the cost of learning which is reducing their fitness score by 0.5. As the population reaches on the global optimum, individuals that achieve the optimum score without learning do not pay that cost. Noticeably the proportion of individuals that use learning falls dramatically to 0.1 when the global optimum is reached. This indicates that 90% of the population have genotypes that consist of all **1**s compared with approximately 15% when there was no cost for learning. The cost of learning has provided a selection pressure between those individuals whose genotypes consist of all **1**s and those that are a Hamming distance of one away. The global optimum in genotype space is once again solely at **111...1** as it was for evolution alone and does not include those genotypes that are a Hamming distance of one away as in the case of cost-free learning. The 10% of the population that are not at the global optimum are those whose parent consisted of all **1**s (in the previous generation) but have been mutated by the genetic operator and are now a Hamming distance of one away. The rank selection ensures that a negligible number of these individuals get passed into the next generation. This interpretation of the results in determining the position of the population in genotype space is supported by the convergence data shown in figure 6c. When the average fitness levels off at the optimal score, the convergence value falls to approximately 0.005. Assuming that the consensus genotype consists of all **1**s, there will be 0.1 of the population that differ by a Hamming distance of one, therefore having 1/20th of their genotype different from the consensus. Thus, over the population the deviation from the consensus is $0.1 \times 1/20 = 0.005$.

The convergence data, together with the dramatic fall in the proportion of learning individuals as the population reaches the global optimum, indicate that the learning that is still in the population in figure 4b (cost-free learning) has been genetically assimilated with the introduction of a cost of learning.

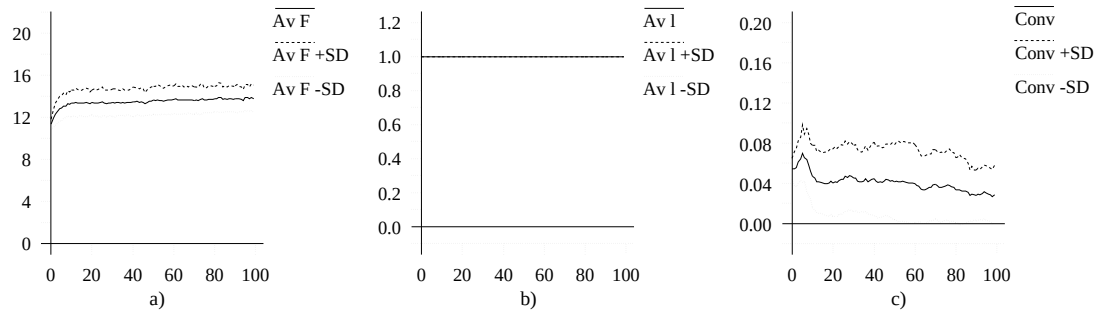

Figure 7: *Evolution and learning with uncorrelated neighbourhoods and a cost for learning. a) Average fitness, b) proportion of individuals using learning, c) convergence.*

Figure 7 shows the combination of learning and evolution on landscapes with uncorrelated neighbourhoods where learning has been penalised with a cost. The introduction of a learning cost has had little effect on the evolution of the population in this case, other than giving an average fitness which is slightly lower in figure 7a than figure 5a. This is due to the fact that all the individuals in the population are using learning but their fitnesses are being reduced by the cost of learning.

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 has available. The arguments put forward in this paper present a mechanism by which the fitness can increase as the population moves closer to genetically specifying the learned trait, i.e. through the reduction of the cost attributed to learning (assuming the closer an individual is to having a trait genetically specified, the less learning it has to do).

The experimental results support this view. Evolution with learning achieved the global solution on landscapes that have the property of neighbourhood correlation in less generations than evolution alone, even when learning was penalised with a cost i.e, the benefits of learning where being exploited. Comparing figures 4 and 6 we can see that when there is a cost to learning, the population converges on the single genotype that can achieve the optimal score rather than containing individuals that can achieve the score through learning. The level of learning that is sustained when there is no cost of learning has been genetically assimilated.

Evolution with and without learning failed to find the optimum phenotype on the landscapes with uncorrelated neighbourhood relations even though learning did improve the fitness scores of individuals. There are no grounds to assume that the inclusion of learning in an evolutionary framework will automatically guide evolution to the optimum, even if that learning algorithm is particularly suited to the phenotypic landscape, as in this case. The nature of the genotype to phenotype mapping is a major factor in determining the suitability of the combination of learning and evolution.

In reality, the constraints placed on the combination of learning and evolution by the neighbourhood correlation idea are less than the results presented in section 7 suggest. Most artificial evolution systems use crossover and mutation for their genetic operators which allow a greater freedom of movement around genotype space than the genetic operator applied here. An individual's neighbourhood in genetic space is, therefore, larger. Also, a degree of neighbourhood correlation can easily be added to a given genotype to phenotype mapping through the use of

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.

Acknowledgements

Many thanks to Inman Harvey, Adrian Thompson, and Dave Cliff for their help and contributions.

Reference

Ackley, D., & Littman, M. (1991). Interaction between learning and evolution. In Langton, C., Taylor, C., Farmer, J., & Rasmussen, S. (Eds.), *Artificial Life II*, pp. 487–509.

Baldwin, J. (1896). A new factor in evolution. *The American Naturalist*, pp. 441–451.

Belew, R. K. (1989). When both individuals and populations search: Adding simple learning to the genetic algorithm. In Schaffer, J. D. (Ed.), *Proceedings of the Third International Conference on Genetic Algorithms*. Morgan Kauffmann.

Chalmers, D. J. (1990). The evolution of learning: An experiment in genetic connectionism. In Touretzky, D. S., Elman, J. L., Sejnowski, T. J., & Hinton, G. E. (Eds.), *Proceedings of the 1990 Connectionist Models Summer School*. Morgan Kaufmann.

Harvey, I. (1993). The puzzle of the persistent question marks: A case for genetic drift. In Forrest, S. (Ed.), *Genetic Algorithms: Proceedings of the Fifth International Conference (GA93)*. Morgan Kaufmann, San Mateo, CA.

Hinton, G. E., & Nowlan, S. J. (1987). How learning can guide evolution. *Complex Systems, 1*, 495 – 502.

Horowitz, P., & Hill, W. (1989). *The Art of Electronics* (2nd edition). Cambridge University Press.

Maynard Smith, J. (1993). *The Theory of Evolution* (3rd edition). Cambridge University Press.

Simpson, G. (1952). The Baldwin Effect. *Evolution, 7*, 110–117.

Waddington, C. (1942). Canalization of development and the inheritance of acquired characters. *Nature*, pp. 563–565.

Waddington, C. (1953a). The "Baldwin Effect," "Genetic Assimilation" and "Homeostasis". *Evolution, 7*, 386–387. In 'Notes and Comments' section.

Waddington, C. (1953b). Genetic assimilation of an acquired character. *Evolution*, *7*, 118–126.

Waddington, C. (1956). Genetic assimilation of the *bithorax* phenotype. *Evolution*, *10*, 1–13.

Whitley, D., Scott Gordon, V., & Mathias, K. (1994). Lamarckian evolution, the Baldwin Effect and functional optimization. In Davidor, Y., Schwefel, H. P., & Manner, R. (Eds.), *Parallel Problem Solving from Nature-PPSN III*, pp. 6–15. Springer-Verlag.