

Learning and Evolution in Neural Networks

Stefano Nolfi
Institute of Psychology
National Research Council
15, Viale Marx, 00137 - Rome
e-mail:nolfi@ip.rm.cnr.it

Introduction

Evolution and learning are two forms of adaptation that operate on different time scales. Evolution is capable of capturing relatively slow environmental changes that might encompass several generations. Learning, instead, allows an individual to adapt to environmental changes that are unpredictable at the generational level. Moreover, while evolution operates on the genotype, learning affects the phenotype and phenotypic changes cannot directly modify the genotype. Recently, the study of artificial neural networks that are subjected both to an evolutionary and a lifetime learning process received an increasing attention. These studies have been conducted with two different purposes: (a) looking at the advantages, in terms of performance, of combining two different adaptation techniques; (b) understanding the role of the interaction between learning and evolution in natural organisms (for a review see, Nolfi and Floreano, 1999). The general picture that emerge from this body of research is that, within an evolutionary perspective, learning has several different adaptive functions:

- It might help and guide evolution by channelling the evolutionary search toward promising directions.
- It might supplement evolution by allowing individuals to adapt to environmental changes that, by occurring during the lifetime of the individual or within few generations, cannot be tracked by evolution.
- It might allow evolution to find more effective solutions and facilitate the ability to scale up to problems that involve large search space.

Learning has also costs. In particular, it might increase the unreliability of evolved individuals (Mayley, 1997). Since the learned abilities are determined also by the learning experiences, learning individuals might fail to acquire the required abilities in unfavorable conditions.

How learning might help and 'guide' evolution

A simple and clear demonstration of how learning might influence evolution even if the characteristics that are learned are not communicated to the genotype has

been provided by Hinton and Nowlan (1987). The authors considered a simple case in which (a) the genotype of the evolving individuals consists of 20 genes that encode the architecture of the corresponding neural networks, and (b) only a single architecture (i.e. only a single combination of gene values) confers added reproductive fitness. Individuals have a genotype with 20 genes that can assume two alternative values (0 or 1). The only combination of genes that provide a fitness value above 0 consists of all ones. In this extreme case, the probability of finding the good combination of genes would be very small given that the fitness surface looks like a flat area with a spike in correspondence of the good combination. Indeed, on such a surface, artificial evolution does not perform better than random search. Finding the right combination is like looking for a needle in a haystack. The fitness surface is a metaphor often used to visualize the search space on an evolutionary algorithm. Any point on the search space corresponds to one of the possible combinations of genetic traits and the height of each point on the fitness surface corresponds to the fitness of the individual with the corresponding genetic traits.

The addition of learning simplify significantly the evolutionary search. One simple way to introduce learning is to assume that, in learning individual, genes can have three alternative values [0, 1, and ?] where question marks indicate modifiable genes whose value is randomly selected within [0, 1] each time step of the individuals' lifetime. By comparing learning and non-learning individuals one can see that performance increases throughout generations much faster in the former than in the latter. The addition of learning, in fact, produces an enlargement and a smoothing of the fitness surface area around the good combination that, in this case, can be discovered much more easily by the genetic algorithm. This is due to the fact that not only the right combination of alleles but also combinations which in part have the right alleles and in part have unspecified (learnable) alleles might report an average fitness greater than 0 (fitness monotonically increases with the number of fixed right values because the time needed to find the right combination is inversely proportional, on the average, to the number of learnable alleles). As claimed by the authors, "it is like searching for a needle in a haystack when someone tells you when you are getting close" (1987, p. 496).

The Hinton and Nowlan's model is an extreme simplified case that can be easily analyzed but that makes several unrealistic assumptions: (1) there is not distinction between genotype and phenotype, (2) learning is modeled as a random process that does not have any directionality, and (3) there is no distinction between the learning task (i.e. the learning functions that individuals try to maximize during lifetime) and the evolutionary task (i.e. the selection criterion that determine the individuals that are allowed to reproduce). Further research conducted by Nolfi et al. (1994) showed how, when these limitations are released,

learning and evolution display also other forms of interactions that are mutually beneficial.

Nolfi et al. (1994) studied the case of artificial neural networks that 'live' in a grid world containing food elements. Networks evolve (to become fitter at one task) at the population level and learn (a different task) at the individual level. In particular, individuals are selected on the basis of the number of food elements that they are able to collect (evolutionary task) and try to predict the sensory consequences of their motor actions during their lifetime (learning task).

The genotype of the evolving individuals encode the initial weights of a feedforward neural network that, each time step, receives sensory information from the environment (the angle and the distance of the nearest food element and the last planned motor action), determines a given motor action selected within four options (move forward, turn left, turn right or stay still) and predicts the next sensory state (i.e. the state of the sensors after the planned action will be executed). Sensory information is used both as input and as teaching input for the output units encoding the predicted state of the sensors (the new sensory state is compared with the predicted state and the difference (error) is used to modify the connection weights through back-propagation. As in the case of the Hinton and Nowlan's model, modification due to learning are not transferred back into the genotype.

The experimental results showed that: (a) after a few generations, by learning to predict, individuals increased their performance not only with respect to their ability to predict but also with respect to their ability to find food (i.e. learning produced a positive effect on evolution even if the learning and the evolutionary tasks were different), and (b) the ability to find food increased faster and achieved better results in the case of learning populations than in the case of control experiments in which individuals were not allowed to learn during lifetime. Further analysis demonstrated that (a) can be explained by considering that evolution tend to select individuals that are located in regions of the search space where the learning and evolutionary task are dynamically correlated (i.e. where changes due to learning that produce an increase in performance with respect to the learning task produce positive effect also with respect to the evolutionary task) and that (b) can be explained by considering that, once learning channel evolution toward solutions in which the learning task and the evolutionary task are dynamically correlated, learning allows individuals to recover from deleterious mutations (Nolfi, 1999).

Adapting to changing conditions on the fly

As we claimed above evolution and learning are two adaptive processes that occur at different time scale. This implies that learning might complement evolution by

providing a mean to master changes that occur too fast to be tracked by the evolutionary process. However, as we will see in this section, the combination of learning and evolution deeply alter both processes so that, in individuals that evolve and learn, adaptive characteristics emerge as the result of the interaction between evolutionary and lifetime adaptation and cannot be traced back to only one of the two processes.

Nolfi and Parisi (1997), evolved neural controllers for a small mobile robot that was asked to explore an arena of 60 x 20 cm surrounded by walls. The robot was provided with 8 infrared sensors that could detect walls up to a distance of about 4 cm and two motors that controlled the two corresponding wheels. The colors of the walls switched from black to white and viceversa each generation. Given that the activity of the infrared sensors is highly affected by the color of the reflecting surface (white reflect much more than black walls), to maximize their exploration behavior, evolved robots should modify their behavior on the fly. In the environment with dark walls, in fact, robots should move very carefully when sensors are activated given that walls are detected only when they are very close. In the environment with white walls, on the contrary, robots should begin to avoid walls only when the sensors are strongly activated in order to explore also the area close to the walls.

Individuals learn during lifetime by means of a self-generated teaching signals. The genotype of the evolving individuals encoded the connection strengths of two neural modules: a teaching module that each time step receives the state of the sensors as input and produce a teaching signal as output and an action module that receives the state of the sensors as input and produce motor actions as output. The self-generated teaching signal is used to modify the connection strengths of the action module (for a similar architecture see also Ackley and Littman, 1991). This implies that not only the initial behavior produced by the evolving individuals but also what individuals learn is the result of the evolutionary process and is not determined by the experimenter.

Evolved robots displayed an ability to discriminate the two types of environments and to modify their behavior accordingly thus maximizing their exploration capability. The analysis of the obtained results revealed that this ability resulted from a complex interaction between the evolutionary and learning process. For example, evolved individuals displayed an inherited ability to behave so to enhance the perceived differences between the two environments. This in turns allows the learning process to progressively modify the behavior of the robots so to adapt to the different environmental conditions.

More generally this and other researches showed that evolution, in the case of individuals that are able to change during lifetime as a result of learning, do not tend to develop directly an ability to solve a problem but rather tend to develop a predisposition to acquire such ability through learning.

Other experiments conducted by co-evolving two competing populations of predator and prey robots (Nolfi and Floreano, 1998) emphasized how lifetime learning might allow evolving individuals to achieve generality, i.e. the ability to produce effective behavior in a variety of different circumstances. Predators consisted of small mobile robots provided with infrared sensors and a linear camera with a view angle of 36° with which they could detect the prey. Prey consisted of mobile robots of the same size provided only with infrared sensors but that had a maximum available speed set to twice that of the predators. Predators were selected for their ability to catch prey while prey were selected for their ability to escape predators.

What is interesting about this experimental situation is that, given that both populations changes throughout generations, predators and prey are facing ever-changing and potentially progressively more complex challenges. Interestingly the authors observed that in this situation, evolution alone displayed severe limitations and progressively more effective solutions could be developed only by allowing evolving individuals to adapt on the fly through a form of lifetime learning. Indeed, any possible fixed strategy was able to master different type of competitors and therefore only by combining learning and evolution the authors were able to synthesize individuals able to deal with competitors adopting qualitatively different strategies. Indeed, by evolving learning individuals, the authors observed the emergence of predators able to detect the current strategy adopted by the prey and to modify their behavior accordingly.

Other advantages

Floreano and Urzelai (in press) conducted a set of experiments in which the genotype of the evolving individuals encoded the learning properties of the neurons of the corresponding neural network. These properties included one of four possible hebbian learning rules, the learning rate, and the sign of all the incoming synapses of the corresponding neuron. When the genotype is decoded into a neural controller, the connection strengths are set to small random values. As reported by the authors, after some generations, the genetically specified configuration of learning rules tend to produce changes in the synaptic strengths that allow individuals to acquire the required competencies through lifetime learning. By comparing the results obtained with this method with a control experiment in which the strength of the synapses were directly encoded into the genotype, the authors observed that evolved controllers able to adapt during lifetime can solve certain tasks faster and better than standard non-adaptive controllers. Moreover they demonstrated that their method scales up well to large neural architectures.

The authors applied this method to evolve neural controllers for a mobile robots. Interestingly, the analysis of the synaptic activity of the evolved controllers showed that several synapses did not reach a stable state but keep changing all the time. In particular, synapses continue to change even when the behavior of the robot became rather stable.

Similar advantages has been reported by Husband et al. (1999) who evolved a type of neural networks in which neurons, that were distributed over a 2D surface, emitted 'gases' that diffused through the network and modulated the transfer function of the neurons in a concentration-dependent fashion thus providing a form of plasticity.

Finally, the experiments performed by Di Paolo (2000), showed how learning could play the role of an homeostatic process whereby evolved neural networks adapt in order to remain stable in the presence of external perturbations.

Discussion

By reviewing the recent literature we demonstrated how the interaction between learning and evolution deeply alters both the evolutionary and the learning process themselves. Evolution in interaction with learning displays dynamics very different from those which are observed in evolution alone. While in non-learning individuals the characters that are selected through evolution directly incorporate an ability to produce successful behaviors, in learning individuals they incorporate a predisposition to learn, i.e. a predisposition to acquire the necessary abilities through learning. This predisposition to learn may consist of:

- 1) the presence of starting conditions that canalize learning in the right direction. Evolution may select initial weight matrices or network architectures that cause a better and/or a faster learning (Belew et al, 1992). This happens either when the learning task and the evolutionary task are the same or when they differs. In the latter case, evolution does not only select individuals that have a predisposition to better learn, but also individuals that, by learning a given task, improve their performance with respect to the evolutionary task.
- 2) an inherited tendency to behave in such a way that the individual is exposed to the appropriate learning experiences. Evolution tends to select characters that produce initial behaviors that enhance the possibility to learn and/or that increase the probability to acquire adaptive characters through learning. In other words evolution tends to select individuals which have an initial behavior suitable for learning and not necessarily for solving the evolutionary task.

Similarly, learning within an evolutionary perspective has quite different characteristics from learning studied in isolation, as in "traditional" connectionist research. While in individuals that learn but are not subjected to an evolutionary

process (e.g., neural networks trained with supervised methods) learning is usually accomplished by ignoring the characters of the individual prior to learning (which are typically generated at random), in evolving plastic individuals learning exploits such starting conditions. Moreover, when the learning process itself (i.e. what it is learn during lifetime) is subjected to evolution and not determined in advance, learning does not necessarily tend to incorporate the right solution to the problem but rather it tends to pull the learning individual in a direction that, given the initial state of the individual, maximizes the chances of acquiring adaptive characters.

References

- Ackley, D.H., and M.L. Littman, 1991. Interaction between learning and evolution, in Proceedings of the Second Conference on Artificial Life, (C.G. Langton et. al eds.), Reading, MA: Addison-Wesley, pp. 487-509.
- Belew, R.K., J. McInerney J., and N.N. Schraudolph N.N., 1992. Evolving networks: using the genetic algorithm with connectionistic learning, in Proceedings of the Second Conference on Artificial Life, (C.G. Langton et. al eds.), Reading, MA: Addison-Wesley.
- Di Paolo, E.A., 2000. Homeostatic adaptation to inversion in the visual field and other sensorimotor disruptions, in From Animals to Animats 6. Proceedings of the VI International Conference on Simulation of Adaptive Behavior, (J-A Meyer, A. Berthoz, D. Floreano, H.L. Roitblat, and S.W. Wilson eds.), Cambridge, MA: MIT Press, pp. 440-449.
- Floreano, D., and J. Urzelai, in press. Evolving robots to face unpredictable change, Robotics and Autonomous Systems.
- *Hinton, G.E., and S.J. Nowlan, 1987. How learning guides evolution, Complex Systems, 1:495-502.
- Husband, P., T. Smith, N. Jakobi, and M. O'Shea, 1999. Better living through chemistry: Evolving GasNets for robot control, Connection Science, 3-4:185-210.
- Mayley, G., 1997. Landscapes, learning costs, and genetic assimilation, Evolutionary Computation, 4:213-234.

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

*Nolfi, S., and D. Floreano, 1999. Learning and evolution, Autonomous Robots, 1: 89-113.

Nolfi S., and D. Floreano, 1998. Co-evolving predator and prey robots: Do 'arm races' arise in artificial evolution?, Artificial Life, 4:311-335.

Nolfi, S., and D. Parisi, 1997. Learning to adapt to changing environments in evolving neural networks, Adaptive Behavior, 1:75-98.

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

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