# Phylogenetic Recapitulation in the Ontogeny of Artificial Neural Networks

Stefano Nolfi Domenico Parisi
Institute of Psychology
National Research Council
15, Viale Marx, 00137 - Rome
voice:0039-6-86894596 fax:0039-6-824737
email: stefano@kant.irmkant.rm.cnr.it - domenico@gracco.irmkant.rm.cnr.it

**Abstract**

The order in which anatomical characteristics develop in an individual from conception to maturity (ontogeny) may repeat the evolutionary history of the individual's species (phylogeny). Phylogenetically older characteristics often develop earlier during ontogeny and more recent phylogenetic characteristics develop later. We show evidence of phylogenetic recapitulation in the development of the nervous system of simple artificial organisms which are evolved using a genetic algorithm technique. In our simulations we found evidence of both acceleration and retardation in the ontogeny of characters acquired during phylogeny. The prevalence of accelerations with respect to retardations explains why characters are usually anticipated during ontogeny and, therefore, why ontogeny may recapitulate philogeny. Such acceleration and retardation phenomena appear to affect only functional characters, i.e. characters that play a role in determining the fitness of individuals.

## Introduction

It has often been proposed that the order in which anatomical characteristics develop in an individual from conception to maturity (ontogeny) may, in general terms, repeat the evolutionary history of its species (phylogeny). More controversial is the level (genes, proteins, physiological and neurological systems, behaviour) and, as a consequence, the disciplines (genetics, molecular biology, physiology and the neurosciences, psychology) that should be concerned with such phenomenon (Medicus,1992).

In addition, if we consider recapitulation as the result of a more general process, the evolutionary alteration of developmental times and rates to produce acceleration and retardation in the ontogenetic development of specific characters (see Gould, 1977; McKinney, McNamara, 1991), the problem of determining which of the two tendencies (acceleration and retardation) dominates over the other remains open.

We simulated the phylogenetic and ontogentic changes that occur in simple artificial organisms (neural networks) by using a genetic algorithm technique (Holland, 1975). The possibility of manipulating these artificial systems with a high degree of freedom and of controlling each parameter of the whole experiment allows us to obtain interesting results that may shed some light on how and why recapitulation phenomena occur in natural living systems.

## The model

Let us begin by assuming that our ultimate goal is to create an organism (O) which is able to find and eat food in its environment. To each O a string of genetic material or genotype corresponds that specifies developmental instructions which generate a certain number of neurons and control the growing and branching process of the neurons' axons (Nolfi, and Parisi, in press). The result of this growing process is a neural network that represents the nervous system of the corresponding O. The architecture and connection weights of such a network determine the way in which O responds to environmental stimuli, i.e. its behavior. O's behavior determines its fitness, i.e. its reproductive chances, through O's interaction with the environment to which it is exposed.

Each O lives in a simulated environment which is a two-dimensional square divided up into cells. At any particular moment O occupies one of these cells. A number of food elements are randomly distributed

in the environment with each food element occupying a single cell. O has a facing direction and a rudimentary sensory system that allows it to receive the angle (relative to where O is currently facing) and the distance of the nearest food element as input. O is also equipped with a simple motor system that provides it with the possibility, in a single action, to turn any angle from 90 degrees left to 90 degrees right and then move from 0 to 5 cells forward. Finally, when O happens to step on a food cell, it eats the food element which disappears.

O's genotype is represented as a string of 0 and 1. It has a fixed length (1600 bits) and is divided up into 40 blocks, each block corresponding to a single neuron which may or may not be expressed in the phenotypic network. Expressed neurons have an axon genetically programmed to grow and branch out at a certain time during O's ontogeny. Each block contains instructions that determine various properties of the corresponding neuron (see Figure 1). The result of the execution of the developmental instructions at a certain stage of development is shown in Figure 2.

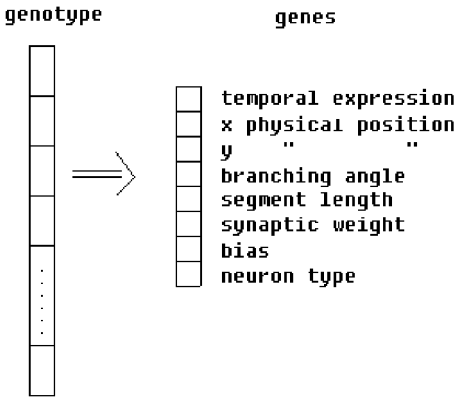


Figure 1: O's genotype. Each block specifies developmental instructions for the corresponding neuron (i.e. time of expression during ontogeny, x and y coordinates of the neuron in the nervous system's bidimensional space, angle of branching and length of each branching segment, weights of the connections departing from the neuron (same weight for all connections), activation bias of the neuron, type of neuron: input, hidden, or output).

When a growing axonal branch of a particular neuron reaches another neuron a connection between the two neurons is established. Figure 2 shows the growth of axons resulting from a random genetic string. Figure 3 shows the connections that are established during this growth. Axonal branchings that do not make connections have been eliminated. Finally, Figure 4 shows the functional network that actually maps sensory input into motor output. The nonfunctional neurons and groups of interconnected neurons of Figure 3 that play no role in this mapping have been removed.

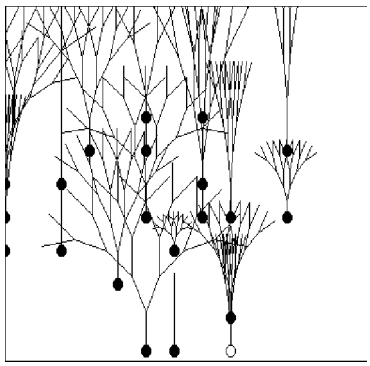

Figure 2: Growth of axons resulting from a randomly generated genetic string. The lowest layer represents sensory neurons, the uppermost layer motor neurons, and the remaining layers internal neurons.

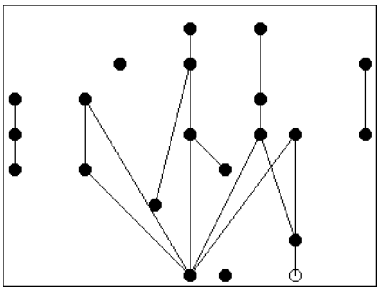

Figure 3: Connections established during the growth process shown in Figure 2.

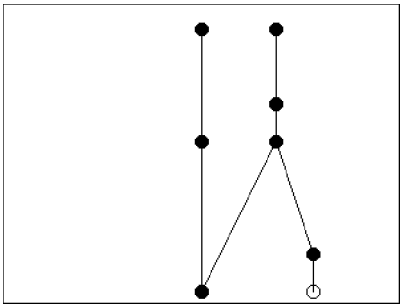

Figure 4: Resulting functional network after elimination of isolated (nonfunctional) neurons and groups of interconnected neurons.

The functional network of Figure 4 determines O's behavior through the interaction with the environment. At each time step, O receives a pattern of activation values on its sensory neurons encoding the position of the nearest food element relative to O. Such an input determines in turn, through a spreading activation process, the activation value of the internal and output neurons. These last neurons determine O's motor reaction to the current input,i.e. O's behavior.

Since Os develop in time, i.e. different neurons

have different expression time during ontogeny, they can react differently to input stimuli in different periods of their life. The genome defines and constructs an initial network at birth but then it continues for a certain time after birth to determine changes in the network's architecture and weights. This period of genetically-based changes during an individual's lifetime can be called its developmental age.

In order to obtain adapted Os, i.e. Os that exhibit an efficient food collecting behavior, we simulate an evolutionary process using a genetic algorithm. (For similar approaches see Harp, Samad, and Guha, 1989; Nolfi, Elman, and Parisi, 1990; Kitano, 1990; Ackley and Littman, 1991). We ran 10 simulations each starting with 100 different randomly generated genotypes. This is generation 0 (G0). Os of G0 are allowed to develop and "live" for 8 epochs, each epoch consisting of 250 actions in 5 different environments (50 actions each). The environment is a grid of 40x40 cells with 10 pieces of food randomly distributed in it. Os are placed in individual copies of these environments, i.e. they live and develop in isolation. At the end of their life (2000 actions) they are allowed to reproduce. However, only the 20 individuals which have accumulated the most food during their life reproduce by generating 5 copies of their genotype. These 20x5=100 new Os constitute the next generation (G1). Random mutations are introduced in the copying process resulting in possible changes in the phenotypic network and/or in the time of development of different subparts of the phenotypic network. The organisms of G1 are also allowed to live for 2000 cycles. The process continues for 1000 generations.

**Results of simulations**

If we look at the fitness of our Os, i.e. at the number of food elements they eat during their life across the 1000 generations, we see that Os are increasingly able to approach food elements. Figure 5 shows the fitness value of the Os of the winning lineage for each of the 1000 generations. The winning lineage is the lineage of the best individual of the last generation. The lineage is constituted by this individual, by its (single) parent in the preceding generation, by its grand parent, its grandgrand parent, etc., until the originator of the lineage in the first generation is reached. The evolutionary increase in fitness implies that generation after generation Os are able to adapt their architecture and synaptic weights to the evolutionary task. The number of food elements available in the entire lifetime of an O is 400. On average the Os of the last generation reach 65% of food elements, with one of the 10 simulations reaching 95%.

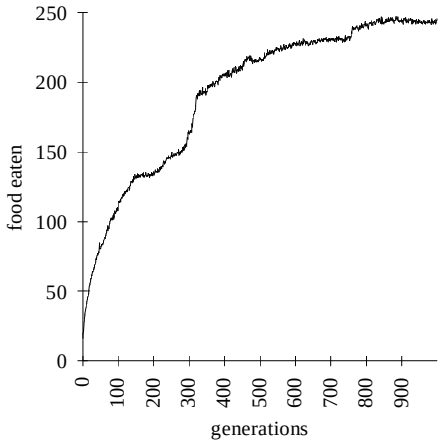


Figure 5: Fitness (=total number of food elements eaten during life) of Os of the winning lineage.The graph represents the average result of 10 simulations starting from initial populations of different randomly generated genotypes.

While Figure 5 shows how the fitness of Os changes phylogenetically, Figure 6 shows how Os' fitness changes ontogenetically. Figure 6 shows the number of food elements Os are able to eat in different epochs of their life. The amount of food eaten increases during life. This implies that the evolutionary process has selected genotypic instructions that result, on the average, in behaviorally useful changes in Os' nervous system after birth.

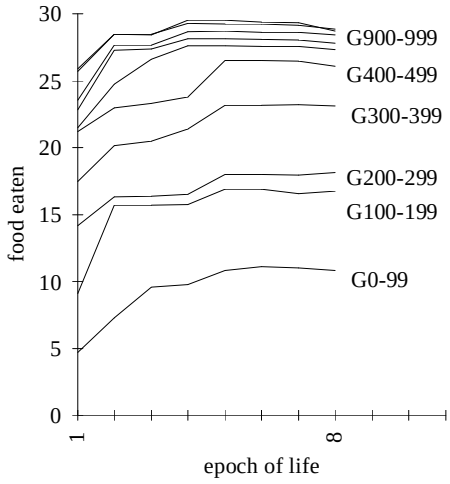


Figure 6. Number of food elements eaten at various epochs of life by Os of the winning lineage. The graph shows the average result of 10 different simulations. Each curve represents the average performance of 100 successive generations.

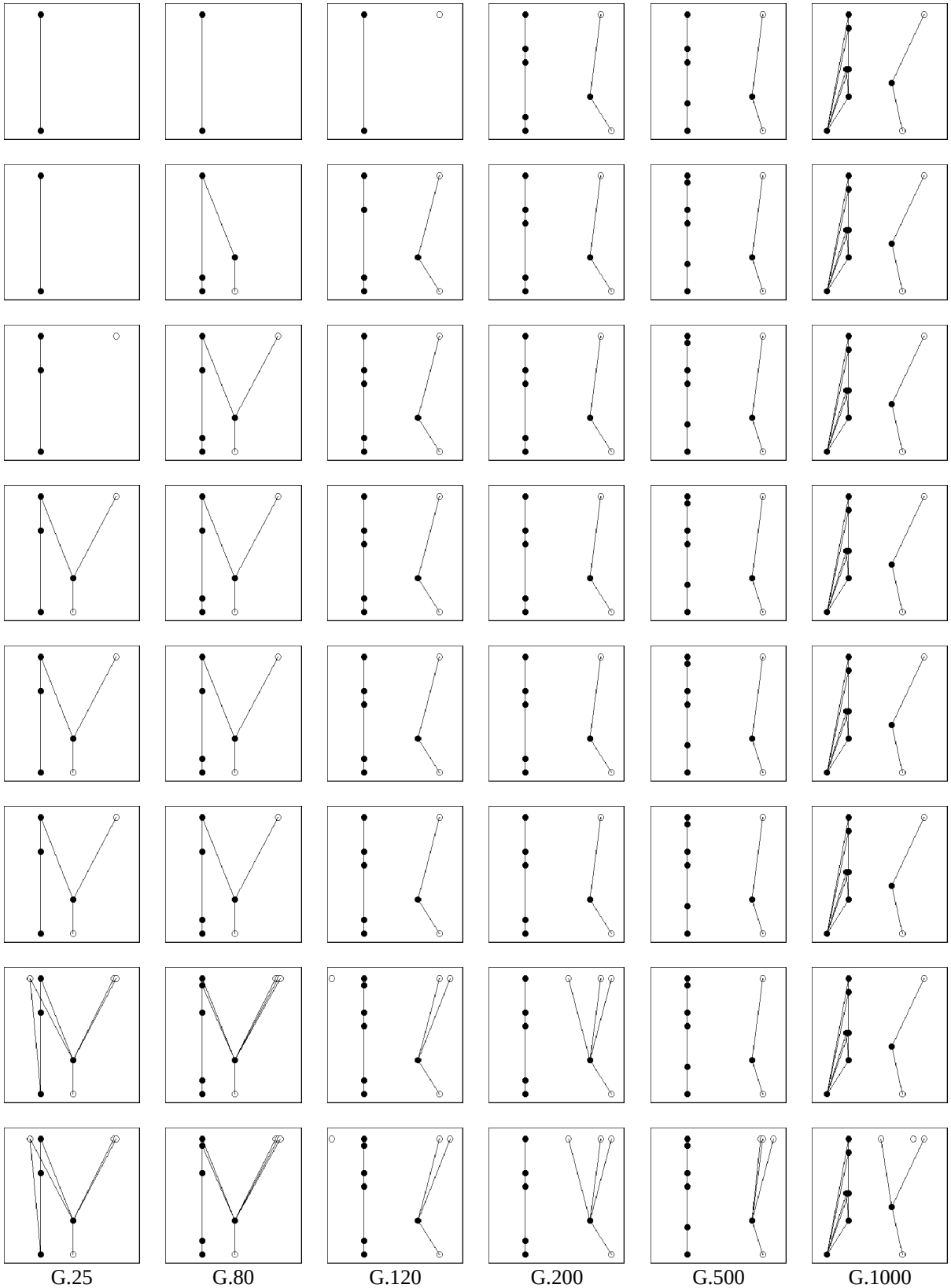


Figure 7. Neural development of 6 individuals belonging to the winning lineage of the most successful simulation. For each of these individuals the functional network controlling the individual's behavior at each of the 8 successive epochs of its life is shown. Hence, the Figure describes neural evolution on the horizontal axis and neural development on the vertical axis. Due to limits of space only 6 Os (belonging to generations 25,80,120,200,500, and 1000, respectively) are shown.

If we analyze how Os' architectures change both ontogenetically and phylogenetically, we find some interesting similarities. Figure 7 shows the changing functional architecture of some Os of the winning lineage of the most successful simulation during the 8 epochs of life. (For reasons of space only 6 Os are shown). Notice that, within the lower layer (i.e. the sensory layer) filled circles represent neurons sensitive to the angle of the nearest food element and empty circles represent neurons sensitive to its distance. In the upper layer (i.e. the motor layer), filled circles represent motor neurons that encode angle of turn and empty circles represent neurons that encode amplitude of the motor step).

Os' architectures appear to structure themselves into two sub-networks. There is a sub-network (on the left side of Os' nervous system) which computes O's direction of movement as a function of the state of the neuron sensitive to food angle, and a second sub-network (on the right side) that controls O's movement speed as a function of food distance. (This second component at first influences also direction of turn through a single connection but this connection is eliminated from generation 120 on). The second sub-network appears at the fourth epoch in generation 25 and it is progressively anticipated in development during the course of evolution (third epoch in generation 80, second epoch in generation 120, and first epoch in generation 200). Similarly, the three internal neurons of the left module that appear at birth in the individual of generation 200 develop after birth in the individuals of preceding generations.

As Gould has observed (Gould, 1977), anticipation of characters in the ontogeny of individuals of successive generations implies a similar succession of stages in ontogeny and philogeny, i.e. some sort of ontogenetic recapitulation of philogeny. In fact, we observe examples of recapitulation in our simulation results. If we analyze the ontogeny of the O of generation 120, for example, we see that its architecture at birth was already present since generation 25. On the other hand, we have to wait until generation 80 to find an O with two internal neurons in the left sub-network that appear in the second stage of its development. In other words, the architecture that appears in the very first stage of development of the O of generation 120 was selected earlier by evolution than the architecture that appears in the second stage of the development of this same individual. Similarly, the architecture that appears in its second stage of development was selected earlier by evolution than the architecture that appears in the third stage (three internal neurons in the left sub-network instead of two).

Selective reproduction based on fitness appears to be the force that causes this developmental anticipation of newly evolved structures. Selective reproduction favours Os that are efficient in all epochs of their life and, therefore, as early as possible during lifetime. (Remember that in the present simulations total fitness is simply the sum of all fitnesses collected in all epochs of life).

On the other hand, this force applies only to those parts of Os' architecture that allow an increase in performance. This implies that only functional neurons and, among these, only the neurons that are really useful, tend to be anticipated. A good example are the additional output neurons that appear in the last two epochs of development in Figure 7. Such additional output neurons, being all connected to the same internal neuron with the same weights, do not change Os' behavior and, as a consequence, are not anticipated during the course of the evolutionary process. (When more than one output neuron of the same type are present in an O's nervous system, O's behavior is determined by the average activation value of all neurons of the same type). On the contrary, the development of the redundant output neurons is postponed from the 7th to the 8th epoch of development after generation 500.

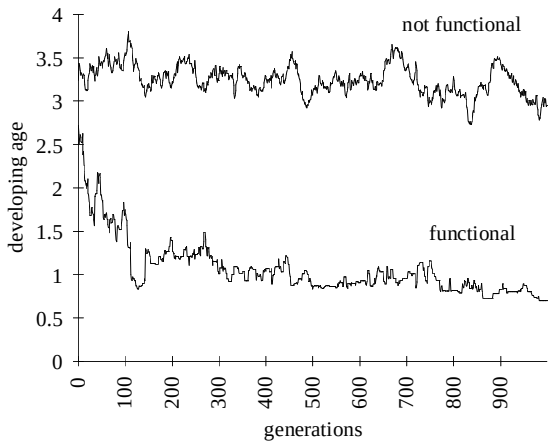


Figure 8. Average epoch of expression of functional and nonfunctional neurons across generations. Each curve represents the average result of 10 different simulations.

That the anticipation process applies only to the functional parts of neural architectures is further shown in Figure 8. This Figure presents the average time of development of functional versus nonfunctional neurons across 1000 generations. Functional neurons appear earlier in development than nonfunctional neurons. Furthermore, while the average time of expression of functional neurons decreases rapidly after a few generations, the average time of expression of nonfunctional neurons does not change significantly across evolutionary time.

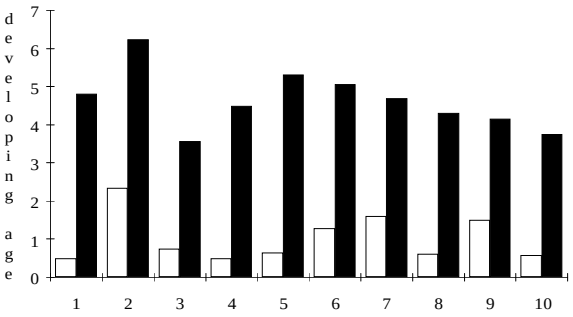


Figure 9. Average epoch of expression of functional neural modules during development for each of the 10 simulations. Filled histograms represent the average epoch in which neural modules appear for the first time during development. Empty histograms represent the average epoch of appearance of neural modules.

Functional neurons that appear at a certain generation or that become functional at a certain generation tend to be anticipated in the development of Os of successive generations (see Figure 8). Similarly, Figure 9 shows that functional neural modules that appear for the first time at a certain generation tend to be anticipated in the development of Os of successive generations. (A neural module is an internal or an output neuron that receives connections from one or more input neurons and/or from other neural modules.)

The fact that old functional neurons and neural modules tend to be anticipated and then preserved is well shown in Figures 10 and 11 that represent the time of expression of the most frequently found functional neurons and neural modules, respectively, for each generation of the most successful simulation (the same shown in Figure 7). The dark lines represent single neurons or neural modules. The x axis represents generations while the y axis represents developmental time for each neuron or neural module. As can be seen, neurons or neural modules that are preserved for many generations tend to be anticipated and to be maintained in the population until the end. On the contrary, neurons and neural modules that are eliminated at some point in evolution tend to have been recently acquired.

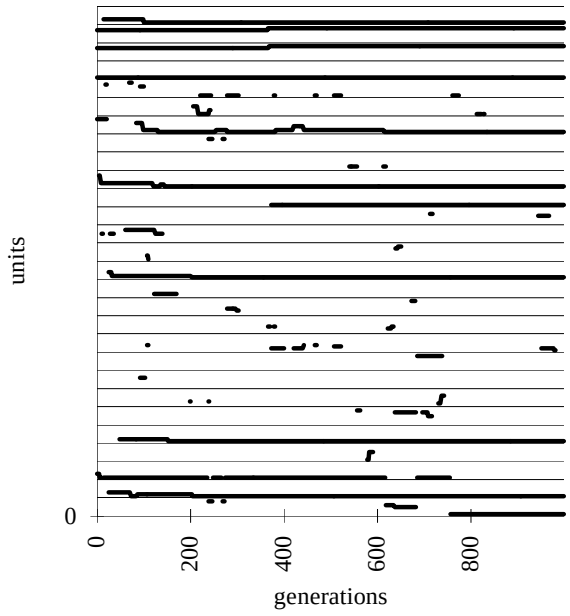


Figure 10. Developmental time of appearence of the 28 neurons that are expressed at least one time in functional networks across generations. The dark lines represent the corresponding neuron and the height with respect to the baseline represent the time needed for such neuron to express itself. The absence of a dark line may indicate that the corresponding neuron is not expressed at all or that it is not included in the functional network of the corresponding O. The two neurons that are retained until the last generation and are postponed at about generation 400 represent the redundant output neurons.

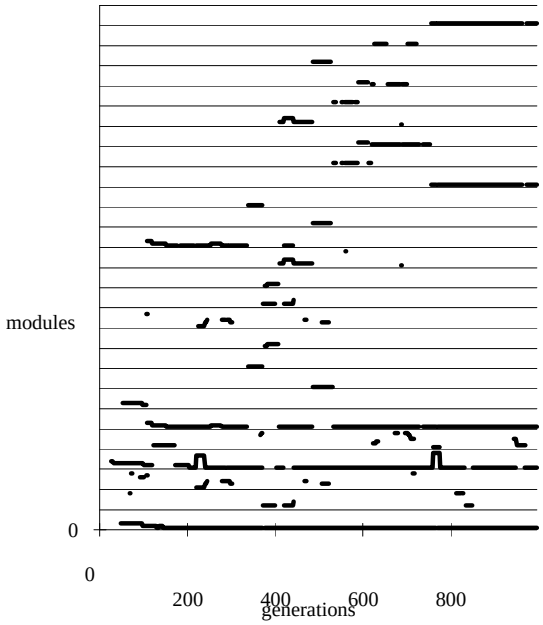


Figure 11. Expression time of the 26 mostly frequently found neural modules across generations. Each dark line represents the corresponding module and the height with respect to the baseline represents the time needed for such a module to express itself. The absence of the dark line may indicate that the corresponding neural module is not expressed at all or that it is not included in the functional network of the corresponding O. Neural modules represent output or hidden neurons that receive connections from one or more input units and/or from other neural modules.

The fact that the anticipation process applies only to functional and useful parts combined with the fact that only neurons and neural modules that are preserved for many generations have the possibility to be anticipated (cf. Figure 10 and 11) ensures that the earlier a neuron or a neural module appears in development the higher the probability that such a neuron or neural model is useful and important in determining a fit behavior and, as a consequence, the higher the probability that it will be preserved in successive generations.

This determines a tendency to preserve long acquired and well tested characters through anticipation and to eliminate useless novelties. Old characters that have been anticipated to early developmental stages are difficult to eliminate because a new character should cause a much better performance in order to replace an old one that appears so early in development. For the same reasons recently acquired characters or characters that have been acquired long ago but have not been anticipated and therefore are unlikely to be very useful, can be easily replaced or eliminated.

However, as can be seen from Figure 7, not all characters that appear early in phylogeny are necessarily preserved in the ontogeny of individuals of later generations. Similarly, characters that appear at different stages of phylogeny and have been preserved may all appear at the same time during the development of Os of successive philogenetical stages. For example, there is no trace in the ontogeny of the O of generation 1000 of the philogenetical stages that gave rise to the left neural module (which is responsible for determining angle of turn). Such older stages were lost and therefore are not recapitulated in the development of the O of generation 1000.

This may be due in part to the fact that in our simulations the expression time of each neuron is encoded by a different fraction of the genotype. This type of encoding implies that neurons and neural modules can be anticipated or postponed in ontogeny independently of each other. This in turn implies that neurons and neural modules cannot assume interphene functions. A character has an interphene function when in addition to having an adaptive value with respect to the external environment it assumes an adaptive value as an intermediate stage which is necessary to induce subsequent developmental characters (see Medicus, 1992). Since our neurons and neural modules cannot have interphene functions because of the genetic encoding we have used, they can be eliminated without any consequence as soon as they lose their adaptive value with respect to the external environment.

**Discussion**

As Gould has observed (Gould, 1977), recapitulation, i.e. the idea that characters that develop in an individual from conception to maturity

may repeat the evolutionary history of the individual's species, implies a fundamental assumption: the length of an ancestral ontogeny is continuously shortened during the subsequent evolution of the lineage. This assumption means that older characters which first appeared at a certain stage of development often develop at an earlier stage during the ontogeny of individuals of successive generations while more recent phylogenetic characters develop later.

However, recapitulation cannot be considered as an absolute law but as a simple tendency because, as several authors have pointed out, acceleration is not general or equal for all characters, new characters can be introduced at any stage of ontogeny (hence, even earlier than older characters), and development can be retarded as well as accelerated (see Gould, 1977). Recapitulation should be considered as the result of a more general process, the evolutionary alteration of developmental times and rates to produce acceleration and retardation in the ontogenetic development of specific characters. It remains to be determined which one of the two tendencies (acceleration and retardation) dominates over the other in particular cases and the implications of such processes of heterochrony (McKinney, McNamara, 1991) for evolution.

In our simulation results we have found evidence of both acceleration and retardation of ontogenetic development although the cases of acceleration largely outnumber the cases of retardation (see Figure 8, 9, 10 and 11). This prevalence of acceleration appears to be due to the selective reproduction mechanism that favors individuals that become efficient in collecting food as soon as possible and therefore individuals that develop their functional nervous system sooner than their ancestors (for a similar explanation see Muller, 1864). On the other hand, we have also shown that the anticipation tendency only applies to the functional characters of individuals, i.e. to characters that imply a direct advantage in the reproductive chances of the individual. In fact, one can expect that, as already observed by Mehnert and Massart at the end of the last century (Menhert, 1898; Massart, 1894), characters which must function first develop first.

As we have already observed, not all the characters that have appeared earlier during phylogeny are necessarily preserved in the ontogeny of individuals of later generations. Characters that become useless as a consequence of the appeareance of new characters tend to be lost and disappear from the ontogeny of individuals of successive generations unless they have assumed an interphene function, i.e. an internal adaptive value as a stage which is necessary to induce subsequent developmental characters. In such a case, the characters may be preserved even after losing their adaptive value.

The fact that ontogeny can be used to infer unknown phylogeny only with a high degree of inaccuracy should not imply that the interaction between ontogeny and philogeny should be ignored. Development appears to be extremely important in explaining evolution despite the fact that the way in which phylogeny and ontogeny interact is far from being well understood (cf. Gottlieb, 1992).

**Acknowledgements**

This research was supported by P.F. "ROBOTICA", C.N.R., Italy

**References**

Ackley, D.E. and Littman, M.L. (1991). Interactions between learning and evolution. *Proceedings of the Second Conference on Artificial Life*. Addison-Wesley: Reading, MA.

Darwin, C. (1859). The origin of species. Reprint from 1979 (J.W.Burrow, Ed.) Harmondsworth: Penguin Books.

Gottlieb, G. (1992). *Individual development and evolution: The genesis of novel behavior*. New York, Oxford University Press.

Gould, S.,J. (1977). *Ontogeny and Phylogeny*. Harward University Press. Cambridge, Massachusetts.

Haeckel, E. (1866). *Generelle morphologie der organismen. Berlin*, Georg Reimer.

Harp, S., Samad, T., and Guha, A. (1989). Toward the genetic synthesis of neural networks. *Proceedings of the Third nternational Conference on Genetic Algorithms.*San Mateo, Kaufmann.

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**, pp. 461-476.

Massart, J. (1894). Le recapitulation et l'innovation en embryologie vegetale. *Bull. Soc. Roy. Bot. Belgique*, **33**, pp.150-247.

Medicus, G. (1992). The inapplicability of the

biogenetic rule to behavioral development. *Human Development* **35**, pp. 1-8

Menhert, E. (1898). *Biomechanik*. Gustav Fisher, Jena.

McKinney, M.L., McNamara, K.J. (1991). *Heterochrony*. Plenum, New York.

Muller, F. (1864). Fur Darwin, in A. Moller, ed., Fritz Muller. (1915). *Werke, Briefe und Leben*. Gustav Fischer, Jena, pp. 200-263.

Nolfi, S., Elman, J, and Parisi, D. (1990). Learning and evolution in neural networks. CRL *Technical Report* 9019. University of California, San Diego.

Nolfi, S. Parisi, D. (in press). "Genotypes" for Neural Networks. In M. A. Arbib (ed.) *The Handbook of Brain Theory and Neural Networks*. Bradford Books, MIT Press.