Learning and Evolution: On the Effects of Directional Learning

Stefano Nolfi   Anne Spalanzani
Division of Neural Systems and Artificial Life,
Institute of Psychology, National Research Council (CNR)
Viale Marx, 15, 00137, Rome, Italy
nolfi@ip.rm.cnr.it, Anne.Spalanzani@imag.fr

Abstract

In this paper we demonstrate that learning tend to have a beneficial effect on evolution even if the characters that are acquired through learning do not have any adaptive advantage by themselves. The beneficial effect is due to the fact that learning force evolution to select individuals that are located in regions of the search space in which the learning and the evolutionary tasks are dynamically correlated. This has two implications: (1) evolving individuals display a predisposition-effect to benefit from learning, i.e. while improving their performance with respect to the learning task during lifetime they also tend to improve their performance with respect to the evolutionary task, and (2) evolving individuals display a shelter-effect, i.e. lifetime learning allows them to partially recover from mutations that are deleterious with respect to the evolutionary task. This, in turn, might allow learning individuals to tolerate higher mutation rate than non-learning individuals thus enhancing the exploratory power of the evolutionary search.

1. Introduction

In the effort to explain evolutionary gaps in the fossil records, more than a century ago James Mark Baldwin (Baldwin, 1896) advanced the idea that evolution could be influenced by learning during life without assuming that learned features could directly modify the genotype (as hypothesized by Lamarck, 1914)[1]. Baldwin's argument was that learning accelerates evolution because sub-optimal individuals can reproduce by acquiring during life necessary features for survival. However, since learning requires time (and might thus be a disadvantage), Baldwin suggested that evolution tends to select individuals who have already at birth those useful features which would otherwise be learned. This latter aspect of Baldwin's effect, namely indirect genetic assimilation of learned traits, has been later supported by scientific evidence and defined by Waddington (1942) as a canalization effect.

   A simple and clear demonstration of how learning influence evolution has been provided by Hinton and Nowlan (1987) with a simple model based on artificial evolution (see section 2). In particular the authors found that: (a) random changes occurring during individuals lifetime, by altering the shape of the search space, might significantly speed up the evolutionary search (we will refer to this as the *guide-effect* of learning on evolution), and (b)

---

[1] Similar views were expressed in the same period by Morgan (Morgan, 1896) and Osborn (Osborn, 1896).

given that learning has a cost, traits that are first acquired through learning tend to be genetically specified in successive generations (we will refer to this as the *assimilation-effect* of previously learned traits).

Few years later Nolfi et al. (1990, 1994) found that learning had positive effects on a population of agents that evolve to get fitter at one task and also learn, during their lifetime, a different task (see also Nolfi and Parisi [1996] and Nolfi [1999]). In particular the authors found that: (a) after few generations evolving individuals, while improving their performance on the learning task, also showed an improvement on the evolutionary task during lifetime (we will refer to this as the *predisposition-effect* to benefit from learning), and (b) learning individuals displayed better performance on the evolutionary task at the population level (we will refer to this as the *shelter-effect* of learning). It should be noticed that the predisposition-effect was also observed in cases in which the learning and the evolutionary tasks were apparently uncorrelated (Parisi et al. 1992).

Recently Harvey (1997) observed a positive effect of learning on evolution in a case in which, as in the case of the experiments described above, the learning and the evolutionary tasks were different. In particular the author observed that: (a) already from the initial generation, individuals improved their performance on the evolutionary task while learning another uncorrelated task, and (b) learning individuals displayed better performance on the evolutionary task. Harvey proposed that both these effects can be explained as form of relearning (and not as a result of the interaction between learning and evolution). We will refer to these effects and to their explanation as the *relearning-effects*.

In this paper, we will demonstrate that the effects observed by Harvey (1997) are due to the fact that the two tasks are dynamically correlated (for a definition of dynamical correlation see below). Beside of that, by relying on a variation of the simple experimental framework designed by Harvey, we will demonstrate that the predisposition-effect and the shelter-effect are a very general phenomena that arises, independently of the learning task, in all cases in which learning has a directionality.

We will start by briefly reviewing the experiments of Hinton and Nowlan's (1987) and of Nolfi et al. (1994) in section 2 and 3 respectively. We will review the experiments conducted by Harvey (1997) and we will present additional analysis on these experiments in section 4. We will present our new results in section 5. Finally, we will discuss the implications of the obtained results in section 6.

2. On the effects of random changes occurring during individuals lifetime

A simple and clear demonstration of how learning[2] influence evolution has been provided by Hinton and Nowlan (1987) with a simple model based on artificial evolution. The authors considered and extreme simple example in which only a single combination of gene values confers added reproductive fitness to a population of evolving individuals. Individuals have a genotype with 20 genes that can assume two alternative values (0 or 1). The only combination of genes that provide a fitness value above 0 consists of all ones. In this extreme case, the probability of finding the good combination of genes would be very small given that the fitness surface looks like a flat area with a spike in correspondence of the good combination (see the thick line in Figure 1, left and right). Indeed, on such a surface, artificial evolution does not perform better than random search. Finding the right combination is like a looking for a needle in a haystack.

---

[2] As in the case of Hinton and Nowlan (1987) we will use the term learning to indicate any form of change occurring to evolving individuals during their lifetime.
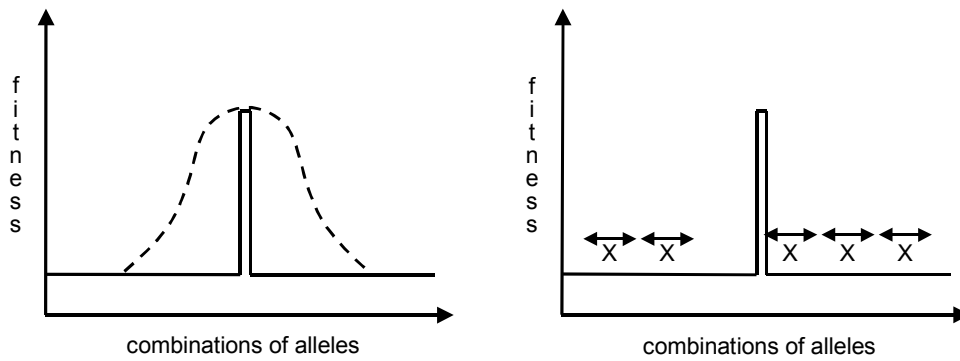
Figure 1. **Left:** Fitness surface with and without learning. In absence of learning, the fitness surface is flat, with a thin spike in correspondence of the good combinations of alleles (thick line). When learning is enabled, the fitness surface has a nice hill around the spike which includes the alleles combination which have in part right fixed values and in part unspecified (learnable) values (thin line). The thick line represents the fitness for each possible combination of two alleles ([0, 1]) while the thin line represents the fitness for each possible combination of three alleles [0, 1, ?]). Redrawn from Hinton & Nowlan (1987). **Right:** The line represents the fitness for each possible combination of two alleles [0, 1]. The cross-marks represent the positions of the individuals at birth. The arrows represent the movements of the individual in the search space corresponding to changes of modifiable alleles.

The situation is much different in the case of learning individuals. One simple way to introduce learning is to assume that, in learning individual, genes can have three alternative values [0, 1, and ?] where question marks indicate modifiable genes whose value is randomly selected within [0, 1] each time step of the individuals' lifetime. In the model proposed by the authors, therefore, learning does not have a directionality and consists in random generating the value of all unspecified genes (i.e. genes whose inherited value is ?). As shown by the authors, performance increases much faster along generations in the case of learning individuals than in their non-learning equivalents. The addition of learning, in fact, produces an enlargement and a smoothing of the fitness surface area around the good combination which can be discovered and easily climbed by the genetic algorithm (see dashed line in Figure 1, left). This is due to the fact that not only the right combination of alleles but also combinations which in part have the right alleles and in part have unspecified (learnable) alleles might report an average fitness greater than 0 (fitness monotonically increases with the number of fixed right values because the time needed to find the right combination is inversely proportional, on the average, to the number of learnable alleles). Learning makes the fitness surface smoother, and this, in turn, simplifies the search which should be performed by evolution. As claimed by Hinton & Nowlan, with learning "it is like searching for a needle in a haystack when someone tells you when you are getting close" (1987, p. 496). This is what we have called the guide-effect of learning.

In the representation adopted in the left part of Figure 1 each individual is represented as a point on the fitness surface with a height corresponding to the average fitness of the individual during its lifetime. This is a static representation in which changes in performance during lifetime cannot be visualized. Another way of representing the individuals in the search space is to imagine that each point in the search space correspond to a given combination of 0 and 1 (see Figure 1, right). In this case changes occurring during individuals' lifetime correspond to movements on the search space (see the arrows in Figure 1, left). By using this convention the advantage of learning can be explained by considering that learning individuals, by moving in the search space during lifetime, are more likely to spend at least few cycles of their lifetime on the spike. In other words we can say that learning allows the evolutionary process to explore the landscape surrounding each candidate for reproduction (Nolfi et al., 1990).

Once individuals that have part of their genes fixed on the right values and part of their genes unspecified (learnable) are selected, individuals with less and less learnable genes tend to be selected given that fitness monotonically increases by decreasing the number of learnable genes (an equilibrium point is eventually reached, see Hinton & Nowlan, 1987). In other words, characters that were first acquired through learning tend to become genetically specified later on. This is what we have called the assimilation effect.

Despite its explicative power, this model has several limitations: (1) learning is modeled as a random process; (2) there is no distinction between the learning task and the evolutionary task; (3) the learning space and the evolutionary space are completely correlated. The two spaces are correlated if genotypes which are close in the evolutionary space correspond to phenotypes which are close in the phenotype space (Mayley, 1997). In this model, learning and evolution operate on the same entities (i.e. the connection weights) with the same operators (i.e. both changes produced by mutations and changes produced by learning correspond to substitutions of genes with new values which are randomly selected). Therefore the two spaces are completely correlated. By systematically varying the cost of learning and the correlation between the learning space and the evolutionary space, Mayley (1997) showed that: (1) the adaptive advantage of learning is proportional to the correlation between the two search spaces; (2) the assimilation of characters first acquired through learning is proportional to the correlation between the two search spaces and to the cost of learning (i.e. to the fitness lost during the first part of the lifetime in which individuals have sub-optimal performance); (3) in certain situations learning costs may exceed learning benefits.

In the next sections we will presents other experimental setting in which part or all these limitations are released.

3. The case of a learning task that differ from the evolutionary task.

Nolfi et al. (1994) have studied the case of artificial agents (also known as animats, see Wilson, 1987) that evolve (to become fitter at one task) at the population level and learn (a different task) at the individual level. In particular, individuals which were selected for their ability to find food in their environment were also asked to learn to predict the sensory consequences of their motor actions during their lifetime.
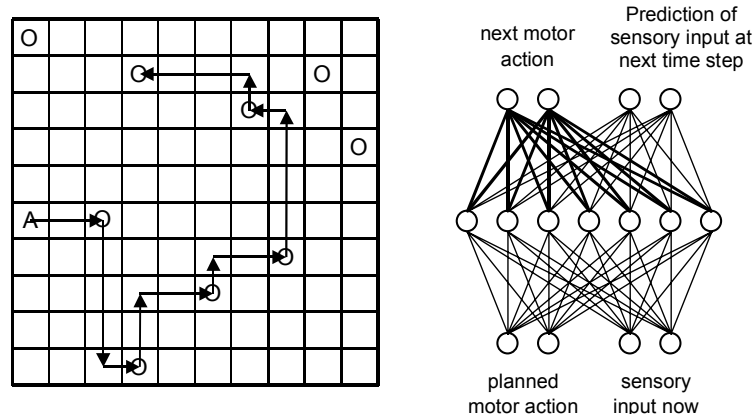

Figure 2. **Left**: The environment containing 10 food tokens (O) and the animat (A). The trace on the terrain represents the trajectory of a typical evolved individual. **Right**: Neural network architecture. All connections weights are inherited; however, connections represented with thin lines are also modified by prediction learning during the lifetime of the individual while connections represented with thick lines are not.

Each individual animat lives in a two-dimensional grid world where a number of food tokens are randomly distributed (Figure 2, left). Each food token occupies one cell; if the animat happens to step on one of these cells, the food token is automatically "eaten" and the animat's fitness is increased. Individuals are equipped with a neural network interfaced to a sensorimotor system that provides input information on the distance and angle (with respect to the facing direction of the animat) of the nearest food token, and on the planned motor action (Figure 2, right). Two input units encode the angle and the distance of the nearest food token and two other units (thresholded to the nearest binary value) encode one of four possible actions: turn 90° right, turn 90° left, move one cell forward, and remain still. At each time step, the neural network receives as input the sensory information on the nearest food token and the current planned motor action and produces as output the next planned action and a prediction of the sensory state after the execution of the current planned action. At this point: (a) the planned action that was used as input is executed and the next planned action is passed as new input; (b) the freshly-gathered sensory information is used both as input and as teaching input for the output units encoding the predicted state of the sensors (the new sensory state is compared with the predicted state and the difference (error) is used to adjust by back-propagation the connection weights between the four input, the seven hidden, and the two prediction units).

Each sensorimotor cycle is repeated for 20 epochs (life span) during which the animat is allowed to spend 50 actions in 5 environments with randomly different food distributions (for a total of 5000 cycles). The initial population is composed of 100 individuals, each with the architecture described in Figure 2 and randomly assigned connection weights in the ± 1.0 interval. At the end of life the 100 individuals are ranked in terms of their fitness (total number of food elements eaten during life) and the best 20 individuals are allowed to reproduce by generating 5 copies each of their connection weights. The inherited original weight matrices (changes due to learning during life are discarded) are mutated by selecting 5 weights at random and perturbing the weight's value by adding a quantity randomly selected in the ± 1.0 interval. The process is repeated for 100 generations.
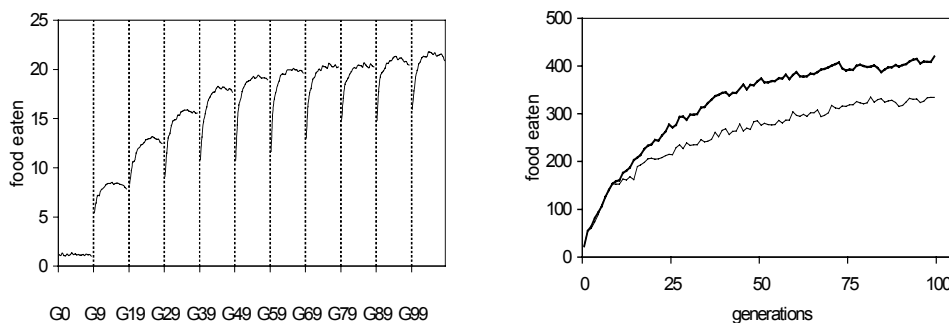

Figure 3. **Left:** Average number of food elements eaten by populations of successive generations which learn to predict. Each curve represents performance prior to learning and then for each of the 20 epochs of life (performance prior to learning are obtained by measuring the number of food tokens eaten by individuals during one epoch of life without updating the weights). For reasons of space, performance are displayed only each 10 generations. **Right:** Average of food tokens eaten by populations of animats throughout generations for experiments with and without learning. **Left** and **Right:** Average results over 10 replications.

The results showed that, after a few generations, individuals learning to predict also increased their ability to find food during life (Figure 3, left). Moreover, by comparing the

results of the experiments described above with another set of experiments in which individuals were not allowed to learn to predict during their lifetime, it was shown that learning populations displayed faster and higher fitness values across generations than populations without learning (Figure 3, right). Similar results were obtained in other cases and in particular in cases in which the learning task and the evolutionary task were apparently uncorrelated (see Parisi et al., 1992).
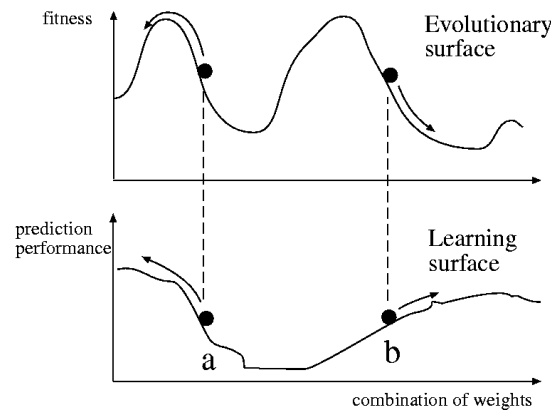

Figure 4. Fitness surface for the evolutionary task and performance surface for the learning task (sensory prediction) for all possible weight combinations. Movements due to learning are represented as arrows. Point *a* is in a region in which the two surfaces are dynamically correlated. Even if *a* and *b* have the same fitness on the evolutionary surface at birth, *a* has more probability to be selected than *b* since it is more likely to increase its fitness during life than *b*.

The first effect, namely the fact that after few generations individuals increase during lifetime their ability to eat while learning to predict, can be explained by considering that evolution will tend to select individuals that are located in regions of the search space where the learning and evolutionary surfaces are dynamically correlated (Nolfi et al., 1994; Parisi and Nolfi, 1996). Changes due to learning produce a movement of the individual phenotype both on the learning and the evolutionary surfaces (see Figure 4). However, because learning tries to maximize performance on the learning task, individuals will move toward the higher area of the learning surface. Given that the way in which individuals move in weight space affects their fitness (the total fitness of the individual is the sum of the fitness values received during such displacements on the weight space) evolution will tend to select individuals located in areas in which, by increasing their performance on the learning task, they also increase their performance on the evolutionary task. The final result is that evolution will have a tendency to progressively select individuals which are located in dynamically correlated regions. In other words, evolution will tend to select individuals that have a predisposition to improve their performance on the evolutionary task by learning, independently from what individuals learn during their lifetime. For this reason we named it predisposition-effect.

The fact that the loss in performance produced by mutations (mutations are counter-adaptive, on the average) can be recovered only in learning individuals may explain the second observed effect, namely that learning individuals outperform non-learning individuals, on the average (see Figure 3, right). The abilities that are lost by the population due to mutations can be restored, at least in part, by learning given that evolving individuals tend to be located in areas where the evolutionary and learning surfaces are dynamically correlated (Nolfi, 1999). For this reason we named this second positive effect of learning on evolution shelter-effect. Further supports to this hypothesis come from the analysis conducted

by Carse and Oreland (2000) on a replication of this experiment. The authors, in fact, observed that the average performance of the top 20 non-learning individuals (i.e. the population of the next generation before mutations were applied) and the average performance of the learning individuals (after mutations were applied) are almost identical. On the basis of this observation the authors claimed "…that learning is not guiding evolution in the sense of finding better individuals but rather that lifetime learning is diminishing the detrimental effect of mutation on the average population fitness" (p.9).

We will come back to these effects in section 5 were we will describe a much more simple experimental framework in which these effects can be observed and more easily analyzed.

4. On the role of re-learning

In a recent paper Harvey (1997) tried to verify if the second effect described in the previous section (namely the fact that learning individual outperform non-learning individuals on the average) could be observed in an experimental framework in which the learning and the evolutionary task were clearly uncorrelated.

The author devised a simple experimental framework in which, as in the case of the simple model described in section 2, there is a direct relation between genotypes and phenotype but in which evolving individuals can display a simple form of directional learning as in the case of the experiments described in the previous section (Harvey, 1997). Genotypes consist in vectors of 50 real numbers that, in the initial population, are randomly drawn from the interval [-1.0, 1.0]. The evolutionary and learning tasks consist in minimizing the distance of the genotype from two fixed target vectors (**E** and **L** respectively) that are randomly drawn from the interval [-1.0, 1.0]. Performance, with respect to the evolutionary and learning task, is given by the inverse of the Euclidean distance of the genotype from **L** and **E** respectively. Learning consist in a single application of the delta rule during individuals lifetime (i.e. the genotype vector is moved once toward the target learning vector by a proportion $\Delta = 0.1$ of its actual distance). The fitness of the genotype consists of the Euclidean distance after the learning step (in the case of the learning individuals). Population consisted of 100 individuals. The top 20 individuals were selected and were allowed to reproduce. Two set of experiments were conducted by allowing each reproducing individual to generate 4 mutated copies and by keeping the 20 elite members unchanged or by allowing each reproducing individual to generate 5 mutated copies. Mutations consist in adding a random value drawn from [-1.0, 1.0] to 5 randomly selected genes. Offspring inherited the genotypes of their parent before learning (i.e. changes due to learning were not inherited).

It should be noticed that the fact that individuals are evaluated at the end of their learning phase has two implications: (1) acquiring a character through learning has no cost and therefore one should not expect a tendency to genetically assimilate characters previously acquired through learning, and (2) learning cannot guide evolution as in the experiment described in section 2. Indeed, although individuals explore two points of the search space during their lifetime (the point corresponding to their genotype at birth and after the movement toward **L**), only the fitness corresponding to the latter point has an effect on the selection process.
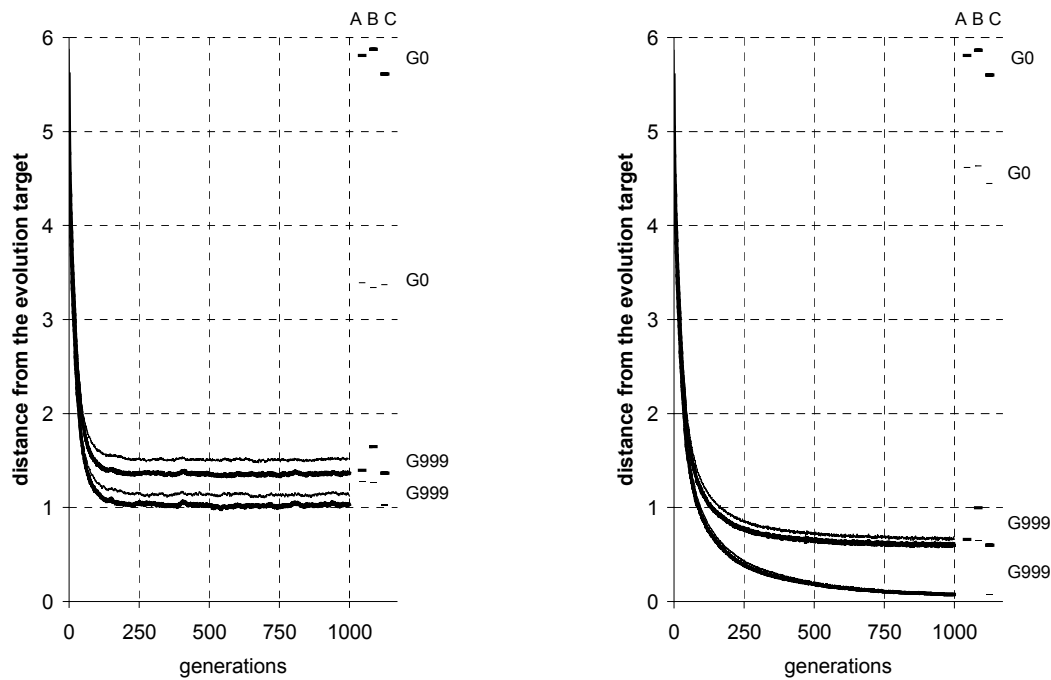
Figure 5. Thin and thick curves represent the distance from the evolutionary target of non-learning and learning individuals respectively. Both the average performance of the population and the performance of the best individual of each generation are shown. A B and C represent the performance before mutations, after mutations, and after learning in the case of learning individuals (thick and thin lines represent the performance of the population and of the best individual respectively). ABC data are shown for generation 0 and 999 only. Average results of 20 replications. **Left:** Performance in the experiments in which mutations were applied to all members of the population. **Right:** Performance in the experiments in which the 20 elite individuals were preserved.

Figure 5 shows the results obtained by replicating the original experiments but by continuing the evolutionary process for 1000 generations instead than 100. Obtained results are consistent with those described in Harvey (1997). It should be noticed, however, that in the experiment in which the 20 elite member of the population were keep unchanged performance continue to increase after the first 100 generations and reach close to optimal performance[3].

By comparing the performance of experiments conducted with and without learning one can see the same two effects that we observed in the experiments described in the previous section: (a) during lifetime learning individuals also tend to increase their performance with respect to the evolutionary task (Figure 5, see BC data), and (b) learning individuals outperform non-learning individuals (Figure 5, see the thick and thin lines corresponding to learning and non learning individuals respectively). The difference is that in this experiment the former effect is observed already from the initial generation (i.e. in absence of any selection and reproduction within the population) while in the experiments described in the previous section it was observed only after few generations. More precisely, as observed in Harvey (1997), already from the first generation the distance from the evolutionary target first increases as a result of mutation and then decreases as a result of lifetime learning (i.e. of movement toward the learning vector [Figure 5, see ABC data]).

---

[3] Similar results can be obtained by mutating all the individuals but by reducing the mutation rate.

Accordingly to Harvey (1997, p. 318), all these effect can be explained by considering that:

> *The effect of lifetime learning is to partially restore degraded performance of networks that have had their weights perturbed (by mutation) away from trained (through evolution) values--a form of re-learning.*

This hypothesis is based on evidence that by perturbing the weights of a neural network previously trained with back-propagation on a set of input-output pairs and then retraining the network on a new training set, uncorrelated with the original one, performance also improves on the patterns belonging to the original training set (Harvey and Stone, 1996). As shown in Figure 6, this form of re-learning can be illustrated with a geometrical argument (Harvey and Stone, 1996; Harvey, 1995, 1997).
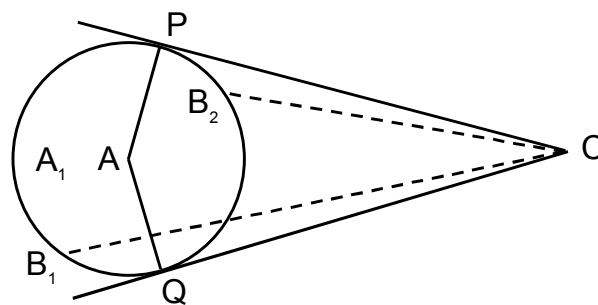

Figure 6. Two dimensional representation of the search space (see text for explanation). Redrawn with modification (A_1 added) from Harvey (1995, 1997).

Assume that A represents the weights of the network trained on the original training set, B_1 and B_2 are two possible positions of the network after perturbation, and C is the position of the network after being trained on the second training set. Finally, assume that performance on the original set is inversely proportional to the distance from point A. Whenever B lies outside the inner arc PQ (e.g., B_1), its trajectory gets closer to A for some time; instead, whenever B lies inside the inner arc PQ (e.g., B_2), its trajectory always goes away from A. Regardless of the position of C, the former situation happens more than 50% of the times for a 2-dimensional weight space and much more often in a high-dimensional weight space; furthermore, it happens 100% of the times when C lies within the circle (Harvey, 1997).

On the basis of this explanation, the author claimed that the beneficial effects of learning can be explained by considering a highly converged evolved population distributed around point A (the evolutionary target) being pulled away by mutations toward B, and then moving toward point C (the learning target) during lifetime learning. Harvey (1995, 1997) also proposed this as an alternative explanation of the dynamical correlation hypothesis described in the previous section (for additional data that demonstrate that the relearning hypothesis cannot explain the results observed in the experiments described in the previous section see Nolfi [1999] and Carse and Oreland [2000]).

If we look at the position of the top 20 individuals in the search space however, we can easily see that the evolving population does not tend to converge around A but around A_1 (see Figure 6), i.e. to a point that ensures that by learning (i.e. by moving toward C of a given Δ) individuals will reach A. Figure 7 shows the Euclidean distance of the top 20 individuals before learning takes place from the evolutionary target **E** and from **E** - Δ**L** (thin and thick lines respectively). Notice that **E** - Δ**L**, i.e. the position from which, by learning, individuals

can reach the point corresponding to the evolutionary target correspond to point A_1 in the case of Figure 6. This result fits nicely the predisposition-effect hypothesis described in the previous section that postulates that evolution tends to select individuals that are located in dynamically correlated regions of the learning and evolutionary surfaces.
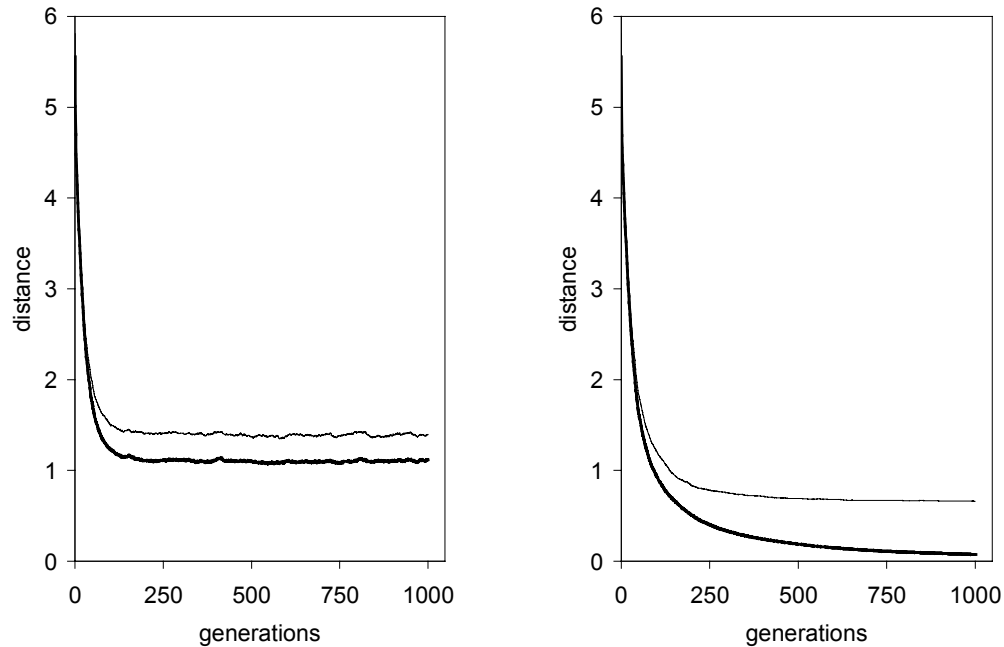

Figure 7. Average distance from **E** and **E** - Δ**L** (thin and thick line respectively) of the 20 elite members throughout generations. **Left:** Performance in the experiments in which mutations were applied to all members of the population. **Right:** Performance in the experiments in which one copy of each elite individuals was preserved by mutations.

It remains to be explained: (a) why in this experiment a positive effect of lifetime learning is observed already in the initial generation, and (b) whether the advantage of learning with respect to non-learning individuals on the average can be explained as a form of shelter-effect or as a form of re-learning (despite population does not tend to converge around A). By carefully analyzing the experimental setting, both effects can be explained by considering that, despite the learning and the evolutionary target vectors are randomly selected, the learning and the evolutionary tasks are dynamically correlated on the average (two tasks are dynamically correlated when changes that result on improvements with respect to one task also produce improvements with respect to the second task, on the average). The fact that the learning and the evolutionary task are dynamically correlated in this case, can be easily explained by considering that the two targets are selected within the central portion of the search space (between -1.0 and 1.0) while mutations tend to move individuals also outside of this portion of the space.
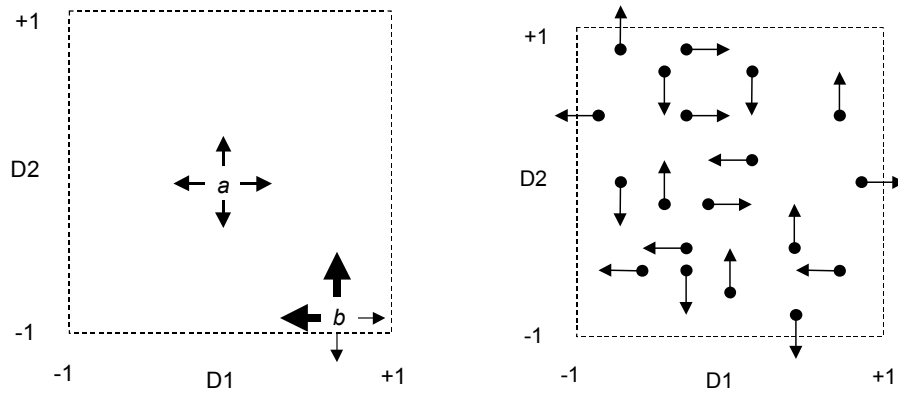
Figure 8. **Left:** The square represents the central portion of the search space (for graphical reason we displayed a 2D search space). *a* and *b* represent two individuals. The thickness of the arrows represents the probability that the two individuals move in the corresponding direction as a result of learning. **Right:** The square represents the central portion of the search space. The filled circles represent the position of the individual at birth (i.e. before learning). Arrows represent the displacement of the individuals in the search space resulting from mutations.

Consider Figure 8, left. For the individual *a* that is located in the very central portion of the square the probability to move toward north, south, east or west are the same given that the learning target might be located in each direction with the same probability. For the individual b however, the probability to move toward north and/or west is much higher than the probability to move toward the other two directions. This can be explained by considering that, the learning target is located in a randomly selected position within the central portion of the space bounded by the square. Given that the portion of the space on the north/west of individual *b* is much larger that the portion of the space on its south-west, it is much more likely that the learning target is located in the former portion of the space. Given that also the evolutionary target is located in a randomly selected position within the square, the probability that the individual *b*, by approximating that learning target, will also approximate the evolutionary target is higher than 50%. The overall picture is that, aside from the case of the individuals located in the very central portion of the space, by learning individuals will also tend to increase their performance on the evolutionary task, on the average, independently of their relative position with respect to the two targets. In other words, the evolutionary and learning surfaces are dynamically correlated overall (two surfaces are dynamically correlated when movements that produce an improvement in the learning surfaces also produce improvements on the evolutionary surface on the average, see Nolfi et al. [1994], Nolfi and Parisi [1996]). Notice that the predisposition-effect described in the previous section does not requires that the two surfaces are dynamically correlated overall but only that there are sub-areas in the search space in which the two surfaces are dynamically correlated. The predisposition-effect, in fact, postulates that evolution will tend to select such sub-areas and therefore that learning will tend to have a beneficial effect with respect to the evolutionary task even if the two surfaces are not dynamically correlated overall.

A further way to demonstrate that the increase in performance as a result of lifetime learning already in the very first generation is due to the fact that the two tasks are dynamically correlated overall, is to conduct another experiment in which the target vectors are carefully chosen so to assure that the two surfaces are not correlated overall. As we will show in the next section, in this case (as in the case of the experiments described in the previous section), the beneficial effect of learning on evolution can be observed only after some generations.

5. Learning X improves your fitness, provided that also your ancestors happened to learn X.

To study the interaction between learning and evolution in a case in which the learning and the evolutionary surfaces are not dynamically correlated, we replicated the experiments described in the previous section by utilizing a toroidal search space. Notice how this is equivalent to postulate a non toroidal and unbounded search space in which there are an infinite number of learning and evolutionary local optima that are separated of a distance of 2.0 in each dimension (i.e. an experimental setting that more naturally resemble the case of an evolving animat such as that described in section 3 in which one can assume that a large variety of solutions exist and that solutions are not necessarily confined within the central portion of the search space). In addition we selected the learning and the evolutionary targets so to assure that the learning and evolutionary surfaces are not dynamically correlated or anti-correlated overall. In particular the evolutionary and the learning targets were separated of a distance of 0.428 along each dimension.

Figure 9 shows an example of how this can be accomplished in a 2D search space. The distance between the learning and evolutionary targets (**L** and **E** respectively) along each dimension is chosen so to assure that half of the search space consist of areas in which the two surfaces are dynamically correlated (see the full circles) and the other half consist of areas in which the two surfaces are dynamically anti-correlated (see the empty circles). The former are areas from which by learning individuals increase their performance with respect to the evolutionary task, the latter are areas from which by learning individuals decrease their performance with respect to the evolutionary task. Therefore the space is dynamically uncorrelated overall. Notice that in this example, in order to maximize their fitness, learning individuals should get born with a genotype that is localized in the north-west area of **E**. Also notice that Figure 9 is a flat representation of the toroidal space. This implies that individuals might travel from opposite extremes of the square representation both as result of learning and as result of mutations.
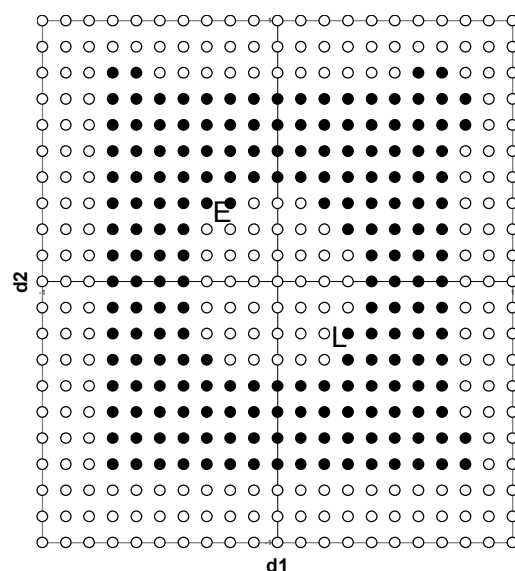

Figure 9. A 2D toroidal search space that is uncorrelated overall. **E** and **L** represent the evolutionary and learning target. Full and empty circles represent areas from which, by reducing the distance of a $\Delta = 0.1$ from **L**, individuals respectively decrease or increase their distance with respect to **E**.

In the case of the toroidal search space, as in the experiment described in section 3, performance of learning individuals with respect to the evolutionary task increase during lifetime at generation 999 but not at generation 0 (Figure 10, BC data). This effect is due to the fact that, as in the case of the experiments described in the previous section, the evolving population tend to converge toward $\mathbf{E} - \Delta \mathbf{L}$ (result not shown), i.e. toward areas in which the learning and the evolutionary surfaces are dynamically correlated. Also notice how the performance of the population at generation 0 does not vary on average as a result of mutations or as a result of lifetime learning (Figure 10, ABC data, top thick lines). This result nicely fits with the predisposition-effect hypothesis we introduced above. Learning influence evolution by pushing the evolving population toward areas of the search space that are dynamically correlated and this, in turn, assures that evolving individuals by learning also improve their performance with respect to the evolutionary task. Notice that this predisposition-effect results from the interaction of learning and evolution and cannot be explained as a form of re-learning.
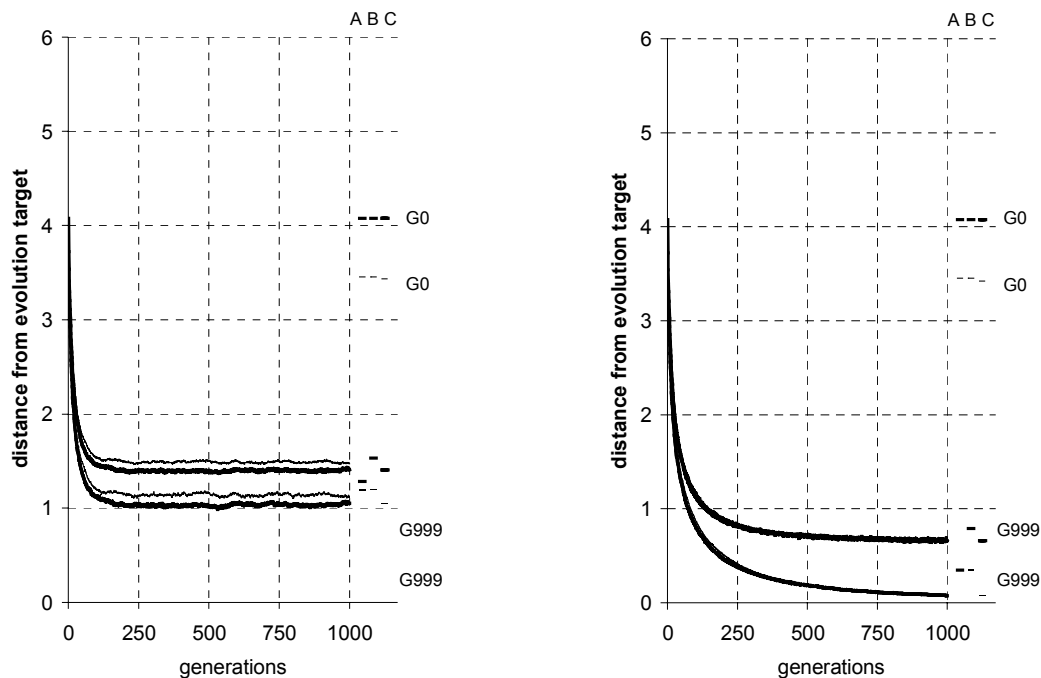

Figure 10. Experiment with the toroidal search space. Thin and thick curves represent the distance from the evolutionary target of non-learning and learning individuals respectively. Both the average performance of the population and the performance of the best individual of each generation are shown. A B and C represent the performance before mutations, after mutations, and after learning in the case of learning individuals (thick and thin lines represent the average performance of the population and of the best individual respectively). ABC data are shown for generation 0 and 999. Average results of 20 replications. **Left:** Performance in the experiments in which mutations were applied to all members of the population. **Right:** Performance in the experiments in which a copy of each reproducing individual was preserved by mutations.

The second thing to notice is that, in the case in which all 100 individuals are affected by mutations, learning individuals outperform non-learning individuals (see Figure 10, left). In the case of the experiment in which a copy of each reproducing individual is preserved by mutations, learning individuals slightly outperform non-learning individuals during the first 250 generations (see Figure 10, right). Also in this case (as in the case of the experiments

described in section 3), learning individuals start to outperform non-learning individuals after some generations.

The fact that the advantage of learning individuals is stronger in the first case in which the population receive more mutations suggest that this effect can be interpreted as a form of shelter-effect (i.e. learning allows evolving individuals to recover from negative mutations). This can be more clearly demonstrated by comparing the loss in performance produced by mutations in two hypothetical population of learning and non-learning individuals that converged to the local optimum[4]. As we demonstrated above, non-learning individuals tend to converge on **E**, and learning individuals on **E-ΔL.** As shown in Figure 11, the average distance from the evolutionary target after mutations is higher in the case of mutated copies of non-learning individuals that are distributed around **E** (see the black histogram) than in the case of mutated copies learning individuals that are distributed around **E-ΔL** (see the white histogram). Given that both learning and non-learning individuals had optimal performance before receiving mutations, we can conclude that learning allows individuals to partially recover the effect of negative mutations. Therefore, this shelter-effect is an indirect consequence of the predisposition-effect given that it require that the population is located in a region of the search space in which the learning and the evolutionary surfaces are dynamically correlated.
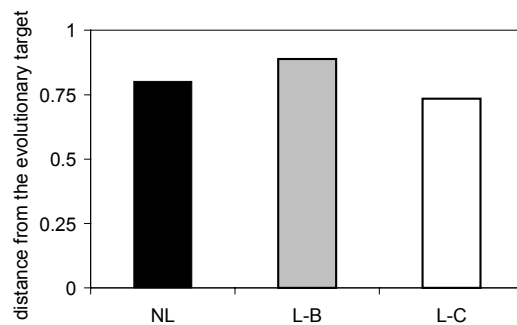

Figure 11. Average loss of performance for non-learning individuals (NL) and learning individuals (L-B and L-C) as a result of mutations. Non learning individuals are mutated copies of an hypothetical individual located on **E**. Learning individuals are mutated copies of an hypothetical individual located on **E-ΔL**. In the case of learning individuals the L-B and L-C histograms represent the distance from the evolutionary target after mutations but prior to learning and after mutations and learning respectively.

6. Discussion

Learning influence evolution in several ways and significantly impact the course of the evolutionary process.

As first demonstrated by Hinton and Nowlan (1987), learning can help and guide evolution by significantly speeding-up the evolutionary search. In particular the guide-effect of learning might favor the acquisition of useful traits during lifetime that might become genetically specified later on as a result of a genetic assimilation effect.

---

[4] Data obtained by setting **E** to all 0.0 and **L** to all 0.428 (i.e. by using a search space in which the learning and the evolutionary surfaces are dynamically uncorrelated overall). The value corresponding to the black histogram has been obtained by computing the average distance from **E** of 100,000 mutated copies of **E**. The value corresponding to the gray and white histogram has been obtained by computing the average distance from **E** of 100,000 mutated copies of **E-ΔL** before and after learning. Learning consisted in moving the mutated copies toward **L** of a Δ = 0.1). Mutations consisted in adding a random value drawn from [-1.0, 1.0] to five selected genes.

As we demonstrated in this paper, learning tend to have a beneficial effect on evolution even if the learning task and the evolutionary task are uncorrelated. This can be explained by considering that learning forces evolution to select individuals that are located in regions of the search space in which the learning and the evolutionary surfaces are dynamically correlated. This, in turns, has two implications: (1) individuals display a predisposition-effect to benefit from learning, i.e. while improving their performance with respect to the learning task during lifetime they also improve their performance with respect to the evolutionary task, and (2) individuals display a shelter-effect, i.e. lifetime learning allows them to partially recover from mutations that are deleterious with respect to the evolutionary task. This in turn might allow learning individuals to tolerate higher mutation rate than non-learning individuals thus enhancing the exploratory power of the evolutionary search.

The only preconditions that should be satisfied in order to observe the predisposition and the shelter effects are: (a) that lifetime learning has a directionality, and (b) that individuals of previous generation where exposed to the same form of learning. In short, by paraphrasing the title of a paper of Harvey and Stone (1995), one might claim that learning to play poker might increase your ability to survive if also your ancestors were poker players.

More generally these results show how learning and evolution cannot be regarded as relatively separated phenomena that can be studied independently. Learning and evolution strongly and deeply influence each other (for an example of how evolution influence learning see Parisi and Nolfi [1996]) and cooperate in ways that might appear subtle and counterintuitive at a first glance.

References

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

Carse B. and Oreland J. (2000). Learning and evolution in neural networks - dynamic correlation, relearning and thresholding. *Technical Report*, University of the West of England, U.K.

Harvey I. (1995). Relearning and evolution in neural networks, *Adaptive Behavior*, 4 (1):81-84.

Harvey I. (1997). Is there another new factor in evolution?, *Evolutionary Computation*, 4 (3):313-329.

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

Hinton G.E. and Nowlan S.J. (1987). How learning guides evolution. *Complex Systems*, 1:495-502.

Hinton G. and Plaut D. (1987). Using fast weights to deblur old memories. In: *Proceedings of the Seventh Annual Conferences of the Cognitive Science Society*, Seattle, U.S.A.

de Lamarck J.B.P.A. (1914). *Zoological Philosophy*. London: MacMillan, 1914. Relevant excerpts reprinted in Belew, R.K, Mitchell, M. (eds.) *Adaptive Individuals in Evolving Populations*. SFI Studies in the Science of Complexity, Vol. XXVI, Addison-Wesley, 1996.

Mayley G. (1997). Landscapes, learning costs, and genetic assimilation, *Evolutionary Computation*, 4 (3):213-234.

Morgan C.L. (1896) *Habit and Instinct*. London: Arnold.

Nolfi S. (1999). How learning and evolution interact: The case of a learning task which differs from the evolutionary task. *Adaptive Behavior*, 7 (2):231-236.

Nolfi S., Elman J.L. and Parisi D. (1990). Learning and Evolution in Neural Networks. *CRL Technical Report 9019*, University of Southern California, San Diego, U.S.A.

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

Osborn H.F. (1896). Ontogenetic and phylogenetic variation. *Science*, 4:786-789.

Parisi D. and Nolfi S. (1996). How learning can influence evolution within a non-Lamarckian framework. In Belew, R.K, Mitchell, M. (eds.) *Adaptive Individuals in Evolving Populations*. SFI Studies in the Science of Complexity, Vol. XXVI, Addison-Wesley.

Parisi D., Nolfi S. and Cecconi F. (1992). Learning, behavior and evolution. In: Varela, F, Bourgine, P. (Eds.), *Toward a Practice of Autonomous Systems*. Cambridge, MA: MIT Press.

Waddington C.H. (1942). Canalization of development and the inheritance of acquired characters. *Nature*, 150:563-565.

Wilson S.W. (1987). Classifier systems and the animat problem. *Machine Learning*, 2:199-228.