Institute of Psychology
C.N.R. - Rome

Learning and evolution in neural networks

Stefano Nolfi*   Jeffrey L. Elman+   Domenico Parisi*
*Institute of Psychology, National Research Council, Rome, Italy
+Department of Cognitive Science, University of California, San Diego, La Jolla.
e-mail: stefano@kant.irmkant.rm.cnr.it
elman@crl.ucsd.edu
domenico@gracco.irmkant.rm.cnr.it

June 1994

Technical Report 94-08

Department of Neural Systems and Artificial Life
15, Viale Marx
00137 - Rome - Italy
voice: 0039-6-86894596
fax: 0039-6-824737

Learning and evolution in neural networks

Stefano Nolfi*   Jeffrey L. Elman+   Domenico Parisi*

*Institute of Psychology
National Research Council, Rome, Italy
+Department of Cognitive Science
University of California, San Diego, La Jolla.
e-mail: stefano@kant.irmkant.rm.cnr.it elman@crl.ucsd.edu   domenico@gracco.irmkant.rm.cnr.it

Abstract

The paper describes simulations on populations of neural networks that both evolve at the population level and learn at the individual level. Unlike other simulations, the evolutionary task (finding food in the environment) and the learning task (predicting the next position of food on the basis of present position and planned network's movement) are different tasks. In these conditions both learning influences evolution (without Lamarckian inheritance of learned weight changes) and evolution influences learning. Average but not peak fitness has a better evolutionary growth with learning than without learning. After the initial generations individuals that learn to predict during life also improve their food finding ability during life. Furthermore, individuals which inherit an innate capacity to find food also inherit an innate predisposition to learn to predict the sensory consequences of their movements. They do not predict better at birth but they do learn to predict better than individuals of the initial generation given the same learning experience. The results are interpreted in terms of a notion of dynamic correlation between the fitness surface and the learning surface. Evolution succeeds in finding both individuals that have high fitness and individuals that although they do not have high fitness at birth end up with high fitness because of learning to predict.

1. Introduction

Most organisms both evolve at the population level and learn at the individual level. Evolutionary change occurs from one generation to the next while learning is change during the lifetime of a single individual. To study how evolution and learning may interact much research has been dedicated recently to applying genetic algorithms (Holland, 1975) to populations of neural networks that learn during life (Yao, 1993; Langton, Taylor, Farmer, & Rasmussen 1991). Neural networks reproduce selectively on the basis of some fitness criterion and offspring inherit some properties from their parent(s). Reproduction can be either sexual (two parents) or agamic (single parent) and in both cases some random mutations introduce themselves during the copying process. Furthermore, individual networks learn some task during their life. Although the changes that result from learning (most often, changes in connection weights) are not inherited, the problem is if and how learning influences the course of evolution (e.g. by channelling evolution) and, viceversa, if evolution influences learning (e.g. by selecting initial states).

In most work that examines evolution together with learning in neural networks there is no distinction between the evolutionary task and the learning task. The fitness criterion is how much an individual network learns a particular task. For example, Miller, Todd, & Hedge (1989) have examined how evolution can shape the network architecture for learning the XOR task. An initial population of networks with randomly generated network architectures and randomly generated initial weights learns the XOR task using backpropagation. The individuals that after a fixed number of learning cycles have a smaller error in the task are more likely to reproduce than other individuals. The offspring of the reproducing individuals inherit a network architecture which is a recombination of complementary parts of their two parents' architectures, plus mutations. Their connection weights are randomly generated at birth. This is repeated for a certain number of generations. Architectures change during evolution and they converge toward one or a few types that appear

to be particularly appropriate for learning the XOR task.

In simulations such as Miller et al.'s the evolutionary task, i.e. the task that determines the fitness of an individual, and the task which is learned during life, are the same task. This makes it possible to study how evolution can influence learning, in particular how evolution can create good initial conditions for learning. In some researches evolution selects good architectures for learning (Miller, Todd, & Hedge, 1989; Kitano, 1990; Nolfi & Parisi, in press). In other cases evolution can select good initial weights or good learning rates or momentums (Belew, McInerney, & Schraudolph, 1991), or even good learning rules (Chalmers, 1990). But if the evolutionary task and the learning task are the same task, it is more difficult to investigate how learning can influence evolution. Hinton & Nowlan (1987) have shown that learning can guide evolution in its exploration of the space of possible genotypes when the learning task and the evolutionary task are the same but their conception of learning is very artificial since learning is equated with random changes in the phenotype.

Notice that there is no a priori reason for assuming that whatever is learned by a given individual during its life automatically increases the individual's reproductive chances. But if the learning performance is directly interpreted as fitness, we are actually making this unjustified assumption. In order to leave the problem of how learning can be related with evolution open, it is necessary to separate the evolutionary task from the learning task. A population of networks reproduce selectively on the basis of each individual's performance on some task (evolutionary task). In addition, each individual learns some other task during its life (learning task). By itself, the performance of each individual in the learning task has nothing to do with the individual's fitness. However, since both what results from the evolutionary process and is genetically inherited and what results from learning reside in the same individual (network), we can examine if and how evolution and learning can influence each other.

In this paper we present some simulations of networks that evolve an ability to find food in their environment. The fitness that determines each individual's reproductive chances is the number of food elements captured during a fixed lifetime (evolutionary task). In addition each individual learns during its life to predict what sensory changes will result from its planned movements (learning task). Individuals reproduce on the basis of their food finding ability (performance in the evolutionary task), not on the basis of how much they learn to predict (performance in the learning task). But, as we will see, what is found is that both evolution influences learning and learning influences evolution in these circumstances.

2. The Problem

Let us begin by assuming that our ultimate goal is to create an organism (O) which is able to find food in its environment (cf. Jefferson, Collins, Cooper, Dyer, & Flowers 1991; Ackley & Littman, 1991). We imagine that O's environment is a two-dimensional grid-world. At any particular moment O occupies one of the cells. A number of food elements are randomly distributed in the environment with each food element occupying a single cell. O has a facing direction and a rudimentary sensory system that allows it to receive as input from the environment the angle (relative to where it is currently facing) and the distance of the nearest food. We shall also equip O with a simple motor system that provides it with a repertoire composed of four possible actions. O may either turn 90 degrees right, turn 90 degrees left, move to the next cell in the facing direction, or do nothing. Finally, when O happens to step on a food cell, it eats the food contained in it (the food disappears).

The O's nervous system is modeled by a feedforward neural network constisting of three layers of units (Figure 1). The input layer contains 2 nodes which encode sensory information from the environment and another 2 nodes encoding the currently planned action. These 4 nodes are fully connected to an intermediate ("hidden") layer of 7 nodes. The 7 hidden nodes are connected to 2 output units encoding a motor action. The sensory information from the environment concerns the angle (measured clockwise from O's facing direction) and the Euclidean distance of the currently nearest food element. Both values are scaled from 0.0 to 1.0. Motor actions are encoded in a binary fashion: 11=one step forward; 10=turn left; 01=turn right; 00=do nothing. The actual analog values produced by the network are thresholded to the nearest binary value.

Planned Action at t1

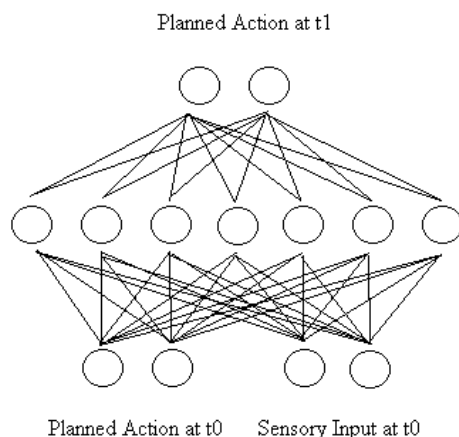Planned Action at t0   Sensory Input at t0

Figure 1. Architecture used in Simulation 1

When O is placed in the sort of environment described above, the following sequence of events will occur. Sensory input is received on two of the input nodes and a planned action (initially chosen at random) is encoded in the other two input nodes. Activation flows through the hidden units to the output nodes. The resulting activation values on the output nodes define the next motor action. At this point, the planned motor action is actually executed (O turns or moves one step forward). This action will determine the next sensory input from the environment. The motor action decided by the output nodes becomes the planned action in the next cycle. Hence, O decides its motor actions on the basis of the current sensory input from the environment (the position of the nearest food element) and its previously decided but not yet executed motor action.

We are interested in the question: How can O develop some purposeful behavior? How can O acquire the ability to seek out food in an efficient manner? (We define efficiency simply as the number of food elements eaten in a given number of actions. Since the connection weights of O are assigned at random O will move haphazardly in the environment, eating food only when it happens by chance to step on a food cell. Its behavior won't seem very purposeful and in any case it is unlikely to be very efficient.

One way of having O develop a better eating ability would be to use a connectionist learning algorithm such as backpropagation of error (Rumelhart, Hinton, and Williams, 1986). But backpropagation learning requires the correct teaching input to be provided for each input/output pair. It seems highly unlikely that nature provides most species with a "teacher" that gives organisms precise feedback at each step and leads them to correct performance. Furthermore, the optimal strategy may be unclear, especially if the environment and/or the task are complex ones. It would be preferable for the network itself to develop useful eating strategies rather than learn a particular strategy devised by us.

This can be obtained if we use natural selection in a population of networks - in a sense, if we "breed" networks. By having a population of different networks reproduce selectively on the basis of their ability to find food and by adding mutations in order to maintain a sufficient basis of variability in the population, we should see a purposeful and efficient food finding behavior emerge across generations.

3. Simulation 1

In this simulation only evolution occurs in the population of networks. There is no learning during life. As we will see, networks of successive generations are born with increasingly appropriate connection weights due to selective reproduction and mutations but these connection weights do not change in any way during life. Hence, the behavior of a particular network (i.e. how the network responds to any specific input) is identical in all phases of its life.

The initial population is composed of 100 individuals, each with the architecture of Figure 1 and a randomly assigned set of connection weights. Each individual network lives alone for a total of 5000 actions in an

environment of 10x10 cells. The environment initially contains 10 randomly distributed food elements but food is periodically reintroduced each 50 actions. 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 reproduce by generating 5 copies each of their connection weights. The inherited weight matrices are mutated by selecting 5 weights at random and perturbing the weight's value by adding a quantity randomly selected in the interval between +1.0 and -1.0. The process is continued for 100 generations.

If we look at the organisms' fitness (i.e. number of food elements eaten) throughout the 100 generations we see that individuals increasingly able to approach food elements evolve. Figure 2 shows the fitness value of the best individual and the average fitness value of the population for each of the 100 generations. Each curve represents the average result of 10 simulations starting with different randomly assigned weights.
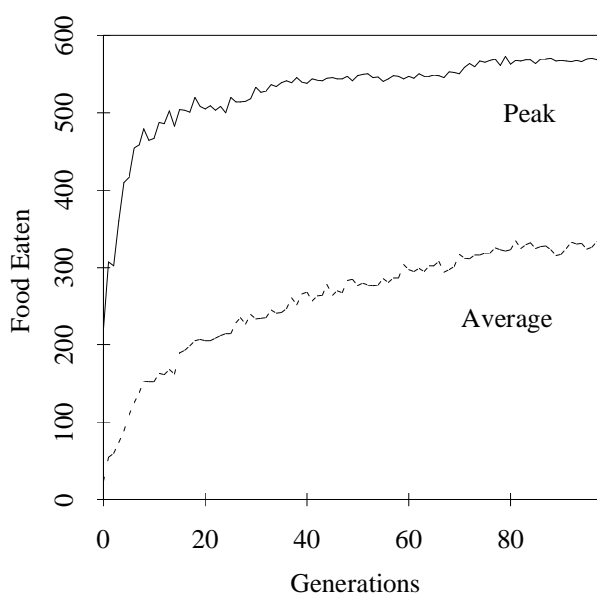

Figure 2. Average and peak food eaten (ordinate) as a function of generation (abscissa).

4. Simulation 2

In Simulation 1 an evolutionary mechanism was used to develop networks which exhibit a desired behavior. The behavior is innate in the sense that it is displayed at birth and does not require any learning or experience. But in real organisms most behavior is not determined by innate knowledge alone but it results from the interaction between innate knowledge (in the current scenario, the weight matrix at birth) and learning (changes in weights through experience). Several researchers have suggested that there is an interaction between genetically transmitted knowledge and learning (Baldwin, 1986; Waddington, 1942; Hinton & Nowlan, 1987; Belew, 1989; Belew, McInerney, & Schraudolph 1990; Parisi, Nolfi, & Cecconi, 1992). In many cases, what is learned may be only indirectly relevant to the pursuit of some endogenously determined behavior. What is learned is not the behavior itself but it may support the behavior.

We are interested in the question whether there may be some useful interaction between a genetically inherited ability to find food and some other, ostensibly unrelated, information about the environment which is learned during life.

If we want to use backpropagation of error as the learning mechanism we should ask where the necessary teaching input comes from. The backpropagation procedure requires that for each input the network is told from outside what is the expected correct output. In this way, the network can compare the actually produced output with this expected output and change the weights to reduce the discrepancy between actual and expected output. However, real organisms cannot generally be expected to have an external "teacher" that

tells them what is the correct response to each environmental input. (For a simulation in which the teaching input for backpropagation learning is generated by the same network that learns on the basis of this teaching input, cf. Nolfi & Parisi, 1993) Hence, we should choose a learning task for our networks for which the assumption of such external "teacher" is a plausible one.

One such task is learning to predict the consequences of one's own actions. The networks of Simulation 1 live in a (simulated) physical environment. (They are "ecological" networks. Cf. Parisi, Cecconi, & Nolfi, 1990.) This implies that at any given cycle the network's input is partially determined by the network's output in the preceding cycle. For example, if the current angle of the nearest food is 90 degrees (angles are measured clockwise from facing direction) and the network responds to this input by turning 90 degrees on the right, the next angle of food will be 0 degrees. On the other hand, if the network's response is a different one, for example turning 90 degrees on the left, the next input will be a food angle of 180 degrees.

We can exploit this tendency of ecological network to determine their own input by making them learn to predict what the next input will be given the current input and a planned motor action. The networks do not learn which specific motor actions are good or bad but simply that there are sensory consequences attendant upon specific movements in the context of specific environmental inputs. While this training involves supervised learning, the information required for supervision is itself available in the environment. For example, if the current food angle is 90 degrees and the network's response is turning 90 degrees on the right, we can ask the network to make a prediction regarding the next food angle just before actually executing the movement. When the movement is executed it will physically cause some consequences that will determine the next input for the network. These actually produced consequences are the teaching input. The network can compare the predicted consequences of its behavior with the actually realized consequences and learn to make better predictions in the future by changing its connection weights using the backpropagation procedure.

In order for the networks of Simulations 1 to learn this prediction task it is necessary to slightly modify their architecture so that it includes additional output units for encoding the network's predictions. As before, there are 4 input units (2 units encode the current position of the nearest food element (angle and distance) and the other 2 units encode the currently planned movement) and 7 hidden units. The output units are 4. Two units encode the motor output of the network (as in Simulation 1). The two new output units encode a prediction on what the future position (angle and distance) of the nearest food element will be given its current position and the currently planned action (cf. Figure 3.)
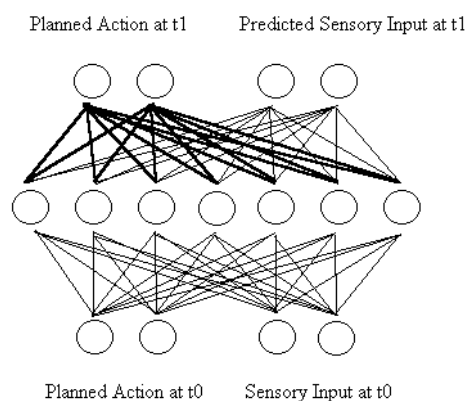

Figure 3. Architecture used in Simulation 2

At any given cycle these networks generate two outputs pertaining to two different tasks. One task it to move in the environment to find food efficiently. The other task is to make correct predictions on the sensory consequences of these movements. Notice that, given the architecture of Figure 3, the two tasks share a set of weights, i.e. the lower weights from the input units to the hidden units. But, in addition to these shared weights, the two tasks use two separate sets of weights. The food finding task uses the weights from the hidden units to the two motor output units. The prediction task uses the weights from the hidden units to the two prediction output units.

It is also important to be clear how these various weights are shaped by evolution or learning. The teaching input for learning to predict changes the weights from the hidden units to the prediction units and the shared lower weights but it leaves untouched the weights from the hidden units to the motor output units. Of course, these changes due to learning are not genetically inherited. An offspring receives at birth the weight matrix of its parent's network as it was at its parent's birth, not as this weight matrix has been modified by its parent's learning to predict during its life. (No inheritance of acquired traits.)

On the other hand, all the weights are shaped by evolution. Therefore, we might expect some influence of evolution on learning to predict. Networks reproduce on the basis of their food finding ability. As a consequence, as we will see, the weight matrices that are inherited at birth by successive generations of networks tend to reflect increasing levels of the ability to find food. This ability is neurally represented in the shared lower weights and in the weights from the hidden units to the motor output units. However, evolution has the possibility also to shape the separate set of weights from the hidden units to the prediction units because these weights too are genetically inherited together with the other weights.

Another consequence of this architecture is that learning to predict during life can have an effect on the behavior of looking for food during life. In Simulation 1 the ability to find food did not change during life. A network responded to environmental input with the same motor output in all stages of its life. The ability to find food changed (increased) from one generation to the next but it remained identical during the life of a single individual because nothing changed in the individual during its life. The situation is different in the present simulation. Since some of the weights that support the behavior of looking for food (the lower weights from the input units to the hidden units) are changed by prediction learning, these weight changes can affect the behavior of looking for food. A particular individual can respond to the same input in different ways in different stages of its life. Hence, we can also expect some influence of learning to predict on the evolution of the capacity to find food in the environment.

We have run a second set of simulations identical to first set except that (a) the network architecture is that of Figure 3, and (b) at each cycle the networks are taught to predict. In each cycle the following events occur. The network receives an encoding of the position of the nearest food and an encoding of the motor action decided in the previous cycle (planned action) as input. The network generates two outputs in response to this input. The motor output units encode a decision on the next motor action (which becomes the planned action input in the next cycle). The prediction output units encode a prediction on the next sensory input. The algorithm that controls the whole process executes the currently planned action by moving the organism in the environment and it computes the new sensory input resulting from this action. This new sensory input is used as teaching input for the prediction output units and as sensory input for the next cycle. The weights that have generated the prediction output are changed using the backpropagation procedure (with a learning rate of 0.2) and then a new cycle begins.

We will analyze the results of these simulations to examine separately the influence of learning on evolution and the influence of evolution on learning.

Figure 4 and 5 show the results regarding the evolution of the food finding ability in this population of networks. The evolutionary increase in fitness (number of food elements eaten) is compared with the corresponding increase in Simulation 1. Figure 4 shows the peak performance and Figure 5 the average performance. Each curve represents the average result of 10 different simulations with different random assignments of initial (G0) weights. As one can see, the simulations with and without prediction learning yield similar peak performance but the simulations with learning yield better average performances on the eating task even though what is learned during the life of an O is not inherited by its offspring. A two-factor analysis of variance was performed on the average performance, the two factors being the two different populations (with and without learning) and generation. The results show significant effects of population (df 1/18, f=9.8, p<0.006), generation (df 4/72, f=214.7, p < 0.001), and interaction between generation and population (df 4/72, f=3.29, p < 0.015). Post-hoc comparisons based on the Duncan test revealed significant differences between the two populations for generations 25,50,75,99 (all p <0.021) but not for generation 0.
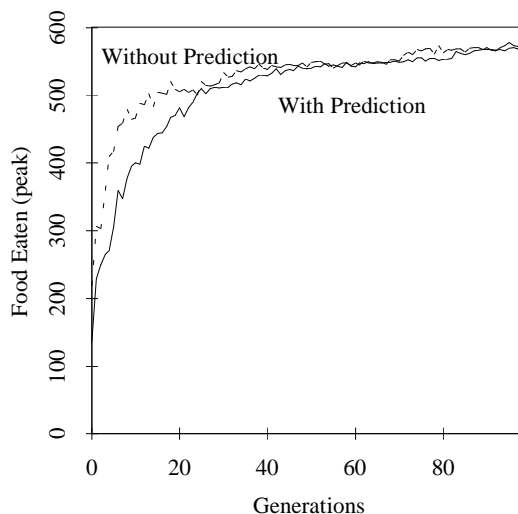
Figure 4. Peak food eaten (ordinate) as a function of generation (abscissa) in the simulations with and without learning.
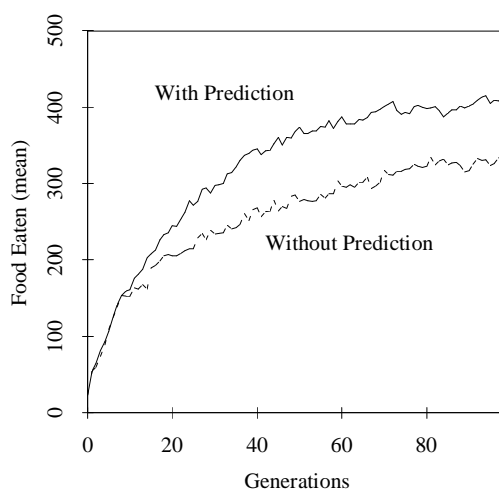

Figure 5. Average food eaten (ordinate) as a function of generation (abscissa) of the simulations with and without learning.

Another way of analyzing the possibile influence of learning to predict on the capacity to find food is to determine if and how the behavior of looking for food changes during the life of an individual while the individual is learning to predict. We know that this behavior is genetically inherited and there is no learning of this behavior during life. However, since some of the weights the generate this behavior (the lower weights) are modified during learning to predict, we might expect a possible influence of learning to predict on the behavior of looking for food. Is there such an influence? Is it a positive or a negative influence?

The answer to these questions is contained in Figure 6. This figure shows the average number of food elements eaten in each epoch of life (an epoch is a succession of 250 actions) by individuals of successive generations. Each curve represents the average performance of 10 simulations. From this figure it is clear that, after the initial generations, learning to predict during life has a beneficial effect on the ability to find food. The weight changes that take place in individual networks that are learning to predict cause the food finding behavior of these network to become more effective. A two-factor analysis of variance was performed with epoch of life and generation as the two factors. The results show significant effects of epoch of life (df=19/1881, f=172.0, p<0.001), generation (df 10/99, f=37.9, p < 0.001), and interaction between epoch of

life and generation (df=190/1881, f=3.12, p < 0.001). Post-hoc comparisons using the Duncan test revealed significant differences between the two populations for generations 9,19,29,..,99 (all p <0.001) but not for generation 0.
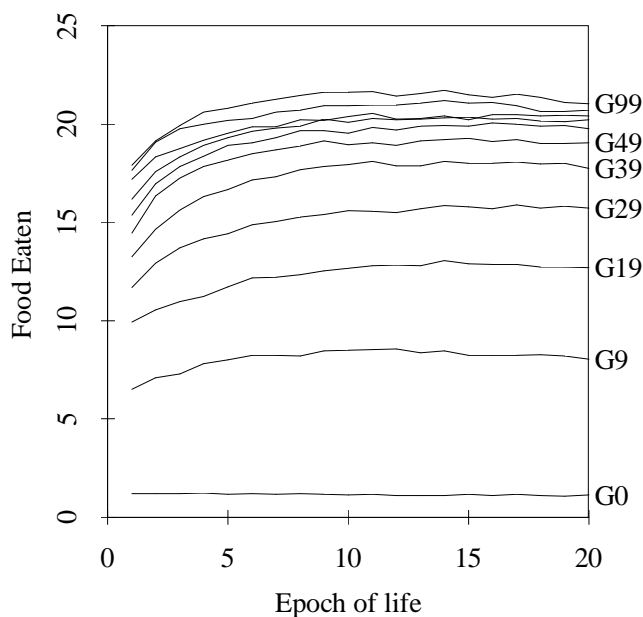

Figure 6. Average food eaten (ordinate) as a function of epoch of life (abscissa).

From these results we can conclude that learning during life to predict how the input from the world changes with Os' movements helps in evolving good eating strategies. Hence, learning influences evolution. But what about the complementary question? Does the evolution of good eating strategies also help in the task of learning to predict? Does evolution influence learning?

To obtain a direct answer to this question we can teach the Os of the first simulation (the one without learning) to predict the sensory consequences of their behavior and see if the Os of the last generation that have evolved an ability to eat food learn to predict faster than the Os of the first generation. (Notice that to teach these networks to predict we had to change their architecture from the architecture of Figure 1 to that of Figure 3.)

However, we could not use the behavior that the networks themselves generate because the Os of the first generation and those of the last generation do not have the same behavior (they execute different movements in the same context) and therefore the results of this test would not tell us much. Completely random Os such as those of the first generation usually exhibit stereotyped behavior (e.g., they tend to execute the same movement in response to all sorts of inputs), whereas the Os of the last generations have evolved more complex behaviors that allow them to eat rather efficiently. And the kind of behavior exhibited by various Os can influence how much they learn from being taught to predict the sensory consequences of this behavior.

To avoid this problem we tested the learning ability of the Os of the first and of the last generation by selecting a list of actions and making them all carry out the same actions instead of letting them move by themselves. (The actions were actually chosen randomly). The varying percentage of correct predictions during 20 epochs of training are shown in Figure 7. The predictions generated by the networks were divided into 4 classes (activation values from 0.0 to 0.25, 0.25 to 0.5, 0.5 to 0.75, and 0.75 to 1.0, of the two prediction output units) and a network's answer was considered as correct if the output value computed by the network was in the same class of the teaching input value.
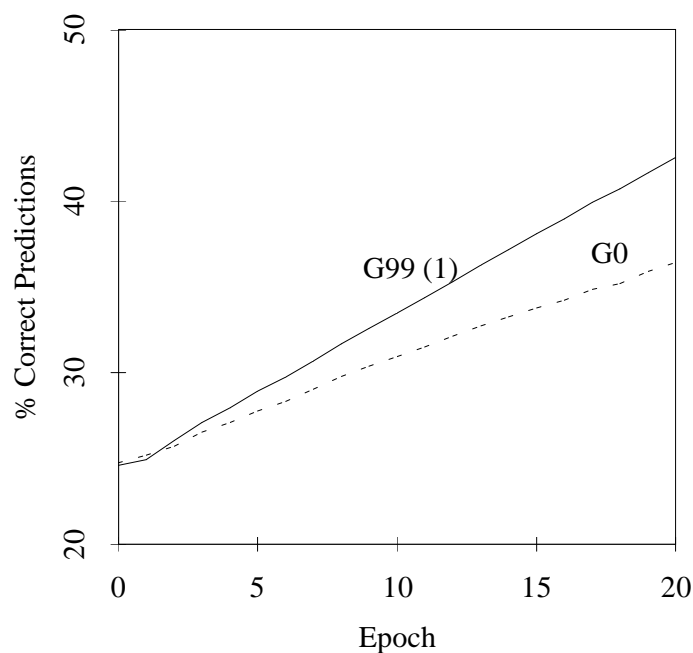
Figure 7. Percentage of correct predictions (ordinate) as a function of epoch of life (abscissa). G0 represents performance of networks with randomly selected weights (i.e. generation 0). G99 represents performance of the last generation of Simulation 1. Performance at epoch 0 were computed without modifing the the networks' connection weights.

As this figure shows, the prediction ability of the Os of the last generation increases faster as a consequence of learning than that of the Os of the first generation. This means that Os that incorporate genetically inherited eating strategies learn to predict better than those which do not incorporate such strategies. Notice however that there is no difference in learning ability at birth. What the evolved networks of the last generation inherit genetically is an innate predisposition to learn to predict, not an innate ability to predict.

Another way of investigating the question if the possession of good eating strategies helps in learning to predict is to examine the Os of our second simulation, i.e. Os that are taught to predict during their life. Even if the changes due to having learned to predict during life are not inherited, networks could learn to predict better if they have ancestors that have also learned to predict during their life. To determine if this is true we must ascertain if the Os of the last generation in our second simulation (the one with learning) learn to predict better than the Os of the first generation. However, even if we find that this the case, we don't know if this happens because the Os of the last generation have developed good eating strategies by evolution or because their ancestors have learned to predict during their life. In fact, we already know that evolving eating strategies by itself helps in learning to predict. But if we find that the improvement in learning to predict is greater than the improvement that we have observed in the simulation without life learning, we must conclude that both the evolution of good eating strategies and the life learning of the preceding generations make the successive generations learn to predict better. As is shown in Figure 8, this is what actually happens.
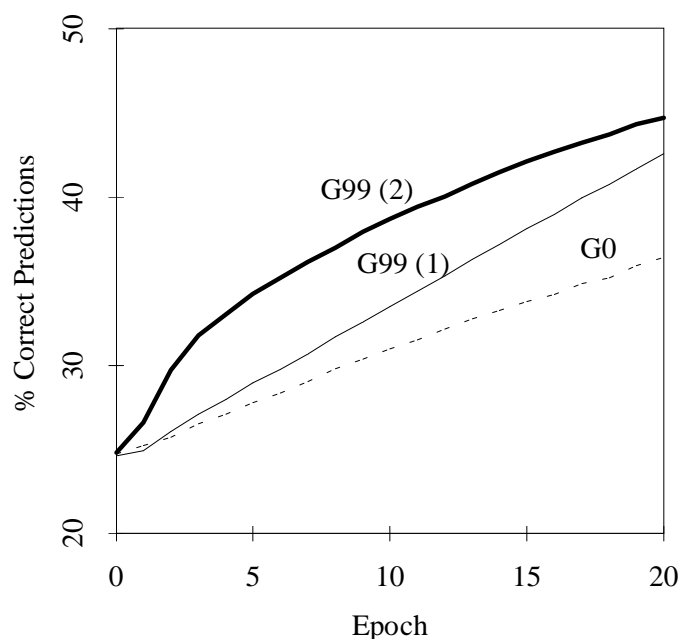
Figure 8. Percentage of correct predictions (ordinate) as a function of epoch of life (abscissa). G0 represents performance of networks with randomly selected weights (i.e. generation 0). G99(1) represents performance of the last generation of Simulation 1. G99(2) represents performance of the last generation of Simulation 2.

A two-factor analysis of variance was performed with the two factors being epoch of life and population. The results indicate significant effects of epoch of life (df 20/520, f=780.1, p<0.001), population (df 2/26, f=53.9, p < 0.001), and interaction between epoch of life and population (df 40/520, f=27.7, p < 0.001). Post-hoc comparisons based on the Duncan test show significant differences between the three different populations at the end of each epoch of life (all p <0.001).

5. Parametric sensitivity

Running simulations of the sort described above requires choosing settings for a variety of parameters that may influence what results are obtained. We carried out a number of simulations in order to understand the sensitivity of the model to different values of these parameters and to determine the robustness of our main result that learning during life has a beneficial effect on evolution.

We varied three parameters: number of mutations, population size, and number of reproducing individuals. In each simulation one parameter was varied and the other two were kept constant (at the value of the simulations described in the preceding sections). In a first set of three simulations the number of mutations was varied: 2, 5, and 8 mutations. (The intermediate value in all simulations was that of the simulations already described.) In a second set population size was varied: 64, 100, and 144 individuals per generation. In a third set the number of reproducing individuals was varied: 16, 20, and 25. We run 10 simulations with different initial random populations for each condition for both populations with and without learning.

The results in terms of the fitness of the best and the average individual of the last generation are summarized in Table 1. From these results it is clear that the beneficial effect of learning to predict on the evolution of the ability to eat is a robust one since it is found in all the conditions tested. The effect concerns the average eating ability while it is practically non-existent (with even some inversions) for the eating ability of the best individual.

Number of mutations						
	2		5		8	
	peak	average	peak	average	peak	average
no-learning	548	285	568	334	596	372
learning	512	370	572	420	547	422
Population size						
	64		100		144	
	peak	average	peak	average	peak	average
no-learning	559	316	568	334	564	324
learning	497	362	572	420	610	428
Reproducing individuals						
	16		20		25	
	peak	average	peak	average	peak	average
no-learning	563	326	568	334	558	355
learning	561	398	572	420	571	447

Table 1. Best and average fitness of the last generation in the populations with and without learning as a function of number of mutations, population size, and number of reproducing individuals.

Another result is that there is a general tendency for all measures in all conditions to increase with number of mutations, population size, and number of reproducing individuals. Again, the effect is most consistent in the case of the average fitness in the population with learning. Since an increase in the value of all three parameters can be interpreted as causing greater population variability, the beneficial effect of learning on evolution appears to be enhanced when there is more variability to play with.

6. Discussion

We have found that evolution can affect learning and learning can affect evolution even if the effects of learning (weight changes) are not genetically inherited and what is learned is an independent task with respect to the task in terms of which individuals are evaluated for their fitness. Our main results are: (a) populations that learn during life to predict the sensory consequences of their actions show a better evolutionary increase in the average ability to find food in the environment; (b) after a certain number of generations learning to predict during life has a beneficial effect on the ability to find food of the individual which is learning; (c) individuals that have evolved an ability to find food do not directly inherit an ability to predict but they do inherit a predisposition to learn to predict; this predisposition is demonstrated by better learning results when a network is exposed to the appropriate learning experiences. Moreover, individuals which belong to a population that learns during life demonstrate a better predisposition to learn at the end of evolution even if what is learned is not inherited.

To try to understand these results it may be useful to reason in terms of the weight space that contains all possible networks (more precisely, all possible weight matrices). This space is an abstract multidimensional space with each dimension representing a particular network connection and each point in the space representing a particular weight matrix. For each particular network (point in weight space) we can measure the performance of the network in the evolutionary task and in the learning task. These two tasks have independent performance surfaces: the fitness surface for the evolutionary task and the performance surface for the learning task. (Good performances in the evolutionary tasks correspond to high values on the fitness

surface and good performances in the learning task correspond to high values on the learning surface.)

Fitness-based reproduction means that individuals with higher values on the fitness surface are more likely to reproduce than individuals with lower values. However, in purely evolutionary simulations without learning (for example, in our Simulation 1) individual points do not move in weight space, although applying mutations at reproduction implies that individual points (parents) are replaced by one or more points (offspring) located near but not in the same position of their parent. (Remember that we are using agamic reproduction in these simulations.) On the other hand, learning means that the weights of the network change during lifetime and therefore individual points (weight matrices) move in weight space. If learning is successful the movement of a network which is learning will be a movement toward positions in weight space that tend to correspond to higher values on the learning surface.

Given a population which evolves and learns at the same time and has an evolutionary task which is distinct from the learning task, we can ask how the fitness surface and the learning surface are reciprocally related. (This problem of course does not arise in simulations in which the fitness of an individual is identical with its performance in the learning task and, therefore, there is a single performance surface.) In evolutionary simulations without learning the point in weight space that represents a particular individual does not move during life and, as a consequence, the individual will have more or less the same fitness in all the epochs of its life. When learning is added to evolution this may not be any longer true. The point representing the individual moves in weight space and, therefore, the fitness of the individual will be a function (in our case, the simple average) of the fitnesses of the various locations in weight space that the individual traverses during learning. It becomes crucial, then, in what direction the individual moves in weight space during life. This depends on learning and on the learning surface. Therefore, learning has the possibility to influence evolution because learning decides which successive positions in weight space are traversed by a given individual and, as a consequence, influences the particular's individual global fitness. More precisely, learning can modify the ranking of the various individuals with respect to the ranking they would have obtained without learning, i.e. by not moving from their initial position in weight space. Notice that this influence of learning on evolution may exist even if there is no genetic transmission of learned changes. 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. But learning may influence evolution by influencing the fitness of individuals on which evolution depends.

If one wishes to compare the fitness surface and the learning surface for a given evolutionary task and a given learning task, respectively, it may be useful to distinguish between a notion of static correlation and a notion of dynamic correlation between the performance surfaces of the two tasks. Two surfaces are statically (positively) correlated if a weight matrix (a point in weight space) which has a given height in one surface tends to have the same (relative) height on the other surface. On the other hand, 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. But the initial (starting) position of the weight matrix need not correspond to the same heights in both surfaces. Hence, even statically uncorrelated surfaces can contains regions of dynamic correlation. Whatever the initial heights on the two surfaces of a given network residing in these dynamically correlated regions, if the network moves in weight space it will tend to move to higher heights on both surfaces.

Let us apply these notions to our population that both evolves a food finding ability across generations and learns during life how to predict. If the fitness surface and the learning surface in our population were statically correlated, we should find that the individuals of the later generations which have a high level of food finding ability at birth also have a high level of prediction ability at birth. As Figure 7 shows, this is not the case. While these individuals genetically inherit an ability to look for food they do not inherit an ability to predict. Hence, high levels on the fitness surface do not necessarily correspond to high levels on the learning surface. The two surfaces do not appear to be statically correlated. A further proof that this is the case is contained in Figure 6. The individuals of the first generation show no increase in the ability to find food in successive epochs of their life although these individuals do learn during their life to make better predictions (cf. Figure 7, G0). This implies that, given a random sample of points in weight space (the initial population), moving these points to new positions in weight space that correspond to higher values on the learning surface does not automatically results in higher values on the fitness surface. Again, the two surfaces do not seem to be statically correlated.

However, there are indications that the two surfaces contain regions of dynamic correlation. One such indication comes from the same Figure 6. After a certain number of generations, learning to predict during life does have the effect of improving the food finding performance. This seems to imply that these individuals are located in regions of weight space that are dynamically correlated. If an individual in these regions moves in weight space toward higher values on the learning surface it automatically moves toward higher values also on the fitness surface.

Our possible interpretation of these results is that in populations that both evolve and learn but have distinct evolutionary and learning tasks evolution progressively selects for individuals that are located in regions of dynamic correlation between the fitness surface and the learning surface. In other words, evolution selects for initial points in weight space such as that when these points move because of learning their movement brings them to new locations that correspond to higher levels not only on the learning surface (this is implicit in their learning the learning task) but also on the fitness surface. This means that these points are located in regions of dynamic correlation between the two surfaces. These regions may not be very numerous globally but evolution may be able to find them.

That evolution is necessary to find them is shown by the results of Figure 7. In the first generation, when evolution has had no chance yet to operate, the movement in weight space caused by learning leads to higher values on the learning surface but not to higher values on the fitness surface. This implies that a random population of points (the initial generation) which occupies the whole weight space evenly is unlikely to be preferentially located in regions of dynamic correlation. However, after a certain number of generations, when evolution has had enough time to play some role, the situation changes. Now learning causes a movement in weight space that brings the points to higher values both on the learning surface and on the fitness surface. The evolved population tends to be preferentially located in regions of dynamic correlation.

This may imply that an effect of learning is that not only individuals which have a high fitness at birth are selected by evolution but also individuals that even if they do not have a high fitness value at birth, may end up having a high fitness value because they have learned some task during life. It is not that learning diminishes evolutionary pressures by making it less important to be born with high fitness. Learning changes the nature of these evolutionary pressures. With learning there are two, possibly alternative, evolutionary pressures. One evolutionary pressure is on being born with high fitness, the other is on being born not with high fitness but in a region of dynamic correlation such that the changes due to learning automatically translate in higher fitness.

But the two evolutionary pressures may both operate at the same time in a given population. One effect of this is that at any given time a larger number of individuals can have high fitness values in a population that both evolves and learns than it would be the case in a population that evolves but does not learn. This effect may explain the results reported in Figure 5. Evolution plus learning causes a better evolutionary increase in fitness with respect to evolution alone for the population average but not for the population's peak performance. This could be explained in the following way. Without learning only one type of individuals tend to reproduce, those that have high fitness at birth. With learning two types of individuals have good reproductive chances: those with high fitness at birth and those that although they may not have particularly high fitness at birth are located in regions of dynamic correlation between the fitness surface and the learning surface. This difference may have no effect on the peak fitness value in each generation but it has an effect on the average fitness. More individuals in each generation can have high fitness values when there are two routes instead than only one to win the evolutionary game. As a consequence, evolution plus learning affects the population average rather than the population peak values in fitness (cf. Figure 5). Williams & Bounds (1993), who made similar simulations, also obtained better average performance in the learning condition but they did not find the difference between the two conditions to be significant. This may be due to the fact that in their simulations the learning task is different and more difficult to learn. Organisms are requested to predict the next sensory input without having in input the planned action as in our simulations (see Figure 3). This may explain why learning was less effective.

This interpretation can also help us understand the results concerning the effect of evolution on the ability to predict. As we have seen (cf. Figure 8), the individuals of the later generation are not born with an ability to predict but they are born with an innate predisposition to learn to predict. At birth they do not predict any better than the individuals of the first generation but, when they are exposed to the same learning experiences of the individuals of the first generation, they learn better than those individuals. We have already interpreted

the lack of an innate ability to predict in these individuals as showing that the fitness surface and the learning surface are not statically correlated. High values on the fitness surface (those of the individuals of the later generations that know how to find food at birth) do not correspond to high values on the learning surface. However, the two surfaces contain regions of dynamic correlation. And there is evidence showing that evolution tends to select for individuals located in these particular regions because to be born in one of these regions means that the changes due to learning automatically translate into higher fitness. But if in these regions improvements in learning performance are correlated with increases in fitness it is likely that evolution will choose those regions where learning is particularly successful. In other words, the individuals of the later generation will be selected for their particular predisposition to learn to predict. As Kolen & Pollack (1990) have demonstrated, backpropagation learning is sensitive to initial conditions (initial weight matrices). Furthermore, Belew, McInerney, & Schraudolph (1991) have shown that evolution can be used to select for good initial weight matrices for learning particular tasks. We conclude that in our populations evolution succeeds in both selecting individuals that are located in dynamically correlated regions of the evolutionary and learning surfaces and in selecting initial points in these regions where learning can give better results.

6. Conclusions

If individuals learn during their life the same task they are selected for in the course of evolution learning and evolution are directly and obviously related. Evolution can help learning by creating good conditions for learning to occur (e.g. good initial weights, good network architectures, good learning rates, etc.) and learning can guide evolution by exploring approximations to the solutions sought by evolution (Hinton & Nowlan, 1987).

However, it is not clear that it is legitimate to make in general the assumption that what is learned during life automatically increases the fitness of the individuals that are learning that particular task. It is more probable that what individuals learn (the learning task) and what individuals are selected for during evolution (the evolutionary task) are more indirectly related. It then becomes an important open question what is that particular (species of) individuals learn during life. Even if we know what they are being selected for, this does not mean that we know as well what they learn.

In the present paper we have not directly tackled this more general question but we have explored some preliminary problems by doing simulations in which the evolutionary task and the learning task are kept distinct. A population of neural networks reproduces on the basis of an ability to find food in the environment and learns to predict the sensory consequences of motor actions during life. We have found that even in these conditions of a separation of the evolutionary task and the learning task, evolution and learning influence each other in complex ways.

The general explanation of this reciprocal influence has not been simply that the fitness surface and the learning surface are statically correlated, i.e. that individual networks (weight matrices) tend to have high (low) values on both the performance surface of the evolutionary task and the performance surface of the learning task. The relation between evolution and learning is more complex and it is rather captured by a notion of dynamic correlation between the two surfaces. Instead of selecting for individuals that are good both at the evolutionary task and at the learning task (there may not be no such individuals), evolution appears to select for individuals that are located in sub-regions of weight space where the changes due to learning during life tend to increase fitness.

A final comment concerns the reproduction scheme used in our simulations. We used agamic (single-parent) reproduction with mutations but no crossover. (For a sexually reproducing population which uses the same task, cf. Menczer & Parisi, 1992). The interaction between learning and evolution with distinct learning and evolutionary tasks in a sexually reproducing population should be the object of further research.

References

Ackley, D.E. & Littman, M.L. (1991). Interactions between learning and evolution. In C.G. Langton, C. Taylor, J.D. Farmer, & S. Rasmussen (Eds.), *Artificial Life II*. Reading, Mass.: Addison-Wesley.

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

Belew, R.K. (1990). Evolution, learning, and culture: computational metaphors for adaptive algorithms. *Complex Systems*, *4*, 11-49.

Belew, R.K., McInerney, J., & Schraudolph, N. (1991). Evolving networks: using the genetic algorithm with connectionist learning. In C.G. Langton, C. Taylor, J.D. Farmer, & S. Rasmussen (Eds.), *Artificial Life II*. Reading, Mass.: Addison-Wesley.

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

Kitano, H. (1990). Designing neural networks using genetic algorithms with graph generation system. *Complex Systems*, *4*, 461-476.

Kolen, J.F. & Pollack J.B. (1990). Back propagation is sensitive to initial conditions. *Complex Systems*, *4*, 269-280.

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

Holland, J.J. (1975). *Adaptation in natural and artificial systems*. Ann Arbor, Mich.: University of Michigan Press.

Yao, X., (1993). Evolutionary artificial neural networks. *International Journal of Neural Systems*, *4*, 203-222

Jefferson, D., Collins, R., Cooper, C. Dyer, M., & Flowers, M. (1991). Evolution as a theme in Artificial Life: the genesys/tracker system. In C.G. Langton, C. Taylor, J.D. Farmer, & S. Rasmussen (Eds.), *Artificial Life II*. Reading, Mass.: Addison-Wesley.

Langton, C.G.,Taylor, C., Farmer, J.D., & Rasmussen, S. (Eds.) (1991). *Artificial Life II*. Reading, Mass.: Addison-Wesley.

Menczer, F. & Parisi, D. (1992). Recombination and unsupervised learning: Effects of crossover on the genetic optimization of neural networks. *Network*, *3*, 423-442.

Miller, G.F., Todd, P.M., & Hedge, S.U. (1989). Designing neural networks using genetic algorithms. In. J.D. Schaffer *(Ed.), Proceedings of the Third International. Conference on Genetic Algorithms.* San Mateo, CA: Morgan Kaufmann.

Nolfi, S., & Parisi, D. (1993). Auto-teaching: networks that develop their own teaching input. In. J.L. Deneubourg, H. Bersini, S. Goss, G. Nicolis, & R. Dagonnier (Eds), *Proceedings of the Second European Conference on Artificial Life*. Brussels, Free University of Brussels.

Nolfi, S., & Parisi, D. (in press). "Genotypes" for Neural Networks. In M. A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks*. Cambridge, Mass.: MIT Press.

Parisi, D., Cecconi, F., & Nolfi, S. (1990). Econets: neural networks that learn in an environment. *Network*, *1*, 149-168.

Parisi, D., Nolfi, S., & Cecconi, F (1992). Learning, behavior, and evolution. In F. Varela & P. Bourgine (Eds.), *Toward a pratice of autonomous systems*. Cambridge, Mass.: MIT Press.

Rumelhart, D.E., Hinton G.E., & Williams, R.J. (1986). Learning internal representations by error propagation. In D.E. Rumelhart & J.L. McClelland (Eds.), *Parallel Distributed Processing. Vol.1: Foundations*. Cambridge, Mass.: MIT Press.

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

Williams, B.V., & Bounds, D.G. (1993). Learning and Evolution in Populations of Backprop Networks. In. J.L. Deneubourg, H. Bersini, S. Goss, G. Nicolis, & R. Dagonnier (Eds), *Proceedings of the Second European Conference on Artificial Life*. Brussels, Free University of Brussels.