
**Institute of Psychology
C.N.R. - Rome**

**Learning to adapt to changing environments in
evolving neural networks**

Stefano Nolfi Domenico Parisi

Institute of Psychology, National Research Council, Rome, Italy.

e-mail: stefano@kant.irmkant.rm.cnr.it

domenico@kant.irmkant.rm.cnr.it

November 1995 (revised September 1996)

Technical Report 95-15

Department of Neural Systems and Artificial Life
15, Viale Marx
00137 - Rome - Italy
voice: 0039-6-86090231
fax: 0039-6-824737

Learning to adapt to changing environments in evolving neural networks

Stefano Nolfi

Domenico Parisi

Institute of Psychology
National Research Council, Rome
stefano@kant.irmkant.rm.cnr.it
domenico@kant.irmkant.rm.cnr.it

Abstract

In order to study learning as an adaptive process it is necessary to take into consideration the role of evolution which is the primary adaptive process. In addition, learning should be studied in (artificial) organisms that live in an independent physical environment in such a way that the input from the environment can be at least partially controlled by the organisms' behavior. To explore these issues we used a genetic algorithm to simulate the evolution of a population of neural networks each controlling the behavior of a small mobile robot that must explore efficiently an environment surrounded by walls. Since the environment changes from one generation to the next each network must learn during its life to adapt to the particular environment it happens to be born in. We found that evolved networks incorporate a genetically inherited predisposition to learn that can be described as: (a) the presence of initial conditions that tend to canalize learning in the right directions; (b) the tendency to behave in a way that enhances the perceived differences between different environments and determines input stimuli that facilitate the learning of adaptive changes, and (c) the ability to reach desirable stable states.

1. The adaptive functions of learning in evolution

From an evolutionary point of view learning has at least three different adaptive functions (cf. Miller and Todd, 1990): it can help and guide evolution, it allows adaptation to environmental changes too fast for genetic change to be able to track them, and it makes it possible to overcome the size limitations of the genotype by exploiting the regularities of the environment. (Learning can also be the basis of cultural transmission and evolution. Cf. Boyd and Richerson, 1985.)

That learning can affect the course of evolution even if acquired characters are not inherited was first claimed by Baldwin (1896) and later elaborated by Waddington (1942). More recently, Hinton and Nowlan (1987) have provided a clear example of how learning can guide evolution using a simulation model. By considering an extreme case in which only a single combination of "genes" has fitness and all the other combinations are equally bad, the

authors show that only by adding a form of learning during life (actually random changes in the genes) evolution can discover the right combination of genes. Once the right combination is found by a particular individual, the individual will be more likely to reproduce. Even if the learned changes are not inherited, the individual's offspring will inherit a combination of genes which is more likely to be close to the right combination and therefore they also will be more likely to reproduce. On this way characters discovered through learning tend to be fixated in the genotype of individuals of successive generations. In other words, learning can provide an easier evolutionary path toward co-adapted alleles and therefore can guide and help evolution. Similar results have been reported for more complex models that do not have some of the simplifications of Hinton and Nowlan's model (Ackley and Littman, 1991; Gruau and Whitley, 1993; Nolfi et al., 1994a; cf. also Parisi and Nolfi, 1995).

That learning allows organisms to adapt to their environment is considered so obvious that

learning is often studied in isolation from natural selection, which appears to be the primary adaptive process. In any case evolution has the same function attributed to learning: adaptation to the environment. Learning supplements evolution in that it makes it possible to adapt to changes in the environment that are too fast for evolution to be able to track them. By being sensitive to environmental conditions that could not be anticipated by evolution learning can incorporate these conditions in the organism's behavior.

Finally, learning can use the regularities of the environment to build more complex phenotypes than would be possible only on the basis of the information contained in the genotype. Environmental regularities can be detected at different levels of the developing phenotype and they can affect the self-organization of the phenotype's structure and behavior. From this point of view learning should be considered as part of development, that is, of the more general process that maps the genotype into the phenotype. Development is a continuously active process which is sensitive to environmental regularities and variabilities. Learning can be defined as that part of development which is most sensitive to environmental influences while maturation tends to be the name for the part which is less sensitive to the environment and is more under the control of the genetically inherited information. (For a model of genotype-to-phenotype mapping which is sensitive to environmental influences, cf. Nolfi et al., 1994b).

2. Learning in ecological conditions

In almost all research on learning in neural networks, learning occurs in a void, or better, in an "environment" that consists of the inputs and teaching inputs arbitrarily decided by the researcher. In contrast with this, to be biologically plausible, a model of learning should consider the fact that learning is a process that arises in ecological conditions, i.e., through the interactions of the individual organism with an independent external environment (Parisi et al, 1990).

One of the most important consequences of behaving and learning in an independent environment is that the motor output of the network partially determines the network's

sensory input. By acting on the environment the individual can change either the environment itself (e.g. it can modify the position of an object in the environment) or it can change its own physical relation to the environment (e.g. by displacing its entire body or some parts of its body, the individual can move to a different environment or it can modify its perception of the environment). Thus, sensory input in ecological networks is a function of both the independent properties of the environment and the individual's behavior. It is the interaction between what the network does and the external environment that will decide which inputs are seen by the network during learning, in what order, with what frequency, etc.

Another consequence of an ecological perspective on network learning is that it becomes necessary to make it explicit how an individual network can extract from the environment the information necessary to adapt to the environment. The environment does not usually provide cues that directly indicate to the individual how it should change in order to produce more adapted behavior. Natural selection is the only source of supervision for many living systems. However, organisms appear to be able to use environmental information, made available to them through their sensors, to trigger changes that make the individual more adapted to the environment.

3. Evolving neural networks that adapt to a changing environment

In this paper we will focus on one of the three adaptive functions of learning described in Section 1: the ability to adapt to fast environmental changes that evolution alone cannot track. To study this question we designed and tested a simulation model in which a population of neural networks, representing the nervous systems of artificial creatures that behave and learn in a physical (simulated) environment, is evolved by using a form of the genetic algorithm (Holland, 1975).

3.1 Background

The problem of adapting to fast environmental changes has been addressed, using a simulation model, by Todd and Miller (1991) who set up a simulated "aquatic" environment containing two distinct patches of

plants. Each patch contains plants of two different colors. In one patch the red plants are "food" while the green plants are "poison". In the other patch the colors are reversed. Red plants are "poison" and green plants are "food". During its entire lifetime a creature lives in one of the two patches. However, the creature's offspring could be placed at birth in the alternative patch.

If a creature eats food, its fitness increases but, if it ingests poison, its fitness is decreased by a comparable amount. Creatures are immobile but since food and poisonous elements move past them, they must decide whether to ingest the particular element sensed at any given time or to ignore it. In addition, while food by itself always smells "sweet" and poison always smells "sour", turbulence in the water causes the smell (but not the color) of nearby material to be erroneously perceived with a given probability.

The behavior of Todd and Miller's creatures is controlled by a neural network with just two input (sensory) units, one for color and the other for scent, and one output (motor) unit for ingestion of nearby material. The genotype of each creature directly specifies, for each connection between the units, if the connection is excitatory or inhibitory and if it is fixed or learnable. Learning occurs via a Hebbian rule according to which correlated firing of connected units increases the strength of the connection. Because what is poisonous (or what is food) can change color from generation to generation there is no advantage in inheriting hardwired connections for poison avoidance (or food ingestion) in terms of color. However, within an individual's lifetime color serves as a cue for discriminating between food and poison which is more accurate than smell.

Todd and Miller report that in their simulations evolved creatures tend to have a hardwired (genetically specified) connection between the smell sensory unit and the eating motor unit and a learnable connection between the color sensory unit and the eating unit. The strength of this last connection is modified during the life of the individual based on which type of patch the individual ends up in at birth. In other words, in their model learning turns out to be evolutionarily adaptive. Moreover, they show that adaptiveness of learning depends critically on smell accuracy. (Smell accuracy varies in various experimental conditions.) If smell is 50% accurate (chance level), so that

food smells sweet half the time and sour the other half, then no useful information can be gained from the smell sensor. If smell accuracy is 100%, there is no need to learn because networks can efficiently rely on the smell sensor and ignore color information. For smell accuracies between 50% and 100% the authors found that the evolutionary time needed to evolve creatures that adapt during their lifetime through learning is shortest for accuracy values around 75% and it increases for both higher and lower values - so that the overall effect is an U-shaped function. This U-shaped curve emerges as the result of a trade-off between the phylogenetic pressure to evolve learning and the ontogenetic usefulness of learning. If smell accuracy is 75% learning is very useful for adaptation and, therefore, there is a strong pressure to evolve learning. In fact, in this condition learning takes less time to evolve. With smell accuracies going to either the 50% or 100% extremes, learning tends to become increasingly less useful and there is less pressure to evolve it. Therefore, the evolution of learning takes progressively more time.

3.3 Our framework

We set up a simulation in which one of the limitations of Todd and Miller's simulation is removed. In their model creatures do not displace themselves in the environment and they do not move in any other way. The creatures can only decide if they will eat or not eat the currently sensed element but they cannot influence the environment in any way. We want to study the case in which creatures can move and, as a consequence, they can partially determine their sensory input with their motor actions. Furthermore, since our creatures have the possibility to learn by extracting useful information from the environment, they can also determine, by acting on the environment, their learning experiences and the type of feedback they receive from the environment.

Our goal was to develop a creature which is able to reach a target area containing food included in its environment. The entire environment is a 60x20 cm arena surrounded by walls. The target area is a circle of 2 cm of diameter and is positioned in a randomly chosen location within the environment. The creature cannot perceive the target area but it must be able to find the target area as quickly as

possible. This implies that our creature should explore the arena efficiently in order to increase its chances to end up in the food area. At the same time the creature should avoid hitting the walls because this will cause it to get stuck into the walls thereby losing all chances to reach the target area.

We assume that the creature's body is Khepera, a small mobile robot (Mondada et al., 1993). Khepera has a cylindrical body shape with a diameter of 5.5 cm and a height of 3 cm, and it weighs 70 g. It is supported by two wheels that are controlled by two DC motors with an incremental encoder (10 pulses per mm of advancement of the robot) and that can move both forward and backward. The robot is provided with 8 infrared proximity sensors which can detect obstacles at a distance that depends on the obstacle's material and color. The 8 sensors are positioned on the periphery of Khepera's body as shown in Figure 1.

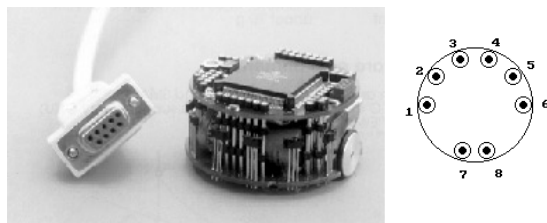


Figure 1. The picture on the left shows Khepera, the miniature mobile robot. The picture on the right shows how the 8 infrared sensors are distributed on the robot's body.

We assume that our creatures can live in two different types of environments: (a) an environment with dark walls, and (b) an environment with bright walls, i.e., walls that reflect six times more light than the dark walls. In the dark environment a sensor is activated within a distance of about 1 cm from the wall while in the light environment this distance is 6 cm. A creature should behave differently in the two environments in order to explore as much as possible of the entire arena without being stuck into the walls. If it lives in environment (a) the creature should move very carefully when sensors are activated because it starts to perceive the dark walls only in their close proximity. In contrast, if it lives in environment (b) a creature can perceive the walls from farther away and therefore it should try to avoid the walls only when the sensors are strongly activated if it wants to explore the portion of the

arena which is close to the walls. Consider however that the creatures do not know in which type of environment they are going to live. Creatures of even generations are placed in environment (a) and creatures of odd generations are placed in environment (b). The creatures living in the bright environment (b) tend to be more stimulated through their infrared sensors than the creatures living in the dark environment (a). Hence, our creatures may adapt to the particular environment in which they happen to live only if they can recognize the environment and change in the appropriate way through some form of learning.

3.2.1 The neural network

Our creatures are controlled by a feedforward neural network consisting of just an input and an output layer (no hidden units) (cf. Figure 2). The input layer includes four units that encode the activation level of Khepera's sensors. To simplify the network the first input unit encodes the average activation level of sensors 1 and 2, the second unit of sensors 3 and 4, etc. (cf. Figure 1). Hence, the network has four receptors: front, back, left, and right. These four input units are connected to four output units. The first two output units represent the two motors of Khepera and they encode the speed of the two wheels. Activation levels above 0,5 encode movements forward of a wheel and activation levels below 0,5 encode movements backward. These motor units control the robot's behavior in the environment. The remaining two output units represent two 'teaching units' that encode a teaching input for the first two output units. (For a more detailed description of the structure and functioning of this type of neural network, cf. Nolfi and Parisi, 1993.) This teaching input is used by the two motor units in order to learn using the backpropagation procedure. In other words, the neural architecture includes two distinct sub-networks that share the same input units but have separate output units. The first sub-network (standard network; cf. the thick connections in Figure 2) determines the creature's motor actions. The second sub-network (teaching network; cf. the thin connections in Figure 2) determines how the standard network changes its connection weights during life. The output of the teaching network is used by the standard network as its teaching input as part of the backpropagation procedure.

While the connection weights of the teaching network (teaching weights) do not change during a creature's life, the connection weights of the standard network (standard weights) do change based on the teaching input provided by the teaching network. Since it is the standard network that controls the creature's behavior in the environment, the behavior also changes.

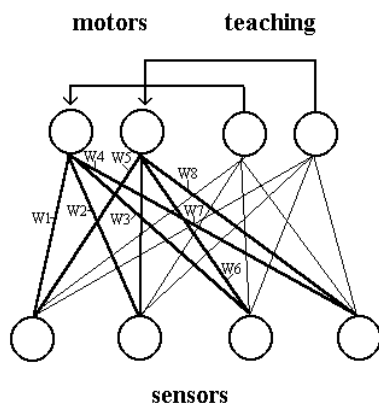


Figure 2. Self-teaching network. The output of the two teaching units is used by the two motor units as teaching input to change the weights of the connections leading to the two motor units.

When a network is placed in the environment described above the following sequence of events will occur. Sensory input is received on the input units. Activation flows up reaching the two motor units and the two teaching units. The activation value of the two motor units is used to move Khepera, thereby changing the sensory input for the next cycle. The activation value of the two teaching units (teaching input) is used to change the weights that connect the input units to the motor units according to the backpropagation procedure (Rumelhart et al., 1986). Then the next cycle begins.

3.2.2 The genetic algorithm

To evolve creatures that are able to reach the target area efficiently we used a genetic algorithm. We begin with 100 randomly generated genotypes each yielding a network with a different set of weights for the standard and teaching sub-networks of 100 creatures. Network architecture and learning rate are fixed and identical for all individuals (although they might have been part of what evolves in the population. Cf. Belew et al., 1989; Kitano, 1990.) This is Generation 0 (G0). G0 networks are allowed to live for 10 epochs, each epoch consisting of 500 input/output cycles. At the

beginning of each epoch both the creature and the target area are randomly placed inside the arena. At the end of life the 20 individuals that have accumulated the most fitness are allowed to reproduce (unisexually) by generating five copies of their genotype. The 100 new creatures constitute the next generation (G1). During the copying process 10% of the weights are mutated by adding a quantity randomly selected in the interval between -1.0 and +1.0 to the weight's current value. The process is repeated for 1000 generations.

In each epoch the fitness of an individual is increased by $500 - N$ units where N is the number of cycles needed to reach the target area in that epoch. In other words, individuals with high fitness are individuals that are able to reach the target area more quickly. The total fitness of an individual is the sum of its fitnesses in the 10 epochs of its life. If in one epoch the individual is unable to reach the target there is no fitness increase for that epoch. If in one epoch an individual happens to hit the wall the epoch is terminated. Therefore, individuals with high fitness tend to be individuals that are able to avoid hitting the wall (at least prior to reaching the target area).

3.2.3 Adaptation to different environments

To test if the model is able to evolve creatures that adapt during their life to the particular environment in which they happen to live we exposed creatures of different generations to different environments. As already said, creatures of even generations lived in an arena with dark walls while creatures of odd generations lived in an arena with bright walls. The creatures had to be able to recognize the particular environment in which they happened to live and to change their behavior (learn) in order to make it more adapted to the particular environment (It must be noticed that learning is only one way to adapt to a non-stationary environment. There may be other ways, e.g., through sensory adaptation or some form of memory.)

The way in which our creatures may adapt to different environments during their life becomes clear if one considers that the output of the teaching network, which functions as teaching input for the standard network, depends on two factors: the connection weights of the teaching network and the activation value of the four

input (sensory) units. While the connection weights of the teaching network are genetically inherited and are completely uninfluenced by the current environment, the sensory input does reflect the external environment. As a consequence, the teaching input generated by the teaching network may be influenced by the external environment and it can teach different things in different environments. Evolution has the possibility to select creatures that are able to adapt to changing environments by selecting teaching weights that produce teaching inputs that are different in different environments and that teach behaviors that are appropriate to the particular environment. More specifically, it should select teaching weights that induce modifications in the standard weights leading to environmentally adapted behavior. In addition, it should be noted that although the backpropagation procedure will try to minimize the discrepancy between the output of the standard network and the output of the teaching network, the two networks' outputs should not necessarily converge (Nolfi and Parisi, 1994). (We will come back to this point in Section 4.3.)

4. Experimental results

We ran two sets of simulations. In one set we allowed individual networks to learn during life. In the other set learning was not allowed. The individuals of the population that learned had the network architecture that has been described in Section 3.2.1, which included a standard and a teaching network. The individuals of the population that did not learn had a simplified network architecture that included only the standard network: four input units encoding the activation level of the eight sensors and two output units encoding the movement of the two wheels. The lifespan was identical in the two populations. All the 16 connection weights plus 4 biases of the population that did learn and the 8 connection weights plus 2 biases of the population that did not learn were subject to evolution. The mutation rate was 10% in both populations.

In this Section (a) we will show that natural selection succeeds in evolving individuals that can learn to adapt to the particular environment in which they happen to live; (b) we will try to analyze how individuals that are allowed to learn are able to discriminate between the two environments; (c) we will discuss the

relationship between the evolutionary process and the learning process and the role of a creature's interaction with the environment during learning.

4.1 Evolution can select for individuals that learn to adapt to the particular environment

The first thing we did was to compare the results of the simulations with and without learning. As Figure 3 shows, individuals increase evolutionarily their ability to efficiently explore the environment in both the learning and the nonlearning condition. They evolve movement strategies that allow them to visit more and more different parts of the environment and to avoid hitting the walls. This allows them to reduce generation after generation the time taken in each epoch of life to discover the target area. However, individuals that are allowed to learn during their life perform better than individuals that do not learn, although the number of input/output cycles is identical in the lifetime of both types of individuals.

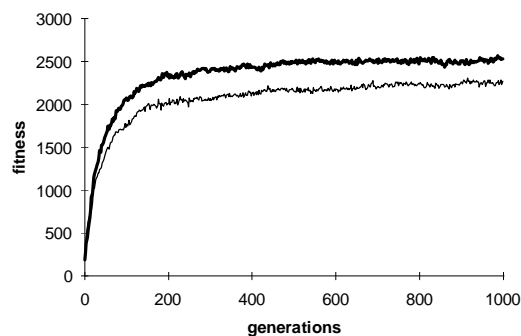


Figure 3. Increase in the fitness of the single best individual of 1,000 successive generations for the population with learning during life (black curve) and for the population without learning (grey curve). Each curve represents the average of 10 replications.

This last result implies that learning has an adaptive function for those creatures that are allowed to learn. To understand how learning can have this adaptive function we can directly inspect the behavior of individuals that are allowed to learn and compare it with the behavior of individuals that do not learn. Figure 4 shows the trace left on the terrain by the movements of a typical individual that learns and by the movements of a typical individual that does not learn, in both the dark and the bright environments. The behavior of both

individuals is rather stereotypical and nonoptimal in the sense that both individuals fail to visit some portion of the environment that may contains the target area. However, the behavior of the learning individual appears to be more efficient than that of the nonlearning individual. The nonlearning individual shown on the top row of Figure 4 is able to avoid hitting the walls in both the dark (left) and the bright (right) environment but this individual is unable to explore the portion of the environment which is close to the walls in the bright environment. Therefore, it can miss the target area in the bright environment if the target area happens to be located near the wall. On the contrary, the individual that learns, shown in the bottom row of Figure 4, is able to travel at about the same close distance from the walls in both the dark and the bright environment.

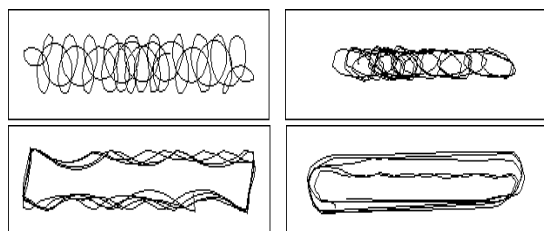


Figure 4. Behavior of two typical evolved individuals of the nonlearning (top row) and learning (bottom row) populations in the dark (left side) and bright (right side) environments.

The reason why the nonlearning individual is unable to visit the zone of the environment near the walls is that this individual ignores in which of the two environments it happens to be born in. It inherits a behavioral strategy that is somewhat adapted to both environments but that prevents the individual to get too close to the wall in the bright environment because the same sensory input in the dark environment would mean risking to hit the wall. (Remember that a strong activation of the infrared sensors means proximity to the wall.) On the contrary, the individual that learns can discover through learning in what environment it happens to live. The inherited teaching weights generate different teaching inputs for the standard network based on the different sensory inputs sent to the network by the two different environments. Since this individual learns that the same level of activation of the sensors means different distances from the wall in the two environments, the individual can generate different behaviors

in the two environments in response to the same input and it can therefore adapt its behavior to the particular environment.

In order to verify if the creatures that learn were actually able to learn to adapt to the particular environment they happen to be born in, we tested 'adult' creatures (that is, creatures at the end of their life and, therefore, of their learning) both in the environment in which they had developed and in an environment which was different from the one they experienced during their life. We made two copies of the weights inherited by the best individual in each generation and we left one copy live and learn in the bright environment and the other copy in the dark environment. At the end of life (learning) the two resulting networks were tested, with their weights frozen, in the opposite environment to that in which they had lived and learned. The results are shown in Figure 5. Individuals perform better (i.e., they obtain more fitness) if the environment in which they are tested is the same environment in which they have lived and learned. This shows that characters acquired through learning are adapted to the particular environment in which the learning takes place. (Similar results have been obtained with phenotypic plasticity, i.e., with inherited genotypes that map into the phenotypic neural networks in a way which is sensitive to the particular environment in which the mapping takes place. Cf. Nolfi et al., 1994b).

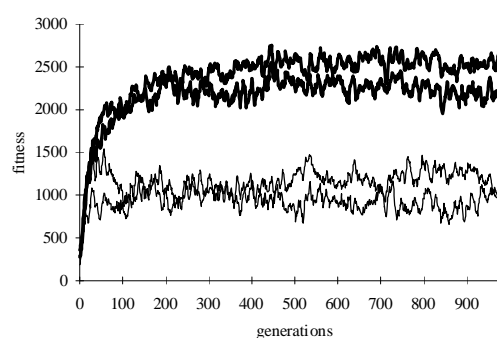


Figure 5. Performance of creatures that have lived and learned in either a dark or a bright environment and are then tested in the same or different environment. The black curves represent the performance of individuals that are tested in the same environment in which they have lived and learned. The grey curves represent the performance of individuals that are tested in a different environment.

4.2 How individuals are able to discriminate between the two environments

In the preceding Section we have seen that evolution is able to evolve individuals that learn to adapt to the particular environment in which they happen to live. Our next question is: How is such an adaptation to the current environment actually accomplished? How can individual networks 'recognize' the type of environment they happen to be born in and how can they modify themselves to adapt to that environment?

If we examine the type of stimuli that the same creature (i.e., the two identical copies of the best individual of each generation) experiences in the dark and in the bright environment we see that these stimuli differ in the two environments both quantitatively and qualitatively. We measured the activation level of the sensors during the entire lifetime of the best individuals of each generation and we discovered that the average activation level was 0.11 for the copy living in the dark environment and 0.24 for the copy living in the bright environment. In addition, we found that the percentage of times each of the four input units (corresponding to the left, right, front, and back pairs of sensors) is the most activated one of all units varies significantly at birth, i.e., prior to learning, between the two environments (see Figure 6). The measurement is obtained by allowing an individual to live for 1 epoch prior to learning in the two environments and measuring the percentage of times each of the four sensors is the most activated one.

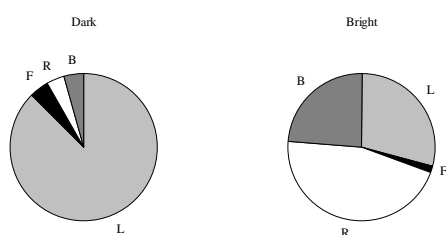


Figure 6. Percentage of time each of the four input units is the most activated one at birth (i.e., prior to learning) in one evolved individual of the simulations with learning in the dark and bright environments. (F=front sensor; B=back sensor; L=left sensor; R=right sensor)

The different types of stimuli the creatures experience in the two environments, by affecting the type of teaching input computed by the teaching network at each time step, allow the creatures to modify their standard weights (i.e.,

the weights that determine their motor behavior) differently in the two environments. Figure 7 shows how the weights of the standard network of a typical evolved individual change in the two environments. The thin curves and the thick curves, that represent the change in weight value in the dark and bright environment, respectively, diverge in many cases. More specifically, in the case of the individual shown, the weights do not change very much when the individual lives and learns in the dark environment whereas they tend to change much more when the same individual lives and learns in the bright environment.

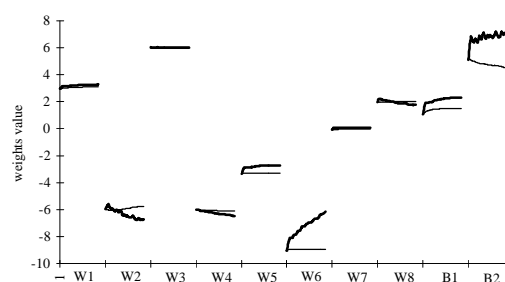


Figure 7. Change in value of the 10 connection weights of an evolved creature that lives and learns either in a dark (thin curves) or in a bright environment (thick curves). (W1-W8 = weights of the input/output connections; B1 and B2 = biases of the two output units)

Figure 7 shows that the changes in weight value are restricted to only a minority of the connection weights. However, these changes are sufficient to determine significant qualitative differences in the behavior of the learning individuals across successive epochs of their life (see below).

4.3 How learning and evolution interact

In the population without learning the inherited standard weights of an evolved individual incorporate an ability to explore the environment efficiently, that is, in such a way that the target area is discovered reasonably quickly in most epochs of life. In other words, an evolved individual is born with a general solution to the problem inscribed in its genes. This general solution is not optimal because it cannot take into account the characteristics of the particular environment in which the individual happens to be born. However, it

allows the individual to perform reasonably well (cf. Figure 3).

What is the role of the inherited standard weights in the case of individuals that are allowed to learn during their life? Are we to expect that the standard weights incorporate the same general solution and that learning is able to refine the inherited strategy by taking into consideration the specificity of the current environment?

If we compare the performance exhibited prior to learning by evolved individuals belonging to the learning population with the performance of individuals belonging to the nonlearning population, we discover that this is not the case (Figure 8). When tested for 10 epochs without any learning, individuals belonging to the learning population perform on the basis of their inherited standard weights less well than individuals of the nonlearning population. (This result is also obtained with evolved self-teaching networks living in a nonchanging environment. Cf. Nolfi and Parisi, 1993). This contrasts with the results of the comparison between the two populations when performance is assessed after learning. In these circumstances the individuals of the learning population outperform those of the nonlearning population (cf. Figure 3).

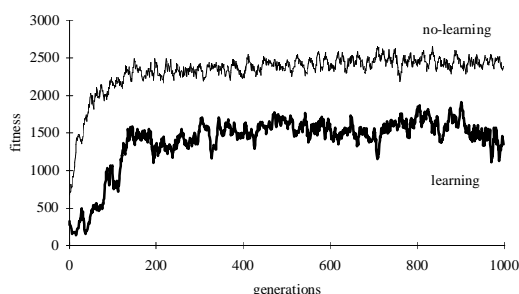


Figure 8. Performance of learning (thick curve) and nonlearning (thin curve) individuals at birth across 1000 generations. The performance of learning individuals has been assessed by letting these individuals live for 10 epochs without any learning. Average of 10 replications.

This result seems to imply that the inherited standard weights of the learning individuals are selected not only in order to allow a good performance in the task (as shown by their performance at birth prior to learning) but also in order to allow learning to produce a good performance. In other words, the genes (i.e., the inherited standard weights plus the inherited teaching weights) of evolved individuals that are

allowed to learn incorporate not a predisposition to behave efficiently but a *predisposition to learn* to behave efficiently. In the next Section we will analyze what a genetically inherited predisposition to learn can mean in the context of our simulations.

5. Genetically inherited predispositions to learn

In order to understand what a predisposition to learn can mean in the case of our creatures we should consider that initial conditions (e.g., initial weights) can determine the course of backpropagation learning (Kolen and Pollack, 1990) and evolution can select for good initial weights for learning during life in nonecological neural networks (Belew et al., 1991). We already know that the initial (inherited) standard weights of our learning creatures incorporate a partially valid innate solution to the evolutionary task since individuals tested at birth prior to learning exhibit some limited ability to find the target area efficiently (cf. Figure 8). However, by conducting some special tests we can show that both the genetically inherited standard weights and the genetically inherited teaching incorporate an innate predisposition to learn the task.

To show that the initial standard weights do incorporate an innate predisposition to learn the task we erase the inherited standard weights and replace them with random values. If we allow our creatures to learn based not on the inherited standard weights but on random initial weights their performance will remain constantly low throughout their life. Although the learning error will progressively decrease, they won't learn anything that improves the efficiency of their exploration of the environment even if the inherited teaching weights that we know can teach efficient behaviors are left intact.

A predisposition to learn to explore more efficiently the environment, therefore, is at least in part incorporated in the inherited standard weights. However, the inherited teaching weights also incorporate a predisposition to learn (or, more precisely, to self-teach). If we allow our creatures to learn based on the genetically inherited standard weights but we randomize the teaching weights, in this case too learning will destroy whatever ability to explore is present at birth rather than increasing that ability.

We conclude that both the standard weights and the teaching weights incorporate a genetically inherited predisposition to learn and that the two sets of weights must co-evolve generation after generation.

However, learning in our ecological networks is different from learning in nonecological networks because ecological networks interact with an independent environment and, therefore, they at least in part control their own input. More specifically, the inputs experienced by an ecological network during learning - what can be called its learning experience - are at least in part a function of the creature's behavior. We can then hypothesize that in ecological networks a predisposition to learn can also mean that the network starts learning with a tendency to behave in such a way that the network is more likely to experience inputs useful for learning than other inputs.

That the standard weights are selected in order to control the type of stimuli a creature experiences during its life has already been shown in previous simulations with a similar framework (self-teaching networks) but with a nonchanging environment (Nolfi and Parisi, 1994). In those simulations it was found that evolved standard weights were able to expose the creatures to sequences of input stimuli that facilitated their adaptation to the environment through learning. (For the role of the inputs experienced during learning in nonecological networks, cf. Plunkett and Marchmann, 1991). In the present simulations learning individuals appear to behave at birth in such a way that they perceive enhanced differences between the two environments with respect to nonlearning individuals. To determine how the two environments differ in the inputs that they make available to the learning and nonlearning organisms, we calculated the percentage of cycles in which each of the four input unit was the most activated one of the four units and we compared these percentages for the two environments for both learning and nonlearning individuals. Since in ecological networks the environmental input is influenced by the network's motor output any discrepancy in how differently the two environments are perceived by learning and nonlearning organisms appears to be entirely due to the different behavior exhibited by the two types of organisms. The differences in the activation level among the four input units in the two environments reflect

the different behaviors of an organism in the two environments. For example, if the left input unit is the most activated one this means that the organism behaves in such a way that it tends to have the wall near its left side. If an individual behaves in such a way that it tends to have the wall near its left side in one environment and near its right side in the other environment, one can say that with its behavior the individual enhances the differences between the inputs perceived in the two environments.

The first column of Figure 9 shows the average difference between the stimuli perceived at birth in the two environments by nonlearning individuals. The average is computed based on the (single) best individual in each of the 1000 generations for 10 replications of the simulation. A value of 0 means that each individual sensor has the same probability to be the most activated one in the two environments while a value of 200 means that each of the four sensors has an opposite probability to be the most activated one. (For example, the left sensor is always the most activated one in one environment and it is never the most activated one in the other environment.) The second column shows the same average difference for the learning individuals at birth, i.e., before any learning. The third column shows the average difference for the learning individuals at the end of life, that is, after learning has had its effect. The Figure indicates that (a) learning individuals perceive at birth the two environments as more different than the nonlearning individuals, and (b) that there may be a tendency for the differences to decrease after learning (although the difference between the second and third column is not statistically significant).

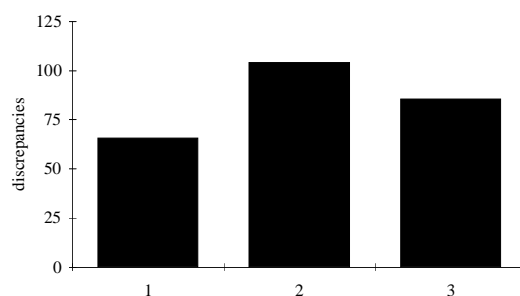
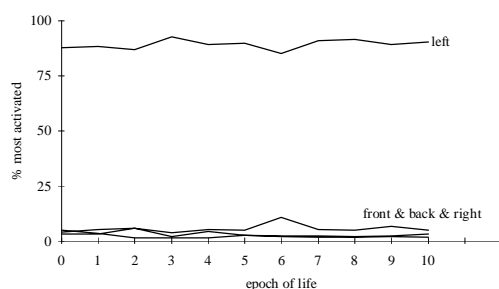


Figure 9. Difference in the percentage of time each of the four input units is the most activated one in the two environments for (1) nonlearning individuals at birth, (2) learning individuals at birth, and (3) learning individuals at the end of

their life. Average results for the best individuals of each generation in 10 replications of the simulation. Differences are measured by allowing an individual to live for 1 epoch prior to learning in the two environments, by measuring the percentage of times each of the four sensors is the most activated, and by summing the differences between the two cases. One-factor analysis of variance revealed that Condition 1 significantly differs from Condition 2 ($df=1/18$, $f=6.842$, $p<0.02$) while Condition 1 and 3 and Condition 2 and 3, respectively, do not significantly differ from one another ($df=1/18$, $f=1.145$, $p=0.29$; $df=1/18$, $f=.992$, $p=0.33$).

We conclude tentatively that the standard weights are selected in the learning individuals to generate behaviors that enhance the perceived differences between the dark and the bright environments in order to allow learning to



produce different adaptive changes in the two environments.

The analysis of the learning population is further complicated by the fact that the input stimuli may change qualitatively during an individual's lifetime because learning modifies the standard weights that determine the behavior of the individual. Figure 10 shows that in a typical evolved individual of the learning population stimuli tend to remain more constant during the course of the individual's life in the dark environment while they tend to change if the individual lives in the bright environment. This leads to the conclusion that also the teaching weights, by determining the effects of learning in the standard weights, have a role in enhancing the perceived differences between the two environments.

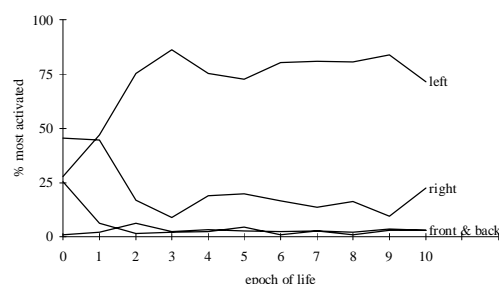


Figure 10. Change in stimuli experienced in successive epochs of life by a typical evolved individual (same individual of Figures 6 and 7). The two graphs show the percentage of times each of the four sensors (front, back, left, right) is the most activated one during life in the dark (left graph) and bright (right graph) environment, respectively.

A further aspect of predispositions to learn in our creatures emerges if one reflects that while learning in classical neural networks has the only goal to reduce the discrepancy between the network's computed output and the teaching input, this is not so in our self-teaching networks. All learning aims at reaching a satisfactory behavior as soon as possible and to maintain that behavior afterwards, with no further weight changes. The adaptive meaning of this for ecological networks is that individuals that have reached a good stable behavior because of learning can accumulate more fitness during their life. Therefore, we should expect that in our creatures both the standard and the teaching weights should be selected for their ability to generate good stable behaviors as early as possible. That stable states are quickly reached by our learning creatures can be shown in various ways. Figure 7 directly

shows that many weights change significantly early in life and then they remain stable for the rest of life. Figure 10 shows that the distribution of stimuli experienced by the individual in any case stabilizes after the third epoch of life. Finally, Figure 11 shows how performance increases very rapidly early in life and then it remains more or less stable.

It is important to notice, however, that in ecological networks a good stable behavior produced by backpropagation learning is not the same thing as zero (or almost zero) error, as is the case for learning in nonecological networks. In ecological networks learning performance, as measured by the discrepancy between computed output and teaching input (error), and fitness are conceptually and operationally distinct entities. In our simulations, although backpropagation tries to minimize the discrepancy between the output of the standard network and the output of

the teaching network, the two outputs never converge. If they were to converge, the teaching network could be used as the network that directly controls the individual's behavior in the environment, i.e., as the individual's standard network, with no need for learning. But, as we saw, a single network controlling the behavior of the organism without learning performs less well than our learning individuals with both a standard network and a teaching network. What the teaching network does is to induce changes in the standard network that are adaptive with respect to the two different environments without directly teaching the appropriate behaviors.

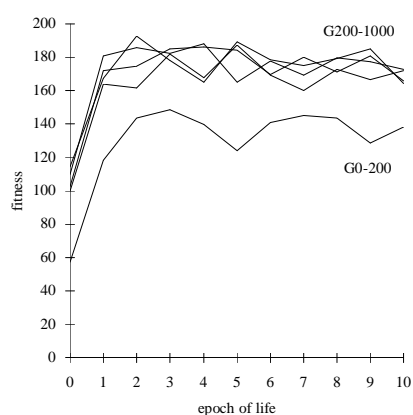


Figure 11. Performance (fitness) of the learning individuals in the 10 successive epochs of their life. Each curve represents the average result of 200 successive generations in 10 replications of the simulation. Performance at epoch 0 was calculated by measuring the fitness of the individual at the end of an entire epoch (500 cycles) without learning.

5. Discussion

To study the ability to adapt to fast evolutionary changes, e.g., changes that occur from one generation to the next, we have run a set of simulations in which a population of neural networks, representing the nervous system of creatures that interact with an independent external environment, is subject to both an evolutionary and a learning process. Creatures are selected for reproduction on the basis of their ability to explore one of two possible environments. The creatures have the possibility to learn using the backpropagation procedure but, since we did not want to provide

them with the correct teaching input at each time step as decided by the researcher, we used a self-teaching architecture (cf. Nolfi and Parisi, 1993). In this architecture the teaching input for the network that controls the individual's behavior (standard network) is provided by a separate network (teaching network) that relies on the same sensors of the standard network and is subject to the evolutionary process in the same way as the standard network. We do not claim that this architecture or the backpropagation learning procedure itself are biologically plausible. We adopted the self-teaching architecture because it allows evolution to select, by selecting the weights of the teaching network, the way in which environmental information modifies the nervous system of our creatures and, as consequence, their behavior during life.

By analyzing the creatures that evolved we found that while in individuals that are not allowed to learn the inherited characters (weights, in our case) are directly selected for their ability to produce successful behaviors, in the case of individuals that can learn, the characters are selected for their ability to incorporate a predisposition to learn. This genetically inherited predisposition to learn consists of several things: (a) the presence of starting conditions at birth (i.e., initial weights for learning) that canalize learning in the right direction; (b) the ability to select experienced inputs that similarly canalize learning in the right direction; (c) the ability to reach and maintain a desirable stable state. This pattern of results is found both in the case the environment is stable (unchanging) (Nolfi and Parisi, 1994) and in the case the environment changes from one generation to the next - the present case. In this second case, however, things are more complex because the starting conditions should be able to canalize learning in different directions and to reach different stable states in different environments. We succeeded in evolving individuals that can recognize the type of environment in which they happen to be born and to learn to adapt to the particular environment. By comparing the stimuli experienced in the two environments by individuals that learn and by individuals that do not learn it can be shown that the weights inherited by the learning individuals are able to enhance the differences between the stimuli experienced in the two environments and,

therefore, to make it easier to learn different behaviors in the two environments.

The simulations described in this paper show that learning cannot be studied, as is done in "classical" connectionism (Rumelhart and McClelland, 1986), independently from evolution. Evolution selects for genetically inherited predispositions to learn in environments that can require different adaptations. These genetically inherited predispositions interact with the sensory input available in the environment to determine the kind of adapted behavior which is appropriate for the particular environment in which the individual happens to develop.

Future directions of research include the study of the effects of different patterns of environmental change across generations (e.g., environments that change randomly at each new generation), of environments that can change during an individual's life, and of more complex environments and behaviors.

References

- Ackley, D.E., & Littman, M.L. (1991). Interaction between learning and evolution. In C.G. Langton et. al (eds.) *Proceedings of the Second Conference on Artificial Life*. Addison-Wesley: Reading, MA.
- Baldwin, J.M. (1896). A new factor in evolution. *American Naturalist*, 30, 441-451.
- Belew, R.K., McInerney, J., & Schraudolph, N.N. (1991). Evolving networks: using the genetic algorithm with connectionistic learning. In C.G. Langton et. al (eds.) *Proceedings of the Second Conference on Artificial Life*. Addison-Wesley: Reading, MA.
- Boyd, R., & Richerson, P.J. (1985). *Culture and the evolutionary process*. Chicago: University of Chicago Press.
- Hinton, G.E., & Nowlan, S.J. (1987). How learning guides evolution. *Complex System*, 1, 495-502.
- Holland, J.J. (1975). *Adaptation in natural and artificial systems*. Ann Arbor, Michigan: University of Michigan Press.
- 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 the initial conditions. *Complex Systems*, 4, 269-280.
- Gruau, F., & Whitley, D. (1993). Adding learning to the cellular development of neural networks: evolution and the baldwin effect. *Evolutionary Computation*, 1(3), 213-233.
- Miller, G.F., & Todd, P.M. (1990) Exploring adaptive agency I: theory and methods for simulating the evolution of learning. In D.S. Touretzky, J.L. Elman, T.J. Sejnowski and G.E. Hinton (eds.), *Proceedings of the 1990 Connectionist Models Summer School*. San Matteo, CA: Morgan Kaufmann.
- Mondada F., Franzi E., & Ienne P. (1993). Mobile Robot miniaturisation: A tool for investigation in control algorithms. In: *Proceedings of the Third International Symposium on Experimental Robotics*, Kyoto, Japan.
- Nolfi, S., Elman, J.L., & Parisi, D. (1994a). Learning and Evolution in Neural Networks. *Adaptive Behavior*, 3, 5-28.
- Nolfi, S., Miglino, O., & Parisi, D. (1994b). Phenotypic plasticity in evolving neural networks. In: D. P. Gaussier and J-D. Nicoud (Eds.) *Proceedings of the Intl. Conf. From Perception to Action*, Los Alamitos, CA: IEEE Press.
- 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.
- Nolfi, S., & Parisi, D. (1994). Desired responses do not correspond to good teaching input in ecological neural networks. *Neural Processing Letters*, 2, 1-4.
- Parisi, D., Cecconi, F., & Nolfi, S. (1990). Econets: neural networks that learn in an environment. *Network*, 1, 149-168.

Parisi, D., & Nolfi, S. (1995). How learning can influence evolution within a non-Lamarckian framework. In Belew, R.K., Mitchell, M. (eds.) *Plastic Individuals in Evolving Populations*. SFI Series, Addison-Wesley

Plunkett, K., & Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: implications for child language acquisition. *Cognition*, 38, 43-102.

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

Rumelhart, D.E., & McClelland, J.L. (1986). *Parallel Distributed Processing*. Mass.: MIT Press.

Todd, P.M., & Miller, G.F. (1991). Exploring adaptive agency II: simulating the evolution of associative learning. In J. A. Meyer and S.W. Wilson (Eds). *From animals to animats*. Proceedings of the First International Conference on Simulation of Adaptive Behavior. Cambridge, MA: MIT Press.

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