Is There Another New Factor in Evolution?

Inman Harvey

School of Cognitive and Computing Sciences

University of Sussex

Brighton BN1 9QH, UK

email: inmanh@cogs.susx.ac.uk

tel: +44 1273 678754

fax: +44 1273 671320

Abstract

For 100 years it has been recognised that interactions between learning and evolution, such as the Baldwin effect (Baldwin, 1896), can be subtle and often counter-intuitive. Recently a new effect has been discussed: it is suggested that evolutionary progress towards one specific goal may be assisted by lifetime learning on a different task which may or may not be 'uncorrelated' (Parisi, Nolfi, & Cecconi, 1992). Here the phenomenon is reproduced in a simple scenario where the tasks are indeed uncorrelated — 'Another New Factor' does indeed exist. The effect is then explained as being due to recovery from weight-perturbations, caused by mutation, in a neural network. It is a special case of a recently discovered relearning effect (Harvey & Stone, 1996), the spontaneous recovery of perturbed associations by learning uncorrelated tasks.

1   Introduction

In 1896 Baldwin proposed 'A New Factor in Evolution', and since then it has often been observed that interactions between learning and evolution can be subtle and counter-intuitive. In recent papers (Parisi et al., 1992; Nolfi, Elman, & Parisi, 1994) another phenomenon was reported, which I shall here call Another New Factor (ANF; I shall be claiming that it is indeed different from the Baldwin effect). This work used simulations of populations of artificial neural networks (NNs) that evolve (to get fitter at one task) at the population level and may also learn (a different task) at the individual level. The ANF result stated was that average fitness at the evolutionary task is improved when lifetime learning at the different task is introduced.

The extent to which the lifetime task 'differs' from the evolutionary task is a matter of some significance. In the later paper (Nolfi et al., 1994) the two tasks are such that it is open to question whether they are completely unrelated; however in the earlier paper (Parisi et al., 1992) separate experiments are also briefly reported (p. 211) where the second task relies on 'randomly generated teaching inputs'. I will distinguish between the 'weak' sense of ANF, in which the relevant tasks are merely different; and a 'strong' sense of ANF in which the relevant tasks can be *uncorrelated* as defined below in section 1.1.

In this paper the emphasis is on the strong sense of ANF: are there interactions between evolution based on fitness at one task, and lifetime learning on an uncorrelated task, in particular such that the latter improves performance on the former? (Nolfi et al., 1994) does not directly address this strong ANF, though the earlier paper implicitly brings up the first reference to such a phenomenon of which I am aware. Such a phenomenon, if it exists, clearly differs from the Baldwin effect which concerns similar tasks being tackled by both evolution and lifetime learning; and which relies on 'nearness in genotype space' to a successful solution of the evolutionary task to be correlated with 'decreased learning effort required' on the lifetime learning task for genetic assimilation to be favoured[1].

I will first describe the simulations in which a weak form of ANF can be observed, drawing on (Nolfi et al., 1994). I then perform a much simpler experiment, in which both evolution and lifetime learning are reduced to movements in a multi-dimensional weight space towards target vectors, and demonstrate similar phenomena in a strong form of ANF. In this simpler situation the parallels with relearning in NNs with perturbed weights can be seen clearly. I then summarise a recently discovered relearning effect (Harvey & Stone, 1996) which can be seen to provide an explanation for ANF (in the strong sense). I conclude by summarising the differences between this explanation for ANF, and the explanations first suggested.

The answer to the question posed in the title of this paper will be: ANF does indeed occur, even in the strong form between uncorrelated tasks. But this is not a new factor unique to evolution, rather it is an example of recovery of perturbed associations by learning unrelated tasks — which may extend to

---

[1] Throughout this paper when I refer to the Baldwin effect I am focusing on the genetic assimilation aspects, where under certain conditions evolution will replace plastic mechanisms of lifetime learning with genetically specified mechanisms that achieve (partly or wholly) similar competencies.

interactions between evolution and learning only under particular conditions.

1.1 Correlation

It was stated above that for the strong form of ANF the tasks may be uncorrelated; to avoid misinterpretation this should be defined[2].

Given a set of real-valued input vectors X and a set of real-valued output vectors Y, I define a behaviour $\mathbf{B}$ as a specific mapping from input vectors to output vectors: $\mathbf{B} : X \rightarrow Y$. I assume that the behaviour $\mathbf{B}$ can be implemented in a neural network with a certain vector of weights W. Let $\mathbf{B}^*$ be the space of possible behaviours, and W^* the space of possible weights associated with them; for each weight vector W in W^*, there is a corresponding behaviour $\mathbf{B}$ in $\mathbf{B}^*$.

I define a task $\mathcal{T}$ as a reward or reinforcement function for behaviours, mapping them onto real-values[3], $\mathcal{T} : \mathbf{B}^* \rightarrow \Re$. Since weight vectors correspond to behaviours, this is equivalent to defining $\mathcal{T}$ as a function mapping weight vectors to real-values: $\mathcal{T} : W^* \rightarrow \Re$. Let $\mathcal{T}^*$ be the space of all possible tasks. Let $\mathbf{P}$ be a probability distribution defined on $\mathcal{T}^* \times \mathcal{T}^*$. If there are distributions $\mathbf{P}_1, \mathbf{P}_2$ on $\mathcal{T}^*$, from which tasks $\mathcal{T}_1, \mathcal{T}_2$ are drawn, such that $\mathbf{P}(\langle \mathcal{T}_1, \mathcal{T}_2 \rangle) = \mathbf{P}_1(\mathcal{T}_1) \times \mathbf{P}_2(\mathcal{T}_2)$, then $\mathcal{T}_1$ and $\mathcal{T}_2$ are independent. I refer to such tasks $\mathcal{T}_1$ and $\mathcal{T}_2$ as *uncorrelated* within this paper. It should be noted that this usage differs somewhat from the statistical usage of 'uncorrelated' which strictly is a property of sequences of samples.

Because of the surprising and counter-intuitive nature of the ANF effect described here, and the underlying effect originally described in (Harvey & Stone, 1996), it has sometimes been assumed that it must require *some* constraints on the tasks or weight vectors. It cannot be emphasised too strongly that this is not the case. For instance, if weight vectors W_1 and W_2 correspond to behaviours $\mathbf{B}_1$ and $\mathbf{B}_2$ that give optimal performance on, respectively, tasks $\mathcal{T}_1$ and $\mathcal{T}_2$ it is irrelevant whether W_1 and W_2 are or are not orthogonal; the possibility that $W_1 = -W_2$ need not be ruled out. Likewise even if behaviour $\mathbf{B}_2$ is chosen such that $\mathbf{B}_2(x) = -\mathbf{B}_1(x)$ for all input vectors $x \in X$, this is still no bar to the effect operating.

2 Learning and Evolution in Neural Networks — Animat Model

In (Parisi et al., 1992; Nolfi et al., 1994) a population of animats (artificial organisms) is simulated; their behaviour depends on an artificial neural network whose weights are genetically specified as real values (see Figure 1). Each animat is separately tested, for a lifetime of 5000 actions, in a grid-like environment of 10×10 cells within which 'food' is randomly scattered at intervals. Two of the NN's inputs are Sensory Inputs which encode the direction of and distance to the currently nearest food element, from

---

[2]I am deeply grateful to an anonymous referee for the precise form of this definition.

[3]One possible reinforcement function would be a penalty function using the non-negative reals which measures the difference between the current behaviour $\mathbf{B}$ of a network and the desired behaviour $\mathbf{B}'$. This difference can be based on a distance metric between current outputs $\mathbf{B}(x)$ and desired or target outputs $\mathbf{B}'(x)$ for input vectors $x \in X$. Such a penalty function should be minimised, giving a zero value only for the desired behaviour. Other classes of reward/reinforcement functions are possible.
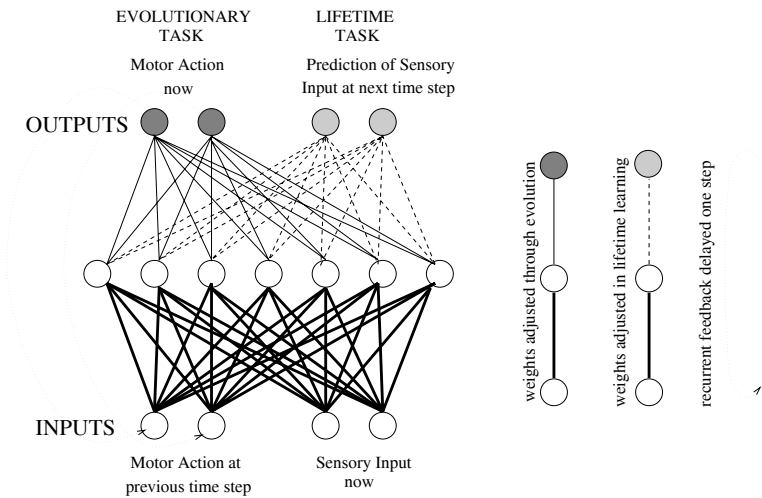
Figure 1: *The architecture used in the simulations of (Nolfi et al., 1994). The weights on the lower connections (bold lines) may be affected by both evolution and lifetime learning.*

the animat's current position. Two of the NN's outputs define the animat's next move: the analog values are thresholded to the nearest binary value, giving two bits which decide between step-forward, turn-left, turn-right, stay-still. These Motor Outputs are also passed back to a further two inputs for the next time step, giving the network an element of recurrency.

Ignoring for the time being the remaining two outputs, an Evolutionary Task can be set, defining a fitness function, which uses the fixed (genetically specified) connection weights between the 4 inputs, the 7 hidden nodes, and the 2 motor outputs. Each animat is assessed on how much food it collects during its lifetime, and on the basis of evaluating the whole population, size 100, an evolutionary algorithm (EA) produces the next generation. The best 20 are each cloned 5 times, asexually, to create 100 offspring; then the inherited weight matrices are each mutated by adding to each of 5 randomly selected weights a different number randomly taken from the interval [-1.0,1.0].

This process was repeated for 100 generations, and performance on the Evolutionary Task was seen to improve. However, a significant new factor was suggested when these non-learning trials were compared with a similar set of trials which involved lifetime learning on a second, different, Lifetime Task.

In these learning trials, in addition to the previous Evolutionary Task, the second pair of outputs (Figure 1) attempted to predict the animat's sensory inputs at the following time step. At each of the animat's 5000 cycles, this prediction was compared with the actual sensory inputs then observed, and the error in prediction used to adjust the connection weights between the 4 inputs, the 7 hidden nodes, and the 2 Prediction outputs. Backpropagation was used to change the weights, of which those between inputs and hidden nodes were simultaneously implicated in the performance on the separate Evolutionary Task.

Hence this lifetime learning influenced the relative fitnesses of different members of the population, which were assessed on performance at the Evolutionary Task. However, when the best were selected to

4

create the next generation, it was their original 'at birth' weights that were inherited, not the weights as altered by learning; there was no inheritance of acquired characteristics. It was demonstrated that there was better performance at the Evolutionary Task after this learning than in the non-learning case, under the same evolutionary selection.

This might at first sight seem unsurprising; after all, presumably there is some correlation between the tasks, within this environment, of finding food and of predicting sensory inputs of food. But success at the first task depends on appropriate outputs encoded in 2 values thresholded to 2 bits and interpreted as forward/left/right/stay, whereas for the second task appropriate outputs are 2 real values interpreted as representing distance and angle; any correlation between these will be complex. Any surmise that the effects observed could be obtained even if the two tasks were completely uncorrelated would be surprising and counter-intuitive — yet this claim was indeed made in an earlier paper (Parisi et al., 1992) p.211:

> We have run some simulations in which individual networks are taught by using randomly generated teaching inputs on the two additional output units . . . , and in this case too there is a positive influence of learning on evolution, even if the improvement is less great than in the case of prediction learning . . .
>
> It remains to be explained why learning a task such as predicting the consequences of one's own actions, which is correlated with the task for which organisms are selected, results in a larger beneficial effect on evolution.

Here I do not pursue the latter question; however I do investigate the surprising claim that the effect occurs between uncorrelated tasks, even though this claim was not repeated in later papers such as (Nolfi et al., 1994). I shall now demonstrate a similar effect in a very simple simulation: I can use two tasks which are clearly uncorrelated, yet the explanation relies on the particular way in which the EA is used.

3   The Simple Model

The EA used is as far as possible identical to that used in the Animat Model above, except that the genotypes consist of vectors of 50 real numbers, which in the initial random population are each randomly drawn from the interval [-1.0,1.0]. The Evolutionary Task $\mathcal{E}$ and the Learning Task $\mathcal{L}$ are defined in terms of minimising the distance from a genotype to a fixed target vector, $\mathbf{E}$ and $\mathbf{L}$ respectively; each component of the 50-dimensional target vectors is similarly drawn at random from [-1.0,1.0][4]. The performance of any individual at either task is given by the Euclidean distance $|\mathbf{G} - \mathbf{E}|$ or $|\mathbf{G} - \mathbf{L}|$ between the vector $\mathbf{G}$ specified by its genotype, and the target vector.

When conducting trials, selection and cloning of the best 20 from 100, and mutation by adding random perturbations drawn from [-1.0,1.0] to 5 randomly chosen weights in each individual genotype, were as above. In any one trial the population was seen to move towards the target vector $\mathbf{E}$ over 100

---

[4]In these tasks input vectors are irrelevant. Since the target vectors are chosen at random, the tasks are uncorrelated in the sense of Section 1.1.
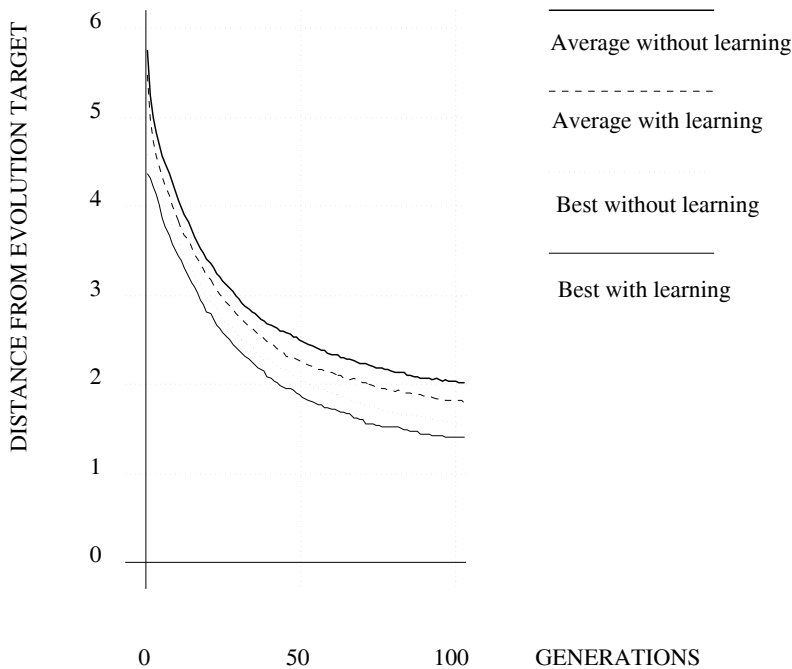
Figure 2: *Results when mutation was applied only to 80 offspring, and not to the 20 elite parents.*

generations. First a non-learning trial was performed, and then a learning trial using the identical starting population and evolutionary target vector $\mathbf{E}$. Learning was implemented using just a single application of the delta rule during an individual genotype's life: the genotype vector was moved once only towards the target learning vector $\mathbf{L}$ by a proportion $\delta = 0.1$ of its initial distance from $\mathbf{L}$: the new vector is $\mathbf{G}' = \mathbf{G} + \delta(\mathbf{L} - \mathbf{G})$. The fitness of the genotype was then assessed as $|\mathbf{G}' - \mathbf{E}|$, its new distance from $\mathbf{E}$; but those selected to pass on genetic material passed on their 'at birth', pre-learning, genetically specified values $\mathbf{G}$. This method of assessing fitness allocates zero cost to the learning process itself, in contrast to the usual models of the Baldwin effect.

To conform with the Animat Model practice, 10 sets of such non-learning and learning trials were made, with fresh vectors $\mathbf{E}$, $\mathbf{L}$, and starting population each time. In the Animat Model reproduction involves keeping the elite 20 individuals and generating 4 mutated copies of each; it was not clear from (Nolfi et al., 1994) whether the elite 20 were kept unchanged, or were also themselves mutated, so both methods were used in turn. In Figure 2 the results are shown using the first method, giving the average and best fitness at the Evolutionary Task at each of the 100 generations. As with the Animat Model, we have the ANF result that average performance of the learning trials is better than the non-learning ones. Peak performance also is improved by learning, which was not demonstrated in the original Animat Model (see Figure 4 of (Nolfi et al., 1994)).

In Figure 3 comparable results are shown for the case when all 100 are mutated; the end results are
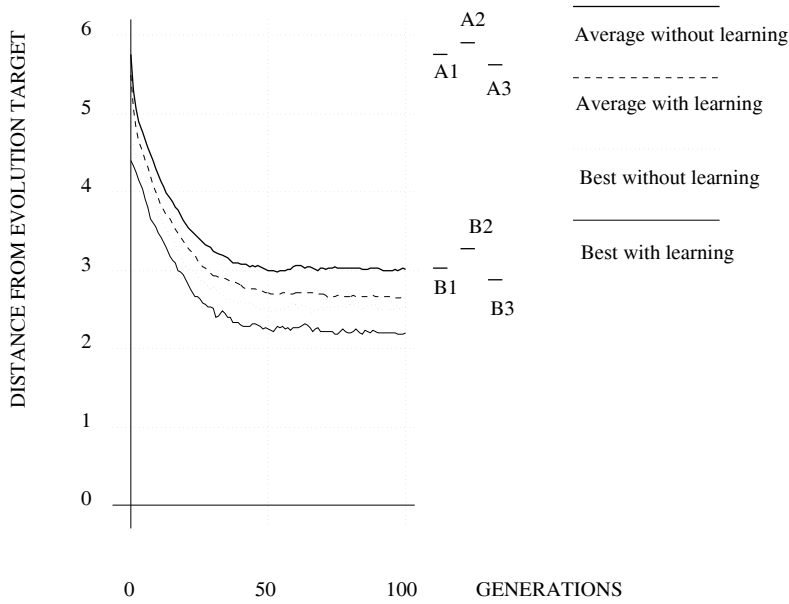
Figure 3: *Results when mutation was applied to all members of the population each generation. See text for explanation of bars A_i and B_i.*

not as good, but the comparisons between learning and non-learning trials remain. On the right side of this graph two sets of three bars are marked, using the same vertical axis as a measure of distance. **A1** is the average distance of the initial random starting population from the Evolutionary target **E**; **A2** is the average after this population has had the mutation operator applied to all members; **A3** is the average after a subsequent learning move towards **L** (in all cases this is averaged over 10 trials). **B1**, **B2** and **B3** are the equivalent figures using the population after a run (with learning) of 100 generations. It can be seen that in each case, *even in the absence of any selection and reproduction within the population*, mutation increases the average distance from **E**, and changes due to learning (towards **L**) then decrease considerably the average distance from **E**. Herein lies a clue leading to the explanation for ANF.

A further clue is given by the fact that the initial random population is broadly confined within genotype space by the initial range of random values; and then the particularly strong truncation selection method used provides an extremely strong convergence pressure, only balanced in this asexual EA by the mutation rate. In the absence of mutation such selection would result in the initial best member taking over the whole population in just 3 generations (from 1% to 5% to 25% to 100%). With mutation the population consists of similar individuals with some values (weights in the NN case) perturbed around those of the elite of the previous generation. To quote from (Nolfi et al., 1994), p. 22:

> The offspring of a reproducing individual occupy initial positions in weight space that are deviations (due to mutations) from the position occupied by their parent at birth (i.e., prior to learning).

7

The effect of lifetime learning is to partially restore degraded performance of networks which have had their weights perturbed (by mutation) away from trained (through evolution) values — a form of relearning.

The Simple Model presented in this section has shown that strong ANF does exist, even when the Evolutionary Task and Lifetime Learning Task are clearly uncorrelated (see Section 1.1) in that they are based on different random vectors. However, for the explanation in terms of relearning we need the result to be presented in the next section.

4 Spontaneous Recovery of Associations by Learning Unrelated Tasks

A spontaneous recovery effect in connectionist nets was first noted in (Hinton & Sejnowski, 1986), and analysed in (Hinton & Plaut, 1987). A NN was first trained on a set of associations, and then its performance on this set was degraded by training on a new set. When retraining was then carried out on a proportion of the original set of associations, performance also improved on the remainder of that set.

In (Harvey & Stone, 1996) a more general effect is demonstrated. A NN is first trained on a set of associations, called task $\mathcal{A}$; and then performance on this task is degraded by random perturbations of the connection weights. Performance on $\mathcal{A}$ is then monitored whilst the NN is trained on a new uncorrelated task $\mathcal{C}$ (in contrast to (Hinton & Plaut, 1987) where a subset of the original task is used). It will be shown that, counter-intuitively, in most cases performance on the original task $\mathcal{A}$ initially improves.

The following is a simplistic analogy, which assumes that this effect carries over to human learning of cognitive tasks. If you have a French examination tomorrow, but you have forgotten quite a lot of French, then a short spell of learning some new task, such as unicycling, can be expected to improve your performance in the French examination. Students of French should be warned not to take this fanciful analogy too literally[5].

I give an informal argument, drawn from (Harvey & Stone, 1996), to explain the underlying geometrical reasons for this effect.

4.1 Geometric Analysis

A number of assumptions will be used here; later their validity will be evaluated.

Learning in connectionist models typically involves a succession of small changes to the connection weights between units. This can be interpreted as the movement of a point W in weight space, the

---

[5]It relies on an assumption that French and unicycling make use of a common subset of neuronal connections. The analogy also assumes that, from the set of all human capacities, ability to stay on a unicycle is uncorrelated (in the sense of Section 1.1) with ability at English-French translation. The relearning effect does indeed happen with correlated tasks (one might expect English-Spanish, even to a lesser extent English-Tagalog, to be correlated with English-French); but the uncorrelated case is the focus of interest here.
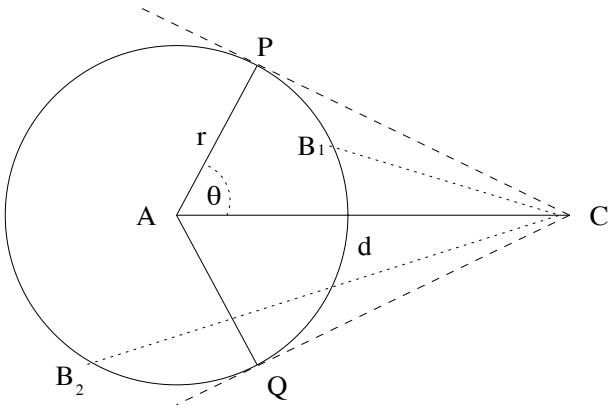
Figure 4: *In this 2-D representation of weight space, A represents weights appropriate for the first task, C for the second task. The circle represents hypersphere $\mathcal{H}$, points perturbed away from A by a distance r. Initial movement from a point B on the circumference towards C has two possible consequences: trajectory $B_1 \rightarrow C$ is outside $\mathcal{H}$, whereas $B_2 \rightarrow C$ intersects $\mathcal{H}$. See text for significance of other symbols used.*

dimensionality of which is the number of weights. For the present, we assume that training on a particular task $\mathcal{A}$ moves W in a straight line towards a point A, where A represents the weights of a NN which performs perfectly on task $\mathcal{A}$; we also assume that distance from A is monotonically related to decrease in performance on task $\mathcal{A}$.

Let A be the position of W after task $\mathcal{A}$ has been learned (see Figure 4). Assume that some 'forgetting' takes place, through random weight changes, which shifts W to a new point B. The point B lies on the surface of $\mathcal{H}$, a hypersphere of radius $r = |A - B|$ centred on A.

We then initiate training on a task $\mathcal{C}$ which is unrelated to task $\mathcal{A}$; under our assumptions, training moves W from B towards a point C, which is distance $d = |A - C|$ from A. If the line connecting B to C passes through the volume of $\mathcal{H}$ then the distance $|W - A|$ initially decreases as W moves towards C. In such cases, training on task $\mathcal{C}$ initially causes improvement in performance on task $\mathcal{A}$.

We assume that point A has been chosen from a bounded set of points $\mathcal{S}$, which may have any distribution; that $\mathcal{H}$ is centred on A; that B is chosen from a uniform distribution over the surface of $\mathcal{H}$; and that C is chosen from $\mathcal{S}$ independently of the positions of A or B. What, then, is the probability that line segment BC passes through $\mathcal{H}$? That is, what is the probability that training on task $\mathcal{C}$ generates spontaneous recovery on task $\mathcal{A}$?

If C lies within $\mathcal{H}$ (i.e. if $d < r$) then recovery is guaranteed. For any point C outside $\mathcal{H}$ there is a probability $p \geq 0.5$ of recovery on task $\mathcal{A}$. Figure 4 demonstrates this for a two-dimensional space. The point B may lie anywhere on the circumference of $\mathcal{H}$. The line segment BC only fails to pass through $\mathcal{H}$ if B lies on the smaller arc PQ; where CP and CQ are tangents to the circle, and hence $cos(\theta) = r/d$. Thus $p \geq 0.5$, and $p \rightarrow 0.5$ as $d \rightarrow \infty$.

Consider the extension to a third dimension, while retaining the same values r, d and θ. The probability $q = (1 - p)$ that BC *fails* to pass through the sphere $\mathcal{H}$ is equal to the proportion of the surface of
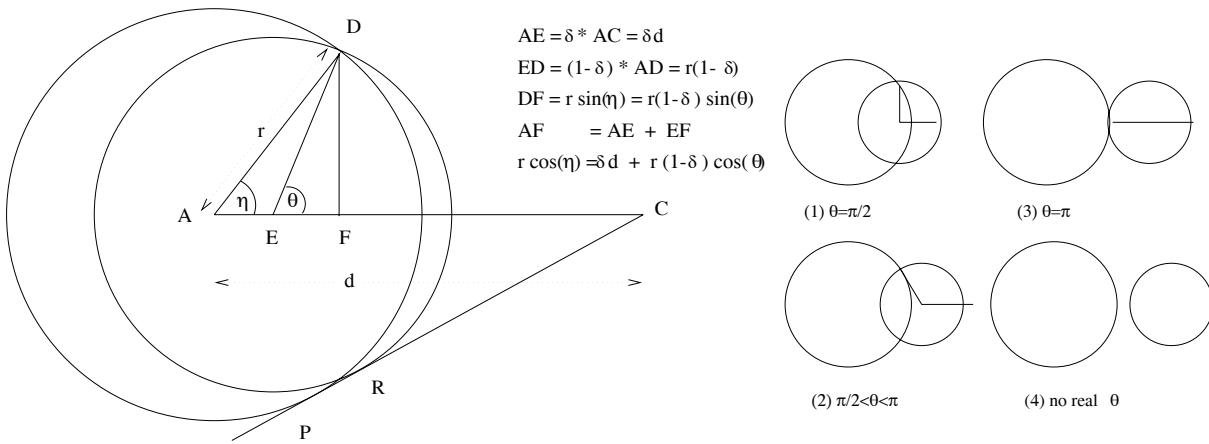
The figure contains the following labels and equations:

$AE = \delta * AC = \delta d$
$ED = (1-\delta) * AD = r(1-\delta)$
$DF = r \sin(\eta) = r(1-\delta) \sin(\theta)$
$AF = AE + EF$
$r \cos(\eta) = \delta d + r(1-\delta)\cos(\theta)$

(1) $\theta = \pi/2$ (3) $\theta = \pi$

(2) $\pi/2 < \theta < \pi$ (4) no real θ

Figure 5: *In this 2-D representation of weight space, points A, C and the larger circle are as in previous figure; the relevant angle θ is now defined within the smaller circle. Learning is assumed to move any points on the larger circle directly towards C a distance δ times their original distance from C; eg. P moves to R. Such points move onto the smaller circle centred on E where $AE = \delta \times AC$. D is on the intersection of these circles, and DF is a perpendicular dropped from D onto AC. On the right are sketched 4 further cases: in (1) $\theta = \pi/2$ and q = probability of failing to improve will be 0.5. In (2) $0.5 < q < 1.0$, in (3) and (4) $q = 1.0$.*

$\mathcal{H}$ which lies within a cone defined by PCQ with apex C. This proportion is considerably smaller in 3-D than it is in 2-D. In (Harvey & Stone, 1996) this is generalised to n-dimensions; and it is demonstrated analytically what can be seen intuitively, namely that for any given $\theta < \pi/2$, as n increases q tends to zero.

Hence for large n, it is almost certain that the line segment BC passes through the hypersphere $\mathcal{H}$. This implies that initially the point W moves from B closer to A. Hence performance improves, at least temporarily, on task $\mathcal{A}$.

Returning to the assumptions stated earlier, we can now examine their validity. First, an irregular error surface ensures that training does not, in general, move W in a straight line (though in the Simple Model above, using the delta rule, the straight line assumption is accurate). Second, perfect performance on task $\mathcal{C}$ may, in general, be associated not with one point C, but with many points which are equivalent in that they each provide a similar mapping from input to output. W may move towards the nearest of many Cs, which is therefore not chosen from $\mathcal{S}$ independently of A. This may alter the probability that W passes through $\mathcal{H}$. In the Simple Model above, where there is a single learning target vector, the assumption of a single point is accurate. For discussion of further issues, and experimental demonstrations of the phenomenon using backpropagation, see (Harvey & Stone, 1996). There it is made clear that this relearning phenomenon is only likely to be significant when the distance between weight vectors A and C is not too large in relation to the perturbations in weight vectors caused, in the current instance, by mutations. We can ensure that this condition holds where the weight vectors arise from any learning process in a feedforward NN with sigmoidal activation functions; since individual units of the NN saturate at finite values, weight vectors after training on different tasks can be expected to lie within a limited
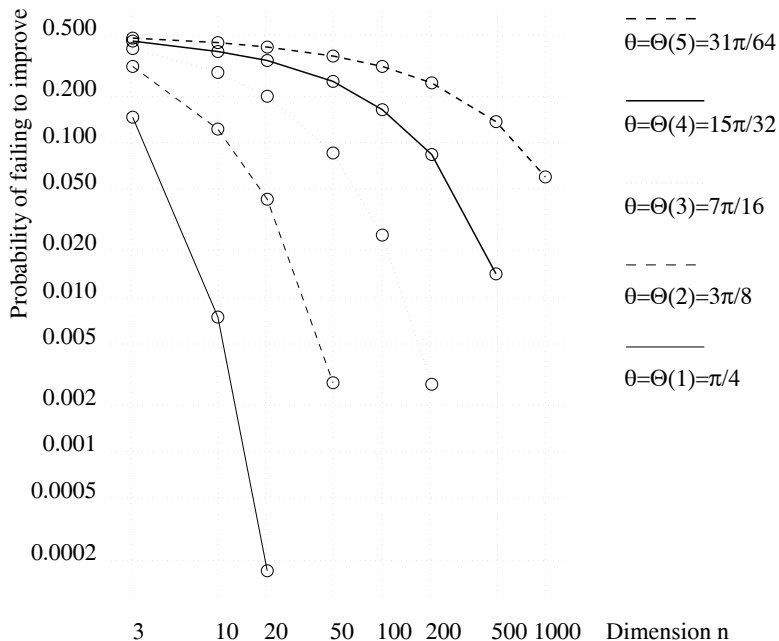
10

Probability of failing to improve

0.500
0.200
0.100
0.050
0.020
0.010
0.005
0.002
0.001
0.0005
0.0002

$\theta=\Theta(5)=31\pi/64$

$\theta=\Theta(4)=15\pi/32$

$\theta=\Theta(3)=7\pi/16$

$\theta=\Theta(2)=3\pi/8$

$\theta=\Theta(1)=\pi/4$

3   10 20   50 100 200   500 1000   Dimension n

Figure 6: *Graph showing probability of failing to improve performance at original task, against dimensionality of weight space, for various values of θ. Both axes are logarithmically scaled.*

(hyper-)volume of weight space. Since there is a bound on the distance A to C, we can make its ratio to perturbation size as small as we like by increasing the latter[6]. This condition does not, however, imply that their respective tasks are correlated — see Section 1.1.

4.2   Effects of distance moved through learning

The above discussion only applies to the first effects of an infinitesimal amount of learning on the new task. If learning continued until the new task was fully learnt, and C had been reached in weight space, then in general C is outside $\mathcal{H}$ and performance on task $\mathcal{A}$ has diminished (even if it had temporarily improved initially). I now quantify the effects of differing amounts of learning.

In Figure 5 it is assumed that learning on $\mathcal{C}$ over a period is equivalent to a single application of the delta rule which reduces error by a proportion δ; hence points on the larger circle move onto the smaller circle. Given that these circles represent the (hyper-)surfaces of hyperspheres centred on A and E, the ratio of interest is the proportion of latter surface which lies outside the former, i.e. the proportion subtended by angle θ. This ratio gives the probability q that after learning on the second task there has *failed* to be improvement in performance on the first task. It is shown in (Harvey & Stone, 1996) that this depends on the dimensionality n of the weight space, and is given by:

$$q = \frac{\int_0^\theta sin^{n-2}(\alpha)\,d\alpha}{\int_0^\pi sin^{n-2}(\alpha)\,d\alpha}$$

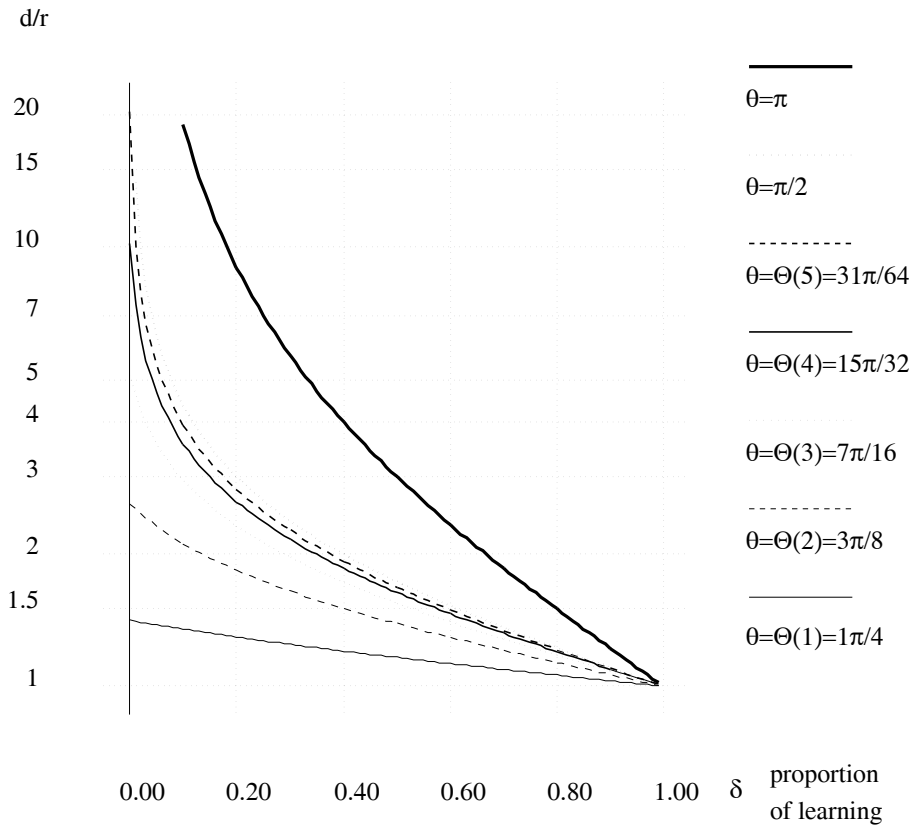---

[6]This point was drawn to my attention by Jim Stone.

Figure 7: *The five lower lines are for values of $\theta < \pi/2$. Each line plots d/r (logarithmically scaled) against δ, the amount of learning equivalent to a single application of the delta learning rule. For specific values of d/r and δ a value of θ can be estimated by interpolation; together with the dimensionality of weight space n this can be used with Figure 6 to estimate q the probability of failing to improve performance. Above are lines for $\theta = \pi/2$, corresponding to $q = 0.5$ for all n, and $\theta = \pi$ corresponding to $q = 1$ for all n.*

While a general analytical solution for all n cannot be found, in Figure 6 data points are plotted for various values of n and 5 specific values of θ: $\pi/4, 3\pi/8, 7\pi/16, 15\pi/32, 31\pi/64$. These values are equivalent to $\Theta(x) = (2^x - 1)\pi/2^{x+1}$ for x from 1 to 5.

From geometrical calculations shown on Figure 5, we have $sin(\eta) = (1 - \delta)sin(\theta)$ and $cos(\eta) = \delta(d/r) + (1 - \delta)cos(\theta)$. Using $sin^2(\eta) + cos^2(\eta) = 1$ this gives:

$$(1 - \delta)^2 sin^2(\theta) + \delta^2(\frac{d}{r})^2 + 2\delta(\frac{d}{r})d(1 - \delta)cos(\theta) + (1 - \delta)^2 cos^2(\theta) = 1$$

$$2\delta(\frac{d}{r})(1 - \delta)cos(\theta) = (1 - (1 - \delta)^2) - \delta^2(\frac{d}{r})^2$$

$$2(\frac{d}{r})(1 - \delta)cos(\theta) = 2 - \delta(1 + (\frac{d}{r})^2)$$

For any particular value of θ we can plot d/r against δ. In Figure 7 this is done for the same 5 values of θ as in Figure 6. Both these graphs are needed to calculate q the probability of failing to improve

12

	d/r = 1.4			d/r = 2			d/r = 4		
δ	Θ	$n = 20$	$n = 50$	Θ	$n = 20$	$n = 50$	Θ	$n = 20$	$n = 50$
0	1	0.0002	$\ll 10^{-4}$	1.6	0.01	0.0002	2.6	0.14	0.035
0.1	1.2	0.0008	$\ll 10^{-4}$	1.9	0.035	0.002	5.0	0.4	0.35
0.2	1.4	0.004	$< 10^{-4}$	2.3	0.08	0.012	$\frac{\pi}{2}<\theta<\pi$	$\frac{1}{2}<q<1$	$\frac{1}{2}<q<1$
0.3	1.6	0.01	0.0002	3	0.2	0.08	$\frac{\pi}{2}<\theta<\pi$	$\frac{1}{2}<q<1$	$\frac{1}{2}<q<1$
0.4	1.8	0.02	0.0008	$\theta = \frac{\pi}{2}$	0.5	0.5	$\theta = \pi$	1	1
0.5	2.0	0.04	0.003	$\frac{\pi}{2}<\theta<\pi$	$\frac{1}{2}<q<1$	$\frac{1}{2}<q<1$	no real θ	1	1

Table 1: *For different values of d/r and δ, the appropriate value of θ is estimated from Figure 7. Where $\theta < \pi/2$, this is estimated in terms of Θ, which then allows for interpolation between the $\Theta(i)$ values on Figure 6. Where $\theta >= \pi/2$, or is there is no real θ, its value is listed explicitly. The approximate probabilities q of failing to improve performance are listed for two dimensionalities of weight space, $n = 20$ and $n = 50$.*

performance on the first task as a consequence of learning δ on the second task, with relevant parameters being d/r and n. The lack of an analytical solution for arbitrary values of n means that in practice interpolation must be used between the few specific values of n calculated.

There are two further values for θ of particular interest, and lines for these values are also plotted in Figure 7. When $\theta = \pi/2$ this gives the equation $\delta = 2/(1 + (\frac{d}{r})^2)$, corresponding to $q = 0.5$. When $\theta = \pi$ this gives $\delta = 2/(1 + \frac{d}{r})$, corresponding to $q = 1$. These latter two results are independent of dimensionality n.

When $\delta = 0$ this gives the probability of failing to improve infinitesimally as a result of an infinitesimal amount of learning, the calculation originally given in (Harvey & Stone, 1996). This is here extended for a delta learning rule for $\delta > 0$; caution should be used in carrying this over to other learning techniques such as back propagation. The general picture emerges from Figure 7 and Table 1 that the ANF effect is strongest for small values of d/r (when $d/r < 0.5$ then even with large amounts of learning such as $\delta = 0.5$ it is still nearly certain) and at these values dimensionality n is of significance. For larger values of d/r, for instance 5, then for $\delta = 0.1$ the effect is less than 50% and for $\delta > 0.333$ it has disappeared completely. For $d/r > 10$ then it just needs $\delta > 0.02$ for the effect to be less than 50%; it should be remembered, however, that for infinitesimal amounts of learning the effect persists out to arbitrary values of d/r.

In typical circumstances for NNs where trained weight vectors can be expected to lie within a limited space (see end of previous section) d will also be bounded. The ANF effect will be most prominent when r is relatively large in comparison; in evolutionary cases r corresponds to perturbations through mutation, and these were indeed relatively large in (Parisi et al., 1992; Nolfi et al., 1994).
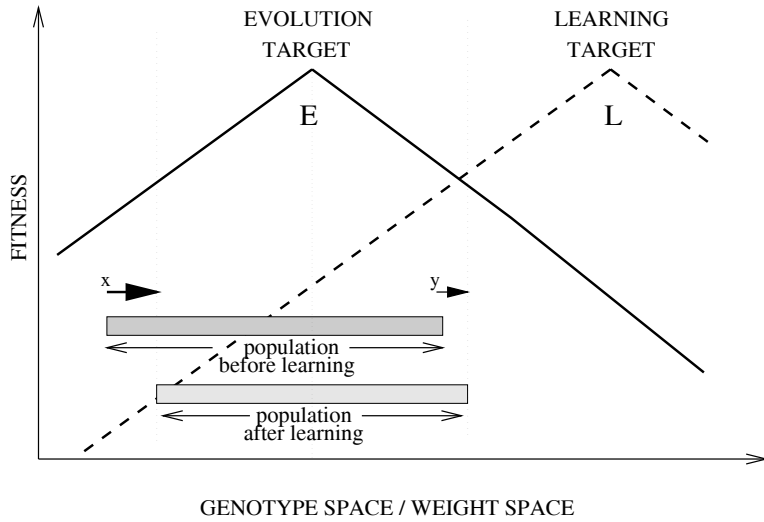
Figure 8: *A sketch of Evolutionary and Learning fitness landscapes for the Simple Model. The 50-dimensional weight space is loosely represented by the horizontal axis. After a few generations the population after learning (i.e. when it is evaluated) will be centred around* **E***, implying that before learning moved it towards* **L** *it was further to the left. Applications of learning rules such as the delta rule towards target vector* **L** *will move points that are left of* **E** *(e.g. x) further than those right of* **E** *(e.g. y).*

5 Discussion

The link between this relearning effect and ANF as shown in the Animat Model was first pointed out in (Harvey, 1996). Here the same ANF has been recreated in a much simpler model where the two tasks are clearly uncorrelated, and it can be seen that this relearning effect is a sufficient explanation for the improvement in performance on the Evolutionary Task. Each individual in the population has inherited its weights from an elite member of the previous generation; these weights have then been perturbed by mutation, but lifetime learning on an uncorrelated task improves performance on the Evolutionary Task.

A different explanation is offered in (Nolfi et al., 1994), where the notion of dynamic correlation between fitness landscapes is introduced (p.22 and p.23):

> ...two surfaces are dynamically correlated if a weight matrix that moves towards higher values on one surface tends also to move toward higher values on the other surface.
> ...evolution progressively selects for individuals that are located in regions of dynamic correlation between the fitness surface and the learning surface.

In the Simple Model which reproduces strong ANF the fitness surfaces are extremely simple, and are sketched in Figure 8. From the earlier graphs it can be seen that the expected distance of the random vector **L** from **E** is, as with the similarly determined random initial population, around 5.5; and after a few generations the average distance of members of the population from **E** is less than 3.0. Hence one can assume that typically the population lies within the region of genotype space indicated in Figure 8, loosely centred around **E** after learning since selection relies on evaluating them after they have learnt.

14

TARGET E   0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

GEN. G   0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0

TARGET L 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1

.... **Loci where G**   - - - ⇒ **Changing G to match L at this locus**
- - - - - **already matches L**   **5%**   **also changes G to match E**
50%

───⇒ **Changing G to match L at this locus**
45%   **disrupts a place where G matched E**

Figure 9: *Binary genotypes: a member* **G** *is shown which is perturbed from the target* **E** *in just 10% of its loci. The second Learning target vector* **L** *typically differs from* **E** *in 50% of loci, as they are assumed uncorrelated. When learning changes loci of* **G** *to match* **L***, an expected 50% of the time there is no change, 45% of the time the match of* **G** *with* **E** *is made worse, and only 5% of the time is it made better.*

This implies that before learning the population was, in the figure, to the left of **E**. One would expect more than 50%, but by no means all, of the population to be born in a region of 'dynamic correlation' between fitness surfaces.

In so far as the 'dynamic correlation' hypothesis refers to a single point representing the population average and its position on the two fitness landscapes, it is no more than a restatement of the phenomenon in different terms. If the hypothesis is intended to cover each individual member of the population, then in general, as in the example given here, it is not correct. The 'relearning' explanation presented here as an alternative provides a sufficient, and completely different, explanation.

It has been suggested[7] that the relearning effect may be related to James-Stein shrinkage (Efron & Morris, 1977; James & Stein, 1961). That is, reducing the variance of the outputs of a NN reduces the squared error at the expense of introducing a bias. It may be that training on the unrelated task **L** incidentally induces shrinkage. Figure 8 illustrates this in one of the n dimensions of weight space; learning adjustments that happen to take points towards **E** (e.g. x) tend to be larger than adjustments which take points further from **E** (e.g. y). Hence when dealing with a large number of cases one can expect the average distance of such points from **L** to decrease, at least initially; the effects of more-than-infinitesimal learning movements were analysed in section 4.2.

---

[7]G. E. Hinton, personal communication.

6 Restrictions on the Phenomenon

The ANF phenomenon exhibited in the Animat and Simple Models relies on both evolution and lifetime learning acting on the same units, in this case weights of a NN or components of a vector. In this respect it is similar to the Baldwin effect. Although the Baldwin effect has been demonstrated with an essentially binary genotype (Hinton & Nowlan, 1987), for this effect real-valued components (or approximations thereof) are essential.

To show that this effect would not work with a binary genotype, consider Figure 9. In this simple example, relating to a binary genotype, a member **G** mutated from the Evolutionary Target genotype **E** is assumed to have 90% of its loci correct, i.e. matching those of **E**. Typically an unrelated Learning vector **L** differs from **E** in around 50% of its loci. Any learning method which tends to change loci of **G** towards the corresponding alleles of **L** will affect its performance at task **E** in the proportions shown — changes for the worse typically greatly outnumber changes for the better.

The effect has only been analysed here for an EA which has no recombination. However when the population is genetically converged one would expect recombination to make little difference; this expectation has been supported by adding recombination to the Simple Model and observing similar results. The explanation relies on the population being in a local region of weight space, around **E**, in which distance from **E** is monotonically related to performance. Hence the strong truncation selection used in this EA, coupled with significant weight changes through mutation, provide ideal conditions for quickly converging a population around the target. In the absence of comparable conditions there is no reason to expect this ANF phenomenon.

7 Conclusion

Interactions between evolution and learning are tricky, and need to be analysed with care. The phenomenon originally demonstrated in the Animats Model (Parisi et al., 1992; Nolfi et al., 1994) showed that, under some circumstances, lifetime learning on a different task can aid progress through evolution on the evolutionary task. Only in the earlier of these two papers was this claimed for two uncorrelated tasks, the effect I have termed strong ANF. This claim has been verified here, and demonstrated in the Simple Model, using target vectors **E** and **L** selected independently at random. The explanation comes not from interactions between evolution and learning *per se*, but rather from a relearning effect on weights perturbed by mutation. Hence the effect can only be expected under very restricted conditions where evolutionary processes of mutation are mimicking the effects of weight perturbation around a learnt set of weights.

The Baldwin effect requires the evolutionary and learning tasks to be similar for genetic assimilation to take place. The strong ANF effect is based on two separate tasks. If one takes the special case where these two tasks are indeed the same, then indeed d/r is at a minimum value of zero and the ANF effect will be at its strongest — but it has been reduced to the trivial and uninteresting statement that

	Baldwin effect with genetic assimilation	ANF effect
Key factor	Genetic assimilation	Improvement on uncorrelated task
Number of tasks	One is sufficient	Two necessary, **E** and **L**
Correlation	Yes	Works when uncorrelated
Improve at-birth performance	Yes	No
Cost of learning	Needed	Unnecessary
Size of mutations	Not crucial	Size compared to d is very significant
Amount of learning	Not crucial	If too large effect typically disappears
Dimensionality	Typically irrelevant	The bigger n is, the stronger the effect
Specific to evolution	Usually assumed so	General, any weight-perturbations
Binary genotype OK?	Yes	No, needs real values

Table 2: *A summary of some differences between the Baldwin effect (with genetic assimilation) and ANF. This list is not meant to be exhaustive.*

'improving performance on $\mathcal{B}$ also improves performance on $\mathcal{A}$ when $\mathcal{A} = \mathcal{B}$'. In a natural context one expects learning to have some cost to the individual; if one artificially manipulates these costs to be zero or less, this will have different consequences for the two effects. The Baldwin effect requires there to be some positive cost of learning, such that individuals that need to learn less are thereby fitter, for genetic assimilation to take place; the ANF effect does not need any cost to learning. For the Baldwin effect the size of mutations, the rate of learning and the dimensionality of weight space are not of crucial significance (as can be seen from consideration of the simulations in (Hinton & Nowlan, 1987)), whereas these have critical relevance to the existence of, or extent of the ANF effect. These differences are summed up in Table 2.

The effect I have termed 'Another New Factor' means that under the circumstances outlined here lifetime learning on one task can indeed improve performance on another task which is the subject of evolutionary selection pressure, even when the two tasks are uncorrelated. The effect is orthogonal to the Baldwin effect. The restrictive conditions for this effect mean that under most circumstances it will not occur in practice; nevertheless in (Parisi et al., 1992) the conditions were right.

Acknowledgments

References

Baldwin, J. M. (1896). A new factor in evolution. *American Naturalist, 30*, 441–451.

Efron, B., & Morris, C. (1977). Stein's paradox in statistics. *Scientific American, 236*(5), 119–127.

Harvey, I. (1996). Relearning and evolution in neural networks. *Adaptive Behavior, 4*(1), 79–82.

Harvey, I., & Stone, J. V. (1996). Unicycling helps your French: Spontaneous recovery of associations by learning unrelated tasks. *Neural Computation, 8*(4), 697–704.

Hinton, G., & Plaut, D. (1987). Using fast weights to deblur old memories. In *Proceedings of the 7th Annual Conference of the Cognitive Science Society, Seattle*.

Hinton, G., & Sejnowski, T. (1986). Learning and relearning in Boltzmann machines. In Rumelhart, D., McClelland, J., & the PDP Research Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations*, pp. 282–317. MIT Press/Bradford Books, Cambridge MA.

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

James, W., & Stein, C. (1961). Estimation with quadratic loss. In *Proc. of the Fourth Berkeley Symposium on Mathematical Statistics and Probability 1*, pp. 361–380 Berkeley, California. University of California Press.

Nolfi, S., Elman, J., & Parisi, D. (1994). Learning and evolution in neural networks. *Adaptive Behavior, 3*(1), 5–28.

Parisi, D., Nolfi, S., & Cecconi, F. (1992). Learning, behavior and evolution. In Varela, F. J., & Bourgine, P. (Eds.), *Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life*, pp. 207–216. MIT Press/Bradford Books, Cambridge, MA.