Developing a Reaching Behaviour in an simulated Anthropomorphic Robotic Arm Through an Evolutionary Technique

Gianluca Massera[1], Angelo Cangelosi[2], Stefano Nolfi[1]
[1]Institute of Cognitive Science and Technologies, National Research Council (CNR), Via S. Martino della Battaglia, 44, 00185, Roma, Italy
[2]Adaptive Behaviour & Cognition Research Group, University of Plymouth, Drake Circus, Plymouth
gianluca.massera@istc.cnr.it, stefano.nolfi@istc.cnr.it, angelo.cangelosi@plymouth.ac.uk

Abstract

In this article we present an evolutionary technique for developing a neural network based controller for an anthropomorphic robotic arm with 4 DOF able to exhibit a reaching behaviour. Evolved neural controllers display an ability to reach targets accurately and generalize their ability to moving targets. This study demonstrates that it is possible to obtain solutions that are extremely parsimonious from the point of view of the control system. Evolutionary training techniques allow us to evolve parameters of the control system on the basis of the global effects that they produce on the dynamics arising from the interaction between the control system, the robot's body and the environment.

1. Introduction

The control of arm and hand movements in human and non-human primates is a fascinating research topic in robotics and cognitive science.

In robotics, the design of adaptive robotic systems able to perform complex object manipulation tasks is one of the most important research issues (Schaal, 2002).

In cognitive science, the relationship between action control and other cognitive functions has been demostrated to be important in the study of cognition (Pulvermuller, 2005; Cangelosi et al., 2005). For example, variour theories of language evolution have focused on the relationship between hand use, tool making and language evolution (Corballis, 2003).

Within arm control, reaching and grasping behaviours represent key abilities since they constitute a prerequisite for any object manipulation. Despite the importance of the topic, the large body of available behavioural and neuropsychological data, and the vast number of studies based a variety of AI and neural network techniques, the issues of how primates and humans learn to display reaching and grasping behaviour still remains highly controversial (Schaal, 2002; Shadmehr, 2002). Similarly, while many of the aspects that makes these problems difficult have been identified, experimental research based on different AI and neural networks techniques does not seem to converge toward the identification of a single general methodology.

In this article we present an evolutionary technique for developing a neural-network based controller for a simulated anthropomorphic robotic arm able to exhibit a reaching behaviour.

In section 2, we define what we mean by reaching behaviour in the context of arm control and we discuss the aspects that make this problem hard to solve. In section 3, we point out the relation of our approach with the other related models. In section 4, we describe our experimental set-up and the method used to develop the control system of a simulated anthropomorphic robotic arm. In section 5, we describe the simulation experiments and results. Finally, in section 6, we will present our conclusions and our future plans.

2. Reaching

Primate arms consist of three segments (the arm, the forearm, and the hand) attached to previous segments (the shoulder, the arm, and the forearm) through three actuated joints (the shoulder, elbow, and wrist joints). Roughly speaking, human arms have seven limited degrees of freedom (DOFs): three in the shoulder, one at the elbow, and three at the wrist. Anthropomorphic robotic arms typically consist of three segments connected through motorized joints. Some models use all the seven DOFs listed above, others may include only part of them.

From the point of view of the control system, reaching consists in producing the appropriate sequence of motor actions (i.e. setting the appropriate torque force for each actuated joint) that, given the current state of the arm and given the current desired target point, will bring the endpoint of the arm in the current desired target position.

Some of the most important issues in the study of reaching behaviour are:

- When the number of DOFs is redundant (as in the case of primate arms), there is an infinite number of trajectories and of final postures for reaching any given target point. This redundancy potentially allows anthropomorphic arms to reach a target point by circumventing obstacles or by overcoming problems due to the limits of the DOFs. However, the redundancy of DOFs, also, implies that the space to be searched during learning is rather vast.

- Anthropomorphic arms are highly non-linear systems. First, small variations in some of the joints might have a huge impact on the end-position of the arm. At the same time, significant variations of other joints might not have any impact. Secondly, due to the limits on the joints' DOFs and due to the interactions between joints, similar target positions might require rather different trajectories and final postures. At the same time, rather different target positions might require similar trajectories and final postures.

- In articulated and suspended structures such as anthropomorphic arms, gravity and inertia play a key role. In primate arms, muscles and associated spinal reflex circuitry seems to confer to the arm the ability to passively settle into a stable position (i.e. an equilibrium point) independently from its previous position. If this hypothesis is true, the contribution of the central nervous system would simply consists in the modification of the current equilibrium point (Shadmehr, 2002).

- Sensors and actuators might be slow and noisy. For instance in humans visual information and proprioceptive information encoding changes of joints positions is available with a delay up to 100ms. Motor commands issued by the central nervous system may take up to 50ms to initiate muscle contraction (Mial, 2002). Moreover, sensors might provide only incomplete information (e.g. the target point might be partially or totally occluded by obstacles and by the arm).

3. The State of the Art

There have been few previous attempts to use evolutionary techniques to develop the controller for a robotic arm.

Bianco and Nolfi (2004) used a similar approach to that described in this paper to develop the controller for a simulated robotic arm with a two-fingered hand and nine DOFs for the ability to grasp objects with different shapes. The arm was only provided with tactile sensors. Evolved robots displayed an ability to grasp objects with different shapes, different orientations, and located in varying positions within a limited area. Evolving robots, however, were not able to deal with larger variations of the objects positions. Indeed, in this paper we used a similar method to solve the reaching problem and we plan to combine the two approaches in future research to develop robotic arms that can effectively reach and grasp objects in a large variety of circumstances.

Buehrmann and Di Paolo (2004) evolved the control system for simulated robotic arm with three DOFs for the ability to reach a fixed object placed on a plane and to track moving objects. The arm was provided with two pan-tilt ”cameras” consisting of a two-dimensional array of ”laser range sensors” placed above the robot arm and on the endpoint of the robotic arm. The controller consisted of several separate neural modules. These receive different sensory information and control different motor joints. The networks are separately evolved for the ability to produce different elementary behaviours (e.g. change the orientation of the above camera so to focus on the object, move the first joint that determines the orientation of the arm so to orient toward the object, approaching the object by controlling the second and the third joint, etc.).

In the work described in this paper, we do not focus on the vision system. Indeed, we assume that a pre-existing vision system can provide to the evolved controller the offset between the target point and the endpoint of the arm. Moreover, rather than on an standard industrial type robotic arm with three DOFs, we study the case of a realistic anthropomorphic arm with four DOFs. This is quite a different system in which each target point can be reached through an infinite number of postures and in which the relation between the joint reference system and the Cartesian reference system are much more complex and indirect.

Finally, rather than relying on an incremental approach in which elementary components of the required behaviour are identified by the experimenter, we select individuals only on the basis of their ability to reach the desired target point by letting them free to develop their own strategy to solve the problem.

4. Experimental set-up

The aim of this study is to develop the control system for an anthropomorphic robotic arm through an evolutionary robotic technique (Nolfi and Floreano, 2000). The arm and the arm/environmental interaction have been simulated using ODE (Open Dynamics Engine www.ode.org), a library for accurately simulating rigid body dynamics and collisions.

The control system consists of a simple neural network that controls directly the direction and the intensity of the forces that are applied to the motorized joints. Neural controllers are selected for their ability to reach the desired target positions and are left free to determine the way in which the problem is solved (i.e. the trajectory and the posture of the arm).

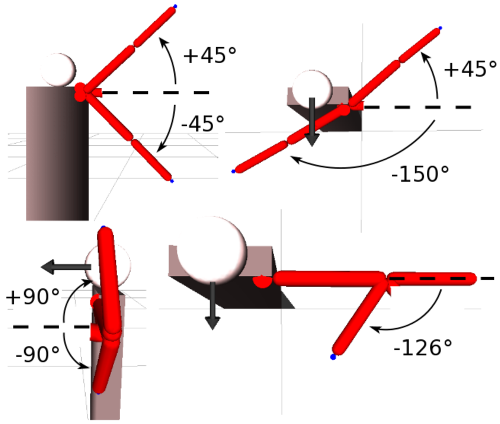

Figure 1: The four DOF of the simulated robotic arm. The two pictures on the top part of the figure indicate the abduction/adduction, extension/flexion of the shoulder joint, respectively from left to right. The bottom figure indicates the rotation of shoulder and the extension/flexion DOF of the elbow. The arrows indicates the frontal direction of robot.

The simulated robotic arm

The simulated robot consists of cylindrical segments articulated by revolute joints, as illustrated in Figure 1. More specifically, the arm consists of two segments (the arm and the forearm) that are attached to the previous segments (the shoulder and the arm) through two joints (the shoulder and the elbow joints). The arm and the forearm have a length of $100cm$ and $80cm$, a diameter of $8cm$ and $7cm$, and a weight of $13kg$ and $8kg$ respectively. The shoulder has three DOF that allow abduction/adduction of $[-45^o, +45^o]$, extension/flexion of $[-150^o, +45^o]$ and rotation of $[-90^o, +90^o]$. The elbow has one DOF that allow extension/flexion of $[-126^o, +0^o]$. Since the robot is only asked to reach a given target position with the endpoint of its arm, we did not modelled the wrist and the wrist joints. Therefore the arm has four motorized joints and four DOF (Figure 1). The acceleration of gravity has been set to $9.8m/s^2$. The robot sensory system includes a simulated vision system that detect the angle and the distance between endpoint of arm (hand) and the target point.

The neural controller

The neural controller consists of a feedforward neural network with 3 sensory neurons directly connected to 4 motor neurons. The four motor neurons are updated on the basis of a standard logistic function. The activation of the sensory and motor neurons is updated every $0.015sec$. The three sensory neurons encode the distance, along the three axes, between the endpoint of the arm and the target point normalized in the range $[-1, +1]$ and up to a maximum distance of $80cm$. The four motor neurons, that are updated on the basis of a standard logistic function, encode the angular velocity of the four corresponding motorized joints. The activation of the output neurons is normalized in the $[-890, +890]\,rpm$ range. The power of motors is set to $326W$.

The evolutionary algorithm

The connection weights of the neural controller have been evolved (Nolfi and Floreano, 2000). The genotype of evolving individuals encodes the connections weights of the neural controller (each connection weights is encoded with 16 bits and normalized in the range $[-10, 10]$). Population size is 100. The 20 best individuals of each generation were allowed to reproduce by generating 5 copies with 1.5% of their bits replaced with a new randomly selected value (reproduction is asexual). The evolutionary process lasted 1000 generations. The experiment was replicated 10 times starting from different, randomly generated, genotypes.

Each individual of the population was tested for 16 trials, with each trial consisting of 300 steps corresponding to $4.5sec$. At the beginning of each trial the arm is set in a random position (i.e. the area of possible angles in the joint-space is divided in 16 non-overlapping sub-areas; for each trials a random joint configuration is picked up from one of that sub-areas) and the target is positioned in front of the robot Figure 1 (at a distance $1m$ and $85cm$ from "head" along the horizontal and vertical planes, respectively). Evolving robots are selected on the basis of their capacity to reach the target point as fast as possible and stay on it. In details, the fitness function selects robots that minimize the cumulative sums over 300 steps of the follow function:

$$dist(x,r) = \begin{cases} 100 & \text{if } x < r \\ 100 \cdot e^{(-0.5*(x-r))} & \text{if } x \geq r \end{cases} \quad (1)$$

where x is the euclidean distance between endeffector of the arm and the target point, and r is a threshold (initially set to 10cm and progressively reduced of 10% during the evolutionary process, each time the average fitness of the individuals overcome 78 units).

5. Results

By running the experiments we observed that, in all replications, evolved agents display an ability to reach the target independently from their initial posture and to produce rather accurate reaching behaviour.

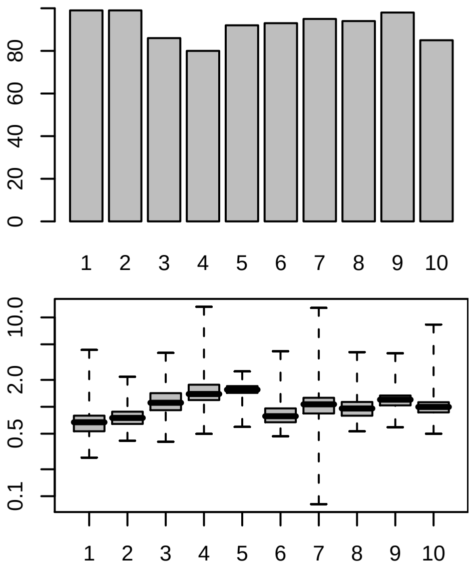

Figure 2: Performance on reaching a fixed target; **Top:** Percentage of trials in which the distance between the endpoint of the arm and the target is below 1*cm*, at the end of the trial. **Bottom:** Average distance between the endpoint of the arm and the target at the end of trials. Each column represents the performance obtained by testing the best evolved individual of each replication for 100 trials. Bold lines, grey hystograms and bars indicate average performance, variance, and mininum and maximum values, respectively

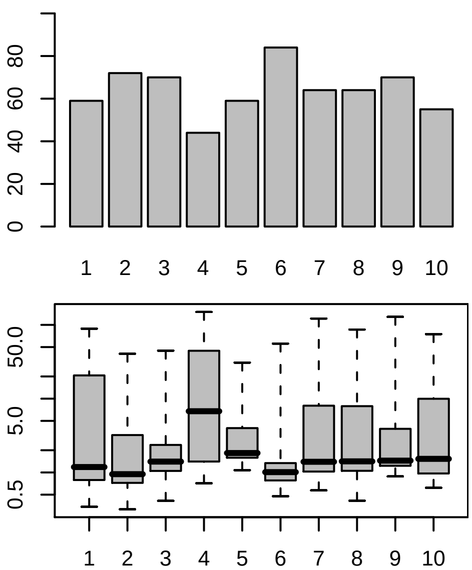

Figure 3: Performance on reaching a random positioned target; **Top:** Percentage of trials in which the distance between the endpoint of the arm and the target is below 1*cm*, at the end of the trial. **Bottom:** Average distance between the endpoint of the arm and the target at the end of trials. Columns, hystograms, bars have the same meanings of Figure 2.

Figure 2 shows, for each replication, the percentage of successful reaching behaviour and the average distance between the endpoint of the arm and the target, at the end of each trial. Reaching behaviour are considered successful when the distance between the target and the endpoint of the arm is less than 1*cm*.

Generalization

The evolved ability also generalize to different positions of the target and to moving targets. Figure 3 shows the performance of evolved robots tested with target placed in randomly selected locations (within a distance of 200*cm* with respect to the fixed location of the target used during the evolutionary process). As shown in the Figure performance significantly vary in different replications. In the case of the best replication, however, performance are only slightly worse with respect to the normal condition (see Figure 2).

Figure 4 shows the results obtained by testing evolved individuals with 125 targets points evenly distributed in front of the robot on a $5 \times 5 \times 5$ grid (for space reason we only report the data for two typical evolved individuals). For each target point individuals have been tested for 5 trials starting from differently, randomly assigned, initial positions. As can be seen performance qualitatively vary in different individuals.

Indeed, the individual represented in the top graph shows slightly better performance in the central and distant areas than in the near area. The individual represented on the bottom graph, instead, shows close to optimal performance in the left area and significantly worse performance in the right area.

This qualitatively different performance can be explained by considering that the four DOF are strongly interdependent. This clearly indicates that strategies that treat each joint as an independent entity (that should be moved so to reduce the distance with respect to the target independently from the current position of the other joints) are insufficient. Evolving robots should select control strategies that minimize the problems resulting from the high interdependence between the DOF.

Figure 5 shows the behaviour produced by one of the best evolved individual that try to reach a target that moves by following a circular and a eight-shaped trajectory. Also in this case, although evolving robots were selected for the ability to reach a fixed target, the robot generalizes their ability to moving targets quite well (Figure 5).

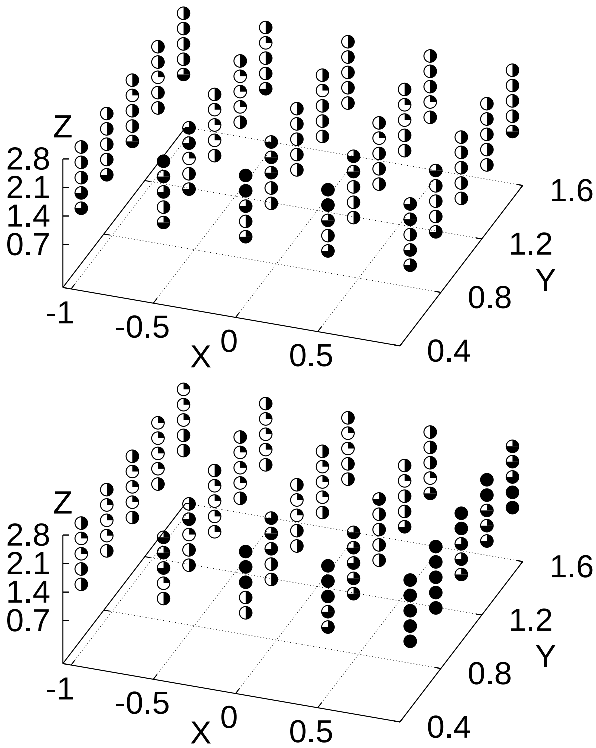

Figure 4: Performance obtained by testing with 125 targets points evenly distributed in front of robot on a $5 \times 5 \times 5$ grid area. The top and bottom graphs report the result obtained by testing two typical evolved individuals. The filled area of each bullet indicates the average distance between the target area and the endpoint of the arm in the following intervals: $< 1cm$ ◔, $[1, 10]cm$ ◑, $[10, 50]cm$ ◕, > 50 ●. The two axis indicate the position of the target points along the vertical and horizontal dimensions in meters.

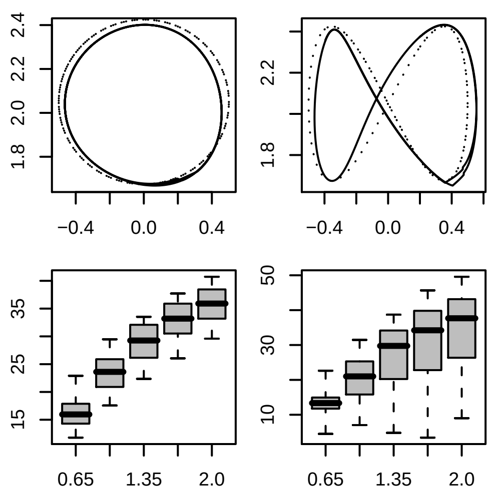

Figure 5: **Top:** trajectory produced by the endpoint of the arm and by a moving target (solid and dotted lines, respectively). Results obtained in two tests in which the target move by displaying a circular and an eight-shape trajectory (left and right picture, respectively). The vertical and horizontal axis indicate the positions of the target and of the end-point of the arm in meters. **Bottom:** average distance between the target point and the end-point of the arm during the tests for target moving at different speed (ranging from 0.65 to $2.0 m/s$).

Finally, by testing evolved individuals in a control condition in which the update of the sensory neurons is delayed, we observed that performance decreases gracefully with delays from 60 to $150ms$ (see Figure 6).

Surprisingly, performance increases with a delay of $30ms$ and remains almost constant with a delay of $15ms$. By replicating the evolutionary process in a condition in which the update of the sensory neurons is delayed of $105ms$, we observed that obtained performance are very similar to those obtained in the first evolutionary experiment without delay. In fact, the percentage of trials in which the distance between the endpoint of the arm and the target is below $1cm$ is 91.2% and the average distance between the target at the endpoint of the arm is $1.34cm$. Without sensory delay these data are 92.2% and $1.31cm$, respectively (see Figure 2). Also in the condition in which the update of the sensory neurons is delayed, evolved robots generalize their ability to target located in varying positions (within limits). In this test condition, the average number of successful reaching behaviour and the average distance between the endpoint of the arm and the target are 62.7% and $6.56cm$, respectively. Performance without sensory delay are 64.1% and $9.81cm$, respectively (see Figure 3). All these data refer to the average performance of the best individuals of the 10 replications of the experiment.

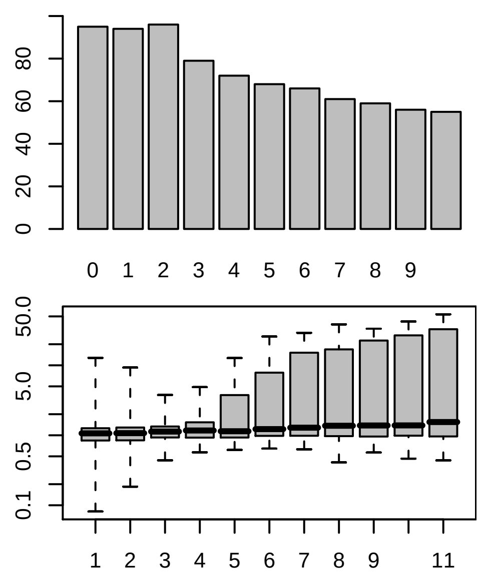

Figure 6: Performances obtained by testing robots evolved in a normal condition in a test condition in which the update of the sensory neurons is delayed. **Top:** Percentage of trials in which the distance between the endpoint of the arm and the target is below $1cm$, at the end of the trial. **Bottom:** Average distance between the endpoint of the arm and the target at the end of trials. Columns and bars have the same meanings of Figure 2. The x axis indicate the sensory delay (in multiples of 15ms)

Analyzing evolved trajectories

To analyze how much the trajectories produced by evolved individuals approximate hand-made trajetories produced by moving the joints toward the values corresponding to the final postures (produced by evolved individuals) we tested evolved robots for 16 trials starting from randomly set initial position (i.e. arm postures). For each trial we:

1. allowed the arm to move on the basis of the evolved neural controller. During this first phase, we recorded the initial and the final posture and the vector of positions of the endpoint of the arm during motion;

2. we placed the arm in the same initial posture of the previous phase and we manually set the desired position of the joints on the basis of the final posture produced in the previous phase. The maximum velocity was sets to $890rpm$, i.e. the same value used for controlling the arm during the first phase. During this second phase, we recorded the vector of positions of the endpoint of the arm during motion;

3. we measured the average difference between the positions produced during the first and the second phase in each time step.

The fact that differences are rather small (Figure 7) indicates that the trajectories produced by evolved robots are quantitatively similar to those that can be obtained by minimizing the movements of the joints.

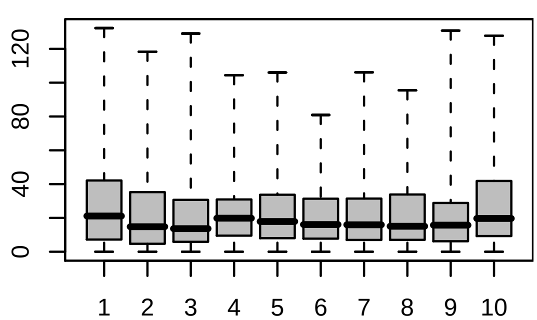

Figure 7: Average distance in cm between the trajectories produced by an evolved neural controller and the trajectories produced by manually setting the desired position of the joints on the basis of the final postures produced by the evolved neural controller. Each column indicates the result obtained for the best individual of a corresponding replication of the experiment. Bold line, grey boxes, and dotted lines indicate the average the variance, and the minimum and maximum values, respectively.

6. Discussion

The problem of controlling a robotic arm is often approached by assuming that the robot should posses, or should acquire through learning, an internal model to: (a) predict how the arm will move and the sensations that will arise, given a specific motor command (direct mapping), and (b) transform a desired sensory consequence into the motor command that would achieve it (inverse mapping) - for a review see Torras (2002).

We do not deny that primates rely on internal models of this form to control their motor behaviour. However, this does not necessarily implies that elementary movements are learned on the basis on a detailed description of the sensory-motor effects of any given motor command and of a detailed specification of the desired sensory states. Direct and inverse mapping might operate at a higher level of organization, for example might play a role in the determination of the specific elementary behaviour to be triggered in a specific circumstance.

Assuming that natural organisms act on the basis of a detailed direct and inverse mapping at the level of micro-actions (i.e. at the level of the elements that constitute elementary behaviours) is implausible for at least two reasons. The first reason is that sensors provide only incomplete and noisy information about the external environment and moreover, muscles have uncertain effects. The former aspect makes the task of producing a detailed direct mapping impossible, given that this would require a detailed description of the actual state of the environment. The latter aspect makes the task of producing an accurate inverse mapping impossible given that the sensory-motor effects of actions cannot be fully predicted. The second reason is that the environment might have its own dynamic and typically this dynamic can be predicted only to a certain extent. For these reasons, the role of the internal models is probably limited to the specification of macro-actions or simple behaviours, rather than to micro-actions that indicate the state of the actuators and the predicted sensory state in any given instant.

This leaves open the question of how simple elementary behaviour might be learned, i.e. how individuals might learn to produce the right micro-actions that lead to a desired elementary behaviour. One possible hypothesis is that elementary behaviours (e.g. reaching a certain class of target points in a certain class of environmental conditions) are produced through simple control mechanisms that exploit the emergent result of fine grained interactions between the control system of the organism, its body and the environment. From this point of view, simple behaviours might be described more effectively through dynamical system methods that identify limit cycle attractors and the effects of parameters variation on the agent/environment dynamics (Sternad and Schaal, 1999).

In this paper we demonstrated how effective reaching behaviours can be developed through a training procedure in which variations, in the parameters of the control system, are retained or discarded on the basis of the global effects that they produce on the dynamics arising from the interaction between the control system, the robot's body and the environment (Nolfi and Floreano, 2000). Moreover, our results indicate that the possibility to discover and retain characters that lead to useful emergent properties (through a process bases on random variation and selection), allow to find solution that are extremely parsimonious from the point of view of the control system.

In future work we plan to: (a) introduce costs in the fitness function which are analogous to well known optimization principles like minimum variance or minimum jerk (Jordan and Wolpert, 1999) by eventually providing the robots with more complex neural controllers, (b) combine the reaching abilities described in this paper with the grasping ability based on tactile information described in Bianco and Nolfi (2004) and (c) extend this model into cognitive robotic agents to investigate the relationship between motor and other linguistic and cognitive capabilities (Marocco et al., 2003; Cangelosi et al., 2005).

Indeed, we believe that the main reason that explain why we obtained such robust and effective results on the basis of extremely simple neural controllers resides in the methodology that we used in which variation in the free parameters of the control system (that regulate the interaction between the agent and the environment at the micro-level) are retained or discarded on the basis of their affects at the macro-level (i.e. the level of behaviour). This methodology, in fact, allow the discovery and the retention of useful properties emerging from the interaction between the robots' controller, its body, and the environment (Nolfi, in press).

Acknowledgments

This research has been supported by MIUR (Italian Ministry of Education, University and Research) within the project "Azione e percezione nella costruzione del mondo cognitivo".

References

Bianco, R. and Nolfi, S. (2004). Evolving the neural controller for a robotic arm able to grasp objects on the basis of tactile sensors. *Adaptive Behavior*, 12(1):37–45.

Buehrmann, T. and Di Paolo, E. A. (2004). Closing the loop: Evolving a model-free visually-guided robot arm. In *Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems (ALIFE9)*. Boston, Cambridge, MA: MIT Press.

Cangelosi, A., Bugmann, G., and Borisyuk, R., editors (2005). *Modeling Language, Cognition and Action: Proceedings of the 9th Neural Computation and Psychology Workshop*. Singapore: World Scientific.

Corballis, M. C. (2003). *From Hand to Mouth: the Origins of Language*. Princeton University Press.

Jordan, M. and Wolpert, D. (1999). Computational motor control. In (Ed.), M. G., editor, *The Cognitive Neurosciences, 2nd edition*. Cambridge, MA, MIT Press.

Marocco, D., Cangelosi, A., and Nolfi, S. (2003). Evolutionary robotics experiments on the evolution of language. *Philosophical Transactions of the Royal Society of London*, A 361:2397–2421.

Mial, R. C. (2002). Motor control, biological and theoretical. In Arbib, M. A., editor, *Handbook of brain theory and neural networks, Second Edition*, pages 110–113. Cambridge, MA: MIT Press.

Nolfi, S. (in press). Behaviour as a complex adaptive system: On the role of self-organization in the development of individual and collective behaviour. *Complex Us*.

Nolfi, S. and Floreano, D. (2000). *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. Cambridge, MA: MIT Press/Bradford Books.

Pulvermuller, F. (2005). Brain mechanisms linking language and action. *Nature review Neuroscience*, 6:576–582.

Schaal, S. (2002). Arm and hand movement control. In (Ed.), M. A. A., editor, *Handbook of brain theory and neural networks, Second Edition*, pages 110–113. Cambridge, MA: MIT Press.

Shadmehr, R. (2002). Equilibrium point hypothesis. In (Ed.), M. A. A., editor, *Handbook of brain theory and neural networks, Second Edition*, pages 409–412. Cambridge, MA: MIT Press.

Torras, C. (2002). Robot arm control. In Arbib, M. A., editor, *Handbook of brain theory and neural networks, Second Edition*, pages 979–983. Cambridge, MA: MIT Press.