How Learning and Evolution Interact:
The Case of a Learning Task which Differs from the Evolutionary Task

Stefano Nolfi
Institute of Psychology, National Research Council
Viale Marx 15, 00137, Roma, Italy
Voice:++39-06-86090231, Fax:++39-06-824737
e-mail:stefano@kant.irmkant.rm.cnr.it
http://kant.irmkant.rm.cnr.it/nolfi.html

Abstract

It has been reported recently that learning has a beneficial effect on evolution even if the learning involved the acquisition of an ability which is different from the ability for which individuals were selected (Nolfi, Elman & Parisi, 1994). This effect was explained as the result of the interaction between learning and evolution. In a successive paper, however, the effect was explained as a form of recovery from weight perturbation caused by mutations (Harvey, 1996, 1997). In this paper I provide additional data that show how the effect, at least in the case considered in the paper, can only be explained as a result of the interaction between learning and evolution as originally hypothesized.

In a recent article Jeffrey Elman, Domenico Parisi, and I reported the results of a set of simulations in which neural networks that evolve (to become fitter at one task) at the population level may also learn (a different task) at the individual level (Nolfi, Elman & Parisi, 1994). In particular, individuals which were selected for their ability to find food in their environment were asked during their lifetime to learn to predict the sensory consequences of their motor actions.

Each individual animat lives in a two-dimensional grid world where 10 food tokens are randomly distributed. Each food token occupies one cell; if the animat happens to enter one of these cells, the food token is automatically "eaten" and the animat's fitness is increased.
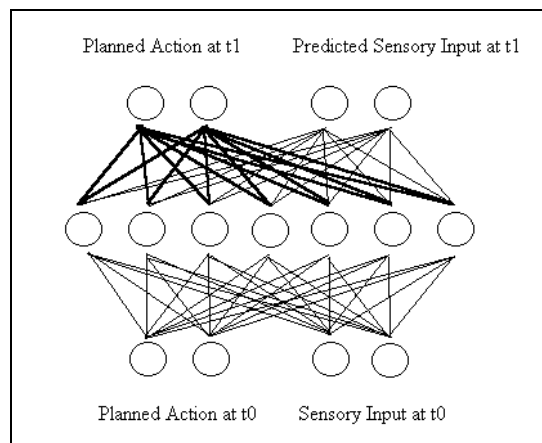

Figure 1. Neural network architecture. All connections are encoded into the genotype and inherited; however, connections represented with thin lines are also modified by prediction learning during the lifetime of the individual while connections represented with thick lines are not.

Each individual is 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. 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 (see Figure 1). 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 new 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). This 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 1 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. All inherited weight matrices are mutated by selecting 5 weights at random and modifying the weight's value by adding a quantity randomly selected in the ± 1.0 interval (the original weights are inherited, i.e. changes introduced by learning do not affect the inherited genotypes). The process is repeated for 100 generations.

As the authors showed, after a few generations, individuals learning to predict also increased their ability to find food during life (see Figure 2, thick lines).
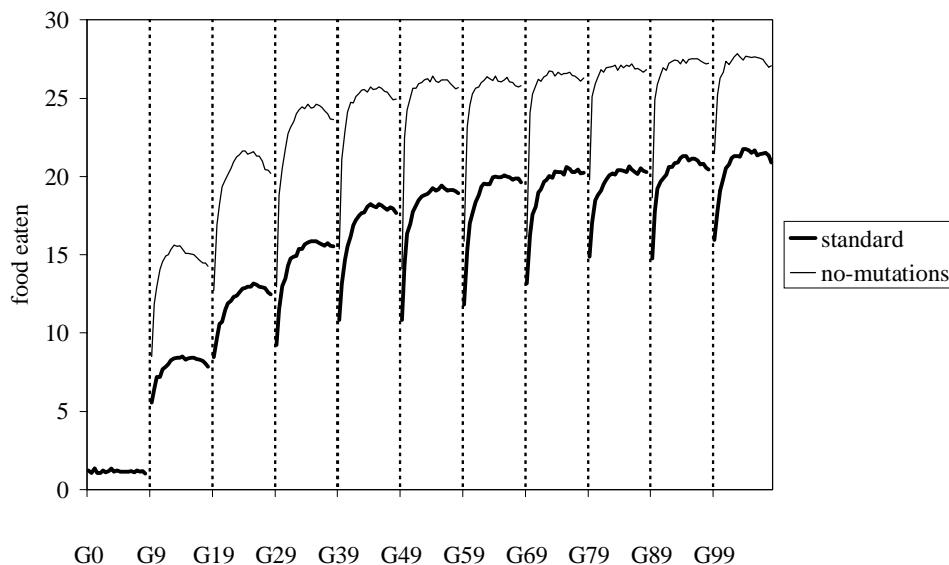

Figure 2. Average number of food elements eaten by populations of successive generations which learn to predict. Thin lines represent performances in a test condition in which mutations have been removed (results not shown for generation 0). Each curve represents performance prior to learning and then for each of the 20 epochs of life (performance prior to learning is obtained by measuring the number of food tokens eaten by individuals during an epoch of life without updating the weights). For reasons of space, performance are displayed only each 10 generations. Average results for 10 replications of the simulation.

Moreover, by comparing the results of the experiments described above with another set experiments in which individuals were not allowed to learn to predict during their lifetime, the authors showed that learning populations systematically and significantly displayed faster and higher fitness values across generations than populations without learning (see Figure 3). The same type of results were obtained in other cases and in particular in cases in which the learning task and the evolutionary task were clearly "uncorrelated" (see Parisi, Nolfi & Cecconi, 1992; Harvey, 1997).
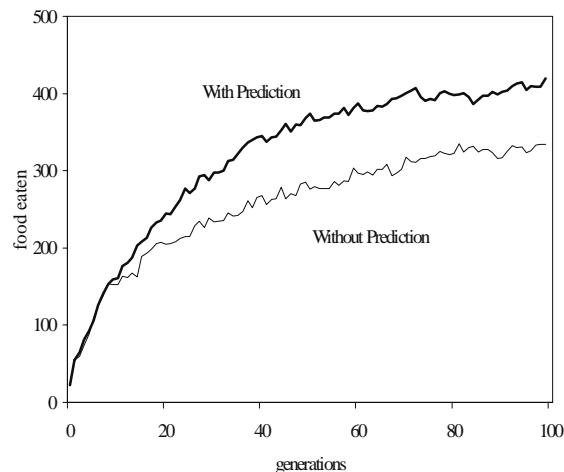

Figure 3. Average of food tokens eaten by populations of animats through out generations for the simulations with and without learning. Each curve is the average result of 10 replications. The difference between the two curves is statistically significant from generation 25 (see Nolfi, Elman, and Parisi, 1994).

In two recent articles, Harvey (1996, 1997) claimed that the improvement in average performance observed in learning individuals (cf. Figure 2 and 3) is not due to an interaction between learning and evolution but "rather from a relearning effect on weights perturbed by mutation" (Harvey, 1997, p.328). Harvey's hypothesis is based on evidence that: (a) 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 subset of the original training set, performance also improves on the other input-output pairs of the original training set not included in the retraining (Hinton, and Plaut, 1987); (b) when good performance on one task is degraded by random perturbations of the weights, it can be observed that even training on any unrelated second task can be expected to improve the performance on the first task (Harvey, and Stone, 1996). On the basis of this, Harvey made two predictions:

In summary, there are two significant ways in which this new explanation of the effects differ from that given in Nolfi et al. (1994). First, the effect is not due specifically to some interaction between learning and evolution. If one substituted for the elite member of a population evolved on the food-finding task one individual trained by back-propagation using an external teacher (or any other learning mechanism) on the same task, then one should expect similar responses after weight perturbations.
Second one should in general expect a decrease in food-finding performance in the elite evolved member (or otherwise trained individual) due to learning on the second task. The only increase in food-finding performance should be seen in those whose weights have been perturbed away from the values of the peak or elite member, by mutation or any other method.

Harvery (1996). pp.83

To verify these predictions we made two tests. First, we measured the performance of individuals of successive generations which were allowed to learn for the first time (i.e. individuals which had the same architecture of learning individuals but which evolved without being exposed to learning during lifetime). Contrary to what Harvey predicted, learning to predict produces a significant decrease in performance in these individuals (see Figure 4) even though their weights have been perturbed by mutations (exactly like the weights of individuals which were exposed to learning in previous generations).
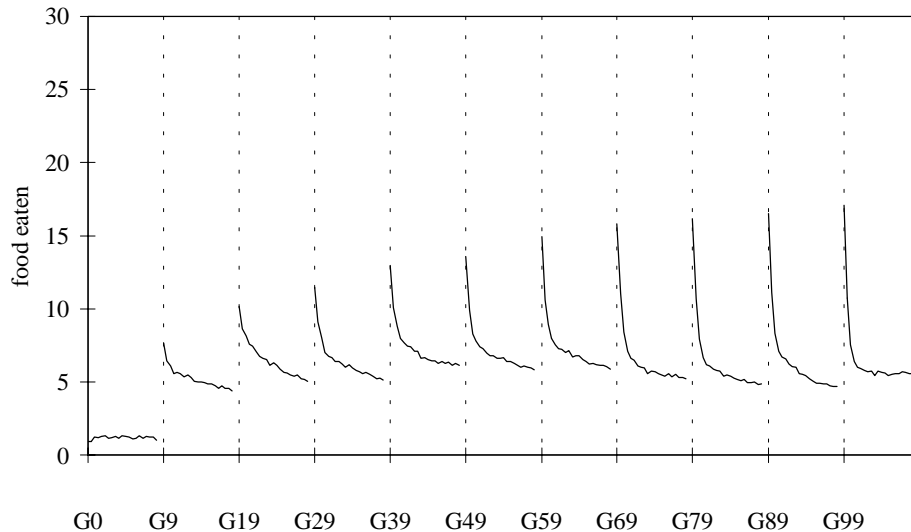

Figure 4. Average number of food elements eaten by population of individuals tested in the learning condition but not subjected to learning in the previous generations. Each curve represents performance prior learning and then for each of the 20 epochs of life (performance prior to learning is obtained by measuring the number of food elements eaten by individuals during an epoch of life without learning). For reasons of space, performance is displayed only each 10 generations. Average results of 10 replications.

These results clearly show how the beneficial effect of learning on the ability to find food tokens depends on the interaction between learning and evolution. A beneficial effect of learning, in fact, is obtained only if individuals of previous generations were subjected to learning during their lifetime. This also explains why a beneficial effect is not observed in the first generation (see Figure 2). Learning a different task by itself does not produce an increase in performance but on the contrary produces damages if performance at birth is above the chance level[1].

In a second test, we measured the performance of the learning individuals after removing the weight perturbations introduced by mutations. The test was conducted by using a set of test generations in which individuals were exact "clones" of the 20 best individuals of the preceding generation. Contrary of what Harvey predicted we observed an increase in performance even in this case (c.f. Figure 2, thin lines). After a few generations performances increase significantly during lifetime even in the absence of weight perturbations caused by mutations. The same type of results can be obtained by preserving the elite members of the

---

[1] Harvey presented a simplified model of the experiments described in this paper in which mutations and learning have respectively a negative and positive effect on the evolutionary task already in the first generation (c.f. Harvey, 1997; Figure 3). However, this seems to be due to the fact that both the learning and the evolutionary target vectors (i.e. the solutions that learning and evolution should approximate) are within the $\pm$ 1.0 region of the genotype space while mutations tend to produce genotypes which lie out of this region. In our case, there is not a single solution for each task and, in any case, solutions are not confined within the central area of the genotype space.

population intact in the next generation. Figure 5 shows the performance obtained during lifetime by replicating the experiments in this condition (i.e. in this case only four of the five offspring of each reproducing individual are affected by mutations).
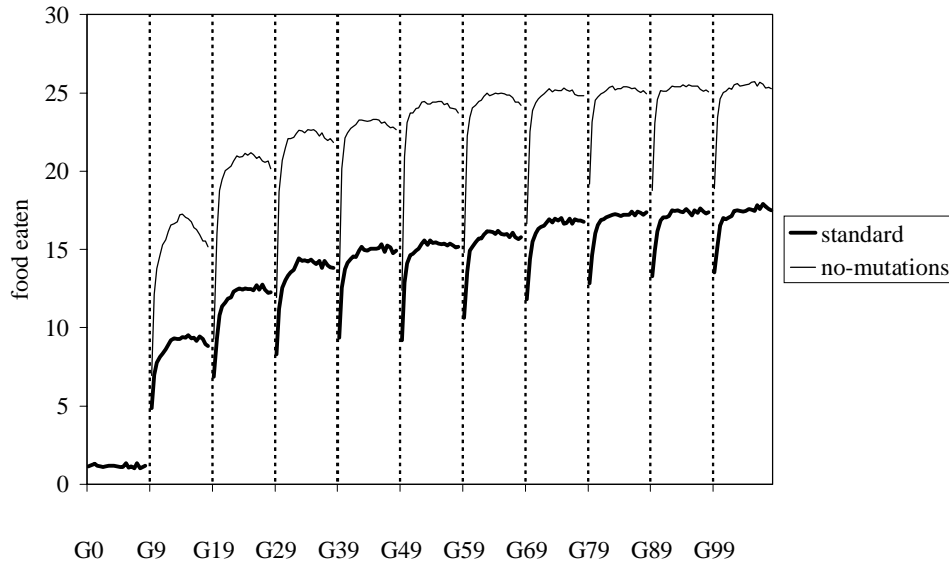

Figure 5. Average number of food elements eaten by populations of successive generations which learn to predict**.** In this case only four of the five offspring of each reproducing individual are affected by mutations. Thin lines represent performances in a test condition in which mutations have been removed (results not shown for generation 0). Each curve represents performance prior to learning and then for each of the 20 epochs of life (performance prior to learning is obtained by measuring the number of food tokens eaten by individuals during an epoch of life without updating the weights). For reasons of space, performance are displayed only each 10 generations. Average results for 10 replications of the simulation.

From these two tests we can conclude that, at least in the case of the experiments described in this paper, the positive effect of prediction learning on the eating task is the result of the interaction between learning and evolution as claimed in Nolfi, Elman and Parisi (1994) and cannot be explained as a form of re-learning as hypothesized by Harvey (1996, 1997). Indeed, the effect can be explained in the following way. Imagine two different search surfaces, an evolutionary surface and a learning surface (Figure 6). Changes due to learning produce a movement of the individual phenotype both on the learning and the evolutionary surfaces. 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 will be 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 individuals, by increasing their performance on the learning task also increase their performance on the evolutionary task (Nolfi, Elman & Parisi, 1994; Parisi & Nolfi, 1996).

Consider two individuals, *a* and *b*, which are located in two distant locations in weight space but have the same fitness at birth; i.e., the two locations correspond to the same height on the fitness surface (cf. Figure 6). However, individual *a* is located in a region in which the fitness surface and the learning surface are dynamically correlated; i.e., a region in which movements that result in an increase in height with respect to the learning surface tend to cause an increase also with respect to the fitness surface. Individual *b*, on the other hand, is located in a region in which the two surfaces are not dynamically correlated. If individual *b* moves in weight space it will go up in the learning surface but not necessarily in the fitness surface. Because of learning, the two individuals will move during their lifetime in a direction

that improves their learning performance, i.e., in a direction in which their height on the learning surface tends to increase. This implies that individual *a*, which is located in a dynamically correlated region, will end up with a higher fitness than individual *b* and, therefore, will have more chance to be selected. The final result is that evolution will have a tendency to progressively select individuals which are located in dynamically correlated regions. In other words, learning forces evolution to select individuals which improve their performance with respect to both the learning and the evolutionary task.
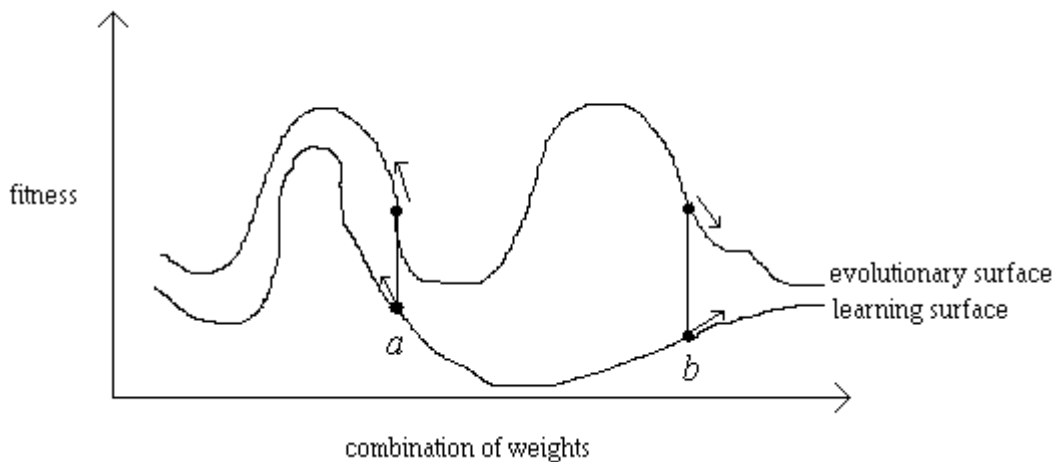

Figure 6. Fitness surface for the evolutionary task and performance surface for the learning task for all possible weights matrices. Movements due to learning are represented as arrows. Point *a* is in a region in which the two surfaces are dynamically correlated. Even if *a* and *b* have the same fitness 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*.

If learning has a cost (as in the experiments described in this paper in which the fitness is the sum of the number of food tokens eaten through out individual lifetime and therefore the time spent to learn may cost in terms of fitness) there is also an evolutionary pressure to improve performance at birth (which in fact increases through out generations, see Figure 2 and Figure 5). In the case of these experiments, however, this pressure does not cause a complete assimilation in the individuals' genotype of the characters first acquired through learning. This may be explained by considering that: (a) the cost of learning can be reduced by increasing the learning speed (as can be observed in Figures 2 and 5, in fact, evolved individuals reach optimal performance in the very first epochs of their lifetime); (b) the learning space and the evolutionary space are not completely correlated in these experiments (on the correlation between the learning space and the evolutionary space see Mayley, 1997). In fact, although learning and evolution operate on the same entities (i.e. the connection weights), the genetic operators and the learning operators are quite different (mutations are accomplished by adding a random selected values to a set of randomly selected connection weights while learning is accomplished by back-propagation). This implies that points which are close in the learning space may be far apart in the evolutionary space. In other words, points, which can be reached with few learning cycles, may require a prohibitively long list of mutations. As shown by Mayley (1997), the probability of assimilating characters first acquired through learning is inversely proportional to the correlation between the learning and the evolutionary space.

Finally, the fact that the loss in performance produced by mutations on the majority of the individuals (mutations are counter-adaptive, on the average) can be recovered only in learning individuals may explain why learning produces a larger benefit to the population on the

average than to the best individuals of each generation which are less affected by counter-adaptive mutations (data shown in Nolfi, Elman & Parisi, 1994)[2]. The abilities which are lost by the population due to mutations can be restored, at least in part, by learning (as claimed by Harvey, 1996, 1997). This, however, does not appear to be the result of a form of re-learning but a side effect of the fact that evolution selects individuals located in areas of the fitness surfaces which are dynamically correlated. The effect of re-learning, if any, may be a contribution in the same direction.

Acknowledgments

References

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

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

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

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

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

---

[2] This has the interesting consequence that plastic individual can tolerate higher mutation rate.