

# Neural Networks in an Artificial Life Perspective

Stefano Nolfi<sup>\*~</sup> Domenico Parisi<sup>\*</sup>

<sup>\*</sup>Institute of Psychology, National Research Council, Rome, Italy

<sup>~</sup>Sony Computer Science Laboratory Inc., Tokyo, Japan

**Abstract.** In the last few years several researchers within the Artificial Life and Mobile Robotics community used Artificial Neural Networks. Explicitly viewing Neural Networks in an Artificial Life perspective has a number of consequences that make research on what we will call Artificial Life Neural Networks (ALNNs) rather different from traditional connectionist research. The aim of the paper is to make the differences between ALNNs and "classical" neural networks explicit.

## 1. Introduction

Although the biological inspiration of neural network models is explicit [1, 2] connectionist research on neural networks is mainly conducted outside the field currently called Artificial Life. This is due to the fact that neural networks tend to be isolated from everything else except the input and teaching input that a researcher may provide to a learning network. Much work on neural networks views them as abstract computational devices or information processing machines which, by adopting a brain style of computation, are capable of input/output mappings (behaviors) that sometimes miraculously resemble the mappings (behaviors) exhibited by humans. But if neural networks are viewed not simply as computational devices or information processing machines but as models of nervous systems one cannot ignore that nervous systems and behaviors are part of organisms which have a body with a size, a shape, physically situated sensory and motor organs, and internal organs and systems beyond the nervous system. The organism lives in a physical environment with which it constantly interacts and which can contain other organisms (including conspecifics), abiotic elements, and various types of artifacts. Moreover individual organisms are member of a biological population of non-identical individuals which are born, live for some time, reproduce, and die, and in the case of human organisms, it is also a member of a cultural population of individuals sharing culturally inherited behaviors and artifacts (for a detailed analysis see [3]).

In this paper we will limit ourselves to the analysis of three important factors: (a) the importance of living in an external environment, (b) the importance of having a body, (c) the importance to adapt to the environment through an evolutionary process.

## 2. ALNNs Live in an Environment

Artificial Life Neural Networks (ALNNs) live in a physical environment. They are ecological networks [4]. This implies that in ALNNs, as in real organisms, it is the

physical environment that assigns a "semantics" to the output of the nervous system. The output of an ALNN primarily encodes movements of the organism's body (or body parts) that cause changes in the environment or in the relationship between the organism's body and the environment. The behavior of ALNNs is the result of the output of the network itself, the environment, and of their interactions.

The physical environment not only transforms the output of a neural network into behavior but it is the source of the input arriving to the neural network. In classical connectionism it is the researcher who arbitrarily decides which input patterns are "seen" by the network, in what order they are "seen", what is the frequency of each pattern, etc. On the contrary, it is the independent physical environment that is responsible for the input patterns encoded in an ALNN's input units. In fact, the input of an ALNN primarily encodes the state of the local physical environment around the organism. Therefore, for ALNNs it is the environment with its independent structure and dynamics, not the researcher, that determines which input patterns are "seen" at any given time by the network.

The physical environment is critical for giving a "semantics" to the activation pattern of the input and output units of a neural network. The input patterns that are received by classical neural networks in their input units have no grounding (in the same sense that symbols in symbol manipulation models are ungrounded [5]). An input pattern can be said to be grounded when it refers to something else beyond itself. The input patterns of classical neural networks do not refer to anything beyond themselves. On the contrary the patterns received by ALNNs are physically caused by the independent environment in which they live and behave. Hence, the input patterns of ALNNs refer to this independent physical environment and therefore they are grounded.

An important consequence of this is that since with their motor output ALNNs can change the environment or their relation to the environment, in both cases the input arriving from the environment to the network tends to be influenced by the network's motor output. The network's output in one cycle can partly determine the network's input in the next cycle. Therefore, ALNNs can at least in part control their own input. If organisms are said to adapt to their environment by constructing internal models of the environment [6], these models are likely to be constructed by learning how the environment responds to the organism's own actions, that is, by observing which changes in the environment or in the relation of the organism to the environment result from which actions. If there is no environment which responds to the network's actions - as in classical neural networks - no such model can be constructed. In fact, classical neural networks are restricted to learning by extracting the regularities inherent in their passively received input. ALNNs learn both in this way and by noticing the regularities in the relationship between their actions and the consequences of these actions. (This active view of knowledge and of knowledge acquisition makes ALNNs closer to Piaget than classical neural networks.)

If ALNNs can control their input they can use this opportunity to be exposed to preferred input. Preferred input can be input to which a network already knows how to respond [7] or it can be input that allows the network to disambiguate or correctly recognize the objects encountered in the environment [8].

Another consequence of living in a physical environment is that learning itself may be different in ALNNs and in classical neural networks. What classical networks learn is simply to associate each input with the correct output. In contrast, since ALNNs partly control their own input with their output they can learn to generate output that allows them to experience the appropriate input during learning [9, 10]. In fact, ALNNs may be learning *two* different things. They can learn to associate each input with the correct output, like classical networks, but they can also learn to have the appropriate learning experiences, i.e., they can modify their behavior so that they are exposed to inputs from which they can learn more than from other inputs.

A further consequence of the fact that ALNNs have some control on their own input is that the succession of inputs experienced by an ALNN tends to have an intrinsic temporal character which is normally absent in the sequence of inputs experienced by a classical network. The intrinsic temporal character of ALNNs also emerges clearly if we consider not the succession of inputs they experience but the succession of outputs they generate. The behavior of an ALNN tends to be evaluated not in terms of single outputs but in terms of entire sequences of outputs. If an ALNN is approaching food, it is only the terminal action of reaching the food that rewards the network but the succession of motor outputs that brings the network to the food is given some sort of temporal unity by this final action.

### **3. ALNNs Have a Body**

Classical neural networks do not have a body. They are abstract "systems" for computing an output given an input. Although connectionism does not subscribe to the functionalism of cognitivism, and of psychology more generally, according to which cognition can and should be studied apart from the details of the physical system underlying it, its physicalism is restricted to the brain and it does not extend to the organism's body. In contrast, Artificial Life can be viewed as an attempt to give a body back to the machine. Neural networks that live in a physical environment cannot be abstract "systems". They must necessarily possess a body that has a physical location in the environment, occupies a given portion of the environment, has physical relations (e.g., distance) with other objects in the environment, etc.

As physical objects ALNNs can only interact *physically* with their environment. All is encoded in an ALNN's input units is the current pattern of various types of energy (light energy, sound energy, mechanical energy, thermal energy, etc.) in the vicinity of the organism's body or inside the organism's body, and all is encoded in an ALNN's output units are changes in the physical position of the organism's body or body parts (movements) or inside the organism's body.

If we consider an ALNN in an extended sense as not just a neural network but as an entire organism with its physical body, the interactions between an ALNN and its environment can be mediated by the neural network but the environment can also affect an ALNN without directly affecting the neural network. For example, an ALNN can bump on an obstacle present in the environment so that it cannot proceed

further. Or, other parts of the body of the ALNN, not its neural network, may be affected by the environment. For example, the ALNN can eat some food and its motivational state (hunger) changes. Conversely, the ALNN's body can cause changes in the environment "unintended" by the ALNN's nervous system. For example when an ALNN is moving in the environment it can inadvertently displace an object encountered on its path. Even the "intended" effects that an ALNN's output has on the environment are filtered by the physical properties of the ALNN's body. For example, if an ALNN must reach a particular position in the environment with the endpoint of its two-segment arm, the length of the arm's two segments is a critical factor.

The implications of having a body and of interacting with an external environment are even more significant when they are real physical objects (as in the case of ALNNs used to control mobile robots). In this case in fact several additional factors (e.g. friction, inertia, ambient light, noise, etc.) arising from the physical properties of the robot and of the environment should be necessarily taken into account [11]. Moreover only realistic types of sensors and actuators (instead of idealized ones that may not respect all physical constraints or can have an infinite precision) can be used. In the case of a physical body and environment, the input and the output of the network should necessarily correspond to physical measures or forces and cannot include any abstract information provided by the experimenter even unconsciously. Finally, only information really available in the environment can be used for training.

#### **4. ALNNs evolve**

Classical neural networks are generally trained starting from scratch (i.e. randomly assigned connection weights). Sometime the experimenter designs the rough architecture of the network by using her/his insights or biological data. However, even in this case, most of the initial conditions are randomly initialized. ALNNs, when subjected to a simulated evolutionary process [12], are born by inheriting their initial state from other network(s) (i.e. their parents) and therefore do not start from scratch. The inherited information is represented in a "genotype" that contains the information necessary to produce a corresponding phenotypical neural network and therefore is a distinct entity from the network itself.

The distinction between genotype and the corresponding phenotypical neural network has several implications.

Firstable, not only the weights of the neural networks but also the body of the agent, the features of the sensory-motor system, and the architecture of the network can adaptively be selected by being encoded in the genotype and subjected to the evolutionary process [13, 14].

Secondly, networks can change both phylogenetically (due to the evolutionary process) and ontogenetically (due to a learning process). In this case positive interaction between the two processes may arise: evolution may select good starting conditions that enhance the learning process or canalize it in the right directions [10]; learning may help evolution to find good solutions [15, 16] and to adapt to fast

changing environments that cannot be tracked by evolution alone [10].

Finally, ALNNs may acquire through evolution an ability to extract from the environment reinforcement learning signals [17] or auto-generated teaching inputs [18] and use them to adapt to the environment during their lifetime.

## References

1. Rumelhart, D.E., and McClelland, J.L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Vol. 1. Cambridge, MA:MIT Press
2. Amit, D.J. (1989). *Modeling brain functions*. Cambridge: Cambridge University Press.
3. Parisi, D., and Nolfi, S. (in preparation). *Neural networks in an Artificial Life perspective*
4. Parisi, D., Cecconi, F., and Nolfi, S. (1990). Econets: Neural networks that learn in an environment. *Network*, 1:149-168.
5. Harnad, S. (1990). Symbol grounding problem, *Physica D*, 42: 335-464.
6. Gallistel, C.R. (1990). *The organization of learning*. MIT Press, MA.
7. Nolfi, S., and Parisi, D. (1993). Self-selection of input stimuli for improving performance, In G. A. Bekey (Ed.) *Neural Networks and Robotics*, Kluwer Academic Publisher
8. Nolfi, S. (in press). Evolving non-trivial behaviors on real robots: a garbage collecting robot. *Journal Robotics and Autonomous System*, special issue on "Robot learning: The new wave"
9. Nolfi, S., and Parisi, D. (1994). Desired responses do not correspond to good teaching input in ecological neural networks. *Neural Processing Letters*, (1) 2:1-4
10. Nolfi, S., and Parisi, D. (1997). Learning to adapt to changing environments in evolving neural networks, *Adaptive Behavior* (5), 1:75-98
11. Brooks, R.A. (1992). Artificial life and real robots. In F. J. Varela, and P. Bourguine (Eds.) *Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life*, Cambridge, MA: MIT Press/Bradford Books.
12. Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor.
13. Gruau, F. (1995). Automatic definition of modular neural networks. *Adaptive Behavior*, 2:151-183.
14. Lee, W.P., Hallam, J., and Lund, H.H. (1996). A Hybrid GP/GA Approach for Co-evolving Controllers and Robot Bodies to Achieve Fitness-Specified Tasks. In *Proceedings of IEEE 3rd International Conference on Evolutionary Computation*. IEEE Press.
15. Hinton, G.E., and Nowlan, S.J. (1987). How learning guides evolution. *Complex System*, 1: 495-502.
16. Parisi, D., and Nolfi, S. (1996). How learning can influence evolution within a non-Lamarckian framework. In R.K. Belew, and M. Mitchell (Eds.) *Adaptive Individuals in Evolving Populations*, SFI Studies in the Science of Complexity, Vol. XXVI, Addison-Wesley
17. Ackley, D.E., and 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.
18. 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.