Investigations on the Robustness of an Evolved Learning Mechanism for a Robot Arm

Gabriel Gómez[1] and Peter Eggenberger Hotz[1,2]
[1]*Artificial Intelligence Laboratory*
Department of Information Technology, University of Zurich, Switzerland
[2]*Emergent Communication Mechanisms Project*
Information Sciences Division, ATR International, Kyoto, Japan
gomez@ifi.unizh.ch, eggen@ifi.unizh.ch

Abstract. This paper reports on a continuing research effort to evolve robot controllers with neural networks and to understand how Nature builds such systems. In an earlier paper [8] we showed that an artificial evolutionary system was able to evolve and grow a neural network, which learned to control a camera to track a light source and showed that the learning mechanisms could perform its task as well as in simulation. In this paper we continue transferring the evolved controller to the real-world and tested the robustness of the learning mechanism as follows: Once the system was able to track a red object with the cameras, we fixed the positions of the cameras and used the same principles to teach a robot arm to move a colored object from an initial location at the periphery of the visual field to the center of it, using visual (e.g., color and movement detection) and proprioceptive (e.g., position of the joints of the arm) feedback. The arm could solve the task not only for two-degrees of freedom, but also for three degrees of freedom. These results suggest that to overcome the ”reality gap” between simulation and the real world one should evolve mechanisms enabling the real-world system to explore its own possibilities instead of specifying the system precisely in simulation.

1 Introduction

In evolutionary robotics the standard approach is to use neural networks as controllers for the agents or robots. Such controllers based on neural networks are evolved by evolutionary algorithms. Most of these approaches use direct genotype-to-phenotype encodings, where each synaptic weight is represented by a genetic parameter. As such an encoding scheme scales quadratically with network size, considerable effort has been put into the development of indirect encoding schemes, where the genome is a developmental program specifying processes which in their turn grow the neural structures and fine-tune the synaptic weights according to the tasks specified by the designer.

In an earlier paper [8] we evolved neural networks using genetic regulatory networks controlling grow processes allowing to connect the neurons to designer defined sensor inputs and motor outputs. At the same time the evolutionary process could not only grow networks, but could also evolve learning mechanisms. In other words the evolutionary process explored not only structures of the neural network, but also how to learn. To be able to explore both, growth and learning, the concept of selective recognition between entities, we call it ”ligand-receptor interactions” played a crucially important role. This concept allows (a) a neuron to recognize more or less specifically another partner neuron, (b) an axon to follow cues in the environment and (c) a neuromodulator to diffuse in the environment and influence specific synapses to change their weights (see figure 2a). In this way learning rules can be evolved and have not to be pre-specified by the designer of the system. The task was that a camera which was actuated by two motors should learn to track a light source. The system was able to find two different principles of how to learn such a task. The AES evolved neuromodulatory systems able to teach the robot about the success of its exploratory movements. The first mechanism localized in the center of the retina cells able to emit neuromodulators when they where activated. This had the effect that the system could learn to move the camera correctly. The second mechanism was more sophisticated and much more efficient: the neuromodulator was emitted if at any given position the movement of the robot brought the light source closer to the center. More details about how this controller was evolved and transferred to the real robot can be found in the following references [7, 8].

In this paper we investigated the evolved neural network for its ability to learn the task in real-world and we asked ourselves how robust the system was. We hypothesize that the incorporation of mechanisms of gene regulation and learning mechanisms potentially leads to more adaptive neural networks, that can help bridging the above mentioned ”reality gap”. To test the robustness of the learning mechanism we had the idea that once the eyes had learned to track a particular colored object, the system should also be able to teach a robot arm to move an object in the center of the visual field. During the experiments it was indeed the case that the robot arm could learn this task for two degrees of freedom as for the eyes, but to test the robustness we increased the degrees of freedom to three and the system was still able to learn its task.

Although in previous work some authors [12, 1, 4, 2, 9, 5, 20, 15, 17] propose to evolve neural network structures or learning rules, none of them could combine developmental processes and learning rules concurrently and could transfer such an evolved working structure on a working robot.

In the next section the methods allowing to combine developmental processes to grow a neural network with learning are summarized (for more details consult earlier work [7, 8]). In section 3 the robotic setup is described, the neural structure and its connections to the robot are described in section 4, then the experiments and results are shown in section 5 and we conclude with a discussion in the last section.

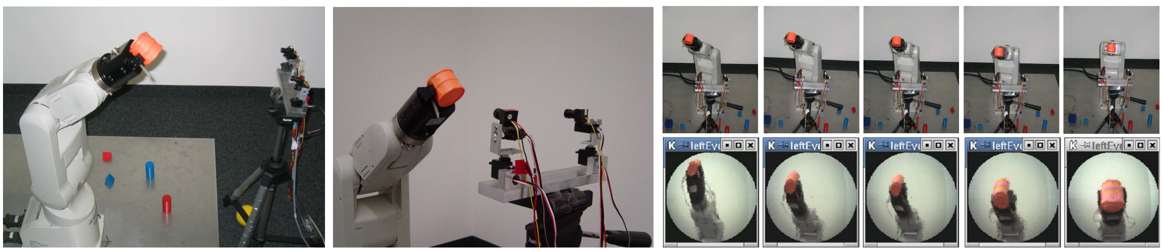

Figure 1: System architecture. From left to right: (a) The six degrees of freedom robot arm. (b) The four degrees of freedom stereo color active vision system. (c) An example of the robot moving an object from the top-left corner of its visual field to the center of it.

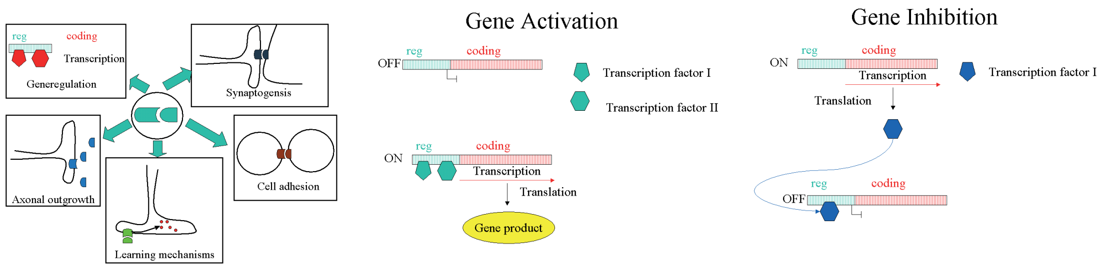

Figure 2: From left to right: (a) Ligand-receptor concept. (b) Gene regulation. By binding to regulatory units of a gene, specific substances, so-called transcription factors, are able to activate or inhibit a gene. If a gene is activated a specific developmental process such as cell division, cell death, axonal outgrowth, synaptogenesis or emission of a neuromodulator is performed.

2 Evolutionary Methods and Developmental Processes

Artificial evolutionary methods are used to optimize a parameter set (usually called a genome) to encode a problem. By simulated mutation, test of sets and selection of the best sets, a problem can be automatically optimized. Here, a more biological approach is propose in order to evolve large neural networks, which are able to grow and learn; in this method the genetic information no longer directly encodes for specific properties (in this case a neural network), but controls developmental processes generating the functioning structures.
Similar to biological cells, the artificial ones can divide, die, move, adhere, produce artificial molecules, and differentiate. In order to control those processes in space and time, each cell contains its own genome and its genes can be activated or inhibited by specific signaling molecules (see figure 2b). Each gene is regulated by several regulatory units and can be activated or inhibited if the impact of incoming signaling molecules is high enough. Each structural gene is linked to a specific function, which is executed if the gene is activated. The cells can become different by exchanging signals and activate different subsets of genes in its genome, which is the basis of cell differentiation. Different cells will perform different functions such as cell division or another function from the above mentioned functions of the cells. Artificial evolution can explore possible interactions between cellular entities such as signaling molecules binding to regulatory units to control gene activities, interactions of adhesion molecules to bind cells together or to build up synapses, or neuromodulators interacting with specific receptors to control the synaptic weights. By selecting those evolved structures with the highest fitness, the AES can learn to coordinate these developmental processes and build structures (neural networks, organisms), which perform designer specified functions. In this paper such processes are exemplified by describing a robotic system with an evolved controller able to learn to track a light source and to control a robotic arm moving a colored object to the center of its visual field with two and three degrees of freedom. Several different concepts had to be developed until the system was able to perform these tasks.

2.1 Growth mechanisms

By varying axonal receptor types different influences can be easily explored by the evolutionary strategy (ES) and growth patterns resulting in topological maps between interneurons and the sensory input emerged. Figure 3 illustrates the mechanisms by which the gene ac-

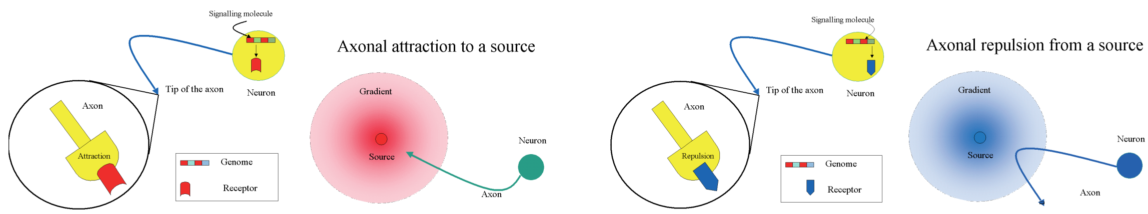

Figure 3: Overview of the most important biological mechanisms used in the proposed artificial evolutionary system. Different cell differentiation will lead to a different receptor expression in the cells and allows to program the cell to be attracted or repulsed from a source depending on the receptor an axon has expressed.

tivities of the cells will determine the types of receptors expressed on their axons and finally determine if a source of signaling molecules has an attraction or repelling influence. Similar ligand-receptor interactions will control which cells will stick together or which will build up synapses or which neuromodulator will change the synaptic weight. By varying the expression of different ligand-receptor interactions the ES can easily explore different connectivity patterns with a small load of genetic information. In figure 5 a somewhat geometric looking overview of the evolved neural net is given, because the paths of the axons searching for their targets are omitted (For other illustrations see figure 4 in paper [8]).

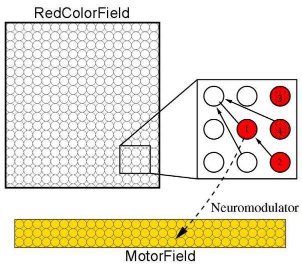

Figure 4: The sensory neurons connect to each other in such a way that the axons point in direction to the center of the retina. If a sensory neuron is activated it will transmit a signaling molecule to its neighbor cell. If this cell is also activated, a neuromodulator is diffused to the motor field and the weights of the active motor neurons are change such that the probability to repeat the same movement in the same situation is increased.

2.2 Learning mechanisms

The exploration of learning mechanisms is based on the expression of artificial receptors on the synapses and the interactions with signaling molecules (often also called neuromodulators), which can specifically combine and control the change of synaptic weights. This mechanism is called ”ligand-receptor concept” [6]. (The concept of neuromodulation was explored by us also in the context of the evolution of a biped robot in simulation [14]). Depending on their specific differentiation, the cells produce and emit signaling molecules (ligands in the context of receptors), which induce a change of the weight of a synapse, if the axon has expressed a receptor able to recognize the signaling molecule. In other words, the synaptic weight changes if ligands emitted from other cells have a strong enough impact on the synapse to elicit a change of its weight. In general the change of these weights depends on the activity of different ligand-receptor complexes. By varying the place and time of these interactions, the ES is able to explore learning mechanisms without the need of the designer to declare a set of learning rules (e.g., Hebbian, Anti-Hebbian), because such rules can be evolved. To explain a specific example how a Hebbian learning rule could be evolved, one possible solution will be discussed in more detail. Neurons can link the emission of a signaling molecule to their neural activity. If now the emitted transmitter has an affinity to one of two receptors on the synapse and the second neuron emits a substance for a second receptor on the synapse, then the weight will be increased in an AND fashion, a Hebbian learning rule is found. As a functioning neuron has already a connection only three new parameters with the right affinities have to be found which is an easy task for an ES. To change now this solution to an Anti-Hebbian rule only the effect on the weight has to be change and can be done by a single mutation. In the beginning of the experiments random weights are assigned to the synapses where the values assigned have a Gaussian distribution. The value system which was evolved is described in figure 4.

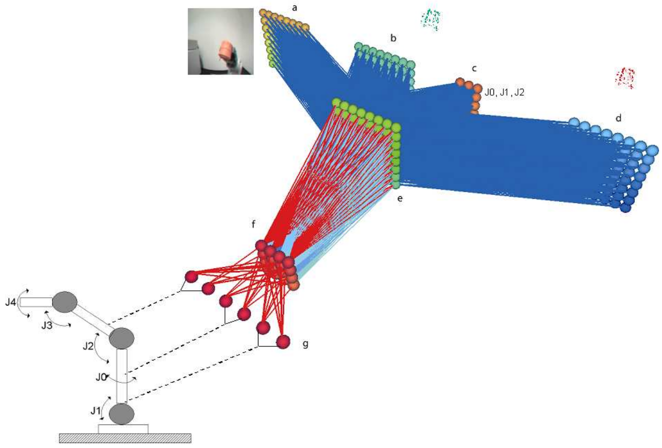

Figure 5: Neural structure and its connections to the robot arm. Neuronal areas: (a) RedColorField. (b) RedMovementToRightField. (c) ProprioceptiveField. (d) RedMovementToLeftField. (e) NeuronalField. (f) MotorField. (g) MotorActivites.

As earlier results [7, 8] showed that the system was also able to find learning mechanisms by influencing other neurons over distance. One of the evolved rules was that active neurons controlling the robot arm were "rewarded" (e.g., if the movement of the arm brought the colored object closer to the center of the visual field) by emitting an artificial molecule signaling the synapses to change their weights. This evolved mechanism (see figure 4) is quite similar to a value system, which are neural structures that are necessary for an organism to modify its behavior based on the salience or value of an environmental cue [3, 10, 19]. In contrast to their hand-designed method, where the value system is located in other brain parts, our evolved value system is located directly in the network performing the task.

2.3 Evolutionary Strategy

To vary and select the genomes an evolutionary strategy (ES) was used. The ES was developed by Rechenberg [18]. In the evolutionary experiments a (μ, λ)-evolution strategy with one global step size σ for each individual was used. (The notation indicates that the procedure generates λ offsprings from μ parents and that it selects the parents only from the best

Table 1: Neural Structure

Field name	Number of neuronal units
RedColorField	64
RedMovementToLeftField	64
RedMovementToRightField	64
ProprioceptiveField	24
NeuronalField	64
MotorField	16
MotorActivities	6

offspring for the next generation).

3 Robotic setup

Our experiments were performed using the experimental setup showed in figure 1. It consists of an industrial robot manipulator (Mitsubishi MELFA RV-2AJ) with six degrees of freedom and a stereo color active vision system. The system architecture consists of two computers Pentium III/600 MHz and the robot arm controller connected together in a private local area network (LAN) based on the TCP/IP protocol, one computer controls the robot arm and the other acquires the visual input from the active vision system using two (Hauppauge) frame grabbers, with a resolution of 128x128 pixels down-sampled at a rate of 20Hz. The communication between the computer and the motor control board that drives the active vision system uses the standard RS-232 protocol. The sensory-motor control board is based on the Hitachi H8/3664 chip from the H8/Tiny series. The operating system is Linux and the programming language is C/C++.
As can be seen in the figure 5, joint J0 (”shoulder”) is responsible for the rotation around the vertical axis; joint J2(”elbow”), joint J1 (”shoulder”) and joint J3 (”wrist”) are responsible for the up and down movements; joint J4(”wrist”) rotates the gripper around the horizontal axis. The additional degree of freedom comes from the gripping manipulator.

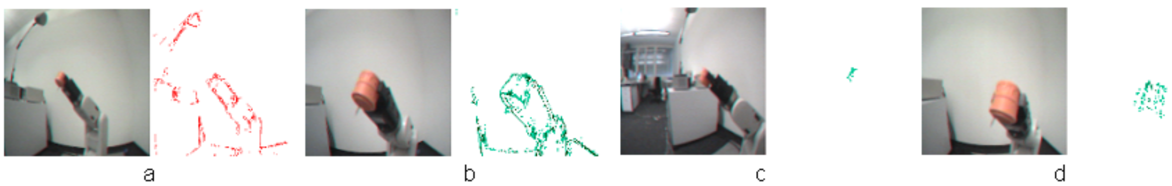

Figure 6: Motion detection. (a) Movement was detected from right to left (negative values are showed in red). (b) Movement was detected from left to right (positive values are showed in green). (c) and (d) Motion detectors reacting only to red objects moving in the environment.

4 Neural structure

The components of the neural structure and its connections to the robot arm can be seen in the figure 5, and the number of neuronal units in each neuronal area can be found in table 1.

4.1 Sensory Field

- *Color information*. Three receptor types were considered: red (r), green (g), and blue (b). An attenuation of the influence of changing lighting conditions was achieved through a color-space transformation, a "broadly" color-tuned channel of size 8x8 was created for red: *R = r - (g + b) / 2*. This channel yields maximal response for the pure, fully-saturated red color, and yields zero response for black (*r=g=b=0*) and white (*r=g=b=255*) inputs. The negative values were set to zero. Each pixel was then mapped directly to neuronal units of area *redColorField*, which activity was calculated according to the following formula:
$$S_i = \begin{cases} 1.0 & : \quad if \;\; R_i > \theta_1 \\ 0.0 & : \quad otherwise \end{cases} \tag{1}$$
Where S_i is the activity of the i-th neuron on area *RedColorField*; R_i is the value of the red color-tuned channel for the pixel i-th; and θ_1 is a threshold value.

- *Motion detection*. We have developed a cheap bio-inspired motion detection mechanism that can be seen in the figure 6. It is based in the well-known elementary motion detector (EMD) of the spatio-temporal correlation type. A description of the model implemented can be found in [13]. Two motion detectors of size 8x8 were created to detect movements of red objects in the environment. Motion detectors reacting to red objects moving to the left side of the image were mapped directly to neuronal units of the area *RedMovementToLeftField*, and the motion detectors reacting to red objects moving to the right side of the image were mapped directly to neuronal units of the area *RedMovementToRightField*. The activities of the neurons in areas *RedMovementToLeftField* and *RedMovementToRightField* were calculated according to the following formula:
$$S_i = \begin{cases} 1.0 & : \quad if \;\; |EMDOutput_i| > \theta_2 \\ 0.0 & : \quad otherwise \end{cases} \tag{2}$$
Where S_i is the activity of the i-th neuron; $EMDOutput_i$ is the output of the motion detector at position i-th; and θ_2 is a threshold value.

- *Propioceptive information*. The neural area *ProprioceptiveField* has a size of 8x3 and its neuronal units encode the position of the joints: J0 ("shoulder"), J1 ("shoulder") and J2("elbow").

4.2 Neuronal Field and Motor Field

The size of the *NeuronalField* is 8x8 and its neuronal units have a sigmoid activation function. The size of the *MotorField* is 4x4 and its neuronal units have a sigmoid activation function, which outputs are passed directly to the *MotorActivities*, which size is 6x1, for controlling the joints of the arm: J0, J1 and J2.

4.3 Synaptic connections

Neuronal units in the areas *RedColorField*, *RedMovementToLeftField*, and *RedMovementToRightField* were connected retinotopically to the neuronal units in area *NeuronalField*. The neuronal units in the area *proprioceptiveField* were fully connected to the neuronal units in area *NeuronalField*. The neuronal units in area *NeuronalField* were fully connected to the neuronal units in area *MotorField*, which in turn were fully connected to the *MotorActivities*. All the initial connection strengths were set randomly. The maximum number of synaptic connections is 14,944. See figure 5 and table 1.

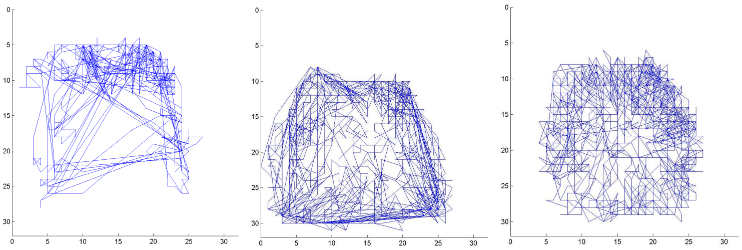

Figure 7: Random experiments. Position of the center of the object in the visual field during random exploration. From left to right: (a) Experiment number 1, with a duration of 881 time steps. (b) Experiment number 2, with a duration of 1,780 time steps. (c) Experiment number 3, with a duration of 2,500 time steps.

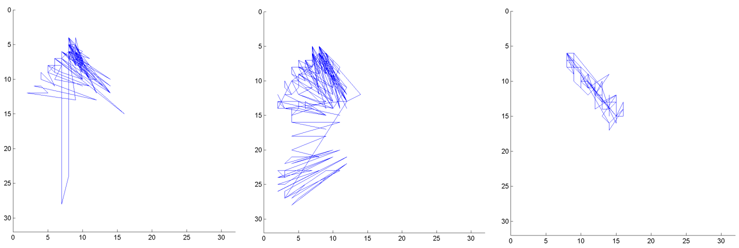

Figure 8: Experiment number 4. From left to right: position of the center of the object in the visual field during the learning cycles in the interval (a) [1, 300]. (b) [301, 600]. (c) [601, 881].

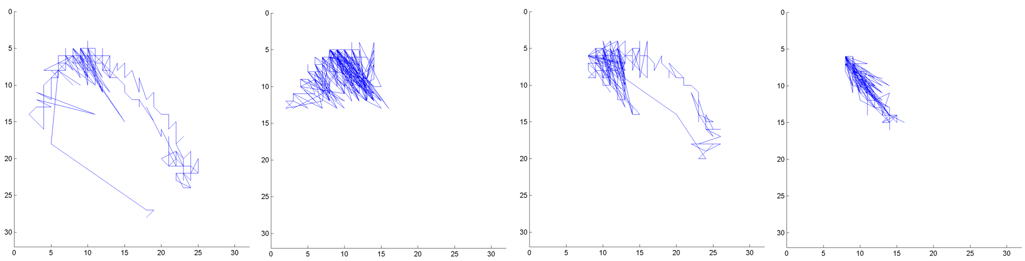

Figure 9: Experiment number 5. From left to right: position of the center of the object in the visual field during the learning cycles in the interval (a) [1, 400]. (b) [401, 800]. (c) [801, 1200]. (d) [1201, 1602].

5 Experiments and results

At the beginning of each experiment the robot arm was initialized at a random position at the periphery of the robot's visual field and a colored object was put on its gripper, as can be seen in figure 1. In order to compare the effects of the learning mechanism different sets of experiments were performed. The figure 7 shows the results of three different non sensory motor

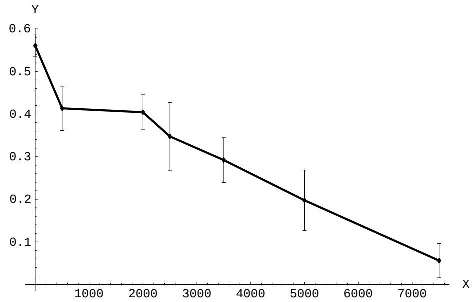

Figure 10: Error function. Distance from the current position of the object to the center of the visual field vs the learning cycles.

coordinated experiments, in these experiments the learning mechanism was not activated and therefore the robot arm made only random movements without being able to bring the object to the center of the visual field.
In contrast when the learning mechanism was activated, the robot was engaged in a sensory motor coordinated loop, the figures 8 and 9 show the results of the experiments number 4 and 5 respectively. In these experiments the robot was able to solve the task bringing the object to the target position several times in a stabilized trajectory. The error function can be seen in figure 10 when the number of degrees of freedom was increased from two (J0 and J2) to three (J0, J1 and J2).

6 Discussion

With the above results we could show that the evolved learning mechanism can be transferred to a robot arm and that the task could be learned by the robot. That the eyes can teach the robot arm to bring an object to the center of the visual field for a possible further and detailed exploration. This was possible because the learning mechanism was able to evaluate at each moment, if a movement was well performed or not. This decreased the learning space considerably and was one of the reasons because the robot could learn its task in a reasonable amount of time (1 hour and 10 minutes for 7500 learning cycles using three degrees of freedom). Furthermore we learned (already from the other paper [8] that different value systems exist and that some are better than others because they can reduce the learning space considerably. Having a neural network in which the value system is part of it, the explanation how value systems evolved can be given more easily. These results point to a new direction of how to overcome the reality gap: the goal should be to evolve learning mechanisms (not fixed neural networks) that can control the performance of a robot interacting with the real-world. If the designer can guarantee that the feedback loop needed to learn a task is functioning, the robot can learn to behave in reality as well as in simulation which we exemplified by our results.

7 Acknowledgements

Peter Eggenberger was sponsored by the EU-Project HYDRA (IST-2001-33060), Gabriel Gómez was supported by the grant NF-10-101827/1 of the Swiss National Science Foundation and the EU-Project ADAPT (IST-2001-37173).

References

[1] Angelo Cangelosi, Domenico Parisi, and Stefano Nolfi. Cell division and migration in a 'genotype' for neural networks. *Network*, 5:497–515, 1994.

[2] Frank Dellaert and Randall D. Beer. A developmental model for the evolution of complete autonomous agents. In [16], pages 393–401, 1996.

[3] Gerald M. Edelman. *Neural Darwinism*. New York:Basic Books, 1987.

[4] Peter Eggenberger. Cell interactions as a control tool of developmental processes for evolutionary robotics. In [16], pages 440–448, 1996.

[5] Peter Eggenberger. Creation of neural networks based on developmental and evolutionary principles. In [11], pages 337–342, 1997.

[6] Peter Eggenberger. Evolving neural network structures using axonal growth mechanisms. In *International Joined Conference on Neural Networks*, 2000.

[7] Peter Eggenberger. On evolving a simulated foveating retina. *3D Forum: The Journal of Three Dimensional Pictures*, 15(4):164–169, 12 2001.

[8] Peter Eggenberger, Gabriel Gómez, and Rolf Pfeifer. Evolving the morphology of a neural network for controlling a foveating retina - and its test on a real robot. In *Standish, R. K., Bedau, M. A., and Abbass, H. A., editors, Artificial Life VIII: Proceedings of The Eight International Symposium on Artificial Life*, pages 243–251, 2002.

[9] Dario Floreano and Francesco Mondada. Evolution of homing navigation in a real mobile robot. In [16], pages 402–4101, 1996.

[10] K. J. Friston, Guilio Tononi, G. N. Reeke, Olaf Sporns, and Gerald M. Edelman. Value-dependent selection in the brain: Simulation in a synthetic neural model. *Neuroscience*, 59(2):229–243, 1994.

[11] Wulfram Gerstner, Alain Germond, Martin Hasler, and Jean-Daniel Nicoud (Eds.), editors. *Seventh International Conference of Artificial Neural Networks (ICANN'97)*, Cambridge, MA, 1997. Sprinter.

[12] Frederic Gruau and D. Whitley. The cellular developmental of neural networks: the interaction of learning and evolution. Technical Report 93-04, Laboratoire de l'Informatique du Parallélisme, Ecole Normale Supérieure de Lyon, France, 1993.

[13] Fumiya Iida and Dimitrios Lambrinos. Navigation in an autonomous flying robot by using a biologically inspired visual odometer. In *Sensor Fusion and Decentralized Control in Robotic System III*, volume 4196, pages 86–97. Proceeding of SPIE, Photonics East, 2000.

[14] Akio Ishiguro, Akinobu Fujii, and Peter Eggenberger Hotz. Neuromodulated control of bipedal locomotion using a polymorphic cpg circuit. *Adaptive Behavior*, 11(1):7–17, 2003.

[15] Jérôme Kodjabachian and Jean-Arcady Meyer. Evolution and development of neural controllers for locomotion, gradient-following, and obstacle-avoidance in artificial insects. *IEEE Transactions on Neural Networks*, 9(5):796–812, 1998.

[16] Pattie Maes, Maja J. Mataric, Jean-Arcady Meyer, Jordan Pollack, and Stewart W. Wilson, editors. *From animals to animats 4: Proceedings of the fourth international conference on simulation of adaptive behavior*. MIT Press, 1996.

[17] Stefano Nolfi and Dario Floreano. Neural synthesis of artificial organisms through evolution. *Trends in Cognitive Science*, 2002.

[18] Ingo Rechenberg. *Evolutionsstrategie' 94*. frommann-holzboog, 1994.

[19] Olaf Sporns and William H. Alexander. Neuromodulation and plasticity in an autonomous robot. *Neural Networks*, 15:761–774, 2002.

[20] Jari Vaario and Katsunori Shimohara. Synthesis of developmental and evolutionary modeling of adaptive autonomous agents. In [11], pages 721–726, 1997.